KRIGING-BASED INTERPOLATORY SUBDIVISION SCHEMES

J. BACCOU * AND J. LIANDRAT †

Abstract. This paper is devoted to the definition, analysis and implementation of a new type of subdivision schemes adapted to data (through a stochastic approach) and to a partition of their support. Its construction combines position-dependent multi-scale approximation ([7]) and Kriging theory ([12]). After a full convergence analysis that requires to extend classical results to this new framework, it is applied to data prediction for uni and bi-variate problems and compared to the Lagrange interpolatory subdivision.

Key words. Kriging, subdivision scheme, non-regular data prediction

AMS subject classifications. 65, 62

1. Introduction. The work presented in this paper lies in the intersection of two fields, stochastic data modeling and multi-scale approximation. It is devoted to the reconstruction of functions, starting from an initial sampling at low rate, with a special attention to piecewise-regular ones that arise in many applications. The main advantages of stochastic data modeling such as Kriging ([12]) stand in the possibility to integrate in the prediction the spatial dependence of the available data and to quantify the precision of the prediction thanks to the model describing this dependence. The counterpart is that the determination of this model is classically performed assuming large enough regularity of the data. Even if a so called nugget effect can be considered to take into account strong local variations and therefore reduce somehow the required regularity, Kriging approach is not tailored to reconstruct piecewise-smooth data. On the contrary, multi-scale approximation and associated subdivision schemes ([14]), even if they are purely deterministic, can be data-dependent, a property that is compulsory to maintain the efficiency in presence of discontinuities. This property however makes the analysis of convergence, smoothness and stability of these schemes more involved ([11]).

This paper therefore deals with the design of new stochastic subdivision schemes that improve the accuracy of the reconstruction of non-regular data. Their construction is performed in two steps: a segmentation of data into different zones of smoothness and a local subdivision based on Kriging prediction according to the information coming from the previous step. Taking apart the problem of data segmentation that refers to the wide literature of edge detector, we focus in the sequel on the construction and on the convergence analysis

*IRSN, Centre de Cadarache, 13115 Saint Paul Les Durance, France, (jean.baccou@irsn.fr).

† Centrale Marseille and LATP, 13451, Marseille, France (jliandrat@centrale-marseille.fr).

of position-dependent Kriging-based subdivision schemes for data prediction i.e. of subdivision schemes derived from Kriging, involving a local strategy depending on a given set of segmentation points and involving an adaptive selection of interpolatory stencils.
Our work is organized as follows: after a brief review of the background related to the Harten's framework for subdivision (Section 2) and to the Kriging theory (Section 3.1) a position-dependent Kriging-based prediction operator is defined in Section 3.2; its convergence is then fully analyzed in Section 4. Numerical applications and a comparison with the position-dependent Lagrange interpolatory subdivision scheme described in [7] are finally presented in Section 5.

2. Harten's framework and subdivision schemes. In this section we provide a short review on Harten's framework including subdivision schemes with emphasis on position-dependent strategy that appears to be suitable for the prediction of piecewise-smooth data.

2.1. Harten's framework. Harten's framework for multi-resolution has been introduced in [5] or [6].

The general setting of Harten's framework is defined by a family of triplets $(V^j, D_j^{j-1}, P_{j-1}^j)$ where $j \in \mathbb{Z}$ is a scale parameter, V^j denotes a separable space of approximation associated to a resolution level 2^{-j}, D_j^{j-1} (resp. P_{j-1}^j) is a decimation (resp. prediction) operator connecting V^j to V^{j-1} (resp. V^{j-1} to V^j); Any element of V^j is a sequence $f^j = \{f_k^j\}_{k\in\mathbb{Z}}$ that stands for a discrete approximation, at scale 2^{-j}, of a function belonging to a suitable space F.

Here, we restrict ourself to sampling, i.e. the sequences f^j coincide with the values of functions on the dyadic grid $\{x_k^j\}_{k\in\mathbb{Z}}$ with $x_k^j = k2^{-j}$. The embedding property linking V^j and V^{j-1} implies here that the decimation operator D_j^{j-1} is a sub-sampling operator. Any right inverse of D_j^{j-1} in V^{j-1} is a suitable prediction operator[1]. Interpolatory subdivision schemes described in [14], are therefore good candidates to define such an operator. They map $\{f_k^{j-1}\}_{k\in\mathbb{Z}}$ to $P_{j-1}^j(f^{j-1}) = f'^j = \{\left(Sf^{j-1}\right)_k\}_{k\in\mathbb{Z}}$ with

$$\left\{ \begin{array}{lcl} f'^j_{2k} & = & f_k^{j-1}, \\ f'^j_{2k-1} & = & \sum_{m=-l}^{r-1} a_{k-2m} f_{k+m}^{j-1}, \end{array} \right. \tag{2.1}$$

where the sequence $\{a_{k-2m}\}_{m\in\mathbb{Z}}$ is called the mask of the subdivision scheme S. The integers l and r stand for the number of left and right points in the interpolatory stencil denoted $S_{l,r}$.
For the treatment of piecewise-smooth functions (that is our concern in this work), we introduce more flexibility in the choice of l and r. More precisely, for any fixed integer $D \geq 0$, we associate to each prediction at position $(2k-1)2^{-j}$

[1] P_{j-1}^j and D_j^{j-1} satisfy $D_j^{j-1}P_{j-1}^j = I_{V^{j-1}}$

a stencil of parameters $(l_{j,2k-1}, r_{j,2k-1})$ with $l_{j,2k-1} + r_{j,2k-1} = D + 1$. This leads to a sequence of masks $\{a^{j,k}_{k-2m}\}_{m\in\mathbb{Z}}$, $(j,k) \in \mathbb{Z}^2$ for the corresponding subdivision scheme. If these masks are independent of f^{j-1}, the subdivision is said to be linear. Moreover, one speaks of stationarity (resp. uniformity) when the masks do not depend on j (resp. on k).

There are many strategies to define the couples $(l_{j,2k-1}, r_{j,2k-1})$ ([13], [3], [11], [7]). Among them, one can mention the two following ones that will be considered in this paper.
The first one is $l_{j,2k-1} = l$, $r_{j,2k-1} = r$ that leads to **translation-invariant stencils**. The corresponding subdivision schemes are then stationary and uniform (and thus linear).

In the second strategy, $(l_{j,2k-1}, r_{j,2k-1})$ depends on an a priori defined segmentation of the real line and on a chosen couple (l, r). More precisely, the couple $(l_{j,2k-1}, r_{j,2k-1})$ is equal to (l, r) except in some regions; these regions and the values of $(l_{j,2k-1}, r_{j,2k-1})$ in these regions are obtained according to a specific rule involving the segmentation and the position. This leads to **position-dependent stencils** and defines a linear but non-stationary and non-uniform subdivision scheme.

The family of ENO interpolation subdivision schemes addressed in [11] corresponds to a third strategy where the definition of $(l_{j,2k-1}, r_{j,2k-1})$ depends on the sequence f^{j-1} and more precisely on its values f^{j-1}_l for index l in the vicinity of k.

The selection of position-dependent stencils is fully explained in the following section.

2.2. Position-dependent stencils. The position-dependent strategy we are interested in is motivated by the approximation of piecewise-smooth functions. This class of functions is defined as follows:

Definition 2.1.
We call $\mathcal{C}(\mathbb{R}\backslash\mathcal{S})$ the space of functions f defined and continuous on $\mathbb{R}\backslash\mathcal{S}_f$ where $\mathcal{S}_f$, when it is non-empty, is a set of reals (depending of f) such that $\forall x \in \mathcal{S}_f, \lim_{y\to x, y>x} f(y) = f^+(x)$ and $\lim_{y\to x, y<x} f(y) = f^-(x)$ exist in $\mathbb{R}$ with $f^+(x) \neq f^-(x)$.
The set $\mathcal{S}_f$ is moreover assumed to be separated i.e. to satisfy

$$\exists \epsilon_{\mathcal{S}_f} > 0 \text{ such that } \forall x, y \in \mathcal{S}_f,\ x \neq y \Rightarrow |x - y| \geq \epsilon_{\mathcal{S}_f}.$$

Accordingly, position-dependent stencils rely on an adaptation of the stencil according to a separated family of segmentation points, $\mathcal{S}_s$, used to define the different expressions of the scheme. Since the segmentation points are supposed to be separated, there exists a scale from which each segmentation point

can be considered as isolated with regards to the stencil length. Therefore we focus in the following definition (proposed in [7]) on a rule involving a single segmentation point y_0 corresponding to $\mathcal{S}_s = \{y_0\}$.

DEFINITION 2.2. *Stencil selection*

Let us define for all $j \in \mathbb{Z}$, the index k_{j-1} s.t $y_0 \in [x^{j-1}_{k_{j-1}-1}, x^{j-1}_{k_{j-1}}]$. For all j and k such that $y_0 \in [x^{j-1}_{-l+k}, x^{j-1}_{r-1+k}]$, the parameters $r_{j,2k-1}$ and $l_{j,2k-1}$ linked by the relation $r_{j,2k-1} + l_{j,2k-1} = D + 1$ are defined as:

- *If $y_0 \in [x^j_{2k_{j-1}-2}, x^j_{2k_{j-1}-1}[$, then*

$$\begin{cases} \text{If } k < k_{j-1} \text{ then } r_{j,2k-1} = k_{j-1} - k, \\ \text{If } k \geq k_{j-1} \text{ then } r_{j,2k-1} = D + 1 + k_{j-1} - k. \end{cases} \tag{2.2}$$

- *If $y_0 \in [x^j_{2k_{j-1}-1}, x^j_{2k_{j-1}}]$, then*

$$\begin{cases} \text{If } k \leq k_{j-1} \text{ then } r_{j,2k-1} = k_{j-1} - k, \\ \text{If } k > k_{j-1} \text{ then } r_{j,2k-1} = D + 1 + k_{j-1} - k. \end{cases} \tag{2.3}$$

Figure 2.1 displays this selection rule when $D = 3$, $l = 2$ and $r = 2$.

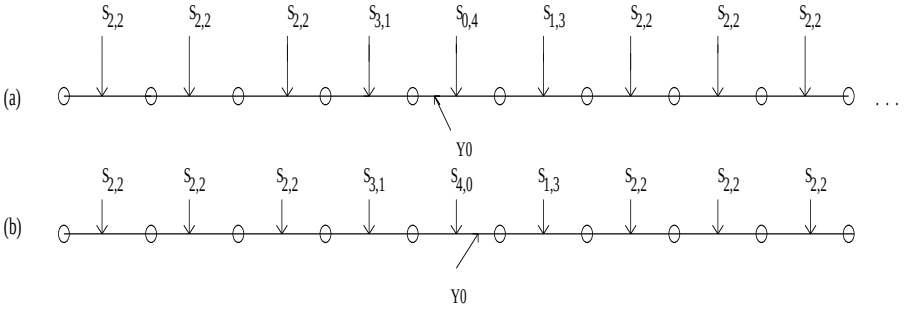

FIG. 2.1. *An example of stencil selection associated to a unique segmentation point $y_0 \in [x^{j-1}_{k_{j-1}-1}, x^{j-1}_{k_{j-1}}]$. (a) $y_0 \in [x^j_{2k_{j-1}-2}, x^j_{2k_{j-1}-1}]$, (b) $y_0 \in [x^j_{2k_{j-1}-1}, x^j_{2k_{j-1}}]$.*

In [7], a first example of Lagrange position-dependent subdivision scheme has been constructed and fully analyzed. It is based on a local polynomial interpolation. More precisely, the sequences $\{a^{j,k}_{k-2m}\}_{m \in \{-l_{j,2k-1},..,r_{j,2k-1}-1\}}$ satisfy:

$$a^{j,k}_{k-2m} = L^{l_{j,2k-1}, r_{j,2k-1}}_m(-\frac{1}{2})$$

where

$$L^{l_{j,2k-1}, r_{j,2k-1}}_m(x) = \Pi^{r_{j,2k-1}-1}_{n=-l_{j,2k-1}, n \neq m} \frac{x-n}{m-n}, \tag{2.4}$$

This family of schemes has provided interesting results for image compression (see [4] for example). However, the computation of the mask does not depend on the data to predict. As a result, it does not integrate their spatial structure, if any, which is not fully satisfactory in many applications. Moreover, for safety assessment studies for instance, the lack of quantification of a prediction error, that ensures the transparency in the communication of the results, can reduce its interest for practical applications. Therefore, in order to circumvent these drawbacks, we propose in the next section a new type of subdivision scheme that combines multi-scale approach and a stochastic model based on Kriging theory.

3. Kriging theory and Kriging-based subdivision. For sake of clarity, the ordinary Kriging interpolation is first described independently of the multi-scale/subdivision framework. We postpone to Section 3.2 the description of the way to combine them in order to construct Kriging-based subdivision schemes.

3.1. Ordinary Kriging. Considering $\{f_i\}_{i=0,\dots,N-1}$, the set of N observations of a function, $f \in F$, on the grid $\{x_i\}_{i=0,..,N-1}$ we address the problem of the prediction of f at a new location $x^\star$. Ordinary Kriging ([12]) belongs to the class of stochastic prediction methods. It is assumed that $\{f_i\}_{i\in\mathbb{Z}}$ are realizations of a subset of random variables $\{\mathcal{F}(x_i), i \in \mathbb{Z}\}$ coming from a random process $\mathcal{F}(x)$ that can be decomposed as:

$$\mathcal{F}(x) = m + \delta(x), \tag{3.1}$$

with m the constant deterministic mean structure of $\mathcal{F}(x)$ and $\delta(x)$ an intrinsically stationary random process, i.e. satisfying:

$$\begin{cases} E\left(\delta(x+h) - \delta(x)\right) = 0, \\ var\left(\delta(x+h) - \delta(x)\right) = 2\gamma(h), \end{cases}$$

where E and var denote the mathematical expectation and the variance respectively[2]. The semi-variogram γ stands for the spatial dependence associated to $\delta(x)$.

This last quantity plays a key role in the prediction of $\mathcal{F}(x)$ since, according to Decomposition (3.1) and the assumption of constant mean structure, it also characterizes the spatial dependence associated to $\mathcal{F}(x)$. Therefore, in the sequel, we first focus on the identification of the semi-variogram γ. It is then used to construct $\mathcal{P}(\mathcal{F}, x^*)$, the ordinary Kriging estimator of $\mathcal{F}(x^\star)$ as a linear combination of $\{\mathcal{F}(x_i)\}_{i=0,\dots,N-1}$. Replacing $\mathcal{F}(x_i)$ by f_i, this estimator leads to an approximation of $f(x^\star)$.

[2]Let X be a real random variable of density f_X. Its mathematical expectation is defined as the integral over the realizations x of X weighted by the density function i.e. $E(X) = \int_{I\!R} x f_X(x) dx$. Moreover, its variance is given by $var(X) = E(X^2) - (E(X))^2$.

3.1.1. Spatial structure identification. Under stationarity assumptions and constant deterministic mean structure of the random process, the semi-variogram is written as:

$$\gamma(h) = \frac{1}{2}E((\mathcal{F}(x+h) - \mathcal{F}(x))^2), \tag{3.2}$$

It is approximated, from the available data, by a least square fit of the discrete experimental semi-variogram,

$$\gamma_{exp}(h) = \frac{1}{2Card(N(h))} \sum_{(m,n)\in N(h)} (f_m - f_n)^2, \tag{3.3}$$

with $N(h) = \{(m,n) \in \{0,...,N-1\},\ |x_m - x_n| = h\}$ and for every h such that $Card(N(h))$, the cardinality of $N(h)$, is sufficiently large. In practice, the number of pairs (x_m, x_n) satisfying $|x_m - x_n| = h$ is usually not sufficient to accurately estimate the semi-variogram. Therefore, the condition $|x_m - x_n| = h$ in $N(h)$ is replaced by $h - \epsilon \leq |x_m - x_n| \leq h + \epsilon$ where ϵ is problem and user-dependent.
The candidates for the experimental semi-variogram fitting have to be chosen among a family of valid semi-variogram models (Figure 3.1), denoted $\mathcal{G}$ (see [12] or [23]).

For the rest of the paper, we focus on the two following families of semi-variograms (that differ according to the behavior of $\gamma(h)$ when h tends to zero) that encompass most of the classical models used in practice.

Definition 3.1.
We introduce $\mathcal{G}^\star_{OD}$ (resp. $\mathcal{G}^\star_{ED}$), the subsets of valid semi-variograms satisfying:

$$\mathcal{G}^\star_{ED} = \{\gamma \in \mathcal{G},\ \gamma(h) =_{h\to 0} \sum_{n=0}^{+\infty} a_n h^{n+1} \text{ with } a_0 \neq 0\}, \tag{3.4}$$

$$\mathcal{G}^\star_{OD} = \{\gamma \in \mathcal{G},\ \gamma(h) =_{h\to 0} \sum_{n=0}^{+\infty} b_n h^{2n+2} \text{ with } \forall n \in I\!N,\ b_n \neq 0\}. \tag{3.5}$$

Note that the linear, spherical and exponential semi-variograms belong to $\mathcal{G}^\star_{ED}$ while the Gaussian, rational quadratic and hole-effect ones belong to $\mathcal{G}^\star_{OD}$ (see Figure 3.1).

Figure 3.2 displays the different steps of the semi-variogram identification.

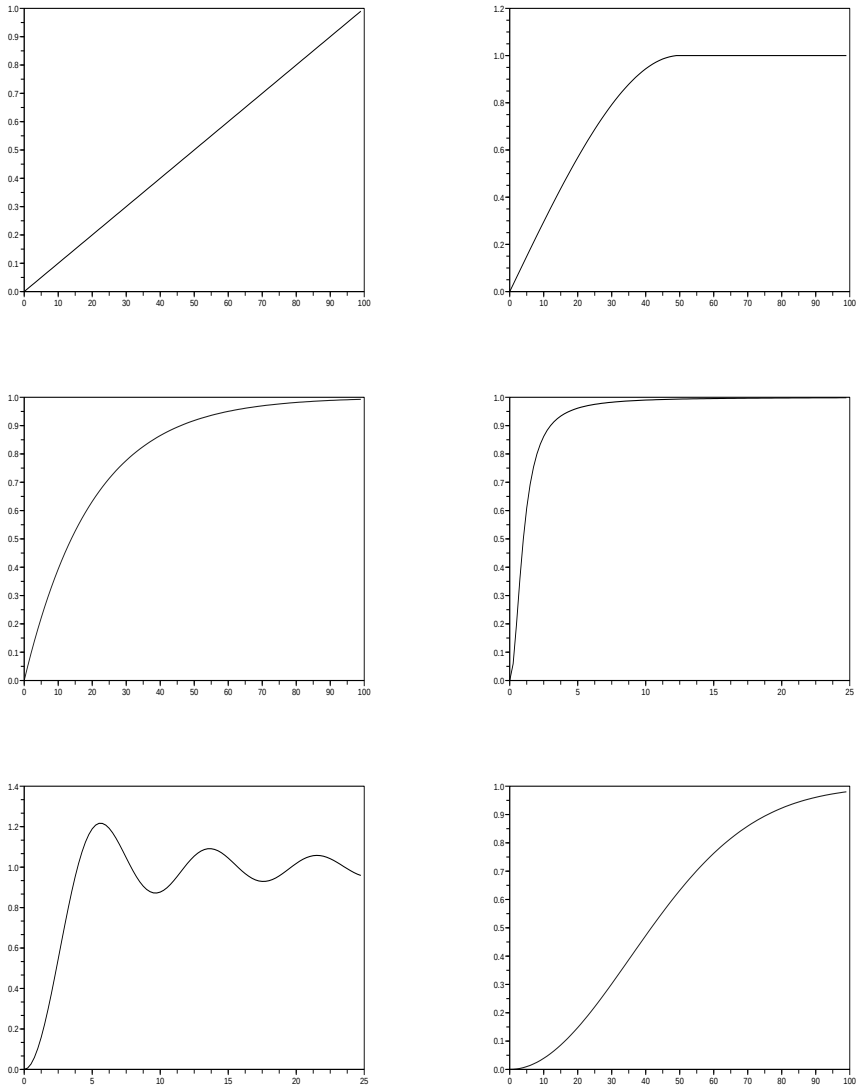

FIG. 3.1. *Examples of valid semi-variograms. From top, left to bottom, right, linear* ($\gamma(h) = \frac{h}{100}$)*, spherical* ($\gamma(h) = \frac{3}{2}\frac{h}{50} - \frac{1}{2}\frac{h^3}{50^3}$ *for* $h \leq 50$ *and* 1*, elsewhere), exponential* ($\gamma(h) = 1 - e^{-\frac{1}{20}h}$)*, rational-quadratic* ($\gamma(h) = \frac{h^2}{1+h^2}$)*, hole-effect* ($\gamma(h) = 1 - 5\frac{sin(0.2\ h)}{h}$) *and Gaussian* ($\gamma(h) = 1 - e^{-\frac{1}{50^2}h^2}$) *semi-variograms.*

REMARK 3.1.

- Observe that each semi-variogram of Figure 3.1 corresponds to different spatial structure of the data. Since it has a zero first derivative at the origin, the Gaussian model translates a stronger dependence between data at small scales than the exponential one. Therefore, it is efficient to represent very regular

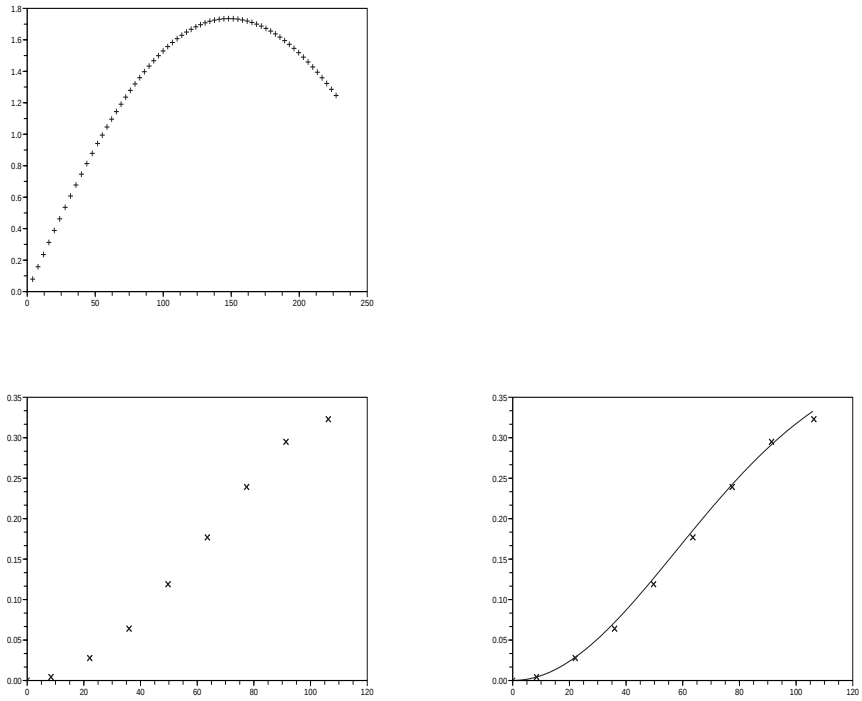

FIG. 3.2. *The different steps of the semi-variogram identification. From top to bottom right, data, experimental semi-variogram, identified semi-variogram model obtained by a least square fit.*

data.

- There exists a more general formulation for theoretical semi-variograms involving an extra-parameter called nugget effect. It corresponds to a discontinuous semi-variogram at the origin and is therefore suitable to model data with strong variations at very small scales. In practice, it is difficult to specify the semi-variogram behavior near the origin since the available set of data is usually not substantial enough to exhibit very small scales dependence. Moreover, the strategy that will be introduced in the next section aims at considering only strongly dependent data for the prediction. Therefore, for the rest of the paper, the semi-variogram is assumed to be continuous at the origin, i.e., $\lim_{h\to 0^+} \gamma(h) = \gamma(0) = 0$.

- Besides Kriging and Lagrange interpolations, an important method for spatial prediction is the spline-based one ([1]). There is a formal connection between splines and Kriging ([12]). However, this kind of prediction is not addressed in this paper for two main reasons. For sake of simplicity, the construction of subdivision schemes is restricted in this paper to the combination of Harten's multi-resolution and ordinary Kriging. In order to work in a unified framework for splines and Kriging, it is required to consider a more general approach than

ordinary Kriging called Kriging with intrinsic random functions ([17]). In this case, the variogram is replaced by a generalized covariance function and all the commonly used spline basis functions appear to be valid generalized covariances ([18]). Moreover, as mentioned in [12], although formally connected, there is an important difference between these two approaches in the spatial structure identification step. In Kriging, one intends to fit, directly from the observed data, a model translating the spatial dependence through the computation of an experimental semi-variogram. In this way, there is no extra assumption related to an a priori choice of the spatial structure which is crucial in industrial applications such as risk analysis studies. This step of "empirical" identification is not considered in the spline-based prediction since each spline basis function corresponds to a given covariance model.

3.1.2. Construction of the ordinary Kriging estimator. The ordinary Kriging estimator of the random process $\mathcal{F}$ at a new location $x^\star$ is denoted $\mathcal{P}(\mathcal{F}, x^\star)$. It is the linear unbiased predictor[3] minimizing the estimation variance $\sigma_K^2 = var(\mathcal{F}(x^\star) - \mathcal{P}(\mathcal{F}, x^\star))$. More precisely, it is written as

$$\mathcal{P}(\mathcal{F}, x^\star) = \sum_{i=0}^{N-1} \lambda_i \mathcal{F}(x_i). \tag{3.6}$$

$\{\lambda_i\}_{i=0,\dots,N-1}$ are the Kriging weights, solutions of the following problem:

$$\begin{bmatrix} \gamma_{0,0} & \dots & \gamma_{0,N-1} & 1 \\ \gamma_{1,0} & \dots & \gamma_{1,N-1} & 1 \\ \dots & \dots & \dots & 1 \\ \gamma_{N-1,0} & \dots & \gamma_{0,0} & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \lambda_2 \\ \dots \\ \lambda_{N-1} \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma_0^\star \\ \gamma_1^\star \\ \dots \\ \gamma_{N-1}^\star \\ 1 \end{bmatrix}, \tag{3.7}$$

where $\gamma_{ij} = \gamma(||x_i - x_j||_2)$ and $\gamma_i^* = \gamma(||x^* - x_i||_2)$ with $||.||_2$ the Euclidean norm and μ the Lagrange multiplier enforcing the unbiasedness of the estimator. Besides the Kriging estimator, the Kriging weights allow one to compute the estimation variance σ_K^2: this quantity can be used to provide an a priori (i.e. depending only on the Kriging weights and on the identified semi-variogram) local prediction error as it is explained at the end of Section 3.2.2.

3.2. Kriging-based subdivision scheme.

[3] $\mathcal{P}(\mathcal{F}, x)$ is an unbiased estimator if $E\left(\mathcal{P}(\mathcal{F}, x) - \mathcal{F}(x)\right) = 0$ where E is the mathematical expectation.

3.2.1. Construction of the scheme. When replacing $x^\star$ by $(2k-1)2^{-j}$, N by $r+l$ and x_i by $(k+i-l)2^{-(j-1)}$, Expression (3.6) can be plugged into the multi-scale framework. By comparison with the second equation of (2.1), it can then be interpreted as a two-scale relation defining a non-stationary and non-homogeneous subdivision scheme.

Assuming that the semi-variogram has been identified from the initial data f^0, the Kriging-based subdivision scheme is then defined in the position-dependent case as follows:

$$\text{(3.8)} \quad \begin{cases} f'^{j}_{2k} = \left(P^{j}_{j-1} f'^{j-1}\right)_{2k} = f'^{j-1}_{k}, \\ f'^{j}_{2k-1} = \left(P^{j}_{j-1} f'^{j-1}\right)_{2k-1} = \sum_{m=-l_{j,2k-1}}^{r_{j,2k-1}-1} \lambda^{l_{j,2k-1},r_{j,2k-1}}_{j,m} f'^{j-1}_{k+m}, \end{cases}$$

with $\forall k \in \mathbb{Z}$, $f'^{0}_{k} = f^{0}_{k}$. The family $\{\lambda^{l_{j,2k-1},r_{j,2k-1}-1}_{j,m}\}_m$ stands for the Kriging weights solution of Problem (3.7) where $x^\star$, N and $\{x_i\}_{i=0,\dots,N-1}$ are replaced by the corresponding quantities in the multi-scale framework (as mentioned at the beginning of this section) and the couple (l,r) is replaced by $(l_{j,2k-1}, r_{j,2k-1})$.

Note that at the first level, the subdivision scheme operates on $\{f^0_k\}_{k\in\mathbb{Z}}$ which are the outcomes of a random process $\mathcal{F}(x)$ associated to the initial mesh grid. However, for any $j>0$, the data involved in the prediction cannot be considered as realizations of the random variables $\{\mathcal{F}(k2^{-j})\}_{k\in\mathbb{Z}}$ since they are computed from previous levels. Therefore, we denote in what follows $\mathcal{F}'^{j}(x)$ the random process associated to $\{f'^{j}_{k}\}_{k\in\mathbb{Z}}$.

The masks of the associated subdivision scheme is defined for even values as

$$\text{(3.9)} \quad \begin{cases} a^{j,2k}_{0} = 1, \\ a^{j,2k}_{m} = 0, \text{ for m} \neq 0, \end{cases}$$

and for odd values as

$$\text{(3.10)} \quad \begin{cases} a^{j,2k-1}_{-2m-1} = \lambda^{l_{j,2k-1},r_{j,2k-1}}_{j,m} \text{ for } m = -l_{j,2k-1}, \dots, r_{j,2k-1}-1, \\ a^{j,2k-1}_{2m} = 0 \text{ otherwise.} \end{cases}$$

Remark 3.2.
Note that ordinary Kriging only requires that the Kriging weights sum up to

1. That means that the above defined Kriging-based subdivision scheme reproduces constants[4]*. They are the only reproduced polynomials. It is possible to increase the degree of polynomial reproduction by considering in Decomposition (3.1) a polynomial (of degree larger than* 1*) deterministic mean structure. In that case, one speaks about universal Kriging ([12]).*

3.2.2. Estimation variance. Coming back to the scheme (3.8) and exploiting the underlying Kriging framework, it is possible to define an estimation variance that reads for each j and k:

$$\left(\sigma_k^{'j}\right)^2 = E\left(\mathcal{F}'^j(k\ 2^{-j}) - \mathcal{F}(k\ 2^{-j})\right)^2, \tag{3.11}$$

It turns out that this variance can be expressed thanks to the identified semi-variogram and the weights of the scheme. More precisely, we have:

PROPOSITION 3.2.
Introducing for any j and k, $\{A^j_{k,m}\}_{m\in I^j_k}$ such that,

$$f_k^{'j} = \sum_{m\in I_k^j} A^j_{k,m} f^0_m, \tag{3.12}$$

then the estimation variance associated to the prediction scheme reads:

$$\left(\sigma_k^{'j}\right)^2 = -\sum_{(n,m)\in I_k^j\times I_k^j} A^j_{k,n} A^j_{k,m}\gamma(|n-m|) + 2\sum_{n\in I_k^j} A^j_{k,n}\gamma(|k2^{-j}-n|). \tag{3.13}$$

Proof. Since the Kriging-based subdivision scheme reproduces constants (Remark 3.2), it is straightforward to get: $\sum_{m\in I_k^j} A^j_{k,m} = 1$. Using (3.12), (3.11) then reads:

$$\begin{aligned}\left(\sigma_k^{'j}\right)^2 &= E\left(\sum_{m\in I_k^j} A^j_{k,m}\left(\mathcal{F}(m) - \mathcal{F}(k2^{-j})\right)\right)^2, \\ &= \sum_{m\in I_k^j} (A^j_{k,m})^2 E\left(\mathcal{F}(m) - \mathcal{F}(k2^{-j})\right)^2\end{aligned}$$

[4] A subdivision scheme S is said to reproduce constants if for any constant C and any scale j, $\{\forall k\in\mathbb{Z}, f^j_k = C\} \Longrightarrow \{(Sf^j)_k = C, \forall k\in\mathbb{Z}\}$.

$$+ \sum_{(n,m)\in I_k^j\times I_k^j, n\neq m} A_{k,n}^j A_{k,m}^j E\left(\mathcal{F}(n)-\mathcal{F}(k2^{-j})\right)\left(\mathcal{F}(m)-\mathcal{F}(k2^{-j})\right).$$

Writing

$$\left(\mathcal{F}(n)-\mathcal{F}(k2^{-j})\right)\left(\mathcal{F}(m)-\mathcal{F}(k2^{-j})\right) = -\frac{1}{2}\left(\mathcal{F}(n)-\mathcal{F}(m)\right)^2 + \frac{1}{2}\left(\mathcal{F}(n)-\mathcal{F}(k2^{-j})\right)^2 + \frac{1}{2}\left(\mathcal{F}(m)-\mathcal{F}(k2^{-j})\right)^2$$

the previous expression becomes:

$$\begin{aligned}\left(\sigma_k^{'j}\right)^2 &= \sum_{m\in I_k^j}(A_{k,m}^j)^2 E\left(\mathcal{F}(m)-\mathcal{F}(k2^{-j})\right)^2 \\ &\quad - \frac{1}{2}\sum_{(n,m)\in I_k^j\times I_k^j, n\neq m} A_{k,n}^j A_{k,m}^j E\left(\mathcal{F}(n)-\mathcal{F}(m)\right)^2 + \sum_{(n,m)\in I_k^j\times I_k^j, n\neq m} A_{k,n}^j A_{k,m}^j E\left(\mathcal{F}(m)-\mathcal{F}(k2^{-j})\right)^2.\end{aligned}$$

From (3.2), we get:

$$\begin{aligned}\left(\sigma_k^{'j}\right)^2 &= 2\sum_{m\in I_k^j}(A_{k,m}^j)^2\gamma(|m-k2^{-j}|) - \sum_{(n,m)\in I_k^j\times I_k^j, n\neq m} A_{k,n}^j A_{k,m}^j \gamma(|n-m|) \\ &\quad + 2\sum_{(n,m)\in I_k^j\times I_k^j, n\neq m} A_{k,n}^j A_{k,m}^j \gamma(|m-k2^{-j}|) \\ &= 2\sum_{n\in I_k^j} A_{k,n}^j \gamma(|n-k2^{-j}|)\left(A_{k,n}^j + \sum_{m\neq n} A_{k,m}^j\right) - \sum_{(n,m)\in I_k^j\times I_k^j, n\neq m} A_{k,n}^j A_{k,m}^j \gamma(|n-m|).\end{aligned}$$

Since $\sum_{m\in I_k^j} A_{k,m}^j = 1$ and $\gamma(0)=0$ (see Remark 3.1), we finally find:

$$\left(\sigma_k^{'j}\right)^2 = 2\sum_{n\in I_k^j} A_{k,n}^j \gamma(|n-k2^{-j}|) - \sum_{(n,m)\in I_k^j\times I_k^j} A_{k,n}^j A_{k,m}^j \gamma(|n-m|),$$

that concludes the proof. ∎

Proposition 3.2 provides an important result for practical applications. Indeed, this proposition allows one to quantify the a priori prediction error associated to each predicted point. More precisely, under normality assumption, the estimation variance can be used to compute, for any $(k,j)\in\mathbb{Z}^2$, the so called confidence intervals $I_\alpha(k,j)$ defined for any confidence level $\alpha\in[0,1]$ by

$$Probability\left(\mathcal{F}^{'j}(k2^{-j}) - \mathcal{F}(k2^{-j}) \in I_\alpha(k,j)\right) = \alpha.$$

This error depends only on the type of identified semi-variogram and on the distance between points involved in the Kriging system. It is an a-priori estimate since its computation does not require to know the true values of f^j.

In the next section, we focus on the convergence analysis of Kriging-based subdivision schemes.

4. Convergence analysis of Kriging-based subdivision scheme. A first result of convergence has been published in [8]; it corresponds to the case $l = r = 2$. This section can then be considered as a generalization of [8] to the case $l = r \in \mathbb{N}^*$.
The analysis is carried out following several steps. The key notions associated to the convergence of subdivision schemes are first recalled. Conditions to ensure the convergence of non-stationary schemes following translation-invariant and position-dependent strategies are then proposed within the matrix formalism with a special attention to the case where an asymptotical connection to a convergent stationary subdivision scheme can be established. Finally, the convergence of the Kriging-based scheme constructed in this paper is fully studied: for that, two important results related to the asymptotical behavior of Kriging weights for the two types of semi-variogram models introduced in Definition 3.1 are provided.

4.1. Definition of convergence. In order to consider a large class of subdivision schemes, we propose a definition of convergence that can be interpreted as a generalization of the classical one ([14]).

According to the definition of $\mathcal{C}(\mathbb{R}\backslash\mathcal{S})$ (Definition 2.1) we propose the following definition for convergence taken from [7].

Definition 4.1. *Convergence of subdivision scheme*

The subdivision scheme S is said to be L^∞-convergent if for any real sequence f^0, there exists a function $f \in \mathcal{C}(\mathbb{R}\backslash\mathcal{S})$ (called the limit function associated to f^0) such that:

$\forall \epsilon, \exists J$ *such that* $\forall j \geq J$ *either*

$$\| S^j f^0 - f^+\left(\frac{\cdot}{2^j}\right) \|_\infty \leq \epsilon,$$

or

$$\| S^j f^0 - f^-\left(\frac{\cdot}{2^j}\right) \|_\infty \leq \epsilon.$$

where $||.||_\infty$ denotes the l^∞-norm defined for any sequence $u = \{u_k\}_{k\in\mathbb{Z}}$ by $||u||_\infty = \sup_{k\in\mathbb{Z}}(|u_k|)$.

Note that if the family of segmentation points, $\mathcal{S}_f$, is empty, f^+ and f^- are replaced by f and Definition 4.1 coincides with the classical convergence definition given in [14].

All the convergence results available nowadays involve, given a subdivision scheme S reproducing constants, the subdivision scheme $\frac{1}{2}S_1$ associated to the subdivision of the sequence δf^j with $\delta f^j_k = f^j_{k+1} - f^j_k, k \in \mathbb{Z}$.
In [14], a necessary and sufficient condition is given (Theorem 3.2 p 53) as: *S is a uniformly convergent subdivision scheme, if and only if $\frac{1}{2}S_1$ uniformly converges to the zero function for all initial data f^0.*

Essentially, the condition requires the existence of a positive integer L and a constant $0 < \mu < 1$ such that

$$|| \left(\frac{1}{2}S_1\right)^L ||_\infty \leq \mu,$$

where for simplicity in the notation, $||.||_\infty$ here denotes the induced norm for operator associated to the l^∞-norm.

This result has been generalized to a sufficient condition for convergence of quasi-linear subdivision schemes[5] in [11] as follows (Theorem 1 p 99): *if the rule for the differences $\frac{1}{2}S_1$ satisfies $\rho_\infty(\frac{1}{2}S_1) < 1$ then the quasi-linear subdivision scheme S is uniformly convergent*, where the joint spectral radius of a subdivision scheme S, denoted $\rho_\infty(S)$ is defined by

$$\rho_\infty(S) = \limsup_{j\to\infty} \sup_{(u^0,u^1,\ldots,u^{j-1})\in(l^\infty)^j} ||S(u^{j-1})...S(u^0)||^{1/j}_\infty.$$

We recall that in the framework of [11] for any sequence u, the operator $S(u)$ is a subdivision scheme.

For applications to specific families of subdivision schemes various versions of these theorems have been derived (independently or not of the previously quoted papers). They imply a matrix formalism ([19], [15], [16], [9]) and a spectral analysis first introduced in [14]. They have been extended to stationary but non-uniform position-dependent subdivision schemes in [7]. We present in the following section an extension of these results to non-stationary translation-invariant and position-dependent schemes. It is then used to analyze the convergence of the Kriging-based subdivision schemes constructed in Section 3.2.

4.2. Convergence analysis of non-stationary translation-invariant and position-dependent subdivision schemes.

[5] A quasi-linear subdivision scheme is the recursive action of a data dependent rule:$Sv := S(v)v$.Applied on an initial sequence f^0 it reads: $f^j := S(f^{j-1})...S(f^0)f^0$.

4.2.1. Matrix formalism. Extending [14] and [7], it turns out that one can associate to a non-stationary scheme, S, a set of refinement matrices for a translation-invariant strategy and a set of refinement and edge matrices (due to subdivision around segmentation points) when choosing a position-dependent stencil. More precisely, we have:

Definition 4.2.
- For a translation-invariant subdivision, let F_k^j be the minimal set of N points at level j that determines the values at dyadic points in the interval $[k2^{-j}, (k+1)2^{-j}]$ at level above j. For each j, the two $N \times N$ refinement matrices, $A_{0,j}$ and $A_{1,j}$, that transform the set F_k^{j-1} into the sets F_{2k}^j and F_{2k+1}^j are defined as follows:

$$F_{2k}^j = A_{0,j} F_k^{j-1} \text{ and } F_{2k+1}^j = A_{1,j} F_k^{j-1}.$$

- For subdivision at level j around a segmentation point y_0 characterized by the index k_j (see Definition 2.2), we introduce G_+^j (resp. G_-^j) the minimal set of M points at level j that determines the values at dyadic points in the interval $[k2^{-j}, (k+1)2^{-j}]$ on the right of y_0 (resp. the values at dyadic points in the interval $[(k-1)2^{-j}, k2^{-j}]$ on the left of y_0). For each j, the two $M \times M$-edge matrices $A_{2,j}^+$ and $A_{2,j}^-$ are defined by:

$$G_-^j = A_{2,j}^- G_-^{j-1} \text{ and } G_+^j = A_{2,j}^+ G_+^{j-1}.$$

Following [7], the non-stationary subdivision scheme is then completely characterized:

- when the translation-invariant strategy is used, by the set of refinement matrices $\{A_{0,j}, A_{1,j}\}_{j\in\mathbb{Z}}$.
- when the position-dependent strategy is used, by the two sets of refinement and edge matrices $\{A_{0,j}, A_{1,j}\}_{j\in\mathbb{Z}}$ and $\{A_{2,j}^-, A_{2,j}^+\}_{j\in\mathbb{Z}}$ and, for any dyadic point $x = k_x 2^{-j_x}$ ($x \neq y_0$), by the length T_x of the transition zone corresponding to the range of scales where the two successive differences, $\delta f_{k2^{j-j_x}}^j$ and $\delta f_{k2^{j-j_x}-1}^j$, are computed mixing refinement and edge matrices.

Assuming that S reproduces constants, in the same way that one defines refinement (resp. refinement and edge) matrices for the subdivision scheme S, we introduce, for each j, $A_{0,j}^{(1)}$ and $A_{1,j}^{(1)}$ (resp. $A_{2,j}^{(1),-}$ and $A_{2,j}^{(1),+}$) the refinement (resp. edge) matrices associated to S_1.

The following convergence theorem then holds, considering a scheme that reproduces constants:

THEOREM 4.1.
Translation-invariant strategy:
If there exists J such that,

$$\limsup_{m\to+\infty} ||\Pi_{k\leq J}\frac{1}{2}A^{(1)}_{\epsilon,m+k}||_\infty \leq \mu_0 < 1, \text{ for all } \epsilon \in \{0,1\}, \tag{4.1}$$

then the translation-invariant subdivision scheme is uniformly convergent.

Position-dependent strategy:
Assume that there exists J such that,

$$\limsup_{m\to+\infty} ||\Pi_{k\leq J}\frac{1}{2}A^{(1)}_{\epsilon,m+k}||_\infty \leq \mu_0 < 1, \text{ for all } \epsilon \in \{0,1\}, \tag{4.2}$$

$$\limsup_{m\to+\infty} ||\Pi_{k\leq J}\frac{1}{2}A^{(1),-}_{2,m+k}||_\infty \leq \mu_1 < 1 \ , \ ||\Pi_{m\leq J}\frac{1}{2}A^{(1),+}_{2,m}||_\infty \leq \mu_2 < 1, \tag{4.3}$$

and that there exists $T < \infty$, such that: $\forall x$, $T_x \leq T$,
then the position-dependent subdivision scheme is uniformly convergent.

Proof. This theorem comes from Theorem 2.1 of [7]. Non stationarity implies to replace quantities of type $||\Pi_{k\leq L}A_k||_\infty$ by $\limsup_{m\to+\infty} ||\Pi_{k\leq J}A_{m+k}||_\infty$. ∎

The sketch of the proof remains the same: Expression (4.1) leads to the convergence in the translation-invariant zones while Expressions (4.2)-(4.3) lead to the punctual convergence at segmentation points. The global and uniform convergence can be proved focussing on the subdivision process at a fixed point $x = k_x 2^{-j_x}$ and exploiting the fact that the length of the transition zone (T_x) is uniformly bounded. ∎

Bounds (4.1) and (4.2)-(4.3) are usually difficult to prove in practice. In the next proposition, we reformulate them in a more tractable form using the eigenvalues of the involved matrices. We note $\rho(A)$ the spectral radius of a matrix A, i.e the supremum of the modulus of its eigenvalues.

PROPOSITION 4.3.
Condition (4.1), resp. Conditions (4.2)-(4.3), can be replaced by:

there exists J' such that,

$$\limsup_{m\to+\infty} \rho\left(\Pi_{k\leq J'}\frac{1}{2}A^{(1)}_{\epsilon,m+k}\right) < 1, \text{ for all } \epsilon \in \{0,1\}, \tag{4.4}$$

resp.:

there exists J' such that,

$$\limsup_{m\to+\infty} \rho\left(\Pi_{k\leq J'} \frac{1}{2} A^{(1)}_{\epsilon,m+k}\right) < 1, \text{ for all } \epsilon \in \{0,1\}, \tag{4.5}$$

$$\limsup_{m\to+\infty} \rho\left(\Pi_{k\leq J'} \frac{1}{2} A^{(1),-}_{2,m+k}\right) < 1, \tag{4.6}$$

$$\limsup_{m\to+\infty} \rho\left(\Pi_{k\leq J'} \frac{1}{2} A^{(1),+}_{2,m+k}\right) < 1, \tag{4.7}$$

Proof. If (4.4) (resp. (4.5)-(4.7)) holds then from norm equivalence, (4.1) (resp. (4.2)-(4.3)) is satisfied for large enough value of J. The reverse is also true. Therefore, the equivalence is proved. ∎

4.2.2. Asymptotical behavior and convergence results. Inside the general class of non-stationary subdivision schemes reproducing constants, we focus on the class of schemes for which the masks converge towards the masks of a stationary scheme SS reproducing constants. We then have the following result:

THEOREM 4.4.
Let S be a non-stationary subdivision scheme defined by its mask $\{a_l^{j,k}\}_{l\in\mathbb{Z}}$, $(j,k) \in \mathbb{Z}^2$. We suppose that there exists two constants $K < K'$, independent of j and k such that $a_l^{j,k} = 0$ for $l > K'$ or $l < K$. If there exists a stationary subdivision scheme SS of mask $\{a_l^k\}_{l\in\mathbb{Z}}$, $k \in \mathbb{Z}$ with $a_l^k = 0$ for $l > K'$ or $l < K$ and satisfying one of the theorems of Sections 4.2.1 and such that

$$\lim_{j\to+\infty} ||a^{j,k} - a^k||_\infty = 0$$

then S is convergent.

Proof. It is a classical result (see for instance [22]) that the function that associates to a matrix $A \in \mathbb{C}^{n\times n}$ its eigenvalues is Lipchitz of exponent $1/n$. By hypothesis, the size of the refinement/edge matrices is bounded. Therefore, the convergence of the coefficients of the mask implies the convergence of the refinement/edge matrices. Finally, since the refinement/edge matrices of the scheme SS satisfy one of the convergence theorems of Sections 4.2.1, it is also true for the non-stationary scheme from a large enough scale j, that concludes the proof. ∎

REMARK 4.1.
Extending the analysis of [11], the convergence of S towards SS under the assumptions of the previous theorem also implies that for any initial sequence f^0, the Hölder exponents of the limit function $S^\infty f^0$ and of the limit function $SS^\infty f^0$ are the same.

In the sequel, we use the matrix formalism and Theorem 4.4 to analyze the convergence of Kriging-based subdivision schemes with parameters (D, r, r).

4.3. Convergence of Kriging-based subdivision scheme with parameters (D, r, r).

4.3.1. Refinement and edge matrices. Before studying the convergence property of the Kriging-based subdivision scheme, the following proposition provides the expressions of the refinement and edge matrices. They are obtained through elementary calculus.

PROPOSITION 4.5.
Refinement matrix:

The two refinement matrices associated to a translation-invariant Kriging-based subdivision with parameters (D, r, r) are the two $(4r-2) \times (4r-2)$ matrices defined by:

$$A_{0,j} = \begin{bmatrix} 0 & \dots & \dots & 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ \lambda_{j,-r}^{r,r} & \lambda_{j,-r+1}^{r,r} & \dots & \lambda_{j,-1}^{r,r} & \lambda_{j,0}^{r,r} & \dots & \lambda_{j,r-1}^{r,r} & .. & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \lambda_{j,-r}^{r,r} & \dots & \lambda_{j,r-1}^{r,r} \end{bmatrix},$$

and

$$A_{1,j} = \begin{bmatrix} \lambda_{j,-r}^{r,r} & \dots & \lambda_{j,0}^{r,r} & \dots & \lambda_{j,r-1}^{r,r} & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & 1 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & \lambda_{j,-r}^{r,r} & \dots & \lambda_{j,0}^{r,r} & \dots & \lambda_{j,r-1}^{r,r} \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 1 & \dots & 0 \end{bmatrix}.$$

Edge matrix:

Moreover, writing for any $j \geq 0$, $y_0 \in [x^{j-1}_{k_{j-1}-1}, x^{j-1}_{k_{j-1}}]$, the vectors G^j_- and G^j_+ introduced in Definition 4.2 are:

$$G^j_- = \{f^j_{k_j+\alpha}, \alpha = -2r, ..., -1\}, \tag{4.8}$$

$$G^j_+ = \{f^j_{k_j+\alpha}, \alpha = 0, ..., 2r-1\}, \tag{4.9}$$

and the four $2r \times 2r$ edge matrices associated to the subdivision near the segmentation point read:

- *if $y_0 \in [x^j_{2k_{j-1}-2}, x^j_{2k_{j-1}-1}[$,*

$$A^-_{2,j} = \begin{bmatrix} \lambda^{r,r}_{j,-r} & ... & ... & \lambda^{r,r}_{j,0} & ... & \lambda^{r,r}_{j,r-1} \\ 0 & ... & ... & 1 & 0 & ... \\ \lambda^{r+1,r-1}_{j,-r-1} & ... & ... & ... & ... & \lambda^{r+1,r-1}_{j,r-2} \\ ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... \\ \lambda^{2r-1,1}_{j,-2r+1} & ... & ... & ... & ... & \lambda^{2r-1,1}_{j,0} \\ 0 & ... & ... & ... & ... & 1 \end{bmatrix},$$

$$A^+_{2,j} = \begin{bmatrix} \lambda^{0,2r}_{j,0} & ... & ... & ... & \lambda^{0,2r}_{j,2r-1} \\ 1 & 0 & ... & ... & 0 \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ \lambda^{r-1,r+1}_{j,-r+1} & ... & \lambda^{r-1,r+1}_{j,0} & ... & \lambda^{r-1,r+1}_{j,r} \\ ... & ... & 1 & ... & 0 \end{bmatrix}.$$

- *if $y_0 \in [x^j_{2k_{j-1}-1}, x^j_{2k_{j-1}}]$*

$$A^-_{2,j} = \begin{bmatrix} 0 & ... & 1 & 0 & ... \\ \lambda^{r+1,r-1}_{j,-r-1} & ... & \lambda^{r+1,r-1}_{j,0} & ... & \lambda^{r+1,r-1}_{j,r-2} \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ 0 & ... & ... & ... & 1 \\ \lambda^{2r,0}_{j,-2r} & ... & ... & ... & \lambda^{2r,0}_{j,-1} \end{bmatrix},$$

$$A^+_{2,j} = \begin{bmatrix} 1 & 0 & ... & ... & ... & ... & 0 \\ \lambda^{1,2r-1}_{j,-1} & ... & ... & ... & ... & ... & \lambda^{1,2r-1}_{j,2r-2} \\ ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... \\ ... & ... & 1 & ... & ... & ... & 0 \\ \lambda^{r,r}_{j,-r} & ... & \lambda^{r,r}_{j,0} & ... & ... & ... & \lambda^{r,r}_{j,r-1} \end{bmatrix}.$$

In [7], a general convergence result has been established in the case of position-dependent Lagrange interpolatory subdivision. Using the property of polynomial reproduction up to degree D, it has been proved that the eigenvalues of the matrices related to the first differences scheme are strictly less than 1. Here, since we are working with ordinary Kriging, the Kriging-based subdivision scheme only reproduces constants. This information is not substantial enough to directly get qualitative insight on the modulus of eigenvalues of each matrix involved in the subdivision process and provided by Proposition 4.5. However, for semi-variograms of types $\mathcal{G}^{\star}_{OD}$ and $\mathcal{G}^{\star}_{ED}$ (Definition 3.1) (that encompass most of the classical semi-variograms used in practice), the uniform convergence can be derived by studying the limit process and its connection to the convergent Lagrange stationary scheme (Theorem 4.4). This is why we study the limit of the Kriging weights when $j \to +\infty$ in the sequel.

4.3.2. $\mathcal{G}^{\star}_{OD}$-type semi-variograms. The following proposition holds.

PROPOSITION 4.6.
The Kriging weights involved in the translation-invariant and position-dependent subdivision schemes with a $\mathcal{G}^{\star}_{OD}$ semi-variogram satisfy:

$$
\begin{gathered}
(\lambda^{r,r}_{j,-r},...,\lambda^{r,r}_{j,r-1}) \to_{j\to+\infty} (L^{r,r}_{-r},...,L^{r,r}_{r-1}),\\
(\lambda^{r-1,r+1}_{j,-r+1},...,\lambda^{r-1,r+1}_{j,r}) = (\lambda^{r+1,r-1}_{j,r-2},...,\lambda^{r+1,r-1}_{j,-r-1}) \to_{j\to+\infty} (L^{r-1,r+1}_{-r+1},...,L^{r-1,r+1}_{r}),\\
\dots\\
(\lambda^{1,2r-1}_{j,-1},...,\lambda^{1,2r-1}_{j,2r-2}) = (\lambda^{2r-1,1}_{j,0},...,\lambda^{2r-1,1}_{j,-2r+1}) \to_{j\to+\infty} (L^{1,2r-1}_{-1},...,L^{1,2r-1}_{2r-2}),\\
(\lambda^{0,2r}_{j,0},...,\lambda^{0,2r}_{j,2r-1}) = (\lambda^{2r,0}_{j,-1},...,\lambda^{2r,0}_{j,-2r}) \to_{j\to+\infty} (L^{0,2r}_{0},...,L^{0,2r}_{2r-1}).
\end{gathered}
$$

where $\{L^{k_1,k_2}_m\}_m$ are the Lagrange interpolatory weights recalled by Formula (2.4).

Proof. We focus on the first expression, the proofs for the remaining ones being similar.
Since we are interested in the behavior of the Kriging weights when $j \to \infty$ (or equivalently when h, the distance between points is small) and remembering the definition of $\mathcal{G}^{\star}_{OD}$ (Definition 3.1), one writes

$$\gamma(h) = P_{2r}(h) + O(h^{2r+2}),$$

with $P_{2r}(h) = \sum_{n=0}^{r-1} b_n h^{2n+2}$. Then, introducing for the rest of the proof the notation $P_{2r,j}(.) = P_{2r}(2^{-j}.)$, the Kriging system from scale j to scale $j+1$ becomes:

$$\left(\Gamma^0_{j,2r+1} + \Gamma^1_{j,2r+1}\right)\mathcal{U}_{j+1} = \gamma^0_{j,2r+1} + \gamma^1_{j,2r+1}, \tag{4.10}$$

where

$$\Gamma^0_{j,2r+1} = \begin{bmatrix} 0 & P_{2r,j}(1) & ... & P_{2r,j}(2r-1) & 1 \\ P_{2r,j}(1) & ... & P_{2r,j}(2r-2) & ... & 1 \\ ... & ... & ... & ... & ... \\ P_{2r,j}(2r-1) & P_{2r,j}(2r-2) & ... & 0 & 1 \\ 1 & 1 & 1 & ... & 0 \end{bmatrix},$$

$$\text{and } \gamma^0_{j,2r+1} = \begin{bmatrix} P_{2r,j}(\frac{2r-1}{2}) \\ P_{2r,j}(\frac{2r-3}{2}) \\ ... \\ P_{2r,j}(\frac{2r-1}{2}) \\ 1 \end{bmatrix}.$$

As for $\Gamma^1_{j,2r+1}$ and $\gamma^1_{j,2r+1}$, it is elementary to show that:

$$||\Gamma^1_{j,2r+1}||_\infty \leq K_{\Gamma^1} 2^{-(2r+2)j},\ ||\gamma^1_{j,2r+1}||_\infty \leq K_{\gamma^1} 2^{-(2r+2)j}. \tag{4.11}$$

The solution of (4.10) can then be formally written as:

$$\begin{aligned} \mathcal{U}_{j+1} &= \left(\Gamma^0_{j,2r+1} + \Gamma^1_{j,2r+1}\right)^{-1} \left(\gamma^0_{j,2r+1} + \gamma^1_{j,2r+1}\right), \\ &= \left(I_{2r+1} + \left(\Gamma^0_{j,2r+1}\right)^{-1} \Gamma^1_{j,2r+1}\right)^{-1} \left(\Gamma^0_{j,2r+1}\right)^{-1} \gamma^0_{j,2r+1} \\ &+ \left(I_{2r+1} + \left(\Gamma^0_{j,2r+1}\right)^{-1} \Gamma^1_{j,2r+1}\right)^{-1} \left(\Gamma^0_{j,2r+1}\right)^{-1} \gamma^1_{j,2r+1}, \end{aligned} \tag{4.12}$$

where I_{2r+1} is the $(2r+1)\times(2r+1)$ identity matrix. Note that to go from the first line to the second one in Expression (4.12), we have assumed that $\left(\Gamma^0_{j,2r+1} + \Gamma^1_{j,2r+1}\right)^{-1} = \left(I_{2r+1} + \left(\Gamma^0_{j,2r+1}\right)^{-1} \Gamma^1_{j,2r+1}\right)^{-1} \left(\Gamma^0_{j,2r+1}\right)^{-1}$, which is satisfied provided $\left(\Gamma^0_{j,2r+1}\right)^{-1}$ and $\left(I_{2r+1} + \left(\Gamma^0_{j,2r+1}\right)^{-1} \Gamma^1_{j,2r+1}\right)^{-1}$ exist. For the moment, we take for granted that this is the case and perform the proof in three steps. The invertibility will be verified in Step 1 and Step 3.

- Step 1: Expansion of $\left(I_{2r+1} + \left(\Gamma^0_{j,2r+1}\right)^{-1} \Gamma^1_{j,2r+1}\right)^{-1}$

The expression of $\left(\Gamma^0_{j,2r+1}\right)^{-1}$ is obtained according to a classical result of matrix calculus. More precisely, $\Gamma^0_{j,2r+1}$ can be written as:

$$\Gamma^0_{j,2r+1} = \begin{bmatrix} B & \mathbb{1}_{2r} \\ \mathbb{1}^T_{2r} & 0 \end{bmatrix}.$$

where $\mathbb{1}_{2r}$ is the $2r$-vector whose elements are equal to 1, the upper-script T denotes the transpose operator and

$$B = 2^{-2j} \begin{bmatrix} 0 & ... & Q_{2r-2,j}(2r-1) \\ ... & ... & ... \\ Q_{2r-2,j}(2r-1) & ... & 0 \end{bmatrix},$$

with $Q_{2r-2,j}(h) = Q_{2r-2}(2^{-j}h) = \sum_{n=0}^{r-1} 2^{-2nj} b_n h^{2n}$.

Introducing the Schur complement $S = -\mathbb{1}_{2r}^T B^{-1} \mathbb{1}_{2r}$, $\left(\Gamma_{j,2r+1}^0\right)^{-1}$ is given by:

$$\left(\Gamma_{j,2r+1}^0\right)^{-1} = \begin{bmatrix} B^{-1}\left(I_{2r} + \mathbb{1}_{2r} S^{-1} \mathbb{1}_{2r}^T B^{-1}\right) & -B^{-1}\mathbb{1}_{2r} S^{-1} \\ -S^{-1}\mathbb{1}_{2r}^T B^{-1} & S^{-1} \end{bmatrix}.$$

From these definitions, it is then straightforward that for sufficiently large j,

$$||B^{-1}\left(I_{2r} + \mathbb{1}_{2r} S^{-1} \mathbb{1}_{2r}^T B^{-1}\right)||_\infty \leq K_1 2^{2j},\ ||B^{-1}\mathbb{1}_{2r} S^{-1}||_\infty \leq K_2,$$
$$||\left(S^{-1}\mathbb{1}_{2r}^T B^{-1}\right)^T||_\infty \leq K_3,\ |S^{-1}| \leq K_4 2^{-2j},$$

where K_1, K_2, K_3 and K_4 are independent of j.
Therefore, $||\left(\Gamma_{j,2r+1}^0\right)^{-1}||_\infty \leq K_{\Gamma^0} 2^{2j}$. Combined with (4.11) it gives

$$||\left(\Gamma_{j,2r+1}^0\right)^{-1} \Gamma_{j,2r+1}^1||_\infty \leq K_{\Gamma^0} K_{\Gamma^1} 2^{-2rj}, \tag{4.13}$$

and finally,
$\left(I_{2r+1} + \left(\Gamma_{j,2r+1}^0\right)^{-1} \Gamma_{j,2r+1}^1\right)^{-1} = \sum_{i=0}^{\infty} (-1)^i \left(\left(\Gamma_{j,2r+1}^0\right)^{-1} \Gamma_{j,2r+1}^1\right)^i$.

Note that this step ensures the invertibility of $I_{2r+1} + \left(\Gamma_{j,2r+1}^0\right)^{-1} \Gamma_{j,2r+1}^1$ provided $\left(\Gamma_{j,2r+1}^0\right)^{-1}$ exists which will be verified in Step 3.
Expression (4.12) becomes:

$$\begin{aligned} \mathcal{U}_{j+1} &= \left(\Gamma_{j,2r+1}^0\right)^{-1} \gamma_{j,2r+1}^0 \\ &+ \left(\sum_{i=1}^{\infty} (-1)^i \left(\left(\Gamma_{j,2r+1}^0\right)^{-1} \Gamma_{j,2r+1}^1\right)^i \left(\Gamma_{j,2r+1}^0\right)^{-1} \gamma_{j,2r+1}^0 \right. \\ &+ \left. \sum_{i=0}^{\infty} (-1)^i \left(\left(\Gamma_{j,2r+1}^0\right)^{-1} \Gamma_{j,2r+1}^1\right)^i \left(\Gamma_{j,2r+1}^0\right)^{-1} \gamma_{j,2r+1}^1 \right). \end{aligned} \tag{4.14}$$

This step also provides the following inequalities that are useful in the sequel:

(4.15) $|| \left(\Gamma^0_{j,2r+1}\right)^{-1} \gamma^0_{j,2r+1}||_\infty \leq \mathcal{K}_0,\ || \left(\Gamma^0_{j,2r+1}\right)^{-1} \gamma^1_{j,2r+1}||_\infty \leq \mathcal{K}_1 2^{-2rj}.$

Coming back to (4.14), in order to conclude the proof, we want to show that the norm of the second term tends to 0 when $j \to +\infty$ and that the first $2r$ components of the first term are equal to $(L^{r,r}_{-r}, ..., L^{r,r}_{r-1})$. This is achieved following the two next steps.

- Step 2: Second term of (4.14)

From (4.13) and (4.15), it is straightforward that:

$$|| \sum_{i=1}^{\infty} (-1)^i \left(\left(\Gamma^0_{j,2r+1}\right)^{-1} \Gamma^1_{j,2r+1} \right)^i \left(\Gamma^0_{j,2r+1}\right)^{-1} \gamma^0_{j,2r+1}||_\infty \leq \mathcal{K}_0 \sum_{i=1}^{\infty} (K_{\Gamma^0} K_{\Gamma^1})^i 2^{-2rij},$$
$$|| \sum_{i=0}^{\infty} (-1)^i \left(\left(\Gamma^0_{j,2r+1}\right)^{-1} \Gamma^1_{j,2r+1} \right)^i \left(\Gamma^0_{j,2r+1}\right)^{-1} \gamma^1_{j,2r+1}||_\infty \leq \mathcal{K}_1 2^{-2rj} \sum_{i=0}^{\infty} (K_{\Gamma^0} K_{\Gamma^1})^i 2^{-2rij},$$

which becomes for sufficiently large j,

$$|| \sum_{i=1}^{\infty} (-1)^i \left(\left(\Gamma^0_{j,2r+1}\right)^{-1} \Gamma^1_{j,2r+1} \right)^i \left(\Gamma^0_{j,2r+1}\right)^{-1} \gamma^0_{j,2r+1}||_\infty \leq \mathcal{K}_0 K_{\Gamma^0} K_{\Gamma^1} \frac{2^{-2rj}}{1 - K_{\Gamma^0} K_{\Gamma^1} 2^{-2rj}},$$
$$|| \sum_{i=0}^{\infty} (-1)^i \left(\left(\Gamma^0_{j,2r+1}\right)^{-1} \Gamma^1_{j,2r+1} \right)^i \left(\Gamma^0_{j,2r+1}\right)^{-1} \gamma^1_{j,2r+1}||_\infty \leq \mathcal{K}_1 \frac{2^{-2rj}}{1 - K_{\Gamma^0} K_{\Gamma^1} 2^{-2rj}}.$$

Therefore

$$\begin{aligned} \lim_{j\to\infty} \quad || \quad & \sum_{i=1}^{\infty} (-1)^i \left(\left(\Gamma^0_{j,2r+1}\right)^{-1} \Gamma^1_{j,2r+1} \right)^i \left(\Gamma^0_{j,2r+1}\right)^{-1} \gamma^0_{j,2r+1} \\ + \quad & \sum_{i=0}^{\infty} (-1)^i \left(\left(\Gamma^0_{j,2r+1}\right)^{-1} \Gamma^1_{j,2r+1} \right)^i \left(\Gamma^0_{j,2r+1}\right)^{-1} \gamma^1_{j,2r+1}||_\infty = 0. \end{aligned}$$

- Step 3: First term of (4.14)

Our goal is to show that the linear equation

$$\Gamma^0_{j,2r+1} U = \gamma^0_{j,2r+1} \tag{4.16}$$

has a unique solution whose first $2r$ components are $(L^{r,r}_{-r}, ..., L^{r,r}_{r-1})$. This will also ensure the invertibility of $\Gamma^0_{j,2r+1}$ which is the underlying assumption used in this proof.

Let us first note that $\sum_{m=-r}^{r-1} L^{r,r}_m = 1$. If $(L^{r,r}_{-r}, ..., L^{r,r}_{r-1})$ are the first $2r$ components of a solution, then the Lagrange multiplier denoted μ in the sequel satisfies the $2r$ following equations:

$$
(4.17)\quad \begin{cases} \mu = \mu_0 = P_{2r,j}(\frac{2r-1}{2}) - \sum_{m=-r}^{r-1} L_m^{r,r} P_{2r,j}(m+r), \\ \mu = \mu_1 = P_{2r,j}(\frac{2r-3}{2}) - \sum_{m=-r}^{r-1} L_m^{r,r} P_{2r,j}(m+r-1), \\ ... \\ \mu = \mu_{2r-1} = P_{2r,j}(-\frac{2r-1}{2}) - \sum_{m=-r}^{r-1} L_m^{r,r} P_{2r,j}(m+r-2r+1), \end{cases}
$$

It then remains to verify that $\forall\ 0 \leq n \leq 2r-2$, $\mu_n - \mu_{n+1} = 0$. Let us focus on $n = 0$, the proof being the same for $n > 0$. Using the expression of P_{2r}, one can write:

$$
\begin{aligned}
\mu_0 - \mu_1 = \sum_{n=0}^{r-1} 2^{-j(2n+2)} b_n \Bigg(& \left(\frac{2r-1}{2}\right)^{2n+2} - \left(\frac{2r-3}{2}\right)^{2n+2} \\
& - \sum_{m=-r}^{r-1} L_m^{r,r} ((m+r)^{2n+2} - (m+r-1)^{2n+2}) \Bigg), \\
= \sum_{n=0}^{r-1} 2^{-j(2n+2)} b_n & \left(R_{2n+1}\left(\frac{2r-1}{2}\right) - \sum_{m=-r}^{r-1} L_m^{r,r} R_{2n+1}(m+r) \right),
\end{aligned} \tag{4.18}
$$

where $R_{2n+1}(x) = x^{2n+2} - (x-1)^{2n+2}$ is a polynomial of degree $(2n+1)$ since after expansion of $(x-1)^{2n+2}$, the terms of order $(2n+2)$ are cancelled. Keeping in mind that $(L_{-r}^{r,r}, ..., L_{r-1}^{r,r})$ are the Lagrange weights associated to the $2r$-points centered interpolatory stencil (Formula (2.4)),for $n \leq r-1$,

$$
\sum_{m=-r}^{r-1} L_m^{r,r} R_{2n+1}(m+r) = R_{2n+1}\left(\frac{2r-1}{2}\right).
$$

Therefore, Expression (4.18) leads to:

$$
\mu_0 - \mu_1 = 0.
$$

In order to conclude this step, we now prove that $(L_{-r}^{r,r}, ..., L_{r-1}^{r,r}, \mu)$ with μ satisfying (4.17) is the unique solution of (4.16).

Define $(U_0, ..., U_{2r-1}, \tilde{\mu})$ as another solution of (4.16) and introduce $\mathcal{V} = (\Delta_0, ..., \Delta_{2r-1}, \Delta\mu)$ the vector of differences between the two solutions and satisfying,

$$
\Gamma_{j,2r+1}^0 \mathcal{V} = 0. \tag{4.19}
$$

Our goal is to show that $\mathcal{V} = 0$.

Let us consider the polynomial

$$K(x) = \Delta_0 P_{2r,j}(x) + \Delta_1 P_{2r,j}(x-1) + ... + \Delta_{2r-1} P_{2r,j}(x-(2r-1)) + \Delta\mu.$$

Note that since $\sum_{i=0}^{2r-1} \Delta_i = 0$ (according to the unbiasedness constraint), the degree of K is $2r-1$. Keeping in mind the expression of $\Gamma^0_{j,2r+1}$, Equation (4.19) means that K has $2r$ roots leading to $K = 0$. By identification and remembering the expression of P_{2r}, we write that each coefficient associated to each power is equal to 0:

$$\begin{cases} \text{for } x^{2r-1} : \sum_{i=0}^{2r-1} i\Delta_i = 0, \\ \text{for } x^{2r-2} : \sum_{i=0}^{2r-1} i^2\Delta_i = 0, \\ ... \\ \text{for } x^2 : \sum_{i=0}^{2r-1} i^{2r-2}\Delta_i = 0, \\ \text{for } x : \sum_{i=0}^{2r-1} i^{2r-1}\Delta_i = 0. \end{cases}$$

Finally, since $\sum_{i=0}^{2r-1} \Delta_i = 0$, $(\Delta_0, ..., \Delta_{2r-1})$ can be interpreted as belonging to the kernel of a $2r \times 2r$ Vandermonde matrix which is invertible. It leads to $\Delta_i = 0,\ \forall i \in \{0, ..., 2r-1\}$, then to $\Delta\mu = 0$ and finally to $\mathcal{V} = 0$. That concludes Step 3. ∎

Proposition 4.6 finally leads to the following convergence result:

Proposition 4.7.
The translation-invariant and position-dependent subdivision scheme of parameter (D, r, r) associated to a $\mathcal{G}^\star_{OD}$-type semi-variogram is uniformly convergent.

Proof. Since the Lagrange interpolatory subdivision scheme with the same translation-invariant and position-dependent strategies as for the Kriging-based one has been proved to uniformly converge in [7], Proposition 4.6 and Theorem 4.4 allows one to conclude. ∎

4.3.3. $\mathcal{G}^\star_{ED}$-type semi-variograms. The following proposition holds:

Proposition 4.8.
When a $\mathcal{G}^\star_{ED}$-type semi-variogram is considered, the Kriging weights involved in the translation-invariant and position-dependent subdivision schemes satisfy:

$$\begin{gathered} (\lambda^{r,r}_{j,-r}, ..., \lambda^{r,r}_{j,r-1}) \to_{j\to+\infty} \tfrac{1}{2}\delta_{r-1} + \tfrac{1}{2}\delta_r \\ (\lambda^{r-1,r+1}_{j,-r+1}, ..., \lambda^{r-1,r+1}_{j,r}) = (\lambda^{r+1,r-1}_{j,r-2}, ..., \lambda^{r+1,r-1}_{j,-r-1}) \to_{j\to+\infty} \tfrac{1}{2}\delta_{r-2} + \tfrac{1}{2}\delta_{r-1}, \\ ... \\ (\lambda^{1,2r-1}_{j,-1}, ..., \lambda^{1,2r-1}_{j,2r-2}) = (\lambda^{2r-1,1}_{j,0}, ..., \lambda^{2r-1,1}_{j,-2r+1}) \to_{j\to+\infty} \tfrac{1}{2}\delta_0 + \tfrac{1}{2}\delta_1, \\ (\lambda^{0,2r}_{j,0}, ..., \lambda^{0,2r}_{j,2r-1}) = (\lambda^{2r,0}_{j,-1}, ..., \lambda^{2r,0}_{j,-2r}) \to_{j\to+\infty} \delta_0. \end{gathered}$$

where δ_{k_1} is the $2r$-vector defined by $\delta_{k_1,m} = 1$ for $m = k_1$ and 0 otherwise.

Proof. The underlying idea is similar to the proof of Proposition 4.6. Focussing on the first limit, we start by writing the expression of $\mathcal{G}^{\star}_{ED}$-type semi-variograms for small h:

$$\gamma(h) = a_0 h + \delta_\gamma \; O(h^2),$$

where δ_γ is equal to 0 or 1 depending on the chosen semi-variogram[6].
Then, the Kriging system becomes:

$$\left(\Gamma^0_{j,2r+1} + \delta_\gamma \Gamma^1_{j,2r+1}\right) \mathcal{U}_{j+1} = \gamma^0_{j,2r+1} + \delta_\gamma \gamma^1_{j,2r+1}, \tag{4.20}$$

where,

$$\Gamma^0_{j,2r+1} = \begin{bmatrix} 0 & a_0 2^{-j} & ... & (2r-1)a_0 2^{-j} & 1 \\ a_0 2^{-j} & 0 & ... & (2r-2)a_0 2^{-j} & 1 \\ ... & ... & ... & ... & ... \\ (2r-1)a_0 2^{-j} & (2r-2)a_0 2^{-j} & ... & 0 & 1 \\ 1 & 1 & 1 & ... & 0 \end{bmatrix},$$

$$\text{and } \gamma^0_{j,2r} = \begin{bmatrix} \frac{2r-1}{2} a_0 2^{-j} \\ \frac{2r-3}{2} a_0 2^{-j} \\ ... \\ \frac{2r-1}{2} a_0 2^{-j} \\ 1 \end{bmatrix}.$$

Similarly to the proof of Proposition 4.6, the solution of (4.20) is first decomposed in two terms (Step 1 leading to Expression (4.14) where $\Gamma^1_{j,2r+1}$ and $\gamma^1_{j,2r+1}$ are replaced by $\delta_\gamma \Gamma^1_{j,2r+1}$ and $\delta_\gamma \gamma^1_{j,2r+1}$ respectively) then one proves that the second term tends to 0 (Step 2). To conclude this proof, it remains to verify that $(\tilde{U}, 0)$ with $\tilde{U} = \frac{1}{2}\delta_{r-1} + \frac{1}{2}\delta_r$ is solution of $\Gamma^0_{j,2r+1}\tilde{U} = \gamma^0_{j,2r+1}$. After a short calculus, one gets the expected result. Note that the uniqueness of this solution is ensured since $\Gamma^0_{j,2r+1}$ is by construction invertible (Kriging matrix associated to the valid linear semi-variogram model). This concludes the proof for all the limits but the last one. In this case, the asymptotical weights are no more the Lagrange ones associated to a two-point stencil for extrapolation. However, a straightforward calculus shows that $(\tilde{U}, \frac{1}{2}a_0)$ with $\tilde{U} = \delta_0$ is solution of the Kriging system for large enough j. ∎

We then have:

[6] For example, $\delta_\gamma = 0$ (resp. $= 1$) for a linear (resp. exponential) model.

PROPOSITION 4.9.
The translation-invariant and position-dependent subdivision scheme of parameter (D, r, r) associated to a $\mathcal{G}^{\star}_{ED}$-type semi-variogram is uniformly convergent.

Proof. The uniform convergence comes again from the uniform convergence of the stationary asymptotical subdivision scheme with the same translation-invariant and position-dependent strategies. Proposition 4.8 combined with Theorem 4.4 allows one to conclude. ∎

5. Numerical tests. This section is devoted to three numerical tests. In the first one, we numerically verify the convergence results of the previous section (Propositions 4.6 and 4.8). In the second and third ones the efficiency of Kriging-based subdivision schemes integrating a position-dependent strategy is illustrated.

5.1. Limit functions of the Kriging-based subdivision scheme. We are focussing on the comparison between the limit functions associated to Kriging subdivision scheme and the ones associated to the corresponding stationary Lagrange scheme (Propositions 4.6 and 4.8) following the same position-dependent strategy (Definition 2.2). For sake of simplicity, we are interested in the limiting process starting from an initial sequence $f^{J_0} = \{\delta_{k,m}\}_{k\in\mathbb{Z}}$, $m \in \mathbb{Z}$ with the parameters characterizing the position-dependence fixed to $D = 3,\ r = 2,\ l = 2$ and $S_s = \{0.5\}$ (the single segmentation point 0.5 separates the interval $[0, 1]$ into two zones).Moreover, we assume that the data exhibit the following structure:

$$\begin{cases} 1 - e^{-ah}\ (\mathcal{G}^{\star}_{ED}\text{-type semi-variogram}),\ \text{in Zone 1}= [0; 0.5] \\ 1 - e^{-(ah)^2}\ (\mathcal{G}^{\star}_{OD}\text{-type semi-variogram}),\ \text{in Zone 2}=]0.5; 1] \end{cases}$$

with $a \in I\!R$. In order to verify our convergence result dealing with the weights of the Kriging subdivision schemes, the limit functions of the Kriging-based subdivision scheme are calculated for different values of a and compared to the Lagrange interpolatory scheme limit functions. A direct consequence of our convergence result is indeed the convergence of the Kriging-based subdivision scheme limit functions towards the Lagrange interpolatory scheme ones when $a \to 0$.
Figures 5.1 and 5.2 display a comparison between some limit functions associated to both subdivision schemes with respect to a. The initial sequence is defined by $J_0 = 5$, $m = 15$ (exponential semi-variogram) and $m = 20$ (Gaussian semi-variogram) respectively.

As expected, when $a \to 0$, each limit function of the position dependent Kriging-based subdivision scheme converges to the limit function of the corresponding Lagrange scheme.

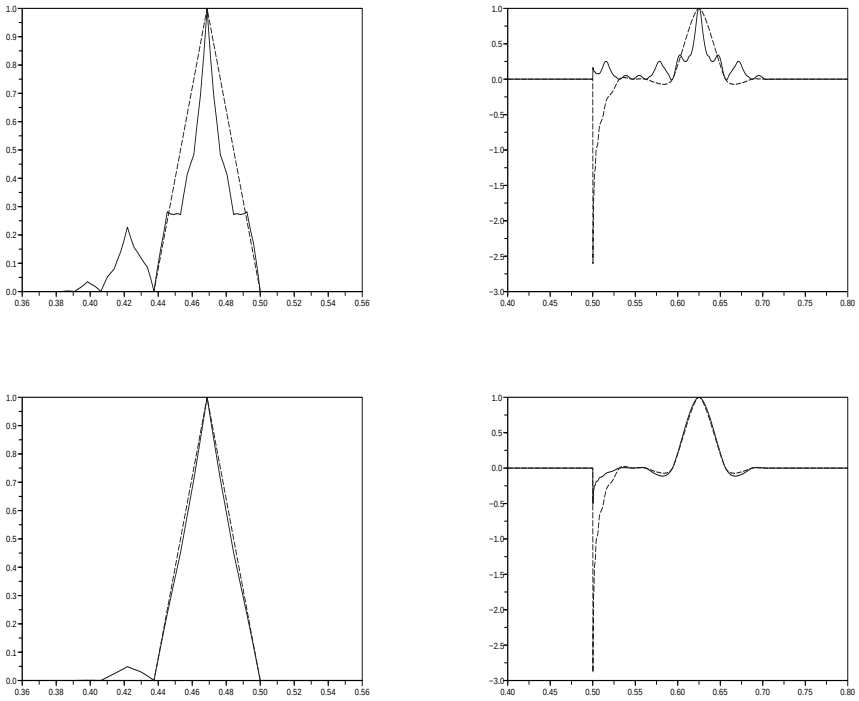

FIG. 5.1. *Examples of limit functions associated to Kriging (solid line) and corresponding Lagrange interpolatory (dashed line) subdivision schemes. For sake of clarity, a zoom of these functions is performed by restriction to an interval containing their support. Left column, exponential case,* $m = 15$ *(Zone 1), from top to bottom,* $a = 0.2$, $a = 0.04$. *Right column, Gaussian case,* $m = 20$ *(Zone 2), from top to bottom,* $a = 0.2$, $a = 0.02$.

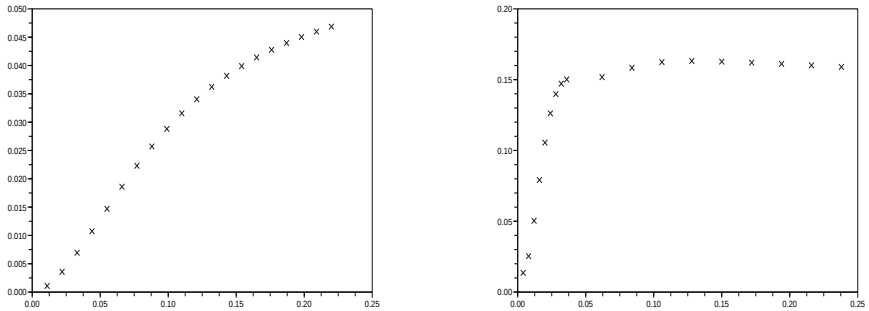

FIG. 5.2. l^2 *error with respect to a between the limit functions associated to Kriging and corresponding Lagrange interpolatory subdivision schemes. Left, exponential case, right, Gaussian case. The crosses correspond to the different values of a for which the error has been evaluated.*

5.2. Data prediction of a univariate discontinuous function. Kriging and Lagrange interpolatory subdivision schemes are here compared on univariate discontinuous data. The test function (thick solid line on Figure 5.3) is defined on the interval $[0, 1000]$ by:

$$f(x) = \left(-2 * \frac{sin(\frac{30}{1000}x)}{2 + \frac{x}{1000}} + 2 \right) \chi_{[255,475]} - \left(2 * \frac{sin(\frac{8}{1000}x)}{1 + \frac{x}{1000}} \right) \chi_{[0,1000]\setminus[255,475]}$$

For sake of simplicity, in the case of a Kriging strategy, we have chosen to decouple the problem of the semi-variogram identification and the subdivision scheme-based prediction. For this simplified univariate example, the semi-variogram in each zone is therefore estimated by considering a sufficiently fine discretization of the test function. Moreover, since our goal is to focus on prediction, we assume for the moment that the segmentation points $\{255, 475\}$ are known at each scale j. The questions related to semi-variogram identification from the initial data and to data segmentation are postponed to Section 5.3 where a bivariate case study is considered.

We apply a position-dependent Kriging and Lagrange interpolatory predictions between $J_0 = 4$ and $J_{max} = 9$.

Table 5.1 provides the l^2-error corresponding to each type of subdivision scheme and Figure 5.3 displays the predicted and exact data in order to visualize the difference between both approaches around the segmentation points.

Method	l^2-error
Lagrange	0.4
Kriging	0.1

TABLE 5.1

l^2-error between predicted and exact data provided by position-dependent Kriging and Lagrange interpolatory subdivision schemes.

It comes out that the position-dependent Kriging-based strategy improves the prediction (the l^2-error is reduced by a factor 4). As for the data in the vicinity of the segmentation points, a strong undershoot appears at $x = 250$ in the Lagrange case whereas it is reduced with the Kriging-based prediction. A similar behavior has already been noticed in the limit function of Figure 5.1, top, right. Even if it has been shown that the two schemes have the same asymptotical behavior, there is a big difference between them for small j especially in the extrapolation process: Lagrange interpolation assigns large weights (in absolute value) to each point of the interpolatory stencil (i.e. $(\frac{35}{16}, -\frac{35}{16}, \frac{21}{16}, -\frac{5}{16})$ in our example) while for Kriging with Gaussian semi-variogram, since it assumes a strong spatial dependence, most of the total

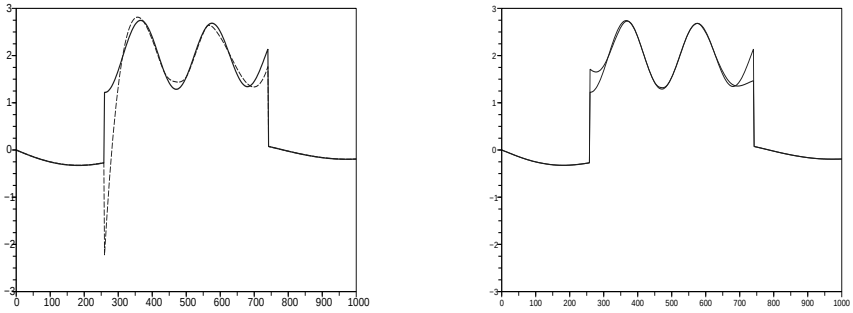

FIG. 5.3. *Predicted data ($J_0 = 4$, $J_{max} = 9$). Left, position-dependent Lagrange (dashed line) interpolation, right, Kriging (solid line) interpolation. The thick solid line corresponds to the exact surface.*

weight is given to the closest point to the predicted one.
Moreover, the improvement of the prediction is not restricted to the area around the segmentation points. Indeed, thanks to the construction of a semi-variogram in each zone, the structure of the data is faithfully identified leading to a better prediction around $x = 470$ for example.

5.3. Data prediction of a bivariate discontinuous surface. The previous univariate construction is extended to the bivariate framework and the following test function (Figure 5.4, top) is considered on the square $[0, 1000]^2$:

$$f(x,y) = \begin{cases} 10\frac{sin(\frac{30}{1000}x)}{2+\frac{x}{1000}}\frac{sin(\frac{10y}{1000})}{5+\frac{y}{1000}} + 2, \text{ in Zone } 1{=}\mathcal{C} \\ 2\frac{sin(\frac{8}{1000}x)}{1+\frac{x}{1000}}\frac{sin(\frac{2y}{1000})}{8+\frac{y}{1000}}, \text{ in Zone } 2{=}[0,1000]^2\backslash\mathcal{C}, \end{cases}$$

where $\mathcal{C} = \{(x,y) \in [0,1000]^2/(\frac{x}{1000} - 0.5)^2/3.0 + (\frac{y}{1000} - 0.5)^2/3.0 \leq 0.02\}$. Contrarily to the previous section, it is assumed that the only available information is provided by the data at the coarse level J_0 where the semi-variogram identification and the segmentation curve detection are performed. Our position-dependent Kriging-based subdivision scheme is here applied from $J_0 = 4$ to $J_0 = 9$. Therefore, we first focus in the sequel on the estimation of the semi-variogram for each zone.

5.3.1. Semi-variogram identification. This identification is performed using the SUNSET software ([10]) developed at IRSN. The two experimental semi-variograms associated to each zone have been first computed from the available data corresponding to the discretization of the test function on the dyadic grid $X^{J_0} = \{(1000 * m2^{-J_0}, 1000 * n2^{-J_0}), (m,n) \in \{0, ..., 2^{J_0}\}^2\}$ with $J_0 = 4$. They are plotted using cross signs on Figure 5.4 bottom. Then, two Gaussian models of type $\gamma(h) = c\left(1 - e^{-(ah)^2}\right)$ have been fit by a least square

method (see Table 5.2 and Figure 5.4).

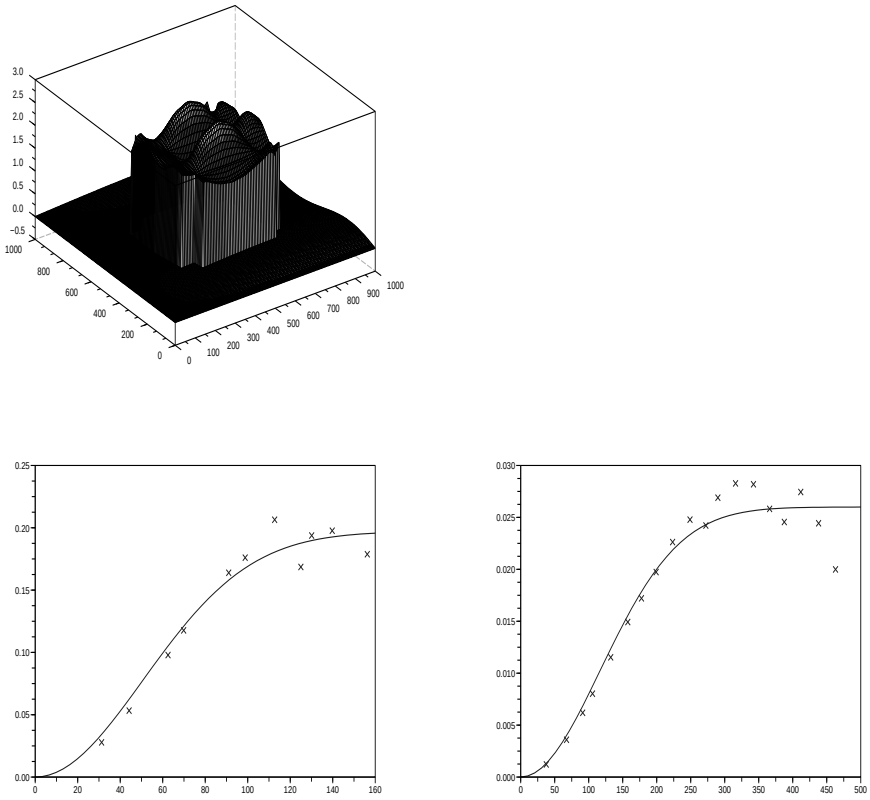

FIG. 5.4. *Top: Test function. Bottom: semi-variograms of Zone 1 (left) and Zone 2 (right).*

Parameters	c	a
Zone 1	0.17	0.014
Zone 2	0.01	0.003

TABLE 5.2
Estimated parameters (c, a) for the two Gaussian semi-variograms.

Obviously, the data exhibit two different spatial structure in each zone. Even if the behavior of the semi-variogram is smooth at the origin for both zones, it is interesting to notice that 95% of the asymptotical sill (the value

of the limit of $\gamma(h)$ when h goes to $+\infty$) is reached when $h = 120$ (resp. $h = 500$) for Zone 1 (resp. 2). This translates the stronger variation at small scales associated to the test function in Zone 1.

5.3.2. Bivariate Kriging-based prediction. The prediction is here performed using a bivariate Kriging-based subdivision scheme constructed by a tensorial product of two univariate schemes ([2], [7]).
Figure 5.5, top, displays two cuts of the predicted data and Table 5.3 (first column) provides the associated l^2-error. Since the segmentation curve is not exactly known at each scale, it appears that the misleading localization of the segmentation curve for $j \geq J_0$ strongly deteriorates the accuracy of the prediction.

Initial grid	X^{J_0}, $J_0 = 4$ (289 points)	X^{J_0,J_0+1}, $J_0 = 4$ (349 points)	X^{J_0}, $J_0 = 5$ (1089 points)
l^2-error	0.42	0.29	0.21

TABLE 5.3
l^2-error with respect to the initial grid.

5.3.3. Grid refinement. In order to circumvent this limitation, we propose a grid refinement algorithm based on the Kriging estimation variance. More precisely, noticing that the estimation variance given by Expression (3.11) is bigger along the segmentation curve (Figure 5.6, left) due to adaptation of the stencil, the algorithm with grid refinement reads as follows:

1) Starting from the initial coarse grid, X^{J_0}, apply the position-dependent Kriging-based prediction until $J_0 + 1$ and compute the associated estimation variance,
2) Construct a new refined grid X^{J_0,J_0+1} by adding to the original coarse grid new points corresponding to the grid points of X^{J_0+1} with the maximal estimation variance, i.e. larger than a given threshold T_h,
3) From the data associated to X^{J_0,J_0+1}, apply the Kriging-based prediction to get the predicted values at level $j > J_0 + 1$.

Figure 5.6, right, provides the refined grid X^{J_0,J_0+1} for $T_h = 0.014$. Moreover, Figure 5.5, bottom, displays cross sections along $x = 375$ and $y = 500$ of the predicted data; the corresponding l^2-error is given by Table 5.3 (second column).

As expected, the adaption of the grid by the addition of 60 points allows one to reduce the l^2-error (of a factor 1.5). In this way, by iterating this algorithm through the scales j, one can limit the impact of a misleading localization of the segmentation curve. Moreover, comparing with the third

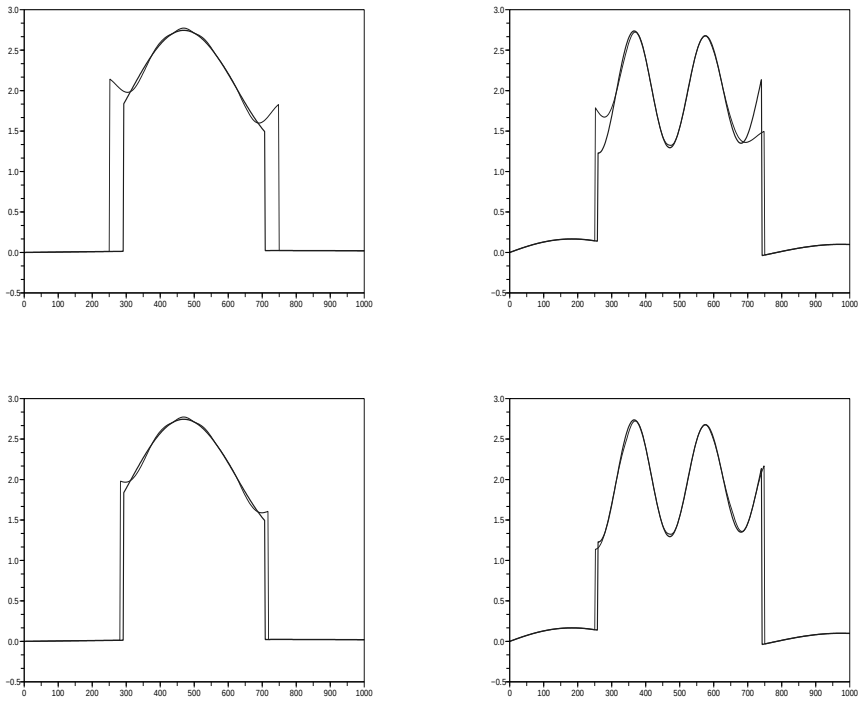

FIG. 5.5. *Cross section of the prediction surface using a position-dependent Kriging-based subdivision scheme (Left $x = 375$, Right $y = 500$), $J_0 = 4$, $J_{max} = 9$. Top, without grid refinement, bottom, with grid refinement. The thin (resp. thick) solid line stand for the predicted (resp. exact) surface.*

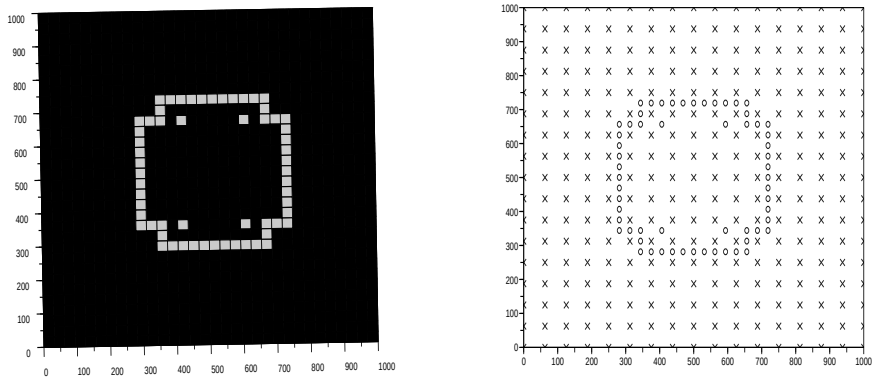

FIG. 5.6. *Grid refinement. Left, estimation variance $\left(\sigma^{J_0+1}\right)^2$ ($J_0 = 4$) (the grey squares stand for the predicted points with an estimation variance larger than 0.014) , right, refined grid X^{J_0,J_0+1} (cross sign for the coarse grid and circle one for the 60 extra points corresponding to the grey squares of the left-hand plot).*

column of Table 5.3, it comes out that the regular grid associated to $J_0 = 5$ and the non-regular one generated by our algorithm lead to similar l^2-error. Therefore, by automatically refining the information in some region of interest (such as where the estimation variance is large), it has been possible to reach the same accuracy in the prediction with a gain of 3 in the number of points in the initial grid. This is an important result for real applications such as risk analysis studies where considering a fine discretization grid is not technically or economically affordable.

REMARK 5.1.
- The previous algorithm is generic and can be applied to any functions exhibiting discontinuities as soon as the data segmentation has been properly performed and the spatial structure in each zone has been identified through the estimation of a semi-variogram. Moreover, it is straightforward to iteratively extend the underlying procedure in order to increase the number of extra points in the new grid coming from finer level, i.e. $j > J_0 + 1$ *if we are interested in a more accurate prediction in the vicinity of the segmentation curve.*
- This last result points out an important advantage of Kriging-based approach. The computation of the estimation variance associated to the prediction provides an extra-information that can be easily integrated to improve the reconstruction by pointing out the zones where to refine the set of known values: it leads to adaptive design of experiments ([21], [20]) . Moreover, this variance is of great interest in practice such as in risk analysis studies since it controls the confidence that one can have in the predicted values ([8]).

6. Conclusion. A new type of subdivision schemes adapted to data and to a partition of their support has been introduced. The originality of its construction lies in the coupling of position-dependent multi-scale approximation and Kriging theory. In this way, it exploits the advantages coming from both approaches:

- the integration of the data spatial dependence (through the estimation of a semi-variogram) in the computation of the mask leading to a relevant prediction,
- the estimation of an a priori prediction error (through the computation of an estimation variance that can be used to exhibit confidence intervals) allowing one to quantify the uncertainty associated to each predicted point and that can be iteratively used in a grid refinement procedure,
- the flexibility in the choice of the interpolatory stencil ensuring a position-dependent strategy to faithfully reconstruct piecewise-regular data.

The proposed Kriging-based subdivision schemes have been fully analyzed by studying their asymptotical limits and extending classical results of convergence for subdivision scheme to this new framework. The numerical tests and the comparisons to the position-dependent Lagrange interpolatory scheme

introduced in [7] have shown a significant improvement for the prediction of non-regular functions which remains a key issue in many industrial problems.

REFERENCES

[1] J.H. Ahlberg, E.N. Nilson, and J.L. Walsh. *The theory of splines and their applications*. Academic Press, 1967.

[2] S. Amat, F. Arandiga, A. Cohen, and R. Donat. Tensor product multiresolution analysis with error control for compact image representation. *Signal Processing*, 4:587–608, 2002.

[3] S. Amat, F. Arandiga, A. Cohen, R. Donat, G. Garcia, and M. von Oehsen. Data compression with ENO schemes: a case study. *Appl. Comp. Harm. Anal.*, 11:273–288, 2001.

[4] F. Arandiga, J. Baccou, M. Doblas, and J. Liandrat. Image compression based on a multi-directional map-dependent algorithm. *Appl. and Comp. Harm. Anal.*, 23(2):181–197, 2007.

[5] F. Arandiga, R. Donat, and A. Harten. Multiresolution based on weighted averages of the hat function I: linear reconstruction techniques. *SIAM J. Sci. Comput.*, 36(1):160–203, 1999.

[6] F. Arandiga, R. Donat, and A. Harten. Multiresolution based on weighted averages of the hat function II: non-linear reconstruction techniques. *SIAM J. Sci. Comput.*, 20(3):1053–1093, 1999.

[7] J. Baccou and J. Liandrat. Position-dependent lagrange interpolating multiresolutions. *Int. J. of Wavelet, Multiresolution and Information*, 5(4):513–539, 2005.

[8] J. Baccou and J. Liandrat. Kriging-based subdivision schemes: application to the reconstruction of non-regular environmental data. *Math. Comput. Simul.*, 81-10:2033–2050, 2011.

[9] A.S. Cavaretta, W. Dahmen, and C.A. Micchelli. Stationary subdivision. In *Memoirs of the American Mathematics Society*, volume 93 (453). Providence, Rhode Island, 1991.

[10] E. Chojnacki and A. Ounsy. Description of the ipsn method for the sensitivity analysis and the associated software: Sunset. *Proc. of AMSE/JSME ICONE*, 3:545–550, 1996.

[11] A. Cohen, N. Dyn, and B. Matei. Quasilinear subdivision schemes with applications to ENO interpolation. *Appl. Comp. Harm. Anal.*, 15:89–116, 2003.

[12] N.A. Cressie. *Statistics for spatial data*. Wiley Series in Probability and Mathematical Statistics, 1993.

[13] G. Deslauriers and S. Dubuc. Interpolation dyadique. In *Fractals, dimensions non entières et applications*, pages 44–45. Masson, Paris, 1987.

[14] N. Dyn. Subdivision schemes in computer-aided geometric design. In W.A Light, editor, *Advances in Numerical analysis II, Wavelets, Subdivision algorithms and Radial Basis functions*. Clarendon Press, Oxford, 1992.

[15] N. Dyn, J.A. Gregory, and D. Levin. A four-point interpolatory subdivision scheme for curve design. *Comp. Aid. Geom. Des.*, 4:257–268, 1987.

[16] N. Dyn, J.A. Gregory, and D. Levin. Analysis of linear binary subdivision schemes for curve design. *Constr. Appr. 7*, 2:127–148, 1991.

[17] G. Matheron. The intrinsic random function and their applications. *Advances in Applied Probability*, 3:439–468, 1986.

[18] C.A. Miccheli. Interpolation of scattered data: Distance matrices and conditionally positive definite functions. *Constr. Appr.*, 2:11–22, 1986.

[19] C.A. Miccheli and H. Prautzsch. Uniform refinement of curves. *Linear Algebra and Applications*, 114/115:841–870, 1989.

[20] V. Picheny, D. Ginsbourger, O. Roustant, and R. T. Haftka. Adaptive designs of

experiments for accurate approximation of a target region. *J. of Mechanical Design*, 132(7), 2010.

[21] J. Sacks, W. J. Welch, T. J. Mitchell, and P. Wynn. Design and analysis of computer experiments. *Statistical Science*, 4(4):409–423, 1989.

[22] G.W Stewart and J.G Sun. *Matrix perturbation theory*. Accademic press, 1990.

[23] H. Wackernagel. *Multivariate geostatistics*. Springer, 1998.