A Unified Approach to State Estimation Problems Under Data and Model Uncertainties

Daniel Sigalov
Applied Mathematics
Technion - IIT
Haifa, Israel

Tomer Michaeli
Electrical Engineering
Technion - IIT
Haifa, Israel

Yaakov Oshman
Aerospace Engineering
Technion - IIT
Haifa, Israel

***Abstract*—We present a unified approach to the problem of state estimation under measurement and model uncertainties. The approach allows formulation of problems such as maneuvering target tracking, target tracking in clutter, and multiple target tracking using a single state-space system with random matrix coefficients. Consequently, all may be solved efficiently using a single IMM algorithm or using a linear optimal filter, derived elsewhere, thus replacing the need for deriving a unique algorithm for each problem.**

***Index Terms*—Maneuvering target tracking, clutter and data association, hybrid systems, multiple target tracking**

I. Introduction

State estimation in dynamical systems with randomly switching coefficients is an important problem with a variety of applications. The most natural examples are maneuvering target tracking and fault detection and isolation (FDI) featured by navigation systems. The standard modeling presumes that the dynamics of the state, being a continuous random variable, is controlled by an evolving mode that takes discrete values. This is the well known concept of hybrid systems [1].

Various problems have been formulated within this modeling. In problems involving uncertain, or intermittent observations, such as [2]–[7], the mode comprises the matrices of the measurement equation. In maneuvering target tracking applications considered in, e.g., [8]–[11] the mode usually comprises the matrices of the dynamics equation.

As is well known, the optimal, in the minimum mean squared error (MMSE) sense, estimator of the state in systems with randomly switching coefficients, or more specifically, in Markov Jump Linear Systems (MJLS), requires exponentially growing resources [9] and is, thus, impractical in most problems of practical interest. Therefore, suboptimal state estimation algorithms attract special interest of both researchers and practitioners. The Interacting Multiple Model (IMM) [10] is perhaps the most famous method that proposes a successful compromise between performance and complexity. Another option is using linear optimal recursive filters [12], [13].

A related class of problems is referred to as data association or data ambiguity. In this family, the estimation process is further complicated by the fact that the acquired data has uncertain origin. For example, in the problem of tracking a single target in clutter, each of the obtained measurements may originate from the true target or may represent a false alarm that does not carry useful information. In multiple target tracking, each of the true detections may be attributed to each of the targets without a-priori labeling.

A well-known Bayesian method for tracking a single non-maneuvering target in clutter is the Probabilistic Data Association Filter (PDAF) [14] in which the optimal state estimate is approximated by fusing, in a Kalman filter's manner, the previously obtained result with a nonlinear combination of the measurements from the current scan. Recently, the problem of tracking in clutter was formulated using a single state-space system and solved in a linear-optimal manner [13]. In the case of multiple targets the joint PDAF (JPDAF) [15] approximates their states by considering all possible association events between the estimates from the previous cycle and the new set of measurements. Treating maneuvering targets in cluttered environment is possible by combining the PDAF/JPDAF with the IMM algorithm.

In this paper we show how some of the above problems may be modeled within the framework of a single, generalized state-space system with randomly switching coefficients. Consequently, all may be solved using a *single* IMM or a linear-optimal filter provided that the set of the system modes is properly chosen. Specifically, we consider the classical problem of tracking a maneuvering target and perform a comparison of IMM and the linear optimal filter. In addition, we discuss the problem of tracking in clutter with both adaptive and linear methods. We extend the treatment of [13] and show that the classical PDA approach is equivalent to an appropriately designed IMM. Finally, we solve the multi-target tracking problem and discuss cons and pros in using IMM and linear algorithms.

The remainder of the paper is organized as follows. In Section II we define and discuss the general modeling and outline the possible generic solutions. Section III is devoted to applications of the unified modeling. We begin with the classical problem of maneuvering target tracking in Section III-A and solve it using IMM and linear methods. Section III-B considers the classical problems of tracking a single target in clutter, where we also show the equivalence between the resulting IMM procedure and the classical PDAF. Section III-C discusses the multiple target tracking problem. Concluding remarks are made in Section IV.

II. The Unified Approach and Existing Solutions

We consider the dynamical system

$$x_{k+1} = A(\theta_k)x_k + C(\theta_k)w_k \tag{1a}$$
$$y_k = H(\theta_k)x_k + G(\theta_k)v_k + F(\theta_k)\hat{x}_{k-1}. \tag{1b}$$

Here x_k is the state vector at time k. The process and measurement noise sequences, $\{w_k\}$ and $\{v_k\}$, are taken to be i.i.d. with zero mean and unit covariance matrices, independent of each other, and $\{A(\theta)\}$, $\{C(\theta)\}$, $\{H(\theta)\}$, $\{G(\theta)\}$, and $\{F(\theta)\}$ are matrix-valued, finite domain functions of a real scalar random variable $\theta \in \mathbb{R}$. The vector $\hat{x}_{k-1}$ is some estimate of x_{k-1} using the measurements $\mathcal{Y}_{k-1}$ where $\mathcal{Y}_k \triangleq \{y_0, \ldots, y_k\}$. For example, it may be the linear MMSE (LMMSE) estimate of the state using the past measurements. The last term in (1b) represents the fact that measurements are not collected from the entire surveillance region but from a validation window, set about the predicted measurement. Without loss of generality, x_0 is taken to be a zero vector with zero covariance matrix. The mode of the system, $\{\theta_k\}$, constitutes a Markov process with a state-space $\{1, \ldots, r\}$, a transition probability matrix (TPM) Π, and an initial distribution ν.

The goal is to obtain an efficient estimate of x_k using the available data $\mathcal{Y}_k$. Since the MMSE estimate is impractical, suboptimal approaches are inevitable. We discuss some of these next.

A. The Interacting Multiple Model Filter

The main idea underlying the IMM algorithm, in the absence of the additional term $F(\theta_k)\hat{x}_{k-1}$ in (1b), is to maintain a bank of primitive Kalman filters, each matched to a different model in the given model set. At step k, the j-th filter produces a local estimate $\hat{x}_j(k)$ with an associated error covariance $P_j(k)$ using its initial estimate $\hat{x}_j^0(k-1)$ and the associated covariance $P_j^0(k-1)$, which are generated externally, and the current measurement y_k, which gets processed by all KFs in the bank. In addition, each filter produces a current value of its own (model-matched) likelihood function $\Lambda_j(k)$. The key element of the IMM scheme is the interaction block that generates, using all local estimates, covariances, and likelihoods from the previous cycle, individual initial conditions for each of the primitive filters in the bank. The steps of the algorithm are summarized as follows.

a) Mixing Probabilities: For $i, j = 1, \ldots, r$ compute

$$\begin{aligned}\mu_{i|j}(k-1) &\triangleq \mathbb{P}\{\theta_{k-1} = i \mid \theta_k = j, \mathcal{Y}_{k-1}\} \\ &= \frac{1}{c_j} p_{ij}\mu_i(k-1),\end{aligned} \tag{2}$$

where c_j is a normalizing constant and $\mu_i(k) \triangleq \mathbb{P}\{\theta_k = i \mid \mathcal{Y}_k\}$.

b) Mixing Step: For $j = 1, \ldots, r$ compute the initial state estimate for the filter matched to m_j

$$\hat{x}_j^0(k-1) = \sum_{i=1}^{r} \hat{x}_i(k-1)\mu_{i|j}(k-1) \tag{3}$$

and the corresponding covariances.

c) Mode-Matched Filtering: For $j = 1, \ldots, r$, using (3) and the corresponding covariance, compute the mode-matched estimate $\hat{x}_j(k)$ and $P_j(k)$ as well as the likelihood $\Lambda_j(k)$, which is approximated as Gaussian

$$\Lambda_j(k) = \mathcal{N}(y_k; \hat{y}_j(k), S_j(k)), \tag{4}$$

where $\hat{y}_j(k)$ and $S_j(k)$ are the predicted measurement and innovation covariance computed by the j-th filter using the initial conditions (3).

d) Mode Probability Update: For $j = 1, \ldots, r$

$$\mu_j(k) = \frac{1}{c}\Lambda_j(k)\sum_{i=1}^{r} p_{ij}\mu_i(k-1), \tag{5}$$

where c is a normalization factor.

e) Output Computation: The algorithm output at time k is obtained as a fused version of the local estimates:

$$\hat{x}(k) = \sum_{j=1}^{r} \hat{x}_j(k)\mu_j(k). \tag{6}$$

The associated covariance is computed in a similar manner.

In the proposed formulation, each primitive KF should be replaced with a corresponding Generalized KF (GKF) incorporating the above feedback term. It is shown in [13] that for a deterministic $\{\theta_k\}$, the GKF estimate of the state defined and observed through (1) is given by

$$\hat{x}_{k+1} = L_k\hat{x}_k + K_k y_{k+1} \tag{7}$$
$$L_k = (I - K_k H)A - K_k F \tag{8}$$
$$K_k = P_{k+1}^- H^T\left(HP_{k+1}^- H^T + GG^T\right)^{-1} \tag{9}$$
$$P_{k+1}^- = A(P_k^- - K_{k-1}HP_k^-)A^T + CC^T. \tag{10}$$

Here, A, C, H, G, and F are functions of the deterministic sequence $\{\theta_k\}$.

B. The Linear Optimal Filter

To the best of the authors' knowledge, linear-optimal algorithms for the generalized system (1) have not been addressed elsewhere. There are, however, special cases for which such filters have been developed. In the absence of the feedback term $F(\theta_k)\hat{x}_{k-1}$ in (1b), the LMMSE algorithm proposed in [12] presumes a finite number of possible modes described by the possible outcomes of θ_k. The idea is to obtain a linear optimal recursive estimate of the following augmented state

$$z_k \triangleq (x_k^T \mathbb{1}_{\{\theta_k=1\}}, \ldots, x_k^T \mathbb{1}_{\{\theta_k=r\}})^T. \tag{11}$$

The estimate of the true state is obtained naturally by summing up the entries of the linear optimal estimate of z_k. The resulting recursive scheme for the estimation of the augmented state vector is summarized below.

$$\begin{aligned}\hat{z}_k = A\hat{z}_{k-1} + (S_k - V_k)H^T\left(H(S_k - V_k)H^T + G_k G_k^T\right)^{-1} \\ \times (y_k - HA\hat{z}_{k-1}),\end{aligned} \tag{12}$$

where S_k, U_k, and V_k are second-order moments of z_k, $\hat{z}_k$, and $A\hat{z}_{k-1}$, respectively. These may be recursively computed from

$$V_k = AU_{k-1}A^T \tag{13}$$

$$U_k = V_k + (S_k - V_k)H^T\left(H(S_k - V_k)H^T + G_k G_k^T\right)^{-1} \times H_k(S_k - V_k) \tag{14}$$

$$S_k = \mathrm{diag}(S_k(j)) \tag{15}$$

$$S_{k+1}(j) = \sum_{i=1}^{r} p_{ij} A(i) S_k(i) (A(i))^T + \sum_{i=1}^{r} p_{ij} \pi_k(i) C(i) (C(i))^T, \tag{16}$$

where $\pi_k(i) = \mathbb{P}\{\theta_k = i\}$ and

$$A \triangleq \begin{pmatrix} p_{11}A(1) & \cdots & p_{r1}A(r) \\ \vdots & \ddots & \vdots \\ p_{1r}A(1) & \cdots & p_{rr}A(r) \end{pmatrix} \tag{17}$$

$$H \triangleq \begin{pmatrix} H(1) & \cdots & H(r) \end{pmatrix} \tag{18}$$

$$G_k \triangleq \begin{pmatrix} G(1)\pi_k^{1/2}(1) & \cdots & G(r)\pi_k^{1/2}(r) \end{pmatrix} \tag{19}$$

It was shown in [16] that when the Markov dynamics of $\{\theta_k\}$ degenerates to an independent process, the LMMSE estimator may be obtained without state augmentation. The algorithm then assumes a convenient Kalman-like form.

It is shown in [13] that adding the additional term $F(\theta_k)\hat{x}_{k-1}$ in (1b), and assuming independent mode transitions, the LMMSE filter comprises the following equations, where we have denoted, for brevity, $A_k \triangleq A(\theta_k)$ and similarly for C_k, H_k, G_k, and F_k.

$$\hat{x}_{k+1} = L_k \hat{x}_k + K_k y_{k+1} \tag{20}$$

$$K_k = \Gamma_{x_{k+1}\tilde{y}_{k+1}} \Gamma^{-1}_{\tilde{y}_{k+1}\tilde{y}_{k+1}} \tag{21}$$

$$L_k = (I - K_k \mathbb{E}[H_{k+1}])\mathbb{E}[A_k] - K_k \mathbb{E}[F_{k+1}], \tag{22}$$

where

$$\Gamma_{x_{k+1}\tilde{y}_{k+1}} = (\Sigma_{k+1} - \mathbb{E}[A_k]\Lambda_k \mathbb{E}[A_k^T])\mathbb{E}[H_{k+1}^T] \tag{23}$$

$$\begin{aligned} \Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}} = {} & \mathbb{E}[H_{k+1}\Sigma_{k+1}H_{k+1}^T] + \mathbb{E}[G_{k+1}G_{k+1}^T] \\ & + \mathbb{E}[F_{k+1}\Lambda_k F_{k+1}^T] - \mathbb{E}[F_{k+1}]\Lambda_k \mathbb{E}[F_{k+1}^T] \\ & - \mathbb{E}[H_{k+1}]\mathbb{E}[A_k]\Lambda_k \mathbb{E}[A_k^T]\mathbb{E}[H_{k+1}^T] \\ & + \mathbb{E}[H_{k+1}\mathbb{E}[A_k]\Lambda_k F_{k+1}^T] \\ & + \mathbb{E}[F_{k+1}\Lambda_k \mathbb{E}[A_k^T] H_{k+1}^T] \\ & - \mathbb{E}[H_{k+1}]\mathbb{E}[A_k]\Lambda_k \mathbb{E}[F_{k+1}^T] \\ & - \mathbb{E}[F_{k+1}]\Lambda_k \mathbb{E}[A_k^T]\mathbb{E}[H_{k+1}^T], \end{aligned} \tag{24}$$

and

$$\Sigma_{k+1} = \mathbb{E}[x_{k+1}x_{k+1}^T] = \mathbb{E}[A_k \Sigma_k A_k^T] + \mathbb{E}[C_k C_k^T] \tag{25}$$

$$\Lambda_{k+1} = L_k \Lambda_k \mathbb{E}[A_k^T] + K_k(\mathbb{E}[F_{k+1}]\Lambda_k \mathbb{E}[A_k^T] + \mathbb{E}[H_{k+1}]\Sigma_{k+1}). \tag{26}$$

III. Applications

In this section we present several classical problems and show how they may be formulated under the considered framework. We begin with the problem of tracking a maneuvering target and discuss the solutions obtained by using the IMM and the linear-optimal method. Next we consider tracking in clutter and extend the study performed in [13] by implementing the PDA approach using IMM. Finally, we discuss the multi-target tracking problem and consider several solutions.

A. Maneuvering Target Tracking

The problem of tracking a maneuvering target modeled as an MJLS is usually solved using the IMM approach. In this section we investigate the performance of linear-optimal trackers and compare it with that of IMM.

We consider the system (1) with $F(\theta_k) = 0$. The state vector $x_k = [p_k \ v_k \ a_k]^T$ comprises the target's position, velocity, and acceleration. At time k, the mode θ_k affects only the dynamics equation We consider the case $r = 2$ such that in the nominal regime, when $\theta_k = 1$, the system obeys the dynamics of the discrete white noise acceleration (DWNA) model [17], specified by the following matrices:

$$A(1) = \begin{pmatrix} 1 & T & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ C(1) = \begin{pmatrix} T^2/2 \\ T \\ 0 \end{pmatrix}\sigma_1, \tag{27}$$

where T is the sampling period corresponding to a single time step of the system (1), σ_1 is the nominal process noise intensity. In the maneuvering regime, when $\theta_k = 2$, the corresponding model is chosen to be the discrete Wiener process acceleration (DWPA) model [17], specified by the following matrices:

$$A(2) = \begin{pmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{pmatrix}, \ C(2) = \begin{pmatrix} T^2/2 \\ T \\ 1 \end{pmatrix}\sigma_2, \tag{28}$$

where σ_2 is the abnormal process noise intensity.

The transitions between the two modes occur according to a Markov process with the following TPM:

$$\Pi = \begin{pmatrix} p & 1-p \\ 1/3 & 2/3 \end{pmatrix}, \tag{29}$$

where p is a deterministic parameter. The initial mode distribution, ν, is taken to be the invariant distribution of (29).

In the experiment below the following common parameters were used: $\sigma_1 = 0.3 \text{ m/s}^2$, $\sigma_2 = 6 \text{ m/s}^2$, $T = 10$ s. These parameters, as well as the second row of the TPM (29), are adopted from [18] and are typical in aircraft tracking. The above system was simulated for 100 time steps.

The measurements are generated according to (1b) where $H(\theta_k) = [1 \ 0 \ 0]$, and $G(\theta_k) = \sigma_v = 1000$ m is the measurement noise standard deviation.

The above formulation allows a direct utilization of IMM and the linear optimal filter defined in Eqs. (11)-(19). Note that the linear filter for independently switching modes defined in Eqs. (21)-(26) is not applicable, since the target maneuvers according to a non-independent switching law. It does however, require smaller computational resources than the above alternatives. We thus compare the performance of the three methods and show that, at least in some cases, the performance loss in using the independence assumption is not drastic in comparison to the LMMSE filter of [12].

By sweeping the parameter p of (29) we compare the performance of the linear filter of Eqs. (21)-(26) (referred to as "IID") with that of the linear optimal filter of Eqs. (11)-(19) (refereed to as "Costa") and the nonlinear IMM. To adapt the IID filter to our Markovian scenario we set the distribution of θ_k, for each k, to the invariant distribution of (29). In addition, we run a "Genie" Kalman filter, that possesses perfect information on the mode value at each time. This ideal (but non-realistic) filter serves to provide an overall performance bound, indicating for each of the compared algorithms how far it is from the (unachievable) optimal performance.

Averaged over 1000 independent runs, the RMS position and velocity estimation errors, versus p, are presented in Fig. 1. It is readily seen that, for $p \in [0, 0.6]$, both linear methods attain similar performance. Moreover, both are only slightly outperformed by the IMM. On the other hand, as indicated by the errors of the Genie filter, there is room for improvement, meaning that the problem is not trivial. We thus conclude that, for this range of p, the independence assumption is reasonable, and the potential improvement that may be achieved by more complex nonlinear techniques is not significant.

For higher values of p, the Costa filter scores better than the IID filter. It should be noted that, although the position estimation accuracy of the IID filter degrades as p increases, at $p = 1$ (not shown here) all three methods coincide with the Genie filter, as in this case, the problem degenerates to a standard, single-mode estimation setting.

B. Target Tracking in Clutter

In this section we consider the problem of tracking a single, nonmaneuvering target in clutter and formulate it in the framework of the proposed generalized model.

1) Modeling: The state evolution is obtained from (1a) by setting $A(\theta_k) = A$ and $C(\theta_k) = C$, which results in

$$x_{k+1} = Ax_k + Cw_k. \tag{30}$$

Here A and C are deterministic matrices representing the state dynamics and process noise covariance, respectively.

At time k the target state is observed via the following linear measurement equation:

$$y_{k,\text{true}} = H_{\text{nom}} x_k + G_{\text{nom}} v_{k,\text{true}}. \tag{31}$$

Here, $y_{k,\text{true}}$ and $v_{k,\text{true}}$ represent the true measurement of the target and the true measurement noise, respectively. It is also assumed that the target may go undetected at some sampling intervals. This phenomenon is captured by the detection probability P_D.

In addition to $y_{k,\text{true}}$, at each time, a number of clutter measurements are obtained. These will be denoted as $y_{k,\text{cl}}$. Clutter measurements originate from false (or ghost) targets and do not carry any information about the target of interest. They are also indistinguishable from true detections. 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 these measurements are uniformly distributed in space.

Instead of scanning the entire surveillance region, the sensor initiates a validation window, centered at the predicted target position, and the algorithm processes the measurements obtained within the window. Since the clutter is uniformly distributed in space, it is also uniformly distributed within the validation window. The probability that the true measurement is inside the validation window is P_G. For simplicity of exposition, we assume that the true measurement is always present in the validation window. In other words, we assume that $P_D = P_G = 1$. In terms of the modeling, this is not a very restrictive assumption, and it may be easily relaxed, as we discuss in the sequel.

In this setting, the acquired measurement vector at time k becomes, for some integer N,

$$y_k = \left((y_k^1)^T, (y_k^2)^T, \ldots, (y_k^N)^T\right)^T. \tag{32}$$

Namely, y_k is a concatenation of N measurements, such that $N-1$ of them originate from false targets, or clutter, and only one originates from the (single) true target. The false measurements are centered around the predicted target state (since this is the center of the validation window), as opposed to the true measurement, which is generated using the true target state (31).

The described observation model follows from (1b) by considering an independent mode sequence $\{\theta_k\}$ taking values in $\{1, \ldots, N\}$ with the corresponding probabilities $\{p_1, \ldots, p_N\}$, and affecting the matrices of the measurement equation in the following manner:

$$\begin{aligned} &\{H(\theta_k), G(\theta_k), F(\theta_k)\} \\ &= \begin{cases} \left\{ \begin{pmatrix} H_{\text{nom}} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \text{diag} \begin{pmatrix} G_{\text{nom}} \\ G_{\text{cl}} \\ \vdots \\ G_{\text{cl}} \end{pmatrix}, \begin{pmatrix} 0 \\ H_{\text{nom}} A \\ \vdots \\ H_{\text{nom}} A \end{pmatrix} \right\}, & \theta_k = 1 \\ \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots & \\ \left\{ \begin{pmatrix} 0 \\ \vdots \\ 0 \\ H_{\text{nom}} \end{pmatrix}, \text{diag} \begin{pmatrix} G_{\text{cl}} \\ \vdots \\ G_{\text{cl}} \\ G_{\text{nom}} \end{pmatrix}, \begin{pmatrix} H_{\text{nom}} A \\ \vdots \\ H_{\text{nom}} A \\ 0 \end{pmatrix} \right\}, & \theta_k = N. \end{cases} \end{aligned} \tag{33}$$

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

Consider, for example, the first realization $\{H(1), G(1), F(1)\}$ in (33). It corresponds to the case where the first of the N acquired measurements is the true target measurement, $y_{k,\text{true}}$, generated according to (31). All

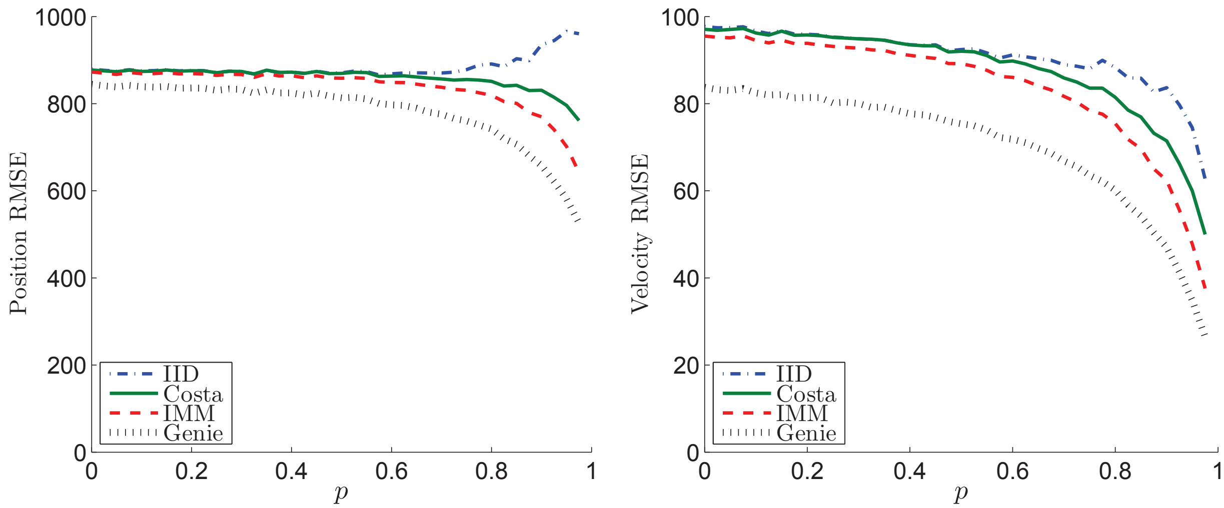

Fig. 1: RMS position (left) and velocity (right) errors vs. p in the maneuvering target tracking application.

other $N-1$ measurements are clutter, the ith of which is generated according to the following model:

$$y^i_{k,\mathrm{cl}} = H_{\mathrm{nom}} A \hat{x}_{k-1} + G_{\mathrm{cl}} v^i_{k,\mathrm{cl}}. \tag{34}$$

Note that $H_{\mathrm{nom}} A \hat{x}_{k-1}$ is the predicted true measurement at time k, which is also the center of the validation window set by the sensor. Thus, clutter measurements acquired at time k may be viewed as generated uniformly around this location.

That exactly one of the N observations is target-originated is reflected in (33) by the fact that exactly one of the blocks of $H(\theta_k)$ is set to H_{nom}, with all others being set to 0. Likewise, all but one block of F_k are taken to be $H_{\mathrm{nom}} A$.

Captured by the matrix G_{cl}, the covariance of the clutter measurements may be different from the true measurement noise covariance, represented by G_{nom}. However, it is not known a-priori which of the concatenated measurements carries useful information. We assume that all possible orderings of the true and clutter observations in the validation window are equiprobable and thus set

$$p_i = \frac{1}{N}, \quad i = 1, \ldots, N. \tag{35}$$

Hence $H(\theta_k)$, $G(\theta_k)$, and $F(\theta_k)$ correspond to random permutations of N possible positions of the true measurement among the clutter measurements. Note that the overall number of validated measurements (i.e., those that are in the validation window), N, is assumed to be known, but may vary in time.

2) Implementing PDA with IMM: We now show that for $P_D = P_G = 1$, the popular PDA algorithm is identical to the IMM scheme with GKF modules. Since the mode sequence is independent, the interaction block of IMM becomes obsolete and the whole procedure reduces to a standard GPB. Hence, we need to show that a standard GPB with modes defined in (33) reduces to the PDA routine.

Consider the final state estimate produced by GPB:

$$\hat{x}^{\mathrm{GPB}}_{k+1} = \sum_{j=1}^{r} \hat{x}^j_{k+1} \mu_j(k+1), \tag{36}$$

where r is the number of modes, $\hat{x}^j_{k+1}$ is the output of the primitive KF (GKF) matched to model j and $\mu_j(k+1) = \mathbb{P}\{\theta_{k+1} = j \mid \mathcal{Y}_{k+1}\}$. In the considered case $r = N$. The j-th mode-conditioned estimate, $\hat{x}^j_{k+1}$, is given by the GKF update defined in Eqs. (7)-(10). Consider the estimate matched to the j-th realization of θ_{k+1} in (33):

$$\hat{x}^j_{k+1} = (A - K^j_k H(j) A - K^j_k F(j)) \hat{x}^{\mathrm{GPB}}_k + K^j_k y_{k+1}. \tag{37}$$

Inspecting (9) it is easy to see that K^j_k has the following form

$$(\underbrace{0 \cdots 0}_{j-1}\; K_{k,\mathrm{nom}}\; \underbrace{0 \cdots 0}_{N-j}),$$

where

$$K_{k,\mathrm{nom}} = P^-_{k+1} H^T_{\mathrm{nom}} \left(H_{\mathrm{nom}} P^-_{k+1} H^T_{\mathrm{nom}} + G_{\mathrm{nom}} G^T_{\mathrm{nom}} \right)^{-1}.$$

Thus, $K^j_k F(j) = 0$ and (36) reads

$$\begin{aligned} \hat{x}^{\mathrm{GPB}}_{k+1} &= (A - K_{k,\mathrm{nom}} H_{\mathrm{nom}} A) \hat{x}^{\mathrm{GPB}}_k \\ &\quad + K_{k,\mathrm{nom}} \sum_{j=1}^{N} \mu_j(k+1) y^j_{k+1}. \end{aligned} \tag{38}$$

Comparing (38) with (7) we note that the former is a KF-like update step with an effective measurement computed as a weighted sum of all the observations acquired at the current time step. The above form is essentially the same as the one used in PDA, meaning that showing equivalence between the two methods boils down to showing that $\{\mu_j(k+1)\}^r_{j=1}$ are the same as the PDA's weighting probabilities. Indeed,

$$\begin{aligned}\mu_j(k+1) &= \mathbb{P}\{\theta_{k+1}=j \mid \mathcal{Y}_{k+1}\} \\ &= \frac{1}{c}\mathrm{p}(y_{k+1} \mid \theta_{k+1}=j,\mathcal{Y}_k)\,\mathbb{P}\{\theta_{k+1}=j \mid \mathcal{Y}_k\} \\ &= \frac{1}{c}\mathrm{p}(y_{k+1} \mid \theta_{k+1}=j,\mathcal{Y}_k)\,p_j. \end{aligned} \tag{39}$$

The term $\mathrm{p}(y_{k+1} \mid \theta_{k+1}=j,\mathcal{Y}_k)$ is the likelihood of the latest measurement set, among which the j-th observation is the true measurement having a (truncated) Gaussian distribution about the true target state, and the rest are clutter measurements uniformly distributed about the predicted measurement in the validation window. Alternatively, the uniform distribution is specified by the covariance matrix G_{cl}. In the present case we assume that the true target is always present in the validation window, meaning that the probability of detection is set to unity and the window is infinite. Hence, G_{cl} must be taken to infinity and (39) is completely specified by the likelihood of the true measurement coinciding with the weighting factors of PDA for $P_D = P_G = 1$. This completes the proof on the equivalence of the two methods.

In the case where the target may either go undetected, or a finite validation window is used, the set of possible modes should be augmented by $\theta_k = 0$, such that

$$\{H(0), G(0), F(0)\} = \{\mathbf{0}, I_N \otimes G_{\mathrm{cl}}, \mathbf{1}_N \otimes H_{\mathrm{nom}}A\}\,. \tag{40}$$

Here, the symbol $\otimes$ 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 prior probability distribution of the mode should be modified in a straightforward manner.

3) Linear Optimal Filter: Since the mode sequence (33) is independent, we utilize the recursive version of the linear optimal filter given in Eqs. (21)-(26). To implement the algorithm one needs to compute the expectations appearing in the above equations. Although these may be computed numerically, via direct summations, in the present application closed form solutions exist as follows [13]:

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

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

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

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

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

where Ξ is defined by

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

Finally,

$$\begin{aligned}&\mathbb{E}\left[H_{k+1}A\Lambda_k F_{k+1}^T\right] \\ &\quad= \frac{1}{N}(\mathbf{1}_N\mathbf{1}_N^T - I_N) \otimes \left(H_{\mathrm{nom}}A\Lambda_k A^T H_{\mathrm{nom}}^T\right). \end{aligned} \tag{47}$$

4) Numerical Example: In this section we demonstrate the performance of the linear optimal filter for tracking a target in clutter by comparing it with that of the standard Nearest Neighbor (NN) [18] method and PDA. We consider a one-dimensional tracking scenario, in which a target is represented via a two dimensional state comprising position and velocity information, $x_k = (p_k,\, v_k)^T$. Starting at $x_0 = (0,\,0)^T$ with $P_0 = \left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$, the target is simulated for 1000 time units by running the dynamical equation (30) with

$$A = \begin{pmatrix}0.95 & 0.2\\ 0 & 0.95\end{pmatrix}, \quad C = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}. \tag{48}$$

The true measurement is generated by computing (31) with

$$H_{\mathrm{nom}} = (1 \quad 0), \quad G_{\mathrm{nom}} = 0.32. \tag{49}$$

As mentioned earlier, for the IMM version of PDA we assume an infinite validation window, taking advantage of the inherent soft validation property of the PDA. For LMMSE such option is not viable since, as shown in [13], the LMMSE for tracking in clutter implements a KF-like procedure on the average measurement in the validation window. Hence, we enforce externally a validation window which is twice as large as the innovation covariance (24).

We use two measures of performance to compare the algorithms. The first is the time until the target is lost, which is defined as the third time the distance between the predicted position and the true state has deviated by more than five standard deviations of the (true) measurement noise. The second measure is the average squared error calculated over the time interval until the first of the three algorithms has lost track. This makes the comparison fair, in the sense that none of the algorithms incorporates meaninglessly large errors corresponding to a lost target.

We define 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 RMS position errors and track loss times, versus ρ, are plotted in Fig. 2. It is readily seen that the linear optimal filter attains the longest track loss time while keeping the estimation errors at a reasonable compromise between the nonlinear IMM/PDA and NN algorithms.

C. Multiple Target Tracking

Here we consider N independently evolving targets, such that the dynamics of target i is described by

$$x_{k+1}^i = A^i x_k^i + C^i w_k^i, \tag{50}$$

where A^i and C^i are target specific matrices, and $\{w_k^i\}$ is the process noise sequence of target i, which is assumed to have zero mean and identity covariance matrix.

This (multitarget) state evolution is obtained from (1a) by defining an *augmented state* x_k as a concatenation of the individual states $x_k^i,\ i=1,\ldots,N$. The dynamics and process noise matrices, A_k and C_k, respectively, are:

$$A(\theta_k) = \mathrm{diag}\left(A^1 \quad A^2 \quad \cdots \quad A^N\right) \tag{51}$$

$$C(\theta_k) = \mathrm{diag}\left(C^1 \quad C^2 \quad \cdots \quad C^N\right). \tag{52}$$

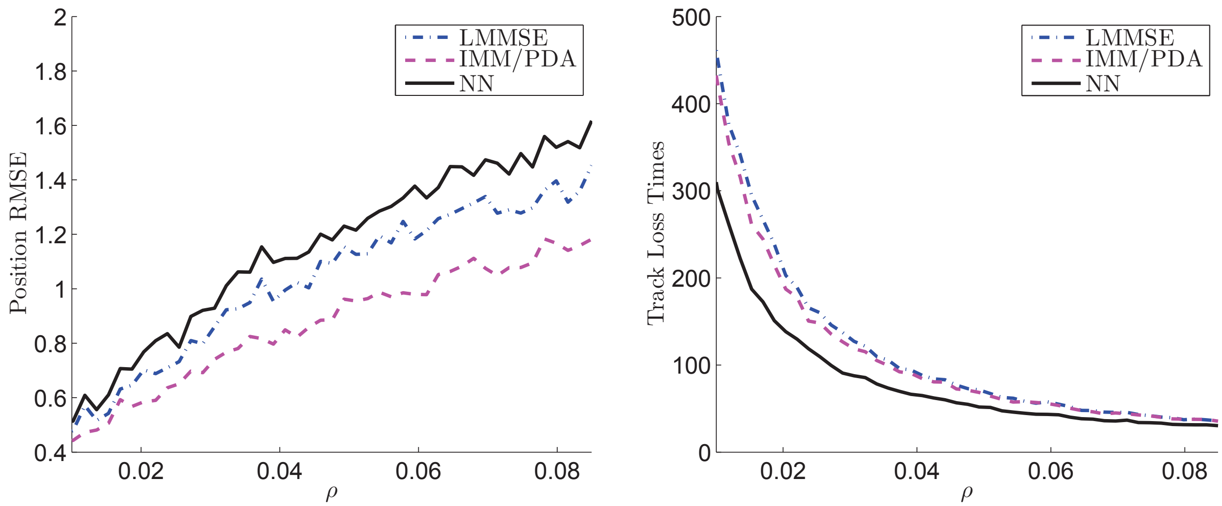

Fig. 2: Position RMSE (left) and track loss times (right) vs. clutter density in the target tracking in clutter application.

The state of target i is observed via the following linear measurement equation:

$$y_k^i = H_{\mathrm{nom}} x_k^i + G^i v_k^i, \tag{53}$$

where H_{nom} is a known deterministic matrix representing sensor geometry and G^i represents the measurement noise covariances, which may depend on the target.

At time k each of the targets generates a single measurement, such the measurement vector at time k is a concatenation of the N measurements. It is not known, however, which measurement corresponds to which target, such that any measurements-to-targets association is possible. Similarly to the case of tracking in clutter, we assume that any such association is equiprobable, and independent of the past ones.

The described observation model follows from (1b) by defining the following set of values taken, with equal probability, by the mode-affected matrices $\{H(\theta_k), G(\theta_k)\}$:

$$\{H(\theta_k), G(\theta_k)\} \in \left\{ \begin{array}{c} \left\{ \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \otimes H_{\mathrm{nom}}, \begin{pmatrix} G^1 & 0 & \cdots & 0 \\ 0 & G^2 & 0 & \cdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & G^N \end{pmatrix} \right\}, \\ \left\{ \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \otimes H_{\mathrm{nom}}, \begin{pmatrix} 0 & G^1 & \cdots & 0 \\ G^2 & 0 & 0 & \cdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & G^N \end{pmatrix} \right\}, \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \left\{ \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \iddots & & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix} \otimes H_{\mathrm{nom}}, \begin{pmatrix} 0 & \cdots & 0 & G^1 \\ 0 & 0 & G^2 & 0 \\ \vdots & \iddots & & \vdots \\ G^N & 0 & \cdots & 0 \end{pmatrix} \right\} \end{array} \right\}. \tag{54}$$

Thus, the set of values taken by $H(\theta_k)$ is obtained by permuting the rows of an $N \times N$ identity matrix and performing a Kronecker product of every such matrix with H_{nom}, and the set of values taken by $G(\theta_k)$ is obtained by permuting the block-rows of the block-diagonal matrix $\mathrm{diag}\begin{pmatrix} G^1 & G^2 & \cdots & G^N \end{pmatrix}$. In the present example we set, for simplicity, $F(\theta_k) = 0$.

Consider, for example, the case where $N = 2$ targets are to be tracked in the linear optimal sense. The first realization of $\{H(\theta_k), G(\theta_k)\}$ in (54) then reads

$$\{H(\theta_k), G(\theta_k)\} = \left\{ \begin{pmatrix} H_{\mathrm{nom}} & 0 \\ 0 & H_{\mathrm{nom}} \end{pmatrix}, \begin{pmatrix} G^1 & 0 \\ 0 & G^2 \end{pmatrix} \right\}. \tag{55}$$

This realization corresponds to the case where the first of the 2 acquired measurements originates from the first target and the second observation originates from the second target. Due to the data ambiguity, each of the two hypotheses has a probability of $\frac{1}{2}$ of being correct. For general N, the number of possible hypotheses is $N!$ and the corresponding probability of each one being correct is $\frac{1}{N!}$.

To track the targets with IMM, we define $N!$ primitive KFs capturing the different association hypotheses. The implementation of the LMMSE filter is straightforward. We note that using the IID filter is possible due to the assumed independence of the modes, and in this case closed form expressions exist for the required expectations:

$$\mathbb{E}\left[H_{k+1}\right] = \frac{1}{N}(\mathbf{1}_N^\top \mathbf{1}_N) \otimes H_{\mathrm{nom}}, \tag{56}$$

$$\mathbb{E}\left[H_{k+1}\Sigma_{k+1}H_{k+1}^T\right] = \frac{1}{N} I \otimes (H_{\mathrm{nom}}(\mathbf{1}_N^T \otimes I)\Sigma_k(\mathbf{1}_N \otimes I)H_{\mathrm{nom}}^T), \tag{57}$$

$$\mathbb{E}\left[G_{k+1}G_{k+1}^T\right] = \mathrm{diag}\begin{pmatrix} G^1(G^1)^T & \cdots & G^N(G^N)^T \end{pmatrix}. \tag{58}$$

To illustrate the feasibility of the approach we track two crossing targets, as shown in Fig. 3, where the position estimates of the LMMSE filter are compared with those of

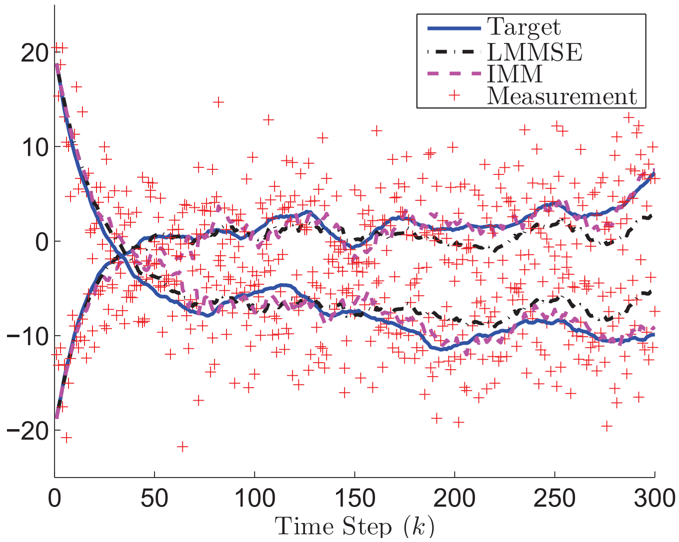

Fig. 3: Target positions accompanied by estimated trajectories and raw measurements vs. time.

IMM. The matrices used in the example are

$$A^i = \begin{pmatrix} 0.95 & 0.2 \\ 0 & 1 \end{pmatrix}, \; C^i = 0.2 \begin{pmatrix} 0.5 \\ 1 \end{pmatrix}, \; i = 1, 2, \tag{59}$$

and

$$H_{\text{nom}} = (1 \quad 0), \quad G_{\text{nom}} = \sqrt{20}, \quad x_0^1 = -x_0^2 = (20 \; -1)^T.$$

It is readily seen that the LMMSE filter is capable of maintaining tracks. It is, however, inferior to the more sophisticated IMM, which may be shown to implement, in this case, the JPDA filter similarly to the implementation of PDA in the previous example. However, IMM and JPDA require, at each time step, the calculation of $N!$ Kalman updates (assuming that N measurements need to be assigned to N targets) and, therefore, they are infeasible even for scenarios with a small number of targets. The computational requirements of the LMMSE filter, on the other hand, are at the bare minimum and it only requires the inversion of an $mN \times mN$ matrix, where m is the measurement dimension.

IV. Conclusion

We presented a unified modeling of state estimation problems under data and model uncertainties. This allowed a representation of several classical problems within a single state-space formulation with random matrix coefficients. Consequently, a standard IMM and LMMSE filters were applied to solve the maneuvering target tracking problem as well as tracking under data ambiguity – target tracking in clutter and multiple target tracking. In the case of tracking in clutter we showed that the resulting IMM scheme is equivalent to the classical PDA approach. We note that additional state estimation problems may be cast in the considered framework. These include maneuvering target tracking in clutter and multiple maneuvering target tracking.

Acknowledgment

This research was supported by the Israeli Science Foundation under grant No. 1608/11 and the Technion's L. Kraus Research Fund.

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

[5] A. Jaffer and S. Gupta, "Optimal sequential estimators of discrete processes with markov interrupted observations," *IEEE Trans. Inf. Theory*, vol. IT-17, no. 6, pp. 614–616, September 1971.

[6] I. Rapoport and Y. Oshman, "A new estimation error lower bound for interruption indicators in systems with uncertain measurements," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3375–3384, 2004.

[7] D. Sigalov and Y. Oshman, "A new formulation of fault-tolerant estimation problems and some solutions," in *Proc. 26th Convention of Electrical and Electronis Engineers in Israel.*, 2011.

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

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

[10] 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.

[11] 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.

[12] 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.

[13] 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.

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

[15] T. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," *IEEE J. Ocean. Eng.*, vol. 8, no. 3, pp. 173–184, 1983.

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

[17] Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York : Wiley, 2001.

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