SOI-KF: Distributed Kalman Filtering With Low-Cost Communications Using the Sign of Innovations

Alejandro Ribeiro, *Student Member, IEEE*, Georgios B. Giannakis, *Fellow, IEEE*, and
Stergios I. Roumeliotis, *Member, IEEE*

Abstract—When dealing with decentralized estimation, it is important to reduce the cost of communicating the distributed observations—a problem receiving revived interest in the context of wireless sensor networks. In this paper, we derive and analyze distributed state estimators of dynamical stochastic processes, whereby the low communication cost is effected by requiring the transmission of a single bit per observation. Following a Kalman filtering (KF) approach, we develop recursive algorithms for distributed state estimation based on the sign of innovations (SOI). Even though SOI-KF can afford minimal communication overhead, we prove that in terms of performance and complexity it comes very close to the clairvoyant KF which is based on the analog-amplitude observations. Reinforcing our conclusions, we show that the SOI-KF applied to distributed target tracking based on distance-only observations yields accurate estimates at low communication cost.

Index Terms—Distributed state estimation, Kalman filter (KF), target tracking, wireless sensor networks.

I. INTRODUCTION

DISTRIBUTED signal processing is a well-appreciated toolbox for decentralized tracking applications involving, e.g., multiple radars, but has received a revived interest recently in the context of wireless sensor networks (WSNs) [5]. Unlike centralized signal processing, observations and the resultant algorithms are physically distributed across sensors in the network, dictating that intersensor communications should be viewed as an integral part of the problem at hand, be it reconstruction, filtering or estimation. For distributed estimation of dynamical stochastic processes dealt with in this paper, only

quantized observations are communicated. Thus, the estimation problem is certainly different from state estimation based on the original (analog-amplitude) observations.

Without explicitly considering quantization, spatial redundancy across sensor observations has been exploited to reduce communication requirements [3], [4], [9], [21], [23], [24]. Accounting for quantization, distributed estimation algorithms were explored in early works, see, e.g., [10] and [17]; and recently in the context of WSNs. The design of quantizers in different scenarios was studied in [22], where the concept of information loss was defined as the relative increase in estimation variance when using quantized observations with respect to the equivalent estimation problem based on analog-amplitude observations. To address the challenge of building suitable noise models for WSNs, universal estimators that work irrespective of the noise distribution were introduced in [18] and shown to have an information loss independent of the network size. Another insight when estimating signals using very noisy sensor data was offered by [25], [26], where it was shown that as the noise variance becomes comparable with the parameter's dynamic range, quantization to a single bit per observation leads to low complexity estimators of *time-invariant deterministic* parameters with minimal information loss. This holds true for a large class of problems, where the noise probability distribution function (pdf) may be parametrically described or even unknown [25].

Taking into account the stringent bandwidth constraints of WSNs, this paper studies state estimation of *dynamical stochastic* processes based on severely quantized observations, whereby low-cost communications restrict sensors to transmit a single bit per observation. The quantization rule manifests itself in a non-linear measurement equation in a Kalman Filtering (KF) setup. While the discontinuous non-linearity precludes application of the extended (E) KF, it can be handled with more powerful techniques such as the unscented (U) KF [14], or the particle filter (PF) [7], [16]—algorithms that have also been applied in the context of filtering [6], [29] and target tracking with a WSN [1], [8]. However, all these approaches are significantly more complex than a KF and, besides, no insight has been provided with regards to their performance degradation when quantized data are used in lieu of the analog-amplitude observations. The contribution of the present paper is precisely to address these two issues with the goal being to construct state estimators based on binary observations so that: i) complexity is rendered comparable to the equivalent KF based on the original observations; and ii) the mean squared error (MSE) of the resultant estimate based on binary observations is close to the MSE of the estimate based on the original observations.

We begin by introducing our WSN setup and formulating the problem in Section II, where we delineate the KF that we will use to benchmark algorithms in the rest of this paper (Section II-A). State estimation based on the sign of innovations (SOI) is considered first for a vector state-scalar observation model in Section III, where we discuss the minimum mean squared error (MMSE) estimator (Section III-A). As the latter may be prohibitive for a resource-limited WSN, we pursue a reduced-complexity approximation in Section III-B which leads to the SOI-KF algorithm whose complexity and performance are surprisingly close to the clairvoyant KF, even when intersensor communication relies on the low-cost transmission of a single bit per sensor. These results are extended to a general vector state-vector parameter model in Section IV. The performance of the SOI-KF is analyzed in Section V, where using the underlying continuous-time physical processes we show that the MSE of the SOI-KF is closely related to the MSE of a KF with measurement noise covariance matrix only $\pi/2$ times the original one. We present a motivating example in Section VI, entailing temperature monitoring with a WSN. Finally, we apply a modified version of the SOI-KF to the canonical problem of distributed target tracking based on binary observations in Section VI-A. Section VII concludes this paper.

Notation: We use $p(x|y;z)$ to denote the probability density function (pdf) of the random variable (r.v.) x given the r.v. y evaluated at z; when using the same letter to denote the r.v. and the argument of the pdf we abbreviate $p(x|y;x) = p(x|y)$. When an r.v. is normally distributed with mean $\boldsymbol{\mu}_x = \mathrm{E}(\mathbf{x})$ and covariance matrix $\mathbf{C}_x = \mathrm{E}(\mathbf{x}\mathbf{x}^T)$, we write $p(\mathbf{x}) = \mathcal{N}(\mathbf{x};\boldsymbol{\mu}_x,\mathbf{C}_x)$, where T stands for transposition. In the particular case of a scalar r.v., we write $p(x) = \mathcal{N}(x;\mu_x,\sigma_x^2)$ and define the Gaussian tail function as $Q(x) := \int_x^\infty \mathcal{N}(u;0,1)du$. We will use $\delta_c(t)$ to denote the Dirac delta function defined by $\delta_c(t) = 0\ \forall t \neq 0$, and $\int_{-\infty}^\infty \delta_c(t)dt = 1$; and $\delta(n)$ to denote the Kronecker delta function defined as $\delta(0) = 1$ and $\delta(n) = 0\ \forall n \neq 0$. For any function $f(t)$ the notation $f(t) = \mathrm{o}(t)$ will imply that $\lim_{t\to 0} f(t)/t = 0$. Throughout this paper, $\mathbf{I}$ will denote the identity matrix, and lower (upper) case boldface letters will stand for column vectors (matrices).

II. PROBLEM STATEMENT AND PRELIMINARIES

We are primarily concerned with so called ad-hoc WSNs in which the network itself is responsible for collecting and processing information; see Fig. 1. Let us consider an *ad-hoc* WSN with K distributed sensors $\{S_k\}_{k=1}^K$ deployed with the objective of tracking a $p \times 1$ real random vector (state) $\mathbf{x}_c(t) \in \mathbb{R}^p$. The state evolution in continuous-time is described by

$$\dot{\mathbf{x}}_c(t) = \mathbf{A}_c(t)\mathbf{x}_c(t) + \mathbf{u}_c(t) \qquad (1)$$

where $\mathbf{A}_c(t) \in \mathbb{R}^{p\times p}$, and the driving input $\mathbf{u}_c(t) \in \mathbb{R}^p$ is a zero-mean white Gaussian process with autocorrelation $\mathrm{E}[\mathbf{u}_c(t_1)\mathbf{u}_c^T(t_2)] = \mathbf{C}_{u_c}(t_1)\delta_c(t_1 - t_2)$. The sensors observe the state $\mathbf{x}_c(t)$ through a linear transformation. Letting $\mathbf{y}_c(t,k) \in \mathbb{R}^M$ denote the observation at sensor S_k, we have

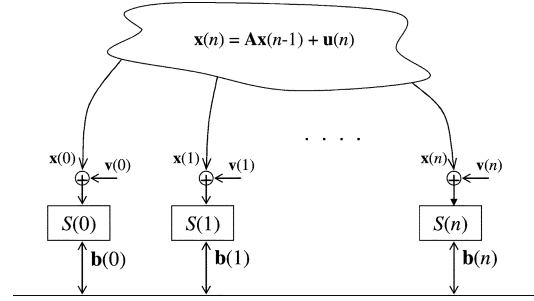$$\mathbf{y}_c(t,k) = \mathbf{H}_c(t,k)\mathbf{x}_c(t) + \mathbf{v}_c(t,k) \qquad (2)$$

Fig. 1. *Ad-ho*c WSN: the network itself is in charge of tracking the state $\mathbf{x}(n)$.

where $\mathbf{H}_c(t,k) \in \mathbb{R}^{M\times p}$ and the observation noise $\mathbf{v}_c(t,k) \in \mathbb{R}^M$ is also a zero-mean Gaussian process with $\mathrm{E}[\mathbf{v}_c(t_1,k_1)\mathbf{v}_c^T(t_2,k_2)] = \mathbf{C}_{v_c}(t_1,k_1)\delta_c(t_1 - t_2)\delta(k_1 - k_2)$, i.e., the noise is uncorrelated across time and sensors.

To track $\mathbf{x}_c(t)$, we consider uniform sampling with period T_s and define the discrete-time state and observations as $\mathbf{x}(n) := \mathbf{x}_c(nT_s)$ and $\mathbf{y}(n,k) := \mathbf{y}_c(nT_s,k)$, respectively. Using the continuous-time model described by (1) and (2) we can obtain an equivalent discrete-time model [20, Sec. 4.9]. Upon defining $\boldsymbol{\Phi}(t_2,t_1) := \exp[\int_{t_1}^{t_2} \mathbf{A}_c(t)dt]$, we can solve the differential equation in (1) between $(n-1)T_s$ and nT_s with initial condition $\mathbf{x}(n-1)$ to obtain

$$\mathbf{x}(n) = \boldsymbol{\Phi}(nT_s,(n-1)T_s)\,\mathbf{x}(n-1)$$
$$+ \int_{(n-1)T_s}^{nT_s} \boldsymbol{\Phi}(nT_s,\tau)\mathbf{u}_c(\tau)d\tau. \quad (3)$$

For simplicity, define the matrix $\mathbf{A}(n) := \boldsymbol{\Phi}(nT_s,(n-1)T_s)$ and the white Gaussian driving noise input $\mathbf{u}(n) := \int_{(n-1)T_s}^{nT_s} \boldsymbol{\Phi}(nT_s,\tau)\mathbf{u}_c(\tau)d\tau$. With these definitions, the resultant discrete-time equivalent model is given by the vector time-varying autoregressive (AR) process

$$\mathbf{x}(n) = \mathbf{A}(n)\mathbf{x}(n-1) + \mathbf{u}(n)$$
$$\mathbf{y}(n,k) = \mathbf{H}(n,k)\mathbf{x}(n) + \mathbf{v}(n,k) \qquad (4)$$

where $\mathbf{H}(n,k) := \mathbf{H}_c(nT_s,k)$ and the observation noise is white Gaussian with pdf $p[\mathbf{v}(n,k)] = \mathcal{N}[\mathbf{v}(n,k);\mathbf{0},\mathbf{C}_v(n,k)]$. Since sampling (2) requires passing $\mathbf{y}_c(t,k)$ through a low- or band-pass filter of bandwidth $1/T_s$, the sampled covariance matrix satisfies $\mathbf{C}_v(n,k) := \mathrm{E}[\mathbf{v}(n,k)\mathbf{v}^T(n,k)] = \mathbf{C}_{v_c}(nT_s,k)/T_s$ [20, Sec. 4.9]. Finally, note that $\mathbf{u}(n)$'s definition implies that $p[\mathbf{u}(n)] = \mathcal{N}[\mathbf{u}(n);\mathbf{0},\mathbf{C}_u(n)]$ with covariance matrix $\mathbf{C}_u(n) := \mathrm{E}[\mathbf{u}(n)\mathbf{u}^T(n)] = \int_{(n-1)T_s}^{nT_s} \boldsymbol{\Phi}(nT_s,\tau)\mathbf{C}_{u_c}(\tau)\boldsymbol{\Phi}^T(nT_s,\tau)d\tau$.

Supposing that $\mathbf{A}(n)$, $\mathbf{C}_u(n)$, $\mathbf{H}(n,k)$ and $\mathbf{C}_v(n,k)$ are available $\forall n,k$, the goal of the WSN is for each sensor S_k to form an estimate of $\mathbf{x}(n)$ to be used in e.g., a habitat monitoring application [19], or, as a first step in e.g., a distributed control setup [13]. In any event, estimating $\mathbf{x}(n)$ necessitates each sensor S_l to communicate $\mathbf{y}(n,l)$ to the remaining sensors $\{S_k\}_{k=1,k\neq l}^K$. This communication takes place over the shared wireless channel that we will assume can afford transmission of a single packet per time slot n, leading to a one-to-one correspondence between time n and sensor index k and allowing

us to drop the sensor argument k in (4). The decision of which sensor $S_k = S(n)$ is active at time n, and consequently which observation $\mathbf{y}(n, k) = \mathbf{y}(n)$ gets transmitted, depends on the underlying scheduling algorithm—see, e.g., [11], [21]. and the references therein—but is assumed given for the purpose of this paper. Digital transmission of $\mathbf{y}(n)$ also implies some form of quantization $\mathbf{q}_n$ to map the analog observations $\mathbf{y}(n)$ into binary data

$$\mathbf{b}(n) := \mathbf{q}_n\left(\mathbf{y}(n)\right), \qquad \text{with} \quad \mathbf{q}_n : \mathbf{R}^M \to \{0, 1\}^M \quad (5)$$

where $\mathbf{b}(n) := [b(n, 1), \ldots, b(n, M)]^T$ is an M-component binary message. Implicit to (5) is the fact that we are restricting the sensors to transmit one bit per scalar observation which effects low-cost communications among sensors. Indeed, the quantization function $\mathbf{q}_n$ partitions $\mathbf{R}^M$ in 2^M regions, implying that on the average each component of $\mathbf{y}(n)$ is quantized to 1 bit. We further suppose that the messages $\mathbf{b}(n)$ are correctly received by all sensors, which assumes deployment of sufficiently powerful error-control codes.

The objective of this paper is to derive and analyze the performance of MMSE estimators of $\mathbf{x}(n)$ based on the messages $\mathbf{b}_{0:n} := [\mathbf{b}^T(0), \ldots, \mathbf{b}^T(n)]^T$ that are available to each and every sensor. It is well known that the MMSE estimator is given by the conditional expectation [15, Ch. 12]; consequently, if we let $\hat{\mathbf{x}}(n|n)$ denote the MMSE estimator of $\mathbf{x}(n)$ given $\mathbf{b}_{0:n}$, we have

$$\hat{\mathbf{x}}(n|n) := \mathrm{E}\left[\mathbf{x}(n)|\mathbf{b}_{0:n}\right] = \int_{\mathbb{R}^p} \mathbf{x}(n)p\left[\mathbf{x}(n)|\mathbf{b}_{0:n}\right] d\mathbf{x}(n). \quad (6)$$

Instrumental to the ensuing derivations are the so called predictors that estimate (predict) the state and observation vectors based on past observations

$$\hat{\mathbf{x}}(n|n-1) := \mathrm{E}\left[\mathbf{x}(n)|\mathbf{b}_{0:n-1}\right] = \mathbf{A}(n)\hat{\mathbf{x}}(n-1|n-1)$$
$$\hat{\mathbf{y}}(n|n-1) := \mathrm{E}\left[\mathbf{y}(n)|\mathbf{b}_{0:n-1}\right] = \mathbf{H}(n)\hat{\mathbf{x}}(n|n-1). \quad (7)$$

For each of the state estimators in (6) and (7), we define the error covariance matrices (ECM) $\mathbf{M}(n|n) := \mathrm{E}[(\hat{\mathbf{x}}(n|n) - \mathbf{x}(n))(\hat{\mathbf{x}}(n|n) - \mathbf{x}(n))^T]$, and $\mathbf{M}(n|n-1) := \mathrm{E}[(\hat{\mathbf{x}}(n|n-1) - \mathbf{x}(n))(\hat{\mathbf{x}}(n|n-1) - \mathbf{x}(n))^T]$ for the filtered and the predicted estimate, respectively. The mean square errors (MSEs) of $\hat{\mathbf{x}}(n|n)$ and $\hat{\mathbf{x}}(n|n-1)$ are given by $\mathrm{tr}[\mathbf{M}(n|n)]$ and $\mathrm{tr}[\mathbf{M}(n|n-1)]$ with these traces being minimum among all possible estimators $\hat{\mathbf{x}}(n|n)$ and $\hat{\mathbf{x}}(n|n-1)$ of $\mathbf{x}(n)$. The ECM of the state predictor can be obtained from the ECM of the state estimator through the recursion

$$\mathbf{M}(n|n-1) = \mathbf{A}(n)\mathbf{M}(n-1|n-1)\mathbf{A}^T(n) + \mathbf{C}_u(n) \quad (8)$$

which we will use in later derivations. Note that the relations between $\hat{\mathbf{x}}(n|n-1)$ and $\hat{\mathbf{x}}(n-1|n-1)$ and $\hat{\mathbf{y}}(n|n-1)$ and $\hat{\mathbf{x}}(n|n-1)$ in (7) and between $\mathbf{M}(n|n-1)$ and $\mathbf{M}(n-1|n-1)$ in (8) follow from the linearity of the expected value operator and are independent of the quantization rule in (5).
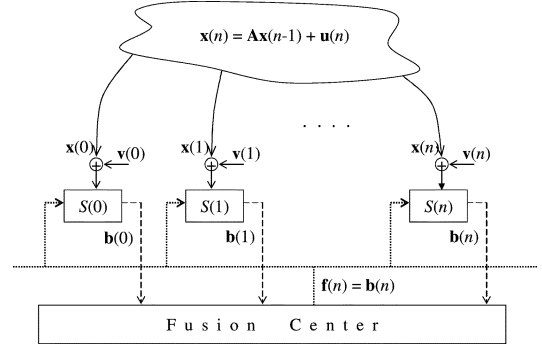

Fig. 2. WSN with a fusion center: The sensors act as data gathering devices.

Remark 1: When a fusion center (FC) is present, the WSN is termed hierarchical in the sense that sensors act as information gathering devices for the FC that is in charge of processing this information; see Fig. 2. Results in this paper also apply to networks of this type provided that the FC feeds back to the sensors packets $\mathbf{f}(n) := \mathbf{b}(n)$. As we will discuss in Sections III-B and IV, the sole condition for applying the proposed method is to have the predicted observation $\hat{\mathbf{y}}(n|n-1)$ available at $S(n)$, a condition that can be met in the hierarchical WSN with feedback $\mathbf{f}(n) = \mathbf{b}(n)$.

A. The Kalman Filter Benchmark

Before considering estimation based on binary observations, let us highlight some properties of the clairvoyant KF that will come handy in subsequent derivations. Consider for simplicity a vector state-scalar observation model described by

$$\mathbf{x}(n) = \mathbf{A}(n)\mathbf{x}(n-1) + \mathbf{u}(n)$$
$$y(n) = \mathbf{h}^T(n)\mathbf{x}(n) + v(n). \quad (9)$$

The model in (9) is a particular case of the general model (4) in which $M = 1$; the observations $\mathbf{y}(n) \leftrightarrow y(n)$, noise $\mathbf{v}(n) \leftrightarrow v(n)$ and noise covariance $\mathbf{C}_v(n) \leftrightarrow \sigma_v^2(n)$ are scalar; and $\mathbf{H}(n) \leftrightarrow \mathbf{h}^T(n) \in \mathbf{R}^{1 \times p}$ is a row vector.

If we had infinite bandwidth available, we could communicate the observations $y(n)$ error-free. This is rightfully a clairvoyant benchmark for our bandwidth-constrained estimators and corresponds to the problem setup of Section II with messages $b(n) = y(n)$. In this case, we have a well known linear Gaussian vector AR estimation problem whose MMSE can be recursively obtained by the KF [15, Ch. 13]. Assuming that the estimate $\hat{\mathbf{x}}(n-1|n-1)$ and the ECM $\mathbf{M}(n-1|n-1)$ are known at step $n-1$, we compute the predicted estimate $\hat{\mathbf{x}}(n|n-1)$ and the corresponding ECM $\mathbf{M}(n|n-1)$ using (7) and (8), respectively. Next, the filtered estimate $\hat{\mathbf{x}}(n|n)$ is obtained by solving the integral in (6) with the posterior pdf computed by means of Bayes' rule

$$p\left[\mathbf{x}(n)|\mathbf{y}_{0:n}\right] = \frac{p\left[y(n)|\mathbf{x}(n), \mathbf{y}_{0:n-1}\right] p\left[\mathbf{x}(n)|\mathbf{y}_{0:n-1}\right]}{p\left[y(n)|\mathbf{y}_{0:n-1}\right]}. \quad (10)$$

The key observation is that because of the linear Gaussian model (9), the posterior $p[\mathbf{x}(n)|\mathbf{y}_{0:n}] = \mathcal{N}[\mathbf{x}(n); \hat{\mathbf{x}}(n|n), \mathbf{M}(n|n)]$ is normal, leading to the so called correction step

$$\hat{\mathbf{x}}(n|n) = \hat{\mathbf{x}}(n|n-1)$$
$$+ \frac{\mathbf{M}(n|n-1)\mathbf{h}(n)}{\mathbf{h}^T(n)\mathbf{M}(n|n-1)\mathbf{h}(n) + \sigma_v^2(n)} \tilde{y}(n|n-1)$$
$$\mathbf{M}(n|n) = \mathbf{M}(n|n-1)$$
$$- \frac{\mathbf{M}(n|n-1)\mathbf{h}(n)\mathbf{h}^T(n)\mathbf{M}(n|n-1)}{\mathbf{h}^T(n)\mathbf{M}(n|n-1)\mathbf{h}(n) + \sigma_v^2(n)} \quad (11)$$

where we defined the innovation sequence $\tilde{y}(n|n-1) := y(n) - \hat{y}(n|n-1)$. Recursive application of (7), (8) and (11) yields the MMSE estimate of $\mathbf{x}(n)$ given $\mathbf{y}_{0:n}$.

III. STATE ESTIMATION USING THE SIGN OF INNOVATIONS

The corrector in (11) depends on the innovation sequence $\tilde{y}(n|n-1) := y(n) - \hat{y}(n|n-1)$ corresponding to the difference between the current observation and the prediction based on past observations. This suggests that a convenient form for the quantization function $\mathbf{q}_n$ is $q_n[y(n) - \hat{y}(n|n-1)]$. We start by considering the vector state-scalar observation model in (9) and define the message $b(n)$ as the SOI

$$b(n) = \text{sign}\,[y(n) - \hat{y}(n|n-1)]$$
$$:= \begin{cases} +1, & \text{if } y(n) \geq \hat{y}(n|n-1) \\ -1, & \text{if } y(n) < \hat{y}(n|n-1) \end{cases}. \quad (12)$$

Notice that the SOI $b(n)$ is *not* a standard quantizer of the data $y(n)$. It can be thought as one that judiciously sets the quantization threshold at the data prediction $\hat{y}(n|n-1)$. The focus of the present section is to study MMSE estimation of $\mathbf{x}(n)$ based on $\mathbf{b}_{0:n} := [b(0), \ldots, b(n)]^T$.

A. Exact MMSE Estimator

To find the MMSE in (6) based on the SOI in (12) we can, in principle, proceed as we described in Section II-A for the KF. However, while we can update the estimate $\hat{\mathbf{x}}(n-1|n-1)$ and its ECM $\mathbf{M}(n-1|n-1)$ using (7) and (8) to obtain the predictor $\hat{\mathbf{x}}(n|n-1)$ and its corresponding ECM $\mathbf{M}(n|n-1)$, the analogy with the KF cannot be pursued any further. The reason is that due to the non-linearity in the definition of $b(n)$ in (12) the distribution $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]$ is no longer normal; and, thus, its description requires additional information besides its mean and variance. This characteristic problem of nonlinear filtering motivates the need for a means of propagating the posterior pdf $p[\mathbf{x}(n)|\mathbf{b}_{0:n}]$ so that the integral in (6) can be evaluated. Such a rule is described in the following proposition.

Proposition 1: Consider the vector state-scalar observation model defined by (9), and the SOI messages defined as in (12). Then, the posterior pdf of $\mathbf{x}(n)$ given the binary observations $\mathbf{b}_{0:n}$ can be obtained using the recursions

$$p\,[\mathbf{x}(n)|\mathbf{b}_{0:n-1}] = \int_{\mathbb{R}^p} p\,[\mathbf{x}(n-1)|\mathbf{b}_{0:n-1}]$$
$$\times \mathcal{N}\,[\mathbf{x}(n); \mathbf{A}(n)\mathbf{x}(n-1), \mathbf{C}_u(n)]\,d\mathbf{x}(n-1) \quad (13)$$
$$p\,[\mathbf{x}(n)|\mathbf{b}_{0:n}] = \alpha_n p\,[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]$$
$$\times Q\left[-b(n)\frac{\mathbf{h}^T(n)\,[\mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1)]}{\sigma_v(n)}\right] \quad (14)$$

where α_n is a normalizing constant ensuring that $\int_{\mathbb{R}^p} p[\mathbf{x}(n)|\mathbf{b}_{0:n}]d\mathbf{x}(n) = 1$.

Proof: The prior pdf $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]$ in (13) follows from the theorem of total probability

$$p\,(\mathbf{x}(n)|\mathbf{b}_{0:n-1}) = \int_{\mathbb{R}^p} p\,[\mathbf{x}(n)|\mathbf{x}(n-1), \mathbf{b}_{0:n-1}]$$
$$\times p\,[\mathbf{x}(n-1)|\mathbf{b}_{0:n-1}]\,d\mathbf{x}(n-1). \quad (15)$$

Note, however, that since $\mathbf{x}(n-1)$ is given in $p[\mathbf{x}(n)|\mathbf{x}(n-1), \mathbf{b}_{0:n-1}]$, conditioning on $\mathbf{b}_{0:n-1}$ is irrelevant and $p[\mathbf{x}(n)|\mathbf{x}(n-1), \mathbf{b}_{0:n-1}] = p[\mathbf{x}(n)|\mathbf{x}(n-1)] = \mathcal{N}[\mathbf{x}(n); \mathbf{A}(n)\mathbf{x}(n-1), \mathbf{C}_u(n)]$ yielding (13). The posterior pdf in (14) can be obtained from Bayes' rule

$$p\,[\mathbf{x}(n)|\mathbf{b}_{0:n-1}, b(n)] = \frac{p\,[b(n)|\mathbf{x}(n), \mathbf{b}_{0:n-1}]\,p\,[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]}{p\,[b(n)|\mathbf{b}_{0:n-1}]}.$$
$$(16)$$

But the term $p[b(n)|\mathbf{x}(n), \mathbf{b}_{0:n-1}] := \Pr\{b(n) = \pm 1|\mathbf{x}(n), \mathbf{b}_{0:n-1}\}$ can be easily expressed in terms of the Gaussian tail function

$$\Pr\{b(n) = \pm 1|\mathbf{x}(n), \mathbf{b}_{0:n-1}\} = \Pr\{y(n) \gtrless \hat{y}(n|n-1)|\mathbf{x}(n)\}$$
$$= \Pr\{v(n) \gtrless \mathbf{h}^T(n)\,[\hat{\mathbf{x}}(n|n-1) - \mathbf{x}(n)]\,|\mathbf{x}(n)\}$$
$$= Q\left[\pm\frac{\mathbf{h}^T(n)\,[\hat{\mathbf{x}}(n|n-1) - \mathbf{x}(n)]}{\sigma_v(n)}\right] \quad (17)$$

where the first equality follows from the definition of the SOI $b(n)$ in (12) and the fact that since $\mathbf{x}(n)$ is given we can ignore the conditioning on $\mathbf{b}_{0:n-1}$; the second equality is obtained by substituting $y(n)$ for the observation model expression in (9); and the last equality is a consequence of the observations' noise distribution, $p[v(n)] = \mathcal{N}[v(n); 0, \sigma_v^2(n)]$.

Substituting (17) into (16) yields (14) after setting $\alpha_n := 1/\Pr\{b(n) = \pm 1|\mathbf{b}_{0:n-1}\}$. ∎

Two recursive algorithms for computing the MMSE $\hat{\mathbf{x}}(n|n)$ can be derived from Proposition 1. Algorithm 1-A is run at the sensors when the scheduling algorithm dictates that is their turn to transmit the SOI. At this time slot, the sensor computes the distribution $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]$ using (13) from where it predicts the state value by numerically evaluating $\hat{\mathbf{x}}(n|n-1) = \int_{\mathbb{R}^p} \mathbf{x}(n)p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]d\mathbf{x}(n)$. Based on this prediction, the sensor evaluates $\hat{y}(n|n-1) = \mathbf{h}^T(n)\hat{\mathbf{x}}(n|n-1)$ as in (7) in order to obtain and transmit the SOI $b(n)$ as defined in (12). Algorithm 1-B is run by all sensors during the life of the network to keep track of the state $\mathbf{x}(n)$ via the filtered estimate (corrector) $\hat{\mathbf{x}}(n|n)$. To this end, all sensors compute the pdf $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]$ using (13), and subsequently apply (14) to find $p[\mathbf{x}(n)|\mathbf{b}_{0:n}]$. The estimate of interest $\hat{\mathbf{x}}(n|n)$ is obtained by numerical integration of the expression in (6).

Albeit optimal, the process described by Algorithms 1-A and 1-B requires numerical integration at three different times. We first have to evaluate the integral necessary to obtain $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]$ as stated in (13) for step 1 of Algorithm 1-A and step 2 of Algorithm 2-B. A second numerical integration in step 2 of Algorithm 1-A is required to compute $\hat{\mathbf{x}}(n|n-1)$ and another one in step 5 of Algorithm 1-B to compute the desired estimate, $\hat{\mathbf{x}}(n|n)$. As these can be prohibitively expensive for a resource limited WSN, we are motivated to pursue a reduced-complexity approximation that we introduce next.

Algorithm 1–A Exact MMSE estimation – Observation and transmission

---

Require: $p[\mathbf{x}(n-1)|\mathbf{b}_{0:n-1}]$

Ensure: $b(n)$

 1: Obtain the distribution $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]$ using (13)

 2: Compute the prediction $\hat{\mathbf{x}}(n|n-1) = \int_{\mathbb{R}^p} \mathbf{x}(n)p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]d\mathbf{x}(n)$

 3: Find $\hat{y}(n|n-1)$ using (7)

 4: Construct $b(n)$ as in (12)

 5: Transmit $b(n)$

Algorithm 1–B Exact MMSE estimation – Reception and estimation

---

Require: prior distribution $p[\mathbf{x}(-1)]$

 1: **for** $n = 0$ to ∞ **do** {repeat for the life of the network}

 2:     Obtain the distribution $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]$ using (13)

 3:     Receive $b(n)$

 4:     Obtain the posterior distribution $p[\mathbf{x}(n)|\mathbf{b}_{0:n}]$ using (14)

 5:     Form the desired estimate, $\hat{\mathbf{x}}(n|n) = \int_{\mathbb{R}^p} \mathbf{x}(n)p[\mathbf{x}(n)|\mathbf{b}_{0:n}]d\mathbf{x}(n)$.

 6: **end for**

B. Approximate MMSE Estimator

A customary simplification in non-linear filtering is to assume that the pdf $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}] = \mathcal{N}[\mathbf{x}(n); \hat{\mathbf{x}}(n|n-1), \mathbf{M}(n|n-1)]$ is Gaussian; see, e.g., [16]. In general, the normal approximation of $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]$ is introduced to reduce the problem of tracking the evolution of a pdf to that of tracking its mean $\hat{\mathbf{x}}(n|n-1)$ and covariance $\mathbf{M}(n|n-1)$. For the problem at hand though, it also leads to a very simple algorithm as asserted by the following proposition.

Proposition 2: Consider the vector state-scalar observation model in (9) and binary observations defined as in (12). If $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}] = \mathcal{N}[\mathbf{x}(n); \hat{\mathbf{x}}(n|n-1), \mathbf{M}(n|n-1)]$, then the MMSE estimator $\hat{\mathbf{x}}(n|n)$ can be obtained from the recursions

$$\hat{\mathbf{x}}(n|n-1) = \mathbf{A}(n)\hat{\mathbf{x}}(n-1|n-1) \tag{18}$$

$$\mathbf{M}(n|n-1) = \mathbf{A}(n)\mathbf{M}(n-1|n-1)\mathbf{A}^T(n) + \mathbf{C}_u(n) \tag{19}$$

$$\hat{\mathbf{x}}(n|n) = \hat{\mathbf{x}}(n|n-1) + \frac{(\sqrt{2/\pi})\mathbf{M}(n|n-1)\mathbf{h}(n)}{\sqrt{\mathbf{h}^T(n)\mathbf{M}(n|n-1)\mathbf{h}(n) + \sigma_v^2(n)}}b(n) \tag{20}$$

$$\mathbf{M}(n|n) = \mathbf{M}(n|n-1) - \frac{(2/\pi)\mathbf{M}(n|n-1)\mathbf{h}(n)\mathbf{h}^T(n)\mathbf{M}(n|n-1)}{\mathbf{h}^T(n)\mathbf{M}(n|n-1)\mathbf{h}(n) + \sigma_v^2(n)}. \tag{21}$$

Proof: The predictor recursions (18) and (19) are identical to (7) and (8), respectively, and are included here for completeness. To establish (20), recall that the conditional mean can be obtained by averaging $\mathbf{x}(n)$ over the posterior pdf $p[\mathbf{x}(n)|\mathbf{b}_{0:n}]$:

$$\hat{\mathbf{x}}(n|n) := \mathrm{E}\left[\mathbf{x}(n)|\mathbf{b}_{0:n}\right] = \int_{\mathbb{R}^p} \mathbf{x}(n)p\left[\mathbf{x}(n)|\mathbf{b}_{0:n}\right]d\mathbf{x}(n). \tag{22}$$

Using Bayes' rule, we can find the posterior $p[\mathbf{x}(n)|\mathbf{b}_{0:n}] = p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}, b(n)]$ as

$$p\left[\mathbf{x}(n)|\mathbf{b}_{0:n-1}, b(n)\right] = \frac{p\left[b(n)|\mathbf{x}(n), \mathbf{b}_{0:n-1}\right]p\left[\mathbf{x}(n)|\mathbf{b}_{0:n-1}\right]}{p\left[b(n)|\mathbf{b}_{0:n-1}\right]}. \tag{23}$$

We now examine the three terms on the right-hand side of (23). The first one is $p[b(n)|\mathbf{x}(n), \mathbf{b}_{0:n-1}] := \mathrm{Pr}\{b(n) = \pm 1|\mathbf{x}(n), \mathbf{b}_{0:n-1}\}$, which after repeating the steps used to establish (17) in the proof of Proposition 1 can be expressed in terms of the Gaussian tail function

$$\mathrm{Pr}\left\{b(n) = \pm 1|\mathbf{x}(n), \mathbf{b}_{0:n-1}\right\}$$
$$= Q\left[\pm \frac{\mathbf{h}^T(n)\left[\hat{\mathbf{x}}(n|n-1) - \mathbf{x}(n)\right]}{\sigma_v(n)}\right]. \tag{24}$$

To obtain an expression for the term $p[b(n)|\mathbf{b}_{0:n-1}] := \mathrm{Pr}\{b(n) = \pm 1|\mathbf{b}_{0:n-1}\}$, we use the normal assumption on the distribution of $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]$ to obtain

$$\mathrm{Pr}\{b(n) = \pm 1|\mathbf{b}_{0:n-1}\} = \mathrm{Pr}\{y(n) \gtrless \hat{y}(n|n-1)|\mathbf{b}_{0:n-1}\} = 1/2 \tag{25}$$

where the first equality follows from the definition of $b(n)$ in (12). To obtain the second equality note that Gaussianity of $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}]$ implies that of $p[y(n)|\mathbf{b}_{0:n-1}]$ since $y(n)$ is a linear transformation of $\mathbf{x}(n)$; and also that the probability of a normal variable to be greater or smaller than its mean equals 1/2.

Substituting (24) and (25) into (23) and using the (assumed) normal distribution $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}] = \mathcal{N}[\mathbf{x}(n); \hat{\mathbf{x}}(n|n-1), \mathbf{M}(n|n-1)]$, we obtain an expression for the posterior distribution $p[\mathbf{x}(n)|\mathbf{b}_{0:n}]$ that we substitute in (22) to arrive at

$$\hat{\mathbf{x}}(n|n) = 2\int_{\mathbb{R}^p} \mathbf{x}(n)\mathcal{N}\left[\mathbf{x}(n); \hat{\mathbf{x}}(n|n-1), \mathbf{M}(n|n-1)\right]$$
$$\times Q\left[\pm \frac{\mathbf{h}^T(n)\left[\hat{\mathbf{x}}(n|n-1) - \mathbf{x}(n)\right]}{\sigma_v(n)}\right]d\mathbf{x}(n). \tag{26}$$

In the Appendix, we prove that the integral in (26) can be reduced to (20).

To obtain (21), we write $\hat{\mathbf{x}}(n|n) = \hat{\mathbf{x}}(n|n-1) + \mathbf{k}(n)b(n)$ with the explicit value of $\mathbf{k}(n)$ as deduced from (20), so that we can write the ECM as

$$\mathbf{M}(n|n) := \mathrm{E}\left[(\mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1) - \mathbf{k}(n)b(n))\right.$$
$$\left.\times (\mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1) - \mathbf{k}(n)b(n))^T\right]$$
$$= \mathbf{M}(n|n-1) + \mathbf{k}(n)\mathbf{k}^T(n)\mathrm{E}\left[b^2(n)\right]$$
$$- 2\mathbf{k}(n)\mathrm{E}\left[b(n)\mathbf{x}^T(n)\right] \tag{27}$$

where the first equality follows by definition and, in the second equality, we used that $\mathbf{M}(n|n-1) := \mathrm{E}[(\mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1))(\mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1))^T]$. The last term in (27) can be further simplified after recalling that $\mathrm{Pr}\{b(n) = \pm 1|\mathbf{b}_{0:n-1}\} = 1/2$, and using the theorem of total probability to obtain

$$\mathrm{E}\left[b(n)\mathbf{x}^T(n)\right] = \frac{1}{2}\mathrm{E}\left[\mathbf{x}^T(n)|b(n) = 1\right] - \frac{1}{2}\mathrm{E}\left[\mathbf{x}^T(n)|b(n) = -1\right]$$
$$= \frac{1}{2}\mathbf{k}^T(n) - \frac{1}{2}\left[-\mathbf{k}^T(n)\right] = \mathbf{k}^T(n). \tag{28}$$

Algorithm 2-A Time invariant vector state - scalar observation SOI-KF – Observation and transmission

---
Require: $\hat{\mathbf{x}}(n-1|n-1)$ and $\mathbf{M}(n-1|n-1)$
Ensure: $b(n)$
1: Compute the prediction $\hat{\mathbf{x}}(n|n-1)$ using (18)
2: Find $\hat{y}(n|n-1)$ using (7)
3: Construct $b(n)$ as in (12)
4: Transmit $b(n)$

Algorithm 2-B Time invariant vector state - scalar observation SOI-KF – Reception and estimation

---
Require: prior estimate $\hat{\mathbf{x}}(-1|-1)$ and ECM $\mathbf{M}(-1|-1)$
1: **for** $n = 0$ to ∞ **do** {repeat for the life of the network}
2:     Compute $\hat{\mathbf{x}}(n|n-1)$ and $\mathbf{M}(n|n-1)$ using (18) and (19)
3:     Receive $b(n)$
4:     Compute $\hat{\mathbf{x}}(n|n)$ and $\mathbf{M}(n|n)$ using (20) and (21)
5: **end for**

Substituting (28) into (27) and noting that $\mathrm{E}[b^2(n)] = 1$, we obtain

$$\mathbf{M}(n|n) = \mathbf{M}(n|n-1) - \mathbf{k}(n)\mathbf{k}^T(n) \qquad (29)$$

which after using the expression for $\mathbf{k}(n)$ leads to (21). ∎

As we commented earlier, the simplification $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}] = \mathcal{N}[\mathbf{x}(n); \hat{\mathbf{x}}(n|n-1), \mathbf{M}(n|n-1)]$ yields the low-complexity SOI-KF that implements distributed state estimation based on single-bit observations using the recursions (18)–(21). To estimate $\mathbf{x}(n)$, we only require a few basic algebraic operations per iteration. Moreover, the SOI-KF recursion is strikingly reminiscent of the KF recursions (7), (8), and (11). The ECM updates in particular are identical except for the $2/\pi$ factor in (21).

The algorithmic description of the SOI-KF is summarized in Algorithm 2-A which is run by the sensors as dictated by the scheduling algorithm; and Algorithm 2-B which is continuously run by all sensors to track $\mathbf{x}(n)$. These algorithms are to be contrasted with their exact MMSE counterparts (Algorithms 1-A and 1-B) to note that the numerical integrations have been replaced by simple algebraic expressions. Indeed, the SOI-KF observation and transmission Algorithm 2-A computes the prediction $\hat{y}(n|n-1)$ by successive application of (18) and (7) to compute and transmit the SOI in (12). The reception and estimation Algorithm 2-B is identical to a KF algorithm except for the (minor) differences in the update equations.

A few remarks are now in order.

Remark 2: It is possible to express the SOI-KF corrector in (20) in a form that exemplifies its link with the KF corrector in (11). Indeed, if we define the SOI-KF innovation as

$$\tilde{b}(n|n-1) := \sqrt{(2/\pi)\left[\mathbf{h}^T(n)\mathbf{M}(n|n-1)\mathbf{h}(n) + \sigma_v^2(n)\right]}b(n) \qquad (30)$$

we can re-write the SOI-KF corrector as

$$\hat{\mathbf{x}}(n|n) = \hat{\mathbf{x}}(n|n-1)$$
$$+ \frac{\mathbf{M}(n|n-1)\mathbf{h}(n)}{\mathbf{h}^T(n)\mathbf{M}(n|n-1)\mathbf{h}(n) + \sigma_v^2(n)}\tilde{b}(n|n-1). \qquad (31)$$

Note that (31) is identical to (11) if we replace $\tilde{b}(n|n-1)$ with the innovation $\tilde{y}(n|n-1) = y(n) - \hat{y}(n|n-1)$. Moreover, note that the units of $\tilde{b}(n|n-1)$ and $\tilde{y}(n|n-1)$ are the same, and that $\mathrm{E}[\tilde{b}(n|n-1)] = \mathrm{E}[\tilde{y}(n|n-1)] = 0$. Even more interesting [cf. (11) and (32)]

$$\mathrm{E}\left[\tilde{b}^2(n|n-1)\right] = (2/\pi)\left[\mathbf{h}^T(n)\mathbf{M}(n|n-1)\mathbf{h}(n) + \sigma_v^2(n)\right]$$
$$= (2/\pi)\mathrm{E}\left[\tilde{y}^2(n|n-1)\right] \qquad (32)$$

which explains the relationship between the ECM corrections for the KF in (11) and for the SOI-KF in (21). The difference between (11) and (31) is that in the SOI-KF the magnitude of the correction at each step is determined by the magnitude of $\mathrm{E}[\tilde{b}^2(n|n-1)]$ and it is the same regardless of how large or small the actual innovation $\tilde{y}(n|n-1)$ is.

Remark 3: As $\sigma_v^2 \to \infty$, the Gaussian tail function $Q[\pm\mathbf{h}^T(n)[\hat{\mathbf{x}}(n|n-1) - \mathbf{x}(n)]/\sigma_v]$ converges uniformly to $1/2$ and, consequently, $p[\mathbf{x}(n)|\mathbf{b}_{0:n}]$ converges uniformly to a normal distribution. Thus, the assumption $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}] = \mathcal{N}[\mathbf{x}(n); \hat{\mathbf{x}}(n|n-1), \mathbf{M}(n|n-1)]$ holds asymptotically as $\sigma_v^2 \to \infty$. For this reason, the SOI-KF yields an accurate approximation of the MMSE estimator at low signal to noise ratios (SNRs). The accuracy of the approximation decreases as the SNR increases.

Remark 4: When matrices in the state–space model (4) are time-invariant, it is well known that the KF converges to the Wiener filter (WF) as $n \to \infty$, which, as is also known, can be approximately implemented by the least-mean-squares (LMS) algorithm. Consequently, one may wonder whether there is a relation between the SOI-KF of this paper and the sign-LMS [28, Sec. 5.7]. Apart from offering low complexity versions of the KF and WF, respectively, the two schemes are fundamentally different. In the sign-LMS algorithm, the SOI sequence is used to recursively update the coefficients of the WF which has as input the analog-amplitude observations. In the SOI-KF, however, the gain coefficients are computed from the KF correlation matrices and the SOI sequence is used as the filter input.

IV. VECTOR STATE-VECTOR OBSERVATION CASE

The general vector state-vector observation case defined by (4) can be reduced to the vector state-scalar observation problem considered in Section III-B. If the noise vectors $\mathbf{v}(n)$ are white, i.e., $\mathbf{C}_v(n) = \mathbf{I}$ then Proposition 2 can be applied verbatim, for $\mathbf{v}(n) := [v_1(n), \ldots, v_M(n)]^T$ is not more than a collection of independent scalar observations $v_1(n), \ldots, v_M(n)$. If $\mathbf{v}(n)$ is not white, the idea is to whiten the observations so that we can rewrite the problem as a sequence of vector state-scalar observation problems. To this end, we define the observation $\mathbf{y}_0(n) := \mathbf{C}_v^{-1/2}(n)\mathbf{y}(n)$ to obtain [c.f. (4)]

$$\mathbf{y}_0(n) = \mathbf{C}_v^{-1/2}(n)\mathbf{H}(n)\mathbf{x}(n) + \mathbf{C}_v^{-1/2}(n)\mathbf{v}$$
$$:= \mathbf{H}_0(n)\mathbf{x}(n) + \mathbf{v}_0(n) \qquad (33)$$

where $\mathrm{E}[\mathbf{v}_0(n)\mathbf{v}_0^T(n)] = \mathbf{I}$. For future reference, we write $\mathbf{y}_0(n) := [y_0(n,1), \ldots, y_0(n,M)]$, $\mathbf{v}_0(n) := [v_0(n,1), \ldots, v_0(n,M)]$ and $\mathbf{H}_0(n) :=$

$[\mathbf{h}_0(n,1),\ldots,\mathbf{h}_0(n,M)]^T$ that allows us to write (33) componentwise as

$$y_0(n,m) = \mathbf{h}_0^T(n,m)\mathbf{x}(n) + v_0(n,m), \qquad m \in [1,M] \quad (34)$$

where the observation noise variance is $\sigma_{v_0}^2 := \mathrm{E}[v_0^2(n,m)] = 1$.

Equation (34) has the same form as (9) in the sense that the state $\mathbf{x}(n)$ is a vector but the observation $y_0(n,m)$ is scalar. Mimicking the treatment in Section III, we define $\mathbf{b}(n, 1 : m) := [b(n,1),\ldots,b(n,m)]^T$ and introduce the MMSE estimator

$$\hat{\mathbf{x}}(n|n-1,m) = \mathrm{E}\left[\mathbf{x}(n)|\mathbf{b}_{0:n-1}, \mathbf{b}(n,1:m)\right] \quad (35)$$

which is the MMSE estimator based on past messages and the first m components of the current message. We adopt the convention $\hat{\mathbf{x}}(n|n-1,0) := \hat{\mathbf{x}}(n|n-1)$, and note that $\hat{\mathbf{x}}(n|n-1,M) := \hat{\mathbf{x}}(n|n)$ with $\hat{\mathbf{x}}(n|n-1)$ as defined in (7) and $\hat{\mathbf{x}}(n|n)$ as defined in (6). From (35), we obtain the MMSE predictor of $y_0(n,m)$ as [c.f. (34) and (35)]

$$\begin{aligned}\hat{y}_0(n,m|n-1,m-1) &:= \mathrm{E}\left[y_0(n,m)|\mathbf{b}_{0:n-1}, \mathbf{b}(n,1:m-1)\right] \\ &= \mathbf{h}_0^T(n,m)\hat{\mathbf{x}}(n|n-1,m-1). \quad (36)\end{aligned}$$

From (36), we define the SOI observations for the vector state-vector observation problem as

$$b(n,m) := \mathrm{sign}\left[y_0(n,m) - \hat{y}_0(n,m|n-1,m-1)\right], \quad m \in [1,M]. \quad (37)$$

Setting aside the necessary differences in notation, the problem of finding the MMSE estimator in (35) based on the observation model (34) when the binary observations are given by (37) is equivalent to a sequence of M MMSE estimation problems for the vector state-scalar observation model in (9) with binary observations as in (12). An approximate MMSE for this problem was summarized in Proposition 2 that, with proper notational modifications, can now be generalized as follows.

Proposition 3: Consider the vector state-vector observation model defined by (4), binary observations defined as in (37) and let $\mathbf{H}_0(n) := [\mathbf{h}_0(n,1),\ldots,\mathbf{h}_0(n,M)]^T$ be defined as $\mathbf{H}_0(n) := \mathbf{C}_v^{-1/2}(n)\mathbf{H}(n)$ [c.f. (33)]. If $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}, \mathbf{b}(n,1:m-1)] = \mathcal{N}[\mathbf{x}(n); \hat{\mathbf{x}}(n|n-1,m-1), \mathbf{M}(n|n-1,m-1)]$, then the MMSE estimate $\hat{\mathbf{x}}(n|n)$ can be obtained from the recursions

$$\hat{\mathbf{x}}(n|n-1) = \mathbf{A}(n)\hat{\mathbf{x}}(n-1|n-1) \quad (38)$$

$$\mathbf{M}(n|n-1) = \mathbf{A}(n)\mathbf{M}(n-1|n-1)\mathbf{A}^T(n) + \mathbf{C}_u(n) \quad (39)$$

$$\mathbf{k}(n,m) = \frac{(\sqrt{2/\pi})\mathbf{M}(n|n-1,m-1)\mathbf{h}_0(n,m)}{\sqrt{1 + \mathbf{h}_0^T(n,m)\mathbf{M}(n|n-1,m-1)\mathbf{h}_0(n,m)}} \quad (40)$$

$$\hat{\mathbf{x}}(n|n-1,m) = \hat{\mathbf{x}}(n|n-1,m-1) + \mathbf{k}(n,m)b(n,m) \quad (41)$$

$$\mathbf{M}(n|n-1,m) = \mathbf{M}(n|n-1,m-1) - \mathbf{k}(n,m)\mathbf{k}^T(n,m) \quad (42)$$

where for each time index n, steps (40) to (42) are repeated for $m \in [1,M]$. We adopt the conventions $\hat{\mathbf{x}}(n|n-1,0) := \hat{\mathbf{x}}(n|n-1)$ and $\mathbf{M}(n|n-1,0) := \mathbf{M}(n|n-1)$, and note that the MMSE estimate and the ECM are given by $\hat{\mathbf{x}}(n|n) = \hat{\mathbf{x}}(n|n-1,M)$ and $\mathbf{M}(n|n) = \mathbf{M}(n|n-1,M)$.

Proof: As pointed out earlier, Proposition 3 follows from repeated application of Proposition 2. Indeed, if we define the vector $\mathbf{x}(n,m) = \mathbf{x}(n) \; \forall m \in [1,M]$ the state equation for $\mathbf{x}(n,m)$ can be written as

$$\begin{aligned}\mathbf{x}(n,1) &= \mathbf{A}(n)\mathbf{x}(n-1,M) + \mathbf{u}(n), \quad m = 1 \\ \mathbf{x}(n,m) &= \mathbf{x}(n,m-1), \qquad\qquad\quad m \neq 1.\end{aligned} \quad (43)$$

Algorithm 3–A SOI-KF – Observation and transmission

Require: $\hat{\mathbf{x}}(n-1|n-1)$ and $\mathbf{M}(n-1|n-1)$
Ensure: $\mathbf{b}(n)$
 1: Compute $\hat{\mathbf{x}}(n|n-1)$ and $\mathbf{M}(n|n-1)$ using (38) and (39).
 2: $\mathbf{y}_0(n) := \mathbf{C}_v^{-1/2}(n)\mathbf{y}(n)$ and $\mathbf{H}_0(n) := \mathbf{C}_v^{-1/2}(n)\mathbf{H}(n)$.
 3: **for** $m = 1$ to M **do**
 4:     Compute $\hat{y}(n,m|n-1,m-1)$ using (36)
 5:     Compute $b(n,m)$ using (37)
 6:     Compute $\mathbf{k}(n,m)$, $\hat{\mathbf{x}}(n|n-1,m)$, and $\mathbf{M}(n|n-1,m)$ using (40) , (41) and (42)
 7: **end for**
 8: Transmit $\mathbf{b}(n) = [b(n,1),\ldots,b(n,M)]$

On the other hand, the whitened observations can be written as [c.f. (34) with $\mathbf{x}(n) = \mathbf{x}(n,m)$]

$$y_0(n,m) = \mathbf{h}_0^T(n,m)\mathbf{x}(n,m) + v_0(n,m). \quad (44)$$

Define now the MMSE estimators $\hat{\mathbf{x}}(n,m|n-1,m) := \mathrm{E}[\mathbf{x}(n,m)|\mathbf{b}_{0:n-1}, \mathbf{b}(n,1:m)]$ and $\hat{\mathbf{x}}(n,m|n-1,m-1) := \mathrm{E}[\mathbf{x}(n,m)|\mathbf{b}_{0:n-1}, \mathbf{b}(n,1:m-1)]$ with corresponding ECM $\mathbf{M}(n,m|n-1,m)$ and $\mathbf{M}(n,m|n-1,m-1)$. Applying Proposition 2, we obtain the prediction recursions for $m = 1$ [c.f. (18), (19), and (43)]

$$\hat{\mathbf{x}}(n,1|n-1,0) = \mathbf{A}(n)\hat{\mathbf{x}}(n-1,M|n-1,0) \quad (45)$$
$$\mathbf{M}(n,1|n-1,0) = \mathbf{A}(n)\mathbf{M}(n-1,M|n-1,0)\mathbf{A}^T(n) + \mathbf{C}_u(n)$$

and for $m \in [2,M]$ [c.f. (18), (19), and (43)]

$$\begin{aligned}\hat{\mathbf{x}}(n,m|n-1,m-1) &= \hat{\mathbf{x}}(n,m-1|n-1,m-1) \\ \mathbf{M}(n,m|n-1,m-1) &= \mathbf{M}(n,m-1|n-1,m-1). \quad (46)\end{aligned}$$

From Proposition 2, we also obtain the correction recursions. Upon defining the gain

$$\mathbf{k}(n,m) = \frac{(\sqrt{2/\pi})\mathbf{M}(n,m|n-1,m-1)\mathbf{h}_0(n,m)}{\sqrt{1 + \mathbf{h}_0^T(n,m)\mathbf{M}(n,m|n-1,m-1)\mathbf{h}_0(n,m)}} \quad (47)$$

the filtered estimate and ECM can be written as [c.f. (20), (21), (44), and (47)]

$$\begin{aligned}\hat{\mathbf{x}}(n,m|n-1,m) &= \hat{\mathbf{x}}(n,m|n-1,m-1) + \mathbf{k}(n,m)b(n,m) \\ \mathbf{M}(n,m|n-1,m) &= \mathbf{M}(n,m|n-1,m-1) - \mathbf{k}(n,m)\mathbf{k}^T(n,m). \quad (48)\end{aligned}$$

Note however, that since $\mathbf{x}(n,m) = \mathbf{x}(n)$ for $m \in [1,M]$, we have that $\hat{\mathbf{x}}(n,m|n-1,m) = \hat{\mathbf{x}}(n|n-1,m)$ and $\hat{\mathbf{x}}(n,m|n-1,m-1) = \hat{\mathbf{x}}(n|n-1,m-1)$; and likewise for the ECMs: $\mathbf{M}(n,m|n-1,m) = \mathbf{M}(n|n-1,m)$ and $\mathbf{M}(n,m|n-1,m-$

Algorithm 3–B SOI-KF – Reception and estimation

Require: prior estimate $\hat{\mathbf{x}}(-1|-1)$ and ECM $\mathbf{M}(-1|-1)$
1: **for** $n = 0$ to ∞ **do** {repeat for the life of the network}
2:     Compute $\hat{\mathbf{x}}(n|n-1)$ and $\mathbf{M}(n|n-1)$ using (38) and (39).
3:     $\mathbf{H}_0(n) := \mathbf{C}_v^{-1/2}(n)\mathbf{H}(n)$
4:     Receive $\mathbf{b}(n)$
5:     **for** $m = 1$ to M **do**
6:         Compute $\mathbf{k}(n,m)$, $\hat{\mathbf{x}}(n|n-1,m)$, and $\mathbf{M}(n|n-1,m)$ using (40) , (41) and (42)
7:     **end for**
8:     $\hat{\mathbf{x}}(n|n) = \hat{\mathbf{x}}(n|n-1,M)$, $\mathbf{M}(n|n) = \mathbf{M}(n|n-1,M)$
9: **end for**

$1) = \mathbf{M}(n|n-1, m-1)$. To obtain (38) and (39), it suffices to substitute the latter into (45). To obtain (40)–(42), we simply make these same substitutions in (48) after plugging (46) into (48). ∎

The algorithmic description of the SOI-KF is summarized in Algorithms 3-A and 3-B. Algorithm 3-A is run at the sensors when the scheduling algorithm dictates that it is their turn to transmit an observation. When this happens, the sensor runs the predictor using (38) and (39) (step 1) and whitens the observation $\mathbf{y}(n)$ (step 2). Subsequently, it recursively computes partial MMSE estimators via (36) and (40)–(42) in order to obtain the binary observations $b(n, m)$ by means of (37). When this process is complete, the message $\mathbf{b}(n)$ is transmitted. Interestingly enough, when the observations are scalar, Algorithm 3-A amounts to sequential application of steps 1, 2, 4, 5, and 8; which is, of course, equivalent to Algorithm 2-A.

Algorithm 3-B is continuously ran by the sensors to estimate the state $\mathbf{x}(n)$. At each time slot n we compute the predictors along with $\mathbf{H}_0(n)$ and move on to process the received message $\mathbf{b}(n)$. Processing of $\mathbf{b}(n)$ entails recursive application of (40)–(42) for the M entries of $\mathbf{b}(n)$. After this process is complete, we obtain the MMSE estimate $\hat{\mathbf{x}}(n|n)$.

V. PERFORMANCE ANALYSIS

By definition, any MMSE estimator minimizes the trace of the corresponding ECM. Thus, to the extent that the approximation $p[\mathbf{x}(n)|\mathbf{b}_{0:n-1}] = \mathcal{N}[\mathbf{x}(n); \hat{\mathbf{x}}(n|n-1), \mathbf{M}(n|n-1)]$ is accurate enough, the SOI-KF in Proposition 2 is optimum in the sense of minimizing $\text{tr}[\mathbf{M}(n|n)]$ and $\text{tr}[\mathbf{M}(n|n-1)]$. However, this optimality does not provide any insight with respect to the performance of the SOI-KF relative to the MMSE based on the original observations which are used by the clairvoyant KF in (7), (8), and (11). In this section, we compare $\text{tr}[\mathbf{M}(n|n)]$ and $\text{tr}[\mathbf{M}(n|n-1)]$ for the SOI-KF with $\text{tr}[\mathbf{M}^K(n|n)]$ and $\text{tr}[\mathbf{M}^K(n|n-1)]$ reserved to denote the corresponding quantities for the KF.

To simplify notation, define $\mathbf{M}(n) := \mathbf{M}(n|n-1)$. Interestingly, $\mathbf{M}(n)$ is independent of the observations $\mathbf{b}_{0:n}$, and regardless of the data we can find $\mathbf{M}(n)$ by solving the discrete-time Ricatti equation that is obtained by substituting the

expression for $\mathbf{M}(n|n)$ in (21) into the ECM update for $\mathbf{M}(n+1) := \mathbf{M}(n+1|n)$ in (19)

$$\mathbf{M}(n+1) = \mathbf{A}(n+1)\mathbf{M}(n)\mathbf{A}^T(n+1) + \mathbf{C}_u(n+1)$$
$$-(2/\pi)\frac{\mathbf{A}(n+1)\mathbf{M}(n)\mathbf{h}(n)\mathbf{h}^T(n)\mathbf{M}(n)\mathbf{A}^T(n+1)}{\mathbf{h}^T(n)\mathbf{M}(n)\mathbf{h}(n) + \sigma_v^2(n)}. \quad (49)$$

Likewise, upon defining $\mathbf{M}^K(n) := \mathbf{M}^K(n|n-1)$, we obtain the discrete-time Ricatti equation for the clairvoyant KF [c.f. (8) and (11)]

$$\mathbf{M}^K(n+1) = \mathbf{A}(n+1)\mathbf{M}^K(n)\mathbf{A}^T(n+1) + \mathbf{C}_u(n+1)$$
$$-\frac{\mathbf{A}(n+1)\mathbf{M}^K(n)\mathbf{h}(n)\mathbf{h}^T(n)\mathbf{M}^K(n)\mathbf{A}^T(n+1)}{\mathbf{h}^T(n)\mathbf{M}^K(n)\mathbf{h}(n) + \sigma_v^2(n)}. \quad (50)$$

Notice that (49) and (50) differ only by the $(2/\pi)$ factor in the numerator of the ratio in (49). A possible performance comparison could be to solve the difference (49) and (50) for specific models and compare $\text{tr}[\mathbf{M}(n)]$ with $\text{tr}[\mathbf{M}^K(n)]$. However, better insight can be gained by recalling the underlying continuous-time model, for which we start with the following definition.

Definition 1: Consider the continuous-time model (1), (2) and a family of corresponding discrete-time models (4) parameterized by T_s. Let $\mathbf{M}(T_s; n|n)$ and $\mathbf{M}(T_s; n|n-1)$ be the ECM of the filtered and predicted estimates of the SOI-KF in Proposition 2 when sampling period T_s is used in (4). Then, the continuous-time ECM $\mathbf{M}_c(t)$ is defined as

$$\mathbf{M}_c(t) := \mathbf{M}_c(nT_s) := \lim_{T_s \to 0} \mathbf{M}(T_s; n|n)$$
$$= \lim_{T_s \to 0} \mathbf{M}(T_s; n|n-1). \quad (51)$$

An equivalent definition can be written for the clairvoyant KF whose continuous-time ECM will be denoted $\mathbf{M}_c^K(t)$ [20]. In general, the continuous-time MSE is easier to analyze but at the same time more general since, being independent of the sampling time, it provides insights about the fundamental properties of the problem. Moreover, it is well known that [20]

$$\text{tr}\left[\mathbf{M}(T_s; n|n)\right] \leq \text{tr}\left[\mathbf{M}_c(nT_s)\right] \leq \text{tr}\left[\mathbf{M}(T_s; n|n-1)\right] \quad \forall T_s. \quad (52)$$

Equation (52) reveals that the continuous-time MSE, $\text{tr}[\mathbf{M}_c(t)]$, serves as an upper (lower) bound for $\text{tr}[\mathbf{M}(T_s; n|n)]$ $(\text{tr}[\mathbf{M}(T_s; n|n-1)])$. The continuous-time ECM $\mathbf{M}_c(t)$ can be obtained by solving a continuous-time Ricatti equation as we show in the next proposition.

Proposition 4: For the SOI-KF introduced in Proposition 2, consider the continuous-time ECM $\mathbf{M}_c(t)$ given by Definition 1. Then, $\mathbf{M}_c(t)$ can be obtained as the solution of the differential equation

$$\dot{\mathbf{M}}_c(t) = \mathbf{A}_c(t)\mathbf{M}_c(t) + \mathbf{M}_c(t)\mathbf{A}_c^T(t) + \mathbf{C}_{u_c}(t)$$
$$-\mathbf{M}_c(t)\mathbf{h}_c(t)\left[\frac{\pi}{2}\sigma_{v_c}^2(t)\right]^{-1}\mathbf{h}_c^T(t)\mathbf{M}_c(t). \quad (53)$$

Proof: Consider a neighborhood around $t := nT_s$. To establish (53), it suffices to subtract $\mathbf{M}(n)$ from both sides of (49),

divide by T_s and let $T_s \to 0$. Indeed, the limit of the left-hand side of (49) is

$$
\begin{aligned}
\mathcal{L} &:= \lim_{T_s \to 0} \frac{\mathbf{M}(n+1) - \mathbf{M}(n)}{T_s} \\
&= \lim_{T_s \to 0} \frac{\mathbf{M}_c[(n+1)T_s] - \mathbf{M}(nT_s)}{T_s} = \dot{\mathbf{M}}_c(t) \quad (54)
\end{aligned}
$$

where the first equality follows from the definition of $\mathbf{M}_c(nT_s)$ in (51) and in the second equality, we used the definition of derivative and set $t := nT_s$. On the right-hand side, we start with the limit shown in

$$
\begin{aligned}
\mathcal{R}_1 &:= \lim_{T_s \to 0} \frac{\mathbf{A}(n+1)\mathbf{M}(n)\mathbf{A}^T(n+1) - \mathbf{M}(n)}{T_s} \quad (55) \\
&= \lim_{T_s \to 0} \frac{\boldsymbol{\Phi}[(n+1)T_s, nT_s]\mathbf{M}_c(t)\boldsymbol{\Phi}^T[(n+1)T_s, nT_s] - \mathbf{M}_c(t)}{T_s}
\end{aligned}
$$

where in the second equality we used the definitions of $\mathbf{A}(n+1) := \boldsymbol{\Phi}[(n+1)T_s, nT_s]$ and $\mathbf{M}_c(t) = \mathbf{M}_c(nT_s)$ in (51). However, since $\boldsymbol{\Phi}[(n+1)T_s, nT_s] = \mathbf{I} + \mathbf{A}_c(nT_s)T_s + \mathrm{o}(T_s)$, we find

$$
\begin{aligned}
\mathcal{R}_1 &= \lim_{T_s \to 0} \left[\frac{T_s \left[\mathbf{A}_c(t)\mathbf{M}_c(t) + \mathbf{M}_c(t)\mathbf{A}_c^T(t) \right]}{T_s} \right. \\
&\qquad \left. + \frac{T_s^2 \mathbf{A}_c(t)\mathbf{M}_c(t)\mathbf{A}_c^T(t) + \mathrm{o}(T_s)}{T_s} \right] \\
&= \mathbf{A}_c(t)\mathbf{M}_c(t) + \mathbf{M}_c(t)\mathbf{A}_c^T(t). \quad (56)
\end{aligned}
$$

Consider now the variance of the driving input whose limit is

$$
\begin{aligned}
\mathcal{R}_2 &:= \lim_{T_s \to 0} \frac{\mathbf{C}_u(n+1)}{T_s} \quad (57) \\
&= \lim_{T_s \to 0} \frac{1}{T_s} \int_{(n-1)T_s}^{nT_s} \boldsymbol{\Phi}(nT_s, \tau)\mathbf{C}_{u_c}(\tau)\boldsymbol{\Phi}^T(nT_s, \tau)d\tau = \mathbf{C}_{u_c}(t)
\end{aligned}
$$

where in obtaining the first equality we used the definition of $\mathbf{C}_u(n+1)$. To obtain the last equality we applied the mean value theorem and wrote $\boldsymbol{\Phi}(nT_s, \tau) = \mathbf{I} + \mathbf{A}_c(nT_s)(\tau - nT_s) + \mathrm{o}(\tau - nT_s)$.

For the remaining term on the right-hand side of (49), we define the limit

$$
\begin{aligned}
\mathcal{R}_3 &:= \lim_{T_s \to 0} \frac{(2/\pi)\mathbf{A}(n+1)\mathbf{M}(n)\mathbf{h}(n)\mathbf{h}^T(n)\mathbf{M}(n)\mathbf{A}^T(n+1)}{T_s \mathbf{h}^T(n)\mathbf{M}(n)\mathbf{h}(n) + T_s \sigma_v^2(n)} \\
&= \lim_{T_s \to 0} \frac{(2/\pi)\mathbf{A}_c(t)\mathbf{M}_c(t)\mathbf{h}(t)\mathbf{h}^T(t)\mathbf{M}_c(t)\mathbf{A}^T(t)}{T_s \mathbf{h}^T(t)\mathbf{M}_c(t)\mathbf{h}(t) + \sigma_{v_c}^2(t)} \\
&= \frac{(2/\pi)\mathbf{A}_c(t)\mathbf{M}_c(t)\mathbf{h}(t)\mathbf{h}^T(t)\mathbf{M}_c(t)\mathbf{A}^T(t)}{\sigma_{v_c}^2(t)} \quad (58)
\end{aligned}
$$

where in the second step the key substitution is $\sigma_v^2(n) = \sigma_{v_c}^2(t)/T_s$; and we also used the fact that $\mathbf{A}(n+1) = \mathbf{I} + \mathbf{A}_c(nT_s)T_s + \mathrm{o}(T_s)$, $\lim_{T_s \to 0} \mathbf{h}(n) = \mathbf{h}_c(t)$ and the definition of $\mathbf{M}_c(t)$ in (51).

Finally, note that according to (49) and the definitions of $\mathcal{L}$, $\mathcal{R}_j$ we have that $\mathcal{L} = \sum_{j=1}^{3} \mathcal{R}_j$. Combining this with the limit expressions in (54), (56), (57) and (58), we obtain (53) after rearranging terms. ∎
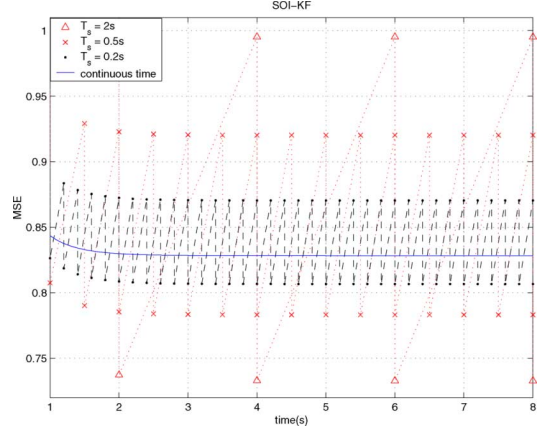

Fig. 3. MSEs $\mathrm{tr}[\mathbf{M}(T_s; n|n)]$ of the estimator and $\mathrm{tr}[\mathbf{M}(T_s; n|n-1)]$ of the predictor converge to the continuous-time MSE $\mathrm{tr}[\mathbf{M}_c(nT_s)]$ as T_s decreases $(\mathbf{A}_c(t) = \mathbf{I}, \mathbf{h}_c(t) = [1, 2]^T, \mathbf{C}_{u_c}(t) = \mathbf{I}, \text{ and } \sigma_{v_c}^2(t) = 1)$.

Either repeating Proposition 4 for the KF, or using standard references for the continuous-time KF, we know that $\mathbf{M}_c^K(t)$ can be obtained as the solution of the Ricatti equation [20]

$$
\begin{aligned}
\dot{\mathbf{M}}_c^K(t) &= \mathbf{A}_c(t)\mathbf{M}_c^K(t) + \mathbf{M}_c^K(t)\mathbf{A}_c^T(t) + \mathbf{C}_{u_c}(t) \\
&\quad - \mathbf{M}_c^K(t)\mathbf{h}_c(t)\left[\sigma_{v_c}^2(t)\right]^{-1}\mathbf{h}_c^T(t)\mathbf{M}_c^K(t) \quad (59)
\end{aligned}
$$

which is identical to (53) with the substitution $\sigma_{v_c}^2(t) \leftrightarrow (\pi/2)\sigma_{v_c}^2(t)$. Thus, the continuous-time MSE of the SOI-KF coincides with the continuous-time MSE of a KF with $\pi/2$ times larger measurement noise variance.

To state an analogous result for the vector state-vector observation SOI-KF of Proposition 3, we will need the definition of the $(\pi/2)$-equivalent system that we introduce next.

Definition 2: Consider a state-observation model as in (1), (2), where the noise autocorrelation is $\mathrm{E}[\mathbf{v}_c(t_1, k_1)\mathbf{v}_c^T(t_2, k_2)] = \mathbf{C}_{v_c}(t_1, k_1)\delta_c(t_1 - t_2)\delta(k_1 - k_2)$. We say that a model with otherwise identical parameters but noise autocorrelation $\mathrm{E}[\mathbf{v}_c(t_1, k_1)\mathbf{v}_c^T(t_2, k_2)] = (\pi/2)\mathbf{C}_{v_c}(t_1, k_1)\delta_c(t_1 - t_2)\delta(k_1 - k_2)$, is $(\pi/2)$-equivalent. For a given sampling period T_s, the KF for this latter model will be henceforth called the $(\pi/2)$-KF. We will denote its filtered and predicted ECM as $\mathbf{M}^{K,\pi/2}(T_s; n|n)$ and $\mathbf{M}^{K,\pi/2}(T_s; n|n-1)$ and the continuous-time ECM as $\mathbf{M}_c^{K,\pi/2}(t)$.

Using Definition 2, we can establish the relationship between the MSEs of the SOI-KF and the KF as follows.

Corollary 1: For the state-observation model in (1), (2), and its corresponding $(\pi/2)$-equivalent system, it holds that

$$
\mathbf{M}_c^{K,\pi/2}(t) = \mathbf{M}_c(t). \quad (60)
$$

Proof: Define the time index $l := Mn + (m-1)$ and apply Proposition 4 to the state-observation model defined by (4) and (34). Observe next that if $\mathbf{M}_c^{K,\pi/2}(t) = \mathbf{M}_c(t)$ holds for a model based on the observations $y_o(n, m)$, it also holds for a model based on $\mathbf{y}(n)$ because $\mathbf{y}_0(n) = \mathbf{C}_v^{-1/2}(n)\mathbf{y}(n)$ implies that the MMSE estimates are equal; i.e., $\mathrm{E}[\mathbf{x}(n)|\mathbf{y}_{0,0:n}] = \mathrm{E}[\mathbf{x}(n)|\mathbf{y}_{0:n}]$. ∎
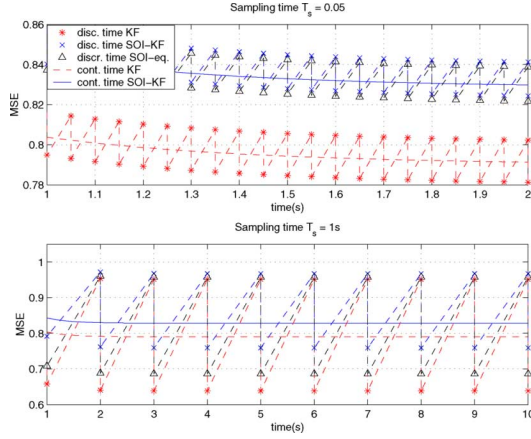
Fig. 4. MSE $\mathrm{tr}[\mathbf{M}(T_s; n|n)]$ of the SOI-KF and the MSE $\mathrm{tr}[\mathbf{M}^{\mathrm{K}, \pi/2}(T_s; n|n)]$ of the $(\pi/2)$-KF (SOI-eq.) are indistinguishable for small T_s; as T_s increases there is a noticeable but still small difference. The penalty with respect to $\mathrm{tr}[\mathbf{M}^{\mathrm{K}}(T_s; n|n)]$ is small for moderate T_s ($\mathbf{A}_c(t) = \mathbf{I}$, $\mathbf{h}_c(t) = [1,2]^T$, $\mathbf{C}_{u_c}(t) = \mathbf{I}$, and $\sigma_{v_c}^2(t) = 1$).

Corollary 1 establishes that the MSE of the SOI-KF is closely related to the MSE of the $(\pi/2)$-KF, since as $T_s \to 0$ the MSEs of these two filters are equal. For a particular example, Fig. 3 depicts $\mathrm{tr}[\mathbf{M}(T_s; n|n-1)]$ and $\mathrm{tr}[\mathbf{M}(T_s; n|n)]$ for different values of T_s illustrating how the gap between these two MSEs narrows as T_s decreases, eventually converging to $\mathrm{tr}[\mathbf{M}_c(nT_s)]$. Fig. 4 compares the KF, the SOI-KF and the $(\pi/2)$-KF for two representative sampling periods T_s. Note that for large T_s, $\mathrm{tr}[\mathbf{M}^{\mathrm{K}, \pi/2}(T_s; n|n)]$ and $\mathrm{tr}[\mathbf{M}(T_s; n|n)]$ are not equal (bottom); but as T_s decreases, these two quantities eventually coincide (top). It is is also worth noting that $\mathrm{tr}[\mathbf{M}_c(nT_s)] = \mathrm{tr}[\mathbf{M}_c^{\mathrm{K}, \pi/2}(nT_s)]$ is a valid upper bound for $\mathrm{tr}[\mathbf{M}(T_s; n|n)]$. We finally stress that the gap between the KF and the SOI-KF is small even for moderate values of T_s.

VI. SIMULATIONS

The SOI-KF can be applied in a number of situations. Consider for instance measuring, e.g., room temperature with a WSN. A common state propagation model is the zero-acceleration model [2, p. 262]

$$\dot{\mathbf{x}}_c(t) := \begin{pmatrix} \dot{\mathcal{T}}(t) \\ \ddot{\mathcal{T}}(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathcal{T}(t) \\ \dot{\mathcal{T}}(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t)$$
$$:= \mathbf{A}_c \mathbf{x}_c(t) + \mathbf{u}_c(t) \qquad (61)$$

where $\mathcal{T}(t)$ is the room's temperature, and $\dot{\mathcal{T}}(t)$ and $\ddot{\mathcal{T}}(t)$ denote first and second derivatives. Consistent with $u(t)$ having variance σ_u^2, the driving input's covariance matrix is $\mathbf{C}_{u_c}(t) = \sigma_u^2[0,1]^T[0,1]$.

Sensor S_k measures the temperature, but due to thermal inertia the observations are given by

$$y(t) = \mathcal{T}(t) - \beta_k \dot{\mathcal{T}}(t) + v(t) \qquad (62)$$

with β_k a sensor dependent constant and $\sigma_{v_c}^2(t) = \sigma_v^2$ denoting the measurement noise variance. For simplicity, we further assume that there are only two sensors that alternate in transmitting their observations.
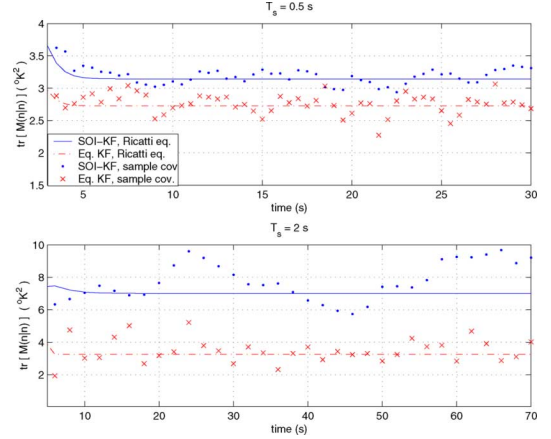

Fig. 5. SOI-KF compared with the $(\pi/2)$-KF. The filtered MSEs of the two filters are indistinguishable for small T_s, but as T_s becomes large, the $(\pi/2)$-KF is not a good predictor of the SOI-KF's performance ($\beta_1 = 0.1$, $\beta_2 = 0.2$, $\sigma_u^2 = 1$ and $\sigma_v^2 = 1$).
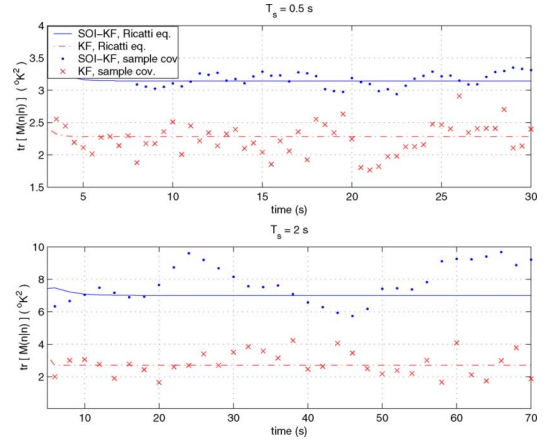

Fig. 6. SOI-KF compared with KF: even for moderate values of T_s, the performance penalty is small ($\beta_1 = 0.1$, $\beta_2 = 0.2$, $\sigma_u^2 = 1$ and $\sigma_v^2 = 1$).

Simulations for this problem are depicted in Figs. 5 and 6, where we can see that the theoretical MSE curves found as the solution of the corresponding Ricatti equations closely match the empirical results. In Fig. 5, we compare the SOI-KF with the $(\pi/2)$-KF for different sampling periods T_s. While for small T_s these two filters yield indeed indistinguishable performance (top), as T_s increases there is a noticeable gap between them (bottom). On the other hand, by inspecting the comparison between SOI-KF and KF in Fig. 6, we deduce that even for relatively large sampling intervals, the MSE penalty paid for quantizing to a single bit per sensor is small.

We finally consider the variation of the predicted and filtered MSEs ($\mathrm{tr}[\mathbf{M}(T_s; n|n)]$ and $\mathrm{tr}[\mathbf{M}(T_s; n|n-1)]$ respectively) with respect to the sampling period T_s. The steady state value as $n \to \infty$ of these quantities is shown in Fig. 7 for the SOI-KF, and corresponding KF and $(\pi/2)$-KF. We can see that as $T_s \to 0$, the gap between $\mathrm{tr}[\mathbf{M}(T_s; n|n)]$ and $\mathrm{tr}[\mathbf{M}(T_s; n|n-1)]$ narrows, and eventually both reach the continuous time MSE $\mathrm{tr}[\mathbf{M}_c(t)]$ obtained as the solution of (53) or (59). More interesting, let us note that in many applications we want T_s sufficiently small so that $\mathrm{tr}[\mathbf{M}^{\mathrm{K}}(T_s; n|n)]$ and $\mathrm{tr}[\mathbf{M}^{\mathrm{K}}(T_s; n|n-1)]$ are not very different. However, for these cases the SOI-KF
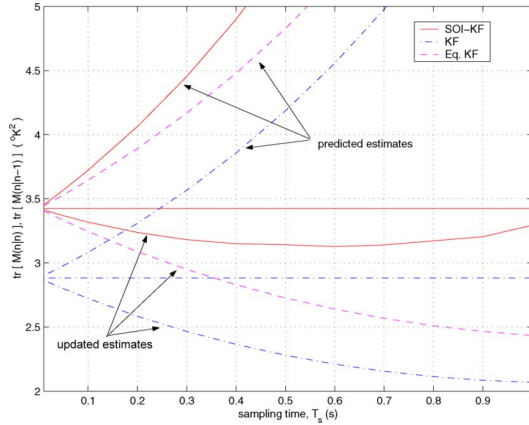
Fig. 7. Variation of estimates and predicted estimates for the SOI-KF, KF, and $(\pi/2)$-KF. For the given parameters we want $T_s < 0.5$ so that $\mathrm{tr}[\mathbf{M}^{\mathrm{K}}(T_s; n|n)]$ and $\mathrm{tr}[\mathbf{M}^{\mathrm{K}}(T_s; n|n-1)]$ are not very different, but for these T_s values the SOI-KF incurs a minimal variance penalty ($\beta_1 = 0.1$, $\beta_2 = 0.2$, $\sigma_u^2 = 1$ and $\sigma_v^2 = 1$).
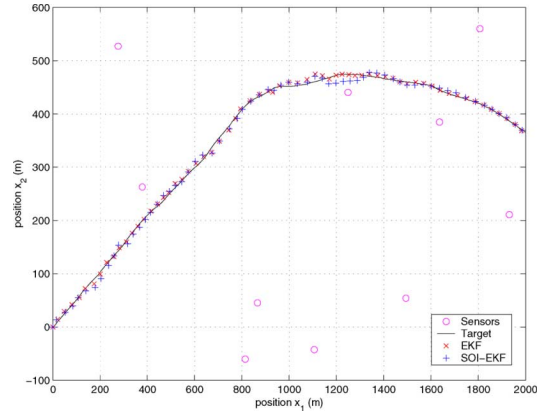

Fig. 8. Target tracking with EKF and SOI-EKF yield almost identical estimates. The scheduling algorithm works in cycles of duration T slots. At the beginning of the cycle, we schedule the sensor S_k closest to the estimate $\hat{\mathbf{x}}(n|n-1)$, next the second closest and so on until we complete the cycle ($T = 4$ slots, $T_s = 1$ s, $L = 2$ km, $K = 100$, $\alpha = 3.4$, $\sigma_u = 0.2$ m/s^2, $\sigma_v = 1$).
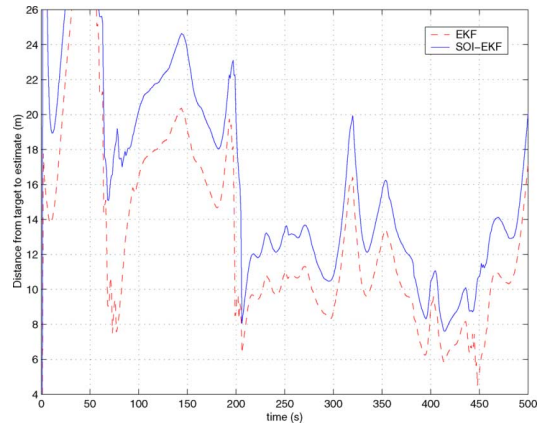

Fig. 9. Standard deviation of the estimates in Fig. 8 are in the order of 10–15 m for both filters.

incurs a negligible performance penalty since we can see that $\mathrm{tr}[\mathbf{M}^{\mathrm{K}}(T_s; n|n)] \approx \mathrm{tr}[\mathbf{M}(T_s; n|n)]$ for $T_s \leq 0.5$ s.

A. Target Tracking With SOI-EKF

Target tracking based on distance-only measurements is a typical problem in bandwidth-constrained distributed estimation with WSNs (see, e.g., [1] and [8]) for which a variation of the SOI-KF appears to be particularly attractive. Consider K sensors randomly and uniformly deployed in a square region of $2L \times 2L$ meters and suppose that sensor positions $\{\mathbf{x}^k\}_{k=1}^K$ are known.

The WSN is deployed to track the position $\mathbf{x}(n) := [x_1(n), x_2(n)]^T$ of a target, whose state model accounts for $\mathbf{x}(n)$ and the velocity $\mathbf{v}(n) := [v_1(n), v_2(n)]^T$, but not for the acceleration that is modelled as a random quantity. Under these assumptions, we obtain the state equation [12]

$$\begin{pmatrix} \mathbf{x}(n) \\ \mathbf{v}(n) \end{pmatrix} = \begin{pmatrix} 1 & 0 & T_s & 0 \\ 0 & 1 & 0 & T_s \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{x}(n-1) \\ \mathbf{v}(n-1) \end{pmatrix} + \begin{pmatrix} T_s^2/2 & 0 \\ 0 & T_s^2/2 \\ T_s & 0 \\ 0 & T_s \end{pmatrix} \mathbf{u}(n) \quad (63)$$

where T_s is the sampling period and the random vector $\mathbf{u}(n) \in \mathbf{R}^2$ is zero-mean white Gaussian, i.e., $p(\mathbf{u}(n)) = \mathcal{N}(\mathbf{u}(n); \mathbf{0}, \sigma_u^2 \mathbf{I})$. The sensors gather information about their distance to the target by measuring the received power of a pilot signal following the path-loss model:

$$y_k(n) = \alpha \log \|\mathbf{x}(n) - \mathbf{x}^k\| + v(n) \quad (64)$$

with $\alpha \geq 2$ a constant, $\|\mathbf{x}(n) - \mathbf{x}^k\|$ denoting the distance between the target and S_k, and $v(n)$ the observation noise with distribution $p(v(n)) = \mathcal{N}(v(n); 0, \sigma_v^2)$.

Mimicking an EKF approach, we linearize (64) in a neighborhood of $\hat{\mathbf{x}}(n|n-1)$ to obtain

$$y_k(n) - y_k^0(n) \simeq \mathbf{h}^T(n)\mathbf{x}(n) + v(n) \quad (65)$$

where $\mathbf{h}(n) := \alpha \hat{\mathbf{x}}(n|n-1)/\|\hat{\mathbf{x}}(n|n-1) - \mathbf{x}^k\|^2$ and $y_k^0(n)$ is an explicit function of α, $\hat{\mathbf{x}}(n|n-1)$ and $\mathbf{x}^k$.

The approximate model in (63)–(65) is of the form (9) and we can apply the SOI-KF outlined in Algorithms 2-A and 2-B to track the target's position $\mathbf{x}(n)$. This procedure amounts to the implementation of an extended SOI-(E)KF which is a low communication cost version of the EKF.

The results of simulating this setup are depicted in Figs. 8 and 9, where we see that the SOI-KF succeeds in tracking the target with distance error for the position estimates of less than 10 m. Similar conclusions can be obtained from Fig. 10 that depicts the error in the first coordinate $x_1(n) - \hat{x}_1(n|n)$ compared with the $\pm 3\sqrt{[\mathbf{M}(n|n)]_{11}}$ curves for the SOI-KF and the KF. As expected, the SOI-KF error matches the KF error with both of them within the corresponding $\pm 3\sqrt{[\mathbf{M}(n|n)]_{11}}$ curves. While this accuracy is just a result of the specific parameters of the experiment, the important point here is that the clairvoyant
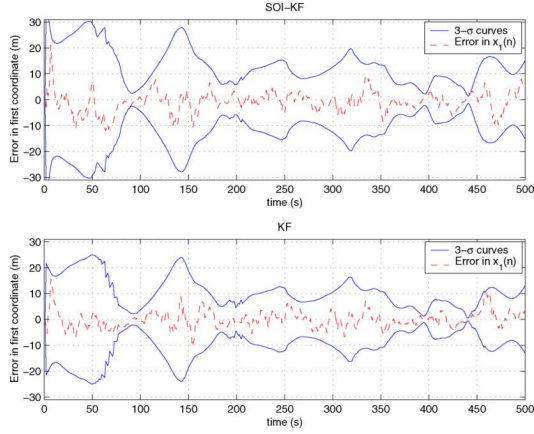
Fig. 10. SOI-KF error in the first component of the position estimate $x_1(n) - \hat{x}_1(n|n)$ (error in $x_1(n)$) is within the $\pm 3\sqrt{[\mathbf{M}(n|n)]_{11}}$ curves ($3 - \sigma$ curves).

EKF and the SOI-EKF yield almost identical performance even when the former relies on analog-amplitude observations and the SOI-EKF on the transmission of a single bit per sensor.

VII. CONCLUDING REMARKS

Relying on the SOI, we considered the problem of distributed state estimation in the context of wireless sensor networks. The binary SOI data render the problem nonlinear and lead to prohibitively complex MMSE state estimation. This motivated an approximation leading to the SOI-KF which constitutes an approximate MMSE estimator whose complexity and performance are very close to that of a KF even when the latter fuses the original (analog-amplitude) observations and the SOI-KF is based on the transmission of a single bit per observation.

Relating the discrete-time KF and SOI-KF with the underlying continuous-time physical process monitored by the WSN, we established that the MSE of the SOI-KF coincides with the MSE of a KF applied to an otherwise equivalent system model with $\pi/2$ larger measurement noise covariance matrix. This result was derived in the limit as the sampling period becomes arbitrarily small; practical simulations confirmed its validity even for moderate-size sampling intervals.

The SOI-KF was applied to a motivating application entailing temperature monitoring and to the canonical target tracking problem based on distance-only measurements. In both cases, we corroborated that at low communication cost the SOI-KF and the SOI-EKF yield estimates that are indistinguishable from the estimates of the clairvoyant KF and EKF for all practical purposes.

Future research directions include SOI-based estimators in different setups. As we pursued the SOI-EKF, one can envision similar combinations with the (SOI-)UKF and the (SOI-)PF in which we trade complexity for performance in highly non-linear state estimation problems. On the other hand, a multi-bit version of the SOI-KF can be devised in which the kth bit of a quantized observation is defined as the SOI relative to the estimator based on the previous $k-1$ bits. In both cases, the goal is to effect distributed state estimation with low-cost communications.

APPENDIX
PROOF OF (20)

To simplify notation, we will drop the time argument to write $\mathbf{h} = \mathbf{h}(n)$ and $\sigma_v^2 = \sigma_v^2(n)$. Due to the symmetry of the problem, it suffices to consider the case $b(n) = 1$. Start with the change of variables $\tilde{\mathbf{x}}(n) := \mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1)$, so that we can write (26) as

$$\hat{\mathbf{x}}(n|n) = \hat{\mathbf{x}}(n|n-1) + 2\int_{\mathbb{R}^p} \tilde{\mathbf{x}}(n)Q\left[-\frac{\mathbf{h}^T\tilde{\mathbf{x}}(n)}{\sigma_v}\right]$$
$$\times \frac{\exp\left[-\frac{1}{2}\tilde{\mathbf{x}}^T(n)\mathbf{M}^{-1}(n|n-1)\tilde{\mathbf{x}}(n)\right]}{(2\pi)^{p/2}\det^{1/2}[\mathbf{M}(n|n-1)]}d\tilde{\mathbf{x}}(n). \quad (66)$$

Introduce a second change of variables $\mathbf{z} := \mathbf{M}^{-1/2}(n|n-1)\tilde{\mathbf{x}}(n)$, and also let $\mathbf{g} := \mathbf{M}^{1/2}(n|n-1)\mathbf{h}$ to obtain

$$\hat{\mathbf{x}}(n|n) = \hat{\mathbf{x}}(n|n-1) + \frac{2\mathbf{M}^{1/2}(n|n-1)}{(2\pi)^{p/2}}$$
$$\times \int_{\mathbb{R}^p} \mathbf{z}\exp\left[-\frac{\mathbf{z}^T\mathbf{z}}{2}\right]Q\left[-\frac{\mathbf{g}^T\mathbf{z}}{\sigma_v}\right]d\mathbf{z} \quad (67)$$

where we recall that $d\mathbf{z} = \det[\mathbf{M}^{-1/2}(n|n-1)]d\tilde{\mathbf{x}}(n) = d\tilde{\mathbf{x}}(n)/\det[\mathbf{M}^{1/2}(n|n-1)]$. Define the integral $\mathbf{l} := [2/(2\pi)^{p/2}]\int_{\mathbb{R}^p} \tilde{\mathbf{x}}(n)\exp[-\mathbf{u}^T\mathbf{u}/2]Q[-\mathbf{g}^T\mathbf{u}/\sigma_v]d\mathbf{u} = [l_1, \ldots, l_p]^T$, that we can express componentwise as

$$l_k = \frac{2}{(2\pi)^{p/2}}\int_{\mathbb{R}^{p-1}}\exp\left[-\frac{\mathbf{z}_{-k}^T\mathbf{z}_{-k}}{2}\right]\int_{\mathbb{R}}z_k\exp\left[-\frac{z_k^2}{2}\right]$$
$$\times \int_{-[\mathbf{g}_{-k}^T\mathbf{u}_{-k}+g_ku_k]}^{\infty}\frac{\exp\left[-\frac{v^2}{2\sigma_v^2}\right]}{\sqrt{2\pi}\sigma_v}d\mathbf{z}_{-k}dz_kdv \quad (68)$$

where we used the definition $Q[-\mathbf{g}^T\mathbf{z}/\sigma_v] := \int_{-\mathbf{g}^T\mathbf{z}}^{\infty}[1/(\sqrt{2\pi}\sigma_v)]\exp[-v^2/2\sigma_v^2]dv$, introduced the notations $\mathbf{z}_{-k} := [z_1, \ldots, z_{k-1}, z_{k+1}, \ldots, z_p]^T$ and $\mathbf{g}_{-k} := [g_1, \ldots, g_{k-1}, g_{k+1}, \ldots, g_p]^T$, and separated the exponent as $\mathbf{z}^T\mathbf{z} = \mathbf{z}_{-k}^T\mathbf{z}_{-k} + z_k^2$.

We can now observe that $z_k\exp[-z_k^2/2] = -(\partial/\partial z_k)\exp[-z_k^2/2]$ and interchange the last two integrals in (68) to obtain

$$l_k = \frac{2\mathrm{sign}(g_k)}{(2\pi)^{p/2}}\int_{\mathbb{R}^{p-1}}\exp\left[-\frac{\mathbf{z}_{-k}^T\mathbf{z}_{-k}}{2}\right]\int_{\mathbb{R}}\frac{\exp\left[-\frac{v^2}{2\sigma_v^2}\right]}{\sqrt{2\pi}\sigma_v}$$
$$\times \int_{-[\mathbf{g}_{-k}^T\mathbf{u}_{-k}+v]/g_k}^{\infty}\frac{\partial}{\partial z_k}\exp\left[-z_k^2/2\right]d\mathbf{z}_{-k}dvdz_k$$
$$= \frac{2}{(2\pi)^{(p+1)/2}\sigma_v}\int_{\mathbb{R}^{p-1}}\exp\left[-\frac{\mathbf{z}_{-k}^T\mathbf{z}_{-k}}{2}\right]$$
$$\times \int_{\mathbb{R}}\exp\left[-\frac{v^2}{2\sigma_v^2}-\frac{(\mathbf{g}_{-k}^T\mathbf{u}_{-k}+v)^2}{2g_k^2}\right]d\mathbf{z}_{-k}dv \quad (69)$$

with the last equality following from the fundamental theorem of calculus. We can further rearrange terms in (69) and interchange the integrals to arrive at

$$
l_k = \frac{2\mathrm{sign}(g_k)}{(2\pi)^{(p+1)/2}\sigma_v} \int_{\mathbb{R}} \exp\left[-\frac{v^2}{2}\left(\frac{1}{\sigma_v^2}+\frac{1}{g_k^2}\right)\right]
$$
$$
\times \int_{\mathbb{R}^{p-1}} \exp\left[-\frac{1}{2}\left(\mathbf{z}_{-k}^T\left(\mathbf{I}+\frac{\mathbf{g}_{-k}\mathbf{g}_{-k}^T}{g_k^2}\right)\mathbf{z}_{-k}\right.\right.
$$
$$
\left.\left.-\frac{2\mathbf{g}_{-k}^T\mathbf{z}_{-k}v}{g_k^2}\right)\right] dv\, d\mathbf{z}_{-k}. \tag{70}
$$

Consider now the quadratic form in the exponent of the second integral, and let us summarize a number of properties about this form in the following lemma.

Lemma 1: If we define the matrix $\mathbf{C}^{-1} := \mathbf{I} + \mathbf{g}_{-k}\mathbf{g}_{-k}^T/g_k^2$, it holds that:

a) the inverse of $\mathbf{C}^{-1}$ is given by $\mathbf{C} = \mathbf{I} - \mathbf{g}_{-k}\mathbf{g}_{-k}^T/\mathbf{g}^T\mathbf{g}$;
b) the determinant of $\mathbf{C}^{-1}$ is $\det(\mathbf{C}^{-1}) = \mathbf{g}^T\mathbf{g}/g_k^2$;
c) the quadratic form in the exponent of the second integral in (70) can be written as

$$
\mathbf{z}_{-k}^T\mathbf{C}^{-1}\mathbf{z}_{-k} - \frac{2\mathbf{g}_{-k}^T\mathbf{z}_{-k}v}{g_k^2}
$$
$$
= (\mathbf{z}_{-k}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{z}_{-k}-\boldsymbol{\mu}) - \frac{v^2\mathbf{g}_{-k}^T\mathbf{g}_{-k}}{g_k^2\mathbf{g}^T\mathbf{g}} \tag{71}
$$

with $\boldsymbol{\mu} := (-v/g_k^2)\mathbf{C}^{-1}\mathbf{g}_{-k}$.

Proof: Statement a) follows from the matrix inversion lemma, and can also be proved by verifying that $\mathbf{C}\mathbf{C}^{-1} = \mathbf{I}$. To prove b), let $\mathbf{w}$ be an eigenvector of $\mathbf{C}^{-1}$; being an eigenvector of $\mathbf{C}^{-1}$, $\mathbf{w}$ must satisfy

$$
\mathbf{C}^{-1}\mathbf{w} = \mathbf{w} + \frac{\mathbf{g}_{-k}\mathbf{g}_{-k}^T\mathbf{w}}{g_k^2} = \lambda_w\mathbf{w} \tag{72}
$$

for some constant λ_w. Note that (72) is satisfied by $\mathbf{w}_1 = \mathbf{g}_{-k}$ with $\lambda_{w_1} = 1 + \mathbf{g}_{-k}^T\mathbf{g}_{-k}/g_k^2$, and by any $\mathbf{w}_j$ perpendicular to $\mathbf{g}_{-k}$ such that $\mathbf{g}_{-k}^T\mathbf{w}_j = 0$ with $\lambda_{w_j} = 1$. Since the determinant can be expressed as the product of the eigenvalues, we have

$$
\det(\mathbf{C}^{-1}) = \prod_{j=1}^{p-1}\lambda_{w_j} = \left(1+\frac{\mathbf{g}_{-k}^T\mathbf{g}_{-k}}{g_k^2}\right)\prod_{j=2}^{p-1}\lambda_{w_j}. \tag{73}
$$

However, the dimension of the subspace perpendicular to $\mathbf{g}_{-k}$ is $p-2$, and thus, $\prod_{j=2}^{p-1}\lambda_j = 1$. Statement b) is obtained by simply rearranging terms.

To prove c), expand the right-hand side and verify that the equality is indeed true. ∎

Using Lemma 1-c), we can rewrite (70) as

$$
l_k = \frac{2\mathrm{sign}(g_k)}{(2\pi)^{(p+1)/2}\sigma_v} \int_{\mathbb{R}} \exp\left[-\frac{v^2}{2}\left(\frac{1}{\sigma_v^2}+\frac{1}{g_k^2}-\frac{\mathbf{g}_{-k}^T\mathbf{g}_{-k}}{g_k^2\mathbf{g}^T\mathbf{g}}\right)\right] dv
$$
$$
\times \int_{\mathbb{R}^{p-1}} \exp\left[-\frac{1}{2}(\mathbf{z}_{-k}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{z}_{-k}-\boldsymbol{\mu})\right] d\mathbf{z}_{-k} \tag{74}
$$

where the two integrals are independent. The second integral is the integral of a $(p-1)$-dimensional Gaussian distribution over

$\mathbb{R}^{p-1}$ which regardless of $\boldsymbol{\mu}$ is equal to $(2\pi)^{(p-1)/2}\det^{1/2}(\mathbf{C})$; given that $\det(\mathbf{C})$ is given by the inverse of the expression in Lemma 1-b), we obtain

$$
\int_{\mathbb{R}^{p-1}} \exp\left[-\frac{1}{2}(\mathbf{z}_{-k}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{z}_{-k}-\boldsymbol{\mu})\right] d\mathbf{z}_{-k}
$$
$$
= (2\pi)^{\frac{p-1}{2}}\left(\frac{g_k^2}{\mathbf{g}^T\mathbf{g}}\right)^{1/2}. \tag{75}
$$

The first integral in (74) is the integral of a Gaussian bell over $\mathbb{R}$ and is thus given by $\sqrt{2\pi}$ times the standard deviation

$$
\int_{\mathbb{R}} \exp\left[-\frac{v^2}{2}\left(\frac{1}{\sigma_v^2}+\frac{1}{g_k^2}-\frac{\mathbf{g}_{-k}^T\mathbf{g}_{-k}}{g_k^2\mathbf{g}^T\mathbf{g}}\right)\right] dv
$$
$$
= \sqrt{2\pi}\left(\frac{\sigma_v^2\mathbf{g}^T\mathbf{g}}{\mathbf{g}^T\mathbf{g}+\sigma_v^2}\right)^{1/2}. \tag{76}
$$

Substituting (75) and (76) into (74), we obtain

$$
l_k = \frac{\sqrt{2/\pi}}{\sqrt{\mathbf{g}^T\mathbf{g}+\sigma_v^2}}g_k. \tag{77}
$$

Placing the components of $\mathbf{l}$ given by (77) into (67) yields the expression

$$
\hat{\mathbf{x}}(n|n) = \hat{\mathbf{x}}(n|n-1) + \frac{\sqrt{2/\pi}}{\sqrt{\mathbf{g}^T\mathbf{g}+\sigma_v^2}}\mathbf{M}^{1/2}(n|n-1)\mathbf{g}. \tag{78}
$$

Recalling that $\mathbf{g} := \mathbf{M}^{1/2}(n|n-1)\mathbf{h}$, (26) follows for $b(n) = 1$. For $b(n) = -1$, the opposite result follows from symmetry.

REFERENCES

[1] J. Aslam, Z. Butler, F. Constantin, V. Crespi, G. Cybenko, and D. Rus, "Tracking a moving object with a binary sensor network," in *Proc. 1st Int. Conf. Embedded Networked Sensor Systems*, Los Angeles, CA, 2003, pp. 150–161.

[2] Y. Bar-Shalom and X. Li, *Estimation and Tracking: Principles, Techniques, and Software*. Norwood, MA: Artech House, 1993.

[3] B. Beferull-Lozano, R. L. Konsbruck, and M. Vetterli, "Rate-distortion problem for physics based distributed sensing," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, Montreal, QC, Canada, May 2004, vol. 3, pp. 913–916.

[4] D. Blatt and A. Hero, "Distributed maximum likelihood estimation for sensor networks," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, Montreal, QC, Canada, May 2004, vol. 3, pp. 929–932.

[5] C. Chong, S. Mori, and K. Chang, *Distributed Multitarget Multisensor Tracking*. Norwood, MA: Artech House, 1990.

[6] R. Curry, W. Vandervelde, and J. Potter, "Nonlinear estimation with quantized measurements—PCM, predictive quantization, and data compression," *IEEE Trans. Inf. Theory*, vol. IT-16, no. 3, pp. 152–161, March 1970.

[7] P. Djuric, J. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. Bugallo, and J. Miguez, "Particle filtering," *IEEE Signal Process. Mag*, vol. 20, no. 9, pp. 19–38, Sep. 2003.

[8] P. Djuric, M. Vemula, and M. Bugallo, "Tracking with particle filtering in tertiary wireless sensor networks," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing*, Philadelphia, PA, Mar. 19–23, 2005, vol. 4, pp. 757–760.

[9] E. Ertin, R. Moses, and L. Potter, "Network parameter estimation with detection failures," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, Montreal, QC, Canada, May 2004, vol. 2, pp. 273–276.

[10] J. Gubner, "Distributed estimation and quantization," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1456–1459, Jul. 1993.

[11] V. Gupta, T. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and dynamic sensor coverage," *Automatica*, vol. 42, no. 2, pp. 251–260, Feb. 2006.

[12] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, and P.-J. Nordlund, "Particle filters for positioning, navigation, and tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 425–437, Feb. 2002.

[13] A. Jadbabaie, J. Lin, and A. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988–1001, Jun. 2003.

[14] S. Julier and J. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[15] S. M. Kay, *Fundamentals of Statistical Signal Processing—Estimation Theory*. Upper Saddle River, NJ: Prentice-Hall, 1993.

[16] J. Kotecha and P. Djuric, "Gaussian particle filtering," *IEEE Trans. Signal Process.*, vol. 51, no. 10, pp. 2602–2612, Oct. 2003.

[17] W. Lam and A. Reibman, "Quantizer design for decentralized systems with communication constraints," *IEEE Trans. Commun.*, vol. 41, no. 8, pp. 1602–1605, Aug. 1993.

[18] Z.-Q. Luo, "An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 735–744, April 2005.

[19] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proc. 1st ACM Int. Workshop on Wireless Sensor Networks and Applications*, Atlanta, GA, 2002, vol. 3, pp. 88–97.

[20] P. S. Maybeck, *Stochastic Models, Estimation and Control—Vol. 1*, 1st ed. New York: Academic, 1979.

[21] A. I. Mourikis and S. I. Roumeliotis, "Optimal sensor scheduling for resource-constrained localization of mobile robot formations," *IEEE Trans. Robotics*, vol. 22, no. 5, pp. 917–931, Oct. 2006.

[22] H. Papadopoulos, G. Wornell, and A. Oppenheim, "Sequential signal encoding from noisy measurements using quantizers with dynamic bias control," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 978–1002, Mar. 2001.

[23] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Process. Mag.*, vol. 19, no. 3, pp. 51–60, Mar. 2002.

[24] M. G. Rabbat and R. D. Nowak, "Decentralized source localization and tracking," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, Montreal, QC, Canada, May 2004, vol. 3, pp. 921–924.

[25] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks, part II: unknown pdf," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2784–2796, Jul. 2006.

[26] ——, "Bandwidth-constrained distributed estimation for wireless sensor networks, Part I: Gaussian case," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1131–1143, Mar. 2006.

[27] A. Ribeiro, G. B. Giannakis, and S. I. Roumeliotis, "SOI-KF: distributed Kalman filtering with low-cost communications using the sign of innovations," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, Toulouse, France, May 14–19, 2006, pp. IV-153–IV-156.

[28] A. H. Sayed, *Fundamentals of Adaptive Filtering*, 1st ed. New York: Wiley-Interscience, 2003.

[29] D. Williamson, "Finite wordlength design of digital Kalman filters for state estimation," *IEEE Trans. Autom. Control*, vol. AC-30, no. 10, pp. 930–939, Oct. 1985.

Georgios B. Giannakis (F'97) received the Diploma in electrical engineering from the National Technical University of Athens, Athens, Greece, 1981. From September 1982 to July 1986, he was with the University of Southern California (USC), where he received the M.Sc. degree in electrical engineering in 1983, the M.Sc. degree in mathematics in 1986, and the Ph.D. degree in electrical engineering in 1986.

After lecturing for one year at USC, he joined the University of Virginia, Charlottesville, in 1987, where he became a Professor of Electrical Engineering in 1997. Since 1999, he has been a Professor with the Department of Electrical and Computer Engineering at the University of Minnesota, Minneapolis, where he now holds an ADC Chair in Wireless Telecommunications. His general interests span the areas of communications and signal processing, estimation and detection theory, time-series analysis, and system identification—subjects on which he has published more than 250 journal papers, 400 conference papers, and two edited books. Current research focuses on diversity techniques for fading channels, complex-field and space-time coding, multicarrier, ultra-wide band wireless communication systems, cross-layer designs, and sensor networks.

Dr. Giannakis is the (co-) recipient of six paper awards from the IEEE Signal Processing (SP) and Communications Societies (1992, 1998, 2000, 2001, 2003, and 2004). He also received Technical Achievement Awards from the SP Society in 2000 and from EURASIP in 2005. He served as Editor in Chief for the IEEE SIGNAL PROCESSING LETTERS, as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, as Secretary of the SP Conference Board, as member of the SP Publications Board, as a member and Vice-Chair of the Statistical Signal and Array Processing Technical Committee, as Chair of the SP for Communications Technical Committee, and as a member of the IEEE Fellows Election Committee. He has also served as a member of the IEEE-SP Society's Board of Governors, the Editorial Board for the PROCEEDINGS OF THE IEEE and the steering committee of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

Alejandro Ribeiro (S'02) received the B.Sc. degree in electrical engineering from the Universidad de la Republica Oriental del Uruguay, Montevideo, Uruguay, in 1998. Since May 2003, he has been working towards the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, where he received the M.Sc. degree in electrical engineering in 2005.

From 1998 to 2003, he was a member of the Technical Staff at Bellsouth Montevideo. His research interests lie in the areas of communication theory, signal processing, and networking. His current research focuses on wireless cooperative communications, random access, wireless ad hoc and sensor networks, and distributed signal processing.

Mr. Ribeiro is a Fulbright Scholar.

Stergios I. Roumeliotis (M'00) received the Diploma in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1995, and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, in 1999 and 2000, respectively.

From 2000 to 2002, he was a Postdoctoral Fellow at the California Institute of Technology, Pasadena. Since 2002, he has been an Assistant Professor in the Department of Computer Science and Engineering at the University of Minnesota, Minneapolis. His research interests include inertial navigation of aerial and ground autonomous vehicles, fault detection and identification, and sensor networks. Recently, his research has focused on distributed estimation under communication and processing constraints and active sensing for reconfigurable networks of mobile sensors.

Dr. Roumeliotis is the recipient of the McKnight Land-Grant Professorship Award and the NASA Tech Briefs Award.