Automatica 61 (2015) 27–34

Contents lists available at ScienceDirect

Automatica

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

Optimal move blocking strategies for model predictive control[☆]

Rohan C. Shekhar[1], Chris Manzie

Department of Mechanical Engineering, The University of Melbourne, Victoria 3010, Australia

ARTICLE INFO

Article history:
Received 18 November 2014
Received in revised form
5 March 2015
Accepted 20 July 2015
Available online 13 August 2015

Keywords:
Move blocking
Model predictive control
Complexity reduction
Parametric optimisation

ABSTRACT

This paper presents a systematic methodology for designing move blocking strategies to reduce the complexity of a model predictive controller for linear systems, with explicit optimisation of the blocking structure using mixed-integer programming. Given a move-blocked predictive controller with a terminal invariant set constraint for stability, combined with an input parameterisation to preserve recursive feasibility, two different optimisation problems are formulated for blocking structure selection. The first problem calculates the maximum achievable reduction in the number of input decision variables and prediction horizon length, subject to the controller's region of attraction containing a specified subset of the state space. Then, for a given fixed horizon length and block count determined by hardware capabilities, the second problem seeks to maximise the volume of an inner approximation to the region of attraction. Numerical examples show that the resulting blocking structures are able to optimally reduce controller complexity and improve region of attraction volume.

1. Introduction

Model Predictive Control (MPC) is an automation paradigm that has been successfully applied to industrial control problems for a number of years (Maciejowski, 2002), owing to its intuitive formulation and unique constraint handling capability. At each time step, MPC aims to find a sequence of inputs to optimise a cost function over a finite time horizon, whilst satisfying operating constraints. With an appropriate selection of the cost function and constraints, recursive feasibility of the optimisation problem and subsequent convergence of the system states can be guaranteed (Mayne, Rawlings, Rao, & Scokaert, 2000).

Although advances in computational power have improved tractability of the MPC optimisation problem for an increasingly large number of systems, a long prediction horizon can prevent real-time MPC implementation, limiting its utility for systems with fast dynamics or high sampling rate requirements. Explicit MPC (Bemporad, Morari, Dua, & Pistikopoulos, 2002) seems to offer a solution by computing the MPC control law offline, but this computation becomes intractable for long horizon lengths and large numbers of constraints (Ferreau, Bock, & Diehl, 2008). Given this challenge, a number of alternative methods have been proposed to simplify the online complexity of MPC optimisation directly. Some approaches exploit the structure of the MPC problem to develop fast optimisation algorithms (Ferreau et al., 2008; Nedelcu, Necoara, & Tran-Dinh, 2014; Patrinos & Bemporad, 2014; Wang & Boyd, 2010), whilst others modify the MPC problem itself to reduce complexity at the expense of optimality. One method of achieving the latter is to assume some form of input parameterisation, curtailing the number of degrees of freedom in the online optimisation problem. Various candidate parameterisations have been proposed, including move blocking (Cagienard, Grieder, Kerrigan, & Morari, 2007; Maciejowski, 2002), linear subspaces (Goebel & Allgöwer, 2014; Ong & Wang, 2014) and Laguerre polynomials (Rossiter & Wang, 2008).

Move blocking is a candidate parameterisation that constrains groups of adjacent-in-time predicted inputs to have the same value. These groups are denoted as *blocks*, from where this parameterisation gets its name. Its advantages over other parameterisation methods include the parameterised input values retaining the same physical meaning as the original inputs, as well as the hardware implementation being straightforward. Whilst move blocking can reduce complexity, providing recursive feasibility guarantees is more challenging, since a "shifted" version of the previously optimal input sequence may no longer be admissible with respect to the blocking structure. In addition, blocking also affects

☆ This work was supported under the Australian Research Council's Discovery Projects funding scheme (Project DP120101830). The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Lalo Magni under the direction of Editor Ian R. Petersen.

E-mail addresses: rshekhar@unimelb.edu.au (R.C. Shekhar), manziec@unimelb.edu.au (C. Manzie).

[1] Tel.: +61 3 83447714.

http://dx.doi.org/10.1016/j.automatica.2015.07.030

the size of the controller's Region of Attraction (ROA). With a recursively feasible controller, the ROA corresponds to the set of states for which an initial feasible solution to the MPC optimisation problem exists.

Different move blocking approaches have been presented in order to maintain recursive feasibility guarantees and attempt to preserve the controller's ROA. Gondhalekar and Imura (2010) introduce an approach based upon calculating an invariant set that is also blocking admissible. Additionally, the approach recovers the ROA of the original unblocked controller by relaxing constraints at future prediction steps. However, this prevents a terminal constraint from being enforced, so explicit guarantees of closed-loop convergence cannot be given. Cagienard et al. (2007) and Shekhar and Maciejowski (2012b) present approaches that utilise time-varying blocking structures, where the changing structure allows a shifted version of the previously optimal input sequence to be feasible at the following time step. Guarantees of both recursive feasibility and closed-loop convergence can then be provided, with an appropriately designed terminal constraint and cost function. However, the ROA of the controller is necessarily reduced by this constraint, making blocking structure selection an important design consideration.

Blocking structure design can be approached from the perspective of satisfying one of two specifications: either a requirement on minimum ROA size, or a fixed complexity requirement specified in terms of the number of blocks (or block count) and the horizon length. If the ROA is required to contain a given subset of the feasible state space, then the amount by which complexity can be reduced through blocking is restricted. In such cases, reducing the number of blocks and the horizon length by the greatest extent possible whilst ensuring the ROA contains this subset allows for faster computation, reduces the hardware footprint and provides opportunities for performance improvements through parallelisation (Longo, Kerrigan, Ling, & Constantinides, 2011).

Conversely, if a controller is to be designed to utilise the maximum capability of a given hardware platform, then an effective limit is placed on the horizon length and number of blocks for a fixed sampling rate. In this scenario, it is desirable to choose a blocking structure that possesses the largest ROA size for this horizon length and block count, especially for regulation problems. This allows the controller to regulate from a large set of initial states whilst making maximum use of hardware capabilities. However there are currently no systematic methodologies to select blocking structures for either of these design objectives.

This paper addresses this research gap by presenting methods for optimally selecting blocking structures in a computationally tractable manner. It first details how a move-blocked predictive controller is formulated, with terminal constraints used to guarantee convergence and a specific input parameterisation that ensures recursive feasibility, similar to that used by Ong and Wang (2014). This parameterisation allows the blocking structure to remain time invariant, by adding to the input an appropriately shifted scalar multiple of the previously optimal input sequence. Two different optimisation problems are then formulated for selecting blocking structures, which form the primary contribution of this paper:

(1) minimising the number of input blocks and the horizon length, whilst ensuring that the ROA contains a desired subset of the state space; and
(2) maximising the volume of an ellipsoidal inner approximation to the ROA, for a specified horizon length and number of blocks.

Finally, numerical examples illustrate the approaches on simple two-dimensional systems, which allow ROAs to be easily visualised and their volumes explicitly computed.

1.1. Nomenclature

The sets of integers and real numbers are denoted $\mathbb{Z}$ and $\mathbb{R}$ respectively. $\mathbb{Z}_{[a,b]}$ denotes the set of numbers $\{a, a+1, \dots, b\}$, where $a, b \in \mathbb{Z}$, $a \le b$. The set of all non-negative integers is denoted $\mathbb{Z}_{\ge 0}$. The vertical concatenation $[X_1^T, X_2^T, \dots, X_N^T]^T$ is written as $[X_1; X_2; \dots; X_N]$. I_n represents the $n \times n$ identity matrix. $\mathbf{0}_n$ represents a vector of zeros having length n, whereas $\mathbf{0}$ is a block matrix of zeros with dimensions to be inferred from context within a larger block matrix. $\mathbf{1}_n$ represents a vector of ones having length n, whereas $\mathbf{1}_{n\times m}$ denotes an $n \times m$ matrix of ones. $\bar{z}_{j|k}$ denotes a prediction of signal $z(k)$ made j-steps in to the future from the current time k. The operator $\otimes$ denotes the Kronecker product. The operator $\oplus$ represents the direct sum. The operator $\mathrm{co}\{\cdot\}$ denotes the convex hull. The volume of a set is denoted $\mathrm{vol}(\cdot)$. $\mathbb{P}_n(\cdot)$ denotes a projection onto $\mathbb{R}^n$. $\|\cdot\|_p$ denotes the p-norm, whereas $|\cdot|$ denotes the element-wise absolute value. The operator $\vee$ denotes an element-wise logical disjunction.

2. Problem formulation

Consider the discrete-time linear system

$$x(k+1) = Ax(k) + Bu(k), \tag{1}$$

where $x(k) \in \mathbb{R}^n$ and $u(k) \in \mathbb{R}^m$. It is assumed that (A, B) is controllable. Defining an output

$$y(k) := Cx(k) + Du(k) \in \mathbb{R}^p, \tag{2}$$

the system is subject to constraints at each sampling instant of the form

$$y(k) \in \mathcal{Y} := \{y \mid Ey \le f\}, \tag{3}$$

for some $E \in \mathbb{R}^{s\times p}$ and $f \in \mathbb{R}^s$ that define polytope $\mathcal{Y}$. The control objective is to asymptotically steer the state of the system to the origin from the initial state $x(0)$, whilst satisfying the constraints (3) at all times. The output equation (2) is chosen by the control designer to incorporate all state, input and cross-constraints.

A model predictive controller is specified to steer the state of the system (1) to the origin, whilst minimising a given finite-horizon cost function, subject to the constraints (3). At each iteration, the MPC optimisation problem takes the form

$$J_N^*(x(k)) := \min_{\bar{\boldsymbol{u}}(k)} J_N(\bar{\boldsymbol{u}}(k), x(k)), \tag{4}$$

subject to

$$\bar{x}_{0|k} = x(k) \tag{5a}$$
$$\bar{x}_{j+1|k} = A\bar{x}_{j|k} + B\bar{u}_{j|k} \tag{5b}$$
$$\bar{y}_{j|k} = C\bar{x}_{j|k} + D\bar{u}_{j|k} \tag{5c}$$
$$\bar{y}_{j|k} \in \mathcal{Y} \tag{5d}$$
$$\bar{x}_{N|k} \in \mathcal{T}, \tag{5e}$$

for all $j \in \mathbb{Z}_{[0,N-1]}$, given prediction horizon N, input prediction variables

$$\bar{\boldsymbol{u}}(k) := \left[\bar{u}_{0|k}; \bar{u}_{1|k}; \dots; \bar{u}_{N-1|k}\right],$$

state prediction variables $\bar{x}_{0|k}, \bar{x}_{1|k}, \dots, \bar{x}_{N|k}$, output predictions $\bar{y}_{0|k}, \bar{y}_{1|k}, \dots, \bar{y}_{N-1|k}$ and cost function $J_N(x, \bar{\boldsymbol{u}})$. For stability purposes, a terminal polytopic constraint-admissible control-invariant set

$$\mathcal{T} := \{x \mid Gx \le h\}$$

is defined, for some $G \in \mathbb{R}^{t\times n}$ and $h \in \mathbb{R}^t$. The invariance of this set implies that for all $x \in \mathcal{T}$, there must exist a μ such that

$$Ax + B\mu \in \mathcal{T} \tag{6a}$$
$$Cx + D\mu \in \mathcal{Y}. \tag{6b}$$