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

Daniel Sigalov, Tomer Michaeli, 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.

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 problems 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]–[5], 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.

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, parts of the state may be estimated optimally while others in a linear optimal manner, 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 a recursive 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 [9]. 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 paper is organized as follows. In Sec. II we describe the proposed modeling and survey some related work. The recursive LMMSE algorithm is derived in Sec. III. An application to target tracking in clutter, followed by a numerical study, is presented in Sec. IV. Concluding remarks are given in Sec. V.

II. SYSTEM MODEL AND RELATED WORK

We consider the dynamical system

\begin{align}
  x_{k+1} &= A_k x_k + B_k u_k + C_k w_k \quad (1a) \\
  y_k &= H_k x_k + G_k v_k + F_k \hat{x}_{k-1}, \quad (1b)
\end{align}

where \( x_k \in \mathbb{R}^n \) and \( y_k \in \mathbb{R}^m \) are the state and measurement vectors at time \( k \), respectively. 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 \( \hat{x}_0 \) 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 \( \hat{F}_k \hat{x}_{k-1} \) 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.

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

\[ \hat{x}_{k+1} = L_k \hat{x}_k + K_k y_{k+1} + J_k u_k \]

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], [6]. 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 V.

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 an LMMSE sense, in [2].
We now compute the covariance terms \( \Gamma_{x_{k+1}\hat{y}_{k+1}} \) and \( \Gamma_{\hat{y}_{k+1}\hat{y}_{k+1}} \).

Let \( \hat{y}_{k+1} \) be the LMMSE estimate of \( y_{k+1} \). Since \( \hat{y}_{k+1} \) is unbiased, and using (15) and (5).

\[
\Gamma_{x_{k+1}\hat{y}_{k+1}} = \mathbb{E} \left[ x_{k+1} (y_{k+1} - \hat{y}_{k+1})^\top \right] = \mathbb{E} \left[ x_{k+1}(H_{k+1}x_{k+1} + G_{k+1}u_{k+1} + F_{k+1}\hat{x}_k)^\top \right] - \mathbb{E} \left[ x_{k+1}(\mathbb{E} [H_{k+1}] + \mathbb{E} [A_k] + \mathbb{E} [F_{k+1}])\hat{x}_k \right]^\top - \mathbb{E} \left[ x_{k+1}(\mathbb{E} [H_{k+1}] + \mathbb{E} [B_k])u_k^\top \right].
\]

Using the independence of \( x_{k+1} \) and \( u_{k+1} \), and canceling out identical terms, (9) becomes

\[
\Gamma_{x_{k+1}\hat{y}_{k+1}} = \mathbb{E} [x_{k+1}x_{k+1}^\top] (\mathbb{E} [A_k] + \mathbb{E} [B_k])^\top \mathbb{E} [A_k]^\top + \mathbb{E} [B_k] \mathbb{E} [B_k]^\top - \mathbb{E} [x_{k+1}u_k^\top] \mathbb{E} [B_k] \mathbb{E} [B_k]^\top. \tag{10}
\]

Before proceeding, we define \( \Sigma_k \triangleq \mathbb{E} [x_kx_k^\top], \Delta_k \triangleq u_ku_k^\top \) and, in addition,

\[
\Lambda_k \triangleq \mathbb{E} [\hat{x}_k\hat{x}_k^\top] = \mathbb{E} [\hat{x}_k\hat{x}_k^\top]; \tag{11}
\]

\[
\Upsilon_k \triangleq \mathbb{E} [x_ku_k^\top] = \mathbb{E} [x_ku_k^\top]. \tag{12}
\]

where the RHS of (11) and (12) follow from the orthogonality principle and from the unbiasedness of \( \hat{x}_k \), respectively. Note that \( \Sigma_k, \Lambda_k, \) and \( \Delta_k \) are symmetric.

Using the independence of \( x_k \) and \( w_k \),

\[
\mathbb{E} [x_{k+1}\hat{x}_k^\top] = \mathbb{E} [x_{k+1}(A_kx_k + B_ku_k + C_kw_k)^\top] = \mathbb{E} [A_k] \Lambda_k + \mathbb{E} [B_k] \Upsilon_k, \tag{13}
\]

which yields for (10)

\[
\Gamma_{x_{k+1}\hat{y}_{k+1}} = (\Sigma_{k+1} - (\mathbb{E} [A_k] \Lambda_k + \mathbb{E} [B_k] \Upsilon_k) \mathbb{E} [A_k]^\top - \mathbb{E} [x_{k+1}u_k^\top] \mathbb{E} [B_k] \mathbb{E} [B_k]^\top) \mathbb{E} [H_{k+1}^\top]. \tag{14}
\]

From (14) we have

\[
\mathbb{E} [x_{k+1}] = \mathbb{E} [A_k]x_k + \mathbb{E} [B_k]u_k + C_kw_k
\]

which, when substituted in (14), leads to

\[
\Gamma_{x_{k+1}\hat{y}_{k+1}} = (\Sigma_{k+1} - (\mathbb{E} [A_k] \Lambda_k \mathbb{E} [A_k]^\top + \Upsilon_k \mathbb{E} [B_k]^\top) + \mathbb{E} [B_k] (\Upsilon_k \mathbb{E} [A_k]^\top + \Delta_k \mathbb{E} [B_k]^\top)) \mathbb{E} [H_{k+1}^\top]. \tag{16}
\]

B. Computation of \( \Gamma_{\hat{y}_{k+1}\hat{y}_{k+1}} \)

Since \( \hat{y}_{k+1} \) is the LMMSE estimate of \( y_{k+1} \) using \( \hat{y}_{k+1} \), \( \hat{y}_{k+1} \) is orthogonal to \( \hat{y}_{k+1} \) and, using (5).

\[
\Gamma_{\hat{y}_{k+1}\hat{y}_{k+1}} = \mathbb{E} [(y_{k+1} - \hat{y}_{k+1})y_{k+1}^\top] = \mathbb{E} [y_{k+1}y_{k+1}^\top] - \mathbb{E} [\hat{y}_{k+1}y_{k+1}^\top] = \mathbb{E} [y_{k+1}y_{k+1}^\top] - \mathbb{E} [H_{k+1}(\mathbb{E} [A_k] + \mathbb{E} [F_{k+1}])\hat{x}_k y_{k+1}^\top] - \mathbb{E} [H_{k+1}]^\top \mathbb{E} [B_k] u_k u_k^\top. \tag{17}
\]

Using (15) and the independence of \( \{x_k, x_{k+1}\} \), \( \{H_{k+1}, G_{k+1}, F_{k+1}\} \) and \( u_{k+1} \), we have

\[
\mathbb{E} [x_k y_{k+1}^\top] = \mathbb{E} [x_k x_{k+1}^\top] + F_{k+1} \mathbb{E} [\hat{x}_k] y_{k+1}^\top] = \mathbb{E} [x_k x_{k+1}^\top] \mathbb{E} [H_{k+1}] + \Lambda_k \mathbb{E} [F_{k+1}] + \Upsilon_k \mathbb{E} [B_k] \mathbb{E} [H_{k+1}], \tag{18}
\]

which, using (13), becomes

\[
\mathbb{E} [x_k y_{k+1}^\top] = \Lambda_k \mathbb{E} [A_k^\top] \mathbb{E} [H_{k+1}^\top] + \mathbb{E} [F_{k+1}] + \Upsilon_k \mathbb{E} [B_k] \mathbb{E} [H_{k+1}]. \tag{19}
\]
Due to the independence of \( \{x_{k+1}, \tilde{x}_k\} \), \( v_{k+1} \), and \( \{H_{k+1}, G_{k+1}\} \)
\[
\mathbb{E}[y_{k+1}^+ y_{k+1}^\top] = \mathbb{E}[H_{k+1} x_{k+1} x_{k+1}^\top H_{k+1}^\top] + \mathbb{E}[G_{k+1} v_{k+1} v_{k+1}^\top G_{k+1}^\top] + \mathbb{E}[F_{k+1} \tilde{x}_k \tilde{x}_k^\top F_{k+1}^\top] + \mathbb{E}[H_{k+1} x_{k+1} \tilde{x}_k^\top F_{k+1}^\top] + \mathbb{E}[F_{k+1} \tilde{x}_k x_{k+1}^\top H_{k+1}^\top].
\]

Consider the last summand. From the smoothing property of the conditional expectation,
\[
\mathbb{E}[F_{k+1} \tilde{x}_k x_{k+1}^\top H_{k+1}^\top] = \mathbb{E}\left[ F_{k+1} \tilde{x}_k x_{k+1}^\top (E_{k+1}^\top | H_{k+1}^\top) \right] = \mathbb{E}[F_{k+1} \mathbb{E}[\tilde{x}_k x_{k+1}^\top | H_{k+1}^\top]],
\]
where we utilized the independence of \( \{H_{k+1}, F_{k+1}\} \) and \( \{x_{k+1}, \tilde{x}_k\} \).

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 \( \mathbb{E}[\tilde{x}_k \tilde{x}_k^\top] = \Lambda_k \), we obtain:
\[
\mathbb{E}[H_{k+1} x_{k+1} x_{k+1}^\top H_{k+1}^\top] = \mathbb{E}[H_{k+1} \Sigma_{k+1} H_{k+1}^\top] \quad (22)
\]
\[
\mathbb{E}[G_{k+1} v_{k+1} v_{k+1}^\top G_{k+1}^\top] = \mathbb{E}[G_{k+1} \Sigma_{k+1} G_{k+1}^\top] \quad (23)
\]
\[
\mathbb{E}[F_{k+1} \tilde{x}_k x_{k+1}^\top F_{k+1}^\top] = \mathbb{E}[F_{k+1} \Lambda_k F_{k+1}^\top].
\]

For future reference, we also note that
\[
\mathbb{E}[A_k x_k x_k^\top A_k^\top] = \mathbb{E}[A_k \Sigma_k A_k^\top] \quad (25)
\]
\[
\mathbb{E}[A_k u_k B_k^\top] = \mathbb{E}[A_k \Upsilon_k B_k^\top] \quad (26)
\]
\[
\mathbb{E}[B_k u_k B_k^\top] = \mathbb{E}[B_k \Delta_k B_k^\top] \quad (27)
\]
\[
\mathbb{E}[C_k w_k w_k^\top C_k^\top] = \mathbb{E}[C_k C_k^\top].
\]

Substituting (13) in (21), and using (21)–(23) in (20),
\[
\mathbb{E}[y_{k+1} y_{k+1}^\top] = \mathbb{E}[H_{k+1} \Sigma_{k+1} H_{k+1}^\top] + \mathbb{E}[G_{k+1} \Sigma_{k+1} G_{k+1}^\top] + \mathbb{E}[F_{k+1} \Lambda_k F_{k+1}^\top] + \mathbb{E}\left[ H_{k+1}(E[A_k] \Lambda_k + E[B_k] \Upsilon_k^\top) F_{k+1}^\top \right] + \mathbb{E}\left[ F_{k+1}(\Lambda_k E[A_k^\top] + \Upsilon_k E[B_k^\top]) H_{k+1}^\top \right].
\]

In addition, we obtain, in a straightforward manner,
\[
\mathbb{E}[y_{k+1}] = (\mathbb{E}[H_{k+1}] \mathbb{E}[A_k] + \mathbb{E}[F_{k+1}]) \mathbb{E}[x_k] + \mathbb{E}[H_{k+1}] \mathbb{E}[B_k] u_k.
\]

Using (22), (23), and (24) in (29), and substituting (19), (29), and (30) in (17), we finally obtain
\[
\Gamma_{y_{k+1} y_{k+1}} = \mathbb{E}[H_{k+1} \Sigma_{k+1} H_{k+1}^\top] + \mathbb{E}[G_{k+1} \Sigma_{k+1} G_{k+1}^\top] + \mathbb{E}[F_{k+1} \Lambda_k F_{k+1}^\top] - \mathbb{E}[H_{k+1}] \mathbb{E}[A_k] \Lambda_k \mathbb{E}[E[H_{k+1}^\top]]
\]
\[
+ \mathbb{E}[H_{k+1}] \mathbb{E}[A_k] \Lambda_k \mathbb{E}[E[H_{k+1}^\top]] + \mathbb{E}\left[ H_{k+1}(E[A_k] \Lambda_k + E[B_k] \Upsilon_k^\top) F_{k+1}^\top \right] + \mathbb{E}\left[ F_{k+1}(\Lambda_k E[A_k^\top] + \Upsilon_k E[B_k^\top]) H_{k+1}^\top \right]
\]
\[
- \mathbb{E}[H_{k+1}] \mathbb{E}[A_k] \Lambda_k \mathbb{E}[E[H_{k+1}^\top]] - \mathbb{E}[F_{k+1}] \Lambda_k \mathbb{E}[E[A_k^\top]] E[H_{k+1}^\top]
\]
\[
- \mathbb{E}[H_{k+1}] \mathbb{E}[A_k] \Lambda_k \mathbb{E}[E[H_{k+1}^\top]] - \mathbb{E}[F_{k+1}] \Lambda_k \mathbb{E}[E[A_k^\top]] E[H_{k+1}^\top]
\]
\[
- \mathbb{E}[H_{k+1}] \mathbb{E}[A_k] \Upsilon_k \mathbb{E}[E[B_k^\top]] E[H_{k+1}^\top]
\]
\[
- \mathbb{E}[F_{k+1}] \Upsilon_k \mathbb{E}[E[B_k^\top]] E[H_{k+1}^\top]
\]
\[
- \mathbb{E}[H_{k+1}] \mathbb{E}[B_k] u_k \mathbb{E}[E[y_{k+1}]].
\]

Since the distribution of \( \mathcal{M}_k \) is known, the expectations of steps (11) and (12) 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
\[
\Gamma_{y_{k+1} y_{k+1}} = (\Sigma_{k+1} - A_k \Lambda_k A_k^\top) H_{k+1}^\top
\]
and
\[
\Gamma_{y_{k+1} y_{k+1}} = H_{k+1} (\Sigma_{k+1} - A_k \Lambda_k A_k^\top) H_{k+1}^\top + G_{k+1} G_{k+1}^\top.
\]

Substituting these in (3) 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 (13), 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 \( \mathcal{Y}_k \). An alternative derivation, based on the orthogonality principle, may be found in [13].

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 (14)

\[
x_{k+1} = A x_k + C w_k.
\]

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 equation

\[
y_k = H_{\text{nom}} x_k + G_{\text{nom}} v_k^\text{true},
\]

where \( v_k^\text{true} \) represents measurement noise. In addition, at each time, a number of clutter detections are obtained. These will be denoted as \( \{ y_{cl, k} \}_{k=1}^{N-1} \), 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 predicted target position, 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 (15) by letting the mode \( \mathcal{M}_k \) be distributed as

\[
\mathcal{M}_k = \{ H_k, G_k, F_k \} = \left\{ \begin{array}{c}
\{ H_{\text{nom}} \} \\
\{ G_{\text{nom}} \} \\
\{ G_{\text{cl}} \} \\
\{ 0 \} \\
\{ 0 \} \\
\{ 0 \} \\
\{ 0 \}
\end{array} \right\} \begin{array}{c}
\text{diag} \left( \frac{G_{\text{nom}}}{\Lambda_{\text{nom}}} \right) \\
\text{diag} \left( \frac{G_{\text{cl}}}{\Lambda_{\text{nom}}} \right) \\
\text{diag} \left( \frac{H_{\text{nom}} A}{\Lambda_{\text{nom}}} \right) \\
0 \\
0 \\
0 \\
0
\end{array}, \quad \text{w.p. } \frac{1}{N}.
\]

\[
\mathcal{M}_k = \left\{ \begin{array}{c}
\{ H_{\text{nom}} \} \\
\{ G_{\text{nom}} \} \\
\{ G_{\text{cl}} \} \\
\{ 0 \} \\
\{ 0 \} \\
\{ 0 \} \\
\{ 0 \}
\end{array} \right\} \begin{array}{c}
\text{diag} \left( \frac{G_{\text{nom}}}{\Lambda_{\text{nom}}} \right) \\
\text{diag} \left( \frac{G_{\text{cl}}}{\Lambda_{\text{nom}}} \right) \\
\text{diag} \left( \frac{H_{\text{nom}} A}{\Lambda_{\text{nom}}} \right) \\
0 \\
0 \\
0 \\
0
\end{array}, \quad \text{w.p. } \frac{1}{N},
\]

where \( G_{\text{cl}} \) 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 (37) corresponds to the scenario in which the first of the \( N \) observations is the true target measurement, \( y_k^{\text{true}} \), generated according to (36), while the other \( N - 1 \) measurements are clutter, each of which is generated according to

\[
y_k^{\text{cl}} = H_{\text{nom}} A \hat{x}_{k-1} + G_{\text{cl}} v_k^{\text{cl}}, \quad i = 2, \ldots, N.
\]

Here, \( H_{\text{nom}} A \hat{x}_{k-1} \) 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, \( v_k^{\text{cl}} \) 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 (2) 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 motion 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

\[
\{ H_k, G_k, F_k \} = \{ 0, I_N \otimes G_{\text{cl}}, I_N \otimes H_{\text{nom}} A \}
\]

needs to be added to the set of possible realizations in (37). Here, \( \otimes \) stands for the Kronecker product, \( I_N \) 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 \((1 - P_D)(1 - P_c)\) where \( P_D \) is the probability of target detection, assumed known, and \( P_c \) is the probability that, upon target detection, the true measurement falls in the validation window. This parameter is defined by the user and, typically, it affects the window size as discussed in the sequel. Note that, when no measurements are available, \( N = 0 \), and (2) becomes (at 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, \( E[ A_k] = A \), \( E[ B_k] = 0 \), \( E[ C_k G_k^\top] = C C^\top \), \( E[ A_k \Sigma_k A_k^\top] = \Lambda_k A \Sigma_k A \top \). Also, according to the distribution defined in (37),

\[
E[ H_{k+1} ] = \frac{1}{N} \left[ \begin{array}{c}
1_N \otimes H_{\text{nom}}
\end{array} \right]
\]

\[
E[ F_{k+1} ] = \frac{N - 1}{N} \left[ \begin{array}{c}
1_N \otimes H_{\text{nom}} A
\end{array} \right].
\]

The remaining terms read

\[
E[ H_{k+1} \Sigma_{k+1} H_{k+1}^\top ] = \frac{1}{N} \left[ \begin{array}{c}
1_N \otimes H_{\text{nom}} \Sigma_{k+1} H_{\text{nom}}^\top
\end{array} \right]
\]

\[
E[ G_{k+1} G_{k+1}^\top ] = \frac{1}{N} \left[ \begin{array}{c}
1_N \otimes G_{\text{nom}} G_{\text{nom}}^\top + (N - 1) G_{\text{cl}} G_{\text{cl}}^\top
\end{array} \right]
\]

\[
E[ F_{k+1} A_k F_{k+1}^\top ] = \Xi \otimes \left( H_{\text{nom}} A \Lambda_k A \top H_{\text{nom}} \right).
\]
where
\[
\Xi = \begin{cases} 
\frac{1}{N} ((N - 2) + 1_N 1_N^T + I_N), & N > 1 \\
0, & N = 1.
\end{cases}
\] (44)

Finally,
\[
\mathbb{E} \left[ H_{k+1} \left( \Xi [ A_k ] A_k^T + \mathbb{E} \left[ B_k \right] Y_k^T \right) F_{k+1}^T \right] \\
= \frac{1}{N} \left( (1_N 1_N^T - I_N) \otimes \left( H_{\text{nom}} A_k A_k^T H_{\text{nom}}^T \right) \right). 
\] (45)

The spatial distribution of clutter is uniform in the validation window, whose size determines \( G_{cl} G_{cl}^T \).

C. Discussion

It is easy to see that, in the present case, \( \Gamma_{\hat{y}_{k+1} \hat{y}_{k+1}} = I_N \otimes D \) where
\[
D = \frac{1}{N} H_{\text{nom}} A_k A_k^T H_{\text{nom}}^T + \frac{1}{N} H_{\text{nom}} \Sigma_{k+1}^1 H_{\text{nom}}^T \\
+ \frac{1}{N} G_{\text{nom}} G_{\text{nom}}^T + \frac{N-1}{N} G_{cl} G_{cl}^T.
\]

Moreover,
\[
\Gamma_{\hat{x}_{k+1} \hat{y}_{k+1}} = \left( \Sigma_{k+1}^1 - A A_k A^T \right) \mathbb{E} \left[ H_{k+1}^T \right] = \frac{1}{N} \left( \Sigma_{k+1}^1 - A A_k A^T \right) \left( H_{\text{nom}}^T \cdots H_{\text{nom}}^T \right)^T, 
\] (46)

and
\[
K_k = \Gamma_{\hat{x}_{k+1} \hat{y}_{k+1}} \Gamma_{\hat{y}_{k+1} \hat{y}_{k+1}}^{-1} = \frac{1}{N} I_N \otimes \left( (\Sigma_{k+1}^1 - A A_k A^T) H_{\text{nom}}^T D^{-1} \right). 
\] (47)

Since \( y_{k+1} \) is a concatenation of all the observations from time \( k+1 \), the product \( K_k y_{k+1} \) in (47) is the average of these measurements, pre-multiplied by \( (\Sigma_{k+1}^1 - A A_k A^T) H_{\text{nom}}^T D^{-1} \). 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) [10] 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, \( x_k = (p_k \ v_k)^T \). Starting at \( x_0 \sim N(x_0, P_0) \) with \( x_0 = (0 \ 0)^T \) and \( P_0 = 30I_2 \), the target is simulated for 400 time units using (35) with \( A = (1 \ 0.2) \) and \( C = \frac{1}{2} (1/4 I_1) \). The process and measurement noises are taken to be Gaussian. The true measurement is generated using (36) with \( \rho = 0.95 \) and the probability that the true observation falls in the validation window is taken to be \( P_C = 0.99 \). A validation window is set about the predicted measurement position. Its size, \( d \), is determined to comply with \( P_C \) (see [17], p.130 for details). Once the window is determined, the clutter variance of (38) is \( G_{cl} G_{cl}^T = d^2/12 \).

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 \( x_0 = \bar{x}_0 \) 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 meaningless 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.

We test the algorithms at a range of clutter densities. Let \( \rho \) 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 \( \rho \), 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.

REFERENCES

[1] E.-K. Boukas and Z.-K. Liu, Deterministic and Stochastic Time-Delay Systems. Boston: 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, December 1979.
[4] N. Nahi and E. Knobe, “Optimal linear recursive estimation with uncertain system parameters,” IEEE Trans. Autom. Control, vol. 21, no. 2, pp. 263–266, 1976.
[5] G. Ackerson and K. Fu, “On state estimation in switching environments,” IEEE Trans. Autom. Control, vol. 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. Michalik, 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. Poliak, 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. 22, no. 3, pp. 376–378, 1976.
[12] D. Magill, “Optimal adaptive estimation of sampled stochastic processes,” IEEE Trans. Autom. Control, vol. 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. Fusion 2010, 13th International Conference on Information Fusion, 2010.
Fig. 1: Position RMSE (left) and track loss time (right) vs. clutter density.

[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. CDC 2011, 50th Conf. on Decision and Control*. IEEE, 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.