

Explicit Approaches to Constrained Model Predictive Control: A Survey

ALEXANDRA GRANCHAROVA*† and TOR ARNE JOHANSEN*

Keywords: Model predictive control, constraints, multi-parametric quadratic programming, piecewise linear controllers

This paper presents a review of the explicit approaches to constrained model predictive control. The main motivation behind the explicit solution is that it avoids the need for real-time optimization, and thus allows implementation at high sampling frequencies in real-time systems with high reliability and low software complexity. The paper is organized as follows. Section 1 includes formulation of the constrained linear quadratic regulation (LQR) problem, summary of the implicit approaches, and the basics of the model predictive control (MPC). Sections 2 and 3 consider respectively the exact and the approximate approaches to explicit solution of constrained MPC problems, together with several examples.

1. Constrained linear quadratic regulation

1.1. Problem formulation

We consider time-invariant linear discrete-time system described by the state-space equation:

$$x(t+k+1) = Ax(t+k) + Bu(t+k), \quad k \geq 0 \quad (1)$$

where A and B are the state transition and input distribution matrices. It is assumed that (A, B) is stabilizable.

The control objective is to regulate the state of the system optimally to the origin. Optimality is defined in terms of a *quadratic objective and a set of inequality constraints* (Scokaert & Rawlings, 1998). The objective is defined over an infinite horizon and is given by:

$$J[x(t), \{u(t), u(t+1), \dots\}] = \sum_{k=0}^{\infty} [x^T(t+k)Qx(t+k) + u^T(t+k)Ru(t+k)] \quad (2)$$

in which $Q \geq 0$ and $R > 0$ are symmetric weighting matrices. The constraints are also defined on an infinite horizon and take the form:

$$Hx(t+k+1) \leq h, \quad k \geq 0 \quad (3)$$

$$Gu(t+k) \leq g, \quad k \geq 0 \quad (4)$$

where vectors h and g (of dimension respectively n_h and n_g) define the constraint levels and H and G are the state and input constraint distribution matrices.

Then, 3 control problems are formulated (Scokaert & Rawlings, 1998).

*Department of Engineering Cybernetics, Norwegian University of Science and Technology, N-7491 Trondheim, Norway. E-mail: Tor.Arne.Johansen@itk.ntnu.no

†Institute of Control and System Research, Bulgarian Academy of Sciences, Acad. G. Bonchev str., Bl.2, P.O.Box 79, Sofia 1113, Bulgaria. E-mail: alexandra@icsr.bas.bg

Problem 1—Unconstrained linear quadratic regulation:

$$\min_{\{u(t), u(t+1), \dots\}} \sum_{k=0}^{\infty} [x^T(t+k)Qx(t+k) + u^T(t+k)Ru(t+k)] \quad (5)$$

subject to:

$$x(t+k+1) = Ax(t+k) + Bu(t+k), \quad k \geq 0 \quad (6)$$

The solution to this problem is the linear feedback control law:

$$u(t+k) = -Kx(t+k), \quad k \geq 0 \quad (7)$$

where the controller gain matrix K can be calculated from the solution of the discrete-time algebraic Riccati equation (Kwakernaak & Sivan, 1972).

Problem 2—Constrained linear quadratic regulation:

$$\min_{\{u(t), u(t+1), \dots\}} \sum_{k=0}^{\infty} [x^T(t+k)Qx(t+k) + u^T(t+k)Ru(t+k)] \quad (8)$$

subject to:

$$x(t+k+1) = Ax(t+k) + Bu(t+k), \quad k \geq 0 \quad (9)$$

$$Hx(t+k+1) \leq h, \quad k \geq 0 \quad (10)$$

$$Gu(t+k) \leq g, \quad k \geq 0 \quad (11)$$

Problem 2 is a natural extension of *Problem 1* that includes constraints. The difficulty associated with *Problem 2* is the infinite number of decision variables in the optimization and the infinite number of constraints.

Problem 3—Model Predictive Control (MPC) Problem:

$$\min_{\{u(t), u(t+1), \dots, u(t+N-1)\}} \sum_{k=0}^{N-1} [x^T(t+k)Qx(t+k) + u^T(t+k)Ru(t+k)] \quad (12)$$

subject to:

$$x(t+k+1) = Ax(t+k) + Bu(t+k), \quad k \geq 0 \quad (13)$$

$$Hx(t+k+1) \leq h, \quad k=0, 1, \dots, N-1 \quad (14)$$

$$Gu(t+k) \leq g, \quad k=0, 1, \dots, N-1 \quad (15)$$

$$u(t+k) = -Kx(t+k), \quad k \geq N \quad (16)$$

This form of MPC has a finite number of decision variables, N , and a finite number of constraints, $N(n_h + n_g)$. It can therefore be solved with standard quadratic programming methods. Here, the unconstrained feedback control law (equation (16)) is added to the finite set of N decision variables.

1.2. Implicit approaches

In Scokaert & Rawlings (1998), Sznaier & Damborg (1987) and Chmielewski & Manousiouthakis (1996) *implicit approaches* have been developed to solve the *constrained linear quadratic regulation problem* (Problem 2). They are implicit in sense that the optimal control does not have the form of feedback control law, but it is obtained in the form of open-loop time trajectory.

In their pioneering work Sznaier and Damborg (1987) showed that *finite-horizon optimization* defined as the *Model Predictive Control problem* also provides the solution to the *infinite-horizon linear quadratic regulation problem with constraints*. This equivalence holds for a certain set of initial conditions, which depends on the length of the finite horizon.

This idea has been developed further by Scokaert and Rawlings (1998) and by Chmielewski and Manousiouthakis (1996).

In Scokaert & Rawlings (1998) the following definition is made and an algorithm is proposed to solve the *constrained LQR problem*:

Definition 1: Let $X_K \subseteq R^n$ denotes the set of states $x(t)$ for which the unconstrained LQR law, $u(t+k) = -Kx(t+k)$, $k \geq 0$, satisfies (1), (3) and (4).

Algorithm 1 (constrained LQR):

Step 0. Choose a finite horizon N_0 , set $N = N_0$.

Step 1. Solve Problem 3 (MPC problem).

Step 2. If $x(t+N) \in X_K$, go to step 4.

Step 3. Increase N , go to step 1.

Step 4. Terminate: $\pi^* = \pi_N$.

Here $\pi_N = \{u(t), u(t+1), \dots, u(t+N-1)\}$ is the optimal control trajectory determined by solving the *MPC problem*, while π^* is the optimal control trajectory that is a solution of the *constrained LQR problem*.

In Scokaert & Rawlings (1998) it has been shown that the algorithm terminates in a finite number of iterations, regardless of the choice of initial horizon in Step 0 and of the heuristics used to increase it in Step 3. In other words, the presented algorithm requires solving a finite number of finite-dimensional positive definite quadratic programs (QP). Also, the *constrained LQR* is shown to be both optimal and stabilizing.

However, the dimension of the QP depends on the initial state and therefore the result is not useful for practical applications where an upper bound on the horizon (respectively of the QP size) which is independent on the initial state is needed for designing the control hardware. In this respect, Chmielewski & Manousiouthakis (1996) describe an algorithm which provides a semi-global upper bound. Namely, for any given compact set X_0 of initial conditions, their algorithm provides the horizon N such that the *finite horizon controller* (solution of Problem 3) solves the *infinite horizon problem* (Problem 2).

1.3. Basics of Model Predictive Control

Model predictive control is an efficient methodology to solve complex constrained multivariable control problems (Mayne *et al.*, 2000, Bemporad & Morari, 1999, Bemporad *et al.*, 2000 and Bemporad *et al.*, 2002). Here the basics of Model Predictive Control (MPC) are given according to Bemporad *et al.* (2002).

Consider the problem of regulating to the origin the discrete-time linear time invariant system:

$$x(t+1) = Ax(t) + Bu(t) \quad (17)$$

$$y(t) = Cx(t) \quad (18)$$

while satisfying the following constraints:

$$y_{\min} \leq y(t) \leq y_{\max}, \quad u_{\min} \leq u(t) \leq u_{\max} \quad (19)$$

at all time instants $t \geq 0$. In (17)–(19), $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$ are the state, input and output vector respectively, y_{\min} , y_{\max} and u_{\min} , u_{\max} are respectively p and m -dimensional vectors and the pair (A, B) is stabilizable.

Model Predictive Control (MPC) solves such a constrained regulation problem in the following way (Bemporad *et al.*, 2002). Assume that a full measurement of the state $x(t)$ is available at the current time t . Then, the optimization problem:

$$\min_{U \equiv \{u_t, u_{t+1}, \dots, u_{t+N_u-1}\}} \left\{ I[U, x(t)] = x_{t+N_y|t}^T P x_{t+N_y|t} + \sum_{k=0}^{N_y-1} [x_{t+k|t}^T Q x_{t+k|t} + u_{t+k}^T R u_{t+k}] \right\} \quad (20)$$

subject to $x_{t|t} = x(t)$ and:

$$y_{\min} \leq y_{t+k|t} \leq y_{\max}, \quad k = 1, \dots, N_c \quad (21)$$

$$u_{\min} \leq u_{t+k} \leq u_{\max}, \quad k = 0, 1, \dots, N_c \quad (22)$$

$$x_{t+N_y|t} \in \Omega \quad (23)$$

$$x_{t+k+1|t} = Ax_{t+k|t} + Bu_{t+k}, \quad k \geq 0 \quad (24)$$

$$y_{t+k|t} = Cx_{t+k|t}, \quad k \geq 0 \quad (25)$$

$$u_{t+k} = -Kx_{t+k|t}, \quad N_u \leq k < N_y \quad (26)$$

is solved at each time t , where $x_{t+k|t}$ denotes the predicted state vector at time $t+k$, obtained by applying the input sequence u_t, \dots, u_{t+k-1} to model (17)–(18) starting from the state $x(t)$.

The name MPC stems from the idea of employing an explicit model of the plant to be controlled which is used to predict the future output behavior. This prediction capability allows solving optimal control problems on line, where tracking error, namely the difference between the predicted output and the desired reference, is minimized over a future horizon (Bemporad *et al.*, 2002).

Further, we assume that $Q = Q^T \geq 0$, $R = R^T > 0$, $P \geq 0$, $(Q^{1/2}, A)$ detectable (for instance $Q = C^T C$ with (C, A) detectable), K is some feedback gain, N_y , N_u , N_c are the output, input and constraint horizons, respectively, with $N_u \leq N_y$ and $N_c \leq N_y - 1$, and Ω is a polyhedral terminal set.

One possibility is to choose $K=0$ and P as the solution of the Lyapunov equation (Bemporad *et al.*, 2002):

$$P = A^T P A + Q \quad (27)$$

The choice $K=0$ implies that after N_u time steps the control is turned off and the system is allowed to settle in an open-loop manner. This is only meaningful when the system is open-loop stable. With P obtained from (27), $J[U, x(t)]$ measures the settling cost of the system from the present time t to infinity under the assumption that the control is turned off after N_u steps.

Alternatively, one can set $K=K_{LQ}$, where K_{LQ} and P are the solution of the unconstrained infinite horizon LQR problem with weights Q and R (Bemporad *et al.*, 2002):

$$K_{LQ} = -(R + B^T P B)^{-1} B^T P A \quad (28)$$

$$P = (A + B K_{LQ})^T P (A + B K_{LQ}) + K_{LQ}^T R K_{LQ} + Q \quad (29)$$

This choice of K implies that after N_u time steps the control is switched to the unconstrained LQR. With P obtained from (29), $J[U, x(t)]$ measures the settling cost of the system from the present time t to infinity under this control assumption.

The MPC control law is based on the following idea (Bemporad *et al.*, 2002): At time t compute the optimal solution to problem (20)–(26) (the optimal input sequence):

$$U^*(t) = \{u_t^*, \dots, u_{t+N_u-1}^*\} \quad (30)$$

and apply to the system only the first input from the sequence:

$$u(t) = u_t^* \quad (31)$$

The remaining optimal inputs are discarded and a new optimal control problem is solved at time $t+1$, based on the new state $x(t+1)$. Such a control strategy is also referred to as *moving or receding horizon*. This idea is illustrated in Figure 1.

As new measurements are collected from the plant at each time t , the receding horizon mechanism provides the controller with the desired feedback characteristics.

The stability of MPC feedback loops was investigated by numerous researchers

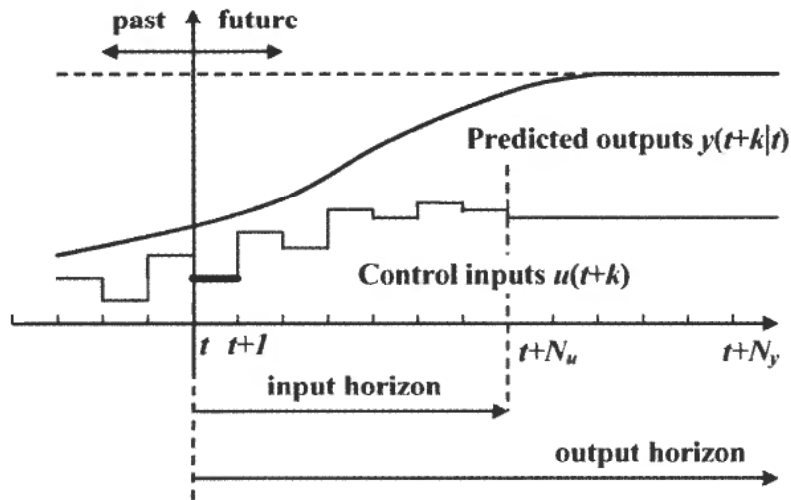


Figure 1. Receding horizon strategy: only the first input of the computed optimal input sequence is implemented.

(Mayne *et al.*, 2000, Bemporad & Morari, 1999 and Bemporad *et al.*, 2002). Stability is in general a complex function of the various tuning parameters N_u , N_y , N_c , P , Q and R (Bemporad *et al.*, 2002). For applications it is most useful to impose some conditions on N_y , N_c and P so that stability is guaranteed for all $Q \geq 0$, $R > 0$. Then Q and R can be freely chosen as tuning parameters to affect performance. The constraint (23) is sometimes called 'stability constraint' and it explicitly forces the state vector to reach an invariant set at the end of the prediction horizon. The stability result is formulated in the following theorem (Bemporad *et al.*, 2002):

Theorem 1:

Let $N_y = \infty$, $K = 0$ or $K = K_{LQ}$, and $N_c < \infty$ be sufficiently large for guaranteeing existence of feasible input sequences at each time step. Then the MPC law (20)–(31) asymptotically stabilizes the system (17)–(18) while enforcing the fulfilment of the constraints (19) from all initial states $x(0)$ such that (20)–(26) is feasible at $t = 0$. ■

2. Exact Approaches to Explicit Solution of MPC Problems

2.1. MPC computation

By substituting (Bemporad *et al.*, 2000 and Bemporad *et al.*, 2002):

$$x_{t+k|t} = A^k x(t) + \sum_{j=0}^{k-1} A^j B u_{t+k-1-j} \quad (32)$$

in the optimization problem (20)–(26), this can be rewritten in the form:

$$V^*[x(t)] = \frac{1}{2} x^T(t) Y x(t) + \min_u \left\{ \frac{1}{2} U^T H U + x^T(t) F U \right\} \quad (33)$$

subject to:

$$G U \leq W + E x(t) \quad (34)$$

where the column vector $U \equiv [u_t^T, \dots, u_{t+N_u-1}^T]^T \in \mathcal{R}^s$, $s = m N_u$, is the optimization vector, $H = H^T > 0$ and H , F , Y , G , W , E are easily obtained from Q , R and (20)–(32).

The optimization problem (33)–(34) is a quadratic program (QP). Because the problem depends on the current state $x(t)$, the implementation of MPC requires the on-line solution of a QP at each time step. Although efficient QP solvers based on active-set methods and interior point methods are available, computing the input $u(t)$ demands significant on-line computation effort. For this reason, the application of MPC has been limited to 'slow' and/or 'small' processes.

In Bemporad *et al.* (2000), Bemporad *et al.* (2002) and Pistikopoulos *et al.* (2000), a new approach for MPC implementation has been proposed, where the computation effort is moved off-line. The MPC formulation described in the previous section provides the control action $u(t)$ as a function of $x(t)$ implicitly defined by (33)–(34). By treating $x(t)$ as a vector of parameters, the goal is to solve (33)–(34) off-line with respect to all the values of $x(t)$ of interest and make this dependence explicit (Bemporad *et al.*, 2000, Bemporad *et al.*, 2002 and Pistikopoulos *et al.*, 2000).

In terms of operations research, mathematical programs which depend only on one scalar parameter are referred to as *parametric programs*, while problems depending on a vector of parameters as *multi-parametric programs*. According to this terminology, (33)–(34) is a *multi-parametric Quadratic Program* (mp-QP).

Once the multi-parametric problem (33)–(34) has been solved off-line, i.e. the solution:

$$U_t^* = U^*[x(t)] \quad (35)$$

of (33)–(34) has been found, the model predictive controller (20)–(26) is available *explicitly*, as the optimal input $u(t)$ consists simply of the first m components of $U^*[x(t)]$:

$$u(t) = [I \ 0 \ \dots \ 0] U^*[x(t)] \quad (36)$$

It has been shown in Bemporad *et al.* (2000), Bemporad *et al.* (2002) and Pistikopoulos *et al.* (2000) that the solution $U^*(x)$ of the mp-QP problem is a *continuous* and *piecewise affine* function of x . Therefore, the same properties are inherited by the controller.

2.2. Exact approach to explicit solution of MPC problems

Bemporad *et al.* (2000) and Bemporad *et al.* (2002) have developed an algorithm to express the solution $U^*(x)$ and the minimum value $V^*(x) = I[U^*(x)]$ as an *explicit* function of the parameters x and to characterize the analytical properties of these functions. In particular they have proved that the solution $U^*(x)$ is a *continuous piecewise affine* function of x in the following sense (Bemporad *et al.*, 2000 and Bemporad *et al.*, 2002):

Definition 2:

A function $z(x): X \rightarrow \mathbb{R}^s$, where $X \subseteq \mathbb{R}^n$ is a polyhedral set, is *piecewise affine* if it is possible to partition X into convex polyhedral regions, CR_i , and $z(x) = H^i x + k^i$, $\forall x \in CR_i$.

Piecewise quadraticity is defined analogously by letting $z(x)$ be a quadratic function $x^T W^i x + H^i x + k^i$.

In Bemporad *et al.* (2000) and Bemporad *et al.* (2002) it is defined:

$$z \equiv U + H^{-1} F^T x(t) \quad (37)$$

where $z \in \mathbb{R}^s$ and the problem (33)–(34) is transformed to the equivalent problem:

$$V_z^*(x) = \min_z \frac{1}{2} z^T H z \quad (38)$$

subject to:

$$Gz \leq W + Sx(t) \quad (39)$$

where $S \equiv E + GH^{-1}F^T$ and $V_z^*(x) = V^*(x) - \frac{1}{2} x^T (Y - FH^{-1}F^T)x$

The solution of mp-QP problems can be approached by employing the principles of parametric nonlinear programming and in particular the first-order Karush-Kuhn-

Tucker (KKT) optimality conditions (Nocedal & Wright, 1999), which lead to the Basic Sensitivity Theorem.

Instead, Bemporad *et al.* (2000) and Bemporad *et al.* (2002) have adopted a more direct approach which exploits the *linearity* of the constraints and the fact that the function to be minimized is *quadratic*. The approach in Bemporad *et al.* (2000) and Bemporad *et al.* (2002) is described as follows. In order to start solving the mp-QP problem, an initial vector x_0 inside the polyhedral set X of parameters is needed, such that the QP problem (38)–(39) is feasible for $x = x_0$. Such a vector can be found for instance by solving the *linear program* (LP) (Bemporad *et al.*, 2000 and Bemporad *et al.*, 2002):

$$\max_{x, z, \varepsilon} \varepsilon \quad (40)$$

subject to:

$$Gz - Sx + \varepsilon \leq W \quad (41)$$

$$\varepsilon \geq 0 \quad (42)$$

$$x \in X \quad (43)$$

If the LP is infeasible, then the QP problem (38)–(39) is infeasible for all $x \in X$. Otherwise, it is fixed $x = x_0$ and the QP problem (38)–(39) is solved in order to obtain the corresponding optimal solution z_0 . Such a solution is unique because $H > 0$ and therefore uniquely determines a set of active constraints $\tilde{G}z_0 = \tilde{S}x_0 + \tilde{W}$ out of the constraints in (38)–(39). Let \tilde{G} , \tilde{S} and \tilde{W} denote the rows of G , S and W corresponding to the active constraints. Then, the following theorem is proved (Bemporad *et al.*, 2000 and Bemporad *et al.*, 2002):

Theorem 2:

Let $H > 0$. Consider a combination of active constraints \tilde{G} , \tilde{S} , \tilde{W} and assume that the rows of \tilde{G} are linearly independent. Let CR_0 be the set of all vectors x for which such a combination is active at the optimum (CR_0 is referred to as critical region). Then, the optimal z and the associated vector of Lagrange multipliers λ are uniquely defined affine functions of x over CR_0 . ■

Proof (Bemporad *et al.*, 2002):

The first-order KKT conditions for the mp-QP are given by:

$$Hz + G^T \lambda = 0, \quad \lambda \in \mathcal{R}^q \quad (44)$$

$$\lambda_i (G^i z - W^i - S^i x) = 0, \quad i = 1, \dots, q \quad (45)$$

$$\lambda \geq 0 \quad (46)$$

where the superscript i denotes the i -th row. Equation (44) is solved for z :

$$z = -H^{-1} G^T \lambda \quad (47)$$

and the result is substituted into (45) to obtain the complementary slackness condition:

$$\lambda (-GH^{-1} G^T \lambda - W - Sx) = 0 \quad (48)$$

Let $\tilde{\lambda}$ and $\tilde{\lambda}$ denote the Lagrange multipliers corresponding to *inactive* and *active* constraints, respectively. For *inactive* constraints $\tilde{\lambda}=0$. For *active* constraints $-\tilde{G}H^{-1}\tilde{G}^T\tilde{\lambda}-\tilde{W}-\tilde{S}x=0$ and therefore:

$$\tilde{\lambda}=-(\tilde{G}H^{-1}\tilde{G}^T)^{-1}(\tilde{W}+\tilde{S}x) \quad (49)$$

where \tilde{G} , \tilde{W} , \tilde{S} correspond to the set of active constraints and $(\tilde{G}H^{-1}\tilde{G}^T)^{-1}$ exists because the rows of \tilde{G} are linearly independent. Thus λ is an affine function of x . By substituting $\tilde{\lambda}$ from (49) into (47), it is obtained:

$$z=H^{-1}\tilde{G}^T(\tilde{G}H^{-1}\tilde{G}^T)^{-1}(\tilde{W}+\tilde{S}x) \quad (50)$$

and it is noted that z is also an affine function of x .

Theorem 2 characterizes the solution only locally in the neighborhood of a specific x_0 , as it does not provide the construction of the set CR_0 where this characterization remains valid. On the other hand, this region can be characterized immediately (Bemporad *et al.*, 2000 and Bemporad *et al.*, 2002). The variable z from (47) must satisfy the constraints (39):

$$GH^{-1}\tilde{G}^T(\tilde{G}H^{-1}\tilde{G}^T)^{-1}(\tilde{W}+\tilde{S}x)\leq W+Sx \quad (51)$$

and by (46) the Lagrange multipliers in (49) must remain nonnegative:

$$-(\tilde{G}H^{-1}\tilde{G}^T)^{-1}(\tilde{W}+\tilde{S}x)\geq 0 \quad (52)$$

as x varies. After removing the redundant inequalities from (51) and (52), a compact representation of CR_0 is obtained. Obviously, CR_0 is a polyhedron in the x -space and represents the largest set of $x\in X$ such that the combination of *active* constraints at the minimizer remains unchanged (Figure 2(a)). Then, the algorithm in Bemporad *et al.* (2000) and Bemporad *et al.* (2002) continues with the division of the parameter space as in Figure 2(b) and (c) by reversing one by one the hyperplanes defining the

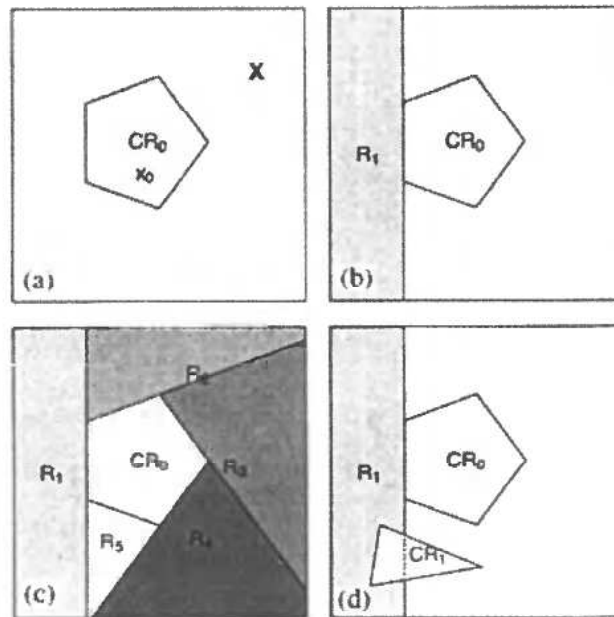


Figure 2. State space exploration strategy of Bemporad *et al.* (2000 and 2002).

critical region CR_0 . Iteratively each new region R_i is subdivided in a similar way as was done with X . As noted in Tøndel *et al.* (2003), the main drawback of this algorithm is that the regions R_i are not related to optimality, as they can split some of the critical regions like CR_1 in Figure 2(d). A consequence is that CR_1 will be detected at least twice.

The properties of the set of feasible parameters $X_f \subseteq X$ (i.e. the set of parameters $x \in X$ such that a feasible solution $z^*(x)$ exists to the optimization problem (38)–(39)), the value function $V_z^*(x)$ and the solution $z^*(x)$ are formulated in the following theorem (Bemporad *et al.*, 2002):

Theorem 3:

Consider the multi-parametric quadratic program (38)–(39) and let $H > 0$, X convex. Then the set of feasible parameters $X_f \subseteq X$ is convex, the optimizer $z^*(x): X_f \mapsto \mathcal{R}^s$ is continuous and piecewise affine and the value function $V_z^*(x): X_f \mapsto \mathcal{R}$ is continuous, convex and piecewise quadratic. ■

Based on the above results, the main steps of the off-line mp-QP solver are outlined in the following algorithm (Bemporad *et al.*, 2002):

Algorithm 2 (exact mp-QP):

- Step 1. Let the current region be the whole set $X \subseteq \mathcal{R}^n$.
- Step 2. Choose a vector x_0 in the current region by solving the linear program (40)–(43).
- Step 3. For $x = x_0$, compute the corresponding optimal solution (z_0, λ_0) by solving a QP.
- Step 4. Determine the set of active constraints when $z = z_0$, $x = x_0$, and build \tilde{G} , \tilde{W} , \tilde{S} .
- Step 5. If $r = \text{rank } \tilde{G}$ is less than the number l of rows of \tilde{G} , take a subset of r linearly independent rows and redefine \tilde{G} , \tilde{W} , \tilde{S} accordingly.
- Step 6. Determine $\tilde{\lambda}(x)$, $z(x)$, from (49) and (50).
- Step 7. Characterize the CR from inequalities (51) and (52).
- Step 8. Define and partition the rest of the region.
- Step 9. For each nonempty new sub-region, go to step 2.
- Step 10. When all regions have been explored, for all polyhedral regions where $z(x)$ is the same and whose union is a convex set, compute such a union.

In conclusion, Algorithm 2 provides the *explicit* solution $u = f(x)$ to the MPC problem (20)–(26), as the piecewise linear (PWL) function:

$$u = K_1^i x + k_2^i \quad \text{if } H^i x \leq h^i, \quad i = 1, \dots, N_{MPC} \quad (53)$$

where the polyhedral sets $\{H^i x \leq h^i\}$, $i = 1, \dots, N_{MPC}$ are a partition of the given set of states X .

Algorithms for iteratively constructing a polyhedral partition of the state space and computing the PWL solution are also given in Pistikopoulos *et al.* (2000), Tøndel *et al.* (2003), Johansen *et al.* (2002) and Seron *et al.* (2000).

2.3. Efficient implementation of the exact approach to explicit solution of MPC problems

2.3.1. Main theoretical result. The approach of Tøndel *et al.* (2003) modifies the *explicit* approach of Bemporad *et al.* (2000 and 2002) by analyzing several properties

of the geometry of the polyhedral partition and its relation to the combination of active constraints at the optimum of the quadratic program. Based on that, they derive a new exploration strategy for sub-dividing the parameter space, which:

1. avoids unnecessary partitioning;
2. avoids the solution to LP problems for determining an interior point in each new region of the parameter space;
3. avoids the solution to the QP problem for such an interior point.

As a consequence, there is a significant improvement of efficiency with respect to the algorithm of Bemporad *et al.* (2000) and Bemporad *et al.* (2002).

Before describing the main idea of the approach (Tøndel *et al.*, 2003), some definitions are made (Tøndel *et al.*, 2003):

Definition 3:

Let $z^*(x)$ be the optimal solution to (38)–(39) for a given x . We define *active constraints* the constraints with $G^i z^*(x) - W^i - S^i x = 0$ and *inactive constraints* the constraints with $G^i z^*(x) - W^i - S^i x < 0$. The *optimal active set* $A^*(x)$ is the set of indices of active constraints at the optimum $A^*(x) = \{i | G^i z^*(x) = W^i + S^i x\}$. We also define as *weakly active constraint* an active constraint with an associated zero Lagrange multiplier λ^i and as *strongly active constraint* an active constraint with a positive Lagrange multiplier λ^i .

Definition 4:

For an active set, we say that the *linear independence constraint qualification (LICQ)* holds if the set of active constraint gradients are linearly independent, i.e. \tilde{G} has full row rank.

Below, the linear expression of the PWL function $z^*(x)$ over the critical region CR_k is denoted by $z_k^*(x)$. In general, a superscript index is used to denote a row of a matrix or element of a vector.

Definition 5:

Let a polyhedron $X \subset \mathbb{R}^n$ be represented by the linear inequalities $A_0 x \leq b$. Let the i -th hyperplane, $A_0^i x = b^i$ be denoted by Ψ . If $X \cap \Psi$ is $(n-1)$ -dimensional then $X \cap \Psi$ is called a *facet* of the polyhedron.

Definition 6:

Two polyhedra are called *neighboring* polyhedra if they have a common facet.

Definition 7:

Let a polyhedron X be represented by $A_0 x \leq b$. We say that $A_0^i x \leq b^i$ is *redundant* if $A_0^i x \leq b^i \forall j \neq i \Rightarrow A_0^i x \leq b^i$ (i.e. it can be removed from the description of the polyhedron). The inequality i is *redundant with degree h* if it is redundant but there exists a h -dimensional subset Y of X such that $A_0^i x = b^i$ for all $x \in Y$.

Let us consider a hyperplane defining the common facet between two polyhedra CR_0 , CR_i in the optimal partition of the state space. There are two different kinds of hyperplanes (Tøndel *et al.*, 2003). The first (Type I) are those described by (51), which represent a non-active constraint that becomes active at the optimum as x moves from CR_0 to CR_i . This means that if a polyhedron is bounded by a hyperplane which originates from inequality (51), the corresponding constraint will be activated

on the other side of the facet defined by this hyperplane. In addition, the corresponding Lagrange multiplier may become positive. The other kind (Type II) of hyperplanes which bounds the polyhedra are those described by (52). In this case, the corresponding constraint will be non-active on the other side of the facet defined by this hyperplane. This is formulated in the following theorem (Tøndel *et al.*, 2003):

Theorem 4:

Consider an optimal active set $\{i_1, i_2, \dots, i_k\}$ and its corresponding n -minimal representation of the critical region CR_0 obtained by (51)–(52) after removing redundant inequalities. Let CR_i be a full-dimensional neighboring critical region to CR_0 and assume LICQ holds on their common facet $\Phi = CR_0 \cap \Psi$ where Ψ is the separating hyperplane between CR_0 and CR_i . Moreover, assume that there are no constraints which are weakly active at the optimizer $z^*(x)$ for all $x \in CR_0$. Then:

Type I. If Ψ is given by $G^{i_{k+1}} z_0^*(x) = W^{i_{k+1}} + S^{i_{k+1}} x$, then the optimal active set in CR_i is $\{i_1, \dots, i_k, i_{k+1}\}$.

Type II. If Ψ is given by $\lambda_0^i(x) = 0$, then the optimal active set in CR_i is $\{i_1, \dots, i_{k-1}\}$. ■

2.3.2. *Example.* The example is taken from Tøndel *et al.* (2003). Consider the double integrator (Johansen *et al.*, 2002):

$$A = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} T_s^2 \\ T_s \end{bmatrix} \quad (54)$$

where the sampling interval is $T_s = 0.05$ and consider the MPC problem over the prediction horizon $N = 2$ with cost matrices:

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad R = 1 \quad (55)$$

The constraints in the system are:

$$-1 \leq u \leq 1 \quad (56)$$

$$-0.5 \leq x_2 \leq 0.5 \quad (57)$$

The mp-QP associated with this problem has the form (38)–(39) with H, F, G, W, S given in Tøndel *et al.* (2003):

$$H = \begin{bmatrix} 1.079 & 0.076 \\ 0.076 & 1.073 \end{bmatrix}, \quad F = \begin{bmatrix} 1.109 & 1.036 \\ 1.573 & 1.517 \end{bmatrix} \quad (58)$$

$$G^T = \begin{bmatrix} 1 & 0 & -1 & 0 & 0.05 & 0.05 & -0.05 & -0.05 \\ 0 & 1 & 0 & -1 & 0 & 0.05 & 0 & -0.05 \end{bmatrix} \quad (59)$$

$$W^T = [1 \quad 1 \quad 1 \quad 1 \quad 0.5 \quad 0.5 \quad 0.5 \quad 0.5] \quad (60)$$

$$S^T = \begin{bmatrix} 1.0 & 0.9 & -1.0 & -0.9 & 0.1 & 0.1 & -0.1 & -0.1 \\ 1.4 & 1.3 & -1.4 & -1.3 & -0.9 & -0.9 & 0.9 & 0.9 \end{bmatrix} \quad (61)$$

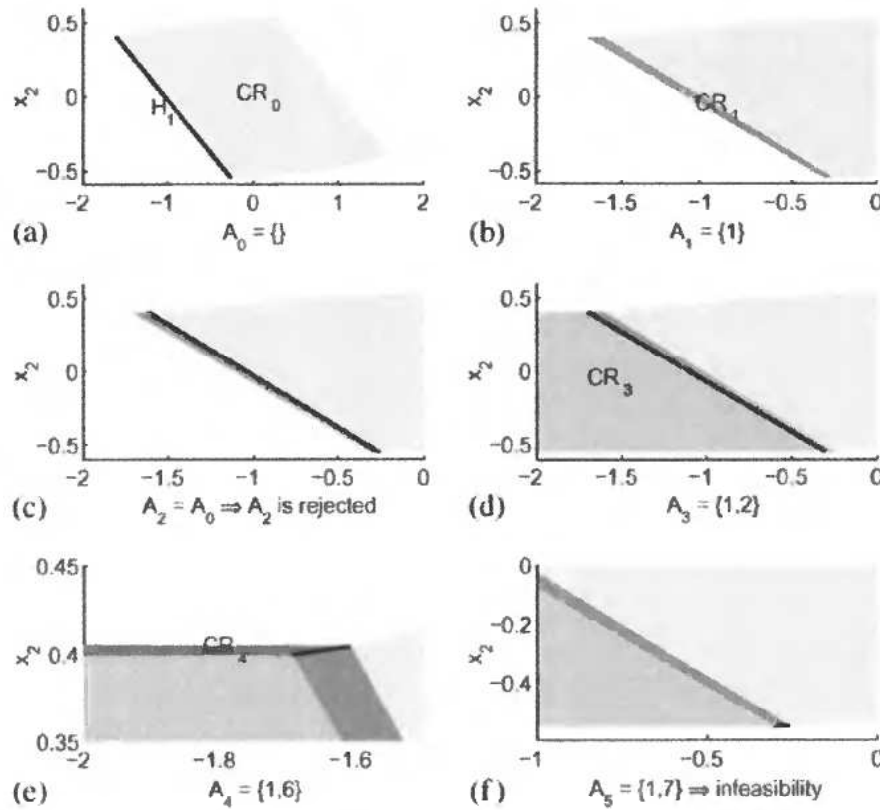


Figure 3. Critical regions for double integrator.

The partitioning starts with finding the region where no constraints are active. As the mp-QP is created from a feasible MPC problem, the empty active set will be optimal in some full-dimensional region ($A_0 = \emptyset$ and \tilde{G} , \tilde{W} and \tilde{S} are empty matrices, $z^*(x) = 0$ and the first component of $U^*(x)$ is the unconstrained LQR gain). This critical region is then described by $0 \leq W + Sx$ which contains 8 inequalities. Two of these inequalities are redundant with degree 0 (#2 and #4), the remaining 6 hyperplanes are facet inequalities of the polyhedron (see Figure 3(a)).

By crossing the facet given by Ψ_1 , defined by inequality 1 and of Type I, as predicted by Theorem 4 the optimal active set across this facet is $A_1 = \{1\}$, which leads to the critical region CR_1 (see Figure 3(b)). After removing redundant inequalities we are left with an n -minimal representation of CR_1 containing 4 facets. The first of these is of Type II, $\lambda^1(x) = 0$. The other three are of Type I. These are inequalities #2, #6 and #7. Consider first the other side of the facet which comes from $\lambda^1(x) = 0$, see Figure 3(c). The region should not have constraint 1 active, so the optimal active set is $A_2 = \emptyset$. This is the same combination of active constraints as A_0 , as expected, so A_2 is not pursued. Next, consider crossing the respective facets of inequalities 2, 6 and 7, see Figures 3(d)–3(f). This results in three different active sets: $A_3 = \{1, 2\}$, $A_4 = \{1, 6\}$ and $A_5 = \{1, 7\}$. The sets A_3 and A_4 lead to new polyhedra as shown in the figures. The combination A_5 leads to an interesting case of ‘degeneracy’. The associated matrix \tilde{G} has linearly dependent rows, which violates the LICQ assumption. In this case, A_5 leads to an infeasible part of the state space.

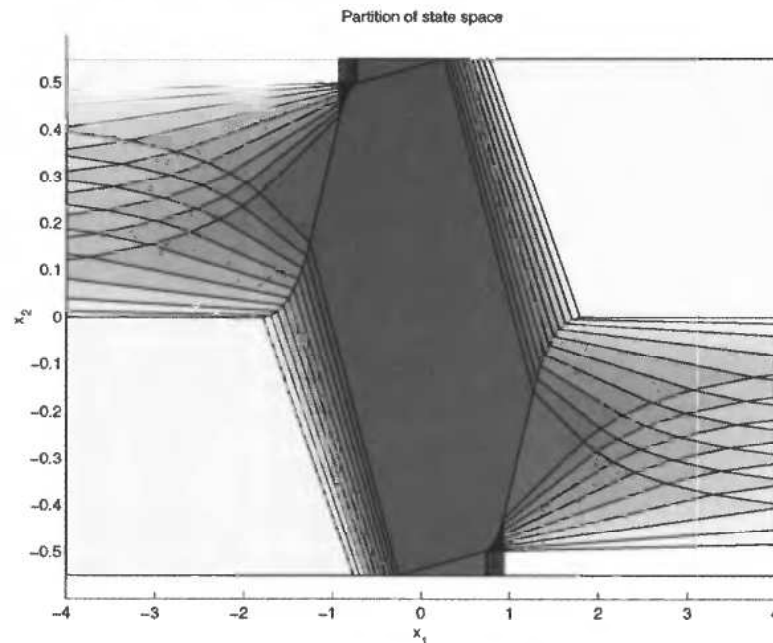


Figure 4. Polyhedral partition of state space for the double integrator with $N=10$.

3. Approximate Approach to Explicit Solution of MPC Problems

3.1. Complexity of the exact approaches

Consider the same double integrator example as in section 2.3.2. Figure 4 shows the partition for horizon $N=10$ corresponding to the exact solution provided by the algorithm (Tøndel *et al.*, 2003). We observe that the exact solution is fairly complex, containing 191 polyhedral critical regions, many of them of very small volume.

3.2. Main idea of the approximate approach

Here we suggest an entirely different approach to compute sub-optimal explicit MPC solutions (Johansen & Grancharova, 2002, 2003 and Grancharova & Johansen, 2002). The idea is to require that the state space partition is represented as a search tree, i.e. to consist of orthogonal hypercubes organized in a hierarchical data-structure that allows *extremely fast real-time search*. The computational complexity with the suggested approach is logarithmic with respect to the number of regions, while a general polyhedral partitioning leads to a computational complexity that is linear with respect to the number of regions, if no additional data structures are built. The optimal solution is computed explicitly using quadratic programming (QP) only at the vertices of these hypercubes, and an approximate solution valid in the whole hypercube is computed based on this data. A hypercube is partitioned into two or more smaller hypercubes only if this is necessary to achieve the desired local accuracy of the solution. This makes the idea similar to storing the pre-computed QP solutions at the various states in a multi-resolution lookup table.

Unlike any other method mentioned above, that all relies on the linearity of the problem to build polyhedral regions and a PWL (piecewise linear) solution, the

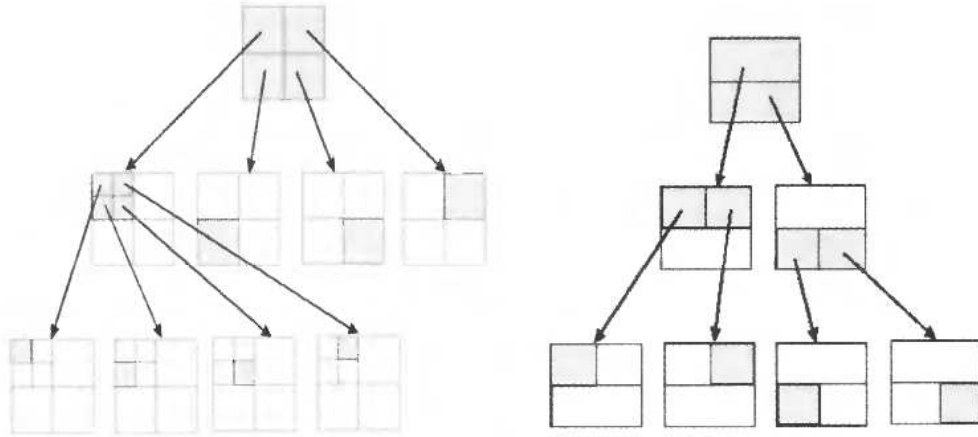


Figure 5. Partition of a rectangular region in a 2-dimensional state space. Left: quad-tree partition. Right: $k-d$ tree partition.

suggested method is straightforward to be extended to nonlinear constrained MPC problems by replacing the QP with a nonlinear program.

3.3. Approximate mp-QP algorithm

We restrict our attention to a hypercube $X \subset \mathbb{R}^n$ where we seek to approximate the optimal PWL solution $z^*(x)$ to the mp-QP problem (38)–(39). In order to minimize the real-time computational complexity we require that the state space partition is orthogonal and can be represented as a search tree (generalized quad-tree (Bentley, 1975), Figure 5 (left)), such that the search complexity is logarithmic with respect to the number of regions.

The orthogonal search tree is a hierarchical data structure where a hypercube can be hierarchically subdivided into smaller hypercubes allowing the local resolution to be adapted, as shown in Figure 5. When searching the tree, only n scalar comparisons are required at each level.

The improved version of the approximate mp-QP algorithm is based on a $k-d$ tree partition of the state space (Figure 5 (right)) as a more flexible and powerful alternative to the generalized quad-tree (Figure 5 (left)). With the $k-d$ tree (Bentley, 1975), a hyper-rectangle is split into two equal parts and thus only *one* scalar comparison is required at each level when searching the tree. Also, the $k-d$ tree allows the incorporation of heuristic rules that split the hyper-rectangle at the axis along which the change of error is maximal (before splitting).

There are two versions of the approximate mp-QP algorithm, based respectively on the cost function approximation error $\varepsilon_{\text{cost}}(x)$ (Johansen & Grancharova, 2002, 2003), and on the control input approximation error $\varepsilon_{\text{input}}(x)$ (Grancharova & Johansen, 2002). The approximation error in the cost function is:

$$\varepsilon_{\text{cost}}(x) = \hat{V}_z(x) - V_z^*(x) \quad (62)$$

where

$$\hat{V}_z(x) = \frac{1}{2} \hat{z}_0^T(x) H \hat{z}_0(x) \quad \text{and} \quad V_z^*(x) = \frac{1}{2} z^{*T}(x) H z^*(x)$$

are respectively the sub-optimal and the optimal costs, and $\hat{z}_0(x)$ and $z^*(x)$ are the sub-optimal and the optimal PWL solutions. The approximation error in the control input is:

$$\varepsilon_{input}(x) = (z^*(x) - \hat{z}_0(x))^T \Sigma (z^*(x) - \hat{z}_0(x)) \quad (63)$$

where $\Sigma \geq 0$ is a weighting matrix which typically has non-zero weight only on the components of the solution corresponding to the first sample of the trajectory.

Initially the algorithm will consider the whole region $X_0 = X$. The main idea of the approximate mp-QP algorithm is to compute the solution of the problem (38)–(39) at the 2^n vertices of a considered hyper-rectangle X_0 by solving up to 2^n QPs. Based on these solutions, a feasible local linear approximation $\hat{z}_0(x)$ to the PWL optimal solution $z^*(x)$, valid in the whole hyper-rectangle X_0 , is computed by using the following result (Bemporad & Filippi, 2003):

Lemma 1:

Consider the bounded polyhedron $X_0 \subseteq X_f$ with vertices $\{v_1, v_2, \dots, v_M\}$ (here X_f is the feasible set: $X_f = \{x(t) \in \mathcal{H}^n | \exists U \text{ satisfying (34)}\}$). If K_0 and g_0 solve the QP:

$$\min_{K_0, g_0} \sum_{i=1}^M (z^*(v_i) - K_0 v_i - g_0)^T H (z^*(v_i) - K_0 v_i - g_0) \quad (64)$$

subject to:

$$G(K_0 v_i + g_0) \leq S v_i + W, \quad i \in \{1, 2, \dots, M\} \quad (65)$$

then the least squares approximation $\hat{z}_0(x) = K_0 x + g_0$ is feasible for the mp-QP (38)–(39) for all $x \in X_0$. ■

If the maximal approximation error ε_0 in the hyper-rectangle X_0 is smaller than some prescribed tolerance $\bar{\varepsilon} > 0$, no further refinement of X_0 is needed. Otherwise, X_0 is partitioned into two hyper-rectangles and the procedure described above is repeated for each of these. If the approximation error in the cost function is considered, the upper bound ε_0 is determined by using the method proposed in Johansen & Grancharova (2002, 2003). If the approximation error in the control input is being used, then the maximal value ε_0 is determined in the way given in Grancharova & Johansen (2002).

In order to reduce the complexity of the partition, the heuristic rule described in Grancharova & Johansen (2002) is applied when splitting the hyper-rectangle X_0 . The rule attempts to split the hyper-rectangle at the axis along which the change of the approximation error is maximal (before splitting), because it is reasonable to hope this is how the largest reduction of the error can be made. The heuristic rule uses information about the error ($\varepsilon_{cost}(x)$ or $\varepsilon_{input}(x)$) in the hyper-rectangle $X_0^d \subset X_0$ that contains a finite number of representative points in X_0 , typically the vertices of one or more hyper-rectangles contained in the interior of X_0 .

Heuristic splitting rule:

Split the hyper-rectangle X_0 by a hyperplane through its center and orthogonal to the axis x_j where the total absolute change of the approximation error measured both at the facet centers of X_0 and the vertices of X_0^d is maximal.

It has been shown in Grancharova & Johansen (2002) that the use of such heuristics reduces the complexity of the partition significantly.

The complexity is further reduced by implementing control input trajectory parameterization as it is described in Tøndel & Johansen (2002). The idea is to use an input trajectory parameterization with less degrees of freedom in order to reduce the dimension of the optimization problem. The most common approach is to pre-determine the time-instants at which the control input is allowed to change (input blocking):

$$N_{change}^u = [1 \quad N_1^u \quad N_2^u \quad \dots \quad N_l^u] \quad (66)$$

The following approximate mp-QP algorithm is taken from Johansen & Grancharova (2003):

Algorithm 3 (approximate mp-QP):

- Step 1.* Initialize the partition to the whole hyper-rectangle, i.e. $P = \{X\}$. Mark the hyper-rectangle X as unexplored.
- Step 2.* Select any unexplored hyper-rectangle $X_0 \in P$. If no such hyper-rectangle exists, go to step 8.
- Step 3.* Compute the solution to the QP (38)–(39) for x fixed to each of the 2^n vertices of the hyper-rectangle X_0 . If all QPs have a feasible solution, go to step 5. Otherwise, go to step 4.
- Step 4.* Compute the size of X_0 using some metric. If it is smaller than some given tolerance, mark X_0 infeasible and explored. Go to step 2. Otherwise, go to step 7.
- Step 5.* Compute an affine state feedback \hat{z}_0 using Lemma 1, as an approximation to be used in X_0 . If no feasible solution was found, go to step 7.
- Step 6.* Compute the error bound ε_0 in X_0 . If $\varepsilon_0 \leq \bar{\varepsilon}$, mark X_0 as explored and feasible and go to step 2.
- Step 7.* Split the hyper-rectangle X_0 into two hyper-rectangles X_1 and X_2 by applying the heuristic splitting rule. Mark them unexplored, remove X_0 from P , add X_1 and X_2 to P , and go to step 2.
- Step 8.* If necessary, split the hyper-rectangles containing the origin such that $z^*(x) = 0$ is optimal everywhere in these hyper-rectangles. Terminate.

This algorithm will terminate with a PWL function that is an approximation to the PWL exact solution and is defined on an inner approximation \underline{X}_f of the set $X \cap X_f$. The set \underline{X}_f is represented as a union of hyper-rectangles.

In Grancharova *et al.* (2003), the approximate explicit MPC approach has been experimentally tested on a two-input two-output laboratory gas-liquid separation plant. The approach achieves performance close to that of conventional MPC, but requires only a fraction of the real-time computational machinery, leading to fast and reliable computations.

A convex nonlinear extension of the approximate explicit MPC approach has recently been developed in Johansen (2004).

3.4. Stability of the PWL approximate solution

It is shown in Johansen & Grancharova (2003) that under some assumptions on the terminal set Ω and the tolerance ε the approximate explicit MPC will make the origin asymptotically stable.

Let Γ be the largest hyper-rectangle containing the origin in its interior where the

solution computed by the approximate explicit MPC is $u^*(x) = Kx$, i.e. exactly the unconstrained LQR feedback. It is straightforward to show that Algorithm 3 leads to a non-empty Γ due to step 8. Let the terminal set Ω be the maximal output admissible set (Gilbert & Tan, 1991) for the linear system $x(t+1) = (A+BK)x(t)$ contained in the polyhedral set:

$$\Psi = \{x \in \Gamma \mid u_{\min} \leq Kx \leq u_{\max}, y_{\min} \leq Cx \leq y_{\max}\} \quad (67)$$

The set Ω is a polyhedron with a finite number of facets and can be easily computed, since $A+BK$ is Hurwitz and Γ is bounded because X is bounded (Gilbert & Tan, 1991). The stability result is formulated in the following theorem (Johansen & Grancharova, 2003):

Theorem 5:

Consider the mp-QP problem (38)–(39) with $H > 0$ defined on a hypercube X such that $X_f \subseteq X$. Define $\Sigma = Q + K^T R K$, assume $\Sigma > 0$, and let γ be the largest positive number for which the ellipsoid $E = \{x \in X_f \mid x^T \Sigma x \leq \gamma\}$ is contained in Ω . Moreover, assume the tolerance $\bar{\epsilon}$ satisfies:

$$0 < \bar{\epsilon} \leq \frac{\gamma + x_0^T \Sigma x_0}{2} \quad (68)$$

where $x_0 = \arg \min_{x \in X_0} x^T \Sigma x$. Then the approximate explicit MPC computed by

Algorithm 3 in closed loop with (17) makes the origin asymptotically stable for all $x(0) \in X_f$, and the state and input trajectories are feasible. ■

3.5. Example

Consider the double integrator from section 2.3.2. With horizon $N=10$ and $\gamma=0.13$, Algorithm 3 gives the quad-tree partition in Figure 6 with 214 regions. The

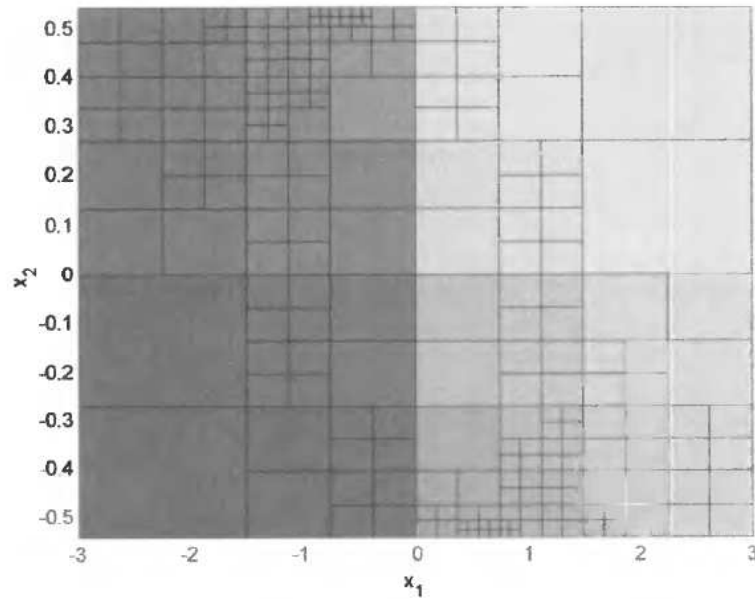


Figure 6. Quad-tree partition for the double integrator with $N=10$.

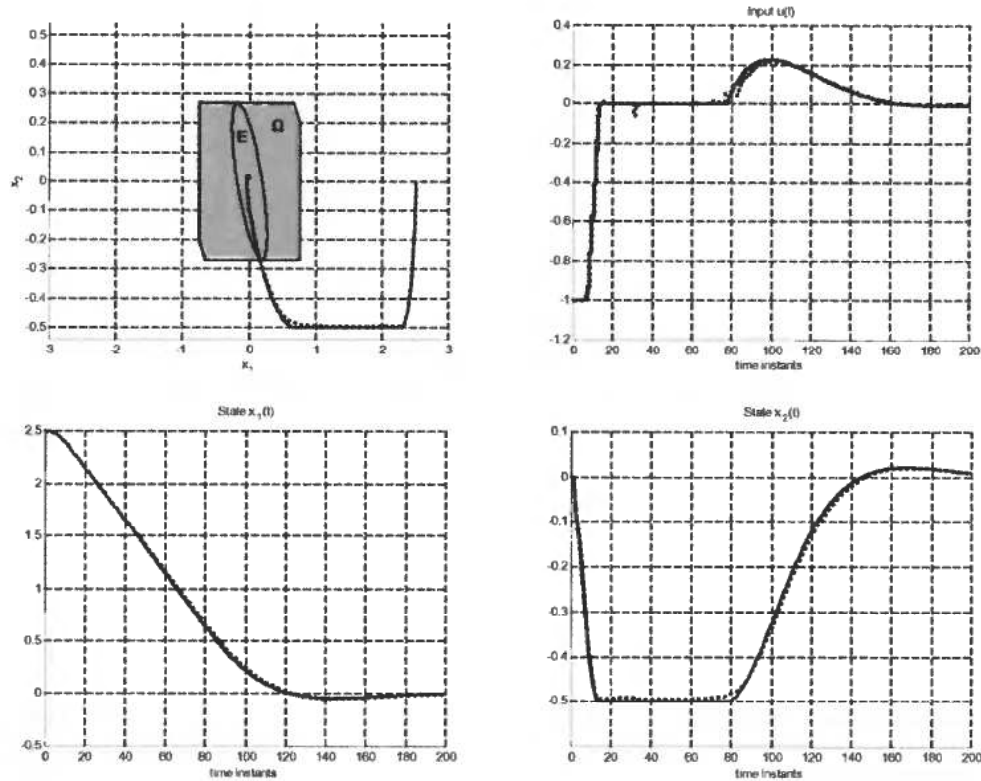


Figure 7. Top, left: Sets E and Ω . Top, right: Control input. Bottom, left: State x_1 . Bottom, right: State x_2 .

method based on the cost function approximation error (Johansen & Grancharova, 2003) is used. The sets E and Ω , together with the control and state trajectories obtained with the exact and the approximate approaches are shown in Figure 7. The solid and dashed curves show an exact and approximate trajectory, respectively. We observe that the discrepancy between them is negligible.

The k - d tree partition for the double integrator, obtained by applying the method based on the control input approximation error (Grancharova & Johansen, 2002) and using the heuristic splitting rule, is given in Figure 8. It can be seen that a significant reduction in complexity is achieved (the state space partition has 97 regions). The relative tolerance in control input error is $\sqrt{\bar{\epsilon}_r} = 0.5$.

It is interesting to compare the structure of the partitions of the approximate PWL explicit MPC feedback laws with the partitions of the exact PWL explicit MPC feedback law, as shown in Figure 4 for the case of horizon $N=10$. In parts of the state space where the exact partition contains several smaller regions while the approximate partition contains only a few large regions, the explanation is that the approximate approach only considers the first sample of the control input and is able to reduce complexity. In parts of the state space where the opposite is true, i.e. the approximate partition is more complex, this is due to a structural mismatch because the orthogonality of the hyperplanes of the approximate partition is enforced.

The exact partition in Figure 4 contains 191 polyhedral regions and is thus of comparable complexity to the approximate partitions. Still, it is clear that there will

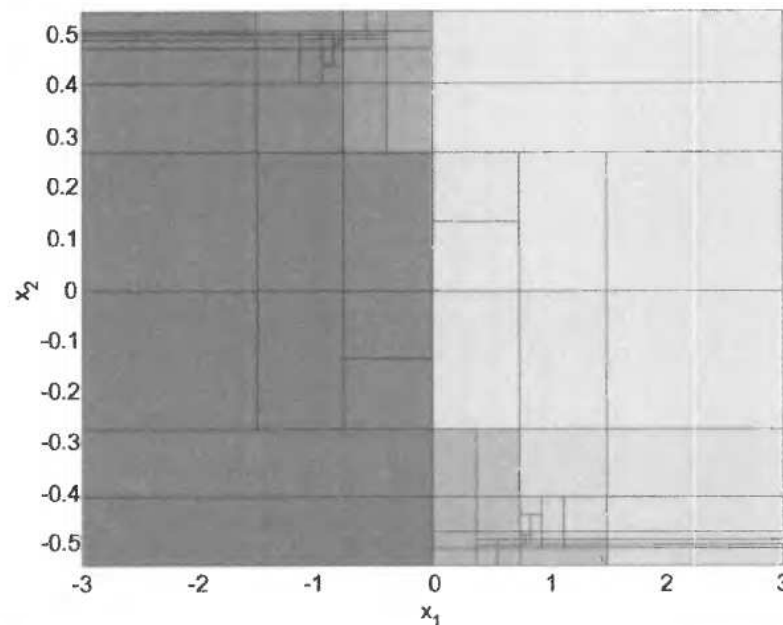


Figure 8. $k-d$ tree partition for the double integrator with $N=10$.

be significantly higher demand for real-time processing capacity and computer memory, since all hyperplanes in the partition are different and they are not orthogonal. This also holds if a search tree is constructed from the exact partition as proposed in Tøndel & Johansen (2002). In this case there will be 9 levels in the tree and 60 arithmetic operations are required to compute the exact solution, while about 3150 numbers must be stored in real-time computer memory. With the suggested approach, 18 arithmetic operations are sufficient, and only about 700 numbers must be stored for the partition with 97 regions. Of course, the price to be paid for this complexity reduction is an approximation error.

As in Johansen & Grancharova (2002) we remark that there is a significant difference between the exact and approximate approaches when the complexity of the partition is viewed as a function of the horizon. While the number of regions with the exact approach seems to give a very rapid growth with N , (Tøndel *et al.*, 2003), the approximate approach gives a partition complexity that is almost independent of the horizon N . One reason for this is that in the approximate approach it is taken into account that we only need the first sample of the input trajectory in order to implement the MPC.

3.6. Robust approximate explicit model predictive control in the presence of bounded disturbances

Some of the exact mp-QP approaches have been further extended to ensure robustness of the explicit MPC controllers against disturbances (Bemporad *et al.*, 2001; Kakalis *et al.*, 2002; Kerrigan & Maciejowski, 2003; Sakizlis *et al.*, 2004). In Kakalis *et al.* (2002) it is assumed that the disturbance input belongs to a compact polyhedral set, and the approach in Grossmann *et al.* (1983) is applied to ensure feasible operation of the MPC controller that minimizes the nominal value of the

performance index. This work has been further extended to proportional integral controllers (Sakizlis *et al.*, 2004). In Bemporad *et al.* (2001), an approach to explicit solution of robust MPC problems based on a min-max formulation with a performance index expressed in ∞ -norm has been proposed. It has to be mentioned however, that solution obtained by optimizing the worst value of the performance criterion can be quite conservative. In Kerrigan & Maciejowski (2003), it is supposed that the uncertainty set is a polytope and it is described how a class of uncertain quadratic and linear optimization problems can be converted to a multi-parametric quadratic programming (mp-QP) or multi-parametric linear programming (mp-LP) problems by solving as many linear programs (LPs) as there are constraints in the optimization problem without uncertainty. It is also shown in Kerrigan & Maciejowski (2003) that if the uncertainty set is given by upper and lower bounds only, then this transformation can be done by simply computing the 1-norms of the rows of the matrix by which the uncertainty enters the constraint.

In this section, an approximate mp-QP approach to explicit solution of constrained linear MPC problems in the presence of bounded disturbances is described (Grancharova *et al.*, 2003). It is based on an orthogonal search tree structure of the state space partition and thus represents an extension of the approximate mp-QP approach Johansen & Grancharova (2003). Like in Kakalis *et al.* (2002), the explicit MPC controller avoids conservativeness by minimizing the nominal value of the performance index and it is robust in the sense that all constraints are satisfied for all possible disturbance realizations within the specified range. Here we consider the case where the set of the disturbance inputs represents a hyper-rectangle that includes the origin in its interior. Based on this assumption, the conditions which guarantee feasible operation of the MPC controller are derived in a way similar to that in Kerrigan & Maciejowski (2003) and the original mp-QP problem with disturbance input is converted into an mp-QP problem without disturbances.

Problem formulation:

Consider the linear discrete-time system:

$$x(t+1) = Ax(t) + Bu(t) + T\theta(t) \quad (69)$$

$$y(t) = Cx(t)$$

where $x(t) \in \mathcal{R}^n$, $u(t) \in \mathcal{R}^m$, and $y(t) \in \mathcal{R}^p$ are the state, input and output variable, $\theta(t)$ is the disturbance input that is assumed to belong to a bounded polyhedral set $\theta(t) \in \Theta^A \subset \mathcal{R}^s$. Also, $A \in \mathcal{R}^{n \times n}$, $B \in \mathcal{R}^{n \times m}$, $C \in \mathcal{R}^{p \times n}$ and $T \in \mathcal{R}^{n \times s}$. Let $\Theta \equiv [\theta_1^T, \dots, \theta_{t+N-1}^T]^T \in \Theta^B$ is a disturbance realization, with $\Theta \in \Theta^B = \{\Theta^A \times \Theta^A \dots \times \Theta^A\} \subset \mathcal{R}^{sN}$. It is assumed that a full measurement of the state $x(t)$ is available at the current time t . Then, for the current $x(t)$, MPC solves the optimization problem:

$$V^*(x(t), \Theta) = \min_{U \equiv \{u_t, \dots, u_{t+N-1}\}} J(U, x(t), \Theta) \quad (70)$$

subject to $x_{t|t} = x(t)$ and:

$$y_{\min} \leq y_{t+k|t} \leq y_{\max}, \quad k=1, \dots, N \quad (71)$$

$$u_{\min} \leq u_{t+k} \leq u_{\max}, \quad k=0, 1, \dots, N-1 \quad (72)$$

$$x_{t+N|t} \in \Omega \quad (73)$$

$$x_{t+k+1|t} = Ax_{t+k|t} + Bu_{t+k} + T\theta_{t+k}, \quad \theta_{t+k} \in \Theta^A, \quad k \geq 0 \quad (74)$$

$$y_{t+k|t} = Cx_{t+k|t}, \quad k \geq 0 \quad (75)$$

with the cost function given by:

$$J(U, x(t), \Theta) = \sum_{k=0}^{N-1} [x_{t+k|t}^T Q x_{t+k|t} + u_{t+k}^T R u_{t+k}] + x_{t+N|t}^T P x_{t+N|t} \quad (76)$$

and symmetric $R > 0$, $Q \geq 0$. We assume (A, B) is stabilizable, (A, \sqrt{Q}) is observable, Ω is a polyhedral terminal set, and the final cost matrix $P > 0$ is the solution of the associated algebraic Riccati equation. It is also assumed $u_{\max} > 0 > u_{\min}$, $y_{\max} > 0 > y_{\min}$, such that the origin is an interior point in the feasible set $X_f = \{x(t) \in \mathbb{R}^n | \exists U \text{ satisfying (71)–(75)}\}$. Here, we consider the nominal optimization criterion:

$$V_{nom}^*(x(t)) = \min_{U \equiv \{u_t, \dots, u_{t+N-1}\}} J(U, x(t), \theta^N) \quad (77)$$

corresponding to $\theta(t) = \theta^N = 0$, where θ^N is the nominal value of the disturbance input. In this problem formulation, the *robustness* is defined in terms of satisfaction of the output and input constraints (71) and (72) under all possible disturbance realizations $\Theta \in \Theta^B$ that influence the state of the system (equation (74)).

By substituting:

$$x_{t+k|t} = A^k x(t) + \sum_{j=0}^{k-1} A^j B u_{t+k-1-j} + \sum_{j=0}^{k-1} A^j T \theta_{t+k-1-j} \quad (78)$$

in the constraints (71)–(75), they can be represented in the form:

$$GU \leq W + E_1 x(t) + E_2 \Theta, \quad \forall \Theta \in \Theta^B \quad (79)$$

where $U \equiv [u_t^T, \dots, u_{t+N-1}^T]^T \in \mathbb{R}^{mN}$ is the optimization vector and $\Theta \in \Theta^B$ is the disturbance realization. Then the nominal optimization criterion (77) is rewritten as:

$$V_{nom}^*(x(t)) = \frac{1}{2} x^T(t) Y x(t) + \min_U \left\{ \frac{1}{2} U^T H U + x^T(t) F U \right\} \quad (80)$$

We apply the same idea as in Kerrigan & Maciejowski (2003) of pre-stabilizing (69) with a linear state feedback gain and optimizing over a sequence of perturbations to this control law. Thus, we define:

$$U \equiv -H^{-1} F^T x(t) + z \quad (81)$$

where $z \in \mathbb{R}^{mN}$ is the control input perturbation. Then, the optimization problem (80) subject to constraint (79) is transformed into the following mp-QP problem:

$$V_{z,nom}^*(x) = \min_z \frac{1}{2} z^T H z \quad (82)$$

subject to:

$$Gz \leq W + S_1 x(t) + S_2 \Theta, \quad \forall \Theta \in \Theta^B \quad (83)$$

Assumption 1:

The disturbance input set:

$$\Theta^A = \{\theta \in \mathcal{R}^s \mid \theta^L \leq \theta \leq \theta^U\} \quad (84)$$

represents a hyper-rectangle that includes the origin in its interior.

Definition 8:

Consider the i th constraint defined by G^i , W^i , S_1^i , S_2^i rows of the matrices G , W , S_1 , S_2 . The worst disturbance realization for the i -th constraint, denoted by $\tilde{\Theta}^i \in \Theta^B$ is one which solves the linear program:

$$S_2^i \tilde{\Theta}^i = \min_{\Theta \in \Theta^B} \{S_2^i \Theta\} \quad (85)$$

Remark:

The linear program (85) can be easily solved by exploiting the fact that the disturbance input set is a hyper-rectangle. Thus:

$$\min_{\Theta \in \Theta^B} \{S_2^i \Theta\} = \min_{\Theta \in \Theta^B} \left\{ \sum_{j=1}^{sN} S_{2,j}^i \Theta_j \right\} = \sum_{j=1}^{sN} \min_{\Theta_j^L \leq \Theta_j \leq \Theta_j^U} \{S_{2,j}^i \Theta_j\} = \sum_{j=1}^{sN} \min \{S_{2,j}^i \Theta_j^L, S_{2,j}^i \Theta_j^U\} \quad (86)$$

where $S_{2,j}^i$ is the j -th element of the row vector S_2^i , Θ_j is the j th element of the column vector of disturbance realization $\Theta \in \Theta^B \subset \mathcal{R}^{sN}$, and Θ_j^L , Θ_j^U are respectively the lower and upper bounds of Θ_j .

Lemma 2:

If there exists an affine function $z(x)$ that satisfies the following constraint:

$$Gz \leq \tilde{W} + S_1 x \quad (87)$$

where the i -th row of the matrix \tilde{W} is determined by:

$$\tilde{W}^i = W^i + S_2^i \tilde{\Theta}^i \quad (88)$$

and where $\tilde{\Theta}^i \in \Theta^B$ is the worst disturbance realization for the i -th constraint, then this implies that $z(x)$ will satisfy constraint (83) for all possible disturbance realizations $\Theta \in \Theta^B$. Such $z(x)$ is referred to as *robustly feasible*. ■

In this way, the constraint (87) which ensures robust feasibility can be easily constructed. Then, the original mp-QP problem (82)–(83) becomes:

$$V_{z,nom}^*(x) = \min_z \frac{1}{2} z^T H z \quad (89)$$

subject to:

$$Gz \leq \tilde{W} + S_1 x(t) \quad (90)$$

where \tilde{W} is determined by (88). Thus the original mp-QP problem with disturbance input (problem (82)–(83)) is reformulated as an mp-QP problem without disturbance (problem (89)–(90)) and therefore the approximate approach (Johansen & Grancharova, 2003) for explicit solution of mp-QP problems can easily be applied to this problem. It has to be stressed that the approximate approach (Johansen & Granch-

arova, 2003) guarantees that the optimal solution is feasible in sense that it will satisfy constraint (90). This directly implies by Lemma 2 above that constraint (83) of the original mp-QP problem will be satisfied for all possible disturbance realizations. This is summarized in the following Lemma:

Lemma 3 (feasible control in the presence of disturbance):

Consider the bounded polyhedron X_0 with vertices $\{v_1, v_2, \dots, v_M\}$. If K_0 and g_0 solve the QP:

$$\min_{K_0, g_0} \sum_{i=1}^M (z^*(v_i) - K_0 v_i - g_0)^T H (z^*(v_i) - K_0 v_i - g_0) \quad (91)$$

subject to:

$$G(K_0 v_i + g_0) \leq \tilde{W} + S_1 v_i, \quad i \in \{1, 2, \dots, M\} \quad (92)$$

then the least squares approximation $\hat{z}_0(x) = K_0 x + g_0$ is *robustly feasible* for the mp-QP (82)–(83) for all $x \in X_0$ and all disturbance realizations $\Theta \in \Theta^B$. ■

Example:

Consider the double integrator:

$$x(t+1) = Ax(t) + Bu(t) + T\theta(t) \quad (93)$$

with:

$$A = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} T_s^2 \\ T_s \end{bmatrix}, \quad T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (94)$$

where the sampling interval is $T_s = 0.3$. Consider the MPC problem with horizon $N = 30$. The cost matrices are $Q = \text{diag}(1, 0)$, $R = 1$, and the matrix $P > 0$ is given as the solution of the algebraic Riccati equation. The constraints are:

$$-1 \leq u \leq 1 \quad (95)$$

$$-0.5 \leq x_2 \leq 0.5 \quad (96)$$

The disturbance vector $\theta = [\theta_1 \theta_2]^T$ has the following bounds:

$$-0.01 \leq \theta_1(t) \leq 0.01 \quad (97)$$

$$-0.015 \leq \theta_2(t) \leq 0.015 \quad (98)$$

The approximation tolerance $\varepsilon > 0$ is chosen according to Theorem 5, with $\gamma = 0.1$. The state space partition of the robust approximate MPC controller is shown in Figure 9. It has 172 regions and 11 levels of search. With one scalar comparison required at each level of the k - d tree, 11 arithmetic operations are required in the worst case to determine which region the state belongs to. Totally, 15 arithmetic operations are needed in real-time to compute the control input with this MPC controller (11 comparisons, 2 multiplications and 2 additions).

In Figure 10, the sets Ω , E and S (S is the terminal region to which the state converges), and disturbance realizations with constant magnitude are given. In Figure 11, the control and state trajectories obtained with the robust MPC under

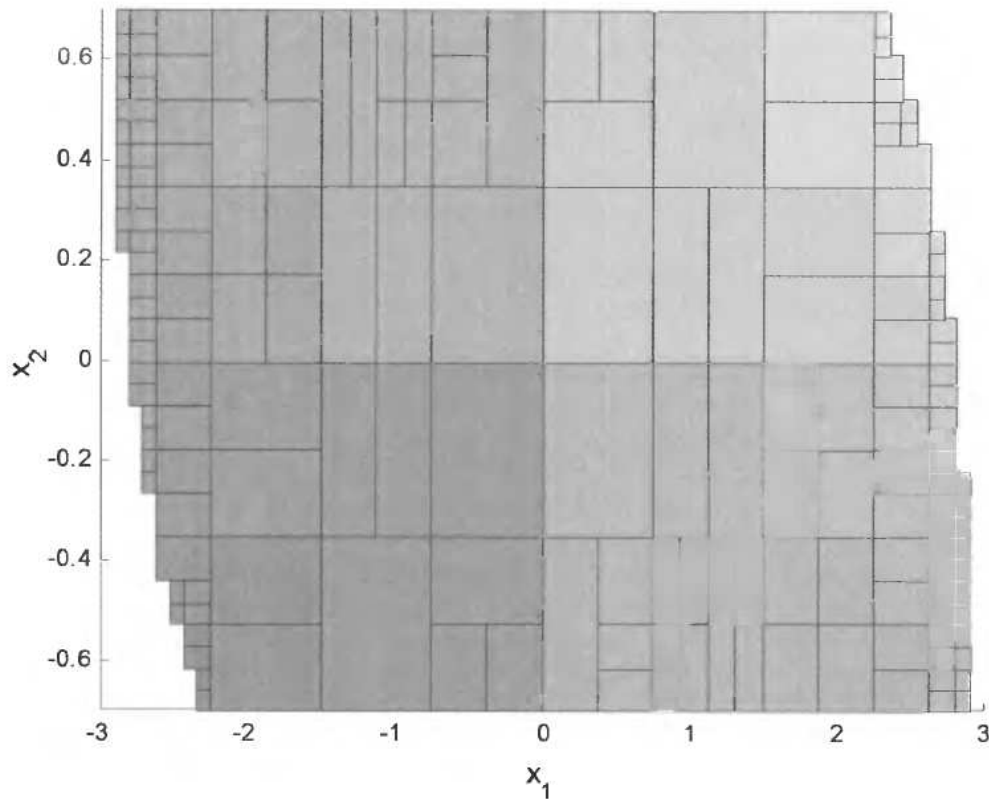


Figure 9. $k-d$ tree partition of the robust MPC.

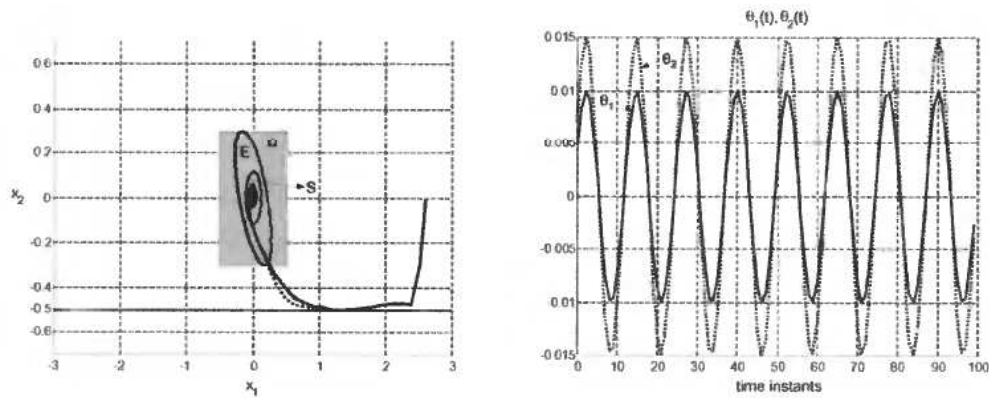


Figure 10. Left: The sets Ω , E , S , the approximate (the solid curve) and the exact (the dashed curve) state trajectories. Right: Disturbance inputs with constant magnitude.

these disturbances are shown (the trajectories with the exact mp-QP approach are given for comparison). The approximate and the exact state trajectories are also depicted in Figure 10, where it can be seen that with the increase of time the state enters and remains in the terminal region S . It can be seen from the above figures that the robust MPC keeps all constraints imposed on the system.

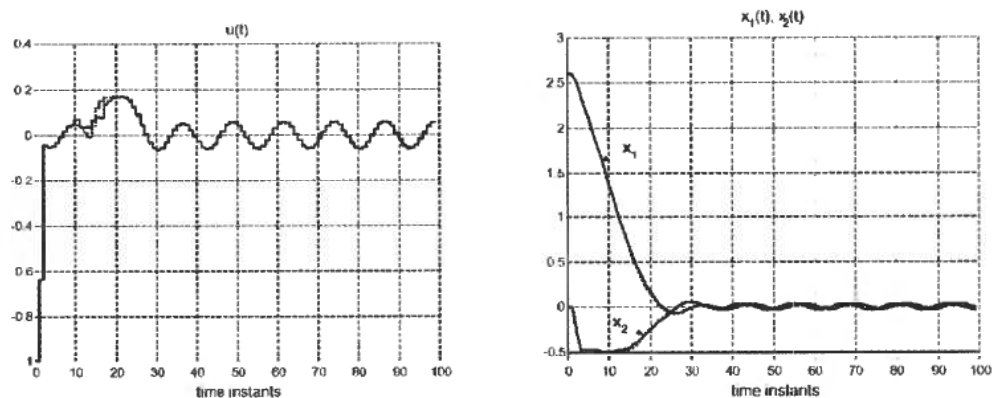


Figure 11. Control input and state trajectories for the robust MPC (the solid curves are with the approximate controller and the dashed curves are with the exact controller).

References

- BEMPORAD, A., BORRELLI, F. & MORARI, M. (2001). Robust model predictive control: Piecewise linear explicit solution, *Proceedings of European Control Conference*, Porto, Portugal, pp. 939–944.
- BEMPORAD, A. & FILIPPI, C. (2003). Suboptimal explicit RHC via approximate quadratic programming, *Optim. Theory Applicat.*, 117, pp. 5–38.
- BEMPORAD, A. & MORARI, M. (1999). Robust model predictive control: A survey, in A. GARULLI, A. TESI, & A. VICINO, eds, *Robustness in Identification and Control*, number 245 in Lecture Notes in Control and Information Sciences, Springer-Verlag, pp. 207–226.
- BEMPORAD, A., MORARI, M., DUA, V. & PISTIKOPOULOS, E. N. (2000). The explicit solution of model predictive control via multiparametric quadratic programming, *Proceedings of the American Control Conference*, Chicago, Illinois, pp. 872–876.
- BEMPORAD, A., MORARI, M., DUA, V. & PISTIKOPOULOS, E. N. (2002). The explicit linear quadratic regulator for constrained systems, *Automatica*, 38, pp. 3–20.
- BENTLEY, J. L. (1975). Multidimensional binary search trees used for associative searching, *Communications of the ACM*, 18, pp. 509–517.
- CHMIELEWSKI, D. & MANOUSIOUTHAKIS, V. (1996). On constrained infinite-time linear quadratic optimal control, *Systems & Control Letters*, 29(3), pp. 121–130.
- GILBERT, E. G. & TAN, K. T. (1991). Linear systems with state and control constraints: The theory and application of maximal output admissible sets, *IEEE Trans. Automatic Control*, 36, pp. 1008–1020.
- GRANCHAROVA, A. & JOHANSEN, T. A. (2002). Approximate explicit model predictive control incorporating heuristics, *Proceedings of IEEE International Symposium on Computer Aided Control System Design*, Glasgow, Scotland, UK, pp. 92–97.
- GRANCHAROVA, A. & JOHANSEN, T. A. (2003). Design of robust explicit model predictive controller via orthogonal search tree partitioning, *Proceedings of European Control Conference*, Cambridge, UK.
- GRANCHAROVA, A., JOHANSEN, T. A. & KOCLIAN, J. (2003). Explicit model predictive control of gas-liquid separation plant, *Proceedings of European Control Conference*, Cambridge, UK.
- GROSSMANN, I. E., HALEMANE, K. P. & SWANEY, R. E. (1983). Optimization strategies for flexible chemical processes, *Computers and Chemical Engineering*, 7, pp. 439–462.
- JOHANSEN, T. A. (2004). Approximate explicit receding horizon control of constrained nonlinear systems, *Automatica*, 40, pp. 293–300.
- JOHANSEN, T. A. & GRANCHAROVA, A. (2002). Approximate explicit model predictive control implemented via orthogonal search tree partitioning, *Proceedings of 15th IFAC World Congress, Barcelona, Spain*, session T-We-M17.
- JOHANSEN, T. A. & GRANCHAROVA, A. (2003). Approximate explicit constrained linear model predictive control via orthogonal search tree, *IEEE Trans. Automatic Control*, 48, pp. 810–815.

- JOHANSEN, T. A., PETERSEN, I. & SLUPPHAUG, O. (2002). Explicit sub-optimal linear quadratic regulation with state and input constraints, *Automatica*, 38, pp. 1099–1111.
- KAKALIS, N. M. P., DUA, V., SAKIZLIS, V., PERKINS, J. D. & PISTIKOPOULOS, E. N. (2002). A parametric optimisation approach for robust MPC, *Proceedings of 15th IFAC World Congress*, Barcelona, Spain.
- KERRIGAN, E. C. & MACIEJOWSKI, J. M. (2003). On robust optimization and the optimal control of constrained linear systems with bounded state disturbances, *Proceedings of European Control Conference*, Cambridge, UK.
- KWAKERNAAK, H. & SIVAN, R. (1972). *Linear optimal control systems*, Wiley-Interscience, New York.
- MAYNE, D. Q., RAWLINGS, J. B., RAO, C. V. & SCOKAERT, P. O. M. (2000). Constrained model predictive control: Stability and optimality, *Automatica*, 36, pp. 789–814.
- NOCEDAL, J. & WRIGHT, S. (1999). *Numerical Optimization*, Springer-Verlag New York, Inc.
- PISTIKOPOULOS, E. N., DUA, V., BOZINIS, N. A., BEMPORAD, A. & MORARI, M. (2000). On-line optimization via off-line parametric optimization tools, *Computers & Chemical Engineering*, 24, pp. 183–188.
- SAKIZLIS, V., KAKALIS, N. M. P., DUA, V., PERKINS, J. D. & PISTIKOPOULOS, E. N. (2004). Design of robust model-based controllers via parametric programming, *Automatica*, 40, pp. 189–201.
- SCOKAERT, P. O. M. & RAWLINGS, J. B. (1998). Constrained linear quadratic regulation, *IEEE Trans. Automatic Control*, 43, pp. 1163–1169.
- SERON, M., DE DONA, J. A., & GOODWIN, G. C. (2000). Global analytical model predictive control with input constraints, *Proceedings of IEEE Conf. Decision and Control*, Sydney, TuA05-2.
- SZNAIER, M. & DAMBORG, J. (1987). Suboptimal control of linear systems with state and control inequality constraints, *Proceedings of 26th IEEE Conference on Decision and Control*, 1, pp. 761–762.
- TØNDEL, P. & JOHANSEN, T. A. (2002). Complexity reduction in explicit linear model predictive control, *Proceedings of 15th IFAC World Congress, Barcelona, Spain*, session T-We-M17.
- TØNDEL, P., JOHANSEN, T. A. & BEMPORAD, A. (2003). An algorithm for multi-parametric quadratic programming and explicit MPC solutions, *Automatica*, 39, pp. 489–497.