HAL
open science

Bayesian Estimation of Smooth Altimetric Parameters: Application to Conventional and Delay/Doppler Altimetry

Abderrahim Halimi, Corinne Mailhes, Jean-Yves Tourneret, Hichem Snoussi

To cite this version:

Abderrahim Halimi, Corinne Mailhes, Jean-Yves Tourneret, Hichem Snoussi. Bayesian Estimation of Smooth Altimetric Parameters: Application to Conventional and Delay/Doppler Altimetry. IEEE Transactions on Geoscience and Remote Sensing, 2016, vol. 54 (n° 4), pp. 2207-2219. 10.1109/TGRS.2015.2497583 . hal-01311346

HAL Id: hal-01311346

https://hal.science/hal-01311346v1

Submitted on 4 May 2016

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

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

Open Archive TOULOUSE Archive Ouverte (OATAO)
OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : http://oatao.univ-toulouse.fr/
Eprints ID : 15731

To link to this article : DOI:10.1109/TGRS.2015.2497583
URL : http://dx.doi.org/10.1109/TGRS.2015.2497583

To cite this version : Halimi, Abderrahim and Mailhes, Corinne and Tourneret, Jean-Yves and Snoussi, Hichem *Bayesian Estimation of Smooth Altimetric Parameters: Application to Conventional and Delay/Doppler Altimetry.* (2016) IEEE Transactions on Geoscience and Remote Sensing, vol. 54 (n° 4). pp. 2207-2219. ISSN 0196-2892

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Bayesian Estimation of Smooth Altimetric Parameters: Application to Conventional and Delay/Doppler Altimetry

Abderrahim Halimi, *Member, IEEE*, Corinne Mailhes, *Member, IEEE*, Jean-Yves Tourneret, *Senior Member, IEEE*, and Hichem Snoussi

***Abstract*—This paper proposes a new Bayesian strategy for the smooth estimation of altimetric parameters. The altimetric signal is assumed to be corrupted by a thermal and speckle noise distributed according to an independent and non-identically Gaussian distribution. We introduce a prior enforcing a smooth temporal evolution of the altimetric parameters which improves their physical interpretation. The posterior distribution of the resulting model is optimized using a gradient descent algorithm which allows us to compute the maximum *a posteriori* estimator of the unknown model parameters. This algorithm has a low computational cost that is suitable for real-time applications. The proposed Bayesian strategy and the corresponding estimation algorithm are evaluated using both synthetic and real data associated with conventional and delay/Doppler altimetry. The analysis of real Jason-2 and CryoSat-2 waveforms shows an improvement in parameter estimation when compared to state-of-the-art estimation algorithms.**

***Index Terms*—Altimetry, Bayesian algorithm, coordinate descent algorithm, delay/Doppler altimetry (DDA), natural gradient, simultaneous autoregression (SAR) altimetry, smooth estimation.**

I. Introduction

A SATELLITE altimeter is a nadir viewing radar that emits regular pulses and records the travel time, the magnitude, and the shape of each return signal after reflection on the Earth's surface. This reflected echo provides information about some physical parameters such as the range between the satellite and the observed scene (denoted by τ), the significant wave height (denoted by SWH), and the wind speed (related to the echo's amplitude P_u). However, altimetric waveforms are corrupted by speckle noise that reduces the quality of the retrieved geophysical parameters. Therefore, many recent studies and missions have been focusing on improving the quality of these estimated parameters by reducing the noise effect. We distinguish between two approaches that have been focusing on improving the altimetric technology or improving the processing of the available data. The first approach is based on reducing the measurement noise of conventional altimeters (such as TOPEX and Poseidon-3) by increasing the number of observations (looks). This can be achieved by keeping the same conventional technology and improving the altimeter characteristics (as for AltiKa [1]–[3]) or by using a new delay/Doppler processing [4] that results in a different echo's shape [5]–[7]. The second approach processes the available data by using improved models for the altimetric echoes [8]–[11] or improved estimation algorithms [12]–[14].

Following this second approach, the goal of this paper is to elaborate a new Bayesian model and its estimation algorithm accounting for the smooth evolution of altimetric signals. Usual methods for parameter estimation are based on maximum-likelihood [15] or weighted [12], [16] and unweighted least squares (LS) approaches [6], [7], [9]. Most existing algorithms consider altimetric echoes independently and estimate one set of parameters per echo. However, it is well known that adjacent echoes share similar characteristics because of their physical nature. The correlation between these adjacent echoes has been analyzed in studies such as [17]–[19]. This correlation has also been considered in [12] which processed a set of echoes by imposing a spline evolution for the epoch parameter τ. However, this method was developed by fixing the significant wave height for the whole set of echoes. In [13], the authors proposed a three-step algorithm that reduces the range noise by first estimating three altimetric parameters (SWH, τ, and P_u), then filtering the estimated SWH using a Gaussian filter, and, finally, re-estimating the epoch τ while fixing SWH. This last algorithm is flexible since it does not enforce the parameter SWH to be constant. However, it requires us to use a sequential estimator whose computational cost can be prohibitive. Moreover, the algorithm of [13] requires the characterization of a Gaussian filter for SWH which has been chosen empirically. Finally, it is interesting to note that the algorithms developed in [12] and [13] have shown interesting properties especially in reducing the variance of the estimated altimetric parameters which encourages the consideration of the echo's correlation to improve parameter estimation.

The first contribution of this paper is the elaboration of a hierarchical Bayesian model that allows smooth estimation of altimetric parameters by considering the correlation between

Manuscript received February 17, 2015; revised June 25, 2015 and August 31, 2015; accepted November 1, 2015. Date of publication December 24, 2015; date of current version March 9, 2016.

A. Halimi and H. Snoussi are with the Institut Charles Delaunay (CNRS), Université de technologie de Troyes, 10010 Troyes, France (e-mail: Abderrahim.Halimi@utt.fr; Hichem.Snoussi@utt.fr).

C. Mailhes and J.-Y. Tourneret are with the University of Toulouse IRIT/INP-ENSEEIHT/TéSA, 31071 Toulouse, France (e-mail: Corinne.Mailhes@enseeiht.fr; Jean-Yves.Tourneret@enseeiht.fr).

successive altimetric echoes. Each altimetric echo is assumed to be corrupted by an additive, independent, and non-identically Gaussian noise [20]. This noise model generalizes the independent and identically distributed (i.i.d.) Gaussian noise that is generally assumed when considering an unweighted LS approach [6], [9]. The proposed approach allows the estimation of the noise parameters from the data which brings additional robustness compared to algorithms requiring a priori known noise parameters. The parameters of interest of the proposed model (i.e., SWH, τ, and P_u) are assigned priors enforcing a smooth evolution between consecutive signals which improves their estimation. These priors are defined from the discrete Laplacian of the different parameters. This kind of prior has shown increasing interest for many problems such as image deconvolution [21], [22], hyperspectral unmixing [23], and medical imaging [24]. Note that the proposed Bayesian hierarchy is generic in the sense that it can be applied to both conventional altimetry (CA) and delay/Doppler altimetry (DDA).

The second contribution of this paper is the derivation of an estimation algorithm for the altimetric and noise parameters associated with the proposed hierarchical Bayesian model. However, the minimum mean square error (MMSE) and maximum *a posteriori* (MAP) estimators of these parameters cannot be easily computed from the obtained joint posterior. This problem has been solved in [25] by drawing samples distributed according to a similar posterior using Markov chain Monte Carlo (MCMC) methods. More precisely, a Hamiltonian Monte Carlo (HMC) algorithm was used for its good mixing properties for high-dimensional vectors [26]. The HMC algorithm provided good estimation results, and it will be briefly recalled in this paper. However, the resulting HMC-based estimators have a high computational complexity which can be penalizing for real-time applications. In this paper, we propose an algorithm that significantly reduces this computational cost by using a gradient descent approach. More precisely, a coordinate descent algorithm [23], [27] is used to sequentially update the noise and altimetric parameters. This algorithm uses a natural gradient approach [28], weighting the gradient direction by using the Fisher information matrix and leading to a fast convergence.[1] The proposed Bayesian model and estimation algorithm are evaluated using synthetic and real echoes acquired during the Jason-2 and CryoSat-2 missions. The obtained results are very promising and confirm the potential of the proposed Bayesian model and its associated algorithm.

This paper is organized as follows. Section II presents the different models used for CA and DDA. The proposed hierarchical Bayesian model and its estimation algorithm are introduced in Sections III and IV. Section V is devoted to testing and validating the proposed technique using synthetic data with controlled ground truth. Section VI shows the results obtained using real data resulting from Jason-2 and CryoSat-2 missions. Conclusion and future work are finally reported in Section VII.

[1]This approach is also known as Fisher scoring [10], [29]. It updates the parameters in a Riemannian space, resulting in a fast convergence to a local minimum of the cost function of interest [30], [31].

II. Altimetric Models

This section introduces the altimetric models used for CA and DDA that are characterized by three altimetric parameters: the significant wave height SWH, the epoch τ, and the amplitude P_u.

A. Conventional and BMs

In CA, the mean power of the altimetric echo $s(t)$ is expressed by a convolution of three terms that are the probability density function (pdf) of the heights of the specular scatterers, the point target response (PTR) of the radar,[2] and the flat surface impulse response (FSIR) as follows [8], [34]:

$$s(t) = \text{FSIR}(t) * \text{PDF}(t) * \text{PTR}_T(t) \tag{1}$$

with

$$\text{FSIR}(t) = P_u \exp\left[-\frac{4c}{\gamma h}(t-\tau_s)\right] U(t-\tau_s) \tag{2}$$

$$\text{PDF}(t) = \frac{1}{\sqrt{2\pi}\sigma_s} \exp\left(-\frac{t^2}{2\sigma_s^2}\right) \tag{3}$$

$$\text{PTR}_T(t) = \left|\frac{\sin\left(\pi\frac{t}{T}\right)}{\pi\frac{t}{T}}\right|^2 \tag{4}$$

where $\sigma_s = \text{SWH}/2c$, t is the time, $\tau_s = 2\tau/c$ is the epoch expressed in seconds whereas τ is expressed in meters, c is the speed of light, T is the time resolution, γ is an antenna beamwidth parameter, h is the minimum satellite-surface distance, and $U(\cdot)$ denotes the Heaviside function. This model has been considered for parameter estimation in many studies such as [6] and [16]. However, the numerical computation of the double convolution (1) leads to a high computational cost when processing a large amount of altimetric data. Therefore, a simplified analytical model known as the Brown model (BM) is often considered. The BM is obtained by assuming a centered Gaussian approximation (with a standard deviation (STD) equal to σ_p) for the squared cardinal sine PTR_T, leading to [8], [9]

$$s(t) = \frac{P_u}{2}\left[1 + \text{erf}\left(\frac{t-\tau_s-\alpha\sigma_c^2}{\sqrt{2}\sigma_c}\right)\right] \exp\left[-\alpha\left(t-\tau_s-\frac{\alpha\sigma_c^2}{2}\right)\right] \tag{5}$$

where

$$\sigma_c^2 = \left(\frac{\text{SWH}}{2c}\right)^2 + \sigma_p^2 \tag{6}$$

and where $\text{erf}(t) = (2/\sqrt{\pi})\int_0^t e^{-z^2}\,dz$ stands for the Gaussian error function and α and σ_p^2 are two known parameters (depending on the satellite and on the Gaussian approximation of the PTR). The model (5) is clearly more attractive than (1) for designing simple estimation algorithms.

[2]Instead of the considered squared cardinal sine PTR, a measured PTR could also be used in CA as in [32] and [33].

B. DDA

DDA aims at reducing the measurement noise and increasing the along-track resolution in comparison with CA. The mean power of a delay/Doppler echo can be expressed as the convolution of three terms: the time/frequency FSIR, the pdf of the heights of the specular scatterers, and the time/frequency PTR as follows [6], [7], [20], [35]:

$$P(t,f) = \text{FSIR}(t,f) * \text{PDF}(t) * \text{PTR}(t,f) \tag{7}$$

with

$$\text{PTR}(t,f) = \text{PTR}_T(t)\text{PTR}_F(f),\ \text{PTR}_F(f) = \left| \frac{\sin\left(\pi \frac{f}{F}\right)}{\pi \frac{f}{F}} \right|^2 \tag{8}$$

where F is the frequency resolution. The FSIR is given by [6]

$$\text{FSIR}(t,q) = \frac{P_u}{\pi} \exp\left[-\frac{4c}{\gamma h}(t-\tau_s)\right] U(t-\tau_s) \times \left[\phi_{t,q+1}(\tau_s) - \phi_{t,q}(\tau_s)\right] \tag{9}$$

with

$$\phi_{t,q}(\tau_s) = \text{Re}\left[\arctan\left(\frac{y_q}{\sqrt{\rho^2(t-\tau_s) - y_q^2}}\right)\right] \tag{10}$$

where $q = 1, \ldots, Q$, with Q being the number of Doppler beams ($Q = 64$ for CryoSat-2), $\rho(t) = \sqrt{hct}$ is the radius of the propagation circles (iso-range circles), $y_q = (h\lambda/2v_s)f_q$ is the ordinate of the qth Doppler beam, $f_q = (q - 32.5)F$ is the qth Doppler frequency (32.5 is introduced to obtain a central beam when $Q = 64$), v_s is the satellite velocity, λ is the wavelength, and $\text{Re}(x)$ denotes the real part of the complex number x [the interested reader is invited to consult [6] for more details about (7)]. The double convolution (7) leads to a 2-D delay/Doppler map (DDM). The analysis of the DDA waveform is based on a multilook echo obtained by applying Doppler processing (slant range correction and multilooking) to the DDM (see [4] and [5]). The resulting multilook delay/Doppler signal can be written as

$$s(t) = \sum_{q=1}^{Q} P(t - \delta t_q, f_q) \tag{11}$$

where δt_q is the delay compensation expressed in seconds. Note finally that the discrete altimetric echo (for CA, BM, and DDA) is gathered in the vector $\boldsymbol{s} = (s_1, \ldots, s_K)^T$, where $K = 128$ gates and $s_k = s(kT)$.

C. Noise Model

The altimetric echoes are corrupted by a speckle noise whose influence is reduced by averaging, onboard the satellite, a sequence of L consecutive echoes. This operation reduces the noise variance by $\sqrt{L}$ when assuming pulse-to-pulse statistical independence between the averaged echoes[3] (see [36] for the introduction of the Walsh criteria). Considering the central limit theorem and using the fact that the averaging is conducted on a large number of echoes, the resulting noise sequence can be approximated by a Gaussian distribution. This approximation has been successfully adopted in the altimetric community [16], [20], [37] and is implicitly used in the well-known LS estimation algorithms [6], [12], [13], [38]. Therefore, this paper considers that the altimetric waveform is corrupted by an additive Gaussian noise as follows:

$$\boldsymbol{y}_m = \boldsymbol{s}_m(\boldsymbol{\Theta}_m) + \boldsymbol{e}_m,\ \text{with } \boldsymbol{e}_m \sim \mathcal{N}(\mu_m \mathbf{1}_K, \boldsymbol{\Sigma}_m) \tag{12}$$

where $\boldsymbol{y}_m$ (respectively, $\boldsymbol{s}_m$) is a $(K \times 1)$ vector representing the mth observed echo (respectively, altimetric model), $\boldsymbol{\Theta}_m = [\theta_1(m), \theta_2(m), \theta_3(m)] = [\text{SWH}(m), \tau(m), P_u(m)]$ is a (1×3) vector containing the 3 altimetric parameters SWH, τ, and P_u for the mth echo, $\boldsymbol{e}_m$ is an independent and non-identically Gaussian noise with mean vector $\mu_m \mathbf{1}_K$ representing the thermal noise (where $\mathbf{1}_K$ is a $K \times 1$ vector of 1), and diagonal covariance matrix $\boldsymbol{\Sigma}_m = \text{diag}(\boldsymbol{\sigma}_m^2)$, with $\boldsymbol{\sigma}_m^2 = (\sigma_{m1}^2, \ldots, \sigma_{mK}^2)^T$ a $(K \times 1)$ vector containing the noise variances of the mth echo.[4] Note that the thermal noise parameter is sometimes represented by an additive constant that is added to the altimetric model $\boldsymbol{s}_m(\boldsymbol{\Theta}_m)$. In the proposed model, this parameter has been included in the noise sequence $\boldsymbol{e}_m$ of (12), where it represents the noise mean μ_m.

III. Hierarchical Bayesian Model

This section introduces a hierarchical Bayesian model for estimating the parameters of M successive altimetric echoes. After defining a statistical model associated with the observed data, the Bayesian approach consists of determining the distribution of the observed data (known as the likelihood) and the priors assigned to the unknown parameters summarizing the knowledge about these parameters. The Bayes theorem allows the likelihood and the priors to be combined to build the posterior distribution of the statistical model. More precisely, if $f(\boldsymbol{x})$ denotes the prior distribution assigned to the parameter $\boldsymbol{x}$, the Bayesian approach computes the posterior distribution of $\boldsymbol{x}$ using Bayes rule

$$f(\boldsymbol{x}|\boldsymbol{Y}) = \frac{f(\boldsymbol{Y}|\boldsymbol{x})f(\boldsymbol{x})}{f(\boldsymbol{Y})} \propto f(\boldsymbol{Y}|\boldsymbol{x})f(\boldsymbol{x}) \tag{13}$$

where $\propto$ means "proportional to" and $f(\boldsymbol{Y}|\boldsymbol{x})$ is the likelihood of the observation vector $\boldsymbol{Y}$. The mean vector and maximum of this posterior can then be computed in order to estimate the unknown parameters $\boldsymbol{x}$. The resulting estimators are referred to as the MMSE and MAP estimators, respectively. The following sections introduce the likelihood and the prior distributions considered in this paper. The unknown parameters of the proposed model include the thermal noise represented by an $(M \times 1)$

[3] This assumption is true for CA echoes, while it is not valid for DDA echoes as explained in [5] and [16].

[4] Note that the noise variance varies with respect to the time instant k and the observed echo m.

vector $\boldsymbol{\mu}$, the $(K \times M)$ matrix $\boldsymbol{\Lambda} = [\boldsymbol{\sigma}_1^2, \ldots, \boldsymbol{\sigma}_M^2]$ containing the noise variances associated with the considered M echoes, and the $(M \times 3)$ matrix gathering the 3 altimetric parameters (SWH, τ, and P_u) of the M echoes.

A. Likelihood

The observation model defined in (12) and the Gaussian properties of the noise sequence $\boldsymbol{e}_m$ yield

$$f(\boldsymbol{y}_m|\boldsymbol{\Theta}_m, \mu_m, \boldsymbol{\Sigma}_m) \propto \left(\frac{1}{\prod_{k=1}^{K} \sigma_{mk}^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\boldsymbol{x}_m^T \boldsymbol{\Sigma}_m^{-1} \boldsymbol{x}_m\right\} \tag{14}$$

where $\boldsymbol{x}_m = \boldsymbol{y}_m - \boldsymbol{s}_m - \mu_m \mathbf{1}_K$ and $\boldsymbol{x}_m(\boldsymbol{\Theta}_m)$ [respectively, $\boldsymbol{s}_m(\boldsymbol{\Theta}_m)$] has been denoted by $\boldsymbol{x}_m$ (respectively, $\boldsymbol{s}_m$) for brevity. Assuming independence between the observed echoes leads to

$$f(\boldsymbol{Y}|\boldsymbol{\Theta}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) \propto \prod_{m=1}^{M} f(\boldsymbol{y}_m|\boldsymbol{\Theta}_m, \mu_m, \boldsymbol{\Sigma}_m) \tag{15}$$

where $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3]$ is an $(M \times 3)$ matrix containing the altimetric parameters of M echoes, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_M)^T$ is an $(M \times 1)$ vector containing the noise means, and $\boldsymbol{\Lambda} = [\boldsymbol{\sigma}_1^2, \ldots, \boldsymbol{\sigma}_M^2]$ is a $(K \times M)$ matrix containing the noise variances.

B. Priors for the Altimetric Parameters

The altimetric parameters are assigned priors enforcing a smooth evolution for successive waveforms. The considered priors constrain the derivatives of the altimetric parameters to have small values. This can be done by assigning a Gaussian distribution to the second derivative of the altimetric parameter $\boldsymbol{\theta}_i$ as follows:

$$f\left(\boldsymbol{\theta}_i|\epsilon_i^2\right) \propto \left(\frac{1}{\epsilon_i^2}\right)^{\frac{M}{2}} \exp\left(-\frac{1}{2\epsilon_i^2}\|\boldsymbol{D}\boldsymbol{\theta}_i\|^2\right) \tag{16}$$

for $i \in \{1, \ldots, 3\}$, where ϵ_i^2 is a hyperparameter, $\|\cdot\|$ denotes the standard l_2 norm such that $\|\boldsymbol{x}\|^2 = \boldsymbol{x}^T\boldsymbol{x}$, and $\boldsymbol{D}$ is the discrete Laplacian operator. This prior has been referred to as simultaneous autoregression (SAR) or conditional autoregression (CAR) in the image deconvolution context [21], [22]. It has also been used for different problems such as spectral unmixing of hyperspectral images [23], medical imaging applications [24], or spectroscopy [39]. The main originality of this work is to consider this prior for the parameters of altimetric waveforms.

C. Prior for the Noise Parameters

The noise parameters introduced in (12) are also estimated by the proposed estimation algorithm. The absence of knowledge about the noise means (which represent the thermal noise) and covariances should be considered when choosing the priors for these parameters. In this paper, we consider the following conjugate Gaussian prior with a large variance for the noise mean:

$$f(\boldsymbol{\mu}|\psi^2) \sim \mathcal{N}(\mathbf{0}_M, \psi^2\mathbf{I}_M) \tag{17}$$

where$\sim$means "is distributed according to," $\mathbf{I}_M$ is the $(M \times M)$ identity matrix, $\mathbf{0}_M$ is an $(M \times 1)$ vector of 0, and ψ^2 is a fixed large variance ($\psi^2 = 10^2$ in the rest of this paper). Considering the noise covariance matrix, we have first to note that, in the case of speckle noise, these variances depend on the echo's shape at each instant, i.e., the variances change from one echo to the other. However, we are generally interested in parameters estimated with a frequency of 1 Hz, i.e., we consider that the physical oceanic parameters do not change significantly in 1 s. This assumption means that the shapes of successive oceanic echoes acquired within a time interval of 1 s do not change significantly[5] and thus that the noise variances are almost the same for these echoes. Therefore, we consider that the noise variance is constant for each $r = 20$ successive echoes.[6] This means that $\sigma_{(n-1)r+1,k}^2 = \cdots = \sigma_{nr,k}^2$ for $n \in \{1, \ldots, N\}$, with $N = M/r$ (note that the general case is obtained by considering $r = 1$). After assuming prior independence between the noise variances $\sigma_{nr,k}^2$, the noninformative Jeffreys prior of $\boldsymbol{\Lambda}$ is defined as

$$f(\boldsymbol{\Lambda}) = \prod_{n=1}^{N}\prod_{k=1}^{K} \frac{1}{\sigma_{nr,k}^2} \mathbb{I}_{\mathbb{R}^+}\left(\sigma_{nr,k}^2\right) \tag{18}$$

where $\mathbb{I}_{\mathbb{R}^+}(.)$ is the indicator of the set $\mathbb{R}^+$.

D. Hyperparameter Priors

The hyperparameters $\epsilon_i^2, i \in \{1, \ldots, 3\}$ are assigned independent inverse gamma distributions as follows:

$$\epsilon_i^2 \sim \mathcal{IG}(a_i, b_i) \tag{19}$$

where a_i and b_i are fixed hyperparameters that depend on the dynamic magnitude of the ith altimetric parameter [39]. Owing to the fast convergence of the proposed algorithm, these parameters have been set by cross-validation to improve the obtained results.

E. Marginalized Posterior Distribution

The proposed Bayesian model includes the parameters $\boldsymbol{\Theta}$, $\boldsymbol{\mu}$, and $\boldsymbol{\Lambda}$ and the hyperparameters $\epsilon_i^2, i \in \{1, \ldots, 3\}$, which is summarized in the directed acyclic graph (DAG) displayed in Fig. 1. The joint posterior distribution of this Bayesian model can be computed from the following hierarchical structure:

$$f(\boldsymbol{\Theta}, \boldsymbol{\mu}, \boldsymbol{\Lambda}, \boldsymbol{\epsilon}|\boldsymbol{Y}) \propto f(\boldsymbol{Y}|\boldsymbol{\Theta}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) f(\boldsymbol{\Theta}, \boldsymbol{\mu}, \boldsymbol{\Lambda}, \boldsymbol{\epsilon}) \tag{20}$$

[5]Note that this assumption is not true in coastal zones where contaminated waveforms are common.

[6]The pulse repetition frequency of Jason-2 is 2 kHz. The echoes are generally averaged to deliver a 20-Hz data, i.e., it delivers 20 echoes for each second of data.

with

$$f(\boldsymbol{\Theta}, \boldsymbol{\mu}, \boldsymbol{\Lambda}, \boldsymbol{\epsilon}) = f(\boldsymbol{\mu}|\psi^2) f(\boldsymbol{\Lambda}) \prod_{i=1}^{3} f\left(\boldsymbol{\theta}_i|\epsilon_i^2\right) f\left(\epsilon_i^2|a_i, b_i\right) \quad (21)$$

where we have assumed *a priori* independence between the altimetric parameters and hyperparameters. This distribution can be marginalized with respect to (w.r.t.) $\boldsymbol{\epsilon}$ as follows:

$$\begin{aligned} f(\boldsymbol{\Theta}, \boldsymbol{\mu}, \boldsymbol{\Lambda}|\boldsymbol{Y}) &\propto f(\boldsymbol{Y}|\boldsymbol{\Theta}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) f(\boldsymbol{\mu}|\psi^2) f(\boldsymbol{\Lambda}) \\ &\quad \times \prod_{i=1}^{3} \left[\int f\left(\boldsymbol{\theta}_i|\epsilon_i^2\right) f\left(\epsilon_i^2|a_i, b_i\right) d\epsilon_i^2 \right] \\ &\propto f(\boldsymbol{Y}|\boldsymbol{\Theta}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) f(\boldsymbol{\mu}|\psi^2) f(\boldsymbol{\Lambda}) \prod_{i=1}^{3} f(\boldsymbol{\theta}_i|a_i, b_i) \end{aligned} \quad (22)$$

where

$$f(\boldsymbol{\theta}_i|a_i, b_i) \propto \left(\frac{\|\boldsymbol{D}\boldsymbol{\theta}_i\|^2}{2} + b_i \right)^{-a_i - M/2}. \quad (23)$$

The advantage of this marginalization is to get rid of the hyperparameter vector $\boldsymbol{\epsilon}$ and thus to make the estimation procedure more robust. However, even with this simplification, the MMSE and MAP estimators associated with the marginalized posterior (22) are not easy to determine mainly because of the nonlinearities relating the noiseless altimetric waveform $\boldsymbol{s}_m$ and the parameters of interest (τ and SWH). In [25], an MCMC algorithm was proposed to sample the joint posterior distribution of a similar Bayesian model.[7] The generated samples were then used to compute the MMSE parameter estimators (see Appendix A for more details). However, this MCMC algorithm has a high computational complexity which can be penalizing for real-time applications. In this paper, we propose an alternative based on an optimization technique maximizing the marginalized posterior (22) w.r.t. the parameters of interest [or equivalently, minimizing the negative log-posterior $-\log[p(\boldsymbol{\Theta}, \boldsymbol{\mu}, \boldsymbol{\Lambda}|\boldsymbol{Y})]$ defined in (22)]. The mode of the posterior distribution (22) is classically referred to as the marginalized MAP estimator of ($\boldsymbol{\Theta}$, $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}$).

IV. Coordinate Descent Algorithm

This section describes a gradient descent algorithm maximizing the marginalized posterior (22) w.r.t. the noise and altimetric parameters, i.e., determining the marginal MAP estimator of ($\boldsymbol{\Theta}$, $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}$). As explained before, maximizing (22) w.r.t. ($\boldsymbol{\Theta}$, $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}$) reduces to minimizing the negative log-posterior defined as (after removing unnecessary constants)

$$\begin{aligned} \mathcal{C}(\boldsymbol{\Theta}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = &\left(\frac{r}{2} + 1\right) \sum_{k=1}^{K} \sum_{n=1}^{N} \log \sigma_{nr,k}^2 + \sum_{m=1}^{M} \frac{\mu_m^2}{2\psi^2} \\ &+ \sum_{i=1}^{3} \left(a_i + \frac{M}{2}\right) \log\left(\frac{\|\boldsymbol{D}\boldsymbol{\theta}_i\|^2}{2} + b_i\right) \\ &+ \sum_{m=1}^{M} \frac{\boldsymbol{x}_m^T \boldsymbol{\Sigma}_m^{-1} \boldsymbol{x}_m}{2}. \end{aligned} \quad (24)$$

Because of the large number of parameters in ($\boldsymbol{\Theta}$, $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}$), we propose a coordinate descent algorithm [23], [27] that sequentially updates the different parameters. More precisely, the algorithm estimates the altimetric and noise parameters iteratively. In each step, the parameters of interest are estimated while holding the other parameters fixed. This process is repeated until the algorithm has converged to a local minimum of the cost function $\mathcal{C}(\boldsymbol{\Theta}, \boldsymbol{\mu}, \boldsymbol{\Lambda})$. The main steps of the resulting algorithm are described in Algorithm 1. The next sections describe the suboptimization procedures maximizing the cost function $\mathcal{C}(\boldsymbol{\Theta}, \boldsymbol{\mu}, \boldsymbol{\Lambda})$ w.r.t. the altimetric and noise parameters. A convergence diagnosis is also provided to study the convergence of the proposed approach.

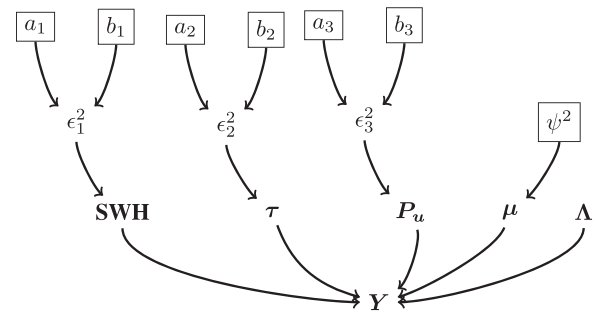

Fig. 1. DAG for the parameter and hyperparameter priors (the user-fixed parameters appear in boxes).

Algorithm 1 Gradient Descent Algorithm

1: Initialization
2: Initialize parameters $\boldsymbol{\theta}_i^{(0)}$, for $i \in \{1, \ldots, 3\}$, $\boldsymbol{\mu}^{(0)}$, $\boldsymbol{\Lambda}^{(0)}$ and $t = 1$
3: conv $= 0$,
4: Parameter update
5: **while** conv $= 0$ **do**
6: Update $\boldsymbol{\theta}_i^{(t)}$, for $i \in \{1, \ldots, 3\}$ according to (25)
7: Update $\boldsymbol{\mu}^{(t)}$ according to (31)
8: Update $\boldsymbol{\Lambda}^{(t)}$ according to (34)
9: Set conv $= 1$ if the convergence criteria are satisfied
10: $t = t + 1$
11: **end while**

1) Natural Gradient for Altimetric Parameters: In order to ensure a fast estimation of the altimetric parameters, we

[7] Note that the Bayesian model in [25] differs slightly from the described model since the parameters ϵ_i are not marginalized and the hyperparameter values are set to $a_i = b_i = 0$ corresponding to Jeffreys prior for ϵ_i.

propose to use the natural gradient algorithm to minimize $\mathcal{C}$ w.r.t. $\boldsymbol{\Theta}$ [28]. This gradient algorithm is also known as Fisher scoring approach since it weights the gradient direction by the Fisher information matrix of the parameters [10], [29]. The resulting descent algorithm updates the parameters in a Riemannian space, resulting in a fast convergence to a local minimum of the objective function [30], [31]. Denote as $\boldsymbol{\gamma} = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T, \boldsymbol{\theta}_3^T)^T$ the $(3M \times 1)$ vector gathering the altimetric parameters. The proposed algorithm updates this vector as follows:

$$\boldsymbol{\gamma}^{(t+1)} = \boldsymbol{\gamma}^{(t)} - \boldsymbol{F}^{-1}\left(\boldsymbol{\gamma}^{(t)}\right)\nabla\mathcal{C}(\boldsymbol{\gamma}^{(t)}) \tag{25}$$

where $\nabla\mathcal{C}(\boldsymbol{\gamma}^{(t)})$ denotes the gradient of the cost function $\mathcal{C}$ given by

$$\nabla\mathcal{C} = \frac{\partial\mathcal{C}}{\partial\boldsymbol{\gamma}} = \left[\left(\frac{\partial\mathcal{C}}{\partial\boldsymbol{\theta}_1}\right)^T, \left(\frac{\partial\mathcal{C}}{\partial\boldsymbol{\theta}_2}\right)^T, \left(\frac{\partial\mathcal{C}}{\partial\boldsymbol{\theta}_3}\right)^T\right]^T \tag{26}$$

with

$$\frac{\partial\mathcal{C}}{\partial\boldsymbol{\theta}_i} = -\left(\frac{\partial\boldsymbol{s}_1^T}{\partial\theta_{1i}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{x}_1, \ldots, \frac{\partial\boldsymbol{s}_M^T}{\partial\theta_{Mi}}\boldsymbol{\Sigma}_M^{-1}\boldsymbol{x}_M\right)^T + \left(a_i + \frac{M}{2}\right)\left(\frac{\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{\theta}_i}{\frac{\|\boldsymbol{D}\boldsymbol{\theta}_i\|^2}{2} + b_i}\right). \tag{27}$$

The Fisher information matrix $\boldsymbol{F}(\boldsymbol{\gamma}^{(t)})$ can be decomposed as the following block matrix:

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}_1 & \boldsymbol{F}_{12} & \boldsymbol{F}_{13} \\ \boldsymbol{F}_{12} & \boldsymbol{F}_2 & \boldsymbol{F}_{23} \\ \boldsymbol{F}_{13} & \boldsymbol{F}_{23} & \boldsymbol{F}_3 \end{bmatrix} \tag{28}$$

with

$$\boldsymbol{F}_{ij} = \operatorname{diag}\left(\frac{\partial\boldsymbol{s}_1^T}{\partial\theta_{1i}}\boldsymbol{\Sigma}_1^{-1}\frac{\partial\boldsymbol{s}_1}{\partial\theta_{1j}}, \ldots, \frac{\partial\boldsymbol{s}_M^T}{\partial\theta_{Mi}}\boldsymbol{\Sigma}_M^{-1}\frac{\partial\boldsymbol{s}_M}{\partial\theta_{Mj}}\right)$$

$$\boldsymbol{F}_i = \boldsymbol{F}_{ii} + \left(a_i + \frac{M}{2}\right)\left[\boldsymbol{D}^T\boldsymbol{D} - \frac{(\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{\theta}_i)(\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{\theta}_i)^T}{\left(\frac{\|\boldsymbol{D}\boldsymbol{\theta}_i\|^2}{2} + b_i\right)^2}\right]. \tag{29}$$

Note that the derivatives required to compute the elements of $\boldsymbol{F}$ are the same as those required to compute $\nabla\mathcal{C}$. Therefore, the only additional computation w.r.t. a steepest descent approach is the matrix inversion in (25). Note also that this matrix inversion is faster than a new evaluation of the altimetric model and its derivatives[8] that are necessary for a new update of $\boldsymbol{\gamma}$ in a steepest descent algorithm. Therefore, it is better to add this inversion and reduce the number of iterations, instead of using a simple steepest descent approach that usually requires a lot of iterations to converge, resulting in a high computational cost.

[8]The evaluation of the altimetric models and their derivatives requires a lot of computations mainly because of the involved nonlinear functions (such as the Gauss error or the exponential functions) and the numerical convolutions for CA and DDA.

2) Updating the Noise Parameters: Maximizing the marginalized posterior distribution (22) w.r.t. the noise parameters (means and variances) can be achieved by considering a coordinate descent algorithm. This approach maximizes the conditional distributions associated with each noise parameter. Using (15) and (17), it can be easily shown that the conditional distribution of the noise mean is the following Gaussian distribution:

$$\mu_m|\boldsymbol{y}_m, \boldsymbol{\Theta}_m, \boldsymbol{\Sigma}_m, \psi^2 \sim \mathcal{N}\left(\overline{\mu_m}, \frac{1}{\psi^{-2} + \sum_{k=1}^K \sigma_{mk}^{-2}}\right) \tag{30}$$

with

$$\overline{\mu_m} = \frac{\sum_{k=1}^K \frac{y_{mk} - s_{mk}}{\sigma_{mk}^2}}{\psi^{-2} + \sum_{k=1}^K \sigma_{mk}^{-2}}. \tag{31}$$

Therefore, the parameter vector $\boldsymbol{\mu}$ can be updated using (31), which is the maximum of the Gaussian distribution (30). Similarly, it can be shown using (15) and (18) that

$$f(\boldsymbol{\Lambda}|\boldsymbol{Y}, \boldsymbol{\Theta}, \boldsymbol{\mu}) = \prod_{n=1}^N \prod_{k=1}^K f\left(\sigma_{nr,k}^2|\boldsymbol{Y}_{:k}, \boldsymbol{\Theta}_{k:}, \boldsymbol{\mu}\right) \tag{32}$$

and that $\sigma_{nr,k}^2|\boldsymbol{Y}_{:k}, \boldsymbol{\Theta}_{k:}, \boldsymbol{\mu}$ is distributed according to the following inverse-gamma distribution:

$$\sigma_{nr,k}^2|\boldsymbol{Y}_{:k}, \boldsymbol{\Theta}_{k:}, \boldsymbol{\mu} \sim \mathcal{IG}\left(\frac{r}{2}, \beta\right) \tag{33}$$

with $\beta = \sum_{m=(n-1)r+1}^{nr} (x_{mk}^2/2)$. Thus, the mode of the inverse gamma distribution (33) is

$$\overline{\sigma_{nk}^2} = \frac{\beta}{\frac{r}{2} + 1}. \tag{34}$$

3) Stopping Criteria: Algorithm 1 is an iterative algorithm that requires the definition of some stopping criteria. In this paper, we have considered three criteria, and the algorithm is stopped if one of them is satisfied. The first criterion compares the new value of the cost function to the previous one and stops the algorithm if the relative error between these two values is smaller than a given threshold, i.e.,

$$\left|\mathcal{C}^{(t+1)} - \mathcal{C}^{(t)}\right| \leq \xi_1\mathcal{C}^{(t)} \tag{35}$$

where $|\cdot|$ denotes the absolute value. The second criterion evaluates the new parameter values and stops the algorithm if the following condition is satisfied [40]:

$$\left\|\boldsymbol{\gamma}^{(t+1)} - \boldsymbol{\gamma}^{(t)}\right\| \leq \xi_2\left(\left\|\boldsymbol{\gamma}^{(t)}\right\| + \xi_2\right). \tag{36}$$

The last criterion is based on a maximum number of iterations $T_{\max}$. Note that the proposed algorithm can be used when considering different altimetric models. Indeed, it just needs the definition of the altimetric model and its derivatives w.r.t.

the altimetric parameters (the altimetric model can be given by a closed-form expression as for BM or by a numerical computation as for CA and DDA). This model flexibility is highlighted in Section VI that considers the BM, CA, and DDA models for processing real data.

It is also interesting to note that the coordinate gradient algorithm converges to a stationary point if the minimum of the cost function (24) w.r.t. ($\boldsymbol{\Theta}$, $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}$) along each coordinate is uniquely reached [27]. This is easily checked for the mean vector $\boldsymbol{\mu}$ since the posterior distribution (22) viewed as a function of $\boldsymbol{\mu}$ is proportional to the Gaussian distribution (30) that has a unique mode. Similarly, the posterior distribution (22) viewed as a function of $\sigma^2_{nr,k}$ is proportional to the inverse gamma distribution (33) that has a unique mode. Considering the altimetric parameters, the convexity of the cost function cannot be theoretically demonstrated mainly because of the nonlinearity of the altimetric models. Therefore, the solution obtained by the considered local estimator might depend on the initial values of the parameters. Thus, we have considered the same realistic values as for [9] and [15] to initialize the parameters (SWH $= 2.9$ m, $\tau = 14.97$ m, and $P_u = 140$) in all what follows (see [41] for more details regarding the initialization of the parameters). Using this initialization, the proposed algorithm always converges to the actual parameters in all of our synthetic simulations (see Section V).

V. Validation on Synthetic Data

This section first introduces the criteria used to evaluate the performances of the proposed estimation algorithm. The second part defines the effective number of looks (ENL) used to exploit the estimated noise covariance. The third part of this section compares the performance of the proposed algorithm with those obtained with state-of-the-art approaches for synthetic data with controlled ground truth.

A. Evaluation Criteria

For synthetic echoes, the true parameters are supposed to be known and can be used to evaluate the estimation performance. More precisely, the bias and STD of the estimator $\widehat{\theta}_i$ are defined as follows:

$$\text{Bias}\left(\widehat{\boldsymbol{\theta}}_i\right) = \frac{1}{M}\sum_{m=1}^{M}\left[\widehat{\theta}_i(m) - \theta_i(m)\right] \tag{37}$$

$$\text{STD}\left(\widehat{\boldsymbol{\theta}}_i\right) = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left[\widehat{\theta}_i(m) - \theta_i(m)\right]^2} \tag{38}$$

for $i \in \{1, \ldots, 3\}$, where $\theta_i(m)$ (respectively, $\widehat{\theta}_i(m)$) is the true (respectively, estimated) parameter for the mth waveform and M is the number of simulated waveforms.

B. ENL

The proposed algorithm estimates the noise covariance $\hat{\boldsymbol{\Lambda}}$ which is used to compute the ENL defined in [5], [16]. This number represents the equivalent number of independent echoes that were averaged to reduce the noise and is defined by

$$N_{\text{eff}}(n,k) = \frac{E^2\left[y(k)\right]}{E\left\{\left[y(k) - E\left(y(k)\right)\right]^2\right\}} = \frac{\left[1/r\sum_{m=(n-1)r+1}^{nr} y_m(k)\right]^2}{\widehat{\sigma^2}_{nr,k}} \tag{39}$$

for $k \in \{1, \ldots, K\}$ and $n \in \{1, \ldots, N\}$. Note that $N_{\text{eff}}(n,k)$ differs from one temporal gate to another and thus depends on k. In order to avoid this dependence, the following averaged number of looks is considered in this paper:

$$\text{ENL}(n) = \frac{1}{K}\sum_{k=1}^{K} N_{\text{eff}}(n,k) \tag{40}$$

for $n \in \{1, \ldots, N\}$. In the case of synthetic data, this estimation will be compared to the number L of averaged echoes in terms of bias and STD.

TABLE I
Performance on Synthetic BM Data (500 Echoes)

		SWH (cm)	τ (cm)	P_u	μ $\times 10^{(-4)}$	ENL
Bias	LS-BM	3.18	1.1	**-0.01**	0.51	-
	CD-BM	0.32	**0.08**	-0.2	0.26	0.97
	HMC-BM	**-0.02**	-0.13	-0.2	-0.1	0.93
STD	LS-BM	44.7	6.1	1.91	11	-
	CD-BM	**2.72**	**1.1**	0.62	12	4.47
	HMC-BM	5.74	1.8	**0.52**	6	4.46
Average time per echo in milliseconds	LS-BM	8.9				
	CD-BM	**3.6**				
	HMC-BM	739.7				

C. Simulation Results on Synthetic Data

The proposed strategy (denoted by CD for coordinate descent) is first investigated when considering $M = 500$ correlated altimetric echoes generated according to the BM defined in (5). The correlation between successive echoes is introduced by considering a smooth evolution of the altimetric parameters. More precisely, we have considered a dynamic evolution of the altimetric parameters defined as follows: SWH$(m) = 2.5 + 2\cos(0.07\text{ m})$, $\tau(m) = 27 + 0.02$ m if $m < 250$ and $\tau(m) = 32 - 0.02$ m if $m \geq 250$ (in unit of gates), and $P_u(m) = 158 + 0.05\sin(0.1\text{ m})$, where m denotes the echo number. The synthetic echoes are then generated by plugging the values of the altimetric parameters in the BM and by corrupting the resulting echo by a thermal noise $\mu_m = 0.025\ \forall m$ and a speckle noise resulting from the averaging of $L = 90$ echoes. The proposed CD strategy is compared to the state-of-the-art LS algorithm described in [6], [9], and [10] and to the Bayesian approach HMC [25] detailed in Appendix A. Table I reports the obtained

biases and STDs when estimating the three parameters of interest and ENL when possible. The considered algorithms are denoted by their names (CD, HMC, and LS), followed by the considered altimetric model (BM, CA, and DDA). For instance, CD-BM represents the CD algorithm when considering BM echoes. Table I shows the reduced biases for τ and SWH when considering the Bayesian CD and HMC methods. Indeed, these two algorithms estimate the non-identically distributed noise, while the LS estimator only estimates the thermal noise and assumes an i.i.d. speckle noise. Thus, the Bayesian algorithms involve a more precise model and provide a lower bias for the altimetric parameters (except for P_u). Note that all algorithms provide a good estimate of the thermal noise $\boldsymbol{\mu}$ as shown in Table I. Moreover, from an STD point of view, both Bayesian CD and HMC approaches lead to lower STDs than when using an LS-based method since they are using the smooth prior defined in (16). Note that CD-BM improves the LS-BM STDs by a factor of 16 for SWH, 5 for τ, and 3 for P_u. Table I also highlights the additional interest of both Bayesian algorithms that are able to estimate the number of averaged independent echoes with a good precision (small bias and STD). Last but not least, Table I gives the averaged computational cost of the three algorithms, for a MATLAB implementation on a 2.4-GHz Intel-i7 personal computer. As expected, the HMC algorithm is highly computationally expensive due to its sampling procedure, while the proposed CD algorithm is the fastest one, which is three times faster than the LS one. Results similar to Table I have been obtained when considering CA and DDA models, but these results are not provided in this section for brevity (these models are investigated in the next section when considering CryoSat-2 echoes). Finally, in the case of synthetic data, the proposed CD approach outperforms the classical LS one, both in terms of bias and STD for the estimated parameters but also in terms of computational cost, allowing, at the same time, the estimation of additional noise parameters. Fig. 2 illustrates the clear improvement of the proposed algorithm when compared to the LS one, in terms of bias and STD. Subfigures (a-top), (c-top), and (e-top) present the actual parameter values (black line) and the estimated ones by considering the LS (red line) and the CD (blue line, almost confused with the black one) algorithms for 500 echoes. When plotting the difference between the actual and estimated parameters in subfigures (a-bottom), (b-bottom), and (c-bottom), the interest of the proposed CD strategy is clearly highlighted.

VI. Results on Real Data

In order to evaluate the performance of the proposed algorithm on real data, we first present our comparison criteria. Then, two kinds of real data are considered: Jason-2 echoes for which BM is applied and CryoSat-2 for which CA and DDA models are more appropriate. Note that the objective of this section is to show the potential of the proposed approach when considering example echoes from CA and DDA. More involved studies should be conducted in the future to better understand the behavior of this algorithm when considering more real data with different ocean characteristics and possible ground truth.

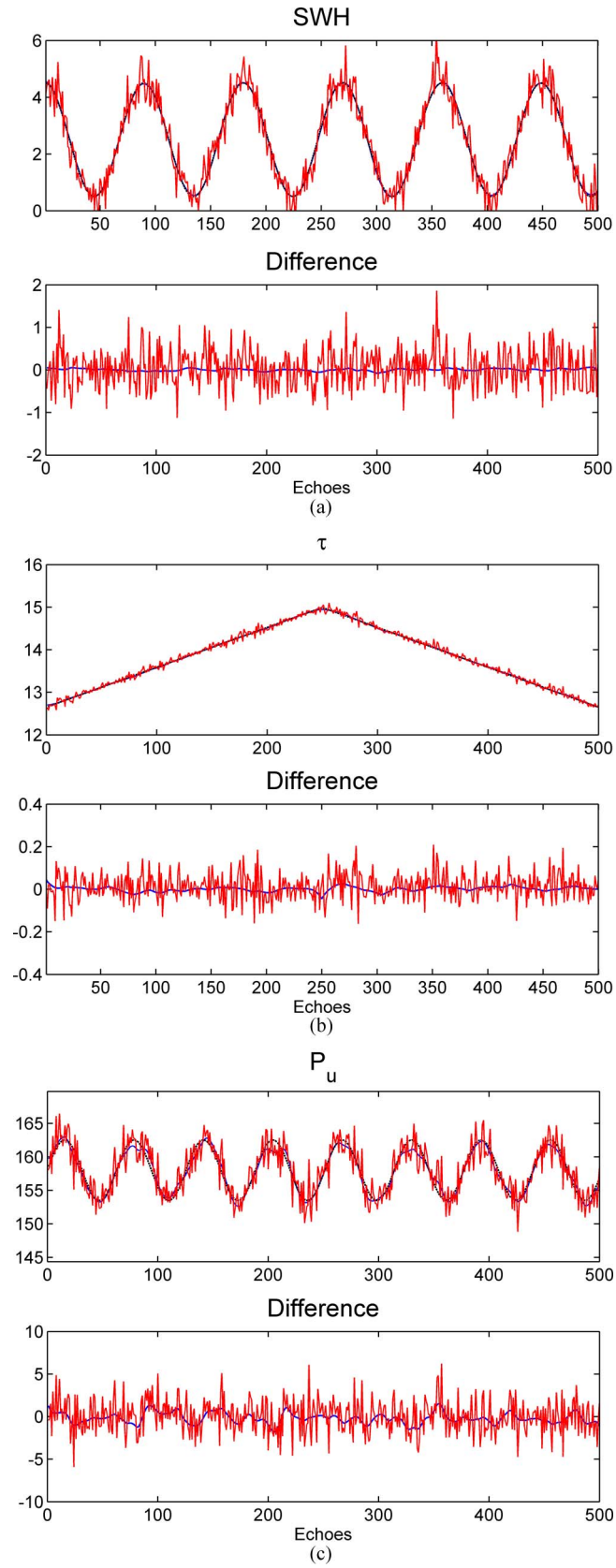

Fig. 2. Actual (black line) and estimated parameters using the LS algorithm (red line) and the proposed CD-BM algorithm (blue line). (a) SWH. (b) τ. (c) P_u. The bottom subfigures show the difference between the estimated and real parameters.

A. Evaluation Criteria

When processing real data without ground truth, it is not possible to compare the performance of the different parameter estimation algorithms using the bias and STD defined in (37) and (38) since the true parameter values are not known. In the case of real waveforms, it is usual to compare the different parameter estimations along a pass by plotting the different obtained results [9], [13] and by considering a modified STD. The modified STD is computed using (38), in which the true parameter value is approached by the mean of the estimated parameters along 20 successive echoes. This modified STD is called "STD at 20 Hz" [6], [42], [43].

B. Analysis of Jason-2 Data With the Brown Model

This section illustrates the performance of the proposed CD algorithm when applied to a real oceanic Jason-2 data set. The considered data last 36 min and consist of 43 000 real echoes that were extracted from the pass 30 of cycle 35. Fig. 3 shows the parameters estimated on 700 successive echoes along the pass when considering the LS-BM (in red), HMC-BM (in green), and CD-BM (in blue) algorithms. As observed for synthetic data in Section V, the LS-BM estimates present a high estimation noise, while both CD and HMC algorithms provide smoother estimates which are physically more consistent. Moreover, the CD and HMC algorithms appear to be more robust to outliers as illustrated for the estimate #890 of P_u (see [41] for more examples). Note that the estimated SWH is slightly larger for the Bayesian algorithms when compared to LS. This difference can be explained by the i.i.d. noise assumption used in the LS algorithm which is not in adequation with the considered data as already discussed in [13] and [16]. The reduction of estimation variance provided by the Bayesian algorithms is particularly noticeable when plotting the spectral power density of SWH and τ as in Fig. 4. This figure shows a clear noise reduction especially for ocean scales smaller than 100 km for SWH and 10 km for τ (note that we have not applied atmospheric and instrumental corrections on the estimated epoch τ) which improves the observation of these scales. More quantitative results are provided in Table II. This table shows a good agreement between the means of the estimated parameters for the LS, HMC, and CD algorithms (except for SWH as explained previously). Bayesian algorithms also estimate the ENL which provides an indication about the noise level corrupting the data. The estimated ENL ranges between 70 and 80 echoes, which is close to the actual value (around 90 or 100 echoes). As for synthetic data, the estimated STDs obtained with the CD algorithm are smaller than those for LS and HMC, which is of great importance for many practical applications related to oceanography such as bathymetry. Indeed, the CD algorithm improves the LS-BM STDs by a factor of about 20 for SWH, 2 for τ, and 3 for P_u. The improvement obtained with the CD algorithm with respect to the HMC algorithm is mainly due to robustness introduced by the marginalization step (that has not been considered for HMC) and to the optimization (by cross-validation) of the hyperparameters a_i and b_i for the CD algorithm (while set to $a_i = b_i = 0$ for HMC). Note that the proposed CD algorithm requires a reduced computational

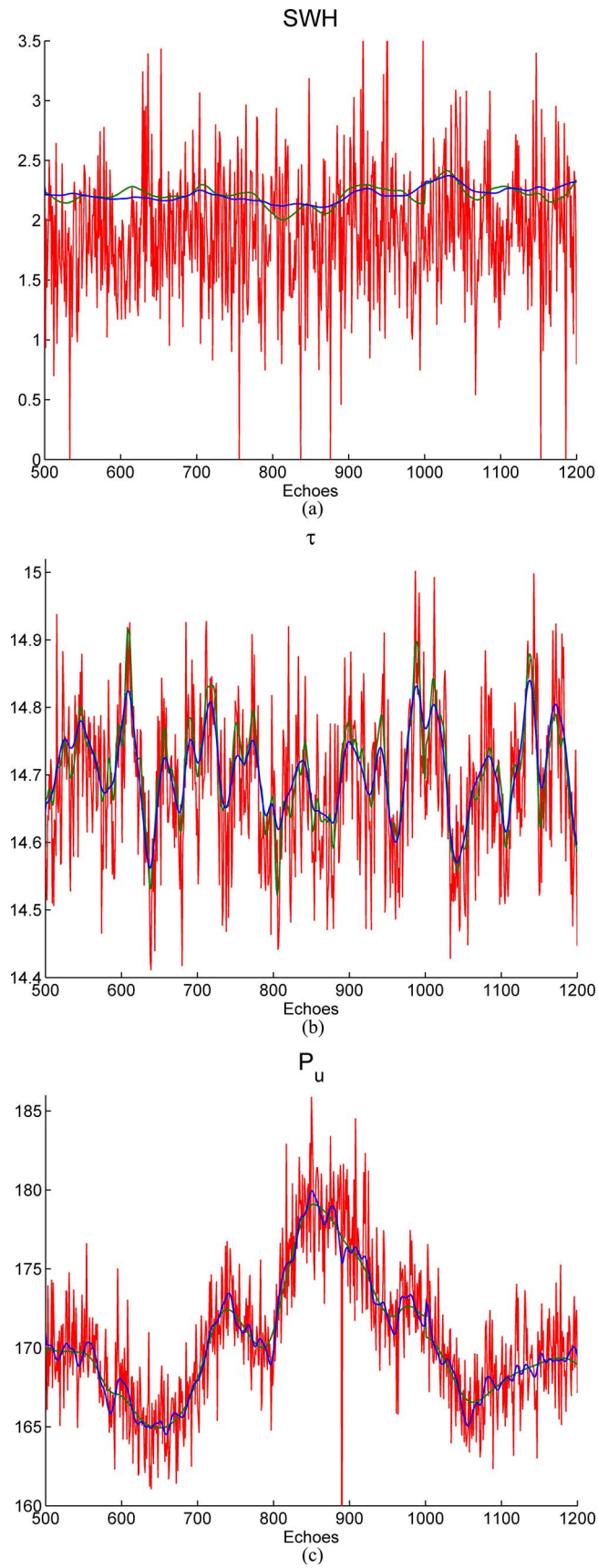

Fig. 3. Estimated parameters using the LS-BM algorithm (red line), the HMC-BM algorithm (green line), and the CD-BM algorithm (blue line) for Jason-2 echoes. (a) SWH. (b) τ. (c) P_u.

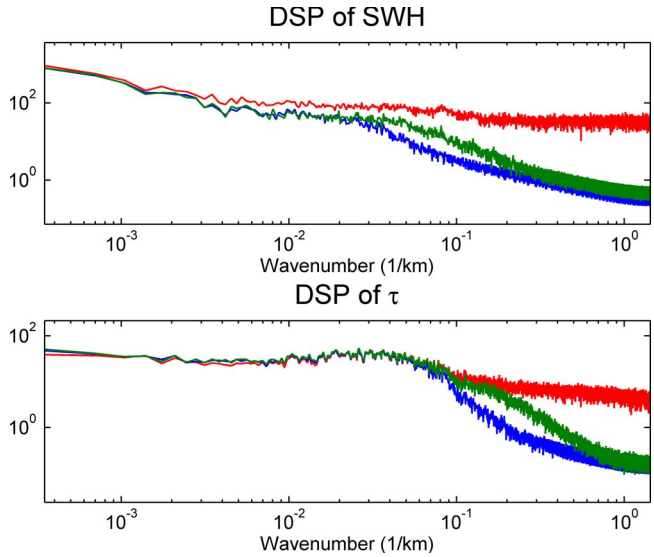

Fig. 4. Estimated spectra of SWH and τ using 43 000 Jason-2 echoes with the LS algorithm (red line), the HMC-BM algorithm (green line), and the CD-BM algorithm (blue line).

TABLE II
PERFORMANCE ON REAL JASON-2 DATA (43 000 ECHOES)

		SWH (cm)	τ (cm)	P_u	μ	ENL
Mean	LS-BM	237	14.67	167.81	0.97	-
	CD-BM	272	14.69	167.52	1.06	74
	HMC-BM	270	14.69	167.53	1.06	80
STD	LS-BM	55.5	11.23	8.18	0.12	-
	CD-BM	**2.37**	**5.57**	**2.79**	0.08	15
	HMC-BM	5.58	7.11	7.64	0.08	9.5
Average time per echo in milliseconds	LS-BM	12.5				
	CD-BM	**6.6**				
	HMC-BM	1142.3				

cost that is two times lower than that of the LS algorithm. As previously explained, this time reduction is due to the parallel scheme considered by the CD algorithm which is more efficient than the independent processing of echoes performed in the LS algorithm. Note finally that one of the main recent concerns in radar altimetry is the retrieval of accurate information near the coast. This goal can be achieved by the proposed CD algorithm when considering a dedicated coastal model such as [10] and [44], which is an interesting issue that will be studied in future works (see [41] for more details regarding the parameter estimation of coastal echoes using the CD-BM algorithm).

C. Comparison Between CA and DDA for CryoSat-2 Data

This section evaluates the performance of the proposed CD algorithm for real oceanic CryoSat-2 waveforms. The CryoSat-2 altimeter called SIRAL presents three modes, which are the following: the low-resolution mode (LRM), the synthetic aperture radar mode (SARM), and the synthetic aperture radar interferometric mode [45]. This paper considers data from the SARM that are used to generate CA[9] and DDA echoes (see [6] for more details about these echoes). The considered data set lasts approximately 100 s (i.e., 2000 echoes), and it was obtained in August 2011 by the CryoSat processing prototype (CPP) developed by CNES [47] (which is doing the level 1 processing). Note that the CNES-CPP uses data preprocessed to full bit rate provided by ESA [45]. For all algorithms, we have considered the same instrumental parameters as in [6] and [7], which are given in Table III. Table IV compares the estimated parameters using the LS algorithm with the CD and HMC algorithms when considering both CA and DDA echoes. To compare with [6], we only present the estimation results of SWH and of the sea surface height anomaly (SSHA) which are the most important parameters (SSHA is obtained by applying the same environmental corrections as in [6] on the estimated epoch τ).[10] Figs. 5 and 6 show the evolution of the estimated SWH and SSHA parameters when considering the three studied algorithms. These figures show a good agreement between the estimated parameters especially between the CD and HMC algorithms, which is confirmed by the mean of the estimated parameters provided in Table IV. Note that, as for Jason-2, a bias seems to appear between the estimated SWH parameter for Bayesian algorithms and the LS approach, which can be explained similarly. Considering the STDs and to compare the results of Table IV with [6], one has to note that the results of [6] were obtained when using the LS algorithm for both CA and DDA. The processing of the CryoSat-2 data further highlights the main advantage of the proposed CD approach which is to reduce drastically the STDs of the estimated parameters. For example, STDs(SWH) [respectively, STD(SSHA)] goes from 91 cm (respectively, 14 cm) for LS-CA to 3 cm (respectively, 4 cm) for CD-CA, which represents a great improvement. The same behavior is observed when considering DDA echoes showing the interest of smoothing the altimetric parameters. The second advantage of the CD algorithm is the reduction of the computational cost especially for the DDA echoes. Indeed, Table IV highlights a time reduction by a factor of 20 w.r.t. LS-DDA that processed the echoes independently while CD-DDA uses a parallel scheme. For the CA echoes, the CD approach shows a time improvement by a factor of 1.5, which is an interesting result. Note finally that the HMC algorithm shows a similar performance with that of the CD algorithm but at a price of a higher computational cost. These results confirm the good performance of the proposed algorithm and show its flexibility w.r.t. the considered altimetric models.

TABLE III
INSTRUMENTAL PARAMETERS

Parameter	Value
Frequency	13.575 GHz
Wavelength (λ)	2.21 cm
Bandwidth (B)	320 MHz
Altitude (h)	730 km
Burst repetition frequency (BRF)	85 Hz
Pulse repetition frequency (PRF)	18182 Hz
3 dB Antenna beam width ($\theta_{3\text{dB}}$)	1.1388 degrees
Velocity (v_s)	7000 m/s
Pulses per burst	64 pulses
Burst length (τ_b)	3.5 ms
Doppler beam width	327 m

[9]These echoes are known under different names: LRM-like [35], pseudo-LRM [18], [46], [47], reduced-SAR [46], [47], or CA-SARM in [6]. These echoes are used to compare the parameters of CA with those of DDA.

[10]We have considered the SSHA instead of τ since different corrections should be applied to CA and DDA. The reader is invited to consult [11] and [48] for more details regarding SSHA computation. The SSB have not been considered for both CA and DDA.

TABLE IV
COMPARISON BETWEEN THE LS, CD, AND HMC ALGORITHMS WHEN CONSIDERING REAL CA AND DDA DATA (2000 CRYOSAT-2 ECHOES)

		CA echoes			DDA echoes		
		LS-CA	CD-CA	HMC-CA	LS-DDA	CD-DDA	HMC-DDA
Means	SWH (m)	2.39	2.59	2.53	2.31	2.70	2.70
	SSHA (m)	-5.12	-5.20	-5.19	-5.13	-5.20	-5.21
STDs at 20 Hz	SWH (cm)	91	**3**	20	45	**3**	8
	SSHA (cm)	14	**4**	**4**	6	**2**	**2**
Average time per echo (s)		0.1	**0.07**	252	0.79	**0.04**	233

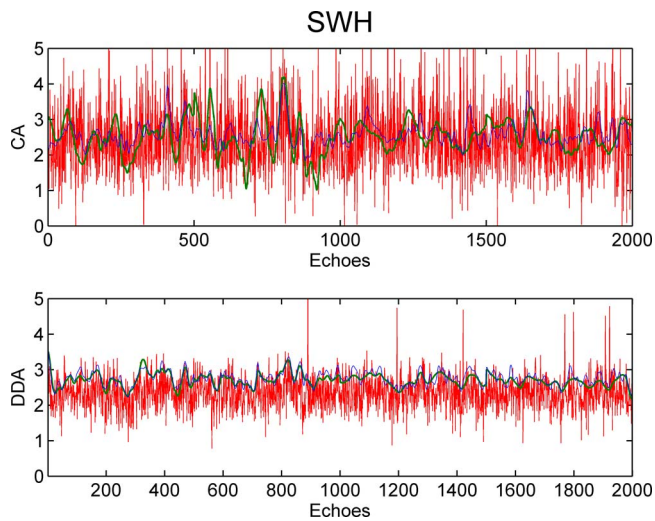

Fig. 5. Estimated SWH using 2000 CryoSat-2 echoes with the LS algorithm (red line), the HMC algorithm (green line), and the CD algorithm (blue line). (Top) CA echoes. (Bottom) DDA echoes.

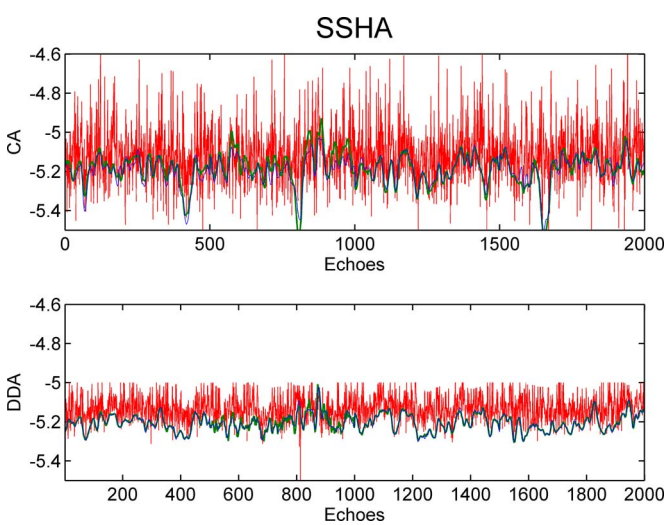

Fig. 6. Estimated SSHA using 2000 CryoSat-2 echoes with the LS algorithm (red line), the HMC algorithm (green line), and the CD algorithm (blue line). (Top) CA echoes. (Bottom) DDA echoes.

VII. CONCLUSION

This paper has proposed a new Bayesian strategy for a smooth parameter estimation of the noise and altimetric parameters. The altimetric parameters were assigned a prior distribution enforcing a smooth temporal evolution. The altimetric signal was assumed to be corrupted by a thermal and speckle noise distributed according to an independent and nonidentically Gaussian distribution. The parameter estimation was achieved using a coordinate descent algorithm that presents a low computational cost which is suitable for real-time applications. The proposed CD algorithm showed good performance and improved the quality of the estimated parameters when applied to both synthetic and real altimetric echoes from conventional and DDA. More precisely, it provided reduced parameter STDs and presented a lower computational cost when compared to the state-of-the-art LS algorithm. Moreover, the CD algorithm is generic, and it showed its efficiency when considering three altimetric models. Future work includes the consideration of other altimetric models such as the DDA-SAMOSA model [49], [50], the coastal models [10], [44], and the CA model based on the true PTR [32], [33]. Considering antenna mispointing is also an interesting issue that is currently under investigation.

APPENDIX
MATHEMATICAL DERIVATIONS

A. Hybrid Gibbs Algorithm

The principle of the Gibbs sampler is to sample according to the conditional distributions of the posterior of interest [51]. The MCMC-based algorithm proposed in [25] uses this principle to sequentially sample the parameters $\boldsymbol{\Theta}$, $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}$, and $\boldsymbol{\epsilon}$. When a conditional distribution cannot be sampled directly, sampling techniques such as the HMC algorithm can be applied (see [26] and [52] for more details about this algorithm). This algorithm has shown better mixing property than independent or random walk Metropolis–Hasting moves especially in high-dimensional problems [26], [52]. Therefore, it has been considered in [25] since the vectors to sample are of size $(M \times 1)$. The Gibbs sampler (including the HMC procedures) used to sample according to the posterior (20) consists of the four steps summarized in Algorithm 2.

Algorithm 2 Hybrid Gibbs Sampler

1: Initialization
2: Initialize parameters $\boldsymbol{\theta}_i^{(0)}$, for $i \in \{1, \ldots, 3\}$, $\boldsymbol{\mu}^{(0)}$, $\boldsymbol{\Lambda}^{(0)}$ and $\boldsymbol{\epsilon}^{(0)}$
3: Update parameters/hyperparameters
4: **for** $t = 1 : N_{\text{bi}} + N_r$ **do**
5: Sample $\boldsymbol{\theta}_i^{(t)}$, for $i \in \{1, 2, 3\}$ according to (41) using HMC
6: Sample $\boldsymbol{\mu}^{(t)}$ according to (30)
7: Sample $\boldsymbol{\Lambda}^{(t)}$ according to (33)
8: Sample $\boldsymbol{\epsilon}^{(t)}$ according to (42)
9: **end for**

This algorithm uses the conditional distribution associated with the altimetric parameters and the hyperparameters that are obtained by considering the likelihood (15) and the priors (16) and (19), leading to

$$f(\boldsymbol{\theta}_i|\boldsymbol{Y},\boldsymbol{\Theta}_{\backslash i},\boldsymbol{\mu},\boldsymbol{\Lambda},\epsilon_i^2)\propto\exp\left(-\frac{1}{2\epsilon_i^2}\|\boldsymbol{D}\boldsymbol{\theta}_i\|^2-\sum_{m=1}^{M}\frac{\boldsymbol{x}_m^T\boldsymbol{\Sigma}_m^{-1}\boldsymbol{x}_m}{2}\right) \tag{41}$$

$$\epsilon_i^2|\boldsymbol{Y},\boldsymbol{\Theta},\boldsymbol{\mu},\boldsymbol{\Sigma}\sim\mathcal{IG}\left(\frac{M}{2}+a_i,\frac{\|\boldsymbol{D}\boldsymbol{\theta}_i\|^2}{2}+b_i\right) \tag{42}$$

for $i\in\{1,\ldots,3\}$. Note that Algorithm 2 draws $N_{\mathrm{MC}}=N_{\mathrm{bi}}+N_r$ samples that are used to approximate the MMSE estimator of the parameters after removing the first N_{bi} burn-in samples. This algorithm benefits from the robust properties of the MCMC methods and allows the estimation of the hyperparameters. However, it presents a high computational cost that prevents its use for real-time applications.

Acknowledgment

The authors would like to thank the editor and anonymous reviewers for their valuable suggestions and comments and CLS, CNES, and ESA for providing real altimetric data.

References

[1] P. Vincent *et al.*, "AltiKa: A Ka-band altimetry payload and system for operational altimetry during the GMES period," *Sensors*, vol. 6, no. 3, pp. 208–234, Mar. 2006.

[2] J. Richard *et al.*, "An advanced concept of radar altimetry over oceans with improved performances and ocean sampling: AltiKa," in *Proc. IEEE IGARSS*, Jul. 2007, pp. 3537–3540.

[3] R. K. Raney and L. Phalippou, "The future of coastal altimetry," in *Coastal Altimetry*, S. Vignudelli, A. G. Kostianoy, P. Cipollini, and J. Benveniste, Eds. Berlin, Germany: Springer-Verlag, 2011, pp. 61–101.

[4] R. K. Raney, "The delay/Doppler radar altimeter," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 5, pp. 1578–1588, Sep. 1998.

[5] D. Wingham, L. Phalippou, C. Mavrocordatos, and D. Wallis, "The mean echo and echo cross product from a beamforming interferometric altimeter and their application to elevation measurement," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2305–2323, Oct. 2004.

[6] A. Halimi, C. Mailhes, J.-Y. Tourneret, P. Thibaut, and F. Boy, "A semi-analytical model for delay/Doppler altimetry and its estimation algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4248–4258, Jul. 2014.

[7] A. Halimi, C. Mailhes, J.-Y. Tourneret, T. Moreau, and F. Boy, "Including antenna mispointing in a semi-analytical model for delay/Doppler altimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 598–608, Feb. 2015.

[8] G. Brown, "The average impulse response of a rough surface and its applications," *IEEE Trans. Antennas Propag.*, vol. AP-25, no. 1, pp. 67–74, Jan. 1977.

[9] L. Amarouche *et al.*, "Improving the Jason-1 ground retracking to better account for attitude effects," *Mar. Geod.*, vol. 27, no. 1/2, pp. 171–197, Aug. 2004.

[10] A. Halimi, C. Mailhes, J.-Y. Tourneret, P. Thibaut, and F. Boy, "Parameter estimation for peaky altimetric waveforms," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 3, pp. 1568–1577, Mar. 2013.

[11] A. Halimi, "From conventional to delay/doppler altimetry," Ph.D. dissertation, Dept. Signal Process. Inst. Nat. Polytech. de Toulouse, Toulouse, France, 2013. [Online]. Available: http://ethesis.inp-toulouse.fr/archive/00002423/01/halimi.pdf

[12] S. Maus, C. M. Green, and J. D. Fairhead, "Improved ocean-geoid resolution from retracked ERS-1 satellite altimeter waveforms," *Geophys. J. Int.*, vol. 134, no. 1, pp. 243–253, Jul. 1998.

[13] D. T. Sandwell and W. H. F. Smith, "Retracking ERS-1 altimeter waveforms for optimal gravity field recovery," *Geophys. J. Int.*, vol. 163, no. 1, pp. 79–89, Oct. 2005.

[14] J. Severini, C. Mailhes, J.-Y. Tourneret, and P. Thibaut, "Bayesian estimation of altimeter echo parameters," in *Proc. IEEE IGARSS*, vol. 3, Boston, MA, USA, Jul. 2008, pp. 238–241.

[15] J.-P. Dumont, "Estimation optimale des paramètres altimétriques des signaux radar Poséidon," Ph.D. dissertation, Dept. Signal Process. Inst. Nat. Polytech. de Toulouse, Toulouse, France, 1985.

[16] A. Halimi, C. Mailhes, and J.-Y. Tourneret, "Cramér–Rao bounds and estimation algorithms for delay/Doppler and conventional altimetry," in *Proc. EUSIPCO*, Marrakech, Marocco, Sep. 1–9, 2013, pp. 1–5.

[17] E. Rodriguez and J. Martin, "Correlation properties of ocean altimeter returns," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 3, pp. 553–561, May 1994.

[18] W. H. F. Smith and R. Scharroo, "Pulse-to-pulse correlation in CryoSat SAR echoes from ocean surfaces: Implications for optimal pseudo-LRM waveform averaging," in *Proc. OSTST*, Venice-Lido, Italy, Sep. 2012, pp. 22–29.

[19] W. Smith and R. Scharroo, "Waveform aliasing in satellite radar altimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1671–1682, Apr. 2015.

[20] C. Martin-Puig and G. Ruffini, "SAR altimeter retracker performance bound over water surfaces," in *Proc. IEEE IGARSS*, Cape Town, South Africa, Jul. 12–17, 2009, pp. 449–452.

[21] P. Campisi and K. Egiazarian, *Blind Image Deconvolution: Theory and Applications*. Boca Raton, FL, USA: Taylor & Francis, 2007.

[22] R. Molina, J. Mateos, and A. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3715–3727, Dec. 2006.

[23] J. Sigurdsson, M. Ulfarsson, and J. Sveinsson, "Hyperspectral unmixing with l_q regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 6793–6806, Nov. 2014.

[24] L. Chaari, T. Vincent, F. Forbes, M. Dojat, and P. Ciuciu, "Fast joint detection-estimation of evoked brain activity in event-related FMRI using a variational approach," *IEEE Trans. Image Process.*, vol. 32, no. 5, pp. 821–837, May 2013.

[25] A. Halimi, C. Mailhes, and J.-Y. Tourneret, "Nonlinear regression using smooth Bayesian estimation," in *Proc. IEEE ICASSP*, Brisbane, QLD, Australia, Apr. 2015, pp. 2634–2638.

[26] S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng, *Handbook of Markov Chain Monte Carlo*, ser. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. New York, NY, USA: Taylor & Francis, 2011.

[27] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 1995.

[28] S. Amari and S. Douglas, "Why natural gradient?" in *Proc. IEEE ICASSP*, May 1998, vol. 2, pp. 1213–1216.

[29] P. J. Green, "Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives," *J. Roy. Stat. Soc., Ser. B (Methodological)*, vol. 46, no. 2, pp. 149–192, 1984.

[30] M. Girolami and B. Calderhead, "Riemann manifold Langevin and Hamiltonian Monte Carlo methods," *J. R. Stat. Soc., Ser. B*, vol. 73, no. 2, pp. 123–214, Mar. 2011.

[31] M. Pereyra, H. Batatia, and S. McLaughlin, "Exploiting information geometry to improve the convergence properties of variational active contours," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 4, pp. 700–707, Aug. 2013.

[32] L. Amarouche *et al.*, "Reduction of the sea surface height spectral hump using a new retracker decorrelating ocean estimated parameters (DCORE)," in *Proc. OSTST*, Lake Constance, Germany, Oct. 27–31, 2014, pp. 1–26.

[33] F. Boy *et al.*, "Towards the Jason-3 waveform processing: Assessment of the numerical retracking performances," in *Proc. OSTST*, Lake Constance, Germany, Oct. 27–31, 2014, pp. 1–13.

[34] G. Hayne, "Radar altimeter mean return waveforms from near-normal-incidence ocean surface scattering," *IEEE Trans. Antennas Propag.*, vol. AP-28, no. 5, pp. 687–692, Sep. 1980.

[35] L. Phalippou and F. Demeestere, "Optimal re-tracking of SAR altimeter echoes over open ocean: From theory to results for SIRAL2," in *Proc. OSTST*, San Diego, CA, USA, Oct. 19–21, 2011, pp. 1–18.

[36] E. J. Walsh., "Pulse-to-pulse correlation in satellite radar altimeters," *Radio Sci.*, vol. 17, no. 4, pp. 786–800, Aug. 1982.

[37] O. Germain and G. Ruffini, "A revisit to the GNSS-R code range precision," in *Proc. GNSS-R*, Noordwijk, The Netherlands, Jun. 14–15, 2006, pp. 1–5.
[38] C. Martin-Puig *et al.*, "SAR Altimetry Over Water Surfaces," presented at the Oceans Space, Venice, Italy, Apr. 26–30, 2010.
[39] V. Mazet, "Joint Bayesian decomposition of a spectroscopic signal sequence," *IEEE Signal Process. Lett.*, vol. 18, no. 3, pp. 181–184, Mar. 2011.
[40] K. Madsen, H. B. Nielsen, and O. Tingleff, "Methods for non-linear least squares problems (2nd ed.)," Informat. Math. Modell., Tech. Univ. Denmark, Kgs. Lyngby, Denmark, 2004, p. 60.
[41] A. Halimi, C. Mailhes, J.-Y. Tourneret, and H. Snoussi, "Bayesian Estimation of Smooth Altimetric Parameters: Application to Conventional and Delay/Doppler Altimetry," Univ. Technol. Troyes, Troyes, France, Tech. Rep., Jun. 2015.
[42] K. Giles, D. Wingham, N. Galin, R. Cullen, and W. Smith, "Precise estimates of ocean surface parameters from CryoSat," in *Proc. OSTST*, Venice-Lido, Italy, Sep. 22–29, 2012, pp. 1–13.
[43] C. Gommenginger *et al.*, "Improved altimetric accuracy of SAR altimeters over the ocean: Observational evidence from CryoSat-2 SAR and Jason-2," in *Proc. OSTST*, San Diego, CA, USA, Oct. 19–21, 2011, pp. 1–35.
[44] J. Gómez-Enri *et al.*, "Modeling Envisat RA-2 waveforms in the coastal zone: Case study of calm water contamination," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 3, pp. 474–478, Jul. 2010.
[45] D. Wingham *et al.*, "CryoSat: A mission to determine the fluctuations in Earth's land and marine ice fields," *Adv. Space Res.*, vol. 37, no. 4, pp. 841–871, Jul. 2006.
[46] C. Gommenginger, P. Cipollini, D. Cotton, S. Dinardo, and J. Benveniste, "Finer, better, closer: Advanced capabilities of SAR altimetry in the open ocean and the coastal zone," in *Proc. OSTST*, Venice-Lido, Italy, Sep. 22–29, 2012, pp. 1–23.
[47] F. Boy *et al.*, "CryoSat processing prototype, LRM and SAR processing on CNES side and a comparison to DUACS SLA," in *Proc. Report OSTST*, Venice-Lido, Italy, Sep. 22–29, 2012, pp. 1–24.
[48] S. Vignudelli, A. G. Kostianoy, P. Cipollini, and J. Benveniste, Eds. *Coastal Altimetry*. Berlin, Germany: Springer-Verlag, 2011.
[49] C. Martin-Puig *et al.*, "Theoretical model of SAR altimeter over water surfaces," in *Proc. IEEE IGARSS*, Boston, MA, USA, Jul. 7–11, 2008, pp. 242–245.
[50] C. Gommenginger, C. Martin-Puig, S. Dinardo, D. Cotton, and J. Benveniste, "Improved altimetric performance of CryoSat-2 SAR mode over the open ocean and the coastal zone," presented at the IEEE Int. Conf. Geoscience Remote Sensing, Munich, Germany, Jul. 22–27, 2012.
[51] C. P. Robert, *The Bayesian Choice: From Decision-Theoretic Motivations to Computational Implementation*, 2nd ed., ser. Springer Texts in Statistics. New York, NY, USA: Springer-Verlag, 2007.
[52] A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Unsupervised unmixing of hyperspectral images accounting for endmember variability," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 1904–4917, Dec. 2015.

Abderrahim Halimi (S'11–M'14) received the Eng. degree in electronics from the National Polytechnic School of Algiers, Algiers, Algeria, in 2009 and the M.Sc. and Ph.D. degrees in signal processing from the Institut National Polytechnique de Toulouse, Toulouse, France, in 2010 and 2013, respectively.

From October 2013 to September 2014, he was a Postdoctoral Research Associate with the Institut de Recherche en Informatique de Toulouse, University of Toulouse, Toulouse. He is currently a Postdoctoral Research Associate with the University of Technology of Troyes, Troyes, France. His research activities focus on statistical signal and image processing, with a particular interest in Bayesian inverse problems with applications to hyperspectral imaging and altimetry.

Corinne Mailhes (M'87) received the Ph.D. degree in signal processing from the University of Toulouse, Toulouse, France, in 1990.

She is currently a Professor with the University of Toulouse, a member of the Institut de Recherche en Informatique de Toulouse, and the Director of the Telecommunications for Space and Aeronautics Laboratory. Her research interests include statistical signal processing, with particular interests in spectral analysis, parameter estimation, and biomedical signal processing.

Jean-Yves Tourneret (SM'08) received the Ingénieur degree in electrical engineering from the Ecole Nationale Supérieure d'Electronique, d'Electrotechnique, d'Informatique, d'Hydraulique et des Télécommunications (ENSEEIHT) de Toulouse, Toulouse, France, in 1989 and the Ph.D. degree from the National Polytechnic Institute of Toulouse, Toulouse, in 1992.

He is currently a Professor with the University of Toulouse (ENSEEIHT) and a member of the IRIT Laboratory (UMR 5505 of the CNRS). His research activities are centered around statistical signal and image processing with a particular interest on Bayesian and Markov chain Monte Carlo methods.

Dr. Tourneret has been involved in the organization of several conferences including the European Conference on Signal Processing EUSIPCO'02 (Program Chair), the international conference ICASSP'06 (plenaries), the Statistical Signal Processing Workshop SSP'12 (international liaisons), the International Workshop on Computational Advances in Multi-Sensor Adaptive Processing CAMSAP 2013 (local arrangements), the Statistical Signal Processing Workshop SSP'2014 (special sessions), and the Workshop on Machine Learning for Signal Processing MLSP'2014 (special sessions). He was the General Chair of the CIMI Workshop on Optimization and Statistics in Image Processing held in Toulouse in 2013 (with F. Malgouyres and D. Kouamé), and he is the General Chair of the International Workshop on Computational Advances in Multi-Sensor Adaptive Processing CAMSAP 2015 (with P. Djuric). He has been a member of different technical committees including the Signal Processing Theory and Methods (SPTM) committee of the IEEE Signal Processing Society (2001–2007 and 2010–present). He has been serving as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING (2008–2011 and 2015–present) and of the *EURASIP Journal on Advances in Signal Processing* (since July 2013).

Hichem Snoussi was born in Bizerta, Tunisia, in 1976. He received the Diploma degree in electrical engineering from the Ecole Superieure d'Electricite (Supelec), Gif-sur-Yvette, France, in 2000, the DEA degree and the Ph.D. degree in signal processing from the University of Paris-Sud, Orsay, France, in 2000 and 2003, respectively, and the HDR from the University of Technology of Compiègne, Compiègne, France, in 2009.

Between 2003 and 2004, he was a Postdoctoral Researcher with the Institut de Recherches en Communications et Cybernétiques de Nantes (IRCCyN), Nantes, France. He has spent short periods as a Visiting Scientist with the Brain Science Institute, RIKEN, Japan, and Olin Neuropsychiatry Research Center, The Institute of Living, USA. Between 2005 and 2010, he was an Associate Professor with the University of Technology of Troyes, Troyes, France, where he has been a Full Professor since September 2010. He is in charge of the regional research program System Security and Safety of the CPER 2007–2013 and the CapSec plateform (wireless embedded sensors for security). He is the Principal Investigator of an ANR-Blanc project (mv-EMD), a CRCA project (new partnership and new technologies), and a GDR-ISIS young researcher project. He is a partner of many ANR projects, GIS, and strategic UTT programs.

Dr. Snoussi received the national doctoral and research supervising award PEDR 2008–2012 and the PES (Scientific Excellence Award) 2013–2017.