



Accurate measurements of self-diffusion coefficients with benchtop NMR using a QM model-based approach

Ellen Steimers¹  | Yevgen Matviychuk² | Daniel J. Holland²  | Hans Hasse¹ | Erik von Harbou³

¹Laboratory of Engineering Thermodynamics (LTD), Technische Universität Kaiserslautern, Erwin-Schrödinger-Straße 44, Kaiserslautern, 67663, Germany

²Department of Chemical and Process Engineering, University of Canterbury, Private Bag 4800, Christchurch, 8140, New Zealand

³Laboratory of Reaction and Fluid Process Engineering, Technische Universität Kaiserslautern, Erwin-Schrödinger-Straße 44, Kaiserslautern, 67663, Germany

Correspondence

Erik von Harbou, Technische Universität Kaiserslautern, Laboratory of Reaction and Fluid Process Engineering, Erwin-Schrödinger-Straße 44, Kaiserslautern 67663, Germany.
Email: lrfpub@mv.uni-kl.de

Funding information

Deutsche Forschungsgemeinschaft, Grant/Award Number: 310714510; New Zealand Ministry for Business Innovation and Employment, Grant/Award Number: UOCX1502

Abstract

The measurement of self-diffusion coefficients using pulsed-field gradient (PFG) nuclear magnetic resonance (NMR) spectroscopy is a well-established method. Recently, benchtop NMR spectrometers with gradient coils have also been used, which greatly simplify these measurements. However, a disadvantage of benchtop NMR spectrometers is the lower resolution of the acquired NMR signals compared to high-field NMR spectrometers, which requires sophisticated analysis methods. In this work, we use a recently developed quantum mechanical (QM) model-based approach for the estimation of self-diffusion coefficients from complex benchtop NMR data. With the knowledge of the species present in the mixture, signatures for each species are created and adjusted to the measured NMR signal. With this model-based approach, the self-diffusion coefficients of all species in the mixtures were estimated with a discrepancy of less than 2 % compared to self-diffusion coefficients estimated from high-field NMR data sets of the same mixtures. These results suggest benchtop NMR is a reliable tool for quantitative analysis of self-diffusion coefficients, even in complex mixtures.

KEYWORDS

benchtop NMR spectrometer, PFG ¹H NMR measurements, quantum mechanical model-based approach

1 | INTRODUCTION

Diffusion plays a significant role in many types of chemical and biological processes. Thus, the measurement of diffusion coefficients has been part of scientific research for many years. High-field nuclear magnetic resonance (NMR) spectroscopy is a powerful tool that allows an accurate measurement of self-diffusion coefficients. However, these measurements are complex, and well-trained users are required. Low-field benchtop NMR

spectrometers could be used, but they place higher demands on the analysis of the acquired NMR signals. This work demonstrates the applicability of benchtop NMR spectrometers for the measurement of self-diffusion coefficients with the use of a recently developed model-based approach that is able to deal with complex benchtop NMR signals with significant peak overlap.

A differentiation is made between transport and self-diffusion, both characterised by their specific diffusion coefficient. Transport diffusion coefficients are needed to

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Magnetic Resonance in Chemistry* published by John Wiley & Sons Ltd.

describe mass transfer, which is a prerequisite for the design of separation processes. Self-diffusion coefficients are important material properties that characterise the mobility of the molecules and contain information on molecule size, shape, or interaction parameters.^[28] At infinite dilution, both the transport and the self-diffusion coefficients of the diluted component merge into each other. For the measurement of self-diffusion coefficients, pulsed-field gradient (PFG) NMR spectroscopy methods are well established. In contrast to other measurement methods, e.g., tracer measurements,^[34] PFG NMR spectroscopy is fast, requires only a small sample volume, and covers a wide temperature and pressure range.^[13]

In PFG NMR experiments, two consecutive pulsed-field gradients are applied. The first gradient pulse causes a dephasing of the spins, whereas the second pulse is intended to refocus the spins back to the initial phase. However, due to molecular motion of the spins, the initial phase is not restored fully. Measurements consist of many molecules, and when the residual phase of many different molecules is combined, it leads to an attenuation of the acquired NMR signal. The attenuation experienced by a spin is used to measure its self-diffusion coefficient. The PFG NMR experiment is usually repeated with incrementing gradient strength. This results in a series of NMR spectra with signals that, depending on their self-diffusion coefficient, show a more or less strong attenuation of their signal intensities (cf. Figure 1).

A general description of the dependence between the gradient strength and the attenuation of the acquired

NMR signal, using a spin-echo pulse sequence, is given by the Stejskal-Tanner Equation (1)^[31]:

$$I_i = I_{0,i} \exp\left(-D_i \gamma^2 g^2 \delta^2 \left(\Delta - \frac{\delta}{3}\right)\right). \quad (1)$$

Therein, I_i denotes the intensity of the signal i , $I_{0,i}$ is the signal intensity with no gradient applied, γ is the gyromagnetic ratio, g is the gradient strength, δ is the duration of the gradient pulse, Δ is the diffusion time, and D_i is the self-diffusion coefficient for signal i . The self-diffusion coefficients are derived by fitting Equation (1) to the series of signal intensities obtained from PFG NMR experiments. A variety of different PFG style pulse sequences exist. Equation (1) has to be adjusted for each pulse sequence, usually by changing the correction to Δ .

Signals that originate from the same molecule show the same attenuation and thus identical values for the self-diffusion coefficients. This characteristic is used in so-called diffusion ordered spectroscopy (DOSY) maps to identify different chemical species in a mixture. In DOSY maps, the signals are plotted as a function of the chemical shift value of the acquired NMR signal and the corresponding self-diffusion coefficient. The width of the plotted signal is determined by the accuracy of the fit of the self-diffusion coefficient, e.g., as obtained from Equation (1).

In principle, self-diffusion coefficients are directly accessible from Equation (1) via the inverse Laplace

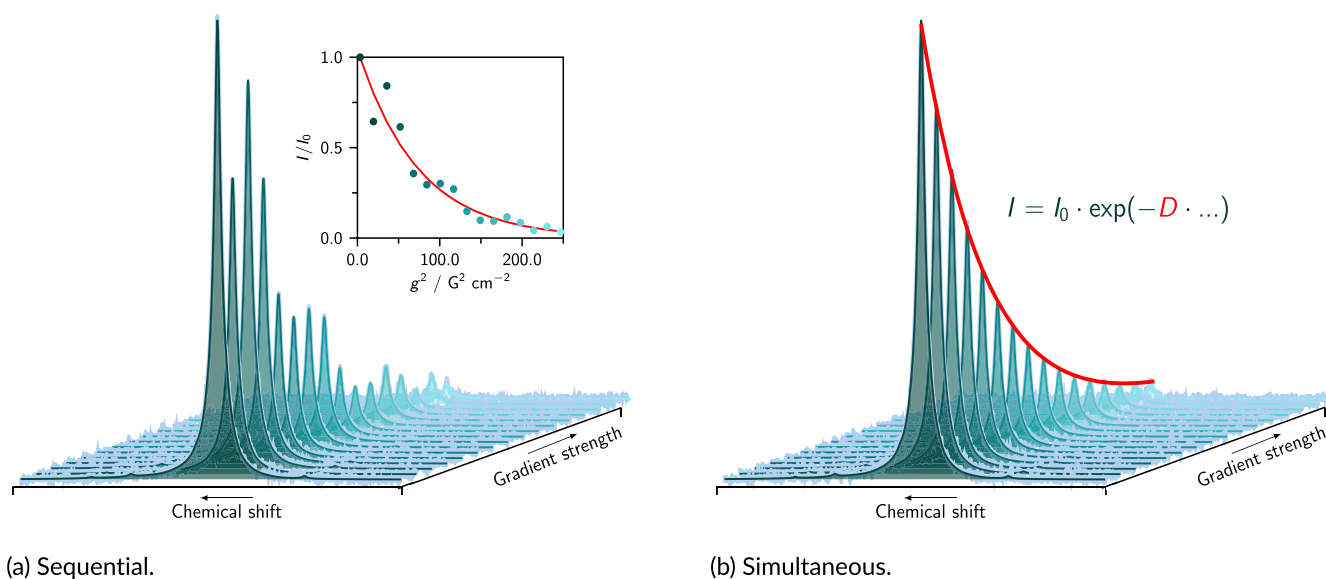


FIGURE 1 Principle of self-diffusion measurements with NMR spectroscopy; left: sequential (standard) approach, i.e., each spectrum is adjusted separately; right: simultaneous approach, i.e., amplitudes are coupled via an higher-level function (here, Equation (5)) and simultaneously adjusted with the meta parameter of the higher level function (here, D)

transform. This method potentially enables the determination of different self-diffusion coefficients when multiple signals overlap. However, the inverse Laplace transform is an ill-posed problem, which can have an infinite number of solutions.^[6] Inverse Laplace transform methods are used, but these require some regularisation functions and the resulting self-diffusion coefficients may not be precise.

Advanced processing methods for the estimation of self-diffusion coefficients from PFG NMR experiments exist that can be classified into univariate and multivariate methods. Results from univariate methods are typically visualised in DOSY maps, whereas multivariate methods yield the self-diffusion coefficients of all species in the mixture together with their individual signature spectra.^[26]

Univariate methods assume that each signal in an NMR spectrum arises from one individual species and thus has a unique self-diffusion coefficient. The signals are approximated by Gauss-Lorentzian curves, and the attenuation of each signal is fitted by a mono-exponential or, if the peaks slightly overlap, by a multi-exponential decay. The most common univariate method is the High-Resolution (HR)-DOSY approach.^[2] For well-resolved NMR spectra and only slight overlapping peaks, HR-DOSY is the method of choice that yields a high resolution for the separation of different species in a mixture.^[26] However, if the spectra suffer from significant peak overlap, as is often the case in benchtop NMR data, more than one species contributes to the attenuation of the acquired NMR signal, and the self-diffusion coefficients estimated with HR-DOSY may be biased.^[26]

Multivariate methods exploit the fact that signals that arise from the same species show identical attenuation of their NMR signal by regarding the spectra (or parts of it) as a whole and fit the NMR spectrum simultaneously.^[26] Thus, these methods are well suited for NMR signals with overlapping peaks. Several methods exist, such as Direct Exponential Curve Resolution Algorithm (DECRA),^[35] Multivariate Curve Resolution (MCR)^[32] or Speedy Component Resolution (SCORE).^[27] The acquired PFG NMR signal is described as the product of the unattenuated signal of each species and the diffusion decay profiles.^[26] The goal is to estimate the contribution from each signal in the frequency domain and the diffusion decay so that the residuals between the modelled and the measured NMR signal become minimal. Most of these methods use prior knowledge of the number of species present in the mixture to improve the fitting process. The results of both the univariate and the multivariate analysis strongly depend on the quality of the given PFG NMR data and the amount of overlap between peaks. Also post-processing steps are required, including phase

and baseline correction, zero-filling, apodisation, and reference deconvolution.^[6]

In order to get a good resolution of the acquired NMR signal, which is a prerequisite for accurate estimation of self-diffusion coefficients and separation of different species, PFG NMR measurements are typically carried out using high-field NMR spectrometers. There are several publications that demonstrate the use of high-field PFG NMR spectroscopy for an accurate estimation of self-diffusion coefficients or its use to gather more information on the studied systems, e.g., dynamic behaviour.^[3,13,14,28] The high resolution, however, is associated with high costs as special laboratory equipment is required. This has driven the development of low-field benchtop NMR spectrometers. These devices are affordable, compact, and user-friendly, which makes NMR spectroscopy more accessible to general laboratories.

Unfortunately, the lower magnetic field entails a lower spectral resolution and sensitivity of the acquired NMR signals, compared to that obtained with high-field NMR spectrometers. In addition, significant peak overlap prevents the use of standard analysis methods. To overcome these challenges, sophisticated analysis methods have been developed that enable an accurate quantitative analysis of benchtop NMR signals, thus extending the application of benchtop NMR as an analytical tool in the field of reaction and process monitoring.^[9,15,23] The use of benchtop NMR spectrometers for PFG NMR experiments has also been reported.^[12,22]

PFG NMR experiments on a benchtop NMR spectrometer have been used to study the dynamic properties of aqueous salt solutions^[10] or the solvent effects in homogeneous catalysis.^[25] Others have used benchtop PFG NMR experiments for the determination of the molar mass of lignins,^[29] the hydrodynamic radii of PEE-G dendrons^[21] or for the characterisation of porous catalytic material.^[11] The treating of benchtop PFG NMR data, however, is especially challenging, as the low resolution impedes accurate estimation of the peak intensities, which are needed for an accurate determination of the self-diffusion coefficients. Assemat et al.^[1] reported the use of a benchtop NMR spectrometer to predict self-diffusion coefficients from drug samples using univariate processing. Although discrepancies between the absolute values of self-diffusion coefficients from benchtop and comparative high-field NMR measurements were obtained, a consistent separation of the different species in the mixture was possible, at least in the absence of significant peak overlap. McCarney et al.^[22] used both the multivariate SCORE algorithm and the HR-DOSY approach to analyse benchtop PFG NMR signals with highly overlapping peaks. In contrast to the univariate HR-DOSY approach, the multivariate processing yielded

consistent self-diffusion coefficients and a clear separation of the individual species.

In this work, we present a model-based approach for the estimation of self-diffusion coefficients from benchtop PFG NMR signals with a target accuracy similar to high-field NMR experiments. In this approach, we combine a quantum mechanical model that allows the estimation of ideal signatures for the individual species in a mixture based on a few model parameters (e.g., chemical shifts or J-coupling constants) with the Bayesian framework. Model-based approaches have already been proven to be particularly suitable for the analysis of benchtop NMR data with low signal-to-noise ratios and overlapping peaks.^[4,7,24] The use of a quantum mechanical model also reduces the number of model parameters by not treating the signals corresponding to the same species independently. The model formulation in terms of the Bayesian statistics allows the inclusion of prior knowledge about the model parameters.

This model formulation offers the possibility to analyse diffusion data in two different ways. As is usually the case, the acquired NMR signals may be treated independently and the signal intensities and the values of the self-diffusion coefficients are estimated sequentially, as shown in Figure 1a. In contrast to that, the inclusion of prior knowledge in the Bayesian framework allows the simultaneous fitting of the peak intensities and the corresponding self-diffusion coefficients by analysing the entire data set. This approach may help to minimise the uncertainties in the determination of self-diffusion coefficients from complex NMR data, for example, measured with benchtop NMR spectrometers. The simultaneous procedure is visualised in Figure 1b.

The QM model fitting has already been successfully established for the analysis of complex 1D ¹H NMR signals acquired with benchtop NMR spectrometers.^[18,30] Even for mixtures with up to 15 components and low signal-to-noise ratio, accurate quantification results can be obtained, as was recently demonstrated for the analysis of wine samples.^[17]

In this work, the application of the method was extended to the estimation of self-diffusion coefficients in complex mixtures with benchtop NMR spectroscopy. In a first study, the inclusion of prior knowledge about the attenuation of the PFG ¹H NMR signal (i.e., the Stejskal-Tanner equation) was tested for synthetic PFG ¹H NMR data sets with known self-diffusion coefficients, overlapping peaks, and low signal-to-noise ratio. Further, PFG ¹H NMR experiments of simple mixtures were acquired on a high-field and a benchtop NMR spectrometer. The high-field NMR data were analysed with the proposed model-based approach and the estimated self-diffusion coefficients were compared to a standard analysis

procedure, the method of direct integration. In addition, the benchtop PFG ¹H NMR data of the simple mixtures were analysed using the model-based approach, and the results were compared to those from high-field PFG ¹H NMR experiments. In a last study, the model-based approach was applied to two aqueous mixtures consisting of glycerol (Glyc), 1,2-propylene glycol (PG), monodiethanolamine (MDEA), ethanol (EtOH), and acetonitrile (ACN). The similar chemical structure of these species makes this mixture a challenging and complex system. The mixtures were measured on both the high-field and the benchtop NMR spectrometer and analysed with the model-based approach. For comparison, the benchtop NMR data were also analysed with quantitative global spectral deconvolution (qGSD),^[8] as standard peak integration was not applicable due to strong overlapping peaks. It should be mentioned that it was not the aim of this work to predict self-diffusion coefficients of species in mixtures, since they depend on the composition and only two different compositions were studied. Instead, the focus was on demonstrating that such measurements are possible using a benchtop NMR instrument. This work shows that benchtop NMR spectroscopy, supported by advanced analysis methods, is a good alternative to expensive high-field NMR spectroscopy. It can significantly help to expand the application field of NMR spectroscopy.

2 | MODEL-BASED QUANTIFICATION

Model-based quantification methods have recently renewed interest in benchtop NMR spectrometers as significant peak overlap and low signal-to-noise ratios of NMR signals acquired with benchtop NMR spectrometers preclude the application of traditional analysis methods, such as direct integration. One class of model-based approaches that were primarily developed for the quantification of 1D NMR signals makes use of the prior knowledge that the species in a mixture are known and that the intensity ratios of peaks pertaining to the same species are locked. These methods describe the measured NMR signal as a sum of signatures for each species present in the mixture. One way to derive these signatures is to approximate the component spectra with lineshape models that are agnostic to the underlying chemical structure, for example, a collection of Gauss-Lorentzian curves as done in the Indirect Hard Modelling method (IHM).^[16] The measured NMR signal of a mixture can then be described as a weighted sum of the signatures of the individual species, where the weighting factors correspond to the amount of each species in the mixture.

Our proposed method uses a similar approach, but the underlying signature models for each species are derived using a quantum mechanical model to reduce the number of free parameters and improve the robustness (e.g., to easily account for changes of chemical shift values with composition). With the signatures of the individual species, the measured NMR signal is described by Equation (2):

$$s_{\text{model}}(\{\theta_k, c_k\}_{k=1}^K, \varphi_0, \tau) = e^{i\varphi_0} \sum_{k=1}^K c_k u_k(\theta_k, \tau). \quad (2)$$

The signature model u_k of a component k is a set of Lorentzian functions, parameterised by the vector θ_k . The weighting factor c_k describes the signal intensity and is thus proportional to the amount of the species k in the sample. The vector θ_k includes the model parameters, such as the chemical shift values, J-coupling constants, and information on the relaxation delay that is proportional to the width of transition peaks. The global phase shift φ_0 and the ringdown delay τ determine the zero- and first-order phase angles of the NMR signal, respectively. Given a 1D NMR signal s the concentration of the species are estimated by minimising the differences between the measured and the modelled NMR signal with respect to the model parameters, for example, in a least-squares sense. The QM model formulation has the advantage that the underlying model is independent from the magnetic field strength and thus its QM parameters (chemical shifts and J-coupling values) can be determined from high-field NMR spectra, or chemical databases, and then applied to benchtop NMR data. More information about the QM model set up can be found in Matviychuk et al.^[20]

The QM NMR model captures a variety of complex peak shapes arising from higher-order coupling effects. However, it is still not a perfect representation of the spectra. For example, we do not necessarily model all long-range couplings, or there may be imperfections in the magnetic field that lead to distortions of the peak shape. To account for these model misspecification, and thus to improve the accuracy of the model fit, a recently developed approach was used for the high-field data.^[19] For the benchtop data, this approach was not used because significant peak overlap led to incorrect distributions of the pure residual signals between the different species. In the following, the set of model parameters for all species $\{\theta_k\}_{k=1}^K$ is referred to as Θ and the set of signal intensities $\{c_k\}_{k=1}^K$ as c .

The major challenge of model-based fitting is to find the best set of model parameters that match the measured NMR signal, as there are often several locally optimal solutions. Casting the parameter estimation problem

in terms of the Bayesian framework allows us to incorporate prior knowledge about the model parameters and thus reduce the possibility of the fitting algorithm being stuck in one of the local optima.

If the model-based fitting is combined with the Bayesian framework, the new target function is the posterior distribution $p(\Theta, c|s)$ that characterises the probability that the set of model parameters produces the measured NMR signal. The posterior is proportional to the product of the likelihood function $p(s|\Theta, c)$, which describes the residuals between the measured and the modelled NMR signal, and the prior distribution of model parameters $p(\Theta, c)$:

$$p(\Theta, c|s) \propto p(s|\Theta, c)p(\Theta, c). \quad (3)$$

By maximising the posterior with respect to the model parameters the most probable set of model parameters is found that has produced the measured NMR signal. The prior term can include prior information about the model parameters that is available even before the measurement. The amplitudes, for example, are described with Gaussian priors; if no information about the intensities is available, we chose uninformative priors, for example, Gaussians with large variances.

Prior information about the model parameters is helpful to improve the model fitting process, especially when analysing benchtop NMR data. As mentioned above, QM parameters (chemical shifts and J-coupling values) can be adopted from high-field measurements. Also, statistical dependencies between model parameters can be modelled using joint priors, as has been demonstrated in Matviychuk et al.^[18] This approach can also be used in kinetic experiments in which the peak amplitudes of the reactants are dependent of the stoichiometry of the reaction as well as in PFG NMR experiments, where the acquired NMR signals are related via the Stejskal-Tanner equation. Prior knowledge about the dependence between peak intensities can be incorporated via some higher-level function, for example, a kinetic or a diffusion model. By fitting the higher-level function to the acquired NMR signal, the NMR signals in a PFG NMR data set are adjusted simultaneously with the corresponding meta parameters such as kinetic constants or self-diffusion coefficients (cf. Figure 1b). Another advantage of working with Bayesian statistics is the possibility to estimate the uncertainty in the quantification with Markov-Chain-Monte-Carlo (MCMC) sampling methods. In MCMC sampling, the entire posterior is sampled with respect to all model parameters of interest, resulting in distributions for all model parameters instead of point estimates that are obtained by maximising the posterior. From the found distributions, the mean, standard

deviations, and credible intervals of the model parameters can be estimated.

3 | MATERIALS AND METHODS

3.1 | Chemicals

Table 1 lists all chemicals that were used for the PFG ^1H NMR measurements together with their chemical structure, suppliers, and purities specified by the suppliers. Ultrapure water was produced with an ultrapurification system of Merck Millipore. In addition, a Bruker standard diffusion sample with a well-defined self-diffusion coefficient at 25°C was used to validate the benchtop PFG ^1H NMR measurements. The Bruker standard diffusion sample (PN Z10906) is composed of 1 % H_2O in D_2O , with 0.1 % GdCl_3 and 0.1 % methanol (^{13}C).

3.2 | Sample preparation and data acquisition

All samples were gravimetrically prepared using a precision balance with an absolute error of 0.001 g specified by the manufacturer. For the study of the simple mixtures, MDEA and glycerol were diluted in D_2O in a volumetric ratio of 1:100. In addition, a second dilution of MDEA in D_2O of 1:2 was prepared. Acetonitrile and water were used as pure components, and the Bruker standard sample was used as supplied by the manufacturer. Two complex mixtures were prepared, consisting of

MDEA, glycerol, 1,2-propylene glycol, ethanol, acetonitrile, and water. The composition of both mixtures is given in Table 2. All mixtures were measured using standard 5 mm NMR tubes. For the pure components acetonitrile and water, special 2.5 mm NMR tubes were used to reduce convection and minimise the risk of radiation damping.

3.2.1 | High-field PFG NMR

High-field PFG ^1H NMR measurements were carried out on a 400 MHz NMR spectrometer from Bruker Biospin, which is equipped with a probe with cryogenically cooled electronics (magnet Ascend 400, console Avance 3 HD 400, probe CryoProbe Prodigy). The PFG ^1H NMR experiments were performed with a stimulated spin-echo pulse sequence using bipolar gradients (stebpgp1s) provided by the Software TopSpin. Accordingly, the following modification of the generalised Stejskal-Tanner equation was used:

$$I_i = I_{0,i} \exp \left[\sum_{n=1}^2 c_n \left(-D_i \gamma^2 g^2 \delta^2 \left(\Delta - \frac{\delta}{3} - \frac{\tau}{2} \right) \right)^n \right]. \quad (4)$$

Therein, τ is a correction constant that accounts for the use of bipolar gradients. The probe specific calibration factors c_1 and c_2 were adopted from Bellaire et al.^[3] and account for the non-linearity of the gradient. A graphical view of the pulse sequence can be found in the supporting information. The diffusion time for every PFG ^1H NMR experiment was set to $\Delta = 50$ ms, and the

TABLE 1 Chemical structure, purities, and suppliers of the chemicals used in this work

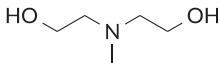
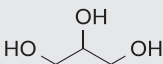
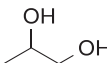


Chemical	Structure	Supplier	Purity / mass-%
MDEA		Sigma Aldrich	≥ 99.00
Glycerol		VWR Chemicals	99.50
1,2-Propylene glycol		Sigma Aldrich	> 99.00
Ethanol		Merck	> 99.90
Acetonitrile		VWR Chemicals	≥ 99.95

TABLE 2 Gravimetrically determined mole fractions of all species in mixtures 1 and 2

Mixture	$x_i / \text{mol mol}^{-1}$					
	MDEA	Glyc	PG	EtOH	ACN	Water
1	0.0057	0.0112	0.0172	0.0086	0.0143	0.9409
2	0.0114	0.0115	0.0114	0.0114	0.0114	0.9429

correction constant for the bipolar gradient was $\tau = 0.2$ ms. The gradient strength was incremented in 16 steps from 2.5 to 48.5 G cm⁻¹ for the simple mixtures and from 5.6 to 48.5 G cm⁻¹ for the complex mixtures. The gradient steps were spaced equally with increasing gradient amplitude squared. The duration of the gradient pulse δ was adjusted for every sample separately to ensure approximately 90–95 % attenuation of the acquired NMR signal for the strongest gradient. The set duration times for both the simple and the complex mixtures can be found in the supporting information. The high-field PFG ¹H NMR spectra were recorded with 16 scans, 32k data points for the simple mixtures and 48k data points for the complex mixtures, an acquisition time of 8 s, and a relaxation delay ranging between 10 and 20 s to ensure sufficient relaxation (at least $4 \times T_1$) of the sample.

All high-field PFG ¹H NMR measurements were performed at 28.7 °C to ensure the same temperature conditions compared to the benchtop NMR measurements. All measurements were repeated 10 times to avoid biased values of the self-diffusion coefficients due to possible random measurement errors, e.g., fluctuations in temperature, and to estimate the uncertainty in the measured self-diffusion coefficients.

3.2.2 | Benchtop PFG NMR

Benchtop NMR measurements were performed on a Magritek Spinsolve Carbon 43 MHz NMR spectrometer, which is equipped with a gradient coil with a maximum field gradient of 15.7 G cm⁻¹. For the PFG ¹H NMR experiments on the benchtop NMR spectrometer, we use a pulsed-field gradient stimulated echo (PGSTE) sequence that is available in the standard operating software Spinsolve (Magritek). The following modification of the general Stejskal-Tanner equation was used to determine the self-diffusion coefficients from the acquired NMR signals:

$$I_i = I_{0,i} \exp\left(-D_i \gamma^2 g^2 \left(\delta_{\text{eff}}^2 \left(\Delta - \frac{\delta_{\text{eff}}^2}{3}\right)\right)\right). \quad (5)$$

Therein, δ_{eff} is the effective duration of the gradient pulse that is composed of the manually set duration and a gradient ramp time that helps to prevent distortions in the acquired NMR signal. As for the high-field NMR experiments, the gradient strength was incremented in 16 steps, which were spaced equally with increasing gradient amplitude squared, and ranged from 1.8 to 15.7 G cm⁻¹. The ramp time was 0.1 ms and was equal for all experiments.

To avoid signal distortions arising from J-modulation, the duration of δ was kept constant at a maximum value

of 2 ms, and the diffusion time Δ was adjusted accordingly to obtain attenuation of approximately 90–95 % of the acquired NMR signals for the strongest gradient strength. Details on the pulse sequence and the acquisition parameters can be found in the supporting information.

The PFG ¹H NMR spectra were acquired with 16 scans, 32k data points, an acquisition time of 6.4 s, and a relaxation delay ranging between 10 to 25 s to ensure sufficient relaxation (at least $4 \times T_1$) of the sample. For the Bruker standard sample, 32 scans were used to ensure approximately the same signal-to-noise ratio conditions as for the other simple mixtures. The benchtop NMR spectrometer is operating at a temperature of 28.7 °C and all measurements were repeated 10 times.

3.3 | Data processing and quantification

The least-squares model-based fitting, described in Section 2, was implemented in a custom software written in Python 3.5.

For the estimation of the signal intensities from an ¹H NMR spectrum, the model parameters (chemical shifts, J-couplings, peak widths, and phasing values) were optimised iteratively with the L-BFGS-B algorithm from the optimisation package of the Python SciPy toolbox.^[5,33] For the analysis of the benchtop ¹H NMR data, initial guesses, that is, prior knowledge, for the model parameters were adopted from the high-field analysis. In addition, zero-filling with a factor of 2 was applied to all benchtop NMR signals.

Each spectrum within a PFG ¹H NMR data set was adjusted separately and the estimated set of signal intensities for each species c_k was used to determine the corresponding self-diffusion coefficient via Equations (4) and (5) for the high-field and benchtop PFG ¹H NMR data sets, respectively. The non-linear least-squares fit of the Stejskal-Tanner equation was performed either in Python with the Levenberg-Marquardt algorithm that is used by the curve fit function or in MATLAB using the lsqnonlin solver that is based on the trust-region-reflective algorithm, both yielding equivalent results. In the following, this procedure is referred to as the standard procedure.

The inclusion of prior knowledge about the signal intensities, i.e., the relation via the Stejskal-Tanner equation, allows the simultaneous fitting of the signal intensities and the corresponding self-diffusion coefficients. This might be especially helpful for the analysis of complex NMR spectra, e.g., with overlapping peaks and low signal-to-noise ratios, as additional information reduces the number of degrees of freedom and thus improves the model-fitting process.

In order to investigate this possibility, synthetic PFG ^1H NMR data sets with known self-diffusion coefficients were generated and analysed with the sequential and the simultaneous fitting procedure. The spectra were modelled with two overlapping peaks (A and B). In addition, random noise was added to the data to achieve low signal-to-noise ratio conditions, which we expect to favour the simultaneous analysis approach. The true values for the self-diffusion coefficients were set to $D_A = 3.5 \cdot 10^{-9} \text{ m}^2\text{s}^{-1}$ and $D_B = 1.5 \cdot 10^{-9} \text{ m}^2\text{s}^{-1}$. In addition, two different scenarios were studied. In the first scenario, the sequential and simultaneous fitting procedures were compared analysing a synthetic PFG ^1H NMR data set with 16 gradient steps. The second, worst case, scenario demonstrates the performance of both procedures if only five gradient steps were measured.

The synthetic data sets of both scenarios were analysed with the sequential procedure as described above. For the simultaneous fitting, Equation (5) was implemented in Python as a Gaussian prior for the decay of the amplitudes of both peaks. Based on the fitting results obtained from the sequential analysis, the self-diffusion coefficients and the amplitudes were adjusted in a next step simultaneously, but iteratively for both peaks.

To estimate the uncertainty of the PFG ^1H NMR measurements, it is necessary to differentiate between different types of errors that contribute to the overall error of the estimated self-diffusion coefficients. First, the values of the self-diffusion coefficients may be erroneous due to differences between the experimental and the modelled NMR signals. For example, inhomogeneities in the magnetic field lead to distortions of the peak shape, and the assumption of ideal signatures is no longer valid. Further, in the standard model-based fitting procedure, the amplitudes are assumed to be Gaussian distributed, and only point estimates of the amplitudes are considered for the estimation of the self-diffusion coefficients. By MCMC sampling the complete distribution of the amplitudes can be estimated. This offers the possibility to calculate the mean values and the variance of the amplitudes, which would give more information on the accuracy of the estimated self-diffusion coefficients. Another source of error is the accuracy of the fit of the Stejskal-Tanner equations (4) and (5) to the PFG ^1H NMR data. This accuracy can be quantified with the standard error of the optimised parameters, that is, the self-diffusion coefficient, that is given by the least-squares solver. Since several repetitions of the PFG ^1H NMR experiments were performed, a variance in the estimated self-diffusion coefficients between the different experiments is also to be expected.

The estimation of the overall error of the self-diffusion coefficients is briefly described below. First,

MCMC sampling was used to calculate the distributions of amplitudes for each signal in each of the 16 spectra from a PFG ^1H NMR data set. Each amplitude was sampled with 4 random walkers and 1,000 steps for the simple mixtures and 250 steps for the complex mixtures. The range for the chemical shift values for MCMC sampling was set to $\pm\sigma$, where σ is the standard deviation of the individual fits, obtained in a least-squares sense, for each spectra in the data set (cf. Section 2). The range for the peak widths was set to ± 0.05 .

Sampling the amplitude of one signal in a PFG ^1H NMR data set results in 16 distributions for the amplitude, one distribution for each spectrum with a different gradient strength. Then, one point was randomly taken from each distribution, resulting in a series of 16 amplitudes one for each gradient strength. For each of these series, the self-diffusion coefficient was determined with the least-squares fit of the corresponding Stejskal-Tanner equation, i.e., Equation (4) or (5). This results in a number of distributions for the self-diffusion coefficients equal to the number of samples drawn from the amplitudes. Each distribution is thereby characterised by the mean and the variance of the self-diffusion coefficient. To obtain the overall distribution of the self-diffusion coefficient, 1,000 random samples were drawn from each distribution. From these final distributions, the mean and the variance of the self-diffusion coefficients were extracted and used to define the confidence intervals for each self-diffusion coefficient, which are shown as error bars in the following figures. These confidence intervals account for the errors of the model fit, the variance of the amplitudes, and the uncertainty of the fit of the Stejskal-Tanner equation.

The determination of signal intensities by direct integration for the high-field PFG ^1H NMR data of the simple mixtures was performed in MestreNova (Mestrelab, v14.2.0). Phase and baseline correction as well as zero-filling and apodisation were done automatically. The integration boundaries were set to encompass all peaks of the species of interest. The qGSD analysis of the benchtop PFG ^1H NMR data of the complex mixtures was also performed in MestreNova. The spectra were manually phase corrected, and the baseline was adjusted using automated algorithms from MestreNova. The qGSD algorithm was applied with eight improvement cycles. With the multiplet manager, overlapping peaks were assigned to the corresponding species.

From the peak intensities determined with direct integration as well as with qGSD, the self-diffusion coefficients were estimated with the Equations (4) and (5). The processing parameters of the high-field and the benchtop PFG ^1H NMR spectra can be found in the supporting information.

4 | RESULTS AND DISCUSSION

4.1 | Test study: Simultaneous approach

The synthetic PFG ^1H NMR spectra for both scenarios are shown in Figure 2, where the five spectra of the second scenario are marked with red rectangles. Both peaks strongly overlap and have a low signal-to-noise ratio, which makes the analysis, especially of peak A with a higher self-diffusion coefficient, challenging.

The PFG ^1H NMR data were analysed using the sequential procedure, i.e., estimation of signal intensities using the model-based approach and fitting Equation (5) to determine D , and the simultaneous procedure, i.e., coupling of amplitudes via Equation (5) and simultaneous fitting of D and the amplitudes. Both peaks show, independent of the number of gradient steps, a significant attenuation of their amplitude, which allows the determination of the self-diffusion coefficients for all data sets. It is also obvious that the signal intensities that were estimated with the sequential model-based approach

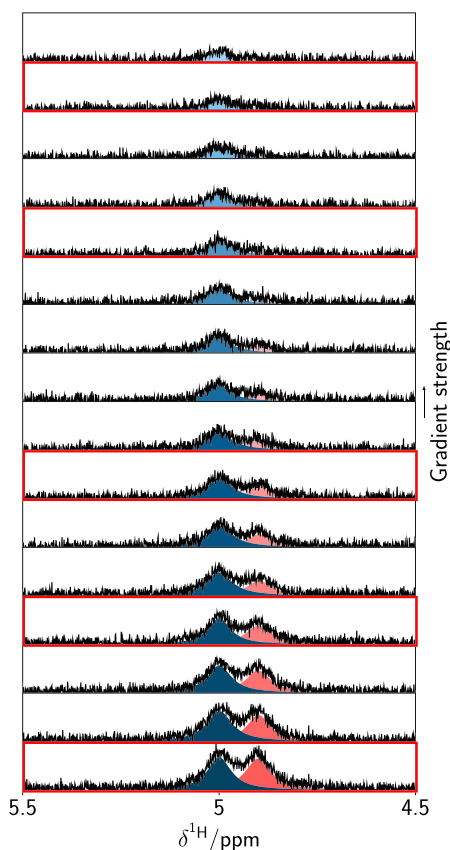


FIGURE 2 Synthetic PFG ^1H NMR data set with two peaks (A \approx 4.8 ppm, B \approx 5 ppm) for comparison of the sequential and simultaneous fitting procedure; scenario 1: 16 spectra, scenario 2: 5 spectra (marked with red rectangles)

(standard procedure) scatter more compared to the signal intensities after the simultaneous fitting. This is due to the fact that in the latter case the amplitudes are correlated via the Stejskal-Tanner equation and an optimisation of the self-diffusion coefficients leads to smooth values for the amplitudes. Figure 3 shows the signal intensities of peak A and B of one synthetic PFG ^1H NMR data set for each scenario, after the fitting with the sequential and the simultaneous procedure. The solid line, represents the Stejskal-Tanner fit of Equation (5).

To compare the performance of each procedure for the estimation of the self-diffusion coefficients, 20 PFG ^1H NMR data sets were synthesised and analysed using the sequential and simultaneous fitting procedure. Table 3 shows the mean values of the self-diffusion coefficients from these analyses and compares these with the true values for the self-diffusion coefficients. Root mean square errors (RMSE) between the true and the estimated self-diffusion coefficients were calculated from

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^N (D_n^{\text{est}} - D^{\text{true}})^2}, \quad (6)$$

where N is the number of synthesised PFG ^1H NMR data sets.

For scenario 1 in which 20 PFG ^1H NMR data sets with 16 gradient steps were analysed with the sequential and the simultaneous procedure, the estimated mean values of the self-diffusion coefficients for both peaks A and B are close to the true values of the self-diffusion coefficients. Also, for scenario 2 in which 20 PFG ^1H NMR data sets with only five gradient steps were analysed, a good agreement between the true and the estimated self-diffusion coefficients is achieved. Due to the higher true value of the self-diffusion coefficient of peak A, which leads to a stronger attenuation of the NMR signal, the overall differences between the true and estimated values of self-diffusion coefficient for peak A are slightly larger than those for peak B. However, there is no significant difference between the sequential and the simultaneous fitting procedure. For both procedures, the mean values of the self-diffusion coefficients were similar. Furthermore, the RMSE values do not differ significantly between the sequential and the simultaneous fitting procedure. Thus, although the simultaneous fitting procedure leads to smooth amplitudes and thus to a visually better fit to the Stejskal-Tanner equation, the final results are comparable to those of the sequential fitting procedure, even for low signal-to-noise ratios. The use of the simultaneous fitting procedure for the analysis of PFG ^1H NMR data was in this work not necessarily found

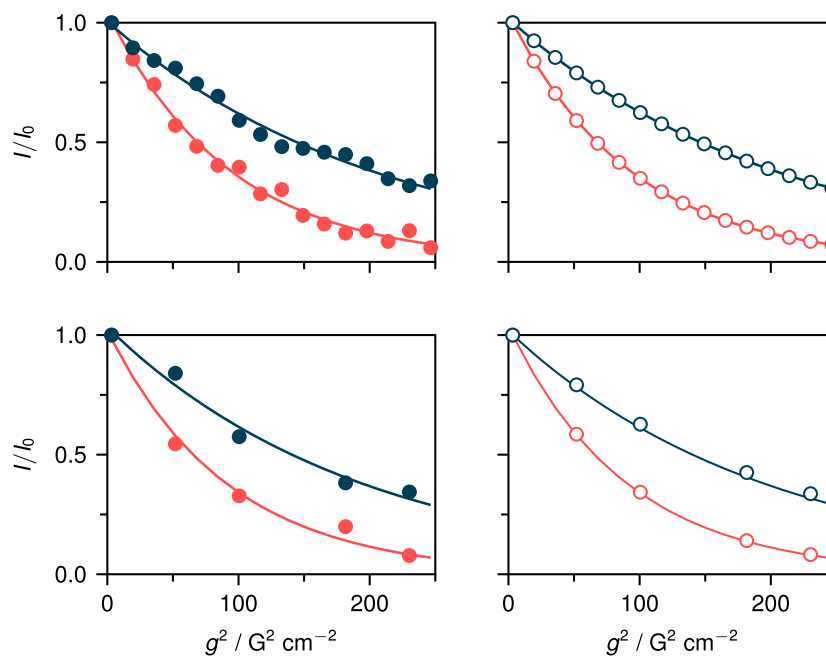


FIGURE 3 Normalised signal intensities of peak A (red) and B (blue) as a function of the gradient strength and the corresponding Stejskal-Tanner fit (solid lines) for one PFG ^1H NMR data set analysed with the sequential (left) and the simultaneous (right) fitting procedure; top: scenario 1, bottom: scenario 2

TABLE 3 Comparison of the sequential and the simultaneous fitting procedure for the estimation of self-diffusion coefficients from synthetic PFG ^1H NMR data sets with two peaks A and B

	Scenario 1				Scenario 2			
	Peak A		Peak B		Peak A		Peak B	
	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
True	3.50	-	1.50	-	3.50	-	1.50	-
Sequential	3.47	0.16	1.52	0.07	3.51	0.23	1.51	0.10
Simultaneous	3.55	0.17	1.51	0.06	3.58	0.26	1.50	0.10

Note: Scenario 1: 16 gradient steps; scenario 2: 5 gradient steps; the mean values of the self-diffusion coefficients of both scenarios were calculated from 20 synthesised PFG ^1H NMR data sets, respectively, and are expressed in units of $10^{-9} \text{ m}^2 \text{ s}^{-1}$.

to be more advantageous compared to the sequential procedure. Thus, only the sequential fitting procedure was used in the following for the analysis of experimental PFG ^1H NMR data. The inclusion of the prior knowledge could be more beneficial for systems with more complex models (e.g., kinetic models or more complex component models), which is part of ongoing research.

4.2 | Simple mixtures

For all simple mixtures, namely, MDEA (1:2), MDEA (1:100), glycerol (1:100), Bruker standard, acetonitrile,

and water, PFG ^1H NMR experiments were acquired on the high-field and on the benchtop NMR spectrometer. The signal intensities of all species were estimated using the model-based approach, and the self-diffusion coefficients were determined via Equations (4) or (5). The measured ^1H NMR signals of all simple mixtures are plotted in Figure 4 together with the modelled NMR signals, which are obtained after the model-based fitting. Each plot shows the ^1H NMR signals acquired on the high-field NMR spectrometer at 400 MHz and the same spectrum acquired on the benchtop 43 MHz NMR spectrometer. Also, the residuals between the measured and the modelled NMR signals are displayed, respectively.

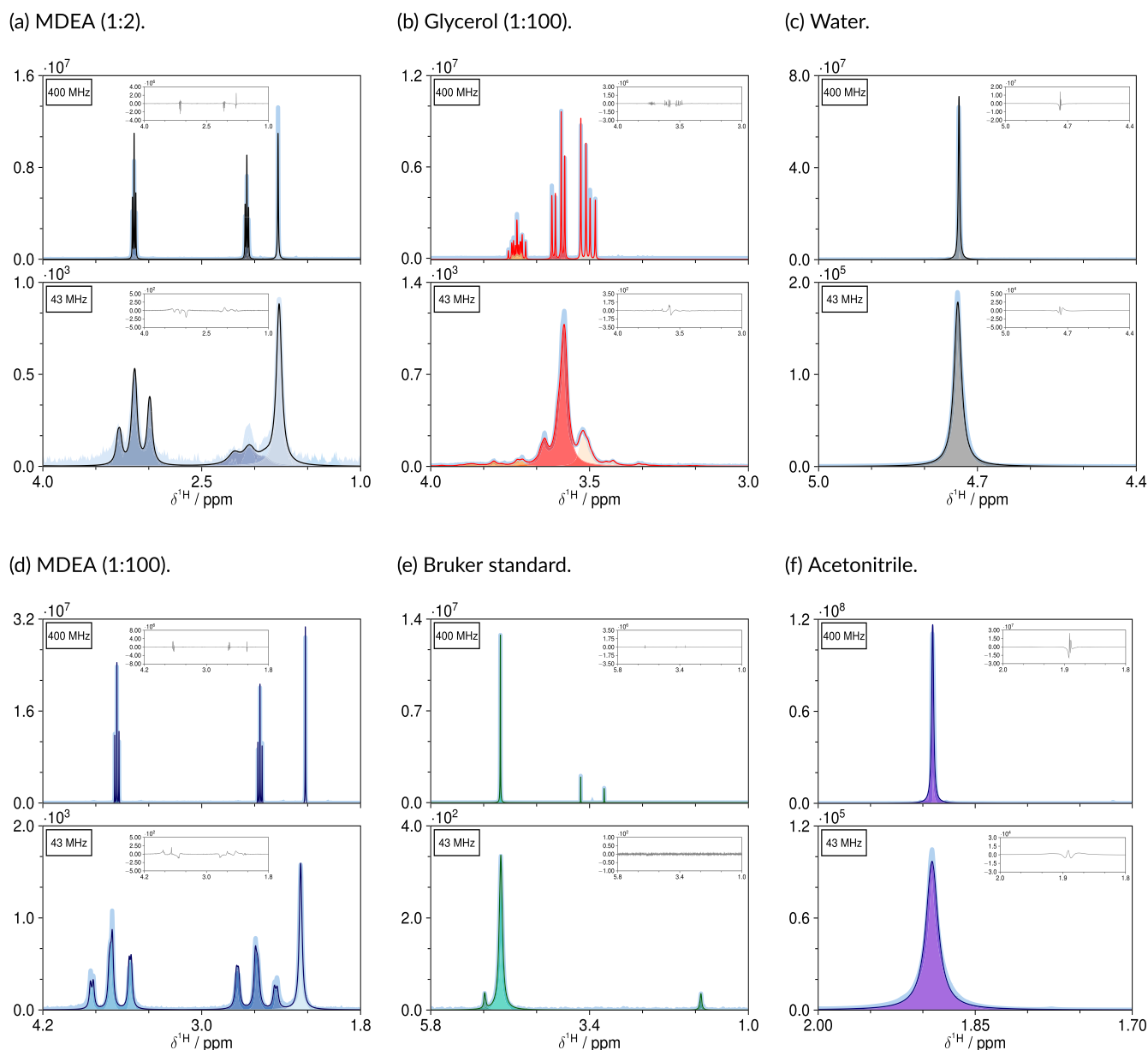


FIGURE 4 ^1H NMR spectra of all simple mixtures acquired with the high-field (top) and benchtop (bottom) NMR spectrometer; measured NMR signal (—), model-based fit (—). The different shaded colours correspond to different peaks of the same species. The insets in each figure show the residual after the model fitting process

The NMR signals acquired with the high-field NMR spectrometer show sharp and well separated peaks. In contrast, on the benchtop NMR spectrometer, the peaks become much broader and merge so that a separation of the different peaks of a species with standard analysis methods, for example, direct integration is no longer possible. With the use of the model-based approach, which uses a quantum mechanical model to describe the individual signatures, it is possible to resolve the different peaks of each species. As can be seen in the Figure 4a,b,d, even the complex MDEA and glycerol NMR signals, in which

many transition peaks occur, are well described with the QM model-based approach.

Also, the dilution ratio of MDEA has an influence on the measured NMR signal and the corresponding self-diffusion coefficient as the viscosity decreases with increasing amount of D_2O . The change in viscosity affects the molecular mobility that causes a significant change in the self-diffusion coefficient and the relaxation rate, so that the peaks are much narrower at a high dilution ratio. For the measurement of self-diffusion coefficients in high-viscosity samples, higher durations for the gradient

pulse and diffusion times are required to obtain the same attenuation as for low-viscosity samples (cf. acquisition parameters of the PFG ^1H NMR experiments in the supporting information). High δ and Δ values, however, lead to phase distortions in the acquired NMR signal due to J-modulation. This is observed in Figure 4a and explains the larger deviations between the modelled and the measured NMR signal of the high concentrated MDEA sample compared to the MDEA sample with high dilution ratio (cf. Figure 4d). Model misspecifications due to J-modulation in the acquired NMR signal is a well-known problem. The use of different, more complicated, relaxation models to set up the chemical signatures could account for this effect. However, in our method, we use a simple relaxation model that provides sufficiently accurate results for our purposes.

In the high-field NMR spectrum of the Bruker standard (cf. Figure 4e), the peak for water appears at ≈ 4.7 ppm, and the peaks of the methanol doublet, resulting from the coupling between the protons from the CH_3 group and the ^{13}C atom, appear at ≈ 3.2 and ≈ 3.5 ppm. Due to the high coupling constant between the protons of the CH_3 group and the ^{13}C atom of approximately 142 MHz, the two peaks of the methanol doublet are spread in the benchtop NMR spectrum and appear at ≈ 1.7 and ≈ 5 ppm. Besides water and methanol, the sample consists of D_2O and GdCl_3 . Both species do not appear in the NMR signal but GdCl_3 with its strong paramagnetic properties leads to an increase in the peak width, which minimises the influence of imperfections in the shim and makes the model fit better, as can be seen in the residual spectra on the top left of each plot.

Water and acetonitrile are in principle the simplest species of these set of chemical components, as both are used as pure components and their individual NMR signals show only one peak. The difficulty, however, of using high-field NMR spectroscopy for the estimation of self-diffusion coefficients of pure components is the high signal intensity that can lead to radiation damping and thus phase distortions in the acquired NMR signals. Radiation damping is not included in our model and hence introduces errors into the model-based approach. To minimise radiation damping, smaller NMR sample tubes with an outer diameter of 2.5 mm were used. However, the occurrence of radiation damping cannot be completely excluded, which may explain the relatively high residuals in the high-field NMR signals of water and acetonitrile that can be observed in Figure 4c,f. Also, small fluctuations in the magnetic field may have an impact on the quality of the high-field NMR signals and thus on the residuals between the measured and the modelled NMR signals. When using benchtop NMR spectrometers, radiation damping and magnetic field

inhomogeneities play a subordinate role, because of the lower magnetic field strength compared to the high-field NMR spectrometer. The application of the model-based approach for the estimation of self-diffusion coefficients from experimental PFG ^1H NMR data was first validated using the high-field NMR experiments of the simple mixtures. Self-diffusion coefficients were estimated using the model-based approach. For comparison, all spectra were analysed with direct integration. If the NMR signal is free from overlapping peaks, direct integration is the method of choice for the estimation of the signal intensities, because this method is robust against lineshape distortions. In contrast, the model-based approach assumes ideal signatures for the different species, and if the lineshapes of the peaks are distorted, the estimated signal intensities may be biased. To overcome this problem in model-based fitting, we use the new approach mentioned in Section 2 to adjust the residuals between the measured and the modelled NMR signal for all high-field NMR signals. For both procedures, model-based fitting and direct integration, the mean values of the self-diffusion coefficients for all species in the simple mixtures were calculated from the 10 repetitions of the PFG ^1H NMR experiments. Figure 5 shows the relative deviation between the mean values of the self-diffusion coefficients estimated with the model-based approach and with direct integration, for each species of the simple mixtures. In general, a good agreement between both analysis method is achieved. The differences between the model-based approach and direct integration for the self-diffusion coefficient of water in the Bruker standard sample are approximately zero. The mean values of the self-diffusion coefficients of glycerol and pure water show the largest deviations.

In the PFG ^1H NMR data set of water, phase errors near the peak edges in the NMR spectrum with the

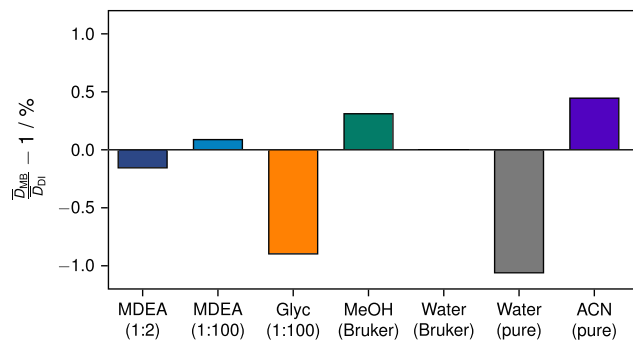


FIGURE 5 Relative deviation between the mean value of the self-diffusion coefficient determined with the model-based approach (MB) and with direct integration (DI); the mean was calculated from 10 high-field PFG ^1H NMR experiments, respectively

lowest gradient strength are observed. This is caused by radiation damping and leads to a reduction of the peak areas, and the estimated self-diffusion coefficient is slightly biased. Radiation damping was observed in the samples of water and acetonitrile, which explains the larger deviations between the analysis methods for these samples.

Also, errors in the model fit or in the adjustment of the residuals can contribute to deviations between the model-based approach and direct integration. Nevertheless the deviation of the mean values of the self-diffusion coefficients for all species are less than 1%. In the following analysis, the high-field results of the model-based approach were used to assess the benchtop PFG ^1H NMR experiments.

Figure 6 shows the self-diffusion coefficients for all species estimated from the 10 benchtop PFG ^1H NMR experiments plotted against the self-diffusion coefficients determined from the high-field PFG ^1H NMR experiments. For a better visualisation, the minor plots show an expansion of the interesting regions for all species. The self-diffusion coefficients of the different species lie in a range between $\approx 0.027 \cdot 10^{-9} \text{ m}^2 \text{ s}^{-1}$ for MDEA (1:2)

and $\approx 4.5 \cdot 10^{-9} \text{ m}^2 \text{ s}^{-1}$ for acetonitrile. Pure water at 28.7°C has a self-diffusion coefficient of $\approx 2.51 \cdot 10^{-9} \text{ m}^2 \text{ s}^{-1}$. The higher the dilution of MDEA and glycerol the higher the self-diffusion coefficients of the species as the viscosity decreases and thus the molecular motion of the spins increases. In this parity plot, the diagonal shows the ideal case, in which the self-diffusion coefficients determined from the benchtop PFG ^1H NMR signals match those from the high-field PFG ^1H NMR experiments. Since hardly any differences between the self-diffusion coefficients determined from the benchtop and the high-field PFG ^1H NMR experiments can be identified in the original view, the results for all species are additionally shown enlarged in the minor plots. Each point in the vertical dimension corresponds to an estimated self-diffusion coefficient from an experiment acquired with the benchtop NMR spectrometer and a horizontal point corresponds to a result of a high-field NMR experiment. In the centre, the mean values of the self-diffusion coefficients of all 10 benchtop and high-field PFG ^1H NMR experiments are plotted. The error bars were calculated for each species by MCMC sampling as described in Section 3.3. It can be seen that the

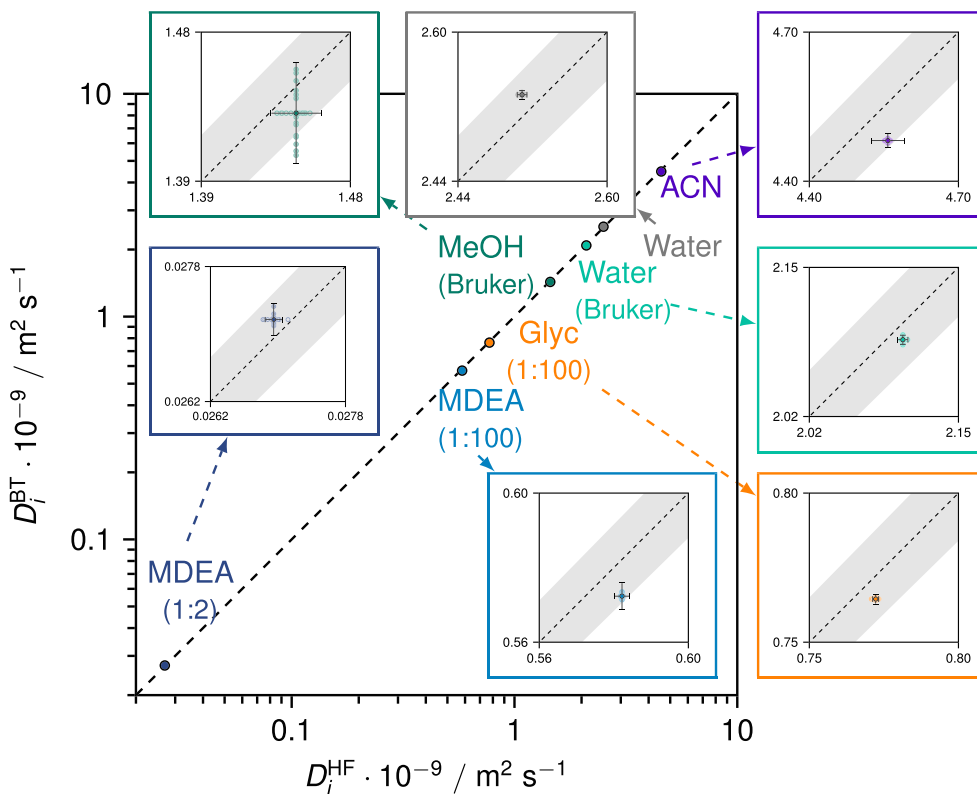


FIGURE 6 Parity plot of the self-diffusion coefficients of MDEA (1:2) (●), MDEA (1:100) (●), glycerol (1:100) (●), methanol (Bruker standard) (●), water (Bruker standard) (●), water (pure) (●), and acetonitrile (pure) (●) at 28.7°C estimated from the 10 benchtop (●) and high-field (○) PFG ^1H NMR experiments, respectively, with the model-based approach; the minor plots show an expansion of the region of interest for each species; the dashed lines correspond to the diagonal of the parity plot; the error bars were calculated according to the procedure described in Section 3.3 and are centred around the mean values of the self-diffusion coefficients determined from the benchtop and high-field PFG ^1H NMR experiments, respectively. The grey shaded area shows the 2% deviation from the diagonal (ideal) line

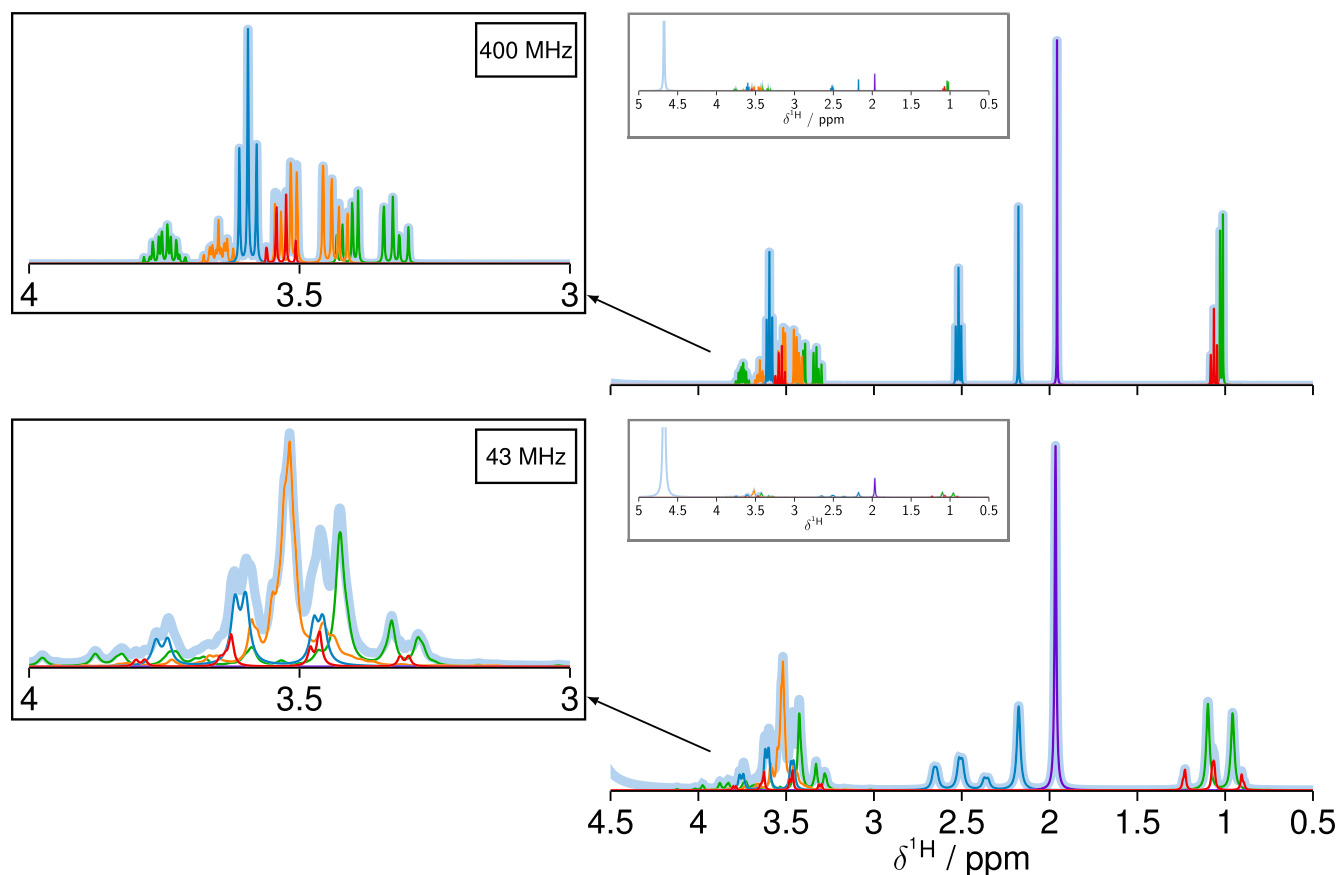


FIGURE 7 ^1H NMR spectrum of the complex mixture 1 measured with the high-field (top) and the benchtop (bottom) NMR spectrometer together with the signatures of MDEA (—), glycerol (—), 1,2-propylene glycol (—), ethanol (—), and acetonitrile (—) obtained with the model-based approach. The plot in the middle shows the full view of each ^1H NMR spectrum, respectively

scatter of the self-diffusion coefficients between the repetitions of the PFG ^1H NMR experiments are covered by the error bars. Overall, a good agreement between the self-diffusion coefficients determined from the benchtop and the high-field PFG ^1H NMR experiments is achieved. All benchtop values show deviations less than 2 % compared to the high-field values for the self-diffusion coefficients. This deviation has an order of magnitude equal to the accuracy of standard NMR measurements. Thus, the model-based approach was found to be well suited for the determination of self-diffusion coefficients from benchtop PFG ^1H NMR data sets and was in the following applied to a more complex mixture.

4.3 | Complex mixtures

After the validation of the model-based approach for the determination of self-diffusion coefficients from benchtop PFG ^1H NMR measurements using the simple mixtures and pure components, a more complex mixture,

consisting of MDEA, glycerol, 1,2-propylene glycol, ethanol, and acetonitrile highly diluted in water, was studied. Two mixtures with different compositions were prepared, and for both mixtures, 10 PFG ^1H NMR experiments were measured on the high-field and on the benchtop NMR spectrometer, respectively. Figure 7 shows a ^1H NMR spectrum of the first mixture, measured with the high-field and the benchtop NMR spectrometer. On the high-field ^1H NMR spectrum, the peaks of all species can be well separated. However, the analysis is not trivial, since peak overlapping at about 3.5 ppm can also be observed. The signal intensity of glycerol, for example, can be hardly determined using standard analysis methods, such as direct integration. The same ^1H NMR spectrum acquired with the benchtop NMR spectrometer is plotted in the bottom view of Figure 7. In the region at about 3.5 ppm, significant overlap of the peaks of glycerol, 1,2-propylene glycol, MDEA, and ethanol is observed. Also, the triplets of ethanol and 1,2-propylene glycol at about 1 ppm strongly overlap. With the model-based approach, it is possible to resolve the signals of all

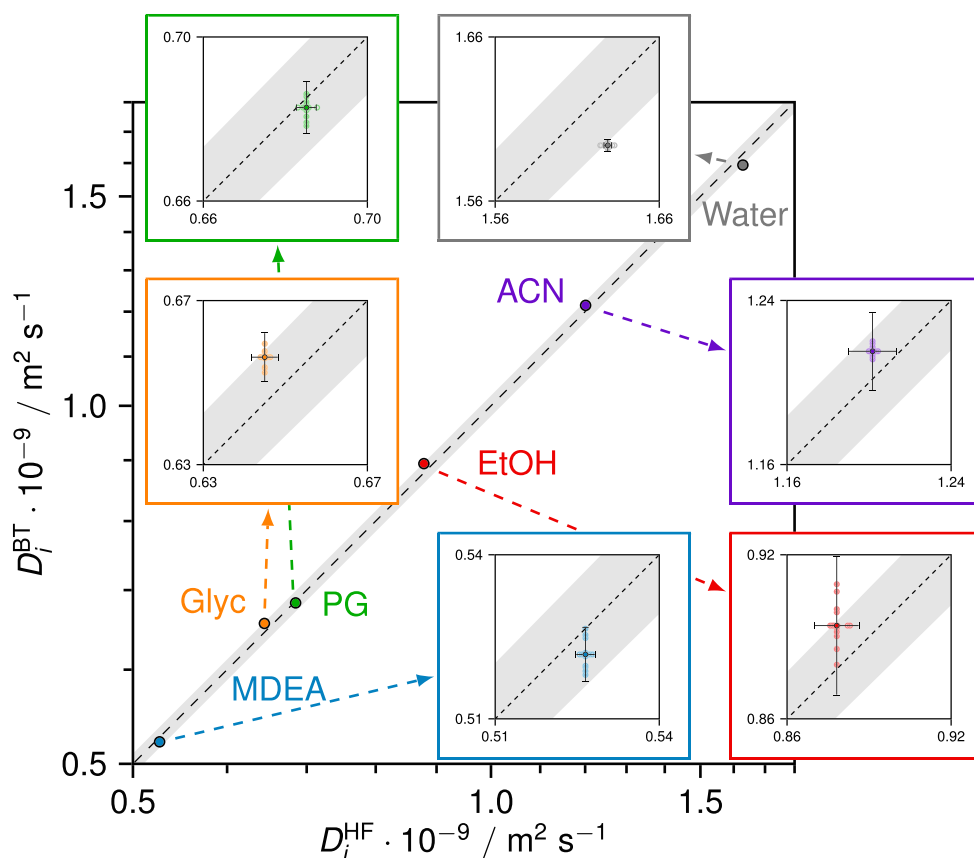


FIGURE 8 Parity plot of the self-diffusion coefficients of MDEA (●), glycerol (●), 1,2-propylene glycol (●), ethanol (●), acetonitrile (●), and water (●) at 28.7 °C estimated from the 10 benchtop (●) and high-field (○) PFG ¹H NMR experiments of the complex mixture 1, respectively, with the model-based approach; the minor plots show an expansion of the region of interest for each species; the dashed line corresponds to the diagonal of the parity plot; the error bars were calculated according to the procedure described in Section 3.3 and are centred around the mean values of the self-diffusion coefficients determined from the benchtop and high-field PFG ¹H NMR experiments, respectively. The grey shaded area shows the 2 % deviation from the diagonal (ideal) line

species as can be seen by the signatures, which are plotted with different colours.

All PFG ¹H NMR data sets of both complex mixtures were analysed with the model-based approach. The self-diffusion coefficients of all species were determined from the signal intensities using Equations (4) and (5) for the high-field and benchtop PFG ¹H NMR data sets, respectively. In Figure 8, the self-diffusion coefficients of each species in complex mixture 1 estimated from benchtop PFG ¹H NMR data sets are plotted against those obtained from the high-field PFG ¹H NMR data sets. The minor plots show an expansion of the interesting region for each species. The vertical and horizontal values correspond to the self-diffusion coefficients, estimated from each of the 10 repetitions of the benchtop and high-field PFG ¹H NMR experiments, respectively. The error bars are estimated with the MCMC method described in Section 3.3 and are centred around the mean values of the self-diffusion coefficients, for all

benchtop and high-field PFG ¹H NMR experiments. All values for the self-diffusion coefficients lie close to the diagonal, for example, a good agreement between the benchtop and the high-field results is achieved. A slight systematic bias between the benchtop and the high-field measurements was observed only for glycerol and water. For water, the differences can be explained by radiation damping that occurred in the high-field NMR signals due to the high concentration of water in the samples. The differences for glycerol can be explained by possible errors in the model fit, which also led to differences between the model-based approach and direct integration (cf. Figure 5). However, the differences between the self-diffusion coefficients estimated with the benchtop and the high-field NMR spectrometers are less than 2 %, even in these cases. Thus, these results demonstrate that the model-based approach is able to obtain quantitatively accurate measurements of the self-diffusion coefficients in this complex mixture,

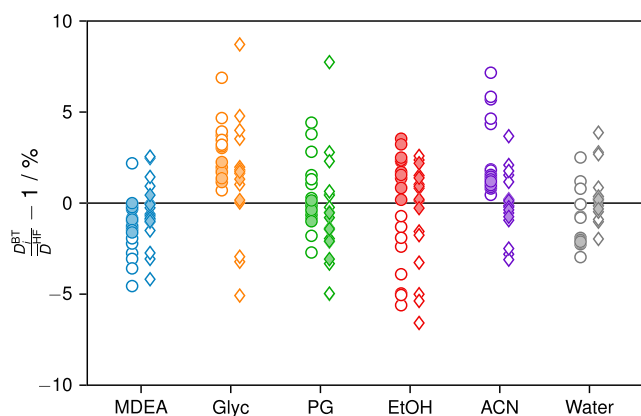


FIGURE 9 Relative deviation between the self-diffusion coefficients of each species in complex mixtures 1 (circles) and 2 (diamonds), determined with the model-based approach (●/◆) and qGSD (○/◇) from the benchtop measurements and the mean value of all high-field PFG ^1H NMR measurements determined with the model-based approach

despite significant overlap in the peaks associated with the different components.

For a better assessment of the use of benchtop NMR spectrometers for the measurement of self-diffusion coefficients in complex mixtures, a second composition of the species in the mixture was studied. In addition, all benchtop PFG ^1H NMR data sets of both mixtures were analysed with qGSD. Since a complete resolving of all peaks by qGSD was not possible, only one peak per species was considered for the qGSD analysis. The intensities of ethanol and 1,2-propylene glycol were determined using the peaks at about 1 ppm, for the intensities of MDEA and acetonitrile the singlets at 2.2 and 2 ppm were analysed and for glycerol the largest peak at about 3.55 ppm was used. The benchtop NMR results of both analysis methods, model-based fitting and qGSD, were compared to the corresponding high-field results. For this comparison, an average value for the self-diffusion coefficients of all species was calculated from the high-field PFG ^1H NMR experiments, for both complex mixtures, respectively. Figure 9 shows the relative deviations between the self-diffusion coefficients estimated from the benchtop PFG ^1H NMR experiments and the mean value of the high-field self-diffusion coefficients for the model-based approach and qGSD and for both complex mixtures.

The filled circles show a comparison between the self-diffusion coefficients of complex mixture 1 estimated from the benchtop and the high-field PFG ^1H NMR experiments, both analysed with the model-based approach, similar to Figure 8. A good agreement between the benchtop and the high-field self-diffusion coefficients is achieved. Also for complex mixture 2, displayed in diamonds, the self-diffusion coefficients estimated with the model-based approach lie close to the high-field self-

diffusion coefficients, that is, low relative deviations are observed. In addition, no systematic bias between the glycerol and water self-diffusion coefficients are observed, suggesting that the deviations that were observed for mixture 1 depend on the sample or may arise from a possible measurement error (e.g., due to a slight difference in temperature in either the high-field or benchtop NMR data). For both mixtures, slightly larger errors are obtained for ethanol and 1,2-propylene glycol as these components show significant peak overlap and lineshape distortions. The lineshape distortions are mainly caused by J-modulation due to the use of field-gradient pulses, and could possibly be corrected with a more advanced relaxation model. However, even for these molecules, the discrepancy between the high-field and benchtop data is small.

The self-diffusion coefficients, estimated from the different benchtop PFG ^1H NMR data sets with qGSD also agree well with the high-field results. However, the self-diffusion coefficients estimated with qGSD are more scattered compared to those obtained with the model-based approach, and more outliers are observed. This is also evident when the average RMS error for both mixtures is compared for both analysis methods. For the analysis with the model-based approach, an average RMS error of 0.011 is obtained, whereas the average RMS error for qGSD is significantly higher at 0.026. Thus, the model-based approach is better suited for the estimation of self-diffusion coefficients from benchtop NMR spectra with high accuracy.

5 | CONCLUSIONS

In this work, we presented a QM model-based approach for the estimation of self-diffusion coefficients from benchtop NMR signals that is capable of dealing with the complex NMR signals often observed when using benchtop NMR instruments.

The model-based quantification is combined with the Bayesian statistics that allows the inclusion of prior knowledge into the model-based fitting process. It has been demonstrated that even overlapping peaks from complex benchtop NMR signals can be resolved, when prior information of the model parameters from high-field measurements is available. However, in the absence of prior information on the model parameters, such as chemical shift values, the quantification results may be less accurate.

For the estimation of the self-diffusion coefficients, the inclusion of prior knowledge about the correlation of signal intensities via the Stejskal-Tanner equation did not turn out to be beneficial, since no significant improvement of the results was obtained and the implementation is even more complex compared to the standard fitting procedure. The inclusion of prior knowledge for a set of

NMR signals might be more advantageous for other types of measurements, such as the determination of kinetic constants, which is part of ongoing research.

With the model-based approach, the estimated self-diffusion coefficients from the benchtop PFG ^1H NMR data sets were in good agreement with the high-field self-diffusion coefficients, which were used to assess the performance of the model-based approach. Both benchtop and high-field analyses showed an agreement of 2 % for the self-diffusion coefficients determined from the simple and complex mixtures. Compared to the self-diffusion coefficients estimated with qGSD, the results obtained with the model-based approach showed less scatter, so that estimation of self-diffusion coefficients with the model-based approach is more robust compared to the qGSD analysis.

The use of benchtop instruments in combination with the model-based approach is a good option when accurate values for self-diffusion coefficients are required. With some prior information about the molecular structure of the species and their assignment, even benchtop signals with many different species can be analysed with low effort. Although knowledge of the species in the mixture is required, the method is particularly suitable for routine measurements with changing composition of the species in a mixture.

ACKNOWLEDGEMENTS

Financial support of the present study by the New Zealand Ministry of Business, Innovation and Employment (grant number UOCX1502) and Deutsche Forschungsgemeinschaft DFG (project number 310714510) is gratefully acknowledged. Open Access funding enabled and organized by Projekt DEAL.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1002/mrc.5300>.

ORCID

Ellen Steimers  <https://orcid.org/0000-0001-9999-3118>

Daniel J. Holland  <https://orcid.org/0000-0002-9192-5396>

REFERENCES

- G. Assemat, B. Gouilleux, D. Bouillaud, J. Farjon, V. Gilard, P. Giraudeau, M. Malet-Martino, *J. Pharm. Biomed. Anal.* **2018**, *160*, 268.
- H. Barjat, G. A. Morris, S. Smart, A. G. Swanson, S. C. R. Williams, *J. Magn. Reson., Ser. B* **1995**, *108*(2), 170.
- D. Bellaire, H. Kiepfner, K. Münnemann, H. Hasse, *J. Chem. Eng. Data* **2020**, *65*(2), 793. <https://doi.org/10.1021/acs.jced.9b01016>
- G. L. Bretthorst, *J. Magn. Reson. (1969)* **1990**, *88*(3), 533.
- R. H. Byrd, P. Lu, J. Nocedal, C. Zhu, *SIAM J. Sci. Comput.* **1995**, *16*(5), 1190. <https://doi.org/10.1137/0916069>
- L. Castañar, G. D. Poggetto, A. A. Colbourne, G. A. Morris, M. Nilsson, *Magn. Reson. Chem.* **2018**, *56*(6), 546. <https://doi.org/10.1002/mrc.4717>
- R. A. Chylla, J. L. Markley, *J. Biomol. NMR* **1995**, *5*(3), 245.
- C. Cobas, S. Sýkora, The bumpy road towards automatic global spectral deconvolution (GSD), in Poster at 50th enc conference, Asilomar (CA, USA), March 29 - April 4 **2009**, 15706.
- F. Dalitz, M. Cudaj, M. Maiwald, G. Guthausen, *Progr. Nuclear Magn. Reson. Spectrosc.* **2012**, *60*, 52. <https://doi.org/10.1016/j.pnmrs.2011.11.003>
- O. Drecun, A. Striolo, C. Bernardini, *Phys. Chem. Chem. Phys.* **2021**, *23*(28), 15224.
- L. Forster, M. Lutecki, H. Fordsmand, L. Yu, C. D'Agostino, *Mol. Syst. Design Eng.* **2020**, *5*(7), 1193. Publisher: The Royal Society of Chemistry.
- B. Gouilleux, J. Farjon, P. Giraudeau, *J. Magn. Reson.* **2020**, *319*, 106810.
- M. Holz, S. R. Heil, A. Sacco, *Phys. Chem. Chem. Phys.* **2000**, *2*(20), 4740.
- C. S. Johnson, *Progr. Nuclear Magn. Reson. Spectrosc.* **1999**, *34*(3-4), 203.
- S. Kern, K. Meyer, S. Guhl, P. Gräber, A. Paul, R. King, M. Maiwald, *Anal. Bioanal. Chem.* **2018**, *410*(14), 3349.
- E. Kriesten, F. Alsmeyer, A. Bardow, W. Marquardt, *Chem. Intell. Lab. Syst.* **2008**, *91*(2), 181.
- Y. Matviychuk, S. Haycock, T. Rutan, D. J. Holland, *Anal. Chimica Acta* **2021**, *1182*, 338944.
- Y. Matviychuk, E. Steimers, E. von Harbou, D. J. Holland, *J. Magn. Reson.* **2020**, *319*, 106814.
- Y. Matviychuk, E. Steimers, E. von Harbou, D. J. Holland, *Magn. Reson.* **2020**, *1*(2), 141.
- Y. Matviychuk, J. Yeo, D. J. Holland, *J. Magn. Reson.* **2019**, *298*, 35. <https://doi.org/10.1016/j.jmr.2018.11.010>
- E. R. McCarney, C. J. Breaux, P. M. Rendle, *Magn. Reson. Chem.* **2020**, *58*(7), 641. <https://doi.org/10.1002/mrc.4997>
- E. R. McCarney, R. Dykstra, P. Galvosas, *Magn. Resonan. Imaging* **2019**, *56*, 103.
- K. Meyer, S. Kern, N. Zientek, G. Guthausen, M. Maiwald, *TrAC - Trends in Anal. Chem.* **2016**, *83*(Part A), 39. <https://doi.org/10.1016/j.trac.2016.03.016>
- M. I. Miller, A. S. Greene, *J. Magn. Resonan. (1969)* **1989**, *83*(3), 525.
- A. Muhammad, G. Di Carmine, L. Forster, C. D'Agostino, *Chem. Phys. Chem.* **2020**, *21*(11), 1101. <https://doi.org/10.1002/cphc.202000267>
- M. Nilsson, *J. Magn. Resonan.* **2009**, *200*(2), 296.
- M. Nilsson, G. A. Morris, *Anal. Chem.* **2008**, *80*(10), 3777. <https://doi.org/10.1021/ac7025833>
- G. Pagés, V. Gilard, R. Martino, M. Malet-Martino, *Analyst* **2017**, *142*(20), 3771. Publisher: The Royal Society of Chemistry.
- J. Rönnols, E. Danieli, H. Freichels, F. Aldaeus, *Holzforschung* **2020**, *74*(2), 226. Publisher: De Gruyter.
- E. Steimers, Y. Matviychuk, A. Friebel, K. Münnemann, E. von Harbou, D. J. Holland, *Magn. Reson. Chem.* **2021**, *59*(3), 221.
- E. O. Stejskal, J. E. Tanner, *The J. Chem. Phys.* **1965**, *42*(1), 288. <https://doi.org/10.1063/1.1695690>

- [32] L. C. M. Van Gorkom, T. M. Hancewicz, *J. Magn. Resonan.* **1998**, *130*(1), 125.
- [33] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, I. Polat, Y. Feng, E. W. Moore, J. Vand erPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, S. Contributors, *Nature Methods* **2020**, *17*, 261.
- [34] J. H. Wang, *J. Am. Chem. Soc.* **1951**, *73*(9), 4181. <https://doi.org/10.1021/ja01153a039>
- [35] W. Windig, B. Antalek, *Chemom. Intell. Lab. Syst.* **1997**, *37*(2), 241.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: E. Steimers, Y. Matviychuk, D. J. Holland, H. Hasse, E. von Harbou, *Magn Reson Chem* **2022**, *60*(12), 1113. <https://doi.org/10.1002/mrc.5300>