

THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Probabilistic modeling and inference for sequential space-varying blur identification

Citation for published version:
Huang, Y, Chouzenoux, E & Elvira, V 2021, 'Probabilistic modeling and inference for sequential space-varying blur identification', *IEEE Transactions on Computational Imaging*, vol. 7, pp. 531 - 546. https://doi.org/10.1109/TCI.2021.3081059

Digital Object Identifier (DOI):
10.1109/TCI.2021.3081059

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
IEEE Transactions on Computational Imaging

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Download date: 28. Mar. 2024

Probabilistic modeling and inference for sequential space-varying blur identification

Yunshi Huang, Émilie Chouzenoux, *Senior Member IEEE*, and Víctor Elvira, *Senior Member IEEE*

***Abstract*—The identification of parameters of spatially variant blurs given a clean image and its blurry noisy version is a challenging inverse problem of interest in many application fields, such as biological microscopy and astronomical imaging. In this paper, we consider a parametric model of the blur and introduce an 1D state-space model to describe the statistical dependence among the neighboring kernels. We apply a Bayesian approach to estimate the posterior distribution of the kernel parameters given the available data. Since this posterior is intractable for most realistic models, we propose to approximate it through a sequential Monte Carlo approach by processing all data in a sequential and efficient manner. Additionally, we propose a new sampling method to alleviate the particle degeneracy problem, which is present in approximate Bayesian filtering, particularly in challenging concentrated posterior distributions. The considered method allows us to process sequentially image patches at a reasonable computational and memory costs. Moreover, the probabilistic approach we adopt in this paper provides uncertainty quantification which is useful for image restoration. The practical experimental results illustrate the improved estimation performance of our novel approach, demonstrating also the benefits of exploiting the spatial structure the parametric blurs in the considered models.**

***Index Terms*—Blur identification; spatially-variant blur; Bayesian estimation; particle filtering.**

I. Introduction

Images produced by optical instruments often suffer from blur caused by light diffraction or object motion. The presence of the blur provokes an infinitesimal point-source to be spread in the acquired image defining the so-called Point Spread Function (PSF). When the PSF is unknown, one can resort to a blind deconvolution strategy [1], [2], [3], [4] to jointly retrieve the image and blur. Another approach consists in first identifying the blur and then removing it from the degraded images using a non-blind restoration strategy [5], [6]. This demands a good accuracy for the PSF estimation, which can be efficiently reached by a preliminary acquisition step of normalized and calibrated objects, such as fluorescent spherical microbeads in microscopy [7], [8] or resolution charts in digital camera calibration [9], [10]. The PSF identification problem is formulated as a least-squares one. On top of serving for image restoration purposes, the identified PSF can also be fitted into a parametric non-linear model in order to determine characteristics of the optical system [11], [12], [13], [14]. In most realistic scenarios, the stationary PSF model is not suitable due to extended depth of field [15], [16], moving objects or cameras [17], [18], anisotropic optical lens aberrations [19], or atmospheric turbulence [20], [21], [22]. Such image degradation sources give rise to a so-called spatially variant blur [23], [24]. The PSF identification then requires the recovery of a PSF map, describing the blur kernel at each location of the spatial plane [25].

In this paper, we address the problem of estimating in a sequential manner the parameters of spatially variant PSFs from calibrated image acquisitions. We consider a flexible piecewise constant parametric model for the space-varying PSF map that allows us to describe smooth variations among PSFs acting on neighbor regions. We formulate a state-space model where each time step corresponds to a different patch location. We adopt a probabilistic approach, aiming at producing posterior distributions of the unknowns. The probabilistic approach allows for dealing with the uncertainty in a systematic manner and the inclusion of prior knowledge about the unknowns. As a consequence, we are able not only to provide uncertainty measures on the unknown parameters, but also to propagate this uncertainty to useful tasks where those estimated parameters are used (e.g., in deblurring). In our approach, the prior probability density function (pdf) of the parameters has a Markovian structure, which allows us to inherit existing inference approaches for the sequential inference. Note that more complicated local dependencies could be modeled, e.g., via Markov random fields, at the expense of losing the aforementioned simplicity in the inference task. Our flexible formulation includes the consideration of potential non-standard observation and transition models. More precisely, we can operate virtually with any non-linear and non-Gaussian model. Due to the intractability of the Bayesian recursions, we resort to particle filtering (PF) for an approximate inference solution. We start by considering existing PF methods, such as the well-known bootstrap PF (BPF) [26]. While the BPF can obtain adequate results in some models, it presents some deficiencies, for instance in scenarios with low observation noise, where the peaky likelihood challenges the diversity among particles (see for instance [27] about the particle degeneracy problem). Due to these limitations, we propose a new PF method that we call *Generalized Interacting Annealed PF* (GIAnPF). The new PF tackles the particle degeneracy by considering a sequence of intermediate distributions, that are annealed versions of the filtering distribution. The considered sequential Bayesian framework provides us with three key advantages, namely (i) low computational cost and limited memory load due to a sequential processing; (ii) a flexible choice of the state-space models that enables us to consider non-linear relations and non-Gaussian noise; and (iii) a measure of statistical uncer-

The authors acknowledge support from the *Agence Nationale de la Recherche* of France under PISCES (ANR-17-CE40-0031-01) and MAJIC (ANR-17-CE40-0004-01) projects.

tainty on the estimated parameters of kernels. The sequential approach can be exploited to build online implementations in applications involving very large scale images available in batches. Dealing with the image in a sequential manner also alleviates the need of large memory capacity, since the information of previously processed patches does not need to be stored.

We demonstrate the good performance of the proposed approach as well as its robustness through several sets of numerical experiments for three representative parametric blur shapes. We also illustrate the validity of the resulting estimations, to solve image restoration problems.

The rest of the paper is organized as follows. Section II introduces the problem of spatially variant blur identification. We review the literature around this topic and discuss the construction of parametric blur models. Section III presents the state-space models for parametric PSF estimation and propose our algorithm for Bayesian inference. Section IV shows abundant experimental validation, and Section V closes the paper with some final remarks.

II. Problem Statement

Let us assume the clean image is $\mathbf{x} \in \mathbb{R}^N$ and its corresponding blurred and noisy image $\mathbf{y} \in \mathbb{R}^N$ that is corrupted by spatially variant blur and noise with given level. Each pair of image $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ are decomposed into a partition of T patches $(\mathbf{x}_t)_{t=1}^T \in \mathbb{R}^P$, $PT = N$, and $(\mathbf{y}_t)_{t=1}^T \in \mathbb{R}^P$. We assume that each patch of the observed image is given by:

$$\mathbf{y}_t = \mathbf{X}_t \mathbf{h}_t + \mathbf{n}_t. \tag{1}$$

Hereabove, $\mathbf{y}_t \in \mathbb{R}^P$ represents the blurry noisy patch, $\mathbf{h}_t \in \mathbb{R}^L$ is the blur kernel to be estimated at patch t, and $\mathbf{n}_t \in \mathbb{R}^P$ models an additive noise. Moreover, $\mathbf{X}_t \in \mathbb{R}^{P \times L}$ is a suitable matrix related to the blur of a patch t of original image $\mathbf{x}$. The model in Eq. (1) identifies with the PSF-interpolation model from [28, Eq. (22)], where the interpolation strategy depends on the choice made for $\mathbf{X}_t$ (see also [29], [24]). In all our practical experiments, for simplicity and complexity reasons, we will adopt the piecewise constant PSF model from [28, Eq.20], that is the order 0 (i.e., nearest neighbor) PSF-interpolation model. Note that several other strategies are available for modeling the space-varying blur operator, for instance in [30], [28], [31], [24], [32], with various modeling accuracy and computational costs. Our choice in Eq. (1) is particularly well suited for fast blur identification due to the separability over patches. This model was also considered in [33], [29], [34], [25]. In many practical situations, a prior over the blur shape is available. It is thus convenient to introduce a parametric model for those blur kernels, with the advantage of reducing the number of unknowns to estimate in the blur identification task. Each kernel is then assumed to read:

$$\mathbf{h}_t = h(\boldsymbol{\rho}_t), \quad t = 1, \ldots, T \tag{2}$$

where h is a known function, not depending on t, that describes the general shape of the kernels, specified by parameters $(\boldsymbol{\rho}_t)_{t=1}^T \in \mathbb{R}^K$, with $K \geq 1$ typically much lower than L. In this paper, we focus on the problem of sequential parametric blur estimation, which amounts to retrieve, for every $t = 1, \ldots, T$, an estimate for $\boldsymbol{\rho}_t$, given the knowledge of the past and present observations $(\mathbf{X}_i, \mathbf{y}_i)_{i=1}^t$ and the function $h(\cdot)$. The main underlying assumption in our approach is the smooth variation of the kernel shape parameters when progressing sequentially along the patch indexes $t = 1, \ldots, T$ (see discussion in Sec. III-B about patch ordering). Therefore, our approach is particularly well suited for space-varying blurs arising from optical aberrations (e.g., phase aberration [35], atmospheric aberration [22]), and smooth motion (e.g., camera motion blur [36]). One can also refer to the smooth varying PSF maps used as illustrative examples in [30], [28]. As described in Section III, the spatial dependence is encoded in the transition density among consecutive patches, which acts as a prior pdf. We note that, as it is often the case in Bayesian inference, vague enough priors are usually adequate particularly when enough informative observations are available. In contrast, this paper does not cover PSF maps with abrupt changes, such as those considered in [17], [15]. As we will show below, we exploit the similarity of kernels that are spatially close during the inference/estimation task. More precisely, unlike in other approaches, here the data of one patch is implicitly used to better estimate the parameter of other patches, which explains the good performance of the proposed modeling and methodology.

A. Related literature and contributions

The problem of space-varying blur identification has been widely studied in the literature of image processing. It is important to distinguish two types of approaches. First, there exists a bunch of methods, for performing image deblurring in the presence of an unknown space-varying blur degradation, thus corresponding to blind image restoration [16], [29], [23], [18] in the case of motion blur. We also refer the reader to the recent work [37] for a review on this topic. Let us also mention [24] for the case of multi-frame blind image restoration. Most of these methods are focused on the restoration task, and not on the quantitative estimation of the blur map itself. This is at the exception of [16], [23], [18], that jointly restore the image and a piecewise-constant map, obtained through a segmentation-based strategy, describing the non-stationary blur effects. Though image restoration is a problem of high interest, in certain applications such as microscopy or astronomical imaging, an accurate qualitative and quantitative knowledge of the blur effects is key for a better understanding and improving (e.g., through calibration) of the imaging device [19], [38], [39], [40]. For performing blur identification, it is necessary to make structural assumptions on the blur map to be estimated. For e.g., [15] assumes a finite set of candidate defocus blurs, to be tested in each location of the image, or [41] considers parametric blurs depending only on a single parameter. The blur estimation can also be facilitated by making use of a calibrated image (e.g., fluorescent bead in microscopy, distant star in astronomical imaging) in order to reduce the ill-posedness of the identification problem. When the PSF is stationary, and no parametric model of it is further assumed, its estimation from calibrated image acquisition can be easily

solved by a penalized least-squared algorithm (see for instance [42] for an efficient method in the case of large size images). The problem raised in this case is similar to the one arising in high dynamic range image fusion, for which wavelet-based methods have shown their efficiency [43], [44]. If, additionally, a parametric model of the PSF is available, non-linear least squares method, such as Levenberg-Macquart [13], or more recent proximity-based algorithms [12], can be employed for estimating directly the sought parameters. However, when PSF is non stationary, the accurate identification of its shift-variant evolution becomes much more challenging to resolve, even when the image is calibrated (i.e., known). This problem is typically addressed using optimization-based methods [25], [34]. We also refer the reader to [38] and references therein, for a review of the problem and recent insights, in the particular context of astronomical imaging. In the aforementioned works, no parametric model of the PSF was assumed though it is highlighted in [38] as a promising research direction. Moreover, most available techniques address the problem in a batch manner, requiring the loading of the full image before the starting of the identification process, which can be at the price of a high memory cost in high resolution imaging or even incompatible with an on-the-fly image acquisition (e.g., satellite imaging). Furthermore, up to our knowledge, no Bayesian-based techniques have been proposed so far, so that available strategies provide only point-wise estimators for the PSF field. We conclude this state-of-the-art review by emphasizing that several works exploit the smooth variability of the PSFs within the field of view, with the aim to reduce the cost of the non-stationary blur operator [28], [30] or to perform PSF field interpolation [45], [46].

In this paper, we propose a Bayesian sampling method, for the identification of parametrized space-varying blurs in the context of calibrated images [1]. Our contribution is threefold: (i) the particle filtering strategy employed for sampling leads to a sequential and fast estimation result; (ii) the proposed method is versatile and easily adapted to diverse parametric blur shapes; and (iii) the Bayesian framework allows us to provide a posterior estimation, including uncertainty quantification, for the PSF map parameters.

B. Relevant parametric PSF models

The general model of Eq. (2) enables to encompass various blur shapes. Hereafter, we present three interesting classes of parametric blur models, that will be used thereby in the experiments of the paper. To simplify the notation, let us ignore the subscript for patch t in this subsection. Without loss of generality, we consider kernels with square support parametrized by a grid with length $\sqrt{L}$ (assumed to be an odd integer), centered and regularly spaced, denoted $\mathcal{G} = \{-\frac{\sqrt{L}-1}{2}, -\frac{\sqrt{L}-1}{2}+1, \ldots, 0, 1, \ldots, \frac{\sqrt{L}-1}{2}\}$. The blur shape is determined by the values of a given function $\mathrm{h}(c_1, c_2)$ for $(c_1, c_2) \in \mathcal{G}^2$. Otherwise stated, (c_1, c_2) are the (signed) distances, in pixels, from the origin (i.e., the center position of the squared kernel) in the horizontal and vertical axis, respectively. The vector $h(\boldsymbol{\rho}) \in \mathbb{R}^L$ is then simply deduced by ordering the entries of matrix $[\mathrm{h}(c_1, c_2; \rho)]_{(c_1,c_2)\in\mathcal{G}^2} \in \mathbb{R}^{\sqrt{L}\times\sqrt{L}}$, following the lexicographic order[2]. The normalization constants and the support size for each presented blur shape are presented in Appendix A.

a) Generalized Gaussian blur: space-varying generalized Gaussian blur shapes have been employed for instance in [48] in the context of out-of-focus image deblurring. Such blur is parametrized by $\boldsymbol{\rho} = (\theta, \mathbf{s})$ with $\theta \in \mathbb{R}$ and $\mathbf{s} = (s_1, s_2) \in (0, +\infty)^2$ the orientation and width parameters, respectively. The parametric model then reads:

$$(\forall (c_1, c_2) \in \mathcal{G}^2),$$
$$\mathrm{h}(c_1, c_2; \boldsymbol{\rho}) = \lambda \exp\left(-\frac{1}{2}\frac{([c_1\ c_2]\boldsymbol{\Sigma}^{-1}(\theta, \mathbf{s})[c_1\ c_2]^\top)^p}{\alpha^p}\right), \tag{3}$$

with $p > 0$ and $\alpha > 0$ the power and the scale of generalized Gaussian model, respectively, and λ the normalization constant (see (30) in App. A). Moreover,

$$\Sigma(\theta, \mathbf{s}) = \mathbf{R}_\theta \begin{bmatrix} s_1^2 & 0 \\ 0 & s_2^2 \end{bmatrix} \mathbf{R}_\theta^\top \tag{4}$$

with $\mathbf{R}_\theta \in \mathbb{R}^{2\times 2}$ the rotation matrix defined by the orientation angle θ,

$$\mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{5}$$

We can ensure that the kernel value in its support corner is less or equal than $a\,\mathrm{h}(0, 0; \boldsymbol{\rho})$, for some $a \in (0, 1)$, as soon as $s_1 \leq s_{\max}$, $s_2 \leq s_{\max}$ with $s_{\max}$ given in (31) in App. A. Note that when $p = \alpha = 1$ in (3), we recover the common Gaussian blur model, used for instance in fluorescence microscopy in [49]. An example is displayed in Fig. 1(left).

b) Defocus blur: Defocus space-varying blurs have been considered for example in [50] in the context of depth estimation in digital camera images. Defocus blur kernel can be parametrized by $\boldsymbol{\rho} = (\theta, \mathbf{s})$ with $\theta \in \mathbb{R}$ and $\mathbf{s} \in (0, +\infty)^2$. The parametric model is,

$$(\forall (c_1, c_2) \in \mathcal{G}^2)$$
$$\mathrm{h}(c_1, c_2; \boldsymbol{\rho}) = \begin{cases} \lambda & \text{if} \quad (c_1, c_2) \in \mathbf{C}(\theta, \mathbf{s}) \\ 0 & \text{otherwise} \end{cases}, \tag{6}$$

with the normalization constant λ defined in (32) in App. A. and the elliptical domain:

$$\mathbf{C}(\theta, \mathbf{s}) = \Big\{(c_1, c_2) \in \mathcal{G}^2 \text{ such that } \frac{(c_1\cos\theta + c_2\sin\theta)^2}{s_1^2} + \frac{(c_1\sin\theta - c_2\cos\theta)^2}{s_2^2} \leq 1\Big\}. \tag{7}$$

The latter can easily be ensured to be included into the kernel support, under the condition that $s_1 \leq s_{\max}$, $s_2 \leq s_{\max}$ with $s_{\max}$ given in (33) in App. A. An example of such blur is displayed in Fig. 1(middle).

[1] A limited version of this work was presented by the authors in the conference paper [47].

[2] The lexicographic ordered vector associated to a matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is the column vector $[a\ b\ c\ d]^\top$.

c) Skew-normal blur: The skew-normal blur shape [51] was first considered in [52] to build synthetic image deblurring problems. Its ability for modeling shift-variant PSFs in real digital imaging sensors was then illustrated in [53]. This blur model depends of five positive parameters $\boldsymbol{\rho} = (s_1, s_2, w, \alpha_1, \alpha_2)$, as follows:

$$(\forall (c_1, c_2) \in \mathcal{G}^2),$$
$$\mathrm{h}(c_1, c_2; \boldsymbol{\rho}) = \lambda \exp\left(-\frac{1}{2}\begin{bmatrix}\gamma_1\\ \gamma_2\end{bmatrix}^\top \mathbf{S}^{-1}(w)\begin{bmatrix}\gamma_1\\ \gamma_2\end{bmatrix}\right) \times \int_{-\infty}^{\alpha_1\gamma_1+\alpha_2\gamma_2} \exp\left(-\frac{z^2}{2}\right)\mathrm{d}z, \quad (8)$$

where

$$\gamma_1 = \frac{c_1}{s_1}, \quad \gamma_2 = \frac{c_2}{s_2}, \quad \mathbf{S}(w) = \begin{bmatrix} 1 & w \\ w & 1 \end{bmatrix}, \quad (9)$$

and the normalization constant λ given in (34), in App. A. As shown in [51, Eq. (2.4)], parameters (w, α_1, α_2) must ensure that the matrix below is positive definite:

$$\begin{bmatrix} 1 & \frac{w-\delta_1\delta_2}{(1-\delta_1^2)^{\frac{1}{2}}(1-\delta_2^2)^{\frac{1}{2}}} \\ \frac{w-\delta_1\delta_2}{(1-\delta_1^2)^{\frac{1}{2}}(1-\delta_2^2)^{\frac{1}{2}}} & 1 \end{bmatrix}. \quad (10)$$

which amounts to satisfying the following technical conditions:

$$\delta_1\delta_2 - \sqrt{(1-\delta_1^2)(1-\delta_2^2)} < w < \delta_1\delta_2 + \sqrt{(1-\delta_1^2)(1-\delta_2^2)},$$
$$\alpha_1 = \frac{\delta_1 - \delta_2 w}{\sqrt{(1-w^2)(1-w^2-\delta_1^2-\delta_2^2+2\delta_1\delta_2 w)}}, \quad (11)$$
$$\alpha_2 = \frac{\delta_2 - \delta_1 w}{\sqrt{(1-w^2)(1-w^2-\delta_1^2-\delta_2^2+2\delta_1\delta_2 w)}},$$

for some $(\delta_1, \delta_2) \in (-1, 1)^2$. Hereagain, the maximal proportion $a \in (0, 1)$ reached on the corner of the kernel support can be imposed by setting s_1 and s_2 no greater than $s_{\max}$ given in (35), in App. A. The skew-normal blur reduces to the Gaussian blur when $\alpha_1 = \alpha_2 = 0$. An example is displayed in Fig. 1(right).

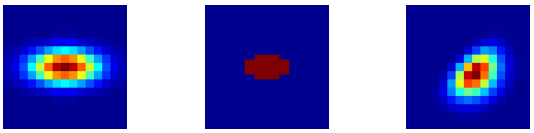


Fig. 1. (left) Gaussian blur with $\theta = 0, \mathbf{s} = [0.2; 0.1]\sqrt{L}, \alpha = 1, p = 1$ and $L = 15^2$; (middle) defocus blur with $\theta = 0, \mathbf{s} = [0.2; 0.1]\sqrt{L}$ and $L = 25^2$; (right) skew-normal blur with $\mathbf{s}_1 = 0.15\sqrt{L}, s_2 = 0.15\sqrt{L}, w = 0, \alpha_1 = 1.5, \alpha_2 = 1.5$ and $L = 15^2$.

III. Space-varying blurs modeling and inference

In this section, we first present our novel modeling approach for space-varying blur maps based on state-space models. We describe the standard bootstrap particle filter (BPF) for inference in such models. After pointing the limitations of the BPF in this context, we propose the *Generalized Interacting Annealed PF* (GIAnPF). Finally, we discuss the properties of GIAnPF and the connections with other methods of the literature.

A. State-space modeling for blur identification

We start by considering a generative state-space model (SSM), where the hidden state represents the hidden (hence unknown) parameters of the variant PSFs. This generative modeling allows for a systematic Bayesian estimation of the unknown parameters, which are considered hidden states in the SSM literature [54]. It allows to explicitly model the spatial smoothness of the PSF field. For example, the variation among neighbor patches can be a small rotation and change of width (see hereafter for a discussion regarding the order of the patches). Let us consider a state-space model given by

$$p(\boldsymbol{\rho}_0), \quad (12)$$
$$p(\boldsymbol{\rho}_t|\boldsymbol{\rho}_{t-1}), \quad (13)$$
$$p(\mathbf{y}_t|\boldsymbol{\rho}_t, \mathbf{X}_t), \quad (14)$$

for $t = 1, ..., T$, where $p(\boldsymbol{\rho}_0)$ is the prior distribution, $p(\boldsymbol{\rho}_t|\boldsymbol{\rho}_{t-1})$ is the transition model that generates the patch parameters t given the previous patch $t-1$, and $p(\mathbf{y}_t|\boldsymbol{\rho}_t, \mathbf{X}_t)$ is the observation model of patch t (that can be seen as the likelihood function when $\mathbf{y}_t$ is observed). The goal is to approximate probabilistically and sequentially the unknown parameters ρ_t for each patch t given subsets of data $\{\mathbf{X}_t, \mathbf{y}_t\}_{t=1}^T$. In particular, one can obtain the so-called filtering posterior distribution $p(\boldsymbol{\rho}_t|\mathbf{X}_{1:t}, \mathbf{y}_{1:t})$, i.e., the posterior on ρ_t conditioning on all data up to t.

B. Patch ordering for a sequential processing

The proposed approach requires setting the order in which the data will be sequentially processed. The parameters of consecutively processed patches must keep certain similarity in such a way the spatial information can be exploited. The underlying assumption in our approach is that the blurs affecting consecutive patches in the sequence have similar shape parameters. Otherwise stating, the parameters of the blurs should not change abruptly when going from patch t to patch $t-1$. Smooth PSF maps can then be identified, as soon as consecutive patches are spatially close. This is actually not the case if the patches are numbered naively following the lexicographic order, as illustrated in Figure 2(left). More suitable ordering must be adopted. If no additional structural assumption is available on the PSF map, we would suggest the zig-zag order, reminiscent from the one used in DCT-based image compression [55] (Fig. 2(middle)). Circular blur maps, as those modeling phase aberrations ([28, Sec.3.4],[35], [53],[34, Chap.4]), would be identified better by using a spiral ordering (Fig. 2(right)). Phase aberrations (and thus, the PSF width) tend to increase when getting further from the center. The spiral ordering is thus more suitable as it implicitly promotes changes of the PSF parameters depending mostly on the distance to the image center. The spiral ordering will be retained in most experiments of the paper.

The smooth variations between consecutive kernels can be easily encoded using a Gaussian distribution for $p(\boldsymbol{\rho}_t|\boldsymbol{\rho}_{t-1})$. In order to ensure that the sampled blurs do not have degenerate shapes, range constraints can furthermore be enforced on some parameters (e.g., strict positivity for the width parameter, bounded PSF support), simply by considering truncated

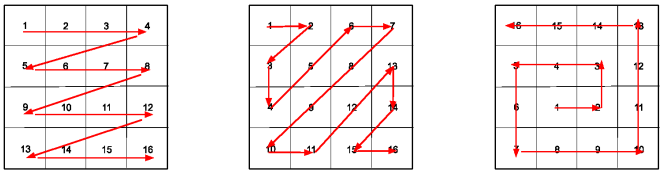


Fig. 2. Patch ordering (in the case of 4×4 structure): (left) lexicographic ; (middle) zig-zag ; (right) spiral.

Gaussian distributions (see more details in Section IV). Note that it would be also possible to learn an optimal pattern to scan/process the data. Intuitively, the modeling/inference results of our proposed method would benefit for finding an order where consecutive kernels would be as similar as possible. This could be done for instance by first running a fast (but inaccurate) optimization-based strategy to choose the ordering in which the SSM would be constructed.

C. Bayesian inference in SSMs through particle filtering

Linear-Gaussian state-space models are often used since they allow for closed-form filtering and smoothing distributions via the well-known Kalman Filter (KF) [56]. However, in this paper we focus on more complex models which broadens the flexibility and applicability of our approach. Unfortunately, in these models the targeted probability density functions (pdf) are intractable and approximations are required. Extensions of the Kalman filter exist, e.g., the extended Kalman filter (EKF) or the unscented Kalman filter (UKF) although they can greatly deviate from the true solution when the model is heavily non-linear non-Gaussian. In order to tackle these limitations, we consider an approach based on Monte Carlo approximations, and in particular on particle filtering (PF) [26].

We start by considering the *bootstrap particle filtering* (BPF) in Table I, the first and arguably the most relevant PF [26]. The BPF departs by simulating M samples (or particles) from the prior distribution. Then, for each patch t, it simulates M particles $\{\overline{\boldsymbol{\rho}}_t^{(m)}\}_{m=1}^M$ from the transition model by conditioning on the previous particles (Step 2a). Then, normalized importance weights are computed as the evaluation of the likelihood at each $\overline{\boldsymbol{\rho}}_t^{(m)}$ in such a way all weights sum up to one (Step 2b). A resampling step is performed (Step 2d) in order to avoid particle degeneracy (see more details in [57]). As a result, the BPF approximates the filtering posterior distribution $p(\boldsymbol{\rho}_t|\mathbf{X}_{1:t}, \mathbf{y}_{1:t})$ for each patch t from the set of M weighted particles $\{\overline{\boldsymbol{\rho}}_t^{(m)}, \overline{w}_t^{(m)}\}_{m=1}^M$.

Note that other existing PFs can be also used in this problem, e.g., the auxiliary PFs (APFs) [58], [59], [60] or the Rao-Blackwellized PF [27], [61], [54]. The research on PF has been very active in the last decade, and new PFs have been proposed in order to overcome existing challenges.

In the specific application we tackle here, we are facing the well-known particle degeneracy problem, appearing due to very informative observations (hence, peaky likelihood). In particular, typical suitable parametric PSF models are very sensitive w.r.t. some of their parameters and the variance of the observation noise can be relatively low. Particle degeneracy refers to the (undesirable) effect where few particles (or even only one) remain after the resampling step. This reduces the diversity of the particle approximation and endangers the estimation of the blur parameters in the following patches. In the next section, we present a new PF that addresses the shortcomings of off-the-shelf PFs in such challenging scenario.

TABLE I
BPF ALGORITHM FOR SPACE-VARYING BLUR IDENTIFICATION.

1) Initialization. Draw M i.i.d. samples, $\{\boldsymbol{\rho}_0^{(m)}\}_{m=1}^M$ from the prior $p(\boldsymbol{\rho}_0)$.
2) Filtering step. For $t = 1, ..., T$:
   a) Simulate
   $$\overline{\boldsymbol{\rho}}_t^{(m)} \sim p(\boldsymbol{\rho}_t|\boldsymbol{\rho}_{t-1}^{(m)}), \quad m = 1, ..., M \tag{15}$$
   b) Compute the normalized weights by
   $$\overline{w}_t^{(m)} \propto p(\mathbf{y}_t|\overline{\boldsymbol{\rho}}_t^{(m)}, \mathbf{X}_t), \quad m = 1, ..., M. \tag{16}$$
   c) Resample M times from $\{\overline{\boldsymbol{\rho}}_t^{(m)}\}_{m=1}^M$ with associated probabilities $\{\overline{w}_t^{(m)}\}_{m=1}^M$, i.e., for $m = 1, ..., M$, let $\boldsymbol{\rho}_t^{(m)} = \overline{\boldsymbol{\rho}}_t^{(j)}$ with probability $\overline{w}_t^{(j)}$, $j = 1, ..., M$
3) Output: The filtering step provides M weighted particles $\{\overline{\boldsymbol{\rho}}_t^{(m)}, \overline{w}_t^{(m)}\}_{m=1}^M$ at each t, for the approximation of the filtering distribution.

D. Generalized Interacting Annealed PF

In this paper we propose the novel *Generalized Interacting Annealed PF* (GIAnPF) algorithm. It implements an annealing mechanism (see for instance [62]) that overcomes the aforementioned limitations. Instead of directly approximating the targeted distribution, a sequence of modified distributions is considered. In such way, the algorithm starts by considering a more convenient distribution (i.e., sufficiently spread) while the last distribution is the true posterior of interest (in our case, the filtering distribution).

The GIAnPF algorithm is described in Table II. It starts by simulating M particles from the prior distribution in Step 1. However, unlike in the BPF or the APF, GIAnPF considers a sequence of Q intermediate steps in the processing of the data at each patch t. In the prediction step, M particles of the previous annealing layer $\{\boldsymbol{\rho}_{t,Q+1}^{(m)}\}_{m=1}^M$ are generated from the transition model by conditioning on the particles at time $t-1$ (Step 2a). Considering a sequence $0 = \beta_{Q+1} < \beta_Q < \cdots < \beta_1 < \beta_0 = 1$ and starting from $q = Q$ with a relatively small β_Q, we first compute the weights by evaluating the tempered likelihood at $\boldsymbol{\rho}_{t,q+1}^{(m)}$ (Step 2b), then each particle is resampled from the whole set of M particles with probability equal to the associated normalized weights (Step 2c). Finally, the new set of particles is simulated according to a transition kernel (Step 2d). We perform one step of Metropolis-Hastings (M-H) algorithm to sample from the transition kernel $\mathcal{T}_{t,q+1}(\cdot)$ with proposal $\mathcal{N}(\cdot, \kappa_{q+1}^2\mathbf{I})$, where κ_{q+1}^2 is the variance of the kernel [63], as it is common in sequential Monte Carlo samplers [64]. Note that in the last step, with $q = 0$, the exponent is $\beta_0 = 1$, and therefore the M-H targets the true posterior pdf of the kernel parameters, which ensures the invariance of the particle approximation.

TABLE II
GENERALIZED INTERACTING ANNEALED PF FOR SPACE-VARYING BLUR IDENTIFICATION.

1) Initialization. Draw M i.i.d. samples, $\{\boldsymbol{\rho}_{0,0}^{(m)}\}_{m=1}^{M}$ from the prior $p(\boldsymbol{\rho}_0)$. Set $Q \in \mathbb{N}$.
2) Bayesian recursion. For $t = 1, ..., T$:
   a) Propagate the particles as

$$\boldsymbol{\rho}_{t,Q+1}^{(m)} \sim p(\boldsymbol{\rho}_{t,Q+1}|\boldsymbol{\rho}_{t-1,0}^{(m)}), \quad m = 1, ..., M \tag{17}$$

   For $q = Q, Q-1, \ldots, 0$,
   b) Compute the normalized tempered weights as:

$$\xi^{(m)} \propto p(\mathbf{y}_t|\boldsymbol{\rho}_{t,q+1}^{(m)}, \mathbf{X}_t)^{\beta_q - \beta_{q+1}}, \quad m = 1, \ldots, M \tag{18}$$

   where $0 = \beta_{Q+1} < \beta_Q < \cdots < \beta_1 < \beta_0 = 1$
   c) resample

$$\overline{\boldsymbol{\rho}}_{t,q+1}^{(m)} = \boldsymbol{\rho}_{t,q+1}^{(j)}, \text{with probability } \xi^{(j)},$$

   for each $m = 1, ..., M$
   d) sample as

$$\boldsymbol{\rho}_{t,q}^{(m)} \sim \mathcal{T}_{t,q+1}(\overline{\boldsymbol{\rho}}_{t,q+1}^{(m)}), \quad m = 1, \ldots, M \tag{19}$$

3) Output: At each t, we provide a set of M unweighted particles $\{\boldsymbol{\rho}_{t,0}^{(m)}\}_{m=1}^{M}$, for the approximation of the filtering distribution.

Rationale and parameter selection. The rationale of GIAnPF is as follows. The first exponent, β_Q, is chosen in such a way the targeted distribution is a sufficiently flattened version of the likelihood. Then, for each annealing layer $q = Q, \ldots, 1$, a larger β_{q-1} than the previous one is adopted so that the likelihood used when evaluating the weights becomes increasingly closer to the original one. The final particles at this time step t, $\{\boldsymbol{\rho}_{t,0}^{(m)}\}_{m=1}^{M}$, are then those used for the approximation of the filtering distribution.

This annealing procedure is especially helpful when the likelihood varies a lot among the particles, so that only few particles would have been chosen if the exact likelihood was used to do the sampling. In Step 2b), we compute the tempered weights to deal with the particle degeneracy problem. In Step 2c), the particle $\overline{\boldsymbol{\rho}}_{t,q+1}^{(m)}$ is resampled from the set of simulated particles, $\{\boldsymbol{\rho}_{t,q+1}^{(j)}\}_{j=1}^{M}$, with associated probabilities $\{\xi^{(j)}\}_{j=1}^{M}$. In Step 2d), we apply one step of the M-H to the re-sampled particles, to ensure that the particle approximation converges to the true targeted distribution (i.e., without the tempering exponent). Note that GIAnPF can be seen as a generalized version of the BPF, since it contains this algorithm as particular case when we set $Q = 0$ and the M-H iteration (Step 2d) is avoided, i.e., with $\kappa_{q+1} = 0$.

Connections to the literature. The GIAnPF connects with several algorithms in the literature. As stated above, it can be seen as a generalization of the BPF [26]. It also holds clear links with annealing schemes such as [62], [65], [66], [67]. In the context of PF, [48] conducts interacting MCMC sampling procedure with the particles obtained from the BPF as the starting point. In [68], the authors incorporate the annealing strategy within a Rao-Blackwellised Particle Filter and update the sample size using Kullback-Leibler Divergence transformation. The algorithm proposed in [69] also bears some similarities with GIAnPF in its use of an annealing scheme in a PF. However, unlike in [69], our proposed method keeps closer ties to the standard BPF by modifying the sampling/weighting and adding an M-H update, which allows the particle approximation to converge to the filtering distribution with M (for any value of Q), and hence constructing consistent estimators of the sought intractable integrals.

IV. EXPERIMENTAL RESULTS

A. Construction of an experimental database

We first describe the experimental database. It is composed of a set of several pairs of clean/degraded images associated to various choices for the patch decomposition, shift variant blur maps, and noise levels.

1) Images: We consider three images of size $N = 512 \times 512$ (displayed in Fig. 4), namely, `Chart`, `Cells`, and `Hubble`. These are representative images of calibrated objects (e.g., resolution charts, fluorescent beads and distant stars) typically used to estimate spatially variant blur parameters. In all our experiments, patches of size 64×64 are used and the blurry noisy version $\mathbf{y}$ is generated following Eqs. (1)-(2), using i.i.d. zero-mean Gaussian noise with standard deviation $\sigma_n > 0$, so that $p(\mathbf{y}_t|\boldsymbol{\rho}_t) = \mathcal{N}(\mathbf{X}_t h(\boldsymbol{\rho}_t), \sigma_n^2 \mathbf{I})$.

2) Shift-variant blur maps: Five realistic space-varying blur maps, associated to the three blur shapes presented in Sec. II-B, are constructed. To this aim, we take inspiration from the examples provided in [30, Fig.8], [6, Fig.7], [6, Fig.7] and [53, Fig.4]. In most of our models, we will consider the existence of an optical center for the device, and that the blur width/orientation depends on the relative position of the patch with respect to this center. Instead of considering deterministic maps, as in the aforementioned works, we include some randomized settings in the expressions so as to simulate more realistic scenarios of an optical device whose effect on the image may slightly vary with respect to external conditions (temperature, laser power, planarity of the disposal). For a pre-specified abscissa-ordinate system in the 2D image plan, we denote, for a given patch t, $\mathbf{o}_t = (o_{1,t}, o_{2,t})$ the coordinates of the patch center, $\hat{\mathbf{o}} = (\hat{o}_1, \hat{o}_2)$ the coordinates of the chosen optical center of the device, and $\ell_t = \|\mathbf{o}_t - \hat{\mathbf{o}}\|$ the Euclidian distance between both. The associated parameters, gathered in vector $\boldsymbol{\rho}_t$, are defined as follows, for the five considered models. Figure 3 displays examples of three realizations of each of those, illustrating the variability among patches, and also among different realizations for the same model.

a) Map 1: We assume Gaussian blurs, with size $L = 15^2$, with width and orientation parameters $\mathbf{s}_t$ and θ_t at each patch t. We opt for the spiral patch ordering and we set, for every t,

$$\theta_t = \arctan\left(\frac{o_{1,t} - \hat{o}_1}{o_{2,t} - \hat{o}_2}\right) + \epsilon_t, \tag{20}$$

where $\epsilon_t \sim \mathcal{N}(\frac{\pi}{2}, \sigma_\epsilon^2)$,

$$s_{1,t} = \delta_{1,t}\ell_t, \quad s_{2,t} = \delta_{2,t}\ell_t. \tag{21}$$

where $\delta_{1,t} \sim \mathcal{U}([\delta_{1,\min}, \delta_{1,\max}]), \delta_{2,t} \sim \mathcal{U}([\delta_{2,\min}, \delta_{2,\max}])$ with $(\sigma_\epsilon, \delta_{1,\min}, \delta_{1,\max}, \delta_{2,\min}, \delta_{2,\max})$ positive scalars.

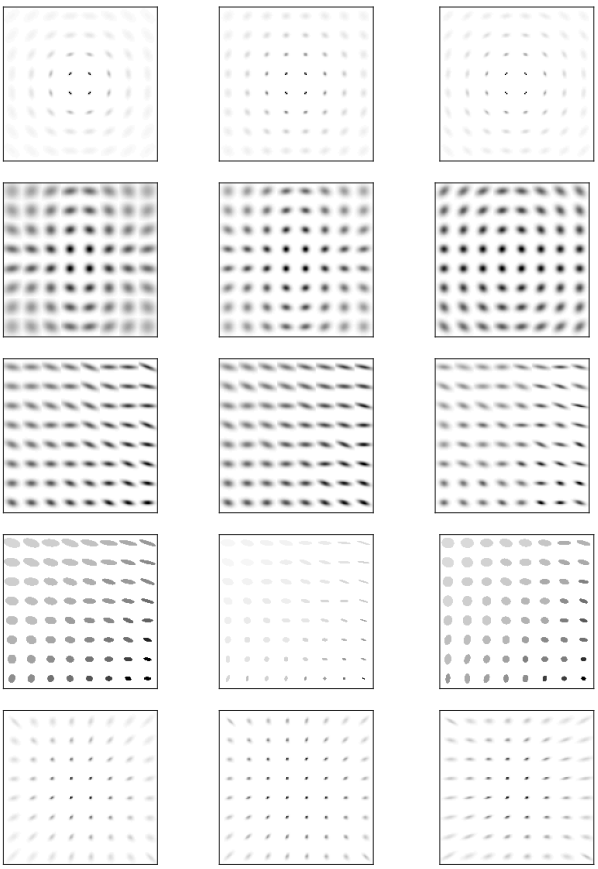


Fig. 3. Three realizations (from left to right) of the space-varying blur Maps 1 to 5 (from top to bottom). Each column displays one realization from the generative model. One can notice that, for each given row (i.e., blur map), the three images displayed share the same global aspect (e.g., circular map in rows 1, 2 or 5) with some slight variability, for instance in the blur kernel widths and orientations.

b) Map 2: We consider Gaussian blurs with size $L = 15^2$ with width and orientation parameters $\mathbf{s}_t$ and θ_t at each patch t. We use the spiral patch ordering and, for every t, we set

$$\theta_t = \arctan\left(\frac{o_{1,t} - \hat{o}_1}{o_{2,t} - \hat{o}_2}\right) + \epsilon_t, \tag{22}$$

where $\epsilon_t \sim \mathcal{N}(\frac{\pi}{2}, \sigma_\epsilon^2)$,

$$\begin{cases} s_{1,t} = \frac{1}{8}(\delta_{1,t}|o_{1,t} - \hat{o}_1| + 1) \\ s_{2,t} = \frac{1}{8}(\delta_{2,t}|o_{2,t} - \hat{o}_2| + 1) \end{cases} \tag{23}$$

with $\delta_{1,t} \sim \mathcal{U}([\delta_{1,\min}, \delta_{1,\max}]), \delta_{2,t} \sim \mathcal{U}([\delta_{2,\min}, \delta_{2,\max}])$ and $(\sigma_\epsilon, \delta_{1,\min}, \delta_{1,\max}, \delta_{2,\min}, \delta_{2,\max})$ some positive scalars.

c) Map 3: We consider Gaussian blurs with size $L = 15^2$ with width and orientation parameters $\mathbf{s}_t$ and θ_t at each patch t. The patches are ordered in lexicographic order and, for every t, we set

$$\theta_t \sim \mathcal{U}([0, \frac{\pi}{8}]), \tag{24}$$

and

$$\begin{cases} s_{1,t} = \frac{1}{8}(\delta_{1,t} - o_{1,t}), \\ s_{2,t} = \frac{1}{15}(\delta_{2,t} - o_{2,t}), \end{cases} \tag{25}$$

where $\delta_{1,t} \sim \mathcal{U}([\delta_{1,\min}, \delta_{1,\max}]), \delta_{2,t} \sim \mathcal{U}([\delta_{2,\min}, \delta_{2,\max}])$, and $(\delta_{1,\min}, \delta_{1,\max}, \delta_{2,\min}, \delta_{2,\max})$ are some positive scalars.

d) Map 4: We assume blurs with defocus shape, with size $L = 25^2$ parametrized by width and orientation parameters $\mathbf{s}_t$ and θ_t at each patch t. We order the patches in lexicographic order and, for every t, we set

$$\theta_t \sim \mathcal{U}([0, \frac{\pi}{8}]), \tag{26}$$

and

$$\begin{cases} s_{1,t} = \frac{1}{3}(\delta_{1,t} - o_{1,t}), \\ s_{2,t} = \frac{1}{6}(\delta_{2,t} - o_{2,t}), \end{cases} \tag{27}$$

with $\delta_{1,t} \sim \mathcal{U}([\delta_{1,\min}, \delta_{1,\max}]), \delta_{2,t} \sim \mathcal{U}([\delta_{2,\min}, \delta_{2,\max}])$, and $(\delta_{1,\min}, \delta_{1,\max}, \delta_{2,\min}, \delta_{2,\max})$ some positive scalars.

e) Map 5: We specify here a generative model for the particular case of the skew-normal blur, with size $L = 15^2$, with the aim to mimic the shift-variant blur map that was considered in the experiments of [53]. We order the patches in spiral order and we set

$$\omega_t = \text{sign}\left(\frac{o_{1,t} - \hat{o}_1}{o_{2,t} - \hat{o}_2}\right)\left(1 - 2\big||o_{1,t} - \hat{o}_1| - |o_{2,t} - \hat{o}_2|\big|\right). \tag{28}$$

We then create a shift-variant map with skew normal shapes of parameters

$$\begin{cases} s_{1,t} = \frac{1}{3}(|o_{2,t} - \hat{o}_2| + \frac{1}{2})(\ell_t + \frac{1}{5}) + \delta_{1,t}, \\ s_{2,t} = \frac{1}{3}(|o_{1,t} - \hat{o}_1| + \frac{1}{2})(\ell_t + \frac{1}{5}) + \delta_{2,t}, \\ w_t = \frac{7}{5}\omega_t\ell_t + \delta_{3,t}, \\ \alpha_{1,t} = \delta_{4,t}(1 - \ell_t), \\ \alpha_{2,t} = \delta_{5,t}(1 - \ell_t), \end{cases} \tag{29}$$

with $\delta_{1,t} \sim \mathcal{U}([0, \delta_{1,\max}]), \delta_{2,t} \sim \mathcal{U}([0, \delta_{2,\max}]), \delta_{3,t} \sim \mathcal{U}([0, \delta_{3,\max}]), \delta_{4,t} \sim \mathcal{U}([\delta_{4,\min}, \delta_{4,\max}]), \delta_{5,t} \sim \mathcal{U}([\delta_{5,\min}, \delta_{5,\max}])$, where $(\delta_{1,\max}, \delta_{2,\max}, \delta_{3,\max}, \delta_{4,\min}, \delta_{4,\max}, \delta_{5,\min}, \delta_{5,\max})$ are positive scalars.

We provide in Table VI in Appendix B the parameter values retained for these five realistic maps.

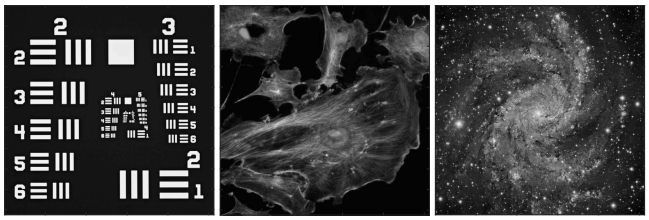


Fig. 4. Test images `Chart`, `Cells` and `Hubble`.

3) Quantitative comparison methodology: We will run several experiments of blur map estimation, using BPF and GIAnPF methods, both on toy examples and on our realistic dataset. The transition models considered in those schemes for describing the spatial blur evolution are those given in Tab. VII in Appendix B. More precisely, Model A will be employed to retrieve Gaussian blur maps (e.g., those resulting from Maps 1, 2 and 3), Model B to retrieve defocus blur maps (e.g., in case of Map 4) and Model C to estimate skew normal blur maps (e.g., Map 5). The upper bounds $s_{\max}$ defined in App. A are

used, to preserve consistency of the estimated shapes with the considered blur support width L. An analysis of robustness to the setting for L will be presented in Sec. IV-D4. In order to quantify the gain of exploiting the spatial structure of the problem, we will compare the filtering pdf approximated by BPF and GIAnPF algorithms, and the posterior distribution that only considers the data corresponding to each patch independently. In the latter case, we use importance sampling (IS) [70] to approximate the intractable posterior. Note that both PF and IS have a comparable computational complexity per sample/particle, while PF allows to exploit the information of previously processed data (i.e., patches), due to an implicit sequential IS structure. We also perform comparisons with two optimization-based methods considering non-parametric models for the kernels. In the so-called NP formulation, the kernels $(\mathbf{h}_t)_{t=1}^T$ are estimated by minimizing a least-squares function under simplex and smoothness constraints, in a fully parallel manner for each patch t. We also compared with a more sophisticated non-parametric formulation where we included, in addition to the NP cost function, the total-variation based spatial regularization among kernels of neighboring patches from [34, Chap.4], yielding NP$^+$ method. Note that both resulting constrained convex minimization problems are solved with the FISTA algorithm [71]. In all experiments, the relative mean squared error (RMSE$_h$) averaged over patches is used to evaluate the numerical performance of the blur estimation, i.e., $\text{RMSE}_h = \frac{1}{T}\sum_{t=1}^{T} \frac{||\overline{\mathbf{h}}_t - \mathbf{h}_t||_2}{||\mathbf{h}_t||_2}$ with $\mathbf{h}_t$ and $\overline{\mathbf{h}}_t$ the original kernels and the estimators respectively at patch t. For methods providing a posterior estimation (i.e., BPF, GIAnPF and IS), we calculate $\overline{\mathbf{h}}_t = h(\overline{\boldsymbol{\rho}}_t)$ with $\overline{\boldsymbol{\rho}}_t$ the mean estimator of the unknown parameters, while for NP and NP$^+$, $\overline{\mathbf{h}}_t$ is directly the solution to the optimization problem, in patch t. We also include the standard deviation as well as the 95% credible interval of the estimators of BPF, GIAnPF and IS methods, to evaluate the statistical accuracy of the results. All the presented results are averaged over 100 random runs, and all the hyperparameters are tuned so as to minimize the RMSE (averaged over 5 random trials). The best results will be marked in bold cases in the tables. The numerical experiments are conducted in a Matlab environment on a computer with an Xeon(R) W-2135 processor (3.7 GHz clock frequency) and 12 GB of RAM.

B. Validation of the proposed method

We first discuss the settings and properties of the proposed method in an illustrative toy synthetic example. Our tests are conducted on the image `Chart` decomposed into $T = 64$ patches, and corrupted by skew-normal blurs whose parameters $(\boldsymbol{\rho}_t)_{t=1}^T$ are generated following the transition model C, with $(\sigma_s, \sigma_w, \sigma_\alpha) = (10^{-1}, 2\cdot 10^{-1}, 10^{-1})$ and, except otherwise stated, a noise standard deviation $\sigma_n = 0.05$. The blur parameter estimations are then conducted with BPF and GIAnPF, assuming transition Model C with known hyperparameters.

a) Setting particle and layer numbers: The BPF method, as well as the proposed GIAnPF method, requires the setting of the number of particles. For GIAnPF, the number of layers Q plays also a role. This is necessary to reach a compromise between a precise target reconstruction and a reasonable computational time, for setting those parameters. We display in Fig. 5 the averaged RMSE on 100 runs with respect to the computational time, for different settings of the parameters of BPF and GIAnPF. The associated values for the parameters $(\beta_q)_{0\leq q\leq Q}$ are chosen on a logarithmic grid between 1 and 10^{-2}, and $\beta_{Q+1} = 0$. For a given number of particles M, the GIAnPF algorithm always generates a better approximation of the unknown blurs than BPF, although at the expense of a slightly increased. Note that, for the same time budget, GIAnPF outperforms BPF. In all the upcoming experiments, we will set $M = 3000$ for BPF, IS, and $Q = 2$ and $M = 500$ for GIAnPF, so that they have comparable time requirement.

b) Alleviating the particle degeneracy: One of the advantages of the novel GIAnPF method is the promotion of the diversity among the particles. In Fig. 6, we display the normalized effective sample size (NESS), computed as $\text{NESS} = \frac{1}{M\sum_{m=1}^{M}\overline{w}_t^{(m)}}$, where $\overline{w}_t^{(m)}$ are the normalized weights (i.e., $0 \leq \text{NESS} \leq 1$). We average the ESS over 100 runs and all the patches, divided by the corresponding number of particles M, for various level of noise σ_n. The results are also averaged over 10 independent generations of the data. It can be seen that GIAnPF algorithm has higher NESS, thus preserves much more diversity in the particles for all analyzed noise levels. Similarly, Fig. 7 displays, for the same experiment, the RMSE of the blur estimate using the mean of the approximate posterior of each filter. We can see that a larger diversity (larger NESS) translates into a smaller error.

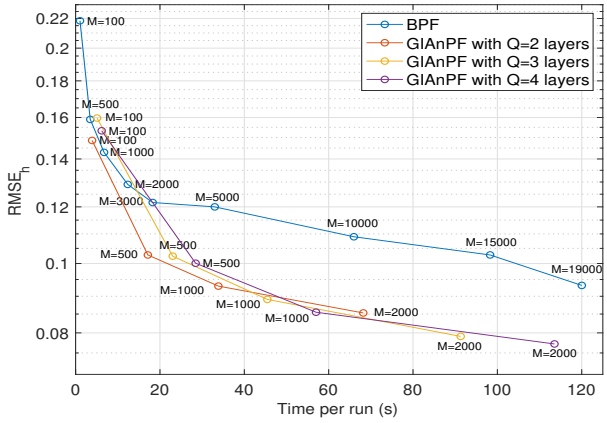


Fig. 5. RMSE$_h$ with respect to computational time for BPF and GIAnPF, using different settings of (M, Q).

C. Comparative performance on the experimental database

In this section, we compare the performance of GIAnPF and other approaches in different scenarios arising from our experimental database.

a) Quantitative performance: Table IV displays the RMSE$_h$ obtained when estimating kernels of Maps 1 to 5, from observations of `Cells` image, degraded by two noise

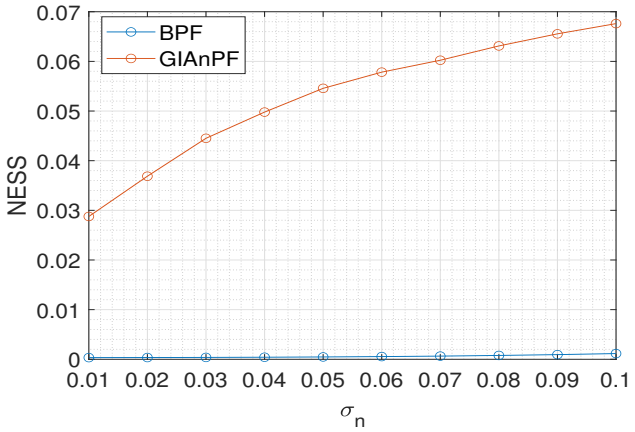


Fig. 6. Normalized effective sample size (NESS) for BPF (blue) and GIAnPF (red) for various noise level σ_n of the blurry noisy image.

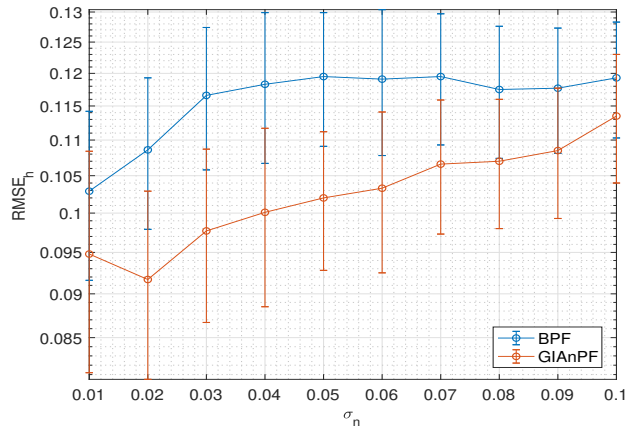


Fig. 7. RMSE in the blur estimation for BPF (blue) and GIAnPF (red) for various noise level σ_n of the blurry noisy image.

levels, namely $\sigma_n = 0.01$ and $\sigma_n = 0.05$. Illustrative examples of results are also provided in Figure 8. In all the Bayesian methods, we display the blur maps resulting from kernel parameters equal to the minimum MSE (MMSE) estimator, i.e., the mean of the posterior distribution given the available data. The methods IS and NP, which do not exploit the spatial smoothness among neighboring kernels, reach the worst performance. A visual inspection of Figure 8 confirms this result. Among the spatially regularized methods, i.e., NP+, BPF and GIAnPF, the former is still far from reaching the quality of estimation of two latter methods though it improves over NP. Such behavior is expected as BPF and GIAnPF incorporate explicit knowledge about the parametric shape of the blur, leading to less diversity and thus less error in the restored blur maps. This can be seen in Figure 8. An analysis of the sensitivity of the methods over an error in the assumed parametric shape will be discussed in Section IV-D1. In all cases presented in Table IV, the proposed method GIAnPF performs similarly or better than the standard BPF. The benefits from the annealing procedure are particularly noticeable in the case of $\sigma_n = 0.01$. This improved performance can be explained by the annealed approach of GIAnPF, which is particularly effective when the posterior pdf of the parameters has the probability mass concentrated in small regions of the space, as it happens with low values of σ_n (due to a *peaky* likelihood).

Image size	$N = 256^2$	$N = 512^2$	$N = 1024^2$	$N = 2048^2$
Patches	$T = 16$	$T = 64$	$T = 256$	$T = 1024$
NP	36.16 (2.26)	135.5 (2.11)	4987 (19.48)	82273 (80.34)
NP$^+$	35.80	127.3	4473	71913
IS	5.21 (0.32)	21.85 (0.34)	91.29 (0.36)	374.7 (0.37)
BPF	5.09	20.15	85.53	345.4
GIAnPF	4.67	19.03	78.81	329.9

TABLE III
AVERAGED COMPUTATIONAL TIME IN SECONDS, FOR THE DIFFERENT METHODS, IN THE CASE OF MAP 5 (I.E., $L = 225$ AND $K = 5$) APPLIED TO VARIOUS IMAGES OF SIZE N, DECOMPOSED INTO T PATCHES. TIME PER PATCH IS ALSO REPORTED FOR IS AND NP, AS THEY PROCESS PATCHES INDEPENDENTLY.

b) Complexity comparison: We perform a scalability analysis of the different methods, by applying those to the estimation of Map 5 blur map, from resized versions of the large scale aerial image from [42]. The same hyperparameter settings as in Table IV were used. In Table III, we report the averaged computational time over 10 random trials. For IS and NP, that are fully separable methods onto patches, we also report the time per patch, that could be reached using parallel implementation on T cores. Both of these methods benefit from fast complexity cost assuming parallel processing of the patches. It is worthy to notice that the cost per patch of NP still increases for the largest images, probably due to memory saturation issues, while IS has a constant time requirement per patch. However, despite their rapidity, let us recall that the qualitative results for both these methods were rather poor on our experiments. The complexity burden dramatically increases for the regularized optimization-based method NP$^+$. In the latter, the data is processed in a batch way. Parallelization over patches for this method is not straightforward, up to our knowledge, due to the non-separable structure of the underlying optimization problem. In contrast, the complexity of the proposed method stays very reasonable even for the largest image example, though involving the probabilistic estimation of $T \times K = 5120$ parameters. As in the previous experiment, BPF and GIAnPF again show a similar computational cost.

c) Stability to the noise level: We display in Fig. 9 a comparison of the RMSE$_h$ reached by the methods, when estimating kernels from observations degraded by increasing noise level σ_n. One can observe that NP and NP+ performance rapidly deteriorates when the noise increases. In contrast, IS, BPF and GIAnPF perform in a relatively stable way. Moreover, the proposed GIAnPF method again outperforms the other competitors. As expected, its superiority over the standard BPF is less visible as σ_n becomes higher, since GIAnPF is particularly dedicated to problems with peaky likelihoods, which arise when the noise level is low.

D. Robustness analysis

In practical contexts, it is frequent that the assumed parametric PSF model is erroneous (e.g., over-simplified with

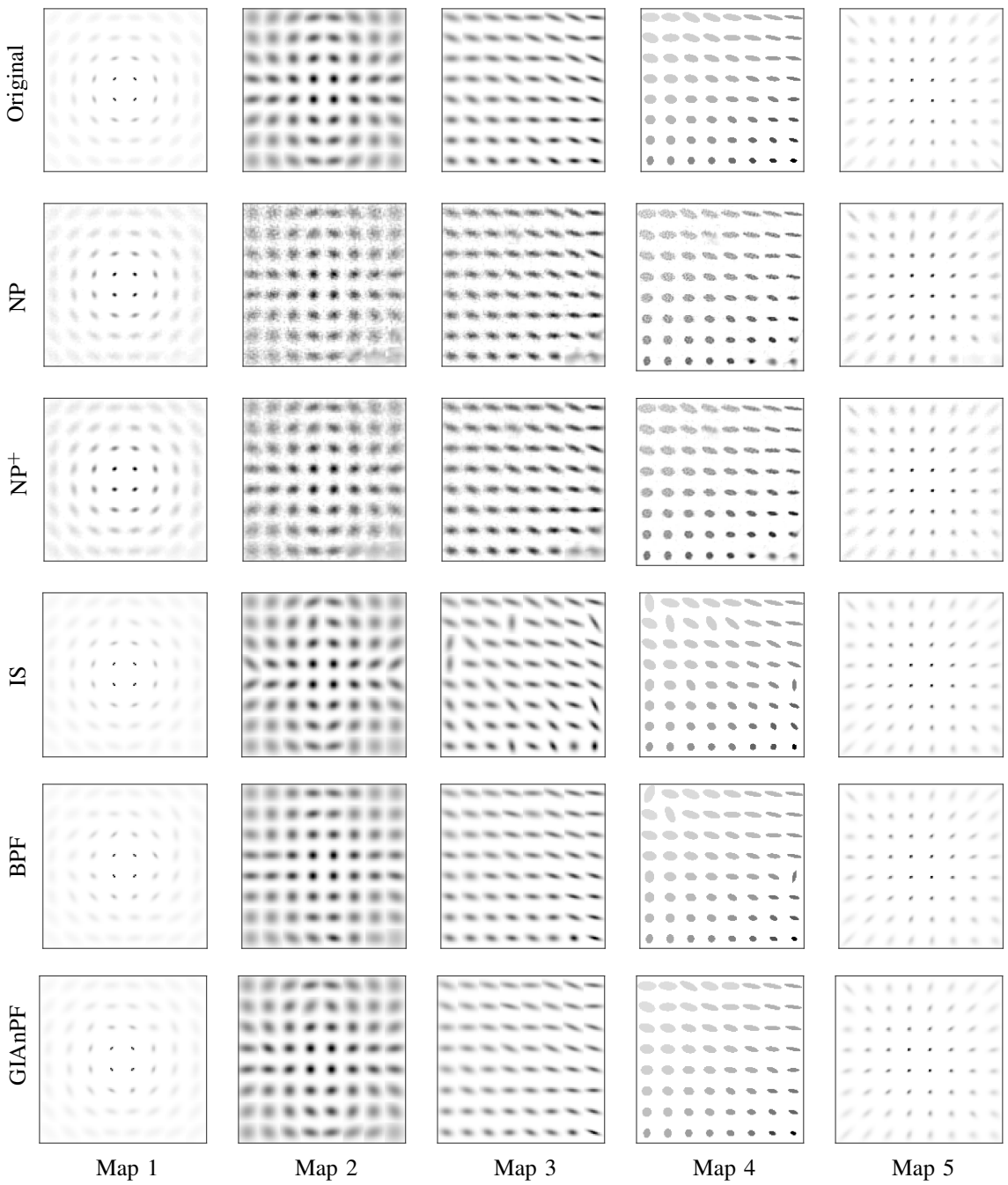


Fig. 8. Original kernels and identification results using `Cells` image with $T = 64$ patches and $\sigma_n = 0.01$. From top to bottom : original kernels, restored kernels with NP, NP$^+$, IS, BPF, GIAnPF.

wrong support size) and/or that the calibrated image is only partially known (e.g., in digital camera imaging, a noisy non-blurry version can be acquired using short exposure settings). Furthermore, the noise level in the blurred image is usually estimated and not known with perfect accuracy. It is thus of main importance to quantify the robustness of our method to such model mismatch. To this aim, we conducted four experiments, aiming at reproducing different realistic scenarios of imperfect knowledge of the acquisition model. Except if specified otherwise, in all the examples of this section, we used the image `Cells`. The degraded image is obtained by applying the space-varying Gaussian blur Map 1, and then adding a noise with level equals to $\sigma_n = 0.05$.

1) Choice of the parametric blur model: We first evaluated the influence of a mismatch between the assumed blur shape, and its actual one. We ran the different methods assuming instead a generalized Gaussian blur shape with different powers p, where $p = 1$ corresponds to the ground truth. The retrieved average RMSE$_h$ of BPF, GIAnPF, IS when taking different values for p are displayed in Fig. 10. We also displayed the results of NP and NP$^+$ approaches, that remain unchanged since no parametric model for the kernels is considered in those methods. As expected, the best performance are obtained for the correct $p = 1$ setting. More interestingly, we can

		Noise level $\sigma_n = 0.01$	Noise level $\sigma_n = 0.05$
Map 1	NP	0.2117	0.3209
	NP$^+$	0.1756	0.2543
	IS	0.1258 (0.0129) 0.1057-0.1545	0.3339 (0.0157) 0.3056-0.3620
	BPF	0.1246 (0.0080) 0.1109-0.1446	0.1964 (0.0204) 0.1696-0.2646
	GIAnPF	**0.0680** (0.0065) 0.0578-0.0837	**0.1564** (0.0122) 0.1347-0.1858
Map 2	NP	0.2474	0.2381
	NP$^+$	0.1383	0.1385
	IS	0.1653 (0.0072) 0.1471-0.1768	0.1653 (0.0072) 0.1471-0.1768
	BPF	0.0917 (0.0093) 0.0812-0.1211	0.1246 (0.0143) 0.1026-0.1661
	GIAnPF	**0.0452** (0.0044) 0.0375-0.0549	**0.1169** (0.0077) 0.1026-0.1307
Map 3	NP	0.1905	0.2723
	NP$^+$	0.1480	0.1608
	IS	0.1380 (0.0128) 0.1143-0.1626	0.4089 (0.0194) 0.3717-0.4503
	BPF	0.0968 (0.0049) 0.0897-0.1063	0.1854 (0.0550) 0.1145-0.3611
	GIAnPF	**0.0396** (0.0034) 0.0345-0.0486	**0.1288** (0.0114) 0.1116-0.1544
Map 4	NP	0.3263	0.4286
	NP$^+$	0.2988	0.3772
	IS	0.2746 (0.0187) 0.2443-0.3185	0.5554 (0.0204) 0.5208-0.6021
	BPF	0.1360 (0.0098) 0.1174-0.1584	0.2847 (0.0236) 0.2459-0.3411
	GIAnPF	**0.0701** (0.0089) 0.0521-0.0860	**0.2640** (0.0100) 0.2449-0.2833
Map 5	NP	0.1830	0.3578
	NP$^+$	0.1415	0.2428
	IS	0.1331 (0.0066) 0.1198-0.1469	0.1852 (0.0079) 0.1247-0.1407
	BPF	0.0910 (0.0067) 0.0799-0.1066	0.1332 (0.0043) 0.1247-0.1407
	GIAnPF	**0.0688** (0.0075) 0.0587-0.0880	**0.1322** (0.0061) 0.1214-0.1470

TABLE IV
RMSE$_h$ FOR NP, NP$^+$, IS, BPF, GIANPF. CELLS IMAGE BLURRED BY SPACE-VARYING KERNELS GENERATED FOLLOWING MAPS 1-5.

observe that, even if a wrong parametric model is adopted, i.e., using $p \neq 1$, our proposed method GIAnPF still performs better than its competitors, and in particular the non parametric method NP$^+$, for a wide range of values for p.

2) Noisy calibrated image: In the previous experiments, we assumed that the original image $\mathbf{x}$ is known in a perfect manner. However, in reality, it is commonly not the case. Here, we consider the situation when only a noisy version of $\mathbf{x}$ is provided in the estimation process, denoted $\tilde{\mathbf{x}} = \mathbf{x} + \mathbf{w}$ where $\mathbf{w}$ is an additive Gaussian noise follows $\mathcal{N}(0, \sigma_x^2 \mathbf{I})$. Eleven different levels for the noise on $\mathbf{x}$ are considered, namely $\sigma_x = \{0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10\}$. Note that this range of values covers from low noise level values (where traditionally the particle-based filters struggle

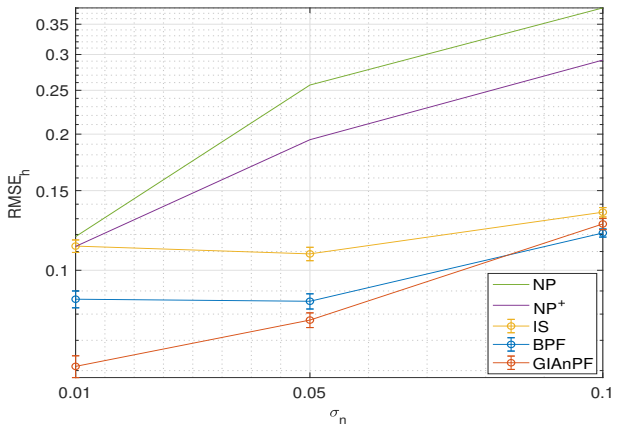


Fig. 9. Mean/variance of the RMSE$_h$ reached by the different methods. `Hubble` image blurred by space-varying kernels generated following Map 5 and three different values for σ_n.

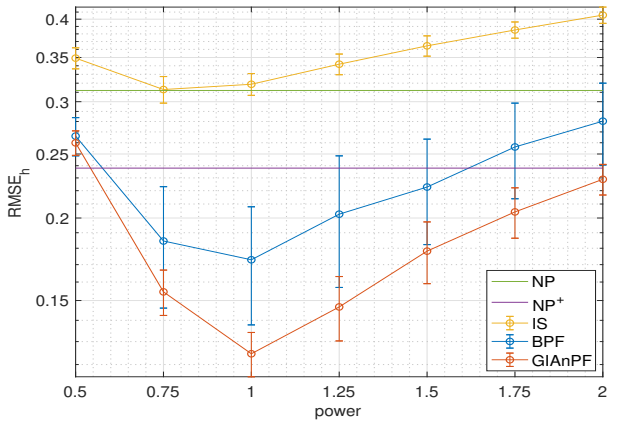


Fig. 10. Mean/variance of the RMSE$_h$ with the different methods assuming different powers within the generalized Gaussian blur shape (the ground truth is $p = 1$).

to operate due to the mass being concentrated in small regions; see for instance [64]) to relatively large noise values (which makes the inference less accurate). We display in Fig. 11 the averaged RMSE over 10 noisy realizations for each σ_x, for the different approaches. The improvement of our proposed method GIAnPF with respect to BPF is more significant at low noise level, since particle degeneracy mostly arises in such context. Moreover, Fig. 11 shows that GIAnPF still outperforms its competitors when the noise level on $\mathbf{x}$ increases.

3) Setting of noise level: In the previous analysis, the exact observation noise variance σ_n was assumed to be known and applied explicitly in our blur identification method. In more realistic situations, such noise level would have to be learned, leading to an estimate $\hat{\sigma}_n$. We ran the different blur identification methods for the five different values $\hat{\sigma}_n = \{0.01, 0.025, 0.05, 0.07, 0.1\}$ (recall that the true observation noise level here is $\sigma_n = 0.05$). Note that the optimization-based methods are not sensitive to such change, so that the results for NP and NP$^+$ are unchanged. We display in Fig. 12 the average RMSE, for the different values for $\hat{\sigma}_n$. We can notice that GIAnPF gives the best performance even when the noise level is poorly estimated. As expected, it reaches minimal mean and variance values when the ground truth noise

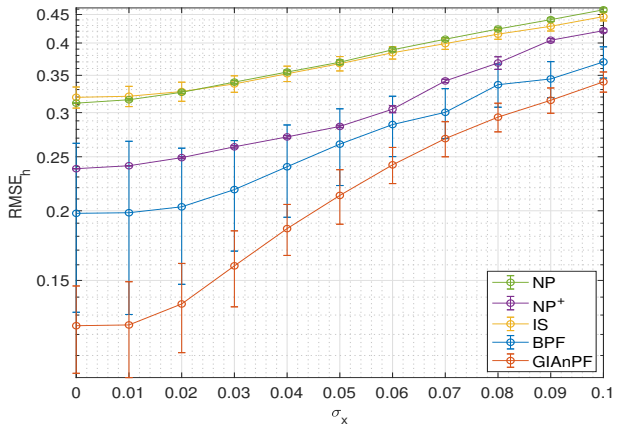


Fig. 11. Mean/variance of the RMSE_h reached by the different methods, for various noise level σ_x deteriorating the input calibrated image.

level is used in the estimation process.

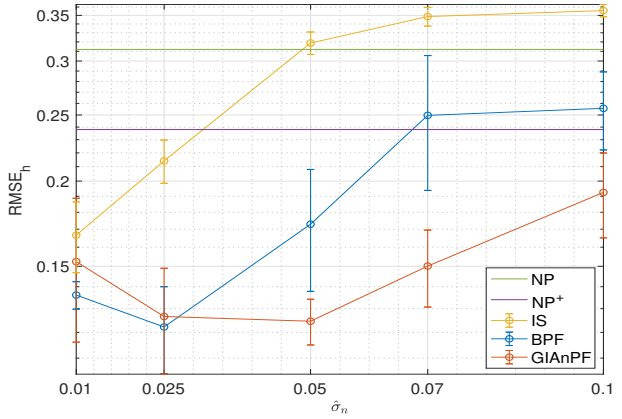


Fig. 12. Mean/variance of the RMSE_h reached by the different methods, when using an erroneous noise level estimate $\hat{\sigma}_n$ (the ground truth is $\sigma_n = 0.05$).

4) Setting of kernel size: In blur estimation, especially when the blur is non-stationary, a challenge lies in the setting a priori, of the kernel support width L. In this experiment we propose to assess the different approaches in the case where an erroneous L is assumed. In particular, we estimate the blur from the image `Cells`, when it is degraded by Map 1, and a noise level $\sigma_n = 0.05$. We run the different methods/models, by setting the size for the kernel width as $L \in \{5, 7, 9, 11, 13, 15, 19, 23, 27\}$, while the groundtruth is $\overline{L} = 15$. All the true/estimated kernels have been extended to the largest tested size $L = 27$ by zero-padding, to allow valid computation of estimation errors. We display in Fig. 13 the obtained RMSE values. One can notice that all methods have stable estimation error, as long as the assumed L is greater or equal to $\overline{L}$, which is indeed expected. In contrast, when $L < \overline{L}$, the results deteriorate for all methods. We observe that the GIAnPF method outperforms its competitors in almost all tested values for L.

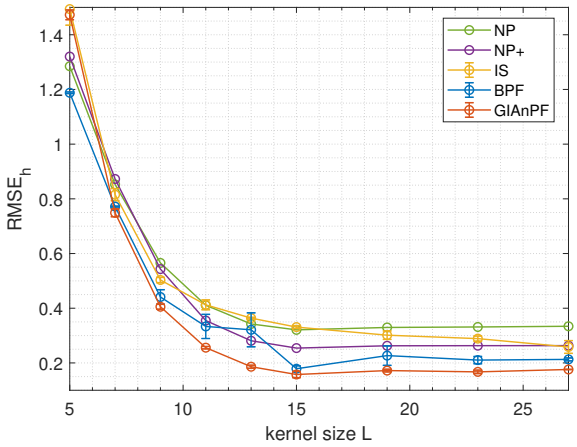


Fig. 13. Mean/variance of the RMSE for various kernel support width L (ground truth size: $\overline{L} = 15$).

E. Image restoration

We conclude this experimental section by illustrating the ability of the obtained kernel estimates to lead to satisfying and interpretable image restoration results. We assume that we have access to a pair $(\mathbf{x}, \mathbf{y})$ of clean/degraded version of a given calibrated image. We focus on the estimation of blurs resulting from our synthetic blurred maps, and then the use of those results for the restoration of several non-calibrated images. Through this example, we also illustrate and discuss the advantage of providing probabilistic blur estimates.

We consider the image `Chart`, and its degraded version corrupted by the space-varying Map 4 with $T = 64$ and noise with standard deviation $\sigma_n = 0.1$. We ran GIAnPF method and NP$^+$ approaches, identified as the two best blur identification procedures in the previous section. As shown in Fig. 14(left), we then perform the restoration of a set of four natural images[3], from their degraded version in Fig. 14(right), following the same blur model, and noise level $\sigma_x = 0$, 0.01, 0.05 and 0.1. The restoration is performed in a non-blind manner, by making use of the fast majorize-minimize memory-gradient (3MG) algorithm from [72], that was recently adapted to the problem of image retrieval under shift-varying blur [73], [74]. Note that the variational formulation adopted in 3MG is very similar to the one in [28]. The regularization parameters of 3MG are finetuned so as to maximize $\text{SNR}_x = 20\log_{10}\left(\frac{||\mathbf{x}||}{||\mathbf{x}-\hat{\mathbf{x}}||}\right)$, with $\mathbf{x}$ the original image, and $\hat{\mathbf{x}}$ the restored one. We present the results in Table V, in terms of SNR_x between $\mathbf{x}$ and the restored images obtained with the blur estimates of either GIAnPF or NP$^+$. In the case of GIAnPF, we ran the restoration with 100 samples from the estimated posterior distribution of the blur maps, for a single run of GIAnPF, which allows us to compute the mean and the 95% confidence intervals for both image quality metrics. In Fig. 15(left), we show the mean of the pixel-wise squared error between the true kernel and these 100 samples obtained from the posterior distribution. This metric, called Bayesian MSE, allows to assess the uncertainty

[3]http://sipi.usc.edu/database/

of GIAnPF on each kernel estimate. One can identify patches with larger estimation errors, corresponding to flat zones in the `Chart` images, thus making the blur estimation less accurate. In contrast, the patches localized in zones with significant content (sharp edges, in particular) benefit from very good estimation quality for the associated blur (pixel-wise error lower than 10^{-6}). Standard deviation is shown in Fig. 15(right), illustrating that the estimated distribution is more spread for difficult patches, while it is rather peaky for the patches that are better estimated. One can see from Table V that GIAnPF outperforms NP$^+$ in almost all examples, with up to 0.4 improvement in SNR score. Examples of visual results obtained from a given GIAnPF sample, are also displayed in Fig. 14(right).

We then illustrate the usefulness of the probabilistic estimation with associated uncertainty quantification that GIAnPF provides, for the image restoration task. Specifically, we focus on the restoration of `Boat` image degraded by Map 4 and no noise (i.e., $\sigma_x = 0$). We display in Fig. 16(left) the uncertainty quantification, defined as the standard deviation per pixel of the 100 restored images, obtained using the 100 samples from the estimated posterior of the blur maps, in a single run of GIAnPF. This uncertainty map can be compared with Fig. 16(right). The latter displays the mean square error map, obtained by averaging the results of 100 restorations performed with the MMSE estimator of the kernel parameters (mean of the approximated posterior), obtained by 100 different independent runs of GIAnPF. It is noticeable that our probabilistic approach with the inference performed by GIAnPF is able to quantify large uncertainty (see Fig. 16(left)) in areas where the squared errors are also large (see Fig. 16(right)). Let us point out that the uncertainty map from Fig. 16(left), is obtained without the need of processing the data multiple times, and does not require the knowledge of the original `boat` image.

Image	σ_x	NP$^+$	GIAnPF
`Boat`	0	20.1051	**20.5247**[20.3207-20.7045]
	0.01	**18.8952**	18.0044 [17.8957-18.1138]
	0.05	17.9529	**17.9610**[17.9485-17.9714]
	0.1	17.2432	**17.2755** [17.2674-17.2848]
`Goldhill`	0	20.7009	**20.9611**[20.6529-21.1522]
	0.01	19.1551	**20.2494** [20.1838-20.3108]
	0.05	18.9202	**18.9990**[18.9728-19.0205]
	0.1	18.2412	**18.3510** [18.3356-18.3630]
`Plane`	0	23.2597	**23.6970** [23.5257-23.8888]
	0.01	**22.1416**	21.8859 [21.7679-21.9899]
	0.05	20.4632	**20.5290**[20.5132-20.5487]
	0.1	19.6010	**19.6535** [19.6445-19.6654]
`Cameraman`	0	20.4253	**20.4839** [19.7894-20.9030]
	0.01	19.3293	**19.8935** [19.7158-20.0245]
	0.05	18.3764	**18.6403** [18.5687-18.6794]
	0.1	17.4996	**17.7183** [17.6861-17.7376]

TABLE V
SNR$_x$ VALUES OF RESTORED IMAGES USING ESTIMATED BLUR DERIVED FROM NP$^+$ AND GIANPF APPLIED ON THE CALIBRATED IMAGE CHART.



Fig. 14. (left) Original images ; (middle) blurry noisy versions with $\sigma_n = 0.01$; (right) example of restored image using the mean blur estimate from one run of GIAnPF.

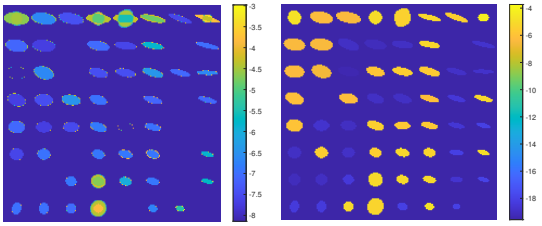


Fig. 15. Results of 100 samples from the posterior obtained with single run of GIAnPF when applied to image `Chart` degraded by Map 4 and $\sigma_n = 0.1$: (left) $\log_{10}$ of the pixel-wise mean squared error over the 100 samples. (right) $\log_{10}$ of the pixel-wise standard deviation over the 100 samples.

V. CONCLUSION

This paper addresses the estimation of the PSF parameters for spatially-varying blurs from calibrated image acquisitions. We propose an original statistical modeling of the problem, accounting for the spatial dependency among neighboring kernels, and we apply a sequential Bayesian inference technique in this context. In order to alleviate the particle degeneracy problem brought by the BPF in some cases, we also propose a new sampling method called the GIAnPF. Our results in different scenarios illustrate the good performance of the approach, including a useful uncertainty quantification. The novel approach opens many possibilities beyond this work. For instance, different noise distributions could be immediately

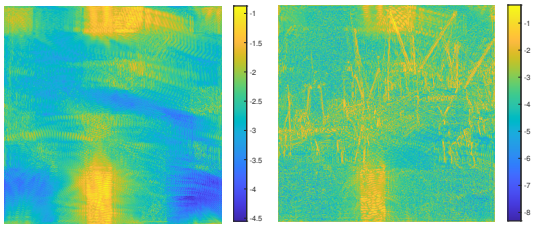


Fig. 16. Uncertainty quantification vs error in the estimation of the kernels by GIAnPF on the image `boat` ; (left) $\log_{10}$ of the pixel-wise standard deviation of the restored image, when using 100 samples from the posterior of the blur parameters, approximated by a single run of GIAnPF. (right) $\log_{10}$ of the mean pixel-wise quadratic error between the true and the restored image when using the posterior mean estimator of the parameters approximated by GIAnPF, averaged over 100 independent runs.

used. Moreover, other state-space models, not necessarily Markovian, could be considered.

APPENDIX A
USEFUL CONSTANTS ON BLUR SHAPES

We list herebelow, for the reader reference, the expression for the normalization constants and support size for the parametric blur shapes described in Section II-B.

a) Generalized Gaussian blur: The normalization constant reads:

$$\lambda = \frac{p}{\pi\Gamma(\frac{1}{p})2^{\frac{1}{p}}\alpha}\frac{1}{|\mathbf{\Sigma}(\theta,\mathbf{s})|^{\frac{1}{2}}}, \tag{30}$$

where $\Gamma(\cdot)$ is the Gamma function. The support size is:

$$s_{\max} = \sqrt{L}\left(2\alpha(-2\ln(a))^{1/p}\right)^{-1/2}. \tag{31}$$

b) Defocus blur: The normalization constant is

$$\lambda = \pi s_1 s_2 L. \tag{32}$$

The support size is:

$$s_{\max} = \max\mathcal{G} = \frac{\sqrt{L}-1}{2}. \tag{33}$$

c) Skew-normal blur: The normalization constant is

$$\lambda = \frac{2}{(2\pi)^{\frac{3}{2}}|\mathbf{S}(w)|^{\frac{1}{2}}}. \tag{34}$$

The support size is:

$$s_{\max} = \frac{\sqrt{L}}{2\sqrt{-\ln(a)}}. \tag{35}$$

APPENDIX B
NUMERICAL SETTINGS

We describe here the detailed numerical settings adopted in the experimental parts for the proposed models and methods. Table VI shows the settings for the realistic blur maps considered in Section IV-A2. Table VII lists the transition models assumed when running BPF and GIAnPF, where $\mathcal{N}_{[s_{\min},s_{\max}]}(s_{t-1},\sigma_s^2)$ denotes a truncated normal distribution, i.e., the pdf is proportional to a normal pdf with mean s_{t-1} and variance σ_s^2 in the support $[s_{\min}, s_{\max}]$, and 0 otherwise.

Blur maps	Parameters
Map 1	$\sigma_\epsilon = 0.01, \delta_{1,\min} = 0.595, \delta_{1,\max} = 0.605,$ $\delta_{2,\min} = 0.295, \delta_{2,\max} = 0.305$
Map 2	$\sigma_\epsilon = 0.01, \delta_{1,\min} = 2.95, \delta_{1,\max} = 3.05,$ $\delta_{2,\min} = 2.95, \delta_{2,\max} = 3.05$
Map 3	$\delta_{1,\min} = 2.45, \delta_{1,\max} = 2.55,$ $\delta_{2,\min} = 1.95, \delta_{2,\max} = 2.05$
Map 4	$\delta_{1,\min} = 1.45, \delta_{1,\max} = 1.55,$ $\delta_{2,\min} = 1.45, \delta_{2,\max} = 1.55$
Map 5	$\delta_{1,\max} = 0.01, \delta_{2,\max} = 0.01, \delta_{3,\max} = 0.1,$ $\delta_{4,\min} = 1.5, \delta_{4,\max} = 2, \delta_{5,\min} = 2, \delta_{5,\max} = 2.5$

TABLE VI
NUMERICAL SETTINGS FOR THE GENERATION OF THE EXPERIMENTAL SHIFT-VARIANT BLUR MAPS.

	Parameters settings	
Model A	$\begin{cases}\theta_t \sim \mathcal{N}(\theta_{t-1},\sigma_\theta^2)\\ s_{1,t} \sim \mathcal{N}_{[s_{\min},s_{\max}]}(s_{t-1},\sigma_s^2)\\ s_{2,t} \sim \mathcal{N}_{[s_{\min},s_{\max}]}(s_{t-1},\sigma_s^2)\end{cases}$	$s_{\min} = 10^{-2}$, $s_{\max}$ set as in (31).
Model B	$\begin{cases}\theta_t \sim \mathcal{N}(\theta_{t-1},\sigma_\theta^2)\\ s_{1,t} \sim \mathcal{N}_{[s_{\min},s_{\max}]}(s_{t-1},\sigma_s^2)\\ s_{2,t} \sim \mathcal{N}_{[s_{\min},s_{\max}]}(s_{t-1},\sigma_s^2)\end{cases}$	$s_{\min} = 4\cdot 10^{-2}$, $s_{\max}$ as in (33).
Model C	$\begin{cases}s_{1,t} \sim \mathcal{N}_{[s_{\min},s_{\max}]}(s_{t-1},\sigma_s^2)\\ s_{2,t} \sim \mathcal{N}_{[s_{\min},s_{\max}]}(s_{t-1},\sigma_s^2)\\ w \sim \mathcal{N}_{[w_{\min},w_{\max}]}(s_{t-1},\sigma_w^2)\\ \alpha_{1,t} \sim \mathcal{N}_{[\alpha_{\min},\alpha_{\max}]}(s_{t-1},\sigma_\alpha^2)\\ \alpha_{2,t} \sim \mathcal{N}_{[\alpha_{\min},\alpha_{\max}]}(s_{t-1},\sigma_\alpha^2)\end{cases}$	$s_{\min} = 5\cdot 10^{-2}$, $s_{\max}$ as in (35), $\{w_{\min}, w_{\max}\} = \{-0.9, 0.9\}$, $\{\alpha_{\min}, \alpha_{\max}\} = \{0.5, 2\}$, so that (11) holds.

TABLE VII
CONSIDERED TRANSITIONS MODELS AS IN EQ. (13).

REFERENCES

[1] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding blind deconvolution algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2354–2367, 2011.

[2] A. S. Carasso, "Direct blind deconvolution," *SIAM Journal on Applied Mathematics*, vol. 61, no. 6, pp. 1980–2007, 2001.

[3] T. Kenig, Z. Kam, and A. Feuer, "Blind image deconvolution using machine learning for three-dimensional microscopy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 2191–204, 12 2010.

[4] S. J. Reeves and R. M. Mersereau, "Blur identification by the method of generalized cross-validation," *IEEE Transactions of Image Processing*, vol. 1, no. 3, 1992.

[5] N. Pustelnik, A. Benazza-Benhayia, Y. Zheng, and J.-C. Pesquet, "Wavelet-based image deconvolution and reconstruction," *Wiley Encyclopedia of Electrical and Electronics Engineering*, 2016.

[6] S. Berisha and J. G. Nagy, *Iterative Methods for Image Restoration*, 2014, ch. 4, pp. 193–247.

[7] H. Kirshner, T. Pengo, N. Olivier, D. Sage, S. Manley, and M. Unser, "A psf-based approach to biplane calibration in 3d super-resolution microscopy," in *Proceedings of the 9th IEEE International Symposium on Biomedical Imaging (ISBI 2012)*, Barcelona, Spain, 2-5 May 2012, pp. 1232–1235.

[8] T. Hortholary, C. Carrion, E. Chouzenoux, J.-C. Pesquet, and C. Lefort, "Multiplex-multiphoton microscopy and computational strategy for biomedical imaging," *Microscopy Research and Technique*, 2021.

[9] F. Mannan and M. S. Langer, "Blur calibration for depth from defocus," in *Proceedings of the 13th Conference on Computer and Robot Vision (CRV 2016)*, Victoria, BC, Canada, 1-3 June 2016, pp. 281–288.

[10] T. Bell, J. Xu, and S. Zhang, "Method for out-of-focus camera calibration," *Applied Optics*, vol. 55, no. 9, pp. 2346–2352, 2016.

[11] S. M. Anthony and S. Granick, "Image analysis with rapid and accurate two-dimensional Gaussian fitting," *Langmuir*, vol. 25, no. 14, pp. 8152–8160, 2009.

[12] E. Chouzenoux, T. Lau, C. Lefort, and J. Pesquet, "Optimal multivariate gaussian fitting with applications to PSF modeling in two-photon

microscopy imaging," *Journal of Mathematical Imaging and Vision*, vol. 61, 2019.

[13] H. Kirshner, F. Aguet, D. Sage, and M. Unser, "3-D PSF fitting for fluorescence microscopy: implementation and localization application," *Journal of Microscopy*, vol. 249, no. 1, pp. 13–25, 2012.

[14] F. Xue and T. Blu, "A novel SURE-based criterion for parametric PSF estimation," *IEEE Transactions on Image Processing*, vol. 24, no. 2, pp. 595 – 607, 2014.

[15] P. Trouvé, F. Champagnat, G. L. Besnerais, and J. Idier, "Single image local blur identification," in *Proceedings of the 18th IEEE International Conference on Image Processing (ICIP 2011)*, Bruxelles, Belgium, 11-14 Sep. 2011, pp. 613–616.

[16] X. Zhang, R. Wang, X. Jiang, W. Wang, and W. Gao, "Spatially variant defocus blur map estimation and deblurring from a single image," *Journal of Visual Communication and Image Representation*, vol. 35, 2016.

[17] H. Zhe, X. Li, and M. Yang, "Joint depth estimation and camera shake removal from single blurry image," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2014)*, Columbus, OH, USA, 23-28 June 2014, pp. 2893–2900.

[18] Y. Zhang and K. Hirakawa, "Blind deblurring and denoising of images corrupted by unidirectional object motion blur and sensor noise," *IEEE Transactions on Image Processing*, vol. 25, no. 9, pp. 4129–4144, 2016.

[19] A. Jezierska, H. Talbot, and J.C.Pesquet, "Spatially variant psf modeling in confocal macroscopy," in *Proceedings of the 15th IEEE International Symposium on Biomedical Imaging (ISBI 2018)*. Washington, DC, USA, 4-7 April 2018, pp. 489–492.

[20] D. Li, R. M. Mersereau, and S. Simske, "Atmospheric turbulence-degraded image restoration using principal components analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 3, pp. 340–344, 2007.

[21] M. A. Schmitz *et al.*, "Euclid: Non-parametric point spread function field recovery through interpolation on a graph laplacian," *arXiv: Instrumentation and Methods for Astrophysics*, 2019.

[22] N. Chimitt and S. H. Chan, "Simulating anisoplanatic turbulence by sampling intermodal and spatially correlated zernike coefficients," *Optical Engineering*, vol. 59, no. 8, p. 083101, 2020.

[23] A. Chakrabarti, T. Zickler, and W. T. Freeman, "Analyzing spatially-varying blur," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2010)*, San Francisco, CA, USA, 13-18 June 2010, pp. 2512–2519.

[24] M. Hirsch, S. Sra, B. Scholkopf, and S. Harmeling, "Efficient filter flow for space-variant multiframe blind deconvolution," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2010*, San Francisco, CA, USA, 13-18 June 2010, pp. 607–614.

[25] R. Tezaur, T. Kamata, H.Li, and S. Slonaker, "A system for estimating optics blur psfs from test chart images," in *Proceedings of SPIE - Digital Photography XI*, vol. 9404, San Francisco, CA, USA, 2015.

[26] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings F - Radar and Signal Processing*, vol. 140, no. 2, pp. 107–113, 1993.

[27] P. M. Djurić, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Míguez, "Particle filtering," *IEEE Signal Processing Magazine*, vol. 20, no. 5, pp. 19–38, 2003.

[28] L. Denis, E. Thiébaut, F. Soulez, J. M. Becker, and R. Mourya, "Fast approximations of shift-variant blur," *International Journal of Computer Vision*, vol. 115, pp. 253–278, 2015.

[29] S. B. Hadj, L. Blanc-Féraud, and G. Aubert, "Space variant blind image restoration," *SIAM Journal on Imaging Sciences*, vol. 7, 2014.

[30] P. Escande and P. Weiss, "Sparse wavelet representations of spatially varying blurring operators," *SIAM Journal on Imaging Sciences*, vol. 8, pp. 2976–3014, 2015.

[31] F. Sroubek, J. Kamenicky, and Y. M. Lu, "Decomposition of space-variant blur in image deconvolution," *IEEE Signal Processing Letters*, vol. 23, no. 3, pp. 346–350, 2016.

[32] J. G. Nagy, K. Palmer, and L. Perrone, "Iterative methods for image deblurring: A matlab object-oriented approach," *Numerical Algorithms*, vol. 36, pp. 73–93, 2004.

[33] L. Bar, N. Sochen, and N. Kiryati, "Restoration of images with piecewise space-variant blur," in *Scale Space and Variational Methods in Computer Vision*, F. Sgallari, A. Murli, and N. Paragios, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 533–544.

[34] R. K. Mourya, "Contributions to image restoration : from numerical optimization strategies to blind deconvolution and shift-variant deblurring," *PhD thesis*, 2016.

[35] C. J. Schuler, M. Hirsh, and S. Harmeling, "Blind correction of optical aberrations," in *Proceedings of the European Conference on Computer Vision (ECCV 2012)*, Florence, Italy, 2012, pp. 187–200.

[36] M. Šorel and J. Flusser, "Space-variant restoration of images degraded by camera motion blur," *IEEE Transactions on Image Processing*, vol. 17, no. 2, pp. 105–116, 2008.

[37] M. Z. Alam, Q. Qian, and B. K. Gunturk, "Space-variant blur kernel estimation and image deblurring through kernel clustering," *Signal Processing: Image Communication*, vol. 76, pp. 41 – 55, 2019.

[38] M. A. Schmitz, J.-L. Starck, F. M. Ngolé Mboula, *et al.*, "Euclid: Nonparametric point spread function field recovery through interpolation on a graph Laplacian," *Instrumentation and Methods for Astrophysics*, vol. 636, 2020.

[39] P. Pankajakshan, Z. Kam, A. Dieterlen, and J.-C. Olivo-Marin, "Characterizing the 3-d field distortions in low numerical aperture fluorescence zooming microscope," *Optics Express*, vol. 20, no. 9, Apr. 2012.

[40] C. Lefort, E. Chouzenoux, L. Magnol, H. Massias, and J.-C. Pesquet, "Multi-parametric 3d-point-spread function estimation in deep multi-photon microscopy with an original computational strategy dedicated to the reconstruction of muscle images," in *Proceedings of SPIE 11354, Optical Sensing and Detection VI*, no. 113541I, 6 April 2020.

[41] A. N. Rajagopalan and S. Chaudhuri, "Mrf model-based identification of shift-variant point spread function for a class of imaging systems," *Signal Processing*, vol. 76, no. 3, pp. 285–299, 1999.

[42] E. Chouzenoux and J.-C. Pesquet, "A stochastic majorize-minimize subspace algorithm for online penalized least squares estimation," *IEEE Transactions on Signal Processing*, vol. 65, no. 18, pp. 4770–4783, 2017.

[43] M. Tico and K. Pulli, "Image enhancement method via blur and noisy image fusion," in *Proceedings of the IEEE International Conference on Image Processing (ICIP 2009)*, Cairo, Egypt, 7-11 Nov. 2009, pp. 1521–1524.

[44] M. Tico, N. Gelfand, and K. Pulli, "Motion-blur-free exposure fusion," in *Proceedings of the IEEE International Conference on Image Processing (ICIP 2010)*, Hong Kong, China, 12-15 Sep. 2010, pp. 3321–3324.

[45] F. M. Ngolé Mboula and J.-L. Starck, "PSF field learning based on optimal transport distances," *SIAM Journal on Imaging Sciences*, vol. 10, pp. 1549–1578, 2017.

[46] Y.-C. Chen, L. R. Furenlid, D. W. Wilson, and H. H. Barrett, "Calibration of scintillation cameras and pinhole SPECT imaging systems," *Small-Animal Spect Imaging*, pp. 195–201, 2005.

[47] Y. Huang, E. Chouzenoux, and V. Elvira, "Particle filtering for online space-varying blur identification," in *Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2019)*, Le Gosier, Guadeloupe, French West Indies, 2019, pp. 544–548.

[48] Y. Liu, X. Du, H. Shen, and S. Chen, "Estimating generalized gaussian blur kernels for out-of-focus image deblurring," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1, 2020.

[49] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," *Applied Optics*, vol. 46, no. 10, pp. 1819–1829, 2007.

[50] P. Trouvé-Peloux, J. Sabater, A. Bernard-Brunel, F. Champagnat, G. Le Besnerais, and T. Avignon, "Turning a conventional camera into a 3D camera with an add-on," *Applied optics*, vol. 57, no. 10, pp. 2553–2563, 2018.

[51] A. Azzalini and A. Valle, "The multivariate skew-normal distribution," *Biometrika*, vol. 83, no. 4, pp. 715–726, 1996.

[52] P. C. Hansen and T. K. Jensen, "Noise propagation in regularizing iterations for image deblurring," *Electronic Transactions on Numerical Analysis*, vol. 31, pp. 204–220, 2008.

[53] R. L. S. Jonathan D. Simpkins, "Parameterized modeling of spatially varying optical blur," *Journal of Electronic Imaging*, vol. 23, no. 1, pp. 1 – 13, 2014.

[54] S. Sarkka, "Bayesian filtering and smoothing," *Cambridge, UK:Cambridge University Press*, 2013.

[55] M. Rabbani and P. W. Jones, *Digital Image Compression Techniques*. Bellingham: Spie Optical Engineering Press, 1991.

[56] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transaction of the ASME—Journal of Basic Engineering*, pp. 35–45, 1960.

[57] L. Tiancheng, B. Miodrag, and D. Petar, "Resampling methods for particle filtering: Classification, implementation, and strategies," *IEEE Signal Processing Magazine*, vol. 32, pp. 70–86, 2015.

[58] M. K. Pitt and N. Shephard, "Auxiliary variable based particle filters," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. Gordon, Eds. Springer, 2001, ch. 13, pp. 273–293.

[59] V. Elvira, L. Martino, M. F. Bugallo, and P. M. Djurić, "In search for improved auxiliary particle filters," in *Proceedings of the 26th European Signal Processing Conference (EUSIPCO 2018)*, Rome, Italy, 3-7 Sep. 2018, pp. 1637–1641.
[60] V. Elvira, L. Martino, M. F. Bugallo, and P. M. Djuric, "Elucidating the auxiliary particle filter via multiple importance sampling [lecture notes]," *IEEE Signal Processing Magazine*, vol. 36, no. 6, pp. 145–152, 2019.
[61] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," *Handbook of Nonlinear Filtering*, vol. 12, no. 656-704, p. 3, 2009.
[62] R. M. Neal, "Annealed importance sampling," *Statistics and Computing*, vol. 11, no. 2, pp. 125–139, 2001.
[63] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. Springer, 2004.
[64] P. Del Moral, A. Doucet, and A. Jasra, "Sequential monte carlo samplers," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 3, pp. 411–436, 2006.
[65] J. Deutscher, A. Blake, and I. Reid, "Articulated body motion capture by annealed particle filtering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2000)*, vol. 2, Hilton Head, SC, 13-15 June 2000, pp. 126–133.
[66] M. Filippone, "Bayesian inference for Gaussian process classifiers with annealing and exact-approximate MCMC," in *Proceedings of the 22nd International Conference on Pattern Recognition (ICPR 2014)*, Stockholm, Sweden, 24-28 Aug. 2014.
[67] Y. Ulker, B. Gunsel, and A. T. Cemgil, "Annealed smc samplers for nonparametric bayesian mixture models," *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 3–6, 2011.
[68] J. Madapura and B. Li, "3d articulated human body tracking using kld-annealed rao-blackwellised particle filter," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME 2007)*, Beijing, China, 2-5 July 2007, pp. 1950 – 1953.
[69] J. Gall, J. Potthoff, and C. Schnörr, "Generalised annealed particle filter - mathematical framework, algorithms and applications," *PhD thesis*, 2005.
[70] V. Elvira, L. Martino, D. Luengo, M. F. Bugallo, *et al.*, "Generalized multiple importance sampling," *Statistical Science*, vol. 34, no. 1, pp. 129–155, 2019.
[71] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, pp. 183–202, 2009.
[72] E. Chouzenoux, A. Jezierska, J.-C. Pesquet, and H. Talbot, "A majorize-minimize subspace approach for l2-l0 image regularization," *SIAM Journal on Imaging Science*, vol. 6, pp. 563–591, 2013.
[73] S. Cadoni, E. Chouzenoux, J.-C. Pesquet, and C. Chaux, "A block parallel majorize-minimize memory gradient algorithm," in *Proceedings of the 23rd IEEE International Conference on Image Processing (ICIP 2016)*, Phoenix, Arizona, 25-28 Sep. 2016, pp. 3194–3198.
[74] M. Chalvidal and E. Chouzenoux, "Block distributed 3mg algorithm and its application to 3d image restoration," in *Proceedings of the 27th IEEE International Conference on Image Processing (ICIP 2020)*, Abu Dhabi, United Arab Emirates, 25-28 Oct. 2020.