

Contribution to Special Issue Dedicated to Paul Heitjans

Marc Christoffer Paulus, Anja Paulus, Rüdiger-Albert Eichel
and Josef Granwehr*

Independent component analysis combined with Laplace inversion of spectrally resolved spin-alignment echo/ T_1 3D ^7Li NMR of superionic $\text{Li}_{10}\text{GeP}_2\text{S}_{12}$

<https://doi.org/10.1515/zpch-2021-3136>

Received October 8, 2021; accepted October 12, 2021; published online November 4, 2021

Abstract: The use of independent component analysis (ICA) for the analysis of two-dimensional (2D) spin-alignment echo- T_1 ^7Li NMR correlation data with transient echo detection as a third dimension is demonstrated for the superionic conductor $\text{Li}_{10}\text{GeP}_2\text{S}_{12}$ (LGPS). ICA was combined with Laplace inversion, or discrete inverse Laplace transform (ILT), to obtain spectrally resolved 2D correlation maps. Robust results were obtained with the spectra as well as the vectorized correlation maps as independent components. It was also shown that the order of ICA and ILT steps can be swapped. While performing the ILT step before ICA provided better contrast, a substantial data compression can be achieved if ICA is executed first. Thereby the overall computation time could be reduced by one to two orders of magnitude, since the number of computationally expensive ILT steps is limited to the number of retained independent components. For LGPS, it was demonstrated that physically meaningful independent components and mixing matrices are obtained, which could be correlated with previously investigated material properties yet provided a clearer, better separation of features in the data.

***Corresponding author: Josef Granwehr**, Institute of Energy and Climate Research – Fundamental Electrochemistry (IEK-9), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany; and Institute of Technical and Macromolecular Chemistry, RWTH Aachen University, 52056 Aachen, Germany, E-mail: j.granwehr@fz-juelich.de

Marc Christoffer Paulus, Institute of Energy and Climate Research – Fundamental Electrochemistry (IEK-9), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany; and Institute of Technical and Macromolecular Chemistry, RWTH Aachen University, 52056 Aachen, Germany

Anja Paulus and Rüdiger-Albert Eichel, Institute of Energy and Climate Research – Fundamental Electrochemistry (IEK-9), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany; and Institute of Physical Chemistry, RWTH Aachen University, 52056 Aachen, Germany, E-mail: r.eichel@fz-juelich.de (R.-A. Eichel)

LGPS from two different batches was investigated, which showed substantial differences in their spectral and relaxation behavior. While in both cases this could be attributed to ionic mobility, the presented analysis may also clear the way for a more in-depth theoretical analysis based on numerical simulations. The presented method appears to be particularly suitable for samples with at least partially resolved static quadrupolar spectra, such as alkali metal ions in superionic conductors. The good stability of the ICA analysis makes this a prospect algorithm for preprocessing of data for a subsequent automatized analysis using machine learning concepts.

Keywords: independent component analysis; ion mobility; Laplace inversion; Li; NMR; solid electrolyte.

Dedicated to Prof. Paul Heitjans on the occasion of his 75th birthday

1 Introduction

Techniques for physical and chemical analysis of materials become ever more powerful. Advances in analytical instrumentation as well as step changes in microelectronics enable equipment for data acquisition and processing that provide large sets of data in a short amount of time. In particular, continuously acquired time series or multidimensional data sets can lead to quickly growing records of raw data [1]. Analyzing these data poses an increasing challenge, leading to the development of new strategies that themselves take advantage of rising processing power of computers. Such strategies include, for example, the use of neural networks or concepts of artificial intelligence to extract physically meaningful information from data, or filtering and data reduction to bring the data in a more easily accessible form for subsequent manual interpretation [2, 3]. Particularly sought after are methods that combine the opportunities of modern computers yet also enable a better understanding of the underlying mechanisms [4].

Nuclear magnetic resonance (NMR) is one example of a technique where instrumental advances enabled new multidimensional experimental protocols with large data sets being collected in a short amount of time [5–7]. Various materials properties, including local structure and dynamics as well as long-range spatial, motional and diffusive features, can be individually encoded and correlated [8, 9]. However, so far statistical data analysis methods have only rarely been employed in NMR of materials [10].

In the spin-alignment echo (SAE) NMR experiment, the decay of quadrupolar order is measured in a stimulated echo experiment as a function of a mixing time

[11–13]. It facilitates the direct acquisition of slow jump auto-correlation functions for ions with a spin $S > 1/2$ in a solid material. SAE has been used, for example, to study Li ion mobility in battery materials, but it is, generally speaking, limited to rather slow jump processes [14]. For ^7Li , SAE is able to resolve jump correlation time constants $\tau_c > 10^{-5}$ s. Faster processes provide a signal at the longitudinal relaxation time constant of quadrupolar order, $T_{1,Q}$. It is, therefore, important to distinguish between jump auto-correlation processes, which may directly provide information regarding diffusion and long-range mobility, and relaxation-limited processes in SAE experiments. In complex materials such as solid electrolytes, typically multiple different environments with different relaxation time constants are found. Therefore, a more selective distinction between jump correlations and relaxation than comparing a 1D SAE with a 1D longitudinal relaxation time NMR experiment is required.

The most common method for the analysis of SAE NMR data employs a fit of the data with one or multiple stretched exponential functions [13]. A more complete distribution of correlation time constants τ_c can be obtained using a Laplace inversion (ILT) [15, 16]. In systems where exchange processes cannot be excluded, in particular in solids, contributions with positive and negative sign in the distribution of time constants may be observed [17, 18]. Using an algorithm that does not employ a non-negativity constraint [19], it was shown that SAE data can be faithfully inverted [20]. Thereby, the residuals of the fit should be consistent with Gaussian white noise without any systematic features, thus the obtained distribution represents the full information available in the data [21].

In battery materials, the ionic mobility is of central importance and often represents a performance-limiting property [22]. Solid electrolytes are of particular interest. They are a necessary component for all-solid-state batteries, which are expected to solve some of the safety and durability problems of current lithium-ion batteries [23]. In terms of performance, their ionic mobility needs to compete with liquid electrolytes, therefore understanding diffusion pathways and exchange processes is of prime importance for material development and optimization. A substantial number of the most promising solid electrolyte candidates are sulfide based [24]. The superionic Li conductor $\text{Li}_{10}\text{GeP}_2\text{S}_{12}$ (LGPS) is one of the best performing material of this class [25]. Its structure and local ion transport properties have been studied in-depth [26, 27]. NMR has proved to be a useful technique due to the high sensitivity of ^7Li NMR and good resolution, which facilitated a large number of different techniques, including pulsed field gradient (PFG) NMR for the investigation of diffusive processes over micrometer lengthscales [28]. Nonetheless, long-range mobility, which includes exchange between crystallites or grains, is not yet fully understood.

Due to the high ionic mobility and associated fast hopping rates, SAE may, at first, appear unsuitable for the investigation of superionic conductors such as LGPS. However, it has been shown that ion conducting solids can show different modes of mobility, with fast local mobility processes causing a residual quadrupolar coupling that is not averaged to zero [29], and slower longer-range processes whose correlation time can be resolved by SAE [20, 30], even for LGPS [31]. Nonetheless, the difficulty of distinguishing between jump correlation time constants τ_c and $T_{1,Q}$ -limited processes still remains.

Taking advantage of the commonly short ^7Li longitudinal relaxation time constants of superionic conductors that allows short experiment repetition times, a straightforward option to identify $T_{1,Q}$ limited processes is presented by a two-dimensional (2D) experiment with an inversion–recovery or a saturation–recovery encoding block combined with an SAE readout block. This allows the correlation of longitudinal relaxation times, T_1 , with τ_c [32, 33]. Signals on the diagonal, where $\tau_c = T_{1,Q}$, cannot be interpreted as jump autocorrelation time constants, but could be either very fast moving ions or very slow ions with jump time constants that are longer than $T_{1,Q}$. Experiments on LGPS revealed a significant signal fraction close to, but slightly below this diagonal. In this case, it would be desirable to have an independent test of the assignment without having to perform additional experiments. In principle, such a test can be obtained by recording the echo transiently [34]. Fourier transform of the measured interferogram provides a spectrum. In combination with a 2D Laplace inversion, for each data point in the $\tau_c - T_1$ correlation map, a ^7Li NMR spectrum is obtained of the nuclei giving rise to the respective signal component [32]. While clear variations are apparent in the NMR spectra between different regions of the 2D correlation map, the quadrupolar coupling constants remain surprisingly similar for LGPS, despite having a mixture of tetragonal and orthorhombic LGPS phases in the material. The most significant variation in the spectra is obtained between the non-central quadrupolar transitions and the central transition. These two features not only show a different shape but also a different signal phase [35].

Here, the applicability of independent component analysis (ICA) in combination with Laplace inversion to separate different spectral components in a 3D τ_c versus T_1 versus ^7Li NMR data set is investigated. 2D maps are obtained for τ_c versus T_1 , whose features are correlated with spectral features from Fourier transform of the detected echo transient. Different protocols are applied to test if ICA can be used for data reduction by providing 2D-SAE/ T_1 maps of statistically as independent NMR spectra as possible.

2 Methodology

This section summarizes the flow of data and aspects relevant for the application of the algorithm. A more in-depth explanation of the mathematical background can be found elsewhere [15, 19, 36].

2.1 Independent component analysis

Independent component analysis (ICA) aims to find a linear representation of non-Gaussian data whose components are statistically as independent as possible [37]. A statistically latent variable model is used to represent the ICA. Observable variables y are considered as linear mixtures of n not directly observable independent variables s [38]. Mathematically this can be expressed as

$$y_j = a_{j,1}s_1 + a_{j,2}s_2 + \dots + a_{j,n}s_n, \quad (1)$$

where $a_{j,k}$ represent mixing elements, $j = 1, \dots, N_y$ with N_y the number of observations, and $k = 1, \dots, n$. More generally, observed signals can be a data vector $\mathbf{y}_j \in \mathbb{R}^{1 \times N_d}$ with N_d elements, such as a spectrum. In this case, the independent component (IC) is also a vector, $\mathbf{s}_k \in \mathbb{R}^{1 \times N_d}$. In matrix notation, this can be rewritten as

$$\mathbf{Y} = \mathbf{A}\mathbf{S}, \quad (2)$$

where $\mathbf{A} \in \mathbb{R}^{N_y \times n}$ denotes the mixing matrix with elements $a_{j,k}$, $\mathbf{Y} \in \mathbb{R}^{N_y \times N_d}$ is the data matrix whose rows represent one observed signal, or mixture, each, and $\mathbf{S} \in \mathbb{R}^{n \times N_d}$ represents a matrix whose rows are the independent components \mathbf{s}_k . The ICA is thus a generative model that describes how observable data is generated by mixtures of the independent components.

The matrices \mathbf{A} and \mathbf{S} are determined considering the general assumptions that (i) the components \mathbf{s}_k are statistically independent; (ii) the components have a non-Gaussian distribution; and (iii) the unknown mixing matrix is square. However, since two unknowns have to be determined, the result is subject to further restrictions: (i) the variance of the independent components cannot be determined, since a scalar multiplier of one of the components \mathbf{s}_k can always be erased by a corresponding multiplier in a column of \mathbf{A} . As a result, the amplitude of the independent components is usually determined by assuming that each of them has unity variance; (ii) the sign cannot be determined because the independent components can always be multiplied by -1 without affecting the model; and (iii) the order of the components cannot be determined.

To calculate \mathbf{A} and \mathbf{S} , an estimate of negentropy is used as contrast function, *i.e.* as a quantitative measure of the non-Gaussianity of a random variable [37]. For its calculation, the observable random variables \mathbf{Y} are first centered, *i.e.* the arithmetic mean $E\{y\}$ is subtracted from the observed mixtures. At the end of the ICA procedure, the centering is undone. As a next preprocessing step, the centered vectors are whitened by a linear transformation that ensures the independent variables are uncorrelated with unity variance. In a further step the dimension of the data is reduced. Robust results were obtained by retaining only as many eigenvalues calculated during the whitening step as independent components are determined. This reduces noise and prevents overlearning, which may sometimes occur for ICA [39]. As the final step, ICA finds a projection for \mathbf{Y} that maximizes non-Gaussianity.

A practical aspect for preventing overlearning when ICA is used in combination with Laplace inversion is to avoid outliers in the density distribution obtained from the ILT step. Sharp spikes in the ILT output can sometimes occur as a consequence of underregularization, insufficiently dense

sampling in the inverted dimension, or data that cannot be modeled with the employed kernel [19]. Before the inversion becomes unstable and produces nonphysical results upon substantial underregularization, sharp spikes may evolve in the distribution, which show the correct integral and position of existing modes, yet with a too narrow width. More pronounced under regularization causes the appearance of spikes with bipolar sign, where overshooting in one direction gets compensated by undershooting in the other direction. Such features in the input data of the employed FastICA algorithm are known to cause noisy and unstable results [39], hence they should be prevented during the ILT step.

2.2 Laplace inversion

On discrete experimental data in one dimension, given a vector $\mathbf{x} \in \mathbb{R}^{N_{x_1} \times 1}$ of N_{x_1} data points that were recorded as a function of an independent variable t_p with $p = 1, \dots, N_{x_1}$, the problem of Laplace inversion, or discrete inverse Laplace transform (ILT), can be written as

$$\mathbf{x} = \mathbf{K}\mathbf{g} + \mathbf{e}. \quad (3)$$

$\mathbf{g} \in \mathbb{R}^{M_1 \times 1}$ is an underlying discretized distribution function of some variable $\tau_{1,l}$, with $l = 1, \dots, M_1$. \mathbf{g} is connected to the data via kernel $\mathbf{K} \in \mathbb{R}^{N_{x_1} \times M_1}$, with matrix elements $K_{p,l} = f(t_p, \tau_l)$ given by kernel function $f(t_p, \tau_l)$. $\mathbf{e} \in \mathbb{R}^{N_{x_1} \times 1}$ represents a vector of additive noise with zero mean, which is commonly assumed to be independent and identically distributed Gaussian white noise. In NMR, where often for practical purposes an assumption of frequency and signal amplitude independent noise is justified, the signal can be scaled without loss of generality such that the noise vector has a unity variance. This data scaling is assumed in the following and simplifies parametrization of the regularization.

An extension to multiple dimensions is mathematically straightforward, but computationally demanding. Assuming an R -dimensional data set $\mathbf{X} \in \mathbb{R}^{N_{x_1} \times \dots \times N_{x_R}}$ with kernel $\mathbf{K}_r \in \mathbb{R}^{N_{x_r} \times M_r}$ along dimension r , an R -dimensional distribution $\mathbf{G} \in \mathbb{R}^{M_1 \times \dots \times M_R}$ and an R -dimensional error function $\mathbf{E} \in \mathbb{R}^{N_{x_1} \times \dots \times N_{x_R}}$, vectorization of \mathbf{X} , \mathbf{G} and \mathbf{E} , along with an outer product of the kernel matrices,

$$\mathbf{x} = \text{vec}(\mathbf{X}) \in \mathbb{R}^{N_x \times 1} \quad (4)$$

$$\mathbf{g} = \text{vec}(\mathbf{G}) \in \mathbb{R}^{M \times 1} \quad (5)$$

$$\mathbf{K} = \mathbf{K}_R \otimes \mathbf{K}_{R-1} \otimes \dots \otimes \mathbf{K}_1 \in \mathbb{R}^{N_x \times M} \quad (6)$$

$$\mathbf{e} = \text{vec}(\mathbf{E}) \in \mathbb{R}^{N_x \times 1} \quad (7)$$

retrieves Eq. (3). However, the dimension of the different vectors and matrices corresponds to the product of the size of the constituting matrices, $N_x = \prod_{r=1}^R N_{x_r}$ and $M = \prod_{r=1}^R M_r$, hence the demand on computational resources may quickly become prohibitive.

The purpose of Laplace inversion is to obtain an estimate $\hat{\mathbf{g}}$ for the unknown distribution function. Since the direct inversion of Eq. (3) is ill-conditioned for kernels such as the exponential function, a generalized Tikhonov regularization is applied to stabilize the inversion, and the functional to be minimized is

$$\hat{\mathbf{g}} = \underset{\mathbf{g}}{\text{argmin}} \{ \|\mathbf{K}\mathbf{g} - \mathbf{x}\|_2^2 + \lambda^2 \|\mathbf{A}\mathbf{g}\|_2^2 \}, \quad (8)$$

where $\|\cdot\|_2$ is the Frobenius norm and λ is a global scaling factor for the regularization. The regularization matrix $\mathbf{A} \in \mathbb{R}^{M \times M}$ contains a uniform penalty term [16] and a penalty for zero crossing, which stabilizes the inversion. Thereby, robust vectorized distributions are obtained without

non-negativity constraint. Parametrization of $\mathbf{\Lambda}$ is done algorithmically as described in Ref. [19] without adding any free parameters, and scaling of the data to unity variance allows setting $\lambda = 1$, unless stated otherwise. Equation (8) was solved iteratively. To obtain 2D correlation maps, the vectorization step of Eq. (5) was reversed on $\tilde{\mathbf{g}}$.

In principle, it would also be beneficial to regularize non-inverted dimensions as well [20]. However, in the case of 2D inversion with a third spectral dimension, the demand on computation resources gets very high, since the kernel and regularization matrices become too large. Hence at this point only regularization along inverted dimensions was employed, and the spectral dimension was sampled point-by-point.

If sampling of the distribution in the inverted dimension is sufficiently dense for a given signal-to-noise ratio of the data [40], the default regularization parametrization generally provides a robust inversion [19]. Occurrence of sharp spikes in the ILT output, as mentioned in Section 2.1, then indicates either a frequency or signal dependent non-uniform noise level, or data that is not compatible with a particular kernel. In magnetic resonance, examples for signals that are incompatible with an exponential kernel are oscillations on top of a relaxing signal [41] or diffraction of diffusion phenomena in PFG NMR [42]. On the other hand, trying to fit a Gaussian signal decay with an exponential kernel may cause oscillations in the distribution, but these are stable towards changes in the regularization and do not show a width of only a single data point. The onset of spikes may be automatically identified by a considerable increase of the norm of the resulting distribution upon changing the regularization, cf. the L-curve method for adjusting the regularization parameter [43].

2.3 Analysis protocol for experimental data

The pulse sequence employed in the presented experiments,

$$\pi - t_d - \left(\frac{\pi}{2}\right)_x - t_p - \left(\frac{\pi}{4}\right)_y - t_m - \left(\frac{\pi}{4}\right) - t_p - \text{echo},$$

contains two independently varied delays, a mixing time t_m that produced a vector $\mathbf{t}_m \in \mathbb{R}^{N_{x1} \times 1}$ as an inner loop, and a recovery delay t_d that produced a vector $\mathbf{t}_d \in \mathbb{R}^{N_{x2} \times 1}$ as an outer loop. As the innermost loop, which does not come with an additional measurement time overhead and which was not inverted by ILT, the echo transient was recorded. After Fourier transform, it provided a discrete spectrum of size N_F . For ILT, along both inverted dimensions a kernel function

$$f(t_p, \tau_i) = \exp\left(-\frac{t_p}{\tau_i}\right) \quad (9)$$

was employed. Inversion of the t_m dimension, the spin-alignment echo dimension, produces a correlation time constant $\tau_c \in \mathbb{R}^{M_1 \times 1}$. Inversion of the t_d dimension, the recovery dimension, produces a distribution of the longitudinal relaxation time constant, $\mathbf{T}_1 \in \mathbb{R}^{M_2 \times 1}$. A simultaneous 2D inversion of both dimension produces a 2D SAE- T_1 correlation map, which can be represented as a $M_1 \times M_2$ matrix. The discrete values of both τ_c and \mathbf{T}_1 are user selectable. Care must be taken that the range is chosen large enough that the distribution shows negligible spectral density outside this region, and that the data points are sufficiently dense that each mode consists of more than just one point.

While nuclei in different environments vary in their NMR spectrum, the same supposedly applies to relaxation and motional correlation times. Therefore some correlation is expected

between the spectral dimension and the relaxation dimensions. To identify statistically independent components, either the spectral or the relaxation domain can be chosen as the observed signal. The relaxation dimensions are regularized, leading to smooth features in the 2D correlation map, which suggests this to be the signal dimension. However, it will also be tested if the spectral dimension, which is built point-by-point by conducting a 2D Laplace inversion individually on each frequency component, could be chosen as the signal dimension as well. In either case the 3D data matrix, where the echo transient is preprocessed with a discrete Fourier transform and phase corrected, is converted into a 2D matrix by vectorizing the τ_{rmc} and the T_1 dimensions, leaving the spectral dimension as the second dimension. This creates a matrix $\mathbf{Y} \in \mathbb{R}^{M_1 M_2 \times N_F}$. If the spectrum is used as the signal dimension, then $\mathbf{Y} = \hat{\mathbf{Y}}$ is to be used in Eq. (2), and the mixing matrix $\mathbf{A} \in \mathbb{R}^{M_1 M_2 \times n}$ contains the vectorized 2D correlation maps as its columns and the independent components in $\mathbf{S} \in \mathbb{R}^{n \times N_F}$ represent ^7Li NMR spectra. If, on the other hand, the correlation map represents the signal dimension, then $\mathbf{Y} = \hat{\mathbf{Y}}^T$, and the meaning and dimensions of \mathbf{A} and \mathbf{S} are interchanged. In both cases, this application of the ICA is referred to as ILT-then-ICA, since the independent component analysis acts on data that is preprocessed with a Laplace inversion.

Another possibility would be to swap the ILT and the ICA step, i.e. first conduct an ICA that is followed by an ILT (ICA-then-ILT). In this case, the t_m and the t_d dimensions of the 3D data matrix are vectorized to create a 2D matrix $\hat{\mathbf{Y}} \in \mathbb{R}^{N_{s_1} N_{s_2} \times N_F}$. Again, the echo transient is Fourier transformed and phase corrected first, although, in principle, the procedure could also be applied directly on the echo transient. Analogous to the ILT-then-ICA case, either the spectral dimension or the SAE/relaxation dimension could be considered the signal dimension. In the former case, $\mathbf{Y} = \hat{\mathbf{Y}}$ in Eq. (2) and $\mathbf{A} \in \mathbb{R}^{N_{s_1} N_{s_2} \times n}$, otherwise $\mathbf{Y} = \hat{\mathbf{Y}}^T$ and $\mathbf{S} \in \mathbb{R}^{n \times N_{s_1} N_{s_2}}$. In both cases, either \mathbf{A} or \mathbf{S} , whichever contains the correlation data, must then be Laplace inverted for each independent component to obtain n independent 2D correlation maps. Since magnetic resonance data are often sparse or contain redundant information [44], the potential for data compression using the ICA-then-ILT approach is significant. Thereby only a reduced number of correlation data sets is retained, which are then further processed with the computationally more demanding Laplace inversion.

3 Experimental

3.1 Sample preparation

Two different LGPS samples were investigated, which were both synthesized as described by Kamaya *et al.* [25]. The sample referred to as sample 1 was from the dominantly tetragonal (*t*-LGPS) batch described in Ref. [32], and the sample referred to as sample 2 was also dominantly tetragonal from the batch described in Ref. [33]. Using ^{31}P NMR and Rietveld refinement of X-ray diffraction data, a tetragonal fraction of 75% was determined for sample 1 and 79% for sample 2.

3.2 NMR experiments

Experiments were conducted using the 2D-SAE/ T_1 correlation sequence with transient detection of the echo, thereby producing a 3D data set [32]. The recovery delay t_d and the mixing time t_m were varied independently in consecutive repetitions of the experiment. The echo time t_p was set to

10 μ s. For sample 2, the inversion π pulse at the beginning of the sequence was replaced with a saturation sequence of 1024 not equally spaced $\pi/2$ pulses, applied during a period of 0.5 s. As variable delays, a series of 32 logarithmically spaced t_m values between 10 μ s and 10 s, and 16 logarithmically spaced t_d values also between 10 μ s and 10 s were applied. Acquisition was started a few data points before the echo maximum, immediately after the end of the spectrometer dead time, to maintain additional flexibility regarding data processing. The acquisition delay was set such that the quadrupolar echo coincided with a data point to simplify phase correction of the data. The experiment on sample 1 was conducted at a magnetic field of 9.4 T and the experiment on sample 2 at a field of 14.1 T, both on Bruker Avance-III spectrometers. The temperature was set to 293 K in both experiments.

3.3 Data processing

Data processing was performed using GNU Octave 4.0.3 [45]. For independent component analysis, FastICA 2.5 was used [36]. Laplace inversion was conducted using the algorithm and parametrization as described previously [19, 20]. For the inversion of the t_d dimension to τ_c as well as for the t_m dimension to T_1 , an exponential kernel was used. A reference spectrum was calculated by a 2D inversion using only the echo maxima. All other analysis procedures were done using the full 3D data set. Unless stated otherwise, all the pseudocolor plots are scaled using the signed square root of the absolute value of the respective data point to enhance weak features.

4 Results and discussion

Figure 1 provides an overview of the data and a reference τ_c versus T_1 map obtained by Laplace inversion of the echo maxima.

A typical spectrum after Fourier transform of the echo transient is shown in Figure 1a. It shows features that have been reported before for superionic conductors, with partially averaged quadrupolar coupling constants and almost negligible asymmetry [29, 46]. Notice the phase shifted central transition, which is characteristic for spin-alignment spectra [34]. Figure 1b shows the decay of the echo amplitude as a function of the SAE mixing time t_m and Figure 1c the signal recovery during t_d . In Figure 1d the 2D-SAE/ T_1 map is shown. The main feature is located at $\tau_c \sim 10^{-2}$ and $T_1 \sim 10^{-0.5}$ s, which had been discussed previously [32, 33]. In addition, a relaxation mode at $T_1 \sim 10^4$ s is visible, which fits the baseline of the recovery signal that does not decay to zero for $t_d \rightarrow \infty$. Finally, a bipolar pair of signals occurs at $\tau_c \approx 10^{-5}$ s. This signal is comparable to the shortest mixing time constant, $t_m = 10^{-5}$ s, hence it depends only on a very small number of data points along this dimension. Its origin will be discussed below.

Before the ICA was carried out, the centered and whitened raw data were subjected to a principal component analysis (PCA). The most intensive PCA eigenvalues and corresponding whitened eigenvectors are shown in Figure 2.

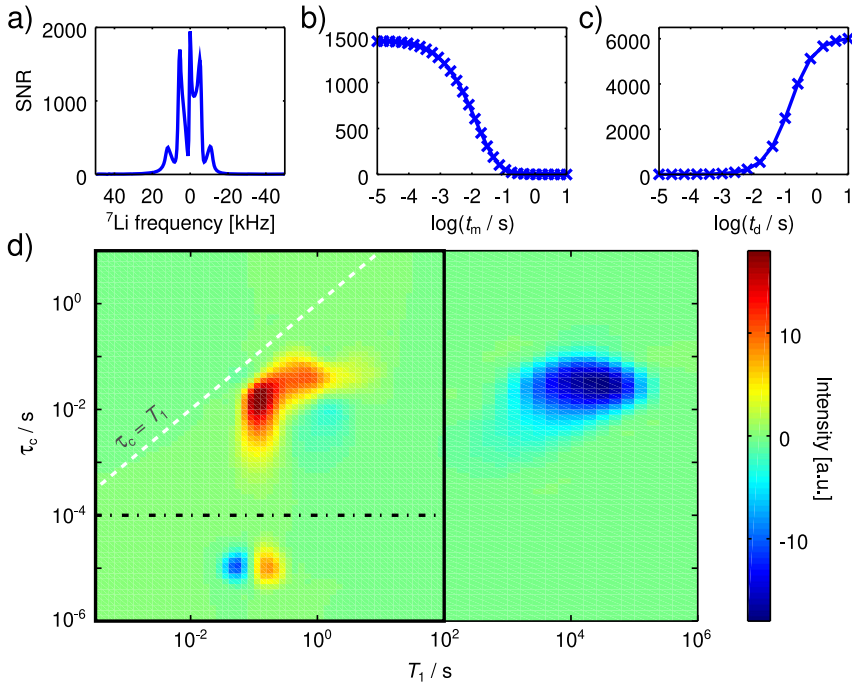


Figure 1: Data and Laplace inversion of 2D-SAE/ T_1 NMR experiment of LGPS sample 2. The data intensity is scaled relative to the respective thermal noise level.

a) ^7Li NMR spectrum obtained by Fourier transform of the echo transient of the repetition with maximum t_d recovery delay and minimum t_m mixing delay. b) Decay of the echo maximum along t_m , summed up along the t_d dimension. c) Recovery of the echo maximum along t_d , summed up along the t_m dimension. d) 2D-SAE/ T_1 pseudocolor map. The baseline of the recovery dimension, where the signal approaches a maximum for long t_d , is fitted by extending the T_1 dimension well into regions where no relaxation time constants are expected anymore. This baseline is responsible for the relaxation mode at around 10^4 s . The color map is scaled linearly. The white dashed line marks data points with $\tau_c = T_1$. The region at $T_1 < 10^2 \text{ s}$, enclosed by a black solid box, represents the default data points used for ICA if ILT precedes the ICA step. The black dash-dotted line separates the regions at short and long τ_c for a separate discussion of the corresponding features.

It can be observed that most of the raw data is covered by the first two to four eigenvalues. Nonetheless, notable differences are observed for different data processing strategies. Figure 2a shows the results when the order of the ICA and the ILT step are changed and for different orientation of the input signal matrix. When the ILT step is conducted before ICA, considerably more PCA eigenvalues are necessary for faithful data replication. A possible reason could be that in relaxation experiments, every signal contributes at each point in time during a decay or a

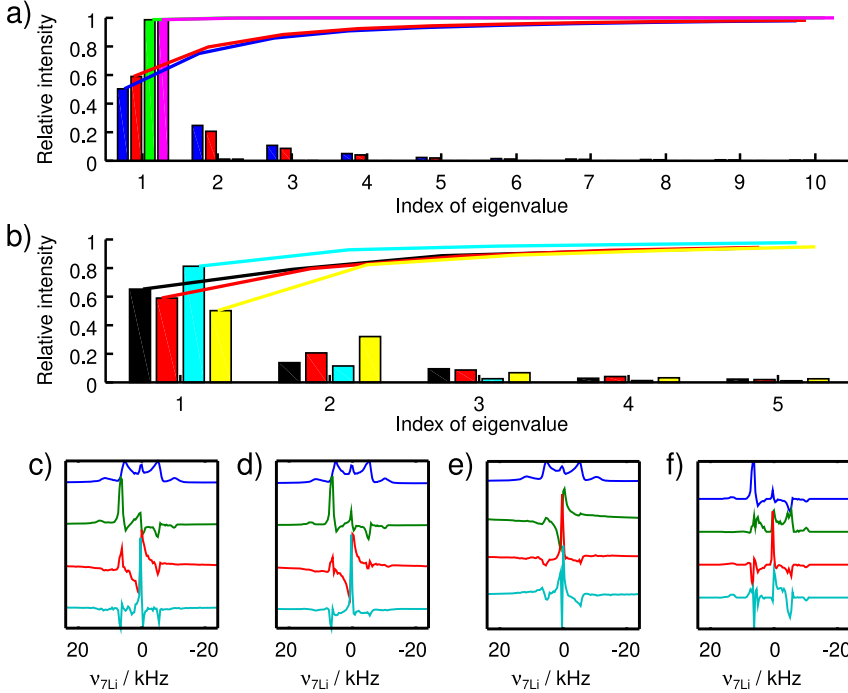


Figure 2: PCA step of ICA analysis of LGPS sample 2.

a) Relative intensity of first 10 principal component eigenvalues (bars) and their cumulative value (lines). Blue: ICA analysis of data after ILT (ILT-then-ICA), ^7Li NMR spectra from Fourier transform of echo transients as observed signals. Red: ILT-then-ICA, vectorization of τ_c and T_1 dimensions for each frequency component as observed signals. Green: Analysis of data after Fourier transform of echo dimension, but without ILT (ICA-then-ILT), spectra as observed signals. Magenta: ICA-then-ILT, vectorization of t_d and t_m dimensions individually for each frequency component as observed signals. b) ILT-then-ICA analysis using different parts of the data set. Black: full τ_c versus T_1 map, including baseline component at very long T_1 . Red: full τ_c data, $T_1 < 100$ s; same as red bars in a). Cyan: $10^{-4} \text{ s} \leq \tau_c \leq 10 \text{ s}$, $T_1 < 100$ s. Yellow: $10^{-6} \text{ s} \leq \tau_c \leq 10^{-4} \text{ s}$, $T_1 < 100$ s. c–f) First four principal component eigenvectors for the different cases shown in b). c) Full τ_c versus T_1 map (black in panel b). d) Full τ_c data, $T_1 < 100$ s (red in panel b). e) $10^{-4} \text{ s} \leq \tau_c \leq 10 \text{ s}$, $T_1 < 100$ s (cyan in panel b). f) $10^{-6} \text{ s} \leq \tau_c \leq 10^{-4} \text{ s}$, $T_1 < 100$ s (yellow in panel b).

recovery, weighted with the corresponding relaxation time constant. This causes a mixing of all the signals across the whole data domain. Laplace inversion, on the other hand, leads to a separation of different signals in a similar manner as Fourier transform causes a separation of resonances when applied to an interferogram. Furthermore, in the presented data set, signals are spread across a wide range of relaxation time constants, hence the relaxation domain provides good spectral

separation of signal components that are overlapping in the frequency dimension, which is beneficial for the ability of PCA to separate different components.

The sequence of ICA and ILT steps is compared more closely in Figure 3.

It is shown that if data is analyzed with retaining two principal component eigenvalues and retaining two ICs, similar independent components, which in this case represent ^7Li NMR spectra, and 2D SAE- T_1 correlation maps are obtained. When comparing the mixing matrix corresponding to the first independent component, somewhat more details are visible for the case of the ILT step preceding the ICA, as was also suggested by the quickly decaying principal component eigenvalues for ICA-then-ILT in Figure 2. On the other hand, conducting first an ICA on the full data set and then only an ILT step on the retained mixing matrices reduces the computation time by more than an order of magnitude. In this case, ICA serves as a filter and for data reduction.

Transposing the signal matrix swaps the meaning of the eigenvectors of the PCA between the ^7Li NMR spectra and the 2D $\tau_c - T_1$ correlation map of the 3D data

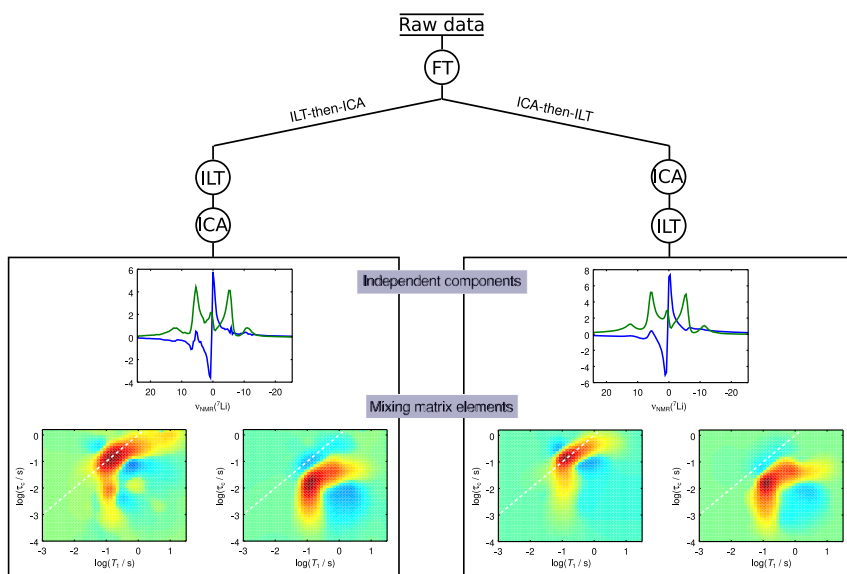


Figure 3: Graphical representation of possible procedures for combining ILT and ICA for the analysis of 2D-SAE/ T_1 data, with experimental results of LGPS sample 2. For ILT-then-ICA as well as for ICA-then-ILT, the signal matrix was oriented that the independent component represents the ^7Li spectrum and the mixing matrix elements the 2D map. For optimum comparability, the IC analysis on the ILT-then-ICA path was conducted using the full data set, including the baseline mode at long T_1 . The analysis was conducted with calculation of the spectra one data point off the echo maximum.

set. If the signal-to-noise ratio (SNR) in both dimensions is similar, which is commonly assumed in NMR experiments unless the SNR is high enough that multiplicative noise starts to dominate [47], this matrix orientation becomes important if contrast for different features in the signal is more pronounced in one dimension over the other. In the present case, both dimensions show a considerable amount of overlap of different features. In the spectral dimension, the quadrupolar coupling constants do not show large variations, while exchange causes mixing of different modes in the correlation time and relaxation dimensions [32]. Since no regularization was applied in the spectral dimension, choosing the $2D \tau_c - T_1$ correlation map as the signal dimension may prove more robust for data with lower SNR, though.

Figure 2b compares the eigenvalues of the PCA step when different regions of the full 2D correlation map are used for the PCA. For each case, the first four eigenvectors are shown in Figure 2c–f. The eigenvectors with and without the inclusion of the baseline feature at around $T_1 \approx 10^4$ s look similar. This is expected since the baseline feature represents a superposition with contributions from all the relaxation components. Considering the second independent component in Figure 3, one notices that it differs substantially from the second principal component eigenvectors, despite only subtle differences in the data analysis protocol. In Figure 2 the transient signal was Fourier transformed with the time origin one point off the echo maximum. Such a significant variation of the result caused by such a small variation of the processing parameter may, at first, indicate that the presented method is not robust enough for routine data analysis. However, this particular variation between independent components or between eigenvectors was observed multiple times when analysis parameters were changed, always providing an eigenvector or an independent component that either matched the second or the third component in Figure 2c, d. The norm of the corresponding contributions to the data was always similar, hence the order was rather random. Therefore, a closer look at the features of these spectra is indicated. One of them represents a central transition, phase shifted with respect to the quadrupolar Pake pattern of the first eigenvector, and the second one an asymmetric pattern with frequencies at the central transition and at the satellites of the quadrupolar spin. Phase shifted central transitions have been discussed for other systems in literature and are common features of SAE spectra (e.g. [34]). The presence of this feature displays the capability of ICA to reliably separate these contributions. The asymmetric feature is less obvious. Asymmetries in the SAE spectra of LGPS had previously been reported and discussed in terms of a high anisotropic mobility of quadrupolar nuclei that leads to a residual quadrupolar interaction [33, 48]. Its origin can be implied by comparing an analysis where PCA was conducted on the

2D $\tau_c - T_1$ correlation data using only either the part with long τ_c (Figure 2e) or the part with $\tau_c \leq 10^{-4}$ s (Figure 2f). The former shows the central transition as the second PCA eigenvector without an asymmetric eigenvector farther down, while the latter shows the asymmetric feature even as the dominant eigenvector. Therefore, the fast decaying signal at $\tau_c \approx 10^{-5}$ s is primarily responsible for the unexpected asymmetric feature.

Before the asymmetry at short τ_c is investigated in more details, the main features in the long τ_c region and the capability of ICA to reliably separate them are examined. Figure 4 shows ICA of the spectrally resolved 2D $\tau_c - T_1$ map with different numbers of eigenvalues retained from the PCA step, using the data at $\tau_c > 10^{-4}$ s and $T_1 < 100$ s. The number of independent components matched the number of PCA eigenvalues.

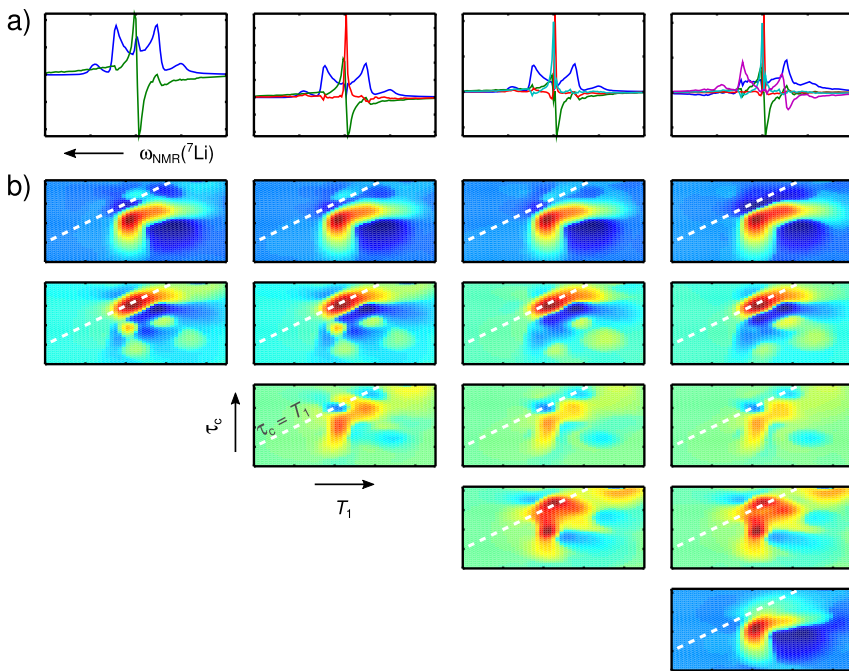


Figure 4: ILT-then-ICA analysis of LGPS sample 2. Independent components represent ^7Li NMR spectra. The number of dimensions retained after the PCA step was varied from two (left row) to five (right row), and the number of ICs was selected according to the dimension.

a) Independent components (blue: first IC; green: second IC; red: third IC; cyan: fourth IC; magenta: fifth IC). b) 2D- τ_c versus T_1 maps obtained from the mixing matrix. The x-axis shows T_1 logarithmically from 10^{-3} to $10^{1.5}$ s, and the y-axis shows τ_c from 10^{-4} to $10^{0.2}$ s. For ICA, data values with $T_1 < 100$ s and values $\tau_c > 10^{-4}$ s were used (within black solid box and above dash-dotted line in Figure 1).

2D Laplace inversion was conducted individually for each frequency of the ^7Li NMR spectrum obtained from Fourier transform of the echo transient of the 3D data set. No regularization in the spectral dimension was applied, hence the good capability to retrieve spectra in the independent components after reassembling the data following the ILT step is an indication of the robustness of the Laplace inversion despite avoiding an overregularization to achieve a smoother distribution and lower fluctuations in dimensions that are not regularized.

Comparing the different independent components, the number of PCA eigenvalues did not cause large variations up to four eigenvalues, only the addition of contributions that had been discarded before. When only two eigenvalues were used, the independent components represent a quadrupolar Pake pattern and a phase-shifted central transition. The 2D correlation maps showed distinct differences, with a dominant mode for the central transition along the $\tau_c = T_1$ diagonal. Hence this mode decays dominantly with T_1 , not with $T_{1,Q}$. Additional features indicate contributions that may be caused by mobility or by exchange. The observation of contributions with negative sign indicates that these features are not only caused by mobility, but also by exchange [17]. The features of the 2D map associated with the Pake-like independent component mirrors the main feature of the full correlation map shown in Figure 1, with a dominant contribution close to the region where $T_{1,Q}$ is expected, and additional signals towards shorter τ_c and increased T_1 , which had been described elsewhere [32]. When adding a third eigenvalue and calculating another independent component, it dominantly shows a central transition that is not phase shifted. Its contribution matches more closely the map of the Pake-like independent component. A fourth independent component showed another signal dominated by a central transition, but slightly shifted compared to the previous component. The 2D correlation map of these two components differed substantially, which may be an indication that another signal with a different chemical shift may be present. In previously published work by Liang *et al.*, such a contribution had been identified by ^7Li magic angle spinning (MAS) NMR [31], with a chemical shift difference reported at about 0.65 ppm. The present data only affords a resolution on the order of about 1.7 ppm, which makes a more in-depth analysis challenging without additional data. The difference between the correlation maps obtained from the ICA mixing matrix may hint at exchange between the two components, yet such an assignment would require independent confirmation.

When a fifth independent component is added, it appears to split the dominant Pake-like component into two separate components with very similar τ_c versus T_1 correlation maps. This is a sign that from this point, adding more independent components mainly provides a better fit of the data, yet the identified components

start to show overlapping features and the attempt of a physical interpretation of the differences may be futile.

The quality of the fit can be assessed by comparing the norm of the residuals between the fit and the experimental data, which is shown in Figure 5 for different values λ of the regularization parameter of the Laplace inversion step in Eq. (8). As a trend, the higher the regularization the smoother the obtained features after inversion and the faster the ICA converges, yet reduced regularization leads to an improved fit in the end. The lowest of the three values of the regularization parameter, $\lambda = 1$, represents the default used in Laplace inversion experiments [20].

In Figure 4, where signal contributions with $\tau_c < 10^{-4}$ s were not included in the ICA, the asymmetric component, which was identified in the PCA eigenvector analysis of Figure 2, was not retrieved even when using five independent components. To investigate this further and prevent discarding any features present in the data, an ICA analysis was performed on the data prior to Laplace inversion, and the 2D maps were calculated afterwards by ILT of the mixing matrix elements. The results are summarized in Figure 6.

In this case, the dominant independent components showed central transitions with similar maps as two of the respective components in Figure 4. Neither of these components showed a feature at $\tau_c < 10^{-4}$ s. The third component showed the

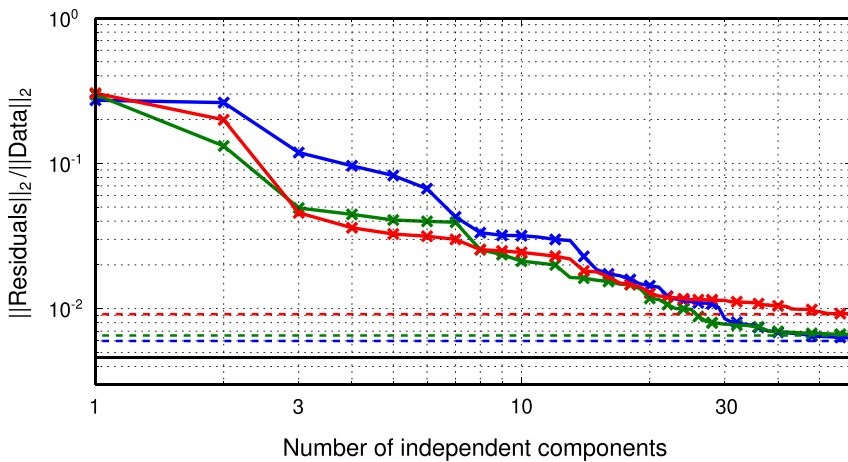


Figure 5: Norm of residuals between experimental data and after ILT-then-ICA fitting of 2D-SAE/ T_1 data of LGPS sample 2 for different ILT regularization. The ICs represent the 2D-SAE/ T_1 map and the mixing matrix the ^7Li NMR spectra. Blue: $\lambda = 1$. Green: $\lambda = 2.5$. Red: $\lambda = 10$. The dashed lines represent the relative norm of the residuals of the ILT step, which represent a lower limit for the subsequent ICA step. The black line marks an estimate of the norm of the random noise of the data after Fourier transform of the echo transient.

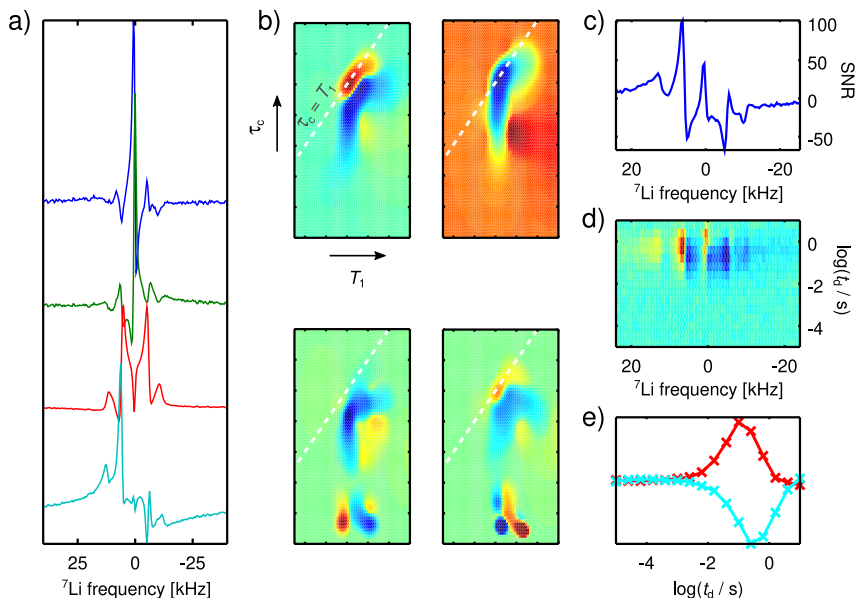


Figure 6: ICA-then-ILT analysis of LGPS sample 2, ^7Li NMR spectra from Fourier transform of echo transients as observed signals. Hence independent components represent ^7Li NMR spectra. a) First four independent components. b) Mixing matrices, reshaped as 2D-SAE/ T_1 correlation maps. Top left corresponds to first IC, top right to second IC, bottom left to third and bottom right to fourth IC. The white dashed line marks data points for which $\tau_c = T_1$. c) ^7Li NMR spectrum obtained from the difference between the first two t_m data points, summed over five relaxation delays with $0.03\text{ s} < t_d < 1.6\text{ s}$. d) t_d versus ^7Li NMR difference spectrum between the first two t_m data points. e) Mixing matrix elements corresponding to the third and the fourth independent components, shown in the same color in a).

Pake-like pattern and the fourth represented the asymmetric features reported previously. The SAE/ T_1 maps of the two latter components showed features at very short τ_c , which were particularly dominant for the asymmetric independent component. Since $\tau_c \approx 10^{-5}\text{ s}$ corresponds to the shortest t_m values employed in this experiment, this signal contribution is mainly encoded in the corresponding data points. By calculating the difference between the spectra obtained with the shortest two t_m values, an estimate for the ^7Li NMR spectrum of the species responsible for this signal can be obtained directly, without employing ICA (Figure 6c). The spectrum does show the asymmetric behavior suggested by PCA and ICA. When plotting this difference spectrum as a function of the recovery time t_d (Figure 6d, e), the origin of the bipolar features along the T_1 dimension (Figures 1 and 6b) is immediately apparent. This signal component does not relax towards a

maximum, but behaves more like an exchange feature that had not been directly excited. The shape of the relaxation curve qualitatively and also semi-quantitatively resembles the evolution of multi-quantum filtered longitudinal orders that was discussed by Paulus *et al.* [33]. It confirms the link established there between the asymmetry of the quadrupolar satellite transitions and the evolution of magnetization into longitudinal multi-quantum order states. The fast decay along t_m indicates that this signal contribution follows a coherence transfer pathway that is not filtered by the employed phase cycle, which may also explain the asymmetry of the spectrum, since the origin of the signal does not coincide with the SAE and changes as a function of t_m . This could also explain why the asymmetry of the measured SAE detected spectra is more pronounced than the relaxation effects simulated in Ref. [33] suggest.

To demonstrate the contrast that can be achieved by using ICA to analyze spectrally resolved 2D SAE/ T_1 data of fast ion conductors, the same protocol was used to process data of an LGPS sample from a different batch, which showed a somewhat different tetragonal-to-orthorhombic ratio. Figure 7 shows the result of an ILT-then-ICA analysis with ^7Li NMR spectra as the independent components.

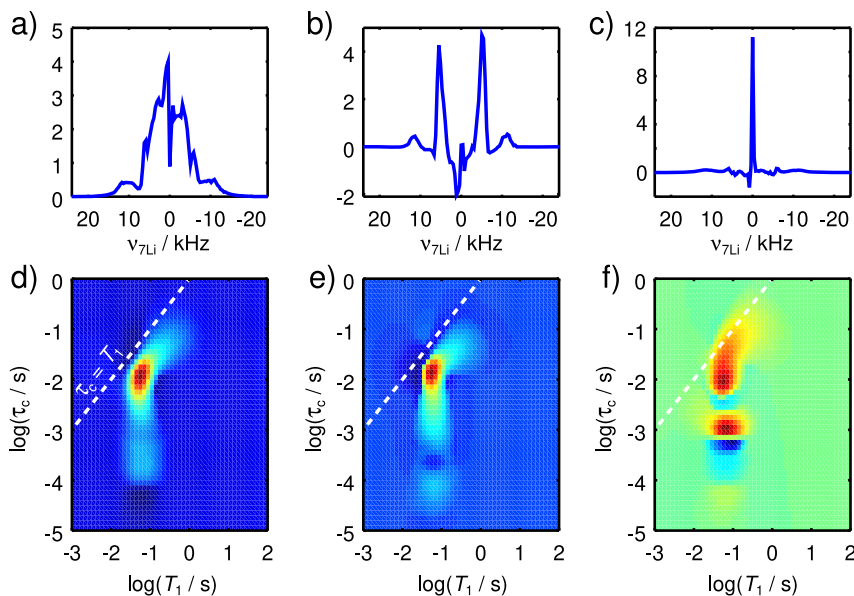


Figure 7: ILT-then-ICA analysis of LGPS sample 1. Independent components represent ^7Li NMR spectra. The top row shows the first (a), second (b) and third (c) independent component. The bottom row (d–f) shows the corresponding 2D-SAE/ T_1 correlation maps, obtained from the mixing matrices. The white dashed lines mark data points for which $\tau_c = T_1$.

These independent components showed markedly different spectra compared to the independent components of sample **2**. The first independent component, which provided the contribution with the largest norm to the data fit, showed a quadrupolar pattern, but with less clearly pronounced features than the dominant Pake-like independent component of the other sample. Such a behavior had been described in literature for species with lower mobility [46], yet it was also reported that with lower mobility the quadrupolar coupling constant increases due to less effective time averaging [29]. Such a behavior can be investigated in more details by temperature dependent measurements [49], which will be reported elsewhere.

For the present data, the second independent component provides a hint regarding the origin of the shape difference. It shows narrow peaks at the frequencies of the principal component of the quadrupole coupling. When summing the first two ICs, a spectrum that resembles more closely the Pake-like IC of the other sample appears. This is confirmed by the overall similarity of the 2D correlation maps for the first two ICs, which suggests a strong partial correlation between them. At the same time, the fact that ICA recognized them as independent components shows that there are statistically significant differences between the two maps. A closer look at the 2D $\tau_c - T_1$ correlation map reveals that the main feature shows minor differences along both dimensions. In addition, the two features differ significantly at short τ_c values and along the $\tau_c = T_1$ line. Since the frequencies of a quadrupolar line are correlated with orientation, this implies orientation-dependent relaxation that can be caused by anisotropic ionic mobility, which is well documented for this class of samples [50–52]. The observation that such an orientation dependence can be resolved indicates a longer residence time in a particular environment than for sample **2**, *i.e.* a slower exchange between crystallites with different orientation. Whether this is beneficial or detrimental for long-range mobility could be investigated, for example, using ^7Li PFG NMR.

For LGPS, an anisotropy of the Li ion dynamics has been suggested in various studies based on a range of different analytical techniques, and it can be justified by structural features of the LGPS lattice [26, 27, 53]. It is also consistent with the results of sample **2**, yet there this anisotropic mobility led to a considerably different NMR response. The presented data are not sufficient for a conclusive explanation, yet a past controversy may give a hint for a possible reason. Vashman *et al.* have concluded from NMR experiments on the superionic compound $\text{Li}_3\text{Sc}_2(\text{PO}_4)_3$ that the formation of Li–Li dimers with a correlated motion over a prolonged period of time could explain their data [54]. Bertermann *et al.* have studied the same material, yet did not find indications for such a dimer formation [29]. If, as a hypothesis, the lifetime of such a dimer would strongly vary even for a particular material, for example due to variations of crystallite domain or grain size, grain boundaries, doping, occurrence and ratio of secondary phases, or

vacancies and defects in the material [55, 56], these experiments would provide a unique tool for at least an indirect assessment of existence and lifetime of Li–Li correlated motion [57]. In the present case, even though both samples were synthesized with a nominally identical route, the ratio of tetragonal-to-orthorhombic LGPS differed between 75% *t*-LGPS for sample **1** and 79% for sample **2**.

The formation of long-lived Li–Li dimers would provide a pathway for dipolar relaxation, which is observed for sample **2** at $\tau_c = T_1$. On the other hand, the third IC of this sample, which represents the central transition resonance, relaxes at values $\tau_c < T_1$ that would be expected for relaxation of quadrupolar order. This could be an indication for a different origin for the relaxation process, be it via exchange or via coupling to a quadrupolar nucleus. While the exact nature of these differences and the consequences for sample optimization will be investigated elsewhere, it shows the capabilities and usefulness of ICA for investigating NMR data of superionic conductors for nuclei with a spin larger than 1/2. In particular, SAE detection appears well suited since it selects coherence pathways that predominantly lead to non-Gaussian features in the spectra as obtained by Fourier transform of the echo transient. While considerable variations of features and sequence of the dominant independent components are observed, the differences between independent components could be justified by sample behavior that had already been investigated previously by other means. Furthermore, for a given analysis protocol the results appear suitably robust for using them as input of a neural network, for example for sample screening. Therefore, it appears promising to look into this possibility in more details in the future.

5 Conclusions

The use of ICA for the analysis of multidimensional NMR experiments with transient spin-alignment echo detection was demonstrated. ICA provided stable, statistically independent correlation maps and the corresponding spectra. It was possible to select the ^7Li NMR spectra as well as the correlation map as the independent component. The former takes advantage of variations in the spectrum of different environments, and the latter of varying dynamics.

The ICA could be conducted before or after a Laplace inversion of the relaxation and SAE dimensions. The former has the advantage that ILT, which is computationally more demanding than ICA, only needs to be conducted on a reduced number of independent components that show significant variations of the NMR spectra. Thereby, ICA acts as a data reduction step. The latter seems to show a better contrast, with density distributed across a larger number of eigenvalues in a PCA.

LGPS material from two different batches was investigated. The results showed substantial differences in their spectral and relaxation behavior, which could be explained with previously known or suspected materials properties, and which could be attributed in some way to ionic mobility. The evolution of different relaxation orders has been observed. The obtained independent components, combined with further analysis based thereon, suggest a slightly different explanation, yet identical origin for asymmetric spectral features as had been previously discussed [33].

ICA was well suited to separate central and satellite transitions of the ^7Li NMR spectrum. Both of them showed a high sensitivity and thereby a good contrast for material variations. This suggests that an ICA based protocol could be used for systematic material optimization and screening applications. Thereby, an important factor appears to be the nature of NMR spectra of quadrupolar nuclei with high mobility in a solid matrix, which show residual quadrupolar coupling rather than a featureless inhomogeneous broadening that approaches a Gaussian distribution. This behavior has been widely observed for superionic compounds [29]. Nonetheless, a wider application of ICA, particularly for nuclei with spin $1/2$ or in combination with multidimensional experiments where a second dimension provides contrast that allows for a more pronounced resolution of spectral features, should not be precluded either and deserves further research.

LGPS shows a diverse dynamic behavior with various different processes coupled by exchange interactions, which makes an identification of individual contributions challenging. A promising strategy could be the combination of temperature dependent NMR experiments with ICA and ILT analysis, intrinsically coupled with numerical *ab initio* simulations [58]. Thereby, certain dynamic processes could be frozen out, which then would allow a better spectroscopic characterization and, by following the different contributions, observation of their evolution and onset of exchange as a function of temperature.

ICA based data analysis may appear rather complex at first for manual analysis. However, the obtained independent components could be related to the physical behavior of the sample. In addition, the obtained features were challenging to extract by more classical analysis techniques, and a tool such as ICA that provides statistically independent features in the data is a welcome addition for the analysis of samples where a variety of different effects influence measurement results. Only the combination of multi-dimensional experimental data with analysis techniques capable of separating hidden correlations may prove successful to provide a full understanding of the obtained experimental results. Furthermore, since the obtained results of a combined ICA and ILT analysis can be correlated with physical properties, this approach provides a promising data preprocessing approach for further automated analysis using machine learning concepts.

Acknowledgments: Helpful discussions during different stages of the project with Philipp Schleker, Christoph Scheurer, Simone Köcher and Paul Heitjans, who also provided us with magnet time to conduct experiments at 14.1 T in his lab at the Leibnitz University Hannover, are gratefully acknowledged. We are thankful for LGPS synthesis support by Peter-Paul Harks and Peter Notten.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: This research has been financially supported by the German Federal Ministry of Education and Research (BMBF), project FestBatt-Charakterisierung (grant number 03XP0176), and by the Ministry of Innovation, Science and Research (MIWF) of the State North Rhine-Westphalia through project “Ionic conductors for efficient energy storage”. Computing resources had been granted by RWTH Aachen University under project rwth0204.

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

1. Kalidindi S., Kalinin S. V., Eds. *Handbook on Big Data and Machine Learning in the Physical Sciences. Volume 1: Big Data Methods in Experimental Materials Discovery*; World Scientific: Singapore, 2020.
2. Zaki M. J., Meira W. *Data Mining and Analysis: Fundamental Concepts and Algorithms*; Cambridge University Press: New York, 2014.
3. Carleo G., Cirac I., Cranmer K., Daudet L., Schuld M., Tishby N., Vogt-Maranto L., Zdeborová L. *Rev. Mod. Phys.* 2019, **91**, 045002.
4. Succi S., Coveney P. V. *Phil. Trans. R. Soc. A Math. Phys. Eng. Sci.* 2019, **377**, 20180145.
5. Frydman L., Scherf T., Lupulescu A. *Proc. Natl. Acad. Sci. U. S. A.* 2002, **99**, 15858–15862.
6. Granwehr J., Panek R., Leggett J., Koeckenberger W. *J. Chem. Phys.* 2010, **132**, 244507.
7. Ahola S., Zhivonitko V. V., Mankinen O., Zhang G., Kantola A. M., Chen H.-Y., Hilty C., Koptiyug I. V., Telkki V.-V. *Nat. Commun.* 2019, **6**, 8363.
8. Ernst R. R., Bodenhausen G., Wokaun A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*; Clarendon Press: Oxford, 1987.
9. Stallmach F., Galvosas P. *Annu. Rep. NMR Spectrosc.* 2007, **61**, 51–131.
10. Monakhova Y. B., Tsikin A. M., Mushtakova S. P. *J. Anal. Chem.* 2016, **71**, 554–560.
11. Spiess H. W. *J. Chem. Phys.* 2008, **72**, 6755–6762.
12. Wilkening M., Küchler W., Heitjans P. *Phys. Rev. Lett.* 2006, **97**, 065901.
13. Böhmer R., Jeffrey K. R., Vogel M. *Prog. Nucl. Magn. Reson. Spectrosc.* 2007, **50**, 87–174.
14. Wilkening M., Amade R., Iwaniak W., Heitjans P. *Phys. Chem. Chem. Phys.* 2007, **9**, 1239–1246.
15. Venkataramanan L., Song Y.-Q., Hürlimann M. D. *IEEE Trans. Signal Process.* 2002, **50**, 1017–1026.
16. Borgia G. C., Brown R. J. S., Fantazzini P. *J. Magn. Reson.* 1998, **132**, 65–77.

17. Rodts S., Bytchenkoff D. J. *Magn. Reson.* 2010, 205, 315–318.
18. Fantazzini P., Galassi F., Bortolotti V., Brown R. J. S., Vittur F. *New J. Phys.* 2011, 13, 065007.
19. Granwehr J., Roberts P. J. J. *J. Chem. Theor. Comput.* 2012, 8, 3473–3482.
20. Graf M. F., Tempel H., Köcher S. S., Schierholz R., Scheurer C., Kungl H., Eichel R.-A., Granwehr J. *RSC Adv.* 2017, 7, 25276–25284.
21. Merz S., Jakes P., Taranenko S., Eichel R.-A., Granwehr J. *Phys. Chem. Chem. Phys.* 2019, 21, 17018–17028.
22. Bachman J. C., Muy S., Grimaud A., Chang H.-H., Pour N., Lux S. F., Paschos O., Maglia F., Lupart S., Lamp P., Giordano L., Shao-Horn Y. *Chem. Rev.* 2016, 116, 140–162.
23. Lau J., DeBlock R. H., Butts D. M., Ashby D. S., Choi C. S., Dunn B. S. *Adv. Energy Mater.* 2018, 8, 1800933.
24. Tatsumisago M., Nagao M., Hayashi A. J. *Asian Ceram. Soc.* 2013, 1, 17–25.
25. Kamaya N., Homma K., Yamakawa Y., Hirayama M., Kanno R., Yonemura M., Kamiyama T., Kato Y., Hama S., Kawamoto K., Mitsui A. *Nat. Mater.* 2011, 10, 682–686.
26. Weber D. A., Senyshyn A., Weldert K. S., Wenzel S., Zhang W., Kaiser R., Berends S., Janek J., Zeier W. G. *Chem. Mater.* 2016, 28, 5905–5915.
27. Kuhn A., Kohler J., Lotsch B. V. *Phys. Chem. Chem. Phys.* 2013, 15, 11620–11622.
28. Kuhn A., Duppel V., Lotsch B. V. *Energy Environ. Sci.* 2013, 6, 3548–3552.
29. Bertermann R., Müller-Warmuth W. Z. *Naturforsch.* 1998, 53, 863–873.
30. Hogrefe K., Minafra N., Zeier W. G., Wilkening H. M. R. J. *Phys. Chem. C* 2021, 125, 2306–2317.
31. Liang X., Wang L., Jiang Y., Wang J., Luo H., Liu C., Feng J. *Chem. Mater.* 2015, 27, 5503–5510.
32. Paulus M. C., Graf M. F., Harks P. P. R. M. L., Paulus A., Schleker P. P. M., Notten P. H. L., Eichel R.-A., Granwehr J. *J. Magn. Reson.* 2018, 294, 133–142.
33. Paulus M. C., Paulus A., Schleker P. P. M., Jakes P., Eichel R.-A., Heitjans P., Granwehr J. *J. Magn. Reson.* 2019, 303, 57–66.
34. Wilkening M., Gebauer D., Heitjans P. J. *Phys. Condens. Matter* 2008, 20, 022201.
35. Storek M., Jeffrey K. R., Böhmer R. *Solid State Nucl. Magn. Reson.* 2014, 59–60, 8–19.
36. Hyvärinen A. *IEEE Trans. Neural Network.* 1999, 10, 626–634.
37. Hyvärinen A., Oja E. *Neural Network.* 2000, 13, 411–430.
38. Comon P. *Signal Process.* 1994, 36, 287–314.
39. Särelä J., Vigário R. J. *Mach. Learn. Res.* 2003, 4, 1447–1469.
40. Brown R. J. S. J. *Magn. Reson.* 1989, 82, 539–561.
41. Ostroff E. D., Waugh J. S. *Phys. Rev. Lett.* 1966, 16, 1097–1098.
42. Callaghan P. T. *Translational Dynamics and Magnetic Resonance: Principles of Pulsed Gradient Spin Echo NMR*; Oxford University Press: Oxford, UK, 2011.
43. Hansen P. C. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*; SIAM: Philadelphia, 1997.
44. Kazimierczuk K., Orekhov V. Y. *Angew. Chem. Int. Ed.* 2011, 50, 5556–5559.
45. Eaton J. W., Bateman D., Hauberg S., Wehbring R. *GNU Octave Version 4.0.0 Manual: A High-Level Interactive Language for Numerical Computations*; Free Software Foundation: Boston, MA, 2015.
46. Qi F., Rier C., Böhmer R., Franke W., Heitjans P. *Phys. Rev. B* 2005, 72, 104301.
47. Granwehr J. *Appl. Magn. Reson.* 2007, 32, 113–156.
48. Petit D., Korb J.-P., Delville A., Grandjean J., Laszlo P. J. *Magn. Reson.* 1992, 96, 252–279.
49. Wilkening M., Heitjans P. *ChemPhysChem* 2012, 13, 53–65.
50. Stöffler H., Zinkevich T., Yavuz M., Senyshyn A., Kulisch J., Hartmann P., Adermann T., Randau S., Richter F. H., Janek J., Indris S., Ehrenberg H. J. *Phys. Chem. C* 2018, 122, 15954–15965.

51. Prutsch D., Gadermaier B., Brandsttter H., Pregartner V., Stanje B., Wohlmuth D., Epp V., Rettenwander D., Hanzu I., Wilkening H. M. R. *Chem. Mater.* 2018, 30, 7575–7586.
52. Harm S., Hatz A. K., Moudrakovski I., Eger R., Kuhn A., Hoch C., Lotsch B. V. *Chem. Mater.* 2019, 31, 1280–1288.
53. Krauskopf T., Culver S. P., Zeier W. G. *Chem. Mater.* 2018, 30, 1791–1798.
54. Vashman A. A., Pronin I. S., Sigaryov S. E. *Solid State Ionics* 1992, 58, 201–215.
55. Heenen H. H., Scheurer S., Reuter K. *Nano Lett* 2017, 17, 3884–3888.
56. Spannenberger S., Miss V., Klotz E., Kettner J., Cronau M., Ramanayagam A., di Capua F., Elsayed M., Krause-Rehberg R., Vogel M., Roling B. *Solid State Ionics* 2019, 341, 115040.
57. Bhandari A., Bhattacharya J. J. *Phys. Chem. C* 2016, 120, 29002–29010.
58. Köcher S. S., Schleker P. P. M., Graf M. F., Eichel R.-A., Reuter K., Granwehr J., Scheurer C. *J. Magn. Reson.* 2018, 297, 33–41.