

## Appendix B: proofs

**Proposition 1.** *Assume that the curve  $C_J$  is included in its osculating simplex on  $[0, 1]$ . If  ${}^t\mathbf{T}\boldsymbol{\alpha} \geq 0$ , then  $\forall x \in [0, 1]$ , we have  $F_\alpha(x) \geq 0$ .*

*Proof.* Because of the linearity of  $Z$  on  $\mathbb{T}$  and because a simplex is convex by definition,  $F_J(x)$  is guaranteed to be positive on  $[0, 1]$ , if it is positive in every vertices of  $S_J$ .  $\square$

### Preliminaries to Theorem 1

Theorem 1 needs the three following preliminary lemmas where we prove that in a small neighborhood of a point on a curve, a smooth curve is included in its osculating simplex.

Let  $T_0$  be a point on the curve  $C_J$  corresponding to  $x_0$  and  $T_J$  corresponding to  $x_J = x_0 + h$ . We denote  $T_0, T_1, \dots, T_J$  the vertices of the osculating simplex between  $x_0$  and  $x_0 + h$  (see Definition 2).

The vectors  $\mathbf{f}^{(1)}(x_0), \dots, \mathbf{f}^{(j)}(x_0)$  are all linearly independent. This results from the definition A1 of an Extended Chebyshev system. In this basis,

$$\overrightarrow{T_0 T_j} = \sum_{k=1}^j \gamma_{k,j} \mathbf{f}^{(k)}(x_0). \quad (1)$$

Similarly,

$$\overrightarrow{T_j T_J} = \sum_{k=j+1}^J \gamma_{k,j} \mathbf{f}^{(k)}(x_J).$$

**Lemma B1.** *Let  $D_j$  be the determinant*

$$D_j = \begin{vmatrix} \mathbf{f}^{(1)}(x_0) & \dots & \mathbf{f}^{(j)}(x_0) & \mathbf{f}^{(1)}(x_J) & \dots & \mathbf{f}^{(J-j)}(x_J) \end{vmatrix}.$$

$D_{j,k}$  is obtained by replacing the  $k$ -th column of  $D_j$  by  $\mathbf{f}(x_J) - \mathbf{f}(x_0)$ . Then, that for all  $j$ ,  $D_j \neq 0$ , and  $\gamma_{k,j} = \frac{D_{j,k}}{D_j}$ .

All the determinants  $D_j$  are strictly positive as a result of the definition A1 of ET systems.

For  $0 < j < J$ ,  $T_j$  belongs to the osculating  $j$ -space at  $T_0$  and simultaneously to the osculating  $J - j$ -space at  $T_J$ . Thus, the vector  $\overrightarrow{T_0 T_j}$  is a linear combination of the first  $j$  derivatives at  $T_0$  and similarly  $\overrightarrow{T_j T_J}$  is a linear combination of the first  $J - j$  derivatives at  $T_J$ . Consequently, the coordinates  $\gamma_{1,j}, \dots, \gamma_{J,j}$  of  $T_j$ , as stated in this lemma, result from the Cramer's rule applied to the linear system of equations:

$$\left( \mathbf{f}^{(1)}(x_0) \dots \mathbf{f}^{(j)}(x_0) \mathbf{f}^{(1)}(x_J) \dots \mathbf{f}^{(J-j)}(x_J) \right) \begin{pmatrix} \gamma_{1,j} \\ \dots \\ \gamma_{J,j} \end{pmatrix} = (\mathbf{f}(x_J) - \mathbf{f}(x_0))$$

**Lemma B2.** *The coefficients  $\gamma_{k,j}$  can be approximated by  $\gamma_{k,j} \sim \frac{h^k}{k!} + o(h^k)$ .*

We start from the previous lemma ?? and the expression of  $D_{j,k}$ .

Since  $x_J = x_0 + h$ , taking the Taylor expansion of  $f_1(x_0+h) - f_1(x_0), \dots, f_J(x_0+h) - f_J(x_0)$ , it can be readily shown that the first non vanishing term in the development of the  $D_{j,k}$  for  $1 \leq k \leq j$  is  $\frac{h^k}{k!} D_j$ . This results in the statement of the lemma.

**Lemma B3.** *Let  $T_0$  and  $T_J$  be two points on the curve  $C_J$  corresponding to  $x_0$  and  $x_J > x_0$ . Then, in the neighborhood of  $x_0$ , for  $h$  small enough, the portion of the curve  $C_J$  corresponding to  $x$  varying in  $(x_0, x_0 + h)$  is strictly included in the cone generated by  $\{\overrightarrow{T_0 T_1}, \dots, \overrightarrow{T_0 T_J}\}$ .*

*Proof.* Our aim is to prove that any point  $T_x = (f_1(x), \dots, f_J(x))$  verifies:

$$\overrightarrow{T_0 T_x} = \sum_{j=1}^J \lambda_j(x) \overrightarrow{T_0 T_j}, \text{ s.t. } \lambda_j(x) \geq 0, \forall j \in [1, j], \forall x \in [x_0, x_0 + h], \quad (2)$$

where, for  $j = 1, J$ ,  $\lambda_j(x)$  are real coefficients depending on  $x$  and  $T_j$  are the vertices of the osculating simplex.

For every  $0 < j < J$ ,  $\overrightarrow{T_0 T_j}$  belongs to the osculating  $j$ -space at  $T_0$ , and  $\overrightarrow{T_0 T_j}$  can be written:  $\overrightarrow{T_0 T_j} = \sum_{k=1}^j \gamma_{k,j} \mathbf{f}^{(k)}(x_0)$ . That is:

$$\begin{aligned} \overrightarrow{T_0 T_1} &= \gamma_{1,1} \mathbf{f}^{(1)}(x_0) \\ \overrightarrow{T_0 T_2} &= \gamma_{1,2} \mathbf{f}^{(1)}(x_0) + \gamma_{2,2} \mathbf{f}^{(2)}(x_0) \\ &\dots \end{aligned}$$

Gathering the coefficients  $\gamma_{k,j}$  for  $k \leq j$  in a matrix  $\mathbf{\Gamma}_h$ , we obtain the system of linear equations:

$$\begin{pmatrix} \overrightarrow{{}^t T_0 T_1} \\ \overrightarrow{{}^t T_0 T_2} \\ \dots \\ \overrightarrow{{}^t T_0 T_J} \end{pmatrix} = \mathbf{\Gamma}_h \begin{pmatrix} {}^t \mathbf{f}^{(1)}(x_0) \\ {}^t \mathbf{f}^{(2)}(x_0) \\ \dots \\ {}^t \mathbf{f}^{(J)}(x_0) \end{pmatrix}. \quad (3)$$

Furthermore, a Taylor expansion of  $\overrightarrow{T_0 T_x}$  gives

$$\overrightarrow{T_0 T_x} = \sum_{j=1}^J \frac{(x - x_0)^j}{j!} \mathbf{f}^{(j)}(x_0) + o((x - x_0)^J). \quad (4)$$

Plugging together equations (??), (??) and (??), we obtain:

$${}^t \mathbf{\Gamma}_h \begin{pmatrix} \lambda_1(x) \\ \dots \\ \lambda_J(x) \end{pmatrix} = \begin{pmatrix} \frac{(x - x_0)}{1!} + o((x - x_0)) \\ \dots \\ \frac{(x - x_0)^J}{J!} + o((x - x_0)^J) \end{pmatrix}. \quad (5)$$

The next step is to solve  $\mathbf{\Gamma}_h$ . With Lemma ??, we have

$$\mathbf{\Gamma}_h = \begin{pmatrix} \frac{h}{1!} + o(h) & 0 & \cdots & \cdots \\ \frac{h^2}{2!} + o(h^2) & \frac{h^2}{2!} + o(h^2) & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \frac{h^J}{J!} + o(h^J) & \cdots & \cdots & \frac{h^J}{J!} + o(h^J) \end{pmatrix}. \quad (6)$$

Equation (??) is rewritten:

$$\mathbf{\Gamma}_h \sim \mathbf{N}_h \mathbf{\Gamma}' (1 + o(1)). \quad (7)$$

The matrix  $\mathbf{N}_h = \begin{pmatrix} \frac{h}{1!} & & & \\ & \ddots & & \\ & & h^J & \\ & & & \frac{1}{J!} \end{pmatrix}$  is diagonal and  $\mathbf{\Gamma}' = \begin{pmatrix} 1 & 0 & \cdots & \cdots \\ 1 & 1 & 0 & \cdots \\ & & \ddots & 0 \\ 1 & 1 & \cdots & 1 \end{pmatrix}$  is lower triangular and does not depend on  $h$ .

Assembling equations (??) and (??), solving  $\mathbf{\Gamma}_h$  when  $h$  is small enough but not equal to zero, and finally simplifying gives:

$$\begin{pmatrix} \lambda_1(x) \\ \cdots \\ \lambda_J(x) \end{pmatrix} = \mathbf{N}_h^{-1} {}^t \mathbf{\Gamma}'^{-1} \begin{pmatrix} \frac{(x-x_0)}{1!} + o((x-x_0)) \\ \cdots \\ \frac{(x-x_0)^J}{J!} + o((x-x_0)^J) \end{pmatrix}.$$

We find an upper band matrix for  $\mathbf{N}_h^{-1} {}^t \mathbf{\Gamma}'^{-1}$ :

$$\mathbf{N}_h^{-1} {}^t \mathbf{\Gamma}'^{-1} = \begin{pmatrix} \frac{1!}{h} & -\frac{1!}{h} & 0 & \cdots & 0 \\ 0 & \frac{2!}{h^2} & -\frac{2!}{h^2} & \cdots & 0 \\ 0 & 0 & \frac{3!}{h^3} & \cdots & \cdots \\ \cdots & \cdots & \cdots & -\frac{(J-1)!}{h^{J-1}} & \cdots \\ 0 & \cdots & 0 & \frac{J!}{h^J} & \cdots \end{pmatrix}.$$

Eventually, we obtain a positive approximation for  $\lambda_j(x)$ ,  $1 \leq j \leq J$ , in the vicinity of  $x_0$ :

$$\lambda_j(x) = \frac{(x-x_0)^j}{(x_J-x_0)^j} \left(1 - \frac{1}{j+1}(x-x_0)\right) + o((x-x_0)^j).$$

□

**Theorem 1.** *Let  $T_0$  and  $T_J$  be two points on the curve  $C_J$ . Under Assumption A, the portion of the curve between  $T_0$  and  $T_J$  is included in its osculating simplex.*

*Proof.* We denote  $T_j$ , for  $0 < j < J$ , the vertex of the osculating simplex defined as the intersection of the osculating  $j$ -space at  $T_0$  and the osculating  $J - j$ -space at  $T_J$  (see Definition 2). We define  $\text{Face}_j$  for  $0 \leq j \leq J$  as the face of the osculating simplex containing all the vertices except  $T_j$ .

Each  $\text{Face}_j$ ,  $0 \leq j \leq J$ , intersects  $C_J$  exactly  $J$  times taking into account the multiplicities. Indeed, the multiplicity of the contact at  $T_0$  between  $C_J$  and  $\text{Face}_j$  is  $J$  since the osculating hyperplane at  $T_0$  is the supporting hyperplane of  $\text{Face}_j$  and  $\text{Face}_j$  contains the first  $J - 1$  derivatives  $\mathbf{f}^{(1)}(x_0), \dots, \mathbf{f}^{(J-1)}(x_0)$ . Thus  $T_0$  is the only contact point between  $C_J$  and  $\text{Face}_j$  (see Corollary A1). The same holds for  $T_J$  and  $\text{Face}_0$ .

For  $\text{Face}_j$ , for  $0 < j < J$ , by construction of the osculating simplex,  $T_0, T_1, \dots, T_{j-1}$  belong to the osculating  $j - 1$ -space at  $T_0$ . Thus the face  $T_0, T_1, \dots, T_{j-1}$  is supported by the vectorial sub-space spanned by the first  $j - 1$  derivatives at  $T_0$ . In the same way,  $T_{j-1}, \dots, T_{j-1}, T_J$  is included in the vectorial sub-space spanned by the first  $J - j - 1$  derivatives at  $T_J$ .

This amounts to saying that the multiplicity of the contact between  $C_J$  and  $\text{Face}_j$  at  $T_0$  is  $j$ . Similarly, the multiplicity of the contact between  $C_J$  and  $\text{Face}_j$  at  $T_J$  is  $J - j$ . Finally,  $C_J$  intersects  $\text{Face}_J$   $J$  times. Due to Corollary A1,  $T_0$  and  $T_J$  are the only intersection points between  $C_J$  and  $\text{Face}_J$ .

As a conclusion, between  $T_0$  and  $T_J$ ,  $C_J$  stays on one side of each of the faces  $\text{Face}_j$  for  $0 \leq j \leq J$ .

$S_J$  can be viewed as the cone generated by  $\{\overrightarrow{T_0T_1}, \dots, \overrightarrow{T_0T_J}\}$  sectioned by the face  $\text{Face}_0$ . From Lemma ??, when  $x_J$  is fixed, in a small neighborhood of  $x_0$ , we know that  $C_J$  is inside the cone. Since  $C_J$  never crosses one of the face  $\text{Face}_j$  except in  $T_0$  and  $T_J$ ,  $C_J$  remains inside this cone.  $\square$

**Preliminaries to Theorem 2** We recall the notations already introduced in the main paper, state two introductory propositions or lemmas, and then prove the theorem.

We consider  $P_{J,K}$  a set of nested simplexes, built so that  $P_{J,K+1} \subset P_{J,K}$ .

Let  $A_J$  be the set of coefficients for which  $\forall x \in [0, 1], F_\alpha(x) \geq 0$ :

$$A_J = \{\boldsymbol{\alpha} \mid \forall x \in [0, 1], F_\alpha(x) \geq 0\}.$$

Similarly, we denote  $A_{J,K}$  the set of possible coefficients at step  $K$ , that is the coefficients for which  ${}^t\mathbf{T}_{\bullet,K}\boldsymbol{\alpha} \geq 0$  where  $\mathbf{T}_{\bullet,K}$  is the matrix of constraints at step  $K$ . Removing the first row, we obtain  $\mathbf{T}_K$  the matrix gathering (in columns) the vertices of  $P_{J,K}$ .

$\tilde{\boldsymbol{\alpha}}_{J,K}$  is the vector of coefficients of the solution to Problem 5 when the constraints match the vertices of  $P_{J,K}$ . The coefficients of the optimal solution to 3 are stored in a vector denoted  $\tilde{\boldsymbol{\alpha}}_J$ .

Let  $\text{cost}(\boldsymbol{\alpha})$  be defined as  $\text{cost}(\boldsymbol{\alpha}) := \sum_{i=1}^I (Y_i - F_\alpha(X_i))^2$ . We have,

$$\text{cost}(\tilde{\boldsymbol{\alpha}}_{J,K}) = \min_{\boldsymbol{\alpha}}(\text{cost}(\boldsymbol{\alpha})), \text{ s.t. } {}^t\mathbf{T}_{\bullet,K}\boldsymbol{\alpha} \geq 0.$$

**Proposition 2.** 1.  $\forall K, A_{J,K} \subset A_{J,K+1} \subset A_J$ .

2.  $A_J$  and all the  $A_{J,K}$  are closed convex cones.

3. The sequence of  $\text{cost}(\tilde{\alpha}_{J,K})$  is decreasing with  $K$ .

*Proof.* **item 1** Thanks to Proposition 1,  $A_{J,K}$  can be seen as

$$A_{J,K} = \{\alpha \mid \forall t \in P_{J,K}, \langle \alpha, t \rangle \geq 0\}.$$

By construction,  $P_{J,K+1} \subset P_{J,K}$ . Indeed, each simplex of  $P_{J,K+1}$  results from cutting in two one of the simplexes in  $P_{J,K}$ , as illustrated on Figure 2.

Thus, if we have  $\langle T, \alpha \rangle \geq 0$  for all the vertices  $T$  of  $P_{J,K}$ , then it is also true for all the vertices of  $P_{J,K+1}$ . This last statement means that  $A_{J,K} \subset A_{J,K+1}$ .

$P_{J,K}$  is a collection of successive osculating simplexes, each of them finishing at the point where the next one begins. Thus  $P_{J,K}$  circumscribes the curve  $C_J$ , and this implies that if  $\alpha$  is in  $A_{J,K}$  then  $\forall x \in [0, 1], F_\alpha(x) \geq 0$ , or equivalently that  $A_{J,K} \subset A_J$ .

**item 2** We only detail this claim for  $A_J$ , similar considerations can be applied to the  $A_{J,K}$ . Indeed, if  $\forall x F(x) \geq 0$  for a given  $\alpha$ , then it is also verified for  $\lambda\alpha$  where  $\lambda$  is real and positive. Thus  $A_J$  is a cone. It is convex: if  $F(x) \geq 0$  for  $\alpha_1$  and  $\alpha_2$ , then it is also non-negative for  $p\alpha_1 + (1-p)\alpha_2$  for any  $p \in [0, 1]$ .

The set  $B = \{\alpha \mid \forall x \in [0, 1], F_\alpha(x) \geq 0\}$  is closed: we consider the application  $g_x$  defined as  $\alpha \in \mathbb{R}^{J+1} \mapsto g_x(\alpha) \in \mathbb{R}, g_x(\alpha) = \langle \mathbf{f}_\bullet(x), \alpha \rangle$ .  $g_x$  is continuous. The inverse image of the open set  $\mathbb{R}^{-*} = (-\infty, 0), g_x^-(\mathbb{R}^{-*})$ , is then open and  $C = \bigcup_{x \in [0,1]} g_x^-(\mathbb{R}^{-*})$  is also

open.  $C$  being the complement of  $B$  in  $\mathbb{R}^J$  is closed.

**item 3** This is a direct consequence of item 1: since  $A_{J,K} \subset A_{J,K+1}$ , the minimum over  $A_{J,K}$  is greater or equal to the minimum over  $A_{J,K+1}$ .  $\square$

We restrict our attention to the sequence  $P_{J,K}$  built as a chain of simplexes  $S_k$  starting at  $x = (k-1)/2^K$  and finishing at  $x = k/2^K$  with  $k$  varying from 1 to  $2^K$ . We first observe that the distance from any point of  $P_{J,K}$  to the curve  $C_J$  can be made as small as needed: more precisely,

**Lemma B4.**  $\forall \epsilon > 0, \exists K \in \mathbb{N}$  such that  $\forall u \in P_{J,K}, \exists t \in C_J$  for which  $\|u - t\| < \epsilon$ .

*Proof.* To prove this claim, we choose  $u$  in  $P_{J,K}$ , and we restrict our attention to the simplex  $S_k$  containing  $u$ . The maximum distance of two points within  $S_k$  is one of the distances between two of its vertices. By means of Equation (??) and Lemma ??, when  $K$  is sufficiently large, calling  $T_{j_1}$  and  $T_{j_2}$  two of the vertices of  $S_k$ , the vector  $\overrightarrow{T_{j_1}T_{j_2}}$  is approximated by

$$\overrightarrow{T_{j_1}T_{j_2}} \sim \sum_{l=j_1+1}^{j_2} \frac{h^l}{l!} \mathbf{f}^{(l)}(x_0) + o(h^{j_1+1}).$$

$\|T_{j_1}T_{j_2}\|$  and then  $\|u - t\|$  are bounded from above by  $\frac{M}{2^K}$  with  $M = \sum_{j=1}^J \sup_{x \in [0,1]} \|\mathbf{f}^{(j)}(x)\|$ .  $\square$

**Theorem 2.** *Under Assumption A, we have  $\lim_{K \rightarrow \infty} \tilde{\alpha}_{J,K} = \tilde{\alpha}_J$ .*

*Proof.* We denote  $B = \overline{\bigcup_{K \in \mathbb{N}} A_{J,K}}$ . Our goal is first to prove that  $B = A_J$ , or in other words that the sequence of sets  $\bigcup_{K \in \mathbb{N}} A_{J,k}$  is dense in  $A_J$ .

The inclusion  $B \subset A_J$  is immediate, as a consequence of items 1 and 2 of Proposition 2. Conversely, we have to prove that every point of  $A_J$  is attained. We choose  $\alpha$  in  $A_J$  and want to show that  $\alpha \in \overline{\bigcup_{K \in \mathbb{N}} A_{J,K}}$ .

Starting from the vector  $\alpha \in A_J$ ,  $\alpha = {}^t(\alpha_0, \alpha_1, \dots, \alpha_J)$ , for any positive integer  $l$  we define  $\alpha_l$  as  $\alpha_l = {}^t(\alpha_0 + \frac{1}{l}, \alpha_1, \dots, \alpha_J)$ .  $\alpha_l$  belongs to  $A_J$ :

$$F_{\alpha_l}(x) \geq F_{\alpha}(x) + \frac{1}{l} > 0.$$

If we exhibit now an index  $K_l$  for which  $\alpha_l$  simultaneously belongs to  $A_{J,K_l}$ , our assertion is proved:  $\alpha$  will be the limit of a sequence of  $\alpha_l$  each of them taken in one  $A_{J,K_l}$ .

The way to achieve this goal is to consider the sequence  $P_{J,K_l}$  of Lemma ???.  $P_{J,K_l}$  is built as a chain of simplexes  $S_k$  for  $k$  varying from 1 to  $2^{K_l}$ . Picking a point  $u$  in  $P_{J,K_l}$  we examine now what is the condition for which  $\langle u, \alpha_l \rangle > 0$ .

We start from the identity

$$\langle u, \alpha_l \rangle = \langle u - t, \alpha_l \rangle + \langle t, \alpha_l \rangle.$$

- We observe that  $\langle t, \alpha_l \rangle > 1/l$ .
- By Cauchy-Schwartz inequality, using Lemma ??,  $\langle u - t, \alpha_l \rangle \geq -\frac{M}{2^{K_l}} \|\alpha_l\|$ .
- By the triangular inequality,  $\|\alpha_l\| \leq \frac{1}{l} + \|\alpha\|$ .

Eventually,

$$\langle u, \alpha_l \rangle \geq -\frac{M}{2^{K_l}} \left( \frac{1}{l} + \|\alpha\| \right) + \frac{1}{l}.$$

For a given  $l$ ,  $K_l$  is chosen so that the right part of the previous inequality be positive.

Since it is true for any  $u \in S_k$  and for any  $k$ ,  $\alpha_l \in A_{J,K}$ , which permits to conclude that  $B = A_J$ .

Thus,  $\tilde{\alpha}_J$  the optimal solution to Problem 3, as an element of  $A_J$ , is the limit of a sequence of vectors  $\alpha_{J,K}$ , each of them taken in one  $A_{J,K}$ . The second step is to extend this first result to the sequence of  $\tilde{\alpha}_{J,K}$ , the solutions to Problem 5.

As  $A_{J,K} \subset A_J$ , we have

$$\sum_{i=1}^I (Y_i - F_{\alpha_{J,K}}(X_i))^2 \geq \sum_{i=1}^I (Y_i - F_{\tilde{\alpha}_{J,K}}(X_i))^2 \geq \sum_{i=1}^I (Y_i - F_{\tilde{\alpha}_J}(X_i))^2.$$

This proves that  $\text{cost}(\alpha_{J,K})$  converges toward  $\text{cost}(\tilde{\alpha}_J)$ .

The function cost is convex. Calling  $\mathbf{X}$  the matrix of the model,  $\mathbf{X} = \begin{pmatrix} \mathbf{f}_\bullet(X_1) \\ \cdots \\ \mathbf{f}_\bullet(X_n) \end{pmatrix}$ , then the hessian matrix of the function cost is simply  ${}^t\mathbf{X}\mathbf{X}$ . Assuming that  ${}^t\mathbf{X}\mathbf{X}$  is definite positive, which is the usual assumption in regression problems, we can infer that  $\tilde{\alpha}_J$  is also the limit of the sequence of the solutions  $\tilde{\alpha}_{J,K}$  of Problem 5. □

**Proposition 3.** *Let  $T_0, T, T_J$  be three points on the curve corresponding to  $x_0 < x < x_J$ . Then the function  $V_{\text{new}} = V(x_0, x) + V(x, x_J)$  has a unique minimum between  $x_0$  and  $x_J$ , where  $V(x_0, x)$  (resp.  $V(x, x_J)$ ) stands for the volume of the simplex between  $x_0$  and  $x$  (resp.  $x$  and  $x_J$ ).*

*Proof.* When cutting the initial simplex at  $x$  the volume of the two new simplexes replacing the old one becomes:  $V_{\text{new}} = V(x_0, x) + V(x, x_J)$ . If  $x = x_0$  or  $x = x_J$  then  $V_{\text{new}} = V(x_0, x_J)$  and is maximum. Due to Rolle's theorem, there exists a  $x$  for which  $V_{\text{new}}$  is minimum. This minimum is unique since by construction  $V(x_0, x)$  is a strictly increasing function while  $V(x, x_J)$  is strictly decreasing. □

**Proposition 4.**  $V(x_0, x_J) = \frac{1}{J!} D_{J,J} \frac{\prod_{j=1}^{J-1} D_{j,j}}{\prod_{j=1}^{J-1} D_j}$ .

*Proof.* Indeed, the volume of a simplex with vertexes  $T_0, \dots, T_J$  is known to be:

$$V(x_0, x_J) = \frac{1}{J!} |\overrightarrow{T_0T_1} \cdots \overrightarrow{T_0T_J}|.$$

Taking the notation of Lemma ??, for  $j < J$ ,  $\overrightarrow{T_0T_j}$  is decomposed in

$$\overrightarrow{T_0T_j} = \sum_{k=1}^j \frac{D_{j,k}}{D_j} \mathbf{f}^{(k)}(x_0).$$

Standard manipulations on determinants give the expected result. □

**Proposition 5.**  $V(x_0, x_J) \propto (x_J - x_0)^{\frac{J(J+1)}{2}}$ .

*Proof.* Restarting from Equation (??), when  $f_j(x) = x^j$ , the coefficients  $\gamma_{k,j}$  of Lemma ?? become exactly  $\gamma_{k,j} = \frac{(x_J - x_0)^k}{k!}$ .

Recalling that

$$\begin{aligned}\overrightarrow{T_0 T_1} &= \gamma_{1,1} \mathbf{f}^{(1)}(x_0) \\ \overrightarrow{T_0 T_2} &= \gamma_{1,2} \mathbf{f}^{(1)}(x_0) + \gamma_{2,2} \mathbf{f}^{(2)}(x_0) \\ &\dots\end{aligned}$$

we see that  $V(x_0, x_J) = |\overrightarrow{T_0 T_1}, \dots, \overrightarrow{T_0 T_J}| = \prod_{j=1}^J \gamma_{j,j} = \frac{(x_J - x_0)^{J(J+1)/2}}{\prod_{j=1}^J j!}$ , which is precisely what we had to establish.  $\square$

**Theorem 3.** *Let the system of functions  $f_j(x)$  be the sequence of monomials  $\{x^j\}_{j=1}^J$ . Then the optimal cut point between  $x_0$  and  $x_J$  is  $\frac{x_0 + x_J}{2}$ .*

*Proof.* Let  $x$  be the parameter of the cut point. From Lemma 3,  $x$  minimizes  $V(x_0, x) + V(x, x_J)$ . Applying Proposition 5, it amounts to find the minimum of  $(x - x_0)^{J(J+1)/2} + (x_J - x)^{J(J+1)/2}$ , which is obviously obtained when  $x = \frac{x_0 + x_J}{2}$ .  $\square$

**Theorem 4.** *Let  $D$  be  $D = [x_{1,0}, x_{1,1}] \times \dots \times [x_{V,0}, x_{V,1}]$ . Under Assumption A*

1. *When  $x$  traverses  $D$ , the corresponding portion of  $C_J$  is included in  $P_J$ .*
2. *If  ${}^t \boldsymbol{\alpha} \mathbf{T}_\bullet \geq 0$ , then  $\forall x \in D$ , we have  $F(x) \geq 0$ .*

*Proof.* We only have to prove item 1. Item 2 is immediate since  $P_J$  is convex by construction.

We recall that  $F_\alpha(x)$  can be expressed by  $F_\alpha(x) = \langle \boldsymbol{\alpha}, \mathbf{f}_\bullet(x) \rangle$ , where  $\mathbf{f}_\bullet(x)$  results from the tensor product

$$\mathbf{f}_\bullet(x) = \mathbf{f}_{1\bullet}(x_1) \otimes \dots \otimes \mathbf{f}_{V\bullet}(x_V). \quad (8)$$

This tensor product gives  $J + 1$  terms. We rewrite  $\mathbf{f}_\bullet(x)$  as

$$\mathbf{f}_\bullet(x) = {}^t (f_0(x), f_1(x), \dots, f_J(x)).$$

$C_J$  is described by  $C_J = (f_1(x), \dots, f_J(x))$ , when  $x$  traverses  $[0, 1]^V$ .

Given a point  $T_{x^*} = {}^t (f_1(x^*), \dots, f_J(x^*))$  of  $C_J$  corresponding to the values  $x^* = (x_1^*, \dots, x_V^*)$  of the variables, we have to show that  $T_{x^*} \in P_J$ .

We call  $T_j$  the vertices of  $P_J$ . By construction, each vertex  $T_j$  can be extracted from the column number  $j$  of the matrix of constraints  $\mathbf{T}_\bullet$  after removing the first coordinate, equal to 1.



Our aim is to exhibit  $J + 1$  non negative coefficients  $\mu_j$ , for  $j = 0, J$ , summing to 1, such that

$$T_{x^*} = \sum_{j=0}^J \mu_j T_j.$$

This equation can be extended to the columns of  $\mathbf{T}_\bullet$ , and is equivalent to

$$\begin{pmatrix} 1 \\ T_{x^*} \end{pmatrix} = \sum_{j=0}^J \mu_j \begin{pmatrix} 1 \\ T_j \end{pmatrix}.$$

For each  $x_v$ , we consider the curve  $C_{v,J_v}$  described by  $(f_{v,1}(x_v), \dots, f_{v,J_v}(x_v))$  and the point  $T_{v,x_v^*}$  corresponding to the value  $x_v^*$  of the variable  $x_v$ .

$C_{v,J_v}$  is included in its osculating simplex. Then we can find  $J_v + 1$  positive coefficients  $\lambda_{v,j_v}$  summing to 1 such that:

$$T_{v,x_v^*} = \sum_{j_v=0}^{J_v} \lambda_{v,j_v} T_{v,j_v},$$

or equivalently

$$\begin{pmatrix} 1 \\ T_{v,x_v^*} \end{pmatrix} = \sum_{j_v=0}^{J_v} \lambda_{v,j_v} \begin{pmatrix} 1 \\ T_{v,j_v} \end{pmatrix}. \quad (9)$$

Stemming from Equation ??, by means of the tensor product, we have:

$$\begin{pmatrix} 1 \\ T_{x^*} \end{pmatrix} = \begin{pmatrix} 1 \\ T_{1,x_1^*} \end{pmatrix} \otimes \dots \otimes \begin{pmatrix} 1 \\ T_{V,x_V^*} \end{pmatrix}. \quad (10)$$

The combination of equations ?? and ?? gives:

$$\begin{pmatrix} 1 \\ T_{x^*} \end{pmatrix} = \sum_{j_1=0}^{J_1} \dots \sum_{j_V=0}^{J_V} \lambda_{1,j_1} \dots \lambda_{V,j_V} \begin{pmatrix} 1 \\ T_{1,j_1} \end{pmatrix} \otimes \dots \otimes \begin{pmatrix} 1 \\ T_{v,j_V} \end{pmatrix},$$

which leads to the desired expression for  $T_{x^*}$  after removing the first row.

Furthermore, one can observe that:

$$\sum_{j_1=0}^{J_1} \dots \sum_{j_V=0}^{J_V} \lambda_{1,j_1} \dots \lambda_{V,j_V} = \prod_{v=1}^V (\lambda_{v,0} + \dots + \lambda_{v,J_v}) = 1.$$

Thus  $T_{x^*}$  is expressed as a linear combination of the vertices of  $P_J$ , where all the coefficients are positive and sum to 1. The proof of item 1 is complete.  $\square$

**Theorem 5.** *We assume that each family of functions  $\{f_{v,j}(x_v)\}_{j=1}^{J_v}$  for  $1 \leq v \leq V$  verifies Assumption A. Then  $\lim_{K \rightarrow \infty} \tilde{\alpha}_{J,K} = \tilde{\alpha}_J$ .*

*Proof.* Theorem 5 is the analog of Theorem 2 for a single variable. Looking closely to the proof of Theorem 2, we can see that it can be readily generalized without any change to the case of more than one variable, except for the maximum distance between two vertices of any simplex.

Indeed, calling  $T_{j_1}$  and  $T_{j_2}$  two of the vertices of one of the simplexes in the univariate case, the vector  $\overrightarrow{T_{j_1}T_{j_2}}$  has been shown to be bounded by  $\frac{M}{2^{K_l}}$  with  $M = \sum_{j=1}^J \sup_{x \in [0,1]} \|\mathbf{f}^{(j)}(x)\|$ .

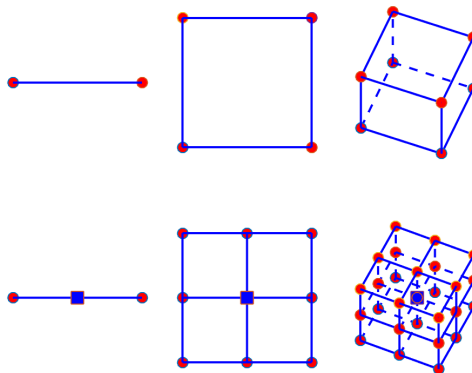
To generalize to the multivariate case, due to the tensorial product,  $\frac{M}{2^{K_l}}$  must be replaced by  $\prod_{v=1}^V \frac{M_v}{2^{K_l}}$ , where each  $M_v$  is taken to be

$$M_v = \sum_{j_v=1}^{J_v} \sup_{x_v \in [0,1]} \|\mathbf{f}_{\mathbf{v}}^{(j_v)}(x_v)\|.$$

□

**Proposition 6.** *When creating a new simplex by subdividing an existing one, the number of constraints is augmented by:*

$$3^V - 2^{2V} + (J + 1) * (2^V - 1).$$



*Proof.*

Figure 1: **number of constraints** the upper row gives the limits of an initial domain ( $[0, 1]^V$  for example) when  $V = 1, 2, 3$ . The lower row gives the new definition domain when a point is added in the previous lattice. For example, when  $V = 3$  each cube corresponds to a simplex.

When adding a point in the center of an initial domain (see figure ??)

- we replace the vertices on the external border: we add  $3^V$  vertices and remove  $2^V$  old ones.
- for the  $2^V$  new simplexes, we add  $2^V(J + 1 - 2^V)$  interior points and remove  $J + 1 - 2^V$  points corresponding to the interior vertexes of the old simplex.

This gives the expected result.

□