The Design of Auxiliary Signals for Robust Active Failure Detection in Uncertain Systems

Stephen L. Campbell
Department of Mathematics
North Carolina State University
Raleigh, NC 27695-8205, USA

Ramine Nikoukhah
INRIA
Rocquencourt BP 105
78153 Le Chesnay Cedex, France.

Abstract

By modeling the normal and the failed behaviors of a process by two or more linear uncertain systems, failure detectability in linear systems can be seen as a linear multi-model identification problem. This paper describes an active approach for multi-model identification using minimal auxiliary signals. Both additive and model uncertainty are included in this approach.

1 Introduction

By modeling the normal and the failed behaviors of a process by two or more linear uncertain systems, failure detectability in linear systems can be seen as a linear multi-model identification problem. Often there is no guarantee that one of the models can be ruled out by simply observing the inputs and outputs of the system. For example if the input to a linear system at rest is zero, the output would remain zero, and zero input/output is consistent with both models. Also, failure may not be apparent during normal operations. For example, a failure of the brakes may not be apparent while driving unless the driver performs a special test, such as tapping the brakes lightly.

For this reason a test signal, usually referred to as *auxiliary signal* [3, 4, 12], is sometimes injected into the system to expose its behavior and facilitate the detection (identification) of the failure. This means that the inputs of the system are partially or completely taken over by the failure detector mechanism for a period of time during which the auxiliary signal is injected into the system and the failure detection test is performed based on the measurements of the inputs and outputs.

In this paper we focus on the case when there are two possible models called Model 0 and Model 1. Let v represent the inputs taken over by the failure detector mechanism, y the outputs of the system, and $\mathcal{A}_i(v)$ the set of outputs y consistent with Model i for a given input v. Then an auxiliary signal v guarantees failure detection if and only if

$$\mathcal{A}_0(v) \cap \mathcal{A}_1(v) = \emptyset. \tag{1.1}$$

We call such a v a *proper auxiliary signal*.

If an auxiliary signal satisfying (1.1) exists, it is usually not difficult to find one, since unreasonably "large" signals often do the job. But such signals cannot be used in practice. For example, if the auxiliary signal is a mechanical force, there are always limitations on it (actuator characteristics). Similarly if it is an electrical signal, there are limitations imposed by the electrical components. But there are even more stringent conditions in practice imposed by the desire that the system continue to operate in a reasonable manner during the test period. Thus the test period should be short and the effect of the auxiliary signal on the system minimal. For example, in most cases, it is required that the test signal have minimum energy.

In this paper we develop a method for constructing optimal auxiliary signals in the two-model case. This paper is part of an ongoing investigation begun in [5, 6, 7]. It differs from previous

papers in several key respects. The only earlier papers to consider model uncertainty in addition to additive uncertainty are [10, 11]. However, this paper differs from [10] in the use of t dependent criteria, and it differs from [10, 11] in the careful consideration of implementation of the test, the discussion of decision making, and the inclusion of proofs. Our approach can be applied to more general situations but space prohibits discussing those extensions here.

2 Basic Theory

The general models that we consider are of the form

$$\begin{aligned} \dot{x}_i &= A_i x_i + B_i v + M_i \nu_i, && \text{(2.2a)} \\ E_i y &= C_i x_i + D_i v + N_i \nu_i && \text{(2.2b)} \end{aligned}$$

where $i = 0, 1$ correspond to normal and failed system models respectively. The v is the auxiliary signal which is computed prior to the test while y are outputs that become known during the test. Since v, y are known they are common to both models. However, y cannot be used to design v since v is computed before the test. The only condition on the system matrices is that the N_i's have full row rank. The unknown initial conditions $x_i(0)$ and noises ν are assumed to satisfy the bounds

$$\mathcal{S}_i(v, s) = x_i(0)^T P_{i0}^{-1} x_i(0) + \int_0^s \nu_i^T J_i \nu_i \, dt < 1, \ \forall s \in [0, T], \tag{2.3}$$

where the J_i's are signature matrices. That is, J_i is a diagonal matrix with 1 and -1 on the diagonal. If J_i has any negative entries, then we also assume that $\nu_i \in L^2$. Coefficients A, B, C, D, M, N can be time varying. Bounds other than 1 are included by rescaling system coefficients.

This formulation includes a number of different problems. For example, it includes the case of purely additive noise where $E_i = I$ and $J_i = I$. In that case we need only consider $s = T$ in (2.3) since the integrand is non-negative and the maximum value of the integral occurs at $s = T$ [2, 6].

2.1 Model uncertainty

Our problem formulation also includes problems with model uncertainty including some of those studied in [9, 8]. To see this, we suppress the i coefficient, and suppose that we have a system with uncertain coefficients and uncertain initial conditions given by (here $\bar{\sigma}$ is the largest singular value)

$$\begin{aligned} \dot{x} &= (\bar{A} + M\Delta G)x + (\bar{B} + M\Delta H)v, && \text{(2.4a)} \\ y &= (\bar{C} + \bar{N}\Delta G)x + (\bar{D} + \bar{N}\Delta H)v && \text{(2.4b)} \\ & \overline{\sigma}(\Delta(t)) \le d_1 && \text{(2.4c)} \\ & x(0)^T P_0^{-1} x(0) < d && \text{(2.4d)} \end{aligned}$$

By adjusting the other terms such as H, G we may assume that $d_1 = 1$. Suppose that system (2.4) is modeled over the period $[0, T]$ as follows:

$$\begin{aligned} \dot{x} &= \bar{A}x + \bar{B}v + \bar{M}\bar{\nu}, && \text{(2.5a)} \\ 0 &= Gx + Hv - z, && \text{(2.5b)} \\ y &= \bar{C}x + \bar{D}v + \bar{N}\bar{\nu} && \text{(2.5c)} \end{aligned}$$

where $\bar{\nu}$ and z are respectively the noise input and noise output representing model uncertainty. Note that (2.5b), (2.5c) can be written in the form

$$\begin{pmatrix} 0 \\ I \end{pmatrix} y = \begin{pmatrix} G \\ \bar{C} \end{pmatrix} x + \begin{pmatrix} H \\ \bar{D} \end{pmatrix} v + \begin{pmatrix} 0 & -I \\ \bar{N} & 0 \end{pmatrix} \begin{pmatrix} z \\ \bar{\nu} \end{pmatrix}$$

which is in the form of (2.2b) with $\nu_i = \begin{pmatrix} z \\ \nu \end{pmatrix}$. We assume that N is full row rank. Letting

$$\nu = \Delta z \tag{2.6}$$

we see that (2.5) become (2.4a), (2.4b). Condition (2.6) and (2.4c) imply that $\|\nu(t)\| \leq \|z(t)\|$ for all $0 \leq t \leq T$. Thus we have that

$$\int_0^s (\|\nu\|^2 - \|z\|^2) < 0, \quad \forall s \in [0, T] \tag{2.7}$$

Combining (2.4d) and (2.7) we get the uncertainty on the initial conditions and ν satisfies

$$x(0)^T P_0^{-1} x(0) + \int_0^s (\|\nu\|^2 - \|z\|^2) dt < d, \ \forall s \in [0, T]. \tag{2.8}$$

which is in the form of (2.3).

The problem we consider is system (2.5) along with the noise bound (2.8). Note that the noise bound (2.8) is more general than (2.4d) and (2.7). Letting $s = 0$ we see that (2.8) implies (2.4d). However, (2.8) does not also imply (2.7). This means that any test derived for the problem we consider here will actually be conservative for the problem (2.4).

By using $Pv = \Delta z$ where P is a projection instead of (2.6) our general formulation includes models with both model uncertainty and additive uncertainty $(I - P)\nu$.

2.2 General problem

Suppose we have y, given a v, consistent with one of the models. We seek an optimal v for which observation of y provides enough information to decide from which model y has been generated. That is, we seek an optimal proper v. The first step is to characterize the proper v. That is those v for which there exist no solution to (2.2) and (2.3) for $i = 0$ and 1 simultaneously.

Note that since N_i's are full row rank, for any L_2 functions v, u and y, there exist L_2 functions ν_i satisfying (2.2). So, the non-existence of solution to (2.2) and (2.3) is equivalent to:

$$\sigma(v, s) \geq 1 \tag{2.9}$$

where

$$\sigma(v, s) = \inf_{\substack{\nu_0, \nu_1, y \\ x_0, x_1}} \max(\mathcal{S}_0(v, s), \mathcal{S}_1(v, s)), \tag{2.10}$$

subject to (2.2), $i = 0, 1$. We need a more computationally amenable criteria. Theorem 4.1 allows us to switch the order of the inf-max (2.10) so that we can reformulate (2.10) as:

$$\sigma(v, s) = \max_{\beta \in [0,1]} \phi_\beta(v, s) \tag{2.11}$$

where

$$\phi_\beta(v,s) = \inf_{\substack{\nu_0,\nu_1,y \\ x_0,x_1}} \beta\mathcal{S}_0(v,s) + (1-\beta)\mathcal{S}_1(v,s) \tag{2.12}$$

subject to (2.2) for $i = 0, 1$. Let

$$x = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix},\ \nu = \begin{pmatrix} \nu_0 \\ \nu_1 \end{pmatrix}, \qquad A = \begin{pmatrix} A_0 & 0 \\ 0 & A_1 \end{pmatrix},\ M = \begin{pmatrix} M_0 & 0 \\ 0 & M_1 \end{pmatrix},$$

$$B = \begin{pmatrix} B_0 \\ B_1 \end{pmatrix},\ D = F_0D_0 + F_1D_1, \qquad C = \begin{pmatrix} F_0C_0 & F_1C_1 \end{pmatrix},\ N = \begin{pmatrix} F_0N_0 & F_1N_1 \end{pmatrix},$$

$$P_\beta^{-1} = \begin{pmatrix} \beta P_{0,0}^{-1} & 0 \\ 0 & (1-\beta)P_{1,0}^{-1} \end{pmatrix}, \qquad J_\beta = \begin{pmatrix} \beta J_0 & 0 \\ 0 & (1-\beta)J_1 \end{pmatrix},$$

where (here $X^\perp$ denotes a maximal row rank left annihilator of X)

$$F = \begin{pmatrix} F_0 & F_1 \end{pmatrix} = \begin{pmatrix} E_0 \\ E_1 \end{pmatrix}^\perp, \tag{2.13}$$

we can reformulate (2.12) as follows:

$$\phi_\beta(v,s) = \inf_{\nu,x} x(0)^T P_\beta^{-1} x(0) + \int_0^s \nu^T J_\beta \nu\, dt \tag{2.14}$$

subject to

$$\dot{x} = Ax + Bv + M\nu \tag{2.15a}$$

$$0 = Cx + Dv + N\nu. \tag{2.15b}$$

Lemma 2.1. *Let $\mathcal{B}$ be the set of all β such that, for all $s \le T$, $\phi_\beta(v,s) > -\infty$. Then the set $\mathcal{B}$ is a sub-interval of $[0,1]$ and is independent of v.*

Proof. Refer to Lemma 4.4 □

If $\mathcal{B} = \emptyset$, then the $\phi_\beta(v,s)$ for all β and consequently $\sigma(v,s)$ is equal to $-\infty$ for some s which means that condition (2.9) can never be satisfied and thus there exists no proper auxiliary signal v. In that case, there is no point in solving the optimization problem. Thus, we need only consider the case where $\mathcal{B}$ is not empty. Theorem 2.1 provides sufficient conditions.

Theorem 2.1. *Suppose for some $\beta \in [0,1]$, that*

$$N_\perp{}^T J_\beta N_\perp > 0, \ \ \forall t \in [0,T] \tag{2.16}$$

and the Riccati equation

$$\dot{P} = (A - S_\beta R_\beta^{-1}C)P + P(A - S_\beta R_\beta^{-1}C)^T - PC^T R_\beta^{-1} CP + Q_\beta - S_\beta R_\beta^{-1} S_\beta^T, P(0) = P_\beta \tag{2.17}$$

has a solution on $[0,T]$ where $\begin{pmatrix} Q_\beta & S_\beta \\ S_\beta^T & R_\beta \end{pmatrix} = \begin{pmatrix} M \\ N \end{pmatrix} J_\beta^{-1} \begin{pmatrix} M \\ N \end{pmatrix}^T$. Then $\beta \in \mathcal{B}$.

This result follows easily from Theorem 4.2. We shall assume from here on that there exists at least one β for which the two conditions of Theorem 2.1 are satisfied.

2.3 Construction of an optimal proper auxiliary signal

To simplify the discussion, we use the optimality criterion of minimizing the L_2 norm of the auxiliary signal. So the problem to solve is:

$$\min_{v} |||v|||^2, \quad \text{subject to} \max_{\substack{\beta\in[0,1]\\ s\in[0,T]}} \phi_\beta(v,s) \geq 1 \tag{2.18}$$

where $|||v|||^2 = \int_0^T \|v\|^2\, dt$. The maximum value of $\phi_\beta(v,s)$ does not always occur at $s = T$, although it often does. We reformulate (2.18) as follows

$$\lambda_{\beta,s} = \max_{v\neq 0} \frac{\phi_\beta(v,s)}{|||v|||^2}. \tag{2.19}$$

So that we end up having to solve the following problem

$$\max_{v} \inf_{\nu,x} x(0)^T P_\beta^{-1} x(0) + \int_0^s \nu^T J_\beta \nu - \lambda \|v\|^2\, dt \tag{2.20}$$

subject to (2.15a) and (2.15b).

Theorem 2.2. *Suppose the two conditions of Theorem 2.1 are satisfied. Then $\lambda_{\beta,s}$ is the infimum of the set of all λ for which the Riccati equation*

$$\dot{P} = (A - S_{\lambda,\beta}R_{\lambda,\beta}^{-1}C)P + P(A - S_{\lambda,\beta}R_{\lambda,\beta}^{-1}C)^T - PC^T R_{\lambda,\beta}^{-1} CP + Q_{\lambda,\beta} - S_{\lambda,\beta}R_{\lambda,\beta}^{-1}S_{\lambda,\beta}^T, \quad P(0) = P_\beta \tag{2.21}$$

where $\begin{pmatrix} Q_{\lambda,\beta} & S_{\lambda,\beta} \\ S_{\lambda,\beta}^T & R_{\lambda,\beta} \end{pmatrix} = \begin{pmatrix} M & B \\ N & D \end{pmatrix} \begin{pmatrix} J_\beta & 0 \\ 0 & -\lambda I \end{pmatrix}^{-1} \begin{pmatrix} M & B \\ N & D \end{pmatrix}^T$ *has a solution on [0,s].*

This result shows that $\lambda_{\beta,s}$ is an increasing function of s and thus it allows us to compute

$$\lambda_\beta = \max_{s\leq T} \lambda_{\beta,s} \tag{2.22}$$

by simply performing a "λ-iteration" by testing for the existence of a solution P to (2.21) over $[0,T]$. Then λ^* and β^* are obtained from

$$\lambda^* = \max_{\beta\in\mathcal{B}} \lambda_\beta \tag{2.23}$$

We define γ^* by $\gamma^* = \frac{1}{\sqrt{\lambda^*}}$. The values of λ^* and β^* are later used to compute the optimal auxiliary signal v^*.

Note that the Riccati equation (2.21) for $\lambda = \lambda^*$ and $\beta = \beta^*$ does not necessarily have an escape time T. Its solution may diverge before the end of the interval. We denote the actual escape time T^*. In most cases T^* is equal to T, but in general we only have that $T^* \leq T$.

Lemma 2.2. *The two-point boundary-value system:*

$$\frac{d}{dt}\begin{pmatrix} x \\ \zeta \end{pmatrix} = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \begin{pmatrix} x \\ \zeta \end{pmatrix} \tag{2.24}$$

with boundary conditions:

$$x(0) = P_{\beta^*}\zeta(0) \tag{2.25a}$$
$$\zeta(\tau) = 0. \tag{2.25b}$$

where

$$\Omega_{11} = -\Omega_{22}^T = A - S_{\lambda^*,\beta^*}R_{\lambda^*,\beta^*}^{-1}C \tag{2.26a}$$
$$\Omega_{12} = Q_{\lambda^*,\beta^*} - S_{\lambda^*,\beta^*}R_{\lambda^*,\beta^*}^{-1}S_{\lambda^*,\beta^*}^T \tag{2.26b}$$
$$\Omega_{21} = C^T R_{\lambda^*,\beta^*}^{-1}C \tag{2.26c}$$

is well-posed for $\tau < T^*$ but it is not well-posed, i.e., it has non trivial solutions (x^*, ζ^*), for $\tau = T^*$.

See Section 4.2 for a proof.

Theorem 2.3. *An optimal auxiliary signal is*

$$v^* = \alpha((B + S_{\lambda^*,\beta^*}R_{\lambda^*,\beta^*}^{-1}D)^T\zeta + D^T R_{\lambda^*,\beta^*}^{-1}Cx) \tag{2.27}$$

where α is a constant such that $|||v^*||| = 1/\gamma^*$.

Proof. An optimal auxiliary signal is obtained by setting to zero the first variation associated to the optimization problem (2.20). Denoting the Lagrange multipliers associated with the constraints (2.15a) and (2.15b) respectively ζ and μ, we obtain

$$\dot{x} = Ax + Bv + M\nu \tag{2.28a}$$
$$0 = Cx + Dv + N\nu \tag{2.28b}$$
$$\dot{\zeta} = -A^T\zeta + C^T\mu \tag{2.28c}$$
$$J_\beta\nu = M^T\zeta - N^T\mu \tag{2.28d}$$
$$-\lambda v = B^T\zeta - D^T\mu \tag{2.28e}$$

which in particular, after replacing λ and β respectively with λ^* and β^*, implies

$$\begin{pmatrix} N & D & 0 \\ J_\beta & 0 & N^T \\ 0 & \lambda^* I & -D^T \end{pmatrix} \begin{pmatrix} \nu \\ v \\ \mu \end{pmatrix} = \begin{pmatrix} Cx \\ M^T\zeta \\ B^T\zeta \end{pmatrix} \tag{2.29}$$

from which we get

$$v = (B + S_{\lambda^*,\beta^*}R_{\lambda^*,\beta^*}^{-1}D)^T\zeta + D^T R_{\lambda^*,\beta^*}^{-1}Cx \tag{2.30a}$$
$$\mu = R_{\lambda^*,\beta^*}^{-1}(\lambda^* Cx - S_{\lambda^*,\beta^*}^T\zeta) \tag{2.30b}$$
$$\nu = J_{\beta^*}^{-1}(M^T\zeta - \lambda^* N^T R_{\lambda^*,\beta^*}(S_{\lambda^*,\beta^*}^T\zeta + Cx)). \tag{2.30c}$$

System (2.24) is then obtained by substituting these expressions for v, μ and ν in (2.28a)-(2.28e). Formula (2.27) is just a renormalization of (2.30a). □

Note that if $T^* < T$, we can reduce our test period by setting $T = T^*$ because the additional time is not of any use to us: it doesn't allow us to improve the separability index. So from here on we assume that $T = T^*$.

3 On-line detection tests

Once the auxiliary signal v is constructed, it can be used for on-line failure detection. Unlike the construction of v which is done off-line and thus can be computationally intensive, the on-line computation burden must be such that real-time implementation can be envisaged for the available hardware.

3.1 Standard solution

One solution to on-line detection problem would be simply to use a realizability test for each of the two models. That is, test whether y is consistent with Model i. In fact, since by construction $\mathcal{A}_0(v^*) \cap \mathcal{A}_1(v^*) = \emptyset$ in a perfect world, we would only have to implement a single realizability test for one of the two models. But in the real world the models are not perfect and the on-line detection test should be based on both realizability tests.

Theorem 3.1. *Consider model (2.2) with uncertainty model (2.3). Suppose that* ${N_i}_\perp{}^T \Gamma N_{i\perp} > 0, \ \forall t \in [0,T]$, *and the following Riccati equation has a solution on [0,T];*

$$\dot{P}_i = (A_i - S_i R_i^{-1} C_i) P_i + P_i (A_i - S_i R_i^{-1} C_i)^T - P_i C_i^T R_i^{-1} C_i P_i + Q_i - S_i R_i^{-1} S_i^T, \ \ P_i(0) = P_{i0} \tag{3.31}$$

where $\begin{pmatrix} Q_i & S_i \\ S_i^T & R_i \end{pmatrix} = \begin{pmatrix} M_i \\ N_i \end{pmatrix} J_i^{-1} \begin{pmatrix} M_i \\ N_i \end{pmatrix}^T$. *Then a realizability test for Model i is:*

$$\gamma_{i,s}(y) < 1, \quad \textit{for all } s \in [0,T] \tag{3.32}$$

where $\gamma_{i,s} = \int_0^s \mu_i^T R_i^{-1} \mu_i \, dt$ *and* μ_i *is the output of the following system*

$$\dot{\hat{x}}_i = A_i \hat{x}_i - (S_i + P_i {C_i}^T) R_i^{-1} \mu_i + B_i v^*, \quad \hat{x}_i(0) = 0 \tag{3.33a}$$

$$\mu_i = C_i \hat{x}_i - (E_i y - D_i v^*). \tag{3.33b}$$

Another advantage of using two realizability tests is that it could result in faster detection since a realizability test can detect "non-realizability" before the end of the test period T. The realizability tests for models (2.2a)-(2.2b) are obtained directly from the application of Theorem 4.2 and its corollary. Note that $y \notin \mathcal{A}_i(v)$ as soon as $\gamma_{i,s} \geq 1$ for some s.

3.2 Hyperplane test

It is sometimes possible to construct a "hyperplane" test to be used for on-line detection. A separating hyperplane doesn't always exist, but it does in many interesting cases when the $\mathcal{A}_i(v)$ are convex. The separating hyperplane test can be expressed as follows:

$$\int_0^T h(t)^T (y - y^*(t)) dt \lesseqgtr 0. \tag{3.34}$$

We have noted explicitly the dependence of h and y^* on t to emphasize that they almost always depend on time.

Theorem 3.2. *Suppose a separating hyperplane exists and that the two conditions of Theorem 2.1 are satisfied. Then, a separating hyperplane can be characterized as follows:*

$$h = \left(F \begin{pmatrix} E_0 \\ -E_1 \end{pmatrix} \right)^T R^{-1}_{\lambda^*,\beta^*} (Cx^* + Dv^* + Q\lambda^*) \tag{3.35}$$

and

$$y^* = \begin{pmatrix} E_0 \\ E_1 \end{pmatrix}^l \left(\begin{pmatrix} C_0 & 0 \\ 0 & C_1 \end{pmatrix} x^* + \begin{pmatrix} D_0 \\ D_1 \end{pmatrix} v^* + \begin{pmatrix} N_0 & 0 \\ 0 & N_1 \end{pmatrix} \nu^* \right). \tag{3.36}$$

Proof. The optimization problem that defines v can be written

$$\inf_{\nu,x,y_0,y_1} x(0)^T P_\beta^{-1} x(0) + \int_0^T \nu^T J_\beta \nu \, dt \tag{3.37a}$$

subject to

$$\begin{aligned} \dot{x} &= Ax + Bv + M\nu && (3.37b) \\ \begin{pmatrix} E_0 & 0 \\ 0 & E_1 \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} &= \begin{pmatrix} C_0 & 0 \\ 0 & C_1 \end{pmatrix} x + \begin{pmatrix} D_0 & 0 \\ 0 & D_1 \end{pmatrix} v + \begin{pmatrix} N_0 & 0 \\ 0 & N_1 \end{pmatrix} \nu && (3.37c) \\ y_0 - y_1 &= 0. && (3.37d) \end{aligned}$$

The Lagrange multiplier associated with the constraint (3.37d) gives h. We are of course interested in the case $v = v^*$. Computing the optimality conditions by performing a first order variation yields the following additional equations:

$$\begin{aligned} \dot{\zeta} + A^T \zeta + \begin{pmatrix} C_0 & 0 \\ 0 & C_1 \end{pmatrix}^T \omega &= 0 && (3.38a) \\ \begin{pmatrix} E_0 & 0 \\ 0 & E_1 \end{pmatrix}^T \omega + \begin{pmatrix} I \\ -I \end{pmatrix} h &= 0 && (3.38b) \\ J_\beta \nu - M^T \zeta - N^T \omega &= 0 && (3.38c) \end{aligned}$$

where ζ and ω are respectively the Lagrange multipliers associated with (3.37b) and (3.37c).

From (3.38b) it follows that $\begin{pmatrix} E_0 \\ E_1 \end{pmatrix}^T \omega = 0$. But thanks to (2.13) we know that columns of F^T form a basis for the null-space of $\begin{pmatrix} E_0 \\ E_1 \end{pmatrix}^T$ so that there exists a σ such that $\omega = F^T \sigma$. It is now straightforward to show that $\sigma = -R^{-1}_{\beta^*}(Cx + Dv + Q^*_\beta \zeta)$.. But $h = \begin{pmatrix} E_0^T & -E_1^T \end{pmatrix} F^T \sigma$ which gives (3.35). Then (3.36) follows easily (3.37c) by letting $y^* = y_0 = y_1$. □

3.3 Numerical issues

All the off-line computations required for the construction of the optimal auxiliary signal and the hyperplane test can be done efficiently using Scilab [1] or Matlab. The computation of $\lambda_{\beta,s}$ requires

a "λ-iteration" which needs the solution of the Riccati equation (2.21). This can be implemented using a standard ODE solver. Note that for a given β, this λ-iteration gives the $\max_s \lambda_{\beta,s}$.

The optimization over the scalar β can be done simply by discretizing the set $\mathcal{B}$ (in the worst case $\mathcal{B} = [0, 1]$). That gives us β^*.

Finally the computation of the optimal auxiliary signal and the hyperplane test (if applicable) requires the solution of the two-point boundary value system (2.24). This problem is not standard because this system is not well-posed. To find the solution, we need first the following result:

Lemma 3.1. *For any solution (x, ζ) of the boundary-value system (2.24), we have, over $[0, T)$,*

$$x = P\zeta \tag{3.39}$$

where P is the solution of the Riccati equation (2.21) with $\beta = \beta^*$ and $\lambda = \lambda^*$.

Proof Suppose that (x, ζ) satisfies (2.24)-(2.25a). Let $\hat{\zeta}$ be the solution of

$$\dot{\hat{\zeta}} = (\Omega_{21}P + \Omega_{22})\hat{\zeta}, \;\; \hat{\zeta}(0) = \zeta(0). \tag{3.40}$$

Let $\hat{x} = P\hat{\zeta}$. It is straightforward to verify that $(\hat{x}, \hat{\zeta})$ is a solution of (2.24)-(2.25a). But (x, ζ) and $(\hat{x}, \hat{\zeta})$ are solutions of (2.24) which satisfy the same initial condition and hence $x = \hat{x}$. □

We have seen that as t goes to T, $P(t)$ diverges so $\bar{P}(T) = \lim_{t\to T} P(t)^{-1}$ is singular. $\bar{P}(T)$ can be computed by inverting $P(t)$ for a t close to T; this may not be numerically reliable. Alternatively, when we get close to T (say at $t = T - \delta$ for some small δ) and $P(t)$ starts diverging, one switches from the Riccati equation (2.21) to the following equivalent Riccati equation (here $\bar{P} = P^{-1}$)

$$-\dot{\bar{P}} = \bar{P}(A - S_{\lambda^*,\beta^*}R^{-1}_{\lambda^*,\beta^*}C) + (A - S_{\lambda^*,\beta^*}R^{-1}_{\lambda^*,\beta^*}C)^T\bar{P} - C^T R^{-1}_{\lambda^*,\beta^*}C + \bar{P}(Q_{\lambda^*,\beta^*} - S_{\lambda^*,\beta^*}R^{-1}_{\lambda^*,\beta^*}S^T_{\lambda^*,\beta^*})\bar{P}, \;\; \bar{P}(T-\delta) = P^{-1}(T-\delta).$$

This equation can be integrated up to T yielding a singular $\bar{P}(T)$. We do not use this Riccati equation from the beginning ($t = 0$) since, as noted earlier, $P(t)$ becomes singular somewhere in the middle of the interval, which means that $\bar{P}$ goes through infinity at that point. Thus δ must be chosen small enough to avoid this singularity, but not too small in order to avoid numerical problems with the inversion of $P(T - \delta)$.

Once we have found $\bar{P}(T)$, noting that $\bar{P}(T)x(T) = \zeta(T) = 0$, we can let $x(T)$ be any non zero vector x_T in the null space of $\bar{P}(T)$. We can now find a non zero solution to (2.24) by taking as boundary (final) condition

$$x(T) = x_T, \;\; \zeta(T) = 0. \tag{3.41}$$

This system is well-posed and has a unique solution. However, since this system is not backward stable, its numerical solution can result in large errors. The way to avoid this problem is to use (2.24) with boundary (final) condition (3.41) simply to solve for (x, ζ) over a short interval, for example $[T - \delta, T]$. Then from $T - \delta$ on, use

$$\dot{\zeta} = (-A^T + C^T R^{-1}_{\lambda^*,\beta^*}CP + C^T R^{-1}_{\lambda^*,\beta^*}S^T_{\lambda^*,\beta^*})\zeta \tag{3.42}$$

to solve for ζ down to zero. This relation is obtained easily from (3.39). The optimal auxiliary signal is then given by (2.27).

4 Needed Theoretical Results

4.1 Optimization of the max of quadratic functionals

In our development the inf-max of noise measures plays a fundamental role. The actual form of the noise measures varies but the following situation covers all cases. We consider the following optimization problem

$$c = \inf_x \max(G_0(x), G_1(x)) \tag{4.1}$$

where $G_i(x)$ are quadratic functionals and $x \in \mathcal{X}$ where $\mathcal{X}$ is a real Hilbert space. That is, $\mathcal{X}$ has an inner product $<,>$ and contains all its limit points. In general $\mathcal{X}$ will be a direct sum of an L^2 space and a finite dimensional space. The functionals may be written

$$G_i(x) =< P_i x, x > + < x, q_i > + r_i \tag{4.2}$$

where P_i is a bounded symmetric operator. In the finite dimensional case (4.2) becomes $x^T P_i x + q_i^T x + r_i$, $i = 0, 1$, for matrices P_i, vectors q_i and scalars r_i.

We will need the following facts.

Lemma 4.1. *For any quadratic functional $G(x)$, if $G(x_0) = \alpha_0 > \alpha_1 = G(x_1)$, then there is a closed interval $[t_0, t_1]$ and a continuous function $x(t)$ on this interval such that $G(x(t_i)) = \alpha_i$ for $i = 0, 1$, $x(t_0) = x_0$, and $G(x(t))$ is monotonically non-increasing on the interval.*

Proof. Note that we do not assert that $x(t_1) = x_1$. In the actual use of this result later we really only need to be able to take $G(x(t_i))$ arbitrarily close to α_i. Thus we can assume without loss of generality that 0 is an isolated point of the spectrum of P. If G is non-negative the proof follows from just having $x(t)$ rescale x_0. For the general case note that P_i symmetric means we may decompose $\mathcal{X}$ into an orthogonal sum $\mathcal{X}_1 \oplus \mathcal{X}_2 \oplus \mathcal{X}_3$ and relative to this direct sum we have $P_i = P_+ \oplus -P_- \oplus 0$ where P_+, P_- are positive definite, and $x = x_+ \oplus x_- \oplus x_o$, $q = q_+ \oplus q_- \oplus q_o$. By a change of variables $x = z + a$ we may assume $q_+ = 0, q_- = 0$. Thus (4.2) becomes $G(x) =< P_+ x_+, x_+ > - < P_- x_-, x_- > + < x_0, q_0 > + r$. The Lemma now follows by considering the individual summands and scaling them separately. □

The optimization problem (4.1) can be re-written as follows. Let

$$c = \inf_x \max_{\beta \in [0,1]} \beta G_0(x) + (1 - \beta) G_1(x) \tag{4.3}$$

Lemma 4.2. *Suppose c is defined by (4.3). Let $\mathcal{S} = \{x | G_0(x) = G_1(x)\}$. If $c \geq 0$, then at least one of the following statements holds*

1. *c is the global minimum of $G_0(x)$*

$$c = \min_x G_0(x), \tag{4.4}$$

2. *c is the global minimum of $G_1(x)$*

$$c = \min_x G_1(x), \tag{4.5}$$

3. or

$$c = \inf_{x \in \mathcal{S}} G_0(x) = \inf_{x \in \mathcal{S}} G_1(x) \tag{4.6}$$

Note that the inf in (4.3) cannot always be replaced with a min even if $c \geq 0$ as can be seen from the example $c = \inf_{x_1,x_2} \max(x_1^2, 1 - x_1 x_2)$. Clearly here $c = 0$ but zero is not achieved by any (x_1, x_2).

Proof. Suppose that $c \geq 0$. Let δ be the inf in (4.6). Clearly $c \leq \delta$. If $c = \delta$ we are done since then (4.6) holds. Suppose then that $c < \delta$. Let $\{x_j\}$ be a sequence such that $\max\{G_0(x_j), G_1(x_j)\}$ decreases monotonically to c. Let

$$S_1 = \{x : G_0(x) > G_1(x)\}, \quad S_2 = \{x : G_1(x) > G_0(x)\}$$

Since $c < \delta$ there must be a subsequence of $\{x_j\}$ which lies in S_1 or S_2. We may assume that it is in S_1. Thus we have $G_0(x_j) \to c$. If c is the global min of G_0, then we are done. Suppose this is not the case. Then there is an $\hat{x} \in S_1$ with $c \leq G_0(\hat{x}) < \delta$ and an $\tilde{x} \in S_2$ with $G_0(\tilde{x}) < c \leq G_0(\hat{x})$. Applying Lemma 4.1 we get an $x(t)$ on $[0, 1]$ such that $x(0) = \hat{x}$ and $G_0(x(1)) = G_0(\tilde{x})$. Clearly $x(1) \in S_2$. But now consider the continuous scalar function $b(t) = G_0(x(t)) - G_1(x(t))$. We have $b(0) > 0$ and $b(1) < 0$ so there is a value $\hat{t}$ so that $b(\hat{t}) = 0$. But then this $x(\hat{t})$ is in S and $G_0(x(\hat{t})) < \delta$ because $G_0(x(t))$ decreases. But this contradicts the definition of δ. □

Lemma 4.3. *If $c \geq 0$, then there exists a $\beta \in [0, 1]$ such that for all x we have*

$$\beta G_0(x) + (1 - \beta) G_1(x) \geq 0 \tag{4.7}$$

Proof. We can suppose that there exists an x such that $G_1(x) < 0$ (otherwise, $G_1(x) \geq 0$ for all x and we can take $\beta = 0$). Also, clearly, for any x, if $G_1(x) < 0$ then $G_0(x) \geq 0$ since otherwise $c < 0$. We can now use Theorem 4.2.1 of [9] which states that for two quadratic functionals $\mathcal{G}_0(x)$ and $\mathcal{G}_1(x)$, if there exists $\bar{x}$ such that $\mathcal{G}_1(\bar{x}) > 0$, then the following conditions are equivalent:

1. $\mathcal{G}_0(x) \geq 0$ for all x such that $\mathcal{G}_1(x) \geq 0$;

2. there exists a constant $\tau \geq 0$ such that $\mathcal{G}_0(x) - \tau \mathcal{G}_1(x) \geq 0$, for all x.

For that, let $\mathcal{G}_1(x) = -G_1(x)$ and $\mathcal{G}_0(x) = G_0(x)$. Then from the above result follows that there exists a constant $\tau \geq 0$ such that $G_0(x) + \tau G_1(x) \geq 0$, for all x. If we let $\tau = \frac{1-\beta}{\beta}$ we obtain (4.7) which is the desired result. □

The preceding results enable us to prove the result which allows us to interchange the inf-max with a max-inf.

Theorem 4.1. *If $c \geq 0$ then*

$$c = \max_{\beta \in [0,1]} \inf_x \beta G_0(x) + (1 - \beta) G_1(x). \tag{4.8}$$

Proof. Since $\max\inf \leq \inf\max$, we know that

$$c \geq \max_{\beta\in[0,1]} \inf_x \beta G_0(x) + (1-\beta)G_1(x). \tag{4.9}$$

So if condition (4.4) holds, then the equality is obtain by setting $\beta = 1$. Similarly, if (4.5) holds, then the equality is obtain by setting $\beta = 0$. So the only case that we need to consider is when

$$c = \inf_{x\in\mathcal{S}} G_0(x) = \inf_{x\in\mathcal{S}} G_1(x). \tag{4.10}$$

Let $\mathcal{B}$ denote the set of all β for which $\beta G_0(x) + (1-\beta)G_1(x)$ is convex in x. This set is not empty because it contains at least one element as shown in Lemma 4.3, and it is closed. Thanks to (4.10), we have

$$\begin{aligned} c &= \inf_x \max_{\beta\in[0,1]} \beta G_0(x) + (1-\beta)G_1(x) = \inf_{x\in\mathcal{S}} \max_{\beta\in[0,1]} \beta G_0(x) + (1-\beta)G_1(x) \\ &= \inf_x \max_{\beta\in\mathcal{B}} \beta G_0(x) + (1-\beta)G_1(x) = \max_{\beta\in\mathcal{B}} \inf_x \beta G_0(x) + (1-\beta)G_1(x). \end{aligned} \tag{4.11}$$

But $\inf_x \beta G_0(x) + (1-\beta)G_1(x)$ is $-\infty$ if β is not in $\mathcal{B}$ proving (4.8). □

Corollary 4.1. *If $c > -\infty$, (4.8) holds.*

Proof. Suppose $c = -p < 0$ and let $G_i(x) = p + G_i(x)$, $i = 0, 1$. Then apply Theorem 4.1. □

We have thus shown that the order of inf and max can be exchanged in problem (4.3) when c is finite and that the max-inf is $-\infty$ when c is $-\infty$. This shows that we can exchange the order of max and inf in every case.

Lemma 4.4. *The set $\mathcal{B}$ consisting of those β for which $\beta G_0(x) + (1-\beta)G_1(x)$ is convex in x is a closed interval inside $[0, 1]$.*

Proof. It is easy to see that $\beta \in \mathcal{B}$ if and only if $M(\beta) \triangleq \beta P_0 + (1-\beta)P_1 \geq 0$. Clearly then $\mathcal{B}$ is closed. Now let $0 \leq \lambda \leq 1$, and let β_1 and β_2 be in $\mathcal{B}$. We can see that $\lambda\beta_1 + (1-\lambda)\beta_2$ is in $\mathcal{B}$ by noting that $M(\lambda\beta_1 + (1-\lambda)\beta_2) = \lambda M(\beta_1) + (1-\lambda)M(\beta_2) \geq 0$. Thus $\mathcal{B}$ is connex. □

Remark The solution to the optimization problem (4.1) can thus be constructed as follows

$$c = \max_{\beta\in\mathcal{B}} \inf_x \beta G_0(x) + (1-\beta)G_1(x). \tag{4.12}$$

The minimization problem is a standard LQ problem, and the maximization is that of a concave function over a finite interval. Note also that $\mathcal{B}$ being not empty does not imply that c is not $-\infty$. For that we need an additional assumption: non-singularity.

Definition 4.1. *The optimization problem (4.1) is called non-singular if there exists a $\beta \in [0, 1]$ such that $\beta G_0(x) + (1-\beta)G_1(x)$ is strictly convex.*

Clearly in the finite dimensional case, non-singularity is equivalent to the existence of $\beta \in [0, 1]$ such that

$$\beta P_0 + (1-\beta)P_1 > 0. \tag{4.13}$$

Lemma 4.5. *If the optimization problem (4.1) is non-singular, then $c > -\infty$.*

Proof. Let β such that (4.13) is satisfied. Then the optimization problem:

$$\inf_x \beta G_0(x) + (1-\beta)G_1(x) \tag{4.14}$$

has a finite solution because we have a strictly convex quadratic functional. Thus $c > -\infty$. □

Note that non-singularity is not a necessary condition for $c > -\infty$, but sufficient. However, it is a necessary and sufficient condition for $c > -\infty$ for all q_i and r_i. Also note that in the non-singular case, $\mathcal{B}$ is not reduced to a single point.

4.2 Continuous-time problem and the Riccati equation

Lemma 4.6. *Suppose the Hamiltonian two-point boundary value system*

$$\dot{\xi} = \begin{pmatrix} A & Q \\ R & -A^T \end{pmatrix} \xi \tag{4.15}$$

$$0 = V_0\xi(0) + V_T\xi(T) \tag{4.16}$$

where $V_0 = \begin{pmatrix} \Pi_i & -I \\ 0 & 0 \end{pmatrix}$, $V_T = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}$, *is well-posed (i.e. has unique zero solution) for all* $T \in [0, \bar{T})$ *but it is not well-posed for* $T = \bar{T}$. *Then the solution of the Riccati equation*

$$\dot{P} = AP + PA^T - PRP + Q, \;\; P(0) = \Pi_i^{-1}, \tag{4.17}$$

diverges at $t = \bar{T}$.

Proof. Let Ψ be the system matrix, i.e, $\Psi(0) = I$ and

$$\frac{d}{dt}\begin{pmatrix} \Psi_1 & \Psi_2 \\ \Psi_3 & \Psi_4 \end{pmatrix} = \begin{pmatrix} A & Q \\ R & -A^T \end{pmatrix}\begin{pmatrix} \Psi_1 & \Psi_2 \\ \Psi_3 & \Psi_4 \end{pmatrix}. \tag{4.18}$$

The well-posedness of the two-point boundary value system over $[0, T]$ can then be characterized in terms of the invertibility of the matrix $V_0 + V_T\Psi(T) = \begin{pmatrix} \Pi_i & -I \\ \Psi_3(T) & \Psi_4(T) \end{pmatrix}$ which after straightforward manipulation can be shown to be equivalent to invertibility of $\Psi_3(T) + \Psi_4(T)\Pi_i$. Let $M(t) = \Psi_3(t) + \Psi_4(t)\Pi_i$. Then we know that $M(t)$ is invertible for $t < \bar{T}$ and singular for $t = \bar{T}$. But it is straightforward to verify that $P(t) = (\Psi_1(t) + \Psi_2(t)\Pi_i)M(t)^{-1}$ satisfies the Riccati equation (4.17), and it diverges at $t = \bar{T}$ because $M(\bar{T})$ is singular. □

Now consider the following optimization problem

$$J(s, a, b) = \min_{x(0),\nu} x(0)^T P_0^{-1} x(0) + \int_0^s \nu^T \Gamma \nu \; dt \tag{4.19}$$

subject to

$$\dot{x} = Ax + B\nu + a \tag{4.20}$$

$$b = Cx + D\nu \tag{4.21}$$

over the interval $[0,T]$. Matrices A, B, C and D have piece-wise continuous-time entries. D is surjective, $P_0 > 0$ and Q is symmetric and invertible but not necessarily sign-definite. Vectors a and b are known continuous functions of time.

Theorem 4.2. *Optimization problem (4.19) has a bounded unique solution for all $s \in (0,T]$ if and only if*

$$D_\perp{}^T \Gamma D_\perp > 0, \quad \forall t \in [0,T] \tag{4.22}$$

where $D_\perp$ is a highest rank matrix such that $DD_\perp = 0$ and the Riccati equation

$$\dot{P} = (A - SR^{-1}C)P + P(A - SR^{-1}C)^T - PC^TR^{-1}CP + Q - SR^{-1}S^T, \quad P(0) = P_0 \tag{4.23}$$

where $\begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} = \begin{pmatrix} B \\ D \end{pmatrix} \Gamma^{-1} \begin{pmatrix} B \\ D \end{pmatrix}^T$ *has a solution on [0,T].*

Corollary 4.2. *If it exists, the unique bounded solution to the optimization problem (4.19) is given by $J(s,a,b) = \int_0^s \mu^T R^{-1} \mu \, dt$ where $\mu = C\hat{x} - b$ and where $\hat{x}$ satisfies $\dot{\hat{x}} = A\hat{x} - (S + PC^T)R^{-1}\mu + a$, $\hat{x}(0) = x_0$ where P is the solution of (4.23).*

Proof. Space prohibits giving anything but an outline of the proof of Theorem 4.2. Recall that in an optimal control problem defined on an interval $[t_0, t_f]$ that the cost to go, $W(x,t)$ is the minimal cost given we start at position x at time t. The principle of optimality from dynamic programming says that the cost to go function satisfies the equation

$$0 = \min_\nu H(x, \nu, W_x) \tag{4.24}$$

where H is the Hamiltonian. The principle (4.24) in turn leads to a partial differential equation on W. Now note that if we change the time direction in (4.19)–(4.21) by letting $t = T - \tau$, then we get a standard terminal cost optimal control problem. Accordingly we shall use the principle (4.24) in reverse. Thus cost (4.19) is the past cost given $x(t), \nu(t)$. Let $V(s,z)$ for a vector z, be the minimum value of (4.19) with the added constraint that $x(s) = z$. We then determine $\hat{x}, \Pi(t), \mu(t)$ such that

$$V(s,x) = (x - \hat{x})^T \Pi(0)(x - \hat{x}) + \int_0^s \mu^T R^{-1} \mu dt \tag{4.25}$$

Note that the existence condition we consider in Theorem 4.2 has to do with a family of optimization problems and in particular optimization problems (4.19) for all $s \in (0,T]$. It turns out the existence condition for the single optimization problem (4.19), $s = T$, is not the same in general. □

5 Conclusion

We have developed some of the theory and algorithms for the design of auxiliary signals for performing model identification in the presence of both additive and model uncertainty. The perturbations satisfy certain bounds and guaranteed indentification is provided by the auxiliary signals.

Acknowledgement: Research supported in part by the National Science Foundation under DMS-0101802, ECS-0114095 and INT-9605114.

References

[1] C. Bunks, J. P. Chancelier, F. Delebecque, C. Gomez (Editor), M. Goursat, R. Nikoukhah, and S. Steer, *Engineering and scientific computing with Scilab*, Birkhauser, 1999.

[2] S. L. Campbell, K. Horton, R. Nikoukhah, and F. Delebecque, "Auxiliary signal design for rapid multimodel identification using optimization," *Automatica*, to appear.

[3] F. Kerestecioğlu, *Change Detection and Input Design in Dynamical Systems*, Research Studies Press, Taunton, U.K. 1993.

[4] F. Kerestecioğlu and M. B. Zarrop, "Input design for detection of abrupt changes in dynamical systems," *Int. J. Control* 59, 1993 1063–1084.

[5] R. Nikoukhah, "Guaranteed active failure detection and isolation for linear dynamical systems," *Automatica* 34, 1998, 1345–1358.

[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* 14, 2000, 701–724.

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

[8] I.R. Petersen and A.V. Savkin, *Robust Kalman filtering for signals and systems with large uncertainties*, Birkhauser, 1999.

[9] I. R. Petersen, V. A. Ugrinovskii, and A. V. Savkin, *Robust Control Design Using H^∞ Methods*, Springer, Communications and Control Engineering, 2000.

[10] R. Nikoukhah, S. L. Campbell, and F. Delebecque, "Auxiliary signal design for failure detection in uncertain systems," *Proc. IEEE Med. Conf. Control and Automation*, 2001.

[11] R. Nikoukhah, and S. L. Campbell, "Active failure detection: Auxiliary signal design and on-line detection," *Proc. IEEE Med. Conf. Control & Automation*, 2002.

[12] X. J. Zhang, *Auxiliary Signal Design in Fault Detection and Diagnosis*, Springer-Verlag, Heidelberg, 1993.