Maximum-Entropy Estimation of Joint Relaxation-Diffusion Distribution Using Multi-TE Diffusion MRI

Lipeng Ning[1,2(✉)]

[1] Brigham and Women's Hospital, Boston, MA 02215, USA
[2] Harvard Medical School, Boston, MA 02215, USA
lning@bwh.harvard.edu

Abstract. Combined relaxation-diffusion MRI (rdMRI) is a technique to probe tissue microstructure using diffusion MRI data with multiple b-values and echo time. Joint analysis of rdMRI data can characterize the joint relaxation and diffusion distribution (RDD) function to examine heterogeneous tissue microstructure without using multi-component models. This paper shows that the problem of estimating RDD functions is equivalent to the multivariate Hausdorff moment problem by applying a change of variables. Three formulations of maximum entropy (ME) estimation problems are proposed to solve the inverse problem to derive ME-RDD functions in different parameter spaces. All three formulations can be solved by using convex optimization algorithms. The performance of the proposed algorithms is compared with the standard methods using basis functions based on simulations and *in vivo* rdMRI data. Results show that the proposed methods provide a more accurate estimation of RDD functions than basis-function methods.

Keywords: Diffusion MRI · T_2 relaxation · maximum entropy

1 Introduction

Diffusion magnetic resonance imaging (dMRI) is sensitive to water diffusion in biological tissue. Analytical models of dMRI signals have played an essential role in quantifying tissue microstructure in clinical studies. Several methods have been developed to use dMRI signals measured with a single or multiple b-values to estimate compartment-specific diffusivity or diffusion propagators [1,20,24]. But these methods are developed using dMRI data acquired with a fixed echo time (TE). Several studies have shown that joint modeling of dMRI with multiple TE can probe TE-dependent diffusivity and [23], tissue-specific T_2 relaxation rate [14] and the joint relaxation diffusion distribution (RDD) in each voxel [2,3,12,19]. More specifically, RDD functions describe the multidimensional distribution of T_2 relaxation rate and diffusivity in each voxel, providing a general framework to characterize heterogeneous tissue microstructure. The RDD functions were first applied to measure the structure of porous media

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14227, pp. 439–448, 2023.
https://doi.org/10.1007/978-3-031-43993-3_43

[4,6,8,11]. It was generalized in [2,3,12] to probe the microstructure of biological tissue using rdMRI. A standard approach for estimating RDD functions is to represent the measurement signal using basis functions of different diffusivity and relaxation rates which may lead to biased estimation results because of the strong coupling between basis signals.

This work introduces the maximum entropy (ME) estimation method for more accurate estimation of RDD functions by adapting theories and techniques developed for the classical Hausdorff moment problems [10,13,15,18]. ME estimation is also a standard approach for high-resolution power spectral estimation of time series data which involves a similar trigonometric moment problem [7,21]. The ME power spectral estimation usually performs better than Fourier transform-based methods [21], which motivates this work to derive ME methods for estimating RDD functions and compare the results with basis function-based methods. To this end, we first show that the problem of estimation RDD function is equivalent to the multivariate Hausdorff moment problem by applying a change of variables. Three formulations of maximum entropy (ME) estimation problems are proposed to estimate ME-RDD functions in different parameter spaces. The performance of these methods is compared with results based on basis functions using simulations and *in vivo* data.

2 Method

2.1 On the Hausdorff Moment Problem

Let $s(b,t)$ denote the dMRI signal with b and t being the b-value and TE, respectively. It is a standard approach to represent $s(b,t)$ as

$$s(b,t) = \int_{I_D \times I_R} e^{-bD-tR} d\rho(D,R), \tag{1}$$

where $\rho(D,R)$ denotes the joint distribution of diffusion and relaxation coefficients and $I_D := [0, D_0]$ and $I_R = [0, R_0]$ denote the finite intervals for D and R. For simplicity, let $\boldsymbol{\theta} := [D,R]$, $\mathbf{x} := [b,t]$ and $\Theta := I_D \times I_R$. Then, (1) is equivalent to

$$s(\mathbf{x}) = \int_{\Theta} e^{-\mathbf{x}\cdot\boldsymbol{\theta}} p(\boldsymbol{\theta}) d\boldsymbol{\theta}, \tag{2}$$

where $\mathbf{x}\cdot\boldsymbol{\theta}$ denotes the inner product between $\mathbf{x}$ and $\boldsymbol{\theta}$. Assume that the signals are sampled at $b = 0, \delta_b, \ldots, n_b\delta_b$ and $t = t_{\min}, t_{\min}+\delta_t, \ldots, t_{\min}+n_t\delta_t$, where $t_{\min}$ denotes the shortest TE. Let $\mathbf{x}_{\min} = [0, t_{\min}]$. Then let $\mathbf{x}_\mathbf{k} := [k_1\delta_b, k_2\delta_t]$ with $\mathbf{k} = [k_1,k_2] \in \mathcal{K}$ where $\mathcal{K} = \{[k_1,k_2] \mid 0 \le k_1 \le n_b, 0 \le k_2 \le n_t\} \subset \mathcal{Z}_+^2$ represents the set of all feasible indices. Next, let $s_\mathbf{k} := s(\mathbf{x}_\mathbf{k})$ which satisfies that

$$s_\mathbf{k} = \int_{\Theta} e^{-\mathbf{x}_\mathbf{k}\cdot\boldsymbol{\theta} - \mathbf{x}_{\min}\cdot\boldsymbol{\theta}} p(\boldsymbol{\theta}) d\boldsymbol{\theta}, \tag{3}$$

$$= \int_{\Theta} e^{-\mathbf{x}_\mathbf{k}\cdot\boldsymbol{\theta}} \hat{p}(\boldsymbol{\theta}) d\boldsymbol{\theta}, \tag{4}$$

where $\hat{p}(\boldsymbol{\theta}) = e^{-\mathbf{x}_{\min}\cdot\boldsymbol{\theta}}p(\boldsymbol{\theta})$ is a scaled RDD function adjusted based on the non-zero minimum TE.

The Hausdorff moment problem focuses on the existence of distribution functions that satisfy a sequence of power moments [10,15]. To change $s_{\mathbf{k}}$ to power moments as in the Hausdorff moment problem, we define $\boldsymbol{\gamma} := [e^{-\delta_b\theta_1}, e^{-\delta_t\theta_2}]$, which takes value in the interval $\Gamma := [e^{-\delta_b D_0}, 1] \times [e^{-\delta_t R_0}, 1]$. For a vector $\mathbf{k}$, we define $\boldsymbol{\gamma}^{\mathbf{k}} = \prod_i \gamma_i^{k_i} = e^{-\mathbf{x}_{\mathbf{k}}\cdot\boldsymbol{\theta}}$ following the convention in [17]. Then Eq. (4) can be expressed using the new variables as

$$s_{\mathbf{k}} = \int_{\Gamma} \boldsymbol{\gamma}^{\mathbf{k}} f(\boldsymbol{\gamma}) d\boldsymbol{\gamma}, \tag{5}$$

with $f(\boldsymbol{\gamma}) = \frac{1}{\delta_b\delta_t\gamma_1\gamma_2}\hat{p}(\boldsymbol{\theta}(\boldsymbol{\gamma}))$. Thus, $s_{\mathbf{k}}$ can be considered as the power moments of the density function $f(\boldsymbol{\gamma})$ on the interval Γ. Therefore the problem of estimating RDD functions using finite rdMRI measurements is equivalent to a multivariate Hausdorff moment problem. We note that the unit interval is usually considered in Hausdorff moment problems. This can be obtained by changing the variable $\boldsymbol{\gamma}$ t $\hat{\boldsymbol{\gamma}} = [\frac{\gamma_1-\alpha_1}{1-\alpha_1}, \frac{\gamma_2-\alpha_2}{1-\alpha_2}]$ where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2] = [e^{-\delta_b D_0}, e^{-\delta_t R_0}]$. Thus $\hat{\boldsymbol{\gamma}}$ takes the value on the unit interval I^2 whose moments can be computed using linear transforms of $s_{\mathbf{k}}$. To simplify notations, the analysis in the following subsection will be based on $s_{\mathbf{k}}$ and the distribution of $\boldsymbol{\gamma}$.

2.2 Maximum-Entropy Estimation

The RDD functions that satisfy the rdMRI data may not be unique. In moment problems, the maximum entropy (ME) method is a standard approach to estimate probability distributions and power spectral density functions [15]. Based on the three representations of rdMRI data in Eqs. (3), (4) and (5), three optimization problems are introduced to estimate ME-RDD functions below.

The first ME problem is developed based on Eq. (3) as below:

$$\begin{aligned} &\max_{p(\boldsymbol{\theta})} \int_{\Theta} -p(\boldsymbol{\theta}) \ln p(\boldsymbol{\theta}) d\boldsymbol{\theta} \\ &\text{s.t.} \int_{\Theta} e^{-\mathbf{x}_{\mathbf{k}}\cdot\boldsymbol{\theta}-\mathbf{x}_{\min}\cdot\boldsymbol{\theta}} p(\boldsymbol{\theta}) d\boldsymbol{\theta} = s_{\mathbf{k}}, \forall \mathbf{k} \in \mathcal{K}, \end{aligned} \tag{6}$$

where the objective function is the Shannon differential entropy of $p(\boldsymbol{\theta})$. The solutions to ME problems have been extensively investigated in moment problems [15]. Using the Lagrangian method, one can derive that the optimal solution has the following form

$$p_{\lambda}^{\text{me1}}(\boldsymbol{\theta}) = \exp\left(-\left(\sum_{\mathbf{k}\in\mathcal{K}} \lambda_{\mathbf{k}} e^{-\mathbf{x}_{\mathbf{k}}\cdot\boldsymbol{\theta}-\mathbf{x}_{\min}\cdot\boldsymbol{\theta}}\right) - 1\right) \tag{7}$$

for some coefficients $\lambda_{\mathbf{k}}$ with $\mathbf{k} \in \mathcal{K}$. The optimal parameters $\lambda_{\mathbf{k}}$ need to be solved to satisfy the constraints in Eq. (6).

Based on Eq. (4), the second ME problem is introduced as below:

$$\max_{\hat{p}} \int_{\Theta} -\hat{p}(\boldsymbol{\theta}) \ln \hat{p}(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{8}$$
$$\text{s.t.} \int_{\Theta} e^{-\mathbf{x_k}\cdot\boldsymbol{\theta}} \hat{p}(\boldsymbol{\theta}) d\boldsymbol{\theta} = s_{\mathbf{k}}, \forall \mathbf{k} \in \mathcal{K},$$

where $\hat{p}(\boldsymbol{\theta})$ is the pre-scaled RDD to adjust for the nonzero $t_{\min}$. The optimal solution to Eq. (8) has the following form

$$\hat{p}_{\lambda}^{\text{me2}}(\boldsymbol{\theta}) = \exp\left(-\sum_{\mathbf{k}\in\mathcal{K}} \lambda_{\mathbf{k}} e^{-\mathbf{k}\cdot\boldsymbol{\theta}}\right). \tag{9}$$

It is noted that Eq. (9) does not include the constant -1 as in Eq. (7) since it can be absorbed by the variable λ_0. Then, $\hat{p}_{\text{ME2}}(\boldsymbol{\theta})$ can be scaled back to the original RDD function by $p_{\lambda}^{\text{me2}}(\boldsymbol{\theta}) = \exp\left(-\left(\sum_{\mathbf{k}\in\mathcal{K}} \lambda_{\mathbf{k}} \cdot e^{-\mathbf{x_k}\cdot\boldsymbol{\theta}}\right) + \mathbf{x}_{\min}\cdot\boldsymbol{\theta}\right)$, which has a different form than the solution in Eq. (7).

The third ME-RDD is estimated based on the new variable $\boldsymbol{\gamma}$ as in Eq. (5) by solving the following problem:

$$\max_{f} \int_{\Gamma} -f(\boldsymbol{\gamma}) \ln f(\boldsymbol{\gamma}) d\boldsymbol{\gamma} \tag{10}$$
$$\text{s.t.} \int_{\Gamma} \boldsymbol{\gamma}^{\mathbf{k}} f(\boldsymbol{\gamma}) d\boldsymbol{\gamma} = s_{\mathbf{k}}, \forall \mathbf{k} \in \mathcal{K}. \tag{11}$$

By changing the variable $\boldsymbol{\gamma}$ back to $\boldsymbol{\theta}$, Eq. (10) is transformed to

$$\max_{\hat{p}} \int_{\Theta} -\hat{p}(\boldsymbol{\theta}) \left(\ln \hat{p}(\boldsymbol{\theta}) - \ln\left(\delta_b \delta_t e^{-\boldsymbol{\delta}\cdot\boldsymbol{\theta}}\right)\right) d\boldsymbol{\theta}$$
$$\text{s.t.} \int_{\Theta} e^{-\mathbf{x_k}\cdot\boldsymbol{\theta}} \hat{p}(\boldsymbol{\theta}) d\boldsymbol{\theta} = s_{\mathbf{k}}, \forall \mathbf{k} \in \mathcal{K},$$

where $\boldsymbol{\delta} = [\delta_b, \delta_t]$. It is interesting to note that the above objective function is equal to minimizing the *Kullback-Leibler* divergence, i.e., the relative entropy, between $\hat{p}(\boldsymbol{\theta})$ and $\delta_b \delta_t e^{-\boldsymbol{\delta}\cdot\boldsymbol{\theta}}$. The optimal solution has the following form

$$\hat{p}_{\lambda}^{\text{me3}}(\boldsymbol{\theta}) = \exp\left(-\left(\sum_{\mathbf{k}\in\mathcal{K}} \lambda_{\mathbf{k}} e^{-\mathbf{k}\cdot\boldsymbol{\theta}}\right) - \boldsymbol{\delta}\cdot\boldsymbol{\theta}\right). \tag{12}$$

Then, $\hat{p}_{\lambda}^{\text{me3}}(\boldsymbol{\theta})$ is scaled back to obtain

$$p_{\lambda}^{\text{me3}}(\boldsymbol{\theta}) = \exp\left(-\left(\sum_{\mathbf{k}\in\mathcal{K}} \lambda_{\mathbf{k}} \cdot e^{-\mathbf{x_k}\cdot\boldsymbol{\theta}}\right) + (\mathbf{x}_{\min} - \boldsymbol{\delta})\cdot\boldsymbol{\theta}\right). \tag{13}$$

It is noted the difference between $p_{\lambda}^{\text{me1}}(\boldsymbol{\theta})$ and $p_{\lambda}^{\text{me2}}(\boldsymbol{\theta})$ is related to the non-zero offset $x_{\min}$. The two solutions are equal if the shortest TE is zero, but it is impossible in practice. The difference between $p_{\lambda}^{\text{me2}}(\boldsymbol{\theta})$ and $p_{\lambda}^{\text{me3}}(\boldsymbol{\theta})$ is related to the sampling rate δ_b and δ_t. The difference is less significant, with a higher sampling rate in the b-value and TEs.

2.3 Dual Energy Minimization Problems

The optimal values of $\lambda_{\mathbf{k}}$ in Eqs. (7), (9) and (12) can be obtained by solving the dual formulation of the ME problems, which can be expressed as energy minimization problems based on the dual formulations.

For the solution in Eq. (7), the corresponding energy function, i.e., the dual objective function, is given by

$$\Delta_1(\boldsymbol{\lambda}) = \int_{\Theta} \exp\left(-\left(\sum_{\mathbf{k}\in\mathcal{K}} \lambda_{\mathbf{k}} e^{-\mathbf{x}_{\mathbf{k}}\cdot\boldsymbol{\theta}-\mathbf{x}_{\min}\cdot\boldsymbol{\theta}}\right) - 1\right) d\boldsymbol{\theta} + \sum_{\mathbf{k}} \lambda_{\mathbf{k}} s_{\mathbf{k}}. \tag{14}$$

It can be shown that

$$\frac{\partial^2 \Delta_1}{\partial \boldsymbol{\lambda}_{\mathbf{k}} \partial \boldsymbol{\lambda}_{\boldsymbol{\ell}}} = \int_{\Theta} e^{-\mathbf{x}_{\mathbf{k}}\cdot\boldsymbol{\theta}-\mathbf{x}_{\boldsymbol{\ell}}\cdot\boldsymbol{\theta}-2\mathbf{x}_{\min}\cdot\boldsymbol{\theta}} p_{\lambda}^{\mathrm{me1}}(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{15}$$

Thus, the Hessian matrix of $\Delta_1(\boldsymbol{\lambda})$ positive definite, indicating that $\Delta_1(\boldsymbol{\lambda})$ is a convex function. If $\Delta_1(\boldsymbol{\lambda})$ has a finite minimizer, then the minimizer satisfies that

$$\frac{\partial \Delta_1}{\partial \lambda_{\mathbf{k}}} = -\int_{\Theta} e^{-\mathbf{x}_{\mathbf{k}}\cdot\boldsymbol{\theta}-\mathbf{x}_{\min}\cdot\boldsymbol{\theta}} p_{\lambda}^{\mathrm{me1}}(\boldsymbol{\theta}) d\boldsymbol{\theta} + s_{\mathbf{k}} = 0. \tag{16}$$

Therefore, the optimal parameters for $p_{\lambda}^{\mathrm{me1}}(\boldsymbol{\theta})$ can be obtained from the minimizer of $\Delta_1(\boldsymbol{\lambda})$.

The optimal λ for Eq. (9) and Eq. (12) can be obtained by minimizing the following to convex energy functions:

$$\Delta_2(\lambda) = \int_{\Theta} \exp\left(-\left(\sum_{\mathbf{k}\in\mathcal{K}} \lambda_{\mathbf{k}} e^{-\mathbf{x}_{\mathbf{k}}\cdot\boldsymbol{\theta}}\right)\right) d\boldsymbol{\theta} + \sum_{\mathbf{k}} \lambda_{\mathbf{k}} s_{\mathbf{k}}, \tag{17}$$

$$\Delta_3(\lambda) = \int_{\Theta} \exp\left(-\left(\sum_{\mathbf{k}\in\mathcal{K}} \lambda_{\mathbf{k}} e^{-\mathbf{x}_{\mathbf{k}}\cdot\boldsymbol{\theta}}\right) - \boldsymbol{\delta}\cdot\boldsymbol{\theta}\right) d\boldsymbol{\theta} + \sum_{\mathbf{k}} \lambda_{\mathbf{k}} s_{\mathbf{k}}. \tag{18}$$

In this paper, the energy minimization problem was solved using a customized Newton algorithm with the Armijo line-search method [5]. The code and data used in this work are available at https://github.com/LipengNing/ME-RDD.

3 Examples

3.1 Synthetic Data

The proposed algorithms were examined using synthetic rdMRI data with an RDD function consisting of three Gaussian components with the mean at (1.5 $\mu m^2/ms$, 10 ms^{-1}), (0.5 $\mu m^2/ms$, 40 ms^{-1}), and (1.5 $\mu m^2/ms$, 40 ms^{-1}) with the volume fraction being 0.2, 0.5 and 0.3, respectively. D and R were uncorrelated

in each component, with the standard deviation being 0.01 $\mu m^2/ms$ and $5ms^{-1}$. Simulated rdMRI signals had b-values at $b = 0, 0.5\ ms/\mu m^2, \ldots, 5\ ms/\mu m^2$ and TEs at $t = 50\ ms, 75\ ms, \ldots, 200\ ms$, which can be achieved using an advanced MRI scanner for in vivo human brains such as the connectom scanner [9]. Then, independently and identically distributed Gaussian noise was added to simulated rdMRI signals with an average signal-to-noise ratio (SNR) from 100 to 600, similar to the range of the SNR of direction-averaged dMRI signals of in vivo human brains that scale according to the square root of the number of directions and voxel size.

3.2 Comparison Methods

For comparison, we applied the basis-function representation method, similar to the methods in [12], by representing the signals as

$$s(\mathbf{x}) = \sum_{n=1}^{N} c_n e^{-\mathbf{x}\cdot\boldsymbol{\theta}_n}, \tag{19}$$

where $\boldsymbol{\theta}_n$ are a set of predefined points on a discrete grid in Θ. We solved a constrained L_2 minimization problem to find the optimal non-negative coefficient c_n with minimum L_2 norm. To evaluate the performances, we computed the error of the center of mass (CE) of the estimated RDD functions in three regions defined using a watershed clustering approach; see the top left figure in Fig. 2. The CE in diffusivity and re Moreover, we also computed the volume-fraction error (VFE), i.e., $\text{VFE} = \sum_{k=1}^{K} |F_{\text{est}}(k) - F_{\text{true}}(k)|$, of the estimated RDD, where $F_{\text{est}}(k)$ and $F_{\text{true}}(k)$ denote the true and estimated volume fraction for each component.

3.3 In Vivo rdMRI

The proposed ME algorithms were applied to an in vivo rdMRI dataset acquired in our previous work [16]. The data was acquired from a healthy volunteer on a 3T Siemens Prisma scanner with the following parameters: voxel size = $2.5\times2.5\times2.5mm^3$, matrix size = $96\times96\times54$, TE = 71, 101, 131, 161 and 191 ms, TR=5.9 s, b = 700, 1400, 2100, 2800, 3500 s/mm^2 along 30 gradient directions together with 6 volumes at b = 0, simultaneous multi-slice (SMS) factor = 2, and iPAT = 2. The pulse width of the diffusion gradients and the diffusion time were fixed across scans. An additional pair of b=0 images with anterior-posterior (AP), and posterior-anterior (PA) phase encoding directions were acquired for distortion correction using FSL TOPUP/eddy. Then the data were further processed using the unring [22] tool.

4 Results

The first two figures in Fig. 1 show the error in the center of mass of D and R. The L2 approach based on basis functions had significantly overestimated D

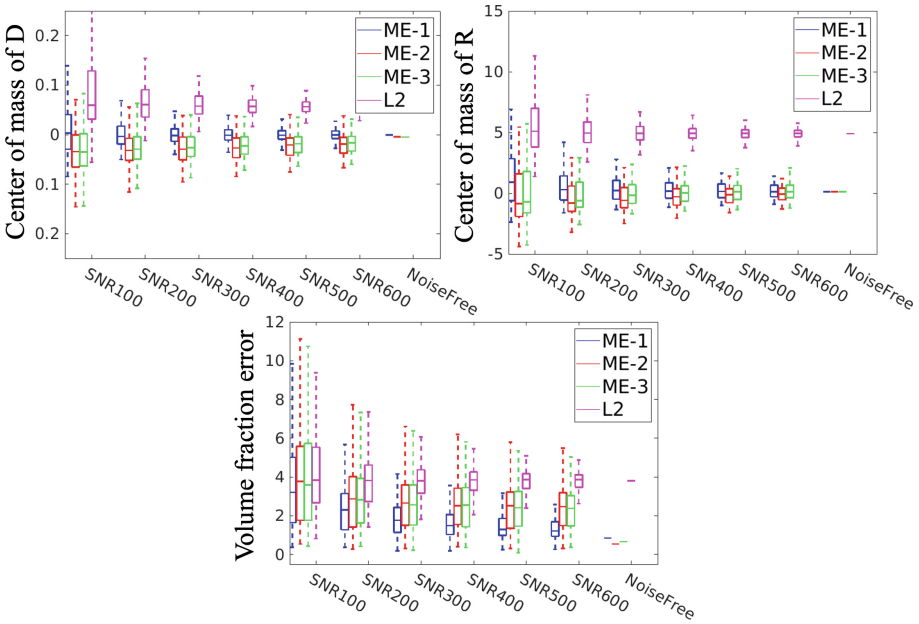

Fig. 1. Performance of ME-RDD and L2-based RDD functions using synthetic data for different simulated SNR levels.

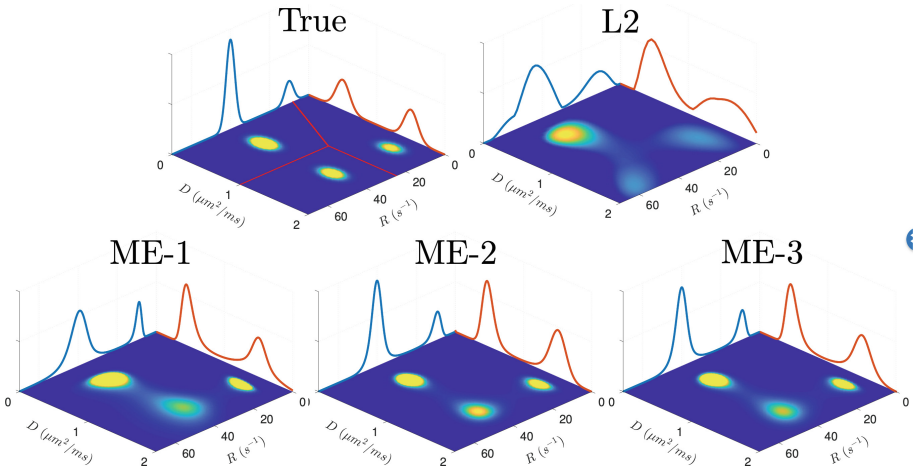

Fig. 2. Comparison of ME-RDD and L2-based functions with noise-free synthetic data.

and R, whereas the three ME methods had much lower estimation error and similar performance to each other. The L2 method also had a biased estimation of the volume fraction, as shown in the third figure. Figure 2 shows the true and estimated RDD functions. The L2-RDD has more spread and biased distributions compared to the true distribution. The three ME-RDD functions were more focal and less biased.

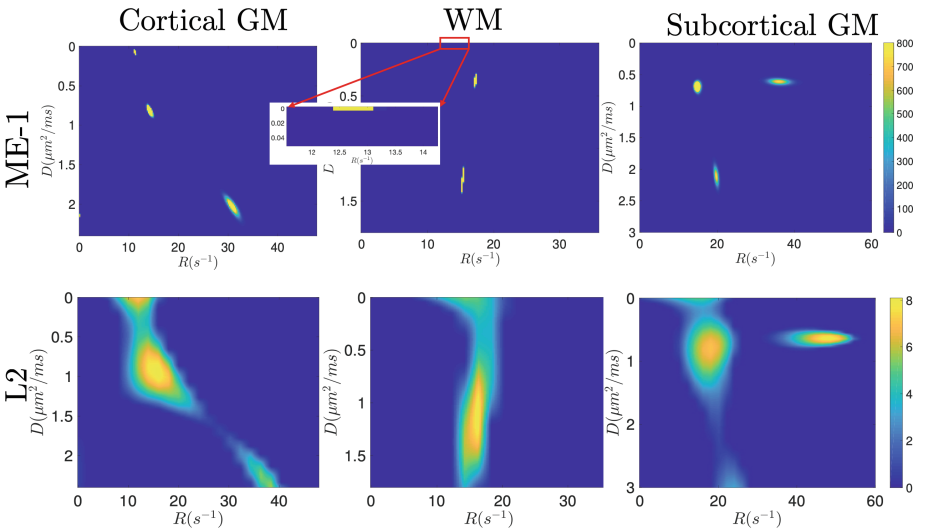

Fig. 3. Comparison of ME-RDD (ME-1)and L2-based functions using *in vivo* data.

Figure 3 compares the ME-RDD (using ME-1) and L2-RDD in the white matter, cortical, and subcortical gray matter of a human brain. The computation time to estimate the ME-RDD for one voxel is about 100 s using an Intel Xeon E5 cpu (2.20 GHz). The results for ME-2 and ME-3 were similar to ME-1. The ME-RDD has shown three separate components for each tissue type. But the L2-RDD has shown two components for cortical GM and subcortical GM with higher R values compared to the components in ME-RDD. For WM, the L2-RDD shows a continuous spectrum without a clear separation of the underlying components.

5 Summary

In summary, this work introduced a maximum-entropy framework for estimating the relaxation-diffusion distribution functions using rdMRI. To our knowledge, this is the first work showing that the estimation of multidimensional RDD functions is equivalent to the classical multivariate Hausdorff moment problem. Although this work focuses on the two dimensional RDD functions, the results

generalize to the special cases for one dimensional relaxation or diffusion distribution functions. The contributions of this work also include the development of three algorithms to estimate RDD functions and the comparisons with the standard basis-function approach. The ME-RDD functions can be estimated using convex optimization algorithms. Experimental results have shown that the proposed methods provide more accurate parameters for each component and more accurate volume fractions compared to the standard basis function methods. Moreover, results based on in vivo data have shown that the proposed ME-RDD can resolve multiple components that cannot be distinguished by the basis function approach. The better performance ME-RDD functions compared to basis-function methods may relate to the superior performance of ME spectral estimation methods compared to Fourier transform-based methods [21]. We expect the improved spectral resolution will be useful in several clinical applications such as lesion and tumor detection. But further theoretical analysis on the performance of the ME methods is needed in future work. Moreover, further histological validations are needed to examine the biological basis of the RDD functions. Finally, we note a limitation of the proposed method is that the optimization algorithm may have a slow convergence speed because the Hessian matrices may not be well conditioned. Moreover, the results may be sensitive to measurement noise. Thus faster and more reliable computation algorithms will be developed in future work.

References

1. Basser, P.J., Mattiello, J., LeBihan, D.: MR diffusion tensor spectroscopy and imaging. Biophys. J. **66**, 259–267 (1994)
2. Benjamini, D., Basser, P.J.: Use of marginal distributions constrained optimization (MADCO) for accelerated 2D MRI relaxometry and diffusometry. J. Magn. Reson. **271**, 40–45 (2016)
3. Benjamini, D., Basser, P.J.: Towards clinically feasible relaxation-diffusion correlation MRI using MADCO. Microporous Mesoporous Mater. **269**, 93–96 (2018)
4. Bernin, D., Topgaard, D.: NMR diffusion and relaxation correlation methods: new insights in heterogeneous materials (2013)
5. Bertsekas, D.P.: Nonlinear Programming. 2nd edn (1999)
6. Callaghan, P.T., Arns, C.H., Galvosas, P., Hunter, M.W., Qiao, Y., Washburn, K.E.: Recent Fourier and Laplace perspectives for multidimensional NMR in porous media. Magn. Reson. Imaging **25**, 441–444 (2007)
7. Cover, T.M., Thomas, J.A.: Elements of Information Theory (1993)
8. De Almeida Martins, J.P., Topgaard, D.: Multidimensional correlation of nuclear relaxation rates and diffusion tensors for model-free investigations of heterogeneous anisotropic porous materials. Sci. Rep. **8**, 2488 (2018)
9. Fan, Q., et al.: Mapping the human connectome using diffusion MRI at 300 mT/m gradient strength: methodological advances and scientific impact. NeuroImage **254**, 118958 (2022)
10. Frontini, M., Tagliani, A.: Hausdorff moment problem and maximum entropy: on the existence conditions. Appl. Math. Comput. **218**(2), 430–433 (2011)

11. Hürlimann, M.D., Venkataramanan, L.: Quantitative measurement of two-dimensional distribution functions of diffusion and relaxation in grossly inhomogeneous fields. J. Magn. Reson. **157**, 31–42 (2002)
12. Kim, D., Doyle, E.K., Wisnowski, J.L., Kim, J.H., Haldar, J.P.: Diffusion-relaxation correlation spectroscopic imaging: a multidimensional approach for probing microstructure. Magn. Reson. Med. **78**, 2236–2249 (2017)
13. Landau, H.J.: Maximum entropy and the moment problem. Bull. Am. Math. Soc. **16**(1), 47–77 (1987)
14. McKinnon, E.T., Jensen, J.H.: Measuring intra-axonal T2 in white matter with direction-averaged diffusion MRI. Magn. Reson. Med. **81**, 2985–2994 (2019)
15. Mead, L.R., Papanicolaou, N.: Maximum entropy in the problem of moments. J. Math. Phys. **25**(8), 2404–2417 (1984)
16. Ning, L., Gagoski, B., Szczepankiewicz, F., Westin, C.F., Rathi, Y.: Joint relaxation-diffusion imaging moments to probe neurite microstructure. IEEE Trans. Med. Imaging **39**, 668–677 (2020)
17. Putinar, M., Scheiderer, C.: Multivariate moment problems: geometry and indeterminateness. Annali della Scuola Normale - Classe di Scienze **5**(2), 137–157 (2006)
18. Schmüdgen, K.: The Moment Problem. GTM, vol. 277. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64546-9
19. Slator, P.J., et al.: Combined diffusion-relaxometry microstructure imaging: current status and future prospects. Magn. Reson. Med. **86**(6), 2987–3011 (2021)
20. Tuch, D.S.: Q-ball imaging. Magn. Reson. Med. **52**, 1358–1372 (2004)
21. Ulrych, T.J., Bishop, T.N.: Maximum entropy spectral analysis and autoregressive decomposition. Rev. Geophys. **13**, 183–200 (1975)
22. Veraart, J., Fieremans, E., Jelescu, I.O., Knoll, F., Novikov, D.S.: Gibbs ringing in diffusion MRI. Magn. Reson. Med. **76**(1), 301–314 (2016)
23. Veraart, J., Novikov, D.S., Fieremans, E.: TE dependent diffusion Imaging (TEdDI) distinguishes between compartmental T2 relaxation times. NeuroImage **182**, 360–369 (2018)
24. Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C.: NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. NeuroImage **61**(4), 1000–1016 (2012)