Technical report 24-011

# Sensitivity analysis for piecewise-affine approximations of nonlinear programs with polytopic constraints*

L. Gharavi, C. Liu, B. De Schutter, and S. Baldi

---

# Sensitivity Analysis for Piecewise-Affine Approximations of Nonlinear Programs with Polytopic Constraints

Leila Gharavi, Changrui Liu, Bart De Schutter *Fellow IEEE*, Simone Baldi *Senior Member IEEE*

*Abstract*—**Nonlinear Programs (NLPs) are prevalent in optimization-based control of nonlinear systems. Solving general NLPs is computationally expensive, necessitating the development of fast hardware or tractable suboptimal approximations. This paper investigates the sensitivity of the solutions of NLPs with polytopic constraints when the nonlinear continuous objective function is approximated by a PieceWise-Affine (PWA) counterpart. By leveraging perturbation analysis using a convex modulus, we derive guaranteed bounds on the distance between the optimal solution of the original polytopically-constrained NLP and that of its approximated formulation. Our approach aids in determining criteria for achieving desired solution bounds. Two case studies on the Eggholder function and nonlinear model predictive control of an inverted pendulum demonstrate the theoretical results.**

*Index Terms*—**Perturbation Analysis, Non-Convex Nonlinear Programming, Piecewise-Affine Functions, Max-Min-Plus-Scaling Systems, Function Approximation**

## I. INTRODUCTION

**N**ONLINEAR Programs (NLPs) are commonly encountered in optimization-based control of nonlinear systems, e.g., Nonlinear Model Predictive Control (NMPC) [1]. Solving non-convex NLPs is intractable, posing a great challenge in applying optimization-based controllers in real-time operations, especially for systems having fast dynamics. Various solutions have been proposed to address this issue, such as adaptive problem formulations [2], learning-based methods [3], and sensitivity analysis of NLPs [4].

PWA approximations are widely used due to their tractability [5], [6]. To obtain a continuous PWA approximation, min and max operators can be used to maintain continuity and to resolve numerical issues in the resulting optimization problem [7]–[9]. The approximated problem can be used to obtain a suboptimal solution [10], whose optimality highly depends on the accuracy of the approximation. For example, a warm start of a non-convex NLP can be obtained by solving

the approximated optimization problem [11], [12]. Optimality guarantees of such approaches can be derived using sensitivity analysis, establishing an upper bound on the distance between the original solutions and the approximated ones. As a result, by finding a subset of the decision space around the approximated solution, one can sample a structured or random warm start to solve the original non-convex NLP more efficiently.

Quantitative bounds on the distance between the original and the approximated solutions have been studied in the sensitivity analysis of quadratic [13] and convex [14] optimization problems. Regarding NLPs, there exist several results on their sensitivity to the parameters in the optimization formulation [15], [16] and the initial solution [17]. In addition, optimality and dissipativity conditions for the perturbed convex NLP problem have also been established [18], [19]. For a more extensive study, the reader can refer to [20]. Recently, sensitivity analysis has also been applied to analyze the infinite-horizon performance of MPC [21]. However, obtaining quantitative bounds on the distance between the solutions of a non-convex NLP and its PWA approximation is still a gap that needs to be filled, and our work addresses this problem.

In this paper, we present a method to bound the solutions of a polytopically-constrained non-convex NLP problem using a continuous PWA approximation of the nonlinear objective function. We employ the Max-Min-Plus-Scaling (MMPS) formulation of continuous PWA functions in [8] to construct a piecewise convex approximation formalism. Leveraging some results in [14], we derive guaranteed bounds on the distance between the original and the approximated optimal solution. Moreover, our approach can not only establish such bounds but also aid in determining necessary criteria during the approximation stage to attain a desired solution bound. To demonstrate the theoretical findings, we present two case studies on the Eggholder function [22], a renowned benchmark optimization problem with multiple local minima, and an NMPC optimization problem for an inverted pendulum [1].

The rest of the paper is organized as follows. In Section II, we present preliminaries regarding the sensitivity analysis of NLPs and the PWA approximation of nonlinear functions. Section III formulates the problem and Section IV elaborates our proposed approach to theoretically compute the confidence radius for the local minima of the corresponding approximated function. In Section V, we then demonstrate the derived confidence radius through a case study on the Eggholder function and we apply our analysis to an NMPC optimization

example. Section VI concludes this paper.

*Notation:* For a positive integer $P$, we use $\mathcal{I}_P$ to denote the set $\{1, 2, \ldots, P\}$. For a connected set $\mathcal{D} \subseteq \mathbb{R}^n$, the diameter of $\mathcal{D}$ is defined as $\mathrm{diam}(\mathcal{D}) := \max_{x_1, x_2 \in \mathcal{D}} \|x_1 - x_2\|$, where $\|\cdot\|$ is the Euclidean or the 2-norm.

## II. PRELIMINARIES

### A. Representation of Continuous PWA Functions

We start with formally defining a continuous PWA function which will be frequently used throughout the paper.

**Definition 1** (Continuous PWA function [23])**.** *A scalar-valued function $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ is said to be a continuous PWA function if and only if the following conditions hold:*

1) *the domain space $\mathcal{D}$ is divided into a finite number of closed polyhedral regions $\mathcal{R}_1, \ldots, \mathcal{R}_R$ with non-overlapping interiors,*
2) *for each $r \in \mathcal{I}_R$, $f$ can be expressed as*

$$f(x) = \alpha_r^T x + \beta_r \quad \text{if} \quad \mathrm{x} \in \mathcal{R}_\mathrm{r},$$

   *with $\alpha_r \in \mathbb{R}^n$ and $\beta_r \in \mathbb{R}$, and*
3) *$f$ is continuous on the boundary between any pair of regions.*

PWA functions can be expressed in different forms, among which MMPS form is powerful for decomposing PWA functions.

**Theorem 1** (MMPS representation [8])**.** *For a scalar-valued continuous PWA function $f$ as in Definition 1, there exist non-empty index sets $\mathcal{I}_P$ and $\mathcal{I}_{Q_p}$ such that*

$$f(x) = \min_{p \in \mathcal{I}_P} \max_{q \in \mathcal{I}_{Q_p}} \left( a_{p,q}^T x + b_{p,q} \right), \quad (1)$$

*for real numbers $b_{p,q}$ and vectors $a_{p,q} \in \mathbb{R}^n$.*

For convenience, we define the $p$-th local convex segment of $f$ as

$$f_p(x) := \max_{q \in \mathcal{I}_{Q_p}} \left( a_{p,q}^T x + b_{p,q} \right), \quad (2)$$

where $f_p$ is convex since it is defined as the maximum of a finite number of affine functions and its domain is also convex. In addition, we define the region $\mathcal{C}_{p,q}$ in which a certain affine function is activated and the region $\mathcal{C}_{p,.}$ in which a convex PWA function is activated, that is,

$$\mathcal{C}_{p,q} := \{x \in \mathcal{D} \mid f(x) = a_{p,q}^T x + b_{p,q}\}, \quad (3a)$$
$$\mathcal{C}_{p,.} := \{x \in \mathcal{D} \mid f(x) = f_p(x)\}. \quad (3b)$$

Further, we have $\mathcal{C}_{p,.} = \bigcup_{q=1}^{Q_p} \mathcal{C}_{p,q}$. Lastly, we define $\sigma_p : \mathcal{C}_{p,.} \to \mathcal{I}_{Q_p}$ as the region index function for $f_p$ as

$$\sigma_p(x) = q \iff x \in \mathcal{C}_{p,q}. \quad (4)$$

### B. Sensitivity of the Convex Optimization Problem

The convexity modulus [14], being used to quantify convexity, is useful in the sensitivity analysis of convex functions. In the following, we define the convexity modulus specifically for $f_p$, the $p$-th convex segment of $f$, and its domain $\mathcal{C}_{p,.}$.

**Definition 2** (Convexity modulus [14])**.** *For a given convex function $f_p$, the convexity modulus $h_1 : [0, +\infty) \to [0, +\infty)$ over the domain $\mathcal{C}_{p,.}$ is defined as*

$$h_1(\gamma) := \begin{cases} \inf\limits_{\substack{v, w \in \mathcal{C}_{p,.} \\ \|v - w\| = \gamma}} J(v, w) & \text{if } \gamma < \mathrm{diam}(\mathcal{C}_{p,.}) \\ +\infty & \text{if } \gamma \geqslant \mathrm{diam}(\mathcal{C}_{p,.}) \end{cases}, \quad (5)$$

*where $v$ and $w$ are two points in $\mathcal{C}_{p,.}$ satisfying $\|v - w\| = \gamma$, and $J(v, w)$ is given as*

$$J(v, w) = \frac{f_p(v) + f_p(w)}{2} - f_p\left(\frac{v + w}{2}\right). \quad (6)$$

**Theorem 2** (Theorem 4.5 in [14])**.** *Suppose that $f_p : \mathcal{C}_{p,.} \to \mathbb{R}$ is a scalar-valued convex function and $\delta_p : \mathcal{C}_{p,.} \to \mathbb{R}$ is an arbitrary function satisfying*

$$\sup_{x \in \mathcal{C}_{p,.}} |\delta_p(x)| = \Delta_p < \infty. \quad (7)$$

*Let $x_p^*$ be any global infimizer of $f_p$ and $\hat{x}_p^*$ be any global infimizer of $\hat{f}_p = f_p + \delta_p$. Then*

$$\|\hat{x}_p^* - x_p^*\| \leqslant h_1^{-1}(2\Delta_p), \quad (8)$$

*where $h_1$ is the convexity modulus in Definition 2.*

For a more compact notation, we call the right-hand side of (8), the confidence radius, defined as follows:

**Definition 3** (Confidence radius)**.** *For a given function $\hat{f}_p : \mathcal{C}_{p,.} \subseteq \mathbb{R}^n \to \mathbb{R}$ approximating the function $f_p : \mathcal{C}_{p,.} \subseteq \mathbb{R}^n \to \mathbb{R}$ with the maximal approximation error $\Delta_p$ from (7), the confidence radius is the upper-bound on the distance between $\hat{x}_p^*$, the global minimizer of $\hat{f}_p$, and $x_p^*$, the global minimizer of $f_p$, and is obtained by*

$$\chi = h_1^{-1}(2\Delta_p),$$

*where $h_1$ is the convexity modulus in Definition 2.*

**Proposition 1** (Proposition 2.2, 2.5 in [14])**.** *Given a convex function $f_p$ on a compact domain $\mathcal{C}_{p,.}$ with convexity modulus $h_1$ defined as (5), the following hold:*

1) *$h_1$ is left-continuous on $(0, \mathrm{diam}(\mathcal{C}_{p,.}))$, and*
2) *for $\gamma_1, \gamma_2 \in [0, \mathrm{diam}(\mathcal{C}_{p,.}))$, if $\gamma_1 < \gamma_2$, then $h_1(\gamma_1)/\gamma_1 \leqslant h_1(\gamma_1)/\gamma_2$.*

### III. CONTINUOUS PWA APPROXIMATION OF NLPS

Consider an NLP with polytopic constraints

$$\min_{x \in \mathcal{X}} F(x), \quad (9)$$

where $F : \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}$ is the nonlinear objective function and $\mathcal{X} \subseteq \mathcal{D}$ is the polytopic feasible region. From now on, we assume that the domain $\mathcal{D}$ is compact. We approximate $F$

by a continuous PWA function $f$ of the MMPS form (1) via solving the approximation problem

$$\min_{\mathcal{A},\,\mathcal{B}} \int_{\mathcal{D}} \left| F(x) - \min_{p\in\mathcal{I}_P}\max_{q\in\mathcal{I}_{Q_p}} \left(a_{p,q}^T x + b_{p,q}\right)\right| dx, \quad (10)$$

to minimize the absolute approximation error, where the ordered sets $\mathcal{A}$ and $\mathcal{B}$ respectively collect $a_{p,q}$ and $b_{p,q}$.

**Example 1.** *Figure 1 shows a 1-dimensional example of approximating a nonlinear objective function $F$ by a continuous PWA function $f$ using the MMPS form (1) as*

$$F(x) \approx f(x) = \min\left(\underbrace{\max(f_{1,1}, f_{1,2})}_{f_{1,\cdot}}, \underbrace{\max(f_{2,1}, f_{2,2})}_{f_{2,\cdot}}, f_{3,\cdot}\right),$$

*with $P = 3$, $Q_1 = Q_2 = 2$, and $Q_3 = 1$. The convex segments of $f$ are shown by $f_{p,\cdot}$ which give the maximum value among $Q_p$ affine functions $f_{p,q}$, $q \in \mathcal{I}_{Q_p}$. The subregions $\mathcal{C}_{p,q}$ are shown in the same color as their corresponding active affine functions, $f_{p,q}$. In this 1-dimensional example, $\mathrm{diam}(\mathcal{C}_{p,\cdot})$ is the distance between the upper and lower bounds of $\mathcal{C}_{p,\cdot}$ on the $x$-axis.*
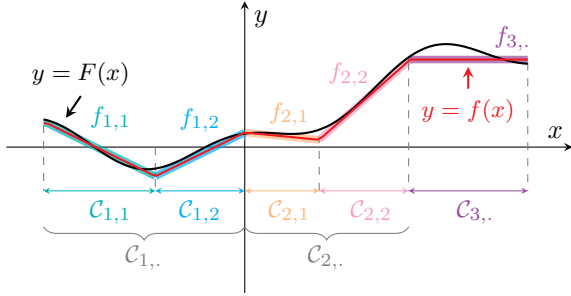


Fig. 1: A conceptual example of approximating a nonlinear function $F$ with a continuous PWA approximation $f$ using the MMPS form in (1).

## IV. Theoretical Analysis

The main result of this section is the sensitivity bound in Theorem 3. Before, we present some building blocks about continuity and boundedness of the convexity modulus (cf. Lemma 1, 2 and 3).

**Lemma 1.** *For a convex PWA function $f_p$ expressed by (2), $\partial h_1/\partial\gamma$ is a piecewise constant function.*

*Proof.* For compactness, let us denote

$$i = \sigma_p(v), \qquad j = \sigma_p(w), \qquad k = \sigma_p\left(\frac{v+w}{2}\right),$$

with $i, j, k \in \mathcal{I}_{Q_p}$. For $f_p$ expressed by (2), the function $J$ in (6) can be written as

$$J(v,w) = \frac{a_{p,i}^T - a_{p,k}^T}{2}v + \frac{a_{p,j}^T - a_{p,k}^T}{2}w + \frac{b_{p,i} + b_{p,j} - 2b_{p,k}}{2}. \quad (11)$$

The necessary Lagrange conditions for optimality at $(v^*, w^*)$ in (5) state that there must exist $\mu \in \mathbb{R}$ that satisfies the following simultaneously:

$$\nabla J(v^*, w^*) + \mu\nabla\left(\|v - w\| - \gamma\right)\Big|_{\substack{v=v^* \\ w=w^*}} = 0, \quad (12a)$$

$$\|v^* - w^*\| = \gamma. \quad (12b)$$

By calculating the gradient of (11), we have that (12) becomes

$$\frac{a_{p,i}^T - a_{p,k}^T}{2} + \mu\frac{v^* - w^*}{\|v^* - w^*\|} = 0, \quad (13a)$$

$$\frac{a_{p,j}^T - a_{p,k}^T}{2} + \mu\frac{w^* - v^*}{\|v^* - w^*\|} = 0, \quad (13b)$$

$$\|v^* - w^*\| = \gamma, \quad (13c)$$

which implies the existence of $\mu \in \mathbb{R}$ satisfying

$$\frac{a_{p,i}^T - a_{p,j}^T}{2} + 2\mu\frac{v^* - w^*}{\gamma} = 0. \quad (14)$$

Note that $\partial h_1/\partial\gamma$ is equal to $\partial J/\partial\gamma$, except where the indices $i$, $j$, and $k$ change. At such points, $h_1$ is not differentiable with respect to $\gamma$, which does not conflict with $\partial h_1/\partial\gamma$ being a piecewise constant function. To find the slope of $h_1$ where it exists, the chain rule can be applied as

$$\frac{\partial h_1}{\partial\gamma} = \frac{\partial J}{\partial v^*}\Big/\frac{\partial\gamma}{\partial v^*} + \frac{\partial J}{\partial w^*}\Big/\frac{\partial\gamma}{\partial w^*} = \frac{a_{p,i}^T - a_{p,j}^T}{2}\frac{\|v^* - w^*\|}{v^* - w^*},$$

which, considering (14), leads to

$$\frac{\partial h_1}{\partial\gamma} = -2\mu, \quad (15)$$

which implies that $\partial h_1/\partial\gamma$ is a function of the $a_{p,i}^T - a_{p,j}^T$ values. □

**Lemma 2.** *For a convex PWA function $f_p$ expressed by (2), the convexity modulus $h_1$ in (5) is continuous on $[0, \mathrm{diam}(\mathcal{C}_{p,\cdot}))$.*

*Proof.* From Proposition 1, we know that $h_1$ is left-continuous on $[0, +\infty)$. Seeking a contradiction, let us assume that $h_1$ is not right-continuous in $\gamma_0 \in [0, \mathrm{diam}(\mathcal{C}_{p,\cdot}))$, hence,

$$\lim_{\gamma\to\gamma_0^+} h_1(\gamma) \neq h_1(\gamma_0).$$

The monotonicity property of $h_1$ in Proposition 1 requires

$$\lim_{\gamma\to\gamma_0^+} h_1(\gamma) > h_1(\gamma_0).$$

Therefore, without loss of generality, we assume there exists a gap $\epsilon_0 > 0$ and a point $\gamma_0 < \gamma_0^+ < \mathrm{diam}(\mathcal{C}_{p,\cdot})$ such that

$$h_1(\gamma_0^+) = h_1(\gamma_0) + \frac{\partial h_1}{\partial\gamma}\Big|_{\gamma=\gamma_0}\left(\gamma_0^+ - \gamma_0\right) + \epsilon_0. \quad (16)$$

Using (5), we define the points $v_0$, $w_0$, and $w_0^+$ such that

$$h_1(\gamma_0) = \inf_{\substack{v,w\in\mathcal{C}_{p,\cdot} \\ \|v-w\|=\gamma_0}} J(v,w) = J(v_0, w_0),$$

and $\|v_0 - w_0^+\| = \gamma_0^+$. Considering the optimality property in (5), we have

$$h_1(\gamma_0^+) \leqslant J(v_0, w_0^+),$$

and knowing that $J$ is a continuous function by definition, we can deduce

$$h_1(\gamma_0^+) \leqslant h_1(\gamma_0) + \rho(w_0^+ - w_0),$$

where $\rho : \mathcal{C}_{p,.} \to \mathbb{R}$ is a function with the following property:

$$\lim_{\nu \to 0^+} \rho(\nu) = 0. \qquad (17)$$

Substituting (16) into (17) and taking the limit on both sides when $\gamma_0^+$ approaches $\gamma_0$ leads to

$$h_1(\gamma_0) + \epsilon_0 \leqslant h_1(\gamma_0),$$

which contradicts the fact that $\epsilon_0 > 0$. Therefore, $h_1$ is right-continuous on $[0, \operatorname{diam}(\mathcal{C}_{p,.}))$. $\qquad\square$

**Proposition 2.** *For a convex PWA function $f_p$ expressed by (2), we have*

$$h_1(\gamma) = 0, \qquad \forall \gamma \leqslant \max_{q \in \mathcal{I}_{Q_p}} \{\operatorname{diam}(\mathcal{C}_{p,q})\}.$$

*Proof.* By seeking a contradiction, let us assume that

$$\exists \gamma_0 \leqslant \max_{q \in \mathcal{I}_{Q_p}} \{\operatorname{diam}(\mathcal{C}_{p,q})\}, \quad \text{such that} \quad h_1(\gamma_0) > 0,$$

with the corresponding optimal points $v_0^*$ and $w_0^*$ from (5) such that

$$h_1(\gamma_0) = \inf_{\substack{v,w \in \mathcal{C}_{p,.} \\ \|v-w\| = \gamma_0}} J(v, w) = J(v_0^*, w_0^*).$$

Let us select two points, $v_0$ and $w_0$, on the largest subregion in $\mathcal{C}_{p,.}$ such that

$$\|v_0 - w_0\| = \gamma_0.$$

which results in $J(v_0, w_0) = 0$. Considering the optimality property in (5), we have

$$J(v_0^*, w_0^*) \leqslant J(v_0, w_0),$$

which contradicts the initial assumption that $h_1(\gamma_0) > 0$ and $h_1(\gamma_0) = J(v_0^*, w_0^*)$. $\qquad\square$

**Lemma 3.** *For a convex PWA function $f_p$ expressed by (2), the convexity modulus $h_1$ is bounded by $\hat{h}_1 \leqslant h_1$, with*

$$\hat{h}_1(\gamma) := \begin{cases} 0 & \text{if } \gamma < \operatorname{diam}(\mathcal{C}_{p,.}) \\ c_1\gamma + c_0 & \text{if } \gamma \geqslant \operatorname{diam}(\mathcal{C}_{p,.}) \end{cases}, \qquad (18)$$

*where*

$$c_1 = \min_{j \in \mathcal{I}_{Q_p}} \left\{ \frac{a_{p,i}^T - a_{p,j}^T}{2} \right\}, \qquad (19a)$$

$$\text{s.t.} \quad i = \arg\max_{q \in \mathcal{I}_{Q_p}} \operatorname{diam}(\mathcal{C}_{p,q}), \qquad (19b)$$

$$\mathcal{C}_{p,i} \cap \mathcal{C}_{p,j} \neq \emptyset, \qquad (19c)$$

*and $c_0 = c_1 \operatorname{diam}(\mathcal{C}_{p,.})$.*

*Proof.* This can be directly deduced from Proposition 2, considering the continuity of $h_1$ from Lemma 2, the piecewise-constant property of $\partial h_1/\partial \gamma$ from Lemma 1, and the increasing property of $h_1$ from Proposition 1. $\qquad\square$

We are now in the position to state our main result:

**Theorem 3.** *Let $F : \mathcal{D} \to \mathbb{R}$ be a scalar-valued objective function and let $f$ be a continuous PWA function as in Definition 1 that approximates $F$ with bounded approximation error $\delta = f - F$. Let $f_p$ in (2) be the local convex segment of $f$ in its MMPS form (1) on the set $\mathcal{C}_{p,.}$, and let $\delta_p : \mathcal{C}_{p,.} \to \mathbb{R}$ be the corresponding approximation error bounded by*

$$\sup_{x \in \mathcal{C}_{p,.}} |\delta_p(x)| = \Delta_p < \infty.$$

*Let $x_p^*$ be any global minimizer of $f_p$ and $\hat{x}_p^*$ be any global minimizer of $F$ on $\mathcal{C}_{p,.}$. Then, the following condition holds:*

$$\|\hat{x}_p^* - x_p^*\| \leqslant \frac{2\Delta_p}{c_1} + \max_{q \in \mathcal{I}_{Q_p}} \{\operatorname{diam}(\mathcal{C}_{p,q})\}, \qquad (20)$$

*where $c_1$ is defined in (19).*

*Proof.* This can be directly concluded by extending Theorem 2 via considering Proposition 2 and Lemma 3. $\qquad\square$

## V. CASE STUDY

To showcase the application of Theorem 3, we select the 1-dimensional cut of the well-known Eggholder test function [22] at $x_2 = 0$ given by

$$F(x) = -47 \sin\left(\sqrt{\left|\frac{x}{2} + 47\right|}\right) - x \sin\left(\sqrt{|x - 47|}\right).$$

We approximate $F$ by a continuous PWA function $f$ that we arbitrarily take as

$$f(x) = \min_{p \in \mathcal{I}_5} (f_{p,.}).$$

with local convex segments

$$f_{1,.} = \max_{q \in \mathcal{I}_3} (f_{1,q}), \qquad \mathcal{C}_{1,.} = [-512, -385], \qquad (21a)$$

$$f_{2,.} = f_{2,1}, \qquad \mathcal{C}_{2,.} = [-385, -330], \qquad (21b)$$

$$f_{3,.} = \max_{q \in \mathcal{I}_3} (f_{3,q}), \qquad \mathcal{C}_{3,.} = [-330, -180], \qquad (21c)$$

$$f_{4,.} = \max_{q \in \mathcal{I}_3} (f_{4,q}) \qquad \mathcal{C}_{4,.} = [-180, 180], \qquad (21d)$$

$$f_{5,.} = f_{5,1}, \qquad \mathcal{C}_{5,.} = [180, 512]. \qquad (21e)$$

Figure 2 shows the plots for the nonlinear objective function $F$ and its PWA approximation $f$. The subregions $\mathcal{C}_{p,.}$ with $p \in \mathcal{I}_5$ are illustrated by different colors. Theorem 3 can be used in two ways:

1) guaranteeing bounds on the distance between the regional minima of $F$ and $f$ on each subregion, given a bound on the approximation error, and
2) finding the required criteria for the approximation to obtain a desired bound on the distance between these minima, which we refer to as the confidence radius.

We discuss each case separately by focusing on the approximation on $\mathcal{C}_{3,.}$.
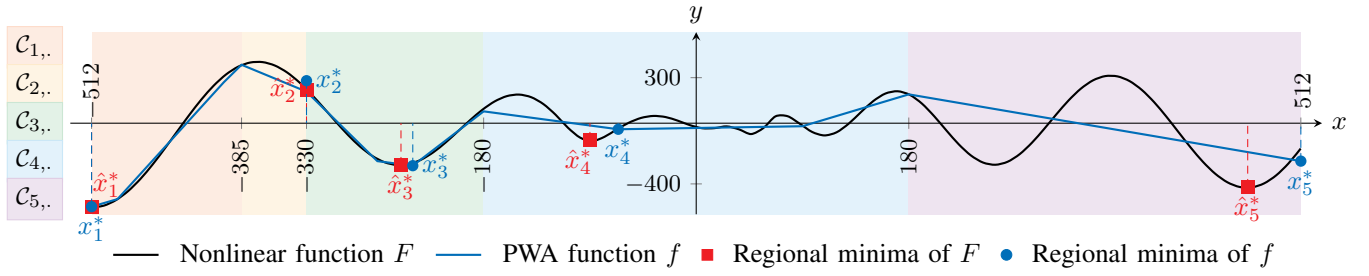
Fig. 2: Plots of the nonlinear objective function $F$ and its PWA approximation $f$.

## A. Case 1: Finding the Confidence Radii

Figure 3a shows $F$ and two PWA approximations on $\mathcal{C}_{p,\cdot}$ with two approximation errors. The first is $f_3^{(1)}$, which divides $\mathcal{C}_{3,\cdot}$ into 3 subregions with maximum approximation error $\Delta_3^{(1)} = 19.9$, and which is given by

$$f_3^{(1)}(x) = \max \left\{ \begin{array}{c} -7.8x - 2365.7 \\ -0.9x - 501.2 \\ 6.1x + 1176.1 \end{array} \right\}.$$

The second approximation is $f_3^{(2)}$ with 8 affine pieces, the maximum error $\Delta_3^{(2)} = 2.6$, and defined as

$$f_3^{(2)}(x) = \max \left\{ \begin{array}{c} -8.6x - 2613.1 \\ -6.8x - 2095.6 \\ -4.6x - 1477.9 \\ -2.2x - 829.8 \\ 0.3x - 191.6 \\ 2.8x + 412.5 \\ 5.1x + 944 \\ 6.9x + 1348.1 \end{array} \right\}.$$

The inverses of the corresponding convexity moduli are shown in Fig. 3b in the same color as their corresponding $f$ in Fig. 3a, where $\chi$ is the confidence radius in Definition 3.

Using Theorem 3, the confidence radius for $f_3^{(1)}$ by is obtained by $\chi^{(1)} = 70$, which is the same value obtained by finding $h_1$ and its inverse function using (5), which is presented in Fig. 3b. The same process can be performed for the second approximation, $f_3^{(2)}$, which gives $\chi^{(2)} = 44.3$. Note that Theorem 3 is more conservative for larger values of $\Delta$, compared to directly using the definition of $h_1$. For instance, if $\Delta_3^{(2)} = 12.5$, employing Theorem 3 leads to $\chi^{(2)} = 65.7$, while the computed confidence radius using $h_1$ is 58.4. The areas within the confidence radii for the PWA approximations are highlighted on the $x$-axis in Fig. 3a as well.

## B. Case 2: Finding the Approximation Criteria

In this case, we approach the problem from another direction: we select a desired confidence radius $\chi^{(3)}$ and find the required criteria for the corresponding approximated function, $f_3^{(3)}$. Let the desired $\chi^{(3)} = 15$; then,

$$\frac{2\Delta_3^{(3)}}{c_1} + \max_{q \in \mathcal{I}_{Q_p}} \left\{ \text{diam}(\mathcal{C}_{3,q}) \right\} \leqslant 10,$$

which means the diameter of the largest subregion $\mathcal{C}_{3,\cdot}$ must be smaller than 10. Firstly, given that $\text{diam}(\mathcal{C}_{3,\cdot}) = 150$, it can be concluded that the PWA approximation requires at least 10



(a) Representation of $f$ and $F$
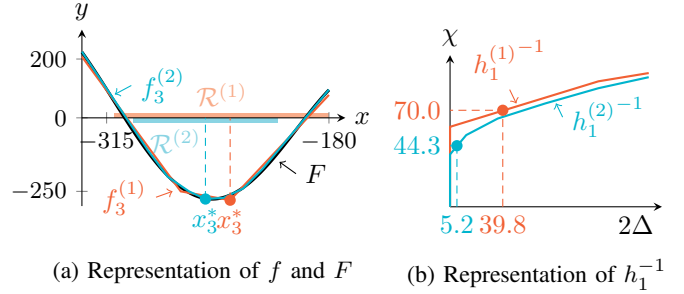


(b) Representation of $h_1^{-1}$

Fig. 3: Comparison of two different PWA approximations of the nonlinear function on $\mathcal{C}_{3,\cdot}$.

partitions. We can then start the approximation by partitioning $\mathcal{C}_{p,\cdot}$ into 15 subregions with the same diameter and find the lowest possible error bound $\Delta_3^{(3)}$ for the approximation, which is obtained as 2.83 with $c_1 = 0.0072$. For this approximation, $\chi^{(3)}$ already exceeds $\text{diam}(\mathcal{C}_{3,\cdot})$.

To improve upon this example, we add another partition to reduce the largest partition diameter further and this time we do not aim at partitions of $\mathcal{C}_{3,\cdot}$ with the same diameter, but require

$$\max_{q \in \mathcal{I}_{16}} \left\{ \text{diam}\left(\mathcal{C}_{3,q}\right) \right\} \leqslant 10.$$

We find $\Delta_3^{(3)} = 2.47$ with $c_1 = 1.03$ and

$$\max_{q \in \mathcal{I}_{16}} \left\{ \text{diam}\left(\mathcal{C}_{3,q}\right) \right\} = 9.4.$$

For this values, we obtain $\chi^{(3)} = 14.24$. In case this value is acceptable, we can use the corresponding PWA approximation while ensuring that the minimizer of $F$ on $\mathcal{C}_{3,\cdot}$ lies in a ball or radius 14.24 around the minimizer of $f^{(3)}$. In case a tighter confidence radius is desired, the same procedure can be followed by adding more subregions.

## C. Application for NMPC Optimization

To showcase the application of our proposed method in PWA approximation to control optimization problems, we use an inverted-pendulum NMPC problem as in [1]. The considered prediction horizon is $N_{\mathrm{p}} = 2$ and the initial rotation velocity is set as $\dot{\theta} = 0$. The objective function $J_{\mathrm{NMPC}}$ can be formulated as a function of the measured pendulum angle $\theta_k$ and the control inputs $u_k$ and $u_{k+1}$. For instance, for $\theta_k = 0$

we have

$$J_{\text{NMPC}}(0, u_k, u_{k+1}) = \sqrt{(0.02u_k + \pi)^2 + 2\pi^2}$$
$$+ 0.02\sqrt{u_k^2 + (u_k + u_{k+1})^2} + 0.01\sqrt{u_k^2 + u_{k+1}^2}.$$

Moreover, the feasible region is defined as the box constraint $|u_{k+i-1}| \leqslant 20N$, $i \in \mathcal{I}_2$ with $\text{diam}(\mathcal{C}_{1,.}) = 56.4$.

We approximate $J_{\text{NMPC}}$ by two convex MMPS forms $f^{(1)}$ and $f^{(2)}$ – with $P^{(1)} = P^{(2)} = 1$ in (1) – with different complexities in terms of the number of affine functions as

$$Q^{(1)} = 4, \quad \Delta^{(1)} = 0.19, \quad \max_{q \in \mathcal{I}_4}\{\text{diam}\left(\mathcal{C}_{1,q}^{(2)}\right)\} = 28.2,$$

$$Q^{(2)} = 24, \quad \Delta^{(2)} = 0.01, \quad \max_{q \in \mathcal{I}_{24}}\{\text{diam}\left(\mathcal{C}_{1,q}^{(2)}\right)\} = 14.6.$$

The inverse of the convexity modulus and the correspond-



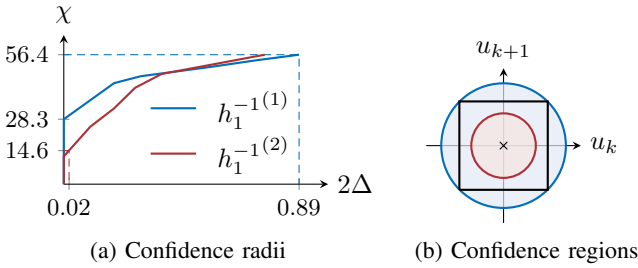(a) Confidence radii      (b) Confidence regions

Fig. 4: Comparison of two PWA approximations of NMPC.

ing confidence regions for both approximations are shown in Fig. 4. While $f^{(1)}$ has a low approximation error, its complexity level does not allow to guarantee a confidence radius lower than the diameter of the feasible region. However, the more accurate approximation $f^{(2)}$ guarantees a smaller confidence radius. Moreover, a general approximation criterion can be obtained, similar to the Eggholder NLP example, for an NMPC problem. In this case, it can be observed from (20) and Fig. 4a that $\chi$ is lower-bounded by the maximum subregion diameter. Therefore, if a particular confidence radius is desired, the approximation problem (10) can be solved while imposing constraints on the diameter of subregions, e.g. an upper bound on the maximum subregion diameter.

### D. Discussion

The Eggholder function and the NMPC case studies are two potential applications with, respectively, 1- and 3-dimensional domains to illustrate the theoretical results. Nevertheless, our proposed approach is applicable in higher dimensions since we do not assume any bounds on the dimension of $\mathcal{D}$ throughout this paper. Moreover, the following aspects should be considered when applying the proposed approach:

- The conservatism of the current approach in obtaining the confidence radii can be high, which can be seen in the case studies.
- The complexity in determining the confidence radii is irrelevant to the dimension of the optimization problem. However, larger dimensions increase the computation

time and range of $\gamma$ values for which (5) needs to be solved. Nevertheless, since this problem is solved offline, higher computation time do not limit the applicability of our approach.

## VI. CONCLUSIONS

This paper has introduced a novel approach for bounding the minimizers of polytopically-constrained NLPs with non-linear continuous objective functions using continuous PWA function approximations. We have leveraged the continuity of the PWA approximations resulting from employing an MMPS formalism to construct a locally-convex representation of the PWA approximation, thus facilitating the derivation of guaranteed bounds on the distance between the original and the approximated optimal solutions of the NLP by considering the maximal approximation error. Our approach offers a practical tool for determining criteria to achieve desired solution bounds. The effectiveness of the method has been demonstrated through two case studies on the Eggholder function and NMPC of an inverted pendulum, highlighting the practical application of the proposed method and its potential impact in optimization and optimal control.

For future work, our primary objectives are conducting an in-depth analysis of the conservatism of our approach, refining our sensitivity analysis, and extending our method to NLPs with non-convex constraints. Moreover, we aim to do more comparative studies to gain insight into the impacts of the improved computational efficiency through PWA approximation in light of the corresponding solution bounds. Finally, investigating the effects of probabilistic error bounds would be an interesting direction to help integrate our approach into learning-based and data-driven applications.

## REFERENCES

[1] S. Gros, M. Zanon, R. Quirynen, A. Bemporad, and M. Diehl. From linear to nonlinear MPC: bridging the gap via the real-time iteration. *International Journal of Control*, 93(1):62–80, 2020.

[2] D. W. Griffith, L. T. Biegler, and S. C. Patwardhan. Robustly stable adaptive horizon nonlinear model predictive control. *Journal of Process Control*, 70:109–122, 2018.

[3] L. Dong, J. Yan, X. Yuan, H. He, and C. Sun. Functional nonlinear model predictive control based on adaptive dynamic programming. *IEEE Transactions on Cybernetics*, 49(12):4206–4218, 2019.

[4] V. M Zavala and L. T. Biegler. The advanced-step NMPC controller: Optimality, stability and robustness. *Automatica*, 45(1):86–93, 2009.

[5] Z. Jin, Q. Shen, and S. Z. Yong. Mesh-based piecewise affine abstraction with polytopic partitions for nonlinear systems. *IEEE Control Systems Letters*, 5(5):1543–1548, 2021.

[6] A. Bemporad, A. Oliveri, T. Poggi, and M. Storace. Ultra-fast stabilizing model predictive control via canonical piecewise affine approximations. *IEEE Transactions on Automatic Control*, 56(12):2883–2897, 2011.

[7] J. Xu, X. Huang, X. Mu, and S. Wang. Model predictive control based on adaptive hinging hyperplanes model. *Journal of Process Control*, 22(10):1821–1831, 2012.

[8] B. De Schutter and T. J.J. van den Boom. MPC for continuous piecewise-affine systems. *Systems and Control Letters*, 52(3-4):179–192, 2004.

[9] L. Gharavi, A. Dabiri, J. Verkuijlen, B. De Schutter, and S. Baldi. Proactive emergency collision avoidance for automated driving in highway scenarios. *arXiv*, (2310.17381), 2023.

[10] Y. Chen, M. Bruschetta, D. Cuccato, and A. Beghi. An adaptive partial sensitivity updating scheme for fast nonlinear model predictive control. *IEEE Transactions on Automatic Control*, 64(7):2712–2726, 2019.

[11] M. Ławryńczuk, P. M. Marusak, P. Chaber, and D. Seredyński. Initialisation of optimisation solvers for nonlinear model predictive control: Classical vs. hybrid methods. *Energies*, 15(7):2483, 2022.

[12] T. Ghandriz, B. Jacobson, P. Nilsson, and L. Laine. Trajectory-following and off-tracking minimisation of long combination vehicles: A comparison between nonlinear and linear model predictive control. *Vehicle System Dynamics*, pages 1–34, 2023.

[13] J. W. Daniel. Stability of the solution of definite quadratic programs. *Mathematical Programming*, 5(1):41–53, 1973.

[14] H. X. Phu. Minimizing convex functions with bounded perturbations. *SIAM Journal on Optimization*, 20(5):2709–2729, 2010.

[15] C. Büskens and H. Maurer. Sensitivity analysis and real-time optimization of parametric nonlinear programming problems. In M. Grötschel, S. O. Krumke, and J. Rambau, editors, *Online Optimization of Large Scale Systems*, pages 3–16. Springer Berlin Heidelberg, 2001.

[16] I. Subotic, A. Hauswirth, and F. Dorfler. Quantitative sensitivity bounds for nonlinear programming and time-varying optimization. *IEEE Transactions on Automatic Control*, 67(6):2829–2842, 2022.

[17] S. Shin and V. M. Zavala. Diffusing-horizon model predictive control. *IEEE Transactions on Automatic Control*, 68(1):188–201, 2023.

[18] H. Pirnay, R. López-Negrete, and L. T. Biegler. Optimal sensitivity based on IPOPT. *Mathematical Programming Computation*, 4(4):307–331, 2012.

[19] M. A. Muller, D. Angeli, and F. Allgower. On necessity and robustness of dissipativity in economic model predictive control. *IEEE Transactions on Automatic Control*, 60(6):1671–1676, 2015.

[20] Y. Puranik and N. V. Sahinidis. Domain reduction techniques for global NLP and MINLP optimization. *Constraints*, 22(3):338–376, 2017.

[21] C. Liu, S. Shi, and B. De Schutter. Stability and performance analysis of model predictive control of uncertain linear systems. *arXiv preprint arXiv:2405.15552*, 2024.

[22] D. Whitley, S. Rana, J. Dzubera, and K. E. Mathias. Evaluating evolutionary algorithms. *Artificial Intelligence*, 85(1):245–276, 1996.

[23] L.O. Chua and A.-C. Deng. Canonical piecewise-linear representation. *IEEE Transactions on Circuits and Systems*, 35(1):101–111, 1988.