ARTICLE

Reaching the sparse-sampling limit for reconstructing a single peak in a 2D NMR spectrum using iterated maps

Robert L. Blum[1] · Jared Rovny[1] · J. Patrick Loria[2,3] · Sean E. Barrett[1]

Abstract

Many of the ubiquitous experiments of biomolecular NMR, including R_2, $R_{1\rho}$, and CEST, involve acquiring repeated 2D spectra under slightly different conditions. Such experiments are amenable to acceleration using non-uniform sampling spectral reconstruction methods that take advantage of prior information. We previously developed one such technique, an iterated maps method (DiffMap) that we successfully applied to 2D NMR spectra, including R_2 relaxation dispersion data. In that prior work, we took a top-down approach to reconstructing the 2D spectrum with a minimal number of sparse samples, reaching an undersampling fraction that appeared to leave some room for improvement. In this study, we develop an in-depth understanding of the action of the DiffMap algorithm, identifying the factors that cause reconstruction errors for different undersampling fractions. This improved understanding allows us to formulate a bottom-up approach to finding the lowest number of sparse samples required to accurately reconstruct individual spectral features with DiffMap. We also discuss the difficulty of extending this method to reconstructing many peaks at once, and suggest a way forward.

Keywords Difference map · DiffMap · Sparse sampling · Reconstruction · Nonuniform sampling

Introduction

NMR experiments aiming to study molecular structure and dynamics will often require the acquisition of large, multidimensional data sets. While these nD experiments (n > 1) can reveal more information than a 1D experiment, they also take much longer, with durations proportional to the number of sampled points in the n-1 indirect dimension(s). In order to use the discrete Fourier transform to generate the spectrum, the points are "densely-sampled," i.e., measuring all time-domain points lying on a Cartesian grid of uniform intervals consistent with the Nyquist-Shannon sampling theorem.

Sparse sampling, also known as non-uniform sampling (NUS), is a common approach to speeding up the acquisition of nD NMR data sets. The speed-up comes from skipping over particular indirect-dimension data points (saving entire 1D scans); the unmeasured points are typically set to zero to fill out the nD grid. Some alternative, non-Fourier processing algorithm must then be used to reconstruct the spectrum. Examples of such reconstruction approaches include maximum entropy (MaxEnt) (Hoch and Stern 1996; Mobli and Hoch 2008; Hyberts et al. 2010; Paramasivam et al. 2012), iterative soft thresholding (IST) (Stern et al. 2007), iteratively reweighted least squares (IRLS) (Schlossmacher 1973; Kazimierczuk and Orekhov 2012), multidimensional decomposition (MDD) (Orekhov and Jaravine 2011), and compressed sensing (CS) (Kazimierczuk and Orekhov 2011; Bostock and Nietlispach 2018), among many others (Stanek and Koźmiński 2010; Ying et al. 2017; Billeter 2017; Rovnyak and Schuyler 2018).

Motivated by the need to accelerate 3D MRI of solids experiments, we recently developed one such alternative reconstruction method, the so-called iterated maps approach, which is based upon Elser's "Difference Map" algorithm

✉ Sean E. Barrett
sean.barrett@yale.edu

[1] Department of Physics, Yale University, 217 Prospect St., New Haven, CT 06511, USA

[2] Department of Chemistry, Yale University, 225 Prospect St., New Haven, CT 06511, USA

[3] Department of Molecular Biophysics and Biochemistry, Yale University, 225 Prospect St., New Haven, CT 06511, USA
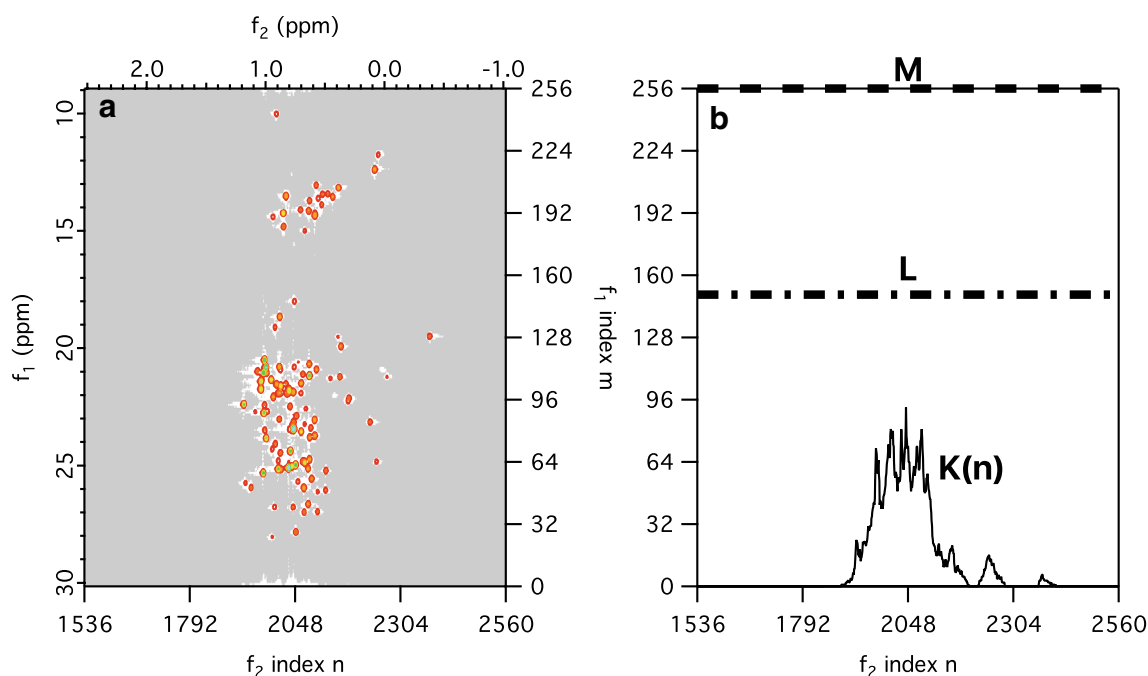
Fig. 1 Background and motivation of this work. **a** Contour plot showing the central part of a $M \times N = 256 \times 4096$ 2D spectrum from the pseudo-3D data set, with the mask($\mathbf{f}$) = 0 regions in gray. Axes show the frequency scaling (left and top) as well as the indices $n = \{0, 1, \dots, 4095\}$ (bottom) and $m = \{0, 1, \dots, 255\}$ (right). **b** Previously achieved and predicted limits to sparse sampling for the data set shown in **a**. Dense sampling corresponds to acquiring all $M = 256$ complex points (dashed line), or 128 complex $t_1 \geq 0$ values. In prior work (Frey et al. 2013), we were able to push a global undersampling scheme down to $L = 150$ (dot-dashed line). However, for each f_2 column of the spectrum, the number of "significant" mask($\mathbf{f}$) $\neq 0$ points $K(n)$ (solid line) is less than L, indicating that even better results may be possible. In this paper, we attempt to push our sampling fraction down towards $K(n)$, by focusing on reconstructing individual peaks in individual columns

(Fienup 1982; Elser 2003; Elser et al. 2007; Frey et al. 2013). In that paper, the iterated maps method (or "DiffMap" for short) was successfully applied to existing 2D NMR spectra (both liquid state and MAS of solids) and 3D MR images of solids. Our iterated maps approach is related to SIFT (Matsuki et al. 2009, 2010, 2011), as well as to POCS (Haacke et al. 1991). All of these methods take advantage of the fact that in many NMR/MRI experiments, the user knows where spectral features "should" and "should not" appear, even if the precise details of the spectrum (e.g., point-by-point amplitude) remain to be determined. This information can be exploited in a deterministic algorithm that nudges a sparsely-sampled data set towards its densely-sampled limit.

Many of the ubiquitous experiments of biomolecular NMR, including R_2, $R_{1\rho}$, and CEST experiments (Palmer et al. 2001; Loria et al. 1999; Neudecker et al. 2009; Palmer and Massi 2006; van Zijl and Yadav 2011; Walinda et al. 2018), are ideal targets for the iterated maps approach, since the series of ~ 20 2D spectra that are typically collected tend to be very similar to each other. Moreover, even if a few slices are densely sampled, sparsely sampling the rest will greatly compress the total duration of the experiment.

In our earlier work, one of the 2D NMR examples was the $\tau_{cp} = 0$ slice of a larger, pseudo-3D data set, acquired as part

of an R_2 dispersion experiment. For that $\tau_{cp} = 0$ slice, we were able to lower the number of sampled indirect time points N_{t_1} all the way down to an undersampling fraction of 58.6% of the dense data ($N_{t_1}/N_{t_1}^{\text{dense}} = 75/128$), while still obtaining a high-quality spectrum. As we will explain below, our approach to making the spectrum entirely real causes M, the number of f_1 rows, to be given by $M = 2N_{t_1}^{\text{dense}} = 256$. So, defining $L = 2 \times 75 = 150$, we had clearly pushed L well below M (Fig. 1). While this was a significant advance, we did not have a deep understanding of why the method broke down at L. In fact, there seemed to be additional room for improvement, since counting the number of significant points in each f_2 column n yielded even smaller values $K(n)$ (i.e., $M > L > K(n)$, see Fig. 1). Could we find a way to push $2N_{t_1}$ below L, and perhaps all the way down to $\sim K(n)$, without hurting the quality of the reconstruction? In other words, what are the fundamental limits of this approach?

To answer this question, we first needed to develop an improved understanding of how the iterated maps reconstruction method breaks down as N_{t_1} is lowered, which included understanding the effects of multiple peaks, spectral tails, noise, and other artifacts. The behavior is very complicated across the whole 2D spectrum. To make

progress, we decided to start with a simpler problem, by focusing on understanding the reconstruction of a single peak in a single f_2 column. We then set out to answer two main questions:

(Q.I) For a high-quality reconstruction of a single peak p in a single f_2 column, what is the minimum number of sparsely-sampled data points, $N_{t_1}^p$?

(Q.II) If we instead wish to reconstruct many (or all) peaks at once, what is the new minimum number of sparsely-sampled data points, $N_{t_1}^*$?

Answering question 1 will take us through most of the paper. As we will show, we can go very close to $K(n)/2$ for most peaks, because we can predict what values of N_{t_1} are susceptible to large errors, as well as the likely range of the "residual" small errors. Question 2 will be addressed briefly at the end, where we show that it is hard to beat $L/2$ once the user tries to reconstruct many peaks simultaneously. A technique to get past this problem for certain types of pseudo-3D experiment is introduced in a companion paper (Rovny et al. 2019), which we will briefly discuss further at the very end of this paper.

Background

The NMR data set used to test our method

In this study of the DiffMap, we focus on reconstructing 2D slices drawn from an existing, pseudo-3D data set, which was previously studied in Frey et al. (2013). This particular pseudo-3D data set is a series of twelve liquid-state 2D NMR spectra, from a $^{13}CH_3$ multiple-quantum CPMG relaxation dispersion experiment of an Isoleucine, Leucine, Valine ^2H, ^{13}C-methyl labeled sample of imidazole glycerol phosphate synthase (IGPS) from *T. maritima* (Lipchock and Loria 2010; Lisi et al. 2018). The data was collected at 14.1 T and 30 °C with 120 t_1 increments and 1889 t_2 increments, using the sequence described by Korzhnev et al. (2004). We zero-fill the data in both dimensions to the next largest power of 2, giving us 128 t_1 points and 2048 t_2 points, which we treat as the set of "measured" points to simplify the reconstruction problem. The ^1H carrier was centered at 0.75 ppm (not 4.7 ppm as incorrectly stated in Frey et al. (2013)) with a spectral width of 8500 Hz, while the ^{13}C carrier was centered at 19.5 ppm with a spectral width of 3200 Hz. The twelve spectra correspond to τ_{cp} values of 0, 0.4167, 0.5, 0.625, 0.7682, 1.0, 1.4286, 2.0, 2.5, 3.333, 5.0, and 10.0 ms. We will refer to these below by their respective indices $i_{\tau_{cp}} = 1, 2, \ldots, 12$. The $\tau_{cp} = 0$ ms slice ($i_{\tau_{cp}} = 1$) of this data set was the subject of Figs. 2, 3 of our earlier paper (Frey

et al. 2013). We subsequently noticed that the spectral tails of that $i_{\tau_{cp}} = 1$ slice are wider than all of the others, so we set that slice aside in the current analysis. Instead, we will assume the $\tau_{cp} = 0.4167$ ms slice ($i_{\tau_{cp}} = 2$) exists as a dense data set to make our $\hat{P}_1$ mask, which will then be used to reconstruct sparsely-sampled versions of that slice and the remaining ten slices, $i_{\tau_{cp}} = 2, 3, 4, \ldots, 12$.

Iterated maps reconstruction (DiffMap)

The general approach to spectral reconstruction using iterated maps has been explained previously (Frey et al. 2013). We quickly review how the method is applied to the 2D NMR data studied here, while noting a few differences from the prior work.

We define the complex vector $\mathbf{T}(\mathbf{t})$ to be the densely-sampled data set in the 2D time-domain ($\mathbf{t} = (t_1, t_2)$), where t_1 and t_2 are the indirect and direct dimensions, respectively. We include the argument $\mathbf{t}$ to clarify that $\mathbf{T}$ is the time-domain data, and will use $T(\mathbf{t})$ to refer to the value of $\mathbf{T}$ at coordinate $\mathbf{t}$. As acquired, a densely sampled "States"-like data set is typically used to fill only the first quadrant of the 2D plane, with ($t_1 \geq 0, t_2 \geq 0$). As in Frey et al. (2013), we instead use the same "States"-like data to fill all four quadrants of the time domain, such that the resulting $M \times N$ grid of complex data points has Hermitian symmetry about the origin (i.e., $T(\mathbf{t}) = T^*(-\mathbf{t})$) (Frey et al. 2013; States et al. 1982; Mayzel et al. 2014). This ensures that the target dense spectrum, $\tilde{\mathbf{T}}(\mathbf{f}) = \text{Ph}(\text{FT}[\mathbf{T}(\mathbf{t})])$, in the 2D frequency-domain ($\mathbf{f} = (f_1, f_2)$) is purely real following a 2D complex discrete Fourier transform FT and a known phase correction Ph (see Supplementary Information). As explained below, forcing the spectrum's phase to be real strengthens the $\hat{P}_1$ mask used in our iterated maps approach, but the sharper spectral features that result from this initial step will also help with other approaches (Zhu and Bax 1990; Mayzel et al. 2014; Ying et al. 2017). For all of the 2D spectra in this paper, $M \times N = 256 \times 4098$, where $M = 2N_{t_1}^{\text{dense}}$ and $N_{t_1}^{\text{dense}} = 128$. In addition to specifying a point in the spectrum by the vector of frequencies $\mathbf{f} = (f_1, f_2)$, we can also use the corresponding indices m and n, which go from 0 to $M - 1$ and $N - 1$, respectively.

To simulate the effect of skipping particular experiments (due to NUS of the indirect dimension in a "States"-like experiment), we construct an initial "sparsely-sampled" data set $\mathbf{S}^0(\mathbf{t}) = \hat{P}_0 \mathbf{T}(\mathbf{t})$, where $\hat{P}_0$ is a projection that sets to zero all skipped points, leaving the rest alone (see Supplementary Information). Thus, for the $t_1 \geq 0$ rows of $\mathbf{S}^0(\mathbf{t})$, the number containing data is N_{t_1}, and the number filled with zeros is $N_{t_1}^{\text{dense}} - N_{t_1}$, where $N_{t_1} \leq N_{t_1}^{\text{dense}}$. Moreover, the pattern of zero-filled rows is symmetric about $t_1 = 0$. For a given N_{t_1},

the choice of which particular t_1 rows are "measured" (i.e., left alone by $\widehat{P}_0$) is specified by the NUS sampling schedule. Many of the popular NUS schedules, such as Poisson Gap sampling or exponentially-biased sampling, sample t_1 rows more densely near $t_1 = 0$. However, since the iterated maps method appears to work better when the gaps between sampled points are kept to a minimum size, we prefer to use NUS schedules that spread the sparsely-sampled points out as uniformly as possible across the densely-sampled grid. In Frey et al. (2013), we used a QUEST schedule ("QUasi-Even Sampling, plus jiTter"). In this paper, we turn off the small random jitter, and use *deterministic* quasi-even sampling, for reasons we explain below (see *Reconstruction error source I: signal inside the mask*).

As N_{t_1} is lowered below $N_{t_1}^{\text{dense}}$, the resulting sparse spectrum $\tilde{\mathbf{S}}^0(\mathbf{f}) = \text{Ph}(\text{FT}[\mathbf{S}^0(\mathbf{t})])$ becomes a poor approximation of the dense spectrum $\tilde{\mathbf{T}}(\mathbf{f})$, with pronounced sparse-sampling artifacts that smear along the f_2 columns (see Supplementary Information). To do better than this, we use iterated maps to deterministically convert $\tilde{\mathbf{S}}^0(\mathbf{f}) \to \tilde{\mathbf{F}}^i(\mathbf{f}) \approx \tilde{\mathbf{T}}(\mathbf{f})$.

The iterated maps method uses two projection operators, one in the frequency domain and one in the time domain, which enforce things that we know to be true about the dense data, similar to SIFT (Matsuki et al. 2009, 2010, 2011). In the frequency domain, the $\widehat{P}_1$ projection makes use of a predefined "mask function" mask($\mathbf{f}$). $\widehat{P}_1$ sets to zero any points in the mask($\mathbf{f}$) = 0 region, where we know that no signal exists. In addition, we exploit our knowledge of the signal's phase and sign within the non-zero support mask regions to further strengthen $\widehat{P}_1$. Specifically, all imaginary parts are set to zero, and the negative (positive) real parts are set to zero where mask($\mathbf{f}$) = +1 (−1). Real signals of either sign are left alone in the "artifact regions," where mask($\mathbf{f}$) = 2. The values of mask($\mathbf{f}$) can either be defined manually, or using an algorithm like this:

- mask($\mathbf{f}$) = 0   if $|\text{Re}(\tilde{T}(\mathbf{f}))| < \tau$
- mask($\mathbf{f}$) = +1   if $\text{Re}(\tilde{T}(\mathbf{f})) \geq \tau$
- mask($\mathbf{f}$) = −1   if $\text{Re}(\tilde{T}(\mathbf{f})) \leq -\tau$
- mask($\mathbf{f}$) = +2   if the sign of $\text{Re}(\tilde{T}(\mathbf{f}))$ is uncertain, and $|\text{Re}(\tilde{T}(\mathbf{f}))| \geq \tau$

where τ is a positive threshold. In our case, the dense spectra from slices 2–12 of our pseudo-3D data set are similar enough that any one of them can be used to create masks for the others. Here, we use slice 2 to create our mask, and take τ to be ≈ 6 times the noise level of f_2 columns far from any signal or artifacts. We will revisit this choice of τ below (see *Can K_p be lowered arbitrarily by increasing the $\widehat{P}_1$ mask threshold τ?*). We also manually chose to use mask($\mathbf{f}$) = 2 for the signal at the f_1 edges of the spectrum, because that signal is an artifact whose sign is inconsistent across slices.

In the time domain, the $\widehat{P}_2$ projection resets measured points to their known values (i.e., the points in $\mathbf{S}^0(\mathbf{t})$ that were not zeroed by $\widehat{P}_0$). While these projections can be altered slightly to allow for noise-handling (Frey et al. 2013), we are not using noise-handling in this paper. This simplification will be helpful later (see *Reconstruction error source II: signal outside the mask*), since it enables us to develop a quantitative prediction of the size of reconstruction errors that arise from non-zero signal in the mask = 0 region.

Rather than simply alternating between these two projections, we combine them in a particular form of Elser's Difference Map algorithm (Elser 2003; Elser et al. 2007) which reproduces Fienup's hybrid input-output map (Fienup 1982). Our map is given by $\widehat{D} = \mathbb{1} + \widehat{P}_1 \cdot (2\widehat{P}_2 - \mathbb{1}) - \widehat{P}_2$, where $\mathbb{1}$ is the identity operator (Elser 2003; Frey et al. 2013), and the iterative mapping becomes $\mathbf{S}^i = \widehat{D}\mathbf{S}^{i-1}$. Fourier transforms and phase corrections are applied as needed to reversibly convert between time and frequency domains. The final output of the iterated maps algorithm after i iterations is given by $\mathbf{F}^i = \widehat{P}_2\mathbf{S}^i$. Most of the reconstructions we calculated required somewhere between 100 and 10,000 iterations to converge (see Supplementary Information for details). The DiffMap calculations were all performed in Igor Pro 8. We have also ported the code to Python 3, and verified that the Python and Igor results agree to within numerical precision. The Python code is available on GitHub at https://github.com/robbyblum/DiffMap-coDiffMap-Python, and will also be available on the NMRbox platform at http://NMRbox.org (Maciejewski et al. 2017).

Because the FT is a linear transformation, and because sparse-sampling is only used in the indirect dimension, the full 2D spectral reconstruction problem can be treated as a series of independent, 1D spectral reconstructions; i.e., each f_2 column can be reconstructed independently of the others. We took advantage of this fact to speed up our study of how the iterated maps method breaks down at low N_{t_1}. At a given N_{t_1}, we refer to the reconstructed 1D spectrum of a particular f_2 column with index n as $\tilde{\mathbf{F}}_n^{i,N_{t_1}}(f_1)$. Furthermore, if we reconstruct a particular peak at f_1 index m well, then $\tilde{F}_{mn}^{i,N_{t_1}} \approx \tilde{T}_{mn}$. We define the reconstruction error of a point as a function of N_{t_1} as $E_{mn}(N_{t_1}) = \tilde{F}_{mn}^{i,N_{t_1}} - \tilde{T}_{mn}$. For a particular peak p at indices (m_p, n_p), we will refer to the reconstruction error as simply $E_p(N_{t_1})$.

What is a reasonable target to shoot for when using NUS?

In many papers that use an NUS approach, the authors report reaching a low sampling fraction $N_{t_1}/N_{t_1}^{\text{dense}}$ (or its (n-1)-dimensional analog for nD experiments). However, as other authors have pointed out, a spectrum that has only

one narrow feature should be easier to reconstruct than one with multiple peaks, even when both have the same $N_{t_1}^{\mathrm{dense}}$ (Shchukina et al. 2017; Hyberts et al. 2014). Thus, a better metric to use when comparing the performance of NUS approaches is the ratio of the number of sparsely-sampled measurements to the number of significant points to be reconstructed. We can think of the number of significant points as the "information content" of the dense spectrum. We will adopt such a metric below.

Following the notation of Shchukina et al. (2017), we define $K(n)$, the number of significant points (relative to a threshold τ) of the nth f_2 column, as the information content of that 1D spectrum $\tilde{\mathbf{T}}_n(f_1)$. We calculate $K(n)$ by counting all of the mask $= -1, +1$, and $+2$ points in the nth column, and define $K_p = K(n_p)$ for a particular peak p in column n_p. It is important to remember that $K(n)$ is much larger than the total number of peaks in a column, since all real peaks have nonzero width. Because we fill all four quadrants of the time domain using the "States"-like data, the relevant number of sparsely-sampled points to compare to $K(n)$ is $2N_{t_1}$, where N_{t_1} is the number of complex points sampled with $t_1 \geq 0$. Thus, we can characterize the compression achieved by NUS using the ratio $\mathcal{C} \equiv 2N_{t_1}/K(n)$.

Even though most NMR spectra have a higher density of non-zero points than the test images typically studied in the compressed sensing (CS) literature, we can compare our results to CS benchmarks. If we insert our parameters into expressions from the CS literature (see Eq. 9.40 of Foucart and Rauhut 2013), we find that CS approaches are likely to yield high-quality reconstructions when:

$$2N_{t_1} > 2K_p \ln(M/K_p) \quad \textbf{(CS limit)}. \tag{1}$$

For $M = 256$, and an average $\overline{K_p} = 53.5$ (over all 114 peaks), the corresponding CS limit on the achievable compression ratio is $\mathcal{C}_{\mathrm{CS}} > 3.1$. This is already a significant improvement over the corresponding value for dense sampling, $\mathcal{C}_{\mathrm{dense}} = 4.8$. Note that the only assumption made by CS approaches is that the spectrum is "sparse."

We may be able to beat this CS limit if we know where spectral features should not appear, since this constraint provides more information about the spectrum. In fact, our iterated maps approach has the potential to push down to the limit imposed by linear algebra (LA): that the number of "constraints" $M - K_p$ must be greater than or equal to the number of "unknowns" $M - 2N_{t_1}$:

$$2N_{t_1} \geq K_p \quad \textbf{(LA limit)}. \tag{2}$$

To test this prediction, we tried to reconstruct an extremely simple, single-peak spectrum (Fig. 2a) that exactly satisfies the constraints of $\hat{P}_1$ and $\hat{P}_2$. In this ideal limit, as we lower N_{t_1} below $N_{t_1}^{\mathrm{dense}}$, there is essentially zero error in the reconstruction of the peak, all the way down to the ideal LA limit

of $\mathcal{C} = 1$ (Fig. 2b). Thus, $\mathcal{C} = 1$ seems to be a reasonable compression target to shoot for with this method.

Features of a realistic spectrum that challenge the iterated maps approach

Compared to the ideal case, a realistic 1D NMR spectrum has many additional features (e.g., Fig. 2c). The most obvious example is the presence of other peaks, which can be either positive or negative in sign, with their own corresponding mask regions (mask($\mathbf{f}$) $= +1$ or -1, respectively). There also can be artifact regions, which are usually unwanted spectral features that do not have a predictable sign, but which must be accounted for (mask($\mathbf{f}$) $= 2$). Finally, even in the regions "without signal" (mask($\mathbf{f}$) $= 0$), any real spectrum will have random noise, as well as spectral tails that fall below the mask cutoff τ. All of these features make it harder for the iterated maps method to work as well as it does in the ideal case. This can be seen by comparing Fig. 2b–d: in the latter, even well above the $\mathcal{C} = 1$ limit, the peak reconstruction error $E_p(N_{t_1})$ is clearly non-zero, and it can swing wildly over narrow regions of N_{t_1}. In fact, these trends are even easier to see in Fig. 3, where we plot the reconstruction errors for the same peak used in Fig. 2c, d (Leu50 δ^2) but for all eleven slices of the pseudo-3D data set. Minor variations in the dense spectrum across different slices lead to large spreads in $E_p(N_{t_1})$ in certain regions of N_{t_1}, but not in others.

Methods

In this section, we will discuss how the reconstruction error arises from two main categories of complications:

I Signal inside the mask (mask($\mathbf{f}$) $\neq 0$)
II Signal outside the mask (mask($\mathbf{f}$) $= 0$)

In each case, we will present a metric that is able to describe the reconstruction error using available information. Category I causes the largest errors, so we discuss this first.

Reconstruction error source I: signal inside the mask

As pointed out by Maciejewski et al. (2009), any NUS schedule leads to intraband aliases of all spectral peaks. These intraband alias features appear to move along the f_1 axis as N_{t_1} is varied. As we studied the performance of our iterated maps method we noted a clear pattern: as N_{t_1} is lowered, the biggest reconstruction errors for a given peak happen when there is a collision between its larger intraband alias features and the nonzero mask regions, or "mask collisions" for short. Qualitatively, at these N_{t_1} values, $\hat{P}_2$ mixes up the
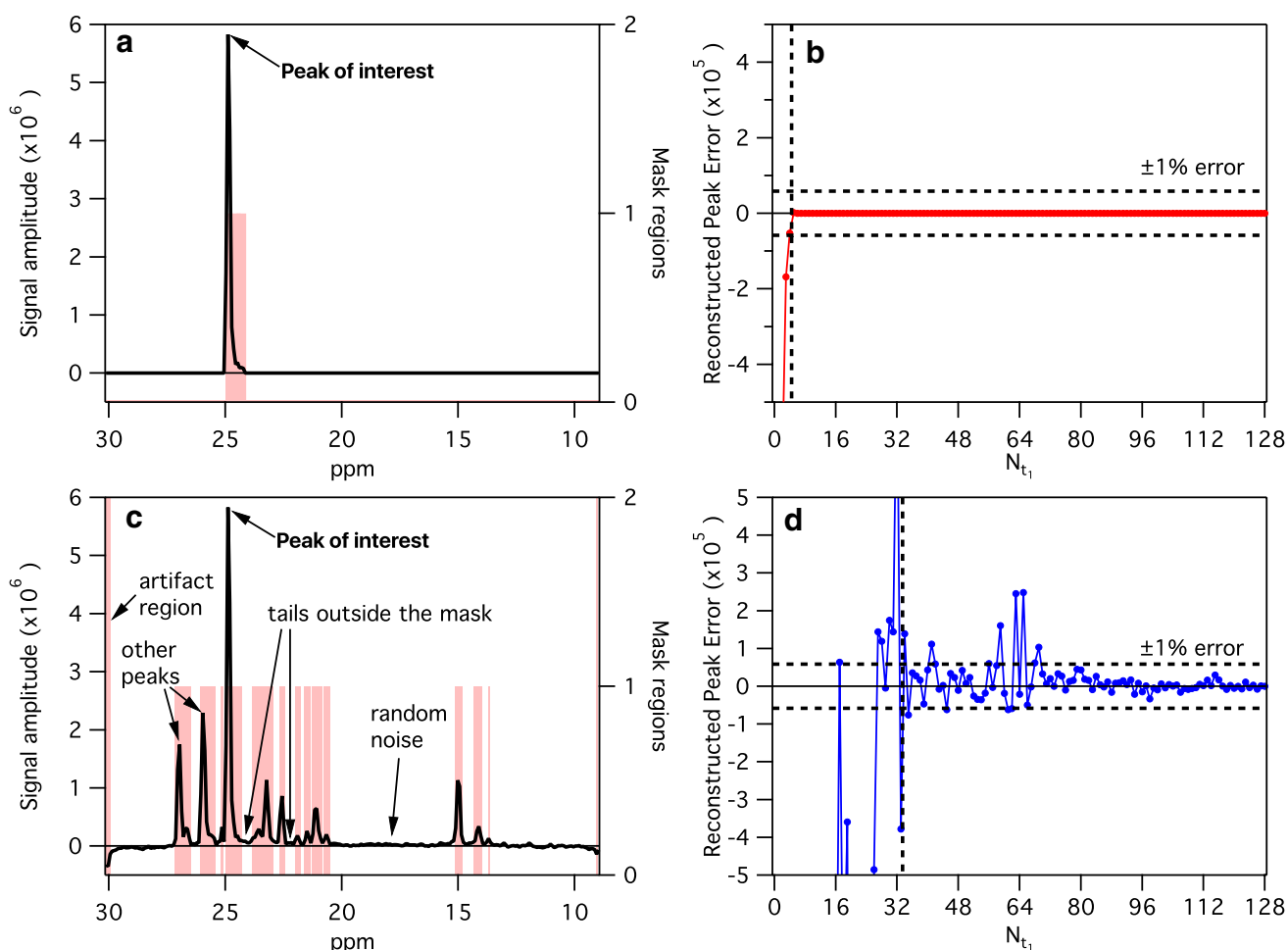
Fig. 2 a "Idealized" spectrum, made by setting most of an actually-acquired spectrum, and its mask, to 0. **b** Error of the reconstructed peak amplitude for the "ideal" peak of interest shown in **a**. Vertical dashed line indicates $K_p/2$. The error is zero (within machine precision) until $N_{t_1} < K_p/2$. **c** The actually-acquired spectrum (for peak Leu50 δ^2, $i_{\tau_{cp}} = 2$) that was simplified in **a**, indicating the important features. **d** Error of the reconstructed peak amplitude for peak Leu50

δ^2, $i_{\tau_{cp}} = 2$. Many complicated features appear that did not arise in the "ideal" case. Notice that K_p has increased because there are more features in the real spectrum than in the simplified spectrum, which increases the size of the non-zero mask region. Note that in **b** and **d** we are graphing the absolute difference between the reconstructed and dense amplitudes of the peak of interest. We use this convention throughout the paper

amplitudes at frequency points connected by intraband aliases, AND $\widehat{P}_1$ is unable to resolve the confusion. Mask collisions seem to matter for any nonzero mask(**f**) value (that is, $+1$, -1, or 2), so we will treat all these cases as being equivalent. The realization that mask collisions cause the worst errors motivated us to switch to a quasi-even sampling schedule.

To understand this further, recall that a non-uniform sampling schedule in the time domain has a corresponding point spread function (PSF) in the frequency domain (see Supplementary Information). The PSF determines how undersampling causes spectral information to be mixed among distinct f_1 points, which results in the intraband aliases. A periodic sampling pattern results in a coherent PSF with sharp sidelobe features, separated by nulls. A randomized

sampling pattern results in an incoherent PSF, smearing the sidelobes, and filling the nulls (Maciejewski et al. 2009). Most NUS reconstruction approaches prefer the less coherent PSFs of sampling schedules with irregular spacings, because they tend to broaden and scramble the intraband alias features, which helps to reduce the extreme errors encountered as N_{t_1} is lowered. Unfortunately, this broadening tends to fill in the nulls in the PSF, which increases the errors elsewhere. Nevertheless, this is a safe strategy that enables N_{t_1} to be picked somewhat arbitrarily, as it smooths out the extreme errors in reconstruction.

However, because we are trying to push so aggressively to the lowest possible N_{t_1} values, we will use the opposite strategy. Namely, we turn off the random jitter of QUEST sampling (Frey et al. 2013), in order to implement *deterministic*
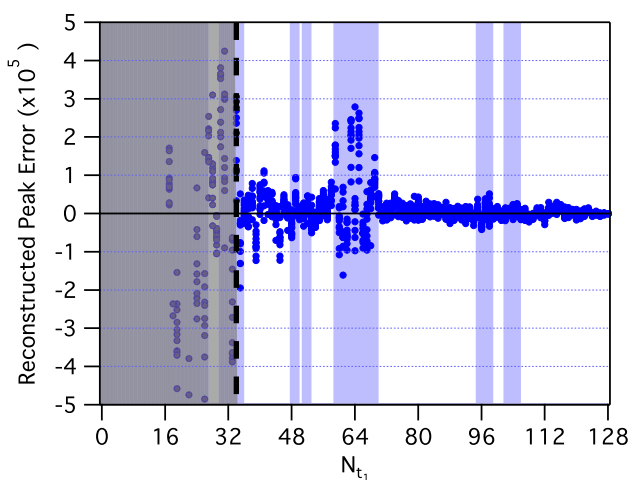
Fig. 3 Error in the reconstructed amplitude for Leu50 δ^2, shown for all 11 slices at each N_{t_1}. The mask-collision metric (blue shading) indicates regions where significant errors are predicted to be possible (see description of "MCM" in main text). The linear algebra limit $K_p/2$ is indicated by the dashed line, below which errors grow very quickly (gray shading). The dense peak amplitude for L50δ2 (averaged over all 11 data slices) is 58.4×10^5, and the mask cutoff is $\tau = 0.9 \times 10^5$
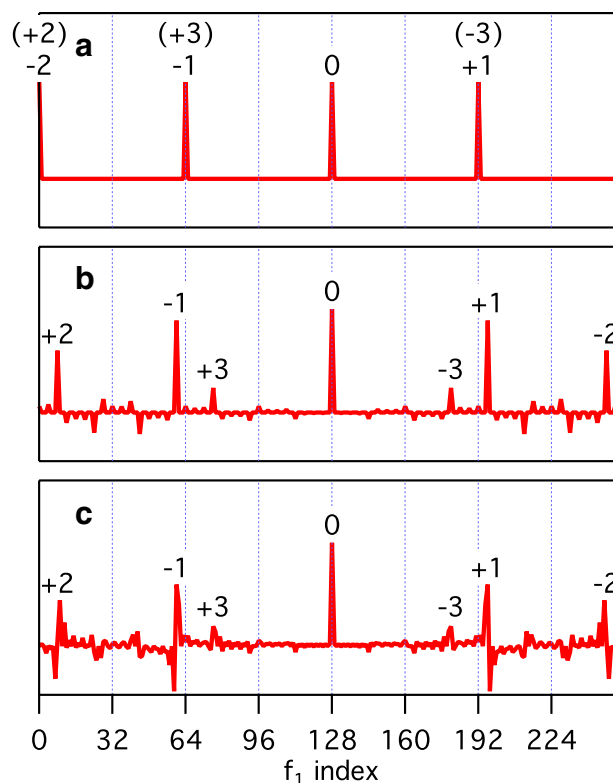

Fig. 4 a Point-spread function of a purely-even undersampling pattern for $N_{t_1} = 32$ and $t_1^{\text{off}} = 0$ (real part). **b** Point-spread function of a quasi-even sampling pattern for $N_{t_1} = 34$ and $t_1^{\text{off}} = 0$ (real part). More harmonics become differentiated. **c** Point-spread function of a sampling pattern for $N_{t_1} = 34$ using quasi-even sampling and $t_1^{\text{off}} = \delta t_1/2$ (real part). In all three plots, harmonics $\pm h = 0, 1, 2, 3$ are indicated

quasi-even sampling, since this sharpens the sidelobes of the corresponding PSF (see Supplementary Information). As N_{t_1} is lowered, this will increase the reconstruction errors due to mask collisions. However, these larger errors will now be restricted to smaller regions of N_{t_1}, which makes it easier to avoid them (see Fig. 3). At the same time, quasi-even sampling will also *decrease* the errors in the absence of mask collisions. As long as we can predict where the small errors will occur as a function of N_{t_1}, this aggressive strategy should result in the best possible reconstructions of the peak of interest.

Given a mask, we can identify potentially dangerous N_{t_1} values for a particular f_1 peak in a particular f_2 column (e.g., the blue bands in Fig. 3 show the so-called danger zones). If we want to achieve a high-quality reconstruction of that peak, we should only use N_{t_1} values in the complementary safe zones (e.g., the white bands in Fig. 3). We describe how we locate the danger zones next.

Finding the N_{t_1} danger zones for a peak of interest

We expect the largest reconstruction errors for a given peak to occur in the danger zones of N_{t_1}, i.e., when significant intraband aliases of that peak collide with the mask. For *a single peak of interest*, we first need to know where its intraband aliases end up at a given N_{t_1}. We then use these intraband alias locations and the $\hat{P}_1$ mask function to define our mask-collision metric (MCM).

For a given N_{t_1}, the $t_1 \geq 0$ points in all sampling schedules discussed here can be described with a single expression:

$$t_1 \in \left\{ t_1^{\text{off}} + \delta t_1 \text{Round}\left(\frac{N_{t_1}^{\text{dense}}}{N_{t_1}} k \right) \right\}, \qquad (3)$$

where $k \in \{0, 1, \ldots, N_{t_1} - 1\}$, δt_1 is the dwell time in the indirect dimension, and t_1^{off} can either be 0 or $\delta t_1/2$. The case $N_{t_1}^{\text{dense}}/N_{t_1} = 1$ is what we mean by dense sampling. A given $t_1 \geq 0$ sampling schedule provides the full sampling vector $s(t_1)$ of length M to match our four-quadrant data, which in turn provides the PSF after a Fourier transform (see Supplementary Information for details) (Maciejewski et al. 2009).

We start by considering the simplest undersampled case: even undersampling (where $N_{t_1}^{\text{dense}}/N_{t_1}$ is an integer greater than one) with $t_1^{\text{off}} = 0$. For this case, sidelobes in the PSF appear at integer multiples of the sparse bandwidth ($MN_{t_1}/N_{t_1}^{\text{dense}}$) away from the central peak at $m_0 = M/2$,
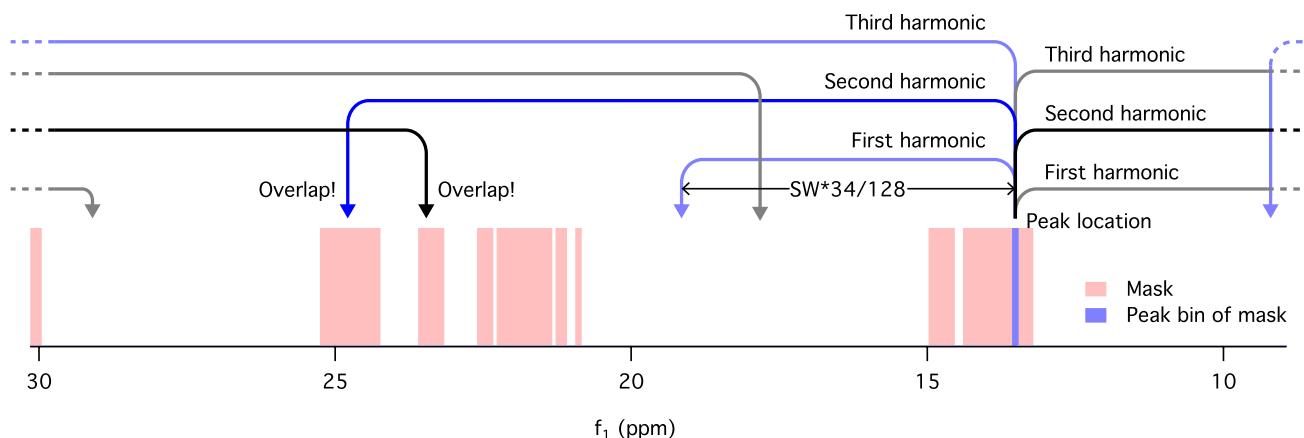
$$m_{\pm h} = m_0 \pm 2N_{t_1}h \mod M, \qquad (4)$$

Fig. 5 Schematic example of a mask collision metric calculation in a 1D spectrum, which requires only the peak-of-interest location (blue shading) and the $\widehat{P}_1$ support mask (pink shading). We wish to check for mask collisions for $N_{t_1}/N_{t_1}^{\text{dense}} = 34/128$. In this case, $h_{\max} = \text{Floor}(3.11) = 3$ (see Eq. (6)), so we need to check three harmonics. We start at the location of the peak of interest ($m_p = 200$) and go out $(SW*34/128)*h$ ($h = 1, 2, 3$) in both directions, where

SW is the sweep width. If this takes us off the edge of the spectrum, we wrap around to the far side of the spectrum. We then check each of the calculated locations for nonzero mask values. In this case, we find nonzero mask values at both the location of the second "up" harmonic (blue arrow) and that of the second "down" harmonic (black arrow), so $\text{Hit}(200, 2, 34) = 2$ (see Eq. (5)), and the Boolean $\text{MCM}(200, 34) = 1$ (see Eq. (7))

where this defines the harmonic number $h = 0, 1, 2, \ldots$. For instance, in Fig. 4a, $N_{t_1} = 32$, so each harmonic is 64 points away from the previous one. If $m_0 \pm 2N_{t_1}h$ falls outside the spectral width, the modulus operation ("mod M") brings it back in by "wrapping it around" to the far side of the window. Our earlier discussion of intraband aliasing (Frey et al. 2013, *S.I.*) included the $\pm 2N_{t_1}h$ term, but we *failed to consider the wraparounds*, which will turn out to be important in what follows.

Truly even sampling is only possible for a few N_{t_1}, so we are usually forced to approximate it with quasi-even sampling ($N_{t_1}^{\text{dense}}/N_{t_1}$ is not an integer, so the rounding in Eq. (3) matters). The PSF for one such case, $N_{t_1} = 34$, is shown in Fig. 4b. While the number of nonzero points is much greater in this case, the locations of the sidelobes are still given by the same $m_{\pm h}$ formula as before.

Finally, for our data set, we do not actually have $t_1^{\text{off}} = 0$; instead, we have $t_1^{\text{off}} = \delta t_1/2$. This complication broadens the absorptive sidelobes of the PSF, making them appear somewhat more dispersive. Fig. 4c shows this effect at $N_{t_1} = 34$. Because the harmonic sidelobes are no longer delta functions at exactly $m_{\pm h}$, we account for their nonzero width below.

Equation (4) gives us the location of the PSF sidelobes, but we are trying to find the location of the intraband aliases. The PSF, when convolved with a dense spectrum, generates a sparse spectrum. Thus, adapting Eq. (4), we know that the intraband aliases of a particular peak of interest at m_p will occur at $(m_p \pm 2N_{t_1}h \mod M)$ (i.e., replace m_0 with m_p in Eq. (4)). From here, to see if an intraband alias is colliding

with the mask, all we need to do is check whether the mask function (in the 1D column of interest) is nonzero at these locations. That is, for a given harmonic order h, we define a "hit function"

$$\text{Hit}(m_p, h, N_{t_1}) = \big(\text{mask}(m_p + 2N_{t_1}h \quad \mod M) \neq 0\big)$$
$$+ \big(\text{mask}(m_p - 2N_{t_1}h \quad \mod M) \neq 0\big).$$
$$(5)$$

There are technically an infinite number of intraband alias harmonics, but in practice, a given harmonic's mask collisions do not seem to noticeably impact the quality of the DiffMap reconstruction until you get to a low enough N_{t_1}, which seems to be proportional to $N_{t_1}^{\text{dense}}/h$. Thus, for a given N_{t_1}, we establish a maximum harmonic number to check for collisions, $h_{\max}$, given by

$$h_{\max} = \text{Floor}\left(\beta \frac{N_{t_1}^{\text{dense}}}{N_{t_1}}\right), \quad (6)$$

where we used $\beta \approx 0.82625$ in this study (see the next section for further discussion).

Putting this all together, we are now able to define the Boolean mask collision metric (MCM). For a peak p at index m_p, $\text{MCM}(m_p, N_{t_1})$ predicts which N_{t_1} are likely to result in a problematic collision between aliases and the mask:

$$\text{MCM}(m_p, N_{t_1}) = \left(\left[\sum_{h=1}^{h_{\max}+1}\sum_{i=-1}^{1} \text{Hit}(m_p + i, h, N_{t_1})\right] > 0\right).$$
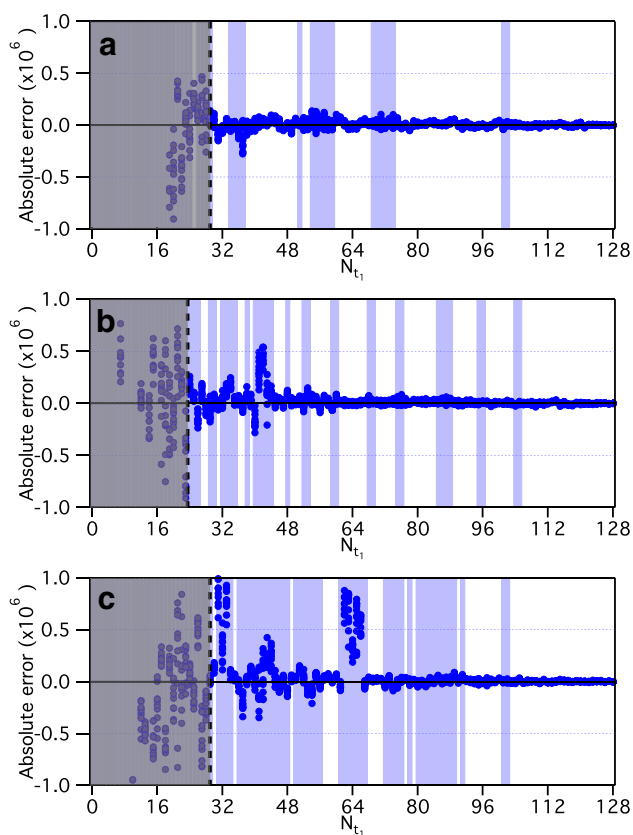$$(7)$$

Fig. 6 Examples of the MCM for three peaks, with reconstruction results. **a** Peak Val56 γ^2, with average dense amplitude 1.46×10^7. **b** Peak Val18 γ^1, with average dense amplitude 5.72×10^6. **c** Peak Ile52 δ^1, with average dense amplitude 5.66×10^6. In each graph, the blue shading indicates which N_{t_1} the metric flags as potentially bad, and the dots represent the error in the reconstructed peak amplitude $E_p(N_{t_1})$, for all 11 slices of the pseudo-3D data set. The black dashed line indicates the linear algebra limit $K_p/2$ for the column

Note that the sum over $i = -1, 0, +1$ is how we are accounting for the nonzero width of the harmonic sidelobes for the $t_1^{\text{off}} = \delta t_1 / 2$ case.

A schematic example of how this calculation works is shown in Fig. 5. The MCM divides the set of possible N_{t_1} into two categories: either a danger zone or a safe zone. Returning to Fig. 3, the blue bars show the danger zones, where $\text{MCM}(m_p, N_{t_1}) = 1$; all other N_{t_1} values are in safe zones, where $\text{MCM}(m_p, N_{t_1}) = 0$. Note that the MCM can be calculated as soon as the mask has been defined.

Testing the predictive power of the mask collision metric

Now that we have defined the mask collision metric, we can easily apply it to all 114 peaks in our spectrum. For example, the danger zones for three additional peaks are shown in Fig. 6. The blue points show the absolute error in

the peak reconstructions, calculated for all 11 slices of the pseudo-3D data set.

We can see some patterns start to emerge across Figs. 3 and 6. First, the reconstruction quality becomes terrible once N_{t_1} goes below the linear algebra limit, $K_p/2$ ($\mathcal{C} = 1$). This makes sense, because in that region we do not expect to have enough information to solve the reconstruction problem uniquely.

Second, above $\mathcal{C} = 1$, the metric does a very good job of locating dangerous N_{t_1} values which result in large reconstruction errors, which often appear as large variances in the error between the slices.

Third, above $\mathcal{C} = 1$, we can see that the mask collision metric is somewhat conservative: it has more false positives than false negatives. For experimentalists, false negatives (predicted safe zones of N_{t_1} that nonetheless lead to bad reconstruction) would be the more significant problem, so we tried to construct the MCM so as to make them as rare as possible. As we explained in the last section, our MCM checks 3 adjacent points at each intraband alias location, which was the most systematic way we found to eliminate almost all false negatives without adding enormous numbers of false positives. We also wanted to limit false positives (predicted danger zones of N_{t_1} that turn out to have good reconstructions), since they exclude more N_{t_1} values than necessary. This motivated how we set β for Eq. (6). In particular, we were able to eliminate a large number of trivial false positives at very high N_{t_1}. There are still remaining false positives which seem very hard to exclude, at least within our current understanding of the method.

In principle, β can be any positive number, but in our model, we have restricted it to the range $0 < \beta < 1$. To choose a value for β, we used a procedure based on the peak-to-sidelobe ratio of the PSF for a given harmonic (see Supplementary Information for details). We were able to tune β to improve the accuracy of the MCM for our data, but we do not know how much it might need to be adjusted for other applications. In general, it is prudent to err on the side of high β, as lowering it too far will introduce an increasing number of false negatives. It is actually perfectly acceptable to use a default value of $\beta = 0.99$ (see Supplementary Information for discussion of why we limit $\beta < 1$).

Before leaving this section, it is worth emphasizing two points. First, as we have defined it so far, $\text{MCM}(m_p, N_{t_1})$ is only predicting dangerous and safe N_{t_1} values for the reconstruction of a *single* peak of interest at f_1 index m_p; it says nothing about reconstruction error elsewhere in the spectrum. For example, as Fig. 7 shows, good (bad) reconstruction performance at m_p can be accompanied by bad (good) performance elsewhere. Second, while the magnitude of the peak of interest at f_1 index m_p has to be larger than our mask threshold τ, that is all that we have assumed
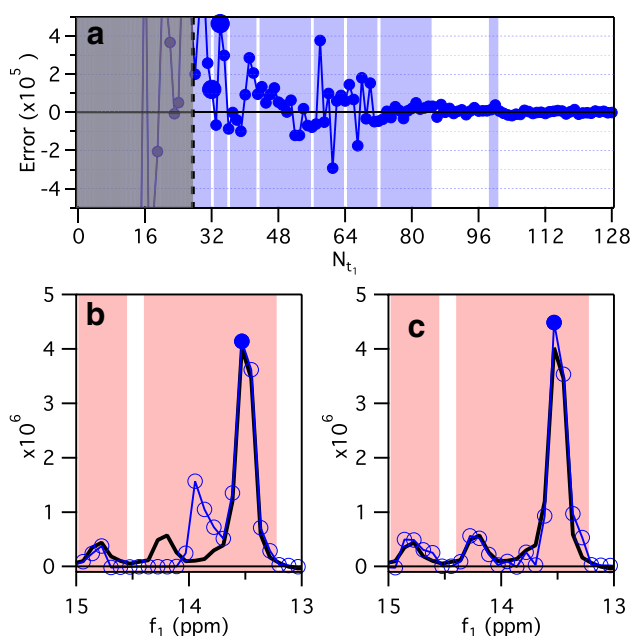
Fig. 7 Reconstruction performance examined in detail. **a** Reconstruction error in the amplitude of the Ile73 δ^1 peak (dots), with danger zones predicted by the MCM (blue shaded region) and the limit $K_p/2$ (dashed line). The results for $N_{t_1} = 32$ and $N_{t_1} = 34$ (larger dots) are examined in more detail in **b**, **c**, respectively. **b** "Good" reconstruction example, $N_{t_1} = 32$. This is a good result despite the significant errors elsewhere since we are only interested in the peak Ile73 δ^1 (filled circle). **c** "Bad" reconstruction example, $N_{t_1} = 34$. Most of the spectrum is reconstructed well, but this is a bad result since there is a noticeable error (more than 10%) at the location of the Ile73 δ^1 peak (filled circle)

to this point. It does not need to be an actual peak, such as a local maximum in the f_2 column, or a peak picked in the 2D spectrum.

Reconstruction error source II: signal outside the mask

As the last section demonstrates, the metric is good at identifying danger zones, or N_{t_1} values where the largest reconstruction errors can occur, due to significant intraband aliases colliding with the mask. The metric also picks out N_{t_1} values that fall in the complementary safe zones.

To understand DiffMap's performance in the safe zones, we need to take into account a second, smaller source of reconstruction error. Our iterated maps method assumes that the mask we use in $\widehat{P}_1$ is correct; that is, there is *NO* signal outside of the support regions. This is just an approximation, since all real NMR spectra have noise. For the 2D data sets under study, the noise is very small compared to the peaks, but it is still nonzero. Furthermore, by making a mask using a threshold τ applied to a dense data set, you basically guarantee that there will be at least some tails of your peaks

that fall outside the mask. Noise fluctuations and tails are actual spectral features of the dense spectrum, just like the significant peaks are. DiffMap treats them the same way as it does the large peaks, with one important distinction: the true location of these outside-the-mask features is *disallowed* by $\widehat{P}_1$. The $\widehat{P}_2$ operation will try to put them back in the right spot, and for large N_{t_1}, the final $\widehat{P}_2$ will mostly fix these features, but $\widehat{P}_1$ will always be trying to drive them out of the mask $= 0$ regions. So where do they end up in the spectrum? These features have intraband aliases just like the large peaks do, so for a given N_{t_1}, they will go into the nonzero mask regions which collide with their intraband aliases. Thus, *the signal outside the mask gets folded into the mask*.

To describe the impact of signal from outside the mask on the reconstruction of points inside the mask, we will use a model in which the spectral features at different points outside the mask are drawn from a normal distribution with mean μ_v and standard deviation σ_v. Note that this is not really a correct description of this signal, because random electronic noise and spectral tails are somewhat different in character. Noise is incoherent, has a mean of zero, and is present everywhere in the spectrum. Spectral tails are coherent, usually positive-amplitude, and localized; in the context of a cutoff-generated mask, they will generally be found just outside the support regions. However, we have found that applying this model with μ_v and σ_v drawn from the calculated mean and standard deviation of the entire mask $= 0$ region in a given spectrum still gives us accurate bounding behavior for E_p. The main difference we see is that the fluctuations end up being more correlated in N_{t_1} than they would be in a purely random case.

Here we briefly describe the results of our model. The full derivations are in the Supplementary Information. We consider a spectrum with a single delta-function peak of amplitude A_{dense} at point m_p, with a corresponding single non-zero mask point. For simplicity, we assume we are in the $t_1^{\text{off}} = 0$ case; simulations of the $t_1^{\text{off}} = \delta_1/2$ case do not show a significant change in behavior. We call the reconstructed amplitude A_{recon}, and calculate the mean μ_E and the standard deviation σ_E of the peak reconstruction error $E \equiv A_{\text{recon}} - A_{\text{dense}}$.

By considering the signal outside the mask in terms of its mean and standard deviation, we can consider the effect of each part separately. To calculate μ_E, we consider a delta-function spectrum with no noise and a DC offset μ_v. Then, the mean error is

$$\mu_E(N_{t_1}) = \mu_v \left(\frac{N_{t_1}^{\text{dense}}}{N_{t_1}} - 1 \right). \tag{8}$$
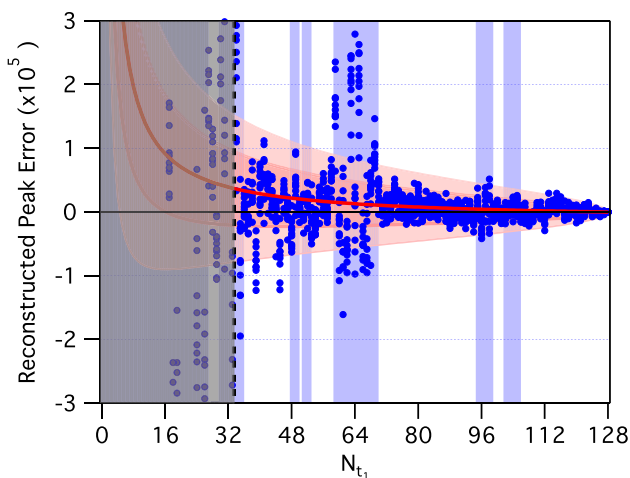
Fig. 8 E_{SZ} prediction bands for peak Leu50 δ^2. $\mu_E(N_{t_1})$ is plotted as the red line, the $\pm 1\sigma_E$ band is the darker red shading, and the $\pm 2\sigma_E$ band is the light red shading. As in Fig. 3, the danger zones of the MCM are plotted as blue bars, and $E_p(N_{t_1})$ for all 11 data sets are plotted as blue dots

To calculate the standard deviation of the error σ_E, we again consider a delta-function spectrum, but now with added noise that has zero mean and standard deviation σ_ν. From this, we find that the standard deviation of the error is

$$\sigma_E(N_{t_1}) = \sigma_\nu \sqrt{\frac{N_{t_1}^{\text{dense}}}{N_{t_1}} - 1} \,. \tag{9}$$

Combining Eqs. 8 and 9, we predict that 95% of the errors for N_{t_1} in the safe zones will lie within the range

$$E_{SZ}(N_{t_1}) = \mu_E(N_{t_1}) \pm 2\sigma_E(N_{t_1}). \tag{10}$$

If we plot this prediction band against the actual reconstruction error for various peaks, we find that it does a very good job of explaining the observed error in the safe zones (see Fig. 8, as well as additional examples in the Supplementary Information). For the N_{t_1} values in the danger zones, the reconstruction error can push well outside this prediction band, which is exactly what we expect.

It is worth pointing out that we have accurately predicted the reconstruction error for all of the 11 sparsely-sampled slices of the pseudo-3D data set, using just the information from a single densely-sampled slice (slice 2). From this spectrum, we were able to calculate the mask, the MCM, and the E_{SZ} prediction bands.
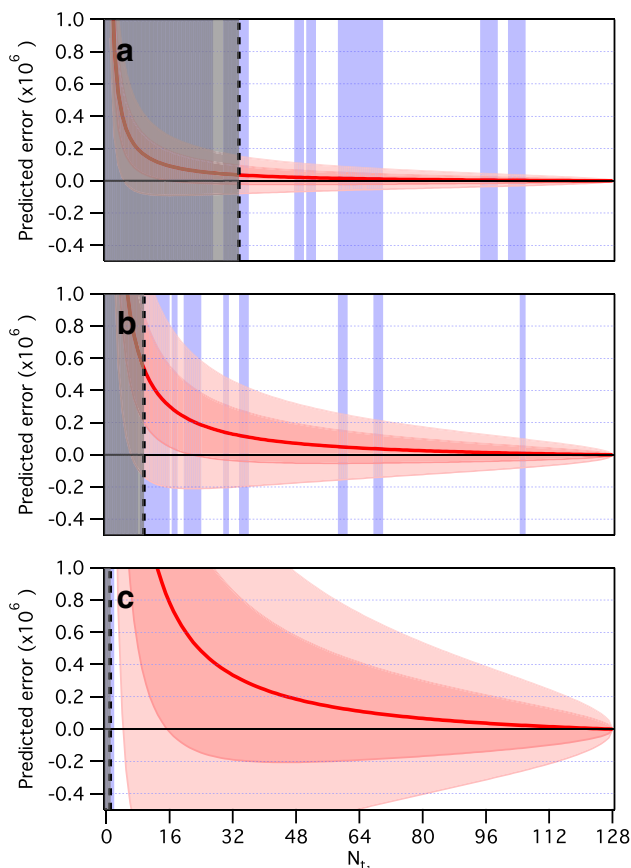

Fig. 9 Metrics for different choices of the mask cutoff τ. **a** $\tau = 90,000$. For a lower mask cutoff, the nonzero mask region is larger (covers more features), increasing the number of alias overlaps with the mask, and decreasing the amount of "noise" (including tails) outside the mask. The linear algebra limit $K_p/2$ is higher (dashed line) because the $\hat{P}_1$ operator can impose fewer constraints. **b** $\tau = 500,000$. Raising the mask cutoff decreases the number of alias overlaps, increases the amount of "noise" outside the mask, and lowers the $K_p/2$ threshold. **c** $\tau = 2,900,000$. Dramatically increasing the mask cutoff unreasonably removes almost all danger zones, sharply increasing the expected errors elsewhere. The dense peak amplitude is approximately 5.8×10^6

Can K_p be lowered arbitrarily by increasing the $\hat{P}_1$ mask threshold τ?

Yes, but the reconstruction error tends to grow as well. The amount of systematic shift of the reconstruction error in the safe zones depends on how you construct your mask. If you set a higher positive cutoff τ, more tails will fall outside the mask. This tends to increase μ_ν and σ_ν of the points outside the mask, which in turn increase μ_E and σ_E. If you set a lower cutoff, the opposite will happen. Thus, choosing a mask involves making trade-offs: a tighter mask (higher τ) will lower K_p and produce more safe zones (fewer intraband alias-mask collisions), but will have more error in those safe
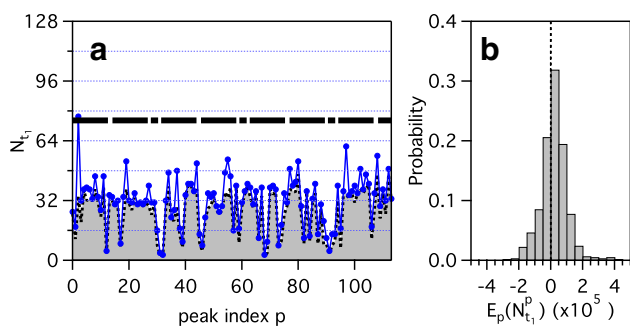
Fig. 10 Performance of the overall metric. **a** $N_{t_1}^p$ for each peak in the spectrum (blue dots), compared to the linear algebra limit $K_p/2$. The limit $L/2 = 75$ achieved in the prior work (Frey et al. 2013) is also shown (dot-dashed line). **b** Histogram of the reconstructed peak errors for all 114 peaks in the spectrum, and all 11 2D slices from the pseudo-3D data set. The histogram shows a slight positive mean, which is anticipated by the fact that the "noise" outside the mask has positive mean μ_E, due to the tails of the actual features. Most of the errors lie within one mask cutoff of zero ($\pm$ 90,000)

Fig. 11 Comparison between $2N_{t_1}^p$ and K_p using our metric. **a** The "bottom-up" compression $\mathcal{C}_{bu} = 2N_{t_1}^p/K_p$ for each peak (red points). The average $\overline{\mathcal{C}_{bu}}$ is 1.25 (blue horizontal line), and the standard deviation is 0.3. The dashed lines show $\mathcal{C}_{dense} = 4.8$, $\mathcal{C}_{CS} = 3.1$, and $\mathcal{C}_{td} = 2.8$. **b** Scatter plot of $2N_{t_1}^p$ vs. K_p for all peaks. A linear fit gives a line with slope 1.043 ± 0.046 and intercept 3.80 ± 1.34 (solid blue line)

zones. A wider mask (lower τ) will have higher K_p and fewer safe zones (more intraband alias-mask collisions), but will have tighter prediction bands. Figure 9 shows the effect of choosing different mask thresholds τ on K_p, the mask collision metric, and the E_{SZ} prediction bands for the peak of interest shown in Fig. 2c. Going from Fig. 9a to 9c lowers the minimum "safe" N_{t_1} for this peak by a factor of 12, but *increases* the mean *and* width of the prediction band at that lower N_{t_1} by a factor of 36.

In some applications, using a tight mask may be preferred, e.g., when the speed-up due to the drop in K_p is much more important than the associated increase in the size of the prediction band.

We actually began our current study with a larger τ inherited from our previous work, but as we came to understand this issue better, we decided that lowering τ to 6 times the noise level enabled us to reach the lowest possible N_{t_1} while maintaining acceptable reconstruction error.

Results

Recipe for finding $N_{t_1}^p$, and overall reconstruction performance

The success of the alias overlap metric for $N_{t_1} \geq K_p/2$ suggests a simple recipe for finding the lowest N_{t_1} that will give an acceptable reconstruction of a given peak. You simply start at $N_{t_1} = K_p/2$, and step up in N_{t_1} until you find the lowest value in a safe zone of the MCM, which we call $N_{t_1}^p$ for peak p. Alternatively, the recipe can be expressed as
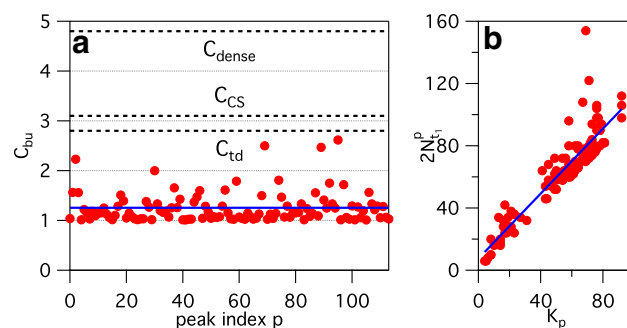
$$N_{t_1}^p = \min(N_{t_1}) \text{ for } \begin{cases} N_{t_1} > K_p/2 \\ \mathrm{MCM}(m_p, N_{t_1}) = 0 \end{cases}. \tag{11}$$

Figure 10a shows the values of $N_{t_1}^p$ found by applying this recipe for each assigned peak in the spectrum (blue points). Many of them are very close to $K_p/2$ (gray region), and only one of them is above $L / 2$ (thick dashed line). Using these $N_{t_1}^p$ values, we build up a histogram of the absolute reconstruction error $E_p(N_{t_1})$ across all 114 peaks, for all 11 slices, as shown in Fig. 10b. Most of the errors are smaller in magnitude than one mask cutoff (90,000), and 88% of the reconstructed peak amplitudes are within 5% of their dense values. We also see a slight positive offset of the center of the distribution, as we expect due to the presence of net-positive tails outside the mask.

Figure 10a plots the $N_{t_1}^p$ values in absolute terms, which naturally suggests a sampling fraction. However, this is not an especially useful metric for comparing different reconstruction methods, because of how much the sampling fraction depends on the original sparseness of the test spectra (Shchukina et al. 2017; Hyberts et al. 2014). Thus, we also plot the "bottom-up" ratio $\mathcal{C}_{bu} = 2N_{t_1}^p/K_p$ in Fig. 11a, as well as plotting $2N_{t_1}^p$ vs. K_p in Fig. 11b. From these plots, we see that even for complicated spectra, we have been able to push down to an average $\overline{\mathcal{C}_{bu}} = 1.25$, which is almost all the way down to our expected limit of $\mathcal{C} = 1$ (Eq. 2).

To further benchmark these results, we compare $\mathcal{C}_{bu}$ to values for dense sampling $\mathcal{C}_{dense} = M/\overline{K_p}$, the CS prediction $\mathcal{C}_{CS} = 2\ln(M/\overline{K_p})$, and the "top-down" limit reached in Frey et al. (2013) $\mathcal{C}_{td} = L/\overline{K_p}$, where the average value $\overline{K_p} = 53.5$ is used for all peaks (Fig. 11a). Every individual-peak $\mathcal{C}_{bu}$ achieved is lower than any of these benchmarks.

The prediction bands described in Eq. (10) can be combined with the recipe (Eq. 11) to predict the scale of

the peak reconstruction error you should expect at $N_{t_1}^p$. Whether that error level is acceptable depends, of course, on your experimental needs. A small modification to the recipe allows it to account for this: if the error range is too large at the initial "$N_{t_1}^p$," simply keep going up through the safe-zone N_{t_1} values until the error range is sufficiently small. This expanded recipe can be written as

$$N_{t_1}^p = \min(N_{t_1}) \text{ for } \begin{cases} N_{t_1} > K_p/2 \\ \text{MCM}(m_p, N_{t_1}) = 0 \ , \\ \max(|E_{\text{SZ}}|) < E_{\max} \end{cases} \quad (12)$$

where $E_{\max}$ is the maximum acceptable noise level. See Supplementary Information for an example of a case when this might be necessary. That said, in Figs. 10 and 11, we simply used the lowest $N_{t_1}^p$ above $K_p/2$ (i.e., we applied Eq. (11) instead of Eq. (12)).

If we build up from the reconstruction of a single peak at a time, to many peaks in parallel, what happens to $N_{t_1}^*$?

In order to answer (Q.I), we reduced the reconstruction problem to its smallest possible target: a single peak of interest in a single f_2 column. Using a deeper understanding of the iterated maps method, we pushed close to the theoretical minimum $N_{t_1}^p$ for each peak.

Turning to (Q.II), we want to see if we can build upon these results to attack the larger problem of reconstructing many, or all, peaks at one time (using a single global sampling number, $N_{t_1}^*$), as we did in our earlier study (Frey et al. 2013).

The mask collision metric MCM is different for each (f_1, f_2) point of interest, but it is easy to extend it to apply to a group of adjacent points in the same f_2 column: simply add together the metric for each point. Any N_{t_1} values for which the combined metric is 0 should be safe for all of the points being considered. It would make sense to do this for a group of adjacent f_1 points, for instance, if one wants to do curve fits to interpolate peak locations. In that case, the combined metric tends to resemble the metric of the center point of the peak.

Extending the metric method beyond this case, however, starts to get trickier. For other non-adjacent points in the same column (a different peak, for instance), the metrics may not overlap very well, even though they use the same mask. The most challenging case is typically peaks in different f_2 columns, since there is no guarantee that the safe zones in the metrics will line up at all. This is simply because the masks in distant columns of a complicated spectrum are not very correlated with each other, so there is no reason to
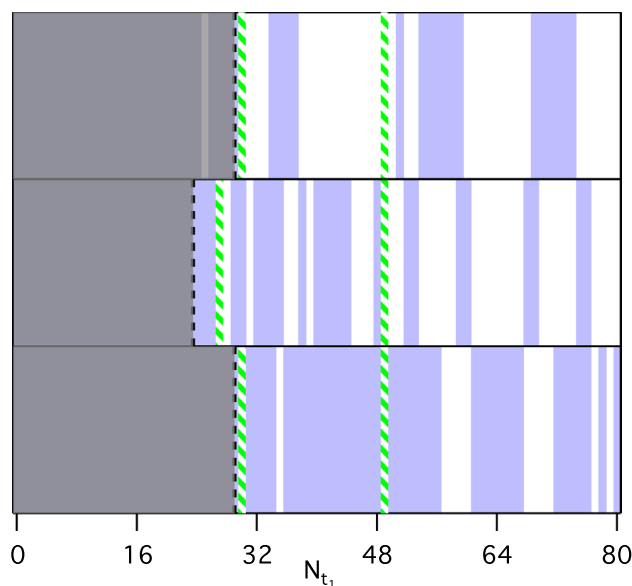

Fig. 12 MCM predictions for the three peaks in Fig. 6 (Val56 γ^2, Val18 γ^1, and Ile52 δ^1). $N_{t_1}^p$ is highlighted for each peak (green dashed lines), calculated according to the recipe in Eq. (11). The peaks Val56 γ^2 and Ile52 δ^1 (top and bottom) both have $N_{t_1}^p = 30$, but in order to simultaneously satisfy the metric condition for those peaks *and* Val18 γ^1 (middle), one must go all the way up to $N_{t_1}^* = 49$. This problem compounds with each added peak

expect that you will be able to use the same N_{t_1} for multiple peaks across the spectrum. Figure 12 shows an example of this issue, using just three peaks of interest. For two of the peaks, their respective mask collision metrics have a safe zone at the same value of $N_{t_1}^p$. But for the third peak, that same N_{t_1} value is in its danger zone. Despite the fact that the third peak's K_p and $N_{t_1}^p$ are lower than that of the other two, the "combined $N_{t_1}^p$" ($N_{t_1}^*$) for all three peaks is much higher.

This problem means that, at least for the relatively complicated spectra analyzed here, the insights we have gleaned into the dependence of the *local* reconstruction performance on N_{t_1} cannot be applied to lower the *global* value of $N_{t_1}^*$ below the already-reached level of $L/2$. In fact, a strict application of the "union of sets" approach described above to the full set of 114 peaks would push $N_{t_1}^*$ well above the level of $L/2$. This is not reasonable, however, because we already know from the top-down approach in Frey et al. (2013) that pushing down to $L/2$ is, in fact, globally acceptable. This apparent contradiction is likely due to the fact that the false positive rate of the MCM grows with increasing N_{t_1}. This shows that while false positives are slightly inconvenient for a single peak, they are crippling when considering a large number of peaks.

Discussion

This paper was motivated by a simple question about the iterated maps approach: could we use a greatly improved understanding of the action of DiffMap on a particular 2D data set to reach an N_{t_1} below the previous minimum number of sparsely-sampled points ($2N_{t_1}^* = L = 150$) achieved for that data set (Frey et al. 2013)? Focusing on this question narrowed the scope of this study. For example, we worked with the data as acquired, so any optimization of our NUS sampling schedules had to be consistent with the existing dense grid of points. Thus, even though the sharper PSF for the case of a $t_1 = 0$ point could lead to improved performance of our method, the current study is limited to the case of a $t_1 = \delta_1/2$ point (see Fig. 4b, c). As another example, we did not directly compare the performance of DiffMap to any other NUS reconstruction methods. Nevertheless, by reporting the minimum compression ratios $\mathcal{C}$ achieved (as opposed to the undersampling fraction, $N_{t_1}^p/N_{t_1}^{\text{dense}}$), we hope to make it easier to compare results across methods. The limited scope allowed us to do a deep dive into the performance of DiffMap.

We started by reducing the 2D problem to a much simpler one: understanding the reconstruction of a single peak of interest in a single f_2 column. Using quasi-even sampling, we developed a metric to predict the danger zones of N_{t_1} values, regions which are correlated with the largest reconstruction errors. We do not know how to predict the size of those errors, but we can avoid them altogether by using only N_{t_1} values in the complementary safe zones.

While reconstruction quality is typically better in the safe zones, it does degrade as N_{t_1} is lowered. We came up with a quantitative model for this smaller source of error, which can be traced back to the signal outside the mask folding back into the mask. Next, we made a recipe for defining and getting to the "minimum acceptable N_{t_1}" for a given peak, called $N_{t_1}^p$. We then applied this procedure to all 114 peaks, peak-by-peak, and it worked very well. In the end, we were able to get high quality reconstructions of each peak (for all 11 slices), while pushing down to $\overline{\mathcal{C}_{\text{bu}}} = 1.25 \pm 0.31$ across all 114 peaks, which is quite close to the fundamental limit on the compression ratio, $\mathcal{C} = 1$.

To contextualize our results, we can compare to a nice recent study of NUS for pseudo-3D data done by Linnet and Teilum (2016). They did a very careful analysis of how many samples they needed to reconstruct a series of spectra, each with different amounts of information content. They define sparsity $\mathcal{S}$ as the fraction of data points in a contiguous (2D) region of the spectrum whose intensities (absolute value of amplitude) are less than six times the noise level, and coverage c as the minimum sampling fraction required to get an accurate reconstruction of a 2D spectrum using MDD.

Conveniently, we also defined τ to be six times the noise level when creating our mask, so $\mathcal{S}$ is equivalent to $1 - \overline{K}/M$, where $\overline{K}$ is the average K over the same 2D region. Meanwhile, c is a sampling fraction, so it is equivalent to what we call $N_{t_1}/N_{t_1}^{\text{dense}}$. Therefore, we can calculate an effective $\mathcal{C}_{\text{LT}}$ for their results which is very similar to $\mathcal{C}$ as we have defined it for our method: $\mathcal{C}_{\text{LT}} = c/(1 - \mathcal{S})$. Across their four spectra (with sparsities ranging from $\mathcal{S} = 92\%$ to $\mathcal{S} = 71\%$), their average compression ratio was $\overline{\mathcal{C}_{\text{LT}}} \approx 2.8$, which is comparable to our previous $\mathcal{C}_{\text{td}}$. Our $\overline{\mathcal{C}_{\text{bu}}}$ is much lower than any of their $\mathcal{C}_{\text{LT}}$ values, but this direct comparison is not entirely fair, because $\overline{\mathcal{C}_{\text{bu}}}$ is the average of a set of separate 1D calculations, each focused on reconstructing a different peak, which do not share the same $N_{t_1}^p$ (or c) and K_p (or $\mathcal{S}$). However, this comparison does illustrate the benefit of focusing the reconstruction problem to a small portion of the spectrum.

Using the refined DiffMap approach we have described in this paper, a *single* peak of interest can be accurately reconstructed, using a number of measurements that is close to the fundamental minimum value. This result is immediately applicable to experiments requiring a particular spectral feature in a 2D spectrum to be repeatedly analyzed, as is done in the CEST experiment. In another example, after running an R_2 dispersion measurement, a user may want to get a higher density of τ_{cp} points for a particularly important residue. Examples such as these can take advantage of our understanding of the single peak reconstruction problem.

Our quantitative model for the smaller source of reconstruction error has an interesting implication: the signal-to-noise ratio (SNR) at the f_1 point[1] of a reconstructed peak of interest is *always worse* after using DiffMap, when compared to the SNR of the dense experiment. To see this, consider the ideal, best-case scenario of just "ideal noise" and no tails outside the mask, which in our error model implies $\mu_v = 0$ and $\sigma_v = \sigma_{\text{min}}$, where σ_{min} is the standard deviation of the noise far from any peaks. In that case, if the SNR of the dense peak is A/σ_{min}, then the SNR of the reconstructed peak should drop to the lower value $A/(\sigma_{\text{min}}\sqrt{N_{t_1}^{\text{dense}}/N_{t_1}})$. This expression is reminiscent of the SNR penalty suffered if fewer repetitions of an experiment are done. Of course, if the time saved by NUS is used to repeat the experiment $N_{t_1}^{\text{dense}}/N_{t_1}$ times, then this boosts the SNR by a factor of $\sqrt{N_{t_1}^{\text{dense}}/N_{t_1}}$. Unfortunately, even in this ideal case, the iterated maps SNR would just match the dense value of A/σ_{min},

---

[1] Note that we are not calculating this SNR by comparing signal amplitude of a peak to noise amplitude far from all peaks after reconstruction, because this is known to be misleading (see e.g. Hyberts et al. (2017)). Instead, we are considering the ratio of the reconstructed amplitude to the expected variation in that amplitude, which we estimate using the E_{SZ} model explained above.

in exactly the same time! In practice, a real spectrum has tails outside the mask, so $\mu_\nu \neq 0$ and $\sigma_\nu > \sigma_{\min}$. Thus, the SNR of the iterated maps reconstruction is almost certainly going to be worse than the dense reconstruction. In our model, we assumed that $\widehat{P}_1$ and $\widehat{P}_2$ do not use noise-handling, so these conclusions should be checked if that is not the case. Still, we think it is safe to say that a peak of interest with the bare-minimum SNR in the dense spectrum is not a good candidate for iterated maps reconstruction.

The low values of $\mathcal{C}_{bu}$ presented here were achieved by analyzing *individual* features in particular 1D spectra. We tried to use our understanding of this case to find the minimum $2N_{t_1}$ required to reconstruct many, or all, peaks simultaneously. However, this bottom-up approach to finding a global $N_{t_1}^*$ did not work. Given the complicated spectrum under study, trying to reconstruct ∼3-4 peaks at once (out of 114) can be enough to push our recipe for $2N_{t_1}^*$ up to the previous value $L = 150$ found in the top-down approach (Frey et al. 2013). It is still possible that a broad search of the parameter space available to DiffMap (e.g., moving away from quasi-even sampling) could reveal new ways to improve global DiffMap performance for $2N_{t_1} < L$, but such a search is beyond the scope of this paper.

For the original problem of global 2D spectral reconstruction using iterated maps, the failure of our bottom-up approach to do any better than our top-down approach left us wondering if we could ever do better. In the end, we decided that there simply is not enough information in a single 2D slice to push below our previous limit of L (Frey et al. 2013). However, the fact that these 2D slices are part of a larger pseudo-3D data set suggests a way forward: use the correlations across 2D data sets as *additional information* that enables a global, high-quality reconstruction. In a companion paper, we develop a modified version of Diff-Map, called coDiffMap (Rovny et al. 2019). Using coDiff-Map, we show that we can improve results *globally* by exploiting the correlations that exist among 2D spectra in particular types of pseudo-3D experiment, adding information to our reconstruction method. This extra information enables us to push global $2N_{t_1}^*$ well below L, and even below $\mathcal{C} = 1$ for certain cases (since the extra information across 2D slices adds to the information content of K_p). This alternate method also changes the scaling of the error that we described earlier in this paper. See Rovny et al. (2019) for more details about coDiffMap.

References

Billeter M (2017) Non-uniform sampling in biomolecular NMR. J Biomol NMR 68(2):65–66. https://doi.org/10.1007/s10858-017-0116-7

Bostock M, Nietlispach D (2018) Compressed sensing: reconstruction of non-uniformly sampled multidimensional NMR data. Concepts Magn Reson Part A 46A(2):e21438. https://doi.org/10.1002/cmr.a.21438

Elser V (2003) Phase retrieval by iterated projections. J Opt Soc Am A 20(1):40–55. https://doi.org/10.1364/JOSAA.20.000040

Elser V, Rankenburg I, Thibault P (2007) Searching with iterated maps. Proc Natl Acad Sci USA 104(2):418–423. https://doi.org/10.1073/pnas.0606359104

Fienup JR (1982) Phase retrieval algorithms: a comparison. Appl Opt 21(15):2758–2769. https://doi.org/10.1364/AO.21.002758

Foucart S, Rauhut H (2013) A mathematical introduction to compressive sensing. Springer, New York

Frey MA, Sethna ZM, Manley GA, Sengupta S, Zilm KW, Loria JP, Barrett SE (2013) Accelerating multidimensional NMR and MRI experiments using iterated maps. J Magn Reson 237:100–109. https://doi.org/10.1016/j.jmr.2013.09.005

Haacke E, Lindskogj E, Lin W (1991) A fast, iterative, partial-Fourier technique capable of local phase recovery. J Magn Reson 92(1):126–145. https://doi.org/10.1016/0022-2364(91)90253-P

Hoch J, Stern A (1996) NMR data processing, 1st edn. Wiley-Liss, New York

Hyberts SG, Takeuchi K, Wagner G (2010) Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. J Am Chem Soc 132(7):2145–2147. https://doi.org/10.1021/ja908004w pMID: 20121194

Hyberts SG, Arthanari H, Robson SA, Wagner G (2014) Perspectives in magnetic resonance: NMR in the post-FFT era. J Magn Reson 241:60–73. https://doi.org/10.1016/j.jmr.2013.11.014

Hyberts SG, Robson SA, Wagner G (2017) Interpolating and extrapolating with hmsist: seeking a tmax for optimal sensitivity, resolution and frequency accuracy. J Biomol NMR 68(2):139–154. https://doi.org/10.1007/s10858-017-0103-z

Kazimierczuk K, Orekhov VY (2011) Accelerated NMR spectroscopy by using compressed sensing. Angew Chem 123(24):5670–5673. https://doi.org/10.1002/ange.201100370

Kazimierczuk K, Orekhov VY (2012) A comparison of convex and non-convex compressed sensing applied to multidimensional NMR. J Magn Reson 223:1–10. https://doi.org/10.1016/j.jmr.2012.08.001

Korzhnev DM, Kloiber K, Kay LE (2004) Multiple-quantum relaxation dispersion NMR spectroscopy probing millisecond time-scale dynamics in proteins: theory and application. J Am Chem Soc 126(23):7320–7329. https://doi.org/10.1021/ja049968b pMID: 15186169

Linnet TE, Teilum K (2016) Non-uniform sampling of NMR relaxation data. J Biomol NMR 64(2):165–173. https://doi.org/10.1007/s10858-016-0020-6

Lipchock JM, Loria JP (2010) Nanometer propagation of millisecond motions in V-type allostery. Structure 18(12):1596–1607. https://doi.org/10.1016/j.str.2010.09.020

Lisi GP, Currier AA, Loria JP (2018) Glutamine hydrolysis by imidazole glycerol phosphate synthase displays temperature dependent allosteric activation. Front Mol Biosci 5:4. https://doi.org/10.3389/fmolb.2018.00004

Loria JP, Rance M, Palmer AG (1999) A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. J Am Chem Soc 121(10):2331–2332. https://doi.org/10.1021/ja983961a

Maciejewski MW, Qui HZ, Rujan I, Mobli M, Hoch JC (2009) Nonuniform sampling and spectral aliasing. J Magn Reson 199(1):88–93. https://doi.org/10.1016/j.jmr.2009.04.006

Maciejewski MW, Schuyler AD, Gryk MR, Moraru II, Romero PR, Ulrich EL, Eghbalnia HR, Livny M, Delaglio F, Hoch JC (2017) NMRbox: a resource for biomolecular NMR computation. Biophys J 112(8):1529–1534. https://doi.org/10.1016/j.bpj.2017.03.011

Matsuki Y, Eddy MT, Herzfeld J (2009) Spectroscopy by integration of frequency and time domain information for fast acquisition of high-resolution dark spectra. J Am Chem Soc 131(13):4648–4656. https://doi.org/10.1021/ja807893k

Matsuki Y, Eddy MT, Griffin RG, Herzfeld J (2010) Rapid three-dimensional MAS NMR spectroscopy at critical sensitivity. Angew Chem Int Ed 49(48):9215–9218. https://doi.org/10.1002/anie.201003329

Matsuki Y, Konuma T, Fujiwara T, Sugase K (2011) Boosting protein dynamics studies using quantitative nonuniform sampling NMR spectroscopy. J Phys Chem B 115(46):13740–13745. https://doi.org/10.1021/jp2081116 pMID: 21992609

Mayzel M, Kazimierczuk K, Orekhov VY (2014) The causality principle in the reconstruction of sparse NMR spectra. Chem Commun 50:8947–8950. https://doi.org/10.1039/C4CC03047H

Mobli M, Hoch JC (2008) Maximum entropy spectral reconstruction of nonuniformly sampled data. Concepts Magn Reson Part A 32A(6):436–448. https://doi.org/10.1002/cmr.a.20126

Neudecker P, Lundström P, Kay LE (2009) Relaxation dispersion NMR spectroscopy as a tool for detailed studies of protein folding. Biophys J 96(6):2045–2054. https://doi.org/10.1016/j.bpj.2008.12.3907

Orekhov VY, Jaravine VA (2011) Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. Prog Nucl Magn Reson Spectrosc 59(3):271–292. https://doi.org/10.1016/j.pnmrs.2011.02.002

Palmer AG, Kroenke CD, Loria JP (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. In: James TL, Dötsch V, Schmitz U (eds) Nuclear magnetic resonance of biological macromolecules—Part B, methods in enzymology, vol 339, Academic Press, New York, Chap 10, pp 204 – 238, https://doi.org/10.1016/S0076-6879(01)39315-1

Palmer AG, Massi F (2006) Characterization of the dynamics of bio-macromolecules using rotating-frame spin relaxation NMR spectroscopy. Chem Rev 106(5):1700–1719. https://doi.org/10.1021/cr0404287 pMID: 16683750

Paramasivam S, Suiter CL, Hou G, Sun S, Palmer M, Hoch JC, Rovnyak D, Polenova T (2012) Enhanced sensitivity by nonuniform sampling enables multidimensional MAS NMR spectroscopy of protein assemblies. J Phys Chem B 116(25):7416–7427. https://doi.org/10.1021/jp3032786 pMID: 22667827

Rovnyak D, Schuyler AD (2018) Advances in alternative sampling and processing. Concepts Magn Reson Part A 46A(2):e21458. https://doi.org/10.1002/cmr.a.21458

Rovny J, Blum RL, Loria JP, Barrett SE (2019) Accelerating 2D NMR relaxation dispersion experiments using iterated maps. J Biomol NMR. https://doi.org/10.1007/s10858-019-00263-3

Schlossmacher EJ (1973) An iterative technique for absolute deviations curve fitting. J Am Stat Assoc 68(344):857–859. https://doi.org/10.1080/01621459.1973.10481436

Shchukina A, Kasprzak P, Dass R, Nowakowski M, Kazimierczuk K (2017) Pitfalls in compressed sensing reconstruction and how to avoid them. J Biomol NMR 68(2):79–98. https://doi.org/10.1007/s10858-016-0068-3

Stanek J, Koźmiński W (2010) Iterative algorithm of discrete Fourier transform for processing randomly sampled NMR data sets. J Biomol NMR 47(1):65–77. https://doi.org/10.1007/s10858-010-9411-2

States D, Haberkorn R, Ruben D (1982) A two-dimensional nuclear overhauser experiment with pure absorption phase in four quadrants. J Magn Reson 48(2):286–292. https://doi.org/10.1016/0022-2364(82)90279-7

Stern AS, Donoho DL, Hoch JC (2007) NMR data processing using iterative thresholding and minimum l_1-norm reconstruction. J Magn Reson 188(2):295–300. https://doi.org/10.1016/j.jmr.2007.07.008

van Zijl PCM, Yadav NN (2011) Chemical exchange saturation transfer (CEST): what is in a name and what isn't? Magn Reson Med 65(4):927–948. https://doi.org/10.1002/mrm.22761

Walinda E, Morimoto D, Sugase K (2018) Resolving biomolecular motion and interactions by R_2 and $R_{1\rho}$ relaxation dispersion NMR. Methods 148:28–38. https://doi.org/10.1016/j.ymeth.2018.04.026

Ying J, Delaglio F, Torchia DA, Bax A (2017) Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. J Biomol NMR 68(2):101–118. https://doi.org/10.1007/s10858-016-0072-7

Zhu G, Bax A (1990) Improved linear prediction for truncated signals of known phase. J Magn Reson 90(2):405–410. https://doi.org/10.1016/0022-2364(90)90150-8