G Model
JJPC-2214; No. of Pages 10
ARTICLE IN PRESS
Journal of Process Control xxx (2017) xxx–xxx
Contents lists available at ScienceDirect
Journal of Process Control
journal homepage: www.elsevier.com/locate/jprocont

Real-time algorithm for self-reflective model predictive control

Xuhui Feng [a,b], Boris Houska [a,*]

[a] School of Information Science and Technology, ShanghaiTech University, China
[b] Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, China

A R T I C L E I N F O

Article history:
Received 13 September 2016
Received in revised form 6 October 2017
Accepted 10 October 2017
Available online xxx

Keywords:
Optimal control
Optimal experiment design
Model predictive control

A B S T R A C T

This paper is about a real-time model predictive control (MPC) algorithm for a particular class of model based controllers, whose objective consists of a nominal tracking objective and an additional learning objective. Here, the construction of the learning term is based on an economic optimal experiment design criterion. It is added to the MPC objective in order to excite the system on purpose thereby improving the accuracy of the state and parameter estimates in the presence of incomplete or noise affected measurements. The focus of this paper is on so-called self-reflective model predictive control schemes, which have the property that the additional learning term can be interpreted as the expected loss of optimality of the controller in the presence of random measurement errors. The main contribution is a formulation-tailored algorithm, which exploits the particular structure of self-reflective MPC problems in order to speed-up the online computation. It is shown that the proposed algorithm can solve the self-reflective optimization problems with reasonable additional computational effort and in real-time.

1. Introduction

The standard variant of model predictive control (MPC) [1,2] relies on the principle of *certainty equivalence*: at every sampling time a nominal control performance objective is optimized subject to dynamic model equations as well as control and state constraints on a finite prediction horizon. Here, the underlying assumption is that there are neither state estimation errors, nor external disturbances, nor any kind of model plant mismatches present, although all these errors and disturbances are the reason why a feedback controller is needed in the first place. After the MPC controller sends the first element of the optimized control input sequence to the real process, the next optimization problem is solved by using the latest state estimate in order to close the loop. The success of such certainty equivalent model predictive controllers, also in industrial applications [3], is to a large part due to the availability of fast and reliable real-time optimal control problem solvers [4,5]. During the last years algorithms as well as mature automatic code generation based software have been developed, which can solve nonlinear model predictive control problems online and within sampling times in the milli- and microsecond range [6,7].

Although one might argue that standard MPC and its more traditional variants do not attempt to achieve a tradeoff between nominal performance and learning objectives, many other controllers from the field of adaptive control, which implement such tradeoffs, have a longer history. One of the pioneers of the so-called dual control problem is A. Feldbaum, who published a whole series of papers on this topic [8]. Numerical algorithms for solving the dual control problem in higher dimensional spaces are often based on approximate dynamic programming, which, however, turn out to be rather expensive [9]. For an overview about other attempts to solve the dual control problem approximately by using techniques from the field of adaptive control the reader is referred to the overview articles [10,11].

As mentioned above, earlier or more traditional variants of MPC usually do not analyze learning objectives explicitly. However, during the last decade this situation has changed and, especially in recent years, there have appeared a significant number of articles about MPC variants that incorporate additional learning terms in order to achieve better future state and parameter estimates. For example, in [12] it is suggested to augment the standard MPC objective by an additional term that penalizes an approximation of the variance of future state estimates. This can be implemented by augmenting the model equations with an extended Kalman filter that can be used to predict the variance of future state and parameter estimates in a linear approximation. Similar extensions of MPC with learning terms have been proposed in [13,14], which augment the nominal MPC objective with optimal experiment design objectives that penalize the predicted variance of the system parameters. In recent years there have appeared a number of articles on

* Corresponding author.
E-mail address: borish@shanghaitech.edu.cn (B. Houska).

https://doi.org/10.1016/j.jprocont.2017.10.003
0959-1524/

Please cite this article in press as: X. Feng, B. Houska, Real-time algorithm for self-reflective model predictive control, J. Process Control (2017), https://doi.org/10.1016/j.jprocont.2017.10.003

persistently exciting MPC [15–19], which all discuss different ways to excite model predictive controllers in order to improve the accuracy of future state and parameter estimates. An even more recent trend is to extend the concept of application oriented optimal experiment design [20] by associated terms in the objective or constraints of an MPC problem [21,22]. Moreover, a recent paper on self-reflective model predictive control [23] proposes a model predictive controller that minimizes a prediction of its own expected loss of control performance in the presence of measurement noise.

A rather apparent drawback of all the above reviewed model predictive control schemes with additional learning objectives is that they are based on introducing additional, typically matrix-valued hyperstates, which are needed for predicting the accuracy of future state and parameter estimates. This leads to an optimal control problem that is from a numerical computation perspective much more difficult to solve than the corresponding optimal control problem without learning terms. Unfortunately, real-time algorithms, which exploit the structure of the model predictive control problems with additional learning terms, are not available to date—let alone implementations and software for solving these problems reliably. Therefore, a principal goal of this paper is to develop a real-time algorithm that can exploit the structure of such problems. Here, we focus on the self-reflective model predictive control formulation, which has been proposed in [23] and which is based on augmenting a nominal MPC objective with an economic optimal experiment design criterion [24].

Section 2 reviews selected MPC schemes, which include additional learning objectives. A particular focus is on self-reflective MPC and Theorem 1 summarizes the main theoretical properties of this self-reflective controller, which have been analyzed originally in [23]. In contrast to the theoretical developments in [23], the focus of the current paper is on the development of numerical algorithms and their implementation. Section 3.3 introduces a novel real-time algorithm for self-reflective MPC, which constitutes the main contribution of this paper. Theorem 3 establishes the corresponding theoretical contraction property of this real-time algorithm. Moreover, Section 4 illustrates the practical performance of the proposed scheme and discusses its advantages and disadvantages compared to more traditional MPC control algorithms. It is shown that the proposed algorithm can solve the self-reflective MPC problem within a sampling time that amounts to less than four times the sampling time of an associated certainty-equivalent real-time MPC scheme [4]. The corresponding numerical implementation is based on CASADI [25,26] and ACADO toolkit [6]. Section 5 concludes the paper.

1.1. Notation

We use the symbol $\mathbb{S}^n_+$ to denote the set of positive semi-definite $n \times n$ matrices. Similarly, $\mathbb{S}^n_{++}$ denotes the set of positive definite $n \times n$ matrices. Throughout this paper, the symbol $k \in \{0, \ldots, N-1\}$ is used as a running index, which, by convention, always runs from 0 to $N-1 \in \mathbb{N}$. For example, if we write a discrete-time system in the form

$$x_{k+1} = f(x_k, u_k),$$

then this means—if not explicitly stated otherwise—that this equation should hold for all $k \in \{0, \ldots, N-1\}$.

2. Model predictive control with learning objectives

2.1. Problem formulation

This paper concerns nonlinear control systems of the form

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

Here, $x_k \in \mathbb{R}^{n_x}$ denotes the state, $u_k \in \mathbb{R}^{n_u}$ the control, and $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}$ the right-hand side function. The variables $w_k \in \mathbb{R}^{n_x}$ denote the process noise.[1] Throughout this paper, the measurements are assumed to have the form

$$\eta_k = Cx_k + v_k, \tag{2}$$

where $C \in \mathbb{R}^{n_\eta} \times \mathbb{R}^{n_x}$ is a given matrix, $v_k \in \mathbb{R}^{n_\eta}$ the random measurement error, and $\eta_k \in \mathbb{R}^{n_\eta}$ the actual measurement. The following sections discuss the advantages and disadvantages of various model predictive controllers whose goal is to minimize a control performance objective of the form,

$$\sum_{k=0}^{N-1} l(x_k, u_k) + m(x_N)$$

on a finite horizon with given length $N \in \mathbb{N}$. Here, $l : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}$ denotes the stage cost and $m : \mathbb{R}^{n_x} \to \mathbb{R}$ the end cost. Throughout this paper, the following assumptions are used.

Assumption 1. We assume that the measurement errors v_k and the process noise w_k are pairwise uncorrelated random variables with bounded support, i.e., $\|v_k\| \le \gamma$ and $\|w_k\| \le \gamma$ for a given radius $\gamma > 0$.

Assumption 2. We assume that the first and second order moments of w_k and v_k, denoted by $W \in S^{n_x}_+$ and $V \in S^{n_\eta}_{++}$, are given,

$$\mathbb{E}\{w_k\} = 0, \ \mathbb{E}\{v_k\} = 0, \ \mathbb{E}\{w_k w_k^T\} = W, \ \mathbb{E}\{v_k v_k^T\} = V. \tag{3}$$

Assumption 3. We assume that the functions f, l, and m are at least three times differentiable in all arguments and all associated third derivatives are locally Lipschitz continuous.

Assumption 4. We have $f(0, 0) = 0$, $l(0, 0) = 0$, $m(0) = 0$, $l(x, u) \ge 0$, $m(x) \ge 0$, and the functions l and m are strongly convex.

In practice, the above control model needs to be augmented with additional control and state constraints, which can be modelled by adding suitable barrier functions to the function l, as for example discussed in the context of interior point based real-time algorithms for MPC [27,28,5]. As the focus of this paper is on the influence of the process noise and measurement errors on the MPC problem, we do not highlight such control- and state constraints in our notation throughout the theoretical developments, but our case study in Section 4 discusses how to deal with such constraints. Finally, Assumption 4 is made for simplicity of presentation and many of the results in this paper can be generalized for non-convex stage-costs. However, from a practical perspective, this condition is not excessively restrictive in the sense that in the context of MPC one often considers least-squares tracking-terms, which satisfy Assumption 4.

Remark 1. The statements in this paper can be extended for the more general case that the measurement function depends nonlinearly on x_k and u_k,

$$\eta_k = h(x_k, u_k) + v_k$$

for a two times Lipschitz continuously differentiable function h by formally replacing C with the state- and control dependent expression

$$C \longleftarrow \nabla_x h(x_k, u_k)^T.$$

[1] Notice that uncertain time-invariant system parameters can be modeled by introducing auxiliary states, which satisfy the trivial recursion $p_{k+1} = p_k$. Such time-invariant uncertainties should not be mixed up with the process noise w_k, which depends on the time-index k.

Please cite this article in press as: X. Feng, B. Houska, Real-time algorithm for self-reflective model predictive control, J. Process Control (2017), https://doi.org/10.1016/j.jprocont.2017.10.003