

A PRIORI MODELING FOR GRADIENT BASED INVERSE SCATTERING ALGORITHMS

S. Nordebo

School of Mathematics and Systems Engineering
Växjö University
Växjö 35195, Sweden

M. Gustafsson

Department of Electrical and Information Technology
Lund University
Box 118, Lund 22100, Sweden

Abstract—This paper presents a Fisher information based Bayesian approach to analysis and design of the regularization and preconditioning parameters used with gradient based inverse scattering algorithms. In particular, a one-dimensional inverse problem is considered where the permittivity and conductivity profiles are unknown and the input data consist of the scattered field over a certain bandwidth. A priori parameter modeling is considered with linear, exponential and arctangent parameter scalings and robust preconditioners are obtained by choosing the related scaling parameters based on a Fisher information analysis of the known background. The Bayesian approach and a principal parameter (singular value) analysis of the stochastic Cramér-Rao bound provide a natural interpretation of the regularization that is necessary to achieve stable inversion, as well as an indicator to predict the feasibility of achieving successful reconstruction in a given problem set-up. In particular, the Tikhonov regularization scheme is put into a Bayesian estimation framework. A time-domain least-squares inversion algorithm is employed which is based on a quasi-Newton algorithm together with an FDTD-electromagnetic solver. Numerical examples are included to illustrate and verify the analysis.

Corresponding author: S. Nordebo (sven.nordebo@vxu.se).

1. INTRODUCTION

Inverse scattering problems of technical interest are almost always ill-posed, see e.g., [2, 5, 9, 11, 15, 16, 19, 22, 29, 31]. With noisy data, a proper regularization has to be incorporated in order to guarantee stable inversion, and there is only a limited resolution attainable, see e.g., [2, 5, 13, 19]. When applicable, the number of degrees of freedom [2, 3, 27, 28] of a linear operator can be used to obtain an estimate of the number of retrievable parameters of an object, and hence a coarse estimate of the resolution. Another approach is to employ a Fisher information analysis and the Cramér-Rao bound to quantify the ill-posedness of the reconstruction and the inherent trade-off between the accuracy and the resolution, see e.g., [13, 24–26]. To regularize the problem, the classical approach is to employ the well-known Tikhonov regularization [16] which controls the modeling error as well as a suitable norm of the image itself.

Recently, there has been an increased interest in Bayesian principles for inverse problems, see e.g., [1, 32, 34, 35]. One possibility is to include the Tikhonov regularization parameter in a Bayesian framework and devise efficient algorithms to determine this parameter from the data, see [1, 32, 34, 35]. These approaches, and the related derivations leading to useful algorithms, usually relies on some form of linearization, such as e.g., with the distorted Born iterative method [4], etc. However, the approach taken here is different. A similar connection is exploited as in e.g., [34], i.e., the connection between the Maximum A Posteriori (MAP) estimate with a Gaussian prior, and the Tikhonov regularization. However, instead of placing a prior on the regularization parameter itself, the MAP criterion is exploited here in a Fisher information analysis setting, which is relating to some known background of interest.

The Fisher information analysis [13, 24] has been employed recently to obtain a robust preconditioner for gradient based inverse scattering algorithms, see [6, 23]. The main idea in [23] is to incorporate a linear parameter scaling such that the scaled Fisher information has a unit diagonal at some known background parameter value, cf., the Jacobi preconditioner in numerical analysis [10, 18]. The purpose of the present paper is to build further on the Fisher information based preconditioner in [23] by integrating it with a Tikhonov type regularization. The aim is to develop an analysis that is able to relate the optimal estimation error (the Cramér-Rao bound), the regularization constant, the noise level and the spatial resolution in a given measurement situation. For this purpose, the Tikhonov regularization scheme is put into a Bayesian estimation framework,

and a principal parameter (singular value) analysis of the stochastic Cramér-Rao bound is used to determine the regularization that is necessary to achieve stable inversion, as well as an indicator to predict the feasibility of achieving successful reconstruction in a given problem set-up.

The Fisher information based preconditioner employed in [23] is furthermore extended here to include nonlinear (a priori) parameter modeling using exponential and arctangential parameter scalings (or transformations), see also [31]. In this way, a priori information such as a lower (and/or upper) parameter bound (or range) can be straightforwardly included in the model. Examples of numerical inversion are used to illustrate that the Fisher information based sensitivity analysis is able to devise suitable regularization constants, as well as to indicate the feasibility of finding a useful solution to the inverse problem at hand, with a given signal to noise ratio, etc. As expected, the incorporation of nonlinear a priori parameter models such as the exponential or arctangential scalings will in some cases (when the background parameter values are close to their bounds) yield inversion results that performs much better than what is indicated by the Fisher information analysis. This is quite natural since the Fisher information analysis does not take upper and lower parameter bounds into account. Furthermore, the Fisher information analysis is effectively based on a linearization (first order differentials) about the known background, and is therefore most reliable for weak scattering problems, and do not take high contrast (nonlinear) effects into account.

The purpose of this paper is to present a Fisher information based Bayesian approach to analysis and design of the regularization and preconditioning parameters used with gradient based inverse scattering algorithms. A time-domain least-squares inversion algorithm [11] based on a quasi-Newton algorithm [8] together with an FDTD-electromagnetic solver [30] has been employed in order to generate the numerical examples. A one-dimensional inverse scattering problem is considered as it offers a natural introduction to parameter identification and wave splitting techniques cf., [11, 21]. Note, however, that the one-dimensional inverse problems are oftenly associated with severe uniqueness and stability problems due to the lack of spatial redundancy in the measured data. This is in contrast to the two- and three-dimensional inverse problems where spatial information can be exploited. Hence, it is only possible to retrieve a limited number of parameters in the one-dimensional inverse problems [11, 15, 21].

The rest of the paper is outlined as follows. In Section 2 is presented a brief review of the one-dimensional inverse scattering

problem with basic gradient expressions, a conditional statistical analysis containing Maximum Likelihood (ML) estimation and the Fisher information analysis. In Section 3, is presented the Fisher information based preconditioning strategy covering linear, exponential and arctangential parameter scalings. In Section 4 is given the Bayesian, or Maximum A Posteriori (MAP) framework for Tikhonov regularization and a principal (SVD) parameter analysis of the stochastic Cramér-Rao bound. Section 5 contains the numerical examples and Section 6 the summary and conclusions.

2. FISHER INFORMATION ANALYSIS FOR THE ONE-DIMENSIONAL INVERSE SCATTERING PROBLEM

Fisher information analysis for one- and two-dimensional inverse scattering problems have been treated in detail in e.g., [13, 23, 24]. A brief review of some of the main results regarding a conditional statistical analysis of the one-dimensional inverse scattering problem is given below to provide a generic framework for the subsequent a priori analysis and modeling.

2.1. The One-dimensional Inverse Scattering Problem

Consider the one-dimensional electromagnetic inverse scattering problem of imaging a finite slab $0 \leq x \leq a$ with relative permittivity and conductivity profiles $\epsilon(x)$ and $\sigma(x)$, respectively, see Fig. 1. The left half space $x < 0$ is free space with $\epsilon = 1$ and $\sigma = 0$, and in the right half space $x > a$ the material parameters ϵ and σ are assumed to take constant values. Let $E(x, t)$ and $H(x, t)$ denote the time domain electric and magnetic field amplitudes of a plane wave propagating in the x -direction.

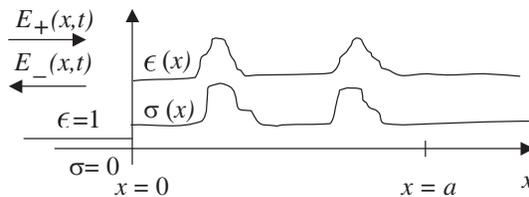


Figure 1. One-dimensional inverse scattering problem for a finite slab $0 \leq x \leq a$ with relative permittivity and conductivity profiles $\epsilon(x)$ and $\sigma(x)$, respectively.

The imaging is based on a known incident field $E_+(0, t)$ and a measurement of the scattered field $E_-^{(m)}(t)$ at the boundary $x = 0$ for $t \in [0, T]$ where T is the length of the observation interval. The following misfit functional is considered

$$\mathcal{J}(\epsilon, \sigma) = \int_0^T |E_-(0, t) - E_-^{(m)}(t)|^2 dt \tag{1}$$

where $E_-(0, t)$ is the modeled scattered field obtained as the solution to the Maxwell's equations [†]

$$\begin{cases} \epsilon \partial_t E - \partial_x H + \sigma E = 0 \\ \partial_t H - \partial_x E = 0 \end{cases} \tag{2}$$

for $x \in \mathbb{R}$ and $t \in [0, T]$, together with the initial condition $E = 0$ for $x \geq 0$ and $t = 0$, and the boundary condition $E_+ = E_+(0, t)$ for $x = 0$ and $t \in [0, T]$. The split fields are defined as $E_{\pm} = (E \mp H)/2$.

Assume that the spatial region $\mathcal{S} = \{x | 0 \leq x \leq a\} = \cup_{i=1}^I \mathcal{S}_i$ is decomposed into a finite set of disjoint intervals \mathcal{S}_i . The relative permittivity and conductivity within the material is discretized according to the finite expansions

$$\begin{cases} \epsilon(x) = \sum_{i=1}^I \epsilon_i \chi_i(x) \\ \sigma(x) = \sum_{i=1}^I \sigma_i \chi_i(x) \end{cases} \tag{3}$$

where ϵ_i and σ_i are the optimization variables and $\chi_i(x)$ the characteristic function for pixel \mathcal{S}_i , i.e., $\chi_i(x) = 1$ if $x \in \mathcal{S}_i$ and $\chi_i(x) = 0$ if $x \notin \mathcal{S}_i$.

The gradients of the misfit functional (1) are given by

$$\begin{cases} \frac{\partial \mathcal{J}}{\partial \epsilon_i} = - \int_{\mathcal{S}_i} \int_0^T \tilde{E}(x, t) \partial_t E(x, t) dt dx \\ \frac{\partial \mathcal{J}}{\partial \sigma_i} = - \int_{\mathcal{S}_i} \int_0^T \tilde{E}(x, t) E(x, t) dt dx. \end{cases} \tag{4}$$

[†] Let $k_0, c_0, \epsilon_0, \mu_0$ and η_0 denote the wave number, the speed of light, the permittivity, the permeability and the wave impedance of free space, respectively. Further, let $e^{i\omega t}$ be the time-convention where ω is the angular frequency. The common SI-unit quantities are normalized as $(t, \omega, \epsilon, \sigma, \mathbf{E}, \mathbf{H}, \mathbf{J})^{\text{norm}} = (c_0 t, \omega/c_0, \epsilon, \eta_0 \sigma, \sqrt{\epsilon_0} \mathbf{E}, \sqrt{\mu_0} \mathbf{H}, \sqrt{\mu_0} \mathbf{J})$ so that the speed of wave propagation is normalized to unity and all fields are measured in the same energy unit (Energy/Volume)^{1/2}.

where the adjoint electric and magnetic fields \tilde{E} and \tilde{H} satisfy the adjoint Maxwell's equations

$$\begin{cases} -\epsilon\partial_t\tilde{E} + \partial_x\tilde{H} + \sigma\tilde{E} = 0 \\ -\partial_t\tilde{H} + \partial_x\tilde{E} = 0 \end{cases} \quad (5)$$

for $x \in \mathbb{R}$ and $t \in [0, T]$, together with the boundary conditions $\tilde{E}_-(0, t) = E_-(0, t) - E_-^{(m)}(t)$ yielded from the solution of (2), see e.g., [7, 11, 12, 23]. Note that (5) is solved backwards in time and the "initial" conditions are $\tilde{E}(x, T) = 0$ for $x \geq 0$.

2.2. Conditional Statistical Analysis

Let $\boldsymbol{\nu} = (\boldsymbol{\epsilon}, \boldsymbol{\sigma})$ denote the finite parameter vector with elements ϵ_i and σ_i as defined in (3) and let \mathbf{x} denote the vector of measurement data with probability density function $p(\mathbf{x}|\boldsymbol{\nu})$. To begin with, the vector \mathbf{x} may be taken as finite, containing either time domain or frequency domain data. By using a limiting process as demonstrated in e.g., [20, 23], the vector \mathbf{x} may then be regarded as a collection of data on a continuous (infinite) time or frequency domain. Below, the angular frequency is denoted ω and the wave number $k_0 = \omega/c_0$. As a notational convention, the arguments of $E(x, t)$ and $E(x, f)$ indicate whether the fields are in the time or frequency domains, respectively.

Under the assumption of uncorrelated Gaussian measurement noise, the negative loglikelihood function for the inverse problem stated in Section 2.1 above can be expressed as

$$\begin{aligned} -\ln p(\mathbf{x}|\boldsymbol{\nu}) &= A + \frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{1}{R_N(\omega)} |E_-(0, \omega) - E_-^{(m)}(\omega)|^2 d\omega \\ &= A + \frac{1}{2N_0} \mathcal{J}(\boldsymbol{\nu}) \end{aligned} \quad (6)$$

where $R_N(\omega)$ is the power spectral density of the measurement noise, A is a constant, and $\mathcal{J}(\boldsymbol{\nu})$ the misfit functional defined in (1), cf., [23]. Here, the power spectral density of the measurement noise is assumed to be a constant $R_N(\omega) = N_0$ over the relevant bandwidth. Hence, with the Gaussian model for the measurement noise, the optimization problem stated in Section 2.1 based on the misfit functional (1) is equivalent to the classical Maximum Likelihood (ML) criterion, see also [14, 16, 17, 23].

The Fisher information matrix [17] for the parameters ϵ_i and σ_i based on the same statistical (Gaussian) measurement model as above,

is given by

$$[\mathcal{I}_{\nu\zeta}]_{ij} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{R_N(\omega)} \frac{\partial E_-^*(0, \omega)}{\partial \nu_i} \frac{\partial E_-(0, \omega)}{\partial \zeta_j} d\omega \quad (7)$$

where $(\cdot)^*$ denotes the complex conjugate and ν and ζ are either ϵ or σ , and $i, j = 1, \dots, I$, cf., [23]. For a homogenous background, it can be shown that the sensitivity field is given by

$$\frac{\partial E_-(0, \omega)}{\partial \nu_i} = -2g_\nu \frac{E_+(0, \omega)}{(1 + \sqrt{\epsilon_c})^2} \frac{\sin(k\Delta x)}{k\Delta x} \Delta x e^{-ik(2i-1)\Delta x} \quad (8)$$

where $g_\nu = 1$ if $\nu = \sigma$ and $g_\nu = i\omega$ if $\nu = \epsilon$, and where $k = k_0\sqrt{\epsilon_c} = \omega\sqrt{\epsilon_c}$, $\epsilon_c = \epsilon - i\sigma/\omega$, $\mathcal{S}_i = [(i-1)\Delta x, i\Delta x]$ and Δx the spatial sampling interval, see [23]. After evaluating (7), the total Fisher information matrix is assembled as

$$\mathcal{I}(\boldsymbol{\nu}) = \begin{pmatrix} \mathcal{I}_{\epsilon\epsilon} & \mathcal{I}_{\epsilon\sigma} \\ \mathcal{I}_{\sigma\epsilon} & \mathcal{I}_{\sigma\sigma} \end{pmatrix}. \quad (9)$$

3. FISHER INFORMATION BASED PARAMETER SCALING

A robust preconditioning strategy based on the Fisher information analysis together with a linear parameter scaling has been treated in [6, 23]. Below, these ideas are extended to incorporate non-linear parameter scalings such as the exponential and arctangential parameter models. The advantage of using a non-linear parameter model (or scaling) is that it has the ability to efficiently incorporate a priori information such as lower and upper parameter bounds, see also [14, 31].

The optimization problem aims at minimizing the mistfit functional (1), or equivalently, to minimize the negative loglikelihood function (6). The Hessian of the negative loglikelihood function is given by

$$[\mathcal{H}(\mathbf{x}|\boldsymbol{\nu})]_{ij} = -\frac{\partial^2 \ln p(\mathbf{x}|\boldsymbol{\nu})}{\partial \nu_i \partial \nu_j} \quad (10)$$

and the Fisher information matrix is defined by

$$[\mathcal{I}(\boldsymbol{\nu})]_{ij} = \mathcal{E}\{[\mathcal{H}(\mathbf{x}|\boldsymbol{\nu})]_{ij}\} = -\mathcal{E}\left\{\frac{\partial^2 \ln p(\mathbf{x}|\boldsymbol{\nu})}{\partial \nu_i \partial \nu_j}\right\}, \quad (11)$$

where $\mathcal{E}\{\cdot\}$ denotes the expectation operator, see e.g., [17].

A robust preconditioner is obtained by incorporating a parameter scaling (or transformation) such that the scaled (or transformed) Fisher information has a unit diagonal at some known background parameter value $\boldsymbol{\nu}$, cf., the Jacobi preconditioner in numerical analysis [10, 18]. Since the Fisher information matrix is the mean value of the Hessian in the corresponding Maximum Likelihood estimation problem, it is expected that such a strategy will stabilize any gradient based numerical inversion algorithm and that the problem with local minima should be alleviated [6, 23]. The preconditioner is robust in the sense that the scaling, i.e., the diagonal Fisher information is virtually invariant to the numerical resolution and the discretization (pixel) model that is employed, see [6, 23] for a detailed study about this issue.

Consider the following linear and nonlinear parameter models

$$\begin{aligned} \nu_i &= \xi_i/\beta_i + \nu_{0i} && \text{Linear} \\ \nu_i &= \alpha_i e^{\xi_i/\beta_i} + \nu_{0i} && \text{Exponential} \\ \nu_i &= \alpha_i \arctan(\xi_i/\beta_i + \bar{\xi}_i) + \nu_{0i} && \text{Arctangential} \end{aligned} \quad (12)$$

where ξ_i is the new optimization variable and β_i the scaling constant. Here, α_i , ν_{0i} , and $\bar{\xi}_i$ are a priori chosen model constants.

With the linear scaling, it is assumed that the known background corresponds to $\xi_i = 0$, or $\nu_i = \nu_{0i}$. The gradient is given by $\frac{\partial}{\partial \xi_i} = G_i \frac{\partial}{\partial \nu_i}$ and the scaled Fisher information is $[\mathcal{I}(\boldsymbol{\xi})]_{ij} = G_i G_j [\mathcal{I}(\boldsymbol{\nu})]_{ij}$ where $G_i = \frac{1}{\beta_i}$. Hence, a robust Fisher information based Jacobi preconditioner with $[\mathcal{I}(\boldsymbol{\xi})]_{ii} = 1$, is given by

$$\begin{cases} \beta_i = \sqrt{[\mathcal{I}(\boldsymbol{\nu})]_{ii}} \\ \frac{\partial}{\partial \xi_i} = \frac{1}{\sqrt{[\mathcal{I}(\boldsymbol{\nu})]_{ii}}} \frac{\partial}{\partial \nu_i} \end{cases} \quad (13)$$

and the resulting scaled Fisher information matrix is given by

$$[\mathcal{I}(\boldsymbol{\xi})]_{ij} = \frac{1}{\sqrt{[\mathcal{I}(\boldsymbol{\nu})]_{ii}} \sqrt{[\mathcal{I}(\boldsymbol{\nu})]_{jj}}} [\mathcal{I}(\boldsymbol{\nu})]_{ij}. \quad (14)$$

With the exponential scaling, it is assumed that the known background corresponds to $\xi_i = 0$, or $\nu_i = \alpha_i + \nu_{0i}$ where ν_{0i} represents a lower parameter bound. The gradient is given by $\frac{\partial}{\partial \xi_i} = G_i \frac{\partial}{\partial \nu_i}$ and the scaled Fisher information is $[\mathcal{I}(\boldsymbol{\xi})]_{ij} = G_i G_j [\mathcal{I}(\boldsymbol{\nu})]_{ij}$ where $G_i = \frac{\alpha_i}{\beta_i} e^{\xi_i/\beta_i}$. The appropriate scaling is then given by

$$\begin{cases} \beta_i = \alpha_i \sqrt{[\mathcal{I}(\boldsymbol{\nu})]_{ii}} \\ \frac{\partial}{\partial \xi_i} = \frac{1}{\sqrt{[\mathcal{I}(\boldsymbol{\nu})]_{ii}}} e^{\xi_i/\beta_i} \frac{\partial}{\partial \nu_i} \end{cases} \quad (15)$$

and the resulting scaled Fisher information matrix at the background level $\xi = \mathbf{0}$ is again given by (14).

With the arctangential scaling, it is assumed that the known background corresponds to $\xi_i = 0$, or $\nu_i = \alpha_i \arctan(\bar{\xi}_i) + \nu_{0i}$. Further, if v_{ui} and v_{li} denotes upper and lower parameter bounds, respectively, it is seen that $2\nu_{0i} = v_{ui} + v_{li}$ and $\pi\alpha_i = v_{ui} - v_{li}$. The gradient is given by $\frac{\partial}{\partial \xi_i} = G_i \frac{\partial}{\partial \nu_i}$ and the scaled Fisher information is $[\mathcal{I}(\xi)]_{ij} = G_i G_j [\mathcal{I}(\nu)]_{ij}$ where $G_i = \frac{\alpha_i}{\beta_i} (1 + (\xi_i/\beta_i + \bar{\xi}_i)^2)^{-1}$. The appropriate scaling is then given by

$$\begin{cases} \beta_i = \frac{\alpha_i}{1 + \bar{\xi}_i^2} \sqrt{[\mathcal{I}(\nu)]_{ii}} \\ \frac{\partial}{\partial \xi_i} = \frac{1 + \bar{\xi}_i^2}{1 + (\xi_i/\beta_i + \bar{\xi}_i)^2} \frac{1}{\sqrt{[\mathcal{I}(\nu)]_{ii}}} \frac{\partial}{\partial \nu_i} \end{cases} \quad (16)$$

and the resulting scaled Fisher information matrix at the background level $\xi = \mathbf{0}$ is again given by (14).

The contrast in the scaled parameter $\Delta \xi_i$ corresponding to a deviation $\Delta \nu_i$ with respect to the known background, is given for the three parameter models in (12) as follows

$$\Delta \xi_i = \sqrt{[\mathcal{I}(\nu)]_{ii}} \Delta \nu_i \quad (17)$$

$$\Delta \xi_i = \alpha_i \sqrt{[\mathcal{I}(\nu)]_{ii}} \ln \left(1 + \frac{\Delta \nu_i}{\alpha_i} \right) \quad (18)$$

$$\Delta \xi_i = \frac{\alpha_i \sqrt{[\mathcal{I}(\nu)]_{ii}}}{1 + \bar{\xi}_i^2} \left[\tan \left(\frac{\Delta \nu_i}{\alpha_i} + \arctan(\bar{\xi}_i) \right) - \bar{\xi}_i \right]. \quad (19)$$

Obviously, the nonlinear models above yield

$$\Delta \xi_i \approx \sqrt{[\mathcal{I}(\nu)]_{ii}} \Delta \nu_i \quad (20)$$

when $\frac{\Delta \nu_i}{\alpha_i}$ is small.

4. A PRIORI MODELING AND REGULARIZATION

4.1. The a Priori Statistics of Tikhonov Regularization

In inverse problem theory and applications, it is common to employ a Tikhonov type of regularization that punish rapid spatial variations in the medium parameters, see e.g., [11]. Hence, the following Tikhonov regularization scheme may be considered

$$\min_{\xi} \left\{ \mathcal{J}(\xi) + \gamma \int_0^a \left(\frac{\partial \xi}{\partial x} \right)^2 dx \right\} \quad (21)$$

where γ is the regularization constant. By employing the boundary conditions $\xi(0) = \xi(a) = 0$ and integrating by parts, the integral above can be approximated in discrete form as

$$\int_0^a \left(\frac{\partial \xi}{\partial x} \right)^2 dx = - \int_0^a \xi(x) \frac{\partial^2 \xi}{\partial x^2} dx \approx \frac{1}{\Delta x} \boldsymbol{\xi}^T \bar{\mathbf{C}}^{-1} \boldsymbol{\xi} \quad (22)$$

where Δx is the discretization interval, $\boldsymbol{\xi}$ an $N \times 1$ sample vector and

$$\bar{\mathbf{C}}^{-1} = \begin{pmatrix} 2 & -1 & 0 & & & \\ -1 & 2 & -1 & & & \\ & & \ddots & & & \\ & & & -1 & 2 & -1 \\ & & & 0 & -1 & 2 \end{pmatrix} \quad (23)$$

an $N \times N$ symmetric Toeplitz matrix. It is noted that the matrix $\bar{\mathbf{C}}^{-1}$ is the inverse of the symmetric matrix $\bar{\mathbf{C}}$ with elements $[\bar{\mathbf{C}}]_{ij} = i(N-j+1)/(N+1)$ for $j \geq i$ and $[\bar{\mathbf{C}}]_{ij} = [\bar{\mathbf{C}}]_{ji}$.

Now, the Tikhonov regularization scheme can be given an unconditional statistical (or Bayesian estimation) interpretation as follows. Assume that the parameter vector $\boldsymbol{\xi}$ has zero mean and known prior Gaussian distribution with probability density function $p(\boldsymbol{\xi})$ where $\ln p(\boldsymbol{\xi}) = d - \frac{1}{2} \boldsymbol{\xi}^T \mathbf{C}^{-1} \boldsymbol{\xi}$, d is a constant and the correlation matrix \mathbf{C} is given by

$$\mathbf{C} = \mathcal{E} \{ \boldsymbol{\xi} \boldsymbol{\xi}^T \} = \frac{\Delta x N_0}{\gamma} \bar{\mathbf{C}} \quad (24)$$

where N_0 is the spectral density of the measurement noise. The spatial variance and correlation coefficient corresponding to the correlation matrix $\bar{\mathbf{C}}$ is depicted in Fig. 2 below.

The Maximum A Posteriori (MAP) criterion [14, 16, 17] is to maximize the posterior conditional density function $p(\boldsymbol{\xi} | \mathbf{x})$ with respect to $\boldsymbol{\xi}$ where \mathbf{x} is the measurement vector, or equivalently, to minimize the function $-\ln p(\mathbf{x} | \boldsymbol{\xi}) - \ln p(\boldsymbol{\xi})$ where $-\ln p(\mathbf{x} | \boldsymbol{\xi})$ is the negative loglikelihood function defined in (6). Hence, the MAP criterion can be stated as

$$\min_{\boldsymbol{\xi}} \left\{ \frac{1}{2N_0} \mathcal{J}(\boldsymbol{\xi}) + \frac{1}{2} \boldsymbol{\xi}^T \mathbf{C}^{-1} \boldsymbol{\xi} \right\} \quad (25)$$

which is equivalent to (21) when $\mathbf{C}^{-1} = \frac{\gamma}{\Delta x N_0} \bar{\mathbf{C}}^{-1}$ according to the definition (24). Note that the gradient corresponding to (25) is given by

$$\frac{1}{2N_0} \frac{\partial}{\partial \boldsymbol{\xi}} \mathcal{J}(\boldsymbol{\xi}) + \mathbf{C}^{-1} \boldsymbol{\xi} \quad (26)$$

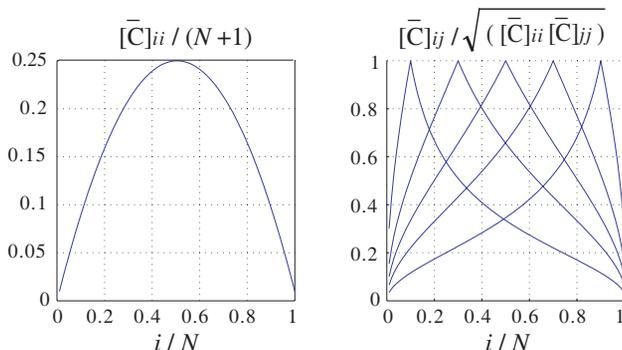


Figure 2. Spatial variance $[\bar{\mathbf{C}}]_{ii}/(N + 1)$ and correlation coefficient $[\bar{\mathbf{C}}]_{ij}/\sqrt{[\bar{\mathbf{C}}]_{ii}[\bar{\mathbf{C}}]_{jj}}$ plotted for large N ($N = 100$). In the right figure, the various curves correspond to $j/N = \{0.1, 0.3, 0.5, 0.7, 0.9\}$.

where $\frac{\partial}{\partial \nu} \mathcal{J}(\nu)$ has been defined in Section 2.1 and $\frac{\partial}{\partial \xi}$ in Section 3.

According to the Bayesian statistical analogue given above, the use of the Tikhonov regularization scheme is equivalent to the assumption of a certain prior Gaussian parameter distribution together with an application of the MAP criterion. As will be demonstrated by using numerical experiments below, even if this parameter model is not physically justified, the statistical analogue is useful for characterizing the balance between the estimation error (the Cramér-Rao bound), the regularization constant, the signal to noise ratio and the spatial resolution in a given measurement situation. The aim is to employ the Fisher information analysis as a tool to predict the feasibility of successful reconstruction as well as to choose a proper regularization constant.

4.2. Principal Parameter Analysis

The Fisher information matrix for a stochastic vector parameter ξ is given by

$$\begin{aligned}
 \mathcal{I} &= -\mathcal{E}_{\mathbf{x}, \xi} \left\{ \frac{\partial^2 \ln p(\xi | \mathbf{x})}{\partial \xi \partial \xi^T} \right\} \\
 &= -\mathcal{E}_{\xi} \mathcal{E}_{\mathbf{x} | \xi} \left\{ \frac{\partial^2 \ln p(\mathbf{x} | \xi)}{\partial \xi \partial \xi^T} \right\} - \mathcal{E}_{\xi} \left\{ \frac{\partial^2 \ln p(\xi)}{\partial \xi \partial \xi^T} \right\} \\
 &= \mathcal{E}_{\xi} \{ \mathcal{I}(\xi) \} + \mathbf{C}^{-1}
 \end{aligned} \tag{27}$$

where $\mathcal{E}_{\mathbf{x},\xi}$ and $\mathcal{E}_{\mathbf{x}|\xi}$ denote the unconditional and the conditional expectation operators, respectively, see e.g., [33]. In (27), $\mathcal{I}(\xi)$ denotes the conditional Fisher information defined in e.g., (7), (11) and (14), and \mathbf{C} the correlation matrix defined in (24). The stochastic Cramér-Rao bound is given by the inverse of \mathcal{I} , see e.g., [33].

In practice, it is extremely costly to compute $\mathcal{E}_{\xi}\{\mathcal{I}(\xi)\}$ in (27), and hence the approximation $\mathcal{I} \approx \mathcal{I}(\xi)|_{\xi=0} + \mathbf{C}^{-1}$ is used where the conditional Fisher information is calculated at the known background. Note that this approximation is asymptotically unbiased and converges in probability as $\gamma \rightarrow \infty$.

The Fisher information matrix \mathcal{I} is extremely ill-conditioned and a calculation of the Cramér-Rao bound for individual pixels ξ_i is virtually impossible (and irrelevant) if the pixel resolution Δx is far below the resolution limit $\Delta x \ll \lambda/2$, cf., [13]. However, a principal parameter analysis using the singular value decomposition (SVD) may be carried out to identify the significant number of retrievable parameters, and hence the resolution. The following notation will be employed

$$\mathcal{I}(\xi)|_{\xi=0} + \mathbf{C}^{-1} = \mathbf{U}\Sigma\mathbf{U}^T \quad (28)$$

where \mathbf{U} contains the singular vectors and Σ the singular values. The principal parameters are defined by $\mathbf{U}^T\xi$ and the corresponding Cramér-Rao bounds are given by the diagonal elements of Σ^{-1} .

5. NUMERICAL EXAMPLES

5.1. High Loss Example

Consider the one-dimensional inverse scattering problem as described in Section 2.1 and depicted in Fig. 1. A conditional Fisher information analysis and related preconditioning was carried out for the linear, exponential and arctangential parameter scalings as described in Sections 2.2 and 3, and an a priori model and principal parameter analysis as described in Section 4. The normalized bandwidth was $B = (\omega_2 - \omega_1)/\omega_0 = 4/3$ ($0 \leq B \leq 2$) where ω_0 is the center frequency ($\omega_0 = 2\pi f_0$ where $f_0 = 6$ GHz in the examples below) and ω_1 and ω_2 the lower and upper frequency band limits, respectively. The calculations were performed for a homogenous background with $\epsilon = 15$ and $\sigma/\omega_0 = 1.2$ (in SI-units $\sigma_{\text{SI}} = 0.4$ S/m). Note that the conductivity parameter has been scaled as σ/ω_0 (where ω_0 is the center frequency) in order for the parameters ϵ and σ/ω_0 to obtain similar sensitivity, cf., [13]. The imaging domain was set to $a = 15\lambda$ where λ denotes the wavelength in the background medium at the upper frequency band

limit ω_2 . The pixel resolution was $\Delta x = 0.1\lambda$. The conditional Fisher information employed is defined in (7).

In Fig. 3, the six subplots show the principal Cramér-Rao bound (singular values) $[\Sigma^{-1}]_n$ as defined in (28), plotted as a function of resolution $2a/n\lambda$ for different regularization parameters γ_0 . The resolution is defined here as the size of the spatial domain in wavelengths (a/λ) divided by the number ($n/2$) of retrieved principal parameters for ϵ and σ , respectively, corresponding to n singular values $[\Sigma]_n$ in decreasing order. In Fig. 3, the regularization parameter

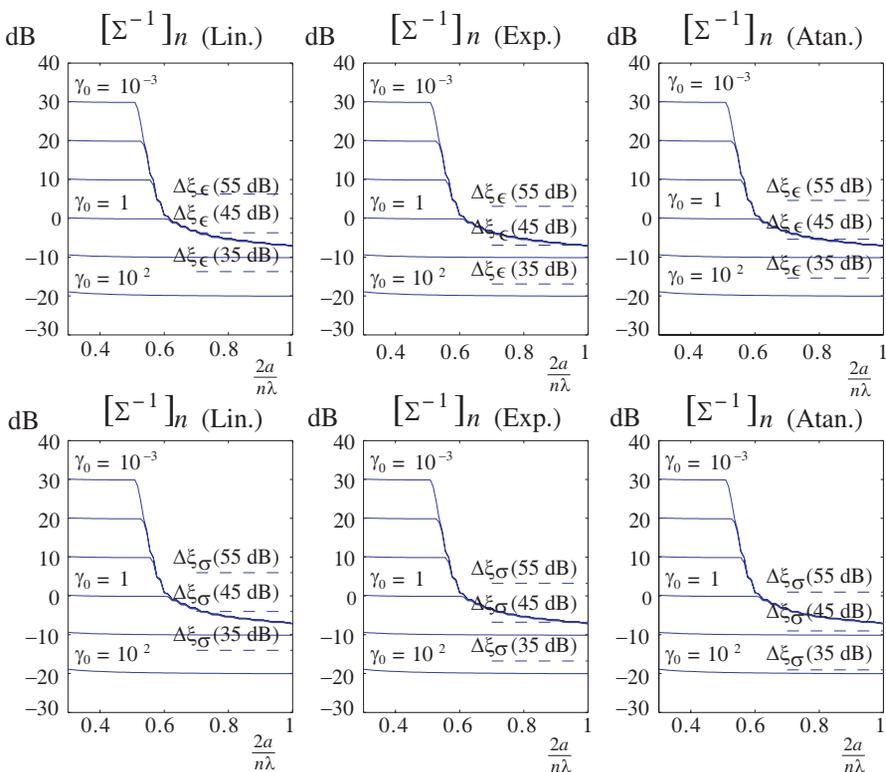


Figure 3. Principal Cramér-Rao bound $[\Sigma^{-1}]_n$ (in dB) as a function of resolution $2a/n\lambda$ for a homogenous background with $\epsilon = 15$ and $\sigma/\omega_0 = 1.2$. Here γ_0 denotes the regularization parameter and $\Delta\xi_\epsilon$ (upper plots) and $\Delta\xi_\sigma$ (lower plots) the contrast levels for the scaled ϵ - and σ -parameters at $x = 6\lambda$ with $\Delta\epsilon = \Delta\sigma = 1$, corresponding to SNR = 35, 45 and 55 dB and linear (Lin.), exponential (Exp.) and arctangential (Atan.) scaling, respectively.

defined in (21) and (24) is given by $\gamma = \gamma_0 2N_0 10^{-4}$ where γ_0 ranges from 10^{-3} to 10^2 , and $\Delta\xi_\epsilon$ (upper plots) and $\Delta\xi_\sigma$ (lower plots) are the contrast levels (17) through (19) for the scaled ϵ - and σ -parameters at $x = 6\lambda$ with $\Delta\epsilon = \Delta\sigma = 1$, corresponding to a signal to noise ratio $\text{SNR} = \frac{1}{4N_0} = 35, 45$ and 55 dB and linear (Lin.), exponential (Exp.) and arctangential (Atan.) parameter scalings, respectively. Note that the scaled Fisher information (14) used in (28) is independent of the noise strength N_0 , and hence motivates the scaling of γ with N_0 as above, see also (24).

In Fig. 3, it can be clearly seen the sharp resolution limit at $2a/n = \lambda/2$, beyond which reconstruction (estimation) becomes virtually unfeasible, cf., also [13]. For a resolution limit not below about 0.6λ and a signal to noise ratio $\text{SNR} = 55$ dB, the analysis predicts that reconstruction is feasible and virtually independent of the regularization constant if $\gamma_0 \leq 1$, i.e., the Cramér-Rao bound is significantly below the required contrast levels $\Delta\xi_\epsilon$ and $\Delta\xi_\sigma$ ($[\Sigma^{-1}]_n \ll \Delta\xi^2$). On the other hand, with a signal to noise ratio $\text{SNR} = 35$ dB, a regularization constant $\gamma_0 = 100$ is required to constrain the Cramér-Rao bound below the required contrast levels. In this case, the singular values as a function of resolution have become perfectly flat, indicating that the regularization has saturated, i.e., the regularization has become so dominating that the only thing that can be retrieved is the a priori information itself, i.e., the background parameters. Hence, the analysis predicts that reconstruction is unfeasible for $\text{SNR} = 35$ dB.

Next, a numerical implementation of the one-dimensional inverse problem is considered with the linear, exponential and arctangential scalings and preconditioning as defined in (12), (13), (15) and (16). A priori model parameters for the linear scaling was; $\epsilon_{0i} = 15$ (background) and $\sigma_{0i} = 1.2$ (background), for the exponential scaling; $\alpha_\epsilon = 1$, $\epsilon_{0i} = 14$ (lower bound) and $\alpha_\sigma = 1.2$, $\sigma_{0i} = 0$ (lower bound), and for the arctangential scaling; $\epsilon_l = 14$ (lower bound), $\epsilon_b = 15$ (background), $\epsilon_u = 17$ (upper bound) and $\sigma_l = 0$ (lower bound), $\sigma_b = 1.2$ (background) and $\sigma_u = 10\sigma_b$ (upper bound). The contrast levels for the simulated objects was $\Delta\epsilon = \Delta\sigma = 1$.

An inversion algorithm was implemented based on a quasi-Newton algorithm using the BFGS formula and Golden section line search, see e.g., [8], together with the gradient calculations that are given by (4), (13), (15), (16) and (26) above. The solution to the related direct and adjoint electromagnetic problems were based on an implementation of the FDTD algorithm, see e.g., [30], where the spatial resolution was 10 points per wavelength. A different spatial grid was used for the generation of input data in order to avoid the ‘‘inverse crime’’ [16, 31]. The signal to noise ratio was $\text{SNR} = \frac{1}{4N_0} = 35$ –

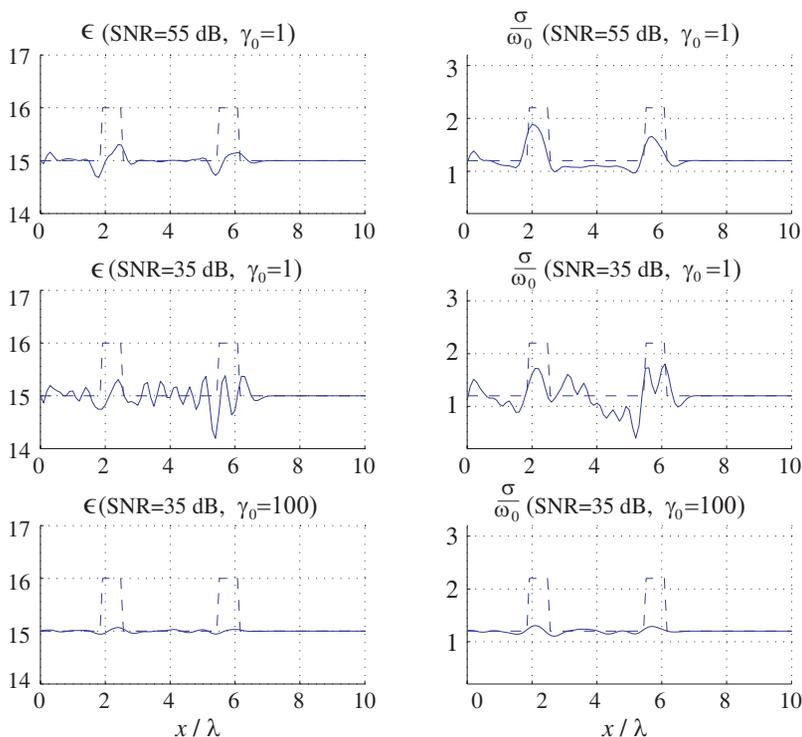


Figure 4. Reconstruction for the one-dimensional inverse problem with linear parameter scaling. The graphs show a reconstruction of the parameters ϵ and σ/ω_0 versus x/λ . The signal to noise ratio is $\text{SNR} = 35$ or $\text{SNR} = 55$ dB and the regularization constant is $\gamma_0 = 1$ or $\gamma_0 = 100$. The true parameter values are shown as a dashed line, where the homogenous background has values $\epsilon = 15$ and $\sigma/\omega_0 = 1.2$ and $\Delta\epsilon = \Delta\sigma = 1$.

55 dB, and artificial noise was added correspondingly prior to the reconstruction.

In Figs. 4 through 6 are shown the numerical reconstructions with linear, exponential and arctangential parameter scalings, respectively. The graphs show a reconstruction of the parameters ϵ and σ/ω_0 versus x/λ . The signal to noise ratio is $\text{SNR} = 35$ or $\text{SNR} = 55$ dB and the regularization constant is $\gamma_0 = 1$ or $\gamma_0 = 100$. The true parameter values are shown as a dashed line.

The reconstruction results in Figs. 4 through 6 should be compared and evaluated against the principal parameter analysis which was discussed above and illustrated in Fig. 3. As was predicted by

the principal parameter analysis, the inversion problem is feasible (reconstruction works reasonably well) when the signal to noise ratio is $\text{SNR} = 55 \text{ dB}$ and is unfeasible (inversion is either unstable or saturated) when $\text{SNR} = 35 \text{ dB}$. Furthermore, with the higher signal to noise ratio $\text{SNR} = 55 \text{ dB}$, the behaviour of the inversion algorithm was (as predicted) rather independent of the regularization constant for $\gamma_0 \leq 1$. In conclusion, the principal parameter analysis worked well as an indicator to whether the inverse problem was feasible or not, as well as an indicator to a proper choice of regularization constant.

However, note also that the reconstruction of the parameter σ/ω_0

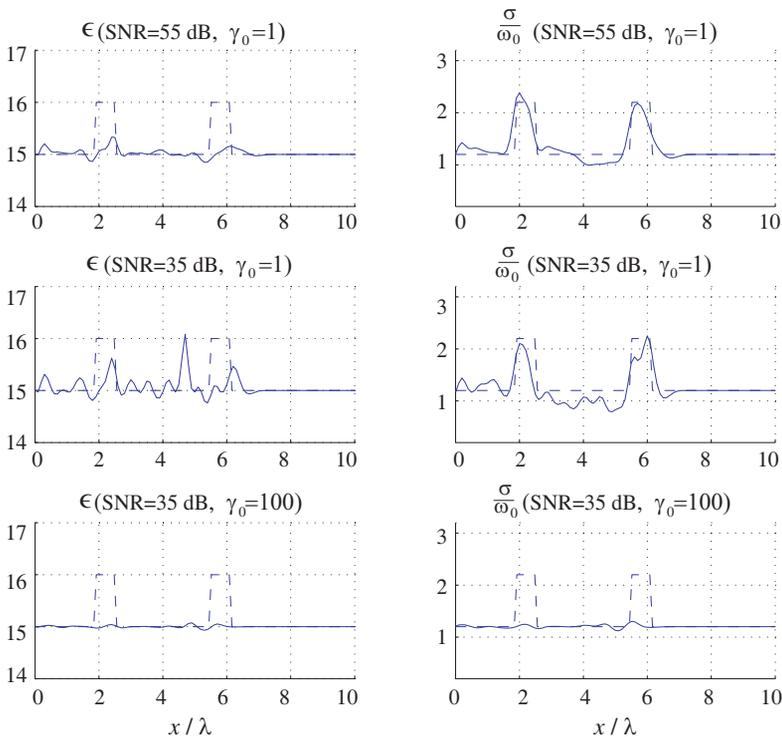


Figure 5. Reconstruction for the one-dimensional inverse problem with exponential parameter scaling. The graphs show a reconstruction of the parameters ϵ and σ/ω_0 versus x/λ . The signal to noise ratio is $\text{SNR} = 35$ or $\text{SNR} = 55 \text{ dB}$ and the regularization constant is $\gamma_0 = 1$ or $\gamma_0 = 100$. The true parameter values are shown as a dashed line, where the homogenous background has values $\epsilon = 15$ and $\sigma/\omega_0 = 1.2$ and $\Delta\epsilon = \Delta\sigma = 1$.

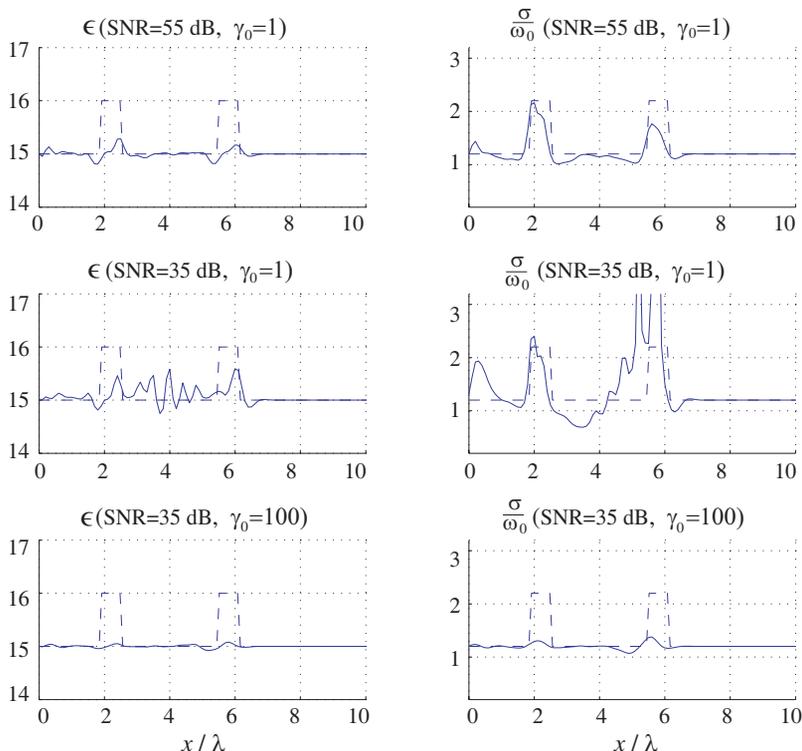


Figure 6. Reconstruction for the one-dimensional inverse problem with arctangential parameter scaling. The graphs show a reconstruction of the parameters ϵ and σ/ω_0 versus x/λ . The signal to noise ratio is SNR = 35 or SNR = 55 dB and the regularization constant is $\gamma_0 = 1$ or $\gamma_0 = 100$. The true parameter values are shown as a dashed line, where the homogenous background has values $\epsilon = 15$ and $\sigma/\omega_0 = 1.2$ and $\Delta\epsilon = \Delta\sigma = 1$.

has better quality than that of ϵ in this numerical example. This is probably due to the fact that the parameter σ/ω_0 has much higher *relative* contrast than ϵ , i.e., $\Delta\sigma/\sigma = 1/0.12$ is much larger than $\Delta\epsilon/\epsilon = 1/15$. As should be expected, this behaviour was not predicted by the principal parameter analysis above. The reason for this is that the principal parameter Fisher information analysis is effectively based on a linearization (first order differentials) about the known background, and is therefore expected to be most reliable for weak scattering problems, and do not take high contrast (nonlinear) effects into account.

As can be seen from the Figs. 4 through 6, the three scaling methods performed rather similar in this numerical example. This may be anticipated from the analysis above since there were no significant differences in the calculated contrast levels $\Delta\xi_\epsilon$ and $\Delta\xi_\sigma$ between the different scaling methods, i.e., the ratio $\Delta\nu_i/\alpha_i$ was reasonable small and the approximation (20) was valid. Hence, the parameter values for the nonlinear exponential and arctangential scalings in this numerical example are mostly active in a region where a linearization

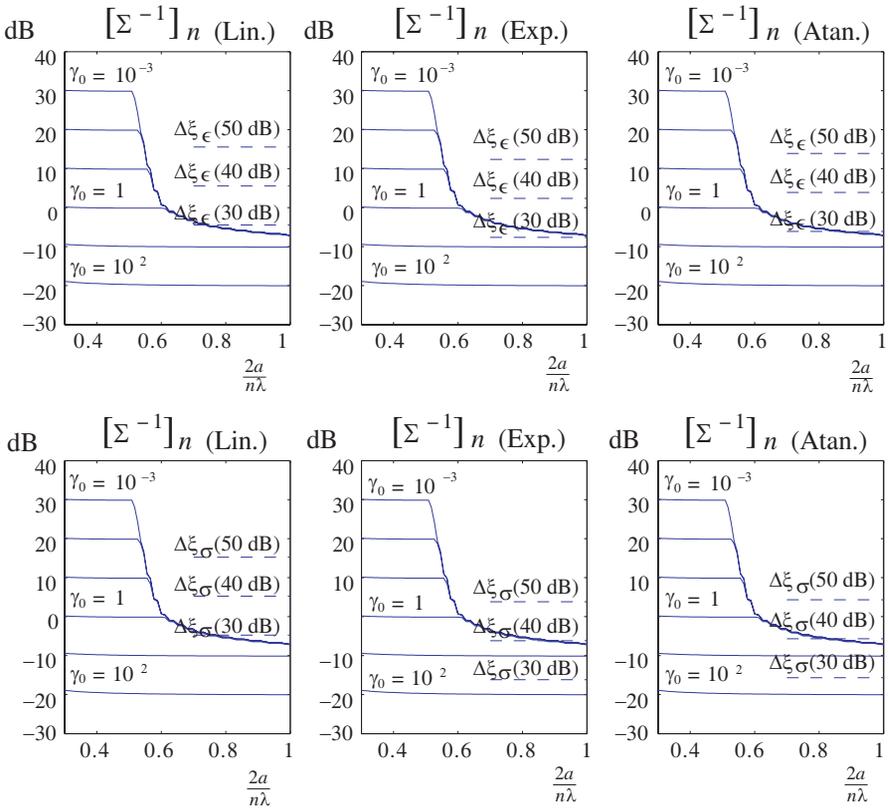


Figure 7. Principal Cramér-Rao bound $[\Sigma^{-1}]_n$ (in dB) as a function of resolution $2a/n\lambda$ for a homogenous background with $\epsilon = 15$ and $\sigma/\omega_0 = 0.12$. Here γ_0 denotes the regularization parameter and $\Delta\xi_\epsilon$ (upper plots) and $\Delta\xi_\sigma$ (lower plots) the contrast levels for the scaled ϵ - and σ -parameters at $x = 6\lambda$ with $\Delta\epsilon = \Delta\sigma = 1$, corresponding to SNR = 30, 40 and 50 dB and linear (Lin.), exponential (Exp.) and arctangential (Atan.) scaling, respectively.

is a reasonable approximation.

5.2. Low Loss Example

An advantage of using a nonlinear parameter scaling with lower (and upper) parameter bounds is that it is a straightforward technique to obtain a robust inversion algorithm that can avoid unphysical solutions such as negative ϵ and σ values and serious parameter divergence problems. This is particularly important when the algorithm is working close to instability, such as with a low signal to noise ratio.

Another situation where the nonlinear scaling is useful is when one of the parameters are very close to its lower (or upper) bound.

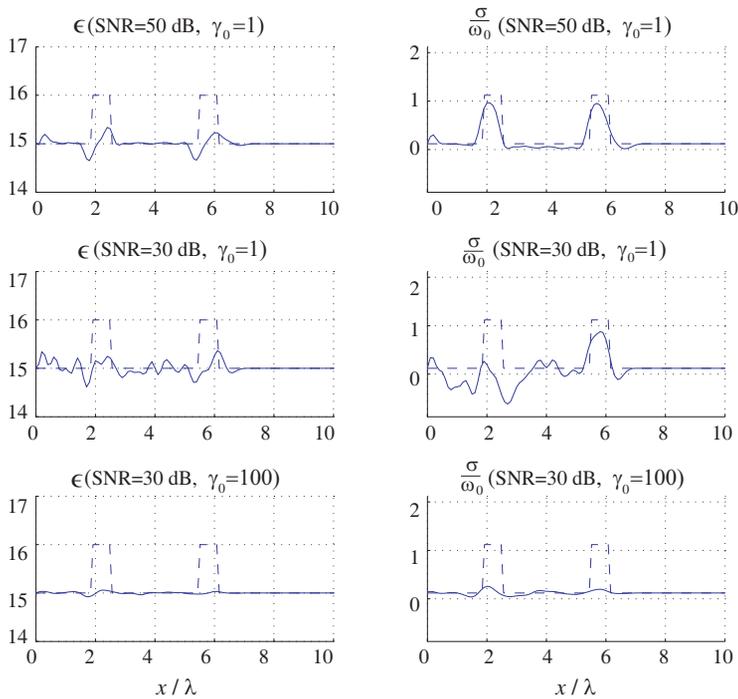


Figure 8. Reconstruction for the one-dimensional inverse problem with linear parameter scaling. The graphs show a reconstruction of the parameters ϵ and σ/ω_0 versus x/λ . The signal to noise ratio is SNR = 30 or SNR = 50 dB and the regularization constant is $\gamma_0 = 1$ or $\gamma_0 = 100$. The true parameter values are shown as a dashed line, where the homogenous background has values $\epsilon = 15$ and $\sigma/\omega_0 = 0.12$ and $\Delta\epsilon = \Delta\sigma = 1$.

To illustrate this, another simulation is presented in Figs. 7 through 10, corresponding to a case with very low losses. Here, the principal parameter analysis and numerical reconstruction is performed precisely as above, except that the conductivity is $\sigma/\omega_0 = 0.04\eta_0/\omega_0 = 0.12$ (in SI-units $\sigma_{SI} = 0.04 \text{ S/m}$).

A priori model parameters for the linear scaling was; $\epsilon_{0i} = 15$ (background) and $\sigma_{0i} = 0.12$ (background), for the exponential scaling; $\alpha_\epsilon = 1$, $\epsilon_{0i} = 14$ (lower bound) and $\alpha_\sigma = 0.12$, $\sigma_{0i} = 0$ (lower bound), and for the arctangential scaling; $\epsilon_l = 14$ (lower bound), $\epsilon_b = 15$ (background), $\epsilon_u = 17$ (upper bound) and $\sigma_l = 0$ (lower bound), $\sigma_b = 0.12$ (background) and $\sigma_u = 100\sigma_b$ (upper bound). The contrast levels for the simulated objects were $\Delta\epsilon = \Delta\sigma = 1$, and the signal to

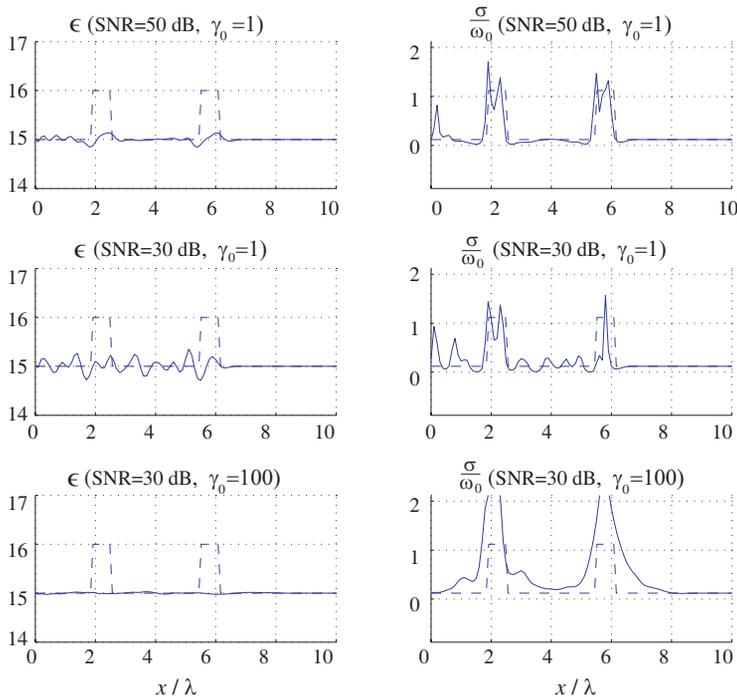


Figure 9. Reconstruction for the one-dimensional inverse problem with exponential parameter scaling. The graphs show a reconstruction of the parameters ϵ and σ/ω_0 versus x/λ . The signal to noise ratio is $\text{SNR} = 30$ or $\text{SNR} = 50$ dB and the regularization constant is $\gamma_0 = 1$ or $\gamma_0 = 100$. The true parameter values are shown as a dashed line, where the homogenous background has values $\epsilon = 15$ and $\sigma/\omega_0 = 0.12$ and $\Delta\epsilon = \Delta\sigma = 1$.

noise ratio ranged between $\text{SNR} = 30$ and $\text{SNR} = 50$ dB.

Again, the principal parameter analysis together with the calculated contrast levels $\Delta\xi_\epsilon$ and $\Delta\xi_\sigma$ in Fig. 7 should be used to evaluate the numerical reconstructions in Figs. 8 through 10. As predicted by the principal parameter analysis, the reconstruction with linear parameter scaling as illustrated in Fig. 8 performs reasonable well with the high signal to noise ratio $\text{SNR} = 50$ dB (since $[\Sigma^{-1}]_n \ll \Delta\xi^2$), and it does not perform well with the low signal to noise ratio $\text{SNR} = 30$ dB (since $[\Sigma^{-1}]_n \approx \Delta\xi_i^2$), i.e., the error variance is high, and the conductivity parameter σ takes on unphysical negative values.

On the other hand, the numerical reconstructions with the

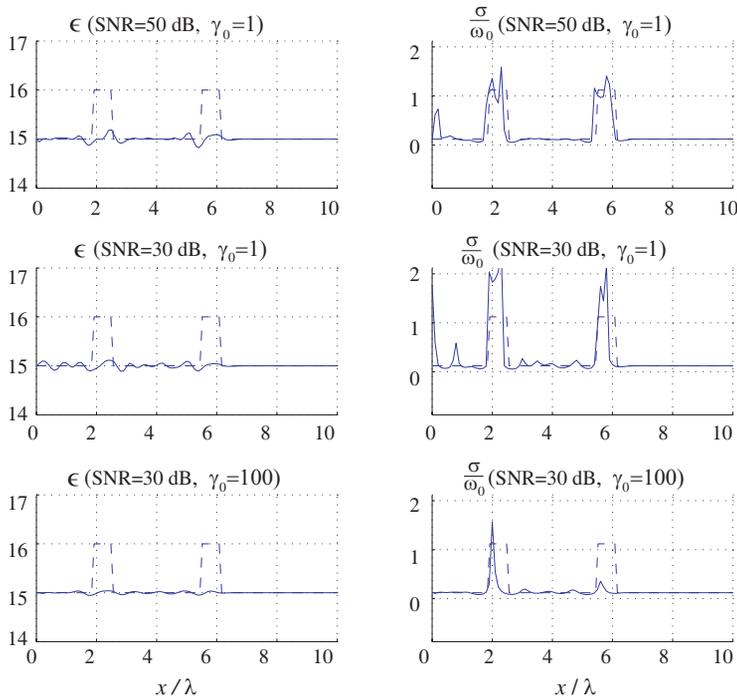


Figure 10. Reconstruction for the one-dimensional inverse problem with arctangential parameter scaling. The graphs show a reconstruction of the parameters ϵ and σ/ω_0 versus x/λ . The signal to noise ratio is $\text{SNR} = 30$ or $\text{SNR} = 50$ dB and the regularization constant is $\gamma_0 = 1$ or $\gamma_0 = 100$. The true parameter values are shown as a dashed line, where the homogenous background has values $\epsilon = 15$ and $\sigma/\omega_0 = 0.12$ and $\Delta\epsilon = \Delta\sigma = 1$.

nonlinear exponential and arctangential scalings as illustrated in Figs. 9 and 10 actually behaves better than what would be expected from the principal parameter analysis illustrated in Fig. 7. The explanation for this is the fact that the principal parameter (Cramér-Rao bound) analysis does not take into account the a priori information that is inherent with the inclusion of a lower (or upper) bound on the parameter values, as is done with the nonlinear parameter scalings. Moreover, this (nonlinear) effect becomes increasingly pronounced as the background parameter values come closer to this lower (or upper) bound.

6. SUMMARY AND CONCLUSIONS

In this paper, a Fisher information based Bayesian approach is presented for analysis and design of the regularization and preconditioning parameters used with gradient based inverse scattering algorithms. In particular, a one-dimensional inverse problem is considered where the permittivity and conductivity profiles are unknown and the input data consist of the scattered field over a certain bandwidth. A priori parameter modeling with linear, as well as nonlinear exponential and arctangential parameter scalings is treated and robust preconditioners are obtained by choosing the related scaling parameters based on a Fisher information analysis of the known background.

The Bayesian approach and a principal parameter (singular value) analysis of the stochastic Cramér-Rao bound is used to investigate the regularization that is necessary to achieve stable inversion, as well as to predict the feasibility of achieving successful reconstruction in a given problem set-up. In particular, the Tikhonov regularization scheme is put into a Bayesian estimation framework. The principal parameter Fisher information analysis is effectively based on a linearization (first order differentials) about the known background, and is therefore expected to be most reliable for weak scattering problems, and do not take high contrast (nonlinear) effects into account.

A time-domain least-squares inversion algorithm based on a quasi-Newton algorithm together with an FDTD-electromagnetic solver has been employed in order to generate the numerical examples. The numerical examples verify the principal parameter analysis by considering low and high noise situations corresponding to feasible and unfeasible inverse problem set-ups, respectively. In a low noise situation, the behaviour of the inversion algorithm is typically independent of the regularization constant if the constant is below a certain limit which can be predicted by the principal parameter

analysis. In a high noise situation, the inverse problem is typically unfeasible (if the noise is high enough) and the regularization will saturate the reconstruction, yielding the a priori known background as an output. As expected, the use of nonlinear a priori parameter models such as the exponential or arctangential scalings will in some cases (when the background parameter values are close to their bounds) yield inversion results that performs much better than what is indicated by the Fisher information analysis. This is quite natural since the Fisher information analysis does not take upper and lower parameter bounds into account.

The Fisher information based principal parameter analysis presented above is based solely on the known background, and it constitutes hence only an indicator to predict the feasibility of achieving successful reconstruction in a given problem set-up. As expected, this indicator is not able to predict the behavior of the algorithms when highly nonlinear effects are present such as with high contrast situations or nonlinear parameter scalings. On the other hand, the Fisher information based technique which is performed here using a uniform background, could be straightforwardly extended to analyze any known background, including high contrast objects.

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support by the Swedish Research Council.

REFERENCES

1. Baussard, A. and O. V. D. Prémel, "A bayesian approach for solving inverse scattering from microwave laboratory-controlled data," *Inverse Problems*, Vol. 17, No. 6, 1659–1669, 2001.
2. Bertero, M., "Linear inverse and ill-posed problems," *Advances in Electronics and Electron Physics*, Vol. 75, 1–120, 1989.
3. Bucci, O. M., L. Crocco, T. Isernia, and V. Pascazio, "Subsurface inverse scattering problems: Quantifying qualifying and achieving the available information," *IEEE Trans. Geoscience and Remote Sensing*, Vol. 39, No. 11, 2527–2538, Nov. 2001.
4. Chew, W. C. and Y. M. Wang, "Reconstruction of twodimensional permittivity distribution using the distorted born iterative method," *IEEE Trans. Med. Imaging*, Vol. 9, No. 2, 218–225, 1990.
5. Colton, D. and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer-Verlag, Berlin, 1992.

6. Fhager, A., M. Gustafsson, S. Nordebo, and M. Persson, "A statistically based preconditioner for two-dimensional microwave tomography," *Proceedings of the Second International Workshop on Computational Advances in Multi-sensor Adaptive Processing*, 173–176, Oct. 2007.
7. Fhager, A. and M. Persson, "Comparison of two image reconstruction algorithms for microwave tomography," *Radio Sci.*, Vol. 40, RS3017, Jun. 2005.
8. Fletcher, R., *Practical Methods of Optimization*, John Wiley & Sons, Ltd., Chichester, 1987.
9. Goharian, M., M. Soleimani, and G. Moran. "A trust region subproblem for 3D electrical impedance tomography inverse problem using experimental data," *Progress In Electromagnetics Research*, PEIR 94, 19–32, 2009.
10. Greenbaum, A., *Iterative Methods for Solving Linear Systems*, SIAM Press, Philadelphia, 1997.
11. Gustafsson, M., *Wave Splitting in Direct and Inverse Scattering Problems*, PhD thesis, Lund University, Department of Electromagnetic Theory, P. O. Box 118, Lund S-22100, Sweden, 2000, <http://www.eit.lth.se>.
12. Gustafsson, M. and S. He, "An optimization approach to twodimensional time domain electromagnetic inverse problems," *Radio Sci.*, Vol. 35, No. 2, 525–536, 2000.
13. Gustafsson, M. and S. Nordebo, "Cramér — Rao lower bounds for inverse scattering problems of multilayer structures," *Inverse Problems*, Vol. 22, 1359–1380, 2006.
14. Habashy, T. M. and A. Abubakar, "A general framework for constraint minimization for the inversion of electromagnetic measurements," *Progress In Electromagnetics Research*, PIER 46, 265–312, 2004.
15. Isakov, V., *Inverse Problems for Partial Differential Equations*, Springer-Verlag, Berlin, 1998.
16. Kaipio, J. and E. Somersalo, *Statistical and Computational Inverse Problems*, Springer-Verlag, New York, 2005.
17. Kay, S. M., *Fundamentals of Statistical Signal Processing, Estimation Theory*, Prentice-Hall, Inc., NJ, 1993.
18. Kelley, C. T., *Iterative Methods for Linear and Nonlinear Equations*, SIAM Press, Philadelphia, 1995.
19. Kirsch, A., *An Introduction to the Mathematical Theory of Inverse Problems*, Springer-Verlag, New York, 1996.
20. Knapp, C. H. and G. C. Carter, "The generalized correlation

- method for estimation of time delay,” *IEEE Trans. Acoustics, Speech, and Signal Process.*, Vol. 24, No. 4, 320–327, 1976.
21. Kristensson, G. and R. J. Krueger, “Direct and inverse scattering in the time domain for a dissipative wave equation. Part 1: Scattering operators,” *J. Math. Phys.*, Vol. 27, No. 6, 1667–1682, 1986.
 22. Marengo, E. A. and R. W. Ziolkowski, “Nonradiating and minimum energy sources and their fields: Generalized source inversion theory and applications,” *IEEE Trans. Antennas Propagat.*, Vol. 48, No. 10, 1553–1562, Oct. 2000.
 23. Nordebo, S., A. Fhager, M. Gustafsson, and M. Persson, “A systematic approach to robust preconditioning for gradient based inverse scattering algorithms,” *Inverse Problems*, Vol. 24, No. 2, 025027, 2008.
 24. Nordebo, S., M. Gustafsson, and B. Nilsson, “Fisher information analysis for two-dimensional microwave tomography,” *Inverse Problems*, Vol. 23, 859–877, 2007.
 25. Nordebo, S., M. Gustafsson, and K. Persson, “Sensitivity analysis for antenna near-field imaging,” *IEEE Trans. Signal Process.*, Vol. 55, No. 1, 94–101, Jan. 2007.
 26. Nordebo, S. and M. Gustafsson, “Statistical signal analysis for the inverse source problem of electromagnetics,” *IEEE Trans. Signal Process.*, Vol. 54, No. 6, 2357–2361, Jun. 2006.
 27. Pierri, R., A. Liseno, and F. Soldovieri, “Shape reconstruction from PO multifrequency scattered fields via the singular value decomposition approach,” *IEEE Trans. Antennas Propagat.*, Vol. 49, No. 9, 1333–1343, Sep. 2001.
 28. Pierri, R. and F. Soldovieri, “On the information content of the radiated fields in the near zone over bounded domains,” *Inverse Problems*, Vol. 14, No. 2, 321–337, 1998.
 29. Soleimani, M., C. N. Mitchell, R. Banasiak, R. Wajman, and A. Adler. “Four-dimensional electrical capacitance tomography imaging using experimental data,” *Progress In Electromagnetics Research*, PEIR 90, 171–186, 2009.
 30. Taflove, A. and S. C. Hagness, *Computational Electrodynamics: The Finite-difference Time-domain Method*, Artech House, Boston, London, 2000.
 31. Tarantola, A., *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, Philadelphia, 2005.
 32. Tipping, M. E., “Sparse Bayesian learning and the relevance

- vector machine,” *J. Mach. Learning Res.*, Vol. 1, No. 2, 211–244, 2001.
33. Van Trees, H. L., *Detection, Estimation and Modulation Theory, Part I*, John Wiley & Sons, Inc., New York, 1968.
 34. Yu, Y. and L. Carin, “Three-dimensional Bayesian inversion with application to subsurface sensing,” *IEEE Trans. Geoscience and Remote Sensing*, Vol. 45, No. 5, 1258–1270, 2007.
 35. Yu, Y., B. Krishnapuram, and L. Carin, “Inverse scattering with sparse Bayesian vector regression,” *Inverse Problems*, Vol. 20, No. 6, S217–S231, 2004.