MAP Moving Horizon state estimation with binary measurements

(Article begins on next page)

03 May 2024

MAP moving horizon state estimation with binary measurements

Giorgio Battistelli   Luigi Chisci   Nicola Forti   Stefano Gherardini

Abstract— The paper addresses state estimation for discrete-time systems with binary (threshold) measurements by following a *Maximum A posteriori Probability* (MAP) approach and exploiting a *Moving Horizon* (MH) approximation of the MAP cost-function. It is shown that, for a linear system and noise distributions with log-concave probability density function, the proposed MH-MAP state estimator involves the solution, at each sampling interval, of a convex optimization problem. Application of the MH-MAP estimator to dynamic estimation of a diffusion field given pointwise-in-time-and-space binary measurements of the field is also illustrated and, finally, simulation results relative to this application are shown to demonstrate the effectiveness of the proposed approach.

I. INTRODUCTION

Binary sensors, whose output just indicates whether the noisy measurement of the sensed variable (analog measurement) exceeds or not a given threshold, are frequently employed in monitoring and control applications [1]-[14]. The idea is that by a multitude of low-cost and low-resolution sensing devices it is possible to achieve the same estimation accuracy that a few (possibly a single one) expensive high-resolution sensors could provide, with significant practical benefits in terms of ease of sensor deployment and minimization of communication requirements. The fact that a binary (threshold) measurement just conveys a minimal amount (i.e. a single bit) of information, while implying communication bandwidth savings and consequently greater energy efficiency, makes of paramount importance to fully exploit the little available information by means of smart estimation algorithms. In this respect, some work has recently addressed system identification [1]-[2], parameter [3]-[6] or state estimation [7]-[14] with binary measurements by following either a deterministic [1]-[2], [7]-[9] or a probabilistic [3]-[6], [10]-[14] approach.

In a deterministic context, the available information is essentially concentrated at the sampling instants in which some binary measurement signal has switched value [7], [8]. As shown in [8], some additional information can be exploited in the other (non switching) sampling instants by penalizing values of the estimated quantity such that the corresponding predicted measurement is on the opposite side, with respect to a binary sensor reading, far away from the threshold. Nevertheless, it is clear that there is no or very little information available for estimation purposes whenever no or very few binary sensor switchings occur. Hence, a

The authors are with Dipartimento di Ingegneria dell'Informazione (DINFO), Università degli Studi di Firenze, via di Santa Marta 3, {giorgio.battistelli,luigi.chisci,nicola.forti,stefano.gherardini}@unifi.it. S. Gherardini is also with Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Firenze.

possible way to achieve high estimation accuracy is to have many binary sensors measuring the same variable with different thresholds as this would clearly increase the number of switchings, actually emulating, when the number of sensors tends to infinity, the availability of a single continuous-valued (analog) measurement.

Conversely, following a probabilistic approach, binary sensor readings could be exploited to infer information about the probability distribution of the variable of interest. To clarify this point, let us assume that a very large number of binary sensors of the same type (i.e. measuring the same variable with the same threshold) be available and the distribution of their measurement noise (e.g. Gaussian with zero mean and given standard deviation) be known. Then, thanks to the numerosity of measurements, the relative frequency of 1 (or 0) values occurring in the sensor readings could be considered as a reasonable estimate of the probability that the sensed variable is above (or below) the threshold and this, in turn, exploiting the knowledge of the measurement noise distribution allows to extract information about the location of the value of the sensed variable with respect to the threshold. If, for example, it is found that the binary measurement is equal to 1 for 70% of the sensors and Gaussian measurement noise is hypothesized, it turns out that the expected measurement of the sensed variable is above the threshold of an amount equal to 0.525 times the standard deviation of the measurement noise. Notice that if the sensors are noiseless, they all provide either 0 or 1 output and, paradoxically, in this case minimal information, i.e. that the sensed variable takes values in a semi-infinite interval (either below or above the threshold), is extracted from the set of binary measurements. The above arguments suggest that, adopting a probabilistic approach to estimation using binary measurements, the presence of measurement noise can be a helpful source of information. In other words, it can be said that noise-aided procedures can be devised for estimation with binary measurements by exploiting the fact that the measurement noise randomly shifts the analog measurement thus making possible to infer statistical information on the sensed variable.

Relying on the above stated *noise-aided* paradigm, this paper presents a novel approach to recursive estimation of the state of a discrete-time dynamical system given binary measurements. The proposed approach is based on a *moving-horizon* (MH) approximation of the *Maximum A-posteriori Probability* (MAP) estimation and extends previous work [3]-[4] concerning parameter estimation to recursive state estimation. A further contribution of the paper is to show that for a linear system the optimization problem arising from

the MH-MAP formulation turns out to be convex and, hence, practically feasible for real-time implementation.

The rest of the paper is organized as follows. Section II introduces the MAP problem formulation of state estimation with binary measurements. Section III presents a MH approximation of MAP estimation, referred to as MH-MAP algorithm, and analyzes the properties of the resulting optimization problem. Section IV discusses a possible application of the proposed approach to the dynamic estimation of a diffusion field from binary pointwise-in-space-and-time field measurements. Section V presents simulation results relative to the dynamic field estimation case-study. Finally, section VI concludes the paper with perspectives for future work.

II. MAXIMUM A POSTERIORI STATE ESTIMATION WITH BINARY SENSORS

The following notation will be used throughout the paper: $col(\cdot)$ denotes the matrix obtained by stacking its arguments one on top of the other; $diag(m^1,\ldots,m^q)$ denotes the diagonal matrix with diagonal entries $m^1,\ldots,m^q$; 0_n, 1_n indicate the $n-$dimensional vectors, respectively, with all zero and unit entries.

Let us consider the problem of recursively estimating the state of the discrete-time nonlinear dynamical system

$$x_{t+1} = f(x_t, u_t) + w_t \qquad (1)$$
$$z_t^i = h^i(x_t) + v_t^i, \quad i = 1,\ldots,l \qquad (2)$$

from a set of measurements provided by binary sensors

$$y_t^i = g^i(z_t^i) = \begin{cases} 1, & \text{if } z_t^i \geq \tau^i \\ 0, & \text{if } z_t^i < \tau^i \end{cases} \qquad (3)$$

where $x_t \in \mathbb{R}^n$ is the state to be estimated, $u_t \in \mathbb{R}^m$ is a known input, and τ^i is the threshold of the $i-$th binary sensor. For the sake of simplicity, we define $z_t = col\left(z_t^i\right)_{i=1}^l \in \mathbb{R}^l$ and $y_t = col\left(y_t^i\right)_{i=1}^l \in \mathbb{R}^l$. The vector $w_t \in \mathbb{R}^n$ is an additive disturbance affecting the system dynamics which accounts for uncertainties in the mathematical model, while $v_t = col\left(v_t^i\right)_{i=1}^l \in \mathbb{R}^l$ is the measurement noise vector.

Let $\mathcal{N}(\mu,\Sigma)$ denote as usual the normal distribution with mean μ and variance Σ. The statistical behaviour of the system is characterized by

$$x_0 \sim \mathcal{N}(\bar{x}_0, P^{-1}), \ w_t \sim \mathcal{N}(0, Q^{-1}), \ v_t \sim \mathcal{N}(0,R) \quad (4)$$

where: $R = diag(r^1,\ldots,r^p)$; $\mathbb{E}[w_j w_k'] = 0$ and $\mathbb{E}[v_j v_k'] = 0$ if $j \neq k$; $\mathbb{E}[w_j v_k'] = 0$, $\mathbb{E}[w_j x_0'] = 0$, $\mathbb{E}[v_j x_0'] = 0$ for any j,k. Notice from (2)-(3) that sensor i produces a binary measurements $y_t^i \in \{0,1\}$ depending on whether the noisy system output z_t^i is below or above the threshold τ^i.

According to the available probabilistic description (4), the problem of estimating the state of system (1) under the binary measurement model (2)-(3) is formulated hereafter in a Bayesian framework by resorting to a maximum a posteriori probability (MAP) criterion. In the remainder of this section, as a preliminary step, the *full-information* MAP state estimation problem is formulated.

To this end, notice that each binary measurement y_t^i provides intrinsically relevant information on the state x_t which can be taken into account by means of the a posteriori probabilities $p(y_t^i|x_t)$. In particular, the binary measurement y_t^i is a Bernoulli random variable such that, for any binary sensor i and any time instant t, the a posteriori probability $p(y_t^i|x_t)$ is given by

$$p(y_t^i|x_t) = p(y_t^i = 1|x_t)^{y_t^i} \ p(y_t^i = 0|x_t)^{1-y_t^i} \qquad (5)$$

where

$$p(y_t^i = 1|x_t) = F^i(\tau^i - h^i(x_t)) \qquad (6)$$

and $p(y_t^i = 0|x_t) = 1 - p(y_t^i = 1|x_t) \triangleq \Phi^i(\tau^i - h^i(x_t))$. The function $F^i(\tau^i - h^i(x_t))$ is the complementary cumulative distribution function (CDF) of the random variable $\tau^i - h^i(x_t)$. Since $v_t^i \sim \mathcal{N}(0,r^i)$, the conditional probability $p(y_t^i = 1|x_t) = F^i(\tau^i - h^i(x_t))$ can be written in terms of the Q-function as follows

$$F^i(\tau^i - h^i(x_t)) = \frac{1}{\sqrt{2\pi r^i}} \int_{\tau^i - h^i(x_t)}^{\infty} e^{-\frac{u^2}{2r^i}} du = Q\left(\frac{\tau^i - h^i(x_t)}{\sqrt{r^i}}\right). \qquad (7)$$

Let us now denote by $Y_t = col(y_0,\ldots,y_t)$ the vector of all binary measurements collected up to time t and by $X_t \triangleq col(x_0,\ldots,x_t)$ the vector of the state trajectory. Further, let us denote by $\hat{X}_{t|t} \triangleq col(\hat{x}_{0|t},\ldots,\hat{x}_{t|t})$ the estimates of X_t to be made at any stage t. Then, at each time instant t, given the a posteriori probability $p(X_N|Y_N)$, the estimate of the state trajectory can be obtained by solving the following MAP estimation problem:

$$\hat{X}_{t|t} = \arg\max_{X_t} p(X_t|Y_t) = \arg\min_{X_t} -\ln p(X_t|Y_t). \qquad (8)$$

From the Bayes rule

$$p(X_t|Y_t) \propto p(Y_t|X_t) \ p(X_t), \qquad (9)$$

where

$$p(X_t) = \prod_{k=0}^{t-1} p(x_{t-k}|x_{t-k-1},\ldots,x_0) \ p(x_0)$$
$$= \prod_{k=0}^{t-1} p(x_{t-k}|x_{t-k-1}) \ p(x_0). \qquad (10)$$

Notice that in the latter equation we have considered the Markov property for the dynamical system state. As x_0 and w_t are normally distributed vectors, we have

$$p(x_0) \propto e^{-\frac{1}{2}\|x_0 - \bar{x}_0\|_P^2} \qquad (11)$$

$$p(x_k|x_{k-1}) \propto e^{-\frac{1}{2}\|x_{k+1} - f(x_k, u_k)\|_Q^2}. \qquad (12)$$

Moreover, the likelihood function $p(Y_t|X_t)$ of the binary measurement vector Y_t can be written as

$$p(Y_t|X_t) = \prod_{k=0}^{t} p(y_k|x_k) = \prod_{k=0}^{t} \prod_{i=1}^{l} p(y_k^i|x_k)$$
$$= \prod_{k=0}^{t} \prod_{i=1}^{l} F^i(\tau^i - h^i(x_k))^{y_k^i} \ \Phi^i(\tau^i - h^i(x_k))^{1-y_k^i} \qquad (13)$$

where in the latter equality we have exploited the statistical independence of the binary sensors. Accordingly, the log-likelihood is

$$
\begin{aligned}
\ln p(Y_t | X_t) = \sum_{k=0}^{t} \sum_{i=1}^{l} & \left[y_k^i \ln F^i(\tau^i - h^i(x_k)) \right. \\
& \left. + (1 - y_k^i) \ln \Phi^i(\tau^i - h^i(x_k)) \right],
\end{aligned}
\tag{14}
$$

and the cost function $-\ln p(X_t | Y_t) = -\ln p(Y_t | X_t) - \ln p(X_t)$ to be minimized in the MAP estimation problem (8) turns out to be, up to additive constant terms,

$$
\begin{aligned}
J_t(X_t) = & \|x_0 - \bar{x}_0\|_P^2 + \sum_{k=0}^{t} \|x_{k+1} - f(x_k, u_k)\|_Q^2 \\
& - \sum_{k=0}^{t} \sum_{i=1}^{l} \left[y_k^i \ln F^i(\tau^i - h^i(x_k)) + (1 - y_k^i) \ln \Phi^i(\tau^i - h^i(x_k)) \right].
\end{aligned}
\tag{15}
$$

Unfortunately, a closed-form expression for the global minimum of (15) does not exist and, hence, the optimal MAP estimate $\hat{X}_{t|t}$ has to be determined by resorting to some numerical optimization routine. With this respect, the main drawback is that the number of optimization variables grows linearly with time, since the vector X_t has size $(t+1)n$. As a consequence, as t grows the solution of the full information MAP state estimation problem (8) becomes eventually unfeasible, and some approximation has to be introduced.

III. MOVING-HORIZON APPROXIMATION

In this section, an approximate solution to the MAP state estimation problem is proposed by resorting to the MHE approach [15]-[21]. Accordingly, by defining a sliding window $\mathfrak{W}_t = \{t-N, t-N+1, \ldots, t\}$, the goal is to find an estimate of the partial state trajectory $X_{t-N:t} \triangleq \mathrm{col}(x_{t-N}, \ldots, x_t)$ by using the information available in $\mathfrak{W}_t$. Then, in place of the full information cost $J_t(X_t)$, at each time instant t the minimization of the following *moving-horizon cost* is addressed:

$$
\begin{aligned}
J_t^{\mathrm{MH}}(X_{t-N:t}) = & \Gamma_{t-N}(x_{t-N}) + \sum_{k=t-N}^{t} \|x_{k+1} - f(x_k, u_k)\|_Q^2 \\
& - \sum_{k=t-N}^{t} \sum_{i=1}^{l} \left[y_k^i \ln F^i(\tau^i - h^i(x_k)) + (1 - y_k^i) \ln \Phi^i(\tau^i - h^i(x_k)) \right]
\end{aligned}
\tag{16}
$$

where the non-negative initial penalty function $\Gamma_{t-N}(x_{t-N})$, known in the MHE literature as *arrival cost* [16], [19], is introduced so as to summarize the past data $y_0, \ldots, y_{t-N-1}$ not explicitly accounted for in the objective function.

As a matter of fact, the form of the arrival cost plays an important role in the behavior and performance of the overall estimation scheme. While in principle $\Gamma_{t-N}(x_{t-N})$ could be chosen so that minimization of (16) yields the same estimate that would be obtained by minimizing (15), an algebraic expression for such a true arrival cost seldom exists, even when the sensors provide continuous (non-binary) measurements [16]. Hence, some approximation must be used. With this respect, a common choice [18], [19], also followed in

the present work, consists of assigning to the arrival cost a fixed structure penalizing the distance of the state x_{t-N} at the beginning of the sliding window from some prediction $\bar{x}_{t-N}$ computed at the previous time instant, thus making the estimation scheme recursive. A natural choice is then a quadratic arrival cost of the form

$$
\Gamma_{t-N}(x_{t-N}) = \|x_{t-N} - \bar{x}_{t-N}\|_\Psi^2,
\tag{17}
$$

which, from the Bayesian point of view, corresponds to approximating the PDF of the state x_{t-N} conditioned to all the measurements collected up to time $t-1$ with a Gaussian having mean $\bar{x}_{t-N}$ and covariance Ψ^{-1}. As for the choice of the weight matrix Ψ, in the case of continuous measurements it has been shown that stability of the estimation error dynamics can be ensured provided that Ψ is not too large (so as to avoid an overconfidence on the available estimates) [18], [19]. Recently, similar results [8] have been proven to hold also in the case of binary sensors in a deterministic context. In practice, Ψ can be seen as a design parameter which has to be tuned by pursuing a suitable tradeoff between such stability considerations and the necessity of not neglecting the already available information (since in the limit for Ψ going to zero the approach becomes a finite memory one).

Summing up, at any stage $t = N, N+1, \ldots$, the following problem has to be solved.

Problem E_t: Given the prediction $\bar{x}_{t-N}$, the input sequence $\{u_{t-N}, \ldots, u_{t-1}\}$, the measurement sequences $\{y_{t-N}^i, \ldots, y_t^i, i = 1, \ldots, l\}$, find the optimal estimates $\hat{x}_{t-N|t}, \ldots, \hat{x}_{t|t}$ that minimize the cost function (16) with arrival cost (17).

Concerning the propagation of the estimation procedure from Problem E_{t-1} to Problem E_t, the prediction $\bar{x}_{t-N}$ is set equal to the value of the estimate of x_{t-N} made at time instant $t-1$, i.e., $\bar{x}_{t-N} = \hat{x}_{t-N|t-1}$. Clearly, the recursion is initialized with the a priori expected value $\bar{x}_0$ of the initial state vector.

In general, solving Problem E_t entails the solution of a non-trivial optimization problem. However, when both equations (1) and (2) are linear, the resulting optimization problem turns out to be convex so that standard optimization routines can be used in order to find the global minimum. To see this, let us consider the following assumption.

A1 The functions $f(\cdot)$ and $h^i(\cdot)$, $i = 1, \ldots, l$, are linear, i.e., $f(x_t, u_t) = Ax_t + Bu_t$ and $h^i(x_t) = C^i x_t$, $i = 1, \ldots, l$, where A, B, C^i are constant matrices of suitable dimensions.

Then, the following result, whose proof is omitted due to lack of space, holds.

Proposition 1: If assumption A1 holds, the CDF $\Phi^i(\tau^i - C^i x_t)$ and its complementary function $F^i(\tau^i - C^i x_t)$ are log-concave. Hence, the cost function (16) with arrival cost (17) is convex.

$\square$

Remark 1: Under assumption A1, the convexity of the cost function (16) is guaranteed also in the more general case in which the statistical behaviour of the random variables x_0, w_t, v_t is described by logarithmically concave distribution functions. Indeed, if a PDF is log-concave, also its cumulative distribution function is log-concave; hence the contribution related to the binary measurements in (16) turns out to be convex.

In the next section we will focus on the case of a discrete-time linear system, in particular considering a diffusion process governed by a *partial differential equation* (PDE) and spatially discretized by means of the *finite element method* (FEM).

IV. DYNAMIC FIELD ESTIMATION

In this section, we consider the problem of reconstructing a two-dimensional diffusion field, sampled with a network of binary sensors arbitrarily deployed over the spatial domain of interest Ω. The diffusion process is governed by the following parabolic PDE:

$$\frac{\partial c}{\partial t} - \lambda \nabla^2 c = 0 \quad \text{in } \Omega \qquad (18)$$

which models various physical phenomena such as the spread of a pollutant in a fluid. In this case, $c(\xi,\eta,t)$ represents the space-time dependent substance concentration, λ denotes the constant diffusivity of the medium, and $\nabla^2 = \partial^2/\partial\xi^2 + \partial^2/\partial\eta^2$ is the Laplace operator, $(\xi,\eta) \in \Omega$ being the 2D spatial variables. Furthermore, let us assume mixed boundary conditions, i.e. a non-homogeneous Dirichlet condition

$$c = \psi \quad \text{on } \partial\Omega_D, \qquad (19)$$

which specifies a constant-in-time value of concentration on the boundary $\partial\Omega_D$, and a homogeneous Neumann condition on $\partial\Omega_N = \partial\Omega \setminus \partial\Omega_D$, assumed impermeable to the contaminant, so that

$$\partial c/\partial\upsilon = 0 \quad \text{on } \partial\Omega_N, \qquad (20)$$

where υ is the outward pointing unit normal vector of $\partial\Omega_N$.

The objective is to estimate the values of the dynamic field of interest $c(\xi,\eta,t)$, given the binary measurements (3). The PDE system (18)-(20) is simulated with a mesh of finite elements over Ω via the Finite Element (FE) approximation described in [22]-[23]. Specifically, the domain Ω is subdivided into a suitable set of non overlapping regions, or elements, and a suitable set of basis functions $\phi_j(\xi,\eta)$, $j = 1, \ldots, n_\phi$ is defined on such elements. The choices of the basis functions ϕ_j and of the elements are key points of the FE method. In the specific case under investigation, the elements are triangle in 2D and define a FE mesh with vertices $(\xi_j, \eta_j) \in \Omega, j = 1, \ldots, n_\phi$. Then each basis function ϕ_j is a piece-wise affine function which vanishes outside the FEs around (ξ_j, η_j) and such that $\phi_j(\xi_j, \eta_i) = \delta_{ij}$, δ_{ij} denoting the Kronecker delta. In order to account for the mixed boundary conditions, the basis functions are supposed to be ordered so that the first n correspond to vertices of the mesh which lie either in the interior of Ω or on $\partial\Omega_N$, while the last $n_\phi - n$ correspond to the vertices lying on $\partial\Omega_D$.

Accordingly, the unknown function $c(\xi,\eta,t)$ is approximated as

$$c(\xi,\eta,t) \approx \sum_{j=1}^{n} \phi_j(\xi,\eta) c_j(t) + \sum_{j=n+1}^{n_\phi} \phi_j(\xi,\eta) \psi_j \qquad (21)$$

where $c_j(t)$ is the unknown expansion coefficient of the function $c(\xi,\eta,t)$ relative to time t and basis function $\phi_j(\xi,\eta)$, and ψ_j is the known expansion coefficient of the function $\psi(\xi,\eta)$ relative to the basis function $\phi_j(\xi,\eta)$. Notice that the second summation in (21) is needed so as to impose the non-homogeneous Dirichlet condition (19) on the boundary $\partial\Omega_D$.

The PDE (18) can be recast into the following integral form:

$$\int_\Omega \frac{\partial c}{\partial t} \varphi \, d\xi d\eta - \lambda \int_\Omega \nabla^2 c \, \varphi \, d\xi d\eta = 0 \qquad (22)$$

where $\varphi(\xi,\eta)$ is a generic space-dependent weight function. By applying Green's identity, one obtains:

$$\int_\Omega \frac{\partial c}{\partial t} \varphi \, d\xi d\eta + \lambda \int_\Omega \nabla^T c \, \nabla\varphi \, d\xi d\eta - \lambda \int_{\partial\Omega} \frac{\partial c}{\partial \upsilon} \varphi \, d\xi d\eta = 0. \qquad (23)$$

By choosing the test function φ equal to the selected basis functions and exploiting the approximation (21), the Galerkin weighted residual method is applied and the following equation is obtained

$$\sum_{i=1}^{n} \int_\Omega \phi_i \phi_j \, d\xi d\eta \, \dot{c}_i(t) + \lambda \sum_{i=1}^{n} \int_\Omega \nabla^T \phi_i \, \nabla\phi_j \, d\xi d\eta \, c_i(t)$$
$$+ \lambda \sum_{i=n+1}^{n_\phi} \int_\Omega \nabla^T \phi_i \, \nabla\phi_j \, d\xi d\eta \, \psi_i = 0 \qquad (24)$$

for $j = 1, \ldots, n$. Notice that in the latter equation the boundary integral of equation (23) is omitted since it is equal to 0 thanks to the homogeneous Neumann condition (20) on $\partial\Omega_N$ and to the fact that, by construction, the basis functions ϕ_j, $j = 1, \ldots, n$ vanish on $\partial\Omega_D$. The interested reader is referred to [24] for further details on the FEM theory, and in particular on how to convert the case of inhomogeneous boundary conditions to the homogeneous one.

By defining the state vector $x = \text{col}(c_1, \ldots, c_n)$ and the vector of boundary conditions $\gamma = \text{col}(\psi_{n+1}, \ldots, \psi_{n_\phi})$, equation (24) can be written in the more compact form

$$M\dot{x}(t) + Sx(t) + S_D\gamma = 0$$

where S is the so-called stiffness matrix representing diffusion, M is the mass matrix, and S_D captures the physical interconnections among the vertices affected by boundary condition (19) and the remaining nodes of the mesh.

By applying for example the implicit Euler method, the latter equation can be discretized in time, thus obtaining the linear discrete-time model

$$x_{t+1} = Ax_t + Bu + w_t \qquad (25)$$

where

$$A = \left[I + \delta t \, M^{-1}S\right]^{-1}$$
$$B = \left[I + \delta t \, M^{-1}S\right]^{-1} M^{-1}\delta t$$
$$u = -S_D \, \gamma$$

δt is the time integration interval, and w_t is the process disturbance taking into account also the space-time discretization errors.

Notice that the linear system (25) has dimension n equal to the number of vertices of the mesh not lying on $\partial\Omega_D$. The linear system (25) is assumed to be monitored by a network of l threshold sensors. Each sensor, before binary quantization is applied, directly measures the pointwise-in-time-and-space concentration of the contaminant in a point of the spatial domain Ω. By exploiting (21), such a concentration can be written as a linear combination of the concentrations on the grid points, so that the resulting output function takes the form

$$z_t^i = C^i x_t + v_t^i, \quad i = 1, \ldots, l \tag{26}$$

and assumption A1 is fulfilled.

V. NUMERICAL RESULTS

In this section, we present the simulation results of the proposed approach applied to the problem of state estimation of spatially distributed processes, discussed in the previous section. We consider the simulated system (25)-(26) with 1695 triangular elements, 915 vertices, $\lambda = 0.01$ $[m^2/s]$, fixed integration step length $\delta t = 1$ $[s]$, $\gamma = 30$ $[g/m^2]$, and initial condition of the field vector $x_0 = 0_n$ $[g/m^2]$. The field of interest is defined over a bounded 2D spatial domain Ω which covers an area of 7.44 $[m^2]$ (see Fig. 1), with boundary condition (19) on the bottom edge and no-flux condition (20) on the remaining portions of $\partial\Omega$. Compared to the
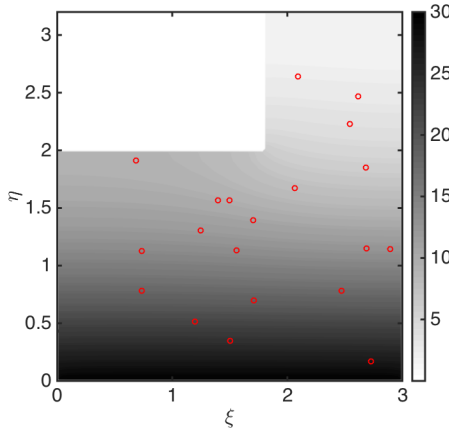

Fig. 1. Concentration field at time $t = 100$ $[s]$ monitored by a random network of 20 binary sensors (red ∘).

ground truth simulator, the proposed MH-MAP estimator implements a coarser mesh (see Fig. 2) of $n_\phi = 97$ vertices ($n = 89$), and runs at a slower sample rate (0.1 $[Hz]$), so that model uncertainty is taken into account. The initial condition of the estimated dynamic field is set to $\bar{x}_0 = 5\ 1_n$ $[g/m^2]$, the moving window has size $N = 5$, and the weight matrices in (4) are chosen as $\Psi = 10^3\ I_n$ and $Q = 10^2\ I_n$.

The *true* concentrations from (25) are first corrupted with a Gaussian noise with variance r^i, then binary observations are obtained by applying a different threshold τ^i for each sensor
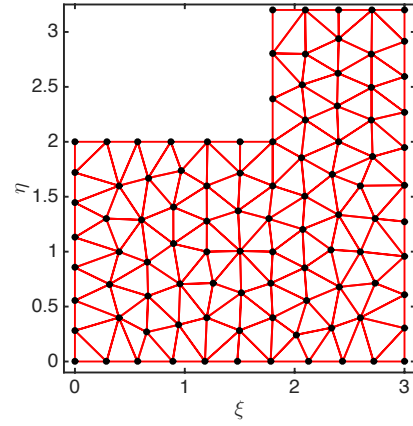

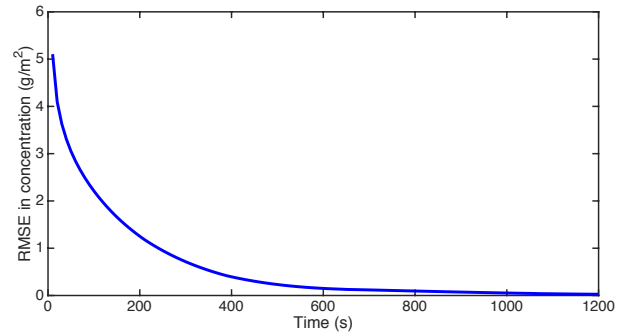Fig. 2. Mesh used by the MH-MAP estimator (152 elements, 97 nodes).

Fig. 3. RMSE in concentration of the MH-MAP state estimator as a function of time, for a random network of 5 threshold sensors.

i of the network. Note that, in order to receive informative binary measurements, τ^i, $i = 1, \ldots, l$ are generated as uniformly distributed random numbers in the interval $(0.05, 29.95)$, being $(0, 30)$ the range of nominal concentration values throughout each experiment. The duration of each simulation experiment is fixed to 1200 $[s]$ (120 samples).

Fig. 3 shows the performance of the novel MH-MAP state estimator implemented in MATLAB, in terms of Root Mean Square Error (RMSE) of the estimated concentration field, i.e.:

$$\text{RMSE}(t) = \left(\sum_{j=1}^{\alpha} \frac{\|e_{t,j}\|^2}{\alpha} \right)^{\frac{1}{2}}, \tag{27}$$

where $\|e_{t,j}\|$ is the norm of the estimation error at time t in the j−th simulation run, averaged over 304 sampling points (evenly spread within Ω) and $\alpha = 100$ independent Monte Carlo realizations. The estimation error is computed at time t on the basis of the estimate $\hat{x}_{t-N|t}$. It can be observed that the proposed estimator successfully estimates the dynamic field, even when observed by a network of $l = 5$ randomly deployed binary sensors, with $r^i = 0.25$ $[g/m^2]$ $\forall i = 1, \ldots, l$.

The effect of measurement noise on the mean value of the RMSE can be seen in Fig. 4, in which it becomes apparent how for certain values of r^i, including an observation noise with higher variance, can actually improve the quality of the overall estimates. The results in Fig. 4 numerically demon-
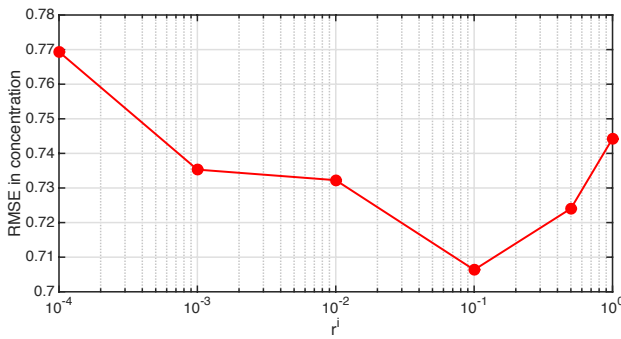
Fig. 4. RMSE in concentration as a function of the measurement noise variance, for a fixed constellation of 20 binary sensors. It is shown here that operating in a noisy environment turns out to be beneficial, for certain values of r^i, to the state estimation problem.
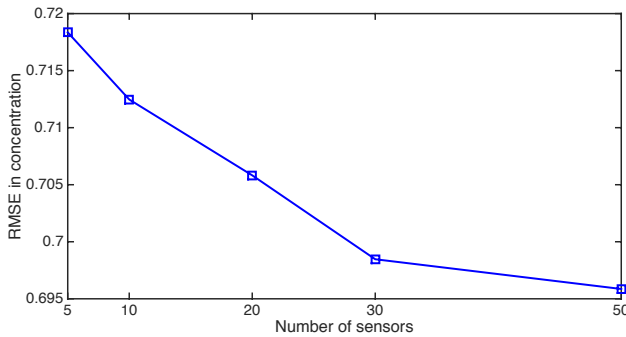

Fig. 5. RMSE of the concentration estimates as a function of the number of sensors deployed over the monitoring area.

strate the validity of the above stated noise-aided paradigm in the recursive state estimation with binary measurements and, thus, represent an interesting contribution of this work. Finally, Fig. 5 shows the evolution of the RMSE as a function of the number of binary observations available at the fusion center.

VI. CONCLUSIONS AND FUTURE WORK

State estimation with binary sensors has been formulated as a *Moving Horizon* (MH) *Maximum A posteriori Probability* (MAP) optimization problem and it has been shown how such a problem turns out to be convex in the linear system case. Simulation results relative to a dynamic field estimation case-study have exhibited the conjectured noise-aided feature of the proposed estimator in that the estimation accuracy improves, starting from a null measurement noise, until the variance of the latter achieves an optimal value beyond which estimation performance deteriorates.

Future work on the topic will concern stability properties of the MH-MAP state estimator and its application to target tracking with binary proximity sensors.

REFERENCES

[1] L.Y. Wang, J.F. Zhang, and G.G. Yin, "System identification using binary sensors", *IEEE Transactions on Automatic Control*, vol. 48, no. 11, pp. 1892–1907, 2003.

[2] L.Y. Wang, G.G. Yin, and J.F. Zhang, "Joint identification of plant rational models and noise distribution functions using binary-valued observations", *Automatica*, vol. 42, no. 4, pp. 543-547, 2006.

[3] B. Ristic, A. Gunatilaka and R. Gailis, "Achievable accuracy in Gaussian plume parameter estimation using a network of binary sensors", *Information Fusion*, vol. 25, pp. 42-48, 2015.

[4] S. Vijayakumaran, Y. Levinbook and T.F. Wong, "Maximum likelihood localization of a diffusive point source using binary observations", *IEEE Trans. on Sign. Process.*, vol. 55, no. 2, pp. 665–676, 2007.

[5] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks - part I: Gaussian case", *IEEE Trans. on Sign. Process.*, vol. 54, no. 3, pp. 1131–1143, 2006.

[6] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks - part II: unknown probability density function", *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2784–2796, 2006.

[7] L.Y. Wang, C. Li, G.G. Yin, L. Guo, and C.-Z. Xu, "State observability and observers of linear-time-invariant systems under irregular sampling and sensor limitations", *IEEE Transactions on Automatic Control*, vol. 56, no. 11, pp. 2639–2654, 2011.

[8] G. Battistelli, L. Chisci and S. Gherardini, "Moving horizon state estimation for discrete-time linear systems with binary sensors", *54th IEEE Conference on Decision and Control*, pp. 2414-2419, 2015.

[9] Er-wei Bai, H. E. Baidoo-Williams, R. Mudumbai and S. Dasgupta "Robust tracking of piecewise linear trajectories with binary sensor networks", *Automatica*, vol. 61, pp. 134-145, 2015.

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

[11] P.M. Djuric, M. Vemula and M.F. Bugallo, "Signal processing by particle filtering for binary sensor networks", *Proc. Digital Signal Processing Workshop*, pp. 263–267, 2004.

[12] P.M. Djuric, M. Vemula and M.F. Bugallo, "Target tracking by particle filtering in binary sensor networks", *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2229–2238, 2008.

[13] J. Teng, H. Snoussi and C. Richard, "Decentralized variational filtering for simultaneous sensor localization and target tracking in binary sensor networks", *IEEE Transactions on Mobile Computing*, vol. 9, no. 10, pp. 1465–1477, 2010.

[14] A. Capponi, I. Fatkullin and Ling Shi, "Stochastic filtering for diffusion processes with level crossings", *IEEE Transactions on Automatic Control*, vol. 56, no. 9, pp. 2201-2206, 2011.

[15] G. Ferrari-Trecate, D. Mignone and M. Morari, "Moving horizon estimation for hybrid systems", *IEEE Transactions on Automatic Control*, vol. 47, no. 10, pp. 1663–1676, 2002.

[16] C.V. Rao, J.B. Rawlings, and D.Q. Mayne, "Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations", *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 246-258, 2003.

[17] A. Alessandri, M. Baglietto, and G. Battistelli, "Receding horizon estimation for discrete-time linear systems", *IEEE Transactions on Automatic Control*, vol. 48, no. 3, pp. 473–478, 2003.

[18] A. Alessandri, M. Baglietto, and G. Battistelli, "Moving horizon state estimation for nonlinear discrete-time systems: new stability results and approximation schemes", *Automatica*, vol. 44, no. 7, pp. 1753–1765, 2008.

[19] A. Alessandri, M. Baglietto, G. Battistelli, and V. Zavala, "Advances in moving horizon estimation for nonlinear systems," *Proc. 49th IEEE Conference on Decision and Control (CDC)*, pp. 5681-5688, 2010.

[20] A. Liu, L. Yu, W.-A. Zhang and M.Z.Q. Chen, "Moving horizon estimation for networked systems with quantized measurements and packet dropouts", *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 7, pp. 1823–1834, 2013.

[21] R. A. Delgado and G. C. Goodwin, "A combined MAP and Bayesian scheme for finite data and/or moving horizon estimation", *Automatica*, vol. 50, no. 4, pp. 1116-1121, 2014.

[22] G. Battistelli, L. Chisci, N. Forti, G. Pelosi, and S. Selleri, "Point source estimation via finite element multiple-model Kalman filtering", *54th IEEE Conference on Decision and Control*, pp. 4984-4989, 2015.

[23] G. Battistelli, L. Chisci, N. Forti, G. Pelosi, and S. Selleri, "Distributed finite element Kalman filter", *14th IEEE European Control Conference*, pp. 3695-3700, 2015.

[24] S.C. Brenner and L.R. Scott, *The mathematical theory of finite element methods*, Springer–Verlag, New York, NY, 1996.

[25] S. Boyd and L. Vandenderghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.