Automatica 50 (2014) 2038–2050

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Probabilistic model validation for uncertain nonlinear systems☆

Abhishek Halder [1], Raktim Bhattacharya

Department of Aerospace Engineering, Texas A&M University, College Station, TX 77843-3141, United States

ARTICLE INFO

Article history:
Received 20 January 2013
Received in revised form
6 February 2014
Accepted 18 April 2014
Available online 17 June 2014

Keywords:
Model validation
Uncertainty propagation
Optimal transport
Wasserstein distance

ABSTRACT

This paper presents a probabilistic model validation methodology for nonlinear systems in time-domain. The proposed formulation is simple, intuitive, and accounts both deterministic and stochastic nonlinear systems with parametric and nonparametric uncertainties. Instead of hard invalidation methods available in the literature, a relaxed notion of validation in probability is introduced. To guarantee provably correct inference, algorithm for constructing probabilistically robust validation certificate is given along with computational complexities. Several examples are worked out to illustrate its use.

1. Introduction

A model serves as a mathematical abstraction of the physical system, providing a framework for system analysis and controller synthesis. Since such mathematical representations are based on assumptions specific to the process being modeled, it is important to quantify the reliability to which the model is consistent with the physical observations. Model quality assessment is imperative for applications where the model needs to be used for prediction (e.g. weather forecasting, stock market) or safety-critical control design (e.g. aerospace, nuclear, systems biology) purposes.

Here it is important to realize that a model can only be validated against experimental observations, not against another model. Thus a *model validation problem* can be stated as: *given a candidate model and experimentally observed measurements of the physical system, how well does the model replicate the experimental measurements*? It has been argued in the literature (Poolla, Khargonekar, Tikku, Krause, & Nagpal, 1994; Popper, 2002; Prajna, 2006; Smith & Doyle, 1992) that the term 'model validation' is a misnomer since it would take infinite number of experimental observations to do so. Hence the term 'model invalidation' or 'falsification' (Brugarolas & Safonov, 2002) is preferred. In this paper, instead of hard invalidation, we will consider the validation/invalidation problem in a probabilistically relaxed sense.

1.1. Related literature

Broadly speaking, there have been three distinct frameworks in which the model validation problem has been attempted till now. *One* is a discrete formulation in *temporal logic framework* (Baier & Katoen, 2008) which has been extended to account probabilistic models (Baier & Katoen, 2008; Ciesinski & Größer, 2004). *Second* is the $\mathcal{H}_\infty$ *control framework* where time-domain (Chen & Wang, 1996; Poolla et al., 1994; Smith & Dullerud, 1996), frequency domain (Smith & Doyle, 1992; Wahlberg & Ljung, 1992) and mixed domain (Xu, Ren, Gu, & Chen, 1999) model validation methods have been studied assuming structured norm-bounded uncertainty in linear dynamics setting. The *third* framework involves deductive inference based on barrier certificates (Prajna, 2006) which was shown to encompass a large class of nonlinear models including differential–algebraic equations (Campbell, 1980), dynamic uncertainties described by integral quadratic constraints (Megretski & Rantzer, 1997), stochastic (Øksendal, 2003) and hybrid dynamics (van der Schaft & Schumacher, 1999).

In statistical setting, model validation has been addressed from system identification perspective (Ljung, 1999; Ljung & Guo, 1997) where the main theme is to validate an identified nominal model through correlation analysis of the residuals. A polynomial chaos framework has also been proposed (Ghanem, Doostan, & Red-Horse, 2008) for model validation. Gevers, Bombois, Codrons, Scorletti, and Anderson (2003) have connected the robust

☆ This research was supported by NSF award #1016299 with Helen Gill as the program manager. The material in this paper was partially presented at the 50th IEEE Conference on Decision and Control (CDC) and European Control Conference (ECC), December 12–15, 2011, Orlando, Florida, USA and the 51st IEEE Conference on Decision and Control (CDC), December 10–13, 2012, Maui, Hawaii, USA. This paper was recommended for publication in revised form by Associate Editor Fabrizio Dabbene under the direction of Editor Roberto Tempo.

E-mail addresses: ahalder@tamu.edu (A. Halder), raktim@tamu.edu (R. Bhattacharya).

[1] Tel.: +1 979 583 6070; fax: +1 979 845 6051.

http://dx.doi.org/10.1016/j.automatica.2014.05.026
0005-1098/

control framework with prediction error based identification for frequency-domain validation of linear systems. In another vein, using Bayesian conditioning, Lee and Poolla (1996) showed that for *parametric* uncertainty models, the statistical validation problem may be reduced to the computation of relative weighted volumes of convex sets. However, for *nonparametric* models: "the situation is significantly more complicated" (Lee & Poolla, 1996) and to the best of our knowledge, has not been addressed in the literature. Recently, in the spirit of weak stochastic realization problem (van Schuppen, 1989), Ugrinovskii (2009) investigated the conditions for which the output of a stochastic nonlinear system can be realized through perturbation of a nominal stochastic *linear* system.

In practice, one often encounters the situation where a model is either proposed from physics-based reasoning or a reduced order model is derived for computational convenience. In either case, the model can be linear or nonlinear, continuous or discrete-time, and in general, it is not possible to make any *a priori* assumption about the noise. Given the experimental data and such a candidate model for the physical process, our task is to answer: "to what extent, the proposed model is valid?" In addition to quantify such degree of validation, one must also be able to demonstrate that the answer is *provably correct* in the face of uncertainty. This brings forth the notion of *probabilistically robust model validation*. In this paper, we will show how to construct such a *robust validation certificate*, guaranteeing the performance of probabilistic model validation algorithm.

1.2. Contributions of this paper

With respect to the literature, the contributions of this paper are as follows.

(1) Instead of interval-valued structured uncertainty (as in $\mathcal{H}_\infty$ control framework) or moment based uncertainty (as in parametric statistics framework), this paper deals with model validation in the sense of nonparametric statistics. Uncertainties in the model are quantified in terms of the probability density functions (PDFs) of the associated random variables. We argue that such a formulation offers several advantages. *Firstly*, we show that model uncertainties in the parameters, initial states and input disturbance, can be propagated accurately by spatio-temporally evolving the joint state and output PDFs. Since experimental data usually come in the form of histograms, it is a more natural quantification of uncertainty than specifying sets (Prajna, 2006) to which the trajectories are contained at each instant of time. However, if needed, such sets can be recovered from the supports of the instantaneous PDFs. *Secondly*, as we will see in Section 4.4, instead of simply invalidating a model, our methodology allows to estimate the probability that a proposed model is valid or invalid. This can help to decide which specific aspects of the model need further refinement. Hard invalidation methods do not cater such constructive information. *Thirdly*, the framework can handle both discrete-time and continuous-time nonlinear models which need not be polynomial. Previous work like (Prajna, 2006) dealt with semialgebraic nonlinearities and relied on the sum of squares (SOS) decomposition (Parrilo, 2000) for computational tractability. From an implementation point of view, the approach presented in this paper does not suffer from such conservatism.
(2) Due to the uncertainties in initial conditions, parameters, and process noise, one needs to compare the output ensembles instead of comparing the individual output realizations. This requires a metric to quantify closeness between the experimental data and the model, in the sense of distribution. We propose *Wasserstein distance* to compare the output PDFs and argue why commonly used information-theoretic notions like *Kullback–Leibler divergence* may not be appropriate for this purpose.
(3) We show that the uncertainty propagation through continuous or discrete-time dynamics can be done via numerically efficient meshless algorithms, even when the model is high-dimensional and strongly nonlinear. Moreover, we outline how to compute the Wasserstein distance in such settings. Further, bringing together ideas from the analysis of randomized algorithms, we outline how sample-complexity bounds can be derived for robust validation inference.

The paper is organized as follows. In Section 2, we describe the problem setup. Then we expound on two main steps of our validation framework, viz. uncertainty propagation, and distributional comparison in Sections 3 and 4, respectively. We provide numerical examples in Section 5, to illustrate the ideas presented in this paper, followed by conclusions in Section 6.

Notation

We use the superscript $^\top$ to denote matrix transpose, $\otimes$ to denote Kronecker product, and the symbol $\wedge$ to denote a minimum of two real numbers. The notation ${}_rF_s\left(a_1, \ldots, a_r; b_1, \ldots, b_s; x\right)$ stands for generalized hypergeometric function. The symbols $\mathcal{N}\left(.,.\right)$, and $\mathcal{U}\left(.\right)$ are used for normal and uniform PDFs, respectively. We use the notation $\xi_0\left(.\right)$ to denote the joint PDF over initial states and parameters. $\xi\left(.,t\right)$ and $\widehat{\xi}\left(.,t\right)$ denote joint PDFs over instantaneous states and parameters, for the true and model dynamics, respectively. Similarly, $\eta\left(.,t\right)$ and $\widehat{\eta}\left(.,t\right)$, respectively denote joint PDFs over output spaces y and $\widehat{y}$ at time t, for the true and model dynamics. The symbol $\widetilde{x}$ is used to denote the extended state vector obtained by augmenting the state (x) and parameter (p) vectors. We use χ to denote indicator function and # to denote cardinality. Unless stated otherwise, $\delta\left(.\right)$ stands for Dirac delta. The symbol I_ℓ denotes the ℓ-by-ℓ identity matrix, ∇_x denotes gradient operator with respect to vector x, $\mathrm{vec}\left(\cdot\right)$ stands for the vectorization operator, and $\|\cdot\|_F$ denotes the Frobenius norm. $\mathrm{tr}\left(\cdot\right)$ and $\det\left(\cdot\right)$ stand for trace and determinant of a matrix. The abbreviations *a.s.* and *i.p.* refer to convergence in *almost sure* and *in probability* sense. The shorthand ∂_α means partial derivative with respect to variable α, $\mathrm{supp}\left(\cdot\right)$ denotes support of a function, and $\mathrm{erf}(\cdot)$ stands for error function.

2. Problem setup

2.1. Intuitive idea

The proposed framework is based on the evolution of densities in the output space, instead of evolution of individual trajectories, as in the Lyapunov framework. Intuitively, characteristics of the input to output mapping is revealed by the growth or depletion of trajectory concentrations in the output space. Growth in concentration, or increased density, defines the regions where the trajectories accumulate. This corresponds to the regions with slow time scale dynamics or time invariance. Similarly, depletion of concentration in a set implies fast time scale dynamics or unstable manifold. We refer the readers to Lasota and Mackey (1994) for an introduction to the analysis of dynamical systems using trajectory densities. This idea of comparing dynamical systems based on density functions, have been presented before by Sun and Mehta (2010) in the context of filtering, and by Georgiou (2007) in the context of matching power spectral densities.

2.1.1. Proposed approach

Given the experimental measurements of the physical system in the form of a time-varying distribution (such as histograms), we propose to compare the *shape* or *concentration profile* of this measured output density, with that predicted by the model. At every