Modelling and Estimation of Multicomponent $T_2$ Distributions

Kelvin J. Layton*, Mark Morelande, David Wright, Peter M. Farrell, Bill Moran, Member, IEEE, and Leigh A. Johnston, Member, IEEE

Abstract—Estimation of multiple $T_2$ components within single imaging voxels typically proceeds in one of two ways; a nonparametric grid approximation to a continuous distribution is made and a regularized nonnegative least squares algorithm is employed to perform the parameter estimation, or a parametric multicomponent model is assumed with a maximum likelihood estimator for the component estimation. In this work, we present a Bayesian algorithm based on the principle of progressive correction for the latter choice of a discrete multicomponent model. We demonstrate in application to simulated data and two experimental datasets that our Bayesian approach provides robust and accurate estimates of both the $T_2$ model parameters and nonideal flip angles. The second contribution of the paper is to present a Cramér-Rao analysis of $T_2$ component width estimators. To this end, we introduce a parsimonious parametric and continuous model based on a mixture of inverse-gamma distributions. This analysis supports the notion that $T_2$ spread is difficult, if not infeasible, to estimate from relaxometry data acquired with a typical clinical paradigm. These results justify the use of the discrete distribution model.

Index Terms—Extended phase graph, multicomponent relaxometry, nonnegative least squares (NNLS), progressive correction, stimulated echo, $T_2$ relaxation.

I. INTRODUCTION

In an ideal environment, spin-spin relaxation can be modelled by a decaying exponential with a characteristic time constant, $T_2$. Accurate estimation of $T_2$ values is important for tissue classification, disease detection and pathology [1]. Although a single exponential decay is a sufficient description for ideal homogenous tissue, in reality spins within a single voxel interact with different molecular environments, leading to a signal decay that arises from integration across a multicomponent distribution of $T_2$ values [2]. Methods to estimate multicomponent $T_2$ distributions can be categorized in two basic classes, nonparametric and parametric, according to the choice of distribution model.

Nonparametric methods model the $T_2$ distribution within a voxel as a pseudo-continuous grid of weights. The method of Whittall and MacKay proposed in [3] approximates the distribution with a large number of delta functions at fixed grid locations and solves the resulting optimization using the nonnegative least squares (NNLS) algorithm. This technique has been used extensively for $T_2$ distribution estimation, e.g., [4]–[7].

The second category of $T_2$ estimation methods, the parametric class, assumes that the $T_2$ distribution can be well described by a small number of components, each characterized by a few parameters. Continuous parametric models include a log-Gaussian mixture used in [8] and [9] and an inverse-Gaussian mixture presented in [10]. These continuous models have been useful for the estimation of nonlocalized $T_2$ distributions, where the signal-to-noise ratio is very high. A popular parametric model for per-voxel estimation is a discrete distribution consisting of three pools proposed and tested in [11] and [12], respectively. This discrete model has also been used to analyze $T_2^*$ decay [13], [14]. In these methods, estimation of the discrete components is typically performed using a gradient based optimization algorithm, equivalent to finding the maximum likelihood estimate (MLE) for additive Gaussian noise. The simulations reported in [12] demonstrated unsurprising robustness, as estimates were initialized at the true values. In general, the discrete multicomponent model has proved unreliable when initialized away from ground truth [15].

Relaxation data for $T_2$ estimation is traditionally acquired using a multi-echo spin echo sequence such as CPMG [16]. The sequence relies on a train of 180° radio-frequency (RF) pulses to refocus the transverse magnetization and produce an echo that is weighted by the $T_2$ component. In practice, the refocusing flip angle can deviate considerably from 180° due to $H_1$ field inhomogeneity, off-resonance effects or nonuniform slice profiles, which leads to the generation of secondary and stimulated echoes. These additional coherence pathways can result in biased estimation of the underlying $T_2$ values [17]. The echo amplitudes resulting from nonideal refocusing pulses can be derived from the extended phase graph (EPG) algorithm [18], [19]. The EPG algorithm tracks the multiple coherence pathways of...
spins after consecutive periods that model precession, relaxation and refocusing. In this way, the problematic stimulated echoes can be accounted for using the improved signal model of the EPG algorithm. Stimulated echo correction for nonideal refocusing flip angles was applied in [20] for estimation of a single discrete $T_2$ component, using a gradient based optimization algorithm to jointly estimate the flip angle and $T_2$ value. In the domain of nonparametric models, the NNLS algorithm was recently extended to compensate for stimulated echoes [21].

The current paper makes two main contributions to the modelling and estimation of multicomponent $T_2$ distributions. We focus on a discrete multicomponent model as in [12], and extend this model to include stimulated echo compensation using the EPG algorithm, analogous to [20] for the single component case. While the theoretical formulation is straightforward, the implementation of an algorithm that robustly and accurately estimates both the weights and locations of multiple discrete $T_2$ components and the parameters of the $B_1$ field is nontrivial. To this end, we present a Bayesian algorithm that overcomes problems of maximum likelihood-based methods associated with low SNR and algorithm initialization. This builds on our previous work in the area for the case of ideal 180° refocusing pulses [22].

The second contribution of this paper is a Cramér-Rao analysis of the $T_2$ distribution models, confirming the idea that, while mode locations of $T_2$ distributions can be estimated with some accuracy, it is not feasible to reliably estimate the width of the modes. This provides a theoretical justification for the use of the discrete distribution model. In order to perform this analysis, we introduce a novel multicomponent $T_2$ distribution model that is both parametric and continuous.

The paper proceeds as follows. In Section II, we present the general signal equation for $T_2$ distribution estimation and review the NNLS pseudo-continuous model along with its estimation algorithm, and the discrete multicomponent model. In Section III, we propose a Bayesian estimation algorithm for discrete multicomponent models that improves on previous MLE-based approaches. An analysis of the ability of algorithms to estimate $T_2$ distribution width is presented in Section IV, employing a novel parametric continuous model for this purpose. We calculate Cramér-Rao lower bounds on the estimation performance and detail the limited information available in the measurement. The remainder of the paper follows with Methods (Section V) and Results (Section VI) that demonstrate application of the Bayesian algorithm to simulated and experimental data sets.

II. BACKGROUND AND SIGNAL MODEL

The distribution of relaxation times is observed through the amplitudes of the acquired echo signals. The measured MRI signal at the $n$th echo is described by the integration of the decaying signals from each contribution

$$ s_n = \int f(\tau) g_n(\tau) e^{-\tau/\tau} \, d\tau $$

where $f$ is the unknown distribution of relaxation times, $g_n$ is a function describing the signal amplitude for a given relaxation rate, and $\phi$ is the signal phase. Notice the complex nature of the signal, $s_n \in \mathbb{C}$. The measurements are noisy samples of the signal described by (1)

$$ y_n = s_n + v_n, \quad n = 1, \ldots, N $$

where $v_n$ is complex Gaussian additive noise and $N$ is the number of echoes.

All existing $T_2$ distribution models, and subsequent estimation algorithms, can be derived from the general signal model form in (1) and (2), according to the choice of $f$, $g_n$, and $v_n$. For example, a single $T_2$ component model is described by

$$ g_n(\tau) = e^{-\tau/\tau} $$

where $\tau_n$ is the echo time and $\tau$ is the $T_2$ time constant.

The effect of $B_1$ inhomogeneity is taken into account in the signal model by modification of the function, $g_n$, to include additional parameters for spin-lattice relaxation, $T_1$, and the variable flip angle, $\alpha$. The distribution of different phase states at the $n$th echo are stored in a vector, $x_n$. The EPG algorithm describes the evolution of this vector via a linear recursive relation, from which the echo amplitudes can be extracted. The algorithm, described in detail in Appendix A, can be summarized as

$$ x_n = E(\tau, T_1, \alpha) x_{n-1} $$

where the vectors $c$ and $x_n$, and the matrix $E$ are defined in Appendix A and the superscript $T$ denotes matrix transpose. Nonuniform slice profiles, such as those in [20], can be considered under the same formalism by definition of appropriate function $g_n$.

All parametric and nonparametric distribution models are specified by choice of the form of $g_n$ together with a model for the distribution, $f$. In Sections II-A and II-B, we detail these choices firstly for the NNLS pseudo-continuous model, and secondly for the multicomponent discrete model. Please note that while “NNLS” refers specifically to an estimation algorithm, we henceforth follow the convention in the literature, and use NNLS to refer to both the pseudo-continuous grid model and the estimation algorithm.

A. NNLS Pseudo-Continuous Model and Estimation

The NNLS approach to $T_2$ distribution estimation models a continuous distribution by values on a grid over the parameter space, defining a large but known sequence of relaxation times, $\tau_1, \ldots, \tau_L$, that cover a physically plausible range [3]. This has been applied extensively in the literature for the case of ideal flip angles by defining $g_n$ as in (3), e.g., [4]–[7]. The distribution is composed of $L$ delta functions

$$ f(\tau) = \sum_{i=1}^L w_i \delta(\tau - \tau_i). $$

This pseudo-continuous model is nonparametric, as minimal assumptions are made about the distribution shape. The extension proposed in [21] to account for stimulated echoes results in $g_n$ defined by the EPG algorithm in (4).
Rician noise can be approximated as Gaussian for sufficiently high SNR [23].

The NNLS signal model is obtained by combining (1), (2), (4), and (5) to give

$$y_n = \sum_{i=1}^{L} w_i g_n(\tau_i, \hat{T}_i, \hat{\alpha}) + v_n.$$  
(6)

Given values of $\hat{T}_i$ and $\hat{\alpha}$, the EPG algorithm gives the expected signal contribution at a grid point $\tau_i$. The problem of estimating the distribution is reduced to estimating the weights.

As the number of measurements is often considerably smaller than the number of unknown weights ($N < L$), the system is underdetermined and will exhibit poor noise performance. The solution proposed in [3] is to regularize the least-squares problem. This amounts to solving the following optimization:

$$\text{minimize } \|y - Aw\|^2 + \lambda \|Cw\|^2$$
$$\text{subject to } w \geq 0$$  
(7)

where $y = [y_1, \ldots, y_N]^T$ is a vector of measurements, $w = [w_1, \ldots, w_L]^T$ is a vector of unknown weights, $A$ is an $N \times L$ matrix with elements $A_{n,i} = g_n(\tau_i, \hat{T}_i, \hat{\alpha})$ and $C$ contains additional constraints (such as smoothness), weighted by the regularization parameter, $\lambda$. This optimization is solved efficiently using the nonnegative least squares algorithm.

It is not practical to estimate $T_1$ directly from the echo data because the signal is only weakly dependent on the $T_1$ value. For example, a typical two component decay curve changes by less than 1% for $T_1$ values above 1 s. Under the assumption that $T_1 \gg T_2$, one can confidently fix $\hat{T}_1 = 1$ s (as in [21]) or $\hat{T}_1 = \infty$ (as in [20]) with minimal effect on the estimation results. The estimation accuracy is expected to degrade for tissues with short $T_1$, for example, when certain contrast agents are used.

Nonideal flip angles, $\hat{\alpha}$, can be estimated in a separate optimization stage which minimizes the sum of squared errors between the measured and predicted signals [21].

### B. Discrete Multicomponent Model

The assumption that the $T_2$ distribution is made up of a small number of discrete components [11] is motivated by the nature of biological tissue, which in the brain for example can consist of myelinated axon tracts, inter/extra-cellular water and cerebrospinal fluid (CSF), each of which has a different relaxation time. The discrete multicomponent distribution is mathematically specified by a small number of weighted delta functions with unknown weights and locations

$$f(\tau) = \sum_{i=1}^{M} w_i \delta(\tau - \tau_i).$$  
(8)

Two signal equations can be derived depending on whether or not the assumed decay function $g_n$ includes stimulated echoes. The case of ideal 180° pulses yields a signal model consisting of a sum of exponential decay functions for each component [12]–[14].

To include stimulated echo compensation, the EPG algorithm is employed. The signal can still be described by a weighted sum of the decay functions but in this case the flip angle must be estimated along with the weights, locations and phase

$$y_n = e^{i\phi} \sum_{i=1}^{M} w_i g_n(\tau_i, \hat{T}_i, \alpha) + v_n.$$  
(9)

Note that (9) is fundamentally different from the nonparametric NNLS form in (6), since the times $\tau_i$ are unknown and must be estimated along with the weights. This is only viable when the distribution has a small number of modes, i.e., $M \ll N$. Equation (9) is a natural extension of the single component case considered in [20].

The discrete multicomponent model (9) does not lend itself well to a simple estimation algorithm despite the simplicity of the model. The nonlinear relationship between the parameters and the signal creates a poorly behaved MLE cost function, with local minima and large regions in parameter space where the cost function is essentially flat. These features are problematic for a naive gradient-based optimization algorithm.

### III. PROPOSED ESTIMATION ALGORITHM FOR DISCRETE MULTICOMPONENT MODEL

As an alternative to gradient-based MLE algorithms, we adopt a Bayesian framework that provides an intuitive way to incorporate prior information [24]. This is particularly useful to describe the physically plausible parameters of biological tissue. We will see that the formulation also leads to the development of a more reliable estimation algorithm. The posterior density given by Bayes’ rule is

$$\pi(\theta | y) = \frac{\ell(\theta | y) \pi_0(\theta)}{\eta}$$  
(10)

where $y$ is a vector of measurements, $\theta$ is a vector of unknown parameters, $\ell(\cdot | \cdot)$ is the likelihood, $\pi_0(\cdot)$ is the prior distribution and $\eta$ is a normalizing constant. The discrete multicomponent model has two parameters per component in addition to a phase and flip angle. To avoid the $\tau = 0$ signal discontinuity, our implementation works with estimation rates, $r = 1/\tau$. The vector of unknown parameters is therefore

$$\theta = [w_1, r_1, \ldots, w_M, r_M, \phi, \alpha]^T.$$  
(11)

The posterior, $\pi$, in (10), cannot be calculated in closed-form due to the nonlinear signal equation and therefore approximations must be sought. There are two main challenges to accurately approximate the posterior. Firstly, approximations degrade for a relatively wide prior, $\pi_0$, and narrow likelihood, $\ell$, representing the realistic situation where one relies on the measurements more than prior knowledge. Secondly, the approximation should be chosen such that Bayes’ rule in (10) can be efficiently computed.

We overcome the first challenge using a technique known as progressive correction [25]. Although we do not use a Monte Carlo approximation, we apply the same principle of ‘flattening’
the likelihood and iteratively correct our estimate of the posterior. We define a schedule of \( P \) corrections, \( \gamma_1, \ldots, \gamma_P \), with the intermediate posterior at the \( j \)th correction step given by

\[
\pi_j(\theta | y) = \frac{L_j(\theta) \pi_{j-1}(\theta | y)}{\eta_j}
\]  

(12)

where \( \eta_j \) is a normalizing constant. The principle of progressive correction provides that when \( \sum_j \gamma_j = 1 \), the final posterior \( \pi_P \) is the posterior specified in (10).

To overcome the second challenge to make the computation of (12) tractable, we use a linearized approximation of the likelihood and a Gaussian prior with mean, \( \mu_0 \), and covariance \( \Psi_0 \). This leads to a closed-form expression for the posterior using the well-known Kalman filter equations. In this case, we do not compute the exact posterior, but rather a Gaussian approximation to it. The approximate likelihood is obtained by linearizing the nonlinear function that describes the EPG signal. Importantly, we use the raw complex-valued data to jointly estimate the phase and relaxation times. This has the advantage of the noise remaining Gaussian instead of Rician as is the case when magnitude data is used [23].

Recall the signal in (9) is

\[
s_n(\theta) = e^{j\phi} \sum_{i=1}^{M} w_i g_n(\tau_i, \hat{T}_1, \alpha).
\]

(13)

The Jacobian of (13) is

\[
J(\theta) = \begin{bmatrix}
\frac{\partial s_1}{\partial \alpha_1} & \frac{\partial s_1}{\partial \tau_1} & \cdots & \frac{\partial s_1}{\partial \alpha_M} & \frac{\partial s_1}{\partial \tau_M} & \frac{\partial s_1}{\partial \alpha} & \frac{\partial s_1}{\partial \tau} \\
\frac{\partial s_2}{\partial \alpha_1} & \frac{\partial s_2}{\partial \tau_1} & \cdots & \frac{\partial s_2}{\partial \alpha_M} & \frac{\partial s_2}{\partial \tau_M} & \frac{\partial s_2}{\partial \alpha} & \frac{\partial s_2}{\partial \tau} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
\frac{\partial s_N}{\partial \alpha_1} & \frac{\partial s_N}{\partial \tau_1} & \cdots & \frac{\partial s_N}{\partial \alpha_M} & \frac{\partial s_N}{\partial \tau_M} & \frac{\partial s_N}{\partial \alpha} & \frac{\partial s_N}{\partial \tau}
\end{bmatrix}
\]

(14)

The partial derivatives can be calculated efficiently using a set of recursive relationships we detail in Appendix B.

The approximate likelihood is

\[
\hat{L}_j(\theta) = \mathcal{N}
\left(
\frac{y - J(\mu_{j-1})}{\gamma_j} + J(\mu_{j-1})(\theta - \mu_{j-1}); \Sigma_j
\right)
\]

(15)

where \( J \) is the Jacobian in (14) and \( \mathcal{N}(\mu, \Sigma) \) denotes a multivariate Gaussian PDF with mean \( \mu \) and covariance \( \Sigma \). Analogous to the Kalman filter, the posterior at each step is Gaussian with mean and variance given by

\[
\mu_j = \mu_{j-1} + K_j (y - s(\mu_{j-1}))
\]

(16a)

\[
\Psi_j = (I - K_j J) \Psi_{j-1}
\]

(16b)

where \( I \) is the identity matrix and \( K_j \) is the gain at the \( j \)th iteration defined as

\[
K_j = \Psi_{j-1} J' \left( J \Psi_{j-1} J' + \Sigma \right)^{-1}
\]

(17)

Initially, this approximation is poor but the likelihood will be wide due to severe flattening, which results in a small update. As the algorithm progresses, the approximate posterior approaches the true posterior and the linearization becomes more accurate. In this way, the correction schedule serves to smoothly transform the prior to the final posterior approximation. The resulting approximation is much more accurate than computing the posterior in one step. The correction schedule should be chosen such that the posterior update at each step is small, to mitigate errors from the linearized approximation. Although a large number of corrections will always satisfy this condition, the choice of a practical schedule will be a trade-off between accuracy and computation. This process is similar to simulated annealing [26] and we will see that the resulting algorithm is very robust to local minima, particularly at low SNR. The algorithm is computationally efficient as matrix inversions can exploit the diagonal nature of the measurement covariance. The final algorithm is described below.

function ESTIMATE_T2(y, \Sigma, \mu_0, \Psi_0, \gamma_1, \ldots, \gamma_P)
for \( j \leftarrow 1 \ldots P \) do
    Calculate Jacobian, \( J \leftarrow J(\mu_{j-1}) \)
    Calculate gain, \( K_j \leftarrow \Psi_{j-1} J' (J \Psi_{j-1} J' + \Sigma / \gamma_j)^{-1} \)
    Update mean, \( \mu_j \leftarrow \mu_{j-1} + K_j (y - s(\mu_{j-1})) \)
    Update covariance, \( \Psi_j \leftarrow (I - K_j J) \Psi_{j-1} \)
end for
return \( \mu_P \)
end function

IV. An Analysis of \( T_2 \) Width Estimation

In this section, we present a Cramér-Rao analysis of mode width estimation. The results confirm that width estimation is especially difficult for clinical imaging paradigms and justifies the use of the discrete model. To facilitate this analysis, we introduce a continuous parametric model that spans the gap between the pseudo-continuous and discrete \( T_2 \) distribution classes. This alternative model, which introduces mode width to the discrete multicomponent distribution, is a mixture of inverse-gamma distributed modes

\[
f(\tau) = \sum_{i=1}^{M} w_i \frac{\beta_i^\alpha_i}{\Gamma(\alpha_i)} \tau^{\alpha_i - 1} e^{-\beta_i \tau}
\]

(18)

where \( \Gamma(\cdot) \) denotes the gamma function and the three parameters, \( w_i, \alpha_i, \beta_i \), characterize the weight, location, and scale of the \( i \)th mode, respectively.

The inverse-gamma distribution is chosen for multiple reasons. Firstly, it can approximate a wide range of distributions. Secondly, the distribution has positive support, i.e., it is only nonzero for values between 0 and \( \infty \), which is suitable for positive relaxation times. Thirdly, it leads to a tractable integration in (1), important for the theoretical analysis presented here. The continuous parametric models presented in [8]–[10] do not fulfill all these requirements.

We assume the ideal exponential decay function for \( y_n \), and together with the inverse-gamma model (18), this leads to a closed-form signal model. Substituting the parametric form (18) into the signal model (1) gives

\[
y_m = e^{j\phi} \sum_{i=1}^{M} w_i \left( \frac{\beta_i}{t_n + \beta_i} \right)^{\alpha_i} + v_n.
\]

(19)
See Appendix C for the proof of (19).

Analogous to the multicomponent discrete model, the inverse gamma mixture distribution requires estimation of $3M + 1$ parameters. It is instructive to consider an alternative parameterization of each mode in the mixture, specified by the mean and variance of the inverse-gamma distribution, given by

$$
\nu_j = \frac{\beta_j}{\alpha_j - 1}, \quad \rho_j^2 = \frac{\beta_j^2}{(\alpha_j - 1)^2(\alpha_j - 2)},
$$

(20)

This parametrization is useful to analyze the fundamental ability to estimate the location and spread of the relaxation times.

### A. Cramèr-Rao Lower Bound Analysis

The minimum variance of an unbiased estimator is given by the Cramèr-Rao Lower Bound (CRLB) [27]. Here, we use the bound to analyze the intrinsic uncertainty in the estimation problems for different distribution models. The CRLB reveals our ability to estimate the model parameters for a given experimental setup.

The CRLB is derived from the Fisher Information Matrix (FIM), $\mathcal{I}$, which we calculate for three different models: the inverse-gamma model in (19) with unknown location and width parameters; a constrained inverse gamma model with unknown location and known width; and the discrete distribution model as in [12]. Note that CRLB analysis is not applicable to the grid model assumed by the NNLS method, and therefore we consider it separately in Section IV-B.

The inverse-gamma mixture is ideal to consider the estimation performance in terms of both the location and width of the unknown distribution components. To this end, we use the relationships in (20) to transform the FIM for this model (initially in terms of $\alpha$; and $\beta$) according to

$$
\mathcal{I}(\nu, \rho) = \mathcal{J}^T \mathcal{I}(\alpha, \beta) \mathcal{J}
$$

(21)

where $\mathcal{J}$ is the Jacobian matrix of the mapping associated with (20).

We evaluate the CRLB of an inverse gamma mixture distribution for two cases: Firstly, a single component distribution with location $\nu_1 = 100$ ms and width $\rho_1 = 10$ ms; secondly, a two-component distribution consisting of a slow mode with location $\nu_1 = 100$ ms and width $\rho_1 = 10$ ms and a fast mode with location $\nu_2 = 20$ ms and width $\rho_2 = 10$ ms. The bounds for the discrete signal model and an inverse-gamma mixture with fixed width were calculated for modes at equivalent locations. All two-component models had weights, $w_1 = 0.7$ and $w_2 = 0.3$ for the slow and fast modes, respectively, while the single component models had a weight $w_1 = 1$. The values for the two-component distribution were chosen to emulate a voxel with a dominant intra/extracellular component and a small myelin water component. The simulated MRI sequence is a multi-echo sequence consisting of $N = 32$ spin echoes with echo times spaced 12 ms apart, i.e., $t_n = (12n)$ ms, $n = 1, \ldots, 32$.

Fig. 1(a) displays the CRLB for the estimation of a single component distribution using different models and Fig. 1(b) shows the CRLB for first mode of a two-component distribution. The figure demonstrates that although width and location can be reasonably well estimated for a single component, the width parameter of the inverse gamma model is exceedingly difficult to estimate for two components. For an SNR of 100, the width of the slow mode can only be estimated with a standard deviation of $\sim 100$ ms, five times greater than the true width parameter. These results also highlight that estimating the location of multiple components is much harder when the width is unknown.

To achieve useful estimates of the weight, location and width parameters of the multicomponent distribution, the SNR would need to be $\sim 10000$, beyond that achievable in a clinical setting. Alternatively we would need to collect in the order of $10^4$ echoes, which is completely impractical. Similar plots for the second mode or different parameter values yield the same conclusions: it is not possible to reliably estimate the width of a voxel’s multicomponent $T_2$ distribution with the SNR and number of echoes currently achievable in a clinical setting.

### B. NNLS and Distribution Width

The inability of the available data to provide information about distribution width is a fundamental property which also affects NNLS. To investigate this further, a single $T_2$ component distribution modelled by a Gaussian was simulated with mean $T_2 = 100$ ms and different distribution widths ranging from 2
Fig. 2. Demonstration of the ability of NNLS to estimate distribution width of a single component. (a) RMSE of the width estimates extracted from NNLS distributions and (b) the true (solid line) and ten example distributions (dashed line) estimated from different noise realizations.

to 20 ms. For each width, NNLS was performed on 1000 independent trials and the estimated width was calculated by fitting a Gaussian to the generated distributions. Fig. 2(a) displays the root mean squared error (RMSE) of the NNLS width estimates, which demonstrates the degraded performance when estimating narrow widths. This is due to the regularization in NNLS necessary to mitigate the influence of noise. Estimated $T_2$ distributions from the first 10 measurement realizations are shown in Fig. 2(b) for a width of 10 ms and SNR of 100. The variability of the width of the estimated distributions is apparent.

Although estimation performance may be reasonable for a single component at high SNR, the per-voxel width estimates are not sufficiently reliable for multicomponent estimation from typical clinical data. As demonstrated in Fig. 1, the theoretical performance of width estimation degrades dramatically for multicomponent distributions. As a consequence, the NNLS results presented here will also degrade substantially for multiple components. Furthermore, extracting consistent width estimates from a multicomponent NNLS distribution is nontrivial. The reason for such poor estimation performance is that a large number of distributions produce very similar measurements. For reasonable noise levels, the differences are indistinguishable and discernment of the true component width is infeasible.

This further motivates the use of the discrete multicomponent model presented in Section II-B and our Bayesian estimation algorithm.

V. METHODS

A. Simulations

To compare the accuracy of the proposed Bayesian algorithm for estimation of discrete multicomponent distribution parameters with the NNLS approach across a range of SNRs, a decay curve was generated from a multicomponent distribution consisting of two discrete modes: a slow component with weight 0.7 and $T_2$ of 100 ms, and a fast component with weight 0.3 and $T_2$ of 20 ms. These general features are typical for white matter tissue in the cortex. A flip angle of 160° was simulated to represent a nonideal $B_1$ field. The signal was calculated from the EPG algorithm to model secondary and stimulated echoes as described by (9) and (4). Simulations consisted of 2000 independent trials where independent and identically distributed (i.i.d.) complex noise was added. The phase angle in (9) was randomly selected each trial from a uniform distribution between $-\pi$ and $\pi$ to demonstrate that the algorithm is insensitive to phase.

Each trial consisted of 1000 1D measurements. The signal was calculated from the EPG algorithm to model secondary and stimulated echoes as described by (9) and (4). Simulations consisted of 2000 independent trials where independent and identically distributed (i.i.d.) complex noise was added. The phase angle in (9) was randomly selected each trial from a uniform distribution between $-\pi$ and $\pi$ to demonstrate that the algorithm is insensitive to phase.

The ability to produce accurate estimates at a wide range of flip angles was examined by another simulation of 2000 trials with an SNR of 300 (as in [21]) for 30 flip angles evenly spaced between 60° and 180°. The MSE of both the estimated flip angle and the location of the slow $T_2$ component was calculated at each flip angle for three algorithms: the Bayesian algorithm and NNLS, as in [21], with 8 and 64 flip angle interpolation points.

The Bayesian algorithm employed a Gaussian prior for the two components, with

$$\mu_0 = \begin{bmatrix} 0.7 & 0.1^{-1} & 0.3 & 0.02^{-1} & 0 & 6\pi/9 \end{bmatrix}^T$$

$$\Psi_0 = \begin{bmatrix} 12 & 100^2 & 0.2^2 & 0 & 35^2 \\ 0 & 100^2 & 0.2^2 & 1 & 35^2 \\ & 0 & 35^2 & 35^2 & \end{bmatrix}$$

Recall the parameter vector defined in (11) consists of a weight and relaxation rate ($\tau^{-1}$) for each component followed by the
phase and flip angle. The prior means are set such that we expect the signal to consist of a dominant slow component and a weaker fast component. The prior covariance is relatively large such that the estimation bias is negligible. To avoid adverse behavior the prior mean should satisfy certain nonrestrictive requirements: different components should have nonidentical relaxation rate means and the flip angle should be strictly less than 180°. The former condition avoids the problem of label switching in mixture estimation [29]; the latter condition ensures the derivative of the signal is nonzero. Under these conditions, our experience is that estimation results are relatively insensitive to the prior mean provided the prior covariance is sufficiently large, which is consistent with Bayesian theory. A similar prior can be created for three or more components.

The correction schedule consists of 30 corrections, logarithmically spaced between $10^{-2}$ and 1 and normalized such that $\sum_{i} \gamma_i = 1$. The logarithmic spacing ensures that small corrections are made initially, when the linearized approximation can be poor, and progressively larger corrections are performed as the estimate becomes more accurate. Other schedules work equally well if a sufficient number of steps are defined.

B. Experiments

Experiments were performed on a 4.7T Bruker BioSpec small bore MRI scanner fitted with a high performance gradient set. A cryogenically cooled surface coil was used to improve SNR. A multi-echo CPMG sequence with 24 echoes was run with a first echo time of 12 ms and an echo spacing of 12 ms. The slice thickness was 1 mm for all experiments.

A sample was constructed consisting of agar gel with a sheep optic nerve fixed parallel to the transverse plane. The $T1R = 2500$ ms, $\gammaVO = 6.4$ mm × 12.8 mm with a matrix size of 64 × 128. $B1$ field mapping was performed using the method in [30]. Data was processed with the proposed algorithm using both the multi-exponential model, which assumes ideal 180° flip angles; and the multicomponent EPG model, which jointly estimates the $T2$ components, weights and $B1$ map.

A mouse brain was scanned with the same 24 echo sequence with $T1R = 2500$ ms, $\gammaVO = 15$ mm × 15 mm and a matrix size of 192 × 192 for an in-plane resolution of 78.1 $\mu$m. A single mid-axial slice of 1 mm was acquired. This data was processed with the proposed algorithm using the EPG model and the NNLS modification in [21]. Both algorithms attempt to estimate multicomponent distributions and account for stimulated echoes by estimating both the $T2$ distribution and the $B1$ map. As described above, $T2$ values were extracted from the NNLS distribution using the geometric mean to obtain location estimates of the fast and slow components as in [6].

To further investigate the reliability of the gradient based MLE algorithm, complex noise was added to the mouse data to yield an SNR of 200. This data was then processed with the proposed algorithm and a gradient based MLE algorithm.

VI. RESULTS

A. Simulations

The Bayesian algorithm proposed in Section III had an average execution time of 0.014 s per voxel compared to the gradient optimization at 0.026 s and NNLS with stimulated echo correction, which averaged 0.16 s per voxel. This represents a two-fold and eleven-fold improvement over existing algorithms.

Fig. 3 displays the performance of the NNLS algorithm, a gradient based MLE algorithm and the proposed Bayesian algorithm. This plot demonstrates that the MLE optimization algorithm achieves the same performance as the Bayesian algorithm for sufficiently high SNR. This is expected for a sufficiently wide prior, when the MLE converges with the maximum a posteriori (MAP) estimate generated by the Bayesian algorithm. However, the MLE algorithm fails to produce reliable estimates at low SNR values, due to an optimization cost function that is poorly behaved (it is relatively flat with local minima). NNLS exhibits suboptimal performance, although it achieves reasonable results across a range of SNR values.

Both the proposed algorithm and NNLS are reasonably accurate in estimating the flip angle. Fig. 4(a) illustrates that the error in the estimated flip angle increases with the true flip angle. This can be understood by noting that the derivative of the signal tends to zero as the flip angle approaches 180°. The 8-point NNLS algorithm as implemented in [21] exhibits elevated error around 80° and 160°. This is due to the algorithm testing only 8 flip angles and using spline interpolation to extract the minimum sum of squares error. A smooth NNLS curve can be obtained by using 64 test angles; however, this increases the computation time eightfold. Fig. 4(b) demonstrates the superior performance in $T2$ estimation of the proposed algorithm compared to NNLS. The RMSE of the proposed algorithm is approximately half that of NNLS. Converse to the error in flip angle, the error in $T2$ for all algorithms decreases as the true flip approaches 180°. This is attributed to the increased signal intensity at larger flip angles.

B. Optic Nerve Experiments

The $B1$ map of the optic nerve sample generated from the proposed Bayesian estimation algorithm is consistent with that generated from a low resolution $H1$ field mapping protocol (Fig. 5). The surface coil was located at one end of the nerve, as indicated by white bar. The surface coil was operated in transmit/receive mode, which creates a dramatic flip angle inhomogeneity.
Fig. 4. RMSE of (a) the flip angle and (b) the slow component, estimated from multiecho data with an SNR of 300. Error is plotted against different values of the true flip angle for the Bayesian algorithm (solid line) and the NNLS algorithm with 8 (dotted line) and 64 (dashed line) interpolation points.

Fig. 5. The $B_1$ map (left) estimated from the multi-echo data and (right) measured from separate spin echo images. White bar illustrates the approximate location of the surface coil.

throughout the sample, with only a small region exhibiting a flip angle close to the prescribed 180°. Note that regions above the 180° are experience a flip angle greater than 180° but are reflected about 180° as this is equivalent in the signal model.

Fig. 6 presents the multicomponent $T_2$ maps generated with different signal models. Fig. 6(a) displays the two components estimated using the EPG algorithm, which models stimulated echoes resulting from nonideal flip angles. The estimated $T_2$ values are similar along the length of the optic nerve, despite the large flip angle variation demonstrated in Fig. 5. Conversely, Fig. 6(b) shows the results of an exponential decay model that neglects flip angle inhomogeneity. The $T_2$ maps in this case exhibit a large variation along the nerve, with the fast component fitted in an attempt fix the disparity of the first few echo amplitudes resulting from stimulated echoes. Fig. 6(c) displays each signal model fit to the two points marked in Fig. 6(a) and (b). As expected the signal is similar in regions corresponding to flip angles close to the ideal 180° In other regions, however, the EPG algorithm provides a much closer fit to the data.

C. Mouse Brain Experiments

Fig. 7 presents results from experimental data, to further explain the poor performance of traditional gradient-based MLE...
optimization algorithms. Fig. 7(a) displays boxplots of the $T_2$ value of the first component for the two algorithms. The plots depict the median, first, and third quartiles, and values considered outliers. The estimates from the MLE optimization and the Bayesian algorithm are similar, as shown by similar medians and first and fourth quartiles. However, the MLE optimization has a higher number of outliers, many of which are far away from the true value. These outliers manifest as erroneous speckled values in the $T_2$ map, as illustrated in Fig. 7(b).

Estimation results from the mouse brain data are displayed in Fig. 8. Estimated values from the proposed algorithm are shown in the top image of each figure part, while the bottom image displays values extracted from the NNLS distribution. The weights and times of the slow component (calculated using the geometric mean between 40 and 200 ms) are displayed in Fig. 8(a) and (b), respectively. The Bayesian and NNLS algorithms produce similar $T_2$ maps for this component. Fig. 8(c) and (d) display the fast component and demonstrate an essential difference between the two algorithms. The maps produced by the proposed algorithm contain anatomical features such as a decreased $T_2$ below the ventricles, consistent with the presence of white matter. Fast components above 40 ms are not displayed. NNLS fast $T_2$ maps were calculated between 0 and 40 ms and exhibit a relatively constant $T_2$ value. NNLS appears to represent the white matter by increasing the weighting of the relatively constant $T_2$. This is due to the regularization of the NNLS algorithm which decreases the estimation variance at the expense of increased bias. On the other hand, the proposed algorithm produces a fast component with both an increased weight and a decreased time constant in the white matter region. It is worthwhile to mention that both algorithms often produce fast component weights close to zero as shown in Fig. 8(c).

VII. DISCUSSION

In this paper, we have presented a Bayesian algorithm for the estimation of discrete multicomponent $T_2$ distribution parameters, motivated by the presence of such discrete models in the literature, and the failure of traditional MLE-based approaches to provide robust and accurate $T_2$ mode estimates. Our algorithm makes use of a signal processing technique known as progressive correction to approach the true posterior density from a schedule of approximations. We have demonstrated the algorithm’s utility in application to both simulations and two sets of experimental data.

One drawback of the discrete multicomponent $T_2$ model is the need to know a priori the number of components. For example, the weights of the fast $T_2$ component were estimated close to zero in the agar gel region of the optic nerve sample (results not shown). This is due to the fact that the relaxation curve of the gel is described accurately by a single decaying component. It would be preferable to only fit a single component in this case, as the second component is simply fit to the noise and has minimal contribution to the signal. Many model order selection techniques exist, such as the Akaike or Bayesian information criteria, that choose an appropriate model order based on a statistical metric, see e.g., [31]. These criteria are applicable to MRI signal decay data although this is beyond the scope of the current paper.

The discrete multicomponent model has been compared with the popular NNLS approach. It is commonly thought that the regularization in NNLS produces “more realistic” distributions, primarily due to the smooth and continuous nature of the resulting curves [21]. Our analysis of the reliability of distribution width estimation with a typical clinical setup suggests however that a regularized distribution, although continuous, is not necessarily closer to the true distribution than the multicomponent discrete counterpart.
The distributions resulting from the NNLS algorithm have been most useful to date by extracting summary statistics, such as the mean intra/extra-cellular peak and the myelin water fraction \[6\]. These measures integrate the distribution and consequently are only weakly dependent on the estimated width. Our analysis justifies the use of these statistics, as these features can be reasonably well estimated, despite minimal knowledge of the distribution width. However it is well known, and our results confirm, that regularization can bias the distribution shape of the fast components, leading to a lack of anatomical information in the \(T_2^*\) maps.

The advantage of the discrete distribution model is that the associated estimation problem is inherently better conditioned than the NNLS pseudo-continuous grid model. This reduces the need for regularization and potentially minimizes the estimation bias. The minimal regularization is apparent in our Bayesian approach since the prior distribution for the unknown parameters is relatively wide. Additionally, the proposed algorithm requires substantially less computation compared to both gradient-based optimization and the NNLS algorithm.

Assuming an exponential decay function for the discrete distribution model, the estimated \(T_2\) components in the presence of nonideal flip angles exhibit more complicated behavior than the single component counterpart \[20\]. In \[20\], the \(T_2\) value was increasingly overestimated as the flip angle deviated from \(180^\circ\). In our case where a two component model was used, the second component is estimated to compensate for the stimulated echoes. This results in a large variation of the second component and relatively small variation of the first component.

**VIII. CONCLUSION**

We have presented a Bayesian algorithm for robust and reliable estimation of discrete multicomponent \(T_2\) models. This algorithm works together with the extended phase graph algorithm to account of nonideal flip angles and the resultant stimulated echoes. We have shown that the Bayesian algorithm exhibits near-optimal performance even at very low SNR, while the gradient-based MLE algorithm alternative was shown to be unsuitable for this purpose. The proposed algorithm jointly estimates the weights and locations of the \(T_2\) components as well as the signal phase and refocusing flip angle, making quantitative multicomponent \(T_2\) analysis viable in the presence of large \(B_1\) inhomogeneity, such as those produced from a transceive surface coil.

**IX. SOURCE CODE**

The source code for the Bayesian estimation algorithm, EPG algorithm and the recursive algorithm for partial derivatives are provided on the author’s website\(^1\). The algorithms are written in MATLAB version 2011b.

**APPENDIX**

The extended phase graph algorithm is a recursive algorithm that tracks the evolution of the phase coherence pathways \[19\]. It describes the spin system at a given time by a distribution over different phase states. There are three types of states, \(F^{(n)}_{k}\) (dephasing), \(F^{(n)}_{-k}\) (rephasing), and \(Z^{(n)}_{k}\) (longitudinal) for each coherence level, \(k\), and echo number, \(n\). After the initial excitation pulse, the magnetization is assumed to be coherent and on the transverse plane so \(F_{0}^{(0)} = 1\) and \(F_{-k}^{(0)} = Z_{k}^{(0)} = 0\) for \(k = 1, \ldots, N\). The evolution around subsequent refocusing pulses is modelled by three periods of relaxation/precession, excitation, followed by relaxation/precession again. The relaxation and precession periods are modelled by the recursive relation

\[
F_{k}^{(n+1)} = F_{k}^{(n)} e^{-\tau/T_2}
\]

\[
Z_{k}^{(n+1)} = Z_{k}^{(n)} e^{\tau/T_1}
\]

where the prime notation indicates the new state. We define a vector containing all the states at the \(n\)th echo as \(x_n = [F_{0}^{(n)}; F_{1}^{(n)}; F_{-1}^{(n)}; Z_{1}^{(n)}; \ldots; F_{N}^{(n)}; F_{-N}^{(n)}; Z_{N}^{(n)}]^{T}\). The relation in (23) can be described by a single matrix, \(P\), consisting of relaxation and coherence level shifts

\[
P = \begin{bmatrix}
e^{-\tau/T_2} & R_0 & R_0 \\
R_0 & R_0 & \ddots \\
\end{bmatrix} S
\]

where \(R_0 = \text{diag}(e^{-\tau/T_2}, e^{-\tau/T_1}, e^{-\tau/T_1})\) and \(S\) is a permutation of the identity matrix that shifts the transverse states up a coherence level.

An RF pulse of flip angle \(\alpha\) mixes states of equal coherence level described by

\[
\begin{bmatrix}
F_{-k}^{(n)} \\
F_{k}^{(n)} \\
Z_{k}^{(n)}
\end{bmatrix} = \begin{bmatrix}
\cos^2(\frac{\alpha}{2}) & \frac{1}{2}\sin\alpha & \sin\alpha \\
\frac{1}{2}\sin\alpha & \cos^2(\frac{\alpha}{2}) & -\frac{1}{2}\sin\alpha \\
-\frac{1}{2}\sin\alpha & \frac{1}{2}\sin\alpha & \cos\alpha
\end{bmatrix} \begin{bmatrix}
F_{-k}^{(n)} \\
F_{k}^{(n)} \\
Z_{k}^{(n)}
\end{bmatrix}
\]

(25)

The mixing in (25) can also be applied to all states simultaneously using a block diagonal matrix, \(T\)

\[
T = \begin{bmatrix}
1 & T_0 \\
T_0 & T_0 \\
\end{bmatrix}
\]

(26)

where \(T_0\) is the matrix defined in (25).

Combining these operators the evolution of the entire state vector between echoes is described as

\[
x_n = PT^n x_{n-1}. \quad (27)
\]

The echo amplitude is given by \(F_{0}^{(n)}\) at each time, \(n\). That is

\[
g_n = c^T x_n. \quad (28)
\]

where \(c = [1, 0, \ldots, 0]^T\).

**Partial Derivatives:** The derivative of the echo amplitude with respect to the unknown parameters is useful for a number

\[^1\text{http://people.eng.unimelb.edu.au/klayton/T2/}\]
function EPG_DERIVATIVES(r1, r2, α)
Calculate matrices, P and T using (24) and (26)
Calculate derivatives, \( P' = \frac{\partial P}{\partial r_2} \) and \( T' = \frac{\partial T}{\partial \alpha} \)
for \( n \leftarrow 1 \ldots N \) do
  > Calculate first term of product rule
  \[ x_n^{(r)} \leftarrow \text{PTP'} x_{n-1} \]
  \[ z_n^{(\alpha)} \leftarrow \text{PTP'} x_{n-1} \]
  > Calculate second term of product rule
  \[ x_n^{(r)} \leftarrow \text{PTP} \frac{\partial x_{n-1}}{\partial r_2} \]
  \[ x_n^{(\alpha)} \leftarrow \text{PTP} \frac{\partial x_{n-1}}{\partial \alpha} \]
  > Combine terms for the derivatives
  \[ \frac{\partial z_n}{\partial r_2} \leftarrow 2x_n^{(r)} + x_n^{(r)} \]
  \[ \frac{\partial z_n}{\partial \alpha} \leftarrow z_n^{(\alpha)} + x_n^{(\alpha)} \]
  > Calculate state evolution
  \[ x_n \leftarrow \text{PTP} x_{n-1} \]
end for
end function

Fig. 9. Recursive algorithm for the derivative of the EPG signal with respect to \( r_2 \) and flip angle, \( \alpha \).

of applications including CRLB analysis, gradient descent optimization and approximations of the likelihood. As such, an efficient means to compute the partial derivatives is included here.

The recursion in (27) is linear in the state vector thus the partial derivatives can also be calculated in a recursive manner. First notice that both \( x_n \) and \( k \) depend on the parameter of interest, \( \theta \), so the product rule must be applied

\[
\frac{\partial x_n}{\partial \theta} = \frac{\partial E}{\partial \theta} x_{n-1} + E \frac{\partial x_{n-1}}{\partial \theta}.
\]

Next we use the decomposition \( P = P(r_1, r_2) \) and \( T = T(\alpha) \). The partial derivative of the matrix elements in \( E = \text{PTP} \) are

\[
\frac{\partial E}{\partial \alpha} - \frac{\partial T}{\partial \alpha} P
\]

\[
\frac{\partial E}{\partial r_2} = \frac{\partial P}{\partial r_2} T + PT \frac{\partial P}{\partial r_2}.
\]

For convenience, we define additional intermediate states \( x_n^{(r)}, x_n^{(\alpha)}, z_n^{(r)}, \) and \( z_n^{(\alpha)} \) to calculate the different terms in (29)–(31). We also let \( P' = \frac{\partial P}{\partial r_2} \) and \( T' = \frac{\partial T}{\partial \alpha} \). The recursive structure is shown in Fig. 9. Note that the operators \( P, P', T, \) and \( T' \) can be executed very efficiently using the inherent structure described in (23) and (25). The derivatives of the EPG signal are simply the first element of the derivatives of the state vector, \( \frac{\partial x_n}{\partial r_2} \) and \( \frac{\partial x_n}{\partial \alpha} \).

Signal Model for Inverse-Gamma Mixture: The measured signal from a distribution consisting of an inverse-gamma mixture is derived as follows:

\[
y_n = e^{i\phi} \sum_{i=1}^{M} w_i \frac{\beta_i}{\Gamma(\alpha_i)} \int_{-\infty}^{0} (t + \beta_i)_{-1} \frac{1}{u} \frac{1}{\Gamma(\alpha_i + 1)} e^{-u} \frac{1}{u^2} du \]

(36)

\[
y_n = e^{i\phi} \sum_{i=1}^{M} w_i \frac{\beta_i}{\Gamma(\alpha_i + 1)} \int_{0}^{\infty} u^2 e^{-u} du \]

(37)

\[
y_n = e^{i\phi} \sum_{i=1}^{M} w_i \frac{\beta_i}{t + \beta_i} \alpha_i \]

(38)

Let \( u = (t + \beta_i)/\tau \). Taking derivatives gives

\[
d\tau = -\frac{1}{u^2} du.
\]

(35)

Changing the variable of integration, and noting the new limits of \( -\infty \) to 0, gives

\[
y_n = e^{i\phi} \sum_{i=1}^{M} w_i \frac{\beta_i}{\Gamma(\alpha_i + 1)} \int_{0}^{\infty} u^2 e^{-u} du \]

(39)

ACKNOWLEDGMENT

The authors would like to thank A. Atik, J. Cheong, and M. Tolcos for supplying the sheep optic nerve used in this study.

REFERENCES


