Preconditioned P-ULA for Joint Deconvolution-Segmentation of Ultrasound Images

Marie-Caroline Corbineau, Denis Kouamé, Emilie Chouzenoux, Jean-Yves Tourneret, Jean-Christophe Pesquet

Preconditioned P-ULA for Joint Deconvolution-Segmentation of Ultrasound Images

Marie-Caroline Corbineau, *Student Member, IEEE*, Denis Kouamé, *Senior Member, IEEE*, Emilie Chouzenoux, *Member, IEEE*, Jean-Yves Tourneret, *Fellow, IEEE*, and Jean-Christophe Pesquet, *Fellow, IEEE*

***Abstract*—Joint deconvolution and segmentation of ultrasound images is a challenging problem in medical imaging. By adopting a hierarchical Bayesian model, we propose an accelerated Markov chain Monte Carlo scheme where the tissue reflectivity function is sampled thanks to a recently introduced proximal unadjusted Langevin algorithm. This new approach is combined with a forward-backward step and a preconditioning strategy to accelerate the convergence, and with a method based on the majorization-minimization principle to solve the inner nonconvex minimization problems. As demonstrated in numerical experiments conducted on both simulated and *in vivo* ultrasound images, the proposed method provides high-quality restoration and segmentation results and is up to six times faster than an existing Hamiltonian Monte Carlo method.**

***Index Terms*—Ultrasound, Markov chain Monte Carlo method, proximity operator, deconvolution, segmentation.**

I. Introduction

IN MEDICAL ultrasound (US) imaging, useful information can be drawn from the statistics of the tissue reflectivity function (TRF) to perform segmentation [1], tissue characterization [2], or classification [3]. Let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ be the vectorized TRF and radio-frequency (RF) image, respectively. The following simplified model is used [4], [5]

$$y = Hx + \omega, \tag{1}$$

where $H \in \mathbb{R}^{n\times n}$ is a linear operator that models the convolution with the point spread function (PSF) of the probe, and $\omega \sim \mathcal{N}(0, \sigma^2 \mathbb{I}_n)$, with $\mathcal{N}$ the normal distribution, and $\mathbb{I}_n$ the identity matrix in $\mathbb{R}^{n\times n}$. This letter assumes that the PSF is known, while $\sigma^2 > 0$ is an unknown parameter to be estimated. The TRF is comprised of K different tissues, which are identified by a hidden label field $z = (z_i)_{1 \leq i \leq n} \in \{1, \ldots, K\}^n$. For every $k \in \{1, \ldots, K\}$, the kth region is modeled by a generalized Gaussian distribution ($\mathcal{GGD}$) [3], [6], which is parametrized by a shape parameter $\alpha_k \in [0, 3]$, related to the scatterer concentration, and a scale parameter $\beta_k > 0$, linked to the signal energy. Given y and H, the aim is to estimate a deblurred image x [7], [8], as well as σ^2, $\alpha = (\alpha_k)_{1\leq k\leq K}$, $\beta = (\beta_k)_{1\leq k\leq K}$, and the label field z. Due to the interdependence of these unknowns, it is beneficial to perform the deconvolution and segmentation tasks in a joint manner [9], [10]. This is achieved in [6] by considering a hierarchical Bayesian model, which is used within a Markov chain Monte Carlo (MCMC) method [11] to sample x, σ^2, α, β, and z according to the full conditional distributions. Despite promising results in image restoration and segmentation, the method in [6] is of significant computational complexity, in particular due to the adjusted Hamiltonian Monte Carlo (HMC) method [12], [13] used to sample the TRF. Recently, efficient and reliable stochastic sampling strategies have been devised [14]–[16] using the proximity operator [17], which is known as a useful tool for large-scale nonsmooth optimization [18]. In this work, we investigate an MCMC algorithm to perform the joint deconvolution and segmentation of US images, where the TRF is sampled with a scheme inspired from the proximal unadjusted Langevin algorithm (P-ULA) [15]. The latter generates samples according to an approximation of the target distribution without acceptance test, while being geometrically ergodic whereas classical unadjusted Langevin algorithms may have convergence issues.

A. Main Contributions

Our contributions include i) the proposition of an original accelerated preconditioned version of P-ULA (PP-ULA), which relies on the use of a variable metric forward-backward strategy [19], [20], ii) an efficient solver based on the majorization-minimization (MM) principle to tackle the involved nonconvex priors, and iii) a new hybrid Gibbs sampler yielding a substantial reduction of the computational time needed to perform joint high-quality deconvolution and segmentation of both simulated and *in vivo* US images.

This article is organized as follows: Section II describes the investigated Bayesian model and sampling strategy. Section III focuses on the proposed TRF sampling method. Numerical experiments are finally presented in Section IV.

II. Bayesian Model

A. Priors

Fig. 1 illustrates the hierarchical model used to perform a joint deconvolution-segmentation of ultrasound images. The

This work was supported in part by Institut Universitaire de France and in part by the Agence Nationale de la Recherche of France under Project MAJIC ANR-17-CE40-0004-01. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ananda Shankar Chowdhury. *(Corresponding author: Marie-Caroline Corbineau.)*

M.-C. Corbineau, E. Chouzenoux, and J.-C. Pesquet are with the CVN, CentraleSupélec, INRIA Saclay, University Paris-Saclay, 91190 Gif-Sur-Yvette, France (e-mail: mariecaroline_corbineau@hotmail.fr; emilie.chouzenoux@univ-mlv.fr; jean-christophe@pesquet.eu).

D. Kouamé is with the IRIT, CNRS UMR 5505, University of Toulouse, 31062 Toulouse, France (e-mail: kouame@irit.fr).

J.-Y. Tourneret is with the IRIT, ENSEEIHT, TéSA, University of Toulouse, 31000 Toulouse, France (e-mail: Jean-Yves.Tourneret@enseeiht.fr).

Digital Object Identifier 10.1109/LSP.2019.2935610

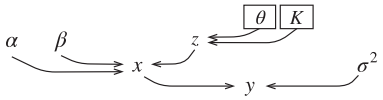


Fig. 1. Hierarchical Bayesian model. Parameters in boxes are fixed in advance.

following likelihood function is derived from (1)

$$p(y|x,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|y-Hx\|^2}{2\sigma^2}\right). \quad (2)$$

The TRF is a mixture of $\mathcal{GGD}$ s which, under the assumption that the pixel values are independent given z, leads to

$$p(x|\alpha,\beta,z) = \prod_{i=1}^{n} \frac{1}{2\beta_{z_i}^{1/\alpha_{z_i}}\Gamma(1+1/\alpha_{z_i})} \exp\left(-\frac{|x_i|^{\alpha_{z_i}}}{\beta_{z_i}}\right). \quad (3)$$

Uninformative Jeffreys priors are assigned to the noise variance and scale parameters, while the shape parameters are assumed to be uniformly distributed between 0 and 3. The labels z are modeled by a Potts Markov random field with prior

$$p(z) = \frac{1}{C(\theta)} \exp\left(\sum_{i=1}^{n}\sum_{j\in\mathcal{V}(i)} \theta\delta(z_i - z_j)\right), \quad (4)$$

with δ the Kronecker function, $C(\theta) > 0$ a normalizing constant, $\theta > 0$ a granularity coefficient, and $\mathcal{V}(i)$ the set of four closest neighbours of the ith pixel.

B. Conditional Distributions

The different variables are sampled according to their conditional distributions, which are provided in this section. The conditional distribution of the noise variance is derived from the Bayes theorem as follows

$$p(\sigma^2|y,x) \propto \mathcal{IG}\left(\frac{n}{2}, \frac{\|y-Hx\|^2}{2}\right), \quad (5)$$

where $\mathcal{IG}$ denotes the inverse gamma distribution. Assuming that the different regions have independent shape and scale parameters for every $k \in \{1,\dots,K\}$, we obtain

$$p(\alpha_k|x,\beta,z) \propto \prod_{i\in\mathcal{I}_k} \frac{\mathbf{1}_{[0,3]}(\alpha_k)}{2\beta_k^{1/\alpha_k}\Gamma\left(1+1/\alpha_k\right)} \exp\left(-\frac{|x_i|^{\alpha_k}}{\beta_k}\right), \quad (6)$$

$$p(\beta_k|x,\alpha,z) \propto \mathcal{IG}\left(\frac{n_k}{\alpha_k}, \sum_{i\in\mathcal{I}_k}|x_i|^{\alpha_k}\right), \quad (7)$$

with $\mathcal{I}_k = \{i \in \{1,\dots,n\} | z_i = k\}$, n_k the number of elements in $\mathcal{I}_k$, and $\mathbf{1}_{[0,3]}$ the characteristic function of [0,3]. Samples for α_k are drawn from (6) using a Metropolis-Hastings (MH) random walk. For every pixel $i \in \{1,\dots,n\}$ and every region $k \in \{1,\dots,K\}$, the Bayes rule applied to the segmentation labels leads to

$$p(z_i = k|x,\alpha,\beta,z_{\mathcal{V}(i)}) \propto \frac{\exp\left(\sum_{j\in\mathcal{V}(i)}\theta\delta(z_j-k) - \frac{|x_i|^{\alpha_k}}{\beta_k}\right)}{2\beta_k^{1/\alpha_k}\Gamma(1+1/\alpha_k)} \quad (8)$$

where $z_{\mathcal{V}(i)}$ denotes the label values in the neighborhood of i. As a consequence, the label z_i is drawn from $\{1,\dots,K\}$ using the above probabilities (suitably normalized).

Algorithm 1: DFB Algorithm to Compute $\mathrm{prox}_f^Q(x)$.

Initialize dual variable $w^{(1)} \in \mathbb{R}^N$;
Set $\rho = |||Q|||^{-1}$, $\epsilon \in]0, \min\{1,\rho\}[$, $\eta \in [\epsilon, 2\rho-\epsilon]$;
for $p = 1, \dots$ **do**
  $u^{(p)} = x - Qw^{(p)}$;
  $w^{(p+1)} = w^{(p)} + \eta u^{(p)} - \eta \mathrm{prox}_{\eta^{-1}f}(\eta^{-1}w^{(p)} + u^{(p)})$
end

III. Preconditioned P-ULA

A. Notation

Let $\mathcal{I}_{<1} = \{i \in \{1,\dots,n\} \mid \alpha_{z_i} < 1\}$ and $\mathcal{I}_{\geq 1} = \{1,\dots,n\}\backslash\mathcal{I}_{<1}$. Let $\mathcal{S}_n$ denote the set of symmetric positive definite matrices in $\mathbb{R}^{n\times n}$, and let $|||\cdot|||$ denote the spectral norm. For every $Q \in \mathcal{S}_n$, let $\|\cdot\|_Q = \langle\cdot, Q\cdot\rangle^{1/2}$. For every function $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, the proximity operator of f at $x \in \mathbb{R}^n$ with respect to the norm induced by $Q^{-1} \in \mathcal{S}_n$ is defined as follows [17],

$$\mathrm{prox}_f^Q(x) \in \mathrm{Argmin}_{u\in\mathbb{R}^n} \frac{1}{2}\|x-u\|_{Q^{-1}}^2 + f(u). \quad (9)$$

If Q is not specified, then $Q = \mathbb{I}_n$. If prox_f is simple to compute, then the solution to (9) for an arbitrary $Q \in \mathcal{S}_n$ can be obtained by using the dual forward-backward (DFB) algorithm [21], summarized in Algorithm 1. If f is proper, lower semicontinuous, and convex, then the sequence $(u^{(p)})_{p\in\mathbb{N}}$ generated by Algorithm 1 converges to $\mathrm{prox}_f^Q(x)$.

B. Sampling the TRF

The conditional distribution of the TRF is

$$\pi(x) = p(x|y,\sigma^2,\alpha,\beta,z) \propto \exp\left(-\frac{\|y-Hx\|^2}{2\sigma^2} - g(x)\right), \quad (10)$$

where $(\forall x \in \mathbb{R}^n)\ g(x) = \sum_{i=1}^n \beta_{z_i}^{-1}|x_i|^{\alpha_{z_i}}$. Let $\gamma > 0$ and let $Q \in \mathcal{S}_n$ be a preconditioning matrix used to accelerate the sampler [22]. Following [15], $\pi(x)$ is approximated by

$$\pi_\gamma(x) \propto \sup_{u\in\mathbb{R}^n} \pi(u) \exp\left(-\frac{\|u-x\|_{Q^{-1}}^2}{2\gamma}\right). \quad (11)$$

As it is detailed in [23], the Euler discretization of the Langevin diffusion equation [24] applied to π_γ with stepsize 2γ and preconditioning matrix Q leads to

$$x^{(t+1)} = \mathrm{prox}_{\gamma g}^Q(\tilde{x}^t) + \sqrt{2\gamma}Q^{\frac{1}{2}}\omega^{(t+1)}, \quad (12)$$

where $\omega^{(t+1)} \sim \mathcal{N}(0,\mathbb{I}_n)$ and

$$\tilde{x}^{(t)} = x^{(t)} - \frac{\gamma}{\sigma^2}QH^\top(Hx^{(t)} - y). \quad (13)$$

Since the proposed sampling strategy is unadjusted, (12) is not followed by an acceptance test. The bias with respect to π increases with γ, as the speed of convergence of the algorithm. A compromise must be found when setting γ. When $\mathcal{I}_{<1}$ is not

Algorithm 2: MM Principle to Compute $\mathrm{prox}^{Q}_{\gamma g}$.

Initialize $u^{(1)} \in \mathbb{R}^n$;
for $q = 1, \ldots$ **do**
  $\mathcal{J}^{(q)} = \{i \in \mathcal{I}_{<1} \mid |u_i^{(q)}| > 0\}$;
  $v^{(q)} = (|u_i^{(q)}|)_{1\leq i\leq n}$;
  $u^{(q+1)} = \mathrm{prox}^{Q}_{\gamma h_{\mathcal{J}^{(q)}}(\cdot, v^{(q)})}(\tilde{x}^{(t)})$ (using Alg. 1)
end

Algorithm 3: Hybrid Gibbs Sampler.

1 Sample the noise variance σ^2 according to (5);
2 Sample the shape parameter α using MH with (6);
3 Sample the scale parameter β according to (7);
4 Sample the hidden label field z using (8);
5 Sample the TRF x using PP-ULA (12)–(13).

empty, we use the MM principle [25] to replace the nonconvex minimization problem involved in the computation of $\mathrm{prox}^{Q}_{\gamma g}$ with a sequence of convex surrogate problems. Let $\mathcal{J} \subset \mathcal{I}_{<1}$. We define $h_{\mathcal{J}}$ at every $(u, v) \in \mathbb{R}^n \times \mathbb{R}^n_{+*}$ by

$$h_{\mathcal{J}}(u,v) = \sum_{i\in\mathcal{I}_{\geq 1}} \frac{|u_i|^{\alpha_{z_i}}}{\beta_{z_i}} + \sum_{j\in\mathcal{J}} \frac{(1-\alpha_{z_j})v_j^{\alpha_{z_j}} + \alpha_{z_j} v_j^{\alpha_{z_j}-1}|u_j|}{\beta_{z_j}}.$$

From concavity, we deduce that, for every $v \in \mathbb{R}^n_{+*}$ and $u \in \mathbb{R}^n$ such that $\mathcal{J} \subset \{i \in \mathcal{I}_{<1} \mid |u_i| > 0\}$, the following majorization property holds

$$h_{\mathcal{J}}(u,v) \geq \sum_{i\in\mathcal{I}_{\geq 1}\cup\mathcal{J}} \frac{|u_i|^{\alpha_{z_i}}}{\beta_{z_i}} = h_{\mathcal{J}}(u, (|u_i|)_{1\leq i\leq n}).$$

Since $h_{\mathcal{J}}(\cdot, v)$ is convex and separable, its proximity operator in the Euclidean metric is straightforward to compute. More precisely, for every $i \in \mathcal{I}_{\geq 1}$, $\eta > 0$ and $s \in \mathbb{R}$, $\mathrm{prox}_{\eta^{-1}|\cdot|^{\alpha_{z_i}}}(s)$ has either a closed form [26] or can be found using a bisection search in $[0, |s|]$. Algorithm 1 can then be called, in order to compute the proximity operator of $h_{\mathcal{J}}(\cdot, v)$ in any metric $Q \in \mathcal{S}_n$. This leads to Algorithm 2 which generates a sequence $(u^{(q)})_{q\in\mathbb{N}}$ estimating $\mathrm{prox}^{Q}_{\gamma g}(\tilde{x}^{(t)})$.

The resulting Gibbs sampler is summarized in Algorithm 3.

IV. Numerical Experiments

A. Experimental Settings

Four experiments are presented. Simu1 and Simu2 refer to simulated images with two and three regions, respectively. Kidney denotes the tissue-mimicking phantom from the Field II simulator [27]. The PSF for these simulations, obtained with Field II, corresponds to a 3.5 MHz linear probe. Finally, Thyroid denotes a real RF image of thyroidal flux obtained *in vivo* with a 7.8 MHz probe. The unknown PSF is identified using the RF image of a wire cross-section acquired with the same probe. Since the diameter of the wire is of the order of a few µm, its cross-section can almost be viewed as a point. Thus, its RF image provides a good approximation of the PSF. The number of regions K is set to 3 for Kidney and Thyroid.

The test settings can be found in Table I. The TRF is initialized using a pre-deconvolved image obtained with a Wiener filter, while the segmentation is initialized by applying a 7×7 median filter and the Otsu method [28] to the B-mode of the initial TRF. Shape and scale parameters are randomly selected in $[0.5,1.5]$, and $[1,200]$, respectively. The granularity parameter θ for the Potts model (4) is adjusted to ensure that the percentage of isolated points in the segmentation, obtained with a 3×3 median filter, is close to 0.05, 0.1, 0.8 and 0.08 for Simu1, Simu2, Kidney and Thyroid, respectively.

TABLE I
Test Settings, Computational Time and MSJ Per s

		Simu1	Simu2	Kidney	Thyroid
Iterations	Size	256×256	256×256	294×354	870×140
	Burn-in	4000	10000	7000	3000
	Total	8000	20000	14000	6000
Time	HMC	1h08min	4h14min	4h23min	2h09min
	PP-ULA	12min	39min	42min	35min
	Speed gain	**5.7**	**6.6**	**6.3**	**3.7**
MSJ (per s)	HMC	173	22	167	175
	PP-ULA	970	793	657	950

B. Comparisons and Evaluation Metrics

All computational times are given for simulations run on Matlab 2018b on an Intel Xeon CPU E5-1650 3.20 GHz. In addition to comparing Algorithm 3 with HMC [6], the quality of the deconvolution is compared with the one obtained with a Wiener filter, where the noise level has been estimated as in [29], and with the solution to the Lasso problem, where the regularization weight is set i) manually when the ground-truth is not available, or ii) using a golden-section search to maximize the peak signal-to-noise ratio (PSNR) defined as (with x^{tr} the true TRF and x^{es} the estimated one)

$$\mathrm{PSNR} = 10\log_{10}(n \max_{1\leq i\leq n}(x_i^{\mathrm{tr}}, x_i^{\mathrm{es}})^2 / \|x^{\mathrm{tr}} - x^{\mathrm{es}}\|^2). \quad (14)$$

We also compare our results with the segmentation given by Otsu's method [28] applied to the Wiener-deconvolved image, and with the SLaT method [30] applied to the Lasso-deconvolved image. PP-ULA is used with $\gamma = 0.09$ and Q an approximation to the inverse of the Hessian of the differentiable term in (10) [31], $Q = \sigma^2(H^\top H + \lambda \mathbb{I}_n)^{-1}$, with $\lambda = 0.1$ so that Q is well-defined.

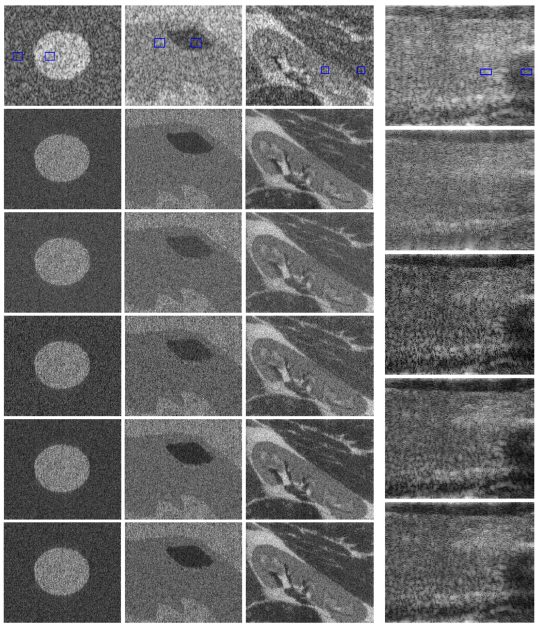
We have also computed the structural similarity measure (SSIM) [32] of the restored TRF and the contrast-to-noise ratio (CNR) [33] between two windows from different regions of the B-mode TRF images. The segmentation is evaluated according to the percentage of correctly predicted labels, or overall accuracy (OA). The minimum mean square error (MMSE) estimators of all parameters in HMC and PP-ULA are computed after the burn-in regime. Moreover, to evaluate the mixing property of the Markov chain after convergence, we compute the mean square jump (MSJ) per second, which is the ratio of the MSJ to the time per iteration. The MSJ is obtained using T samples of the TRF $(x^{t_0+1}, \ldots, x^{t_0+T})$ generated after the burn-in period, i.e. $\mathrm{MSJ} = (\frac{1}{T-1}\sum_{t=1}^{T-1} \|x^{(t_0+t)} - x^{(t_0+t+1)}\|^2)^{1/2}$.

C. Results and Discussion

The convergence speed of Algorithm 3 is empirically observed in all experiments, as illustrated in Fig. 3, where we also display the results of the non-preconditioned P-ULA, for which $Q = \mathbb{I}_n$ and $\gamma = 1.99\sigma^2/|||H|||^2$. P-ULA is 12.2 and 4.8 times slower than PP-ULA on Simu1 and Simu2, respectively [23]. Fig. 4 and Table II show that P-ULA is also more biased, clearly emphasizing the benefits of preconditioning in this example. From Table I, PP-ULA is between 3.7 and 6.6 times faster than HMC and has better mixing properties, as shown by the MSJ

TABLE II
MMSE ESTIMATES OF THE NOISE VARIANCE AND $\mathcal{GGD}$ PARAMETERS

	Simu1					Simu2						
	σ^2	α_1	β_1	α_2	β_2	σ^2	α_1	β_1	α_2	β_2	α_3	β_3
True	0.013	1.5	1.0	0.60	1.0	33	1.5	100	1.0	50	0.50	4.0
P-ULA	0.041	2.0	0.5	0.59	1.0	122	2.0	330	2.0	3186	0.48	3.4
HMC	0.013	1.8	1.2	0.61	1.0	34	1.4	66	1.1	111	0.54	5.2
PP-ULA	0.013	1.4	0.9	0.62	1.1	35	2.3	2676	1.2	122	0.55	5.8



(a) (b)

Fig. 2. B-mode visualization. (a) Left to right: Simu1, Simu2, Kidney. Top to bottom: RF image, TRF: ground-truth, Wiener, Lasso, HMC, PP-ULA. (b) Thyroid. Top to bottom: RF image, TRF: Wiener, Lasso, HMC, PP-ULA. Blue boxes indicate regions used for the CNR.

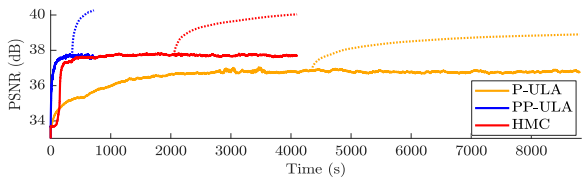


Fig. 3. PSNR along time for Simu1. Dotted lines indicate the PSNR of the MMSE estimator of the TRF after the burn-in regime.

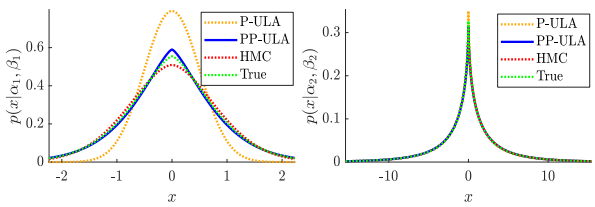


Fig. 4. Simu1, $\mathcal{GGD}$ distributions (3) of regions 1 (left) and 2 (right).

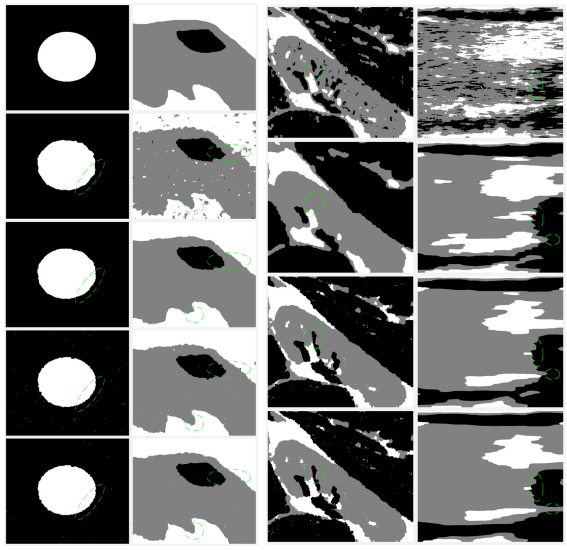
per second. From Table II and Fig. 4, we deduce that PP-ULA samples correctly the target distributions. Visual results from Fig. 2 and CNR values in Tables III and IV show that the contrast obtained with PP-ULA is better than with competitors on almost all test images. In addition, the PSNR and SSIM values from Tables III and IV obtained with PP-ULA are equivalent or higher

TABLE III
PSNR, SSIM, CNR AND OA RESULTS ON SIMULATED DATA

	Simu1				Simu2			
	PSNR	SSIM	CNR	OA	PSNR	SSIM	CNR	OA
Wiener-Otsu	37.1	0.57	1.26	99.5	35.4	0.63	0.97	96.0
Lasso-SLaT [30]	39.2	0.60	1.15	99.6	37.8	0.70	0.99	98.3
P-ULA	38.9	0.45	**1.82**	98.7	37.1	0.57	1.59	94.9
HMC	40.0	**0.62**	1.47	**99.7**	36.4	0.64	1.59	98.5
PP-ULA	**40.3**	**0.62**	1.51	**99.7**	**38.6**	**0.71**	**1.64**	**98.7**

TABLE IV
RESULTS ON TISSUE-MIMICKING AND REAL DATA

	Kidney			Thyroid
	PSNR	SSIM	CNR	CNR
Wiener	27.6	0.58	0.66	0.56
Lasso	28.5	0.59	0.67	0.99
HMC	**29.5**	**0.62**	1.10	1.52
PP-ULA	29.3	**0.62**	**1.14**	**1.56**



(a) (b)

Fig. 5. Segmentation. (a) Left: Simu1, right: Simu2, top to bottom: ground-truth, Otsu, SLaT, HMC, PP-ULA. (b) Left: Kidney, right: Thyroid, top to bottom: Otsu, SLaT, HMC, PP-ULA. Main differences are circled in green.

than all competitors for all experiments. Visual segmentation results are shown in Fig. 5, and OA values can be found in Table III. For the simulated images, more pixels are correctly labeled with PP-ULA than with competitors. Regarding Kidney and Thyroid, the segmentation based on the Potts model (PP-ULA and HMC) gives more homogeneous areas than Otsu, and recovers more details than SLaT.

V. CONCLUSION

In this letter we investigated a new method based on a preconditioned proximal unadjusted Langevin algorithm for the joint restoration and segmentation of ultrasound images, which showed faster convergence than an existing Hamiltonian Monte Carlo algorithm. A direction for future work is to extend this framework to a spatially variant PSF.

REFERENCES

[1] M. A. Pereyra, N. Dobigeon, H. Batatia, and J.-Y. Tourneret, "Segmentation of skin lesions in 2D and 3D ultrasound images using a spatially coherent generalized Rayleigh mixture model," *IEEE Trans. Med. Imag.*, vol. 31, no. 8, pp. 1509–1520, Aug. 2012.

[2] O. Bernard, J. D'hooge, and D. Friboulet, "Statistics of the radio-frequency signal based on K distribution with application to echocardiography," *IEEE Trans. Ultrasonics, Ferroelectrics, Freq. Control*, vol. 53, no. 9, pp. 1689–1694, Sep. 2006.

[3] M. Alessandrini *et al.*, "A restoration framework for ultrasonic tissue characterization," *IEEE Trans. Ultrasonics, Ferroelectrics, Freq. Control*, vol. 58, no. 11, pp. 2344–2360, Nov. 2011.

[4] J. A. Jensen, J. Mathorne, T. Gravesen, and B. Stage, "Deconvolution of in-vivo ultrasound B-mode images," *Ultrasonic Imag.*, vol. 15, no. 2, pp. 122–133, 1993.

[5] J. Ng, R. Prager, N. Kingsbury, G. Treece, and A. Gee, "Modeling ultrasound imaging as a linear, shift-variant system," *IEEE Trans. Ultrasonics, Ferroelectrics, Freq. Control*, vol. 53, no. 3, pp. 549–563, Mar. 2006.

[6] N. Zhao, A. Basarab, D. Kouamé, and J.-Y. Tourneret, "Joint segmentation and deconvolution of ultrasound images using a hierarchical Bayesian model based on generalized Gaussian priors," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3736–3750, Aug. 2016.

[7] J. A. Jensen, "Deconvolution of ultrasound images," *Ultrasonic Imag.*, vol. 14, no. 1, pp. 1–15, 1992.

[8] O. Michailovich and A. Tannenbaum, "Blind deconvolution of medical ultrasound images: A parametric inverse filtering approach," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 3005–3019, Dec. 2007.

[9] H. Ayasso and A. Mohammad-Djafari, "Joint NDT image restoration and segmentation using Gauss–Markov–Potts prior models and variational Bayesian computation," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2265–2277, Sep. 2010.

[10] A. Pirayre, Y. Zheng, L. Duval, and J.-C. Pesquet, "HOGMep: Variational Bayes and higher-order graphical models applied to joint image recovery and segmentation," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2017, pp. 3775–3779.

[11] M. Pereyra *et al.*, "A survey of stochastic simulation and optimization methods in signal processing," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 2, pp. 224–241, Mar. 2016.

[12] R. M. Neal, "MCMC using Hamiltonian dynamics," in *Handbook of Markov Chain Monte Carlo*, vol. 2, S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng, Eds. Boca Raton, Florida, USA: Chapman Hall/CRC, 2011, pp. 113–162.

[13] C. P. Robert, V. Elvira, N. Tawn, and C. Wu, "Accelerating MCMC algorithms," *Wiley Interdisciplinary Rev.: Comput. Statist.*, vol. 10, no. 5, 2018, Art. no. e1435.

[14] A. Durmus, E. Moulines, and M. Pereyra, "Efficient Bayesian computation by proximal Markov chain Monte Carlo: When Langevin meets Moreau," *SIAM J. Imag. Sci.*, vol. 11, no. 1, pp. 473–506, 2018.

[15] M. Pereyra, "Proximal Markov chain Monte Carlo algorithms," *Statist. Comput.*, vol. 26, no. 4, pp. 745–760, 2016.

[16] A. Schreck, G. Fort, S. Le Corff, and E. Moulines, "A shrinkage-thresholding Metropolis adjusted Langevin algorithm for Bayesian variable selection," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 2, pp. 366–375, Mar. 2016.

[17] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Basel, Switzerland: Springer Nature Switzerland AG, 2017.

[18] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-point Algorithms for Inverse Problems in Science and Engineering*, H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. Russell Luke, and H. Wolkowicz, Eds. Basel, Switzerland: Springer Nature Switzerland AG, 2011, pp. 185–212.

[19] A. M. Stuart, J. Voss, and P. Wilberg, "Conditional path sampling of SDEs and the Langevin MCMC method," *Commun. Math. Sci.*, vol. 2, no. 4, pp. 685–697, 2004.

[20] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, "Variable metric forward–backward algorithm for minimizing the sum of a differentiable function and a convex function," *J. Optim. Theory Appl.*, vol. 162, no. 1, pp. 107–132, 2014.

[21] P. L. Combettes, D. Dũng, and B. C. Vũ, "Proximity for sums of composite functions," *J. Math. Anal. Appl.*, vol. 380, no. 2, pp. 680–688, 2011.

[22] Y. Marnissi, E. Chouzenoux, A. Benazza-Benyahia, and J.-C. Pesquet, "Majorize-minimize adapted Metropolis-Hastings algorithm. Application to multichannel image recovery,' in *Proc. 22nd European Signal Processing Conference (EUSIPCO)*, Lisbon, 2014, pp. 1332–1336.

[23] M.-C. Corbineau, D. Kouamé, E. Chouzenoux, J.-Y. Tourneret, and J.-C. Pesquet, "Preconditioned P-ULA for joint deconvolution-segmentation of ultrasound images – Extended version," Tech. Rep., 2019. Available: https://arxiv.org/abs/1903.08111

[24] G. O. Roberts and O. Stramer, "Langevin diffusions and Metropolis-Hastings algorithms," *Methodology Comput. Appl. Probability*, vol. 4, no. 4, pp. 337–357, 2002.

[25] E. D. Schifano, R. L. Strawderman, and M. T. Wells, "Majorization-minimization algorithms for nonsmoothly penalized objective functions," *Electron. J. Statist.*, vol. 4, pp. 1258–1299, 2010.

[26] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs, "A variational formulation for frame-based inverse problems," *Inverse Problems*, vol. 23, no. 4, pp. 1495–1518, Jun. 2007.

[27] J. A. Jensen, "Simulation of advanced ultrasound systems using Field II," in *Proc. 2nd IEEE Int. Symp. Biomed. Imag.: Nano to Macro*. IEEE, 2004, pp. 636–639.

[28] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybernet.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.

[29] S. Mallat, *A Wavelet Tour of Signal Processing*. Amsterdam, Netherlands: Elsevier, 1999.

[30] X. Cai, R. Chan, M. Nikolova, and T. Zeng, "A three-stage approach for segmenting degraded color images: Smoothing, lifting and thresholding (SLaT)," *J. Scientific Comput.*, vol. 72, no. 3, pp. 1313–1332, 2017.

[31] S. Becker and J. Fadili, "A quasi-Newton proximal splitting method," in *Proc. 25th Int. Conf. Neural Inf. Process. Syst.*, vol. 2. Red Hook, NY, USA: Curran Associates Inc., 2012, pp. 2618–2626.

[32] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[33] S. Krishnan, K. W. Rigby, and M. O'donnell, "Improved estimation of phase aberration profiles," *IEEE Trans. Ultrasonics, Ferroelectrics, Freq. Control*, vol. 44, no. 3, pp. 701–713, May 1997.