Ultrahigh-Resolution NMR with No Signal Loss

Simon G. Hulse,[a] Mohammadali Foroozandeh*[a], Gareth A. Morris[b], Mathias Nilsson[b]

We present CUPID (Computer-assisted Undiminished-sensitivity Protocol for Ideal Decoupling), a technique which uses parametric estimation to produce pure shift spectra from easy-to-acquire J-resolved datasets without loss of sensitivity, and with absorption-mode lineshapes. We show that it is effective even in scenarios where state-of-the-art pure shift experimental techniques are too insensitive, and where undesired solvent/impurity peaks overlap with signals of interest. The method also enables the extraction of individual multiplets from overlapping spectra, based on a simple criterion involving the relationship between a particular signal's direct and indirect dimension frequencies. The functionality required to use CUPID is provided by NMR Estimation in Python (NMR-EsPy), an open-source package with a simple-to-use API and supporting graphical interface.

2D J-resolved (2DJ) spectroscopy[1,2] is a valuable technique for separating chemical shift and scalar coupling information, and can be used to acquire broadband homodecoupled (pure shift) spectra. This is achieved by applying a 45° shear to the spectral data such that $S(f^{(1)}, f^{(2)}) \rightarrow S(f^{(1)}, f^{(2)} - f^{(1)})$. This transformation leaves chemical shifts and scalar couplings dispersed orthogonally to each other, along the $F^{(2)}$ and $F^{(1)}$ axes respectively. A summation along $F^{(1)}$ ought then to yield the pure shift spectrum.

Unfortunately, J-resolved datasets are hypercomplex (i.e. phase-modulated with respect to both time dimensions), making it impossible to generate a spectrum with desirable absorption-mode Lorentzian lineshapes in this way. Direct summation of the spectrum, in which signals have "phasetwist" lineshapes, leads in fact to zero signal, as positive and negative contributions cancel each other out. It is therefore necessary to process 2DJ spectra using magnitude (absolute value) mode, leading to pure shift spectra in which peaks have broad wings that contain both absorption- and dispersion-mode contributions. Further distortions arise where the dispersion-mode tails of peaks interefere. Post-processing techniques such as pseudo-echo reshaping[3] and sine-bell apodisation[4] can help suppress the dispersive components, but only at a high cost in both sensitivity and resolution.

There has been considerable development of other pulse sequences for pure shift NMR, which typically involve concatenating the initial "chunks" of FIDs acquired in a 2D mode experiment. Prominent examples involve the use of the Zangger-Sterk (ZS)[5,6], Bilinear rotational Decoupling (BIRD)[7,8], and Pure Shift Yielded by Chirp Excitation (PSYCHE)[9,10] pulse sequence elements. These techniques all incur a significant cost in sensitivity, albeit for different reasons, as only a subset of the available spin magnetisation contributes to the final signal (see[11] for an overview).

We introduce here a new method to generate pure shift spectra, which utilises NMR Estimation in Python (NMR-EsPy), a free and open source package we recently presented[12]. By estimating the parameters that describe the 2DJ FID, it is possible to generate pure shift spectra with absorption-mode lineshapes, with similar sensitivity to a simple pulse-acquire experiment. We name the technique CUPID (Computer-assisted Undiminished-sensitivity Protocol for Ideal Decoupling). Using parametric estimation as a route to pure shift spectra also provides extra benefits, such as the ability to identify individual multiplet structures, and to remove unwanted signals such as solvent peaks or peaks that arise from strong coupling effects[13,14].

A 2DJ dataset can be modelled as a summation of exponentially damped hypercomplex oscillators, with each oscillator defined by six parameters: amplitude (a), phase (ϕ), two frequencies ($f^{(1)}, f^{(2)}$), and two damping factors ($\eta^{(1)}, \eta^{(2)}$), such that a signal comprising M oscillators can be fully described by $6M$ parameters. A technique has been previously presented which estimates the $4M$ parameters associated with 1D NMR signals in the software package NMR-EsPy, using numerical optimisation to minimise a cost function quantifying how well a given set of parameters describes the underlying data[12]. NMR-EsPy has here been extended to support the estimation of hypercomplex 2D datasets; more detail of the underlying theory is provided in Section A of the supplementary material. One noteworthy feature of the technique is that it typically requires the generation of frequency-filtered "sub-FIDs" in the direct dimension, as the computational requirements scale rapidly with (a) the number of data points, and, especially, (b) the model order M. A similar segmentation of frequency space is used, for the same reasons, in the Filter Diagonalisation Method [add citation of JMR 234,125].

The key goal of CUPID here is to generate a 1D FID from the estimated parameters, named for historical reasons the "$-45°$ signal" [1] $\boldsymbol{y}_{-45°} \in \mathbb{C}^N$:

$$\boldsymbol{y}_{-45°}[n] = \sum_{m=1}^{M} a_m e^{\mathrm{i}\phi_m} e^{\left(2\pi\mathrm{i}\left(f_m^{(2)} - f_m^{(1)}\right) - \eta_m^{(2)}\right) n\Delta t}, \qquad (1)$$

$\forall n \in \{0, \cdots, N-1\}$. This signal takes the same mathematical form as a conventional FID acquired by a pulse-acquire experiment, except that the usual multiplet component frequencies $\{f_m^{(2)}\}$ $\forall m = 1, \cdots, M$ are replaced with the chemical shifts $\{f_m^{(2)} - f_m^{(1)}\}$. Oscillators which belong to a particular multiplet have frequencies $f_m^{(1)} = f_\mathrm{d}$ and $f^{(2)} = f_\mathrm{c} + f_\mathrm{d}$ in 2DJ datasets, where f_c is the multiplet's

[a] Simon G. Hulse, Mohammadali Foroozandeh*
Chemistry Research Laboratory, University of Oxford, 12 Mansfield Rd, Oxford OX1 3TA
E-mail: mohammadali.foroozandeh@chem.ox.ac.uk

[b] Gareth A. Morris, Mathias Nilsson
School of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, UK

[1] See Section A.6 of the supplementary material for further details.

central frequency (the chemical shift in frequency units), and f_d is the displacement of a given oscillator from the central frequency[2]. Hence, all oscillators belonging to the same multiplet will contribute a component with the multiplet's central frequency (f_c) to the $-45°$ signal. It is worth restating here that under the assumption of successful extraction of the signal parameters of the 2DJ dataset, Equation 1 leads to spectra with absorption Lorentzian lineshapes with complete conservation of signal intensity.

The identification of multiplet structures can be achieved by grouping together oscillators in an estimation result, such that any pair of oscillators $\{i, j\}$ that are part of the same multiplet lie along a $+45°$ diagonal in frequency space and therefore satisfy

$$\left|\left(f_i^{(2)} - f_i^{(1)}\right) - \left(f_j^{(2)} - f_j^{(1)}\right)\right| < \epsilon, \qquad (2)$$

where ϵ is a threshold to account for uncertainty in the estimation. A sensible threshold to set for ϵ is the sampling rate in the more poorly resolved dimension. In practice, this can be a little too optimistic, and ϵ sometimes needs to be slightly increased to resolve multiplets effectively.

The process we describe has similarities with the work of Nuzillard[15,16], and of Mutzenhardt et al.[17], in which 1D estimation is used to determine signal parameters for each slice in either $t^{(1)}$ (Nuzillard) or $t^{(2)}$ (Mutzendardt et al.). These parameters are then used to back-propagate the experimental data into negative $t^{(1)}$ or $t^{(2)}$, generating a "full echo". After Fourier transformation to produce an absorption-mode 2D spectrum this can be sheared and summed as described earlier. The use of holistic 2D, rather than 1D, estimation here opens up the possibility of (a) estimating more complex datasets, as individual oscillators are more easily distinguished, and (b) identifying individual 1D multiplet structures, which requires knowledge of oscillator frequencies in both dimensions.

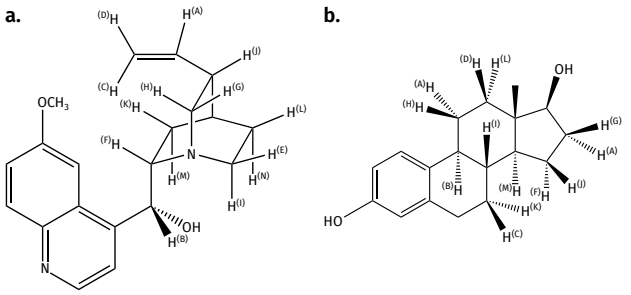


Figure 1. The structures of **a.** quinine, **b.** 17β-estradiol.

We present here two examples of the application of CUPID, to the following samples (Figure 1):

- Quinine in CD_3OD (Figure 2).
- Low concentration (2 mM) 17β-estradiol in DMSO-d_6 (Figure 3). A PSYCHE spectrum was also acquired for comparison.

Further examples are presented in Section B of the supplementary material, including results for simulated data that show that in the limit of a "perfect" estimation result, a spectrum which agrees exactly with the expected pure shift spectrum is achievable. After the generation of each result, certain oscillators were removed from the set of parameters. The oscillators in question were typically spurious and of low-intensity. These tend to arise in situations where a slight over-fitting of the data has occurred. Most of these could be removed automatically, based on a couple of simple criteria outlined in the supplementary material (Section A.7). The updated results re-underwent optimisation after editing. While these edits did improve the results by removing spurious oscillators, the original results still successfully generated homodecoupled spectra with only a small differences to the final spectra presented here.

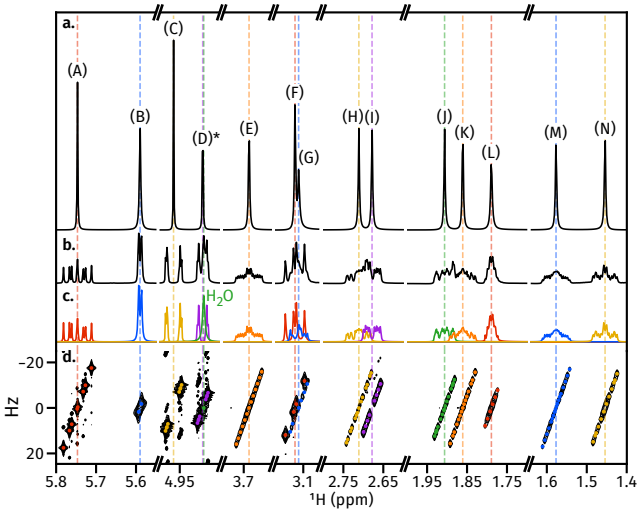


Figure 2. Application of CUPID to the the non-aromatic regions of a quinine dataset. **a.** Pure shift spectrum generated using CUPID, with the signal arising from residual water, overlapping with the multiplet of proton (D), removed. **b.** Conventional 1D spectrum (produced by Fourier transformation of the first increment of the 2DJ dataset). **c.** Multiplets extracted with CUPID using a threshold $\epsilon = 0.92$ Hz, the digital resolution in the direct dimension. **d.** Experimental 2DJ spectrum, displayed in magnitude mode, with the chemical shifts of the assigned peaks marked by coloured vertical lines.

The quinine data of Figure 2 provide an example of the suppression of an undesired signal in the final pure shift spectrum. An intense broad singlet arising from residual water was detected with a very similar frequency to the central frequency of the multiplet arising from proton D. The set of parameters describing the water resonance was simply removed prior to construction of the pure shift spectrum, leaving only signals derived from proton D. A similar approach can also be applied to remove strong coupling artefacts from the final spectrum. An example of this is shown in the Supplementary Material, where a 2DJ dataset for camphor is considered.

The application of CUPID to the dilute estradiol example shows that it is able to cope with highly complex overlapping multiplet structures, even in a signal-to-noise regime where PSYCHE is approaching the point of being too insensitive to generate usable spectra.

CUPID works even in scenarios where the estimation routine has under-fit a particular multiplet structure. However, the resulting pure shift peak tends to be broadened beyond the "true" linewidth in these situations. This effect can be seen for peak (B) in the estradiol result. These under-fitting phenomena do not tend to significantly perturb the integrals of the peaks however.

[2] for weak coupling, f_d will be some linear combination of the scalar couplings associated with a given spin, with all coefficients in the combination being $\pm 1/2$

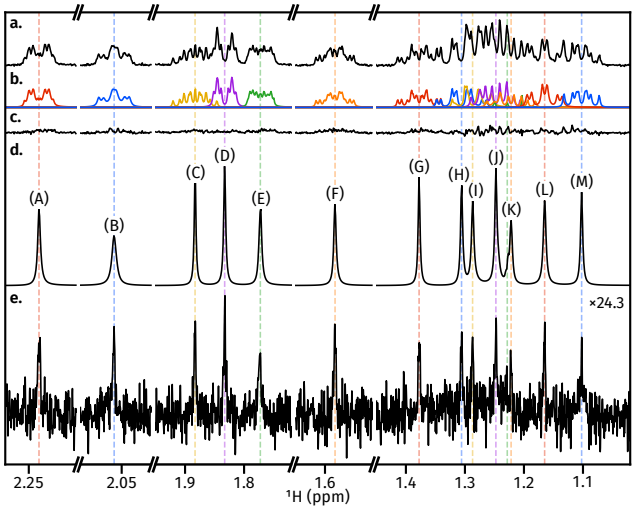


Figure 3. The result of applying CUPID to a 2DJ dataset for 2 mM estradiol in DMSO-d_6. **a.** Conventional 1D spectrum, generated from the FT of the first increment of the 2DJ dataset. **b.** Multiplets extracted ($\epsilon = 2$ Hz). **c.** Residual for the first increment of the dataset, obtained by subtracting the sum of the signals in panel b from the signal in panel a. **d.** Pure shift spectrum generated by FT of the $-45°$ signal. **e.** PSYCHE spectrum for the same sample, scaled so that its most intense point has same magnitude as the equivalent point in the CUPID spectrum.

All the results presented in this paper were generated using the NMR-EsPy Python package. The typical CUPID workflow is as follows:

1. Acquire a conventional J-resolved dataset for the sample of interest. As a rule of thumb, it is recommended to have an indirect dimension of at least 64 incerments. If time permits, or for particularly complex cases, such as estradiol, more points are advisable.
2. Import the data[3], phase the direct dimension for the first increment, and (optionally) baseline correct [18].
3. Select direct-dimension regions of interest, generate filtered sub-FIDs for each region, and estimate the parameters for each sub-FID.
4. (Optional) Edit to correct spurious features in the estimation result if required (removing spurious oscillators), and re-run the optimisation.
5. Construct the pure shift spectrum and, if desired, spectra for individual multiplets, and export to the desired format[3].

From version 2.0, NMR-EsPy makes it easy to perform all the above data processing (points 2 to 5), either with Python scripts making use of the `Estimator2DJ` object, or via a graphical user interface which can be loaded either from the command line or from within TopSpin. The reader is directed to the documentation, at https://foroozandehgroup.github.io/NMR-EsPy/, for details of usage. The project's source code is hosted on GitHub at https://github.com/foroozandehgroup/NMR-EsPy/. Queries on usage, bug reports, and suggestions for features/improvements to add are all welcome.

This work has presented CUPID, a procedure for generating pure shift spectra from 2DJ datasets by parametric estimation. CUPID can produce spectra with absorption-mode lineshapes with zero signal loss, and can also extract multiplet structures from crowded spectral regions. It has been shown to be effective even where state-of-the-art pure shift methods are too insensitive to provide useful results.

This work was funded by the Royal Society (grant numbers URF\R1\180233 and RGF\EA\181018) and by the Engineering and Physical Sciences Research Council (grant numbers EP/L018500 and EP/M013820). SGH would like to thank Jonathan Yong and James Montgomery for assistance in acquiring the datasets, and Timothy Claridge and Fay Probert for useful discussions.

Keywords: Pure shift NMR • Estimation • Python

Author Contributions

Simon Hulse: Methodology, Software, Validation, Formal analysis, Data Curation, Writing - Original Draft, Visualization. **Mohammadali Foroozandeh:** Conceptualization, Methodology, Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition. **Gareth Morris:** Initial conceptualization, Writing - Review & Editing. **Mathias Nilsson:** Initial conceptualization, Writing - Review & Editing.

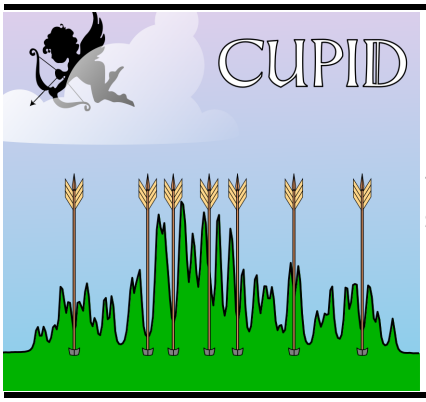
References

[1] W. Aue, J. Karhan, R. Ernst, *The Journal of Chemical Physics* **1976**, *64*, 4226.
[2] G. A. Morris, *Two-Dimensional J -Resolved Spectroscopy*, John Wiley & Sons, Ltd **2009**.
[3] A. Bax, R. Freeman, G. A. Morris, *Journal of Magnetic Resonance (1969)* **1981**, *43*, 333.
[4] J. Lindon, A. Ferrige, *Progress in Nuclear Magnetic Resonance Spectroscopy* **1980**, *14*, 27.
[5] K. Zangger, H. Sterk, *Journal of Magnetic Resonance* **1997**, *124*, 486.
[6] J. Aguilar, S. Faulkner, M. Nilsson, G. Morris, *Angewandte Chemie International Edition* **2010**, *49*, 3901.
[7] J. Garbow, D. Weitekamp, A. Pines, *Chemical Physics Letters* **1982**, *93*, 504.
[8] A. Bax, *Journal of Magnetic Resonance (1969)* **1983**, *53*, 517.
[9] M. Foroozandeh, R. W. Adams, N. J. Meharry, D. Jeannerat, M. Nilsson, G. A. Morris, *Angewandte Chemie International Edition* **2014**, *53*, 6990.
[10] M. Foroozandeh, G. A. Morris, M. Nilsson, *Chemistry – A European Journal* **2018**, *24*, 13988.
[11] K. Zangger, *Progress in Nuclear Magnetic Resonance Spectroscopy* **2015**, *86-87*, 1.
[12] S. G. Hulse, M. Foroozandeh, *Journal of Magnetic Resonance* **2022**, *338*, 107173.
[13] G. Wider, R. Baumann, K. Nagayama, R. R. Ernst, K. Wüthrich, *Journal of Magnetic Resonance (1969)* **1981**, *42*, 73.
[14] M. J. Thrippleton, R. A. Edden, J. Keeler, *Journal of Magnetic Resonance* **2005**, *174*, 97.
[15] J.-M. Nuzillard, *Journal of Magnetic Resonance, Series A* **1996**, *118*, 132.
[16] A. Martinez, F. Bourdreux, E. Riguet, J.-M. Nuzillard, *Magnetic Resonance in Chemistry* **2012**, *50*, 28.

[3] Currently, only data in Bruker format is supported

[17] P. Mutzenhardt, F. Guenneau, D. Canet, *Journal of Magnetic Resonance* **1999**, *141*, 312.
[18] J. Carlos Cobas, M. A. Bernstein, M. Martín-Pastor, P. G. Tahoces, *Journal of Magnetic Resonance* **2006**, *183*, 145.

Entry for the Table of Contents



Pure shift spectra are generated from 2D J-resolved data, with no signal loss and with absorption mode peaks.