

# Reconstruction of Backscatter and Extinction Coefficients in Lidar: A Stochastic Filtering Approach

José M. Bioucas Dias, *Member, IEEE*, José M. N. Leitão, *Member, IEEE*, and Elsa S. R. Fonseca

**Abstract**—Reconstruction of the backscatter and extinction coefficients is a crucial step in many quantitative remote sensing applications, such as radar, light detection and ranging (lidar), and sonar. We present a novel stochastic filtering approach for the estimation of the backscatter and extinction coefficients from time–range elastic-backscatter lidar data. The Bayesian perspective is adopted; we take as prior a causal first-order autoregressive Gauss–Markov random field tailored to enforce smoothness on time and range dimensions. By using a reduced-order state-space representation of the prior, we derive a suboptimal stochastic filter that recursively computes the backscatter and extinction coefficients at each range–time inversion cell. The estimator is a kind of adaptive extended Kalman filter, being efficient from the computational point of view. A set of experiments illustrates the effectiveness of the proposed approach, namely its advantage over the classical Klett deterministic approach.

**Index Terms**—Backscatter and extinction coefficients, Bayesian, Kalman filter, light detection and ranging (lidar), reduced-order model, stochastic filter.

## I. INTRODUCTION

**A** WAVE PROPAGATING in a random medium suffers attenuation and scattering. If the multiple scattering among scatters is negligible and there is no wavelength shift (elastic scattering), then the *first-order multiple scattering* solution [1, ch. 8] to the backscattered mean power received from range  $z$  is

$$P(z) = \frac{C}{z^2} \beta(z) \exp\left(-2 \int_0^z \alpha(z') dz'\right) \quad (1)$$

where  $C$  is a system dependent constant, and  $\beta$  and  $\alpha$  are, respectively, the backscatter and the extinction coefficients (herein termed *optical parameters*).

An important class of remote sensing applications aims at the measurement of the backscatter and extinction coefficients in the volume being scanned. Relevant examples are reflectivity weather radars [2], [3] and elastic-backscatter light detection and ranging (lidar) [4]–[14].

Fig. 1 illustrates a typical pulsed lidar remote sensing scenario. At regular time intervals the laser emits light pulses that

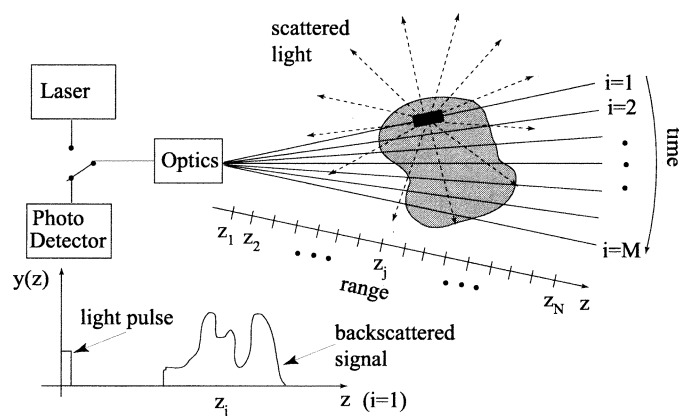


Fig. 1. Lidar remote sensing scenario. The light pulses transmitted by the laser are scattered by the clutter, and part of the energy reaches the photodetector yielding the signal  $y(z)$ .

are steered by an optical assembly toward a given region of the space; usually, the lateral coordinates of the beam are constant during a given number of pulses; these coordinates may, however, vary linearly from pulse to pulse to decrease the scan time, as it happens, for example, in weather radar. Part of the light propagating through the medium is scattered by the optical clutter and reaches the photodetector originating the received signal  $y(z)$ .

Determining the optical parameters is a hard inverse problem, namely, owing to the following features.

- 1) Mean power  $P$  is not known, and it has to be estimated from a series of noisy echoes.
- 2) The pair  $(\alpha, \beta)$  cannot be determined (even if the mean power  $P$  is exactly known), since (1) is a many-to-one mapping.
- 3) Optical parameters are usually nonhomogeneous and nonstationary, introducing additional complexity in the processing scheme.

### A. Classical Approaches

Relevant approaches for the inversion of optical parameters are the slope and the exponential fitting methods (e.g., see [8] and [15]), the Klett's stable implementation [5], [7] of the deterministic solution proposed by Hitschfeld and Bordan [16] (in the context of rain intensity measurements by weather radars), the extended Kalman filters [17], [18], and the nonlinear stochastic filter [19].

The exponential fitting technique applies to homogeneous clutter, i.e., the backscatter and extinction coefficients are

Manuscript received August 12, 2002; revised May 12, 2003. This work was supported by the Fundação para a Ciência e Tecnologia, under Projects POSI/3471/CPS/2000 and POSI/EEI/1580/2001-FASE II.

J. M. Bioucas Dias and J. M. N. Leitão are with the Instituto de Telecomunicações, Instituto Superior Técnico, 1049-001 Lisbon, Portugal (e-mail: bioucas@lx.it.pt; jleitao@lx.it.pt).

E. S. R. Fonseca is with the Departamento de Física, Universidade da Beira Interior, 6201-001 Covilhã, Portugal (e-mail: efonseca@mercury.ubi.pt).

Digital Object Identifier 10.1109/TGRS.2003.817216

assumed constant; the optical parameters are found by minimizing the square error between the observed data and the mean power given by (1). The slope method is also a least square fitting technique, but applied to

$$S(z) \equiv \log(z^2 P(z)).$$

The exponential fitting technique provides better results than the slope method [15], at the expense of a higher computationally burden associated with a nonlinear fitting problem.

For homogeneous clutter, the slope and the exponential fitting methods are robust in the sense that they do not assume any prior relation between  $\alpha$  and  $\beta$  (herein termed *constitutive relation*). However, the homogeneous assumption severely limits the application scope of these methodologies.

In order to invert function (1) with respect to optical parameters in the case of nonhomogeneous clutter, there is need for a constitutive relation. Several experimental and theoretical studies [20]–[22] have concluded that the power-law relation

$$\beta(z) = B_0 \alpha^c(z) \quad (2)$$

models with good accuracy situations where Mie scattering is dominant over Rayleigh scattering, and the multiple scattering is negligible [22] (e.g., hazy, cloudy, or foggy conditions or generally for infrared (IR) wavelengths). Parameters  $B_0$  (the so-called *backscatter-to-extinction ratio*) and the power law exponent  $c$  depend on the type of aerosol and on the laser wavelength. Constitutive relation (2) has been assumed, e.g., in [5], [15], and [17]–[19]. In [18], the backscatter-to-extinction ratio  $B_0$  is allowed to vary smoothly with range  $z$ . This makes sense whenever the light pulses propagate through optical clutters with different scattering characteristics. Herein, we assume that  $B_0$  does not vary with range.

Under the constitutive relation (2), the exact solution for the extinction coefficient is (e.g., see [5])

$$\alpha(z) = \frac{\exp\left[\frac{(S(z)-S_0)}{c}\right]}{\alpha_0^{-1} - \frac{2}{c} \int_{z_0}^z \exp\left[\frac{(S(z')-S_0)}{c}\right] dz'} \quad (3)$$

with  $S_0 \equiv S(z_0)$  and  $\alpha_0 \equiv \alpha(z_0)$ , where  $z_0$  is a given range. Klett [5] shows that the backward implementation of expression (3) (integrating from the farthest range backward) is more robust to errors on  $\alpha_0$  than the forward one, mainly for scenarios where the maximum range-integrated extinction coefficient (the so-called *optical thickness*) is *high*. According to his words: changing from forward to backward implementation “makes a very significant difference in the behavior of the solution. As  $z$  decreases,  $\alpha$  is now determined as the ratio of two numbers that each becomes progressively larger, so that the dependence of the solution on  $\alpha_0$  decreases with decreasing  $z$ .”

The application scope of expression (3), which we at times term the *deterministic solution*, is limited, as this solution depends on the mean power  $P$  [recall that  $S(z) = \log(z^2 P(z))$ ]. In practical applications,  $P$  is inferred by time and range averaging many samples. But how many samples should be averaged and what range and time weighting should be applied in order to have a good balance between bias and error covariance of the

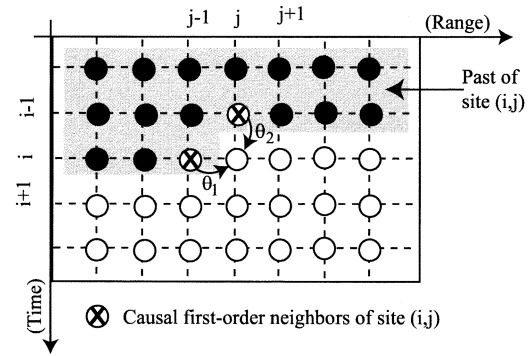


Fig. 2. Illustration of the 2-D scheme of sites where the observations are made. According to the adopted model, the crossed sites are the neighbors of site  $(i, j)$ , given the past.

solution? This problem is solved by the stochastic filter we develop in the remaining part of the paper; this filter applies time and range smoothing adaptively, as function of the observed data likelihood and of the prior.

Rocadenbosch *et al.* [17], [18] propose extended Kalman filters whose state-vectors include the optical parameters associated with a given range interval. The state-vector equation is a first-order autoregressive vector process; basically, it imposes, in a probabilistic sense, smoothness between two consecutive time estimates at a given range. The smoothness among components within the state-vector is achieved by imposing nonzero correlation on these components. The referred works embody the better features of the deterministic solution and of the least square fitting procedures: they do not assume homogeneous optical parameters and they address the problem at hand under the stochastic framework. Moreover, [18] identifies the extinction-to-backscattering constitutive parameter  $B_0$  by including it in the state-vector. Their major drawbacks are the lack of robustness of the estimates with respect to the state-space noise correlation matrix and its complexity, in the computational sense. For example, if the optical parameters are to be estimated within a range size of 5 km with a resolution of 20 m, the implementation of the extended Kalman filter involves operations with matrices of size  $500 \times 500$ .

## B. Proposed Approach

Assume that the observed data and the optical parameters are arranged into two-dimensional (2-D) fields (images); data along rows is associated with the echo of a given time pulse (range coordinate varying, time coordinate fixed), whereas data along columns is associated with a given range (time coordinate varying, range coordinate fixed). Fig. 2 schematizes a 2-D field where each site<sup>1</sup> is associated with a random variable with a given time and range.

We propose a stochastic filtering solution [23], [24] to the optical parameter estimation at hand. It relies on the probability density function (pdf) of the observed data, given the parameters, and on the state-space equation modeling smoothness among neighboring sites. The observed data are assumed to be random variables with pdfs depending on the optical parameters. In this paper, we consider only the Gaussian pdf, which

<sup>1</sup>By *site* we mean a range–time inversion cell.

models, for example, large aperture lidar applications [25]. As shown in the Appendix, the methodology is, however, easily adapted to other statistics, such as the exponential and Gamma pdfs, covering a wide range of scenarios in radar and lidar. The assumed space-state equation is the *reduced-order model* (ROM) [26] description of a causal first-order autoregressive (AR) Gauss–Markov random field (GMRF) [27]–[29]. The state variables are the extinction coefficient and the path-optical thickness coefficient (optical thickness).

Stochastic filtering is a Bayesian approach where the state-space equation plays the role of prior. In this paper, the prior enforces smooth variation of the optical parameters. This is attained by properly constraining the regression parameters of the AR model. The reason underlying the choice of a causal random field ROM description of the GMRF is that it leads to an efficient (in terms of computational complexity) stochastic filter for the problem at hand.

The estimation problem at hand is nonlinear in the following two senses.

- 1) The optical parameters do not relate linearly to the observed data.
- 2) The variance of the observation noise depends on the state-space vector.

Implementation of nonlinear recursive filters is known to be a hard problem because each recursion involves products and convolutions between pdfs. Finite filter representation is therefore a crucial aspect in nonlinear filtering. For the problem at hand, the propagation of the involved pdfs was derived and implemented with a point mass representation in [19]. In spite of the good results obtained therein, the computational complexity of the filter is unbearable in practical applications.

In this work, aiming at an efficient solution in computational terms, we represent the likelihood function as the Gaussian pdf that minimizes the *Kulback distance* [30] to the true likelihood function. The Gaussian representation allows a fruitful reinterpretation of the observation model: the observed data is a noisy version of the mean power (1), where the noise is additive Gaussian with mean and variance not depending on the state-vector, but rather on the observed data.

The filtering problem we were led to is still nonlinear, as the optical parameters relate in a nonlinear fashion with the observed data. By expanding the observation equation to first-order about the predicted state-vector, we obtain a linear approximation to the original function. The final filter obtained is of extended Kalman–Bucy type having *low* complexity. We stress, however, that the overall algorithm is not an extended Kalman–Bucy filter (EKBF), since the observation equation does not have the canonical form: nonlinear observations immersed in Gaussian additive noise with known statistics.

To assess the effectiveness of the proposed approach, we compare the derived filter with the standard particle filter (e.g., see [31] and [32]). Particle filters are sequential Monte Carlo approaches to nonlinear recursive filtering, where the posterior density is represented by a set of random samples with associated weights. As the number of samples becomes very large, the filter approaches the optimal Bayesian behavior. We give evidence that the proposed filter provides nearly optimum

performance, yet being much light, from the computational point of view, than the particle filter.

The parameters of the AR model and of the power-law constitutive relation are not known and have to be identified. To this purpose, we exploit the characteristics of pseudoinnovations process, a sequence of random variables internal to the proposed stochastic filter: the variance of this variable is minimized when the filter is optimal. Based on this fact, we propose a computationally efficient scheme to identify the filter parameters.

### C. Paper Overview

The paper is organized as follows. Section II introduces formally the optical parameter estimation and presents the observation mechanism, the prior model, and the ROM representation. Section III elaborates on the stochastic filtering algorithm and on its implementation, namely on the approximation of the likelihood function with a Gaussian pdf. Section IV presents the algorithm aimed at the identification of the filter parameters. Finally, Section V presents experimental results focusing on attenuation degree and on SNR, on robustness to model mismatch, and on the comparison with the deterministic method and with the particle filter.

## II. PROBLEM FORMULATION

Consider a pulsed remote sensing system operating at a fixed pulse rate. Let  $\mathcal{Z} \equiv \{(i, j) \mid i = 1, \dots, M, j = 1, \dots, N\}$  be a set of sites, where  $j$  refers to the range  $z_j = z_0 + j\delta$  ( $z_0$  and  $\delta$  are known constants), and  $i$  refers to the  $i$ th pulse, also termed time index  $i$ . Define also the lexicographical ordering  $n = j + (i - 1)N$ , associated with the index set  $\mathcal{Z}$ .

Define  $y_{i,j}$  and  $\alpha_{i,j}$  as, respectively, the observed signal and the extinction coefficient, both at site  $(i, j)$ . Define also the optical thickness  $\gamma_{i,j}$  as

$$\gamma_{i,j} \equiv \int_0^{z_j} \alpha_i(z') dz' \quad (4)$$

where  $\alpha_i(z)$  denotes the extinction coefficient at range  $z$  and pulse  $i$ .

Since the mapping from the bidimensional indexes to the one-dimensional ones is injective, we use the two indexing formats wherever necessary, i.e., we can write  $y_{i,j}$  or  $y_n$  to denote the same entity.

### A. Observation Model

In remote sensing lidar applications, the signal  $y_n$  at the output of a photodiode-based receiver has mean value and variance given by (e.g., see [25])

$$\bar{y}_n = P_n + v_d \quad (5)$$

$$\sigma_{y_n}^2 = \frac{P_n^2}{m} + b(P_n + v_d) + \sigma_{\text{th}}^2 \quad (6)$$

where  $P_n$  is the mean backscattered power;  $v_d$ ,  $P_n^2/m$ ,  $b(P_n + v_d)$ , and  $\sigma_{\text{th}}^2$  are, respectively, the variances of the dark current noise, of the speckle noise, of the shot noise, and of the thermal noise. Constant  $m$  in (6) is the speckle count [25], which depends on the size of the lidar aperture. According to expression

(1) and assuming the power law (2), the mean backscattered power  $P_n$  is given by

$$P_n = \frac{CB_0}{z_n^2} \alpha_n^c \exp(-2\gamma_n) \quad (7)$$

where  $C$  is a system-dependent constant,  $B_0$  and  $c$  are the power-law parameters, and  $z_n$ ,  $\alpha_n$ , and  $\gamma_n$  are, respectively, the range, the extinction coefficient, and the optical thickness, all at site  $n$ .

Herein we assume  $P_n^2/m \ll b(P_n + v_d)$ , meaning that the speckle noise is negligible. This scenario accurately models large aperture lidars with high speckle count figure, where the shot noise is predominant. We also assume that the filtered Poisson process at the output of the photodiode-based receiver is Gaussian. The central limit theorem underlies this approximation (e.g., see [33, ch. 4] and [34]).

In this paper, we adopt the state-vector  $\mathbf{x}_n \equiv [\alpha_n, \alpha_{n-N+1}, \gamma_n, \gamma_{n-N+1}]^T$ . The details of this choice will be given shortly. The conditional pdf  $p(y_n|\mathbf{x}_n)$  is

$$p(y_n|\mathbf{x}_n) \equiv \mathcal{N}[y_n - \bar{y}_n, \sigma_{y_n}^2] \quad (8)$$

with  $\mathcal{N}(y - m, \sigma^2)$  standing for a Gaussian pdf with mean  $m$  and variance  $\sigma^2$ . Note that  $p(y_n|\mathbf{x}_n)$  depends on the state-vector via the mean power  $P_n = P_n(\mathbf{x}_n)$ .

According to density (8), the random variable  $y_n$  can be reinterpreted as being the sum

$$y_n = \bar{y}_n + v_n \quad (9)$$

where  $v_n$  is a zero-mean Gaussian random variable of variance  $\sigma_{y_n}^2$ . We stress that  $\sigma_{y_n}^2$  depends on the mean power  $P_n$  and consequently on the state-vector. This precludes the direct application of Kalman–Bucy algorithm to the present problem, let alone the nonlinear nature of  $P(\mathbf{x}_n)$ .

## B. Prior Model

Fig. 2 schematizes the 2-D index set  $\mathcal{Z}$ . The shadowed region denotes the past, in a lexicographical sense, of site  $(i, j)$ . The crossed sites  $(i, j - 1)$  and  $(i - 1, j)$  are the first-order neighbors of site  $(i, j)$ , given the past.

We model the extinction coefficient  $\alpha_n$  and the optical thickness  $\gamma_n$  as causal first-order AR-GMRF [27]–[29] given by

$$\alpha_{i,j} = \theta_1 \alpha_{i,j-1} + \theta_2 \alpha_{i-1,j} + w_{\alpha(i,j)} \quad (10)$$

$$\gamma_{i,j} = \gamma_{i,j-1} + \delta \alpha_{i,j} + w_{\gamma(i,j)} \quad (11)$$

where  $\{w_{\alpha(i,j)}\}$  and  $\{w_{\gamma(i,j)}\}$  are mutually independent, zero-mean, independent Gaussian white fields with variances  $\sigma_{\alpha}^2$  and  $\sigma^2$ , respectively, and  $\theta_1, \theta_2 \geq 0$ , with  $\theta_1 + \theta_2 \leq 1$ .

In (10) and (11), when  $i = 1$  or  $j = 1$  or both, some boundary conditions have to be assumed, since indexes  $(i, j - 1)$  and/or  $(i - 1, j)$  take values outside the index set  $\mathcal{Z}$ . Herein, we adopt the so-called *free boundary condition*, which amounts to reducing the AR support at the boundary, in order not to include sites not in  $\mathcal{Z}$ .

Regression (10) gives the next extinction coefficient  $\alpha_{i,j}$  as a weighted mean of the past neighbors  $\alpha_{i,j-1}$  and  $\alpha_{i-1,j}$  plus a random component. Therefore, the proposed AR prior enforces smoothness in a statistical sense, whose strength is controlled by

$\sigma_{\alpha}^2$ . Regression (11) is a discrete approximation of the integral relation (4); noise  $w_{i,j}$  models the approximation error.

By introducing  $\alpha_{i,j}$  into (11), we obtain

$$\gamma_{i,j} = \gamma_{i,j-1} + \delta \theta_1 \alpha_{i,j-1} + \delta \theta_2 \alpha_{i-1,j} + w_{\gamma(i,j)} \quad (12)$$

where  $\{w_{\gamma(i,j)}\} = \{\delta w_{\alpha(i,j)} + w_{\gamma(i,j)}\}$  is a zero-mean Gaussian white field of variance  $\sigma_{\gamma}^2 = \delta^2 \sigma_{\alpha}^2 + \sigma^2$ . We note that  $w_{\alpha(i,j)}$  and  $w_{\gamma(i,j)}$  have cross correlation given by  $E[w_{\alpha(i,j)} w_{\gamma(i,j)}] = \delta \sigma_{\alpha}^2$ .

1) *State-Space Formulation*: In the recursive stochastic filtering setup, the prior has necessarily a state-space description. The AR model [(10)–(11)] admits a state-space description. However, its state-vector would include  $2(N + 1)$  components [ $(N + 1)$  for  $\alpha$  and  $(N + 1)$  for  $\gamma$ ] (see [29] and [35]). Values of the order of hundreds for  $N$  are usual in lidar applications. To avoid this huge state-vector, and the implicit computational complexity, we adopt the ROM proposed in [26], where the state-vector contains only the components whose indexes are in the support of the AR model.

The state-vector is therefore  $[\alpha_n, \alpha_{n+1-N}, \gamma_n, \gamma_{n+1-N}]^T$  and the ROM state-space formulation of (10) and (11) is expressed as

$$\underbrace{\begin{bmatrix} \alpha_{n+1} \\ \alpha_{n+2-N} \\ \gamma_{n+1} \\ \gamma_{n+2-N} \end{bmatrix}}_{\mathbf{x}_{n+1}} = \underbrace{\begin{bmatrix} \theta_1 & \theta_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \delta \theta_1 & \delta \theta_2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{\mathbf{A}_n} \underbrace{\begin{bmatrix} \alpha_n \\ \alpha_{n+1-N} \\ \gamma_n \\ \gamma_{n+1-N} \end{bmatrix}}_{\mathbf{x}_n} + \underbrace{\begin{bmatrix} 0 \\ \alpha_{n+2-N} \\ 0 \\ \gamma_{n+2-N} \end{bmatrix}}_{\mathbf{u}_n} + \underbrace{\begin{bmatrix} \omega_{\alpha(n)} \\ 0 \\ \omega_{\gamma(n)} \\ 0 \end{bmatrix}}_{\mathbf{w}_n} \quad (13)$$

with

$$\mathbf{Q}_n \equiv E[\mathbf{w}_n \mathbf{w}_n^T] = \begin{bmatrix} \sigma_{\alpha}^2 & 0 & \delta \sigma_{\alpha}^2 & 0 \\ 0 & 0 & 0 & 0 \\ \delta \sigma_{\alpha}^2 & 0 & \delta^2 \sigma_{\alpha}^2 + \sigma^2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (14)$$

where  $E[\cdot]$  denotes the mean value operator.

At the boundaries  $\{(1, j), (i, 1), i = 1, \dots, M; j = 1, \dots, N\}$ , matrix  $\mathbf{A}_{i,j}$  is set to

$$\mathbf{A}_{i,j} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \delta & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{A}_{i,1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (15)$$

Notice that vector  $\mathbf{u}_n$  in (13) plays the role of a deterministic input. This is, in fact, what happens in the ROM state-space formulation: past state-vector components not in the AR support act as known (already estimated) deterministic inputs.

The filter based on the ROM state-space representation exchanges optimality with computational complexity. A pertinent question is, of course, what is the degradation introduced by the ROM approach. In [36], experiments comparing both filters produced comparable visual results as well as comparable mean-square errors. This issue is, however, beyond the scope of this paper. For a deeper treatment, see [26], [29], [36], and [37].

### III. STOCHASTIC FILTERING SOLUTION

In the previous section, we derived the observation (9) and proposed the ROM representation (13) of the AR-GMRF (10) and (11) as prior for the clutter parameters. In summary, we have

$$\mathbf{x}_{n+1} = \mathbf{A}_n \mathbf{x}_n + \mathbf{u}_n + \mathbf{w}_n \quad (16)$$

$$y_n = h(\mathbf{x}_n) + v_n \quad (17)$$

where  $h(\mathbf{x}_n) \equiv \bar{y}_n$  is given by (5), and  $\{v_n\}$  is a zero-mean white Gaussian sequence of variance  $\sigma_{y_n}^2$  given by (6).

Stochastic recursive filtering is a Bayesian estimation technique that, based on a state-space recursive model (not necessarily linear) and on an observation equation, recursively propagates the pdf  $p(\mathbf{x}_n | \mathbf{Y}_n)$  (the so-called *filtering density*), where  $\mathbf{Y}_n \equiv \{y_n, y_{n-1}, \dots, y_1\}$ . It is worth noting that, from a Bayesian point of view, the filtering density conveys all the information about  $\mathbf{x}_n$ , the entity to be estimated.

The recursive propagation of the filtering density is accomplished by implementing the following steps (for a detailed explanation, see [24]).

#### 1) Prediction

$$p(\mathbf{x}_{n+1} | \mathbf{Y}_n) = \int_{\mathbb{R}^4} p(\mathbf{x}_{n+1} | \mathbf{x}_n) p(\mathbf{x}_n | \mathbf{Y}_n) d\mathbf{x}_n. \quad (18)$$

#### 2) Filtering

$$p(\mathbf{x}_{n+1} | \mathbf{Y}_{n+1}) \propto p(y_{n+1} | \mathbf{x}_{n+1}) p(\mathbf{x}_{n+1} | \mathbf{Y}_n) \quad (19)$$

where  $p(\mathbf{x}_{n+1} | \mathbf{x}_n)$  and  $p(y_{n+1} | \mathbf{x}_{n+1})$  are the so-called *convolution kernel* and *observation factor*, respectively.

When both the state equation and the observation equation are linear, the state and observation noises are independent of the state-vector and Gaussian, and the initial filtering density  $p(\mathbf{x}_0)$  is also Gaussian, the recursive propagation of the filtering density is computed by the Kalman–Bucy filter [38]. This filter exploits the fact that the convolution and the product of Gaussian pdfs (prediction and filtering steps) yields Gaussian shapes [24]. Therefore, the prediction and filtering steps are given by simple linear operations involving mean vectors and covariance matrices (the first two moments).

When any of the above conditions fails, at least one of the pdfs involved is not Gaussian, and the problem falls into the general setup of stochastic nonlinear filtering [23], [24], [31], [32], [38]–[41]. In this case, a second-order representation might be far from optimal.

In the problem we are addressing, the observation (17) and the variance  $\sigma_{y_n}^2$  of the observation noise  $v_n$  are nonlinear functions of the state-vector  $\mathbf{x}_n$ . Accordingly, the problem is nonlinear. A solution for the problem at hand has been developed and tested within the nonlinear framework in [19] using a point mass filter [42]. This solution consists, basically, in approximating the pdfs involved in the filtering process with point mass densities.

Although the point mass filter might produce results close to the exact ones (providing that the sampling interval used in representing the pdfs is *small*), it exhibits quite often complexity unbearable to most applications.

In the remaining part of this section, we develop a suboptimal linear filtering scheme exhibiting far less complexity than the exact nonlinear filter, and nevertheless able to produce results

close to the optimal ones. The goodness of the estimates is illustrated in Section IV, by comparing results obtained with the standard particle filter (e.g., see [31] and [32]).

#### A. Extended Kalman–Bucy-Type Filter

According to the rationale developed in Section II-A, the observation factor  $p(y_n | \mathbf{x}_n)$  is the Gaussian density  $\mathcal{N}(y_n - \bar{y}_n, \sigma_{y_n}^2)$ . As stated above, the problem at hand is nonlinear, since: 1) the function  $P_n(\mathbf{x}_n)$ , given by (7), is nonlinear; and 2) the shape of  $p(y_n | P_n)$  as function of  $P_n$  is not Gaussian. To lighten the filtering complexity of the problem at hand, we derive in the Appendix, a Gaussian approximation to  $p(y_n | P_n)$  interpretable,<sup>2</sup> as the filtering step is concerned, in terms of the observation equation

$$z_n = \tilde{h}(\mathbf{x}_n) + \tilde{v}_n \quad (20)$$

where  $z_n \equiv |y_n + \sigma_{\text{th}}^2/b|$ ,  $\tilde{h}(\mathbf{x}) \equiv P(\mathbf{x}) - b + v_d + \sigma_{\text{th}}^2/b$ , and  $\tilde{v}_n$  is a zero-mean white Gaussian sequence of variance  $\sigma_n^2 \equiv b(2b + z_n)$ , independent of the state noise  $\mathbf{w}_n$ .

The EKBF for the nonlinear model (16) and (20) is the Kalman–Bucy filter for the linear model

$$\mathbf{x}_{n+1} = \mathbf{A}_n \mathbf{x}_n + \mathbf{u}_n + \mathbf{w}_n \quad (21)$$

$$z_n = \tilde{h}(\hat{\mathbf{x}}_{n|n-1}) + \mathbf{H}_n(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1}) + \tilde{v}_n \quad (22)$$

where (22) is the observation equation linearized about the state prediction  $\hat{\mathbf{x}}_{n|n-1}$  and  $\mathbf{H}_n = \nabla \tilde{h}(\hat{\mathbf{x}}_{n|n-1})$  is given by [see expression of  $P$  given by (7)]

$$\mathbf{H}_n = P(\hat{\mathbf{x}}_{n|n-1}) \begin{bmatrix} \frac{c}{\hat{\alpha}_{n|n-1}}, 0, -2, 0 \end{bmatrix}. \quad (23)$$

As in the nonlinear case, the Kalman–Bucy filter implements, in a recursive fashion, a prediction step and a filtering step (for a detailed explanation, see [24]).

#### 1) Prediction

$$\begin{aligned} \hat{\mathbf{x}}_{n|n-1} &= \mathbf{A}_{n-1} \hat{\mathbf{x}}_{n-1|n-1} + \mathbf{u}_{n-1} \\ \Sigma_{n|n-1} &= \mathbf{A}_{n-1} \Sigma_{n-1|n-1} \mathbf{A}_{n-1}^T + \mathbf{Q}_{n-1}. \end{aligned} \quad (24)$$

#### 2) Filtering

$$\begin{aligned} \hat{\mathbf{x}}_{n|n} &= \hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n \left[ z_n - \tilde{h}(\hat{\mathbf{x}}_{n|n-1}) \right] \\ \mathbf{K}_n &= \Sigma_{n|n-1} \mathbf{H}_n^T (\mathbf{H}_n \Sigma_{n|n-1} \mathbf{H}_n^T + \sigma_n^2)^{-1} \\ \Sigma_{n|n} &= (\mathbf{I} - \mathbf{K}_n \mathbf{H}_n) \Sigma_{n|n-1} \end{aligned} \quad (25)$$

where  $\hat{\mathbf{x}}_{n|n-1} \simeq E[\mathbf{x}_n | \mathbf{Z}_{n-1}]$ ,  $\hat{\mathbf{x}}_{n|n} \simeq E[\mathbf{x}_n | \mathbf{Z}_n]$ ,  $\Sigma_{n|n-1}$  and  $\Sigma_{n|n}$  are approximations<sup>3</sup> of the conditional error covariance matrixes of  $\hat{\mathbf{x}}_{n|n-1}$  and of  $\hat{\mathbf{x}}_{n|n}$ , respectively,  $\sigma_n^2 = b(2b + z_n)$ , and  $\mathbf{Z}_n \equiv \{z_n, \dots, z_1\}$ . The filter is initialized with  $\hat{\mathbf{x}}_{0|0} = E[\mathbf{x}_0]$  and  $\Sigma_{0|0} = \text{Cov}(\mathbf{x}_0)$ .

<sup>2</sup>We stress that this interpretation is only symbolic. It just means that  $p(y_n | P_n) \simeq \mathcal{N}(z_n - \tilde{h}(P_n))$ , both as functions of  $P_n$ . See the Appendix for details.

<sup>3</sup>The estimates  $\hat{\mathbf{x}}_{n|n-1}$  and  $\hat{\mathbf{x}}_{n|n}$  are approximations (possibly poor) of  $E[\mathbf{x}_n | \mathbf{Z}_{n-1}]$  and  $E[\mathbf{x}_n | \mathbf{Z}_n]$ , respectively, as linearization (22) is an approximation to  $z_n$  and noise  $\tilde{v}_n$  is not Gaussian. The same rationale applies also to  $\Sigma_{n|n-1}$  and  $\Sigma_{n|n}$ .

### B. Backward Filtering

In Section II, we have introduced the lexicographical ordering  $n = j + N(i - 1)$ , where  $j$  refers to the range  $z_j = z_0 + j\delta$  ( $z_0$  and  $\delta$  are known constants), and  $i$  refers to the  $i$ th pulse. The proposed filter evolves forward, from the closest range  $z_0$  toward the farthest range  $z_m$ . In this section, we assume that  $z_j = z_m - (j-1)\delta$ , i.e., the filter evolves backward, from the farthest range  $z_m$  toward the closest range  $z_0$ . To distinguish both filters, we term the former *forward filter* and the latter *backward filter*.

Assume that the mean power  $P_{i,N}$  and the extinction coefficient  $\alpha_{i,N}$ , both corresponding to the  $i$ th time-pulse and to the farthest range  $N$ , are available. Dividing the observation (39) derived in the Appendix by the the mean power  $P_{i,N}$  corresponding to the farthest range, we get

$$z_{0n} \equiv \underbrace{P_0(\mathbf{x}_n) - b_0 + v_{0d}}_{\tilde{h}_0(\mathbf{x}_n)} + \underbrace{\frac{\sigma_{th0}^2}{b}}_{\tilde{v}_{0n}} - v_{P_{0n}} \quad (26)$$

where the subscript 0 refers to the same entities of (39) (see the Appendix) normalized by the mean power. The normalized power  $P_0$  is given by

$$P_0(\mathbf{x}_{i,j}) \equiv \frac{P_{i,j}}{P_{i,N}} = \left(\frac{z_N}{z_j}\right)^2 \left(\frac{\alpha_{i,j}}{\alpha_{i,N}}\right)^c \exp(2\gamma_{i,j}) \quad (27)$$

with  $\gamma_{i,j}$  being now defined as

$$\gamma_{i,j} \equiv \int_{z_j}^{z_N} \alpha_i(r) dr. \quad (28)$$

The derivation of the backward filter parallels the forward one with the exception of gradient (23), which is now given by

$$\mathbf{H}_n = P_0(\hat{\mathbf{x}}_{n|n-1}) \left[ \frac{c}{\hat{\alpha}_{n|n-1}}, 0, 2, 0 \right]. \quad (29)$$

Therefore, the backward filter is also implemented according to the recursive steps (24) and (25).

From (27) and (29), we see that the backward filter does not depend neither on the system constant  $C$  nor on the backscatter-to-extinction ratio  $B_0$ . In addition, the backward filter is more robust to numerical errors, particularly for large values of both the optical thickness and the SNR. The reason is basically the same underlying the advantage of the backward implementation of deterministic solution (3) over the forward one.

### C. Parameter Identification

The filter parameters are unknown and must be estimated. The forward filter is parameterized with the prior parameters  $\theta_1, \theta_2, \sigma_\alpha, \sigma_\gamma$ , the constitutive parameters  $B_0, c$ , and the receiver parameter  $C$ , whereas the backward filter is parameterized only with the prior parameters  $\theta_1, \theta_2, \sigma_\alpha, \sigma_\gamma$  and the constitutive parameter  $c$ . Notice that, in the case of the forward filter, we need only to identify the product  $CB_0$ .

The partitioned filter [43], [44] is a possible line of attack to stochastic filtering with unknown parameters. This method was already applied to lidar return power and log power estimation [45]. In order to apply the partitioning scheme, one has to build a

bank of EKBFs, each one matched to an appropriate member of the parameters set. In the present setup, any reasonable number of filters would lead to a computational complexity much higher than we wish.

Aiming at parameter identification, we propose a simple scheme based on the *expected* properties of the pseudoinnovations process  $\{\tilde{z}_n = z_n - \tilde{h}(\hat{\mathbf{x}}_{n|n-1})\}$ . Assuming that the filter operation is optimum, i.e.,  $\tilde{h}(\hat{\mathbf{x}}_{n|n-1}) = E[z_n | z_{n-1}, z_{n-2}, \dots, z_1]$ , then the variance  $\sigma_{\tilde{z}}^2$  of  $\tilde{z}_n$  underbounds the variance of the innovation process generated by any suboptimal filter [46, chap. 10]. Based on this rationale and on the assumption that the optical parameters are quasi-stationary on the time coordinate, we propose the following procedure:

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} J(\boldsymbol{\theta}) \quad (30)$$

with

$$J(\boldsymbol{\theta}) = \frac{1}{M_1 N_1} \sum_{i=1}^{M_1} \sum_{j=1}^{N_1} \tilde{z}_{i,j}^2(\boldsymbol{\theta}) \quad (31)$$

where  $\boldsymbol{\theta} = (\theta_1, \theta_2, \sigma_\alpha, \sigma_\gamma, CB_0)$  for the forward filter and  $\boldsymbol{\theta} = (\theta_1, \theta_2, \sigma_\alpha, \sigma_\gamma)$  for the backward filter. Minimization (30) should take into account the constraints  $\theta_1, \theta_2 \geq 0$ ,  $\theta_1 + \theta_2 = 1$ , and  $\sigma_\alpha, \sigma_\gamma \geq 0$ .

Notice that parameter  $c$  is not included in  $\boldsymbol{\theta}$ . The reason is that it is not identifiable: the inversion of (7) with respect to  $(B_0, c, \alpha_n)$ , for  $n = 1, \dots, MN$ , is not one-to-one: there are many triplets  $(B_0, c, \alpha_n)$  yielding the same mean power  $P_n$ . We cope with this difficulty by setting  $c = 1$  and identifying the product  $CB_0$  that best explains the observed data. This setting is justified by the following.

- The extinction coefficient  $\alpha_n$  exhibits low sensitivity to  $c$ . This can be drawn from (3), and it is illustrated in Section IV-B.
- Theoretical and experimental studies have led to the conclusion that  $c \in [0.67, 1]$  when Mie scattering is dominant over Rayleigh scattering, and the multiple scattering is negligible [22] (e.g., for hazy, clouds, or foggy conditions or generally for IR wavelengths). The exact value depends on the laser wavelength and on the aerosol.

The computation of  $\hat{\boldsymbol{\theta}}$  given by (30) is done using a coordinate descent technique. We have found that minimizations with respect to  $(\theta_1, \theta_2)$ ,  $CB_0$ , and  $\sigma_\alpha, \sigma_\gamma$  are practically decoupled. This greatly simplifies the inference of vector  $\boldsymbol{\theta}$ .

## IV. EXPERIMENTAL RESULTS

The backward recursive stochastic filter derived in the previous section is now applied to simulated data. The robustness of this filter to model mismatches, referred to in Section III-B, underlies this choice.

To apply the backward filter one needs the mean power  $P_{i,N}$  and the extinction coefficient  $\alpha_{i,N}$ , both at the farthest range  $N$ , for  $i = 1, \dots, M$ . We have computed estimates of these entities based on the set  $\{y_{i,j}, i = 1, \dots, M, j = N - 9, \dots, N\}$ , i.e., based on a small number of columns of  $y_{i,j}$  close to the farthest range. Concerning the power  $P_{i,N}$ , we averaged these columns

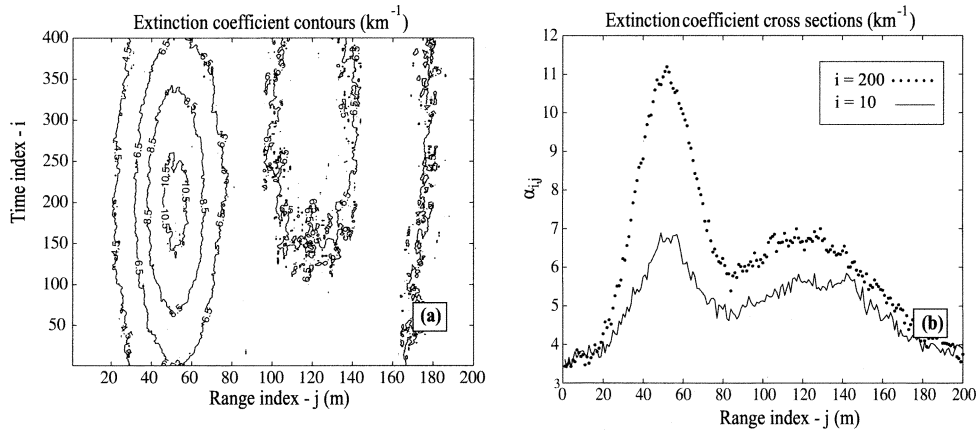


Fig. 3. Simulated extinction coefficient. (a) Contour plot. (b) Two cross sections of (a).

and then applied a lowpass Gaussian filter in the time direction. Concerning the extinction coefficient  $\alpha_{i,N}$ , we applied the slope method to  $\{y_{i,j}, i = 1, \dots, M, j = N - 9, \dots, N\}$ . The estimates of  $P_{i,N}$  and  $\alpha_{i,N}$  are, of course rough and might contain larger errors. Nevertheless, as it is next illustrated, the backward filter is able to recover from them most of the times.

Fig. 3(a) shows contour plots of a simulated extinction coefficient surface. It consists of a random component generated according to the recursive model [(10)–(11)] plus two Gaussian elevations with their major axes parallel to the time coordinate. This orientation reflects quasi stationarity of the optical parameters on the time dimension, a typical situation on pulsed remote sensing systems. Fig. 3(b) shows cross sections of the extinction coefficient profile at times  $i = 10$  and  $i = 200$ . Since we set  $\delta = 1$  m, the range index  $j$  is associated with range  $z = j$  m. The noise standard deviations and the AR parameters of model (10) and (11) were set to  $(\sigma_\alpha = 0.07, \sigma_\gamma = 0)$  and  $(\theta_1 = 0.1, \theta_2 = 0.9)$ , respectively. The value  $\sigma_\gamma = 0$  means that we are neglecting the error in computing the optical thickness from discrete samples of the extinction coefficient, what is acceptable assuming a *small* sampling interval. The values  $\theta_1 = 0.1, \theta_2 = 0.9$  account for the above-referred quasi stationarity of the optical parameters.

For the extinction coefficient shown in Fig. 3, the maximum optical thickness is  $\gamma_{\max} = 1.25$  and occurs at time  $i = 200$ . In order to study the behavior of the proposed method with respect to the optical thickness, we also present estimation results for a more severe attenuation scenario with a maximum optical thickness of  $\gamma_{\max} = 2.5$ . The extinction coefficient data for this experiment is the image shown in Fig. 3 scaled by a factor of 2.

The extinction coefficient shown in Fig. 3 is not exactly generated according to the recursive model [(10)–(11)], since we are adding a deterministic component. Our aim in doing this is to illustrate the robustness of the proposed filters against the presence of deterministic components, as long as the magnitude of their increasing/decreasing rates is not much higher than the standard deviation  $\sigma_\alpha$  of the white noise driving the AR-GMRF (10).

Table I summarizes the simulation parameters. Constant  $b$  determines the magnitude of the observed data for a given SNR; we set it to  $b = 1$ , since it does not affect the filter performance. The system parameter  $C$  is selected in order to have a given SNR at a given range.

TABLE I  
SIMULATION PARAMETERS

Constitutive	Model	System					
$B_0$ ( $\text{sr}^{-1}$ )	$c$	$\theta_1$	$\theta_2$	$C$	$\sigma_{i_b}^2$ ( $\text{Wkm}^3$ )	$b$	$v_d$
$5 \times 10^{-2}$	1.0	0.1	0.9	-	100	1	0

TABLE II  
IDENTIFIED PARAMETERS

	$\gamma_{\max} = 1.25$		$\gamma_{\max} = 2.5$	
	Simul.	Estim.	Simul.	Estim.
$\alpha_0$	3.5	4.0	7.5	7.7
$\theta_1$	0.1	0.11	0.1	0.12
$\sigma_\alpha$	0.07	0.09	0.14	0.16

The forthcoming simulation results are divided into four parts. In Section IV-A, we conduct experiments to show how the filters are affected by the optical thickness magnitude and by the SNR. In Section IV-B, we study the filter robustness to model mismatches. In Section IV-C, we present comparisons of the proposed filters with the deterministic method, and finally, in Section IV-D, we compare the proposed stochastic approach with the *sampling importance resampling (SIR) filter* proposed in [47].

#### A. Attenuation Degree and SNR

Table II displays, for two attenuation scenarios, the parameters used in the simulation and their estimated values according to the methodology proposed in Section III-C. In order to speed up the identification procedure, we have computed the pseudoinnovation and its covariance only in the set of sites  $\{(i, j) \mid i = 1, \dots, M, j = 1, \dots, 50\}$ , a small subset of  $\mathcal{Z}$ .

Fig. 4 shows filtering results. The underlying observed data were generated with basis on the attenuation map shown in Fig. 3. The maximum optical thickness is  $\gamma_{\max} = 1.25$ . Fig. 4(a) plots a cross section at time index  $i = 200$ ; Fig. 4(b) and (d) plots cross sections at range indexes 50 and 120, respectively; and Fig. 4(c) plots the sample square estimation error of the extinction coefficient shown in Fig. 4(a). This sample error is of course a rough estimate of the mean square error, since the successive time estimate errors for a given range are not stationary. However, given that the optical parameters vary smoothly along

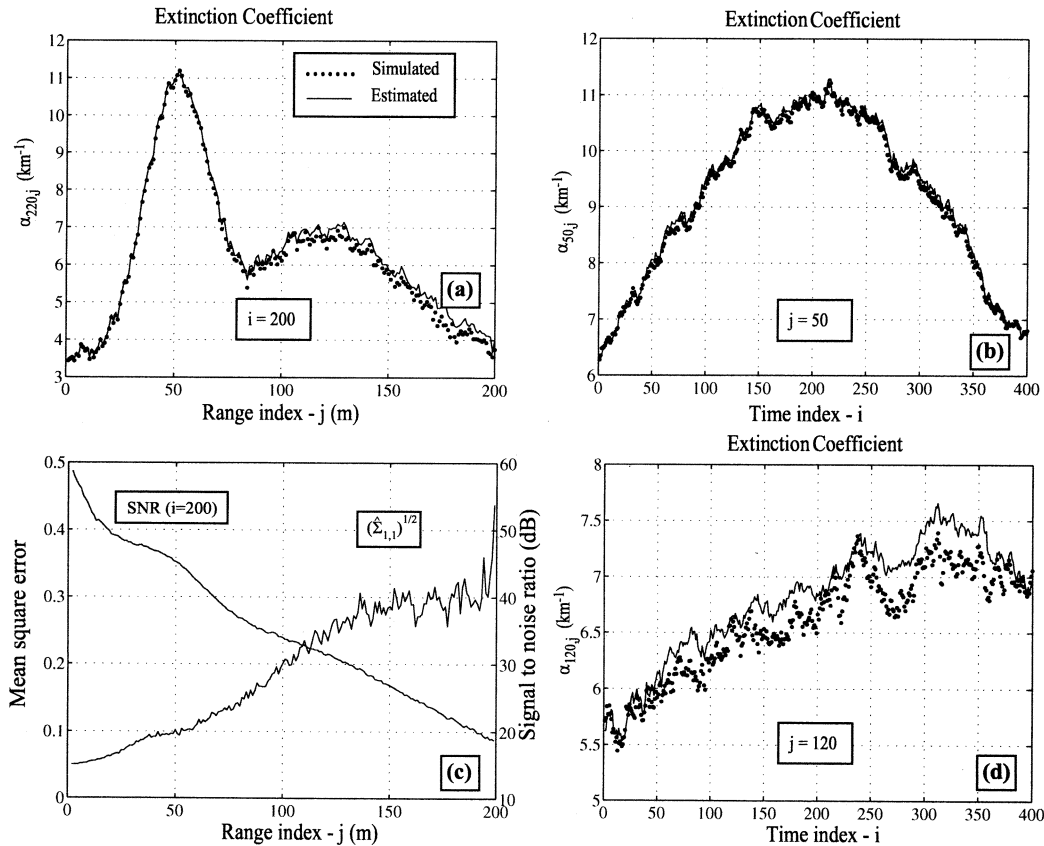


Fig. 4. Filtering results (maximum optical thickness  $\gamma_{\max} = 1.25$ ). (a), (b), and (d) plot cross sections of the extinction coefficient. (c) Plot of the SNR and the sample mean square error of the estimate shown in (a).

the time coordinate, one can expect that the sample square estimation error be indicative of the mean square error (letting alone the finite sample size issue). We have also plotted the SNR evolution at time index  $i = 200$ .

The backward filter was initialized with

$$\mathbf{x}_0 = [\hat{\alpha}_0, 0, 0, 0]^T \quad \Sigma_0 = \text{diag}([10, 0, 0, 0]).$$

Thus, the optical thickness and its error covariance are correctly initialized with 0, while the extinction coefficient and its covariance are initialized with  $\hat{\alpha}_0$  and 10, respectively. In order to get good filter performance at the farthest ranges, it is mandatory to have good estimates of  $\alpha_0$ . Nevertheless, if the optical thickness is *large*, the filter is able to recover even from large errors in the estimates of  $\alpha_0$ . This is illustrated in the next section.

The estimates are very close to the simulated data. As expected, the filter performance degrades along range as the SNR decreases. The larger mean square error occurs in the neighborhood of range  $j = 200$ , this being due to estimation errors of  $\alpha_0$  (error of  $\hat{\alpha}_0$  is 12.5%).

Fig. 5 shows estimation results obtained by applying the proposed filter to data generated with basis on the attenuation map shown in Fig. 3 scaled by a factor of 2. The maximum optical thickness is now  $\gamma_{\max} = 2.5$ . Fig. 5(a)–(d) is as in Fig. 4(a)–(d). Comparing Figs. 4(c) and 5(c), we conclude that the filter performance tends to be better where the SNR is higher, regardless of the attenuation scenario. The exceptions occur about range  $j = 60$ , as it can be read from Fig. 5(c), which is due to strong model mismatches.

### B. Model Mismatch Robustness

Fig. 6 illustrates the robustness of the proposed scheme to mismatches on the filter parameters. Solid and dotted curves represent estimated and simulated data, respectively. Estimates are obtained with filter parameters given in the box. Parameter  $\beta$  parameterizes the error on the initial extinction coefficient  $\alpha_m$ . The maximum optical thickness is  $\gamma_{\max} = 2.5$ , and the minimal SNR is  $\text{SNR}_{\min} = 10$  dB.

The proposed filter is robust with respect to errors on the prior parameters  $c$ ,  $\theta_1$ , and  $\sigma_\alpha$ . Notice that errors of, respectively, 10%, 100%, and 100% on these parameters do not lead to visible errors on the estimates, as it can be read from the curve parameterized with  $\beta = 1.0$ . This is not the case with respect to the parameter  $\alpha_m$ . At ranges near  $z_m$ , the estimate of the backscattering coefficient is close to  $\hat{\alpha}_m$ . Therefore, large errors on  $\hat{\alpha}_m$  imply large error on the estimates of  $\alpha_n$  in the neighborhood of  $z_m$ . However, the filter sensitivity to errors on  $\hat{\alpha}_m$  decreases as the range decreases. For example, an error of 50% on  $\alpha_m$  causes practically no impact on filter estimates for ranges less than 100 m.

The explanation for the filter behavior with respect to errors on  $\alpha_m$  can be found in the deterministic solution (3). In the backward solution, the function of  $z$

$$-\frac{2}{c} \int_{z_m}^z \exp \left[ \frac{(S(z') - S_0)}{c} \right] dz'$$

which is present in the denominator of (3), is positive and increasing. Therefore, if it takes values much larger than  $\alpha_m^{-1}$  for



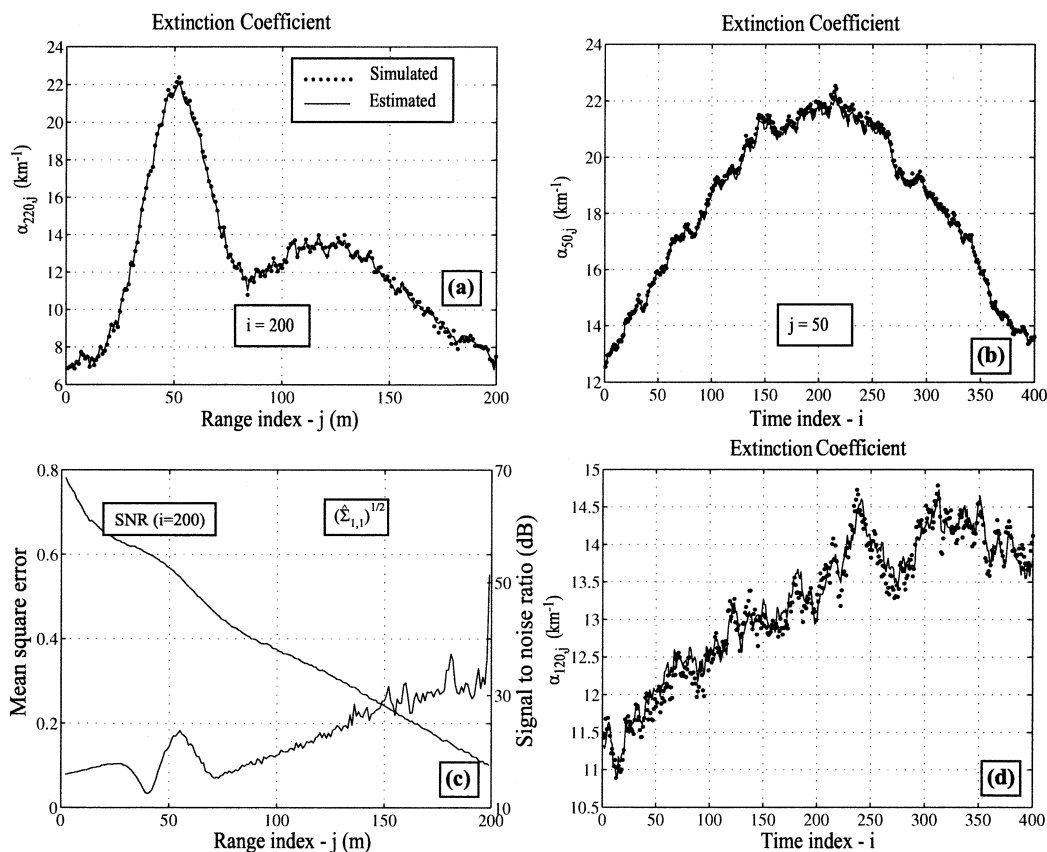


Fig. 5. As in Fig. 4, for  $\gamma_{\max} = 2.5$ .

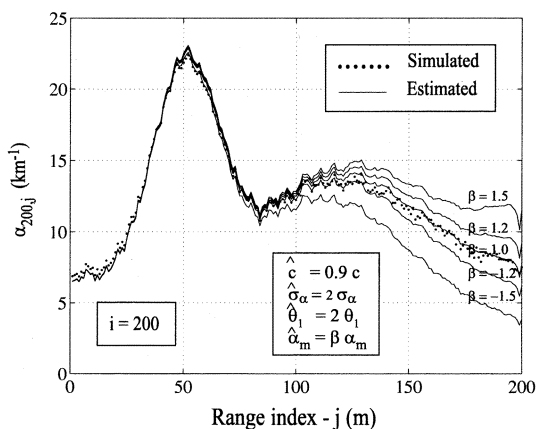


Fig. 6. Robustness to filter parameters mismatches. Solid and dotted curves represent estimated and simulated data, respectively. Estimates are obtained with filter parameters given in the box. Parameter  $\beta$  parameterizes the error on the initial extinction coefficient  $\alpha_m$ . The maximum optical thickness is  $\gamma_{\max} = 2.5$ , and the minimal SNR is  $\text{SNR}_{\min} = 10$  dB.

some  $z'$ , then the dependency of  $\alpha(z)$  on  $\alpha_m^{-1}$  is very small for  $z \lesssim z'$ . As a rule of thumb, we conclude, therefore, that the larger the optical thickness is, the more robust the filter estimates are at a given range.

### C. Comparison With the Deterministic Method

Fig. 7 shows the extinction coefficient estimates generated by the proposed backward filter (gray solid curves) and by the deterministic method (black solid curve), for  $\gamma_{\max} = 1.25$ ,

$\text{SNR}_{\max} = 50$  dB,  $\text{SNR}_{\min} = -5$  dB, and  $\sigma_{\text{th}} = 1000$ . These values were selected in order to have  $\text{SNR} \leq 10$  dB, approximately, for  $i > 100$  m. Note that we wish to compare both filters at low SNR, for the stochastic filter estimates tends to the deterministic solution (3) as  $\text{SNR} \rightarrow \infty$ . Both filters were initialized with the correct value of  $\alpha_m = \alpha_{i,N}$ .

In Fig. 7(a), the solution (3) was applied directly to the observed data at time index  $i = 200$ , whereas in Fig. 7(b), it was applied to lowpass filtered data obtained with a rectangular sliding window of size  $20(\text{time}) \times 1(\text{range})$ . Fig. 7(c) and (d) is as Fig. 7(b) and (c), but for range index  $j = 50$ .

The stochastic approach produces much better results than the deterministic one, even when the former is applied to smoothed data. Of course one could argue that if the data smoothness was increased, the deterministic solution would improve; but how much smoothness should be imposed in order not to bias the solution? And how? Should it be space-invariant or not? These questions do not arise in the proposed stochastic solution because it adapts the amount of smoothness depending on the the observed data SNR and on the prior strength controlled by the variance  $\sigma_\alpha$ .

As the  $\text{SNR}_{\min}$  increases the differences between the deterministic and the stochastic solutions vanishes, as the weight of the prior vanishes.

### D. Comparison With the SIR Filter

In this section, we compare the proposed stochastic approach with the SIR filter proposed in [47]. The SIR algorithm is a

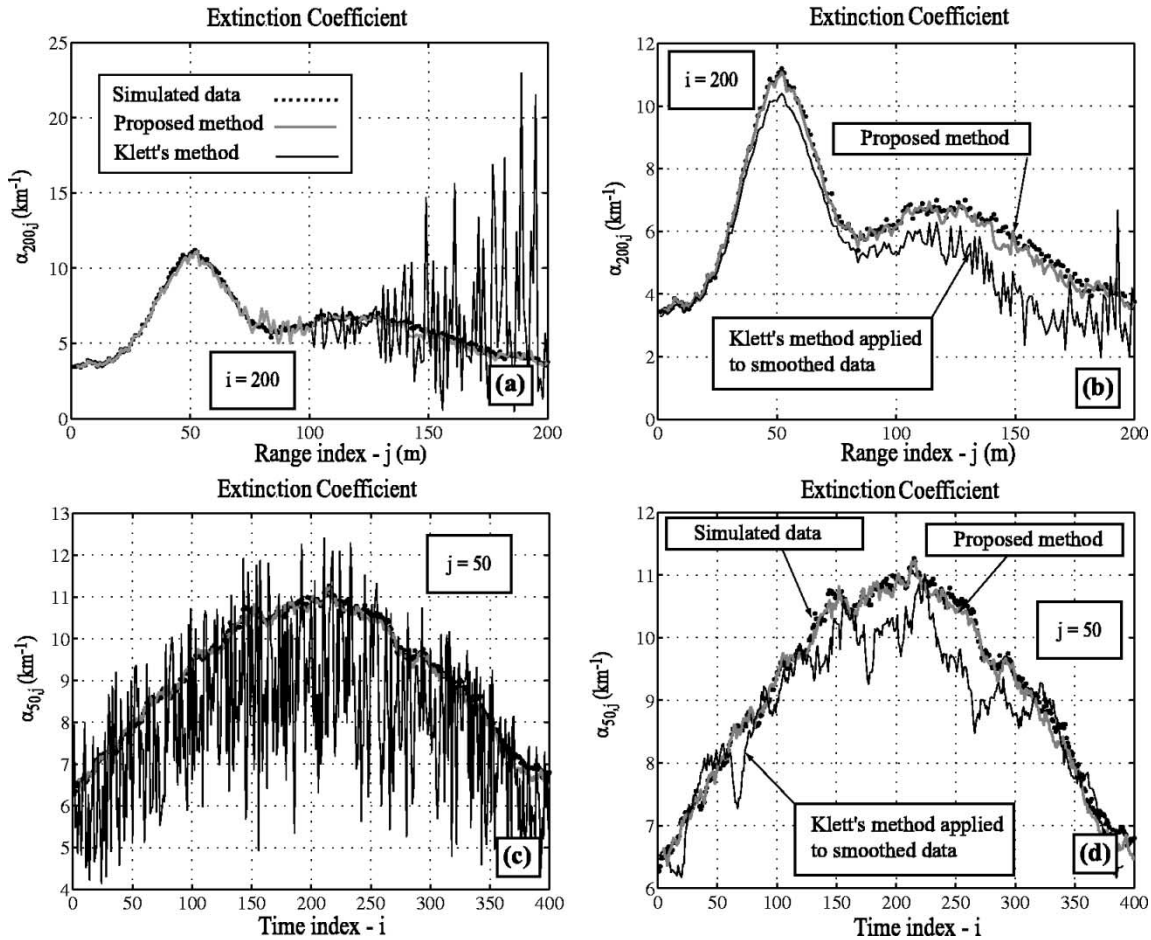


Fig. 7. Comparison with the deterministic solution (3) for  $\gamma_{\max} = 1.25$ . Dotted curves represent true data, black solid curves represent estimates given by (3), and gray solid curves represent estimates obtained by the proposed backward filter.

sequential Monte Carlo technique aimed at the computation of the posterior  $p(\mathbf{x}_n | \mathbf{Y}_n)$ . This density is approximated by a set of random samples, the so-called *particles*, with associated weights. As the number of particles goes to infinity, the information content in the particles and respective weights is equivalent to that of density  $p(\mathbf{x}_n | \mathbf{Y}_n)$ . Therefore, we use a large number of particles in the SIR filter and compare the estimates based on the particles with those provided by the proposed filter. An example of application of particle filters in refractivity estimation from sea clutter radar returns, roughly similar to the lidar problem herein considered, may be found in [48].

The SIR filter is a particularization of the *sequential importance sampling* algorithm, also known as bootstrap filter, condensation algorithm, particle filter, interacting particle approximations, and survival of the fittest (see [32] and references therein). Herein, we have implemented the SIR filter according to [32, Alg. 4]. The sampling step of the  $i$ th particle at time  $n$ ,  $\mathbf{x}_n^i \sim p(\mathbf{x}_n^i | \mathbf{x}_{n-1}^i)$ , has been performed according to the state (21); as in the proposed stochastic filter,  $\mathbf{u}_n$  is the conditional mean estimate of  $[0, \alpha_{n+2-N}, 0, \gamma_{n+2-N}]^T$  [see (13)], already available at time  $n$ .

Fig. 8(a) shows the extinction coefficient estimates generated by the proposed forward filter (gray solid curves) and by the SIR

filter using 100 particles (black solid curve), for  $\gamma_{\max} = 1.25$ ,  $\text{SNR}_{\max} = 50$  dB,  $\text{SNR}_{\min} = 5$  dB, and  $\sigma_{\text{th}} = 100$ . In both filters, the constitutive parameters, the prior parameters, and the initial extinction coefficient were set to their true values.<sup>4</sup>

Both estimates are close to the true data. The square root of the sample mean square error is  $0.15 \text{ km}^{-1}$  for the proposed filter and  $0.2 \text{ km}^{-1}$  for the SIR filter with ten particles. For 100 particles or more (we have tested up to 1000 particles), the error achieves the lower bound of  $0.16 \text{ km}^{-1}$ . As for complexity, the particle filter implementation with 100 particles needs, roughly, 40 times the number of floating-point operations used by the proposed filter.

One could argue that a more sophisticated particle filter could lead to better estimates. That might be the case. Better schemes to combat the so-called *degeneracy phenomenon* are the optimal importance sampling introduced in [50] and the auxiliary particle filters introduced in [51]. *Particle impoverishment* can also be minimized by using better resampling schemes, such as the systematic resampling proposed in [52], and the *resample-move* introduced in [53]. These approaches achieve the performance

<sup>4</sup>We note that it is also possible to jointly estimate the sequence of hidden state vectors and the unknown model parameters in a true Bayesian fashion (e.g., see [49]).

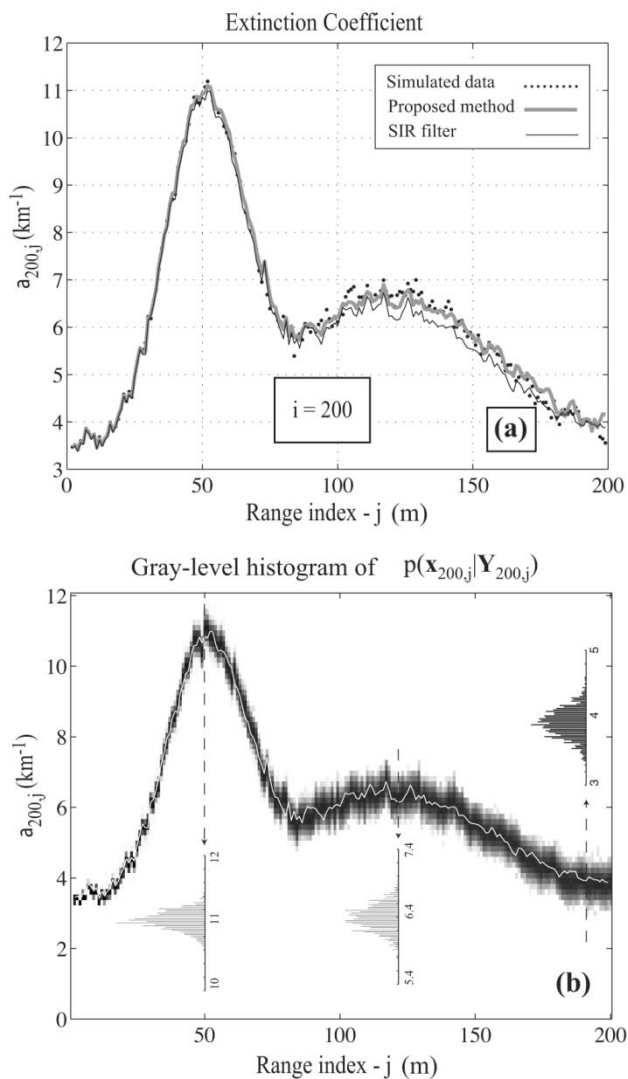


Fig. 8. Comparison with the SIR filter for  $\gamma_{\max} = 1.25$ . (a) Cross section of the filtering results. Dotted curve represents true data. The black solid curves represents estimates given by SIR filter using 100 particles, and gray solid curves represent estimates given by the proposed forward filter. Part (b) shows in gray levels and for each range index the histogram associated with  $p(\mathbf{x}_{200,j} | \mathbf{Y}_{200,j})$ . A detailed view of the histogram is presented for  $j = 50, 125$ , and  $190$ . The SIR estimate is also plotted (white solid line).

of SIR filter with fewer particles and consequently with lower computational complexity. We call the attention, however, for Fig. 8(b), where it is plotted, in gray levels and for range indexes  $j = 1, \dots, 200$ , the histogram associated with  $p(\mathbf{x}_{200,j} | \mathbf{Y}_{200,j})$  obtained with 1000 particles. We see that the *particle impoverishment* effect is not severe, indicating that the estimates are not compromised by this effect. We conclude then that the performance of proposed filter is nearly optimal. Since its complexity is comparable with that of a SIR filter with roughly three particles, it would be very difficult for any sort of particle filter with this low number of particles to achieve the performance of the proposed adaptive extended Kalman filter.

We call the attention to the detailed histograms presented in Fig. 8(b), for  $j = 50, 125$ , and  $190$ ; they are unimodal and exhibit roughly a Gaussian shape. This fact justifies, at least partially, the direction we have followed: approximating the posterior density with a Gaussian shaped function.

## V. CONCLUDING REMARKS

A new stochastic filtering technique for lidar reconstruction of the backscattering and the extinction coefficients was developed; we adopted a first-order autoregressive Gauss–Markov random field as Bayesian prior model for the observed image of extinction coefficients. In order to derive an extended Kalman type filtering structure, we adopted the following strategy.

- 1) Approximate the space-state equation with the *reduced-order model* description of a causal first-order autoregressive Gauss–Markov random field. For this to be possible, we chose the extinction coefficient and the optical thickness as state variables.
- 2) Approximate the observed data random variable with a Gaussian one, using the Kullback divergency criterion.

After these approximations, we ended up with a linear state-space equation and with a nonlinear observation equation. Extended Kalman filtering was then applied. Since the state-vector has only four components (due to the reduced-order model description of the GMRF), the complexity of this filter is *low*. For example, an image of size  $400(\text{time samples}) \times 200(\text{range samples})$  was processed in a few seconds, using an ordinary personal computer.

In deriving the stochastic filter, we considered two variants: the forward filter, which evolves toward the farthest range, and the backward filter, which evolves toward the nearest range. The latter is more robust to initialization and numerical errors and does not depend neither on the backscatter-to-extinction ratio nor on the system constant. These advantages underlay the choice of the backward filter in Section IV.

The filter was applied to simulated data generated according to the adopted GMRF to which a deterministic component consisting of two Gaussian elevations was added. The performance of the proposed method was studied with respect to the attenuation degree and SNR, to model mismatch robustness, and to the deterministic model and to the sampling importance resampling filter (SIR) comparisons. Roughly, the proposed stochastic filter yielded effective estimates if  $\text{SNR} \geq 5$  dB. The list below points out the main findings relative to the filter robustness with respect to the model parameters.

- 1) The filter sensitivity to the parameter backscatter-to-extinction ratio  $B_0$  is zero and is moderate to the power-law parameter  $c$ . For example, a model mismatch of 10% on  $c$  produced a marginal increment on the error estimates.
- 2) The filter sensitivity to prior parameters (i.e., to  $\theta_1$ ,  $\theta_2$ ,  $\sigma_\alpha$ , and  $\sigma_\gamma$ ) is very small. For example, a model mismatch 100% on  $\theta_1$  and on  $\sigma_\alpha$ , produced marginal increments on the error estimates.
- 3) The filter sensitivity to the initial extinction at the farthest range is high, decreasing as the filter evolves toward the closest range. The sensitivity decreases faster on scenarios with large optical thickness.

Concerning the comparison with the deterministic method, we concluded that the stochastic approach yielded much better results than the deterministic solution. On the basis of this behavior is the adaptive nature of the proposed algorithm: the amount of smoothness depends on the the observed data SNR and on the prior strength controlled by variance of the extinction

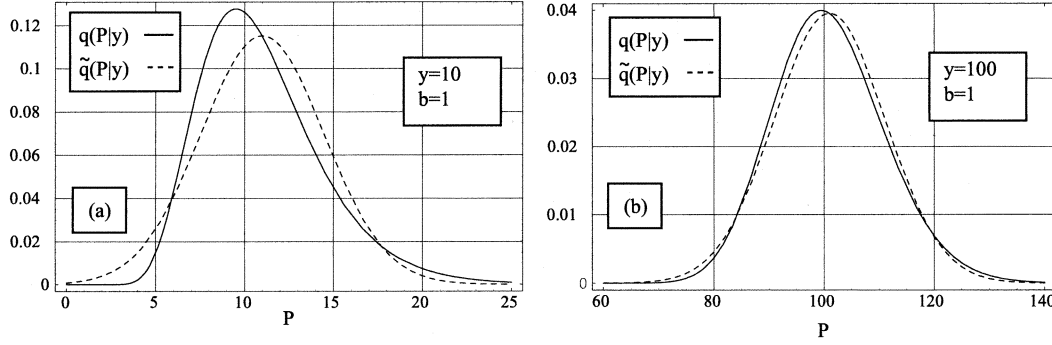


Fig. 9. Gaussian approximation of the observation factor  $p(y|P)$ . Solid and dashed lines represent the shape of the observation factor and of the Gaussian approximation, respectively. (a) Low SNR. (b) Medium SNR.

coefficient. As the  $\text{SNR}_{\min}$  increases, the differences between the deterministic and the stochastic solutions vanishes, as the weight of the prior vanishes.

Concerning the comparison with the SIR filter, we concluded that the proposed algorithm yields similar results with far less computational burden. The approximated Gaussian shape of the posterior density underlies this scenario, i.e., since the posterior is well approximated by a Gaussian shape, there is no point in applying Monte Carlo techniques to propagate the posterior pdf.

Regarding future work, we foresee the replacement of the proposed causal Gauss–Markov random field with a noncausal one. In this way, all data image (past and future) is used to infer the optical parameters at a given site, thus obtaining better estimates. This can be achieved by writing the prior density of the complete image in terms of the so-called *clique potentials* (see [54]) and then maximizing the posterior pdf.

#### APPENDIX APPROXIMATING THE OBSERVATION FACTOR WITH A GAUSSIAN DENSITY

Consider the pdf  $q(P|y) \propto p(y|P)$  defined as (to ease the exposition, subindexes were dropped)

$$q(P|y) \equiv \frac{p(y|P)}{\int_{-\infty}^{\infty} p(P|y) dP} \quad (32)$$

where  $p(y|P) \equiv \mathcal{N}(y - \bar{y}, \sigma_y^2)$ ,  $\bar{y} = P + v_d$ , and  $\sigma_y^2 = b(P + v_d) + \sigma_{\text{th}}^2$ . Given that the denominator of (32) does not depend on  $P$ , finding a Gaussian shape that best fits  $p(y|P)$  or  $q(P|y)$ , as functions of  $P$ , are equivalent problems.

Let us define  $\tilde{q}(P|m_P, \sigma_P^2) \equiv \mathcal{N}(P - m_P, \sigma_P^2)$ . Parameters  $m_P$  and  $\sigma_P$  are to be chosen such that  $\tilde{q}(P|m_P, \sigma_P^2)$  approximates  $q(P|y)$ . As in [55], we adopt the minimum Kullback divergency [30] criterion

$$(m_P, \sigma_P^2) = \arg \min_{m', \sigma'^2} \underbrace{\int_{-\infty}^{\infty} q(P|y) \ln \frac{q(P|y)}{\tilde{q}(P|m', \sigma'^2)} dP}_{D(q||\tilde{q})} \quad (33)$$

where  $D(q||\tilde{q})$  is the Kullback divergency [30], which measures the dissimilarity between density functions  $q$  and  $\tilde{q}$ . The solution to the minimization (33) is (e.g., see [30])

$$\begin{cases} m_P = E_q[P] \\ \sigma_P^2 = E_q[(P - m_P)^2] \end{cases} \quad (34)$$

which applied to pdf  $q(P|y)$  leads to

$$\begin{cases} m_P = b + z - v_d - \frac{\sigma_{\text{th}}^2}{b} \\ \sigma_P^2 = b(2b + z) \end{cases} \quad (35)$$

where  $z \equiv |y + \sigma_{\text{th}}^2/b|$ .

In (35), the mean value  $m_P$  and the variance  $\sigma_P^2$  depend on  $z = |y + \sigma_{\text{th}}^2/b|$ . Although not immediately evident, this is a consequence of  $q(P|z) = q(P| -z)$ .

Fig. 9(a) shows the shape of  $q(P|y)$  and  $\tilde{q}(P|y)$  for  $\sigma_{\text{th}}^2 = 0$ ,  $z = 10$ , and  $b = 1$ ; Fig. 9(b) shows the same functions for  $\sigma_{\text{th}}^2 = 0$ ,  $z = 100$  and  $b = 1$ . The Kullback divergencies are  $D(q||\tilde{q}) = 0.065$  for Fig. 9(a) and  $D(q||\tilde{q}) = 0.0065$  for Fig. 9(b).

Defining the SNR as

$$\text{SNR} = \frac{E^2[y]}{\sigma_y^2} = \frac{(P + v_d)}{b + \sigma_{\text{th}}^2(P + v_d)^{-1}} \quad (36)$$

we can loosely refer to Fig. 9(a) and (b) as a low SNR scenario and a medium SNR scenarios, respectively. Note that shape of the observation factor tends to Gaussian as  $z/b$  increases, what is in accordance with the ten times lower Kullback divergency of scenario Fig. 9(b).

The Gaussian approximation to  $p(y|P)$  is then given by

$$q(P|y) = \frac{1}{\sqrt{2\pi}\sigma_P} e^{-(P - m_P)^2/2\sigma_P^2} \quad (37)$$

$$= \frac{1}{\sqrt{2\pi}\sigma_P} e^{-(z - \tilde{h})^2/2\sigma_P^2} \quad (38)$$

where  $\tilde{h} = P - b + v_d + \sigma_{\text{th}}^2/b$ . From (38), we see that  $q(P|y) = \mathcal{N}(z - \tilde{h}, \sigma_P^2)$ . Thus, one might tempt to interpret  $z$  as a Gaussian random vector with mean  $\tilde{h}$  and variance  $\sigma_P^2$ . In an equivalent form

$$z = \underbrace{P(\mathbf{x}) - b + v_d + \frac{\sigma_{\text{th}}^2}{b}}_{\tilde{h}(\mathbf{x})} + \tilde{v} \quad (39)$$

where  $\tilde{v}$  is zero-mean Gaussian noise of variance  $\sigma_P^2$ . This interpretation is, however, incorrect because  $\mathcal{N}(z - \tilde{h}, \sigma_P^2)$  is not Gaussian as function  $z$ . We adopted it, however, as a symbolic representation of the density (38). It plays the the role of the observation equation in our stochastic filtering problem. Notice that it is still nonlinear on  $\mathbf{x}_n$ , but the variance of the additive

noise does not depend now on the state-vector, but rather on the observation  $z$ .

The methodology proposed to approximate the observation factor with a Gaussian density may be applied to any density  $p(y|P)$ . For example, if the speckle noise is dominant and assuming that the thermal noise is negligible, the observation factor is Gamma distributed, thus given by

$$p(y|P) = \frac{1}{\Gamma(m)} \left(\frac{m}{P}\right)^m y^{m-1} e^{-my/P}$$

where  $E[y] = P$  and  $\sigma_y^2 = P^2/m$ , and  $m$  is the speckle count. The solution to the minimization (33), assuming that  $m$  is greater than 3, is now given by

$$m_P = y \frac{m}{m-2}$$

$$\sigma_P^2 = y^2 \frac{Pm^2}{(m-2)(m-3)}.$$

#### REFERENCES

- [1] A. Ishimaru, *Wave Propagation and Scattering in Random Media*. Piscataway, NJ: IEEE Press, 1997.
- [2] R. J. Doviak and D. S. Zrnic, "The Doppler weather radar," in *Aspects of Modern Radar*, E. Brookner, Ed. Norwood, MA: Artech House, 1988.
- [3] —, *Doppler Radar and Weather Observations*. New York: Academic, 1993.
- [4] F. Fernald, B. Herman, and J. Reagan, "Determination of aerosol height distributions by lidar," *J. Appl. Meteorol.*, vol. 11, pp. 482–489, 1972.
- [5] J. D. Klett, "Stable analytical inversion solution for processing lidar returns," *Appl. Opt.*, vol. 20, pp. 211–220, 1981.
- [6] F. Fernald, "Analysis of atmospheric lidar observations: Some comments," *Appl. Opt.*, vol. 23, no. 5, pp. 652–653, 1984.
- [7] J. D. Klett, "Lidar inversion with variable backscatter/extinction ratios," *Appl. Opt.*, vol. 24, pp. 1638–1985, 1985.
- [8] G. Kunz and G. Leeuw, "Inversion of lidar signals with the slope method," *Appl. Opt.*, vol. 32, no. 18, pp. 3249–3256, 1993.
- [9] A. I. Carswell, "Lidar remote sensing of atmospheric aerosols," *Proc. SPIE*, vol. 1312, pp. 206–220, 1990.
- [10] Y. Sasano, E. Browell, and S. Ismail, "Error caused by using a constant extinction/backscattering ratio in the lidar solution," *Appl. Opt.*, vol. 24, pp. 3929–3932, 1985.
- [11] V. Kovalev and H. Moosmüller, "Distortion of particulate extinction profiles measured with lidar in a two-component atmosphere," *Appl. Opt.*, vol. 33, pp. 6499–6507, 1994.
- [12] M. Matsumoto and N. Takeuchi, "Effects of misestimated far-end boundary values on two common lidar inversion solutions," *Appl. Opt.*, vol. 33, no. 27, pp. 6451–6456, 1994.
- [13] V. Matthias, C. Böckmann, V. Freudenthaler, G. Pappalardoand, J. Bösenberg, V. Amiridis, A. Amodio, A. Ansmannand, D. Balis, A. Bosellia, A. Chaykovski, G. Chourdakis, A. Comeron, A. Delaval, F. De Tomasi, R. Eixmann, M. Frioud, A. Hagard, M. Iarlori, L. Komguem, S. Kreipl, G. Larcheveque, H. Linne, R. Matthey, I. Mattis, A. Papavannis, J. Pelon, R. M. Perrone, R. Persson, D. P. Resendes, V. Rizi, F. Rocadenbosch, J. A. Rodriguez, L. Sauvage, J. Schneider, R. Schumacher, V. Shcherbakov, V. Simeonov, U. Wandinger, X. Wang, M. Wiegner, and C. Zerefos, "Lidar Intercomparisons on algorithm and system level in the frame of EARLINET," Max-Planck-Institut. Meteorologie, Stuttgart, Germany, MPI-Rep. 337, 2002.
- [14] C. Böckmann, U. Wandinger, A. Ansmann, J. Bösenberg, V. Amiridis, A. Boselli, A. Delaval, F. De Tomasi, M. Frioudand, A. Hagard, M. Iarlori, L. Komguem, S. Kreipl, G. Larcheveque, V. Matthias, A. Papavannis, G. Pappalardo, F. Rocadenbosch, J. Schneider, V. Shcherbakov, and M. Wiegner, "EARLINET: Backscatter lidar algorithm intercomparison," in *Proc. 21st Int. Laser Radar Conf.*, G. Roy, L. Bissonnette, and G. Vallée, Eds., 2002, pp. 353–356.
- [15] F. Rocadenbosch, A. Comeron, and D. Pineda, "Assessment of lidar inversion errors for homogeneous atmospheres," *Appl. Opt.*, vol. 37, no. 12, pp. 2199–2206, 1998.
- [16] W. Hitschfeld and J. Bordan, "Errors inherent in the radar measurement of rainfall at attenuating wavelengths," *J. Meteorol.*, vol. 11, pp. 58–67, 1954.
- [17] F. Rocadenbosch, G. Vazquez, and A. Comeron, "Adaptive filter solution for processing lidar returns: Optical parameter estimation," *Appl. Opt.*, vol. 37, no. 30, pp. 7019–7034, 1998.
- [18] F. Rocadenbosch, C. Soriano, A. Comeron, and J. Baldasano, "Lidar inversion of atmospheric backscatter and extinction-to-backscatter ratios by use of a Kalman filter," *Appl. Opt.*, vol. 38, no. 15, pp. 3175–3189, 1999.
- [19] J. Dias and E. Susana, "Nonlinear stochastic filtering technique for radar/lidar inversion," *Proc. SPIE*, vol. 3809, pp. 51–62, July 1999.
- [20] J. Curcio and G. Knestric, *J. Opt. Soc. Amer.*, no. 48, p. 686, 1958.
- [21] R. Pinnik, S. Jennings, P. Chylek, C. Ham, and W. Grandy, Jr, "Backscatter and extinction in water clouds," *J. Geophys. Res.*, vol. 88, pp. 6787–6796, 1983.
- [22] S. Twomey and H. Howell, "The relative merit of white and monochromatic light for the determination of visibility by backscattering measurements," *Appl. Opt.*, vol. 4, pp. 501–506, 1965.
- [23] J. Moura, "Stochastic filtering: Linear and nonlinear," in *Signal Processing*, J. L. Lacoume and R. Stora, Eds. Amsterdam, The Netherlands: North Holland, 1987.
- [24] R. S. Bucy, *Lectures on Discrete Time Filtering*. New York: Springer-Verlag, 1994.
- [25] B. J. Rye, "Power ratio estimation in incoherent backscatter lidar: Direct detection with Gaussian noise," *Appl. Opt.*, vol. 28, no. 17, pp. 3639–3646, 1989.
- [26] D. Angwin and H. Kaufman, "Image restoration using reduced order models," *Signal Process.*, vol. 16, pp. 21–28, 1989.
- [27] J. Woods, "Two-dimensional discrete Markovian fields," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 232–240, Mar. 1972.
- [28] R. Chellappa, "Two-dimensional discrete Gaussian Markov random field models for image processing," in *Progress in Pattern Recognition*, L. Kanal and A. Rosenfeld, Eds. Amsterdam, The Netherlands: Elsevier, 1985.
- [29] J. Woods and C. Radewan, "Kalman filtering in two dimensions," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 473–482, July 1977.
- [30] S. Kullback, *Information Theory and Statistics*. New York: Peter Smith, 1978.
- [31] A. Doucet, J. Freitas, and N. Gordon, "A introduction to sequential Monte Carlo methods in practice," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, J. Freitas, and N. Gordon, Eds. New York: Springer-Verlag, 2001, ch. 1.
- [32] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Processing*, vol. 50, pp. 174–188, Feb. 2002.
- [33] D. L. Snyder, *Random Point Process*. New York: Wiley, 1975.
- [34] A. Papoulis, "High density shot noise and Gaussianness," *J. Appl. Prob.*, vol. 8, no. 1, pp. 118–127, 1971.
- [35] D. Angwin and H. Kaufman, "Nonhomogeneous image identification and restoration procedures," in *Digital Image Restoration*, A. Katsaggelos, Ed. Berlin, Germany: Springer-Verlag, 1991, ch. 7, pp. 177–208.
- [36] D. Angwin, "Adaptive image restoration using reduced order model based Kalman filter," Ph.D dissertation, Rensselaer Polytech. Inst., Troy, NY, 1989.
- [37] J. Kim and J. Woods, "A new interpretation of romkf," *IEEE Trans. Image Processing*, vol. 6, pp. 599–601, Aug. 1997.
- [38] A. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970.
- [39] R. Bucy, C. Hecht, and K. D. Senne, "An engineer's guide to buiding nonlinear filters," Frank J. Seiler Res. Lab., Colorado Springs, CO, Tech. Rep. SRL-TR-72-0004, vol. 1 and 2, 1989.
- [40] R. S. Bucy and K. D. Senne, "Nonlinear filtering algorithms for vector processing machines," in *Computation and Mathematics With Applications*. New York: Pergamon, 1980, vol. 6, pp. 317–338.
- [41] A. Strandlie and J. Zerubia, "Particle tracking with iterated Kalman filters and smoothers: The PMHT algorithm," *Comput. Phys. Commun.*, vol. 123, pp. 77–86, 1999.
- [42] R. S. Bucy and K. D. Senner, "Digital synthesis of nonlinear filters," *Automatica*, vol. 7, pp. 287–298, 1971.
- [43] D. Lainiotis, "Partitioning: A unifying framework for adaptive systems, I: Estimation," *Proc. IEEE*, vol. 64, pp. 1126–1143, 1976.
- [44] —, "Partitioned estimation algorithms II: Nonlinear estimation," *J. Inform. Sci.*, vol. 7, pp. 202–235, 1974.
- [45] D. Lainiotis, P. PapaparasKeva, G. Kothapalli, and K. Plataniotis, "Adaptive filter applications to lidar: Return power and log power estimation," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 886–891, July 1996.
- [46] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

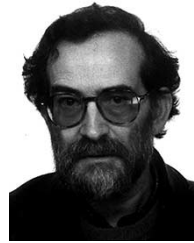
- [47] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear and non-Gaussian state estimation," *Proc. Inst. Elect. Eng. F*, vol. 140, no. 2, pp. 107–113, 1993.
- [48] S. Vasudevan and J. Krolik, "Refractivity estimation from radar clutter by sequential importance sampling with a Markov model for microwave propagation," in *Proc. ICASSP*, 2001, pp. 2905–2908.
- [49] J. Liu and M. West, "Combined parameter and state estimation in simulation-based filtering," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, J. Freitas, and N. Gordon, Eds. New York: Springer-Verlag, 2001, ch. 10.
- [50] A. Doucet, "On sequential Monte Carlo methods for Bayesian filtering," Dept. Eng., Univ. Cambridge, Cambridge, U.K., 1998.
- [51] M. Pitt and N. Shephard, "Filtering via simulation: Auxiliary particle filters," *J. Amer. Statist. Assoc.*, vol. 94, no. 446, pp. 590–599, 1999.
- [52] G. Kitagawa, "Monte Carlo filter for smoother and non-Gaussian nonlinear state-space models," *J. Comput. Graph. Statist.*, vol. 5, no. 1, pp. 1–25, 1996.
- [53] W. Gilks and C. Berzuini, "Following a moving target—Monte Carlo inference for Bayesian dynamic models," *J. R. Statist. Soc. B*, vol. 63, pp. 127–146, 2001.
- [54] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. R. Statist. Soc. B*, vol. 36, no. 2, pp. 192–225, 1974.
- [55] J. Leitão and M. Figueiredo, "Absolute phase image reconstruction: A stochastic nonlinear filtering approach," *IEEE Trans. Image Processing*, vol. 7, pp. 868–882, June 1998.



**José M. Bioucas Dias** (S'87–M'95) received the E.E., M.Sc., and Ph.D. degrees in electrical and computer engineering, all from Instituto Superior Técnico (IST), Technical University of Lisbon, Lisbon, Portugal, in 1985, 1991, and 1995, respectively.

He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, IST. He is also a Researcher with the Communication Theory and Pattern Recognition Group, IST. His research interests include remote sensing, signal and

image processing, pattern recognition, and communications.



**José M. N. Leitão** (M'95) received the E.E. and Ph.D. degrees in 1970 and 1983, respectively, both in electrical engineering, and the "Agregado" degree in electrical and computer engineering in 1992, all from the Instituto Superior Técnico (IST), Technical University of Lisbon, Lisbon, Portugal.

He was with the Laboratory of Physiology, Instituto Gulbenkian de Ciência, Oeiras, Portugal, from 1970 to 1972. After spending three years at the University of Tübingen, Tübingen, Germany, he joined IST in 1976, where he is currently a Full

Professor with the Department of Electrical and Computer Engineering. He is also the Coordinator of the Communication Theory and Pattern Recognition Group, Institute of Telecommunications, IST. His main research interests are communication and information theory, pattern recognition, and signal and image processing.



**Elsa S. R. Fonseca** received the undergraduate degree in physics engineering in 1995, and the M.S. degree in electrical and computer engineering in 1998, both from Instituto Superior Técnico, Lisbon, Portugal. She is currently pursuing the Ph.D. degree in opto-thermal radiometry at the Universidade da Beira Interior, Covilhã, Portugal.

She is currently a Teaching Assistant with the Physics Department, Universidade da Beira Interior. Her research interests include applied optics in remote sensing and biomedical applications,

opto-thermal radiometry, and biomedical optics.