Solving Fredholm Integrals of the First Kind With Tensor Product Structure in 2 and 2.5 Dimensions

Lalitha Venkataramanan, *Member, IEEE*, Yi-Qiao Song, and Martin D. Hürlimann

***Abstract*—We present an efficient algorithm to solve a class of two- and 2.5-dimensional (2-D and 2.5-D) Fredholm integrals of the first kind with a tensor product structure and nonnegativity constraint on the estimated parameters of interest in an optimization framework. A zeroth-order regularization functional is used to incorporate *a priori* information about the smoothness of the parameters into the problem formulation. We adapt the Butler–Reeds–Dawson (BRD) algorithm to solve this optimization problem in three steps. In the first step, the data are compressed using singular value decomposition (SVD) of the kernels. The tensor-product structure of the kernel is exploited so that the compressed data is typically a thousand fold smaller than the original data. This size reduction is crucial for fast optimization. In the second step, the constrained optimization problem is transformed to an unconstrained optimization problem in the compressed data space. In the third step, a suboptimal value of the smoothing parameter is chosen by the BRD method. Steps 2 and 3 are iterated until convergence of the algorithm. We demonstrate the performance of the algorithm on simulated data.**

***Index Terms*—Data compression with SVD, Fredholm integral, Laplace inversion, NMR, tensor kernal.**

I. Introduction

We present an adaptation of Butler–Reeds–Dawson (BRD) algorithm [2] to solve a class of two- and 2.5-dimensional (2-D and 2.5-D) Fredholm integrals of the first kind. Such problems are frequently encountered in diverse fields such as image restoration and nuclear magnetic resonance (NMR) [1]. In these applications, the data M are measured at a matrix of sampling times τ_1 and τ_2. The measured data are related to the parameters of interest $\mathcal{F}(x,y)$ by a Fredholm integral of the first kind with a tensor product kernel

$$M(\tau_1,\tau_2)=\int\int k_1(x,\tau_1)k_2(y,\tau_2)\mathcal{F}(x,y)\,dxdy+\epsilon(\tau_1,\tau_2) \tag{1}$$

where $\epsilon(\tau_1,\tau_2)$ is a noise term that is assumed to be additive, white, and Gaussian. In most applications, the kernels k_1 and k_2 are smooth, continuous, known functions. The objective is to estimate $\mathcal{F}(x,y)$ from the measured data $M(\tau_1,\tau_2)$. It is well known that solving Fredholm integrals with smooth kernels is an ill-conditioned problem since the singular values of the kernels decay quickly to zero [9]. Thus, the solution does not vary smoothly with the data.

Manuscript received May 3, 2001; revised January 16, 2002. The associate editor coordinating the review of this paper and approving it for publication was Dr. Hamid Krim.

The authors are with Schlumberger Doll Research, Ridgefield, CT 06877 USA (e-mail: lalitha@ridgefield.sdr.slb.com).

Publisher Item Identifier S 1053-587X(02)03282-8.

Here, we consider applications where $\mathcal{F}(x,y)$ denotes the joint probability density function of random variables x and y. For example, in image restoration, the kernels in (1) represent the blurring phenomena [3], [11]. In NMR, the study of 2-D density functions can provide insights into the molecular processes of relaxation by providing an additional dimension for resolving the spin system. Fredholm integrals of this kind are also encountered in 2-D inverse Laplace transforms.

Our study in 2.5 dimensions has been guided by NMR applications in biological systems and porous media where experimental protocols have been developed to measure one or more 2-D data sets [4], [10], [17]. From the measured R independent data sets, the objective is to estimate R different projections or cross-sections of a three-dimensional (3-D) density function. The observed data are related to the underlying density function by a Fredholm integral of the first kind with a tensor product structure

$$M_r(\tau_1,\tau_2)=\int\int k_1(x,\tau_1)k_2(y,\tau_2)\mathcal{F}_r(x,y)\,dx\,dy+\epsilon_r(\tau_1,\tau_2)\quad r=1,\ldots,R \tag{2}$$

where $\epsilon_r(\tau_1,\tau_2)$ denotes additive, white Gaussian noise. Here, the random variables x and y typically represent relaxation times or time-dependent diffusion constant. The objective is to estimate $\mathcal{F}_r(x,y)$ subject to non-negativity constraint

$$\mathcal{F}_r(x,y)\geq 0\quad \forall x \text{ and } y,\quad r=1,\ldots,R. \tag{3}$$

In the special case when $R=1$, (2) reduces to (1) and corresponds to a 2-D inversion [4], [17]. When $R>1$, $\mathcal{F}_r(x,y)$ corresponds to the joint density function of random variables x and y given experimentally controlled variable r [10]. Thus, when there is more than one data set, R different cross-sections or projections of a 3-D density function are estimated and the inversion is considered to be 2.5-D.

In this manuscript, we present an adaptation of the algorithm presented in [2] to solve (2) for the density function $\mathcal{F}_r$ from the observed data $M_r, r=1,...R$. Let $x^{(i)}, i=1,\ldots,N_x$ and $y^{(i)}, i=1,\ldots,N_y$ denote the discretization of parameters x and y. In matrix notation, (2) can be rewritten as

$$M_r=K_1F_rK_2'+E_r,\quad r=1,\ldots,R \tag{4}$$

where matrix $M_r\in\Re^{N_1\times N_2}, r=1,\ldots,R$ contains the rth 2-D data set, matrices $K_1\in\Re^{N_1\times N_x}$ and $K_2\in\Re^{N_2\times N_y}$ contain entries corresponding to continuous kernels k_1 and k_2, and matrix $F_r\in\Re^{N_x\times N_y}$ is the discretized version of continuous density function $\mathcal{F}_r(x,y)$. Without loss of generality, we assume that the Gaussian noise in matrix E_r has zero mean and

unit variance. If the noise variances are unequal for different observations, the kernel matrices and data can be scaled by the reciprocal of the noise variance such that each equation in the resulting system of linear equations is given equal weight in a least squares sense. Thus, the problem of estimating the density function $\mathcal{F}_r(x,y)$ in (2) can be approximated to solving the system of equations in (4) for matrix F_r subject to

$$F_r(i,j) \geq 0, \quad i = 1,\ldots,N_x, \quad j = 1,\ldots,N_y \quad r = 1,\ldots,R. \quad (5)$$

However, as is well known, this problem is ill conditioned. Matrices K_1 and K_2 are usually Vandermonde and are rank deficient. Their singular values decay quickly toward zero, and their condition numbers are infinite.

In the literature, there have been two methods to solve Fredholm integrals. The first approach uses truncated singular value decomposition (TSVD). In this approach, a closed-form expression for F_r can be found in a straightforward manner in terms of nonzero singular values and singular functions of matrices K_1 and K_2 [3], [8], [18]. The number of singular values is used to control the quality of the solution [8], [9], [13]. This approach has two main drawbacks that make it unattractive in our application. First, the estimated density function is not guaranteed to be non-negative. Second, choosing the number of significant singular values usually relies on the availability of the stochastic properties of the true density function, which is unknown in many applications.

In this paper, we have followed a second approach where the non-negativity constraint on the parameters are easily implemented. In this approach, a regularization functional is used to incorporate *a priori* information about the unknown density function in an optimization framework [15]. This *a priori* information is typically a measure of the smoothness of the desired solution and is weighted by a smoothing parameter. The problem is posed in an optimization framework and is solved in three steps. In the first step, the data are compressed using SVD of the kernels. In this step, we exploit the separable kernel structure in (2) to solve the problem efficiently in real time with no large memory requirements. In the second step, the density function is estimated for a given value of the smoothing parameter. This is achieved by transforming the constrained optimization problem to an unconstrained optimization problem in the compressed data space using the BRD method. In the third step, we present an extension of the BRD method to choose a sub-optimal value of the smoothing parameter. Steps 2 and 3 are applied iteratively until convergence of the algorithm to an optimal smoothing parameter occurs. We demonstrate the performance of the algorithm on simulated data in the context of nuclear magnetic resonance measurements.

In principle, the 2-D problem in (2) can be reduced to a 1-D problem

$$m_r = K_0 f_r + \epsilon_r, \quad r = 1,\ldots,R. \quad (6)$$

Here, the vectors m_r, f_r and ϵ_r are obtained by lexicographically ordering matrices M_r, F_r and E_r, respectively, and

$$K_0 = K_1 \otimes K_2 \quad \text{where } K_0 \in \Re^{(N_1N_2)\times(N_xN_y)} \quad (7)$$

where the symbol $\otimes$ denotes the tensor product. Standard nonnegative least squares (nnls) routines can then be used to solve for f_r with zeroth-order regularization [4]. However, implementation of (6) is computationally intensive. For a 2-D inversion with $R = 1$ and for typical values such as $N_1 = 30, N_2 = 4000, N_x = 100$, and $N_y = 100$, matrix K_0 has $[120\,000 \times 10\,000]$ elements. Thus, storage of K_0 is expensive in comparison with individual matrices K_1 and K_2, which have $[30 \times 100]$ and $[4000 \times 100]$ elements respectively. Hence, an nnls algorithm to solve (6) is usually impractical in real time. For example, English *et al.* [4] used a CRAY supercomputer to estimate a coarsely discretized density function ($N_x = N_y = 35$), even from a small ($R = 1, N_1 = 50, N_2 = 50$) data set. To overcome the huge computational burden of solving (6) subject to the non-negativity constraint on f_r and to be able to invert larger data sets, we propose an adaptation of algorithm in [2] and [6] to solve the problem in real time without large memory requirements. This algorithm can be easily implemented on standard PCs. In our simulations, the data were easily analyzed within a few minutes.

This paper is organized as follows. In the next section, we present the algorithm to estimate the discrete joint density function $F_r(x,y)$. In Section III, we present the simulation results. The results are summarized in the last section.

II. Estimation of Density Function

A. Algorithm

In an optimization framework, the problem is posed as

$$\min \sum_{r=1}^{R} ||M_r - K_1 F_r K_2'||^2 + \alpha ||F_r||^2 \quad (8)$$

where $||\cdot||^2$ denotes the Frobenius norm of a matrix. Since the terms inside the summation sign are independent, minimizing (8) is equivalent to minimizing R 2-D cost functions separately

$$\hat{F}_r = \arg\min_{F_r \geq 0} ||M_r - K_1 F_r K_2'||^2 + \alpha ||F_r||^2, \quad r = 1,\ldots,R. \quad (9)$$

In (9), the first term is a measure of the difference between the data and the fit. The second term summed over the different 2-D data sets is a measure of the desired smoothness of the 2-D discretized density function $F_r, r = 1,\ldots,R$. Recalling that $F_r, r = 1,\ldots,R$ corresponds to R different cross-sections of a 3-D density function of variables x, y, and r, the smoothing parameter α is chosen to be the same for all the R 2-D inversion problems.

A number of different regularization schemes have been considered in the literature [13]. We use Tikhonov regularization for two reasons. First, the quadratic nature of the second term in (9) guarantees the existence of an unique solution. Second, a zeroth-order regularization functional is appealing for its mathematical simplicity. It captures the smoothness desired in the solution while providing a framework to apply the algorithm in [2] to estimate F_r in an efficient manner. In (9), the parameter α denotes the tradeoff between the least-squares error and the

desired smoothness. Determination of the regularization parameter α is discussed in the later part of this section.

We describe below the modification of the algorithm described in [2] and [6] to solve the optimization problem in (9) in three steps. In the first step, the data M_r are compressed using SVD of K_1 and K_2. The size of the compressed data $\tilde{M}_r$ depends on the kernels. As will be shown in the simulations section, when the kernels are smooth, the data are highly redundant, and this step reduces the size of the data typically about a thousand fold. In the second step, it is assumed that the optimal value of smoothing parameter α is known. The algorithm presented in [2] is then applied to convert a constrained optimization problem into an unconstrained optimization problem. In the unconstrained optimization framework, we solve for a vector $c_r, r = 1, \ldots, R$ whose dimension equals the size of the compressed data. As will be shown below, the vectors c_r are then related to the parameters of interest F_r in a straightforward manner. In practice, for smooth kernels, the size of c_r is much smaller than the size of F_r, thus providing an elegant method to efficiently estimate the underlying density function. A sub-optimal value of the smoothing parameter α is computed in the third step from intermediate values of parameters calculated in the second step. The second and third steps are repeated iteratively until the algorithm converges to an optimal smoothing parameter. We describe below the three steps in further detail.

1) Step 1: Data Compression: Let the SVD of K_1 be

$$K_1 = U_1 \Sigma_1 V_1'. \tag{10}$$

Here, Σ_1 is a diagonal matrix with the singular values in a decreasing order along the diagonal. When K_1 represents a smooth kernel, the singular values decay quickly toward zero. Thus, the condition number of K_1 is ∞. Let s_1 denote the rank or, equivalently, the number of nonzero singular values of K_1. In this case, $\Sigma_1 \in \Re^{s_1 \times s_1}$ and contains only the nonzero distinct singular values along the diagonal. In the presence of noise, the number of singular values is chosen such that noise magnification is not significant [16]. Guidelines to choose the number of significant values is presented in the simulations section.

The matrices $U_1 (U_1 \in \Re^{N_1 \times s_1})$ and $V_1 (V_1 \in \Re^{N_x \times s_1})$ are unitary. The columns of V_1 form an orthonormal set of basis vectors that span the row space of K_1, and the columns of U_1 form an orthonormal set of basis vectors that span the range of K_1. Matrix $U_1 U_1'$ is the projector onto the range space of K_1. Similarly, let the SVD of K_2 be

$$K_2 = U_2 \Sigma_2 V_2' \tag{11}$$

where $\Sigma_2 \in \Re^{s_2 \times s_2}$ is a diagonal matrix with s_2 nonzero distinct singular values along the diagonal, $U_2 (U_2 \in \Re^{N_2 \times s_2})$ and $V_2 (V_2 \in \Re^{N_y \times s_2})$ are unitary, and $U_2 U_2'$ is the projector of K_2. Thus, the projection of the data M_r onto the range space is given in terms of unitary matrices U_1 and U_2 as

$$\tilde{\tilde{M}}_r = U_1 U_1' M_r U_2 U_2' \quad \text{where } \tilde{\tilde{M}}_r \in \Re^{N_1 \times N_2}. \tag{12}$$

The matrix $\tilde{\tilde{M}}$ is referred to as projected or uncompressed data. Making use of the unitary properties of U_1 and U_2, it can be shown that for $r = 1, \ldots, R$

$$\begin{aligned} \hat{F}_r &= \arg \min_{F_r \geq 0} ||M_r - K_1 F_r K_2'||^2 + \alpha ||F_r||^2 \\ &= \arg \min_{F_r \geq 0} ||\tilde{\tilde{M}}_r - U_1 U_1' (K_1 F_r K_2') U_2 U_2'||^2 \\ &\quad + ||M_r||^2 - ||\tilde{\tilde{M}}_r||^2 + \alpha ||F_r||^2. \end{aligned} \tag{13}$$

Let

$$\tilde{M}_r = U_1' M_r U_2 \quad \text{where } \tilde{M}_r \in \Re^{s_1 \times s_2}. \tag{14}$$

The matrix $\tilde{M}_r$ is referred to as compressed data of the rth 2-D data set. Let

$$\tilde{K}_1 = \Sigma_1 V_1', \quad \tilde{K}_1 \in \Re^{s_1 \times N_x} \quad \text{and} \quad \tilde{K}_2 = \Sigma_2 V_2' \quad \tilde{K}_2 \in \Re^{s_2 \times N_y}. \tag{15}$$

Then, (13) reduces to

$$\begin{aligned} \hat{F}_r = \arg \min_{F_r \geq 0} &||\tilde{\tilde{M}}_r - U_1 \tilde{K}_1 F_r \tilde{K}_2' U_2'||^2 \\ &+ ||M_r||^2 - ||\tilde{\tilde{M}}_r||^2 + \alpha ||F_r||^2. \end{aligned} \tag{16}$$

Since matrices U_1 and U_2 are unitary and the second and third terms in (16) are independent of F_r, (16) reduces to

$$\hat{F}_r = \arg \min_{F_r \geq 0} ||\tilde{M}_r - \tilde{K}_1 F_r \tilde{K}_2'||^2 + \alpha ||F_r||^2. \tag{17}$$

Thus, the problem can be reformulated as solving for F_r from (17) subject to the non-negativity constraint in (5). Comparing (9) and (17), data compression does not modify the structure of the problem. In practice, however, it does reduce the size of the original data at least a thousand fold for smooth kernels. Data compression is illustrated in Fig. 1 on an example of NMR spin relaxation correlation data measured in a laboratory on bulk water similar to the data analyzed in [4]. The measured data is shown in Fig. 1(a). In this case, the kernels are $k_1(x, \tau_1) = 1 - 2e^{-\tau_1/x}$ and $k_2(y, \tau_2) = e^{-\tau_2/y}$. Here, the variables x and y correspond to two different relaxation times T_1 and T_2, respectively. The different traces correspond to different values of τ_1. Each trace corresponds to T_2 decay measured at successive values of τ_2. The singular values of K_1 and K_2 decay quickly toward zero, as is shown in Fig. 1(b). Therefore, it is sufficient to project the data on a much smaller dimensional subspace spanned by the largest singular vectors. Choosing 30 of the largest singular vectors, the data projected onto the range space of K_0 is shown in Fig. 1(c). The compressed data contains 30 points and is shown in Fig. 1(d).

Data compression before optimization of the parameters offers two distinct advantages. First, it helps in the fast computation of the first term in (17). Second, as will be described in the next step, it helps in the efficient estimation of the discretized density function F_r. It is worthwhile to point out some properties of the data compression. Both compression and uncompression are linear transformations of M_r. The Frobenius norm of the matrix is preserved during compression or uncompression

$$||\tilde{M}_r||^2 = ||\tilde{\tilde{M}}_r||^2, \quad r = 1, \ldots, R. \tag{18}$$

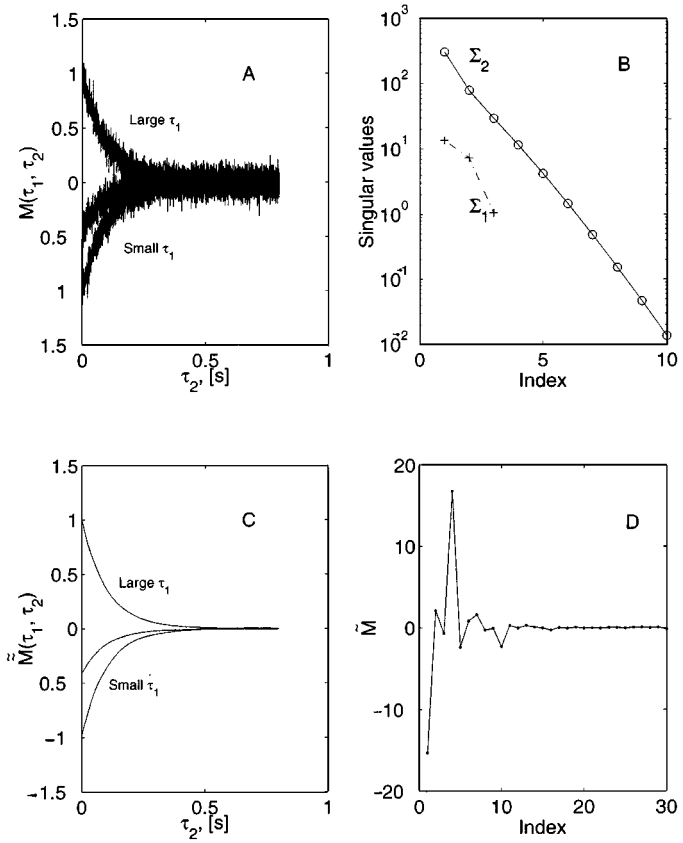

Fig. 1. Data compression. (a) Two-dimensional data set with $N_1 = 3$ and $N_2 = 4000$. The kernels are $k_1 = 1 - 2e^{-\tau_1/x}$ and $k_2 = e^{-\tau_2/y}$. Measured data are shown for $\tau_1 = 500\ \mu$s, 0.07 s and 4 s and $\tau_2 = 200$ k μs, where $k = 1, 2, \ldots, 4000$. (b) Singular values of K_1 and K_2 decay quickly toward zero. (c) Projected data $\tilde{M}$ with 4000 × 3 data points. (d) Compressed data $\tilde{M}$ contains only 30 data points.

However, uncompression of the compressed data does not yield the original data but the projection of the data M_r onto the range space. Similarly, once the data is compressed or projected, it is not possible to reconstruct the original data from the compressed or uncompressed data. Compression or uncompression does not change the noise statistics. In (4), we assume that the noise is independently and identically distributed (iid) with mean zero and unit variance. In this case, the compressed noise $\tilde{\epsilon}_r = U_1' \epsilon_r U_2, r = 1, \ldots, R$ is also iid with zero mean and unit variance.

2) Step 2: Optimization: For a given value of α, (17) has a unique solution for F_r since the second term in (17) is quadratic. In this step, the constrained optimization problem for estimating F_r in (17) is mapped onto an unconstrained optimization problem for estimating a vector c_r [2]. This vector belongs to a space whose dimension equals the size of the compressed data. The solution vector c_r is then mapped to the matrix F_r in a straightforward manner. Since the size of the compressed data is in practice much smaller than the dimensions of F_r, this second step plays a central role in reducing the computational intensity of the algorithm.

As before, (17) can be reformulated as

$$\hat{f}_r = \arg\min_{f_r \geq 0} Q = ||\tilde{m}_r - \tilde{K}_0 f_r||^2 + \alpha ||f_r||^2 \qquad (19)$$

where the vectors $\tilde{m}_r$ and f_r are obtained by lexicographically ordering matrices $\tilde{M}_r$ and F_r, respectively, and

$$\tilde{K}_0 = \tilde{K}_1 \otimes \tilde{K}_2 \quad \text{where } \tilde{K}_0 \in \Re^{(s_1 s_2) \times (N_x N_y)}. \qquad (20)$$

Let $f_{i,r}, i = 1, \ldots, (N_x N_y), r = 1, \ldots, R$ denote the ith element of vector f_r. The Kuhn–Tucker conditions for the minimum are

$$\nabla Q(f_{i,r}) = 0, \quad \text{if } f_{i,r} > 0, \quad i = 1, \ldots, (N_x N_y) \qquad (21)$$

$$\nabla Q(f_{i,r}) \geq 0, \quad \text{if } f_{i,r} = 0 \qquad (22)$$

where ∇ represents the derivative operator. The derivative of (19) is given by

$$\frac{\partial Q}{\partial f_{i,r}} = \tilde{K}_{0,i}'(\tilde{K}_0 f_r - \tilde{m}_r) + \alpha f_{i,r} \qquad (23)$$

where $\tilde{K}_{0,i}$ is the ith column of $\tilde{K}_0$. From (21) and (23), we get

$$\alpha f_{i,r} = -\tilde{K}_{0,i}'(\tilde{K}_0 f_r - \tilde{m}_r), \quad \text{if } f_{i,r} > 0 \qquad (24)$$

Therefore

$$f_{i,r} = \tilde{K}_{0,i}' c_r, \quad \text{if } f_{i,r} > 0 \qquad (25)$$

where

$$c_r = \frac{\tilde{K}_0 f_r - \tilde{m}_r}{-\alpha}, \quad c_r \in \Re^{(s_1 s_2) \times 1}. \qquad (26)$$

Direct implementation of (25) and (26) to estimate f_r and c_r iteratively from starting estimates can lead to divergence of the algorithm. Often, this divergence is due to round-off errors that arise when the largest eigenvalue of the relevant recursive matrix for estimation of f_r is very close to unity. Butler *et al.* [2] provide an elegant method described below to circumvent this problem. From (22), (23), and (25), we get

$$f_r = \max(0, \tilde{K}_0' c_r). \qquad (27)$$

Substituting (27) in (26), we have

$$\tilde{K}_0[\max(0, \tilde{K}_0' c_r)] - \tilde{m}_r + \alpha c_r = 0 \qquad (28)$$

$$(G(c_r) + \alpha I) c_r = \tilde{m}_r, \quad I \in \Re^{s_1 s_2 \times s_1 s_2} \qquad (29)$$

where

$$G(c_r) = \tilde{K}_0 \begin{bmatrix} H(\tilde{K}_{0,1}' c_r) & 0 & \ldots & 0 \\ 0 & H(\tilde{K}_{0,2}' c_r) & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & H\left(\tilde{K}_{0_{N_x \times N_y}}' c_r\right) \end{bmatrix} \tilde{K}_0' \qquad (30)$$

and where $H(\cdot)$ denotes the Heavyside function. The matrix $G(c_r)$ is symmetric and semi-positive definite. Using the theorem $\partial G(c_r)/\partial c_{k,r} \cdot c_r = 0, k = 1, \ldots, s_1 s_2$ [2], it can be shown that the vector c_r that satisfies (29) can be estimated by minimizing the function

$$\chi(c_r) = \frac{1}{2} c_r' [G(c_r) + \alpha I] c_r - c_r' \tilde{m}_r \qquad (31)$$

where $I \in \Re^{(s_1 s_2)\times(s_1 s_2)}$ is an identity matrix. The optimization for c_r can be performed using the inverse Newton method since the first and second derivative of $\chi(c_r)$ are available in closed form as

$$\nabla\chi(c_r) = (G(c_r) + \alpha I)c_r - \tilde{m}_r \tag{32}$$

$$\nabla\nabla\chi(c_r) = G(c_r) + \alpha I. \tag{33}$$

Once the vector c_r is estimated, f_r can be estimated from (25). The matrix F_r is then be estimated by reordering f_r into the matrix notation.

In the absence of the positivity constraint on F_r, (27) and (30) reduce to standard well-known equations. In this case

$$G(c_r) = \tilde{K}_0\tilde{K}_0' \tag{34}$$

and is independent of c_r. From (31), an optimal value of c_r is

$$c_r = (\tilde{K}_0\tilde{K}_0' + \alpha I)^{-1}\tilde{m}_r. \tag{35}$$

In this case, (27) reduces to the well-known minimum norm solution

$$f_r = \tilde{K}_0' c_r = \tilde{K}_0'(\tilde{K}_0\tilde{K}_0' + \alpha I)^{-1}\tilde{m}_r \tag{36}$$

and the covariance matrix of the estimated parameters is obtained in closed form as

$$\text{cov}(f_r) = \tilde{K}_0'(\tilde{K}_0\tilde{K}_0' + \alpha I)^{-2}\tilde{K}_0. \tag{37}$$

3) Step 3: Choosing α: In the previous step, it has been assumed that the optimal value of the regularization parameter is known. A number of different strategies have been proposed in the literature to determine α. These include cross-validation [19], the Morozov discrepancy principle, minimum-bound method, predictive mean square estimation [13], L-curve analysis, and the S-curve method [7]. In this paper, we have chosen α by extension of the BRD method [2]. The application of this method to 2-D and 2.5-D data is straightforward and is briefly described below.

Case 1: Application of BRD Method to 2-D Data: Let F_T denote the true but unknown underlying density function of the 2-D data set. Let $\hat{F}(\alpha)$ denote the estimated density using the algorithm described in the previous step. The parameter α can then be found by minimizing

$$S(\alpha) = ||F_T - \hat{F}(\alpha)||^2 - ||F_T||^2 \tag{38}$$

with respect to α. Under assumptions of domain containment [2], it can be shown that the lack of knowledge of F_T can be traded for the ignorance in the additive noise

$$S(\alpha) \approx c'G(c)c - 2c'\tilde{m} + 2c'\tilde{\epsilon} \tag{39}$$

where parameter $\tilde{\epsilon}$ refers to the unknown compressed noise. Recalling that the noise variance is unity, the last term in (39) can be approximated by its norm, yielding

$$S(\alpha) \approx H(\alpha) = c'G(c)c - 2c'\tilde{m} + 2||c||\sqrt{s_1 s_2}. \tag{40}$$

Using properties of matrix G [2], a closed-form expression for the derivative of $H(\alpha)$ can be derived. Equating the derivative of $H(\alpha)$ to zero yields the optimal value of α

$$\alpha_{\text{opt}} = \frac{\sqrt{s_1 s_2}}{||c||}. \tag{41}$$

Given a starting point of α, the vector c can be easily found using the algorithm described in the previous section. From this vector c, the parameter α can be updated from (41). This iterative method for choosing α is easy to implement and typically converges in ten to 20 iterations, irrespective of the initial value of α. Equation (41) also has a ready intuitive explanation. It can be shown that

$$||\tilde{m} - \tilde{K}_0 f||^2 = \alpha^2||c||^2. \tag{42}$$

Over the different noise realizations, the residual $||\tilde{m} - \tilde{K}_0 f||^2$ has a chi-squared distribution with mean $s_1 s_2$ and variance $2s_1 s_2$. The BRD method proposes to choose α such that the norm of the residual at the optimal value of α equals its expected value, which is similar to the approach in [14]. An example comparing the approximate parameter bias $H(\alpha)$ to the true bias $S(\alpha)$ will be shown in the simulations section.

Case 2: Application of BRD Method to 3-D Data: Let $F_{T,r}$ denote the true but unknown underlying density function of the rth $(r = 1, \ldots, R)$ 2-D data set. Let $\hat{F}_r(\alpha)$ denote the estimated density using the algorithm described in the previous step. The parameter α can then be found by minimizing

$$S(\alpha) = \sum_{r=1}^{R} ||F_{T,r} - \hat{F}_r(\alpha)||^2 - ||F_{T,r}||^2 \tag{43}$$

with respect to α. From (39)

$$S(\alpha) \approx \sum_{r=1}^{R} c_r' G(c_r)c_r - 2c_r'\tilde{m}_r + 2c_r'\tilde{\epsilon}_r \tag{44}$$

where parameter $\tilde{\epsilon}_r$ refers to the unknown compressed noise of the rth data set. As before, the last term in (44) can be approximated by its norm, yielding

$$S(\alpha) \approx H(\alpha) = \sum_{r=1}^{R} c_r' G(c_r)c_r - 2c_r'\tilde{m}_r + 2||c_r||\sqrt{s_1 s_2}. \tag{45}$$

Let

$$T(c_r) = (G(c_r) + \alpha I)^{-1}. \tag{46}$$

Along the lines of [2], using properties of matrix $G(c_r)$, a closed-form expression for the derivative of $H(\alpha)$ can be derived as

$$\frac{\partial H(\alpha)}{\partial \alpha} = \sum_{r=1}^{R} c_r' T(c_r)c_r \left[\alpha - \frac{\sqrt{s_1 s_2}}{||c_r||}\right]. \tag{47}$$

Equating the derivative to zero yields the optimal value of α

$$\alpha_{\text{opt}} = \sqrt{s_1 s_2}\frac{\sum_{r=1}^{R} \frac{c_r' T(c_r)c_r}{||c_r||}}{\sum_{r=1}^{R} c_r' T(c_r)c_r}. \tag{48}$$

Given a starting point of α, the vector c_r can be easily found using the algorithm described in the previous section. From this vector c_r, the parameter α can be updated from (48). This iterative method for choosing α converges in ten to 20 iterations, irrespective of the initial guess of α.

B. Issues in Practical Implementation

1) Discretization: The integral in (2) is approximated by finite-dimensional vectors and summation operator leading to (4). The approximation to the integral can be done in several ways. However, a simple approximation serves our purpose because error caused by this approximation is much smaller than error caused by regularization and noise in the data. Therefore, we use a simple quadratic approximation to replace the integral by the summation. A common fallacy is to believe that the finer the discretization and the higher the number of x's and y's, the more parameters need to be estimated and, therefore, more ill-conditioned the problem. This is, however, not true.

The ill-conditioned nature of the problem is a function of the intrinsic nature of the problem formulation. It is characterized by the condition numbers of the matrices K_1 and K_2 and is independent of the measured data. As the discretization becomes finer and the number of x's and y's is larger, the singular values of matrices K_1 and K_2 gradually converge to the singular values of the respective continuous kernels. Thus, in a Tikhonov regularization framework, the condition number is controlled only by the regularization parameter and is independent of the discretization for large number of x's and y's. In our application with exponential kernels, we typically tend to use $N_x = N_y = 100$, where x and y represent time constants and are discretized logarithmically between a few milliseconds to a few seconds.

2) Choosing Singular Values: In our simulations, the number of singular values were chosen such that the condition number of matrix K_0 equals a predetermined value. The operation is performed directly on matrix K_0 and not on matrices K_1 and K_2 since for some kernels $s_1 s_2$, the largest singular values of K_0 do not necessarily correspond to the product of s_1 largest singular values of K_1 with the s_2 largest singular values of K_2. In this case, the compressed data can be found efficiently, taking advantage of the tensor product structure of the kernels as follows. Let $\sigma_1^{\max}$ and $\sigma_2^{\max}$ correspond to the largest singular values of K_1 and K_2, respectively. Let the ith and jth singular values and vectors of matrices K_1 and K_2 be denoted by $\sigma_1^{(i)}, u_1^{(i)}, \sigma_2^{(j)}$, and $u_2^{(j)}$, respectively. If $(\sigma_1^{\max}\sigma_2^{\max}/\sigma_1^{(i)}\sigma_2^{(j)})$ is less than a predetermined value, then the corresponding compressed data point in the rth data set is given by $u_1^{(i)'} M_r u_2^{(j)}$.

3) Estimating Noise Variance: In practice, an *a priori* estimate of the noise variance depends on the data and the data acquisition system. For example, in NMR applications, noise variance can be estimated from quadrature channel after phase rotation [5]. Alternatively, an estimate of noise variance can be obtained from the data channel if the data that are well modeled by exponentially decay as in [4]. Assuming that the decay times are longer than the data spacing of τ_1 and τ_2, an estimate can be found by fitting the data with a straight line over three successive data points

$$\sigma_r^2 = \frac{1}{6kN_1}\sum_{\tau_1}\sum_{i=0}^{k-1}\left[M_r\left(\tau_1,\tau_2^{3i+1}\right) - 2M_r\left(\tau_1,\tau_2^{3i+2}\right) + M_r\left(\tau_1,\tau_2^{3i+3}\right)\right]^2, \quad r = 1,\ldots,R \tag{49}$$

where $k = \text{floor}(N_2/3)$.

III. Simulation Results

In this section, we present results on three sets of simulations. In the first set of simulations, we demonstrate the performance of the algorithm on 2-D data at different SNRs using three models for the density function $\mathcal{F}(x,y)$. The second simulation compares the performance of the algorithm on very small and large values of the smoothing parameter α. The third simulation investigates the results of simulation when the number of significant singular values is changed.

A. Data Simulation

In the simulations, we studied applications with specific kernels $k_1(x,\tau_1) = 1 - 2e^{-\tau_1/x}$ and $k_2(\tau_2,y) = e^{-\tau_2/y}$. These kernels are relevant in nuclear magnetic resonance measurements of inversion-recovery-CPMG experiments [4]. In these applications, x and y correspond to longitudinal and transverse relaxation times.

We analyze the inversion of simulated data for three different forms of the density function $F(x,y)$. In model A, $F(x,y)$ is a Gaussian with a relatively small spread. In model B, x and y are strongly positively correlated, giving a ridge-like characteristic to the joint density function. In model C, the density function has two distinct features: a Gaussian with a finite spread and a ridge obtained from negatively correlating x and y. From the true density function $F_T(x,y)$ for each model, we generate simulated data by

$$S = K_1 F_T K_2'. \tag{50}$$

White Gaussian noise ϵ was added to signal S to simulate data matrix M. Following [12], the signal-to-noise ratio (SNR) of the data is defined as

$$\text{SNR} = 10\log\frac{\|S+\epsilon\|^2}{\|\epsilon\|^2}\ \text{dB}. \tag{51}$$

In the evaluation of (50), we define the true density function at discrete values of x and y on a finely sampled logarithmic grid. The simulated data were sampled at N_1 numbers of τ_1, evenly spaced logarithmically between $\tau_{1,\min}$ and $\tau_{1,\max}$ and N_2 numbers of τ_2s linearly spaced with a sampling time of $\Delta\tau_2$. The values of $N_1, N_2, \tau_{1,\min}, \tau_{1,\max}$ and $\Delta\tau_2$ for the three models are given in Table I.

B. Data Inversion Results

1) Model A: Gaussian Density Function: Fig. 2(a) shows the contour plot of the true density function $F_T(x,y)$ that was assumed for model A. The contours are equidistant on a linear

TABLE I
VALUES OF $N_1, N_2, \tau_{1,\min}, \tau_{1,\max}$, AND $\Delta\tau_2$ FOR THE THREE MODELS

Model	N_1	N_2	$\tau_{1,min}$	$\tau_{1,max}$	$\Delta\tau_2$
			s	s	s
A	30	4000	5×10^{-4}	4	2×10^{-4}
B	30	8000	5×10^{-4}	4	2.5×10^{-4}
C	50	4000	10^{-4}	10	3.5×10^{-4}

scale. The circular contours in the $x-y$ plane show that x and y are uncorrelated. The density function is centered at $[\bar{x}\ \bar{y}] = [0.2s\ \ 0.1s]$. For a maximum signal amplitude of one unit, the SNR of 15, 10, and 5 dB corresponds to a noise standard deviation of 0.007, 0.02, and 0.07 units, respectively.

The simulated data were inverted with the algorithm presented in the previous section in two ways. In the first case, the density functions were estimated from the data using a wide range of the smoothing parameter α. This allows us to study parameter bias and residual error as a function of the smoothing parameter. In the second case, the data at different SNRs were analyzed with the optimal value of smoothing chosen by the BRD method. This allows us to study parameter bias as a function of SNR of the data. The results of both of these simulations are presented in the following.

2) Parameter Bias: The parameter bias for a wide range of the smoothing parameter α is illustrated in Fig. 2(b). The true parameter bias $S(\alpha)$ defined in (38) is a measure of the difference between the true density function F_T and estimated density function $\hat{F}(x,y)$ for a given value of regularization α and is shown by the solid curve in Fig. 2(b). As will be shown more clearly below, the parameter bias $S(\alpha)$ is large at small values of α when the estimated $\hat{F}(x,y)$ contains spurious features. The bias is also high at large values of α when the estimated $\hat{F}(x,y)$ becomes featureless and is independent of the data with $\hat{F}(x,y) \to 0$. This demonstrates that there is an optimal value of α denoted by $\alpha_{\rm True}$ when the parameter bias is minimized. However, in practice, $S(\alpha)$ cannot be calculated since the true underlying density function F_T is unknown. The function $H(\alpha)$ in (39) serves to approximate the parameter bias. This function evaluated over a range of αs is shown by the dashed curve in Fig. 2(b). It is seen that $H(\alpha)$ also has a distinct minimum indicated in the figure by $\alpha_{\rm opt}$. It slightly overestimates the true minimum $\alpha_{\rm True}$. The overestimation will result in extra smoothing of the features of the density function. However, we find that this is often more desirable than underestimation of the $\alpha_{\rm True}$, which can lead to false features in the estimated function.

3) Effect of Noise: Results for the inverted density functions with the optimal value of α chosen by the BRD method are shown in Fig. 3(a)–(c) for different levels of noise in the simulated data. Comparing the estimated density functions with the true density function in Fig. 2(a), it can be seen that the estimated density function is unbiased with mean $[\bar{x}\ \bar{y}]$ being close to the true value of $[0.2s\ \ 0.1s]$. Moreover, the volume under the

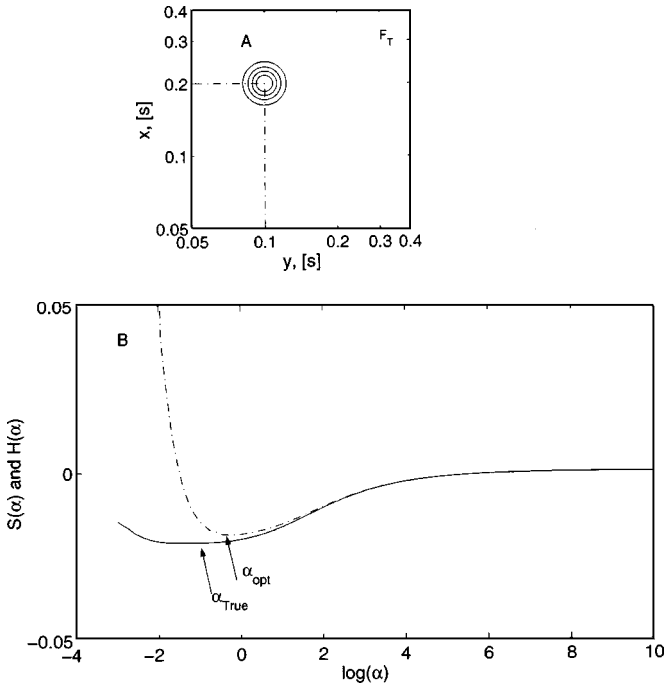

Fig. 2. Model A. (a) True density function is assumed to be a Gaussian with a small spread in the $x-y$ domain. (b) Approximate parameter bias $H(\alpha)$ (dashed curve) is compared to the true parameter bias $S(\alpha)$ (solid curve). In this simulation, the data was generated from Model A with SNR of 15 dB. The value $\alpha_{\rm opt}$ that minimizes $H(\alpha)$ is close to $\alpha_{\rm True}$, which minimizes $S(\alpha)$.

density function at the different SNRs matches the true volume. As expected, with increasing noise, the resulting density function has a larger support corresponding to larger standard deviation of the estimated parameters. The fit error in the compressed data domain

$$\chi(\alpha) = \|\tilde{m} - \tilde{K}_0 \hat{f}(\alpha)\| = \alpha \|c\| \tag{52}$$

is plotted as a function of the regularization parameter α in Fig. 3(d). For small values of α, the estimated density function fits the data closely but has a large variance and, therefore, low reproducibility for different noise realizations with the same noise standard deviation. For large values of α, $\|\hat{f}(\alpha)\| \to 0$; the fit is independent of the data and results in large residual error. In this case, the estimated density function is strongly biased but extremely stable. Thus, the fit error follows an "S" shaped curve. The S-curve method discussed in [6] proposes to choose α at the heel of this curve. The optimal values of α chosen by the BRD method at the three SNRs are indicated in the figure. They are close to the heel of the respective S-curves and are thus qualitatively consistent with the idea of choosing α that balances the tradeoff between the residual error and parameter variance.

4) Timing: To assess the speed of our algorithm, we have compared it with the standard non-negative least squares optimization with regularization implemented on MATLAB. For this comparison, the parameters x and y were discretized coarsely, taking 20 values each. The simulated data was obtained from Model A with a SNR of 30 dB. The data was first compressed to 39 data points. The value of α was fixed at the optimal value found previously. In our algorithm, the optimization in this 39-dimensional space converged in 8 cpu

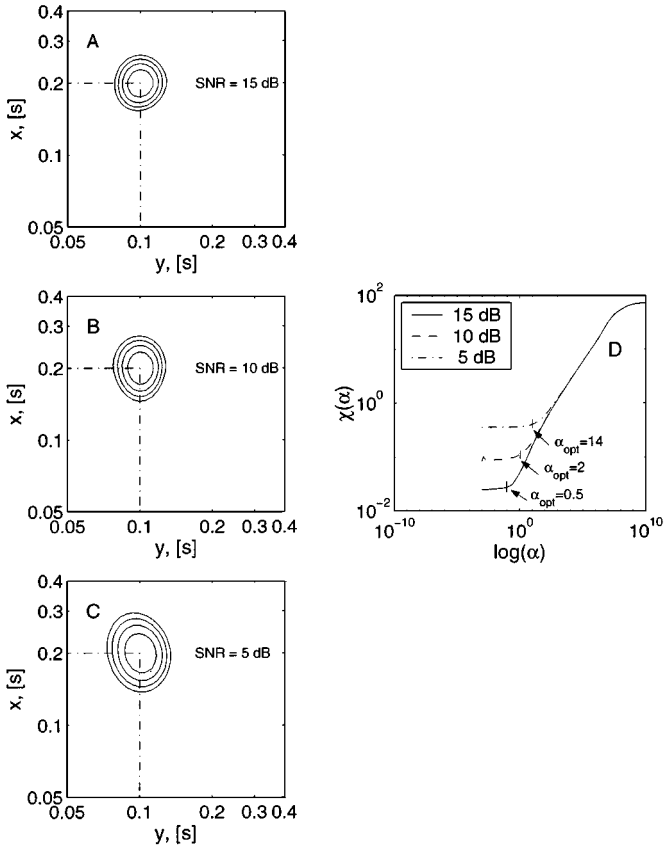

Fig. 3. Model A inversion results. (a)–(c) Estimated density function for simulated data with SNR of 15, 10, and 5 dB, respectively. The parameter $\alpha_{\rm opt}$ for the three SNRs was found from the BRD method to be 0.5, 2, and 14, respectively. The estimated density functions are similar to the true density function shown in Fig. 2(a). The residual error χ shows a characteristic S-shaped dependence on α. As shown in (d), the optimal values of α picked by the BRD method are close to the heels of these curves.

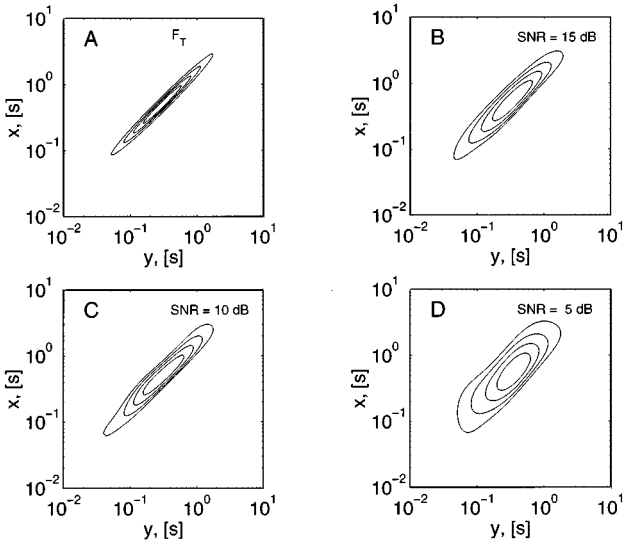

Fig. 4. Model B inversion results. (a) True density function $F_T(x,y)$ for model B showing a positive correlation between x and y. (b)–(d) Estimated $\hat{F}(x,y)$ from the simulated data with SNR of 15 dB, 10 dB, and 5 dB, respectively. They show the same characteristic as F_T with a larger spread at lower SNRs. For a maximum signal amplitude of one unit, the SNR corresponds to a noise standard deviation of 0.04, 0.07, and 0.15 units, respectively. The parameter $\alpha_{\rm opt}$ for the three SNRs was found from the BRD method to be 166.8, 147, and 3630, respectively.

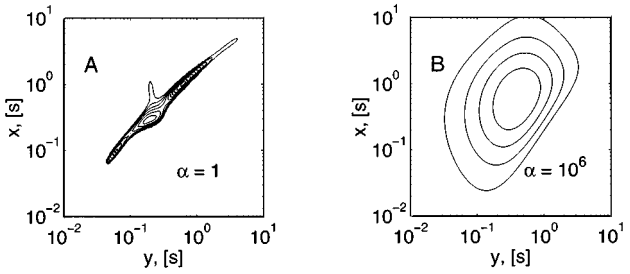

Fig. 5. Dependence of estimated density function $\hat{F}(x,y)$ on smoothing parameter α for simulated data from Model B with SNR $=$ 10 dB: In (a), $\alpha = 1 \ll \alpha_{\rm opt} = 147$. Comparison with Fig. 4(a) shows visible false features that are not reproducible for different noise realizations. In (b), $\alpha = 10^6 \gg \alpha_{\rm opt}$, and the resulting $\hat{F}(x,y)$ is much broader than for $\alpha_{\rm opt} = 147$ in Fig. 4(c) but is extremely stable.

seconds. On the other hand, standard constrained optimization on the same compressed data requires optimization for vector f in a 400-dimensional space and converged only after 10 cpu hours. This demonstrates the ease of solving Fredholm integrals with the new algorithm in real time on standard computers.

5) Model B—Ridge Like Density Function With Positive Correlation: The true density function of model B from which the data were simulated is shown in Fig. 4(a). For a maximum signal amplitude of one unit, the SNR of 15, 10, and 5 dB corresponds to a noise standard deviation of 0.04, 0.07, and 0.15 units, respectively. The data were analyzed as before with $\alpha_{\rm opt}$ being chosen by the BRD method and found to be 167, 147, and 3630, respectively. The resulting density functions are shown in Fig. 4(b)–(d). It is seen that the estimated density functions successfully recover the same ridge-like characteristic with no false features. As expected, the distributions become broader with increasing noise in the data.

6) Dependence on α: Fig. 5 shows the performance of the algorithm for very small and large values of the smoothing parameter α on data from this model at SNR of 10 dB. In Fig. 5(a), $\alpha = 1$ and is much smaller than the optimal value of 147. In this case, the estimated density function fits the data very well. The norm of the residual is 0.19 and comparable with the norm of the residual for $\alpha_{\rm opt}$ of 0.27. This explains the flat portion of the fit error S-curve for low values of α in Fig. 3(d). However, since the problem is ill-conditioned, the estimated density function has a large variance. Thus, false features are evident in Fig. 5(a). These features are not reproducible for different noise realizations with the same noise standard deviation. On the other hand, when α is much larger than the optimal value, the estimated density function $\hat{F} \to 0$ and is therefore strongly biased but extremely stable. This is illustrated in Fig. 5(b) with $\alpha = 10^6$. The norm of the residual is 9.8 and much larger than the norm of the residual for $\alpha_{\rm opt}$ of 0.27, thus explaining the flat portion of the "S" shaped curve of the fit error for large values of α in Fig. 3(d). In our applications, we have found that the main features and moments of the density function are not critically dependent on the exact determination of the optimal value of α. However, in applications wherein $\alpha_{\rm opt}$ plays a critical role, it may be meaningful to compare different strategies to choose α [13].

7) Model C—Density Function With Point-Like and Ridge-Like Features: Fig. 6(a) shows the true density function

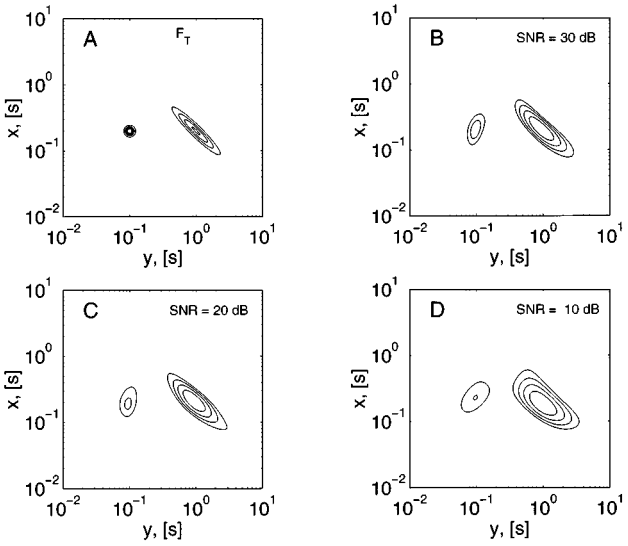

Fig. 6. Model C inversion results: (a) True density function $F_T(x, y)$ for model C showing two features: a negative correlation between x and y as seen in the ridge-like feature and a Gaussian with a small spread. (b)–(d) Estimated density functions $\hat{F}(x, y)$ from simulated data with SNR of 30 dB, 20 dB, and 10 dB, respectively. The estimated density functions all recover the essential characteristic of the true density function with a smaller resolution at lower SNRs. The smoothing parameter α_{opt} for the three SNRs was found from the BRD method to be 16, 38, and 474, respectively.

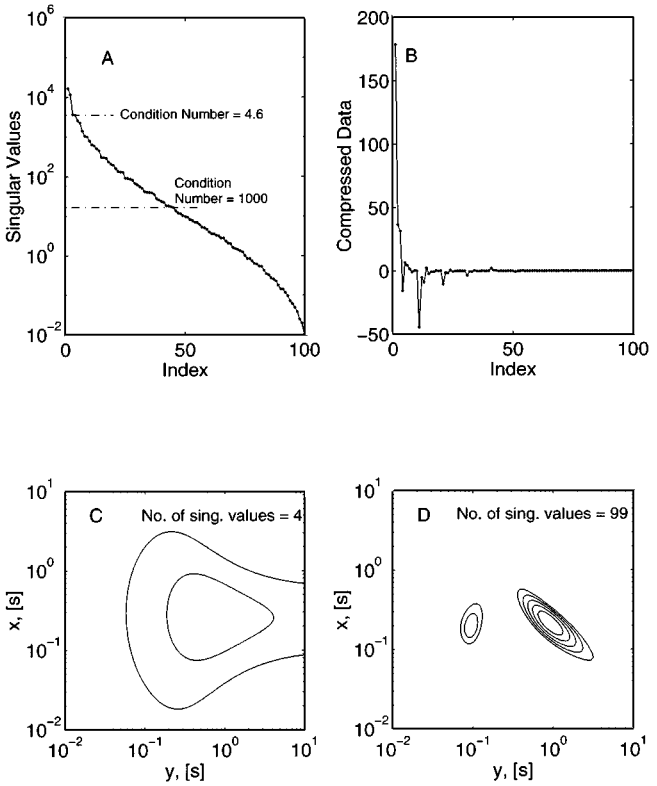

Fig. 7. Dependence of the estimated density functions $\hat{F}(x, y)$ for simulated data of model C with SNR of 20 dB on the number of singular values. (a) First 100 singular values. (b) Compressed data corresponding to the first 100 singular values shown in (a). (c) Estimated density function $\hat{F}(x, y)$ when only the first four singular values are considered. There is strong bias as significant information is discarded. (d) Estimated density function $\hat{F}(x, y)$ using 99 singular values.

of model C. To ensure that the analysis does not have an inherent bias toward a positive ridge-like characteristic, the data are simulated here from a density function where x and y are in part negatively correlated, leading to a ridge with a negative slope in the $x - y$ domain. In addition, the density function has a secondary feature—a Gaussian with a finite support. For a maximum signal amplitude of one unit, SNR of 30, 20, and 10 dB corresponds to a noise standard deviation of 0.01, 0.04, and 0.14 units, respectively. The analysis of the data are shown in Figs. 6(b)–(d). On comparison of the estimated and true joint density functions, it is seen that the parameters are unbiased and reproduce the two main features. At SNR of 10 dB, the resolution between the two features is poorer, although the features are still well distinguishable.

8) Dependence on the Number of Singular Values: In the previous analysis, the number of singular values was chosen to be such that the condition number was 1000. This corresponded to 45 significant singular values. We analyzed the effect of changing the number of singular values using simulated data from Model C at SNR of 20 dB at optimal value of α. Results are shown in Fig. 7. When the number of significant singular values considered is reduced, the compressed data has fewer points. If the number is too small, the parameter estimates will be strongly biased, as shown in Fig. 7(c). In this case, only the four largest singular values were considered, resulting in a condition number of 4.6. Thus, only the first four points of the compressed data shown in Fig. 7(b). This clearly results in a significant loss of information. As a consequence, the estimated density function is biased, and the two features cannot be resolved.

On the other hand, increasing the number of significant singular values to include singular values above 45 (the number used before) does not add significant new information. This is evident from Fig. 7(b). It shows that the compressed data points for indices larger than 45 are very close to zero. In Fig. 7(d), we display the estimated density function using 99 compressed data points, corresponding to a condition number of 8×10^5. The estimated density function in Fig. 7(d) is very similar to the one shown in Fig. 6(b), which was obtained using less than half the number of singular values. In fact, since the singular values decrease rapidly, increasing the number of singular values increases the noise magnification term and is equivalent to undersmoothing or choosing suboptimal values of α in the regularization framework [16].

Shim *et al.* [16] have proposed a method to choose the optimal number of singular values that balances the tradeoff between bias and variance of the estimated parameters. This method assumes that some statistics of the estimated parameters are known. In many applications, this information may not be available. In our simulations, the number of singular values of K_0 is chosen such that the ratio of the largest to the smallest singular value is about a thousand fold. As a general rule of thumb, the number of significant singular values is too small if the norm of the residual is much larger than that expected from random noise with known variance.

In general, we have found that most smooth features in the density function are resolved easily by our algorithm. However, sharp features and discontinuities are hard to recover from the data in our problem formulation. In these applications, other regularization functionals optimized for the expected features in the density function should be used in the optimization framework.

The implementation of the algorithm was done on a SPARC Ultra 60 computer. From a given starting point for α, the algorithm took 5–10 min to converge to a final estimate. Thus, the algorithm exploiting the separable kernel structure in (2) can be used to solve the inversion problem in real time without large memory requirements.

IV. Summary

In this paper, we have presented an efficient algorithm to solve a class of two and 2.5-dimensional Fredholm integrals of the first kind with a tensor product structure and non-negativity constraint on the estimated parameters of interest. This is an ill-conditioned problem since the singular values of the kernels decay quickly to zero. Problems of this kind are frequently encountered in computing a density function from its moments or in inverse Laplace transforms in diverse applications such as image restoration and nuclear magnetic resonance.

We have presented an adaptation of the algorithm described in [2] to solve two and 2.5-dimensional inversion problems. The algorithm consists of three main steps. In the first step, the data are compressed using SVD of the kernels. In the second step, an optimization problem with a non-negativity constraint is solved in a low-dimensional space. The dimensionality of this space is dictated by the number of significant singular values. In the third step, a suboptimal value of smoothing parameter α is found by the BRD method. Steps 2 and 3 are iterated until convergence of the algorithm to an optimal value of the smoothing parameter is reached.

We have demonstrated the performance of the algorithm on simulated data corresponding to different density functions and different levels of additive noise. We have investigated the effect of choice of the smoothing parameter and degree of data compression on the extracted density function and have studied the resulting parameter bias and stability. The application of this theory to NMR data, sensitivity analysis, and the interpretation of the results are presented in [10] and [17].

Acknowledgment

The authors wish to thank A. Sezginer for helpful discussions on data inversion.

References

[1] I. K. Abu-Shumays and L. D. Marinelli, "A smoothing solution (unfolding) of a two dimensional density function from its spectrum," *J. Comput. Phys.*, vol. 7, pp. 219–238, 1971.

[2] J. P. Butler, J. A. Reeds, and S. V. Dawson, "Estimating solutions of the first kind integral equations with nonnegative constraints and optimal smoothing," *SIAM J. Numer. Anal.*, vol. 18, no. 3, pp. 381–397, June 1981.

[3] L. Eldén and I. Skoglund, "Algorithms for the regularization of ill-conditioned least squares problems with tensor product structure, and application to space-variant image restoration," Dept. Math., Linköping Univ., Linköping, Sweden, Rep. LiTH-MAT-R-1982-48.

[4] A. E. English, K. P. Whittal, M. L. G. Joy, and R. M. Henkelman, "Quantitative two-dimensional time correlation relaxometry," *Magn. Reson. Med.*, vol. 22, pp. 425–434, 1991.

[5] T. C. Farrar and E. D. Becker, *Pulse and Fourier Transform NMR*. New York: Academic, 1971.

[6] E. J. Fordham, A. Sezginer, and L. D. Hall, "Imaging multiexponential relaxation in the $(y, \log_e T_1)$ plane, with application to clay filtration in rock cores," *J. Magn. Reson. A*, vol. 113, pp. 139–150, 1995.

[7] P. C. Hansen, "Analysis of discrete ill-posed problems by means of the L-curve.," *SIAM Rev.*, vol. 34, no. 4, pp. 561–580, Dec. 1992.

[8] ——, "The truncated SVD as a method for regularization," *BIT*, vol. 27, pp. 534–553, 1987.

[9] R. J. Hanson, "A numerical method for solving Fredholm integral equations of the first kind using singular values," *SIAM J. Numer. Anal.*, vol. 8, no. 3, pp. 616–622, Sept. 1971.

[10] M. D. Hurlimann and L. Venkataramanan, "Quantitative measurement of two dimensional distribution functions of diffusion and relaxation in strongly inhomogenous fields," *J. Magn. Reson.*, 2001, submitted for publication.

[11] J. Kamm and J. Nagy, "Kronecker product and SVD approximations in image restoration," *Linear Algebra Appl.*, vol. 284, pp. 177–192, 1998.

[12] R. L. Lagendijk and J. Biemond, *Iterative Identification and Restoration of Images*. Boston, MA: Kluwer, 1991.

[13] M. A. Lucas, "Comparisons of parameter choice methods for regularization with discrete noisy data," *Inv. Prob.*, vol. 14, pp. 161–184, 1998.

[14] D. L. Phillips, "A technique for the numerical solution of certain integral equations of the first kind," *J. Assoc. Comput. Mach.*, vol. 9, pp. 84–97, 1962.

[15] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*. Cambridge, U.K.: Cambridge Univ. Press, 1992.

[16] Y. S. Shim and Z. H. Cho, "SVD Pseudoinversion image reconstruction," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 904–909, Aug. 1981.

[17] Y.-Q. Song, L. Venkataramanan, M. D. Hürlimann, M. Flaum, P. Frulla, and C. Straley, "T_1-T_2 Correlation spectra obtained using a fast two-dimensional laplace inversion," *J. Magn. Reson.*, vol. 154, pp. 261–268, 2002.

[18] P. Xu, "Truncated SVD methods for discrete ill-posed problems," *Geophy. J. Int.*, vol. 135, pp. 505–514, 1998.

[19] G. Wahba, "Practical approximate solutions to linear operator equations when the data are noisy," *SIAM J. Num. Anal.*, vol. 15, pp. 651–667, 1977.

Lalitha Venkataramanan (M'99) received the M.S. degree is electrical engineering from the Indian Institute of Science, Bangalore, in 1992. She received the Ph.D. degree in electrical engineering from Yale University, New Haven, CT, in 1998.

She is currently a Senior Research Scientist at Schlumberger Doll Research, Ridgefield, CT. Her current interests include time series modeling, protein modeling, and theoretical and applied digital signal processing.

Yi-Qiao Song received the B.S. degree from Peking University, Beijing, China, in 1985 and the Ph.D. degree in physics from Northwestern University, Evanston, IL, in 1991.

He was a Miller Research Fellow at University of California, Berkeley, where he studied highly polarized xenon for nuclear magnetic resonance and magnetic resonance imaging. In 1997, he joined Schlumberger Doll Research, Ridgefield, CT, to investigate applications of nuclear magnetic resonance technologies to the study of porous media and well logging.

Dr. Song is a member of the American Physical Society.

Martin D. Hürlimann received the diploma in natural sciences from the Swiss Federal Institute of Technology (ETH), Zürich, Switzerland, in 1982. He received the Ph.D. degree in physics from the University of British Columbia (UBC), Vancouver, BC, Canada, in 1989.

He spent two years at the University of California, Berkeley, on a joint research project with the Department of Physics and the Department of Chemistry. In 1992, he joined Schlumberger Doll Research, Ridgefied, CT, where he is currently Principal Research Scientist. His research interests include the development and application of advanced MNR techniques to study porous media and the development of new NMR techniques for measurements in very inhomogeneous fields.