Journal of Magnetic Resonance **140,** 153–161 (1999)
Article ID jmre.1999.1837, available online at http://www.idealibrary.com on IDEAL®

The Inversion of NMR Log Data Sets with Different Measurement Errors

Keh-Jim Dunn and Gerald A. LaTorraca

Chevron Petroleum Technology Company, P.O. Box 446, La Habra, California 90633

Received December 14, 1998

We present a composite-data processing method which simultaneously processes two or more data sets with different measurement errors. We examine the role of the noise level of the data in the singular value decomposition inversion process, the criteria for a proper cutoff, and its effect on the uncertainty of the solution. Examples of processed logs using the composite-data processing method are presented and discussed. The possible usefulness of the apparent T_1/T_2 ratio extracted from the logs is illustrated.

INTRODUCTION

Since the introduction of pulsed nuclear magnetic resonance (NMR) logging tools (*1–3*), there have been extensive studies on the inversion of NMR data to obtain T_2 amplitude distributions. These T_2 distributions provide crucial information on porosity, irreducible water saturation, oil viscosity, permeability, etc., which are very important to the petroleum industry. Hence, it is essential to use a proper inversion procedure to ensure that the T_2 distribution we obtain faithfully reflects the information acquired from the measurements.

In the oil industry, a suite of well logging tools is used in a newly drilled well to perform physical measurements to find where the oil is. These tools measure such things as electron density, acoustic velocity, neutron density, electrical resistivity, and, most recently, nuclear magnetic resonance T_2 relaxation. The purpose is to determine reservoir properties such as porosity (void fraction), the relative amounts of hydrocarbon and brine in the reservoir, and how permeable the reservoir is to fluid flow. In measuring T_2 relaxation, NMR logging tools obtain measures of porosity (proportional to signal amplitude) and pore sizes (proportional to relaxation time constants). These measures are used, in turn, to infer permeability and the minimum brine saturation that the reservoir can have.

To improve the signal-to-noise (S/N) ratio without significantly increasing the logging time, Prammer *et al.* (*4*) introduced the idea of following each CPMG T_2 sequence with a series of CPMG "bursts" of echoes separated by short delays between bursts. The purpose of the added bursts is to improve the precision of the signal amplitude determination at short relaxation times. The problem that this approach raises, however, is how do we combine two data sets with different S/N ratios, echo spacings, and repeat delay times in the calculation of signal amplitude and T_2 distribution. The purpose of this paper is to describe a means of combining these data sets in a way that accounts for the differences in a self-consistent manner. The approach results in an improvement in the precision of the porosity and the resolution of short T_2 time constants over what could be achieved without the bursts.

THE INVERSION OF NMR LOG DATA

The NMR logging tools typically acquire T_2 spin-echo trains of several hundreds to a few thousands of echoes. The inversion process involved is to find a set of T_2 amplitudes, f_j, from a set of measurements, g_i,

$$g_i = M(t_i)/M_0 = \sum_{j=1}^{m} f_j e^{-t_i/T_j} + \epsilon_i, \quad i = 1, \ldots, n, \quad [1]$$

by minimizing the error ϵ_i. Here we assume that there are n echoes each with an amplitude g_i measured at t_i and m relaxation times T_j preselected to be equally spaced on a logarithmic scale. Thus, it is a linear inversion problem. The least-squares fit is used to minimize the sum

$$\min\left\{ \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\sum_{j=1}^{m} f_j e^{-t_i/T_j} - g_i \right)^2 \right\}. \quad [2]$$

This problem is well known to be ill-posed. Small fluctuations in the measured data can lead to drastically different sets of f_j's which all give reasonably good fits. This problem has been studied extensively in the past. There are various ways of handling the problem. One of the approaches is to add a penalty function to smooth or "regularize" the solution. Then an optimum regularization parameter commensurate with the

1090-7807/99 $30.00

measurement error is determined and used as the optimally regularized solution.

We found, however, that the singular value decomposition (SVD) method is particularly attractive for the following reasons: (1) it does not require a penalty term and a search for the optimum regularization parameter, (2) it provides a clear picture of the characteristics of the matrix and the overall properties of the inversion procedure, and (3) it offers a straightforward and easily comprehensive method of obtaining the regularized solution by removing the singular values smaller than the measurement error.

In general, we assume the measurement error for each spin echo is the same. Thus, σ_i in Eq. [2] is the same for all i's. However, the procedure becomes somewhat unclear when we need to process simultaneously two or more NMR data sets with very different measurement errors. In fact, one encounters a similar problem when the NMR time domain data are compressed or averaged within different window lengths. This windowing leads to different improved precisions.

In the following, we take a closer look at the procedure used in the SVD inversion as to how the noise level affects the solution and elucidate the process with a geometric picture.

We first present a composite-data processing method which processes two or more data sets with different measurement errors simultaneously. We then examine the role of the noise level of the data in the SVD inversion process and its effect on the uncertainty of the solution. The results from the composite-data processing method are presented and some of the issues concerning the T_1/T_2 ratio are discussed. Finally, we review some of the general properties of the linear inversion for NMR data.

COMPOSITE-DATA PROCESSING

Recently, Numar introduced a total porosity tool (*4*), MRIL-CTP, which measures the "clay-bound" water. It produces two data sets: the regular T_2 measurement (time between echoes, TE = 1.2 ms) with several hundred echoes and a complete recovery time, and the short burst mode (TE = 0.6 ms) which stacks 48 bursts of 10 echo trains each with an incomplete recovery time of 20 ms. The short burst mode has a sevenfold improvement of measurement precision compared with the regular mode. Figure 1 shows how the raw echoes for the two sets of data look. The short burst mode has a high measurement precision which helps accurate determination of the short T_2 components, whereas the regular mode covers the medium to long T_2 relaxation times. This is an excellent combination to measure the T_2 relaxation behavior of the formation over a large time scale. The challenge is how to invert these two data sets simultaneously to obtain as accurate a T_2 distribution as possible.

More recent logging tool developments (*5, 6*) involve a nine-frequency data acquisition scheme and three sets of echo trains with different measurement precisions. However, the challenge to properly invert all data sets simultaneously is the same. Without loss of generality, we shall limit our following discussion to two data sets with different measurement errors.

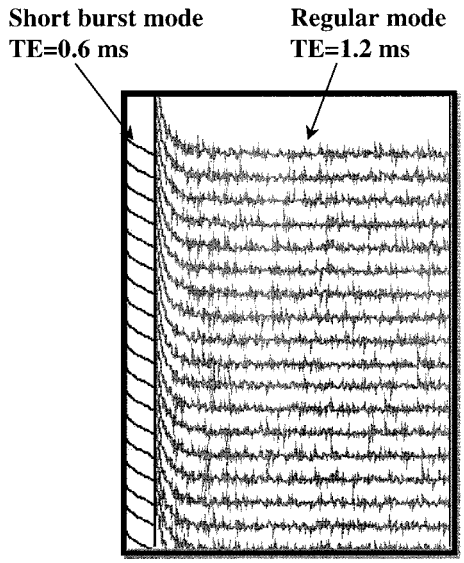

FIG. 1. Two sets of CPMG spin-echo trains measured by an NMR logging tool with different measurement errors, one is the short burst mode with echo spacing of 0.6 ms and the other is the regular mode with echo spacing of 1.2 ms.

We proposed a composite-data processing method (*7*) to handle the problem. Basically, we need to minimize the following two sets of data simultaneously:

$$\min\left\{\sum_{i=1}^{p}\frac{1}{\sigma_1^2}\left(\sum_{j=1}^{m}k_{ij}y_j-b_i\right)^2\right\} \qquad [3]$$

$$\min\left\{\sum_{i=1}^{q}\frac{1}{\sigma_2^2}\left(\sum_{j=1}^{m}l_{ij}y_j-c_i\right)^2\right\}, \qquad [4]$$

where

$$k_{ij}=(1-e^{-\mathrm{TR_l}/rT_j})e^{-t_i/T_j} \qquad [5]$$

$$l_{ij}=(1-e^{-\mathrm{TR_s}/rT_j})e^{-t_i/T_j}. \qquad [6]$$

$\mathrm{TR_l}$ and $\mathrm{TR_s}$ represent, respectively, long (complete) and short (incomplete) TR (recovery time), σ_1 and σ_2, the regular mode measurement error and the high precision, short burst mode measurement error, and p and q, the number of echoes (or data points) in the regular and the short burst modes.

Hence, to set up the combined data sets as one least-squares fit problem, we have

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ \vdots & & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

$$= \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1m} \\ \vdots & & & \vdots \\ \frac{\sigma_1}{\sigma_2} l_{11} & \frac{\sigma_1}{\sigma_2} l_{12} & \cdots & \frac{\sigma_1}{\sigma_2} l_{1m} \\ \vdots & & & \vdots \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

$$= \begin{bmatrix} b_1 \\ \vdots \\ \frac{\sigma_1}{\sigma_2} c_1 \\ \vdots \end{bmatrix} = \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix}, \qquad [7]$$

where A_{ij} represents the combined matrix with $n = p + q$, and g the combined data vector with n elements.

Note that the high precision data as well as its corresponding matrix elements have now been scaled by a factor of σ_1/σ_2, such that the combined data set has a common noise level of σ_1. The latter will be used for the cutoff in the SVD inversion procedure.

Once the combined data sets are set up as one least-squares fit problem with the proper weighting taken care of, the method for solving the equation need not to be limited to the SVD technique. The combined data set need not be restricted to two data sets, either.

In fact, we encounter a similar situation of varying variances when compressing the time domain data into several windows to increase the computation speed. For example, Eq. [2] can be transformed to

$$\sum_{i=1}^{s} \frac{1}{N_i \sigma^2} \left(\sum_{j=1}^{m} f_j K_{ij} - G_i \right)^2, \qquad [8]$$

where n echoes have been partitioned into s windows ($i = 1, \ldots, s$) with N_i echoes in the ith window, and

$$n = N_1 + \cdots + N_{i-1} + N_i + \cdots + N_s,$$

$$r_i = N_1 + \cdots + N_{i-1}, \quad r_1 = 0,$$

$$K_{ij} = \sum_{k=r_i+1}^{r_i+N_i} e^{-t_k/T_j},$$

$$G_i = \sum_{k=r_i+1}^{r_i+N_i} g_k. \qquad [9]$$

Now the variance for G_i is $N_i\sigma^2$, and each window has a different variance.

Similarly, we can also average the data within each window, assuming the linear approximation is valid, to obtain

$$\sum_{i=1}^{s} \frac{1}{(\sigma/\sqrt{N_i})^2} \left(\sum_{j=1}^{m} f_j e^{-t_i/T_j} - \bar{g}_i \right)^2, \qquad [10]$$

where $\bar{g}_i$ is simply the average of the data within the ith window, and t_i is an appropriate midpoint within the ith window. The noise level for the ith window is now reduced to $\sigma/\sqrt{N_i}$.

Regardless of time domain data sum or average, it is easy to rescale the matrix elements and the data vectors so that all windows have the same noise level, as we did in the composite-data processing. This common noise level can then be used as the cutoff for the SVD inversion process.

CUTOFF FOR SINGULAR VALUE DECOMPOSITION

Now that the common noise level for a composite-data set can be established, we need to understand how to choose the signal level for the inversion process and reexamine the criterion and the effect for the cutoff of singular values in the SVD inversion process. Recall that the singular value decomposition theorem in linear algebra (*8, 9*) states that any real $n \times m$ matrix $\mathbf{A}$ ($n \geq m$) can be written as the product of an $n \times n$ orthonormal matrix $\mathbf{U}$, an $n \times m$ diagonal matrix $\mathbf{\Lambda}$ with positive or zero diagonal elements, and the transpose of an $m \times m$ orthonormal matrix $\mathbf{V}$, i.e.,

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^{\mathrm{T}}, \qquad [11]$$

where

$$\mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I}, \qquad \mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I} \qquad [12]$$

$$\sum_{i=1}^{n} U_{ip}U_{iq} = \delta_{pq}, \quad 1 \leq p, q \leq n \qquad [13]$$

$$\sum_{i=1}^{m} V_{ip}V_{iq} = \delta_{pq}, \quad 1 \leq p, q \leq m \qquad [14]$$

and

$$\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m). \qquad [15]$$

The diagonal elements $\lambda_1 > \lambda_2 > \cdots > \lambda_m \geq 0$ are called the singular values of matrix $\mathbf{A}$.

The straightforward Householder transformations used for the decomposition of $\mathbf{A}$ need not produce all positive

singular values λ's in descending order. However, it is easy to make them all positive by making those that are negative positive and changing the sign of the corresponding eigenvectors, either the column vectors u_i or the row vectors v_i. It is also straightforward to rearrange them in a descending order by interchanging the corresponding column and row vectors. Hence, once the program is written in such a way to produce a $\mathbf{\Lambda}$ with all the singular values positive and in descending order, the information is contained in both **U** and **V** matrices. When the operator **A** is transformed to the coordinates of the principal axes, the **U** and **V** matrices also transform the solution vector **y** and data vector **b** into the corresponding principal axes. If we change the order of equations in the set of linear equations represented by **A**, the decomposed diagonal matrix $\mathbf{\Lambda}$ will always have the same singular values in the same descending order, whereas the **U** and **V** matrices will be different.

Let us consider the matrix equation

$$\mathbf{Ay} = \mathbf{b}, \tag{16}$$

where the matrix $\mathbf{A} \in \Re^{n\times m}$ operates on a vector $\mathbf{y} \in \Re^m$ to produce a vector $\mathbf{b} \in \Re^n$, and **A** has all properties described in Eqs. [11–15]. Then

$$\|\mathbf{Ay}\|^2 = \mathbf{y}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{Ay} = \mathbf{y}'^{\mathrm{T}}\mathbf{\Lambda}^{\mathrm{T}}\mathbf{\Lambda y}' = \|\mathbf{\Lambda y}'\|^2, \tag{17}$$

where $\mathbf{A} = \mathbf{U\Lambda V}^{\mathrm{T}}$, $\mathbf{y} = \mathbf{Vy}'$, and $\|\ \ \|$ represents the norm of a vector.

The largest (smallest) value of $\|\mathbf{Ay}\|^2$ occurs when **y** is along the direction of the eigenvector corresponding to the largest (smallest) singular value λ_1 (λ_m). Hence, for an arbitrary vector **y**, the following inequality is established:

$$\lambda_1^2 \geq \frac{\|\mathbf{Ay}\|^2}{\|\mathbf{y}\|^2} \geq \lambda_m^2, \tag{18}$$

which means for the vector **b**,

$$\lambda_1^2 \geq \frac{\|\mathbf{b}\|^2}{\|\mathbf{y}\|^2} \geq \lambda_m^2. \tag{19}$$

Frequently, we measure **b** and try to determine **y**. There is always measurement error involved in obtaining **b**. Therefore, instead of **b** and **y**, we have $\mathbf{B} = \mathbf{b} + \delta\mathbf{b}$ and $\mathbf{Y} = \mathbf{y} + \delta\mathbf{y}$, and we are faced with a problem of

$$\mathbf{A}(\mathbf{y} + \delta\mathbf{y}) = \mathbf{b} + \delta\mathbf{b} \tag{20}$$

rather than

$$\mathbf{Ay} = \mathbf{b}, \tag{21}$$

where **y** and **b** represent, respectively, the ideal solution and the ideal data vector without measurement errors, and $\delta\mathbf{y}$ and $\delta\mathbf{b}$ represent, respectively, the deviation in the solution and the measurement error of the data vector.

To satisfy Eqs. [20] and [21], we must have

$$\mathbf{A}\delta\mathbf{y} = \delta\mathbf{b}. \tag{22}$$

Assuming that the matrix **A** is of rank m, its form in the principal coordinates is represented by $\mathbf{\Lambda}$ with m nonvanishing descending singular values $\lambda_1, \ldots, \lambda_m$, and the normalized left and right singular vectors are denoted by $u_1, \ldots, u_n$, and $v_1, \ldots, v_m$, respectively, we can then express $\mathbf{y}'$ and $\mathbf{b}'$ as

$$\mathbf{y}' = \mathbf{V}^{\mathrm{T}}\mathbf{y} \tag{23}$$

$$\mathbf{b}' = \mathbf{U}^{\mathrm{T}}\mathbf{b}, \tag{24}$$

where the lengths of both **y** and **b** are not influenced by the orthogonal transformation, i.e., $\|\mathbf{y}\| = \|\mathbf{y}'\|$ and $\|\mathbf{b}\| = \|\mathbf{b}'\|$.

Note that since the operation of **A** is restricted to the subspace of **U** (i.e., spanned by $u_1, \ldots, u_m$), both **b** and $\mathbf{b}'$ should be in that subspace if the model for inversion is correct. The portion of both **b** and $\mathbf{b}'$ in the subspace spanned by $u_{m+1}, \ldots, u_n$ is a measure of how inaccurate the model is.

When **A** is transformed to the coordinates of the principal axes, we have

$$\mathbf{\Lambda y}' = \mathbf{\Lambda V}^{\mathrm{T}}\mathbf{y} = \mathbf{U}^{\mathrm{T}}\mathbf{b} = \mathbf{b}' \tag{25}$$

$$\mathbf{\Lambda}\delta\mathbf{y}' = \mathbf{\Lambda V}^{\mathrm{T}}\delta\mathbf{y} = \mathbf{U}^{\mathrm{T}}\delta\mathbf{b} = \delta\mathbf{b}'. \tag{26}$$

Assuming that different b_i's have errors that are statistically uncorrelated and have the same σ, then

$$\langle \delta b_i \delta b_j \rangle = \sigma^2 \delta_{ij} \tag{27}$$

and

$$\langle \delta b_i' \delta b_j' \rangle = \langle \sum_k U_{ki}\delta b_k \sum_l U_{lj}\delta b_l \rangle = \sigma^2\delta_{ij} \tag{28}$$

$$\langle \delta y_i' \delta y_j' \rangle = \left\langle \frac{\delta b_i'}{\lambda_i}\frac{\delta b_j'}{\lambda_j} \right\rangle = \frac{\sigma^2}{\lambda_i^2}\delta_{ij} \tag{29}$$

$$\langle \delta y_i \delta y_j \rangle = \langle \sum_k V_{ik}\delta y_k' \sum_l V_{jl}\delta y_l' \rangle = \sum_k \frac{\sigma^2}{\lambda_k^2} V_{ik}V_{jk}. \tag{30}$$

The requirement for a good solution, $\langle \delta y_i^2 \rangle \ll y_i^2$, is thus translated into

$$\sum_k \frac{\sigma^2}{\lambda_k^2} V_{ik}^2 \ll y_i^2 = (\sum_k V_{ik} y'_k)^2$$

$$= \left(\sum_k \frac{V_{ik}}{\lambda_k} b'_k \right)^2 = \left(\sum_{kl} \frac{V_{ik} U_{lk}}{\lambda_k} b_l \right)^2 . \quad [31]$$

A necessary condition for this to be true for every i is that the sum obeys the same inequality

$$\sum_{i=1}^m \langle \delta y_i^2 \rangle \ll \sum_{i=1}^m y_i^2 . \quad [32]$$

The r.h.s is

$$\sum_{i=1}^m y_i^2 = \sum_{i=1}^m y_i'^2 = \sum_{i=1}^m \frac{b_i'^2}{\lambda_i^2} \quad [33]$$

and the l.h.s is

$$\sum_{i=1}^m \langle \delta y_i^2 \rangle = \sum_{i=1}^m \langle \delta y_i'^2 \rangle = \sum_{i=1}^m \frac{\langle \delta b_i'^2 \rangle}{\lambda_i^2} = \sum_{i=1}^m \frac{\sigma^2}{\lambda_i^2} . \quad [34]$$

Thus we need to satisfy

$$\sum_{i=1}^m \frac{\sigma^2}{\lambda_i^2} \ll \sum_{i=1}^m \frac{b_i'^2}{\lambda_i^2} . \quad [35]$$

This equation can be satisfied if $\sigma^2 \ll b_i'^2$ for all i. However, it can fail if the m supposedly independent equations are actually nearly nonindependent which results in some very small singular values. In particular, those values may dominate the two sums of [35]. If for those small λ_i's, $|b'_i|$ is not much larger than σ, then [35] will not be satisfied. We can remedy this by discarding a sufficient number of small λ_i, e.g., all λ_i, $i > r$, such that the more restricted sums (i.e., for $i = 1, \ldots, r$; $r \leq m$) satisfy the inequality to a desired degree

$$\sum_{i=1}^r \frac{\sigma^2}{\lambda_i^2} = \epsilon^2 \sum_{i=1}^r \frac{b_i'^2}{\lambda_i^2} , \quad [36]$$

where ϵ is a suitably chosen small parameter.

Apparently, a data set with a very high signal-to-noise ratio will allow us to choose a small ϵ. Instead of choosing a constant ϵ, it is convenient to choose an ϵ commensurate with the noise level. From the discussion of Eqs. [18] and [19], it seems appropriate to choose

$$\frac{\lambda_1}{\lambda_r} = \frac{\|\mathbf{b}'\|}{\sigma} = \frac{\|\mathbf{b}\|}{\sigma} \approx \frac{\|\mathbf{b} + \delta\mathbf{b}\|}{\sigma} , \quad [37]$$

where λ_r is the cutoff and is determined by $\|b + \delta b\|/\sigma$. Such a choice corresponds to an ϵ^2 given by

$$\epsilon^2 \equiv \frac{\|b'\|^2/\lambda_1^2}{\sum_{i=1}^r b_i'^2/\lambda_i^2} \frac{\sum_{i=1}^r 1/\lambda_i^2}{1/\lambda_r^2} . \quad [38]$$

Once a proper cutoff such as [37] is determined, the uncertainty δy_i for each y_i can be calculated from [31]. The choice we made for the cutoff, i.e., Eq. [37], is a consistent and reasonable one. On the other hand, it is also somewhat subjective as the requirement of [32] is not very well defined. A slightly lower cutoff can be used and will provide more information at the risk of admitting components with higher uncertainties, and vice versa.

The singular values smaller than λ_r (i.e., all λ_i, $i > r$) which are discarded in the restricted sums are effectively replaced by 0 in the inverse matrix $\mathbf{\Lambda}^{-1}$. This means matrix $\mathbf{\Lambda}^{-1}$ is effectively a smaller matrix $\mathbf{\Lambda}_r^{-1}$ which has only r nonzero diagonal elements, and both $\mathbf{U}$ and $\mathbf{V}$ are replaced by smaller $\mathbf{U}_r$ and $\mathbf{V}_r$ spanned by $u_1, \ldots, u_r$, and $v_1, \ldots, v_r$ ($r \leq m$), respectively.

Hence the solution $\mathbf{Y} = \mathbf{y} + \delta\mathbf{y}$ is given by

$$\mathbf{Y} = \mathbf{A}^{-1}\mathbf{B} = \mathbf{V}_r \mathbf{\Lambda}_r^{-1} \mathbf{U}_r^{\mathrm{T}} (\mathbf{b} + \delta\mathbf{b}), \quad [39]$$

where $\mathbf{A}^{-1}$ is a pseudo inverse operating only in the subspace spanned by $\mathbf{V}_r$ and $\mathbf{U}_r$. The measurement error σ forces us to cut off from λ_r. This means that part of $\mathbf{b}$ that is in the $m - r$ dimensional space is masked by noise and becomes part of $\delta\mathbf{b}$. The remaining $\mathbf{b}$ (which we refer to as $\mathbf{b}$) is in the r dimensional space, and $\delta\mathbf{b}$ is in the $m - r$ dimensional space which is always normal to the subspace $\mathbf{U}_r$.

Therefore, we have

$$\|\mathbf{b} + \delta\mathbf{b}\|^2 = \|\mathbf{b}\|^2 + \|\delta\mathbf{b}\|^2 \quad [40]$$

and

$$\|\mathbf{AY} - \mathbf{B}\|^2 = \|\mathbf{AY} - \mathbf{b}\|^2 + \|\delta\mathbf{b}\|^2, \quad [41]$$

and the l.h.s. is at its minimum when $\mathbf{AY} - \mathbf{b} = 0$. When there is no constraint on $\mathbf{Y}$, a solution for $\mathbf{AY} - \mathbf{b} = 0$ can always be found. However, if some of the elements of $\mathbf{Y}$ happen to be negative, and the nonnegativity constraint is imposed, then $\mathbf{Y}$ is no longer at its calculus minimum and $\|\mathbf{AY} - \mathbf{b}\|^2 \geq 0$.

GEOMETRIC PICTURE FOR SVD

In fact, we can try to understand the properties of the SVD inversion process with a geometric picture. From [25], we get

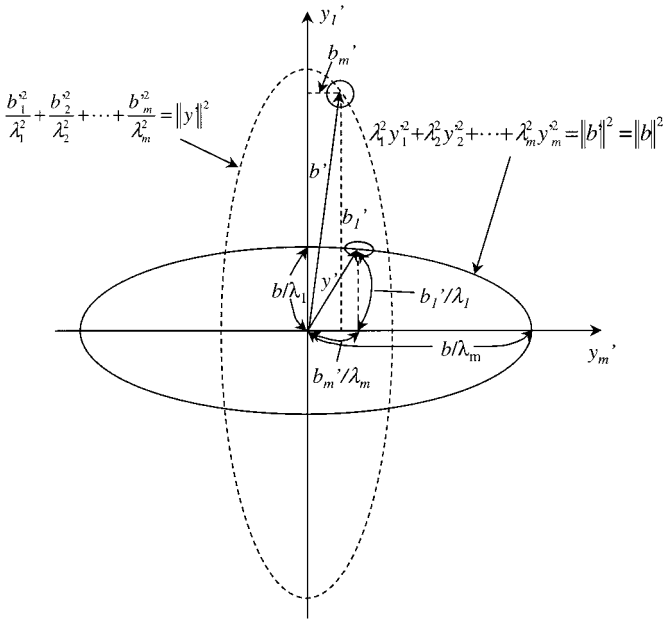

FIG. 2. The m-dimensional hyperellipsoids depicting relationships between the data and solution vectors.

$$\frac{b_1'^2}{\lambda_1^2}+\frac{b_2'^2}{\lambda_2^2}+\cdots+\frac{b_m'^2}{\lambda_m^2}=\|\mathbf{y}'\|^2 \qquad [42]$$

and

$$\lambda_1^2 y_1'^2+\lambda_2^2 y_2'^2+\cdots+\lambda_m^2 y_m'^2=\|\mathbf{b}'\|^2. \qquad [43]$$

As shown in Fig. 2, **A** operates on a vector $\mathbf{y}'$ and produces an ideal data vector $\mathbf{b}'$. The vector $\mathbf{b}'$ should have a relationship with $\mathbf{y}'$ according to [25], and $\mathbf{b}'$ and $\mathbf{y}'$ should be situated on the surface of m-dimensional hyperellipsoids described by [42] and [43], respectively.

Then the inversion process would be, given a data vector $\mathbf{b}'$ with measurement error (depicted in the figure as a sphere about the vector $\mathbf{b}'$, illustrating the random nature of the noise), and try to determine a vector $\mathbf{y}'$ (also depicted in the figure with an ellipsoid representing uncertainties about the solution). Note the lengths of the radius vectors along the principal axes for the b'-ellipsoid are proportional to singular values λ_i, whereas those for the y'-ellipsoid are inversely proportional to singular values λ_i.

The inversion process is equivalent to squeezing the b'-ellipsoid along all its principal axes by a factor proportional to λ_i, where $\mathbf{b}'$ will now become $\mathbf{y}'$, and $\mathbf{y}'$ will be located on the surface of the y'-ellipsoid. Naturally, any error in $\mathbf{b}'$ (the spherical shade) will be translated to the error in $\mathbf{y}'$ (the ellipsoidal shade). Note that the uncertainty of $\mathbf{y}'$ along the axis associated with the largest singular value, λ_1, is relatively small compared with that associated with the smallest singular value, λ_m. In fact, the error in $\mathbf{y}'$ along the principal axes (i.e., $v_1, \ldots, v_m$) is inversely proportional to the singular values λ_i. That's why the sphere, representing the isotropic noises in the data space, translates to an ellipsoid representing the uncertainties in the solution space.

To illustrate how the discarding of small singular values affects the uncertainty of the solution **y**, we draw in Fig. 3 an ellipsoid representing the uncertainty of $\mathbf{y}'$ in the solution space. The equation for the ellipsoid

$$\lambda_1^2\delta y_1'^2+\lambda_2^2\delta y_2'^2+\cdots+\lambda_m^2\delta y_m'^2=\|\delta\mathbf{b}'\|^2 \qquad [44]$$

is obtained from [26], where the lengths of radius vectors are inversely proportional to λ_i. By discarding small singular values, i.e., λ_2, as shown in Fig. 3, one effectively collapses the y_2' axis to zero (i.e., replacing $1/\lambda_2$ by 0 in the inverse matrix $\mathbf{\Lambda}^{-1}$). As a result, the uncertainties in y_1 and y_2 are greatly reduced.

RESULTS AND DISCUSSION

A typical example for implementing [36] is shown in Fig. 4. Track 1 shows the porosity and the uncertainties computed from [34]. Track 2 shows the parameter ϵ (EPSILON) computed from [38] when the cutoff is determined by [37]. This is equivalent to a value of ϵ^2 around 0.01, a reasonably small number. Track 2 also shows another parameter, UXP, which is defined as

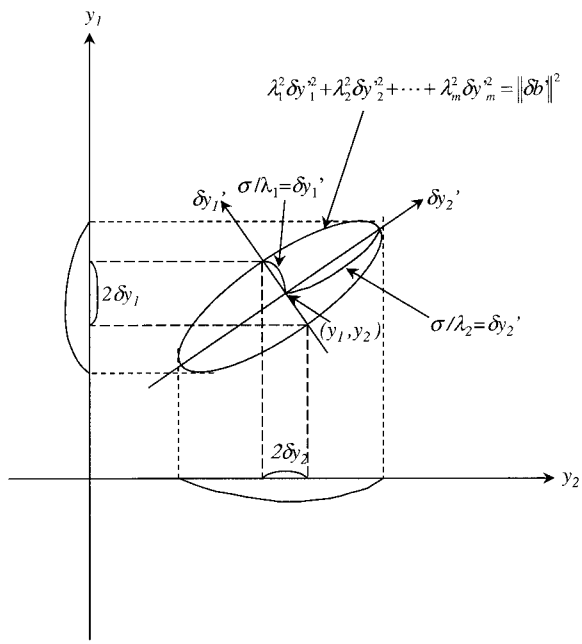

FIG. 3. The hyperellipsoid of the uncertainty of the solution. Discarding the small singular values has the effect of reducing the uncertainty of the solution vector.

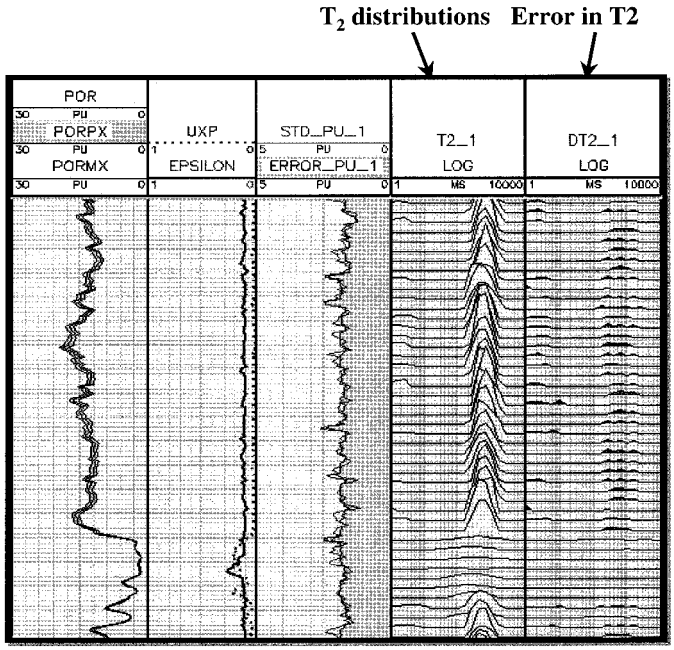

FIG. 4. The log plot showing estimated porosity, T_2 distributions, and error in each T_2 component from an NMR logging tool. See text for description of each track.

$$\mathtt{UXP} = \sqrt{\frac{\|b + \delta b\|^2 - \|b'\|^2}{\|b + \delta b\|^2}}, \qquad [45]$$

where $\|b'\|$ represents only the portion of the data vector in space spanned by $u_1, \ldots, u_r$. Thus the quantity, UXP, is a measure of how accurate the model is. A small value of UXP means only a very small component of the data vector is in a space not associated with the matrix **A**. Hence, the inversion model, namely **A**, is quite accurate. On the other hand, if UXP is large, then the model is not accurate.

Track 3 shows the standard deviations of the noise channel (ERROR) and the fit of the signal (echo train) channel (STD). Track 4 shows the T_2 distributions, and Track 5 shows the uncertainties of all T_2 components computed from [31] at the same scale. Note that there are some short T_2 components in Track 4 which have the same order of magnitude as the uncertainties shown in Track 5. This is an indication that these short T_2 components are not to be trusted.

The uncertainty of the solution which we discuss here refers to the precision, not the accuracy, of the solution, whereas the accuracy of the model is addressed by the parameter, UXP. Note that while the standard deviation of the fit, STD, is relatively constant for different porosities, the uncertainty of the solution gets larger for larger porosity as shown in Track 1. The relative uncertainty, i.e., the ratio of uncertainty to porosity, gets smaller as the porosity gets larger, but the absolute uncertainty in porosity units gets larger for larger porosities. This can be understood in terms of the mathematical procedure of removing the singular values commensurate with the noise level. When the noise level is low, more singular values are included in the inversion, hence $\|\delta y\|^2$ increases. As the noise level increases, more singular values are discarded, hence $\|\delta y\|^2$ decreases. In the limiting case, the uncertainty approaches zero for zero signal. But this is also when the model breaks down, and the parameter, UXP, approaches 1.

Figure 5 shows the results of composite-data processing method when applied to a typical data set of Numar's MRIL-CTP. Track 1 shows the results of a splicing technique where the raw data of regular mode and the short burst mode were processed separately and then spliced together (*7*). A constant cliff in the T_2 distribution as a function of depth is clearly visible. This discontinuity is an apparent artifact from the processing. Track 2 shows the results of the composite-data processing where both data sets were processed simultaneously, resulting in a smooth T_2 distribution. There are several places where the result of composite-data processing show a very different T_2 behavior.

We made, however, an important assumption in the composite-data processing method, i.e., the knowledge of T_1/T_2 ratio. The result of composite-data processing shown in Fig. 5 assumed a constant T_1/T_2 ratio of 1.5 for the entire log. In fact, such an assumption is not necessary. One can extract the T_1/T_2 ratio from the raw data, simply by scanning through the possible T_1/T_2 ratios, and choose the one with the best fit. The result of such an exercise is shown in Fig. 6.

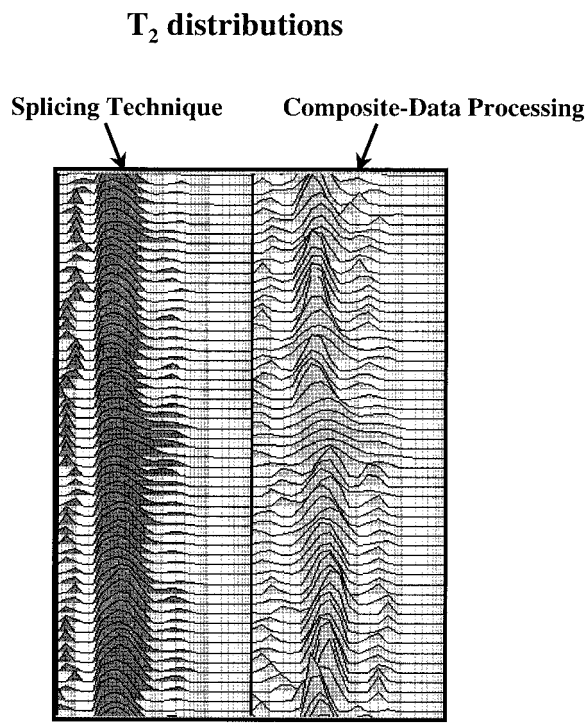

FIG. 5. The T_2 distributions obtained from the splicing technique and composite-data processing method. The constant cliff shown in the splicing technique is a processing artifact.

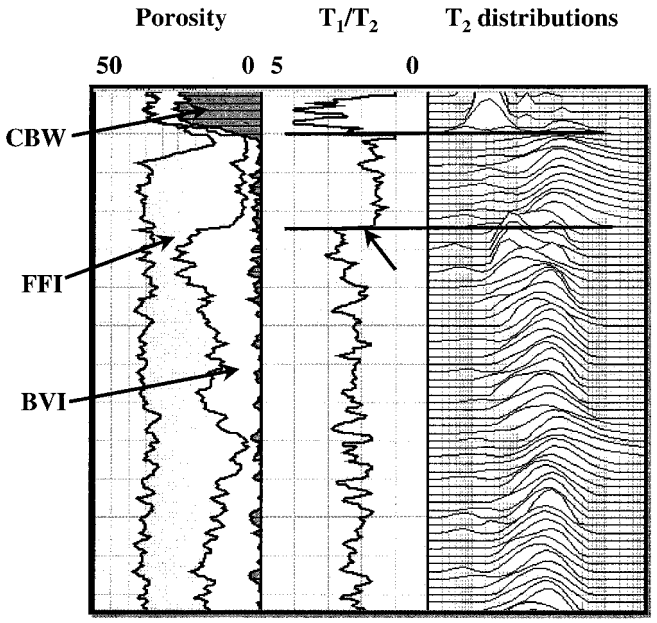

FIG. 6. The log plot showing the T_2 distribution and T_1/T_2 ratio obtained from an NMR log using the composite-data processing method.

Even though it is commonly accepted (*10*) that such a ratio exists for the entire time scale, there are certainly exceptions. A good example would be when there are both oil and water present in the pore space, each having very different T_1/T_2 ratios. However, such exceptions can be used to our advantage as a diagnostic tool for the log interpretation.

For water in a water–wet interface situation, we have

$$\frac{1}{T_1} \approx \frac{1}{T_{1s}} \qquad \frac{1}{T_2} \approx \frac{1}{T_{2s}} + \frac{1}{T_{2D}} \qquad \frac{1}{T_{2D}} = \frac{1}{3}\,\gamma^2 G^2 D_w \tau^2 \quad [46]$$

where $1/T_{1s}$ and $1/T_{2s}$ are the surface relaxation terms, $1/T_{2D}$ is from the diffusion effect in a gradient field, γ is the gyromagnetic ratio, G is the gradient, D_w is the diffusion coefficient for water, and τ is half the echo spacing. The apparent T_1/T_2 ratio is then

$$\frac{T_1}{T_2} = \frac{T_{1s}}{T_{2s}}\left(1 + \frac{T_{2s}}{T_{2D}}\right), \quad [47]$$

where T_{1s}/T_{2s} is the ratio in a homogeneous field and the factor $(1 + T_{2s}/T_{2D})$ reflects the effect of a gradient field on the ratio.

Thus, in a gradient field (as is the case for NMR logging tools being used nowadays) and in a water zone, a formation with predominantly long T_2 components would have a larger apparent T_1/T_2 ratio than that with predominantly short T_2 components. When going into the oil zone, such T_1/T_2 may undergo an abrupt change, presumably to a smaller value because the diffusion coefficient for oil is much smaller. Hence, this apparent T_1/T_2 ratio can be used for diagnostic purposes.

In Track 1 of Fig. 6, we show the total porosity, clay-bound water, capillary-bound water, and free fluid index. In Track 2, we show the apparent T_1/T_2 ratio, and in Track 3, the T_2 distribution. There is a zone where the apparent T_1/T_2 ratio made a jump to a smaller value. This can be interpreted as an indication of the presence of oil. However, since the T_1/T_2 ratio is not a very well-defined quantity, care should be taken in the use and the interpertation of this apparent quantity derived from logs.

As pointed out earlier in the introduction, when the T_2 relaxation times are preselected to be equally spaced logarithmically, the inversion is a linear regression problem. The linear inversion has many advantages. It is very robust, computationally efficient, and ideal for log data processing. It also provides a common reference, i.e., the preselected logarithmically equally spaced T_2 relaxation times, for the discussion of T_2 amplitude (pore size) distribution, T_2 cutoff for irreducible water saturation, etc.

However, a set of preselected logarithmically equally spaced T_2 relaxation times may not always be a good model. In some of the data sets we analyzed, we found that the portion of the data vector $\mathbf{b} + \delta\mathbf{b}$ that is in the space spanned by $u_{m+1}, \ldots, u_n$ is not insignificant.

In fact, the characteristic relaxation times of a fluid-saturated porous medium are usually not very many, and they need not be equally spaced logarithmically, either. However, to treat the relaxation times as unknowns is a nonlinear inversion problem. It is inherently more difficult and sometimes unstable to use for log processing. Without having a common reference of a suite of preselected T_2 relaxation times, it is also difficult to compare physical quantities on a comparative basis. For certain applications, however, nonlinear inversion may still be needed to reveal certain characteristics lost in the regularization procedure.

ACKNOWLEDGMENTS

The authors acknowledge helpful discussions with David J. Bergman, Robert J. S. Brown, and Simon W. Stonard, especially DJB, whose many valuable comments are included in this report.

REFERENCES

1. M. N. Miller, Z. Paltiel, M. E. Gillen, J. Granot, and J. C. Bouton, SPE Paper 20561, 65th SPE Annual Tech. Conf. and Exhib., New Orleans, LA (1990).
2. R. L. Kleinberg, A. Sezginer, D. D. Griffin, and M. Fukuhara, *J. Magn. Reson.* **97,** 466 (1992).
3. R. Freedman and C. E. Morriss, SPE Paper 30560, 70th SPE Annual Tech. Conf. and Exhib., Dallas, TX (1995).

4. M. G. Prammer, E. D. Drack, J. C. Bouton, and J. S. Gardner, SPE Paper 36522, 71st SPE Annual Tech. Conf. and Exhib., Denver, CO (1996).

5. M. G. Prammer, J. C. Bouton, R. N. Chandler, E. D. Drack, and M. N. Miller, SPE Paper 49011, 73rd SPE Annual Tech. Conf. and Exhib., New Orleans, LA (1998).

6. S. Menger and M. G. Prammer, SPE Paper 49013, 73rd SPE Annual Tech. Conf. and Exhib., New Orleans, LA (1998).

7. K. J. Dunn, D. J. Bergman, G. A. LaTorraca, S. W. Stonard, and M. B. Crowe, M. B., Paper JJ, 39th Annual Logging Symposium of SPWLA, Keystone, CO, May 26–29 (1998).

8. C. Lanczos, "Linear Differential Operators," Van Nostrand, London, Chapter 3 (1961).

9. C. L. Lawson and R. J. Hanson, "Solving Least Squares Problems," Prentice–Hall, Englewood Cliffs, NJ (1974).

10. R. L. Kleinberg, C. Straley, W. E. Kenyon, R. Akkurt, and S. A. Farooqui, SPE Paper 26470, 68th SPE Annual Tech. Conf. and Exhib., Houston, TX (1993).