

Subject Section

RTEExtract: Time-series NMR spectra quantification based on 3D surface ridge tracking

Yue Wu¹, Michael T. Judge², Jonathan Arnold^{1,2}, Suchendra M. Bhandarkar⁴ and Arthur S. Edison^{1,2,3,5*}

¹Institute of Bioinformatics, University of Georgia, Athens, GA, USA, ²Department of Genetics, University of Georgia, Athens, GA, USA, ³Complex Carbohydrate Research Center, University of Georgia, Athens, GA, USA, ⁴Department of Computer Science, University of Georgia, Athens, GA, USA, ⁵Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA, USA.

*To whom correspondence should be addressed.

Associate Editor: XXXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Motivation: Time-series NMR has advanced our knowledge about metabolic dynamics. Before analyzing compounds through modeling or statistical methods, chemical features need to be tracked and quantified. However, because of peak overlap and peak shifting, the available protocols are time consuming at best or even impossible for some regions in NMR spectra.

Results: We introduce RTEExtract (Ridge Tracking based Extract), a computer vision-based algorithm, to quantify time-series NMR spectra. The NMR spectra of multiple time points were formulated as a 3D surface. Candidate points were first filtered using local curvature and optima, then connected into ridges by a greedy algorithm. Interactive steps were implemented to refine results. It was tested on both simulated and experimental data sets composed of time-series NMR spectra and exhibited high accuracy in peak tracking for overlapping regions.

Availability: Source code is freely available within Metabolomics toolbox GitHub repository (https://github.com/artedison/Edison_Lab_Shared_Metabolomics_UGA) and is implemented in MATLAB and R.

Contact: aedison@uga.edu

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Deleted: ing

1 Introduction

Experimental approaches have been developed in time-series metabolic measurements by both NMR (Nuclear Magnetic Resonance) and MS (Mass Spectrometry) (Judge, et al., 2019; Koczula, et al., 2016; Link, et al., 2015; Montana, et al., 2011). These experimental methods provide opportunities to understand metabolic dynamics, including metabolic changes under variation in carbon sources or oxygen levels (Judge, et al., 2019; Link, et al., 2014; Link, et al., 2015). Among existing approaches, the CIVM-NMR (continuous *in vivo* monitoring of metabolism by NMR) method provided high time-resolution, *in vivo* measurements of metabolites in *Neurospora crassa* under aerobic and anaerobic conditions (Judge, et al., 2019). These measurements covered a large proportion of pathways

in central metabolism, and interesting dynamics in compound concentration were observed.

NMR provides a highly reproducible way to identify and quantify compounds. In an NMR spectrum, different metabolites are represented by different peaks (features), and peak height (intensity) is proportional to compound concentration. Peak resonance frequency is sensitive to the local electronic structure and some environmental variables. Resonance frequency is reported as chemical shift, δ , which is derived by dividing the frequency in Hz by the spectrometer frequency in MHz and thus has units of parts per million (ppm). The dependence on local electronic structure allows for reliable complete identification. Additionally, some metabolites are sensitive to changes in the local chemical environment (e.g. pH or metal ion concentration) and systematically change their chemical shift,

providing a useful way to measure these environmental factors (Takis, et al., 2017; Tredwell, et al., 2016).

Metabolism yields changes in metabolite concentration and local pH, changing peak intensity and chemical shift. These changes provide important information about metabolic dynamics but also complicate feature extraction. Moreover, peaks in NMR spectra often overlap, which affects both compound annotation and quantification. The combination of systematic chemical shift, overlap, and amplitude changes makes peak tracking and quantification a difficult problem. A practical, stable computational approach is needed to track and quantify peaks over time, regardless of overlap, amplitude and chemical shift changes.

Traditional alignment-based methods are popular for processing NMR spectra from different samples and aligning shifting peaks. However, these methods often introduce artifacts and are unreliable for the regions where peaks cross (Csenki, et al., 2007; Vu and Laukens, 2013). In CIVM-NMR data, the pattern of peak shifting is less noisy and more continuous than in discrete extracted samples by traditional methods. These properties provide new information for quantifying crossing peaks as discussed later.

Multiple methods have been implemented to track peaks in time-series NMR spectra. The TSA Tool can track a peak through time by a predefined function describing peak shifting trajectory (Koczula, et al., 2016; Ludwig and Gunther, 2011). This method, though capable of tracking individual peaks, does not provide a general solution for quantifying multiple peaks efficiently. In our initial CIVM-NMR study, peak tracking was achieved by a smoothing filter to reduce noise (filtering step) and hierarchical clustering (connection step) to connect candidate peaks (Judge, et al., 2019). While this method tracked peaks with chemical shift variation, substantial manual effort was needed in parameter tuning to accommodate different regions of the spectra. For instance, the proper scaling factor for the extent of chemical shift variation and the number of expected clusters were crucial parameters but difficult to optimize. A few days of work were needed for the quantification in the original CIVM-NMR publication, which can be a significant bottleneck and cost (Judge, et al., 2019). Additionally, none of the aforementioned methods can deal with crossing or severely overlapped peaks.

Computer vision methods have been adapted to solve other spectroscopy problems (Klukowski, et al., 2018; Klukowski, et al., 2015) and can also be implemented here to promote efficiency. The steps of both filtering candidate points and subsequent connection can be improved by treating NMR peak extraction as a ridge tracking problem. Time-series NMR data can be viewed as a 2D matrix (or a 3D surface if we treat matrix elements as height) with each row being a spectrum at one time point and each column being the intensity of a particular resonance frequency across time. As the same peaks change continuously through time, they can be conceptualized as surface ridges, for which efficient detection algorithms exist (Suk and Bhandarkar, 1992). Surface segmentation techniques have been implemented in computer vision to classify 3D surface points based on their local curvature into qualitative surface types: *inter alia*, ridge, peak, and valley (Supplementary Figure 1) (Besl and Jain, 1986; Besl and Jain, 1988; Suk and Bhandarkar, 1992).

In RTExtract, to filter candidate peak points, we combined ridge classification with other information such as local maxima. This combination of multiple filters provided cleaner results with fewer false positives, and tuning parameters were fewer and more intuitive. Candidate points were then connected by a 2-step greedy method, which is composed of simple

local optimal connections without global evaluations and are possible because of the better filter on candidate points. Additionally, refinement steps were introduced to expand flexibility in tracking and increase tracking accuracy.

In this paper, we present our new method (RTExtract) to extract and quantify time-series NMR spectra. We simulated time-series NMR data specifically presenting the challenges that limited previous methods. We also conducted a direct comparison of our previous method and RTExtract on experimental datasets (Judge, et al., 2019) and found that RTExtract was faster and manually easier than our previous approach. Previous tracking results were reproduced in less than two hours instead of several days. Additionally, we were also able to track complex spectral regions, such as those with high amounts of overlap and crossing peaks, that were impossible with previous published methods. RTExtract therefore significantly expands the utility of the rich data collected in CIVM-NMR and accelerates its analysis.

2 Methods

2.1 Ridge point classification

Local curvature was used to classify ridge points and functions as one of the filters for candidate points. Including a ridge point filter with local optima filters reduced noise levels in selecting candidate points and increased accuracy in ridge tracking. The following section describes the ridge point filter.

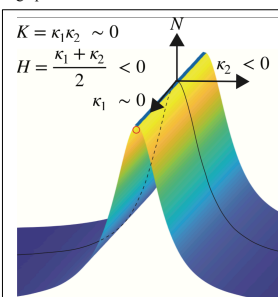


Figure 1: Illustration of the concept of H and K curvature. N is the normal vector and the black curve is one of the intersect curve. For every point in the curve, there is a tangent vector (as indicated by the two horizontal black arrows) and based on its derivative, curvature of the curve can be represented. The two principal curvatures, the maximum curvature (κ_1) and the minimum curvature (κ_2), correspond to the two vectors. Gaussian curvature (K) and mean curvature (H) are represented by the principal curvatures as shown above. For the ridge surface shown here, $\kappa_2 < 0$ and $\kappa_1 \approx 0$ which results in $H < 0$ and $K \approx 0$. Details regarding formulation and computation of H and K curvature can be found in Methods 2.1. Other surface types are illustrated in Supplementary Figure 1.

For each point on the 3D surface (Figure 1), a normal vector (N) can be defined. All planes that contain N (the normal planes) will intersect the 3D surface along a curve, and for each such curve at the point of interest, the curvature can be computed. The maximum and minimum curvature values, denoted by κ_1 and κ_2 respectively, correspond to two mutually orthogonal orientations of the normal planes and are referred to as the principle curvatures of the 3D surface at that point. From κ_1 and κ_2 , the Gaussian curvature (K) and mean curvature (H) are defined (Equation [3] & [2] and Figure 1) (Besl and Jain, 1986; Besl and Jain, 1988; Suk and Bhandarkar, 1992). The curvatures H and K can be used to classify 3D

Deleted: d

Formatted: Font: Not Italic

Deleted: 1

Formatted: Font: Not Italic

Deleted: 2

NMR feature quantification by ridge tracking

surface points locally into qualitative types, including, *inter alia*, peak, ridge, and valley. Specifically, when $K \approx 0$ and $H < 0$, the surface is classified as a ridge surface, and the central point of the surface is the candidate point (Supplementary Figure 1).

$$K = \kappa_1 \kappa_2 \quad [1]$$

$$H = \frac{\kappa_1 + \kappa_2}{2} \quad [2]$$

As an alternative to Equations [1] and [2], the values of H and K can also be derived through the fundamental form matrices G and B , which provide a practical computation process (Stoker, 1969). Let $z = f(x, y)$ be the surface and the point $X = (x, y, f(x, y))$ be a point on the surface. The first fundamental form G of the surface and the second fundamental form B of the surface can be computed from partial derivatives of the surface and the unit surface normal vector (\mathbf{n}) (Equations [3], [4], [5]) (Suk and Bhandarkar, 1992).

$$G = \begin{bmatrix} X_x \cdot X_x & X_x \cdot X_y \\ X_y \cdot X_x & X_y \cdot X_y \end{bmatrix} \quad [3]$$

$$B = \begin{bmatrix} \mathbf{n} \cdot X_{xx} & \mathbf{n} \cdot X_{xy} \\ \mathbf{n} \cdot X_{xy} & \mathbf{n} \cdot X_{yy} \end{bmatrix} \quad [4]$$

$$\mathbf{n} = \frac{X_x \times X_y}{\|X_x \times X_y\|} \quad [5]$$

We use a discrete biorthogonal second-order Chebychev polynomial with the interaction term ignored to approximate the local 3D surface within a 7 by 7 window (Besl and Jain, 1986; Haralick, et al., 1983). Using biorthogonal polynomials instead of a more general fitting process increased computational speed. As the surface of interest was large (e.g. ~50 spectra \times 35000 points for each spectrum in our experimental data set) (Judge, et al., 2019), this approximation was necessary to incorporate real-time analysis input within the workflow (wait time < 5 seconds). From the biorthogonal polynomial approximation, the first and second order derivatives, the fundamental forms, and the curvatures (H and K) were computed in that order (Suk and Bhandarkar, 1992).

Multiple surface types were generated from a second-order polynomial with the interaction term in a 101×101 window (Equation [6], Supplementary Figure 1). The parameters A , B , and C in Equation [6] were varied to produce different surface types, including saddle ridge, minimal surface, saddle valley, ridge, flat surface, valley, peak, and pit. The curvatures H and K were computed for the central point in the window to check with the expected values. The same algorithm for ridge tracking (see the previous paragraph) was used in the computation. The expected H and K curvatures were derived based on Equation [6] and computed with Equations [7] and [8].

$$Z = AX^2 + BXY + CY^2 \quad [6]$$

$$H = A + C \quad [7]$$

$$K = 4AC - B^2 \quad [8]$$

2.2 Feature quantification by ridge tracking

The entire workflow of RTEExtract is presented in Figure 2. The steps include filtering candidate points, connecting candidate points into initial ridges, ridge refinement, and manual ridge selection. The initial ridge tracking process will be covered in this section and refinement in the next section.

The tested experimental data sets contain ~50 spectra acquired at different time points (~11 hours), and each of them is comprised of ~35000 points in chemical shift resolution. The original time-series data sets were collected with finer time resolution, and the averaged (denoised) data sets were used to evaluate RTEExtract on the same criteria as published (Judge, et al., 2019). In NMR spectra, even a small region can exhibit high

complexity, and peaks of interest also differ considerably from each other in intensity (Figure 2A, green and blue boxes). A region of interest (ROI) (ppm [1.3, 1.35] (Orange block in Figure 2A) was selected as an example to illustrate the computational pipeline (Figure 2B-E). The ROI is presented as a surface, in which different intensities can be visualized as different colors like a topographic map (Figure 2B). Each row of the surface matrix is a single spectrum acquired at one time point. To filter candidate points, information from curvature, local maxima, and a controlled number (N_{max}) of global maxima were combined (Figure 2C). Points on the surface were classified as ridge points (Set S_R) if they satisfied the curvature criteria in Equation [9], for which no changes in the thresholds (1 and 0) were needed to accommodate different spectral regions. Besides ridge points, candidate points were also supplied through N_{max} (the number of highest local maxima to add for each row), which define the Set $S_G(N_{max})$. For each selected ROI, N_{max} more points are added as candidate points on each row. These candidate points were then intersected with local maxima (Set S_L) to filter out points which did not correspond to true peaks. The combination of the three criteria ($(S_R \cup S_G(N_{max})) \cap S_L$) helped identify most ridges and improved accuracy.

$$\begin{aligned} |K| &< 1 \\ H &< 0 \end{aligned} \quad [9]$$

A two-step greedy connection procedure was implemented to connect these candidate points into ridges for quantifying individual peaks through time (Figure 2D). This procedure assumes that chemical shift variation of peaks at nearby time points is local and continuous, which is typically the case in time-series measurements. First, points adjacent in time and with the closest chemical shift distance within L_{gap} (largest step size in chemical shift dimension) were connected into segments. Second, these segments were connected into ridges to cover the entire time-range for the peak. The segment connection was based on the shortest distance and a user-adjustable threshold on angle ($\leq 60^\circ$ default) between them. The angle threshold ensured smooth shift pattern in ridges. The order of the segment connection was ranked from high to low based on their average intensity.

In the ridge tracking process, only the parameters N_{max} and L_{gap} required tuning. The remaining parameters in the program required no modification for the simulated and experimental data sets we tested. Choices for N_{max} and L_{gap} values were also intuitive. In the majority of cases, we recommend the same small L_{gap} for most regions. The parameter L_{gap} can be increased when there is peak shifting and can be decreased when there are peaks that are close to each other. For N_{max} , we recommend using $N_{max}=1$ plus the number of ridges expected but not yet tracked. The values used in the script ($N_{max} = 1$, $L_{gap} = 10$) can be used as an initial guess for other data sets.

2.3 Refinement of the tracking results

While most peaks can be tracked without refinement, in some cases, the tracking is imperfect, which can be solved in the refinement step (Figure 2E). Chemical expertise adds value to this step, especially in regions with a low signal-to-noise ratio (SNR). Besides removing short ridges (default minimum ridge length is 5 time points), the refinement steps also include retracking for small regions, manual ridge selection, and removal of imperfect ridge ends (Supplementary File 1).

Deleted: 1

Formatted: Font: Not Italic

Deleted: 2

Formatted: Font: Not Italic

Deleted: 3

Deleted: 4

Deleted: 5

Deleted: 1

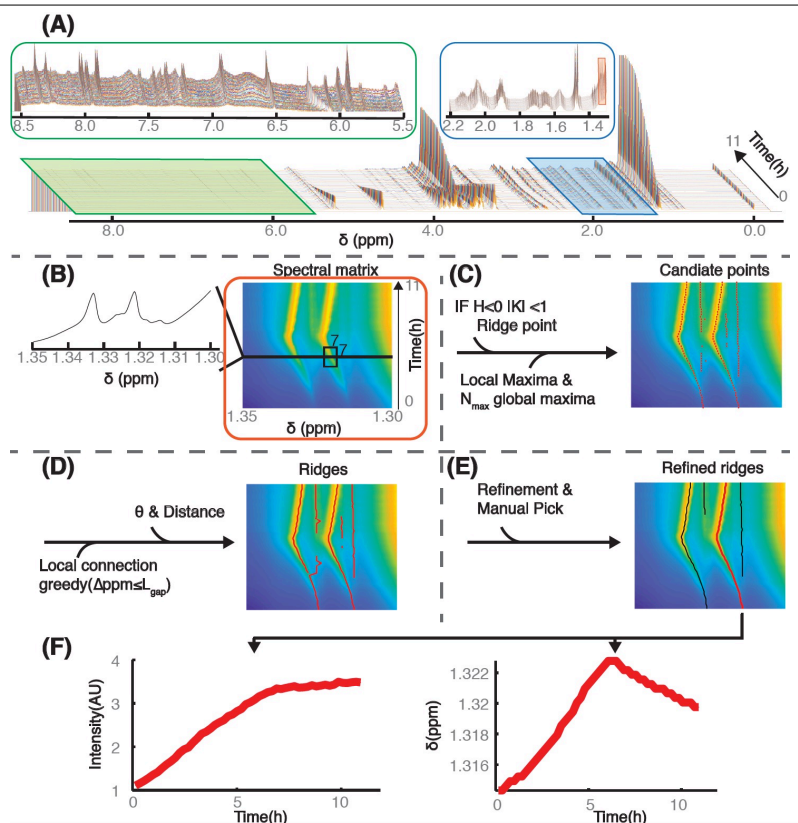


Figure 2: Illustration of the RTExtract algorithm. The algorithm is presented step by step based on an example time-series NMR data set, which was measured under aerobic conditions (A) (Judge, et al., 2019). The stacked spectral plot (A) shows changes of whole NMR spectra (X-axis chemical shift δ) through time (Y-axis), and each time point is distinguished by a different line color. Two regions (ppm [5.5, 8.5] and [1.3, 2.2]) are expanded here to show complexity of the spectra in scales and dynamics. The time-series spectra can be treated as a matrix, and each row of the matrix is a single NMR spectrum measured at one time point. (B) A column slice (ppm [1.30, 1.35]) and the entire time range, indicated as orange box in (A) of the matrix is chosen as show case for RTExtract (B-F). The selected sub-matrix can be treated as a surface as in Figure 1, and peaks changing through time will be ridges on this surface. (C) Based on the matrix, H and K curvatures were computed by an biorthogonal polynomial based on the 7 by 7 window (B) around each point, which was used to define ridge points ($H < 0$ and $|K| < 1$). These ridge points are intersected with local maxima. These peaks are also combined with the first N_{max} global maximal points for each spectrum. The candidate points are shown in red on the surface in (C). (D) These points were connected in two steps by a greedy method to form segments and then ridges. The first step was based on chemical shift distance (L_{gap}) between points, and the second step was based on distance and angles between segments from the last step. These ridges are shown as red lines on the surface in (D). (E) Ridges produced by the algorithm were refined and manually selected for feature quantification. In the refinement process, correct ridges that did not cover the entire time range can be extended, and imperfect ends of ridges can be removed. For the red ridge in (E), intensity and chemical shift (δ) were extracted and plotted against time (F). The intensity is measured in arbitrary units (AUs). Details on the RTExtract algorithm and tuning parameters can be found in Method 2.1 and 2.2. Time is measured in hours and chemical shift (δ) is measured in ppm.

When multiple peaks exhibit overlap and shifting, the greedy connection method (Methods 2.2) had difficulty deciding which direction to continue through peak crossing. This was ameliorated by local retracking. In retracking, we imposed a more stringent constraint that the peaks tend to maintain their original directions when they cross. For each time point, a new small search window (length 5) for connecting next candidate points was centered at the linear extrapolation of previous (last 5 time points)

chemical shift values. That is, ridges are assumed to be locally linear, which is a reasonable constraint for small windows. In the global tracking process (larger windows), however, there are indeed rapid changes in chemical shift, so in this case the stringent constraint is not imposed. Combining different procedures for global and local tracking increases flexibility when necessary.

Table 1: Complexity metrics of spectral data sets from simulations and experimental measurements.

Data source	SNR	Shift complexity	Scale complexity	Dynamics complexity
aerobic experimental data set	2.47E+02	4.53E-07	6.76E-02	5.61E-02
anaerobic experimental data set	6.49E+02	7.25E-08	3.55E-02	3.32E-02
simulated data set	1.84E+03	2.62E-07	1.67E-01	7.57E-02

Experimental and simulated data sets were compared for different types of complexity. SNR: signal to noise ratio for the DSS peak in the spectra. Shift complexity (C_{shift}): measurements of how much peaks shift in chemical shift dimension. Scale complexity (C_{scale}): measurements of how nearby peaks differ in intensity. Dynamics complexity ($C_{dynamics}$): measurements of how dynamic peak intensity changes through time. The higher the SNR, the less complex the spectral data sets. The higher the other complexity measurement, the more complex the spectral data sets. The complexity measurements were compared based on both spectra matrix and peak intensity. Details in computing these criteria can be found in Method 2.6.

The automatic ridge tracking procedure often generates false positives and these false ridges can be easily distinguished by the analyst. Hence, the interactive step (manual ridge selection) boosted performance by allowing analysts to select peaks with high confidence according to their knowledge. Moreover, compound quantification can also be improved by selecting peaks with good peak shapes, minimal overlap (with other peaks), and high SNR. The user can also record the annotated compound name and indicate whether the tracked peaks should be used for compound quantification in later steps.

When the peak intensity decreases to near 0, the ridge tracking sometimes extends to noisy regions with no peak for a few time points. These imperfect ridge ends can also be removed by the ridge end removal interactive step.

The feature quantification workflow provides an interface to walk the user through ridge tracking, retracking for overlapping regions, manual ridge picking, and manual end removal. The user can skip certain steps if needed. All information related to the tracking process, both explicit (parameter choices) and implicit (manual tuning records), is logged in data structure for documentation and reproducibility. Subsequently, peak intensity and chemical shift can be plotted through time (Figure 2F). More details can be found in the tutorial (Supplementary File 1).

2.4 Feature mapping and quantification

Tracked ridge features need to be mapped to compounds and quantified. In the simulated spectra, after ridge tracking, the time-series data for each peak were automatically mapped back to compounds by differences in chemical shift. The difference (D) was evaluated as Equation [10], and only the pairs with the smallest difference for all ridges and all compound peaks were selected as the final matches. In Equation [10], Ω is a set of time points overlapping between the simulated compound peaks and the tracked ridge peaks, and L is the size of the set Ω . The variables v_{ppm}^{compd} (simulated compound peak) and v_{ppm}^{ridge} (tracked ridge peak) are the corresponding chemical shift vectors within Ω . D is the sum of the squared differences between extracted and simulated chemical shifts within the overlapping time range and is normalized by the range size (L). Compounds were quantified by peak intensities normalized by the intensity of DSS (3-(Trimethylsilyl)-1-propanesulfonic acid sodium salt, a chemical shift reference and intensity internal standard) peak. We also compute RMSD (Root Mean Square Deviation) between simulated and extracted chemical shift (δ_{sim} and δ_{ex} , N is the length of the ridge) for each ridge (Equation [11]). The annotation and quantification of the experimental data sets follows our published methods (Judge, et al., 2019).

$$D = \frac{1}{L} \sum_{l \in \Omega} (v_{ppm}^{compd} - v_{ppm}^{ridge})^2 \quad [10]$$

$$RMSD = \sqrt{\frac{1}{N} \sum_{l=1}^N (\delta_{sim} - \delta_{ex})^2} \quad [11]$$

2.5 NMR spectral simulation

Time-series data sets were simulated according to the known challenges that limited previous methods. The simulation include following procedures and quantification of their complexity is defined in the next section (Methods 2.6).

Reference spectra for simulation were selected from GISSMO database (<http://gissmo.nmr.fam.wisc.edu>) because there is no noise or solvent peaks as in experimental measured spectra (Supplementary table 1). This is important, especially when multiple reference spectra are summed to simulate the mixture spectra, where measurement noise can accumulate. The GISSMO database also provides accurate simulation of NMR spectra at different field strengths, making it feasible to simulate spectra of different compounds under the same field strength as in the experimental data (600 MHz) (Dashti, et al., 2017; Judge, et al., 2019).

Peak intensity is simulated to change through time by changing compound concentration linearly. Some features increased through time, some decreased, and a few (e.g. DSS as the reference) were kept constant (Supplementary table 1). For each time point, compound spectra were multiplied by their concentration and summed to make the *in silico* mixture.

Besides changes in peak intensity, changes in chemical shift due to pH is also simulated in acetate and formate based on published titration data (Ackerman, et al., 1996; Tredwell, et al., 2016; Ye, et al., 2018). The pH varied from 4.0 to 6.0. The simulation was based on the Henderson-Hasselbalch equation and corresponding parameters, including pKa, chemical shift for the protonated form, and the unprotonated form (Supplementary table 1) (Ackerman, et al., 1996; Tredwell, et al., 2016; Ye, et al., 2018).

After summing all composed spectra, white noise was added (Gaussian distribution, $\mu = 0$, $\sigma = 5$). Exponential line broadening was introduced to simulate a line width similar to experimental measurements. The line broadening effect on NMR spectra (frequency domain) was implemented by multiplying the time domain data with an exponential decay function $e^{\lambda t}$ ($\lambda = -0.00035$), where t is the same in length with the original spectra.

Deleted: 10

Deleted: 10

Deleted: and are indexed by time

Formatted: Font: Not Italic

Deleted: 11

In addition to standard metabolite peaks, we added extra peaks (arbitrary peaks) to increase complexity and test the algorithm. More variation in chemical shift and peak intensity is introduced through these peaks. Arbitrary peaks are also introduced to overlap with shifting peaks, such as the acetate peak.

In summary, simulation of time-series NMR spectra is composed of these steps in order: varying concentration, shifting peak according to pH, adding arbitrary peak, adding noise, and applying exponential line broadening.

2.6 Metrics of spectral complexity

The complexity of experimental and simulated spectra is measured based on key peak qualities, including SNR, overlap, change in intensity, and chemical shift.

The metrics are calculated from both the spectral matrix and the ridge intensity matrix. In the ridge intensity matrix X , each row $X(i,:)$ indicates ridge points at one time point, and each column $X(:,j)$ indicates a distinct ridge, in which $i \in 1 \dots N$ and $j \in 1 \dots M$, where N is the total number of time points and M is the total number of ridges. The vector V_{ppm} denotes the whole chemical shift vector for the entire NMR spectrum, and v'_{ppm} denotes the chemical shift vector for a single ridge j through existing time points. The ridge matrix was shifted to be nonnegative (X_i) before the following calculations. A spectral region with no peaks (ppm [-0.4, -0.2]) was selected as a sample matrix (X_N) for noise level calculation.

SNR is the mean ratio of the DSS peak intensity to the standard deviation (sd) in a region with no peaks (Equation [12]). The higher the noise level, the lower the SNR value. Annotation and quantification of peaks are more difficult in region with low SNR.

$$SNR = \frac{1}{N} \sum_{i=1}^N \frac{X_i(l_{DSS})}{sd(X_N(i,:))} \quad [12]$$

Shift complexity (C_{shift}) measures the complexity in chemical shift variation and is computed through Equations [13] and [14]. Equation [13] centered and scaled the chemical shift vector for each peak ($v'_{ppm}(j)$), and in Equation [14] an average of the normalized sum of squares (NSS , Equation [15]) of $v'_{ppm}(j)$ is used to measure the extent of peak shift for each data set. The more chemical shift varied for each individual peak, the larger C_{shift} . In the calculation of NSS , V is a vector, s is the index in the vector, V_s is one element in the vector, and N_L is the length of the vector.

$$v'_{ppm}(j) = \frac{v_{ppm}^j - \min(V_{ppm})}{\max(V_{ppm}) - \min(V_{ppm})} \quad [13]$$

$$C_{shift} = \frac{1}{M} \sum_j NSS(v'_{ppm}(j)) \quad [14]$$

$$NSS(V) = \frac{1}{N_L} \sum_s \left(V_s - \frac{1}{N_L} \sum_s V_s \right)^2 \quad [15]$$

Scale complexity (C_{scale}) measures the extent to which peak intensities differ from their closest neighbors (Equation [16]). N_{pair} is the number of closest neighbor ridge pairs with average differences in chemical shift less than 0.05. N_{diff} is the number of peak pairs both counted in N_{pair} and with fold change in intensity larger than 10. N is the total number of time points. The greater the difference in intensity between neighboring peaks, the higher the value of C_{scale} . Considerable differences in intensity between nearby peaks result in imperfect peaks shapes and

quantification. Sometimes, the smaller peak is only discernible in a portion of time points because of overlap.

$$C_{scale} = \frac{1}{N(N_{pair})} \sum_i^N N_{diff} \quad [16]$$

Dynamic complexity ($C_{dynamics}$) measures the complexity in intensity variation through time and is computed with Equation [17]. For each ridge, the intensity vector is scaled by its maximum. NSS of the scaled intensity can indicate time dynamics for each ridge, and the value is averaged for all ridges. The more the intensity changes for each peak through time, the higher $C_{dynamics}$. Data set with higher $C_{dynamics}$ have an interesting underlying variation in metabolism, such as metabolic adaptation under different carbon sources (Judge, et al., 2019).

$$C_{dynamics} = \frac{1}{M} \sum_j NSS \left(\frac{X_i(:,j)}{\max(X_i(:,j))} \right) \quad [17]$$

2.7 Data and software

Programs were programmed in MATLAB and R. They are shared through the Edison lab metabolomics toolbox GitHub repository (https://github.com/artedison/Edison_Lab_Shared_Metabolomics_UGA). The used experimental data can be found in Metabolomics Workbench (<https://www.metabolomicsworkbench.org> PR000738) and other used data can be found in Supplementary Data. We also provide a tutorial on the workflow (Supplement File 1).

The programs were extensively run and tested in MATLAB 2018b and R (RStudio Version 1.1.456 and R Version 3.5.1) on a macOS (Mojave 10.14.5) system.

3 Results

3.1 Comparison of simulated and experimental time-series NMR spectra

Complexity in time-series NMR spectra was evaluated in SNR, peak intensity, and chemical shift variation (Table 1). Besides the SNR value in the main simulation, multiple SNR levels were tested, and the workflow can still track ridges accurately in lower SNR levels (Supplementary Figure 6A-D). The peak with changing chemical shift was tracked automatically for $SNR \geq 99.24$ (Supplementary Figure 6A-C) and needed some manual tuning for $SNR = 19.97$ (Supplementary Figure 6D). In practice, most peaks in the experimental spectra possessed good enough SNR for tracking if they could be visually identified. As an example, the valine multiplet centered at ppm 2.267 had low SNR, and 6 peaks of the multiplet could still be tracked (Supplementary Figure 6E). The ridge tracking method had robust performance under a large range of different SNR levels.

For the complexity metrics in variation of intensity (C_{scale} and $C_{dynamics}$) and chemical shift (C_{shift}), the values are similar between experimental and simulated data sets (Table 1). It is also interesting to see that the aerobic sample seems to be more complex than the anaerobic sample, which agrees with our observation that NMR spectra in aerobic data set has larger peak shifting and the intensities are changing more rapidly (Judge, et al., 2019).

Formatted: Font: Not Italic

Deleted: 17

Deleted: 13

Formatted: Font: Not Italic

Deleted: 14

Deleted: 13

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Deleted: 14

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Deleted: 15

Formatted: Font: Not Italic

Deleted: 16

3.2 Performance evaluation of ridge tracking on the simulated data sets

RTEExtract was first tested on the simulated data sets (Supplementary Figure 2). Time-series NMR spectra were simulated with known concentration and chemical shift (simulated value), which were used to evaluate the ridge tracking result (extracted value). Extractions were evaluated by RMSD in chemical shift and peaks were tracked accurately (low RMSD) in simulated data sets (Supplementary Figure 3 and Supplementary Table 2).

Concentration estimation was also evaluated. We plotted extracted concentrations ($[C]_{ex}$) against simulated concentrations ($[C]_{sim}$) and peaks without overlap (red) were all around the diagonal (slope=1) (Figure 3A). Different peaks are affected by line broadening slightly differently and so there are deviations from the diagonal. Nearly half of the ridges were not overlapping and could be accurately quantified.

For those peaks with overlap, quantification was affected. Both spectral changes through time (Figure 3B) and peak composition at selected time points (Figure 3C) are presented. For the alanine peak with overlap (peak 1 in Figure 3A-C), $[C]_{ex}$ was overestimated with a linear curve. This can be explained by the overlap with neighboring glycerol peaks, and both of

Commented [YW1]: Art: What does this mean? This will cause confusion to an NMR person, because the term "line broadening" means something specific.

Is this clear?

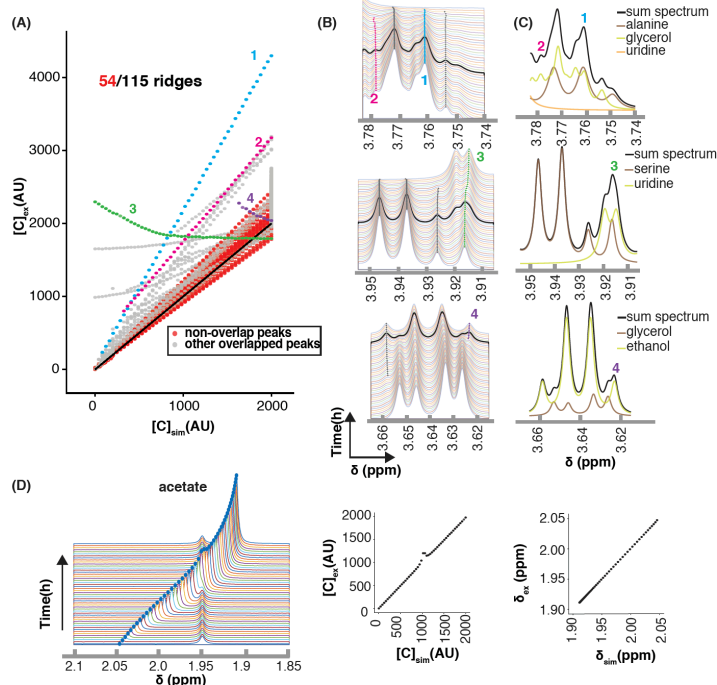


Figure 3: Evaluation of the RTEExtract algorithm on simulated data sets. Performance in estimating intensity and chemical shift was tested with simulated data sets. (A) Compound quantification for all ridge points were plotted. For each ridge peak, extracted compound concentrations normalized to DSS (Y-axis, $[C]_{ex}$) were plotted against simulated compound concentration (X-axis, $[C]_{sim}$). The black diagonal is with slope 1 and works as the perfect quantification reference. Red points are peaks without overlap and points of other colors are with overlap. Among the overlapping ridge peaks, four peaks mapped to four compounds (1: alanine (blue); 2: glycerol (pink); 3: serine (green); 4: ethanol (purple)) were selected as examples of quantification problems when there is overlap (B). Number and color of example peaks are used correspondingly in A-C. All other ridge peaks with overlap are in grey. (B) Tracked ridges of example compounds were plotted in stack spectra. One spectrum (at one specific time point) was highlight in black, which was then decomposed into compound peaks involved in the overlap (C). In the decomposition plot (C), the black line is the real spectrum and the sum of all compound peaks. Other compound spectra are presented in different colors. (D) Effects of overlapping on estimating intensity and chemical shift for the acetate peak. In the first column, the acetate peak was simulated with chemical shift variation under different pH conditions and overlapped with another peak. In the second column, extracted concentrations ($[C]_{ex}$) were plotted against simulated concentrations ($[C]_{sim}$). In the third column, extracted chemical shifts (δ_{ex}) were plotted against simulated chemical shift (δ_{sim}). Performance of concentration and chemical shift estimation with more compounds can be found in Supplementary Figure 4 and Supplementary Figure 5. Accuracy in peak tracking was also evaluated by chemical shift differences between tracked peaks and simulated peaks (Supplementary Table 2). Compound concentration was simulated with arbitrary unit (AU) and chemical shift δ was evaluated in ppm.

them were decreasing through time. A similar explanation exists for the glycerol peaks with overlap (peak 2 in Figure 3A-C). Overlaps with glycerol also caused quantification of the ethanol peak to change in the opposite direction (peak 4 in Figure 3A-C). In this case, the intensity variation of the small side peak from ethanol is dominated by the variation in glycerol peak.

Besides intensity estimation, overlap can affect chemical shift (peak 3 mapped to serine in Figure 3A-C). In this overlapping region, the uridine concentration increased through time, and serine concentration decreased. Even though the ridge was mapped to serine by Equation (10), it was actually mixture of peaks from both compounds (uridine and serine), and there were different relative intensity contributions from both sides at different time points. This kind of continuous shift between two features not only caused incorrect quantification but also a fake chemical shift variation. Even though neither uridine nor serine had clear peak shift for the pH range considered ([4.0, 6.0]), the overlapped peak shifted smoothly. Through the changes of relative intensities of the two peaks, chemical shift of the summed peak changed. This could be mistaken for chemical shift change due to pH variation and seems to occur on the

right-side peak at Figure 5D and glucose peaks (ppm region [5.2, 5.26]) under the aerobic condition in experimental data sets.

In the case that the peaks are separated enough to be distinguishable, ridge tracking is often accurate in chemical shift estimation (Figure 3D). For the acetate peak, concentration ($[C]_{\text{ex}}$) was over-estimated in the overlapped region, but chemical shift (δ_{ex}) was estimated accurately. Both peaks could be tracked for the entire range through the overlapped region, and relative intensity between the two overlapping peaks did not affect tracking capability. Comparison of extracted and simulated values in concentration and chemical shift for more compounds can be found in Supplementary Figure 4 and 5.

3.3 Performance evaluation of ridge tracking on the experimental data sets

The ridge tracking method was next tested on experimental data sets and compared with the approach used in the initial CIVM publication (Judge, et al., 2019). We first assessed the agreement on quantification of regions without much overlap (Figure 4 and Supplementary Figure 7). For quantification under aerobic and anaerobic conditions, the residuals between the two methods were close to zero for most compounds, and RTExtract reproduced the earlier results with much less time and manual input. The

Formatted: Font: 8 pt, Not Italic

Deleted: 10

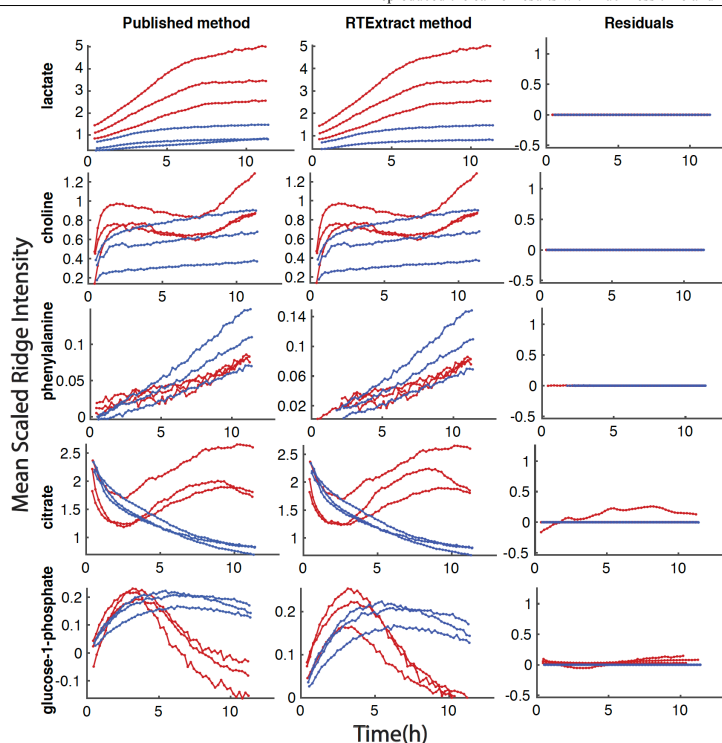


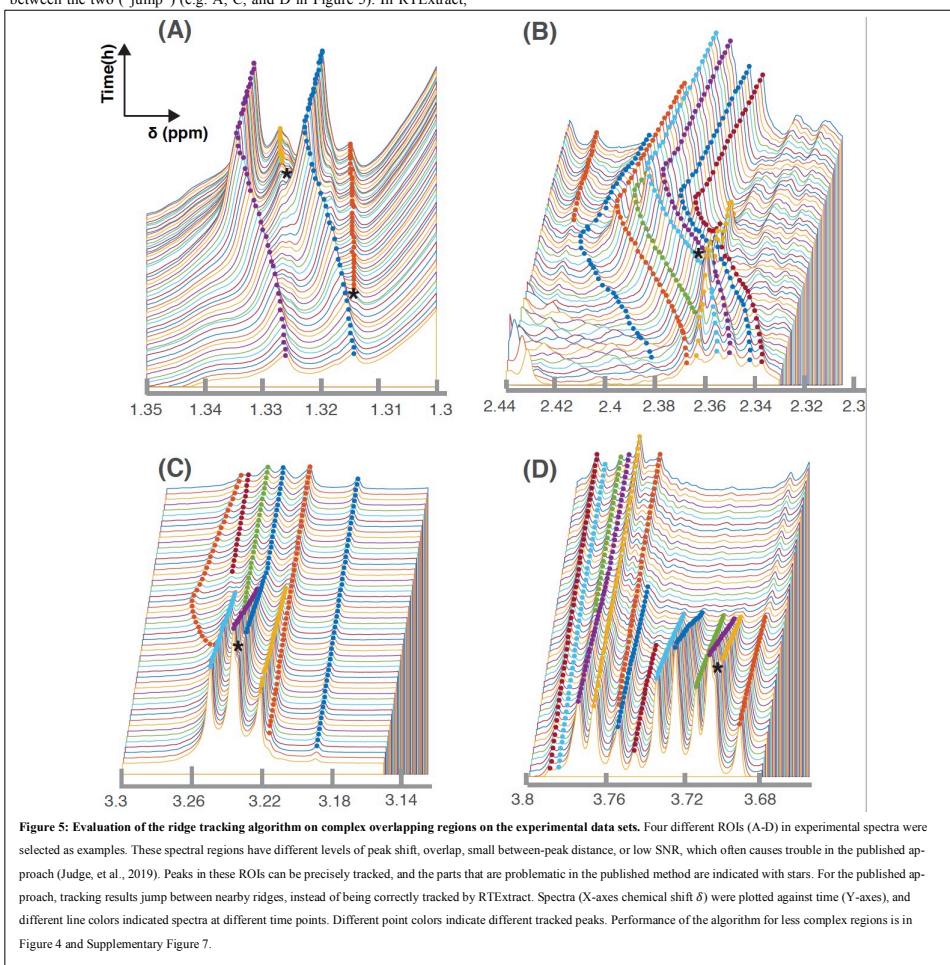
Figure 4: Reproduction of compound quantification results from the published method. In each plot, the X-axis indicates time, and the Y-axis indicates Mean Scaled Ridge Intensity. Red curves are from aerobic data sets, and blue curves are from anaerobic data sets. Details in computing Mean Scaled Ridge Intensity can be found in the previous publication (Judge, et al., 2019). Each row indicates quantification for one compound, including the published and the RTExtract methods and residuals (differences between the quantifications by the two methods). A value near zero in residuals indicates agreement between results of the two methods. Comparison for more compounds can be found in Supplementary Figure 7.

few differences were attributed to how negative values were dealt with between the two methods in computing Mean Scale Ridge Intensity (Figure 4 and Supplementary Figure 7). In RTExtract, time-series data with negative intensities are shifted to positive values first, which was not done in the original publication (Judge, et al., 2019).

Besides on peaks with good peak shapes and little overlap, RTExtract also worked for more complex regions which were difficult for the published method (Figure 5) (Judge, et al., 2019). The complex regions might contain one or more of the following complications: overlap, peak shifting, and peaks with short distances (in chemical shift) to each other. When two peaks are close to each other, especially when one is on the shoulder of the other, the original method often produced tracking results that move between the two ("jump") (e.g. A, C, and D in Figure 5). In RTExtract,

these peaks were tracked with no jumps, resulting in fewer errors in chemical shift and intensity estimation. Parameter tuning, particularly in complex regions, was difficult in the original publication but is much easier in RTExtract. The glutamate region (ppm [2.3, 2.44], Figure 5B) was another difficult case, in which the six peaks from glutamate shifted with pH and overlap with an unknown peak (yellow). By the previous method, only a small side peak in the multiplet was tracked for glutamate, so the quantification had a low SNR (Judge, et al., 2019). By RTExtract, all the six glutamate peaks in the multiplet can be tracked with the retracing approach.

4 Discussion



RTExtract, a computer vision-based approach is introduced in this paper to quantify time-series NMR spectra. RTExtract takes less time and exhibits better performance on complex regions than our original, less automated, approach (Judge, *et al.*, 2019). It provides a more practical way to process time-series NMR spectra and analyze *in vivo* metabolic dynamics of an organism.

RTExtract is an improvement from multiple perspectives. First, we reduced the number of tuning parameters from six to two, which reduces the interactive time and is more intuitive to optimize. Second, the refinement steps allowed fine-tuning of the ridge tracking process and easily removed imperfect regions. Instead of exploring a huge parameter space, the user can fix the imperfect regions through simple manual steps. With these two improvements, the published results can be replicated within 2 hours by RTExtract instead of a few days by the original method (Judge, *et al.*, 2019). Finally, RTExtract is also capable of dealing with more complex regions, especially with peak overlap and peak shifting (Figure 5). It is now possible to track peaks in these difficult regions, without merging multiple peaks into one. Subsequently, more tracked features can be used for downstream statistical analyses.

We still offer the option of manual interaction in the workflow, which helps produce accurate results but still requires time and manual effort. Future versions of the workflow will incorporate statistical filters accompanied by higher degrees of automation. A clustering-based method can be implemented to remove artifact ridges, which are characterized by random changes in intensity and chemical shift. Implementing this step might fully remove the manual procedure and make the full process much faster.

The RTExtract can also combine with spectral deconvolution for overlapping feature quantification. From RTExtract, chemical shift and intensity of individual overlapping peaks can be obtained and subsequently fed into the deconvolution methods. Based on the information of intensity and chemical shift, a Bayesian-based deconvolution approach can compute the underlying peak intensity (Krishnamurthy, 2013).

In principle, as long as peaks are changing in a continuous manner, they can be tracked by RTExtract. The experimental data tested in this paper is from the CIVM method (Judge, *et al.*, 2019), and provides dense, continuous, time-series measurements. Other time-series NMR methods, such as flow NMR can also provide proper candidate measurements (Foley, *et al.*, 2014). Possible applications go beyond time-series measurements as long as the continuity constraint is met between neighboring spectra.

5 Conclusion

RTExtract is introduced in this paper to quantify dense time-series NMR spectra by ridge tracking. It is faster, easier to use, and can deal with more complex regions than previously published methods. The extraction is accurate even in complex overlapping regions. As the ridge tracking method relies on the continuity of peaks at neighboring spectra, it can be further applied to other suitable data types.

Acknowledgements

We thank Heinz-Bernd Schuttler, Leidong Mao, and Juan B. Gutierrez for helpful discussion on the computational method, Peter Kner for discussion on image processing method, John N. Glushka for discussion on NMR spectral processing. Goncalo Gouveia, Amanda Shaver, Mario Uchimiyi, and Nicole Holderman have

provided good suggestions on figures. Laura Morris has provided technical support on computers.

Funding

This work was primarily supported by NSF 1713746. Arthur S. Edison was also supported by the Georgia Research Alliance.

Conflict of Interest: none declared.

References

- Ackerman, J.J., *et al.* The NMR chemical shift pH measurement revisited: analysis of error and modeling of a pH dependent reference. *Magn Reson Med* 1996;36(5):674-683.
- Besl, P.J. and Jain, R.C. Invariant Surface Characteristics for 3d Object Recognition in Range Images. *Comput Vision Graph* 1986;33(1):33-80.
- Besl, P.J. and Jain, R.C. Segmentation through Variable-Order Surface Fitting. *Ieee T Pattern Anal* 1988;10(2):167-192.
- Csenki, L., *et al.* Proof of principle of a generalized fuzzy Hough transform approach to peak alignment of one-dimensional 1H NMR data. *Anal Bioanal Chem* 2007;389(3):875-885.
- Dashti, H., *et al.* Spin System Modeling of Nuclear Magnetic Resonance Spectra for Applications in Metabolomics and Small Molecule Screening. *Analytical chemistry* 2017;89(22):12201-12208.
- Foley, D.A., *et al.* NMR flow tube for online NMR reaction monitoring. *Analytical chemistry* 2014;86(24):12008-12013.
- Haralick, R.M., Watson, L.T. and Laffey, T.J. The Topographic Primal Sketch. *Int J Robot Res* 1983;2(1):50-72.
- Judge, M.T., *et al.* Continuous *in vivo* Metabolism by NMR. *Frontiers in Molecular Biosciences* 2019;6(26).
- Klukowski, P., *et al.* NMRNet: a deep learning approach to automated peak picking of protein NMR spectra. *Bioinformatics* 2018;34(15):2590-2597.
- Klukowski, P., *et al.* Computer vision-based automated peak picking applied to protein NMR spectra. *Bioinformatics* 2015;31(18):2981-2988.
- Koczula, K.M., *et al.* Metabolic plasticity in CLL: adaptation to the hypoxic niche. *Leukemia* 2016;30(1):65-73.
- Krishnamurthy, K. CRAFT (complete reduction to amplitude frequency table) - robust and time-efficient Bayesian approach for quantitative mixture analysis by NMR. *Magn Reson Chem* 2013;51(12):821-829.
- Link, H., Christodoulou, D. and Sauer, U. Advancing metabolic models with kinetic information. *Curr Opin Biotechnol* 2014;29:8-14.
- Link, H., *et al.* Real-time metabolome profiling of the metabolic switch between starvation and growth. *Nature methods* 2015;12(11):1091-1097.
- Ludwig, C. and Gunther, U.L. MetaboLab-advanced NMR data processing and analysis for metabolomics. *BMC bioinformatics* 2011;12:366.
- Montana, G., Berk, M. and Ebbers, T. Modelling short time series in metabolomics: a functional data analysis approach. *Adv Exp Med Biol* 2011;696:307-315.
- Stoker, J.J. Differential geometry. New York: Wiley-Interscience; 1969.
- Suk, M. and Bhandarkar, S.M. Three-dimensional object recognition from range images. Tokyo ; New York: Springer-Verlag; 1992.
- Takis, P.G., *et al.* Deconvoluting interrelationships between concentrations and chemical shifts in urine provides a powerful analysis tool. *Nature Communications* 2017;8(1):1662.
- Tredwell, G.D., *et al.* Modelling the acid/base (1)H NMR chemical shift limits of metabolites in human urine. *Metabolomics* 2016;12(10):152.

NMR feature quantification by ridge tracking

Vu, T.N. and Laukens, K. Getting your peaks in line: a review of alignment methods for NMR spectral data. *Metabolites* 2013;3(2):259-276.

Ye, L., De Iorio, M. and Ebbels, T.M.D. Bayesian estimation of the number of protonation sites for urinary metabolites from NMR spectroscopic data. *Metabolomics* 2018;14(5):56.