



HAL
open science

MultiNMRFit: A software to fit 1D and pseudo-2D NMR spectra

Pierre Millard, Loïc Le Grégam, Svetlana Dubiley, Thomas
Gosselin-Monplaisir, Guy Lippens, Cyril Charlier

► **To cite this version:**

Pierre Millard, Loïc Le Grégam, Svetlana Dubiley, Thomas Gosselin-Monplaisir, Guy Lippens, et al..
MultiNMRFit: A software to fit 1D and pseudo-2D NMR spectra. 2025. hal-04871541

HAL Id: hal-04871541

<https://hal.science/hal-04871541v1>

Preprint submitted on 7 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

MultiNMRFit: A software to fit 1D and pseudo-2D NMR spectra

Pierre Millard^{1,2,*}, Loïc Le Grégam^{1,2}, Svetlana Dubiley¹, Thomas Gosselin-Monplaisir¹,
Guy Lippens¹, Cyril Charlier^{1,*}

¹Toulouse Biotechnology Institute (TBI), Université de Toulouse, CNRS, INRA, INSA, Toulouse, 31077, France

²MetaToul-MetaboHUB, National Infrastructure of Metabolomics and Fluxomics, Toulouse, 31077, France

Summary. Nuclear Magnetic Resonance (NMR) is widely used for quantitative analysis of metabolic systems. Accurate extraction of NMR parameters – such as chemical shift, intensity, coupling constants, and linewidth – is essential for obtaining information on the structure, concentration, and isotopic composition of metabolites. We present MultiNMRFit, an open-source software designed for high-throughput analysis of one-dimensional NMR spectra, whether acquired individually or as pseudo-2D experiments. MultiNMRFit extracts signal parameters (e.g. intensity, area, chemical shift, and coupling constants) by fitting the experimental spectra using built-in or user-defined signal models that account for multiplicity, providing high flexibility along with robust and reproducible results. The software is accessible both as a Python library and via a graphical user interface, enabling intuitive use by end-users without computational expertise. We demonstrate the robustness and flexibility of MultiNMRFit on datasets collected in metabolomics and isotope labeling studies.

Availability and Implementation. MultiNMRFit is implemented in Python 3 and was tested on Unix, Windows, and MacOS platforms. The source code and the documentation are freely distributed under GPL3 license at <https://github.com/NMRTeamTBI/MultiNMRFit/>.

Supplementary data. Supplementary data are available online.

Contact. Cyril Charlier (charlier@insa-toulouse.fr) and Pierre Millard (millard@insa-toulouse.fr)

32 **1. Introduction**

33
34 Nuclear Magnetic Resonance (NMR) spectroscopy is a powerful tool for investigating metabolism.
35 Key applications include identifying and quantifying metabolites in metabolomics, determining the
36 isotopic content of metabolites in ^{13}C -fluxomics, and identifying and characterizing protein-
37 metabolite interactions [1–4]. Extracting meaningful information from NMR spectra requires precise
38 characterization of each signal of interest, including their position (chemical shift), intensity (and
39 area) and shape (coupling constants and linewidth) [5]. Signal parameters can be estimated by fitting
40 spectra using various mathematical models that relate these parameters to the observed signals. This
41 fitting step is often achieved with in-house scripts, which, while functional, may lack the robustness
42 and reproducibility necessary for large-scale or complex NMR studies of metabolic systems.
43 Commercial software, such as TopSpin (Bruker) and Mnova (Mestrelab), provide alternative
44 solutions, offering a user-friendly interface but functioning as proprietary “black boxes” and restricted
45 to a limited set of signal models. Open-source R and Python packages (such as *MetaboDeconID* [6]
46 and *BATMAN* [7,8]) have also been developed to enhance accessibility and reproducibility. However,
47 they are designed for programmatic use and do not include certain features essential for users without
48 computational expertise. Such features include statistical methods for calculating confidence intervals
49 for extracted parameters, a graphical user interface, and a logging system to save calculation
50 parameters and process information along with the results. Another key challenge is that most existing
51 tools treat each peak as an independent signal. However, NMR signals often consist of multiple peaks
52 arising from homo- and heteronuclear scalar couplings, which are intrinsic properties crucial for
53 identifying metabolites and quantifying their isotopic content. Considering signal multiplicity can also
54 improve parameter estimation in cases of overlapping peaks. However, most current tools do not
55 account for this complexity, which can limit their effectiveness in extracting and utilizing this
56 information. Consequently, the detailed analysis of NMR spectra remains a low throughput task,
57 typically performed by expert users, which can constrain the analysis of large datasets, such as those
58 generated in real-time NMR experiments. The development of user-friendly, open-source software
59 that incorporates signal multiplicity and enables flexible implementation of tailored signal models
60 would significantly enhance the scope, efficiency, and reproducibility of NMR-based metabolic
61 studies.

62 To address these challenges, we have developed MultiNMRFit, a versatile Python software for fitting
63 one dimensional NMR spectra, whether acquired individually or as pseudo-2D datasets in time-course
64 experiments. MultiNMRFit can be applied to any nucleus (^1H , ^{13}C , ^{15}N , ^{31}P , *etc.*), any type of 1D or
65 pseudo-2D NMR experiment, and any type of signal. It includes built-in signal models for common
66 multiplets and is easily extendable with user-defined models to process less conventional signals.
67 MultiNMRFit is accessible both as a Python library and via a graphical user interface, allowing
68 biologists without computational expertise to use it effectively. We demonstrate the robustness of

69 MultiNMRFit on synthetic data and through its application to real-time metabolomics and isotope
70 labeling experiments.

71

72 **2. Methods and implementation**

73 The general workflow of MultiNMRFit is shown in Figure 1. MultiNMRFit is a Python-based library
74 that operates within common web browsers via a Streamlit-powered graphical user interface. Users
75 can load 1D and pseudo-2D Bruker spectra pre-processed beforehand (*i.e.* referenced, phase-corrected
76 and baseline-corrected). Spectra can also be provided as text files containing chemical shifts and
77 intensities. MultiNMRFit performs automated peak picking within a specified spectral region, and
78 users can add additional peaks manually. One or several peaks can then be grouped into a signal, and
79 a model must be selected for each signal. As detailed in the next section, each model contains
80 equations to simulate the corresponding signal based on parameters. MultiNMRFit automatically
81 suggests a list of models based on the number of peaks (e.g., *doublet of doublets* and *quartet* for
82 signals containing four peaks) and initializes parameters based on the peak list. Note that a given peak
83 can be included in multiple signals to account for overlapping peaks. For spectrum fitting, parameters
84 p are estimated by minimizing the cost function $c(p)$, which is defined as the sum of squared
85 residuals between the experimental intensities (a_i) and the values simulated by the model $y_i(p)$:

$$c(p) = \sum_i (a_i - y_i(p))^2$$

86 The cost function is minimized using the L-BFGS-B method [9], with an optional initial refinement
87 step performed by the differential evolution algorithm [10]. Standard deviations on parameters are
88 estimated from the covariance matrix. For batch processing of spectra, users can select a spectrum as
89 reference to automatically initialize signals and parameters values for the next spectra, thereby
90 automating the processing of multiple spectra for high-throughput analysis. Finally, MultiNMRFit
91 provides a range of interactive plots for visual inspection of fitting results. The results can be exported
92 as *csv* files containing (*i*) parameters values (with their confidence intervals) and (*ii*) chemical shifts
93 and intensities of experimental and fitted spectra.

94 The complete process is automatically saved as a pickle file, which can be reloaded in MultiNMRFit
95 to continue previous work or reanalyze calculation results. This file can also be shared, enabling
96 collaborative analysis across multiple users and ensuring reproducibility. Detailed usage of
97 MultiNMRFit can be found in the online documentation (<https://MultiNMRFit.readthedocs.io>).

98

99 **3. Signal models**

100 A signal model consists of (*i*) equations that describe signal intensity as a function of chemical shift
101 (used for simulations) and (*ii*) a list of parameters with their default values and bounds (used for
102 optimization). Various mathematical functions, including Lorentzian, Gaussian, and Voigt functions,

103 can be used to simulate the NMR spectrum of a given signal. For example, the intensity y_i of a
104 doublet at chemical shift x_i can be described using the following Gaussian function:

105

$$y_i = I/2 \cdot \left(e^{-(x_i - x_0 + J/2)^2 / (2 \cdot l_w^2)} + e^{-(x_i - x_0 - J/2)^2 / (2 \cdot l_w^2)} \right)$$

106

107 with parameters x_0 (chemical shift of the signal), I (total signal intensity), J (coupling constant), and
108 l_w (linewidth at half maximum). An NMR spectrum containing multiple signals can be simulated by
109 summing the spectra of individual signals.

110 By default, MultiNMRFit provides mixed Gaussian-Lorentzian models for common multiplets,
111 including singlet, doublet, triplet, doublet of doublets, and quadruplet. Advanced users can create
112 custom models to handle less conventional signals, such as signals with other coupling patterns or
113 described using alternative mathematical functions. A model template is included with MultiNMRFit
114 for this purpose, and a tutorial on model construction is available in the documentation.

115

116 **4. Validation and application to metabolomics and isotopic datasets**

117 We validated MultiNMRFit using synthetic datasets (Supplementary data). Results confirm the
118 capability of the software to handle complex spectral data and deliver robust parameter estimation
119 even under challenging conditions (such as low signal-to-noise ratio or significant signal overlaps).

120 Then, we used MultiNMRFit to monitor the dynamic conversion of glucose through the initial steps
121 of glycolysis. Glucose was added to a buffer solution containing TSP (sodium trimethylsilyl
122 propionate, used as internal standard), ATP and the first three enzymes of glycolysis: glucokinase,
123 glucose-6-phosphate isomerase, and phosphofructokinase (Figure 2A, Supplementary data). The time-
124 course signals of glucose (GLC), glucose-6-phosphate (G6P), fructose-6-phosphate (F6P), fructose-
125 1,6-bisphosphate (FBP), ATP and ADP were monitored in real-time using a pseudo-2D experiment
126 consisting in 1024 1D ^1H spectra. All signal parameters were successfully extracted, including for
127 overlapped signals of ADP and ATP (Figure 2B). The integrals of signals specific to each metabolite
128 were normalized to the TSP signal (Figure 2C). A progressive decrease in GLC signal was observed,
129 accompanied by a transient accumulation of G6P and F6P, which were eventually converted into
130 FBP. The high precision and resolution on the estimated dynamics illustrate the value of
131 MultiNMRFit for high-throughput processing of large datasets, with the relative concentrations of the
132 six metabolites of interest quantified in less than three hours for the 1024 spectra.

133 MultiNMRFit can also exploit the multiplicity resulting from the presence of isotopes in metabolites
134 to quantify their isotopic content – a key capability for applications such as detecting counterfeit
135 chemicals, identifying metabolic pathways, and quantifying metabolic fluxes [11,12]. To demonstrate
136 this, we conducted a ^{13}C -isotope labeling experiment by growing *Escherichia coli* on a mixture of
137 glucose (15 mM, 80% $1\text{-}^{13}\text{C}_1\text{-glucose}$ + 20% $\text{U-}^{13}\text{C}_6\text{-glucose}$) and acetate (8 mM, 100% $2\text{-}^{13}\text{C}_1\text{-}$

138 acetate). Samples were collected each hour to quantify the dynamics of all isotopic forms of acetate
139 (Supplementary data). In the ^1H NMR spectra, the methyl protons of acetate may give rise to nine
140 peaks corresponding to the four isotopic species of this compound: a singlet for $^{12}\text{C}_2$ -acetate, a doublet
141 with a small (6 Hz) 2J coupling constant for $1\text{-}^{13}\text{C}_1$ -acetate, a doublet with a large (127 Hz) 1J
142 coupling constant for $2\text{-}^{13}\text{C}_1$ acetate, and a doublet of doublet for $\text{U-}^{13}\text{C}_2$ -acetate. The results showed
143 that the total acetate concentration remained stable throughout the experiment (8.5 ± 0.1 mM), which
144 could be interpreted as an acetate flux close to zero (Figure 2D-G). However, while the provided
145 isotopic form ($2\text{-}^{13}\text{C}_1$ -acetate) initially represents 99% of the total acetate pool, its concentration
146 decreased over time (from 8.5 to 6.5 mM). Concurrently, the concentration of the other labeled forms
147 (derived from glucose) increased (from 0.1 to 1.9 mM). These results support previous findings [13]
148 that acetate is continuously exchanged between cells and their environment, even in the absence of net
149 acetate accumulation in the medium. This example illustrates the applicability of MultiNMRFit for
150 extracting quantitative isotopic information on metabolites, providing meaningful biological insights.

151

152 **Conclusion**

153 We present MultiNMRFit, an open-source Python software designed to extract detailed quantitative
154 parameters from NMR spectra. MultiNMRFit is applicable to spectra acquired from any nucleus,
155 whether as 1D or pseudo-2D spectra, and can handle any type of multiplicity through a flexible
156 system for implementing NMR signal models. MultiNMRFit is accessible to non-expert users through
157 a graphical user interface and can also be used as a Python package, enabling seamless integration
158 into metabolomics and fluxomics workflows. Furthermore, MultiNMRFit holds significant potential
159 for other NMR applications, including protein studies, the pharmaceutical industry, biotechnology,
160 and chemistry.

161

162 **Acknowledgements**

163 We thank Marin Martin Mawah, Pauline Rouan and Dr. Valeria Gabrielli for testing previous versions
164 of the software and for insightful discussions. We thank Dr. Edern Cahoreau and Lindsay Peyriga
165 (MetaToul, Toulouse metabolomics & fluxomics facilities, www.metatoul.fr) for access to NMR
166 facilities. MetaToul is part of MetaboHUB, the French National Infrastructure for Metabolomics and
167 Fluxomics (www.metabohub.fr).

168

169 **Author contributions**

170 **PM** & **CC** Software, Project administration, Supervision, Conceptualization, Writing - review &
171 editing; **LL** Software; **SD** & **TM** Resources; **GL** Conceptualization, Writing - review & editing.

172

173 **Competing interests**

174 The authors declare that they have no conflict of interest.

175

176 References

177

- 178 [1] K.M. Koczula, C. Ludwig, R. Hayden, L. Cronin, G. Pratt, H. Parry, D. Tennant, M. Drayson, C.M. Bunce,
179 F.L. Khanim, U.L. Günther, Metabolic plasticity in CLL: adaptation to the hypoxic niche, *Leukemia* 30
180 (2016) 65–73. <https://doi.org/10.1038/leu.2015.187>.
- 181 [2] S. Moco, Studying Metabolism by NMR-Based Metabolomics, *Front. Mol. Biosci.* 9 (2022) 882487.
182 <https://doi.org/10.3389/fmolb.2022.882487>.
- 183 [3] S. Nargund, M.E. Joffe, D. Tran, V. Tugarinov, G. Sriram, Nuclear Magnetic Resonance Methods for
184 Metabolic Fluxomics, in: H.S. Alper (Ed.), *Systems Metabolic Engineering*, Humana Press, Totowa, NJ,
185 2013: pp. 335–351. https://doi.org/10.1007/978-1-62703-299-5_16.
- 186 [4] M. Diether, Y. Nikolaev, F.H. Allain, U. Sauer, Systematic mapping of protein–metabolite interactions in
187 central metabolism of *Escherichia coli*, *Molecular Systems Biology* 15 (2019) e9008.
188 <https://doi.org/10.15252/msb.20199008>.
- 189 [5] A.S. Edison, M. Colonna, G.J. Gouveia, N.R. Holderman, M.T. Judge, X. Shen, S. Zhang, NMR: Unique
190 Strengths That Enhance Modern Metabolomics Research, *Anal. Chem.* 93 (2021) 478–499.
191 <https://doi.org/10.1021/acs.analchem.0c04414>.
- 192 [6] M. Häckl, P. Tauber, F. Schweda, H.U. Zacharias, M. Altenbuchinger, P.J. Oefner, W. Gronwald, An R-
193 Package for the Deconvolution and Integration of 1D NMR Data: MetaboDecon1D, *Metabolites* 11 (2021)
194 452. <https://doi.org/10.3390/metabo11070452>.
- 195 [7] J. Hao, W. Astle, M. De Iorio, T.M.D. Ebbels, BATMAN—an R package for the automated quantification
196 of metabolites from nuclear magnetic resonance spectra using a Bayesian model, *Bioinformatics* 28 (2012)
197 2088–2090. <https://doi.org/10.1093/bioinformatics/bts308>.
- 198 [8] J. Hao, M. Liebeke, W. Astle, M. De Iorio, J.G. Bundy, T.M.D. Ebbels, Bayesian deconvolution and
199 quantification of metabolites in complex 1D NMR spectra using BATMAN, *Nat Protoc* 9 (2014) 1416–
200 1427. <https://doi.org/10.1038/nprot.2014.090>.
- 201 [9] R.H. Byrd, P. Lu, J. Nocedal, C. Zhu, A Limited Memory Algorithm for Bound Constrained Optimization,
202 *SIAM J. Sci. Comput.* 16 (1995) 1190–1208. <https://doi.org/10.1137/0916069>.
- 203 [10] R. Storn, K. Price, Simple and Efficient Heuristic for global Optimization over Continuous Spaces, *Journal*
204 *of Global Optimization* 11 (1997) 341–359. <https://doi.org/10.1023/A:1008202821328>.
- 205 [11] M.V.D.M. Ribeiro, N. Boralle, H.R. Pezza, L. Pezza, Authenticity assessment of anabolic androgenic
206 steroids in counterfeit drugs by ¹H NMR, *Anal. Methods* 10 (2018) 1140–1150.
207 <https://doi.org/10.1039/C8AY00158H>.
- 208 [12] B. Christensen, A. Karoly Gombert, J. Nielsen, Analysis of flux estimates based on ¹³C–labelling
209 experiments, *European Journal of Biochemistry* 269 (2002) 2795–2800. <https://doi.org/10.1046/j.1432-1033.2002.02959.x>.
- 210 [13] B. Enjalbert, P. Millard, M. Dinclaux, J.-C. Portais, F. Létisse, Acetate fluxes in *Escherichia coli* are
211 determined by the thermodynamic control of the Pta-AckA pathway, *Sci Rep* 7 (2017) 42135.
212 <https://doi.org/10.1038/srep42135>.
- 213
- 214
- 215

216 Figure Legends

217

218 Figure 1. Flow chart of MultiNMRFit.

219 **Figure 2. Applications of MultiNMRFit to metabolomics and isotope labeling studies.** (A)
220 Scheme of the biochemical reactions monitored in a real-time NMR experiment. (B) Dynamic of
221 integrals of glucose (GLC, blue), glucose-6-phosphate (G6P, red), fructose-6-phosphate (F6P,
222 orange), fructose-1,6-bisphosphate (FBP, purple), adenosine triphosphate (ATP, gray), and adenosine
223 diphosphate (ADP, green). Integrals were normalized with the TSP signal. (C) Examples of signals
224 processed with MultiNMRFit for each molecule of interest, where individual signals are shown by the
225 shaded areas and the experimental and fitted spectra are represented by dots and lines, respectively.
226 (D) Isotopic forms of acetate. (E) Theoretical ¹H spectrum of the methyl protons of acetate, with a
227 singlet for ¹²C₂-acetate (blue), a doublet for 1-¹³C₁-acetate (orange), a doublet for 2-¹³C₁ acetate
228 (green), and a doublet of doublets for U-¹³C₂-acetate (red). (F) Example of spectrum processed with
229 MultiNMRFit with the experimental spectrum shown as dots, the fitted spectrum as a line, and
230 individual signals as shaded areas. (G) Time-course concentrations of the four isotopic forms of
231 acetate (using the color code shown in panel E) and of the total acetate concentration (in grey).

232



