Explicit-Model Predictive Control: A simulation based scalability study

Arun Gupta*, Sharad Bhartiya**, P.S.V. Nataraj***

*ABB-GISL, Research, Bangalore, India 560048 (Tel: +91-9731029788; e-mail: a.gupta@abb.co.in).
**Department of Chemical Engineering, IIT Bombay, Mumbai, India (Tel: +91-22-25767225; e-mail:bhartiya@che.iitb.ac.in )
*** Systems and Control Engineering, IIT Bombay, Mumbai, India (Tel: +91-22-25767887;e-mail: nataraj@sc.iitb.ac.in)

Abstract: The main impediment in the applicability of explicit model predictive control (e-MPC) is that the number of regions in the parametric space increases dramatically with an increase in the dimension of the parameter and decision spaces as well as with an increase in the number of constraints. This work explores the scalability issues by simulation in which e-MPC is used for control of quadruple tank system. Explicit-MPC problem formulation from a standard MPC (for finite and infinite prediction horizon) using multi parametric quadratic programming (mp-QP) is presented. The results confirm that while the number of critical regions increases exponentially with the control horizon, the online calculations are computationally inexpensive.

1. INTRODUCTION

Since implementation of model predictive control (MPC) requires solution of an online optimization problem, its applications are typically limited to systems with sampling periods of the order of few seconds or minutes as commonly encountered in the process industry. In order to extend the range of MPC to applications with sampling periods in the milli/micro-second range, such as in mechatronics, it imperative to employ efficient optimization algorithms. For example, Wang and Boyd (2010) exploit the structure of the quadratic program (QP), which is solved using interior point methods to accomplish online optimization in the order of milliseconds for reasonably large sized problems. Interior point methods have also been used by other researchers (Rao et al., 1998). On the other hand, Ferreau et al. (2008) used customized active set methods and showed that these are at least an order of magnitude faster than conventional approaches for solution of QPs.

In this context, a radically different approach to solving the optimization problem called multiparametric (mp) programming (Gupta et al., 2011; Sjoqvold et al., 2006; Tondel et al., 2003a; Dua et al., 2002) has emerged as a promising tool which is particularly suited for online optimization. Multiparametric programming solves optimization problems by computing a parameter-dependent solution offline and subsequently integrating the pre-computed solution with parameters whose values become apparent online. In MPC, these parameters are past inputs, measurements and setpoint values. The offline solution consists of determining i) a set of critical regions that exhaustively partitions the parameter space, and ii) a set of functions corresponding to each critical region, whose solution yields the optimized decision variables (that is, the optimal input profile). The online implementation requires i) a search algorithm to determine the critical region, where the current values of parameters lie, and ii) a subsequent evaluation of the optimized decision variables using the corresponding function. For linear and quadratic programming problems that arise in linear MPC, these functions are linear and the optimized decision variables can be evaluated explicitly. Since the online component of the algorithm is typically simple, potential exists to implement linear MPC through low-cost hardware on systems with dynamic response time of the order of 10μs to 100ms (Pistikopoulos, 2009). Such an approach for solving the MPC optimization problem using multi-parametric programming has been referred to as explicit-MPC or e-MPC. However, a well-known impediment in successful implementation of e-MPC for medium to large-scale problems is that the number of critical regions becomes prohibitively large. Thus, a large online computational expense is incurred in searching for the applicable critical region thereby nullifying the benefits of computational efficiency. Various efforts have been presented in literature to overcome this so-called point location problem (Bayat et al., 2011; Monnigmann and Kastian, 2011), which have yielded online computational times of less than a microsecond. However, unlike the fast MPC methods using interior point and active set methods, which can potentially solve large practical problems, a criticism of e-MPC is that it can be used only for small-state and input dimension problems particularly due to the large offline computational load in addition to the online burden of the point location problem. Gupta et al., (2011) recently presented an active set mp-QP approach based on implicit enumeration of active sets and use of a pruning criterion, which they showed to scale well with the size of the problem. In this work, we benchmark the implicit enumeration mp-QP method for linear e-MPC using a simulation study in which the complexity of the mp-QP can be varied. The naive approach of sequential search among the polyhedral partition of the parametric space is used for the point location problem and no attempt is made to use state-of-art approaches in this regard.

This paper briefly reviews a novel algebraic approach to multi-parametric quadratic programming (mp-QP) that eliminates lacunae with previous geometric approaches (Gupta et al., 2011) and subsequently explores scalability issues in using e-MPC for a simulation case study consisting
of model predictive control of a benchmark quadruple tank system by varying the dimension of the decision and parametric spaces. Both, finite and infinite horizon e-MPC are used as a means of varying the complexity of the optimization problem. The paper is organized as follows: section 2 outlines the linear MPC problem for linear time invariant dynamical system; section 3 briefly reviews the methods to solve multiparametric programming problems; section 4 presents the e-MPC simulation case study on a quadruple tank system, and conclusions are presented in section 5.

2. LINEAR MODEL PREDICTIVE CONTROL

Model Predictive Control (MPC) is well established standard in the process industry, which considers operating constraints together with multivariable interactions. For an excellent review on the historical development of MPC, the reader is referred to Lee (2011). Here, we briefly review the MPC problem from a multi-parametric standpoint. In particular, we identify those parameters of the control algorithm that form the parameters of the multiparametric program.

Linear MPC uses a linear process model to predict the future process behaviour,

$$x_{k+1} = Ax_k + Bu_k$$

$$y_k = Cx_k$$

where, $x \in \mathbb{R}^n$ is the vector of states, $u \in \mathbb{R}^m$ is the vector of inputs and $y \in \mathbb{R}^l$ is the vector of outputs. The receding horizon MPC solves a constrained quadratic optimization problem to calculate the input vector over a finite control horizon $(N)$ by minimizing the deviation of the future process outputs from a desired reference trajectory over either a finite or an infinite prediction horizon $(p)$ subject to process and safety constraints. Thus, the MPC regulator problem is as follows,

$$\text{minimize } \sum_{j=0}^{p} (y_{k+j}^T Q y_{k+j} + u_{k+j}^T R u_{k+j} + \Delta u_{k+j}^T S \Delta u_{k+j})$$

such that,

$$y_{\text{min}} \leq y_{k+i} \leq y_{\text{max}}, \quad \forall \ i \in 1 ... p$$

$$u_{\text{min}} \leq u_{k+i} \leq u_{\text{max}}, \quad \forall \ i \in 0 ... N - 1$$

$$\Delta u_{\text{min}} \leq \Delta u_{k+i} \leq \Delta u_{\text{max}}, \quad \forall \ i \in 0 ... N - 1$$

and $u_{N-1} = u_N = \cdots = u_p$

where, $Q > 0, R \geq 0, S \geq 0$ are the weighing matrices of appropriate dimensions. In the above, we have re-parameterized the inputs by computing the input profile over only $N (\leq p)$ samples. This has been done to change the complexity of the resulting QP problem in the simulation case study. Other strategies of controlling the complexity of the QP have also been reported in literature (Cagnienard et al., 2004). Note that as $N, p \to \infty$, the resulting system exhibits closed-loop nominal stability (but not practically implementable). Various approaches exist to treat the infinite horizon feedback law by solving a finite horizon problem while retaining the stability property (Mayne et al., 2000; Chmielewski and Manousiouthakis, 1996). Note that predicted outputs $y_{k+j/k}$ use feedback information of the output at the current instant $k$. Muske and Rawlings (1993) provide various approaches for incorporating feedback information using models augmented by disturbance states or parameters. The MPC - optimization problem in Eq. (2) and (3) can be re-written in the standard multi-parametric form as follows,

$$\text{min}_{U=u_k \ldots u_{k+N-1}} U^T H U + 2U^T (G x_k - F u_{k-1})$$

$$\text{s.t. } AU \leq b + E x_k$$

where, $H, G, F, A, b$ and $E$ are the constant matrices derived from Eq. (2,3) (see for example, Seron et al, 2000). For extensions to servo problem and various approaches of incorporating feedback, the reader is referred to Muske and Rawlings (1993). In such a case, the set-points become a part of the parametric space. The next section presents the multi-parametric formulation of the regulatory problem.

3. EXPLICIT-MPC AS AMULTIPARAMETERIC QUADRATIC PROGRAM

Multiparametric programming provides an analytical solution of the constrained optimization problem as a function of the parameters associated with it. The MPC problem in Eq. (4,5) can be written as the following multiparametric quadratic program (mp-QP):

$$\min_{U} \frac{1}{2} U^T H U + c^T U + \theta^T PU$$

$$\text{s.t. } AU \leq b + E \theta$$

$$\theta \in \Theta \subset \mathbb{R}^q$$

The mp-QP in Eq. (6) is solved within the defined range of a closed polyhedral set $\Theta \subset \mathbb{R}^q$ which defines the bounds on the parameter set $\theta$. Comparing mp-QP problem in Eq. (6) with a standard MPC problem in Eq. (4 - 5):

$$c \equiv [0]; \theta \equiv \begin{bmatrix} x_k \\ [u_{k-1}] \end{bmatrix}; P \equiv \begin{bmatrix} G & F \end{bmatrix}$$

where, $U$ is the set of decision variables (stacked manipulated variables over the control horizon), $\theta$ is the set of parameter vector containing the current state of the system and previous control input. The term $\theta^T PU$ in the objective function can be eliminated by the following transformation $U = Z - H^{-1} P \theta$, (Dua et al., 2002) which converts the strictly convex mp-QP to the following standard form,

$$\min_{Z} \frac{1}{2} Z^T H Z + c^T Z$$

$$\text{s.t. } AZ \leq b + F \theta$$

$$\theta \in \mathcal{E} + \mathcal{A} H^{-1} P \theta$$

Equation (7) represents e-MPC problem with the matrices defined above.

The parameter dependent solution is based on the knowledge of optimal active constraints. Let $\mathbb{M}$ refer to the set of indices of all constraints of the QP,

$$\mathbb{M} \triangleq \{1, 2, \ldots, p\}$$


and the active set $\mathcal{A}(Z,\theta)$ denote the set of indices of the constraints that are active at $(Z,\theta)$,
\[ \mathcal{A}(Z,\theta) \triangleq \{ i \in \mathbb{M} | A_i Z - \delta_i - \tau_i \theta = 0 \} \quad (9) \]
where, $\mathcal{A}$ denotes the $i^{th}$ row of matrix $\mathcal{A}$. The maximum possible number of active sets which can be constructed from $\mathbb{M}$ is given by its power set,
\[ \mathcal{P}(\mathbb{M}) \triangleq \{ A_1 = \{ \}, A_2 = \{1\}, \ldots, A_{p+1} = \{1,2\}, \ldots, A_{2p} = \{1,2,\ldots,p\} \} \quad (10) \]
The set of indices which are not in the active set $\mathcal{A}$ are members of the inactive set $\mathcal{I}$,
\[ \mathcal{I}(u,\theta) = \mathbb{M} \setminus \mathcal{A}(Z,\theta) \quad (11) \]
In order to determine set of critical regions which exhaustively map the parameter space $\Theta$, an implicit enumeration technique is employed (Gupta et al., 2011), wherein all combinations of optimal active set are considered. Corresponding to each optimal active set, the Lagrangian multipliers are used to determine the sensitivity of the optimized decisions variables to change in parameters namely, the current state $x_t$ and past inputs $x_{k-1}$. This yields the explicit control law. This is stated in form of a theorem due to Fiacco (1983).

**Basic Sensitivity Theorem** (Fiacco, 1983): Let $\Theta'$ be a parameter vector in the convex mp-QP in Eq. (7) and $(Z'(\theta'), \lambda'(\theta'))$ be the corresponding Karush-Kuhn-Tucker (KKT) pair. Also assume that (i) strict complementary slackness (SCS) is satisfied, and (ii) the linear independence constraint qualification (LICQ) holds, then in the neighborhood of $\theta'$, there exists a unique, once continuously differentiable function $[Z(\theta'), \lambda(\theta')]$ satisfying the KKT conditions in Eq. (8) where $Z(\theta')$ is a unique isolated minimizer for the mp-QP in Eq. (2) such that
\[
\begin{bmatrix} Z(\theta') \\ \lambda(\theta') \end{bmatrix} = -M^{-1}N(\theta - \theta') + \begin{bmatrix} Z'(\theta') \\ \lambda'(\theta') \end{bmatrix} \quad (12)
\]

Thus, the parametric solution of the e-MPC in Eq. (7) can be obtained as a set of piecewise affine functions of parameters as shown in Eq. (12), where each function is valid in a closed polyhedron called as critical region (Dua et al., 2002; Tondel et al., 2003; Spjotvold et al., 2006; Gupta et al., 2011). Note that satisfying LICQ and SCS guarantees that the Jacobian $M$ is invertible for $H > 0$. The result of enumeration provides the optimal active sets $\mathcal{A}'$ and the corresponding optimal inactive sets from the solution of Karush-Kuhn-Tucker (KKT) conditions, which can be characterized as follows,
\[ \mathcal{A}_i Z(\theta) \leq \delta_i - \tau_i \theta, \ i \in \mathcal{I}'(\theta') \quad (13) \]
\[ \delta_i(\theta) \geq 0, \ i \in \mathcal{A}'(\theta') \quad (14) \]
where, $\theta'$ is fixed value of $\theta \in \Theta$ and $\delta_i$ represents the Lagrange multipliers that correspond to the $i^{th}$ constraint. These inequalities along with the original parameter bounds $\Theta$, after removal of redundant constraints, represent a polyhedron in the parameter space, termed as Critical Region (CR$_\mathcal{A}$) and corresponds to the active set $\mathcal{A}$,
\[ CR_\mathcal{A} \triangleq \Delta \{ \theta \in \Theta \leq \mathbb{R}^q : A_i Z(\theta) - b_i - F_i \theta \leq 0, \ \lambda(\theta) \geq 0 \} \quad (15) \]
where, $\Delta$ is a notional operator that eliminates redundant inequalities. Knowing the set of active constraints $\mathcal{A}$, the optimal solution of Eq. (7) can be represented as follows:
\[ Z_\mathcal{A} = \Omega_\mathcal{A} \theta + \omega_\mathcal{A} \quad (16) \]
where, $\Omega_\mathcal{A} \in \mathbb{R}^{m \times q}$ and $\omega_\mathcal{A} \in \mathbb{R}^m$ are constant matrices for each critical region. For detailed derivation of $\Omega_\mathcal{A}$ and $\omega_\mathcal{A}$ matrices readers are referred to Gupta et al., (2011).

**3.1 Robustness of mp-QP for e-MPC**

Two key issues arise due to degeneracy in mp-QP are (i) redundant partitioning of parameter space, and (ii) unexplored parameter regions in the parameter space. Redundant partitioning increases the size of look-up table thus increasing the online computational burden whereas unexplored regions can lead to abrupt termination of online algorithm. Thus, for e-MPC implementation it is vital that the parameter space be exclusively and exhaustively partitioned into critical regions, i.e. union of all critical regions constitutes the parameter space, as follows:
\[ \sum_i CR_{\mathcal{A}_i} = \sum_i \theta_{\mathcal{A}_i} = \theta \quad (17) \]
\[ \mathcal{A}_i \neq \mathcal{A}_j \ \forall \ i \neq j, \{\mathcal{A}_i, \mathcal{A}_j\} \in \mathcal{P} \quad (18) \]

All multi-parametric algorithms attempt to ensure exclusive and exhaustive partitioning of the parametric space and provide a unique solution in decision space corresponding to each critical region in the parameter space. Dua et al., (2002) first proposed the multiparametric quadratic algorithm to explicitly identify the solution of mp-QP as affine function of parameters. However, the algorithm introduces artificial cuts in the parameter space which result in redundant partitioning. Tondel et al., (2003) addressed the issue of redundant partitioning by assuming that a certain facet-to-facet property holds. However, Spjotvold et al. (2006) later showed that upon failure of this property, the algorithm may result in unexplored regions of the parameter space. Spjotvold et al., (2006) then provided a combination of the algorithms by Dua (2002) and Tondel (2003a), such that it exhaustively explores the parameter space and reduces the degree of redundant partitioning to a large extent. While all above algorithms appealed to geometric approaches for navigating in the parameter space, Gupta et al. (2011) proposed an algebraic approach based on implicit enumeration of all active sets, which guarantees that the parameter space is exhaustively explored with unique optimal affine functions in decision space even in the case of degeneracy. Like the other algorithms, the implicit enumeration approach also may provide overlapping partitions in presence of degeneracies. The e-MPC formulation presented in this work uses the algorithm by Gupta et al. (2011) to generate the critical regions and optimal solution functions forming an online look-up table. An active set enumeration approach has also
been mentioned in early mp-QP literature (Tondel et al., 2003a,b). An obvious pitfall of an enumeration based approach is that one may need to visit all active sets to completely partition the parameter space. However, in practice only a small number of active sets actually belong to polyhedral partition of the parametric space.

3.2 Scalability and complexity of e-MPC

The set of critical regions characterized by Eq. (15) and the optimal solution in Eq. (16) represent the analytical solution of Eq. (7) and can be thought as a look-up table for each critical region corresponding to the optimal active constraints. For a medium to large scale e-MPC problem, the size of look-up table can be prohibitively large. The number of critical regions (or size of look-up table) is proportional to the complexity of the optimization problem in Eq. (7), which can be judged by size of decision variables, parameters and number of constraints. The number of decision variable \( U \in \mathbb{R}^{m \times k} \) changes with the control horizon, \( N \) and remains unchanged with prediction horizon, \( p \) for both finite and infinite prediction horizon formulations. The size of parameter vector, \( \theta \) is independent of control and prediction horizon but changes if the weights and constraints on \( \Delta u \) are present in Eq. (2) and Eq. (3) respectively. Omitting the \( \Delta u \) term in MPC formulation will release the parameter \( u_{k-1} \in \mathbb{R}^m \) from the optimization problem thereby reducing the size of e-MPC problem. In this work, it is assumed that in the infinite horizon MPC, enforcing the constraints over the control horizon \( N \) ensures that the output constraints are satisfied over the infinite horizon. In case of finite horizon formulation however, the number of output constraints equals the prediction horizon, \( p \) as per Eq. (3). A simulation case study is presented next to explore the effect of change in control horizon, prediction horizon and number of parameters on the size of look-up table for finite and infinite horizon e-MPC problem.

4. SIMULATION CASE STUDY – FOUR TANK SYSTEM

Consider a quadruple tank system consisting of four interconnected tanks as shown in Fig. 1 (Johansson, 2000). The objective is to control the level in the two lower tanks with two pumps. The process inputs, \( v_1 \) and \( v_2 \), represents input voltages to the two pumps and the states \([h_1, h_2, h_3, h_4]\) representing the level of water in four tanks. The outputs \( y_1 = h_1 \) and \( y_2 = h_2 \) (water level in the two bottom tanks) are the control variables. At nominal operating condition with input voltages in both pumps being 3V, the corresponding values of level in tank 1 and tanks 2 are 12.4 and 12.7 cm respectively. To simulate level control, the first principles model (Johansson, 2000) is used as plant and the controller model shown in Eq. 19 is its linearized version obtained at nominal operation, with input voltages in both pumps being 3V and the corresponding values of level in Tank 1 and Tank 2 being 12.4 and 12.7 cm, respectively, where \( x_i = h_i - h_i^0 \) and \( u_i = v_i - v_i^0 \), deviations from the steady-state. The sampling time of 5 sec, with constraints on deviation variables for level \( x = \pm [5,5,1,1] \) cm, input voltage, \( u = \pm [1,1] \) V and \( \Delta u = \pm [0.1,0.1] \) V are selected. The constraints on \( x \) and \( u \) also define the parameter space to be explored.

![Fig. 1 Schematic diagram of the quadruple-tank process. (Johansson, 2000)](image)
Table 2, summarizes the effect of varying control horizon and effect of presence of constraints on rate of change of input on the computational cost of an infinite prediction horizon e-MPC with a finite control horizon (Muske and Rawlings, 1993) formulation. In addition to stabilizing properties, the infinite horizon formulation eliminates a user-specified parameter namely, the prediction horizon. Similar to the finite horizon case, an exponential increase in number of regions is noted with increase in control horizon. As expected, constraints on the input rate, also drastically increases the number of critical region.

It is interesting to note that the number of critical regions is typically fewer in infinite horizon than that of finite horizon formulation, exception being the last row of Table 2 when compared to last three rows of Table 1. Also the offline computational efforts for the infinite horizon formulation are lower than that for the finite horizon, even in case where infinite formulation generates more critical regions. The simulation case study thus suggests that a greater incentive exists for implementing the infinite horizon e-MPC than the finite horizon for comparable problem sizes. Additional work needs to be done to delineate reasons for the above suggestion.

Table 1: Finite horizon regulator with constraints on u, Δu and y

<table>
<thead>
<tr>
<th>Constraints on u, Δu and y</th>
<th>Control-Prediction horizon</th>
<th>No. of region</th>
<th>Offline CPU time (sec)</th>
<th>Average online CPU time (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-3</td>
<td>53</td>
<td>28.25</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>1-4</td>
<td>53</td>
<td>33.64</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>1-5</td>
<td>69</td>
<td>48.14</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>2-3</td>
<td>367</td>
<td>476.78</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>2-4</td>
<td>419</td>
<td>813.67</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>2-5</td>
<td>507</td>
<td>1413.8</td>
<td>0.391</td>
</tr>
<tr>
<td></td>
<td>3-3</td>
<td>1201</td>
<td>57612</td>
<td>0.391</td>
</tr>
<tr>
<td></td>
<td>3-4</td>
<td>1667</td>
<td>15890</td>
<td>0.391</td>
</tr>
<tr>
<td></td>
<td>3-5</td>
<td>2138</td>
<td>36349</td>
<td>0.396</td>
</tr>
</tbody>
</table>

Table 2: Infinite horizon servo with and without constraints on Δu

<table>
<thead>
<tr>
<th>Constraints on u, Δu and y</th>
<th>Control horizon</th>
<th>Number of region</th>
<th>Offline CPU time (sec)</th>
<th>Average online CPU time (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>25</td>
<td>15.1</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>313</td>
<td>338.3</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1978</td>
<td>5178</td>
<td>0.214</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Constraints on u, and y</th>
<th>Control horizon</th>
<th>Number of region</th>
<th>Offline CPU time (sec)</th>
<th>Average online CPU time (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>9</td>
<td>2.2</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>49</td>
<td>24.5</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>225</td>
<td>207</td>
<td>0.195</td>
</tr>
</tbody>
</table>

Fig. 2: Input/output control profile (a) output heights (solid: $h_1$, dashed: $h_2$), (b) input volts (solid: $v_1$, dashed: $v_2$).
5. CONCLUSIONS

The current work reports a simulation case study of e-MPC implementation. In particular, we studied the scalability of the algorithm with the prediction and control horizons. The results indicate that while the number of regions varies widely depending on the prediction and control horizons, the online burden was typically small. In case the online computational burden becomes significant, one may consider a sub-optimal approach such as use of shorter control horizons. Further, use of constraints on rate of inputs increases the problem size and may be dispensed with if possible. The computation times using a naive sequential search algorithms indicate that potential exists of reducing the sampling period further. This may aid in extending the domain on constrained optimal control through e-MPC implementation to applications in variable frequency drive control, position and velocity control in robotic systems, and crane control among others. The online computational time depends on the computational power and search algorithm used, than that of problem size. Advanced search algorithms or use of lower level language like C, can further reduce the online computational time. Furthermore, implementation on chip based hardware will reduce the online implementation time.

ACKNOWLEDGEMENTS

Authors thank Department of Science and Technology, India, for financial assistance under grant 09DST010.

REFERENCES


