MCMC and variational approaches for Bayesian inversion in diffraction imaging

H Ayasso, Bernard Duchêne, Ali Mohammad-Djafari

Contents

Chapter 1

MCMC and Variational Approaches for Bayesian Inversion in Diffraction Imaging

1.1. Introduction

The term "diffraction imaging" is meant, herein, in the sense of an "inverse scattering problem" where the goal is to build up an image of an unknown object (a mapping of physical features such as dielectric permittivity) from measurements of the scattered field that results from its interaction with a known probing wave (denoted as incident wave). This type of problem occurs in many imaging and non-destructive testing applications. It corresponds to the situation where looking for a good trade-off between the image resolution (the smallest observable detail) and the penetration of the incident wave in the probed medium, leads to choosing the frequency of the latter in such a way that its wavelength lies in the "resonance" domain, in the sense that it is approximately of the same order of magnitude as the characteristic dimensions of the inhomogeneities of the inspected object. In this situation the wave-object interaction gives rise to important diffraction phenomena. This is the case for the two applications considered herein, where the interrogating waves are electromagnetic waves with wavelengths in the microwave and optical domains, whereas the characteristic dimensions of the sought object are 1 cm and 1 μm, respectively.

The solution of an inverse problem obviously requires previous construction of a forward model that expresses the scattered field as a function of the parameters of the sought object. In this model, diffraction phenomena are taken into account by means

---

Chapter written by Hacheme AYASSO, Bernard DUCHÊNE and Ali MOHAMMAD-DJAFARI.

of domain integral representations of the electric fields. The forward model is then described by two coupled integral equations, whose discrete versions are obtained using a method of moments and whose inversion leads to a non-linear problem.

Concerning inversion, at the beginning of the 1980s, accounting for the diffraction phenomena has been the subject of much attention in the field of acoustic imaging for applications in geophysics [DEV 84], non-destructive testing or biomedical imaging [MUE 80]. It led to techniques such as diffraction tomography, a term that denotes [SCH 92] "applications that employs diffracting wavefields in the tomographic reconstruction process" , but which generally implies reconstruction processes based on the generalized projection-slice theorem, an extension to the diffraction case of the projection-slice theorem of the classical computed tomography (CT) whose forward model is given by a Radon transform . This theorem is based upon first-order linearizing assumptions such as the Born's or Rytov's approximations. So, the term diffraction tomography was paradoxically used to describe reconstruction techniques adapted to weakly scattering environments that do not provide quantitative information on highly contrasted dielectric objects [AZI 83, SLA 84] such as those encountered in the applications considered herein, where multiple diffraction cannot be ignored.

Furthermore, the resolution of these techniques is limited because evanescent waves are not taken into consideration. These limitations have led researchers to develop inversion algorithms able to deal with non-linear problems, at the beginning of the 1990s for microwave imaging and more recently for optical imaging [BEL 03]. Many studies have focused on the development of deterministic methods, such as the Newton-Kantorovich algorithm [JOA 91], the modified gradient method (MGM, [KLE 92]) or the contrast-source inversion technique (CSI, [BER 97]), where the solution is sought for by means of an iterative minimization by a gradient method of a cost functional that expresses the difference between the scattered field and the estimated model output. But, in addition to be non-linear, inverse scattering problems are also known to be ill-posed, which means that their resolution requires a regularization which generally consists in introducing prior information on the sought object. In the present case, for example, we look for man-made objects that are composed of homogeneous and compact regions made of a finite number of different materials, and with the aforementioned deterministic methods, it is not easy to take into account such prior information because it must be introduced into the cost functional to be minimized.

On the contrary, the probabilistic framework of Bayesian estimation [IDI 13], basis of the model presented herein, is especially well suited for this situation. Prior information is appropriately introduced via a probabilistic Gauss-Markov-Potts model [PIE 03, TIE 94]. The marginal contrast distribution is modeled as a mixture of Gaussians [FÉR 02], where each Gaussian distribution represents a class of materials and the compactness of the regions is taken into account using a hidden Markov

model. Estimation of the unknowns and parameters introduced into the prior model is performed via an unsupervised joint approach.

Two iterative algorithms are proposed. The first one, denoted as the MCMC algorithm (*Monte-Carlo Markov Chain*), is rather classic ; it consists in expressing all the joint posterior or conditional distributions of all the unknowns and, then, using a Gibbs sampling algorithm [ROB 04] for estimating the posterior mean of the unknowns. This algorithm yields good results, however, it is computationally intensive mainly because Gibbs sampling requires a significant number of samples.

The second algorithm is based upon the variational Bayesian approximation (VBA) [SMÍ 06]). The latter was first introduced in the field of Bayesian inference for applications to neural networks [HIN 93], learning graphic models[JOR 99] and model parameter estimation [JAA 00]. Its appearance in the field of inverse problems is relatively recent, starting with source separation [CHO 02] and image restoration [LIK 04]. It consists in approximating the joint posterior distribution of all the unknowns by a free-form separable distribution that minimizes, with respect to the posterior law, the Kullback-Leibler divergence [KUL 51] which has interesting properties for optimization and leads to an implicit parametric optimization scheme. Once the approximate distribution is built up, the estimator can be easily obtained.

A solution to this functional optimization problem can be found in terms of exponential distributions whose shape parameters are estimated iteratively. It can be noted that, at each iteration, the updating expression for these parameters is similar to the one that could be obtained if a gradient method was used to solve the optimization problem. Moreover, the gradient and the step size have an interpretation in terms of statistical moments (means, variances, etc.) [SAT 01].

Both algorithms introduced herein are applied to two quite different configurations. The one related to microwave imaging is quasi-optimal: data are quasi-complete and frequency diverse. This means that the scattered fields are measured all around the object for several directions of illumination and several frequencies. The configuration used in optical imaging is less favorable since only aspect-limited data are available at a single frequency. This means that illuminations and measurements can only be performed in a limited angular sector. This limited aspect reinforces the ill-posedness of the inverse problem and makes essential the introduction of prior information. However, it will be shown that, in both cases, satisfactory results are obtained.

This chapter consists of four parts. The first one describes the experimental configurations corresponding to the two studied applications : microwave imaging and optical imaging. The second part is devoted to the direct problem, its discrete formulation and its validation against laboratory controlled data. The third part deals with a Bayesian approach of inversion and with the description of the MCMC and

VBA algorithms. The fourth section presents the results obtained using the two algorithms in microwave imaging and optical imaging and draws some conclusions.

1.2. Measurement configuration

The experimental set-ups considered in this chapter were developed at the Fresnel Institute (Marseille). They are presented in more detail in [GEF 05] for microwave imaging and in [MAI 10] for optical imaging. Both devices have common characteristics that are modeled as follows. The objects are considered as cylinders with infinite extension along the Oz axis and of arbitrary cross-section Ω. These objects are illuminated by a harmonic incident wave, with an angular frequency ω, whose implicit time dependence is chosen as exp(-i ωt). This incident wave can be regarded as a plane wave whose electric field is polarized along the Oz axis so that a 2D configuration is considered in a transverse magnetic polarization case which leads to a scalar formulation of the electric fields. The different media are assumed to be non magnetic and lossless (however, it can be noted that the case of lossy media is not particularly problematic) and are characterized by their propagation constants k_m ($m = 1, 2$ or Ω) such that $k_m^2 = \omega^2 \varepsilon_0 \varepsilon_m \mu_0$, where ε_0 ($\varepsilon_0 = 8.854 \times 10^{-12}$ F.m^{-1}) and μ_0 ($\mu_0 = 1.256 \times 10^{-6}$ H.m^{-1}) are the dielectric permittivity and magnetic permeability of vacuum and ε_m is the relative dielectric permittivity of medium $\mathcal{D}_m$. The object is assumed to be contained in a test domain $\mathcal{D}$ ($\mathcal{D} \subset \mathcal{D}_1$) in which we define a normalized contrast function χ, such that $\chi(\boldsymbol{r}) = (k^2(\boldsymbol{r}) - k_1^2)/k_1^2 = \varepsilon(\boldsymbol{r}) - 1$, representative of the electromagnetic parameters of the object and null outside Ω.

1.2.1. *The microwave device*

The microwave set-up consists of a network analyzer coupled to two horn antennas, one of which is for emission and the other one for reception. The device operates at $N_f = 9$ frequencies ranging from 2 to 10 GHz. The studied object $\mathcal{M}$ consists of three dielectric cylinders with circular cross-sections, two of which are made of plastic and are identical, with a diameter of 3 cm and a relative permittivity of $\varepsilon_\Omega = 3$, and the third one consists of foam and has a diameter of 8 cm and a relative permittivity of $\varepsilon_\Omega = 1.45$ (figure 1.1). This object is placed in air (medium $\mathcal{D}_1$) and is illuminated under an incidence angle θ_1 which may vary around the object. Thus, $N_v = 18$ views are performed at varying θ_1, each view consisting of scattered field measurements at a fixed distance $r = 1.67$ m for $N_r = 241$ different observation directions θ, uniformly distributed in a sector of 240° ($\theta \in \theta_1 \pm 120°$).

1.2.2. *The optical device*

The optical set-up consists of a helium-neon laser coupled to a microscope operating in the reflection mode and equipped with an interferometer that can provide

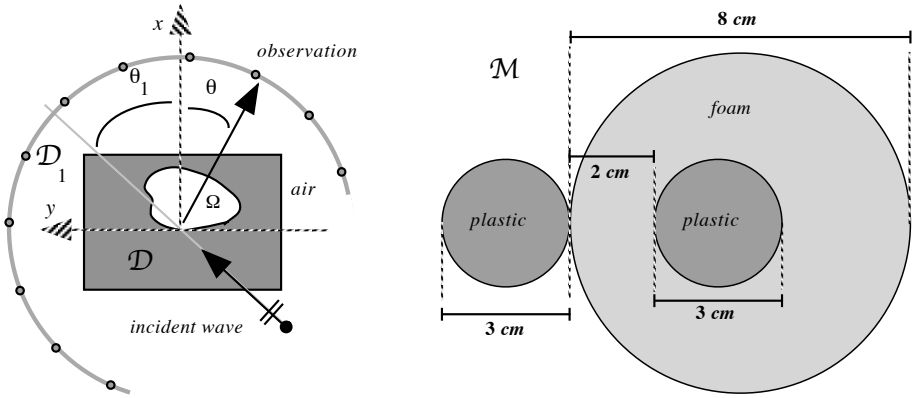

Figure 1.1. *Measurement configuration for microwave imaging (left) and geometry of object $\mathcal{M}$ (right)*

accurate estimations of the phase of the measured electric fields. This device operates at a frequency of 473.6 THz, which corresponds to a wavelength of 633 nm in vacuum. The sought object is typical of those encountered in the field of nanotechnology. It consists of two resin rods with rectangular cross-sections, arranged in parallel on a silicon substrate, whose dimensions are large compared to the rod cross-sections so that it can be regarded as of infinite extension. The measurement configuration can be modeled as follows (figure 1.2): an object $\mathcal{O}$ consisting of resin, whose relative permittivity is $\varepsilon_\Omega = 2.66$, is located in the top layer of a stratified medium, consisting of two semi-infinite half-spaces separated by a horizontal planar interface γ_{12}. The top half-space $\mathcal{D}_1$ is air ($\varepsilon_1 = 1$) while the lower one $\mathcal{D}_2$ consists of silicon with a relative permittivity of $\varepsilon_2 = 15.07$. The rods are of same height ($0.14\ \mu$m) but they are of different widths ($1\ \mu$m and $0.5\ \mu$m, respectively) and are spaced $0.5\ \mu$m apart from one another. This object is illuminated at an incidence θ_1 which can vary in a sector of $\pm 32^\circ$. Thus, $N_v = 8$ views are performed at varying θ_1, each view consisting of measurements of the reflected scattered field at a fixed distance r and for $N_r = 611$ different observation angles θ in a sector of $\pm 46^\circ$.

The data collected at the Fresnel Institute using these experimental configurations (courtesy of K. Belkebir and M. Saillard for microwave data and G. Maire, K. Belkebir and A. Sentenac for optical data) will allow us to validate the forward model and to test the inversion algorithms without committing an "inverse crime" in the sense of [COL 92] that would consist in testing the latter on synthetic data obtained by solving the forward problem with the help of a method closely related to the one used to solve the inverse problem.

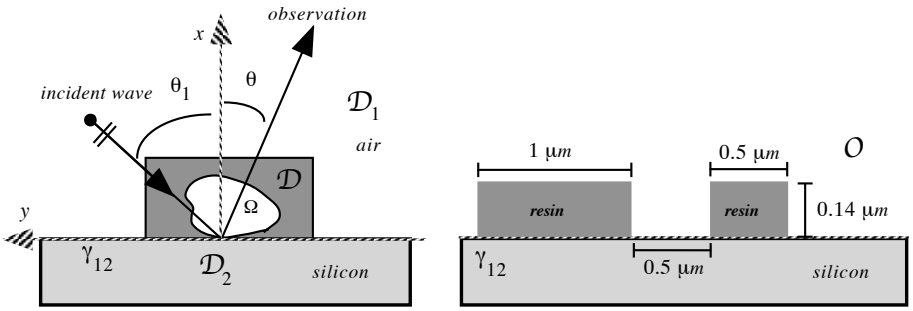

Figure 1.2. *Measurement configuration for optical imaging (left) and geometry of object $\mathcal{O}$ (right)*

1.3. The forward model

The forward model is based upon domain integral representations of the electric fields obtained by applying the Green's theorem to the Helmholtz wave equations satisfied by the fields and by accounting for continuity and radiation conditions [CHE 95, COL 92]. This yields two coupled integral equations, whose first one, denoted as coupling or state equation, links the total electric field E in $\mathcal{D}$ to the Huygens type sources $w(\boldsymbol{r}')$ induced in the object by the incident wave, *i.e.,* $w(\boldsymbol{r}') = \chi(\boldsymbol{r}')E(\boldsymbol{r}')$ where χ is the contrast function. This equation reads:

$$E(\boldsymbol{r}) = E^{\text{inc}}(\boldsymbol{r}) + k_1^2 \int_{\mathcal{D}} \mathcal{G}(\boldsymbol{r},\boldsymbol{r}')w(\boldsymbol{r}')\ \mathrm{d}\boldsymbol{r}', \boldsymbol{r} \in \mathcal{D}. \quad [1.1]$$

E^{inc} and $\mathcal{G}(\boldsymbol{r},\boldsymbol{r}')$, which will be shown in more detail later, represent the incident field and the Green's function in the considered medium, respectively. The second equation, denoted as observation equation, is a Fredholm integral equation of the first kind. It links the scattered field E^{dif} observed in the measurement domain $\mathcal{S}$ to the induced sources $w(\boldsymbol{r}')$:

$$E^{\text{dif}}(\boldsymbol{r}) = k_1^2 \int_{\mathcal{D}} \mathcal{G}(\boldsymbol{r},\boldsymbol{r}')w(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}', \boldsymbol{r} \in \mathcal{S}. \quad [1.2]$$

This model describes the two configurations considered here. They only differ in the incident field and Green's function expressions because in one case (microwave) the sought object is located in a homogeneous medium, while in the other case (optical) it is embedded in a stratified medium.

1.3.1. *The microwave case*

Since the object is placed in a homogeneous medium, the incident field and Green's function expressions are fairly simple. Thus, the incident field, which is a

plane wave propagating in the direction θ_1, reads at point $\boldsymbol{r} = (x, y)$:

$$E^{\text{inc}}(\boldsymbol{r}) = \exp(\mathrm{i}k_1(x\cos\theta_1 + y\sin\theta_1)), \qquad [1.3]$$

while the Green's function $\mathcal{G}(\boldsymbol{r}, \boldsymbol{r}')$, which represents the radiation of a line-source placed at $\boldsymbol{r}'$ and observed at $\boldsymbol{r}$ in the absence of object, is given by:

$$\mathcal{G}(\boldsymbol{r}, \boldsymbol{r}') = \frac{\mathrm{i}}{4} H_0^1(k_1 |\boldsymbol{r} - \boldsymbol{r}'|), \qquad [1.4]$$

where H_0^1 is the first kind Hankel function of order zero.

1.3.2. *The optical case*

As the object is embedded in a stratified medium, these expressions become more complicated because they must take into account the reflections on the interface γ_{12}. Thus, the incident field becomes:

$$E^{\text{inc}}(\boldsymbol{r}) = \exp(-\mathrm{i}k_1(x\cos\theta_1 + y\sin\theta_1)) + R\exp(\mathrm{i}k_1(x\cos\theta_1 - y\sin\theta_1))$$
$$R = \frac{k_1\cos\theta_1 - k_2\cos\theta_2}{k_1\cos\theta_1 + k_2\cos\theta_2}, \text{ with } \theta_2 \text{ such that } k_1\sin\theta_1 = k_2\sin\theta_2,$$

while the Green's function is expressed in the spectral domain associated with y, the axis parallel to γ_{12}, as a spectrum of plane waves with variable incidence [LES 91, SOU 96]:

$$\mathcal{G}(\boldsymbol{r}, \boldsymbol{r}') = \frac{1}{2\pi}\int_{-\infty}^{+\infty} g(x, x', \alpha)\exp\left(\mathrm{i}\alpha(y - y')\right)\mathrm{d}\alpha. \qquad [1.5]$$

Each plane wave is reflected or transmitted at the interface γ_{12} and finally the contributions of elementary plane waves are summed at the observation point. Here, accounting for the fact that the source ($\boldsymbol{r}'$) and observation ($\boldsymbol{r}$) points are both located in medium $\mathcal{D}_1$, the plane wave spectrum $g(x, x', \alpha)$ reads:

$$g(x, x', \alpha) = \frac{\mathrm{i}}{2\beta_1}\left(\exp(\mathrm{i}\beta_1 |x - x'|) + \frac{\beta_1 - \beta_2}{\beta_1 + \beta_2}\exp(\mathrm{i}\beta_1(x + x'))\right), \qquad [1.6]$$
$$\beta_m = \sqrt{k_m^2 - \alpha^2}\,, \qquad \Im m(\beta_m) \geq 0\,, \qquad m = 1, 2.$$

It consists of two terms: the first one represents the direct contribution, i.e. the spectral expansion of free space Green's function in medium $\mathcal{D}_1$ [1.4], and the second takes into consideration the reflections on the interface γ_{12}.

1.3.3. *The discrete model*

Assuming that both the contrast χ and the incident field E^{inc} are known, the resolution of the forward problem consists of first solving the coupling equation [1.1] for the sources w and then solving the observation equation [1.2] for the scattered field E^{dif}. This is done by using the discrete versions of these equations, obtained by applying the method of moments with pulse basis and Dirac delta testing functions [GIB 07]. That amounts to partition the test domain $\mathcal{D}$ into $N_{\mathcal{D}}$ elementary pixels small enough in order to consider the electric field and the contrast as constant over each of them. Rewriting equation [1.1] in terms of the induced sources leads to the two following linear systems:

$$w(\boldsymbol{r}_i) = \chi(\boldsymbol{r}_i)E^{\text{inc}}(\boldsymbol{r}_i) + \chi(\boldsymbol{r}_i)\sum_{j=1}^{N_{\mathcal{D}}} \boldsymbol{\mathcal{H}}_{ij}^{\mathcal{D}} w(\boldsymbol{r}_j), \quad i = 1, \ldots, N_{\mathcal{D}}, \qquad [1.7]$$

$$E^{\text{dif}}(\theta_n) = \sum_{j=1}^{N_{\mathcal{D}}} \boldsymbol{\mathcal{H}}_{nj}^{\mathcal{S}} w(\boldsymbol{r}_j), \quad n = 1, \ldots, N_r, \qquad [1.8]$$

where the matrix elements $\boldsymbol{\mathcal{H}}_{nj}^{\mathcal{S}}$ and $\boldsymbol{\mathcal{H}}_{ij}^{\mathcal{D}}$ result from the integration of the observation and coupling Green's functions, respectively, over the elementary pixels.

1.3.3.1. *The observation matrix*

The calculation of elements $\boldsymbol{\mathcal{H}}_{nj}^{\mathcal{S}}$ is relatively because in the microwave case, as well as in the optical case, an approximate analytical expression of $\boldsymbol{\mathcal{H}}_{nj}^{\mathcal{S}}$ can be obtained. Thus, in the microwave case, it is common to replace the integral over the square pixel by an integral over a disc of same surface [RIC 65] and this yields:

$$\boldsymbol{\mathcal{H}}_{nj}^{\mathcal{S}} = \frac{\mathrm{i}\pi\Delta k_1}{2}\, H_0^1(k_1\,|\boldsymbol{r}_n - \boldsymbol{r}_j|)\, J_1(k_1\Delta), \qquad [1.9]$$

where J_1 is the Bessel function of the first kind and $\Delta = 2a/\sqrt{\pi}$, a being the half-side of the pixel.

In the optical case, the scattered field being measured at points $\boldsymbol{r} = (r, \theta) \in \mathcal{S}$ located in the far field, with a fixed r and in directions θ such that $x > x'$, an approximate analytical expression $\mathcal{G}^o$ of $\mathcal{G}$ can be obtained in the spatial domain using the stationary phase method [CLE 66] by introducing the asymptotic expansion of H_0^1

for large arguments:

$$\begin{aligned}
\mathcal{G}^o(\theta, \boldsymbol{r}') &= \mathrm{i}\mathcal{Q}(\theta, x')\frac{\exp\left(-\mathrm{i}(\alpha(\theta)y' + \pi/4)\right)}{\sqrt{8\pi k_1}}, \\
\mathcal{Q}(\theta, x') &= \exp(-\mathrm{i}\beta_1(\theta)x') + \frac{\beta_1(\theta) - \beta_2(\theta)}{\beta_1(\theta) + \beta_2(\theta)}\,\exp(\mathrm{i}\beta_1(\theta)x'), \qquad [1.10] \\
\alpha(\theta) &= k_1\sin(\theta), \quad \beta_m(\theta) = \sqrt{k_m^2 - \alpha(\theta)^2} \quad \text{pour } m = 1, 2,
\end{aligned}$$

where a constant term $\exp(-\mathrm{i}k_1 r)/\sqrt{r}$ was omitted as it has been taken into account elsewhere in the normalization constants. By integration, this leads to the coefficients:

$$\boldsymbol{\mathcal{H}}^{\mathcal{S}}_{nj} \;=\; \mathrm{i}\sqrt{\frac{2k_1^3}{\pi}}\;\frac{\sin\left(\beta_1(\theta_n)a\right)\sin\left(\alpha(\theta_n)a\right)}{\alpha(\theta_n)\beta_1(\theta_n)}\;\mathcal{Q}(\theta_n,x_j)\exp\left(-\mathrm{i}\left(\alpha(\theta_n)y_j + \frac{\pi}{4}\right)\right).$$

1.3.3.2. *The coupling matrix*

Regarding the coupling matrix elements $\boldsymbol{\mathcal{H}}^{\mathcal{D}}_{ij}$, in the microwave case they are obtained in a similar way as those of the observation matrix, but with a different expression for the diagonal term obtained when source and observation are at the same location:

$$\boldsymbol{\mathcal{H}}^{\mathcal{D}}_{ij} = \begin{cases} \mathrm{i}\pi\Delta k_1 H_1^1(k_1\Delta)/2 - 1 & \text{if } i = j, \\ \mathrm{i}\pi\Delta k_1 H_0^1(k_1\,|\boldsymbol{r}_i - \boldsymbol{r}_j|)J_1(k_1\Delta)/2 & \text{otherwise.} \end{cases} \qquad [1.11]$$

In the optical case, things are a little bit more complicated because the Green's function $\mathcal{G}$ is known in the spectral domain and approximation [1.10] cannot be used for the coupling matrix elements because the observation point is located in the near field. However, knowing $\mathcal{G}$ in the spectral domain might not necessarily be disadvantageous because this suggests solving equation (1.7) by a CG-FFT type method (*Conjugate Gradient Fast Fourier Transform method* [SAR 86]), which significantly saves time by calculating in the spectral domain the convolution and correlation products that appear in this equation and by solving the system using a conjugated gradient method; such calculations are detailed in [LES 91] for a configuration similar to the one considered herein. It must be noted that the two terms of the Green's function [1.6] must be treated separately because the direct contribution presents a singularity at $\boldsymbol{r} = \boldsymbol{r}'$. The elements $\boldsymbol{\mathcal{H}}^{\mathcal{D}_s}_{ij}$ corresponding to the singular contribution are given by [1.11] and are known in the spatial domain. Regarding the

coefficients $\boldsymbol{\mathcal{H}}_{ij}^{\mathcal{D}_{\rm ns}}$ corresponding to the non-singular part of the Green's function, their spectral counterparts $\boldsymbol{h}_{ij}^{\mathcal{D}_{\rm ns}}$ read:

$$\boldsymbol{h}_{ij}^{\mathcal{D}_{\rm ns}} = \frac{2\mathrm{i}k_1^2 \sin(\alpha a)\sin(\beta_1 a)}{\alpha\beta_1^2} \frac{\beta_1-\beta_2}{\beta_1+\beta_2} \exp\left(\mathrm{i}\beta_1(x_i+x_j)\right). \quad [1.12]$$

1.3.4. *Validation of the forward model*

Figure 1.3 displays the results obtained in this way for the objects $\mathcal{M}$ and $\mathcal{O}$. The domain $\mathcal{D}$ is partitioned into $N_{\mathcal{D}}$ square pixels with half-side a and is illuminated at a frequency f and in direction θ_1 such that $N_{\mathcal{D}} = 64\times64$, $a = 1.17$ mm, $f = 4$ GHz and $\theta_1 = 40°$ for $\mathcal{M}$ and $N_{\mathcal{D}} = 512\times32$, $a = 3.7$ nm, $f = 473.6$ THz and $\theta_1 = -22.24°$

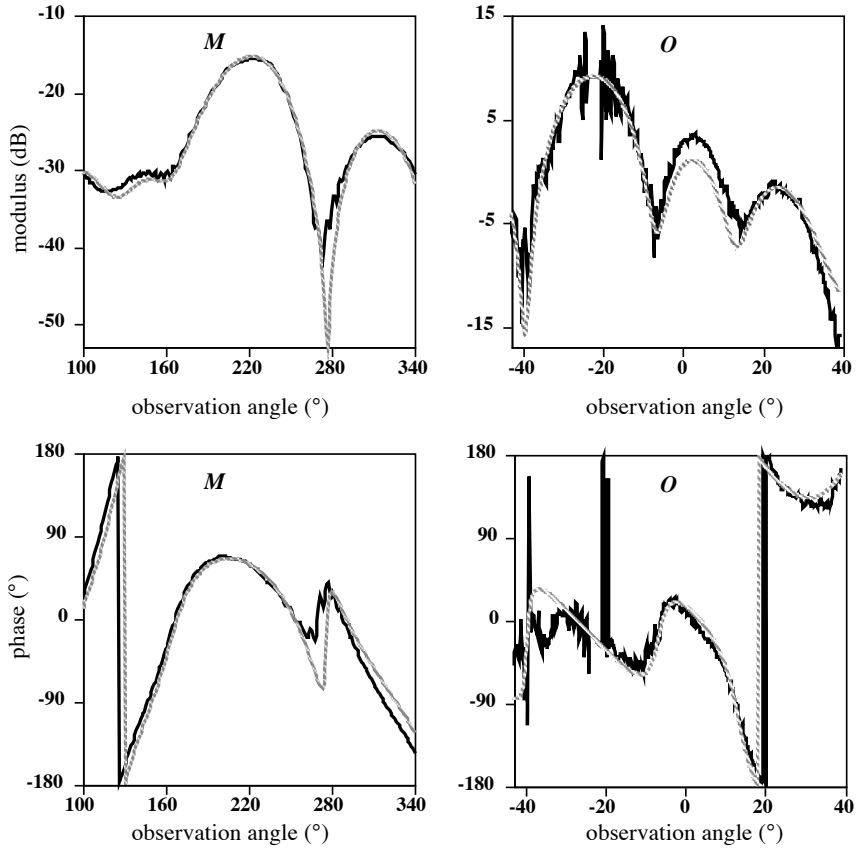

Figure 1.3. *The module (top) and the phase (bottom) of the computed (dotted gray line) and measured (continuous black line) scattered fields for the objects $\mathcal{M}$ (left) and $\mathcal{O}$ (right)*

for $\mathcal{O}$. Generally, the scattered fields are relatively well described. However, it must be noted that the fields measured with the optical device are very noisy, especially near the specular directions, and that data are missing in their immediate vicinity. This can be explained by the fact that the scattered field is negligible compared to the incident field. Therefore, accurate determination of the former is difficult to perform in those directions because it is obtained from the difference between the total field and the incident field.

1.4. Bayesian inversion approach

Now, let us focus on the inverse problem which consists in estimating the contrast $\boldsymbol{\chi}$ from measurements of the scattered field $\boldsymbol{E}^{\text{dif}}$, the incident field $\boldsymbol{E}^{\text{inc}}$ being known. Note that the induced sources $\boldsymbol{w}$, also unknown, should be estimated simultaneously with the contrast. First, let us define two vectors, $\boldsymbol{\epsilon}$ and $\boldsymbol{\xi}$, that represent all the errors, *i.e.,* the measurement uncertainties and the model errors due to the discretization and other approximations. Let us assume that these errors are i.i.d Gaussian, centered and with variances ρ_ϵ^2 and ρ_ξ^2, respectively. Accounting for the different views $v = 1, \ldots, N_v$ and the different frequencies (microwave case) $f = 1, \ldots, N_f$ using the index $u = 1, \ldots, N_v \times N_f$ and introducing the errors defined previously, the forward discrete model (equations (1.7) and (1.8)) can be rewritten, using matrix notation, as follows:

$$\boldsymbol{E}_u^{\text{dif}} = \boldsymbol{\mathcal{H}}_u^{\mathcal{S}} \boldsymbol{w}_u + \boldsymbol{\epsilon}_u, \qquad [1.13]$$

$$\boldsymbol{w}_u = \boldsymbol{\chi} \boldsymbol{E}_u^{\text{inc}} + \boldsymbol{\chi} \boldsymbol{\mathcal{H}}_f^{\mathcal{D}} \boldsymbol{w}_u + \boldsymbol{\xi}_u, \qquad [1.14]$$

where $\boldsymbol{E}_u^{\text{dif}}$, $\boldsymbol{E}_u^{\text{inc}}$ and $\boldsymbol{w}_u$ are the complex vectors that contain the scattered fields data, the incident fields and the induced sources corresponding to the different views and frequencies; $\boldsymbol{\chi}$ is a real vector that contains the contrast values at the pixel centers; $\boldsymbol{\mathcal{H}}_u^{\mathcal{S}}$ and $\boldsymbol{\mathcal{H}}_f^{\mathcal{D}}$ are operators mapping $L_2(\mathcal{D})$ into $L_2(\mathcal{S})$ and $L_2(\mathcal{D})$ into itself, respectively, that are represented by large matrices whose elements are $\boldsymbol{\mathcal{H}}_{nj}^{\mathcal{S}}$ and $\boldsymbol{\mathcal{H}}_{ij}^{\mathcal{D}}$, respectively.

As noted previously, it is now necessary to take into consideration the prior information available on the sought object. We know that the latter consists of a finite number N_k of different materials. This prior information is introduced through a hidden variable $z(\boldsymbol{r})$ associated with each pixel $\boldsymbol{r}$. This label defines the different material classes and the pixels that belong to a given class k can be characterized by a contrast that follows a Gaussian distribution:

$$p(\chi(\boldsymbol{r}) \,|\, z(\boldsymbol{r}) = k) = \mathcal{N}(m_k, \rho_k^2), \qquad k = 1, \ldots, N_k, \qquad [1.15]$$

with a mean value m_k and a variance ρ_k^2. Prior information that the different materials are distributed into compact regions is modeled by using a Markov-Potts field for $\boldsymbol{z}$

which describes the spatial dependence between neighboring pixels:

$$p(\boldsymbol{z} \,|\, \Upsilon) = \frac{1}{\boldsymbol{\Xi}} \exp\left(\frac{\Upsilon}{2} \sum_{i=1}^{N_{\mathcal{D}}} \sum_{\boldsymbol{r}_j \in \mathcal{V}(\boldsymbol{r}_i)} \delta\left[z(\boldsymbol{r}_i) - z(\boldsymbol{r}_j)\right] \right), \quad [1.16]$$

where $\boldsymbol{\Xi}$ is a normalization constant, $\delta(0) = 1$ and $\delta(t) = 0$ if $t \neq 0$ and Υ is a parameter that determines the degree of dependence or correlation between neighboring labels; herein $\Upsilon = 2$ and the probability $p(\boldsymbol{z} \,|\, \Upsilon)$ will be simply denoted as $p(\boldsymbol{z})$. $\mathcal{V}(\boldsymbol{r}_i)$ is a neighborhood of $\boldsymbol{r}_i$ which consists of the four closest pixels.

The probability distributions defined previously involve several different parameters, such as ρ_ϵ^2, ρ_ξ^2, m_k and ρ_k^2 ($k = 1, \ldots, N_k$), that, from now on, will be denoted as hyperparameters and gathered in a vector $\boldsymbol{\psi}$. On the other hand, the different means m_ϱ and variances ρ_ϱ^2 ($\varrho = \chi, w$ or k) are grouped in vectors $\boldsymbol{m}_\varrho = \{m_{\varrho(\boldsymbol{r}_i)},\, i = 1, 2, \ldots, N_{\mathcal{D}} \text{ or } N_k\}$) and $\boldsymbol{v}_\varrho = \{\rho^2_{\varrho(\boldsymbol{r}_i)},\, i = 1, 2, \ldots, N_{\mathcal{D}} \text{ or } N_k\}$), respectively, and in diagonal matrices $\mathrm{M}_\varrho = \mathrm{Diag}\,(\boldsymbol{m}_\varrho)$ and $\mathrm{V}_\varrho = \mathrm{Diag}\,(\boldsymbol{v}_\varrho)$ whose elements are the components of these vectors. It can be noted that an unsupervised approach is adopted: the contrast $\boldsymbol{\chi}$, the induced currents $\boldsymbol{w}$, the segmentation $\boldsymbol{z}$ and the hyperparameters of the model $\boldsymbol{\psi}$ are estimated together. By using the Bayesian rule, we get:

$$p(\boldsymbol{\chi},\boldsymbol{w},\boldsymbol{z},\boldsymbol{\psi} \,|\, \boldsymbol{E}^{\text{dif}}) \propto p(\boldsymbol{E}^{\text{dif}} \,|\, \boldsymbol{w},\boldsymbol{\psi})\, p(\boldsymbol{w} \,|\, \boldsymbol{\chi},\boldsymbol{\psi})\, p(\boldsymbol{\chi} \,|\, \boldsymbol{z},\boldsymbol{\psi})\, p(\boldsymbol{z})\, p(\boldsymbol{\psi}). \quad [1.17]$$

In this equation, $p(\boldsymbol{\chi} \,|\, \boldsymbol{z},\boldsymbol{\psi})$ and $p(\boldsymbol{z})$ are given by equations [1.15] and [1.16], whereas $p(\boldsymbol{E}^{\text{dif}} \,|\, \boldsymbol{w},\boldsymbol{\psi})$ and $p(\boldsymbol{w} \,|\, \boldsymbol{\chi},\boldsymbol{\psi})$ are obtained from the observation and coupling equations, respectively. These distributions read:

$$\begin{aligned} p(\boldsymbol{E}^{\text{dif}} \,|\, \boldsymbol{w},\boldsymbol{\psi}) &= \prod_u \left(\frac{1}{2\pi\rho_\epsilon^2} \right)^{N_r/2} \exp\left(-\frac{1}{2\rho_\epsilon^2} \left\| \boldsymbol{E}_u^{\text{dif}} - \boldsymbol{\mathcal{H}}_u^{\mathcal{S}} \boldsymbol{w}_u \right\|_{\mathcal{S}}^2 \right), \\ p(\boldsymbol{w}_u \,|\, \boldsymbol{\chi},\boldsymbol{\psi}) &\propto \exp\left(-\frac{1}{2\rho_\xi^2} \left\| \boldsymbol{w}_u - \boldsymbol{\chi}\, \boldsymbol{E}_u^{\text{inc}} - \boldsymbol{\chi}\, \boldsymbol{\mathcal{H}}_f^{\mathcal{D}}\, \boldsymbol{w}_u \right\|_{\mathcal{D}}^2 \right), \end{aligned} \quad [1.18]$$

where $\|\cdot\|_{\mathcal{A}}$ represents the norm associated with the scalar product $\langle \cdot,\cdot \rangle_{\mathcal{A}}$ in $L_2(\mathcal{A})$ ($\mathcal{A} = \mathcal{S}$ or $\mathcal{D}$). As for $p(\boldsymbol{\psi})$, it is a set of conjugate priors [BER 94]. This means that the variances and the means follow inverse-gamma ($\mathcal{IG}$) and Gaussian distributions, respectively:

$$\begin{aligned} p(\rho_\varrho^2) &= \mathcal{IG}(\eta_\varrho,\phi_\varrho) \propto \rho_\varrho^{-2(\eta_\varrho+1)} \exp\left(-\phi_\varrho/\rho_\varrho^2\right), \qquad \varrho = \epsilon, \xi, k \\ p(m_k) &= \mathcal{N}(\mu_k, \tau_k), \qquad k = 1, \ldots, N_k, \end{aligned} \quad [1.19]$$

with meta-hyperparameters $(\boldsymbol{\eta}, \boldsymbol{\phi}, \boldsymbol{\mu}, \tau)$ fixed in a way that satisfies a non-informative flat distribution.

All the terms on the right-hand side of equation [1.17] are known, and this allows us to obtain the left-hand side, *i.e.,* the unnormalized joint posterior distribution of all the unknowns. From this expression, different inferences can be made about these unknowns. The classic way is to define a point estimator, such as the maximum *a posteriori* (MAP) or the posterior mean (PM), but it is usually very difficult to get tractable expressions for such estimators. Therefore, an approximation of the posterior distribution is sought for, either numerically by using an MCMC sampling algorithm, or analytically by using the VBA method, which will be discussed in more detail thereafter.

1.4.1. *The MCMC sampling method*

Among the so-called MCMC sampling methods [ROB 04], one consists in drawing samples according to the conditional posterior distributions, which corresponds to a Gibbs sampling algorithm. Therefore, several of these conditional distributions must be determined. The distributions $p(\mathrm{V}_k \,|\, \boldsymbol{E}^{\text{dif}}, \boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{m}_k)$, $p(\boldsymbol{m}_k \,|\, \boldsymbol{E}^{\text{dif}}, \boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \mathrm{V}_k)$, $p(\rho_\epsilon^2 \,|\, \boldsymbol{E}^{\text{dif}}, \boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z})$ and $p(\rho_\xi^2 \,|\, \boldsymbol{E}^{\text{dif}}, \boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z})$ can be easily sampled because the choice of conjugate priors for hyperparameters allows them to stay in the same family, *i.e.,* Gaussian for the means and inverse-gamma for the variances:

$$\begin{aligned} p(\rho_\varrho^2 \,|\, \boldsymbol{E}^{\text{dif}}, \boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}) &\propto p(\boldsymbol{E}^{\text{dif}} \,|\, \boldsymbol{\chi}, \boldsymbol{w}, \rho_\varrho^2)\, p(\boldsymbol{w} \,|\, \boldsymbol{\chi}, \rho_\varrho^2)\, p(\rho_\varrho^2)\, p(\boldsymbol{\chi}, \boldsymbol{z}) \\ &\propto p(\rho_\varrho^2)\, p(\boldsymbol{\chi}, \boldsymbol{z}) \prod_u p(\boldsymbol{E}_u^{\text{dif}} \,|\, \boldsymbol{\chi}, \boldsymbol{w}_u, \rho_\varrho^2)\, p(\boldsymbol{w}_u \,|\, \boldsymbol{\chi}, \rho_\varrho^2) \\ &= \mathcal{IG}\left(\check{\eta}_\varrho, \check{\phi}_\varrho\right), \quad \varrho = (\epsilon, \xi), \end{aligned} \tag{1.20}$$

$$p(\mathrm{V}_k \mid \boldsymbol{E}^{\text{dif}}, \boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{m}_k) = p(\rho_k^2 \mid \boldsymbol{E}^{\text{dif}}, \boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{m}_k) = \mathcal{IG}(\check{\eta}_k, \check{\phi}_k), \qquad [1.21]$$

$$p(\boldsymbol{m}_k \mid \boldsymbol{E}^{\text{dif}}, \boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \mathrm{V}_k) = p(\boldsymbol{m}_k \mid \rho_k^2, \boldsymbol{E}^{\text{dif}}, \boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}) = \mathcal{N}(\check{\boldsymbol{\mu}}_k, \check{\tau}_k). \qquad [1.22]$$

By introducing the following notation:

$$R_k = \{\boldsymbol{r}; z(\boldsymbol{r}) = k\}, n_k = \mathrm{Card}(R_k), \overline{\chi}_k = \sum_{R_k} \frac{\boldsymbol{\chi}(r)}{n_k},\ s_k^2 = \sum_{R_k} \frac{(\boldsymbol{\chi}(r) - \boldsymbol{m}_k)^2}{n_k},$$

the tilded parameters in the conditional posterior distribution read:

$$\check{\mu}_k = \frac{n_k \overline{\chi}_k + \mu_k \rho_k^2/\tau_k}{n_k + \rho_k^2/\tau_k}, \qquad \check{\tau}_k = \frac{\rho_k^2}{n_k + \rho_k^2/\tau_k}, \qquad \check{\eta}_k = \eta_k + \frac{n_k}{2},$$

$$\check{\phi}_k = \phi_k + \frac{n_k\, s_k^2}{2}, \qquad \check{\eta}_\epsilon = \eta_\epsilon + \frac{N_u(N_r + N_{\mathcal{D}})}{2}$$

$$\check{\phi}_\epsilon = \phi_\epsilon + \frac{1}{2}\sum_u \left\| \boldsymbol{E}_u^{\text{dif}} - \boldsymbol{\mathcal{H}}_u^{\mathcal{S}} \boldsymbol{w}_u \right\|_{\mathcal{S}}^2 + \frac{1}{2\gamma}\sum_u \left\| \boldsymbol{\chi} \boldsymbol{E}_u^{\text{inc}} - \boldsymbol{w}_u + \boldsymbol{\chi} \boldsymbol{\mathcal{H}}_f^{\mathcal{D}} \boldsymbol{w}_u \right\|_{\mathcal{D}}^2$$

$$\check{\eta}_\xi = \eta_\xi + \frac{N_u N_{\mathcal{D}}}{2}, \qquad \check{\phi}_\xi = \phi_\xi + \frac{1}{2}\sum_u \left\| \boldsymbol{\chi} \boldsymbol{E}_u^{\text{inc}} - \boldsymbol{w}_u + \boldsymbol{\chi} \boldsymbol{\mathcal{H}}_f^{\mathcal{D}} \boldsymbol{w}_u \right\|_{\mathcal{D}}^2$$

with $\gamma = \rho_\xi^2/\rho_\epsilon^2$.

The posterior distribution $p(\boldsymbol{z} \mid \boldsymbol{E}^{\text{dif}}, \boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{\psi})$ of the segmentation is a Markov field with the same neighborhood as before (four pixels); sampling of this distribution can be done using a two step procedure[BES 74]. First of all, the pixels are divided up like a chessboard and the sets of white and black pixels are denoted as $\boldsymbol{z}_w$ and $\boldsymbol{z}_b$, respectively. Then, it can be noted that the four neighbors of each white pixel are black and *vice versa*. So, given the black pixels $\boldsymbol{z}_b$, the white pixels $\boldsymbol{z}_w$ are independent and can be sampled simultaneously and *vice versa*. The sampling of $p(\boldsymbol{z} \mid \boldsymbol{E}^{\text{dif}}, \boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{\psi})$ is then carried out using a Gibbs sampling algorithm, by alternatively drawing $\boldsymbol{z}_w$, given $\boldsymbol{z}_b$, and $\boldsymbol{z}_b$, given $\boldsymbol{z}_w$.

Regarding $\boldsymbol{w}_u$, its *a posteriori* distribution reads:

$$p(\boldsymbol{w}_u \mid \boldsymbol{E}_u^{\text{dif}}, \boldsymbol{\chi}, \boldsymbol{z}, \boldsymbol{\psi}) \propto p(\boldsymbol{E}_u^{\text{dif}} \mid \boldsymbol{w}_u, \rho_\epsilon^2)\, p(\boldsymbol{w}_u \mid \boldsymbol{\chi}, \rho_\epsilon^2) \propto \exp\left(-\frac{J_u(\boldsymbol{w}_u)}{2\rho_\epsilon^2} \right), \qquad [1.23]$$

$$\text{with } J_u(\boldsymbol{w}_u) = \left\| \boldsymbol{E}_u^{\text{dif}} - \boldsymbol{\mathcal{H}}_u^{\mathcal{S}} \boldsymbol{w}_u \right\|_{\mathcal{S}}^2 + \frac{1}{\gamma} \left\| \boldsymbol{\chi} \boldsymbol{E}_u^{\text{inc}} - \boldsymbol{w}_u + \boldsymbol{\chi} \boldsymbol{\mathcal{H}}_f^{\mathcal{D}} \boldsymbol{w}_u \right\|_{\mathcal{D}}^2.$$

Thus the posterior distribution of $\boldsymbol{w}_u$ is Gaussian. Its sampling requires the knowledge of its mean and covariance matrix. The calculation of the latter requires the inversion

of a large matrix. To avoid this, the sampling of this distribution is performed using a perturbed gradient method [ORI 12], which amounts to minimize the criterion J_u (see (1.23)).

The distribution $p(\boldsymbol{\chi} \,|\, \boldsymbol{E}^{\text{dif}}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\psi})$ can be obtained by using the same method:

$$p(\boldsymbol{\chi} \,|\, \boldsymbol{E}^{\text{dif}}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\psi}) \propto \exp\left(-\frac{1}{2\rho_\xi^2} \sum_u \|\boldsymbol{\chi}\, \boldsymbol{E}_u - \boldsymbol{w}_u\|_{\mathcal{D}}^2 - \frac{1}{2} \boldsymbol{B}(\boldsymbol{\chi}) \right), \qquad [1.24]$$

where $\boldsymbol{E}_u = \boldsymbol{E}_u^{\text{inc}} + \boldsymbol{\mathcal{H}}_f^{\mathcal{D}}\, \boldsymbol{w}_u$ and $\boldsymbol{B}(\boldsymbol{\chi}) = (\boldsymbol{\chi} - \boldsymbol{m}_\chi)^{\text{t}}\, \mathrm{V}_\chi^{-1}\, (\boldsymbol{\chi} - \boldsymbol{m}_\chi)$, while superscript t stands for the transposed vector.

1.4.2. *The VBA method*

The idea, here, is to approximate the posterior law $p(\boldsymbol{w},\boldsymbol{\chi},\boldsymbol{z},\boldsymbol{\psi} \,|\, \boldsymbol{E}^{\text{dif}})$ using a free-form separable distribution $q(\boldsymbol{w},\boldsymbol{\chi},\boldsymbol{z},\boldsymbol{\psi})$ that minimizes the Kullback-Leibler divergence $KL(q||p) = \int q \ln (q/p)$. Let us first define a separation form:

$$q(\boldsymbol{w},\boldsymbol{\chi},\boldsymbol{z},\boldsymbol{\psi}) = \prod_i q(w_i) \prod_j q(\chi_j) \prod_l q(z_l) \prod_k q(\rho_k^2) q(m_k) q(\rho_\epsilon^2) q(\rho_\xi^2). \qquad [1.25]$$

Then, the optimal form of q that minimizes the Kullback-Leibler divergence is sought for. With the chosen forward model (equations [1.13] and [1.14]), the Gaussian laws defined by [1.18] and the conjugated distributions [1.19] chosen for the hyperparameters, this approach leads to the following parametric distributions:

$$\begin{aligned} &q(\boldsymbol{w}) = \mathcal{N}(\tilde{\boldsymbol{m}}_w, \tilde{\mathrm{V}}_w), \quad q(\boldsymbol{\chi}) = \mathcal{N}(\tilde{\boldsymbol{m}}_\chi, \tilde{\mathrm{V}}_\chi), \quad q(\boldsymbol{z}) = \prod_{\boldsymbol{r}} \tilde{\zeta}_k(\boldsymbol{r})\,, \\ &q(m_k) = \mathcal{N}(\tilde{\mu}_k, \tilde{\boldsymbol{\tau}}_k), \quad q(\rho_k^2) = \mathcal{IG}(\tilde{\eta}_k, \tilde{\phi}_k), \quad k = 1, \ldots, N_k\,, \\ &q(\rho_\varrho^2) = \mathcal{IG}(\tilde{\eta}_\varrho, \tilde{\phi}_\varrho), \quad \varrho = \epsilon, \xi\,, \end{aligned} \qquad [1.26]$$

where the tilded parameters are mutually dependent and are calculated iteratively. The expressions of these parameters at iteration step n are detailed further on. It must be noted that, for the sake of clarity, superscript $(n-1)$ is omitted in the following, which means that the parameter values obtained at iteration $(n-1)$ are indicated without superscript. The following expressions are then obtained for the distributions of:

– the hidden field $\boldsymbol{z}$:

$$q(\boldsymbol{z}) \propto \prod_{\boldsymbol{r}} \exp\left(-\frac{1}{2} \tilde{\zeta}_k(\boldsymbol{r}) \right)\,, \qquad [1.27]$$

$$\tilde{\zeta}_k(\boldsymbol{r}) = \Psi(\tilde{\eta}_k) + \ln \tilde{\phi}_k + \overline{\rho_k^{-2}} \left((\tilde{m}_\chi(\boldsymbol{r}) - \tilde{\mu}_k)^2 + \tilde{\tau}_k + \tilde{\boldsymbol{v}}_\chi(\boldsymbol{r}) \right) - \Upsilon \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} \tilde{\zeta}_k(\boldsymbol{r}')\,,$$

where Ψ is the digamma function and overbar indicates the expectation of the variable with respect to q (*i.e.,* $\bar{a} = \mathrm{E}_q(a)$);

– the variance of the observation noise ρ_ϵ^2:

$$q(\rho_\epsilon^2) = \mathcal{IG}(\tilde{\eta}_\epsilon,\tilde{\phi}_\epsilon), \qquad \tilde{\phi}_\epsilon = \phi_\epsilon + N_r/2\,, \tag{1.28}$$
$$\tilde{\eta}_\epsilon = \eta_\epsilon + \frac{1}{2}\left(\|\boldsymbol{E}^{\text{dif}}\|_{\mathcal{S}}^2 + \left\|\mathcal{H}^{\mathcal{S}}\tilde{\boldsymbol{m}}_w\right\|_{\mathcal{S}}^2 - 2\Re e\left\{(\boldsymbol{E}^{\text{dif}})^\dagger \mathcal{H}^{\mathcal{S}}\tilde{\boldsymbol{m}}_w\right\} + \|\mathcal{H}^{\mathcal{S}^2}\tilde{\boldsymbol{v}}_w\|_1\right)\,,$$

where superscript † stands for the conjugate transpose, $\|\cdot\|_1$ denotes the L_1 norm and where the elements of $\mathcal{H}^{\mathcal{S}^2}$ are the squared elements of $\mathcal{H}^{\mathcal{S}}$;

– the variance of the coupling noise ρ_ξ^2:

$$q(\rho_\xi^2) = \mathcal{IG}(\tilde{\eta}_\xi,\tilde{\phi}_\xi), \qquad \tilde{\phi}_\xi = \phi_\xi + N_{\mathcal{D}}/2\,, \tag{1.29}$$
$$\tilde{\eta}_\xi = \eta_\xi + \frac{1}{2}\left(\|\tilde{\boldsymbol{m}}_w\|_{\mathcal{D}}^2 + \|\tilde{\mathrm{V}}_w\|_1 + \|(\tilde{\mathrm{M}}_\chi^2 + \tilde{\mathrm{V}}_\chi)\overline{\boldsymbol{E}^2}\|_1 - 2\Re e\left\{\tilde{\boldsymbol{m}}_\chi^\dagger \overline{\boldsymbol{w}\boldsymbol{E}^*}\right\}\right)\,,$$

where * denotes the conjugate complex and $\tilde{\mathrm{M}}_\chi^2 = \tilde{\mathrm{M}}_\chi^*\tilde{\mathrm{M}}_\chi$;

– the variances of the classes ρ_k^2:

$$q(\rho_k^2) = \mathcal{IG}(\tilde{\eta}_k,\tilde{\phi}_k), \quad k = 1,\ldots,N_k\,, \qquad \tilde{\phi}_k = \phi_k + \sum_{i=1}^{N_{\mathcal{D}}} \tilde{\zeta}_k(\boldsymbol{r}_i)/2\,, \tag{1.30}$$
$$\tilde{\eta}_k = \eta_k + \frac{1}{2}\sum_{i=1}^{N_{\mathcal{D}}} \tilde{\zeta}_k(\boldsymbol{r}_i)\left(\tilde{m}_\chi^2(\boldsymbol{r}_i) + \tilde{\boldsymbol{v}}_\chi(\boldsymbol{r}_i) + \tilde{\mu}_k^2 + \tilde{\tau}_k - 2\tilde{\mu}_k\tilde{m}_\chi(\boldsymbol{r}_i)\right)\,,$$

– the means of the classes m_k:

$$q(m_k) = \mathcal{N}(\tilde{\mu}_k,\tilde{\tau}_k), \quad k = 1,\ldots,N_k\,,$$
$$\tilde{\tau}_k = \left(\tau_k^{-1} + \overline{\rho_k^{-2}}\sum_{i=1}^{N_{\mathcal{D}}} \tilde{\zeta}_k(\boldsymbol{r}_i)\right)^{-1}, \quad \tilde{\mu}_k = \tilde{\tau}_k\left(\frac{\mu_k}{\tau_k} + \overline{\rho_k^{-2}}\sum_{i=1}^{N_{\mathcal{D}}} \tilde{\zeta}_k(\boldsymbol{r}_i)\tilde{m}_\chi(\boldsymbol{r}_i)\right), \tag{1.31}$$

– the induced current $\boldsymbol{w}$:

$$q(\boldsymbol{w}) = \prod_{i=1}^{N_{\mathcal{D}}} q(w(\boldsymbol{r}_i)) = \mathcal{N}(\tilde{\boldsymbol{m}}_w,\tilde{\mathrm{V}}_w), \qquad \tilde{\mathrm{V}}_w^n = \left(\text{Diag}\left(\overline{\rho_\epsilon^{-2}}\,\Gamma^{\mathcal{S}} + \overline{\rho_\xi^{-2}}\,\overline{\Gamma_\chi^{\mathcal{D}}}\right)\right)^{-1},$$
$$\tilde{\boldsymbol{m}}_w^n = \tilde{\boldsymbol{m}}_w + \left(\tilde{\mathrm{V}}_w^n\right)\left(\overline{\rho_\epsilon^{-2}}\mathcal{H}^{\mathcal{S}\dagger}\left(\boldsymbol{E}^{\text{dif}} - \mathcal{H}^{\mathcal{S}}\tilde{\boldsymbol{m}}_w\right) + \overline{\rho_\xi^{-2}}\Big(\tilde{\mathrm{M}}_\chi\boldsymbol{E}^{\text{inc}} + \tilde{\mathrm{M}}_\chi\mathcal{H}^{\mathcal{D}}\tilde{\boldsymbol{m}}_w\right.$$
$$\left. - \tilde{\boldsymbol{m}}_w - \mathcal{H}^{\mathcal{D}\dagger}\tilde{\mathrm{Y}}_\chi\boldsymbol{E}^{\text{inc}} + \mathcal{H}^{\mathcal{D}\dagger}\tilde{\mathrm{M}}_\chi^\dagger\tilde{\boldsymbol{m}}_w - \mathcal{H}^{\mathcal{D}\dagger}\tilde{\mathrm{Y}}_\chi\mathcal{H}^{\mathcal{D}}\tilde{\boldsymbol{m}}_w\Big)\right), \tag{1.32}$$

where $\tilde{\mathrm{Y}}_\chi = (\tilde{\mathrm{M}}_\chi^2 + \tilde{\mathrm{V}}_\chi)$ and $\Gamma^{\mathcal{S}}$ and $\overline{\Gamma_\chi^{\mathcal{D}}}$ are such that:

$$\Gamma^{\mathcal{S}}(\boldsymbol{r}_j) = \sum_{i=1}^{N_r} |\mathcal{H}_{ij}^{\mathcal{S}}|^2 \,,$$

$$\overline{\Gamma_\chi^{\mathcal{D}}}(\boldsymbol{r}_j) = 1 - 2\Re e\left\{\mathcal{H}_{jj}^{\mathcal{D}} \tilde{m}_\chi(\boldsymbol{r}_j)\right\} + (|\tilde{m}_\chi(\boldsymbol{r}_j)|^2 + \tilde{\boldsymbol{v}}_\chi(\boldsymbol{r}_j)) \sum_{i=1}^{N_{\mathcal{D}}} |\mathcal{H}_{ij}^{\mathcal{D}}|^2 \,;$$

– and the contrast $\boldsymbol{\chi}$:

$$q(\boldsymbol{\chi}) = \prod_{i=1}^{N_{\mathcal{D}}} q(\chi(\boldsymbol{r}_i)) = \mathcal{N}(\tilde{\boldsymbol{m}}_\chi, \tilde{\mathrm{V}}_\chi) \,, \qquad [1.33]$$

$$\tilde{\boldsymbol{m}}_\chi^n = (\tilde{\mathrm{V}}_\chi^n)\left(\sum_k \overline{\rho_k^{-2}} \tilde{\mu}_k \tilde{\boldsymbol{\zeta}}_k + \overline{\rho_\xi^{-2}}\, \overline{\boldsymbol{w}\boldsymbol{E}^*}\right), \qquad \tilde{\mathrm{V}}_\chi^n = \left(\overline{\rho_\xi^{-2}}\, \overline{E^2} + \overline{\mathrm{V}_\chi^{-1}}\right)^{-1} \,,$$

where $\overline{\boldsymbol{w}\boldsymbol{E}^*}$, the expectation of vector $\boldsymbol{w}\boldsymbol{E}^*$, is such that:

$$\overline{\boldsymbol{w}\boldsymbol{E}^*}(\boldsymbol{r}_i) = \sum_1^{N_u} E_u^{inc*}(\boldsymbol{r}_i)\tilde{m}_w(\boldsymbol{r}_i) + \tilde{m}_w(\boldsymbol{r}_i) \sum_{j=1}^{N_{\mathcal{D}}} \mathcal{H}_{ij}^{\mathcal{D}*} \tilde{m}_w^*(\boldsymbol{r}_j) + \mathcal{H}_{ii}^{\mathcal{D}*}\, \tilde{\boldsymbol{v}}_w(\boldsymbol{r}_i)$$

and $\overline{\mathrm{V}_\chi^{-1}}$ and $\overline{E^2}$ are diagonal matrices whose elements read:

$$\left(\overline{\mathrm{V}_\chi^{-1}}\right)_{ii} = \sum_k \tilde{\boldsymbol{\zeta}}_k(\boldsymbol{r}_i)\overline{\rho_k^{-2}}, \qquad \left(\overline{E^2}\right)_{ii} = \sum_1^{N_u} \left|E_u^{\text{inc}}(\boldsymbol{r}_i)\right|^2 + A \,,$$

$$A = 2\Re e\left\{\boldsymbol{E}^{\text{inc}\,*}(\boldsymbol{r}_i)\mathcal{H}_{ii}^{\mathcal{D}} \tilde{\boldsymbol{m}}_w(\boldsymbol{r}_i)\right\} + \left|\sum_{j=1}^{N_{\mathcal{D}}} \mathcal{H}_{ij}^{\mathcal{D}} \tilde{m}_w(\boldsymbol{r}_j)\right|^2 + \sum_{j=1}^{N_{\mathcal{D}}} \left|\mathcal{H}_{ij}^{\mathcal{D}}\right|^2 \tilde{\boldsymbol{v}}_w(\boldsymbol{r}_j) \,.$$

1.4.3. *Initialization, progress and convergence of the algorithms*

The two algorithms described above are iterative processes that require an initialization of the different variables ($\boldsymbol{\chi}^{(0)}$,$\boldsymbol{w}^{(0)}$,$\boldsymbol{z}^{(0)}$,$\boldsymbol{\psi}^{(0)}$). The initial estimation of the sources $\boldsymbol{w}^{(0)}$ is obtained by back-propagating the scattered field from the measurement domain $\mathcal{S}$ onto the test domain $\mathcal{D}$ [DUC 04]:

$$\boldsymbol{w}_u^{(0)} = \Gamma \boldsymbol{\mathcal{H}}_u^{\mathcal{S}\dagger} \boldsymbol{E}_u^{\text{dif}} \,, \qquad [1.34]$$

where Γ is a constant obtained by minimizing $\sum_u \left\| \boldsymbol{E}_u^{\text{dif}} - \Gamma \boldsymbol{\mathcal{H}}_u^{\mathcal{S}} \boldsymbol{\mathcal{H}}_u^{\mathcal{S}\dagger} \boldsymbol{E}_u^{\text{dif}} \right\|_{\mathcal{S}}^2$ and $\boldsymbol{\mathcal{H}}^{\mathcal{S}\dagger}$ is the operator adjoint to $\boldsymbol{\mathcal{H}}^{\mathcal{S}}$ that acts from $L_2(\mathcal{S})$ onto $L_2(\mathcal{D})$ and such that:

$$\left\langle \boldsymbol{w}_u, \boldsymbol{\mathcal{H}}_u^{\mathcal{S}\dagger} \boldsymbol{E}_u^{\text{dif}} \right\rangle_{\mathcal{D}} = \left\langle \boldsymbol{\mathcal{H}}_u^{\mathcal{S}} \boldsymbol{w}_u, \boldsymbol{E}_u^{\text{dif}} \right\rangle_{\mathcal{S}} .$$

The field $\boldsymbol{E}_u^{(0)}$ follows immediately *via* the coupling equation:

$$\boldsymbol{E}_u^{(0)} = \boldsymbol{E}_u^{\text{inc}} + \boldsymbol{\mathcal{H}}_f^{\mathcal{D}}\, \boldsymbol{w}_u^{(0)} \qquad [1.35]$$

and $\boldsymbol{\chi}^{(0)}$ is then obtained by minimizing the error in the induced source constitutive relationship $\boldsymbol{w} = \boldsymbol{\chi}\boldsymbol{E}$, or, more precisely, by minimizing the regularized criterion:

$$\sum_u \left(\left\| \boldsymbol{\chi}\, \boldsymbol{E}_u^{(0)} - \boldsymbol{w}_u^{(0)} \right\|_{\mathcal{D}}^2 + \sigma_u \left\| \boldsymbol{\chi} \right\|_{\mathcal{D}}^2 \right), \qquad [1.36]$$

where σ_u is a regularization constant empirically set at $\sigma_u = 0.1 \left| \boldsymbol{E}_u^{\text{inc}} \right|$ in the optical case and at zero in the microwave case. It must be noted that the presence of a regularization term is required in the optical case because arbitrarily small values of $\boldsymbol{E}_u^{(0)}$ can be obtained near the interface γ_{12}, which would lead to infinite initial values for the contrast in the absence of regularization. Taking into account that the contrast is real and positive, this leads to:

$$\boldsymbol{\chi}^{(0)} = \frac{\sum_u \Re e(\boldsymbol{w}_u^{(0)} \boldsymbol{E}_u^{\star(0)})}{\sum_u \left(\left| \boldsymbol{E}_u^{(0)} \right|^2 + \sigma_u \right)} \mathbb{1}_{\boldsymbol{\chi} \geq 0}\,, \qquad [1.37]$$

where $\mathbb{1}_{\boldsymbol{\chi} \geq 0} = 1$ if $\boldsymbol{\chi} \geq 0$ and $\mathbb{1}_{\boldsymbol{\chi} \geq 0} = 0$ if $\boldsymbol{\chi} < 0$. $\boldsymbol{\chi}^{(0)}$ and $\boldsymbol{w}^{(0)}$ being known, the classification $\boldsymbol{z}$ and the hyperparameters can be initialized by means of a segmentation method. The *K-means* algorithm [MAC 67] is used here with an empirical estimator for estimating the classes and their centers so that their variances would be minimal, given the number of classes.

Once initialized, the algorithms can then proceed as summarized in table 1.1. Starting from the contrast $\boldsymbol{\chi}^{(n-1)}$, the sources $\boldsymbol{w}^{(n-1)}$, the hidden field $\boldsymbol{z}^{(n-1)}$ and the hyperparameters $\boldsymbol{\psi}^{(n-1)}$ determined at iteration step $(n-1)$, the variables of the second column are either sampled according to the distributions indicated in the third column and given by the equations of the fourth column (MCMC algorithm), or updated from the shape parameters of the fifth column, calculated by means of the equations given in the sixth column (VBA algorithm).

Steps 1 to 7 are iterated until convergence is reached. This convergence can be evaluated by means of several methods, *e.g.,* an auto-correlation test for MCMC or the negative free energy for VBA. Herein, an empirical criterion based on the evolution of the hyperparameters is used. Convergence is illustrated in figure 1.4 through the evolution of some of the hyperparameters. For VBA, there is no lower limit to the number of iterations to be performed and, in the case considered herein,

		MCMC		VBA	
step	variable	sampling of	Eq.	updating	Eq.
1	$\widehat{\boldsymbol{z}}^{(n)}$	$p(\boldsymbol{z} \mid \boldsymbol{E}^{\text{dif}}, \widehat{\boldsymbol{\chi}}, \widehat{\boldsymbol{\psi}})$		$\tilde{\boldsymbol{\zeta}}_k^n$	[1.27]
2	$\widehat{\rho_\epsilon^2}^{(n)}$	$p(\rho_\epsilon^2 \mid \boldsymbol{E}^{\text{dif}}, \widehat{\boldsymbol{\chi}}, \widehat{\boldsymbol{w}}, \widehat{\boldsymbol{z}}^{(n)})$	(1.20)	$\tilde{\phi}_\epsilon^n$ and $\tilde{\eta}_\epsilon^n$	[1.28]
3	$\widehat{\rho_\xi^2}^{(n)}$	$p(\rho_\xi^2 \mid \boldsymbol{E}^{\text{dif}}, \widehat{\boldsymbol{\chi}}, \widehat{\boldsymbol{w}}, \widehat{\boldsymbol{z}}^{(n)})$	(1.20)	$\tilde{\phi}_\xi^n$ et $\tilde{\eta}_\xi^n$	[1.29]
4	$\widehat{\mathrm{V}}_k^{(n)}$	$p(\mathrm{V}_k \mid \boldsymbol{E}^{\text{dif}}, \widehat{\boldsymbol{\chi}}, \widehat{\boldsymbol{w}}, \widehat{\boldsymbol{z}}^{(n)})$	(1.21)	$\tilde{\phi}_k^n$ and $\tilde{\eta}_k^n$	[1.30]
5	$\widehat{\boldsymbol{m}}_k^{(n)}$	$p(\boldsymbol{m}_k \mid \boldsymbol{E}^{\text{dif}}, \widehat{\boldsymbol{\chi}}, \widehat{\boldsymbol{w}}, \widehat{\boldsymbol{z}}^{(n)}, \widehat{\mathrm{V}}_k^{(n)})$	(1.22)	$\tilde{\tau}_k^n$ et $\tilde{\mu}_k^n$	[1.31]
6	$\widehat{\boldsymbol{w}}^{(n)}$	$p(\boldsymbol{w} \mid \boldsymbol{E}^{\text{dif}}, \widehat{\boldsymbol{\chi}}, \widehat{\boldsymbol{z}}^{(n)}, \widehat{\boldsymbol{\psi}}^{(n)})$	(1.23)	$\tilde{\mathrm{V}}_w^n$ et $\tilde{\boldsymbol{m}}_w^n$	[1.32]
7	$\widehat{\boldsymbol{\chi}}^{(n)}$	$p(\boldsymbol{\chi} \mid \boldsymbol{E}^{\text{dif}}, \widehat{\boldsymbol{w}}^{(n)}, \widehat{\boldsymbol{z}}^{(n)}, \widehat{\boldsymbol{\psi}}^{(n)})$	(1.24)	$\tilde{\mathrm{V}}_{\boldsymbol{\chi}}^n$ and $\tilde{\boldsymbol{m}}_{\boldsymbol{\chi}}^n$	[1.33]

Table 1.1. *The MCMC and VBA algorithms and the reference equations*

the convergence is reached after almost a hundred iterations (see figure 1.4 (bottom)). However, with the MCMC algorithm, theoretically, many iterations must be carried out, first without keeping track of the samples until a certain convergence level is reached (which corresponds to the burn-in time of the Gibbs sampler) and, then, keeping track of them in order to compute their means and variances. However, in the present case, it can be observed that the variables and hyperparameters do not change significantly after approximately 250 iterations (see figure 1.4 (top)); thus, the maximum number of iterations has been set to 500 and the posterior mean is estimated with the last hundred samples.

1.5. Results

Figure 1.5 shows the results obtained with the two algorithms for the objects $\mathcal{M}$ and $\mathcal{O}$. In the microwave case, the domain $\mathcal{D}$ consists of 51×51 pixels with half side $a = 1.75$ mm, *i.e.,* $\mathcal{D} = 17.85 \times 17.85\,\text{cm}^2$, while, in the optical case, it is made of 32×512 pixels with half side $a = 3.7\,\text{nm}$, *i.e.,* $\mathcal{D} = 0.237 \times 3.789\ \mu\text{m}^2$. In all cases, the two algorithms succeed to reconstruct homogeneous regions that correspond to different materials with relatively accurate contrast values, as shown in Figure 1.5 (bottom) which displays the profiles reconstructed along the axis of symmetry of object $\mathcal{M}$, or at a height of $x = 0.1\,\mu$m for object $\mathcal{O}$, compared to the real profiles. The shape of the reconstructed objects is sometimes slightly different from the actual shape but the results are far more accurate than those obtained using deterministic methods such as CSI [ABU 05, AYA 11] or MGM [DUB 04, MAI 09].

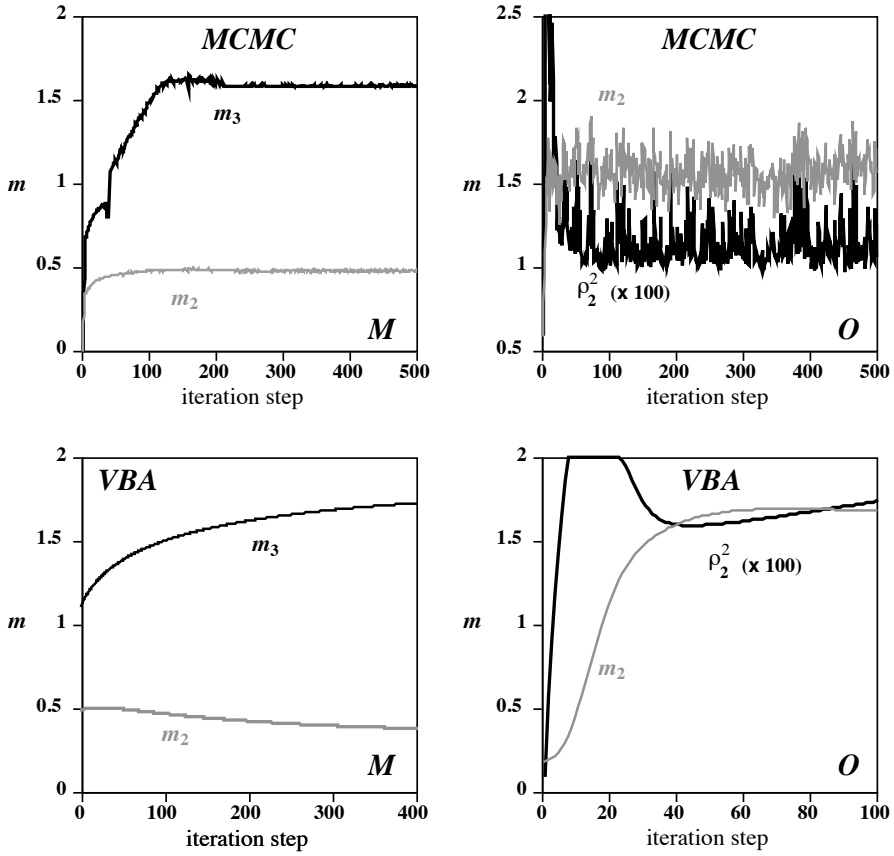

Figure 1.4. *The evolution of the hyperparameters for the MCMC (top) and VBA (bottom) algorithms applied to the microwave* [*left, mean of the class 2 (foam: gray line) and 3 (plastic: black line)*] *and optical* [*right, mean (gray line) and variance (black line) of class 2 (resin)*] *configurations.*

1.6. Conclusions

In this chapter, we addressed microwave and optical imaging as inverse scattering problems that are known as non-linear ill-posed problems. The objects considered here have a significant dielectric contrast which rules out the use of small perturbation approximations, such as the Born or Rytov ones, which linearize the inverse problem. The latter is a non-linear problem described by two coupled integral equations that link the measured scattered fields to the sources induced within the object by the incident wave, the induced sources and the contrast being then unknown.

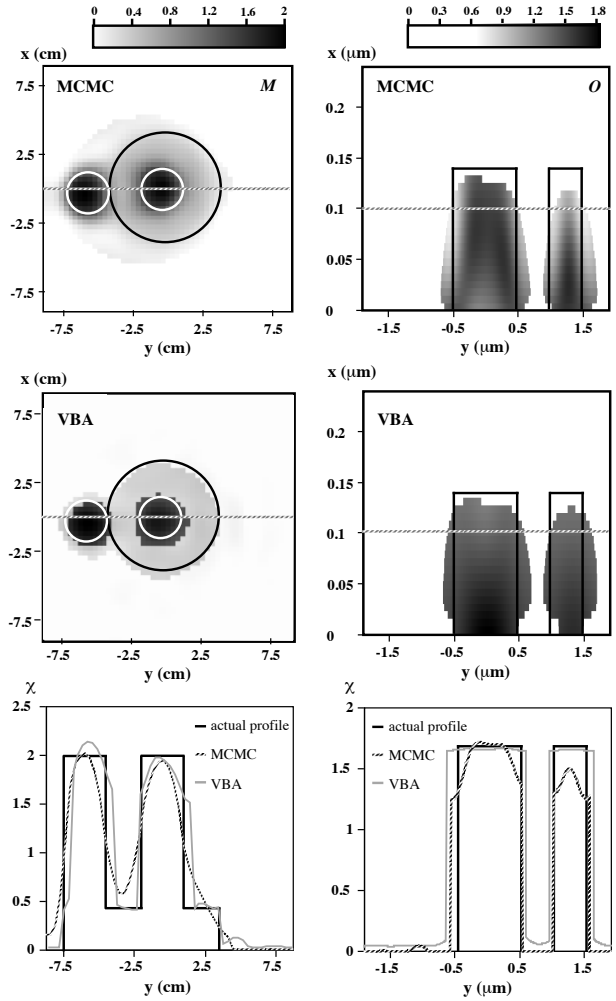

Figure 1.5. *The contrast of objects $\mathcal{M}$ (left) and $\mathcal{O}$ (right) obtained by means of the MCMC (top) and VBA (middle) algorithms and the profiles reconstructed along the dashed lines (bottom)*

As for the ill-posedness, the inverse problem must be regularized prior to its resolution and this regularization generally consists in introducing prior information about the sought solution. Such information is particularly necessary in the case of optical imaging, where the aspect-limited nature of data enhances the ill-posedness of the inverse problem. The latter consists in imaging objects composed of a finite number of different materials, which is an important prior information. This means that the unknown image is composed of a finite number of homogeneous regions.

This prior information is taken into account using a Gauss-Markov-Potts model for the contrast distribution, developed in a Bayesian estimation framework.

The Bayesian approach presented herein yields better results, both in microwave and optical imaging, than those obtained using deterministic iterative methods, such as CSI or MGM. Moreover, compared to the latter, it has the advantage of providing not only an estimation of the contrast distribution, but also a segmentation into regions and the values of the parameters (means, variances) of each class of materials. In some applications, this segmentation is even more important than the reconstruction itself.

Finally, two variants of this approach have been examined: the classic method (denoted as MCMC,) based on a Gibbs sampling algorithm, which results in a numerical approximation of the joint posterior of the unknowns, and the variational Bayesian approach (denoted as VBA) which allows an analytical approximation of the latter. Although these two variants lead to comparable results, the latter converges much faster than the former and this saved time is important when studying complicated configurations like the one considered in optical imaging where the object is embedded in a stratified medium, which requires a significant amount of calculations.

1.7. Bibliography

[ABU 05] ABUBAKAR A., VAN DEN BERG P. M., HABASHY T. M., "Application of the multiplicative regularized contrast source inversion method on TM- and TE-polarized experimental Fresnel data", *Inverse Problems*, vol. 21, no. 6, p. S65–S80, 2005.

[AYA 11] AYASSO H., DUCHÊNE B., MOHAMMAD-DJAFARI A., "Optical diffraction tomography within a variational Bayesian framework", *Inverse Problems in Science and Engineering*, 2011, DOI:10.1080/17415977.2011.624620.

[AZI 83] AZIMI M., KAK A. C., "Distorsion in diffraction tomography caused by multiple scattering", *IEEE Transactions on Medical Imaging*, vol. MI-2, p. 176-195, 1983.

[BEL 03] BELKEBIR K., SENTENAC A., "High-resolution optical diffraction microscopy", *Journal of the Optical Society of America (A)*, vol. 20, no. 7, p. 1223–1229, 2003.

[BER 94] BERNARDO J. M., SMITH A. F. M., *Bayesian Theory*, Wiley, Chichester, UK, 1994.

[BER 97] VAN DEN BERG P. M., KLEINMAN R. E., "A contrast source inversion method", *Inverse Problems*, vol. 13, p. 1607–1620, 1997.

[BES 74] BESAG J. E., "Spatial interaction and the statistical analysis of lattice systems (with discussion)", *Journal of the Royal Statistical Society B*, vol. 36, no. 2, p. 192-236, 1974.

[CHE 95] CHEW W., *Waves and Fields in Inhomogeneous Media*, IEEE Press, New York, 1995.

[CHO 02] CHOUDREY R., Variational methods for Bayesian independent component analysis, PhD thesis, University of Oxford, 2002.

[CLE 66] CLEMMOW P. C., *The Plane Wave Spectrum Representation of Electromagnetic Fields*, Pergamon Press, Oxford, 1966.

[COL 92] COLTON D., KRESS R., *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer Verlag, New York, 1992.

[DEV 84] DEVANEY A. J., BEYLKIN G., "Diffraction tomography using arbitrary source-receiver surfaces", *Ultrasonic Imaging*, vol. 6, p. 181-193, 1984.

[DUB 04] DUBOIS A., BELKEBIR K., SAILLARD M., "Retrieval of inhomogeneous targets from experimental frequency diversity data", *Inverse Problems*, vol. 21, no. 6, p. S65–S80, 2004.

[DUC 04] DUCHÊNE B., JOISEL A., LAMBERT M., "Nonlinear inversions of immersed objects using laboratory-controlled data", *Inverse Problems*, vol. 20, p. S81–98, 2004.

[FÉR 02] FÉRON O., MOHAMMAD-DJAFARI A., "Image fusion and joint segmentation using an MCMC algorithm", *Journal of Electronic Imaging*, vol. 14, no. 2, p. 1–12, 2002, ID: 023014.

[GEF 05] GEFFRIN J.-M., SABOUROUX P., EYRAUD C., "Free space experimental scattering database continuation: experimental set-up and measurement precision", *Inverse Problems*, vol. 21, no. 6, p. S117–S130, 2005.

[GIB 07] GIBSON W. C., *The Method of Moments in Electromagnetics*, Chapman & Hall/CRC, Boca Raton, 2007.

[HIN 93] HINTON G. E., VAN CAMP D., "Keeping the neural networks simple by minimizing the description length of the weights", 6^{th} *Annual Conference on Computational Learning Theory*, ACM New York, NY, USA, p. 5–13, 1993.

[IDI 13] IDIER J., *Bayesian approach to inverse problems*, John Wiley & Sons, 2013.

[JAA 00] JAAKKOLA T. S., JORDAN M. I., "Bayesian parameter estimation via variational methods", *Statistics and Computing*, vol. 10, no. 1, p. 25–37, 2000.

[JOA 91] JOACHIMOWICZ N., PICHOT C., HUGONIN J.-P., "Inverse scattering: An iterative numerical method for electromagnetic imaging", *IEEE Transactions on Antennas and Propagation*, vol. AP-39, no. 12, p. 1742-1752, 1991.

[JOR 99] JORDAN M. I., GHAHRAMANI Z., JAAKKOLA T. S., SAUL L. K., "An introduction to variational methods for graphical models", *Machine Learning*, vol. 37, no. 2, p. 183–233, 1999.

[KLE 92] KLEINMAN R. E., VAN DEN BERG P. M., "A modified gradient method for two-dimensional problems in tomography", *Journal of Computational and Applied Mathematics*, vol. 42, p. 17-35, 1992.

[KUL 51] KULLBACK S., LEIBLER R. A., "On information and sufficiency", *The Annals of Mathematical Statistics*, vol. 22, p. 79-86, 1951.

[LES 91] LESSELIER D., DUCHÊNE B., "Buried, 2-D penetrable objects illuminated by line-sources: FFT-based iterative computations of the anomalous field", SARKAR T. (DIR.), *Application of Conjugate Gradient Methods to Electromagnetics and Signal Analysis*, vol. 5 of *PIER* , p. 400–438, Elsevier, New York, 1991.

[LIK 04] LIKAS A. C., GALATSANOS N. P., "A variational approach for Bayesian blind image deconvolution", *IEEE Transactions on Signal Processing*, vol. SP-52, no. 8, p. 2222–2233, 2004.

[MAC 67] MACQUEEN J., "Some methods for classification and analysis of multivariate observations", *5th Berkeley Symposium on Mathematical Statistics and Probability*, Los Angeles, Western Management Sciences Institute UCLA, p. 281-297, 1967.

[MAI 09] MAIRE G., DRSEK F., GIRARD J., GIOVANNINI H., TALNEAU A., KONAN D., BELKEBIR K., CHAUMET P. C., SENTENAC A., "Experimental demonstration of quantitative imaging beyond Abbe's limit with optical diffraction tomography", *Physical Review Letters*, vol. 102, may 2009.

[MAI 10] MAIRE G., GIRARD J., DRSEK F., GIOVANNINI H., TALNEAU A., BELKEBIR K., CHAUMET P. C., SENTENAC A., "Experimental inversion of optical diffraction tomography data with a nonlinear algorithm in the multiple scattering regime", *Journal of Modern Optics, special issue on Digital Optical Microscopy*, vol. 57, no. 9, p. 746–755, 2010.

[MUE 80] MUELLER R., KAVEH M., INVERSON R., "A new approach to acoustic tomography using diffraction techniques", METHERELL A. (DIR.), *Acoustical Imaging*, vol. 8, p. 615–628, Plenum Press, New York, 1980.

[ORI 12] ORIEUX F., FÉRON O., GIOVANNELLI J.-F., "Sampling high-dimensional Gaussian distributions for general linear inverse problems", *IEEE Signal Processing Letters*, vol. 19, no. 5, p. 251–254, 2012.

[PIE 03] PIECZYNSKI W., "Modèles de markov en traitement d'images", *Traitement du Signal*, vol. 20, no. 3, p. 255-278, 2003.

[RIC 65] RICHMOND J., "Scattering by a dielectric cylinder of arbitrary cross-section shape", *IEEE Transactions on Antennas and Propagation*, vol. AP-13, no. 3, p. 334–341, 1965.

[ROB 04] ROBERT C. P., CASELLA G., *Monte Carlo Statistical Methods*, Springer Texts in Statistics, Springer Verlag, New York, 2004.

[SAR 86] SARKAR T. K., ARVAS E., RAO S. M., "Conjugate gradient method for the solution of electromagnetic radiation from electrically large and small conducting bodies", *IEEE Transactions on Antennas and Propagation*, vol. AP-34, no. 5, p. 635–640, 1986.

[SAT 01] SATO M., "Online model selection based on the variational Bayes", *Neural Computation*, vol. 13, no. 7, p. 1649–1681, 2001.

[SCH 92] SCHATZBERG A., DEVANEY A., "Super-resolution in diffraction tomography", *Inverse Problems*, vol. 8, p. 149–164, 1992.

[SLA 84] SLANEY M., KAK A. C., LARSEN L. E., "Limitations of imaging with first-order diffraction tomography", *IEEE Transactions on Microwave Theory and Techniques*, vol. 32, no. 8, p. 860-874, 1984.

[SMÍ 06] SMÍDL V., QUINN A., *The Variational Bayes Method in Signal Processing*, Springer Verlag, Berlin, 2006.

[SOU 96] SOURIAU L., DUCHÊNE B., LESSELIER D., KLEINMAN R. E., "A modified gradient approach to inverse scattering for binary objects in stratified media", *Inverse Problems*, vol. 12, no. 4, p. 463–481, 1996.

[TIE 94] TIERNEY L., "Markov chain for exploring posterior distribution", *Annals of Statistics*, vol. 22, no. 4, p. 1701-1762, 1994.