Explicit Model Predictive Control via Nonlinear Piecewise Approximations ⋆

Van-Vuong Trinh [*,**] **Mazen Alamir** [*] **Patrick Bonnay** [**] **François Bonne** [**]

[*] *GIPSA-lab, CNRS, University of Grenoble Alpes, 11 rue des Mathématiques, 38402 Saint Martin d'Hères, France (e-mail: {van-vuong.trinh,mazen.alamir}@ gipsa-lab.fr)*
[**] *CEA-INAC-SBT, 38054 Grenoble, France (e-mail: {patrick.bonnay,francois.bonne}@cea.fr)*

Abstract: In this paper, a novel identification methodology is proposed to capture general multivariate nonlinear relationships, with focus on the bounded-error approximation of model predictive control for constrained (non)linear systems. The formulation of the identification problem takes, at each iteration, the form of a constrained linear (or quadratic) optimization problem that is mathematically feasible as well as numerically tractable. The efficiency of the proposed method for the derivation of low-complexity explicit model predictive controllers is demonstrated via the constrained control of a thermodynamic power plant.

Keywords: Explicit Model Predictive Control, Nonlinear Approximation.

1. INTRODUCTION

Model Predictive Control (MPC, see [Mayne et al. (2000)]) is a widely used control method which involves the solution at each sampling instant of a finite horizon optimal control problem subject to the system dynamics and input and state constraints. Nevertheless, the on-line solution of an optimization problem is often time consuming and requires expensive computation. This makes the real-time MPC implementation usually limited to processes where the sampling time and hardware capability are sufficient to support the computational needs.

To overcome the aforementioned implementation issues, several approaches to explicit MPC (EMPC) have been recently pursued, where the on-line computation would simply be point location search and function evaluation. An interesting survey on this topic can be found in [Alessio and Bemporad (2009)]. In the context of linear systems, the classical problem of obtaining EMPC controllers with quadratic cost and linear constraints can be solved by using parametric quadratic programming techniques (see, for instance, [Bemporad et al. (2002)]). As for nonlinear systems, deriving the true optimal nonlinear MPC (NMPC) control law is generally not possible and hence approximate approaches have to be followed. Following this direction, several techniques have been devised such as artificial neural networks [Pin et al. (2013)], set memberships identification [Canale et al. (2009), Fagiano et al. (2012)] and piecewise affine (or linear) approximators [Johansen (2004), Grancharova and Johansen (2012), Grancharova and Olaru (2014)]. Nevertheless, while piecewise affine approaches normally suffer from the "curse of dimensionality", standard nonlinear approximators offer universal capabilities at the price of non-convex optimization schemes. Hence, the investigation of EMPCs with nonlinear piecewise representations is of interests.

In this paper, we introduces an extension of the multi-input single-output identification method recently proposed in [Alamir (2013)] which has been used in several applications [Alamir et al. (2014a), Alamir et al. (2014b), Rahmani et al. (2015)]. This method is suitable for the component-wise approximation of MPC control laws, i.e. each control input will be identified independently. The obtained approximate EMPC is represented as piecewise nonlinear approximators, allowing to reduce the number of regions with respect to the piecewise affine approaches. Compared to the neural networks approach, the methodology takes advantages of efficient computation of constrained linear (or quadratic) programming problems. Bounds on admissible errors are also given in order to maintain the closed-loop performance as well as to trade off complexity and approximation error.

This paper is organized as follows. In Section 2, the design of a nonlinear piecewise approximator is presented. Section 3 describes a heuristic procedure for approximating general nonlinear functions as well as MPC control laws based on the designed approximator. Simulation results are reported in Section 4 before drawing the conclusions in the final section.

Notation

- $\mathbb{I}_{a:b} := \{a, a+1, \ldots, b-1, b\}$, $a, b \in \mathbb{N}$, $a < b$;
- p_i, $i \in \mathbb{I}_{1:n}$: components of a vector $p \in \mathbb{R}^n$;
- $\text{card}(\mathcal{S})$: the cardinality of the set $\mathcal{S}$, i.e. the number of elements of the set;
- $N(P)$: the neighborhood of a point, a set or a topological space P within an appropriate tolerance;
- ∂P: the boundary of a topological space P.

⋆ This work was sponsored within the project CRYOGREEN, financed by the French National Research Agency (ANR).

2. A NONLINEAR PIECEWISE APPROXIMATOR

2.1 Problem statement and mathematical formulation

Consider the following nonlinear identification problem:
Problem 1. *Given the data* $\mathcal{D} = \{(q(k), Z(k)\}_{k\in\mathbb{I}_{1:N}}$ *where* $Z \in \mathcal{H} \subseteq \mathbb{R}^{n_z}$ *is the regressor and* $q \in [\underline{q}, \overline{q}] \subset \mathbb{R}$ *is the output. Find a map* $F : \mathbb{R}^{n_z} \to \mathbb{R}$ *of the form:*

$$F(Z) := \Gamma^{-1}(L^T Z) \text{ where } \Gamma(\cdot) \text{ is strictly increasing} \quad (1)$$

such that the following approximation holds:

$$q \approx F(Z) \quad (2)$$

The structure (1) is a Wiener-like model with nonlinear piecewise strictly monotonic static output mapping which can be parametrized using a finite function basis:

$$\Gamma(q) = \sum_{j=1}^{n_b} [B^{(j)}(\xi(q))]\mu_j = B(\xi(q))\mu;\ \xi(q) = \frac{q-\underline{q}}{\overline{q}-\underline{q}} \quad (3)$$

where the basis functions are given by:

$$\{B^{(j)}\}_{j\in\mathbb{I}_{1:n_b}} := \{1\} \cup \{B_1^{(i)}\}_{i\in\mathbb{I}_{1:n_m-1}} \cup \{B_2^{(i)}\}_{i\in\mathbb{I}_{1:n_m}} \quad (4)$$

The number of functions is $n_b = 2n_m$ while $B_1^{(i)}$ and $B_2^{(i)}$ are defined as:

$$B_1^{(i)}(\eta) := (1+\alpha_i)\frac{\eta}{1+\alpha_i\eta};\ B_2^{(i)}(\eta) := \frac{\eta}{1+\alpha_i(1-\eta)} \quad (5)$$

The coefficients α_i are given by $\alpha_i := e^{\beta(i-1)} - 1$ for some constant $\beta > 0$ (see Alamir (2013)).

Denote $\mu \in \mathbb{R}^{n_b}$ and $L \in \mathbb{R}^{n_z}$ as the parameters of $F(Z)$. The total number of parameters is $n_p = n_b + n_z$. The basic idea of the above formulation is to solve Problem 1 by finding $\mu \in \mathbb{R}^{n_b}$ and $L \in \mathbb{R}^{n_z}$ such that the approximation $B(\xi(q))\mu \approx L^T Z$ holds. Hence, consider the following linear program (LP) with a positive constant ϵ:

$$\begin{aligned} \min_{\mu,L} \quad & \max_{(q,Z)\in\mathcal{D}} |w(q,Z)\cdot(B(q)\mu - Z^T L)| \\ \text{s.t.} \quad & [\frac{dB}{d\xi}(\xi)]\mu \geq \epsilon,\ \forall \xi \in [0,1] \end{aligned} \quad (6)$$

where the indicator $w(q,Z)$: $\mathbb{R}\times\mathbb{R}^{n_z} \to \mathbb{R}_{>0}$ is introduced in (6) to enforce specific precision depending on the problem. We define this weight as a function of the learning data:

$$w(q,Z) = \begin{cases} \rho_i & \text{if } (q,Z) \in \mathcal{W}^{(i)} \\ 1 & \text{otherwise} \end{cases} \quad (7)$$

where ρ_i are some positive constants and $\mathcal{W}^{(i)}$ are disjoint subspaces. This formulation obviously recalls the known *weighted norm approximation* where ∞-norm is employed. Alternative formulation, based on the L_2-norm, can also be adopted leading to a quadratic programming (QP) problem.

The constraint expresses the fact that Γ has to be strictly monotonic in order to guarantee the existence of the inverse map $\Gamma^{-1}(\cdot)$. Note that by defining a sufficiently dense grid of ξ over the interval $[0,1]$, namely $\{\xi_i\}_{i\in\mathbb{I}_{1:n_\xi}}$, this constraint can be transformed into a finite number of linear inequalities.

Remark 1. The normalization constraint in the formulation of [Alamir (2013)] has been removed as the imposed constraint guarantees that $\mu \neq 0$, i.e. the trivial solution ($\mu = 0$, $L = 0$) is never admissible.

2.2 Preliminary analysis

For convenience, we denote $p_{id} = (n_m, \beta, \epsilon, \{\xi_i\}_{i\in\mathbb{I}_{1:n_\xi}}, w(\cdot))$ as *identification parameters* and *model parameters* as $p = \begin{bmatrix}\mu \\ L\end{bmatrix} \in \mathbb{R}^{n_p}$. Let the slack variable ζ be the minimizing cost, the LP (6) is rewritten as follows

$$\begin{aligned} \mathcal{L}_{(p_{id},\mathcal{D})}(p,\zeta): \quad & \min_{p,\zeta} \quad \zeta \\ & \text{s.t.} \quad A(p_{id},\mathcal{D})\cdot\begin{bmatrix}p \\ \zeta\end{bmatrix} \leq b(p_{id},\mathcal{D}) \end{aligned} \quad (8)$$

where $A(p_{id},\mathcal{D})$ and $b(p_{id},\mathcal{D})$ are easily derived.

The feasibility of $\mathcal{L}_{(p_{id},\mathcal{D})}(p,\zeta)$ is as follows:
Proposition 1. *The LP* $\mathcal{L}_{(p_{id},\mathcal{D})}(\cdot)$ *is feasible.*

Proof. According to the definition of the functional basis, $[\frac{dB}{d\xi}(\xi)] = \left[\frac{dB^{(j)}}{d\xi}(\xi_i)\right]_{i\in\mathbb{I}_{1:n_\xi}, j\in\mathbb{I}_{1:n_b}}$ is an $n_\xi \times n_b$ matrix with the elements of the first column being zeros while the remaining being positive. Hence, it is obvious that there exists μ such that $[\frac{dB}{d\xi}(\xi)]\mu > 0$. Thus, the feasibility of $\mathcal{L}_{(p_{id},\mathcal{D})}(\cdot)$ is guaranteed with any $\epsilon > 0$. □

The identification residual can be characterized as follows:
Proposition 2. *If* (μ, L) *is a feasible solution of the LP* $\mathcal{L}_{(p_{id},\mathcal{D})}(\cdot)$ *with corresponding cost* J*, the identification residual of any learning data point* (q, Z) *is such that:*

$$|q - F(Z)| \leq \frac{1}{w(q,Z)}\cdot\frac{1}{\epsilon}\cdot(\overline{q}-\underline{q})\cdot J \quad (9)$$

Proof. For any $(q,Z) \in \mathcal{D}$, the continuity of $\Gamma(\cdot)$ implies the existence of $\hat{q} = F(Z)$ such that $B(\xi(\hat{q}))\mu = L^T Z$. One clearly has

$$\begin{aligned} J &\geq w(q,Z)|B(\xi(q))\mu - L^T Z| \\ &= w(q,Z)|B(\xi(q))\mu - B(\xi(\hat{q}))\mu| \\ &\geq \min_{\xi\in[0,1]}\left[(\frac{dB}{d\xi}(\xi)\mu)\right] w(q,Z)|\xi(\hat{q}) - \xi(q)| \\ &\geq \epsilon w(q,Z)|\xi(\hat{q}) - \xi(q)| = \epsilon w(q,Z)\frac{|\hat{q}-q|}{\overline{q}-\underline{q}} \end{aligned} \quad (10)$$

which is equivalent to (9). □

Remark 2. Proposition 2 implies that the desired fit can be obtained with a sufficiently small cost and appropriate identification parameters. It is also possible to adjust the relative fit between subsets of data through $w(q,Z)$. Nevertheless, the upper bound (9) on the identification residual can be very conservative in some cases.

3. NONLINEAR PIECEWISE APPROXIMATIONS AND MODEL PREDICTIVE CONTROL

3.1 Problem statement and proposed methodology

The lack of universal property of the proposed nonlinear approximator (1) has been underlined in [Alamir (2013)]. In order to overcome this structural limitation, we consider the following identification problem:

Problem 2. *Given the data* $\mathcal{D} = \{(q(k), Z(k)\}_{k\in\mathbb{I}_{1:N}}$ *where* $Z \in \mathcal{H} \subseteq \mathbb{R}^{n_z}$ *is the regressor and* $q \in [\underline{q}, \overline{q}] \subset \mathbb{R}$ *is the output. Find* s *maps* $F^{(i)} : \mathbb{R}^{n_z} \to \mathbb{R}$ *of the form:*

$$F^{(i)}(Z) := {\Gamma^{(i)}}^{-1}({L^{(i)}}^T Z);\ \Gamma^{(i)}(\cdot) \text{ is strictly increasing} \tag{11}$$

and corresponding regions $\{\mathcal{R}^{(i)}\}_{i\in\mathbb{I}_{1:s}}$ *forming a partition of* $\mathcal{H}$ *such that the following approximation holds:*

$$q \approx \begin{cases} F^{(1)}(Z) & \text{if } Z \in \mathcal{R}^{(1)} \\ \vdots & \\ F^{(s)}(Z) & \text{if } Z \in \mathcal{R}^{(s)} \end{cases} \tag{12}$$

The structure (12) will be referred to as *nonlinear piecewise monotonic approximations* while s is called the *model complexity*. This structure is obviously a generalization of piecewise affine identification frameworks.

Each map $\Gamma^{(i)}$ is parametrized as in (3) over $[0, 1]$, using the same normalization $\xi(q) = \dfrac{q - \underline{q}}{\overline{q} - \underline{q}}$. For convenience, define model parameters as

$$p^{(i)} = \begin{bmatrix} p_1^{(i)} \\ \vdots \\ p_{n_p}^{(i)} \end{bmatrix} = \begin{bmatrix} \mu^{(i)} \\ L^{(i)} \end{bmatrix} \in \mathbb{R}^{n_p} \tag{13}$$

In order to limit the number of submodels and obtain sufficient fitting, a methodology is proposed to solve Problem 2, consisting of two stages (see Fig. 1):

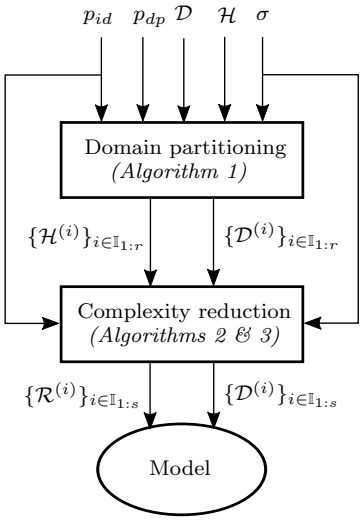

Fig. 1. Flow diagram of identification methodology

(1) *Domain partitioning.* This stage seeks partitions $\{\mathcal{H}^{(i)}\}_{i\in\mathbb{I}_{1:r}}$ of $\mathcal{H}$ and corresponding interior data clusters $\{\mathcal{D}^{(i)}\}_{i\in\mathbb{I}_{1:r}}$ where r is sufficient for data fit.
(2) *Complexity reduction.* This stage aims to obtain $s \leq r$ regions by merging the partitions into $\{\mathcal{R}^{(i)}\}_{i\in\mathbb{I}_{1:s}}$ and reassignment of data into clusters $\{\mathcal{D}^{(i)}\}_{i\in\mathbb{I}_{1:s}}$ if that does not cause a major increase in the minimal cost.

In the identification procedure, a bound σ on the identification residual is imposed and is used to trade off fit and complexity. This method can be considered as a bounded-error approach where the construction of regions, the reassignment of data clusters and model approximation are carried out simultaneously. Its two stages are detailed in the next sub-sections.

3.2 Domain partitioning

We restrict our attention to the hyper-rectangular regression domain $\mathcal{H} \subset \mathbb{R}^{n_z}$ defined by

$$\mathcal{H} = \{Z \in \mathbb{R}^{n_z} : \underline{h}_i \leq Z_i \leq \overline{h}_i, i \in \mathbb{I}_{1:n_z}\} \tag{14}$$

The non uniform hyper-rectangular partitions of $\mathcal{H}$, namely $\mathcal{P} = \{\mathcal{H}^{(j)}\}_{j\in\mathbb{I}_{1:r}}$ are defined as

$$\mathcal{H}^{(j)} = \{Z \in \mathbb{R}^{n_z} : \underline{h}_i^{(j)} \leq Z_i \leq \overline{h}_i^{(j)}, i \in \mathbb{I}_{1:n_z}\} \tag{15}$$

Let $\mathcal{D}^{(j)}$ be the data cluster belonging to $\mathcal{H}^{(j)}$ and $\sigma^{(j)}$ be the residual error of the corresponding submodel obtained by solving of the LP $\mathcal{L}_{(p_{id},\mathcal{D}^{(j)})}(\cdot)$.

Given $\sigma > 0$ and the *domain partitioning parameter* $p_{dp} = (\{\underline{\epsilon}_i\}_{i\in\mathbb{I}_{1:n_z}}, \underline{N})$ where $\underline{\epsilon}_i$ are the minimum allowed size of each hypercube along each dimension while $\underline{N}$ is the minimum allowed number of data in each partition, our objective is to find the smallest number r such that the following conditions hold for all $\mathcal{H}^{(j)}$, $\mathcal{D}^{(j)}$ and $\sigma^{(j)}$:

(1) Size condition
$$\overline{h}_i^{(j)} - \underline{h}_i^{(j)} \geq \underline{\epsilon}_i, \text{ for all } i \in \mathbb{I}_{1:n_z} \tag{16}$$
These lower bounds are imposed by the machine precision.
(2) Cardinality condition
$$\text{card}(\mathcal{D}^{(j)}) \geq \underline{N} \tag{17}$$
This condition avoids over-fitting phenomenon.
(3) Fitting condition
$$\sigma^{(j)} \leq \sigma \tag{18}$$
Generally, the lower σ, the finer resulting partitions.

In general, the fulfilment of conditions (16), (17) and (18) is nontrivial, especially with small bounding error, low-dimensional spaces or noisy data. Hence, their relaxation may be needed. We propose a data-based splitting rule and a domain partitioning algorithm which are motivated from [Johansen (2004)] where the hypercubes is partitioned into smaller hypercubes in order to achieve the desired fit.

The data-based splitting rule: Split a hypercube by a hyperplane through the Euclidean center of its interior data cluster and is orthogonal to one regressor-axis such that the reduction of identification error is maximal while both size condition and cardinality condition hold. Nevertheless, alternative splitting rules can also be considered with any splitting hyperplanes that through the interior data cluster.

Algorithm 1 Domain partitioning.

Parameters: p_{dp}, p_{id}, σ
Input: $\mathcal{D}$, $\mathcal{H}$
Step 1: Initialize the partition to a rough initial partition, i.e. $\mathcal{P} = \mathcal{P}^{(0)}$ such that $\mathcal{P}^{(0)}$ admits size condition and cardinality condition. It can be the whole regression domain, i.e. $\mathcal{P}^{(0)} = \{\mathcal{H}\}$. Mark all hypercubes of $\mathcal{P}$ as unexplored and go to step 2.
Step 2: Select any unexplored hypercube $\mathcal{H}_0 \in \mathcal{P}$. If no such hypercube exists, the algorithm terminates with the partition $\mathcal{P}$. Otherwise, go to step 3.

Step 3: Based on the data cluster $\mathcal{D}_0$ belonging to the hypercube $\mathcal{H}_0$, compute and validate the corresponding submodel to obtain its residual error σ_0. If σ_0 admits fitting condition, mark $\mathcal{H}_0$ as explored and go to step 2. Otherwise, go to step 4.
Step 4: Split the hypercube $\mathcal{H}_0$ into two hypercubes $\mathcal{H}_1$ and $\mathcal{H}_2$ based on the data-based splitting rule, and go to step 5. In the case of splitting failure, the algorithm unsuccessfully terminates.
Step 5: Remove $\mathcal{H}_0$ from $\mathcal{P}$, mark $\mathcal{H}_1$ and $\mathcal{H}_2$ as unexplored, and add them to $\mathcal{P}$. Go to step 2.
Output: $\mathcal{P} = \{\mathcal{H}^{(i)}\}_{i\in\mathbb{I}_{1:r}}$

3.3 Complexity reduction

If the domain partitioning stage leads to an overestimation of the model complexity needed for data fit, this number can be reduced by forcing the parameters of some partitions to be identical. The *merged submodel* is characterized by a nonempty index set $\mathcal{I} \subseteq \mathbb{I}_{1:r}$. Its region is $\bigcup_{i\in\mathcal{I}} \mathcal{H}^{(i)}$ while its parameter is solution of $\mathcal{L}_{(p_{id},\bigcup_{i\in\mathcal{I}} \mathcal{D}^{(i)})}(\cdot)$. Let $\sigma^{(\mathcal{I})}$ be its residual error over reassigned cluster $\bigcup_{i\in\mathcal{I}} \mathcal{D}^{(i)}$.

Given $\sigma > 0$, our objective is to find the smallest number s such that there exists a collection of s disjoint index sets $\{\mathcal{I}_1, \ldots, \mathcal{I}_s\}$ which covers all elements of $\mathbb{I}_{1:r}$, each $\mathcal{I}_i \subseteq \mathbb{I}_{1:r}$ and such that $\sigma^{(\mathcal{I}_i)} \leq \sigma$. This is a combinatorial problem with the trivial solution $\{\mathcal{I}_1 = \{1\}, \ldots, \mathcal{I}_{s=r} = \{r\}\}$. Thus, heuristic approaches can be suitable to find an approximate solution within a reasonable time.

The basic idea of our complexity reduction procedure is that given a nonempty index set $\mathcal{K}$ with $\sigma^{(\mathcal{K})} > \sigma$, we try to find the subset $\mathcal{I} \subseteq \mathcal{K}$ such that card($\mathcal{I}$) is as large as possible and $\sigma^{(\mathcal{I})} \leq \sigma$. The set $\mathcal{I}$ is called as *maximum mergeable subset* of $\mathcal{K}$. This can be done by considering the following LP, namely $\mathcal{L}^{(\mathcal{K},\mathcal{I})}_{(p_{id},\mathcal{D})}(\cdot)$:

$$\begin{aligned} \min_{p^{(i)},\zeta^{(i)},\tilde{p}} \quad & \sum_{i\in\mathcal{K}} \zeta^{(i)} \\ \text{s.t.} \quad & A(p_{id},\mathcal{D}^{(i)}) \cdot \begin{bmatrix} p^{(i)} \\ \zeta^{(i)} \end{bmatrix} \leq b(p_{id},\mathcal{D}^{(i)}) \quad \forall i \in \mathcal{K} \\ \text{and} \quad & p^{(i)} = \tilde{p} \quad \forall i \in \mathcal{I} \end{aligned} \tag{19}$$

The slack variable $\zeta^{(i)}$ is the minimizing cost of the LP $\mathcal{L}_{(p_{id},\mathcal{D}^{(i)})}(\cdot)$ while the constraint matrices $A(\cdot)$, $b(\cdot)$ are defined as in (8). The additional equality constraints are called the *merging constraint* while $\tilde{p}$ is parameter of merged submodel. Note that it can be transformed into:

$$||p^{(i)} - \tilde{p}|| = 0 \tag{20}$$

where $||\cdot||$ is ∞-norm.

The feasibility of (19) results from Proposition 1:

Proposition 3. *The LP $\mathcal{L}^{(\mathcal{K},\mathcal{I})}_{(p_{id},\mathcal{D})}(\cdot)$ is feasible.*

Let $\lambda^{(i)}_{(\mathcal{K},\mathcal{I})}$ be the Lagrangian multiplier of $\mathcal{L}^{(\mathcal{K},\mathcal{I})}_{(p_{id},\mathcal{D})}(\cdot)$ corresponding to the equality constraint (20). Thanks to the availability of these local sensitivities, the following greedy-like procedure is proposed to search for the maximum mergeable subset.

Algorithm 2 Search of maximum mergeable subset.

Parameters: p_{id}, σ
Input: $\mathcal{K}$
Initialize $\mathcal{I} \leftarrow \mathcal{K}$
repeat
  Solve $\mathcal{L}^{(\mathcal{K},\mathcal{I})}_{(p_{id},\mathcal{D})}(\cdot)$ to compute $\{\lambda^{(i)}_{(\mathcal{K},\mathcal{I})}\}_{i\in\mathcal{I}}$ and $\sigma^{(\mathcal{I})}$
  if $\sigma^{(\mathcal{I})} > \sigma$ **then**
    Update $\mathcal{I} \leftarrow \mathcal{I} \setminus \{\underset{i\in\mathcal{I}}{\text{argmax}}\, \lambda^{(i)}_{(\mathcal{K},\mathcal{I})}\}$
  end if
until $\sigma^{(\mathcal{I})} \leq \sigma$
Output: $\mathcal{I}$

Thus, the simple heuristic procedure Algorithm 3 is proposed for complexity reduction. This procedure stops when the number of regions cannot be reduced.

Algorithm 3 Complexity reduction.

Input: $\{\mathcal{H}^{(i)}\}_{i\in\mathbb{I}_{1:r}}$, $\{\mathcal{D}^{(i)}\}_{i\in\mathbb{I}_{1:r}}$
Initialize $s_{[0]} = r$, $\mathcal{R}^{(i)}_{[0]} = \mathcal{H}^{(i)}$, $\mathcal{D}^{(i)}_{[0]} = \mathcal{H}^{(i)}$, $\mathcal{K} = \mathbb{I}_{1:r}$
$k = 0$
repeat
  $k = k + 1$, $i = 0$
  repeat
    $i = i + 1$
    Find $\mathcal{I}_i$ is maximum mergeable subset of $\mathcal{K}$
    Update $\mathcal{K} = \mathcal{K} \setminus \mathcal{I}_i$
  until $\mathcal{K} = \emptyset$
  Update $s_{[k]} = i$, $\mathcal{K} = \mathbb{I}_{1:s_{[k]}}$
  Merge $\mathcal{R}^{(i)}_{[k]} = \bigcup_{j\in\mathcal{I}_i} \mathcal{R}^{(j)}_{[k-1]}$, $\mathcal{D}^{(i)}_{[k]} = \bigcup_{j\in\mathcal{I}_i} \mathcal{D}^{(j)}_{[k-1]}$
until $s_{[k]} = s_{[k-1]}$
Output: $\{\mathcal{R}^{(i)}\}_{i\in\mathbb{I}_{1:s}}$, $\{\mathcal{D}^{(i)}\}_{i\in\mathbb{I}_{1:s}}$

3.4 Approximate explicit model predictive control

The identification of a MPC law presents some peculiar aspects with respect to the approximation of a generic nonlinear function. Precisely, if the identified variable is one of the control inputs $q = u_i$, the construction of the weight indicator $w(\cdot)$ is worth considering:

(1) *Relaxation of input saturation.* Assume that q has saturation behavior at $q_{\text{sat}} \in \{\underline{q}, \overline{q}\}$, i.e. $q = q_{\text{sat}}$ forall $Z \in \mathcal{R}_{\text{sat}}$. In such saturation region, the approximation $B(\xi(q_{\text{sat}}))\mu \approx L^T Z$ cannot tightly holds since the r.h.s. stays unchanged while the l.h.s. varies. Hence, a low weight should be imposed for such data, i.e.

$$w(q, Z) = \rho_{\text{sat}} < 1 \text{ if } q \approx q_{\text{sat}} \tag{21}$$

(2) *Guarantee of stability at operating points.* The data set $\mathcal{D}_{\text{st}} = \{(q_{\text{st}}(k), Z_{\text{st}}(k)\}_{k=1,2,\ldots}$ including steady state points should be added into learning data. A typically high weight is imposed for such data to obtain nearly zero steady-state errors while a moderate weight for data in the neighborhood to guarantee stability, i.e.

$$w(q, Z) = \rho_{\text{st}} >> 1 \text{ if } (q, Z) \in \mathcal{D}_{\text{st}} \tag{22}$$

$$\rho_{\text{st}} > w(q, Z) = \rho_{\text{stab}} > 1 \text{ if } (q, Z) \in N(\mathcal{D}_{\text{st}}) \tag{23}$$

(3) *Guarantee of state constraints satisfaction.* Similarly, the satisfaction of some critical constraints on states

can be guaranteed by imposing a high weight for the corresponding data, i.e.

$$w(q, Z) = \rho_{\text{cstr}} > 1 \text{ if } (q, Z) \in \mathcal{D}_{\text{cstr}} \tag{24}$$

(4) *Reducing switching effect.* The proposed method does not guarantee the smoothness of MPC control laws as it is. To reduce this effect, a moderate weight is imposed on the data in the neighborhood of the separating hyperplane between two regions to obtain less discontinuities, i.e.

$$w(q, Z) = \rho_{\text{sw}} > 1 \text{ if } Z \in \bigcup_{i \in \mathbb{I}_{1:r}} N(\partial \mathcal{H}_i) \tag{25}$$

4. APPLICATION: CONTROL OF A STIRLING ENGINE BASED POWER GENERATION SYSTEM

The efficiency of the proposed method was tested in the case of the constrained control of a fast nonlinear power electronics system. All computations were performed using MATLAB, CPLEX and ACADO Toolkit [Houska et al. (2011)] on a platform with 2.6 GHz Intel Core i7, Windows OS and 16 GB of RAM.

4.1 System description and NMPC formulation

Consider the following system which belongs to a stand-alone hybrid Stirling engine / supercapacitor power plant [Alamir et al. (2014), Rahmani et al. (2013)]:

$$\begin{aligned} \dot{x}_1 &= -a_1 x_1 - a_3 x_2 + a_2 \\ \dot{x}_2 &= -a_4 x_2 + a_6 x_1 - a_7 x_3 \\ \dot{x}_3 &= a_8 (x_2 - k x_4 u) \\ \dot{x}_4 &= a_9 (-x_5^{\text{st}} + k x_3 u) \end{aligned} \tag{26}$$

where the states should satisfy positivity constraints $x \geq 0$ since the used DC/DC full bridge is not reversible. The control variable u corresponding to the duty ratio of this converter and has a strong saturation $u \in [0\ 1]$. The sampling time is $T_s = 100\ \mu s$.

The control objective is to manipulate the states x_4 tracks a reference signal x_4^r while guaranteeing state and input constraints. In Rahmani et al. (2013), it has been shown that this is a challenging problem as (26) contains highly oscillatory modes that induce constraints violation if the latter are not explicitly addressed. These makes the NMPC methodology particularly suitable. This can be achieved by minimizing the following cost function:

$$\begin{aligned} J(\mathbf{u}, x_k, x^r_{k+\cdot}, u^r_{k+\cdot}) = \sum_{i=1}^{N_p} (&||x_{k+i} - x^r_{k+i}||^2_Q \\ &+ ||u_{k+i-1} - u^r_{k+i-1}||^2_R) + \rho \sigma^2 \end{aligned} \tag{27}$$

with the prediction horizon $N_p = 3$, the cost matrices $Q = I$, $R = 1$ and the weight penalizing constraint violations $\rho = 10^4$ while fulfilling the hard constraint $0 \leq u \leq 1$, the soft constraints $[0 - \sigma,\ 4.5 - 0.1 \cdot \sigma,\ 55 - \sigma,\ 1 - \sigma]^T \leq x_k \leq [40 + \sigma,\ 5.5 + 0.1 \cdot \sigma,\ 200 + \sigma,\ 25 + \sigma]^T$ with $\sigma \geq 0$. The prediction simulator is sampled with a sampling period of $\tau_u = 1$ ms.

4.2 Computation and performance of EMPCs

The fast dynamics of the considered system would require the design of an explicit MPC controller (EMPC). In order to do this, the learning data is generated by simulating the closed loop system with the on-line NMPC during a 10 s scenario. A sparse data is extracted with the cardinality of $N = 24205$ that are then used to build the regressor instances $Z(k) = [x(k),\ x^r(k), u^r(k),\ x(k) - x^r(k)]^T \in \mathbb{R}^{13}$ as well as the corresponding instances of $q(k) = u(k)$. This learning data includes 48 data points corresponding to the steady state of the control loop. Note that the reference tracking error is incorporated into the regressor in order to further reduce the partition complexity of the explicit controller. The regression domain is defined from learning data as $\mathcal{H} = [18, 27] \times [4.4, 5.6] \times [70, 140] \times [0, 25] \times [18, 27] \times [4.738, 4.752] \times [70, 140] \times [7, 13] \times [0.35, 0.7] \times [-6, 6] \times [-0.4, 0.8] \times [-50, 30] \times [-15, 15]$. The identification parameters have been chosen as $\beta = 0.5$, $\epsilon = 1$, $n_m = 10$ and a uniform grid of $n_\xi = 50$ points. The weight indicator is defined by $\rho_{\text{st}} = 10^2$, $\rho_{\text{stab}} = 10$ and $\rho_{\text{sw}} = 2$. Several approximate controllers have been identified with different values of small bounded-error σ. Fig. 2 provides a complexity assessment of identified controllers which illustrate the low-complexity of EMPCs as well as the efficiency of the complexity reduction procedure. The nonlinearity as well as the strict monotonicity of several scalar maps corresponding to a 17-regions 0.005-error EMPC are illustrated in Fig. 3.

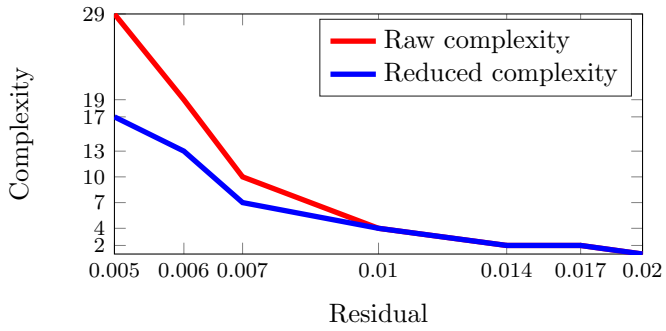

Fig. 2. Complexity versus residual

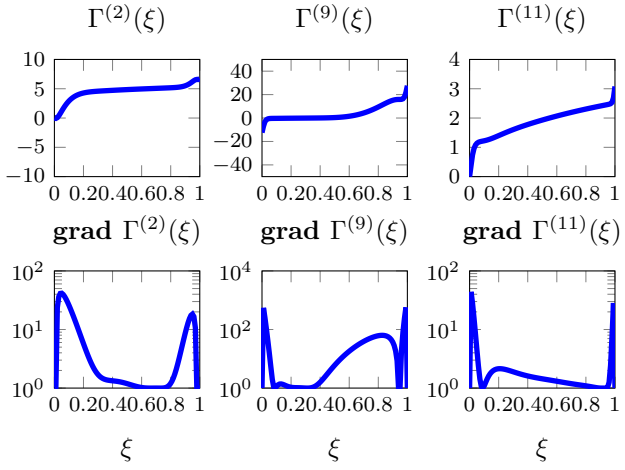

Fig. 3. Nonlinearity of 17-regions 0.005-error EMPC

The performance of two EMPCs is evaluated by simulations and is compared with the closed-loop by the on-line implicit NMPC delivered by ACADO. Both the on-line

NMPC and the EMPCs have been compiled in C language. Fig. 4 shows the evolution of control inputs, region indexes of the EMPCs and the system states. It is clear that the tracking performance of those EMPCs is acceptable while the constraint satisfaction of the on-line NMPC is inherited. Regarding real-time aspect, Fig. 5 illustrates the computation efficiency of the EMPCs as their computation time vary between 10 μs and 60 μs, is much less than that of ACADO and obviously suitable for this application.

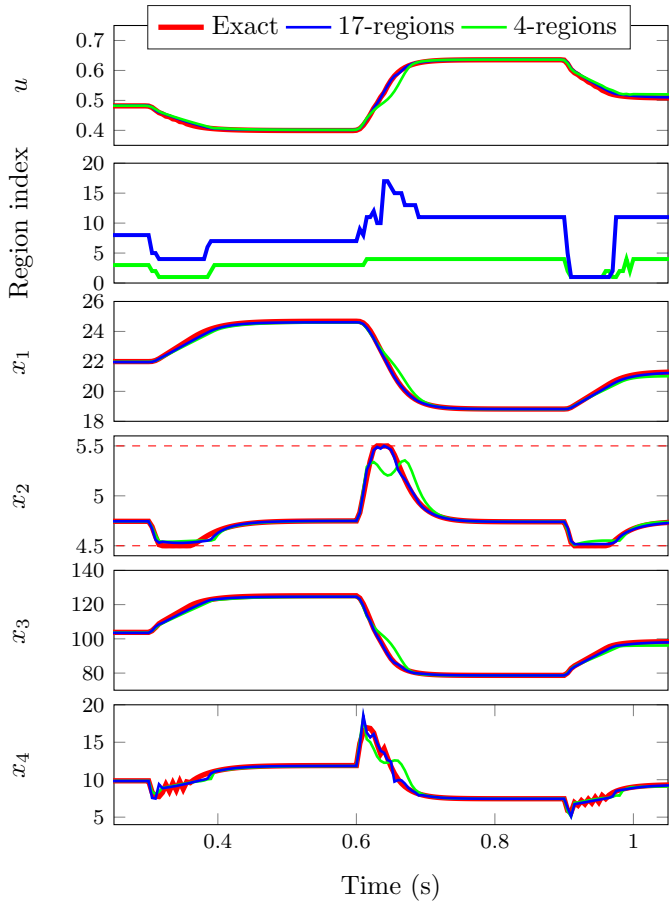

Fig. 4. Performance comparison between controllers

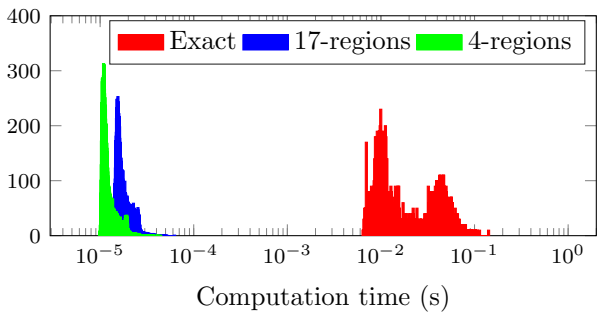

Fig. 5. Histogram of computation time

5. CONCLUSION AND FUTURE WORK

In this paper, a new approach is proposed for nonlinear piecewise approximations of MPC control laws. Undergoing work addresses the extension of the approach to semi-explicit MPC and distributed MPC settings.

REFERENCES

D. Q. Mayne, J. B. Rawlings, C. V. Rao and P. O. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*. 36(6): 789-814, 2000.

A. Alessio and A. Bemporad. *A Survey on Explicit Model Predictive Control*, chapter Nonlinear Model Predictive Control: Towards New Challenging Applications, pages 345–369, Springer-Verlag, 2009.

A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos. The Explicit Linear Quadratic Regulator for Constrained Systems. *Automatica*, 38(1): 3-20, 2002.

T. A. Johansen. Approximate Explicit Receding Horizon Control of Constrained Nonlinear Systems. *Automatica*. 40(2): 293-300, 2004.

A. Grancharova and T. A. Johansen. *Explicit Nonlinear Model Predictive Control: Theory and Applications*. Lecture Notes in Control and Information Sciences. Springer-Verlag, 2012.

M. Canale, L. Fagiano, and M. Milanese. Set Membership Approximation Theory For Fast Implementation of Model Predictive Control. *Automatica*, 45(1): 45–54, 2009.

L. Fagiano, M. Canale, and M. Milanese. Set Membership Approximation of Discontinuous Nonlinear Model Predictive Control Laws. *Automatica*, 48(1): 191-197, 2012.

G. Pin, M. Filippo, F. A. Pellegrino, G. Fenu, and T. Parisini. Approximate Model Predictive Control Laws for Constrained Nonlinear Discrete-Time Systems: Analysis and Offline Design, *International Journal of Control*, 86(5): 804-820, 2013.

A. Grancharova and S. Olaru. Explicit Interpolation-Based Nonlinear Model Predictive Control With A Convex Approximation Of the Feasible set. IEEE Conference on Control Applications, CCA2014.

M. Alamir. A New Identification Framework for Off-Line Computation of Moving-Horizon-Observers. *IEEE Transactions on Automatic Control*, 58(7): 1877-1882, 2013.

M. Alamir, D. Alberer and L. Del Re. Identification of A Class of Nonlinear Dynamic Relationships: Application to The Identification of Engine Emission Models. *International Journal of Engine Research*, 15(8): 898-905, 2014.

M. Alamir, M. A. Rahmani, D. Gualino. Generic Model Identification Framework for Thermodynamic Engines for Use In Hybrid Power Stations Control And Simulation. *Journal of Process Control*, 24(6): 966-974, 2014.

M. A. Rahmani, M. Alamir, and D. Gualino. Nonlinear Dynamic Model Identification and MPC control of an Organic Rankine Cycle (ORC) based Solar Thermal Power Plant. European Control Conference, ECC2015.

B. Houska and H.J. Ferreau and M. Diehl. ACADO Toolkit – An Open Source Framework for Automatic Control and Dynamic Optimization. *Optimal Control Applications and Methods*. 32(3): 298-312, 2011.

M. Alamir, M. A. Rahmani and D. Gualino. Constrained Control Framework for A Stand-Alone Hybrid (Stirling Engine) / Supercapacitor Power Generation System. *Applied Energy*. 118(1): 192-206, 2014.

M. A. Rahmani, M. Alamir and D. Gualino. Control Strategy for An Off-Grid Hybrid Stirling Engine / Supercapacitor Power Generation System. American Control Conference, ACC2013.