



**HAL**  
open science

# Nonsmooth Convex Joint Estimation of Local Regularity and Local Variance for Fractal Texture Segmentation

Barbara Pascal, Nelly Pustelnik, Patrice Abry

► **To cite this version:**

Barbara Pascal, Nelly Pustelnik, Patrice Abry. Nonsmooth Convex Joint Estimation of Local Regularity and Local Variance for Fractal Texture Segmentation. 2019. hal-02346159v1

**HAL Id: hal-02346159**

**<https://hal.science/hal-02346159v1>**

Preprint submitted on 4 Nov 2019 (v1), last revised 3 Apr 2021 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Nonsmooth Convex Joint Estimation of Local Regularity and Local Variance for Fractal Texture Segmentation

Barbara Pascal, Nelly Pustelnik and Patrice Abry<sup>\*</sup>

October 11, 2019

## 1 Introduction

**Context.** Segmentation (partitioning images into *homogeneous* regions) constitutes an on-going and major challenge for image analysis and computer vision. There exist numerous well-established and efficient methods to partition images according to their geometrical properties (cartoon) into homogeneous color (or grayscale) regions (cf. e.g., [17, 21, 30, 38]). Segmentation of images in terms of piecewise homogeneous textures [29] though remains particularly challenging, when both the characteristics of the textures in each regions and the boundaries between regions have to be estimated jointly. Textured images appear as *natural* models in a large variety of applications: Biologic tissues analysis [28], tomography-based pathology diagnostic [18, 27], art painting expertise [1], microfluidics [34]. In these examples, accuracy in interface detection strongly impacts medical diagnosis or physical interpretation. Along another line, texture segmentation in real-world applications often entails processing both large-size images and large-size databases, typically sequences of several hundreds of  $10^7$  pixel images, cf. e.g. [34], thus precluding the use of most state-of-the-art yet too computationally demanding procedures, thus calling for solutions combining both high performance and controlled computational costs.

**Related work.** Most bottom-up segmentation procedures rely on first computing local features representing textures, such as Gabor or short-term Fourier transform coefficients [16, 20], fractal dimension [10], Amplitude/Frequency Modulation models [22]. These local features are then aggregated within regions where they show homogeneous statistics. For example, in [47], local Gabor coefficient histograms were spatially grouped using matrix factorization ; in [3], textons were combined with brightness and color features to yield a multiscale contour detection procedure. More recently, fractal (or scale-free) features were also involved in texture segmentation (cf. e.g., [46]). Notably, *local regularity* was used to quantify the fluctuations of roughness along the texture. Local regularity is quantified as an *optimal* local power-law behavior across scales

---

<sup>\*</sup>B. Pascal, N. Pustelnik and P. Abry are with the Université de Lyon, ENS de Lyon, CNRS, Laboratoire de Physique, Lyon, France, e-mail: `firstname.lastname@ens-lyon.fr`

<sup>†</sup>This work was supported by Defi Imag'in SIROCCO and by ANR-16-CE33-0020 Multi-Fracs, France.

for some multiscale quantities [46]. Modulus of wavelet coefficients [25] were initially used as multiscale quantities, followed by continuous wavelet transform modulus maxima [26, 31]. More recently, *wavelet leaders* (local suprema of discrete wavelet coefficients), used here, were shown to permit a theoretically accurate and practically robust estimation of local regularity [46], and successfully involved in texture characterization in e.g., [1, 27, 32]. Wavelet leaders are used to estimate local regularity mostly by linear regressions in log-log coordinates. Local estimates are further grouped together in piecewise homogeneous regions. For instance, in [37], it was proposed to *force* piecewise constant estimates by regularization, using Rudin-Osher-Fatemi (ROF) model [41] (i.e. *total variation* (TV) denoising). Further in [37], it was also attempted to combine both steps into a single one by incorporating the regression weight estimation into the optimization procedure. The high computational burden implied by iteratively solving optimization problems as well as by tuning the regularization parameter has tentatively been addressed by *block splitting approaches* as suggested in [39] and explored in [33]. Strong convexity [9] constitutes another recently proposed track to address iterative optimization acceleration. However, the proposed procedures, despite showing satisfactory and state-of-the-art estimation/segmentation performance suffered from two major limitations: i) Their being computationally very intensive preclude their use of large-size images and databases ; ii) While based on a key fractal feature, local regularity, they neglect changes in power, a potentially relevant information for texture segmentation, notably to extract accurate region boundaries.

**Goals, contributions and outline.** Aiming to address the two above limitations, the goals of the present contribution are first to investigate the benefits in texture segmentation potentially brought by the joint estimations of local variance and local regularity, by means of nonsmooth optimization designed to favor piecewise regions consisting of homogeneous textures, and second to devise fast and efficient algorithms, for minimizing well-designed functionals. The key concepts and definitions related to local regularity, wavelet leaders and corresponding state-of-the-art linear regression and TV-based estimation procedures are recalled in Section 2. Two objective functions for the estimation of both local variance and regularity, referred to as *joint* and *coupled*, are proposed in Section 3, based on a same data fidelity term but on two different total-variation based regularization strategies. Studying strong convexity and duality gaps, two classes of fast iterative algorithms (*Primal-Dual* and *Dual Forward-backward*), are devised. Piecewise homogeneous fractal textures are defined in Section 4, thus enabling performance assessment from Monte-Carlo simulations on well-controlled synthetic textures, as reported in Section 5. Conducting such simulations requires to address issues related to regularization parameter selection, iterative algorithm stopping criterion, and performance metrics. It permits relevant answers to the final aim of assessing the actual benefits of using both local variance and regularity in texture segmentation with respect to issues such as sensitivity to fractal parameter changes, computational costs, impact of the different optimization formulations. A MATLAB toolbox implementing the analysis and synthesis procedures devised here will be made freely and publicly available at the time of publication.

## 2 Local regularity

### 2.1 Local regularity and Hölder exponent

Local regularity can classically be assessed by means of Hölder exponent [19,24], defined as follows:

**Definition 1** Let  $f: \Omega \rightarrow \mathbb{R}$  denote a 2D real field defined on an open set  $\Omega \subset \mathbb{R}^2$ . The Hölder exponent  $h(\underline{z}_0)$  at location  $\underline{z}_0 \in \Omega$  is defined as the largest  $\alpha > 0$  such that there exists a constant  $\chi > 0$ , a polynomial  $\mathcal{P}_{\underline{z}_0}$  of degree lower than  $\alpha$  and a neighborhood  $\mathcal{V}(\underline{z}_0)$  satisfying:

$$(\forall \underline{z} \in \mathcal{V}(\underline{z}_0)) \quad |f(\underline{z}) - \mathcal{P}_{\underline{z}_0}(\underline{z})| \leq \chi \|\underline{z} - \underline{z}_0\|^\alpha$$

where  $\|\cdot\|$  denotes the Euclidian norm.

Definition 1 does not provide a practical way to estimate local regularity. Thus, the practical assessment of  $h(\underline{z}_0)$  usually relies on the use of multiscale quantities, such as wavelet coefficients, or wavelet leaders [19,24,46].

### 2.2 Local regularity and wavelet leaders

Because the practical aim is to analyse digitized images, definitions are given in a discrete setting. Let  $X = (X_{\underline{n}})_{\underline{n} \in \Upsilon} \in \mathbb{R}^{|\Upsilon|}$  denote the digitized version of the 2D real field  $f$  on a finite grid  $\Upsilon = \{1, \dots, N\}^2$ . Let  $d_j = \mathbf{W}_j X$  denote the discrete wavelet transform (DWT) coefficients of  $X$ , at resolution  $j \in \{j_1, \dots, j_2\}$ , with  $\mathbf{W}_j: \mathbb{R}^{N \times N} \rightarrow \mathbb{R}^{M_j \times M_j}$  the operator formulation of the DWT.

Further, let the wavelet leader  $\mathcal{L}_{j,\underline{k}}$ , at scale  $2^j$  and location  $\underline{n} = 2^j \underline{k}$ , be defined as the local supremum of modulus of wavelet coefficients in a small neighborhood across all finer scales, [19,24,46]  $\mathcal{L}_{j,\underline{k}} = \sup_{\lambda_{j',\underline{k}'} \subset 3\lambda_{j,\underline{k}}} |2^j d_{j',\underline{k}'}|$ , with

$$\lambda_{j,\underline{k}} = [\underline{k}2^j, (\underline{k} + 1)2^j] \text{ and } 3\lambda_{j,\underline{k}} = \bigcup_{p \in \{-1,0,1\}^2} \lambda_{j,\underline{k}+p}.$$

It was proven in [19,24,46] that wavelet leaders provide multiscale quantities intrinsically tied to Hölder exponents, insofar as when  $X$  has Hölder exponent  $h_{\underline{n}}$  at location  $\underline{n}$ , when the wavelet leaders  $\mathcal{L}_{j,\underline{k}}$  such that  $\underline{n} \in \lambda_{j,\underline{k}}, \forall j$  satisfy:

$$\mathcal{L}_{j,\underline{k}} \simeq \eta_{\underline{n}} 2^{jh_{\underline{n}}} \text{ as } 2^j \rightarrow 0, \text{ with } \underline{n} = 2^j \underline{k}, \quad (1)$$

with  $\eta_{\underline{n}}$  proportional to local variance of  $X$  around pixel  $\underline{n}$ .

### 2.3 Local regularity estimation

#### 2.3.1 Linear regressions

Eq. (1) naturally leads to estimate  $\mathbf{h}$  and  $\mathbf{v} = \log_2 \boldsymbol{\eta}$  by means of linear regressions in log-log coordinates [19,24,46], denoted  $(\widehat{\mathbf{v}}_{\text{LR}}, \widehat{\mathbf{h}}_{\text{LR}})$ :

$$(\forall \underline{n} \in \Upsilon) \quad \begin{pmatrix} \widehat{\mathbf{v}}_{\text{LR},\underline{n}} \\ \widehat{\mathbf{h}}_{\text{LR},\underline{n}} \end{pmatrix} = \mathbf{J}^{-1} \begin{pmatrix} \mathcal{S}_{\underline{n}} \\ \mathcal{T}_{\underline{n}} \end{pmatrix} \quad (2)$$

$$\text{with } \mathbf{J} = \begin{pmatrix} R_0 & R_1 \\ R_1 & R_2 \end{pmatrix} \text{ and } R_m = \sum_j j^m, \quad (3)$$

$$\text{and } \mathcal{S}_{\underline{n}} = \sum_j \log_2 \mathcal{L}_{j,\underline{n}}, \quad \mathcal{T}_{\underline{n}} = \sum_j j \log_2 \mathcal{L}_{j,\underline{n}}. \quad (4)$$

The sums  $\sum_j$  implicitly stand for  $\sum_{j=j_1}^{j_2}$  with  $(j_1, j_2)$  the range of octaves involved in the estimation. Though linear regressions provide estimates both for  $\mathbf{h}$  and  $\log_2 \boldsymbol{\eta}$ , the later remains to date rarely used. Estimates of  $\mathbf{h}$  were for instance used in [33, 37]. Bayesian based extension of linear regressions were also and alternatively proposed to estimate  $\boldsymbol{\eta}$  and  $h$  [45].

A sample of piecewise homogeneous fractal texture is displayed in Fig. 1b, synthesized using the mask sketched in Fig. 1a. The corresponding linear regression based estimates  $\hat{\boldsymbol{\eta}}_{\text{LR}} \equiv 2^{\hat{\mathbf{v}}_{\text{LR}}}$ , and  $\hat{\mathbf{h}}_{\text{LR}}$  (Fig. 1(c-d)) show too poor performance (high variability) to permit to detect the two regions and the corresponding boundary.

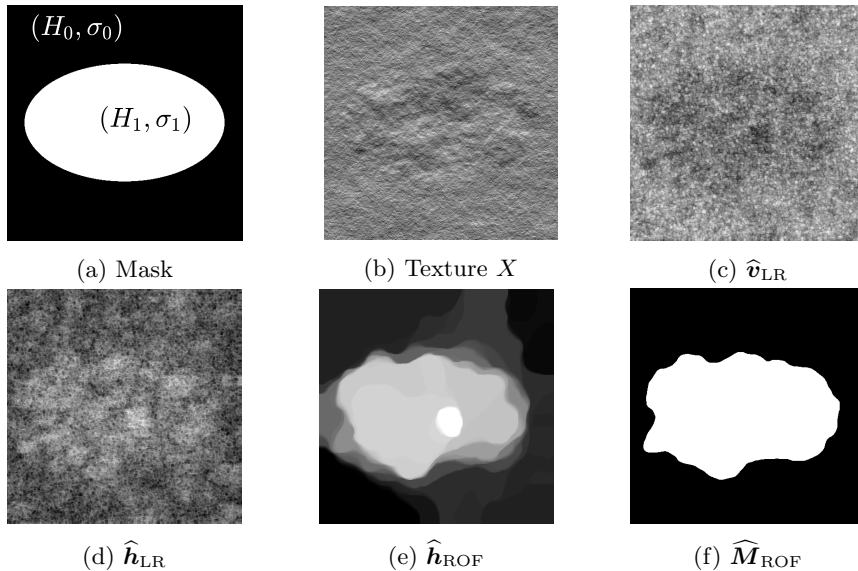


Figure 1: **Piecewise homogeneous fractal texture and local variance and regularity estimates.** (a) Synthesis mask  $(H_0, \sigma_0) = (0.5, 0.6)$  and  $(H_1, \sigma_1) = (0.8, 0.65)$ ; (b) sample texture (see Section 4), (c) and (d) linear regression based estimates of local variance and regularity  $\hat{\mathbf{v}}_{\text{LR}}$  and  $\hat{\mathbf{h}}_{\text{LR}}$ ; (e) and (f) Total variation based estimates of local regularity  $\hat{\mathbf{h}}_{\text{ROF}}$  and segmentation  $\hat{\mathbf{M}}_{\text{ROF}}$  obtained with Alg. 1.

### 2.3.2 Total variation based estimates

To address the poor estimation performance achieved by linear regressions for local estimation, it has been proposed to *denoise*  $\hat{\mathbf{h}}_{\text{LR}}$  by using an optimisation framework involving TV regularization, as a mean to enforce piecewise constant estimates [37]:

$$\hat{\mathbf{h}}_{\text{ROF}} = \underset{\mathbf{h} \in \mathbb{R}^{|\mathcal{Y}|}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{h} - \hat{\mathbf{h}}_{\text{LR}}\|_2^2 + \lambda \operatorname{TV}(\mathbf{h}). \quad (5)$$

The above functional balances a data fidelity term versus a total variation penalization defined as:

$$\text{TV}(\mathbf{h}) = \|\mathbf{D}\mathbf{h}\|_{2,1}, \quad (6)$$

where the operator  $\mathbf{D} : \mathbb{R}^{|\Upsilon|} \rightarrow \mathbb{R}^{2 \times |\Upsilon|}$ , consists of horizontal and vertical increments:

$$(\mathbf{D}\mathbf{h})_{n_1, n_2} = \begin{pmatrix} h_{n_1, n_2+1} - h_{n_1, n_2} \\ h_{n_1+1, n_2} - h_{n_1, n_2} \end{pmatrix} \quad (7)$$

and where the mixed norm,  $\|\cdot\|_{2,1}$ , defined for  $\mathbf{z} = [\mathbf{z}_1; \dots; \mathbf{z}_I]^T \in \mathbb{R}^{I \times |\Upsilon|}$  as:

$$\|\mathbf{z}\|_{2,1} = \sum_{n_1=1}^{N_1-1} \sum_{n_2=1}^{N_2-1} \sqrt{\sum_{i=1}^I (z_i)_{n_1, n_2}^2} \quad (8)$$

both ensures an isotropic and piecewise constant behaviour of the estimates [37].

While showing substantial benefits compared to linear regression based estimates and although fast, the TV-based procedure above, proposed in [37], suffers from several significant limitations: i) It is restricted to the estimation of local regularity only and neglects local variance ; ii) It consists of a two-step process (first apply linear regression to obtain  $\hat{\mathbf{h}}_{\text{LR}}$ , second apply TV to  $\hat{\mathbf{h}}_{\text{LR}}$  to obtain  $\hat{\mathbf{h}}_{\text{ROF}}$ ) and is hence potentially not optimal. Section 3 below will propose several solutions that address these two limitations by performing the estimate of both local variance and regularity in a single pass at similar computational cost than the two-step procedure.

### 2.3.3 A posteriori segmentation and local regularity global estimation

When the number of regions is known *a priori*, it has been shown in [6, 7] that a fast iterative thresholding post-processing procedure can be applied to  $\hat{\mathbf{h}}_{\text{ROF}}$  to obtain a segmentation of the TV denoised estimate. In [6, 7] were provided theoretical guarantees linking this Threshold-ROF (T-ROF) strategy to Mumford-Shah-like segmentation. Alg. 1 explicitly describes the procedure applied to  $\hat{\mathbf{h}}_{\text{ROF}}$  to obtain a two-region T-ROF segmentation  $\hat{\mathbf{M}}_{\text{ROF}}$  with  $\hat{M}_{\text{ROF}, \underline{n}} = 0$  if  $n \in \Upsilon_0$  and 1 otherwise. Figs. 1e, 1f report examples of TV denoised estimate  $\hat{\mathbf{h}}_{\text{ROF}}$  and T-ROF segmentation  $\hat{\mathbf{M}}_{\text{ROF}}$ .

From T-ROF segmentation, Texture X can be interpreted as the concatenation of two fractal textures  $X_0$  and  $X_1$ , each with its own uniform regularity and variance. Therefore, a posteriori global estimates,  $\hat{H}_{0, \text{ROF}}$  and  $\hat{H}_{1, \text{ROF}}$ , for the local regularity of each region,  $\Upsilon_0$  and  $\Upsilon_1$ , can be obtained by performing linear regressions applied to the logarithm of multiscale quantities  $\mathcal{L}_j$  averaged over  $\Upsilon_0$  (resp.  $\Upsilon_1$ ) (cf. [46]).

---

**Algorithm 1:** T-ROF: iterative thresholding of  $\widehat{\mathbf{h}}_{\text{ROF}}$

---

**Input:**  $\mathbf{h}$

**Initialization:**  $m_0^{[0]} = \min_{n \in \Upsilon} h_n$ ,  $m_1^{[0]} = \max_{n \in \Upsilon} h_n$ .

**for**  $t \in \mathbb{N}^*$  **do**

*Compute the threshold:*

$$T^{[t-1]} = (m_0^{[t-1]} + m_1^{[t-1]}) / 2$$

*Threshold  $\mathbf{h}$ :*

$$\Upsilon_0^{[t]} = \{n \mid h_n \leq T^{[t]}\}, \quad \Upsilon_1^{[t]} = \{n \mid h_n > T^{[t]}\}$$

*Update region mean:*

$$m_0^{[t]} = 1/|\Upsilon_0| \sum_{n \in \Upsilon_0} h_n, \quad m_1^{[t]} = 1/|\Upsilon_1| \sum_{n \in \Upsilon_1} h_n.$$

**Output:**  $\Upsilon_0 = \Upsilon_0^{[\infty]}$ ,  $\Upsilon_1 = \Upsilon_1^{[\infty]}$

---

### 3 Total variation based estimation of local variance and regularity

#### 3.1 Design of the objective function

##### 3.1.1 Linear regressions as functional minimization

Linear regressions can be viewed as optimization schemes. Indeed, setting  $\mathbf{v} = \log_2 \boldsymbol{\eta}$  the following strictly convex functional (in variables  $(\mathbf{v}, \mathbf{h})$ ):

$$\Phi(\mathbf{v}, \mathbf{h}; \mathcal{L}) = \frac{1}{2} \sum_j \|\mathbf{v} + j\mathbf{h} - \log_2 \mathcal{L}_j\|_{\text{Fro}}^2, \quad (9)$$

has a unique minimum corresponding to the linear regression estimates in 2.

##### 3.1.2 Penalization

To improve on the poor performance of linear regressions, a generic optimization framework can be proposed as  $(\widehat{\mathbf{v}}, \widehat{\mathbf{h}}) = \arg \min_{\mathbf{v} \in \mathbb{R}^{|\Upsilon|}, \mathbf{h} \in \mathbb{R}^{|\Upsilon|}} \Phi(\mathbf{v}, \mathbf{h}; \mathcal{L}) + \Psi(\mathbf{v}, \mathbf{h})$ , where

the data fidelity term  $\Phi$  is that of the linear regression (cf.(9)), and where the penalization term  $\Psi$  favors piecewise constancy of  $\mathbf{v}$  and  $\mathbf{h}$ . Two different strategies are proposed:

(i) *Joint estimation:*

$$(\widehat{\mathbf{v}}_J, \widehat{\mathbf{h}}_J) = \arg \min_{\mathbf{v}, \mathbf{h} \in \mathbb{R}^{|\Upsilon|} \times \mathbb{R}^{|\Upsilon|}} \Phi(\mathbf{v}, \mathbf{h}; \mathcal{L}) + \Psi_J(\mathbf{v}, \mathbf{h}), \quad (10)$$

$$\text{with } \Psi_J(\mathbf{v}, \mathbf{h}) = \lambda(\text{TV}(\mathbf{v}) + \alpha \text{TV}(\mathbf{h})), \quad (11)$$

that does not favor changes in  $\mathbf{v}$  and  $\mathbf{h}$  that occur at same location;

(ii) *Coupled estimation:*

$$(\widehat{\mathbf{v}}_C, \widehat{\mathbf{h}}_C) = \arg \min_{\mathbf{v}, \mathbf{h} \in \mathbb{R}^{|\Upsilon|} \times \mathbb{R}^{|\Upsilon|}} \Phi(\mathbf{v}, \mathbf{h}; \mathcal{L}) + \Psi_C(\mathbf{v}, \mathbf{h}), \quad (12)$$

$$\begin{aligned} \text{with } \Psi_C(\mathbf{v}, \mathbf{h}) &= \lambda \text{TV}_\alpha(\mathbf{v}, \mathbf{h}), \\ &= \lambda \|\mathbf{D}\mathbf{v}; \alpha \mathbf{D}\mathbf{h}\|_{2,1}^\top \end{aligned} \quad (13)$$

where  $\text{TV}_\alpha$  couples spatial variations of  $v_{\underline{n}}$  and  $h_{\underline{n}}$  and thus favor their occurrences at same location.

For both constructions,  $\lambda$  and  $\alpha$  constitute regularization hyperparameters that need to be selected.

Further, the thresholding procedure in [6, 7] is generalized to “*joint*” and “*coupled*” estimation strategies: Alg. 1, applied to  $\hat{\mathbf{h}}_J$  (resp.  $\hat{\mathbf{h}}_C$ ) provides “*joint*” (resp. “*coupled*”) segmentation  $\widehat{\mathbf{M}}_J$  (resp.  $\widehat{\mathbf{M}}_C$ ).

Finally, from the two region segmentation  $\widehat{\mathbf{M}}_J$  (resp.  $\widehat{\mathbf{M}}_C$ ) obtained with Alg. 1, one can use global techniques [46] to obtain *a posteriori* estimates of the Hölder exponents of each region  $\widehat{H}_{0,J}$  and  $\widehat{H}_{1,J}$  (resp.  $\widehat{H}_{0,C}$  and  $\widehat{H}_{1,C}$ ) (cf. Sec. 2.3.2).

## 3.2 Minimization algorithms

The data fidelity function  $\Phi$  is Lipschitz differentiable. This is the case neither for  $\Psi_J$  nor for  $\Psi_C$ . Therefore, gradient descent methods are not appropriate to solve (10) or (12). Instead, the use of proximal algorithms is adapted to the minimization of nonsmooth functions [13]. While the non-differentiability stems from the mixed norm  $\|\cdot\|_{2,1}$ , appearing both in  $\Psi_J$  and in  $\Psi_C$ , the linear operator  $\mathbf{D}$  into the nonsmooth term makes the computation of the proximal operator of total variation difficult [11, 36]. Below two algorithms are devised: the dual FISTA and the primal-dual.

### 3.2.1 Dual FISTA

The well-known Fast Iterative Shrinkage Thresholding Algorithm (FISTA) algorithm [8] is here customized to Pb (10) and (12), to achieve faster convergences than with the basic dual forward-backward. Corresponding iterations are detailed in Algorithms 2 and 4. Convergence guarantees are specified in Theorems 1 and 2.

**Theorem 1 (Convergence of FISTA<sub>J</sub>)** *The sequence  $(\mathbf{v}^{[t]}, \mathbf{h}^{[t]})$  generated by Algorithm 4 converges towards a solution of the joint estimation problem (10).*

**Theorem 2 (Convergence of FISTA<sub>C</sub>)** *The sequence  $(\mathbf{v}^{[t]}, \mathbf{h}^{[t]})$  generated by Algorithm 2 converges towards a solution of coupled estimation problem (12).*

**Proof 1** *The proofs of Theorems 1 and 2 stem directly from the choice of descent parameters proposed in Algs. 2 and 4 following the reasoning in [8].*

### 3.2.2 Primal-dual

Primal-dual algorithms [9, 14, 23, 43] can also be customized to solve Problems (10) and (12).

This requires the derivation of a closed form expression for the proximal operator associated with the quadratic data fidelity term  $\Phi$ , provided in Proposition 1.



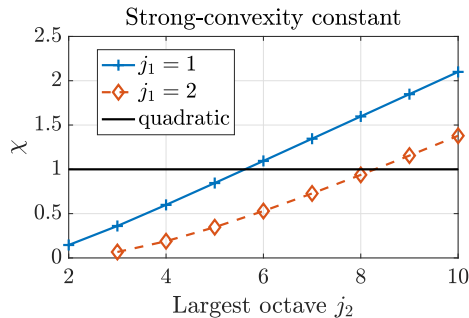


Figure 2: **Strong convexity constant** as a function of the range of scales involved in the estimation.

**Proposition 1 (Computation of  $\text{prox}_{\Phi}$ )** For every  $(\mathbf{v}, \mathbf{h}) \in \mathbb{R}^{|\mathcal{T}|} \times \mathbb{R}^{|\mathcal{T}|}$ , denoting  $(\mathbf{p}, \mathbf{q}) = \text{prox}_{\Phi}(\mathbf{v}, \mathbf{h}) \in \mathbb{R}^{|\mathcal{T}|} \times \mathbb{R}^{|\mathcal{T}|}$  one has

$$\begin{cases} \mathbf{p} = \frac{(1+R_2)(\mathcal{S}+\mathbf{v})-R_1(\mathcal{T}+\mathbf{h})}{(1+R_0)(1+R_2)-R_1^2}, \\ \mathbf{q} = \frac{(1+R_0)(\mathcal{T}+\mathbf{h})-R_1(\mathcal{S}+\mathbf{v})}{(1+R_0)(1+R_2)-R_1^2}. \end{cases}$$

with  $\mathcal{S}$  and  $\mathcal{T}$  defined in (4) and  $R_0, R_1, R_2$  defined in (3).

Further, in [9], it was described how to take advantage of the strong convexity<sup>1</sup> of the objective function to obtain fast implementation for primal-dual algorithms, with linear convergence rates. In Proposition 2 below, we prove that the data fidelity term  $\Phi$  is strongly convex and derive the closed form expression of the strong convexity coefficient.

**Proposition 2** Function  $\Phi(\mathbf{v}, \mathbf{h}; \mathcal{L})$  is  $\mu$ -strongly convex w.r.t the variables  $(\mathbf{v}, \mathbf{h})$ , with  $\mu = \chi$  where  $\chi > 0$  is the lowest eigenvalue of the symmetric and positive definite matrix  $\mathbf{J}$  defined in Eq. (3).

Indeed, since  $\nabla\Phi(\mathbf{v}, \mathbf{h}; \mathcal{L}) = \mathbf{J}(\mathbf{v}, \mathbf{h})^\top$  is linear, the condition for  $\Phi$  to be  $\mu$ -strongly convex can be recasted as:

$$\begin{aligned} \langle \nabla\Phi(\mathbf{v}, \mathbf{h}; \mathcal{L}), (\mathbf{v}, \mathbf{h}) \rangle &= \langle \mathbf{J}(\mathbf{v}, \mathbf{h})^\top, (\mathbf{v}, \mathbf{h})^\top \rangle \\ &\geq \mu \|(\mathbf{v}, \mathbf{h})\|^2 \end{aligned} \quad (14)$$

Intuitively a function with a large strong-convexity constant  $\mu$  is very stitched around its minimum, hence yielding a good theoretical convergence rate toward the minimizer. The strong-convexity parameter is represented in Fig. 2 as a function of the range of scales involved in the estimation procedure (cf. (9)).

The iterations of the devised fast primal-dual algorithms are detailed in Algorithms 3 and 5. Convergence guarantees are specified in Theorem 3 (resp. Theorem 4).

<sup>1</sup>A function  $\varphi : \mathcal{H} \rightarrow \mathbb{R}$ , defined on an Hilbert space  $\mathcal{H}$ , is said to be  $\mu$ -strongly convex, for a given  $\mu > 0$  if the function  $y \mapsto \varphi(y) - \frac{\mu}{2} \|y\|_2^2$  is convex. When the function  $\varphi$  is differentiable,  $\varphi$  is  $\mu$ -strongly convex if and only if  $(\forall (y, z) \in \mathcal{H} \times \mathcal{H}) \quad \langle \nabla\varphi(y) - \nabla\varphi(z), y - z \rangle \geq \mu \|y - z\|^2$  where  $\langle \cdot, \cdot \rangle$  denotes the Hilbert scalar product and  $\|\cdot\|$  the associated scalar product.

**Theorem 3 (Convergence of PD<sub>J</sub>)** *The sequence  $(\mathbf{v}^{[l]}, \mathbf{h}^{[l]})$  generated by Algorithm 3 converges towards a solution of the joint estimation problem (10).*

**Theorem 4 (Convergence of PD<sub>C</sub>)** *The sequence  $(\mathbf{v}^{[l]}, \mathbf{h}^{[l]})$  generated by Algorithm 5 converges towards a solution of the joint estimation problem (12).*

**Proof 2** *The proofs of Theorems 3 and 4 come directly from the choice of descent parameters proposed in Algs. 3 and 5 following the reasoning of [9].*

### 3.3 Duality gap

To ensure fair comparisons between the different algorithms, we construct a stopping criterion based on the *duality gap*. Let  $\mathcal{H}, \mathcal{G}$  Hilbert spaces,  $\Theta : \mathcal{H} \rightarrow ]-\infty, +\infty]$ ,  $\Xi : \mathcal{G} \rightarrow ]-\infty, +\infty]$  and  $\mathbf{L} : \mathcal{H} \rightarrow \mathcal{G}$  a bounded linear operator. From a primal optimization problem of the general form  $\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \mathcal{H}} \Theta(\mathbf{x}) + \Xi(\mathbf{L}\mathbf{x})$ , and its associated dual problem  $\hat{\mathbf{y}} = \arg \min_{\mathbf{y} \in \mathcal{G}} \Theta^*(-\mathbf{L}^*\mathbf{y}) + \Xi^*(\mathbf{y})$ , we define:

$$\Gamma(\mathbf{x}; \mathbf{y}) := \Theta(\mathbf{x}) + \Xi(\mathbf{L}\mathbf{x}) + \Theta^*(-\mathbf{L}^*\mathbf{y}) + \Xi^*(\mathbf{y}), \quad (15)$$

$\Theta^*$  (resp.  $\Xi^*$ ), being the Fenchel conjugate of  $\Theta$  (resp.  $\Xi$ ).

**Definition 2** *The duality gap  $\delta\Gamma$  associated to primal and dual optimization problems is defined as the infimum:*

$$\delta\Gamma := \inf_{(\mathbf{x}, \mathbf{y}) \in \mathcal{H} \times \mathcal{G}} \Gamma(\mathbf{x}; \mathbf{y}). \quad (16)$$

Under some loose assumptions on  $\Theta$ ,  $\Xi$  and  $\mathbf{L}$ , referred as *strong duality* in [4], the duality gap  $\delta\Gamma$  is zero and the respective solutions  $\hat{\mathbf{x}}$  and  $\hat{\mathbf{y}}$  of primal and dual problems are characterized by  $\Gamma(\hat{\mathbf{x}}; \hat{\mathbf{y}}) = \delta\Gamma = 0$ .

Primal-dual algorithms presented above built a minimizing sequence  $(\mathbf{x}^{[t]}, \mathbf{y}^{[t]})_{t \in \mathbb{N}}$  for  $\Gamma$ , therefore converging towards the unique point achieving the infimum of Eq. (16):  $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ . Then the convergence of  $\mathbf{x}^{[t]}$  (resp.  $\mathbf{y}^{[t]}$ ) towards  $\hat{\mathbf{x}}$  (resp.  $\hat{\mathbf{y}}$ ) can be measured evaluating the primal-dual functional  $\Gamma$ , as:

$$\Gamma(\mathbf{x}^{[t]}; \mathbf{y}^{[t]}) \xrightarrow[t \rightarrow \infty]{} 0. \quad (17)$$

The “*joint*” (10) and “*coupled*” (12) optimization problems, with primal variable  $\mathbf{x} = (\mathbf{v}, \mathbf{h}) \in \mathbb{R}^{|\Upsilon|} \times \mathbb{R}^{|\Upsilon|}$ , share the same data fidelity term  $\Theta(\mathbf{x}) = \Phi(\mathbf{v}, \mathbf{h}; \mathcal{L})$  and a penalization based on the mixed norm  $\Xi = \lambda \|\cdot\|_{2,1}$ . Yet, the linear operator  $\mathbf{L}$  differ between *joint* and *coupled* formulations, as:

- i)  $\mathbf{L}\mathbf{x} = [\mathbf{D}\mathbf{v}, \alpha\mathbf{D}\mathbf{h}] \in \mathbb{R}^{2 \times 2|\Upsilon|}$ , for “*joint*” (10),
- ii)  $\mathbf{L}\mathbf{x} = [\mathbf{D}\mathbf{v}; \alpha\mathbf{D}\mathbf{h}] \in \mathbb{R}^{4 \times |\Upsilon|}$ , for “*coupled*” (12).

For Pb. (10) (resp. (12)), denoting  $\mathbf{y} = (\mathbf{u}, \boldsymbol{\ell})$  the dual variable, the evaluation of  $\Gamma_{\text{J (resp. C)}}(\mathbf{v}, \mathbf{h}; \mathbf{u}, \boldsymbol{\ell})$  requires computing:

- i)  $\mathbf{L}^*$  (straightforward from  $\mathbf{D}^*$ ),
- ii)  $\Xi^* = (\lambda \|\cdot\|_{2,1})^* = \iota_{\|\cdot\|_{2,+\infty} \leq \lambda}$  (direct computation),
- iii)  $\Theta^* = \Phi^*$  which is devised in Proposition 3.

**Proposition 3** Let  $\mathbf{J}$  and  $(\mathcal{S}, \mathcal{T})$  defined in (3) and (4),

$$\begin{aligned}\Phi^*(\mathbf{v}, \mathbf{h}; \mathcal{L}) &= \frac{1}{2} \langle (\mathbf{v}, \mathbf{h})^\top, \mathbf{J}^{-1}(\mathbf{v}, \mathbf{h})^\top \rangle \\ &\quad + \langle (\mathcal{S}, \mathcal{T})^\top, \mathbf{J}^{-1}(\mathbf{v}, \mathbf{h})^\top \rangle + \mathcal{C}, \\ \mathcal{C} &= \frac{1}{2} \langle (\mathcal{S}, \mathcal{T})^\top, \mathbf{J}^{-1}(\mathcal{S}, \mathcal{T})^\top \rangle - \frac{1}{2} \sum_j (\log_2 \mathcal{L}_j)^2.\end{aligned}\tag{18}$$

**Proof 3** By definition of the Fenchel conjugate,

$$F^*(\mathbf{v}, \mathbf{h}; \mathcal{L}) = \sup_{\tilde{\mathbf{v}} \in \mathbb{R}^{|\Omega|}, \tilde{\mathbf{h}} \in \mathbb{R}^{|\Omega|}} \langle \tilde{\mathbf{v}}, \mathbf{v} \rangle + \langle \tilde{\mathbf{h}}, \mathbf{h} \rangle - F(\tilde{\mathbf{v}}, \tilde{\mathbf{h}}; \mathcal{L}).\tag{19}$$

The supremum is obtained at  $(\bar{\mathbf{v}}, \bar{\mathbf{h}})$  such that, for every  $\underline{n} \in \Omega$ ,

$$\begin{cases} v_{\underline{n}} - \sum_j (\bar{v}_{\underline{n}} + j\bar{h}_{\underline{n}} - \log_2 \mathcal{L}_{j,\underline{n}}) = 0 \\ h_{\underline{n}} - \sum_j j (\bar{v}_{\underline{n}} + j\bar{h}_{\underline{n}} - \log_2 \mathcal{L}_{j,\underline{n}}) = 0. \end{cases}\tag{20}$$

or equivalently,

$$\begin{cases} R_0 \bar{v}_{\underline{n}} + R_1 \bar{h}_{\underline{n}} = v_{\underline{n}} + \mathcal{S}_{\underline{n}} \\ R_1 \bar{v}_{\underline{n}} + R_2 \bar{h}_{\underline{n}} = h_{\underline{n}} + \mathcal{T}_{\underline{n}} \end{cases}\tag{21}$$

that yields to

$$\begin{pmatrix} \bar{v}_{\underline{n}} \\ \bar{h}_{\underline{n}} \end{pmatrix} = \mathbf{J}^{-1} \begin{pmatrix} v_{\underline{n}} + \mathcal{S}_{\underline{n}} \\ h_{\underline{n}} + \mathcal{T}_{\underline{n}} \end{pmatrix}\tag{22}$$

and it is then necessary to re-inject this expression into the explicit expression of  $\Psi$ .

## 4 Piecewise homogeneous fractal textures

**Homogeneous fractal textures.** Numerous models of homogeneous textures were proposed in the literature, mostly consisting of 2D extensions (referred to as fractional Brownian field fBf) of fractional Brownian motion, the unique univariate Gaussian exactly selfsimilar process with stationary increments [12]. However, in most applications, real world textures are better modeled by stationary processes, hence by increments of fBf, referred to as fractional Gaussian fields. Further, to construct a formally relevant definition for piecewise homogeneous fractal textures, as originally proposed here, it is easier to *manipulate* stationary processes.

For Gaussian processes, space or frequency domain definitions of fBf are theoretically equivalent and several variations were proposed in either domain, mostly consisting in different tuning of the 2D extension of the fractional integration kernel underlying selfsimilar textures (cf. e.g., [2, 12, 35]). Yet, in practice, numerical issues needs to be accounted for, such as discrete sampling from a continuous process, or computation of integrals defined from a theoretically infinite support kernel and related border effects. Further, fast circulant embedding matrix algorithms were proposed, that yet do not permit to reach

---

**Algorithm 2:** FISTA<sub>J</sub>: Joint estimation (Pb. (10))
 

---

**Initialization:** Set  $\mathbf{u}^{[0]} \in \mathbb{R}^{2 \times |\Upsilon|}$ ,  $\bar{\mathbf{u}}^{[0]} = \mathbf{u}^{[0]}$ ;  
 Set  $\boldsymbol{\ell}^{[0]} \in \mathbb{R}^{2 \times |\Upsilon|}$ ,  $\bar{\boldsymbol{\ell}}^{[0]} = \boldsymbol{\ell}^{[0]}$ ;  
 Let  $(\mathcal{S}_{\underline{n}}, \mathcal{T}_{\underline{n}})$  defined in (4);  
 Let  $\mathbf{J}$  defined in (3);  
 Set  $(\forall \underline{n}) \begin{pmatrix} v_{\underline{n}}^{[0]} \\ h_{\underline{n}}^{[0]} \end{pmatrix}^\top = \mathbf{J}^{-1} (\mathcal{S}_{\underline{n}}, \mathcal{T}_{\underline{n}})^\top$ ;  
 Set  $b > 2$  and  $\tau_0 = 1$ ;  
 Set  $\alpha > 0$  and  $\lambda > 0$ ;  
 Set  $\gamma > 0$  such that  $\gamma \|\mathbf{J}^{-1}\| \|\mathbf{D}\|^2 < 1$ ;

**for**  $t \in \mathbb{N}$  **do**

*Dual variable update:*

$$\begin{aligned} \mathbf{u}^{[t+1]} &= \text{prox}_{\gamma(\lambda\|\cdot\|_{2,1})^*} (\bar{\mathbf{u}}^{[t]} + \gamma \mathbf{D} \mathbf{v}^{[t]}) \\ \boldsymbol{\ell}^{[t+1]} &= \text{prox}_{\gamma(\lambda\alpha\|\cdot\|_{2,1})^*} (\bar{\boldsymbol{\ell}}^{[t]} + \gamma \mathbf{D} \mathbf{h}^{[t]}) \end{aligned}$$

*FISTA parameter update*

$$\tau_{t+1} = \frac{t+b}{b}$$

*Auxiliary variable update*

$$\begin{aligned} \bar{\mathbf{u}}^{[t+1]} &= \mathbf{u}^{[t+1]} + \frac{\tau_t - 1}{\tau_{t+1}} (\mathbf{u}^{[t+1]} - \mathbf{u}^{[t]}) \\ \bar{\boldsymbol{\ell}}^{[t+1]} &= \boldsymbol{\ell}^{[t+1]} + \frac{\tau_t - 1}{\tau_{t+1}} (\boldsymbol{\ell}^{[t+1]} - \boldsymbol{\ell}^{[t]}) \end{aligned}$$

*Primal variable update*

$$\begin{pmatrix} \mathbf{v}^{[t+1]} \\ \mathbf{h}^{[t+1]} \end{pmatrix} = \begin{pmatrix} \mathbf{v}^{[t]} \\ \mathbf{h}^{[t]} \end{pmatrix} - \mathbf{J}^{-1} \begin{pmatrix} \mathbf{D}^* (\mathbf{u}^{[t+1]} - \mathbf{u}^{[t]}) \\ \mathbf{D}^* (\boldsymbol{\ell}^{[t+1]} - \boldsymbol{\ell}^{[t]}) \end{pmatrix}$$


---

---

**Algorithm 3:** PD<sub>J</sub>: Joint estimation (Pb. (10))

---

**Initialization:** Set  $\mathbf{v}^{[0]} \in \mathbb{R}^{|\Upsilon|}$ ,  $\mathbf{u}^{[0]} = \mathbf{D}\mathbf{v}^{[0]}$ ,  $\bar{\mathbf{u}}^{[0]} = \mathbf{u}^{[0]}$ ;  
Set  $\mathbf{h}^{[0]} \in \mathbb{R}^{|\Upsilon|}$ ,  $\boldsymbol{\ell}^{[0]} = \mathbf{D}\mathbf{h}^{[0]}$ ,  $\bar{\boldsymbol{\ell}}^{[0]} = \boldsymbol{\ell}^{[0]}$ ;  
Set  $\alpha > 0$  and  $\lambda > 0$ ;  
Set  $(\delta_0, \nu_0)$  such that  $\delta_0\nu_0\|\mathbf{D}\|^2 < 1$ ;

**for**  $t \in \mathbb{N}^*$  **do**

*Primal variable update:*

$$\begin{pmatrix} \mathbf{v}^{[t+1]} \\ \mathbf{h}^{[t+1]} \end{pmatrix} = \text{prox}_{\delta_t \Phi} \left( \begin{pmatrix} \mathbf{v}^{[t]} \\ \mathbf{h}^{[t]} \end{pmatrix} - \delta_t \begin{pmatrix} \mathbf{D}^* \bar{\mathbf{u}}^{[t]} \\ \mathbf{D}^* \bar{\boldsymbol{\ell}}^{[t]} \end{pmatrix} \right)$$

*Dual variable update:*

$$\begin{aligned} \mathbf{u}^{[t+1]} &= \text{prox}_{\nu_t(\lambda\|\cdot\|_{2,1})^*} (\mathbf{u}^{[t]} + \nu_t \mathbf{D}\mathbf{v}^{[t]}) \\ \boldsymbol{\ell}^{[t+1]} &= \text{prox}_{\nu_t(\lambda\alpha\|\cdot\|_{2,1})^*} (\boldsymbol{\ell}^{[t]} + \nu_t \mathbf{D}\mathbf{h}^{[t]}) \end{aligned}$$

*Descent steps update:*

$$\vartheta_t = (1 + 2\mu\delta_t)^{-1/2}, \delta_{t+1} = \vartheta_t \delta_t, \nu_{t+1} = \nu_t / \vartheta_t$$

*Auxiliary variable update:*

$$\begin{pmatrix} \bar{\mathbf{u}}^{[t+1]} \\ \bar{\boldsymbol{\ell}}^{[t+1]} \end{pmatrix} = \begin{pmatrix} \mathbf{u}^{[t+1]} \\ \boldsymbol{\ell}^{[t+1]} \end{pmatrix} + \vartheta_t \left( \begin{pmatrix} \mathbf{u}^{[t+1]} \\ \boldsymbol{\ell}^{[t+1]} \end{pmatrix} - \begin{pmatrix} \mathbf{u}^{[t]} \\ \boldsymbol{\ell}^{[t]} \end{pmatrix} \right)$$


---

the full range  $0 < H < 1$  for the selfsimilarity parameter [42]. Therefore, we will make use of a self-customized construction that combines an effective implementation scheme with excellent theoretical and practical control of the local variance and regularity, while permitting direct extension to piecewise homogeneous construction.

Following e.g., [5, 15, 40], we start from an harmonizable representation of fBf  $B_{\underline{n}}$ ,  $\underline{n} \in \mathbb{R}^{|\Upsilon|}$ ,  $\Upsilon = \{1, \dots, N\}^2$ :

$$B_{\underline{n}} = \frac{\sigma}{C(H)} \int \frac{e^{-i\underline{f} \cdot \underline{n}} - 1}{\|\underline{f}\|^{H+1}} d\tilde{G}(\underline{f}), \quad (23)$$

with  $d\tilde{G}(\underline{f})$  the Fourier transform of a white Gaussian noise

$$\text{and } C(H) = \frac{\pi^{1/2} \Gamma(H + 1/2)}{2^{d/2} H \Gamma(2H) \sin(\pi H) \Gamma(H + d/2)}.$$

The covariance function of the zero mean process  $B_{\underline{n}}$  reads:

$$\mathbb{E} [B_{\underline{n}} B_{\underline{m}}] = \frac{\sigma^2}{2} (\|\underline{n}\|^{2H} + \|\underline{m}\|^{2H} - \|\underline{n} - \underline{m}\|^{2H}). \quad (24)$$

Texture  $Y_{\underline{n}}$  is constructed from increments of  $B_{\underline{n}}$  as (with  $\mathbf{e}_1, \mathbf{e}_2$ , unitary vectors in horizontal and vertical directions):

$$Y_{\underline{n}} = \frac{\sigma}{2\delta^H \sqrt{1 - 2^{H-2}}} \left( \underbrace{B_{\underline{n} + \delta \mathbf{e}_1} - B_{\underline{n}}}_{\text{horizontal increment}} + \underbrace{B_{\underline{n} + \delta \mathbf{e}_2} - B_{\underline{n}}}_{\text{vertical increment}} \right), \quad (25)$$

---

**Algorithm 4:** FISTA<sub>C</sub>: Coupled estimation (Pb. (12))
 

---

**Initialization:** Set  $\mathbf{u}^{[0]} \in \mathbb{R}^{2 \times |\Upsilon|}$ ,  $\bar{\mathbf{u}}^{[0]} = \mathbf{u}^{[0]}$ ;  
 Set  $\boldsymbol{\ell}^{[0]} \in \mathbb{R}^{2 \times |\Upsilon|}$ ,  $\bar{\boldsymbol{\ell}}^{[0]} = \boldsymbol{\ell}^{[0]}$ ;  
 Let  $(\mathcal{S}_{\underline{n}}, \mathcal{T}_{\underline{n}})$  defined in (4);  
 Let  $\mathbf{J}$  defined in (3);  
 Set  $(\forall \underline{n}) \begin{pmatrix} v_{\underline{n}}^{[0]} \\ h_{\underline{n}}^{[0]} \end{pmatrix}^\top = \mathbf{J}^{-1} (\mathcal{S}_{\underline{n}}, \mathcal{T}_{\underline{n}})^\top$ ;  
 Set  $b > 2$  and  $\tau_0 = 1$ ;  
 Set  $\alpha > 0$  and  $\lambda > 0$ ;  
 Set  $\gamma > 0$  s. t.  $\gamma \max(1, \alpha) \|\mathbf{J}^{-1}\| \|\mathbf{D}\|^2 < 1$ ;

for  $t \in \mathbb{N}$  do

*Dual variable update:*

$$\begin{pmatrix} \mathbf{u}^{[t+1]} \\ \boldsymbol{\ell}^{[t+1]} \end{pmatrix} = \text{prox}_{\gamma(\lambda \|\cdot\|_{2,1})^*} \begin{pmatrix} \bar{\mathbf{u}}^{[t]} + \gamma \mathbf{D} \mathbf{v}^{[t]} \\ \bar{\boldsymbol{\ell}}^{[t]} + \gamma \alpha \mathbf{D} \mathbf{h}^{[t]} \end{pmatrix}$$

*FISTA parameter update*

$$\tau_{t+1} = \frac{t+b}{b}$$

*Auxiliary variable update*

$$\begin{aligned} \bar{\mathbf{u}}^{[t+1]} &= \mathbf{u}^{[t+1]} + \frac{\tau_t - 1}{\tau_{t+1}} (\mathbf{u}^{[t+1]} - \mathbf{u}^{[t]}) \\ \bar{\boldsymbol{\ell}}^{[t+1]} &= \boldsymbol{\ell}^{[t+1]} + \frac{\tau_t - 1}{\tau_{t+1}} (\boldsymbol{\ell}^{[t+1]} - \boldsymbol{\ell}^{[t]}) \end{aligned}$$

*Primal variable update*

$$\begin{pmatrix} \mathbf{v}^{[t+1]} \\ \mathbf{h}^{[t+1]} \end{pmatrix} = \begin{pmatrix} \mathbf{v}^{[t]} \\ \mathbf{h}^{[t]} \end{pmatrix} - \mathbf{J}^{-1} \begin{pmatrix} \mathbf{D}^* (\mathbf{u}^{[t+1]} - \mathbf{u}^{[t]}) \\ \alpha \mathbf{D}^* (\boldsymbol{\ell}^{[t+1]} - \boldsymbol{\ell}^{[t]}) \end{pmatrix}$$


---

---

**Algorithm 5:** PD<sub>C</sub>: Coupled estimation (Pb. (12))

---

**Initialization:**

Set  $\mathbf{v}^{[0]} \in \mathbb{R}^{|\Upsilon|}$ ,  $\mathbf{u}^{[0]} = \mathbf{D}\mathbf{v}^{[0]}$ ,  $\bar{\mathbf{u}}^{[0]} = \mathbf{u}^{[0]}$ ;  
Set  $\mathbf{h}^{[0]} \in \mathbb{R}^{|\Upsilon|}$ ,  $\boldsymbol{\ell}^{[0]} = \alpha\mathbf{D}\mathbf{h}^{[0]}$ ,  $\bar{\boldsymbol{\ell}}^{[0]} = \boldsymbol{\ell}^{[0]}$ ;  
Set  $\alpha > 0$  and  $\lambda > 0$ .  
Set  $(\delta_0, \nu_0)$  such that  $\delta_0\nu_0 \max(1, \alpha)\|\mathbf{D}\|^2 < 1$ ;

**for**  $t \in \mathbb{N}^*$  **do**

Primal variable update:

$$\begin{pmatrix} \mathbf{v}^{[t+1]} \\ \mathbf{h}^{[t+1]} \end{pmatrix} = \text{prox}_{\delta_t \Phi} \left( \begin{pmatrix} \mathbf{v}^{[t]} \\ \mathbf{h}^{[t]} \end{pmatrix} - \delta_t \begin{pmatrix} \mathbf{D}^* \bar{\mathbf{u}}^{[t]} \\ \alpha \mathbf{D}^* \bar{\boldsymbol{\ell}}^{[t]} \end{pmatrix} \right)$$

Dual variable update:

$$\begin{pmatrix} \mathbf{u}^{[t+1]} \\ \boldsymbol{\ell}^{[t+1]} \end{pmatrix} = \text{prox}_{\nu_t(\lambda\|\cdot\|_{2,1})^*} \left( \begin{pmatrix} \mathbf{u}^{[t]} + \nu_t \mathbf{D}\mathbf{v}^{[t]} \\ \boldsymbol{\ell}^{[t]} + \nu_t \alpha \mathbf{D}\mathbf{h}^{[t]} \end{pmatrix} \right)$$

Descent steps update:

$$\vartheta_t = (1 + 2\mu\delta_t)^{-1/2}, \quad \delta_{t+1} = \vartheta_t \delta_t, \quad \nu_{t+1} = \nu_t / \vartheta_t$$

Auxiliary variable update:

$$\begin{pmatrix} \bar{\mathbf{u}}^{[t+1]} \\ \bar{\boldsymbol{\ell}}^{[t+1]} \end{pmatrix} = \begin{pmatrix} \mathbf{u}^{[t+1]} \\ \boldsymbol{\ell}^{[t+1]} \end{pmatrix} + \vartheta_t \left( \begin{pmatrix} \mathbf{u}^{[t+1]} \\ \boldsymbol{\ell}^{[t+1]} \end{pmatrix} - \begin{pmatrix} \mathbf{u}^{[t]} \\ \boldsymbol{\ell}^{[t]} \end{pmatrix} \right)$$


---

**Proposition 4** *The field  $Y = (Y_{\underline{n}})_{\underline{n} \in \Upsilon}$  is a zero-mean Gaussian process with variance  $\mathbb{E}[Y_{\underline{n}}^2] = \sigma^2$  and covariance:*

$$\begin{aligned} \mathbb{E}[Y_{\underline{n}+\Delta\underline{n}}Y_{\underline{n}}] &= \frac{\sigma^2\delta^{-2H}}{4-2H}(\|\Delta\underline{n} + \delta\underline{e}_1\|^{2H} \\ &+ \|\Delta\underline{n} - \delta\underline{e}_1\|^{2H} + \|\Delta\underline{n} + \delta\underline{e}_2\|^{2H} + \|\Delta\underline{n} - \delta\underline{e}_2\|^{2H} - 3\|\Delta\underline{n}\|^{2H} \\ &- \frac{1}{2}\|\Delta\underline{n} + \delta\underline{e}_1 - \delta\underline{e}_2\|^{2H} - \frac{1}{2}\|\Delta\underline{n} - \delta\underline{e}_1 + \delta\underline{e}_2\|^{2H}). \end{aligned}$$

**Piecewise homogeneous fractal textures.** Piecewise homogeneous fractal textures  $X = (X_{\underline{n}})_{\underline{n} \in \Upsilon}$  are defined as the concatenation of  $M$  (pairwise disjoint) regions, denoted  $(\Upsilon_m)_{0 \leq m \leq M-1}$ , such that  $\Upsilon = \cup_m \Upsilon_m$ , with textures in each region consisting of an homogeneous fractal  $Y_{\underline{n}}^{(m)}$  ( $0 \leq m \leq M-1$ ), with variance and regularity  $(\sigma_m^2, H_m)$ :

$$X_{\underline{n}} = Y_{\underline{n}}^{(m)}, \text{ when } \underline{n} \in \Upsilon_m.$$

Examples of piecewise fractal textures are shown in Fig. 3.

## 5 Estimation/Segmentation performance

### 5.1 Monte-Carlo simulation set-up

The proposed *joint* and *coupled* segmentations are now compared in terms of performance and computational costs to the state-of-the-art T-ROF procedure,

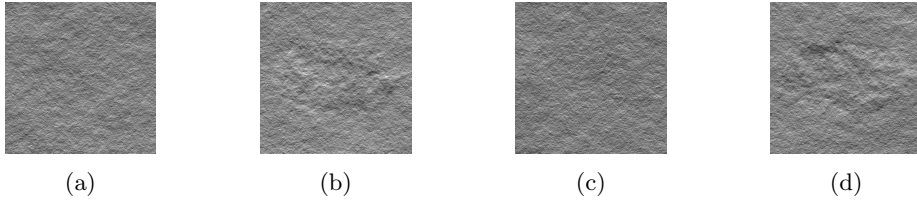


Figure 3: **Piecewise homogeneous fractal textures**  $X$  generated using the mask displayed in Fig. 1a, with parameters  $(H_0, \sigma_0) = (0.5, 0.6)$  and  $(H_1, \sigma_1)$  chosen as: (a)  $(0.5, 0.6)$ , no change ; (b)  $(0.5, 0.65)$ , change in  $\sigma$  ; (c)  $(0.8, 0.6)$  change in  $H$  ; (d)  $(0.8, 0.65)$ , change in  $\sigma$  and  $H$ .

in the context of a two-region segmentation, by means of Monte-Carlo simulations.

### 5.1.1 Synthetic textures

Piecewise homogeneous fractal textures are generated as defined in Section 4, using the mask shown in Fig. 1a, with  $N = 512$ . They consist of two regions: i) a background, with variance and local regularity  $(\sigma_0^2, H_0) = (0.6, 0.5)$  kept fixed, ii) a central ellipse, for which variance and local regularity  $(\sigma_1^2, H_1)$  are varied, as illustrated in Fig. 6. For each configuration, 5 realizations of homogeneous fractal are generated and analyzed ; performance are reported as averages over realizations.

### 5.1.2 Wavelet transform

2D-Wavelet decompositions are performed using tensor product wavelets constructed from 1D-Daubechies orthonormal least asymmetric wavelets with  $N_\psi = 3$  vanishing moments [25]. Wavelet leaders are computed as in Section 2.2. Local estimate  $\hat{h}_{LR}$  in T-ROF is computed as in Sec. 2.3. Estimates from all procedures involve octaves  $(j_1, j_2) = (2, 5)$ . Octave  $j_1 = 1$  is a priori excluded as leaders are biased [44, 46].

### 5.1.3 Optimization algorithm parameters

To achieve best convergence of the optimization schemes, descent steps are chosen as large as permitted:

- Alg. 2 (FISTA<sub>J</sub>):  $\gamma = 0.99 / (\|\mathbf{J}^{-1}\| \|\mathbf{D}\|^2)$ ,
- Alg. 4 (FISTA<sub>C</sub>):  $\gamma = 0.99 / (\max(1, \alpha) \|\mathbf{J}^{-1}\| \|\mathbf{D}\|^2)$ ,
- Alg. 3 (PD<sub>J</sub>):  $\delta_0 = \nu_0 = 0.99 / \|\mathbf{D}\|$ .
- Alg. 5 (PD<sub>C</sub>):  $\delta_0 = \nu_0 = 0.99 / (\max(1, \sqrt{\alpha}) \|\mathbf{D}\|)$ ,

where  $\|\mathbf{D}\| = 2\sqrt{2}$ , with  $\mathbf{D}$  defined in Eq. (7), and where  $\|\mathbf{J}^{-1}\|$  depends on the octaves involved in estimation ( $\mathbf{J}$  defined in Eq. (3)): With  $(j_1, j_2) = (2, 5)$ ,  $\|\mathbf{J}^{-1}\| \simeq 2.88$ . For FISTA, the inertia parameter is set to  $b = 4$ .



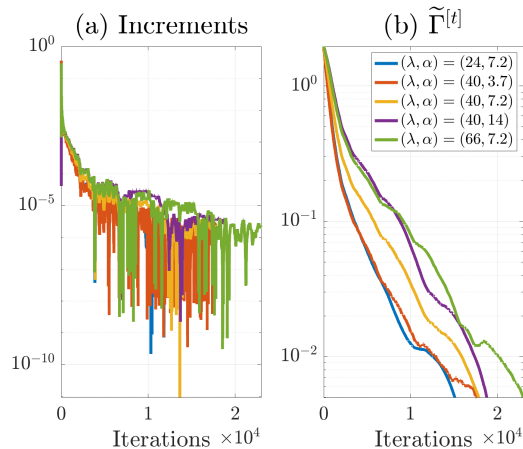


Figure 4: **Stopping criteria for proximal algorithms.** Convergence of Alg. 3 solving *joint* problem (10) for five pairs of hyperparameters  $(\lambda, \alpha)$  evaluated with: (a) (Normalized) increments of objective function, (b) Normalized primal-dual functional  $\tilde{\Gamma}_J^{[t]}$ .

## 5.2 Issues in performance and algorithm comparison

### 5.2.1 Stopping criteria for proximal algorithms

To ensure fair comparisons either between algorithms or between *joint* and *coupled* formulations (Pb. (10) vs. Pb. (12)), an effective stopping criterion is needed. Fig. 4 illustrates the convergence toward zero of two potential candidates, for the particular case, chosen as representative example, of the *joint* estimation using primal-dual Alg. 3, and for several choices of hyperparameters.

First, the normalized increments of objective function, as a function of the number of iterations, are observed (Fig. 4a) to decrease with an extremely irregular behavior, which makes them ill-suited to serve as effective stopping criterion.

Second, the *normalized* primal-dual functional (with notations as in Sec. 3.3)

$$\tilde{\Gamma}^{[t]} := \frac{\Gamma^{[t]}}{|\Theta(\mathbf{x}^{[t]}) + \Xi(\mathbf{L}\mathbf{x}^{[t]})| + |\Theta^*(-\mathbf{L}^*\mathbf{y}^{[t]}) + \Xi^*(\mathbf{y}^{[t]})|},$$

is observed (Fig. 4b) to decrease smoothly. Systematic inspections of such decreases together with that of the corresponding achieved solutions lead us to devise a stopping criterion, which is effective both for primal-dual and forward-backward algorithms and for *joint* and *coupled* estimations, as,

$$\tilde{\Gamma}_{\bullet}^{[t]} < 5 \cdot 10^{-3}, \text{ for } \bullet \in \{\text{ROF}, \text{J}\}, \text{ and } \tilde{\Gamma}_{\text{C}}^{[t]} < 10^{-4}. \quad (26)$$

The use of the normalized quantities,  $\tilde{\Gamma}^{[t]}$ , makes this stopping criterion robust to variations of the hyperparameters. For practical purposes, we also impose an upper limit on the number of iterations:  $t < 2.5 \cdot 10^5$ .

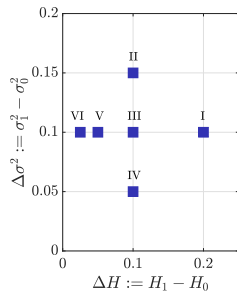


Figure 5: Explored  $\Delta\sigma^2$  and  $\Delta H$ .

		I	II	III	IV	V	VI
		$\Delta\sigma^2 = 0.1$ $\Delta H = 0.2$	$\Delta\sigma^2 = 0.15$ $\Delta H = 0.1$	$\Delta\sigma^2 = 0.1$ $\Delta H = 0.1$	$\Delta\sigma^2 = 0.05$ $\Delta H = 0.1$	$\Delta\sigma^2 = 0.1$ $\Delta H = 0.05$	$\Delta\sigma^2 = 0.1$ $\Delta H = 0.025$
T-ROF	Score	$86.7 \pm 2.1\%$	$79.5 \pm 1.2\%$	$78.5 \pm 1.1\%$	$77.5 \pm 2.9\%$	$69.9 \pm 7.1\%$	$59.5 \pm 2.4\%$
	$\widehat{\Delta H}$	$0.21 \pm 0.07$	$0.05 \pm 0.02$	$0.05 \pm 0.06$	$0.07 \pm 0.04$	$0.01 \pm 0.06$	$0.05 \pm 0.07$
T- <i>joint</i>	Score	$91.6 \pm 1.7\%$	$91.5 \pm 2.0\%$	$90.2 \pm 1.9\%$	$84.2 \pm 4.5\%$	$84.3 \pm 3.2\%$	$74.7 \pm 8.2\%$
	$\widehat{\Delta H}$	$0.21 \pm 0.06$	$0.07 \pm 0.03$	$0.10 \pm 0.02$	$0.04 \pm 0.07$	$0.05 \pm 0.02$	$0.11 \pm 0.28$
T- <i>coupled</i>	Score	$91.7\% \pm 1.7\%$	$91.9\% \pm 4.0\%$	$91.1\% \pm 1.5\%$	$85.5\% \pm 3.8\%$	$86.1\% \pm 4.3\%$	$74.3\% \pm 8.2\%$
	$\widehat{\Delta H}$	$0.20 \pm 0.05$	$0.06 \pm 0.04$	$0.10 \pm 0.02$	$0.08 \pm 0.04$	$0.05 \pm 0.02$	$0.06 \pm 0.04$

Table 1: **Optimal segmentation performance for different configurations of fractal textures, averaged over 5 realizations.** Piecewise fractal textures are characterized by  $(\sigma_0^2, H_0) = (0.6, 0.5)$  and different  $(\Delta\sigma^2, \Delta H)$  as sketched in Fig. 1a. First row: T-ROF segmentation. Second row: T-*joint* segmentation. Third row: T-*coupled* segmentation.

### 5.2.2 Choice of regularization hyperparameters

The choice of regularization parameters  $\lambda$  and  $\alpha$  appearing in Problems (5), (10), (12) is of prime importance as  $\lambda$  tunes the trade-off between fidelity to the fractal model (1) and expected piecewise constancy, while  $\alpha$  controls the relative weight given to local wavelet log-variance  $\mathbf{v}$  compared to local regularity  $\mathbf{h}$ , in the (*joint* (11) and *coupled* (13)) total variation penalization. The automated choice of the regularization parameters is a difficult issue, beyond the scope of the present work. In this study, a grid search strategy is used to find the parameters  $\lambda$  and  $\alpha$  achieving the best segmentation. In practice, logarithmically spaced ranges are used, from  $10^{-1}$  to  $10^3$  for  $\lambda$  and from  $10^{-2}$  to  $10^3$  for  $\alpha$ .

### 5.2.3 Performance assessment

A natural performance criterion consists in comparing the achieved classification, denoted  $\widehat{\mathbf{M}}_{\text{ROF}}$ ,  $\widehat{\mathbf{M}}_{\text{J}}$  and  $\widehat{\mathbf{M}}_{\text{C}}$  respectively for the three segmentation procedures compared here, (*ROF*, *joint* and *coupled*), to the mask in Fig 1a, regarded as ground truth. It leads to define the *classification score* as the percentage of correctly labeled pixels. Classification scores for  $\widehat{\mathbf{M}}_{\text{ROF}}$ ,  $\widehat{\mathbf{M}}_{\text{J}}$  and  $\widehat{\mathbf{M}}_{\text{C}}$ , applied to different configurations of piecewise fractal textures are reported in Table 1, together with the difference between the a posteriori global estimates obtained for each segmented regions  $\Upsilon_0$  and  $\Upsilon_1$  (cf. Section 2.3.3):

$$\widehat{\Delta H}_{\bullet} := \widehat{H}_{1,\bullet} - \widehat{H}_{0,\bullet}, \text{ for } \bullet \in \{\text{ROF}, \text{J}, \text{C}\}.$$

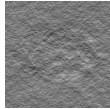

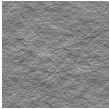
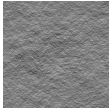
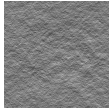
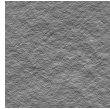


















	I	II	III	IV	V	VI
	$\Delta\sigma^2 = 0.1$	$\Delta\sigma^2 = 0.15$	$\Delta\sigma^2 = 0.1$	$\Delta\sigma^2 = 0.05$	$\Delta\sigma^2 = 0.1$	$\Delta\sigma^2 = 0.1$
	$\Delta H = 0.2$	$\Delta H = 0.1$	$\Delta H = 0.1$	$\Delta H = 0.1$	$\Delta H = 0.05$	$\Delta H = 0.025$
Texture						
T-ROF						
T-joint						
T-coupled						

Figure 6: **Compared optimal segmentation.** Piecewise fractal textures are characterized by  $\sigma_0^2 = 0.6$ ,  $H_0 = 0.5$  and different  $\Delta\sigma^2$ ,  $\Delta H$  as sketched in Fig. 1a. First row: T-ROF segmentation  $\widehat{M}_{\text{ROF}}$ . Second row: T-joint segmentation  $\widehat{M}_{\text{J}}$ . Third row: T-coupled segmentation  $\widehat{M}_{\text{C}}$ .

### 5.3 Performance comparisons

#### 5.3.1 Segmentation and estimation performance

Fig. 6 and Table 1 report segmentation and estimation performance for 7 different configurations and for the optimal set of hyperparameters (i.e., those that maximize the classification scores).

Configurations I, III, V, VI correspond to a decrease in the difference between the regularity of each region of the piecewise fractal texture:  $\Delta H = H_1 - H_0$ , (hence to an increase in difficulty) for a fixed  $\Delta\sigma^2 = 0.1$ . While the segmentation performance of the three procedures (T-ROF, T-joint and T-coupled) are comparable for *easy* configuration, those of T-ROF decrease drastically when  $\Delta H$  decreases while those of T-joint and T-coupled decrease significantly less. Along the same line, the estimation of  $\Delta H$  remains more satisfactory at small  $\Delta H$  for T-joint and T-coupled than for T-ROF. It can also be observed that the performance of T-coupled degrade slightly less than those T-joint.

Configurations II, III, IV correspond to a decrease in variance,  $\Delta\sigma^2$ , (hence to an increase in difficulty) for a fixed  $\Delta H = 0.1$  which can already be regarded as a difficult case. As expected, T-ROF is not helped by the increase of variance between IV and II as estimation of local regularity does not depend on variance [37, 46], and T-ROF segmentation results are not satisfactory. It can also be observed that performance of T-joint and T-coupled improve when  $\Delta\sigma^2$  increase, and again that the improvement is slightly larger for T-coupled than for T-joint.

These results permit to draw two clear conclusions. First, there are quantifiable benefits in using the side information brought by  $\Delta\sigma^2$ , notably when the

		Configuration I			Configuration III		
		T-ROF	T-joint	T-coupled	T-ROF	T-joint	T-coupled
Iterations ( $10^3$ it.)	DFB	$96 \pm 48$	$> 250$	$> 250$	$241 \pm 18$	$> 250$	$> 250$
	FISTA	$1.7 \pm 0.4$	$50.2 \pm 21.0$	$231 \pm 37$	<b><math>3.7 \pm 0.7</math></b>	$48.1 \pm 3.4$	$> 250$
	PD	$31.8 \pm 17.0$	$> 250$	$> 250$	$201 \pm 69$	$> 250$	$> 250$
	AcPD	<b><math>1.5 \pm 0.4</math></b>	<b><math>31.4 \pm 4.6</math></b>	<b><math>125 \pm 67</math></b>	$45.2 \pm 43$	<b><math>40.5 \pm 2.8</math></b>	<b><math>121 \pm 42</math></b>
Time (s)	DFB	$1,090 \pm 520$	$4,840 \pm 15$	$4,210 \pm 76$	$2,010 \pm 73$	$4,810 \pm 215$	$4,200 \pm 76$
	FISTA	$16 \pm 4$	$1,030 \pm 410$	$4,800 \pm 560$	<b><math>30 \pm 5</math></b>	$989 \pm 64$	$5,110 \pm 340$
	PD	$297 \pm 150$	$4,180 \pm 69$	$4,110 \pm 43$	$1,580 \pm 490$	$4,150 \pm 18$	$4,100 \pm 15$
	AcPD	<b><math>15 \pm 4</math></b>	<b><math>619 \pm 96</math></b>	<b><math>2,420 \pm 1,300</math></b>	$349 \pm 330$	<b><math>785 \pm 59</math></b>	<b><math>2,320 \pm 790</math></b>

Table 2: Number of iterations and computational time necessary to reach Condition (26) for the different proximal algorithms investigated, illustrated on two configurations I ( $\Delta H = 0.2$ ,  $\Delta\sigma^2 = 0.1$ ) and III ( $\Delta H = 0.1$ ,  $\Delta\sigma^2 = 0.1$ ). **DFB**: Dual Forward-Backward, **FISTA**: inertial acceleration of DFB, **PD**: primal-dual, **AcPD**: strong-convexity based acceleration of PD.

changes in regularity become small (low  $\Delta H$ ): T-joint and T-coupled outperform T-ROF. Second, T-coupled, that, by principle, favor co-localized changes in regularity and variance, shows overall better performance than T-joint, that does not favor co-localized changes. This is a satisfactory outcome as all the configurations chosen follow the a priori intuition, relevant for real-world applications, that changes of textures naturally imply co-localized changes in local variance and local regularity.

### 5.3.2 Computational costs

Comparisons in terms of computational costs both between the three approaches, and between the two classes of proximal algorithms, dual forward-backward, standard and accelerated (FISTA), vs. primal-dual, standard and accelerated by strong convexity (cf. Sec. 3), are reported in Table 2, for configurations I and III (regarded as *easy* and *difficult*) considered as representative. Computational costs are reported in number of iterations actually used to reach the stopping criterion and in real time, for the optimal set of hyperparameters and averaged over 5 realizations.

Table 2 shows first that, as expected, accelerated algorithms are always require less iterations than non accelerated ones, thus generally leading to lower computational times (though this is not always the case with FISTA whose complexity per iteration is larger). Also, T-ROF shows always lower computational costs compared to T-joint and T-coupled. This is expected as T-ROF only works with the regularity and do not use variance.

*FISTA vs. Accelerated primal-dual.* For T-ROF, FISTA is overall preferable to the accelerated primal-dual algorithm, as both show equivalent computational costs for Configuration I but FISTA is ten times faster (both in number of iterations and computation time) for Configuration III. For T-joint and T-coupled, for both configurations, accelerated primal-dual is faster than FISTA. For T-coupled, in configuration III, FISTA has actually not converged when meeting the the upper limit of iterations. Therefore, FISTA is to be preferred for T-ROF, while accelerated primal-dual algorithms are more relevant for T-joint and T-coupled.

*T-joint vs. T-coupled.* Focusing on T-joint and T-coupled and thus on the

accelerated primal-dual algorithm that is faster for these two methods, Table 2 shows that *T-joint* is solved 3 to 4 times faster (both in number of iterations and computational cost) than *T-coupled*.

### 5.3.3 Overall comparison

As an overall conclusion, results reported above show that there are benefits to use together local regularity and variance, compared to regularity only, when the changes in regularity become small. This implies switching from accelerated dual forward-backward algorithms (FISTA, for T-ROF) to accelerated primal-dual algorithms (for *T-joint* and *T-coupled*).

For difficult configurations, *T-coupled* (slightly) outperforms *T-joint* in terms of segmentation performance, at the price of non negligible increases of computational costs. In that sense, *T-joint* can be considered a reasonable trade-off between too poor segmentation performance (as those of T-ROF) and too large computational costs (as those of *T-coupled*).

## 6 Conclusion and future work

The present article has significantly advanced the state-of-the-art in the segmentation of piecewise fractal textures.

First, it has been proposed to base the segmentation of fractal textures not only on the estimation of the sole local regularity parameter, but to use an additional local parameter, the log-wavelet variance, tightly related to the local variance of the textures. Two variations were investigated, *coupled* and *joint*, that respectively enforce or not co-localized changes in regularity and variance. It has been shown, using large size Monte Carlo simulations, that the use of this additional features improves drastically segmentation performance when the difference in regularity becomes negligible. This yet comes at the price of a non negligible increase in computational costs.

Therefore, a second contribution has been to construct accelerated primal-dual algorithms, requiring the explicit calculation of the strong convexity constant underlying the data fidelity term form. The achieved substantial reduction in computational costs has turned critical both to be able to conduct large size Monte Carlo simulation and to perform the greedy search of the optimal set of hyper-parameters. This low computational cost is also crucial for application on real-world data.

The investigations reported here have permitted to show that accelerated primal-dual algorithms outperform accelerated dual forward-backward (FISTA-type) algorithms for piecewise fractal texture segmentation as soon as the joint use of regularity and variance is required. Further, they showed that the *coupled* formulation, that favor co-localized changes in regularity and variance, performs better than the *joint* formulation, yet at the price of a significantly larger computational cost. Thus, depending on budget constraints on time and requested quality of the solution, the *joint* formulation can be regarded as an effective trade-off.

The proposed theoretical formulations for piecewise fractal texture segmentation and the corresponding accelerated algorithms are matured enough for applications on real-world data. Application to the segmentation of multipha-

sic flows is under current investigations. The automation of the tuning of the hyperparameters is also been investigated. Extensions to piecewise multifractal textures are also targeted.

A MATLAB toolbox implementing the analysis and synthesis procedures devised here will be made freely and publicly available at the time of publication.

## References

- [1] P. Abry, S. Jaffard, and H. Wendt. When Van Gogh meets Mandelbrot: Multifractal classification of painting's texture. *Signal Process.*, 93(3):554–572, 2013.
- [2] P. Abry and F. Sellan. The wavelet-based synthesis for fractional brownian motion proposed by F. Sellan and Y. Meyer: Remarks and fast implementation, 1996.
- [3] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE Trans. Pattern Anal. Match. Int.*, 33(5):898–916, 2011.
- [4] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.
- [5] H. Biermé, M. Meerschaert, and H. Scheffle. Operator scaling stable random fields. *Stochastic Processes and their Applications*, 117(3):312–332, 2007.
- [6] X. Cai, R. Chan, C.-B. Schonlieb, and T. Steidl, G.and Zeng. Linkage between piecewise constant mumford-shah model and rof model and its virtue in image segmentation. *arXiv preprint arXiv:1807.10194*, 2018.
- [7] X. Cai and G. Steidl. Multiclass segmentation by iterated rof thresholding. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 237–250. Springer, 2013.
- [8] A. Chambolle and C. Dossal. On the convergence of the iterates of "FISTA". *J. Optim. Theory Appl.*, 166(3):25, August 2015.
- [9] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imag. Vis.*, 40(1):120–145, 2011.
- [10] B. B. Chaudhuri and N. Sarkar. Texture segmentation using fractal dimension. *IEEE Trans. Pattern Anal. Match. Int.*, 17(1):72–77, 1995.
- [11] C. Chaux, P.L. Combettes, and V.R. Pesquet, J.-C. Wajs. A variational formulation for frame-based inverse problems. *Inverse Problems*, 23(4):1495–1518, Jun. 2007.
- [12] S. Cohen and J. Istas. *Fractional fields and applications*. Springer, 2013.

- [13] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, editors, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pages 185–212. Springer-Verlag, New York, 2011.
- [14] L. Condat. A primal-dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms. *J. Optim. Theory Appl.*, 158(2):460–479, 2013.
- [15] G. Didier, M. M. Meerschaert, and V. Pipiras. Domain and range symmetries of operator fractional brownian fields. *Stochastic Processes and their Applications*, 128(1):39–78, 2018.
- [16] D. Dunn, W. E. Higgins, and J. Wakeley. Texture segmentation using 2-d gabor elementary functions. *IEEE Trans. Pattern Anal. Match. Int.*, 16(2):130–149, 1994.
- [17] P. Getreuer. Chan-veese segmentation. *Image Processing On Line*, 2:214–224, 2012.
- [18] M. A. Ibrahim, O. A. Ojo, and P. A. Oluwafisoye. Identification of emphysema patterns in high resolution computed tomography images. *Journal of Biomedical Engineering and Informatics*, 4(1):16, 2017.
- [19] S. Jaffard. Wavelet techniques in multifractal analysis. *Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot*, M. Lapidus and M. van Frankenhuijsen Eds., *Proceedings of Symposia in Pure Mathematics*, 72(2):91–152, 2004.
- [20] A. K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern Recogn.*, 24(12):1167–1186, 1991.
- [21] M. Jung, G. Peyré, and L. D. Cohen. Non-local active contours. In *Proc. SSVM*, volume 6667, pages 255–266, Ein-Gedi, Israel, 2011. Springer.
- [22] I. Kokkinos, G. Evangelopoulos, and P. Maragos. Texture analysis and segmentation using modulation features, generative models, and weighted curve evolution. *IEEE Trans. Pattern Anal. Match. Int.*, 31(1):142–157, 2009.
- [23] N. Komodakis and J.-C. Pesquet. Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems. *IEEE Signal Process. Mag.*, 32(6):31–54, Nov. 2015.
- [24] R. Leonarduzzi, H. Wendt, P. Abry, S. Jaffard, C. Melot, S. G. Roux, and M. E. Torres. p-exponent and p-leaders, Part II: Multifractal analysis. relations to detrended fluctuation analysis. *Physica A*, 448:319–339, 2016.
- [25] S. Mallat. *A wavelet tour of signal processing*. Academic Press, San Diego, USA, 1997.
- [26] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Trans. Pattern Anal. Match. Int.*, 14(7):710–732, 1992.

- [27] Z. Marin, K. A. Batchelder, B. C. Toner, L. Guimond, E. Gerasimova-Chechkina, A. R. Harrow, A. Arneodo, and A. Khalil. Mammographic evidence of microenvironment changes in tumorous breasts. *Medical physics*, 44(4):1324–1336, 2017.
- [28] M. T. McCann, D. G. Mixon, M. C. Fickus, C. A. Castro, J. A. Ozolek, and J. Kovacevic. Images as occlusions of textures: A framework for segmentation. *IEEE Trans. Image Process.*, 23(5):2033–2046, 2014.
- [29] Majid Mirmehdi. *Handbook of texture analysis*. Imperial College Press, 2008.
- [30] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Applied Math.*, 42(5):577–685, 1989.
- [31] J.F. Muzy, E. Bacry, and A. Arneodo. Wavelets and multifractal formalism for singular signals: Application to turbulence data. *Phys. Rev. Lett.*, 67(25):3515–3518, 1991.
- [32] J.D.B. Nelson, C. Naornita, and A. Isar. Semi-local scaling exponent estimation with box-penalty constraints and total-variation regularization. *IEEE Trans. Image Process.*, 25(7):3167–3181, 2016.
- [33] B. Pascal, N. Pustelnik, P. Abry, and J.-C. Pesquet. Block-coordinate proximal algorithms for scale-free texture segmentation. In *Proc. Int. Conf. Acoust., Speech Signal Process.*, Calgary, Alberta, Canada, Apr. 15-20 2018.
- [34] B. Pascal, N. Pustelnik, P. Abry, M. Serres, and V. Vidal. Joint estimation of local variance and local regularity for texture segmentation. application to multiphase flow characterization. In *Proc. Int. Conf. Image Process.*, pages 2092–2096, Athens, Greece, 2018. IEEE.
- [35] R. M. Pereira, C. Garban, and L. Chevillard. A dissipative random velocity field for fully developed fluid turbulence. *Journal of Fluid Mechanics*, 794:369–408, 2016.
- [36] N. Pustelnik, C. Chaux, and J.-C. Pesquet. Parallel ProXimal Algorithm for image restoration using hybrid regularization. *IEEE Trans. Image Process.*, 20(9):2450–2462, 2011.
- [37] N. Pustelnik, H. Wendt, P. Abry, and N. Dobigeon. Combining local regularity estimation and total variation optimization for scale-free texture segmentation. *IEEE Trans. Computational Imaging*, 2(4):468–479, 2016.
- [38] S. Ray and R. H. Turi. Determination of number of clusters in K-means clustering and application in colour image segmentation. In *Proc. Int. Conf. Adv. Pattern Recognition and Digital Techniques*, pages 137–143. Calcutta, India, 1999.
- [39] A. Repetti, E. Chouzenoux, and J.-C. Pesquet. A parallel block-coordinate approach for primal-dual splitting with arbitrary random block selection. In *Proc. Eur. Sig. Proc. Conference*, pages 235–239, Nice, France, Aug. 31-Sept. 4, 2015.



- [40] S.G. Roux, M. Clausel, B. Vedel, S. Jaffard, and P. Abry. Self-similar anisotropic texture analysis: the hyperbolic wavelet transform contribution. *IEEE Trans. Image Process.*, 22(11):4353 – 4363, 2013.
- [41] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, 1992.
- [42] M. L. Stein. Fast and exact simulation of fractional brownian surfaces. *Journal of Computational and Graphical Statistics*, 11(3):587–599, 2002.
- [43] B.C. Vũ. A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Adv. Comput. Math.*, 38(3):667–681, Apr. 2013.
- [44] H. Wendt, P. Abry, and S. Jaffard. Bootstrap for empirical multifractal analysis. *IEEE Signal Process. Mag.*, 24(4):38–48, 2007.
- [45] H. Wendt, S. Combexelle, Y. Altmann, J.-Y. Tournet, S. McLaughlin, and P. Abry. Multifractal analysis of multivariate images using gamma markov random field priors. *SIAM J. Imaging Sci.*, 11(2):1294–1316, 2018.
- [46] H. Wendt, S. G. Roux, P. Abry, and S. Jaffard. Wavelet leaders and bootstrap for multifractal analysis of images. *Signal Process.*, 89(6):1100–1114, 2009.
- [47] J. Yuan, D. Wang, and A. M. Cheriadat. Factorization-based texture segmentation. *IEEE Trans. Image Process.*, 24(11):3488–3497, November 2015.