

Non-Parametric Field Estimation with Randomly Deployed, Noisy, Binary Sensors¹

Ye Wang and Prakash Ishwar
 Department of Electrical and Computer Engineering
 Boston University, Boston, MA
 {yw, pi}@bu.edu

Abstract—The reconstruction of a deterministic data field from binary-quantized noisy observations of sensors which are randomly deployed over the field domain is studied. The study focuses on the extremes of lack of deterministic control in the sensor deployment, lack of knowledge of the noise distribution, and lack of sensing precision and reliability. Such adverse conditions are motivated by possible real-world scenarios where a large collection of low-cost, crudely manufactured sensors are mass-deployed in an environment where little can be assumed about the ambient noise. A simple estimator that reconstructs the entire data field from these unreliable, binary-quantized, noisy observations is proposed. Technical conditions for the almost sure and integrated mean squared error (MSE) convergence of the estimate to the data field, as the number of sensors tends to infinity, are derived and their implications are discussed. For finite-dimensional, bounded-variation, and Sobolev-differentiable function classes, specific integrated MSE decay rates are derived. For the first and third function classes these rates are found to be minimax order optimal with respect to infinite precision sensing and known noise distribution.

Keywords: nonparametric regression; Monte-Carlo sampling; dithered scalar quantization; minimax rate of convergence; almost sure convergence; oversampled analog-to-digital conversion; distributed source coding; sensor networks; scaling law;

I. INTRODUCTION

In a recent paper [1] we considered the problem of reconstructing a bounded deterministic multidimensional data field $f : [0, 1]^p \rightarrow [a, -a]$, $0 < a < \infty$, from noisy dithered binary-quantized observations collected by n sensors randomly deployed over the field domain. The random sensor deployment model was based on uniform Monte Carlo sampling locations where n sensors are independently and identically distributed (iid) uniformly over the field domain² $[0, 1]^p$. A simple estimator that reconstructs the entire data field from these unreliable, binary-quantized, noisy observations was proposed in [1] and an upper bound on the integrated MSE of the estimator was derived. Using this bound, the integrated

MSE convergence of the estimator to the actual field as the number of sensors $n \rightarrow \infty$ was established.

In the present paper we expand and complete the development of results in [1]: (i) In Section III-B we expand the results of [1] to general deployment distributions. We establish a general upper bound to the integrated MSE which highlights the interaction of the deployment distribution and the orthonormal basis used for non-parametric field estimation (Theorem 3.1). (ii) We then derive sufficient conditions on the deployment distribution, the orthonormal basis, and the dimension of the field estimate which ensure the asymptotic (as $n \rightarrow \infty$) integrated MSE consistency of the proposed estimator. Implications for desirable deployment distributions are also discussed. (iii) In Section III-C we comprehensively investigate the asymptotic (as $n \rightarrow \infty$) almost sure consistency of the proposed estimator. The highlight of this section is Theorem 3.2 which provides an interesting set of sufficient conditions on the deployment distribution, the orthonormal basis, and the dimension of the field estimate which ensures asymptotic almost sure consistency of the estimation error. The implications of Theorem 3.2 are explored in detail through Proposition 3.1 and Corollary 3.2 and are of independent interest.

For the finite-dimensional, bounded-variation, and Sobolev-differentiable function classes, explicit achievable decay rates for the integrated MSEs are provided in Section IV. Specifically, for fields that belong to a finite-dimensional function space, the integrated MSE decays as³ $O(1/n)$ (Corollary 4.1). For fields of bounded-variation, the integrated MSE decays as $O(1/\sqrt{n})$ (Corollary 4.2). For fields that are s -Sobolev smooth (see IV-C), the integrated MSE decays as $O(n^{-\frac{2s}{2s+1}})$ (Corollary 4.3).

One of the highlights of this work is that for multidimensional fields living in rich function spaces, the minimax rate of convergence, of the integrated MSE, even with randomly deployed sensors, unknown noise statistics, and binary dithered scalar quantization (a highly nonlinear operation), can match the minimax rate of convergence with infinite-precision real-valued samples and known noise statistics.

The application context of this work is distributed sensing and coding for field reconstruction in wireless sensor networks as in [1]. The focus is on the extremes of lack of control in

¹This material is based upon work supported by the US National Science Foundation (NSF) under award (CAREER) CCF-0546598. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. A part of this work was presented at the 2007 International Symposium on Information Theory (ISIT).

²The field domain $[0, 1]^p$ is used for clarity and ease of exposition. However, the results can be generalized to compact subsets of \mathbb{R}^p .

³Landau's asymptotic notation: $f(n) = O(g(n)) \Leftrightarrow \limsup_{n \rightarrow \infty} |f(n)/g(n)| < \infty$; $f(n) = \Omega(g(n)) \Leftrightarrow g(n) = O(f(n))$; $f(n) = \Theta(g(n)) \Leftrightarrow f(n) = O(g(n))$ and $g(n) = O(f(n))$.

the sensor deployment, arbitrariness and lack of knowledge of the noise distribution, and low-precision and unreliability in the sensors. These adverse conditions are motivated by possible real-world scenarios where a large collection of low-cost, crudely manufactured sensors are mass-deployed in an environment where little can be assumed about the ambient noise. Each sensor measures a noisy sample of the field at its location under iid zero-mean, bounded amplitude, additive noise. The statistical distribution of the noise is *unknown* to the sensors and the fusion center, and the results in this paper hold for *arbitrary* distributions satisfying these assumptions. Each noisy sensor sample is quantized to a binary value by comparison with a random threshold (1-bit dithered scalar quantization). The binary-quantization models the extreme of low-precision quantization. The random thresholds are assumed to be iid across the sensors and uniformly distributed over the sample dynamic range, modeling the extreme unreliability in the quantization across sensors due to manufacturing process variations and environmental conditions at different sensor locations. Such extreme modeling assumptions are considered to demonstrate what is still achievable under adverse conditions.

The communication channel issues are abstracted away by assuming that the underlying sensor communication network is able to handle the modest payload of transmitting one bit (the binary-quantized observation) per sensor to the fusion center. The focus of this work is on reconstructing a single time snapshot of the field at a fusion center. The reconstruction of multiple time snapshots of the field can also be accommodated within the framework of this work as in [2] but is omitted for clarity. In fact, this can be achieved with time-sharing sensors, vanishing per-sensor rate, and vanishing sensor location “overheads”⁴ (see [2]). It is also assumed that the fusion center has access to the physical locations of the sensors and can correctly associate messages with their points of origin. This may be justifiable by possible models for the underlying wireless transmission where triangulation of sensors is inherently performed. The problem setup is illustrated in Figure 1.

The available literature on distributed field estimation which simultaneously treats binary-sensing, random sensor deployment, and unknown observation noise distribution is limited. The early works in [3], [4] consider the problem of reconstructing a signal from binary-quantized samples acquired with random thresholds, but do not consider arbitrary additive noise with an unknown distribution and only consider fixed deterministic sampling locations (deployment). The work in [5] is limited to the estimation of a *constant* field and does not explicitly address sampling precision (sensing) constraints. A recent work [2] provides pointwise MSE decay rates in terms of the local and global modulus of field continuity by building upon the techniques in [3], [4], [5]. However, [2] does not consider random sensor deployment and requires local field continuity for pointwise MSE convergence. The present work incorporates random sensor deployment,

binary-sensing, and unknown noise distribution while studying almost sure and integrated MSE convergence of the field estimate. The integrated MSE convergence for the bounded-variation, Sobolev-differentiable, and finite-dimensional function classes are explored in detail. Our results expose the effects of field “smoothness”, deployment randomness, and observation/sensing noise on the integrated MSE scaling behavior.

For field estimation approaches which are not constrained by finite sensing precision and sensing unreliability, such as those involving “uncoded” analog joint sampling-transmission, there is a growing body of literature now available (e.g., see [6], [7], [8], [9], [10], [11], [12] and references therein). Related to the distributed field reconstruction problem is the so-called CEO problem studied in the Information Theory community in which the distortion is averaged over multiple field snapshots over time (e.g., see [13], [14], [15] and references therein). There is also a significant body of work on oversampled A-D conversion (e.g., see [16] and references therein), which is loosely related to the results of the present work concerning finite-dimensional fields. However, these are different problem formulations and are not the focus of the present work.

The rest of this paper is organized as follows. The problem formulation with detailed modeling assumptions are presented in Section II. The core technical results are then summarized and discussed in Section III. The core results are then used to derived explicit expressions of the decay rate of the integrated MSE for three rich function classes in Section IV. The proofs of all the core technical results are presented in Section V and concluding remarks are made in Section VI.

II. PROBLEM FORMULATION

Field Model: We model the field as a real-valued, bounded, deterministic function $f : \mathcal{D} \rightarrow [-a, +a]$ belonging to a non-parametric function class⁵ \mathcal{F} , that is, $f \in \mathcal{F}$, where \mathcal{F} is a set of measurable functions mapping \mathcal{D} to $[-a, +a]$. The domain of the field \mathcal{D} is assumed to be a compact subset of \mathbb{R}^d , the d -dimensional Euclidean space. The objective is to reconstruct this function with high fidelity from binary-quantized noisy observations collected by a network of non-cooperative⁶ sensors that are randomly deployed over the domain \mathcal{D} .

Random Sensor Deployment: We assume that the n sensors are independently and identically randomly deployed over the domain \mathcal{D} according to a known distribution p_X . If $X_i \in \mathcal{D}$ denotes the location of the i^{th} sensor for $i \in \{1, \dots, n\}$, then $X_i \sim \text{iid } p_X$ captures the lack of control in sensor deployment. We assume that the support of p_X is \mathcal{D} and that p_X is a non-singular distribution⁷.

Additive Noise: Each sensor takes a sample of the field under additive noise. The noisy samples are given by $Y_i =$

⁵The number of parameters that specify a non-parametric function class is not fixed a priori and is possibly infinite.

⁶The sensors do not exchange information or otherwise collaborate at the time of or before taking measurements.

⁷A random variable with a non-singular distribution takes values in a subset of \mathcal{D} with Lebesgue measure 0 with probability 0.

⁴Network overheads refer to additional bits of information that must be attached to each message to identify the point of origin of the message.

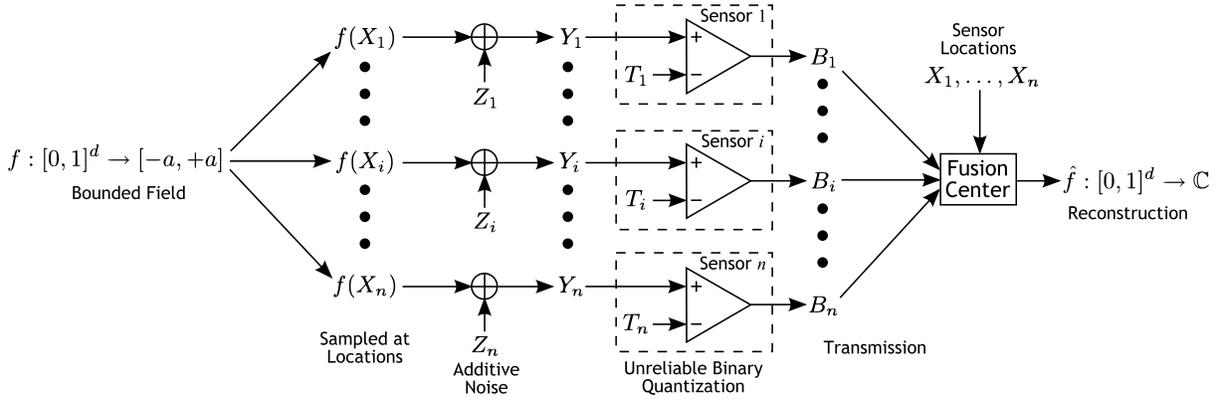


Fig. 1. **Problem Setup:** The field (in a single snapshot) is sampled by n sensors at their respective locations under additive noise. Each sample is unreliably quantized to a binary value by a comparison with a random threshold. These binary values are transmitted to the fusion center which reconstructs the field.

$f(X_i) + Z_i$, for $i \in \{1, \dots, n\}$, where the noise variables $Z_i \sim \text{iid } p_Z$ and are independent of the sensor locations. We assume that each Z_i is zero-mean and is bounded in amplitude by a constant $b > 0$, that is, the support of p_Z is contained in $[-b, +b]$. However, besides these assumed conditions, the distribution p_Z is *unknown* to either the sensors or the fusion center, and the results and methods of this paper hold for *arbitrary* noise distributions satisfying these conditions. We let \mathcal{P}_Z denote the set of all noise distributions satisfying these assumptions. Note that since both the field and the noise are bounded, the noisy samples are bounded: $|Y_i| \leq c := a + b$. We assume that the value of c is known. The values of a and b can remain unknown to the sensors and the fusion center.

Unreliable, Binary Quantization: We assume that in the sensor hardware frontend, the noisy sample is quantized by an unreliable, low-precision analog-to-digital converter. Specifically, we consider one-bit (binary), dithered, scalar quantization implemented as a comparison to a random threshold that is uniformly distributed over the sample dynamic range $[-c, +c]$. The binary-quantized observations are given by

$$\begin{aligned} B_i &= \text{sgn}(Y_i - T_i) && \text{for } i \in \{1, \dots, n\} \\ &:= \begin{cases} +1 & Y_i > T_i, \\ -1 & Y_i \leq T_i, \end{cases} = \begin{cases} +1 & f(X_i) + Z_i > T_i, \\ -1 & f(X_i) + Z_i \leq T_i, \end{cases} \end{aligned}$$

where $T_i \sim \text{iid Unif}[-c, +c]$ are the uniform random thresholds. The thresholds are independent of the sensor locations and the noise. The value of B_i is finally the observation that sensor i has access to.

Transmission: We abstract away communication channel issues and assume that the underlying communication network of the sensors is able to handle the modest payload of transmitting one bit per sensor to the fusion center. We also assume that the fusion center has access to the physical locations of the sensors and can correctly associate messages with their points of origin. Thus, we assume that through this abstracted communications network, the sensor location and quantized observation pairs $\{(X_i, B_i)\}_{i=1}^n$ are reliably made available to the fusion center. The reconstruction of multiple time snapshots of the field with time-sharing sensors, vanishing per-sensor rate and sensor location ‘‘overheads’’ can

also be accommodated within the framework of this work as in [2] but is omitted for clarity.

Reconstruction and Distortion Criterion: Given $\{(X_i, B_i)\}_{i=1}^n$, the fusion center constructs the field estimate $\hat{f}_{X_1, \dots, X_n, B_1, \dots, B_n}: \mathcal{D} \rightarrow \mathbb{C}$. For notational convenience, the explicit dependence on $\{(X_i, B_i)\}_{i=1}^n$ will be suppressed and the estimator will simply be denoted by \hat{f}_n . The performance criterion is the integrated MSE given by

$$D(f, \hat{f}_n) := \mathbb{E} \left[\|f - \hat{f}_n\|^2 \right] = \mathbb{E} \left[\int_{\mathcal{D}} |f(x) - \hat{f}_n(x)|^2 dx \right],$$

where the expectation is taken with respect to the random noise, thresholds, and the sensor locations. The objective is to design an estimator \hat{f}_n that minimizes the integrated MSE D . The problem setup is shown in Figure 1.

Minimax Integrated MSE: For a given field subclass $\mathcal{F}_{\text{sub}} \subset \mathcal{F}$, of interest are the corresponding upper, lower, and minimax rates of convergence of the integrated MSE. A positive sequence γ_n is an upper rate of convergence if there exists a constant $C < \infty$ and an estimator \hat{f}_n^* such that

$$\limsup_{n \rightarrow \infty} \sup_{p_Z \in \mathcal{P}_Z} \sup_{f \in \mathcal{F}_{\text{sub}}} \gamma_n^{-1} D(f, \hat{f}_n^*) \leq C.$$

A positive sequence γ_n is a lower rate of convergence if there exists a constant $C > 0$ such that

$$\liminf_{n \rightarrow \infty} \inf_{\hat{f}_n} \inf_{p_Z \in \mathcal{P}_Z} \sup_{f \in \mathcal{F}_{\text{sub}}} \gamma_n^{-1} D(f, \hat{f}_n) \geq C,$$

where the $\inf_{\hat{f}_n}$ is the infimum over all field estimators. The upper rate represents the asymptotic worst-case performance achieved by a given estimator. The lower rate represents a fundamental limit on the asymptotic performance of any estimator. A positive sequence γ_n that is both a lower rate and an upper rate of convergence is called the minimax rate of convergence and the corresponding estimator \hat{f}_n^* that achieves the upper rate is called a minimax order optimal estimator.

Note that showing $D(f, \hat{f}_n^*) = O(\gamma_n)$ for all $f \in \mathcal{F}_{\text{sub}}$ and $p_Z \in \mathcal{P}_Z$ for a particular estimator \hat{f}_n^* is equivalent to showing that \hat{f}_n^* achieves γ_n as an upper rate of convergence of the integrated MSE. If it can be further shown that $D(f, \hat{f}_n) = \Omega(\gamma_n)$ for a particular $f \in \mathcal{F}_{\text{sub}}$, a particular $p_Z \in \mathcal{P}_Z$, and for all estimators \hat{f}_n , then γ_n is the minimax rate of convergence of the integrated MSE.

III. MAIN RESULTS

In this section, we describe our proposed field estimator and analyze its performance. We show that under suitable technical conditions, the field estimate is asymptotically integrated MSE consistent, that is, as $n \rightarrow \infty$, $\mathbb{E}[\|f - \hat{f}_n\|^2] \rightarrow 0$. We also show that under suitable technical conditions, the field estimate is asymptotically almost sure consistent, that is, as $n \rightarrow \infty$, almost surely $\hat{f}_n \rightarrow f$ pointwise almost everywhere on \mathcal{D} . We also provide an upper bound to the integrated MSE which is used in Section IV to derive achievable integrated MSE decay rates for specific function classes. The proofs of all theorems are presented in Section V.

Let \mathcal{F} denote the set of all bounded, measurable functions $f: \mathcal{D} \rightarrow [-a, +a]$. Note that $\mathcal{F} \subseteq \mathbf{L}^2(\mathcal{D})$. Let $\mathcal{B} = \{\phi_j\}_{j=0}^\infty$, with $\phi_j: \mathcal{D} \rightarrow \mathbb{C}$, denote an indexed orthonormal (Schauder) basis (e.g. Fourier, wavelet, etc.) of $\mathbf{L}^2(\mathcal{D})$. Any $f \in \mathcal{F}$ can be decomposed as

$$f \stackrel{\mathbf{L}^2}{=} \sum_{j=0}^{\infty} \langle f, \phi_j \rangle \phi_j =: \sum_{j=0}^{\infty} \alpha_j \phi_j, \quad (1)$$

where $\alpha_j := \langle f, \phi_j \rangle$ denotes the coefficients (projections onto the basis functions) of the expansion. The m -term approximation of f with respect to an orthonormal basis $\mathcal{B} = \{\phi_j\}_{j=0}^\infty$ is given by

$$f_m := \sum_{j=0}^{m-1} \langle f, \phi_j \rangle \phi_j. \quad (2)$$

The corresponding m -term approximation error is given by

$$\varepsilon[f, m, \mathcal{B}] := \|f - f_m\|^2 = \sum_{j=m}^{\infty} |\langle f, \phi_j \rangle|^2 = \sum_{j=m}^{\infty} |\alpha_j|^2, \quad (3)$$

which is a non-negative, non-increasing sequence of m that converges to zero for all $f \in \mathcal{F}$ [17, Chapter 9].

A. Proposed estimator

Our proposed estimator first estimates the first m coefficients $\{\alpha_j\}_{j=0}^{m-1}$ of (1) with respect to a given orthonormal basis \mathcal{B} , according to

$$\hat{\alpha}_j := \frac{c}{n} \sum_{i=1}^n \frac{\phi_j^*(X_i)}{p_X(X_i)} B_i, \quad (4)$$

for $j \in \{0, \dots, m-1\}$. A general tunable field estimate is given by the m -term approximation,

$$\hat{f}_{n,m} := \sum_{j=0}^{m-1} \hat{\alpha}_j \phi_j, \quad (5)$$

where m is the tunable design parameter which can be chosen to depend on n to optimize the rate of decay of the integrated MSE for specific function classes. The final field estimate is given by specifying m as a function of n ,

$$\hat{f}_n := \sum_{j=0}^{m(n)-1} \hat{\alpha}_j \phi_j. \quad (6)$$

The specification of $m(n)$ for specific function classes is discussed in Section IV. The dependence of m on n needs

to satisfy certain conditions to ensure that the estimate is asymptotically consistent. These conditions are described in Section III-B and Section III-C.

B. Integrated MSE upper bounds and convergence results

The following theorem, whose proof appears in Section V-A, upper bounds the integrated MSE as the sum of two terms. The first term is due to the variance of the coefficient estimates. The second term is due to the bias caused by the finite-term series approximation.

Theorem 3.1: (Integrated MSE Upper Bound) *Let \mathcal{F} , \mathcal{P}_Z , and p_X be as given in Section II. Let $\hat{f}_{n,m}$ be given by (4) and (5), where $\mathcal{B} = \{\phi_j\}_{j=0}^\infty$ is any orthonormal Schauder basis of $\mathbf{L}^2(\mathcal{D})$. Then, $\forall f \in \mathcal{F}$ and $\forall p_Z \in \mathcal{P}_Z$, the integrated MSE is upper bounded by*

$$D = \mathbb{E} \left[\|f - \hat{f}_{n,m}\|^2 \right] \leq \frac{c^2}{n} \sum_{j=0}^{m-1} \int_{\mathcal{D}} \frac{|\phi_j(x)|^2}{p_X(x)} dx + \varepsilon[f, m, \mathcal{B}], \quad (7)$$

where $\varepsilon[f, m, \mathcal{B}]$, given by (3), is a non-negative, non-increasing sequence that converges to 0 as $m \rightarrow \infty$.

In light of Theorem 3.1 we now examine conditions on $m(n)$, \mathcal{B} , and p_X which ensure that the estimator is asymptotically consistent in the integrated MSE sense, that is $D \rightarrow 0$ as $n \rightarrow \infty$. The following corollary specifies conditions that immediately ensures that the integrated MSE converges to 0.

Corollary 3.1: (Integrated MSE Convergence of the Field Estimate) *Under the same setup of Theorem 3.1, if $m(n)$, $\mathcal{B} = \{\phi_j\}_{j=0}^\infty$, and p_X satisfy*

$$m(n) \rightarrow \infty, \quad \text{as } n \rightarrow \infty, \quad (8)$$

$$\frac{1}{n} \sum_{j=0}^{m(n)-1} \int_{\mathcal{D}} \frac{|\phi_j(x)|^2}{p_X(x)} dx \rightarrow 0, \quad \text{as } n \rightarrow \infty, \quad (9)$$

then the estimate converges in the integrated MSE sense to the field, that is,

$$\forall f \in \mathcal{F} \text{ and } \forall p_Z \in \mathcal{P}_Z, \quad D \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Condition (8) is sufficient (and often necessary) to ensure that $\varepsilon[f, m, \mathcal{B}]$ converges to 0. Condition (9) is equivalent to the first term of the integrated MSE upper bound, given in (7), converging to 0. For some deployment distributions p_X , condition (9) may not be attainable for many orthonormal bases. For example, let the domain $\mathcal{D} = [0, 1]$ with the deployment distribution $p_X(x) = 2x$ over $[0, 1]$. Then for any orthonormal basis in which $\phi_0(x) = 1$ over $[0, 1]$, e.g., Fourier, Harr wavelets, Legendre polynomials, etc., the first term of the summation in (9) is given by

$$\int_{\mathcal{D}} \frac{|\phi_0(x)|^2}{p_X(x)} dx = \int_0^1 \frac{1}{2x} dx = \infty.$$

Thus integrated MSE upper bound becomes useless. This implies that in general the deployment distributions and orthonormal bases have to be appropriately *matched* as a design consideration in order to satisfy condition (9). However, condition (9) is ensured for any orthonormal basis if the deployment distribution p_X has a strictly positive infimum over \mathcal{D} , that is,

$$\inf_{x \in \mathcal{D}} p_X(x) = \nu > 0. \quad (10)$$

Sensor deployment distributions over compact domains which are useful for high-resolution field reconstruction would satisfy such a condition. Given (10), we have that

$$\begin{aligned} \sum_{j=0}^{m-1} \int_{\mathcal{D}} \frac{|\phi_j(x)|^2}{p_X(x)} dx &\leq \sum_{j=0}^{m-1} \int_{\mathcal{D}} \frac{|\phi_j(x)|^2}{\nu} dx \\ &= \sum_{j=0}^{m-1} \frac{1}{\nu} \|\phi_j\|^2 = \frac{m}{\nu}, \end{aligned}$$

and $\forall f \in \mathcal{F}$ and $\forall p_Z \in \mathcal{P}_Z$, the corresponding integrated MSE upper bound becomes

$$D = \mathbb{E} \left[\|f - \hat{f}_{n,m}\|^2 \right] \leq \frac{c^2 m}{n\nu} + \varepsilon[f, m, \mathcal{B}]. \quad (11)$$

The upper bound in (11) converges to 0 as $n \rightarrow \infty$ when $m(n)$ satisfies the following two conditions

$$\begin{aligned} m(n) &\rightarrow \infty, \quad \text{as } n \rightarrow \infty, \\ \frac{m(n)}{n} &\rightarrow 0, \quad \text{as } n \rightarrow \infty. \end{aligned}$$

C. Almost sure convergence results

In this subsection we establish sufficient conditions for the field estimate to be asymptotically almost sure consistent, that is, as $n \rightarrow \infty$, almost surely $\hat{f}_n \rightarrow f$ pointwise almost everywhere on \mathcal{D} . First, we establish a key theorem that gives sufficient conditions for the convergence of the pointwise errors of the estimate with respect to the truncated approximation of the field. The proof of this theorem appears in Section V-B.

Theorem 3.2: (Almost Sure Convergence of Estimate Errors) *Let p_X be the deployment distribution described in Section II satisfying (10), $\mathcal{B} = \{\phi_j\}_{j=0}^{\infty}$ be an orthonormal Schauder basis of $\mathbf{L}^2(\mathcal{D})$, $\hat{f}_{n,m}$ be the field estimate given by (4) and (5), and f_m be the m -term approximation to the field given by (2). Let*

$$S_{n,m}(x) := \hat{f}_{n,m}(x) - f_m(x),$$

for all $x \in \mathcal{D}$. If there exists a non-negative, increasing sequence of real numbers $\{\Lambda_m\}_{m=1}^{\infty}$, and a non-negative, increasing sequence of positive integers $\{m(n)\}_{n=1}^{\infty}$ which satisfy the following three conditions

$$\forall x, y \in \mathcal{D}, \quad \left| \sum_{j=0}^{m-1} \phi_j(x) \phi_j^*(y) \frac{1}{p_X(y)} \right| \leq C_1 \Lambda_m, \quad (12)$$

$$\forall f \in \mathcal{F}, \forall x \in \mathcal{D}, \quad \left| \sum_{j=0}^{m-1} \langle f, \phi_j \rangle \phi_j(x) \right| \leq C_2 \Lambda_m, \quad (13)$$

$$\forall \epsilon > 0, \quad \sum_{n=1}^{\infty} \exp\left(\frac{-\epsilon^2 n}{\Lambda_m^2}\right) < \infty, \quad (14)$$

where $C_1, C_2 > 0$ are some constants, then $\forall f \in \mathcal{F}$ and $\forall x \in \mathcal{D}$ except on a set of Lebesgue measure zero, as $n \rightarrow \infty$, almost surely,

$$S_n(x) := S_{n,m(n)} \rightarrow 0.$$

Conditions (12) and (13) impose constraints on the basis functions $\mathcal{B} = \{\phi_j\}_{j=0}^{\infty}$ and the deployment distribution p_X . Condition (14) implies that as $n \rightarrow \infty$, $\Lambda_{m(n)}^2/n \rightarrow 0$. This places a constraint on how fast $m(n)$ can go to infinity. In particular it requires that in relation to Λ_m , $m(n)$ not grow too fast with n .

We now examine some special choices of $\{\Lambda_m\}_{m=1}^{\infty}$ for which conditions (12) and (13) will hold. For $m \in \{1, 2, \dots\}$, define auxiliary functions:

$$g_m(x, y) := \frac{1}{\Lambda_m} \sum_{j=0}^{m-1} \phi_j(x) \phi_j^*(y) \frac{1}{p_X(y)}, \quad (15)$$

$$h_m(x) := \frac{1}{\Lambda_m} \sum_{j=0}^{m-1} \langle f, \phi_j \rangle \phi_j(x), \quad (16)$$

for $x, y \in \mathcal{D}$. The following proposition, whose proof appears in Section V-C, gives two sets of conditions on $\{\Lambda_m\}_{m=1}^{\infty}$, $\mathcal{B} = \{\phi_j\}_{j=0}^{\infty}$, and p_X , for which conditions (12) and (13) will hold.

Proposition 3.1: *Let $\{\Lambda_m\}_{m=1}^{\infty}$ be as in Theorem 3.2 and g_m, h_m be given by (15) and (16) respectively.*

(i) *If*

$$\frac{m}{\Lambda_m^2} \rightarrow 0, \quad \text{as } m \rightarrow \infty, \quad (17)$$

and for $x, y \in \mathcal{D}$ almost everywhere, the limits

$$g_{\infty}(x, y) := \lim_{m \rightarrow \infty} g_m(x, y), \quad (18)$$

$$h_{\infty}(x) := \lim_{m \rightarrow \infty} h_m(x) \quad (19)$$

exist, then the limits are zero almost everywhere and the conditions (12) and (13) are satisfied for some constants $C_1, C_2 > 0$.

(ii) *If the basis functions are uniformly amplitude bounded, that is, $\forall j \in \{0, 1, \dots\}$ and $\forall x \in \mathcal{D}$,*

$$|\phi_j(x)| \leq \beta < \infty,$$

then conditions (12) and (13) are satisfied for $\Lambda_m = m$ with constants $C_1 = \beta^2/\nu$ and $C_2 = a\beta\sqrt{\text{vol}(\mathcal{D})}$.

Part (i) of Proposition 3.1 shows that if the limits of the auxiliary functions (15) and (16) as $m \rightarrow \infty$ exist, then for any Λ_m such that (17) is satisfied, e.g., $\Lambda_m = m^{\gamma/2}$, for any $\gamma > 1$, conditions (12) and (13) are satisfied for some constants. Part (ii) of Proposition 3.1 shows that if the basis functions are uniformly bounded as, for example, in the orthonormal Fourier and Legendre bases, then conditions (12) and (13) are satisfied for $\Lambda_m = m$ and given constants.

We now examine conditions on $\{\Lambda_m\}_{m=1}^{\infty}$ under which (14) will be satisfied. According to Ermakoff's series convergence test [18], if for some non-negative, non-increasing, real function $q(t)$, $t \geq 1$,

$$\lim_{t \rightarrow \infty} \frac{e^t q(e^t)}{q(t)} < 1,$$

where e is the base of the natural logarithm, then

$$\sum_{n=1}^{\infty} q(n) < \infty.$$

Let $q(t) = \exp\left(\frac{-\epsilon^2 t}{t^\psi}\right)$, $t \geq 1$, where $\psi \in (0, 1)$ and $\epsilon > 0$. Then

$$\frac{e^t q(e^t)}{q(t)} = \frac{e^t \exp\left(\frac{-\epsilon^2 e^t}{e^{t\psi}}\right)}{\exp\left(\frac{-\epsilon^2 t}{t^\psi}\right)} = \exp\left(t - \epsilon^2 e^{t-t\psi} - \epsilon^2 t^{1-\psi}\right) \longrightarrow 0,$$

as $t \longrightarrow \infty$. By Ermakoff's test, for all $\psi \in (0, 1)$ and all $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \exp\left(\frac{-\epsilon^2 n}{n^\psi}\right) < \infty.$$

Thus condition (14) will be satisfied if $\Lambda_{m(n)}^2 = n^\psi$ for any $\psi \in (0, 1)$.

Combining the above result with Proposition 3.1 yields possible forms of the design parameters $\{m(n)\}_{n=1}^{\infty}$ and $\{\Lambda_m\}_{m=1}^{\infty}$ such that the conditions for almost sure convergence (12), (13), and (14) are all simultaneously satisfied. Choosing $m(n) = \Theta(n^\psi)$, where $\psi \in (0, 1)$, and $\Lambda_m = m^{\gamma/2}$, for some $\gamma \in (1, 1/\psi)$, yields $\Lambda_{m(n)}^2 = n^{\psi'}$, where $\psi' = \gamma\psi \in (\psi, 1)$, which satisfies (14) and (17) simultaneously. With these choices, Proposition 3.1 shows that conditions (12) and (13) will be satisfied as well if the limits (18) and (19) of the auxiliary functions (15) and (16) respectively can be assumed to exist. Thus, for any $m(n)$ of the form $m(n) = \Theta(n^\psi)$, where $\psi \in (0, 1)$, we can choose $\{\Lambda_m\}_{m=1}^{\infty}$ such that conditions (12), (13), and (14) are simultaneously satisfied, if the limits (18) and (19) exist.

Due to the properties of an orthonormal basis, as $m \longrightarrow \infty$, the m -term approximation, f_m given by (2), converges in \mathbf{L}^2 -norm to f for any $f \in \mathcal{F}$. Although, it is not guaranteed that for general orthonormal bases f_m will converge pointwise almost everywhere to a specific function. However, if f_m does converge almost everywhere to some f_∞ , then f_∞ must be equal to f almost everywhere. This can be seen by writing

$$\begin{aligned} 0 &\leq \int_{\mathcal{D}} |f(x) - f_\infty(x)|^2 dx \\ &= \int_{\mathcal{D}} \liminf_{m \rightarrow \infty} |f(x) - f_m(x)|^2 dx \\ &\leq \liminf_{m \rightarrow \infty} \int_{\mathcal{D}} |f(x) - f_m(x)|^2 dx = 0, \end{aligned}$$

where the inequality follows due to Fatou's lemma [19]. Thus $\int_{\mathcal{D}} |f(x) - f_\infty(x)|^2 dx = 0$, so $|f(x) - f_\infty(x)| = 0$ for $x \in \mathcal{D}$ almost everywhere. For example, it is well known that for any $f \in \mathcal{F} \subset \mathbf{L}^2([0, 1])$ the m -term Fourier series approximation, f_m converges to f almost everywhere [20].

Corollary 3.2: (Almost Sure Convergence of the Field Estimate) *Within the context of Theorem 3.2, if conditions (12), (13), and (14) hold and if as $m \longrightarrow \infty$, the m -term approximation f_m converges almost everywhere to some function f_∞ , then $\forall x \in \mathcal{D}$, the pointwise error of the field estimate satisfies*

$$\begin{aligned} |\hat{f}_n(x) - f(x)| &\leq |\hat{f}_n(x) - f_{m(n)}(x)| + |f_{m(n)}(x) - f(x)| \\ &= |S_n(x)| + |f_{m(n)}(x) - f(x)| \\ &\xrightarrow{\text{a.s.}} 0, \quad \text{as } n \longrightarrow \infty. \end{aligned}$$

Thus, for $x \in \mathcal{D}$ almost everywhere, as $n \longrightarrow \infty$, almost surely $\hat{f}_n(x) \longrightarrow f(x)$.

IV. ACHIEVABLE INTEGRATED MSE DECAY RATES

In this section, we use the integrated MSE upper bound (7) derive explicit expressions for the achievable upper rates of convergence of the integrated MSE for three specific function classes, namely, finite-dimensional $\mathcal{F}_{\mathcal{B}_k}$, bounded-variation \mathcal{F}_{BV} , and s -Sobolev differentiable \mathcal{F}_s . Throughout this section, we assume that (10) holds.

The general approach for deriving such rates of convergence for functions living in a function class $\mathcal{F}_{\text{sub}} \subseteq \mathcal{F}$ is select an appropriate basis $\mathcal{B} = \{\phi_j\}_{j=0}^{\infty}$ in which the m -term approximation error given by $\varepsilon[f, m, \mathcal{B}]$ in (3) can be upper bounded by an explicit function of m for all $f \in \mathcal{F}_{\text{sub}}$. Then m in (7) can be chosen to depend on n to optimize the convergence rate. Thus given the appropriate function approximation theoretic results that upper bound $\varepsilon[f, m, \mathcal{B}]$, this approach establishes achievable upper rates of convergence of the integrated MSE for the corresponding function class.

A. Functions in a finite-dimensional subspace of \mathcal{F}

The first function class represents the scenario where the fusion center has an exact prior knowledge of the finite-dimensional space in which the function lives. Let $\mathcal{F}_{\mathcal{B}_k}$ denote the subset of \mathcal{F} that is composed of functions that are linear combinations of a given set of k orthonormal functions $\mathcal{B}_k = \{\phi_j\}_{j=0}^{k-1}$. Note that for any $f \in \mathcal{F}_{\mathcal{B}_k}$, $f = \sum_{j=0}^{k-1} \langle f, \phi_j \rangle \phi_j$. Thus the function approximation at the truncation point $m = k$ is exact, that is, $f_m = f$ for $m = k$ so that,

$$\forall f \in \mathcal{F}_{\mathcal{B}_k}, \quad \varepsilon[f, k, \mathcal{B}_k] = 0. \quad (20)$$

Combining (7) with (20) yields the following corollary.

Corollary 4.1: (Decay rate of integrated MSE for $\mathcal{F}_{\mathcal{B}_k}$) *Let \mathcal{B}_k and $\mathcal{F}_{\mathcal{B}_k}$ be as given above and \mathcal{P}_Z and p_X be as given in Section II. Let $\hat{f}_{n,m}$ be given by (4) and (5) with \mathcal{B}_k as the basis. If p_X satisfies (10), then $\forall f \in \mathcal{F}_{\mathcal{B}_k}$ and $\forall p_Z \in \mathcal{P}_Z$, the integrated MSE of $\hat{f}_{n,m}$ with the truncation point m set to k is upper bounded as follows*

$$D = \mathbb{E} \left[\|f - \hat{f}_{n,m}\|^2 \right] \leq \frac{c^2 k}{n\nu} = O\left(\frac{1}{n}\right).$$

Therefore, $\forall f \in \mathcal{F}_{\mathcal{B}_k}$ and $\forall p_Z \in \mathcal{P}_Z$, an achievable upper rate of convergence of the integrated MSE for fields in a finite-dimensional subspace is given by

$$D = \mathbb{E} \left[\|f - \hat{f}_{n,m}\|^2 \right] = O\left(\frac{1}{n}\right).$$

It should be noted that for this function class, the field estimation problem for integrated MSE is equivalent to a finite-dimensional parameter estimation problem with conditionally independent noisy observations. Under the choice of an appropriate, well-behaved noise distribution⁸ $p_Z \in \mathcal{P}_Z$, the Cramér-Rao lower bound for the integrated MSE decay rate for finite-dimensional parameter estimation from iid noisy observations asymptotically behaves as $D = \Omega(1/n)$ for all asymptotically integrated MSE consistent estimators [21]. Hence the estimator is minimax order optimal for $\mathcal{F}_{\mathcal{B}_k}$ and achieves the minimax rate of convergence $\gamma_n = (1/n)$.

⁸A noise distribution is chosen such that the observation model satisfies the Cramér-Rao regularity conditions [21].

B. Functions of bounded-variation on Domain $\mathcal{D} = [0, 1]$

Let \mathcal{F}_{BV} denote the subset of \mathcal{F} which is composed of functions on $\mathcal{D} = [0, 1]$ of bounded-variation. Formally,

$$\mathcal{F}_{BV} := \left\{ f \in \mathcal{F} \left| \lim_{\delta \rightarrow 0} \int_0^1 \frac{|f(x) - f(x - \delta)|}{|\delta|} dx < +\infty \right. \right\}.$$

A function in \mathcal{F}_{BV} has a derivative (at points for which it exists) which is uniformly bounded and the sum of the amplitudes of its discontinuous jumps is finite. The bounded-variation condition represents a minimal ‘‘smoothness’’ assumption since a restriction is placed on the amount of total discontinuous jumps.

It is well known that for the Fourier basis,

$$\mathcal{B}_{\text{Fourier}} = \left\{ \phi_j(x) = \begin{cases} e^{+\pi j x \sqrt{-1}}, & j \text{ even,} \\ e^{-\pi(j+1)x\sqrt{-1}}, & j \text{ odd} \end{cases} \right\}_{j=0}^{\infty}, \quad (21)$$

the m -term approximation error (3) is upper bounded as follows,

$$\forall f \in \mathcal{F}_{BV}, \quad \varepsilon[f, m, \mathcal{B}_{\text{Fourier}}] \leq \frac{\sigma}{m}, \quad (22)$$

where $\sigma > 0$ is a constant [17, Chapter 9]. Combining (7) with (22) yields the following corollary.

Corollary 4.2: (Decay rate of integrated MSE for \mathcal{F}_{BV})

Let \mathcal{F}_{BV} be as given above and \mathcal{P}_Z and p_X be as given in Section II. Let $\hat{f}_{n,m}$ be given by (4) and (5) with $\mathcal{B}_{\text{Fourier}}$ as given in (21). If p_X satisfies (10), then $\forall f \in \mathcal{F}_{BV}$ and $\forall p_Z \in \mathcal{P}_Z$, the integrated MSE of $\hat{f}_{n,m}$ is upper bounded as follows

$$D = \mathbb{E} \left[\|f - \hat{f}_{n,m}\|^2 \right] \leq \frac{c^2 m}{n\nu} + \frac{\sigma}{m},$$

where $\sigma > 0$ is a constant. Setting $m(n) = \sqrt{n}$ to optimize the decay rate of the upper bound yields the following achievable upper rate of convergence of the integrated MSE $\forall f \in \mathcal{F}_{BV}$ and $\forall p_Z \in \mathcal{P}_Z$:

$$D = \mathbb{E} \left[\|f - \hat{f}_{n,m}\|^2 \right] = O \left(\frac{1}{\sqrt{n}} \right).$$

C. Sobolev differentiable functions on Domain $\mathcal{D} = [0, 1]$

This function class includes functions which are differentiable in a generalized sense to a degree of differentiability parameterized by s which can take non-integer values. The value of s can be considered as a measure of smoothness. For $s > 1/2$, let \mathcal{F}_s denote the subset of \mathcal{F} which is composed of functions on $\mathcal{D} = [0, 1]$ that are s -times Sobolev differentiable. Formally,

$$\mathcal{F}_s := \left\{ f \in \mathcal{F} \left| \int_{-\infty}^{+\infty} |\omega|^{2s} |\tilde{f}(\omega)|^2 d\omega < +\infty \right. \right\}, \quad (23)$$

where $\tilde{f}(\omega)$ denotes the Fourier transform of f . Note that the condition in (23) (for integer values of s) corresponds to the s^{th} derivative of f belonging to $\mathbf{L}^2([0, 1])$. Thus, this set includes functions that are $\lfloor s \rfloor$ -times differentiable.

It is well known that for $s > 1/2$,

$$\forall f \in \mathcal{F}_s, \quad \varepsilon[f, m, \mathcal{B}_{\text{Fourier}}] \leq \frac{\sigma}{m^{2s}}, \quad (24)$$

where $\sigma > 0$ is a constant [17, Chapter 9]. Combining (7) with (24) yields the following corollary.

Corollary 4.3: (Decay rate of integrated MSE for \mathcal{F}_s) Let \mathcal{F}_s be as given above and \mathcal{P}_Z and p_X be as given in Section II. Let $\hat{f}_{n,m}$ be given by (4) and (5) with $\mathcal{B}_{\text{Fourier}}$ as given in (21). If p_X satisfies (10), then $\forall f \in \mathcal{F}_s$ and $\forall p_Z \in \mathcal{P}_Z$, the integrated MSE of $\hat{f}_{n,m}$ is upper bounded as follows

$$D = \mathbb{E} \left[\|f - \hat{f}_{n,m}\|^2 \right] \leq \frac{c^2 m}{n\nu} + \frac{\sigma}{m^{2s}},$$

where $\sigma > 0$ is a constant. Setting $m(n) = n^{\frac{1}{2s+1}}$ to optimize the decay rate of the upper bound yields the following achievable upper rate of convergence of the integrated MSE $\forall f \in \mathcal{F}_s$ and $\forall p_Z \in \mathcal{P}_Z$:

$$D = \mathbb{E} \left[\|f - \hat{f}_{n,m}\|^2 \right] = O \left(n^{\frac{-2s}{2s+1}} \right).$$

It is well known that the exact minimax rate of convergence of the integrated MSE for non-parametric regression, based on full-resolution, real-valued, noisy observations in an s -Sobolev space is given by $\gamma_n = n^{\frac{-2s}{2s+1}}$ [22], [23]. In non-parametric regression, the field estimate is based directly on the full-resolution real-valued noisy observations $\{Y_i\}_{i=1}^n$, whereas in our problem the field estimate is based on only the binary-quantized observations $\{B_i\}_{i=1}^n$. In both setups, the corresponding sensor locations are known. Thus, it is interesting to observe that our proposed estimator is minimax order optimal even with respect to the case in which the observations have not been quantized.

V. PROOFS

A. Proof of Theorem 3.1

We first establish some results regarding the estimated coefficients of (4).

Lemma 5.1: The expected value of an approximated coefficient is given by

$$(i) \quad \mathbb{E}[\hat{\alpha}_j] = \alpha_j = \langle f, \phi_j \rangle, \quad (25)$$

and the integrated MSE of the coefficient estimates satisfies

$$(ii) \quad \mathbb{E}[|\hat{\alpha}_j - \alpha_j|^2] \leq \frac{c^2}{n} \int_{\mathcal{D}} \frac{|\phi_j(x)|^2}{p_X(x)} dx. \quad (26)$$

The approximated coefficients also have the following convergence property

$$(iii) \quad \hat{\alpha}_j \xrightarrow{\text{a.s.}} \alpha_j, \quad n \rightarrow \infty. \quad (27)$$

Proof: (i) The expectation of the coefficient estimates can be evaluated as follows

$$\begin{aligned} \mathbb{E}[\hat{\alpha}_j] &= \mathbb{E} \left[\frac{c}{n} \sum_{i=1}^n \frac{\phi_j^*(X_i)}{p_X(X_i)} B_i \right] \\ &= \frac{c}{n} \sum_{i=1}^n \mathbb{E} \left[\frac{\phi_j^*(X_i)}{p_X(X_i)} \text{sgn}(f(X_i) + Z_i - T_i) \right] \\ &= c \mathbb{E} \left[\frac{\phi_j^*(X_1)}{p_X(X_1)} \text{sgn}(f(X_1) + Z_1 - T_1) \right], \quad (28) \end{aligned}$$

where the last equality follows since the terms are iid. This last expectation can be evaluated as follows

$$\begin{aligned}
& \mathbb{E} \left[\frac{\phi_j^*(X_1)}{p_X(X_1)} \text{sgn}(f(X_1) + Z_1 - T_1) \right] \\
&= \int_{\mathcal{D}} p_X(x) \int_{-b}^{+b} p_Z(z) \int_{-c}^{+c} \frac{1}{2c} \frac{\phi_j^*(x)}{p_X(x)} \text{sgn}(f(x) + z - t) dt dz dx \\
&= \int_{\mathcal{D}} \int_{-b}^{+b} p_Z(z) \phi_j^*(x) \frac{1}{2c} \left(\int_{-c}^{f(x)+z} dt - \int_{f(x)+z}^{+c} dt \right) dz dx \\
&= \frac{1}{c} \int_{\mathcal{D}} \int_{-b}^{+b} p_Z(z) \phi_j^*(x) (f(x) + z) dz dx \\
&= \frac{1}{c} \int_{\mathcal{D}} \phi_j^*(x) f(x) dx \\
&= \frac{1}{c} \langle f, \phi_j \rangle = \frac{\alpha_j}{c}, \tag{29}
\end{aligned}$$

where the second to last line follows from the assumption that p_Z is a zero-mean distribution. Combining (28) and (29), we have (25). (ii) Thus,

$$\mathbb{E}[|\hat{\alpha}_j - \alpha_j|^2] = \mathbb{E}[|\hat{\alpha}_j - \mathbb{E}[\hat{\alpha}_j]|^2] = \text{Var}[\hat{\alpha}_j]. \tag{30}$$

Using standard properties of variance and the fact that the terms $\{\phi_j^*(X_i)B_i/p_X(X_i)\}_{i=1}^n$ are iid, we obtain the following

$$\begin{aligned}
\text{Var}[\hat{\alpha}_j] &= \frac{c^2}{n^2} \sum_{i=1}^n \text{Var} \left[\frac{\phi_j^*(X_i)}{p_X(X_i)} B_i \right] \\
&= \frac{c^2}{n} \text{Var} \left[\frac{\phi_j^*(X_1)}{p_X(X_1)} B_1 \right] \\
&= \frac{c^2}{n} \mathbb{E} \left[\left| \frac{\phi_j^*(X_1)}{p_X(X_1)} B_1 \right|^2 \right] - \frac{c^2}{n} \left| \mathbb{E} \left[\frac{\phi_j^*(X_1)}{p_X(X_1)} B_1 \right] \right|^2 \\
&\leq \frac{c^2}{n} \mathbb{E} \left[\left| \frac{\phi_j^*(X_1)}{p_X(X_1)} \right|^2 \right], \\
&= \frac{c^2}{n} \int_{\mathcal{D}} \frac{|\phi_j(x)|^2}{p_X^2(x)} p_X(x) dx, \\
&= \frac{c^2}{n} \int_{\mathcal{D}} \frac{|\phi_j(x)|^2}{p_X(x)} dx. \tag{31}
\end{aligned}$$

Combining (30) and (31), we arrive at (26). (iii) The coefficient estimates

$$\begin{aligned}
\hat{\alpha}_j &= \frac{c}{n} \sum_{i=1}^n \frac{\phi_j^*(X_i)}{p_X(X_i)} B_i \\
&= \frac{c}{n} \sum_{i=1}^n \frac{\phi_j^*(X_i)}{p_X(X_i)} \text{sgn}(f(X_i) + Z_i - T_i) \\
&\xrightarrow{\text{a.s.}} c \mathbb{E} \left[\frac{\phi_j^*(X_1)}{p_X(X_1)} \text{sgn}(f(X_1) + Z_1 - T_1) \right], \tag{32}
\end{aligned}$$

as $n \rightarrow \infty$ by Kolmogorov's strong law of large numbers since each term in the summation is iid and has a first moment bounded by $\sqrt{\text{vol}(\mathcal{D})}$:

$$\mathbb{E} \left[\left| \frac{\phi_j(X_1)}{p_X(X_1)} \right| \right] = \|\phi_j\|_1 \leq \sqrt{\text{vol}(\mathcal{D})} \|\phi_j\|_2 = \sqrt{\text{vol}(\mathcal{D})},$$

where the last inequality follows from the Cauchy-Schwartz inequality. Combining (29) and (32), we obtain (27), concluding the proof of the Lemma 5.1. \blacksquare

For any orthonormal basis $\mathcal{B} = \{\phi_j\}_{j=0}^{\infty}$ and for any field $f \in \mathcal{F}$, the integrated MSE of the estimate can be written as follows

$$\begin{aligned}
D &= \mathbb{E}[\|f - \hat{f}_{n,m}\|^2] \\
&= \mathbb{E} \left[\left\| \sum_{j=0}^{\infty} \alpha_j \phi_j - \sum_{j=0}^{m-1} \hat{\alpha}_j \phi_j \right\|^2 \right] \\
&= \sum_{j=0}^{m-1} \mathbb{E}[|\hat{\alpha}_j - \alpha_j|^2] + \sum_{j=m}^{\infty} |\alpha_j|^2 \\
&\leq \frac{c^2}{n} \sum_{j=0}^{m-1} \int_{\mathcal{D}} \frac{|\phi_j(x)|^2}{p_X(x)} dx + \underbrace{\sum_{j=m}^{\infty} |\alpha_j|^2}_{=:\varepsilon[f,m,\mathcal{B}]}, \tag{33}
\end{aligned}$$

where in the last step we used the bound given in (26). Thus we have (7), concluding the proof of Theorem 3.1. \blacksquare

B. Proof of Theorem 3.2

The pointwise errors of the field estimate with respect to the m -term approximation can be written as

$$\begin{aligned}
S_n(x) &:= \hat{f}_n(x) - f_{m(n)}(x) \\
&= \sum_{j=0}^{m(n)-1} \left(\frac{1}{n} \sum_{i=1}^n \frac{c\phi_j^*(X_i)B_i}{p_X(X_i)} - \alpha_j \right) \phi_j(x) \\
&= \frac{1}{n} \sum_{i=1}^n \sum_{j=0}^{m(n)-1} \left(\frac{c\phi_j^*(X_i)B_i}{p_X(X_i)} - \alpha_j \right) \phi_j(x) \\
&= \frac{1}{n} \sum_{i=1}^n U_i(x),
\end{aligned}$$

where for $i \in \{1, \dots, n\}$,

$$U_i(x) := \sum_{j=0}^{m(n)-1} \left(\frac{c\phi_j^*(X_i)B_i}{p_X(X_i)} - \alpha_j \right) \phi_j(x).$$

Note that $U_i(x)$ is iid and that it is zero-mean due to (25) of Lemma 5.1. However, almost sure convergence of $S_n(x)$ cannot be directly deduced from the standard strong law of large numbers since the distribution of $U_i(x)$ itself depends on n because it is the summation of $m(n)$ terms. Instead, we leverage a more fundamental condition for almost sure convergence [24, p. 206]: if for all $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \mathbb{P}[|S_n(x)| \geq \epsilon] < \infty,$$

then $S_n(x) \xrightarrow{\text{a.s.}} 0$ as $n \rightarrow \infty$.

Associated with $S_n(x)$, is a martingale $\{V_k(x)\}_{k=0}^n$ given by $V_0 := 0$, and for $k \in \{1, \dots, n\}$,

$$V_k(x) := \sum_{i=1}^k \frac{1}{n} U_i(x).$$

$V_0(x), \dots, V_n(x)$ is a martingale since $\{U_i(x)\}_{i=1}^n$ is iid with zero-mean. Note that $V_n(x) = S_n(x)$ and that $|V_k(x) - V_{k-1}(x)| \leq |\frac{1}{n}U_k(x)|$. For each $i \in \{1, \dots, n\}$,

$$|U_i(x)| \leq c \left| \sum_{j=0}^{m(n)-1} \frac{\phi_j(x)\phi_j^*(X_i)}{p_X(X_i)} \right| + \left| \sum_{j=0}^{m(n)-1} \alpha_j \phi_j(x) \right|,$$

by the triangle inequality. Under the assumptions that the conditions given by (12) and (13) hold and that the deployment distribution p_X is non-singular, there exists some constant $C > 0$ such that for all $i \in \{1, \dots, n\}$,

$$|U_i(x)| \leq C\Lambda_{m(n)},$$

with probability 1 for $x \in \mathcal{D}$ almost everywhere. Thus

$$|V_k(x) - V_{k-1}(x)| \leq \frac{C\Lambda_{m(n)}}{n}. \quad (34)$$

According to the Azuma–Hoeffding inequality (see [25, p. 303]), if for all $k \in \{1, \dots, n\}$, $|V_k(x) - V_{k-1}(x)| \leq C_k$, then for all $\epsilon > 0$,

$$\mathbb{P}[|V_n(x)| \geq \epsilon] \leq 2 \exp\left(\frac{-\epsilon^2}{2 \sum_{k=1}^n C_k^2}\right).$$

Applying this inequality with $C_k = C\Lambda_{m(n)}/n$ for all $k \in \{1, \dots, n\}$ (see (34)) and $V_n(x) = S_n(x)$ we obtain the following upper bound

$$\begin{aligned} \mathbb{P}[|S_n(x)| \geq \epsilon] &\leq 2 \exp\left(\frac{-\epsilon^2}{2 \sum_{k=1}^n \frac{C^2 \Lambda_{m(n)}^2}{n^2}}\right) \\ &= 2 \exp\left(\frac{-\epsilon^2 n}{2C^2 \Lambda_{m(n)}^2}\right), \end{aligned}$$

for $x \in \mathcal{D}$ almost everywhere. Therefore,

$$\sum_{n=1}^{\infty} \mathbb{P}[|S_n(x)| \geq \epsilon] \leq \sum_{n=1}^{\infty} 2 \exp\left(\frac{-\epsilon^2 n}{2C^2 \Lambda_{m(n)}^2}\right),$$

which is less than infinity $\forall \epsilon > 0$ and $x \in \mathcal{D}$ almost everywhere, due to condition (14). Thus, as $n \rightarrow \infty$, almost surely $S_n(x) \rightarrow 0$, for $x \in \mathcal{D}$ almost everywhere. ■

C. Proof of Proposition 3.1

Part (i): Note that if $|g_\infty(x, y)| = 0$ for $x, y \in \mathcal{D}$ almost everywhere and $|h_\infty(x)| = 0$ for all $f \in \mathcal{F}$ and for $x \in \mathcal{D}$ almost everywhere, then conditions (12) and (13) hold with

some constants $C_1, C_2 > 0$. For g_m , we can write

$$\begin{aligned} &\iint_{\mathcal{D} \times \mathcal{D}} p_X^2(y) |g_\infty(x, y)|^2 dx dy = \\ &= \iint_{\mathcal{D} \times \mathcal{D}} \liminf_{m \rightarrow \infty} p_X^2(y) |g_m(x, y)|^2 dx dy \\ &\stackrel{(a)}{\leq} \liminf_{m \rightarrow \infty} \iint_{\mathcal{D} \times \mathcal{D}} p_X^2(y) |g_m(x, y)|^2 dx dy \\ &= \liminf_{m \rightarrow \infty} \iint_{\mathcal{D} \times \mathcal{D}} p_X^2(y) \frac{1}{\Lambda_m^2} \sum_{j=0}^{m-1} \sum_{k=0}^{m-1} \phi_j(x) \phi_j^*(y) \cdot \\ &\quad \phi_k^*(x) \phi_k(y) \frac{1}{p_X^2(y)} dx dy \\ &= \liminf_{m \rightarrow \infty} \frac{1}{\Lambda_m^2} \sum_{j=0}^{m-1} \sum_{k=0}^{m-1} \underbrace{\int_{\mathcal{D}} \phi_j(x) \phi_k^*(x) dx}_{=\delta_{j-k}} \cdot \\ &\quad \underbrace{\int_{\mathcal{D}} \phi_k(y) \phi_j^*(y) dy}_{=\delta_{j-k}} \\ &= \liminf_{m \rightarrow \infty} \frac{1}{\Lambda_m^2} \sum_{j=0}^{m-1} 1 = \liminf_{m \rightarrow \infty} \frac{m}{\Lambda_m^2}, \end{aligned}$$

where the inequality (a) is due to Fatou's lemma [19] and δ_k is the Kronecker delta function. Thus for Λ_m such that (17) is satisfied we have that

$$\iint_{\mathcal{D} \times \mathcal{D}} p_X^2(y) |g_\infty(x, y)|^2 dx dy = 0,$$

which implies that $|g_\infty(x, y)| = 0$ for $x, y \in \mathcal{D}$ almost everywhere due to (10). For h_m , we can write

$$\begin{aligned} \int_{\mathcal{D}} |h_\infty(x)|^2 dx &= \int_{\mathcal{D}} \liminf_{m \rightarrow \infty} |h_m(x)|^2 dx \\ &\leq \liminf_{m \rightarrow \infty} \int_{\mathcal{D}} |h_m(x)|^2 dx \\ &= \liminf_{m \rightarrow \infty} \int_{\mathcal{D}} \frac{1}{\Lambda_m^2} \sum_{j=0}^{m-1} \sum_{k=0}^{m-1} \alpha_j \alpha_k^* \phi_j(x) \phi_k^*(x) dx \\ &= \liminf_{m \rightarrow \infty} \frac{1}{\Lambda_m^2} \sum_{j=0}^{m-1} \sum_{k=0}^{m-1} \alpha_j \alpha_k^* \underbrace{\int_{\mathcal{D}} \phi_j(x) \phi_k^*(x) dx}_{=\delta_{j-k}} \\ &= \liminf_{m \rightarrow \infty} \frac{1}{\Lambda_m^2} \sum_{j=0}^{m-1} |\alpha_j|^2 \\ &\leq \liminf_{m \rightarrow \infty} \frac{\|f\|^2}{\Lambda_m^2} \\ &\leq \liminf_{m \rightarrow \infty} \frac{a^2 \text{vol}(\mathcal{D})}{\Lambda_m^2}, \quad \forall f \in \mathcal{F}, \end{aligned}$$

where the first inequality follows from Fatou's lemma [19] and the last inequality is due to f being amplitude-bounded by a over the support \mathcal{D} . Thus for Λ_m such that (17) is satisfied, we have that

$$\forall f \in \mathcal{F}, \quad \int_{\mathcal{D}} |h_\infty(x)|^2 dx = 0,$$

which implies that $|h_\infty(x)| = 0$ for all $f \in \mathcal{F}$ and for $x \in \mathcal{D}$ almost everywhere.

Part (ii): Applying the triangle inequality, we can write

$$\begin{aligned} & \left| \sum_{j=0}^{m-1} \phi_j(x) \phi_j^*(y) \frac{1}{p_X(y)} \right| \\ & \leq \sum_{j=0}^{m-1} \frac{|\phi_j(x)| |\phi_j^*(y)|}{p_X(y)} \\ & \leq \sum_{j=0}^{m-1} \frac{\beta^2}{\nu} = \frac{\beta^2}{\nu} m = C_1 \Lambda_m, \end{aligned}$$

which shows that condition (12) is satisfied for $\Lambda_m = m$ and $C_1 = \beta^2/\nu$. Again, applying the triangle inequality, we can write

$$\begin{aligned} & \left| \sum_{j=0}^{m-1} \langle f, \phi_j \rangle \phi_j(x) \right| \\ & \leq \sum_{j=0}^{m-1} |\langle f, \phi_j \rangle| |\phi_j(x)| \\ & \leq \sum_{j=0}^{m-1} \|f\| \beta = m \|f\| \beta \\ & \leq m a \beta \sqrt{\text{vol}(\mathcal{D})} = C_2 \Lambda_m, \end{aligned}$$

which shows that condition (13) is satisfied for $\Lambda_m = m$ and $C_2 = a\beta\sqrt{\text{vol}(\mathcal{D})}$. ■

VI. CONCLUDING REMARKS

The principal contribution of this work is a systematic treatment of (i) binary–sensing, (ii) random sensor deployment, and (iii) unknown observation noise distribution for high–resolution distributed sensing and estimation of multidimensional fields using dense sensor networks. A key finding of this work is that the rate of convergence of the integrated MSE for field estimation is extremely robust to the apparent limitations of ultra–poor sensing precision, random sensor deployment, and lack of knowledge of observation noise statistics. In some cases, the convergence rate exactly matches the minimax rate of convergence with infinite–precision real–valued samples and known noise statistics. Interesting directions for future work include (i) establishing the exact rate of convergence of the integrated MSE and a central limit theorem for the estimate, (ii) analysis of the sensitivity of the integrated–MSE to sensor location uncertainty, (iii) unbounded–amplitude signal and noise models, and (iv) general dither distributions.

ACKNOWLEDGMENT

The authors would like to thank Professor Elias Masry, Department of Electrical and Computer Engineering at the University of California San Diego, for his encouraging comments.

REFERENCES

- [1] Y. Wang and P. Ishwar, “On non-parametric field estimation using randomly deployed, noisy, binary sensors,” in *Proc. IEEE International Symposium on Information Theory*, Nice, France, Jun. 2007.
- [2] Y. Wang, N. Ma, M. Zhao, P. Ishwar, and V. Saligrama, “On Universal Distributed Estimation of Noisy Fields with One-bit Sensors,” in *Proc. Allerton Conference on Communications, Control, and Computing*, Sep. 2006.
- [3] E. Masry, “The reconstruction of analog signals from the sign of their noisy samples,” *IEEE Trans. Info. Theory*, vol. IT–27, pp. 735–745, Nov. 1981.
- [4] E. Masry and S. Cambanis, “Consistent estimation of continuous–time signals from nonlinear transformations of noisy samples,” *IEEE Trans. Info. Theory*, vol. IT–27, pp. 84–96, Jan. 1981.
- [5] Z. Q. Luo, “Universal decentralized estimation in a bandwidth constrained sensor network,” *IEEE Trans. Info. Theory*, vol. IT–51, pp. 2210–2219, Jun. 2005.
- [6] M. Gastpar and M. Vetterli, “Source–channel communication in sensor networks,” *Lecture Notes in Computer Science*, vol. 2634, pp. 162–177, Apr. 2003.
- [7] M. Gastpar, B. Rimoldi, and M. Vetterli, “To Code, or not to code: Lossy source–channel communication revisited,” *IEEE Trans. Info. Theory*, vol. IT–49, pp. 1147–1158, May 2003.
- [8] R. Nowak, U. Mitra, and R. Willet, “Estimating inhomogenous fields using wireless sensor networks,” *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 999–1006, Aug. 2004.
- [9] K. Liu, H. El–Gamal, and A. Sayeed, “On optimal parametric field estimation in sensor networks,” in *Proc. IEEE/SP 13th Workshop on Statistical Signal Processing*, Jul. 2005, pp. 1170–1175.
- [10] W. Bajwa, A. Sayeed, and R. Nowak, “Matched source–channel communication for field estimation in wireless sensor networks,” in *Proc. Fourth Intl. Symposium Information Processing in Sensor Networks*, Apr. 2005, pp. 332–339.
- [11] N. Liu and S. Ulukus, “Optimal distortion–power tradeoffs in Gaussian sensor networks,” in *Proc. IEEE International Symposium on Information Theory*, Seattle, WA, USA, Jul. 2006, pp. 1534–1538.
- [12] M. Gastpar, “Uncoded transmission is exactly optimal for a simple Gaussian sensor network,” in *Proc. Information Theory and Applications Workshop*, Jan. 29–Feb. 2, 2007.
- [13] T. Berger, Z. Zhang, and H. Viswanathan, “The CEO problem [multiterminal source coding],” *IEEE Trans. Info. Theory*, vol. IT–42, pp. 887–902, May. 1996.
- [14] H. Viswanathan and T. Berger, “The quadratic Gaussian CEO problem,” *IEEE Trans. Info. Theory*, vol. IT–43, pp. 1549–1559, Sep. 1997.
- [15] V. Prabhakaran, D. Tse, and K. Ramchandran, “Rate–region of the quadratic Gaussian CEO problem,” in *Proc. IEEE International Symposium on Information Theory*, Chicago, IL, Jun. 2004, p. 119.
- [16] Z. Cvetković, “Resilience properties of redundant expansions under additive noise and quantization,” *IEEE Trans. Info. Theory*, vol. IT–49, pp. 644–656, Mar. 2003.
- [17] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, 1998.
- [18] K. Knopp, *Theory and Application of Infinite Series*. Dover Publications, 1990.
- [19] W. Rudin, *Real and Complex Analysis*, 2nd ed. McGraw–Hill, 1974.
- [20] L. Carleson, “On convergence and growth of partial sums of Fourier series,” *Acta Mathematica*, vol. 116, no. 1, pp. 135–157, Dec. 1966.
- [21] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, 1st ed. Upper Saddle River, NJ: Prentice–Hall, 1993.
- [22] A. P. Korostelev and A. B. Tsybakov, *Minimax Theory of Image Reconstruction*. Springer Verlag, New York, 1993.
- [23] G. Golubev and M. Nussbaum, “A Risk Bound in Sobolev Class Regression,” *The Annals of Statistics*, vol. 18, no. 2, pp. 758–778, 1990.
- [24] R. Ash, *Basic Probability Theory*. John Wiley and Sons, 1970.
- [25] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.