Model (In)validation 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, critical infrastructure control (electric power, water resources, and communications systems), defense systems, and smart structures [1], [2]. Therefore, detecting anomalies and faults in such systems in real-time is an important challenge [3], [4]. 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.

Originally, model invalidation was developed by robust control community as a tool to build trust in the models obtained from system identification or to improve those models [5]. More recently, model invalidation approaches have been developed for continuous-time polynomial systems [6], polynomial implicit difference equations [7] and switched auto-regressive models [8]–[10].

Fault detection has been an important field of research in different communities. From the control perspective, most of the fault detection approaches are based on residual generation and evaluation [11]. The residuals could be generated by parameter estimation techniques [12] or output and state observers [11], [13]. As an alternative to residual generation, set-membership fault detection methods have been proposed both for passive [14] and active [15]–[17] fault detection. Our approach falls into the category of set-membership passive fault detection approaches but instead of computing explicit set representations, we keep implicit constraints to represent sets. In our earlier work [18], we have considered invalidation of switched state-space models. Additionally, we have introduced the notion of T-detectability for a fault model that enables applying model invalidation in a receding horizon fashion to detect faults, if a fault model is known a priori. In this paper, we extend these results to polynomial state-space models. Although conceptually similar, the extension requires totally different computational techniques. The computational techniques used in [18] are satisfiability modulo theory (with linear arithmetic) and mixed integer linear programming, which are not applicable when the system dynamics are polynomial. In comparison to the earlier approaches for invalidation of polynomial models, we provide asymptotically tight conditions as opposed to the sufficient conditions in [6], and our approach incorporates all the data through a trajectory (crucial for extensions to run-time fault detection), whereas the approach in [6] uses only initial and final states of a trajectory.

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 invlidation 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.

This work is supported in part by DARPA grant N66001-14-1-4045.

The authors are with the Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI, 48109. {harirchi,zlou,necmiye}@umich.edu

II. PRELIMINARIES

In this section, we introduce some notation as well as some basic concepts that are essential for the proposed algorithm.

A. Notation

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

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:

$$p_K^* := \min_{\mathbf{x}\in K} p(\mathbf{x}) \tag{P1}$$

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

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

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 $\mathbf{M}$ and localization $\mathbf{L}(q_k\mathbf{m})$ matrices constructed from the moments $\mathbf{m}$ of distributions supported on K [1], and the objective function becomes a linear function of the moments. Although $\mathbf{M}$ and $\mathbf{L}$ are infinite dimensional matrices, they can be truncated to obtain a relaxation. In particular, the truncated moment and localization matrices up to order r are defined as follows:

$$\mathbf{M}_r^{i,j}(\mathbf{m}) = m_{\boldsymbol{\alpha}^{(i)}+\boldsymbol{\alpha}^{(j)}} \text{ for all } i,j \leq S_r \tag{1}$$

$$\mathbf{L}_r^{i,j}(q_k\mathbf{m}) = \sum_{\boldsymbol{\beta}} q_{k,\boldsymbol{\beta}^{(l)}} m_{\boldsymbol{\beta}^{(l)}+\boldsymbol{\alpha}^{(i)}+\boldsymbol{\alpha}^{(j)}} \text{ for all } i,j \leq S_{r-\left\lfloor \frac{d_k}{2}\right\rfloor} \tag{2}$$

where $S_r = \binom{r+n}{n}$ is the number of monomials in $\mathbb{R}^n$ up to order r, $q_{k,\boldsymbol{\beta}}$ denotes the coefficient of monomial $\mathbf{x}^{\boldsymbol{\beta}}$ in polynomial $q_k(\mathbf{x})$ with degree d_k. The main result of [19] is the following.

[1]Recall that for a given multi-sequence $\boldsymbol{\alpha} = [\alpha_1,\dots,\alpha_n]$ of nonnegative integers, the $\boldsymbol{\alpha}^{th}$ moment of a distribution μ supported on K is given by $m_{\boldsymbol{\alpha}} = \mathbb{E}_\mu(\mathbf{x}^{\boldsymbol{\alpha}}) := \int_K \mathbf{x}^{\boldsymbol{\alpha}}\mu(dx)$, where $\mathbf{x}^{\boldsymbol{\alpha}}$ is the monomial $x_1^{\alpha_1}\dots x_n^{\alpha_n}$.

Theorem 1: Let

$$\begin{aligned} p_r^* = \min_{\mathbf{m}} & \sum_{\boldsymbol{\alpha}} p_{\boldsymbol{\alpha}} m_{\boldsymbol{\alpha}} \\ \text{s.t. } & \mathbf{M}_r(\mathbf{m}) \succeq 0 \\ & \mathbf{L}_r(q_k\mathbf{m}) \succeq 0,\ k = 1,\dots,c. \end{aligned} \tag{3}$$

As r goes to infinity, p_r^* approaches p_K^* 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 has 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 $\mathbf{x}$, with $\bigcup_{i=0}^{l'} I_i = \{x_1, x_2, \dots, x_n\}$. Assume that the objective function can be partitioned as $p(\mathbf{x}) = \sum_{j=1}^{l} p_j(\mathbf{x})$, where each $p_j(\mathbf{x})$ depends only on the variables in I_i for some i (of j). In addition, each of the polynomials $q_k(\mathbf{x})$, defining K, depends also 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

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

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

$$\begin{aligned} \bar{p}_r^* = \min_{\mathbf{m}} & \sum_{j=1}^{l} \sum_{\boldsymbol{\alpha}(j)} p_{j,\boldsymbol{\alpha}(j)} m_{\boldsymbol{\alpha}(j)} \\ \text{s.t. } & \mathbf{M}_r(\mathbf{m}_{I_i}) \succeq 0,\ i = 0,\dots,l' \\ & \mathbf{L}_r(q_k\mathbf{m}_{I_{i(k)}}) \succeq 0,\ k = 1,\dots,c, \end{aligned} \tag{5}$$

where $p_{j,\boldsymbol{\alpha}(j)}$ denotes the coefficient of the $\boldsymbol{\alpha}(j)^{th}$ monomial in the polynomial p_j, and $\mathbf{M}_r(\mathbf{m}_{I_i})$ and $\mathbf{L}_r(\mathbf{m}_{I_{i(k)}})$ 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_K^* 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:

$$G(\mathcal{X},\mathcal{U},\mathcal{E},f) \tag{6}$$

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 $\mathcal{E}_m \subset \mathbb{R}^{n_y}$, the set of possible process and measurement noise, and finally, f is the polynomial defining the state equations as follows:

$$\mathbf{x}(k+1) = f\big(\mathbf{x}(k), \mathbf{u}(k)\big) + E_p \boldsymbol{\eta}_p(k) \tag{7}$$

$$\mathbf{y}(k) = \mathbf{x}(k) + E_m \boldsymbol{\eta}_m(k) \tag{8}$$

where $\mathbf{x}, \mathbf{u}$ and $\mathbf{y}$ denote the state, input and output vectors, respectively. The process and measurement noise vectors are represented by $\boldsymbol{\eta}_p$ and $\boldsymbol{\eta}_m$, 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, ϵ_p and ϵ_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

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

We call $\mathcal{B}^N_{poly}(G)$ 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 $\big\{\mathbf{u}(k), \mathbf{y}(k)\big\}_{k=0}^N$, 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

$$\big\{\mathbf{u}(k), \mathbf{y}(k)\big\}_{k=0}^N \in \mathcal{B}^N_{poly}(G). \tag{9}$$

B. Approach

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

$$\phi_k(\mathbf{y}, \mathbf{u}, \boldsymbol{\eta}_m, \boldsymbol{\eta}_p,) := f\big(\mathbf{y}(k) - E_m \boldsymbol{\eta}_m(k), \mathbf{u}(k)\big) + E_p \boldsymbol{\eta}_p(k) - \mathbf{y}(k+1) + E_m \boldsymbol{\eta}_m(k+1) \quad \forall k \in [0, N-1], \tag{10}$$

by combining the state and output equations. Note that $\phi_k(\mathbf{y}, \mathbf{u}, \boldsymbol{\eta}_m, \boldsymbol{\eta}_p)$ is a vector of functions of the size of states. Consider the following optimization problem:

$$\begin{aligned} \epsilon^* = \min_{\boldsymbol{\eta}_m, \boldsymbol{\eta}_p, \epsilon}\ & \epsilon \qquad (\mathrm{P_{MI}}) \\ \text{s. t. } & \phi_k(\mathbf{y}, \mathbf{u}, \boldsymbol{\eta}_m, \boldsymbol{\eta}_p) = 0,\ \forall k \in [0, N-1] \\ & \|\mathbf{y}(k) - \boldsymbol{\eta}_m(k)\|_\infty \le M,\ \forall k \in [0, N] \\ & \|\mathbf{u}(k)\|_\infty \le U,\ \forall k \in [0, N] \\ & \|\boldsymbol{\eta}_p(k)\|_\infty \le \epsilon_p,\ \|\boldsymbol{\eta}_m(k)\|_\infty \le \epsilon,\ \forall k \in [0, N]. \end{aligned}$$

Then, we can state the following proposition:

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

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 ($\mathrm{P_{MI}}$) is not necessarily a convex problem, because of the constraints defined by ϕ functions, which can be any arbitrary polynomial. We leverage moments-based relaxation techniques to solve this problem but first, we show that ($\mathrm{P_{MI}}$) satisfies the running intersection property.

Proposition 2: Problem ($\mathrm{P_{MI}}$) satisfies running intersection property.

Proof: Define $I_k = \{\epsilon, \boldsymbol{\eta}_m(k), \boldsymbol{\eta}_m(k+1), \boldsymbol{\eta}_p(k)\}, \forall k \in [0, N-1]$, and $I_N = \{\epsilon, \boldsymbol{\eta}_m(k)\}$. 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,

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

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

Let $P^r_{\text{MI,rel}}$ be the moments-based relaxation of problem ($\mathrm{P_{MI}}$) of the form (5) with relaxation order r.

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

Proof: From Theorem 2, the optimal value of $P^r_{\text{MI,rel}}$ converges to ϵ^* from below, so we have $\epsilon_r \le \epsilon^*$. 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 $\{\mathbf{u}(k), \mathbf{y}(k)\}_{k=0}^N$ is called *abnormal* for a polynomial model G if and only if $\{\mathbf{u}(k), \mathbf{y}(k)\}_{k=0}^N \notin \mathcal{B}^N_{poly}(G)$.

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_{MI}) 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 starts evolving according to them.

Definition 5: A fault model G^f for a polynomial system G is called *T-step detectable* if $\mathcal{B}^N_{poly}(G) \cap \mathcal{B}^N_{poly}(G^f) = \emptyset$ 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_{MI}) satisfies $\epsilon^* \geq \epsilon_m$ for input-output sequence $\{\mathbf{u}(k), \mathbf{y}(k)\}_{k=k^*-T}^{k^*}$ 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:

$$
\begin{aligned}
\tilde{\epsilon} = \min_{\boldsymbol{\eta}, \bar{\boldsymbol{\eta}}, \mathbf{u}, \mathbf{y}, \epsilon} \ & \epsilon \qquad (\text{P}_T) \\
\text{s.t. } & f\big(\mathbf{y}(k) - E_m \boldsymbol{\eta}_m(k), \mathbf{u}(k)\big) + E_p \boldsymbol{\eta}_p(k) - \mathbf{y}(k+1) \\
& \quad + E_m \boldsymbol{\eta}_m(k+1) = 0, \ \forall k \in [0, T-1] \\
& f^f\big(\mathbf{y}(k) - E^f_m \bar{\boldsymbol{\eta}}_m(k), \mathbf{u}(k)\big) + E^f_p \bar{\boldsymbol{\eta}}_p(k) - \mathbf{y}(k+1) \\
& \quad + E^f_m \bar{\boldsymbol{\eta}}_m(k+1) = 0, \ \forall k \in [0, T-1] \\
& \|\mathbf{u}(k)\|_\infty \leq \min(U, U^f), \ \forall k \in [0, T] \\
& \|\mathbf{y}(k) - E_m \boldsymbol{\eta}_m(k)\|_\infty \leq M, \ \forall k \in [0, T] \\
& \|\mathbf{y}(k) - E^f_m \bar{\boldsymbol{\eta}}_m(k)\|_\infty \leq \bar{M}, \ \forall k \in [0, T] \\
& \|\boldsymbol{\eta}_p(k)\|_\infty \leq \epsilon_p, \ \|\boldsymbol{\eta}_m(k)\|_\infty \leq \epsilon, \ \forall k \in [0, T] \\
& \|\bar{\boldsymbol{\eta}}_p(k)\|_\infty \leq \bar{\epsilon}_p \ \|\bar{\boldsymbol{\eta}}_m(k)\|_\infty \leq \bar{\epsilon}_m, \ \forall k \in [0, T],
\end{aligned}
$$

where $\boldsymbol{\eta}$ 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 $\tilde{\epsilon} > \epsilon_m$.

Proof: Direct consequence of definitions. ∎

Similar to problem (P_{MI}), 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 $P^r_{\text{T,rel}}$ be the moments-based relaxation of (P_T) of the form (5) with relaxation order r, and denote the optimal value of $P^r_{\text{T,rel}}$ with $\tilde{\epsilon}_r$.

Proposition 6: Given a model G, fault model G^f and T, let $\tilde{\epsilon}_r$ denote the optimal value of $P^r_{\text{T,rel}}$. If $\tilde{\epsilon}_r > \epsilon_m$, then G^f is T-detectable.

Proof: Follows from Theorem 2. ∎

For a given T, if the solution of $P^r_{\text{T,rel}}$ results in an objective value greater than ϵ_m, the fault is $T-$detectable. That is, if one can solve the invalidation problem (P_{MI}) 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 $P^r_{\text{MI,rel}}$ at run time. This is established next.

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

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

$$
\begin{aligned}
\epsilon^* := \min_{\{\mathbf{u}(k), \mathbf{y}(k)\}_{k=t}^{t+T}} \ & \epsilon_o(\{\mathbf{u}(k), \mathbf{y}(k)\}_{k=t}^{t+T}) \qquad (11) \\
\text{s.t.} \quad & \{\mathbf{u}(k), \mathbf{y}(k)\}_{k=t}^{t+T} \in \mathcal{B}^T_{poly}(G^f).
\end{aligned}
$$

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

$$
\begin{aligned}
\epsilon^* := \min_{\bar{\boldsymbol{\eta}}, \mathbf{u}, \mathbf{y}} \ & \epsilon_o(\{\mathbf{u}(k), \mathbf{y}(k)\}_{k=t}^{t+T}) \qquad (12) \\
\text{s.t.} \ & f^f\big(\mathbf{y}(k) - E^f_m \bar{\boldsymbol{\eta}}_m(k), \mathbf{u}(k)\big) + E^f_p \bar{\boldsymbol{\eta}}_p(k) - \\
& \quad \mathbf{y}(k+1) + E^f_m \bar{\boldsymbol{\eta}}_m(k+1) = 0, \ \forall k \in [t, t+T-1] \\
& \quad \|\mathbf{u}(k)\|_\infty \leq U^f, \ \forall k \in [t, t+T] \\
& \quad \|\mathbf{y}(k) - E^f_m \bar{\boldsymbol{\eta}}_m(k)\|_\infty \leq \bar{M}, \ \forall k \in [t, t+T] \\
& \quad \|\bar{\boldsymbol{\eta}}_p(k)\|_\infty \leq \bar{\epsilon}_p \ \|\bar{\boldsymbol{\eta}}_m(k)\|_\infty \leq \bar{\epsilon}_m, \ \forall k \in [t, t+T].
\end{aligned}
$$

A relaxation P' of (12) can be obtained by replacing the monomials in $\bar{\boldsymbol{\eta}}, \mathbf{u}, \mathbf{y}$ with moment variables of order upto r. The objective value ϵ' of P' satisfies $\epsilon' \leq \epsilon^*$. Note that P' has bilinear terms containing moments of $\bar{\boldsymbol{\eta}}, \mathbf{u}, \mathbf{y}$ multiplying those of $\boldsymbol{\eta}$ inside $\epsilon_o(\cdot)$. 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 ϵ'. Finally, relaxing the independence assumption and replacing the bilinear terms with cross-moment variables lead to $P^r_{\text{T,rel}}$. Since $P^r_{\text{T,rel}}$ is obtained as a relaxation of P'', we have $\tilde{\epsilon}_r \leq \epsilon'$. Moreover, since T-detectability certificate was obtained

via $P^r_{\text{T,rel}}$ (Proposition 6), we have $\tilde{\epsilon}_r > \epsilon_m$. 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_{MI}) 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:

$$f(\mathbf{x}, \mathbf{u}) = \mathbf{c}^T \mathbf{q}, \tag{13}$$

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

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

$$\tilde{f}(\mathbf{x}, \mathbf{u}, \boldsymbol{\Delta c}) = (\mathbf{c} + \boldsymbol{\Delta c})^T \mathbf{q}, \tag{14}$$

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

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

$$\tilde{\phi}_k(\mathbf{y}, \mathbf{u}, \boldsymbol{\eta}_m, \boldsymbol{\eta}_p, \Delta \mathbf{c}) := \tilde{f}\big(\mathbf{y}(k) - E_m \boldsymbol{\eta}_m(k), \mathbf{u}(k), \Delta \mathbf{c}\big) + E_p \boldsymbol{\eta}_p(k) - \mathbf{y}(k+1) + E_m \boldsymbol{\eta}_m(k+1) \quad \forall k \in [0, T]. \tag{15}$$

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

$$\begin{aligned} \sigma^* = \min_{\boldsymbol{\eta}_m, \boldsymbol{\eta}_p, \Delta \mathbf{c}, \sigma} \ & \sigma \qquad (\text{P}_\Delta) \\ \text{s. t. } & \tilde{\phi}_k(\mathbf{y}, \mathbf{u}, \boldsymbol{\eta}_m, \boldsymbol{\eta}_p, \Delta \mathbf{c}) = 0, \ \forall k \in [0, N] \\ & \|\mathbf{y}(k) - \boldsymbol{\eta}_m(k)\|_\infty \leq M, \ \forall k \in [0, N] \\ & \|T_c \Delta \mathbf{c}\|_\infty \leq \sigma, \ \|\mathbf{u}(k)\|_\infty \leq U, \ \forall k \in [0, N] \\ & \|\boldsymbol{\eta}_p(k)\|_\infty \leq \epsilon_p, \ \forall k \in [0, N] \\ & \|\boldsymbol{\eta}_m(k)\|_\infty \leq \epsilon_m, \ \forall k \in [0, N]. \end{aligned}$$

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.

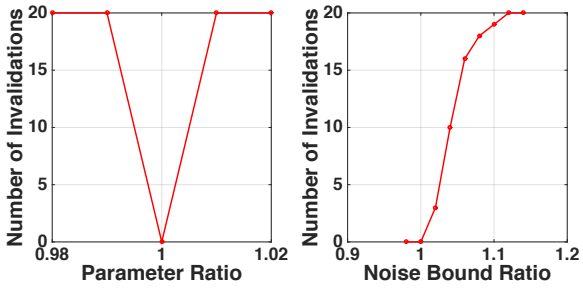

Fig. 1. The invalidation results over 20 trials (left) when there is a mismatch between the parameters of the model that generated the data is β times those of the a priori model, and (right) when the data is generated with noise, bounded by α times the a priori noise bound.

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[2] toolbox.

A. Numerical Examples

Consider a discrete-time pendulum model with friction:

$$\begin{aligned} x_1(k+1) &= x_1(k) + 0.3 x_2(k) \\ x_2(k+1) &= 0.9 x_2(k) - 0.3 \sin(x_1(k)). \end{aligned} \tag{16}$$

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 α values. The run is repeated for 20 times for each α and the number of invalidations are illustrated in Fig. 1. In the second set of examples, we fix α at 1, and change the parameter β, 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.

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

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.

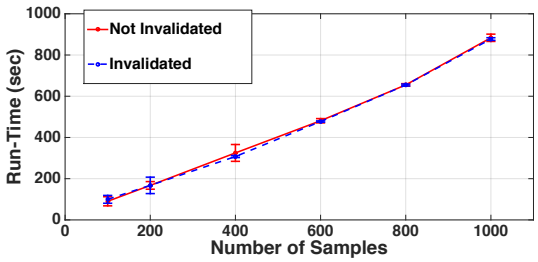

Fig. 2. Model invalidation run-time increase almost linearly with the increase in time horizon for both valid and invalid data.

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 optimal value obtained from $P^{r}_{\text{MI,rel}}$ as a measure of how "far" the model is from being valid. In this example, the $\sin(x)$ term is replaced with its Taylor series expansion of orders $2, 4, 6$ and 8 around zero, and $P^{4}_{\text{MI,rel}}$ is solved for the data generated from the nonlinear model (16) while the a priori model is the polynomial approximation obtained from the Taylor series expansion. For each order, an output data sequence of length 20 is used. Fig. 3 shows that by increasing the order of Taylor expansion, the optimal solution to $P^{4}_{\text{MI,rel}}$, ϵ_4 decreases, which means the model becomes closer to being valid. The ϵ_4 values can be used as a measure of how valid is a polynomial model to represent a nonlinear model.

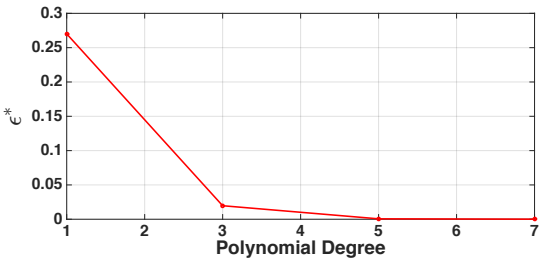

Fig. 3. The optimal value of $P^{4}_{\text{MI,rel}}$ versus the order of the Taylor series expansion.

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:

$$\begin{aligned}\Phi(k+1) &= \Phi(k) - T_s\big(\Psi(k) + \frac{3}{2}\Phi^2(k) + \frac{1}{2}\Phi^3(k)\big) + \sigma \\ \Psi(k+1) &= \Psi(k) + 3T_s\Phi(k) - T_s\Psi(k)\end{aligned} \tag{17}$$

where, Φ is the mass flow and Ψ is the pressure rise. The measurements are the noisy states with noise vector $\boldsymbol{\eta}_m$.

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:

$$\begin{aligned}\Phi(k+1) &= \Phi(k) - T_s\big(\Psi(k) + \frac{3}{2}\Phi^2(k) + \frac{1}{2}\Phi^3(k)\big) + \sigma \\ &\quad + b_f \\ \Psi(k+1) &= \Psi(k) + 3T_s\Phi(k) - T_s\Psi(k),\end{aligned} \tag{18}$$

where b_f indicates the faulty bias. The output of faulty model is noisy measurements (with $\bar{\boldsymbol{\eta}}_m$) 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 $\boldsymbol{\eta}_m$ and $\bar{\boldsymbol{\eta}}_m$. 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.

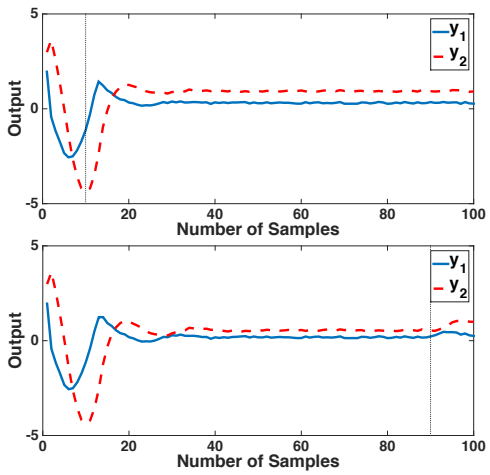

Fig. 4. The outputs of jet engine model when fault occurs at time sample 10 (top), and when it occurs at time sample 90 (bottom).

The two vertical dotted lines indicate where the fault is occurred for each case. The goal of this example is to illustrate the advantage of knowing a given fault is $T-$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

are invalidated. The same thing happens for the case 2, at W_{92}. 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 model being valid or invalid was obtained, which was further shown to be asymptotically tight. Moreover, we introduced the notion of T-detectability of a fault with a polynomial state-space representation, with respect to a given polynomial system model, and presented an optimization-based approach that can be used to compute T. 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 run-time, therefore decentralized fault detection algorithms are required.

References

[1] E. Lee. Cyber physical systems: Design challenges. In *11th IEEE International Symposium on Object Oriented Real-Time Distributed Computing (ISORC)*, pages 363–369. IEEE, 2008.

[2] M. Sznaier, O. Camps, N. Ozay, and C. Lagoa. Surviving the upcoming data deluge: A systems and control perspective. In *51st IEEE Conference on Decision and Control (CDC)*, Dec 2014.

[3] V. Calderaro, C. N. Hadjicostis, A. Piccolo, and P. Siano. Failure identification in smart grids based on Petri Net modeling. *IEEE Transactions on Industrial Electronics*, 58(10):4613–4623, Oct 2011.

[4] F. Pasqualetti, F. Dorfler, and F. Bullo. Attack detection and identification in cyber-physical systems. *IEEE Transactions on Automatic Control*, 58(11):2715–2729, Nov 2013.

[5] R. S. Smith and J. C. Doyle. Model validation: A connection between robust control and identification. *IEEE Transactions on Automatic Control*, 37(7):942–952, 1992.

[6] S. Prajna. Barrier certificates for nonlinear model validation. *Automatica*, 42(1):117–126, 2006.

[7] P. Rumschinski, S. Borchers, S. Bosio, R. Weismantel, and R. Findeisen. Set-base dynamical parameter estimation and model invalidation for biochemical reaction networks. *BMC systems biology*, 4(1):69, 2010.

[8] Y. Cheng, Y. Wang, M. Sznaier, N. Ozay, and C. Lagoa. A convex optimization approach to model (in)validation of switched arx systems with unknown switches. In *IEEE 51st Annual Conference on Decision and Control (CDC)*, pages 6284–6290, Dec 2012.

[9] N. Ozay, M. Sznaier, and C. Lagoa. Model (in) validation of switched arx systems with unknown switches and its application to activity monitoring. In *49th IEEE Conference on Decision and Control (CDC)*, pages 7624–7630, Dec 2010.

[10] N. Ozay, M. Sznaier, and C. Lagoa. Convex certificates for model (in)validation of switched affine systems with unknown switches. *IEEE Transactions on Automatic Control*, 59(11):2921–2932, Nov 2014.

[11] P. M. Frank. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy-a survey and some new results. *Automatica*, 26(3):459–474, May 1990.

[12] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri. A review of process fault detection and diagnosis: Part i: Quantitative model-based methods. *Computers and Chemical Engineering*, 27(3):293 – 311, 2003.

[13] R. Isermann. Supervision, fault-detection and diagnosis methods a short introduction. In *Fault-Diagnosis Applications*, pages 11–45. Springer Berlin Heidelberg, 2011.

[14] P. Rosa, C. Silvestre, J. S. Shamma, and M. Athans. Fault detection and isolation of LTV systems using set-valued observers. In *IEEE Conference on Decision and Control (CDC)*, pages 768–773. IEEE, 2010.

[15] R. Nikoukhah. Guaranteed active failure detection and isolation for linear dynamical systems. *Automatica*, 34(11):1345–1358, 1998.

[16] R. Nikoukhah and S. Campbell. Auxiliary signal design for active failure detection in uncertain linear systems with a priori information. *Automatica*, 42(2):219–228, 2006.

[17] J. K. Scott, R. Findeisen, R. D Braatz, and D. M. Raimondo. Input design for guaranteed fault diagnosis using zonotopes. *Automatica*, 50(6):1580–1589, 2014.

[18] F. Harirchi and N. Ozay. Model invalidation for switched affine systems with applications to fault and anomaly detection. In *IFAC Conference on Analysis and Design of Hybrid Systems (ADHS)*, Oct 2015.

[19] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, March 2000.

[20] J. B. Lasserre. Convergent SDP-relaxations for polynomial optimization with sparsity. 4151:263–272, 2006.

[21] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. *Emerging Applications of Algebraic Geometry*, 149(1):157–270, Springer, 2009.

[22] S. X. Ding. *Model-based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*. Springer, 2008.

[23] J. Löfberg. Yalmip : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[24] H. Waki, S. Kim, M. Kojima, M. Muramatsu, and H. Sugimoto. Algorithm 883: Sparsepop– a sparse semidefinite programming relaxation of polynomial optimization problems. *ACM Trans. Math. Softw.*, 35(2):15:1–15:13, July 2008.

[25] M. Krstic, I. Kanellakopoulos, and P. Kokotovic. Nonlinear and adaptive control design, 1995. *John Wiley & Sons*, pages 21–86.

[26] E. Aylward, P. Parrilo, and J. Slotine. Stability and robustness analysis of nonlinear systems via contraction metrics and SOS programming. *Automatica*, 44(8):2163–2170, 2008.