Relative Entropy Rate Based Design for Linear Hybrid System Models

Onvaree Techakesari and Jason J. Ford

***Abstract*— Hybrid system representations have been applied to many challenging modeling situations. In these hybrid system representations, a mixture of continuous and discrete states is used to capture the dominating behavioural features of a nonlinear, possible uncertain, model under approximation. Unfortunately, the problem of how to best design a suitable hybrid system model has not yet been fully addressed. This paper proposes a new joint state-measurement relative entropy rate based approach for this design purpose. Design examples and simulation studies are presented which highlight the benefits of our proposed design approaches.**

I. INTRODUCTION

Hybrid system models involve a mixture of continuous (base) states and discrete (mode) states [1], [2]. Over the past few decades, hybrid system model filters have received a large amount of interest (see [2] for a review of hybrid system model filtering techniques). This interest has occurred because hybrid system model filter can be applied to a range of important problems, for example: road networks [3], target tracking and air traffic surveillance [4], [5], and fault detection and isolation [6].

The successful application of hybrid system models to nonlinear filtering problems involves both the selection of a suitable model and the application of a filter based on the selected model. Hence, there are two basic avenues to improve the performance of hybrid system model filtering approach: development of a better filtering algorithm, or selection of a better model representation [7]. In this paper, we consider the task of selecting a suitable hybrid system model representation.

In recent years, three model selection approaches have emerged [7]: the minimum-mismatch design method, the minimum-distance design method, and the equal-distance design method. In the minimum-mismatch method, a hybrid system model is designed in a manner to ensure that each mode is equally likely to be active [7], [8]. That is, the range of possible mode values is divided into equal probable regions of mode values, and a set of modes are selected such that each mode is the median of each region. Alternatively, the minimum-distance method selects a hybrid model in a manner that minimises the expected distance metric between the chosen set of mode values, and the underlying parameter values [7], [9]. That is, a set of modes is selected so that the likelihood of each mode within the chosen set matches the likelihood of the nominal parameter values the mode is representing. Finally, the equal-distance method proposes that modes are equally spaced over the range of parameter values [7], [10]. Hence, there is equal distance between adjacent modes. Unfortunately, none of these three design techniques is completely suitable in problems where the density of the model parameters is unknown.

In related work, relative entropy concepts (or the Kullback-Leibler distance concepts) have been shown to provide powerful tools to quantify the similarity of two model descriptions [11], [12]. Most importantly, this work showed that the relative entropy rate (RER) between the joint state and measurement processes of two hidden Markov models (HMMs) is related to the probabilistic distance between the HMMs (suggesting a connection between RER and the filter performance). In [12], this relationship between the RER concepts and the probabilistic distance is exploited in the design of HMM to approximate uncertain dynamic models [12]. Thus, these relative entropy concepts seem suitable for the above hybrid system model selection problem.

In this paper, we present a new joint state-measurement relative entropy rate based hybrid system model design approach which selects a model in a manner that minimises the RER between the nominal system and the hybrid system model under design. Importantly, the proposed design does not require *a priori* information about density of uncertain parameters. Several design examples and simulation studies are presented which illustrate and highlight the benefits of our presented design techniques.

This paper is organised as follows: Section II presents our nominal dynamics and our linear hybrid system model used for approximation. In Section III, relative entropy concepts are introduced and hybrid system model design approaches are proposed. Section IV presents some implementation issues and performance measures. Section V provides some design examples that illustrate the application of our design approaches. Finally, some concluding remarks are presented in Section VI.

II. PROBLEM FORMULATION

A. Nominal Dynamics

Consider an abstract complete probability space $(\Omega, \mathcal{F}, P)$. For $k > 0$, we will consider a nominal state process $x_k \in \mathbb{R}^n$ and a measurement process $y_k \in \mathbb{R}^m$ described by

$$\begin{aligned} x_k &= f(x_{k-1}, v_k, \theta_k) \\ y_k &= h(x_k, w_k) \end{aligned} \tag{1}$$

where x_k has an initial value x_0 and $\theta_k \in S_\theta \subset \mathbb{R}^d$ is an unknown (possibly time-varying) parameter process. Here, v_k and w_k are sequences of independent and identically distributed *i.i.d.* random variables with densities $\phi_v(\cdot)$ and $\phi_w(\cdot)$, respectively. We assume that w_k, v_k and x_0 are

O. Techakesari and J.J. Ford are with the School of Engineering Systems, Queensland University of Technology, Brisbane 4001, Australia o.techakesari@qut.edu.au, j2.ford@qut.edu.au

mutually independent for all k. We will now introduce some shorthand notation that will help with our presentation. We will use $x_{[a,b]}$ to denote the process x_k from $k = a$ to $k = b$, and we will define $y_{[a,b]}$ and $\theta_{[a,b]}$ similarly. We will let $S_{\theta|[1,k]}$ denote a set of possible parameter sequences. Throughout this paper, we will use the notation $\lambda^0(\theta_{[1,k]})$, or simply λ^0, to denote our model of the nominal state and measurement processes x_k and y_k.

In this paper, we are interested in the estimation problem of determining the conditional mean state estimates $E[x_k|y_{[1,k]}]$ in the situation where there is uncertainty about the system parameter $\theta_{[1,k]}$. We will consider a linear hybrid system based estimation approach for this problem, and we will focus our attention on how to best select an appropriate hybrid system model.

B. Linear Hybrid System Models

We will now introduce a linear hybrid system model that will serve as a possible approximation for the processes x_k and y_k. This hybrid representation is understood to involve a base state process in x_k, a mode process in $X_k \in S_X$, and a measurement process in y_k. The mode process describes the way the system switches between a finite set of N elementary dynamic behaviours that describe the evolution of the base state. We will let $e_i = [0, ..., 0, 1, 0, ...0]' \in \mathbb{R}^N$ denote an indicator vector with 1 in the ith position and zeros elsewhere, and let $S_X = \{e_1, ..., e_N\}$ denote the set of allowed indicator vectors. For $k > 0$, let us consider the linear hybrid system [13]:

$$\begin{aligned} x_k &= F_k(X_{k-1})\, x_{k-1} + G_k v_k^a \\ X_k &= A X_{k-1} + V_k^a \\ y_k &= H_k x_k + w_k^a \end{aligned} \tag{2}$$

where $F_k(X_k) \in \mathbb{R}^{n\times n}$, for all $X_k \in S_X$, is a mode-dependent base state transition matrix, $H_k \in M^{m\times n}$ is an output mapping matrix, and $G_k \in \mathbb{R}^{n\times v}$ is the state noise mapping matrix. We highlight that in more general formulations other model parameters (for example H_k and G_k) might also depend on the mode process X_k. However, for simplicity of presentation, we will consider the case where only $F_k(\cdot)$ is mode-dependent in this paper. Here, the state noise process $v_k^a \in \mathbb{R}^v$ and the measurement noise process $w_k^a \in \mathbb{R}^m$ are assumed to be zero-mean Gaussian noise processes with densities $\phi_v^a(x) = \mathcal{N}(x; 0, Q^a)$ and $\phi_w^a(x) = \mathcal{N}(x; 0, R^a)$, respectively, where $\mathcal{N}(x; \bar{x}, \Phi) \triangleq |2\pi\Phi|^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\bar{x})'\Phi^{-1}(x-\bar{x})}$ denotes the Gaussian probability density function (pdf) with the argument x, mean $\bar{x}$, and covariance Φ [1, p.51]. We highlight that the superscript a denotes that these processes are in the approximating model. Further, A is the time-invariant transition probability matrix for our mode process with ijth element $A^{ij} = P(X_{k+1} = e_i | X_k = e_j)$ and we understand V_k^a to be a martingale increment in the sense that $E[V_k^a | X_k] = 0$, see [14]. To complete the description, we assume the state process x_k has initial value x_0 and the Markov chain X_k has initial probabilities described by π_0 with ith element $\pi_0^i = P(X_0 = e_i)$.

Let us now introduce some model-related notation. We will use the notation $\lambda_E^i = (F_k(e_i), G_k, Q^a, x_0)$ to denote the ith elementary dynamic behaviour, and will assume that $\lambda_E^i \in S_E$, where S_E denotes all candidate elementary dynamics. We will also let $C_E = \{\lambda_E^1, \dots, \lambda_E^N\}$ denote a collection of dynamic behaviours present in a given hybrid model, and we will assume $C_E \in S_C$, where S_C will denote the set of all behaviour collections under consideration. Note that C_E does not include a characterisation of how mode transitions occur, and hence is only a partial description of the hybrid model. Finally, we will use the notation $\lambda = (F_k(\cdot), H_k, G_k, Q^a, R^a, A, \pi_0, x_0)$ to denote a hybrid system model, and will use S_λ to denote the set of all candidate hybrid system models under consideration.

The basic idea is that a suitable hybrid system model might lead to a more suitable filtering solution than the complex (perhaps intractable) nonlinear filter developed from the Duncan-Mortensen-Zakai equation [14, p. 269] applied directly to the nonlinear dynamics described by (1) [2]. The basic premise of this approach is that a hybrid model can be found that provides a reasonable representation of behaviours described by the nonlinear model being approximated.

III. RELATIVE ENTROPY CONCEPTS AND HYBRID SYSTEM MODEL SELECTION

Let us consider the joint state and measurement probability densities $p^\lambda(x_{[0,k]}, y_{[1,k]})$ and $p^{\overline{\lambda}}(x_{[0,k]}, y_{[1,k]})$ corresponding to two different models λ and $\overline{\lambda}$, respectively. The relative entropy, $\mathcal{D}_k^{KL}(\cdot||\cdot)$, between these densities is defined as (3) [15], [16] where we will use the conventions that $0/0 = 1$, $0\log 0 = 0$, $0 \times \infty = 0$, and that the relative entropy $\mathcal{D}_k^{KL}\left(p^\lambda(x_{[0,k]}, y_{[1,k]}) || p^{\overline{\lambda}}(x_{[0,k]}, y_{[1,k]}) \right)$ will be defined to be infinite whenever the ratio $\frac{p^\lambda(x_{[0,k]}, y_{[1,k]})}{p^{\overline{\lambda}}(x_{[0,k]}, y_{[1,k]})}$ is infinite for any feasible $x_{[0,k]}$, $y_{[1,k]}$ sequences.

When interested in dynamic systems, it is often more useful to consider the relative entropy rate (RER), $\mathcal{R}\left(p^\lambda(x_{[0,\infty]}, y_{[1,\infty]}) || p^{\overline{\lambda}}(x_{[0,\infty]}, y_{[1,\infty]}) \right)$, which is defined as [15, p. 134]

$$\begin{aligned} &\mathcal{R}\left(p^\lambda(x_{[0,\infty]}, y_{[1,\infty]}) || p^{\overline{\lambda}}(x_{[0,\infty]}, y_{[1,\infty]}) \right) \\ &\triangleq \lim_{k\to\infty} \frac{1}{k} \mathcal{D}_k^{KL}\left(p^\lambda(x_{[0,k]}, y_{[1,k]}) || p^{\overline{\lambda}}(x_{[0,k]}, y_{[1,k]}) \right). \end{aligned}$$

We will use the shorthand $\mathcal{R}\left(\lambda || \overline{\lambda}\right)$ to denote the joint state-measurement RER between models λ and $\overline{\lambda}$ (induced from their densities).

We now propose our hybrid system model design approaches that are based on the joint state-measurement RER between the nominal model and a candidate hybrid system model under consideration.

A. Min-Max Joint RER Hybrid System Model Design

Let us consider the problem of selecting hybrid system model to represent the nominal process model λ^0 with unknown parameter process $\theta_{[1,\infty]}$. We will aim to select a hybrid model that produces reasonable filter estimates

$$\mathcal{D}_k^{KL}\left(p^{\lambda}\left(x_{[0,k]},y_{[1,k]}\right)||\,p^{\overline{\lambda}}\left(x_{[0,k]},y_{[1,k]}\right)\right)\triangleq\int\log\left(\frac{p^{\lambda}\left(x_{[0,k]},y_{[1,k]}\right)}{p^{\overline{\lambda}}\left(x_{[0,k]},y_{[1,k]}\right)}\right)p^{\lambda}\left(x_{[0,k]},y_{[0,k]}\right)dx_{[0,k]}dy_{[0,k]} \tag{3}$$

regardless of the true parameter process $\theta_{[1,\infty]}$. For this purpose, for a candidate hybrid system model $\lambda \in S_\lambda$, let us define the max joint RER $J^M(\lambda)$ (worst-case RER) as

$$J^M(\lambda) \triangleq \sup_{\theta_{[1,\infty]}\in S_{\theta|[1,\infty]}} \mathcal{R}\left(\lambda^0(\theta_{[1,\infty]})||\lambda\right). \tag{4}$$

The min-max joint RER design will be defined as the hybrid system model $\lambda^M \in S_\lambda$ that satisfies

$$J^M(\lambda^M) = \inf_{\lambda\in S_\lambda} J^M(\lambda). \tag{5}$$

B. Special Cases - Static Hybrid System Model

We now consider some special cases that arise when the mode process X_k is unknown but constant for all k corresponding to constant $\theta_k = \theta$, and we denote the corresponding model as $\lambda^0(\theta)$. Although this static model description can be solved using the general min-max joint RER approach described by (5), this special design task can also be approached as a simplified problem of selecting a collection of elementary dynamic behaviours $C_E = \{\lambda_E^1, \ldots, \lambda_E^N\}$ (rather than selection of a full hybrid system model).

We now consider two approaches for this special case: a worst-mode design approach and a conditional-mode design approach.

1) Worst-Mode Design: Consider design of a behaviour collection C_E for the nominal model $\lambda^0(\theta)$ where θ is constant but unknown. In this case, we propose to select C_E so that for every value of $\theta \in S_\theta$, the nominal model $\lambda^0(\theta)$ is close (in a joint RER sense) to at least one mode $\lambda_E^i \in C_E$. For this purpose, consider the worst-mode criteria,

$$J^W(C_E) \triangleq \sup_{\theta\in S_\theta}\left[\min_{i\in[1,N]} \mathcal{R}\left(\lambda^0(\theta)||\lambda_E^i\right)\right]. \tag{6}$$

Hence, the static hybrid system model selection problem can be posed as the problem of finding the worst-mode collection $C_E^W \in S_C$ that satisfies

$$J^W(C_E^W) = \inf_{C_E\in S_C} J^W(C_E). \tag{7}$$

2) Conditional-Mode Design: Alternatively, when considering design of a behaviour collection C_E for a nominal model $\lambda^0(\theta)$ where θ is constant and has a known *a priori* density $p(\theta)$, then we can consider the following conditional RER concept. Let us introduce the conditional joint state-measurement RER, $\mathcal{R}^C(\lambda^0(\theta)||\lambda_E^i)$, which is defined as

$$\mathcal{R}^C\left(\lambda^0(\theta)||\lambda_E^i\right) \triangleq \int_{\theta\in S_\theta} p(\theta)\mathcal{R}\left(\lambda^0(\theta)||\lambda_E^i\right)d\theta \tag{8}$$

Related to this conditional joint state-measurement RER, we introduce the following conditional–mode criteria:

$$\begin{aligned} J^C(C_E) &\triangleq \min_{i\in[1,N]} \mathcal{R}^C\left(\lambda^0(\theta)||\lambda_E^i\right) \\ &= \int_{\theta\in S_\theta} p(\theta)\min_{i\in[1,N]} \mathcal{R}\left(\lambda^0(\theta)||\lambda_E^i\right)d\theta. \end{aligned} \tag{9}$$

We can then propose the conditional–mode model collection $C_E^C \in S_C$ as satisfying

$$J^C(C_E^C) = \inf_{C_E\in S_C} J^C(C_E). \tag{10}$$

Remark 1: Our presented conditional-mode criteria $J^C(C_E)$ is related to the minimum-distance criteria [7].

Remark 2: We highlight that the joint state-measurement RER is non-symmetric and that the order of RER arguments in (4), (6), and (9) is important.

IV. SIMULATION STUDIES

In this section, we will discuss some implementation issues and filter performance measures used in our simulation studies.

A. Calculation of Relative Entropy Rate

We remind that our joint state-measurement RER design approaches require the calculation of the joint state-measurement density $p^\lambda\left(x_{[0,k]}, y_{[1,k]}\right)$. For this purpose, note that we can establish efficient recursions for calculating $p^\lambda\left(x_{[0,k]}, y_{[1,k]}\right)$ by considering a fictitious HMM in which X_k is the state process and x_k is the observation process. Let $\alpha_k(j) \triangleq p^\lambda\left(X_k = e_j, x_{[0,k]}\right)$ be the probability of the observation sequence until time k, $x_{[0,k]}$, and the current state value X_k, then $\alpha_k(j)$ can be calculated via the usual forward recursion (or the unnormalised HMM filter), for $k > 0$,

$$\alpha_{k+1}(j) = \left[\sum_{i=1}^N \alpha_k(i)A^{ji}\right] b_j(x_{k+1}|x_k), \text{ for } j = 1, ..., N \tag{11}$$

with the initial value $\alpha_1(i) = \pi_0^i b_i(x_1|x_0)$ and the observation probability densities $b_i(x_k|x_{k-1}) \triangleq p^\lambda(x_k|x_{k-1}, X_k = e_i) = \mathcal{N}\left((x_k - F(e_i)x_{k-1}); 0, G_k Q^a G_k'\right)$, see [18]. We highlight that even though the interpretation is unusual, all these quantities are defined, and this filter exists. Using this fictitious HMM, the joint probability law of the state and measurement processes $p^\lambda(x_{[0,k]}, y_{[1,k]})$ can be calculated from α_k using Bayes' rule as follows:

$$p^\lambda(x_{[1,k]}, y_{[1,k]}) = p^\lambda(y_{[1,k]}|x_{[0,k]})\sum_{i=1}^N \alpha_k(i) \tag{12}$$

where $p^\lambda(y_{[1,k]}|x_{[0,k]}) = \prod_{i=1}^k \mathcal{N}\left((y_i - H_i x_i); 0, R^a\right)$ and we note that $\sum_{i=1}^N \alpha_k(i) = p^\lambda(x_{[0,k]})$. This fictitious HMM allows us to marginalise out the dynamics due to the mode process X_k and hence, provides an efficient finite dimensional algorithm to calculate the joint state-measurement probability densities $p^\lambda\left(x_{[0,k]}, y_{[1,k]}\right)$. The joint state-measurement RER can then be calculated by exploiting Monte Carlo technique [19].

B. Filter Performance Measure

In our simulation studies, we will use two root-mean-square error (RMSE) concepts to quantitative performance: average RMSE (ARMSE) and maximum RMSE (MRMSE). For $i = 1, \dots, L$, consider a set of state processes $x_k^{(i)}$ which are sampled to match the statistics provided for the nominal system, and corresponding state estimate $\hat{x}_k^{(i)}$, then average RMSE is defined as

$$\text{ARMSE} = \sqrt{\frac{1}{LN}\sum_{i=1}^{L}\sum_{k=1}^{N}\left(x_k^{(i)} - \hat{x}_k^{(i)}\right)'\left(x_k^{(i)} - \hat{x}_k^{(i)}\right)}.$$

Now let us consider a new set of state processes which are generated according to a specific θ value. Let us denote these processes as $x_k^{\theta|i}$. The maximum RMSE can be used to quantify filter performance for the worst parametric value, and is defined as

$$\text{MRMSE} = \max_{\theta \in S_\theta} \sqrt{\frac{1}{LN}\sum_{i=1}^{L}\sum_{k=1}^{N}\left(x_k^{(\theta|i)} - \hat{x}_k^{(i)}\right)'\left(x_k^{(\theta|i)} - \hat{x}_k^{(i)}\right)}.$$

C. The Static Multiple Model Filter

In our following design examples, the static multiple model (SMM) filter is the appropriate state estimator [2]. The SMM filter is a bank of N parallel Kalman filters where each Kalman filters is designed on one of the possible values of $F_k(X_k)$ (see [1], [2] for detailed implementation and structure of SMM filter and Kalman filter equations).

V. DESIGN EXAMPLES

We will now illustrate our hybrid system model design techniques in aircraft target tracking problems. For all design examples in this section, we will consider the following coordinated turn dynamics model $\bar{F}(\cdot)$ and position-only measurement $\bar{H}$ [1], [7]:

$$\bar{F}(\omega) = \begin{bmatrix} 1 & \frac{\sin(\omega T)}{\omega} & 0 & -\frac{1-\cos(\omega T)}{\omega} \\ 0 & \cos(\omega T) & 0 & -\sin(\omega T) \\ 0 & \frac{1-\cos(\omega T)}{\omega} & 1 & \frac{\sin(\omega T)}{\omega} \\ 0 & \sin(\omega T) & 0 & \cos(\omega T) \end{bmatrix} \text{ and}$$

$$\bar{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

where $T = 5s$ is the sampling period.

We will now present two hybrid system model design examples. In the first example, we consider the case where the uncertain parameter is constant but unknown, and there is no *a priori* information about the density of the uncertain parameter. This lack of *a priori* information means that the minimum-distance and minimum-mismatch approaches cannot be applied to these problems. Whilst the equal-distance design method can be applied, the method leads to rudimentary designs. Hence, this example will illustrate the benefits of our RER based design methods when there is a lack of *a priori* information about θ. Our second example involves a problem that can be approached using the existing techniques. That is, the density of the uncertain parameter is known *a priori* (this problem is also considered in [7]). This example provides a comparison between our proposed conditional-mode design method and the existing approaches. We will show that our method selects the same model as the minimum-distance method, and the selected model leads to a filter with reasonable performance.

Example 1: Parameterised Linear Dynamics (Density of Parameter Unknown), SMM Problem

For $k > 0$, consider a $2D$ maneuvering target tracking problem described by discrete-time Cartesian state process $x_k = [x_k^c, \dot{x}_k^c, y_k^c, \dot{y}_k^c]'$ with the nominal dynamics (see, [1, pp. 467-468]),

$$\begin{aligned} x_k &= \bar{F}(\omega^0)x_{k-1} + v_k \\ y_k &= \bar{H}x_k + w_k \end{aligned} \qquad (13)$$

where $\omega^0 \in \mathbb{R}^n$ is a constant nominal turn rate, $x_0 = [1000, 100, 200, 120]'$, and v_k and w_k are zero-mean Gaussian noise processes with covariances $Q = 0.01I_4$ and $R = 100I_2$, respectively. Let us assume that $|\omega^0| \leq 6^o/s$.

In this example, we are interested in estimating position and velocity of the target when lacking the knowledge of the density and the true value of ω^0. We construct a filter based on a three-element ($N = 3$) hybrid system model described by (2) with $F_k(X_k) = \bar{F}(V_x X_k)$, $G_k = I_4$, $H_k = \bar{H}$, and $A = I_4$, where $V_x = [\omega^m, 0, -\omega^m]$ is the mode mapping vector based on turn rate ω^m. Here, the noise processes v_k^a and w_k^a have covariances $Q^a = I_4$ and $R^a = R$, respectively. We also assume that the models have uniform initial probabilities $\pi_0 = [1/3, 1/3, 1/3]$. Note that in other application V_x could have independent elements, but this symmetric structure makes sense in this example.

In our design, we consider a selection of behaviour collection C_E with different V_x. We will use $C_E^i = [\lambda_E^{1|(i)}, \lambda_E^{2|(i)}, \lambda_E^{3|(i)}]$ to denote the ith candidate collection where $\lambda_E^{j|(i)} = \{\bar{F}(V_x e_j), G_k, Q^a, x_0\}$ for $j = 1, 2, 3$. Here, we apply our worst-mode design to a set of five candidate collections $S_C = \{C_E^1, C_E^2, C_E^3, C_E^4, C_E^5\}$ with the candidate turn rates $V_x^1 = [-2, 0, 2]^o/s$, $V_x^2 = [-3, 0, 3]^o/s$, $V_x^3 = [-4, 0, 4]^o/s$, $V_x^4 = [-5, 0, 5]^o/s$, and $V_x^5 = [-6, 0, 6]^o/s$.

Simulation studies involving these candidate models and corresponding filters were conducted by generating 241 samples of ω^0 equally spaced on the set $S_\theta = [-6, -5.5, -5, \dots, 6]^o/s$. For each ω^0, 1000 sets of $x_{[0,1000]}$ and $y_{[1,1000]}$ were generated according to (13). The initial state estimate $\hat{x}_0 = x_0$ and the initial filter covariance $P_0 = 100I_4$ are assumed in all these candidate filters.

Table I illustrates our worst-mode criteria, $J^W(\lambda)$, of different candidate hybrid system models. As shown by Table I, the model with behaviour collection C_E^3 (with the turn rate $\omega^m = 4^o/s$) solves our worst-mode design problem (7) with respect to the candidate set S_C. To experimentally validate this conclusion, we implemented SMM filters to compare the performance of the filters corresponding to each of the candidate hybrid system models. The simulated SMM

TABLE I
WORST-MODE CRITERIA AND SMM FILTER PERFORMANCE

Candidate Turn Rate $\omega^m (^o/s)$	Worst-Mode Criteria, $J^W(\lambda)$	Simulated MRMSE	Simulated ARMSE
2	10714	318.35	158.70
3	6036	246.94	116.16
4	2692	169.55	99.19
5	4202	211.52	114.55
6	6010	255.18	148.02

TABLE II
CONDITIONAL-MODE CRITERIA AND SMM FILTER PERFORMANCE

Candidate Turn Rate $\omega^m (^o/s)$	Conditional-Mode Criteria, $J^C(\lambda)$	Simulated ARMSE
2	592.43	287.01
3	480.48	271.50
4	692.80	328.87
5	1113.70	414.35
2.7	472.93	267.01

performance is also shown in Table I and the result suggests that the worst-mode design C_E^3 outperforms the other candidate filters under consideration. Importantly, the result suggests that the filter performance predicted by our worst-mode design approach corresponds well with the simulated filter performance.

Example 2: Parameterised Linear Dynamics (Density of Parameter Known), SMM Problem

We now consider another version of the previous example in which the density of ω^0 is known *a priori*. This example will allow us to compare our proposed design technique with the existing techniques. Consider the nominal dynamics described by (13) but now we will assume that the nominal turn rate ω^0 is distributed by Gaussian-mixture,

$$p(\omega^0) = c_0 \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(\omega^0)^2}{2\sigma_0^2}} + c\left[\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\omega^0+\omega_s)^2}{2\sigma^2}} + \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\omega^0-\omega_s)^2}{2\sigma^2}}\right] \tag{14}$$

where $c_0 = 2/3, c = 1/6, \sigma_0 = \sigma = 1$, and $\omega_s = 3^o/s$ (this problem is also considered in [7]).

The same type of hybrid system models as in the previous example were considered and our conditional-mode design was applied to a set of five candidate collections $S_C = \{C_E^1, C_E^2, C_E^3, C_E^4, C_E^5\}$ with the candidate turn rates $V_x^1 = [-2, 0, 2]^o/s$, $V_x^2 = [-3, 0, 3]^o/s$, $V_x^3 = [-4, 0, 4]^o/s$, $V_x^4 = [-5, 0, 5]^o/s$, and $V_x^5 = [-2.7, 0, 2.7]^o/s$. We note that the design parameter $2.7^o/s$ is included because this is the design identified in [7].

Similar to the previous example, 100,000 nominal state and measurement sequences, $x_{[0,1000]}$ and $y_{[1,1000]}$, were generated according to (13) where ω^0 is selected from the density (14). Again, we implemented SMM filters to compare the performance of the filters corresponding to each of the candidate hybrid system models. Here, the initial state estimate $\hat{x}_0 = x_0$ and the initial filter covariance $P_0 = 100 I_4$ are assumed in all these candidate filters.

Table II illustrates the conditional-mode criteria (9) between the nominal model and the candidate hybrid system model, and the simulated SMM performance. It is clear that the candidate collection C_E^5 (with the turn rate $2.7^o/s$) solves the conditional-model design problem. Further, the results suggest that the filter corresponding to the conditional-mode design collection C_E^5 outperforms the other candidate filters under consideration. We highlight that the model with C_E^5 is the same model selected by the existing minimum-distance design approach [7]. Hence, this design example illustrates that our design approach is consistent with the existing technique when the density of uncertain parameter is available, and the filter corresponding to the designed model outperforms other filters under consideration.

A. Summary of Simulation Studies

In these simulation studies, we illustrated the benefits of our proposed hybrid system model design methods. Our proposed conditional-mode design method was shown to lead to the same hybrid system model as the minimum-distance method when there is *a priori* information about the density of the unknown parameter. When the density of the uncertain parameter is not available, our studies suggested that the filter resulting from our designed hybrid system model outperformed other candidate filters. Importantly, the filter performance predicted by our RER designs corresponds well with the simulated filter performance. We highlight that we are currently investigating the use of these design tools in the design of interacting multiple model (IMM) filters (see [1], [2] for structure and implementation process of IMM algorithm).

VI. CONCLUSION

This paper presented a novel joint state-measurement relative entropy rate based approach for the design of hybrid system representations. These hybrid system descriptions are useful in developing filters for uncertain nonlinear dynamics. The proposed approach is superior to previous design techniques which require more information about the nominal system. A number of illustrative design examples were presented and the results were examined in simulation studies.

ACKNOWLEDGMENTS

This research was supported under Australian Research Council's Linkage Projects funding scheme (project number LP100100302) and the Smart Skies Project, which is funded, in part, by the Queensland State Government Smart State Funding Scheme. Computational resources and services used in this work were provided by the High Performance Computing and Research Support Group, Queensland University of Technology, Brisbane, Australia.

REFERENCES

[1] Y. Bar-Shalom, X. Rong Li, and K. Thiagalingan, *Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software*, 2nd ed., New York: Wiley, 2001.

[2] X. Rong Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part V: multiple-model methods," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 4, pp. 1255-1321, Oct. 2005.

[3] T. Kirubarajan, Y. Bar-Shalom, K. R. Pattipati, and I. Kadar, "Ground target tracking with variable structure IMM estimator," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 1, pp. 26-46, Jan. 2000.

[4] X. R. Li and Y. Bar-Shalom, "Design of an interacting multiple model algorithm for air traffic control tracking," *IEEE Trans. Contr. Syst. Technol.*, vol. 1, no. 3, pp. 186-194, Sep. 1993.

[5] C. E. Seah and I. Hwang, "Stochastic linear hybrid systems: modeling, estimation and application in air traffic control," *IEEE Trans. Contr. Syst. Technol.*, vol. 17, no. 3, pp. 563-575, May 2009.

[6] J. Ru and X. Rong Li, "Variable-structure multiple-model approach to fault detection, identification, and estimation," *IEEE Trans. Contr. Syst. Technol.*, vol. 16, no. 5, pp. 1029-1038, Sep. 2008.

[7] X. R. Li, Z. Zhao, and X.-B. Li, "General model-set design methods for multiple-model approach," *IEEE Trans. Autom. Control*, vol. 50, no. 9, pp. 1260-1276, Sep. 2005.

[8] X. R. Li, Z. L. Zhao, P. Zhang, and C. He, "Model-set design for multiple-model estimation. Part II: examples," in *Proc. 2002 Int. Conf. Information Fusion*, Annpolis, MD, Jul. 2002, pp.1347-1354.

[9] M. Yang, J. F. Ru, H. M. Chen, X. R. Li, and N. S.V. Rao, "Predicting internet end-to-end delay: a statistical case study," in *Annual Review of Communications*, vol.58, Chicago, IL, 2005.

[10] X. Rong Li, V. P. Jilkov, and J. Ru, "Multiple-model estimation with variable structure. Part VI: expected-mode augmentation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 3, pp. 853-867, Jul. 2005.

[11] L. Xie, V. A. Ugrinovskii, and I. R. Petersen, "Probabilistic distances between finite-state finite-alphabet hidden Markov models," *IEEE Trans. Autom. Control*, vol. 50(4), pp. 505-511, Apr. 2005.

[12] J. Lai and J. J. Ford, "Relative entropy rate based multiple hidden Markov model approximation," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 165-174, Jan. 2010.

[13] R. J. Elliott, F. Dufour, and D. D. Sworder, "Exact Hybrid Filters in Discrete Time", *IEEE Trans. Autom. Control*, vol. 41, no. 12, pp. 1807-1810, Dec. 1996.

[14] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models: Estimation and Control*, New York: Springer-Verlag, 1995.

[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed., New York: Wiley, 2005.

[16] R. M. Gray, *Entropy and Information Theory*, New York: Springer-Verlag, 1990.

[17] P. Billingsley, *Probability and Measure*, 3rd ed., New York: Wiley, 1995.

[18] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257-286, Feb. 1989.

[19] M. N. Do, "Fast approximation of Kullback-Leibler distance for dependence trees and hidden Markov models," *IEEE Signals Process. Lett.*, vol. 10, no. 4, pp. 115-118, Apr. 2003.