A Variational Model for Multiplicative Structured Noise Removal

Paul Escande · Pierre Weiss · Wenxing Zhang

Abstract We consider the problem of restoring images impaired by noise that is simultaneaously *structured* and *multiplicative*. Our primary motivation for this setting is the Selective Plane Illumination Microscope (SPIM) which often suffers from severe inhomogeneities due to light absorption and scattering. This type of degradation arises in other imaging devices such as ultrasonic imaging. We model the multiplicative noise as a stationary process with known distribution. This leads to a novel convex image restoration model based on a maximum a posteriori estimator. After establishing some analytical properties of the minimizers, we finally propose a fast optimization method on GPU. Numerical experiments on 2D fluorescence microscopy images demonstrate the usefulness of the proposed models in practical applications.

Keywords Denoising, Multiplicative noise, Stationary noise, Variational method, Convex optimization

AMS Classification 65K10, 68U10, 94A08, 49M29, 92C55

P. Escande
DISC, ISAE, Université de Toulouse, France, E-mail: paul.escande@gmail.com. This author is pursuing a Ph.D. degree supported by the MODIM project funded by the PRES of Toulouse University and Midi-Pyrénées region.

P. Weiss
ITAV and IMT, Université de Toulouse, CNRS, 31106 Toulouse, France, E-mail: pierre.weiss@itav-recherche.fr. Supported by the OPTIMUS Project (fondation RITC, France).

W. Zhang
School of Mathematical Sciences, University of Electronic Science and Technology of China, 611731 Chengdu, China, E-mail: zhangwx@uestc.edu.cn. Support by the NNSFC Grant 11301055

1 Introduction

Noise reduction is a long-standing problem which plays a fundamental role to simplify image analysis. Denoising algorithms should be designed by accounting carefully for the noise properties. Many works focus on additive white noise (i.e. noise that is independent pixelwise). This problem is arguably close to have reached a mature state, at least in the case of white Gaussian noise [22]. In this paper, we consider a setting where the noise is made of spatially correlated patterns that affect the image multiplicatively. To the best of our knowledge, this problem is left unexplored until now.

Our initial motivation is an instance of fluorescence microscopy called Selective Plane Illumination Microscopy (SPIM) [19]. In this modality, samples are excited by a light sheet that coincides with the focal plane of a camera. While passing through the sample, the light sheet is diffracted and absorbed. As a result, images suffer from stripes parallel to the direction of light propagation. A typical image obtained with a SPIM is displayed in Figure 1. As can be seen, large stripes make automatic image analysis difficult. The proposed ideas, though initially developed for the SPIM, are likely to be useful for other imaging devices such as confocal microscopes, ultrasound imaging (speckle noise) [26], hyperspectral remote sensing imaging (waterfall effects) [16, 19] or nanotomography and synchrotron based X-ray microtomography ("ripple" effect) [4, 17, 21, 24, 25, 33].

1.1 Related works

In this section, we provide a brief and incomplete review of existing methods to treat multiplicative noise and additive structured noise.
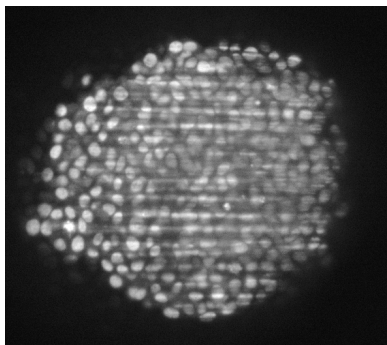
Fig. 1: SPIM acquisition of a HCT116 tumor spheroid. Large stripes strongly impair image quality.

1.1.1 Multiplicative noise

Probably the simplest way to treat multiplicative noise simply consists of applying a logarithm to the image and then use standard techniques for additive white noise. This however introduces a bias that needs to be corrected with dedicated methods [12]. The first variational methods that attempted to take the multiplicative nature of noise into account more finely are probably due to Rudin, Lions, Osher in [28] and later by Shi and Osher [30]. These methods were however not based on a clear statistical framework. Roughly at the same time, Aubert and Aujol [3] proposed a nonconvex variational model based on a Maximum A Posteriori (MAP) estimator. The nonconvexity usually makes it impossible to find global minimizers. This motivated Teuber and Steidl [31] to develop convex variational methods based on I-divergences. More recently, some authors extended the ideas from patch-based and dictionary learning restoration to the multiplicative case [10, 18], with significant improvements. It is however unclear how to apply patch-based restoration with correlated noise. Finally, let us mention that a few authors also studied deconvolution problems with multiplicative noise [11, 28, 35]. These models are closely related to the proposed approach, since deconvolution can be seen as a denoising problem with structured noise after a direct deconvolution. The main difference is that the noise structure is completely different in the proposed approach, since its Fourier transform does not blow up in high frequencies.

1.1.2 Structured noise

In recent years, treating structured noise gained importance in fields such as life sciences, astronomy or seismology. Many works are dedicated to a particular imaging modality, since noise structure is highly dependent on the physical acquisition principle. To name a few, Münch et al [24] devised a wavelet-Fourier based filter for removing stripe and ring artifacts in ion beam nanotomography and computed tomography (CT) X-Ray microtomography. Aizenberg and Butakoff [1] proposed to use median filters in the Fourier domain to reduce quasi-periodic noise. Boas and Fleischmann [4] reviewed variant structured artifacts in computed tomography (CT) and developed an iterative reconstruction approach. Cornelis et al [9] designed double filters (a smoothing Wiener filter and an adaptive filter) approach for digitally removing periodic canvas patterns superimposed on painting artworks. Chen and Pellequer [8] developed a "divide-and-conquer" approach, where the Fourier spectrum of the image is divided into central and off-center regions for noisy pixels detection and intensity restoration, to remove heavy and fine stripes in atomic force microscopy bio-molecular images. Anas et al [2] provided a detailed statistical analysis including classification, detection and correction of ring artifacts in flat panel detectors based micro-CT images. Kim et al [20] suggested to reduce ring artifacts in sinogram data by calculating the line-ratio and equalizing detector element in sinogram space. Chang et al [7] developed a variational destriping model by combining the unidirectional total variation and framelet regularization. Fitschen et al [15] proposed similar ideas to remove the curtaining effect and stripes in Ion Beam nanotomography images. Finally, Sur and Grédicac [32] proposed an automatized algorithm based on natural image statistics to remove quasi-periodic noise from strain images.

To finish, let us mention previous works by one of the authors [13, 14] (which played a role in the emergence of subsequent methods such as [7,15,32]). Therein, a generic framework was proposed to treat additive structured noise within a variational framework. The noise was modeled as the convolution of a given filter with a random vector with i.i.d. components. The motivation behind this modeling is that in many applications, the noise is stationary: its distribution should be invariant to translations. The MAP principle then leads to a convex restoration model solved with efficient first order methods. The proposed methodology was proved to be very useful in different image modalities (SPIM, nanotomography, atomic force microscopy, bathymetry, satellite imaging...). It is used on a daily basis in the imaging facility of our laboratory. However, it leads to some artifacts and the restoration is imperfect when images suffer from with very large or dark stripes.

1.2 Our contribution

This work is built upon the previous paper [14]. The first contribution is a phenomenological model to de-

scribe the random patterns that appear in Figure 1. The proposed statistical model reproduces the degradations observed in practice rather faithfully. Its second important feature, is that the maximum a posteriori principle leads to convex restoration models. Convexity allows us determining some analytical properties of the minimizers and to design an efficient minimization procedure. We finally apply the proposed algorithm to synthetic and real images, to demonstrate the interest of the proposed approach in biological imaging. The proposed algorithm is implemented on GPU, interfaced with Fiji [29] and distributed freely on our webpage.

1.3 Paper structure

We first provide some background information in section 2. In section 3, we design a first image formation model. The MAP leads to a convex variational problem that contains four parameters. In section 4, we analyze the proposed model and show that it can be simplified and contain only one regularization parameter. We then provide some analytical properties of the simplified model minimizers and design a minimization algorithm. Finally, we propose some numerical results on synthetic and real images in section 6.

2 Preliminaries

2.1 Notation

Let E denote the vector space of images defined on $\Omega = \{1, \ldots, n_1\} \times \ldots \times \{1, \ldots, n_d\}$. The total number of pixels is therefore $n = n_1 \cdot \ldots \cdot n_d$. Let E_+ denote the space of images on Ω with non-negative entries. The pixels of the image are identified by a multi-index $\mathbf{i} = (i_1, \ldots, i_d) \in \Omega$. For any image $\mathbf{u} \in E$ and scalar $p \in [1, +\infty)$, we let $\|\mathbf{u}\|_p := \left(\sum_{\mathbf{i} \in \Omega} |\mathbf{u}[\mathbf{i}]|^p \right)^{1/p}$ and $\|\mathbf{u}\|_\infty := \max_{\mathbf{i} \in \Omega} |\mathbf{u}[\mathbf{i}]|$ denote the standard l^p and l^∞-norm, respectively. The dot product of $\mathbf{u}_1$ and $\mathbf{u}_2 \in E$ is defined by $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle := \sum_{\mathbf{i} \in \Omega} \mathbf{u}_1[\mathbf{i}]\mathbf{u}_2[\mathbf{i}]$. The symbol $\mathbf{1}$ stands for the element of E with all values equal to 1. The identity operator is denoted $\mathbf{I}$. The notation diag($\mathbf{u}$) indicates a diagonal operator with diagonal elements equal to the entries of $\mathbf{u}$. Let $V = E^d$ denote the space of discrete vector fields on Ω. For any $\mathbf{q} = (\mathbf{q}_1, \ldots, \mathbf{q}_d) \in V$, $|\mathbf{q}|$ denotes an element of E with $\mathbf{i}$-th entry defined by

$$|\mathbf{q}|[\mathbf{i}] := \sqrt{\sum_{l=1}^d \mathbf{q}_l[\mathbf{i}]^2}. \tag{1}$$

2.2 Convex sets and functions

For simplicity, we identify E with its dual E^*, the space of linear forms on E. Let $f : E \to (-\infty, +\infty]$ be a function. The domain of f is defined by dom$f :=$ $\{\mathbf{x} \in E \mid f(\mathbf{x}) < +\infty\}$. Function f is proper if it is not identically equal to $+\infty$. It is closed if its epigraph epi$f := \{(\mathbf{x}, z) \in E \times \mathbb{R} \mid f(\mathbf{x}) \leq z\}$ is closed in $E \times \mathbb{R}$. The indicator of a set $D \subset E$ is defined by

$$\chi_D(\mathbf{x}) = \begin{cases} 0 & \text{if } \mathbf{x} \in D, \\ +\infty & \text{otherwise.} \end{cases} \tag{2}$$

The proximity operator of the proper convex function f is defined by

$$\text{prox}_f(\mathbf{x}) := \arg\min_{\mathbf{y} \in E} \left\{ f(\mathbf{y}) + \tfrac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 \right\}, \quad \forall \mathbf{x} \in E. \tag{3}$$

The conjugate of f is the function $f^* : E \to (-\infty, +\infty]$ defined by

$$f^*(\mathbf{y}) := \sup_{\mathbf{x} \in E} \left\{ \langle \mathbf{x}, \mathbf{y} \rangle - f(\mathbf{x}) \right\}. \tag{4}$$

Let $D \subseteq E$ be a convex and closed subset of E. The normal cone to D at point $\boldsymbol{\lambda}$ is denoted $\mathcal{N}_D(\boldsymbol{\lambda})$. It is defined for all $\boldsymbol{\lambda} \in D$ by:

$$\mathcal{N}_D(\boldsymbol{\lambda}) = \{\boldsymbol{\eta} \in E, \langle \boldsymbol{\eta}, \boldsymbol{\lambda}' - \boldsymbol{\lambda} \rangle \leq 0, \forall \boldsymbol{\lambda}' \in D\}. \tag{5}$$

Notice that since $\mathcal{N}_D(\boldsymbol{\lambda})$ is a cone:

$$t\mathcal{N}_D(\boldsymbol{\lambda}) = \mathcal{N}_D(\boldsymbol{\lambda}), \ \forall t > 0, \ \forall \boldsymbol{\lambda} \in E. \tag{6}$$

2.3 Discretization of differential operators

The discrete gradient of $\mathbf{u} \in E$ is defined by:

$$\nabla \mathbf{u} = (\partial_1 \mathbf{u}, \ldots, \partial_d \mathbf{u}) \in V. \tag{7}$$

The partial derivatives are defined as in [5] by:

$(\partial_l \mathbf{u})[\mathbf{i}] =$

$\begin{cases} \mathbf{u}[\ldots, i_l + 1, \ldots] - \mathbf{u}[\ldots, i_l, \ldots] & \text{if } i_l < n_l, \\ 0 & \text{otherwise.} \end{cases}$

The adjoint operator of ∂_l^* is the unique operator satisfying:

$$\langle \partial_l \mathbf{u}, \mathbf{q}_l \rangle = \langle \mathbf{u}, \partial_l^* \mathbf{q}_l \rangle, \ \forall \mathbf{u}, \mathbf{q}_l \in E. \tag{8}$$

It is easy to establish that:

$(\partial_l^* \mathbf{q}_l)[\mathbf{i}] =$

$\begin{cases} \mathbf{u}[\ldots, i_l, \ldots] - \mathbf{u}[\ldots, i_l - 1, \ldots] & \text{if } 1 < i_l < n_l, \\ \mathbf{u}[\ldots, i_l, \ldots] & \text{if } i_l = 1, \\ -\mathbf{u}[\ldots, i_l - 1, \ldots] & \text{if } i_l = n_l. \end{cases}$

The adjoint of the gradient operator is defined by:

$$\nabla^* : V \to E$$
$$\mathbf{q} \mapsto \sum_{l=1}^{d} \partial_l^* \mathbf{q}_l. \qquad (9)$$

Let $\mathbf{A} : E \to V$ denote a linear operator. Its spectral norm is defined by

$$\|\mathbf{A}\| = \sup_{\mathbf{x} \in E, \|\mathbf{x}\|_2 \leq 1} \|\mathbf{A}\mathbf{x}\|_2. \qquad (10)$$

The point-wise product between two elements $\mathbf{u}$ and $\mathbf{v}$ of E is denoted $\mathbf{u} \odot \mathbf{v}$:

$$(\mathbf{u} \odot \mathbf{v})[\mathbf{i}] = \mathbf{u}[\mathbf{i}]\mathbf{v}[\mathbf{i}]. \qquad (11)$$

The point-wise division between two elements $\mathbf{u}$ and $\mathbf{v}$ of E is denoted $\mathbf{u} \oslash \mathbf{v}$:

$$(\mathbf{u} \oslash \mathbf{v})[\mathbf{i}] = \mathbf{u}[\mathbf{i}]/\mathbf{v}[\mathbf{i}]. \qquad (12)$$

The convolution product $\mathbf{w} = \mathbf{u} \star \mathbf{v}$ between $\mathbf{u}$ and $\mathbf{v}$ is defined by:

$$\mathbf{w}[\mathbf{i}] = \sum_{\mathbf{j} \in \Omega} \mathbf{u}[\mathbf{i} - \mathbf{j}]\mathbf{v}[\mathbf{j}]. \qquad (13)$$

In this definition, we assumed periodic boundary conditions. The discrete Fourier transform operator on E is denoted $\mathbf{F}$. It satisfies the fundamental relationship:

$$\mathbf{F}(\mathbf{u} \star \mathbf{v}) = \mathbf{F}(\mathbf{u}) \odot \mathbf{F}(\mathbf{v}). \qquad (14)$$

The discrete Dirac mass is denoted $\boldsymbol{\delta}$. It satisfies $\mathbf{u} \star \boldsymbol{\delta} = \mathbf{u}$ for all $\mathbf{u} \in E$.

2.4 The (inverse)-gamma distribution

A random variable X that is gamma-distributed is denoted by $X \sim \mathrm{Gamma}(a, b)$. Its probability density function (p.d.f.) is defined by:

$$\mathbb{P}(x) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx), \quad \forall x > 0,$$

where $\Gamma(\cdot)$ is the gamma function. The parameter $a > 0$ is called shape parameter, while $b > 0$ is called inverse scale parameter. If $X \sim \mathrm{Gamma}(a, b)$, then $\mathbb{E}(X) = a/b$ and $\mathrm{var}(X) = a/b^2$. For any $c > 0$, if $X \sim \mathrm{Gamma}(a, b)$, then $cX \sim \mathrm{Gamma}(a, cb)$. Let

$$X_i \sim \mathrm{Gamma}(a_i, b), \quad i \in \{1, \dots, n\} \qquad (15)$$

be independent random variables, then

$$\sum_{i=1}^{n} X_i \sim \mathrm{Gamma}\left(\sum_{i=1}^{n} a_i, b\right). \qquad (16)$$

A random variable X that is inverse-gamma distributed is denoted by $X \sim \mathrm{InvGamma}(a, b)$. Its p.d.f. is:

$$\mathbb{P}(x) = \frac{b^a}{\Gamma(a)} x^{-a-1} \exp(-b/x), \quad \forall x > 0.$$

If $X \sim \mathrm{Gamma}(a, b)$ then $1/X \sim \mathrm{InvGamma}(a, 1/b)$. Moreover, $\mathbb{E}(X) = \frac{b}{a-1}$ for $a > 1$ and $\mathrm{var}(X) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ for $\alpha > 2$.

3 Noise model

3.1 Modeling the noise

Designing a precise image formation model in the examples given in Figure 1 is a hard task. Many different physical phenomena such as diffraction, absorption and scattering with unknown physical quantities are involved. It is in fact likely that good forward models of light propagation are too complicated to use in the frame of inverse problems. This observation leads us to develop a simple phenomenological model. From a practical point of view, a good model should:

i) Reproduce approximately what is observed in real experiments.
ii) Lead to optimization problems that can be solved with a reasonable complexity. In this paper, we will be particularly interested in designing *convex problems*.

Let $\mathbf{u}$ denote the ideal image we wish to recover and $\mathbf{u}_0$ denote the degraded image. The first observation is that the noise is *multiplicative* since it is due to attenuation of the excitation light through the sample: the light emitted by a fluorescent sample is approximately proportional to the excitation intensity. The standard modeling of such noise consists of writing $\mathbf{u}_0 = \mathbf{u} \odot \boldsymbol{\xi}$, where $\boldsymbol{\xi}$ is some random vector. This equation however leads to serious numerical troubles since the maximum a posteriori principle leads to problems of the form:

$$\min_{(\mathbf{u}, \boldsymbol{\xi}) \in E_+ \times E_+, \ \mathbf{u}_0 = \mathbf{u} \odot \boldsymbol{\xi}} f(\mathbf{u}) + g(\boldsymbol{\xi}), \qquad (17)$$

where functions f and g are priors on the image and the noise respectively. The constraint set $\{(\mathbf{u}, \boldsymbol{\xi}) \in E_+ \times E_+, \ \mathbf{u}_0 = \mathbf{u} \odot \boldsymbol{\xi}\}$ is nonconvex and - except for specific f and g - finding the global minimizer of problem (17) is therefore out of reach.

To avoid this pitfall, we instead write that:

$$\mathbf{u}_0 = \mathbf{u} \oslash \boldsymbol{\eta}, \qquad (18)$$

where $\boldsymbol{\eta}$ is the realization of some random vector. The constraint set obtained by using a division instead of

a multiplication becomes the linear subspace $\{(\mathbf{u}, \boldsymbol{\eta}) \in E_+ \times E_+, \; \boldsymbol{\eta} \odot \mathbf{u}_0 = \mathbf{u}\}$. The whole minimization problem

$$\min_{(\mathbf{u}, \boldsymbol{\eta}) \in E_+ \times E_+, \; \boldsymbol{\eta} \odot \mathbf{u}_0 = \mathbf{u}} f(\mathbf{u}) + g(\boldsymbol{\eta}), \tag{19}$$

therefore becomes convex as soon as f and g are convex.

The main difficulty is now to construct a probability distribution function for the random vector $\boldsymbol{\eta}$. Similarly to [13,14], we assume that the noise is *stationary*, meaning that all translated versions of $\boldsymbol{\eta}$ have the same likelihood as $\boldsymbol{\eta}$. A simple way to generate a stationary random vector consists of writing

$$\boldsymbol{\eta} = \boldsymbol{\psi} \star \boldsymbol{\lambda}, \tag{20}$$

where $\boldsymbol{\psi}$ is a convolution filter that depends on the noise structure and $\boldsymbol{\lambda}$ is the realization of a random vector in E with *independent* entries.

In order to specify our model completely, we still need to define a distribution for $\boldsymbol{\lambda}$. Since fluorescence images have nonnegative values, all components of $\boldsymbol{\eta}$ need to be positive. A simple way to ensure this is to set a convolution filter $\boldsymbol{\psi} \neq 0$ with nonnegative entries and to define $\boldsymbol{\lambda}$ as a positive random vector. By doing so, the event $\boldsymbol{\eta}[\mathbf{i}] = 0$ occurs with null probability. Among the positive, log-concave distributions, the gamma distribution has two parameters, allowing to control its mean and variance. This is important to be able varying the attenuation amplitude.

We now specified every piece of our image formation model. Let us recollect everything below.

Proposed image formation model.
We assume that:

$$\mathbf{u}_0 = \mathbf{u} \oslash (\boldsymbol{\psi} \star \boldsymbol{\lambda}), \quad \text{where:} \tag{21}$$

- $\mathbf{u}_0 \in E_+$ is the observed, noisy image.
- $\mathbf{u} \in E_+$ is the clean image we wish to recover.
- $\boldsymbol{\psi} \in E_+$ is a convolution filter.
- $\boldsymbol{\lambda} \in E_+$ is the realization of a random vector with i.i.d. components following a Gamma(a, b) distribution.

3.2 A few noise properties

3.2.1 The case $\boldsymbol{\psi} = \boldsymbol{\delta}$.

In the specific case $\boldsymbol{\psi} = \boldsymbol{\delta}$, our model is adapted to i.i.d. multiplicative noise. The generative model (21) can be rewritten as

$$\mathbf{u}_0 = \mathbf{u} \odot \boldsymbol{\varrho}, \tag{22}$$

where $\boldsymbol{\varrho}$ follows an *inverse-gamma* distribution of parameters a and b^{-1}. If we follow a Bayesian point of view, our model is therefore well suited to multiplicative inverse-gamma noise. This departs from the usual modeling used in SAR imaging [3]. However, we will see that this model also performs well in the case $\mathbf{u}_0 = \mathbf{u} \odot \boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is a white gamma noise.

3.2.2 Indicator functions

Let $\omega \subseteq \Omega$ denote a subset of cardinality $|\omega|$. Set

$$\boldsymbol{\psi}[\mathbf{i}] = \begin{cases} 1/|\omega| & \text{if} \quad \mathbf{i} \in \omega, \\ 0 & \text{otherwise.} \end{cases} \tag{23}$$

Then, a marginal $\boldsymbol{\eta}[\mathbf{i}]$ is just a sum of i.i.d. random variables with distribution Gamma$(a, b/|\omega|)$. Therefore $\boldsymbol{\eta}[\mathbf{i}] \sim \text{Gamma}(|\omega|a, b/|\omega|)$ and

$$1/\boldsymbol{\eta}[\mathbf{i}] \sim \text{InvGamma}(|\omega|a, |\omega|/b). \tag{24}$$

This simple observation allows to evaluate the mean and variance of the marginals. We get:

$$\mathbb{E}\left(\frac{1}{\boldsymbol{\eta}[\mathbf{i}]}\right) = \frac{|\omega|}{b(|\omega|a - 1)} \tag{25}$$

and

$$\text{var}\left(\frac{1}{\boldsymbol{\eta}[\mathbf{i}]}\right) = \frac{|\omega|^2}{b^2(|\omega|a - 1)^2(|\omega|a - 2)}. \tag{26}$$

Since we have to set $a > 1$ in order to preserve convexity, we get that for large $|\omega|$, $\mathbb{E}\left(\frac{1}{\boldsymbol{\eta}[\mathbf{i}]}\right) \simeq \frac{1}{ba}$ and $\text{var}\left(\frac{1}{\boldsymbol{\eta}[\mathbf{i}]}\right) \simeq 0$. As a conclusion, the proposed model cannot reproduce large attenuation dynamics, whenever the patterns have a large support. This is a limitation of the model.

3.2.3 The generic case.

When $\boldsymbol{\psi}$ is an arbitrary nonnegative filter, studying the statistical properties of $\mathbf{1} \oslash \boldsymbol{\eta}$ is significantly more complicated. Let $\boldsymbol{\eta}[\mathbf{i}] = \sum_{\mathbf{j} \in \Omega} \boldsymbol{\lambda}[\mathbf{j}] \boldsymbol{\psi}[\mathbf{i} - \mathbf{j}]$ be a sum of independent random variables. Each term $\boldsymbol{\lambda}[\mathbf{j}] \boldsymbol{\psi}[\mathbf{i} - \mathbf{j}]$ in the sum follows a Gamma$(a, b\,\boldsymbol{\psi}[\mathbf{i} - \mathbf{j}])$ distribution. There is no simple analytical expression for such a sum, see e.g. [23]. We therefore do not investigate further the statistical properties of our model in the general case.

Fig. 2: Examples of a random process of type $\mathbf{1} \oslash (\boldsymbol{\psi} \star \boldsymbol{\lambda})$, where $\boldsymbol{\lambda}[\mathbf{i}] \sim \text{Gamma}(a, a)$ are independent random variables. First column: filter $\boldsymbol{\psi}$. From the second to the last column: realizations with different values of a : 0.2, 1.1 and 2.

3.3 Noise simulations

Figure 2 illustrates some realizations of a random vector of type $\mathbf{1} \oslash (\boldsymbol{\psi} \star \boldsymbol{\lambda})$ for various $\boldsymbol{\psi}$ and various a with $b = a$. As can be seen, the noise amplitude cannot vary too much for $a > 1$, when the filter size is large. This property was explained in the previous section. For $0 < a < 1$, the variance of the inverse-gamma distribution is undefined: the stochastic process $\boldsymbol{\lambda}$ can take isolated high values that dominate all the others. For instance, the top left realization took a single huge value exceeding the largest value tolerated by the computer. This explains why it looks completely gray.

Figure 3 illustrates different examples of noisy images that can be generated by model (18) for $a > 1$. As can be seen, various degradations resembling what is observed in SPIM imaging can be reproduced.
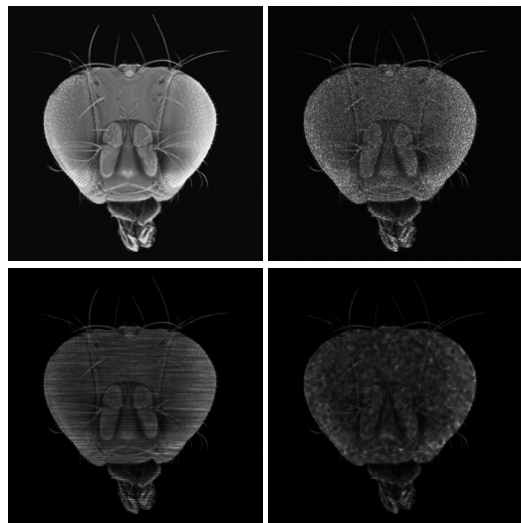

Fig. 3: Examples of noisy images generated by model (18). For all images, $\boldsymbol{\lambda}[\mathbf{i}] \sim \text{Gamma}(a, a)$. Top-Left: original image. Top-Right: $\boldsymbol{\psi} = \boldsymbol{\delta}$ and $a = 2$. Bottom-Left: $\boldsymbol{\psi}$ is a segment and $a = 1.1$. Bottom-Right: $\boldsymbol{\psi}$ is an indicator function of a 4×4 square and $a = 1.1$.

4 Restoration model

4.1 MAP based restoration models

In this paragraph, we propose a MAP restoration model. We aim at solving:

$$\max_{\mathbf{u}, \boldsymbol{\lambda}} \mathbb{P}(\mathbf{u}, \boldsymbol{\lambda} | \mathbf{u}_0). \tag{27}$$

By using Bayes rule, this amounts to:

$$\max_{\mathbf{u}, \boldsymbol{\lambda}} \frac{\mathbb{P}(\mathbf{u}_0 | \mathbf{u}, \boldsymbol{\lambda}) \mathbb{P}(\mathbf{u}, \boldsymbol{\lambda})}{\mathbb{P}(\mathbf{u}_0)}. \tag{28}$$

Assuming that $\mathbf{u}$ and $\boldsymbol{\lambda}$ are *independent*, we get that $\mathbb{P}(\mathbf{u}, \boldsymbol{\lambda}) = \mathbb{P}(\mathbf{u})\mathbb{P}(\boldsymbol{\lambda})$. Moreover, by setting

$$\varXi = \left\{ (\mathbf{u}, \boldsymbol{\lambda}) \in E_+ \times E_+, (\boldsymbol{\psi} \star \boldsymbol{\lambda}) \odot \mathbf{u}_0 = \mathbf{u} \right\}, \tag{29}$$

we get

$$\mathbb{P}(\mathbf{u}_0 | \mathbf{u}, \boldsymbol{\lambda}) = \begin{cases} 0 \text{ if } (\mathbf{u}, \boldsymbol{\lambda}) \notin \varXi, \\ 1 \text{ otherwise.} \end{cases} \tag{30}$$

By taking the negative log in (28), we therefore get that the maximizers of (28) coincide with the minimizers of:

$$\min_{(\mathbf{u}, \boldsymbol{\lambda}) \in \varXi} -\log(\mathbb{P}(\mathbf{u})) - \log(\mathbb{P}(\boldsymbol{\lambda})). \tag{31}$$

In this paper, we use the standard hypothesis that images have a low total variation, which can be expressed as:

$$\mathbb{P}(\mathbf{u}) \propto \exp\left(-c \|\|\nabla \mathbf{u}\|\|_1\right), \tag{32}$$

for some $c > 0$. Since $\boldsymbol{\lambda}$ follows a gamma distribution, the variational problem we end with reads:

$$\min_{(\mathbf{u}, \boldsymbol{\lambda}) \in \Xi} c \|\|\nabla \mathbf{u}\|\|_1 + \langle b\boldsymbol{\lambda} - (a - 1) \log \boldsymbol{\lambda}, \mathbf{1} \rangle. \tag{33}$$

Finally, by expressing $\mathbf{u}$ in terms of $\boldsymbol{\lambda}$, we obtain:

$$\min_{\boldsymbol{\lambda} \in E_+} c \|\|\nabla(\mathbf{u}_0 \odot (\boldsymbol{\psi} \star \boldsymbol{\lambda}))\|\|_1 + \langle b\boldsymbol{\lambda} - (a-1) \log \boldsymbol{\lambda}, \mathbf{1} \rangle. \tag{34}$$

This functional is convex for $b > 0$ and $a \geq 1$.

4.2 Simplifying the model

Model (34) contains three explicit parameters (a, b, c) and one implicit parameter $\|\boldsymbol{\psi}\|_1$: the ℓ^1-norm of $\boldsymbol{\psi}$. Tuning four parameters is very hard in practice and we propose to simplify parameter estimation in this section.

Proposition 1 *Problem* (34) *admits a unique minimizer if $c \geq 0$, $a > 1$ and $b > 0$.*

Proof Set

$$g(\boldsymbol{\lambda}) = c \|\|\nabla(\mathbf{u}_0 \odot (\boldsymbol{\psi} \star \boldsymbol{\lambda}))\|\|_1 + \langle b\boldsymbol{\lambda} - (a - 1) \log \boldsymbol{\lambda}, \mathbf{1} \rangle. \tag{35}$$

First notice that the set E_+ is convex closed. Moreover, $g(\boldsymbol{\lambda})$ is strictly convex and closed. Indeed, for $a > 1$, the function defined on E_+ by $\boldsymbol{\lambda} \mapsto -(a-1)\langle \log \boldsymbol{\lambda}, \mathbf{1} \rangle$ is strictly convex since its Hessian $(a-1)\mathrm{diag}\left(\mathbf{1} \oslash (\boldsymbol{\lambda}^2) \right)$ is positive definite for all $\boldsymbol{\lambda} \in E_+$. Problem (34) consists of minimizing a strictly convex function over a convex closed set. This is sufficient to ensure uniqueness of a minimizer if it exists. The existence of a solution is granted by noticing that $g(\boldsymbol{\lambda})$ is continuous on the interior of its domain since it is convex. Moreover, it is coercive since the 1D function $t \mapsto -(a-1)\log(t) + bt$ goes to $+\infty$ as $t \to +\infty$.

In order to reduce the number of parameters, it is possible to divide the cost function in (34) by $c\|\boldsymbol{\psi}\|_1$. By doing so, it is straightforward to see that the minimizer of (34) is equal to the minimizer of:

$$\min_{\boldsymbol{\lambda} \in E_+} \|\|\nabla(\mathbf{u}_0 \odot (\tilde{\boldsymbol{\psi}} \star \boldsymbol{\lambda}))\|\|_1 + \langle \beta\boldsymbol{\lambda} - \alpha \log \boldsymbol{\lambda}, \mathbf{1} \rangle. \tag{36}$$

where $\tilde{\boldsymbol{\psi}} = \frac{\boldsymbol{\psi}}{\|\boldsymbol{\psi}\|_1}$, $\alpha = (a-1)/(c\|\boldsymbol{\psi}\|_1)$ and $\beta = b/c\|\boldsymbol{\psi}\|_1$. Model (36) still contains two parameters. The following proposition shows that it can be further reduced to one parameter.

Proposition 2 *Let $\boldsymbol{\lambda}(\alpha, \beta)$ denote the minimizer of* (36). *Then, for all $t > 0$*

$$\boldsymbol{\lambda}(t\alpha, \beta) = t\boldsymbol{\lambda}(\alpha, \beta). \tag{37}$$

Proof Let $\tilde{\boldsymbol{\lambda}} = t\boldsymbol{\lambda}(\alpha, \beta)$. The minimizer $\boldsymbol{\lambda}(\alpha, \beta)$ of (36) satisfies:

$$0 \in \tilde{\mathbf{A}}^* \partial_{\|\cdot\|_1}(\tilde{\mathbf{A}}\boldsymbol{\lambda}(\alpha, \beta)) + \beta\mathbf{1} - \alpha\mathbf{1} \oslash \boldsymbol{\lambda} + \mathcal{N}_{E_+}(\boldsymbol{\lambda}), \tag{38}$$

where $\tilde{\mathbf{A}}$ is the linear operator defined for all $\boldsymbol{\lambda} \in E$ by

$$\tilde{\mathbf{A}}\boldsymbol{\lambda} = \nabla(\mathbf{u}_0 \odot (\tilde{\boldsymbol{\psi}} \star \boldsymbol{\lambda})). \tag{39}$$

Therefore $\tilde{\boldsymbol{\lambda}}$ obeys:

$$0 \in \tilde{\mathbf{A}}^* \partial_{\|\cdot\|_1}(\tilde{\mathbf{A}}\tilde{\boldsymbol{\lambda}}/t) + \beta\mathbf{1} - \alpha t\mathbf{1} \oslash \tilde{\boldsymbol{\lambda}} + \mathcal{N}_{E_+}(\tilde{\boldsymbol{\lambda}}). \tag{40}$$

Since $\partial_{\|\cdot\|_1}(\tilde{\mathbf{A}}\tilde{\boldsymbol{\lambda}}/t) = \partial_{\|\cdot\|_1}(\tilde{\mathbf{A}}\tilde{\boldsymbol{\lambda}})$, this is still equivalent to:

$$0 \in \tilde{\mathbf{A}}^* \partial_{\|\cdot\|_1}(\tilde{\mathbf{A}}\tilde{\boldsymbol{\lambda}}) + \beta\mathbf{1} - \alpha t\mathbf{1} \oslash \tilde{\boldsymbol{\lambda}} + \mathcal{N}_{E_+}(\tilde{\boldsymbol{\lambda}}). \tag{41}$$

This inclusion characterizes the minimizer of problem (36), where α is replaced by $t\alpha$.

Proposition 2 indicates that only β really matters. The value α only allows to set the amplitude of the solution. In fluorescence microscopy, this amplitude depends on the excitation intensity which is usually unknown. Fluorescence images are therefore not used for quantitative measures, and only contrasts matter. As a conclusion, the model could be further simplified by setting $\alpha = 1$. We will see in Proposition 3 that setting $\beta = \alpha$ allows preserving the harmonic mean of $\mathbf{u}_0$ and this choice therefore seems more natural. The following model summarizes the simplifications we have proposed in this section.

Proposed variational formulation.
The variational problem we study in this paper reads:

$$\min_{\boldsymbol{\lambda} \in E_+} \|\|\nabla(\mathbf{u}_0 \odot (\boldsymbol{\psi} \star \boldsymbol{\lambda}))\|\|_1 + \alpha \langle \boldsymbol{\lambda} - \log(\boldsymbol{\lambda}), \mathbf{1} \rangle, \tag{42}$$

where $\boldsymbol{\psi} \in E_+$, $\|\boldsymbol{\psi}\|_1 = 1$ and $\alpha > 0$.

4.3 Duality and mean preservation

In this paragraph, we prove that model (42) preserves the *harmonic* mean of the input image $\mathbf{u}_0$ in the specific case $\boldsymbol{\psi} = \boldsymbol{\delta}$. In the generic case, we prove that the weighted harmonic mean of $\boldsymbol{\lambda}$ is constant.

Proposition 3 *Assume that $\mathbf{u}_0 > 0$ and let $\bar{\boldsymbol{\lambda}}$ be any vector such that $\boldsymbol{\psi} \star \bar{\boldsymbol{\lambda}} = \mathbf{1} \oslash \mathbf{u}_0$ [1]. Then*

$$\sum_{\mathbf{i} \in \Omega} \frac{\bar{\boldsymbol{\lambda}}[\mathbf{i}]}{\boldsymbol{\lambda}[\mathbf{i}]} = \sum_{\mathbf{i} \in \Omega} \bar{\boldsymbol{\lambda}}[\mathbf{i}], \tag{43}$$

---
[1] A sufficient condition for existence of such a $\bar{\boldsymbol{\lambda}}$ is that the Fourier transform $\hat{\boldsymbol{\psi}}$ does not vanish.

In the particular case $\boldsymbol{\psi} = \boldsymbol{\delta}$, this yields:

$$\sum_{\mathbf{i}\in\Omega} \frac{1}{\mathbf{u}[\mathbf{i}]} = \sum_{\mathbf{i}\in\Omega} \frac{1}{\mathbf{u}_0[\mathbf{i}]}. \tag{44}$$

Proof Let us first use Fenchel-Rockafellar duality [27], in order to derive the dual problem of (42). Let $\mathbf{A}$ be the linear operator defined, for all $\boldsymbol{\lambda} \in E$, by $\mathbf{A}\boldsymbol{\lambda} = \nabla(\mathbf{u}_0 \odot (\boldsymbol{\psi} \star \boldsymbol{\lambda}))$. We have:

$$
\begin{aligned}
&\min_{\boldsymbol{\lambda}\in E_+} \||\mathbf{A}\boldsymbol{\lambda}|\|_1 + \alpha\left\langle\boldsymbol{\lambda} - \log(\boldsymbol{\lambda}), \mathbf{1}\right\rangle\\
&= \min_{\boldsymbol{\lambda}\in E_+} \sup_{\mathbf{q}\in V, \|\|\mathbf{q}\|\|_\infty\le 1} \left\langle\mathbf{A}\boldsymbol{\lambda}, \mathbf{q}\right\rangle + \alpha\left\langle\boldsymbol{\lambda} - \log(\boldsymbol{\lambda}), \mathbf{1}\right\rangle\\
&= \sup_{\mathbf{q}\in V, \|\|\mathbf{q}\|\|_\infty\le 1} \inf_{\boldsymbol{\lambda}\in E_+} \left\langle\mathbf{A}\boldsymbol{\lambda}, \mathbf{q}\right\rangle + \alpha\left\langle\boldsymbol{\lambda} - \log(\boldsymbol{\lambda}), \mathbf{1}\right\rangle.
\end{aligned}
$$

The minimizer of the inner-problem satisfies:

$$\mathbf{A}^*\mathbf{q} + \alpha\mathbf{1} - \alpha\mathbf{1}\oslash\boldsymbol{\lambda} + \mathcal{N}_{E_+}(\boldsymbol{\lambda}) \ni 0. \tag{45}$$

Since the logarithm is a barrier function for the set E_+, the minimizer belongs to the interior of E_+ and the optimality conditions simplify to:

$$\mathbf{A}^*\mathbf{q} + \alpha\mathbf{1} - \alpha\mathbf{1}\oslash\boldsymbol{\lambda} = 0. \tag{46}$$

Therefore, the primal-dual pair $(\boldsymbol{\lambda}, \mathbf{q})$ satisfies:

$$\boldsymbol{\lambda} = \alpha\mathbf{1}\oslash(\mathbf{A}^*\mathbf{q} + \alpha\mathbf{1}). \tag{47}$$

By injecting this result in the inner-problem, we obtain the dual problem:

$$\sup_{\|\|\mathbf{q}\|\|_\infty\le 1} \left\langle\log\left(\alpha\mathbf{1} + \mathbf{A}^*\mathbf{q}\right), \mathbf{1}\right\rangle. \tag{48}$$

Then, remark that:

$$\mathbf{1}\oslash\boldsymbol{\lambda} = \frac{\mathbf{A}^*\mathbf{q}}{\alpha} + \mathbf{1}. \tag{49}$$

By construction, $\bar{\boldsymbol{\lambda}} \in \mathrm{Ker}(\mathbf{A})$. Since $\mathrm{Ker}(\mathbf{A}) = \mathrm{Im}(\mathbf{A}^*)^\perp$, we get that $\left\langle\mathbf{A}^*\mathbf{q}, \bar{\boldsymbol{\lambda}}\right\rangle = 0$ for all $\mathbf{q} \in V$. Therefore:

$$\left\langle\mathbf{1}\oslash\boldsymbol{\lambda}, \bar{\boldsymbol{\lambda}}\right\rangle = \left\langle\mathbf{1}, \bar{\boldsymbol{\lambda}}\right\rangle. \tag{50}$$

If $\boldsymbol{\psi} = \boldsymbol{\delta}$, then $\bar{\boldsymbol{\lambda}} = \mathbf{1}\oslash\mathbf{u}_0$ and $\boldsymbol{\lambda} = \mathbf{u}\oslash\mathbf{u}_0$. This yields (44).

4.4 Relationships to other models in the white noise regime

When $\boldsymbol{\psi} = \boldsymbol{\delta}$ is a Dirac mass, our model is adapted to white noise denoising. Problem (42) can be rephrased in terms of $\mathbf{u}$ as:

$$\min_{\mathbf{u}\in E_+} \||\nabla\mathbf{u}|\|_1 + \alpha\left\langle\mathbf{u}\oslash\mathbf{u}_0 - \log(\mathbf{u}\oslash\mathbf{u}_0), \mathbf{1}\right\rangle. \tag{51}$$

Comparatively, the convex Shi-Osher or Teuber-Steidl model proposed in [30, 31] is:

$$\min_{\mathbf{u}\in E_+} \||\nabla\mathbf{u}|\|_1 + \alpha\left(\left\langle\mathbf{u}, \mathbf{1}\right\rangle - \left\langle\mathbf{u}_0, \log(\mathbf{u})\right\rangle\right). \tag{52}$$

The nonconvex Aubert-Aujol model proposed in [3] is:

$$\min_{\mathbf{u}\in E_+} \||\nabla\mathbf{u}|\|_1 + \alpha\left\langle\mathbf{u}_0\oslash\mathbf{u} + \log(\mathbf{u}), \mathbf{1}\right\rangle. \tag{53}$$

All models are proved to preserve the geometric mean, while our preserves the harmonic mean. It is unclear to us what property is better. We will however see in the numerical experiments that our model better preserves contrasts.

5 Solvers for the proposed model (42)

In this section we provide an algorithm to solve problem (42). A large amount of numerical approaches were developed recently to solve non-smooth convex problems of that type. Problem (42) can be rewritten as:

$$\min_{\boldsymbol{\lambda}\in E} J(\boldsymbol{\lambda}) = F(\mathbf{A}\boldsymbol{\lambda}) + G(\boldsymbol{\lambda}), \tag{54}$$

where $\mathbf{A} : E \to V$ defined by $\mathbf{A}\boldsymbol{\lambda} = \nabla(\mathbf{u}_0 \odot (\boldsymbol{\psi} \star \boldsymbol{\lambda}))$, $F : V \to \mathbb{R}$ is defined as $F(\mathbf{q}) = \||\mathbf{q}|\|_1$ and $G : E \to \mathbb{R}\cup\{+\infty\}$ is defined as $G(\boldsymbol{\lambda}) = \alpha\left\langle\boldsymbol{\lambda} - \log\boldsymbol{\lambda}, \mathbf{1}\right\rangle + \chi_{E_+}(\boldsymbol{\lambda})$. This reformulation perfectly fits the framework of the first-order primal-dual algorithm proposed in [6]. It is described in Algorithm 1.

Algorithm 1: First order primal-dual algorithm [6]

Input	: ϵ : the desired precision,
	$(\mathbf{x}_0, \mathbf{y}_0) \in E_+ \times V$: a starting point.
Output	: an approximate solution of problem (54).
Init.	: Choose τ, $\sigma > 0$ such that $\sigma\tau\|\mathbf{A}\|^2 < 1$,
	set $\theta = 1$, $k = 0$ and $\bar{\mathbf{x}}_0 = \mathbf{x}_0$.

1. **while** *Convergence criterion* $> \epsilon$ **do**
2. $\mathbf{y}_{k+1} = \mathrm{prox}_{\sigma F^*}(\mathbf{y}_k + \sigma\mathbf{A}\bar{\mathbf{x}}_k)$
3. $\mathbf{x}_{k+1} = \mathrm{prox}_{\tau G}(\mathbf{x}_k - \tau\mathbf{A}^*\mathbf{y}_{k+1})$
4. $\bar{\mathbf{x}}_{k+1} = \mathbf{x}_{k+1} + \theta(\mathbf{x}_{k+1} - \mathbf{x}_k)$
5. $k = k + 1$
6. **end**

Algorithm 1 generates a feasible sequence of iterates $(\mathbf{x}_k)_{k\in\mathbb{N}}$ that converges to the minimizer $\mathbf{x}^*$ of (54). Moreover the iterates satisfy $J(\mathbf{x}_k) - J(\mathbf{x}^*) = O\left(\frac{1}{k}\right)$ which is arguably optimal. The proximity operators of F^* and G have a simple closed form expression:

$$\mathrm{prox}_{\sigma F^*}(\mathbf{q})[\mathbf{i}] = \frac{\mathbf{q}[\mathbf{i}]}{\max(1, |\mathbf{q}|[\mathbf{i}])} \tag{55}$$

and

$$\mathrm{prox}_{\tau G}(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x} - \tau\alpha\mathbf{1} + \sqrt{(\mathbf{x} - \tau\alpha\mathbf{1})^2 + 4\tau\alpha\mathbf{1}}\right). \tag{56}$$

This algorithm was implemented on a GPU using CUDA. Except convolutions, all operators appearing in this algorithm are pixel-wise operations which are the perfect setting for GPU. To implement convolutions we used the cuFFT library. In all experiments, the GPU implementation has been run in double precision on a NVIDIA Tesla K20c containing 2496 CUDA cores and 5GB internal memory.

6 Numerical experiments

In this section, we first provide some numerical results on synthetic images, where ground-truth is available and then turn to real images.

6.1 The case of white noise

In this paragraph, we first generate noisy images by the formula:

$$\mathbf{u}_0 = \mathbf{u} \odot \boldsymbol{\eta}, \tag{57}$$

where $\boldsymbol{\eta} \in E$ is a random vector with i.i.d. components. The distribution of the marginals $\boldsymbol{\eta}[\mathbf{i}]$ is set as a gamma distribution $\mathrm{Gamma}(a, a)$ in Figure 4 and as an inverse-gamma distribution in Figure 5. Our model is supposed to be well adapted only to the inverse-gamma distribution. We provide comparisons with the denoising model (52) proposed in [30, 31]. The minimizer of (42) is referred to as MSNR, for Multiplicative Stationary Noise Removal.

In order to compare the restoration results, we evaluate the rescaled Signal to Noise Ratio, denoted SNRr and defined by:

$$\mathrm{SNRr}(\mathbf{u}, \mathbf{u}_0) = -\min_{a\in\mathbb{R}} \log_{10}\left(\frac{\|a\mathbf{u} - \mathbf{u}_0\|_2^2}{\|\mathbf{u}_0\|_2^2}\right). \tag{58}$$

The reason to use this image quality measure is that the models only yield results valid up to a multiplicative

(a) Original   (b) Noisy
SNRr = 17.17dB
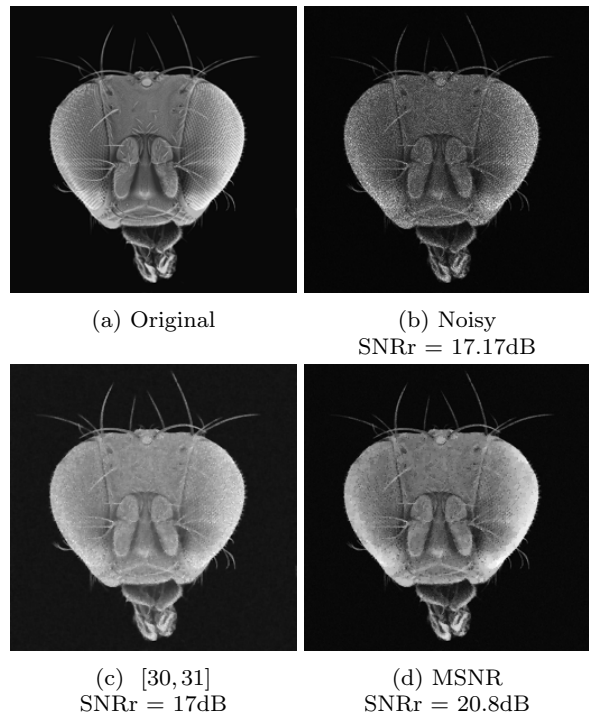
(c) [30, 31]   (d) MSNR
SNRr = 17dB   SNRr = 20.8dB

Fig. 4: Denoising experiment for image formation model $\mathbf{u}_0 = \mathbf{u}\odot\boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is a gamma distributed white noise with $\boldsymbol{\eta}[1] \sim \mathrm{Gamma}(50, 50)$. The SNRr for MSNR is stable after 50 iterations. Computing time MATLAB - CPU: 0.15s.

	a	10	20	50
Degraded	SNRr (dB)	10.6	13.2	17.1
	SSIMr	0.69	0.78	0.87
[30, 31]	SNRr (dB)	15.5	16	16.6
	SSIMr	0.78	0.81	0.83
MSNR	SNRr (dB)	17.2	18.7	20.7
	SSIMr	0.79	0.85	0.90

Table 1: SNRr and SSIMr for image formation model $\mathbf{u}_0 = \mathbf{u}\odot\boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is a gamma distributed white noise with $\boldsymbol{\eta}[1] \sim \mathrm{Gamma}(a, a)$.

constant. We also compute another image quality measure denoted SSIMr, defined as the Standard Structural Similarity Index Measure [34] applied to the rescaled image $a\mathbf{u}$ where a is optimized in (58). In all experiments, the model parameters are set so as to maximize the SNRr.

As can be seen in Figures 4 and 5 and Tables 1 and 2, our model provides significantly higher rescaled SNRs and SSIMs than model (52). When having a close look at the images, it can be seen however that the details are preserved similarly in both approaches. The main difference lies in the fact that model (52) does not preserve the contrasts as well as the proposed model.
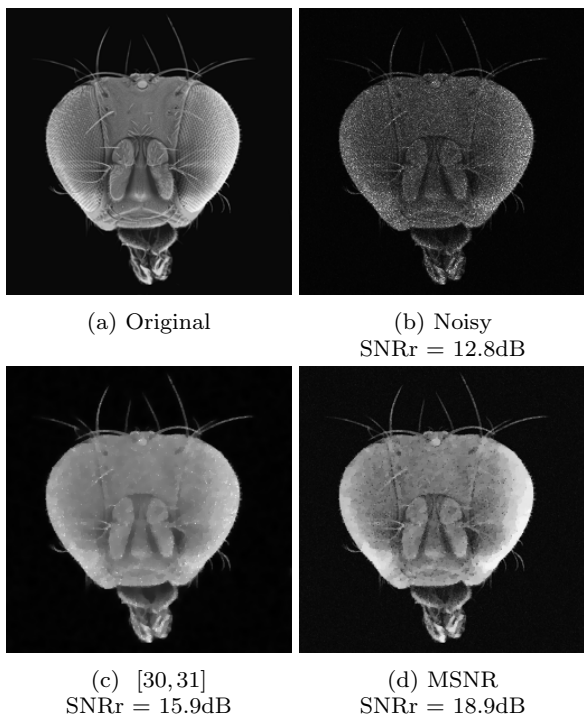
(a) Original

(b) Noisy
SNRr = 12.8dB

(c) [30, 31]
SNRr = 15.9dB

(d) MSNR
SNRr = 18.9dB

Fig. 5: Denoising experiment for image formation model $\mathbf{u}_0 = \mathbf{u} \oslash \boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is a gamma distributed white noise with $\boldsymbol{\eta}[1] \sim \mathrm{Gamma}(20, 20)$. The SNRr for MSNR is stable after 50 iterations. Computing time MATLAB - CPU: 0.15s.

		10	20	50
	a			
Degraded	SNRr (dB)	9.84	12.8	16.9
	SSIMr	0.66	0.77	0.87
[30, 31]	SNRr (dB)	15.2	16	16.8
	SSIMr	0.79	0.81	0.85
MSNR	SNRr (dB)	17.4	18.9	20.8
	SSIMr	0.79	0.85	0.91

Table 2: SNRr and SSIMr for image formation model $\mathbf{u}_0 = \mathbf{u} \oslash \boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is a gamma distributed white noise with $\boldsymbol{\eta}[1] \sim \mathrm{Gamma}(a, a)$.

6.2 The case of structured noise

We now investigate the efficiency of the method proposed in this paper versus VSNR (Variational Stationary Noise Remover) [13,14]. We use the 256×256 drosophila image see Figure 3 top left, corrupted with noise generated using two different filters $\boldsymbol{\psi}$. The first one is an anisotropic Gaussian and the second one is the indicator of a square. The data term parameter α is set so as to maximize the SNRr. Results are displayed Figures 6 and 7. The proposed method provides higher rescaled SNR and images containing finer details and more natural contrasts. The computing times are also very appealing.

(a) Noisy – SNRr = 15.1dB

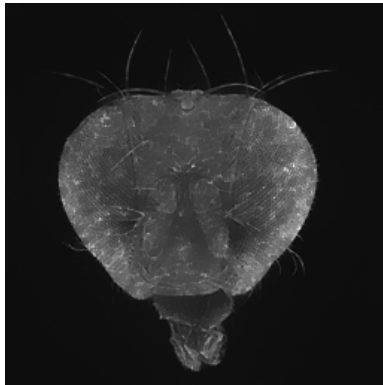(b) VSNR – SNRr = 19.1dB

(c) MSNR – SNRr = 19.9dB

Fig. 6: Denoising simulation with structured noise. The image is corrupted with $\boldsymbol{\psi}$ a segment and $a = 1.1$ see Figure 3 bottom-left. The SNRr is stable after 50 iterations. Computing time GPU: 0.03s.

## 6.3 A few results on real images

To end up with our numerical validation, we perform experiments on real images. We consider two different images and compare the results on these images obtained with VSNR and the method proposed in this paper, see Figures 8 and 10. Following the previous works [13,14], we set $\boldsymbol{\psi}$ as an anisotropic Gaussian filter
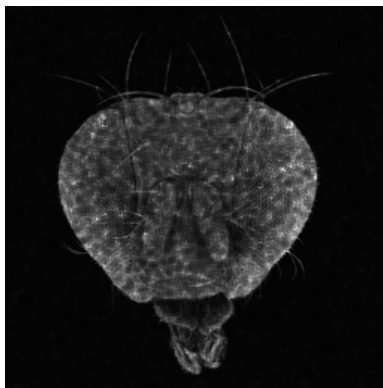
(a) Noisy – SNRr = 12.5dB

(b) VSNR – SNRr = 14.8dB
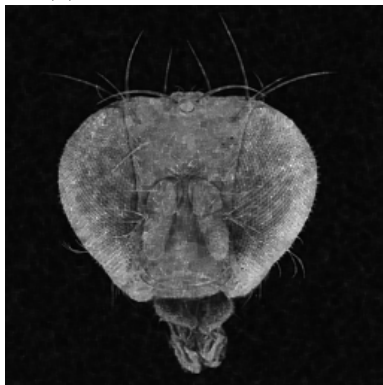
(c) MSNR – SNRr = 15.4dB

Fig. 7: Denoising simulation with structured noise. The image is corrupted with ψ the indicator function 4×4 square and $a = 1.1$ see Figure 3 bottom-right. The SNRr is stable after 40 iterations. Computing time GPU: 0.025s.

elongated in the stripes direction. This choice allows to accurately describe the noise in the frequency domain. In particular, the filter's smoothness ensures that only low frequencies will be changed by the variational model (42). This is a key feature to preserve fine details in the image.

The results can only be compared qualitatively since no ground truth is available. A quick inspection of the results advocates for the method proposed in this paper. Images restored with VSNR suffer from smear artifacts at the locations of stripes, especially in dark region. Images proposed by the MSNR do not suffer from these artifacts and overall have a better contrast see Figure 9.

Acknowledgment

References

1. I. Aizenberg and C. Butakoff. A windowed Gaussian notch filter for quasi-periodic noise removal. *Image and Vision Computing*, 26(10):1347–1353, 2008.
2. E. M. A. Anas, S. Y. Lee, and M. Kamrul Hasan. Classification of ring artifacts for their effective removal using type adaptive correction schemes. *Computers in Biology and Medicine*, 41(6):390–401, 2011.
3. G. Aubert and J.-F. Aujol. A variational approach to removing multiplicative noise. *SIAM Journal on Applied Mathematics*, 68(4):925–946, 2008.
4. F. E. Boas and D. Fleischmann. CT artifacts: causes and reduction techniques. *Imaging in Medicine*, 4(2):229–240, 2012.
5. A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical imaging and vision*, 20(1-2):89–97, 2004.
6. A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.
7. Y. Chang, H. Fang, L. Yan, and H. Liu. Robust destriping method with unidirectional total variation and framelet regularization. *Optics Express*, 21(20):23307–23323, 2013.
8. S.-W. Chen and J.-L. Pellequer. Destripe: frequency-based algorithm for removing stripe noises from afm images. *BMC Structural Biology*, 11(1):7, 2011.
9. B. Cornelis, A. Dooms, J. Cornelis, and P. Schelkens. Digital canvas removal in paintings. *Signal Processing*, 92(4):1166–1171, 2012.
10. C. Deledalle, L. Denis, and F. Tupin. Iterative weighted maximum likelihood denoising with probabilistic patch-based weights. *IEEE Transactions on Image Processing*, 18(12):2661–2672, 2009.
11. Y. Dong and T. Zeng. A convex variational model for restoring blurred images with multiplicative noise. *SIAM Journal on Imaging Sciences*, 6(3):1598–1625, 2013.
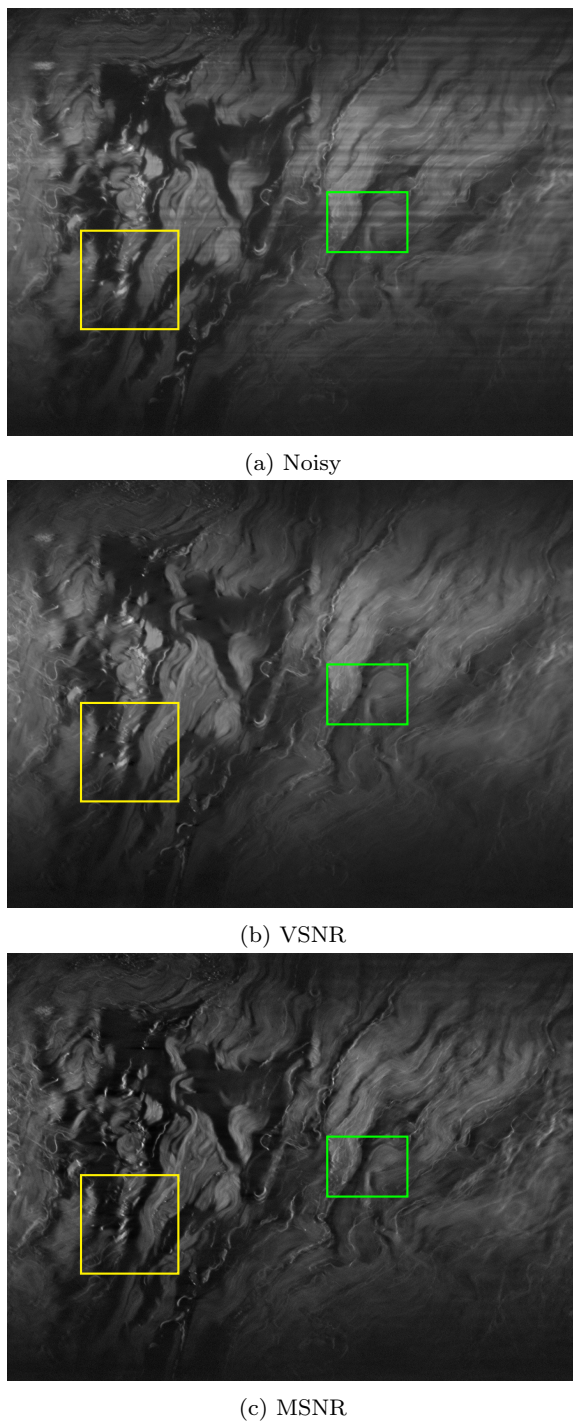
(a) Noisy

(b) VSNR

(c) MSNR

Fig. 8: Denoising experiment on real data: 1040×1390. The degraded image is a XY plane of an skin dermis acquired using a SPIM. The image is used here by courtesy of Claire Jardet from Genoskin. The iterates stabilize visually after 300 iterations. Computing time GPU: 3.3s (about 11ms per iteration)
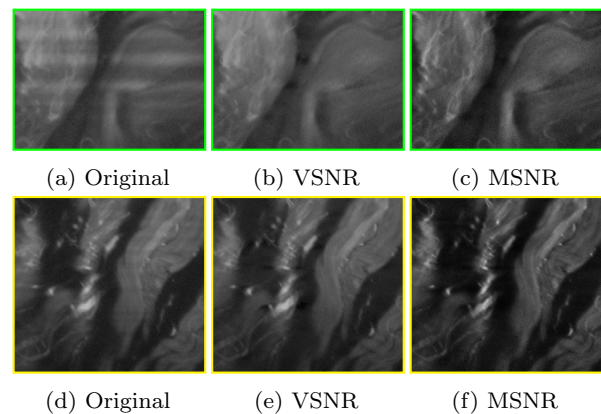

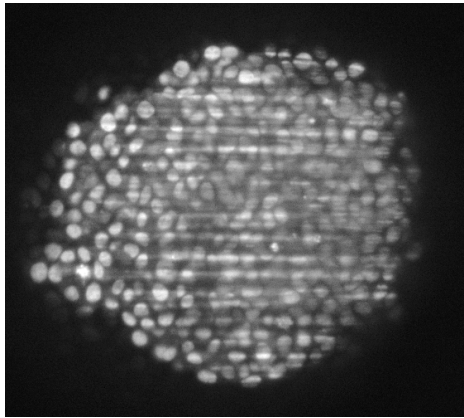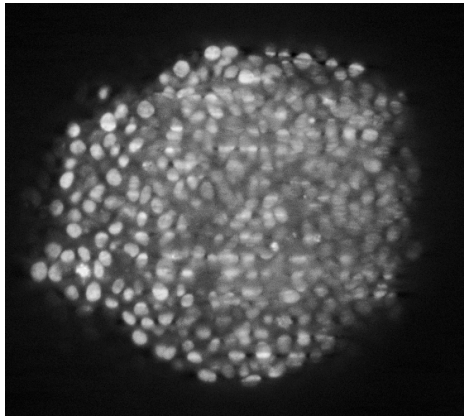| (a) Original | (b) VSNR | (c) MSNR |
| (d) Original | (e) VSNR | (f) MSNR |

Fig. 9: Zoom on Figure 8. Observe in the top figure that VSNR algorithm smears the black region. In the bottom figure, observe that the contrasts seem slightly more pronounced in MSNR.
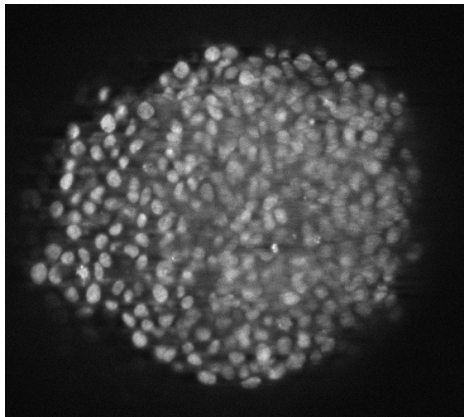
12. S. Durand, J. Fadili, and M. Nikolova. Multiplicative noise removal using l1 fidelity on frame coefficients. *Journal of Mathematical Imaging and Vision*, 36(3):201–226, 2010.

13. J. Fehrenbach and P. Weiss. Processing stationary noise: model and parameter selection in variational methods. *SIAM Journal on Imaging Sciences*, 7(2):613–640, 2013.

14. J. Fehrenbach, P. Weiss, and C. Lorenzo. Variational algorithms to remove stationary noise: applications to microscopy imaging. *Image Processing, IEEE Transactions on*, 21(10):4420–4430, 2012.

15. J. H. Fitschen, J. Ma, and S. Schuff. Removal of curtaining effects by a variational model with directional first and second order differences. *arXiv preprint arXiv:1507.00112*, 2015.

16. L. Gómez-Chova, L. Alonso, L. Guanter, G. Camps-Valls, J. Calpe, and J. Moreno. Correction of systematic spatial noise in push-broom hyperspectral sensors: application to chris/proba images. *Applied Optics*, 47(28):F46–F60, 2008.

17. J. Hsieh. Computed tomography: principles, design, artifacts, and recent advances. SPIE Bellingham, WA, 2009.

18. Y.-M. Huang, L. Moisan, M. K. Ng, and T. Zeng. Multiplicative noise removal via a learned dictionary. *Image Processing, IEEE Transactions on*, 21(11):4534–4543, 2012.

19. J. Huisken, J. Swoger, F. Del Bene, J. Wittbrodt, and E. H. Stelzer. Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science*, 305(5686):1007–1009, 2004.

20. Y. Kim, J. Baek, and D. Hwang. Ring artifact correction using detector line-ratios in computed tomography. *Optics Express*, 22(11):13380–13392, 2014.

21. A. Kryvanos, J. Hesser, and G. Steidl. Nonlinear image restoration methods for marker extraction in 3D fluorescent microscopy. In *Electronic Imaging 2005*, pages 432–443. International Society for Optics and Photonics, 2005.

22. A. Levin and B. Nadler. Natural image denoising: Optimality and inherent bounds. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2833–2840. IEEE, 2011.

(a) Noisy

(b) VSNR

(c) MSNR

Fig. 10: Denoising experiment on real data : 499×561. The degraded image is a XY plane of spheroid acquired using a SPIM. The iterates stabilize visually after 500 iterations. Computing time GPU: 3s (about 6ms per iteration)

23. P. G. Moschopoulos. The distribution of the sum of independent gamma random variables. *Annals of the Institute of Statistical Mathematics*, 37(1):541–544, 1985.
24. B. Münch, P. Trtik, F. Marone, and M. Stampanoni. Stripe and ring artifact removal with combined wavelet Fourier filtering. *Optics express*, 17(10):8567–8591, 2009.
25. V. Panin, G. Zeng, and G. Gullberg. Total variation regulated EM algorithm. *Nuclear Science, IEEE Transactions on*, 46(6):2202–2210, 1999.
26. A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy. A versatile wavelet domain noise filtration technique for medical imaging. *Medical Imaging, IEEE Transactions on*, 22(3):323–331, 2003.
27. R. T. Rockafellar. *Convex analysis*. Princeton university press, 2015.
28. L. Rudin, P.-L. Lions, and S. Osher. Multiplicative denoising and deblurring: theory and algorithms. In *Geometric Level Set Methods in Imaging, Vision, and Graphics*, pages 103–119. Springer, 2003.
29. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, et al. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682, 2012.
30. J. Shi and S. Osher. A nonlinear inverse scale space method for a convex multiplicative noise model. *SIAM Journal on Imaging Sciences*, 1(3):294–321, 2008.
31. G. Steidl and T. Teuber. Removing multiplicative noise by Douglas-Rachford splitting methods. *Journal of Mathematical Imaging and Vision*, 36(2):168–184, 2010.
32. F. Sur and M. Grédiac. Automated removal of quasiperiodic noise using frequency domain statistics. *Journal of Electronic Imaging*, 24(1):013003–013003, 2015.
33. F. Tsai and W. W. Chen. Striping noise detection and correction of remote sensing images. *Geoscience and Remote Sensing, IEEE Transactions on*, 46(12):4122–4131, 2008.
34. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on*, 13(4):600–612, 2004.
35. X.-L. Zhao, F. Wang, and M. K. Ng. A new convex optimization model for multiplicative noise and blur removal. *SIAM Journal on Imaging Sciences*, 7(1):456–475, 2014.