Computers and Chemical Engineering 64 (2014) 138–152

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Identification-based optimization of dynamical systems under uncertainty

Siyun Wang, Michael Baldea*

McKetta Department of Chemical Engineering, The University of Texas at Austin, 200 East Dean Keeton Street, Stop C0400, Austin, TX 78712, USA

ARTICLE INFO

Article history:
Received 1 September 2013
Received in revised form 20 January 2014
Accepted 2 February 2014
Available online 10 February 2014

Keywords:
Dynamical systems
Optimization
Uncertainty
Random sequences
System identification

ABSTRACT

The operation of chemical processes is inherently subject to uncertainty. Traditionally, uncertainties have been accounted for in system design by discretizing the uncertainty space and considering the resulting ensemble of scenarios in solving the design optimization problem. Scenario-based approaches are computationally demanding and can rapidly become intractable. We propose identification-based optimization (IBO) as a novel framework for the optimal design of dynamical systems under uncertainty. Our method originates in nonlinear system identification theory, and is predicated on representing uncertain variables as pseudo-random multi-level signals (PRMSs), which are imposed on the system model during each time integration step of a dynamic optimization. The uncertainty space is thus efficiently sampled without using computationally expensive scenario sets. We establish a procedure for generating PRMSs for uncertain variables based on their probability density functions. The computational benefits of IBO are illustrated through comparative case studies.

1. Introduction

Fluctuating market conditions and the need for switching between an increasingly diverse portfolio of conventional and renewable feedstock are at the origin increased uncertainty in the operation of chemical and energy generation processes (Baldea & Daoutidis, 2012). This translates into fluctuations in the process variables (flow rates, pressures, temperatures, etc.) in the entire process system. These fluctuations can lower process efficiency and can lead to unsatisfactory product quality. It is therefore essential that such uncertainty be accounted for at the system design stage.

The challenge of designing process systems that are capable of coping with uncertainty has been recognized at an early stage in process systems engineering (Biegler & Grossmann, 2004). Initial research has focused on steady-state design, where a two-stage approach was adopted from the operations research literature (Ierapetritou, Acevedo, & Pistikopoulos, 1996; Pistikopoulos, 1995; Pistikopoulos & Ierapetritou, 1995). The two-stage formulation was subsequently extended to multi-stage approaches (Cheng, Subrahmanian, & Westerberg, 2003; Masoumeh, Mustapha, & Daoud, 2009; Pereira & Pinto, 1991; Verderame, Elia, Li, & Floudas, 2010). These formulations focus on minimizing the expected value of the design objective function. In a different vein, chance-constrained optimization was proposed, centered on ensuring that the probability of violating design constraints remains within desired limits (Arellano-Garcia & Wozny, 2009; Lapteva, Ziyatdinov, Ostrovskii, & Pervukhin, 2010; Li, Garcia, & Wozny, 2008; Wendt, Li, & Wozny, 2002). These methodologies have found applications in the design of typical unit operations (e.g., binary distillation columns Wendt et al., 2002), as well as in process-wide settings, such as the optimization of heat exchanger networks and reactor–separator systems (Grossmann & Sargent, 1978).

The design of dynamical systems under uncertainty has also been studied. In this case, the uncertain variables need to be considered as functions of time, and the goal is to find a design that minimizes the objective function and satisfies both end-point constraints and path constraints subject to time-dependent fluctuations in the uncertain variables. Both small-scale lumped-parameter systems and large-scale distributed-parameter systems, involving both continuous and integer decision variables have been investigated using methods based on the aforementioned steady-state optimization ideas (Bansal, Perkins, & Pistikopoulos, 2002; Bansal, Sakizlis, Ross, Perkins, & Pistikopoulos, 2003; Mohideen, Perkins, & Pistikopoulos, 1996, 1997).

In its most general form, the problem of optimizing a system with uncertain parameters (expressed, e.g., in terms of the moments of a distribution), is infinite-dimensional. Thus, solution methods (for both steady state and dynamic problems) rely on the discretization of the stochastic variables and the generation of an ensemble of scenarios that must be considered simultaneously. Discretization augments the size of the problem, with a corresponding increase in computational effort. Attempts at mitigating computational challenges have spurred developments

* Corresponding author. Tel.: +1 512 471 1281.
E-mail address: mbaldea@che.utexas.edu (M. Baldea).

0098-1354/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compchemeng.2014.02.001

Nomenclature

J	objective function
$\mathbf{x}$	state vector of the system
$\mathbf{u}$	vector of manipulated variables
$\boldsymbol{\theta}$	vector of uncertain variables
t	time
$\mathbf{h}(\cdot)$	equality constraints
$\mathbf{g}(\cdot)$	inequality constraints
T_o	time horizon of the design problem
$\mathbf{d}$	vector of design variables
$\mathbf{z}$	vector of decision variables
$P(\cdot)$	probability
α	significance level
I	number of inequality constraints
N_{SAA}	number of samples for the SAA approach
k	sample index
$\bar{\boldsymbol{\theta}}(t)$	a PRMS of θ
q	size of a Galois field
n	order of a primitive polynomial
s_i	the ith term of an m-sequence
c_j	coefficients of the primitive polynomial
T_s	switching time
ω_{max}	maximum frequency of the uncertain variable
ω_B	process bandwidth
κ	design parameter for frequency band
$s(t)$	intermediate pseudo-random multi-level signal
$\bar{\boldsymbol{\theta}}_k$	discretized realization of the uncertain variable vector
w_k	probability of $\bar{\boldsymbol{\theta}}_k$
ν_k	quota of levels
μ	number of switches
$M(\cdot)$	mapping from $s(t)$ to $\bar{\boldsymbol{\theta}}(t)$
$\Phi(\mathbf{x}, t)$	a term in the objective function or constraints, the integral of which depends on T_o
$\tilde{\Phi}(\mathbf{x}, t)$	an equivalent form to the integral of $\Phi(\mathbf{x}, t)$ in the IBO formulation
$\mathbf{x}_{ss}$	steady state of the system
N_m	length of an m-sequence
$\mathbf{k}(\cdot)$	control law
$\mathbf{p}$	controller parameter vector
t'	shifted time variable
T_{PRMS}	length of a PRMS

at the algorithmic level using, e.g., generalized Benders decomposition strategies (Benders, 1962) and, more recently, parallel computing (Zhu, Word, Siirola, & Laird, 2009), as well as a quest for more efficient sampling methods that can reduce the number of scenarios required to accurately capture uncertainty. However, scenario-based approaches remain time- and resource-consuming, and the development of efficient methods for optimizing the design of systems (in particular, dynamical systems) under uncertainty remains an important and open problem.

This contribution introduces identification-based optimization (IBO) as a novel framework for the optimal design of dynamical systems under uncertainty. Our approach originates in nonlinear system identification theory, and is predicated on representing uncertain variables (whose probability distributions are assumed to be known, and under additional ergodicity assumptions) as pseudo-random multi-level signals (PRMSs), which are imposed on the system model during each time integration step of a dynamic optimization. This allows the system to efficiently sample the entire span of the uncertainty space in an efficient manner without the need for generating computationally expensive scenario sets. The paper is organized as follows: we begin by formally stating the problem of process optimization under uncertainty, and provide a brief review of solution approaches published in the literature to date. We then describe the principle of IBO and establish a general procedure for the generation of PRMS for any uncertain variable given its (not necessarily normal) probability distribution. Finally, we illustrate the proposed framework with two process system case studies, demonstrating significant computational benefits compared to conventional scenario-based approaches.

2. Preliminaries

The solution of optimization problems for systems operating under uncertainty entails, (i) formulating the (infinite-dimensional) problem, followed by (ii) converting the problem into a deterministic, finite-dimensional (mixed-integer) (non)linear program, which, (iii) can be solved using available optimization algorithms. In this section, we provide a brief account of corresponding formulations and solution methods; the reader is invited to consult available literature reviews (e.g., Sahinidis, 2004) for further information.

2.1. Problem definition

This step involves developing a mathematical representation for a design optimization problem from the physical realm. In particular, a dynamic model of the system under consideration should be constructed. Information regarding the uncertain variables should also be captured at this stage. Once these are available, a generic formulation of the dynamic stochastic program can be written as:

$$\begin{aligned} \min_{\mathbf{z}} \quad & J(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{z}, \boldsymbol{\theta}, t) \\ \text{s.t.} \quad & \mathbf{h}(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{z}, \boldsymbol{\theta}, t) = 0 \\ & \mathbf{g}(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{z}, \boldsymbol{\theta}, t) \le 0 \\ & t \in [0, T_o] \end{aligned} \tag{1}$$

where $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^x$ are the state variables, $\mathbf{z} \in \mathcal{Z} \subset \mathbb{R}^z$ are the (possibly time-dependent) decision variables, t is time, $\boldsymbol{\theta}$ is the vector of uncertain variables, $J \in \mathbb{R}$ is the objective function, T_o is the time horizon, and $\mathbf{h}$ and $\mathbf{g}$ are vector fields capturing equality and inequality constraints; the former include the model equations (typically in differential-algebraic equation form), while the latter amount to design constraints. We note that solving the generic problem (1) poses several challenges. First, it is difficult (or impossible) to find a unique $\mathbf{z}$ such that the equality constraints hold for *all* the realizations of $\boldsymbol{\theta}$. Second, it is possible that the "strictest" inequality constraint will only be active for a small number of the possible realizations of $\boldsymbol{\theta}$, which, from a practical perspective, amounts for optimizing the system for the worst case scenario. These challenges can be dealt with using several techniques, which are reviewed below.

2.1.1. Multi-stage formulations

Staged formulations rely on separating the decision variables $\mathbf{z}$ into two categories, the time independent *design* variables $\mathbf{d}$, and the time-dependent controls $\mathbf{u}$. Typical examples of the former include equipment sizes, while the latter are represented by manipulated variables such as material flow rates. The essence of the *two-stage* formulation is the use of control variables to reject the impact of time-dependent uncertainties, such that excessively large values of the design variables (i.e., equipment overdesign) are avoided (Pistikopoulos & Ierapetritou, 1995). The two-stage stochastic problem formulation is described mathematically as