Computer Science and Artificial Intelligence Laboratory

Technical Report

MIT-CSAIL-TR-2006-013 February 28, 2006

Finite Horizon Control Design for Optimal Discrimination between Several Models

Lars Blackmore and Brian Williams

massachusetts institute of technology, cambridge, ma 02139 usa — www.csail.mit.edu

Finite Horizon Control Design for Optimal Discrimination between Several Models

Lars Blackmore and Brian Williams

Abstract— **Multiple-Model fault detection is a powerful method for detecting changes, such as faults, in dynamic systems. In many cases, the ability of such a detection scheme to distinguish between possible models for the system dynamics depends critically on the control inputs applied to the system. Prior work has therefore aimed to design control inputs in order to improve fault detection. We previously developed a new method that uses constrained finite horizon control design to create control inputs that minimize an upper bound on the probability of model selection error. This method is limited, however, to the problem of selection between two models. In this paper we describe a new method that extends this approach to handle an arbitrary number of models. By optimizing subject to hard constraints, the new method can ensure that a defined task is fulfilled, while optimally discriminating between models. This means that the discrimination power of the designed control input can be much greater than that created by other approaches, which typically design 'auxiliary' signals with limited power so that the effect on the system state is small. Furthermore, the optimization criterion, which is an upper bound on the probability of model selection error, has a more meaningful interpretation than alternative approaches that are based on information gain, for example.**

We demonstrate the method using an aircraft fault detection scenario and show that the new method significantly reduces the bound on the probability of error when compared to a manually generated identification sequence and a fuel-optimal sequence.

I. Introduction

In multiple-model (MM) fault detection it is necessary to select the most likely model from a finite set, given observations [1][2]. For example, fault detection can be posed as a problem of deciding whether the dynamic model describing the nominal behavior of a system or the dynamic model describing a failure is most likely[3]. Well-known methods exist for selecting a model given a set of observations[3][4]. For this paper we assume a Bayesian decision rule, which selects the most likely model given the observations and a prior distribution over the models. The ability of any detection method to discriminate between different competing dynamic system models is highly dependent on the control inputs applied to the system. By designing control inputs to discriminate between models, the performance of the detection system can be enhanced greatly. In this paper we present a new method for designing system inputs that discriminate between the possible models in an optimal sense.

The problem of control design for model discrimination has received a great deal of attention[5][6][7][8]. Typical approaches design *auxiliary* signals that are added to the nominal control signal for the purpose of model discrimination[7][8]. The auxiliary signal has limited power so that its effect on the system state is limited. This, however, also restricts the discrimination power of the signal. In our previous work, we developed a new method that used constrained finite horizon control design to create discrimination signals with respect to hard constraints[9]. These constraints can be used to ensure that a certain task, defined in terms of the system state, is fulfilled, or that hard constraints such as actuator saturation are not violated. By optimizing subject to hard constraints, the method can generate signals that are far more effective in discrimination than a limited power auxiliary signal.

This method was, however, limited to the problem of discrimination between two models. In this paper, we extend the approach described in [9] to handle an arbitrary number of different models.

Previous approaches to the problem of control design for model discrimination have suggested a number of different criteria for optimization, such as information gain or the distance between the distributions of predicted observations[7]. These criteria typically do not have a meaningful interpretation in the context of the model selection problem. By contrast, consistent with a Bayesian approach, we use an upper bound on the probability of model selection error as the optimization criterion.

The key insight behind the new method is to create a tractable bound that applies to multiple models. The Battacharrya bound[10] used in [9] applies only to the case of selection between two models. We derive a new upper bound that applies to an arbitrary number of models; this bound is expressed in terms of the Battacharrya bound between pairs of models, leading to a sum-of-Gaussians form. Then we pose the problem of designing a finite sequence of control inputs to minimize this bound, subject to constraints, as a finite horizon trajectory design problem. Lastly, we show that in the case of linear constraints this problem can be solved using Sequential Quadratic Programming[12].

The result of this work is a new algorithm that generates a finite sequence of control inputs that minimize an upper bound on the probability of model selection error. These sequences are designed subject to state and control constraints. The algorithm can be used to ensure that a given task, defined in terms of constraints on the expected state, is fulfilled while optimally detecting failures; the constraints are used for control, while the optimization is used for model discrimination.

We present results for the example of an aircraft that has a number of possible failure modes. Compared to a typical

sequence designed by a human, and a sequence optimized to minimize fuel consumption, the method dramatically reduces the upper bound on the probability of model selection error.

A. Aircraft Fault Detection Scenario

In this paper, the detection of failures on an aircraft is used as a motivating example. The discrete time approximation to the longitudinal dynamics of the aircraft, linearized about the trim state, is shown in Fig. 1.

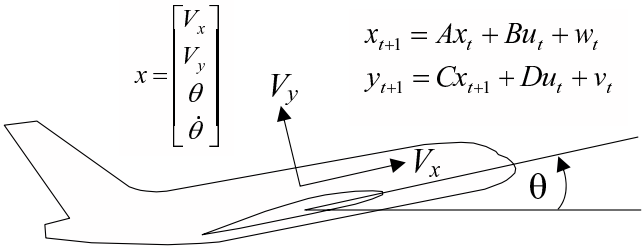

Fig. 1. Discrete-time aircraft model linearized about the trim state. Here x_t is the state of the system at time step t while y_t is the observed output of the system at time t, which is taken to be the pitch rate $\dot{\theta}$ and the vertical velocity V_y. The input is denoted u_t, and is taken to be the requested elevator angle. We assume that the elevator actuator saturates at $\pm 0.25 rad$. The terms w_t and v_t are the *process noise* and *observation noise* respectively. The noise at any time step is assumed to be independent of the noise at any other time step, and w_t and v_t are independent of each other with normal distributions $\mathcal{N}(0, Q)$ and $\mathcal{N}(0, R)$ respectively. The initial state of the system is also assumed to be normally distributed.

In the model selection task, we must determine which system described by the system matrices $\{A, B, C, D\}$ best models the data. Here the situation under consideration has an arbitrary, but finite, number of candidate models. Under Hypothesis H_i, the system is described by $\{A_i, B_i, C_i, D_i\}$. For the aircraft example, we consider three single-point failures; the pitch rate sensor may fail, the vertical velocity sensor may fail, or the elevator actuator may fail. This gives the following alternative models:

- H_0: Nominal (no faults)
- H_1: Faulty pitch rate sensor
- H_2: Faulty vertical velocity sensor
- H_3: Faulty elevator actuator

In the case of sensor faults, we assume that the sensor reading is zero mean white noise, while in the case of the actuator fault we assume that the elevator exerts no control effort. Consistent with a Multiple-Model fault detection framework, we assume that the system matrices $\{A, B, C, D\}$ are fully known for each of the possible faults.

II. Hypothesis Selection and Bayes Risk

Here we assume models are selected by Bayesian hypothesis selection. Bayesian hypothesis selection between an arbitrary number of models can be expressed as follows:

$$\text{Select } H_i \text{ where } i = \arg\max_j p(H_j|\mathbf{y}, \mathbf{u}).$$

Using Bayes' rule, this selection is given by:

$$\text{Select } H_i \text{ where } i = \arg\max_j p(\mathbf{y}|H_j, \mathbf{u})p(H_j).$$

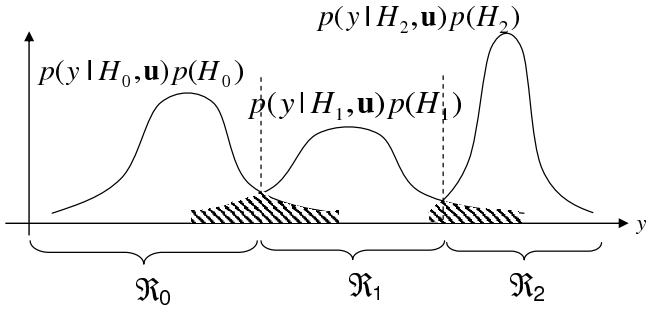

Fig. 2. Selection between multiple models given an observation y and a prior. In general Bayesian selection between multiple hypotheses yields a number of *decision regions* in the space of possible observations. If the observation y falls into set $\Re_i$ then the classifier selects H_i. Even with Bayes optimal selection there is a finite probability of error given by the *Bayes Risk*, denoted by the shaded region.

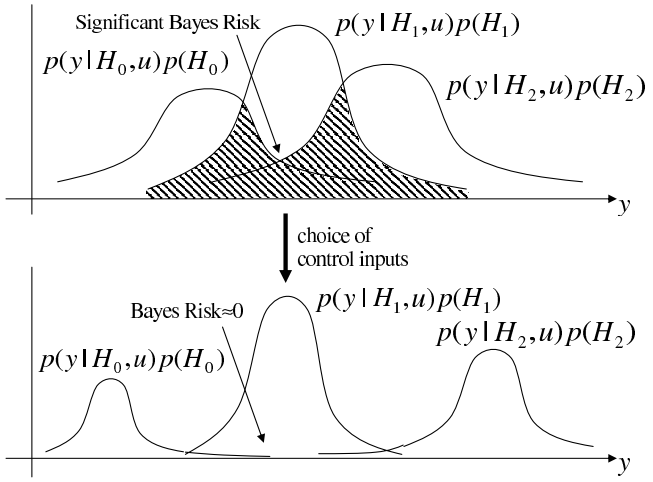

Fig. 3. Graph showing $p(y|H_0, u)$, $p(y|H_1, u)$ and $p(y|H_1, u)$ for two different choices of u. In the upper figure, the predicted distributions overlap significantly, leading to a large Bayes risk. In the lower figure, a different selection of u has separated the distributions, meaning that when the observation y is made, the correct hypothesis can be selected with high confidence. The Bayes risk is very low, meaning that the probability of error is very low.

The term $p(H_j)$ represents prior probabilities for hypothesis j. The priors for each hypothesis can be calculated in a number of different ways: there may be explicit knowledge about how *a priori* likely the different hypotheses are, or the prior can represent the belief state created by an estimator. As shown in Fig. 2, Bayesian selection yields a number of decision regions $\Re_i$ defined such that:

$$\Re_i = \{\mathbf{y}|p(H_i, \mathbf{y}|\mathbf{u}) > p(H_l, \mathbf{y}|\mathbf{u}) \quad \forall l \neq i\} \tag{1}$$

Hypothesis i is selected if the observation $\mathbf{y}$ falls in region $\Re_i$. This Bayesian selection rule minimizes the likelihood of selecting an incorrect hypothesis given the available information. As shown in Fig. 2, the Bayesian optimal classifier has a finite probability of selecting the incorrect hypothesis,

known as the *Bayes Risk*. The Bayes risk is given by:

$$\begin{aligned} P(error) &= \sum_i \sum_{j \neq i} P(\mathbf{y} \in \Re_j, H_i|\mathbf{u}) \\ &= \sum_i \sum_{j \neq i} P(\mathbf{y} \in \Re_j|H_i, \mathbf{u}) P(H_i) \\ &= \sum_i \sum_{j \neq i} \int_{\Re_j} p(\mathbf{y}|H_i, \mathbf{u}) P(H_i) d\mathbf{y} \end{aligned} \quad (2)$$

Since the Bayes Risk is the probability of error when using the optimal classifier, we would like to optimize our control inputs to the system to minimize this measure. The effect of input choice on the Bayes Risk is illustrated in Fig. 3.

III. Bounding the Bayes Risk for Multiple Hypotheses

The Bayes Risk is unsuitable as an optimization criterion, since the finite integral in (2) cannot, in general, be evaluated in closed form. In [9], we used instead a tractable upper bound on the Bayes Risk, known as the Bhattacharrya bound[10]. This bound does not, however, apply to more than two hypotheses. In this section we therefore introduce a new bound that applies to an arbitrary number of hypotheses.

Theorem 1. *For Gaussian observation distributions, such that $p(\mathbf{y}|H_i) = \mathcal{N}(\mu_i, \Sigma_i)$ and $p(\mathbf{y}|H_j) = \mathcal{N}(\mu_j, \Sigma_j)$, the Bayes Risk when performing hypothesis selection between an arbitrary number of hypotheses is upper bounded as follows:*

$$P(error) \leq \sum_i \sum_{j>i} P(H_i)^{\frac{1}{2}} P(H_j)^{\frac{1}{2}} e^{-k(i,j)} \quad (3)$$

where:

$$\begin{aligned} k(i,j) &= \frac{1}{4}(\mu_\mathbf{j} - \mu_\mathbf{i})^T \left[\Sigma_i + \Sigma_j\right]^{-1} (\mu_\mathbf{j} - \mu_\mathbf{i}) \\ &\quad + \frac{1}{2} ln \frac{\left|\frac{\Sigma_i + \Sigma_j}{2}\right|}{\sqrt{|\Sigma_i||\Sigma_j|}}. \end{aligned} \quad (4)$$

Proof: The key idea is to express the n-hypothesis Bayes Risk in terms of the Bayes Risk between pairs of hypotheses, and then to bound the term for each pair in a manner analogous to the Bhattacharrya bound.

The n-hypothesis Bayes Risk given by (2) can be expressed in terms of hypothesis pairs as follows:

$$\begin{aligned} &\textstyle\sum_i \sum_{j \neq i} P(\mathbf{y} \in \Re_j, H_i|\mathbf{u}) \\ &= \textstyle\sum_i \sum_{j>i} \Big(P(\mathbf{y} \in \Re_j, H_i|\mathbf{u}) + P(\mathbf{y} \in \Re_i, H_j|\mathbf{u}) \Big) \end{aligned} \quad (5)$$

Here, each term in the summation is the Bayes Risk between Hypothesis i and Hypothesis j. This can be written exactly as:

$$\begin{aligned} &\Big(P(\mathbf{y} \in \Re_j, H_i|\mathbf{u}) + P(\mathbf{y} \in \Re_i, H_j|\mathbf{u}) \Big) = \\ &\textstyle\int_{\Re_j} p(\mathbf{y}|H_i, \mathbf{u}) P(H_i) d\mathbf{y} + \int_{\Re_i} p(\mathbf{y}|H_j, \mathbf{u}) P(H_j) d\mathbf{y} \end{aligned} \quad (6)$$

We now define two regions $\Re_A$ and $\Re_B$ as follows:

$$\begin{aligned} \Re_A &= \{\mathbf{y}|p(H_i, \mathbf{y}|\mathbf{u}) > p(H_j, \mathbf{y}|\mathbf{u})\} \\ \Re_B &= \{\mathbf{y}|p(H_i, \mathbf{y}|\mathbf{u}) > p(H_j, \mathbf{y}|\mathbf{u})\} \end{aligned} \quad (7)$$

We can relate these regions to the decision regions $\Re_i$ and $\Re_j$. From the definition of hypothesis selection given in Section II, decision region $\Re_i$ is where the likelihood of Hypothesis H_i is greater than all other hypotheses:

$$\begin{aligned} \Re_i &= \{\mathbf{y}|p(H_i, \mathbf{y}|\mathbf{u}) > p(H_l, \mathbf{y}|\mathbf{u}) \quad \forall l \neq i\} \\ \Re_j &= \{\mathbf{y}|p(H_j, \mathbf{y}|\mathbf{u}) > p(H_l, \mathbf{y}|\mathbf{u}) \quad \forall l \neq j\}. \end{aligned} \quad (8)$$

It is clear from (7) and (8) that the decision regions are subsets of the regions $\Re_A$ and $\Re_B$, such that $\Re_i \subseteq \Re_A$ and $\Re_j \subseteq \Re_B$. We can therefore bound the 2-hypothesis Bayes Risk term in (6) as follows:

$$\begin{aligned} &\textstyle\int_{\Re_j} p(\mathbf{y}|H_i, \mathbf{u}) P(H_i) d\mathbf{y} + \int_{\Re_i} p(\mathbf{y}|H_j, \mathbf{u}) P(H_j) d\mathbf{y} \leq \\ &\textstyle\int_{\Re_B} p(\mathbf{y}|H_i, \mathbf{u}) P(H_i) d\mathbf{y} + \int_{\Re_A} p(\mathbf{y}|H_j, \mathbf{u}) P(H_j) d\mathbf{y} \end{aligned} \quad (9)$$

The key idea in the derivation of the Bhattacharyya bound is to express two integrals over different decision regions, as a single integral over the entire space of $\mathbf{y}$. The integral over decision regions is intractable, because it involves integration of Gaussian distributions with definite limits. The integral over the entire space, on the other hand, can be evaluated in closed form. In an analogous manner, we now note that the union of the regions $\Re_A$ and $\Re_B$ is the entire space of $\mathbf{y}$. Using the definitions in (7), we can therefore write the integrals in (9) as a single integral over the entire space of $\mathbf{y}$.

$$\begin{aligned} &\int_{\Re_B} p(\mathbf{y}|H_i, \mathbf{u}) P(H_i) d\mathbf{y} + \int_{\Re_A} p(\mathbf{y}|H_j, \mathbf{u}) P(H_j) d\mathbf{y} \\ &= \int_\mathbf{y} \min\{p(\mathbf{y}|H_i, \mathbf{u}) P(H_i), p(\mathbf{y}|H_j, \mathbf{u}) P(H_j)\} d\mathbf{y} \end{aligned} \quad (10)$$

We now use the following inequality[10]:

$$\min\{a, b\} \leq a^{\frac{1}{2}} b^{\frac{1}{2}}, \quad (11)$$

which means that the integral over $\mathbf{y}$ can be bounded as follows:

$$\begin{aligned} &\int_\mathbf{y} \min\{p(\mathbf{y}|H_i, \mathbf{u}) P(H_i), p(\mathbf{y}|H_j, \mathbf{u}) P(H_j)\} d\mathbf{y} \\ &\leq P(H_i)^{\frac{1}{2}} P(H_j)^{\frac{1}{2}} \int_\mathbf{y} p^{\frac{1}{2}}(\mathbf{y}|H_i, \mathbf{u}) p^{\frac{1}{2}}(\mathbf{y}|H_j, \mathbf{u}) d\mathbf{y} \end{aligned} \quad (12)$$

Furthermore, if the distributions are Gaussian, such that $p(\mathbf{y}|H_i) = \mathcal{N}(\mu_i, \Sigma_i)$ and $p(\mathbf{y}|H_j) = \mathcal{N}(\mu_j, \Sigma_j)$, then this integral can be evaluated in closed form to give:

$$\int_\mathbf{y} p^{\frac{1}{2}}(\mathbf{y}|H_i, \mathbf{u}) p^{\frac{1}{2}}(\mathbf{y}|H_j, \mathbf{u}) d\mathbf{y} = e^{-k(i,j)} \quad (13)$$

where:

$$\begin{aligned} k(i,j) &= \frac{1}{4}(\mu_\mathbf{j} - \mu_\mathbf{i})^T \left[\Sigma_i + \Sigma_j\right]^{-1} (\mu_\mathbf{j} - \mu_\mathbf{i}) \\ &\quad + \frac{1}{2} ln \frac{\left|\frac{\Sigma_i + \Sigma_j}{2}\right|}{\sqrt{|\Sigma_i||\Sigma_j|}}. \end{aligned} \quad (14)$$

We therefore have:

$$P(\mathbf{y} \in \Re_j, H_i|\mathbf{u}) + P(\mathbf{y} \in \Re_i, H_j|\mathbf{u}) \leq P(H_i)^{\frac{1}{2}} P(H_j)^{\frac{1}{2}} e^{-k(i,j)}, \quad (15)$$

which leads to a new upper bound on the probability of model selection error:

$$P(error) \leq \sum_i \sum_{j>i} P(H_i)^{\frac{1}{2}} P(H_j)^{\frac{1}{2}} e^{-k(i,j)} \quad (16)$$

where $k(i,j)$ is defined in (14). □

We have therefore introduced a new upper bound on the probability of model selection error between an arbitrary number of models. This bound can be evaluated in closed form, and is a summation over Gaussian-like forms. In Section IV we describe how this bound can be used as an optimization criterion for finite horizon control design.

IV. Finite Horizon Formulation

As noted by [9], the design of control inputs to minimize the Bayes risk is a problem of optimal trajectory design for dynamic systems. In typical optimal trajectory design problems, an optimized sequence of control inputs is designed so that a system passes through a sequence of predicted states that minimize some cost function. In the model discrimination problem, we would like to design an optimal sequence of control inputs so that the system passes through a trajectory of state *distributions* that minimize the probability of model selection error (Bayes Risk).

Optimal trajectory design for linear, discrete-time systems can be posed as a constrained optimization problem[11]. Given a tractable optimization criterion, this problem can be solved using techniques such as Sequential Quadratic Programming (SQP)[12]. In this section we pose the problem of model discrimination between an arbitrary number of models as a finite horizon trajectory design problem, using a similar approach to [9], and show that it can be solved using SQP.

A. A Tractable Cost Function for Model Discrimination

We would like to plan a finite sequence of inputs in order to minimize the upper bound on the probability of error. This form of planning is known as *finite horizon* planning. In this case, if the horizon is of length k, we are concerned with a sequence of observations $\mathbf{y}_{t+1}...\mathbf{y}_{t+k}$ and a sequence of inputs $\mathbf{u}_t...\mathbf{u}_{t+k-1}$. We define:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_{t+1}^T & \mathbf{y}_{t+2}^T & \cdots & \mathbf{y}_{t+k}^T \end{bmatrix}^T$$
$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_t^T & \mathbf{u}_{t+1}^T & \cdots & \mathbf{u}_{t+k-1}^T \end{bmatrix}^T \quad (17)$$

We assume that model selection is carried out by a Bayes optimal classifier that makes its decision based on the horizon of all k observations. Due to uncertainty in the initial state and noise, the future observations $\mathbf{y}_{t+1}...\mathbf{y}_{t+k}$ are random variables, which we denote $Y_{t+1}...Y_{t+k}$. Under the assumptions in Section I-A, Y_{t+i} is normally distributed given a sequence of inputs $\mathbf{u}$ and given a model H_l. We now define $\mu_{t+i,l}$ and $\Sigma_{t+i,l}$ for time steps $i = 1, ..., k$ and hypotheses H_l such that:

$$p_{Y_{t+i}}(\mathbf{y}|H_l, \mathbf{u}) = \mathcal{N}(\mu_{t+i,l}, \Sigma_{t+i,l}) \quad (18)$$

Then the vector of all observations $Y = [Y_{t+1}^T ... Y_{t+k}^T]^T$ is a vector of normally distributed random variables given a sequence of inputs and a hypothesis. We define μ_l and Σ_l to be the mean and covariance of the vector of all observations such that:

$$p_{Y_l}(\mathbf{y}|H_l, \mathbf{u}) = \mathcal{N}(\mu_l, \Sigma_l) \quad (19)$$

From the above definitions the distribution of Y_l is given by:

$$\mu_l = \left[\mu_{t+1,l}^T ... \mu_{t+k,l}^T\right]^T \quad (20)$$

$$[\Sigma_l]_{i,j} = E\Big[\left([Y]_i - [\mu_l]_i\right)\left([Y]_j - [\mu_l]_j\right)\Big|H_l\Big] \quad (21)$$

Here $[\cdot]_i$ denotes the value at index i into the vector, and similarly $[\cdot]_{i,j}$ denotes the value at index (i,j) into the matrix.

Having defined μ_l and Σ_l for all l, the bound given in (3) provides an upper bound for the probability of error when using the entire sequence of observations from time $t+1$ to time $t+k$. We therefore use this bound as a cost function on the finite horizon optimization formulation:

$$J = \sum_i \sum_{j>i} P(H_i)^{\frac{1}{2}} P(H_j)^{\frac{1}{2}} e^{-k(i,j)} \quad (22)$$

where $k(i,j)$ is defined in (14).

Given a hypothesis H_l, the system equations in Fig. 1 are fully known. Hence the distribution $p(\mathbf{y}|H_l, \mathbf{u})$ can be calculated for all l. Explicit expressions for μ_l and Σ_l are given in [9]. For the sake of brevity, they are not repeated here, however we note two important properties:

1) The equation for the mean of the predicted distribution of Y is linear in the control inputs $\mathbf{u}$.
2) The covariance of the predicted distribution of Y is not a function of the control inputs $\mathbf{u}$.

These properties mean that the criterion in (22) has a particularly tractable form, enabling it to be used in a constrained optimization formulation.

B. Linear Constraints

A powerful aspect of the constrained finite horizon formulation is that optimal input sequences can be found subject to hard constraints on the control inputs. This can be used to model actuator saturation, for example, by constraining $u_{min} < u_{t+i} < u_{max}$. In [9] we showed that constraints on the expected system state and inputs can be expressed as linear constraints on the control inputs. The expected system state conditioned on a hypothesis H_l is a linear function of the control inputs:

$$E[\mathbf{x}_{t+i}|H_l] = \sum_{j=0}^{i-1} A_l^{i-j-1}(B_l \mathbf{u}_{t+j}) + A_l^i \hat{\mathbf{x}}_t, \quad (23)$$

where $\hat{\mathbf{x}}_t$ is the mean of the initial system state, Hence constraints on the expected system state of the form

$E[\mathbf{x}_{t+i}|H_l] = goal$ or $\mathbf{x}_{min} < E[\mathbf{x}_{t+i}|H_l] < \mathbf{x}_{max}$ are linear constraints in the control inputs. As noted in [9], by imposing such constraints, we can:

1) Ensure that a certain task, defined in terms of the expected system state, is fulfilled
2) Ensure that the mean of system stays within a 'safe' operating region or within a valid linearization region
3) Ensure that the system ends the experiment in the same region as it started.

Here, we have restricted out attention to linear constraints, since these lead to a tractable optimization problem. The general formulation, however, applies to nonlinear constraints.

C. Summary

We have shown that the problem of designing a sequence of optimal control inputs to discriminate between an arbitrary number of models can be posed as a finite-horizon trajectory design problem. The resulting optimization minimizes a novel, tractable upper bound on the probability of model selection error, and imposes constraints on the expected system state and control inputs to ensure that a defined task is fulfilled and that actuator limits are not violated. This optimization can be solved efficiently using existing methods such as Sequential Quadratic Programmming,

V. Simulation

In this section we present results from a number of discrimination tasks. We use the aircraft example presented in Section I-A, where we consider single-point failures of the pitch rate sensor, vertical velocity sensor and elevator actuator. In each of these tasks we control the the system by constraining the expected state. In Section V-A we constrain the aircraft to remain within an altitude envelope. For the sake of comparison, in Section V-B we show results from a manually generated identification sequence. In Section V-C the aircraft carries out an altitude change maneuver, while optimally detecting faults.

A. Altitude Envelope

Fig. 4 and Fig.5 show results from a fault detection scenario where the aircraft in Section I-A is constrained to remain within a flight envelope around an altitude of $100m$. The elevator angle is constrained to be at most $0.25rad$ in magnitude. The prior probabilities of models H_0 through H_3 are 0.75, 0.1, 0.05 and 0.1 respectively. We assume that these priors have been generated by a Multiple Model estimator running up until time t. The optimized control input yields an upper bound on the probability of error of 0.0013. The *bang-bang* nature of the optimized control input highlights the fact that by optimizing up against hard constraints the discrimination power of the input signal can be much greater than a power-bounded auxiliary signal.

B. Manually Generated Sequence

In order to identify the longitudinal dynamics of an aircraft, pilots typically use a doublet control input[13]. Fig. 6 and Fig. 7 show the results for such a control input,

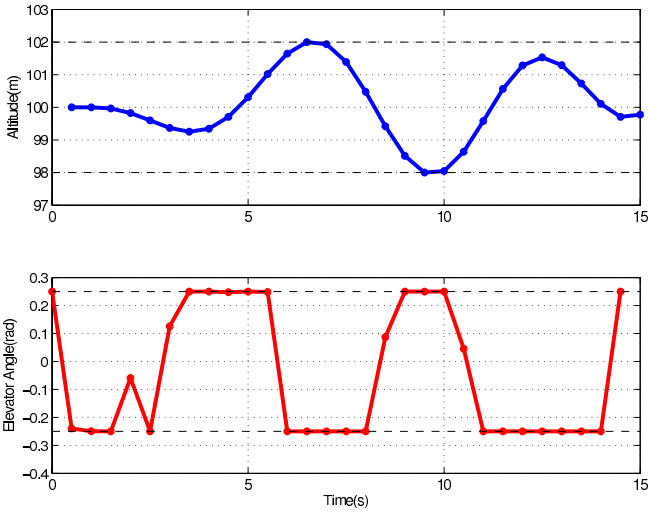

Fig. 4. Optimized input design for aircraft flight envelope scenario. Top: Altitude of aircraft given nominal operation. Bottom: Optimized sequence of control inputs. The optimal control discriminates between the different models while ensuring that in the nominal case, the aircraft altitude remains between 98 and $102m$. The resulting upper bound on the probability of model selection error is 0.0013.

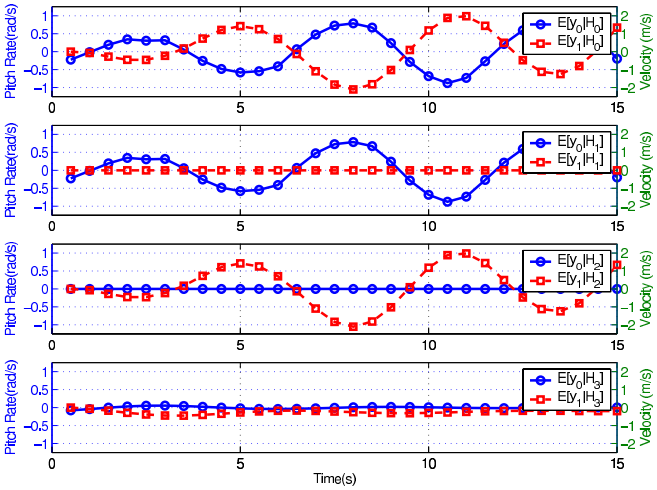

Fig. 5. Expected observations for aircraft flight envelope scenario with optimized control input. The top plot shows the nominal case, where there are no faults. In the second and third plots, the pitch rate sensor and vertical velocity sensors are faulty, respectively. In the bottom plot, the elevator actuator is faulty. The optimized control input ensures that the observation sequences in each case are as different as possible, in order to minimize the probability of model selection error.

with the same actuator limits as for the optimized control sequence. The resulting upper bound on the probability of error is 0.063. Hence the optimized sequence in Fig. 4 has significantly greater discrimination power than the manually generated sequence.

C. Altitude Change Maneuver

The new method can use constraints to ensure that a given control task is performed, while optimizing with respect to discrimination. This is demonstrated in Fig. 8, where the aircraft carries out a maneuver that changes its altitude from $100m$ to $120m$, conditioned on the elevator actuator being

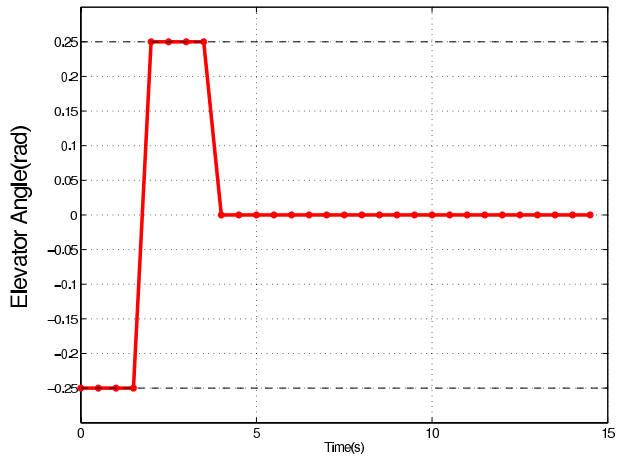

Fig. 6. Typical manually generated identification sequence. This doublet form is used by pilots to perform aircraft system identification. The resulting upper bound on the probability of error is 0.063.

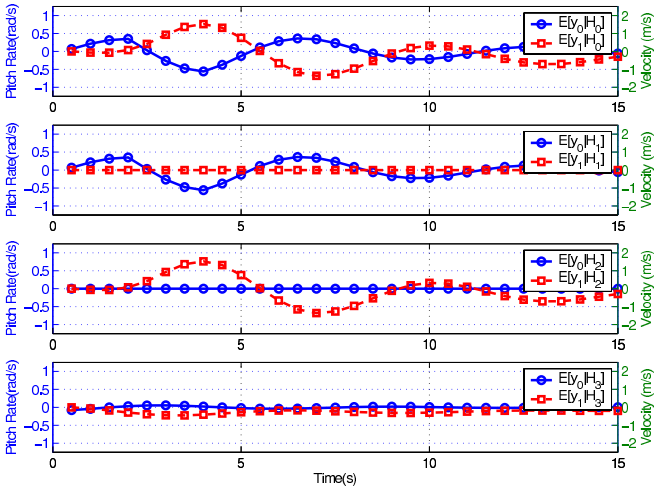

Fig. 7. Expected observations for aircraft flight envelope scenario with manually control input. The top plot shows the nominal case, where there are no faults. In the second and third plots, the pitch rate sensor and vertical velocity sensors are faulty, respectively. In the bottom plot, the elevator actuator is faulty. The pitch rate failure and elevator actuator failure cases yield somewhat similar observations, hence the upper bound on the probability of failure is larger than for the optimized control input.

functional. The discrimination-optimal control sequence is compared to the fuel-optimal one. Optimizing with respect to discrimination yields an upper bound on the probability of error of 0.0011, while optimizing with respect to fuel gives an upper bound on the probability of error of 0.12. Hence a significant improvement in fault detection can be achieved by using control inputs designed for model discrimination, rather than those designed to optimize some other criterion and employing only passive model selection.

VI. Conclusion

This paper presents a novel algorithm for model discrimination that uses a constrained finite horizon control approach to ensure that a defined control objective is satisfied while optimizing with regard to model discrimination. A

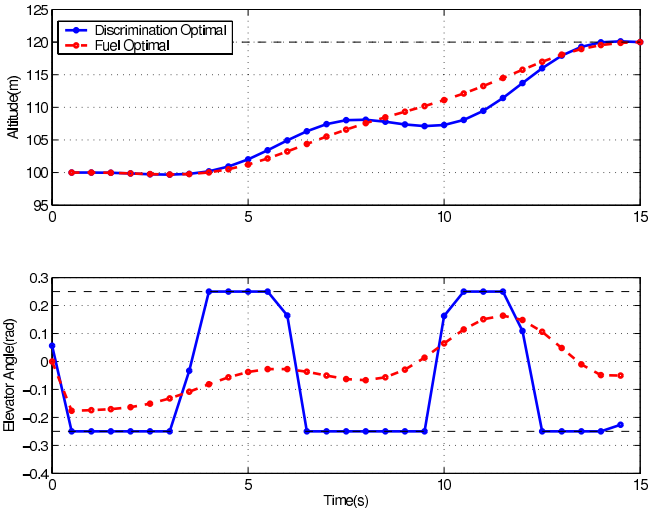

Fig. 8. Discrimination-optimal and fuel-optimal control design for altitude change maneuver. The discrimination-optimal sequence gives an upper bound on the probability of error of 0.0011, while the fuel-optimal sequence gives a bound of 0.12.

new, tractable upper bound on the probability of model selection error allows the problem to be solved using existing optimization methods. In contrast to prior work, the model discrimination method presented here applies to an arbitrary number of linear dynamic models. Simulation results show that compared to a typical human-generated identification sequence, and a control sequence optimized with regard to fuel consumption, the new method significantly reduces the upper bound on the probability of error.

References

[1] A. Pouliezos and G. Stavrakakis, *Real Time Fault Monitoring of Industrial Processes*. Boston: Kluwer Academic, 1994.

[2] J. Chen and R. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Boston: Kluwer Academic, 1999.

[3] P. Hanlon and P. Maybeck, "Multiple-model adaptive estimation using a residual correlation kalman filter bank," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, pp. 393–406, Apr. 2000.

[4] A. Willsky, "A survey of design methods for failure detection in dynamic systems," *Automatica*, vol. 12, pp. 601–611, 1976.

[5] R. Nikoukhah, S. L. Campbell, K. Horton, and F. Delebecque, "Auxiliary signal design for robust multi-model identification," *IEEE Trans. Automat. Contr.*, vol. 47, pp. 158–163, 2002.

[6] R. Nikoukhah, S. L. Campbell, and F. Delebecque, "Detection signal design for failure detection: a robust approach," *Int. J Adaptive control and Signal Processing*, vol. 14, pp. 701–724, 2000.

[7] X. J. Zhang, "Auxiliary signal design in fault detection and diagnosis," *NASA STI/Recon Technical Report A*, vol. 90, pp. 24 257–+, 1989.

[8] F. Keresteciogl u, *Change Detection and Input Design in Dynamical Systems*. New York: Wiley, 1993.

[9] L. Blackmore and B. C. Williams, "Finite horizon control design for optimal model discrimination," in *Proc. Control and Decision Conference 2005*, Seville, Spain, Dec. 2005.

[10] R. Duda, P. Hart, and D. Stork, *Pattern Classification (2nd ed.)*. Wiley Interscience, 2000.

[11] T. Schouwenaars, B. D. Moor, E. Feron, and J. How, "Mixed integer programming for multi-vehicle path planning," in *Proc. European Control Conference*, 2001.

[12] J. Nocedal and S. Wright, *Numerical Optimization*. Springer, 1999.

[13] J. Jang and C. Tomlin, "Longitudinal stability augmentation system design of the stanford dragonfly uav using a single gps receiver," in *Proceedings of the AIAA GNC conference*, Austin, Texas, Aug. 2003.