Transforming NMR Spectroscopy: Extraction of Multiplet Parameters with Deep Learning

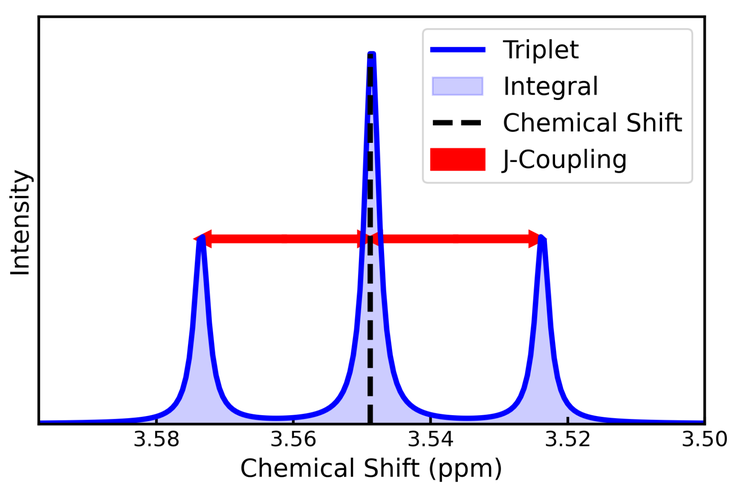
N. Schmid[1,2], M.Wanner[3], G. Fischetti[4], M. Meshkian[1], S. Bruderer[5], A. Henrici[1], J.D.Wegner[2], R.K.O.Sigel[2], B. Heitmann[4], E. Konukoglu[3], and D. Wilhelm[1]

[1]Zurich University of Applied Sciences (ZHAW), Switzerland
[2]University of Zurich (UZH), Switzerland
[3]Swiss Federal Institute of Technology (ETH), Switerland
[4]Ca' Foscari University of Venice, Italy
[5]Bruker BioSpin, Switzerland

Introduction

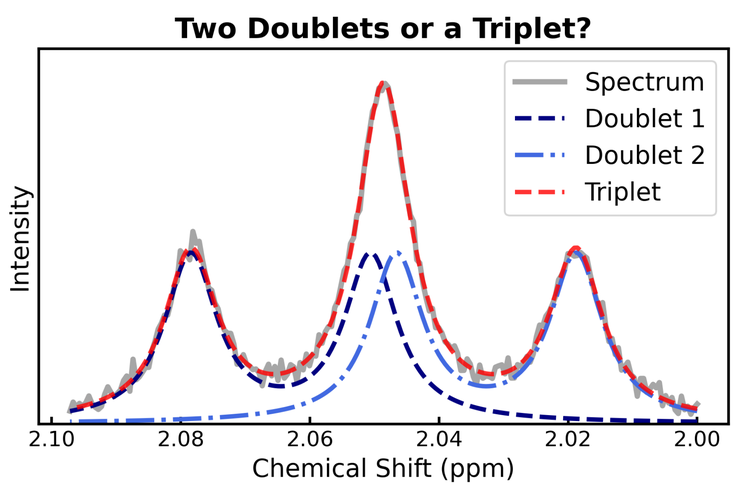
Accurate extraction of multiplet parameters, such as J-couplings and chemical shifts, play a vital role in small molecule analysis using nuclear magnetic resonance (NMR) spectroscopy. These parameters provide essential quantitative information about molecular structures, interatomic interactions, and chemical environments, enabling precise characterization of small organic compounds. This poster presents an innovative computational approach that utilizes state-of-the-art deep learning techniques, specifically detection transformers [1], to automate and optimize the extraction of multiplet parameters from 1D NMR spectra of small molecules. By incorporating these advanced computational methods, experimenters can achieve improved efficiency, accuracy, and speed in analyzing and characterizing small organic compounds using NMR spectroscopy.

1. Multiplet Parameter Extraction



Goal: Determine the J-coupling, chemical shift, proton number and type of multiplet.

2. Ill-Posed Problem



Problem: Multiplet parameter extraction does not have a single solution. The overall patterns of two doublets or a triplet are, in this example, difficult or unattainable to differentiate.
Solution: Define which solution is favored through **regularization**.

3. Combined Loss Function

The **combined loss function** L_{Com} integrates various loss components, including **parameter regression** L_R, **intersection over union** L_{IOU}, and **multiplet classification** L_C, to effectively capture diverse multiplet characteristics across different scales.

$$L_{Com} = L_R + L_{IOU} + L_C$$

Thus, the approach encompasses **accurate parameter estimation**, **precise alignment**, and **confident classification**.

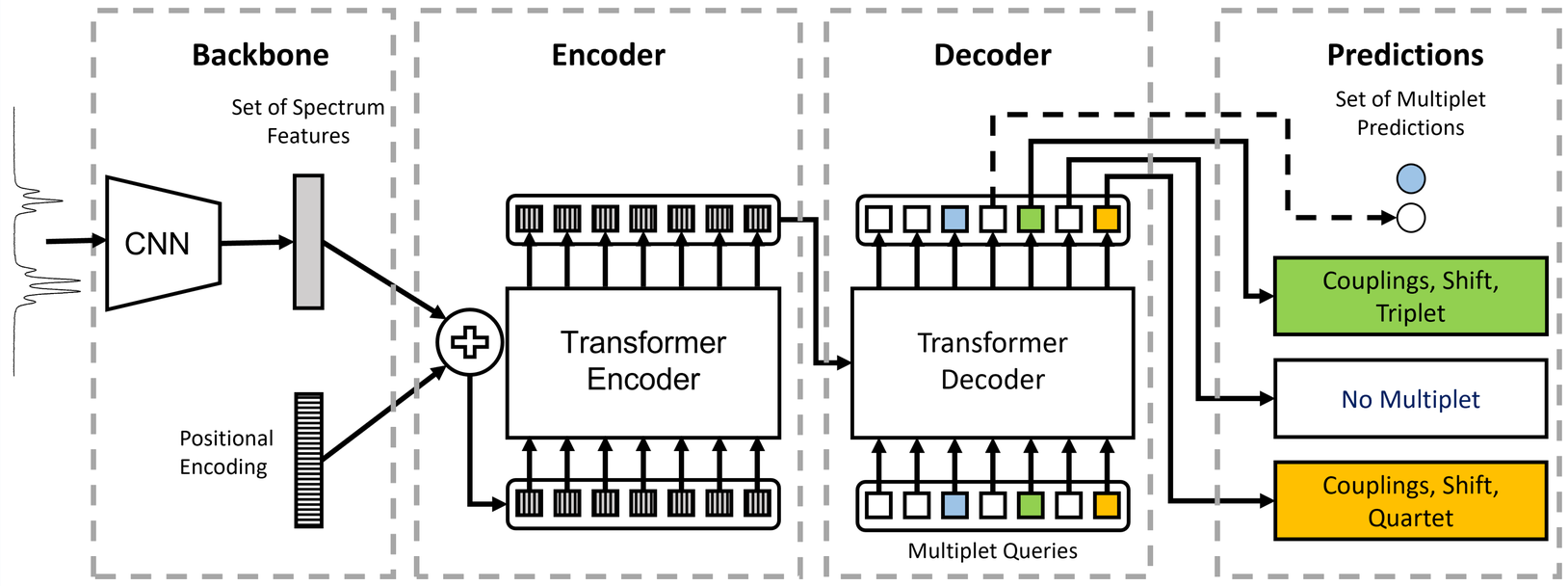
Reference

[1] Carion et al., End-to-End Object Detection with Transformers, ECCV 2020

Acknowledgement

This work was supported by Innosuisse Innovation project 44786.1 1P-ENG.

4. Detection Transformer for NMR Spectra

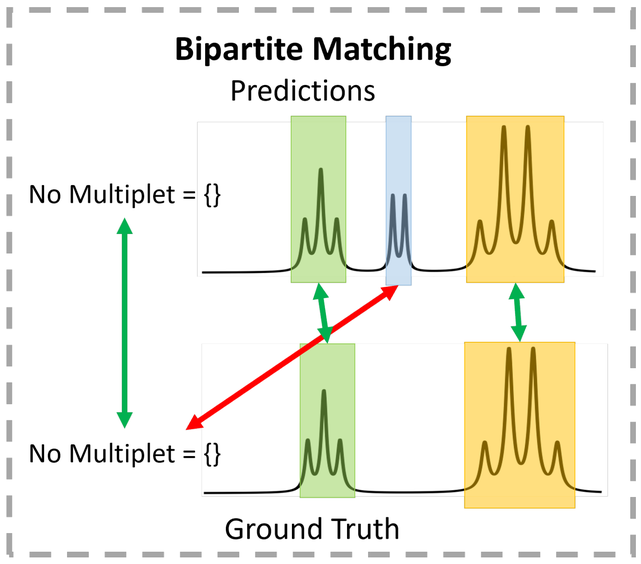


Backbone: Create spectral embeddings to represent the input spectrum, incorporating positional encoding to capture spatial information.

Encoder: Utilize the encoded spectral embeddings to establish features, capturing relationships between aspects of different multiplets, such as chemical shifts, J-couplings or splitting patterns.

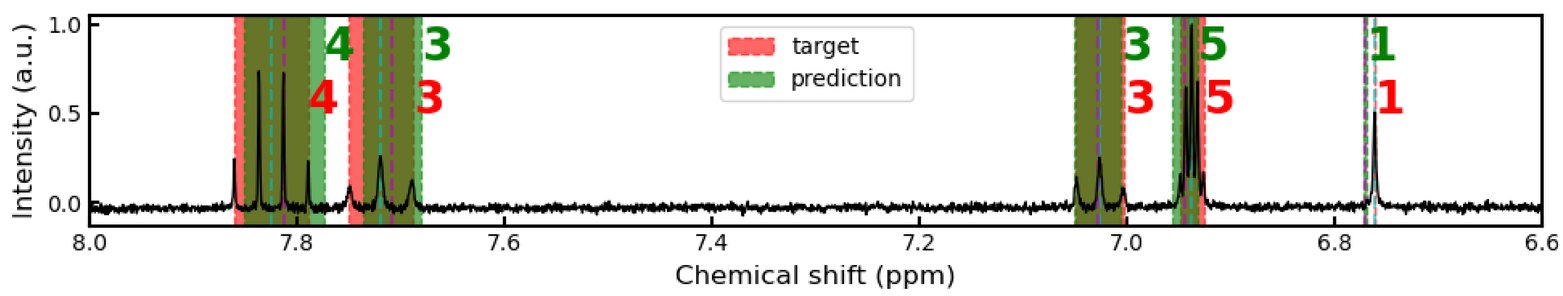
Decoder: Retrieve essential information from the encoded features using specific queries, with each query focusing on specific types and characteristics of multiplets.

Prediction: Predict multiplet parameters (e.g., chemical shifts, J-couplings) and determine the multiplet type, providing insights into their properties.



Training: Match predicted multiplets with ground truth multiplets using bipartite matching (Hunagrian algorithm), minimizing the matching cost and enabling pair-wise loss calculation.

5. Results: Parameter Extraction in Practice



The algorithm showcases **high multiplet identification rate** of over 95% in distorted and low signal-to-noise semi-experimental spectra. It also achieves **precise prediction** of **J-couplings** and **chemical shifts**, with errors well below 1 Hz. Additionally, when combined as a prior with classical local fitting, it further enhances the accuracy of parameter estimates.

6. Conclusions

- The method acts as a **virtual spectroscopist**, capturing and correlating information of and between diverse multiplets, enabling comprehensive analysis of complex NMR spectra.
- It provides **accurate predictions of multiplet parameters** through advanced global reasoning with transformers, enhancing spectral resolution and enabling characterization of partially overlapping multiplet patterns.
- The method's potential for automation and improved analysis in NMR holds **significant promise**, facilitating tasks such as structural elucidation, metabolomics, and drug discovery.