Efficient model predictive control for nonlinear systems via function approximation techniques

Massimo Canale, *Member, IEEE,* Lorenzo Fagiano, *Member, IEEE,* and Mario Milanese, *Senior Member, IEEE,*

***Abstract*—This note deals with the problem of fast implementation of nonlinear model predictive control using approximated control laws. At first, accuracy properties of a generic approximated controller $\hat{\kappa}$ are introduced together with their influence on closed loop stability and performance. Then, exploiting such results, it is shown how Set Membership (SM) function approximation theory can be systematically employed to improve the accuracy performance of $\hat{\kappa}$. The resulting controller, given by the sum of $\hat{\kappa}$ with a SM approximating function, satisfies the above–mentioned properties even if they are not met by $\hat{\kappa}$ alone. A nonlinear oscillator example shows the effectiveness of the proposed methodology.**

***Index Terms*—Nonlinear model predictive control, approximate predictive control, constrained control.**

A viable solution to the problem of fast implementation of a given Nonlinear Model Predictive Control (NMPC) law κ (see e.g. [1]) is the use of an approximated control law $\hat{\kappa} \approx \kappa$, with lower computational effort, derived using a finite number ν of exact control moves computed off–line. A first contribution along this line was given in [2], using a neural network approximation of κ. However, no means of evaluating the guaranteed approximation error were given, thus no a priori or a posteriori stability analysis could be systematically performed. Another methodology to approximate a NMPC controller has been introduced in [3], using an off–line approximate multi–parametric programming technique for the construction of a piecewise affine (PWA) approximation. Similar approaches have been applied also in the case of linear systems (see e.g. [4], [5]) and extended to the robust min–max case in [6], using a piecewise nonlinear (PWNL) approximation. A bound on the approximation error, in terms of difference between the exact and approximated optimal cost functions, can be obtained in these cases, as well as stability and constraint satisfaction properties. Finally, a further technique has been recently introduced in [7] using Set Membership (SM) function approximation theory, obtaining approximated MPC laws with guaranteed stability, constraint satisfaction and state regulation to an arbitrarily small neighborhood of the origin. In [7] it has been shown that such stabilizing properties rely on the satisfaction of three assumptions by function $\hat{\kappa}$, namely i) input constraint satisfaction, ii) the possibility to estimate a bound ζ on the worst–case approximation error, directly in terms of the difference $\kappa - \hat{\kappa}$, and iii) the ability to reduce ζ to any desired guaranteed accuracy level. In this note we provide theoretical contributions and practical methods that deal with the problems of a) finding out sufficient conditions for a generic approximated NMPC law $\hat{\kappa}$ to satisfy properties i)–iii), b) providing a method to compute the value of ζ and c) deriving a technique able to improve the accuracy of a given preliminary approximating function $\hat{\kappa}$, obtaining lower (eventually minimal) values of ζ and satisfying assumptions i)–iii) even if they are not met by $\hat{\kappa}$ alone.

The authors are with the Dipartmento di Automatica e Informatica, Politecnico di Torino, Torino, Italy. e-mail: {massimo.canale, lorenzo.fagiano, mario.milanese}@polito.it

I. Approximated Nonlinear Model Predictive Control

A. NMPC: nominal formulation

Consider the following nonlinear state space model:

$$x_{t+1} = f(x_t, u_t) \tag{1}$$

where $x_t \in \mathbb{R}^n$ and $u_t \in \mathbb{R}^m$ are the system state and control input respectively. It is assumed that function f in (1) is continuous over $\mathbb{R}^n \times \mathbb{R}^m$ and that the control objective is to regulate the system state to the origin under some input and state constraints represented by a compact set $\mathbb{U} \subseteq \mathbb{R}^m$ and a convex set $\mathbb{X} \subseteq \mathbb{R}^n$ respectively, both containing the origin in their interiors. In NMPC, the control move is computed on–line by solving the following nonlinear program (NLP) at each time step, according to the Receding Horizon (RH) principle (see e.g. [1]):

$$\begin{aligned} &\min_{U} \quad J(U, x_{t|t}) \\ &\text{s. t.} \\ &\begin{cases} x_{t+k|t} = f(x_{t+k-1|t}, u_{t+k-1|t}),\ k = 1, \dots, N_p \\ x_{t+k|t} \in \mathbb{X},\ k = 1, \dots, N_p \\ u_{t+k|t} \in \mathbb{U},\ k = 0, \dots, N_p - 1 \\ \textit{stabilizing constraints} \end{cases} \end{aligned} \tag{2}$$

where $J(U, x_{t|t}) = \Phi(x_{t+N_p|t}) + \sum_{j=0}^{N_p-1} L(x_{t+j|t}, u_{t+j|t})$, $x_{t|t} = x_t$ and $U = [u_{t|t}^T, \dots, u_{t+N_c-1|t}^T]^T$ is the vector of the control moves to be optimized. N_p and $N_c \leq N_p$ are the prediction horizon and the control horizon respectively. The remaining predicted control moves $[u_{t+N_c|t}, \dots, u_{t+N_p-1|t}]$ can be computed according to different strategies, e.g. by setting $u_{t+j|t} = u_{N_c-1|t}$ or $u_{t+j|t} = K\, x_{t+j|t}$, $\forall j \in [N_c, N_p - 1]$, where K is a suitable matrix. The cost functions L and Φ have to be suitably chosen and tuned according to the desired control performance (see e.g. [1] for details). Indeed, possible additional "*stabilizing constraints*" (e.g. state contraction, terminal set) have been included in (2) in order to ensure stability of the controlled system. It is assumed that the optimization problem (2) is feasible over a set $\mathcal{F} \subseteq \mathbb{R}^n$. The application of such RH procedure implicitly defines the predictive controller as a nonlinear static function κ of the state, i.e. $u_t = \kappa(x_t)$. In the sequel, function κ will be denoted as the "exact"

or "nominal" control law. It is assumed that the nonlinear autonomous system obtained by applying control law κ to the system (1) is asymptotically stable at the origin for any initial state value $x_0 \in \mathcal{F}$ (see e.g. [8]).

B. NMPC approximation: stability results

A limitation in the practical use of NMPC is the presence of fast plant dynamics, for which the required sampling time may be too low for the real-time solution of (2). A viable solution to this problem is the use of an approximated control function $\hat{\kappa} \approx \kappa$, derived off–line, whose on–line computational load is lower (see [2]–[7]). A crucial point arising when the approximated function $\hat{\kappa}$ is employed for feedback control is about the stability properties of the resulting closed loop system. In this Section, the sufficient conditions for stability introduced in [7] will be briefly resumed. It is considered that $\hat{\kappa}$ is defined over a compact set $\mathcal{X} \subseteq \mathcal{F}$, containing the origin in its interior, i.e. $\hat{\kappa} : \mathcal{X} \rightarrow \mathbb{R}^m$, $\mathcal{X} \subseteq \mathcal{F}$. $\mathcal{X}$ is the compact set where the approximation is carried out and in the practice it is usually chosen as a set of state values of interest for the particular control problem, e.g. by considering the closed loop trajectories obtained through numerical simulation tests. Moreover, it is supposed that the following assumption holds:

Assumption 1 (A1): *the nominal control law κ is continuous over $\mathcal{X}$.*

Assumption **A1** depends on the characteristics of the NLP (2): results on this aspect can be found e.g. in [9] and [10] and the references therein. Function $\hat{\kappa}$ is computed on the basis of the knowledge of a finite number ν of exact control moves, i.e.:

$$\tilde{u}^k = \kappa(\tilde{x}^k), k = 1, \ldots, \nu \tag{3}$$

where the state values $\tilde{x}^k$ are suitably chosen and give rise to the set $\mathcal{X}_\nu = \{\tilde{x}^k,\, k = 1, \ldots, \nu\} \subseteq \mathcal{F}$. It is assumed that $\mathcal{X}_\nu$ is chosen such that the following property holds:

$$\lim_{\nu \rightarrow \infty} d(\mathcal{X}, \mathcal{X}_\nu) = 0 \tag{4}$$

where $d(\mathcal{X}, \mathcal{X}_\nu) = \sup_{x \in \mathcal{X}} \inf_{\tilde{x} \in \mathcal{X}_\nu} (\|x - \tilde{x}\|_2)$. For example, uniform gridding over $\mathcal{X}$ satisfies condition (4). Moreover, $\hat{\kappa}$ is supposed to enjoy the following *key properties*:

I) input constraint satisfaction. For the sake of simplicity of presentation, in this paper it will be assumed that $\mathbb{U} = \{u \in \mathbb{R}^m : \underline{u}_i \leq u_i \leq \overline{u}_i,\, i = 1, \ldots, m\}$. Thus, the considered property is the following:

$$\underline{u}_i \leq \hat{\kappa}_i(x) \leq \overline{u}_i,\ \forall i \in [1, m],\ \forall x \in \mathcal{X} \tag{5}$$

II) the pointwise approximation error $\Delta_{\hat{\kappa}}(x) \doteq \kappa(x) - \hat{\kappa}(x)$ is norm bounded:

$$\|\Delta_{\hat{\kappa}}(x)\| \leq \zeta,\ \forall x \in \mathcal{X} \tag{6}$$

III) the bound $\zeta(\nu)$ converges to zero as the number ν of the off–line computed solutions increases:

$$\lim_{\nu \rightarrow \infty} \zeta(\nu) = 0 \tag{7}$$

In Theorem 2 of [7], which can be straightforwardly extended to the case of nonlinear systems, it is shown that if the approximated controller $\hat{\kappa}$ enjoys (5)–(7), then closed loop trajectory boundedness, guaranteed bounded trajectory error, state constraint satisfaction (after a finite number of time steps) and trajectory convergence to an arbitrarily small neighborhood of the origin can be achieved. Indeed, several studies in the literature (see e.g. [11] and the references therein) indicate that, if κ is stabilizing and continuous, the possibility to achieve an arbitrarily small approximation error bound ζ is sufficient for closed loop trajectory boundedness and convergence to a neighborhood of the origin. Thus, this paper focuses on giving methods that can be practically applied to evaluate and improve the guaranteed accuracy of a given function $\hat{\kappa}$.

II. Approximated NMPC laws: accuracy results

In this Section, sufficient conditions for a generic approximation $\hat{\kappa}$ to satisfy properties (5)– (7), i.e. to be able to guarantee closed loop stability, will be given. The same result also allows to compute an estimate of the worst–case approximation error, which can be practically employed to evaluate the accuracy of $\hat{\kappa}$ and to compare different approximating functions. In the remaining of the paper, it is implicitly meant that any i is considered and notation "$\forall i : i = 1, \ldots, m$" is omitted for simplicity of reading. The results of this paper are derived in the framework of SM theory. In this context, the available prior information on the function κ_i to be approximated has to be considered and it will be now resumed. Since κ_i is continuous over the compact $\mathcal{X}$ and its image set $\mathbb{U}$ is also compact, κ_i is Lipschitz continuous over $\mathcal{X}$ with Lipschitz constant $L_{\kappa,i}$. Thus, the available information on κ_i defines the following Feasible Function Set (FFS): $\kappa_i \in FFS_{L_{\kappa,i}} \doteq \{\kappa_i : \mathcal{X} \rightarrow [\underline{u}_i, \overline{u}_i] : \kappa_i \in \mathcal{A}_{L_{\kappa,i}};\ \kappa_i(\tilde{x}) = \tilde{u}_i,\ \forall \tilde{x} \in \mathcal{X}_\nu\}$, where:

$$\mathcal{A}_{L_{\kappa,i}} = \{\kappa_i : |\kappa_i(x^1) - \kappa_i(x^2)| \leq L_{\kappa,i}\|x^1 - x^2\|_2, \quad \forall x^1,\, x^2 \in \mathcal{X}\} \tag{8}$$

Estimates $\hat{L}_{\kappa,i}$ of $L_{\kappa,i}$ can be derived as follows:

$$\hat{L}_{\kappa,i} = \inf\left(\tilde{L}_i : \tilde{u}_i^h + \tilde{L}_i\|\tilde{x}^h - \tilde{x}^k\|_2 \geq \tilde{u}_i^k,\ \forall k, h = 1, \ldots, \nu\right) \tag{9}$$

Convergence of (9) to $L_{\kappa,i}$ has been showed in [7]. The same procedure as (9) can be employed to practically estimate also the other Lipschitz constants involved in the following theoretical developments. Sufficient conditions are now given for any approximating function $\hat{\kappa}_i$ to obtain a bound ζ_i on the pointwise approximation error norm $|\Delta_{\hat{\kappa},i}(x)| = |\kappa_i(x) - \hat{\kappa}_i(x)|$ and, consequently, to satisfy property (6). From the knowledge of the ν exact control moves computed off–line (3), the exact values of $\Delta_{\hat{\kappa},i}(\tilde{x})$ are known: $\Delta_{\hat{\kappa},i}(\tilde{x}) = \tilde{u}_i - \hat{\kappa}_i(\tilde{x}),\ \forall \tilde{x} \in \mathcal{X}_\nu$. The following Theorem shows how to compute a bound on $|\Delta_{\hat{\kappa},i}(x)|$ on the basis of the knowledge of $\Delta_{\hat{\kappa},i}(\tilde{x})$.

Theorem 1: Suppose that $\kappa_i \in FFS_{L_{\kappa,i}}$ and that **I)** $\hat{\kappa}_i$ satisfies property (5) and **II)** it is continuous over $\mathcal{X}$ (thus it is Lipschitz continuous over $\mathcal{X}$ with constant $L_{\hat{\kappa},i}$), then:
i) the approximation error $\Delta_{\hat{\kappa},i}$ is a Lipschitz continuous function over $\mathcal{X}$, with constant $L_{\Delta_{\hat{\kappa},i}}$

ii) $|\Delta_{\hat{\kappa},i}(x)|$ is bounded: $|\Delta_{\hat{\kappa},i}(x)| \leq \zeta_i,\ \forall x \in \mathcal{X}$
iii) a bound ζ_i can be computed as:

$$\zeta_i = \sup_{x\in\mathcal{X}} \max\left(\overline{\Delta}_{\hat{\kappa},i}(x), -\underline{\Delta}_{\hat{\kappa},i}(x)\right) \tag{10}$$

where

$$\begin{aligned}\overline{\Delta}_{\hat{\kappa},i}(x) &\doteq \min[\overline{u}_i - \hat{\kappa}_i(x), \min_{\tilde{x}\in\mathcal{X}_\nu} \left(\Delta_{\hat{\kappa},i}(\tilde{x}) + L_{\Delta_{\hat{\kappa},i}}\|x-\tilde{x}\|_2\right)]\\ \underline{\Delta}_{\hat{\kappa},i}(x) &\doteq \max[\underline{u}_i - \hat{\kappa}_i(x), \max_{\tilde{x}\in\mathcal{X}_\nu} \left(\Delta_{\hat{\kappa},i}(\tilde{x}) - L_{\Delta_{\hat{\kappa},i}}\|x-\tilde{x}\|_2\right)]\end{aligned} \tag{11}$$

iv) if $L_{\hat{\kappa},i} \leq L_{\Delta_{\hat{\kappa},i}}$, the bound ζ_i (10) is tight according to the available information on κ_i.
In addition, if **III)** $\hat{\kappa}$ satisfies the data interpolation property:

$$\hat{\kappa}_i(\tilde{x}) = \kappa_i(\tilde{x}) = \tilde{u}_i,\ \forall \tilde{x} \in \mathcal{X}_\nu \tag{12}$$

the following results hold:
v) the bound ζ_i on the approximation error can be computed as:

$$\begin{aligned}&\zeta_i = \sup_{x\in\mathcal{X}} \min\left[\max\left(\overline{u}_i - \hat{\kappa}_i(x), -\underline{u}_i + \hat{\kappa}_i(x)\right), \chi_i(x)\right],\\ &\text{where } \chi_i(x) = \min_{\tilde{x}\in\mathcal{X}_\nu} \left(L_{\Delta_{\hat{\kappa},i}}\|x-\tilde{x}\|_2\right)\end{aligned} \tag{13}$$

vi) $\zeta_i(\nu)$ converges to zero: $\lim_{\nu\to\infty} \zeta_i(\nu) = 0$

Proof: See the Appendix. ■

Theorem 1 can be used to compute an upper bound $\zeta_i(\nu)$ on the worst–case approximation error $e(\kappa_i, \hat{\kappa}_i) = \|\kappa_i - \hat{\kappa}_i\|_\infty = \sup_{x\in\mathcal{X}} |\kappa_i(x) - \hat{\kappa}_i(x)|$ obtained using the approximated control law $\hat{\kappa}$. The error bound (6) related to function $\hat{\kappa}$ can be computed as $\zeta = \sqrt{\sum_{i=1}^{m} (\zeta_i)^2}$. Then, if also (12) holds, it is possible to "tune" ν (using an iterative procedure) in order to obtain a given desired guaranteed accuracy. Practical computation of ζ_i can be performed using the results of [12].

III. Set membership local approximation of NMPC laws

In this Section, the problem of improving the accuracy of a generic approximating function $\hat{\kappa}_i$ fulfilling the hypotheses **I)**–**II)** of Theorem 1 is studied. Standard methods, e.g. based on expansions in term of suitable basis functions (polynomials, sigmoids, wavelets, etc.) can be used to compute $\hat{\kappa}$. However in general, as the number of basis functions is increased in order to achieve the interpolation condition (12), the approximation error $e(\kappa_i, \hat{\kappa}_i)$ may become very large. On the other hand, the approximation $\hat{\kappa}_i$ may give useful information on κ_i. Such information, together with the prior information given by $\kappa_i \in FFS_{L_{\kappa,i}}$, can be exploited in the SM framework to derive an approximated control law with better accuracy with respect to $\hat{\kappa}_i$ itself, giving lower (possibly minimal) worst–case error bound ζ_i and satisfying condition (12) even if $\hat{\kappa}_i$ alone does not satisfies it. Following this idea, a new SM technique for approximate NMPC, denoted as the "local" SM approximation κ_i^{LOC}, is now presented.
Consider the residue function $\Delta_{\hat{\kappa},i} = \kappa_i - \hat{\kappa}_i$ which, on the basis of Theorem 1, is Lipschitz continuous over $\mathcal{X}$ with Lipschitz constant $L_{\Delta_{\hat{\kappa},i}}$. The available information on κ_i and on $\hat{\kappa}_i$ can be summarized by the following set $FFS_{\Delta,i}$:

$$\begin{aligned}FFS_{\Delta,i} = \{\kappa_i : \mathcal{X} \to [\underline{u}_i, \overline{u}_i],\ &(\kappa_i - \hat{\kappa}_i) \in \mathcal{A}_{L_{\Delta_{\hat{\kappa},i}}},\\ &\kappa_i(\tilde{x}) = \tilde{u}_i,\ \forall \tilde{x} \in \mathcal{X}_\nu\}\end{aligned} \tag{14}$$

where

$$\begin{aligned}\mathcal{A}_{L_{\Delta_{\hat{\kappa},i}}} = \big\{\Delta_i : \mathcal{X} \to \mathbb{R},\ &|\Delta_i(x^1) - \Delta_i(x^2)| \leq\\ &L_{\Delta_{\hat{\kappa},i}}\|x^1 - x^2\|_2,\ \forall x^1, x^2 \in \mathcal{X}\big\}.\end{aligned}$$

Define the following function:

$$\kappa_i^{\text{LOC}} \doteq \hat{\kappa}_i + \Delta_{\hat{\kappa},i}^{\text{OPT}},\ \text{where } \Delta_{\hat{\kappa},i}^{\text{OPT}}(x) \doteq \frac{1}{2}[\overline{\Delta}_{\hat{\kappa},i}(x) + \underline{\Delta}_{\hat{\kappa},i}(x)] \tag{15}$$

with $\overline{\Delta}_{\hat{\kappa},i}(x)$ and $\underline{\Delta}_{\hat{\kappa},i}(x)$ defined in (11).

Theorem 2: For any given approximating function $\hat{\kappa}_i$ of $\kappa_i \in FFS_{L_{\kappa,i}}$ such that **I)** $\hat{\kappa}_i$ satisfies property (5) and **II)** it is continuous over $\mathcal{X}$, the corresponding function κ_i^{LOC} (15) enjoys the following properties:
i) κ_i^{LOC} is Lipschitz continuous over $\mathcal{X}$, satisfies property (5) and interpolates the off–line computed data: $\kappa_i^{\text{LOC}}(\tilde{x}) = \tilde{u}_i,\ \forall \tilde{x} \in \mathcal{X}_\nu$
ii) the quantity

$$\zeta_i^{\text{LOC}} \doteq \sup_{x\in\mathcal{X}} \frac{1}{2}\left(\overline{\Delta}_{\hat{\kappa},i}(x) - \underline{\Delta}_{\hat{\kappa},i}(x)\right) \tag{16}$$

is a bound on the worst–case approximation error $e(\kappa_i, \kappa_i^{\text{LOC}})$.
iii) $\zeta_i^{\text{LOC}} \leq \zeta_i$, where ζ_i is the error bound related to the preliminary approximating function $\hat{\kappa}$, computed according to (10).
Additionally, if **III)** $L_{\hat{\kappa},i} \leq L_{\Delta_{\hat{\kappa},i}}$ then:
iv) $\kappa_i^{\text{LOC}} \in FFS_{\Delta,i}$
v) κ_i^{LOC} is an optimal approximation of κ_i with respect to the information $\kappa_i \in FFS_{\Delta,i}$, i.e.:

$$\begin{aligned}&\sup_{\kappa_i\in FFS_{\Delta,i}} e(\kappa_i, \kappa_i^{\text{LOC}}) =\\ &\inf_{\tilde{\kappa}_i\in FFS_{\Delta,i}} \sup_{\kappa_i\in FFS_{\Delta,i}} e(\kappa_i, \tilde{\kappa}_i) = r_{\Delta,\infty,i}\end{aligned}$$

where $r_{\Delta,\infty,i}$ is the ∞-norm radius of information of $FFS_{\Delta,i}$ (see [13]).

Proof: See the Appendix. ■

Theorem 2 shows how SM theory can be employed to improve the performance of a given approximating function $\hat{\kappa}_i$. In fact, result **iii)** of Theorem 2 shows that the error bound ζ_i^{LOC} of the approximated NMPC law κ_i^{LOC} is lower than that of $\hat{\kappa}_i$. Moreover, from result **i)**, κ_i^{LOC} satisfies the data interpolation condition (12), even if $\hat{\kappa}_i$ does not satisfy it. Finally, under assumption **III)** of Theorem 2, κ_i^{LOC} also gives the minimal worst–case error bound according to the considered prior information (result **v)** of Theorem 2). The error bound (6) related to function $\kappa^{\text{LOC}} = [\kappa_1^{\text{LOC}}, \ldots, \kappa_m^{\text{LOC}}]^T$ can be computed as $\zeta^{\text{LOC}} = \sqrt{\sum_{i=1}^{m} \left(\zeta_i^{\text{LOC}}\right)^2}$. Without efficient implementations, the computational effort required to compute κ^{LOC} grows linearly with ν. Such a growth can be made logarithmic by using suitable techniques like those proposed e.g. by [14], with the only differences that [14] does not considers constraints and employs the simplicial distance between two points x_1, x_2 in

the Lipschitz assumption, while in this paper the Euclidean norm is considered (see (8)).

IV. Numerical example

Consider the discrete–time, two-dimensional nonlinear oscillator obtained by forward difference approximation of the Duffing equation (see e.g. [15]) using sampling time $T_s = 0.05$ s:

$$x_{t+1} = \begin{bmatrix} 1 & T_s \\ -T_s & (1-0.6\,T_s) \end{bmatrix} x_t + \begin{bmatrix} 0 \\ T_s \end{bmatrix} u_t + \begin{bmatrix} 0 & 0 \\ -T_s & 0 \end{bmatrix} x_t^3$$

The input and state constraint sets are $\mathbb{U} = \{u \in \mathbb{R} : |u| \leq 5\}$ and $\mathbb{X} = \{x \in \mathbb{R}^2 : \|x\|_\infty \leq 3\}$ respectively. In the NLP (2), horizons $N_p = 30$, $N_c = 20$ and functions $L(x,u) = x^T Q x + u^T R u$, $\Phi = 0$ have been chosen, with $Q = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$, and $R = 2$. The terminal state constraint $x_{t+N_p|t} = 0$ has been also included to guarantee nominal closed loop stability (see [1]). The state prediction has been performed by setting $u_{t+j|t} = u_{t+N_c-1|t},\ j = N_c, ..., N_p - 1$. Fig. 1 shows the set $\mathcal{X}$ considered for the approximation, together with the constraint set $\mathbb{X}$ and the level curves of the optimal cost function $J^*(x) = \min_U J(U,x)$. It can be noted that $J^*(x)$ is not convex, thus the technique proposed in [3] can be applied only with modifications like heuristic splitting–rules (see e.g. [6] and the references therein). On the other hand, continuity of κ was checked by numerical inspection, thus the techniques proposed in this paper can be employed. Two neural approximations of κ have been computed, denoted

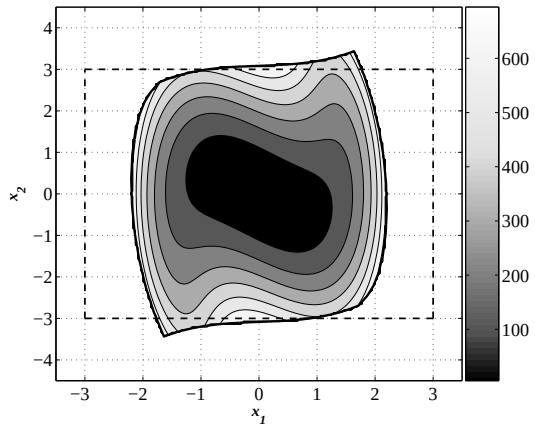

Fig. 1. Set $\mathcal{X}$ (solid), state constraint set $\mathbb{X}$ (dashed) and level curves of the optimal cost function $J^*(x)$.

as $\hat{\kappa}^{\text{NN},1}$ and $\hat{\kappa}^{\text{NN},2}$. Both neural networks are composed of two layers with l nodes in the first layer (with sigmoidal activating function) and a single linear node in the second layer. Values of $l = 3$ and $l = 10$ have been used for $\hat{\kappa}^{\text{NN},1}$ and $\hat{\kappa}^{\text{NN},2}$ respectively. In order to satisfy condition (5), the outputs of $\hat{\kappa}^{\text{NN},1}$ and $\hat{\kappa}^{\text{NN},2}$ have been saturated to $[-5, 5]$. The neural networks have been trained using 10^4 off–line computed control moves, chosen with uniform gridding over $\mathcal{X}$. Then, the local SM approximations $\kappa^{\text{LOC},1}$ and $\kappa^{\text{LOC},2}$, derived using as preliminary approximations $\hat{\kappa}^{\text{NN},1}$ and $\hat{\kappa}^{\text{NN},2}$ respectively, have been computed using different values of ν, considering uniform gridding over $\mathcal{X}$. The following estimated values of the involved Lipschitz constants have been obtained using the procedure (9): $\hat{L}_{\hat{\kappa}^{\text{NN},1}} = 3.022$, $\hat{L}_{\hat{\kappa}^{\text{NN},2}} = 4.118$, $\hat{L}_{\Delta_{\hat{\kappa}^{\text{NN},1}}} = 3.944$ and $\hat{L}_{\Delta_{\hat{\kappa}^{\text{NN},2}}} = 1.912$. The obtained accuracy results are reported in Tables I and II, in terms of the worst–case error bounds $\zeta^{\text{NN},1}$, $\zeta^{\text{NN},2}$ of the neural networks (obtained applying (10)), of the worst–case error bounds $\zeta^{\text{LOC},1}$, $\zeta^{\text{LOC},2}$ of the corresponding local SM approximations (obtained with (16)) and of the maximal approximation error $\Delta^{\text{MAX}}_{\hat{\kappa}} = \max_{r=1,\ldots,\xi} |\kappa(\check{x}^r) - \hat{\kappa}(\check{x}^r)|$ computed by considering for each of the approximated controllers the same set $\mathcal{X}_\xi = \{\check{x}^r,\, r = 1, \ldots, 10^4\} \subseteq \mathcal{X}$, containing a number $\xi = 10^4$ points $\check{x}^r \in \mathcal{X}$ chosen with uniformly random distribution. The practical computation of the worst–case bounds has been carried out using the techniques of [12]. As expected, with

TABLE I
Performance results of the two neural network approximations

$\hat{\kappa}^{\text{NN},1}$				
$\zeta^{\text{NN},1}$	$\Delta^{\text{MAX}}_{\hat{\kappa}^{\text{NN},1}}$	$\bar{t}^{\text{NN},1}$ (s)	$\bar{d}^{\text{NN},1}_{\%}$ (%)	$\bar{d}^{\text{NN},1}$
1.7	1.5	$3.0\,10^{-5}$	$8.8\,10^{1}$	$5.6\,10^{-2}$
$\hat{\kappa}^{\text{NN},2}$				
$\zeta^{\text{NN},2}$	$\Delta^{\text{MAX}}_{\hat{\kappa}^{\text{NN},2}}$	$\bar{t}^{\text{NN},2}$ (s)	$\bar{d}^{\text{NN},2}_{\%}$ (%)	$\bar{d}^{\text{NN},2}$
$2.0\,10^{-1}$	$4.6\,10^{-2}$	$3.0\,10^{-5}$	5.0	$3.4\,10^{-3}$

TABLE II
Performance results of the two local SM approximations

$\kappa^{\text{LOC},1}$					
ν	$\zeta^{\text{LOC},1}$	$\Delta^{\text{MAX}}_{\kappa^{\text{LOC},1}}$	$\bar{t}^{\text{LOC},1}$ (s)	$\bar{d}^{\text{LOC},1}_{\%}$ (%)	$\bar{d}^{\text{LOC},1}$
$1.0\,10^{3}$	$5.4\,10^{-1}$	$4.1\,10^{-1}$	$1.1\,10^{-4}$	5.8	$3\,10^{-3}$
$8.0\,10^{3}$	$1.5\,10^{-1}$	$7.1\,10^{-2}$	$2.0\,10^{-3}$	1.6	$1.2\,10^{-3}$
$3.0\,10^{4}$	$9.1\,10^{-2}$	$5.0\,10^{-2}$	$6.0\,10^{-3}$	$3.1\,10^{-1}$	$3.5\,10^{-4}$
$7.5\,10^{4}$	$4.9\,10^{-2}$	$3.0\,10^{-2}$	$1.9\,10^{-2}$	$2.0\,10^{-1}$	$1.6\,10^{-4}$
$\kappa^{\text{LOC},2}$					
ν	$\zeta^{\text{LOC},2}$	$\Delta^{\text{MAX}}_{\kappa^{\text{LOC},2}}$	$\bar{t}^{\text{LOC},2}$ (s)	$\bar{d}^{\text{LOC},2}_{\%}$ (%)	$\bar{d}^{\text{LOC},2}$
$1.0\,10^{3}$	$1.5\,10^{-1}$	$3.0\,10^{-2}$	$1.1\,10^{-4}$	1.2	$0.8\,10^{-3}$
$8.0\,10^{3}$	$8.2\,10^{-2}$	$8.4\,10^{-3}$	$2.0\,10^{-3}$	$3.2\,10^{-1}$	$2.0\,10^{-4}$
$3.0\,10^{4}$	$3.0\,10^{-2}$	$3.4\,10^{-3}$	$6.0\,10^{-3}$	$7.0\,10^{-2}$	$1.0\,10^{-4}$
$7.5\,10^{4}$	$2.0\,10^{-2}$	$1.1\,10^{-3}$	$1.9\,10^{-2}$	$3.0\,10^{-2}$	$7.5\,10^{-5}$

all the considered approximations the maximal computed error is lower than the corresponding worst–case bound. Moreover, it can be noted that, according to the presented results, the guaranteed error bounds related to functions $\hat{\kappa}^{\text{NN},1}$ and $\kappa^{\text{LOC},1}$ are much tighter since the condition $\hat{L}_{\hat{\kappa}^{\text{NN},1}} \leq \hat{L}_{\Delta_{\hat{\kappa}^{\text{NN},1}}}$ is satisfied. As regards the improvements obtained with the local SM approach, notice that with any of the considered values of ν, the proposed technique achieves better accuracy than either $\hat{\kappa}^{\text{NN},1}$ or $\hat{\kappa}^{\text{NN},2}$, in terms both of theoretical worst–case bounds and "actual" maximal approximation errors. Moreover, it can be remarked that, as expected, the worst–case bounds $\zeta^{\text{LOC},1}$ and $\zeta^{\text{LOC},2}$ decrease as ν increases. Thus, this numerical example shows that, according to the theoretical results of the paper, **a)** a bound on the worst–case error of a given continuous preliminary approximating function can be computed, **b)** the local SM technique can be employed to systematically improve the performance of any continuous preliminary approximating function, **c)** an arbitrary small guaranteed worst–case error can be obtained with the local SM approach and **d)** a tradeoff

between accuracy and complexity can be achieved by tuning ν and by choosing different preliminary approximating functions (e.g. neural networks with different numbers of nodes).

All of the approximated controllers always satisfy the input constraints and satisfy the state constraints after the first step. In order to evaluate, in relative terms, the closed loop performance and computational times of the considered control laws, Monte Carlo simulations have been performed starting from different initial conditions chosen inside $\mathcal{X}$. All the computations have been performed using MatLab® 7 with an AMD Athlon(tm) 64 3200+ with 1 GB RAM. The maximal and average computational times, over all time steps of all simulations, of the nominal controller were $0.7\,\mathrm{s}$ and $0.1\,\mathrm{s}$, while the maximal computational times $\overline{t}$ obtained with each of the approximated controllers are reported in Table II. As a measure of control system performance, at each time step the ratio $d_{\%} = 100\,\dfrac{d_t}{\|\phi_t\|_2}$ has been considered, where ϕ_t is the state trajectory obtained with the nominal controller and d_t is the Euclidean distance between ϕ_t and the trajectory obtained with any of the approximated controllers. When$\|\phi_t\|_2 \leq 0.05$, the value of d has been considered to measure the performance instead of $d_{\%}$. Then, the maximal values $\overline{d}_{\%}$ and $\overline{d}$ of $d_{\%}$ and d respectively over all time steps of all simulations have been computed. The obtained results are reported in Table II too. Note that the neural network approximations $\hat{\kappa}^{\mathrm{NN},1}$ and $\hat{\kappa}^{\mathrm{NN},2}$ achieve the lowest value of $\overline{t}$ (about 14000 times lower than on–line optimization), however, as expected, their performance in terms of $\overline{d}_{\%}$ and $\overline{d}$ are worse than those obtained with the related local SM approximating functions with any value of ν. By increasing the value of ν, better performance are obtained with the SM approach at the cost of higher computational times and memory usage. In practice, the tradeoff between accuracy, computational time and memory requirements has to be chosen on the basis of the particular control problem and on the available hardware. Finally, as regards the scalability of the approach, a state dimension of 7–8 may be still tractable with the proposed technique and a suitable choice of the set $\mathcal{X}$, depending also on the considered problem, on the available hardware and on the influence of each state variable on the control inputs. Similar limitations affect the other existing approximation techniques and probably represent the limit beyond which explicit/approximate NMPC does not give any advantage with respect to on–line optimization.

V. Conclusions

The theoretical results presented in this paper rely on SM theory and provide sufficient conditions for a generic approximated NMPC law to achieve arbitrary good accuracy. The proposed techniques allow to a) compute a bound on the worst–case approximation error of a given continuous approximated NMPC law and b) reduce such a worst–case error bound to an arbitrarily small value, using a local SM approximation approach. The results are based on the knowledge of a finite number ν of exact control moves, computed off–line and stored. Since the guaranteed error bound obtained with the local SM approach decreases with increasing ν, a tradeoff can be achieved between accuracy, on–line computational time and memory usage. A numerical example evidenced the effectiveness of the results.

Appendix A
Proofs

The following Lemma is instrumental for the proofs of Theorems 1–2

Lemma 1: Let $h : \mathcal{X} \to \mathbb{R}$ be an unknown function defined over a compact domain $\mathcal{X} \in \mathbb{R}^n$. Let the prior information available on h be described by: $h \in FFS_h = \{\tilde{h} \in \mathcal{A}_{L_h} : \tilde{h}(\tilde{x}) = \tilde{g},\ \forall \tilde{x} \in \mathcal{X}_\nu, \underline{g}(x) \leq h(x) \leq \overline{g}(x),\ \forall x \in \mathcal{X}\}$ where $\mathcal{A}_{L_h}$ is the set of Lipschitz continuous functions with Lipschitz constant L_h, $\mathcal{X}_\nu = \{\tilde{x}^k \in \mathcal{X} : h(\tilde{x}^k) = \tilde{g}^k, \tilde{g}^k \text{ is known}, k = \ldots, \nu\}$ and $\underline{g}, \overline{g} : \mathcal{X} \to \mathbb{R}$ are Lipschitz continuous functions with Lipschitz constant L_g. Define the functions:

$$\begin{aligned} \overline{h}(x) &\doteq \min[\overline{g}(x), \min_{\tilde{x}\in\mathcal{X}_\nu} (h(\tilde{x}) + L_h\|x-\tilde{x}\|_2)], \\ \underline{h}(x) &\doteq \max[\underline{g}(x), \max_{\tilde{x}\in\mathcal{X}_\nu} (h(\tilde{x}) - L_h\|x-\tilde{x}\|_2)] \end{aligned} \tag{17}$$

Then:

i) $\overline{h}(x) \geq \sup_{\tilde{h}\in FFS_h} \tilde{h}(x)\,,\ \underline{h}(x) \leq \inf_{\tilde{h}\in FFS_h} \tilde{h}(x)$

ii) if $L_g \leq L_h$, then the bounds $\overline{h}, \underline{h} \in FFS_h$ and they are tight: $\overline{h}(x) = \max_{\tilde{h}\in FFS_h} \tilde{h}(x)\,,\ \underline{h}(x) = \min_{\tilde{h}\in FFS_h} \tilde{h}(x)$

Proof. **i)** By contradiction. Suppose that a function $h^a \in FFS_h$ exists such that, for a certain $x^1 \in \mathcal{X}$,

$$h^a(x^1) > \min[\overline{g}(x^1), \min_{\tilde{x}\in\mathcal{X}_\nu} \big(h(\tilde{x}) + L_h\|x^1-\tilde{x}\|_2\big)] = \overline{h}(x^1) \tag{18}$$

Denote by $\tilde{x}^b$ a value of $\tilde{x} \in \mathcal{X}_\nu$ such that: $h(\tilde{x}^b) + L_h\|x^1-\tilde{x}^b\|_2 = \min_{\tilde{x}\in\mathcal{X}_\nu} \big(h(\tilde{x}) + L_h\|x^1-\tilde{x}\|_2\big)$ If $h(\tilde{x}^b) + L_h\|x^1-\tilde{x}^b\|_2 \geq \overline{g}(x^1)$, it means that $h^a(x^1) > \overline{g}(x^1) \Rightarrow h^a \notin FFS_h$ Otherwise, if $h(\tilde{x}^b) + L_h\|x^1-\tilde{x}^b\|_2 < \overline{g}(x^1)$, it means that $h^a(x^1) > h(\tilde{x}^b) + L_h\|x^1-\tilde{x}^b\|_2$. Since it was assumed that $h^a \in FFS_h$, $h^a(\tilde{x}^b) = h(\tilde{x}^b)$. Thus, $h^a(x^1) - h(\tilde{x}^b) = h^a(x^1) - h^a(\tilde{x}^b) > L_h\|x^1-\tilde{x}^b\|_2$. Moreover, since $h^a(x^1) > h(\tilde{x}^b) + L_h\|x^1-\tilde{x}^b\|_2$, then $h^a(x^1) > h(\tilde{x}^b) \Rightarrow h^a(x^1) - h(\tilde{x}^b) > 0$. Then, $h^a(x^1) - h^a(\tilde{x}^b) = |h^a(x^1) - h^a(\tilde{x}^b)| > L_h\|x^1-\tilde{x}^b\|_2 \Rightarrow h^a \notin FFS_h$. Therefore, there is no function $h^a \in FFS_h$ with the characteristics specified in (18), i.e. $\overline{h}(x) \geq h(x),\ \forall x \in \mathcal{X},\ \forall h \in FFS_h$. A similar proof holds for the lower bound $\underline{h}$. **ii)** It will be now shown that $\overline{h}$ belongs to FFS_h. Conditions $\overline{h}(x) \leq \overline{g}(x),\ \forall x \in \mathcal{X}$, and $\overline{h}(\tilde{x}) = \tilde{g},\ \forall \tilde{x} \in \mathcal{X}_\nu$, are satisfied by definition. Condition $\overline{h}(x) \geq \underline{g}(x)$ is also satisfied, since $L_g \leq L_h$ and $\overline{h}(x) = \min[\overline{g}(x), \min_{\tilde{x}\in\mathcal{X}_\nu} (h(\tilde{x}) + L_h\|x-\tilde{x}\|_2)] \geq \min[\overline{g}(x), \min_{\tilde{x}\in\mathcal{X}_\nu} (\underline{g}(\tilde{x}) + L_g\|x-\tilde{x}\|_2)] \geq \underline{g}(x),\ \forall x \in \mathcal{X}$. About the Lipschitz continuity of $\overline{h}$, for any $x^1 \in \mathcal{X}$ consider a value $\tilde{x}^b \in \mathcal{X}_\nu$ such that $h(\tilde{x}^b) + L_h\|x^1-\tilde{x}^b\|_2 = \min_{\tilde{x}\in\mathcal{X}_\nu} \big(h(\tilde{x}) + L_h\|x^1-\tilde{x}\|_2\big)$. If $h(\tilde{x}^b) + L_h\|x^1-\tilde{x}^b\|_2 \geq \underline{g}(x^1)$, it means that $\overline{h}(x^1) = \overline{g}(x^1)$, thus for any $x^2 \in \mathcal{X}$, since $\overline{h}(x^2) \leq \overline{g}(x^2)$, the following holds: $\overline{h}(x^2) - \overline{h}(x^1) \leq \overline{g}(x^2) - \overline{g}(x^1) \leq L_g\|x^2-x^1\|_2 \leq L_h\|x^2-x^1\|_2$. Otherwise, if $h(\tilde{x}^b) + L_h\|x^1-\tilde{x}^b\|_2 < \overline{g}(x^1)$, it means that $\overline{h}(x^1) = h(\tilde{x}^b) + L_h\|x^1-\tilde{x}^b\|_2$ and, for any $x^2 \in \mathcal{X}$, it can be noted that $\overline{h}(x^2) = \min[\overline{g}(x^2), \min_{\tilde{x}\in\mathcal{X}_\nu} (h(\tilde{x}) + L_h\|x^2-\tilde{x}\|_2)] \leq h(\tilde{x}^b) + L_h\|x^2-\tilde{x}^b\|_2 \leq h(\tilde{x}^b) + L_h\|x^2-x^1\|_2 + L_h\|x^1-\tilde{x}^b\|_2 = \overline{h}(x^1) + L_h\|x^2-x^1\|_2 \Rightarrow \overline{h}(x^2) - \overline{h}(x^1) \leq L_h\|x^2-x^1\|_2$ In a similar way, it can be shown that $\overline{h}(x^2) - \overline{h}(x^1) \geq -L_h\|x^2-x^1\|_2$ Therefore, since $\overline{h}(x^2) - \overline{h}(x^1) \leq L_h\|x^2-x^1\|_2$ and $\overline{h}(x^2) - \overline{h}(x^1) \geq -L_h\|x^2-x^1\|_2$, $|\overline{h}(x^2) - \overline{h}(x^1)| \leq L_h\|x^2-x^1\|_2,\ \forall x^1, x^2 \in \mathcal{X} \Rightarrow h \in \mathcal{A}_{L_h}$ Thus, if $L_g \leq L_h$ function $\overline{h}$ belongs to FFS_h and, considering result (i) of this Lemma, it is a tight upper bound for $\tilde{h}(x),\ \forall x \in \mathcal{X},\ \forall \tilde{h} \in FFS_h$. A similar proof holds for the

tight lower bound $\underline{h}$. ■

Proof of Theorem 1. **i)** Trivial application of Lipschitz continuity properties of κ_i and $\hat{\kappa}_i$.

ii)–vi) Note that the pointwise value of $\Delta_{\hat{\kappa},i}$ is bounded: $\forall x \in \mathcal{X},\ \underline{u}_i \leq \kappa_i(x) \leq \overline{u}_i \Rightarrow \underline{u}_i - \hat{\kappa}_i(x) \leq \kappa_i(x) - \hat{\kappa}_i(x) = \Delta_{\hat{\kappa},i}(x) \leq \overline{u}_i - \hat{\kappa}_i(x)$ and that the bounds $\underline{u}_i - \hat{\kappa}_i,\ \overline{u}_i - \hat{\kappa}_i : \mathcal{X} \to \mathbb{R}$ are Lipschitz continuous functions with Lipschitz constant $L_{\hat{\kappa},i}$. Thus, the prior information on $\Delta_{\hat{\kappa},i}$ is summarized by: $\Delta_{\hat{\kappa},i} \in \mathcal{D}_i$, where $\mathcal{D}_i = \{\Delta_i \in \mathcal{A}_{L_{\Delta_{\hat{\kappa},i}}} : \Delta_i(\tilde{x}) = \tilde{u}_i - \hat{\kappa}_i(\tilde{x}) = \Delta_{\hat{\kappa},i}(\tilde{x}),\ \forall \tilde{x} \in \mathcal{X}_\nu,\ \underline{u}_i - \hat{\kappa}_i(x) \leq \Delta_i(x) \leq \overline{u}_i - \hat{\kappa}_i(x),\ \forall x \in \mathcal{X}\}$. Thus, Lemma 1 can be used to compute the bounds $\overline{\Delta}_{\hat{\kappa},i},\ \underline{\Delta}_{\hat{\kappa},i}$ of $\mathcal{D}_i$, given by (11). On the basis of these bounds, it can be noted that $|\Delta_{\hat{\kappa},i}(x)| \leq \max(\overline{\Delta}_{\hat{\kappa},i}(x), -\underline{\Delta}_{\hat{\kappa},i}(x))$. Thus, $\forall x \in \mathcal{X},\ |\Delta_{\hat{\kappa},i}(x)| \leq \sup_{x\in\mathcal{X}} \max\left(\overline{\Delta}_{\hat{\kappa},i}(x), -\underline{\Delta}_{\hat{\kappa},i}(x)\right) = \zeta_i(\nu)$. Moreover, if $L_{\hat{\kappa},i} \leq L_{\Delta_{\hat{\kappa},i}}$, due to Lemma 1 the bound $\zeta_i(\nu)$ (10) is the tightest on the basis of the available prior information on κ_i, since it is computed on the basis of functions $\overline{\Delta}_{\hat{\kappa},i},\ \underline{\Delta}_{\hat{\kappa},i}$ which tightly bound the set $\mathcal{D}_i$. Furthermore, if property (12) holds it can be noted that $\Delta_{\hat{\kappa},i}(\tilde{x}) = \kappa_i(\tilde{x}) - \hat{\kappa}_i(\tilde{x}) = \tilde{u}_i - \tilde{u}_i = 0,\ \forall \tilde{x} \in \mathcal{X}_\nu$. Then, by substituting $\Delta_{\hat{\kappa},i}(\tilde{x}) = 0$ and $\chi_i(x) = \min_{\tilde{x}\in\mathcal{X}_\nu}\left(L_{\Delta_{\hat{\kappa},i}}\|x - \tilde{x}\|_2\right)]$ in the computation of ζ_i given in (10), by straightforward manipulations it can be shown that $\zeta_i = \sup_{x\in\mathcal{X}} \min[\max(\overline{u}_i - \hat{\kappa}_i(x), \hat{\kappa}_i(x) - \underline{u}_i), \chi_i(x)]$. Additionally, note that $\chi_i(x) = \min_{\tilde{x}\in\mathcal{X}_\nu}\left(L_{\Delta_{\hat{\kappa},i}}\|x - \tilde{x}\|_2\right) = L_{\Delta_{\hat{\kappa},i}} \min_{\tilde{x}\in\mathcal{X}_\nu}(\|x - \tilde{x}\|_2) \leq L_{\Delta_{\hat{\kappa},i}}\, d_H(\mathcal{X}, \mathcal{X}_\nu)$ and that, due to its formulation, $\chi_i(x) \geq 0$. Then, due to property (4) it is obtained that $0 \leq \lim_{\nu\to\infty} \chi_i(x) \leq \lim_{\nu\to\infty} L_{\Delta_{\hat{\kappa},i}}\, d_H(\mathcal{X}, \mathcal{X}_\nu) = 0 \Rightarrow \lim_{\nu\to\infty} \chi_i(x) = 0$. Note also that $\overline{u}_i - \hat{\kappa}_i(x) \geq 0$ and $\hat{\kappa}_i(x) - \underline{u}_i \geq 0$, because $\hat{\kappa}_i$ satisfies the input saturation constraints by assumption. Thus, the value of ζ_i (13) is such that $\zeta_i = \sup_{x\in\mathcal{X}} \min\left[\max\left(\overline{u}_i - \hat{\kappa}_i(x), -\underline{u}_i + \hat{\kappa}_i(x)\right), \chi_i(x)\right] \geq 0$ and it can be noted that $0 \leq \lim_{\nu\to\infty} \zeta_i = \lim_{\nu\to\infty} \sup_{x\in\mathcal{X}} \min(\max(\overline{u}_i - \hat{\kappa}_i(x), -\underline{u}_i + \hat{\kappa}_i(x)), \chi_i(x)) = \sup_{x\in\mathcal{X}} \min(\max(\overline{u}_i - \hat{\kappa}_i(x), -\underline{u}_i + \hat{\kappa}_i(x)), \lim_{\nu\to\infty} \chi_i(x)) = 0 \Rightarrow \lim_{\nu\to\infty} \zeta_i = 0$ ■

Proof of Theorem 2. **i–iii)** Property (5) and Lipschitz continuity of κ_i^{LOC} are trivially satsified by construction. For any $\tilde{x}^h \in \mathcal{X}_\nu$, note that $\min_{\tilde{x}\in\mathcal{X}}(\Delta_{\hat{\kappa},i}(\tilde{x}) + L_{\Delta_{\hat{\kappa},i}}\|\tilde{x}^h - \tilde{x}\|_2) = \Delta_{\hat{\kappa},i}(\tilde{x}^h)$ and $\max_{\tilde{x}\in\mathcal{X}}(\Delta_{\hat{\kappa},i}(\tilde{x}) - L_{\Delta_{\hat{\kappa},i}}\|\tilde{x}^h - \tilde{x}\|_2) = \Delta_{\hat{\kappa},i}(\tilde{x}^h)$. Moreover, since by assumption (5) $\hat{\kappa}_i$ satisfies the input constraints, it can be noted that $\Delta_{\hat{\kappa},i}(\tilde{x}^h) = \kappa_i(\tilde{x}^h) - \hat{\kappa}_i(\tilde{x}^h) \leq \overline{u}_i - \hat{\kappa}_i(\tilde{x}^h)$ and $\Delta_{\hat{\kappa},i}(\tilde{x}^h) = \kappa_i(\tilde{x}^h) - \hat{\kappa}_i(\tilde{x}^h) \geq \underline{u}_i - \hat{\kappa}_i(\tilde{x}^h)$. Thus, the following result is obtained: $\overline{\Delta}_{\hat{\kappa},i}(\tilde{x}^h) = \underline{\Delta}_{\hat{\kappa},i}(\tilde{x}^h) = \Delta_{\hat{\kappa},i}(\tilde{x}^h)$. As a consequence, $\Delta_{\hat{\kappa},i}^{\text{OPT}}(\tilde{x}^h) = \frac{1}{2}(\overline{\Delta}_{\hat{\kappa},i}(\tilde{x}^h) + \underline{\Delta}_{\hat{\kappa},i}(\tilde{x}^h)) = \Delta_{\hat{\kappa},i}(\tilde{x}^h),\ \forall \tilde{x}^h \in \mathcal{X}_\nu$. Therefore, it can be noted that: $\kappa_i^{\text{LOC}}(\tilde{x}) = \hat{\kappa}_i(\tilde{x}) + \Delta_{\hat{\kappa},i}^{\text{OPT}}(\tilde{x}) = \hat{\kappa}_i(\tilde{x}) + \Delta_{\hat{\kappa},i}(\tilde{x}) = \kappa_i(\tilde{x}) - \hat{\kappa}_i(\tilde{x}) + \hat{\kappa}_i(\tilde{x}) = \kappa_i(\tilde{x}) = \tilde{u},\ \forall \tilde{x} \in \mathcal{X}_\nu$. As it has been shown in the proof of Theorem 1, the prior information on the approximation error $\Delta_{\hat{\kappa},i}$ is summarized by $\Delta_{\hat{\kappa},i} \in \mathcal{D}_i$. Thus, according to Lemma 1, $\underline{\Delta}_{\hat{\kappa},i}(x) \leq \Delta_{\hat{\kappa},i}(x) \leq \overline{\Delta}_{\hat{\kappa},i}(x),\ \forall x \in \mathcal{X}$. Therefore, by straightforward manipulations it can be noted that $|\kappa_i(x) - \kappa_i^{\text{LOC}}(x)| \leq \frac{1}{2}(\overline{\Delta}_{\hat{\kappa},i}(x) - \underline{\Delta}_{\hat{\kappa},i}(x)),\ \forall x \in \mathcal{X} \Rightarrow |\kappa_i(x) - \kappa_i^{\text{LOC}}(x)| \leq \sup_{x\in\mathcal{X}} \frac{1}{2}(\overline{\Delta}_{\hat{\kappa},i}(x) - \underline{\Delta}_{\hat{\kappa},i}(x)) = \zeta_i^{\text{LOC}},\ \forall x \in \mathcal{X}$. Moreover, due to Theorem 1, the approximation error $\Delta_{\hat{\kappa},i}$ is bounded by (10) $\zeta_i = \sup_{x\in\mathcal{X}} \max(\overline{\Delta}_{\hat{\kappa},i}(x), -\underline{\Delta}_{\hat{\kappa},i}(x))$ and it can be noted that $\frac{1}{2}(\overline{\Delta}_{\hat{\kappa},i}(x) - \underline{\Delta}_{\hat{\kappa},i}(x)) \leq \max(\overline{\Delta}_{\hat{\kappa},i}(x), -\underline{\Delta}_{\hat{\kappa},i}(x)),\ \forall x \in \mathcal{X}$, thus $\zeta_i^{\text{LOC}} = \sup_{x\in\mathcal{X}} \frac{1}{2}(\overline{\Delta}_{\hat{\kappa},i}(x) - \underline{\Delta}_{\hat{\kappa},i}(x)) \leq \sup_{x\in\mathcal{X}} \max(\overline{\Delta}_{\hat{\kappa},i}(x), -\underline{\Delta}_{\hat{\kappa},i}(x)) = \zeta_i$.

iv)–v) The considered prior information on κ_i is given by (14). For any generic function $\tilde{\kappa}_i$, consider the corresponding error function $\tilde{\Delta}_i = \tilde{\kappa}_i - \hat{\kappa}_i$. From (14) it can be noted that: $\tilde{\kappa}_i \in FFS_{\Delta,i} \Rightarrow \tilde{\Delta}_i \in \mathcal{A}_{L_{\Delta_{\hat{\kappa},i}}}$; $\tilde{\kappa}_i \in FFS_{\Delta,i} \Rightarrow \tilde{\Delta}_i(\tilde{x}) = \tilde{u}_i - \hat{\kappa}_i(\tilde{x}),\ \forall \tilde{x}_i \in \mathcal{X}_\nu$ and $\tilde{\kappa}_i \in FFS_{\Delta,i} \Rightarrow \underline{u}_i - \hat{\kappa}_i(x) \leq \tilde{\Delta}_i(x) \leq \overline{u}_i - \hat{\kappa}_i(x),\ \forall x \in \mathcal{X}$. Thus, the following necessary condition is obtained: $\tilde{\kappa}_i \in FFS_{\Delta,i} \Rightarrow \tilde{\Delta}_i \in \mathcal{D}_i$. On the other hand, if $\tilde{\Delta}_i \in \mathcal{D}_i$ then $\underline{u}_i - \hat{\kappa}_i(x) \leq \tilde{\Delta}_i(x) \leq \overline{u}_i - \hat{\kappa}_i(x) \Rightarrow \underline{u}_i \leq \tilde{\Delta}_i(x) + \hat{\kappa}_i(x) \leq \overline{u}_i \Rightarrow \underline{u}_i \leq \tilde{\kappa}_i(x) \leq \overline{u}_i$. Moreover, $\tilde{\Delta}_i \in \mathcal{D}_i \Rightarrow \tilde{\kappa}_i(\tilde{x}) = \hat{\kappa}_i(\tilde{x}) + \tilde{\Delta}_i(\tilde{x}) = \hat{\kappa}_i(\tilde{x}) + \tilde{u}_i - \hat{\kappa}_i(\tilde{x}) = \tilde{u}_i,\ \forall \tilde{x} \in \mathcal{X}_\nu$ and, since $\Delta_{\hat{\kappa},i} \in \mathcal{D}_i$, $\tilde{\kappa}_i - \hat{\kappa}_i = \tilde{\Delta}_i \in \mathcal{A}_{L_{\Delta_{\hat{\kappa},i}}}$. Thus the following sufficient condition is also obtained: $\tilde{\kappa}_i \in FFS_{\Delta,i} \Leftarrow \tilde{\Delta}_i \in \mathcal{D}_i$ Therefore, $\tilde{\kappa}_i \in FFS_{\Delta,i} \iff \tilde{\Delta}_i \in \mathcal{D}_i$.Moreover, note that $e(\kappa_i, \tilde{\kappa}_i) = \|\kappa_i - \hat{\kappa}_i - \tilde{\Delta}_i\|_\infty = \|\Delta_{\hat{\kappa},i} - \tilde{\Delta}_i\|_\infty = e(\Delta_{\hat{\kappa},i}, \tilde{\Delta}_i) \Rightarrow E(\kappa_i, \tilde{\kappa}_i) = \sup_{\kappa_i\in FFS_{\Delta,i}} e(\kappa_i, \tilde{\kappa}_i) = \sup_{\Delta_{\hat{\kappa},i}\in\mathcal{D}_i} e(\Delta_{\hat{\kappa},i}, \tilde{\Delta}_i) = E(\Delta_{\hat{\kappa},i}, \tilde{\Delta}_i)$. Therefore, since $\tilde{\kappa}_i \in FFS_{\Delta,i} \iff \tilde{\Delta}_i \in \mathcal{D}_i$ and $E(\kappa_i, \tilde{\kappa}_i) = E(\Delta_{\hat{\kappa},i}, \tilde{\Delta}_i)$, finding an optimal approximation $\kappa_i^{\text{LOC}} = \hat{\kappa}_i + \Delta_{\hat{\kappa},i}^{\text{OPT}} \approx \kappa_i$ such that $\kappa_i^{\text{LOC}} \in FFS_{\Delta,i}$, considering the information $\kappa_i \in FFS_{\Delta,i}$, is equivalent to finding an optimal approximation $\Delta_{\hat{\kappa},i}^{\text{OPT}} \approx \Delta_{\hat{\kappa},i}$ such that $\Delta_{\hat{\kappa},i}^{\text{OPT}} \in \mathcal{D}_i$, considering the information $\Delta_{\hat{\kappa},i} \in \mathcal{D}_i$, i.e. $E(\kappa_i, \kappa_i^{\text{LOC}}) = \inf_{\tilde{\kappa}_i\in FFS_{\Delta,i}} E(\kappa_i, \tilde{\kappa}_i) = \inf_{\tilde{\Delta}_i\in\mathcal{D}_i} E(\Delta_{\hat{\kappa},i}, \tilde{\Delta}_i) = E(\Delta_{\hat{\kappa},i}, \Delta_{\hat{\kappa},i}^{\text{OPT}}) = r_{\Delta,\infty,i}$. Thus, the aim is to show that $\Delta_{\hat{\kappa},i}^{\text{OPT}} = 1/2\,(\overline{\Delta}_{\hat{\kappa},i} + \underline{\Delta}_{\hat{\kappa},i})$ (15) belongs to $\mathcal{D}_i$ and is an optimal approximation of $\Delta_{\hat{\kappa},i}$. This result can be straightforwardly derived from the proof of Theorem 4 of [7]. ■

REFERENCES

[1] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, pp. 789–814, 2000.

[2] T. Parisini and R. Zoppoli, "A receding-horizon regulator for nonlinear systems and a neural approximation," *Automatica*, vol. 31, no. 10, pp. 1443–1451, 1995.

[3] T. Johansen, "Approximate explicit receding horizon control of constrained nonlinear systems," *Automatica*, vol. 40, pp. 293–300, 2004.

[4] T. Johansen and A. Grancharova, "Approximate explicit constrained linear model predictive control via orthogonal search tree," *IEEE Transactions on Automatic Control*, vol. 48, no. 5, pp. 810–815, May 2003.

[5] D. Muñoz de la Peña, A. Bemporad, and C. Filippi, "Robust explicit MPC based on approximate multiparametric convex programming," *IEEE Transactions on Automatic Control*, vol. 51, no. 8, pp. 1399–1403, August 2006.

[6] A. Grancharova and T. A. Johansen, "Computation, approximation and stability of explicit min–max nonlinear model predictive control," *Automatica*, vol. 45, no. 5, pp. 1134–1143, 2009.

[7] M. Canale, L. Fagiano, and M. Milanese, "Set membership approximation theory for fast implementation of model predictive control laws," *Automatica*, vol. 45, no. 1, pp. 45–54, 2009.

[8] M. Vidyasagar, *Nonlinear systems analysis*. SIAM, 2002.

[9] J. Spjøtvold, P. Tøndel, and T. A. Johansen, "Continuous selection and unique polyhedral representation of solutions to convex parametric quadratic programs," *Journal of Optimization Theory and Applications*, vol. 134, pp. 177–189, 2007.

[10] M. Baes, M. Diehl, and I. Necoara, "Every continuous nonlinear control system can be obtained by parametric convex programming," *IEEE Transactions on Automatic Control*, vol. 53, no. 8, pp. 1963–1967, 2008.

[11] G. Grimm, M. J. Messina, S. E. Tuna, and A. R. Teel, "Examples when nonlinear modelo predictive control is nonrobust," *Automatica*, vol. 40, pp. 1729–1738, 2004.

[12] M. Milanese and C. Novara, "Computation of local radius of information in SM-IBC-identification of nonlinear systems," *Journal of Complexity*, vol. 23, no. 6, pp. 937–951, 2007.

[13] J. Traub and H. Woźniakowski, *A General Theory of Optimal Algorithms*. New York: Academic Press, 1980.

[14] G. Beliakov, "Interpolation of Lipschitz functions," *Journal of Computational and Applied Mathematics*, vol. 196, pp. 20–44, 2006.

[15] D. W. Jordan and P. Smith, *Nonlinear ordinary differential equations*. Oxford University Press, 1987.