A systematic selection of an alternative parameterisation for predictive control

Bilal Khan
Automatic Control and Systems Engineering
The University of Sheffield, UK
Email: b.khan@sheffield.ac.uk

John Anthony Rossiter
Automatic Control and Systems Engineering
The University of Sheffield, UK
Email: j.a.rossiter@sheffield.ac.uk

Abstract—Alternative parametrisations have been shown to improve the feasible region for a predictive control law when the number of degrees of freedom is limited. One question yet to be resolved is: which alternative parametrisation is best for a particular problem and what choice of parameter(s) within each parameterisation will lead to an improved feasible region with good performance? This paper tackles this question and demonstrates two systematic approaches to select the best alternative parameterisations. These approaches are based on multiojective optimisation and a pragmatic selection. Numerical examples demonstrate the efficacy of both methods.

Keywords: Alternative parameterisation, Feasibility, multiobjective optimisation.

I. INTRODUCTION

Model Based Predictive Control (MPC) or receding horizon control (RHC) or embedded optimisation or moving horizon or predictive control [2], [3], [1], is the general name for different computer control algorithms that use past information of the inputs and outputs and a mathematical model of the plant to optimise predicted future behaviour. MPC has been developed widely both in the process industry and control research community and has reached a high degree of maturity in its linear variant. Currently, MPC research focuses on stochastic and nonlinear scenarios, robustness and fast optimisation or related computational aspects.

All algorithms to some extent form a trade-off between feasibility, optimality and inexpensive optimisation, but may not have systematic tools for doing this trade-off. The success of earlier industrial heuristic MPC algorithms motivated the research community to develop several algorithms with improved performance and feasibility. There are several successful theoretical approaches but few of them are exploited commercially for real-time implementation. One important issue for real-time implementation is the ability to do a systematic trade-off between feasibility, performance and computational burden when choosing from currently available algorithms.

In recent years many authors have focussed on parametric solutions [5] or efficient implementations of online optimisers (e.g. for quadratic programming). However, this paper will follow a different route and consider the underlying structure of the MPC algorithm. Specifically, this paper proposes a systematic selection of an alternative parameterisation to enlarge the feasible region without too much detriment to performance (or optimality) and the computational burden.

In an MPC problem formulation, stability can be ensured using dual mode prediction by including a terminal constraint at the expense of the computational load [16]. In dual mode prediction, different formulations of the degrees of freedom (d.o.f.) or decision variables have been utilized using interpolation techniques [8] to enlarge feasible region without detriment to performance. There are different variants of interpolation, however, these methods are limited to small dimensional systems. Another approach in the literature is a concept of triple mode control [6], [7]; the triple mode strategy introduces an extra mode which may reduce the online computational burden with good performance and feasibility. However, in this strategy, the challenge is to find a suitable linear time varying (LTV) control law which enlarge feasible region with improved performance. One pragmatic solution to finding this law is tackled in [12] which used Laguerre parameterisations. However, this paper does not pursue Triple mode approaches as the offline complexity and decision making is substantially increased as compared to dual mode approaches.

More recently, alternative parameterisation based on Laguerre [17] and Kautz functions have been proposed to simplify the trade-off without increasing computational burden within dual mode prediction paradigm [9], [10], [11]. The main idea is to form the degrees of freedom in the predictions as a combination of either Laguerre or Kautz functions. These functions have proven to be a very effective for improving the volume of the feasible region with a limited number of d.o.f. These alternative parameterisations are generalised using orthonormal basis functions with Laguerre and Kautz functions as special cases [13].

The key question left to be resolved is, what is the best prediction dynamics to assume for predicted inputs, that is the d.o.f., to allow for a large feasible region without detriment to the closed loop performance and computational burden? In [15], different ways of parameterising the d.o.f. were investigated and a mechanism based on a Monte Carlo approach to define the best parameterisation using optimal sequences for numerous search directions was considered. However, a systematic choice for the underlying dynamics remains an open question. This paper focuses on systematic approaches to identify the best parameterisation dynamics using: (i) a multiobjective optimisation and (ii) a pragmatic choice.

This paper assumes that Laguerre, Kautz and Generalised
parameterisation are able to achieve large feasible regions while maintaining local optimality and a relatively low computational complexity [9], [10], [11], [12], [13], [14] and extends the earlier studies in [11], [13] in order to identify a systematic selection of the best parameterisation dynamics. Section II will give the necessary background about an optimal MPC, Laguerre MPC, generalised parameterisation for optimal MPC. Section III discusses two proposed schemes to identify the best parameterisation dynamics based on a multi-objective optimisation and a pragmatic approach. Numerical examples are presented in section IV and paper finishes with conclusions and future work in section V.

II. BACKGROUND

This section will summarise the background information related to nominal dual-mode MPC and the use of alternative parameterisations within MPC.

A. Problem formulation for MPC

Assume a discrete time linear time invariant (LTI) state space model of the form

\[ x_{k+1} = Ax_k + Bu_k \]  \hspace{1cm} (1)

where \( x_k \in \mathbb{R}^{n_x} \) and \( u_k \in \mathbb{R}^{n_u} \) which are the state vectors and the plant input respectively. Assume that the states and inputs at all time instants should fulfill the following constraints:

\[ u \leq u_k \leq \bar{u}; \quad \Delta u \leq \Delta u_k \leq \bar{\Delta} u; \quad x \leq x_k \leq \bar{x} \]  \hspace{1cm} (2)

B. Nominal MPC algorithm

In dual-mode MPC it is assumed that one has total freedom in the choice of the input signal \( u_k \) up to horizon \( n_c \) subject to constraints. Beyond horizon \( n_c \), a terminal control law with an asymptotic stabilising optimal feedback gain \( K \) is assumed. The ’predicted’ control law [3], [16] takes the form

\[ u_k = -K x_k + c_k, \quad k = 0, \ldots, n_c - 1, \]
\[ u_k = -K x_k, \quad k \geq n_c, \]  \hspace{1cm} (3)

where only first control move is ever implemented, \( c_k \) are degrees of freedom (d.o.f.) available for constrained handling. The terminal control law defines an invariant set \( \chi_0 \) for state vector \( x_k \) [8]. This invariant set is also known as maximum admissible set (MAS) which satisfies all polytopic constraints with recursive use of the terminal control law \( u_k = -K x_k \in \chi_0 \). The MAS is defined as:

\[ \chi_0 = \{ x_0 \in \mathbb{R}^{n_x} : x \leq x_k \leq \bar{x}, u \leq -K x_k \leq \bar{u}, x(k+1) = Ax(k) + Bu(k), \forall k \geq 0 \} \]  \hspace{1cm} (4)

In compact form defined as \( \chi_0 = \{ x_k : M x_k \leq b \} \) for suitable \( M \) and \( b \). Performance, either predicted or actual, will be assessed by the cost

\[ J = \sum_{k=0}^{\infty} x_k^T Q x_k + u_k^T R u_k \]  \hspace{1cm} (5)

Substituting the nominal model and predicted control values (3) into (5) and ignoring terms that do not depend upon the d.o.f., one finds from [3] that the optimisation problem in (5) can be reformulated as:

\[ \min_C J_c = C^T S C \text{ s.t. } M x_k + N C \leq b; \]  \hspace{1cm} (6)

where \( C = [c_k^T, \ldots, c_{k+n_c-1}^T] \). Details are in the literature [2], [3], [4]. The maximal control admissible set (MCAS) \( \chi_c \), the feasible set for optimal control problem in (6) that satisfies all polytopic constraints, is defined as:

\[ \chi_c = \{ x_k : \exists C, M x_k + N C \leq b \} \]  \hspace{1cm} (7)

The optimal MPC (OMPC) algorithm with guarantees of recursive feasibility and convergence, for the nominal case is given by solving QP optimisation (6) at every sampling instance then implementing the first component of \( C \), that is \( c_k \) in the control law of (3). The algorithm is formulated as:

Algorithm 2.1: OMPC [16], [3]

\[ e_k^* = \text{arg min}_{e_k} J_c \text{ s.t. } M x_k + N c_k \leq b; \]

Implement \( u_k = -K x_k + e_k^* \) where \( e_k^* = [I, 0, \ldots, 0] \). When the initial states \( x_0 \in \chi_0 \) then the optimising \( e_k^* \) is zero so the terminal control law \( u_k = -K x_k \) is implemented.

Remark 2.1: There is a well understood set of potentially conflicting objectives using OMPC e.g. between the desire for good performance and large feasible regions with the equally important desire to keep the number of d.o.f. small. In order for OMPC to obtain a large feasible region and good performance, a large number of d.o.f. or \( n_c \) is required.

C. Generalised parameterisation for Optimal MPC

Alternative parameterisation techniques for the d.o.f. in the future control values have been developed to improve the feasible region in nominal case. Laguerre, Kautz and generalised parameterisations have been proposed in [9], [11], [13] as an effective alternative to the standard basis set. This section will summarise Generalised optimal MPC.

1) Generalised functions: The generalised parameterisation [13] is defined using a higher order network such as:

\[ G_i(z) = G_{i-1}(z) \frac{(z^{-1} - a_1) \ldots (z^{-1} - a_n)}{(1 - a_1 z^{-1}) \ldots (1 - a_n z^{-1})}; \]  \hspace{1cm} (8)

\[ 0 \leq a_k < 1, \quad k = 1 \ldots n \]

With \( G_i(z) = \frac{\sqrt{(1-a_1^2) \ldots (1-a_n^2)}}{(1-a_1 z^{-1}) \ldots (1-a_n z^{-1})} \). The generalised function with \( a_k, \forall k = 1, \ldots, n \) gives [13]

\[ \text{Laguerre : } G_i = L_i, \text{ if } a_k = [a] \]
\[ \text{Kautz : } G_i = K_i, \text{ if } a_k = [a, b]. \]  \hspace{1cm} (9)

The Laguerre function is a special case of a generalised function and may be computed using the following state-space
dynamics model:
\[
\mathcal{G}(k+1) = \begin{pmatrix} a & 0 & 0 & 0 & \ldots \\ \beta & a & 0 & 0 & \ldots \\ -a\beta & \beta & a & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}_{A_G} \mathcal{G}(k);
\]
\[
\mathcal{G}(0) = \sqrt{1-a^2} [1, -a, a^2, \ldots]^T; \quad \beta = 1 - a^2
\]

2) GOMPC: generalised functions and MPC: The predictions using d.o.f. based on generalised functions [13] are:
\[
C = \begin{pmatrix} c_k \\ \vdots \\ c_{k+n_G-1} \end{pmatrix} = \begin{pmatrix} \mathcal{G}(0)^T \\ \mathcal{G}(1)^T \\ \vdots \end{pmatrix}_\rho = \mathcal{H}_G \rho
\]
where \(\rho\) is the \(n_G\) dimension decision variable when one uses the first \(n_G\) column of \(H_G\). The only difference between Laguerre MPC and Kautz or generalised MPC is that of \(H_G\) matrix. For further details readers are referred to [11] and [13].

**Algorithm 2.2: GOMPC**
\[
\rho^* = \text{arg} \min_{\rho} \rho^T \sum_{i=0}^{\infty} A_i^G \mathcal{G}(0) S \mathcal{G}(0)^T (A_i^G)^T \rho \\
\text{s.t.} \quad M x_k + NH \rho \leq b
\]
Define \(c_k^i = H_G \rho_k^i\) and implement \(u_k = -K x_k + c_k^T e_k^i\), where \(e_i^k\) is the \(i\)-th standard basis vector. 

**Remark 2.2:** It is straightforward to show, with conventional arguments, that all algorithms (i.e. LOMPC, KOMPC, GOMPC) using terminal constraints within MPC problem formulation provides recursive feasibility and Lyapunov stability.

### III. Best alternative parameterisation selection

In generalised parameterisation there are two clear choices within the future input predictions. First one can choose the order of the dynamics, that is the number of poles \(a_i\) in \(G_i(z)\), and second is the parameter selection(s), that is the actual values of \(a_i\). This section discusses the systematic selection of the parameterisation dynamics by asking what impact this choice has on feasibility and performance?

**A. The best choice for order of the prediction dynamics**

The prediction dynamics for the 3rd order prediction dynamics from (8) (which can easily be extended to \(n\)th order) can be defined as:
\[
\mathcal{G}(k+1) = \begin{pmatrix} b & 0 & 0 & 0 & \ldots \\ -ab & c & 0 & 0 & \ldots \\ ab^2 -b(1-ac) & a & \ldots & \vdots & \vdots \end{pmatrix}_{A_G} \mathcal{G}(k)
\]
\[
\mathcal{G}(0) = \gamma [1, 1, 1, -a, ab, -abc, a^2bc, \ldots]^T,
\]
\[
\gamma = \sqrt{(1-a^2)(1-b^2)(1-c^2)},
\]
The dynamic structure of the prediction in (13) is quite generic with distinct eigenvalues. To fulfill the algebraic relations in (11), \(A_G\) used in (13) must be a square matrix with a maximum number of dimensions the same as \(n_G\). Moreover a key observation from the prediction matrix structure is that \(\text{dim}(A_G) = n_G\) is an upper bound on dynamic dimensions. In fact, one could select \(\text{dim}(A_G) \leq n_G\).

**B. Multiobjective solution to select best parameterisation**

The main objective of this paper is to formulate a systematic method for handling the compromise between feasibility, performance and computational burden. This section shows how one can produce trade-off curves between feasible volume \(v\), performance loss \(\beta\) and computational burden (implicitly linked to \(n_c\)) of the parameterised MPC problem. Such trade-off curves allow us to formulate a multi-objective optimisation problem as a function of parameterisation parameters \(\alpha = [a_1, \ldots, a_m] \in (0, 1)^m\):
\[
J(n_c, n) := \min_{\alpha} \beta, \max_{\alpha} v \\
\text{s.t.} \quad M x + NH \rho \leq b, \\
\frac{\text{vol}(P_H)}{\text{vol}(P_{opt})}, \\
\beta = \frac{1}{n} \sum_j J_H(x) - J_{opt}(x)
\]
where \(P_{opt} := \{(x, c) | M x + N c \leq b\}\) represents the MCAS with \(n_c = 20\) (used as an approximate for the global maximum MCAS) for comparison, \(P_H = \{(x, c) | M x + N H \rho \leq b\}\) is the polytope sliced by the matrix \(H\), \(\text{vol}(\cdot)\) is the volume, \(J_{opt}(x) = \text{avg}. \{J_c(x, c)| (x, c) \in P_{opt}\}\) and \(J_H(x) = \text{avg}. \{J_c(x, H \rho)| (x, \rho) \in P_H\}\) for convex function \(J_c(x, c)\).

1) Volume approximation: Computing the volume of a high dimensional polytope is a complex task, and can, in the worst case be exponential in the size of the data \(M\) and \(N\) or \(NH\). Consequently, this paper approximates the volume. First select a large number of equi-spaced (by solid angle) or random directions in the state space i.e. \(x = (x_1, \ldots, x_n)\) and then, for each direction, the distance from the origin to the boundary of MCAS is determined by solving a linear programming (LP) and clearly the larger the distance, hereafter denoted as radius, the better the feasibility. Finally radii are normalised against the radii obtained for \(P_{opt}\). Although this might be considered somewhat arbitrary, it seems a pragmatic way of indicating relative volumes of different feasible regions compared to the best feasible shape with a large \(n_c\). One might argue that a precise volume measurement would, in some sense, be an equally arbitrary comparison.

2) Performance approximation: Similarly in case of calculating the performance loss, it is converted from integral to
sampled sum i.e.
\[
\beta := \frac{1}{n} \sum J_H(x) J_{opt}(x) - 1
\] (15)

That is, one computes the predicted performance for all the given state directions and compares each of these to the ‘global’ optimum. Equation (15) normalises these so that the best achievable performance would correspond to a \(\beta\) measure of unity.

By deploying explicit numeric measures of volume and performance, the multi-objective optimisation problem is able to generate trade-off curves between the d.o.f. \(n_c\), average radii and average performance loss of the parameterised problem with different parameterisation settings.

C. Pragmatic selection

In practice it is known from parameterisation insight that the optimal \(\rho_k\) is highly nonlinear in terms of its dependence upon the current state \(x_k\); more over the solution is to some extent unpredictable. The following gives a simple approach to identify a pragmatic selection of pole location(s) and order selection of parameterisation dynamics.

1) Order selection: The generalised functions with higher order orthonormal functions have more flexibility to improve feasibility while retaining good performance with a limited number of d.o.f [13]. So \(dim(A_G) = n_c\) is a sub-optimal choice to select order parameterisation dynamics.

2) Parameter selection: It is observed from simulation results that there is a relation between the closed loop poles and good locations of parameterisation dynamics. The pole locations of generalised functions can be selected to be equal to poles of the closed loop system using an optimal gain \(K\); this may be sub-optimal but is efficient.

D. Summary

In summary, there are two choices: (i) the order of parameterisation dynamics \(dim(A_G)\) and (ii) parameter value(s) or pole location(s) \([\alpha_1, \ldots, \alpha_n]\). One can use multi-objective optimisation to find \(dim(A_G)\) and \([\alpha_1, \ldots, \alpha_n]\). Another key observation is that the maximum \(dim(A_G) \leq n_c\) and closed loop eigenvalues or pole locations with an optimal gain \(K\) are a pragmatic choice to tune the parameterisation dynamics.

IV. Numerical Examples

In this section three numerical examples are presented to illustrate the efficacy of the proposed approach. The aim is to compare two aspects: (i) feasible regions; (ii) performance loss for the generalised parameterisation dynamics with different parameter(s) or pole location(s) selection by varying the d.o.f. or \(n_c\). The trade-off curves are shown between average radius and performance loss as a function of the different parameterisation parameter(s) and d.o.f. or \(n_c\). The parallel coordinates are also shown to represent an alternative variation of the parameter(s) and the effects on average performance loss and average feasibility gain. The multi-objective optimisation is done using the NSG-II Matlab toolbox.

A. Examples

Three numerical examples are simulated using weightings \(Q = C^T C\), \(R = 2\) and are simulated by varying \(n_c = [2, \ldots, 6]\) and with different parameter values. Trade-off curves between average radius and average performance loss are plotted by solving the multi-objective optimisation with varying parameter values. Similarly the effect on average performance loss and average radius are shown by varying \(n_c\). Conversely, Table 1 shows the average performance loss and average radius using a pragmatic approach to parameter selection.

1) Example 1:
\[
A = [1.2]; B = [1]; C = [1];
-0.2 \leq u_k \leq 0.2; -0.2 \leq \Delta u_k \leq 0.2; -5 \leq C x_k \leq 5;
\]

2) Example 2:
\[
A = [0.6, -0.4]; B = [0.2]; C = [1, -2.2];
-0.8 \leq u_k \leq 1.5; -0.4 \leq \Delta u_k \leq 0.4; -5 \leq C x_k \leq 5;
\]

3) Example 3:
\[
A = \begin{bmatrix} 1 & 0.1 & 0.3 \\ 0 & 0.1 & 0 \end{bmatrix}; B = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix};
C = [1, 0, 0]
-2 \leq u_k \leq 2; -1 \leq \Delta u_k \leq 1; -5 \leq C x_k \leq 5;
\]

B. Optimal selection based on multiobjective solution

The multi-objective procedure was run for different sample directions, which were chosen uniformly on the unit circle. The parameters \(n_c\) representing d.o.f. were chosen in the range \(\{2, \ldots, 6\}\). The results for Examples 1, 2, 3 are shown in Figures 1-2, 3-5 and 6-7 respectively. In all plots, colour circles represent the parameter values \(\alpha\).

Figures 1, 3 and 6 show the parallel coordinates and trade-off between average performance loss and average radius by varying \(n_c\) with different parameter values \(\alpha\). Parallel coordinates are used to plot all axes representing different objectives parallel to each other. Only the Laguerre function parameterisation simulation results are shown in the figures, but there are similar results for higher order parameterisations i.e. Kautz and 3rd order functions. These trade-off curves may be used as a selection criteria to choose best function parameterisation \(\alpha\).

Figures 2, 4 and 7 show the trade-off curves between average performance loss and average radius for \(n_c = 4\) with different parameterisation dynamics. It is observed that 3rd order function parameterisation improves both the feasible region without too much detriment to performance. An average performance loss and feasible region gain for example 2 is shown in figure 5 by varying \(n_c \in \{2, \ldots, 6\}\) for a Laguerre parameterisation. One can see that best parameter selection may improve the feasible region and performance with a limited number of d.o.f. \(n_c\).
C. Sub optimal selection based on closed loop eigenvalues

Table I, II and III show the pragmatic selection of parameterisation dynamics using closed loop eigenvalues with an optimal gain $K$. It is clear from these that pragmatic choices are sub optimal. The selection of parameterisation is based on closed loop eigenvalues so parameterisation order is selected based on example dimensions.

Laguerre parameterisation is the only choice for example 1. With a sub optimal choice of $n_c = 2$ achieves 70% of feasible set and 22% of performance loss. Similarly in example 2, there are two parameterisation choices Laguerre and Kautz functions. The Laguerre function achieved 90% of feasible set with 6% performance loss. On the otherhand the Kautz function achieved 95% of feasible set with 1% performance loss with $n_c = 2$. In example 3, there are three parameterisation choices i.e. Laguerre, Kautz and 3rd order functions. All achieved more than 90% global feasible set and maximum of 8% of performance loss.

Remark 4.1: It is clear from figures 1-4, 6 and 7 that the pragmatic choice has solution in vicinity of an optimal pareto solution.

Remark 4.2: The optimisation structure of an approximate global optimal control is different to the parameterised one, so in few cases performance may even improve.

V. CONCLUSION

The main contribution of this paper was to present a systematic approach to selecting the best alternative parameterisation for MPC. Two techniques were discussed based on a multi-objective optimisation and a pragmatic approach using closed loop poles. Examples demonstrate that in many cases nominal closed loop poles are a good sub optimal choice to tune the parameterisation dynamics. However an optimal selection may be made using a multi-objective optimisation technique. In future work there is a need to further investigate in parallel issues such as: will these choices lead to an efficient online optimisation QP structure and/or low complexity multi-parametric solution? Finally, it is an interesting future area to consider these choices within a robust scenario.

<table>
<thead>
<tr>
<th>Parameterisation</th>
<th>$n_c$</th>
<th>Avg. perf. loss</th>
<th>Avg. radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laguerre</td>
<td>2</td>
<td>22%</td>
<td>0.6904</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>22%</td>
<td>0.8011</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>22%</td>
<td>0.8720</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>22%</td>
<td>0.9237</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>22%</td>
<td>0.9378</td>
</tr>
</tbody>
</table>
REFERENCES


TABLE II
PRAGMATIC SELECTION FOR $\alpha$

<table>
<thead>
<tr>
<th>Example 2</th>
<th>Parameterisation</th>
<th>$n_c$</th>
<th>Avg. perf. loss</th>
<th>Avg. radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laguerre</td>
<td>$a=0.72$</td>
<td>2</td>
<td>6 %</td>
<td>0.9098</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>3 %</td>
<td>0.9548</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>2 %</td>
<td>0.9885</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>2 %</td>
<td>0.9799</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>2 %</td>
<td>0.9396</td>
</tr>
<tr>
<td>Kautz</td>
<td>$(a,b)=0.72\pm0.37$</td>
<td>2</td>
<td>1 %</td>
<td>0.9408</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>0 %</td>
<td>0.9584</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>0.5 %</td>
<td>0.9829</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>2 %</td>
<td>0.9885</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>2 %</td>
<td>0.9907</td>
</tr>
</tbody>
</table>