

TOPSPIN

Automatic Peak Picking in 2D NMR Spectra Using Neural Networks

Dominik Graf, Michael Fey, Leila Mohammadzadeh, Manuel Cordova, Simon Bruderer, Bruker Switzerland AG, 8117 Fällanden, Switzerland.

Innovation with Integrity

Abstract

An important step in analysing NMR spectra is the identification of peaks in multi-dimensional NMR spectroscopy. In multi-dimensional spectra, signal-to-noise of peaks is typically low, and artifacts are prevalent. This makes peak identification a challenging task. Despite a wealth of computational approaches to automatize the peak-picking task, manual work is usually still required to improve the resulting peak-list. The biggest difficulty faced by peak-picking methods is to differentiate between real signals and various types of artifacts. Recent advances in machine learning algorithms offer new opportunities to solve this problem. Deep learning algorithms have proven to reach human-level performance in various pattern recognition tasks.

Here, we introduce a new deep learning-based approach for peak picking in 2D NMR experiments in our software TopSpin. The approach is available starting from version 4.4.1 as command *pp2dml*. We show that the method performs better than the currently implemented method in TopSpin and can differentiate between different types of peaks, such as compound peaks or various types of artifacts.

Introduction

Multi-dimensional NMR spectroscopy is an important tool for structure elucidation and verification, molecular dynamics studies, structural studies, and many other applications^[1,2]. One of the first steps of most NMR studies is to identify peaks in the acquired spectrum. Creating a peak-list is particularly important if the spectrum is automatically processed.

Peak picking is still often done by hand. In 1D spectroscopy, the challenge is to separate overlapping peaks to obtain a meaningful peak-list. In multi-dimensional spectroscopy, the main challenge is to distinguish between real NMR signals and artifacts generated due to limitations in acquisition time and imperfections in the pulse sequence.

However, this can be a tedious and time-consuming process depending on the measured sample and the type of experiment performed. In particular, high-throughput and routine experiments may strongly benefit from an automated peak picking process.

There are numerous algorithms developed to automate peak picking [3-9]. However, such methods are typically designed for certain types of spectra, as for example the DEEP picker neural network model by Li et al. which is optimized for protein spectra, and does not generalize well on different types of experiments [9]. More general methods such as the 2D peak picker which is currently implemented in TopSpin (TopSpin command: *pp2d*) [10] work based on intensity-thresholding. These algorithms are not able to distinguish between artifacts and real NMR signals if their amplitudes are of the same magnitude.

Here, we introduce a neural network-based method for automatic peak picking in 2D NMR spectroscopy. We show that our algorithm performs better than previous methods in TopSpin, creating a comprehensive peak-list whilst avoiding picking artifacts. The new TopSpin command *pp2dml* adds a first AI-based command to process multi-dimensional NMR spectra to the TopSpin AI commands *sigreg*, *apbk* and *mlcon* [11-13].

Methods

Because 2D spectra are sparser than their 1D counterparts and contain much more datapoints, the application of fully deep learning-based methods is computationally expensive. The approach introduced here should provide a peak-list within a few seconds. For computational efficiency we thus structure the approach in three steps: First, possible peak candidates are identified. Then, for each candidate around 30 features such as the signal-to-noise ratio of the peak or intensity patterns in regions related to the peak are extracted from the spectrum. These features are fed into a neural network to distinguish between real NMR peaks and artifacts such as T1-noise peaks or truncation artifacts. Finally, the real NMR peaks are annotated as either peaks from the compound, impurity, solvent, or special types of artifacts (e.g., HMQC sidebands).

A dedicated neural network was developed for COSY, HSQC and HMBC spectra to ensure an optimal identification of real NMR peaks for these 2D experiments. For training, 5000 synthetic spectra have been used. Experimental data is only available in limited quantities and needs to be annotated manually, which can introduce inconsistencies in the peak-list. Synthetic data is available in large quantities and the peak-list is known precisely. The synthetic spectra are tailored to cover a wide range of sizes and zero filling ratios and contain artifacts relevant to the individual experiment. The neural networks developed for the different experiments consist of several fully connected layers with a total of around 800k weights.

Real NMR signals found by the network are classified into compound, impurity, solvent, and other artifacts using a rule-based method. Compound peaks are distinguished from impurities peaks by an amplitude threshold with respect to the most intense peak in the spectrum. Experimental artifacts such as the so-called COSY artifacts in HSQC spectra are identified by searching for patterns in the peak-list. To find COSY artifacts, the intersections between the ¹H and ¹³C frequencies of two different real NMR signals are checked for weaker signals. Solvent peaks are found by searching a pre-defined list of known solvent signals. The full list of peak types is given in the TopSpin manual entry of *pp2dml* [10]. The peak types are presented by TopSpin as peak annotations and can be used to filter out undesired peaks (e.g., remove all impurities and artifacts).

To test how well the approach can distinguish between real NMR peaks and artifacts, we use an experimental dataset containing 11 COSY, 24 HSQC and 16 HMBC spectra with manually created peak-lists. In addition, for each experiment type 100 synthetic spectra with exactly known peak positions are used. To evaluate the algorithm results, we compare the known peak positions (ground truth) with the peaks picked by our algorithm, using a tolerance threshold to match peaks with slightly different chemical shifts. We distinguish between true positives, false negatives, and false positives. True positives are peaks found in both the ground truth peak-list and by the algorithm (expected peaks). False negatives are peaks present in the ground truth peak-list but not found by the algorithm (missing peaks). False positives are peaks found by the algorithm that are not present in the ground truth peak-list (over-picked peaks).

As a quantitative measure for the performance of our peak-picking algorithm, we define two different metrics. Metric 1 (sensitivity) is a measure of how often ground truth peaks are found, compared to the number of peaks in the ground truth peak-list:

$$m_1 = \frac{\text{\#true positives}}{\text{\#ground truth}}$$

Metric 2 indicates how much the algorithm over-picks, comparing incorrectly picked peaks to the total number of picked peaks:

$$m_2 = 1 - \frac{\text{\#false positives}}{\text{\#total positives}}$$

The values of both metrics are between 0 and 1 and a higher value indicates a better result.

Results and Discussion

Presented in **Figure 1** are the results of *pp2dml* in a region of a COSY spectrum. The peaks are shown together with the peak annotations found by the algorithm. In this region, the two main diagonal doublets at F1 = F2 = 6.2 ppm and F1 = F2 = 6.7 ppm produce four cross-peaks (two off-diagonal doublets). Additionally, their satellites produce four diagonal and eight off-diagonal doublets due to scalar couplings. A low intensity peak which is attributed to an impurity in the sample, is measured close to each of the four main peaks. This spectral region is a good test for our algorithm due to the presence of truncation artifacts in F1 direction, and T1 noise at the frequencies of the main doublets whose intensity is higher than the satellite peaks. In this case, it is crucial to pick the low-intensity satellite and impurity peaks while avoiding picking the noise. Only real peaks were picked here, showing the high performances of *pp2dml*.

In **Figure 2**, the same section of the spectrum as in **Figure 1** is shown, but with the peaks found by the existing TopSpin command *pp2d*. There are no annotations because the *pp2d* command does not distinguish between compound and other types of peaks. The intensity thresholds for peak acceptance must be set manually for *pp2d*. If a high threshold value is chosen to avoid picking truncation artifacts and T1 noise, only the main doublets are being picked (red stars in **Figure 2**). This setting thus misses the satellite and impurity peaks present in the spectrum. On the other hand, if the threshold is set low enough to pick the satellites and their cross-peaks, many truncation artifacts and T1 noise peaks are picked along with the desired ones (black together with red stars in **Figure 2**). In this case, most of the picked peaks are not real NMR peaks, and the low-intensity impurity peaks – which might still be desired depending on the application – are still missed by the picker. Thus, independent of the threshold setting, *pp2d* does not reach the performance of *pp2dml* for this example.

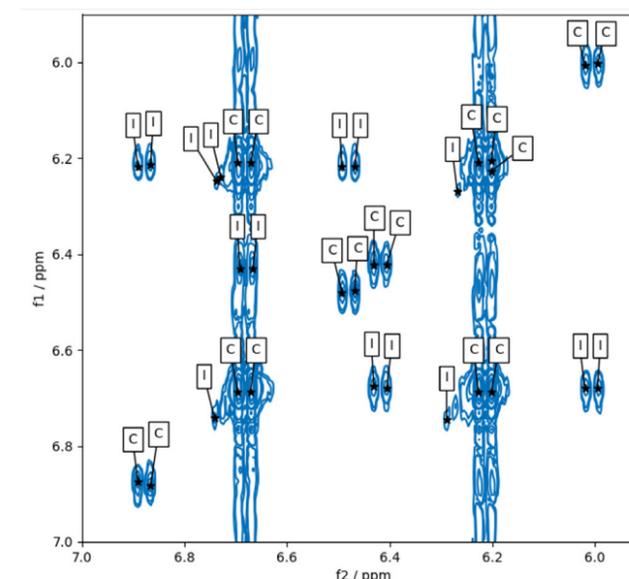


Figure 1: Results of *pp2dml* on the aromatic region in the COSY spectrum of Santonin. Contour lines are shown in blue and picked peaks as black stars. Peak annotations found by the algorithm are given in boxes (C=compound, I=impurity).

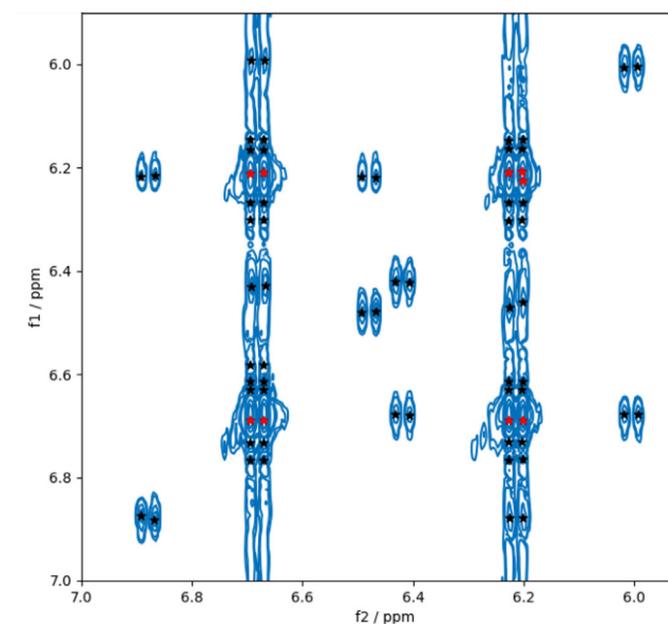


Figure 2: Peaks found by the existing TopSpin command *pp2d* on the same region as shown in Figure 1. Two different peak-lists are shown: One for a high value of the threshold (red stars) and one for a low value of the threshold (black stars).

The performance of pp2dml on the experimental and synthetic datasets is presented in **Figure 3** split by experiment type (HMBC, HSQC and COSY). It shows the fraction of ground truth peaks found (metric m1) and a measure of over-picking (metric m2). For the synthetic dataset, both m1 and m2 are mostly close to one, indicating a large fraction of peaks found with little over-picking. An exception are COSY spectra, where m1 is lower due to the weak signals of which some are missed. The weak signals result from the multiplet patterns in both dimensions prevalent in COSY but not in the other experiment types. The high values confirm that the training of the network succeeded, and the characteristics of the synthetic NMR spectra could be learned.

For the experimental dataset m1 values are close to one while m2 values are a bit lower. This is, because the algorithm was tuned to score higher on m1 compared to m2, since we assume that missing a peak is worse than over-picking. We note that some of the experimental HSQC spectra have only a handful of peaks and a single over-picked peak can substantially reduce the metric value. A caveat of this study is that the ground truth peaks were meticulously picked by hand, but some remaining residual errors in the peak-list can still affect the metrics. Additionally, an exact definition of whether a peak should be considered as a real NMR peak or not is elusive and subject to discussion and might also depend on the purpose the peak-list is used for.

A comparison between pp2dml and pp2d for the experimental dataset is presented in **Figure 4** using a diagram similar to a ROC (receiver operating characteristic) curve. For the figure, the metric averaged over all HSQC, HMBC, COSY spectra or the entire dataset is calculated both for pp2dml and pp2d. Since pp2dml is fully automatized and does not require any user input, its result is a single combination of average m1 and m2 values. For pp2d, a minimum intensity threshold (parameter name in TopSpin: MI) must be set per spectrum, resulting in different m1 and m2 depending on this threshold value. Thus, pp2d is represented by a curve in the diagram, depending on the threshold value. If the intensity threshold is chosen high, little over-picking occurs, but weak peaks are missed. This results in a low value of m1 but a high value of m2. For **Figure 4**, the intensity threshold of pp2d is defined as either a

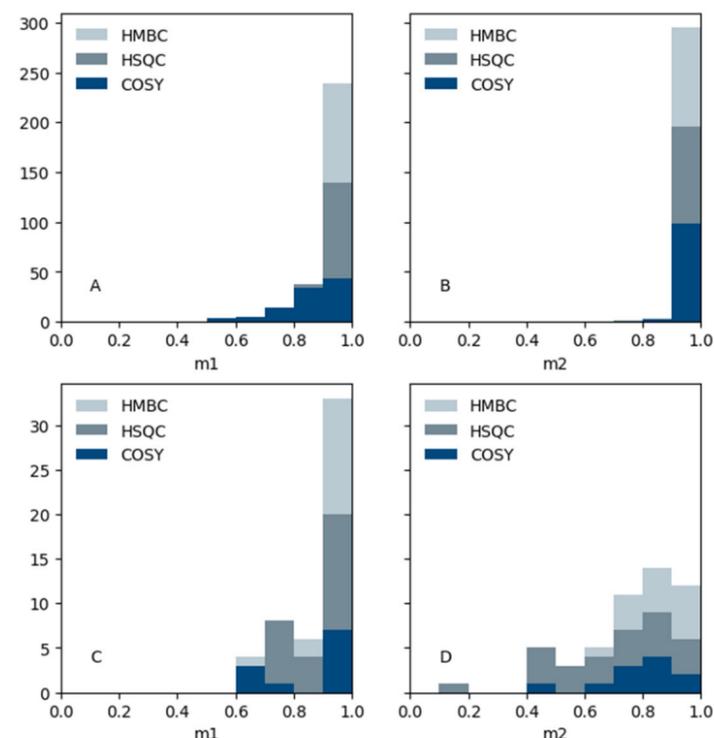


Figure 3: Distribution of the metric m1 (fraction of ground truth peaks found) and metric m2 (indication of over-picking) of pp2dml results on the synthetic dataset (upper panels A and B) and the experimental dataset (lower panels C and D) vs the number of spectra.

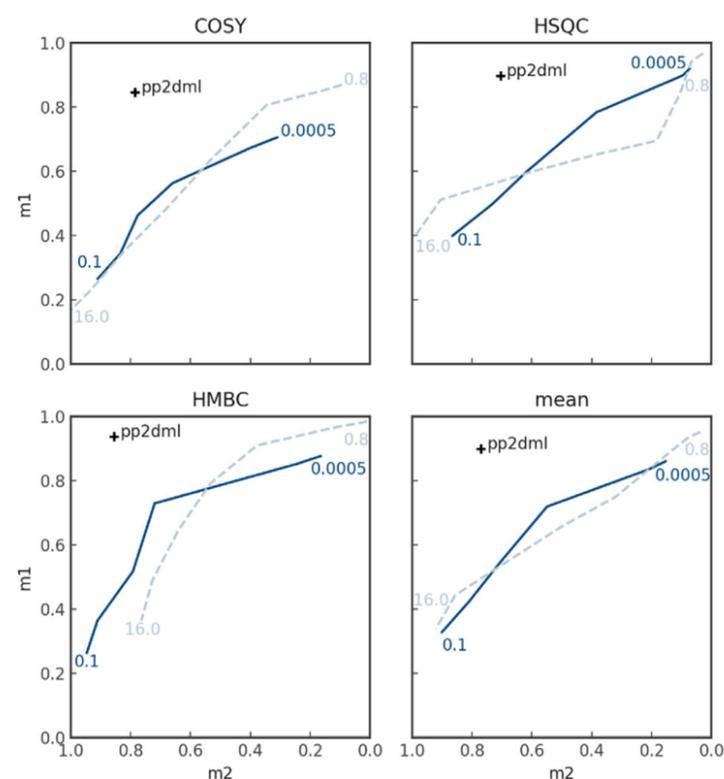


Figure 4: Metric values of pp2dml (black cross) compared to pp2d (blue lines) averaged over all COSY, HSQC and HMBC spectra or the entire dataset (mean). For pp2d MI has been varied per spectrum as a fraction of the maximum amplitude of the spectrum (solid line) or a multiple of the noise level (dashed line).

fraction of the highest peak of a spectrum (i.e., between 0.05 % and 10 % of the maximum intensity) or a multiple of the noise level of the spectrum (i.e., between 0.8 and 16 times the noise level). A perfect peak picker would be at the top left corner of each panel in **Figure 4**, with m1 and m2 both equal to 1. The values of pp2dml are closer to this corner independent of the selection of the threshold value. We thus conclude that the performance of pp2dml is superior to that of pp2d.

Options to run pp2dml

The new approach can be run in full automation without any parameter needed to be selected. However, to tailor the results to a specific application, optional parameters can be set. A full list of the run parameters is provided in the TopSpin manual entry of *pp2dml*^[10].

In **Figure 5** a region of a COSY spectrum is shown which appears asymmetric because of axial artifacts. In the default settings, the off-diagonal peaks at the bottom right of the spectrum are not picked, since they do not have a symmetry partner and no corresponding diagonal peak. However, if pp2dml is started with the option “-nosymmetryfilter”, those peaks are picked. This option is useful if weak peaks in COSY spectra need to be detected, close to the detection threshold.

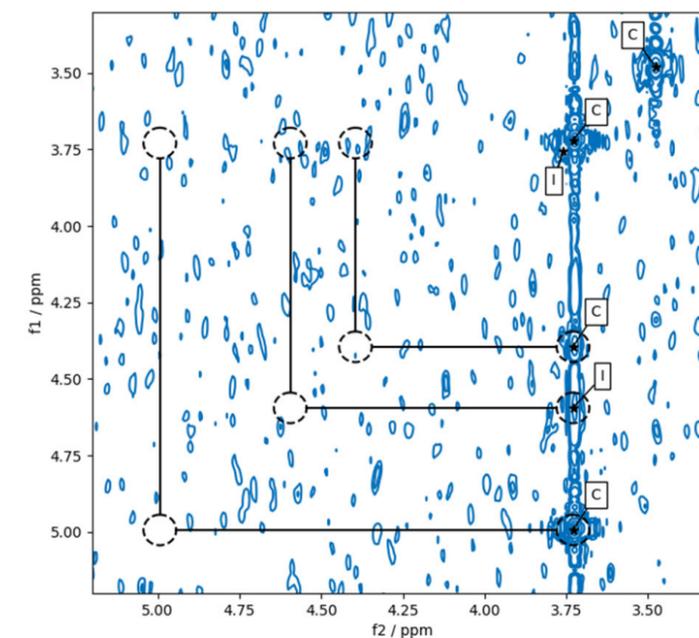


Figure 5: Results of pp2dml run with the option “-nosymmetryfilter” on a COSY spectrum. With that option set, peaks are picked despite a missing symmetry partner and no corresponding diagonal peak present. Dashed circles show the peaks only picked with the option “-nosymmetryfilter” set and positions checked by the symmetry-filter.

Settings for pp2dml	m1	m2
No option set (default)	0.85	0.79
Symmetry filter disabled (“-nosymmetryfilter”)	0.92	0.67
Only compound peaks (“-onlycompound”)	0.67	0.85

Table 1: Metric results of pp2dml on the experimental COSY spectra

Some applications require only compound peaks for the analysis. To remove all peaks not annotated as compound like impurities or solvent peaks, the option “-onlycompound” can be set. To find peaks not at a local maximum or minimum, the option “-shoulder” is available. Information from external projections can be used to filter out peaks that have no counterpart in the 1D spectrum. This filter is activated with the option “-useprojection” and internally runs the AI command mldcon on the 1D spectrum to find possible peak positions.

In **Table 1**, the metric values for the experimental COSY spectra are shown, when either the option “-nosymmetryfilter” or “-onlycompound” is set. As expected, the option “-nosymmetryfilter” increases the fraction of peaks detected (m1 increases) but at the price of a larger fraction of over-picked peaks (m2 lower). The option “-onlycompound” on the other hand acts in the opposite direction, increasing the metric for over-picking and decreasing the detection fraction.

Running pp2dml on Fourier 80 spectra

In **Table 2** the results of pp2dml acquired with the benchtop instrument Bruker Fourier 80 are compared to those measured at higher field strengths. Both for HSQC and HMBC spectra, pp2dml is found to perform equally well for the low- and high-field spectra, showing the capability to use the approach also for lower field spectra. For COSY spectra however, pp2dml shows a lower detection fraction m1 in Fourier 80 spectra compared to higher fields. This is again mostly due to generally lower signal-to-noise ratio in spectra measured using the Fourier 80 spectrometer in our test set. Then the symmetry filter often removes peaks due to a missing symmetry partner. As mentioned before, m1 can be improved using the option “-nosymmetryfilter”, but this leads to some over-picking.

Settings for pp2dml	m1	m2
COSY (n=11)		
All spectra	0.85	0.79
Higher-field spectra (n=8)	0.91	0.75
Fourier 80 Spectra (n=3)	0.67	0.89
HSQC (n=24)		
All spectra	0.90	0.70
Higher-field spectra (n=20)	0.90	0.71
Fourier 80 Spectra (n=4)	0.91	0.68
HMBC (n=16)		
All spectra	0.94	0.85
Higher-field spectra (n=12)	0.92	0.84
Fourier 80 Spectra (n=4)	0.97	0.90

Table 2: Metric results of pp2dml on Fourier 80 spectra compared to higher field strength (Proton base frequency > 300 MHz).

Peak annotation

Some experiments create real NMR peaks that are however often not used in the interpretation of the spectrum. For example HSQC spectra show weak peaks resulting from multiple bond correlations between a proton and a carbon. These 3J long-range coupling effects are called COSY artifacts and are often not needed for the interpretation of the spectrum. Currently, peak annotations can only be derived for COSY, HSQC and HMBC spectra.

Shown in **Figure 6** is a region of a HSQC spectrum with peaks and peak annotations found by pp2dml. The peaks at the top right and bottom left are annotated as compound peaks. In addition, the two low intensity peaks at the top left and bottom right are annotated as COSY artifacts. They originate from the two compound peaks and are characterized by their symmetric arrangement.

For **Figure 7**, pp2dml has been run on a HMBC spectrum. The two peaks at the bottom of the region are correctly annotated as compound peaks. The two peaks at the top of the region are correctly annotated as 1J coupling artifacts or otherwise referred to as HMQC responses. They are characterized by their almost identical F1 positions at a specific f2 distance and are symmetrically arranged in F2 around a compound peak which can have any F1 position.

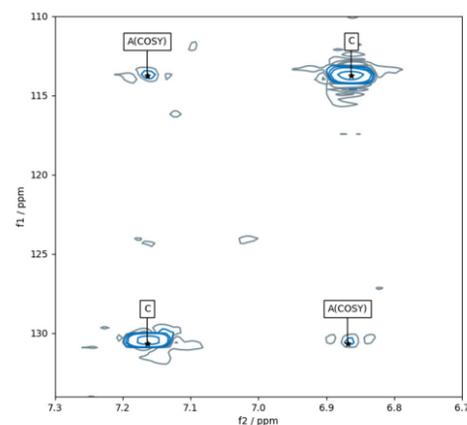


Figure 6: Result of pp2dml on a region of a HSQC spectrum. Contour lines are shown in blue and grey for positive and negative amplitudes, respectively). The peak annotation “C” stands for compound and “A(COSY)” for a COSY artifact.

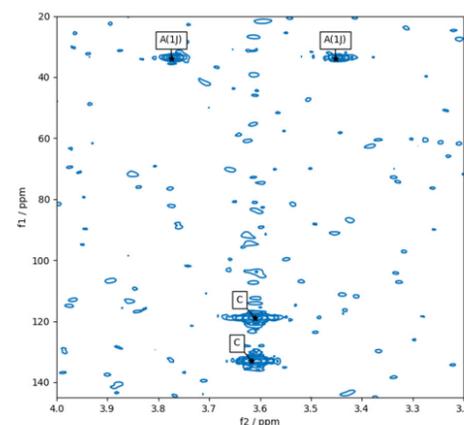


Figure 7: Result of pp2dml on a region of a HMBC spectrum. The peak annotation “C” stands for compound and “A(1J)” for an artifact due to 1J coupling.

Conclusion

We have presented a fully automated algorithm to pick peaks in 2D NMR spectra which can run without any user input. Using neural networks, the method first picks the peaks and then classifies them into compound peaks and different artifacts applying a rule-based algorithm. The corresponding command **pp2dml** is available in TopSpin version 4.4.1 and higher.

The test results on experimental and synthetic spectra show that *pp2dml* performs better than the currently available TopSpin command *pp2d*. The most important difference is that compared to the rule-based pp2d algorithm, the neural network in pp2dml learned to distinguish real NMR peaks from artifacts.

Practical tips

- The algorithm is optimized for COSY, HSQC and HMBC spectra. By default, it will not run on other types of spectra. To force the algorithm to run on other types of 2D experiments the option -f can be used (**pp2dml -f**), but satisfying results cannot be guaranteed for other types of spectra. We plan to support more types of spectra in the future.
- Peak annotations are currently only available for COSY, HSQC or HMBC spectra.
- While pp2dml shows to run on spectra acquired with non-uniform sampling (NUS), it has not been optimized to find low-SINO peaks in such spectra.
- If external projections are defined for a 2D spectrum there is the option to use the command **pp2dml -useprojections**. With this option, only peaks in regions containing signals in the 1D projections are picked.
- To reduce over-picking there are two options. With the command **pp2dml -ppmnum=X** only the X most intense peaks are picked. Otherwise, the command **pp2dml -onlycompound** can be used to display only the peaks classified as compound.
- If there are peaks missing in a COSY spectrum, the command **pp2dml -nosymmetryfilter** can be used. Peaks will then be added to the peak-list even if they have no symmetric or diagonal counterpart.
- To access the pp2dml page in the TopSpin manual use the command **help pp2dml**.

References

1. R. R. Ernst, G. Bodenhausen, and A. Wokaun, «Principles of Nuclear Magnetic Resonance in One and Two Dimensions», Clarendon Press, 1990.
2. F.J.M. van de Ven, «Multidimensional NMR in Liquids: Basic Principles and Experimental Methods», Wiley, 1996.
3. R. Koradi, M. Billeter, M. Engeli, P. Guntert, and K. Wuthrich, «Automated peak picking and peak integration in macromolecular NMR spectra using AUTOPSY», Journal of Magnetic Resonance, Bd. 135, Nr.2, pp.288-297, 1998.
4. Z. Liu, A. Abbas, BY. Jing, and X. Gao, «WaVPeak: picking NMR peaks through wavelet-based smoothing and volume-based filtering», Bioinformatics, Bd.28, Nr. 7, pp.914-920, 2012.
5. B. Alipanahi, X. Gao, E. Karakoc, L. Donaldson, and M. Li, «PICKY: a novel SVD-based NMR spectra peak picking method», Bioinformatics, Bd.25, Nr.12, pp.i268-275, 2009.
6. Y. Cheng, X. Gao, and F. Liang, «Bayesian peak picking for NMR spectra», Genomics Proteomics & Bioinformatics, Bd.12, Nr.1, pp.39-47, 2014.
7. P. Klukowski, M. Augoff, M. Zieba, M. Drwal, A. Gonczarek, and M. J. Walczak, « NMRNet: a deep learning approach to automated peak picking of protein NMR spectra», Bioinformatics, Bd.34, Nr. 15, pp.2590-2597, 2018.
8. N. Kobayashi et al., «Noise peak filtering in multi-dimensional NMR spectra using convolutional neural networks», Bioinformatics, Bd.34, Nr. 24, Pp.4300-4301, 2018.
9. DW. Li, A. L. Hansen, C. Yuan, L. Brüscheweiler-Li, and R. Brüscheweiler, «DEEP picker is a deep neural network for accurate deconvolution of complex two-dimensional NMR spectra», Nature Communications, Bd. 12, Nr. 5229, pp. 1-13, 2021.
10. «Processing Commands and Parameters - TopSpin User Manual» TopSpin 4.4.1 Bruker, 2024.
11. Schmid, N., Bruderer, S., Paruzzo, F., Fischetti, G., Toscano, G., Graf, D., Fey, M., Henrici, A., Ziebart, V., Heitmann, B. and Grabner, H., Wegner J.D., Sigel R.K.O., Wilhelm D.; Deconvolution of 1D NMR Spectra: A Deep Learning-Based Approach. Journal of Magnetic Resonance, p.107357, 2022
12. Bruderer, S., Paruzzo, F., and Bolliger, C.; Deep learning-based phase and baseline correction of 1D 1H NMR Spectra, Bruker, URL: <https://www.bruker.com/en/products-and-solutions/mr/nmr-software/topspin.html>, 2021
13. F. Paruzzo, S. Bruderer, Y. Janjar, B. Heitmann and C. Bolliger, «Automatic Signal Region Detection in 1H NMR Spectra Using Deep Learning» Bruker Whitepaper, <https://www.bruker.com/en/products-and-solutions/mr/nmr-software/topspin.html>, 2020

Bruker BioSpin
info@bruker.com

bruker.com

Customer Support
<https://www.bruker.com/en/services/support.html>

Online information
[bruker.com/](https://www.bruker.com)

