The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Statistical learning of NMR tensors from 2D isotropic/anisotropic correlation nuclear magnetic resonance spectra

Cite as: J. Chem. Phys. 153, 134201 (2020); doi: 10.1063/5.0023345
Submitted: 29 July 2020 • Accepted: 9 September 2020 •
Published Online: 5 October 2020

Deepansh J. Srivastava[a] and Philip J. Grandinetti[b]

AFFILIATIONS

Department of Chemistry, Ohio State University, 100 West 18th Avenue, Columbus, Ohio 43210, USA

[a]Electronic mail: srivastava.89@osu.edu
[b]Author to whom correspondence should be addressed: grandinetti.1@osu.edu. URL: http://www.grandinetti.org/

ABSTRACT

Many linear inversion problems involving Fredholm integrals of the first kind are frequently encountered in the field of magnetic resonance. One important application is the direct inversion of a solid-state nuclear magnetic resonance (NMR) spectrum containing multiple overlapping anisotropic subspectra to obtain a distribution of the tensor parameters. Because of the ill-conditioned nature of this inverse problem, we investigate the use of the truncated singular value decomposition and the smooth least absolute shrinkage and selection operator based regularization methods, which (a) stabilize the solution and (b) promote sparsity and smoothness in the solution. We also propose an unambiguous representation for the anisotropy parameters using a piecewise polar coordinate system to minimize rank deficiency in the inversion kernel. To obtain the optimum tensor parameter distribution, we implement the k-fold cross-validation, a statistical learning method, to determine the hyperparameters of the regularized inverse problem. In this article, we provide the details of the linear-inversion method along with numerous illustrative applications on purely anisotropic NMR spectra, both synthetic and experimental two-dimensional spectra correlating the isotropic and anisotropic frequencies.

I. INTRODUCTION

Nuclear magnetic resonance (NMR), like other spectroscopies, is a probe of the local structure. To obtain these local structures from an experimental spectrum, one must (1) determine the underlying distribution of parameters $\mathbf{R}$ that quantify the tensorial interactions of the nuclear moments with local magnetic and electric fields and (2) have quantitative relationships for mapping the distribution of nuclear moment interaction parameters into local structure distributions.

Determining the parameters, $\mathbf{R}$, from an inhomogeneous NMR spectrum,[1] $s(\nu)$, usually begins by modeling the spectrum as a linear combination of N simulated subspectra, each associated with a nuclear spin system with a specific set of nuclear moment interaction parameters, $\mathbf{R}$. In the case of an ordered sample, this linear combination can be written as

$$s(\nu) = \sum_{i=1}^{N} f_i \mathcal{K}(\nu, \mathbf{R}_i), \tag{1}$$

where ν is the spectroscopic dimension and $\mathcal{K}(\nu, \mathbf{R}_i)$ represents a simulated subspectrum of a nuclear spin system with a given set of parameters $\mathbf{R}_i$. The population f_i and parameters $\mathbf{R}_i$ of each spin system contributing to the spectrum are determined by a numerical least-squares analysis, comparing the simulated and experimental spectra.

In crystalline materials, this is traditionally performed as a non-linear least-squares analysis due to the non-linear dependence of $\mathcal{K}(\nu, \mathbf{R}_i)$ on $\mathbf{R}$, with N usually fixed to the minimum number of spin systems needed to obtain a good fit of the spectrum—a value expected to be consistent with the number of magnetically inequivalent sites in the crystal. In this traditional approach, each subspectrum is convolved with a Lorentzian line shape to account for finite transition lifetimes. More often than not, an additional *ad hoc* line broadening convolution with a Gaussian line shape is applied to each subspectrum to model the effects of structural disorder on the spectrum. Such structural disorder effects on a spectrum, however, are more realistically modeled with a continuous distribution of $\mathbf{R}$,

written as

$$s(v) = \int_{\mathbf{R}} f(\mathbf{R})\mathcal{K}(v,\mathbf{R})d\mathbf{R}, \quad (2)$$

where $f(\mathbf{R})$ is the continuous and multi-dimensional distribution of nuclear spin system parameters present within a sample.

Equations (1) and (2) suggest turning the traditional non-linear least-squares problem into a linear one by considering only a basis set of pre-calculated subspectra, $\mathcal{K}(v,\mathbf{R})$, associated with some unknown parameter distribution, $f(\mathbf{R})$. Such linear problems[2–4] are frequently encountered in a variety of biological[4] and geological[2,3] NMR applications, particularly relaxation measurements—spin–lattice T_1, spin–spin T_2, diffusion D, and multi-dimensional correlations[5,6] among them. The general challenge with this approach is that Eq. (2) is a Fredholm integral of the first kind,[7] and the inverse of the forward computation, i.e., calculating $f(\mathbf{R})$ from $s(v)$, is often an ill-posed problem.[8] Both Eq. (1) and, using a quadrature rule,[7] Eq. (2) can be written as a matrix equation

$$\mathbf{s} = \mathbf{K}\cdot\mathbf{f}, \quad (3)$$

where $\mathbf{s}$ is a spectrum vector of length m, $\mathbf{f}$ is the model parameter vector of length n, and $\mathbf{K}$ is an $m \times n$ kernel matrix. With d parameters in $\mathbf{R}$, one can imagine constructing a d-dimensional grid of amplitudes and a corresponding d-dimensional grid of basis subspectra. In this linear inversion problem, the d-dimensional grid of amplitudes are stacked into a one-dimensional model vector, $\mathbf{f}$, with the basis subspectra placed in the corresponding columns of the kernel matrix, $\mathbf{K}$.

In one of the earliest attempts to solve the inverse problem with an NMR spectra, Peterson *et al.*[9–11] obtained the distribution of ^{11}B quadrupolar coupling constants from the ^{11}B static spectrum of a B_2O_3 glass using truncated singular value decomposition (TSVD).[8] In this pioneering but overlooked work, they reduced the size of $\mathbf{f}$ with an unrealistic assumption for the basis set of subspectra—allowing only the quadrupolar coupling constant, C_q, to vary, while all other parameters such as the quadrupolar asymmetry and the shielding tensor parameters were ignored or held fixed. Obviously, a successful forward or inverse calculation requires having a complete basis set of $\mathcal{K}(v,\mathbf{R})$ for accurately modeling the spectrum. Numerous successful efforts when $\mathbf{R}$ contains only one parameter can be found in the solid-state NMR literature, such as the de-Paking of one-dimensional NMR spectra[12–17] or obtaining one-dimensional angle distributions from two-dimensional NMR spectra.[18]

In the case of solid-state NMR spectra of dilute (isolated) spin 1/2 nuclei, $\mathbf{R}$ would correspond to the three principal components of the symmetric part of the nuclear shielding tensor. These three parameters are re-expressed as the isotropic chemical shift, $\delta_{iso}^{(cs)}$, and two other parameters describing the anisotropy of the interaction. In the Haeberlen notation,[19,20] these latter two parameters are represented by the nuclear shielding anisotropy, ζ_σ, and asymmetry, η_σ. The form of $\mathcal{K}(v,\mathbf{R})$ would also depend on whether the measurement takes place under static, variable-angle spinning (VAS) or magic-angle spinning (MAS) conditions. The MAS basis subspectra become less distinguishable as spinning sidebands are eliminated with increasing rotor speeds. This leads to greater rank deficiency in the kernel, $\mathbf{K}$, and, as expected, increased uncertainty in determining the distribution of the tensor parameters. On the other hand, the great advantage of faster MAS speeds is that $\mathbf{R}$ gets reduced to a single parameter, $\delta_{iso}^{(cs)}$, resulting in a simplified spectral analysis, albeit with a loss of the local structure information contained in the anisotropy parameters.

For dilute quadrupolar nuclei, $\mathbf{R}$ would include not only the three principal components of the symmetric part of the nuclear shielding tensor but also two principal components of the surrounding traceless electric field gradient (EFG) tensor as well as the three Euler angles specifying the relative orientation between the two tensors. For the subspectra of quadrupolar nuclei experiencing only first-order anisotropies, $\mathbf{R}$ gets reduced to a single parameter, $\delta_{iso}^{(cs)}$, when spinning sidebands are eliminated at sufficiently high MAS speeds. When second-order anisotropies are present, high MAS speeds can only reduce $\mathbf{R}$ down to $\delta_{iso}^{(cs)}$, the quadrupolar coupling constant, C_q, and the quadrupolar asymmetry parameter, η_q.

A major advantage of multi-dimensional NMR methods is that the influence of different subsets of nuclear moment parameters on the subspectra can be altered along different spectral dimensions. Thus, for the same set of $\mathbf{R}$, the multi-dimensional basis subspectra tend to be more distinguishable compared to the one-dimensional basis subspectra, leading to less rank deficiency in the kernel, $\mathbf{K}$, and decreased uncertainty when determining the distribution of $\mathbf{R}$.[21,22] Two-dimensional NMR methods that separate and correlate the isotropic and the anisotropic frequency contributions into orthogonal dimensions are often used to determine NMR tensor parameters. For uncoupled spin $I = 1/2$ nuclei, this includes techniques such as 2D One Pulse (TOP) MAS,[23,24] phase adjusted spinning sidebands (PASS),[25–27] magic-angle turning (MAT),[28,29] extended chemical shift (XCS) modulation,[30,31] magic-angle hopping (MAH),[32] magic-angle flipping (MAF),[33] Recoupling Of Chemical Shift Anisotropy (ROCSA),[34] and Variable Angle Correlation Spectroscopy (VACSY).[35] For half-integer quadrupolar nuclei, this includes techniques such as dynamic-angle spinning (DAS),[36–38] multiple-quantum magic-angle spinning (MQ-MAS),[39,40] and satellite-transition MAS (ST-MAS).[41,42] A key feature of all these 2D isotropic/anisotropic correlation spectra—either as acquired or after a shear transformation—is that the subspectral basis set needed to model an anisotropic cross section,

$$s(v|\delta_{iso}) = \int_{\mathbf{R}} \mathcal{K}(v,\mathbf{R})f(\mathbf{R}|\delta_{iso})d\mathbf{R}, \quad (4)$$

depends only on two anisotropy parameters. For the anisotropic cross sections from the spin $I = 1/2$ techniques,[23–33,35,43] $\mathbf{R}$ is reduced down to the two parameters describing the nuclear shielding anisotropy. For the half-integer quadrupolar nuclei techniques,[36–42] $\mathbf{R}$ can be reduced down to the two parameters describing the nuclear quadrupolar coupling. Note, in this latter case, the cross section's correlated isotropic shift, δ_{iso}, is a linear combination of a second-order isotropic quadrupolar shift and an isotropic chemical shift.[44]

In previous attempts to apply linear inversion methods to the ^{11}B DAS spectrum of B_2O_3 glass[45] and the ^{23}Na and ^{27}Al MQ-MAS spectra of borosilicate glasses,[46] the asymmetry parameter, η_q, was

unrealistically held constant while determining a bivariate distribution of isotropic, $\delta_{iso}^{(cs)}$ and C_q, parameters. Similarly, the shielding asymmetry parameter, η_σ, was held constant without justification in a recent application of linear inversion methods to determine the distribution of shielding anisotropies in the ^{29}Si 2D PASS NMR spectra of silica-based nuclear waste glasses.[22] Although not explicitly stated, it is likely that the regularization used in these studies, TSVD or the related ridge regression[47] (aka Tikhonov), performed poorly and a more complete subspectral basis set could not be used to obtain a meaningful trivariate distribution of tensor parameters when modeling these 2D isotropic/anisotropic correlation spectra.

In this paper, we take advantage of recent advances in linear inversion and statistical learning algorithms as well as better model selection methods[48–52] to obtain a robust inversion method for determining the trivariate distribution of NMR tensor parameters from 2D isotropic/anisotropic correlation spectra. Here, we focus primarily on determining the distribution of ^{29}Si nuclear shielding tensors from the experimental 2D MAF and MAT spectra of silicate glasses.

To minimize rank deficiency in the inversion kernel and obtain a more robust inversion, we also propose a new and unambiguous representation for the two anisotropy parameters using a piecewise polar coordinate system, where the magnitude of ζ forms the radial dimension and η forms the angular dimension. In this more natural representation, the ζ and η parameters are replaced with two Cartesian parameters denoted with x and y (*vide infra*). This x–y notation overcomes the two serious shortcomings of the Haeberlen ζ–η notation[20] that (1) η becomes undefined when ζ goes to zero and (2) ζ has a sign degeneracy when $\eta = 1$. In the Herzfeld–Berger (also known as Maryland) span and skew notation, the skew parameter similarly becomes undefined when span goes to zero.[53]

Finally, we have created the open-source python software package, *mrinversion*, for implementing the solid-state NMR spectrum inversions described here.[54] Details for obtaining *mrinversion*, along with the example datasets discussed in this article, are given at the end.

II. THEORETICAL BACKGROUND AND METHODS

A. Nuclear shielding interaction

In this article, we follow the International Union of Pure and Applied Chemists (IUPAC) definitions for nuclear shielding and chemical shift interactions.[20] The isotropic nuclear shielding is defined as the trace of the nuclear shielding tensor,

$$\sigma_{iso} = \frac{1}{3}(\sigma_{zz} + \sigma_{yy} + \sigma_{xx}), \tag{5}$$

where σ_{zz}, σ_{yy}, and σ_{xx} are the components of the second-rank shielding tensor. The isotropic chemical shift is defined as

$$\delta_{iso}^{(cs)} = \frac{\sigma_{ref} - \sigma_{iso}}{1 - \sigma_{ref}}, \tag{6}$$

where σ_{ref} is the nuclear shielding of the reference compound, tetramethyl silane (TMS) for this study.

The nuclear shielding tensor can be visualized as a three-dimensional ellipsoid, and the deviations, positive or negative, of this ellipsoid from a sphere are best understood by examining the elements of the traceless symmetric part of the nuclear shielding tensor, given by

$$S_{ik}^{(\sigma)} = \frac{1}{2}(\sigma_{ik} + \sigma_{ki}) - \sigma_{iso}. \tag{7}$$

The principal axis system of the second-rank symmetric tensor, $\mathbf{S}^{(\sigma)}$, is defined as the coordinate system, where $\mathbf{S}^{(\sigma)}$ is diagonal with principal components, $\lambda_{zz}^{(\sigma)}$, $\lambda_{yy}^{(\sigma)}$, and $\lambda_{zz}^{(\sigma)}$ ordered, according to the Haeberlen convention,[20] such that

$$|\lambda_{zz}^{(\sigma)}| \geq |\lambda_{xx}^{(\sigma)}| \geq |\lambda_{yy}^{(\sigma)}|. \tag{8}$$

Additionally, we define the second-rank symmetric tensor anisotropy, ζ_σ, and asymmetry parameter, η_σ, according to

$$\zeta_\sigma = \lambda_{zz}^{(\sigma)} \quad \text{and} \quad \eta_\sigma = \frac{\lambda_{yy}^{(\sigma)} - \lambda_{xx}^{(\sigma)}}{\zeta_\sigma}. \tag{9}$$

Similarly, following the IUPAC convention,[20] the notation ζ_δ represents the chemical shift anisotropy, given by

$$\zeta_\delta = -\zeta_\sigma, \tag{10}$$

however, in this article, we will keep our discussion in terms of the nuclear shielding anisotropy, ζ_σ. Furthermore, for the remainder of this article, we shall drop the σ subscript and indicate the nuclear shielding anisotropy and asymmetry parameters with ζ and η, respectively.

The first-order contribution of the nuclear shielding interaction to the NMR transition ($m_i \rightarrow m_f$) frequency is given by

$$\Omega_\sigma^{(1)}(\Theta, m_i, m_f) = -\omega_0\,\sigma_{iso}\,(m_f - m_i) - \omega_0\,\zeta\,\mathbb{D}^{\{\sigma\}}(\Theta)\,(m_f - m_i), \tag{11}$$

where $\omega_0 = -\gamma_I B_0$ is the Larmor frequency, γ_I is the gyromagnetic ratio of the nucleus, B_0 is the external magnetic flux density, and the orientation dependence of the shielding anisotropy is given by

$$\mathbb{D}^{\{\sigma\}}(\Theta) = P_2^0(\cos\beta) - \frac{\eta}{6}P_2^2(\cos\beta)\cos 2\alpha, \tag{12}$$

where α and β define the orientation of the shielding tensor relative to the external magnetic field vector, $\mathbf{B}$, and $P_2^0(\cos\beta)$ and $P_2^2(\cos\beta)$ are the associated Legendre polynomials.

B. Linear inverse method

A one-dimensional linear problem is of the form

$$\underbrace{s_{true}(x_1) + e(x_1)}_{s(x_1)} = \int \mathcal{K}(x_1, x_2)\, f(x_2)\, dx_2, \tag{13}$$

where $\mathcal{K}(x_1, x_2)$ is a known *kernel*, $s(x_1)$ is the observed *signal*, and $f(x_2)$ is the desired unknown underlying probability distribution,

referred to as the *model*. The terms, $s_{\text{true}}(x_1)$ denotes the true signal and $e(x_1)$ is the measurement noise, assumed to be white Gaussian with the standard deviation, σ_e. Here, the inverse problem is solving for the unknown model, $f(x_2)$. An equivalent representation of Eq. (13) in the matrix notation follows

$$\mathbf{s} = \mathbf{K} \cdot \mathbf{f}, \tag{14}$$

where $\mathbf{K} \in \mathbb{R}^{m\times n}$ is a matrix, $\mathbf{s} \in \mathbb{R}^m$, $\mathbf{f} \in \mathbb{R}^n$ are the column vectors, and m and n are the sampling size of the vectors $\mathbf{s}$ and $\mathbf{f}$, respectively.

A classical approach in solving the problem in Eq. (14) is the ordinary least-squares (OLS) estimator, which solves for $\mathbf{f}$ by minimizing the mean square error,

$$\mathbf{f}_{\text{OLS}} = \underset{\mathbf{f}}{\text{argmin}}\left(\frac{1}{m}\|\mathbf{K}\cdot\mathbf{f} - \mathbf{s}\|_2^2\right), \tag{15}$$

where $\|\mathbf{z}\|_2 = \sqrt{\sum_i z_i^2}$ denotes the ℓ_2-norm of vector $\mathbf{z}$ and $\mathbf{f}_{\text{OLS}}$ is the ordinary least-squares solution. When $m = n$ and $\mathbf{K}$ is nonsingular, the solution to the OLS estimator has a closed-form expression, given as $\mathbf{K}^{-1}\cdot\mathbf{s}$. In practice, however, most problems suffer from a near-singular and ill-conditioned kernel, where $n \gg m$. In such a situation, the OLS estimator in Eq. (15) is ill-posed because there exist infinite least-squares solutions, and the solutions are unstable in the presence of noise, i.e., a small perturbation in $\mathbf{s}$, from the noise, results in a massive fluctuation in the solution, $\mathbf{f}_{\text{OLS}}$.

A common approach in solving such ill-posed problems is to employ a regularization[45,55–59] method of the form

$$\mathbf{f}_g = \underset{\mathbf{f}}{\text{argmin}}\left(\frac{1}{m}\|\mathbf{K}\cdot\mathbf{f} - \mathbf{s}\|_2^2 + g(\mathbf{f})\right), \tag{16}$$

where $g(\mathbf{f})$ is the regularization or the *penalty* function that stabilizes the solution and $\mathbf{f}_g$ is a regularized solution that satisfies the criterion of the penalty function. In doing so, however, the solution becomes dependent on the regularization term, and choosing a correct regularization, therefore, becomes crucial.

In this work, our objective is to determine the distribution of the nuclear shielding tensor parameters from a 2D NMR measurement correlating the isotropic with the anisotropic frequencies. In choosing the penalty function, we consider the following—first, the NMR measurement may suffer from a low signal-to-noise ratio and, second, the probability distribution of the shielding tensors is likely to be smooth with correlated amplitudes for amorphous materials and sparse for crystalline materials. Therefore, our choice of the penalty functions includes the Truncated Singular Value Decomposition (TSVD) method, which suppresses the effect of noise in the solution, and the Smooth Least Absolute Shrinkage and Selection Operator (S-LASSO), a method that promotes both smoothness and sparsity in the solution. Note, our choice of the penalty is based on what we believe is best for describing the NMR tensor parameters and does not suggest a unique penalty function.

1. TSVD regularized least-squares problem

The truncated singular value decomposition (TSVD) is a frequently used regularization method[57,60] in solving the ill-posed linear-inverse problems. A singular value decomposition (SVD) of the kernel $\mathbf{K}$ is

$$\mathbf{K} = \mathbf{U}\cdot\boldsymbol{S}\cdot\mathbf{V}^T, \tag{17}$$

where the columns of $\mathbf{U} \in \mathbb{R}^{m\times m}$, denoted as $\mathbf{u}_i$, and $\mathbf{V} \in \mathbb{R}^{n\times n}$, denoted as $\mathbf{v}_i$, form an orthonormal basis set of the left and right singular vectors, respectively. The diagonal matrix, $\boldsymbol{S} \in \mathbb{R}^{m\times n}$, contains non-negative real singular values, ς_i, arranged in the decreasing order. The superscript T is the matrix transpose.

The singular value decomposition is a powerful method because it is stable, regardless of the form of $\mathbf{K}$, and the singular values readily determine the *ill-conditioned* nature of the kernel. If at least one singular value is zero, the kernel is *singular* with a rank p, where p is the number of non-zero singular values. In this case, there exists a *non-trivial null space* of $\mathbf{K}$ that is spanned by the last $(n - p)$ columns of $\mathbf{V}$, denoted here by $\mathbf{V}_0$. Any arbitrary vector $\mathbf{f}_0$ from the null space, $\mathbf{f}_0 = \mathbf{c}_0\cdot\mathbf{V}_0$, where $\mathbf{c}_0$ is an arbitrary diagonal matrix, will always satisfy the condition $\mathbf{K}\cdot\mathbf{f}_0 = \mathbf{0}$, and therefore, determining these vectors from the measurement, $\mathbf{s}$, is not feasible. In such cases, $\mathbf{K}$ is approximately expressed in a subspace, called the *range space*, by truncating the singular values to the first p largest values, following

$$\mathbf{K} \approx \mathbf{U}_p\cdot\boldsymbol{S}_p\cdot\mathbf{V}_p^T, \tag{18}$$

where the columns of $\mathbf{U}_p \in \mathbb{R}^{m\times p}$ and $\mathbf{V}_p \in \mathbb{R}^{n\times p}$ form the basis vectors that span the *range space* of $\mathbf{K}$ and $\mathbf{K}^T$, respectively. The diagonal square matrix $\boldsymbol{S}_p \in \mathbb{R}^{p\times p}$ contains the first p singular values. The columns of $\mathbf{U}_p$ and $\mathbf{V}_p$ are the first p columns of $\mathbf{U}$ and $\mathbf{V}$, respectively. In this case, the SVD solution to the problem in Eq. (14) has a simple analytical form

$$\mathbf{f}_\dagger = \mathbf{K}^\dagger\cdot\mathbf{s} = \mathbf{V}_p\cdot\boldsymbol{S}_p^{-1}\cdot\mathbf{U}_p^T\cdot\mathbf{s} = \mathbf{V}_p\cdot\boldsymbol{S}_p^{-1}\cdot\mathbf{U}_p^T\cdot(\mathbf{s}_{\text{true}} + \mathbf{e}), \tag{19}$$

where $\mathbf{f}_\dagger$ is the TSVD solution, $\mathbf{K}^\dagger$ is the pseudoinverse of $\mathbf{K}$, and $\boldsymbol{S}_p^{-1}$ is obtained by taking the inverse of all diagonal elements in $\boldsymbol{S}_p$. The above expression can be re-expressed as a linear combination of the singular vectors, $\mathbf{v}_i$,

$$\mathbf{f}_\dagger = \sum_{i=1}^{p}\left(\frac{\mathbf{u}_i^T\cdot\mathbf{s}_{\text{true}}}{\varsigma_i}\right)\mathbf{v}_i + \sum_{i=1}^{p}\left(\frac{\mathbf{u}_i^T\cdot\mathbf{e}}{\varsigma_i}\right)\mathbf{v}_i. \tag{20}$$

An interesting case arises when the singular values quickly decay to zero. This happens when $\mathbf{K}$ is *ill-conditioned*. The term *condition number* is defined as the ratio of the largest to the smallest singular value of $\mathbf{K}$ in the range space,

$$\text{cond}(\mathbf{K}) = \frac{\varsigma_1}{\varsigma_p}, \tag{21}$$

and provides a maximum bound on the relative error in the solution, $\mathbf{f}_\dagger$, given a relative error in measurement. In Eq. (20), note that the singular vectors, $\mathbf{v}_i$, corresponding to small singular values, ς_i, contribute massively to the solution $\mathbf{f}_\dagger$, and therefore, any perturbations arising from the noise $\mathbf{e}$ are amplified by these small singular values. A kernel with a large condition number is, thus, more susceptible to an unstable solution.

To regularize the solution, the original ill-conditioned kernel, $\mathbf{K}$, is replaced with an approximate well-conditioned kernel through the truncation of the singular values. If the first $r < p$ largest singular values are considered, then the reduced space approximate kernel follows

$$\mathbf{K}_{\text{TSVD}} = \mathbf{U}_r \cdot \mathbf{S}_r \cdot \mathbf{V}_r^T, \tag{22}$$

where $\mathbf{U}_r \in \mathbb{R}^{m\times r}$, $\mathbf{V}_r \in \mathbb{R}^{n\times r}$, and $\mathbf{S}_r \in \mathbb{R}^{r\times r}$ are the truncated matrices. In this case, the regularization parameter is $(p - r)$. The index r is chosen such that the condition number, $(\varsigma_1/\varsigma_r)$, of the truncated kernel, $\mathbf{K}_{\text{TSVD}}$, is relatively small. In this study, we employ the maximum entropy-based criterion[48] for selecting r (see Subsection 1 of the Appendix for details).

The truncated singular value decomposition (TSVD) method is very efficient in projecting the ordinary least-squares (OLS) problem onto a smaller subspace,[61] which dramatically reduces the computation time and resources, primarily when m is large. This process is also referred to as *data compression*. Having determined the optimal truncation index $r < p$, the linear problem in Eq. (14) can be approximately re-defined on the reduced subspace following

$$\tilde{\mathbf{s}} = \tilde{\mathbf{K}} \cdot \mathbf{f}, \tag{23}$$

where $\tilde{\mathbf{K}} = \mathbf{S}_r \cdot \mathbf{V}_r^T \in \mathbb{R}^{r\times n}$ is the "nearby" well-conditioned kernel and $\tilde{\mathbf{s}} = \mathbf{U}_r^T \cdot \mathbf{s} \in \mathbb{R}^r$ is the compressed signal. Similarly, the regularized least-squares estimator from Eq. (16) reformulates to

$$\underset{\mathbf{f}}{\text{argmin}}\left(\frac{1}{r}\|\tilde{\mathbf{K}} \cdot \mathbf{f} - \tilde{\mathbf{s}}\|_2^2 + g(\mathbf{f})\right). \tag{24}$$

Note, here, we use r to calculate the mean square error (MSE) of the reduced subspace problem.

2. Smooth-LASSO, S-LASSO

The Smooth Least Absolute Shrinkage and Selection Operator[50] (S-LASSO) estimator in a variant of the Elastic-Net estimator[49] in which the penalty function for a d-dimensional solution $\mathbf{f}$ is a combination of the quadratic and LASSO penalties of form

$$g(\mathbf{f}) = \alpha \sum_{i=1}^{d} \|\mathbf{J}_i \cdot \mathbf{f}\|_2^2 + \lambda\|\mathbf{f}\|_1, \tag{25}$$

where the first term is the quadratic penalty term that promotes a smooth solution and the second term is the LASSO penalty term that promotes sparsity in the solution. The multipliers, $\alpha > 0$ and $\lambda > 0$, are the *regularization* parameters that control the smoothness and sparsity of the solution, respectively. They are also known as the *hyperparameters*, and sometimes called the *tuning* parameters. The notation $\|\mathbf{f}\|_1 = \sum_i |f_i|$ denotes the ℓ_1-norm of $\mathbf{f}$. The matrix, $\mathbf{J}_i$, typically reflects some underlying geometry or the structure in the true solution. Here, $\mathbf{J}_i$ is defined to promote smoothness along the ith dimension of a d-dimensional solution $\mathbf{f}$ and is given as

$$\mathbf{J}_i = \mathbf{I}_{n_1} \otimes \cdots \otimes \mathbf{A}_{n_i} \otimes \cdots \otimes \mathbf{I}_{n_d}, \tag{26}$$

where $\mathbf{I}_{n_i} \in \mathbb{R}^{n_i\times n_i}$ is an identity matrix, $\mathbf{A}_{n_i}$ is a first difference matrix given by

$$\mathbf{A}_{n_i} = \begin{pmatrix} 1 & -1 & 0 & \cdots & \vdots \\ 0 & 1 & -1 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & \cdots & 0 & 1 & -1 \end{pmatrix} \in \mathbb{R}^{(n_i-1)\times n_i}, \tag{27}$$

and the symbol $\otimes$ is the Kronecker product. The terms, $(n_1, n_2, \ldots, n_d)$, are the number of points along the solution dimensions, with the constraint that $\prod_{i=1}^{d} n_i = n$. Note, the penalty function in Eq. (25) reduces to the Elastic-Net penalty when $\mathbf{J}_i$ is the identity matrix and further reduces to a LASSO penalty when $\alpha = 0$.

In the case of a one-dimensional ($d = 1$) solution, i.e., when $\mathbf{f}$ is a vector, the matrix

$$\mathbf{J}_1 = \mathbf{A}_{n_1} \in \mathbb{R}^{(n_1-1)\times n_1} \tag{28}$$

promotes a smooth solution by penalizing the squared difference between the adjacent elements of the vector $\mathbf{f}$. In the case of a two-dimensional ($d = 2$) solution, e.g., when $\mathbf{f}$ is an image, the $\mathbf{J}_1$ and $\mathbf{J}_2$ matrices,

$$\mathbf{J}_1 = \mathbf{A}_{n_1} \otimes \mathbf{I}_{n_2} \in \mathbb{R}^{(n_1-1)n_2\times n_1 n_2}, \tag{29}$$

$$\mathbf{J}_2 = \mathbf{I}_{n_1} \otimes \mathbf{A}_{n_2} \in \mathbb{R}^{(n_2-1)n_1\times n_2 n_1}, \tag{30}$$

promotes smoothness by penalizing the squared differences between the adjacent elements of $\mathbf{f}$ along both dimensions.

The advantage of having a quadratic penalty term in Eq. (25) is that it is easily integrable into the OLS term of Eq. (24) by defining an augmented dataset,

$$\mathbf{K}_d = \begin{pmatrix} \tilde{\mathbf{K}} \\ \sqrt{r\alpha}\,\mathbf{J}_1 \\ \vdots \\ \sqrt{r\alpha}\,\mathbf{J}_d \end{pmatrix} \quad \text{and} \quad \mathbf{s}_d = \begin{pmatrix} \tilde{\mathbf{s}} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}, \tag{31}$$

where $\mathbf{K}_d$ is the augmented kernel, $\mathbf{s}_d$ is the augmented signal, and $\mathbf{0}$ is a vector of zeros. Thereafter, the solution to the problem in Eq. (24), with the penalty term given by Eq. (25), is equivalent to solving the following LASSO problem[62–71] over the augmented dataset,

$$\mathbf{f}_g = \underset{\mathbf{f}\geq 0}{\text{argmin}}\left(\frac{1}{r}\|\mathbf{K}_d \cdot \mathbf{f} - \mathbf{s}_d\|_2^2 + \lambda\|\mathbf{f}\|_1\right). \tag{32}$$

In this study, we solve for a two-dimensional solution, $\mathbf{f}$, and implement the LASSO estimator from the scikit-learn[51,52] python package with a non-negative constraint on $\mathbf{f}$ to solve Eq. (32). A non-negative constraint is essential for representing a pragmatic solution by ensuring that the probability of a tensor cannot be negative. In determining the optimum hyperparameters α^* and λ^*, we employ the ten-fold cross-validation method (see Subsection 2 of the Appendix for further details).

3. Role of the hyperparameters

In the TSVD-S-LASSO hybrid linear inversion method, there are three hyperparameters, $(p - r)$, α, and λ. The first hyperparameter, $(p - r)$, is the number of truncated singular values whose role is to limit the effect of the measurement noise on the solution and reduce the dimensionality of the problem. The hyperparameters α and λ control the smoothness and sparsity of the solution. The role of α and λ becomes fairly intuitive when we re-express Eq. (25) as

$$g(\mathbf{f}) = a\left(b\sum_{i=1}^{d}\|\mathbf{J}_i\cdot\mathbf{f}\|_2^2 + (1-b)\|\mathbf{f}\|_1\right), \tag{33}$$

where

$$a = \alpha + \lambda \quad \text{and} \quad b = \alpha/(\alpha+\lambda). \tag{34}$$

Here, the parameter a controls the amount of regularization used in the inversion and the parameter b controls the relative amount of ℓ_1 and ℓ_2 contributions. Generally, the optimum value of a depends on the noise standard deviation, i.e., greater noise leads to a larger a. A larger value of b promotes a smooth and blurred solution, while a smaller value promotes a sparse and sharp solution. The optimum value of b will also depend on the 2D inversion grid interval (*vide infra*), i.e., course grids tend to favor sparse solutions, while fine grids tend to favor smooth solutions.

C. Practical challenges

1. Defining the 2D inversion grid

A common choice for representing a distribution of ζ and η parameters is a 2D Cartesian grid where the coordinates ζ and η span the orthogonal dimensions. As mentioned earlier, this coordinate system can be problematic since η becomes undefined as $\zeta \to 0$; this is indicated as black-filled circles in Fig. 1(a). Additionally, in the case of the nuclear shielding interaction, the anisotropic line shapes are invariant of the sign of ζ when $\eta = 1$, forming a degenerate system of line shapes; this is depicted with colored circles in Fig. 1(a). To overcome these issues, we adopt an approach similar to Czjzek[72] by re-expressing the coordinates ζ and η in the first quadrant of a polar coordinate system, (r_ζ, θ), where

$$r_\zeta = |\zeta| \quad \text{and} \quad \theta = \begin{cases} \frac{\pi}{4}\eta, & \zeta \le 0 \\ \frac{\pi}{2}\left(1 - \frac{\eta}{2}\right), & \zeta > 0. \end{cases} \tag{35}$$

We choose an x–y Cartesian grid over the r–θ polar grid because Cartesian grids are more manageable for implementing the inversion algorithm. It is given by

$$x = r_\zeta \cos\theta \ \text{ and } \ y = r_\zeta \sin\theta, \quad 0 \ge \theta \ge \pi/2. \tag{36}$$

As shown in Fig. 1(b), the magnitude of ζ forms the radial dimension, while η forms the angular dimension. The line $|x| = |y|$ corresponds to $\eta = 1$. When progressing toward the x or y-axis from this line, η uniformly decreases from 1 to 0, where $\eta = 0$ is along the x or y-axis, depending on the sign of ζ. The undefined condition for η when $\zeta \to 0$ is true, irrespective of the choice of the coordinate system; however, this representation confines $\zeta = 0$ to a single point, located at the origin. Additionally, since the positive and negative values of ζ correspond to the same (r, θ) coordinate when $\eta = 1$, the x–y representation also removes the anisotropic line shape degeneracy associated with the ζ–η grid.

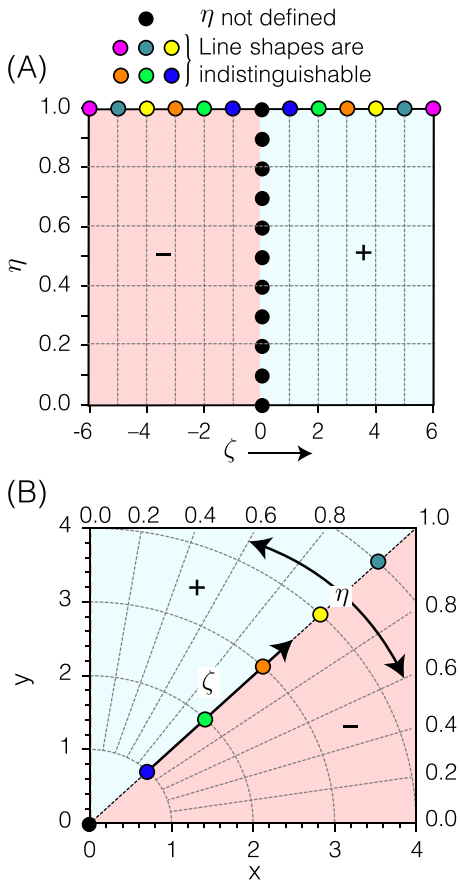

FIG. 1. (a) A schematic representation of the ζ–η Cartesian grid system depicting the undefined regions, shown in black circles, and redundant indistinguishable regions, shown in colored circles, of the coordinate system. (b) The positive quadrant of the x–y grid. Here, the magnitude of ζ is the radial dimension, while η is the angular dimension. The blue and red shading indicate the regions of positive and negative ζ, respectively.

In the x–y grid representation, the integral in Eq. (4) is expressed as

$$s(\nu|\delta_{\text{iso}}) = \oint_{x,y} \mathcal{K}_{xy}(\nu, x, y) f(x, y|\delta_{\text{iso}})\partial x \partial y, \tag{37}$$

with the corresponding discretized approximation

$$s(\nu|\delta_{\text{iso}}) \approx \sum_{i=0}^{n_1-1}\sum_{j=0}^{n_2-1} K_{xy}(\nu, x_i, y_j) f(x_i, y_j|\delta_{\text{iso}})\Delta x \Delta y, \tag{38}$$

where n_1, n_2, and Δx, Δy are the number of points and increments along the x and y dimensions, respectively. The discretized coordinates, (x_i, y_j), span a predefined grid of x and y coordinates, respectively.

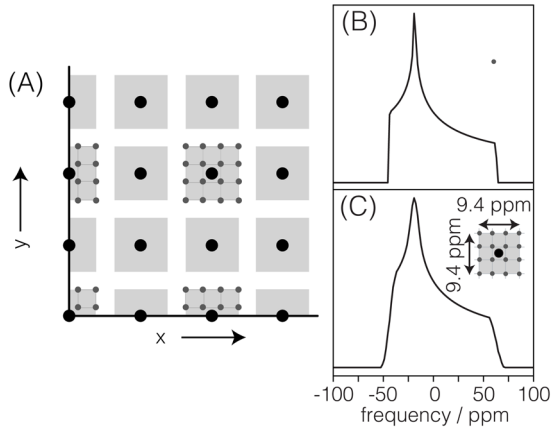

FIG. 2. (a) In this illustration of the supersampling scheme, the square region around each base grid cell coordinate, indicated by a bold circle, is sub-divided into a 4 × 4 sub-grid. (b) The subspectrum at each sub-grid point, indicated by the smaller circle, is averaged together to create (c) a basis subspectrum associated with the corresponding base grid cell coordinate. Note, the base grid cells along the axes are averaged over one-half of the cell area, except for the cell at the origin, which is averaged over one-fourth of the cell area.

When determining the grid resolution, it is necessary to ensure that the grid increments are sufficiently small such that the approximation in Eq. (38) remains valid. This requires a massive grid system of basis subspectra—10^4–10^6 grid points—which can be computationally expensive. An alternative is to use a supersampling scheme, illustrated in Fig. 2, in which each grid cell area is subsampled into an $n_x \times n_y$ sub-grid, with sufficiently small sub-grid intervals chosen to obtain a valid approximation of Eq. (37). The $n_x \times n_y$ subspectra, associated with each point on a sub-grid, are then averaged together to create the basis subspectrum associated with each grid cell.

2. Defining the kernel

While the probability distribution in Eq. (38), $f(x_i, y_j|\delta_{\text{iso}}) \in \mathbb{R}^{n_1 \times n_2}$, is two-dimensional, when solving the linear inverse problem in Eq. (16), $\mathbf{f} \in \mathbb{R}^n$, is required to be a vector. Therefore, for all calculations, we define $\mathbf{f}$ as the row-major vectorization of $f(x_i, y_j|\delta_{\text{iso}})$, where the $(c = jn_1 + i)$th element of $\mathbf{f}$ is $f(x_i, y_j|\delta_{\text{iso}})$. Similarly, the cth column of the matrix $\mathbf{K}$ is $K_{xy}(v, x_i, y_j) \in \mathbb{R}^m$ and holds the subspectrum corresponding to the coordinates (x_i, y_i).

Having determined the kernel $\mathbf{K}$, all further operations are subject to the TSVD method, followed by smooth-LASSO optimization, as described in Sec. II B. The optimum inverse solution $\mathbf{f} \in \mathbb{R}^{n_1 n_2}$ is then interpreted as an $n_1 \times n_2$ matrix.

III. RESULTS AND DISCUSSION

A. Synthetic datasets

In this section and in the supplementary material, we examine the inversion of synthetic purely anisotropic 1D spectra generated from the known (ground truth) bivariate distributions of nuclear shielding anisotropy parameters, i.e., $f(x, y)$. Here, the term "purely anisotropic spectrum" refers to situations where $\delta_{\text{iso}}^{(cs)} = 0$ for every subspectrum contributing to a spectrum, which is the case for the anisotropic cross sections of the spectra from isotropic/anisotropic correlation methods, such as 2D MAF, PASS, or MAT, either as acquired or after an appropriate shear transformation is applied.[73–75] The anisotropic cross sections in 2D MAF are purely anisotropic variable angle spinning (90°) spectra, and in the 2D PASS or 2D MAT spectrum are purely anisotropic MAS sideband spectra.

All NMR simulations presented here were performed for a ^{29}Si nuclide at 9.4 T using the python *mrsimulator* package.[76] The synthetic VAS (90°) spectra use 96 points spanning over 20 kHz and spinning at 14 kHz. Except where indicated, the synthetic MAS sideband spectra used up to 32 sideband orders with a MAS spinning speed of 625 Hz.

1. Uni-modal distributions

Figure 3 depicts the results of the smooth-LASSO inversion method on spectra obtained from four different ground truth unimodal distributions shown in the first column. On the top and right of each ground truth distribution contour plot are the projections of the respective distribution onto the x and y coordinate axes, respectively.

The ground truth distribution labeled U-1 is a bivariate normal distribution generated in the x and y coordinates, with mean values of $\mu_x = 35$ ppm and $\mu_y = 75$ ppm, standard deviations of $\sigma_x = 5$ ppm and $\sigma_y = 5$ ppm, and a correlation coefficient of $r_{xy} = 0.12$. The ground truth distribution labeled U-2 in Fig. 3 is also a bivariate normal distribution but generated in the ζ and η coordinates and mapped onto the x and y grid. This distribution has mean values of $\mu_\zeta = -75$ ppm and $\mu_\eta = 0.5$, standard deviations of $\sigma_\zeta = 3.873$ ppm, $\sigma_\eta = 0.173$, and a correlation coefficient of $r_{\zeta,\eta} = 0$.

The ground truth distribution labeled U-3 is a Czjzek probability distribution[77,78] generated in the ζ and η coordinates using the expression

$$p^{(c)}(\zeta, \eta) = \frac{\zeta^4 \eta}{\sqrt{2\pi}\sigma_c^5}\left(1 - \frac{\eta^2}{9}\right)\exp\left\{-\frac{\zeta^2\left(1 + \frac{\eta^2}{3}\right)}{2\sigma_c^2}\right\}, \quad (39)$$

with $\sigma_c = 6.62$ ppm, and mapped onto the x and y grids. The Czjzek distribution—originally developed to model random distributions of electric field gradients (EFGs) in glasses—is a model for anisotropic line shapes arising from random deviations from a mean anisotropy of zero. The Czjzek distribution assumes uncorrelated Gaussian distributions of second-rank spherical tensor components with a single width parameter and gives the distribution of anisotropy parameters, ζ and η. It is a valuable model for identifying anisotropic line shapes arising from a random distribution of a second-rank NMR tensors.

The ground truth distribution labeled U-4 is generated from an extended Czjzek probability distribution, $p^{(xc)}(\zeta, \eta)$, using the numerical approach outlined by Le Caër *et al.*[79] It has a mean anisotropy of $\mu_\zeta = -25$ ppm and a mean asymmetry of $\mu_\eta = 0.2$ and uses a perturbation factor of $\epsilon = 0.35$.

In the second and fourth columns of Fig. 3 are the corresponding synthetic VAS (90°) and MAS sideband spectra, respectively, shown in black, obtained after adding Gaussian noise of $\sigma_e = 0.005$

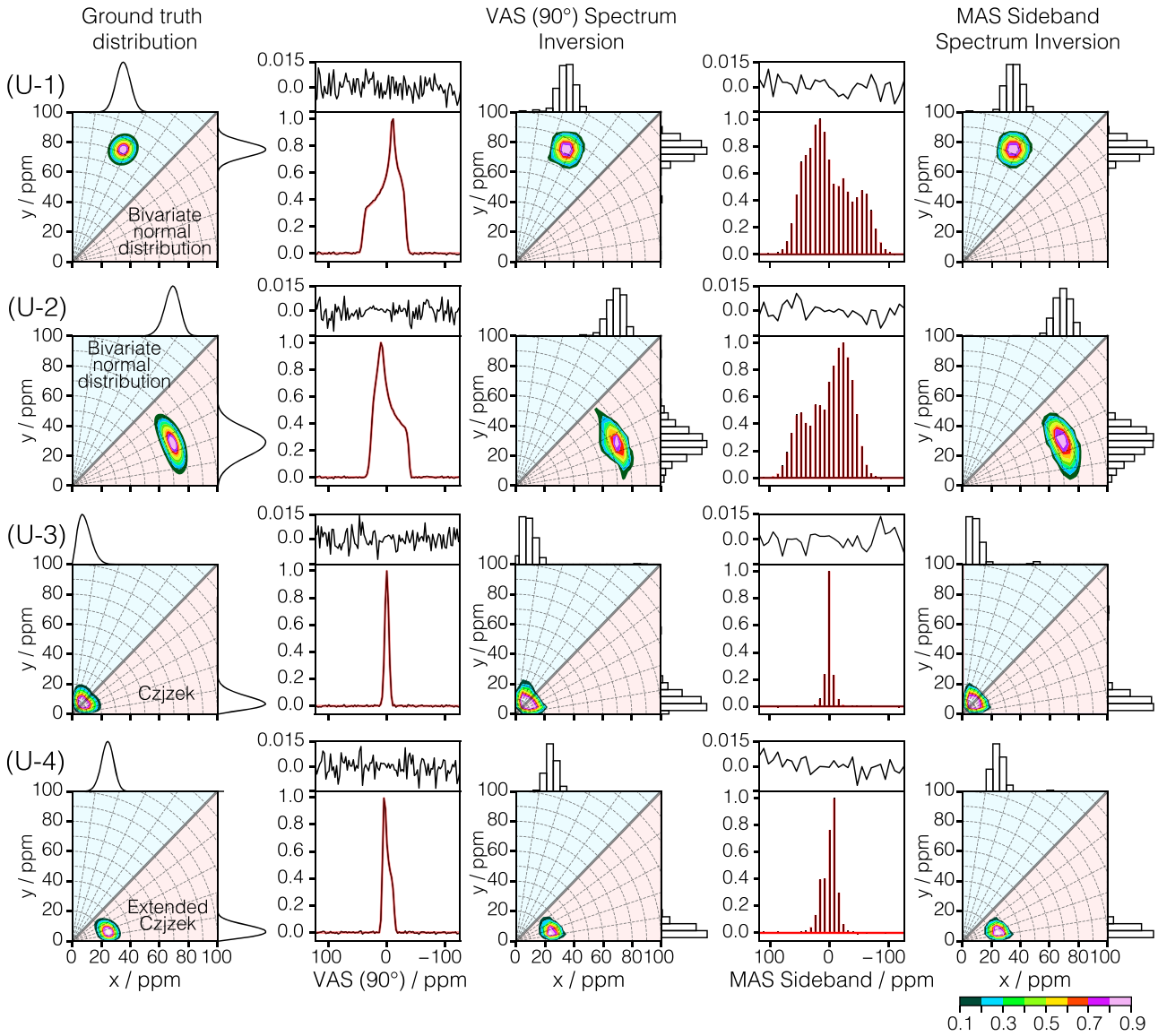

FIG. 3. Comparison of smooth-LASSO inversion of synthetic purely anisotropic spectra originating from four different unimodal bivariate shielding tensor distributions along rows labeled U-1, U-2, U-3, and U-4. In the first column are the ground truth distributions along with 1D projections onto the coordinate axes. In the second and fourth columns are the corresponding synthetic VAS (90°) spectrum and MAS sideband spectrum, respectively, with noise, σ_e = 0.005, along with the best fit spectrum shown in red. Above each spectrum are the residuals. In the third and fifth columns are the corresponding inversion solutions for the shielding tensor distributions shown with histogram projections onto the coordinate axes. The contours are drawn at every 10%.

to each spectrum simulated from its ground truth shielding tensor distribution. The overlaid spectra, shown in red in the second and fourth columns of Fig. 3, are the fits after the inversion with the fit residuals shown in the plot above each spectrum. Note that the synthetic VAS (90°) line shapes arising from a Czjzek distribution, U-3 in Fig. 3, are Gaussian.

In order to obtain a realistic synthetic dataset, each ground truth distribution is sampled over a high resolution (128 × 128) x–y grid with a cell resolution of 0.942 × 0.942 ppm^2. The inversion of each synthetic anisotropic spectrum is carried out over a lower resolution (25 × 25) x–y grid, where each 4.65 × 4.65 ppm^2 grid cell (basis) subspectrum is created by supersampling on a 5 × 5 subgrid, as detailed in Fig. 2. In the third and fifth columns of Fig. 3 are the contour plots of the best-fit model parameters, i.e., the nuclear shielding tensor distributions, obtained from the smooth-LASSO inversion of the corresponding synthetic VAS (90°) and MAS

sideband spectra, respectively, with histograms on the top and right showing the projection of the best-fit distribution onto the x and y coordinate axes. Good agreement between the ground truth and inverted distributions is obtained in all cases, as can be seen in both the contour plots and the histogram projections.

While a cell-by-cell comparison of the ground truth and inverted distributions cannot be made due to the differences in sampling grid resolution, we can still quantify the inversion performance by comparing the mean and standard deviation of the ground truth and inverted distributions, calculated in the x–y coordinates, shown in Table S1 of the supplementary material. In the case of U-1, the mean positions of the inverted distributions along both x and y directions are within 0.5 ppm of the ground truth distribution. This is well within the inversion grid cell resolution and even within the resolution of the sub-grid (0.93 ppm) used for supersampling. The inverted distribution widths (standard deviations) along the x and y directions are within 0.4 ppm of the ground truth distribution value of 5 ppm. In the U-2, U-3, and U-4 case, the mean positions, standard deviations, and correlation coefficients of the inverted distributions along the x and y directions are also in reasonably good agreement with their corresponding ground truth distributions. While higher moment analyses could be used for a more detailed comparison of the U-2, U-3, and U-4 distribution shapes, these moments tend to be less useful as they are more susceptible to noise.

2. Cross-validation and hyperparameters

The optimum hyperparameters, λ^* and α^*, for each inversion are determined by examining the cross-validation error metric from the ten-fold cross-validation method on a 20 × 20 grid of predefined λ and α values. Further details on the cross-validation technique are given in Subsection 2 of the Appendix. A total of 4000 solutions are assessed from which the model parameters with the least cross-validation error are selected. Typical contour plots of the cross-validation error, in the case of the inversion of the U-1 synthetic VAS (90°) and MAS sideband spectra, are shown in Figs. 4(a) and 4(b), respectively, with the optimum λ^* and α^* values marked with an "x." The optimum λ^* and α^* values from each inversion are given in Table S2. Additionally, Table S2 of the supplementary material lists the alternative hyperparameters, a^* and b^*, as defined in Eq. (34). Recall that a reflects the amount of regularization used in the inversion, while the hyperparameter b indicates the relative amount of ℓ_1 and ℓ_2 contributions. To illustrate these roles, we examine, in the supplementary material, the smooth-LASSO inversion method on spectra as a function of increasing noise standard deviations and increasing width of the ground truth distribution.

3. Regularization comparisons

To highlight the performance improvement and need for the smooth-LASSO model selection method, we inverted the synthetic VAS (90°) spectrum generated from the U-1 distribution with noise of $\sigma_e = 0.005$ and compared the results of (1) truncated singular value decomposition, (2) ℓ_1 regularization, (3) elastic-net, and (4) smooth-LASSO regularization approaches, as shown in the different rows of Fig. 5. When comparing the performance of these different regularization methods, it is important to examine variations with changing solution grid sizes, displayed at the top of each column in Fig. 5.

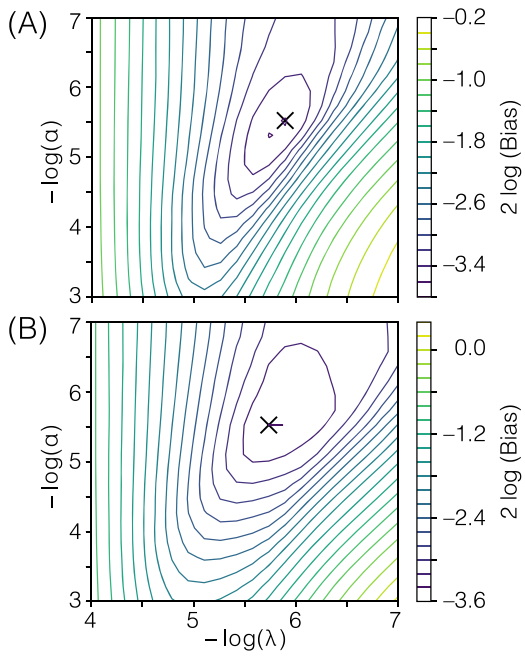

FIG. 4. The 20 × 20 α–λ hyperparameter grids displaying the contours of the cross-validation error [see Eq. (A7) in Subsection 2 of the Appendix], resulting from the ten-fold cross-validation for the inversion of (a) the synthetic VAS (90°) spectrum and (b) the MAS sideband spectrum, both generated from the U-1 ground truth distribution in Fig. 3. The marker "x" designates the optimum λ^* and α^* values.

The first row shows the TSVD solutions corresponding to the $r = 40$ largest singular values. Using only TSVD, regardless of the grid resolution, leads to broad and unphysical distributions with finite amplitudes extending over the entire solution grid as well as unphysical negative solutions. The choice of $r = 40$ was arbitrary in the sense that it was chosen to give best agreement to the ground truth distribution. An "optimal" truncation index of $r = 60$ was determined by the maximum entropy method but gave a noisy solution with little resemblance to the ground truth distribution. One might expect similarly poor performance from ridge regression (also known as Tikhonov regularization), as it also performs a weighted reduction in the singular values.[47]

The second row shows the result of the ℓ_1-norm regularization with a non-negative constraint on the solution. This type of regularization is well known for promoting sparse solutions, which is immediately obvious in comparison to TSVD solutions in Fig. 5. In this case, cross-validation was performed to determine the hyperparameter controlling the amount of ℓ_1-norm regularization. These values are available in Table S3 of the supplementary material. A challenge with ℓ_1 regularization, however, is that its solutions can become unreliable as the grid-resolution increases, as seen in the columns going from left to right in Fig. 5.

The third row shows the result from an elastic net regularization. As described in Sec. II B 2, this approach combines ℓ_1-norm and ℓ_2-norm regularization with two hyperparameters

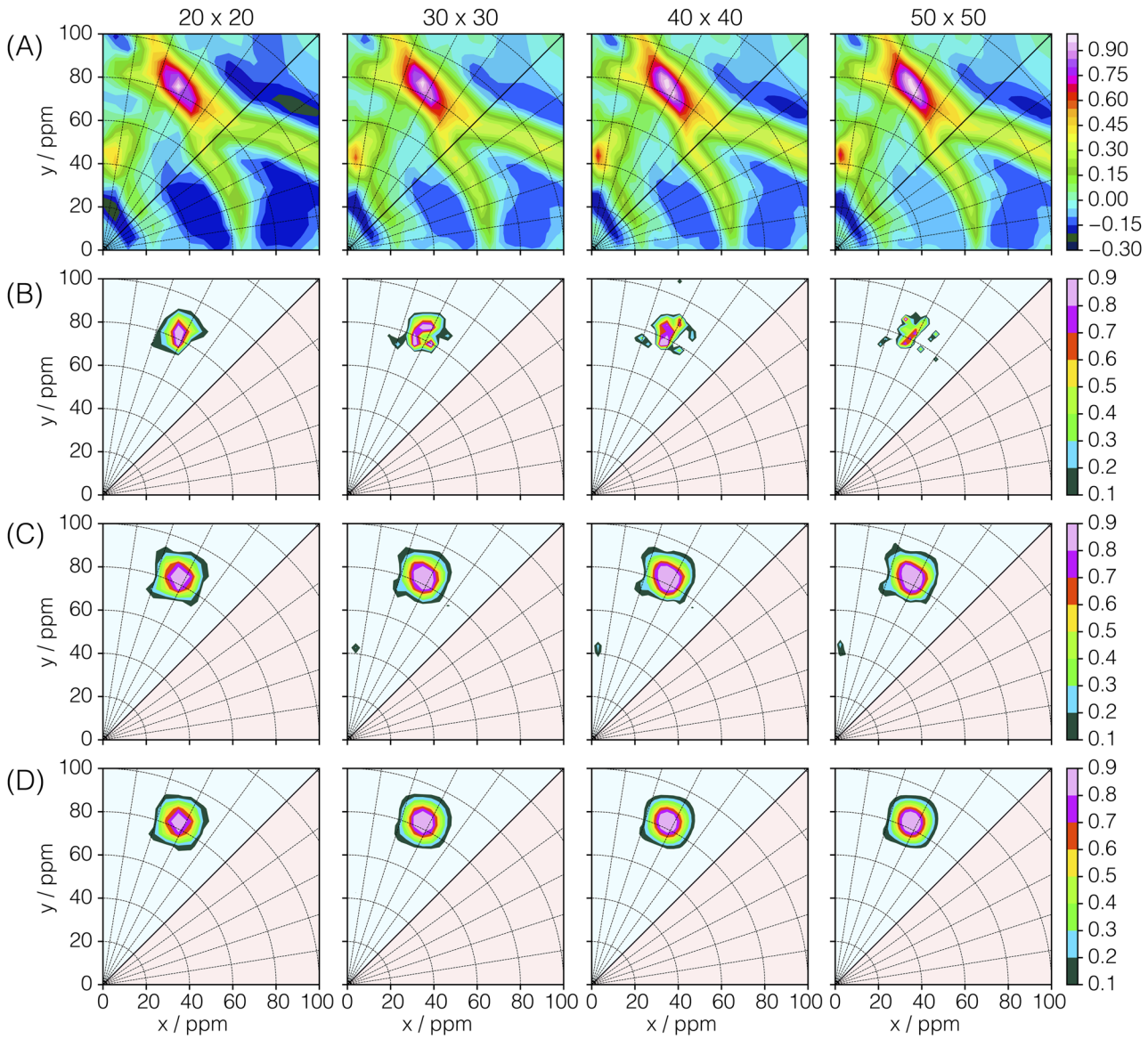

FIG. 5. Comparison of the inversion results on a synthetic VAS (90°) spectrum generated from the U-1 distribution with the noise of σ_e = 0.005 using (a) truncated singular value decomposition, (b) ℓ_1 regularization, (c) elastic-net, and (d) smooth-LASSO. From left to right, the columns represent solutions with increasing grid sizes of 20 × 20, 30 × 30, 40 × 40, and 50 × 50 displayed above each column. The TSVD solutions use the first 40 largest singular values. The contours are drawn at 5% for data in the first row and at 10% for the second, third, and fourth rows.

determining their respective weights. The difference between elastic net and the smooth-LASSO regularization is that the former employs no smoothing of the solution, i.e., the $\mathbf{J}_i$ in the penalty function, Eq. (25), are identity matrices. Compared to ℓ_1 regularization, the elastic net provides a more consistent inversion onto the various grid-resolutions.

Finally, the results in the fourth row are from the smooth-LASSO regularization. Like the elastic net, a consistent inversion solution is seen across different grid resolutions. For the most part, the elastic net and smooth-LASSO act in the same way as seen in their respective inversion results. An additional feature of the smooth-LASSO is the use of the first-difference matrix, i.e., Eq. (27), which promotes a smooth solution. This smoothness penalty suppresses any isolated spike in the solution amplitudes that may result from the noise in the spectrum. This can be seen by noting that the extraneous small signal around $x \approx 2$ ppm and $y \approx 40$ ppm in the elastic net solution is suppressed in the smooth-LASSO solution.

4. Dependence on MAS speed

As mentioned earlier, the MAS sideband basis subspectra become less distinguishable with increasing MAS rotor speeds, and we can expect greater uncertainty in the determined tensor parameter distribution. This is illustrated in Fig. 6 with a comparison of smooth-LASSO inversion on synthetic MAS sideband spectra as a function of rotor frequency or equivalently the number of sidebands. The corresponding optimum inversion hyperparameters are given in Table S3 of the supplementary material. The spectra in the first and third columns are generated from the U-1 distribution (μ_ζ of 6.6 kHz or 82.8 ppm) with noise of $\sigma_e = 0.0001$ using a range of rotor frequencies from 1.1875 Hz to 8 kHz. The corresponding inversion results, depicted in the second and fourth columns, show that the ground truth distribution can be reliably reconstructed from the spinning sideband spectrum as long as the number of sidebands is 12 or higher. Less than 12 sidebands and our results show that the information loss is considerable for a complete reconstruction of the ground truth distribution—a finding consistent with Hodgkinson and Emsley.[80] As also expected, the rank deficiency of the inversion

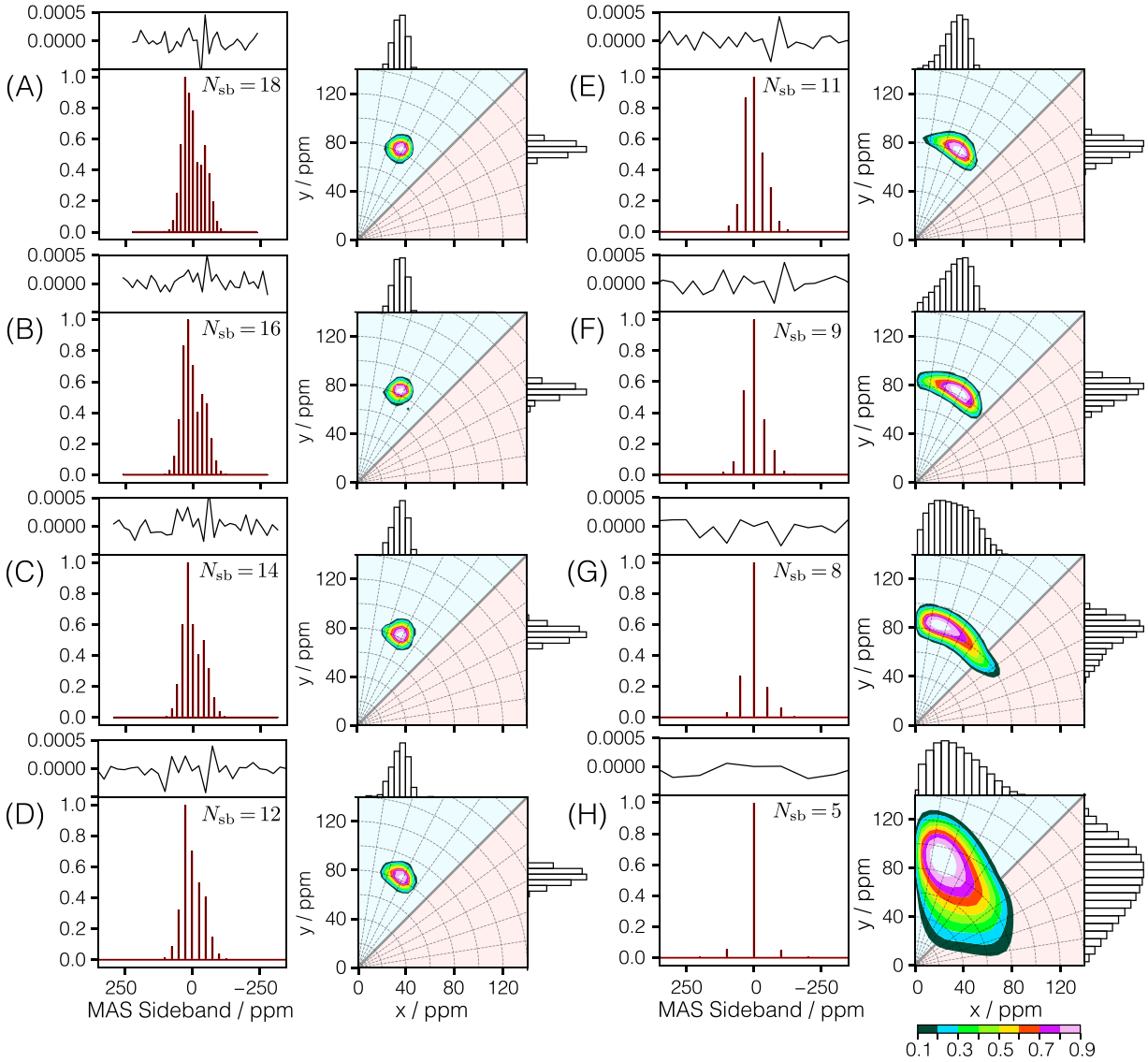

FIG. 6. Comparison of smooth-LASSO inversions, shown in second and fourth columns, on synthetic MAS sideband spectra, shown in the first and third columns, generated from the U-1 distribution with noise of $\sigma_e = 0.0001$ as a function of rotor frequency or equivalently the number of sidebands. The best fit spectra are shown in red. Rotor frequencies are (a) 1.1875 kHz, (b) 1.375 kHz, (c) 1.5625 kHz, (d) 2 kHz, (e) 2.5 kHz, (f) 3 kHz, (g) 4 kHz, and (h) 8 kHz. The number of sidebands, N_{sb}, displayed on the top-right corner of the spectrum, is calculated from the resulting spectrum as the number of points with amplitude greater than $6\sigma_e = 0.0006$.

kernel increases with increasing rotor frequency, causing a steady drop in the optimal truncation index, r, from a maximum possible value of $r = 32$ to $r = 31$ at rotor frequency of 1.1875 kHz to a value of $r = 12$ at 8 kHz (see Table S3 of the supplementary material).

B. Experimental ^{29}Si NMR datasets

In silicate networks there are five types of SiO_4 tetrahedra, each characterized by their connectivity, i.e., the number of oxygen atoms that are corner-linked to other tetrahedra.[81,82] These are designated with the notation Q^n, where $n(\sim 0–4)$ represents the number of bridging oxygen per tetrahedron. Quantifying the populations of Q^n species in a silicate glass and their degrees of connectivity can provide considerable insight into the structure of glass[83,84] as well as physical properties,[85–91] such as the glass transition temperature, viscosity, mechanical properties, and ionic (modifier) transport.

Under high-speed MAS conditions where anisotropic broadening (i.e., information on ζ and η) is removed, those five structural units can only be identified by their isotropic position, $\delta_{iso}^{(cs)}$, which reflects the mean silicon environment and ranges from approximately −70 ppm for Q^0 to −110 ppm for Q^4 with some considerable overlap.[92] In the majority of NMR studies of network forming glasses, where high-speed MAS spectra contain a number of resolved "resonances," spectroscopists focus almost entirely on using MAS spectra to identify and quantify populations of polyhedral units and polyhedral linkages.[93] By referring to "resonances" in quotes, we are highlighting the fact that these resolved "resonances" are inhomogeneously broadened, that is, inside each "resonance" is a mix of homogeneous resonances (subspectra) from numerous structurally distinct sites.

Under static conditions, off-magic-angle spinning, or slow MAS, each Q^n site exhibits a characteristic anisotropic ^{29}Si NMR line shape from the magnetic shielding originating from the anisotropic electron cloud surrounding the nucleus.[94,95] Thus, the shielding tensor can reveal the nature and directionality of the bonding, with ζ measuring the extent of the electron cloud distortion and η expressing the departure from a cylindrically symmetric environment [$\eta = 0$ (symmetric) → 1 (asymmetric)]. The Q^0 and Q^4 sites have the smallest anisotropy due to their highly symmetric tetrahedral environment. The remaining sites have ζ values of relatively larger magnitude, with Q^3 and Q^1 sites having low asymmetry parameters reflecting an environment close to axial symmetry and positive and negative signs of ζ, respectively, and with Q^2 site having intermediate asymmetry parameters.

As mentioned earlier, numerous NMR methods exist for obtaining a 2D spectrum correlating isotropic and anisotropic frequencies. In this section, we perform smooth-LASSO inversions on experimental 2D spectra, previously published in the 2D MAF and 2D MAT ^{29}Si studies of silicate glasses,[73–75,96,97] to obtain the 3D distributions of ^{29}Si shielding tensor principal components for the Q^n sites. For all 2D spectra presented in this section, an active shear[44,97] is applied to obtain signal correlating pure isotropic and anisotropic dimensions, leaving all anisotropic cross sections centered at 0 Hz. For the inversion of these 2D spectra, the kernel, $\mathbf{K} \in \mathbb{R}^{m\times n}$, is the same as before, but now the signal in Eq. (14) is two-dimensional, $\mathbf{s} \in \mathbb{R}^{m\times m_{iso}}$, where m_{iso} is the number of isotropic cross sections. Computational details for the smooth-LASSO inversion of each experimental 2D spectrum in this section are given in Table I.

The first set of experimental examples, shown in the first column of Fig. 7, are three ^{29}Si 2D MAF spectra for the silicate glass compositions $Cs_2O \cdot 4.72SiO_2$, obtained by Jardón-Álvarez *et al.*,[75] and $Na_2O \cdot 4.74SiO_2$ and $Rb_2O \cdot 2.25SiO_2$, obtained by Baltisberger *et al.*[74] Applying the smooth-LASSO inversion to each 1D anisotropic cross section along the isotropic dimension gives the 3D distributions of shielding tensor parameters for the corresponding 2D spectra shown in the second column of Fig. 7. The contours of the 2D projections from this 3D distribution onto the axes are shown along the face normal to the axis of projection. Similarly, the 1D projections from the respective 2D projections are shown on top of the respective 2D projection planes. Below each 2D spectrum in the first column is the best fit residuals obtained with its inversion solution. As shown in Fig. 7, the inversions yield bi-modal 3D tensor parameter distributions for all three 2D spectra. The distribution modes near the origin in x–y are assigned to the Q^4 sites. This assignment is also consistent with their projection onto the isotropic chemical shift dimension, appearing in the −100 ppm to −110 ppm range. The projections of this mode onto the x–y plane are also consistent with a Czjzek distribution, as expected for Q^4 sites with random structural variations away from a highly symmetric local structure

TABLE I. Computational details for smooth-LASSO inversion of experimental MAF and MAT spectra. The standard deviation of the measurement noise is σ_e. All inversions are performed on an x–y solution grid with the indicated size and resolution using a 5 × 5 sub-grid supersampling scheme. The optimum singular value truncation index, r, is determined from the maximum entropy method[48] (briefly described in Appendix). The optimum hyperparameters, α^* and λ^*, are determined over a 20 × 20 α–λ grid using the cross-validation error metric from the ten-fold cross-validation method. The signal to noise ratio is the reciprocal of σ_e.

Measurement		Inversion				Cross-validation			
Sample	σ_e	Kernel	Grid size	Grid resolution (ppm^2)	r	$\alpha^*/10^{-6}$	$\lambda^*/10^{-6}$	$a^*/10^{-6}$	b^*
$Cs_2O \cdot 4.72SiO_2$	0.002	MAF	25 × 25	5.66 × 5.66	98	0.834	0.616	1.45	0.58
$Na_2O \cdot 4.74SiO_2$	0.003	MAF	25 × 25	5.03 × 5.03	87	0.207	3.79	4.0	0.05
$Rb_2O \cdot 2.25SiO_2$	0.0045	MAF	25 × 25	5.03 × 5.03	88	0.886	3.79	4.68	0.19
$MgO \cdot SiO_2$	0.015	MAF	28 × 28	5.03 × 5.03	63	119.9	4.55	124.5	0.96
$CaO \cdot SiO_2$	0.0012	MAF	25 × 25	5.03 × 5.03	45	28.0	8.86	36.8	0.76
$KMg_{0.5}O \cdot 4SiO_2$	0.0007	MAT	25 × 25	4.65 × 4.65	31	3.79	11.1	15.1	0.25

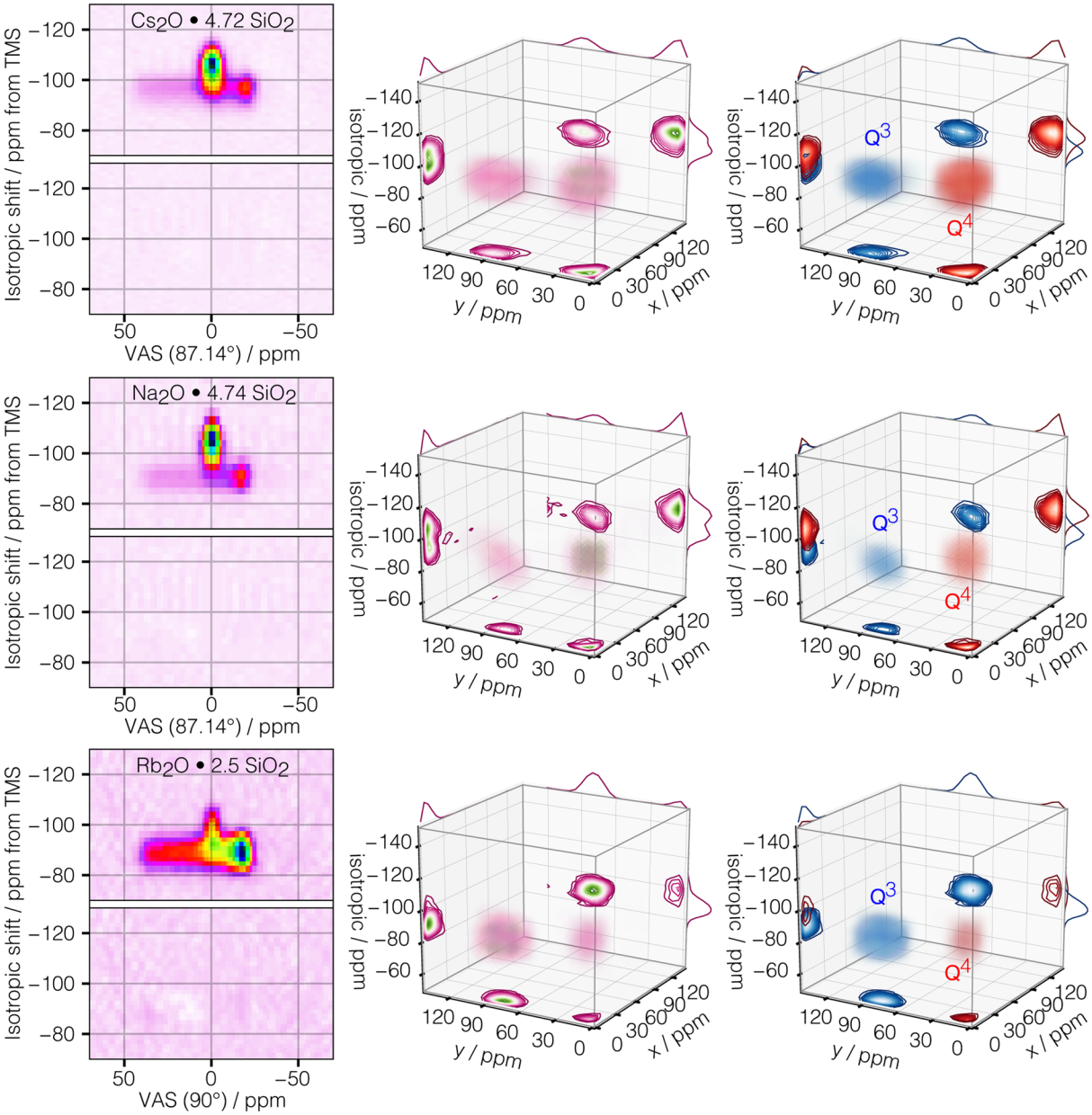

FIG. 7. Smooth-LASSO inversions of the ^{29}Si 2D MAF spectra for the $Cs_2O \cdot 4.72SiO_2$, $Na_2O \cdot 4.74SiO_2$, and $Rb_2O \cdot 2.25SiO_2$ glasses,[74,75] ordered from top to bottom, into the corresponding tri-variate tensor parameter distribution. In the 1st column are the 2D MAF spectra of silicate glass (top) and the corresponding residuals from the best fit (bottom). In the second column are the corresponding inversion solutions for the shielding tensor distributions. The contours of the 2D projections from this 3D distribution onto the axes are shown along the face normal to the axis of projection. Similarly, the 1D projections from the respective 2D projections are shown on top of the respective 2D projection planes. The third column is similar to the second, except the individual distributions for the Q^4 and Q^3 sites are highlighted in red and blue colors, respectively. The contours of the 2D projections are drawn at 5%.

with a mean anisotropy of $\zeta = 0$. The distribution modes at a position shifted further away from the x–y origin are assigned to the Q^3 sites. This assignment is again consistent with their projection onto the isotropic chemical shift dimension, appearing in the −100 ppm to −90 ppm range. The well-separated Q^4 and Q^3 distribution modes can be isolated within rectangular cuboids and are highlighted in red and blue colors, respectively, in the third column of Fig. 7.

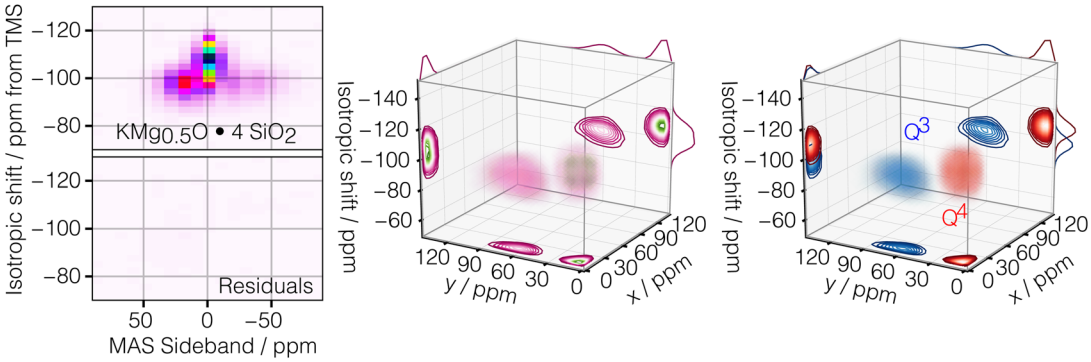

FIG. 8. Smooth-LASSO inversion of the ^{29}Si 2D MAT-PIETA spectra of the $KMg_{0.5}O \cdot 4SiO_2$ glass[73] into the corresponding tri-variate tensor parameter distribution. The spectrum was acquired at a MAS speed of 790 Hz. On the left is the 2D spectrum (top) and the corresponding residuals from the best fit (bottom). In the middle is the corresponding inversion solution for the shielding tensor distribution. The contours of the 2D projections from this 3D distribution onto the axes are shown along the face normal to the axis of projection. Similarly, the 1D projections from the respective 2D projections are shown on top of the respective 2D projection planes. The distribution on the right is similar to the middle, except the Q^4 and Q^3 distributions are highlighted in red and blue colors, respectively. The contours of the 2D projections are drawn at 5%.

Another example, shown in the first column of Fig. 8, is the ^{29}Si 2D MAT-PIETA spectrum for the silicate glass composition $KMg_{0.5}O \cdot SiO_2$, obtained by Walder *et al.*[73] In this 2D spectrum, the anisotropic cross sections are MAS sideband spectra acquired at a spin rate of 790 Hz. This speed produces a sufficient number of sideband amplitudes across the Q^3 and Q^4 regions, ~14 and ~5, respectively, for obtaining the distribution of anisotropy parameters. As before, applying the smooth-LASSO inversion to each 1D anisotropic cross section along the isotropic dimension gives the 3D distribution of shielding tensor parameters shown in the second column of Fig. 8 along with 2D and 1D projections onto the corresponding planes and axes. Similarly, below the 2D spectrum is the best fit residuals obtained with the inversion solution. For this glass composition, the inversion of the 2D spectrum also yields a bimodal distribution, with the well-separated mode near the *x*–*y* origin assigned to the Q^4 sites and mode shifted further away from the *x*–*y* origin assigned to the Q^3 sites. The shapes of the Q^4 and Q^3 distributions are similar to those shown in Fig. 7. The well-separated Q^4 and Q^3 distribution modes, isolated within rectangular cuboids, are highlighted in red and blue colors, respectively, in the third column of Fig. 8.

The results of a moment analysis on the isolated Q^4 and Q^3 distribution modes shown in Figs. 7 and 8 are given in Table II. Also shown in Table II are the results from the previously published

TABLE II. Moment analysis up to the second moment of the isolated Q^4 and Q^3 distribution modes obtained from the smooth-LASSO inversion of the ^{29}Si 2D NMR spectra shown in Figs. 7 and 8. The values for the isotropic chemical shift δ_{iso}, *x*, and *y* are reported as the mean ± standard deviation for the distribution modes. The previously determined isotropic chemical shift values δ_{iso} are also the mean ± standard deviation for the distribution modes. In previous work, only the distribution mean values $\bar{\zeta}$ and $\bar{\eta}$ could be determined.

	Q^4 statistics from inversion (this work)				Q^4 statistics from previous work		
Composition	%	δ_{iso} (ppm)	*x* (ppm)	*y* (ppm)	%	δ_{iso} (ppm)	References
$Cs_2O \cdot 4.72SiO_2$	59.23	−104.5 ± 5.4	9.6 ± 6.4	10.9 ± 6.9	57.7 ± 0.4	−104.7 ± 5.2	75
$Na_2O \cdot 4.74SiO_2$	60.46	−103.7 ± 5.4	8.5 ± 4.8	9.0 ± 4.9	57.8 ± 0.1	−103.7 ± 5.31	74
$Rb_2O \cdot 2.25SiO_2$	11.87	−98.0 ± 5.4	8.2 ± 4.3	8.8 ± 4.4	11.0 ± 0.3	−98.0 ± 5.64	74
$KMg_{0.5}O \cdot 4SiO_2$	55.53	−107.3 ± 5.5	8.7 ± 4.6	9.0 ± 4.8	Not analyzed		73

	Q^3 statistics from inversion (this work)						Q^3 statistics from previous work				
Composition	%	δ_{iso} (ppm)	*x* (ppm)	*y* (ppm)	$\bar{\zeta}$ (ppm)	$\bar{\eta}$	%	δ_{iso} (ppm)	$\bar{\zeta}$ (ppm)	$\bar{\eta}$	References
$Cs_2O \cdot 4.72SiO_2$	40.77	−96.1 ± 3.9	11.0 ± 7.2	87.7 ± 9.9	88.4	0.16	42.3 ± 0.4	−96.1 ± 4.0	89.0	0	75
$Na_2O \cdot 4.74SiO_2$	39.54	−90.6 ± 4.3	10.2 ± 6.1	79.1 ± 7.8	79.8	0.16	42.2 ± 0.2	−90.5 ± 4.29	79.8	0	74
$Rb_2O \cdot 2.25SiO_2$	88.13	−88.9 ± 4.4	10.1 ± 6.5	79.9 ± 8.1	80.5	0.16	89.0 ± 0.1	−89.5 ± 4.65	80.7	0	74
$KMg_{0.5}O \cdot 4SiO_2$	44.47	−97.0 ± 4.8	9.8 ± 4.4	62.5 ± 10.7	63.3	0.20	Not analyzed				73

spectral analysis[74,75] carried out using a forward non-linear least-squares approach where an *ad hoc* Gaussian line shape convolution was applied to the anisotropic cross sections to model the structural disorder around the Q^n sites. In these earlier constrained analyses, the ζ and η values were assumed to be zero for the Q^4 sites, and the mean ζ values were determined for the Q^3 sites with an assumption of $\eta = 0$. As listed in Table II, the Q^3 and Q^4 populations, the mean and standard deviation of the Q^3 and Q^4 isotropic chemical shift distributions are in good agreement with the previously published spectral analyses.[74,75] While there is also good agreement with the previously published mean ζ value for Q^3, the inversion gives the non-zero mean values for η in contrast to the assumed value of $\eta = 0$ in the previous analyses. Additionally, it should be noted that the model-free inverse approach can yield a more accurate Q^n population as the model constraints in the forward non-linear least-squares approach generally lead to larger residuals and more systematic biases in quantifying integrated intensities. Most importantly, the overly restrictive assumptions on the anisotropy parameters for both Q^3 and Q^4 sites in these earlier works are eliminated in the smooth-LASSO inversion analysis. This creates new opportunities for more detailed analyses of the distribution of anisotropy parameters.

In Fig. 9 is a comparison of the bi-variate distributions of isotropic chemical shift, $\delta_{iso}^{(cs)}$, and nuclear shielding anisotropy, ζ. These distributions are obtained by projecting out the asymmetry parameter dependence of the 3D tensor parameter distributions from the smooth-LASSO inversions of the ^{29}Si 2D NMR spectra shown in Figs. 7 and 8. The near-symmetric pair of peaks about $\zeta = 0$ occurring at the more negative isotropic chemical shifts are assigned to the Q^4 sites. The lower silica content in the $Rb_2O \cdot 2.25SiO_2$ glass composition is reflected in the significantly lower Q^4 distribution intensity. While it is often the first coordination sphere geometry that determines the second-rank anisotropy, there are situations where the mean first coordination sphere geometry has a symmetry that leads to no anisotropy. Such is the case with the tetrahedral symmetry around ^{29}Si in Q^4 (and also Q^0) sites. Thus, the origin of this near-symmetric pair of peaks in ^{29}Si shielding anisotropy for Q^4 likely arises from random structural deviations away from this high symmetry in the first-coordination sphere of silicon. Again, this type of ζ distribution is consistent with the Czjzek distribution. As noted earlier, the synthetic VAS (90°) line shape arising from a Czjzek distribution is Gaussian. For this reason, Baltisberger *et al.*[74] modeled the anisotropic ^{29}Si line shape of the Q^4 sites in MAF spectra with a Gaussian line shape. They found that this anisotropic Gaussian line width increases with increasing isotropic chemical shift. Close examination of the Q^4 distributions in Fig. 9 also reveals an increase in the ζ distribution width when moving toward less negative isotropic chemical shifts, $\delta_{iso}^{(cs)}$, in agreement with this earlier observation. Baltisberger *et al.*[74] explained this increasing width by assigning the narrowest anisotropic cross sections at the most negative chemical shifts to $Q^{4,4444}$ sites and the increasingly wider anisotropic cross sections at less negative chemical shits to Q^4 sites in anionic clusters with Q^3, such as $Q^{4,4443}$, $Q^{4,4433}$, $Q^{4,4333}$, and $Q^{4,3333}$. Here, the notation $Q^{4,4444}$ represents a Q^4 unit connected to four other Q^4 units, as in silica glass SiO_2, while $Q^{4,4333}$ represents a Q^4 unit connected to one Q^4 units and three Q^3.

The widths of the ζ distribution for Q^4 in the $Na_2O \cdot 4.74SiO_2$ and $KMg_{0.5}O \cdot 4SiO_2$ glass compositions are narrower than in the

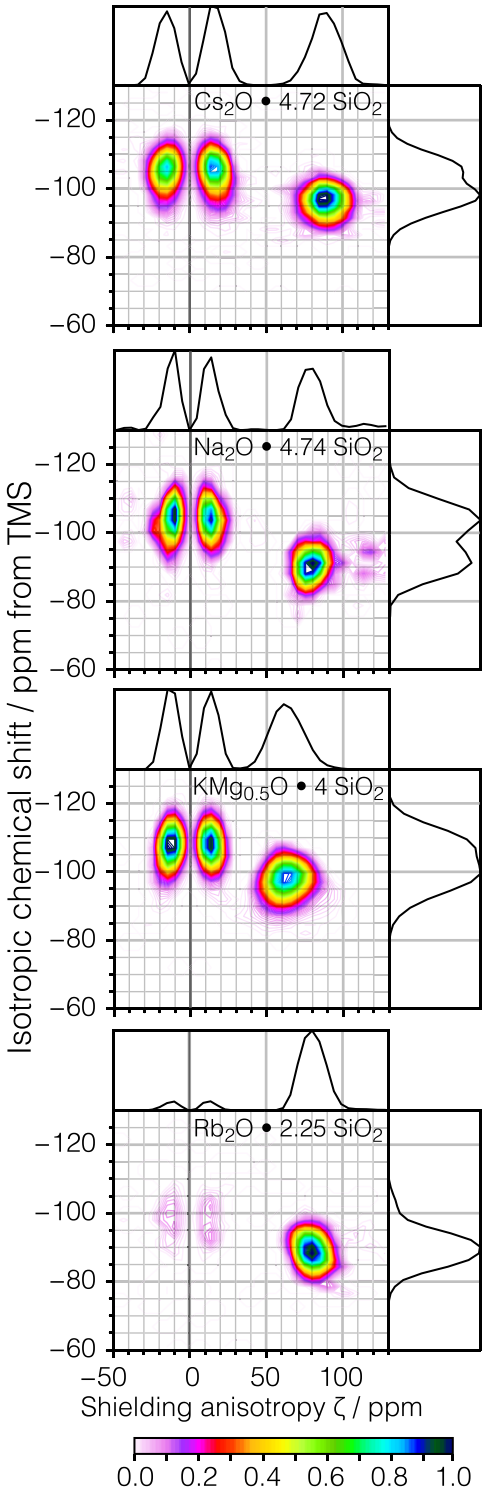

FIG. 9. Comparison of the bi-variate distributions of isotropic chemical shift, $\delta_{iso}^{(cs)}$, and nuclear shielding anisotropy, ζ, obtained by projecting out the η dependence of the 3D distributions from the smooth-LASSO inversions of the ^{29}Si 2D NMR spectra shown in Figs. 7 and 8.

$Cs_2O \cdot 4.72SiO_2$ glass. This can be similarly explained with the commonly accepted view that the Na^+ and Mg^{2+}, with a higher cation potential, i.e., higher charge to radius ratio (Z/r), tend to cluster and lead to more silica-rich, i.e., $Q^{4,4444}$, regions, thus leading to the narrower ζ distributions. Overall, these results represent the first quantification of the distribution of ^{29}Si shielding anisotropy for Q^4 sites in a silicate glass and suggest possible correlations with structural features such as depolymerization of the network and modifier cation potential.

The Q^3 modes in Fig. 9 occur at the lower negative isotropic chemical shifts and at the larger positive-only values of ζ. While the Q^3 distribution shapes also reveal slight correlations between ζ and isotropic chemical shifts—most noticeable in the $Na_2O \cdot 4.74SiO_2$ and $Rb_2O \cdot 2.25SiO_2$ glasses—it is difficult to speculate on the origin of these correlations without further investigations. Projecting out the isotropic chemical shift gives the 1D ζ distributions for the Q^3 sites in the four silicate glass compositions shown together in Fig. 10. Baltisberger *et al.*[74] found that the mean Q^3 shielding anisotropy was linearly dependent on the silicon–non-bridging oxygen (NBO) bond length with decreasing anisotropy as the Si–NBO bond lengthened. Furthermore, they found that the Si–NBO bond length, in turn, is linearly dependent on the cation potential and the number of coordinating modifier cations. That is, both higher cation potentials and high modifier cation coordination lead to longer Si–NBO bond lengths. This linear dependence of the shielding anisotropy on the Si–NBO length is well known and dates back to work by Grimmer and co-workers[94,95] in the early 1980s and similarly confirmed by Kirkpatrick and co-workers[98,99] in phosphate glasses.

Jardón-Álvarez *et al.*[75] proposed an approximate expression for the variations in the mean Q^3 nuclear shielding ζ values given by

$$\bar{\zeta} \approx m(\langle d_{\mathrm{Si-BO}} \rangle - \langle d_{\mathrm{Si-NBO}} \rangle), \tag{40}$$

where $\langle d_{\mathrm{Si-BO}} \rangle$ and $\langle d_{\mathrm{Si-NBO}} \rangle$ are the mean Si–O lengths involving the bridging (BO) and non-bridging oxygen on a Q^3, respectively. A Si–O length typically varies from ~1.50 Å to ~1.60 Å in going from Si–NBO to Si–BO. Assuming $\langle d_{\mathrm{Si-BO}} \rangle$ to be 1.60 Å, they obtained an approximate slope of m = 1500 ppm/Å. Using Eq. (40), we can approximately map the ζ distribution obtained from the smooth-LASSO inversion into the full distribution of Q^3 Si–NBO bond lengths, $d_{\mathrm{Si-NBO}}$, as shown along the top of the plot in Fig. 10. Each distribution peaks at a bond length consistent with known variations with composition.[74] That is, the observed ordering in the Si–NBO bond length peaks, from shortest to longest, is a result of increasing network modifier cation potential and increasing coordination of the NBO by modifier cations. Cesium, with the lowest cation potential, results in the shortest mean Si–NBO bond length, shifting the ζ distribution toward larger values. The broader ζ distribution in the cesium silicate glass is also a result of its low cation potential, which leads to a more random distribution of cesium cations in the glass, i.e., a greater variation in cesium coordination numbers around non-bridging oxygen,[75] and hence a larger distribution of Si–NBO bond lengths. While rubidium has a lower cation potential than sodium, similar mean Si–NBO bond lengths in the rubidium and sodium silicate glasses arise from a higher NBO coordination by rubidium than sodium—a simple consequence of the higher rubidium content in the glass composition. The high cation potential of Mg^{2+} leads to the longer Si–NBO bond lengths in the $KMg_{0.5}O \cdot 4SiO_2$ glass, shifting the ζ distribution toward smaller values. One might be tempted to interpret the broader distribution of ζ as a broader distribution in Si–NBO bond lengths. One could even further suggest that such a broad Si–NBO bond length distribution is a result of a variety of potassium and magnesium modifier coordination environments around the non-bridging oxygen. This explanation, however, is at odds with an earlier ^{17}O DAS study[84] on this same composition, which suggests a high degree of ordering of K and Mg around the non-bridging oxygen. Instead, we believe that the broader distribution of ζ in the $KMg_{0.5}O \cdot 4SiO_2$ glass arises from greater distortions of the SiO_4 tetrahedra, in this case the Q^3 sites, caused by the stronger oxygen coordination requirements around Mg^{2+}. Thus, while Eq. (40) appears to be a reasonable model for interpreting the mean $\bar{\zeta}$ in terms of the difference in mean distances, $\langle d_{\mathrm{Si-BO}} \rangle$ and $\langle d_{\mathrm{Si-NBO}} \rangle$, some care must be taken in using it to interpret the ζ distribution width.

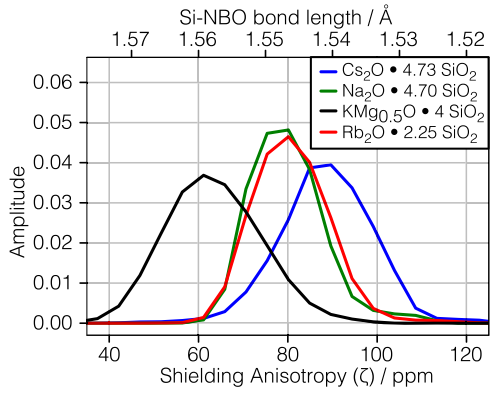

FIG. 10. A comparison of the distributions of the Q^3 anisotropy parameter, ζ, from the smooth-LASSO inversion of the ^{29}Si 2D NMR spectra shown in Figs. 7 and 8. Each Q^3 distribution has been normalized to the unit area.

In Fig. 11 are the final experimental examples: the ^{29}Si 2D MAF spectra of $CaO \cdot SiO_2$ glass, obtained by Zhang *et al.*,[96] and $MgO \cdot SiO_2$ glass, obtained by Davis *et al.*[97] These two spectra have less discernible features than the previously discussed 2D spectra of alkali silicate glasses for reasons mainly attributable to the higher cation potentials of the network modifying cations, Ca^{2+} and Mg^{2+}. While Q^2 is the predominant anionic species present in both glass compositions, the higher cation potentials of the alkaline earth Ca^{2+} and Mg^{2+} cations lead to an increase in the disproportionation reactions

$$2Q^n \rightleftharpoons Q^{n-1} + Q^{n+1} \tag{41}$$

and

$$2Q^0 \rightleftharpoons 2Q^1 + O^{2-} \tag{42}$$

in the melt from which the glass was formed and therefore more disorder in the types of anionic species present in the glass.[96,97,100,101] Thus, compared to the previously discussed alkali silicate glass spectra, we observe a wider range of isotropic ^{29}Si chemical shifts due

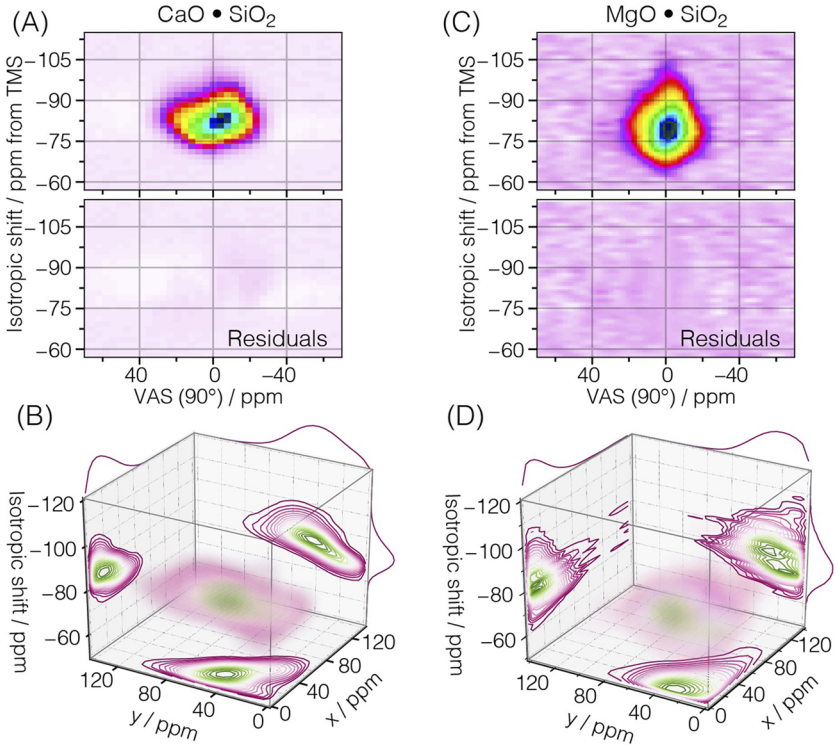

FIG. 11. Smooth-LASSO inversions of the ^{29}Si 2D MAF spectra[96,97] for the (a) $CaO \cdot SiO_2$ and (c) $MgO \cdot SiO_2$ glasses, into the corresponding tri-variate tensor parameter distribution. In (a) and (c) are the 2D MAF spectra of silicate glass (top) and the corresponding residuals from the best fit (bottom). In (b) and (d) are the corresponding inversion solutions for the shielding tensor distributions. The contours of the 2D projections from this 3D distribution onto the axes are shown along the face normal to the axis of projection. Similarly, the 1D projections from the respective 2D projections are shown on top of the respective 2D projection planes. The contours of the 2D projections are drawn at 5%.

to the presence of all five possible Q^n environments. This anionic species disorder is even greater for Mg^{2+} than Ca^{2+} given its relatively higher cation potential, and this is also reflected in the corresponding larger isotropic line width of the $MgO \cdot SiO_2$ spectrum in Fig. 11. As the high cation potentials of Ca^{2+} and Mg^{2+} lead to Si–NBO bonds approaching Si–BO bond lengths, there is a systematic drop in the magnitude of the mean ^{29}Si shielding anisotropy, ζ, for the Q^1, Q^2, and Q^3 sites.[96,97] The effect is less for Ca^{2+} than Mg^{2+} given its relatively lower cation potential. For this reason, it is somewhat easier to discern the variations in the slightly larger Q^n anisotropic line shapes in the $CaO \cdot SiO_2$ spectrum in Fig. 11 with changing isotropic chemical shift. Finally, the stronger oxygen coordination requirement around Ca^{2+} and, even more so for Mg^{2+}, is expected to lead to greater intra-tetrahedral distortions of the SiO_4, and overall broader distributions of shielding tensor parameters for all Q^n sites.

Applying the smooth-LASSO inversion to each 1D anisotropic cross section along the isotropic dimension gives the 3D distribution of shielding tensor parameters shown in Figs. 11(b) and 11(d) along with 2D and 1D projections onto the corresponding planes and axes. Similarly, below the 2D spectrum is the best fit residuals obtained with the inversion solution. While we know the underlying Q^n tensor parameter distributions for these spectra are multi-modal, with modes associated with each of the five possible Q^n environments, we expect the various modes to separate along the isotropic dimension such that the inversion of a given 1D anisotropic cross section yields primarily a unimodal or bimodal distribution. To a certain degree, this is observed in the smooth-LASSO inversion of both the $CaO \cdot SiO_2$ and $MgO \cdot SiO_2$ spectra; however, as one can see in Figs. 11(b) and 11(d), the individual distribution modes are strongly overlapping—although the situation is slightly better in the $CaO \cdot SiO_2$ case due to its larger and more discernible anisotropic line shapes. Unfortunately, given the extent of overlapping distributions, any statistical analysis of individual Q^n distribution modes in these two spectra would require additional modeling and is beyond the scope of this work.

On the left in Figs. 12 and 13 are the bi-variate distributions of isotropic chemical shift, $\delta_{iso}^{(cs)}$, and nuclear shielding anisotropy, ζ, obtained by projecting out the η dependence of the 3D distributions from the smooth-LASSO inversions of the ^{29}Si 2D NMR spectra shown in Fig. 11, respectively. On the right in Figs. 12 and 13 are five x–y cross sections from the smooth-LASSO inversion, labeled (a)–(e), taken at isotropic chemical shifts corresponding to the approximately most probable Q^4, Q^3, Q^2, Q^1, and Q^0 regions, respectively. Overlaid on each of the five x–y cross sections from the smooth-LASSO inversion is a small white circular marker whose position corresponds to the previously reported mean ζ and η values. These were obtained using a forward non-linear least-squares approach with an *ad hoc* Gaussian line shape convolution applied to the anisotropic cross sections to model the structural disorder around the Q^n sites. Again, in these earlier analyses, the ζ and η values were assumed to be zero for the Q^4 and Q^0 sites along with sample-specific constraints for the mean ζ and η values of the other Q^n sites.[96,97]

The stronger mode intensity of all Q^n inversion cross sections relative to the Q^2 cross section in the $MgO \cdot SiO_2$ vs the $CaO \cdot SiO_2$

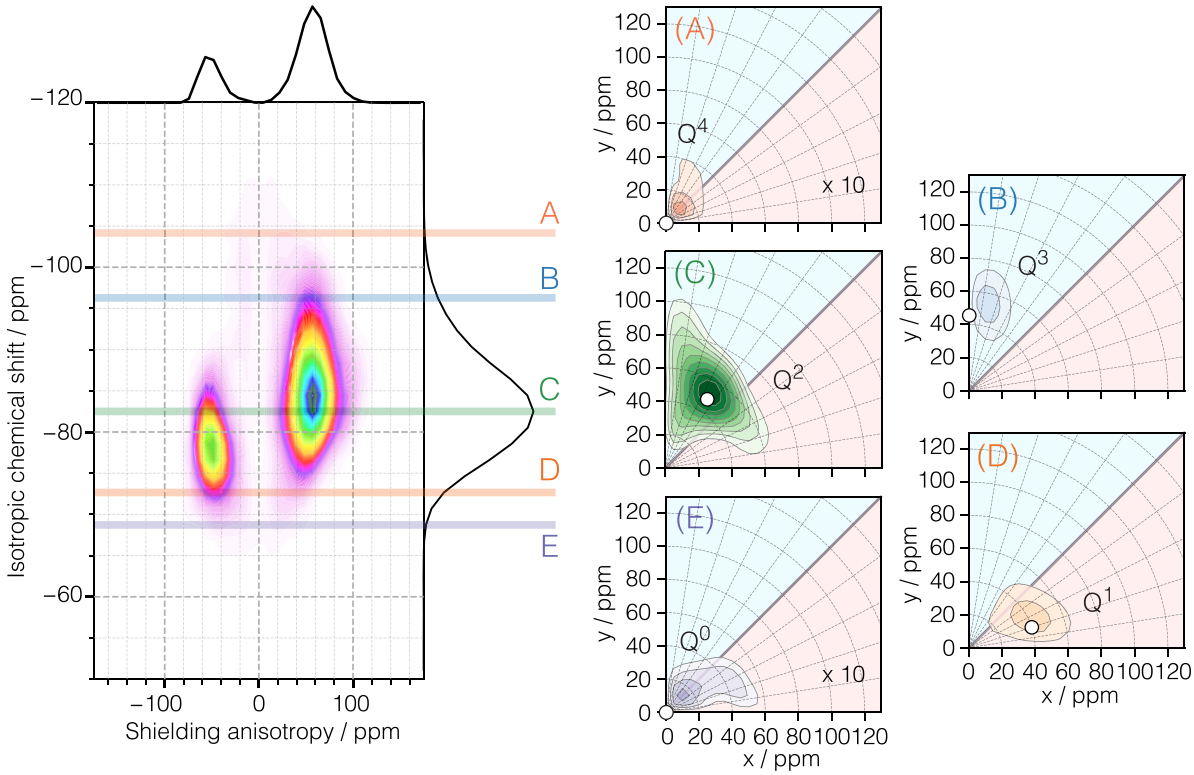

FIG. 12. On the left is the bi-variate distribution of isotropic chemical shift, $\delta_{iso}^{(cs)}$, and nuclear shielding anisotropy, ζ, in $CaO \cdot SiO_2$ glass obtained by projecting out the η dependence of the 3D distributions from the smooth-LASSO inversions of the ^{29}Si NMR spectra shown in Fig. 11. On the right are five smooth-LASSO inversion x–y cross sections, labeled (a)–(e), taken at isotropic chemical shifts corresponding to the approximately most probable Q^4, Q^3, Q^2, Q^1, and Q^0 regions, respectively. The white circular marking over the distributions is the reported values for the corresponding label site. The contours are drawn at every 10%.

case is consistent with the expected increased disproportionation of anionic species. Furthermore, the widths of the distribution in the x–y cross sections are greater for Q^3, Q^2, and Q^1 in $MgO \cdot SiO_2$ compared to $CaO \cdot SiO_2$. Again, this is as expected based on the increased intra-tetrahedral distortions induced by the higher cation potential of Mg^{2+}.

Unrelated to the glass structural chemistry, a possible factor contributing to the broad overlapping Q^n tensor parameter distributions in these particular samples, unlike those in Figs. 7 and 8, is that both samples are made from SiO_2 that was highly enriched (>95%) in ^{29}Si. The homonuclear ^{29}Si dipolar couplings in these samples, particularly at low external magnetic field strengths, lead to an increased spectral broadening of the anisotropic cross sections. It is far from trivial to expand the kernel subspectral basis to include the variation due to these homonuclear ^{29}Si dipolar couplings. Thus, the inversion of these spectra, using a kernel intended only for modeling the spectra of dilute spin 1/2 nuclei, i.e., not accounting for broadenings from dipolar couplings or finite transition lifetimes, will be biased toward broader solutions.

We close this section by noting that the inversion kernel of subspectra assumes an ideal response free from artifacts, i.e., an undistorted pure absorption-mode spectrum with no convolution from signal truncations due to delayed acquisition or insufficient acquisition lengths. For example, while many 2D isotropic/anisotropic correlation methods can benefit from the higher sensitivity obtained with echo train acquisition methods, signal artifacts are introduced if there is signal truncation due to the finite acquisition time between π pulses. Additional signal artifacts may be introduced if the π pulse spacing is not synchronous with the sample rotation rate. In some cases, the kernel of subspectra can incorporate certain systematic distortions, e.g., receiver dead times or echo train finite acquisition time windows, and reduce some broadening of the inversion solution due to these artifacts. Each 2D isotropic/anisotropic correlation method may also have its own specific artifacts that need to be suppressed or, if possible, incorporated into the inversion kernel. For MAS sideband spectra, the experimental spinning speed must be accurately known and stable throughout the measurement. In the magic-angle flipping technique, a high enough rotor speed is needed to make sure the spectrum is sideband free. Additionally, if there is a difference in the external magnetic field strength, then the pure anisotropic dimension obtained after shearing the 2D MAF spectrum may still include a small isotropic component. This issue can be corrected by recording the difference in the field strength at the two angles and correcting the sheared MAF dataset using this value.

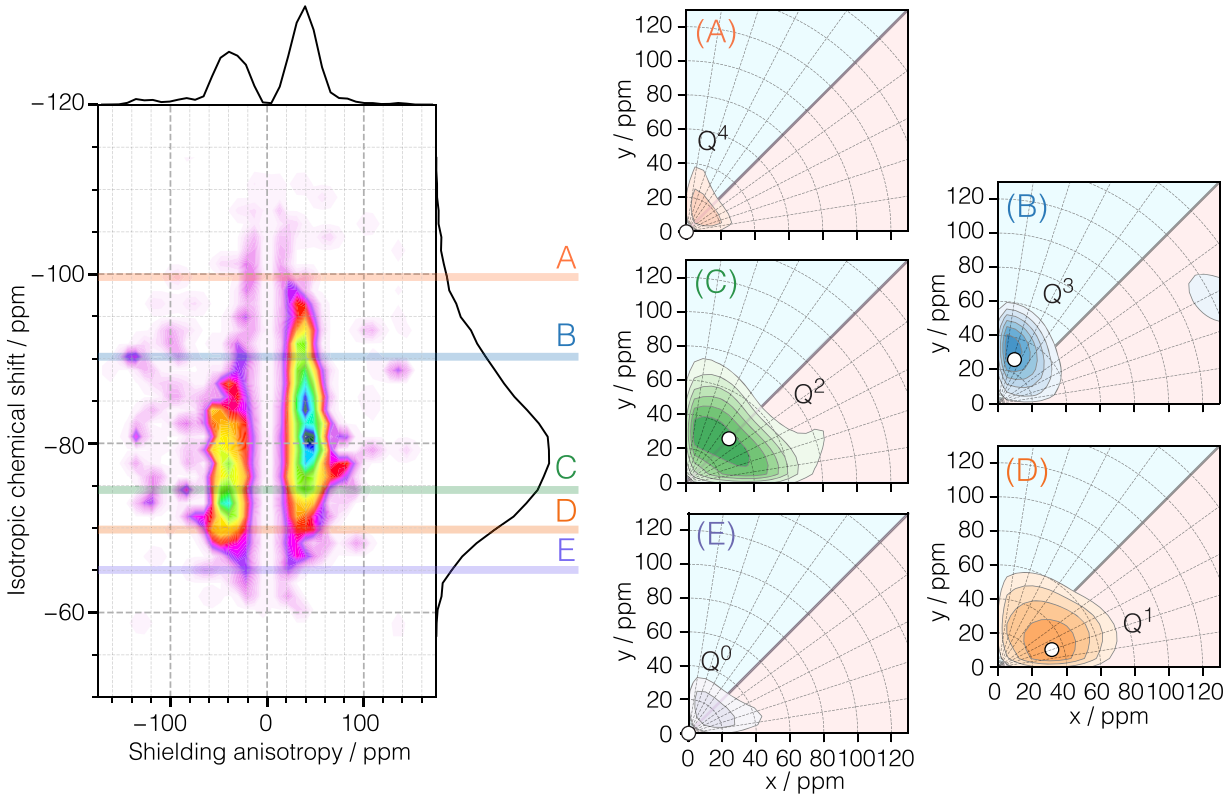

FIG. 13. On the left is the bi-variate distribution of isotropic chemical shift, $\delta_{iso}^{(cs)}$, and nuclear shielding anisotropy, ζ, in $MgO \cdot SiO_2$ glass obtained by projecting out the η dependence of the 3D distributions from the smooth-LASSO inversions of the ^{29}Si NMR spectra shown in Fig. 11. On the right are five smooth-LASSO inversion x–y cross sections, labeled (a)-(e), taken at isotropic chemical shifts corresponding to the approximately most probable Q^4, Q^3, Q^2, Q^1, and Q^0 regions, respectively. The white circular marking over the distributions is the reported values for the corresponding label site. The contours are drawn at every 10%.

IV. SUMMARY

In this work, we have shown that the linear inversion of a pure anisotropic spectrum into a two-dimensional distribution of the second-rank nuclear shielding tensor anisotropy parameters can be successfully performed using the Smooth Least Absolute Shrinkage and Selection Operator[50] (S-LASSO) estimator, a variant of the Elastic-Net estimator,[49] as a penalty function along with 10-fold cross-validation.

We also propose a piecewise polar coordinate system—where the magnitude of NMR tensor anisotropy parameter, ζ, forms the radial dimension and the asymmetry parameter, η, forms the angular dimension—to minimize the rank deficiency of the kernel and obtain a more robust and unambiguous inversion. This polar coordinate system also facilitates the classification of structural motifs present in amorphous materials, e.g., the Q^n species present in silicate glasses, through their characteristic NMR tensor parameter distributions.

We find the smooth-LASSO performance to be superior to other commonly used regularization approaches and able to regularize the inverse solutions of spectra from both amorphous and crystalline materials. It performs reliably with a variety of synthetic purely anisotropic 1D spectra [VAS (90°) and MAS sidebands] generated from known (ground truth) bivariate uni-modal distributions of nuclear shielding anisotropy parameters having different shapes and widths. In the supplementary material, we also show similarly good performance for the smooth-LASSO inversion of various synthetic purely anisotropic 1D spectra arising from more challenging bivariate bi- and tri-modal distributions of nuclear shielding anisotropy parameters.

When applied to experimental 2D solid-state NMR spectra correlating isotropic and anisotropic frequencies, the smooth-LASSO inversion provides a general approach for obtaining the trivariate distribution of isotropic shifts and the nuclear interaction tensor anisotropy and asymmetry parameters. We present the results from applying the smooth-LASSO inversion to the previously published experimental 2D ^{29}Si NMR magic-angle flipping or magic-angle turning spectra of six different alkali and alkaline earth silicate glass compositions. We find that these results not only agree with the previously performed non-linear forward least-squares spectral analyses but further provide an unparalleled level of quantitative details about the distribution of ^{29}Si nuclear shielding tensor parameters

in these glasses. Here, we show how such results allow us to probe the intra-tetrahedral distortions of Q^n and the distance distributions for silicon—non-bridging oxygen bonds on Q^3 sites. More generally, the smooth-LASSO-determined trivariate distributions of nuclear shielding tensor parameters in a variety of glass compositions can provide a wealth of new data for both supervised and unsupervised machine learning efforts to classify structural motifs present in both crystalline and amorphous solids.

We emphasize that the primary advantage of the smooth-LASSO inversion over forward least-squares analyses is that it is model free. While the forward least-squares analysis has an advantage of mathematical simplicity, a main drawback is that it requires an explicit model for the probability distribution of NMR tensor parameters. Depending on the form of this explicit model, a forward least-squares analysis can lead to solutions biased toward narrower or broader parameter distributions. In contrast, the smooth-LASSO method biases solutions toward broader distributions of NMR tensor parameters. We show that such broadening of the inversion solutions of NMR spectra can occur from increased noise, increased MAS speeds, and spectral broadenings from spin relaxation or homogeneous dipolar couplings.

The smooth-LASSO inversion presented in this study is general and can be extended to anisotropic spectra arising from other subspectral bases, such as the central transition of second-order broadened quadrupolar nuclei. It can, for example, be readily applied to 2D DAS spectra. In the case of 2D MQ-MAS, where there is a non-uniform excitation and detection of spectral frequencies, knowledge of the excitation and mixing radio frequency field strengths will be needed to generate an accurate kernel of subspectra. Of course, one-dimensional Bloch decay spectra of solids are not purely anisotropic—neither static nor MAS spectra. In principle, one could extend the smooth-LASSO inversion to such spectra with a suitable expansion of the subspectral basis to include the variation due to isotropic shifts as well as other NMR interaction parameters. The increasing size of the kernel with the increasing dimensionality of the multi-variate distribution of NMR parameters, however, could lead to greater uncertainties in the inversion solution as well as make the approach more computationally intense.

SUPPLEMENTARY MATERIAL

See the supplementary material for an outline of the smooth-LASSSO algorithm, additional examples of linear inversion, and tables of tensor distribution moment analyses and optimum hyperparameters for the inversion of all synthetic datasets.

ACKNOWLEDGMENTS

This material is based on the work supported, in part, by the U.S. National Science Foundation under Grant Nos. DIBBS OAC 1640899 and CHE 1807922. D.J.S. would like to acknowledge Dr. Jay Baltisberger for sharing the glass MAF datasets and providing step-wise instructions on MAF data processing. D.J.S. would also like to acknowledge Dr. Brennan Walder for sharing the MAT dataset. P.J.G. thanks Dr. Philip Schniter for helpful discussions.

APPENDIX: DETERMINING HYPERPARAMETERS

1. SVD truncation parameter, *r*

The hyperparameter, r, can be determined from the noise standard deviation[57–59] or from more sophisticated methods such as Akaike Information Criterion[6,104] and entropy theory.[48,105] Here, we implement the maximum entropy[48] based method. Let the dataset entropy $(0 \le E \le 1)$ be

$$E = -\frac{1}{\log(l)} \sum_{j=1}^{l} \mathcal{S}_j \log(\mathcal{S}_j), \tag{A1}$$

where $l = \min(m, n)$ and $\mathcal{S}_j$ is the jth normalized eigenvalue of the matrix $\mathbf{K}\mathbf{K}^T$ given by

$$\mathcal{S}_j = \frac{\varsigma_j^2}{\sum_{k=1}^{l} \varsigma_k^2},$$

where ς_j is the jth singular value. The contribution of the ith feature to the total entropy E is defined by leave one out comparison, which follows

$$\nabla E_i = E - E_i = E - \frac{E\log(l) + \mathcal{S}_i \log(\mathcal{S}_i)}{\log(l-1)}, \tag{A2}$$

where E_i is the entropy with the ith contribution removed. Let c and d be the mean and the standard deviation of ∇E_i's, then the optimum truncation index is given as

$$r = \underset{i}{\operatorname{argmin}}(\nabla E_i - c + d). \tag{A3}$$

2. Hyperparameters, λ and α

When the number of parameters in the model space, $\mathbf{f} \in \mathbb{R}^n$, exceeds the number of points in the data space, or more specifically, the reduced data space, r, classic overfitting of the OLS problem is a real issue. Given the ill-conditioned nature of the problem, the condition $n \gg r$ is frequently encountered. In such cases, the regularization parameters $\alpha, \lambda > 0$, in Eq. (32), can be interpreted as the tuning parameters. Setting the tuning parameters too small leads to the classic overfitting of the OLS problem with the resulting model also fitting the noise. If the tuning parameters are set too large, the model may never adequately describe the data. In practice, one wants to avoid both extremes, and therefore, the challenge is finding the values of λ and α that is just right.

Cross-validation[106] is a widely used method in determining the optimum hyperparameters. Cross-validation is a statistical learning method for evaluating the generalized performance and stability of the model. Our approach uses a stratified k-fold cross-validation method: the signal $\mathbf{s}$ is divided into k subsets of roughly equal size, $m_i \approx m/k$, called *folds*, where k is an integer number, typically chosen as 5 or 10. In addition, a range of hyperparameter values, in this case, $\lambda \in \mathbb{R}_+^{n_\lambda}$ and $\alpha \in \mathbb{R}_+^{n_\alpha}$, are chosen on a uniform log scale, where the subscript + denotes positive numbers and n_λ and n_α are the total number of λ and α values. Out of k folds, one fold is set aside and is called the *test set*. The remaining $(k-1)$-folds are collectively called the *training set*. In the following notation, we represent the ith test set with $\mathbf{s}^i$ and the corresponding training set with $\mathbf{s}^{-i}$. Similarly, the

corresponding kernels are designated with $\mathbf{K}^i$ and $\mathbf{K}^{-i}$, respectively. For every training set, $\mathbf{s}^{-i}$, the model $\mathbf{f}_i(\lambda, \alpha)$ is evaluated, following

$$\mathbf{f}_i(\lambda,\alpha) = \underset{\mathbf{f}\geq 0}{\operatorname{argmin}}\left(\frac{1}{m-m_i}\|\mathbf{K}^{-i}\cdot\mathbf{f}-\mathbf{s}^{-i}\|_2^2+\lambda\|\mathbf{f}\|_1\right) \tag{A4}$$

for the range of λ and α values. Next, the mean square error (MSE), called the *test error* of the test set, is evaluated using

$$\epsilon_i(\lambda,\alpha) = \frac{1}{m_i}\sum\left(\mathbf{K}^i\cdot\mathbf{f}_i(\lambda,\alpha)-\mathbf{s}^i\right)^2, \tag{A5}$$

where the summation runs over the elements of the test set, $\mathbf{s}^i$. This process is repeated k times while assigning a different fold as the *test set* each time. The cross-validation error is then given as

$$CV(\lambda,\alpha) = \frac{1}{m}\sum_{i=1}^{k} m_i\,\epsilon_i(\lambda,\alpha). \tag{A6}$$

The optimal λ^* and α^* hyperparameters are determined as the argument that minimizes the cross-validation error, given as

$$\lambda^*,\alpha^* = \underset{\lambda,\alpha}{\operatorname{argmin}}\left(\left|CV(\lambda,\alpha)-\sigma_e^2\right|\right), \tag{A7}$$

where σ_e is the standard deviation of the noise.

DATA AVAILABILITY

The open-source python package, *mrinversion*, for implementing the smooth-LASSO inversion described here along with documentation for its installation and use is made available in Zenodo at http://doi.org/10.5281/zenodo.3964643 (Ref. 54). The documentation for *mrinversion* also includes example scripts for obtaining inversions of all the experimental datasets presented in this work, as well as numerous synthetic datasets. The experimental CSDM[102] compliant datasets and their inversion solutions that support the findings of this study are openly available in Zenodo at http://doi.org/10.5281/zenodo.3964530 (Ref. 103).

REFERENCES

[1] M. Mehring, *High Resolution NMR Spectroscopy in Solids* (Springer-Verlag, Berlin, 1983), Vol. 11.

[2] G. Coates, L. Xiao, and M. Prammer, *NMR Logging: Principles and Applications* (Haliburton Energy Services, 1999).

[3] P. T. Callaghan, *Translational Dynamics and Magnetic Resonance: Principles of Pulsed Gradient Spin Echo NMR* (Oxford University Press, 2011).

[4] R. Bai, C. G. Koay, E. Hutchinson, and P. J. Basser, "A framework for accurate determination of the T_2 distribution from multiple echo magnitude MRI images," J. Magn. Reson. **244**, 53–63 (2014).

[5] Y.-Q. Song, L. Venkataramanan, M. D. Hürlimann, M. Flaum, P. Frulla, and C. Straley, "T_1–T_2 correlation spectra obtained using a fast two-dimensional Laplace inversion," J. Magn. Reson. **154**, 261–268 (2002).

[6] X. Ge, Y. Fan, H. Chen, S. Deng, Y. Cao, and M. A. Zahid, "Probing the influential factors of NMR T_1–T_2 spectra in the characterization of the kerogen by numerical simulation," J. Magn. Reson. **260**, 54–66 (2015).

[7] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes* (Cambridge University Press, New York, NY, 1992).

[8] P. Hansen, *Discrete Inverse Problems: Insight and Algorithms* (Society for Industrial and Applied Mathematics, Philadelphia, 2010).

[9] G. E. Peterson, C. R. Kurkjian, and A. Carnevale, "Informational content of boron NMR in amorphous materials," J. Non-Cryst. Solids **21**, 283–285 (1976).

[10] G. E. Peterson, A. Carnevale, and C. R. Kurkjian, "Singular-value decomposition and boron NMR spectra in glass," J. Non-Cryst. Solids **23**, 243–259 (1977).

[11] G. E. Peterson, A. Carnevale, and C. R. Kurkjian, "A fast and stable technique for the analysis of B^{11} NMR spectra in glass," Phys. Chem. Glasses **18**, 41–48 (1977).

[12] M. Bloom, J. H. Davis, and A. L. Mackay, "Direct determination of the oriented sample NMR spectrum from the powder spectrum for systems with local axial symmetry," Chem. Phys. Lett. **80**, 198–202 (1981).

[13] K. P. Whittall, E. Sternin, M. Bloom, and A. L. Mackay, "Time- and frequency-domain "dePakeing" using inverse theory," J. Magn. Reson. **84**, 64–71 (1989).

[14] H. Schäfer, B. Mädler, and F. Volke, "De-Pake-ing of NMR powder spectra by nonnegative least-squares analysis with Tikhonov regularization," J. Magn. Reson., Ser. A **116**, 145–149 (1995).

[15] H. Schafer and R. Stannarius, "Computation of orientational distributions of partially ordered samples from NMR spectra," J. Magn. Reson., Ser. B **106**, 14–23 (1995).

[16] H. Schäfer, B. Mädler, and E. Sternin, "Determination of orientational order parameters from ^{2}H NMR spectra of magnetically partially oriented lipid bilayers," Biophys. J. **74**, 1007–1014 (1998).

[17] F. G. Vogt, D. J. Aurentz, and K. T. Mueller, "Determination of internuclear distances from solid-state nuclear magnetic resonance: Dipolar transforms and regularization methods," Mol. Phys. **95**, 907–919 (1998).

[18] D. Grabowski and J. Honerkamp, "The determination of the reorientational angle distribution in two-dimensional exchange nuclear magnetic resonance spectroscopy on powder samples," J. Chem. Phys. **96**, 2629–2632 (1992).

[19] U. Haeberlen, *High Resolution NMR in Solids: Selective Averaging*, Advances in Magnetic Resonance (Academic Press, New York, 1976).

[20] R. K. Harris, E. D. Becker, S. M. C. De Menezes, P. Granger, R. E. Hoffman, and K. W. Zilm, "Further conventions for NMR shielding and chemical shifts, IUPAC recommendations 2008," Solid State Nucl. Magn. Reson. **33**, 41–56 (2008).

[21] M. Utz, "Measurement of structural distribution functions in disordered systems: A general approach for sensitivity estimation," J. Chem. Phys. **109**, 6110–6124 (1998).

[22] D. Sakellariou and T. Charpentier, "Shift anisotropy tensors in amorphous natural-abundance solids: High-resolution ^{29}Si chemical shift anisotropy distributions under very slow sample rotation," Appl. Magn. Reson. **32**, 583–594 (2007).

[23] B. Blümich, P. Blümler, and J. Jansen, "Presentation of sideband envelopes by two-dimensional one-pulse (TOP) spectroscopy," Solid State Nucl. Magn. Reson. **1**, 111–113 (1992).

[24] M. C. Davis, K. M. Shookman, J. D. Sillaman, and P. J. Grandinetti, "TOP-PASS: A processing algorithm to reduce 2D PASS acquisition time," J. Magn. Reson. **210**, 51–58 (2011).

[25] W. T. Dixon, "Spinning-sideband-free NMR-spectra," J. Magn. Reson. **44**, 220–223 (1981).

[26] W. T. Dixon, "Spinning-sideband-free and spinning-sideband-only NMR spectra in spinning samples," J. Chem. Phys. **77**, 1800 (1982).

[27] O. N. Antzutkin, S. C. Shekar, and M. H. Levitt, "Two-dimensional sideband separation in magic-angle spinning NMR," J. Magn. Reson., Ser. A **115**, 7–19 (1995).

[28] J. Z. Hu, D. W. Alderman, C. H. Ye, R. J. Pugmire, and D. M. Grant, "An isotropic chemical shift-chemical shift anisotropy magic-angle slow-spinning 2D NMR experiment," J. Magn. Reson., Ser. A **105**(1), 82–87 (1993).

[29] S. L. Gann, J. H. Baltisberger, and A. Pines, "Dynamic-angle spinning without sidebands," Chem. Phys. Lett. **210**, 405–410 (1993).

[30] T. Gullion, "Extended chemical-shift modulation," J. Magn. Reson. **85**, 614–619 (1989).

[31] I. Hung and Z. Gan, "An efficient amplification pulse sequence for measuring chemical shift anisotropy under fast magic-angle spinning," J. Magn. Reson. **213**, 196–199 (2011).

[32]N. M. Szeverenyi, A. Bax, and G. E. Maciel, "Magic-angle hopping as an alternative to magic-angle spinning for solid-state NMR," J. Magn. Reson. **61**, 440 (1985).
[33]A. Bax, N. M. Szeverenyi, and G. E. Maciel, "Chemical shift anisotropy in powdered solids studied by 2D FT NMR with flipping of the spinning axis," J. Magn. Reson. **55**, 494 (1983).
[34]J. C. C. Chan and R. Tycko, "Recoupling of chemical shift anisotropies in solid-state NMR under high- speed magic-angle spinning and in uniformly ^{13}C-labeled systems," J. Chem. Phys. **118**, 8378–8389 (2003).
[35]L. Frydman, G. C. Chingas, Y. K. Lee, P. J. Grandinetti, M. A. Eastman, G. A. Barrall, and A. Pines, "Variable-angle correlation spectroscopy in solid-state nuclear magnetic resonance," J. Chem. Phys. **97**(7), 4800–4808 (1992).
[36]B. F. Chmelka, K. T. Mueller, A. Pines, J. F. Stebbins, Y. Wu, and J. W. Zwanziger, "Oxygen-17 NMR in solids by dynamic-angle spinning and double rotation," Nature **339**, 42–43 (1989).
[37]K. T. Mueller, B. Q. Sun, G. C. Chingas, J. W. Zwanziger, T. Terao, and A. Pines, "Dynamic-angle spinning of quadrupolar nuclei," J. Magn. Reson. **86**, 470 (1990).
[38]P. J. Grandinetti, J. H. Baltisberger, A. Llor, Y. K. Lee, U. Werner, M. A. Eastman, and A. Pines, "Pure absorption-mode lineshapes and sensitivity in two-dimensional dynamic angle spinning NMR," J. Magn. Reson., Ser. A **103**, 72–81 (1993).
[39]L. Frydman and J. S. Harwood, "Isotropic spectra of half-integer quadrupolar spins from bidimensional magic-angle spinning NMR," J. Am. Chem. Soc. **117**, 5367–5369 (1995).
[40]A. Medek, J. S. Harwood, and L. Frydman, "Multiple-quantum magic-angle spinning NMR: A new method for the study of quadrupolar nuclei in solids," J. Am. Chem. Soc. **117**, 12779–12787 (1995).
[41]Z. H. Gan, "Satellite transition magic-angle spinning nuclear magnetic resonance spectroscopy of half integer quadrupolar nuclei," J. Chem. Phys. **114**, 10845–10853 (2001).
[42]H.-T. Kwak and Z. H. Gan, "Double-quantum filtered STMAS," J. Magn. Reson. **164**(2), 369–372 (2003).
[43]L. Frydman, G. C. Chingas, Y. K. Lee, P. J. Grandinetti, M. A. Eastman, G. A. Barrall, and A. Pines, "Correlation of isotropic and anisotropic chemical-shifts in solids by 2-dimensional variable-angle-spinning NMR," Isr. J. Chem. **32**(2-3), 161–164 (1992).
[44]P. J. Grandinetti, J. T. Ash, and N. M. Trease, "Symmetry pathways in solid-state NMR," Prog. Nucl. Magn. Reson. Spectrosc. **59**, 121–196 (2011).
[45]J. W. Zwanziger, "Interpreting nuclear magnetic resonance spectra of disordered materials: Direct inversion of powder patterns," Solid State Nucl. Magn. Reson. **3**, 219–229 (1994).
[46]F. Angeli, T. Charpentier, P. Faucon, and J.-C. Petit, "Structural characterization of glass from the inversion of ^{23}Na and ^{27}Al 3Q-MAS NMR spectra," J. Phys. Chem. B **103**, 10356–10364 (1999).
[47]P. C. Hansen, *Translational Dynamics and Magnetic Resonance: Principles of Pulsed Gradient Spin Echo NMR* (Society for Industrial and Applied Mathematics, Philadelphia, 1998).
[48]R. Varshavsky, A. Gottlieb, M. Linial, and D. Horn, "Novel unsupervised feature filtering of biological data," Bioinformatics **22**, e507–e513 (2006).
[49]H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," J. R. Stat. Soc.: Ser. B **67**, 301–320 (2005).
[50]M. Hebiri and S. A. V. D. Geer, "The smooth-lasso and other $\ell_1 + \ell_2$-penalized methods," Electron. J. Statist. **5**, 1184–1226 (2011).
[51]F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in python," J. Mach. Learn. Res. **12**, 2825–2830 (2011).
[52]L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux, "API design for machine learning software: Experiences from the scikit-learn project," in ECML PKDD Workshop: Languages for Data Mining and Machine Learning, 2013; arXiv:1309.0238.
[53]J. Herzfeld and A. E. Berger, "Sideband intensities in NMR spectra of samples spinning at the magic angle," J. Chem. Phys. **73**, 6021 (1980).
[54]D. Srivastava and P. J. Grandinetti (2020). "mrinversion: v0.1.0: A python package for statistical learning of NMR tensors from 2D isotropic/anisotropic correlation nuclear magnetic resonance spectra," Zenodo. https://doi.org/10.5281/zenodo.3964644
[55]M. E. Kilmer and D. P. O'Leary., "Choosing regularization parameters in iterative methods for ill-posed problems," SIAM J. Matrix Anal. Appl. **22**, 1204–1221 (2001).
[56]A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill-Posed Problems* (Wiley, New York, 1977).
[57]Y.-Q. Song, L. Venkataramanan, and L. Burcaw, "Determining the resolution of Laplace inversion spectrum," J. Chem. Phys. **122**, 104104 (2005).
[58]K.-J. Dunn and G. A. LaTorraca, "The inversion of NMR log data sets with different measurement errors," J. Magn. Reson. **140**, 153–161 (1999).
[59]W. Wang, P. Li, and C. Ye, "Multi-exponential inversions of nuclear magnetic resonance relaxation signal," Sci. China, Ser. A: Math. **44**, 1477–1484 (2001).
[60]P. C. Hansen, "The truncated SVD as a method for regularization," BIT Numer. Math. **27**, 534–553 (1987).
[61]L. Venkataramanan, Y. Q. Song, and M. D. Hurlimann, "Solving Fredholm integrals of the first kind with tensor product structure in 2 and 2.5 dimensions," IEEE Trans. Signal Process. **50**, 1017–1026 (2002).
[62]I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," Commun. Pure Appl. Math. **57**, 1413–1457 (2004).
[63]A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," SIAM J. Imaging Sci. **2**, 183–202 (2009).
[64]X. Zhou, G. Su, L. Wang, S. Nie, and X. Ge, "The inversion of 2D NMR relaxometry data using L1 regularization," J. Magn. Reson. **275**, 46–54 (2017).
[65]P. D. Teal and C. Eccles, "Adaptive truncation of matrix decompositions and efficient estimation of NMR relaxation distributions," Inverse Probl. **31**, 045010 (2015).
[66]M. Urbańczyk, D. Bernin, W. Koźmiński, and K. Kazimierczuk, "Iterative thresholding algorithm for multiexponential decay applied to PGSE NMR data," Anal. Chem. **85**, 1828–1833 (2013).
[67]M. Urbańczyk, M. Nowakowski, W. Koźmiński, and K. Kazimierczuk, "Joint non-uniform sampling of all incremented time delays for quicker acquisition in protein relaxation studies," J. Biomol. NMR **68**, 155–161 (2017).
[68]M. Urbańczyk and K. Kazimierczuk, "A method for joint sparse sampling of time and gradient domains in diffusion-ordered NMR spectroscopy," in *2013 Signal Processing Symposium (SPS)* (IEEE, Serock, Poland, 2013).
[69]M. Guerquin-Kern, J. Baritaux, and M. Unser, "Efficient image reconstruction under sparsity constraints with application to MRI and bioluminescence tomography," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, Prague, Czech Republic, 2011).
[70]M. Teboulle and A. Beck, "Gradient-based algorithms with applications to signal-recovery problems," in *Convex Optimization in Signal Processing and Communications* (Cambridge University Press, Cambridge, 2009).
[71]J. D. Lee, Y. Sun, and M. A. Saunders, "Proximal Newton-type methods for minimizing composite functions," SIAM J. Optim. **24**, 1420–1443 (2014).
[72]G. Czjzek, "Distribution of nuclear quadrupole splittings in amorphous materials and the topology of the (V_{zz}, η)-parameter space," Hyperfine Interact. **14**, 189–194 (1983).
[73]B. J. Walder, K. K. Dey, D. C. Kaseman, J. H. Baltisberger, and P. J. Grandinetti, "Sideband separation experiments in NMR with phase incremented echo train acquisition," J. Chem. Phys. **138**, 174203 (2013).
[74]J. H. Baltisberger, P. Florian, E. G. Keeler, P. A. Phyo, K. J. Sanders, and P. J. Grandinetti, "Modifier cation effects on ^{29}Si nuclear shielding anisotropies in silicate glasses," J. Magn. Reson. **268**, 95–106 (2016).
[75]D. Jardón-Álvarez, K. J. Sanders, P. Phyo, J. H. Baltisberger, and P. J. Grandinetti, "Cluster formation of network-modifier cations in cesium silicate glasses," J. Chem. Phys. **148**, 094502 (2018).
[76]D. Srivastava, M. C. Venetos, P. J. Grandinetti, and S. Dwaraknath (2020). "mrsimulator: v0.4.0: A fast solid-state NMR spectrum simulation/analysis library," Zenodo. https://doi.org/10.5281/zenodo.4014009

[77]G. Czjzek, J. Fink, F. Götz, H. Schmidt, J. M. D. Coey, J.-P. Rebouillat, and A. Liénard, "Atomic coordination and the distribution of electric field gradients in amorphous solids," Phys. Rev. B **23**, 2513–2530 (1981).
[78]J.-B. d'Espinose de Lacaillerie, C. Fretigny, and D. Massiot, "MAS NMR spectra of quadrupolar nuclei in disordered solids: The Czjzek model," J. Magn. Reson. **192**, 244–251 (2008).
[79]G. Le Caër, B. Bureau, and D. Massiot, "An extension of the Czjzek model for the distributions of electric field gradients in disordered solids and an application to NMR spectra of ^{71}Ga in chalcogenide glasses," J. Phys.: Condens. Matter **22**, 065402 (2010).
[80]P. Hodgkinson and L. Emsley, "The reliability of the determination of tensor parameters by solid-state nuclear magnetic resonance," J. Chem. Phys. **107**, 4808–4816 (1997).
[81]E. A. Porai-Koshits, "Genesis of concepts on structure of inorganic glasses," J. Non-Cryst. Solids **123**, 1–13 (1990).
[82]A. C. Wright, "Neutron scattering from vitreous silica. V. The structure of vitreous silica: What have we learned from 60 years of diffraction studies?," J. Non-Cryst. Solids **179**, 84–115 (1994).
[83]G. N. Greaves and S. Sen, "Inorganic glasses, glass-forming liquids and amorphizing solids," Adv. Phys. **56**, 1–166 (2007).
[84]I. Farnan, P. J. Grandinetti, J. H. Baltisberger, J. F. Stebbins, U. Werner, M. A. Eastman, and A. Pines, "Quantification of the disorder in network-modified silicate glasses," Nature **358**, 31–35 (1992).
[85]A. K. Varshneya, *Fundamentals of Inorganic Glasses* (Academic Press, Inc., San Diego, CA, 1994).
[86]G. N. Greaves, S. J. Gurman, C. R. A. Catlow, A. V. Chadwick, S. Houde-Walter, C. M. B. Henderson, and B. R. Dobson, "A structural basis for ionic diffusion in oxide glasses," Philos. Mag. A **64**, 1059–1072 (1991).
[87]M. D. Ingram, "Ionic conductivity and glass structure," Philos. Mag. B **60**, 729 (1989).
[88]P. Maass, A. Bunde, and M. D. Ingram, "Ion transport anomalies in glasses," Phys. Rev. Lett. **68**(20), 3064–3067 (1992).
[89]J. O. Isard, "The mixed alkali effect in glass," J. Non-Cryst. Solids **1**, 235 (1969).
[90]D. E. Day, "Mixed alkali glasses—Their properties and uses," J. Non-Cryst. Solids **21**, 343 (1976).
[91]P. Maass, "Towards a theory for the mixed alkali effect in glasses," J. Non-Cryst. Solids **255**, 35–46 (1999).
[92]G. Engelhardt and D. Michel, *High-Resolution Solid-State NMR of Silicates and Zeolites* (John Wiley & Sons, Chichester, 1987).
[93]M. Edén, "NMR studies of oxide-based glasses," Annu. Rep. Prog. Chem., Sect. C: Phys. Chem. **108**, 177–221 (2012).
[94]A.-R. Grimmer, E. F. Gechner, and G. Molgedey, "High resolution ^{29}Si NMR in solid silicates. correlations between shielding tensor and Si-O bond length," Chem. Phys. Lett. **77**, 331–335 (1981).
[95]A.-R. Grimmer, "Correlation between individual Si-O bond lengths and the principal values of the ^{29}Si chemical-shift tensor in solid silicates," Chem. Phys. Lett. **119**, 416–420 (1985).
[96]P. Zhang, P. J. Grandinetti, and J. F. Stebbins, "Anionic species determination in $CaSiO_3$ glass using two-dimensional ^{29}Si NMR," J. Phys. Chem. B **101**(20), 4004–4008 (1997).
[97]M. C. Davis, K. J. Sanders, P. J. Grandinetti, S. J. Gaudio, and S. Sen, "Structural investigations of magnesium silicate glasses by ^{29}Si magic-angle flipping NMR," J. Non-Cryst. Solids **357**, 2787–2795 (2011).
[98]G. L. Turner, K. A. Smith, R. J. Kirkpatrick, and E. Oldfieldt, "Structure and cation effects on phosphorus-31 NMR chemical shifts and chemical-shift anisotropies of orthophosphates," J. Magn. Reson. **70**, 408–415 (1986).
[99]R. K. Brow, C. C. Phifer, G. L. Turner, and R. J. Kirkpatrick, "Cation effects on ^{31}P MAS NMR chemical shifts of metaphosphate glasses," J. Am. Ceram. Soc. **74**, 1287–1290 (1991).
[100]J. B. Murdoch, J. F. Stebbins, and I. S. E. Carmichael, "High-resolution ^{29}Si NMR study of silicate and aluminosilicate glasses: The effect of network-modifying cations," Am. Mineral. **70**, 332–343 (1985).
[101]G. Libourel, C. A. Geiger, L. Merwin, and A. Sebald, "^{29}Si and ^{27}Al MAS-NMR spectroscopy of glasses in the system $CaSiO_3$-$MgSiO_3$-Al_2O_3," Chem. Geol. **96**, 387–397 (1992).
[102]D. J. Srivastava, T. Vosegaard, D. Massiot, and P. J. Grandinetti, "Core scientific dataset model: A lightweight and portable model and file format for multi-dimensional scientific data," PLoS One **15**, e0225953 (2020).
[103]D. J. Srivastava and P. J. Grandinetti (2020). "Silicon-29 NMR experimental datasets used in statistical learning of NMR tensors from 2D isotropic/anisotropic correlation nuclear magnetic resonance spectra," Zenodo. https://doi.org/10.5281/zenodo.3964531
[104]X. Ge, H. Wang, Y. Fan, Y. Cao, H. Chen, and R. Huang, "Joint inversion of T_1–T_2 spectrum combining the iterative truncated singular value decomposition and the parallel particle swarm optimization algorithms," Comput. Phys. Commun. **198**, 59–70 (2016).
[105]X. Ge, H. Chen, Y. Fan, J. Liu, J. Cai, and J. Liu, "An improved pulse sequence and inversion algorithm of T_2 spectrum," Comput. Phys. Commun. **212**, 82–89 (2017).
[106]G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, Springer Texts in Statistics Vol. 103, 1st ed. (Springer New York, NY, 2013).

Supplementary Material for Statistical Learning of NMR tensors from 2D Isotropic/Anisotropic Correlation Nuclear Magnetic Resonance Spectra

Deepansh J. Srivastava and Philip J. Grandinetti
Department of Chemistry, Ohio State University,
100 West 18th Avenue, Columbus, OH 43210, USA

September 4, 2020

Contents

S1 Smooth-LASSO inversion algorithm

A summary of the smooth-LASSO inversion computational approach is given in Algorithm 1.

S2 Role of Hyperparameters

To illustrate the role of the a hyperparameter, we apply the smooth-LASSO inversion method on spectra with increasing noise standard deviations as shown in Fig. S1. In this investigation, the U-1 unimodal ground truth distribution of Fig. 3 is used to generate synthetic spectra with increasing noise standard deviations labeled as (E-1) $\sigma_e = 0.0005$, (E-2) $\sigma_e = 0.001$, (E-3) $\sigma_e = 0.01$, and (E-4) $\sigma_e = 0.02$ in Fig. S1. These synthetic VAS (90°) and MAS sideband spectra are shown in the first and third columns of Fig. S1, respectively. Eash synthetic spectrum and the corresponding fit are depicted in black and red, respectively, with the residuals depicted above each spectrum. The second and fourth columns show the best-fit nuclear shielding tensor distributions obtained from the inversions of the synthetic VAS (90°) and MAS sideband spectra, respectively. The histogram on the top and right of each distribution is the projection on to the respective coordinate axes.

TSVD compression algorithm

Input $\mathbf{K}$, $\mathbf{s}$

$\mathbf{U},\ \boldsymbol{S},\ \mathbf{V}^T = \text{SVD}(\mathbf{K})$

Determine the TSVD truncation parameter, r.

Calculate truncated matrices, $\mathbf{U}_r$, $\boldsymbol{S}_r$, and $\mathbf{V}_r$ from r.

Evaluate the compressed kernel, $\tilde{\mathbf{K}}$, and compressed signal, $\tilde{\mathbf{s}}$.

10-fold cross-validation algorithm

Import: $\tilde{\mathbf{K}}$, $\tilde{\mathbf{s}}$

Set: $\lambda = [\lambda_l]_{l=1}^{n_\lambda}$ and $\alpha = [\alpha_j]_{j=1}^{n_\alpha}$ as the pre-defined range of λ and α values, respectively.

Divide the compressed signal, $\tilde{\mathbf{s}}$, into $[\tilde{\mathbf{s}}^i]_{i=1}^{10}$ test sets of roughly equal size.

Similarly, divide the compressed kernel, $\tilde{\mathbf{K}}$, into $[\tilde{\mathbf{K}}^i]_{i=1}^{10}$.

Create the training sets $[\tilde{\mathbf{s}}^{-i}]_{i=1}^{10}$ by leaving out the i^{th} set.

Create the corresponding training set kernels, $[\tilde{\mathbf{K}}^{-i}]_{i=1}^{10}$.

for $i \leftarrow 1$ **to** 10 **do**

  set: testSet $\leftarrow \tilde{\mathbf{K}}^i, \tilde{\mathbf{s}}^i$; trainingSet $\leftarrow \tilde{\mathbf{K}}^{-i}, \tilde{\mathbf{s}}^{-i}$

  for $j \leftarrow 1$ **to** n_α **do**

    Calculate the augmented kernel, $\mathbf{K}_d^{-i}$, and augmented signal, $\mathbf{s}_d^{-i}$, from $\tilde{\mathbf{K}}^{-i}$ and $\tilde{\mathbf{s}}^{-i}$, respectively, and the hyperparameter α_j,

    for $l \leftarrow 1$ **to** n_λ **do**

      Train

      Solve the following LASSO estimator over the augmented dataset using the hyperparameter λ_l, $\mathbf{f}_i(\lambda_l, \alpha_j) = \underset{\mathbf{f}}{\text{argmin}} \left(\left\langle \|\mathbf{K}_d^{-i}\mathbf{f} - \mathbf{s}_d^{-i}\|_2^2 \right\rangle + \lambda_l \|\mathbf{f}\|_1 \right)$.

      Test

      Calculate the mean square error (MSE) of the test set using the model $\mathbf{f}_i(\lambda_l, \alpha_j)$, $\epsilon_i(\lambda_l, \alpha_j) = \left\langle \left(\mathbf{s}^i - \mathbf{K}^i \mathbf{f}(\lambda_l, \alpha_j)\right)^2 \right\rangle$.

    end

  end

end

Calculate the cross-validation error.

Calculate the optimum hyperparameters.

Solution

Calculate the augmented kernel, $\mathbf{K}_d$, and augmented signal, $\mathbf{s}_d$, from the complete kernel $\tilde{\mathbf{K}}$ and signal $\tilde{\mathbf{s}}$, respectively, and the optimum hyperparameter α^*.

Solve the following LASSO estimator over the augmented dataset using the optimum hyperparameter λ^*, $\mathbf{f} = \underset{\mathbf{f}}{\text{argmin}} \left(\left\langle \|\mathbf{K}_d\mathbf{f} - \mathbf{s}_d\|_2^2 \right\rangle + \lambda^* \|\mathbf{f}\|_1 \right)$,

where, $\mathbf{f}$ is the optimum solution.

Algorithm 1: TSVD-S-LASSO hybrid algorithm with the ten-fold cross-validation.

The mean positions, standard deviations, and correlation coefficients of the inverted distributions along the x and y directions, along with the U-1 ground truth distribution values, are shown in Table S1. While the inversion of the spectra remains stable with increasing noise, the quality of inversion decreases somewhat as the noise in the corresponding spectrum increases. As seen in Table S1, the inverted distributions from the lower noise E-1 and E-2 spectra show good agreement, with mean and the standard deviation within 0.1 ppm of the ground truth distribution values. In comparison, the inverted distributions from the higher noise E-3 and E-4 spectra have mean positions that are displaced slightly away from the ground truth values, and also have larger standard deviations, an indication of a greater uncertainty in the inversion of the anisotropic spectrum.

In comparing the hyperparameters from the E series of inversions in Table S2, we see that the optimum a^* value increases systematically with increasing noise added to the VAS (90°) and MAS sideband spectra. No significant variation in b^* is observed, as expected, since the ground truth distribution is identical for all cases in Fig. S1 and the same x-y grid representation, i.e., basis subspectra, is used for the inversions.

To illustrate the role of the b hyperparameter, we performed the linear inversion method on synthetic spectra as a function of the width of the ground truth distribution, as shown in Fig. S2. The ground truth distributions shown in the first column are bivariate normal distributions with mean values of $\mu_x = 35$ ppm, $\mu_y = 75$ ppm, a correlation coefficient of $r_{xy} = 0.12$, and increasing standard deviations labeled as (S-1) $\sigma_x = \sigma_y = 1.25$ ppm, (S-2) $\sigma_x = \sigma_y = 2.5$ ppm, (S-3) $\sigma_x = \sigma_y = 6.25$ ppm, and (S-4) $\sigma_x = \sigma_y = 7.5$ ppm. As was the case in Figs. 3 and S1, each ground truth distribution is sampled over a high resolution x-y grid and the inversion of each synthetic anisotropic spectrum is carried out over a lower resolution (25×25) x-y grid with each grid cell (basis subspectrum) supersampled on a 5×5 sub-grid. The distributions in the third and the fifth columns are from the inversion of the corresponding synthetic VAS (90°) and MAS sidebands spectra, respectively. The mean positions, standard deviations, and correlation coefficients of these inverted distributions along the x and y directions, shown in Table S1, are in good agreement with their corresponding ground truth distribution values. As all inversions are performed on identical (25×25) x-y grids, we see a systematic increase in the optimum b^* hyperparameter in Table S2 as the sparsity of the solution decreases with increasing distribution width.

S3 Bimodal Distribution

In this section we give additional examples of smooth-LASSO inversions of synthetic datasets generated from various bimodal ground truth distributions. The distribution labeled B-1 in the first column of Fig. S3 consists of a linear combination of Czjzek and extended Czjzek distributions calculated in the ζ and η coordinates according to

$$f_{\text{B-1}}(\zeta,\eta) = \tfrac{1}{2}p_A^{(C)}(\zeta,\eta) + \tfrac{1}{2}p_B^{(XC)}(\zeta,\eta),$$

and mapped onto the x and y grid. The $p_A^{(C)}(\zeta,\eta)$ parameters are $\sigma_{c,A} = 6.64$ ppm and the $p_B^{(XC)}(\zeta,\eta)$ parameters are $\mu_{\zeta,B} = 65$ ppm, $\mu_{\eta,B} = 0$, and $\epsilon_B = 0.21$.

The bimodal distribution labeled B-2 consists of a linear combination of two extended Czjzek distributions calculated in the ζ and η coordinates according to

$$f_{\text{B-2}}(\zeta,\eta) = \tfrac{1}{3}p_A^{(XC)}(\zeta,\eta) + \tfrac{2}{3}p_B^{(XC)}(\zeta,\eta),$$

and mapped onto the x and y grid. The $p_A^{(XC)}(\zeta,\eta)$ distribution parameters are $\mu_{\zeta,A} = 75$ ppm, and $\mu_{\eta,A} = 0.5$, $\epsilon_A = 0.1$ and the $p_B^{(XC)}(\zeta,\eta)$ parameters are $\mu_{\zeta,B} = 55$ ppm, $\mu_{\eta,B} = 0$, and $\epsilon_B = 0.13$.

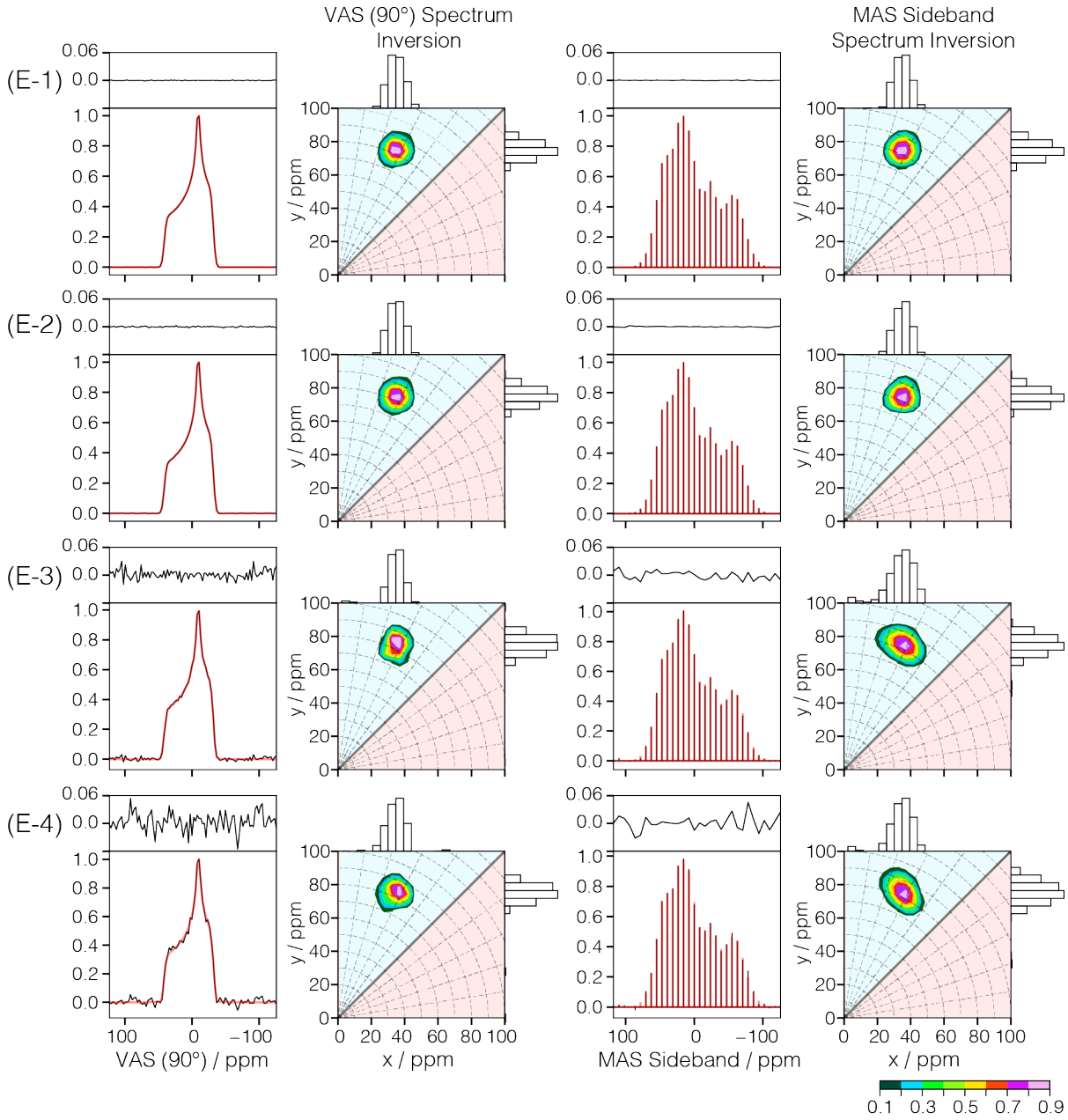

Figure S1: Comparison of smooth-LASSO inversion of synthetic purely anisotropic spectra originating from the (U-1) unimodal bivariate shielding tensor distribution as a function of noise standard deviations, (E-1) $\sigma_e = 0.0005$, (E-2) $\sigma_e = 0.001$, (E-3) $\sigma_e = 0.01$, and (E-4) $\sigma_e = 0.02$. In the first and third columns are the corresponding synthetic VAS (90°) spectrum and MAS sideband spectrum, respectively, along with the best fit spectrum shown in red. Above each spectrum are the residuals. In the second and fourth columns are the corresponding inversion solutions for the shielding tensor distributions along with histogram projections onto the coordinate axes. The contours are drawn at every 10%.

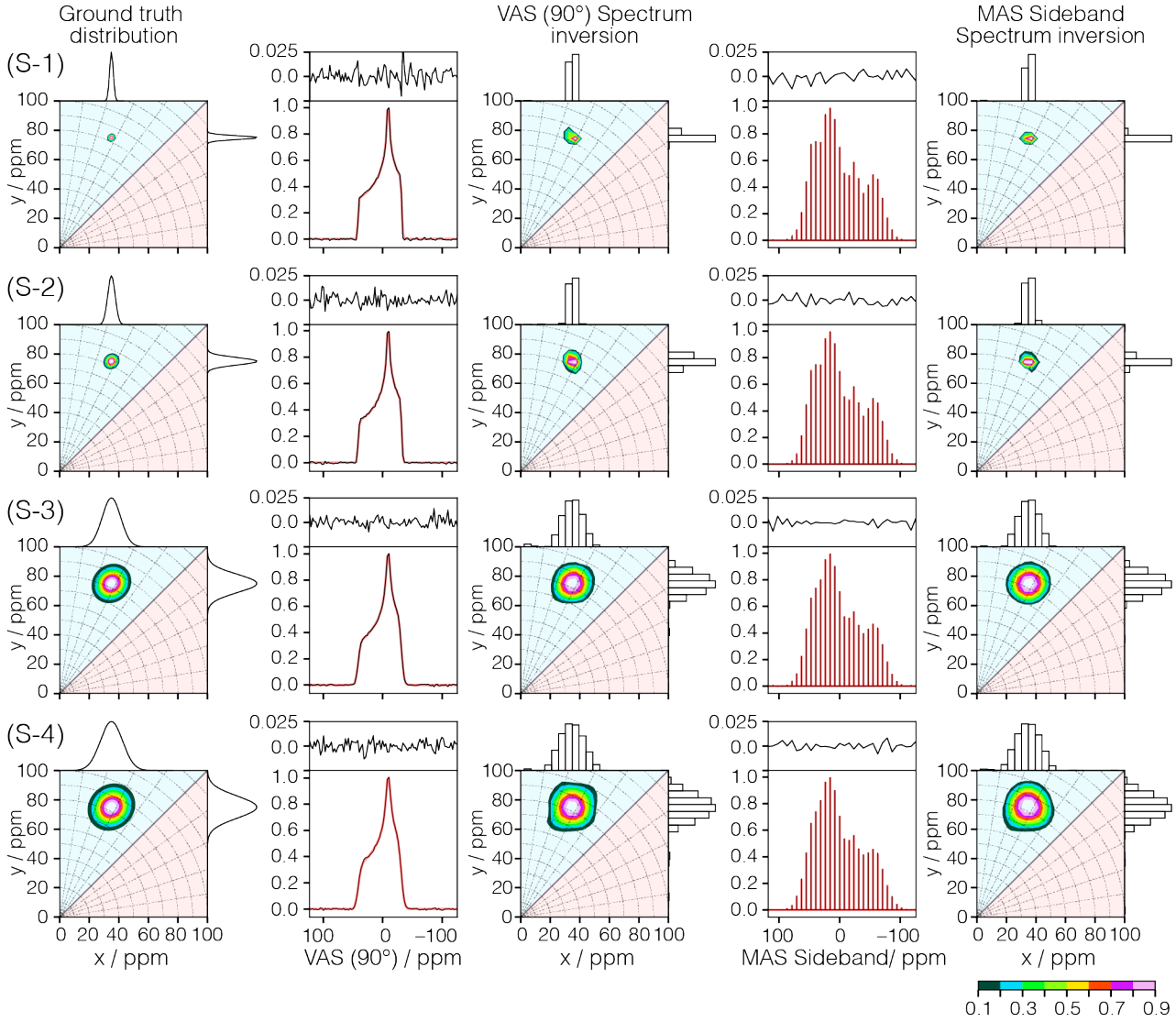

Figure S2: Comparison of smooth-LASSO inversions as a function of the width of the ground truth distribution. The true shielding tensor distributions are shown in the first column. In the second and fourth columns are the corresponding synthetic VAS (90°) spectrum and MAS sideband spectrum, respectively, along with the best fit spectrum shown in red. The inversion solutions for the shielding tensor distributions from the VAS (90°), and MAS sideband spectra are shown in the third and fifth columns, respectively. From top to bottom, the rows are arranged in the order of increasing width of the true distribution, where each distribution is a bivariate normal distribution with mean ($\mu_x = 35$ ppm , $\mu_y = 75$ ppm) and correlation coefficient, $r_{xy} = 0.12$. The standard deviations, $\sigma_x = \sigma_y$ of the distributions are (S-1) 1.25 ppm, (S-2) 2.5 ppm, (S-3) 6.25 ppm, and (S-4) 7.5 ppm, respectively. The contours are drawn at 10%.

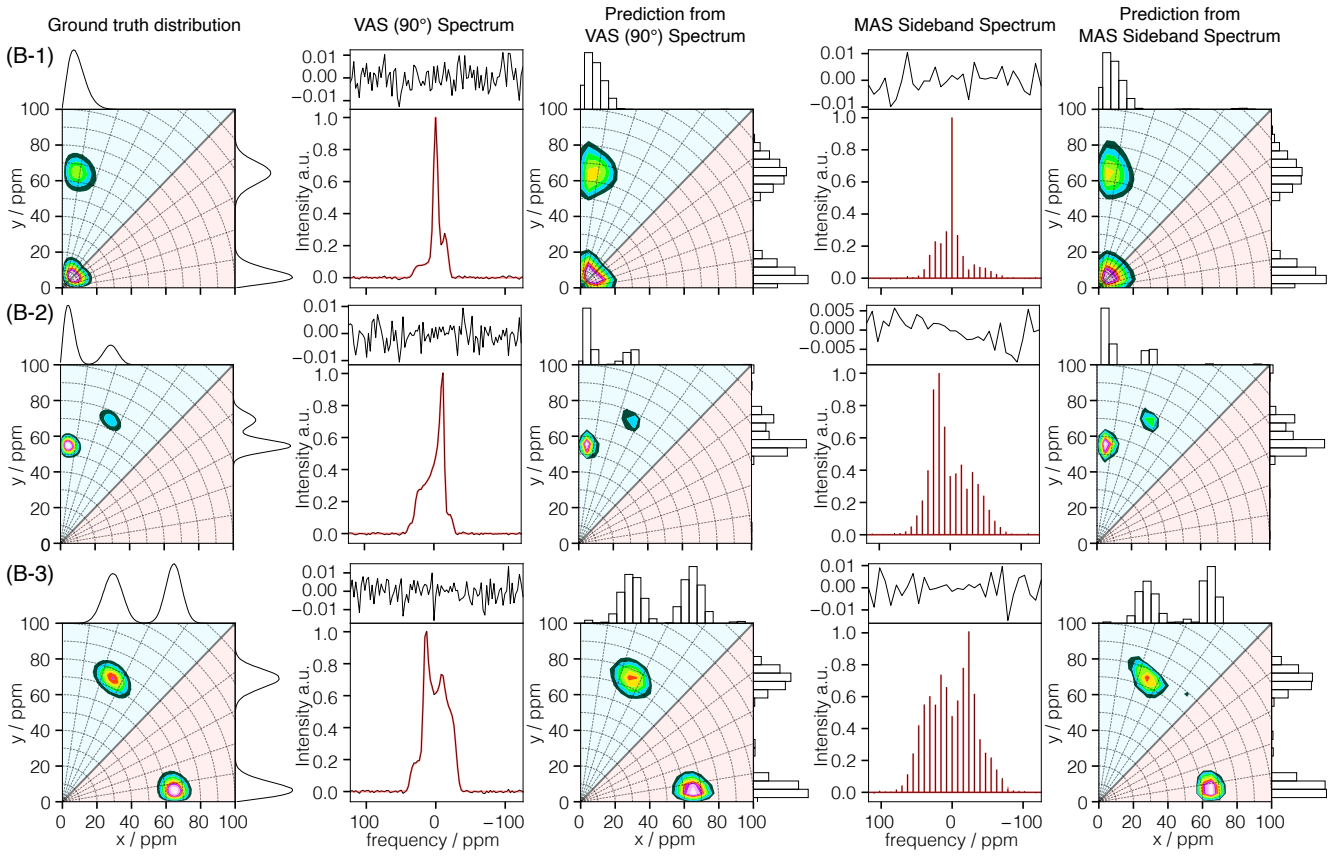

Figure S3: Comparison of smooth-LASSO inversion of synthetic purely anisotropic spectra originating from three different bimodal bivariate shielding tensor distributions along rows labeled B-1, B-2, and B-3. In the first column are the ground truth distributions along with 1D projections onto the coordinate axes. In the second and fourth columns are the corresponding synthetic VAS (90°) spectrum and MAS sideband spectrum, respectively, with noise, $\sigma_e = 0.005$, along with the best fit spectrum shown in red. Above each spectrum are the residuals. In the third and fifth column are the corresponding inversion solutions for the shielding tensor distributions shown with histogram projections onto the coordinate axes. The contours are drawn at every 10%.

The bimodal distribution labeled B-3 consists of a linear combination of two extended Czjzek distributions calculated in the ζ and η coordinates according to

$$f_{\text{B-3}}(\zeta,\eta) = \tfrac{1}{2}p_A^{(XC)}(\zeta,\eta) + \tfrac{1}{2}p_B^{(XC)}(\zeta,\eta)$$

and mapped onto the x and y grid. The $p_A^{(XC)}(\zeta,\eta)$ distribution parameters are $\mu_{\zeta,A} = 75$ ppm, $\mu_{\eta,A} = 0.5$, and $\epsilon_A = 0.13$ and the $p_B^{(XC)}(\zeta,\eta)$ distribution parameters are $\mu_{\zeta,B} = -65$ ppm, $\mu_{\eta,B} = 0$, and $\epsilon_B = 0.16$.

In the second and fourth columns of Fig. S3 are the pure anisotropic 1D VAS 90° and MAS simulations of the respective bimodal distribution in the first column. The plot is black in the simulation with added noise of $\sigma_e = 0.005$, and the plot in red is the fit. On top of the plots is the fit residuals. In the third and fifth columns are the solution form the Smooth-LASSO inversion method applied to the VAS and MAS spectra, respectively. The statistics of the bimodal distribution from each inversion is shown in Table S1.

S4 Tri-modal distributions

Figure S4 depicts the results of the smooth-LASSO inversion on spectra originating from various ground truth trimodal distributions. The trimodal distribution labeled T-1 in Fig. S4 consists of a linear combination of two extended Czjzek distributions and one Czjzek distribution calculated in the ζ and η coordinates according to

$$f_{\mathrm{T1}}(\zeta,\eta) = \tfrac{4}{9}p_A^{(xc)}(\zeta,\eta) + \tfrac{1}{3}p_B^{(xc)}(\zeta,\eta) + \tfrac{2}{9}p_C^{(c)}(\zeta,\eta), \tag{1}$$

and mapped onto the x and y grid. The $p_A^{(xc)}(\zeta,\eta)$ parameters are $\mu_{\zeta,A} = 65$ ppm, $\mu_{\eta,A} = 0$, $\epsilon_A = 0.13$, $p_B^{(xc)}(\zeta,\eta)$ parameters are $\mu_{\zeta,B} = 85$ ppm, $\mu_{\eta,B} = 0.8$, $\epsilon_B = 0.13$, and $p_C^{(c)}(\zeta,\eta)$ parameters are $\sigma_{c,C} = 5.77$ ppm.

The trimodal distribution labeled T-2 consists of a linear combination of three extended Czjzek distributions calculated in the ζ and η coordinates according to

$$f_{\mathrm{T2}}(\zeta,\eta) = \tfrac{1}{4}p_A^{(xc)}(\zeta,\eta) + \tfrac{1}{2}p_B^{(xc)}(\zeta,\eta) + \tfrac{1}{4}p_C^{(xc)}(\zeta,\eta), \tag{2}$$

and mapped onto the x and y grid. The $p_A^{(xc)}(\zeta,\eta)$ parameters are $\mu_{\zeta,A} = 45$ ppm, $\mu_{\eta,A} = 0.1$, $\epsilon_A = 0.2$, $p_B^{(xc)}(\zeta,\eta)$ parameters are $\mu_{\zeta,B} = 85$ ppm, $\mu_{\eta,B} = 0.3$, $\epsilon_B = 0.1$, and $p_C^{(xc)}(\zeta,\eta)$ parameters are $\mu_{\zeta,C} = -55$ ppm, $\mu_{\eta,C} = 0$, $\epsilon_C = 0.2$.

The trimodal distribution labeled T-3 consists of a linear combination of two extended Czjzek distributions and one Czjzek distribution calculated in the ζ and η coordinates according to

$$f_{\mathrm{T3}}(\zeta,\eta) = \tfrac{3}{5}p_A^{(xc)}(\zeta,\eta) + \tfrac{1}{5}p_B^{(xc)}(\zeta,\eta) + \tfrac{1}{5}p_C^{(c)}(\zeta,\eta) \tag{3}$$

and mapped onto the x and y grid. The $p_A^{(xc)}(\zeta,\eta)$ parameters are $\mu_{\zeta,A} = 80$ ppm, $\mu_{\eta,A} = 0.6$, $\epsilon_A = 0.17$, $p_B^{(xc)}(\zeta,\eta)$ parameters are $\mu_{\zeta,B} = -70$ ppm, $\mu_{\eta,B} = 0.2$, $\epsilon_B = 0.2$, and $p_C^{(c)}(\zeta,\eta)$ parameters are $\sigma_{c,C} = 6.93$ ppm.

The plots are organized in the same manner as Fig. S3. As before, Gaussian noise of $\sigma_e = 0.005$ was used in generating each synthetic spectrum in Fig. S4. In all three cases, the inversion of the synthetic VAS (90°) and MAS sideband spectra recovers trimodal distributions with resolved modes and is consistent with the respective ground truth distributions. Table S1 lists the statistics—probability, mean, and standard deviation—of the individual modes of the bivariate distributions for the ground truth and recovered distributions. In all cases, good agreement is obtained with the mean values of the ground truth distributions.

Mathematically, the integration volume of the $f(x,y)$ distribution is conserved and is equal to the area under the spectrum, that is,

$$\int f(x,y)dxdy = \int s(\nu)d_\nu.$$

Note, the above integral equation is valid because the kernel, $\mathcal{K}$, is normalized such that every basis spectrum within the kernel is of unit area. Although the total volume of the predicted multimodal distribution is conserved, the volume under the individual modes may not match the ground truth distribution. Nonetheless, in these trimodal inversion examples, reasonably good agreement is obtained for the volume ratio of the three modes, as seen in Table S1. Of course, quantifying integrated mode volumes becomes more challenging when distributions are close or overlapping. The quality of the inversion will decrease with increasing noise, making higher modal distributions harder to invert. As the inversion gets more challenging, the integrated volume of the solution will start to disperse into extraneous solutions as can be seen in the T-3 distribution inversions of Fig. S4.

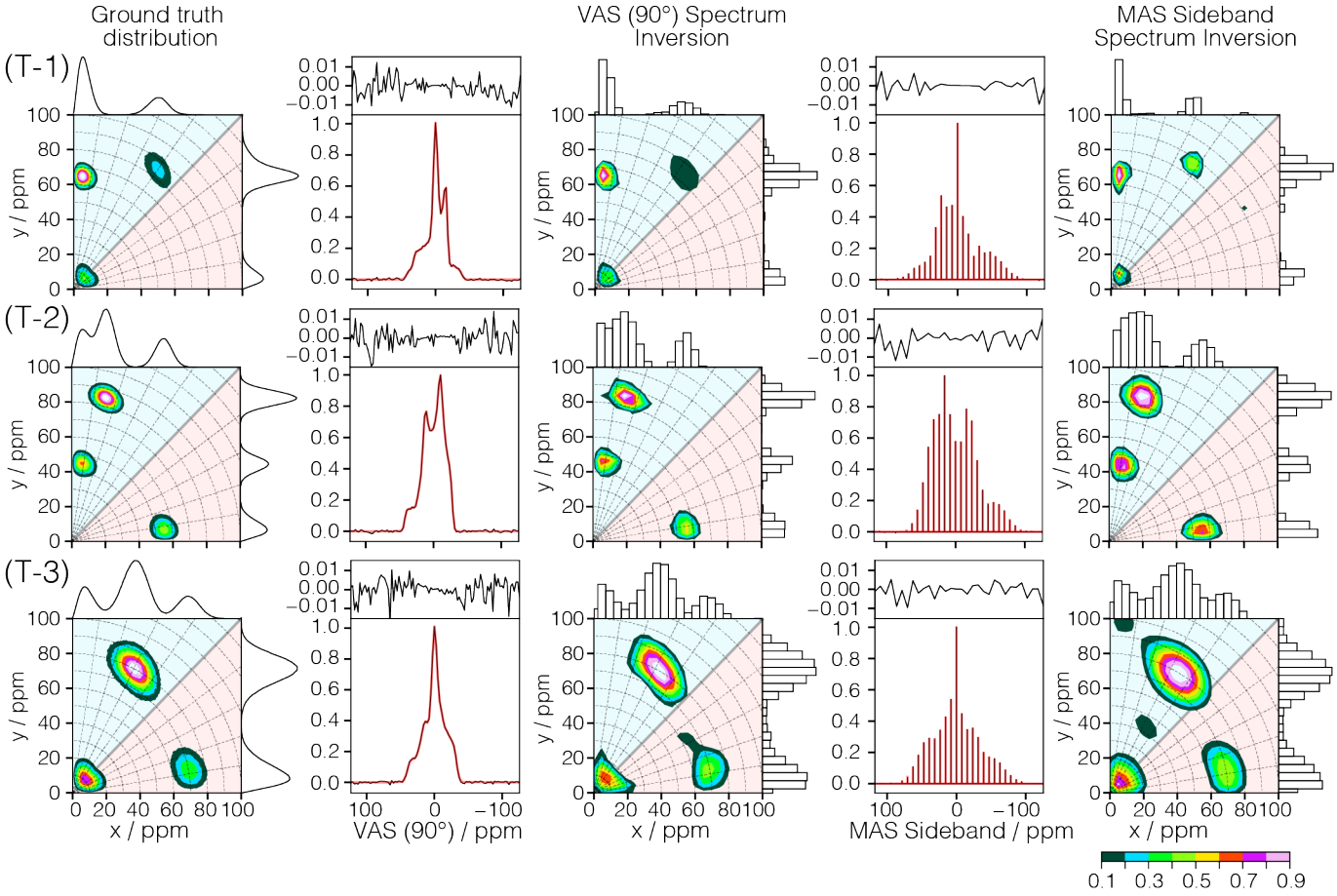

Figure S4: Comparison of smooth-LASSO inversion of synthetic purely anisotropic spectra originating from three different trimodal bivariate shielding tensor distributions along rows labeled T-1, T-2, and T-3. In the first column are the ground truth distributions along with 1D projections onto the coordinate axes. In the second and fourth columns are the corresponding synthetic VAS (90°) spectrum and MAS sideband spectrum, respectively, with noise, $\sigma_e = 0.005$, along with the best fit spectrum shown in red. Above each spectrum are the residuals. In the third and fifth columns are the corresponding inversion solutions for the shielding tensor distributions shown with histogram projections onto the coordinate axes. The contours are drawn at every 10%.

Distribution		Ground truth				VAS (90°) inversion				MAS sidebands inversion			
		%	$\langle A\rangle_x$/ppm	$\langle A\rangle_y$/ppm	r_{xy}	%	$\langle A\rangle_x$/ppm	$\langle A\rangle_y$/ppm	r_{xy}	%	$\langle A\rangle_x$/ppm	$\langle A\rangle_y$/ppm	r_{xy}
U-1	$p^{(n)}(x,y)$	100	35.0 ± 5.0	75.0 ± 5.0	0.12	100	35.1 ± 5.4	75.3 ± 5.2	0.05	100	35.2 ± 5.3	75.5 ± 5.2	0.05
U-2	$p^{(n)}(\zeta,\eta)$	100	68.7 ± 5.3	28.5 ± 9.4	−0.61	100	68.7 ± 5.8	28.4 ± 9.1	−0.65	100	68.9 ± 5.6	28.3 ± 9.4	−0.54
U-3	$p^{(c)}(\zeta,\eta)$	100	8.7 ± 4.5	8.7 ± 4.5	−0.21	100	8.7 ± 4.0	8.7 ± 4.0	−0.28	100	8.8 ± 4.4	8.6 ± 4.2	−0.11
U-4	$p^{(xc)}(\zeta,\eta)$	100	24.1 ± 4.3	7.3 ± 3.6	−0.28	100	24.1 ± 4.0	7.5 ± 3.3	−0.37	100	24.3 ± 3.9	7.1 ± 2.9	−0.24
E-1	$p^{(n)}(x,y)$	100	35.0 ± 5.0	75.0 ± 5.0	0.12	100	34.9 ± 5.1	75.0 ± 5.0	0.11	100	35.0 ± 4.9	75.0 ± 4.9	0.11
E-2	$p^{(n)}(x,y)$	100	35.0 ± 5.0	75.0 ± 5.0	0.12	100	34.8 ± 5.0	75.1 ± 5.2	0.10	100	34.9 ± 5.1	75.0 ± 4.9	0.07
E-3	$p^{(n)}(x,y)$	100	35.0 ± 5.0	75.0 ± 5.0	0.12	100	34.6 ± 5.7	75.1 ± 5.7	−0.14	100	33.8 ± 8.4	75.0 ± 8.4	−0.09
E-4	$p^{(n)}(x,y)$	100	35.0 ± 5.0	75.0 ± 5.0	0.12	100	34.7 ± 5.9	75.6 ± 5.9	0.35	100	34.0 ± 7.5	75.7 ± 8.2	−0.24
S-1	$p^{(n)}(x,y)$	100	35.0 ± 1.3	75.0 ± 1.3	0.12	100	34.9 ± 3.3	75.1 ± 3.3	0.33	100	35.0 ± 3.6	75.0 ± 3.5	−0.5
S-2	$p^{(n)}(x,y)$	100	35.0 ± 2.5	75.0 ± 2.5	0.12	100	34.8 ± 3.5	75.1 ± 4.2	−0.16	100	35.2 ± 2.8	75.1 ± 3.1	−0.18
S-3	$p^{(n)}(x,y)$	100	35.0 ± 6.25	75.0 ± 6.25	0.12	100	34.6 ± 7.5	75.3 ± 7.5	0.04	100	34.9 ± 6.7	75.0 ± 6.8	0.08
S-4	$p^{(n)}(x,y)$	100	35.0 ± 7.5	75.0 ± 7.5	0.12	100	34.5 ± 8.0	75.0 ± 8.2	0.17	100	34.3 ± 8.1	75.4 ± 9.1	0.0
B1	$p_A^{(c)}(\zeta,\eta)$	50	7.6 ± 3.9	7.6 ± 3.9	−0.21	51	7.4 ± 5.0	7.4 ± 5.2	−0.35	52	7.6 ± 5.2	7.5 ± 5.1	−0.1
	$p_B^{(xc)}(\zeta,\eta)$	50	10.3 ± 5.3	64.3 ± 6.2	−0.15	49	8.3 ± 5.5	64.9 ± 7.3	−0.03	48	8.5 ± 5.2	65.1 ± 8.0	−0.1
B2	$p_A^{(xc)}(\zeta,\eta)$	33.3	29.0 ± 4.2	68.9 ± 4.2	−0.41	30	29.4 ± 3.5	69.6 ± 3.9	−0.46	28	30.3 ± 2.3	68.8 ± 3.0	−0.37
	$p_B^{(xc)}(\zeta,\eta)$	66.6	6.1 ± 4.2	55.3 ± 4.7	0.49	70	5.5 ± 2.5	55.0 ± 3.8	0.1	72	5.9 ± 2.1	54.9 ± 3.7	0.07
B3	$p_A^{(xc)}(\zeta,\eta)$	50	29.1 ± 5.6	68.8 ± 5.4	−0.41	50	29.4 ± 5.7	68.9 ± 5.3	−0.31	50	29.3 ± 7.0	68.6 ± 5.4	−0.66
	$p_B^{(xc)}(\zeta,\eta)$	50	64.5 ± 4.7	7.9 ± 4.1	−0.12	50	64.7 ± 5.0	7.5 ± 3.5	−0.13	50	64.5 ± 3.4	8.3 ± 3.4	−0.08
T-1	$p_A^{(xc)}(\zeta,\eta)$	44	6.4 ± 3.3	64.6 ± 3.8	−0.09	45	6.5 ± 3.0	64.9 ± 3.7	−0.19	42	5.6 ± 1.9	65.2 ± 4.0	0.24
	$p_B^{(xc)}(\zeta,\eta)$	33	50.1 ± 6.0	68.4 ± 7.5	−0.49	32	51.4 ± 6.7	67.5 ± 7.3	−0.64	35	50.6 ± 8.5	70.2 ± 7.4	−0.86
	$p_C^{(xc)}(\zeta,\eta)$	22	7.6 ± 3.9	7.6 ± 3.9	−0.21	23	7.8 ± 3.3	7.9 ± 3.9	−0.03	23	7.8 ± 5.4	8.8 ± 3.4	0.31
T-2	$p_A^{(xc)}(\zeta,\eta)$	25	7.5 ± 4.0	44.5 ± 4.4	−0.05	26	8.4 ± 3.9	44.9 ± 4.2	−0.28	26	8.5 ± 3.6	44.1 ± 4.1	−0.11
	$p_B^{(xc)}(\zeta,\eta)$	50	20.4 ± 5.0	82.2 ± 4.3	−0.26	49	20.0 ± 5.8	83.1 ± 4.8	−0.43	50	19.7 ± 5.3	82.9 ± 5.1	−0.18
	$p_C^{(xc)}(\zeta,\eta)$	25	54.5 ± 5.0	8.3 ± 4.3	−0.15	24	54.8 ± 4.1	8.8 ± 3.8	−0.10	24	54.9 ± 5.4	7.4 ± 3.1	0.06
T-3	$p_A^{(xc)}(\zeta,\eta)$	60	36.8 ± 7.6	70.8 ± 8.0	−0.45	59	37.9 ± 8.0	70.6 ± 8.8	−0.53	61	39.0 ± 10.2	69.0 ± 10.1	−0.44
	$p_B^{(xc)}(\zeta,\eta)$	20	68.3 ± 6.9	14.7 ± 7.2	−0.30	20	65.9 ± 8.9	16.5 ± 9.0	−0.62	19	67.6 ± 7.1	14.1 ± 8.0	−0.39
	$p_C^{(c)}(\zeta,\eta)$	20	9.1 ± 4.7	9.1 ± 4.7	−0.21	21	8.6 ± 5.7	9.5 ± 7.1	−0.30	20	9.1 ± 6.7	9.9 ± 7.8	0.32

Table S1: Comparison of integrated volumes (%), means, standard deviations, and correlation coefficients, r_{xy}, of ground truth and inverted distributions from the VAS (90°) and MAS sidebands spectra. Here, $\langle A\rangle_x$ and $\langle A\rangle_y$ denotes the mean ± standard deviation of the distribution with respect to the x and y coordinates, respectively.

		90° spinning spectrum				MAS sideband spectrum			
Test	σ_e	$\alpha^*/10^{-6}$	$\lambda^*/10^{-6}$	$a^*/10^{-6}$	b^*	$\alpha^*/10^{-6}$	$\lambda^*/10^{-6}$	$a^*/10^{-6}$	b^*
U-1	0.005	2.98	1.27	4.25	0.70	2.98	1.83	4.81	0.63
U-2	0.005	1.83	1.27	3.11	0.59	12.74	1.83	14.58	0.87
U-3	0.005	11.3	69.5	80.8	0.14	69.5	42.8	112.3	0.62
U-4	0.005	0.83	3.6	4.4	0.2	5.8	3.6	9.4	0.62
E-1	0.0005	0.043	0.06	0.10	0.41	0.18	0.13	0.31	0.59
E-2	0.001	0.48	0.38	0.86	0.56	0.18	0.13	0.31	0.59
E-3	0.01	1.13	1.83	2.96	0.38	7.85	2.64	10.48	0.75
E-4	0.02	2.98	5.46	8.43	0.35	7.85	3.79	11.64	0.67
S-1	0.005	0.0003	2.64	2.64	0.0001	0.0018	1.27	1.27	0.001
S-2	0.005	0.023	2.64	2.87	0.08	0.0004	1.83	1.83	0.0002
S-3	0.005	5.46	1.83	7.29	0.75	3.36	0.62	3.98	0.85
S-4	0.005	14.38	1.27	15.7	0.92	14.4	0.89	15.3	0.94
T-1	0.005	0.89	1.83	2.72	0.33	0.43	1.83	2.26	0.19
T-2	0.005	0.51	0.62	1.13	0.45	7.85	1.83	9.68	0.81
T-3	0.005	12.7	1.73	14.5	0.88	162.4	2.64	165.0	0.98

Table S2: Comparison of the optimum hyperparameters, α^* and λ^*, as well as the alternative optimum hyperparameters, a^* and b^*, obtained from cross-validation of the inversion of the VAS (90°) and MAS sidebands spectra. The σ_e value is the noise standard deviation added to each simulated spectrum. The α^* and λ^* values are determined over a 20×20 α-λ grid using a ten-fold cross-validation method for the VAS (90°) and MAS sidebands spectrum inversion. The alternative hyperparameters, a^* and b^*, are derived from α^* and λ^* according to Eq. (34) of the main text. The SVD truncation index, r, depends on the kernel $\mathcal{K}$, which in turn depends on the x-y grid dimensions and the anisotropic dimension of the spectrum. Values of $r = 62$ for all VAS (90°) inversions and $r = 31$ for all MAS sideband inversions are chosen using the maximum entropy method[1] outlined in Appendix A1 of the main text. These values remain the same since the x-y grid and the anisotropic dimensions are identical for all synthetic dataset inversions.

ν_r / kHz	r	$\alpha^*/10^{-6}$	$\lambda^*/10^{-6}$	$a^*/10^{-6}$	b^*
0.8125*	31	0.15	0.15	0.3	0.5
1.0000*	31	0.08	0.08	0.16	0.5
1.1875	30	0.22	0.1	0.32	0.68
1.3750	27	0.1	0.46	5.64	0.18
1.5625	24	0.22	0.46	0.68	0.32
2.0000	21	0.22	0.1	0.32	0.68
2.5000	18	0.31	0.06	0.37	0.85
3.0000	18	0.56	17.8	18.3	0.03
4.0000	15	3.2	0.32	3.48	0.91
8.0000	12	10	0.32	10.32	0.97

Table S3: The hyperparameters of the optimum solution of the MAS sidebands inversion as a function of spinning speed. Here, r, is the number of singular values retained in the solution. The maximum number of singular values is 32. Synthetic datasets and inversions not shown in main text are marked with an asterisk.

S5 Spinning Speed Hyperparameters

Table S3 gives the optimum hyperparameters for the smooth-LASSO inversions of Fig. 8 in the main text.

S6 Comparison of Regularization Methods

In Fig. 7 of the main manuscript, we show the comparison of different regularization methods. Table S4 gives the optimum hyperparameters obtained. In those examples implement the solvers from the scikit-learn[2] library, which minimizes the following objective functions

$$\underset{\mathbf{f}\geq 0}{\operatorname{argmin}} \left(\frac{1}{2m} \|\mathbf{s} - \mathbf{K}\cdot\mathbf{f}\|_2^2 + \lambda_l \|\mathbf{f}\|_1 \right), \tag{4}$$

and

$$\underset{\mathbf{f}\geq 0}{\operatorname{argmin}} \left(\frac{1}{2m} \|\mathbf{s} - \mathbf{K}\cdot\mathbf{f}\|_2^2 + \alpha_e l_{1,\text{ratio}} \|\mathbf{f}\|_1 + \frac{1}{2}\alpha_e (1 - l_{1,\text{ratio}}) \|\mathbf{f}\|_2^2 \right), \tag{5}$$

for ℓ_1 and elastic net regularization, respectively. Here, λ_l is hyperparameter for the ℓ_1 term in Eq. (4), and α_e, $\ell_{1,\text{ratio}}$ are the total regularization and relative fraction of ℓ_1 term in the elastic net regularization of Eq. (5), respectively. The term, m, is the sampling size of the sample vector $\mathbf{s}$, $\mathbf{K}$ is the kernel, and $\mathbf{f}$ is the solution.

References

[1] R. Varshavsky, A. Gottlieb, M. Linial, and D. Horn. Novel unsupervised feature filtering of biological data. *Bioinformatics*, 22:e507 – e513, Mar-07-2007 2006.

[2] Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton,

Jake VanderPlas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, 2013.

	Grid size	resolution/ppm^2	$\lambda_l^*/10^{-6}$			
ℓ_1	20×20	5.81×5.81	1.23			
	30×30	3.87×3.87	0.35			
	40×40	2.91×2.91	0.35			
	50×50	2.32×2.32	0.19			
	size	resolution/ppm^2	$\alpha_e^*/10^{-6}$	$l_{1,\text{ratio}}^*$		
Elastic net	20×20	5.81×5.81	16.9	0.7		
	30×30	3.87×3.87	23.4	0.5		
	40×40	2.91×2.91	33.0	0.35		
	50×50	2.32×2.32	46.3	0.25		
	size	resolution/ppm^2	$\alpha^*/10^{-6}$	$\lambda^*/10^{-6}$	$a^*/10^{-6}$	b^*
S-LASSO	20×20	5.81×5.81	2.85	2.31	5.16	0.55
	30×30	3.87×3.87	15.2	1.23	16.4	0.93
	40×40	2.91×2.91	35.8	0.66	35.77	0.98
	50×50	2.32×2.32	81.1	0.66	81.77	0.99

Table S4: The hyperparameters corresponding to the optimum solution for the ℓ_1, elastic net, and smooth-LASSO regularized linear problems as a function of the grid-resolution.