BILEVEL PROGRAMMING FOR ANALYSIS OF REDUCED MODELS FOR USE IN MODEL PREDICTIVE CONTROL

Henrik Manum and Sigurd Skogestad

Dept. of Chemical Engineering, Norwegian University of Science and Technology,
N-7491 Trondheim, Norway,
e-mail: skoge@ntnu.no

Abstract: In this paper we use bilevel programming to find the maximum difference between a model predictive controller (MPC) using a full model and an MPC using a reduced model. The results apply to MPC with quadratic cost function and linear model with linear constraints.

Keywords: model reduction, bilevel programming, linear MPC

1 INTRODUCTION AND PRELIMINARIES

Consider the linear system

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + B_d d_k, \quad k \in \{0,1,2,\dots\} \\ y_k &= Cx_k + Du_k + D_d d_k \end{aligned} \tag{1}$$

with constraints

$$x_k \in \mathbf{X} \subset \mathbb{R}^{n_x},\ y_k \in \mathbf{Y} \subset \mathbb{R}^{n_y},\ u_k \in \mathbf{U} \subset \mathbb{R}^{n_u}, \tag{2}$$

where x_k note states, y_k are measurements, u_k are controlled inputs, and d_k are disturbances. Further $\mathbf{X}, \mathbf{Y}, \mathbf{U}$ are polytopes.

In addition we have a "reduced" model

$$\begin{aligned} x_{k+1}^{\text{red}} &= A^{\text{red}} x_k^{\text{red}} + B^{\text{red}} u_k + B_d^{\text{red}} d_k, \quad k \in \{0,1,2,\dots\} \\ y_k^{\text{red}} &= C^{\text{red}} x_k^{\text{red}} + D^{\text{red}} u_k + D_d^{\text{red}} d_k \end{aligned} \tag{3}$$

where $x_k^{\text{red}} \in R^{n_{x^{\text{red}}}}$ with $n_{x^{\text{red}}} < n_x$. The reduced model (3) is assumed to be found by some model reduction scheme such as balanced truncation, balanced residualization or optimal Hankel norm reduction [Skogestad and Postlethwaite, 2005].

In this paper we consider model predictive control (MPC) [Mayne et al., 2000]. The question we want to answer is: *What is the worst-case difference between an MPC using the full model* (1) *and an MPC using the reduced model* (3)*?* This question is related to the problem of "closed loop analysis of reduced order models for use in MPC".

An important feature of MPC is its possibility to handle constraints. However, if there are no constraints, several methods exists to analyse the performance of control based on the reduced order model. In time domain one may consider measures such as rise-time, settling time, overshoot, decay ratio, steady state offset and total variance [Skogestad and Postlethwaite, 2005]. In frequency domain one may consider gain and phase margins but also peaks on sensitivity functions as a more general measure. These methods are however mostly limited to single input single output systems (SISO). In the general case of a multiple input multiple output (MIMO) plant we recommend to use robust stability and performance, through the μ- analysis, as discussed in detail by Skogestad and Postlethwaite [2005].

Hovland et al. [Hovland et al., 2006, Hovland and Gravdahl, 2008] proprose a scheme to use reduced model in explicit MPC. They perform a two-step procedure to analyse the reduced order controller: First, they analyse the model reduction using an open loop evaluation of the model mismatch (evaluated by the $\mathscr{H}_2$-norm). Then, they make a table of model order and resulting number of regions, and choose the model order that gives a satisfactory low number of regions combined with a low model mismatch.

In this paper we evaluate the performance of the reduced-order controller by addressing the following problem:

$$\begin{aligned} &\max_{d\in\mathscr{D}} \text{distance}(u_k, u_k^{\text{red}}) \\ \text{subject to } &u_k = \arg\min\{\text{MPC forumlation with full model}\} \\ &u_k^{\text{red}} = \arg\min\{\text{MPC formulation with reduced model}\} \end{aligned} \tag{4}$$

The goal of problem (4) is to find the maximum difference between the full-order controller and the low-order controller. Note that we do not use an explicit formulation of the controllers, rather we simply express them as solutions to optimization problems. We will show that problem (4) can be rewritten as a mixed-integer linear program (MILP) and solved using standard software.

Remark 1 *In this paper we treat the distance between the controllers as* $\|u_k - u_k^{\text{red}}\|_\infty$. *However, we could also have used difference in outputs, i.e.* $\|Q_y(y_k - y_k^{\text{red}})\|_\infty$, *or a combination of both. We use the infinity norm* $\|\cdot\|_\infty$ *because then the problem can be reformulated as an MILP.*

1.1 Notation and assumptions

We use "full-order controller" to indicate an MPC based on the full model (1) and "low-order controller" for MPC based on the reduced model (3).

In this paper we follow the normal way of letting the initial state x_0 represent the disturbances, i.e. in the following we do not consider the effect of B_d and C_d as they appear in model (1). This is mostly to ease the presentation, but in general we recommend to keep the disturbances d_k in the problem formulation.

1.2 Organization of the paper

We first review a model reduction technique that we later will use in an example. This gives a map $x_k^{\text{red}} = T_l x_k$ which represents the model reduction. We then review how bilevel optimization problems can under some assumptions be reformulated to MILP problems, and thereafter show how the linear quadratic MPC fits into this framework, and finally how we can formulate problem (4) as an MILP.

2 MODEL REDUCTION BY BALANCED TRUNCATION

We here review model reduction by balanced truncation [Moore, 1981] as an example of a model reduction scheme that can be analyzed with the proposed method. We follow Dones et al. [2010].

Consider a continuous linear system

$$\begin{gathered} \dot{x} = A^c x + B^c u, \quad y = C^c x + D^c u \\ x \in \mathbb{R}^{n_x},\ y \in \mathbb{R}^{n_y},\ u \in \mathbb{R}^{n_u}. \end{gathered} \tag{5}$$

The model reduction by balanced truncation consists of two steps: First, we find a balanced representation of system (5), then we remove the states corresponding to the smallest Hankel singular values of the balanced representation.

2.1 Balanced representation

The controllability and observability gramians of a linear system are defined as

$$A^c W_c + W_c A^{c\prime} + B^c B^{c\prime} = 0 \tag{6}$$

$$A^{c\prime} W_o + W_o A^c + C^{c\prime} C^c = 0 \tag{7}$$

A balanced representation of system (5) is obtained through a transformation matrix T, such that $\bar{W}_c$ and $\bar{W}_o$ (of the transformed system) are equal. Let z denote the states of the balanced system, i.e. $z = Tx$. It can be shown that

$$\begin{aligned} \bar{W}_c = \bar{W}_o = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_{n_x}) \\ \bar{W}_c = T W_c T^{-1} \\ \bar{W}_o = (T^{-1})' W_o T^{-1} \end{aligned} \tag{8}$$

where σ_i, $k = 1, 2, \dots, n_x$ are the Hankel singular values of the balanced representation, ordered according to

$$\sigma_1 > \sigma_2 > \cdots > \sigma_{n_x} \geq 0.$$

2.2 Truncation

Let $z' = [z_1' \; z_2']$. In balanced truncation we simply delete z_2 from the vector of balanced states z.

Denote T_l and T_r as

$$T = \begin{bmatrix} \overbrace{\begin{bmatrix} T_{11} & \dots & T_{1n} \\ \vdots & & \vdots \\ T_{\tilde{n}1} & \dots & T_{\tilde{n}n} \end{bmatrix}}^{T_l} \\ \begin{matrix} \vdots & \quad & \vdots \\ T_{n1} & \dots & T_{nn} \end{matrix} \end{bmatrix}, \quad T^{-1} = \begin{bmatrix} \underbrace{\begin{bmatrix} T^{-1}_{11} & \dots & T^{-1}_{1\tilde{n}} \\ \vdots & & \vdots \\ T^{-1}_{n1} & \dots & T^{-1}_{n\tilde{n}} \end{bmatrix}}_{T_r} \begin{matrix} \dots & T^{-1}_{1n} \\ & \vdots \\ \dots & T^{-1}_{nn} \end{matrix} \end{bmatrix} \tag{9}$$

We can now express the balanced and truncated result as

$$\begin{aligned} \dot{z}_1 &= T_l A^c T_r + T_l B^c u \\ \bar{y} &= C^c T^r z_1 + Du, \end{aligned} \tag{10}$$

and we note that the map from the full state vector x to the balanced and truncated system (10) is given by $z_1 = T_l x$.

3 BILEVEL OPTIMIZATION

Here we give an introduction to bilevel optimization and solution methods, following Jones and Morari [2009]. For more background details the reader is referred to a recent survey [Colson et al., 2005].

Bilevel problems are hierarchical in that the optimization variables (y,z) are split into upper y and lower z parts, with the lower level variables constrained to be an optimal solution to a secondary optimization problem:

$$\begin{aligned} &\min_{y} \; V_U(y,z) \\ \text{subject to } &G_U(y,z) \leq 0 \\ &z = \arg\min_{z} V_L(y,z) \\ &\qquad \text{subject to } G_L(y,z) \leq 0 \end{aligned} \tag{11}$$

In this paper we will only consider problems where the lower-level problem has an unique optimizer. Moreover, we will have two low-level problems, one for the full-order controller and one for the low-order controller.

3.1 Solution methods

If the lower level problem is convex and regular, then it can be replaced by its necessary and sufficient Karush-Kuhn-Tucker (KKT) conditions, yielding a standard single-level optimization problem:

$$\begin{aligned} &\min_{y,z,\lambda} \; V_U(y,z) \\ \text{subject to } &G_U(y,z) \leq 0 \\ &G_L(y,z) \leq 0 \\ &\lambda \geq 0 \\ &\lambda' G_L(y,z) = 0 \\ &\nabla_z \mathscr{L}(y,z,\lambda) = 0 \end{aligned} \tag{12}$$

where $\mathscr{L}(y,z,\lambda) := G_L(y,z) + \lambda' G_L(y,z)$ is the Lagrangian function associated with the lower-level problem. For the special case of linear constraints and a quadratic cost, all constraints of (12) are linear and the complimentary condition $\lambda' G_L(y,z) = 0$ is a set of disjunctive linear constraints, which can be described using binary variables, and thus leads to a mixed-integer linear problem.

4 APPLICATION TO ANALYSIS OF MPC CONTROLLERS

4.1 MPC formulation

Consider the following semi-infinite horizon optimal control problem [Jones and Morari, 2009]:

$$\begin{aligned}\min_{\mathbf{x},\mathbf{u}} \; J(\mathbf{x},\mathbf{u}) &= \frac{1}{2}x_N' P x_N + \frac{1}{2}\sum_{i=0}^{N-1} u_i' R u_i + x_i' Q x_i,\\ \text{subject to } x_{i+1} &= A x_i + B u_i, \quad \forall i = 0,\dots,N-1,\\ x_i &\in \mathbf{X}, \quad \forall i = 1,\dots,N-1,\\ u_i &\in \mathbf{U}, \quad \forall i = 0,\dots,N-1,\\ x_N &\in \mathbb{X}_N,\\ x_0 &= x.\end{aligned} \tag{13}$$

Here $\mathbb{X}_N = \{x \mid Hx \le h\} \subset \mathbb{X}$ is a polytopic invariant set for the system $x^+ = Ax + B\mu(x)$ for some given control law $\mu : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_u}$. Further $P \in \mathbb{R}^{n_x \times n_x}$ and $Q \in \mathbb{R}^{n_x \times n_x}$ are positive definite matrices and $R \in \mathbb{R}^{n_u \times n_u}$ is a positive semi-definite matrix. We define $\mathscr{X} \subset \mathbb{R}^{n_x}$ to be the set of states x for which there exists a feasible solution to (13).

If $\mathbf{u}^*(x)$ is the optimal input sequence of (13) for the state x, and $u_0^*(x)$ is the resulting control law, then stability of the system $x^+ = Ax + Bu_0^*(x)$ can be established under the assumption that $V_N(x) = x'Px$ is a Lyapunov function for the system $x^+ = Ax + B\mu(x)$ and that the decay rate of V_N is greater than the stage cost $l(u,x) = u'Ru + x'Qx$ within the set $\mathbb{X}_N$.

By using $x_k = A^k x_0 + \sum_{j=0}^{k-1} A^j B u_{k-1-j}$ the MPC problem (13) can be rewritten as [Bemporad et al., 2002]:

$$V(x_0) = \frac{1}{2}x_0' Y x_0 + \min_U \{\frac{1}{2}U'HU + x_0' F U, \quad \text{subject to } \; GU \le W + E x_0\}, \tag{14}$$

where $U' = \begin{bmatrix} u_0' \; u_1' \; \cdots \; u_{N-1}' \end{bmatrix}$.

We want to use (14) as a lower-level problem in bilevel programming. The following equations define the KKT conditions for this problem:

$$\begin{aligned} HU + F'x_0 + G'\lambda &= 0\\ GU - W - Ex_0 &\le 0\\ \lambda &\ge 0\\ \lambda &\le Ms\\ GU - W - Ex_0 &\ge -M(1-s)\end{aligned} \tag{15}$$

Here $s \in \{0,1\}^{n_W}$, where n_W is the number of inequality constraints in (14). The two last equations in (15) correspond to the complementary condition $\lambda' G_L(y,z) = 0$ in the general bilevel problem, here described with binary variables s. M is a constant that is large enough such that the solution to (15) corresponds to the solution of (14). (This is called a "big-M" formulation.)

4.2 First input analysis problem

Let $(H^{\text{full}}, F^{\text{full}}, G^{\text{full}}, W^{\text{full}}, E^{\text{full}})$ correspond to (H,F,G,W,E) in (15) for an MPC using the full-order model (1) and further let $(H^{\text{red}}, F^{\text{red}}, G^{\text{red}}, W^{\text{red}}, E^{\text{red}})$ be the corresponding matrices to the reduced-order model (3).

Further we denote $\mathrm{KKT}(\mathrm{MPC}^{\mathrm{full}}, x_0^{\mathrm{full}})$ as the set of equations (15) evaluated at $(H, F, G, W, E) = (H^{\mathrm{full}}, F^{\mathrm{full}}, G^{\mathrm{full}}, W^{\mathrm{full}}, E^{\mathrm{full}})$ and $x_0 = x_0^{\mathrm{full}} \in \mathbb{R}^{n_x}$. Correspondingly we let $\mathrm{KKT}(\mathrm{MPC}^{\mathrm{red}}, x_0^{\mathrm{red}})$ describe the KKT-conditions in equations (15) for the reduced order controller, with $x_0 = x_0^{\mathrm{red}} \in \mathbb{R}^{n_x^{\mathrm{red}}}$.

We define the one-step problem as:

$$\begin{aligned} &\max_{x \in \mathscr{X}} \| B(U^{\mathrm{full}}(1:n_u) - U^{\mathrm{red}}(1:n_u)) \|_\infty \\ \text{subject to } &\mathrm{KKT}(\mathrm{MPC}^{\mathrm{full}}, x) \\ &\mathrm{KKT}(\mathrm{MPC}^{\mathrm{red}}, T_l x) \end{aligned} \tag{16}$$

The notation $U^q(1:n_u)$, $^q \in \{^{\mathrm{full}}, ^{\mathrm{red}}\}$ means the first n_u elements of the vector U^q. This is the input from the MPC that is actually implemented in the plant.

The polytope $\mathscr{X}$ is the search space for the MILP. This can either be the set of feasible initial states for the full-order controller, or a set of initial states that the engineer find interesting.

Using the reformulations show earlier this can be rewritten as an MILP.

Remark 2 *We observe that the objective function renders* (16) *non-convex due to the term* $\max \|t\|_\infty$ *(where t is a convex function of* (u, u^{red})*) . However, the problem may be converted into a mixed integer linear program (MILP) using a standard technique (e.g. [Löfberg, 2004]), in which we introduce binary variables* n_i, p_i *for each element of t and add the condition that the binary variable* p_i *is one if* $\|t\|_\infty = t_i$ *and* n_i *is one if* $\|t\|_\infty = -t_i$*. The method adds only linear and binary conditions to* (16) *and therefore the overall problem remains a MILP [Jones and Morari, 2009].*

Remark 3 *We evaluate the input difference in the direction B because this is the direction a "wrong" input (due to the reduced model) will influence the states.*

5 EXAMPLE: DISTILLATION

We here consider MPC for "column A" by Skogestad [1997]. This is a 82-state nonlinear model which we linearize around a nominal operating point and discretize with sample time $T_s = 1$. The model has two inputs (reflux and boilup) and two output (mole fractions in the top and bottom of the column). The model of the column is available on Prof. Skogestads homepage (google "`skogestad`").

The physical meaning of inputs and outputs is not important for the purpose of displaying the methods in this paper, hence we will simply treat them as generic variables

$$u_k \in \mathbb{R}^2, \quad y_k \in \mathbb{R}^2, \quad k = 0, 1, 2, \ldots. \tag{17}$$

In order to simplify calculations we first reduce the linearized model to 16 states, and we consider this to be our base case. Using balanced truncation, as described in section 2, we generate a set of models consisting of 1 to 15 states.

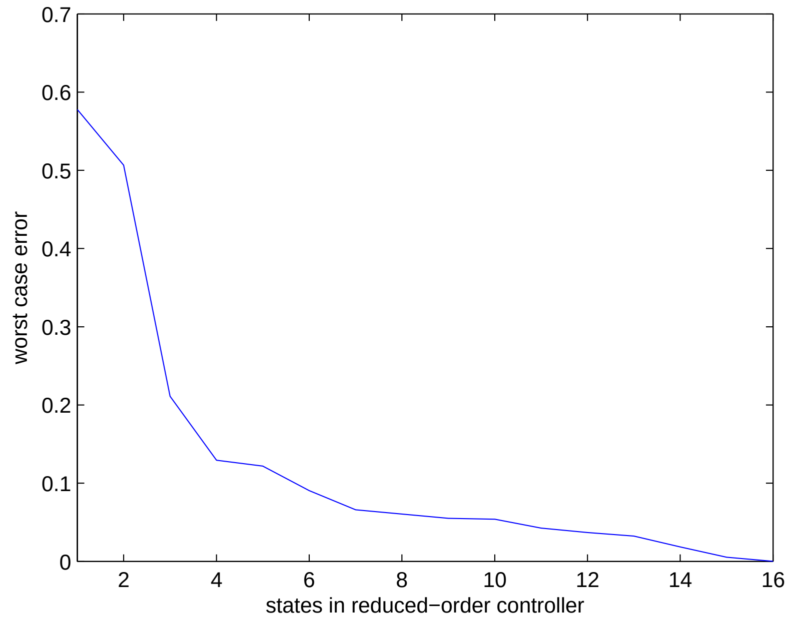

Figure 1: Solutions for a set of different reduced order models.

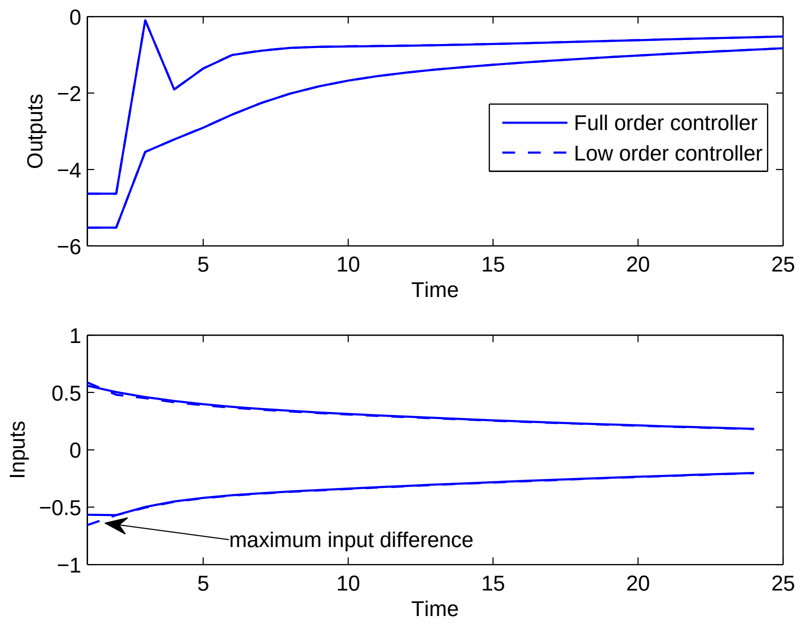

Figure 2: Closed loop simulation for the full order controller ($n_{\text{full}} = 16$) and low-order controller with $n_{\text{red}} = 6$.

The MPC problem we consider is the following one:

$$\begin{aligned}\min_{u_0,\ldots,u_7} & \; y_8' y_8 + \sum_{i=0}^{7} y_i' y_i + u_i' u_i \\ \text{subject to } & x_{k+1} = A x_k + B u_k, \quad k = 0,1,\ldots,7 \\ & y_k = C x_k, \quad k = 0,1,\ldots,7 \\ & -\mathbf{1} \le u_k \le \mathbf{1}, \quad k = 0,1,\ldots,7. \end{aligned} \tag{18}$$

Here $\mathbf{1}' = [1\ 1]$. The only difference between the MPC using full order model and MPC using reduced order model is the internal model represented by the matrices (A, B, C) and the dimension of the state vector x_k.

Thereafter we use problem (16) to calculate the maximum difference $\|B(u^{\text{red}} - u^{\text{full}})\|_\infty$ applied to the plant. The search space $\mathscr{X}$ is given by the box constraint

$$\mathscr{X} = \left\{ x \in \mathbb{R}^{16} \;\middle|\; \|x\|_\infty \le 10 \right\}. \tag{19}$$

The resulting differences are shown in figure 1. In figure 2 we show a closed loop simulation for a reduced controller using 6 internal states together with the controller using 16 states. The system we are simulating is the 16-state system. We start the simulation from the worst possible initial state, which in this case was

$$x_0 = 10 \cdot [-1 \quad 1 \quad 1 \quad -1 \quad 1 \quad 1 \quad -1 \quad 1 \quad 1 \quad -1 \quad -1 \quad -1 \quad 1 \quad 1 \quad 1 \quad -1]'. \tag{20}$$

We observe that for this initial point the maximum input difference is 0.0905, which is in agreement with figure 1. Further, the performance of the two controllers are very similar in terms of bringing the system from this initial state to the reference point $x = 0$.

We used Yalmip [Löfberg, 2004] under Matlab™ to set up the optimization model and GLPK[1] to solve the problem. For this example, using a PC with 2000 MHz CPU with 2 GB memory, it took about 1-2 seconds to solve each problem.

6 DISCUSSION

Infinity norm We used the infinity norm in the calculation of the difference between the full- and the reduced- order controller. Using this norm we can rewrite the problem as an MILP, and hence renders the problems solvable. Further this norm should be a natural norm in order to evaluate the *maximum* difference between two functions, as discussed by Jones and Morari [2009].

Implicit representation of MPC We use the KKT conditions to describe the MPC controllers, as the solution to the KKT system contains the optimal input from the MPC controller. An alternative method could be to find the MPC controllers explicitly by solving a parametric program [Kvasnica et al., 2004]. However, this would be a lot more complicated (need a binary variable for each region in the MILP formulation), and for the example discussed in this report (16 states with an input horizon of 8) it would most likely take a very long time even to find an explicit representation of the controller.

[1] can be found at `http://www.gnu.org/software/glpk/`

Planned activity In the near future we want to investigate different ways of solving this problem, rather than formulating it as an MILP [Bard, 1998]. In addition we want to extend the method to check trajectories of inputs, rather than only the first input. This can be done simply by stacking problems on the type of the "first-input difference" in equation (16).

Perhaps more interesting is that rather than treating the initial state vector as a disturbance, we include only disturbances that have a physical meaning (i.e. actual disturbances to the system). This should fit into the formulation quite easily as long as the disturbances enter linearly as is the case for system (1).

In a recent paper [Manum et al., 2009] we tried to use the same problem as described in this paper in order to show nominal stability of a low-complexity controller by comparing it to a robust MPC for the same system. The same methodology could be used to prove stability of MPC with a reduced order model.

7 CONCLUSIONS

An MILP framework for analysis of closed-loop performance of MPC using reduced order model has been presented. The method was demonstrated on a 16-state linear system.

Acknowledgements

The authors acknowledge the fruitful discussions with Dr. Colin N. Jones.

References

J.F. Bard. *Practical Bilevel Optimization*. Kluwer Academic Publishers, Dodrecht/Boston/London, 1998.

A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos. The explicit linear quadratic regulator for constrained systems. *Automatica*, 38:3–20, 2002. See also corrigendum 39(2003), pages 1845-1846.

B. Colson, P. Marcotte, and G. Savard. Bilevel programming: A survey. *4OR: A Quarterly Journal of Operations Research*, 3(2):87–107, June 2005.

I. Dones, S. Skogestad, and H.A. Preisig. Application of balanced truncation to nonlinear systems. Submitted to Computers and Chemical Engineering, 2010.

S. Hovland and J.T. Gravdahl. Complexity Reduction in Explicit MPC through Model Reduction. In *Proceedings of the 17th World Congress*, pages 7711–7716, Seoul, Korea, July 2008. The International Federation of Automatic Control.

S. Hovland, K. Willcox, and J.T. Gravdahl. MPC for Large-Scale Systems via Model Reduction and Multiparametric Convex Programming. In *Proceedings of the IEEE Conference on Decision and Control*, San Diego, USA, Dec 2006.

C.N. Jones and M. Morari. Approximate Explicit MPC using Bilevel Optimization. In *European Control Conference*, Budapest, Hungary, August 2009. URL `http://control.ee.ethz.ch/index.cgi?page=publications;action=details;id=3339`.

M. Kvasnica, P. Grieder, and M. Baotić. Multi-Parametric Toolbox (MPT), 2004. URL `http://control.ee.ethz.ch/ mpt/`.

J. Löfberg. Yalmip : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. URL `http://control.ee.ethz.ch/ joloef/yalmip.php`.

H. Manum, C.N. Jones, J. Löfberg, M. Morari, and S. Skogestad. Bilevel programming for analysis of low-complexity control of linear systems with constraints. In *Conference on Decision and Control*, pages 946–951, Shanghai, China, 2009.

D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36:789–814, 2000.

B. Moore. Principal Component Analysis in Linear Systems: Controllability, Observability and Model Reduction. *IEEE Transactions on Automatic Control*, 26(1):17–32, 1981.

S. Skogestad. Dynamics and control of distillation columns - a tutorial introduction. *Trans IChemE, Part A*, 75:539–562, September 1997.

S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control*. Wiley, 2005.