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, Denis Kouamé, *Senior Member, IEEE,* Emilie Chouzenoux, *Member, IEEE,* Jean-Yves Tourneret, *Fellow, IEEE,* 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

SCATTERING frequently occurs during the acquisition of medical ultrasound images, which produces speckle in the resulting radio-frequency (RF) data. Useful information can be drawn from the statistics of the echo envelope to perform segmentation [1], tissue characterization [2], or classification [3]. The vectorized RF image $y \in \mathbb{R}^n$ is modeled as follows

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

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, $x \in \mathbb{R}^n$ is the vectorized tissue reflectivity function (TRF), 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 paper 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 $z = (z_i)_{1\leq i\leq n} \in \{1,\dots,K\}^n$. For every $k \in \{1,\dots,K\}$, the k^{th} region is modeled by a generalized Gaussian distribution ($\mathcal{GGD}$) [3], [4], 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 [5], [6], 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 [7], [8]. This is achieved in [4] by considering a hierarchical Bayesian model, which is used within a Markov chain Monte Carlo (MCMC) method [9] 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 [4] is of significant computational complexity, in particular due to the adjusted Hamiltonian Monte Carlo (HMC) method [10], [11] used to sample the TRF. Recently, efficient and reliable stochastic sampling strategies have been devised [12], [13], [14] using the proximity operator [15], which is known as a useful tool for large-scale nonsmooth optimization [16].

In this work, we propose to accelerate the MCMC algorithm of [4] by replacing the HMC step used to sample the TRF, with a proximal unadjusted Langevin algorithm (P-ULA) proposed in [13]. P-ULA generates samples according to an approximation of the target distribution. One advantage is that P-ULA is geometrically ergodic whereas classical unadjusted Langevin algorithms may have convergence issues. We propose the introduction of a preconditioning matrix in P-ULA for acceleration purposes [17], [18], as well as an efficient solver based on the majorization-minimization (MM) principle to solve the involved nonconvex subproblems. As shown in numerical experiments conducted on simulated and *in vivo* ultrasound images, the proposed preconditioned P-ULA (PP-ULA) yields a substantially lower computational time than P-ULA and HMC while achieving similar, and sometimes better, image restoration and segmentation quality.

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 PP-ULA. 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 following likelihood function is derived from (1)

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

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

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

M.-C. Corbineau, E. Chouzenoux and J.-C. Pesquet are with the CVN, CentraleSupélec, INRIA Saclay, University Paris-Saclay, Gif-Sur-Yvette, France (e-mail: first-name.last-name@centralesupelec.fr)

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

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

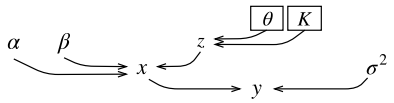

Fig. 1. Hierarchical Bayesian model. Parameters in boxes are fixed in advance.

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

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

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 i^{th} 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 $\mathcal{IG}(\cdot,\cdot)$ 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

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

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

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).

III. Preconditioned P-ULA

A. Notation

Let $\mathcal{I}_{<1} = \{i \in \{1,\dots,n\} \mid \alpha_{z_i} < 1\}$ and $\mathcal{I}_{\geq 1} = \{1,\dots,n\} \setminus \mathcal{I}_{<1}$. 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 [15],

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

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 [19], 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 $\text{prox}_f^Q(x)$.

Algorithm 1: DFB algorithm to compute $\text{prox}_f^Q(x)$

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**
  $u^{(p)} = x - Qw^{(p)}$;
  $w^{(p+1)} = w^{(p)} + \eta u^{(p)} - \eta\text{prox}_{\eta^{-1}f}(\eta^{-1}w^{(p)} + u^{(p)})$
end

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 [20]. Following [13], $\pi(x)$ is approximated by

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

The Euler discretization of the Langevin diffusion equation [21] applied to π_γ with stepsize 2γ leads to

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

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 empty, we use the MM principle [22] to replace the nonconvex minimization problem involved in the computation of prox_g^Q 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 majoration 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 [23] 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 $\mathrm{prox}_g^Q(\tilde{x}^{(t)})$.

Algorithm 2: MM principle to compute prox_g^Q.

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

The resulting Gibbs sampler is summarized in Algorithm 3.

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).

IV. Numerical experiments

A. Experimental settings

Four experiments are implemented in Matlab. Simu1 and Simu2 refer to simulated images with two and three regions, respectively. Kidney denotes the tissue-mimicking phantom from the Field II simulator [24]. The PSF for these simulations is obtained with Field II using 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 the cross section of a wire acquired with the same probe. 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 [25] 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 0.2 for Simu1, Simu2, Kidney and Thyroid, respectively.

B. Comparisons and evaluation metrics

We compare the proposed Algorithm 3 with the HMC approach from [4]. We also provide the deconvolution results obtained by a Wiener filter (Wnr), where the noise level is estimated as in [26], and the segmentation results given by the Otsu method [25] applied to Wnr (Otsu). PP-ULA is used with $\gamma = 0.09$ and Q an approximation to the inverse of the Hessian of the differentiable term in (10) [27], $Q = \sigma^2(H^\top H + \lambda \mathbb{I}_n)^{-1}$, with $\lambda = 0.1$ so that Q is well-defined. When the true TRF x^{tr} is available, the estimate x^{es} is evaluated by using the peak signal-to-noise ratio

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

We also compute the contrast-to-noise ratio $\mathrm{CNR} = |\mu_1 - \mu_2|/(\nu_1 + \nu_2)^{1/2}$, with the means (μ_1, μ_2) and variances (ν_1, ν_2) in two windows from different regions of the B-mode image. 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,

$$\mathrm{MSJ} = \left(\frac{1}{T-1} \sum_{t=1}^{T-1} \|x^{(t_0+t)} - x^{(t_0+t+1)}\|^2 \right)^{1/2}. \quad (15)$$

		Simu1	Simu2	Kidney	Thyroid
	Size	256×256	256×256	294×354	880×196
Iterations	Burn-in	4000	10000	7000	3000
	Total	8000	20000	14000	6000
	HMC	1h19min	4h32min	4h41min	2h50min
Time	PP-ULA	13min	41min	46min	48min
	Speed gain	**6.2**	**6.7**	**6.1**	**3.5**
	HMC	140	22	155	0.5
MSJ (per s)	PP-ULA	915	762	599	137

TABLE I
Test settings, computational time and MSJ per s.

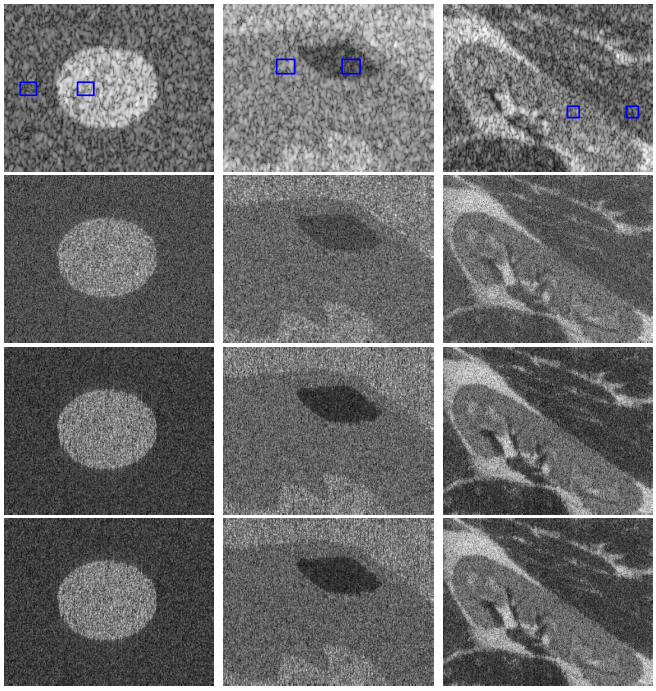

Fig. 2. First row: B-mode of RF image. Second to last row: B-mode of the TRF: Wnr, HMC, PP-ULA. Left to right: Simu1, Simu2, Kidney (blue boxes indicate regions used to compute the CNR).

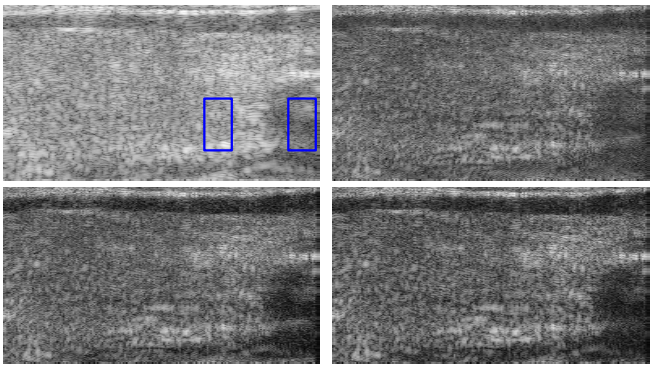

Fig. 3. B-mode of Thyroid. Left to right, top to bottom: RF image (blue boxes indicate regions used to compute the CNR), TRF: Wnr, HMC, PP-ULA.

C. Results and discussion

The convergence speed of Algorithm 3 is empirically observed in all experiments, as illustrated in Fig. 4. 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 needs more iterations and more time to converge than PP-ULA, and it is also more biased, clearly emphasizing the benefits of preconditioning in this example. From Table I, PP-ULA is up to 6.7 times faster than HMC and has better mixing properties as shown by the MSJ per second, while sampling correctly the target distributions. As shown in Table II and Fig. 5, the model parameters given by PP-ULA are close to the true ones. Visual results from Figs. 2 and 3, and CNR values in Table III show that the contrast obtained with HMC and PP-ULA is better than with Wnr.

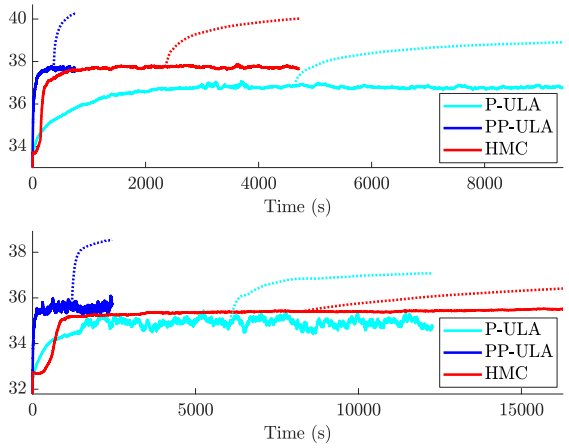

Fig. 4. PSNR along time. Dotted lines indicate the PSNR of the MMSE estimator of the TRF after the burn-in regime. Top: Simu1. Bottom: Simu2.

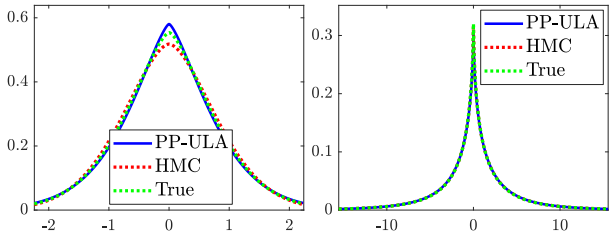

Fig. 5. Simu1, $\mathcal{GGD}$ distributions of regions 1 (left) and 2 (right).

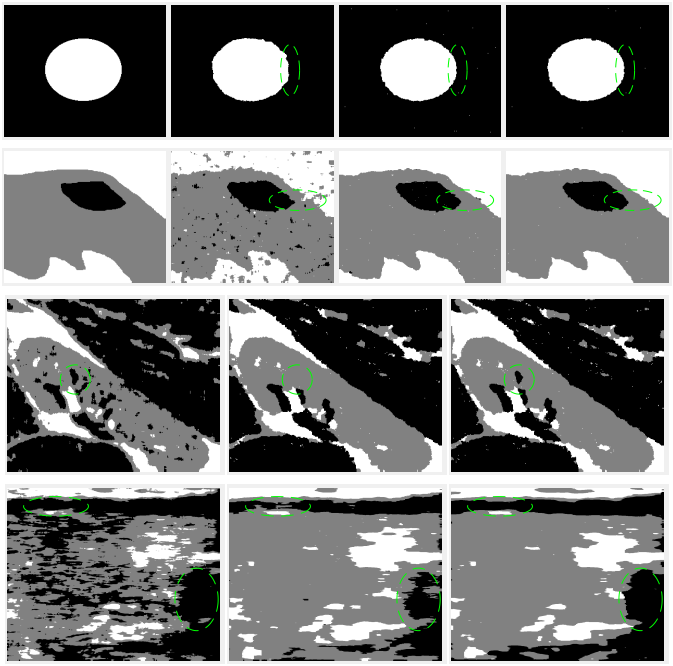

Fig. 6. Segmentation. Top to bottom: Simu1, Simu2, Kidney, Thyroid. Rows 1 and 2 left to right: ground-truth, Ostu, HMC, PP-ULA. Rows 3 and 4 left to right: Ostu, HMC, PP-ULA. Main differences are circled in green.

The PSNR and CNR values obtained with PP-ULA are equivalent or higher than both competitors on almost all test images. Visual segmentation results are shown in Fig. 6, and OA values can be found in Table III. For the simulated images, more pixels are correctly labeled with PP-ULA than with HMC or Otsu. Regarding Kidney and Thyroid, the segmentation based on the Potts model gives more homogeneous areas than the one based on pixel intensities.

	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
HMC	0.013	1.7	1.1	0.60	1.0	34	2.5	3642	1.1	100	0.54	5.2
PP-ULA	0.013	1.4	0.9	0.62	1.1	35	2.1	1068	1.2	118	0.55	5.7

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

	Simu1			Simu2			Kidney		Thyroid
	PSNR	OA	CNR	PSNR	OA	CNR	PSNR	CNR	CNR
Wnr-Otsu	37.1	99.4	1.26	35.4	93.8	0.97	27.6	0.66	0.80
HMC	40.0	**99.7**	1.49	36.4	98.6	1.59	**29.5**	1.11	1.20
PP-ULA	**40.3**	**99.7**	**1.52**	**38.6**	**98.7**	**1.66**	29.3	**1.14**	**1.43**

TABLE III
DECONVOLUTION AND SEGMENTATION RESULTS: PSNR, CNR AND OA.

V. CONCLUSION

This letter studied 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, possibly unknown, 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 Transactions on Medical Imaging*, vol. 31, no. 8, pp. 1509–1520, 2012.

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

[3] M. Alessandrini, S. Maggio, J. Porée, L. De Marchi, N. Speciale, E. Franceschini, O. Bernard, and O. Basset, "A restoration framework for ultrasonic tissue characterization," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 58, no. 11, 2011.

[4] 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 Transactions on Image Processing*, vol. 25, no. 8, pp. 3736–3750, 2016.

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

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

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

[8] 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 *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 3775–3779.

[9] M. Pereyra, P. Schniter, E. Chouzenoux, J.-C. Pesquet, J.-Y. Tourneret, A. O. Hero, and S. McLaughlin, "A survey of stochastic simulation and optimization methods in signal processing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 2, pp. 224–241, 2016.

[10] R. M. Neal, "MCMC using Hamiltonian dynamics," *Handbook of Markov Chain Monte Carlo*, vol. 2, no. 11, pp. 2, 2011.

[11] C. P. Robert, V. Elvira, N. Tawn, and C. Wu, "Accelerating MCMC algorithms," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 10, no. 5, 2018.

[12] A. Durmus, E. Moulines, and M. Pereyra, "Efficient Bayesian computation by proximal Markov chain Monte Carlo: when Langevin meets Moreau," *SIAM Journal on Imaging Sciences*, vol. 11, no. 1, pp. 473–506, 2018.

[13] M. Pereyra, "Proximal Markov chain Monte Carlo algorithms," *Statistics and Computing*, vol. 26, no. 4, pp. 745–760, 2016.

[14] A. Schreck, G. Fort, S. Le Corff, and E. Moulines, "A shrinkage-thresholding Metropolis adjusted Langevin algorithm for Bayesian variable selection," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 2, pp. 366–375, 2016.

[15] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, Springer, 2017.

[16] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-point algorithms for inverse problems in science and engineering*, pp. 185–212. Springer, 2011.

[17] A. M. Stuart, J. Voss, and P. Wilberg, "Conditional path sampling of SDEs and the Langevin MCMC method," *Communications in Mathematical Sciences*, vol. 2, no. 4, pp. 685–697, 2004.

[18] 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," *Journal of Optimization Theory and Applications*, vol. 162, no. 1, pp. 107–132, 2014.

[19] P. L. Combettes, D. Dũng, and B. C. Vũ, "Proximity for sums of composite functions," *Journal of Mathematical Analysis and applications*, vol. 380, no. 2, pp. 680–688, 2011.

[20] Y. Marnissi, E. Chouzenoux, A. Benazza-Benyahia, and J.-C. Pesquet, "Majorize-minimize adapted Metropolis-Hastings algorithm," *HAL preprint HAL:01909153*, 2018.

[21] G. O. Roberts and O. Stramer, "Langevin diffusions and Metropolis-Hastings algorithms," *Methodology and computing in applied probability*, vol. 4, no. 4, pp. 337–357, 2002.

[22] E. D. Schifano, R. L. Strawderman, and M. T. Wells, "Majorization-minimization algorithms for nonsmoothly penalized objective functions," *Electronic Journal of Statistics*, vol. 4, pp. 1258–1299, 2010.

[23] 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, June 2007.

[24] J. A. Jensen, "Simulation of advanced ultrasound systems using Field II," in *4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2004, pp. 636–639.

[25] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[26] S. Mallat, *A wavelet tour of signal processing*, Elsevier, 1999.

[27] S. Becker and J. Fadili, "A quasi-Newton proximal splitting method," in *Advances in Neural Information Processing Systems*, 2012, pp. 2618–2626.