BULLETIN OF THE
GREEK MATHEMATICAL SOCIETY
Volume 58, 2010 - 2015 (43–66)

A Bayesian approach to the inverse source problem for the parabolic approximation to the Maxwell equations

I. G. Stratis and A. N. Yannacopoulos

Received 28/05/2015

Abstract

We consider a class of inverse source problems for the parabolic approximation to the Maxwell equations using a Bayesian approach.

Keywords: Maxwell equations, parabolic approximation, inverse source problems, Bayesian techniques.
(AMS) subject Classification: 35Q61, 35R30, 49J20, 62F15, 62P30.

0. Introduction

An important class of inverse problems, are the so called inverse source problems where by observation of the solution of the problem at some particular instants of time the determination of an unknown source is sought. Inverse source problems have been extensively studied for a variety of PDE problems and find interesting applications in a wide class of fields, see [1]. The inverse source problem for the time-harmonic Maxwell equations is considered in [2], where it is shown that the problem of finding a volume current density from surface measurements does not have a unique solution (while, with additional a priori information, uniqueness can be established). The inverse source problem for the eddy current approximation of the time-harmonic Maxwell equations is studied in [3]: nonuniqueness and uniqueness results are furnished, and the results are applied for the localisation of brain activity from electroencephalography and magnetoencephalography measurements. In [4] it is established that the eddy current model approximates the full Maxwell equations up to the second order with respect to frequency if and only if an additional condition is imposed on the current source (otherwise it is a first order approximation); further,

Research supported by the programme THALES, Code number 3570 (“Analysis, Modeling and Simulations of Complex and Stochastic Systems” University of Crete - Scientific coordinator: M. Katsoulakis) funded by the European Social Fund and National Hellenic Sources.

the well-posedness of the eddy current model is proved. Based on a new Carleman estimate, the inverse source problem for the nonstationary Maxwell equations is treated in [5]. Inverse source problems for the heat equation are recently studied by different approaches, e.g. by the enclosure method in [6], for a fractional diffusion equation in [7], for a one-dimensional evolution linear transport equation with spatially varying coefficients with an application to surface water pollution in [8]. Let us finally mention that the connection between inverse problems and controllability has a long history and a lot of a research activity has been and is performed on this connection.

It is the aim of this paper to present some comments on a class of inverse source problems related to the parabolic approximation to the Maxwell's equations. In Section 2, we present the parabolic approximation to the Maxwell's equation. In Section 3, we study a Bayesian approach to the inverse source problem. Finally, in Section 4, we make some comments on related problems for complex electromagnetic media. The novelty of the present work is twofold: (a) on the abstract level, to the best of our knowledge, this is the first treatment of the inverse source problem for a parabolic system using the Bayesian approach. There is of course extensive work on the problem of determining the initial data, using the Bayesian approach (see e.g, [9] and references therein, or [10]), which is a similar problem, but with distinctive differences in the compactness properties of the operators involved, a fact that is reflected on the permissible correlation operators for the prior and noise distribution, which in turn leads to distinctive differences on the regularization properties related to the Bayesian approach (see Theorem 2.1 for details). (b) As far as we know, the inverse source problem for the eddy current approximation to the Maxwell equations and the extensions to complex media (Section 4) have not been treated so far, let alone using the Bayesian techniques.

1. Parabolic approximation to the Maxwell equations

Consider a bounded, simply connected domain $\mathcal{O} \subset \mathbb{R}^3$, with sufficiently smooth boundary, filled with a (possibly complex) electromagnetic medium. The evolution of the electromagnetic field is governed by the Maxwell equations

$$
\begin{aligned}
&\mathrm{curl} H = \frac{\partial D}{\partial t} + J,\\
&\mathrm{curl} E = -\frac{\partial B}{\partial t},\\
&\mathrm{div} D = 0, \;\; \mathrm{div} B = 0,
\end{aligned}
$$

where D, B are the electric and magnetic flux densities, E, H are the electric and magnetic fields, J is the electric current density and we have assumed that the density of the electric charge is zero. This system is complemented with a constitutive law, modelling a (in general) complex electromagnetic medium, which is a linear (or nonlinear) relation connecting (D, B) with (E, H). The constitutive law can be given in terms of a nonlocal in time operator, involving convolutions which model various dispersion effects taking place in the medium. To fix ideas we consider the case of a dielectric, so that $D = \varepsilon E$ and $B = \mu H$, where ε is the electric conductivity and μ the magnetic permeability of the medium, expressed by matrix valued functions satisfying appropriate assumptions that are specified in the sequel. This system is complemented with the perfect conductor boundary condition $n \times E = 0$ on $\partial\mathcal{O}$, where n is the outer unit normal vector on the boundary of the domain. Finally, $J = \sigma E + J_0$, σ being the conductivity, where we have used the "standard" Ohm's law; J_0 is an externally imposed current, which - in Section 3 - will be used as a control procedure.

In the parabolic approximation to Maxwell equations we assume that the displacement current $\frac{\partial}{\partial t}(\varepsilon E)$ can be neglected, an assumption which is compatible with the assumption of slowly varying electromagnetic fields. Following [11] (see, also, [12]) this assumption yields

$$\begin{aligned} \mathrm{curl} H &= \sigma E + J_0, \\ \mathrm{curl} E &= -\frac{\partial}{\partial t}(\mu H) \end{aligned}$$

where upon elimination of the magnetic field H we obtain

$$\begin{aligned} &\frac{\partial}{\partial t}(\sigma E) + \mathrm{curl}(\mu^{-1}\mathrm{curl} E) = f, \text{ in } \mathcal{O}, \text{ where } f = -\frac{\partial J_0}{\partial t}, \\ &\mathrm{div}(\varepsilon E) = 0, \text{ in } \mathcal{O}, \\ &n \times E = 0 \text{ on } \partial\mathcal{O}, \\ &E(0, x) = E_0(x), \ x \in \mathcal{O}. \end{aligned} \tag{1}$$

The above problem (1) is called the parabolic approximation to Maxwell equations, endowed with the so called electric boundary conditions. A closely related system it the following,

$$\begin{aligned} &\frac{\partial}{\partial t}(\sigma E) + \mathrm{curl}(\mu^{-1}\mathrm{curl} E) = f, \text{ in } \mathcal{O}, \text{ where } f = -\frac{\partial J_0}{\partial t}, \\ &\mathrm{div}(\varepsilon E) = 0, \text{ in } \mathcal{O}, \\ &n \times (\mu^{-1}\mathrm{curl} E) = 0 \text{ on } \partial\mathcal{O}, \\ &n \cdot E = 0 \text{ on } \partial\mathcal{O}, \\ &E(0, x) = E_0(x), \ x \in \mathcal{O}. \end{aligned} \tag{2}$$

which is called the parabolic approximation to Maxwell equations, endowed with the so called magnetic boundary conditions. Both systems have similar properties from the mathematical point of view.

The parabolic approximation (1) or (2) has attracted the attention of the mathematical modelling community as a useful approximation to the full (hyperbolic)

Maxwell equation, and is sometimes referred to as the transient eddy current approximation. For a rigorous mathematical justification of the eddy current model in the low-frequency (time-harmonic) case see [13], where the related theory, algorithms and many applications are also presented; see, also, [4]. An important feature of this equation is that in certain cases, σ can vanish for certain parts of $\mathcal{O}$, thus turning the above system into an elliptic-parabolic system. The eddy current equations are of parabolic-elliptic type: in insulating regions, the field instantaneously adapts to the excitation (elliptic behaviour), while in conducting regions, this adaptation takes some time due to the induced eddy currents (parabolic behaviour). Since the main focus of this work is on issues related to controllability, we will assume throughout this paper that $\sigma \neq 0$, everywhere in $\mathcal{O}$, and in particular without loss of generality we will assume that $\sigma = 1$ (in appropriately scaled variables).

Let

$$H_0(\mathrm{curl}; \mathcal{O}) = \{u \in (L^2(\mathcal{O}))^3 \ : \ \mathrm{curl}\, u \in (L^2(\mathcal{O}))^3, \ n \times u = 0 \text{ on } \partial\mathcal{O}\},$$
$$H(\mathrm{div}(\varepsilon\, 0); \mathcal{O}) = \{u \in (L^2(\mathcal{O}))^3 \ : \ \mathrm{div}(\varepsilon u) = 0\},$$

and define the function spaces

$$\begin{aligned} \mathfrak{H} &:= H(\mathrm{div}(\varepsilon\, 0); \mathcal{O}) \\ \mathfrak{V} &:= H_0(\mathrm{curl}; \mathcal{O}) \cap H(\mathrm{div}(\varepsilon\, 0); \mathcal{O}). \end{aligned} \tag{3}$$

It is known that $\mathfrak{V}$ is dense in $\mathfrak{H}$.

Problem (1) can be expressed in abstract functional form, through the definition of the modified Maxwell operator M defined by $\mathsf{M}u = \mathrm{curl}(\mu^{-1}\mathrm{curl} u)$. The operator M is an unbounded operator with domain $D(\mathsf{M}) = \{u \in \mathfrak{V}, \ \mathrm{curl}\,\mathrm{curl} u \in \mathfrak{H}\}$. In terms of this operator the original problem assumes the abstract form

$$u' + \mathsf{M}u = f, \ \ u(0) = u_0 := E_0 \tag{4}$$

where $f = -\frac{\partial J_0}{\partial t}$, is considered as a function from $[0, T] \to \mathfrak{V}^*$. This is considered as an ODE in the function space $\mathfrak{H}$.

Throughout this work we make the following assumptions on the data:

Assumption 1.1 $\mu^{-1} \in W^{1,\infty}(\mathcal{O}; \mathbb{R}^{3\times 3})$ is a symmetric matrix valued function, and there exist constants κ_1, κ_2, $0 < \kappa_1 < \kappa_2$, such that

$$\kappa_1 |\xi|^2 \leq (\mu^{-1}\xi) \cdot \xi \leq \kappa_2 |\xi|^2, \ \ \forall \ \xi \in \mathbb{R}^3.$$

Assumption 1.2 $\varepsilon \in W^{1,\infty}(\mathcal{O}; \mathbb{R}^{3\times 3})$ is a symmetric matrix valued function, and there exist constants k_1, k_2, $0 < k_1 < k_2$, such that

$$k_1 |\xi|^2 \leq (\varepsilon\xi) \cdot \xi \leq k_2 |\xi|^2, \ \ \forall \ \xi \in \mathbb{R}^3.$$

Assumption 1.3 $J_0 \in H^1([0,T];\mathfrak{V}^*)$ and $E_0 \in \mathfrak{H}$.

It is known, [14], that under the above assumptions problem (4) has a unique solution u such that $u \in L^2([0,T];\mathfrak{V})$ and $u' \in L^2([0,T];\mathfrak{V}^*)$, $0 \le T < \infty$. Furthermore, by standard arguments (see e.g. [15]) $-\mathsf{M}$ is the generator of an analytic semigroup in $\mathfrak{H}$. We define by $\{S(t)\}_{t\in[0,T]}$ the analytic semigroup generated by $-\mathsf{M}$ and by $\{S^*(t)\}_{t\in[0,T]}$ its adjoint semigroup.

Comment 1.4 We further define the operators

$$\begin{aligned} L &= -\frac{\epsilon_1}{\epsilon_2}\int_0^T S(T-r)S^*(T-r)dr, \\ N &= S(T), \end{aligned}$$

and

$$F = \int_0^T S(T-r)\phi_p dr.$$

Assumption 1.5 The operator M is symmetric and positive definite with discrete and real spectrum and an eigensystem $\{\lambda_n, e_n\}$ such that $\lambda_n \to \infty$ as $n \to \infty$ and $\{e_n\}$ being an orthonormal basis of $\mathfrak{H}$. Furthermore, we assume that there exists an $\nu > 0$, which is the smallest possible number such that $\sum_n \lambda_n^{-\nu} < \infty$.

Remark 1.1 If μ is assumed symmetric and the domain $\mathcal{O}$ is bounded and has sufficiently smooth boundary, the above assumption holds for the operator $\mathsf{M} = \mathrm{curl}\left(\mu^{-1}\mathrm{curl}\right)$, see e.g., [16].

2. A Bayesian approach to inverse source problems

Measurements related to inverse problems are often subject to noise, thus leading to uncertain, or incomplete, knowledge of the measured data. This leads to the need of construction of inversion methods, that can handle the noisy data and provide trustworthy results concerning the quantities of interest. There is a well established literature concerning the treatment of general classes of such problems using the methods and techniques of Bayesian statistics (see, e.g., [17], [9] and references therein). Here, we sketch the application of this class of methods to a particular class of inverse source problems for the parabolic approximation of the Maxwell equation. For the appearing notions and results of Measure Theory on can consult [18] (see, also, [19]), while for elements of Operator and Spectral Theory see, e.g., [20].

Consider the inverse source problem: Given (u_0, u_T, j), such that $j(x) = J(0,x)$, find J, such that

$$u' + \mathsf{M}u = f \tag{5}$$

where $f = -\frac{\partial}{\partial t}(J)$ and $u(0) = u_0$, $u(T) = u_T$. Using the semigroup generated by $-\mathsf{M}$, $S(t) = e^{-\mathsf{M}t}$, we have that

$$u_T = S(T)u_0 + \int_0^T S(T-t)f(t)dt,$$

The data u_T and u_0 are subject to uncertainty, so we may assume that our true measurement is not u_T and u_0 but rather y_T and y_0, which is a noisy version of the true measurement, i.e., that $y_T = u_T + \xi_T$ and $y_0 = u_0 + \xi_0$, where ξ_0, ξ_T are appropriately chosen random variables taking values in appropriately chosen function spaces. Therefore, the model may be rearranged in terms of the actual measurements to yield

$$y_T - S(T)y_0 = \int_0^T S(T-t)f(t)dt + \eta, \tag{6}$$

where $\eta = \xi_T - S(T)\xi_0$ is a random variable taking values in the same function space as ξ_0, ξ_T. Furthermore, the distribution of η and be determined if the distribution of ξ_0, ξ_T is known.

Equation (6) is a random integral equation, the solution of which will yield the unknown function f and upon integration the unknown source J. This is an ill posed problem, even in the absence of noise. This can be written in the general operator form $y = \mathcal{G}f + \eta$, where $\mathcal{G}$ is a given bounded operator*, $y = y_T - S(T)y_0$ are the measurements and η is the noise term.

In what follows we will need the operator $\bar{\mathcal{G}} : H \to H$, where $H = \mathfrak{H}$, defined by

$$\bar{\mathcal{G}} = \int_0^T S(T-r)a(r)ds. \tag{7}$$

This is a bounded operator. If Assumption 1.5 holds, this operator can be defined through its action on any $w = \sum_n w_n e_n \in H$ as

$$\bar{\mathcal{G}}w = \sum_n g_n w_n e_n, \quad g_n = \int_0^T e^{-\lambda_n(T-t)}a(t)dt. \tag{8}$$

We may further assume that we look for sources of the special type $f(t) = a(t)w$ where $a(t)$ is a known scalar function of time and $w \in \mathfrak{H}$ is an unknown, spatially dependent only term, which is to be determined. This particular class of problems s of the general form studied by a number of authors, see e.g., [21] or [22]. For this particular class of problems the operator equation (6) yields

$$y_T - S(T)y_0 = \left(\int_0^T S(T-t)a(t)dt\right) w + \eta, \tag{9}$$

*In the particular model $\mathcal{G}f = \int_0^T S(T-t)f(t)dt$.

or in compact form,

$$y = \bar{\mathcal{G}} w + \eta, \tag{10}$$

where $\bar{\mathcal{G}}$ is given as in (7), or equivalently as (8) under Assumption 1.5. The solution of the inverse source problem for the special type of sources $f(t) = a(t)w$ where $a : [0, T] \to \mathbb{R}$ is a known function is equivalent to solving the random operator equation (10). However, this equation is ill posed, and problematic even in the case where noise is absent. The situation is made worse when noise and uncertainty are present.

To handle this situation we employ a Bayesian inversion technique in the spirit of [9]. In [9] an inverse problem for the recovery of the initial condition for the heat equation is treated in detail. This problem can be expressed in a very similar form as problem (10) but with the operator $N = S(T)$ in place of $\bar{\mathcal{G}}$. This problem is similar to the one treated here (as the parabolic approximation of the Maxwell equation shares a lot of common features with the heat equation), however there are important differences when treating the inverse source problem rather than the recovery of the initial condition problem which are related to the spectral behaviour of the operator $\bar{\mathcal{G}}$ as opposed to the spectral behaviour of the operator N. This leads to differences on the choice of priors as well as for the noise distribution. This comment will be made precise when stating and proving Theorem 2.1, below.

To fix ideas, and to illustrate the various problems arising with the operator equation (10), consider that $a(t)$ is a temporally decaying function, e.g., $a(t) = e^{-\mu t}$. Then, a straightforward calculation yields that

$$g_n = \frac{e^{-\mu T} - e^{-\lambda_n T}}{\lambda_n - \mu}.$$

These coefficients decay for large n as $\frac{e^{-\mu T}}{\lambda_n - \mu} \simeq \lambda_n^{-1}$, and in general present a polynomial decay in n (and not an exponential decay in n as is the case for the initial condition problem). Assuming that y and η have expansions in the basis $\{e_n\}$ of the form

$$y = \sum_n y_n e_n, \quad \eta = \sum_n \eta_n e_n,$$

allows us to rewrite the operator equation (10) as a sequence of scalar equations of the form

$$y_n = g_n w_n + \eta_n, \quad n \in \mathbb{N}.$$

The decay of the coefficients g_n turns this problem to an ill-posed one, which requires regularisation techniques, or a Bayesian approach. For a large class of problems of this form, a Bayesian approach using priors of the Gaussian type is equivalent to quadratic regularisation schemes (see e.g., [9]).

According to the Bayesian approach to inverse problems, since the operator equation we address is a random operator equation, the solution w is a random variable, taking values in a Hilbert space H. We are going to choose H as defined in (3). The solution to the inverse problem means characterising the probability distribution of the Hilbert space valued random variable w that satisfies (10). We assume that we have prior beliefs concerning w, which is given by a measure μ_0 defined on the Borel σ-algebra of H. This measure quantifies our initial belief concerning the solution to the problem. We then acquire data y concerning the problem and we update our a priori beliefs by taking the data into consideration. This will lead us to a new probablity measure for the random variable w, the posterior measure μ_y,which will allow us to estimate the solution. The determination of this posterior measure needs a model for the noise as well, i.e., for the distribution of the Hilbert space valued random variable η. We will denote the relevant probability measure by μ_N. The posterior measure μ_y will then be determined using Bayes formula from the measures μ_0, μ_N and the measurements y. To motivate this, assume momentarily that the state space is finite dimensional and that the probability measures μ_0, μ_N, μ_y, are absolutely continuous with respect to the Lebesgue measure having densities ρ_0, ρ_N, ρ_y respectively. Then the probabilty of y given w has density $\rho(y \mid w) = \rho_N(y - \bar{\mathcal{G}}w)$ and a straightforward application of Bayes formula yields that $\rho_y(w) = Z\rho_N(y - \bar{\mathcal{G}}w)\rho_0(w)$, where $Z \in \mathbb{R}$ is a normalisation factor. Using this posterior density, we may then make estimates for the solution of the inverse source problem. Clearly, this approach is not directly extendable to infinite dimensional Hilbert spaces, on account of the various intricacies related to measure theory in infinite dimensional spaces, but the basic elements of the theory are fortunately feasible in this context as well. We will not consider the general case, but rather consider the special case where the prior measure and the noise measure are Gaussian measures on H, a case which allows us to derive analytic results, and at the same time is reasonable from the modelling point of view, using arguments based on the central limit theorem.

We need the following definitions.

Definition 2.1 If Assumption 1.5 holds, the powers of the operator M are determined through the spectral decomposition of M, in the standard fashion

$$\mathsf{M}^{-\alpha} w = \sum_n \lambda_n^{-\alpha} w_n e_n, \;\; w = \sum_n w_n e_n.$$

Definition 2.2 We may define Sobolev type spaces

$$\mathcal{H}^s := \left\{ \sum_n u_n e_n \in H \,:\, \sum_n \lambda_n^s u_n^2 < \infty \right\}.$$

The space $\mathcal{H}^s$ can be understood as the domain of the operator $\mathsf{M}^{\frac{s}{2}}$.

We will assume that the measure μ_0 is a Gaussian measure on $(H, \mathcal{B}(H))$, with mean $m_0 \in H$ and covariance operator $\mathcal{C}_0 : H \to H$. The mean m_0 is an estimate of the solution to the inverse problem, before any measurement is taken. The covariance operator $\mathcal{C}_0$ is an indication of how credible our a priori knowledge is. We also assume that the noise is distributed according to a measure μ_N on $(H, \mathcal{B}(H))$, which is also a Gaussian measure on H, with mean $0 \in H$ and covariance operator $\mathcal{C}_N : H \to H$. It is assumed that η is independent of y. We will assume that $\mathcal{C}_0 = \beta \mathsf{M}^{-\alpha}$ and $\mathcal{C}_N = \delta \mathsf{M}^{-\gamma}$, for appropriate choice of $\alpha, \gamma > 0$. If Assumption 1.5 holds, and since M is a positive operator, we have that $\lambda_n \to \infty$ as $n \to \infty$, and also that M is a diagonalisable operator. For any $\alpha > 0$ and $\gamma > 0$ the eigenvalues of the operators $\mathcal{C}_0$ and $\mathcal{C}_N$ tend to 0 as $n \to \infty$, so $\mathcal{C}_0$ and $\mathcal{C}_N$ are compact operators, which have a smoothing effect on $u = \sum_n u_n e_n$, when considered as a function of the spatial variable. The covariance operators have to be self-adjoint, positive semi-definite and trace class. The first two properties are obvious since M has these properties as long as μ is symmetric. The third property requires proper choice of the powers α, β. Another notion which is important with Gaussian measures in Hilbert space, $\mu = \mathcal{N}(m, \mathcal{C})$, is the notion of the Cameron-Martin space (see, e.g., [9]), which is identified as $E = Im(\mathcal{C}^{\frac{1}{2}})$. The Cameron-Martin space may be identified as the subset of H, which is the intersection of all subspaces of H for which the Gaussian measure μ has full measure. As we will see the Cameron-Martin space will play an important role in treating the inverse source problem.

The following very simple and classical lemma (whose proof is only included to facilitate the reader), will be useful concerning the choice of smoothing parameter for the prior and for the characterisation of the Cameron-Martin space for the various Gaussian measures used.

Lemma 2.1 *Let Assumption 1.5 hold.*

(a) The choice $\mathcal{C} = A\,\mathsf{M}^{-s}$, where $s > 0$, and A is a constant, is a suitable choice for a covariance operator for a Gaussian measure if $s > \nu$.

(b) Consider the Gaussian measure $\mu \sim \mathcal{N}(m, \mathcal{C})$ on $(H, \mathcal{B}(H))$ where $\mathcal{C} = A\,M^{-s}$, $s > \nu$. The Cameron-Martin space $E = Im(\mathcal{C}^{\frac{1}{2}})$ for the Gaussian measure μ is identified by $\mathcal{H}^s = D(\mathsf{M}^{\frac{s}{2}})$.

Proof. (a) Since $\mathcal{C}$ is symmetric and positive semi-definite by the properties of M, it remains to check that $\mathcal{C}$ is a trace class operator. $\mathcal{C}$ s a diagonalisable operator with eigenvalues $\{\lambda_n^{-s}\}$, and therefore, $\mathcal{C}$ is trace class if $\sum_n \lambda_n^{-s} < \infty$, so $s > \nu$.

(b) We need to characterise $Im(\mathcal{C}^{\frac{1}{2}}) = Im(\mathsf{M}^{-\frac{s}{2}})$. This requires characterising all h such that there exists $y \in H$ for which $\mathsf{M}^{-\frac{s}{2}} y = h$. If $y = \sum_n y_n e_n$ such that $\sum_n |y_n|^2 < \infty$, then a straight forward calculation yields that $h = \sum_n h_n e_n$ is characterised by $h_n = \lambda_n^{-\frac{s}{2}} y_n$, so that

$$\infty > \sum_n |y_n|^2 = \sum_n \lambda_n^s |h_n|^2$$

and this means that $h \in \mathcal{H}^s$. This implies that $Im(\mathcal{C}^{\frac{1}{2}}) = D(\mathsf{M}^{\frac{s}{2}}) = \mathcal{H}^s$. □

By Lemma 2.1(a), we see that we must choose the measures μ_0,μ_N as

$$\begin{aligned} &\mu_0 \sim \mathcal{N}(m_0, \beta \mathsf{M}^{-\alpha}), \quad \alpha > \nu, \\ &\mu_N \sim \mathcal{N}(0, \delta \mathsf{M}^{-\gamma}), \quad \gamma > \nu. \end{aligned} \tag{11}$$

The basic result of this section is the following:

Theorem 2.1 *Let Assumption 1.5 hold and assume that the prior and noise distribution are as in* (11)*, with* α *and* γ *satisfying*

$$\begin{cases} \alpha > \nu \text{ and } \nu < \gamma < \dfrac{\alpha}{2} + 1 \text{ if } \nu < 1, \\ \alpha > 2\nu - 1 \text{ and } \nu < \gamma < \dfrac{\alpha}{2} + 1 \text{ if } 1 < \nu < 2, \\ \alpha > 2\nu - 1 \text{ and } \nu < \gamma < \dfrac{\alpha}{2} + 2 - \dfrac{\nu}{2} \text{ if } 2 < \nu < 3, \\ \alpha > 3\nu - 4 \text{ and } \nu < \gamma < \dfrac{\alpha}{2} + 2 - \dfrac{\nu}{2} \text{ if } 3 < \nu. \end{cases} \tag{12}$$

Then, the posterior distribution for w *is the Gaussian measure* $\mu_y = \mathcal{N}(m_y, \mathcal{C}_y)$ *where*

$$\begin{aligned} &m_y = m_0 + \varrho \mathsf{M}^{\gamma-\alpha} \bar{\mathcal{G}} \mathsf{C}(y - \bar{\mathcal{G}} m_0), \\ &\mathcal{C}_y = \beta\, \mathsf{M}^{-\alpha} \mathsf{C} \end{aligned}$$

where

$$\mathsf{C} := \left(I + \varrho \bar{\mathcal{G}}^2 \mathsf{M}^{\gamma-\alpha}\right)^{-1}, \quad \varrho = \frac{\beta}{\delta}.$$

Furthermore, the measures μ_0 *and* μ_y *are equivalent.*

Proof. A straightforward calculation shows that the correlation operator of (w, y) considered as an element of the Hilbert space $H \otimes H$ is the matrix operator

$$\mathcal{C} = \begin{pmatrix} \mathcal{C}_{11} & \mathcal{C}_{12} \\ \mathcal{C}_{21} & \mathcal{C}_{22} \end{pmatrix} := \begin{pmatrix} \mathbb{E}\left[(w - \mathbb{E}[w]) \otimes (w - \mathbb{E}[w])\right] & \mathbb{E}\left[(w - \mathbb{E}[w]) \otimes (y - \mathbb{E}[y])\right] \\ \mathbb{E}\left[(y - \mathbb{E}[y]) \otimes (w - \mathbb{E}[w])\right] & \mathbb{E}\left[(y - \mathbb{E}[y]) \otimes (y - \mathbb{E}[y])\right] \end{pmatrix}$$

where $\mathbb{E}[]$ is the expectation operator with respect to a probability measure μ on $(H, \mathsf{B}(H))$. Under the assumptions imposed here $\mathcal{C}_{11} = \mathcal{C}_0$, $\mathcal{C}_{12} = \mathcal{C}_{21} = \mathcal{C}_0 \bar{\mathcal{G}}$, $\mathcal{C}_{22} = \bar{\mathcal{G}} \mathcal{C}_0 \bar{\mathcal{G}} + \mathcal{C}_N$, where it must be taken into account that $\mathcal{C}_0$, $\mathcal{C}_N$ and $\bar{\mathcal{G}}$ commute. Then it can be shown that (see Theorem 6.20 in [9]) that the posterior measure μ_y is also a Gaussian measure on H, with mean and covariance

$$\begin{aligned} m_y &= m_0 + \mathcal{C}_{12}\mathcal{C}_{22}^{-1}(y - \bar{\mathcal{G}}m_0), \\ \mathcal{C}_y &= \mathcal{C}_{11} - \mathcal{C}_{12}\mathcal{C}_{22}^{-1}\mathcal{C}_{21}. \end{aligned}$$

By the commutation properties of the operators $\mathcal{C}_0$, $\mathcal{C}_N$ and $\bar{\mathcal{G}}$, it can be seen that

$$\mathcal{C}_{22}^{-1} = \frac{1}{\delta}\mathsf{M}^{\gamma}\left(I + \frac{\beta}{\delta}\bar{\mathcal{G}}^2\mathsf{M}^{\gamma-\alpha}\right)^{-1} \tag{13}$$

and this provides the expression for m_y. Substituting (13) in the expression for $\mathcal{C}_y$, and using the commutativity property we obtain

$$\mathcal{C}_y = \beta\mathsf{M}^{-\alpha}\left\{I - \frac{\beta}{\delta}\bar{\mathcal{G}}^2\mathsf{M}^{\gamma-\alpha}\left(I + \frac{\beta}{\delta}\bar{\mathcal{G}}^2\mathsf{M}^{\gamma-\alpha}\right)^{-1}\right\}$$

and using the operator identity

$$I - K(I+K)^{-1} = (I+K)^{-1}, \tag{14}$$

for the choice $K = \frac{\beta}{\delta}\bar{\mathcal{G}}^2\mathsf{M}^{\gamma-\alpha}$, we obtain the stated result for $\mathcal{C}_y$.

We need to check that $\mathcal{C}_y$ is a proper covariance operator. Evidently, it is symmetric and positive definite, so it remains to check that $\mathcal{C}_y$ is a trace class operator. The operator C is diagonalisable in the basis of H that consists of the eigenfunctions of M and can be expressed through its action on any $h = \sum_n h_n e_n$ by the formula

$$\mathsf{C}h = \sum_n c_n h_n e_n, \;\; c_n := \frac{1}{1 + \varrho\, g_n^2 \lambda_n^{\gamma-\alpha}}, \tag{15}$$

where $\varrho := \frac{\beta}{\delta}$, and as a consequence $\mathcal{C}_y$ is also a diagonalisable operator, with eigenvalues $\mu_n = \beta\lambda_n^{-\alpha}c_n$. Since $\lambda_n > 0$ for all n, clearly $0 < c_n < 1$ for all n, therefore $\mu_n < \beta\,\lambda_n^{-\alpha}$, therefore the $\sum_n \mu_n \leq \beta\sum_n \lambda_n^{-\alpha} < \infty$ since we have chosen $\alpha > \nu$. Therefore, $\mathcal{C}_y$ is a trace class operator.

It remains to show the equivalence of μ_y and μ_0. Recall (see e.g. Theorem 6.13 in [9]) that two Gaussian measures $\mu_1 = \mathcal{N}(m_1, \mathcal{C}_1)$ and $\mu_2 = \mathcal{N}(m_2, \mathcal{C}_2)$ are equivalent if and only if (a) $Im(\mathcal{C}_1^{\frac{1}{2}}) = Im(\mathcal{C}_2^{\frac{1}{2}})$, (b) $m_1 - m_2 \in \overline{Im(\mathcal{C}_1^{\frac{1}{2}})}$ and (c) the operator $T = (\mathcal{C}_1^{-\frac{1}{2}}\mathcal{C}_2^{\frac{1}{2}})(\mathcal{C}_1^{-\frac{1}{2}}\mathcal{C}_2^{\frac{1}{2}})^* - I$ is Hilbert-Schmidt in $\overline{Im(\mathcal{C}_1^{\frac{1}{2}})}$.

We will apply this theorem for $\mathcal{C}_1 = \mathcal{C}_0$ and $\mathcal{C}_2 = \mathcal{C}_y$. Recall the representation (15) for the operator C and consider the asymptotic behaviour of c_n. For large n, $g_n \sim \lambda_n^{-1}$ therefore $c_n \sim \frac{1}{1+\varrho\lambda_n^{\gamma-\alpha-2}}$ where $\varrho = \frac{\beta}{\delta}$. Since $\lambda_n > 0$, and $\lambda_n \to \infty$, if $\gamma - \alpha - 2 < 0$ then $\sup_n\left(\varrho g_n^2\lambda_n^{\gamma-\alpha}\right) =: k < \infty$ so that $\inf_n c_n \geq \frac{1}{1+k}$ and as a consequence of that the operator C is bounded and invertible, with all eigenvalues in $\left[\frac{1}{1+k}, 1\right]$. Imposing the condition

$$\gamma - \alpha - 2 < 0, \tag{16}$$

by the definition of $\mathcal{C}_y$ in the statement of the proposition,

$$\frac{1}{1+k}\langle h, \mathcal{C}_0 y\rangle \leq \langle h, \mathcal{C}_y y\rangle \leq \langle h, \mathcal{C}_0 h\rangle$$

Using a standard result from the theory of positive definite and bounded linear operators, according to which $Im(\mathcal{C}_1^{\frac{1}{2}}) \subset Im(\mathcal{C}_2^{\frac{1}{2}})$ if and only if there exists a constant $c > 0$ such that $\langle h, \mathcal{C}_1 h\rangle \leq c\langle h, \mathcal{C}_2 h\rangle$ for every $h \in H$, applied twice for $\mathcal{C}_1 = \mathcal{C}_0$, $\mathcal{C}_2 = \mathcal{C}_y$ and $\mathcal{C}_1 = \mathcal{C}_y$, $\mathcal{C}_2 = \mathcal{C}_0$, we obtain that $Im(\mathcal{C}_0^{\frac{1}{2}}) = Im(\mathcal{C}_y^{\frac{1}{2}}) = \mathcal{H}^\alpha$. This shows that condition (a) is satisfied.

We now check condition (b). For that we first need to characterise $Im(\mathcal{C}_0^{\frac{1}{2}}) = Im(\mathsf{M}^{-\frac{\alpha}{2}})$. By Lemma 2.1(b) $Im(\mathcal{C}_0^{\frac{1}{2}}) = \mathcal{H}^\alpha$. We now check whether $m_y - m_0 \in \mathcal{H}^\alpha$ for all $m_0, y \in H$. From the statement of the proposition it is seen that $m_y - m_0 = \varrho \mathsf{M}^{\gamma-\alpha}\bar{\mathcal{G}}\mathsf{C}(y - \bar{\mathcal{G}}m_0)$ so it remains to check whether $\mathsf{M}^{\gamma-\alpha}\bar{\mathcal{G}}\mathsf{C}(y - \bar{\mathcal{G}}m_0) \in \mathcal{H}^\alpha$ for every $y, m_0 \in H$. We first deal with the first part of this term, $\mathsf{M}^{\gamma-\alpha}\bar{\mathcal{G}}\mathsf{C}y$: Consider any $y \in H$, so that $y = \sum_n y_n e_n$, and $\sum_n |y_n|^2 < \infty$. Then,

$$q := \mathsf{M}^{\gamma-\alpha}\bar{\mathcal{G}}\mathsf{C}y = \sum_n \lambda_n^{\gamma-\alpha} g_n c_n y_n e_n,$$

and $q \in \mathcal{H}^\alpha$ if

$$Q := \sum_n \lambda^\alpha \lambda_n^{2(\gamma-\alpha)} g_n^2 c_n^2 |y_n|^2 < \infty.$$

Since $c_n < 1$ and $g_n \sim \lambda_n^{-1}$ we see that

$$Q := \sum_n \lambda^\alpha \lambda_n^{2(\gamma-\alpha)} g_n^2 c_n^2 |y_n|^2 < C_1 \sum_n \lambda_n^{2\gamma-\alpha-2} |y_n|^2,$$

for an appropriately chosen constant C_1. Since $\lambda_n \to \infty$ as $n \to \infty$, we see that if $2\gamma - \alpha - 2 < 0$, then there exists a constant C_2 such that

$$Q < C_1 \sum_n |y_n|^2 < \infty,$$

where the last inequality follows since we have assumed $y \in H$. This leads us to a second condition on α and γ,

$$2\gamma - \alpha - 2 < 0. \tag{17}$$

The second contribution to $m_y - m_0$ is $p := \bar{\mathcal{G}}\mathsf{M}^{\gamma-\alpha}\mathsf{C}\bar{\mathcal{G}}m_0$. To see whether $p \in \mathcal{H}^\alpha$ for every $m_0 \in H$, take $m_0 = \sum_n m_{0,n}e_n$, $\sum_n |m_{0,n}|^2 < \infty$, and calculate

$$p = \sum_n \lambda_n^{\gamma-\alpha} g_n^2 c_n m_{0,n} e_n,$$

so that $p \in \mathcal{H}^\alpha$ if

$$\sum_n \lambda_n^\alpha \lambda_n^{2(\gamma-\alpha)} g_n^4 c_n^2 |m_{0,n}|^2 < \infty.$$

Similar arguments as above, taking into account that $c_n < 1$ and $g_n \sim \lambda_n^{-1}$, show that if $2\gamma - \alpha - 4 < 0$ then $p \in \mathcal{H}^\alpha$. But clearly this condition holds as long as condition (17) holds, so we do not need to impose any further restrictions on α and γ to account for the second term. This concludes the verification of condition (b).

We now check condition (c). Since the operators M and its powers commute with the operator C we see that $T = (I + \varrho\bar{\mathcal{G}}^2\mathsf{M}^{\gamma-\alpha})^{-1} - I$ and it is immediate to check, using the operator identity (14), that the operator T assumes the form

$$T = -\varrho\mathsf{C}\mathsf{M}^{\gamma-\alpha}\bar{\mathcal{G}}^2,$$

whose action on the basis $\{e_n\}$ is

$$Te_n = t_n e_n, \;\; t_n = -\varrho\lambda_n^{\gamma-\alpha} c_n g_n^2.$$

The operator $T : \mathcal{H}^\alpha \to \mathcal{H}^\alpha$ is a a Hilbert-Schmidt operator[†] if $\sum_n ||Te_n||^2_{\mathcal{H}^\alpha} < \infty$. Recalling the definition of the norm of $\mathcal{H}^\alpha$, the condition for T defined as above, to be a Hilbert-Schmidt operator becomes

$$P := \sum_n \lambda_n^\alpha t_n^2 = \sum_n \lambda_n^\alpha \lambda_n^{2(\gamma-\alpha)} c_n^2 g_n^4 < \infty,$$

which since $c_n < 1$ and $g_n \sim \lambda_n^{-1}$ holds as long as the condition

$$\sum_n \lambda_n^{2\gamma-\alpha-4} < \infty,$$

holds. Since $\lambda_n \to \infty$ as $n \to \infty$ this condition can hold only if $2\gamma-\alpha-4$ is sufficiently negative. Condition (17) guarantees that $2\gamma - \alpha - 4 < 0$, but this may not be enough for the convergence of the series related to the Hilbert-Schmidt norm of T. The exact value of the exponent $2\gamma - \alpha - 4$ depends on the asymptotics of the spectrum of the operator M on the chosen domain. Let $\nu > 0$ be the smallest real number such that $\sum_n \lambda_n^{-\nu} < \infty$. Then, the operator $T : \mathcal{H}^\alpha \to \mathcal{H}^\alpha$ is Hilbert-Schmidt if

$$2\gamma - \alpha - 4 < -\nu. \tag{18}$$

[†]Recall, from the theory of Hilbert-Schmidt operators, that this condition is independent of the choice of basis.

The exact value of ν depends on the asymptotic behaviour of the spectrum of M, if $\lambda_n \sim n^{\zeta}$ as $n \to \infty$, then ν has to be chosen so that $-\nu\zeta < -1$, i.e., $\nu > \frac{1}{\zeta}$. Condition (17) implies that $2\gamma - \alpha - 4 < -2$, if $-2 < -\nu$, i.e. $\nu > 2$, then condition (17) implies condition (18).

We complete the proof by clarifying the conditions on α and γ, under which our analysis holds. Conditions (16), (17) and (18) must hold simultaneously, and by Lemma 2.1 it must also hold that $\gamma > \nu$. Fixing $\alpha > \nu$, we solve them in terms of γ to obtain that it must simultaneously hold that

$$\nu < \gamma < \alpha + 2, \ \ \nu < \gamma < \frac{\alpha}{2} + 1, \ \ \nu < \gamma < \frac{\alpha}{2} + 2 - \frac{\nu}{2}. \tag{19}$$

Clearly, these may hold if α is chosen so that

$$\nu < \alpha + 2, \ \ \nu < \frac{\alpha}{2} + 1, \ \ \nu < \frac{\alpha}{2} + 2 - \frac{\nu}{2},$$

i.e., as long as α is such that

$$\alpha > 3\nu - 4, \ \ \alpha > \nu, \ \ \alpha > 2\nu - 1. \tag{20}$$

If $\nu < 1$ then $\nu = \max\{3\nu - 4, \nu, 2\nu - 1\}$ and (20) hold as long as $\alpha > \nu$. If $1 < \nu < 3$ then $2\nu - 1 = \max\{3\nu - 4, \nu, 2\nu - 1\}$ and (20) hold as long as $\alpha > 2\nu - 1$. Finally, if $\nu > 3$, then $3\nu - 4 = \max\{3\nu - 4, \nu, 2\nu - 1\}$ and (20) hold as long as $\alpha > 3\nu - 4$. Summarising, α should be chosen as

$$\begin{cases} a > \nu \ \text{ if } \ \nu < 1, \\ a > 2\nu - 1 \ \text{ if } \ 1 < \nu < 3, \\ a > 3\nu - 4 \ \text{ if } \ \nu > 3. \end{cases} \tag{21}$$

We now consider (19), in terms of γ. If $\nu < 2$, then $\frac{\alpha}{2} + 1 = \min\{\alpha + 2, \frac{\alpha}{2} + 1, \frac{\alpha}{2} + 2 - \frac{\nu}{2}\}$, so (19) hold if $\nu < \gamma < \frac{\alpha}{2} + 1$. If $\nu > 2$, then $\frac{\alpha}{2} + 2 - \frac{\nu}{2} = \min\{\alpha + 2, \frac{\alpha}{2} + 1, \frac{\alpha}{2} + 2 - \frac{\nu}{2}\}$, so (19) hold if $\nu < \gamma < \frac{\alpha}{2} + 2 - \frac{\nu}{2}$. Therefore, γ must be such that

$$\begin{cases} \nu < \gamma < \dfrac{\alpha}{2} + 1 \ \text{ if } \ \nu < 2. \\ \nu < \gamma < \dfrac{\alpha}{2} + 2 - \dfrac{\nu}{2} \ \text{ if } \ \nu > 2. \end{cases} \tag{22}$$

Combining (21) and (22) we finally obtain

$$\begin{cases} \alpha > \nu \ \text{ and } \ \nu < \gamma < \dfrac{\alpha}{2} + 1 \ \text{ if } \ \nu < 1, \\ \alpha > 2\nu - 1 \ \text{ and } \ \nu < \gamma < \dfrac{\alpha}{2} + 1 \ \text{ if } \ 1 < \nu < 2, \\ \alpha > 2\nu - 1 \ \text{ and } \ \nu < \gamma < \dfrac{\alpha}{2} + 2 - \dfrac{\nu}{2} \ \text{ if } \ 2 < \nu < 3, \\ \alpha > 3\nu - 4 \ \text{ and } \ \nu < \gamma < \dfrac{\alpha}{2} + 2 - \dfrac{\nu}{2} \ \text{ if } \ 3 < \nu. \end{cases}$$

This concludes the proof. □

The following corollary provides estimates concerning the balance between the regularisation effects on the prior and on the noise, as quantified by the choice of the coefficients α and γ.

Corollary 2.1 *Suppose the spectrum of* M *follows the same asymptotics as the Laplace operator. Then,*

(a) For dimensions $d = 2$ *Theorem 2.1 holds for the choice* $\alpha > 1$ *and* $1 < \gamma < \frac{\alpha}{2}+1$.

(b) For dimensions $d = 3$ *Theorem 2.1 holds for the choice* $\alpha > \frac{3}{2}$ *and* $\frac{3}{2} < \gamma < \frac{\alpha}{2}+1$.

Proof. The asymptotics for the spectrum of the Laplace operator are $\lambda_n \sim n^{\frac{2}{d}}$, where d is the dimension of the domain. Then,

$$\sum_n \lambda_n^{-\nu} < C \sum_n n^{-\frac{2\nu}{d}}$$

for an appropriate constant C, and this series converges if $-\frac{2\nu}{d} < -1$, i.e. if $\nu > \frac{d}{2}$. The result follows by direct application of Theorem 2.1. □

We conclude this section by showing that the Radon-Nikodym derivative of the measures μ_0, μ_y can be expressed as the exponential of a potential function. This potential function is defined as

$$\mathfrak{I}_b(w; y) := \frac{1}{2}||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}u||^2 - \langle \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}u, \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}y\rangle$$

We will show, that under appropriate conditions the Radon-Nikodym derivative of the equivalent measures μ_0 and μ_y can be expressed in exponential form as

$$\frac{d\mu_y}{d\mu_0}(w) = C(y)\exp(-\mathfrak{I}_b(w; y)) \tag{23}$$

where $C(y)$ is an appropriate normalisation constant. This means that the most likely values for the solution of the inverse problem w would correspond to the minimum of the functional $\mathfrak{I}_b(w; y)$ with respect to w (given the data y). In this spirit, the Bayesian approach can be thought of as an alternative probabilistic regularisation procedure.

The following discussion serves as a motivation for this result. Since we impose a prior measure on the solution w, $\mu_0 = \mathcal{N}(m_0, \mathcal{C}_0)$ and a measure on the distribution of the noise η, $\mu_N = \mathcal{N}(0, \mathcal{C}_N)$, by the assumption of independence of η and w, it is seen that the measure or the random variable (w, y) taking values in $X \times Y$ will be the product measure $\varpi := \mu_0 \otimes \mu_N$. The aim of the Bayesian method is to "update" this belief by observing some of the data y and keeping the assumption of independence of η and w, as well as the assumption that noise is distributed via the Gaussian measure μ_N, obtain the conditional distribution $w \mid y$, which will be denoted by μ_y. This can be understood as the conditional distribution of a measure μ on $(X\times Y, \mathcal{B}(X\times Y))$, which corresponds to the joint distribution of (w, y). To obtain this we would informally work as follows: Since $y = \bar{\mathcal{G}}w + \eta$ thus $\eta = y - \bar{\mathcal{G}}w$, and we consider the distribution of noise as known (μ_N) we essentially know the distribution of $y - \bar{\mathcal{G}}w$. Keeping technicalities aside for the moment, the independence would mean that the joint distribution of $y - \bar{\mathcal{G}}w$ and w will be given by a measure μ such that $\mu(w \in A, y - \bar{\mathcal{G}}w \in B) = \mu_0(w \in A)\mu_N(y - \bar{\mathcal{G}}w \in B)$. From that we wish to extract information on $w \mid y$, i.e., on w given that we have obtained the measurements y. This requires obtaining the conditional measure μ_y, which is the conditional measure of μ defined above. To get this conditional measure we need to work using an infinite dimensional version of Bayes theorem. Informally, the above means that once w is known, the conditional distribution of $y \mid w$ is known since the distribution of $y - \bar{\mathcal{G}}w$ is known and w is known. Therefore, the conditional distribution of $y \mid w$, let us call it $P(y \in B \mid w) = \mu_w(B) = \mu_N(B - \bar{\mathcal{G}}w)$ where $B - \bar{\mathcal{G}}w = \{u - \bar{\mathcal{G}}w,\ u \in Y\}$. Therefore, the conditional measure which provides $y \mid w$, μ_w is nothing else but the shift of the measure μ_N in the direction $\bar{\mathcal{G}}w$. On the other hand, we have to consider the probability that we do get w as a solution. Under the chosen prior $P(w \in A) = \mu_0(w \in A)$, so we may define the measure μ on $(X \times Y, \mathcal{B}(X \times Y))$ so that

$$\mu(dw, dy) = \mu_0(dw)\mu_w(dy),$$

which corresponds to the joint distribution for w and y, once some knowledge for the solution w is given.

Our first observation is that in the construction of the measure μ is involved the quantity $\mu_N(y - \bar{\mathcal{G}}w \in B)$ for any set in $\mathcal{B}(Y)$. This is related to the translation of the measure μ_N along the direction $\bar{\mathcal{G}}w$ in the Hilbert space Y. As is well known, the translation properties of Gaussian measures in infinite dimensional spaces are rather subtle, the directions along which this is possible are limited to the Cameron-Martin space. Using the Cameron-Martin-Girsanov theorem [23], we may rewrite the shifted measure μ_w in terms of an exponential measure change,

$$\frac{d\mu_w}{d\mu_N}(y) = \exp\left(-\frac{1}{2}|\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w|^2 + \langle \mathcal{C}_N^{-\frac{1}{2}}y, \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w\rangle\right) = \exp(-\mathfrak{I}_b(w; y)),$$

if and only if $\bar{\mathcal{G}}w \in \mathcal{C}_N^{\frac{1}{2}}(H) = Im(\mathcal{C}_N^{\frac{1}{2}})$. The first observation then requires w to be such that $\bar{\mathcal{G}}w \in Im(\mathcal{C}_N^{\frac{1}{2}})$.

By the definition of the measures μ and ϖ we may see that $\mu << \varpi$, and in fact,

$$\begin{aligned}\frac{d\mu}{d\varpi}(w, y) &= \frac{d\mu_w}{d\mu_N}(y)\\ &= \exp\left(-\frac{1}{2}|\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w|^2 + \langle \mathcal{C}_N^{-\frac{1}{2}}y, \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w\rangle\right) = \exp(-\mathfrak{I}_b(w; y)). \quad (24)\end{aligned}$$

We now apply Bayes' theorem in the infinite dimensional setting. This comes from the application of Theorem 6.29 in [9], which states that, given any two probability measures μ, ϖ: if $\mu << \varpi$, $\frac{d\mu}{d\varpi} = \phi$ and the conditional distribution of $w \mid y$ under ϖ, $\varpi_y(dw)$ exists then the conditional distribution of $w \mid y$ under μ, $\mu_y(dw)$ exists, is absolutely continuous with respect to ϖ_y and has Radon-Nikodym derivative

$$\frac{d\mu_y}{d\varpi_y}(w) = \begin{cases} \frac{1}{C(y)}\phi(w,y) & \text{if} \quad c(y) > 0, \\ 1 & \text{else} \quad , \end{cases}$$

where $C(y) = \int_H \phi(w,y)d\varpi_y(w)$, for all $y \in H$. We apply this theorem to the measures μ and ϖ defined as above. Since ϖ is a product measure the conditional measure that provides the distribution of $w \mid y$, $\varpi_y = \mu_0$. An application of the theorem thus guarantees the existence of the conditional distribution of $w \mid y$ under the measure μ, i..e, existence of the conditional measure μ_y, which has a Radon-Nikodym derivative, given in exponential form, with respect to $\varpi_y = \mu_0$, hence,

$$\frac{d\mu_y}{d\mu_0}(w) = \exp\left(-\frac{1}{2}|\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w|^2 + \langle \mathcal{C}_N^{-\frac{1}{2}}y, \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w\rangle\right) = \exp(-\mathfrak{J}_b(w;y))$$

Another issue which is perhaps the issue that makes the Bayesian approach so attractive, is its possibility to deal with the effect that the result will have when the data are altered slightly, i.e., robustness of the estimators with respect to the data y. Recall, that the issue of robustness is a crucial point in inverse problems, and a number of deterministic regularisation techniques have been devised to deal with it (see e.g. [17] and references within). The Bayesian method provides robustness with respect to the data, in the sense that in the right function space setting, given two different sets of data y_1, y_2, and calculating the conditional distributions of $w \mid y_1$, $w \mid y_2$, μ_{y_1} and μ_{y_2} respectively, these probability measures are close together in terms of the Hellinger distance $d_{Hel}(\mu_{y_1}, \mu_{y_2})$ defined by

$$d_{Hel}(\mu_{y_1}, \mu_{y_2})^2 := \frac{1}{2}\int_H \left(\left(\frac{d\mu_{y_1}}{d\mu_0}\right)^{\frac{1}{2}} - \left(\frac{d\mu_{y_2}}{d\mu_0}\right)^{\frac{1}{2}}\right)^2 d\mu_0, \tag{25}$$

which is clearly simplified considerably if the exponential measure change (23) holds. The Hellinger distance between the two measures is an important quantity as it controls the variability of the moments of random variables which are functionals of w, under the probability measure μ_y, with respect to the observed data y.

The above informal steps can only be made rigorous by proper choice of the function spaces X and Y. In particular, these spaces are considered as subspaces of the original Hilbert space H, chosen so that the functional $\mathfrak{J}_b : X \times Y \to \mathbb{R}$, satisfies certain properties that guarantee that all the above steps are well defined. These properties have been collected in abstract form in [9] (see also [24]) and fall into two main groups:

Assumption 2.1 The space X is chosen so that $\mu_0(X) = 1$ and the functional $\mathfrak{J}_b : X \times Y \to \mathbb{R}$ satisfies

(i) For every $\epsilon > 0$, $r > 0$ these exists $M = M(\epsilon, r)$ such that for all $w \in X$, $y \in Y$ with $||y||_Y < r$,

$$\mathfrak{J}_b(w; y) \geq M - \epsilon ||u||_X^2;$$

(ii) For every $r > 0$ there exists $L = L(r) > 0$ such that for all $w_1, w_2 \in X$, $y \in Y$, with $\max\{||w_1||_X, ||w_2||_X, ||y||_Y\} < r$ it holds that

$$|\mathfrak{J}_b(w_1; y) - \mathfrak{J}_b(w_2; y)| \leq L(r)||w_1 - w_2||_X.$$

Assumption 2.2 The space X is chosen so that $\mu_0(X) = 1$ and the functional $\mathfrak{J}_b : X \times Y \to \mathbb{R}$ satisfies

(i) For every $r > 0$ there exists $L = L(r) > 0$ such that for all $(w, y) \in X \times Y$, with $\max\{||w||_X, ||y||_Y\} < r$ it holds that

$$\mathfrak{J}_b(w; y) \leq L(r);$$

(ii) For every $\epsilon > 0$, $r > 0$ these exists $M = M(\epsilon, r)$ such that for all $y_1, y_2 \in Y$ with $\max\{||y_1||_Y, ||y_2||_Y\} < r$ and for all $w \in X$,

$$|\mathfrak{J}_b(w; y_1) - \mathfrak{J}_b(w; y_2)| \leq \exp(\epsilon ||w||_X^2 + M)\, ||y_1 - y_2||_Y.$$

In the abstract setting of the problem explained in detail in [9], if Assumption 2.1 holds for the functional $\mathfrak{J}_b$ then the measure μ_y given by the exponential formula (23) is a well defined probability measure in H, whereas if Assumption 2.2 holds, then the measure μ_y enjoys continuity properties with respect to the data y.

We now formulate and prove the exponential measure change result.

Proposition 2.1 *Let Assumption 1.5 hold and assume further that $m_0 \in \mathcal{H}^\alpha$ and the solution and data space are chosen as $X = \mathcal{H}^r$ and $Y = \mathcal{H}^s$ respectively with r and s satisfying*

$$\begin{aligned} \gamma - 1 < r < \alpha - \nu, \\ \gamma - 1 < s < \gamma - \nu \end{aligned} \tag{26}$$

Then,

(i) The exponential measure change formula (23) *holds.*

(ii) μ_y is locally Lipschitz continuous with respect to the data $y \in Y$, in the sense that for all $y_1, y_2 \in Y$, such that $\max\{||y_1||_Y, ||y_2||_Y\} < r$,

$$d_{Hel}(\mu_{y_1}, \mu_{y_2}) \leq C(r)\, ||y_1 - y_2||_Y,$$

where d_{Hel} is the Hellinger distance (25) *between the two measures, therefore, $\phi(y) := \mathbb{E}_{\mu_y}[f(w)]$ is a continuous function of the data, for any polynomially bounded function $f : X \to \mathbb{R}$.*

Proof. Since $r \in (\gamma - 1, \alpha - \nu)$ and $s > \gamma - 1$, by Lemma 2.2 we have that the functional $\mathfrak{J}_b$ satisfies Assumptions 2.1 and 2.2.

(i) By the abstract results of [9] (see Theorem 4.1 op. cit.) μ_y given by the measure change (23) is a well defined probability measure on H. For that we employ Assumption 2.1(ii) which is a continuity result and which guarantees measurability of $\mathfrak{J}_b$, and Assumption 2.1(i) which provides by application of Fernique's theorem integrability of the Radon-Nikodym derivative. In all these arguments it is crucial to work in a subset of the original space which is of full measure. By a conditioning argument, based on the infinite dimensional version of Bayes' theorem (see e.g. Theorem 6.31 in [9]) we conclude that μ_y is in fact the distribution of $w \mid y$ we seek for. The inequality $s < \gamma - \nu$ is needed if we require Y to be of full measure with respect to μ_N (i.e., $\varpi(X \times Y) = 1$), or equivalently if we require some regularity for the data.

(ii) By the abstract results of [9] (see Theorem 4.2 op. cit.) we obtain the local Lipschitz continuity with respect to y of the Hellinger distance, hence the robustness of the solution with respect to the data. □

Remark 2.1 The first inequality in (26)f implies that the choice of γ depends on the choice of α and it must hold that $\gamma < \alpha + 1 - \nu$. The second of these inequalities imposes the condition that $\nu < 1$.

Lemma 2.2 *Let Assumption 1.5 hold. If we choose $X = \mathcal{H}^r$, $Y = \mathcal{H}^s$, with $\gamma - 1 < r < \alpha - \nu$ and $\gamma - 1 < s$, then the functional $\mathfrak{J}_b : X \times Y \to \mathbb{R}$ satisfies Assumptions 2.1 and 2.2.*

Proof. We first consider (i). Since $\mu_0 = \mathcal{N}(m_0, \mathcal{C}_0)$ with $\mathcal{C}_0 = \beta\,\mathsf{M}^{-\alpha}$, standard results (see e.g., Lemma 6.27 in [9]) imply that $\mu_0(X) = 1$ if $0 \le r < \alpha - \nu$. Properties (ii) and (iii) follow in a straightforward manner if X and Y are chosen so that the operators $\mathcal{C}_N^{-\frac{1}{2}} \bar{\mathcal{G}}^{\frac{1}{2}} = \delta^{-1}\mathsf{M}^{\gamma} : X \to H$ and $\mathcal{C}_N^{-\frac{1}{2}} \bar{\mathcal{G}}^{\frac{1}{2}} = \delta^{-1}\mathsf{M}^{\gamma} : Y \to H$ are continuous (equivalently bounded). We only check the second, the first being similar. This requires the following calculation: Take any $y = \sum_n y_n e_n \in Y = \mathcal{H}^r$ and calculate

$$||\mathcal{C}_N^{-\frac{1}{2}} \bar{\mathcal{G}}^{\frac{1}{2}} y||^2 = \delta^{-1} \sum_n \lambda_n^{\gamma} g_n y_n^2 \le C \sum_n \lambda_n^{\gamma - 1 - s} \lambda_n^s y_n^2 \le C \sum_n \lambda_n^s y_n^2 = C\,||y||_{\mathcal{H}^s},$$

as long as $\gamma - 1 - s < 0$, i.e., $s > \gamma - 1$. In the above, we have used the asymptotic behavior of $g_n \sim \lambda_n^{-1}$. A similar argument holds for the choice of the space $X = \mathcal{H}^r$ which leads to the choice $r > \gamma - 1$. The rest follows by elementary arguments based on the Cauchy-Schwarz inequality, simply stated here for the comfort of the reader.

Observe that for every $\theta > 0$, and every $(w, y) \in X \times Y$, with X, Y chosen as above, we have

$$\mathfrak{J}_b(w;y) = \frac{1}{2}||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w||^2 - \langle \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}y, \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}w\rangle \geq -|\langle \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}y, \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}w\rangle|$$
$$\geq -\frac{\theta^2}{2}||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}w||^2 - \frac{1}{2\theta^2}||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}y||^2 \geq -\frac{C\theta^2}{2}||w||_X^2 - \frac{C}{2\theta^2}||y||_Y^2,$$

where in the last inequality we used the continuity properties of the operator $\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}\frac{1}{2}$ in the chosen functional setting. Therefore, if we restrict to $y \in Y$ such that $||y||_Y < r$, Assumption 2.1(i) clearly holds for $\epsilon = \frac{C\theta^2}{2}$ and $M = -\frac{Cr^2}{2\theta^2}$.

By another application of the Cauchy-Schwarz inequality we obtain that for every $\theta > 0$,

$$\begin{aligned}
\mathfrak{J}_b(w;y) &= \frac{1}{2}||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w||^2 - \langle \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}y, \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}w\rangle \\
&\leq \frac{1}{2}||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w||^2 + |\langle \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}y, \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}w\rangle| \\
&\leq \frac{1}{2}||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w||^2 + \frac{\theta^2}{2}||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}w||^2 + \frac{1}{2\theta^2}||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}y||^2 \\
&\leq \left(C_1 + \frac{\theta^2}{2}\right)||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}w||^2 + \frac{1}{2\theta^2}||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}y||^2 \\
&\leq C\left(C_1 + \frac{\theta^2}{2}\right)||w||_X^2 + \frac{C}{2\theta^2}||y||_Y 2,
\end{aligned}$$

where we first used the fact that the operator $\bar{\mathcal{G}}^{\frac{1}{2}}$ is bounded and then the continuity of the operators $\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}} : X \to \mathbb{R}$ and $\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}} : Y \to \mathbb{R}$. Assumption 2.2(i) follows from the above estimate.

The Lipschitz continuity properties of Assumptions 2.1(ii) and 2.2(ii) follow from the quadratic nature of the functional $\mathfrak{J}_b$ and the continuity properties of the operator $\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}$ for the chosen function space setting. For example,

$$\begin{aligned}
&\mathfrak{J}_b(w_1;y) - \mathfrak{J}_b(w_2;y) = \frac{1}{2}||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w_1||^2 - \frac{1}{2}||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w_2||^2 \\
&-\langle \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}(w_1 - w_2), \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}y\rangle \\
&= \langle \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w_1, \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}(w_1 - w_2)\rangle \\
&+\langle \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}(w_1 - w_2), \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w_2\rangle - \langle \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}(w_1 - w_2), \mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}y\rangle
\end{aligned}$$

therefore,

$$|\mathfrak{J}_b(w_1;y) - \mathfrak{J}_b(w_2;y)| \leq \left(||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w_1|| + ||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}w_2|| + ||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}}y||\right)||\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}(w_1 - w_2)||$$
$$\leq C\left(||w_1||_X + ||w_2||_X + ||y||_Y\right)||w_1 - w_2||_X,$$

where we have used the continuity properties of the operators $\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}} : X \to \mathbb{R}$ and $\mathcal{C}_N^{-\frac{1}{2}}\bar{\mathcal{G}}^{\frac{1}{2}} : Y \to \mathbb{R}$. Therefore, Assumption 2.1(ii) holds. The proof of Assumption 2.2(ii) follows similarly. □

3. Extensions to complex media

One may further consider more complex materials, for example assuming that there holds a nonlocal in time version of Ohm's law of the form $J = \sigma E + \sigma_1 \star E + J_0$, where by $\sigma_1 \star E$ we denote the temporal convolution $(\sigma_1 \star E)(t, x) = \int_0^t \bar{\sigma}_1(t-s, x)E(s, x)dx$, $\bar{\sigma}_1$ being a matrix kernel. Then the quasi-static (parabolic) approximation becomes

$$\frac{\partial}{\partial t}(\sigma E + \sigma_1 \star E) + \mathrm{curl}(\mu^{-1}\mathrm{curl}E) = -\frac{\partial}{\partial t}(J_0),$$

or by assuming weak differentiability of the kernel σ_1 and setting $\bar{\sigma}_1(0) = \vartheta$, we get the equivalent form

$$\frac{\partial}{\partial t}(\sigma E) + \mathrm{curl}(\mu^{-1}\mathrm{curl}E) = -\frac{\partial}{\partial t}(J_0) + K \star E + \vartheta E, \tag{27}$$

where $K = (\bar{\sigma}_1)'$. Similar corrections will be introduced if one assumes more complex constitutive laws, involving convolutions in the description of the magnetic displacement B. Furthermore, we will denote by K the convolution operator related to the kernel $\bar{\sigma}_1'$, which is to be considered as an operator $\mathsf{K} : \mathfrak{H} \to \mathfrak{H}$. In terms of the operators defined above the equation assumes the abstract form

$$u' + \mathsf{M}u = f + \mathsf{K}u + \vartheta u. \tag{28}$$

The case $\mathsf{K} = 0$ corresponds to materials without dispersion effects. For a discussion of various possible constitutive laws see e.g. [25], or [26], where possible qualitative conditions for the material parameters are given. For example, for certain materials it is conceivable that $\vartheta < 0$.

We conclude this section with a few remarks on how the dispersion effects can be included, i.e., how we may treat the inverse source problem in the case where the linear operator $\mathsf{K} \neq 0$. We restrict ourselves to the case where the source of of a particular type $f(t, x) = a(t)w(x)$, where a is a known function of time and w is the spatial part of the source to be determined, and we see that the problem assumes the abstract form,

$$u' + \mathsf{M}u = \mathsf{K}u + a\, w, \tag{29}$$

where $\mathsf{K} : \mathfrak{H} \to \mathfrak{H}$ is a convolution operator $(\mathsf{K}u)(t) = \int_0 tK(t-s)u(s)ds$ (which is a bounded and linear operator). Then, the results of e.g., Grimmer [27] or Prüss [28] guarantee the existence of a unique family of bounded operators $R(t, s)$ which simplifies to the form $R(t,s) = R(t-s)$ because of the fact that K is of convolution type, called the family of resolvent operators, satisfying the conditions

$$\frac{\partial}{\partial t}R(t-s)v + \mathsf{M}R(t-s)v = \int_s^t K(t-r)R(r-s)vdr,$$
$$\frac{\partial}{\partial s}R(t-s)v - R(t-s)\mathsf{M}v = -\int_s^t R(t-r)K(r-s)vdr,$$

for every $v \in D(\mathsf{M})$ ($R(t-s)v$ is strongly continuously differentiable for every $v \in D(M)$). This family also enjoys strong continuity properties and $R(s,s) = I$. By Theorem 2.5 of [27], we have a variation of constants representation for the solution of (29) in terms of the resolvent operator as

$$u(t) = R(t)u(0) + \int_0^t R(t-r)a(r)wdr,$$

which formally is the same as the semigroup representation, only that now R involves the effects of the convolution operator K. In fact if $K = 0$ i.e., $\mathsf{K} = 0$, the operator family $R(t-s)$ coincides with the analytic semigroup $S(t-s) = \exp(-(t-s)\mathsf{M})$. By setting $t = T$ and assuming uncertainty in the data we obtain the following random operator equation for the solution of the inverse source problem

$$y = \bar{\mathcal{G}}w + \eta,$$

where y and η are as before, but now the operator $\bar{\mathcal{G}}$ is defined as

$$\bar{\mathcal{G}}w := \left(\int_0^T R(T-r)a(r)\right)w$$

where now the evolution family takes the role of the analytic semigroup generated by $-\mathsf{M}$. The resolvent operator may be computed by a Neumann series which is convergent under certain assumptions (see e.g., [28] where the existence proof of the operator is actually based upon the convergence of this series). In fact, as for the problem in question the function K generating the convolution operator K and which models dispersion effects is small in the sense of an appropriate norm, this series representation is expected to converge quite fast. Therefore, the Bayesian approach may proceed at least on a formal level without many differences at most point. However, there is an important difference. Unless the matrix function $K = \bar{\sigma}_1'$ is of a very special form the eigenfunctions of M may not coincide with the eigenfunctions of $\bar{\mathcal{G}}$, and this may complicate some of the arguments, which are facilitated considerably if the eigenbasis of M diagonalises $\bar{\mathcal{G}}$ as well. One way to bypass this difficulty is to choose the covariance of the prior and the noise distribution related to $\bar{\mathcal{G}}$ rather than M. However, all these considerations are beyond the scope of the present work and will be pursued in future work.

4. Conclusion

In this work we study the inverse source problem for the eddy current approximation to the Maxwell equations. This is a parabolic problem, extensively used in various applications in electromagnetic theory (in classical or complex media). In particular, we formulate the problem in abstract form as an inverse problem which (i) due to its inherent ill posedness and (ii) due to possible noise contamination of the available data is treated using statistical inversion techniques (a direction in which there has been extensive research activity during the last decade or so). In particular we apply the Bayesian inversion approach developed for infinite dimensional systems by A. M. Stuart [9] to this problem and provide detailed estimates on the choice of correlation operators for the prior and noise data, related to the spectrum of the Maxwell operator. These estimates provide important information of the regularization procedure needed in order to address this problem. Extensions to complex electromagnetic media, and the necessary modifications are also discussed.

Acknowledgement

We thank Professor Markos Katsoulakis for useful discussions and suggestions.

References

1. Victor Isakov. *Inverse Source Problems.* American Mathematical Society, 1990.
2. Richard Albanese and Peter B Monk. The inverse source problem for maxwell's equations. *Inverse Problems*, 22(3):1023, 2006.
3. Ana Alonso Rodríguez, Jessika Camano, and Alberto Valli. Inverse source problems for eddy current equations. *Inverse Problems*, 28(1):015006, 2012.
4. Habib Ammari, Annalisa Buffa, and Jean-Claude Nédélec. A justification of eddy currents model for the Maxwell equations. *SIAM Journal on Applied Mathematics*, 60(5):1805–1823, 2000.
5. Shumin Li and Masahiro Yamamoto. An inverse source problem for maxwell's equations in anisotropic media. *Applicable Analysis*, 84(10):1051–1067, 2005.
6. Masaru Ikehata. An inverse source problem for the heat equation and the enclosure method. *Inverse Problems*, 23(1):183, 2007.
7. Ying Zhang and Xiang Xu. Inverse source problem for a fractional diffusion equation. *Inverse Problems*, 27(3):035010, 2011.
8. Adel Hamdi and Imed Mahfoudhi. Inverse source problem in a one-dimensional evolution linear transport equation with spatially varying coefficients: application to surface water pollution. *Inverse Problems in Science and Engineering*, (ahead-of-print):1–25, 2013.
9. Andrew M Stuart. Inverse problems: a bayesian perspective. *Acta Numerica*, 19(1):451–559, 2010.
10. Bartek T Knapik, Aad W van der Vaart, and J Harry van Zanten. Bayesian recovery of the initial condition for the heat equation. *Communications in Statistics-Theory and Methods*, 42(7):1294–1313, 2013.
11. Lilian Arnold and Bastian Harrach. A unified variational formulation for the parabolic-elliptic eddy current equations. *SIAM Journal on Applied Mathematics*, 72(2):558–576, 2012.

12. Lilian Arnold and Bastian Harrach. Unique shape detection in transient eddy current problems. *Inverse Problems*, 29(9):095004, 2013.
13. Ana Alonso Rodríguez and Alberto Valli. *Eddy Current Approximation of Maxwell Equations.* Springer-Verlag Italia, Milano, 2010.
14. Florian Bachinger, Ulrich Langer, and Joachim Schöberl. Numerical analysis of nonlinear multiharmonic eddy current problems. *Numerische Mathematik*, 100(4):593–616, 2005.
15. Robert Dautray and Jacques-Louis Lions. *Mathematical Analysis and Numerical Methods for Science and Technology, Vol. 5: Evolution Problems I.* Springer, Berlin, 1992.
16. Mikhail Shlemovich Birman and Michael Solomyak. The selfadjoint maxwell operator in arbitrary domains. *Leningrad Mathematical Journal*, 1:95–115, 1990.
17. Jari P Kaipio and Erkki Somersalo. *Statistical and Computational Inverse Problems.* Springer, 2005.
18. Vladimir I Bogachev. *Measure Theory.* Springer Verlag, 2006 (2 volumes set).
19. Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions.* Cambridge University Press, 2008 (corrected version).
20. William Arveson. *A Short Course on Spectral Theory.* Springer, New York, 2002.
21. Alemdar Hasanov and Marián Slodička. An analysis of inverse source problems with final time measured output data for heat conduction equation: A semigroup approach. *Applied Mathematics Letters*, 26(2):207–214, 2012.
22. Aleksey I Prilepko, Dmitry G Orlovsky, and Igor A Vasin. *Methods for Solving Inverse Problems in Mathematical Physics.* Marcel Dekker, Inc., New York, 2000.
23. Vladimir I Bogachev. *Gaussian Measures.* American Mathematical Society, 1998.
24. Simon L Cotter, Masoumeh Dashti, and Andrew M Stuart. Approximation of bayesian inverse problems for pdes. *SIAM Journal on Numerical Analysis*, 48(1):322–345, 2010.
25. Gary F Roach, Ioannis G Stratis, and Athanasios N Yannacopoulos. *Mathematical Analysis of Deterministic and Stochastic Problems in Complex Media Electromagnetics*. Princeton University Press, 2012.
26. Jaime E Munoz Rivera, Maria Grazia Naso, and Elena Vuk. Asymptotic behaviour of the energy for electromagnetic systems with memory. *Mathematical Methods in the Applied Sciences*, 27(7):819–842, 2004.
27. Ronald C Grimmer. Resolvent operators for integral equations in a banach space. *Transactions of the American Mathematical Society*, 273(1):333–349, 1982.
28. Jan Prüss. On linear volterra equations of parabolic type in banach spaces. *Transactions of the American Mathematical Society*, 301(2):691–721, 1987.

◇ I. G. Stratis
Department of Mathematics,
National and Kapodistrian
University of Athens,
Athens, 15784
Greece
istratis@math.uoa.gr

◇ A. N. Yannacopoulos
Department of Statistics,
Athens University of Economics and Business,
Greece
ayannaco@aueb.gr