

Retaining both discrete and smooth features in 1D and 2D NMR relaxation and diffusion experiments

A. Reci, A. J. Sederman^{a)}, L. F. Gladden

Department of Chemical Engineering and Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge CB3 0AS, United Kingdom

^{a)} Corresponding author

Name: Andrew J. Sederman
Address: Department of Chemical Engineering and Biotechnology
University of Cambridge
Philippa Fawcett Drive
Cambridge CB3 0AS, University of Cambridge
Email: ajs40@cam.ac.uk
Tel: (+44) 1223 766338

^{a)} Author to whom correspondence should be addressed. Electronic mail: ajs40@cam.ac.uk.

Abstract

A new method of regularization of 1D and 2D NMR relaxation and diffusion experiments is proposed and a robust algorithm for its implementation is introduced. The new form of regularization, termed the Modified Total Generalized Variation (MTGV) regularization, offers a compromise between distinguishing discrete and smooth features in the reconstructed distributions. The method is compared to the conventional method of Tikhonov regularization and the recently proposed method of L_1 regularization, when applied to simulated data of 1D spin-lattice relaxation, T_1 , 1D spin-spin relaxation, T_2 , and 2D T_1 - T_2 NMR experiments. A range of simulated distributions composed of two lognormally distributed peaks were studied. The distributions differed with regard to the variance of the peaks, which were designed to investigate a range of distributions containing only discrete, only smooth or both features in the same distribution. Three different signal-to-noise ratios were studied: 2000, 200 and 20. A new metric is proposed to compare the distributions reconstructed from the different regularization methods with the true distributions. The metric is designed to penalise reconstructed distributions which show artefact peaks. Based on this metric, MTGV regularization performs better than Tikhonov and L_1 regularization in all cases except when the distribution is known to only comprise of discrete peaks, in which case L_1 regularization is slightly more accurate than MTGV regularization.

Keywords: NMR relaxation and diffusion, Modified Total Generalized Variation, 1D and 2D inverse Laplace transformation

1. Introduction

Measurements of spin-lattice relaxation, T_1 , spin-spin relaxation, T_2 , and diffusion coefficient, D , distributions have found applications in many areas. 1D T_1 and T_2 distributions are used in obtaining the pore structure of rocks [1, 2] and probing water compartments in cells [3], while 1D distributions of D are used to characterize polymer molecular weight [4] and emulsion droplet size distributions [5]. These distributions can also be obtained jointly from 2D NMR correlation experiments. For example, T_1 - T_2 experiments are used in characterizing the pore structure of rocks [6], probing adsorbate-adsorbent interactions [7], and discriminating between tissues in medicine [8], while T_2 - T_2 experiments are used in probing exchange rates between different environments [9-13]. D - T_2 experiments are used in characterizing fluids in porous media [14], correlating D and T_2 with the viscosity of heavy oils [15], and in analysing pore geometry [16], while D - D experiments have found application in studying anisotropic diffusion in surfactant systems [17, 18] and drug delivery systems [19] and in distinguishing between restricted and bulk diffusion in porous media [20].

Central to the processing of the NMR data in these experiments are the 1D and 2D Inverse Laplace Transforms. It is well known that these transforms are ill-conditioned; that is, slight variations in the acquired signal, caused by the random nature of noise, can lead to significantly different reconstructions of the required distributions. The challenge in processing the NMR data has attracted much attention [21-27]. The most common method of tackling the ill-conditioning has been to perform Tikhonov regularization [22] and successful algorithms have been proposed to invert this regularized problem [23, 24]. Although Tikhonov regularization is very stable to noise, its main disadvantage is the inherent assumption that the distribution is smooth; that is there is a gradual transition from one feature to another. Therefore, the method suffers from a low spectral resolution [28-30]. Indeed, the generally accepted limitation is that Tikhonov regularization cannot distinguish between features in the distribution which differ from each other by a factor of ~ 3 or less in the relaxation time constants or diffusion coefficients [28]. As a result, use of Tikhonov regularization is more appropriate if there is a prior knowledge that the true distribution is smooth. If the true distribution is expected to have only a few non-zero entries (sparse), a new form of regularization, L_1 regularization, has been recently proposed [31-34]. In particular, it was shown in the work of Reci *et al.* [34] that L_1 regularization can resolve components in the distribution which have relaxation constants and diffusion coefficients that differ by as little as 10%. However, the method does not perform well if the true distribution map is smooth. Both methods of regularization, Tikhonov and L_1 , depend on a prior knowledge of what the true distribution should look like. Therefore, choosing the most appropriate method to use will depend on what the distribution is expected to be. However, in many practical cases, such knowledge may not be available, or it is exactly what one is trying to find out. Therefore, imposing an incorrect prior knowledge can lead to choosing the incorrect method which will give an inaccurate reconstruction.

We propose a new form of regularization, based on an adaptation of Total Generalized Variation (TGV) regularization [35], which will be termed Modified Total Generalized Variation (MTGV) regularization. TGV regularization has been successfully used in image denoising and deblurring, as well as MRI image reconstruction [36-39]. The regularization method proposed offers a compromise in distinguishing between smooth and discrete features in the distribution. Therefore, it is no longer required to know *a priori* whether the distribution is expected to be smooth or sparse, since MTGV

regularization will retain both features. We also introduce an algorithm that is robust in reconstructing the 1D and 2D distributions using MTGV regularization and provide some practical guidance in implementing it.

This work focuses on simulated 1D and 2D NMR signals used to obtain the distributions of the longitudinal relaxation time constant, T_1 , and the transverse relaxation time constant, T_2 . These distributions can be obtained either by performing dedicated 1D experiments or by performing a T_1 - T_2 experiment and projecting the 2D distribution map in each dimension. The T_1 and T_2 distributions reconstructed from MTGV regularization are compared to the results obtained from Tikhonov and L_1 regularization. Simulations are performed on a range of different input T_1 and T_2 distributions, differing with respect to the variance of the peaks in the distributions, and at three different signal-to-noise ratios (SNR) of 2000, 200 and 20. A new metric is proposed to compare the reconstructed distributions from each method with the input distributions, which allows a quantitative comparison to be made between Tikhonov, L_1 and MTGV regularizations. The comparison of the different regularization techniques is performed with regards to simulated rather than experimental data because in experimental data the true distribution is not known. This renders the objective comparison between the different regularization techniques from experimental data impossible.

The paper is structured as follows: In section 2, Tikhonov and L_1 regularization methods are defined. In section 3, MTGV regularization is introduced and an overview of the algorithm proposed to solve the 1D and 2D MTGV regularized problems is given. Section 4 describes the simulations performed and the results are presented in section 5.

2. Conventional methods

The most common 1D NMR experiments used to measure the distribution of relaxation time constants T_1 and T_2 are, respectively, the Inversion Recovery (IR) and CPMG experiments, the signal from which can be formulated as:

$$S_1(t_1) = \int_0^\infty F_1(T_1)(1 - 2e^{-t_1/T_1}) dT_1 + E_1(t_1), \quad (1)$$

$$S_2(t_2) = \int_0^\infty F_2(T_2) e^{-t_2/T_2} dT_2 + E_2(t_2), \quad (2)$$

where F_1 , F_2 are the distribution functions, t_1 , t_2 are the time parameters varied during the experiments and E_1 , E_2 represent the unknown noise assumed inherent in the signals.

A 2D T_1 - T_2 experiment stacks the 1D experiments in series and the signal acquired from such an experiment can be formulated as:

$$S(t_1, t_2) = \int_0^\infty \int_0^\infty F(T_1, T_2)(1 - 2e^{-t_1/T_1}) e^{-t_2/T_2} dT_1 dT_2 + E(t_1, t_2), \quad (3)$$

where F is the joint T_1 - T_2 distribution. Individual T_1 or T_2 distributions can be obtained either by performing a 1D NMR experiment and inverting its signal or by performing a 2D NMR experiment, inverting its signal and projecting the reconstructed 2D distribution in the appropriate dimension.

In a discrete form, Eq. (1-3) can all be written as:

$$\underline{S} = \underline{K} \underline{F} + \underline{E}, \quad (4)$$

where \underline{S} is either the discretized vector of S_1 or S_2 or the discretized matrix of S , written in a vector form by stacking its column vectors (vectorization). \underline{F} is either the discretized vector of F_1 or F_2 or the discretized matrix of F , vectorized. \underline{K} is a kernel matrix and \underline{E} is

either the discretized vector of E_1 or E_2 or the discretized matrix of E , vectorized. The reconstruction problem is formulated as: *Given \underline{S} and \underline{K} , estimate \underline{F} .*

The kernel matrix, \underline{K} , is ill-conditioned because of the exponentially decaying functions that it is composed of. Therefore, tackling the reconstruction problem is challenging. If the distribution \underline{F} is expected to be smooth, the method of choice to tackle the ill-conditioning is to perform a Tikhonov regularization. This consists of estimating \underline{F} from the following minimization problem:

$$\underline{F} = \arg \min_{\underline{F} \geq 0} \left(\frac{\alpha}{2} \|\underline{K} \underline{F} - \underline{S}\|_2^2 + \frac{1}{2} \|\underline{R} \underline{F}\|_2^2 \right), \quad (5)$$

where \underline{R} is a regularization matrix and α is a regularization parameter. The first term is called the fidelity term, while the second term is called the penalty term. For 1D experiments, \underline{R} is typically a matrix that performs the second derivative of \underline{F} , therefore constraining \underline{F} to be smooth. This particular choice of \underline{R} is sometimes referred to as the Phillips-Twomey method [40, 41]. For 2D experiments, \underline{R} is typically the identity matrix, \underline{I} , [42], constraining \underline{F} to have a small Euclidean norm. This particular choice of \underline{R} is made because of the difficulty of defining the second derivative of a matrix in both dimensions. Tikhonov regularization can only be reliably used if it is known *a priori* that \underline{F} will be smooth. Otherwise, it tends to smooth features which are discrete.

If the distribution \underline{F} is expected to be discrete, the recent trend is to use another form of regularization, L_1 regularization. This consists of obtaining the distribution \underline{F} from the following minimization problem:

$$\underline{F} = \arg \min_{\underline{F} \geq 0} \left(\frac{\alpha}{2} \|\underline{K} \underline{F} - \underline{S}\|_2^2 + \|\underline{F}\|_1 \right). \quad (6)$$

$\|\underline{F}\|_1$ refers to the L_1 norm of the distribution, where the L_p norm of a vector \underline{A} , with entries a_1, a_2, \dots, a_n is defined as:

$$L_p(\underline{A}) = \|\underline{A}\|_p = (\sum_{i=1}^n |a_i|^p)^{1/p}. \quad (7)$$

The inclusion of the L_1 penalty term gives sparser reconstructions of \underline{F} . The L_1 regularization is only suitable when it is known *a priori* that \underline{F} will be discrete. For distributions which are smooth, it gives artefacts in the reconstructed distributions.

3. Proposed method

The motivation for the present work is to develop a method of processing 1D and 2D NMR relaxation and diffusion data which does not rely on prior information of whether the distribution is discrete, smooth, or contains both types of features.

- 1.
- 2.
- 3.

In order to deal with such cases, a new regularization method is proposed which is based on an adaptation of Total Generalized Variation (TGV) regularization, termed the Modified Total Generalized Variation (MTGV) regularization. It consists of estimating the distribution \underline{F} from the following minimization problem:

$$(\underline{F}, \underline{W}) = \arg \min_{\underline{F} \geq 0, \underline{W}} \left(\frac{\alpha}{2} \|\underline{K} \underline{F} - \underline{S}\|_2^2 + \|\underline{F} - \underline{W}\|_1 + \beta \|\underline{D}_2 \underline{W}\|_1^* \right), \quad (8)$$

where \underline{W} is an auxiliary vector and \underline{D}_2 is a matrix that performs some form of second derivative of the vector it is applied on. The $\|(\dots)\|_1^*$ norm is the L_1 norm for 1D NMR experiments. The definition of the $\|(\dots)\|_1^*$ norm for 2D NMR experiments is slightly different in order to take into account the fact that the second derivative of a matrix can be taken in different directions; the proper definition is given in the appendix.

In overview and with reference to section 2, it is seen that the penalty term in MTGV regularization is composed of two parts: the first one enforces discrete features in the reconstructed distribution, while the second term enforces smooth features, with the weight between these two terms controlled by the regularization parameter β . Therefore, MTGV regularization offers a compromise between discrete and smooth features. The trade-off between fidelity and regularization is controlled by the other regularization parameter, α . In the original application of TGV regularization [35], the term that enforced sparsity was $\|\underline{D} \underline{F} - \underline{W}\|_1$, where \underline{D} performs the first derivative of \underline{F} . This was done because the workers were interested in obtaining piece-wise constant images, which is sparse in the first derivative domain. However, in NMR relaxation and diffusion experiments, the emphasis is in discrete distributions, which are sparse themselves, hence the formulation as proposed in Eq. (8). A similar approach exists for the L_1 regularization of 2D NMR correlation experiments [31-34].

The overview of the steps involved in the development of an algorithm, accompanied by a pseudocode to numerically solve Eq. (8) for both 1D and 2D NMR data is given in the appendix. The algorithm is based on the Primal-Dual Hybrid Gradient Method (PDHGM) [43]. The appendix also contains practical considerations about the implementation of the algorithm.

4. Simulations

The performance of Tikhonov, L_1 and MTGV regularization in the reconstruction of individual T_1 and T_2 distributions are compared to each other by performing the following numerical experiments:

- a) Tikhonov regularization applied to 1D NMR T_1 and T_2 data, referred to as 1D Tikhonov regularization.
- b) L_1 regularization applied to 1D NMR T_1 and T_2 data, referred to as 1D L_1 regularization.
- c) MTGV regularization applied to 1D NMR T_1 and T_2 data, referred to as 1D MTGV regularization.
- d) Tikhonov regularization applied to 2D T_1 - T_2 data, with the reconstructed map projected in each dimension, referred to as 2D Tikhonov regularization.
- e) L_1 regularization applied to 2D T_1 - T_2 data, with the reconstructed map projected in each dimension, referred to as 2D L_1 regularization.
- f) MTGV regularization applied to 2D T_1 - T_2 data, with the reconstructed map projected in each dimension, referred to as 2D MTGV regularization.

In order to investigate the performance of the different regularization methods, 7 simulated 2D T_1 - T_2 distributions and their corresponding T_1 and T_2 projections were used. All these distributions were 32×32 logarithmically spaced maps that had two lognormally distributed peaks of the same magnitude and centred at $T_1 = T_2 = 0.05$ s and $T_1 = T_2 = 0.15$ s. The distributions, labelled A-G, are illustrated in Fig. 1. For distributions A-F both peaks were of similar smoothness, with the level of smoothness increasing from distributions A to F. Distribution G is an example of a distribution

where one of the peaks is discrete and the other is smooth. The lognormal distribution details for all of the peaks are given in Table 1.

From the 1D and 2D distributions, noiseless 1D and 2D NMR signals were simulated according to Eqs. (1-3). 32 logarithmically spaced t_1 steps and 256 linearly spaced t_2 steps from 0 to 0.75 s were used. Random Gaussian noise was then added to the signals such that 2D NMR experiments at SNR values of 2000, 200 and 20 were studied. For the respective 1D signals, the standard deviation of the noise was chosen according to the method described by Celik *et al.* [44]. Before processing, the 2D NMR data were truncated according to the standard technique described by Venkataramanan *et al.* [45]. The distribution maps were then reconstructed from the noisy simulated signals using the different regularization methods.

In order to compare the reconstructed distributions with the true distributions, the following metric is proposed: If $F_{\text{true},i}$ is the true distribution's i -th entry and $F_{\text{rec},i}$ the reconstructed distribution's i -th entry, then a metric of how close the reconstructed distribution is to the true distribution is:

$$\chi = \sum_i \frac{(F_{\text{rec},i} - F_{\text{true},i})^2}{\max(10^{-4}, F_{\text{true},i})}. \quad (9)$$

One reason for choosing such a metric instead of the more commonly used mean square error or peak-signal-to-noise ratio (PSNR) is that the latter do not give any structural information. In particular, any extra peaks in the reconstructed distribution, as compared to the true distribution are not heavily penalized. This issue is addressed in the proposed metric by the division operation. A threshold of 10^{-4} was used for the division process, in order to account for distribution entries with $F_{\text{true},i} \approx 0$. Simulations showed that reducing the threshold further made no difference to the calculation of χ . The proposed metric is similar to the goodness of fit measure which is used in hypothesis testing in statistics [46]. The smaller the value of χ , the closer the reconstructed map is to the true distribution. The perfect reconstruction has $\chi = 0$.

For Tikhonov and L_1 regularization, χ was estimated for a range of the regularization parameter α , from 10^{-4} to 10^4 and the smallest value (corresponding to the best reconstruction) was recorded. For MTGV regularization, χ was estimated for a range of the regularization parameter α , from 10^{-7} to 10^2 and a range of the regularization parameter β , from 10^{-10} to 10^{-1} and the smallest value was recorded. A comparison of the lowest χ from the different regularization methods is then a comparison of their performance, with the best method being the one with the smallest χ .

5. Results and discussion

The reconstructed T_1 and T_2 distributions from the different regularization methods, Tikhonov, L_1 and MTGV are compared to each other using the metric χ . Initially, the comparison is performed on distributions A-F which contain two peaks that are of similar smoothness, with smoothness increasing from distribution A to F. Three values of SNR: 2000, 200 and 20 are investigated. The results are shown in Figs. 2-4. In each case, the smaller the value of χ , the better the reconstruction of the distributions by that particular regularization method.

Comparing the performance of Tikhonov and L_1 regularization with MTGV regularization in processing the 1D NMR T_1 or T_2 data at SNR = 2000 in Fig. 2, it is clearly seen that MTGV regularization is always superior, apart from the case when the

true distribution is very discrete (distribution A), in which case L_1 regularization is slightly more accurate. This is valid for both T_1 and T_2 distributions. Further, applying MTGV regularization to 2D NMR T_1 - T_2 data is always superior to applying Tikhonov or L_1 regularization to 2D NMR T_1 - T_2 data, apart from the case when the distribution is very discrete (distribution A), in which case L_1 regularization is again slightly more accurate.

The same conclusions can be drawn for the reconstructions obtained at lower signal-to-noise ratios shown in Fig. 3 (SNR = 200) and Fig. 4 (SNR = 20). At all levels of noise, MTGV outperforms Tikhonov and L_1 regularization when processing 1D or 2D NMR data, apart from when the distribution is very discrete, in which case L_1 regularization is slightly more accurate.

The results in Figs. 2-4 also confirm the well-known strengths of Tikhonov and L_1 regularization; it is seen that when the true distribution is discrete (distributions A and B), L_1 regularization (whether performed on 1D or 2D data) performs better than Tikhonov regularization. When the true distribution is smooth (distributions C to F), Tikhonov regularization performs better than L_1 regularization. It is also seen that when Tikhonov regularization is applied to the T_1 - T_2 discrete distributions A and B and the reconstructed maps are projected in each dimension, this gives better results than applying Tikhonov regularization to the individual 1D data. Therefore, 2D experiments give a better spectral resolution than 1D experiments, when processed with Tikhonov regularization. This is in agreement with the observations made by Celik *et al.* [44]. However, it is observed that when the true distribution is smooth (distributions C to F), this conclusion is no longer valid, and 1D experiments processed with Tikhonov regularization perform better than 2D experiments processed with Tikhonov regularization and then projected in each dimension.

Considering the implementation of the MTGV regularizer in more detail, it was discussed in section 3 that the regularization parameter β controls the relative amount of smoothness and discreteness imposed in the reconstruction. A method for choosing the regularization parameter β is now introduced. This method constrains β to be a function of the other regularization parameter α . Therefore, MTGV regularization becomes effectively a one-parameter regularization method, similar to Tikhonov and L_1 regularization. In Figs. 2-4, the value of the metric χ quoted for the MTGV regularization was that obtained by varying α and β independently in Eq. (8) and finding the minimum χ obtained in this process. In addition to recording this minimum χ , the values of α and β at which this minimum is obtained were also recorded. Fig. 5 shows these values of α and β for all the 1D and 2D MTGV regularization data displayed in Figs. 2-4. It can be observed that while the optimal α and β vary from one experiment to the other, there seems to be a correlation between $\log(\alpha)$ and $\log(\beta)$, and a relationship of the form $\beta = c\alpha$ is proposed, where c is a constant. The value of c is expected to depend mainly on the type of distribution (discrete or smooth) and the noise level in the data. Fig. 5 shows that $c = 2 \times 10^{-5}$ for 1D MTGV regularization and $c = 0.3$ for 2D MTGV regularization are a good estimate for a range of distributions and noise levels. Therefore, similar values of c are expected when MTGV regularization is applied to other 1D and 2D NMR relaxation and diffusion data. The choice of the only remaining free regularization parameter, α , can be done in similar ways to the methods that exist for choosing the regularization parameter in Tikhonov regularization [42].

In order to investigate the effect that constraining β has on the performance of MTGV regularization, similar simulations to the ones described in section 4 were performed on

distributions A-F by constraining the regularization parameter β to $\beta = 2 \times 10^{-5} \alpha$ in 1D MTGV regularization and $\beta = 0.3 \alpha$ in 2D MTGV regularization. An example of the comparison between the performance of Tikhonov, L_1 and MTGV regularization (with constrained β) at SNR = 20 is shown in Fig. 6 for both T_1 and T_2 distributions. A negligible difference is observed between the results in Fig. 6 and the results in Fig. 4, which compared the performance of Tikhonov, L_1 and MTGV regularization for an unconstrained β at SNR = 20. Similar conclusions can be made at SNR = 2000 and SNR = 200; the simulation results for these cases are not shown here. These results indicate that rendering MTGV regularization as a one-parameter regularization method by constraining β does not change any of the conclusions made in this work when β was considered a free parameter.

The comparison between the different regularization techniques has so far been made on distributions A-F, which contain two peaks of similar smoothness, with the degree of smoothness increasing from A to F. An example of the comparison between Tikhonov, L_1 and MTGV regularization for a distribution which contains a mixture of discrete and smooth features is shown in Fig. 7. Fig. 7 compares the reconstructions of the T_1 and T_2 distributions of the true distribution G, illustrated in Fig. 1, at three different SNR: 20, 200 and 2000. MTGV regularization reconstructions were performed with a constrained regularization parameter β . It is clearly seen that 1D MTGV regularization always outperforms 1D Tikhonov and 1D L_1 regularizations and 2D MTGV regularization always outperforms 2D Tikhonov and 2D L_1 regularizations.

The results presented in this section suggest that MTGV regularization gives the best reconstructions of 1D and 2D NMR relaxation data, regardless of the type of distribution and the SNR in the NMR data. The only exception is the case when the true distribution is known to only comprise of discrete peaks, in which case L_1 regularization gives slightly more accurate reconstructions.

Although in this paper the capabilities of MTGV regularization have been shown with respect to 1D and 2D NMR relaxation data, the methods can be easily adapted to 1D NMR diffusion data [4, 5] and 2D NMR correlation experiments which are diffusion encoded [14-20].

6. Conclusions

In this paper, a new method of regularization of 1D and 2D NMR relaxation and diffusion data was proposed, termed Modified Total Generalized Variation (MTGV) regularization. Unlike other methods of regularization of Tikhonov and L_1 , MTGV regularization offers a compromise between preserving smooth and discrete features in the reconstructed distributions. This eliminates the requirement of knowing *a priori* what the distribution should look like before selecting the appropriate regularization technique to process the data. The improvements offered by MTGV regularization were demonstrated by applying it to simulated 1D T_1 , 1D T_2 and 2D T_1 - T_2 NMR data. MTGV regularization always outperforms Tikhonov and L_1 regularization, apart from the case when the true distribution is known to only comprise of discrete peaks, in which case L_1 regularization performs slightly better.

Acknowledgments

A. R. acknowledges the Gates Trust Cambridge for financial support. L. F. G. and A. J. S. thank EPSRC (EP/K039318/1) and Microsoft Research for funding.

Appendix: 1D and 2D MTGV regularization algorithm

An algorithm is presented for the MTGV regularization of 1D and 2D NMR data. The steps involved in the development of the algorithm are similar for the 1D and 2D cases; the main difference is in the handling of derivatives and norms of vectors and vectorized matrices. The overview of steps involved in the development of the algorithm is presented below. Where applicable, differences between the 1D and 2D MTGV regularization implementations are emphasized.

- a) For the 1D MTGV algorithm, the 1D NMR signal is collected in a $p \times 1$ vector, \underline{S} . \underline{K} is a $p \times n$ kernel matrix and \underline{F} is the $n \times 1$ distribution vector to be found. \underline{W} is an $n \times 1$ auxiliary vector.

For the 2D MTGV algorithm, the 2D NMR signal is collected in a $p \times q$ matrix and its column vectors are stacked into a $pq \times 1$ vector, \underline{S} . \underline{K} is a $pq \times mn$ kernel matrix and \underline{F} is an $mn \times 1$ vector to be found. Once \underline{F} is found, it will be reshaped into an $m \times n$ matrix by the inverse process of stacking; this is the reconstructed distribution. \underline{W} is an $mn \times 1$ auxiliary vector. In order to make the explanation easier, it will be assumed that $m = n$.

- b) For the 1D MTGV algorithm, define \underline{D}_2 as the $n \times n$ matrix that performs the second derivative of a 1D $n \times 1$ signal. \underline{D}_2 is a tri-diagonal matrix with entries 1, -2 and 1 in diagonals -1, 0 and 1. Neumann boundary conditions were used, which practically translates into the last two rows of \underline{D}_2 being composed only of zeros. This ensures that the derivative is performed within the domain of the vector it acts upon.

For the 2D MTGV algorithm, define \underline{D}_1 as the $n \times n$ matrix that performs the first derivative of a 1D $n \times 1$ signal and \underline{I}_n be the $n \times n$ identity matrix. \underline{D}_1 is a bi-diagonal matrix with entries -1 and 1 in diagonals 0 and 1. Neumann boundary conditions were used, which practically translates into the last row of \underline{D}_1 being composed only of zeros. Then, $\underline{D}_v = \underline{I}_n \otimes \underline{D}_1$ is an $n^2 \times n^2$ matrix that performs the first derivative in the vertical direction of a vectorized $n^2 \times 1$ matrix, while $\underline{D}_h = \underline{D}_1 \otimes \underline{I}_n$ is an $n^2 \times n^2$ matrix that performs the first derivative in the horizontal direction of a vectorized $n^2 \times 1$ matrix. As a result, the $2n^2 \times n^2$ matrix $\underline{D} = [\underline{D}_v^T \ \underline{D}_h^T]^T$ performs the 1D derivatives of a vectorized $n^2 \times 1$ matrix in both directions and stores the result in a $2n^2 \times 1$ vector. Define the $3n^2 \times 2n^2$ matrix \underline{G} :

$$\underline{G} = \begin{bmatrix} -\underline{D}_v^T & \underline{0} \\ -\underline{D}_h^T & -\underline{D}_v^T \\ \underline{0} & -\underline{D}_h^T \end{bmatrix}. \quad (\text{A1})$$

The $3n^2 \times n^2$ matrix $\underline{D}_2 = \underline{G} \underline{D}$ performs the symmetrical second order derivative of a vectorized $n^2 \times 1$ matrix and stores it in a $3n^2 \times 1$ vector.

- c) For the 1D MTGV algorithm, define $\|\underline{A}\|_1^*$ for an $n \times 1$ vector \underline{A} as the L_1 norm, according to Eq. (7).

For the 2D MTGV algorithm, define $\|\underline{A}\|_1^*$ for a $3n^2 \times 1$ vector \underline{A} with entries $a_1, a_2, \dots, a_{3n^2}$ as:

$$\|\underline{A}\|_1^* = \sum_{i=1}^{n^2} \sqrt{\sum_{j=0}^2 a_{i+jn^2}^2}. \quad (\text{A2})$$

- d) For the 1D MTGV algorithm, let \underline{Y}_1 and \underline{Y}_2 be $n \times 1$ vectors, while for the 2D MTGV algorithm, let \underline{Y}_1 be a $n^2 \times 1$ vector and \underline{Y}_2 be a $3n^2 \times 1$ vector. It can be shown that for both 1D and 2D MTGV regularization algorithms, Eq. (8) is equivalent to:

$$(\underline{F}, \underline{W}, \underline{Y}_1, \underline{Y}_2) = \arg \min_{\underline{F} \geq 0, \underline{W}} \max_{\underline{Y}_1, \underline{Y}_2} \left(\frac{\alpha}{2} \|\underline{K} \underline{F} - \underline{S}\|_2^2 + \underline{Y}_1^T (\underline{F} - \underline{W}) + \underline{Y}_2^T \underline{D}_2 \underline{W} - h(\underline{Y}_1) - h\left(\frac{\underline{Y}_2}{\beta}\right) \right), \quad (\text{A3})$$

where $h(\underline{Y}_1)$ is the indicator function [47] defined as:

$$h(\underline{Y}_1) = \begin{cases} 0 & \|\underline{Y}_1\|_\infty \leq 1 \\ +\infty & \|\underline{Y}_1\|_\infty > 1 \end{cases}. \quad (\text{A4})$$

The problem in Eq. (A3) is a primal-dual problem, which was chosen to be tackled by the Primal-Dual Hybrid Gradient Method [43]. A pseudocode of the algorithm implemented is described in the following paragraph. In the pseudocode: `reshape(A, n1, n2)` reshapes the $n_1 n_2 \times 1$ vector \underline{A} into an $n_1 \times n_2$ matrix; `sum(A, i)` sums the column or row vectors of matrix \underline{A} along the i -th dimension; and `repmat(A, k, i)` stacks k copies of \underline{A} in the i -th dimension.

Step 1. Choose algorithm step parameters τ, σ and the regularization parameters, α, β .

Step 2. Set the convergence tolerance, TOL .

Step 3. Calculate $\underline{B} = (\underline{I} + \tau \alpha \underline{K}^T \underline{K})^{-1}$.

Step 4. Initialize $\underline{Y}_1^{(0)} = \underline{Y}_2^{(0)} = \underline{0}$, $\underline{F}^{(0)} \neq \underline{0}$, $\underline{W}^{(0)} \neq \underline{0}$, $\tilde{\underline{F}}^{(0)} = \underline{F}^{(0)}$ and $\tilde{\underline{W}}^{(0)} = \underline{W}^{(0)}$.

Step 5. Initialize count number, $k = 1$ and convergence tracker, $\epsilon^{(0)} = 1$.

Step 6. **while** $\epsilon^{(k-1)} > TOL$ **do**

a. $\tilde{\underline{Y}}_1^{(k)} \leftarrow \underline{Y}_1^{(k-1)} + \sigma (\tilde{\underline{F}}^{(k-1)} - \tilde{\underline{W}}^{(k-1)})$

b. $\underline{Y}_1^{(k)} \leftarrow \frac{\tilde{\underline{Y}}_1^{(k)}}{\max(1, |\tilde{\underline{Y}}_1^{(k)}|)}$ All operations in this step are element-wise.

c. $\underline{F}^{(k)} \leftarrow \underline{B} (\underline{F}^{(k-1)} - \tau \underline{Y}_1^{(k)} + \tau \alpha \underline{K}^T \underline{S})$

d. $\underline{F}^{(k)} \leftarrow \max(0, \underline{F}^{(k)})$ All operations in this step are element-wise.

e. $\tilde{\underline{Y}}_2^{(k)} \leftarrow \underline{Y}_2^{(k-1)} + \sigma \underline{D}_2 \tilde{\underline{W}}^{(k-1)}$

f. The actions in this step are element-wise and differ between algorithms

i. $\underline{X} \leftarrow \sqrt{\text{sum}(\tilde{\underline{Y}}_2^{(k)})^2}$ 1D MTGV algorithm

ii. $\underline{X} \leftarrow \sqrt{\text{sum}(\left| \text{reshape}(\left(\tilde{\underline{Y}}_2^{(k)}\right)^2, n^2, 3) \right|, 2)}$ 2D MTGV algorithm

g. The actions in this step are element-wise and differ between algorithms

i. $\underline{Y}_2^{(k)} \leftarrow \tilde{\underline{Y}}_2^{(k)} / \max\left(1, \frac{\underline{X}}{\beta}\right)$ 1D MTGV algorithm

ii. $\underline{Y}_2^{(k)} \leftarrow \tilde{\underline{Y}}_2^{(k)} / \max\left(1, \text{repmat}(\underline{X}, 3, 1) / \beta\right)$ 2D MTGV algorithm

h. $\underline{W}^{(k)} \leftarrow \underline{W}^{(k-1)} + \tau (\underline{Y}_1^{(k)} - \underline{D}_2^T \underline{Y}_2^{(k)})$

i. $\tilde{\underline{F}}^{(k)} \leftarrow 2\underline{F}^{(k)} - \underline{F}^{(k-1)}$

- j. $\tilde{W}^{(k)} \leftarrow 2W^{(k)} - W^{(k-1)}$
- k. $\epsilon^{(k)} \leftarrow \frac{\|F^{(k)} - F^{(k-1)}\|_2}{\|F^{(k-1)}\|_2}$
- l. $k \leftarrow k + 1$

end while

The condition for the convergence of the MTGV algorithm is similar to the condition for the TGV algorithm to converge, as has been discussed by Valkonen *et al.* [48]. The condition depends primarily on the choice of the two algorithmic parameters, τ and σ , which are described in the pseudocode. Although there is a relationship that they must obey, the particular choice of τ and σ is largely heuristic. The choice of $\tau = \sigma = 0.1$ was found to offer a good compromise between convergence speed and stability in our particular application. The values of τ and σ will depend slightly on how the NMR signal is scaled, determined by the largest data point in the acquired signal. To avoid this, it is best to normalize the signal to a maximum of 1. In the present work, the number of iterations required to arrive at a reasonable convergence was 10,000. The time in which this convergence was achieved for a 1D 32×1 distribution map with a 2.0 GHz Intel® Core™ i5-4590T CPU and 16.4 GB RAM, was approximately 6 s, while the speed required to achieve this convergence for a 2D 32×32 distribution map was approximately 23 s.

References

- [1] W. E. Kenyon, P. I. Day, C. Straley and J. F. Willemsen, A three-part study of NMR longitudinal relaxation properties of water-saturated sandstones, SPE Formation Evaluation, 3, September 1988.
- [2] R. L. Kleinberg and M. A. Horsfield, Transverse relaxation processes in porous sedimentary rock, J. Magn. Reson. 88 (1990), 9-19.
- [3] J. E. Snaar and H. van As, Probing water compartments and membrane permeability in plant cells by ^1H NMR relaxation measurements, Biophys. J. 63 (1992), 1654-1658.
- [4] A. Chen, D. Wu and C. S. Johnson Jr., Determination of molecular weight distributions for polymers by diffusion-ordered NMR, J. Am. Chem. Soc. 117 (1995), 7965-7970.
- [5] K. J. Packer and C. Rees, Pulsed NMR studies of restricted diffusion. I. Droplet size distributions in emulsions, J. Colloid Interface Sci. 40 (1972), 206-218.
- [6] M. Fleury, M. Romero-Samiento, Characterization of shales using T_1 - T_2 maps, J. Petrol. Sci. Eng. 137 (2016), 55-62.
- [7] D. Weber, J. Mitchell, J. McGregor, L. F. Gladden, Comparing strengths of surface interactions for reactants and solvents in porous catalysts using two-dimensional NMR relaxation correlations, J. Phys. Chem. C 113 (2009), 6610-6615.
- [8] A. E. English, K. P. Whittall, M. L. G. Joy, R. M. Henkelman, Quantitative two-dimensional time correlation relaxometry, Magn. Reson. Med. 22 (1991), 425-434.
- [9] J. Lee, C. Labadie, C. S. Springer Jr., G. S. Harbison, Two-dimensional inverse Laplace transform NMR: Altered relaxation times allow detection of exchange correlation, J. Am. Chem. Soc. 115 (1993), 7761-7764.
- [10] L. Monteilhet, J.-P. Korb, J. Mitchell, P. J. McDonald, Observation of exchange of micropore water in cement pastes by two-dimensional T_2 - T_2 nuclear magnetic resonance relaxometry, Phys. Rev. E 74 (2006), 061404.
- [11] Y. Song, L. Zielinski, S. Ryu., Two-dimensional NMR of diffusion systems, Phys. Rev. Lett. 100 (2008), 248002.
- [12] K. E. Washburn, P. T. Callaghan, Tracking pore to pore exchange using relaxation exchange spectroscopy, Phys. Rev. Lett. 97 (2006), 175502.
- [13] R. Song, Y.-Q. Song, M. Vembusubramanian, J. L. Paulsen, The robust identification of exchange from T_2 - T_2 time-domain features, J. Magn. Reson. 265 (2016), 164-171.
- [14] M. D. Hürlimann, L. Venkataramanan, Quantitative measurement of two-dimensional distribution functions of diffusion and relaxation in grossly inhomogeneous fields, J. Magn. Reson. 157 (2002), 31-42.
- [15] J.-P. Korb, N. Vorapalawut, B. Nicot, R. G. Bryant, Relation and correlation between NMR relaxation times, diffusion coefficients, and viscosity of heavy crude oils, J. Phys. Chem. C 119 (2015), 24439-24446.

- [16] Y. Zhang, B. Blümich, Spatially resolved D - T_2 correlation NMR of porous media, *J. Magn. Reson.* 242 (2014), 41-48.
- [17] P. T. Callaghan, I. Furó, Diffusion-diffusion correlation and exchange as a signature for local order and dynamics, *J. Chem. Phys.*, 120 (2004), 4032-4038.
- [18] P. L. Hubbard, K. M. McGrath, P. T. Callaghan, A study of anisotropic water self-diffusion and defects in lamellar mesophase, *Langmuir* 21 (2004), 4340-4346.
- [19] Y. Qiao, P. Galvosas, T. Adalsteinsson, M. Schönhoff, P. T. Callaghan, Diffusion exchange NMR spectroscopic study of dextran exchange through polyelectrolyte multilayer capsules, *J. Chem. Phys.* 122 (2005), 214912.
- [20] J. L. Paulsen, Y.-Q. Song, Two-dimensional diffusion time correlation experiment using a single direction gradient, *J. Magn. Reson.* 244 (2014), 6-11.
- [21] C. L. Lawson and R. J. Hanson, Solving least squares problems, SIAM, 1995.
- [22] A. N. Tikhonov, V. Y. Arsenin, Solutions of ill-posed problems, V. H. Winston and Sons, 1977.
- [23] S. J. Provencher, A constrained regularization method for inverting data represented by linear algebraic or integral equations, *Comput. Phys. Commun.* 27 (1982), 213-227.
- [24] S. J. Provencher, CONTIN: A general purpose constrained regularization program for inverting noisy linear algebraic and integral equations, *Comput. Phys. Commun.* 27 (1982), 229-242.
- [25] G. C. Borgia, R. J. S. Brown and P. Fantazzini, Uniform-penalty inversion of multiexponential decay data, *J. Magn. Reson.* 122 (1998), 65-77.
- [26] L. Ambrosone, A. Ceglie, G. Colafemmina and G. Palazzo, General methods for determining the droplet size distribution in emulsion systems, *J. Chem. Phys.* 110, (1999), 797-804.
- [27] G. Su, X. Zhou, L. Wang, Y. Wang, S. Nie, An inversion method of 2D NMR relaxation spectra in low fields based on LSQR and L-curve, *J. Magn. Reson.* 265 (2016), 146-152.
- [28] P. T. Callaghan, Translational Dynamics and Magnetic Resonance, Oxford University Press, 2011.
- [29] K. P. Whittall, A. L. MacKay, Quantitative interpretation of NMR relaxation data, *J. Magn. Reson.* 84 (1989), 134-152.
- [30] Y. Song, L. Venkataramanan and L. Burcaw, Determining the resolution of Laplace inversion spectrum, *J. Chem. Phys.* 122 (2005), 104104.
- [31] D. Benjamini and P. J. Basser, Use of marginal distributions constrained optimization (MADCO) for accelerated 2D MRI relaxometry and diffusometry, *J. Magn. Reson.* 271 (2016), 40-45.
- [32] D. Benjamini and P. J. Basser, Towards clinically feasible relaxation-diffusion correlation MRI using MADCO, *Micropor. Mesopor. Mat.* (2017) 10.1016/j.micromeso.2017.02.001.
- [33] X. Zhou, G. Su, L. Wang, S. Nie, X. Ge, The inversion of 2D NMR relaxometry data using L_1 regularization, *J. Magn. Reson.* 275 (2017), 46-54.

- [34] A. Reci, A. J. Sederman, L. F. Gladden, Obtaining sparse distributions in 2D inverse problems, *J. Magn. Reson.* 281 (2017), 188-198.
- [35] K. Bredies, K. Kunisch, T. Pock, Total generalized variation, *SIAM J. Imaging Sci.* 3 (2010), 492-526.
- [36] F. Knoll, K. Bredies, T. Pock, R. Stollberger, Second order total generalized variation (TGV) for MRI, *Magn. Reson. Med.* 65 (2011), 480-491.
- [37] M. Benning, C. Brune, M. Burger, J. Müller, Higher-order TV methods - enhancement via Bergman iteration, *J. Sci. Comput.* 54 (2013), 269-310.
- [38] M. Benning, L. F. Gladden, D. Holland, C. Schönlieb, T. Valkonen, Phase reconstruction from velocity-encoded MRI measurements - A survey of sparsity-promoting variational approaches, *J. Magn. Reson.* 238 (2014), 26-43.
- [39] S. Ono, I. Yamada, Second-order total generalized variation constraint, *IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)* (2014), 4938-4942.
- [40] B. L. Phillips, A technique for the numerical solution of certain integral equations of the first kind, *J. ACM* 9 (1962), 84-97.
- [41] S. Twomey, On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature, *J. ACM* 10 (1963), 97-101.
- [42] J. Mitchell, T. C. Chandrasekera, and L. F. Gladden, Numerical estimation of relaxation and diffusion distributions in two dimensions, *Prog. Nucl. Magn. Reson. Spectrosc.* 62 (2012), 34-50.
- [43] A. Chambolle T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, *J. Math. Imaging Vis.* 40 (2011), 120-145.
- [44] H. Celik, M. Bouhrara, D. A. Reiter, K. W. Fishbein, R. G. Spencer, Stabilization of the inverse Laplace transform of multiexponential decay through introduction of a second dimension, *J. Magn. Reson.* 236 (2013), 134-139.
- [45] L. Venkataramanan, Y.-Q. Song, M. D. Hürlimann, Solving Fredholm integrals of the first kind with tensor product structure in 2 and 2.5 dimensions, *IEEE Trans. Signal Process.* 50 (2002), 1017-1026.
- [46] D. C. Montgomery, G. C. Runger, *Applied Statistics and Probability for Engineers*, 6th Edition, Wiley, 2013.
- [47] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [48] T. Valkonen, K. Bredies, F. Knoll, Total generalized variation in diffusion tensor imaging, *SIAM J. Imaging Sci.* 6 (2013), 487-525.

Figure captions

Fig. 1. T_1 - T_2 input distributions used in simulations. The distributions in figures (a)-(g) will be labelled accordingly A-G. All peaks are lognormally distributed and the details are described in Table 1.

Fig. 2. Comparison of the performance of the different regularization techniques in reconstructing (a) the T_1 and (b) T_2 distributions for a range of smooth and discrete true distributions at SNR = 2000. Distributions A-F are shown in Fig. 1 and the metric χ is defined in Eq. (9).

Fig. 3. Comparison of the performance of the different regularization techniques in reconstructing (a) the T_1 and (b) T_2 distributions for a range of smooth and discrete true distributions at SNR = 200. Distributions A-F are shown in Fig. 1 and the metric χ is defined in Eq. (9).

Fig. 4. Comparison of the performance of the different regularization techniques in reconstructing (a) the T_1 and (b) T_2 distributions for a range of smooth and discrete true distributions at SNR = 20. Distributions A-F are shown in Fig. 1 and the metric χ is defined in Eq. (9).

Fig. 5. Variation of the best α and β for (a) 1D MTGV regularization and (b) 2D MTGV regularization in the experiments described in Figs. 2-4. The line of best fits (a) $\beta = 2 \times 10^{-5} \alpha$ and (b) $\beta = 0.3 \alpha$ are estimated by fitting a straight line with gradient 1 to the data set $\log_{10} \beta$ against $\log_{10} \alpha$. 108 points are plotted in both maps but some of them are not distinguishable because they overlap with other points.

Fig. 6. Comparison of the performance of the different regularization techniques in reconstructing (a) the T_1 and (b) T_2 distributions for a range of smooth and discrete true distributions at SNR = 20. Distributions A-F are shown in Fig. 1 and the metric χ is defined in Eq. (9). The value of β for MTGV regularization is constrained to (a) $\beta = 2 \times 10^{-5} \alpha$ for 1D MTGV and (b) $\beta = 0.3 \alpha$ for 2D MTGV regularization.

Fig. 7. Comparison of the performance of the different regularization techniques in reconstructing (a) the T_1 and (b) T_2 distributions for the true distribution G shown in Fig. 1 at three SNR: 20, 200 and 2000. The metric χ is defined in Eq. (9). The value of β for MTGV regularization is constrained to (a) $\beta = 2 \times 10^{-5} \alpha$ for 1D MTGV and (b) $\beta = 0.3 \alpha$ for 2D MTGV regularization.

Tables

Table 1: Details of the peaks of distributions A-G shown in Fig. 1. Each distribution is composed of two peaks; the peak at the lower relaxation times is referred to as peak 1 while the peak at the higher relaxation times is referred to as peak 2. The distribution in each dimension, T_1 or T_2 , is described by $\text{lognpdf}(T_i, \mu_1, \sigma_1) + \gamma \text{lognpdf}(T_i, \mu_2, \sigma_2)$, for $i = 1, 2$, where μ, σ refer to the mean and standard deviation of each peak. Distributions A-F were designed to have the same area under each peak in a logarithmic scale. Distribution G was designed such that both peaks have the same maximum amplitude when projected in each dimension.

distribution	μ_1 (s)	σ_1 (s)	μ_2 (s)	σ_2 (s)	γ
A	0.050	0.003	0.150	0.010	10
B	0.050	0.005	0.150	0.016	10
C	0.050	0.008	0.150	0.025	10
D	0.050	0.013	0.150	0.040	10
E	0.050	0.020	0.150	0.063	10
F	0.050	0.032	0.150	0.100	10
G	0.050	0.032	0.150	0.010	2

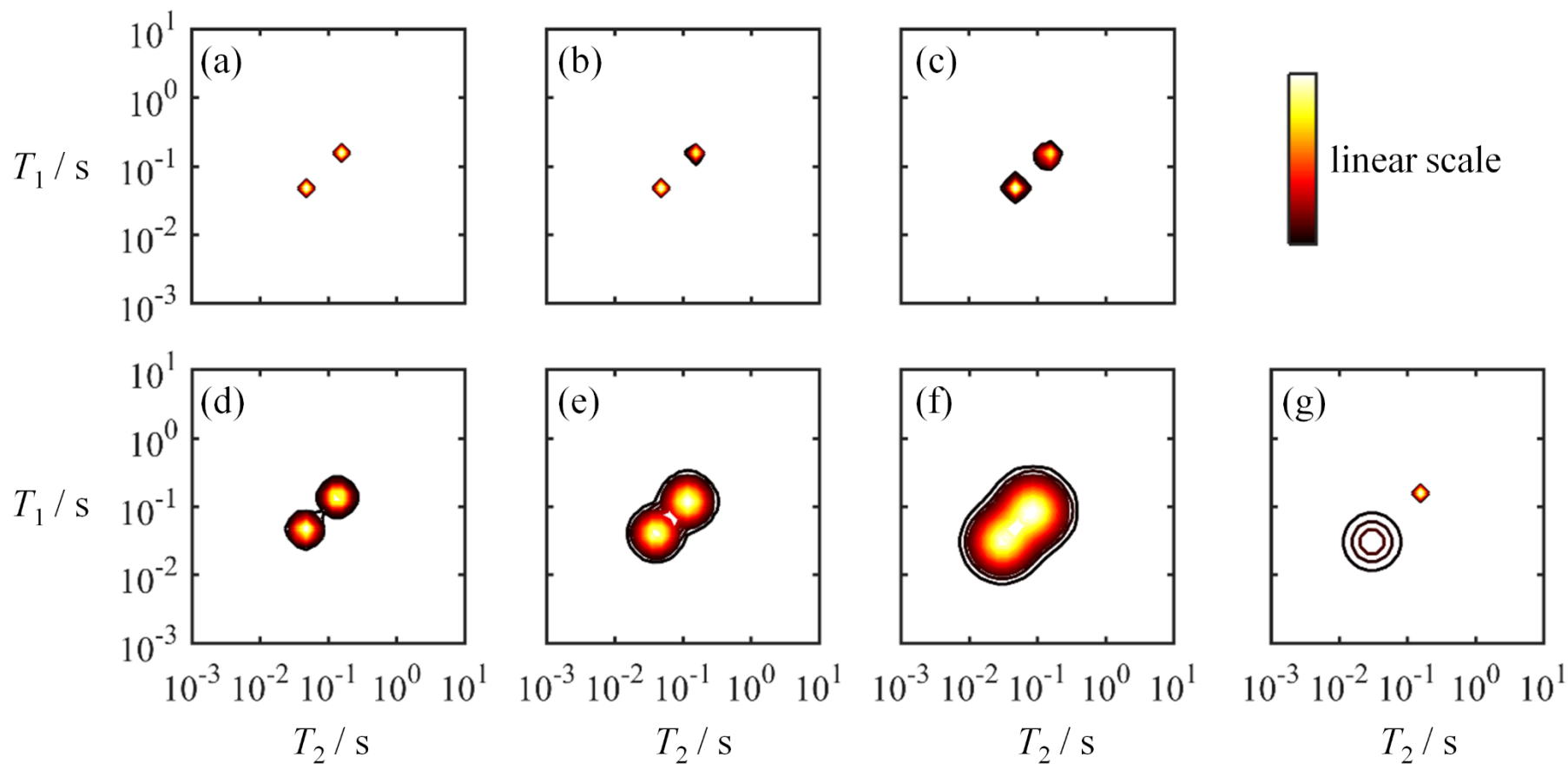


Fig. 1. T_1 - T_2 input distributions used in simulations. The distributions in figures (a)-(g) will be labelled accordingly A-G. All peaks are lognormally distributed and the details are described in Table 1.

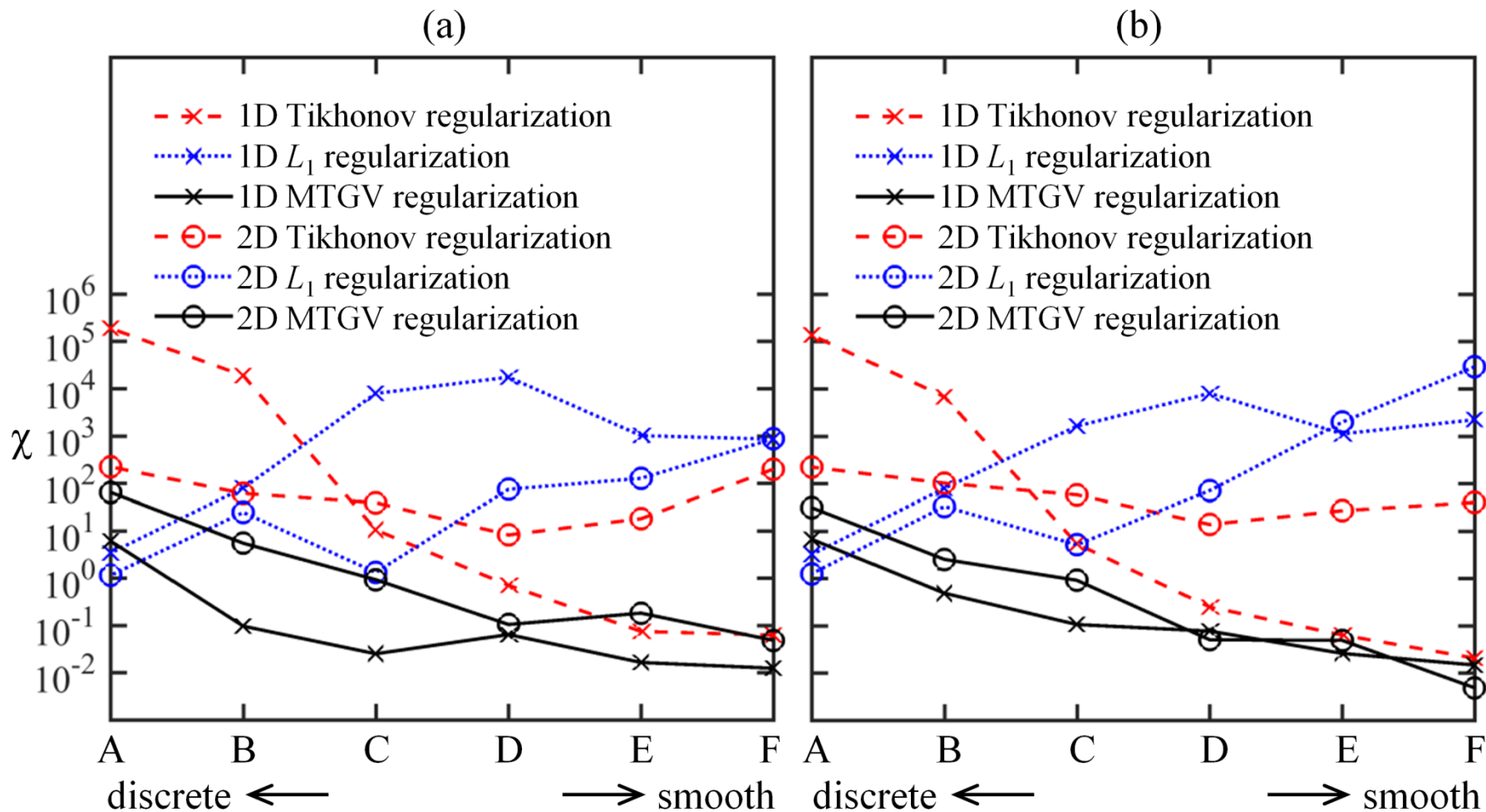


Fig. 2. Comparison of the performance of the different regularization techniques in reconstructing (a) the T_1 and (b) T_2 distributions for a range of smooth and discrete true distributions at SNR = 2000. Distributions A-F are shown in Fig. 1 and the metric χ is defined in Eq. (9).

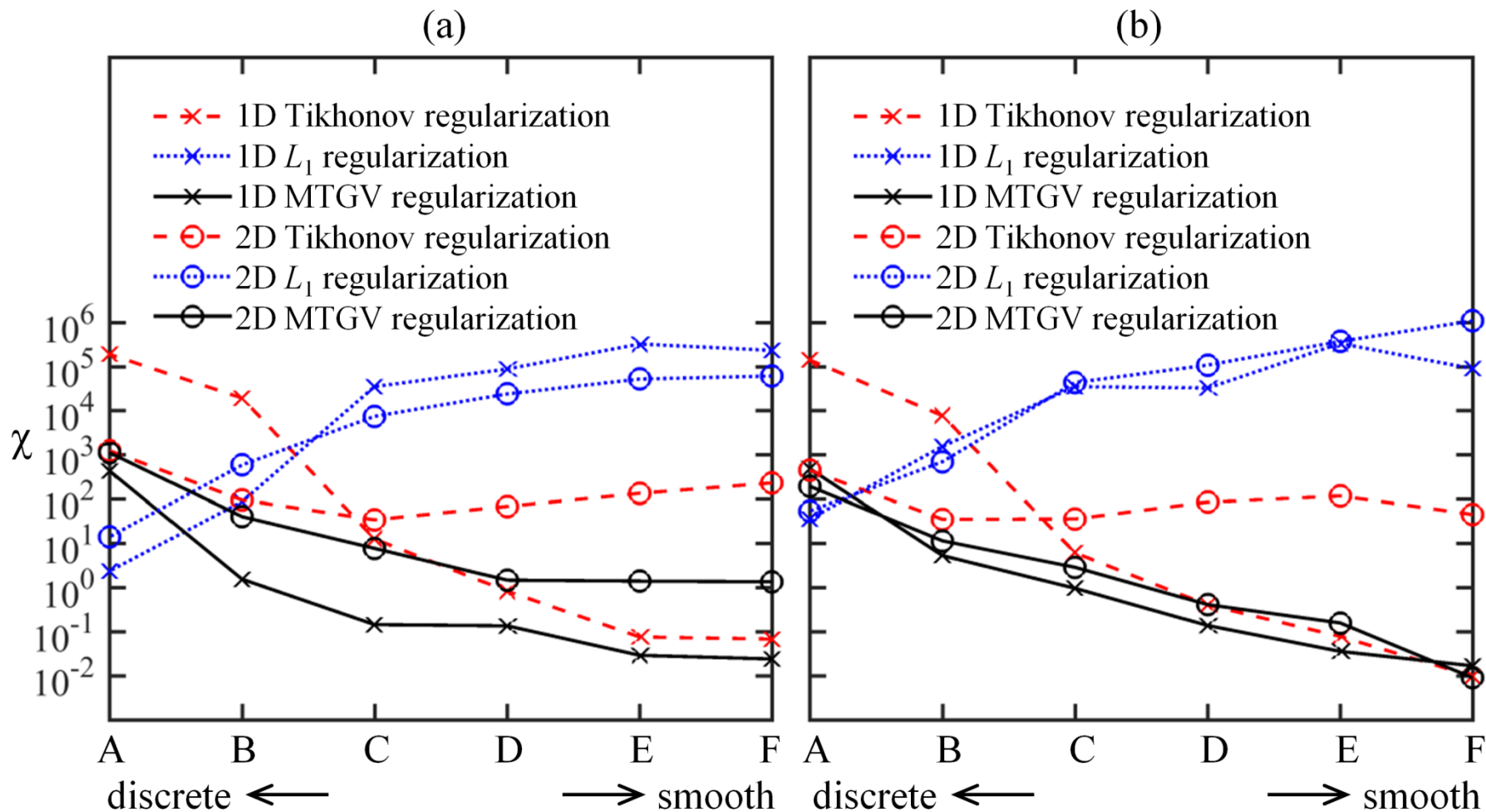


Fig. 3. Comparison of the performance of the different regularization techniques in reconstructing (a) the T_1 and (b) T_2 distributions for a range of smooth and discrete true distributions at SNR = 200. Distributions A-F are shown in Fig. 1 and the metric χ is defined in Eq. (9).

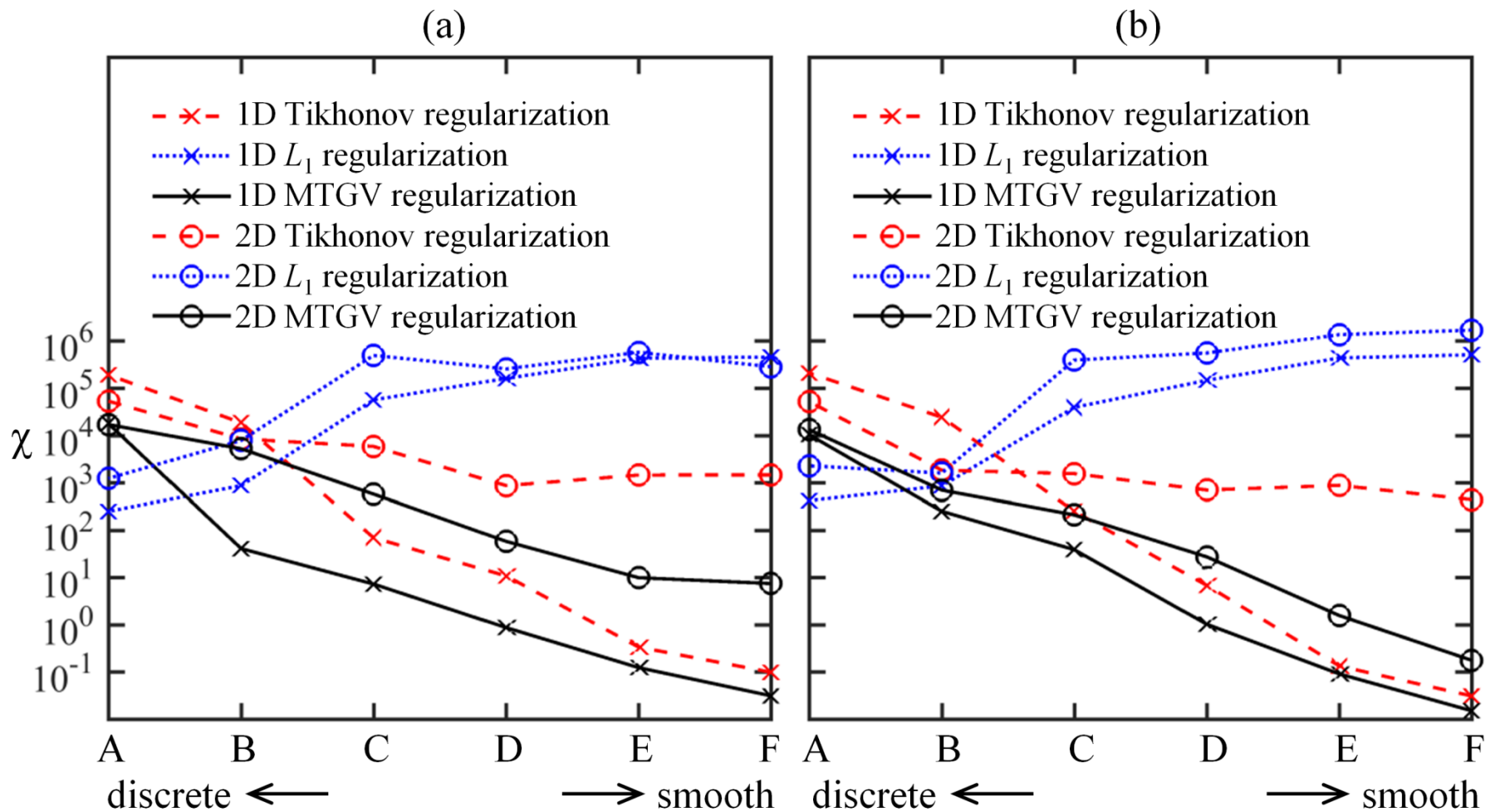


Fig. 4. Comparison of the performance of the different regularization techniques in reconstructing (a) the T_1 and (b) T_2 distributions for a range of smooth and discrete true distributions at SNR = 20. The distributions A-F are shown in Fig. 1 and the metric χ is defined in Eq. (9).

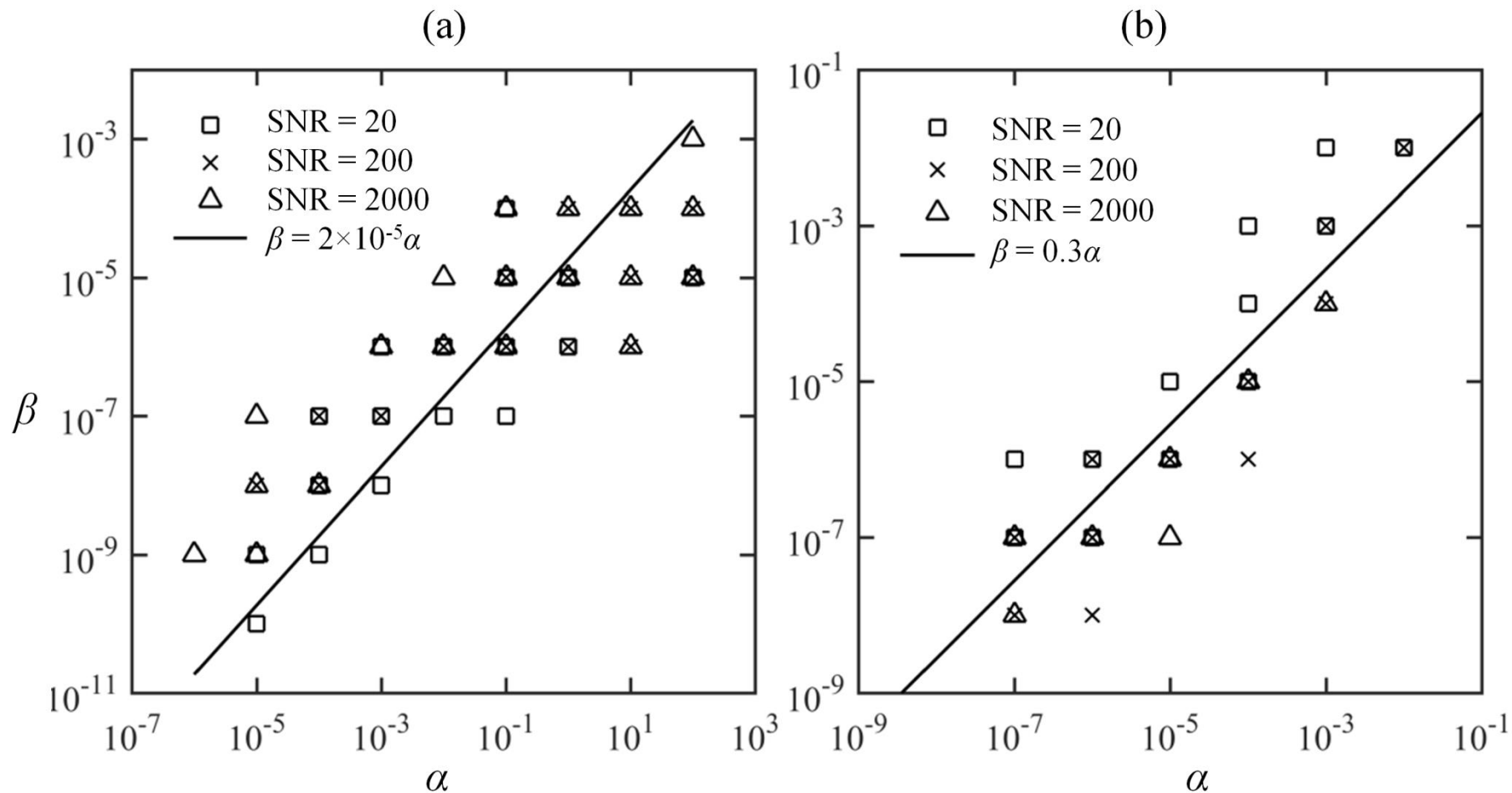


Fig. 5. Variation of the best α and β for (a) 1D MTGV regularization and (b) 2D MTGV regularization in the experiments described in Figs. 2-4. The line of best fits (a) $\beta = 2 \times 10^{-5} \alpha$ and (b) $\beta = 0.3 \alpha$ are estimated by fitting a straight line with gradient 1 to the data set $\log_{10} \beta$ against $\log_{10} \alpha$. 108 points are plotted in both maps but some of them are not distinguishable because they overlap with other points.

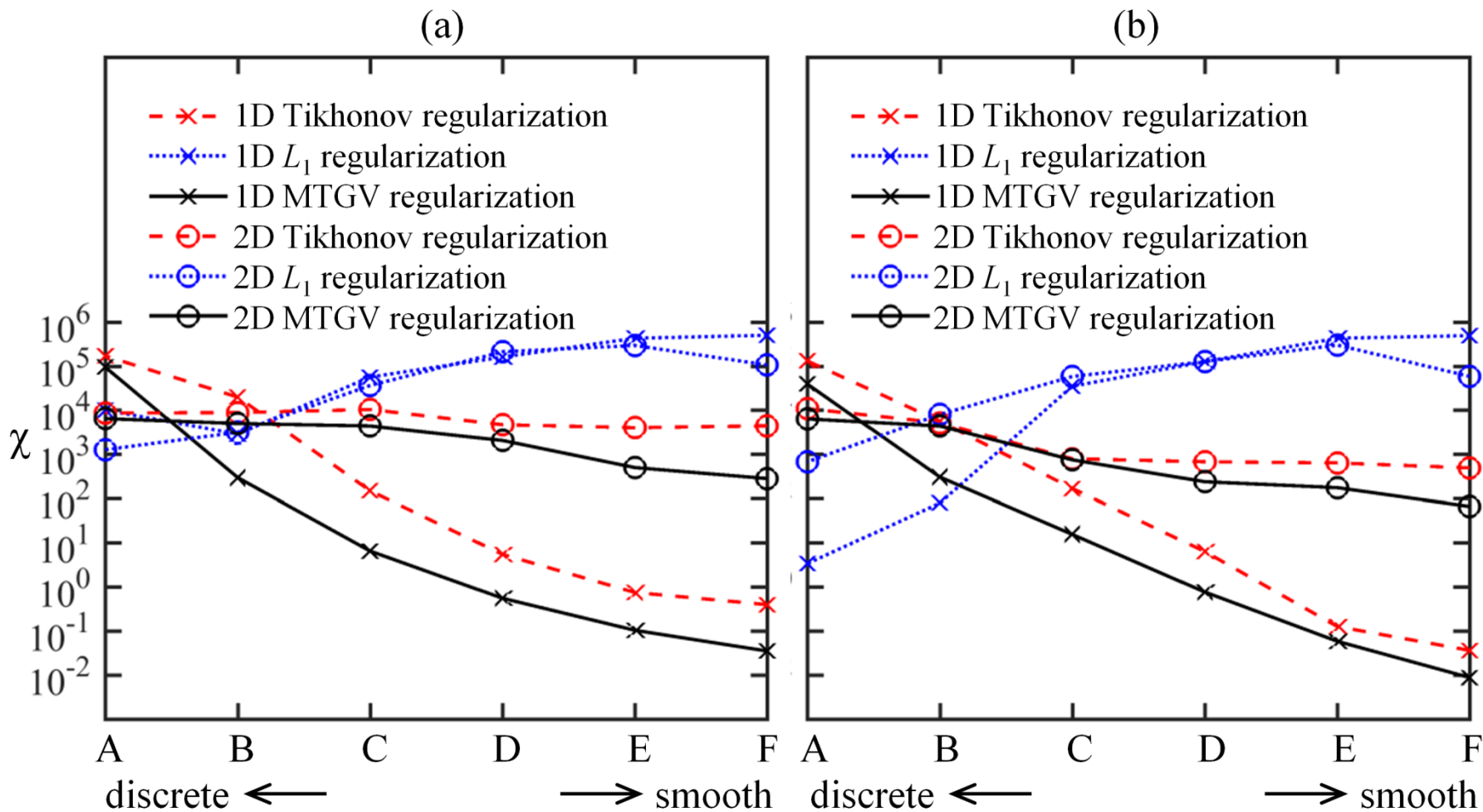


Fig. 6. Comparison of the performance of the different regularization techniques in reconstructing (a) the T_1 and (b) T_2 distributions for a range of smooth and discrete true distributions at SNR = 20. Distributions A-F are shown in Fig. 1 and the metric χ is defined in Eq. (9). The value of β for MTGV regularization is constrained to (a) $\beta = 2 \times 10^{-5} \alpha$ for 1D MTGV and (b) $\beta = 0.3 \alpha$ for 2D MTGV regularization.

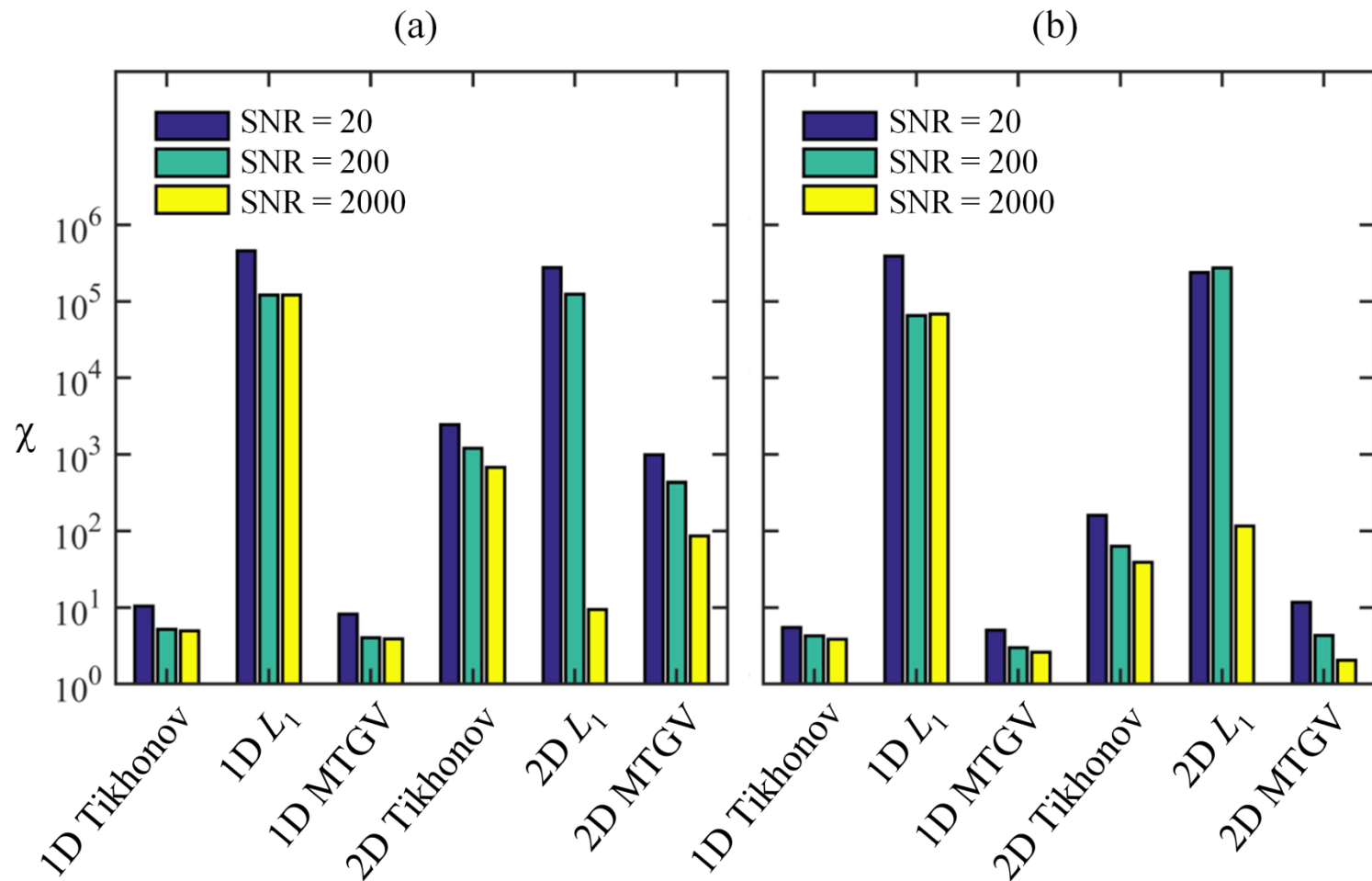


Fig. 7. Comparison of the performance of the different regularization techniques in reconstructing (a) the T_1 and (b) T_2 distributions for the true distribution G shown in Fig. 1 at three SNR: 20, 200 and 2000. The metric χ is defined in Eq. (9). The value of β for MTGV regularization is constrained to (a) $\beta = 2 \times 10^{-5} \alpha$ for 1D MTGV and (b) $\beta = 0.3 \alpha$ for 2D MTGV regularization.