Simultaneous Gaussian and exponential inversion for improved analysis of shales by NMR relaxometry

Authors: Kathryn E. Washburn, Endre Anderssen, Sarah J. Vogt, Joseph D. Seymour, Justin E. Birdwell, Catherine M. Kirkland, & Sarah L Codd

NOTICE: this is the author’s version of a work that was accepted for publication in Journal of Magnetic Resonance. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Journal of Magnetic Resonance, 250, January 2015. DOI# http://dx.doi.org/10.1016/j.jmr.2014.10.015.


Made available through Montana State University’s ScholarWorks scholarworks.montana.edu
Simultaneous Gaussian and exponential inversion for improved analysis of shales by NMR relaxometry

Kathryn E. Washburn, Endre Anderssen, Sarah J. Vogt, Joseph D. Seymour, Justin E. Birdwell, Catherine M. Kirkland, Sarah L Codd

Abstract

Nuclear magnetic resonance (NMR) relaxometry is commonly used to provide lithology-independent porosity and pore-size estimates for petroleum resource evaluation based on fluid-phase signals. However, in shales, substantial hydrogen content is associated with solid and fluid signals and both may be detected. Depending on the motional regime, the signal from the solids may be best described using either exponential or Gaussian decay functions. When the inverse Laplace transform, the standard method for analysis of NMR relaxometry results, is applied to data containing Gaussian decays, this can lead to physically unrealistic responses such as signal or porosity overcall and relaxation times that are too short to be determined using the applied instrument settings. We apply a new simultaneous Gaussian-Exponential (SGE) inversion method to simulated data and measured results obtained on a variety of oil shale samples. The SGE inversion produces more physically realistic results than the inverse Laplace transform and displays more consistent relaxation behavior at high magnetic field strengths. Residuals for the SGE inversion are consistently lower than for the inverse Laplace method and signal overcall at short T2 times is mitigated. Beyond geological samples, the method can also be applied in other fields where the sample relaxation consists of both Gaussian and exponential decays, for example in material, medical and food sciences.

Introduction

Nuclear magnetic resonance (NMR) has been used to characterize petroleum reservoirs for decades [1–3]. Because measurement of the transverse relaxation rate via the Carr–Purcell–Meiboom–Gill (CPMG) pulse sequence [4,5] is robust and fast, it is a staple technique in NMR evaluation of petroleum resources [6]. While carbon and sodium nuclei have been investigated at high magnetic fields [7–11], due to the low field strength of the logging tools, NMR on a commercial basis in the oil industry has focused almost exclusively on the hydrogen nucleus. In conventional petroleum reservoirs, the NMR signal is assumed to arise solely from liquids located in the pore space and, as such, NMR logging can provide lithology-independent porosity estimates for reservoir evaluation [12]. Because conventional reservoirs consist primarily of sand-stone and carbonates, which contain little hydrogen in the matrix itself, porosities determined by NMR are accurate enough to be useful. However, growing global energy demands and declining discoveries of conventional reservoirs have led to increased evaluation of and production from unconventional petroleum resources in recent years. Some of the most productive unconventional resources are shales, which may produce either oil, gas or both. Unlike conventional reservoirs, shales may contain significant quantities of clay and organic matter, which are often rich in solid or immobile (clay-bound) hydrogen [13–15]. Therefore, the resulting signal can no longer be assumed to arise solely from fluid porosity [16,17] and this will affect the results in several ways.

NMR relaxometry has been used to estimate pore sizes in reservoir rocks [18], allowing estimation of system permeability [19–21]. Bulk fluids have a relaxation rate dependent on the internal interactions of their molecules [22]. When fluids are placed inside
a porous medium, interactions between the fluid molecules and pore surfaces cause an increase in the rate of relaxation [18]. Because smaller pores have a larger surface area-to-volume ratio, fluid molecules present in small pores will encounter the pore surface more frequently than in larger pores, leading to an increase in relaxation rate proportional to the pore size. The relaxation rate is described by:

\[
\frac{1}{T_2} = \frac{1}{T_{2\text{bulk}}} + \rho S
\]  

where \( T_{2\text{bulk}} \) is the bulk transverse relaxation time constant of the fluid (\( \mu \text{s} \)), \( \rho \) is the surface relaxation (m/\( \mu \text{s} \)) and \( S/N \) (m\(^2\)/m\(^3\)) is the surface area-to-volume ratio of the pore. For conventional reservoirs, \( \rho \) is associated with paramagnetic impurities [23] and does not contain significant temperature dependence. However in unconventional samples, surface relaxivity may arise from homonuclear dipolar coupling as well [24, 25], which will cause it to have a dependence on temperature.

The measured data are a sum of all the relaxing components in the system. In order to interpret the results, an inverse Laplace transform is applied to the decay to obtain a distribution of characteristic relaxation times that, in fluid-saturated conventional reservoir rocks, will reflect the distribution of pore sizes present. There are a variety of methods available for this analysis. The most common way of implementing an inverse Laplace transform in the oil industry is to apply the Butler–Reed–Dawson (BRD) method [26], though approaches such as a non-negative least squares (NNLS) [27] or a uniform penalty (UPEN) function [28] are sometimes applied. Other direct measurement methods have been implemented to analyze relaxometry results [29–31].

Analysis of NMR shale relaxometry is more challenging than for conventional reservoirs because of the addition of solids signal that does not arise from fluid filled porosity and therefore does not follow Eq. (1), which relates relaxation rate to pore size. An added challenge in the interpretation of NMR shale results is that solids and liquids have different transverse relaxation behavior. While solid signal decays are readily described using an exponential decay function, solids may decay according to Gaussian or exponential decay functions, depending on the motional regime [22]. The motional regime is determined by a combination of the applied magnetic field strength and the rotational correlation time of the molecules. Therefore, when hydrogen-rich solids or solids and liquids exist within the same sample, the measured relaxation may be a combination of Gaussian and exponential decays described by:

\[
M(t) = \sum_i A_i e^{-(t/T_{2i})} + B_i e^{-(t/T_{3i})}
\]

where \( M(t) \) is the measured signal, \( A_i \) is the Gaussian intensity of the \( T_{2i} \) component, and \( B_i \) is the exponential intensity of the \( T_{3i} \) component. This confounds the standard inverse Laplace transform methods used to analyze CPMG data. Previous work [24] has shown that attempts to invert Gaussian decays using exponentials will lead to significant overcall in the signal and relaxation times that are erroneously short, sometimes producing physically unrealistic results; significant peaks may appear at times well below the echo spacing of the CPMG. Despite the fact that it is well established in the literature that organic matter may produce relaxation consisting of both exponential and Gaussian decays [32–37], analysis of organic rich shale NMR relaxation results by the standard inverse Laplace transform methods is still the norm [17, 38–40] due to lack of alternative inversion methods. Therefore, the development of new inversion methods is needed to analyze shales to avoid the introduction of artifacts to the results.

In order to address this issue, we have developed a generalized inversion to simultaneously solve for Gaussian and exponential decays. Previous inversions required a priori knowledge about the number of components to be fitted [41, 42]. In recent work by Zhu et al. [43] a generalized inversion was developed that did not require input of a fixed number of relaxation components. However, this technique employs a hard cutoff time; everything below a certain relaxation time is assumed to belong to Gaussian components and everything at longer times to exponential components. For complex materials such as shale, applying a hard cutoff time has several problems. Organic matter in shales, such kerogen and bitumen, consists of large organic molecules with no fixed molecular structure and their composition may widely vary between reservoirs. In addition to differences between samples, measurement frequency and sample temperature will also affect the Gaussian-exponential relaxation transition in the sample [34]. In complex samples with many constituents, if there are other effects influencing the relaxation rate, e.g. spin diffusion, more than one transition time between the Gaussian and exponential domains can potentially exist. The association of different organic molecules with the mineral structure may also impact the relaxation behavior. All these issues create a challenge in determining the optimal cutoff time between the Gaussian and exponential regimes. In contrast, we incorporate a sigmoid penalty function into the analysis, decreasing the need for exact knowledge of the Gaussian-exponential transition and allowing for the possibility of overlapping Gaussian and exponential components:

\[
R = M(t) - M(t) + \sum_i A_i \times \text{Sig}_A + B_i \times \text{Sig}_B
\]

where \( R \) is the residual to be minimized, \( M(t) \) is the measured data, \( M(t) \) is the calculated data, \( \text{Sig}_A \) is the sigmoid penalty function ramping up and \( \text{Sig}_B \) is the penalty function ramping down. \( \text{Sig}_A \) and \( \text{Sig}_B \) are vectors of length of \( A \) and \( B \) respectively and are logistic functions. \( \text{Sig}_A \) is used to penalize results that produce Gaussian fits at times above the sigmoid center while \( \text{Sig}_B \) penalizes results that produce exponential fits below the sigmoid center. \( \text{Sig}_B \) is described by:

\[
\text{Sig}_B = \frac{1}{1 + \exp(-\text{Sig}_{\text{center}} \times \text{Sig}_{\text{width}})} \times \text{Sig}_{\text{weight}} + \alpha
\]

where \( \text{Sig}_{\text{center}} \) is the center of the sigmoid penalty function. It is expressed as an index of the desired threshold time in the vector of the \( T_2 \) values to be solved over. \( \text{Sig}_{\text{width}} \) is the width of the sigmoid and is applied over the logarithmically spaced \( T_2 \) values. \( \text{Sig}_{\text{weight}} \) is the amount of penalty for a peak to be located beyond the threshold. \( \alpha \) is the smoothing factor. Similarly, the penalty function for the exponential relaxation is given by:

\[
\text{Sig}_B = \left( 1 - \frac{1}{1 + \exp(-\text{Sig}_{\text{center}} \times \text{Sig}_{\text{width}})} \right) \times \text{Sig}_{\text{weight}} + \alpha
\]

As Gaussian signal is more likely to appear at shorter \( T_2 \) times than exponential signal, a sigmoid penalty function was applied to drive the inversion to select solutions that reflect that physical situation, but does not completely restrict Gaussians to short relaxation times. Similarly, the penalty function encourages exponential peaks at longer \( T_2 \) times, but will also allow for exponential relaxation at short relaxation times. This allows the inversion to function when there is not a clear delineation between Gaussian and exponential relaxation times. For example, in the case where a system may contain both organic-rich solids and fluids trapped in very small pores, both Gaussian and exponential decay signals may be present at similar short relaxation times.

The combined SGE inversion was written in the R programming language (Free Software Foundation Inc., Boston, MA) using a sum of squares method. The inversion code is listed in Appendix A. The optimization was performed using the standard R optimization
package, optim, with the Nelder–Mead method [44]. However, there is no reason the inversion needs to be restricted to this particular approach and most nonlinear optimization methods should function adequately. The optimization operates by altering an input to minimize a function. For the SGE inversion, the input is a vector of A and B, the Gaussian and exponential $T_2$ peak distributions, and the function is $R$, the sum of the squares of the difference between the measured data and the reconstructed data plus the penalty function. The reconstructed data are created from A and B using the time axis of the measured data and range of $T_2$ times supplied by the user. The optimization searches the solution space of A and B until convergence upon a minimum value of $R$ is reached.

This inversion was applied to a range of simulated data and high field NMR transverse relaxometry results collected on oil shale samples and the results compared to standard processing methods. The method developed herein focuses on shales and geological samples, but the inversion technique has potential for application to a wide range of fields such as polymer science, food science, agriculture and soil science. Many diverse materials such as rubber [45,46], cement [47], wood [48], bone marrow [49], ice cream [50], and cheese [51] all contain heterogeneous mixtures of solid and liquid components which can exhibit combined Gaussian and exponential relaxation behavior. While not attempted here, the SGE inversion could also be adapted for two-dimensional data such as $T_1$–$T_2$ or $T_2$–diffusion correlations.

2. Experimental

All transverse relaxation measurements were made with the standard CPMG pulse sequence, acquiring 3000 echoes. No baseline offset nor oscillations in the initial echoes of the CPMG were observed in the raw data. Measurements were performed on a Bruker 300 MHz Avance III spectrometer using a 5 mm rf coil in a Micro2.5 probe (Karlsruhe, Germany) with an echo spacing of 22 $\mu$s. The rf pulse length was 4.5 $\mu$s for the 90° pulse and 9 $\mu$s for the 180°. Acquisition dead time was 0.025 $\mu$s.

Samples tested included a variety of oil shales, or organic-rich, thermally immature sedimentary rocks [14], from around the world. These oil shales represent different geologic ages and depositional environments as well as a range of organic richness and mineralogy. Qualitative mineralogy was determined by attenuated total reflectance Fourier transform infrared spectroscopy using a method described previously [52]. Samples included the Cambrian Alum shale (marine, quartz-illite rich, Sweden), Jurassic Kimmeridgian Blackstone (marine, clay-rich, UK), Tertiary Condor oil shale (lacustrine, clay-rich, Australia), Permian Glen Davis torbanite (lacustrine, quartz-rich, Australia), Eocene Green River Mahogany zone oil shale (lacustrine, dolomitic marlstone, Colorado, USA), Cretaceous Juliana Creek shale (marine, carbonate-rich, Australia), Ordovician Narva-E kükse (marine, carbonate-rich, Estonia), Permian phosphorite shale (marine, fluorapatite–quartz, Montana, USA), and Cretaceous Timaidoh oil shale (marine, clay-carbonate, Morocco). These shales contain small quantities of highly viscous bitumen, organic matter which is soluble in chloroform, and large quantities of kerogen, solid organic matter insoluble in most organic solvents. Quantification of the organic material in the samples was performed by several methods. Geochemical characterization included programmed pyrolysis (Weatherford Laboratories, Golden, CO) using a Source Rock Analyzer (SRA) which generates both a total organic carbon (TOC) value and Rock–Eval parameters, including thermally-distilled hydrocarbons at 300 °C (S1), and pyrolysis-generated hydrocarbon between 300 and 550 °C (S2). Units for TOC, S1 and S2 are given in milligrams per gram rock. The values of S1 and S2 are generally assumed to correlate with bitumen and kerogen content in a sample. In addition, the Hydrogen Index is calculated from the ratio of S2 to TOC and is given as milligrams of pyrolysis-generated hydrocarbon per gram TOC. Bitumen content was also estimated through extraction of organic matter (EOM) using chloroform. Total organic matter content, bitumen content plus kerogen content, was also determined through low temperature ashing (LTA) the samples in an oxygen plasma at 60 °C for 96 h with daily sample stirring. Geochemical parameters of the samples from the different characterization methods are given in Table 1.

A suite of simulated Gaussian, exponential and combined Gaussian–exponential decays were created to test the behavior of the inversion. The data were simulated with an echo spacing of 22 $\mu$s and 3000 echoes to match the experimental parameters of the measured data. All data were normalized to a signal intensity of one for ease of comparison between measurements. Gaussian white noise was added to the simulated data to obtain an SNR of 3500 ± 100, a value similar to the SNR of the measured data. There were three single Gaussian decays of 25, 79 and 158 $\mu$s simulated as well as a dual Gaussian system of equally weighted peaks at 25 and 79 $\mu$s. Three exponential decays of 7, 158, and 501 $\mu$s were simulated and a dual exponential system of equally weighted peaks at 158 and 501 $\mu$s. Again, four examples were simulated that included both Gaussian and exponential decays. The first case included a 25 $\mu$s Gaussian and a 501 $\mu$s exponential. The second case involved peaks that were closer to the center of the penalty function, with a 79 $\mu$s Gaussian and a 158 $\mu$s exponential. The third case was for a 100 $\mu$s Gaussian peak and 100 $\mu$s exponential peak. The final case included four peaks; two Gaussian peaks centered at 25 $\mu$s and 79 $\mu$s and two exponential peaks located at 158 $\mu$s and 501 $\mu$s.

The standard BRD inversion was performed in Prospa (Magritek, Wellington, New Zealand) using the Lexus inversion module. Simulated and measured data were logarithmically compressed before inversion for both methods to decrease computation time. The data were pruned to 400 points for the simultaneous Gaussian-Exponential (SGE) calculation. Because the SGE inversion has twice as many degrees of freedom due to fitting the relaxation data to distributions of Gaussian and exponential decays, the data for the BRD inversion was pruned to 800 points in order to maintain the same number of degrees of freedom between the two inversions. Spot checks between compressed and uncompressed results did not show significant differences, indicating no major loss of information during compression.

The center of the sigmoid for the SGE inversion was kept constant at 100 $\mu$s for the simulated data. This value was chosen based on previous relaxometry work on Gaussian-exponential decays [43]. However, the important factor for the simulated data is the location of the sigmoid center relative to the simulated decays, not its absolute value. The proper amount of smoothing was determined by the L-curve method [53], such that the smoothing parameter was selected to minimize the residuals, but not forced to an arbitrarily low value to prevent over-fitting. The location for the sigmoid center in the experimental oil shale data was optimized by minimizing the residuals for a given sample. This location varied between samples and the values are listed in Table 5. $\text{Sig}^{\text{weighted}}$ and $\text{Sig}^{\text{unweighted}}$ were kept constant at an order of magnitude larger than $\lambda$. The sigmoid weights should avoid being placed too high as this will force the solutions to be Gaussian below and exponential above the threshold, as well as resulting in larger residuals and unstable delta function like peaks. Future work may benefit from optimizing the sigmoid weight through the L-curve method like the smoothing parameter. If there is a known physical justification, the weight of one sigmoid can be adjusted to favor one type of relaxation over the other. As we had no reason to prefer one type of relaxation over another, the weights of the two sigmoid penalties were kept equal.

The value of $\text{Sig}^{\text{weighted}}$ of the sigmoid function was kept constant at 1 for all inversions. This is reasonable given that significant over-
Table 1
Geochemical properties of oil shale samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>LTA(^a)</th>
<th>EOM(^b)</th>
<th>SRA TOCc</th>
<th>S1</th>
<th>S2</th>
<th>Hydrogen Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alum shale</td>
<td>139</td>
<td>0.7</td>
<td>124</td>
<td>0.8</td>
<td>60.4</td>
<td>487</td>
</tr>
<tr>
<td>Kimmeridgian Blackstone</td>
<td>622</td>
<td>22.1</td>
<td>538</td>
<td>15.4</td>
<td>366.4</td>
<td>681</td>
</tr>
<tr>
<td>Condor oil shale</td>
<td>273</td>
<td>ND(^d)</td>
<td>203</td>
<td>1.8</td>
<td>108.5</td>
<td>534</td>
</tr>
<tr>
<td>Glen Davis torbanite</td>
<td>670</td>
<td>10.2</td>
<td>562</td>
<td>3.7</td>
<td>552.2</td>
<td>983</td>
</tr>
<tr>
<td>Green River Mahogany zone</td>
<td>305</td>
<td>30.7</td>
<td>234</td>
<td>12.5</td>
<td>277.2</td>
<td>1091</td>
</tr>
<tr>
<td>Julia Creek shale</td>
<td>203</td>
<td>ND(^d)</td>
<td>122</td>
<td>1.7</td>
<td>71.1</td>
<td>501</td>
</tr>
<tr>
<td>Narve-E Kukersite</td>
<td>629</td>
<td>1.1</td>
<td>474</td>
<td>3.4</td>
<td>498</td>
<td>1114</td>
</tr>
<tr>
<td>Phosphoria shale</td>
<td>283</td>
<td>11.3</td>
<td>197</td>
<td>7.2</td>
<td>132.6</td>
<td>673</td>
</tr>
<tr>
<td>Timahdit oil shale</td>
<td>136</td>
<td>2.8</td>
<td>95</td>
<td>2.5</td>
<td>59.6</td>
<td>627</td>
</tr>
</tbody>
</table>

All units are in mg/g.

\(^a\) Low temperature ashing.
\(^b\) Extractable organic matter.
\(^c\) Total organic carbon.
\(^d\) Not determined.

Table 2
Comparison of signal intensities determined for simulated data sets using the BRD and SGE inversion methods. Gaussian decays are indicated by a G, Exponential decay are indicated by an E, and the number indicates the characteristic time, e.g. G25 & E501 consists of a Gaussian decay of 25 \(\mu\)s and an exponential decay of 501 \(\mu\)s.

<table>
<thead>
<tr>
<th>Sample</th>
<th>BRD total</th>
<th>BRD &lt;100 (\mu)s</th>
<th>BRD &gt;100 (\mu)s</th>
<th>SGE total</th>
<th>Gaussian</th>
<th>Exponential</th>
</tr>
</thead>
<tbody>
<tr>
<td>G25 (\mu)s</td>
<td>3.58</td>
<td>3.58</td>
<td>-</td>
<td>1.08</td>
<td>1.08</td>
<td>-</td>
</tr>
<tr>
<td>G79 (\mu)s</td>
<td>1.47</td>
<td>1.47</td>
<td>-</td>
<td>1.00</td>
<td>0.98</td>
<td>0.02</td>
</tr>
<tr>
<td>G158 (\mu)s</td>
<td>1.33</td>
<td>0.42</td>
<td>0.91</td>
<td>1.06</td>
<td>0.77</td>
<td>0.29</td>
</tr>
<tr>
<td>G25 &amp; G79 (\mu)s</td>
<td>1.28</td>
<td>1.28</td>
<td>-</td>
<td>1.04</td>
<td>1.03</td>
<td>1.01</td>
</tr>
<tr>
<td>E79 (\mu)s</td>
<td>1.00</td>
<td>0.87</td>
<td>0.13</td>
<td>0.89</td>
<td>0.70</td>
<td>0.19</td>
</tr>
<tr>
<td>E158 (\mu)s</td>
<td>1.00</td>
<td>0.06</td>
<td>0.54</td>
<td>1.00</td>
<td>0.06</td>
<td>0.84</td>
</tr>
<tr>
<td>E501 (\mu)s</td>
<td>1.00</td>
<td>-</td>
<td>1.00</td>
<td>0.95</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>E158 &amp; E501 (\mu)s</td>
<td>1.00</td>
<td>0.05</td>
<td>0.95</td>
<td>1.00</td>
<td>0.04</td>
<td>0.96</td>
</tr>
<tr>
<td>G25 &amp; E501 (\mu)s</td>
<td>1.91</td>
<td>1.41</td>
<td>0.50</td>
<td>1.04</td>
<td>0.54</td>
<td>0.50</td>
</tr>
<tr>
<td>G79 &amp; E158 (\mu)s</td>
<td>1.20</td>
<td>0.87</td>
<td>0.33</td>
<td>1.00</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>G100 &amp; E100 (\mu)s</td>
<td>1.18</td>
<td>1.16</td>
<td>0.02</td>
<td>0.99</td>
<td>0.65</td>
<td>0.34</td>
</tr>
<tr>
<td>G25, G79, E158, E501 (\mu)s</td>
<td>1.27</td>
<td>0.88</td>
<td>0.39</td>
<td>1.02</td>
<td>0.54</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Table 3
Fit quality parameters determined for simulated the data sets by the BRD and SGE inversion methods. Gaussian decays are indicated by a G, Exponential decay are indicated by an E, and the number indicates the characteristic time, e.g. G25 & E501 consists of a Gaussian decay of 25 \(\mu\)s and an exponential decay of 501 \(\mu\)s. F-test values above 1 indicate the BRD method produces a better fit while values below 1 indicate the SGE method produces a better fit to the data. For a \(p\)-score of < 1e-6, the value has reached the level of the computer precision 2e-16.

<table>
<thead>
<tr>
<th>Sample</th>
<th>BRD residual ((\times 10^{-4}))</th>
<th>SGE residual ((\times 10^{-4}))</th>
<th>F-test value</th>
<th>(p)-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>G25 (\mu)s</td>
<td>1.06</td>
<td>0.49</td>
<td>0.47</td>
<td>&lt;|</td>
</tr>
<tr>
<td>G79 (\mu)s</td>
<td>217</td>
<td>2.80</td>
<td>0.01</td>
<td>&lt;|</td>
</tr>
<tr>
<td>G158 (\mu)s</td>
<td>683</td>
<td>68.9</td>
<td>0.10</td>
<td>&lt;|</td>
</tr>
<tr>
<td>G25 &amp; G79 (\mu)s</td>
<td>5.87</td>
<td>1.48</td>
<td>0.25</td>
<td>&lt;|</td>
</tr>
<tr>
<td>E79 (\mu)s</td>
<td>1.40</td>
<td>3.67</td>
<td>2.61</td>
<td>&lt;|</td>
</tr>
<tr>
<td>E158 (\mu)s</td>
<td>2.14</td>
<td>2.43</td>
<td>1.14</td>
<td>5.1 \times 10^{-4}</td>
</tr>
<tr>
<td>E501 (\mu)s</td>
<td>2.33</td>
<td>2.37</td>
<td>1.02</td>
<td>0.642</td>
</tr>
<tr>
<td>E158 &amp; E501 (\mu)s</td>
<td>2.07</td>
<td>2.39</td>
<td>1.16</td>
<td>4.8 \times 10^{-5}</td>
</tr>
<tr>
<td>G25 &amp; E501 (\mu)s</td>
<td>2.56</td>
<td>1.64</td>
<td>0.64</td>
<td>&lt;|</td>
</tr>
<tr>
<td>G79 &amp; E158 (\mu)s</td>
<td>33.5</td>
<td>2.54</td>
<td>0.08</td>
<td>&lt;|</td>
</tr>
<tr>
<td>G100 &amp; E100 (\mu)s</td>
<td>60.3</td>
<td>2.17</td>
<td>0.04</td>
<td>&lt;|</td>
</tr>
<tr>
<td>G25, G79, E158, E501 (\mu)s</td>
<td>1.95</td>
<td>1.90</td>
<td>0.98</td>
<td>0.537</td>
</tr>
</tbody>
</table>

Table 4
Fit quality parameters determined for the measured oil shale relaxation data by the BRD and SGE inversion methods. F-test values above 1 indicate the BRD method produces a better fit while values below 1 indicate the SGE method produces a better fit to the data. For a \(p\)-score of < 1e-6, the value has reached the level of the computer precision 2e-16.

<table>
<thead>
<tr>
<th>Sample</th>
<th>BRD residual ((\times 10^{-4}))</th>
<th>SGE residual ((\times 10^{-4}))</th>
<th>F-test value</th>
<th>(p)-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alum shale</td>
<td>6.57</td>
<td>4.51</td>
<td>0.68</td>
<td>&lt;|</td>
</tr>
<tr>
<td>Kimmeridgian Blackstone</td>
<td>4.49</td>
<td>3.30</td>
<td>0.74</td>
<td>&lt;|</td>
</tr>
<tr>
<td>Condor oil shale</td>
<td>17.1</td>
<td>6.76</td>
<td>0.39</td>
<td>&lt;|</td>
</tr>
<tr>
<td>Glen Davis torbanite</td>
<td>3.13</td>
<td>2.28</td>
<td>0.67</td>
<td>&lt;|</td>
</tr>
<tr>
<td>Green River Mahogany zone</td>
<td>10.2</td>
<td>6.02</td>
<td>0.59</td>
<td>&lt;|</td>
</tr>
<tr>
<td>Julia Creek shale</td>
<td>8.67</td>
<td>5.48</td>
<td>0.63</td>
<td>&lt;|</td>
</tr>
<tr>
<td>Narve-E Kukersite</td>
<td>1.16</td>
<td>0.78</td>
<td>0.68</td>
<td>&lt;|</td>
</tr>
<tr>
<td>Phosphoria shale</td>
<td>6.14</td>
<td>3.14</td>
<td>0.44</td>
<td>&lt;|</td>
</tr>
<tr>
<td>Timahdit oil shale</td>
<td>5.48</td>
<td>4.86</td>
<td>0.51</td>
<td>&lt;|</td>
</tr>
</tbody>
</table>
lap was not expected in the samples as they contain no appreciable quantities of liquid. Provided the weight of the sigmoid functions is not extremely high, increasing the width will allow for more overlap in the Gaussian and exponential peaks. If the weight of the sigmoids are very high, this will instead lead to high residuals and a dead zone where it is numerically unfavorable to have peaks.

Optimization of the sigmoid width may lead to further reduction of residuals. The standard tolerance of $1 \times 10^{-6}$ was used as the criterion for convergence of the algorithm. The initial input for the distribution of Gaussian and exponential peaks to the function was a vector containing only zeroes. The range of $T_2$ times spanned from $1$ to $10^5 \mu$s logarithmically spaced.

### 3. Results

Quality of the inversions was evaluated in several ways. First, the traditional method of examining the sum of the squares of the difference between the measured and fitted results was determined to obtain the total residual used for both the BRD and the SGE methods. A second, more thorough method, the $F$-test, was also applied to evaluate the quality of the fits. The $F$-test is a method to compare the variance of two populations and is commonly used to evaluate how well a proposed regression method fits to the data. In its application presented in this manuscript, the residuals for all the points fit by the standard BRD model are in the denominator, and the residuals from the SGE model are in the numerator. Therefore, an $F$-test value below $1$ indicates that the SGE model provides a better fit to the data. The smaller the value, the better the new model fits to the data compared to the standard model.

Conversely, an $F$-test value above $1$ indicates the standard BRD model is a better fit to the data. The $p$-score indicates confidence that the difference between the fits is a real effect and not just a statistical aberration. This approach compares the overall quality of fit for all the points, and as such is a more sensitive method for evaluating the quality of the fit and the certainty in the $p$-value becomes more reliable. In addition, because the total signal intensity is known for simulated results, the inverted intensity was compared to the true signal intensity to determine if the inversion methods over- or underestimated the total signal. For the oil shale samples, the true NMR intensity of hydrogen in the samples is not known, however total signal intensity generated by the BRD and SGE inversions were compared to assess any systematic differences between them. Signal intensities for the simulated data from both the BRD and SGE inversions are presented in Table 2.

#### 3.1. Simulated data

Results of the SGE and BRD inversions for the simulated Gaussian signals are shown in Fig. 1, panels A1, B1, C1, and D1 and A2, B2, C2, and D2, respectively. The dashed lines indicate the proper location of the simulated relaxation times. Inversion quality evaluation is presented in Table 3. As expected, the SGE inversion manages to recover the Gaussian signal accurately for the two cases where the peaks are below the sigmoid. The peak locations are accurate and the inverted signal intensity is very close to the actual intensity. A small signal spike is seen in the exponential region for the $79 \mu$s peak. In contrast, BRD inversion greatly overcalls the signal for the $25 \mu$s Gaussian and the peak is located at an unphysical $12 \mu$s given the $22 \mu$s echo spacing used. The peak location for the $79 \mu$s Gaussian peak is still too low at $56 \mu$s for the BRD inversion, but is at least a physically realistic value given the echo spacing used. The BRD signal intensity for the $79 \mu$s peak is overcalled, but less so than the $25 \mu$s. Residuals were much lower for the SGE inversion for both cases. In particular, in the case of the $79 \mu$s Gaussian peak, the residuals are substantially lower for the SGE inversion. This is because as the relaxation time of the Gaussian becomes significantly longer than the echo spacing, the Gaussian character of the relaxation curve becomes more apparent and the BRD inverse Laplace transform is unable to fit as well to the curve as compared to Gaussian relaxation times closer to the echo spacing of the measurement. This indicates a combined SGE inversion will be particularly useful as NMR equipment with increasingly shorter echo spacing becomes available.

For the Gaussian peak past the $100 \mu$s center of the sigmoid penalty function, the SGE inversion manages to obtain the correct Gaussian relaxation time, however some signal was allocated to exponential decay as well. The resulting inversion has a much higher residual than the other simulations, for both the BRD and SGE inversions when compared to the inverted peaks below the sigmoid penalty function. For the dual Gaussian case, the BRD inversion produces a very weak peak at $7 \mu$s and a strong peak at $44 \mu$s. The BRD signal intensity is substantially overestimated. The SGE manages to recover both peaks with accurate peak locations and the peak weights are also close to the correct values, although there is slight bias toward the shorter peak ($0.548$) compared to the longer peak ($0.480$). There is also a weak artifact in the exponential component at around $178 \mu$s.

Results for the exponential decay simulations are shown in Fig. 2. For the purely exponential decays, the BRD and SGE inversions produced similar results when the exponentials were longer than the central sigmoid time. Peak locations and intensities were accurate. The residuals were slightly higher for the SGE inversion, though not substantially so, indicating a similar quality of fit between the two types of inversions. Further refinement of the SGE inversion method is expected to achieve parity between the two methods. However, for the case where the exponential peak is located below the sigmoid penalty function center ($79 \mu$s), the standard BRD inversion is clearly superior. The resulting SGE inversion produces a very poor fit (Fig. 2, panel A1), with highly inaccurate peak locations and larger residuals. As such, comparison of the two types of results indicates that when the cutoff time is uncertain, it may be best to err on the side of a lower sigmoid center value. If the sigmoid penalty function center is changed to a value

### Table 5

Comparison of signal intensities determined for oil shale samples using the BRD and SGE inversion methods.

<table>
<thead>
<tr>
<th>Sample</th>
<th>BRD total</th>
<th>BRD &lt;SGE sigmoid</th>
<th>BRD &gt;SGE sigmoid</th>
<th>SGE sigmoid (μs)</th>
<th>SGE total</th>
<th>Gaussian intensity</th>
<th>Exponential intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alum shale</td>
<td>7.73</td>
<td>5.96</td>
<td>1.77</td>
<td>71</td>
<td>4.44</td>
<td>2.49</td>
<td>1.95</td>
</tr>
<tr>
<td>Kimmeridgian Blackstone</td>
<td>16.59</td>
<td>14.99</td>
<td>1.60</td>
<td>200</td>
<td>7.87</td>
<td>5.94</td>
<td>1.93</td>
</tr>
<tr>
<td>Condor oil shale</td>
<td>17.57</td>
<td>12.69</td>
<td>4.88</td>
<td>71</td>
<td>10.79</td>
<td>5.73</td>
<td>5.06</td>
</tr>
<tr>
<td>Glen Davis torbanite</td>
<td>24.94</td>
<td>21.58</td>
<td>3.36</td>
<td>158</td>
<td>12.70</td>
<td>9.68</td>
<td>3.02</td>
</tr>
<tr>
<td>Green River Mahogany</td>
<td>16.22</td>
<td>13.42</td>
<td>2.80</td>
<td>200</td>
<td>9.04</td>
<td>6.16</td>
<td>2.88</td>
</tr>
<tr>
<td>Julia Creek shale</td>
<td>7.98</td>
<td>6.21</td>
<td>1.77</td>
<td>71</td>
<td>4.82</td>
<td>2.89</td>
<td>1.94</td>
</tr>
<tr>
<td>Narve-E Kukersite</td>
<td>29.25</td>
<td>25.21</td>
<td>4.04</td>
<td>71</td>
<td>13.41</td>
<td>9.16</td>
<td>4.25</td>
</tr>
<tr>
<td>Phosphoria shale</td>
<td>16.42</td>
<td>12.67</td>
<td>3.75</td>
<td>200</td>
<td>9.35</td>
<td>5.78</td>
<td>3.56</td>
</tr>
<tr>
<td>Timahdit oil shale</td>
<td>8.45</td>
<td>6.62</td>
<td>1.83</td>
<td>200</td>
<td>5.14</td>
<td>3.26</td>
<td>1.88</td>
</tr>
</tbody>
</table>
Fig. 1. Purely Gaussian decays inverted by the SGE (panel 1) and the BRD (panel 2) methods: (A) 25 μs Gaussian decay; (B) 79 μs Gaussian decay; (C) 158 μs Gaussian decay; (D) 25 μs and 79 μs Gaussian decays. In the SGE inversions, the Gaussian component is shown as black lines and the exponential component in grey. The dashed line indicates the proper location of the simulated relaxation times.

Fig. 2. Purely exponential decays inverted by the SGE (panel 1) and the BRD (panel 2) methods: (A) 79 μs exponential decay; (B) 158 μs exponential decay; (C) 501 μs exponential decay; (D) 158 μs and 501 μs exponential decays. In the SGE inversions, the Gaussian component is shown as black lines and the exponential component in grey. The dashed line indicates the proper location of the simulated relaxation times.
Fig. 3. Combined Gaussian and exponential decays inverted by the SGE (panel 1) and the BRD (panel 2) methods: (A) 25 µs Gaussian and 501 µs exponential; (B) 79 µs Gaussian and 158 µs exponential; (C) 100 µs Gaussian and 100 µs exponential; (D) 25 µs Gaussian, 79 µs Gaussian, 158 µs exponential, and 501 µs exponential. In the SGE inversions, the Gaussian component is shown as black lines and the exponential component in grey. The dashed line indicates the proper location of the simulated relaxation times.

Fig. 4. Oil Shale $T_2$ data inverted by the SGE (panel 1) and the BRD (panel 2) methods: (A) Alum; (B) Blackstone; (C) Condor; (D) Glen Davis; (E) Green River Mahogany; (F) Julia Creek; (G) Kukersite; (H) Phosphoria; (I) Timahdit shales. In the SGE inversions, the Gaussian component is shown as black lines and the exponential component in grey.
below the exponential relaxation time, the SGE inversion is able to
again accurately recover the peak (results not shown). This sug-
gests that minimization of the residuals is a valid option for select-
ing the correct center for the sigmoid penalty function if it is not
known a priori. Inversion of the dual-exponential simulation
showed similar results for both methods, although the BRD had
slightly better residuals.

Finally, simulated results containing both Gaussian and expo-
nential components were analyzed and the results are shown in
Fig. 3. The SGE inversion is quite clearly able to distinguish the
Gaussian and exponential peaks for the first case of the 25 \(\mu s\)
Gaussian and a 501 \(\mu s\) exponential peaks. Peak intensities and
locations are accurate. In contrast, the BRD inversion significantly
overcalls the intensity of the Gaussian peak and the relaxation time
indicated is much too short at 13 \(\mu s\). The position of the exponen-
tial peak is accurate for the BRD inversion. As the two peaks come
closer with the 79 \(\mu s\) Gaussian and 158 \(\mu s\) exponential peaks, the
SGE inversion is still able to distinguish both peaks with accurate
signal intensity, while the BRD inversion reproduces one peak with
a shoulder and an overcalled intensity. The residuals of the BRD
inversion are significantly higher. When the Gaussian and exponen-
tial peaks overlap, the SGE inversion is able to resolve both with accu-
rate locations, but again there appears to be bias in signal intensity
toward the Gaussian peak (0.64) over the exponential peak (0.35).
The BRD inversion only produced a single peak located at a shorter
relaxation time (79 \(\mu s\)) and a slight signal intensity overcall. For

the four component simulation, the SGE inversion is able to accu-
rately recover the peak positions and the signal intensities are very
good. The BRD inversion greatly overcalls the signal intensity of the
79 \(\mu s\) Gaussian peak and both Gaussian peaks are located at
shorter relaxation times (10 and 50 \(\mu s\)) than the simulated data
indicated. Surprisingly, the residuals for the two inversions are
remarkably similar.

3.2. Measured data

Measured results for the oil shale samples were inverted using
both the BRD and SGE methods and the \(T_2\) distributions are shown
in Fig. 4. This comparison of the BRD and SGE inversion methods
shows that the most intense, short relaxation time peaks are
located at physically realistic locations using the SGE method,
while the BRD is located at values below the CPMG echo spacing
for the BRD method. Fit quality evaluation of the two inversion
methods for the oils shale samples are shown in Table 4. The SGE
and BRD total intensities were well correlated \((R^2 = 0.928)\) with a
slope of 1.907 (intercept forced to zero), reflecting the substantially
higher signal intensities generated by the BRD method for the
shales. Other intensity correlations relating the Gaussian compo-
nent of the SGE inversion results to BRD intensities below the sig-
moid value determined for the SGE model (see Table 5) and the
exponential component of the SGE to the sum of BRD values above
the sigmoid were also determined. Both were well correlated

![Comparison between geochemistry measurements and NMR results](image)

**Fig. 5.** Comparison between geochemistry measurements and NMR results (A) Total organic matter from LTA to the total integrated SGE signal (solid circles) and the total integrated BRD signal (open square) (B) Total hydrocarbon intensity \((S1 + S2)\) to the total integrated SGE signal (solid circles) and the total integrated BRD signal (open square) (C) Kerogen content measured via total organic matter from LTA minus the quantity of extractable organic matter to the integrated Gaussian SGE signal (solid circles) and the integrated BRD signal below the sigmoid center (open square) (D) Kerogen content measured via SRA to the integrated Gaussian SGE signal (solid circles) and the integrated BRD signal below the sigmoid center (open square).
The Gaussian component correlation with the BRD < sigmoid signal had a slope similar to that of the total signal correlation (slope = 2.372, intercept forced to zero), while the slope relating the exponential component and BRD > sigmoid signal was very close to unity (slope = 0.979, intercept forced to zero). It is not possible to know the true signal intensity for these samples, however the consistently higher total intensities generated by the BRD method in the simulated data and these intensity correlations for the two inversion methods indicate that the SGE results may represent more realistic values and, at the very least, are unlikely to overcall the signal intensity at short times where much of the signal from the kerogen is expected.

Fig. 5 shows the relationships found between the NMR and geochemistry results. Total signal intensities for both the SGE and BRD inversions tended to increase with increasing total organic matter content as measured by low temperature ashing. Similarly, the total signal intensities for both methods also increased with the combined kerogen and bitumen content as measured by SRA. Based on the geochemical parameters, it is clear that solid organic matter represents the majority of the signal observed in the T₂ distributions. SGE Gaussian component intensities and BRD intensities below the SGE sigmoid increased with kerogen content as measured by both methods. However, previous analysis using T₁–T₂ correlation methods under similar conditions [54] indicates that signal generated below ~100 μs includes contributions from multiple phases or domains within the organic matter, therefore this interpretation from the T₂ distribution results alone should be treated with caution. No relationship was observed between bitumen content from either measurement method and the exponential component of the SGE or the total BRD intensities above the SGE sigmoid. This is not unexpected, as previous work at high field shows that bitumen may relax with both Gaussian and exponential contributions [34]. This may also be the result of strong interactions between bitumen and the kerogen and mineral phases present in the samples. Based on similar results at low field [54], the relaxation distributions are less complex and it may be feasible that the SGE inversion could potentially be used to discriminate kerogen through the Gaussian signal and the bitumen through the exponential signal from the T₂ relaxation alone at the NMR frequencies commonly used for lab and logging measurements. However, the SGE inversion combined with more extensive NMR measurements, such as T₁–T₂ correlations or binomial editing [55], may still be required at low field in order to accurately separate the two types of organic matter.

4. Conclusions

A simultaneous Gaussian-Exponential inversion for improved analysis of NMR transverse relaxation measurements on shales has been developed and preliminary testing of the method has been completed using simulated data and measurements on geologic samples. The combination of Gaussian and exponential decays leads to more realistic results for organic hydrogen-rich solids present in shales when compared to the widely used BRD inversion approach. Average residuals were significantly lower for most results analyzed by the SGE inversion and the resulting relaxation times consistently appeared at more physically realistic values given the measurement acquisition parameters. Results on simulated data indicate that signal overcall occurs in the BRD inversion and this artifact is mitigated by the SGE approach. Both the BRD and SGE inversion methods generated total signal intensities that were well correlated with the concentration of hydrogen-rich phases in the shale samples, but the physically unrealistic peak positions generated by the BRD inversion at short relaxation times make the consistently higher intensities this method produces suspect. The SGE inversion method has applications other than petroleum and geological samples and may be useful to analyze data from other fields where combined Gaussian and exponential relaxation decays occur, such as food, material and soil sciences. As the combined inversion has more parameters that can be adjusted in the inversion process, caution should be exercised to ensure that inverted results are consistent with the physics of the system.

Acknowledgments

The authors thank Michael Lewan for sample access, William Betterton for assistance with low temperature ashing, and Kirk Feindel for helpful discussion. JDS acknowledges the U.S. NSF MRI Program and the M.J. Murdock Charitable Trust for equipment funding. Any use of trade, product or firm names is for descriptive purposes only and does not imply endorsement by the US government.

Appendix A

The function that computes the difference between the measured and reconstructed data plus the sigmoid penalty function is given by:

$$\text{SGE} \leftarrow \text{function}(X, M, \text{to, T2, smoothing, sig, sigwidth, sigweightA, sigweightB})$$

```
{ halfX ← length(X)/2
  A ← X[1:halfX]
  B ← X[(halfX + 1):length(X)]
  sigcent ← c(-sig:((sig + (halfX - 1))
  sigup ← ((1/(1 + exp(-sigcent * sigwidth)))) * sigweightA + smoothing)
  sigdown ← (smoothing + (1 - (1/(1 + exp(-sigcent * sigwidth)))))) * sigweightB
  Mhat ← rep(0, length(M))
  for (ii in 1:halfX)
  { Mhat ← Mhat + A[ii] * exp(-(to/T2[ii])^2)) + B[ii] * exp(-to/T2[ii]) }
  SS ← sum((M - Mhat)^2) + sum(A * sigup) + sum(B * sigdown)
  return(SS) }
```
An example of the command for minimization of the SGE function is given by:

\[
\text{SGE.fit -- optim}(\text{rep}(0, \text{length}(T2 \text{ Values}) = 2), \ fn = \text{SGE.M} = \text{Measured}. \text{ Data}, \ to = \text{Time. Axis}, \ T2 = T2 \text{Values}, \ \text{smoothing} = 0.0001, \ sig = 38, \ \text{sigwidth} = 1.0, \ \text{sigweightA} = 0.001, \ \text{sigweightB} = 0.001, \ \text{lower} = \text{Min}, \ \text{upper} = \text{Max}, \ \text{method} = "\text{L-BFGS-B}")
\]

X is a vector from the combined A and B vectors. The initial input used here for X is all zeroes. \( fn \) is the inputted function into the minimization, here the SGE kernel. \( M \) is the input for the measured data, \( to \) is the input for the time axis associated with the measured data. \( T2 \) is the inputted range of \( T2 \) values for the inversion to solve over. It is recommended to input the \( T2 \) values as a logarithmic distribution. Smoothing is the smoothing constant \( \alpha \), \( \sigma \) is the center of the sigmoid as an index value of a \( T2 \) time from \( T2 \), \( \text{Sigwidth} \) is the width of the sigmoid applied logarithmically. \( \text{SigweightA} \) and \( \text{SigweightB} \) are the associated penalty functions for the Gaussian and exponential components. \( \text{Lower} \) is the minimum expected intensity for a given \( T2 \) time, usually zero. This should not be set to negative values to avoid the possibility of negative peaks. \( \text{Upper} \) is the maximum expected value for peaks in the \( T2 \) distribution. This should be chosen to be well above any expected intensity to avoid inadvertent peak clipping. \text{Method} relates to the different options of the optimization and full explanation can be found in the \text{R} optim help file [56].

References