LMMSE Filtering in Feedback Systems With White Random Modes: Application to Tracking in Clutter

Daniel Sigalov, Tomer Michaeli, *Member, IEEE*, and Yaakov Oshman, *Fellow, IEEE*

***Abstract*—A generalized state space representation of dynamical systems with random modes switching according to a white random process is presented. The new formulation includes a term, in the dynamics equation, that depends on the most recent linear minimum mean squared error (LMMSE) estimate of the state. This can model the behavior of a feedback control system featuring a state estimator. The measurement equation is allowed to depend on the previous LMMSE estimate of the state, which can represent the fact that measurements are obtained from a validation window centered about the predicted measurement and not from the entire surveillance region. The LMMSE filter is derived for the considered problem. The approach is demonstrated in the context of target tracking in clutter and is shown to be competitive with several popular nonlinear methods.**

***Index Terms*—Clutter and data association, state estimation, target tracking.**

I. Introduction

State estimation in dynamical systems with randomly switching coefficients is an important problem in many applications. Natural examples are maneuvering target tracking and fault detection and isolation algorithms, featured, e.g., in aerospace navigation systems. In the standard modeling the dynamics of the continuously-valued state, and, possibly, its measurement equation, are controlled by a discrete evolving mode. This is the well known concept of hybrid systems [1].

Various problems have been formulated using the hybrid systems framework. In cases involving uncertain observations, such as [2], [3], the mode affects the matrices of the measurement equation. In target tracking applications, considered in, e.g., [4]–[6], the mode usually affects the dynamics equation.

We consider a state space representation of dynamical systems with random coefficients that constitute a white stochastic sequence, accompanied by the following feedback terms. First, we allow the system input to depend on the latest estimate of the state, as is common practice in closed loop control systems. In this work, the state estimate is taken to be the linear minimum mean squared error (LMMSE) estimate.

In addition, the measurement equation is also set to depend on the latest LMMSE state estimate. This can represent the fact that observations are not taken in the entire feasible space, but, rather, in a small validation window set about the predicted measurement of the state.

Manuscript received September 18, 2012; revised June 1, 2013 and December 31, 2013, and January 27, 2014; accepted February 18, 2014. Date of publication February 26, 2014; date of current version August 20, 2014. This research was supported by the Israeli Science Foundation under grant No. 1608/11. Recommended by Associate Editor A. Chiuso.

D. Sigalov is with the Department of Mathematics, Technion-Israel Institute of Technology, 32000 Haifa, Israel (e-mail: sigaldan@gmail.com).

T. Michaeli is with the Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, 7610001 Rehovot, Israel (e-mail: tomer.michaeli@weizmann.ac.il).

Y. Oshman is with the Department of Aerospace Engineering, Technion-Israel Institute of Technology, 32000 Haifa, Israel (e-mail: yaakov.oshman@technion.ac.il).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2014.2308601

It is well known [5] that, even for the case of independently switching modes, the optimal estimate of the state cannot be obtained without resorting to exhaustive enumeration. Therefore, significant efforts have been dedicated to developing suboptimal approaches for state estimation in hybrid systems and, especially, for the important subclass of jump linear systems (JLS). The most popular nonlinear methods include the generalized pseudo-Bayesian (GPB) filter [5] and the interacting multiple model (IMM) algorithm [6].

Alternatively, one may consider optimality within the narrower family of linear filters. Among these we mention [2] and [3] that considered estimation with uncertain observations, [7] that derived a Kalman filter-like (KF) algorithm for a JLS with independently switching modes and uncorrelated matrices within each time step, and [8] that derived an LMMSE scheme for a Markov JLS by means of state augmentation. In addition, in some cases, the state may be estimated optimally within the family of filters which are linear in some of the measurements and nonlinear in the rest, as was shown in [9].

In this paper we concentrate on feedback JLS with independent mode transitions and consider optimal estimation within the family of linear filters.

We derive an LMMSE algorithm that may be conveniently implemented in a recursive form, eliminating the need for unbounded memory. Unlike [7], we do not assume that the matrices within each time step are uncorrelated. This allows tackling a wider variety of problems, such as tracking in clutter, which cannot be modeled directly within the framework of [7]. On the other hand, since we still treat the easier case of independent, rather than Markov, mode transitions, we do not require state augmentation, as does the algorithm of [8]. Our filter reduces to several previously reported results when the parameters of the underlying problem are appropriately adjusted. As an illustration, we formulate the problem of target tracking in clutter within the proposed framework and show that the resulting filter is competitive with several classical nonlinear methods.

The remainder of the paper is organized as follows. In Section II we describe the proposed modeling and survey some related work. The recursive LMMSE algorithm is derived in Section III. An application to target tracking in clutter, followed by a numerical study, is presented in Section IV. Concluding remarks are given in Section V.

II. System Model and Related Work

We consider the dynamical system

$$x_{k+1} = A_k x_k + B_k u_k + C_k w_k \tag{1a}$$

$$y_k = H_k x_k + G_k v_k + F_k \hat{x}_{k-1}, \tag{1b}$$

where $x_k \in \mathbb{R}^n$ and $y_k \in \mathbb{R}^m$ are, respectively, the state and measurement vectors at time k. The processes $\{w_k\}$ and $\{v_k\}$ constitute zero-mean unity-covariance strictly white sequences, and x_0 is a random vector (RV) with mean $\bar{x}_0$ and second-order moment P_0.

We consider two variants for the modeling of u_k. In the first case, u_k is a known deterministic input. However, because in some cases u_k serves as a closed loop control signal, it is common practice to let it depend on the most recent estimate of the state. Thus, in the second variant we set $u_k = \hat{x}_k$, where $\hat{x}_k$ is the LMMSE estimate of x_k using the measurement history $\mathcal{Y}_k \triangleq \{y_1, \ldots, y_k\}$.

Likewise, the term $\hat{x}_{k-1}$ in the measurement equation is the LMMSE estimate of x_{k-1} based on the measurement history $\mathcal{Y}_{k-1}$. Affecting the measurement at time k, the term $F_k \hat{x}_{k-1}$ can be used to represent the fact that observations are not taken in the entire space, but, rather, in a small validation window, set about the predicted measurement.

0018-9286 © 2014 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

The system mode, $\mathcal{M}_k \triangleq \{A_k, B_k, C_k, H_k, G_k, F_k\}$, is a strictly white random process with known distribution. The quantities $\{w_k\}$, $\{v_k\}$, $\{\mathcal{M}_k\}$, and x_0 are assumed to be independent.

We seek to obtain the LMMSE estimate $\hat{x}_{k+1}$ using the measurements $\mathcal{Y}_{k+1}$. It will be shown in the sequel that, in our setting, $\hat{x}_{k+1}$ conveniently possesses the recursive form

$$\hat{x}_{k+1} = L_k\hat{x}_k + K_k y_{k+1} + J_k u_k \tag{2}$$

thus avoiding the need to store the entire measurement sequence. When $u_k = \hat{x}_k$, the terms $L_k\hat{x}_k$ and $J_k\hat{x}_k$ in (2) may be grouped together.

Note that the described problem does not require the system mode to assume values in a discrete domain as opposed to, e.g., [2], [3], [8]. In addition, the above formulation allows evolution not only of the entries of the mode matrices, but also of their dimensions [10]. This observation allows treatment of problems that, to the best of our knowledge, have not been previously considered in the context of LMMSE algorithms. One such example is given in Section IV.

For the setting without feedback terms, several variants and special cases of the presented problem have been considered in the past. Independent measurement faults were treated, in the LMMSE sense, in [2]. De Koning [7] considered a more general case of independently switching modes where, however, the mode elements are assumed uncorrelated, and Costa [8] developed, by means of state augmentation, a recursive LMMSE filter for systems with discrete modes obeying Markov dynamics. Additional contributions include [3], that considered correlated faults, [11], that allowed correlations between subsequent fault variables, and [4], that proposed an LMMSE filter for the static multiple model problem [12]. Related nonlinear solutions were proposed in [5], [6], [13] and references therein.

Besides the novel introduction of the feedback terms, this paper contains several additional contributions. First, we derive a recursive LMMSE algorithm without assuming uncorrelatedness of the mode elements, as done in [7]. This assumption precludes the utilization of the algorithm of [7] even for the simple problem of uncertain observations where measurement noise has a higher variance when faults occur, not to mention more involved settings, such as tracking in clutter. In addition, our algorithm is derived without state augmentation and without assuming discrete modes, as done in [8]. Finally, the approach allows a broader class of problem to be formulated within a single state-space model. Specifically, the new feedback terms allow the application of the idea to the problem of tracking in clutter.

III. Linear Optimal Recursive Estimation

We begin the derivation with deterministic u_k. The stochastic case is treated in Section III-E.

Let Y_k be the RV obtained by concatenating the elements of $\mathcal{Y}_k$. We derive the result using the following lemma, which follows from [14, p. 190] and the linearity of the MMSE estimator in the Gaussian case.

Lemma: Let x, y and z be RVs and let $\hat{x}(z)$ and $\hat{x}(y,z)$ denote, respectively, the LMMSE estimates of x using z, and using both y and z. Let $\hat{y}(z)$ be the LMMSE estimate of y using z. Then

$$\hat{x}(y,z) = \hat{x}(z) + \Gamma_{x\tilde{y}}\Gamma_{\tilde{y}\tilde{y}}^{-1}\tilde{y}, \tag{3}$$

where $\tilde{y} = y - \hat{y}(z)$ and Γ_{ab} is the cross-covariance matrix between the RVs a and b.

Letting $z \triangleq Y_k$, $y \triangleq y_{k+1}$ and using the lemma, the LMMSE estimate of x_{k+1} using $\mathcal{Y}_{k+1}$ is

$$\hat{x}_{k+1} = \hat{x}_{k+1}^{-} + \Gamma_{x_{k+1}\tilde{y}_{k+1}}\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}^{-1}\tilde{y}_{k+1} \tag{4}$$

where $\hat{x}_{k+1}^{-}$ is the LMMSE estimate of x_{k+1} using $\mathcal{Y}_k$, $\tilde{y}_{k+1} \triangleq y_{k+1} - \hat{y}_{k+1}^{-}$, and $\hat{y}_{k+1}^{-}$ is the LMMSE estimate of y_{k+1} using $\mathcal{Y}_k$. If $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}$ is singular, (4) still holds with the inverse replaced by the Moore-Penrose pseudo-inverse. It is easily verified that

$$\hat{x}_{k+1}^{-} = \mathbb{E}[A_k]\hat{x}_k + \mathbb{E}[B_k]u_k \tag{5}$$

$$\begin{aligned}\hat{y}_{k+1}^{-} &= \mathbb{E}[H_{k+1}]\hat{x}_{k+1}^{-} + \mathbb{E}[F_{k+1}]\hat{x}_k\\ &= \left(\mathbb{E}[H_{k+1}]\mathbb{E}[A_k] + \mathbb{E}[F_{k+1}]\right)\hat{x}_k + \mathbb{E}[H_{k+1}]\mathbb{E}[B_k]u_k.\end{aligned} \tag{6}$$

Plugging (5) in (4) we identify the desired matrix coefficients K_k, L_k, and J_k of (2) as follows:

$$K_k = \Gamma_{x_{k+1}\tilde{y}_{k+1}}\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}^{-1} \tag{7}$$

$$L_k = \left(I - K_k\mathbb{E}[H_{k+1}]\right)\mathbb{E}[A_k] - K_k\mathbb{E}[F_{k+1}] \tag{8}$$

$$J_k = \left(I - K_k\mathbb{E}[H_{k+1}]\right)\mathbb{E}[B_k]. \tag{9}$$

We now compute the covariance terms $\Gamma_{x_{k+1}\tilde{y}_{k+1}}$ and $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}$.

A. Computation of $\Gamma_{x_{k+1}\tilde{y}_{k+1}}$

Since $\hat{y}_{k+1}^{-}$ is unbiased, and using (1b) and (6)

$$\begin{aligned}\Gamma_{x_{k+1}\tilde{y}_{k+1}} &= \mathbb{E}\left[x_{k+1}\left(y_{k+1} - \hat{y}_{k+1}^{-}\right)^\top\right]\\ &= \mathbb{E}\left[x_{k+1}(H_{k+1}x_{k+1} + G_{k+1}v_{k+1} + F_{k+1}\hat{x}_k)^\top\right]\\ &\quad - \mathbb{E}\left[x_{k+1}\left(\left(\mathbb{E}[H_{k+1}]\mathbb{E}[A_k] + \mathbb{E}[F_{k+1}]\right)\hat{x}_k\right)^\top\right]\\ &\quad - \mathbb{E}\left[x_{k+1}\left(\mathbb{E}[H_{k+1}]\mathbb{E}[B_k]u_k\right)^\top\right].\end{aligned} \tag{10}$$

Using the independence of x_{k+1} and v_{k+1}, and canceling out identical terms, (10) becomes

$$\begin{aligned}\Gamma_{x_{k+1}\tilde{y}_{k+1}} =& \mathbb{E}\left[x_{k+1}x_{k+1}^\top\right]\mathbb{E}\left[H_{k+1}^\top\right] - \mathbb{E}\left[x_{k+1}\hat{x}_k^\top\right]\mathbb{E}\left[A_k^\top\right]\mathbb{E}\left[H_{k+1}^\top\right]\\ &- \mathbb{E}[x_{k+1}]u_k^\top\mathbb{E}\left[B_k^\top\right]\mathbb{E}\left[H_{k+1}^\top\right].\end{aligned} \tag{11}$$

Before proceeding, we define $\Sigma_k \triangleq \mathbb{E}[x_k x_k^\top]$, $\Delta_k \triangleq u_k u_k^\top$ and, in addition

$$\Lambda_k \triangleq \mathbb{E}\left[\hat{x}_k\hat{x}_k^\top\right] = \mathbb{E}\left[\hat{x}_k x_k^\top\right] \tag{12}$$

$$\Upsilon_k \triangleq \mathbb{E}[x_k]u_k^\top = \mathbb{E}[\hat{x}_k]u_k^\top \tag{13}$$

where the RHS of (12) and (13) follow from the orthogonality principle and from the unbiasedness of $\hat{x}_k$, respectively. Note that Σ_k, Λ_k, and Δ_k are symmetric.

Using the independence of $\hat{x}_k$ and w_k

$$\begin{aligned}\mathbb{E}\left[x_{k+1}\hat{x}_k^\top\right] &= \mathbb{E}\left[(A_k x_k + B_k u_k + C_k w_k)\hat{x}_k^\top\right]\\ &= \mathbb{E}[A_k]\Lambda_k + \mathbb{E}[B_k]\Upsilon_k^\top\end{aligned} \tag{14}$$

which yields for (11)

$$\begin{aligned}\Gamma_{x_{k+1}\tilde{y}_{k+1}} =& \Big(\Sigma_{k+1} - \left(\mathbb{E}[A_k]\Lambda_k + \mathbb{E}[B_k]\Upsilon_k^\top\right)\mathbb{E}\left[A_k^\top\right]\\ &- \mathbb{E}[x_{k+1}]u_k^\top\mathbb{E}\left[B_k^\top\right]\Big)\mathbb{E}\left[H_{k+1}^\top\right].\end{aligned} \tag{15}$$

From (1a), we have

$$\begin{aligned}\mathbb{E}[x_{k+1}] &= \mathbb{E}[A_k x_k + B_k u_k + C_k w_k]\\ &= \mathbb{E}[A_k]\mathbb{E}[x_k] + \mathbb{E}[B_k]u_k\end{aligned} \tag{16}$$

which, when substituted in (15), leads to

$$\Gamma_{x_{k+1}\tilde{y}_{k+1}} = \Big(\Sigma_{k+1} - \Big(\mathbb{E}[A_k]\left(\Lambda_k \mathbb{E}\left[A_k^\top\right] + \Upsilon_k \mathbb{E}\left[B_k^\top\right]\right) + \mathbb{E}[B_k]\left(\Upsilon_k^\top \mathbb{E}\left[A_k^\top\right] + \Delta_k \mathbb{E}\left[B_k^\top\right]\right)\Big)\Big)\mathbb{E}\left[H_{k+1}^\top\right]. \quad (17)$$

B. Computation of $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}$

Since $\hat{y}_{k+1}^-$ is the LMMSE estimate of y_{k+1} using $\mathcal{Y}_k$, $\tilde{y}_{k+1}$ is orthogonal to $\hat{y}_{k+1}^-$ and, using (6)

$$\begin{aligned}\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}} &= \mathbb{E}\left[\left(y_{k+1} - \hat{y}_{k+1}^-\right) y_{k+1}^\top\right] \\ &= \mathbb{E}\left[y_{k+1}y_{k+1}^\top\right] - \mathbb{E}\left[\hat{y}_{k+1}^- y_{k+1}^\top\right] \\ &= \mathbb{E}\left[y_{k+1}y_{k+1}^\top\right] - \left(\mathbb{E}[H_{k+1}]\mathbb{E}[A_k] + \mathbb{E}[F_{k+1}]\right)\mathbb{E}\left[\hat{x}_k y_{k+1}^\top\right] \\ &\quad - \mathbb{E}[H_{k+1}]\mathbb{E}[B_k]u_k\mathbb{E}\left[y_{k+1}^\top\right]. \end{aligned} \quad (18)$$

Using (1b) and the independence of $\{\hat{x}_k, x_{k+1}\}$, $\{H_{k+1}, G_{k+1}, F_{k+1}\}$ and v_{k+1}, we have

$$\begin{aligned}\mathbb{E}\left[\hat{x}_k y_{k+1}^\top\right] &= \mathbb{E}\left[\hat{x}_k (H_{k+1}x_{k+1} + F_{k+1}\hat{x}_k)^\top\right] \\ &= \mathbb{E}\left[\hat{x}_k x_{k+1}^\top\right]\mathbb{E}\left[H_{k+1}^\top\right] + \Lambda_k \mathbb{E}\left[F_{k+1}^\top\right] \end{aligned} \quad (19)$$

which, using (14), becomes

$$\mathbb{E}\left[\hat{x}_k y_{k+1}^\top\right] = \Lambda_k \left(\mathbb{E}\left[A_k^\top\right]\mathbb{E}\left[H_{k+1}^\top\right] + \mathbb{E}\left[F_{k+1}^\top\right]\right) + \Upsilon_k \mathbb{E}\left[B_k^\top\right]\mathbb{E}\left[H_{k+1}^\top\right]. \quad (20)$$

Due to the independence of $\{x_{k+1}, \hat{x}_k\}$, v_{k+1}, and $\{H_{k+1}, G_{k+1}\}$

$$\begin{aligned}\mathbb{E}\left[y_{k+1}y_{k+1}^\top\right] &= \mathbb{E}\left[H_{k+1}x_{k+1}x_{k+1}^\top H_{k+1}^\top\right] + \mathbb{E}\left[G_{k+1}v_{k+1}v_{k+1}^\top G_{k+1}^\top\right] \\ &\quad + \mathbb{E}\left[F_{k+1}\hat{x}_k\hat{x}_k^\top F_{k+1}^\top\right] + \mathbb{E}\left[H_{k+1}x_{k+1}\hat{x}_k^\top F_{k+1}^\top\right] \\ &\quad + \mathbb{E}\left[F_{k+1}\hat{x}_k x_{k+1}^\top H_{k+1}^\top\right]. \end{aligned} \quad (21)$$

Consider the last summand. From the smoothing property of the conditional expectation

$$\begin{aligned}\mathbb{E}\left[F_{k+1}\hat{x}_k x_{k+1}^\top H_{k+1}^\top\right] &= \mathbb{E}\left[\mathbb{E}\left[F_{k+1}\hat{x}_k x_{k+1}^\top H_{k+1}^\top \mid F_{k+1}, H_{k+1}\right]\right] \\ &= \mathbb{E}\left[F_{k+1}\mathbb{E}\left[\hat{x}_k x_{k+1}^\top\right] H_{k+1}^\top\right] \end{aligned} \quad (22)$$

where we utilized the independence of $\{H_{k+1}, F_{k+1}\}$ and $\{x_{k+1}, \hat{x}_k\}$.

Similarly, since $\mathbb{E}[x_{k+1}x_{k+1}^\top] = \Sigma_{k+1}$, $\mathbb{E}[v_{k+1}v_{k+1}^\top] = I$, and $\mathbb{E}[\hat{x}_k\hat{x}_k^\top] = \Lambda_k$, we obtain

$$\mathbb{E}\left[H_{k+1}x_{k+1}x_{k+1}^\top H_{k+1}^\top\right] = \mathbb{E}\left[H_{k+1}\Sigma_{k+1}H_{k+1}^\top\right] \quad (23)$$

$$\mathbb{E}\left[G_{k+1}v_{k+1}v_{k+1}^\top G_{k+1}^\top\right] = \mathbb{E}\left[G_{k+1}G_{k+1}^\top\right] \quad (24)$$

$$\mathbb{E}\left[F_{k+1}\hat{x}_k\hat{x}_k^\top F_{k+1}^\top\right] = \mathbb{E}\left[F_{k+1}\Lambda_k F_{k+1}^\top\right]. \quad (25)$$

For future reference, we also note that

$$\mathbb{E}\left[A_k x_k x_k^\top A_k^\top\right] = \mathbb{E}\left[A_k \Sigma_k A_k^\top\right] \quad (26)$$

$$\mathbb{E}\left[A_k x_k u_k^\top B_k^\top\right] = \mathbb{E}\left[A_k \Upsilon_k B_k^\top\right] \quad (27)$$

$$\mathbb{E}\left[B_k u_k u_k^\top B_k^\top\right] = \mathbb{E}\left[B_k \Delta_k B_k^\top\right] \quad (28)$$

$$\mathbb{E}\left[C_k w_k w_k^\top C_k^\top\right] = \mathbb{E}\left[C_k C_k^\top\right]. \quad (29)$$

Substituting (14) in (22), and using (22)–(25) in (21),

$$\begin{aligned}\mathbb{E}\left[y_{k+1}y_{k+1}^\top\right] &= \mathbb{E}\left[H_{k+1}\Sigma_{k+1}H_{k+1}^\top\right] + \mathbb{E}\left[G_{k+1}G_{k+1}^\top\right] \\ &\quad + \mathbb{E}\left[F_{k+1}\Lambda_k F_{k+1}^\top\right] \\ &\quad + \mathbb{E}\left[H_{k+1}\left(\mathbb{E}[A_k]\Lambda_k + \mathbb{E}[B_k]\Upsilon_k^\top\right)F_{k+1}^\top\right] \\ &\quad + \mathbb{E}\left[F_{k+1}\left(\Lambda_k\mathbb{E}\left[A_k^\top\right] + \Upsilon_k\mathbb{E}\left[B_k^\top\right]\right)H_{k+1}^\top\right]. \end{aligned} \quad (30)$$

In addition, we obtain, in a straightforward manner

$$\mathbb{E}[y_{k+1}] = \left(\mathbb{E}[H_{k+1}]\mathbb{E}[A_k] + \mathbb{E}[F_{k+1}]\right)\mathbb{E}[x_k] + \mathbb{E}[H_{k+1}]\mathbb{E}[B_k]u_k. \quad (31)$$

Using (23), (24), and (25) in (30), and substituting (20), (30), and (31) in (18), we finally obtain

$$\begin{aligned}\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}} &= \mathbb{E}\left[H_{k+1}\Sigma_{k+1}H_{k+1}^\top\right] + \mathbb{E}\left[G_{k+1}G_{k+1}^\top\right] \\ &\quad + \mathbb{E}\left[F_{k+1}\Lambda_k F_{k+1}^\top\right] - \mathbb{E}[F_{k+1}]\Lambda_k\mathbb{E}[F_{k+1}^\top] \\ &\quad - \mathbb{E}[H_{k+1}]\mathbb{E}[A_k]\Lambda_k\mathbb{E}\left[A_k^\top\right]\mathbb{E}\left[H_{k+1}^\top\right] \\ &\quad + \mathbb{E}\left[H_{k+1}\left(\mathbb{E}[A_k]\Lambda_k + \mathbb{E}[B_k]\Upsilon_k^\top\right)F_{k+1}^\top\right] \\ &\quad + \mathbb{E}\left[F_{k+1}\left(\Lambda_k\mathbb{E}\left[A_k^\top\right] + \Upsilon_k\mathbb{E}\left[B_k^\top\right]\right)H_{k+1}^\top\right] \\ &\quad - \mathbb{E}[H_{k+1}]\mathbb{E}[A_k]\Lambda_k\mathbb{E}\left[F_{k+1}^\top\right] \\ &\quad - \mathbb{E}[F_{k+1}]\Lambda_k\mathbb{E}\left[A_k^\top\right]\mathbb{E}\left[H_{k+1}^\top\right] \\ &\quad - \mathbb{E}[H_{k+1}]\mathbb{E}[A_k]\Upsilon_k\mathbb{E}\left[B_k^\top\right]\mathbb{E}\left[H_{k+1}^\top\right] \\ &\quad - \mathbb{E}[F_{k+1}]\Upsilon_k\mathbb{E}\left[B_k^\top\right]\mathbb{E}\left[H_{k+1}^\top\right] \\ &\quad - \mathbb{E}[H_{k+1}]\mathbb{E}[B_k]u_k\mathbb{E}\left[y_{k+1}^\top\right]. \end{aligned} \quad (32)$$

Notice, that a sufficient condition for the nonsingularity of $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}$ is $\mathbb{E}[G_{k+1}G_{k+1}^\top] \succ 0$. To see this, recall that, by definition, $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}$ is positive semi-definite for any choice of $\mathbb{E}[G_{k+1}G_{k+1}^\top]$ and, in particular, for $G_{k+1} = 0$. But this means that the matrix on the RHS of (32) minus $\mathbb{E}[G_{k+1}G_{k+1}^\top]$ is positive semi-definite, rendering $\mathbb{E}[G_{k+1}G_{k+1}^\top] \succ 0$ a sufficient condition for the non-singularity of $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}$.

C. Computation of the Second-Order Moments

Utilizing the independence of x_k, w_k and $\{A_k, B_k, C_k\}$, and (26)–(29), Σ_{k+1} is given by

$$\begin{aligned}\Sigma_{k+1} &= \mathbb{E}\left[x_{k+1}x_{k+1}^\top\right] \\ &= \mathbb{E}\left[(A_kx_k + B_ku_k + C_kw_k)(A_kx_k + B_ku_k + C_kw_k)^\top\right] \\ &= \mathbb{E}\left[A_k\Sigma_kA_k^\top\right] + \mathbb{E}\left[A_k\Upsilon_kB_k^\top\right] + \mathbb{E}\left[B_k\Upsilon_k^\top A_k^\top\right] \\ &\quad + \mathbb{E}\left[B_k\Delta_kB_k^\top\right] + \mathbb{E}\left[C_kC_k^\top\right]. \end{aligned} \quad (33)$$

Next, consider Λ_{k+1}. Direct computation yields:

$$\begin{aligned}\Lambda_{k+1} &= \mathbb{E}\left[\hat{x}_{k+1}x_{k+1}^\top\right] \\ &= \left(L_k + K_k\mathbb{E}[F_{k+1}]\right)\mathbb{E}\left[\hat{x}_k x_{k+1}^\top\right] \\ &\quad + K_k\mathbb{E}[H_{k+1}]\Sigma_{k+1} + J_ku_k\mathbb{E}\left[x_{k+1}^\top\right]. \end{aligned} \quad (34)$$

Using (14), the latter becomes

$$\Lambda_{k+1} = \left(L_k + K_k\mathbb{E}[F_{k+1}]\right)\left(\Lambda_k\mathbb{E}\left[A_k^\top\right] + \Upsilon_k\mathbb{E}\left[B_k^\top\right]\right) + J_k\left(\Upsilon_k^\top\mathbb{E}\left[A_k^\top\right] + \Delta_k\mathbb{E}\left[B_k^\top\right]\right) + K_k\mathbb{E}[H_{k+1}]\Sigma_{k+1}. \quad (35)$$

Finally, $\Upsilon_{k+1} = \mathbb{E}[x_{k+1}]u_{k+1}^\top$. Note that Δ_k is known for all k.

D. Algorithm Summary

a) Initialization: $\hat{x}_0 = \bar{x}_0$, $\Sigma_0 = P_0 + \bar{x}_0\bar{x}_0^\top$, $\Lambda_0 = \bar{x}_0\bar{x}_0^\top$, $\Upsilon_0 = \bar{x}_0 u_0^\top$, $\Delta_0 = u_0 u_0^\top$.

b) Recursion: For $k = 1, 2, \ldots$ perform the routine of Alg. 1.

Algorithm 1

Input: y_{k+1}, u_{k+1}, $\hat{x}_k$, $\mathbb{E}[x_k]$, Σ_k, Λ_k, Υ_k, Δ_k

1: Compute $\mathbb{E}[A_k]$, $\mathbb{E}[B_k]$, $\mathbb{E}[C_k C_k^\top]$, $\mathbb{E}[A_k \Sigma_k A_k^\top]$, $\mathbb{E}[A_k \Upsilon_k B_k^\top]$, and $\mathbb{E}[B_k \Delta_k B_k^\top]$.

2: Compute $\mathbb{E}[x_{k+1}]$ and Σ_{k+1} using Eqs. (16) and (33).

3: Compute $\mathbb{E}[H_{k+1}]$, $\mathbb{E}[G_{k+1}G_{k+1}^\top]$, $\mathbb{E}[F_{k+1}]$, $\mathbb{E}[F_{k+1}\Lambda_k F_{k+1}^\top]$, $\mathbb{E}[H_{k+1}\Sigma_{k+1}H_{k+1}^\top]$, and $\mathbb{E}[H_{k+1}(\mathbb{E}[A_k]\Lambda_k + \mathbb{E}[B_k]\Upsilon_k^\top)F_{k+1}^\top]$.

4: Compute $\Gamma_{x_{k+1}\tilde{y}_{k+1}}$ and $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}$ using Eqs. (17) and (32).

5: Compute K_k, L_k, and J_k using Eqs. (7), (8), and (9), and $\hat{x}_{k+1}$ using Eq. (2).

6: Compute Λ_{k+1} using Eq. (35) and Υ_{k+1} by plugging $\mathbb{E}[x_{k+1}]$ into (13).

Output: $\hat{x}_{k+1}$, $\mathbb{E}[x_{k+1}]$, Σ_{k+1}, Λ_{k+1}, Υ_{k+1}

Since the distribution of $\mathcal{M}_k$ is known, the expectations of steps 1 and 3 of Alg. 1 may be calculated by, e.g., direct summations in case of discrete modes. In some cases, as demonstrated in Section IV, closed form expressions exist for the above expectations.

We note that the standard KF for a system with no inputs should be obtained when $\{\mathcal{M}_k\}$ is a deterministic sequence with $B_k = 0$, $F_k = 0$. In this setting we have

$$\Gamma_{x_{k+1}\tilde{y}_{k+1}} = \left(\Sigma_{k+1} - A_k\Lambda_k A_k^\top\right) H_{k+1}^\top$$

and

$$\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}} = H_{k+1}\left(\Sigma_{k+1} - A_k\Lambda_k A_k^\top\right) H_{k+1}^\top + G_{k+1}G_{k+1}^\top.$$

Substituting these in (4) we indeed obtain the standard KF in the form where the time and measurement updates are combined together. The error covariances follow in a similar manner.

E. Random Inputs

In the second variant of (1a), in which $u_k = \hat{x}_k$, it turns out that the roles played by A_k and B_k are identical. Specifically, after replacing u_k with $\hat{x}_k$, at each step of the derivation of Section III, A_k and B_k are multiplied by the same quantities. Thus, the filter for the modified problem is obtained from the one described in Alg. 1 by replacing A_k with $A_k + B_k$ and nullifying u_k and Υ_k. An alternative derivation, based on the orthogonality principle, may be found in [15].

IV. Application to Target Tracking in Clutter

In this section we demonstrate the proposed concept by casting the classical problem of tracking in clutter within our formulation, and applying the LMMSE filter of Section III.

A. System and Clutter Models

Consider a single target obeying a linear model. Setting $A_k = A$, $B_k = 0$, and $C_k = C$ in (1a)

$$x_{k+1} = Ax_k + Cw_k. \tag{36}$$

Here A and C are deterministic matrices, accounting for the state dynamics and process noise covariance, respectively, and $\{w_k\}$ is a scalar process noise sequence. The target state is observed via the the equation

$$y_k^{\mathrm{true}} = H_{\mathrm{nom}}x_k + G_{\mathrm{nom}}v_k^{\mathrm{true}} \tag{37}$$

where v_k^{true} represents measurement noise. In addition, at each time, a number of clutter detections are obtained. These will be denoted as $\{y_{k,i}^{\mathrm{cl}}\}_{i=1}^{N-1}$, where N is the total number of detections. Clutter measurements do not carry any information about the target of interest. They are, however, indistinguishable from true detections in the sense that they carry information of the same type (say, position). At each time, the clutter measurements are assumed to be independent of each other, of the clutter measurements at other times, and of the true state and observation. In addition, we assume that they are uniformly distributed in space.

To correctly model the distribution of the clutter detections, we note that, typically, at each scan, the sensor initiates a validation window centered about the next predicted target detection, and the algorithm processes only those measurements obtained within the window. Since the clutter detections are uniformly distributed in space, they are also uniformly distributed within the validation window.

We define the measurement vector y_k to be the concatenation of all measurements from time k, $N-1$ of which correspond to clutter, and one originating from the true target. The location of the true measurement within this concatenated vector is, of course, unknown to the algorithm.

This setting can be modeled using (1b) by letting the mode $\mathcal{M}_k$ be distributed as

$$\mathcal{M}_k = \{H_k, G_k, F_k\}$$
$$= \begin{cases} \left\{\begin{pmatrix} H_{\mathrm{nom}} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathrm{diag}\begin{pmatrix} G_{\mathrm{nom}} \\ G_{\mathrm{cl}} \\ \vdots \\ G_{\mathrm{cl}} \end{pmatrix}, \begin{pmatrix} 0 \\ H_{\mathrm{nom}}A \\ \vdots \\ H_{\mathrm{nom}}A \end{pmatrix}\right\}, & \text{w.p. } \frac{1}{N} \\ \quad\vdots \qquad\qquad\quad \vdots \qquad\qquad\quad \vdots & \\ \left\{\begin{pmatrix} 0 \\ \vdots \\ 0 \\ H_{\mathrm{nom}} \end{pmatrix}, \mathrm{diag}\begin{pmatrix} G_{\mathrm{cl}} \\ \vdots \\ G_{\mathrm{cl}} \\ G_{\mathrm{nom}} \end{pmatrix}, \begin{pmatrix} H_{\mathrm{nom}}A \\ \vdots \\ H_{\mathrm{nom}}A \\ 0 \end{pmatrix}\right\}, & \text{w.p. } \frac{1}{N} \end{cases} \tag{38}$$

where G_{cl} is the square-root of the covariance matrix associated with the clutter.

For example, the first realization of $\{H_k, G_k, F_k\}$ in (38) corresponds to the scenario in which the first of the N observations is the true target measurement, y_k^{true}, generated according to (37), while the other $N-1$ measurements are clutter, each of which is generated according to

$$y_{k,i}^{\mathrm{cl}} = H_{\mathrm{nom}}A\hat{x}_{k-1} + G_{\mathrm{cl}}v_{k,i}^{\mathrm{cl}}, \quad i = 2, \ldots, N. \tag{39}$$

Here, $H_{\mathrm{nom}}A\hat{x}_{k-1}$ is the predicted true measurement at time k, which is also the center of the validation window, so that clutter measurements at time k are uniformly distributed around this quantity. Namely, $v_{k,i}^{\mathrm{cl}}$ has a uniform distribution. The overall number of measurements in the validation window, N, is assumed to be known, but may vary in time. Thus, the dimensions of H_k, G_k, and F_k may depend on k.

It is readily observed that the matrices $\{H_k, G_k, F_k\}$ are correlated in this setting. This renders the approach of [7] inapplicable in the current scenario. Furthermore, it can be seen that without the feedback term in the measurement equation, it is impossible to account for the fact that clutter is uniformly distributed in a window centered about the predicted measurement. In fact, any linear method disregarding this term, such as [7], [8], must assume that clutter measurements are distributed about 0.

Notice that we assumed, for simplicity, that the true measurement is always present in the validation window. To account for the possibility that the true measurement does not fall in the validation window, the option

$$\{H_k, G_k, F_k\} = \{\mathbf{0}, I_N \otimes G_{\mathrm{cl}}, \mathbf{1}_N \otimes H_{\mathrm{nom}} A\} \tag{40}$$

needs to be added to the set of possible realizations in (38). Here, $\otimes$ stands for the Kronecker product, $\mathbf{1}_N$ is an $N \times 1$ vector comprising all ones, and I_N is the $N \times N$ identity matrix. The probability of this outcome is $(1 - P_D)(1 - P_G)$ where P_D is the probabilty of target detection, assumed known, and P_G is the probability that, upon target detection, the true measurement falls in the validation window. This parameter is defined by the user and, typically, affects the window size as discussed in the sequel. Note that, when no measurements are available, $N = 0$, and (2) becomes (in the absence of u_k) $\hat{x}_{k+1} = L_k \hat{x}_k$, which corresponds to a simple prediction (time update) without consecutive measurement update, as expected.

B. Matrix Computations

To invoke the algorithm presented in Section III we need to compute the expectations of Steps 1 and 3 of Alg. 1. Although these may be evaluated numerically, via direct summations, in the present example closed-form expressions exist, as we show next for the simple setting in which the true measurement is always present in the validation window (extensions are straightforward.)

As the matrices of the dynamics equation are deterministic, $\mathbb{E}[A_k] = A$, $\mathbb{E}[B_k] = 0$, $\mathbb{E}[C_k C_k^\top] = CC^\top$, $\mathbb{E}[A_k \Upsilon_k B_k^\top] = 0$, $\mathbb{E}[B_k \Delta_k B_k^\top] = 0$, and $\mathbb{E}[A_k \Sigma_k A_k^\top] = A\Sigma_k A^\top$. Also, according to the distribution defined in (38),

$$\mathbb{E}[H_{k+1}] = \frac{1}{N} \mathbf{1}_N \otimes H_{\mathrm{nom}} \tag{41}$$

$$\mathbb{E}[F_{k+1}] = \frac{N-1}{N} \mathbf{1}_N \otimes H_{\mathrm{nom}} A. \tag{42}$$

The remaining terms read

$$\mathbb{E}\left[H_{k+1} \Sigma_{k+1} H_{k+1}^\top\right] = \frac{1}{N} I_N \otimes H_{\mathrm{nom}} \Sigma_{k+1} H_{\mathrm{nom}}^\top \tag{43}$$

$$\mathbb{E}\left[G_{k+1} G_{k+1}^\top\right] = \frac{1}{N} I_N \otimes \left(G_{\mathrm{nom}} G_{\mathrm{nom}}^\top + (N-1) G_{\mathrm{cl}} G_{\mathrm{cl}}^\top\right) \tag{44}$$

$$\mathbb{E}\left[F_{k+1} \Lambda_k F_{k+1}^\top\right] = \Xi \otimes \left(H_{\mathrm{nom}} A \Lambda_k A^\top H_{\mathrm{nom}}^\top\right) \tag{45}$$

where

$$\Xi = \begin{cases} \frac{1}{N}\left((N-2)\mathbf{1}_N \mathbf{1}_N^\top + I_N\right), & N > 1 \\ 0, & N = 1. \end{cases} \tag{46}$$

Finally,

$$\mathbb{E}\left[H_{k+1}\left(\mathbb{E}[A_k]\Lambda_k + \mathbb{E}[B_k]\Upsilon_k^\top\right) F_{k+1}^\top\right] = \frac{1}{N}\left(\mathbf{1}_N \mathbf{1}_N^\top - I_N\right) \otimes \left(H_{\mathrm{nom}} A \Lambda_k A^\top H_{\mathrm{nom}}^\top\right). \tag{47}$$

The spatial distribution of clutter is uniform in the validation window, whose size determines $G_{\mathrm{cl}} G_{\mathrm{cl}}^\top$.

C. Discussion

It is easy to see that, in the present case, $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}} = I_N \otimes D$ where

$$D = \frac{1}{N} H_{\mathrm{nom}} A \Lambda_k A^\top H_{\mathrm{nom}}^\top + \frac{1}{N} H_{\mathrm{nom}} \Sigma_{k+1} H_{\mathrm{nom}}^\top + \frac{1}{N} G_{\mathrm{nom}} G_{\mathrm{nom}}^\top + \frac{N-1}{N} G_{\mathrm{cl}} G_{\mathrm{cl}}^\top.$$

Moreover

$$\begin{aligned} \Gamma_{x_{k+1}\tilde{y}_{k+1}} &= (\Sigma_{k+1} - A\Lambda_k A^\top)\mathbb{E}\left[H_{k+1}^\top\right] \\ &= \frac{1}{N}(\Sigma_{k+1} - A\Lambda_k A^\top)\left(H_{\mathrm{nom}}^\top \cdots H_{\mathrm{nom}}^\top\right)^\top \end{aligned} \tag{48}$$

and

$$\begin{aligned} K_k &= \Gamma_{x_{k+1}\tilde{y}_{k+1}} \Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}^{-1} \\ &= \frac{1}{N} \mathbf{1}_N^\top \otimes \left((\Sigma_{k+1} - A\Lambda_k A^\top) H_{\mathrm{nom}}^\top D^{-1}\right). \end{aligned} \tag{49}$$

Since y_{k+1} is a concatenation of all the observations from time $k+1$, the product $K_k y_{k+1}$ in (2) is the average of these measurements, pre-multiplied by $(\Sigma_{k+1} - A\Lambda_k A^\top) H_{\mathrm{nom}}^\top D^{-1}$. Consequently, the LMMSE estimator for tracking a target in clutter is a KF-like algorithm, operating on the average of all detections in the validation window. In this respect, its mode of operation resembles classical methods. For example, the probabilistic data association (PDA) [16] method implements a KF driven by the weighted average of all measurements in the window, and the nearest neighbor (NN) filter [17] is a KF driven by the measurement nearest to the prediction assigning it a weight of 1 and assigning 0 to the rest of the measurements.

D. Numerical Study

We consider a one-dimensional tracking scenario, in which the state comprises position and velocity information, $x_k = (p_k \ v_k)^\top$. Starting at $x_0 \sim \mathcal{N}(\bar{x}_0, P_0)$ with $\bar{x}_0 = (0 \ 0)^\top$ and $P_0 = 30 I_2$, the target is simulated for 400 time units using (36) with $A = \left(\begin{smallmatrix} 1 & 0.2 \\ 0 & 0.95 \end{smallmatrix}\right)$ and $C = (0.25 \quad 0.5)^\top$. The process and measurement noises are taken to be Gaussian.

The true measurement is generated using (37) with $H_{\mathrm{nom}} = (1 \ 0)$ and $G_{\mathrm{nom}} = \sqrt{30}$. The target is detected with probability $P_D = 0.95$ and the probability that the true observation falls in the validation window is taken to be $P_G = 0.99$. A validation window is set about the predicted measurement position. Its size, d, is determined to comply with P_G (see [17, p. 130] for details). Once the window is determined, the clutter variance of (39) is $G_{\mathrm{cl}} G_{\mathrm{cl}}^\top = d^2/12$.

The derived algorithm is compared with NN and PDA filters, that are equipped with the same windowing logic and parameters. All algorithms are initialized with $\hat{x}_0 = \bar{x}_0$ and the initial error covariance matrix is taken to be P_0.

When dealing with tracking in clutter, using the MSE as the only performance measure may result in misleading conclusions, since, eventually, the estimate will draw away from the true measurement and follow the clutter, and the errors will become meaninglessly large.

We thus use two measures of performance to evaluate the algorithms. The first is the time until the target is lost, defined as the third consecutive time when the measurement of a detected target falls outside the validation window. The second measure is the root MSE (RMSE) calculated over the time interval until the first of the three algorithms loses track.

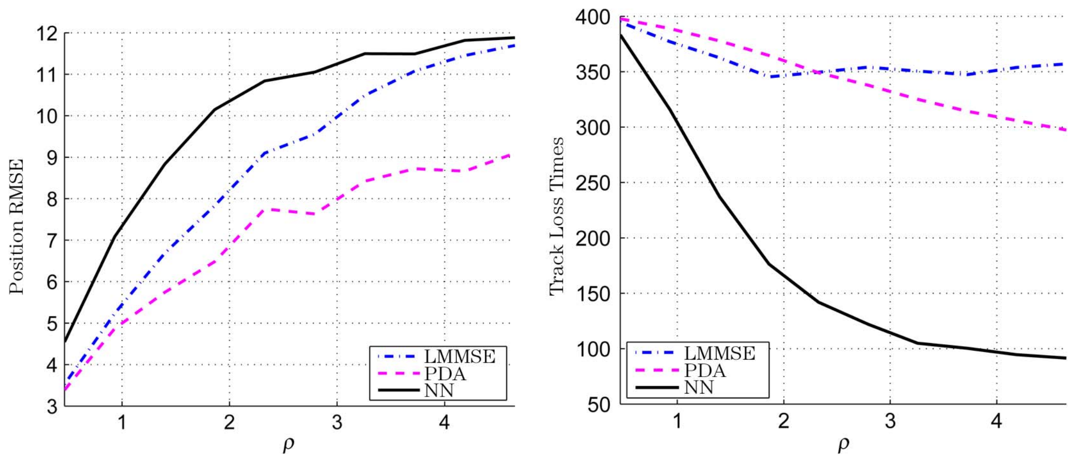

Fig. 1. Position RMSE (left) and track loss time (right) versus clutter density.

We test the algorithms at a range of clutter densities. We define the clutter density ρ to be the average number of clutter measurements falling in an interval of one standard deviation of the (true) measurement noise. Averaged over 1000 independent Monte Carlo runs, the average position RMSE and track loss times are plotted, versus ρ, in Fig. 1.

It is readily seen that the LMMSE filter attains competitive performance relatively to the nonlinear algorithms. Specifically, for heavy clutter regimes it maintains longest track loss times. It is not very surprising that the errors of PDA are better, since these are calculated before the first of the three algorithms has lost track (NN in all cases). During this period the PDA performs a more efficient, nonlinear manipulation on the measurements. However, for high clutter rates, it is probable that clutter measurements will be assigned higher weights than the true detection, eventually leading to a track loss. In this case, it is better to simply average the measurements, as the linear filter does.

V. Conclusion

We proposed a new formulation of JLS, where the dynamics and measurement equations are allowed to depend on previous estimates of the state representing closed-loop control input and measurement validation window. We derived an LMMSE recursive algorithm for this setting, and illustrated the approach in the context of tracking in clutter. In this case, our filter demonstrates competitive performance, when compared with classical, nonlinear methods.

Acknowledgment

The authors would like to thank Dr. Omer Bobrowski for fruitful discussions on the earlier versions of the paper.

References

[1] E.-K. Boukas and Z.-K. Liu, *Deterministic and Stochastic Time-Delay Systems*. Boston, MA: Birkhäuser, 2002.

[2] N. Nahi, "Optimal recursive estimation with uncertain observation," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 4, pp. 457–462, 1969.

[3] M. Hadidi and S. Schwartz, "Linear recursive state estimators under uncertain observations," *IEEE Trans. Autom. Control*, vol. AC-24, no. 6, pp. 944–948, Dec. 1979.

[4] N. Nahi and E. Knobbe, "Optimal linear recursive estimation with uncertain system parameters," *IEEE Trans. Autom. Control*, vol. AC-21, no. 2, pp. 263–266, 1976.

[5] G. Ackerson and K. Fu, "On state estimation in switching environments," *IEEE Trans. Autom. Control*, vol. AC-15, no. 1, pp. 10–17, 1970.

[6] H. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Autom. Control*, vol. 33, no. 8, pp. 780–783, 1988.

[7] W. De Koning, "Optimal estimation of linear discrete-time systems with stochastic parameters," *Automatica*, vol. 20, no. 1, pp. 113–115, 1984.

[8] O. Costa, "Linear minimum mean square error estimation for discrete-time Markovian jump linear systems," *IEEE Trans. Autom. Control*, vol. 39, no. 8, pp. 1685–1689, 1994.

[9] T. Michaeli, D. Sigalov, and Y. Eldar, "Partially linear estimation with application to sparse signal recovery from measurement pairs," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2125–2137, 2012.

[10] T. Yuan, Y. Bar-Shalom, P. Willett, E. Mozeson, S. Pollak, and D. Hardiman, "A multiple IMM estimation approach with unbiased mixing for thrusting projectiles," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 3250–3267, 2012.

[11] R. Jackson and D. Murthy, "Optimal linear estimation with uncertain observations (Corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-22, no. 3, pp. 376–378, 1976.

[12] D. Magill, "Optimal adaptive estimation of sampled stochastic processes," *IEEE Trans. Autom. Control*, vol. AC-10, no. 4, pp. 434–439, 1965.

[13] D. Sigalov and Y. Oshman, "State estimation in hybrid systems with a bounded number of mode transitions," in *Proc. 13th Int. Conf. Information Fusion*, 2010.

[14] J. Mendel, *Lessons in Digital Estimation Theory*: Prentice-Hall, Inc., 1986.

[15] D. Sigalov, T. Michaeli, and Y. Oshman, "Linear optimal state estimation in systems with independent mode transitions," in *Proc. 50th IEEE Conf. Decision and Control*, 2011.

[16] Y. Bar-Shalom and E. Tse, "Tracking in a cluttered environment with probabilistic data association," *Automatica*, vol. 11, no. 5, pp. 451–460, 1975.

[17] Y. Bar-Shalom and X. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS Publishing, 1995.