

# On Fusion of Sensor Measurements and Observation with Uncertain Timestamp for Target Tracking

Clas Veibäck, Gustaf Hendeby and Fredrik Gustafsson

**Linköping University Post Print**



N.B.: When citing this work, cite the original article.

Original Publication:

Clas Veibäck, Gustaf Hendeby and Fredrik Gustafsson, On Fusion of Sensor Measurements and Observation with Uncertain Timestamp for Target Tracking, 2016, Proceedings of the 19th International Conference on Information Fusion, Pages 1268-1275.

<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7528030>

Copyright: ISIF <http://isif.org>

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-130349>

# On Fusion of Sensor Measurements and Observation with Uncertain Timestamp for Target Tracking

Clas Veibäck, Gustaf Hendeby, Fredrik Gustafsson  
Dept. Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden.  
Email: `firstname.lastname@liu.se`

**Abstract**—We consider a target tracking problem where, in addition to the usual sensor measurements, accurate observations with uncertain timestamps are available. Such observations could, e.g., come from traces left by a target or from witnesses of an event, and have the potential in some scenarios to improve the accuracy of an estimate significantly. The Bayesian solution to the smoothing problem for one observation with uncertain timestamp is derived for a linear Gaussian state space model. The joint and marginal distributions of the states and uncertain time are derived, as well as the minimum mean squared error (MMSE) and maximum a posteriori (MAP) estimators. To attain an intuition for the problem in consideration a simple first-order example is presented and its posterior distributions and point estimators are compared and examined in some depth.

## I. INTRODUCTION

We consider a seemingly novel formulation of the general target tracking problem with mixed measurements of the following kinds:

- 1) Usual position measurements with precise timestamps.
- 2) Observations that are fairly precise in position but come with a relatively large uncertainty in time.

To explain a specific case where this occurs, we start with an application to sprint orienteering. Fig. 1 illustrates the information that is available together with an estimated trajectory as presented in [1]. This kind of global positioning system (GPS) data is available for download from all major races. Also the map is available to download. Unfortunately, the time of stamping at each control is not saved to the server. Since we know that the runner was not disqualified, we have indirect information that the runner must have passed control number 11, just as all other controls. We also know that the controls are passed in ascending order, which gives a rough prior on the ordered list of uncertain timestamps. Clearly, the official track marked in red is unreliable, and this deficiency happens all the time for runners in city races, because of GPS multipath problems. Apparently, this limits the viewing experience of the on-line and off-line spectators.

A control is a kind of point of attraction that can be used in target tracking algorithms to force the trajectory to pass the control. However, classical theory does not cover this case where we have observations of the second kind above.

There are many other potential applications. In fact, our interest in this problem originally comes from an idea for *fusion of radar and bullshit*. In the project Smart Savannahs [2] we are developing tracking algorithms for following rhinos inside a sanctuary. This can be done with sensors of the first



Fig. 1. A snapshot from the world championship in sprint orienteering in Venice 2014. Soren Bobach is on his way to victory, and the live public can watch the red trajectory based on map matching of GPS data sampled with several seconds sampling time. The blue line shows his own reconstruction (can be seen as ground truth), while the green one is based on a particle filter supported by map information, see [1].

kind above, complemented with manual observations from the rangers. These rangers spend a considerable time on patrol, recording all indirect rhino sightings, including droppings, foot prints, rest places and browsing damages. These are all of the second kind above, and they are particularly useful in the labelling problem, *i.e.*, which track belongs to which rhino.

Another application is in crime scene investigations. The place of the crime is often known precisely, but not the time, and witness statements often contain uncertainty in both time and place, while surveillance cameras are precise in time.

Parameter and state estimation in dynamical models with uncertain and stochastic timestamps is to our knowledge a less studied problem in literature. Sampling jitter is one exception, see for instance [3] where estimation problems are formulated based on parametric models mixed with prior distributions of the sampling times. However, this idea has seemingly not been explored in target tracking.

One attempt would be to consider the uncertain timestamps as latent variables and use for example the powerful expectation-maximization (EM) algorithm [4]. For instance, [5] discusses missing data in state space models, using the EM

algorithm, and this is further developed in *e.g.* [6] and [7]. Further, [8] discusses using the EM algorithm to estimate synchronization. The latter is related to timing issues, but still we have not found any related study to our problem formulation.

One can see observations of the second kind as out of sequence (OOS) measurements. As presented in literature, these have precise timestamps, but they are available at a time later than the actual measurement, and have to be processed in a smoothing framework similar to ours. General theory is developed in [9] and [10]. In this context, [11] discusses acceleration bias caused by synchronization issues between GPS and inertial measurement unit (IMU), which is further elaborated on in [12]. Further, [13] discusses time delays caused by propagation.

Our approach is based on extending the optimal Bayes filter to uncertain stochastic timestamps. Given a prior on the timestamps, the position posterior conditioned on observations of both the first and second kind is derived. This leads to a posterior that can be expressed as a Gaussian mixture, where the number of modes for each observation of the second kind equals the support of the timestamp prior expressed in basic sampling intervals. Depending on the model, a Kalman filter (KF), extended Kalman filter (EKF) or unscented Kalman filter (UKF) can be applied for each hypothesis of timestamp, and the final posterior is a weighted sum of Gaussian distributions, where each distribution needs a smoothing step similar to the OOS problem. The total number of modes increases exponentially with the number of observations of the second kind. Here, there is an abundance of methods for limiting the number of modes in a Gaussian mixture, see for instance [14]. However, in this contribution we limit the number of observations with uncertain timestamps to one, motivated by the fact that measurements of the second kind are quite sparse compared to the measurements of the first kind, and thus can be processed one at a time without too much loss of optimality.

The outline of the paper is as follows: An illustrative example is presented in Sec. II, providing an intuition for the problem. Sec. III generalizes the example to linear Gaussian state space models and discusses some objectives that might be of interest. Sec. IV and V proceed by deriving posterior distributions and point estimators, which are then computed in Sec. VI for the example in Sec. II. A discussion regarding the results is presented in Sec. VII and in Sec. VIII the paper is summarized and future work is proposed.

## II. ILLUSTRATIVE EXAMPLE

Our tracking problem is most interesting in two- and three-dimensional spaces, where we want to force the estimated trajectory into the points of attractions defined by the observations of the second kind. However, the numerical example is preferably defined in one dimension, where the joint distribution of position and uncertain time can be plotted, as well as the position posterior as a function of time. We consider a trajectory that is measured to start at position 0 at time 0 and

end in 1 at time 1. If we then are given the observation 0.5 with a flat prior on the time, we can of course argue that any physical quantity that starts at 0 and ends at 1 must pass 0.5, so this observation contains no information. This is not quite true, however, since the posterior distribution will be shown to narrow down in an interesting way. We will also consider the case of an observation 1.5, which can be seen as an illustration of what happens in the  $y$ -dimension in Fig. 1.

We define *measurements* and *observations*, respectively, in the following way:

- 1)  $y(t_k)$  are the usual measurements occurring at time  $t_k$  with an uncertainty  $\mathcal{N}(0, R^y)$ .
- 2)  $z$  is the observation taking place at the uncertain time  $\tilde{\tau}$  with prior  $p(\tilde{\tau})$  and an (amplitude) uncertainty  $\mathcal{N}(0, R^z)$ .

In principle the problem can be solved in continuous time; however, as the involved integrals are difficult to compute even in the simplest of cases we choose to grid  $\tilde{\tau}$ , effectively substituting the involved integrals with sums. This results in a discretized state space model, with sampling time  $T_s$ , where the sample time ensures sufficient temporal resolution in  $\tilde{\tau}$ .

The model for the illustrative example, after discretization with  $T_s = 1/N$  and  $\tilde{\tau} = T_s\tau$ , is defined by

$$x_k = x_{k-1} + v_k, \quad v_k \sim \mathcal{N}(0, Q), \quad (1a)$$

for  $k \in \{1, \dots, N\}$  and

$$y_j = x_j + e_j^y, \quad e_j^y \sim \mathcal{N}(0, R^y) \quad (1b)$$

where the two measurements of the first kind are  $y_1$  and  $y_N$ . The observation of the second kind is modelled as

$$z = x_\tau + e^z, \quad e^z \sim \mathcal{N}(0, R^z) \quad (1c)$$

sampled at an uncertain time  $\tau \in \{1, \dots, N\}$ , where the prior on  $\tau$  is discretized as

$$p(\tau) \propto p(\tilde{\tau} = T_s\tau). \quad (2)$$

Our aim is to compute the posterior distribution of the position  $p(x_k|y_1, y_N, z)$ . This distribution is given by the marginalization of the joint posterior distribution of  $x_k$  and  $\tau$ ,

$$\begin{aligned} p(x_k|y_1, y_N, z) &= \sum_{\tau=1}^N p(x_k, \tau|y_1, y_N, z) \\ &= \sum_{\tau=1}^N p(x_k|y_1, y_N, z, \tau)p(\tau|y_1, y_N, z), \end{aligned} \quad (3)$$

where the first factor is recognized as the smoothed posterior distribution of the position given the measured and observed positions and the uncertain time  $\tau$ , which is readily obtained using a Kalman smoother [15] as  $p(x_k|y_1, y_N, z, \tau) = \mathcal{N}(x_k|\hat{x}_k^\tau, P_k^\tau)$ . The superscript  $\tau$  denotes that the quantity is conditioned on  $\tau$ , as well as the measurements and observation. The second factor can be computed using Bayes' theorem as

$$\begin{aligned} w_\tau &\triangleq p(\tau|y_1, y_N, z) \propto p(z|\tau, y_1, y_N)p(\tau|y_1, y_N) \\ &= \mathcal{N}(z|\hat{x}_\tau^y, P_\tau^y + R^z)p(\tau), \end{aligned} \quad (4)$$

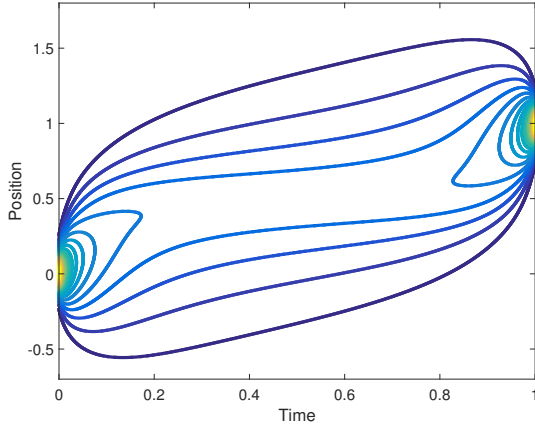


Fig. 2. Contour plot of the posterior distribution for the first scenario.

where  $p(x_k|y_1, y_N) = \mathcal{N}(x_k|\hat{x}_k^y, P_k^y)$  is obtained using a Kalman smoother given only the measurements, see Sec. IV for details. The superscript  $y$  denotes a quantity conditioned only on the measurements with precise timestamps.

The posterior distribution in (3) therefore simplifies to the Gaussian mixture

$$p(x_k|y_1, y_N, z) = \sum_{\tau=1}^N w_{\tau} \mathcal{N}(x_k|\hat{x}_k^{\tau}, P_k^{\tau}). \quad (5)$$

The minimum mean squared error (MMSE) estimate [16] and its approximate covariance, see Sec. V-A, are given by

$$\hat{x}_k = \sum_{\tau=1}^N w_{\tau} \hat{x}_k^{\tau} \quad \text{and} \quad P_k = \sum_{\tau=1}^N w_{\tau} (P_k^{\tau} + (\hat{x}_k^{\tau} - \hat{x}_k)^2). \quad (6)$$

Two scenarios are considered and outlined here, see Sec. VI-A for more details. In the first scenario the measured positions are  $y_1 = 0$  and  $y_N = 1$ , and the observation of the second kind is  $z = 0.5$  and the time  $\tilde{\tau}$  is very uncertain, which is modelled as a flat prior distribution. The posterior distribution of the position is shown in Fig. 2. The MMSE estimate of the position given the measurements and observation, and its uncertainty, are shown in Fig. 3, together with the estimate given only the measured positions for comparison. It is clear that the uncertainty in the estimate decreases with the observed position, despite the uncertain time. Furthermore, the estimate tends towards the observation.

In the second scenario the measured positions are the same as before, but the observed position is  $z = 1.5$  and, in addition, the timestamp is quite certain at about  $\tilde{\tau} = 0.5$ , modelled as  $p(\tilde{\tau}) = \mathcal{N}(\tilde{\tau}|0.5, 0.01)$ . The posterior distribution and MMSE estimate are shown in Fig. 4 and 5, respectively.

### III. PROBLEM DESCRIPTION

The example in Sec. II is a special case of a more general problem that could be nonlinear and include several observations with uncertain timestamps. In Sec. III-A the slightly more general linear Gaussian model is defined, still only with a single observation with uncertain timestamp. Some posterior

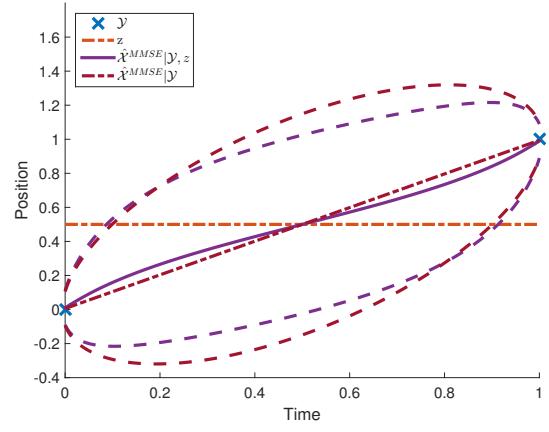


Fig. 3. The MMSE estimate and its uncertainty for the first scenario compared to the MMSE estimate given only measurements.

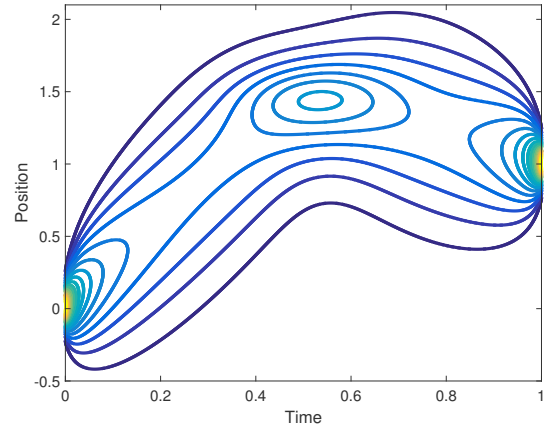


Fig. 4. Contour plot of the posterior distribution for the second scenario.

distributions and point estimators related to the problem are discussed in Sec. III-B.

#### A. Model

Consider a linear state space model with independent additive Gaussian noise, discretized with sampling time  $T_s$  chosen to ensure sufficient temporal resolution,

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{v}_k, \quad \mathbf{v}_k \sim \mathcal{N}(0, \mathbf{Q}_k) \quad (7a)$$

$$\mathbf{y}_j = \mathbf{H}_j^y \mathbf{x}_j + \mathbf{e}_j^y, \quad \mathbf{e}_j^y \sim \mathcal{N}(0, \mathbf{R}_j^y) \quad (7b)$$

$$\mathbf{z} = \mathbf{H}^z \mathbf{x}_{\tilde{\tau}} + \mathbf{e}^z, \quad \mathbf{e}^z \sim \mathcal{N}(0, \mathbf{R}^z) \quad (7c)$$

with prior distributions  $p(\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_0|\bar{\mathbf{x}}_0, \mathbf{P}_0)$  and  $p(\tau)$ , defined as in (2), where

$$\mathbf{x}_k \in \mathbb{R}^n, \quad k \in \mathcal{K} \triangleq \{1, \dots, N\} \quad (8a)$$

are the states and

$$\mathbf{y}_j \in \mathbb{R}^{m^y}, \quad j \in \mathcal{J} \subseteq \mathcal{K}, \quad \text{and} \quad \mathbf{z} \in \mathbb{R}^{m^z} \quad (8b)$$

are measurements of the first kind with precise timestamps and observation of the second kind with uncertain timestamp, respectively, denoted *measurements* and *observation* in the sequel. The observation time is limited to the interval  $\tau \in \mathcal{K}$ .

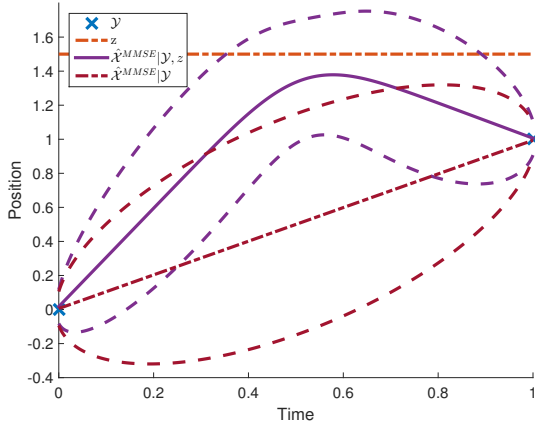


Fig. 5. The MMSE estimate and its uncertainty for the second scenario compared to the MMSE estimate given only measurements.

To simplify notation,  $\mathcal{X} \triangleq \{\mathbf{x}_k\}_{k \in \mathcal{K}}$  and  $\mathcal{Y} \triangleq \{\mathbf{y}_j\}_{j \in \mathcal{J}}$  will be used to denote the ordered sets of all states and measurements, respectively, where both will be treated as column vectors where applicable.

### B. Posterior Distributions and Point Estimators

Given the model, measurements and observation, one is in general interested in estimating the states, the uncertain time or both. The following related posteriors are considered:

- the joint posterior of all states  $p(\mathcal{X}|\mathcal{Y}, \mathbf{z})$ ;
- the posterior of the states  $p(\mathbf{x}_k|\mathcal{Y}, \mathbf{z})$ ,  $\forall k \in \mathcal{K}$ ;
- the joint posterior of all states and the uncertain time  $p(\mathcal{X}, \tau|\mathcal{Y}, \mathbf{z})$ ;
- the posterior of the uncertain time  $p(\tau|\mathcal{Y}, \mathbf{z})$ ; and
- the partially maximized distribution  $\max_{\mathcal{X}} p(\mathcal{X}, \tau|\mathcal{Y}, \mathbf{z})$  of the uncertain time.

Given the posterior distributions, common point estimators are the MMSE and maximum a posteriori (MAP) estimators [16], where the following are considered:

- the MMSE estimator of  $p(\mathcal{X}|\mathcal{Y}, \mathbf{z})$ ;
- the MMSE estimator of  $p(\mathbf{x}_k|\mathcal{Y}, \mathbf{z})$ ,  $\forall k \in \mathcal{K}$ ;
- the MAP estimator of  $p(\mathcal{X}|\mathcal{Y}, \mathbf{z})$ ;
- the MAP estimator of  $p(\mathbf{x}_k|\mathcal{Y}, \mathbf{z})$ ,  $\forall k \in \mathcal{K}$  as an approximation to the above estimator;
- the joint MAP (JMAP) estimator of  $p(\mathcal{X}, \tau|\mathcal{Y}, \mathbf{z})$ ;
- the MAP estimator of  $p(\tau|\mathcal{Y}, \mathbf{z})$ ; and
- the MAP estimator of  $\max_{\mathcal{X}} p(\mathcal{X}, \tau|\mathcal{Y}, \mathbf{z})$ .

## IV. POSTERIOR DISTRIBUTIONS

In this section the posterior distributions discussed in Sec. III-B are derived for the model in Sec. III-A.

The joint posterior distribution of all states is given by marginalization as

$$p(\mathcal{X}|\mathcal{Y}, \mathbf{z}) = \sum_{\tau \in \mathcal{K}} p(\mathcal{X}, \tau|\mathcal{Y}, \mathbf{z}), \quad (9)$$

where the joint posterior of all states and the uncertain time can be computed up to a constant factor using Bayes' theorem and the independence assumption as

$$\begin{aligned} p(\mathcal{X}, \tau|\mathcal{Y}, \mathbf{z}) &\propto p(\tau)p(\mathbf{x}_0)p(\mathbf{z}|\mathbf{x}_\tau, \tau) \\ &\quad \times \prod_{j \in \mathcal{J}} p(\mathbf{y}_j|\mathbf{x}_j) \prod_{k \in \mathcal{K}} p(\mathbf{x}_k|\mathbf{x}_{k-1}) \\ &= p(\tau)\mathcal{N}(\mathbf{x}_0|\bar{\mathbf{x}}_0, \mathbf{P}_0)\mathcal{N}(\mathbf{z}|\mathbf{H}^z\mathbf{x}_\tau, \mathbf{R}^z) \\ &\quad \times \prod_{j \in \mathcal{J}} \mathcal{N}(\mathbf{y}_j|\mathbf{H}_j^y\mathbf{x}_j, \mathbf{R}_j^y) \\ &\quad \times \prod_{k \in \mathcal{K}} \mathcal{N}(\mathbf{x}_k|\mathbf{F}_k\mathbf{x}_{k-1}, \mathbf{Q}_k) \\ &\propto p(\tau) \exp \left[ -\frac{1}{2} \left( \|\mathbf{x}_0 - \bar{\mathbf{x}}_0\|_{\mathbf{P}_0}^2 \right. \right. \\ &\quad \left. \left. + \|\mathbf{z} - \mathbf{H}^z\mathbf{x}_\tau\|_{\mathbf{R}^z}^2 \right. \right. \\ &\quad \left. \left. + \sum_{j \in \mathcal{J}} \|\mathbf{y}_j - \mathbf{H}_j^y\mathbf{x}_j\|_{\mathbf{R}_j^y}^2 \right. \right. \\ &\quad \left. \left. + \sum_{k \in \mathcal{K}} \|\mathbf{x}_k - \mathbf{F}_k\mathbf{x}_{k-1}\|_{\mathbf{Q}_k}^2 \right) \right], \quad (10) \end{aligned}$$

where  $\|\mathbf{b}\|_{\mathbf{A}}^2 = \mathbf{b}^T \mathbf{A}^{-1} \mathbf{b}$  is the Mahalanobis distance.

Alternatively, (10) can be factorized as

$$p(\mathcal{X}, \tau|\mathcal{Y}, \mathbf{z}) = p(\mathcal{X}|\tau, \mathcal{Y}, \mathbf{z})p(\tau|\mathcal{Y}, \mathbf{z}). \quad (11)$$

For the model in (7) the first distribution is Gaussian, which can be obtained using weighted least squares [17] as

$$p(\mathcal{X}|\tau, \mathcal{Y}, \mathbf{z}) = \mathcal{N}(\mathcal{X}|\hat{\mathcal{X}}^\tau, \mathbf{P}^\tau). \quad (12)$$

The posterior of  $\tau$  in (11) can be computed using Bayes' theorem and marginalization as

$$\begin{aligned} w_\tau &\triangleq p(\tau|\mathcal{Y}, \mathbf{z}) \propto p(\tau)p(\mathbf{z}|\tau, \mathcal{Y}) \\ &= p(\tau) \int_{\mathbf{x}_\tau} p(\mathbf{z}|\mathbf{x}_\tau, \tau)p(\mathbf{x}_\tau|\mathcal{Y}) d\mathbf{x}_\tau. \quad (13) \end{aligned}$$

In (13) the first factor in the integrand is given by (7c) and the second factor is the smoothed posterior of the state given only measurements with precise timestamps, which is a Gaussian distribution,  $\mathcal{N}(\mathbf{x}_\tau|\hat{\mathbf{x}}_\tau^y, \mathbf{P}_\tau^y)$ , that can be computed efficiently using the Rauch-Tung-Striebel smoother [15] or the Forward-Backward smoother [18]. Evaluating the integral in (13) results in

$$\begin{aligned} w_\tau &\propto p(\tau) \int_{\mathbf{x}_\tau} \mathcal{N}(\mathbf{z}|\mathbf{H}^z\mathbf{x}_\tau, \mathbf{R}^z)\mathcal{N}(\mathbf{x}_\tau|\hat{\mathbf{x}}_\tau^y, \mathbf{P}_\tau^y) d\mathbf{x}_\tau \\ &= p(\tau)\mathcal{N}(\mathbf{z}|\mathbf{H}^z\hat{\mathbf{x}}_\tau^y, \mathbf{H}^z\mathbf{P}_\tau^y(\mathbf{H}^z)^T + \mathbf{R}^z) \\ &= p(\tau)\mathcal{N}(\mathbf{z}|\hat{\mathbf{z}}_\tau, \mathbf{S}_\tau) \quad (14) \end{aligned}$$

and combining (9), (11), (12) and (14) results in the Gaussian mixture

$$p(\mathcal{X}|\mathcal{Y}, \mathbf{z}) = \sum_{\tau \in \mathcal{K}} w_\tau \mathcal{N}(\mathcal{X}|\hat{\mathcal{X}}^\tau, \mathbf{P}^\tau). \quad (15)$$

Similarly, the marginalization of the posterior in (15) can be derived as

$$p(\mathbf{x}_k|\mathcal{Y}, \mathbf{z}) = \sum_{\tau \in \mathcal{K}} w_\tau \mathcal{N}(\mathbf{x}_k|\hat{\mathbf{x}}_k^\tau, \mathbf{P}_k^\tau), \quad (16)$$

where  $\mathcal{N}(\mathbf{x}_k|\hat{\mathbf{x}}_k^\tau, \mathbf{P}_k^\tau)$  is the smoothed posterior of the state conditional on  $\tau$ .

## V. POINT ESTIMATORS

In many situations one is interested in point estimates for the posterior distributions in Sec. IV, *e.g.* for presentation, where visualizing the posterior distribution often is difficult. Point estimators are straightforward for Gaussian distributions [16], because they are symmetric and unimodal. The posteriors derived in Sec. IV are Gaussian mixtures, that in general are neither symmetric nor unimodal, making obtaining point estimates more difficult. Two common options are the MMSE estimator and the MAP estimator, which are derived for the model in (7).

### A. Minimum Mean Squared Estimator

The MMSE estimator, derived in [16], of the posterior  $p(\mathcal{X}|\mathcal{Y}, \mathbf{z})$  in (15) is given by

$$\begin{aligned}\hat{\mathcal{X}}^{MMSE} &= \arg \min_{\hat{\mathcal{X}}} E_{\mathcal{X}|\mathcal{Y}, \mathbf{z}} [(\hat{\mathcal{X}} - \mathcal{X})^T (\hat{\mathcal{X}} - \mathcal{X})] \\ &= E_{\mathcal{X}|\mathcal{Y}, \mathbf{z}} [\mathcal{X}] = \sum_{\tau \in \mathcal{K}} w_\tau \hat{\mathcal{X}}^\tau,\end{aligned}\quad (17)$$

which also minimizes the mean squared error (MSE) for each component, resulting in an identical MMSE estimator for the marginalized posterior for each state,  $p(\mathbf{x}_k|\mathcal{Y}, \mathbf{z})$ .

The classical MSE is

$$\begin{aligned}\mathbf{P}^{MMSE} &= \sum_{\tau \in \mathcal{K}} w_\tau \left( \mathbf{P}^\tau \right. \\ &\quad \left. + (\hat{\mathcal{X}}^\tau - \hat{\mathcal{X}}^{MMSE})(\hat{\mathcal{X}}^\tau - \hat{\mathcal{X}}^{MMSE})^T \right),\end{aligned}\quad (18)$$

which is an approximate estimate of the performance of the estimator. Note that this is the classical MSE, as described in [16], rather than the Bayesian MSE, which is more complicated to compute due to the additional expectation with respect to the measurements and observation.

### B. Maximum A Posteriori Estimator

The MAP estimator, derived in [16], of the joint posterior of all states,  $p(\mathcal{X}|\mathcal{Y}, \mathbf{z})$ , is given by

$$\begin{aligned}\hat{\mathcal{X}}^{MAP} &= \arg \max_{\mathcal{X}} p(\mathcal{X}|\mathcal{Y}, \mathbf{z}) \\ &= \arg \max_{\mathcal{X}} \sum_{\tau \in \mathcal{K}} p(\mathcal{X}, \tau|\mathcal{Y}, \mathbf{z}),\end{aligned}\quad (19)$$

where  $p(\mathcal{X}, \tau|\mathcal{Y}, \mathbf{z})$  is preferably computed using (10) to avoid computing the smoothed joint posterior of all states conditional on all  $\tau$ . Even for the model in Sec. III-A it is often difficult to maximize (19) considering the high dimensionality and that Gaussian mixtures in general have multiple local maxima.

An approximation is to assume independence between the states and compute the MAP estimate of  $p(\mathbf{x}_k|\mathcal{Y}, \mathbf{z})$  in (16) for each state separately,

$$\begin{aligned}\hat{\mathbf{x}}_k^{MAP} &= \arg \max_{\mathbf{x}_k} p(\mathbf{x}_k|\mathcal{Y}, \mathbf{z}) \\ &= \arg \max_{\mathbf{x}_k} \sum_{\tau \in \mathcal{K}} w_\tau \mathcal{N}(\mathbf{x}_k|\hat{\mathbf{x}}_k^\tau, \mathbf{P}_k^\tau).\end{aligned}\quad (20)$$

Maximization of a Gaussian mixture is still required, but the dimensionality is significantly reduced.

To estimate the uncertain times, the MAP estimate of (14) can be computed as

$$\begin{aligned}\hat{\tau}^{MAP} &= \arg \max_{\tau} \{p(\tau) \mathcal{N}(\mathbf{z}|\hat{\mathbf{z}}_\tau, \mathbf{S}_\tau)\} \\ &= \arg \max_{\tau} \{2 \ln p(\tau) - \|\mathbf{z} - \hat{\mathbf{z}}_\tau\|_{\mathbf{S}_\tau}^2 - \ln |\mathbf{S}_\tau|\},\end{aligned}\quad (21)$$

where the monotonically increasing property of logarithms is used in the second equality.

### C. Joint Maximum A Posteriori Estimate

The JMAP estimator of the joint posterior distribution  $p(\mathcal{X}, \tau, |\mathcal{Y}, \mathbf{z})$  is given by

$$\begin{aligned}\{\hat{\mathcal{X}}^{JMAP}, \hat{\tau}^{JMAP}\} &= \arg \max_{\mathcal{X}, \tau} p(\mathcal{X}, \tau|\mathcal{Y}, \mathbf{z}) \\ &= \arg \max_{\mathcal{X}, \tau} p(\mathcal{X}|\tau, \mathcal{Y}, \mathbf{z}) p(\tau|\mathcal{Y}, \mathbf{z}) \\ &= \arg \max_{\mathcal{X}, \tau} w_\tau \mathcal{N}(\mathcal{X}|\hat{\mathcal{X}}^\tau, \mathbf{P}^\tau),\end{aligned}\quad (22)$$

where (11)–(13) are used, which further reduces to

$$\begin{aligned}\hat{\tau}^{JMAP} &= \arg \max_{\tau} \left\{ \max_{\mathcal{X}} w_\tau \mathcal{N}(\mathcal{X}|\hat{\mathcal{X}}^\tau, \mathbf{P}^\tau) \right\} \\ &= \arg \max_{\tau} \frac{w_\tau}{\sqrt{|\mathbf{P}^\tau|}} \\ &= \arg \max_{\tau} \left\{ w_\tau \sqrt{|\mathbf{R}^z + \mathbf{H}^z \mathbf{P}_\tau (\mathbf{H}^z)^T|} \right\} \\ &= \arg \max_{\tau} \left\{ p(\tau) \mathcal{N}(\mathbf{z}|\hat{\mathbf{z}}_\tau, \mathbf{S}_\tau) \sqrt{|\mathbf{S}_\tau|} \right\} \\ &= \arg \max_{\tau} \left\{ 2 \ln p(\tau) - \|\mathbf{z} - \hat{\mathbf{z}}_\tau\|_{\mathbf{S}_\tau}^2 \right\},\end{aligned}\quad (23a)$$

$$\hat{\mathcal{X}}^{JMAP} = \hat{\mathcal{X}}^{\hat{\tau}^{JMAP}},\quad (23b)$$

where Proposition 1 and (14) are used. Note that the function no longer depends on the joint covariance of all states, only the covariance of each state.

**Proposition 1.** *Given a joint posterior distribution  $p(\mathcal{X}|\mathcal{Y})$  of a linear Gaussian state space model described by (7), where  $\mathbf{P} = \text{cov}(\mathcal{X}|\mathcal{Y})$ ,  $\mathbf{P}_k = \text{cov}(\mathbf{x}_k|\mathcal{Y})$  and  $\mathbf{P}$  and  $\mathbf{R}^z$  are non-singular, the determinant of  $\mathbf{P}^\tau = \text{cov}(\mathcal{X}|\mathcal{Y}, \mathbf{z}, \tau)$  is given by*

$$|\mathbf{P}^\tau| = |\mathbf{P}| |\mathbf{R}^z| |\mathbf{R}^z + \mathbf{H}^z \mathbf{P}_\tau (\mathbf{H}^z)^T|^{-1}.\quad (24)$$

*Proof.* Let the observation be  $\mathbf{z} = \mathbf{H}^z \mathbf{x}_\tau + \mathbf{e}^z = \bar{\mathbf{H}}_\tau^z \mathcal{X} + \mathbf{e}^z$ , then

$$\begin{aligned}|\mathbf{P}^\tau|^{-1} &= |[\mathbf{P}^{-1} + (\bar{\mathbf{H}}_\tau^z)^T (\mathbf{R}^z)^{-1} \bar{\mathbf{H}}_\tau^z]^{-1}|^{-1} \\ &= |\mathbf{P}^{-1} + (\bar{\mathbf{H}}_\tau^z)^T (\mathbf{R}^z)^{-1} \bar{\mathbf{H}}_\tau^z| \\ &= |\mathbf{P}^{-1} \mathbf{I} + (\mathbf{R}^z)^{-1} \bar{\mathbf{H}}_\tau^z \mathbf{P} (\bar{\mathbf{H}}_\tau^z)^T| \\ &= |\mathbf{P}^{-1} |\mathbf{R}^z|^{-1} |\mathbf{R}^z + \mathbf{H}^z \mathbf{P}_\tau (\mathbf{H}^z)^T|,\end{aligned}\quad (25)$$

where Millman's formula [19] was used in the first equality and Sylvester's identity [20] in the third equality, concluding the proof.  $\square$

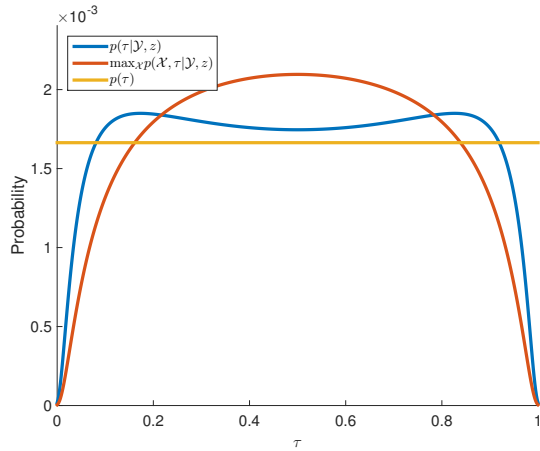


Fig. 6. The distributions  $p(\tau|y_1, y_N, z)$ ,  $\max_{\mathcal{X}} p(\mathcal{X}, \tau|y_1, y_N, z)$  and  $p(\tau)$  for the first scenario.

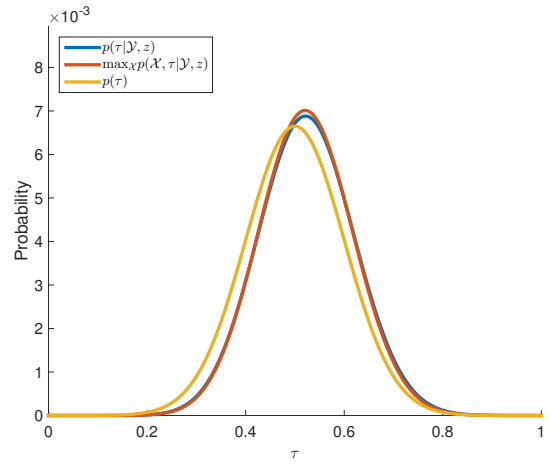


Fig. 7. The distributions  $p(\tau|y_1, y_N, z)$ ,  $\max_{\mathcal{X}} p(\mathcal{X}, \tau|y_1, y_N, z)$  and  $p(\tau)$  for the second scenario.

## VI. RESULTS

Two different scenarios are defined in Sec. VI-A and their posterior distributions and point estimators are presented in Sec. VI-B and VI-C.

### A. Model

To illustrate the problem, the model described in Sec. II is assumed with the resolution  $N = 600$  and  $T_s = 1/N$ , the noise covariances  $Q = 1.67/N$ ,  $R^y = 0.01$  and  $R^z = 0.01$ , the position prior  $\bar{x}_0 = 0$  and  $P_0 = 10^9$ , and the measurements  $y_1 = 0$  and  $y_N = 1$ . Two different scenarios are considered and compared with the following parameters:

- 1)  $z = 0.5$  and  $p(\tilde{\tau}) = 1$ .
- 2)  $z = 1.5$  and  $p(\tilde{\tau}) = \mathcal{N}(\tilde{\tau}|0.5, 0.01)$ .

The priors  $p(\tilde{\tau})$  are discretized using (2).

### B. Posterior Distributions

The marginal posterior distribution  $p(x_k|y_1, y_N, z)$  is shown in Fig. 2 and 4 for the first and second scenario respectively and the distributions  $\max_{\mathcal{X}} p(\mathcal{X}, \tau|y_1, y_N, z)$ ,  $p(\tau|y_1, y_N, z)$  and  $p(\tau)$  are shown in Fig. 6 and 7.

### C. Point Estimators

The MMSE estimate is shown in Fig. 3 and 5 for the first and second scenario, respectively, together with the uncertainty and the estimate using only measurements with precise time-stamps.

The MAP estimates of  $p(\mathcal{X}|y_1, y_N, z)$  and  $p(x_k|y_1, y_N, z)$  are shown in Fig. 8 and 9 together with the credible interval for  $p(x_k|y_1, y_N, z)$ . The estimates are computed with the quasi-Newton numerical optimization method [21] on (9), using (10), and (16), respectively, without a guarantee on convergence to the global maximum.

The joint MAP estimates  $\hat{\mathcal{X}}^{JMAP}$  are shown in Fig. 10 and 11 together with the uncertainty of the state estimate given that the estimated uncertain time is correct, *i.e.*, the uncertainty

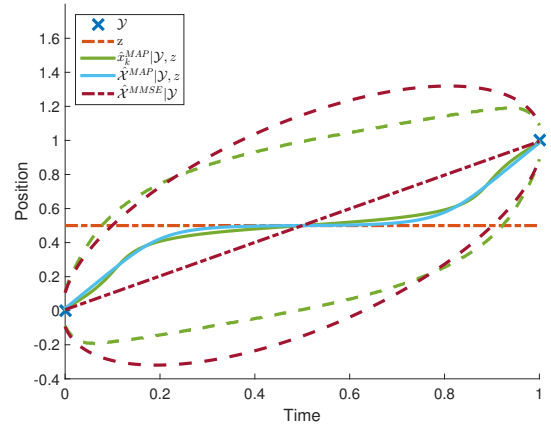


Fig. 8. The MAP estimates of  $p(\mathcal{X}|y_1, y_N, z)$  and  $p(x_k|y_1, y_N, z)$  for the first scenario and the MMSE estimate given only the measurements.

in time is disregarded, resulting in an underestimated credibility interval. The estimated uncertain times  $\hat{\tau}^{JMAP}$  are 0.5 and 0.52, respectively.

## VII. DISCUSSION

The posterior distributions for the scenarios in Sec. VI are examined and discussed in Sec. VII-A. Advantages and disadvantages of the different point estimators are discussed in Sec. VII-B and the resulting estimates for the two scenarios are compared in Sec. VII-C.

### A. Posterior Distributions

Assuming noise-free measurements and observation, and a continuous underlying system, the intermediate value theorem [22] for the first model already gives the information that  $x(\tilde{\tau}) = z$  for some continuous  $\tilde{\tau}$ , so at a first glance the observation does not seem to add any new information. However, since  $\tau$  is random, the state is more likely to be in the vicinity of the observation, explaining the shape of the posterior distribution in Fig. 2. For the second model in Fig. 4 the prior is more informative and the observation is

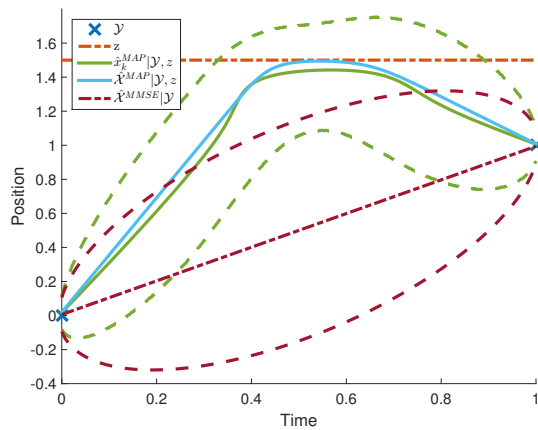


Fig. 9. The MAP estimates of  $p(\mathcal{X}|y_1, y_N, z)$  and  $p(x_k|y_1, y_N, z)$  for the second scenario and the MMSE estimate given only the measurements.

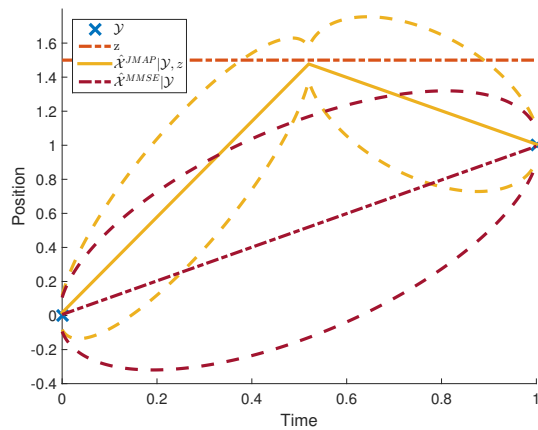


Fig. 11. Joint MAP state estimate of  $p(\mathcal{X}, \tau|y_1, y_N, z)$  for the second scenario and the MMSE estimate given only the measurements.

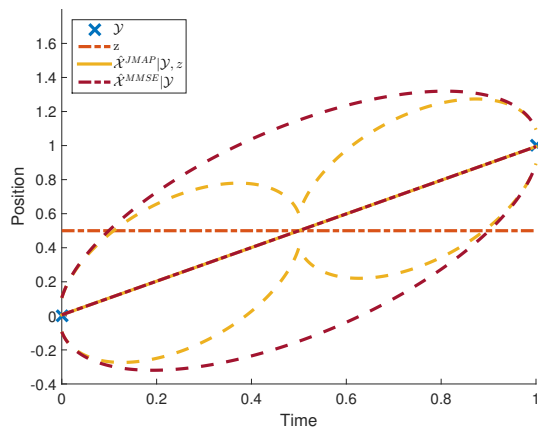


Fig. 10. Joint MAP state estimate of  $p(\mathcal{X}, \tau|y_1, y_N, z)$  for the first scenario and the MMSE estimate given only the measurements.

also not one guaranteed by the intermediate value theorem, thus providing more information than in the first scenario.

In Fig. 7 the distributions do not differ much, but are slightly shifted to account for the observation being closer to the last measurement than the first measurement. Fig. 7 is less intuitive. As can be seen in (23a) the distribution  $\max_{\mathcal{X}} p(\mathcal{X}, \tau|y_1, y_N, z)$  depends on the prior, which is constant in this case, and the Mahalanobis distance between the observation and its prediction, while in (21) it can be seen that  $p(\tau|y_1, y_N, z)$  in addition depends on the determinant of the innovation covariance, in effect penalizing uncertainty in the likelihood of  $z$ . By considering  $\max_{\mathcal{X}} p(\mathcal{X}, \tau|y_1, y_N, z)$ , the reduction of uncertainty in the joint posterior is also taken into account, while only the likelihood of the observation given the measurements is considered in  $p(\tau|y_1, y_N, z)$ .

### B. Comparison of Point Estimators

An advantage of the MMSE estimator is that it is unbiased, however for multimodal distributions a disadvantage is that the estimate might fall between modes, producing an unlikely estimate. The MAP and joint MAP estimators will on the other hand produce likely estimates, but at the cost of being biased.

Often, the reason for using the MAP estimator over the MMSE estimator is that it is cheaper to compute since it avoids an expensive integration over the states, but this is not the case for the Gaussian distribution, as well as for Gaussian mixtures, where this integration is analytical. Both the MAP and MMSE estimators, on the other hand, require a summation over all possible  $\tau$ , for which the complexity increases exponentially with the number of observations with uncertain timestamps. However, this is not a problem for the joint MAP estimator, which lends itself more easily to be computed using efficient methods, *e.g.*, expectation-maximization [23] or stochastic methods [24].

Another disadvantage of the MAP estimator is that it requires numerical optimization methods for Gaussian mixtures. In the case of  $p(\mathcal{X}|y_1, y_N, z)$  the dimension increases with decreased sampling time  $T_s$ , which in general is chosen small for good resolution, and already the simple model considered in Sec. II is difficult. As an alternative, the posterior distribution of each state can be maximized, resulting in multiple low-dimensional optimization problems, but as can be seen in Fig. 8 and 9 and as is argued in [16] the estimators do not produce the same result. The reason for this is that the cross-covariances are ignored. Another issue is that Gaussian mixtures in general are multimodal and it might be difficult to find optimization methods that guarantee finding the global optimum.

### C. Comparison of Point Estimates

The effects of including an observation with uncertain timestamp is most visible for the first scenario. Since the observation coincides with the previous trajectory, *i.e.*, the estimate given only the measurements with precise timestamps, the joint MAP estimator in Fig. 10 will estimate  $\tau$  at the time of coincidence and result in the same trajectory. The uncertainty is underestimated because it disregards the uncertainty in time.

The MMSE estimator in Fig. 3 produces a trajectory that slightly tends towards the observation compared to the previous trajectory. The reason for this behaviour can be found

in Fig. 6 where most weights, given by  $p(\tau|y_1, y_N, z)$ , are seen to be similar except near the measurements with precise timestamps. The trajectory for each weight passes near the observation at its corresponding time, attracting the weighted average of the trajectories in the MMSE estimate towards it. It should also be noted that the uncertainty for the estimate reduces over the previous trajectory.

This behaviour is even more pronounced for the MAP estimators in Fig. 8. For the MAP estimator of  $p(x_k|y_1, y_N, z)$ , this can intuitively be understood by realizing that at each time within the region of similar weights in Fig. 6 the heights of the peaks will be approximately proportional to the inverse of the covariance, and the smallest covariance is given by the Gaussian where the observation was used to update the posterior at the current time. This will attract the trajectory even more towards the observation than the MMSE estimate. Although harder to visualize, due to the high dimension, a similar intuition applies to the MAP estimator for  $p(\mathcal{X}|y_1, y_N, z)$ . The MAP estimate of  $p(\mathcal{X}|y_1, y_N, z)$  is also smoother than the MAP estimate of  $p(x_k|y_1, y_N, z)$ , which can be explained by the discarded cross-covariances in the latter. The credible interval of  $p(x_k|y_1, y_N, z)$  also improves significantly over the previous trajectory.

For the second scenario the estimates are drastically different since the observation does not coincide with the previous trajectory. The other difference compared to the first scenario is that the prior  $p(\tilde{\tau})$  is no longer flat, see Fig. 7, giving much more influence to trajectories corresponding to the peaked region. The joint MAP estimator in Fig. 11 produces the trajectory that corresponds to the maximum of  $p(\mathcal{X}, \tau|y_1, y_N, z)$ , and as before, the MMSE estimate in Fig. 5 is a weighted average of trajectories, thus not reaching all the way to the observation. The MAP estimates  $p(x_k|y_1, y_N, z)$  and  $p(\mathcal{X}|y_1, y_N, z)$  tend even more to the observation, explained with similar intuition as for the first scenario.

## VIII. SUMMARY AND FUTURE WORK

In this paper the problem of an observation with uncertain timestamp has been considered and a simple one-dimensional example has been examined to attain intuition for its effects. Posterior distributions for the states and the uncertain time have been derived as well as some point estimators for a conditionally linear Gaussian state space model.

The implications for several observations with uncertain timestamp need to be investigated further, especially methods for handling the exponentially increasing computational cost. There has been much research into the similar problem of jump Markov linear systems that most likely will apply directly. The similar problems of filtering and prediction would also be interesting to investigate, as well as the extension to multi-dimensional and nonlinear models.

## ACKNOWLEDGEMENT

This work has been supported by funding from Vinnova Industry Excellence Center LINK-SIC, the Swedish strategic

research center Security Link and the Swedish Research Council through the project Scalable Kalman Filters.

## REFERENCES

- [1] M. Hallmén, "Map-aided GPS tracking in urban areas: Application to runner tracking in sprint orienteering," Master's thesis, Linköping University, 2015.
- [2] (2016) Smart savannahs. [Online]. Available: <http://wildlifesecurity.se/smart-savannahs/>
- [3] F. Eng and F. Gustafsson, "Identification with stochastic sampling time jitter," *Automatica*, vol. 44, p. 637–646, Mar. 2008.
- [4] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, vol. 39, no. 1, pp. 1–38, 1977.
- [5] R. H. Shumway and D. S. Stoffer, "An approach to time series smoothing and forecasting using the EM algorithm," *Journal of Time Series Analysis*, vol. 3, no. 4, pp. 253–264, 1982.
- [6] G. Icaza and R. Jones, "A state-space EM algorithm for longitudinal data," *Journal of Time Series Analysis*, vol. 20, no. 5, pp. 537–550, 1999.
- [7] A. Naranjo, A. A. Trindade, and G. Casella, "Extending the state-space model to accommodate missing values in responses and covariates," *Journal of the American Statistical Association*, vol. 108, no. 501, pp. 202–216, 2013.
- [8] C. N. Georghiadis and D. L. Snyder, "The expectation-maximization algorithm for symbol unsynchronized sequence detection," *IEEE Trans. Commun.*, vol. 39, no. 1, pp. 54–61, Jan. 1991.
- [9] Y. Bar-Shalom, "Update with out-of-sequence measurements in tracking: exact solution," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 3, pp. 769–777, Jul 2002.
- [10] K. Zhang, X. R. Li, and Y. Zhu, "Optimal update with out-of-sequence measurements," *IEEE Trans. Signal Process.*, vol. 53, no. 6, pp. 1992–2004, Jun. 2005.
- [11] I. Y. Bar-Itzhack and Y. Vitek, "The enigma of false bias detection in a strapdown system during transfer alignment," *Journal of Guidance, Control, and Dynamics*, vol. 8, no. 2, pp. 175–180, 1985.
- [12] I. Skog and P. Handel, "Effects of time synchronization errors in GNSS-aided INS," in *IEEE/ION Position, Location and Navigation Symposium*, May 2008, pp. 82–88.
- [13] U. Orguner and F. Gustafsson, "Particle filtering with propagation delayed measurements," in *IEEE Aerospace Conference*, Mar. 2010, pp. 1–9.
- [14] K. Granström and U. Orguner, "On the reduction of Gaussian inverse Wishart mixtures," in *15th International Conference on Information Fusion (FUSION)*, Jul. 2012, pp. 2162–2169.
- [15] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA Journal*, vol. 3, no. 8, pp. 1445–1450, Aug. 1965.
- [16] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.
- [17] C. C. Paige and M. A. Saunders, "Least squares estimation of discrete linear dynamic systems using orthogonal transformations," *SIAM J. Numer. Anal.*, vol. 14, pp. 180–193, 1977.
- [18] D. Fraser and J. Potter, "The optimum linear smoother as a combination of two optimum linear filters," *IEEE Trans. Autom. Control*, vol. 14, no. 4, pp. 387–390, Aug 1969.
- [19] J. Ajgl, M. Šimandl, and J. Dunfk, "Millman's formula in data fusion," in *Proceedings of the 10th International PhD Workshop Young Generation Viewpoint*, Hluboka and Vltavou, Czech Republic, Sep. 2009, pp. 1–6.
- [20] A. G. Akritas, E. K. Akritas, and G. I. Malaschonok, "Various proofs of Sylvester's (determinant) identity," *Math. Comput. Simul.*, vol. 42, no. 4-6, pp. 585–593, Jan. 1996.
- [21] W. C. Davidon, "Variable metric method for minimization," Argonne National Laboratory, Argonne, IL, Tech. Rep. ANL-5990, 1959.
- [22] R. A. Adams, *Calculus: A Complete Course*, 5th ed. Pearson Education Canada, 2003.
- [23] A. Logothetis and V. Krishnamurthy, "Expectation maximization algorithms for MAP estimation of jump Markov linear systems," *IEEE Trans. Signal Process.*, vol. 47, no. 8, pp. 2139–2156, Aug. 1999.
- [24] A. Doucet, A. Logothetis, and V. Krishnamurthy, "Stochastic sampling algorithms for state estimation of jump Markov linear systems," *IEEE Trans. Autom. Control*, vol. 45, pp. 200–0, 2000.