Generalized Innovation and Inference Algorithms for Hidden Mode Switched Linear Stochastic Systems with Unknown Inputs

Sze Zheng Yong[a] Minghui Zhu[b] Emilio Frazzoli[a]

Abstract— **In this paper, we propose inference algorithms for simultaneously estimating the mode, input and state of hidden mode switched linear stochastic systems with unknown inputs. First, we define the generalized innovation for the recently proposed optimal filter for simultaneous input and state estimation [1] and show that the sequence is a Gaussian white noise. Then, we utilize this whiteness property of the generalized innovation, which reflects the estimation quality to form the likelihood function of the system model. Consequently, we employ the multiple model (MM) approach based on the likelihood function for inferring the hidden mode of switched linear stochastic systems. Algorithms for both static and dynamic MM estimation are presented and compared using a simulation example of vehicles at an intersection with switching driver intentions.**

I. Introduction

Most autonomous systems must operate without knowledge of the intention and the decisions of other non-communicating systems or humans. Thus, in many instances, these intention and control decisions need to be inferred from noisy measurements. This problem can be conveniently considered within the framework of *hidden mode hybrid systems* (HMHS, see, e.g., [2], [3] and references therein) with unknown inputs, in which the system *state* dynamics is described by a finite collection of functions. Each of these functions corresponds to an intention or *mode* of the hybrid system, where the mode is unknown or *hidden* and mode transitions are autonomous. In addition, by allowing unknown inputs in this framework, both deterministic and stochastic disturbance inputs and noise can also be considered. There are a large number of applications, such as urban transportation systems, target tracking and fault detection, in which it is not realistic to assume knowledge of the mode and disturbance inputs or they are too costly to measure.

Literature review. The filtering problem of hidden mode hybrid systems without unknown inputs have been extensively studied (see, e.g., [4], [5] and references therein), especially in the context of target tracking applications. These filtering algorithms, which use a multiple model approach, consist of three components: (i) a bank of filters for each mode, (ii) a likelihood-based approach to determine the probability of each mode, and (iii) a hypothesis management algorithm to trade off between computational cost and estimation quality. Oftentimes, the Kalman filter [6] or a variation thereof is used as the filtering algorithm, while the likelihood-based mode association typically uses the whiteness property of the innovation [7], [8]. On the other hand, some popular hypothesis merging algorithms include the first- and second-order generalized pseudo-Bayesian (GPB1 and GPB2) as well as the interacting multiple model (IMM) algorithms [4], [9]. However, oftentimes the disturbance inputs cannot be modeled as a zero-mean, Gaussian white noise, which gives rise to a need for an extension of the existing algorithms to hidden mode hybrid systems with unknown inputs. To the best of our knowledge, this problem has not been addressed in the literature.

To develop a multiple model estimation approach for switched linear stochastic systems with unknown inputs, each component of the approach has to be reinvented. A filter algorithm for stochastic systems with unknown inputs is needed for the first component; thus, a relevant set of literature is that of optimal filters that simultaneously estimate inputs and states of linear stochastic systems with unknown inputs. Research on simultaneous input and state estimation has been gaining momentum during the past years, largely stimulated by its wide applications. Notable algorithms include [1], [10]–[12]. Of all these algorithms, the optimal filter in [1] is in the most general form and is hence the most suitable for the problem at hand. However, the remaining two major components of the multiple model estimation approach are still lacking at present.

Contributions. In this paper, we present a novel multiple model approach for simultaneous estimation of mode, input and state of switched linear stochastic systems with unknown inputs. As with multiple model estimation of systems without unknown inputs, we also present static and dynamic variants (i.e., decoupled and with cooperation between models, respectively) of the estimation algorithm. In both variants, a bank of optimal input and state filters [1], one for each mode, is run in parallel. Next, we devise a likelihood-based mode association algorithm to determine the probability of each mode. This involves the definition of a *generalized innovation* which we prove is a Gaussian white noise. Then, we use this whiteness property to form a likelihood function, which is used in hypothesis testing. For the dynamic variant (with cooperating models), to manage the growing number of hypotheses, we employ a similar approach to the interacting multiple model estimator [9] which mixes the initial conditions based on mode transition matrix probabilities.

II. Motivating Example

To motivate the problem considered in this paper, we consider the scenario of vehicles crossing a 4-way intersection where each vehicle does not have any information

[a] S.Z Yong and E. Frazzoli are with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA (e-mail: szyong@mit.edu, frazzoli@mit.edu).

[b] M. Zhu is with the Department of Electrical Engineering, Pennsylvania State University, 201 Old Main, University Park, PA 16802, USA (e-mail: muz16@psu.edu).

about the intention of the other vehicles. To simplify the problem, we consider the case with two vehicles: Vehicle A is human driven (uncontrolled) and Vehicle B is autonomous (controlled), with dynamics described by $\ddot{x}_A = -0.1\dot{x}_A + d_1$ and $\ddot{x}_B = -0.1\dot{x}_B + u$, where x and $\dot{x}$ are vehicle positions and velocities. We assume[1] that Vehicle A approaches the intersection with a default intention, i.e., without considering the presence of Vehicle B. Then, at the intersection, the driver of Vehicle A can choose between three intentions:

- to continue while ignoring the other vehicle with an unknown input d_1 (Inattentive Driver, default mode),
- to attempt to cause a collision (Malicious Driver), or
- to stop (Cautious Driver).

Then, once either vehicle completes the crossing of the intersection, Vehicle A returns to the default intention.

Thus, if we assume the presence of noise, this intersection-crossing scenario is an instance of a hidden mode switched linear stochastic system with an unknown input. The intention of driver A is a hidden mode and the actual input of vehicle A is an unknown input (which is not restricted to a finite set). The objective is to simultaneously estimate the intention (mode), input and state of the vehicles for safe navigation through the intersection.

III. Problem Statement

We consider a hidden mode switched linear stochastic system with unknown inputs:

$$\begin{aligned}(x_{k+1}, q_k) &= (A_k^{q_k} x_k + B_k^{q_k} u_k^{q_k} + G_k^{q_k} d_k^{q_k} + w_k^{q_k}, q_k), & x_k \in \mathcal{C}_{q_k}\\ (x_k, q_k)^+ &= (x_k, \delta^{q_k}(x_k)), & x_k \in \mathcal{D}_{q_k}\\ y_k &= C_k^{q_k} x_k + D_k^{q_k} u_k^{q_k} + H_k^{q_k} d_k^{q_k} + v_k^{q_k} & \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the continuous system state and $q_k \in \{1, 2, \ldots, N\}$ the hidden discrete state or *mode*. For each mode q_k, $u_k^{q_k} \in U_{q_k} \subset \mathbb{R}^m$ is the known input, $d_k^{q_k} \in \mathbb{R}^p$ the unknown input, $y \in \mathbb{R}^l$ the output, $\delta^{q_k}(\cdot)$ the mode transition function, $\mathcal{C}_{q_k}$ and $\mathcal{D}_{q_k}$ are flow and jump sets, while the process noise $w_k^{q_k} \in \mathbb{R}^n$ and the measurement noise $v_k^{q_k} \in \mathbb{R}^l$ are assumed to be mutually uncorrelated, zero-mean, Gaussian white random signals with known covariance matrices, $Q_k^{q_k} = \mathbb{E}[w_k^{q_k} w_k^{q_k \top}] \succeq 0$ and $R_k^{q_k} = \mathbb{E}[v_k^{q_k} v_k^{q_k \top}] \succ 0$, respectively. The matrices $A_k^{q_k}$, $B_k^{q_k}$, $G_k^{q_k}$, $C_k^{q_k}$, $D_k^{q_k}$ and $H_k^{q_k}$ are known. x_0 is assumed to be independent of $v_k^{q_k}$ and $w_k^{q_k}$ for all k. No prior 'useful' knowledge of the dynamics of $d_k^{q_k}$ is assumed (uncorrelated with $\{d_j^{q_j}\}$, $\forall j \neq k$, as well as $\{w_j^{q_j}\}$ and $\{v_j^{q_j}\}$, $\forall j$) and $d_k^{q_k}$ can be a signal of any type.

Moreover, the mode jump process is assumed to be left-continuous and hidden mode systems refer to systems in which q_k is not directly measured and the mode transitions are autonomous. We assume that in each mode, the system has *strong detectability*, i.e., the initial condition x_0 and the unknown input sequence $\{d_j^{q_j}\}_{j=0}^{r-1}$ can be uniquely determined from the measured output sequence $\{y_i\}_{j=0}^{r}$ of a sufficient number of observations, i.e., $r \geq r_0$ for some $r_0 \in \mathbb{N}$ (see [1, Section 3.2] for necessary and sufficient conditions for this property) and the required rank condition for the existence of a stable filter [1, Theorem 9] is satisfied.

[1]The assumed permutation of intentions is for illustrative purposes only and was not a result of any limitations on the proposed algorithms.

The objective of this paper[2] is to design an optimal recursive filter algorithm which simultaneously estimates the system state x_k, the unknown input $d_k^{q_k}$ and the hidden mode q_k based on the measurements up to time k, $\{y_0, y_1, \ldots, y_k\}$.

IV. Preliminary Material

In this section, we present a brief summary of the minimum-variance unbiased filter for linear systems with unknown inputs. For detailed proof and derivation of the filter, the reader is referred to [1]. Moreover, we define a *generalized innovation* and show that it is a Gaussian white noise. These form an essential part of the multiple model estimation algorithm that we will describe in Section V. The algorithm runs a bank of $\mathfrak{N}$ filters (one for each mode) in parallel and each of the the filter are in essence the same except for the different sets of matrices and signals $\{A_k^{q_k}, B_k^{q_k}, C_k^{q_k}, D_k^{q_k}, G_k^{q_k}, H_k^{q_k}, Q_k^{q_k}, R_k^{q_k}, u_k^{q_k}, d_k^{q_k}\}$. Hence, to simplify notation, the conditioning on the mode q_k is omitted in the entire Section IV.

A. Minimum-Variance Unbiased Filter

As is shown in [1, Section 3.1], the system for each mode after a similarity transformation is given by:

$$x_{k+1} = A_k x_k + B_k u_k + G_{1,k} d_{1,k} + G_{2,k} d_{2,k} + w_k \tag{2}$$

$$z_{1,k} = C_{1,k} x_k + D_{1,k} u_k + \Sigma_k d_{1,k} + v_{1,k} \tag{3}$$

$$z_{2,k} = C_{2,k} x_k + D_{2,k} u_k + v_{2,k}. \tag{4}$$

The transformation essentially decomposes the unknown input d_k and the measurement y_k each into two orthogonal components, i.e., $d_{1,k} \in \mathbb{R}^{p_{H_k}}$ and $d_{2,k} \in \mathbb{R}^{p-p_{H_k}}$; as well as $z_{1,k} \in \mathbb{R}^{p_{H_k}}$ and $z_{2,k} \in \mathbb{R}^{l-p_{H_k}}$, where $p_{H_k} = \text{rank}(H_k)$. Then, given measurements up to time $k-1$, the optimal three-step recursive filter in the minimum-variance unbiased sense can be summarized as follows:

Unknown Input Estimation:

$$\begin{aligned}\hat{d}_{1,k} &= M_{1,k}(z_{1,k} - C_{1,k}\hat{x}_{k|k} - D_{1,k}u_k)\\ \hat{d}_{2,k-1} &= M_{2,k}(z_{2,k} - C_{2,k}\hat{x}_{k|k-1} - D_{2,k}u_k)\\ \hat{d}_{k-1} &= V_{1,k-1}\hat{d}_{1,k-1} + V_{2,k-1}\hat{d}_{2,k-1}\end{aligned} \tag{5}$$

Time Update:

$$\begin{aligned}\hat{x}_{k|k-1} &= A_{k-1}\hat{x}_{k-1|k-1} + B_{k-1}u_{k-1} + G_{1,k-1}\hat{d}_{1,k-1}\\ \hat{x}^\star_{k|k} &= \hat{x}_{k|k-1} + G_{2,k-1}\hat{d}_{2,k-1}\end{aligned} \tag{6}$$

Measurement Update:

$$\hat{x}_{k|k} = \hat{x}^\star_{k|k} + \overline{L}_k \Gamma_k (z_{2,k} - C_{2,k}\hat{x}^\star_{k|k} - D_{2,k}u_k) \tag{7}$$

where $\hat{x}_{k-1|k-1}$, $\hat{d}_{1,k-1}$, $\hat{d}_{2,k-1}$ and $\hat{d}_{k-1}$ denote the optimal estimates of x_{k-1}, $d_{1,k-1}$, $d_{2,k-1}$ and d_{k-1}; $\Gamma_k \in \mathbb{R}^{p_{\tilde{R}} \times l - p_{H_k}}$ is a design matrix that is chosen to project the "innovation" $\overline{\nu}_k := z_{2,k} - C_{2,k}\hat{x}^\star_{k|k} - D_{2,k}u_k$ onto a vector of $p_{\tilde{R}}$ independent random variables, while $\overline{L}_k \in \mathbb{R}^{n \times p_{\tilde{R}}}$, $M_{1,k} \in \mathbb{R}^{p_{H_k} \times p_{H_k}}$ and $M_{2,k} \in \mathbb{R}^{(p-p_{H_k}) \times (l-p_{H_k})}$ are filter gain matrices that minimize the state and input error covariances. For the sake of completeness, the optimal input and state filter in [1] is reproduced in Algorithm 1.

[2]Due to space limitation, a technical characterization of the inference algorithm will be presented in an upcoming companion paper [13].

Algorithm 1 Opt-Filter (q_k, $\hat{x}^{0,q_k}_{k-1|k-1}$, $\hat{d}^{0,q_k}_{1,k-1}$, $P^{x,0,q_k}_{k-1|k-1}$, $P^{d,0,q_k}_{1,k-1}$) [superscript q_k omitted in the following]

▷ Estimation of $d_{2,k-1}$ and d_{k-1}

$\hat{A}_{k-1} = A_{k-1} - G_{1,k-1}M_{1,k-1}C_{1,k-1}$;

$\hat{Q}_{k-1} = G_{1,k-1}M_{1,k-1}R_{1,k-1}M^\top_{1,k-1}G^\top_{1,k-1} + Q_{k-1}$;

$\tilde{P}_k = \hat{A}_{k-1}P^{x,0}_{k-1|k-1}\hat{A}^\top_{k-1} + \hat{Q}_{k-1}$;

$\tilde{R}_{2,k} = C_{2,k}\tilde{P}_kC^\top_{2,k} + R_{2,k}$;

$P^d_{2,k-1} = (G^\top_{2,k-1}C^\top_{2,k}\tilde{R}^{-1}_{2,k}C_{2,k}G_{2,k-1})^{-1}$;

$M_{2,k} = P^d_{2,k-1}G^\top_{2,k-1}C^\top_{2,k}\tilde{R}^{-1}_{2,k}$;

$\hat{x}_{k|k-1} = A_{k-1}\hat{x}^0_{k-1|k-1} + B_{k-1}u_{k-1} + G_{1,k-1}\hat{d}^0_{1,k-1}$;

$\hat{d}_{2,k-1} = M_{2,k}(z_{2,k} - C_{2,k}\hat{x}_{k|k-1} - D_{2,k}u_k)$;

$\hat{d}_{k-1} = V_{1,k-1}\hat{d}^0_{1,k-1} + V_{2,k-1}\hat{d}_{2,k-1}$;

$P^d_{12,k-1} = M_{1,k-1}C_{1,k-1}P^{x,0}_{k-1|k-1}A^\top_{k-1}C^\top_{2,k}M^\top_{2,k} - P^{d,0}_{1,k-1}G^\top_{1,k-1}C^\top_{2,k}M^\top_{2,k}$;

$P^d_{k-1} = V_{k-1}\begin{bmatrix} P^{d,0}_{1,k-1} & P^d_{12,k-1} \\ P^{d\top}_{12,k-1} & P^d_{2,k-1} \end{bmatrix} V^\top_{k-1}$;

▷ Time update

$\hat{x}^\star_{k|k} = \hat{x}_{k|k-1} + G_{2,k-1}\hat{d}_{2,k-1}$;

$P^{\star x}_{k|k} = G_{2,k-1}M_{2,k}R_{2,k}M^\top_{2,k}G^\top_{2,k} + (I - G_{2,k-1}M_{2,k}C_{2,k})\tilde{P}_k(I - G_{2,k-1}M_{2,k}C_{2,k})^\top$;

$\tilde{R}^\star_{2,k} = C_{2,k}P^{\star x}_{k|k}C^\top_{2,k} + R_{2,k} - C_{2,k}G_{2,k-1}M_{2,k}R_{2,k} - R_{2,k}M^\top_{2,k}G^\top_{2,k-1}C^\top_{2,k}$;

▷ Measurement update

$\tilde{L}_k = (P^{\star x}_{k|k}C^\top_{2,k} - G_{2,k-1}M_{2,k}R_{2,k})\tilde{R}^{\star\dagger}_{2,k}$;

$\hat{x}_{k|k} = \hat{x}^\star_{k|k} + \tilde{L}_k(z_{2,k} - C_{2,k}\hat{x}^\star_{k|k} - D_{2,k}u_k)$;

$P^x_{k|k} = (I - \tilde{L}_kC_{2,k})G_{2,k-1}M_{2,k}R_{2,k}\tilde{L}^\top_k + \tilde{L}_kR_{2,k}M^\top_{2,k}G^\top_{2,k-1}(I - \tilde{L}_kC_{2,k})^\top + (I - \tilde{L}_kC_{2,k})P^{\star x}_{k|k}(I - \tilde{L}_kC_{2,k})^\top + \tilde{L}_kR_{2,k}\tilde{L}^\top_k$;

▷ Estimation of $d_{1,k}$

$\tilde{R}_{1,k} = C_{1,k}P^x_{k|k}C^\top_{1,k} + R_{1,k}$;

$M_{1,k} = \Sigma^{-1}_k$;

$P^d_{1,k} = M_{1,k}\tilde{R}_{1,k}M^\top_{1,k}$;

$\hat{d}_{1,k} = M_{1,k}(z_{1,k} - C_{1,k}\hat{x}_{k|k} - D_{1,k}u_k)$;

B. Properties of the Generalized Innovation Sequence

In Kalman filtering, the innovation reflects the difference between the measured output at time k and the optimal output forecast based on information available prior to time k. The *a posteriori* (updated) state estimate is then a linear combination of the *a priori* (predicted) estimate and the weighted innovation. In the same spirit, we generalize this notion of innovation to linear systems with unknown inputs by defining a *generalized innovation* given by:

$$\begin{aligned}\nu_k &:= \Gamma_k(z_{2,k} - C_{2,k}\hat{x}^\star_{k|k} - D_{2,k}u_k) := \Gamma_k\overline{\nu}_k \qquad (8)\\ &= \Gamma_k(I - C_{2,k}G_{2,k-1}M_{2,k})(z_{2,k} - C_{2,k}\hat{x}_{k|k-1} - D_{2,k}u_k)\end{aligned}$$

which, similar to the conventional innovation, is weighted by $\overline{L}_k$ and combined with the predicted state estimate $\hat{x}^\star_{k|k}$ to obtain the updated state estimate $\hat{x}_{k|k}$ as seen in (7). This definition differs from the conventional innovation in that the generalized innovation uses a subset of the measured outputs, i.e. $z_{2,k}$. In addition, the matrix Γ_k is any matrix whose rows are independent of each other and are in the range space of $\mathbb{E}[\overline{\nu}_k\overline{\nu}_k^\top]$ that removes dependent components of $\overline{\nu}_k$ (a consequence of [1, Lemma 17]), which further lowers the dimension of the generalized innovation. An intuition for this is that the information contained in the 'unused' subset is already exhausted for estimating the unknown inputs. Moreover, the optimal output forecast that is implied in (8) is a function of $\hat{x}^\star_{k|k}$ which contains information from the measurement at time k. Nonetheless, it is clear from (8) that when there are no unknown inputs, both $C_{2,k}$ and $G_{2,k-1}$ are empty and Γ_k can be chosen to be the identity matrix, in which case the definitions of generalized innovation and (conventional) innovation coincide.

In the following theorem, we establish that the generalized innovation, like the innovation, is a Gaussian white noise.

Theorem 1. *The generalized innovation, ν_k given in* (8) *is a Gaussian white noise with zero mean and a variance of $S_k = \Gamma_k(I - C_{2,k}G_{2,k-1}M_{2,k})\tilde{R}_{2,k}(I - C_{2,k}G_{2,k-1}M_{2,k})^\top\Gamma^\top_k$.*

Proof. Substituting (4) into (8), we have

$$\nu_k = \Gamma_k(C_{2,k}\tilde{x}^\star_{k|k} + v_{2,k}). \qquad (9)$$

Since $\mathbb{E}[\tilde{\hat{x}}^\star_{k|k}] = 0$ and $\mathbb{E}[v_{2,k}] = 0$ for all k as is proven [1, Lemma 13], it follows that the generalized innovation has zero mean, i.e., $\mathbb{E}[\nu_k] = 0$, with covariance

$$\mathbb{E}[\nu_k\nu_j^\top] = \mathbb{E}[\Gamma_k(C_{2,k}\tilde{\hat{x}}^\star_{k|k} + v_{2,k})(C_{2,j}\tilde{\hat{x}}^\star_{j|j} + v_{2,j})^\top\Gamma^\top_j].$$

We first show that the above covariance is zero when $k \neq j$. Without loss of generality, we assume that $k > j$. In this case, from the properties of the filter, we have $\mathbb{E}[v_{2,k}\tilde{x}^{\star\top}_{j|j}] = \mathbb{E}[v_{2,k}v^\top_{2,j}] = 0$, thus the covariance reduces to

$$\mathbb{E}[\nu_k\nu_j^\top] = \Gamma_kC_{2,k}(\mathbb{E}[\tilde{\hat{x}}^\star_{k|k}\tilde{\hat{x}}^{\star\top}_{j|j}]C^\top_{2,k} + \mathbb{E}[\tilde{\hat{x}}^\star_{k|k}v^\top_{2,j}])\Gamma^\top_j. \qquad (10)$$

Next, to evaluate $\mathbb{E}[\tilde{\hat{x}}^\star_{k|k}\tilde{\hat{x}}^{\star\top}_{j|j}]$ and $\mathbb{E}[\tilde{\hat{x}}^\star_{k|k}v^\top_{2,j}]$, we first evaluate the *a priori* estimation error:

$$\begin{aligned}&\tilde{\hat{x}}^\star_{k+1|k+1} = x_{k+1} - \hat{x}^\star_{k+1|k+1}\\ &= \overline{A}_k(I - \overline{L}_k\Gamma_kC_{2,k})\hat{x}^\star_{k|k} + (I - G_{2,k}M_{2,k+1}C_{2,k+1})w_k\\ &\quad - G_{2,k}M_{2,k+1}v_{2,k+1} + G_{2,k}M_{2,k+1}C_{2,k+1}G_{1,k}M_{1,k}v_{1,k}\\ &\quad - \overline{A}_k\overline{L}_k\Gamma_kv_{2,k} := \Phi_k\tilde{x}^\star_{k|k} + v'_k, \qquad (11)\end{aligned}$$

where Φ_k and v'_k are defined by the above while $\overline{A}_k := (I - G_{2,k}M_{2,k+1}C_{2,k+1})\hat{A}_k$ and $\hat{A}_k := A_k - G_{1,k}M_{1,k}C_{1,k}$. With the state transition matrix of the error system given by

$$\Phi_{k|j} = \begin{cases} \Phi_{k-1}\Phi_{k-1}\dots\Phi_j = \Phi_{k|j+1}\Phi_j, & k > j \\ I, & k = j, \end{cases}$$

the state estimate error is given by

$$\tilde{x}^\star_{k|k} = \Phi_{k|j}\tilde{x}^\star_{j|j} + \textstyle\sum_{\ell=j}^{k-1}\Phi_{k|\ell+1}v'_\ell. \qquad (12)$$

Thus, from (11), we obtain $\mathbb{E}[v'_\ell\tilde{x}^{\star\top}_{j|j}] = 0$ and $\mathbb{E}[v'_\ell v^\top_{2,j}] = 0$ when $\ell > j$ (i.e., future noise is uncorrelated with current estimate error and current noise) while when $\ell = j$, $\mathbb{E}[v'_j\tilde{x}^{\star\top}_{j|j}] = \overline{A}_j\overline{L}_j\Gamma_jR_{2,j}M^\top_{2,j}G^\top_{2,j-1}$ and $\mathbb{E}[v'_jv^\top_{2,j}] = \overline{A}_j\overline{L}_j\Gamma_jR_{2,j}$. With this and from (10), we can evaluate $\mathbb{E}[\tilde{x}^\star_{k|k}\tilde{x}^{\star\top}_{j|j}]$, $\mathbb{E}[\tilde{x}^\star_{k|k}v^\top_{2,j}]$ and $\mathbb{E}[\nu_k\nu_j^\top]$ as follows:

$$\begin{aligned}\mathbb{E}[\tilde{x}^\star_{k|k}\tilde{x}^{\star\top}_{j|j}] &= \Phi_{k|j+1}(\Phi_jP^{\star x}_{j|j} + \overline{A}_j\overline{L}_j\Gamma_jR_{2,j}M^\top_{2,j}G^\top_{2,j-1})\\ \mathbb{E}[\tilde{x}^\star_{k|k}v^\top_{2,j}] &= -\Phi_{k|j+1}(\Phi_jG_{2,j-1}M_{2,j}R_{2,j} + \overline{A}_j\overline{L}_j\Gamma_jR_{2,j})\\ \Rightarrow \mathbb{E}[\nu_k\nu_j^\top] &= \Gamma_kC_{2,k}\Phi_{k|j+1}(\overline{A}_j\overline{L}_j\Gamma_jR_{2,j}M^\top_{2,j}G^\top_{2,j-1}C^\top_{2,j}\\ &\quad + \Phi_jP^{\star x}_{j|j}C^\top_{2,j} - \Phi_jG_{2,j-1}M_{2,j}R_{2,j} - \overline{A}_j\overline{L}_j\Gamma_jR_{2,j})\Gamma^\top_j\\ &= \Gamma_kC_{2,k}\Phi_{k|j+1}\overline{A}_j(P^{\star x}_{j|j}C^\top_{2,j} - G_{2,j-1}M_{2,j}R_{2,j}\\ &\quad - \overline{L}_j\Gamma_j\tilde{R}^\star_{2,j})\Gamma^\top_j = 0, \qquad (13)\end{aligned}$$

where $\tilde{R}^\star_{2,j} = C_{2,j}P^{\star x}_{j|j}C^\top_{2,j} + R_{2,j} - R_{2,j}M^\top_{2,j}G^\top_{2,j-1}C^\top_{2,j} -$

$C_{2,j}G_{2,j-1}^\top M_{2,j}R_{2,j}$ and for the final equality, we substituted the filter gain from [1, Section 5.4]:

$$\overline{L}_j = (P_{j|j}^{\star x}C_{2,j}^\top - G_{2,j-1}M_{2,j}R_{2,j})\Gamma_j^\top(\Gamma_j\tilde{R}_{2,j}^\star\Gamma_j^\top)^{-1}.$$

Finally, for $j = k$, we can find $S_k := \mathbb{E}[\nu_k\nu_k^\top]$ as

$$\begin{aligned} S_k &= \Gamma_k(C_{2,k}P_{k|k}^{\star x}C_{2,k}^\top - C_{2,k}G_{2,k-1}M_{2,k}R_{2,k} \\ &\quad -R_{2,k}M_{2,k}^\top G_{2,k-1}^\top C_{2,k}^\top + R_{2,k})\Gamma_k^\top \\ &= \Gamma_k(I - C_{2,k}G_{2,k-1}M_{2,k})\tilde{R}_{2,k}(I - C_{2,k}G_{2,k-1}M_{2,k})^\top\Gamma_k^\top \\ &= \Gamma_k\tilde{R}_{2,k}^\star\Gamma_k^\top. \end{aligned}$$

Furthermore, from (12) and (9), since we assumed that w_k and v_k for all k and x_0 are Gaussian, the generalized innovation ν_k is a linear combination of Gaussian random variables and is thus itself Gaussian. Therefore, we have shown that ν_k is a Gaussian white noise with zero mean and covariance S_k. Moreover, S_k is positive definite since Γ_k is chosen such that S_k is invertible [1, Section 5.4]. ■

Remark 1. *The whiteness property of the generalized innovation provides an alternative approach to derive the filter gain $\overline{L}_k$ in [1] as can be seen by setting* (13) *to zero.*

C. Likelihood Function

Since the generalized innovation ν_k is a Gaussian white noise with zero mean and covariance S_k (Theorem 1), the conditional probability density function of ν_k given all measurements prior to time k, Z^{k-1}, is given by

$$\begin{aligned} f_{\nu_k|Z^{k-1}}(\nu_k|Z^{k-1}) &= \frac{\exp(-\nu_k^\top S_k^{-1}\nu_k/2)}{(2\pi)^{p_{\tilde{R}}/2}|S_k|^{1/2}} \\ &= \frac{\exp(-\overline{\nu}_k^\top\Gamma_k^\top(\Gamma_k\tilde{R}_{2,k}^\star\Gamma_k^\top)^{-1}\Gamma_k\overline{\nu}_k/2)}{(2\pi)^{p_{\tilde{R}}/2}|S_k|^{1/2}} \end{aligned} \tag{14}$$

where $p_{\tilde{R}} := \text{rank}(\tilde{R}_{2,k}^\star)$, $\tilde{\nu}_k := z_{2,k} - C_{2,k}\hat{x}_{k|k-1} - D_{2,k}u_k$, $\overline{\nu} = (I - C_{2,k}G_{2,k-1}M_{2,k})\tilde{\nu}_k$, $\tilde{\Gamma}_k := \Gamma_k(I - C_{2,k}G_{2,k-1}M_{2,k})$ and we have applied the identity $\tilde{R}_{2,k}^\star := \mathbb{E}[\overline{\nu}_k\overline{\nu}_k^\top] = (I - C_{2,k}G_{2,k-1}M_{2,k})\tilde{R}_{2,k}^\star(I - C_{2,k}G_{2,k-1}M_{2,k})^\top$ which results from the idempotence of $(I - C_{2,k}G_{2,k-1}M_{2,k})$ [1, cf. Lemma 17]. Furthermore, the idempotence of $(I - C_{2,k}G_{2,k-1}M_{2,k})$ implies that, similar to [1, Lemma 17], we can apply [14, Fact 3.12.9 and Proposition 2.6.3] to obtain $p_{\tilde{R}} := \text{rank}(\tilde{R}_{2,k}^\star) = l - p \leq l - p_{H_k}$ with equality only when $p = p_{H_k}$, i.e., H_k has full rank. Next, we note that if Γ_k is chosen as a matrix with orthonormal rows, $\Gamma_k^\top(\Gamma_k\tilde{R}_{2,k}^\star\Gamma_k^\top)^{-1}\Gamma_k$ is the generalized inverse and $|S_k|$ the pseudo-determinant of $\tilde{R}_{2,k}^\star$ [15, pp. 527-528]. From the above reference, for the case $p_{\tilde{R}} = l - p < l - p_{H_k}$, we also see that (14) represents the Gaussian distribution of $\overline{\nu}_k \in \mathbb{R}^{l-p_{H_k}}$ whose base measure is restricted to the $p_{\tilde{R}}$-dimensional affine subspace where the Gaussian distribution is supported. When H_k has full rank, the Gaussian distribution is fully supported in $\mathbb{R}^{l-p}$ and no restriction is necessary.

Therefore, with the above base measure, we obtain the conditional probability density function of $z_{2,k}$ conditioned on the system mode, q_k, and all prior measurements, Z^{k-1}, which we define as the *likelihood function* at time k:

$$\begin{aligned} \mathcal{L}(q_k|z_{2,k}) &:= f_{z_{2,k}|Z^{k-1},q_k}(z_{2,k}|Z^{k-1},q_k) \\ &= f_{\nu_k|Z^{k-1},q_k}(\nu_k|Z^{k-1},q_k) = \mathcal{N}(\nu_k^{q_k};0,S_k^{q_k}). \end{aligned} \tag{15}$$

As proposed in [1, Section 5.3], we can choose $\Gamma_k = \begin{bmatrix}0 & I_{p_{\tilde{R}}}\end{bmatrix}\tilde{U}_k^\top\hat{R}_k^{-\frac{1}{2}}$ where $\tilde{U}_k$ matrix is obtained from the singular value decomposition of $\hat{R}_k^{-\frac{1}{2}}C_kG_{2,k-1}$. A second choice bypasses the explicit expression of Γ_k with the use of Moore-Penrose pseudoinverse (†) and pseudodeterminant ($|\cdot|_+$), i.e., $\Gamma_k^\top(\Gamma_k\tilde{R}_{2,k}^\star\Gamma_k^\top)^{-1}\Gamma_k = \tilde{R}_{2,k}^{\star\dagger}$ and $|S_k| = |\tilde{R}_{2,k}^\star|_+$.

V. Multiple Model Approach

From the perspective of a hybrid system, the multiple model (MM) approach implements a *bank of filters* in parallel, with each corresponding to a system mode. The objective is then to decide which model/mode is the best representation of the current system mode as well as to estimate the state of the system based on this decision. In a nutshell, a Bayesian framework is used to compute the mode probabilities $\mu_k := P(q_k|Z^k)$ given all measurements up to time k. Using Bayes' rule, we can recursively find the mode probability $\mu_k^j := P(q_k = j|Z^k)$ at step k for each mode j, given $Z^k = \{z_{1,i}, z_{2,i}\}_{i=0}^k$ as

$$\begin{aligned} \mu_k^j &= P(q_k = j|z_{1,k}, z_{2,k}, Z^{k-1}) = P(q_k = j|z_{2,k}, Z^{k-1}) \\ &= \frac{f_{z_{2,k}|q_k,Z^{k-1}}(z_{2,k}|q_k{=}j,Z^{k-1})P(q_k{=}j|Z^{k-1})}{\sum_{\ell=1}^{\mathfrak{N}} f_{z_{2,k}|q_k,Z^{k-1}}(z_{2,k}|q_k{=}\ell,Z^{k-1})P(q_k{=}\ell|Z^{k-1})} \\ &= \frac{\mathcal{N}(\nu_k^j;0,S_k^j)P(q_k{=}j|Z^{k-1})}{\sum_{\ell=1}^{\mathfrak{N}}\mathcal{N}(\nu_k^\ell;0,S_k^\ell)P(q_k{=}\ell|Z^{k-1})}, \end{aligned} \tag{16}$$

where we have substituted (15) for the final equality and assumed that the probability of $q_k = j$ is independent of the measurement $z_{1,k}$. The rationale is that since we have no knowledge about $d_{1,k}$ and the $d_{1,k}$ signal can be of any type, the measurement $z_{1,k}$ provides no 'useful' information about the likelihood of the system mode (cf. (3)).

This section presents two types of multiple model estimators—static and dynamic. The static MM estimator assumes that the system mode remains constant/static, with a heuristic modification to keep the mode probabilities non-zero, while the dynamic MM estimator captures the mode-switching phenomenon by assuming that a Markov process can describe the mode jump process. In both cases, we assume that the prior mode probabilities are given

$$P(q_0 = j|Z^0) = \mu_0^j, \quad \forall\ 1, 2, \ldots, \mathfrak{N}, \tag{17}$$

where Z^0 is the prior information at time $k = 0$ and $\sum_{j=1}^{\mathfrak{N}}\mu_0^j = 1$, while $P(q_k = j|Z^{k-1})$ in (16) differs for both estimator variants. Both multiple model estimator variants have a fixed number of models. For better performance, modifications of the algorithms in this paper could be carried out to allow for a variable structure (cf. [16] for a discussion on model selection and implementation details).

A. MM Estimation: Static Variant

The static MM estimator (Algorithm 2) implements a bank of $\mathfrak{N}$ mode-conditioned simultaneous input and state filters (described in Section IV-A) in parallel with the assumption that the true system mode is fixed. As such, the state estimates of other potentially mismatched models will not be beneficial and thus, the bank of filters is run independently from each other. However, in order to apply the static MM estimator to the switched linear systems, some heuristic

Algorithm 2 Static-MM-Estimator ( )

Initialize for all $j \in \{1, 2, \dots, \mathfrak{N}\}$: $\hat{x}^j_{0|0}$; μ^j_0;
$\hat{d}^j_{1,0} = \Sigma^{j\,-1}_0 (z^j_{1,0} - C^j_{1,0} \hat{x}^j_{0|0} - D^j_{1,0} u_0)$;
$P^{d,j}_{1,0} = \Sigma^{j\,-1}_0 (C^j_{1,0} P^{x,j}_{0|0} C^{j\top}_{1,0} + R^j_{1,0}) \Sigma^{j\,-1}_0$;
for $k = 1$ to K **do**
  for $j = 1$ to $\mathfrak{N}$ **do**
    ▷ Mode-Matched Filtering
    Run Opt-Filter(j,$\hat{x}^j_{k-1|k-1}$, $\hat{d}^j_{1,k-1}$, $P^{x,j}_{k-1|k-1}$, $P^{d,j}_{1,k-1}$; f)
    $\overline{\nu}^j_k := z^j_{2,k} - C^j_{2,k} \hat{x}^{j\star}_{k|k} - D^j_{2,k} u_k$;
    $\mathcal{L}(j|z^j_{2,k}) = \frac{1}{(2\pi)^{p^j_{\tilde{R}}/2} |\tilde{R}^{j,\star}_{2,k}|^{1/2}_+} \exp\left(-\frac{\overline{\nu}^{j\top}_k \tilde{R}^{j,\star\dagger}_{2,k} \overline{\nu}^j_k}{2}\right)$;
  end for
  for $j = 1$ to $\mathfrak{N}$ **do**
    ▷ Mode Probability Update (small $\epsilon > 0$)
    $\overline{\mu}^j_k = \max\{\mathcal{L}(j|z^j_{2,k}) \mu^j_{k-1}, \epsilon\}$;
  end for
  for $j = 1$ to $\mathfrak{N}$ **do**
    ▷ Mode Probability Update (normalization)
    $\mu^j_k = \frac{\overline{\mu}^j_k}{\sum_{\ell=1}^{\mathfrak{N}} \overline{\mu}^\ell_k}$;
    ▷ Output
    Compute (18);
  end for
end for

modifications of the static MM estimator are necessary. Firstly, to keep all modes 'alive' such that they can be activated when appropriate, an artificial lower bound needs to be imposed on the mode probabilities. Moreover, to deal with unacceptable growth of estimate errors of mismatched filters, reinitialization of the filters may be needed, oftentimes with estimates from the most probable mode.

For each filter matched to mode j, the posterior mode probabilities can be computed with (16) where $P(q_k = j|Z^{k-1}) = P(q_{k-1} = j|Z^{k-1}) = \mu^j_{k-1}$, since the model is fixed. Then, these posterior mode probabilities are used to determine the most probable mode at each time k and the associated state and input estimates and covariances as:

$$\begin{aligned} j^* &= \arg\max \mu^j_k \\ \hat{x}_{k|k} &= \hat{x}^{j^*}_{k|k}, \quad \hat{d}_k = \hat{d}^{j^*}_k, \ P^x_{k|k} = P^{x,j^*}_{k|k}, \ P^d_k = P^{d,j^*}_k. \end{aligned} \tag{18}$$

Remark 2. *Alternatively, the output of the filter can be computed via combination of the mode-conditioned estimates:*

$$\begin{aligned} \hat{x}_{k|k} &= \textstyle\sum_{j=1}^{\mathfrak{N}} \mu^j_k \hat{x}^j_{k|k}, \ \hat{d}_k = \sum_{j=1}^{\mathfrak{N}} \mu^j_k \hat{d}^j_k, \\ P^x_{k|k} &= \textstyle\sum_{j=1}^{\mathfrak{N}} \mu^j_k [(\hat{x}^j_{k|k} - \hat{x}_{k|k})(\hat{x}^j_{k|k} - \hat{x}_{k|k})^\top + P^{x,j}_{k|k}] \\ P^d_k &= \textstyle\sum_{j=1}^{\mathfrak{N}} \mu^j_k [(\hat{d}^j_k - \hat{d}_k)(\hat{d}^j_k - \hat{d}_k)^\top + P^{d,j}_k] \end{aligned}$$

However, in this case, the estimate $\hat{d}_k$ was observed (in simulation) to deviate significantly from the true value. This is consistent with the fact that with a mismatched model, the estimate of $\hat{d}^j_k$ can be arbitrarily large. Hence, this choice of filter output for $\hat{d}_k$ is not recommended.

B. MM Estimation: Dynamic Variant

In contrast with the static MM estimator, the dynamic variant (Algorithm 3) assumes that the true mode switches in a Markovian manner with known (and possibly state dependent) transition probabilities

$$P(q_k = j|q_{k-1} = i, x_{k-1}) = p_{ij}(x_{k-1}), \ \forall\, i, j \in 1, \dots, \mathfrak{N}.$$

In fact, the mode transition probabilities can serve as estimator design parameters (cf. [4]). Therefore, the dynamic MM estimator design is more flexible and have the ability to integrate prior information of the mode switching process into the estimator. For conciseness and without loss of generality, we shall assume that the state transition probabilities are state independent, i.e., $p_{ij}(x_{k-1}) = p_{ij}$. The incorporation of the state dependency for stochastic guard conditions is rather straightforward, albeit lengthy and interested readers are referred to [17] for details and examples.

With the Markovian setting, the mode can change at each time step. As a result, the number of hypotheses (mode history) grows exponential with time. Therefore, an optimal multiple model filter is computationally intractable. We thus resort to suboptimal filters that manage the hypotheses in an efficient way. The simplest technique is *hypothesis pruning* in which a finite number of most likely hypotheses are kept, whereas the *hypothesis merging* approach keeps only last few of the mode histories, and combine hypotheses that differ in earlier steps (cf. [4] for approaches designed for switched linear systems without unknown inputs). In the following, we propose a recursive algorithm similar to the interactive multiple model (IMM) algorithm [4] which maintains $\mathfrak{N}$ number of estimates and $\mathfrak{N}$ number of filters at each time k.

Each iteration/cycle of the algorithm consists of three major components—initial condition mixing, mode-matched filtering and mode probability update. In the *initial condition mixing* step, for each input and state filter matched to mode j at time k, we compute the probability that the system was in mode i at time $k-1$ conditioned on Z^{k-1}:

$$\begin{aligned} \mu^{i|j}_k &:= P(q_{k-1} = i|q_k = j, Z^{k-1}) \\ &= \frac{P(q_k=j|q_{k-1}=i, Z^{k-1}) P(q_{k-1}=i|Z^{k-1})}{\sum_{\ell=1}^{\mathfrak{N}} P(q_k=j|q_{k-1}=\ell, Z^{k-1}) P(q_{k-1}=\ell|Z^{k-1})} \\ &= \frac{p_{ij}\mu^i_{k-1}}{P(q_k=j|Z^{k-1})} = \frac{p_{ij}\mu^i_{k-1}}{\sum_{\ell=1}^{\mathfrak{N}} p_{\ell j}\mu^\ell_{k-1}}. \end{aligned} \tag{19}$$

Then, with this mixing probabilities $\mu^{i|j}_k$ for all $i = \{1, \dots, \mathfrak{N}\}$, we mix the initial conditions for the filter matched to $q_k = j$ for all $j = \{1, \dots, N\}$ according to

$$\hat{x}^{0,j}_{k-1|k-1} = \textstyle\sum_{i=1}^{\mathfrak{N}} \mu^{i|j}_k \hat{x}^i_{k-1|k-1} \tag{20}$$

$$\hat{d}^{0,j}_{1,k-1} = \textstyle\sum_{i=1}^{\mathfrak{N}} \mu^{i|j}_k \hat{d}^i_{1,k-1} \tag{21}$$

$$\begin{aligned} P^{x,0,j}_{k-1|k-1} &= \textstyle\sum_{i=1}^{\mathfrak{N}} \mu^{i|j}_k [(\hat{x}^i_{k-1|k-1} - \hat{x}^{0,j}_{k-1|k-1}) \\ &\quad (\hat{x}^i_{k-1|k-1} - \hat{x}^{0,j}_{k-1|k-1})^\top + P^{x,i}_{k-1|k-1}] \\ P^{d,0,j}_{1,k-1} &= \textstyle\sum_{i=1}^{\mathfrak{N}} \mu^{i|j}_k [(\hat{d}^i_{1,k-1} - \hat{d}^{0,j}_{1,k-1}) \\ &\quad (\hat{d}^i_{1,k-1} - \hat{d}^{0,j}_{1,k-1})^\top + P^{d,i}_{1,k-1}] \end{aligned} \tag{22}$$

Note that there is no mixing of the $\hat{d}_{2,k}$ and its corresponding covariances because they are computed for a previous step and are not initial conditions for the bank of filters.

Next, in the *mode-matched filtering* step, a bank of $\mathfrak{N}$ simultaneous input and state filter (described in Section IV-A) is run in parallel using the mixed initial conditions computed in (20), (21) and (22). In addition, the likelihood function $\mathcal{L}(q_k = j|z_{2,k})$ corresponding to each filter matched to mode

Algorithm 3 Dynamic-MM-Estimator ( )

Initialize for all $j \in \{1, 2, \ldots, \mathfrak{N}\}$: $\hat{x}^j_{0|0}$; μ^j_0;
$\hat{d}^j_{1,0} = \Sigma_0^{j\,-1}(z^j_{1,0} - C^j_{1,0}\hat{x}^j_{0|0} - D^j_{1,0}u_0)$;
$P^{d,j}_{1,0} = \Sigma_0^{j\,-1}(C^j_{1,0}P^{x,j}_{0|0}C^{j\top}_{1,0} + R^j_{1,0})\Sigma_0^{j\,-1}$;
for $k = 1$ to K **do**
  for $j = 1$ to $\mathfrak{N}$ **do**
    ▷ Initial Condition Mixing
    $p^j_k = \sum_{\ell=1}^{\mathfrak{N}} p_{\ell j}\mu^\ell_{k-1}$;
    for $i = 1$ to $\mathfrak{N}$ **do**
      $\mu^{i|j}_k = \frac{p_{ij}\mu^i_{k-1}}{p^j_k}$;
    end for
    Compute (20), (21) and (22);
    ▷ Mode-Matched Filtering
    Run Opt-Filter(j,$\hat{x}^{0,j}_{k-1|k-1}$, $\hat{d}^{0,j}_{1,k-1}$, $P^{x,0,j}_{k-1|k-1}$, $P^{d,0,j}_{1,k-1}$);
    $\overline{\nu}^j_k := z^j_{2,k} - C^j_{2,k}\hat{x}^{j\star}_{k|k} - D^j_{2,k}u_k$;
    $\mathcal{L}(j|z^j_{2,k}) = \frac{1}{(2\pi)^{p^j_{\tilde{R}}/2}|\tilde{R}^{j,\star}_{2,k}|_+^{1/2}} \exp\left(-\frac{\overline{\nu}^{j\top}_k \tilde{R}^{j,\star\dagger}_{2,k}\overline{\nu}^j_k}{2}\right)$;
  end for
  for $j = 1$ to $\mathfrak{N}$ **do**
    ▷ Mode Probability Update
    $\mu^j_k = \frac{\mathcal{L}(j|z^j_{2,k})p^j_k}{\sum_{\ell=1}^{\mathfrak{N}}\mathcal{L}(j|z^\ell_{2,k})p^\ell_k}$;
    ▷ Output
    Compute (18);
  end for
end for

j is obtained by (15). Finally, in the *mode probability update* step, the posterior probability of mode j given measurements up to time k, $\mu^j_k := P(q_k = j|Z^k)$, can be found by substituting $P(q_k = j|Z^{k-1}) = \sum_{i=1}^{\mathfrak{N}} p_{ij}\mu^i_{k-1}$ from (19) into (16). For output purposes only (not a major step in the algorithm), the combined estimates and covariances can be computed as given in (18) (cf. Remark 2).

VI. Simulation Example

We return to the motivating example in Section II of two vehicles crossing an intersection. Using the hidden mode system model with state $x = \begin{bmatrix} x_A, \dot{x}_A, x_B, \dot{x}_B \end{bmatrix}$, each intention corresponds to a mode $q \in$ {I , M, C} with the following set of parameters and inputs:

• Inattentive Driver ($q =$ I), with an unknown time-varying d_1 (uncorrelated with x_B and $\dot{x}_B$, otherwise unrestricted):

$$A^I_c = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -0.1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -0.1 \end{bmatrix}, B^I_c = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, G^I_c = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$C^I_c = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, D^I_c = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, H^I_c = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0.1 \\ 0 & 1 \end{bmatrix}.$$

• Malicious Driver ($q =$ M), i.e., with $d_1 = K_p(x_B - x_A) + K_d(\dot{x}_B - \dot{x}_A)$ where $K_p = 2$ and $K_d = 4$:

$$A^M_c = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -K_p & -0.1 - K_d & K_p & K_d \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -0.1 \end{bmatrix}, H^I_c = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -1 \end{bmatrix}$$
$$B^M_c = B^I_c, G^M_c = G^I_c, C^M_c = C^I_c, D^M_c = D^I_c.$$

• Cautious Driver ($q =$ C), i.e., with $d_1 = -K_p x_A - K_d \dot{x}_A$ where $K_p = 2$ and $K_d = 4$:

$$A^M_c = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -K_p & -0.1 - K_d & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -0.1 \end{bmatrix}, H^I_c = \begin{bmatrix} 0 & 0 \\ 0 & -1 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$
$$B^M_c = B^I_c, G^M_c = G^I_c, C^M_c = C^I_c, D^M_c = D^I_c.$$

Furthermore, the velocity measurement of the vehicle is corrupted by an unknown time-varying bias d_2. Thus, the switched linear system is described by

$$\dot{x} = A^q_c x + B^q_c u + G^q_c d + w^q, \; y = C^q_c x + D^q_c u + H^q_c d + v^q,$$

where $d = \begin{bmatrix} d_1 & d_2 \end{bmatrix}^\top$, the intensities of the zero mean, white Gaussian noises, $w = \begin{bmatrix} 0 & w_1 & 0 & w_2 \end{bmatrix}^\top$ and v, are:

$$Q_c = 10^{-4}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1.6 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.9 \end{bmatrix}; R_c = 10^{-4}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.16 & 0 & 0 \\ 0 & 0 & 0.9 & 0 \\ 0 & 0 & 0 & 2.5 \end{bmatrix}.$$

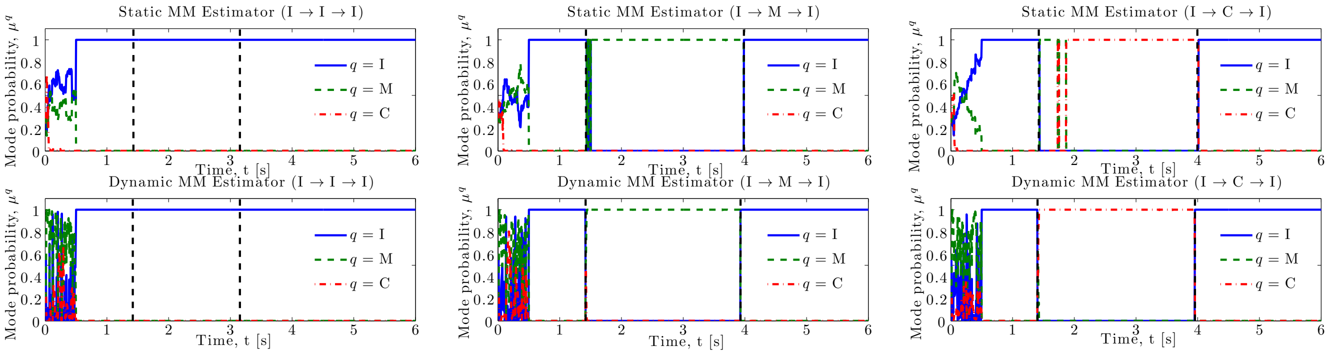
Since the proposed filter is for discrete-time systems, we employ a common conversion algorithm to convert the continuous dynamics to a discrete equivalent model with sample time $\triangle t = 0.01s$, assuming zero-order hold for the known and unknown inputs, u and d.

From Figure 1, we observe that both the static and dynamic MM estimators were successful at inferring the hidden modes of the system in the cases when the vehicle remains in the 'Inattentive' mode, or switches modes according to I→M→I or I→C→I. The performance of the static MM estimator is slightly worse than the dynamic variant, as can be seen in Figure 1(c). On the other hand, the changes in the mode probability estimate of the dynamic MM estimator are quicker which could be interpreted as having a higher 'sensitivity' to mode changes.

Taking a closer look at the 'I→M→I' scenario (the others are omitted due to space limitations) depicted in Figures 2a and 2b, we observe that both variants of the MM estimators performed satisfactorily in the estimation of states and unknown inputs. Similar to the observation of the mode probabilities, we note that the estimates of the static MM estimator (Figure 2a) are slightly inferior to that of the dynamic variant (Figure 2b). As aforementioned, this is because the dynamic MM estimator allows for mode transitions through a Markovian jump process where the transition matrix can be used as a design tool or incorporate prior knowledge about the mode switching process. In this example, the transition matrix is chosen as $P_T = \begin{bmatrix} 0.7 & 0.15 & 0.15 \\ 0.399 & 0.6 & 0.001 \\ 0.399 & 0.001 & 0.6 \end{bmatrix}$.

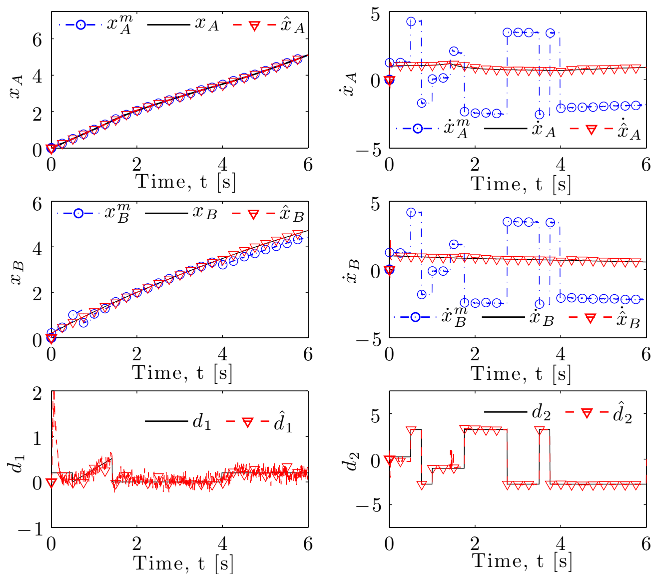
VII. Conclusion

This paper presented multiple model estimation algorithms for simultaneously estimating the mode, input and state of hidden mode switched linear stochastic systems with unknown inputs. We defined the notion of a generalized innovation sequence, which we then show to be a Gaussian white noise. Next, we exploited the whiteness property of the generalized innovations to form a likelihood function for determining mode probabilities. Simulation results for

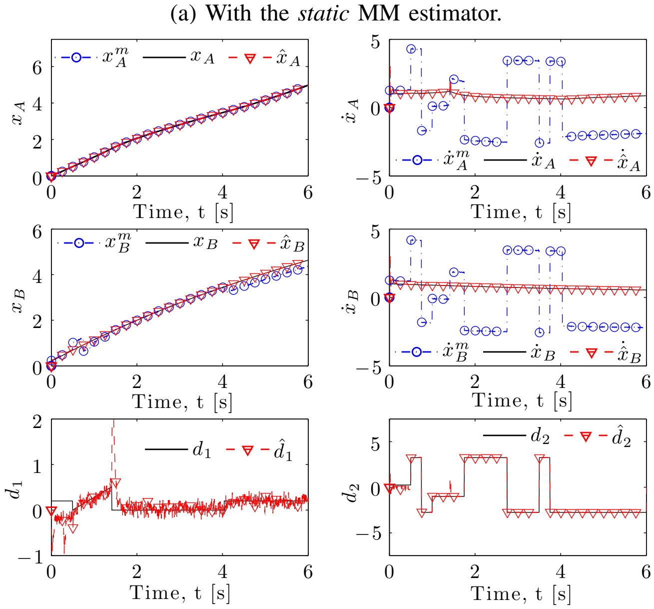


(a) Vehicle remain in 'I' mode. (b) Vehicle switches intentions 'I→M→I'. (c) Vehicle switches intentions 'I→C→I'.

Fig. 1: Mode probabilities for each mode with static (top) and dynamic (bottom) MM estimators.



(a) With the *static* MM estimator.



(b) With the *dynamic* MM estimator.

Fig. 2: Measured (superscript 'm', unfiltered), actual and estimated states and unknown inputs for the 'I→M→I' case.

vehicles at an intersection with switching driver intentions demonstrated the effectiveness of the proposed algorithms.

ACKNOWLEDGMENTS

The work presented in this paper was supported by the National Science Foundation, grant #1239182. M. Zhu is partially supported by ARO W911NF-13-1-0421 (MURI).

REFERENCES

[1] S.Z. Yong, M. Zhu, and E. Frazzoli. A unified filter for simultaneous input and state estimation of linear discrete-time stochastic systems. *Automatica*, 2014. Provisionally accepted. Available from: `http://arxiv.org/abs/1309.6627`.

[2] R. Verma and D. Del Vecchio. Safety control of hidden mode hybrid systems. *IEEE Transactions on Automatic Control*, 57(1):62–77, 2012.

[3] S.Z. Yong and E. Frazzoli. Hidden mode tracking control for a class of hybrid systems. In *Proceedings of the American Control Conference*, pages 5735–5741, June 2013.

[4] Y. Bar-Shalom, T. Kirubarajan, and X.R. Li. *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., New York, NY, USA, 2002.

[5] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan. Interacting multiple model methods in target tracking: a survey. *IEEE Transactions on Aerospace and Electronic Systems*, 34(1):103–123, Jan 1998.

[6] R.E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

[7] T. Kailath. An innovations approach to least-squares estimation–part i: Linear filtering in additive white noise. *Automatic Control, IEEE Transactions on*, 13(6):646–655, Dec 1968.

[8] P.D. Hanlon and P.S. Maybeck. Multiple-model adaptive estimation using a residual correlation kalman filter bank. *Aerospace and Electronic Systems, IEEE Transactions on*, 36(2):393–406, Apr 2000.

[9] H.A. Blom and Y. Bar-Shalom. The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Transactions on Automatic Control*, 33(8):780–783, August 1988.

[10] S. Gillijns and B. De Moor. Unbiased minimum-variance input and state estimation for linear discrete-time systems. *Automatica*, 43(1):111–116, January 2007.

[11] S. Gillijns and B. De Moor. Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough. *Automatica*, 43(5):934–937, 2007.

[12] S.Z. Yong, M. Zhu, and E. Frazzoli. Simultaneous input and state estimation for linear discrete-time stochastic systems with direct feedthrough. In *Conference on Decision and Control (CDC)*, pages 7034–7039, 2013.

[13] S.Z. Yong, M. Zhu, and E. Frazzoli. Resilient state estimation against switching attacks on stochastic cyber-physical systems. In *Proceedings of the American Control Conference*, 2015. Submitted.

[14] D.S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas (Second Edition)*. Princeton University Press, 2009.

[15] C.R. Rao. *Linear statistical inference and its applications*. Wiley, 1973.

[16] X.R. Li and Y. Bar-Shalom. Multiple-model estimation with variable structure. *IEEE Transactions on Automatic Control*, 41(4):478–493, 1996.

[17] C.E. Seah and I. Hwang. State estimation for stochastic linear hybrid systems with continuous-state-dependent transitions: an IMM approach. *IEEE Transactions on Aerospace and Electronic Systems*, 45(1):376–392, 2009.