Annual Reviews in Control 000 (2017) 1–11

Contents lists available at ScienceDirect

Annual Reviews in Control

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

Review article

Stochastic model predictive control with active uncertainty learning: A Survey on dual control☆

Ali Mesbah

Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA

ARTICLE INFO

Article history:
Received 4 August 2017
Revised 25 October 2017
Accepted 3 November 2017
Available online xxx

Keywords:
Stochastic optimal control
Stochastic model predictive control
Active uncertainty learning
Implicit dual control
Explicit dual control

ABSTRACT

This paper provides a review of model predictive control (MPC) methods with active uncertainty learning. System uncertainty poses a key theoretical and practical challenge in MPC, which can be aggravated when system uncertainty increases due to the time-varying nature of system dynamics. For uncertain systems with stochastic uncertainty, this paper presents the stochastic MPC (SMPC) problem in the dual control paradigm, where the control inputs to an uncertain system have a probing effect for active uncertainty learning and a directing effect for controlling the system dynamics. The complexity of the SMPC problem with dual control effect is described in connection to stochastic dynamic programming as well as Bayesian estimation for its output feedback implementation. Further, implicit and explicit dual control methods for approximating the receding-horizon control problem with dual control effect are surveyed and analyzed with the intent to discuss the key challenges and opportunities in SMPC with dual control effect.

Contents

1. Introduction

Model predictive control (MPC), also known as *receding-horizon control*, has demonstrated exceptional success for the high-performance control of technical systems in a variety of applications such as automotive applications, building climate control, microgrids, process systems control, and robotics and vehicle path planning (Lee, 2011; Mayne, 2014; Morari & Lee, 1999). The ability to systematically cope with multivariable system dynamics and system constraints has made MPC an attractive optimal control strategy (Morari & Lee, 1999). A key theoretical and practical challenge in MPC, however, arises from handling system uncertainty under closed loop. Even though MPC exhibits some degree of robustness to sufficiently small uncertainties due to its receding-horizon implementation, a marginal robust performance may not be adequate in many practical situations. This consideration has led to development of robust MPC (Bemporad & Morari, 1999; Mayne, 2014) and stochastic MPC (SMPC) (Kouvaritakis & Cannon, 2016; Mesbah, 2016) strategies that account for system uncertainties of, respectively, deterministic, bounded-set and probabilistic nature. SMPC is generally intended to guarantee robust stability and performance of the closed-loop system in a probabilistic sense by explicitly incorporating a probabilistic description of model uncertainty into an optimal control problem. Specifically, SMPC allows for handling system constraints probabilistically using *chance constraints* (Mesbah, 2016).

The MPC framework, however, cannot *actively* learn about the system uncertainty. In particular, the MPC framework is incapable

☆ This paper was not presented at any IFAC meeting.
E-mail address: mesbah@berkeley.edu

https://doi.org/10.1016/j.arcontrol.2017.11.001
1367-5788/

of proactively coping with system changes that may occur due to the time-varying nature of system dynamics, which can in turn increase uncertainty even more, leading to control performance degradation (e.g., see MacGregor and Cinar (2012); Mesbah, Bombois, Forgione, Hjalmarsson, and den Hof (2015); Patwardhan and Shah (2002); Qin (2012); Zagrobelny, Ji, and Rawlings (2013) for performance monitoring and diagnosis). Thus, active learning of system uncertainty and regular adaptation of uncertainty descriptions, via system re-identification (Ljung, 1999) or Bayesian inference (Chen, 2003), in the optimal control problem is crucial for maintaining the MPC performance for uncertain systems.

The problem of parametric model uncertainty in model-based control has inspired several important research directions in the field of (stochastic) *adaptive control*, which generally relies on the basic notion of regular model adaptation under closed loop (e.g., see Åström and Wittenmark (2008); Dumont and Huzmezan (2002); Filatov and Unbehauen (2004); Landau, Lozano, M'Saad, and Karimi (2011); Sastry and Bodson (2011); Wittenmark (1995)). Historically, adaptive control is based on the *separation principle*, which involves separation of parameter estimation/system identification and control design (Patchell & Jacobs, 1971; Witsenhausen, 1971). That is, a system model is first identified and, subsequently, used for designing a *certainty equivalence* (CE) controller, as if the model was an exact representation of the system (Bar-Shalom & Tse, 1974). In optimal control, when the control cost function is quadratic and the system is linear in uncertain parameters that are Gaussian processes (i.e., linear-quadratic-Gaussian control), the separation principle is exact and the resulting deterministic optimal control law is equivalent to the stochastic optimal control law (Stengel, 1986). However, neither the separation principle is optimal for general systems, nor the deterministic framework of CE control can account for the model uncertainty. These shortcomings can result in severe robustness problems in optimal control of uncertain systems such as the well-known bursting problem (Anderson, 1985).

Alternatively, explicit incorporation of model uncertainty into a control design problem leads to *cautious control*, that is, the control inputs will become small (cautious) when the uncertainty is large (Wittenmark, 1975b). Cautious control action arises from predicting the impact of control inputs on the future system uncertainty. In the limiting case of a one-step-ahead predictor, however, cautious control can yield exceedingly small control inputs when the uncertainty grows. Small control inputs will in turn generate less information about the uncertain system and, consequently, the uncertainty will be increased further, eventually yielding zero control inputs (Åström & Wittenmark, 1971). This is because, like CE control, the control inputs in cautious control do not have a probing effect. Therefore, control inputs are only *passively adaptive* since uncertainty learning is accidental, merely due to the feedback action of the controller.

The seminal work of Feldbaum was the first to recognize that control inputs to an uncertain system must have a *probing effect* for active learning of system uncertainty and a *directing effect* for controlling the system dynamics (Feldbaum, 1960a,b, 1961a,b). That is, control inputs should have the *dual control effect* of influencing not only the system states, but also the uncertainty associated with the states (Bar-Shalom & Tse, 1974). The dual control paradigm provides a unified framework for stochastic optimal control and model uncertainty handling. Dual control maintains an optimal balance (in the sense of the principle of optimality (Bellman, 1957)) between the probing activity and control activity of control inputs, which are naturally in conflict. This arises from systematically accounting for the possibility of poor transient control performance due to probing in order to achieve improved control performance in future because of reduced system uncertainty. Despite its conceptually appealing features, dual control relies on stochastic dynamic programing (DP) (Bellman, 1957), which is computationally intractable even for moderately-sized systems (Åström & Wittenmark, 2008). To address the computational complexity of dual control, various approximate solutions have been proposed that can be broadly categorized into (Filatov & Unbehauen, 2000): (i) *implicit dual control* that involves direct approximation of the stochastic DP problem and (ii) *explicit dual control* that involves reformulation of the dual control problem into a tractable optimal control problem with some form of heuristic-based probing effect for active uncertainty learning.

The main objective of this paper is to provide a comprehensive review of receding-horizon control methods with active uncertainty learning. First, the SMPC problem with dual control effect is presented, which naturally incorporates system probing and output feedback to enhance the information content of system observations for stochastic uncertainty handling (Section 2). The complexity of the SMPC problem with dual control effect is discussed in light of different classes of control inputs and their relation to solution of Bellman equation for stochastic DP (Section 3). Implicit and explicit dual control methods for approximating the receding-horizon control problem with dual control effect are then surveyed (Section 4). The paper is intended to provide the reader with an impression of the key theoretical issues in receding-horizon control with dual control effect without undue mathematical complexity.

Notation. $\mathbb{R}$ and $\mathbb{N} = \{1, 2, \dots\}$ denote the sets of real and natural numbers, respectively; $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.$(\Omega, \mathsf{F}, \mathsf{P})$ denotes a probability space, where random variables are F-measurable functions of ω that is a generic element of the sample space Ω, F is the σ-algebra of sets in Ω, and P is the probability measure on F. $\mathsf{P}[\cdot|\cdot]$ denotes the conditional probability.

2. SMPC With dual control effect

Consider a discrete-time, uncertain system

$$x_{k+1} = f(x_k, u_k, w_k, \theta), \tag{1a}$$

$$y_k = h(x_k, v_k), \tag{1b}$$

where $k \in \mathbb{N}_0$ is the time index; $x_k \in \mathbb{R}^{n_x}$, $u_k \in \mathbb{R}^{n_u}$, and $y_k \in \mathbb{R}^{n_y}$ denote the system states, inputs, and outputs, respectively; $\theta \in \mathbb{R}^{n_\theta}$ denotes the system parameters; $w_k \in \mathbb{R}^{n_w}$ denotes stochastic process noise; $v_k \in \mathbb{R}^{n_v}$ denotes measurement noise; and $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \times \mathbb{R}^{n_\theta} \to \mathbb{R}^{n_x}$ and $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \to \mathbb{R}^{n_y}$ are (possibly nonlinear) system state and output equations, respectively.[1] The uncertain initial states x_0 are described by the known probability distribution function (pdf) $\mathsf{P}[x_0]$. The random noise sequences $\{w_k\}$ and $\{v_k\}$ over different time instants are generated based on the known probability distributions $\mathsf{P}[w]$ and $\mathsf{P}[v]$, respectively. The random variables x_0, $\{w_k\}$, and $\{v_l\}$ are mutually independent on their respective probability space $(\Omega, \mathsf{F}, \mathsf{P})$ for all $k, l \geq 0$.[2]

When the parameters θ are known, the system (1) represents a Markov decision process (Kumar & Varaiya, 2016). Let $\mathcal{I}_k$ denote the vector of system information that is causally available at time k

$$\mathcal{I}_k := [y_k, \dots, y_0, u_{k-1}, \dots, u_0],$$

with $\mathcal{I}_0 := [y_0]$. Define *hyperstate* $\xi_{k|k}$ as the conditional probability of states x_k given $\mathcal{I}_k$, i.e., $\xi_{k|k} := \mathsf{P}[x_k|\mathcal{I}_k]$. Hyperstate describes

[1] The problem of model structure uncertainty is not considered in this paper; see Heirung and Mesbah (2017).

[2] The treatment of the stochastic optimal control problem in this paper does not consider temporally correlated stochastic noise. Non-white stochastic noise sequences can be handled through the application of a whitening filter (e.g., see Lewis, Xie, & Popa (2008)).