Full inference for the anisotropic fractional Brownian field

Paul Escande, Frédéric Jp Richard

Theor. Probability and Math. Statist.
No. 0, XXXX, Pages 000–000

FULL INFERENCE FOR THE ANISOTROPIC FRACTIONAL BROWNIAN FIELD

UDC 517.927.25+517.928.5

PAUL ESCANDE AND FRÉDÉRIC RICHARD

Abstract. The anisotropic fractional Brownian field (AFBF) is a non-stationary Gaussian random field which has been used for the modeling of textured images. In this paper, we address the open issue of estimating the functional parameters of this field, namely the topothesy and Hurst functions. We propose an original method which fits the empirical semi-variogram of an image to the semi-variogram of a turning-band field that approximates the AFBF. Expressing the fitting criterion in terms of a separable non-linear least square criterion, we design a minimization algorithm inspired from the variable projection approach. This algorithm also includes a coarse-to-fine multigrid strategy based on approximations of functional parameters. Compared to existing methods, the new method enables to estimate both functional parameters on their whole definition domain. On simulated textures, we show that it has a low estimation error, even when the parameters are approximated with a high precision. We also apply the method to characterize mammograms and sample images with synthetic parenchymal patterns.

1. Introduction

The anisotropic fractional Brownian field (AFBF) is a non-stationary Gaussian random field with stationary increments; see [10] and Section 2 for details. The finite-dimensional probabability distributions of an AFBF Z are completely determined by a semi-variogram of the form

$$\begin{aligned} \forall\, h \in \mathbb{R}^2,\ v_0(h;\tau,\beta) &= \frac{1}{2}\mathbb{E}\left(\left(Z(x+h)-Z(x)\right)^2\right), \\ &= \frac{1}{2}\int_{\mathbb{R}^2} \left|1-e^{i\langle h,w\rangle}\right|^2 \tau(\arg(w))|w|^{-2\beta(\arg(w))-2}dw. \end{aligned} \tag{1.1}$$

In this expression, the functions τ and β are the parameters of the AFBF. They are called the topothesy and Hurst functions, respectively. They are non-negative π-periodic functions whose values only depend on the direction $\arg(w)$ of w in $\mathbb{R}^2$. The Hurst function β further ranges in $(0,1)$. In this paper, we propose an original method to estimate these functional parameters from a single realization of an AFBF.

The AFBF has been used for the modeling of image textures [8, 9, 33, 35, 34, 36, 38]. Figure 1 shows some textures together with the Hurst and topothesy functions of the AFBF they are generated from. On the first column, constant topothesy and Hurst functions are associated to an isotropic texture, *i.e.* a texture whose properties

2020 *Mathematics Subject Classification.* Primary 62M40; Secondary 78M50.

Key words and phrases. inference from random field, anisotropic fractional Brownian fields, image texture analysis, image texture synthesis.

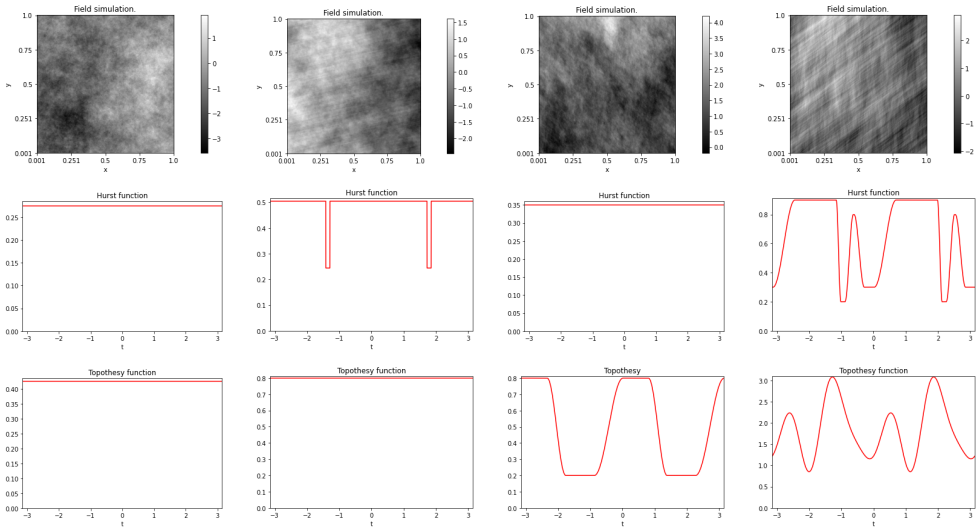


FIGURE 1. Textures generated from AFBF. First row: the realization of the AFBF, second and third rows: representations of their hurst and topothesy functions, respectively. The textures were generated using the PyAFBF package [6, 37].

are invariant to any image rotation. By contrast, as observed on other columns, non-constant topothesy or Hurst functions result in anisotropic textures whose properties vary according to the direction. Let us also notice that the texture anisotropy manifests differently depending on the shape of the topothesy and Hurst functions.

Actually, these two functional parameters are critical to characterize the texture. Such a characterization may serve for the classification of textures. In [33], a feature was derived from the topothesy function to measure a degree of texture anisotropy. This feature was used in conjunction with a measure of texture regularity to classify photographic films with respect to their paper type. In [35, 34], another measure of anisotropy depending both on the topothesy and the Hurst functions was designed and applied to detect or diagnose lesions in mammograms. The AFBF can be further extended to an heterogeneous field where the topothesy and Hurst function both vary in space [44, 32, 5]. Due to this model, we can also consider characterizing locally the texture through the topothesy and Hurst functions. Such a characterization can be used to segment images, *i.e.* partition the image into several regions of homogeneous textures. In [36], biological microscopic images were segmented using local features related to the topothesy and Hurst functions. The work presented in this paper is mainly motivated by these applications of AFBF to the classification or the segmentation of image textures. By proposing a method to estimate topothesy and Hurst functions, we enable the computation of texture descriptors that are critical to these two image-processing tasks. For instance, descriptors may include variances of the topothesy and Hurst functions that would measure a degree of anisotropy.

Another important motivation concerns the simulation of textures. In [6], AFBF were simulated with a method which was later implemented in a Python package [37]. Currently, this package enables to generate textures from AFBF whose parameters are randomly sampled. One of our goals would be to extend this package to the simulation

of realistic textures. Such a simulation could be obtained from AFBF whose parameters were previously estimated from one or several real examples.

However, the realistic simulation, as well as the texture classification, require a complete estimation of the Hurst and topothesy functions. Methods proposed in the literature only give a partial solution to that estimation issue. In [7, 38], a method was developed for the estimation of the Hurst function. It relies upon a theoretical relationship that exists between an AFBF and its Radon projection in a direction orthogonal to θ [10]: in dimension $d = 2$, this projection is a fractional Brownian motion of Hurst index $H = \beta(\theta) + \frac{1}{2}$. Hence, by estimating the Hurst index of the projections, it is theoretically possible to recover the Hurst function of the field. However, in practice, such an approach is only feasible in the vertical and horizontal directions where the Radon transform can be discretized.

In [33], another method was designed for the estimation of the so-called asymptotic topothesy. This function is equal to the topothesy function in directions where the Hurst function is minimal, and vanishes elsewhere. The estimation approach is based on quadratic variations (averages of squared increments) computed from the image at different scales and in different directions. Asymptotically, the logarithm of these variations is linearly related to the Hurst index of the field (minimal value of the Hurst function). This linear relationship includes some intercepts that can be estimated by linear regression. These intercepts can further be expressed as a convolution of the asymptotic topothesy with some known functions. Due to this expression, the asymptotic topothesy can be recovered from the intercepts by solving an inverse problem.

The construction of our estimation method takes its root in the work in [6]. Following [6], the semi-variogram of an AFBF can be expressed in a polar form as

$$v_0(h;\tau,\beta) = \frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \tilde{\tau}(\theta)|\langle h, u(\theta)\rangle|^{2\beta(\theta)}d\theta, \tag{1.2}$$

where $u(\theta) = (\cos\theta, \sin\theta)$ and $\tilde{\tau}(\theta) = \tau(\theta)\int_0^{+\infty} \left|1 - e^{i\rho}\right|^2 \rho^{-2\beta(\theta)-1}d\rho$. Given some ordered angles $(\theta_m)_{m=1}^M$ in $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and appropriate positive weights $(\lambda_m)_{m=1}^M$, it can be further approximated by

$$v(h;\tilde{\tau},\beta) = \frac{1}{2}\sum_{m=1}^{M} \lambda_m\, \tilde{\tau}(\theta_m)|\langle h, u(\theta_m)\rangle|^{2\beta(\theta_m)}. \tag{1.3}$$

The function v corresponds to the semi-variogram of a turning band-field defined as

$$Z_M(x) = \sum_{m=1}^{M} \sqrt{\lambda_m\, \tilde{\tau}(\theta_m)}Y_m(\langle x, u(\theta_m)\rangle),$$

where the random processes Y_m are independent standard fractional Brownian motions of Hurst index $\beta(\theta_m)$. In [6], it was shown that v converges to v_0 and Z_M to Z as the maximal spacing between successive angles θ_m tends to 0. This result justifies using the turning-band field Z_M for simulating the AFBF Z and the semi-variogram v for approximating v_0. In this work, we thus propose to estimate the topothesy and Hurst functions of an AFBF by fitting the semi-variogram v of its approximating turning-band field to the empirical semi-variogram of a textured image.

2. Preliminary considerations about the AFBF

A random field is a collection $Z = \{Z(x), x \in \mathbb{R}^d\}$ of random variables $Z(x)$ indexed by a position x of the space $\mathbb{R}^d$ of dimension d. In dimension $d = 2$ or $d = 3$, such a field can represent an image, $Z(x)$ being in this case the grey-level value of the image at a pixel x.

A random field is Gaussian if any finite linear combination $\sum_i \lambda_i Z(x_i)$ of its variables is a Gaussian random variable. The finite-dimensional probability distributions of a Gaussian random field can be characterized by two functions: an expectation function $x \to \mathbb{E}(Z(x))$ giving the mean value of the field at each position x and the auto-covariance function $(x, y) \to \mathrm{cov}(Z(x), Z(y))$ describing the interactions of the field variables between pairs (x, y) of positions.

A Gaussian random field is stationary if its expectation function is constant and its auto-covariance function at (x, y) only depends on $x - y$. Such a field may describe images whose properties are spatially homogeneous. When it is not the case, we can resort to a weaker stationarity assumption which is stated on field increments rather than the field itself. An increment W of a field Z is a field defined for any position $x \in \mathbb{R}^d$ as $W(x) = Z(x + h) - Z(x)$ given some fixed lag $h \in \mathbb{R}^d$. We say that a field has stationary increments whenever its increment fields are stationary. The finite-dimensional probability distributions of such a field can be characterized by the semi-variogram $v(h) = \frac{1}{2}\mathbb{E}((Z(y + h) - Z(y))^2)$ of the field. The AFBF forms a family of random fields with stationary increments which is defined by semi-variograms of the form (1.1).

3. Problem formulation

The estimation context is the following. We observe the gray-level values $Y[i]$ of an image Y at points i of a grid $[\![1, I]\!]^2$. We assume that this image can be modeled as the sum

$$Y[i] = Z\left(\frac{i}{I}\right) + W[i]$$

of a turning-band field Z having a semi-variogram v of the form (1.3) and a noise W formed by independent centered Gaussian variables $W[i]$ of variance τ_0. It follows that the semi-variogram of the image is given by

$$w(h; \tau, \beta) = \tau_0 + v\left(\frac{h}{I}; \tau, \beta\right). \tag{3.1}$$

In this equation, the topothesy and Hurst functions τ and β are unknown. The problem is to estimate these functional parameters.

For that, we compute the empirical semi-variogram $(\widehat{w}_n)_{n=1}^N$ at some lags $(h_n)_{n=1}^N$ of $[\![1, I]\!]^2$ (see Section 4 for the choice of these lags):

$$\widehat{w}_n = \frac{1}{N_n} \sum_i (Y[i + h_n] - Y[i])^2, \tag{3.2}$$

N_n being the number of summed differences. We then propose to fit the theoretical semi-variogram w to the empirical one $\widehat{w}$ by minimizing the least-square criterion

$$\mathcal{L}(\tau, \beta) = \frac{1}{2} \sum_{n=1}^{N} (w(h_n; \tau, \beta) - \widehat{w}_n)^2. \tag{3.3}$$

To minimize the least-square criterion, the topothesy and Hurst functions are expanded into two finite-dimensional subspaces $\mathcal{T}$ and $\mathcal{B}$ of the space of square integrable π-periodic functions. Given some bases $(T_j)_{j=1}^J$ and $(B_k)_{k=1}^K$ of $\mathcal{T}$ and $\mathcal{B}$, the topothesy

and Hurst functions are represented as

$$\begin{aligned} \tau(\theta) &= \sum_{j=1}^{J} \tau_j T_j(\theta), \\ \beta(\theta) &= \sum_{k=1}^{K} \beta_k B_k(\theta). \end{aligned} \tag{3.4}$$

Using these representations, the semi-variogram of the image may be written as

$$w(h_n;\tau,\beta) = \tau_0 + \frac{1}{2}\sum_{j=1}^{J}\tau_j \sum_{m=1}^{M} T_j(\theta_m)\lambda_m \left|\left\langle \frac{h_n}{I}, u(\theta_m)\right\rangle\right|^{2\beta(\theta_m)} = F(\beta)_n\tau, \tag{3.5}$$

where τ is a notation abuse for the column-vector $(\tau_j)_{j=0}^{J}$ and F_n is a row-vector-valued function whose components are

$$F(\beta)_{nj} = \frac{1}{2}\sum_{m=1}^{M} T_j(\theta_m)\lambda_m \left|\left\langle \frac{h_n}{I}, u(\theta_m)\right\rangle\right|^{2\beta(\theta_m)} = v(h_n;T_j,\beta), \tag{3.6}$$

for $j = 1, \cdots, J$ and $F(\beta)_{n0} = 1$. We will denote $F(\beta)$ the matrix of terms $F(\beta)_{nj}$.

Consequently, the least-square criterion can be reformulated as

$$\mathcal{L}(\tau,\beta) = \frac{1}{2}\sum_{n=1}^{N}(\epsilon_n(\tau,\beta))^2. \tag{3.7}$$

where $\epsilon_n(\tau,\beta)$ are residuals defined by

$$\epsilon_n(\tau,\beta) = F_n(\beta)\tau - \widehat{w}_n. \tag{3.8}$$

In the optimization literature, such a criterion is known as a separable nonlinear least square (SNLLS) criterion [16].

4. Problem resolution

There is a wide literature devoted to the minimization of SNLLS criteria such as the one defined by Equations (3.7) and (3.8); see, for instance, [16, 24, 17, 39, 29]. The construction of minimization algorithms is based on the variable projection method introduced in [16]. Let

$$\mathcal{K}(\beta) = \mathcal{L}(\tau^*(\beta),\beta),$$

where $\tau^*(\beta)$ minimizes $\mathcal{L}$ for a fixed β,

$$\tau^*(\beta) \in \arg\min_{\tau} \mathcal{L}(\tau,\beta).$$

The variable projection method consists in reducing the minimization of $\mathcal{L}$ with respect to the couple (τ,β) to the minimization of $\mathcal{K}$ with respect to the single variable β.

The function $\mathcal{K}$ is usually minimized using some variants of the Newton algorithm involving approximations of the Hessian of $\mathcal{K}$ (*e.g.* the Gauss-Newton algorithm or the Levenberg-Marquardt algorithm [26, 27]). The approximation of the Hessian are specifically derived according to the form of the Hessian of a SNLLS [16, 24, 17, 39, 29]. Here, the Hessian of $\mathcal{K}$ can be written as follows (see Sections A and B of the appendix

for the computation details). Denote

$$
\begin{aligned}
A_{11}(\beta) &= \sum_{n=1}^{N} F_n(\beta)^T F_n(\beta), \\
A_{22}(\beta) &= (\tau^*(\beta))^T \left(\sum_{n=1}^{N} DF_n(\beta) DF_n(\beta)^T \right) \tau^*(\beta), \\
E_{22}(\beta) &= \sum_{n=1}^{N} \epsilon_n(\tau^*(\beta), \beta) \sum_{j=1} \nabla^2 F_{nj}(\beta) \tau^*(\beta)_j, \\
A_{12}(\beta) &= \sum_{n=1}^{N} F_n(\beta)^T \, DF_n(\beta)^T \tau^*(\beta), \\
E_{12}(\beta) &= \sum_{n=1}^{N} \epsilon_n(\tau^*(\beta), \beta) \, DF_n(\beta).
\end{aligned}
\tag{4.1}
$$

where ϵ_n are residuals given by Equation (3.8), $F_n(\beta)$ is the function defined by Equation (3.6), $DF_n(\beta)$ its Jacobian and $\nabla^2 F_{nj}(\beta)$ the Hessian of its j component F_{nj}. Then, the Hessian of $\mathcal{K}$ can be expressed as

$$
\nabla^2 \mathcal{K}(\beta) = A_{22}(\beta) + E_{22}(\beta) - (A_{12}(\beta) + E_{12}(\beta))^T (A_{11}(\beta))^{-1} (A_{12}(\beta) + E_{12}(\beta)),
\tag{4.2}
$$

Approximations of the Hessian usually exclude terms that depends on the second derivatives of the function (here, the term E_{22}). In our algorithm, we use a simple approximation of $\mathcal{K}$ by the matrix $A_{22}(\beta) + \lambda I$; for $\lambda = 0$, this corresponds to the Algorithm III pointed out in [39]. The minimization algorithm was implemented in Python using the method *least_squares* of the optimization toolbox of the Scipy library [43]. At iteration $t+1$, the algorithm updates the current value $\beta^{(t)}$ of β by computing

$$
\beta^{(t+1)} = \beta^{(t)} - \left(A_{22}(\beta^{(t)}) + \lambda I \right)^{-1} \nabla \mathcal{K}(\beta^{(t)}),
$$

where λ is a non-negative damping factor adjusted at each iteration, and the gradient of $\mathcal{K}$ is given by

$$
\nabla \mathcal{K}(\beta) = \sum_{n=1}^{N} DF_n(\beta)^T \tau^*(\beta) \, \epsilon_n(\tau, \beta).
\tag{4.3}
$$

Besides, the solution $\tau^*(\beta)$ minimizes a linear least square problem, and is characterized as the solution of the linear system:

$$
F(\beta)^T F(\beta) \tau = F(\beta)^T \widehat{w}.
$$

In our implementation, it is found using the method *lsq_linear* of the Scipy library, which is robust to ill-conditioned systems.

To design our algorithm, we set up a coarse-to-fine multigrid strategy: the parameters (τ, β) are successively approximated in a series of embedded subspaces $(\mathcal{T}_s \times \mathcal{B}_s)_{s=0}^{S}$. In the current implementation of the algorithm, the subspaces $\mathcal{T}_s$ and $\mathcal{B}_s$ are both defined as the space of piecewise constant functions on the intervals

$$
\left\{ \left[-\frac{\pi}{2} + \frac{m}{2^s}\pi, -\frac{\pi}{2} + \frac{m+1}{2^s}\pi \right), m = 0, \cdots, 2^s - 1 \right\}.
$$

Moreover, some constraints are added to these spaces to ensure that the topothesy function is non-negative and the Hurst function ranges in $(0, 1)$. This is directely implemented in the *least_squares* and *lsq_linear* methods that use a Trust-Region-Reflective algorithm to deal with these constraints [11].

Set the algorithm parameters S, T and η (by default, $S = 3$, $T = 10000$ and $\eta = 1e{-}3$).
Initialization:
Set $L_0 = +\infty$
for $h \in \{0.1, \cdots, 0.9\}$ **do**
  Set $\beta \in \mathcal{B}_0$ equal to h.
  Find

$$\tau^*(\beta) = \arg\min_{\tau \in \mathcal{T}_0} \mathcal{L}(\tau, \beta)$$

  using *lsq_linear (method TRF)* .
  if $\mathcal{L}(\tau^*(\beta), \beta) < L_0$ **then**
    Set $L_0 = \mathcal{L}(\tau^*(\beta), \beta)$,
    Set $\tau_0 = \tau^*(\beta)$ and $\beta_0 = \beta$.
  end if
end for
Main iterations:
Set $\tau = \tau_0$ and $\beta = \beta_0$.
Set iter = 0.
for $s = 0$ **to** S **do**
  Project the current functions τ, β into the space $\mathcal{T}_s \times \mathcal{B}_s$.
  Set iter = 0.
  repeat
    iter $\leftarrow$ iter + 1.
    Set $\tau_0 = \tau$ and $\beta_0 = \beta$
    Fix β and find

$$\tau^*(\beta) = \arg\min_{\tau \in \mathcal{T}_s} \mathcal{L}(\tau, \beta).$$

    using *lsq_linear (method TRF)*.
    Fix $\tau = \tau^*(\beta)$ and approach

$$\arg\min_{\beta \in \mathcal{B}_s} \mathcal{K}(\beta).$$

    using a few iterations of the algorithm *least_squares (method TRF)*.
  until $\frac{\mathcal{L}(\tau_0, \beta_0) - \mathcal{L}(\tau, \beta)}{\mathcal{L}(\tau_0, \beta_0)} < \eta$ **or** iter $> T$.
end for

FIGURE 2. Algorithm for the estimation of the topothesy and Hurst functions.

The overall algorithm is summarized in Figure 2.

In practice, we observed that the choice of the lags where the semi-variograms are evaluated had an influence on the minimization. In particular, we believe that having two orthogonal lags with a same module can make the problem ill-posed with permutable solutions. So we arranged the lags to avoid such pathological configuration. Moreover, we spread the lags uniformly in all directions to avoid privileging a few directions. A typical set of lags is shown on Figure 3.

Besides, we set angles θ_m of the turning-band field representation using the dynamic programming algorithm described in [6] with $M = 657$ and set $\lambda_m = \theta_m - \theta_{m-1}$.

5. Numerical study

In order to assess the performance of the proposed method, we conducted experiments on field realizations generated by the PyAFBF package [6, 37]. The topothesy and Hurst functions of the simulated fields were represented as piecewise constant functions on

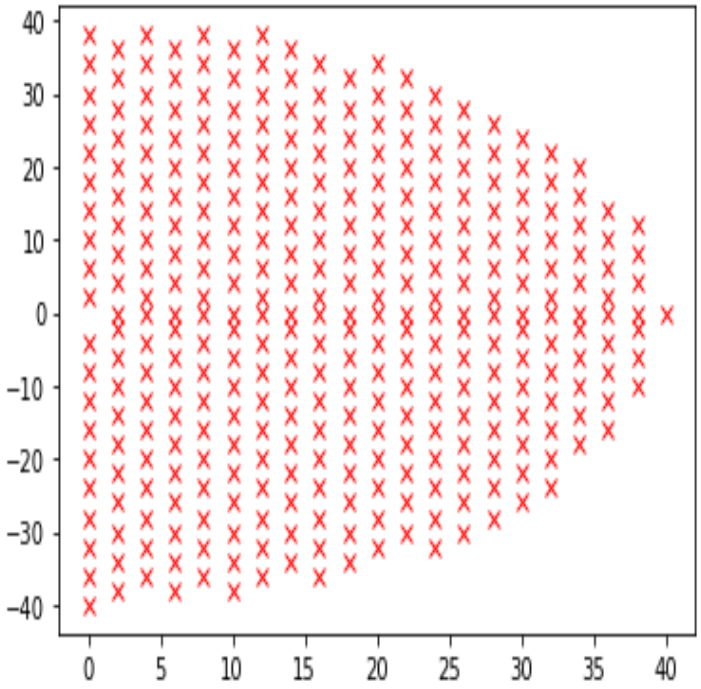


FIGURE 3. Lags where the semi-variograms are computed.

uniformly spaced intervals of $]-\frac{\pi}{2},\frac{\pi}{2}]$. In a first series of experiments, we made the number of constants vary so as to study the effect of the model complexity on the estimation precision. We built seven sets of hundred realizations. Realizations were images of size 1024×1024. On each set, they were obtained from different fields having a same number of parameters. Depending on the set, the number of parameters was $M = 2^m$ for $m = 0, \cdots, 6$ for the topothesy and Hurst functions. These parameters correspond to the constants of the topothesy and Hurst functions. The higher the number of these constants, the more precise is the directional description of the field. To account for this precision, we defined a so-called radial precision, expressed in percent, as

$$r = (1 - 1/M) * 100,$$

with 0 corresponding to the worst precision and 100 to the best one. The estimation error was evaluated on the Hurst function using a L_1 error, expressed in percent, as

$$e = \frac{100}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} |\beta(s) - \hat{\beta}(s)| ds,$$

β and $\hat{\beta}$ being the true and estimated Hurst functions, respectively. We averaged the estimation errors over the hundred realizations of each set. To obtain the estimates, we applied the method with the coarse-to-fine multigrid strategy (see Section 4), making the size parameter s vary from 0 to m. The maximal number of iterations and the tolerance were set to $T = 10000$ and $\eta = 1e-3$, respectively. All experiments can be reproduced or modified using the PyAFBF package. The results are shown on Figure 4. The estimation error was quite low (below 10 percent), even when the radial precision was high. The method outperformed the one in [7, 38] which could only achieved the

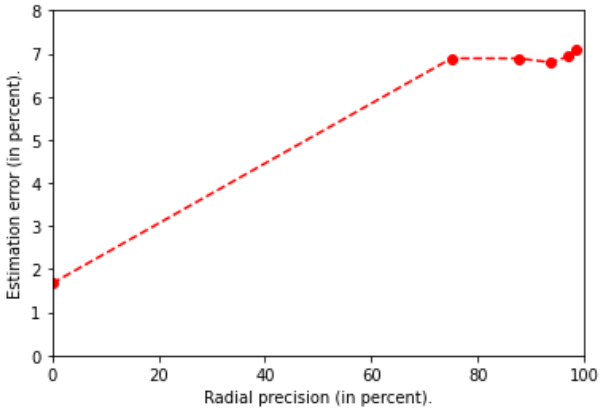


FIGURE 4. Estimation error on the Hurst function as a function of the radial precision.

estimation with a low radial precision of 50 % (4 parameters) and with an error above 10 %. The coarse-to-fine multigrid strategy was critical for the estimation procedure. Without this strategy, the estimation error was above 10 % for a radial precision of 87.5 %.

Besides, the estimation error depended on the error made when estimating the field semi-variogram with the empirical semi-variogram (see Equation (3.2)): after having replaced the empirical semi-variogram by the field one in the fitting criterion, the estimation error dropped down below 2 % for a radial precision of 87.5%. The estimation quality of the empirical semi-variogram depends on the size of the field realizations. Hence, to see the effect of this size on the estimation of model parameters, we conducted some other experiments on several sets of hundred realizations of a same size $M \times M$ with a radial precision of 87.5%. On Figure (5), estimation errors are reported as a function of the parameter size M.

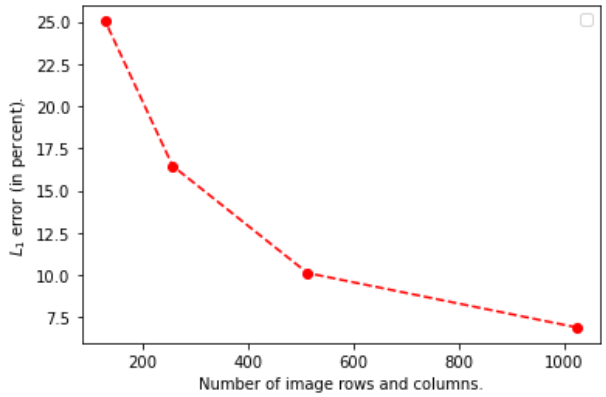


FIGURE 5. Estimation error on the Hurst function with respect to the realization size.

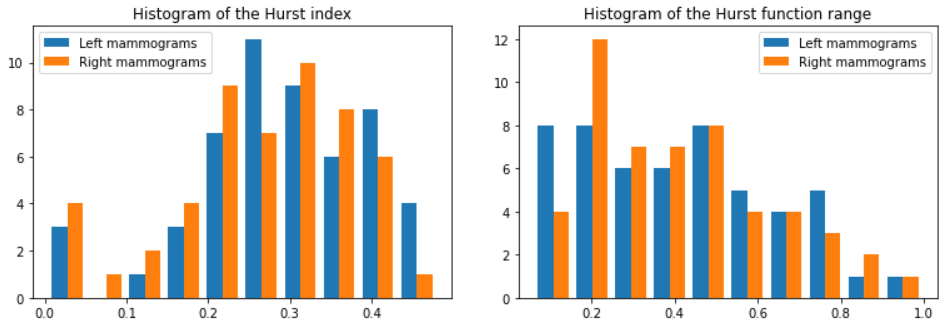
We notice that the error significantly increased as the size of the image decreased. For a radial precision of 87.5%, a realization larger than 512 is required to ensure that the estimation error remains below 10 %.

6. Illustration

To illustrate the interest of our estimation procedure, we applied it to mammograms. Mammograms are projective images of the breast obtained by capturing part of an X-ray beam that passes through the adipose and fibro-glandular tissues organized in hierarchical matrix of fibrous compartments. Radiographic projections of this matrix create the mammographic texture, so called, parenchymal pattern. Mammographic texture has a critical effect on cancer detection, as it can hide early tumors (causing missed cancers) or mimic tumors (causing false positive findings). The statistical nature of the parenchymal pattern was acknowledged by several authors. In [12, 20, 23], a power-law noise model was proposed to characterize mammogram textures and study either the lesion detectability or assess the breast cancer risk. This model accounts for the self-similarity of mammographic textures, but it is global and stationary. In [25, 1], another model was design using a wavelet-based multifractal formalism to locally characterize mammogram textures. In [38], the AFBF were applied to mammograms to account for and analyze their anisotropy.

In this part, we propose to revisit the study of [38] by taking the same dataset. This dataset was provided by the Department of Radiology of the University of Pennsylvania. Images of this dataset were acquired in medio-lateral oblique position using a Senographe 2000D (General Electric Medical Systems, Milwaukee, WI), with a spatial resolution of 0.1 mm $\times$ 0.1 mm per pixel (image size: 1914 $\times$ 2294 pixels). In each image of the database, a region of interest of size 512 $\times$ 512 was manually extracted within the densiest region of the breast. Originally, the dataset included a total of 58 cases, each case being composed of the left and right breasts of a woman. For this study, we removed 12 images that contained a part of the background (area outside the breast), causing an estimation failure.

We applied to these images our procedure to estimate the Hurst and topothesy functions represented as piecewise constant functions with 16 parameters each. We then studied several field features extracted from these functional estimates.



(a) (b)

Figure 6. Histograms of (a) the Hurst index and (a) Hurst function range computed on the mammogram dataset.

The first feature was the minimal value H of the Hurst function (called the Hurst index), that measures a degree of texture regularity. The repartitions of this index over the mammograms of left and right breasts are shown on histograms of Figure 6 (a). These histograms are similar to the ones published in [38]: the Hurst index is also distributed in a same way on left and right breasts. The mean value is around 0.3; it is of 0.28 with the standard deviation of 0.11. In the previous distributions of [38], there were much

fewer index values in the interval $(0.3, 0.5)$ than in the present distributions. Contrarily to the approach proposed in this paper, the model used in [38] did not take into account the image noise. This probably led to an underestimation of the Hurst index.

The second feature was the length of the Hurst function range measured by the difference between its maximal and minimal values. In [38], such a feature could not be precisely computed due to the fact that the estimation of the Hurst function was limited to a few directions. It could only be underestimated by taking the difference between the maximum of the hurst values in vertical and horizontal directions and the Hurst index (histogram (c) of Figure 4). The distribution of this previously estimated index was concentrated around the value of 0.2 with very few values exceeding 0.3. This is in sharp constrast with the distribution of the new length estimates shown on Figure 6 (b). We can see that the length is more uniformly distributed over the interval $(0, 1)$, and often above 0.3. This length is an indicator of the degree of texture anisotropy: the larger this length, the more anisotropic the texture. Hence, the result obtained in this study shows that the degree of anisotropy of mammograms is even larger than expected in [38].

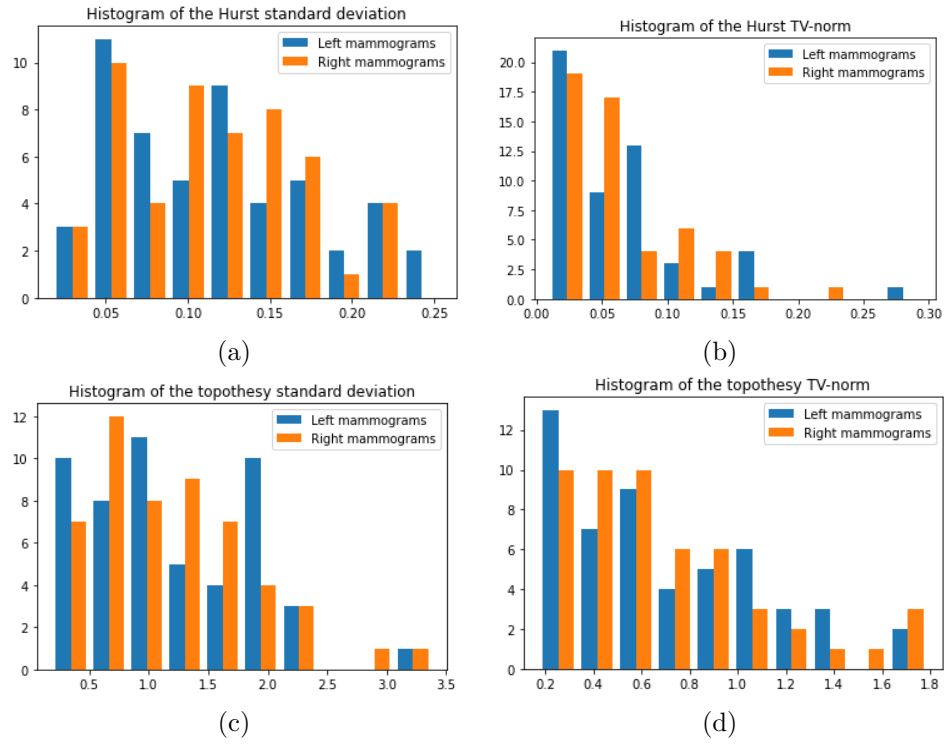


FIGURE 7. Histograms of (a) the standard deviation and (a) TV-norm of the Hurst function computed on the mammogram dataset. Histograms of (c) the standard deviation and (d) TV-norm of the topothesy function.

We further propose to describe the mammogram anisotropy through some original features that could not have been computed nor approached in the previous work [38]. These features are the standard deviations of the Hurst and topothesy functions, and their TV-norms. The standard deviation accounts for the importance of function variations and the TV-norm for the sharpness of these variations. For the topothesy function, these features were normalized by the mean function value to make these features independent of a linear contrast change of the image. According to histograms on Figure 7 (a) and

(b), variations of the Hurst function can be large (standard variation often reaches values around of 0.15) but are rather smooth than sharp (the TV-norm remains quite low). As observed on Figure 7 (c) and (d), variations of the topothesy function can be large and seems to be sharper than those of the Hurst functions.

On Figure 8, some images of the datasets are shown together with their corresponding topothesy and Hurst functions. We can observe a typical form of the Hurst function which is peaked in one direction. The peak location indicates a direction where image variations have lowest frequencies. The peaks are also present in the topothesy functions at a same place as the ones of the Hurst function. The peaks of the Hurst and topothesy functions seem to vary in sharpness and height depending on the image anisotropy.

On Figure 9, we further show simulations of several original images sampled from the estimated AFBF. These illustrations account for the ability of the simulated model to reproduce the mammogram aspect, including its regularity, directionality, anisotropy, trends and details. We observe that some fine structures present for instance in the fifth example could not be reproduced truthfully.

7. Comparison with state-of-the-art methods

In this section, we investigate the ability of state-of-the art texture models based on stationary Gaussian field to estimate and characterize mammograms.

7.1. **Methods.** As argued in [14], there exist mainly two approaches to simulate stationary Gaussian fields: Random Phase Noise (RPN) and Asymptotic Discrete Spot Noise (ADSN). Even though these two approaches generate different types of stochastic process, they provide similar results for most textures.

The presentations of each method is inspired from [14]. Both methods assume that textures are periodic images and require an input image $u \in \mathbb{R}^{I^2}$ for the estimation of the field. They rely upon a so-called texton ψ of size $L \times L$ (with $L \leq I$) which a subimage selected from the original image and represents the underlying stationary field.

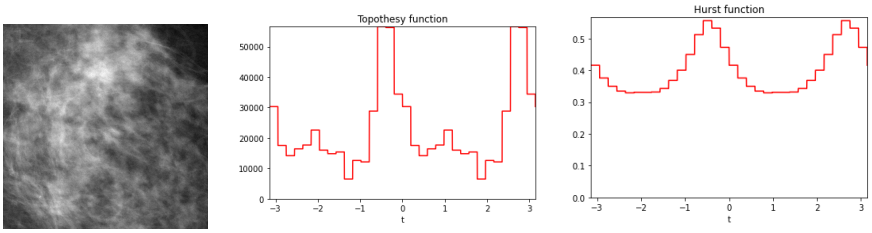
To do our experiments, we used the code provided by authors of [15].

Random Phase Noise (RPN): The generation of new textures from an input texton $\psi \in \mathbb{R}^{L^2}$ through the RPN approach consists in simulating the field $Z \in \mathbb{R}^{I^2}$ with Fourier transform defined as $\widehat{Z} = |\widehat{\psi}|e^{i\theta}$ where θ is a random odd vector in $(-\pi, \pi]^{I^2}$ and $\widehat{\psi} \in \mathbb{C}^{I^2}$ is the Fourier transform of ψ zero-padded. In other words, new textures are simulated by keeping the modulus of ψ but using random phases.

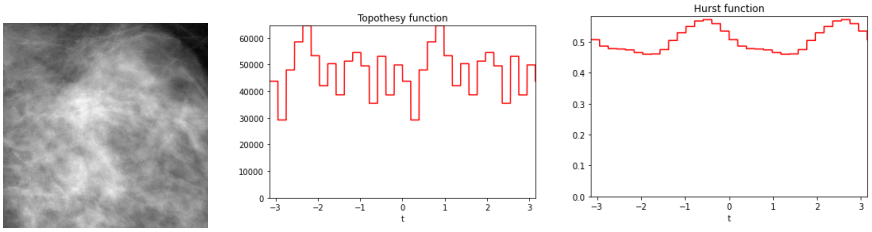
Asymptotic Discrete Spot Noise (ADSN): This simulation method consists in simulating the field $Z = \psi \star Y$ where $Y \sim \mathcal{N}(0, \mathrm{Id}_{I^2})$ and $\psi \in \mathbb{R}^{L^2}$ is a user-determined texton with mean 0 and properly normalized.

It is worth mentioning that the autocovariance of the processes Z are fully determined by ψ and is given by $\widehat{c} = |\widehat{\psi}|^2$. The estimation of the texton ψ from the input image u is a crucial step and is performed using the method of [15]. In practice, the input image u may not be periodic, which leads to estimation artifacts due to the implicit periodic extension of the domain. We will therefore run each method on the periodic component of u obtained via the periodic+smooth decomposition of [28].

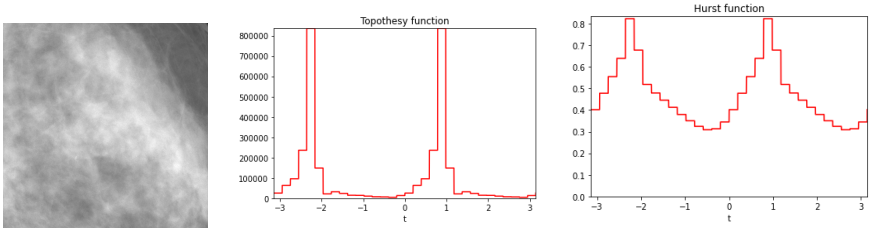
7.2. **Estimation of mammograms.** The textons associated to each mammogram are estimated with sizes $L = 63$ and $L = 255$. The ADSN processes associated to each texton are then simulated and are displayed on Figure 10. The results with the RPN are not displayed here since very similar to the one obtained with the ADSN (see next subsection).



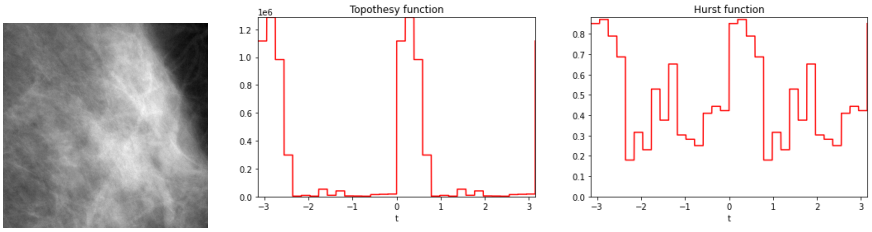
$H = 0.33$, $H_l = 0.23$, $\beta_{std} = 0.08$, $\beta_{tv} = 0.03$, $\tau_{std} = 0.63$, $\tau_{tv} = 0.32$.



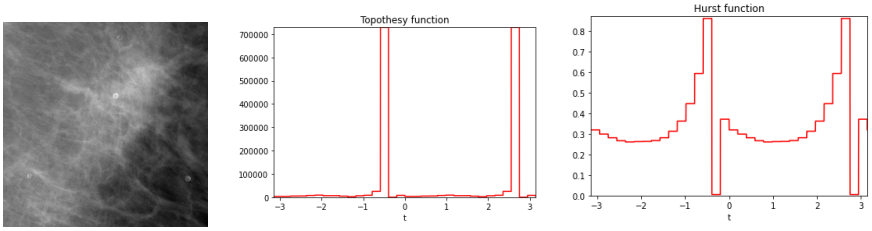
$H = 0.46$, $H_l = 0.11$, $\beta_{std} = 0.04$, $\beta_{tv} = 0.01$, $\tau_{std} = 0.19$, $\tau_{tv} = 0.23$.



$H = 0.31$, $H_l = 0.51$, $\beta_{std} = 0.14$, $\beta_{tv} = 0.06$, $\tau_{std} = 2.07$, $\tau_{tv} = 1.09$.



$H = 0.18$, $H_l = 0.69$, $\beta_{std} = 0.22$, $\beta_{tv} = 0.02$, $\tau_{std} = 1.8$, $\tau_{tv} = 0.26$.



$H = 0.005$, $H_l = 0.86$, $\beta_{std} = 0.18$, $\beta_{tv} = 0.12$, $\tau_{std} = 3.38$, $\tau_{tv} = 1.79$.

FIGURE 8. Estimation of the topothesy (τ) and Hurst (β) functions from some mammograms. First column: image ROI, second and third columns, the graphs of τ and β estimated from the ROI of the same row. Below the graphs, the values of the different field features are also reported: H is Hurst index, H_l the range of τ ; β_{std} and β_{tv} are the standard deviation and the TV-norm of β respectively. Similarly, τ_{std} and τ_{tv} are the standard deviation and the TV-norm of τ respectively.

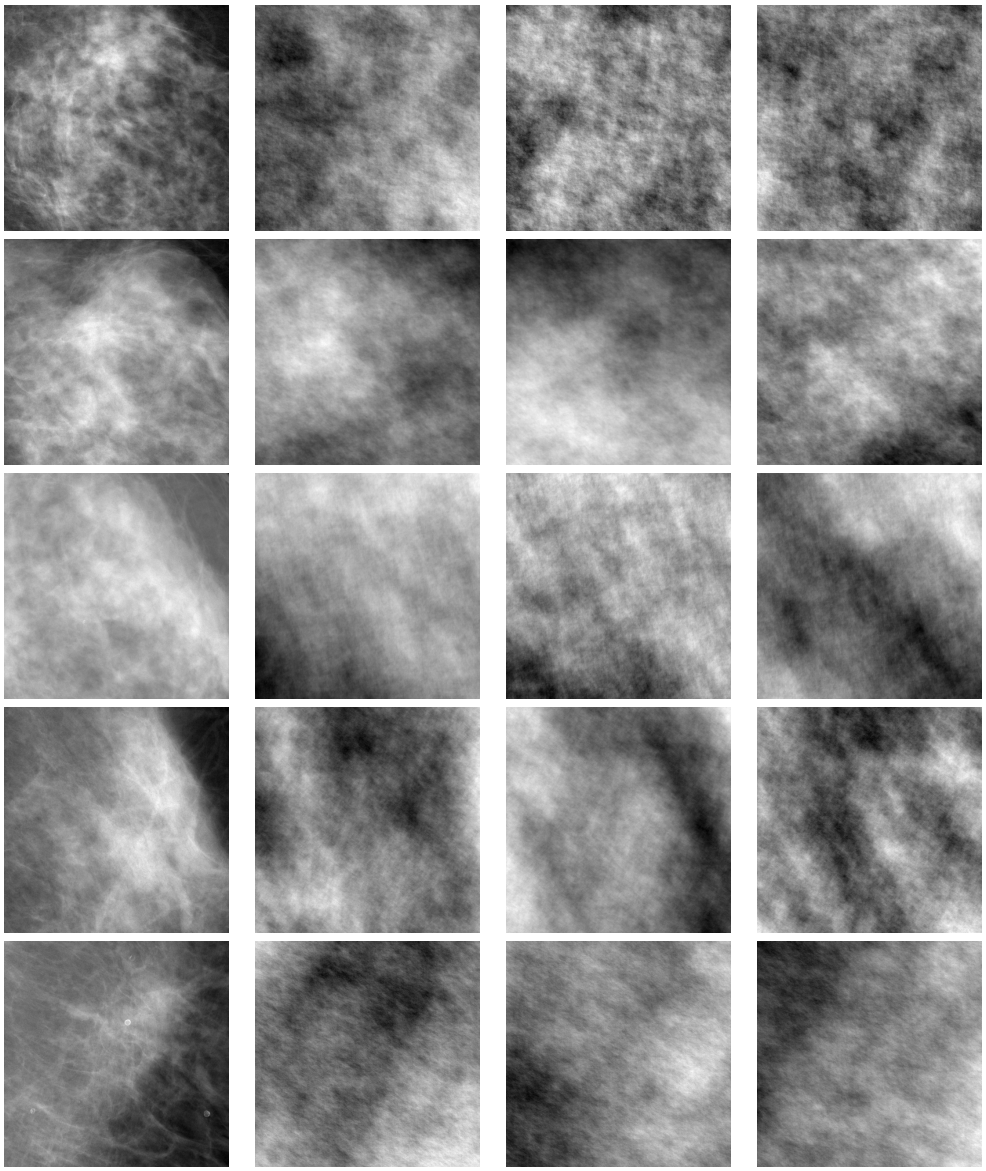


FIGURE 9. Simulations of mammograms: First column: the original image, second to fourth column: three synthetic images sampled from the AFBF estimated from the original image of the same row.

This illustration shows that state-of-the-art methods are not able to characterize real textures from mammograms. In order to understand this failure, we may wonder if these state-of-the-art methods are able to correctly approximate AFBFs. This question is investigated in the next paragraph.

7.3. **Estimation of AFBFs.** In order to illustrate the performance of the RPN and ADSN on AFBFs, we generate three different AFBF realizations u of size 512×512 displayed on Figure 11:

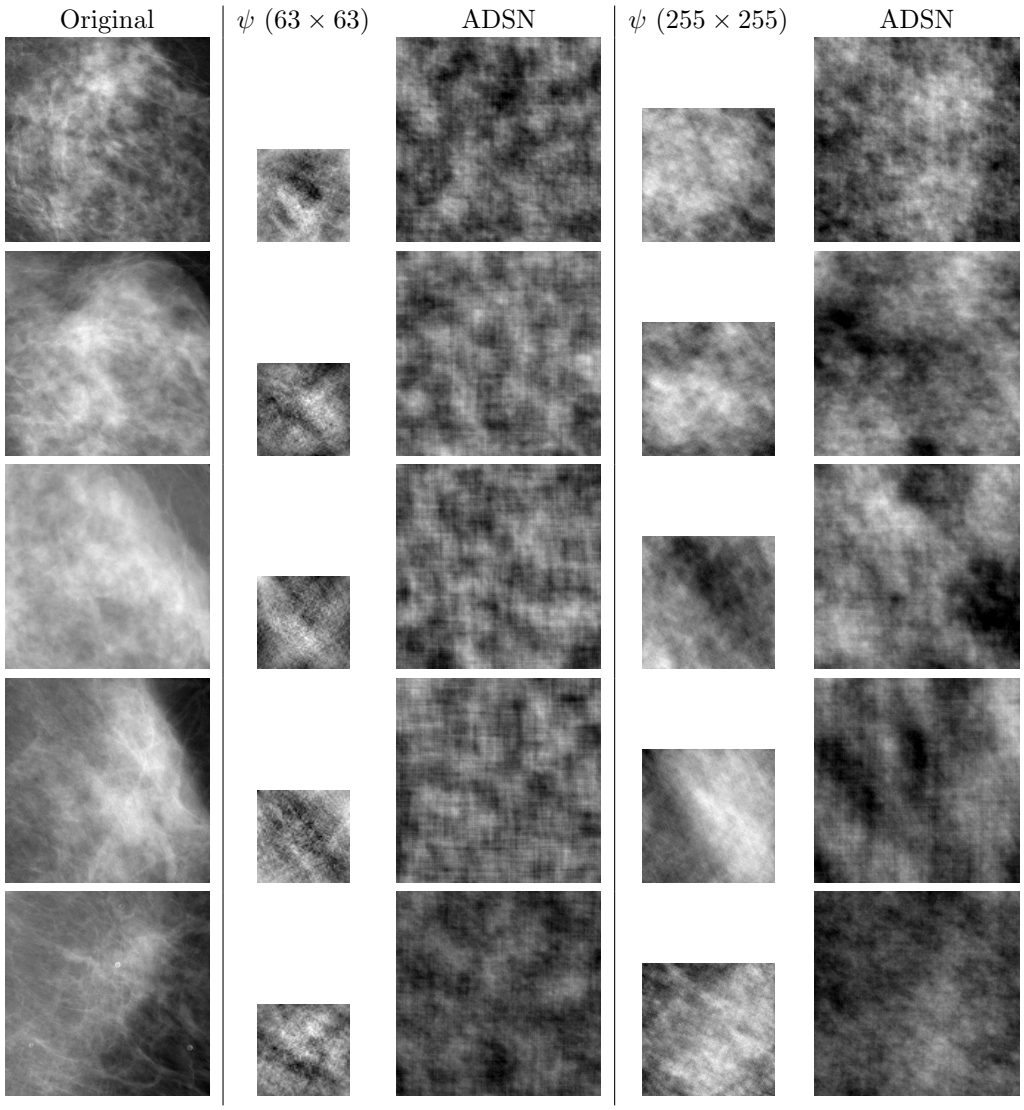


FIGURE 10. Realization of the ADSN for each mammogram of Section 6. First column: the periodic component of the mammogram on a 512×512 grid. Second and third column: the estimated texton ψ of size 63×63 (resp. 255×255) and one realization of the associated process.

Iso1: An isotropic AFBF with its Hurst parameter $\beta < 1/2$ i.e. without long-range dependency.

Iso2: An isotropic AFBF with its Hurst parameter $\beta > 1/2$ i.e. with long-range dependencies.

Aniso: An anisotropic AFBF with its Hurst function $\beta(\theta_1) > 1/2$ in direction θ_1 and $\beta(\theta_2) < 1/2$ in the other direction.

The texton are estimated with sizes $L = \{63, 255, 512\}$ from which the RPN and ADSN processes are simulated and displayed on Figures 12, 13 and 14 respectively.

When the AFBF contains long-range dependencies (Iso2 and Aniso), its estimation can only be performed with a texton of size $L = I = 512$. However, when the AFBF

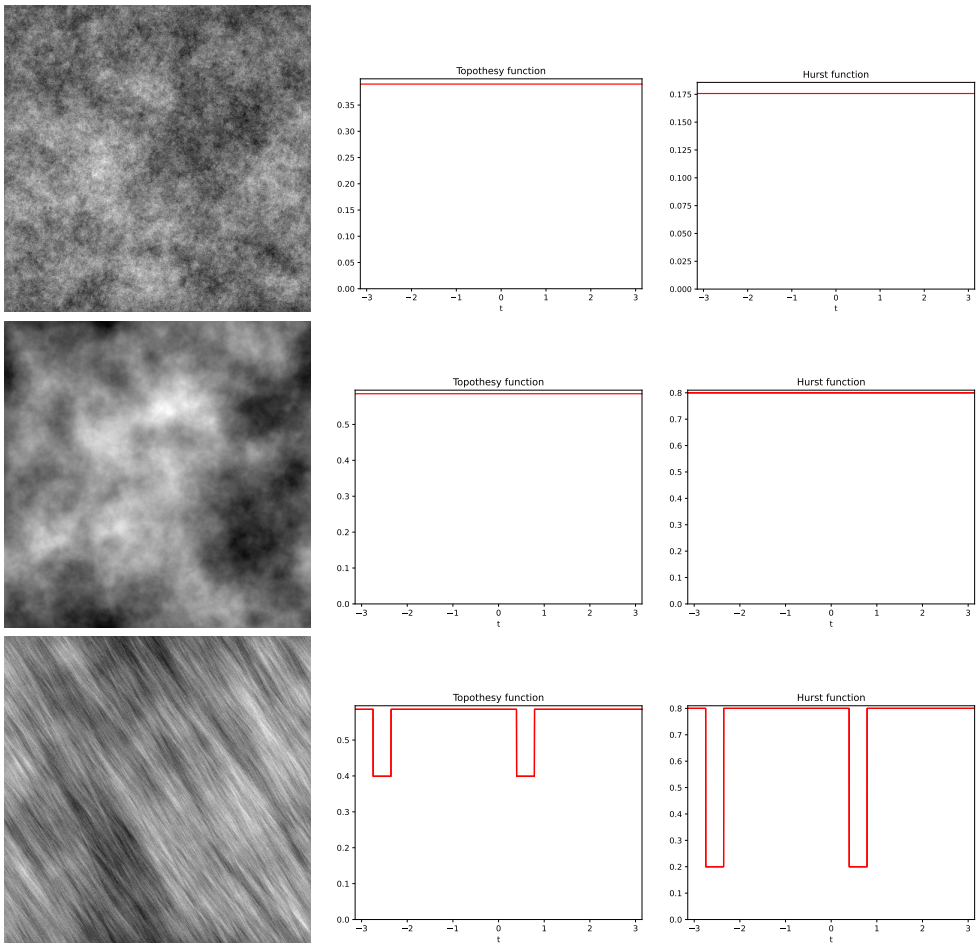


FIGURE 11. The three AFBFs used: Iso1, Iso2, and Aniso form top to bottom. First column: the periodic component of one realization of the AFBF u on a 512×512 grid. Second and third columns: the Hurst and topothesy functions, respectively.

does not contain long-range dependencies (Iso1), a relatively small L is enough to obtain a visually good estimation ($L = 63$).

8. Discussion

We designed a method for the estimation of the functional parameters of the anisotropic fractional Brownian field, namely the topothesy and Hurst functions. This method consisted of fitting the empirical semi-variogram computed from an image to the semi-variogram of a turning-band field that approximates the AFBF. The fitting criterion was formulated in terms of a separable non-linear least square criterion, the linear part relating to the topothesy function and the non-linear part to the Hurst function. We then proposed an algorithm relying on a variable projection to minimize this criterion. This algorithm was combined with a coarse-to-fine multigrid strategy to improve its convergence. We presented a numerical study of the algorithm on simulated textures. In this

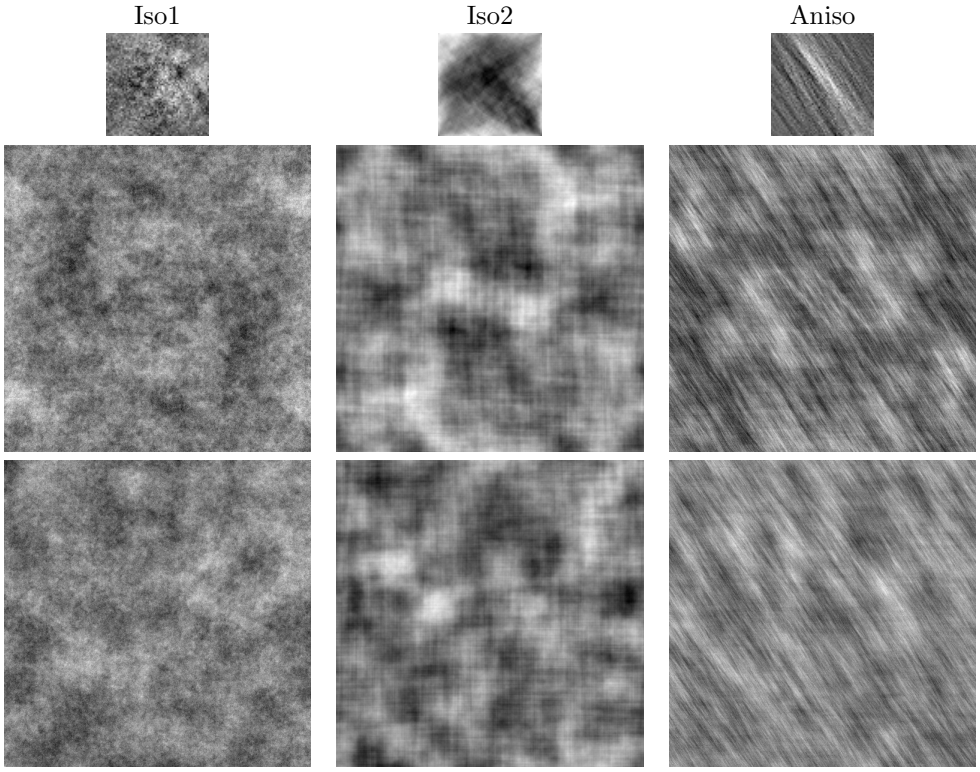


FIGURE 12. Realizations of the RPN and ADSN processes for each of the AFBFs of Figure 11 and their estimated texton ψ of size 63×63. From top to bottom: the texton ψ, the RPN realization and the ADSN one.

study, the estimation error was evaluated as a function of a radial precision, directly related to the number of parameters appearing in the discrete representations of the Hurst and topothesy functions. This error was below ten percent for the finest radial precision, when the empirical semi-variogram was computed on large images. This error increased as the image size descreased. It reached values above 10 % when the image size was below 400×400. For illustration, the estimation method was applied to mammograms of a private dataset studied in a previous work. From the estimated functional parameters, we derived some field features that described the mammogram texture. We also sampled the estimated fields to generate some realistic simulations of mammograms.

This work represents a breakthrough concerning the estimation of parameters of AFBF, which had only been partly and inaccurately achieved. Features can now be derived from the estimated field and used to classify textures. In applications to mammograms, we plan to develop features that could help assessing a breast cancer risk, or evaluate the BIRADS grade of breast density [21, 22], which is a well-known cancer risk factor. Following the approach in [35, 34], we also envisage to use features to detect lesions in mammograms or breast tomosynthesis. However, the actual method is not accurate enough to envisage its application to small-image classification or texture segmentation. Indeed, such an application require estimating model parameters in a local way on small neighborhoods of image pixels. One way to overcome this limitation would

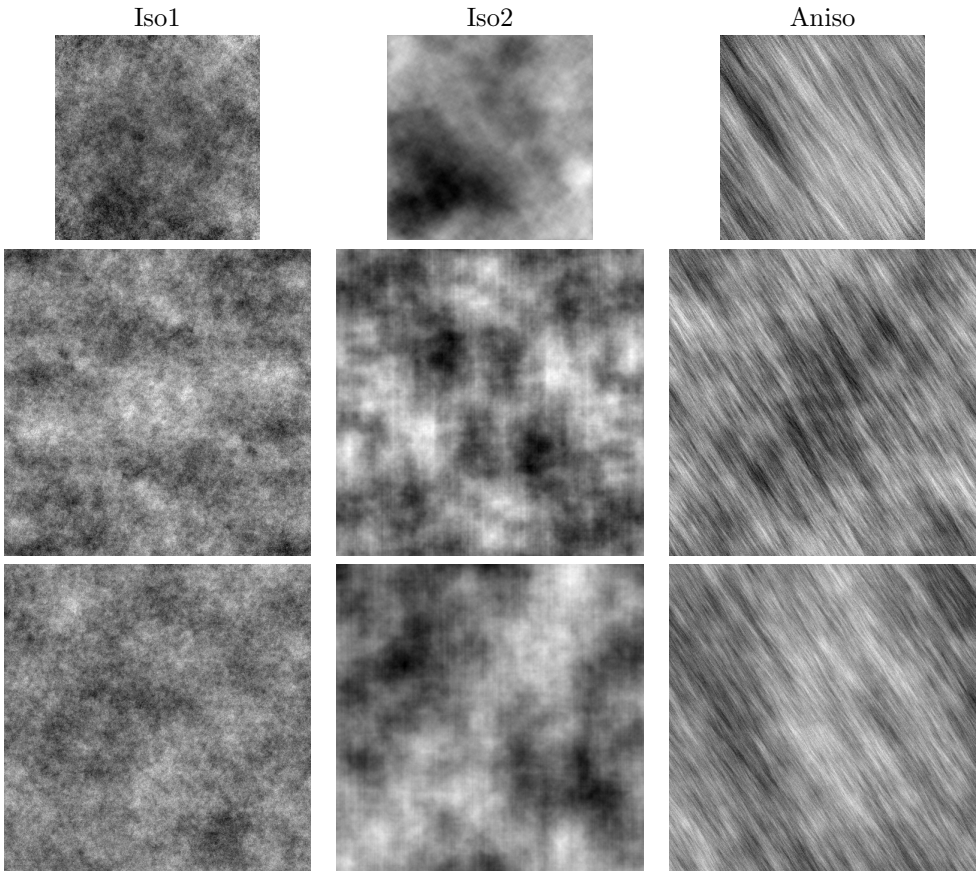


FIGURE 13. Realizations the RPN and ADSN processes for each of the AFBFs of Figure 11 and their estimated texton ψ of size 256×256. From top to bottom: the texton ψ, the RPN realization and the ADSN one.

be to introduce penalizations of parameters within the fitting criterion. Such an approach was adopted for the estimation of the local Hurst index in [31, 30]. Penalty terms would constrain the model parameters in shape or space. For instance, they could be a TV-norm on topothesy and Hurst functions that would regularize the shape of these functions. As the penalty terms could be non-differentiable, our algorithm would no longer be appropriate for the minimization of the fitting criterion. It could however be adapted following the approach developed in [42] for separable non-linear least square criterion or, alternately, specifying generic algorithms for the minimization of non-smooth functions [45].

Besides, the facility brought by this work to sample realistic textures from AFBF is also of interest for applications. For instance, synthetic textures could be used for data augmentation in machine learning [40]: they could supply examples for learning the neural network architectures in domains where there is a lack of data to achieve this task. This could be the case in the medical domain of breast images where, to our knowledge, there are a few public database [41, 19, 18] and, within these database, even fewer cases representing rare lesion types. As shown by illustrations, AFBF seems to be appropriate to randomly reproduce mammograms. However, it fails to reproduce some

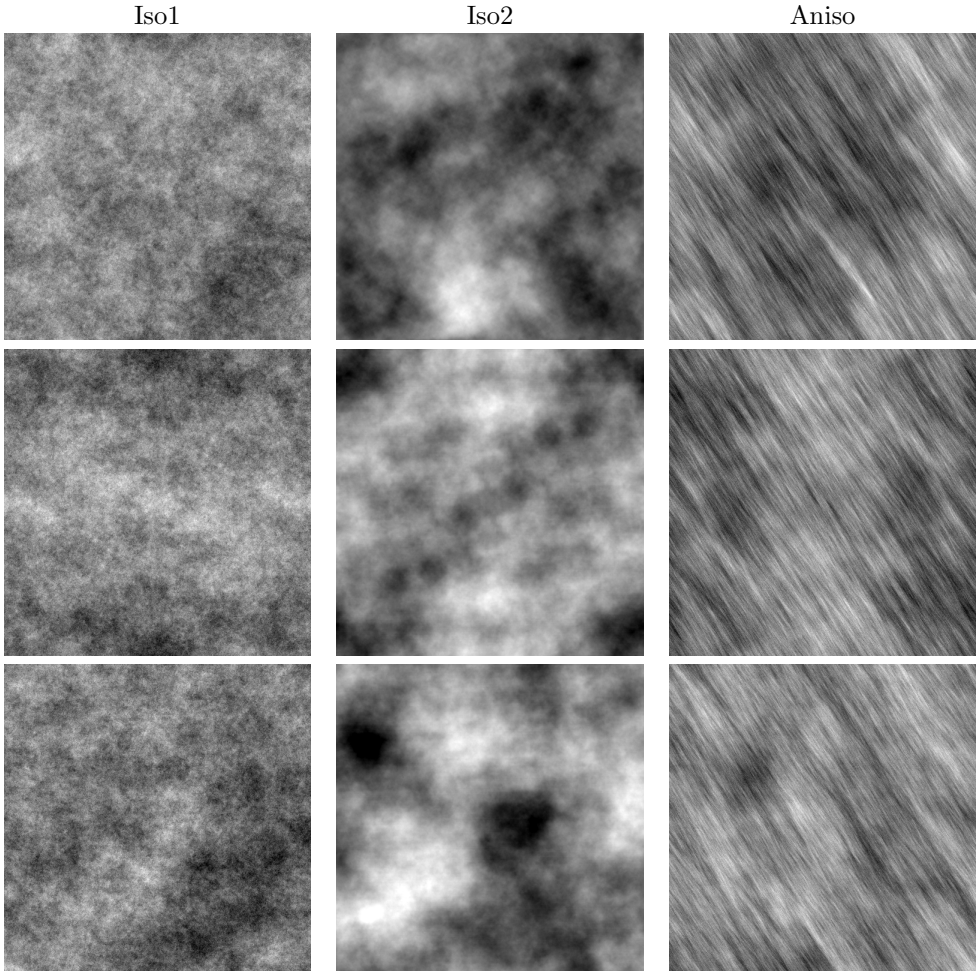


FIGURE 14. Realizations of the RPN and ADSN processes for each of the AFBFs of Figure 11 and their estimated texton ψ of size 512×512. From top to bottom: the texton ψ, the RPN realization and the ADSN one.

fine structured details of the image. To remove this limitation, the AFBF simulation could be combined with simulation methods based on a physical modeling of breast tissues and image acquisition [2, 3, 4, 13].

Synthetic mammograms generated from the AFBF model had a better aspect than those simulated with the state-of-the-art methods ADSN and RPN. Contrarily to these texton-based methods, the AFBF simulation could render long-range field dependencies which occur in mammograms. Fields with long range dependencies are non-stationary and have a variogram which tends to infinity as the lag module tends to infinity. Texton-based methods are not well suitable to simulate such fields. Their application is limited to stationary fields, whose variograms are bounded. On the other hand, the AFBF model is not appropriate for the modeling of fields with bounded variograms. Hence, the AFBF model should be used in complement to state-of-the-art methods for simulating images which do not fulfill the stationary assumption.

Bibliography

1. A. Arneodo, N. Decoster, P. Kestener, et al., *A wavelet-based multifractal image analysis: from theoretical concepts to experimental applications*, Adv. Imaging Electron Phys. (2003), no. 126, 1–92.
2. P. Bakic, M. Albert, D. Brzakovic, and A. Maidment, *Mammogram synthesis using a 3D simulation. I. breast tissue model and image acquisition simulation*, Med. Phys. **29** (2002), 2131–9.
3. ———, *Mammogram synthesis using a 3D simulation. II. evaluation of synthetic mammogram texture*, Med. Phys. **29** (2002), no. 9, 2140–2151.
4. ———, *Mammogram synthesis using a 3D simulation. III. modeling and evaluation of the breast ductal network*, Med. Phys. **30** (2003), no. 7, 1914–1925.
5. A. Benassi, S. Jaffard, and D. Roux, *Elliptic Gaussian random processes*, Rev. Mathem. Iberoamericana **13** (1997), no. 1, 19–89.
6. H. Biermé, M. Moisan, and F. Richard, *A turning-band method for the simulation of anisotropic fractional Brownian field*, J. Comput. Graph. Statist. **24** (2015), no. 3, 885–904.
7. H. Biermé and F.J.P. Richard, *Estimation of anisotropic Gaussian fields through Radon transform*, ESAIM: Probab. Stat. **12** (2008), no. 1, 30–50.
8. H. Biermé and F.J.P. Richard, *Analysis of texture anisotropy based on some Gaussian fields with spectral density*, Mathematical Image Processing (M. Bergounioux, ed.), Springer Proceedings, 2011, pp. 59–73.
9. H. Biermé, F.J.P. Richard, M. Rachidi, and C.L. Benhamou, *Anisotropic texture modeling and applications to medical image analysis*, Mathematical Methods for Imaging and Inverse Problems (H. Ammari, ed.), ESAIM Proc., vol. 26, 2009, pp. 100–122.
10. A. Bonami and A. Estrade, *Anisotropic analysis of some Gaussian models*, J. Fourier Anal. Appl. **9** (2003), 215–236.
11. M. A. Branch, T. F. Coleman, and Y. Li, *A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems*, J. Sci. Comput. **21** (1999), no. 1, 1–23.
12. A. Burgess, F. Jacobson, and P. Judy, *Human observer detection experiments with mammograms and power-law noise*, Med. Phys. **28** (2001), no. 4, 419–437.
13. A.-K. Carton, P. Bakic, C. Ullberg, H. Derand, and A. Maidment, *Development of a physical 3D anthropomorphic breast phantom*, Med. Phys. **38** (2011), no. 2, 891–896.
14. Bruno Galerne, Yann Gousseau, and Jean-Michel Morel, *Random phase textures: Theory and synthesis*, IEEE Transactions on image processing **20** (2010), no. 1, 257–267.
15. Bruno Galerne, Arthur Leclaire, and Lionel Moisan, *A texton for fast and flexible gaussian texture synthesis*, 2014 22nd European Signal Processing Conference (EUSIPCO), IEEE, 2014, pp. 1686–1690.
16. G. Golub and V. Pereyra, *The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate*, J. Numer. Anal. **10** (1973), no. 2, 413–432.
17. ———, *Separable nonlinear least squares: the variable projection method and its applications*, Inverse Probl. **19** (2003), no. 2, R1.
18. M. Heath, K. Bowyer, D. Kopans, et al., *Current status of the digital database for screening mammography*, Digital Mammo, Springer, 1998, pp. 457–460.
19. M. Heath, K. Bowyer, D Kopans, et al., *The digital database for screening mammography*, Proc. of the Fifth International Workshop on Digital Mammography, 2001, pp. 212–218.
20. J. Heine, S. Deine, R. Velthuizen, et al., *On the statistical nature of mammograms*, Med. Phys. **26** (1999), no. 11, 2254–2269.
21. J. Heine and P/ Malhotra, *Mammographic tissue, breast cancer risk, serial image analysis, and digital mammography: Part 1. tissue and related risk factors*, Acad. Radiol. **9** (2002), no. 3, 298–316.
22. J/ Heine and P. Malhotra, *Mammographic tissue, breast cancer risk, serial image analysis, and digital mammography: Part 2. serial breast tissue chage and related temporal influences*, Acad. Radiol. **9** (2002), no. 3, 317–335.
23. J. Heine and R. Velthuizen, *Spectral analysis of full field digital mammography data*, Med. Phys. **29** (2002), no. 5, 647–661.
24. L. Kaufman, *A variable projection method for solving separable nonlinear least squares problems*, BIT Numer. Math. **15** (1975), no. 1, 49–57.
25. P. Kestener, J.-M. Lina, P. Saint-Jean, et al., *Wavelet-based multifractal formalism to assist in diagnosis in digitized mammograms*, Image Anal. Stereol. **20** (2001), 169–174.

26. K. Levenberg, *A method for the solution of certain non-linear problems in least squares*, Q. Appl. Math. **2** (1944), no. 2, 164–168.
27. D. W. Marquardt, *An algorithm for least-squares estimation of nonlinear parameters*, J. Soc. Ind. Appl. Math. **11** (1963), no. 2, 431–441.
28. L. Moisan, *Periodic plus smooth image decomposition*, J. Math. Imaging Vis. **39** (2011), no. 2, 161–179.
29. D. P O'Leary and B. W Rust, *Variable projection for nonlinear least squares problems*, Comput. Optim. Appl. **54** (2013), no. 3, 579–593.
30. B. Pascal, N. Pustelnik, and P. Abry, *Strongly convex optimization for joint fractal feature estimation and texture segmentation*, Appl. Comput. Harmon. Anal. **54** (2021), 303–322.
31. B. Pascal, S. Vaiter, N. Pustelnik, and P. Abry, *Automated data-driven selection of the hyperparameters for total-variation-based texture segmentation*, J. Math. Imaging Vis. **63** (2021), no. 7, 923–952.
32. K. Polisano, M. Clausel, V. Perrier, and L. Condat, *Texture modeling by Gaussian fields with prescribed local orientation*, Int Conf on Image Processing (ICIP), 2014 IEEE, 2014, pp. 6091–6095.
33. F. Richard, *Anisotropy of Hölder Gaussian random field: characterization, estimation and application to image textures*, Stat. Comput. **28** (2018), no. 6, 1155–1168.
34. F.J.P. Richard, *Analysis of anisotropic Brownian textures and application to lesion detection in mammograms*, Procedia Environ. Sci. **27** (2015), 16–20.
35. ———, *Some anisotropy indices for the characterization of Brownian textures and their application to breast images*, Spat. Stat. **18** (2016), 147–162.
36. ———, *Tests of isotropy for rough textures of trended images*, Stat. Sinica **26** (2016), no. 3, 1279–1304.
37. ———, *PyAFBF: a Python library for sampling image textures from the anisotropic fractional Brownian field.*, J. Open Source Softw. **7** (2022), no. 75, 3821.
38. F.J.P. Richard and H. Biermé, *Statistical tests of anisotropy for fractional Brownian textures. Application to full-field digital mammography*, J. Math. Imaging Vis. **36** (2010), no. 3, 227–240.
39. A. Ruhe and P. Wedin, *Algorithms for separable nonlinear least squares problems*, SIAM Rev. **22** (1980), no. 3, 318–337.
40. C. Shorten and T. Khoshgoftaar, *A survey on image data augmentation for deep learning*, J. Big data **6** (2019), no. 1, 1–48.
41. J. Suckling et al., *The mammographic image analysis society digital mammogram database*, Digital Mammo (1994), 375–386.
42. T. Van Leeuwen and A. Aravkin, *Variable projection for nonsmooth problems*, SIAM J. Sci. Comput. **43** (2021), no. 5, S249–S268.
43. P. Virtanen, R. Gommers, T. E. Oliphant, et al., *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, Nat. Methods **17** (2020), 261–272.
44. T.H.L. Vu and F.J.P. Richard, *Statistical tests of heterogeneity for anisotropic multifractional Brownian fields*, Stoch. Proc. Appl. **130** (2020), no. 8, 4667–4692.
45. Y. Wang, W. Yin, and J. Zeng, *Global convergence of ADMM in nonconvex nonsmooth optimization*, J. Sci. Comput. **78** (2019), no. 1, 29–63.

Appendix A. Derivatives of $\mathcal{L}$.

Let $\epsilon_n(\tau, \beta) = F_n(\beta)\tau - \widehat{w}_n$ be the n-th residual of the function $\mathcal{L}$.
Gradients of ϵ_n with respect to τ and β are

$$\begin{aligned} \nabla_\tau \epsilon_n(\tau, \beta) &= F_n(\beta)^T, \\ \nabla_\beta \epsilon_n(\tau, \beta) &= DF_n(\beta)^T \tau. \end{aligned} \tag{A.1}$$

where $DF_n(\beta)$ is the jacobian matrix of $F_n(\beta)$ whose terms are

$$DF_{njk}(\beta) = \frac{\partial F_{nj}}{\partial \beta_k}(\beta) = \frac{1}{2} \sum_{m=1}^{M} T_j(\theta_m) B_k(\theta_m) \, \log\left(|\langle x_n, u(\theta_m)\rangle|^2\right) |\langle x_n, u(\theta_m)\rangle|^{2\beta(\theta_m)}. \tag{A.2}$$

The Hessians of ϵ_n are given by

$$\begin{aligned}
\nabla^2_{\tau\tau}\epsilon_n(\tau,\beta) &= 0,\\
\nabla^2_{\beta\beta}\epsilon_n(\tau,\beta) &= \sum_{j=1} \nabla^2 F_{nj}(\beta)\tau_j,\\
\nabla^2_{\beta\tau}\epsilon_n(\tau,\beta) &= DF_n(\beta)^T,\\
\nabla^2_{\tau\beta}\epsilon_n(\tau,\beta) &= DF_n(\beta),
\end{aligned} \tag{A.3}$$

where $\nabla^2 F_{nj}(\beta)$ is the Hessian matrix of F_{nj} whose terms are

(A.4)

$$\begin{aligned}
\nabla^2 F_{njkl}(\beta) &= \frac{\partial^2 F_{nj}}{\partial\beta_l \partial\beta_k}(\beta)\\
&= \frac{1}{2}\sum_{m=1}^{M} T_j(\theta_m)B_k(\theta_m)B_l(\theta_m)\ \left|\log\left(|\langle x_n, u(\theta_m)\rangle|^2\right)\right|^2 |\langle x_n, u(\theta_m)\rangle|^{2\beta(\theta_m)}.
\end{aligned}$$

The first-order derivatives of $\mathcal{L}$ are

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial\tau_j}(\tau,\beta) &= \sum_{n=1}^{N} \frac{\partial\epsilon_n}{\partial\tau_j}(\tau,\beta)\ \epsilon_n(\tau,\beta),\\
\frac{\partial\mathcal{L}}{\partial\beta_k}(\tau,\beta) &= \sum_{n=1}^{N} \frac{\partial\epsilon_n}{\partial\beta_k}(\tau,\beta)\ \epsilon_n(\tau,\beta).
\end{aligned} \tag{A.5}$$

The corresponding gradients of $\mathcal{L}$ with respect to the two variables (τ,β) are

$$\begin{aligned}
\nabla_\tau\mathcal{L}(\tau,\beta) &= \sum_{n=1}^{N} \nabla_\tau\epsilon_n(\tau,\beta)\ \epsilon_n(\tau,\beta),\\
\nabla_\beta\mathcal{L}(\tau,\beta) &= \sum_{n=1}^{N} \nabla_\beta\epsilon_n(\tau,\beta)\ \epsilon_n(\tau,\beta).
\end{aligned} \tag{A.6}$$

Using Equation (A.1), it follows that

$$\begin{aligned}
\nabla_\tau\mathcal{L}(\tau,\beta) &= \sum_{n=1}^{N} F_n(\beta)^T\ \epsilon_n(\tau,\beta),\\
\nabla_\beta\mathcal{L}(\tau,\beta) &= \sum_{n=1}^{N} DF_n(\beta)^T\tau\ \epsilon_n(\tau,\beta).
\end{aligned} \tag{A.7}$$

The second-order partial derivatives of $\mathcal{L}$ are

$$\begin{aligned}
\frac{\partial^2\mathcal{L}}{\partial\tau_m\partial\tau_k}(\beta) &= \sum_{n=1}^{N}\left(\frac{\partial\epsilon_n}{\partial\tau_m}(\tau,\beta)\ \frac{\partial\epsilon_n}{\partial\tau_k}\epsilon_n(\tau,\beta) + \frac{\partial^2\epsilon_n}{\partial\tau_m\partial\tau_k}(\tau,\beta)\ \epsilon_n(\tau,\beta)\right),\\
\frac{\partial^2\mathcal{L}}{\partial\beta_m\partial\beta_k}(\beta) &= \sum_{n=1}^{N}\left(\frac{\partial\epsilon_n}{\partial\beta_m}(\tau,\beta)\ \frac{\partial\epsilon_n}{\partial\beta_k}\epsilon_n(\tau,\beta) + \frac{\partial^2\epsilon_n}{\partial\beta_m\partial\beta_k}(\tau,\beta)\ \epsilon_n(\tau,\beta)\right),\\
\frac{\partial^2\mathcal{L}}{\partial\tau_m\partial\beta_k}(\beta) &= \sum_{n=1}^{N}\left(\frac{\partial\epsilon_n}{\partial\tau_m}(\tau,\beta)\ \frac{\partial\epsilon_n}{\partial\beta_k}\epsilon_n(\tau,\beta) + \frac{\partial^2\epsilon_n}{\partial\tau_m\partial\beta_k}(\tau,\beta)\ \epsilon_n(\tau,\beta)\right).
\end{aligned} \tag{A.8}$$

The Hessians of $\mathcal{L}$ with respect to the two variables (τ,β) are thus given by

$$\begin{aligned}
\nabla^2_{\tau\tau}\mathcal{L}(\tau,\beta) &= \sum_{n=1}^{N} \nabla_\tau \epsilon_n(\tau,\beta)^T \nabla_\tau \epsilon_n(\tau,\beta) + \epsilon_n(\tau,\beta)\, \nabla^2_{\tau\tau}\epsilon_n(\tau,\tau),\\
\nabla^2_{\beta\beta}\mathcal{L}(\tau,\beta) &= \sum_{n=1}^{N} \nabla_\beta \epsilon_n(\tau,\beta)^T \nabla_\beta \epsilon_n(\tau,\beta) + \epsilon_n(\tau,\beta)\, \nabla^2_{\beta\beta}\epsilon_n(\tau,\beta),\\
\nabla^2_{\tau\beta}\mathcal{L}(\tau,\beta) &= \sum_{n=1}^{N} \nabla_\tau \epsilon_n(\tau,\beta)^T \nabla_\beta \epsilon_n(\tau,\beta) + \epsilon_n(\tau,\beta)\, \nabla^2_{\tau\beta}\epsilon_n(\tau,\beta).
\end{aligned} \tag{A.9}$$

Therefore, the Hessians of $\mathcal{L}$ are

$$\begin{aligned}
\nabla^2_{\tau\tau}\mathcal{L}(\tau,\beta) &= \sum_{n=1}^{N} F_n(\beta)^T F_n(\beta),\\
\nabla^2_{\beta\beta}\mathcal{L}(\tau,\beta) &= \tau^T \left(\sum_{n=1}^{N} DF_n(\beta) DF_n(\beta)^T \right) \tau + \sum_{n=1}^{N} \epsilon_n(\tau,\beta) \sum_{j=1} \nabla^2 F_{nj}(\beta)\tau_j,\\
\nabla^2_{\tau\beta}\mathcal{L}(\tau,\beta) &= \sum_{n=1}^{N} F_n(\beta)^T\, DF_n(\beta)^T \tau + \sum_{n=1}^{N} \epsilon_n(\tau,\beta)\, DF_n(\beta).
\end{aligned} \tag{A.10}$$

Appendix B. Derivatives of $\mathcal{K}$.

Using the chain-rule, the first-order partial derivatives of $\mathcal{K}$ are

$$\frac{\partial \mathcal{K}}{\partial \beta_k}(\beta) = \sum_{j=1}^{J} \frac{\partial \mathcal{L}}{\partial \tau_j}(\tau^*(\beta),\beta) \frac{\partial \tau_j^*}{\partial \beta_k}(\beta) + \frac{\partial \mathcal{L}}{\partial \beta_k}(\tau^*(\beta),\beta).$$

But, since $\tau^*(\beta)$ minimizes $\mathcal{L}(\cdot,\beta)$, we have

$$\nabla_\tau \mathcal{L}(\tau^*(\beta),\beta) = 0.$$

So,

$$\frac{\partial \mathcal{K}}{\partial \beta_k}(\beta) = \frac{\partial h}{\partial \beta_k}(\tau^*(\beta),\beta).$$

Hence, the gradient of $\mathcal{K}$ is given by

$$\nabla \mathcal{K}(\theta) = \nabla_\beta \mathcal{L}(\tau^*(\theta),\theta). \tag{B.1}$$

Using Equation (A.7), the expression of the gradient given in Equation (4.3) follows. Furthermore, second-order partial derivatives of $\mathcal{K}$ are given by

$$\frac{\partial^2 \mathcal{K}}{\partial \beta_m \partial \beta_k}(\beta) = \sum_{j=1}^{J} \frac{\partial^2 \mathcal{L}}{\partial \tau_j \partial \beta_k}(\tau^*(\beta),\beta)\, \frac{\partial \tau_j^*}{\beta_m}(\beta) + \frac{\partial^2 \mathcal{L}}{\partial \beta_m \partial \beta_k}(\tau^*(\beta),\beta).$$

Hence, the Hessian of $\mathcal{K}$ is

$$\nabla^2 \mathcal{K}(\theta) = \nabla^2_{\beta\beta}\mathcal{L}(\tau^*(\theta),\theta) + \nabla^2_{\tau\beta}\mathcal{L}(\tau^*(\theta),\theta) D\tau^*(\beta)^T,$$

where $D\tau^*(\beta)$ is the Jacobian of τ^* at β.
Moreover, since

$$u(\beta) = \nabla_\tau \mathcal{L}(\tau^*(\beta),\beta) = 0,$$

we have $\nabla u(\beta) = 0$. Hence, using the chain-rule,

$$\nabla^2_{\tau\tau}\mathcal{L}(\tau^*(\beta), \beta) D\tau^*(\beta)^T + \nabla^2_{\beta\tau}\mathcal{L}(\tau^*(\beta), \beta) = 0$$

Therefore,

$$D\tau^*(\beta)^T = (\nabla^2_{\tau\tau}\mathcal{L}(\tau^*(\beta), \beta))^{-1} \nabla^2_{\beta\tau}\mathcal{L}(\tau^*(\beta), \beta).$$

Consequently, the Hessian of $\mathcal{K}$ can be expressed as

$$\nabla^2\mathcal{K}(\theta) = \nabla^2_{\beta\beta}\mathcal{L}(\tau^*(\theta), \theta) - \nabla^2_{\tau\beta}\mathcal{L}(\tau^*(\theta), \theta)(\nabla^2_{\tau\tau}\mathcal{L}(\tau^*(\beta), \beta))^{-1} \nabla^2_{\beta\tau}\mathcal{L}(\tau^*(\beta), \beta). \tag{B.2}$$

Using Equation (A.10), the expression of the Hessian of $\mathcal{K}$ given by Equation (4.2) follows.

Aix Marseille University, CNRS, Centrale Marseille, I2M, UMR 7373, Marseille, France
Email address: paul.escande@univ-amu.fr

Aix Marseille University, CNRS, Centrale Marseille, I2M, UMR 7373, Marseille, France
Email address: frederic.richard@univ-amu.f

Received January 2, 2023