Advanced Interacting Sequential Monte Carlo Sampling for Inverse Scattering

F Giraud[1,2], P Minvielle[1] and P Del Moral[2]

[1] CEA, DAM, CESTA, F-33114 Le Barp, FRANCE

[2] INRIA Bordeaux Sud-Ouest & Institut de Mathématiques, Université Bordeaux I, 33405 Talence cedex, France

E-mail: francois.giraud@cea.fr, pierre.minvielle@cea.fr, pierre.del_moral@inria.fr

Abstract. The following electromagnetism (EM) inverse problem is addressed. It consists in estimating local radioelectric properties of materials recovering an object from global EM scattering measurements, at various incidences and wave frequencies. This large scale ill-posed inverse problem is explored by an intensive exploitation of an efficient 2D Maxwell solver, distributed on high performance computing machines. Applied to a large training data set, a statistical analysis reduces the problem to a simpler probabilistic metamodel, from which Bayesian inference can be performed. Considering the radioelectric properties as a hidden dynamic stochastic process that evolves according to the frequency, it is shown how advanced Markov Chain Monte Carlo methods – called Sequential Monte Carlo (SMC) or interacting particles – can take benefit of the structure and provide local EM property estimates.

1. Introduction

Inverse scattering is a topic of major importance; it encompasses various applications [1, 2, 3] in acoustics, optics and electromagnetism, e.g. medical imaging, tomography, ionospheric sounding or SAR (Synthetic Aperture Radar). In electromagnetism (EM), the direct scattering problem is the determination of the scattered field, due to the scattering of an incident wave in the presence of inhomogeneities, when the geometrical and physical properties of the scatterer are known. Conversely, inverse scattering is defined as "inferring information on the inhomogeneity from knowledge of the far-field pattern..." [2]; it is an inverse problem. In this paper, we focus on a specific, though worthwhile, EM inverse scattering issue. The aim is to estimate the electromagnetic properties of materials from global microwave scattering measurements. Related applications can be located at the crossroads of non-destructive testing, quality control and material measurement. Many EM material characterization techniques have been developed in the domain of agricultural and food materials, radar absorbers [4], etc. Most of these techniques, from the transmission lines to the admittance tunnel method, require small-scale material test samples. For instance, transmission lines enclosed

samples inside the conductors of a transmission-line sample holder. Although the EM properties (i.e. permeability and permittivity) can be measured, they can differ significantly from the final product's ones, when the materials are assembled and placed on the full-scaled object or system [4]. The so-called free-space RCS (Radar Cross Section: scalar that quantifies reflectivity) methods [4] can overcome this pitfall by measuring the monostatic reflectivity of a large planar sample. The sample is then located inside an anechoic chamber, in the far field of the transmitting and receiving antennas. The reflectivity is measured at various arrival angles of the incident wave. Besides, let mention the classic bistatic alternative in near field, known as the NRL arch method [4]. In this paper, we focus on the following challenging inverse scattering problem: the control and evaluation of EM properties of a full-scaled objet or mock-up from the global reflectivity measurements in a free-space RCS device. Deviations of microwave properties, such as permeability and permittivity, are to be determined along the object.

Nearly 50 years ago, a closely related issue was formerly outlined in [5]. Least-square optimization was applied to determine the dielectric constants that made the analytically computed RCS fit with measurements. This issue reemerged in a slightly different way in [6]: both complex permittivity and permeability of a lossless plane stratified medium were evaluated. More recently, [7] considers the reconstruction in microwave tomography of the dielectric properties of a strongly inhomogeneous object by a stochastic global optimization algorithm, based on simulated annealing. Similarly, [8] develops a pseudoinversion algorithm for 2D imaging, with the aim of locating and estimating the dielectric permittivities of unknown inhomogeneous dielectric cylindrical objects. It corresponds to an important class of inverse scattering problems, which differs markedly from our topic and consists in determining the constitutive parameters of bounded objects embedded in a known medium. Measurements are performed at different points of the scattered field, that results from a known single-frequency wave illumination. On the whole, it is known to be an ill-posed inverse problem. Like image reconstruction and many other imaging inverse problems [9, 10], it necessitates, at some step, a regularization procedure: it tends to eliminate the artificial oscillations resulting from the ill-posedness of the problem. According to [2], the procedures can be partitioned into the next two families: the non-linear optimization schemes and the weak scattering linearization approximation methods, such as physical optics and Born approximation. When it is possible, linearization considerably simplifies and accelerates the inversion. Yet, Born approximation, that assumes that the electric field inside the integral equation can be approximated by the incident field, is only appropriate when the material contrast is small and the object size is not too large compared to the wavelength. Consequently, to overcome the limitations imposed by first-order methods, many deterministic nonlinear methods have been developed to solve different types of inverse scattering problems efficiently. First of all, let us mention the Contrast Source Inversion (CSI) method [11, 12]: it minimizes a cost functional

by iteratively updating unknown contrast sources. Many improvements have been proposed: the introduction of a total variation regularization term to preserve edges [12] (with the difficulty to determine the regularization parameters), the extended Born approximation which seems to be more accurate for a larger range of material contrasts and object sizes [11], the distorted Born approach [13] where the Green's function is updated at every iteration step, the diagonalized contrast source approach [14], etc. Another approach is the modified gradient method [15], based on gradient scheme and a successive over-relaxation approach to solve the direct problem. Finally, let us emphasize the efficient linear sampling method in 3D shape reconstruction of obstacles due to local inhomogeneities [16, 2].

In this paper, a global statistical approach is developed to examine the "free space RCS" inverse scattering and solve the large scale ill-posed inverse problem. In some way, the approach can be considered to be part of the two aforementioned procedure families. It involves an approximation method. Intensive Maxwell solver computations, distributed on high performance computing (HPC) machines, results in a surrogate likelihood model. It is the starting point of a complete statistical dynamic model framework that leads to an efficient inference scheme, close to optimization. It stems from statistical signal processing and advanced Monte Carlo sampling (e.g. Markov Chain Monte Carlo). Bayesian inference is performed by a sequential Monte Carlo (SMC) stochastic algorithm (see for instance [17] or [18], respectively applied to tomographic imaging and magnetoencephalography). These algorithms are called "interacting particles" [19] or particle filtering in adaptive filtering and sequential estimation. They are used to provide, in addition to microwave properties estimates of materials, the very significant information of the associated uncertainties. From the seminal work of Geman and Geman [20], stochastic methods have been commonly used in inverse scattering and, more generally, in image inverse problems: simulated annealing for image reconstruction [21], expectation-maximization algorithm for radar imaging [22], etc. In microwave imaging, [23] points out genetic algorithms and stochastic heuristics, such as differential evolution methods, memetic algorithms, particle swarm optimizations, ant colonies, etc. In short, many attempts have been made in electromagnetism to apply stochastic methods to tricky inverse problems or non-convex optimization (such as multilayered radar absorbing coatings [24, 25]). Though powerful, stochastic inverse methods often come up against high-dimensional curse. In the approach, advantage is taken of the problem structure to achieve a Rao-Blackwellisation strategy [26, 19] of Monte Carlo variance reduction and to design a powerful stochastic inversion method that overcomes the high-dimensional obstacle.

Compared to the above-mentioned deterministic methods, such as the contrast source inversion method, our statistical inverse approach has several advantages in in dealing with the "free space RCS" inverse scattering problem. As it is stressed by [3], it offers a powerful and convenient framework to solve ill-posed inverse problem

via a Bayesian regularization resulting from the probabilistic modeling of the prior knowledge and measurement uncertainties, i.e. the information. It provides a frequency band reconstruction algorithm that jointly considers the measurements at different frequencies. And above all, the approach provides confidence intervals or uncertainties on the estimated microwave properties, with a set of possible properties in coherence with the measurements. It can be especially useful in the context of quality control or material measurement. Besides, unlike Born or Rytov approximations, notice that the approach approximations rely on an exact Maxwell solver that is able to catch all the complex object-wave interactions.

This paper is organized as follows. In section 2, free-space RCS material measurements is introduced and the inverse scattering problem is developed. Next, section 3 describes the probabilistic modeling , from the surrogate likelihood model to the overall statistical dynamic modeling framework and, at its core, a hidden Markov model (HMM). The inversion Rao-Blackwellised stochastic algorithm is developed in section 4. It is evaluated in section 5 where its statistical performance is assessed.

2. The inverse scattering problem

2.1. Electromagnetic scattering measurement

EM scattering measurements have been achieved ever since radar invention [4]. Briefly speaking, EM scattering is the standard phenomenon that occurs when an object is exposed to an EM wave and disperses incident energy in all directions (scattering is this spatial distribution of energy). Some energy is scattered back to the source of the wave. It constitutes the radar echo of the object, the intensity of which results from the radar cross section (RCS) of the object. More precisely, RCS is defined by:

$$\sigma_s = \lim_{R \to +\infty} 4\pi R^2 \frac{|\mathbf{E}_{\text{scat}}|^2}{|\mathbf{E}_{\text{inc}}|^2} \tag{1}$$

It quantifies the scattering power of an object, i.e. the ratio between the scattered power density $\mathbf{E}_{\text{scat}}$ at the receiver and the power density of the incident wave at the target (with R the radar-object range). It depends on the wave polarization and frequency. The $4\pi R^2$ term takes into account the radiated spherical wave. Implicitly, (1) requires that the incident wave is planar ($R \to +\infty$). Practically, it is possible to measure the RCS at limited ranges with a sufficient accuracy. It is usually achieved in indoor RCS test chambers, also called anechoic chambers. There, interferences can be limited by microwave absorbing materials (see figure 1).

In the article, we consider that an object or mock-up is illuminated by a radar, i.e. a single antenna or a more complex device (such as the antenna array of figure 1) that fulfills to a certain extent directivity and far-field conditions [27]. Herein the radar system is monostatic, which means that the transmitter and receiver are collocated.

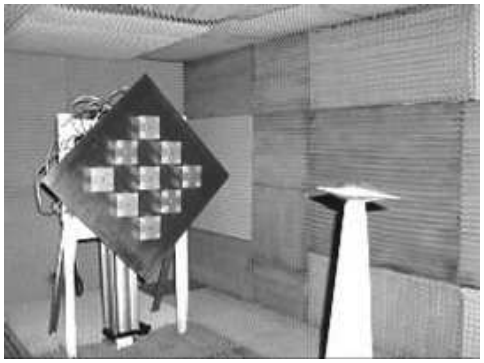

Figure 1. RCS measurement inside an anechoic chamber

Its common principle is described in figure 2. Considering that the radar illuminates the object at a given incidence with a quasi-planar monochromatic continuous wave (CW) of frequency f (incident electric field $\mathbf{E}_{\text{inc}}$), the object backscatters a CW to the radar (scattered electric field $\mathbf{E}_{\text{scat}}$) at the same frequency. With an appropriate instrumentation system (radar, network analyzers, etc.) and a calibration process, it is possible to measure the complex scattering coefficient, which can be roughly defined by: $\mathcal{S} = \frac{\mathbf{E}_{\text{scat}}}{\mathbf{E}_{\text{inc}}}$. It sums up the EM scattering, indicating the wave change in amplitude and phase. $\mathcal{S}$ is closely linked to the RCS, with: $\sigma_s = |\mathcal{S}|^2$. It is important to notice that the scattering coefficient quantifies a global characteristic of the whole object-EM wave interaction in specific conditions (incidence, frequency, etc.). It is possible to measure the scattering coefficient for different transmitted and received polarizations.

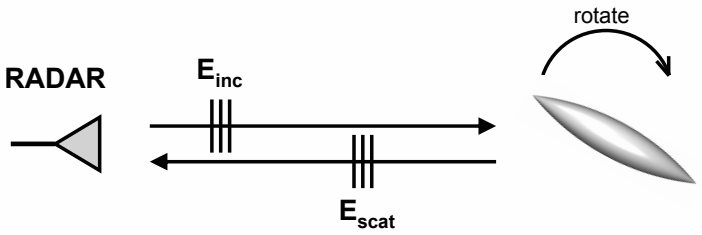

Figure 2. Monostatic scattering measurement principle

Let us assume the following conventional RCS acquisition mode, widely used in Inverse Synthetic Aperture Radar (ISAR) imaging. It consists in measuring various complex scattering coefficients $\mathcal{S}$:

- at different wave frequencies: $f \in \{f_1, f_2, \cdots, f_{K_f}\}$, for K_f successive discrete frequencies. Basically, it consists in a series of transmitted narrow-band pulses, commonly known as SFCW (Stepped Frequency Continuous Wave) burst [28].
- at different incidence angles: $\theta \in \{\theta_1, \theta_2, \cdots, \theta_{K_\theta}\}$, for K_θ different incidence angles (object rotation with a motorized rotating support).

- at different (transmitted and received) linear polarizations: pol $\in \{HH, VV\}$, meaning respectively, horizontally and vertically polarized both at microwave emission and reception.

Let's call $\mathcal{M}$ the complete measurement, set of $2 \cdot K_f \cdot K_\theta$ elementary complex scattering coefficients: $\mathcal{M} = \{\mathcal{S}^{f,\theta,\text{pol}}\}$, for $f \in \{f_1, \cdots, f_{K_f}\}$, $\theta \in \{\theta_1, \cdots, \theta_{K_\theta}\}$ and pol $\in \{HH, VV\}$.

2.2. Nondestructive testing

In this article, we are interested in an industrial control issue, that can be assimilated to nondestructive testing (NDT). Unlike usual EM material characterization techniques [4], the point is to determine or check radioelectric properties (i.e. relative dielectric permittivity and magnetic permeability) of materials that are assembled and placed on the full-scaled object or system. Is it possible from the above complete measurement $\mathcal{M}$? Is it possible to extract some local information on the material properties along the object from the global scattering measurement information?

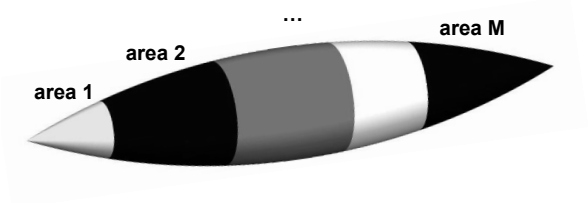

Figure 3. The object coated by N_a material areas

In order to circumscribe the investigation, the article is restricted to a metallic axisymmetric object, which is coated by N_a material areas, each area corresponding to a rather homogeneous material, with its associated isotropic radioelectric properties weakly varying within the area. It is illustrated in figure 3, with an ogival shape taken from the RCS benchmark [29]. Consequently, the aim is to determine, from the global scattering measurement $\mathcal{M}$, the unknown isotropic local EM properties $(\epsilon_1, \mu_1), (\epsilon_2, \mu_2), \cdots, (\epsilon_N, \mu_N)$ along the object, where N is the number of different elementary zones (cf. Figure 4).

2.3. An inverse problem for Maxwell's equations

Naturally, there is no direct model that is able to compute the radioelectric properties from global scattering information. On the contrary, the forward scattering model based on the resolution of Maxwell's equations can determine the scattering coefficients, given the EM properties, the object geometry and acquisition conditions (i.e. wave frequency, incidence, etc.). It lies in the resolution of Maxwell's equations, partial derivative equations that represent the electromagnetic scattering problem of an inhomogeneous

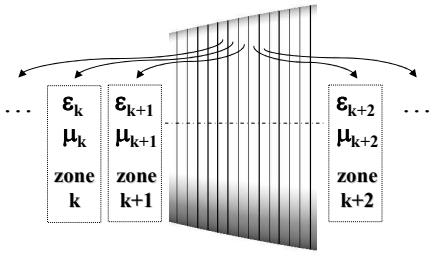

Figure 4. Elementary mesh zones

obstacle. It is performed by an efficient parallelized harmonic Maxwell solver, an exact method that combines a volume finite element method and integral equation technique, taking benefit from the axisymmetrical geometry of the shape [30]. Discretization is known to lead to problems of very large sizes, especially when the frequency is high. Furthermore, as it is shown further on, the solver is to be run many times for the inversion purpose. Hence, it necessitates high performance computing (HPC): a massive supercomputing system, with nearly 20,000 processors and a performance higher than 1 petaflops (million billion operations per second).

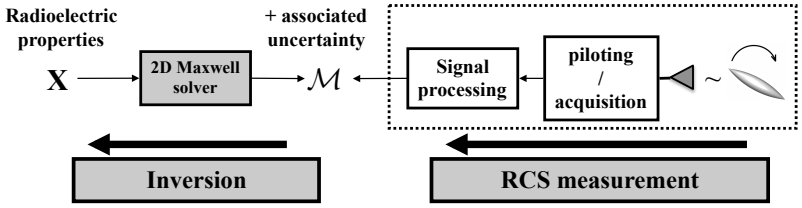

Figure 5. The inverse scattering problem

Figure 5 sums up the entire inverse scattering problem. On the one hand, the RCS measurement process, that includes acquisition, signal processing, calibration, etc., provides the complex scattering measurement $\mathcal{M}$, with uncertainties. On the other hand, it would be useful to "row upstream" the Maxwell solver, in order to determine the unknown radioelectric properties, denoted by $\mathbf{x}$. Yet, even with recourse to HPC, there is no direct way to solve what turns out to be a high dimensional ill-posed inverse problem, like imaging inverse problems [10]. Next, we propose a global statistical inference approach, which is able to take prior information into account and achieve the required inversion. Like Tikhonov regularization, it tends to eliminate artificial oscillations due to the ill-posedness of the problem.

3. The statistical problem formulation

The global statistical approach is introduced gradually, from its formulation at a given frequency f_k to the whole stochastic model at the various frequencies $f_1, f_2, \cdots, f_{K_f}$.

3.1. The problem statement at a single frequency f_k

Consider a given frequency f_k of the SFCW burst. Let us define the two main modeling components at f_k: the system state $\mathbf{x}_k$, the observation $\mathbf{y}_k$ and the probabilistic link between them, i.e. the likelihood model $p(\mathbf{y}_k|\mathbf{x}_k)$. To lighten the notations, they are denoted respectively $\mathbf{x}$, $\mathbf{y}$ and $p(\mathbf{y}|\mathbf{x})$ in this section.

3.1.1. System state $\mathbf{x} = \left[\underline{\boldsymbol{\epsilon}}' \quad \underline{\boldsymbol{\epsilon}}'' \quad \underline{\boldsymbol{\mu}}' \quad \underline{\boldsymbol{\mu}}''\right]^T$ includes the relative permittivity and permeability components of the N elementary zones, where $'$ and $''$ denote respectively the real and imaginary parts ‡ (at frequency f_k). The four components can be developed as: $\underline{\boldsymbol{\epsilon}}' = [\epsilon'_1 \cdots \epsilon'_N]^T$, $\underline{\boldsymbol{\epsilon}}'' = [\epsilon''_1 \cdots \epsilon''_N]^T$, $\underline{\boldsymbol{\mu}}' = [\mu'_1 \cdots \mu'_N]^T$ and $\underline{\boldsymbol{\mu}}'' = [\mu''_1 \cdots \mu''_N]^T$. $\mathbf{x}$ is in a system space of dimension $4N$; it includes all the unknown parameters that are to be estimated.

3.1.2. Observation $\mathbf{y} = [\Re(\mathcal{S}_{\mathbf{HH}}) \quad \Im(\mathcal{S}_{\mathbf{HH}}) \quad \Re(\mathcal{S}_{\mathbf{VV}}) \quad \Im(\mathcal{S}_{\mathbf{VV}})]^T$ contains the real ($\Re(\cdot)$) and imaginary ($\Im(\cdot)$) parts of the complex scattering coefficients $\mathcal{S}_{\mathbf{HH}}$ and $\mathcal{S}_{\mathbf{VV}}$ measured at the K_θ angles $\theta_1, \cdots, \theta_{K_\theta}$ (at frequency f_k). The two complex terms $\mathcal{S}_{\mathbf{HH}}$ and $\mathcal{S}_{\mathbf{VV}}$ can be detailed: $\mathcal{S}_{\mathbf{HH}} = \left[\mathcal{S}^{f_k,\theta_1,\mathrm{HH}} \quad \mathcal{S}^{f_k,\theta_2,\mathrm{HH}} \quad \cdots \quad \mathcal{S}^{f_k,\theta_{K_\theta},\mathrm{HH}}\right]^T$ and $\mathcal{S}_{\mathbf{VV}} = \left[\mathcal{S}^{f_k,\theta_1,\mathrm{VV}} \quad \mathcal{S}^{f_k,\theta_2,\mathrm{VV}} \quad \cdots \quad \mathcal{S}^{f_k,\theta_{K_\theta},\mathrm{VV}}\right]^T$. The observation space dimension is $4 \cdot K_\theta$.

3.1.3. Likelihood model $p(\mathbf{y}|\mathbf{x})$ describes the probabilistic relation between the system state $\mathbf{x}$ and the observation $\mathbf{y}$ (at frequency f_k). In other words, it provides the probability distribution of the observation $\mathbf{y}$, given a known system state $\mathbf{x}$. It is a key element of the knowledge that needs to be taken into account. Our inference goal is going to inverse this statistical relation. The likelihood model can be expressed as a multidimensional Gaussian of mean $\mathcal{F}_{\mathrm{Maxwell}}(\mathbf{x})$ and covariance matrix $\mathbf{R}_{\mathbf{m}}$:

$$\mathbf{y}|\mathbf{x} \sim \mathcal{N}(\mathcal{F}_{\mathrm{Maxwell}}(\mathbf{x}), \mathbf{R}_{\mathbf{m}}) \tag{2}$$

where $\mathcal{F}_{\mathrm{Maxwell}}$ is the direct model, from the state space to the observation space, that relies on the aforementioned Maxwell solver. Taking into account measurement uncertainties, the likelihood model results from the following considerations.

- The Maxwell solver, based on a direct method, is exact, i.e. extremely precise. $\mathcal{F}_{\mathrm{Maxwell}}$ is assumed to compute the "perfect observations", meaning without measurement noise, bias, etc. Implicitly, it is assumed that the shape object is perfectly known and that, conditionally to radioelectric properties, uncertainty only comes from measurement.
- From previous measurement uncertainty analysis (see metrology guideline [31]), it has been shown that the measurement uncertainty can be reasonably modeled by an additive Gaussian noise ($\mathbf{y} = \mathcal{F}_{\mathrm{Maxwell}}(\mathbf{x}) + \mathbf{v}_{\mathbf{m}}$, $\mathbf{v}_{\mathbf{m}} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{\mathbf{m}})$) with the quantified covariance matrix $\mathbf{R}_{\mathbf{m}}$.

‡ In other words, $\boldsymbol{\epsilon} = \boldsymbol{\epsilon}' + j\boldsymbol{\epsilon}''$ and $\boldsymbol{\mu} = \boldsymbol{\mu}' + j\boldsymbol{\mu}''$ (for time dependence convention $e^{j\omega t}$).

Consequently, the likelihood model can be expressed as (with $\nu = 4 \cdot K_\theta$):

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{(2\pi)^{\frac{\nu}{2}}\sqrt{\det \mathbf{R}}} e^{-\frac{1}{2}(\mathbf{y}-\mathcal{F}_{\text{Maxwell}}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}-\mathcal{F}_{\text{Maxwell}}(\mathbf{x}))} \tag{3}$$

At first sight, just considering a single frequency f_k, numerous evaluations of $p(\mathbf{y}|\mathbf{x})$, i.e. of the Maxwell solver $\mathcal{F}_{\text{Maxwell}}(\mathbf{x})$, are required in order to solve the inverse problem; they can be far too time-consuming, even with high performance computing. To avoid heavy $\mathcal{F}_{\text{Maxwell}}$ computations, a statistical learning approach has been achieved. Its basic principle is to build a surrogate model, i.e. an approximation of $\mathcal{F}_{\text{Maxwell}}$ that is acceptable in the limited domain of interest. In a way, it is related to weak scattering linearization approximation methods of [2] in inverse scattering, and among them, the former mentioned and widely used Born approximation [1, 2]. Here, the statistical linearization is not performed from truncation of physical interactions, but from full Maxwell solution computations that take multiple interactions, creeping waves, etc. into account. The system, i.e. the high dimension state space of $\mathbf{x}$ and the associated system response $\mathcal{F}_{\text{Maxwell}}(\mathbf{x})$, is explored by random sampling, according to a prior knowledge about the expected radioelectric properties (prior distribution $p(\mathbf{x})$). The computations are massively distributed on HPC machines, each computation involving the parallelized Maxwell solver. The computation number depends mainly on the state space dimension. The Monte Carlo simulation process leads to the following training set:

$$\mathcal{B} = \{(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), (\mathbf{x}^{(2)}, \mathbf{y}^{(2)}), \cdots, (\mathbf{x}^{(N_S)}, \mathbf{y}^{(N_S)})\} \tag{4}$$

where $\mathbf{x}^{(k)} \sim p(\mathbf{x})$ ($\sim$ for realization of) and $\mathbf{y}^{(k)} = \mathcal{F}_{\text{Maxwell}}(\mathbf{x}^{(k)})$ (for $k = 1 \cdots N_S$), N_S being the number of samples. Multidimensional linear regression provides a straightforward and efficient way to build a linear model $\mathbf{y} = f(\mathbf{x}) + \mathbf{v_l}$ ($\mathbf{v_l}$ is an linearization error term) with:

$$f(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x} + \mathbf{y}^0 \text{ or } f(\mathbf{x}) = \mathbf{A}^\star \cdot [1 \quad \mathbf{x}] \text{ , } \mathbf{A}^\star = [\mathbf{y}^0 \quad \mathbf{A}] \tag{5}$$

$\mathbf{A}^\star$ is the least square (LS) estimates of the matrix of parameters that minimizes the errors to linearity (δ_l), is given by the solution to the normal equations:

$$\mathbf{A}^\star = (\mathcal{X}_\mathcal{B}^T \cdot \mathcal{X}_\mathcal{B})^{-1} \mathcal{X}_\mathcal{B}^T \mathcal{Y}_\mathcal{B} \text{ with } \mathcal{X}_\mathcal{B} = \begin{bmatrix} 1 & \mathbf{x}^{(1)} \\ 1 & \mathbf{x}^{(2)} \\ \cdots & \cdots \\ 1 & \mathbf{x}^{(N_S)} \end{bmatrix}, \mathcal{Y}_\mathcal{B} = \begin{bmatrix} \mathbf{y}^{(1)} \\ \mathbf{y}^{(2)} \\ \cdots \\ \mathbf{y}^{(N_S)} \end{bmatrix} \tag{6}$$

where $\mathcal{X}_\mathcal{B}$ is the $(4N \times N_S)$ input matrix and $\mathcal{Y}_\mathcal{B}$ the $(\ 4K_\theta \times N_S)$ response matrix, from the training set $\mathcal{B}$. For numerical stability, a QR decomposition of $\mathcal{X}_\mathcal{B}$ is introduced. By residual analysis, it is then possible to assess the linear model fitness, i.e. to determine the discrepancy between the data and the model in the domain of interest. In principle, the covariance matrix ($\mathbf{R}_l$) evaluation of the linearization error $\mathbf{v_l}$ may require a supplementary data set or cross-validation methods. Remark that additional statistical analysis can be achieved to extract reduced models, removing useless explanatory variables, i. e. permittivity or permeability components of zone subsets. That depends on the wave interaction, especially on the frequency band.

Back to the likelihood model (2), it leads to an overall error term $\mathbf{v} = \mathbf{v_l} + \mathbf{v_m}$ of covariance matrix $\mathbf{R}$ (in our context, the linearization error turns out to be negligible compared to the RCS measurement uncertainties: $\mathbf{R} + \mathbf{R}_l \simeq \mathbf{R}$) and to the following linear Gaussian (LG) likelihood model (reintroducing the subscript k for frequency f_k):

$$\mathbf{y}_k|\mathbf{x}_k \sim \mathcal{N}(\mathbf{A}_k \cdot \mathbf{x}_k + \mathbf{y}_k^0, \mathbf{R}_k) \text{ or } \mathbf{y}_k = \left[\mathbf{A}_k \cdot \mathbf{x}_k + \mathbf{y}_k^0\right] + \mathbf{v_k} \tag{7}$$

with $\mathbf{A}_k$ and $\mathbf{y}_k^0$ learned from the training set $\mathcal{B}_k$. It is illustrated in figure 6 for the ogival shape of figure 3 ($N = 137$, $f = 1.5$ GHz, $\theta = 0° : 1° : 180°$ - exploration: 1000 HPC $\mathcal{F}_{\text{Maxwell}}$ simulations). Inside each bloc, the pattern can be explained by the coherent contribution of each elementary zone.

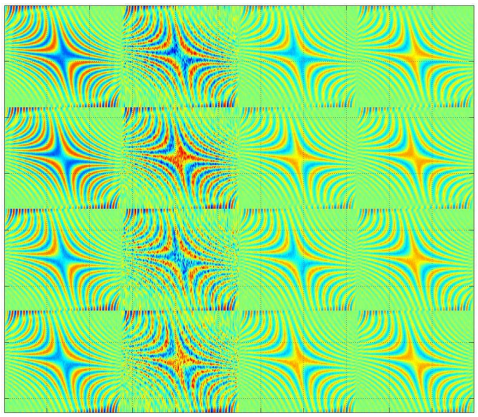

Figure 6. Matrix $\mathbf{A}_k$ illustration

3.1.4. Bayesian approach If such an inversion at a single frequency f_k could be solved by classical regularization methods [10], Bayesian estimation could offer a convenient and powerful framework. Let us probabilize the unknown state vector $\mathbf{x}_k$ and consider a prior probability distribution $p(\mathbf{x}_k)$. It is possible to model the priori knowledge with a Gaussian distribution: $\mathbf{x}_k \sim \mathcal{N}(\mathbf{m}_k, \mathbf{P}_k)$.

The mean $\mathbf{m}_k$ (dimension N) defines the reference radioelectric properties for the N_a areas that divide the object (cf. figure 3).

$$\mathbf{m}_k = \left[\mathbf{m}_k^{\epsilon'} \quad \mathbf{m}_k^{\epsilon''} \quad \mathbf{m}_k^{\mu'} \quad \mathbf{m}_k^{\mu''}\right]^T \tag{8}$$

where $\mathbf{m}_k^{\epsilon'} = [\underbrace{\epsilon'_k(1)\cdots\epsilon'_k(1)}_{\text{area 1}} \quad \underbrace{\epsilon'_k(2)\cdots\epsilon'_k(2)}_{\text{area 2}} \quad \cdots \quad \underbrace{\epsilon'_k(N_a)\cdots\epsilon'_k(N_a)}_{\text{area } N_a}]^T$, $\epsilon'_k(i)$ being the reference real permittivity of area i ($i = 1 \cdots N_a$). Similar construction for $\mathbf{m}_k^{\epsilon''}$, $\mathbf{m}_k^{\mu'}$ and $\mathbf{m}_k^{\mu''}$.

The covariance $\mathbf{P}_k$ (dimension $N \times N$) quantifies the prior uncertainty around $\mathbf{m}_k$. $\mathbf{P}_k$ is block-diagonal: $\mathbf{P}_k = \texttt{diag}(\mathbf{P}_k^{\epsilon'}, \mathbf{P}_k^{\epsilon''}, \mathbf{P}_k^{\mu'}, \mathbf{P}_k^{\mu''})$. It means that the properties (ϵ', ϵ'', μ', μ'') are assumed to be uncorrelated. Each property block is block-structured itself. For instance, $\mathbf{P}_k^{\epsilon'} = \texttt{diag}(\mathbf{P}_k^{\epsilon'}(1), (\mathbf{P}_k^{\epsilon'}(2), \cdots, (\mathbf{P}_k^{\epsilon'}(N_a))$, expressing

the assumed property independence between areas. Focusing on one block $\mathbf{P}_k^{\epsilon'}(i)$, a squared exponential covariance expresses the spatial homogeneity (of the given property) between components, i.e. elementary zones of the object that belong to the same i^{th} material area :

$$\mathbf{P}_k^{\epsilon'}(i) = \left[\sigma_k^{\epsilon'}(i)\right]^2 \times \begin{bmatrix} 1 & \rho_S & \rho_S^2 & \cdots & \rho_S^{n-1} \\ \rho_S & 1 & \rho_S & & \vdots \\ \rho_S^2 & \rho_S & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \rho_S \\ \rho_S^{n-1} & \cdots & \cdots & \rho_S & 1 \end{bmatrix} \tag{9}$$

where $\left[\sigma_k^{\epsilon'}(i)\right]^2$ is the spatial variance of i^{th} area and $\rho_S \in [0,1]$ the normalized spatial correlation parameter (e.g. $\rho_S = 0.95$). With this Markovian property, commonly used in Gaussian field modeling, correlation decreases geometrically with the distance between components. $\mathbf{P}_k^{\epsilon''}$, $\mathbf{P}_k^{\mu'}$ and $\mathbf{P}_k^{\mu''}$ are similarly constructed.

With linear Gaussian structure, i.e. Gaussian prior and linear Gaussian likelihood, Bayesian inversion can be performed straightforwardly, with closed-form solutions [3]. In our problem, it is a part of the more complex global problem that encompasses the frequency variation.

3.2. The global problem statement

Radioelectric properties are known to vary according to the wave frequency [4]. They can be quite different from the lower band frequency f_1 to the higher band one f_K. The basic idea is to maintain the former statistical modeling at each frequency f_k while introducing additional a priori information about the dynamic in frequency, i.e. how quickly a property can vary with frequency, what the correlation is between two different frequencies, etc. This regularity information can be quite different from one EM property (ϵ', ϵ'', μ', μ'') to another, as well as from one material to another,

3.3. Generalized Auto-Regressive random process

The statistical modeling extension consists in modeling the whole sequence $(\mathbf{x}_k, k \in \{1,\ldots,K_f\})$ by a generalized autoregressive (AR) random process:

$$\begin{aligned} \mathbf{x}_1 &\sim \mathcal{N}(\mathbf{m}_1, \mathbf{P}_1) \\ \mathbf{x}_{k+1} &= \mathbf{m}_{k+1} + \mathbf{D}_\rho \cdot \mathbf{H}_{k+1} \cdot \mathbf{H}_k^{-1} \cdot (\mathbf{x}_k - \mathbf{m}_k) + \sqrt{\mathbf{I_d} - \mathbf{D}_\rho^2} \cdot \mathbf{H}_{k+1} \cdot \mathbf{V}_k \end{aligned} \tag{10}$$

where $\mathbf{H}_k$ is the square root of the covariance matrix $\mathbf{P}_k$ §. $(\mathbf{V}_k, k \in \{1,\ldots,K\})$ are i.i.d. (independent, identically distributed) $\mathcal{N}(0, \mathbf{I_d})$ and $\mathbf{D}_\rho$ is a positive diagonal matrix commuting with $\mathbf{H}_k$. The dynamic model expresses the linear Gaussian correlation structure. It can be checked that the marginal distribution of $\mathbf{x}_k$ is still

§ unique symmetric definite positive matrix such as: $\mathbf{H}_k \cdot \mathbf{H}_k^T = \mathbf{P}_k$.

$\mathcal{N}(\mathbf{m}_k, \mathbf{P}_k)$. More generally, it can be shown that the distribution of concatenated vector $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_{K_f})$ is Gaussian with mean $\mathbf{m} = (\mathbf{m}_1, \ldots, \mathbf{m}_{K_f})$ and covariance matrix:

$$\mathbf{P} = \mathcal{H} \cdot \begin{bmatrix} \mathbf{I_d} & \mathbf{D}_\rho & \mathbf{D}_\rho^2 & \cdots & \mathbf{D}_\rho^{K_f-1} \\ \mathbf{D}_\rho & \mathbf{I_d} & \mathbf{D}_\rho & & \vdots \\ \mathbf{D}_\rho^2 & \mathbf{D}_\rho & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \mathbf{D}_\rho \\ \mathbf{D}_\rho^{K_f-1} & \cdots & \cdots & \mathbf{D}_\rho & \mathbf{I_d} \end{bmatrix} \cdot \mathcal{H}^T \tag{11}$$

where $\mathcal{H}$ is the block diagonal matrix $\mathcal{H} = \mathtt{diag}(\mathbf{H}_1, \ldots, \mathbf{H}_{K_f})$. Basically, every joint distribution $(\mathbf{x}_i, \mathbf{x}_j)$ is expressed .

The matrix $\mathbf{D}_\rho$ takes the frequential correlations of the EM properties $\mathbf{x}_1 \cdots \mathbf{x}_{K_f}$ into account; it refers to a hyper-parameter ρ. According to the frequency correlation prior knowledge, the following alternatives can be considered:

(i) The frequency correlation doesn't depend on the material and the EM property (ϵ', ϵ'', μ' or μ''): ρ is scalar ($\in [0,1]$) and $\mathbf{D}_\rho = \rho.\mathbf{I_d}$.

(ii) It depends on the material: ρ is N_a-dimensional ($\in [0,1]^{N_a}$), and $\mathbf{D}_\rho$ is the block-diagonal matrix made up of N_a terms $\rho_i.\mathbf{I_d}$.

(iii) It depends on both: ρ is $4.N_a$-dimensional and $\mathbf{D}_\rho$ is the block-diagonal matrix made up of $4.N_a$ terms $\rho_i.\mathbf{I_d}$.

AR models are frequently used to express dynamically Gaussian field modeling. Starting from a multidimensional Gaussian distribution, we design an autoregressive model that complies with the marginal distributions at each frequency, with the spatial Markovian structure, and integrates the frequential correlations. Let's emphasize how the complete stochastic prior AR modeling is really adapted to our problem. Indeed, it does not need too much information, roughly speaking it is not too constrained. Concerning a material area, it requires only to give information about the microwave properties, about the supposed evolution of properties according to the frequency and about the spatial homogeneity. Basically, it is a probabilistic way to fix the regularization. It is made by means of very understandable and common terms, such as mean, variance and correlations. Notice that alternative approaches from spatial statistics could surely be chosen to deal with the Gaussian field. The chosen dynamic modelling provides an efficient sequential way to take it into account. Furthermore, as it is done for spatiotemporal modelling [32], it could integrate decompositions on basis functions, in order to reduce the problem dimension.

3.4. A conditionally hidden dynamic Markov process

The generalized AR random processes include the linear Gaussian models at the various frequencies f_k ($k = 1 \cdots K_f$). It provides a spatial and frequential correlation structure.

Assuming that the material areas are known to be quite homogeneous, the spatial correlation parameter can be fixed (typically $\rho_S = 0.95$). Quite the reverse, frequency correlations can not be really known; they are to be determined by the inversion process. Back to Bayesian statistics, it is chosen to probabilize the unknown hyper-parameter ρ. Finally, the combination of the AR dynamic model (11) with the likelihood model (7) end in the following state-space model, observed at "times" f_k $(k = 1, \cdots, K)$:

$$\mathbf{x}_{k+1} = \mathbf{M}_k^{\rho} \cdot \mathbf{x}_k + \mathbf{w}_k \qquad \mathbf{y}_k = \left[\mathbf{A}_k \cdot \mathbf{x}_k + \mathbf{y}_k^0\right] + \mathbf{v}_k \tag{12}$$

assuming the initial state $\mathbf{x}_1 \sim \mathcal{N}(\mathbf{m_1}, \mathbf{P_1})$. $\mathbf{M}_k^{\rho}$ is a transition matrix and $\mathbf{w}_k$ a Gaussian model noise ($\mathbb{E}(\mathbf{w}_k) \neq 0$). Both directly arise from (11); they are not detailed here for clearness.

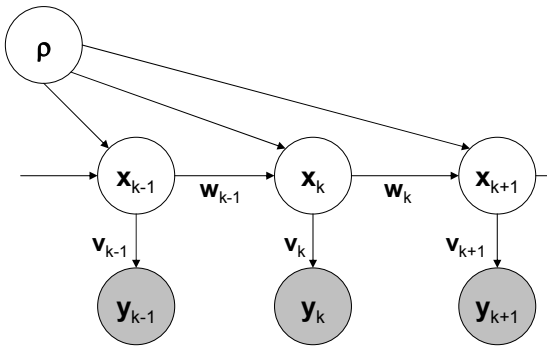

Figure 7. A graphical representation

Again, let us emphasize that the dynamic model involves that each marginal complies with $\mathbf{x}_k \sim \mathcal{N}(\mathbf{m_k}, \mathbf{P_k})$. On the other hand, it is important to remark that, conditionally to the frequential correlation parameter ρ, the model is a classic linear Gaussian hidden dynamic Markov process. A graphical representation of the entire model is given in figure 7. Given a value of ρ, the lower part describes a linear Gaussian system. The idea is to make the most of this specific structure.

4. Advanced Sequential Monte Carlo inversion

4.1. The Rao-Blackwellized Approach

As already mentioned, the unknown hyper-parameter ρ is probabilized, and so it is given a prior distribution $p(\rho)$, assumed calculable (up to a normalizing constant) and easy to sample. The posterior distribution $p(\mathbf{x}, \rho|\mathbf{y})$ can be decomposed as:

$$p(\mathbf{x}, \rho|\mathbf{y}) = p(\mathbf{x}|\rho, \mathbf{y}) \cdot p(\rho|\mathbf{y}) \tag{13}$$

Since the system is linear Gaussian conditionally to ρ, the conditional distributions $p(\mathbf{x}_k|\rho, \mathbf{y})$ can be straightforwardly computed by classic Kalman filtering. This forward algorithm can be completed by backward smoothing, in this off-line context; the overall

is often called "Kalman smoother". On the other hand, the term $p(\rho|\mathbf{y})$ can be decomposed as:

$$\begin{aligned} p(\rho|\mathbf{y}) &\propto p(\rho)\cdot p(\mathbf{y}|\rho) \\ &\propto p(\rho)\cdot\prod_{k=1}^{K_f}\underbrace{p\left(\mathbf{y}_k|\rho,\mathbf{y}_1,\ldots,\mathbf{y}_{k-1}\right)}_{:=J_k(\rho)}. \end{aligned} \tag{14}$$

Again, for any hyper-parameter ρ, the quantities $J_k(\rho)$ can be evaluated from the likelihood terms provided by the Kalman filter [3, 33]. Eventually, it is possible to exploit this conditional system structure, with Kalman smoothers [3, 33] that can be applied and integrated in the following interacting particle approach. In a first step, a stochastic algorithm (described in section 4.2) gives an approximation of $p(\rho|\mathbf{y})$. It estimates the frequential correlations (i.e. regularity) of the EM properties $\epsilon'(f), \epsilon''(f), \mu'(f), \mu''(f)$. In a second step, the first moments of $\mathbf{x}_k$ can be evaluated (for each frequency f_k) by the theoretical conditioning relations:

$$\mathbb{E}(\mathbf{x}_k|\mathbf{y}) = \mathbb{E}\left[\mathbb{E}(\mathbf{x}_k|\rho,\mathbf{y})|\mathbf{y}\right] \tag{15}$$

$$\mathbb{V}\mathrm{ar}(\mathbf{x}_k|\mathbf{y}) = \mathbb{E}\left[\mathbb{V}\mathrm{ar}(\mathbf{x}_k|\rho,\mathbf{y})|\mathbf{y}\right] + \mathbb{V}\mathrm{ar}\left[\mathbb{E}(\mathbf{x}_k|\rho,\mathbf{y})|\mathbf{y}\right] \tag{16}$$

Note that Kalman recursions [33] are used both in the first step for calculating the likelihood of the hyper-parameter ρ (up to a normalizing constant) and in the second step for determining the quantities $\mathbb{E}(\mathbf{x}_k|\rho,\mathbf{y})$ and $\mathbb{V}\mathrm{ar}(\mathbf{x}_k|\rho,\mathbf{y})$. This idea of mixing analytic integration (here Kalman evaluation of $p(\mathbf{x}|\rho,\mathbf{y})$) with stochastic sampling (here to approximate $p(\rho|\mathbf{y})$) is a variance reduction approach, known as Rao-Blackwellisation [26].

Let us denote by $\eta(d\rho)$ the probability measure associated with the marginal distribution $p(\rho|\mathbf{y})$, for a fixed observation vector $\mathbf{y}$. Similarly to [26], we choose to implement, for the first step, an efficient interacting particle approach, called Sequential Monte Carlo (SMC), in order to estimate η. We now give a brief but general description of these methods.

4.2. The SMC algorithm

Sequential Monte Carlo is a stochastic algorithm to sample from complex high-dimensional probability distributions. The principle (see, e.g., [19]) is to approximate a sequence of target probability distributions (η_n) by a large cloud of random samples termed particles $(\zeta_n^k)_{1\leq k\leq N_p} \in E^{N_p}$, E being called the state space. Between "times" $n-1$ and n, the particles evolve in the state space E according to two steps (see figure 8):

(i) **A selection step**: every particle ζ_{n-1}^i is given a weight ω_i defined by a selection function $G_n : E \to (0,+\infty)$ (i.e. $\omega_i = G_n(\zeta_{n-1}^i)$). By resampling (stochastic or deterministic), low-weighted particles vanish and are replaced by replicas of high-weighted ones.

(ii) **A mutation step**: each selected particle $\widehat{\zeta}^i_{n-1}$ moves, independently from the others, according to a Markov kernel $M_n : E \to E$.

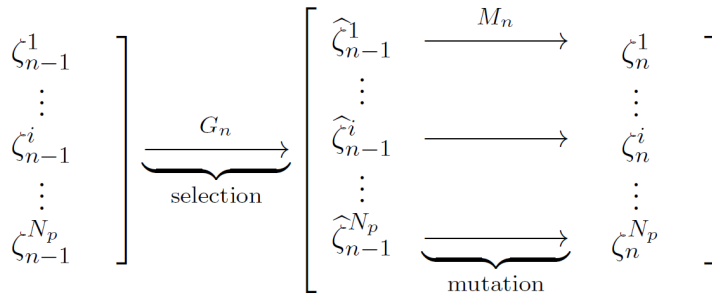

Figure 8. The SMC 2-step evolution

Evolving this way, the cloud of particles, and more precisely the occupation distribution $\eta_n^{N_p} := \frac{1}{N_p}\sum_{k=1}^{N_p} \delta_{\zeta_k^n}$ (sum of Dirac distributions), approximates for each n the theoretical distribution η_n defined recursively by the Feynman-Kac formulae. It is associated with the potentials G_n and kernels M_n (see [34] for further details). More precisely, this sequence η_n is defined by an initial probability measure η_0 and the recursion:

$$\eta_n = \Psi_{G_n}(\eta_{n-1}).M_n \tag{17}$$

where $\Psi_{G_n}(\eta_{n-1})$ is the probability measure defined by $\Psi_{G_n}(\eta_{n-1})(dx) \propto G_n(x).\eta_{n-1}(dx)$ and, for any probability measure μ, $\mu.M_n$ is the measure so that $\mu.M_n(A) = \int_E M_n(x, A)\mu(dx)$.

The SMC approach is often used for solving sequential problems, such as filtering (e.g., [35, 36, 37]). In other problems, like ours, this algorithm also turns out to be efficient to sample from a single target measure η. In this context, the central idea is to find a judicious interpolating sequence of probability measures $(\eta_n)_{0\leq k\leq n_f}$ with increasing sampling complexity, starting from some initial distribution η_0, up to the final target one $\eta_{n_f} = \eta$. Consecutive measures η_n and η_{n+1} are to be sufficiently similar to allow for efficient importance sampling and/or acceptance-rejection sampling. The sequential aspect of the approach is then an "artificial way" to solve the sampling difficulty gradually. More generally, a crucial point is that large population sizes allow to cover several modes simultaneously. This is an advantage compared to standard MCMC (Monte Carlo Markov Chain) methods that are more likely to be trapped in local modes. These sequential samplers have been used with success in several application domains, including rare events simulation [38], stochastic optimization and, more generally, Boltzmann-Gibbs measures sampling [39].

From a theoretical viewpoint, the stochastic convergence performance of SMC algorithms has been mostly analyzed using asymptotic (i.e. when number of particles N_p tends to infinity) techniques, notably through fluctuation theorems and large deviation

principles (see for instance [40, 41], and [42] for an overview). Some non-asymptotic theorems have been recently developed [38, 43, 44, 45]. They lead to some biais and variance estimations, L^p-error bounds and exponential concentration inequalities. Roughly speaking, one can show that under some stability properties, the accuracy of the method is of order $\left|\eta_n^{N_p} - \eta_n\right| = \mathcal{O}(\frac{1}{\sqrt{N_p}})$ (see for instance [43], Theorem 12).

4.3. Interpolating sequences of measures

Back to our objective of sampling from $\eta(d\rho)$, let us denote by E the state space of the variable ρ (i.e. $E = [0,1]$, $[0,1]^{N_a}$ or $[0,1]^{4N_a}$). We have to define a sequence of distributions $(\eta_n)_{0\leq k\leq n_f}$ from the initial distribution $\eta_0(d\rho) = p(\rho)d\rho$ (easy to sample) to the target one $\eta_{n_f}(d\rho) = \eta(d\rho) = p(\rho|\mathbf{y})d\rho$.

4.3.1. The guiding principle With this in mind, we first define an interesting class of Markov kernels on E: let h be a positive, bounded function on E, and let $Q(x, dy)$ be a Markov kernel on E, assumed reversible w.r.t. the Lebesgue measure on E. The Metropolis-Hastings kernel $K_{h,Q}(x, dy)$ associated with h and Q is given by the following formula:

$$K_{h,Q}(x, dy) = Q(x, dy).\min\left(1, \tfrac{h(y)}{h(x)}\right) \qquad \forall y \neq x$$

$$K_{h,Q}(x, \{x\}) = 1 - \int_{y\neq x} Q(x, dy).\min\left(1, \tfrac{h(y)}{h(x)}\right)$$

Using an acceptance/rejection method, this kernel is easy to sample as soon as one can sample $Q(x, dy)$ and calculate the ratios $h(y)/h(x)$. Here is a crucial property: if μ_h denotes the probability measure defined by $\mu_h(d\rho) \propto h(\rho)d\rho$, then it is well known (see, e.g., [46]) that $K_{h,Q}$ admits μ_h as an invariant measure:

$$\mu_h.K_{h,Q} = \mu_h \quad \left(\Longleftrightarrow \int_E K_{h,Q}(\rho, A)\mu_h(d\rho) = \mu_h(A)\ ,\ \forall A \subset E\right)$$

More generally, this property is satisfied for the iterated kernel $K_{h,Q}^m$, i.e. $\mu_h.K_{h,Q}^m = \mu_h$ (for any integer m).

Let η_n be a sequence of probability measures defined with some positive, bounded functions h_n so that: $\eta_n(d\rho) \propto h_n(\rho).d\rho$. Then, for any sequence of reversible Markov kernels Q_n and any sequence of integers m_n, η_n satisfies the Feynman-Kac formula (17) with potentials $G_n := h_n/h_{n-1}$ and Markov kernels $M_n := K_{h_n,Q_n}^{m_n}$ (K_{h_n,Q_n} iterated m_n times). Practically, the consequence is that such a sequence η_n can be approximated using a SMC algorithm as soon as one can calculate the functions h_n up to a normalizing constant. Similarly to traditional MCMC or simulated annealing methods, this algorithm is all the more robust when the iteration numbers m_n are large, since the kernels K_{h_n,Q_n} are just defined and used to stabilize the system.

4.3.2. Design of bridging measure sequences From these considerations, we propose three scheme variants of interpolating sequences of measures.

(i) The annealed scheme: the sequence η_n is defined by the positive, bounded functions

$$h_n(\rho) = p(\mathbf{y}|\rho)^{\alpha_n} \cdot p(\rho)$$

where $(\alpha_n)_{1\leq n\leq n_f}$ is a sequence of numbers increasing from 0 to 1 (arbitrarily chosen). In this situation, the potentials $G_n(\rho)$ used in the selection are equal to $p(\mathbf{y}|\rho)^{\alpha_n-\alpha_{n-1}}$. Thus, α_n is to be chosen to control the selectivity of these functions, which is important in practice. Annealing or tempering is frequently used in SMC (see [47, 19] and [48] in video tracking); it is related to simulated annealing (with inhomogeneous sequence of MCMC kernels).

(ii) The data tempered scheme: for all $n \in \{0, 1, \ldots, K_f\}$, η_n is the probability measure associated with: $h_n(\rho) = p(\rho) \cdot \prod_{k=1}^{n} \underbrace{p\left(\mathbf{y}_k|\rho, \mathbf{y}_1, \ldots, \mathbf{y}_{k-1}\right)}_{=J_k(\rho)}$. In other words, at each generation n, the selection potential $G_n(\rho)$ that is applied to the particles is the term $p\left(\mathbf{y}_n|\rho, \mathbf{y}_1, \ldots, \mathbf{y}_{n-1}\right)$, i.e. the likelihood of the n-th observation vector given the previous ones. This allows the algorithm to work "online", since it treats the observations sequentially. According to [47], it is efficient for problems that exhibit a natural order (e.g. hidden Markov models). Yet, when these potentials turn out to be too selective, the SMC algorithm turns out to perform poorly since the cloud of particles loses its diversity at each selection step. It is substituted for the next scheme that overcomes this drawback.

(iii) The hybrid scheme: similarly to the previous one, this scheme incorporates the observations one after the other, but each likelihood function $J_k(\rho)$ is handled as a product:

$$J_k(\rho) = \prod_{i=1}^{n_k} J_k(\rho)^{(\alpha_i^{(k)}-\alpha_{i-1}^{(k)})}$$

where for all $k \in \{1, \ldots, K_f\}$, $(\alpha_i^{(k)})_{1\leq i\leq n_k}$ is a sequence $0 \nearrow 1$. Then, if $n = (n_1 + \cdots + n_{r-1}) + s$, the function h_n is given by:

$$h_n(\rho) = p(\rho) \cdot \left(\prod_{k=1}^{r-1} J_k(\rho)\right) \cdot J_r(\rho)^{\alpha_s^{(r)}}$$

Note that the selection potential $G_n = J_r^{(\alpha_s^{(r)}-\alpha_{s-1}^{(r)})}$ can be arbitrarily controlled.

For each of these interpolating schemes, the functions h_n are calculable up to a normalizing constant (Kalman equations), so that the Metropolis-Hastings kernels (possibly iterated) can be used to perform the mutation steps.

4.4. The global estimation

To sum up, the joint distribution $p(\mathbf{x}, \rho|\mathbf{y})$ can be decomposed and evaluated as follows:

$$p(\mathbf{x}, \rho|\mathbf{y}) = \underbrace{p(\mathbf{x}|\rho, \mathbf{y})}_{\text{KF (+ smoothing)}} \cdot \underbrace{\overbrace{p(\rho|\mathbf{y})}^{\propto \overbrace{p(\mathbf{y}|\rho)}^{\text{KF output}} \cdot \overbrace{p(\rho)}^{\text{prior}}}}_{\text{SMC}}$$

As previously mentioned, the SMC algorithm of section 4.2 provides in the first stage an evaluation of the frequency correlations $p(\rho|\mathbf{y})$ (i.e. an approximation $\hat{\eta} = \eta_{n_f}^{N_p}$ of η). It is computed from the last generation of particles $(\rho^{(1)}, \ldots, \rho^{(N_p)}) := (\zeta_{n_f}^1, \ldots, \zeta_{n_f}^{N_p})$. In the second stage, estimators of EM properties are straightforwardly computed from conditioning relations (15) and (16) (see details in annex 6); it consists in approximations of the mean and covariance matrix of the system state $\mathbf{x}_k$. Focusing on a given frequency or on a fixed zone, the SMC method provides useful information:

- For any frequency f_k, it computes an approximation of the mean and covariance matrix of the system state $\mathbf{x}_k$. Roughly speaking, one can sample from the posterior distribution $p(\mathbf{x}_k|\mathbf{y})$ by picking a $\rho^{(i)}$ from the final cloud of particles and computing associated samples of $\mathbf{x}_k$ by a Kalman smoother conditionally to $\rho^{(i)}$ (see further illustration figure 12 page 22).
- For any fixed zone, the method provides estimators of the mean and marginal variance for every frequency, so that the results can be presented as frequential profiles, with marginal uncertainties (using the diagonal values of $\hat{\mathbf{\Sigma}}_k$) (see further illustration figure 13 page 22).

5. Applications

In this section, the inverse scattering approach is applied to EM scattering measurements of a metallic ogival-shaped object. The validation is achieved with simulated data in a wide frequency band from $f = 200$ MHz to 8 GHz. Section 5.1 describes the reference nondestructive testing scenario. Next, section 5.2 describes the inversion process and illustrates some results. A detailed performance analysis is developed in Section 5.3. Then, in Section 5.4, we briefly analyze some variants of the approach.

5.1. Nondestructive testing scenario

The metallic object We consider the metallic axisymmetric object, previously shown in figure 3; its ogival shape, derived from the RCS benchmark [29], is perfectly known. The 2 m long object is coated by $N_a = 5$ material areas, the isotropic radioelectric properties weakly varying within each area. For each material area, the true EM properties $\mathbf{x}_{\text{true}}(f)$ undergo the following model: $\mathbf{x}_{\text{true}}(f) = \mathbf{x}_{\text{ref}}(f) + c \cdot \Lambda(f)$.

At each frequency f, the true (unknown) vector $\mathbf{x}_{\text{true}}(f)$ is $4N = 76$-dimensional, where:

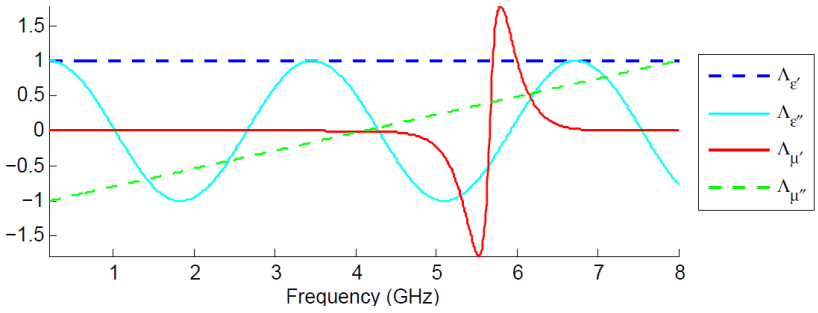

Figure 9. The functions Λ

- $\mathbf{x}_{\text{ref}}(f)$ is a reference frequency profile, depending on the area and on the radioelectric component $(\epsilon', \epsilon'', \mu', \mu'')$. Note that these $4N_a = 20$ reference profiles are chosen regular and with typical orders of magnitude (i.e. non-negative and ≤ 20).
- $\Lambda(f)$ is a perturbation function depending on the radioelectric component. Thus, the 4 functions $\Lambda_{\epsilon'}, \Lambda_{\epsilon''}, \Lambda_{\mu'}, \Lambda_{\mu''}$ define the perturbation shapes . As shown in figure 9, they are chosen more or less regular (in order to test the inversion capabilities).
- c is a simple scaling factor, depending on the area. To examine the perturbation amplitude influence, increasing values of c are chosen: $\{0.5, 1, 2, 4, 8\}$, related to the 5 successive areas.

(Simulated) scattering measurements According to the conventional RCS acquisition mode described in section 2.1, complex scattering coefficients are measured for both polarizations HH and VV, at $K_f = 20$ regularly spaced frequencies ($f_1 = 0.2$ GHz, $\cdots, f_{K_f} = 8$ GHz) and at $K_\theta = 23$ regularly spaced incidence angles ($\theta_1 = 0°, \cdots, \theta_{K_\theta} = 180°$).

The observation data $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_{K_f})$ is simulated from the likelihood model (2). That involves to run the parallelized harmonic Maxwell solver ($\mathcal{F}_{\text{Maxwell}}$) and to draw an additive white Gaussian noise of marginal standard deviation $\sigma_{\text{n}} = 10^{-3}$. Note that each of the 20 observation vectors $\mathbf{y}_k$ is $4 \times K_\theta = 92$-dimensional. The data is represented in figure 10. On the amplitude representations, note the high specular reflections when the ogival object is turned perpendicularly to the wave propagation direction. Concerning the signal-to-noise ratio (SNR), it is high, around 40 dB ($\sim 1\%$), for the specular reflexion angles (high RCS). At the opposite, the SNR is very low, much less than 0 dB, when the ogival-shaped object is illuminated at small incidences (low RCS). There, the signal is hidden by the noise and less informative.

5.2. Inversion process

The goal is to estimate the radioelectric properties, the $\mathbf{x}_{\text{true}}$ term function of the frequency f, from the scattering measurements. In this section, we give a few

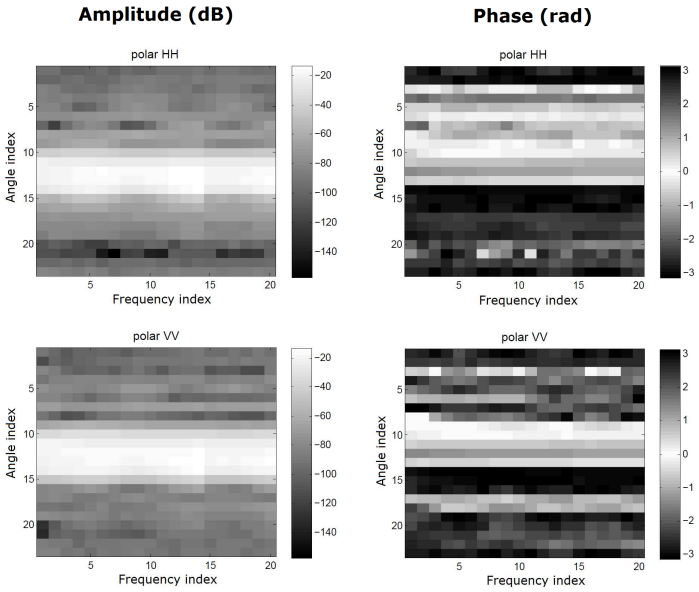

Figure 10. Observation hologram, amplitude and phase (polar HH and VV)

implementation details regarding the application context.

State space The state space dimension stems from the wave frequency number and from the discretization of the object in elementary mesh zones. In order to limit it, the cutting up of the object is here restricted to $N = 19$ elementary zones.

Prior information The prior information (see section 3) needs to be detailed in this context. Concerning the prior spatial information $p(\mathbf{x}_k)$, its means $\mathbf{m}_k$ are given, for each k, by the former reference frequency profiles $\mathbf{x}_{\text{ref}}(f_k)$. Around them, the uncertainties are given by the block-structured covariance matrices $\mathbf{P}_k$ of (9) with: $\rho_S = 0.95$ and $\sigma_k(i) = 1+0.15\times\mathbf{m}_k(i)$ for any elementary zone i. In other words, we assume a minimum standard deviation of 1 that increases proportionally to the reference amplitude value. Regarding the prior frequential information, we assume that ρ depends on both area and EM property $(\epsilon', \epsilon'', \mu', \mu'')$, so that it is 20-dimensional. As for its prior distribution $p(\rho)$, we set:

$$p(\rho) = \prod_{i=1}^{20} p(\rho_i)$$

where all the marginal prior distributions $p(\rho_i)$ are identical and presented on figure 11. Note that this distribution $p(\rho)$ can be sampled straightforwardly by sampling independently each component ρ_i using, e.g., an acceptance/rejection method.

Likelihood model The surrogate likelihood model (7) has been formerly learned: $\mathbf{A}_k$ and $\mathbf{y}_k^0$ are known (see figure 6), as well as the marginal standard deviation σ_{n} which is in conformity with the measurement noise of the above observation simulation.

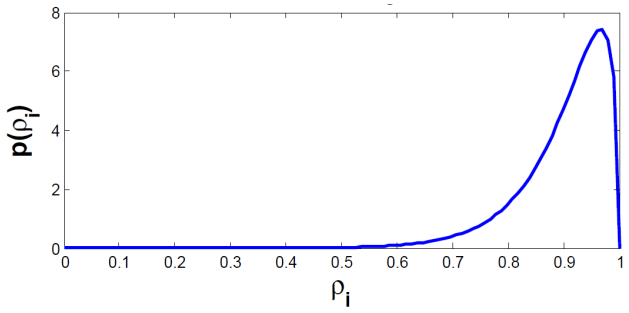

Figure 11. Marginal prior distribution $p(\rho_i)$

SMC tuning The sequence of probability measures η_n is standardly defined by the annealed scheme (see section 4.3). To ensure a stable behavior of the SMC algorithm (i.e. keep a good approximation $\eta_n^{N_p} \simeq \eta_n$ until the end), we chose the following efficient adaptive strategies (that make it possible to limit the number of particles to $N_p = 100$):

- *selection step*: as mentioned, the increment $\Delta\alpha_n = \alpha_n - \alpha_{n-1}$ controls the selectivity degree. If $\Delta\alpha_n$ is too small, every particle is given approximately the same weight, and there is no selection among them. If $\Delta\alpha_n$ is too large, the majority of the particles are killed, the cloud loses all its diversity, and the SMC algorithm performs poorly. Therefore, instead of choosing beforehand $\Delta\alpha_n$, it is defined adaptively *so that* the selection step kills around 25% of the particle population. This is a way to ensure a reasonable selection.
- *mutation step*: the mutation step is crucial since it allows the particles to explore the state space E. We use Markov kernels M_n defined as being the composition of several Metropolis-Hastings kernels $K_n^{(i)}$ whose proposition kernels $Q_n^{(i)}(x, dy)$ are uniform, centered in x, and associated with a window size $\sigma_{\text{prop},n}^{(i)}$. To be sure that the particles move in a well-sized neighborhood, (i.e. large enough to explore E and small enough to converge), the sequence $(\sigma_{\text{prop},n}^{(i)})_i$ always starts with large values and decreases geometrically. Once more, we use an adaptive criteria to stop the process.

Results In the context of this reference study, the inversion process takes about 30 minutes with a current standard processor. Note that the higher the dimension space is, the longer the inversion. In figure 12, we show the estimations of μ' for all the zones of the object, with their associated uncertainties, compared with the true values, at a fixed frequency $f_{14} = 5.6$ GHz. Note that the EM property deviation is important in our example (see figure 9). As already mentioned, it is possible to provide some samples of the posterior distribution $p(\mathbf{x}_{14}|\mathbf{y})$ to determine the uncertainty on the estimators. The EM radioelectric properties are correctly inferred all along the ogival object and its 5 material areas. The uncertainty recovers more or less the real profiles.

Figure 13 presents frequential profiles for a fixed elementary zone (the 18^{th}). All the components $(\epsilon', \epsilon'', \mu', \mu'')$ are represented. Each of them is quite accurately estimated.

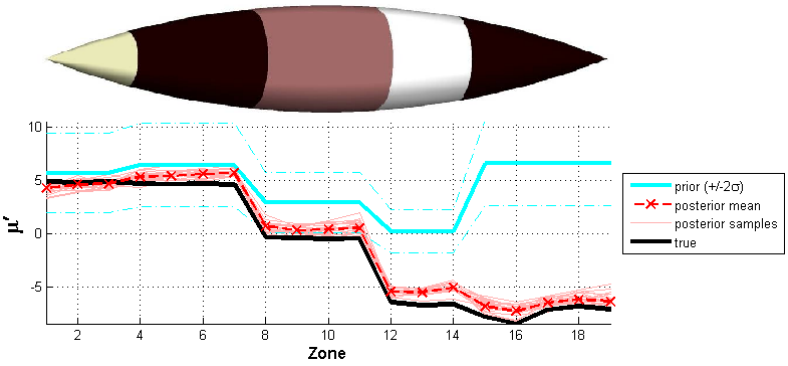

Figure 12. EM estimated properties at frequence $f = 5.6$ GHz

The results are good, even when the perturbations (i.e. the difference between the prior and real profiles) are large and irregular. This robustness is due to the adaptive estimation of ρ's components. Next it is confirmed by several thorough analysis.

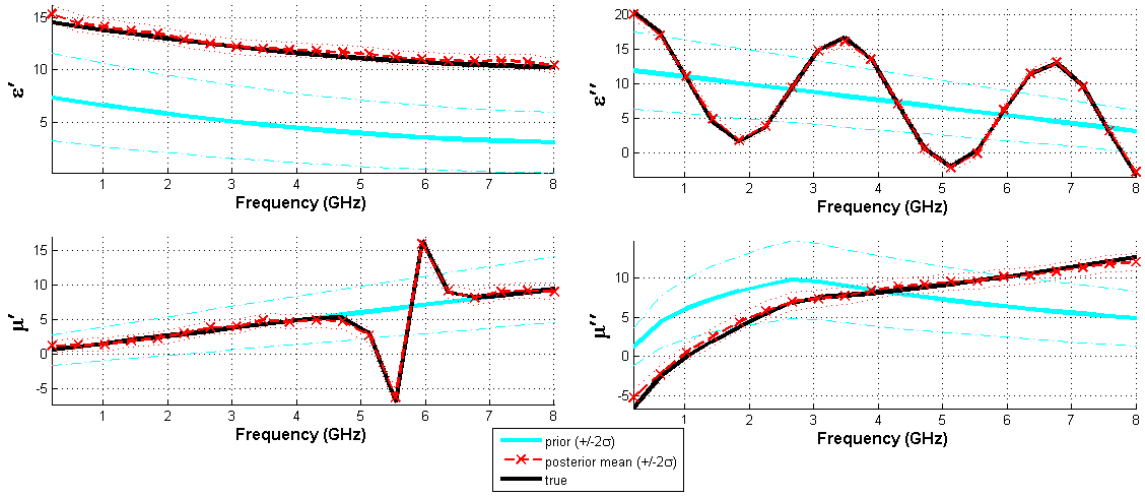

Figure 13. Estimated EM properties of the 18^{th} elementary zone

5.3. Performance analysis

To extend the results, we propose a statistical performance analysis of the inversion process. It is lead in the same context of section 5.1. As the developed interacting particle approach is partly stochastic, two different aspects must be studied. Firstly, for a single given data $\mathbf{y}$, the variance of our estimators $\hat{\mathbf{x}}_k$ and $\hat{\mathbf{\Sigma}}_k$, only due to the random feature of the method. Secondly, the average variance of our method for several data $\mathbf{y}^{(i)}$.

5.3.1. Stochastic variation For a given data $\mathbf{y}$, our method mainly provides 2 sequences of estimators. The posterior mean estimators $(\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_{K_f})$, and the posterior covariance matrices estimators $(\hat{\mathbf{\Sigma}}_1, \ldots, \hat{\mathbf{\Sigma}}_{K_f})$. As with all stochastic algorithms, one has to check that despite random, it always gives the same result, or at least that its own variance is negligible.

Let $\hat{\mathbf{x}}$ denote the concatenation of the vectors $\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_{K_f}$. Let $\hat{\sigma}$ denote the concatenation of the estimated marginal uncertainties (square root of the $\hat{\mathbf{\Sigma}}_k$'s diagonal values). Defined in this way, $\hat{\mathbf{x}}$ and $\hat{\sigma}$ can be considered as 2 matrices of size 76×20, and the 2 main estimators of our method. To quantify the stochastic variance, we simulate an observation data $\mathbf{y}$, and we perform the inversion method 30 times. At the end, we get 30 pairs of estimators $\left\{(\hat{\mathbf{x}}^{(1)}, \hat{\sigma}^{(1)}), \ldots, (\hat{\mathbf{x}}^{(30)}, \hat{\sigma}^{(30)})\right\}$. For any pair of index $(i,k) \in \{1, \ldots, 76\} \times \{1, \ldots, 20\}$, we consider the mean values of the estimators and their RMS (root mean square) values:

$$\bar{\hat{\mathbf{x}}}(i,k) := \frac{1}{30}\sum_{r=1}^{30} \hat{\mathbf{x}}^{(r)}(i,k) \quad \text{and} \quad \bar{\hat{\sigma}}(i,k) := \frac{1}{30}\sum_{r=1}^{30} \hat{\sigma}^{(r)}(i,k)$$

$$\mathrm{RMS}\left(\hat{\mathbf{x}}\right)(i,k) := \left(\frac{1}{30}\sum_{r=1}^{30}\left(\hat{\mathbf{x}}^{(r)}(i,k) - \bar{\hat{\mathbf{x}}}(i,k)\right)^2\right)^{1/2}$$
$$\mathrm{RMS}\left(\hat{\sigma}\right)(i,k) := \left(\frac{1}{30}\sum_{r=1}^{30}\left(\hat{\sigma}^{(r)}(i,k) - \bar{\hat{\sigma}}(i,k)\right)^2\right)^{1/2}$$

The numerical results, taken over all the pairs of index (i,k), are summed up in table 1. Two points can be clearly emphasized. First, the standard deviation of the $\hat{\mathbf{x}}^{(r)}$ is very small in an absolute way ($\simeq 10^{-2}$). Moreover, it is negligible compared to the estimated variance of our estimators (at least 1 decimal). Secondly, the standard deviation of the $\hat{\sigma}^{(r)}$ is even smaller ($\simeq 10^{-3}$) and negligible compared to the values of the $\hat{\sigma}^{(r)}$ themselves (at least 2 decades). Consequently, there exists a stochastic variance, but it is far negligible compared to the uncertainty inherent to the inverse problem, including measurements.

mean $\mathrm{RMS}\left(\hat{\mathbf{x}}\right)$	max $\mathrm{RMS}\left(\hat{\mathbf{x}}\right)$	mean $\frac{\mathrm{RMS}(\hat{\mathbf{x}})}{\bar{\hat{\sigma}}}$	max $\frac{\mathrm{RMS}(\hat{\mathbf{x}})}{\bar{\hat{\sigma}}}$
$4.16\ \ 10^{-3}$	$4.11\ \ 10^{-2}$	$1.08\ \ 10^{-2}$	$9.87\ \ 10^{-2}$
mean $\mathrm{RMS}(\hat{\sigma})$	max $\mathrm{RMS}(\hat{\sigma})$	mean $\frac{\mathrm{RMS}(\hat{\sigma})}{\bar{\hat{\sigma}}}$	max $\frac{\mathrm{RMS}(\hat{\sigma})}{\bar{\hat{\sigma}}}$
$1.10\ \ 10^{-3}$	$7.56\ \ 10^{-3}$	$2.99\ \ 10^{-3}$	$1.84\ \ 10^{-2}$

Table 1. RMS results of $\hat{\mathbf{x}}$ and $\hat{\sigma}$

5.3.2. Average precision The average precision is analyzed on several cases. For this purpose, 30 independent observation data $\left\{\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(30)}\right\}$ are simulated. For each of

these observation vectors $\mathbf{y}^{(r)}$, the inversion algorithm computes the pair of estimators $(\hat{\mathbf{x}}^{(r)}, \hat{\sigma}^{(r)})$. The comparison with the true values of $\mathbf{x}$ is quantified by the following root mean square error (RMSE) :

$$\mathrm{RMSE}(i,k) := \left(\frac{1}{30} \sum_{r=1}^{30} \left(\hat{\mathbf{x}}^{(r)}(i,k) - \mathbf{x}_{\mathrm{true}}(i,k) \right)^2 \right)^{1/2}$$

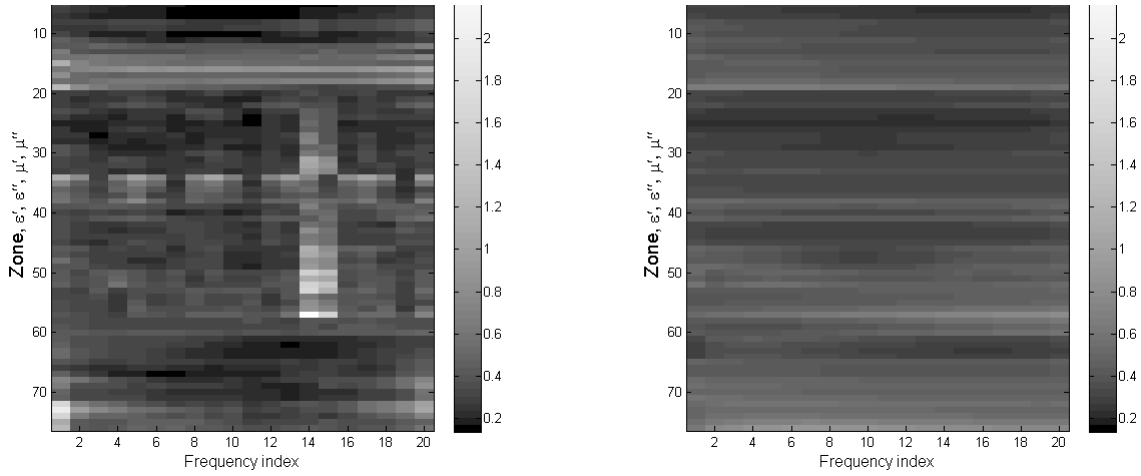

Figure 14. RMSE (left) and estimated marginal uncertainties (right)

These errors are shown on figure 14, where they can be compared to the estimated errors $\bar{\hat{\sigma}}$. From these results, the following conclusions can be drawn. Despite the large amplitude and irregularity of the perturbations, far from the assumed prior model, the estimators $\hat{\mathbf{x}}^{(r)}$ give a good approximation of $\mathbf{x}_{\mathrm{true}}$ (note that the mean RMSE $= 3.68\ \ 10^{-1}$). Moreover, the RMSE values are comparable to the marginal uncertainties given by $\bar{\hat{\sigma}}$, which proves that the estimated posterior variances make sense.

In this inversion process, the role of ρ's estimation is very interesting. Roughly speaking, it is as if it could give in advance the shape type of each of the unknown true frequencial profile, by estimating its regularity. On figure 15, we show the results given by $(\hat{\mathbf{x}}^{(1)}, \hat{\sigma}^{(1)})$ for the zones number 2, 9 and 17 and the permeability μ'. On the right part, the histograms represent the posterior distribution of ρ. As predictable, the difficulty increases from zone 2 to zone 17, due to the perturbation which is larger and larger, as well as irregular. On the right side of the figure, we show the histograms of all the particles $(\rho^{(1)}, \ldots, \rho^{(100)})$ (each particle being represented by its associated component). We clearly see that the more irregular the true signal is, the smaller the $\rho^{(i)}$ are, which is quite coherent since ρ quantifies frequential correlation. Meanwhile, we check in the center of the figure that, in spite of the increasing difficulty, the mean RMSE remains stable. Again, let us stress that the adaptive behavior of ρ estimation is essential to the algorithm robustness.

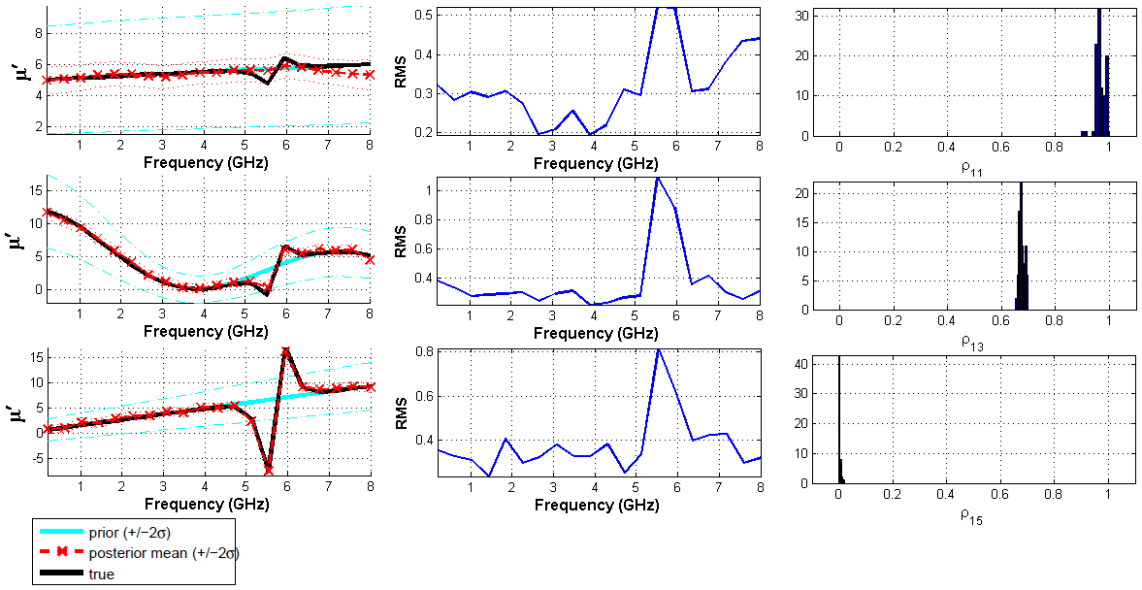

Figure 15. μ' estimators for zones 2, 9 and 17

5.4. Additional analyses

We propose now to briefly analyze the influence of other parameters, that can come from the context or from the inversion process itself.

5.4.1. Influence of the Processing Parameters The inversion process we have described in section 4 admits several qualitative and quantitative degrees of freedom, particularly in the SMC step. We propose here our empirical remarks about some of them.

The number of particles N_p Like a classic i.i.d. (independent and identically distributed) sampling method, the SMC algorithm precision is proportional to $N_p^{-1/2}$. However, in our problem, the main objective is not to have a precise estimation of η, but of $\mathbf{x}$. As the impact of a local variance of ρ on $\mathbf{x}$ is rather small, the crucial point is that the global cloud of particles reaches the correct area in E. From this point of view, the important condition is the stability of the Feynman-Kac flow (see [34]), which ensures that the particles don't get lost in E. This is precisely the purpose of the adaptive strategies inside the selection and mutation steps). That's why it seems useless (and time consuming) to use a high number of particles. Note that below $N_p \simeq 40$, the SMC approach may be trapped by some local modes.

The interpolating scheme η_n In addition to the annealed probability measure scheme, the hybrid one has been tested. It can assimilate the observations one by one, and update the estimators progressively. Moreover, it manages the computational problems of selectivity that affects the data tempered scheme. The results are good, nearly identical to those obtained with the annealed scheme. And yet, the SMC algorithm

lasts around 4 times longer than before.

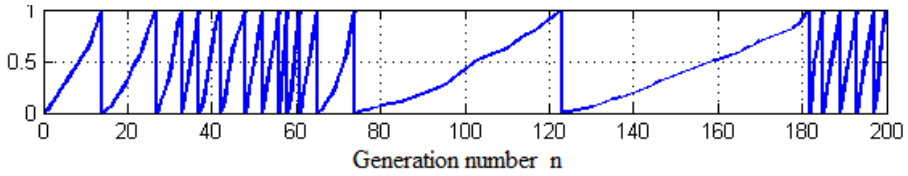

Figure 16. Annealing parameter $\alpha_i^{(k)}$

This behavior can be easily interpreted by figure 16. It appears that many observations do not bring any new information, so that the associated annealing sequence $\alpha_i^{(k)}$ takes value 1 at once. On the contrary, when a new observation provides information in contradiction with the previous ones, the particles have to migrate from an area of E to another, which takes a longer time (i.e. more steps).

The parameter ρ The prior distribution $p(\rho)$ of figure 11 has a limited impact on the final estimation of η. Corresponding to a prior knowledge of frequency regularity, it is arbitrarily chosen in order to penalize the small values and favor regular profiles. But in practice, this penalization term $p(\rho)$ is less determining than the likelihood one $p(\mathbf{y}|\rho)$. Besides, ρ can be defined 5-dimensional. In this case, the SMC algorithm performs quicker. However, the underlying hypothesis – i.e. the frequential correlation is the same for ϵ', ϵ'', μ', μ'' – is not necessarily fulfilled in practice.

5.4.2. Context influence As we have mentioned, the method is very robust concerning the amplitude and the irregularity of the perturbation (i.e. the deviation from the reference profiles).

Measurement noise However, it is naturally sensitive to the observation noise magnitude. Its performance degrades when the observation noise is too high (low SNR). That is clearly a matter of information. Numerically, it can be explained by considering the accurate approximation given by the surrogate model. Indeed, the $\mathbf{A}_k$ matrices are ill-conditioned. In particular, μ' and ϵ'' components are highly correlated; it is the same for μ'' and ϵ'. Notice that only more regularization, i.e. the introduction of more prior knowledge (precision, correlation, etc.), could counterbalance a low SNR; it is possible in our Bayesian framework.

In figure 17, we give the estimations of these 2 quantities in case the amplitude noise $\sigma_{\mathrm{n}} = 10^{-2}$ (instead of 10^{-3}), corresponding to a lower SNR (20 dB instead of $40dB$ at high RCS incidence). One can then see that the unknown perturbations of μ'_{true} and $\epsilon''_{\mathrm{true}}$ are correctly detected by the process, but improperly distributed between μ' and ϵ''. Notice that the initial SNR requirements are not really demanding, according to what is usually performed in RCS measurement inside an indoor anechoic chamber

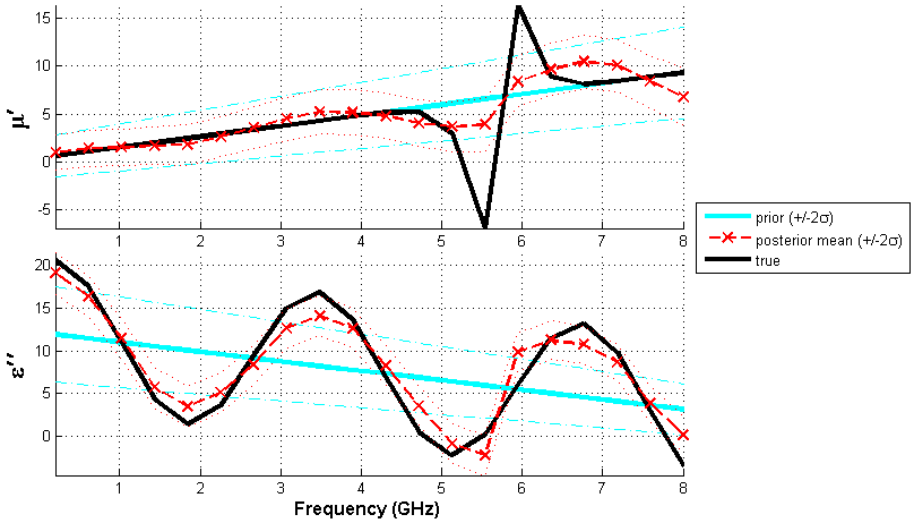

Figure 17. Estimation of μ' and ϵ'', $\sigma_{\mathrm{n}} = 10^{-2}$

(with the use of absorbers, range gates, time/frequency domain cancellation, etc.) [49]. Again, let us repeat that when the SNR is very low, no information can emerge.

Computational time Concerning the computation time, it depends on the problem, i.e. both on the dimension and the SMC algorithm convergence rate. To look further into this matter, a set of numerical experiments have been designed, from the former described application of section 5.1. The 14 cases are summed up in the table of figure 18. The problem dimension, i.e. the state space dimension N and the measurement dimension $\nu = 4 \cdot K_\theta$, is chosen in ascending order. Notice that the number K_f of regularly spaced frequency acquisitions K_f is here unchanged (and equal to 20). The associated results are performed by averaging tens of runs. The results are given in the right part of the table. For each case, according to the chosen N and ν dimensions, the averaged following values are given in the 3 last columns: the CPU time (the mean processing duration of the Rao-Blackwellised SMC approach), the mean number of likelihood evaluations (computations of $p(\mathbf{y}|\rho)^{\alpha_n}$ in the annealing scheme of section 4.2) and the number of adaptive steps (on $\Delta\alpha_n$).

As both N and ν increases, its computational requirements turn out to be slightly higher than quadratic. The Kalman filter complexity is known to be approximately cubic [33], considering both the state and measurement dimensions. Actually, it is around $N^{2.5}$ in our specific application. All in all, since it is involved in the numerous likelihood evaluations, it dominates the other contributions. Indeed, the numbers of likelihood evaluations and adaptive steps are relatively stable. Their slight increase comes from the adaptive behavior that requires more steps when there is more information and the posterior distribution lives in a smaller part of the space. Furthermore, the overall computation time is roughly linear according to the frequency number K_f. Besides,

Figure 18. Computation time and problem dimension

Case	N	ν	CPU time (s)	Mean likelihood #	Mean adaptive step #
1	20	20	146.3	21680	37.2
2	24	24	272.0	30710	53.0
3	32	28	291.6	25720	43.6
4	40	40	654.6	36510	63.0
5	44	44	768.2	35760	62.1
6	52	52	1042.7	38490	66.4
7	56	64	1364.7	40380	69.6
8	76	84	2092.7	42060	72.2
9	88	104	2683.6	41950	71.5
10	100	124	3546.7	43200	73.0
11	112	148	4644.0	43950	74.5
12	144	184	7523.9	39700	66.0
13	188	244	15199.0	47200	79.0
14	276	364	36364.7	45900	76.0

notice that for a given frequency correlation length, more frequencies mean smaller frequency intervals. It leads to a matrix $\mathbf{D}_\rho$, that expresses higher correlations and takes part in Kalman filter steps (see the AR dynamic model of section 3.2).

6. Conclusion

An efficient statistical inference approach has been applied. From global EM scattering measurements, it enables to estimate local radioelectric properties of materials assembled and placed on the full-scaled object, assuming that its shape is perfectly known in the context of a quality control application. The inverse problem is solved by combining intensive computations with high performance computing (HPC), surrogate modeling and advanced sequential Monte Carlo techniques dedicated to frequency dynamic estimation. It takes advantage of the problem structure to achieve a Rao-Blackwellisation strategy of Monte Carlo variance reduction. On top of that, the Bayesian approach quantifies the uncertainties around the estimates, with limited computation time.

Future research could take into account bistatic scattering measurement to enlarge acquisition information, similarly to the classic NRL arch method [49]. On the other hand, it could be rather straightforward to complexify the prior knowledge on the material microwave properties. It would consist in introducing a hyperparametric model that would be part of the inference process. Another promising perspective is to deal with the computation time issue which arises for large dimension problems, especially

where the number of elementary areas is around a few hundreds and more. To tackle this problem and achieve the estimation goal for larger dimensions, different strategies can be applied. The Rao-Blackwellised SMC approach can be parallelized at the fine-grained level of each Kalman filter and its required linear algebra manipulations. It can be also parallelized at a coarse-grained level: the particles and the likelihood evaluations can be distributed massively. From this highly parallelization/distribution potential on HPC, it would be possible to tackle high-dimensional problems as well as 3D geometries. Another less computer resource demanding approach could be to substitute the workhorse, i.e. the Kalman filter, by a fast approximate algorithm, such as the Ensemble Kalman (EnKF) [50]. Related to particle filtering, EnKF has been developed for large data assimilation problems in geophysical and weather forecasting. Finally, close stochastic techniques, such as "interacting Kalman filters", could be applied on condition that a loosen formulation is chosen.

Acknowledgments

We greatfully acknowledge help from P. Bonnemason and F. Caron, about various electromagnetism and computational statistics points.

Appendix: estimation of $\mathbf{x}_k$ and conditioning

For a given $\mathbf{y}$, it is possible to define judicious estimators of $\mathbf{x}_k$ (for each k). Indeed, the first moments $\hat{\mathbf{x}}_k$ and $\hat{\mathbf{\Sigma}}_k$) can be determined from the theoretical conditional expectation $\bar{\mathbf{x}}_k := \mathbb{E}[\mathbf{x}_k|\mathbf{y}]$ and covariance matrix $\mathbf{\Sigma}_k := \mathbb{V}\mathrm{ar}[\mathbf{x}_k|\mathbf{y}]$.

For all $\rho \in E$, let set: $\hat{\mathbf{x}}_k(\rho) := \mathbb{E}\left[\mathbf{x}_k|\rho, \mathbf{y}\right]$ and $\hat{\mathbf{\Sigma}}_k(\rho) := \mathbb{V}\mathrm{ar}\left[\mathbf{x}_k|\rho, \mathbf{y}\right]$, i.e. the main quantities provided by the Kalman smoother. Under this notation, we combine $\hat{\eta} \simeq \eta$ together with the equation (15), and derive a natural choice for the estimator $\hat{\mathbf{x}}_k$:

$$\begin{aligned}
\bar{\mathbf{x}}_k \;=\; \mathbb{E}[\mathbf{x}_k|\mathbf{y}] \;&=\; \mathbb{E}[\underbrace{\mathbb{E}(\mathbf{x}_k|\rho,\mathbf{y})}_{\hat{\mathbf{x}}_k(\rho)}|\mathbf{y}] \\
&= \int_{\rho\in E} \hat{\mathbf{x}}_k(\rho)\eta(d\rho) \;\simeq\; \int_{\rho\in E} \hat{\mathbf{x}}_k(\rho)\hat{\eta}(d\rho) \\
&= \underbrace{\frac{1}{N_p}\sum_{i=1}^{N_p} \hat{\mathbf{x}}_k(\rho^{(i)})}_{=:\hat{\mathbf{x}}_k}
\end{aligned}$$

Regarding the covariance estimator $\hat{\mathbf{\Sigma}}_k$, under the same notation and according to (16), we have:

$$\mathbf{\Sigma}_k = \underbrace{\mathbb{E}\left(\hat{\mathbf{\Sigma}}_k(\rho)|\mathbf{y}\right)}_{\mathbf{\Sigma}_k^{(1)}} + \underbrace{\mathbb{V}\mathrm{ar}\left(\hat{\mathbf{x}}_k(\rho)|\mathbf{y}\right)}_{\mathbf{\Sigma}_k^{(2)}}.$$

We estimate $\mathbf{\Sigma}_k^{(1)}$ and $\mathbf{\Sigma}_k^{(2)}$ separately:

(i) Evaluation of $\boldsymbol{\Sigma}_k^{(1)} = \mathbb{E}\left(\hat{\boldsymbol{\Sigma}}_k(\rho)|\mathbf{y}\right)$

$$\boldsymbol{\Sigma}_k^{(1)} = \int_{\rho\in E} \hat{\boldsymbol{\Sigma}}_k(\rho)\eta(d\rho) \simeq \int_{\rho\in E} \hat{\boldsymbol{\Sigma}}_k(\rho)\hat{\eta}(d\rho)$$
$$= \underbrace{\frac{1}{N_p}\sum_{i=1}^{N_p} \hat{\boldsymbol{\Sigma}}_k(\rho^{(i)})}_{=:\hat{\boldsymbol{\Sigma}}_k^{(1)}}$$

(ii) Evaluation of $\boldsymbol{\Sigma}_k^{(2)} = \mathbb{E}\left[\left(\hat{\mathbf{x}}_k(\rho) - \bar{\mathbf{x}}_k\right)\left(\hat{\mathbf{x}}_k(\rho) - \bar{\mathbf{x}}_k\right)^T |\mathbf{y}\right]$

$$\boldsymbol{\Sigma}_k^{(2)} = \int_{\rho\in E} \left(\hat{\mathbf{x}}_k(\rho) - \bar{\mathbf{x}}_k\right)\left(\hat{\mathbf{x}}_k(\rho) - \bar{\mathbf{x}}_k\right)^T \eta(d\rho) \simeq \int_{\rho\in E} \left(\hat{\mathbf{x}}_k(\rho) - \hat{\mathbf{x}}_k\right)\left(\hat{\mathbf{x}}_k(\rho) - \hat{\mathbf{x}}_k\right)^T \hat{\eta}(d\rho)$$
$$= \underbrace{\frac{1}{N_p}\sum_{i=1}^{N_p} \left(\hat{\mathbf{x}}_k(\rho^{(i)}) - \hat{\mathbf{x}}_k\right)\left(\hat{\mathbf{x}}_k(\rho^{(i)}) - \hat{\mathbf{x}}_k\right)^T}_{=:\hat{\boldsymbol{\Sigma}}_k^{(2)}}$$

Finally, the estimator of $\boldsymbol{\Sigma}_k$ is given by: $\hat{\boldsymbol{\Sigma}}_k := \hat{\boldsymbol{\Sigma}}_k^{(1)} + \hat{\boldsymbol{\Sigma}}_k^{(2)}$.

References

[1] K I Hopcraft and P R Smith. *An introduction to electromagnetic inverse scattering*. Kluwer Academic Publishers, 1992.

[2] D Colton, H Haddar, and M Piana. The linear sampling method in inverse electromagnetic scattering theory. *Inverse problems*, 19:S105, 2003.

[3] J P Kaipio and E Somersalo. *Statistical and computational inverse problems*, volume 160. Springer Science+ Business Media, 2005.

[4] E F Knott, J F Shaeffer, and M T Tuley. *Radar cross section*. SciTech Publishing, 2004.

[5] R Abbato. Dielectric constant measurements using RCS data. *Proceedings of the IEEE*, 53(8):1095–1097, 1965.

[6] S Coen. Inverse scattering of the permittivity and permeability profiles of a plane stratified medium. *Journal of Mathematical Physics*, 22:1127, 1981.

[7] L Garnero, A Franchois, J P Hugonin, C Pichot, and N Joachimowicz. Microwave imaging-complex permittivity reconstruction-by simulated annealing. *IEEE Transactions on Microwave Theory and Techniques*, 39(11):1801–1807, 1991.

[8] S Caorsi, G L Gragnani, and M Pastorino. Two-dimensional microwave imaging by a numerical inverse scattering solution. *IEEE Transactions on Microwave Theory and Techniques*, 38(8):981–980, 1990.

[9] A Ribes and F Schmitt. Linear inverse problems in imaging. *Signal Processing Magazine, IEEE*, 25(4):84–99, 2008.

[10] G Demoment. Image reconstruction and restoration: Overview of common estimation structures and problems. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(12):2024–2036, 1989.

[11] Q H Liu, Z Q Zhang, T T Wang, J A Bryan, G A Ybarra, L W Nolte, and W T Joines. Active microwave imaging: 2-d forward and inverse scattering methods. *Microwave Theory and Techniques, IEEE Transactions on*, 50(1):123–133, 2002.

[12] P M Abubakar. Contrast source inversion method: State of art. *Progress In Electromagnetics Research*, 34:189–218, 2001.

[13] W C Chew and Y M Wang. Reconstruction of two-dimensional permittivity distribution using the distorted born iterative method. *Medical Imaging, IEEE Transactions on*, 9(2):218–225, 1990.

[14] A Abubakar, T M Habashy, P M Van den Berg, and D Gisolf. The diagonalized contrast source approach: an inversion method beyond the born approximation. *Inverse problems*, 21(2):685, 2005.

[15] R E Kleinman and P M Berg. A modified gradient method for two-dimensional problems in tomography. *Journal of Computational and Applied Mathematics*, 42(1):17–35, 1992.

[16] F Collino, M B Fares, and H Haddar. Numerical and analytical studies of the linear sampling method in electromagnetic inverse scattering problems. *Inverse Problems*, 19:1279, 2003.

[17] Jiang Wan and Nicholas Zabaras. A Bayesian approach to multiscale inverse problems using the sequential monte carlo method. *Inverse Problems*, 27(10):105004, 2011.

[18] C Campi, A Pascarella, A Sorrentino, and M Piana. A rao–blackwellized particle filter for magnetoencephalography. *Inverse Problems*, 24(2):025023, 2008.

[19] P Del Moral, A Doucet, and A Jasra. Sequential Monte Carlo for Bayesian computation. *Bayesian Statistics*, 8(1):34, 2007.

[20] S Geman and D Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6):721–741, 1984.

[21] M C Robini, T Rastello, and I E Magnin. Simulated annealing, acceleration techniques, and image restoration. *IEEE Transactions on Image Processing*, 8(10):1374–1387, 1999.

[22] A D Lanterman. Statistical radar imaging of diffuse and specular targets using an expectation-maximization algorithm. In *SPIE Proceedings*, pages 20–31. Citeseer, 2000.

[23] M Pastorino. Stochastic optimization methods applied to microwave imaging: A review. *IEEE Transactions on Antennas and Propagation*, 55(3):538–548, 2007.

[24] E Michielssen, J M Sajer, S Ranjithan, and R Mittra. Design of lightweight, broad-band microwave absorbers using genetic algorithms. *IEEE Transactions on Microwave Theory and Techniques*, 41(6):1024–1031, 1993.

[25] B Chambers and A Tennant. Optimised design of radar absorbing materials using a genetic algorithm. In *IEE Radar, Sonar and Navigation Proceedings*, volume 143, pages 23–30. IET, 1996.

[26] J S Liu and R Chen. Sequential Monte Carlo methods for dynamic systems. *Journal of the American statistical association*, pages 1032–1044, 1998.

[27] P Minvielle, E Tantar, A Tantar, and P Berisset. Sparse antenna array optimization with the cross-entropy method. *IEEE Transactions on Antennas and Propagation*, 59(8):2862–2871, 2011.

[28] D R Wehner. High resolution radar. *Norwood, MA, Artech House, Inc., 1995*, 1995.

[29] A C Woo, H T G Wang, M J Schuh, and M L Sanders. EM programmer's notebook-benchmark radar targets for the validation of computational electromagnetics programs. *Antennas and Propagation Magazine, IEEE*, 35(1):84–89, 1993.

[30] B Stupfel, R Le Martret, P Bonnemason, and B Scheurer. Combined boundary-element and finite-element method for the scattering problem by axisymmetrical penetrable objects. *Mathematical and numerical M*, 2:78–17, 1991.

[31] IEEE Recommended Practice for Radar Cross-Section Test Procedures. *IEEE Std 1502-2007*, pages 1 –60, 7 2007.

[32] Jonathan R Stroud, P Müller, and B Sansó. Dynamic models for spatiotemporal data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(4):673–689, 2001.

[33] Y Bar-Shalom, X R Li, and T Kirubarajan. *Estimation with applications to tracking and navigation: theory algorithms and software*. Wiley-Interscience, 2004.

[34] P Del Moral. *Feynman-Kac formulae. Genealogical and interacting particle approximations*. Springer New York, 2004.

[35] O Cappé, E Moulines, and T Rydén. *Inference in hidden Markov models*. Springer Series in

Statistics. Springer, New York, 2005.

[36] A Doucet, N De Freitas, and N Gordon. *Sequential Monte Carlo Methods in Practice*. Springer New York, 2001.

[37] P Del Moral. Nonlinear filtering: Interacting particle solution. *Markov Processes and Related Fields*, 2(4):555–579, 1996.

[38] F Cérou, P Del Moral, T Furon, and A Guyader. Sequential monte carlo for rare event estimation. *Statistics and Computing*, 22(3):795–808, 2012.

[39] P Del Moral, A Doucet, and A Jasra. Sequential Monte Carlo Samplers. *Journal of the Royal Statistical Society, Series B*, 68:411–436, 2006.

[40] P Del Moral and A Guionnet. Central limit theorem for nonlinear filtering and interacting particle systems. *Annals of Applied Probability*, pages 275–297, 1999.

[41] N Chopin. Central limit theorem for sequential monte carlo methods and its application to Bayesian inference. *The Annals of Statistics*, 32(6):2385–2411, 2004.

[42] P Del Moral. *Feynman-Kac Formulae*. Springer, 2004.

[43] F Giraud and P Del Moral. Non-asymptotic analysis of adaptive and annealed Feynman-Kac particle models. *arXiv preprint arXiv:1209.5654*, 2012.

[44] N Schweizer. *Non-asymptotic error bounds for Sequential MCMC methods*. PhD thesis, Ph. D. thesis, Universität Bonn, 2011.

[45] N Whiteley. Sequential monte carlo samplers: error bounds and insensitivity to initial conditions. *Stochastic Analysis and Applications*, 30(5):774–798, 2012.

[46] P Del Moral and A Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. *Annales de l'Institut Henri Poincaré*, 37(2):155–194, 2001.

[47] A Jasra, D A Stephens, and C C Holmes. On population-based simulation for static inference. *Statistics and Computing*, 17(3):263–279, 2007.

[48] P Minvielle, A Doucet, A. Marrs, and S Maskell. A Bayesian approach to joint tracking and identification of geometric shapes in video sequences. *Image and Vision Computing*, 28(1):111–123, 2010.

[49] E F Knott. *Radar cross section measurements*. SciTech Publishing, 2006.

[50] G Evensen. The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343–367, 2003.