Model (In)vaildation and Fault Detection for Systems with Polynomial State-Space Models

Farshad Harirchi, Zheng Luo, Necmiye Ozay

Abstract—This paper addresses the problem of (in)validation of polynomial state-space models, that is, checking whether a discrete-time uncertain polynomial state-space model can explain noisy experimental input/output data. We first recast this problem as a polynomial optimization problem and present asymptotically tight invalidation certificates by appealing to well-known moments-based relaxations. In the second part of the paper, we show how a model-based run-time fault detection algorithm can be developed based on a notion of $T$-detectability, which enables the proposed model invalidation approach to be applied in receding horizon fashion to detect faults. The efficacy of the proposed methods are illustrated with some numerical and practical examples.

I. INTRODUCTION

In recent years, there is a growing interest in safety, reliability and security of Cyber-Physical Systems (CPS). CPS are integrations of networks and embedded computers with physical processes, which are controlled with feedback loops. Our every-day life depends on the reliability of cyber-physical systems such as traffic control systems, advanced automotive systems, environmental control systems, and smart structures. Therefore, detecting anomalies and faults in such systems in real-time is an important challenge. Many physical systems can be represented or approximated by polynomial state-space models. As a result, these models play an important role in the analysis of cyber-physical systems. In this work, we initially investigate the model invalidation problem for polynomial state-space models, and consequently, propose a theoretical framework that allows us to utilize a model invalidation approach to detect faults and anomalies in real-time.

Our contributions in this paper are two-fold. First, we address the model invalidation problem for polynomial models, by recasting it as a polynomial optimization problem and leveraging moment-based relaxation techniques, which take into account the sparse structure of the problem. This is similar in spirit to the techniques used in [7], [10]. Second, we adapt the concept of $T$-detectability of a fault model to polynomial models, and propose a computational framework to calculate $T$, if it is finite. This framework is also based on polynomial optimization that is relaxed to a semidefinite programming problem, which provides certificates for $T$-detectability. Moreover, the relaxations used for invalidation are shown to be consistent, in a sense made precise later in the paper, with the $T$-detectability certificates, therefore the detection guarantees are preserved even when one solves the relaxed invalidation problem online using a sliding window.

The paper structure is as follows. In Section II, some preliminaries and notations are described. Section III explains the model invalidation approach proposed for polynomial models. The relation between model invalidation and fault detection as well as the concept of $T$-detectability are highlighted in Section IV. We represent the effect of uncertainty in Section V, and illustrative examples are provided in Section VI. Finally, conclusions are made in Section VII.
In this section, we introduce some notation as well as some basic concepts that are essential for the proposed algorithm.

A. Notation

$x \in \mathbb{R}^n$ denotes a vector and $M \in \mathbb{R}^{n \times m}$ represents a matrix, where $M_{i,j}$ indicates the element on $i$th row and $j$th column of the matrix $M$. The infinity norm of a vector $x$ is denoted by $\|x\|_{\infty}$. $M \succeq N$ indicates that the matrix $M - N$ is positive semi-definite. $p(x)$ denotes a polynomial in $x$.

B. Polynomial optimization and moments-based relaxations

The approach proposed in this paper is built on some results from polynomial optimization and problem of moments literature [19]-[21]. Consider the polynomial optimization problem:

$$p_K := \min_{x \in K} p(x)$$  \hspace{1cm} (P1)

over a compact semialgebraic set, $K \subset \mathbb{R}^n$, defined by $c$ polynomial inequalities. That is, $K := \{x \in \mathbb{R}^n \mid q_k(x) \geq 0, k = 1, \ldots, c\}$. Problem (P1) is typically non-convex and NP-hard. Alternatively, consider the following infinite dimensional convex problem:

$$\tilde{p}_K := \min_{\mu \in \mathcal{P}(K)} \mathbb{E}_\mu[p(x)]$$  \hspace{1cm} (P2)

where $\mathcal{P}(K)$ is the set of positive Borel measures on the set $K$ with $\mu(K) = 1$. In [19], it is shown that problems (P1) and (P2) have the same optimal value. Moreover, an alternative (yet still infinite dimensional) characterization of the feasible set of the functional optimization problem (P2) exists in terms of positive semi-definiteness of the so-called moment $M$ and localization $L(q_k, m)$ matrices constructed from the moments $m$ of distributions supported on $K$.

The truncated moment and localization matrices associated with the moments $M$ and localization $L(q_k, m)$ are defined as follows:

$$M^i_j(m) = m_{\alpha^i + \alpha^j}$$ for all $i, j \leq S_r$$  \hspace{1cm} (1)

$$L^i_j(q_k m) = \sum_{\beta} q_k \beta_{(i)} m_{\beta_{(i)} + \alpha^i + \alpha^j}$$

for all $i, j \leq S_r - \left\lfloor \frac{n}{2} \right\rfloor$$  \hspace{1cm} (2)

where $S_r = \left( \frac{r + n}{2} \right)$ is the number of monomials in $\mathbb{R}^n$ up to order $r$, $q_k \beta_{(i)}$ denotes the coefficient of monomial $x^\beta$ in polynomial $q_k(x)$ with degree $d_k$. The main result of [19] is the following.

Theorem 1: Let

$$p_r^* = \min_m \sum_{\alpha} p_{\alpha} m_{\alpha}$$  \hspace{1cm} (3)

$$s.t. M_r(m) \succeq 0$$

$$L_r(q_k m) \succeq 0, k = 1, \ldots, c.$$  \hspace{1cm} (4)

As $r$ goes to infinity, $p_r^*$ approaches $p_r^*$ from below.

C. Running Intersection Property and Sparse Structure

In general, even the truncated matrices in (1)-(2) can have a large size if the original polynomial optimization problem involves too many variables. On the other hand, if the polynomials involved have a sparse structure satisfying the so-called running intersection property, this structure can be utilized to obtain multiple smaller sized linear matrix inequality constraints, where most of the moment variables corresponding to cross moments are eliminated [20].

Definition 1: Consider problem (P1), and a finite collection $\{I_i\}_{i=0}^l$, where each $I_i$ is a subset of variables in $x$, with $\bigcup_{i=0}^l I_i = \{x_1, x_2, \ldots, x_n\}$. Assume that the objective function can be partitioned as $p(x) = \sum_{j=1}^l p_j(x)$, where each $p_j(x)$ depends only on the variables in $I_i$ for some $i$ (of $k$). Then, the problem (P1) is said to satisfy the running intersection property if the collection $\{I_i\}_{i=0}^l$ satisfies

$$I_{i+1} \cap \bigcup_{j=0}^i I_j \subseteq I_r \text{ for some } r \leq i.$$  \hspace{1cm} (4)

If the running intersection property is satisfied, the following hierarchy of semidefinite programs can be constructed:

$$p_r^* = \min_m \sum_{j=1}^l \sum_{\alpha(j)} p_j \alpha(j) m_{\alpha(j)}$$  \hspace{1cm} (5)

$$s.t. M_r(m_{I_{i(k)}}) \succeq 0, \text{ } i = 0, \ldots, l'$$

$$L_r(q_k m_{I_{i(k)}}) \succeq 0, k = 1, \ldots, c,$$

where $p_j \alpha(j)$ denotes the coefficient of the $\alpha(j)$th monomial in the polynomial $p_j$, and $M_r(m_{I_{i(k)}})$ and $L_r(m_{I_{i(k)}})$ indicate the moment and localization matrices associated with variables in $I_i$, and $I_{i(k)}$, respectively. Similar to Theorem 1, we have the following.

Theorem 2: Consider problems (P1) and (5). As $r$ goes to infinity, $p_r^*$ approaches $p_r^*$ from below.

For more details on moments-based polynomial optimization and the relation to the running intersection property, see [20], [21].

III. MODEL INVALIDATION

A. Problem Definition

In this section, the problem of model invalidation for discrete-time polynomial state-space models is described. We consider polynomial models of form:

$$G(X, U, E, f)$$  \hspace{1cm} (6)
where $\mathcal{X} \subset \mathbb{R}^n$ is the set of states, $\mathcal{U} \subset \mathbb{R}^{n_u}$ is the set of admissible inputs, $\mathcal{E}$ captures both $\mathcal{E}_p \subset \mathbb{R}^n$ and $\mathcal{E}_m \subset \mathbb{R}^{n_x}$, the set of possible process and measurement noise, and finally, $f$ is the polynomial defining the state equations as follows:

$$x(k + 1) = f(x(k), u(k)) + E_p \eta_p(k)$$

$$y(k) = x(k) + E_m \eta_m(k)$$

where $x$, $u$ and $y$ denote the state, input and output vectors, respectively. The process and measurement noise vectors are represented by $\eta_p$ and $\eta_m$, and $E_p$ and $E_m$ denote how the noise vectors affect the states and outputs.

For simplicity we consider all the sets $\mathcal{X}, \mathcal{U}, \mathcal{E}_p$ and $\mathcal{E}_m$ to be infinity norm balls, with sizes $M, U, \epsilon_p$ and $\epsilon_m$, respectively. By the definition above, we assume that the noisy states are being observed as outputs. The results of this work can be easily extended to the case in which the output is a polynomial function of the states. In order to state the model invalidation problem, we first define the behavior of the polynomial model.

**Definition 2:** The length-$N$ behavior associated with a polynomial system $G$ is the set of all input-output trajectories in the interval $[0, N]$ compatible with $G$, given by the set

$$B^N_{\text{poly}}(G) := \{ (u(k), y(k))_{k=0}^N | u(k) \in \mathcal{U} \text{ and } \exists x(k) \in \mathcal{X}, \eta_p(k) \in \mathcal{E}_p, \eta_m(k) \in \mathcal{E}_m \text{ s.t. } (7), (8) \text{ holds for } k = 0, \ldots, N \}. $$

We call $B^N_{\text{poly}}(G)$ just the behavior of the system $G$ throughout this paper.

Now we can state the model invalidation problem for polynomial systems. In words, given an input-output data sequence and a polynomial model, model invalidation problem is to determine whether or not the data is compatible with the model. This can be formally stated in terms of behaviors as follows:

**Problem 1:** Given $\{(u(k), y(k))\}_{k=0}^N$ an input-output sequence, and a polynomial model $G$, determine whether or not the input-output sequence is contained in the behavior of $G$. That is, whether or not the following is true

$$\{(u(k), y(k))\}_{k=0}^N \in B^N_{\text{poly}}(G).$$

**B. Approach**

An optimization-based approach is proposed in this work to tackle the model invalidation problem. Let us first define

$$\phi_k(y, u, \eta_m, \eta_p) := f(y(k) - E_m \eta_m(k), u(k)) + E_p \eta_p(k) - y(k + 1) + E_m \eta_m(k + 1), \forall k \in [0, N - 1],$$

by combining the state and output equations. Note that $\phi_k(y, u, \eta_m, \eta_p)$ is a vector of functions of the size of states. Consider the following optimization problem:

$$\epsilon^* = \min_{\eta_m, \eta_p} \epsilon$$

s. t. $\phi_k(y, u, \eta_m, \eta_p) = 0, \forall k \in [0, N - 1]$

$$\|y(k) - \eta_m(k)\|_{\infty} \leq M, \forall k \in [0, N]$$

$$\|u(k)\|_{\infty} \leq U, \forall k \in [0, N]$$

$$\|\eta_p(k)\|_{\infty} \leq \epsilon_p, \|\eta_m(k)\|_{\infty} \leq \epsilon, \forall k \in [0, N].$$

Then, we can state the following proposition:

**Proposition 1:** Given the input-output sequence, $\{(u(k), y(k))\}_{k=0}^N$, and the polynomial model $G$, the model is invalidated by input-output sequence, if $\epsilon^* > \epsilon_m$, and vice versa if $\epsilon^* \leq \epsilon_m$.

**Proof:** Direct consequence of definitions.

The choice of measurement noise bound to be a variable is arbitrary. One can as well choose the process noise bound, input or state bounds as an optimization variable. Even though the problem can also be stated as a polynomial feasibility problem, the optimization form gives a quantitative notion of invalidation. That is, by comparing the optimal objective value and the a priori bounds, it is possible to assess how “far” the model is from being valid or invalid. Similar optimization problems can also be used to derive bounds on uncertain parameters if such bounds are not known a priori.

Problem (P$_{\text{Ms}}$) is not necessarily a convex problem, because of the constraints defined by $\phi$ functions, which can be any arbitrary polynomial. We leverage moment-based relaxation techniques to solve this problem but first, we show that (P$_{\text{Ms}}$) satisfies the running intersection property.

**Proposition 2:** Problem (P$_{\text{Ms}}$) satisfies running intersection property.

**Proof:** Define $I_k = \{ \epsilon, \eta_m(k), \eta_m(k+1), \eta_p(k) \}, \forall k \in [0, N - 1]$, and $I_N = \{ \epsilon, \eta_m(k) \}$. Then at each time, $k \in [0, N]$, the only variables which appear in the objective function and constraints are listed in $I_k$. Since,

$$\{ \epsilon, \eta_m(k+1) \} \subseteq I_k \forall k \in [0, N - 1],$$

condition (4) is satisfied and therefore running intersection property holds.

Let $P'_{\text{Ms},\text{rel}}$ be the moments-based relaxation of problem (P$_{\text{Ms}}$) of the form (5) with relaxation order $r$.

**Proposition 3:** Given input-output data sequence, let $\epsilon_r$ denote the optimal value of $P'_{\text{Ms},\text{rel}}$. If $\epsilon_r > \epsilon_m$, then the model is invalid.

**Proof:** From Theorem 2, the optimal value of $P'_{\text{Ms},\text{rel}}$ converges to $\epsilon^*$ from below, so we have $\epsilon_r \leq \epsilon^*$. Therefore, $\epsilon_m < \epsilon^*$.

**IV. FAULT DETECTION**

In this section, the application of the proposed model invalidation scheme in fault and anomaly detection is presented. Let us first recall the definition of anomaly and fault from [18].

**Definition 3:** An input-output sequence $\{(u(k), y(k))\}_{k=0}^N$ is called abnormal for a polynomial model $G$ if and only if $\{(u(k), y(k))\}_{k=0}^N \notin B^N_{\text{poly}}(G)$. 


**Definition 4:** A fault model for a polynomial system $G = (\mathcal{X}, \mathcal{U}, \mathcal{E}, f)$ is another polynomial system $G_f = (\mathcal{X}_f, \mathcal{U}_f, \mathcal{E}_f, f_f)$ with the same number of states, inputs and outputs.

As seen from Def. 3, if the a priori model $G$ captures the behavior of the nominal system, the data being invalid is equivalent to an abnormal behavior occurring in the system. This allows us to apply the proposed model invalidation to detect generic abnormalities that might happen in the system, which can include faults, attacks or failures. However, the size of $(P_{M})$ increases by time-horizon. The next question is that if we have some information about the fault such as a model as defined by Def. 4, can we detect it more efficiently? In the next step, we answer this question, but first we make the following assumption on faults.

**Assumption 1:** We assume that all the faults are persistent, that is once they occur, the system stays evolving according to them.

**Definition 5:** A fault model $G_f$ for a polynomial system $G$ is called $T$-step detectable if $\mathcal{B}_{\text{poly}}^T(G) \cap \mathcal{B}_{\text{poly}}^N(G_f) = \emptyset$ for all $N \geq T$, where $T$ is a positive integer.

It is clear from the definition that if a fault model is $T$-step detectable, it is also $T$'-step detectable for all $T' \geq T$.

**Proposition 4:** Given a $T$-step detectable fault model $G_f$ for a polynomial model $G$, under Assumption 1, the existence of the fault can be verified by checking, at each time $k^+$, if the solution of $(P_{M})$ satisfies $\epsilon^T \geq \epsilon_m$ for input-output sequence $\{u(k), y(k)\}_{k=0}^{\epsilon_m}$ or not.

**Proof:** The proof is similar to that of Proposition 3 in [18].

If an explicit fault model exists and it is $T$-detectable for some finite $T$, Proposition 4 enables us to solve an optimization problem of size $T$ at each time step in order to detect that particular fault. Next, we propose an optimization-based approach to verify whether for a given $T$, a fault is $T$-detectable for a system or not. Consider the following optimization problem:

$$
\hat{\epsilon} = \min_{\eta, u, y, \epsilon} \epsilon \quad (P_T)
$$

s.t. $f(y(k) - E_m \eta_m(k), u(k)) + C_p \eta_p(k) - y(k + 1) + E_m \eta_m(k + 1) = 0, \forall k \in [0, T - 1]$

$f(y(k) - E_m \eta_m(k), u(k)) + C_p \eta_p(k) - y(k + 1) + E_m \eta_m(k + 1) = 0, \forall k \in [0, T - 1]$

$\|u(k)\|_\infty \leq \min(U, U_f), \forall k \in [0, T]$

$\|y(k) - E_m \eta_m(k)\|_\infty \leq M, \forall k \in [0, T]$

$\|y(k) - E_m \eta_m(k)\|_\infty \leq M, \forall k \in [0, T]$

$\|\eta_p(k)\|_\infty \leq \epsilon_p, \|\eta_m(k)\|_\infty \leq \epsilon, \forall k \in [0, T]$

$\|\eta_p(k)\|_\infty \leq \epsilon_p, \|\eta_m(k)\|_\infty \leq \epsilon_m, \forall k \in [0, T]$,

where $\eta$ represents both process and measurement noise variables. Then, we can state the following proposition:

**Proposition 5:** Given the a priori and fault models, $G$ and $G_f$, the fault model is $T$-detectable for $G$ if and only if $\hat{\epsilon} \geq \epsilon_m$.

**Proof:** Direct consequence of definitions.

Similar to problem $(P_{M})$, it can be shown that optimization problem $(P_T)$ also satisfies the running intersection property. Therefore, $T$-detectability certificates can be obtained using a relaxation of the form (5). Let $P_{T, \text{rel}}$ be the moments-based relaxation of $(P_T)$ of the form (5) with relaxation order $r$, and denote the optimal value of $P_{T, \text{rel}}$ by $\epsilon_r$.

**Proposition 6:** Given a model $G$, fault model $G_f$ and $T$, let $\epsilon_r$ denote the optimal value of $P_{T, \text{rel}}$. If $\epsilon_r > \epsilon_m$, then $G_f$ is $T$-detectable.

**Proof:** Follows from Theorem 2.

For a given $T$, if the solution of $P_{T, \text{rel}}$ results in an objective value greater than $\epsilon_m$, the fault is $T$-detectable. That is, if one can solve the invalidation problem $(P_{M})$ exactly, then, by Propositions 4 and 5, it is enough to use only data from a window of size $T$ to declare the existence of the fault. One question that remains is that whether this is still true if one solves a relaxed version $P_{M, \text{rel}}$ at run time. This is established next.

**Theorem 3:** If a fault model $G_f$ is shown to be $T$-detectable using $P_{T, \text{rel}}$ and if this fault occurs at time $t^*$, then using data $\{u(k), y(k)\}_{k=t^*-T}^{t^*}$, $P_{M, \text{rel}}$ will have an objective value greater than $\epsilon_m$, that is, under Assumption 1, the fault is guaranteed to be detected using the relaxed problem $P_{M, \text{rel}}$ in receding horizon manner.

**Proof:** Let $P_{M, \text{rel}}(\{u(k), y(k)\}_{k=t^*-T}^{t^*})$ be the moments-based relaxation of problem $(P_{M})$ of the form (5) with relaxation order $r$ parametrized by data $\{u(k), y(k)\}_{k=t^*-T}^{t^*}$, and let $\epsilon_r(\{u(k), y(k)\}_{k=t^*-T}^{t^*})$ be its objective value. Let

$$
\epsilon^* := \min_{\{u(k), y(k)\}_{k=t^*-T}^{t^*}} \epsilon_r(\{u(k), y(k)\}_{k=t^*-T}^{t^*}) \quad (11)
$$

s.t. $\{u(k), y(k)\}_{k=t^*-T}^{t^*} \in \mathcal{B}_{\text{poly}}^T(G_f)$.

We need to show $\epsilon^* > \epsilon_m$. Problem (11) can be rewritten as:

$$
\epsilon^* := \min_{\eta, u, y} \epsilon_r(\{u(k), y(k)\}_{k=t^*-T}^{t^*}) \quad (12)
$$

s.t. $f(y(k) - E_m \eta_m(k), u(k)) + C_p \eta_p(k) - y(k + 1) + E_m \eta_m(k + 1) = 0, \forall k \in [t, t + T - 1]$

$\|u(k)\|_\infty \leq \min(U, U_f), \forall k \in [t, t + T]$

$\|y(k) - E_m \eta_m(k)\|_\infty \leq M, \forall k \in [t, t + T]$

$\|\eta_p(k)\|_\infty \leq \epsilon_p, \|\eta_m(k)\|_\infty \leq \epsilon, \forall k \in [t, t + T]$

$\|\eta_p(k)\|_\infty \leq \epsilon_p, \|\eta_m(k)\|_\infty \leq \epsilon_m, \forall k \in [t, t + T]$

A relaxation $P'$ of (12) can be obtained by replacing the monomials in $\eta, u, y$ with moment variables of order up to $r$. The objective value $\epsilon'$ of $P'$ satisfies $\epsilon' \leq \epsilon^*$. Note that $P'$ has bilinear terms containing moments of $\eta, u, y$ multiplying those of $\eta$ inside $\epsilon_r(\cdot)$. This corresponds to an independence assumption between the distributions these moments belong to. Adding redundant equalities in $P'$ and taking Kronecker products with appropriately constructed moment matrices, result in a new problem $P''$ with the same objective value $\epsilon'$. Finally, relaxing the independence assumption and replacing the bilinear terms with cross-moment variables lead to $P_{T, \text{rel}}'$. Since $P_{T, \text{rel}}$ is obtained as a relaxation of $P''$, we have $\epsilon_r \leq \epsilon'$. Moreover, since $T$-detectability certificate was obtained.
via $P_{m}^c$ (Proposition 6), we have $\epsilon_r > \epsilon_m$. Therefore, $\epsilon^* > \epsilon_m$.

Remark 1: Note that in general, if multiple faults can occur, running model invalidation in receding horizon fashion, we can only conclude the occurrence of a fault but might not be able to differentiate them. This is related to the concepts of fault identifiability and isolability [22]. This can potentially be done by checking detectability across different fault models, but left for future work.

V. UNCERTAINTY

In this section, we consider uncertainty in the parameters of the polynomial models and propose necessary changes to the model invalidation framework ($P_{m}^c$) so that it can handle uncertain models. Let us first define what we mean by the uncertainty in polynomial models. We can write any polynomial a product of two vectors as follows:

$$ f(x, u) = c^T q, $$  \hspace{1cm} (13)

where $q$ is the vector of monomials in $x$ and $u$; and $c$ is the vector of corresponding coefficients.

Definition 6: The polynomial model, $G(X, U, E, \tilde{f})$ has uncertainty if $\tilde{f}(x, u, \Delta c)$ has the following form:

$$ \tilde{f}(x, u, \Delta c) = (c + \Delta c)^T q, $$  \hspace{1cm} (14)

where $\Delta c$ is the uncertainty vector, $[T_c, \Delta c] < 1\bar{c}$, $T_c$ is a diagonal matrix that scales the bound on different elements of uncertainty vector, $1$ is a vector of ones of the appropriate size, and $\bar{c}$ is a constant.

For simplicity of notation, let us define $\tilde{\phi}_k$ for all $k$ as follows:

$$ \tilde{\phi}_k(y, u, \eta_m, \eta_p, \Delta c) := \tilde{f}(y(k) - E_m \eta_m(k), u(k), \Delta c) + E_p \eta_p(k) - y(k + 1) + E_m \eta_m(k + 1) \forall k \in [0, T]. $$  \hspace{1cm} (15)

With this notation, if the data matches the system equations at time $k$, $\tilde{\phi}_k(y, u, \eta_m, \eta_p, \Delta c) = 0$. In order to address the model invalidation problem, one needs to take into account the effect of the uncertainty vector. Consider the following problem:

$$ \sigma^* = \min_{\eta_m, \eta_p, \Delta c, \sigma} \sigma \quad \text{(P)} $$

$$ \text{s. t. } \tilde{\phi}_k(y, u, \eta_m, \eta_p, \Delta c) = 0, \forall k \in [0, N] $$
$$ \|y(k) - \eta_m(k)\|_\infty \leq M, \forall k \in [0, N] $$
$$ \|T_c\Delta c\|_\infty \leq \sigma, \|u(k)\|_\infty \leq U, \forall k \in [0, N] $$
$$ \|\eta_p(k)\|_\infty \leq \epsilon_p, \forall k \in [0, N] $$
$$ \|\eta_m(k)\|_\infty \leq \epsilon_m, \forall k \in [0, N]. $$

The model is invalidated by data, if $\sigma^* > \bar{\sigma}$ and vice versa, if $\sigma^* \leq \bar{\sigma}$. Proceeding along the same lines as in Section IV, an optimization problem to certify $T$-detectability of an uncertain system can be defined.

VI. ILLUSTRATIVE EXAMPLES

In this section, we consider one set of numerical examples and an application-motivated example to illustrate the efficacy of the methods proposed in this paper. All the examples are implemented on a 3.5 GHz machine with 32 GB of memory running Ubuntu. For the implementation of model invalidation approach and finding $T$ for $T$-detectability, we used Yalmip [23] and SparsePOP [24]. All the approaches and examples are implemented in Matlab, and are available as part of MI4Hybrid toolbox.

A. Numerical Examples

Consider a discrete-time pendulum model with friction:

$$ x_1(k + 1) = x_1(k) + 0.3x_2(k) $$
$$ x_2(k + 1) = 0.9x_2(k) - 0.3\sin(x_1(k)). $$  \hspace{1cm} (16)

First, we replace the $\sin(x_1)$ with its fourth order Taylor series expansion around zero. The outputs of the system are noisy measurements of the states. The measurement noise bound is assumed to be 0.1. In addition, we assume both states are bounded in the range $[-2, 2]$.

1) Model Invalidation Example: In this example, the tightness of model invalidation is illustrated by slightly changing the measurement noise bound. The output sequence is generated from (16), while the measurement noise is generated uniformly from the set $\alpha[-0.1, 0.1]$. The model invalidation algorithm is applied to input-output sequence of size 50, generated from (16) with different $\alpha$ values. The run is repeated for 20 times for each $\alpha$ and the number of invalidations are illustrated in Fig. 1. In the second set of examples, we fix $\alpha$ at 1, and change the parameter $\beta$, which is multiplied with Taylor series expansion of the right-hand side of (16). The run is repeated for 20 times for noise that is randomly generated with bound 0.1. Fig. 1 illustrates that the model invalidation approach is tight, that is a slight mismatch in the parameters or noise bounds can be detected even with low relaxation orders. Note also that when the noise bound ratio is less than or equal to 1, the data is generated by a valid model. Even though the data sequences are generated by a slightly different model than the a priori model, they can still be valid behaviors of the original system, which can be the case for the trials that are not invalidated in Fig. 1.

2https://github.com/data-dynamics/MI4Hybrid

Fig. 1. The invalidation results over 20 trials (left) when there is a mismatch between the parameters of the model that generated the data is $\beta$ times those of the a priori model, and (right) when the data is generated with noise, bounded by $\alpha$ times the a priori noise bound.
2) Computation Time Results: In this example, we investigate how the computation time scales as the data length increases. Two cases are considered: (i) when the input-output sequence is valid for model (16) and (ii) when it is invalid. The model invalidation algorithm is applied to both cases 10 times for each time horizon. Fig. 2 illustrates the mean and standard deviation of the run-time results obtained.

![Fig. 2](image_url)

Fig. 2. Model invalidation run-time increase almost linearly with the increase in time horizon for both valid and invalid data.

3) Quantitative Measure of the Validity of Data: Consider the discrete-time pendulum model in (16). The goal of this example is to demonstrate how we can practically use the discrete-time pendulum model in (16). The goal of this example is to illustrate the advantage of knowing a given fault is detectable, and showing how the invalidation approach proposed here can be employed for receding horizon fault detection. By applying the $T$-detectability approach, we discover that the fault model (18) is $T$-detectable for a priori model (17) with $T = 6$. Now, we gradually move a window of size 7 over all the samples, starting at time 7. We refer to the window at time $k$ as $W_k = [k-6,k]$. The model invalidation approach then is applied to each window. For case 1, the model is not invalidated until $W_{12}$, and all the instances after that

![Fig. 3](image_url)

Fig. 3. The optimal value of $P_{ML,rel}^4$ versus the order of the Taylor series expansion.

B. Moore-Greitzer Jet Engine Model

Jet engine compression systems are subject to control studies, which are aimed to analyze and avoid their instability against rotating stall and surge [25]. A polynomial state space model with uncertainty for the stabilized jet engine compression system with no stall is obtained in [26], and we discretized it using the Euler’s discretization method as follows:

$$\begin{align*}
\Phi(k+1) &= \Phi(k) - T_s (\Psi(k) + \frac{3}{2} \Phi^2(k) + \frac{1}{2} \Phi^3(k)) + \sigma \\
\Psi(k+1) &= \Psi(k) + 3T_s \Phi(k) - T_s \Psi(k)
\end{align*}$$  \hspace{1cm} (17)

where, $\Phi$ is the mass flow and $\Psi$ is the pressure rise. The measurements are the noisy states with noise vector $\eta_n$. Fault Description: We assume that the fault model for this system is caused by the bias in the mass flow rate that can be an influence of rotating stall in compression system. The rotating stall in this system shows itself as a region of severely reduced flow that rotates at a fraction of the rotor speed. The faulty model, then is described as follows:

$$\begin{align*}
\Phi(k+1) &= \Phi(k) - T_s (\Psi(k) + \frac{3}{2} \Phi^2(k) + \frac{1}{2} \Phi^3(k)) + \sigma + b_f \\
\Psi(k+1) &= \Psi(k) + 3T_s \Phi(k) - T_s \Psi(k) + b_f
\end{align*}$$  \hspace{1cm} (18)

where $b_f$ indicates the faulty bias. The output of faulty model is noisy measurements (with $\eta_m$) of the states.

The following assumptions are made on the model: the sampling time $T_s = 0.2$, the faulty bias $b_f = 0.1$, the maximum possible uncertainty $\sigma = 0.12$, and the infinity norm bound on the noise is equal to 0.05 for both $\eta_n$ and $\eta_m$. Fig. 4 illustrates the outputs of the two following cases with initial condition $[2 \ 3]^T$ for the two states:

1. The fault occurs at time sample 10.
2. The fault occurs at time sample 90.

![Fig. 4](image_url)

Fig. 4. The outputs of jet engine model when fault occurs at time sample 10 (top), and when it occurs at time sample 90 (bottom).
are invalidated. The same thing happens for the case 2, at \( W_{02} \). This indicates that the proposed approach is able to detect faults just two samples after it occurs. One can ask “why the \( T \)–detectability approach finds \( T = 6 \), but the faults can be detected in the windows of size 3?”. To answer this question we note that first, the relaxation order to find \( T = 6 \) is 3 in this example, so we might find smaller \( T \) with higher relaxation orders. Second, the \( T \)–detectability approach finds \( T \) such that no inputs, outputs, measurement and process noise can fall into the behavior of the system, however, in the model invalidation problem, input-output data sequence is fixed. In other words, the \( T \)–detectability approach checks the worst case scenario and finds \( T \) such that even the worst case cannot be explained by the model. Hence, even though we obtain a \( T \), often it is possible to detect the fault in a window of the size smaller than \( T \).

The mean time to run model invalidation approach on a window of size 7 is 2.4131 seconds. Suppose we do not know that the fault is 6–detectable and apply invalidation over the window of size 100 samples, then the run-time is 287.3850 seconds. Hence, the results on \( T \)–detectability of faults play an important role on being able to use the model invalidation approach for real-time fault detection.

VII. CONCLUSION

In this work, we proposed a tractable model invalidation approach for possibly uncertain polynomial state-space models, and discussed how it can be used for fault or anomaly detection. By recasting the model invalidation problem as a polynomial optimization problem, a quantitative notion of a detectability approach for polynomial state-space models is introduced. The algorithm proposed for obtaining a \( T \)–detectability certificate is shown to be consistent with the relaxation used for model invalidation. This is important mainly because it allows us to apply model invalidation on a window of size \( T \) to detect faults, which are \( T \)-detectable. This paves the way toward real-time fault detection. Finally, the performance of the proposed methods is verified on numerical and application-motivated examples.

As future work, we will investigate how to use \( T \)–detectability to guide sensor selection. We are also interested in extending some of the ideas presented in this paper to networked systems where sensory data is collected locally and partially shared between different nodes at runtime, therefore decentralized fault detection algorithms are required.

REFERENCES


