

On Distributed Sampling in Dense Sensor Networks: a “bit-conservation” principle

Prakash Ishwar, Animesh Kumar, and Kannan Ramchandran
Department of Electrical Engineering and Computer Sciences
University of California at Berkeley
{ishwar, animesh, kannanr}@eecs.berkeley.edu

Abstract

We address the problem of sampling bandlimited sensor fields in a distributed, limited-precision, communication-constrained, processing environment, where it is desired for a central intelligent unit to reconstruct the sensor field to a desired accuracy. We show the feasibility of having a flexible tradeoff between the oversampling rate (sensor density) and the Analog to Digital (A/D) quantizer precision while achieving an exponential accuracy in the number of bits per Nyquist-interval. This exposes a key underlying “*conservation of bits*” principle. That is, the bit budget per Nyquist-interval (the rate) can be distributed along the amplitude-axis (sensor precision) and space (sensor density) in an almost arbitrary discrete-valued manner, while retaining the same error-rate characteristics. Interestingly, this is possible in a highly localized communication setting with only nearest-neighbor communication, making it well suited for dense sensor networks operating under stringent inter-node communication constraints. Our analysis also leads to an understanding of the “information-density” of sensor fields: if D is the desired distortion, the number of bits per Nyquist-interval (or equivalently, the logarithm of the number of fixed-precision sensors N) grows as $\log(1/D)$. Alternatively, the (mean square) distortion goes down as $1/N^2$ while the information per Nyquist-interval increases as $\log N$. The bits per sensor, however, goes down as $\frac{1}{N} \log N$. For a fixed, *nonzero* target distortion, the number of fixed-precision sensors needed is always *finite*.

I. INTRODUCTION

Remote sensing of physical phenomena of interest using an embedded network of sensors has aspects that relate intimately to the classical problem of sampling of continuous signals, a mature topic in signal processing that has accumulated a rich knowledge-base over the past several decades [2], [3]. Although this analogy with sampling theory is striking, the sensor network scenario brings with it a curious set of attributes that impose challenging constraints on the classical sampling paradigm. These attributes arise from the physical constraints of the sensing devices (low-precision, low-power, unreliable) as well as the communication constraints (low range, multihop) associated with the network in which they live. This paper is accordingly driven by the goal of addressing the sampling problem within the context of sensor networks and their associated constraints in terms of both computation and communication.

At an information processing level, one can broadly classify many sensor network problems of interest as consisting of two functional tasks: (i) information sensing/data-acquisition or sampling of the sensor field, and (ii) information transport, where the information is disseminated across a network or to a data-gathering unit. While both aspects are important (and inter-twined), this paper will focus primarily on the first task, suggesting and analyzing a novel “distributed” sampling framework that is well suited for sensor networks. An in-depth analysis of the information transport necessarily involves an accurate modeling of the communication channel and devising efficient network management protocols which are beyond the scope of this work.

Consider the scenario where a large number of sensors are deployed over a region of interest in order to collect and return sensor measurements to a central processing unit (CPU) with the goal of reconstructing the sensor field to maximum accuracy. Many physical signals are approximately bandlimited (the physical channel propagation laws often provide a natural smoothing effect that attenuates high frequencies) and can be reconstructed in a stable manner from samples taken slightly above the Nyquist-rate on a uniform lattice. In practice, however, the samples of the signal are quantized due to the finite precision of A/D converters, leading to unavoidable signal reconstruction errors. When signals are uniformly sampled at the critical Nyquist-rate, which we will refer to as the Nyquist-sampling setup, the reconstruction error decays exponentially with the bit-rate (measured in number of bits per Nyquist-interval¹) of the A/D converters [4]. Hence, increasingly precise sensors are needed to achieve increasingly precise field reconstructions. However, high-precision A/D operations are expensive (this is true even outside the sensor network world). This leads to the following question: is it possible to trade-off A/D converter resolution in terms of bits per sample for (average) oversampling rate (attained through denser oversampling but using poor, fixed-resolution sensors) while maintaining the same reconstruction error performance as a function of the number of bits per Nyquist-interval? In the sequel we answer this question in the affirmative, and show how it is possible to compensate for lack of precision in the A/D elements via spatial oversampling without compromising the asymptotic reconstruction accuracy. Even with one-bit sensors, arbitrary accuracy can be attained by oversampling sufficiently.

Our main contributions in this work are two-fold:

- (i) “*Bit-conservation*” principle: First, we tackle the sensor acquisition (sampling) problem where we identify fundamental tradeoffs between sensor A/D precision and sensor-field reconstruction accuracy, uncovering in the process a key “*conservation of bits*” principle that underlies sensor fields. While this applies even in the classical sampling scenario (i.e., devoid of any communication constraints), we further relate this to a highly distributed processing scenario with a novel “bit-rippling” protocol that is highly communication-efficient and integrated into the sampling algorithm.
- (ii) “*Information density*”: Secondly, our analysis provides fundamental insights into some relevant “scaling laws” that dictate how performance (sensor field reconstruction accuracy at a centralized data-collecting unit) scales with the size/density of the sensor network. This helps clarify the notion of “information density” of sensor-fields.

This work has several important consequences for sensor-networks:

A. The pioneering work of Gupta and Kumar in [5] has led to recent interest in scaling laws of ad hoc networks. These scaling laws however hold in a regime where the network nodes produce “independent” data, no matter what the scale of the network. This independent data model is obviously at odds with the sensor network scenario, where the “information density” (related to the sensor field) remains fixed regardless of the “network density” (related to the network size). This results in an easier network requirement as the data is highly correlated, which can be exploited in a variety of ways [6], [1], [7]. Specifically, we show that if D is the desired distortion, the number of bits per Nyquist-interval (or equivalently, the logarithm of the number of fixed-precision sensors N) grows as $\log(1/D)$. Alternatively, the (mean square) distortion goes down as $1/N^2$ while the information per Nyquist-interval increases as $\log N$. The bits per sensor, however, goes down as $\frac{1}{N} \log N$.

We would like to note that a scaling law (i.e., growth of information per unit length with increasing sensor density) comes into effect only when the desired distortion is zero, i.e. when one considers the high-quality or high-rate asymptotics. For a fixed *non-zero* target

¹We are interested in reconstructing the field over an *infinite* region and accordingly the scaling laws (bits, sensors, precision) would be in terms of density, i.e., per unit length (for 1-D signals).

distortion, as will become clear in the sequel, the number of fixed-precision sensors needed to achieve the target distortion is always *finite*². This is exactly in the spirit of what was conjectured recently by Marco et al. in [8] when their analysis (which applies to a very broad class of signals including amplitude-limited and bandlimited signals) lead them to a rather pessimistic conclusion wherein, even with the number of sensors going to infinity, the distortion did not go to zero. This underscores the importance of using the “right kind” of sensor data acquisition and transport framework in order to fully exploit the underlying correlation in sensor-fields.

B. A direct fallout of the “bit-conservation” principle is the ability to do spatially adaptive sampling: we can have critically sampled Nyquist sampling using higher resolution A/D converters when the sampling density is forced to be light (i.e. near the Nyquist-rate), e.g., due to terrain difficulties in the placement of sensors, and use proportionately lower resolution sensors when the sampling density can be high. The bit-conservation principle also implies a certain degree of robustness to node failures. *Node failures reduce the average sampling density and have the same effect as loss of amplitude resolution.* This leads to a *graceful degradation* of reconstruction quality with node failures. This also has bearing on A/D precision versus inter-sensor communication cost tradeoffs. Densely spaced low-resolution sensors would need to communicate fewer bits while sparsely spaced high-resolution sensors need to to communicate more bits. However, the total number of bits exchanged in a Nyquist-interval (bit-meters) would be about the same.

C. A side-benefit (that we do not emphasize in this work) is that there is a measure of “security” against eavesdropping provided by the specific data-acquisition strategy proposed in our work (namely, the use of a dithering function which can be kept secret).

Related work: The data acquisition component of this paper is founded on the deterministic 1-bit A/D work of Cvetkovic and Daubechies [9] that represents the current state-of-the-art in the classical oversampling literature. Our work builds on this in three ways: first, we generalize their results to arbitrary precision A/D; secondly, we extend their framework to the stochastic setting; and thirdly, we “port” these results from the classical centralized setup to the distributed setting that is more relevant to the sensor network problem. The scaling laws that result from our bit conservation principle are inspired by the seminal work on scaling laws by Gupta and Kumar, which however, as was previously mentioned, applies in an independent data network setting unlike ours. Other recent works addressing scaling laws applicable to sensor networks include the works of Scaglione and Servetto [6] and Marco et al. [8].

II. SAMPLING BANDLIMITED SIGNALS

A. Sensor-field models

Deterministic model: Let $f(t)$, $t \in \mathbb{R}$ be a deterministic, finite-energy, bandlimited signal. Without loss of generality (WLOG) in what follows, we will assume that the dynamic range of f is $[-1, 1]$ and the spectral support is contained in $[-\pi, \pi]$.

Stochastic model: Correspondingly, let $X(t)$, $t \in \mathbb{R}$ be a wide-sense stationary (WSS) stochastic process (w.r.t. some common probability space $(\Omega, \mathcal{F}, \mathcal{P})$) with a square-integrable, autocorrelation function $R_X(t)$, $t \in \mathbb{R}$ and the power spectral density (p.s.d.) $S_X(\omega)$, $\omega \in \mathbb{R}$ that is bandlimited to $[-W, W]$, with $0 < W < \infty$. We further assume that $X := \{X(t)\}_{t \in \mathbb{R}}$ is amplitude limited, i.e., $|X(t)| \leq A$, $\forall t \in \mathbb{R}$ where $A < \infty$ is some nonnegative constant.

The following proposition alleviates the technical difficulties of dealing with random processes, and helps us to develop the theory for processes using the same ideas as for deterministic signals,

²In practice, it might be undesirable to try to recover the sensor-field with *zero* distortion due to the presence of ambient noise.

Proposition 2.1 [10] *For every WSS, amplitude-limited, bandlimited process, X with a square-integrable autocorrelation function R_X and every $W' \in (W, \infty)$, there exists a jointly WSS, amplitude-limited, bandlimited process X_{BL} with the following properties:*

- (i) $|X_{BL}(t)| \leq A_{BL}$, where $A_{BL} > A$ is some constant that depends on W, W' , and A ,
- (ii) $X_{BL}(t)$ is differentiable everywhere and $|X'_{BL}(t)|$ is uniformly bounded,
- (iii) $R_{X_{BL}}(t) = R_X(t), \forall t \in \mathbb{R}$, and hence X and X_{BL} are energy equivalent: $X(t) \stackrel{\mathcal{L}^2}{=} X_{BL}(t), \forall t \in \mathbb{R}$, where $\stackrel{\mathcal{L}^2}{=}$ stands for mean-square (m.s.) equivalence,
- (iv) X and X_{BL} agree with probability one on any countable set, i.e., for any countable subset of real-valued random variables: $\{t_l\}_{l \in \mathbb{Z}}, \mathcal{P}(\{\exists l \in \mathbb{Z} : X(t_l) \neq X_{BL}(t_l)\}) = 0$.

Hence, by a suitable normalization of the amplitude and spatial axes we may assume WLOG that $W' = \pi$ and $A_{BL} = 1$. Proposition 2.1 helps in treating X exactly like its deterministic counterpart f when the reconstruction quality is measured by the mean³ squared error (MSE). Hence the machinery developed for analyzing deterministic signals (that of Section II-C in particular) will carry over almost without any change to the stochastic setting of interest.

1) *Distortion criteria:* The performance of different sampling schemes will be characterized by deriving rate-dependent upper bounds for the worst-case/mean (\mathcal{L}^p) error.

Deterministic Case: For deterministic signals, the worst-case, pointwise, signal reconstruction error, $\|e(t)\|_\infty$, will be bounded. These bounds will continue to remain valid for the so-called local average⁴ \mathcal{L}^p norm of the reconstruction error $e(t)$ defined by: $\|e\|_p(t, T) := \left\{ \frac{1}{T} \int_{|s-t| < T/2} |e(s)|^p ds \right\}^{\frac{1}{p}}, p \in (0, \infty)$, because $\|e\|_p(t, T) \leq \|e\|_\infty, \forall t \in \mathbb{R}, T > 0$.

Stochastic Case: For stochastic signals, worst case pointwise MSE, i.e., $\|E(e^2(t))\|_\infty$ will be bounded. As will be demonstrated in the sequel, similar bounds for the mean \mathcal{L}^p error can be obtained through an application of the following elementary proposition for all $p \in (0, \infty)$.

Proposition 2.2 *If $X \stackrel{\mathcal{L}^q}{=} \tilde{X}$ for any positive q , i.e., $E|X - \tilde{X}|^q = 0$, then (i) for any positive p , $X \stackrel{\mathcal{L}^p}{=} \tilde{X}$ and (ii) $E|X - Y|^p = E|\tilde{X} - Y|^p$ for any Y .*

Proposition 2.2 is key to conveniently deriving bounds on the mean \mathcal{L}^p energy of the reconstruction error $(X - Y)$ through the agency of $\tilde{X} = X_{BL}$ of Proposition 2.1, for which error bounds are straightforward to evaluate using results for the deterministic setting.

B. Nyquist-sampling at near-Nyquist-rates

Deterministic Case: $f(t)$ can be reconstructed from the samples $\{f(l)\}_{l \in \mathbb{Z}}$ according to the well known interpolation formula [2]:

$$f(t) = \lim_{L \rightarrow \infty} \sum_{l=-L}^L f(l) \text{sinc}(t-l), \quad \forall t \in \mathbb{R}, \quad (1)$$

where $\text{sinc}(t) := \frac{\sin(\pi t)}{\pi t}$ for $t \neq 0$ and $\text{sinc}(0) := 1$. The series in (1) is not absolutely convergent which makes the reconstruction unstable in presence of bounded perturbations of the samples. However the instability can be overcome by taking samples slightly above the Nyquist-rate:

Fact 2.3 [11] *For each $\lambda > T_{NQ} = 1$, there exists an absolutely integrable kernel $\phi_\lambda(t)$*

³Here, mean stands for mathematical expectation.

⁴Here the average is over space and not with respect to any probability distribution.

bandlimited to $[-\pi\lambda, \pi\lambda]$ such that $C_\lambda := \sup_{t \in \mathbb{R}} \left(\sum_{l \in \mathbb{Z}} |\phi_\lambda(t - \frac{l}{\lambda})| \right) < \infty$ and

$$f(t) = \lim_{L \rightarrow \infty} \sum_{l=-L}^L f\left(\frac{l}{\lambda}\right) \frac{1}{\lambda} \phi_\lambda\left(t - \frac{l}{\lambda}\right), \quad \forall t \in \mathbb{R}, \quad (2)$$

where the convergence of the series in (2) holds pointwise (and in \mathcal{L}^2), absolutely, and uniformly on all compact subsets of \mathbb{R} .

Let $Q_k(\cdot)$ denote the k -bit uniform scalar quantizer on $[-1, 1]$ [12] so that $|z - Q_k(z)| \leq 2^{-k}$ for all $|z| < 1$. The bitrate in bits per Nyquist interval used to quantize the signal is given by $R = k\lambda$. Let the bitrate- R quantized Nyquist reconstruction be defined as $\widehat{f}_R^{\text{NQ}}(t) := \lim_{L \rightarrow \infty} \sum_{l=-L}^L Q_k(f(l/\lambda)) \frac{1}{\lambda} \phi\left(t - \frac{l}{\lambda}\right)$. Using Fact 2.3 and the triangle inequality, it immediately follows that

$$\|f - \widehat{f}_R^{\text{NQ}}\|_\infty \leq \frac{C_\lambda}{\lambda} \cdot 2^{-\frac{1}{\lambda}R}. \quad (3)$$

Stochastic Case: The following propositions are the stochastic counterparts of the deterministic case,

Proposition 2.4 For each $\lambda > T_{\text{NQ}} = 1$, $p \in (0, \infty)$, and $\phi_\lambda(\cdot)$ as in Fact 2.3,

$$X(t) \stackrel{\mathcal{L}^p}{=} \lim_{L \rightarrow \infty} \sum_{l=-L}^L X\left(\frac{l}{\lambda}\right) \frac{1}{\lambda} \phi_\lambda\left(t - \frac{l}{\lambda}\right), \quad \forall t \in \mathbb{R},$$

where the mean- \mathcal{L}^p -convergence of the above series holds pointwise, absolutely, and uniformly on all compact subsets of \mathbb{R} .

This is in fact a special case of the following more general result whose proof is outlined in [10]

Proposition 2.5 [10] Let $\{t_l \in \mathbb{R}\}_{l \in \mathbb{Z}}$ be a set of points such that $\sup_l |t_l - \frac{l}{\lambda}| < C < \infty$ and $\delta := \inf_{j,l,j \neq l} |t_j - t_l| > 0$ then there exist a set of interpolating functions $\{\psi_l(t)\}_{l \in \mathbb{Z}}$ such that

$$X(t) \stackrel{\mathcal{L}^p}{=} \lim_{L \rightarrow \infty} \sum_{l=-L}^L X(t_l) \psi_l(t - t_l),$$

where the mean- \mathcal{L}^p -convergence of the above series holds pointwise, absolutely, and uniformly on all compact subsets of \mathbb{R} .

Let the bitrate- R quantized Nyquist reconstruction of X be defined as in the deterministic case. It is easy to show, using Propositions 2.1 and 2.4, that

$$\|E|X - \widehat{X}_R^{\text{NQ}}|^p\|_\infty \leq \left(\frac{C_\lambda}{\lambda}\right)^p 2^{-\frac{p}{\lambda}R}. \quad (4)$$

C. 1-bit A/D dithered oversampling

The reconstruction accuracy for Nyquist sampling set-up can be improved only by improving the precision of A/D converters of the sensors. However, sensors are low precision devices available in large volumes. In this section, we study the sampling of bandlimited fields using *many* 1-bit sensors. Samples collected using 1-bit sensors might be available at non-uniform locations. This motivates the need for the following result from non-harmonic Fourier analysis (c.f. Fact 2.3),

Fact 2.6 [9] Let $\lambda > 1$ be fixed. Let $\{t_l\}_{l \in \mathbb{Z}}$ be a set of points such that $\delta := \inf_{j,l,j \neq l} |t_j - t_l| > 0$ and $\kappa := \sup_l |t_l - \frac{l}{\lambda}| < \infty$ then there exist absolutely integrable interpolating functions $\psi_l(t)$ with $C' := \sup_{t \in \mathbb{R}} [\sum_{l \in \mathbb{Z}} |\psi_l(t - t_l)|] < \infty$ such that support

$$f(t) = \lim_{L \rightarrow \infty} \sum_{l=-L}^L f(t_l) \psi_l(t - t_l),$$

with the convergence holding pointwise (and in \mathcal{L}^2), absolutely, and uniformly on all compact subsets of \mathbb{R} .

To develop the framework for sampling using 1-bit A/D converters, a key-component is the dither function, $d(t)$, having the following properties:

1. $1 < |d(\frac{l}{\lambda})| =: \gamma < \infty, \forall l \in \mathbb{Z}$;
2. $\text{sgn}[d(\frac{l}{\lambda})] = -\text{sgn}[d(\frac{l+1}{\lambda})], \forall l \in \mathbb{Z}$;
3. $d(t)$ is differentiable (except possibly at $\{(\frac{l}{\lambda})\}_{l \in \mathbb{Z}}$) and $\Delta := \sup_{t \in \mathbb{R}} |d'(t)| < \infty$;

For example, $d(t) = \gamma \cos(\lambda \pi t)$ with $|\gamma| > 1$ is a valid dither function. It is easy to check that $f(t) + d(t)$ will change signs in every Nyquist interval and has a uniformly bounded slope.

Let $N = 2^k$, one-bit A/D converters be placed uniformly in every Nyquist-interval to record the sign of the dithered signal $f(t) + d(t)$, i.e., sensors are placed at the locations $\{m\tau_k\}_{m \in \mathbb{Z}}$ where $\tau := \frac{(1/\lambda)}{2^k}$ is the uniform oversampling period. Let $m_l \in \{0, \dots, 2^k - 1\}$ be the smallest index for which $[f + d](\frac{l}{\lambda} + m_l \tau)$ and $[f + d](\frac{l}{\lambda} + (m_l + 1)\tau)$ have opposite signs in $[\frac{l}{\lambda}, \frac{l+1}{\lambda}]$. Let $t_l := [\frac{l}{\lambda} + (m_l + \frac{1}{2})\tau]$. By an application of the Lagrange mean-value theorem it is easy to see that [10] $|f(t_l) - (-d(t_l))| \leq (\frac{\pi + \Delta}{2}) \tau = (\frac{\pi + \Delta}{2\lambda}) 2^{-k}$. Thus uniform oversampling of the dithered signal using 1-bit A/D converters gives samples of f having linear precision in τ (exponential in k) at the nonuniformly spaced points $\{t_l\}_{l \in \mathbb{Z}}$. There are 2^k one-bit sensors uniformly distributed over an interval of length $\frac{1}{\lambda}$. It takes k bits or a bitrate of $R = k\lambda$ bits/Nyquist-interval to index the location of the first zero-crossing. Hence the sample errors will decay with rate R as follows

$$|f(t_l) - (-d(t_l))| \leq \left(\frac{\pi + \Delta}{2\lambda}\right) 2^{-\frac{1}{\lambda}R}, \quad \forall l \in \mathbb{Z}. \quad (5)$$

It can be verified that the sequence $\{t_l\}_{l \in \mathbb{Z}}$ will satisfy the conditions of Fact 2.6 (see [10]), and hence there will exist interpolating functions $\{\psi_l(t)\}_{l \in \mathbb{Z}}$ for which the decay of the worst-case interpolated reconstruction error will also have a similar decay with rate:

$$\|f - \hat{f}_R^{1\text{-bit}}\|_{\infty} \leq C' \left(\frac{\pi + \Delta}{2\lambda}\right) 2^{-\frac{1}{\lambda}R}. \quad (6)$$

where $\hat{f}_R^{1\text{-bit}}(t) := \sum_{l \in \mathbb{Z}} (-d(t_l)) \psi_l(t - t_l)$. Observe that the reconstruction accuracy, in terms of bitrate, is similar to the Nyquist-sampling scheme, i.e., it is exponentially decaying in rate with the same exponent. The reconstruction accuracy can be improved by reducing τ , i.e., by packing more sensors inside each Nyquist-interval.

Stochastic Case: As in the deterministic case, sign changes in $[X + d](t)$ can be located. Let m_l be the index of the sensor just prior to the sign-change and denote $t_l := [\frac{l}{\lambda} + (m_l + \frac{1}{2})\tau]$. X agrees with X_{BL} with probability one (w.p. 1) simultaneously on all $(\frac{l}{\lambda} + m_l \tau)$ and $(\frac{l+1}{\lambda} + m_l \tau)$. Hence X_{BL} will also have all the first sign changes at the same locations as X w.p.1. Hence, (5) will apply to X_{BL} (The slope of X_{BL} is uniformly bounded, see Proposition 2.1(ii)) and we can assert that

$$|X_{BL}(t_l) - (-d(t_l))| \leq \left(\frac{\pi + \Delta}{2\lambda}\right) 2^{-\frac{1}{\lambda}R}, \quad \forall l \in \mathbb{Z}, \quad \text{w.p.1.} \quad (7)$$

Reconstructing $\widehat{X}_R^{1\text{-bit}}(t)$ as in deterministic case and using Propositions 2.1, 2.2 and 2.5 with (7) one can obtain [10]

$$\|E|X - \widehat{X}_R^{1\text{-bit}}|^p\|_\infty \leq (C'')^p 2^{-\frac{p}{\lambda}R}, \quad (8)$$

where $C'' := C'(\pi + \Delta)/(2\lambda)$. Collating (3), (4), (6) and (8), it is clear that

Theorem 2.7 *For a fixed $\lambda, \gamma > 1$, the \mathcal{L}^∞ norm of the point-wise error in the deterministic Nyquist reconstruction or 1-bit dithered reconstruction decays exponentially in the bitrate R with exponent $\frac{1}{\lambda}$. The worst-case, mean \mathcal{L}^p norm of the error in the stochastic Nyquist reconstruction or one-bit dithered reconstruction also decays exponentially in the bitrate R with exponent $\frac{p}{\lambda}$ for $p \in (0, \infty)$. Similar rate-error characteristics hold for the average distortion criteria of Section II-A.1. Hence, if D denotes distortion, the ‘‘information density’’ grows as $R \propto \log(1/D)$ as D decreases to 0. The number of sensors N , for dithered sampling, grows as $\log(1/D)$. Equivalently, the mean \mathcal{L}^p distortion decreases as $\frac{1}{N^p}$ while the information per Nyquist-interval grows as $\log N$. The bits per sensor decreases as $\frac{1}{N} \log N$.*

III. b -BIT DITHERED SCHEME AND ‘‘BIT-CONSERVATION’’ PRINCIPLE

This section explains how k -bit Nyquist-sampling accuracy can be achieved using b -bit A/D converters and an appropriate dither-based oversampling scheme for any $1 < b < k$. This leads to a ‘‘bit-conservation principle’’– a trade-off between the oversampling factor and A/D precision for ‘‘similar’’ asymptotic reconstruction accuracy [1], [10]. We present the deterministic case in detail and only state the results for the stochastic case.

Deterministic Case: Let b -bit A/D converters be placed at the locations $\{m\tau\}_{m \in \mathbb{Z}}$. Each sensor can detect $2^b - 1$ distinct level-crossings: $\{0, \pm \frac{1}{2^{b-1}}, \dots, \pm (1 - \frac{1}{2^{b-1}})\}$. One can design a b -bit dither function $d_b(t)$ such that $f(t) + d_b(t)$ will cross a level in every interval of the form $[A_l, B_l] := [\frac{l}{\lambda}, \frac{l}{\lambda} + (2^{k-b+1} - 1)\tau] \subset [\frac{l}{\lambda}, \frac{l+1}{\lambda}]$ which covers 2^{k-b+1} b -bit sensors. At most b -bits are needed for indexing the level that is crossed and $(k - b + 1)$ -bits for indexing the location of the first level crossing in each Nyquist interval. Let q_l be the level crossed and t_l be the mid-point of $A_l + m_l\tau$ and $A_l + (m_l + 1)\tau$, where level change happens between m_l and $m_l + 1$ sensor. Then it can be shown that [1], [10],

$$|f(t_l) - (q_l - d_b(t_l))| \leq \left(\frac{\pi + \Delta_b}{\lambda}\right) 2^{-\frac{1}{\lambda}R}, \quad \forall l \in \mathbb{Z}.$$

where $R = (k + 1)\lambda$ and $\Delta_b := \sup_t |d'_b(t)| < 2\Delta (1 + \frac{\pi}{\lambda})$. If the b -bit reconstruction is given by $\widehat{f}_R^{b\text{-bit}}(t) := \sum_{l \in \mathbb{Z}} (q_l - d_b(t_l)) \psi_l(t - t_l)$ then

$$\|f - \widehat{f}_R^{b\text{-bit}}\|_\infty \leq C' \left(\frac{\pi + \Delta_b}{\lambda}\right) 2^{-\frac{1}{\lambda}R},$$

Hence, the distortion-rate asymptotics is the same as in the Nyquist sampling and the 1-bit dithered sampling set-ups. This phenomenon can be summarized as follows:

‘‘Conservation of bits’’ principle: *Let k be the number of bits available per Nyquist-interval. For each $1 \leq b \leq k$ there exists a (dither-based) sampling scheme with not more than 2^{k-b+1} , b -bit A/D converters per Nyquist-interval achieving a worst-case pointwise reconstruction accuracy of the order of 2^{-k} .*

Stochastic Case: X can be sampled by adding $d_b(t)$ with same properties as earlier. Let m_l be the index of the sensor just prior to the first level-change and denote $t_l := [\frac{l}{\lambda} + (m_l + \frac{1}{2})\tau]$ and q_l as the level crossed. As earlier, X_{BL} will also have all the first level changes at the same locations as X w.p. 1. Also (5) applies to X_{BL} , therefore, denoting the reconstruction as $\widehat{X}_R^{b\text{-bit}}(t) := \sum_{l \in \mathbb{Z}} (q_l - d_b(t_l)) \psi_l(t - t_l)$, we can assert that

$$\|E|X - \widehat{X}_R^{b\text{-bit}}|^p\|_\infty \leq (C''')^p 2^{-\frac{p}{\lambda}R} \quad (9)$$

where $C''' := C' \left(\frac{\pi + \Delta_b}{\lambda}\right)$.

IV. DISTRIBUTED SAMPLING IN SENSOR NETWORKS

We have so far described, in an application-independent context, how an underlying bit-conservation principle admits flexible tradeoffs between A/D quantizer resolution and the oversampling rate. In this section, we show how the proposed framework is particularly germane to the sensor network application as motivated in the introduction. Our discussion centers around 1-bit sensors but easily extends to the general b -bit case.

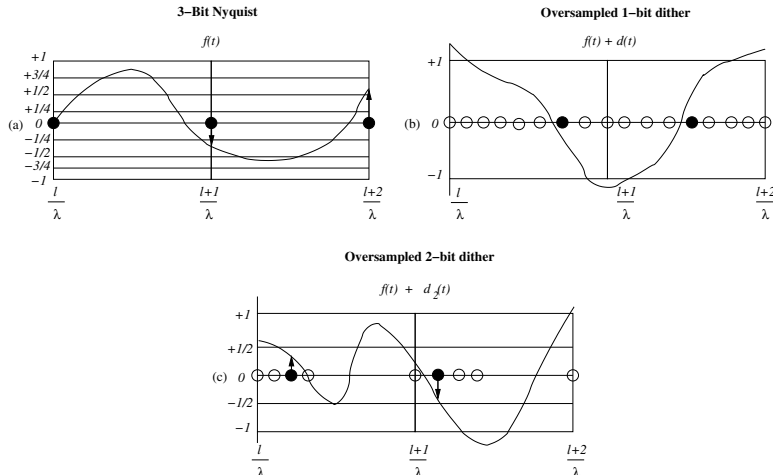


Fig. 1. Illustrating the amplitude precision and oversampling rate tradeoffs in conventional and dither-based sampling frameworks. The top figure depicts the conventional Nyquist-sampling using 3-bit A/D converters in which the entire budget of 3 bits is exhausted at a single sample point. The middle figure shows a dither-based sampling scheme that uses eight, 1-bit A/D converters uniformly distributed over a Nyquist-interval to locate the zero crossing of the dithered signal. The bottom figure shows how to achieve a flexible tradeoff between these extremes. Four, 2-bit A/D converters uniformly distributed over half the Nyquist-period detect level crossings at 0 and $\pm \frac{1}{2}$. The three schemes have similar exponential error accuracy in bitrate.

- Let N , one-bit sensors be placed uniformly at every $\tau = \frac{1}{\lambda N}$ in every Nyquist interval $[\frac{l}{\lambda}, \frac{(l+1)}{\lambda} - \tau]$ (see Fig.(1)). The sensors at $\frac{l}{\lambda}$ are the starting nodes.
- Periodically, the sensors take snapshots of the 1-D spatio-temporal random field by comparing the field value to the dither value at their respective locations (the dither values are assumed to be pre-stored during sensor deployment).
- Corresponding to each temporal snapshot of the field, each starting node passes a message to its neighbor (say it's right neighbor), indicating two things: (i) whether or not a zero-crossing has already been found by some preceding sensor and (ii) if a zero-crossing has not yet been found, the sign of the field plus dither at its location.
- The first sensor in each Nyquist interval that detects a sign mismatch between what its left neighbor reports and its own reading records a one. Other sensors record a zero for that snapshot. The local communication for detecting the zero-crossings need not be done in real-time. The sensors can store the signs of the field plus dither over many snapshots and locate the zero-crossings later.
- Finally, each sensor encodes the zero-crossing information using principles of distributed compression [13], [14].

Growth of information To understand how the sensor data-rate grows with N one needs to specify, at some level, a temporal model for the evolution of the random field. However, the temporal sequence of the first zero-crossing locations in every Nyquist-interval contains all the information pertaining to the field. Suppose that the sensors within each Nyquist-interval could collaborate without penalty to jointly encode the zero-crossing information over many temporal snapshots. By design, there is exactly one sensor per Nyquist-interval

per snapshot that will record the value one. Hence, the N -length vector of “zero-crossings” in any Nyquist-interval can take only N distinct values in each snapshot. Irrespective of the exact statistical structure of the temporal evolution, in the worst case, no more than $R_{Nyquist} = \log_2(N)$ bits per snapshot per Nyquist-interval are needed to encode the N -length vectors of “zero-crossings” (being the maximum entropy of a source over an alphabet of size N). This represents the worst-case joint entropy of all the (binary-valued) sensor outputs (zero-crossing information) in any Nyquist-interval. It is possible to achieve a compression efficiency of $\log_2(N)$ bits per Nyquist-interval per snapshot *without any sensor collaboration* using principles of distributed source coding theory [13], [14] which exploits the correlation between sensor outputs. In fact, it is possible for each sensor to encode its zero-crossing data at the rate $R_{sensor} = \frac{1}{N} \log_2(N)$ bits snapshot without talking to other sensors in the same Nyquist-interval and the central unit will be able to recover the encoded information. Recall that the distortion (maximum expected point error) decreases as $D = O(2^{-R_{Nyquist}})$. Hence the distortion goes to zero as $D = O(\frac{1}{N})$ quite unlike the setup in [8] and [6] where the distortion does not go to zero even as N goes to infinity. Equivalently, for a fixed $D > 0$, the number of 1-bit sensors per Nyquist interval needed is finite with $N \propto \frac{1}{D}$.

Local inter-sensor communication cost: For a reconstruction accuracy of the order of $\tau = \frac{1}{\lambda N}$, 1-bit precision sensors, and the above local communication protocol, each sensor transmits 2-bits to its right neighbor over a distance of τ . Hence, the local sensor communication cost is $\frac{2}{N\lambda}$ bit-meters. In general, for b -bit sensors ($b > 1$) the local sensor communication cost is no more than $(b+1)/(\lambda N_b 2^{b-1})$ where $N_b = \frac{N}{2^{b-1}}$ is the number of b -bit sensors needed to match the asymptotic reconstruction error performance of 1-bit sensors. Here, the b -bit sensors are placed uniformly every $\tau = \frac{1}{\lambda N} = \frac{1}{\lambda N_b 2^{b-1}}$ in $[\frac{l}{\lambda}, \frac{l}{\lambda} + (N_b - 1)\tau]$. The total local communication cost is no more than $\frac{2}{\lambda}$ bit-meters per Nyquist-interval. Thus, with limited local communication cost, the sampling task can be nicely distributed among the sensors. We believe, that even this limited local inter-sensor communication can be completely done-away with by exploiting the structure of band-limited fields. For instance by design of a suitable dither function, when ameliorated over many Nyquist intervals, there will only be a bounded number of zero-crossings of the field plus the dither per Nyquist interval. If Z is the maximum average number of zero-crossings per Nyquist interval then the distributed encoding of the pure sensor observations (the signs of the field plus dither) over say L Nyquist intervals will not need more than $\frac{1}{L \cdot N} \log_2 \left(\binom{LZ}{LZ} \right)$ bits per snapshot per sensor.

Sensor distribution: As we increase the A/D precision (b increases), we get a reduction in the number of sensors according to the “conservation of bits” principle (N_b decreases exponentially as b increases) while maintaining the same asymptotic error decay profile. Sensors need to be placed only in intervals of the form $[\frac{l}{\lambda}, \frac{l}{\lambda} + (N_b - 1)\tau]$. This leaves intervals over which there is no need to sample the signal at all. Hence for a given reconstruction quality the number of sensing units goes down exponentially with b : 1-bit dithered sampling needs N sensors, b -bit ($b > 1$) dithered sampling needs $N_b = N/2^{b-1}$ sensors, and Nyquist-sampling needs only one sensor per Nyquist-interval. Since the scheme naturally allows “inactive” regions in oversampling, we can have bunched irregular sampling using sensors. This allows for design flexibility in sensor deployment. For example, in rugged terrain or in the presence of occluding obstacles, one would need to use higher precision sensors. Where sensors can be deployed in large numbers it is sufficient to use cheap 1-bit sensors. More sophisticated non-uniform sampling using heterogeneous sensors can also be realized by using Fact 2.3.

Robustness: The dither-based oversampling method also offers robustness to node failures in terms of a graceful degradation of reconstruction error. For example, if every alternate node fails, the effective inter-node separation would increase to 2τ . This has the same effect as halving the resolution of the A/D converters by the bit-conservation principle. The same dither function will continue to work because it was designed for a higher spatial density.

For local communication, the sensors would need to use more power to ensure that their message gets across a distance of 2τ as opposed to τ when all nodes were functioning.

Security: The design of the underlying dither function is very flexible. This allows for a secure sampling by selecting a covert dither function. The dither can be implementation-specific and can even be different on different Nyquist-intervals. However, the choice of the dither function also affects the reconstruction accuracy (through the slope of the dither function). Hence, a larger distortion for an eavesdropper will also imply a larger reconstruction error in general. Quantifying the tradeoff between reconstruction error and security is part of our ongoing research.

V. CONCLUDING REMARKS

This paper is but the first step towards understanding the fundamentals of distributed sampling theory. The setup of a 1-D bandlimited signal model is somewhat simplistic from the point of view of sensor networks but a necessary first step in probing further. Extensions to non-bandlimited signals, signals with a finite rate of innovation, and 2-D spatio-temporal fields under the severe communication and processing constraints associated with sensor networks are all exciting avenues for taking this work further. Preliminary results for non-bandlimited fields are presented in [10].

ACKNOWLEDGMENT

The authors would like to thank Vinod. M. Prabhakaran (EECS dept. UC Berkeley) and Prof. Sandeep Pradhan (EECS dept. UMich. Ann Arbor) for several inspiring discussions on the stochastic aspects of distributed sampling.

REFERENCES

- [1] P. Ishwar, A. Kumar, and K. Ramchandran, "Distributed Sampling for Dense Sensor Networks: a "bit-conservation" principle," in *Information Processing in Sensor Networks (IPSN), Proceedings of the Second International Workshop, Palo Alto, CA, USA, April 22-23, 2003*, Lecture Notes in Computer Science edited by L. J. Guibas and F. Zhao, Springer, New York, 2003, pp. 17–31.
- [2] R. J. Marks, II, *Introduction to Shannon Sampling and Interpolation Theory*. New York, USA: Springer-Verlag, 1990.
- [3] J. Higgins, *Sampling Theory in Fourier and Signal Analysis: Foundations*. USA: Clarendon Press, 1996.
- [4] Z. Cvetković and M. Vetterli, "Error-rate Characteristics of Oversampled Analog-to-Digital Conversion," *IEEE Trans. on Information Theory*, vol. 44, pp. 1961–1964, sep 1998.
- [5] P. Gupta and P. Kumar, "Capacity of wireless networks," *IEEE Trans. on Information Theory*, vol. 46, pp. 388–404, Mar 2000.
- [6] A. Scaglione and S. D. Servetto, "On the interdependence of routing and data compression in multi-hop sensor networks," in *Proceedings of the eighth annual international conference on Mobile computing and networking*, pp. 140–147, ACM Press, 2002.
- [7] J. Chou, D. Petrovic, and K. Ramchandran, "A Distributed and Adaptive Signal Processing Approach to Reducing Energy Consumption in Sensor Networks," in *Proc. IEEE Infocom*, (San Francisco, CA), March 2003.
- [8] 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 *Information Processing in Sensor Networks (IPSN), Proceedings of the Second International Workshop, Palo Alto, CA, USA, April 22-23, 2003*, Lecture Notes in Computer Science edited by L. J. Guibas and F. Zhao, Springer, New York, 2003, pp. 1–16.
- [9] Z. Cvetković and I. Daubechies, "Single Bit oversampled A/D conversion with exponential accuracy in bit rate," *Proceeding DCC*, pp. 343–352, March 2000.
- [10] P. Ishwar, A. Kumar, and K. Ramchandran, "On Distributed Sampling in Dense Sensor Networks: a "bit-conservation" Principle," Journal preprint available at <http://www.eecs.berkeley.edu/~animesh/jsac03ishwartetal.pdf>.
- [11] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia: SIAM, 1992.
- [12] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. on Information Theory*, vol. IT-44, pp. 2325–2383, Oct. 1998.
- [13] D. Slepian, and J. K. Wolf, "Noiseless Coding of Correlated Information Sources," *IEEE Trans. on Information Theory*, vol. 19, pp. 471–480, july 1973.
- [14] S. S. Pradhan and K. Ramchandran, "Distributed Source Coding Using Syndromes (DISCUS): Design and Construction," *IEEE Trans. on Information Theory*, vol. 49, pp. 626–643, Mar 2003.