PERGAMON

Automatica 38 (2002) 879–884

automatica

www.elsevier.com/locate/automatica

Technical Communique

Who needs QP for linear MPC anyway?☆

B. Kouvaritakis[a,*], M. Cannon[a], J.A. Rossiter[b]

[a] *Department of Engineering Science, Oxford University, Parks Road, Oxford OX1 3PJ, UK*
[b] *Department of Mathematical Sciences, Loughborough University, Loughborough, Leic. LE11 3TP, UK*

Received 5 April 2000; received in revised form 11 June 2001; accepted 26 October 2001

Abstract

Conventional MPC uses quadratic programming (QP) to minimise, on-line, a cost over n linearly constrained control moves. However, stability constraints often require the use of large n thereby increasing the on-line computation, rendering the approach impracticable in the case of fast sampling. Here, we explore an alternative that requires a fraction of the computational cost (which increases only linearly with n), and propose an extension which, in all but a small class of models, matches to within a fraction of a percent point the performance of the optimal solution obtained through QP. The provocative title of the paper is intended to point out that the proposed approach offers a very attractive alternative to QP-based MPC. © 2002 Published by Elsevier Science Ltd.

Keywords: Predictive control; Constraints; Optimization; Quadratic programming; Invariant sets

1. Introduction

Stability constraints (e.g. Lee, Kwon, & Choi, 1998) provide a convenient means of ensuring stability in model-based predictive control (MPC); n free control moves are deployed to steer the predicted state into invariant sets while optimising performance. However, computational complexity becomes an issue, especially when n needs to be large, or it is necessary to propagate the effects of uncertainty over the horizon of the n moves (Kouvaritakis, Rossiter, & Schuurmans, 2000). It is possible to deploy a polyhedral partition of the state-space in order to trade off on-line complexity with off-line computation (Bemporad, Morari, Dua, & Pistikopoulos, 2000; Seron, De Dona, & Goodwin, 2000). However: (i) the off-line computation requires the solution of an NP-hard multi-parametric QP (mp-QP), and this implies that for large n the approach can only be used when all setpoint changes are known in advance; (ii) the upper bound on the number of required partitions grows exponentially with n and thus can raise inordinate memory demands; (iii) the approach cannot handle model uncertainty. Alternatives which avoid these difficulties (albeit at the expense of optimality) exist, e.g. input blocking which constraints the control input to have the same value for several consecutive samples, or (Wredenhagen & Belanger, 1994) which is based on an on-line tuning of a state feedback controller, or (Kouvaritakis et al., 2000) which is based on the closed-loop paradigm (thereby allowing for a wider prediction class) and associated autonomous state-space formulation. The free moves in (Kouvaritakis et al., 2000) are introduced as perturbations c on an optimal closed-loop system; c is used merely to avoid constraint violations, but otherwise c should be kept as small as possible. An augmentation of state that introduces c as part of the new state of an autonomous formulation allows stability constraints to be applied at current time, rather than at the end of a prediction horizon. This, combined with the use of ellipsoidal sets (Section 2) results in an on-line optimisation that is univariate and can be solved trivially (Kouvaritakis et al., 2000).

This note explores (Section 3) the properties of the univariate optimisation and shows that techniques such as Newton–Raphson can be used efficiently and are guaranteed to converge. Furthermore, extensions are developed (Section 4) to allow the optimisation to search outside the ellipsoidal set thereby removing most of the conservativism incurred through use of ellipsoidal rather than

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Rick Middleton under the direction of Editor Paul Van den Hof.

* Corresponding author. Tel.: +44-1865-273-100; fax: +44-1865-273-906.

E-mail addresses: basil.kouvaritakis@eng.ox.ac.uk (B. Kouvaritakis), mark.cannon@eng.ox.ac.uk (M. Cannon), j.a.rossiter@lboro.ac.uk (J.A. Rossiter).

0005-1098/02/$ - see front matter © 2002 Published by Elsevier Science Ltd.
PII: S0005-1098(01)00263-1

maximal polyhedral sets. Monte Carlo simulations demonstrate (Section 5) that this extension, in all but a small class of models, matches the performance of QP-based MPC to within a percent point yet achieves that at a fraction of the computational cost. Unlike QP-based MPC whose computational complexity increases significantly with n, that of the new algorithm grows only linearly with n. This can be a very significant advantage in fast sampling applications for which QP-based MPC becomes intractable whenever n has to be chosen to be large. Thus, the new algorithm provides a very attractive alternative to the currently available MPC algorithms.

2. Earlier work

Consider the system model, control law and system constraints

$$x_{k+1} = Ax_k + Bu_k, \tag{1a}$$

$$y_k = Cx_k, \tag{1b}$$

$$u_k = -Kx_k, \tag{1c}$$

$$Hx_k - h \leqslant 0, \tag{1d}$$

$$H = \begin{bmatrix} D \\ -FK \end{bmatrix}, \tag{1e}$$

$$h = \begin{bmatrix} d \\ g \end{bmatrix}, \tag{1f}$$

where $x \in R^p,\ u \in R^m, y \in R^l$, and in the absence of system constraints K minimises the infinite horizon LQ cost, J_{LQ}, with weights $Q > 0, R > 0$. To guarantee stability in MPC it is convenient to deploy "dual mode" predictions: (i) MODE 1 comprises the first n control moves, $u_{k+i}, i = 0, \ldots, n-1$, which are free; (ii) MODE 2 adopts the LQ control law of (1c) from the nth prediction onwards. The relevant MPC, referred to as QPMPC, solves on-line a QP which minimises a finite horizon cost with a terminal weight term, subject to (1d) and for n sufficiently large (Scokaert & Rawlings, 1998) minimises J_{LQ}. MODE 2 is required to be feasible (i.e. the control moves in MODE 2 must satisfy constraints (1d)) and this is equivalent to the stability constraint that x_{k+n} lies in the maximal admissible (Gilbert & Tan, 1991) polytopic set, P_0. Instead, it is possible to use ellipsoidal sets, but membership of these sets implies quadratic constraints which in turn require the on-line use of semi-definite programming (SDP); the relevant MPC will be referred to as SDMPC. Ellipsoidal sets are subsets of P_0 and hence the stability constraint of SDMPC will in general require the use of larger horizons n. This, compounded by the more demanding nature of SDP renders SDMPC a less attractive proposition.

Computational complexity can be reduced dramatically through the use of the "closed-loop" paradigm. According to this (1c) is replaced by

$$u_{k+i} = -Kx_{k+i} + c_i, \quad c_i = 0 \quad \forall i > n, \tag{2}$$

where the predicted sequence $c_i,\ i = 1, \ldots, n$ represents perturbations on the LQ optimal state feedback system. For $\{c_i\} = 0$ the closed-loop system is optimal in the unconstrained LQ sense, but may violate constraints (1d); $\{c_i\}$ is the instrument through which to achieve feasibility, and otherwise should be kept as small as possible. This suggests the alternative (but equivalent) on-line optimisation problem of minimising (Kouvaritakis et al., 2000) $J_f = f_k^{\mathrm{T}} f_k$, where f_k is a block vector with elements $Wc_i,\ i = 1, \ldots, n$; $W = B^{\mathrm{T}} P_{\mathrm{LYAP}} B + R$, where P_{LYAP} is the solution of the Lyapunov equation $P_{\mathrm{LYAP}} - (A - BK)^{\mathrm{T}} P_{\mathrm{LYAP}} (A - BK) = Q + K^{\mathrm{T}} RK$ and where Q, R are the usual weights in the LQ cost. [1] Thus, it is possible to reformulate the problem so that membership of an invariant set is invoked *at the current time* rather than n steps ahead; propagating the effects of uncertainty over n future steps is computationally intensive and leads to conservative results. One way to avoid this is to remove the degrees of freedom c_i but allow for time-varying K (Kothare, Balakrishnan, & Morari, 1996). A much more convenient/optimal way is via the autonomous formulation of the prediction dynamics $z_{k+1} = \Psi z_k$, $z = [x^{\mathrm{T}} \ \ f^{\mathrm{T}}]^{\mathrm{T}}$ (for details see Kouvaritakis et al., 2000). The ellipsoidal set and corresponding invariance condition

$$S_z = \{z \,|\, z^{\mathrm{T}} P z \leqslant 1\}, \quad \Psi^{\mathrm{T}} P \Psi \leqslant P,$$

$$P^{\mathrm{T}} = P, \quad P > 0, \quad \Psi = \begin{bmatrix} A - BK & BE \\ 0 & \mathrm{T} \end{bmatrix} \tag{3}$$

then imply that the predictions $z_{k+i},\ \forall i \geqslant 1$ lie inside S_z whenever $z_k \in S_z$; E is a row block matrix of zeros with the first block being W^{-1} and T is a block matrix of zeros with the identity appearing along its first superdiagonal blocks. To ensure feasibility, it is possible to follow the analysis of Kothare, Balakrishnan, and Morari (1996) in conjunction with (3) to get the sufficient condition:

$$\begin{aligned} |H_j^{\mathrm{T}} z|^2 &= |H_j^{\mathrm{T}} P^{-1/2} P^{1/2} z|^2 \\ &\leqslant H_j^{\mathrm{T}} P^{-1} H_j z^{\mathrm{T}} P z \\ &\leqslant H_j^{\mathrm{T}} P^{-1} H_j \leqslant h_j^2, \end{aligned} \tag{4}$$

where $H_j^{\mathrm{T}},\ h_j$ denote the jth row and element of H, h; H has to be suitably augmented (Kouvaritakis et al., 2000) to account for the inclusion of f as part of z, namely to account for the fact that $u = -Kx$ is replaced by $u = -Kx + c = [-K \ \ E]z$. Thus, for $P > 0$ satisfying (3) and (4), one need no longer consider constraints (1d) explicitly, but rather perform the optimisation

$$\min_{f_k} \quad f_k^{\mathrm{T}} f_k, \tag{5a}$$

$$\text{s.t.} \quad z_k^{\mathrm{T}} P z_k = x_k^{\mathrm{T}} P_{11} x_k + 2 f_k^{\mathrm{T}} P_{21} x_k + f_k^{\mathrm{T}} P_{22} f_k \leqslant 1, \tag{5b}$$

[1] The equivalence between J_{LQ} and J_f follows from the fact that from (2) and for $\underline{c}_k$ being the block vector of predicted c_i, we have $J_{\mathrm{LQ}} = \underline{c}_k^{\mathrm{T}} S \underline{c}_k + \beta^{\mathrm{T}} \underline{c}_k + \gamma$ where γ is a constant and where β must be zero, given that the unconstrained optimal solution for $\underline{c}_k$ is known to be zero; simple algebra then gives that $S = diag\{W\}$.

where P_{ij} are partitions of P conformal to the partition of z. For given x_k, (5b) defines an ellipsoid, $S_f(x_k)$, in f_k. If $0 \in S_f(x_k)$ then $f_k = 0$ is optimal, otherwise the optimum is (Kouvaritakis et al., 2000):

$$f = \lambda M P_{21} x_k,$$
$$\phi(\lambda) = x_k^{\mathrm{T}} P_{12}[M P_{22}^{-1} M - P_{22}^{-1}] P_{21} x_k + x_k^{\mathrm{T}} P_{11} x_k - 1 = 0; \quad M = (I - \lambda P_{22})^{-1}, \tag{6}$$

where λ is the unique negative real root of ϕ. The solution of (6) is trivial (especially so if one exploits the eigenvalue/vector properties of P_{22} which can be computed off-line), irrespective of whether (1) is subject to uncertainty or not. However, replacing the polytopic constraints with the ellipsoidal constraint (5b) has the effect of reducing the set of initial conditions over which (5) can be applied. A partial remedy is to choose P so as to maximise the volume of S_x, the projection of S_z onto x-space; this constitutes a convex optimisation problem (Boyd, Ghaoui, Feron, & Balakrishnan, 1996). A further remedy is to increase the number of degrees of freedom n; as discussed below this can be implemented at minimal extra computational cost. Yet another solution is to deploy (whenever necessary) "triple mode" predictions (Rossiter, Kouvaritakis, & Cannon, 2000) in the sense that some of the degrees of freedom can be reserved for the purpose of driving x into S_x.

3. Properties of the on-line optimisation

Lemma 3.1. *From the definition of M Eqn.* (6) *it follows that*

$$\frac{\mathrm{d}M^i}{\mathrm{d}\lambda} = iM^{i+1}P_{22} \quad \forall i \geqslant 1. \tag{7}$$

Proof. This follows from $\mathrm{d}M^i/\mathrm{d}\lambda \equiv iM^{i-1}\mathrm{d}M/\mathrm{d}\lambda \equiv iM^{i-1}(-1)M(-P_{22})M = iM^{i+1}P_{22}$. □

Theorem 3.1. *For $\lambda \leqslant 0$ all the derivatives of the function $\phi(\lambda)$ of* (6) *are positive.*

Proof. Using Lemma 3.1 and a process of induction it is easy to show that

$$\frac{\mathrm{d}^i\phi(\lambda)}{\mathrm{d}\lambda^i} = (i+1)!x_k^{\mathrm{T}}P_{12}R\Lambda^{i+1}(I - \lambda\Lambda^{i+2})R^{\mathrm{T}}P_{21}x_k, \tag{8}$$

where R, Λ are defined by the eigenvalue/vector decomposition of $P_{22} = R\Lambda R^{\mathrm{T}}$. Clearly, the RHS of (8) is positive for all $\lambda \leqslant 0$ given that $P = P^{\mathrm{T}}$ and $P > 0$ from which $\Lambda > 0$. □

Thus the Newton–Raphson (NR) method is guaranteed to converge (quadratically) to the solution of (6) when initialised at $\lambda = 0$. On-line use of NR can be used to replace the QP of QPMPC or the SDP of SDMPC to give the very efficient "Newton–Raphson MPC" (NRMPC). It is easy to show that, providing use is made of the eigenvalue/vector decomposition of P_{22} (which can be computed off-line), the computational burden of solving (6) grows linearly with n.

4. The extension of the NRMPC

The numerical advantages of NRMPC over QPMPC or SDMPC are obvious given that the NRMPC complexity grows only linearly with n. This follows from the fact that M and P_{22}^{-1} share a common eigenvalue/vector decomposition which can be performed once off-line and does not need to be re-evaluated for the different values of λ. Therefore for fast sampling applications and/or applications for which feasibility considerations require large n, NRMPC is an attractive alternative to the usual MPC based on QP or SDP. Were it to be the case that the degree of optimality of NRMPC matched that of QPMPC or SDMPC there would be little reason to ever use QPMPC or SDMPC. However, QPMPC may still be the algorithm of choice because it uses (implicitly) as its target set the maximal admissible set and thus provides the optimal solution (assuming that on-line optimisation is possible for n large enough). Here, we propose an extension to NRMPC which reduces sub-optimality to such an extent that it makes it the algorithm of choice for all but very slow sampling applications.

Sub-optimality in NRMPC is due to the constraint that $z_k \in S_z$ (i.e. $f_k \in S_f(x_k)$). Clearly there are $f_k \notin S_f(x_k)$ which: (i) satisfy constraints (1d) over the prediction horizon n; and (ii) place x_{k+n} in the maximal admissible set P_o, thereby guaranteeing feasibility from the $(k+n)$th prediction instant onwards. Since both u_{k+i} $i = 0, \ldots, n-1$ and x_{k+n} depend affinely on f_k, it is clear that conditions (i) and (ii) above imply a set of linear inequalities on f_k. Thus for feasibility f_k must lie in the polyhedral set $\Pi_n(x_k)$ defined by these linear inequalities. It follows that $S_f(x_k) \subseteq \Pi_n(x_k)$ and thus the NRMPC choice for f_k, say f_k^*, will not necessarily be the feasible solution of minimum norm. Thus, what is needed is a low complexity extension to NRMPC that enables the algorithm to search for smaller f_k outside $S_f(x_k)$ while retaining feasibility and recovering optimality (i.e. $f = 0$) whenever possible. Here, we exploit scaling (i.e. $f_k = \mu f_k^*$, $\mu \leqslant 1$) subject to the constraint that the "feasibility now" implies "feasibility at the next time instant".

Lemma 4.1. *Let $Q_{11}, Q_{12}, Q_{21}, Q_{22}$ denote the blocks of $\Psi^{\mathrm{T}}P\Psi$ generated by partition conformal to the partition of $z = [x^{\mathrm{T}} \;\; f^{\mathrm{T}}]^{\mathrm{T}}$, let optimisation* (5) *be feasible at time k, and let the minimiser be denoted by f_k^*. Then the inequality below admits solutions $0 \leqslant \mu \leqslant 1$:*

$$a\mu^2 + b\mu + c \leqslant 0, \tag{9a}$$
$$a = f_k^{*\mathrm{T}} Q_{22} f_k^*, \tag{9b}$$
$$b = 2x_k^{\mathrm{T}} Q_{12} f_k^*, \tag{9c}$$
$$c = x_k^{\mathrm{T}} Q_{11} x_k - 1. \tag{9d}$$

Proof. For $z_k(\mu) = [x_k^{\mathrm{T}} \ \mu f_k^*]^{\mathrm{T}}$, (9a) is equivalent to $z_{k+1}^{\mathrm{T}} P z_{k+1} \leqslant 1$ for $z_{k+1} = \Psi z_k(\mu)$; this is true for $\mu = 1$: by assumption $z_k(\mu) \in S_z$ and S_z is invariant. The existence of solutions $0 \leqslant \mu \leqslant 1$ follows from the continuity of $z_k^{\mathrm{T}}(\mu) P z_k(\mu)$, $z_k^{\mathrm{T}}(\mu)\Psi^{\mathrm{T}} P \Psi z_k(\mu)$ in μ. □

Theorem 4.1. *Let μ_k^* be the smallest value of μ which satisfies constraint* (9) *and*

$$Hz_k(\mu) - h \leqslant 0. \tag{10}$$

Then under the assumptions of Lemma 4.1, *$f_k(\mu) = \mu f_k^*$ satisfies the condition*

$$f_k^{\mathrm{T}}(\mu_k^*) f_k(\mu_k^*) \leqslant f_k^* f_k^* \tag{11}$$

and $u_k = -Kx_k + Ef_k(\mu_k^*)$ ensures the feasibility of (5b) *at $k+1$.*

Proof. By the definition of f_k^* there exist $\mu \leqslant 1$ such that (10) holds true, and by Lemma 4.1 there exist $\mu \leqslant 1$ such that (9) holds true. Therefore, given the definition of μ_k^* we have $\mu_k^* \leqslant 1$ so that (11) holds true. Also by Lemma 4.1 we have that satisfaction of (9) ensures that $u_k = -Kx_k + Ef_k(\mu_k^*)$ results in $Tf_k(\mu_k^*) \in S_f(\Phi x_k + BEf_k(\mu_k^*))$ (with T as defined below (3)) which of course guarantees the feasibility of optimisation (5) at the next sampling instant. Thus it is possible to step outside $S_f(x_k)$, to get smaller values for the cost of (5) without losing the guarantee of feasibility at the next (and all subsequent) time instants. □

This suggests the extended Newton–Raphson MPC (ENRMPC) algorithm stated below.

Algorithm 4.1 (ENRMPC).

Off-line: *solve a SDP to maximise the volume of S_x subject to* (3) *and* (4).

On-line, *Step* 1: *Compute the negative root of $\phi(\lambda)$, and thus determine f_k^* from* (6).

On-line, *Step* 2: *Compute μ_k^*, the smallest value of μ which satisfies* (9–10)

On-line, *Step* 3: *Implement $u_k = -Kx_k + Ef_k(\mu_k^*)$; at the next instant go back to Step* 1.

Theorem 4.2. *Given that optimisation* (5) *is feasible at start time, the origin of the closed-loop system under ENRMPC is asymptotically stable and ENRMPC will revert to the LQ control law as soon as that becomes feasible* (*i.e. as soon as x_k enters P_o*).

Proof. Given feasibility at start time, Theorem 4.1 implies that feasibility is guaranteed at all future times. Furthermore by the invariance of S_z and condition (4), we have that $u_k = -Kx_k + Ef_k$ implies that Tf_k, though not the minimising solution of (5), does nonetheless satisfy constraints (1d). Use of the minimising solution instead of Tf_k reduces the cost further, since $f_{k+1}^{\mathrm{T}} f_{k+1} \leqslant f_k^{\mathrm{T}} TT^{\mathrm{T}} f_k = f_k^{\mathrm{T}} f_k - f_k^{\mathrm{T}} E^{\mathrm{T}} E f_k, k = 0, 1, \ldots$. It follows that $Ef_k \to 0$, and the law of (2) will tend to the LQ optimal of (1c). Clearly, $f_k = 0$ defines the optimal solution to (5) just as soon as LQ law is feasible w.r.t. (1d), and from then onwards all optimal f_k will be zero. □

The extra computation involved in Step 2 is trivial: the computation of μ_k^* can be performed explicitly. All that is involved is the solution of one quadratic and a set of linear equations in one variable. The low complexity is due to the simplicity of the scaling operation which allows the reduction of cost by stepping outside $S_f(x_k)$. This reduction allows for a significant improvement on NRMPC, so much so that in many instances ENRMPC performance is indistinguishable to the optimum achieved (at very much greater computational cost) through QPMPC. An analytical comparison of the two algorithms is difficult, hence below we use Monte Carlo simulations which show that for all but a small class of models ENRMPC performance matches very closely that of QPMPC.

5. Comparison of QPMPC with ENRMPC

5.1. A simple example

Consider the state-space model and constraints

$$A = \begin{bmatrix} 1 & 0.1 \\ 0 & 1 \end{bmatrix}, \tag{12a}$$

$$B = \begin{bmatrix} 0 \\ 0.0787 \end{bmatrix}, \tag{12b}$$

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}, \tag{12c}$$

$$-1 \leqslant u \leqslant 1 \tag{12d}$$

for which $K = [2.828 \ \ 2.826]$ is the LQ controller for $Q = C^{\mathrm{T}}C, \lambda = 0.1$. The maximal admissible set P_o is shown in Fig. 1 (solid line) together with the projection Σ_n of S_z onto x-space for $n = 0$ and $n = 9$ (dashed lines). As

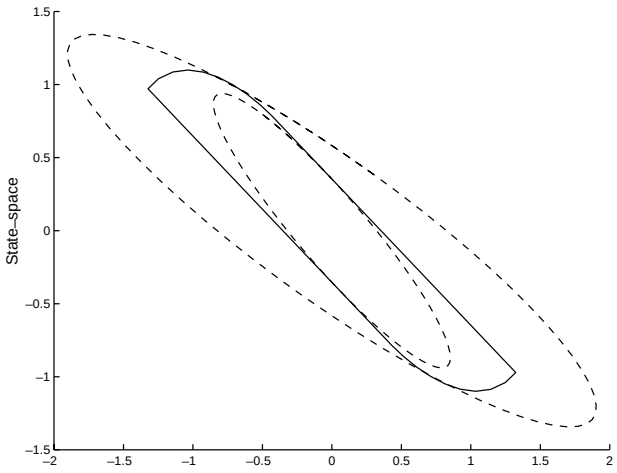

Fig. 1. Invariant sets.

Table 1
Closed-loop costs

	NRMPC	ENRMPC	QPMPC
$x_a = -[0.4\ \ 0.3]^{\mathrm{T}}$	3.5102	3.3498	3.3498
$x_b = [-1.4\ \ 0.5]^{\mathrm{T}}$	15.7432	14.8771	14.8771

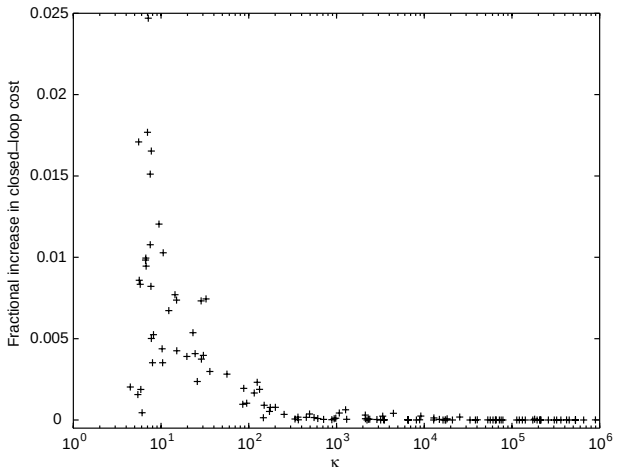

Fig. 2. Fractional increase in closed-loop cost vs observability matrix condition number κ for Group A models.

expected $\Sigma_0 \subset P_o$ but the extra 9 degrees of freedom results in Σ_9 which is significantly larger and contains P_o as a subset. The simulations were performed for two initial conditions and the resulting closed-loop costs are listed in Table 1. Clearly NRMPC did not do as well as QPMPC, but the ENRMPC and QPMPC results are indistinguishable.

5.2. Monte Carlo simulations

Two groups, A and B, were used comprising 2400 second and 2400 fifth order models, all subject to $|u| \leqslant 1$, and selected randomly with the constraint that the moduli of the eigenvalues of A should lie between 0.1 and 10. ENRMPC and QPMPC were applied with $n = 5$ (Group A) and $n = 10$ (Group B) and each model was released from randomly selected feasible initial conditions (for which $f_0 = 0$ was not feasible). For each run the infinite horizon closed-loop ENRMPC and QPMPC costs were computed and the relative difference is plotted against the condition number κ of the

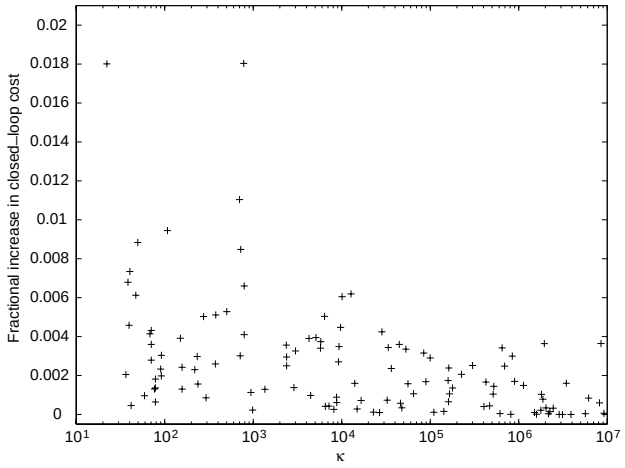

Fig. 3. Fractional increase in closed-loop cost vs observability matrix condition number κ for Group B models.

observability matrix for the pair $(\Psi,\ [K\ \ E])$ in Figs. 2 and 3 for Groups A and B. The frequency of the various values of κ for the two groups were also computed but for brevity are not shown here. From these plots it can be confirmed that the relative error never rose above 2% and indeed was less than about 1% for about 97% of the Group A models and about 96% of the Group B models. The improvement for ENRMPC over NRMPC follows the same trend but is an order of magnitude greater than that for QPMPC over ENRMPC and for both Groups A and B ranges from 0% (for large κ) to about 15% (for small κ)! Note that this range of values κ may appear excessive but is not when one considers the fact that the observability matrices involved are 7 and 15 dimensional. From the figures it is apparent that for larger κ ENRMPC matched closely the QPMPC cost. Fig. 4 provides the clue for this behaviour by demonstrating that there is a high correlation between κ and the condition number η of P. Large η correspond to the case of degenerate ellipsoids for which S_f gives a good approximation to the feasible set for QPMPC thereby reducing considerably the degree of sub-optimality in ENRMPC.

The results presented thus far are all based on SISO models. MIMO simulations (which for brevity are not plotted) presented a similar picture. For example, results based on 2000 randomly chosen models and initial conditions, representing systems with $n_x = 10$ states, $n_u = 4$ inputs and with

Table 2
CPU average times (using Matlab) per sampling instant

n_x	n_u	n	d	NRMPC	ENRMPC	QPMPC (warm started)
2	1	5	8	0.0023	0.0031	0.0385
5	1	10	16	0.0029	0.0044	0.048
5	2	10	17	0.0026	0.0032	0.0345
10	4	12	26	0.0029	0.0030	0.0402

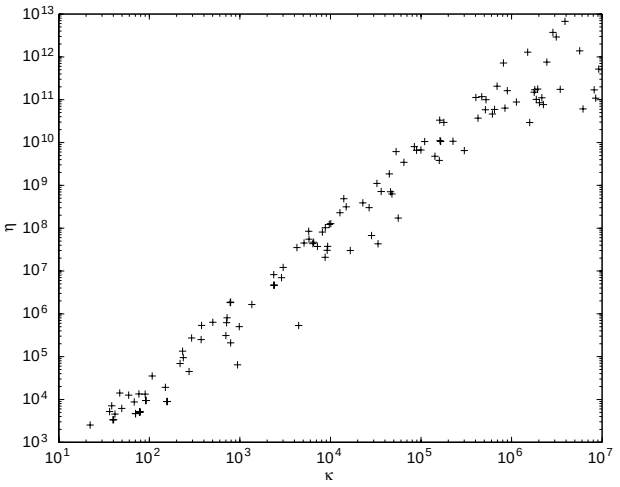

Fig. 4. Condition number η vs κ for Group B models.

$n = 12$, show that the relative error never rose above 2.8%, and for about 90% of the cases the error was $< 1.8\%$; once again, the figure of 90% was obtained from a plot of the frequency distribution of the condition number κ. A further 2000 simulations confirmed that for $n_x = 5, n_u = 2, n = 10$, the relative error never rose above 2.2%. In conclusion, therefore, the worse case relative error rises linearly with dimensionality $d = n_x + n_u + n$. The key point to note here is that this is achieved at a fraction of the QPMPC cost. In fact, this is true even for the case when QPMPC uses the previous optimal solution to warm-start the on-line optimisation as demonstrated by Table 2 which lists the computation times (in s) per sampling instant (averaged over the entire runtime and over 20 randomly selected initial conditions).

Thus even with warm starts, QPMPC requires more than 10 times the computational effort demanded by ENRMPC.

6. Conclusions

Ellipsoidal approximations to admissible sets provide a convenient means with which to reduce dramatically the computational burden of linear MPC. This dramatic reduction can be achieved through the use of a simple univariate optimisation (such as the Newton–Raphson iteration) which was shown to always converge. The results thus obtained are somewhat conservative on account of the use of the approximate ellipsoidal sets. This was overcome through ENRMPC, an extension which allows a search outside such ellipsoidal sets. The implementation of this extension requires a minimal amount of additional computation, yet, for the Monte Carlo simulations afforded an improvement of up to 15% over NRMPC. Indeed, for all but a small class of examples ENRMPC gave results which are within a fraction of a percent point of QPMPC. On the other hand, use of QP in fast sampling applications often will limit the number of degrees that can be introduced into the on-line optimisation, and in instances like this, ENRMPC will outperform QPMPC and will provide a very attractive alternative.

Acknowledgements

Financial support for this work was provided by the EPSRC.

References

Bemporad, A., Morari, M., Dua, V., & Pistikopoulos, E. N. (2000). *The explicit linear quadratic regulator for constrained systems.* Technical Report AUT99-16, ETH Zurich.

Boyd, S., El Ghaoui, L., Feron, E., & Balakrishnan, V. (1996). *Linear matrix inequalities, systems and control theory*. Philadelphia: SIAM.

Gilbert, C. E., & Tan, K. T. (1991). Linear systems with state and control constraints: the theory and practice of maximal admissible sets. *IEEE Transactions on Automatic Control*, *36*(9), 1008–1020.

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

Kouvaritakis, B., Rossiter, J. A., & Schuurmans, J. (2000). Efficient robust predictive control. *IEEE Transactions on Automatic Control*, *45*(8), 1545–1550.

Lee, J-W., Kwon, W. H., & Choi, J. (1998). On stability of constrained receding horizon control with finite terminal weighting matrix. *Automatica*, *34*(12), 1607–1612.

Rossiter, J. A., Kouvaritakis, B., & Cannon, M. (2000). *Triple Mode Control in MPC*. ACC00. Chicago, pp. 3753–3757.

Scokaert, P. O. M., & Rawlings, J. B. (1998). Constrained linear quadratic regulation. *IEEE Transactions on Automatic Control*, *43*(8), 1163–1169.

Seron, M., De Dona, J. A., Goodwin, G. C. (2000). *Global analytical model predictive control with input constraints*. CDC00, Sydney.

Wredenhagen, G.F., & Belanger, P. R. (1994). Piecewise-linear LQ control for systems with input constraints. *Automatica*, *30*(3), 403–416.