An algorithm for PWL approximations of nonlinear functions

A. Alessio, A. Bemporad, B. Addis, and A. Pasini

September 24, 2005

Abstract

In this report we provide some technical details for some of the results appeared in [Alessio et al.(2005)]. In the first section we provide the proof of continuity of the PPWA function computed with the "squaring the circle" algorithm stated in ACC 06. Then, we analyze the complexity of the previous algorithm, in terms of the desired level of accuracy in the approximation of the PPWA function.

1 Definitions and Notations

A subset $\mathcal{P}$ of $\mathbb{R}^n$ is called a *convex polyhedron* if it is the set of solutions to a finite system of linear inequalities, and called a *convex polytope* if it is convex and bounded. For the sequel, we might omit convex for convex polytopes and convex polyhedra, and call them simply polytopes and polyhedra. Let $\mathcal{P}$ be a polyhedron (or a polytope). A linear inequality $c^T x \leq d$ is called *valid* for $\mathcal{P}$, if $c^T x \leq d$ holds for all $x \in \mathcal{P}$. A subset F of a polyhedron (or a polytope) $\mathcal{P}$ is called a *face* if it is represented as

$$F = \mathcal{P} \cap \{x : c^T x = d\} \tag{1}$$

for some valid inequalities. By this definition, both the empty set $\emptyset$ and the whole set $\mathcal{P}$ are faces. The faces of dimension $0, 1, dim(\mathcal{P})-1$ are called *vertices, edges* and *facets* respectively. The vertices are defined as points which cannot be represented as convex combinations of two other points of $\mathcal{P}$.
Let $\mathcal{P} \subseteq \mathbb{R}^n$ be a polyhedral set. Given $(v^0, .., v^n)$ affinely independent points of $\mathcal{P}$, we define a simplex S as

$$S \triangleq \{x \in \mathbb{R}^n \; | x = \sum_{l=0}^{n} \mu_l v^l, \sum_{l=0}^{n} \mu_l = 1, \\ \mu_l \geq 0 \text{ for } l = 0, 1, \ldots, n\}.$$

A simplex S in $\mathbb{R}^n$ has exactly $n+1$ vertices and $n+1$ facets.

In the sequel we will refer to $\Omega \subseteq \mathbb{R}^n$ and $\Theta \subseteq \mathbb{R}$ as two open sets.

Definition 1 *Given a definite positive matrix A, a generic ellipsoidal set $\mathcal{E}_{(A,a,r)}$ is defined as $\{x \in \Omega | x'Ax + ax \leq r^2\}$, where $a \in R^{1\times n}$, $r \in \mathbb{R}$.*

Definition 2 *A function $f : \Omega \to \Theta$ is* quadratic *if $f(x) = x'Ax + ax + \alpha$, with $A \in R^{n\times n}$, $a \in R^{1\times n}, \alpha \in \Theta$. In addition, f is convex, if and only if $A > 0$.*

So, given the matrices $(A > 0, a, r)$ defining an ellipsoidal set $\mathcal{E}_{(A,a,r)} \subset \Omega$, the function $f_{\mathcal{E}} : \Omega \to \Theta$ defined as

$$f_{\mathcal{E}}(x) = x'Ax + ax + \alpha_{\mathcal{E}}, \qquad x \in \Omega, \alpha \in \Theta \tag{2}$$

is quadratic and convex. Note that the level sets of $f_{\mathcal{E}}(x)$ are ellipsoidal sets $\mathcal{E}_{(A,a,r)}$. In fact, once a level $f_0 > f_{min} = -\frac{a^T A^{-1} a}{4} + \alpha_{\mathcal{E}}$ [1] is fixed, the curve $\gamma : x'Ax + ax = f^0 - \alpha$, defines an ellipse, $\forall x \in \Omega$. So, the set defined as $\{x \in \Omega | x'Ax + ax \leq f^0 - \alpha\}$ is an ellipsoid $\mathcal{E}_{(A,a,r)}$, where $r = (f^0 - \alpha)^{\frac{1}{2}}$.

Definition 3 *A function $\bar{f} : \Omega \to \Theta$ is piecewise affine (PWA) if there exists a partition $R_1, ..., R_N$ of Ω where $\bar{f}(x) = c_i x + d_i$, $\forall x \in R_i$, $i = 1, ..., N$.*

Definition 4 *A function $\bar{f} : \Omega \to \Theta$, where $\Omega \subset R^n$ is piecewise affine on polyhedra $(PPWA)$ if there exists a polyhedral partition $R_1, ..., R_N$ of Ω where $\bar{f}(x) = c_i x + d_i$, $\forall x \in R_i$, $i = 1, ..., N$.*

Given a PPWA function $\bar{f}(x) = c_i x + d_i$, defined over a polyhedral partition $\{R_i\}_{i=1}^N$, $(R_i \subset \Omega,\ i = 1, .., N)$, where the generic region R_i is defined by the inequalities $A_i x \leq b_i$, for $x \in \Omega$, the following proposition states how to compute the level sets of $\bar{f}(x)$ over $\bigcup_{i=1}^N R_i$.

Proposition 5 *The level sets of a PPWA function $\bar{f} : \Omega \to \Theta$, where $\bar{f}(x) = c_i x + d_i$ is defined over a polyhedral partition $\{R_i\}_{i=1}^N$, $(R_i \subset \Omega,\ i = 1, .., N)$, are* polyhedrons *$\mathcal{P} \subset \Omega$ defined as*

$$\mathcal{P} = \{x \in R_i | c_i x \leq \bar{f}_z - d_i\}, \tag{3}$$

where $\bar{f}_z \in \Theta$ is the desired fixed level.

Note that, while in the quadratic and convex case the boundary of the level sets obtained by *cutting* the initial function f at a desired level are given by a curve γ that defines an *ellipse* in the space Ω, in the piecewise linear case the boundary is defined by a closed curve $\bar{\gamma} : c_i x = \bar{f}_z - d_i$, $\forall x \in R_i$.

Definition 6 *A function $\bar{f} : \Omega \to \Theta$ is a* PWA approximation *of $f : \Omega \to \Theta$ on polyhedra, if there exists a polyhedral partition $R_1, ..., R_N$ of Ω where $\bar{f}(x) = c_i x + d_i$, $\forall x \in R_i$, $i = 1, ..., N$, and $\forall \epsilon > 0, \|\bar{f}(x) - f(x)\| \leq \epsilon$, $\forall x \in \Omega$.*

[1] If $f^0 \leq f_{min} = -\frac{a^T A^{-1} a}{4} + \alpha_{\mathcal{E}}$ the level sets of the function are either $\{0\}$ or $\emptyset$.

Definition 7 *A function $\bar{f} : \Omega \to \Theta$ is a* PWA over-approximation *of $f : \Omega \to \Theta$ on polyhedra, if there exists a polyhedral partition $R_1, ..., R_N$ of Ω where $\bar{f}(x) = c_i x + d_i$, $\forall x \in R_i$, $i = 1, ..., N$, is a PWA approximation of f and $\bar{f}(x) \geq f(x)$ $\forall x \in \Omega$.*

So, given two closed ellipsoidal sets $\beta\mathcal{E}_{(A,a,r)} (= \mathcal{E}_{(A,a,\beta^{\frac{1}{2}} r)})$ and $\mathcal{E}_{(A,a,r)}$, where $\beta\mathcal{E}_{(A,a,r)} \subset \mathcal{E}_{(A,a,r)}$, as $\beta < 1$, and their associated quadratic functions $f_{\beta\mathcal{E}}(x)$ and $f_{\mathcal{E}}(x)$ as defined in (2), the problem of fitting a polyhedron $\mathcal{P}$ in between the two closed sets $\beta\mathcal{E}$ and $\mathcal{E}$, amounts of finding a PWA over-approximation $\bar{f}(x)$ of $f_{\mathcal{E}}(x)$, within a given bounded error ϵ where $\epsilon = \min_x (f_{\beta\mathcal{E}}(x) - f_{\mathcal{E}}(x))$, $x \in \Omega$. This leads to the fact that

$$f_{\mathcal{E}}(x) \leq \bar{f}(x) < f_{\beta\mathcal{E}}(x), \ \forall x \in \Omega \tag{4}$$

The first inequality comes from the fact that $\bar{f}$ is a PWA over-approximation of $f_{\mathcal{E}}$, while the validity of the second one is due to the fact that since $\bar{f}(x) - f_{\mathcal{E}}(x) < \epsilon \leq f_{\beta\mathcal{E}}(x) - f_{\mathcal{E}}(x)$, $\forall x \in \Omega$, then $\bar{f}(x) - f_{\beta\mathcal{E}}(x) < 0$, $\forall x \in \Omega$.

Note that since (4) is valid, the level sets of the three functions denoted by $\mathcal{E}, \mathcal{P}, \beta\mathcal{E}$, respectively will have the property that $\beta\mathcal{E} \subset \mathcal{P} \subset \mathcal{E}$. So, the desired polytope $\mathcal{P}$ between the two ellipsoidal sets $\mathcal{E}, \beta\mathcal{E}$ is the level set of a PPWA function not greater than $f_{\beta\mathcal{E}}$, that over approximates $f_{\mathcal{E}}$.

Note that once $\beta\mathcal{E}$ and $\mathcal{E}$ are given, the two associated quadratic functions $f_{\beta\mathcal{E}}(x)$ and $f_{\mathcal{E}}(x)$ as defined in (2), are not unique. Their representation depends in fact on the two constants $\alpha_{\mathcal{E}}$ and $\alpha_{\beta\mathcal{E}}$. In the following sections, we will give a more detailed description of the procedure stated for obtaining such polyhedron $\mathcal{P}$ between two closed ellipsoidal sets $\beta\mathcal{E}$ and $\mathcal{E}$. In the sequel we will simplify our notation omitting the explicit dependence on the triple (A, a, r) of the generic ellipsoid $\mathcal{E}$, as we will refer to our particular case, where $\mathcal{E} := \{x \in \Omega | x'Ax + ax \leq r^2\}$ and $\beta\mathcal{E} := \{x \in \Omega | x'Ax + ax \leq \beta r^2\}$

2 Continuity of the PPWA function

Given $f : \Omega \to \Theta$, where Ω and Θ are open sets of $\mathbb{R}^n$ and $\mathbb{R}$ respectively, and a non empty polytope $\mathcal{P} \subset \Omega$, we can split the polytope $\mathcal{P}$ into a certain number of simplices S_i, $i = 1, .., s$, using, for example, the *Delaunay Triangulation*. The set of simplices so obtained has the following properties

$$(i) \quad int(S_i) \cap int(S_j) \quad = \emptyset, \quad \forall i, j = 1, .., s, \quad (i \neq j); \tag{5}$$

$$(ii) \quad \bigcup_{i=1}^{s} S_i \quad = \mathcal{P}. \tag{6}$$

Every single simplex S_i has $n+1$ vertices $[v_i^0, .., v_i^n]$, for $i = 1, .., s$. Define the matrices M_i, $i = 1, .., s$ as follows

$$M_i \triangleq \begin{bmatrix} 1 & 1 & .. & 1 \\ v_i^0 & v_i^1 & .. & v_i^n \end{bmatrix}, \tag{7}$$

and the vectors

$$V_i \triangleq [\ f_{\mathcal{E}}(v_i^0)\ f_{\mathcal{E}}(v_i^1)\ ..\ f_{\mathcal{E}}(v_i^n)\ ]^\top, i = 1, .., s.$$

Theorem 8 (Continuity of the PPWA function) *Given a simplex $\mathcal{S} \subset \mathbb{R}^n$ and a set $\{S_i\}_{i=0}^l$ of simplices obtained by the* Delaunay Triangulation *of $\mathcal{S}$, the PPWA $\bar{f} : \mathcal{S} \to \mathbb{R}$ such that $\bar{f}(x) = \begin{cases} \bar{f}_i(x) & \forall x \in S_i \\ 0, & elsewhere, \end{cases}$ where $\bar{f}_i(x) = V_i^T M_i^{-1} \left[\begin{smallmatrix} 1 \\ x \end{smallmatrix}\right], \forall x \in S_i$ is continuous $\forall x \in \mathcal{S}$.*

proof

By definition of V_i, M_i^{-1}, the function $\bar{f}_i$ is an affine function of x continuous $\forall x \in S_i$, $i = 0, .., l$. To show the continuity of the entire function $\bar{f}$ over $\mathcal{S}$, we need to prove that given two simplices S_i, S_j such that $S_i \cap S_j \neq \emptyset$,

$$\bar{f}_i(x) = \bar{f}_j(x), \forall x \in S_i \cap S_j. \tag{8}$$

Since, by construction, the set of simplices $\{S_i\}_{i=0}^l$ is obtained through the Delaunay Triangulation of $\mathcal{S} \subset \mathbb{R}^n$, two simplices S_i, S_j can share at most a *face* [2], that is a lower-dimensional subspace of $\mathcal{S}_i$ (or $\mathcal{S}_j$). Since a *face* is, in general, the convex hull of a certain number p of vertices (that, in this case, can be at most $(n-1)$), in order to prove that $\bar{f}(x)$ is continuous $\forall x \in \mathcal{S}$, is sufficient to show that condition (8) is true on the vertices in common between S_i and S_j. That is

$$\bar{f}_i(v_k) = \bar{f}_j(v_k) = \bar{f}(v_k), \forall v_k \in S_i \cap S_j. \tag{9}$$

But the value of $\bar{f}_i(v_k)$ does not depend on the simplex S_i. In fact, if we rewrite the initial definition of a generic simplex $\mathcal{S}$ so that

$$\mathcal{S} = \{x \in \mathbb{R}^n, \mu \in \mathbb{R}^{n+1} | M^{-1} \left[\begin{smallmatrix} 1 \\ x \end{smallmatrix}\right] = \mu\}, \tag{10}$$

where $\mu = [\mu_0, .., \mu_n]'$ is a vector of coefficients such that given $x \in \mathcal{S}$,

$$x = \sum_{l=0}^{n} \mu_l v^l, \qquad \sum_{l=0}^{n} \mu_l = 1. \tag{11}$$

The value of $\bar{f}_i$ on the generic vertex v_k of $S_i \cap S_j$ is given by

$$\bar{f}(v^k) = V^T M^{-1} \left[\begin{smallmatrix} 1 \\ v^k \end{smallmatrix}\right]. \tag{12}$$

Since $M^{-1} \left[\begin{smallmatrix} 1 \\ v^k \end{smallmatrix}\right] = \begin{bmatrix} 0 \\ \vdots \\ \mu_k \\ \vdots \\ 0 \end{bmatrix} = \mu$, we have that $\sum_{l=0}^n \mu_l = \mu_k = 1$. The value of $\bar{f}_i(v_k)$ is function of v_k only, and does not depend on the simplex S_i. In fact

$$\bar{f}_i(v^k) = V^T \begin{bmatrix} 0 \\ \vdots \\ \mu_k \\ \vdots \\ 0 \end{bmatrix} = V^T \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = f_{\mathcal{E}}(v^k).$$

[2]The faces of dimension 0, 1, $dim(S_i) - 1$ are called *vertices, edges* and *facets*, respectively.

Since $f_{\mathcal{E}}(v^k)$ depends only on the vertex v_k and not on the generic simplex $\mathcal{S}_i$, the values assumed by $\bar{f}_i$ and $\bar{f}_j$ on each common vertex v_k will be equal to $f_{\mathcal{E}}(v_k)$. That is,

$$\bar{f}_i(v_k) = \bar{f}_j(v_k) = \bar{f}(v_k) = f_{\mathcal{E}}(v_k), \forall v_k \in \mathcal{S}_i \cap \mathcal{S}_j, i, j = 0, .., l, i \neq j. \tag{13}$$

In consequence of (13), $\bar{f}_i(x) = \bar{f}_j(x)\ \forall x \in \mathcal{S}_i \cap \mathcal{S}_j$.

3 Complexity

In this section we will analyze the complexity of the algorithm described in [Alessio et al.(2005)], in terms of the desired level of accuracy in the approximation of the quadratic and convex function $f_{\mathcal{E}}$ with the PPWA function $\bar{f}$. Define $\epsilon_0 = \max_{x \in S_i} \bar{f}_i(x) - f_{\mathcal{E}},\ (i = 1, .., s)$, as the initial maximum error committed at the *first* step of the algorithm in [Alessio et al.(2005)], and $\epsilon_f = \min_{x \in S_i} f_{\beta\mathcal{E}}(x) - f_{\mathcal{E}}(x) = \alpha_{\beta\mathcal{E}} - \alpha_{\mathcal{E}}$. So, algorithm in [Alessio et al.(2005)] computes a PPWA function $\bar{f}$ such that

$$\max_{x \in S_i} \bar{f}_i(x) - f_{\mathcal{E}}(x) < \epsilon_f,\ (i = 1, .., s). \tag{14}$$

Note that conditions (14) and (4) are equivalent, but since condition (4) is useful to describe the relation between the functions and their respective level sets, condition (14) makes the study of the algorithm complexity easier. So, in the following, we will study the complexity of the computation of a PPWA function $\bar{f}$, described in section 3, that approximates a convex and quadratic function $f_{\mathcal{E}}$ within a given bounded error ϵ_f.

3.1 Computation of a 1-step error (lower bound)

Given a function $f_{\mathcal{E}} : \mathbb{R}^n \to \mathbb{R}$ that associates $x \in \mathbb{R}^n \to x^T A x + a^T x + \alpha$, where A is a *symmetric positive definite* matrix, $a \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$. Define $v_0, .., v_n \in \mathbb{R}^n$ as $n+1$ geometric independent points of $\mathbb{R}^n$, that is $v_1 - v_0, .., v_n - v_0$ are linearly independent.
Furthermore define a function $\bar{f} : \mathbb{R}^n \to \mathbb{R}$ so that $x \in \mathbb{R}^n \to c^T x + d$, where $c \in \mathbb{R}^n \backslash \{0\}$ and $f_{\mathcal{E}}(v_i) = \bar{f}(v_i)$ for $i = 0, .., n$.
Define v_{max} as the $\arg\max_{x \in S} f_{\mathcal{E}}(x) - \bar{f}(x)$, where S is the simplex with vertices $v_0, .., v_n$, and $\delta = \max_{x \in S} f_{\mathcal{E}}(x) - \bar{f}(x)$.

Now, substitute the initial simplex $S = [v_0, .., v_n]$ with the $n+1$ simplices $S_k = [v_0, .., v_{k-1}, v_{max}, v_{k+1} .., v_n]$ obtained through the substitution of one vertex at time with the vector obtained before and denoted as v_{max}.
At this point we compute $\bar{f}$ again starting from the vertices of S_k, and then the new vectors v'_{max}, δ' with the new function $f_{\mathcal{E}} - \bar{f}$. Now we want to state the dependence between v'_{max}, v_{max} and δ', δ.

Note that $v_{max} \in S$, as A is positive definite. Moreover, it can happen that v_{max} lies on a facet of S, and not only into its interior. In this case, if $[v_{i_1}, .., v_{i_s}]$ is the minimal facet which contains v_{max}, we will substitute v_{max} with v_k, where $k = i_1, .., i_s$. However, v_{max} cannot lie on any vertex of S.

For sake of simplicity, we make a change of coordinates putting the origin into v_0. In this way, $\alpha_1 = \alpha$ and the function $f_{\mathcal{E}} - \bar{f}$ becomes:

$$f_{\mathcal{E}}(x) - \bar{f}(x) = \frac{1}{2}x^T Ax - b^T x \tag{15}$$

where $b = c - a$. So, we have that $v_{max} = A^{-1}b$ and $\delta = \frac{1}{2}b^T A^{-1}b$

Assume that v_{max} lies into the interior of S, otherwise if v_{max} lies into a facet of S, we will have the same following arguments.

Define $S' = [v_0 = 0, v_{max}, v_2, .., v_n]$ and $(c')^T x + \alpha$ as the function that takes the role of the new f. So, $v'_{max} = A^{-1}b'$ is the new point of maximum and $\delta' = \frac{1}{2}(b')^T A^{-1}b'$ is the new maximum, and $b' = c' - a$. So, we want now to obtain explicitly c' (and so b').

Define

$$V = \begin{bmatrix} v_2^T \\ .. \\ v_n^T \end{bmatrix}, \tilde{V} = \begin{bmatrix} v_1^T \\ .. \\ v_n^T \end{bmatrix}, \tilde{V}' = \begin{bmatrix} v_{max}^t \\ v_2^T \\ .. \\ v_n^T \end{bmatrix}, \tag{16}$$

and

$$\gamma = \begin{bmatrix} \frac{1}{2}v_1^T Av_1 \\ .. \\ \frac{1}{2}v_n^T Av_n \end{bmatrix}, \gamma' = \begin{bmatrix} \frac{1}{2}v_{max}^T Av_{max} \\ \frac{1}{2}v_2^T Av_2 \\ .. \\ \frac{1}{2}v_n T Av_n \end{bmatrix}. \tag{17}$$

The condition $f_{\mathcal{E}}(v_i) = \bar{f}(v_i), i = 0, .., n$ is now stated as $\tilde{V}b = \gamma$, $\tilde{V}'b' = \gamma'$. Furthermore, we have that $Vb = Vb' = \gamma_1$, where $\gamma_1 = [\frac{1}{2}v_2^T Av_2 .. \frac{1}{2}v_n^T Av_n^T]$. Follows that the two vectors b and b' belongs to the same subspace (of dimension 1), that is

$$b' = tb_0 + b, \tag{18}$$

where $b_0 \in ker(V)$, that is $Vb_0 = 0$.

We want to determine t. From the condition $\tilde{V}'b' = \gamma'$ we obtain that

1. $v_{max}^T(tb_0 + b) = \frac{1}{2}v_{max}^T Av_{max}$
2. $b^T A^{-1}(tb_0 + b) = \frac{1}{2}b^T A^{-1}AA^{-1}b = \frac{1}{2}b^T A^{-1}b$
3. $tb^T A^{-1}b_0 = -\frac{1}{2}b^T A^{-1}b$
4. $t = -\frac{\delta}{b^T A^{-1}b_0}$

So

$$b' = -\frac{\delta}{b^T A^{-1}b_0}b_0 + b \tag{19}$$

Let's compute δ':

$$\delta' = \frac{1}{2}(b')^T A^{-1} b' = \frac{1}{2}[(-\frac{\delta}{b^T A^{-1} b_0} b_0^T + b^T) A^{-1} (-\frac{\delta}{b^T A^{-1} b_0} b_0 + b)], \tag{20}$$

It follows that

$$\frac{1}{2}[\frac{\delta^2}{(b^T A^{-1} b_0)^2} b_0^T A^{-1} b_0 - 2\delta + 2\delta] = \frac{\delta}{2} \frac{\delta b_0^T A^{-1} b_0}{(b^T A^{-1} b_0)^2}. \tag{21}$$

In the end we obtain

$$\delta' = \frac{\delta}{4} \frac{b^T A^{-1} b b_0^T A^{-1} b_0}{(b^T A^{-1} b_0)^2} \tag{22}$$

Now: A^{-1} is positive definite as A, so the generic form $\alpha(x, y) := x^T A^{-1} y$ can be interpreted as a scalar product. If we define θ as the angle between b and b_0 in the metric defined by α we obtain that

$$\frac{(b^T A^{-1} b_0)^2}{b^T A^{-1} b b_0^T A^{-1} b_0} = \cos^2 \theta \tag{23}$$

and so

$$\delta' = \frac{\delta}{4} \frac{1}{cos^2 \theta}. \tag{24}$$

But $cos^2 \theta \leq 1$, so

$$\delta' \geq \frac{\delta}{4} \tag{25}$$

3.2 Computation of a 1-step error (upper bound)

Now compute $(c^T - (c')^T) A^{-1} b'$. This quantity is greater than zero for sure, this is due to the convexity of the function $f_{\mathcal{E}}$ and to the fact that $f'(x) = (c')^T x + d$ is a tighter approximation of $f_{\mathcal{E}}$ within the new simplex $S_1 = [v_0 = 0, v_{max}, .., v_n]$ obtained by the substitution of the first vertex v_1 with v_{max}.

Now, since $c = b + a, c' = b' + a$, we know that $c - c' = b - b'$. So,

$$(b^T - (b')^T) A^{-1} b' > 0 \tag{26}$$

and

$$(b')^T A^{-1} b' = 2\delta'. \tag{27}$$

We already know that $b' = t b_0 + b$, where $t = -\frac{\delta}{b^T A^{-1} b_0}$, then

$$b^T A^{-1} b' = b^T A^{-1} (t b_0 + b) = t b^T A^{-1} b_0 + b^T A^{-1} b = -\delta + 2\delta = \delta \tag{28}$$

So

$$b^T A^{-1} b' - 2\delta' > 0 \leftrightarrow \delta - 2\delta' > 0, \tag{29}$$

that is $\delta > 2\delta'$ or $\delta' < \frac{\delta}{2}$.

So, in the end

$$\frac{\delta}{4} \leq \delta' < \frac{\delta}{2}. \tag{30}$$

The algorithms builds recursively a tree, where at every node stores the vertices of the current simplex and the vector (c_i, α_i) which define the function $\bar{f}$ in that simplex. Every node has at most $n+1$ leaves, because if the approximation obtained is not below a desired error ϵ_f, it splits the simplex into at most $n+1$ simplices. Every one differs from the other only for the single vertex v_{max}. So, if we start from an initial error ϵ_0, at the first step of the procedure, we will add the first level to the tree, we will split the initial simplex into at most $n+1$ leaves and we are sure that the value of the error obtained at this level ϵ_1, belongs to the interval $[\frac{\epsilon_0}{4}, \frac{\epsilon_0}{2})$.
So at the $k-th$ level of the tree, the value of the error $\epsilon_k \in [\frac{\epsilon_0}{4^k}, \frac{\epsilon_0}{2^k})$. The algorithm stops when $k : \epsilon_k \leq \epsilon_f$, where ϵ_f is the desired approximation error. And so the height of the tree h is such that $\log_4 \frac{\epsilon_f}{\epsilon_0} \leq h < \log_2 \frac{\epsilon_f}{\epsilon_0}$.

References

[Alessio et al.(2005)] Alessio, A., Bemporad, A., Addis, B., Pasini, A., 2005. An algorithm for PWL approximations of nonlinear functions. Tech. rep., Universita di Siena, Italy, web: http://control.dii.unisi.it/research/ABAP05b.pdf.