European Funds for Smart Economy
Republic of Poland
Co-funded by the European Union
FNP Foundation for Polish Science

Network flow methods for alignment and deconvolution of multidimensional spectra

Maria Bochenek[1], Michał Startek[1], Krzysztof Kaźmierczuk[2], and Anna Gambin[1]

[1] Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097, Warsaw, Poland
[2] Centre of New Technologies, University of Warsaw, Banacha 2c, 02-097, Warsaw, Poland

Multidimensional nuclear magnetic resonance (NMR) spectroscopy is an analytical technique widely used in proteomics, metabolomics, and structural biology. It reduces peak overlap and improves interpretability by spreading information across multiple chemical-shift axes.

However, comparative or replicate experiments often introduce frequency shifts, hindering accurate matching of corresponding signals. In some experiments, such as variable-temperature NMR, the shifts are intentionally triggered, and frequency tracking provides important information. Accurate spectra alignment is crucial for comparative analysis, automated quantification, and structure elucidation.

Another application of NMR spectroscopy is mixture analysis and reaction monitoring. In this context, deconvolution methods seek to identify individual compounds and estimate their relative abundances within complex mixtures. Reliable identification and quantification are frequently challenged by peak overlap, signal distortions, limited spectral resolution, missing signals, and discrepancies in peak positions between the observed mixture and reference compound libraries. Addressing these issues requires robust computational approaches.

In previous work, we addressed one-dimensional alignment and deconvolution using optimal transport theory. By representing spectra as discrete probability measures, we formulated both tasks as mass-transport problems and quantified spectral dissimilarity using the truncated Wasserstein distance. While effective in one dimension, these methods do not scale well to multidimensional spectra due to the high computational cost of evaluating Wasserstein distances.

To overcome this limitation, we develop a network-flow-based framework for multidimensional spectral analysis. We show that the Wasserstein distance can be efficiently approximated by solving a minimum-cost flow problem on a network whose nodes represent spectral signals from replicate spectra in the alignment setting and from mixture and reference compounds in the deconvolution setting. This formulation substantially improves scalability while retaining the robustness of optimal-transport-based approaches.