

# Optimising sampling patterns for bi-exponentially decaying signals

A. Reci, M.I. Ainte, A.J. Sederman<sup>a)</sup>, M.D. Mantle, L.F. Gladden

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

<sup>a)</sup> 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

---

<sup>a)</sup> Author to whom correspondence should be addressed. Electronic mail: ajs40@cam.ac.uk.

## Abstract

A recently reported method, based on the Cramér-Rao Lower Bound theory, for optimising sampling patterns for a wide range of nuclear magnetic resonance (NMR) experiments is applied to the problem of optimising sampling patterns for bi-exponentially decaying signals. Sampling patterns are optimised by minimizing the percentage error in estimating the most difficult to estimate parameter of the bi-exponential model, termed the objective function. The predictions of the method are demonstrated in application to pulsed field gradient NMR data recorded for the two-component diffusion of a binary mixture of methane/ethane in a zeolite. It is shown that the proposed method identifies an optimal sampling pattern with the predicted objective function being within 10% of that calculated from the experiment dataset. The method is used to advise on the number of sampled points and the noise level needed to resolve two-component systems characterised by a range of ratios of populations and diffusion coefficients. It is subsequently illustrated how the method can be used to reduce the experiment acquisition time while still being able to resolve a given two-component system.

Keywords: sampling pattern, bi-exponential model, Cramér-Rao Lower Bound theory, PFG NMR diffusion

# 1. Introduction

Bi-exponentially decaying signals are common in many areas of science and engineering, such as in photoluminescence [1, 2] and reaction kinetics [3, 4]. In NMR applications, bi-exponential models have been assumed to be a good approximation for  $T_2$  and  $T_{1\rho}$  signal decays and signal attenuations recorded during diffusion measurements [5-9]. The fitting of bi-exponential decays is a long-standing challenge and remains a subject of debate [10-14] because of the relatively large uncertainty related to the estimated parameters of the model.

In a recent work [15], a systematic, statistical approach, based on the Cramér-Rao Lower Bound (CRLB) theory [16], was described for optimising sampling patterns in a wide range of NMR experiments. The findings of the CRLB theory were validated against pulsed field gradient (PFG) NMR diffusion experimental data of an emulsion system, characterised by a lognormal distribution of droplet sizes. In this work, the same method is applied to the problem of optimising the sampling pattern for the most accurate estimation of the parameters of a bi-exponential model. The findings of the CRLB theory are validated against PFG NMR diffusion experimental data acquired for a binary gaseous mixture of methane/ethane adsorbed in a zeolite. Zeolites are increasingly being used to separate gaseous hydrocarbon mixtures in the gas processing industry [17, 18]. The CRLB theory is subsequently used to determine the conditions under which a system characterised by a bi-exponentially decaying signal is resolvable.

The paper is structured as follows. Section 2 outlines the application of the CRLB theory to a bi-exponentially decaying signal. The experimental methods are described in Section 3. Section 4 presents the results and discussion.

## 2. Theory

A noisy, bi-exponentially decaying signal,  $\hat{y}_i$ , is described as:

$$\hat{y}_i = y_i + \epsilon_i = w_1 \exp(-b_i D_1) + w_2 \exp(-b_i D_2) + \epsilon_i, \quad (1)$$

where  $y_i$  is the noise-free signal;  $b_i$  represents the sampling pattern;  $w_1$  and  $w_2$  are the populations of the two components;  $D_1$  and  $D_2$  are the decay rate constants, where for

convenience  $D_1 < D_2$ ; and  $\epsilon_i$  represents the noise related to each signal data point. It is assumed that  $\epsilon_i$  are random and independently normally distributed with zero mean and variance  $\sigma^2$  (Gaussian noise). The noise variance depends on the scaling of the signal; in this work, the stated value of  $\sigma$  corresponds to a scaled signal of maximum 1, which corresponds to  $w_1 + w_2 = 1$ . The noise-free signal,  $y_i$ , depends on 4 parameters,  $\theta_l$  ( $l = 1, 2, 3, 4$ ):  $w_1$ ,  $w_2$ ,  $D_1$  and  $D_2$ . In the presence of noise, the estimate of these parameters becomes uncertain and there is an uncertainty associated with the estimate of each parameter. A common characteristic of this uncertainty is the standard deviation,  $\text{std}(\theta_l)$ , a lower limit of which is given by the Cramér-Rao Lower Bound (CRLB) theory:

$$\text{std}(\theta_l) \geq \sqrt{(\underline{F}^{-1})_{l,l}}, \quad (2)$$

where  $\underline{F}$  is the Fisher Information Matrix [19]. Given the assumption that the noise random variables are independently normally distributed with zero mean and variance  $\sigma^2$ , the Fisher Information Matrix is simplified to [20]:

$$F_{l_1, l_2} = \frac{1}{\sigma^2} \left( \sum_i \left( \frac{\partial y_i}{\partial \theta_{l_1}} \frac{\partial y_i}{\partial \theta_{l_2}} \right) \right). \quad (3)$$

The elements of  $\underline{F}$ , as described by Eq. (3), are obtained through partial differentiation of the noise-free signal defined in Eq. (1):

$$\frac{\partial y_i}{\partial w_1} = \exp(-b_i D_1), \quad (4a)$$

$$\frac{\partial y_i}{\partial w_2} = \exp(-b_i D_2), \quad (4b)$$

$$\frac{\partial y_i}{\partial D_1} = -w_1 b_i \exp(-b_i D_1), \quad (4c)$$

$$\frac{\partial y_i}{\partial D_2} = -w_2 b_i \exp(-b_i D_2). \quad (4d)$$

It follows that the uncertainty associated with the estimate of each parameter depends on  $\sigma$ ,  $w_1$ ,  $w_2$ ,  $D_1$ ,  $D_2$  and, most importantly, on the sampling pattern,  $b_i$ . Given the noise standard deviation,  $\sigma$ , and an experimental system (characterised by  $w_1$ ,  $w_2$ ,  $D_1$  and  $D_2$ ), the error associated with the estimate of each parameter depends solely on the sampling pattern,  $b_i$ . Therefore, by optimizing the sampling pattern, it is possible to decrease to a minimum the uncertainty associated with the estimate of a given parameter, hence, making the estimate of that parameter as accurate as possible.

The sampling pattern that minimises the uncertainty related to one of the parameters does not necessarily minimise the uncertainty related to the other parameters. Therefore, an objective function to be minimised needs to be defined. The method presented in this work can be easily adapted to different objective functions. Here, it is chosen to minimise:

$$\chi = \max_l [\text{std}(\theta_l)/\theta_l] \times 100\% , \quad (5)$$

which practically means minimising the percentage error of the most difficult to estimate parameter of the model.

**1.**

**2.**

## **3. Materials and methods**

### *3.1. Sample preparation and data acquisition*

A PFG NMR diffusion experiment was conducted on a binary gaseous mixture adsorbed on a porous material. The gases used were methane (BOC, >99% purity) and ethane (Air Liquide, >99.99%). The porous material used was a microporous  $\beta$ -zeolite (BEA), supplied by Tosoh Corporation. Zeolite crystals of mean crystallite size of 15  $\mu\text{m}$  and Si:Al ratio 124 were used. The crystals were packed into a cylindrical column of height 10 mm and diameter 5 mm. The adsorption of the gases in the zeolite was carried out on a Schlenk vacuum line; further details are found elsewhere [21]. Each individual gas adsorbed in the  $\beta$ -zeolite, methane or ethane, is characterised by a diffusion attenuation that is well described by a single exponential decay [21].

The PFG NMR experiment was performed on a Bruker DMX300 spectrometer, operating at a resonant frequency of 300.13 MHz for  $^1\text{H}$  observation, at a temperature of  $20 \pm 0.5$  °C. The maximum gradient amplitude available was 1176 G  $\text{cm}^{-1}$  and the radiofrequency (r.f.) coil had a diameter of 5 mm. The 13-interval bipolar gradient stimulated echo PFG NMR pulse sequence [22], illustrated in Fig. 1(a), was used for

diffusion measurements. The relevant NMR acquisition parameters used, annotated in Fig. 1(a), were  $T = 100$  ms,  $\delta = 0.25$  ms and  $\tau = 0.46$  ms. The echo time,  $\tau$ , was minimised to increase the signal-to-noise ratio and to reduce relaxation weighting. The small remaining relaxation weighting does not affect the following analysis. 248 gradient strength values,  $g_i$ , spaced equidistantly in the range 1.5-50 G cm<sup>-1</sup>, were used. This dataset was used in two ways. First, the full dataset provided an estimate of the ground truth values of the populations and the diffusion coefficients for the methane/ethane mixture in the zeolite. Second, sub-sets of the acquired data points, which form sampling patterns, were used to test the performance of CRLB method in being able to predict the correct objective function, as defined in Eq. (5). The diffusion data cannot be resolved spectroscopically and, therefore, the signal attenuation data reported in Fig. 1(b) include the attenuating signal for both chemical species present. The attenuation is well described by the bi-exponential model described in Eq. (1), with  $b_i = \gamma^2 g_i^2 \delta^2 (4T + 6\tau - 2/3\delta)$  [22], where  $\gamma$  is the gyromagnetic ratio of <sup>1</sup>H nuclei. The estimated model parameters, which are used as ground truth for the following analysis, are:  $w_1 = 0.199 \pm 0.001$  (ethane population),  $D_1 = (9.24 \pm 0.07) \times 10^{-9}$  m<sup>2</sup> s<sup>-1</sup> (ethane diffusion coefficient),  $w_2 = 0.801 \pm 0.001$  (methane population) and,  $D_2 = (1.42 \pm 0.01) \times 10^{-7}$  m<sup>2</sup> s<sup>-1</sup> (methane diffusion coefficient). The quoted uncertainty corresponds to the standard deviation of the bi-exponential fit. The residuals of the fit presented in Fig. 1(b) do not show any systematic deviation and correspond to a noise standard deviation of  $\sigma = 0.0032$ . The uncertainty related to each of the parameters is small because of the low  $\sigma$  and the large number of points in the decay.

### 3.2. Generation of sampling patterns

A sampling pattern for the acquisition of the PFG NMR data consists of the choice of the sampling points,  $b_i$ , at which the attenuation of the NMR signal is acquired. This involves choosing the number of data points,  $n$ , and their respective values. For a given number of points,  $n$ , the optimal choice of  $b_i$  would require setting the partial derivatives of the expression in Eq. (5) with respect to each individual  $b_i$  to zero and solving the set of resulting  $n$  simultaneous non-linear equations; this is computationally intractable. As a result, this paper focuses on finding the optimal sampling pattern for the class of linear sampling patterns, defined as:

$$b_i = b_1 + (b_n - b_1) \left( \frac{i-1}{n-1} \right)^r, \quad (6)$$

where  $b_1$  and  $b_n$  are the first and last sampling points and  $r$  determines the density distribution of the sampling points. The class of logarithmic sampling patterns [15] was also investigated, but the optimised logarithmic sampling pattern had the same value of  $\chi$  as the optimised linear sampling pattern; this is consistent with the work of Reci *et al.* [15]. Therefore, only the class of linear sampling patterns is considered here.

Sampling a larger number of points,  $n$ , leads to a more accurate estimate of the model parameters. Therefore, the comparison between the predictions of the CRLB theory and experimental results is performed with respect to sampling patterns of a particular number of sampling points,  $n = 32$ . Other values of  $n$  could have been chosen, but  $n = 32$  is a commonly used experimental value. Using the smallest possible value for  $b_1$  is supported by the results of other works [23, 24]. Therefore, the smallest experimentally used value,  $b_1 = 4.9 \times 10^5 \text{ m}^{-2} \text{ s}$ , was used when validating the predictions of the CRLB theory against experimental data. Having fixed  $n$  and  $b_1$ , optimising the linear sampling pattern consists of optimising  $b_n$  and  $r$ .

The comparison between the predictions of the CRLB theory and experimental results was performed as follows. Using a range of  $b_n$  from  $5.4 \times 10^7$  to  $4.8 \times 10^8 \text{ m}^{-2} \text{ s}$  and a range of  $r$  from 0.6 to 10,  $10^4$  linear sampling patterns were constructed using Eq. (6). For each of these constructed sampling patterns, the CRLB theory was used to predict the objective function,  $\chi$ , defined in Eq. (5). This was compared to the objective function obtained from the experimental data which was calculated as follows. The closest experimental 32  $b_i$  points (out of the total 248 experimental sampling points acquired) to the numerically constructed sampling patterns were selected. The experimental attenuation data corresponding to these 32 points was fitted to a bi-exponential decay and confidence intervals for each parameter,  $\theta_l$ , were extracted. The standard deviations in the estimation of each parameter,  $\text{std}(\theta_l)$ , were then calculated from the confidence intervals using a  $z$ -test and these values were used to calculate the experimental objective function,  $\chi$ .

3.

## 4. Results and discussion

In Section 4.1, the predictions of the CRLB theory (described in Section 2) are validated against the experimental data acquired for the methane/ethane system, using the procedure outlined in Section 3.2. The CRLB theory is then used in Section 4.2 to advise on which bi-exponential systems, characterised by  $w_1$  and  $D_2/D_1$ , can be resolved and under which conditions they can be resolved, in terms of the noise standard deviation,  $\sigma$ , and the number of sampled points,  $n$ .

### 4.1. Validation of the CRLB theory against experimental data

The comparison of the objective function,  $\chi$ , as predicted by the CRLB theory and as calculated from the experimental data is shown in Fig. 2. Small values of  $\chi$  indicate a good sampling pattern. Fig. 2(a) shows the contour map of  $\chi$  as a function of the linear sampling pattern parameters  $b_n D_1$  and  $r$ , as predicted by the CRLB theory. Fig. 2(b) shows the corresponding contour map obtained from the experimental data. There is very good agreement between the CRLB and experimental contour maps. On average, there is < 10% difference between the corresponding  $\chi$  values of the two maps, thereby validating the application of the CRLB theory to the optimization of sampling patterns for bi-exponentially decaying signals.

For a given value of  $r$ , for example  $r = 1$  (corresponding to equidistant points), the value of the objective function is large at small  $b_n D_1$  because late decay points are not sampled and, therefore, the uncertainty related to estimating the slower diffusing component (ethane) is large. The value of the objective function is also large at very large  $b_n D_1$  because early decay points are not sampled properly. Therefore, the uncertainty related to estimating the faster diffusing component (methane) is large. As a result, for any given  $r$ , there is an optimum  $b_n D_1$  which minimises the objective function. For a given  $b_n D_1$ , for example  $b_n D_1 = 2.5$ , the value of the objective function is large at small  $r$  because early decay points are sampled sparsely. Therefore, the uncertainty related to estimating the faster diffusing component (methane) is large. The value of the objective



function is also large at very large  $r$  because late decay points are sampled sparsely. Therefore, the uncertainty related to estimating the slower diffusing component (ethane) is large. As a result, for any given  $b_n D_1$ , there is an optimum  $r$  which minimises the objective function. In summary, there is an optimum combination  $(b_n D_1, r)$  at which the objective function obtains a minimum, which corresponds to the best linear sampling pattern. The best linear sampling pattern for 32 sampled points is predicted by the CRLB theory at  $r = 1.07$  and  $b_n D_1 = 2.34$ , at which point  $\chi = 1.9\%$ . Therefore, using the best linear sampling pattern, the maximum percentage error for any of the estimated parameters is 1.9%.

For completeness, the objective function of a commonly used sampling pattern with 32 equidistantly sampled points ( $r = 1$ ), with the largest sampled point being  $b_n D_1 = 5$  (typical rule of thumb), is  $\chi = 3.1\%$ . This value of  $\chi$  is >50% larger than the value achieved using the optimised linear sampling pattern ( $\chi = 1.9\%$ ). The calculation was performed using the CRLB theory for the two-component system under investigation and assuming  $\sigma = 0.0032$ .

The use of the proposed method for optimising the sampling pattern relies on a prior knowledge about the diffusion coefficients and the populations. In this work, this knowledge was obtained by sampling a large number of points (248) at a low noise standard deviation,  $\sigma$  (0.0032). In practice, the acquisition of such a large number of points defeats the purpose of optimising the sampling pattern. It is therefore advised to sample a small number of points at a high noise standard deviation (through less signal averaging). A fit to these experimental data would provide a crude prior knowledge about the diffusion coefficients and the populations, which could then be used to optimise the sampling pattern using the proposed method. To answer how crude the prior knowledge needs to be, a sensitivity analysis similar to the analysis described by Reci *et al.* [15] can be performed.

#### 4.2. Resolution limit of bi-exponential decays

In Section 4.1, the CRLB theory predictions were validated against an experimentally acquired dataset for the methane/ethane gas mixture in a zeolite; this is a system exhibiting two-component diffusion characterised by  $D_2/D_1 = 15$  (ratio of the largest to the smallest diffusion coefficients) and  $w_1 = 0.20$  (population of slowest diffusing

component), as stated in Section 3.1. The aim of this section is to use the CRLB theory to investigate which systems characterised by bi-exponential signal attenuation data, characterised by  $w_1$  and  $D_2/D_1$ , can be resolved and under which conditions they can be resolved, in terms of the noise standard deviation,  $\sigma$ , and the number of sampled points,  $n$ .

This question is investigated as follows. For a given experimental system, characterized by  $w_1$  and  $D_2/D_1$ , and a noise standard deviation,  $\sigma$ , a linear sampling pattern with a given number of points,  $n$ , and  $b_1 D_1 = 0$  is optimised in terms of  $r$  and  $b_n D_1$ . This is performed using a similar method to the method used to produce Fig. 2(a). For the optimised linear sampling pattern, the value of  $\chi$  is recorded. This is repeated for 100 values of  $w_1$  in the range 0.0-1.0 (non-inclusive) and 100 values of  $D_2/D_1$  in the range 1-1000 (non-inclusive at the lower limit). A contour map of the optimal  $\chi$  as a function of  $w_1$  and  $D_2/D_1$ , for the chosen noise standard deviation,  $\sigma$ , and sampling number of points,  $n$ , is then constructed. At this point, the resolution limit needs to be defined; in this work the resolution limit is taken to be  $\chi < 30\%$ ; i.e., if the percentage error in the estimate of all the parameters is  $< 30\%$ , the bi-exponential decay is defined as being resolvable. It is noted that a different resolution limit definition is easily implemented. With the definition of the resolution limit, the contour map is divided into two regions: resolvable and non-resolvable, for the chosen values of  $\sigma$  and  $n$ . The whole procedure is performed for  $\sigma = 0.001, 0.005, 0.02, 0.05$  and  $n = 16, 32, 64$ . The results are presented in Fig. 3.

The first observation from Fig. 3 is that the contour maps are not symmetric about  $w_1 = 0.5$ . The asymmetry suggests that it is easier to resolve bi-exponential decays when the slowest diffusing component has the largest population than when the slowest diffusing component has the smallest population. This is explained by the fact that the most information about the slow diffusing component is contained at the high  $b_i$  points. However, at high  $b_i$  points, the signal becomes comparable to the noise. Therefore, a high population of the slow diffusing component is needed if it is to be resolved. Another qualitative observation from Fig. 3 is that to resolve components of similar diffusion coefficients,  $D_2/D_1 \sim 1$ , many points at a high signal-to-noise ratio need to be acquired.

The maps shown in Fig. 3 can be used quantitatively in many ways. For example, the methane/ethane gaseous mixture investigated in this paper ( $D_2/D_1 = 15$ ,  $w_1 = 0.20$ ) was easily resolved ( $\chi = 1.9\%$ ) using  $n = 32$  points because the noise standard deviation was relatively low ( $\sigma = 0.0032$ ). According to Fig. 3(b), the limit of resolution of these components for  $n = 32$  points is at  $\sigma \sim 0.05$ . Therefore, this analysis shows that the PFG NMR data acquisition time could be accelerated by a factor of  $\sim 4$  (i.e., employing less signal averaging), while still keeping the components resolvable (assuming acquisition time  $\propto \sqrt{1/\sigma}$  [25]). Accelerating the PFG NMR data acquisition time could be important if temporally changing phenomena are being investigated.

## 5. Conclusions

A recently reported method, based on the CRLB theory, was used to design optimal sampling patterns for bi-exponentially decaying signals. The optimal sampling pattern was defined as the sampling pattern that minimizes the percentage error in estimating the most difficult to estimate parameter of the bi-exponential model, termed the objective function. The capability of the method to predict the optimal sampling pattern is validated with respect to PFG NMR diffusion data of a binary gaseous mixture of methane/ethane adsorbed in a zeolite. The predictions of the method about the variation of the objective function with the sampling pattern were within 10% of the value of the objective function calculated from the experimental dataset. The method was subsequently used to determine under which conditions systems characterised by a bi-exponentially decaying signal can be resolved. An example was given about how the method can be used to reduce the experiment acquisition time, while still being able to resolve a two-component system.

## Acknowledgements

AR acknowledges Gates Trust Cambridge for financial support. MDM and MA wish to thank Johnson Matthey PLC and the EPSRC for funding.



## References

- [1] Gokus T, Cognet L, Duque JG, Pasquali M, Hartschuh A, Lounis B. Mono- and biexponential luminescence decays of individual single-walled carbon tubes. *J Phys Chem* 2010;114:14025-14028.
- [2] Coppens P, Sokolow J, Trzop E, Makal A, Chen Y. On the biexponential decay of the photoluminescence of the two crystallographically-independent molecules in crystals of [Cu(I)(phen)(PPh<sub>3</sub>)<sub>2</sub>][BF<sub>4</sub>]. *J Phys Chem Lett* 2013;4:579-582.
- [3] Harding D, Ford MS, Walsh TR, Mackenzie SR. Dramatic size effects and evidence of structural isomers in the reactions of rhodium clusters, Rh<sub>n</sub><sup>±</sup>, with nitrous oxide. *Phys Chem Chem Phys* 2007;9:2130-2136.
- [4] Mondal P, Meuwly M. Solvent composition drives the rebinding kinetics of nitric oxide to microperoxidase. *Sci Rep* 2018;8:5281.
- [5] Niendorf T, Dijkhuizen RM, Norris DG, van Lookeren Campagne M, Nicolay K. Biexponential diffusion attenuation in various states of brain tissue: Implications for diffusion-weighted imaging. *Magn Reson Med* 1996;36:847-857.
- [6] Mulkern R, Zengingonul HP, Robertson RL, Bogner P, Zou KH, Gudbjartsson H, Guttman CRG, Holtzman D, Kyriakos W, Jolesz FA, Maier SE. Multi-component apparent diffusion coefficients in human brain: Relationship to spin-lattice relaxation. *Magn Reson Med* 2000;44:292-300.
- [7] Clark CA, Le Bihan D. Water diffusion compartmentation and anisotropy at high b values in the human brain. *Magn Reson Med* 2000;44:852-859.
- [8] Ababneh Z, Beloeil H, Berde CB, Gambarota G, Maier SE, Mulkern RV. Biexponential parametrization of diffusion and T<sub>2</sub> relaxation decay curves in a rat muscle edema model: Decay curve components and water compartments. *Magn Reson Med* 2005;54:524-531.
- [9] Sharafi A, Chang G, Regatte RR. Bi-component T<sub>1ρ</sub> and T<sub>2</sub> relaxation mapping of skeletal muscle in-vivo. *Sci Rep* 2017;7:14115.
- [10] Nilsson M, Connell MA, Davis AL, Morris GA. Biexponential fitting of diffusion-ordered NMR data: Practicalities and limitations. *Anal Chem* 2006;78:3040-3045.

- [11] Shazeeb MS, Sotak CH. Limitations in biexponential fitting of NMR inversion-recovery curves. *J Magn Reson* 2017;276:14-21.
- [12] Mulkern RV, Balasubramanian M, Maier SE. On the perils of multiexponential fitting of diffusion MR data. *J Magn Reson Imaging* 2016;45:1545-1547.
- [13] Pagès G, Bonny A, Gilard V, Malet-Martino M. Pulsed field gradient NMR with sigmoid shape gradient sampling to produce more detailed diffusion ordered spectroscopy maps of real complex mixtures: Examples with medicine analysis. *Anal Chem* 2016;88:3304-3309.
- [14] Bai R, Benjamini D, Cheng J, Basser PJ. Fast, accurate 2D-MR relaxation exchange spectroscopy (REXSY): Beyond compressed sensing. *J Chem Phys* 2016;145:154202.
- [15] Reci A, Sederman AJ, Gladden LF. Optimising magnetic resonance sampling patterns for parametric characterisation. *J Magn Reson* 2018;294:35-43.
- [16] Cramér H. *Mathematical Methods of Statistics*. New Jersey: Princeton University Press; 1963.
- [17] Lauerer A, Binder T, Chmelik C, Miersemann E, Haase J, Ruthven DM, Kärger J. Uphill diffusion and overshooting in the adsorption of binary mixtures in nanoporous solids. *Nat Comm* 2015;6:7697.
- [18] Yang Y, Burke N, Ali S, Huang S, Lim S, Zhu Y. Experimental studies of hydrocarbon separation on zeolites, activated carbons and MOFs for applications in natural gas processing. *RSC Adv* 2017;7:12629-12638.
- [19] Fisher RA. *The Design of Experiments*. Edinburgh: Oliver and Boyd; 1953.
- [20] Norton JP. *An Introduction to Identification*. London: Academic Press; 1986.
- [21] Ainte MI, *NMR techniques for measuring transport phenomena in microporous materials [PhD thesis]*, Cambridge: University of Cambridge; 2017.
- [22] Cotts RM, Hoch MJR, Sun T, Marker JT. Pulsed field gradient stimulated echo methods for improved NMR diffusion measurements in heterogeneous systems. *J Magn Reson* 1989;83:252-266.

- [23] Weiss GH, Gupta RK, Ferretti JA, Becker ED. The choice of optimal parameters for measurement of spin-lattice relaxation times. I. Mathematical formulation. *J Magn Reson* 1980;37:369-379.
- [24] Jones JA, Hodgkinson P, Barker AL, Hore PJ. Optimal sampling strategies for the measurement of spin-spin relaxation times. *J Magn Reson Series B* 1996;113:25-34.
- [25] Callaghan PT. *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford: Oxford University Press; 1993.

## Figure captions

Fig. 1. (a) Schematic of the 13-interval bipolar gradient stimulated echo PFG NMR pulse sequence used in this work. (b) Experimental PFG NMR diffusion data of a binary gaseous mixture of methane/ethane in zeolite BEA, and the corresponding bi-exponential fit. The sampling points,  $b_i$ , have been non-dimensionalised by the diffusion coefficient of the slowest diffusing component, ethane,  $D_1 = 9.24 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$ . Embedded in the figure are the residuals of the fit.

Fig. 2. (a) Contour map of the variation of the objective function,  $\chi$ , with respect to the parameters of the family of linear sampling patterns with  $n = 32$ , as predicted by the CRLB theory. (b) Contour map of the variation of the objective function,  $\chi$ , with respect to the parameters of the class of linear sampling patterns with  $n = 32$ , obtained from the experimental data shown in Fig. 1. The best linear sampling pattern in (a) is obtained for  $r = 1.07$  and  $b_n D_1 = 2.34$ , and is identified as +, at which point  $\chi = 1.9\%$ .

Fig. 3. Resolution limit of bi-exponential decays, defined as  $\chi < 30\%$ , for a range of experimental systems characterised by  $w_1$  and  $D_2/D_1$  and noise standard deviation,  $\sigma$ , for (a)  $n = 16$ , (b)  $n = 32$ , and (c)  $n = 64$  sampling points. The arrows in the map point towards the resolvable region.



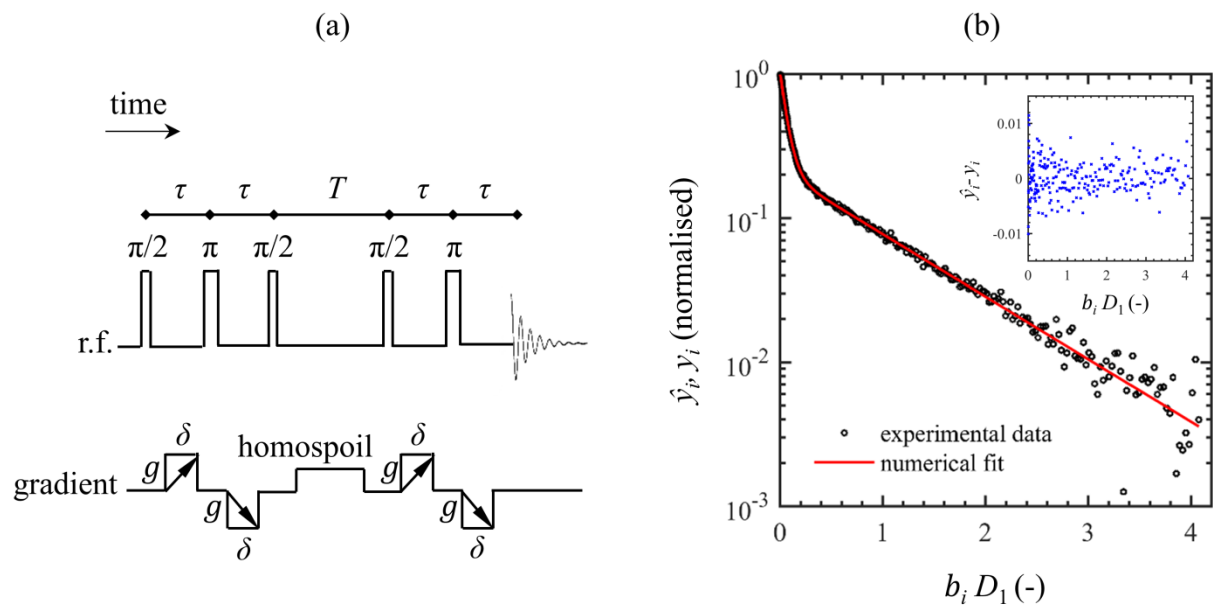


Figure 1

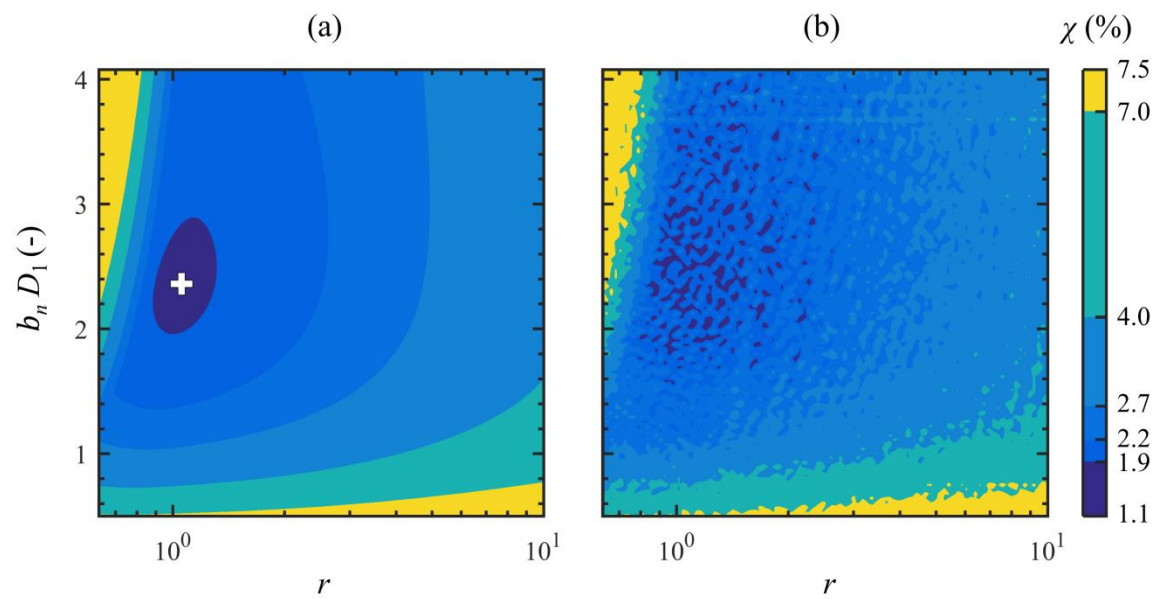


Figure 2

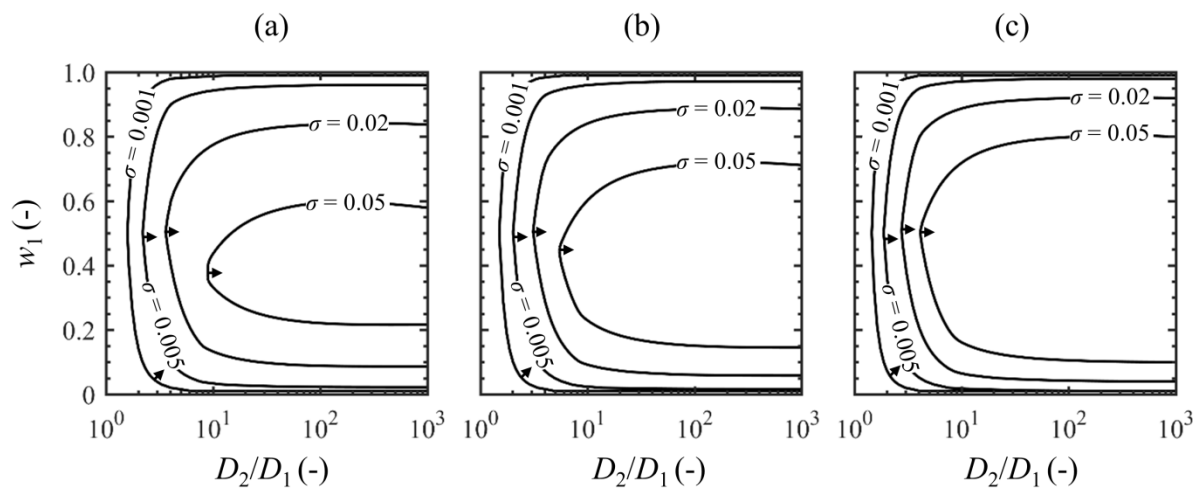


Figure 3