Proceedings of the American Control Conference
Chicago, Illinois • June 2000

Triple Mode Control in MPC

J.A. Rossiter
Department of Mathematical Sciences
Loughborough University, Loughborough
LE11 3TU, UK
j.a.rossiter@lboro.ac.uk

B. Kouvaritakis, M. Cannon
Department of Engineering Science
Oxford University, Oxford
OX1 3PJ, UK
basil.kouvaritakis@eng.ox.ac.uk

Abstract

Dual mode control with ellipsoidal invariant sets leads to efficient implementations of MPC. Despite their convenience, ellipsoidal sets are not maximal in volume and may result in small stabilizable sets. Here we achieve an enlargement of these sets at a small additional computational load by front-ending the application of dual mode by a third mode which uses polytopic rather than ellipsoidal constraints.

1 Introduction

Dual mode control [8, 10] prescribes the use of: (i) n free control moves followed by (ii) a fixed term control law, $\mathbf{u} = -K\mathbf{x}$. The combined use of these two modes in Model Based Predictive control (MPC) provides a guarantee of closed-loop stability. In the interest of optimality, K should be highly tuned, but feasibility constraints arising out of limits on actuation (and/or states), may require the use of large horizons n, thereby increasing the complexity of on-line computation, i.e. the on-line solution of a Quadratic Program (QP).

The trade-off between optimality and computational efficiency has been the object of recent research. For example when the latter is paramount it is possible to reduce the degrees of freedom in predictions to two [5] or one [6] while still retaining good performance by interpolating between prescribed predicted input trajectories. Other recent work [7, 4] has shown that, through an alternative use [9] of dual mode control in conjunction with ellipsoidal feasible invariant sets, it is possible to: (i) avoid the need for QP; (ii) reduce dramatically the complexity of on-line optimization in a manner that is independent of n.

Despite their convenience, feasible invariant ellipsoidal sets are necessarily subsets of the maximal admissible set [1] (ie. the largest set of initial states for which stability can be guaranteed) which, for a linear system subject to linear input/state constraints is polytopic. Consequently ellipsoidal stabilizable sets may be overly conservative. It is the purpose of the present paper to propose a "triple mode" control strategy by front-ending the "dual mode" paradigm by a "third mode" involving m control moves which are subject to polytopic rather than ellipsoidal constraints. The incorporation of dual mode control in the new strategy implies that m can be kept small, thereby limiting on-line computational complexity. Given that constraints are usually active only during transients, the insertion of the "third mode" has potentially a very significant effect on the size of the stabilizable sets.

The structure of the paper is as follows. Section 2 gives a brief description of background material and notation. The "triple mode" control strategy is outlined in Section 3 and is followed by comparisons with existing algorithms in Section 4. Conclusions together with a discussion of further improvements in terms of computational complexity are discussed in Section 5.

2 Notation and concepts

Consider the discrete-time model

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k, \tag{1}$$

$\mathbf{x} \in \mathbb{R}^p$, $\mathbf{u} \in \mathbb{R}^q$. We denote the usual LQ cost as

$$J = \sum_{k=0}^{\infty} \mathbf{x}_k^T Q \mathbf{x}_k + \mathbf{u}_k^T R \mathbf{u}_k, \tag{2}$$

and define the associated optimal control by

$$\mathbf{u} = -K\mathbf{x} \tag{3}$$

Under (3) we have $J = \mathbf{x}_0^T P \mathbf{x}_0$, and the closed-loop transition matrix is given by $\Phi = A - BK$. Input and state constraints are denoted

$$C_u u_k - \mathbf{d}_u \leq 0, \quad C_\mathbf{x} \mathbf{x}_k - \mathbf{d}_\mathbf{x} \leq 0; \quad k = 0, 1, \ldots \tag{4}$$

and yield the maximal admissible set [1] for (1,3,4):

$$T_0 = \{\mathbf{x} : C_{\max}\mathbf{x} - \mathbf{d}_{\max} \leq 0\}. \tag{5}$$

The subscript "0" indicates that the problem formulation so far does not involve any degrees of freedom, ie. $n = 0$.

2.1 Predictive control

Consider the dual mode control strategy (eg. [10]) which defines future control moves via (3) but allows the first n moves to be free and selects them so as to minimize the cost (2) subject to constraints (4). This minimization problem can be reformulated as:

$$\begin{array}{l} \displaystyle\min_{\mathbf{c}_k, k<n} J = \sum_{k=0}^{\infty} \mathbf{x}_k^T Q \mathbf{x}_k + \mathbf{u}_k^T R \mathbf{u}_k \\ \text{s.t.} \left\{ \begin{array}{l} \mathbf{u}_k = -K\mathbf{x}_k + \mathbf{c}_k,\ k < n \\ \mathbf{u}_k = -K\mathbf{x}_k,\ k \geq n \\ (1,4),\ k = 0, 1, \ldots \end{array} \right. \end{array} \tag{6}$$

It is easy to show [9] that the objective J can be equivalently expressed:

$$J = \mathbf{x}_0^T P \mathbf{x}_0 + J_2; \quad J_2 = \sum_{k=0}^{n-1} \mathbf{c}_k^T W \mathbf{c}_k = \mathbf{C}^T \hat{W} \mathbf{C} \tag{7}$$

where $\mathbf{C} = [\mathbf{c}_0^T, \dots, \mathbf{c}_{n-1}^T]^T$, $\hat{W} = \text{diag}[W, \dots, W]$, and $W > 0$ is given by $W = B^T P B + R$. Thus (6) becomes

$$\min_{\mathbf{C}} \mathbf{C}^T \hat{W} \mathbf{C} \quad \text{s.t.} \quad \begin{cases} \mathbf{u}_k = -K\mathbf{x}_k + \mathbf{c}_k, \; k < n \\ (1,4), \; k < n \\ \mathbf{x}_n \in T_0. \end{cases} \tag{8}$$

For convenience we refer to the MPC that arises from the receding horizon application of (8) as Dual Mode Polytopic (DMP) where "polytopic" denotes to the implicit use of T_0.

Remark 1. *Small n will usually result in small stabilizable sets which may not include all likely initial states; this problem is especially relevant if K is highly tuned. Increasing n on the other hand will have a significant effect on the computational load of the QP in (8), which is to be solved on-line.*

2.2 Ellipsoidal feasible invariant sets

A set S is invariant under (1,3) iff $\mathbf{x} \in S$ implies $\Phi\mathbf{x} \in S$. An ellipsoidal set $S_0 = \{\mathbf{x} : \mathbf{x}^T Q_0 \mathbf{x} \le 1\}$, $Q_0 > 0$, will therefore be invariant (for (1,3)), iff

$$\Phi^T Q_0 \Phi - Q_0 \le 0. \tag{9}$$

If additionally $S_0 \subseteq T_0$, then S_0 is both invariant and feasible since constraints (4) are satisfied for all $\mathbf{x} \in S_0$. Thus $\mathbf{x} \in S_0$ implies convergence to the origin of (1,3,4) (assuming strictly stable Φ). Procedures for selecting Q_0 are considered in the literature (eg. [3, 7]) and will not be discussed here; instead it is assumed that Q_0 has been selected so as to maximize the volume of S_0 subject to the constraints $S_0 \subseteq T_0$ and (9).

The linear feedback law (3) does not cater for constraints. To enable systematic constraint handling, MPC introduces degrees of freedom into future inputs — in DMP this is done via the variable $\mathbf{C}$. Alternatively one can find an ellipsoidal feasible invariant set for the augmented state $\mathbf{X} = [\mathbf{x}^T, \mathbf{C}^T]^T$ and compute its projection to $\mathbf{x}$-space, [7, 2]. The maximum volume of this projection is greater than that of the maximal feasible invariant ellipsoidal set for $\mathbf{C} = 0$ and hence the introduction of d.o.f. enables handling of larger sets of initial states. A brief account of this development is given below.

From (1) and $\mathbf{u}_k = -K\mathbf{x}_k + \mathbf{c}_k$, $\mathbf{c}_k = 0$ $k \ge n$, we derive the autonomous formulation:

$$\begin{aligned} \mathbf{X}_{k+1} &= \Psi \mathbf{X}_k; \\ u_k &= \tilde{K} \mathbf{X}_k; \end{aligned} \quad \Psi = \left[\begin{array}{c|cccc} \Phi & B & 0 & \dots & 0 \\ \hline 0 & 0 & & I & \\ \hline 0 & 0 & 0 & \dots & 0 \end{array} \right]$$

$$\mathbf{X}_{k+1} = \begin{bmatrix} \mathbf{x}_{k+1} \\ \hline \mathbf{c}_1 \\ \mathbf{c}_2 \\ \vdots \\ \mathbf{c}_{n-1} \\ \hline 0 \end{bmatrix}, \; \mathbf{X}_k = \begin{bmatrix} \mathbf{x}_k \\ \hline \mathbf{c}_0 \\ \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_{n-2} \\ \hline \mathbf{c}_{n-1} \end{bmatrix}, \; \tilde{K}^T = \begin{bmatrix} K^T \\ \hline I \\ 0 \\ \vdots \\ 0 \\ \hline 0 \end{bmatrix} \tag{10}$$

where I is the identity matrix of conformal dimensions. Then using the invariance condition $\Psi^T \tilde{Q}_n \Psi - \tilde{Q}_n \le 0$ in conjunction with (4) (re-written in terms of the augmented state vector), one can find a feasible invariant set for the augmented state vector $\tilde{S}_n = \{\mathbf{X} : \mathbf{X}^T \tilde{Q}_n \mathbf{X} \le 1\}$. From this it is trivial to project back to $\mathbf{x}$-space to derive a new stabilizable set $S_n = \{\mathbf{x} : \mathbf{x}^T Q_n \mathbf{x} \le 1\}$, where the subscript n is used to denote that this set is obtained through the use of n degrees of freedom. In the following development we assume that $\tilde{Q}_n$ is computed so as to maximize the volume of the projection S_n, and straightforward algebra therefore establishes that the volume of S_n is greater than that of S_{n-1} for all $n \ge 1$.

Earlier work [7] has shown that the computational burden of predictive control can be significantly reduced through the use of ellipsoidal feasible invariant sets. The approach is similar to the DMP strategy discussed in Section 2.1, but replaces the transient constraints of (4) for $k < n$ and the terminal constraint that $\mathbf{x}_n$ lies in the polytope T_0 by the requirement that the augmented state $\mathbf{X}$ lies in the ellipsoid $\tilde{S}_n$. The advantage of using this augmented ellipsoidal set is that the QP optimization of DMP can be replaced by the problem of determining the shortest distance from the origin to an ellipse. This reduces to a univariate problem (irrespective of n) of finding the only negative real root of a known polynomial, and hence the dramatic reduction in computational complexity. The algorithm can be summarized as:

$$\min_{\mathbf{C}} \mathbf{C}^T \hat{W} \mathbf{C} \quad \text{s.t.} \quad \mathbf{X}_1^T \tilde{Q}_n \mathbf{X}_1 \le 1 \tag{11}$$

where $\mathbf{X}_1$ denotes the augmented state at the next time instant. The solution $\mathbf{C}_{\text{opt}}$ to this problem may not be the same as the optimal solution given by DMP, but can be further improved (see [4]), by scaling: $\mathbf{C} = \mu \mathbf{C}_{\text{opt}}$, $\mu \le 1$. Solving for the optimal μ once again constitutes a single variable problem (irrespective of the value of n) in the form of a linear program which is computationally undemanding. The degree of residual sub-optimality (after scaling) is usually minimal. The associated MPC algorithm will be referred to as Dual Mode Ellipsoidal (DME).

The DMP and DME algorithms differ only in the constraints employed in the optimization of predicted inputs. To compare the properties of DMP and DME it is therefore instructive to compare their respective sets of stabilizable initial states. The stabilizable set for DMP with n degrees of freedom (which we denote T_n) is simply the projection onto $\mathbf{x}$-space of the maximal admissible set for (10,4). From the definition of the maximal admissible set as the largest set of stabilizable initial states, it follows that S_n is a subset of, and necessarily smaller than T_n. This is illustrated in Figure 1 which compares the polytope T_n and ellipsoid S_n for the example system of Section 3.1, with $n = 2$ degrees of freedom.

The aim of the current paper is to reduce/remove the restriction on the initial state which arises from the use of ellipsoidal feasible invariant sets, thereby extending the applicability of DME while retaining the computational benefits of (10) in conjunction with $\tilde{S}_n$. As stated above, the com-

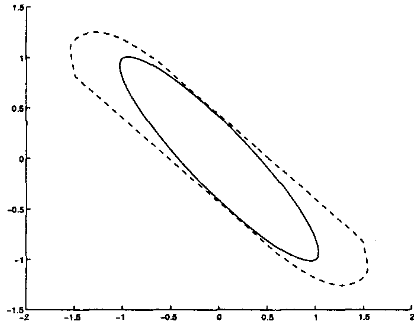

Figure 1: Stabilizable sets: T_2 (dashed) S_2 (solid line)

putational complexity of the on-line optimization in DME is independent of n whereas the volume of S_n is monotonically nondecreasing in n, and so the applicability of DME could be widened by increasing n to some value $N > n$. However even for large N some states (eg. those corresponding to and/or close to vertices of T_n) may lie outside S_N, hence removing the advantages of ellipsoids over polytopes. This is motivates the new strategy outlined in the following section.

3 Introducing the third mode

In order to combine the larger volume of polytopes with the efficiency of ellipsoidal representations, we consider the following set

$$M_{n,m} = \{\mathbf{x}_0 : \mathbf{x}_m \in S_n \quad \text{s.t.} \quad (1,4),\ k < m\}. \tag{12}$$

Here $M_{n,m}$ is the set from which it is possible to drive in m steps the current state $\mathbf{x}_0$ into S_n. This definition introduces a further set of m degrees of freedom and constitutes a "triple mode" prediction strategy comprising the two modes of "dual mode" (ie. the fixed term state feedback law preceded by the n free moves) front-ended by a further m free moves. The set $M_{n,m}$ can be used with the following Triple Mode Predictive Control (TMPC) algorithm.

Algorithm 1 (TMPC). *At each sampling instant:*

1. *If* $\mathbf{x} \in S_n$, *perform the minimization of (11).*
2. *If* $\mathbf{x} \notin S_n$, *solve the following QP:*

$$\min_{\mathbf{c}_k, k<m} J_{\text{ep}} = \mathbf{x}_m^T Q_n \mathbf{x}_m \quad \text{s.t.} \quad \begin{cases} \mathbf{u}_k = -K\mathbf{x}_k + \mathbf{c}_k,\ k < m \\ (1,4),\ k < m. \end{cases} \tag{13}$$

3. *Implement* $\mathbf{u} = -K\mathbf{x} + \mathbf{c}_0$.

Theorem 1. *Under TMPC,* $\mathbf{x} = 0$ *is asymptotically stable for all* $\mathbf{x} \in M_{n,m}$.

Proof. If $\mathbf{x} \in S_n$, the proof follows directly from the use of minimization (11) — see [4]. If $\mathbf{x}$ lies outside S_n but $\mathbf{x} \in M_{n,m}$, then from the definition of $M_{n,m}$ (12), there exists a sequence $\mathbf{c}_k$, $k = 0,\ldots,m-1$ such that constraints are satisfied over those m steps and the predicted value of $\mathbf{x}_m$ lies in S_n. Now S_n is defined via $\mathbf{x}_m^T Q_n \mathbf{x}_m \leq 1$. The choice of objective function J_{ep} in (13) therefore ensures that $\mathbf{x}_m \in S_n$ is obtained under the optimal predicted input sequence with respect to the minimization of step 2. Hence the state necessarily converges to S_n. □

Remark 2. *Step 2 of TMPC involves only transient constraints (which are few in number) and does not have a terminal condition.*

3.1 Comparing DMP and TMPC

Any comparison of the efficiency of DMP and TMPC will depend on constraints, the feedback gain K, the size of state space region that is to be controlled and the system dynamics. This section provides an illustration of how the size of the sets M, S, T can be influenced through the choice of n and m.

Consider the model and constraints

$$A = \begin{bmatrix} 1 & 0.1 \\ 0 & 1 \end{bmatrix},\ B = \begin{bmatrix} 0 \\ 0.0787 \end{bmatrix}; \quad -1 \leq \mathbf{u} \leq 1, \tag{14}$$

for which $Q = \text{diag}\{1,0\}$ and $R = 0.1$ in (2) yield $K = [2.828, 2.826]$. The choice $n = 10, m = 5$ results in the sets $T_{10}, M_{10,5}$ plotted in Figure 2: clearly the two sets are of comparable volume.

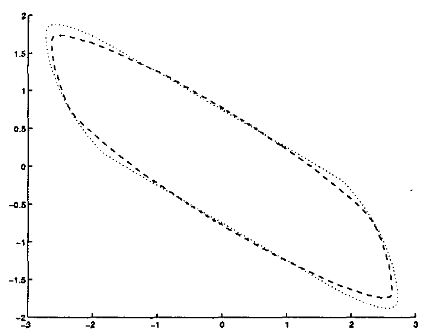

Figure 2: Stabilizable sets: $M_{10,5}$ (dashed), T_{10} (dotted line)

It is reasonable to compare the computational loads of DMP with $n = 10$ and TMPC with $n = 10$, $m = 5$. If $\mathbf{x} \in S_{10}$ then TMPC defaults to DME which is far more efficient. If $\mathbf{x} \notin S_{10}$, then TMPC solves a 5-dimensional QP with only transient constraints whereas DMP solves 10-dimensional QP with far more linear constraints (as required by the definition of T_0). Consequently TMPC is far more efficient. On the other hand, balancing the computational loads by reducing the value of n for DMP from 10 to 5 results in the much smaller stabilizable set T_5 shown in Figure 3.

4 Simulation examples

The potential of TMPC is illustrated here via simulations using the model (14) with initial conditions: $\mathbf{x}_0 =$

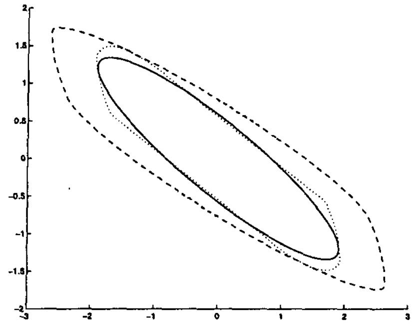

Figure 3: $M_{10,5}$ (dashed), S_{10} (solid), T_5 (dotted line)

$[-0.7, -0.33]^T$ (run 1), and $\mathbf{x}_0 = [-2.47, 1.55]^T$ (run 2). With $n = 10$, $m = 4$ (TMPC) and $n = 10$ (DMP), the corresponding simulations are plotted in Figures 4–7 (run 1) and 8–11 (run 2) for the states, inputs, outputs and d.o.f. $\mathbf{c}$ respectively. The plots for TMPC are in solid line and for DMP in dashed line. The dotted lines in Figures 4 and 8 show the sets S_{10} and S_0. Table 1 gives the costs of (2) for the closed-loop responses of the two algorithms.

Table 1: Closed-loop costs

Algorithm	TMPC	DMP
Run 1	9.33	9.33
Run 2	42.3	37.9

For run 1 the plots are indistinguishable — despite its significantly reduced computational complexity, TMPC did not exhibit any degradation in performance. For run 2, DMP affords slightly better performance. This example shows clearly how TMPC may become suboptimal. The state $\mathbf{x}$ is driven as fast as possible into S_{10} (Figure 8). Once inside S_{10} the control law switches from step 2 to step 1 of algorithm 12; this is notable in the change of gradient in phase-plane trajectory and also in the change in u (see Figure 9 at the 4th sampling instant). On the other hand, DMP minimizes the cost of (2) but converges more slowly (see Figure 11).

5 Conclusions

Appropriate use of ellipsoidal invariant sets in dual mode MPC reduces the on-line optimization to a univariate problem and therefore affords very significant reductions in computational complexity. Polytopic invariant target sets on the other hand have large volume and thus allow for greater applicability (in terms of size of stabilizable sets). Here a triple mode predictive control algorithm has been formulated to take advantage of the computational efficiency afforded by ellipsoidal invariant sets while at the same time removing the potential loss in applicability as compared with the earlier dual mode MPC algorithms. Numerical examples demon-

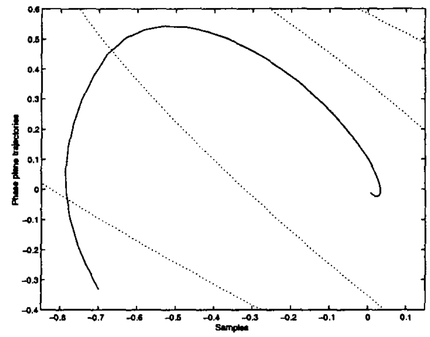

Figure 4: State responses, run 1

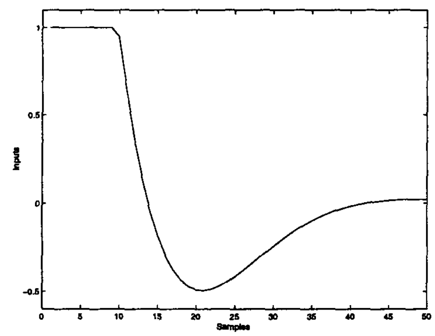

Figure 5: Input responses, run 1

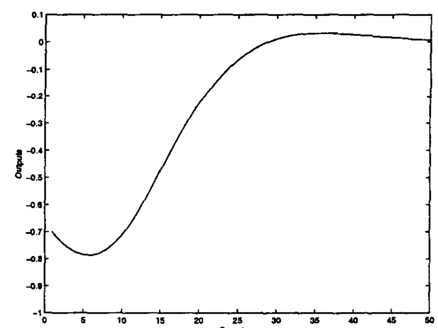

Figure 6: Output responses, run 1

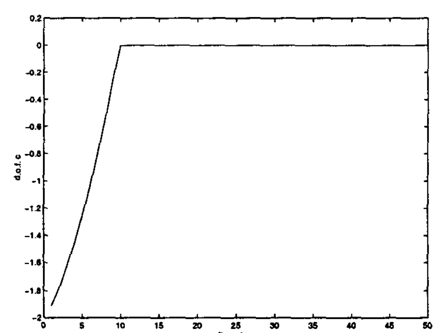

Figure 7: D.o.f. responses, run 1

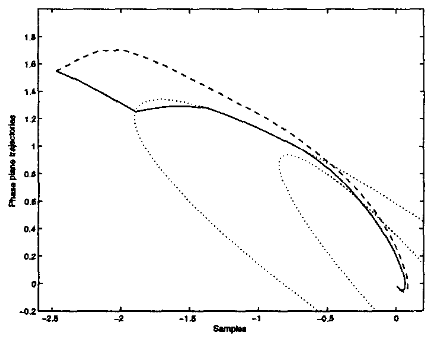

Figure 8: State responses, run 2

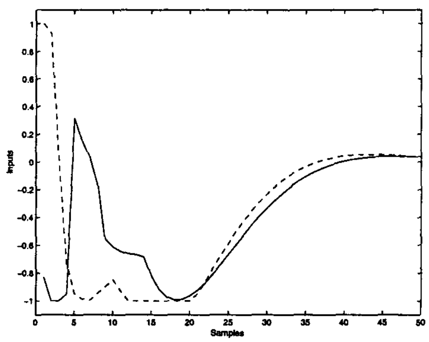

Figure 9: Input responses, run 2

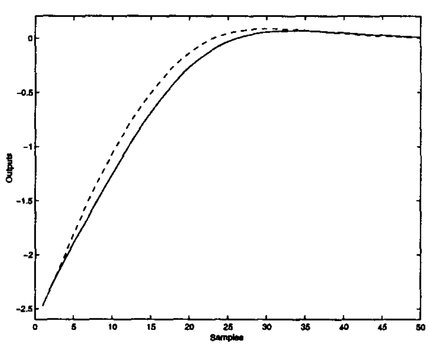

Figure 10: Output responses, run 2

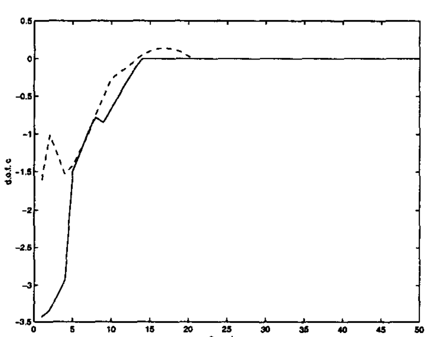

Figure 11: D.o.f. responses, run 2

strate the potential benefits of triple mode control but also highlight some weakness arising from the definition of the associate cost J_{ep}. There is evidence (currently under research) that it is possible to use the cost of (11) and still be able to convert the associated on-line optimization to a convex problem which can be solved efficiently, thereby removing the remaining weakness in TMPC.

Acknowledgments The authors thank EPSRC for financial support.

References

[1] E. G. Gilbert and K. T. Tan. Linear systems with state and control constraints: the theory and practice of maximal admissible sets. *IEEE Trans. AC*, 36(9):1008–1020, 1991.

[2] P. O. Gutman and M. Cwikel. An algorithm to find maximal state constraint sets for discrete time linear dynamical systems with bounded controls and states. *IEEE Trans. AC*, 32(4):251–254, 1987.

[3] M. V. Kothare, V. Balakrishnan, and M. Morari. Robust constrained model predictive control using linear matrix inequalities. *Automatica*, 32(10):1361–1379, 1996.

[4] B. Kouvaritakis, M. Cannon, and J. A. Rossiter. Removing the need for QP in constrained predictive control. Submitted to Adchem 2000.

[5] B. Kouvaritakis and J. A. Rossiter. State-space approaches to interpolation in MPC. To appear in J. Nonlinear Robust Control.

[6] B. Kouvaritakis, J. A. Rossiter, and M. Cannon. Linear quadratic feasible predictive control. *Automatica*, 34(12):1583–1592, 1998.

[7] B. Kouvaritakis, J. A. Rossiter, and J. Schuurmans. Efficient robust predictive control. In *Proc. ACC*, 1999. Submitted to IEEE Trans. AC.

[8] H. Michalska and D. Q. Mayne. Robust receding horizon control of constrained nonlinear systems. *IEEE Trans. AC*, 38(11):1623–1633, 1993.

[9] J. A. Rossiter, B. Kouvaritakis, and M. J. Rice. Numerically robust state-space approach to stable predictive control strategies. *Automatica*, 34(1):65–73, 1998.

[10] P. O. M. Scokaert and J. B. Rawlings. Infinite horizon linear quadratic control with constraints. In *Proc. IFAC World Congr.*, volume M., pages 109–114, San Francisco, 1996.