Spectral Unmixing of Multispectral Lidar Signals

Yoann Altmann, *Member, IEEE*, Andrew Wallace, and Steve McLaughlin, *Fellow, IEEE*

Abstract—In this paper, we present a Bayesian approach for spectral unmixing of multispectral Lidar (MSL) data associated with surface reflection from targeted surfaces composed of several known materials. The problem addressed is the estimation of the positions and area distribution of each material. In the Bayesian framework, appropriate prior distributions are assigned to the unknown model parameters and a Markov chain Monte Carlo method is used to sample the resulting posterior distribution. The performance of the proposed algorithm is evaluated using synthetic MSL signals, for which single and multi-layered models are derived. To evaluate the expected estimation performance associated with MSL signal analysis, a Cramer-Rao lower bound associated with model considered is also derived, and compared with the experimental data. Both the theoretical lower bound and the experimental analysis will be of primary assistance in future instrument design.

Index Terms—Bayesian estimation, estimation performance, Markov chain Monte Carlo, multispectral lidar, remote sensing, spectral unmixing.

I. INTRODUCTION

LASER altimetry (or Lidar) is an acknowledged tool for extracting spatial structures from three-dimensional (3D) scenes, including forest canopies [1], [2]. Using time-of-flight to create a distance profile, signal analysis can recover tree and canopy heights, leaf area indices (LAIs) and ground slope by analyzing the reflected photons from a target. Conversely, passive multispectral (MSI) (dozen of wavelengths) and hyperspectral images (HSI) (hundreds of wavelengths) are widely used to extract spectral information about the scene which, for forest monitoring, can also provide useful parameters about the canopy composition and health (tree species, leaf chlorophyll content, water content, stress, among others) [3], [4]. The most natural evolution to extract spatial and spectral information from sensed scenes is to couple Lidar data and multi/hyperspectral images [5], [6]. Although the fusion of Lidar data and HSIs can improve scene characterization, there are problems of data synchronization in space (alignment, resolution) and time (dynamic scene, change of observation conditions, etc). For these reasons, multispectral Lidar (MSL) has recently received attention from the

remote sensing community for its ability to extract both structural and spectral information from 3D scenes [7], [8].

The key advantage of MSL is the ability to provide information on the vertical distribution of spectra, used to infer physiological processes directly linked to actual carbon sequestration as well as carbon stocks [9]. For example, a key capability is to directly measure and classify ground-based shrub infestation, which is difficult with conventional HSI as the lower spectral response is obscured by the 'first' signals returned from the top of the canopy. Preliminary trials with existing commercial LiDAR systems with three operational wavelengths have already taken place [10], but as we shall demonstrate here, such systems cannot yet provide all the necessary spectral and structural discrimination to directly measure these processes.

Another motivation for MSL is that HSI, even when fully synchronised, can only integrate the spectral response along the path of each optical ray, not measure the spectral response as a function of distance, e.g., depth into a forest canopy. Multiple scattering effects cannot be neglected in some scenes with relief or containing multi-layered objects (such as trees), which complicates surface detection and quantification. Most existing spectral unmixing (SU) techniques rely on a linear mixture assumption [11]–[15] to identify the components (so-called endmembers) of an image and their proportions (abundances). However, it has been shown that the classical linear mixing model (LMM) can be inaccurate when multiple scattering effects occur [16], [17]. Although polynomial models [18], [19] (including bilinear models [20]–[23]) can be adapted for "long range" multiple scattering (in contrast to intimate mixtures models [24]), motivating the additional parameter constraints and designing accurate nonlinear unmixing methods are still challenging problems [25]. Since MSL data present an additional dimension when compared to HSIs, we can expect a reduction of the SU problem complexity and thus an improvement in the target characterization performance.

In [26], the authors proposed a Bayesian algorithm to estimate the positions and amplitudes in Lidar signals associated with a multi-layer target. This method has been extended in [9], [27] to MSL by first estimating the positions of the peaks (*i.e.*, the layers), which were assumed to be the same in all spectral bands, then estimating the amplitudes of the peaks for each wavelength, and finally relating the estimated amplitude peaks to areas associated with each material that make up the target based on a linear mixing model. The method has been applied to real MSL signals (four wavelengths) to analyze a conifer structure and has been shown to aid the recovery of bark and needle areas of the tree assuming it is modeled as a set of irregularly spaced layers.

In this paper, we propose to investigate a SU problem applied to single- and multi-layered targets to estimate the positions and proportions of each (known) material. This problem extends the

supervised SU problem of HSIs by also estimating the target position. Estimating the spectral variation of material signatures (unsupervised SU problem) is also of interest. For example, the estimation of physiological content, such as Chlorophyll concentration, in forest canopies is of key concern, and this results in variable spectra for leaves and needles. Extracting these spectra from MSL signals is a more challenging problem (probably more difficult than in HSIs due to the statistical properties of the noise and observation model) which is out of scope of this paper, but will require particular attention in future work.

In contrast with the classical additive Gaussian noise assumption used for HSIs, a Poisson noise model is more appropriate for MSL signals. Indeed, Lidar and thus MSL systems usually record, for each pixel/region of the scene, a histogram of time delays between emitted laser pulses and the detected photon arrivals. Within each histogram bin, the number of detected photons follows a discrete distribution which can be approximated by a Poisson distribution due to the particle nature of light. Using a Bayesian approach, appropriate prior distributions are chosen for the unknown parameters of the model considered here, *i.e.*, the material areas, the surface positions and the background parameters. The joint posterior distribution of these parameters is then derived. Since the classical Bayesian estimators cannot be easily computed from this joint posterior (mainly due to the model complexity), a Markov chain Monte Carlo (MCMC) method is used to generate samples according to the posterior of interest.

More precisely, following the principles of the Gibbs sampler, samples are generated according to the conditional distributions of the posterior. Due to the possibly high correlations between the material proportions/areas, we propose to use a Hamiltonian Monte Carlo (HMC) [28] method to sample according to some of the conditional distributions. HMCs are powerful simulation strategies based on Hamiltonian dynamics which can improve the convergence and mixing properties of classical MCMC methods (such as the Gibbs sampler and the Metropolis-Hastings algorithm) [29], [30]. These methods have received growing interest in many applications, especially when the number of parameters to be estimated is large [31], [32]. Classical HMC can only be used for unconstrained variables. However, new HMC methods have been recently proposed to handle constrained variables [29, Chap. 5], [33], [34] which allow HMCs to sample according to the posterior of the Bayesian model proposed for SU. Finally, as in any MCMC method, the generated samples are used to compute minimum mean square error (MMSE) estimators as well as measures of uncertainties such as confidence intervals.

Predicting the parameter estimation performance is of prime interest for designing an estimation procedure but also assists with the instrument design. Indeed, since MSL is a recent modality, it is important to identify the key parameters that have an influence on the material estimation performance (such as the necessary signal to background levels, the number of bands to be considered and the associated wavelengths). To guide future instrument design, we consider a Cramer-Rao lower bound (CRLB) associated with the observation model and show that it can be used to approximate the estimation errors of the proposed Bayesian algorithm.

The remainder of the paper is organized as follows. Section II introduces the observation model associated with MSL returns for a single-layered object to be analyzed. Section III presents the hierarchical Bayesian model associated with the proposed model and its posterior distribution. The hybrid MCMC method used to sample from the posterior of interest is detailed in Section IV. Section V investigates the expected SU performance using Cramer-Rao lower bounds. Experimental results are shown and discussed in Section VI. Conclusions are reported in Section VII.

II. PROBLEM FORMULATION

This section introduces the observation statistical model associated with MSL returns for a single-layered object which will be used in Sections III and IV to solve the SU problem of MSL data. Precisely, we consider a set of L observed Lidar waveforms $\mathbf{y}_\ell = [y_{\ell,1}, \ldots, y_{\ell,T}]^T, \ell \in \{1, \ldots, L\}$ where T is the number of temporal (corresponding to range) bins. To be precise, $y_{\ell,t}$ is the photon count within the tth bin of the ℓth spectral band considered. Let t_0 be the position of a complex object surface or surfaces at a given range from the sensor, composed of R materials whose spectral signatures (observed at L wavelengths) are denoted as $\mathbf{m}_r = [m_{1,r}, \ldots, m_{\ell,r}]^T, r \in \{1, \ldots, R\}$. According to [26], each photon count $y_{\ell,t}$ is assumed to be drawn from the following Poisson distribution

$$y_{\ell,t} \sim \mathcal{P}\left(\sum_{r=1}^{R} w_r m_{\ell,r} g_{0,\ell}(t) + b_\ell\right) \quad (1)$$

where $g_{0,\ell}(\cdot)$ is the photon impulse response whose shape can differ between wavelength channels, w_r denotes the area of the rth material composing the object (relative to a reference area) and b_ℓ is a background and dark photon level, which is constant in all bins at a given wavelength. Without loss of generality the photon impulse responses are assumed to be the same and modeled by the following piece-wise exponential

$$g_{0,\ell}(t) = \beta \begin{cases} \exp^{-\frac{T_1^2}{2\sigma^2}} \exp^{\frac{t+T_1-t_0}{\tau_1}} & \text{if } t - t_0 < -T_1 \\ \exp^{-\frac{(t-t_0)^2}{2\sigma^2}} & \text{if } T_1 \leq t - t_0 < T_2 \\ \exp^{-\frac{T_2^2}{2\sigma^2}} \exp^{\frac{t-T_2-t_0}{\tau_2}} & \text{if } T_2 \leq t - t_0 < T_3 \\ \exp^{-\frac{T_2^2}{2\sigma^2}} \exp^{\frac{T_2-T_3}{\tau_2}} \exp^{\frac{t-T_3-t_0}{\tau_3}} & \text{else} \end{cases}$$

where $\boldsymbol{\phi} = [T_1, T_2, T_3, \tau_1, \tau_2, \tau_3, \sigma^2, \beta]^T$ is a positive hyperparameter vector. In this study, we assume that the shape parameters of the model (which is appropriate for our own lidar sensors) are fixed and known from the instrumental response. As seen later in Fig. 3 the shape is slightly asymmetrical when compared to the more common Gaussian model [35] which is used for higher power, longer pulse duration lidar systems. The consideration of a piece-wise exponential approximation is motivated by the actual shape of our instrumental impulse response but it is worth mentioning that any positive analytical approximation, even different approximations for the different wavelengths, could be used instead without modifying the proposed Bayesian estimation procedure. The consideration of a single impulse response model is done only for ease of reading. While
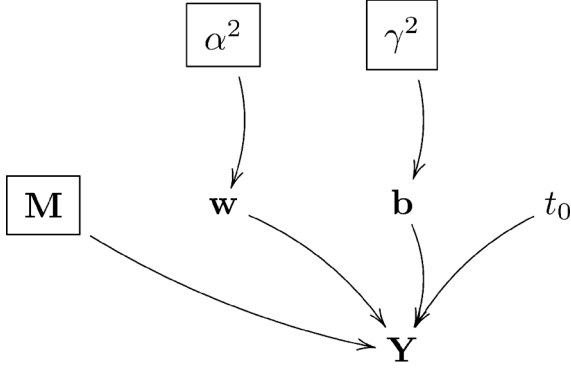
Fig. 1. DAG for the parameter and hyperparameter priors (the fixed parameters appear in boxes).

a Gaussian approximation can lead to a slight bias in depth estimation, we shall use a Gaussian approximation in Section V to compute the Cramer-Rao bounds since the Gaussian approximation is twice differentiable (in contrast to the piece-wise exponential approximation), which simplifies the derivation of the bounds.

Due to physical considerations, the relative areas are assumed to satisfy the following positivity constraints

$$w_r \geq 0, \quad r = 1, \dots, R. \tag{2}$$

Similarly, the average background level in each channel satisfies $b_\ell \geq 0, \ell = 1, \dots, L$. Spectral unmixing of the MSL data consists of estimating the position (*i.e.*, t_0) of the target, the area w_r of each component $\mathbf{m}_r$ of the target, as well as the background levels from the observed data in $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_L]$. The next section studies a Bayesian model to estimate the unknown parameters in (1) while ensuring the positivity constraints mentioned above.

III. BAYESIAN MODEL

The unknown parameter vector associated with (1) contains the relative areas $\mathbf{w} = [w_1, \dots, w_R]^T$, the position of the target t_0 and background levels $\mathbf{b} = [b_1, \dots, b_L]^T$ (satisfying positivity constraints). This section summarizes the likelihood and the parameter priors associated with these parameters.

A. Likelihood

Assuming that the detected photon counts/noise realizations, conditioned on their mean in all bins and for all wavelengths, are independent, (1) leads to

$$P(\mathbf{Y}|\mathbf{M}, \mathbf{w}, \mathbf{b}, t_0) = \prod_{\ell=1}^{L} \prod_{t=1}^{T} \frac{\lambda_{\ell,t}^{y_{\ell,t}}}{y_{\ell,t}!} \exp^{-\lambda_{\ell,t}} \tag{3}$$

where $\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_R]$, $\lambda_{\ell,t} = \mathbf{m}_{\ell,:}\mathbf{w}g_{0,\ell}(t) + b_\ell$ and $\mathbf{m}_{\ell,:}$ denotes the ℓth row of $\mathbf{M}$.

B. Prior for the Target Position

Since we don't have prior information about the position of the target, the following uniform distribution

$$t_0 \sim \mathcal{U}_{(1;T)}(t_0) \tag{4}$$

is assigned to t_0. Note that the position t_0 is a real variable that is not restricted to the integer values in $(1; T)$.

C. Prior for the Relative Areas

To reflect the lack of prior knowledge about the areas, the following truncated Gaussian prior

$$w_r|\alpha^2 \sim \mathcal{N}_{\mathbb{R}^+}(w_r; 0, \alpha^2) \tag{5}$$

is assigned to each area w_r, where $\mathcal{N}_{\mathbb{R}^+}(\cdot; 0, \alpha^2)$ denotes the Gaussian distribution restricted to $\mathbb{R}^+$, which hidden mean and variance 0 and α^2, respectively. The hyperparameter α^2 is shared by all the parameters w_r and is arbitrarily fixed to a large value to ensure a weakly informative prior. Assuming prior independence for the unknown areas yields

$$f(\mathbf{w}|\alpha^2) \propto \left(\frac{1}{\alpha^2}\right)^{\frac{R}{2}} \exp^{-\frac{1}{2\alpha^2}\mathbf{w}^T\mathbf{w}} \mathbf{1}_{(\mathbb{R}^+)^R}(\mathbf{w}) \tag{6}$$

where $\propto$ means "proportional to" and $\mathbf{1}_{(\mathbb{R}^+)^R}(\cdot)$ is the indicator function defined on $(\mathbb{R}^+)^R$.

D. Prior for the Background Level

Similarly, assigning a truncated Gaussian prior to each background level and assuming prior independence between these parameters leads to

$$f(\mathbf{b}|\gamma^2) \propto \left(\frac{1}{\gamma^2}\right)^{\frac{L}{2}} \exp^{-\frac{1}{2\gamma^2}\mathbf{b}^T\mathbf{b}} \mathbf{1}_{(\mathbb{R}^+)^L}(\mathbf{b}) \tag{7}$$

where the hyperparameter γ^2 is shared by all the parameters b_ℓ and is arbitrarily fixed to a large value to ensure a weakly informative prior.

In this paper, we assume that prior knowledge about the relative areas and background levels is limited. We use weakly informative Gaussian priors restricted to the positive orthant to reflect this lack of knowledge and reduce the estimation bias while ensuring the positivity constraints are satisfied. However any proper weakly informative prior distributions could have been used, leading to similar estimation results. Similarly, if useful information about the unknown parameters is available, more informative (possibly physics-based) priors can be used instead of (4) and (6) (target position and composition) and (7) (background levels).

E. Joint Posterior Distribution

The resulting directed acyclic graph (DAG) associated with the proposed Bayesian model is depicted in Fig. 1. The joint posterior distribution of the unknown parameter vector $\boldsymbol{\theta} = \{\mathbf{w}, \mathbf{b}, t_0\}$ can be computed using

$$f(\boldsymbol{\theta}|\mathbf{Y}, \mathbf{M}, \alpha, \beta) \propto P(\mathbf{Y}|\mathbf{M}, \mathbf{w}, \mathbf{b}, t_0)f(\boldsymbol{\theta}|\alpha, \beta) \tag{8}$$

where $P(\mathbf{Y}|\mathbf{M}, \mathbf{w}, \mathbf{b}, t_0)$ has been defined in (3) and $f(\boldsymbol{\theta}|\alpha, \beta) = f(\mathbf{w}|\alpha^2)f(\mathbf{b}|\beta^2)f(t_0)$.

Unfortunately, it is difficult to obtain closed form expressions for standard Bayesian estimators associated with (8). In this paper, we propose to use efficient Markov Chain Monte Carlo (MCMC) methods to generate samples asymptotically distributed according to (8), employing a Gibbs sampler. The principle of the Gibbs sampler is to sample according to

the conditional distributions of the posterior of interest [30, Chap. 10]. Due to the high correlation between the elements of $\mathbf{w}$ (which will be discussed later in the paper), we use an Hamiltonian Monte Carlo (HMC) method which improves the mixing properties of the sampling procedure (compared to a classical Metropolis-within-Gibbs sampler using Gaussian random walks). This method is detailed in the next section.

IV. BAYESIAN INFERENCE USING A CONSTRAINED HAMILTONIAN MONTE CARLO METHOD

In this Section, we develop an HMC-within Gibbs sampler to generate samples according to (8) and estimate the unknown parameters involved in (1) in order to solve the SU problem of MSL data for a single-layered object. The proposed sampler consists of three steps to update sequentially $\mathbf{w}, t_0$ and $\mathbf{b}$ and is summarised in Algo. 1.

Algorithm 1: Gibbs Sampling Algorithm (single layer)

---

1 *Fixed input parameters:* $\mathbf{M}, \boldsymbol{\phi}, \alpha^2, \gamma^2$, number of burn-in iterations N_{bi}, total number of iterations N_{MC}
2 *Initialization* $(i = 0)$
   Set $\mathbf{w}^{(0)}, t_0^{(0)}, \mathbf{b}^{(0)}$
3 *Iterations* $(1 \leq i \leq N_{\mathrm{MC}})$
4 Sample $\mathbf{w}^{(i)}$ from the pdf in **(9)** and CHMC
5 Sample $t_0^{(i)}$ from the pdf in **(11)** and a Gaussian random walk
6 Sample $\mathbf{b}^{(i)}$ from the pdfs in **(13)** and Gaussian random walks
7 Set $i = i + 1$.

A. Sampling the Areas

The full conditional distribution of $\mathbf{w}$ is given by

$$f(\mathbf{w}|\mathbf{Y}, \mathbf{M}, t_0, \mathbf{b}, \alpha^2) \propto P(\mathbf{Y}|\mathbf{M}, \mathbf{w}, \mathbf{b}, t_0) f(\mathbf{w}|\alpha^2) \quad (9)$$

and is not a standard distribution which is easy to sample. Consequently, it is the norm to use an accept/reject procedure to update $\mathbf{b}$. A classical and simple approach would use a multivariate Gaussian random walk. However, in practice the elements of $\mathbf{w}$ can be highly correlated, especially when the materials present similar spectral signatures (which will be the case for vegetation targets). Consequently, a Hamiltonian Monte Carlo method [29, Chap. 5] is preferred to improve the mixing properties of the sampler. The principle of HMCs is to introduce auxiliary, or momentum, variables, and perform a Metropolis-Hastings move in a higher dimensional parameter space. The proposal distribution is then built to take into account the shape of the target distribution (9). Due to the area constraints (2), a constrained HMC must be used. In this paper, we used a constrained HMC (CHMC) scheme similar to that described in [19] and thus consider the following potential energy function

$$U(\mathbf{w}) = -\sum_{\ell,t} y_{\ell,t} \log(\lambda_{\ell,t}) + \lambda_{\ell,t} + \frac{\mathbf{w}^T \mathbf{w}}{2\alpha^2} \quad (10)$$

to simulate Hamiltonian dynamics and compute the appropriate acceptance ratio (see [19], [29] for technical details). This choice of potential energy function ensures that

$U(\mathbf{w}) = -\log\left(f(\mathbf{w}|\mathbf{Y}, \mathbf{M}, t_0, \mathbf{b}, \alpha^2)\right) + c$ where c is a positive constant. Note that the proposed CHMC step can be applied for any differentiable relative area prior. Metropolis-Hastings or more complex HMC moves should be used instead when considering non-differentiable priors instead of (6).

B. Sampling the Target Position

It can be shown from (8) that

$$f(t_0|\mathbf{Y}, \mathbf{M}, \mathbf{w}, \mathbf{b}, \alpha^2) \propto P(\mathbf{Y}|\mathbf{M}, \mathbf{w}, \mathbf{b}, t_0) f(t_0) \quad (11)$$

is not a standard distribution and an accept/reject procedure must be used to update the target position t_0. We use a Gaussian random walk to update this parameter. More precisely, a doubly truncated Gaussian proposal is considered to ensure each candidate belongs to the admissible set $(1; T)$ and the variance of the proposal is adjusted during the burn-in period of the sampler to obtain an acceptance rate close to 0.45, as recommended in [36, p. 8].

C. Sampling the Background Levels

It can be shown from (8) that

$$f(\mathbf{b}|\mathbf{Y}, \mathbf{M}, \mathbf{w}, t_0, \alpha^2) = \prod_{\ell=1}^{L} f(b_\ell|\mathbf{y}_\ell, \mathbf{m}_{\ell,:}, \mathbf{w}, t_0, \alpha^2), \quad (12)$$

where

$$f(b_\ell|\mathbf{y}_\ell, \mathbf{m}_{\ell,:}, \mathbf{w}, t_0, \alpha^2)$$
$$\propto \prod_{t=1}^{T} \lambda_{\ell,t}^{y_{\ell,t}} \exp^{-\lambda_{\ell,t}} \exp^{-\frac{b_\ell^2}{2\gamma^2}} \mathbf{1}_{\mathbb{R}^+}(b_\ell). \quad (13)$$

Sampling from (13) is again not straightforward and Gaussian random walks are used to update the background levels, similar to the position t_0. However the background levels are a posteriori independent and can be updated in parallel. Similar to the target position update, the variances of the L parallel Gaussian random walk procedures are set during the burn-in period of the sampler to obtain an acceptance rate close to 0.45.

After generating N_{MC} samples using the procedures detailed above and removing N_{bi} iterations associated with the burn-in period of the sampler (N_{bi} has been set from preliminary runs), the MMSE estimator of the unknown parameters can be approximated by computing the empirical averages of the remaining samples. The minimal length of the burn-in period can be determined using several convergence diagnostics [36] but the number of initial samples to be discarded generally varies between data sets and can be thus difficult to adjust in advance without overestimating it. In previous work on parallel acceleration of RJMCMC algorithms we have investigated the application of such diagnostics to our monochromatic LiDAR signals [37], but as absolute speed of processing is not a major concern here, the length of the burn-in period is assessed visually from the preliminary runs and fixed to ensure that the sampler has converged. The total number of iterations has then been set to obtain accurate approximations (computed over 4000 samples in Section VI) of the MMSE estimators.

V. PREDICTING UNMIXING PERFORMANCE

This section studies a Cramer-Rao lower bound associated with the observation model (1) which can be used to assess the performance of methods that aim to solve the SU problem of MSL data considered in this paper as well as to assist with the future instrument design. Precisely, Section V-A recalls the definition of the CRLB for the problem of interest and Section V-B discusses the impact of several key parameters on the expected parameter estimation performance of SU methods.

A. Cramer-Rao Lower Bound

Prediction of the unmixing performance is necessary to assist in the design of a lidar system for a specific application, identifying those parameters that have the most significant impact. In deriving a CRLB, we propose firstly to relax the impulse response model in (2) by considering the following Gaussian approximation

$$\tilde{g}_{0,\ell}(t) = \beta \exp^{-\frac{(t-t_0)^2}{2\sigma^2}} \approx g_{0,\ell}(t) \quad \forall t \in (1; T), \forall \ell. \quad (14)$$

This simplifies the CRLB derivation and does not significantly bias the prediction as the piece-wise exponential impulse response can be accurately approximated by a Gaussian function. The CRLB associated with any unbiased estimator $\hat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}$ involved in the mixing model (1) (having replaced $g_{0,\ell}(t)$ by $\tilde{g}_{0,\ell}(t)$) and constructed from $\mathbf{Y}$ is given by

$$\mathrm{CRLB}(\boldsymbol{\theta}) = \mathbf{J}_\mathrm{F}^{-1} \quad (15)$$

where $\mathbf{J}_\mathrm{F}$ is the Fisher information matrix whose elements are[1]

$$[\mathbf{J}_\mathrm{F}]_{i,j} = -\mathrm{E}_{\mathbf{Y}|\boldsymbol{\theta}} \left[\frac{\partial^2 \log f(\mathbf{Y}|\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right], i,j = 1,\ldots,R+L+1.$$

The ith diagonal element of the CRLB matrix in (15) provides a lower bound for the variance of $\hat{\theta}_i$, given that $\hat{\boldsymbol{\theta}}$ is an unbiased estimator of $\boldsymbol{\theta}$. Of course, the Bayesian estimation procedure proposed in Sections III and IV does not provide a strictly unbiased estimator of $\boldsymbol{\theta}$ and a Bayesian CRLB should have been considered instead of (15). However, due to the weakly informative priors chosen in Section III, when the actual vector $\boldsymbol{\theta}$ is far enough from the boundaries of the admissible set (defined by the positivity constraints), the bias of the proposed estimation procedure can be neglected and the resulting estimator achieves the CRLB (as will be shown in Section VI). Moreover, the CRLB in (15) can also provide information about the estimation performance of a potential optimization method that could be proposed to estimate $\boldsymbol{\theta}$ based on maximum likelihood estimation (MLE).

B. Performance Analysis

We first investigate the performance of spectral unmixing of MSL signals for a single layered, artificial target composed of $R = 3$ materials, specifically needles, bark and soil. These materials have been chosen because of our interests in forest canopy monitoring using MSL signals [9]. The reflectance spectra of these materials, observed at equally spaced spectral bands ranging from 400 nm to 2500 nm are depicted in Fig. 2. It is interesting to note the strong similarity between the bark and needles spectral signatures, which can make their discrimination difficult and highlights the need of multiple wavelengths

[1]The Fisher information matrix $\mathbf{J}_\mathrm{F}$ is derived in the Appendix
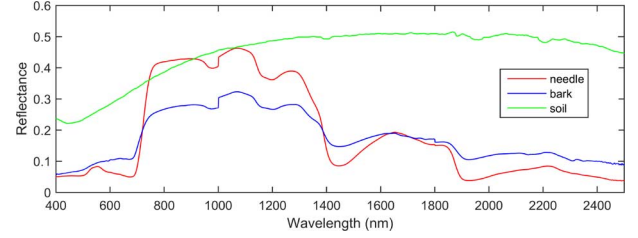

Fig. 2. Spectral signatures of needle (red), bark (blue) and soil (green) considered in the synthetic data.
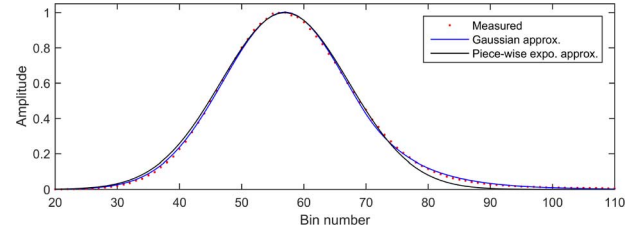

Fig. 3. Real normalized photon impulse response (red dots) and approximations using piece-wise exponential (blue line) and Gaussian (black line) functions.

to quantify these materials. The maximum number of spectral bands considered in this paper has been set to $L_{\max} = 32$, which is a realistic value for a short-term real measurement campaign (the current instrument uses only 4 wavelengths). The relative area of needles (resp. bark and soil) has been arbitrarily set to $w_n = 0.2$ (resp. $w_b = 0.3$ and $w_s = 0.4$). These relative areas do not necessarily sum to one as we consider that part of the laser light can penetrate through the target (semi-transparent target or a significant gap fraction) and may not be further reflected (single hit assumption). The fixed model parameters have been fixed from experiments to

$$\begin{cases} T = 2500 \\ \beta = 3000 \\ t_0 = 1000 \\ b_\ell = b = 10, \quad \forall \ell. \end{cases} \quad (16)$$

These parameters will be fixed in the remainder of the paper unless otherwise specified. The impulse response parameters have been set to $T_1 = 402$, $T_2 = 12.5$, $T_3 = 239$, $\tau_1 = 395$, $\tau_2 = 7.9$, $\tau_3 = 1595$ and $\sigma^2 = 105.82$ for the piece-wise exponential approximation in (2) and $\sigma^2 = 105.68$ for the Gaussian approximation in (14). These parameters have been obtained by fitting the experimental impulse response measured in [26]. Fig. 3 shows that the Gaussian approximation provides a good estimate of the experimental impulse response, although the piece-wise exponential better fits the experimental curve. Fig. 4 shows an example of MSL data generated using the parameters in (16) and the impulse response approximation in (2). Table I shows the predicted estimation performance for the unknown parameters of interest (*i.e.*, $\mathbf{w} = [w_n, w_b, w_s]^T$ and t_0) assuming the Gaussian approximation of the photon impulse response (14). The relative errors provided in the bottom row of Table I are computed by dividing the square root of the CRLBs by the actual values of the parameters. The relative estimation errors of the background levels (not presented here) are lower than 1%.

These results show that the estimation errors associated with the target position are usually much lower than those associ-
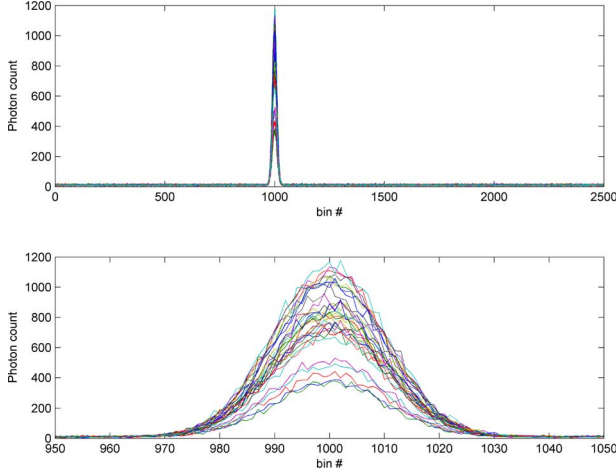
Fig. 4. Example of MSL signals for $L = 32$ spectral bands and a target located at $t_0 = 1000$ (the bottom subplot is a zoom around the actual target location). The different colors correspond to different spectral bands.

TABLE I
ESTIMATION PERFORMANCE ($L = 32$).

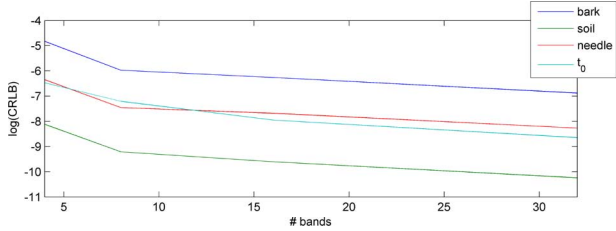|	w_n	w_b	w_s	t_0
Actual value	0.2	0.3	0.4	1500
CRLB	2.6×10^{-4}	0.001	3.6×10^{-5}	1.8×10^{-4}
Rel. err. (%)	8.0	10.7	1.5	$< 10^{-5}$

Fig. 5. Evolution of the CRLB for $\mathbf{w}$ and t_0 as a function of the number of bands $L(\beta = 3000, b_\ell = 10, \forall\ell)$.

ated with the material areas. Moreover, although the amplitude of the peak for each wavelength is much larger than the background level (see Fig. 4), the CRLB predicts possibly large estimation errors, especially for the needle and bark areas. This can be partially explained by the fact that the soil reflectance spectrum presents an average energy which is higher than those of needles and bark (and thus $w_s > w_b > w_n$) but also and mainly because of the high correlation between the bark and needle spectra, which complicates their quantification. Fig. 5 to Fig. 7 show the predicted estimation errors for different values of the key parameters β (which is related to the number of photons emitted by the laser sources), the number of spectral bands (equally spaced between 400 nm and 2500 nm) and the average background level b (assumed to be the same in all spectral bands and bins). Fig. 5 shows that the estimation performance generally increases with the number of spectral bands. This result is well known for SU of multispectral and hyperspectral images and is here demonstrated for MSL signals. However, the performance improvement can vary, depending on the additional spectral bands considered. For instance, additional bands containing similar spectral information may not significantly improve the unmixing results (e.g., several wavelengths between 2000 nm
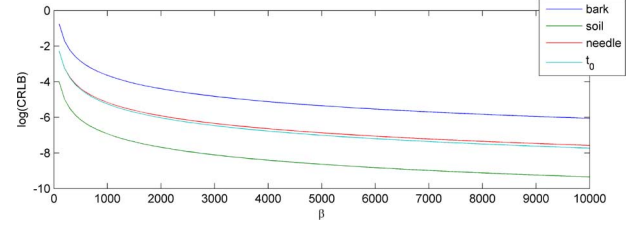

Fig. 6. Evolution of the CRLB for $\mathbf{w}$ and t_0 as a function of $\beta(L = 32, b_\ell = 10)$.
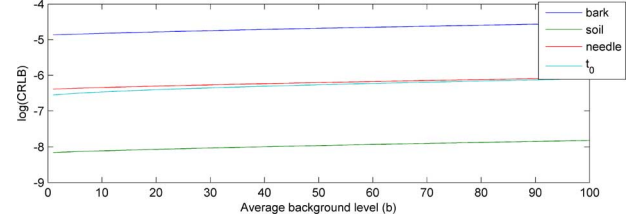

Fig. 7. Evolution of the CRLB for $\mathbf{w}$ and t_0 as a function of $b = b_\ell, \forall\ell(L = 32, \beta = 3000)$.

and 2500 nm for the three materials in Fig. 2). Fig. 6 shows that the parameter β has a significant impact on the SU performance. This parameter increases with the amplitude and the number of laser source pulses used to acquire the data and decreases with the average distance between the source and the target (as the probability of recording a reflected photon decreases). Increasing β leads to better area estimation but can require a longer target exposure (which can be problematic for airborne sensors for instance). Finally, Fig. 7 shows that for relatively small values of background levels (compared to β), the SU performance is not significantly degraded when the background increases. This background level (which depends primarily on the background (e.g., solar) radiation as well as the instrument design) is expected to be quite small in practice.

VI. EXPERIMENTS

In this section, we first apply (Section VI-A) the proposed SU method to synthetic MSL data associated with a single-layered target and compare its estimation performance to that predicted by the CLRB presented in Section V and to that of a state-of-the art method [9]. In Section VI-B, we investigate the case of a multi-layered target whose layer positions are assumed to be known. Although the current algorithm does not estimate the positions of multiple surfaces, Section VI-B aims to illustrate how the proposed method can be extended to analyze more complex objects. The main steps of the methodology (proposed to analyze single and multi-layer target) are summarised in Fig. 8.

A. Single Layer Target

In this section, we evaluate the performance of the proposed spectral unmixing algorithm on synthetic MSL signals generated using the parameters used in Section V-B ((16)) and impulse response approximation (2).

The number of spectral bands has been chosen as $L = 4, 8, 16, 32$, equally spaced between 400 nm and 2500 nm. For each scenario, the number of iterations has been set to $N_{\mathrm{MC}} = 8000$ (including $N_{\mathrm{bi}} = 4000$ burn-in samples). The
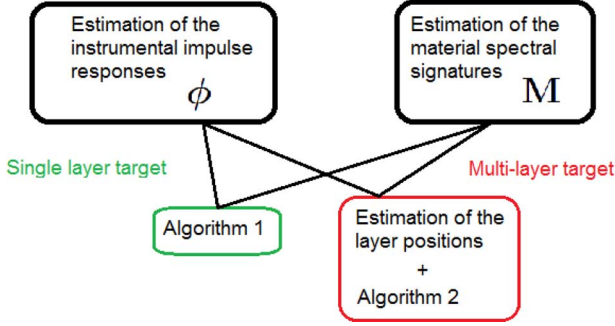
Fig. 8. Main steps of the method proposed for spectral unmixing of MSL data associated with single and multi-layer targets.

TABLE II
ESTIMATION PERFORMANCE: MSE.

Actual parameter value		w_n	w_b	w_s
		0.2	0.3	0.4
$L = 32$	CRLB	2.6×10^{-4}	1.0×10^{-3}	3.6×10^{-5}
	Proposed Algo.	2.7×10^{-4}	1.1×10^{-3}	3.8×10^{-5}
	State-of-the-art	3.0×10^{-4}	1.2×10^{-3}	4.1×10^{-5}
$L = 16$	CRLB	4.6×10^{-4}	1.9×10^{-3}	6.8×10^{-5}
	Proposed Algo.	4.8×10^{-4}	2.0×10^{-3}	7.0×10^{-5}
	State-of-the-art	5.3×10^{-4}	2.2×10^{-3}	7.8×10^{-5}
$L = 8$	CRLB	5.8×10^{-4}	2.6×10^{-3}	1.0×10^{-4}
	Proposed Algo.	5.7×10^{-4}	2.5×10^{-3}	1.0×10^{-4}
	State-of-the-art	6.1×10^{-4}	2.7×10^{-3}	1.1×10^{-4}
$L = 4$	CRLB	1.8×10^{-3}	8.0×10^{-3}	3.0×10^{-4}
	Proposed Algo.	1.6×10^{-3}	7.4×10^{-3}	2.9×10^{-4}
	State-of-the-art	1.8×10^{-3}	8.6×10^{-3}	3.3×10^{-4}

hyperparameters have been fixed to $(\alpha^2, \gamma^2) = (10^6, 10^6)$. The estimation performance of the proposed algorithm is evaluated by comparing the CRLB defined in Section V to the mean square error (MSE) defined as

$$\text{MSE}_i = \text{E}\left[\left(\hat{\theta}_i - \theta_i\right)^2\right], \quad i = 1, \dots, R + L + 1 \quad (17)$$

where θ_i and $\hat{\theta}_i$ are the actual and estimated ith element of the unknown parameter vector $\boldsymbol{\theta}$ and the expectation is approximated using 900 Monte Carlo runs. The performance of the proposed algorithm is compared to the CRLB (including the Gaussian approximation (14)) and to the performance of the algorithm developed in [9], denoted as "State-of-the-art" and which consists of estimating sequentially the position, amplitudes of the peaks for each wavelength and finally the relative areas.

Table II shows that the MSEs obtained for $L \in \{4, 8, 16, 32\}$ by the two algorithms are close to the CLRBs, although the proposed algorithm generally outperforms the method proposed in [9] due to the joint estimation of the target position and the material areas. Note that for $L = 4$, the variance of the proposed estimator is slightly lower than the CRLB. This can be mainly explained by the fact that the estimator is no longer unbiased in that case. Table III compares the average relative estimation errors obtained by the two algorithms and computed by dividing the square root of the MSEs by the actual values of the parameters. These results show that increasing the number of spectral bands from $L = 4$ to $L = 32$ almost divides the area estimation errors by three (*e.g.*, from $\approx 30\%$ to $\approx 10\%$ for the bark

TABLE III
ESTIMATION PERFORMANCE: AVERAGE RELATIVE ERRORS (IN %).

Actual parameter value		w_n	w_b	w_s
		0.2	0.3	0.4
$L = 32$	Theo. (%)	8.0	10.7	1.5
	Proposed Algo.	8.2	11.0	1.5
	State-of-the-art	8.6	11.5	1.6
$L = 16$	Theo. (%)	10.8	14.6	2.0
	Proposed Algo.	11.0	14.8	2.1
	State-of-the-art	11.6	15.7	2.2
$L = 8$	Theo. (%)	12.1	16.8	2.5
	Proposed Algo.	12.0	16.8	2.5
	State-of-the-art	12.4	17.4	2.6
$L = 4$	Theo. (%)	21.0	29.9	4.3
	Proposed Algo.	20.0	28.8	4.2
	State-of-the-art	21.4	30.8	4.5

area), which highlights the benefits of increasing the number of wavelengths.

B. Extension to a Multi-Layer Target

In this section we extend the model (1) by considering a target composed of D layers, leading to

$$y_{\ell,t} \sim \mathcal{P}\left(\sum_{d=1}^{D} \mathbf{m}_{\ell,:} \mathbf{w}_d g_{d,\ell}(t) + b_\ell\right) \quad (18)$$

where $g_{d,\ell}(\cdot)$ is the photon impulse response of the dth layer located at t_d and $\mathbf{w}_d = [w_{1,d}, \dots, w_{R,d}]^T \succeq 0$ is the relative area vector associated with the dth layer. In this scenario, we assume that the number and positions of the layers are known. Consequently, the unmixing problem reduces to estimating the relative areas of the R known components for the D layers (and the background levels). The joint spectral and spatial unmixing problem is more challenging and will be discussed in Section VI-C. The Bayesian model presented in Section III has been extended by considering the same prior (6) for the area vectors of each layer. The sampling procedure studied in Section IV has been modified in order to update sequentially the D area vectors and the target position update step has been removed since the surface positions are assumed to be fixed in this paragraph (see Algo. 2).

Algorithm 2: Gibbs Sampling Algorithm (multi-layer)

---

1 *Fixed input parameters* : $\mathbf{M}, \boldsymbol{\phi}, \alpha^2, \gamma^2$, number of layers D, layer positions $\{t_d\}_{d=1,\dots,D}$, number of burn-in iterations N_{bi}, total number of iterations N_{MC}

2 *Initialization* $(i = 0)$
   Set $\left\{\mathbf{w}_d^{(0)}\right\}_{d=1,\dots,D}, \mathbf{b}^{(0)}$

3 *Iterations* $(1 \leq i \leq N_{\text{MC}})$

4 **for** d $= 1 : D$ **do**

5 Sample $\mathbf{w}_d^{(i)}$ from its conditional distribution and CHMC

6 **end for**

7 Sample $\mathbf{b}^{(i)}$ from its conditional distribution and Gaussian random walks
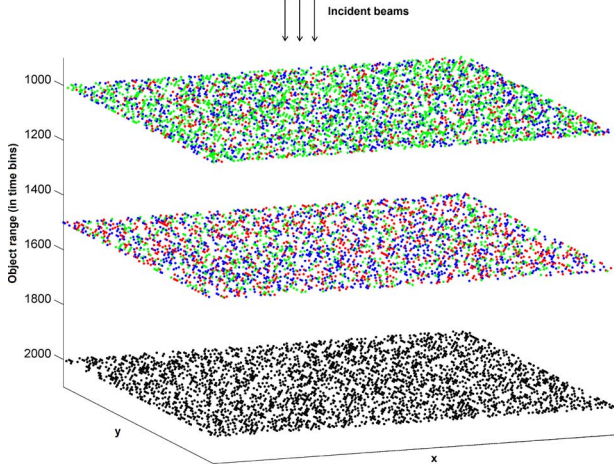
8 Set $i = i + 1$.

Fig. 9. Artificial 3-layered target composed of needles (red), bark (green), soil (blue) and pure reflective material (black).
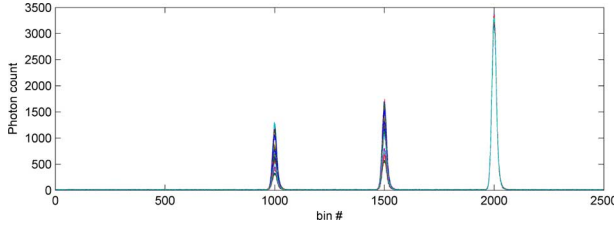

Fig. 10. MSL signals ($L = 32$) associated with the artificial 3-layered target ($t_1 = 1000, t_2 = 1500, t_1 = 2000$). The different colors correspond to different spectral bands.

1) Target Description: We evaluate the unmixing performance of the extended Bayesian algorithm using an artificial target composed of $D = 3$ layers. More precisely, the multi-layer target is assumed to be far enough from the source to ensure that the laser rays hitting the target have the same incident direction. Thus the photons reflected onto the dth layers lead to the same impulse response (*i.e.*, same t_d). The first two layers are located at $t_1 = 1000$ and $t_2 = 1500$ and are composed of needles, bar and soil. It is important to note that the 3 materials composing the first two layers are assumed to randomly distributed, without overlap. The third layer modeling the reference spectralon is located at $t_3 = 2000$. The associated relative areas are presented in Table IV and the location of the $D = 3$ layers is depicted in Fig. 9. Note that in this particular scenario, the relative areas correspond to areas visible by the source (no occlusion) and satisfy $\sum_{r=1}^{R} \sum_{d=1}^{D} w_{r,d} = 1$. However, this constraint is not ensured by the proposed estimation procedure. MSL data associated with this scene have been generated according to (18) with $\beta = 10^4, b_\ell = b = 10, \forall \ell$ and $L = 32$ and are depicted in Fig. 10

2) Estimation Performance: The proposed Bayesian algorithm extended to account for multiple layers has been applied to the MSL signals with $N_{\mathrm{MC}} = 8000$ (including $N_{\mathrm{bi}} = 4000$ burn-in samples). The layer positions have been fixed to their actual values. Table IV shows the estimated relative areas for each layer and $L \in \{4, 8, 16, 32\}$. This table shows that the proposed algorithm provides accurate area estimates for each layer and the more wavelengths the better the estimation, as already observed with the single-layered target in Section VI-A.

TABLE IV
ESTIMATION PERFORMANCE: MULTI-LAYER TARGET.

		w_n	w_b	w_s	spectralon
Actual	Layer 1	0.099	0.099	0.102	0
	Layer 2	0.080	0.200	0.120	0
	Layer 3	0	0	0	0.30
Estimated ($L=32$)	Layer 1	0.0928	0.0780	0.1185	0.0004
	Layer 2	0.0669	0.1953	0.1161	0.0006
	Layer 3	0.0019	0.0011	0.0011	0.3274
Estimated ($L=16$)	Layer 1	0.0932	0.0773	0.1181	0.0006
	Layer 2	0.0620	0.1961	0.1179	0.0006
	Layer 3	0.0024	0.0018	0.0014	0.3270
Estimated ($L=8$)	Layer 1	0.0957	0.0764	0.1166	0.0008
	Layer 2	0.0640	0.1954	0.1172	0.0006
	Layer 3	0.0024	0.0028	0.0016	0.3268
Estimated ($L=4$)	Layer 1	0.1286	0.0687	0.1013	0.0015
	Layer 2	0.0975	0.1863	0.1014	0.0016
	Layer 3	0.0023	0.0019	0.0015	0.3271

C. Towards Real Data Analysis

Due to the current limitations of the instrument, the proposed algorithm has only been applied to synthetic MSL signals whose underlying model has been shown to be in good agreement with our previous real MSL measurements using fewer wavelengths [9], [26]. Using this real data, we have managed to investigate the precision of area estimation in a multi-layer model; ground truth was available for the endmembers, but unfortunately we did not have structural truth. As stated above, real MSL signal acquisition campaigns are underway with 3 wavelengths [10] but we need to extend these and have simultaneous ground truth measurements to better evaluate the accuracy of the model and the estimation algorithm presented in this paper. Moreover, the number and positions of the layers which the multi-layer target is composed of were assumed to be known in Section VI-B. Estimating these parameters, especially the number of layers, for MSL data is a challenging problem that can be addressed in the Bayesian framework using reversible jump MCMC methods, in a similar fashion to [26]. This issue is however out of scope of this work and will also be addressed in a future paper.

VII. CONCLUSION

In this paper, we have proposed and developed a Bayesian model and an MCMC method for spectral unmixing of multispectral Lidar data. First, we evaluated our method on a single layer simulated target composed of known materials so that the MSL returns consisted of the sum of the individual contributions of the different components. The algorithm estimated the target position, the relative areas of the materials and the noise statistical properties.

The Cramer-Rao lower bound associated with the observation model was derived to identify the key parameters involved in the expected performance of the SU algorithm. It was shown that this bound can be used to help design future multispectral/hyperspectral Lidar systems (*e.g.,* number of spectral bands and associated wavelengths, laser power ...) for specific applications. It was also shown that the performance of the proposed SU

strategy is close to the results provided by the Cramer-Rao lower bound, although the bound considered in the paper only holds for deterministic parameters. The proposed Bayesian model was then extended to handle multi-layer targets (assuming the layer positions are known) and the simulation results provided interesting results for solving the more challenging joint spectral and spatial unmixing problem.

The model use in this paper does not take into account possible multiple scattering effects (which are likely to occur when analyzing multi-layer scenes). Such effects have already been observed in HSIs and Lidar signals over canopies. Developing non-linear models for MSL signal analysis is a challenging problem to be addressed in future studies.

APPENDIX
FISHER INFORMATION MATRIX

The likelihood of the observation matrix $\mathbf{Y}$ can be expressed as

$$f(\mathbf{Y}|\mathbf{w}, \mathbf{M}, t_0, \mathbf{b}) = \prod_{t=1}^{T} \prod_{\ell}^{L} \frac{(\lambda_{\ell,t})^{y_{\ell,t}}}{y_{\ell,t}!} \exp^{-\lambda_{\ell,t}} \qquad (19)$$

where $\lambda_{\ell,t} = \mathbf{m}_{\ell,:} \mathbf{w} \tilde{g}_{0,\ell}(t) + b_\ell$. The corresponding log-likelihood $P = \log f(\mathbf{Y}|\mathbf{w}, \mathbf{M}, t_0, \mathbf{b})$ is given by

$$P = \sum_{t=1}^{T} \sum_{\ell=1}^{L} y_{\ell,t} \log(\lambda_{\ell,t}) - \log(y_{\ell,t}!) - \lambda_{\ell,t!}. \qquad (20)$$

The partial derivatives of P with respect to (w.r.t.) the unknown model parameters are

$$\frac{\partial P}{\partial \mathbf{w}} = \sum_{\ell=1}^{L} \sum_{t=1}^{T} \left(\frac{y_{\ell,t} \tilde{g}_{0,\ell}(t)}{\lambda_{\ell,t}} - \tilde{g}_{0,\ell}(t) \right) \mathbf{m}_{\ell,:}$$

$$\frac{\partial P}{\partial b_\ell} = \sum_{t=1}^{T} \frac{y_{\ell,t}}{\lambda_{\ell,t}} - 1$$

$$\frac{\partial P}{\partial t_0} = \sum_{t=1}^{T} \sum_{\ell=1}^{L} \left(\frac{y_{\ell,t}}{\lambda_{\ell,t}} - 1 \right) \mathbf{m}_{\ell,:} \mathbf{w} \tilde{g}_{0,\ell}(t) \frac{t - t_0}{\sigma^2}.$$

Straightforward computations lead to

$$\mathrm{E}\left[\frac{\partial^2 P}{\partial \mathbf{w}^2} \right] = -\sum_{t=1}^{T} \sum_{\ell=1}^{L} \frac{\tilde{g}_{0,\ell}^2(t)}{\lambda_{\ell,t}} \mathbf{m}_{\ell,:}^T \mathbf{m}_{\ell,:}$$

$$\mathrm{E}\left[\frac{\partial^2 P}{\partial b_\ell \partial b_{\ell'}} \right] = \begin{cases} -\sum_{z=1}^{N_z} \frac{1}{\lambda_{\ell,t}} & \text{if } \ell = \ell' \\ 0 & \text{else} \end{cases}$$

$$\mathrm{E}\left[\frac{\partial^2 P}{\partial \mathbf{w} \partial b_\ell} \right] = -\sum_{t=1}^{T} \frac{\tilde{g}_{0,\ell}(t)}{\lambda_{\ell,t}} \mathbf{m}_{\ell,:}$$

$$\mathrm{E}\left[\frac{\partial^2 P}{\partial t_0^2} \right] = -\sum_{t=1}^{T} \sum_{\ell=1}^{L} \frac{[(t - t_0) \mathbf{m}_{\ell,:} \mathbf{w} \tilde{g}_{0,\ell}(t)]^2}{\sigma^4 \lambda_{\ell,t}}$$

$$\mathrm{E}\left[\frac{\partial^2 P}{\partial t_0 \partial \mathbf{w}} \right] = -\sum_{t=1}^{T} \sum_{\ell=1}^{L} \frac{(t - t_0) \mathbf{m}_{\ell,:} \mathbf{w} \tilde{g}_{0,\ell}^2(t)}{\sigma^2 \lambda_{\ell,t}} \mathbf{m}_{\ell,:}$$

$$\mathrm{E}\left[\frac{\partial^2 P}{\partial t_0 \partial b_\ell} \right] = -\sum_{t=1}^{T} \frac{(t - t_0) \mathbf{m}_{\ell,:} \mathbf{w} \tilde{g}_{0,\ell}(t)}{\sigma^2 \lambda_{\ell,t}}$$

using the fact fact $\mathrm{E}[y_{\ell,t}] = \mathbf{m}_{\ell,:} \mathbf{w} \tilde{g}_{0,\ell}(t) + b_\ell$

REFERENCES

[1] C. Mallet and F. Bretar, "Full-waveform topographic lidar: State-of-the-art," *ISPRS J. Photogram. Remote Sens.*, vol. 64, no. 1, pp. 1–16, 2009.

[2] M. van Leeuwen and M. Nieuwenhuis, "Retrieval of forest structural parameters using lidar remote sensing," *Eur. J. Forest Res.*, vol. 129, no. 4, pp. 749–770, 2010.

[3] G. le Maire, C. François, K. Soudani, D. Berveiller, J.-Y. Pontailler, N. Bréda, H. Genet, H. Davi, and E. Dufrêne, "Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3846–3864, 2008.

[4] R. Hernandez-Clemente, R. M. Navarro-Cerrillo, L. Suarez, F. Morales, and P. J. Zarco-Tejada, "Assessing structural effects on {PRI} for stress detection in conifer forests," *Remote Sens. Environ.*, vol. 115, no. 9, pp. 2360–2375, 2011.

[5] M. Dalponte, L. Bruzzone, and D. Gianelle, "Fusion of hyperspectral and lidar remote sensing data for classification of complex forest areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1416–1427, May 2008.

[6] S. J. Buckley, T. H. Kurz, J. A. Howell, and D. Schneider, "Terrestrial lidar and hyperspectral data fusion products for geological outcrop analysis," *Compu. Geosci.*, vol. 54, pp. 249–258, 2013.

[7] A. M. Wallace, C. J. Nichol, and I. H. Woodhouse, "Recovery of forest canopy parameters by inversion of multispectral lidar data," *Remote Sens.*, vol. 4, no. 2, pp. 509–531, 2012.

[8] T. Hakala, J. Suomalainen, S. Kaasalainen, and Y. Chen, "Full waveform hyper-spectral lidar for terrestrial laser scanning," *Opt. Exp.*, vol. 20, no. 7, pp. 7119–7127, Mar. 2012.

[9] A. M. Wallace, A. McCarthy, C. Nichol, X. Ren, S. Morak, D. Martinez-Ramirez, I. Woodhouse, and G. S. Buller, "Design and evaluation of multispectral lidar for the recovery of arboreal parameters," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4942–4954, Aug. 2014.

[10] S. Fleming, A. Cottin, and I. Woodhouse, "The first spectral map of a forest understory from multispectral lidar," *LiDAR News Mag.*, vol. 5, pp. 26–30, 2015.

[11] M. Craig, "Minimum volume transforms for remotely sensed data," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 3, pp. 542–552, May 1994.

[12] D. C. Heinz and C.-I. Chang, "Fully constrained least-squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 29, no. 3, pp. 529–545, Mar. 2001.

[13] O. Eches, N. Dobigeon, C. Mailhes, and J.-Y. Tourneret, "Bayesian estimation of linear mixtures using the normal compositional model," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1403–1413, Jun. 2010.

[14] L. Miao, H. Qi, and H. Szu, "A maximum entropy approach to unsupervised mixed-pixel decomposition," *IEEE Trans. Image Process.*, vol. 16, no. 4, pp. 1008–1021, Apr. 2007.

[15] Z. Yang, G. Zhou, S. Xie, S. Ding, J.-M. Yang, and J. Zhang, "Blind spectral unmixing based on sparse nonnegative matrix factorization," *IEEE Trans. Image Process.*, vol. 20, no. 4, pp. 1112–1125, Apr. 2011.

[16] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[17] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[18] Y. Altmann, A. Halimi, N. Dobigeon, and J. Tourneret, "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 21, no. 6, pp. 3017–3025, Jun. 2012.

[19] Y. Altmann, N. Dobigeon, and J. Tourneret, "Unsupervised post-nonlinear unmixing of hyperspectral images using a Hamiltonian Monte Carlo algorithm," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2663–2675, Jun. 2014.

[20] B. Somers, K. Cools, S. Delalieux, J. Stuckens, D. V. der Zande, W. W. Verstraeten, and P. Coppin, "Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards," *Remote Sens. Environ.*, vol. 113, no. 6, pp. 1183–1193, 2009.

[21] J. M. P. Nascimento and J. M. Bioucas-Dias, L. Bruzzone, C. Notarnicola, and F. Posa, Eds., "Nonlinear mixture model for hyperspectral unmixing," in *Proc. SPIE Image Signal Process. Remote Sens. XV SPIE*, 2009, vol. 7477, p. 74770I.

[22] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *Remote Sens. Environ.*, vol. 30, no. 11, pp. 2951–2962, Jun. 2009.

[23] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.

[24] B. W. Hapke, "Bidirectional reflectance spectroscopy. I. theory," *J. Geophys. Res.*, vol. 86, pp. 3039–3054, 1981.

[25] N. Dobigeon, L. Tits, B. Somers, Y. Altmann, and P. Coppin, "A comparison of nonlinear mixing models for vegetated areas using simulated and real hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1869–1878, Jun. 2014.

[26] S. Hernandez-Marin, A. M. Wallace, and G. J. Gibson, "Bayesian analysis of lidar signals with multiple returns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 12, pp. 2170–2180, Dec. 2007.

[27] D. Ramirez, G. S. Buller, A. McCarthy, X. Ren, A. M. Wallace, S. Morak, C. Nichol, and I. Woodhouse, "Developing hyperspectral lidar for structural and biochemical analysis of forest data," presented at the Adv. Geosci., Proc. EARSEL Conf., Mykonos, Greece, May 2012.

[28] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, "Hybrid Monte Carlo," *Phys. Lett. B*, vol. 195, pp. 216–222, Sep. 1987.

[29] S. Brooks, *Handbook of Markov Chain Monte Carlo*, ser. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. London, U.K.: Taylor & Francis, 2011.

[30] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York, NY, USA: Springer-Verlag, 2004.

[31] R. M. Neal, *Bayesian Learning for Neural Networks*. New York, NY, USA: Springer-Verlag, 1996.

[32] M. N. Schmidt, A. P. Danyluk, L. Bottou, and M. L. Littman, Eds., "Function factorization using warped Gaussian processes," in *Proc. Int. Conf. Mach. Learn., ser. ACM Int. Conf. Proc. Series ACM*, 2009, vol. 382, p. 116.

[33] C. Hartmann and C. Schütte, "A constrained hybrid Monte-Carlo algorithm and the problem of calculating the free energy in several variables," *ZAMM—J. Appl. Math. Mechan./Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 85, no. 10, pp. 700–710, 2005.

[34] M. A. Brubaker, M. Salzmann, and R. Urtasun, "A family of MCMC methods on implicitly defined manifolds," *J. Mach. Learn. Res.—Proc. Track*, vol. 22, pp. 161–172, 2012.

[35] W. W. A. Ullrich, V. Ducic, T. Melzer, and N. Studnick, "Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner," *ISPRS J. Photogram. Remote Sens.*, vol. 60, no. 2, pp. 100–112, 2006.

[36] C. P. Robert and D. Cellier, "Convergence control of MCMC algorithms," in *Discretization and MCMC Convergence Assessment*, C. P. Robert, Ed. New York, NY, USA: Springer-Verlag, 1998, pp. 27–46.

[37] J. Ye, AM. Wallace, A. Zain, and J. Thompson, "Parallel Bayesian inference of range and reflectance from ladar profiles," *J. Parallel Distrib. Comput.*, vol. 73, pp. 383–399, 2013.
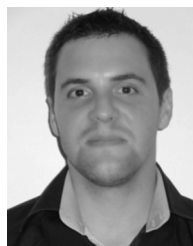
Yoann Altmann (S'10–M'14) was born in Toulouse, France, in 1987. He received the Engineering degree in electrical engineering from Ecole Nationale Supérieure d'Electronique, d'Electrotechnique, d'Informatique, d'Hydraulique et des Télécommunications, Toulouse, and the M.Sc. degree in signal processing from the National Polytechnic Institute of Toulouse (INP Toulouse), Toulouse, both in 2010, and the Ph.D. degree from INP Toulouse in 2013.

Since 2014, he has been with the Heriot-Watt University, Edinburgh, as a Post-Doctoral Researcher. He conducts his research within the Institute of Sensors, Signals and Systems, School of Engineering and Physical Sciences. His current research activities focus on statistical signal and image processing, with a particular interest in Bayesian inverse problems with applications to remote sensing and biomedical imaging.

Andrew Wallace received his BSc and PhD degrees from the University of Edinburgh in 1972 and 1975 respectively.

He is a Professor of Signal and Image Processing at Heriot-Watt University, with particular interests in LiDAR signal processing, video analytics and parallel, many core architectures. He has published extensively; receiving a number of best paper and other awards, and has secured funding from EPSRC, the EU and other industrial and government sponsors. He is a chartered engineer and a Fellow of the Institute of Engineering Technology.

Steve McLaughlin (S'90–M'92–SM'99–F'11) was born in Clydebank Scotland in 1960. He received the B.Sc. degree in Electronics and Electrical Engineering from the University of Glasgow in 1981 and the Ph.D. degree from the University of Edinburgh in 1990.

From 1981 to 1984 he was a Development Engineer in industry involved in the design and simulation of integrated thermal imaging and fire control systems. From 1984 to 1986 he worked on the design and development of high frequency data communication systems. In 1986 he joined the Dept. of Electronics and Electrical Engineering at the University of Edinburgh as a research fellow where he studied the performance of linear adaptive algorithms in high noise and nonstationary environments. In 1988 he joined the academic staff at Edinburgh, and from 1991 until 2001 he held a Royal Society University Research Fellowship to study nonlinear signal processing techniques. In 2002 he was awarded a personal Chair in Electronic Communication Systems at the University of Edinburgh. In October 2011 he joined Heriot-Watt University as a Professor of Signal Processing and Head of the School of Engineering and Physical Sciences. His research interests lie in the fields of adaptive signal processing and nonlinear dynamical systems theory and their applications to biomedical, energy and communication systems.

Prof. McLaughlin is a Fellow of the Royal Academy of Engineering, of the Royal Society of Edinburgh, of the Institute of Engineering and Technology and of the IEEE.