Journal of Computational and Applied Mathematics 236 (2012) 4264–4275

Contents lists available at SciVerse ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

A maximum product criterion as a Tikhonov parameter choice rule for Kirsch's factorization method

Fermín S.V. Bazán [a], J.B. Francisco [a], Koung Hee Leem [b], G. Pelekanos [b,*]

[a] Department of Mathematics, Federal University of Santa Catarina, Florianópolis SC, Santa Catarina, CEP 88040-900, Brazil

[b] Department of Mathematics and Statistics, Southern Illinois University, Edwardsville, IL 62026, USA

ARTICLE INFO

Article history:
Received 18 February 2011
Received in revised form 3 February 2012

Keywords:
Factorization method
Fixed points
Regularization
SVD-tail

ABSTRACT

Kirsch's factorization method is a fast inversion technique for visualizing the profile of a scatterer from measurements of the far-field pattern. We present a Tikhonov parameter choice approach based on a maximum product criterion (MPC) which provides a regularization parameter located in the concave part of the L-curve on a log–log scale. The performance of the method is evaluated by comparing our reconstructions with those obtained via the L-curve, Morozov's discrepancy principle and the SVD-tail. Numerical results that illustrate the effectiveness of the MPC in reconstruction problems involving both simulated and real data are reported and analyzed.

1. Introduction

The linear sampling method introduced in [1] and further clarified in [2] is one of the major visualization algorithms for solving inverse obstacle scattering problems in the resonance region. The method involves the solution of a linear Fredholm equation of the first kind, the far-field equation, which is written for each point inside the scatterer and whose integral kernel is the far-field pattern, i.e. far-field data that are usually contaminated with significant noise. In addition, it works independently of the kinds of boundary conditions involved and requires no a priori knowledge of the physical properties of the scatterer. Further, unlike for nonlinear methods, factors that can compromise robustness such as local minima and an inaccurate choice of the initial guess are not present. A difficulty with the linear sampling method is that the far-field equation in general is not solvable; however the problem of what happens when the sampling point z is in the exterior of the scatterer was addressed in [3]. This difficulty was remedied by [4] who introduced a new linear sampling method named the factorization method. A rigorous convergence analysis for the case of acoustic scattering by sound-soft obstacles can be found in [4]. For more recent contributions, the reader is referred to [5] and references therein.

It is well known that every numerical implementation of an inverse scattering method requires at some point regularization in order to cope with the ill-posedness of the problem, and the factorization method is not an exception. In most numerical applications of the factorization method, Tikhonov regularization has been employed and the regularization parameter was computed via Morozov's generalized discrepancy principle, which involves the computation of the zeros of the discrepancy function at each point of the grid [2]. In addition, the noise level in the data should be known a priori, something that in real life applications is not the case in general. There exist some alternative parameter selection criteria that avoid knowledge of the noise level and consequently give rise to the so-called heuristic parameter choice rules; these include the generalized cross-validation introduced in [6], the fixed-point method in [7] and the L-curve in [8]; an

* Corresponding author.
E-mail addresses: fermin@mtm.ufsc.br (F.S.V. Bazán), juliano@mtm.ufsc.br (J.B. Francisco), kleem@siue.edu (K.H. Leem), gpeleka@siue.edu (G. Pelekanos).

0377-0427/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2012.05.008

application of the L-curve method in inverse elastic scattering can be found in [9]. More recently, Fares et al. [10] developed a new heuristic algorithm, the SVD-tail one, based on the combined presence of error in the operator and eigenvalue clusters corresponding to a singular subspace associated with a few small singular values. The main feature of the SVD-tail procedure is that the pointwise solution of the far-field equation is never explicitly constructed, and hence the method is computationally efficient; one of its disadvantages though is that the quality of the reconstruction depends on the chosen dimension of the singular subspace. For a fast reconstruction algorithm that rejects the smallest singular values of the finite far-field operator and applies the fixed-point method for selecting the regularization parameter, the reader is referred to Leem et al. [11].

In this work we propose a new Tikhonov parameter choice rule for Kirsch's factorization method named the maximum product criterion (MPC). The new criterion needs computation of the regularized solution norm and the corresponding residual norm and selects the parameter which maximizes the product of these norms as a function of the regularization parameter; its main virtue is that it constructs regularized solutions of either large or small norm depending on whether a certain inclusion condition is satisfied or not. In addition, it does not depend on user specified input parameters (like subspace dimension or truncating parameter) and requires no a priori knowledge of the noise level.

We organize our paper as follows. Section 2 will be devoted to the formulation of the problem and a brief description of the factorization method. Subsequently, Section 3 will deal with the description of MPC as a parameter choice rule, concentrating on certain properties of the L-curve that explain why the criterion is expected to work well. In Section 4, we will be concerned with theoretical properties on which MPC relies as well as with its implementation within the framework of the factorization method. In order to show the effectiveness of our method, in Section 5, we will present numerical examples for the case of impenetrable and penetrable scatterers and we will compare the reconstructions obtained via MPC with the ones obtained by means of the L-curve criterion, the Morozov generalized discrepancy principle and the SVD-tail. In our experiments we will use simulated data obtained by means of the Nyström method [12] as well as real data. The real data are made available by the *Electromagnetics Technology Division, Sensors Directorate, Air Force Research Laboratory, Hanscom Air Force Base, Massachusetts*, and are known by the name of *The Ipswich Data* [13]. We will finally list our conclusions in Section 6.

2. A brief description of Kirsch's factorization method

It is well known that the propagation of time-harmonic acoustic fields in a homogeneous medium, in the presence of a sound-soft obstacle D, is modeled by the exterior boundary value problem (direct obstacle scattering problem)

$$\triangle_2 u(x) + k^2 u(x) = 0, \quad x \in \mathbb{R}^2 \setminus \bar{D} \tag{1}$$

$$u(x) + u^i(x) = 0, \quad x \in \partial D \tag{2}$$

where k is a real positive wavenumber and u^i is a given incident field, that in the presence of D will generate the scattered field u. In addition, the scattered field u will satisfy the Sommerfeld radiation condition

$$\lim_{r\to\infty} \sqrt{r}\left(\frac{\partial u}{\partial r} - iku\right) = 0 \tag{3}$$

where $r = |x|$, $x \in \mathbb{R}^2 \setminus \bar{D}$, and the limit is taken uniformly for all directions $\hat{x} = x/|x|$.

The Green's formula implies that the solution u of the direct obstacle scattering problem above has the asymptotic behavior [12]

$$u(x) = u_\infty(\hat{x})\frac{e^{ikr}}{\sqrt{r}} + O(r^{-3/2}) \tag{4}$$

for some analytic function u_∞, called the far-field pattern of u, given by

$$u_\infty(\hat{x}) = \frac{-e^{i\pi/4}}{\sqrt{8\pi k}} \int_{\partial D} \frac{\partial u}{\partial n}(y) e^{-ik\hat{x}\cdot y}\, ds(y) \tag{5}$$

for $\hat{x} = x/|x|$ on the unit sphere Ω. In the case of the inverse problem, it represents the measured data. In particular, the inverse problem that will be considered here is the problem of finding the shape of D from a complete knowledge of the far-field pattern.

We now define the far-field equation

$$(Fg_z)(\hat{x}) = \frac{e^{i\pi/4}}{\sqrt{8\pi k}} e^{-ik\hat{x}\cdot z}, \quad z \in \mathbb{R}^2 \tag{6}$$

where the right hand side is the far-field pattern of the fundamental solution of the Helmholtz equation given by

$$\Phi(x, z) = \frac{i}{4} H_0^{(1)}(k|x - z|), \quad x \neq z \tag{7}$$

in which $H_0^{(1)}$ is the Hankel function of order zero and of the first kind. Moreover $F : L^2(\Omega) \to L^2(\Omega)$ is given by

$$(Fg)(\hat{x}) = \int_{\Omega} u_\infty(\hat{x}; \hat{d})\, g(\hat{d})\, ds(\hat{d}), \quad \hat{d} \in \Omega. \tag{8}$$

It is well known that the first version of the linear sampling method [1] solves the linear operator Eq. (6) on the basis of the numerical observation that its solution will have a large L^2-norm outside and close to ∂D. Hence, reconstructions are obtained by plotting the norm of the solution. However, the problem is that the right hand side does not in general belong to the range of the operator F. Kirsch [4] was able to overcome this difficulty with the introduction of a new version of the linear sampling method based on appropriate factorization of the far-field operator F. In this method, Kirsch elegantly uses the spectral properties of the operator F to characterize the obstacle. In particular, the following linear operator equation is now used in place of Eq. (6):

$$(F^\star F)^{1/4} g_z = \frac{e^{i\pi/4}}{\sqrt{8\pi k}} e^{-ik\hat{x}\cdot z} \tag{9}$$

and the spectral properties of F are used for the reconstructions. To be more specific, since F is normal and compact, which guarantees the existence of a singular system $\{\sigma_j^c, u_j^c, v_j^c\}, j \in \mathbb{N}$, of F with $v_j^c = s_j u_j^c$ and $s_j \in \mathbb{C}$ with $|s_j| = 1$, then the characterization of the object depends on a range test as described in the following theorem due to Kirsch [4].

Theorem 2.1. *For any $z \in \mathbb{R}^2$ let Φ_z denote the right hand side of Eq. (9). Assume that k^2 is not a Dirichlet eigenvalue of $-\Delta_2$ in D, i.e. the corresponding homogeneous problem has only the trivial solution. Then a point $z \in \mathbb{R}^2$ belongs to D if and only if the series*

$$\sum_{j=1}^{\infty} \frac{|(\Phi_z, v_j^c)|^2}{\sigma_j^c} \tag{10}$$

converges, or equivalently, $z \in D$ if and only if Φ_z belongs to the range of the operator $(F^*F)^{1/4}$.

Moreover, from the main result of [14], if one considers the partial sum of the first n terms of the series, it follows that for large n the partial sum must be large for z outside D and small for z inside. Hence, the characterization of the object can be done by inspecting those z for which the partial sums are large; the same idea is followed when Kirsch's factorization method is implemented in a finite dimensional framework, i.e., in the finite case we deal with a finite dimensional counterpart of (9):

$$(\widetilde{F}_q^\star \widetilde{F}_q)^{1/4} g = r_z, \tag{11}$$

and the object is identified by computing the finite series (or equivalently the squared solution norm of (11))

$$\sum_{j=1}^{n} \frac{|(r_z, v_j)|^2}{\sigma_j}, \tag{12}$$

separating large values from small ones. Here $\widetilde{F}_q$ is a perturbation of a finite dimensional approximation to the far-field operator: $\widetilde{F}_q = F_q + E \in \mathbb{C}^{n\times n}$, where F_q has singular values decaying quickly to zero, n denotes the number of observed incident fields and the number of outgoing directions, and $\{\sigma_j, u_j, v_j\}_{j=1}^n$ denotes a singular system of $\widetilde{F}_q$.

However, since the far-field data are corrupted by noise, the approach may lead to a significant loss of accuracy when the method is applied without regularization. In the next section, we will show that we can overcome this difficulty by using Tikhonov regularization in conjunction with an appropriate parameter choice rule.

3. The maximum product criterion for the selection of the Tikhonov regularization parameter

We start with some notation concerning the Tikhonov method for finite dimensional problems of the form

$$\min_{f\in\mathbb{C}^n} \|b - Af\|_2, \quad A \in \mathbb{C}^{m\times n}\ (m \geq n),\ b \in \mathbb{C}^m, \tag{13}$$

where A is ill-conditioned and has singular values decaying to zero without a particular gap in the singular value spectrum. In its simplest form, Tikhonov's method amounts to replacing problem (13) by

$$\min_{f\in\mathbb{C}^n}\{\|b - Af\|_2^2 + \lambda^2 \|f\|_2^2\} \tag{14}$$

where $\lambda > 0$ is the regularization parameter. Solving (14) is equivalent to solving the *regularized* normal equations

$$(A^*A + \lambda^2 I_n) f = A^* b, \tag{15}$$

from which we obtain the so-called regularized solution $f_\lambda = (A^*A + \lambda^2 I_n)^{-1} A^* b$. Here I_n is the $n \times n$ identity matrix and * denotes the complex conjugate transpose.

Let the singular value decomposition (SVD) of A be

$$A = U\Sigma V^*$$

where $U = [u_1, \ldots, u_m]$, $V = [v_1, \ldots, v_n]$ are orthonormal matrices, and $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$, with $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_p > \sigma_{p+1} = \cdots = 0$, $p = \mathrm{rank}(A) \leq n$. Define $\alpha_i = |u_i^* b|^2$ (the squared Fourier coefficient of b), and $\delta_0 = \|(I - UU^*)b\|_2$ (the size of the incompatible component of b that lies outside the column space of A). Define also $y(\lambda) = \|f_\lambda\|_2^2$, and $x(\lambda) = \|b - Af_\lambda\|_2^2$. Then

$$x(\lambda) = \sum_{i=1}^{p} \frac{\lambda^4 \alpha_i}{(\sigma_i^2 + \lambda^2)^2} + \delta_0^2, \qquad y(\lambda) = \sum_{i=1}^{p} \frac{\sigma_i^2 \alpha_i}{(\sigma_i^2 + \lambda^2)^2}, \tag{16}$$

and the derivative with respect to λ of these (for $\lambda > 0$) is

$$x'(\lambda) = 4\lambda^3 \sum_{i=1}^{p} \frac{\sigma_i^2 \alpha_i}{(\sigma_i^2 + \lambda^2)^3} > 0, \qquad y'(\lambda) = -4\lambda \sum_{i=1}^{p} \frac{\sigma_i^2 \alpha_i}{(\sigma_i^2 + \lambda^2)^3} < 0. \tag{17}$$

Thus

$$dy/dx = -1/\lambda^2. \tag{18}$$

We now turn our attention to Tikhonov regularization in connection with Kirsch's method in a finite dimensional framework. In this case, the problem to be solved is (see (11))

$$\min_{g \in \mathbb{C}^n} \{\|r_z - \widetilde{A}_q g\|_2^2 + \lambda^2 \|g\|_2^2\} \tag{19}$$

for a proper regularization parameter, where $\widetilde{A}_q = (\widetilde{F}_q^* \widetilde{F}_q)^{1/4}$. The choice of λ has been made via Morozov's generalized discrepancy principle (GDP) [1,2,4] where the regularization parameter is shown to depend on the noise in the data; the difficulty with GDP is that in real world problems such a noise level is not always available. In these cases, GDP may lead to significant loss of accuracy when the noise level is not correctly estimated.

Our goal is thus to propose a method for the determination of a regularization parameter without relying on any knowledge concerning the noise level. It is obvious that for a selection parameter rule to be successful in our context, the corresponding regularized solution norm should behave similarly to the solution norm of (11) in the case where the far-field matrix is free of noise. That is, it would be important to select the regularization parameter in such a way that

$$\left[\begin{array}{l} \text{For } z \text{ outside and close to } D \text{ the regularized solution } g_{\lambda,z} \text{ of (19) has} \\ \text{a large norm (12), while for } z \text{ inside } D, g_{\lambda,z} \text{ has a small norm.} \end{array}\right. \tag{P}$$

Our method, to be stated below, relies on the following assumptions:

A1. All singular values of $\widetilde{F}_q$ are distinct and $\widetilde{F}_q$ is non-singular.
A2. The available data contain reliable information in the sense that

$$\|A_q - \widetilde{A}_q\|_2 \ll \|A_q\|_2.$$

Assumption **A1** is supported by the fact that perturbed matrices are rarely singular and rarely have multiple singular values. **A2** essentially means that the noise is much smaller than the signal, a practically necessary condition for any successful calculation.

In order to proceed we shall now introduce some definitions and notation. In fact, for fixed z and $\lambda \geq 0$, let the squared residual norm and the squared solution norm be denoted by $\mathbf{x}(\lambda)$ and $\mathbf{y}(\lambda)$, respectively, that is,

$$\mathbf{x}(\lambda) = \|r_z - \widetilde{A}_q g_{\lambda,z}\|_2^2, \qquad \mathbf{y}(\lambda) = \|g_{\lambda,z}\|_2^2. \tag{20}$$

These functions are monotonic (by virtue of (17)) and satisfy [8]

$$\textbf{(i)}\ 0 < \mathbf{y}(\lambda) \leq \|\widetilde{g}_z\|_2^2, \qquad \textbf{(ii)}\ 0 \leq \mathbf{x}(\lambda) \leq \|r_z\|_2^2. \tag{21}$$

Here $\widetilde{g}_z$ is the unregularized solution of (11). Consider now the function

$$\Psi(\lambda) = \mathbf{x}(\lambda)\mathbf{y}(\lambda), \quad \lambda \geq 0 \tag{22}$$

and observe that because $\mathbf{x}(0) = 0$ due to **A1** and that $\lim_{\lambda \to \infty} \mathbf{y}(\lambda) = 0$ due to (16), it holds that

$$\Psi(0) = 0 = \lim_{\lambda \to \infty} \Psi(\lambda).$$

Since $\mathbf{y}(\lambda)$ is decreasing and $\mathbf{x}(\lambda)$ is increasing, a consequence of the result above concerning Ψ is that there exists a point λ at which Ψ is maximized. An analysis concerning optimality conditions for maximizing Ψ is postponed to the next section; here we restrict ourselves to formally stating/discussing our parameter choice rule, called the *maximum product criterion* (MPC), which consists of selecting as the regularization parameter the maximizer of the function Ψ.

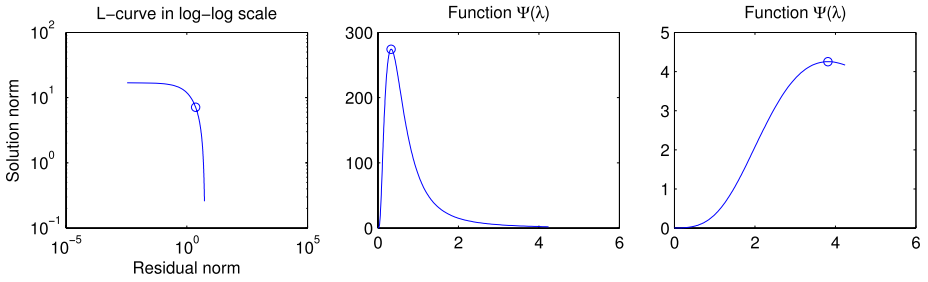

Fig. 1. The L-curve for an inverse scattering problem for a grid point z outside D (left), curve $(\lambda, \Psi(\lambda))$ for a grid point z outside D (center), and curve $(\lambda, \Psi(\lambda))$ for a grid point z inside D (right). The point corresponding to the maximum of Ψ is denoted by a small circle.

As we will see later, the main virtue of MPC is that it delivers small regularization parameters for z outside D (which results in regularized solutions of large norm) and relatively large regularization parameters for z inside (which results in regularized solutions of small norm), that is, the main virtue of MPC is that it enjoys property (P). This can be explained by the fact that there is a close relationship between the curve $(\lambda, \Psi(\lambda))$ and the L-curve on a log–log scale. Let $m_L(\lambda)$ denote the slope of the L-curve at the point $(u(\lambda) = \log \mathbf{x}(\lambda), v(\lambda) = \log \mathbf{y}(\lambda))$. Then it is easy to see that [7]

$$m_L(\lambda) = \frac{dv}{du} = \begin{cases} -\dfrac{\mathbf{x}(\lambda)}{\lambda^2 \mathbf{y}(\lambda)} & \text{if } \lambda > 0, \\ 0 & \text{if } \lambda = 0 \end{cases}$$

where $m_L(0) = 0$ because of assumption **A1**. As a consequence, the L-curve will be flat when $|m_L(\lambda)|$ is small and almost vertical when $|m_L(\lambda)|$ becomes large. Indeed, for the problem under consideration it can readily be proved that the L-curve is concave and flat for λ ranging from zero to the smallest singular value of $\widetilde{A}_q$, and concave for λ larger than or equal to $\|\widetilde{A}_q\|_2$—see, e.g., [15]; a typical L-curve for our problem for z outside D is shown in Fig. 1.

Due to the behavior of the slope m_L it is therefore clear that for values of λ corresponding to the flat part, relatively small changes in the solution norm correspond to relatively large changes in the residual norm, while for values of λ corresponding to the vertical part, large changes in the solution norm correspond to small changes in the residual norm. Hence, the choice of a reasonable regularization parameter should be focused on the region of the L-curve where the slope of the graphs begins to decrease more rapidly and the L-curve becomes vertical, or equivalently, on the region where both norms remain almost unchanged or equilibrated. Our criterion relies on the observation that these norms remain almost unchanged in the region where Ψ is maximized; see Fig. 1. Note that the maximum is relatively large for z outside D and of moderate size for z inside. In addition, note that while the maximizer of Ψ depicted in the center gets small, the maximizer of Ψ depicted on the right gets large.

On the basis of the discussion above, we see that two qualitatively similar imaging parameters can be used in order to identify the scatterer. The first one is given by the mapping

$$z \to \lambda(z) \tag{23}$$

and the second one is given by the mapping

$$z \to W(z) = 1/\|g_{\lambda,z}\|^2. \tag{24}$$

Both imaging parameters are small for z outside D and large for z inside. Hence, the identification of the object is done by plotting the imaging parameters looking for those z for which the parameters are large. A typical behavior of the parameter λ determined according to MPC for an inverse scattering problem, to be discussed in detail later, is shown in Fig. 2.

4. Analysis of MPC

In this section we analyze some supporting properties of MPC. The assertions of following lemma are a consequence of (18) and differentiation with respect to λ.

Lemma 4.1. *For fixed z and $\lambda \geq 0$ define the functions ϕ and φ by*

$$\phi(\lambda) = \frac{\|r_z - \widetilde{A} g_{\lambda,z}\|_2}{\|g_{\lambda,z}\|_2}, \qquad \varphi(\lambda) = -\lambda^2 \mathbf{y}(\lambda) + \mathbf{x}(\lambda).$$

Then the function Ψ in (22) satisfies

$$\Psi'(\lambda) = \varphi(\lambda)\mathbf{y}'(\lambda), \tag{25}$$

$$\Psi''(\lambda) = [-2\lambda + 2\phi(\lambda)\phi'(\lambda)]\mathbf{y}(\lambda)\mathbf{y}'(\lambda) + [-\lambda^2 + \phi^2(\lambda)][\mathbf{y}(\lambda)\mathbf{y}'(\lambda)]'. \quad \square \tag{26}$$

This results in necessary and sufficient conditions for a maximum of Ψ.

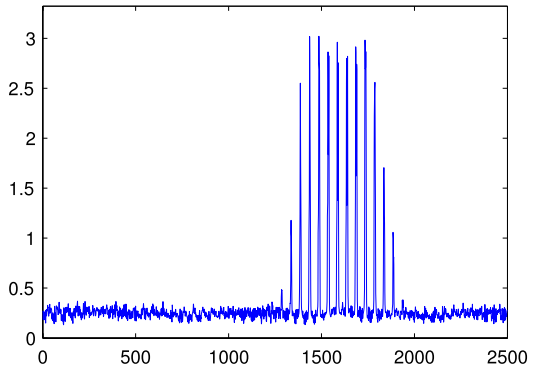

Fig. 2. MPC regularization parameter (in columnwise form) for an inverse scattering problem.

Theorem 4.2. *Let the far-field matrix $\widetilde{F}_q$ have a finite singular system $\{\sigma_j, u_j, v_j\}$. Then the following assertions hold:*

(a) *Function Ψ has at least one critical point inside $(\sqrt{\sigma_n}, \sqrt{\sigma_1})$ and at most one critical point inside $(\frac{\sqrt{3}}{3}\sqrt{\sigma_1}, \sqrt{\sigma_1})$.*
(b) *There exists a local maximum of Ψ at $\widehat{\lambda}$ if and only if $\varphi(\widehat{\lambda}) = 0$ and $\phi'(\widehat{\lambda}) > 1$.*

Proof. Note that because $\widetilde{A}_q$ depends on $\widetilde{F}_q$, its singular system is of the form $\{\sqrt{\sigma_j}, v_j\}$. Hence, the squared solution norm of the Tikhonov problem (19) and the corresponding squared residual norm are (see (16))

$$\mathbf{y}(\lambda) = \sum_{i=1}^{p} \frac{\sigma_i \alpha_i}{(\sigma_i + \lambda^2)^2}, \qquad \mathbf{x}(\lambda) = \sum_{i=1}^{p} \frac{\lambda^4 \alpha_i}{(\sigma_i + \lambda^2)^2}, \tag{27}$$

where $\alpha_i = |v_i^* r_z|^2$. Thus

$$\varphi(\lambda) = -\lambda^2 \mathbf{y}(\lambda) + \mathbf{x}(\lambda) = \sum_{j=1}^{n} \frac{\lambda^2(\lambda^2 - \sigma_i)\alpha_i}{(\sigma_i + \lambda^2)^2}.$$

But since the squared solution norm $\mathbf{y}(\lambda)$ decreases with λ, which implies that $\mathbf{y}'(\lambda) \neq 0,\ \lambda > 0$, by virtue of (17), it turns out that the critical points of Ψ, if any, are roots of φ due to (25). Now note that $\varphi(\lambda) \geq 0$ if $\lambda \geq \sqrt{\sigma_1}$ and $\varphi(\lambda) \leq 0$ if $\lambda \leq \sqrt{\sigma_n}$. Note also that $\sqrt{\sigma_1}$ and $\sqrt{\sigma_n}$, respectively, cannot be roots of φ since by assumption **A1** all singular values are distinct. This implies that φ must have at least a root inside $(\sqrt{\sigma_n}, \sqrt{\sigma_1})$. On the other hand, using (18) and (27) we get

$$\begin{aligned} \varphi'(\lambda) &= -2\lambda \mathbf{y}(\lambda) - \lambda^2 \mathbf{y}'(\lambda) + \mathbf{x}'(\lambda) \\ &= -2[\lambda \mathbf{y}(\lambda) + \mathbf{x}'(\lambda)] = -2 \sum_{j=1}^{n} \frac{\lambda \sigma_i (3\lambda^2 - \sigma_i)\alpha_i}{(\sigma_i + \lambda^2)^3}, \end{aligned}$$

where the last equality is obtained by computing $\mathbf{x}'(\lambda)$ using (27) and then rearranging the terms. This shows that $\varphi'(\lambda)$ does not change sign inside $(\frac{\sqrt{3}}{3}\sqrt{\sigma_1}, \sqrt{\sigma_1})$. Thus, if Ψ has a critical point inside this interval, this critical point is unique and the proof of (a) is complete.

To prove assertion (b), observe from (25) that $\widehat{\lambda}$ is a critical point of Ψ if and only if $\varphi(\widehat{\lambda}) = 0$, this being true if and only if $\phi(\widehat{\lambda}) = \widehat{\lambda}$. Therefore, at the point $\widehat{\lambda}$, by (26) we have $\Psi''(\widehat{\lambda}) < 0$ iff $\phi'(\widehat{\lambda}) > 1$, and (b) is proved. □

Corollary 1. *Assume that Ψ has a local maximum at $\widehat{\lambda}$. Then there exists a vicinity of $\widehat{\lambda}$ on which the L-curve on a log–log scale is locally concave.*

Proof. We know from [16] that the L-curve is locally concave at λ if and only if $\phi'(\lambda) > \frac{\phi(\lambda)}{\lambda}$. Now observe that since Ψ has a local maximum at $\widehat{\lambda}$, then $\phi(\widehat{\lambda}) = \widehat{\lambda}$ (as $\varphi(\widehat{\lambda}) = 0$) and $\phi'(\widehat{\lambda}) > 1 = \frac{\phi(\widehat{\lambda})}{\widehat{\lambda}}$, which proves the corollary. □

For problem (14) with general matrix A satisfying $\delta_0 = 0$, it is known that the associated L-curve is concave on $(0, \sigma_n) \cup (\sigma_1, \infty)$ [15]; nothing very specific can be said about concavity/convexity when $\lambda \in [\sigma_n, \sigma_1]$ [16,15]. However, for the inverse scattering problem under consideration, Corollary 1 indicates that the L-curve is concave on $(0, \sqrt{\sigma_n}) \cup (\sqrt{\sigma_1}, +\infty)$ and locally concave on $[\sqrt{\sigma_n}, \sqrt{\sigma_1}]$. This suggests that in inverse scattering the L-curve may be concave in all domains and justifies the behavior of the L-curve plot displayed in Fig. 1.

We now state a theorem that relates the regularized solution norm associated with MPC and the one determined by the generalized discrepancy principle. For this, set $\epsilon = \|\widetilde{A}_q - A_q\|_2$ and use singular value perturbation theory [17] to obtain

$$\sqrt{\sigma_i^{(e)}} - \epsilon \leq \sqrt{\sigma_i} \leq \sqrt{\sigma_i^{(e)}} + \epsilon, \quad i = 1, \ldots, n$$

where $\sigma_i^{(e)}$ denotes the ith singular value of F_q. Next, since $\sigma_n^{(e)}$ usually vanishes for moderate n, deduce that a reasonable lower bound for ϵ is $\sqrt{\sigma_n} \le \epsilon$. Finally, on the basis of Theorem 4.2-(a) conclude that, if $\sqrt{\sigma_n} < \epsilon$ then the maximizer of Ψ belongs either to $(\sqrt{\sigma_n}, \epsilon]$ or to $[\epsilon, \sqrt{\sigma_1})$.

Theorem 4.3. *For fixed z let $\lambda_{\mathrm{MPC}}^{(z)}$ and $\lambda_{\mathrm{GDP}}^{(z)}$ denote the regularization parameter associated with MPC and GDP, respectively. Assume that $\sqrt{\sigma_n} < \epsilon$. Then*

$$\|g_{\lambda_{\mathrm{MPC}}^{(z)}}\|_2 \le \|g_{\lambda_{\mathrm{GDP}}^{(z)}}\|_2 \le \|g_z\|_2, \quad \text{if } \lambda_{\mathrm{MPC}}^{(z)} \ge \epsilon, \tag{28}$$

where g_z denotes the solution of (11) when the far-field matrix is free of noise, while if $\lambda_{\mathrm{MPC}}^{(z)} < \epsilon$, there exists a number $K = K(z)$, $K > 1$, such that

$$\|g_{\lambda_{\mathrm{GDP}}^{(z)}}\|_2 \le \|g_{\lambda_{\mathrm{MPC}}^{(z)}}\|_2 \le \sqrt{\frac{1+K^2}{2}}\,\|g_z\|_2. \tag{29}$$

Proof. The right inequality in (28) is a well-known result attributed to Morozov's discrepancy principle that holds independently of the condition $\lambda_{\mathrm{MPC}}^{(z)} \ge \epsilon$ [4]. So it remains to prove the left inequality. For this, observe that while for DP we have [4] $\phi(\lambda_{\mathrm{GDP}}^{(z)}) = \|\widetilde{A}_q - A_q\|_2$, for MPC we have $\phi(\lambda_{\mathrm{MPC}}^{(z)}) = \lambda_{\mathrm{MPC}}^{(z)}$. Thus $\phi(\lambda_{\mathrm{MPC}}^{(z)}) \ge \phi(\lambda_{\mathrm{GDP}}^{(z)})$, where we have used the assumption that $\lambda_{\mathrm{MPC}}^{(z)} \ge \epsilon$. But since ϕ is an increasing function [16], then $\lambda_{\mathrm{MPC}}^{(z)} \ge \lambda_{\mathrm{GDP}}^{(z)}$ and the required inequality follows on noting that the regularized solution norm is decreasing.

We now prove the right inequality in (29). In fact, since $\varphi(\lambda_{\mathrm{MPC}}^{(z)}) = 0$ implies

$$[\lambda_{\mathrm{MPC}}^{(z)}]^2 \|g_{\lambda_{\mathrm{MPC}}^{(z)}}\|_2^2 = \|r_z - \widetilde{A}_q\, g_{\lambda_{\mathrm{MPC}}^{(z)}}\|_2^2,$$

and since the Tikhonov functional (19) is minimized at $g_{\lambda_{\mathrm{MPC}}^{(z)}}$, we have

$$\begin{aligned} 2[\lambda_{\mathrm{MPC}}^{(z)}]^2 \|g_{\lambda_{\mathrm{MPC}}^{(z)}}\|_2^2 &= \|r_z - \widetilde{A}_q\, g_{\lambda_{\mathrm{MPC}}^{(z)}}\|_2^2 + [\lambda_{\mathrm{MPC}}^{(z)}]^2 \|g_{\lambda_{\mathrm{MPC}}^{(z)}}\|_2^2 \\ &\le \|r_z - \widetilde{A}_q\, g_z\|_2^2 + [\lambda_{\mathrm{MPC}}^{(z)}]^2 \|g_z\|_2^2 \\ &= \|(A_q - \widetilde{A}_q) g_z\|_2^2 + [\lambda_{\mathrm{MPC}}^{(z)}]^2 \|g_z\|_2^2, \end{aligned}$$

which implies

$$2[\lambda_{\mathrm{MPC}}^{(z)}]^2 \|g_{\lambda_{\mathrm{MPC}}^{(z)}}\|_2^2 \le (\epsilon^2 + [\lambda_{\mathrm{MPC}}^{(z)}]^2) \|g_z\|_2^2. \tag{30}$$

Thus, if $\lambda_{\mathrm{MPC}}^{(z)} < \epsilon$, there exists a constant $K > 1$ such that $\epsilon = K\lambda_{\mathrm{MPC}}^{(z)}$, and the required inequality follows from (30). Finally, if $\epsilon > \lambda_{\mathrm{MPC}}^{(z)} = \phi(\lambda_{\mathrm{MPC}}^{(z)})$, as ϕ is an increasing function $\lambda_{\mathrm{MPC}}^{(z)} \le \lambda_{\mathrm{GDP}}^{(z)}$, and the left inequality in (28) follows from the monotonic behavior of the regularized solution norm. □

Recall that the identification of the object via Kirsch's method depends on a separation of large imaging parameters from small ones. Hence, if we consider the imaging parameters (23) and (24) associated with MPC and GDP, respectively, a consequence from Theorem 4.3 is that such a separation is more appealing in the MPC case than in the GDP case. To be more precise, assume first that $\lambda_{\mathrm{MPC}}^{(z)} < \epsilon$ (which is likely to occur for z outside D). Then it follows from (29) that the MPC's imaging parameters are smaller than those corresponding to GDP. On the other hand, if $\lambda_{\mathrm{MPC}}^{(z)} > \epsilon$ (which is likely to occur for z inside D), then inequality (28) indicates that the MPC's imaging parameters are larger than those corresponding to GDP. A typical behavior of the imaging parameters (23) for MPC and GDP, respectively, obtained when reconstructing a kite, to be described later, is illustrated in Fig. 3.

We end the section with the observation that the maximizer of Ψ can be determined as a root of the function $\varphi(\lambda)$ introduced in Lemma 4.1 by using, e.g., the Newton method, the secant method or another approach. Alternatively, as the maximizer of Ψ is a fixed point of ϕ, this maximizer can be determined by using the improved fixed-point method described in [16]. A further alternative is to use optimization methods in several ways. For example, if we define for each z the function

$$h(\lambda) = (\lambda - \phi(\lambda))^2, \qquad \sqrt{\sigma_n} \le \lambda \le \sqrt{\sigma_1}, \tag{31}$$

then the maximizer of Ψ can be determined by finding the minimum of h. We will always rely on the assumption that the L-curve is strictly concave in all domains, as illustrated in Fig. 1; when this is the case, following the analysis in [16] it is easy to prove that Ψ has a unique maximizer that is a nonzero fixed point of ϕ. For further details concerning convexity/concavity properties of the L-curve, the reader is referred to [16].

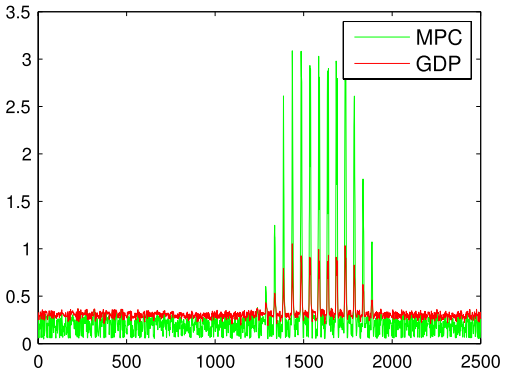

Fig. 3. Regularization parameters (in columnwise form) for an inverse scattering problem.

5. Numerical applications

In this section we shall illustrate the effectiveness of an implementation of MPC where the maximizer of Ψ is determined as a root of $\varphi(\lambda)$ by using the regula falsi method.

5.1. Reconstructions using synthetic data

In order to simulate perturbed data, we generate Gaussian random matrices E_1, E_2 and use a far-field matrix defined by

$$\widetilde{F}_q = F_q + \delta(E_1 + iE_2)\|F_q\|_2 \quad \delta > 0,$$

where δ is given and where F_q is constructed by using the Nyström method [12]. As indicated in [12], the Nyström method not only requires less computational effort compared to collocation and Galerkin methods but also is generically stable in the sense that it preserves the condition of the integral equation. We first report an image reconstruction experiment of a sound-soft obstacle. In this case the object to be reconstructed is a kite located in a grid of 50×50 points, the far-field matrix $\widetilde{F}_q$ is 21×21 (i.e., we use 21 incident and observed directions), and the relative noise level in $\widetilde{F}_q$ is 1% (which implies a relative noise level in $\widetilde{A}_q$ of approximately 8%).

The profile of the kite and its reconstructed version obtained by the L-curve method are shown in Fig. 4(a) and (b). The regularization parameter for the L-curve is determined via the function l_curve which is included in the Regularization Toolbox in [18]. Reconstructions of the kite via MPC are also shown in Fig. 4(c) and (d), in which, for comparison purposes, we have included reconstructions obtained through the imaging parameters (23) and (24). Note that the only difference between the two objects shown at the bottom of Fig. 4 is their size, whereas the qualities of the reconstructions are essentially the same.

We also illustrate the effectiveness of MPC when noise contamination is high. To this end we consider the same reconstruction problem but with the noise level in $\widetilde{F}_q$ increased to 20% (which implies a relative noise level in $\widetilde{A}_q$ of 38%). This example is also useful for illustrating that Morozov's discrepancy principle (GDP) can fail if the noise level is not correctly estimated. The results of the experiment displayed in Fig. 5 show that GDP leads to a significant loss of accuracy, as expected, a result that does not occur with MPC which still manages to deliver a reasonable reconstruction.

So far we have presented numerical experiments in which MPC was compared with relatively old parameter selection algorithms like the L-curve and the Morozov discrepancy principle. For the sake of completeness we will now compare MPC with the recently proposed SVD-tail procedure [10]. For the convenience of the reader and ease of presentation of the results we will start with a brief review of the SVD-tail procedure. Essentially, the SVD-tail procedure selects the dimension d of a subspace spanned by the d left singular vectors of $\widetilde{F}_q$ corresponding to the d smallest singular values, and uses an imaging parameter defined by the map

$$z \to \tau(z) = 1/\|\theta_1(z), \ldots \theta_d(z)\|_2, \tag{32}$$

where the $\theta_j(z)$ are the Fourier coefficients of r_z with respect to the basis of the chosen subspace. In order to motivate the use of $\tau(z)$ as an imaging parameter, Fares et al. [10] argue that the discrete Picard condition cannot be assumed in this case and proceed by remarking that for large j the values of $|\theta_j(z)|$ stabilize, remaining small for z inside the object and significantly nonzero for z outside; see Figure 1 in [10].

Fig. 6 shows the reconstruction of the kite from 300 observed incident directions and data contaminated by 5% noise. We illustrate the method for $d = 5, 10, 15, 20, 25$. It becomes apparent that the SVD-tail procedure does not yield satisfactory results. In order for the SVD-tail procedure to work it is important for the Fourier coefficients θ_j for z inside the object to differ significantly in magnitude from those for z outside. In particular, the Fourier coefficients for interior points should be very close to zero but significantly larger for exterior ones [10]. A closer look at the Fourier coefficients for the kite problem (see Fig. 7) reveals that there is no significant change in their magnitude for either interior or exterior z. In addition, the

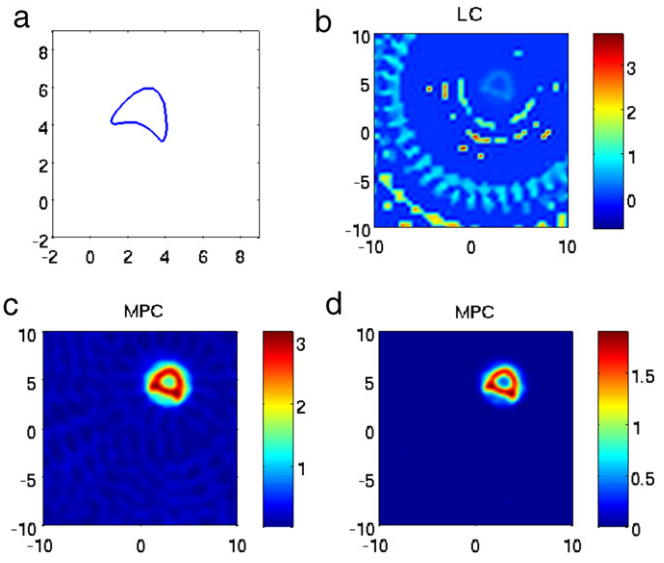

Fig. 4. (a) Profile of a kite. (b) Image of the map $z \to 1/\|g_{\lambda_{LC}^{(z)}}\|_2^2$. (c) Image of the map $z \to \lambda_{MPC}^{(z)}$. (d) Image of the map $z \to 1/\|g_{\lambda_{MPC}^{(z)}}\|_2^2$.

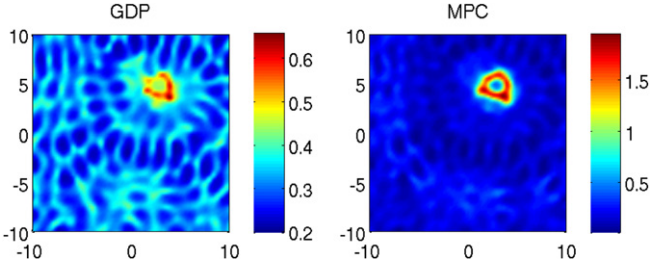

Fig. 5. Results from Morozov's generalized discrepancy criterion and MPC for highly noisy data. For GDP we use 0.15ϵ as an estimate of the noise level.

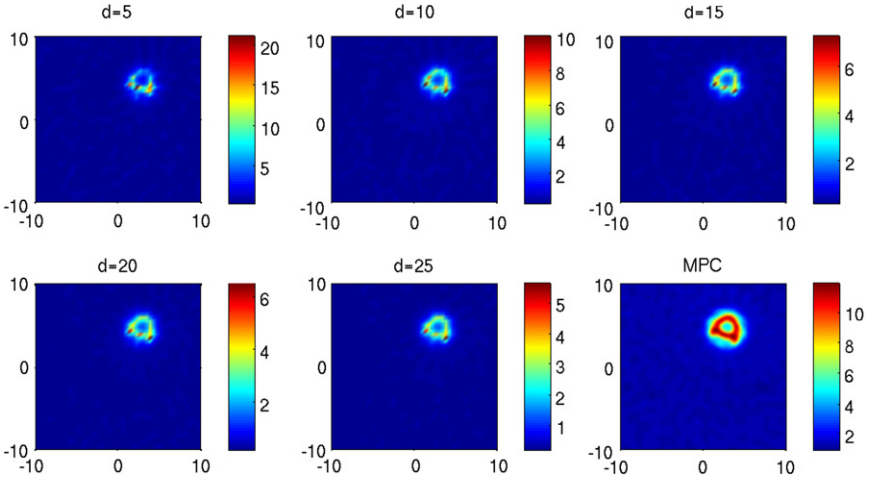

Fig. 6. Comparison of the SVD-tails for different values of the smallest d singular values of the coefficient matrix and MPC for the kite.

coefficients corresponding to an interior point z are a considerable distance from 0. This is what the authors believe is the reason for the SVD-tail procedure not producing favorable reconstructions. We now continue our numerical experiments by attempting to reconstruct two sound-soft obstacles. In this case the objects to be reconstructed are two circles located in a grid of 50×50, the far-field matrix $\widetilde{F}_q$ is again 21×21, and the relative noise level in $\widetilde{F}_q$ is 25% (which implied a relative noise level in $\widetilde{A}_q$ of 38%). In this experiment we consider GDP in two distinct circumstances: (i) when the exact error norm $\|\widetilde{A}_q - A_q\|_2 = \epsilon$ is used as input data and (ii) when the error norm is underestimated and set to 0.08ϵ. The results are depicted

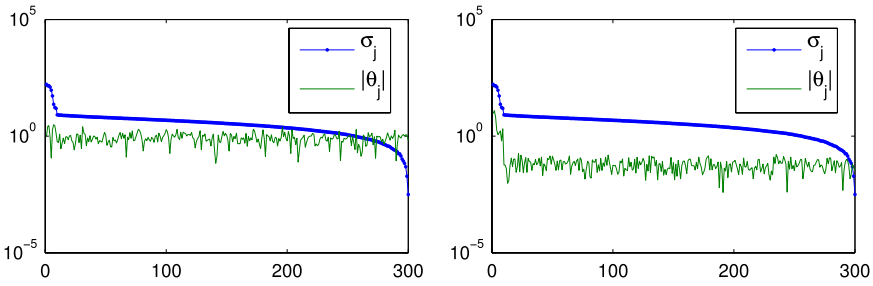

Fig. 7. Fourier coefficients $\theta_j = u_j^\star r_z$ and σ_j as a function of j for a point external (left) or internal (right) with respect to the kite.

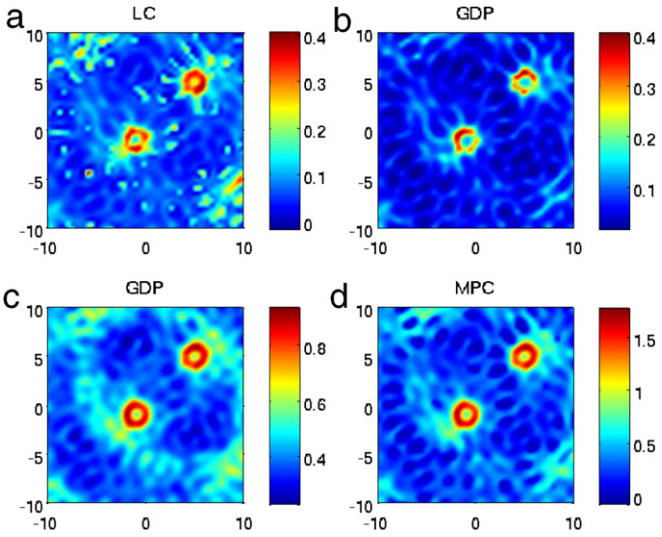

Fig. 8. (a) Image of the map $z \to 1/\|g_{\lambda_{LC}^{(z)}}\|_2^2$. (b) Image of the map $z \to 1/\|g_{\lambda_{GDP}^{(z)}}\|_2^2$ (case (ii)) (c) Image of the map $z \to \lambda_{GDP}^{(z)}$ (case (i)). (d) Image of the map $z \to \lambda_{MPC}^{(z)}$.

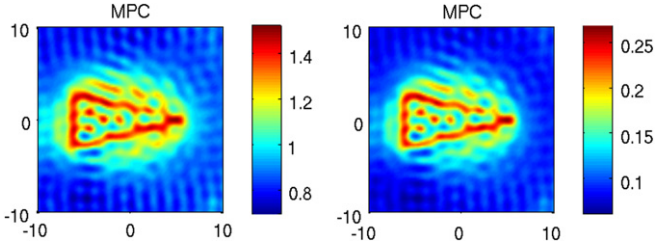

Fig. 9. Images of the maps $z \to \lambda_{MPC}^{(z)}$ (left) and $z \to 1/\|g_{\lambda_{MPC}^{(z)}}\|_2^2$ (right).

in Fig. 8. We note that while both MPC and GDP (case (i)) produce reasonable reconstructions and outperform the L-curve method (as expected), GDP (case (ii)) yields a reconstruction of poor quality due to the lack of prior information about the noise level, and once more we see that MPC may outperform GDP if the noise level is not correctly estimated.

We end this section with the observation that the SVD-tail was also used in the two-object case. The quality of the results was very similar to that of those for the kite case, which explains why the results are not shown here.

5.2. Reconstruction using real data

We will now consider real data sets (The Ipswich Data) produced by using an echo-free chamber, a fixed transmitter and a receiver rotating around the scatterer. The incident and observation angles number 36 for both experiments. Initially we will attempt to reconstruct an aluminum triangle (IPS009) whose outer circle has radius equal to 6 cm using MPC and the SVD-tail. Like in the above, we will illustrate the SVD-tail method for values of $d = 5$, 10, 15, 20, 25 and 30.

Examining the series of reconstructions that appear in the bottom of Fig. 10 we see that the quality of the image improves as we increase the dimension d; however. comparing those reconstructions with the ones obtained via MPC (see Fig. 9) it

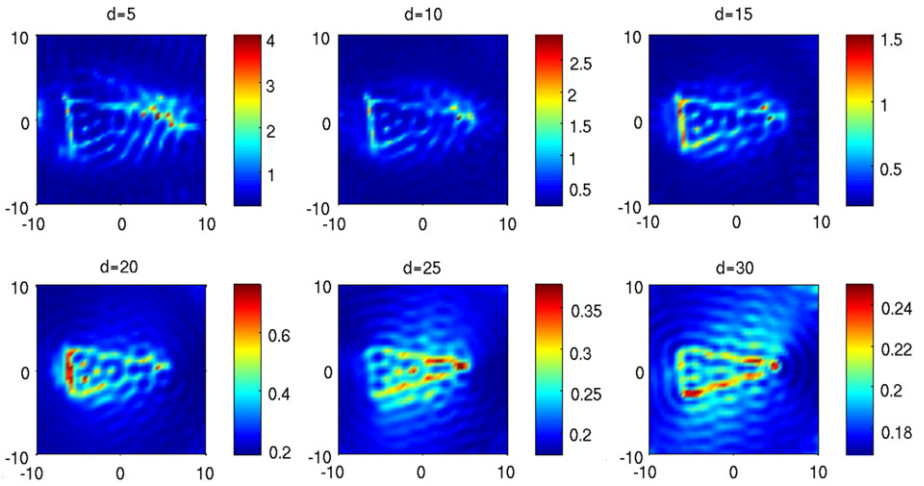

Fig. 10. Images obtained using the SVD-tail for different values of the smallest d singular values of the far-field matrix.

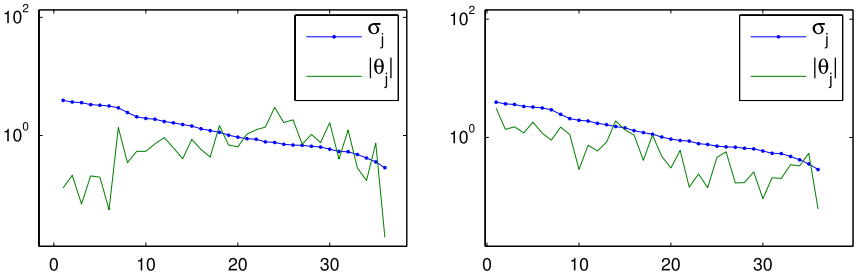

Fig. 11. Fourier coefficients $\theta_j = u_j^* r_z$ and σ_j as a function of j for a point external (left) or internal (right) with respect to the triangle.

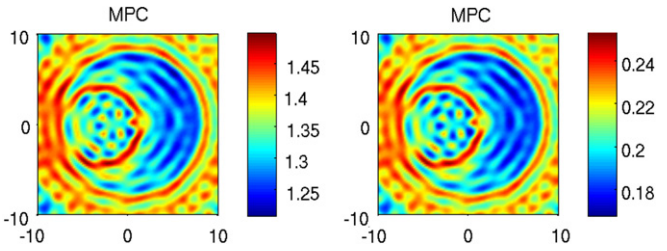

Fig. 12. Top: images of the maps $z \to \lambda_{\mathrm{MPC}}^{(z)}$ (left) and $z \to 1/\|g_{\lambda_{\mathrm{MPC}}^{(z)}}\|_2^2$ (right).

becomes apparent that the reconstructions obtained via the MPC capture more details of the image. For example, the base and the top side of the triangle appear more distinct in the MPC case. Examining the behavior of the Fourier coefficients (see Fig. 11) we observe that for an interior point z the Fourier coefficients assume values significantly larger than zero, which contradicts an essential requirement of the SVD-tail procedure. Moreover, the authors believe that for small size problems (like the one above), eliminating Fourier coefficients corresponding to large singular values may result in losing significant information, and hence the use of the SVD-tail may not be appropriate.

We finally attempt the reconstruction of the "mystery" object given by the data set (IPS007), with the a priori information that the object is penetrable and lies inside a circle of radius 7.5 cm. The reconstruction results appear in Figs. 12 and 13 and, as clearly indicated by the MPC, the object was a circular tube, with a smaller one in its interior. Unfortunately, the SVD-tail approach failed to clearly reconstruct the profiles of the two circles. It is worth mentioning here that reconstructions via Morozov's discrepancy principle are difficult to obtain for real data cases due to the fact that the level of error in the experimental far-field matrix is not available or difficult to estimate.

6. Conclusions

We have introduced a new technique for determining the regularization parameter for Kirsch's factorization method that works well without any a priori information on the noise level in the far-field matrix. Numerical experiments indicate

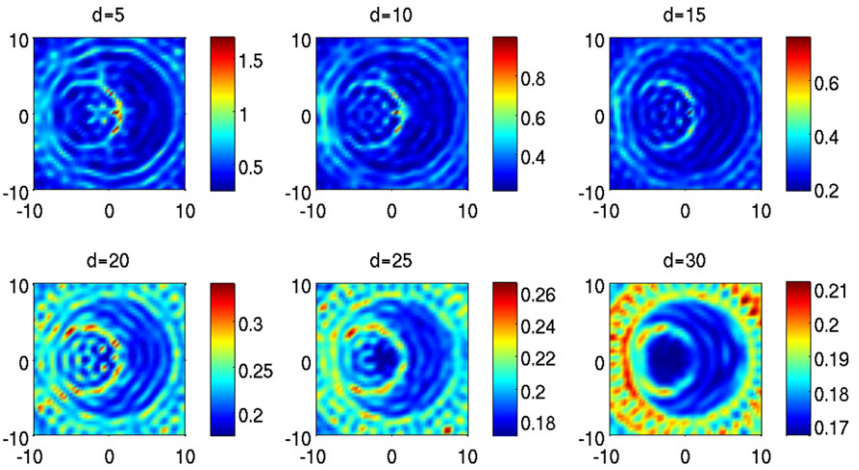

Fig. 13. Images obtained using the SVD-tail for different values of the smallest d singular values of the far-field matrix.

that the computing cost is modest (comparing to that for the L-curve) whereas the quality of reconstructions is very good. In particular, we showed that our approach outperforms that based on Morozov's discrepancy principle if the noise level is not correctly estimated, a situation which is very likely in real world problems. In addition, reconstructions from real data were presented and the results compared with the ones obtained via the SVD-tail. Future research will include further modifications of the MPC such that it can be applied to three-dimensional reconstructions.

Acknowledgment

The authors would like to thank Dr. Michele Piana for kindly providing them with the real data sets.

References

[1] D. Colton, A. Kirsch, A simple method for solving the inverse scattering problems in the resonance region, Inverse Problems 12 (1996) 383–393.
[2] D. Colton, M. Piana, R. Potthast, A simple method using Morozov's discrepancy principle for solving inverse scattering problems, Inverse Problems 13 (1999) 1477–1493.
[3] F. Caconi, D. Colton, On the mathematical basis of the linear sampling method, Georg. Math. J. 10 (2003) 911–925.
[4] A. Kirsch, Characterization of the shape of a scattering obstacle using the spectral data of the far-field operator, Inverse Problems 14 (1998) 1489–1512.
[5] A. Kirsch, N. Grinberg, The Factorization Method for Inverse Problems, Oxford University Press, Oxford, 2008.
[6] G. Wahba, Spline Models for Observational Data, SIAM, Philadelphia, 1990.
[7] F.S. Viloche Bazán, Fixed-point iterations in determining the Tikhonov regularization parameter, Inverse Problems 24 (2008) 1–15.
[8] P.C. Hansen, Rank-Deficient and Discrete Ill-Posed Problems, SIAM, Philadelphia, 1998.
[9] G. Pelekanos, V. Sevroglou, Shape reconstruction of a 2D-elastic penetrable object via the L-curve method, Journal of Inverse and Ill-Posed Problems 14 (2006) 1–16.
[10] M. Fares, S. Gratton, P. Toint, SVD-tail: a new linear-sampling reconstruction method for inverse scattering problems, Inverse Problems 25 (2009) 1–19.
[11] K.H. Leem, G. Pelekanos, F.S. Viloche Bazán, Fixed-point iterations in determining a Tikhonov regularization parameter in Kirsch's factorization method, Applied Mathematics and Computations 216 (2010) 3747–3753.
[12] D. Colton, R. Kress, Inverse Acoustic and Electromagnetic Scattering Theory, Springer-Verlag, New York, 1992.
[13] R. McGahan, R. Kleinman, Third annual special session on image reconstruction using real data, part 1, IEEE Antennas Propagation Magazine 41 (1999) 34–36.
[14] A. Lechleiter, A regularization technique for the factorization method, Inverse Problems 22 (2006) 1605–1625.
[15] T. Regińska, A regularization parameter in discrete ill-posed problems, SIAM Journal on Scientific Computing 3 (1996) 740–749.
[16] F.S. Viloche Bazán, J.B. Francisco, An improved fixed-point algorithm for determining a Tikhonov regularization parameter, Inverse Problems 25 (2009).
[17] G.H. Golub, C.F. Van Loan, Matrix Computations, third ed., The Johns Hopkins University Press, Baltimore, MD, 1996.
[18] P.C. Hansen, Regularization tools: a matlab package for analysis and solution of discreet ill-posed problems, Numerical Algorithms 6 (1994) 1–35.