Exact classification of NMR spectra from NMR signals

Problem Formulation

In **Nuclear Magnetic Resonance (NMR)** a sample is plunged into an **external magnetic field** of magnitude B_0 and is excited with an external **radio-frequency (RF) wave**. The response produces the **NMR signal**:

$$u(t) = \sum_{k=1}^{p} c_k e^{\lambda_k t} + N(t)$$

where:

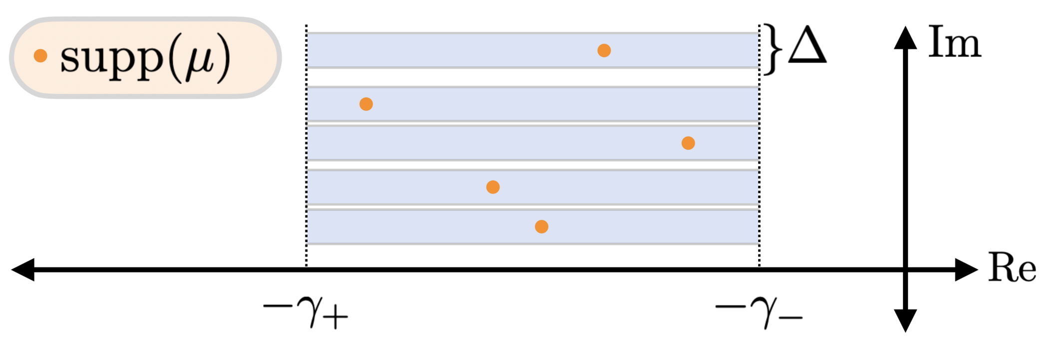
- $\{\mathrm{im}(\lambda_k)\}$ are the sample-dependent **frequencies**.
- $\{\mathrm{re}(\lambda_k)\}$ are the **decay rates** proportional to B_0^{-1}.
- $\{c_k\}$ are the **amplitudes** and **phases**.
- $N(t)$ is the **additive noise** on the signal.

The NMR spectrum and NMR signal are related as follows:

- The **NMR spectrum** is encoded in the **complex measure**

  $$\mu = \sum_{k=1}^{p} c_k \delta_{\lambda_k}$$

  for a known and fixed p. We assume that the decay rates are bounded and that the frequencies are separated, that is $-\mathrm{re}(\lambda_k) \in [\gamma_-, \gamma_+]$ and $|\mathrm{im}(\lambda_k - \lambda_\ell)| \geq \Delta$ for some $\gamma_+, \gamma_-, \Delta > 0$.



- The noiseless **NMR signal** is represented as the image of the spectrum under the injective linear operator S

  $$t > 0: \quad S\mu(t) := \int_{\mathbb{C}} e^{\lambda t}\, d\mu(\lambda)$$

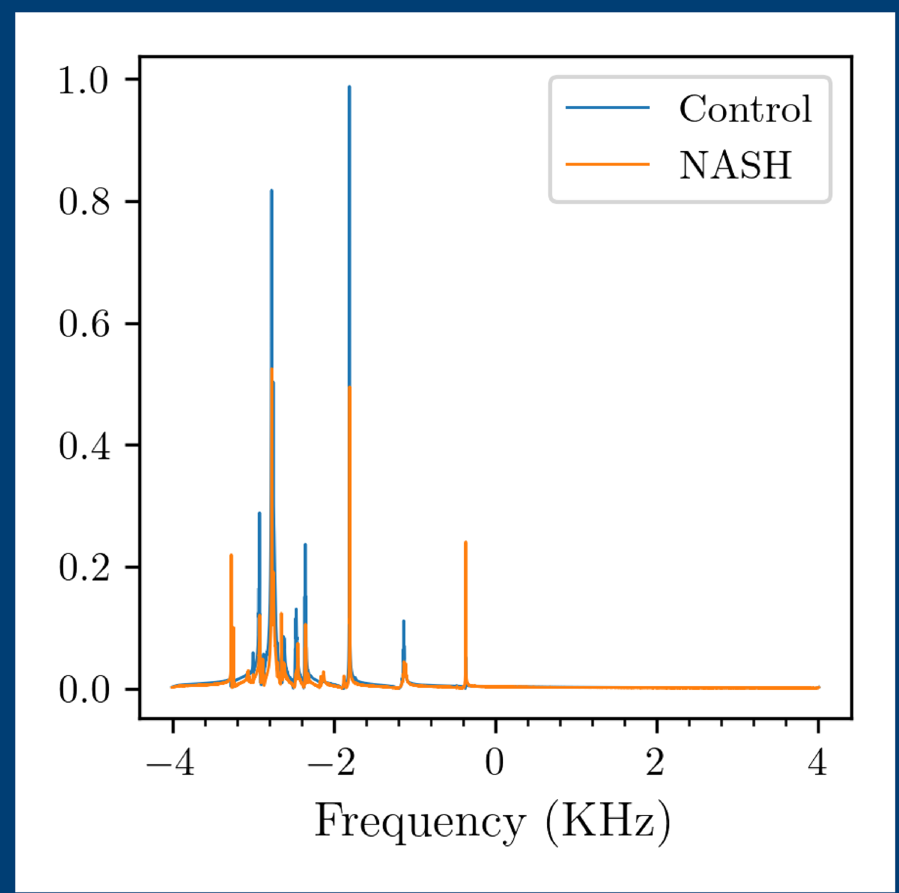
- The noisy NMR signal becomes $u(t) = S\mu(t) + N(t)$.

To compare samples belonging to two classes we quantify the difference between:

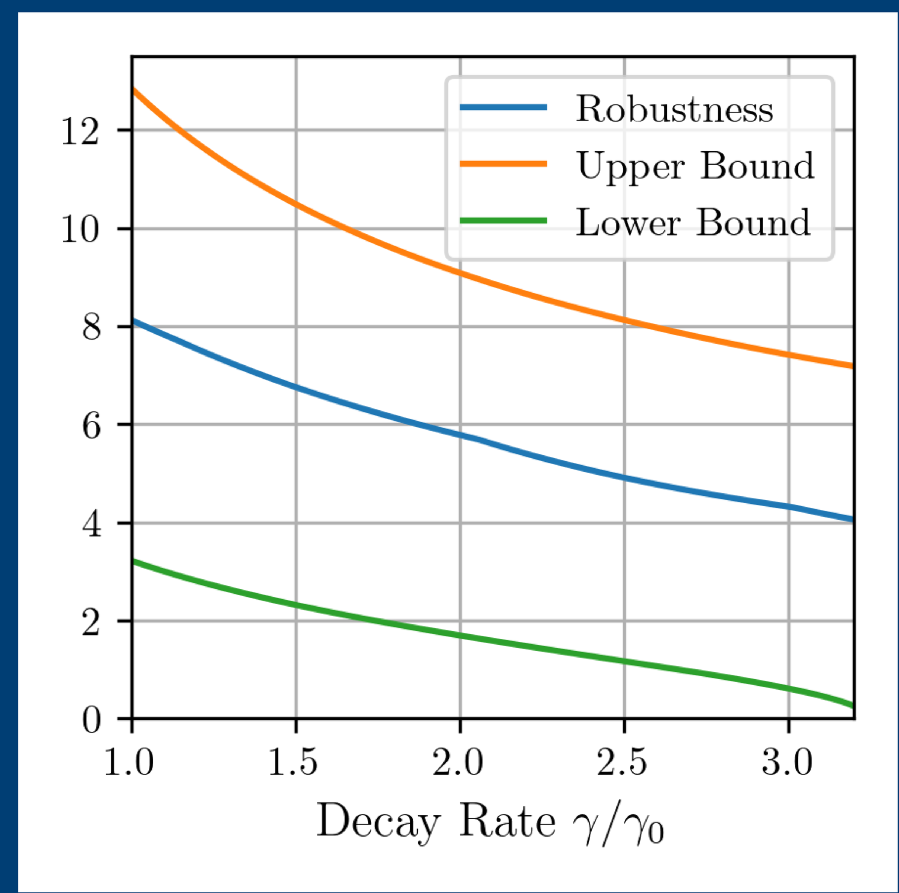
- NMR spectra using the **total variation distance** d_{TV}.
- NMR signals using the L^2 **distance** d_{L^2}.

How **robust to adversarial additive noise** is the **end-to-end classification** of two disjoint sets O_1 and O_2 of NMR **spectra** from their NMR **signals** SO_1 and SO_2?

The **NMR spectrum** can be **exactly classified** from the **NMR signal** as long as its **SNR** is, at most, proportional to the **square root of the external magnetic field.**



NMR signals of a control group mouse u_1 and of a mouse with NASH u_{13}.



Theoretical upper and lower bounds, and empirical robustness.

Pedro Izquierdo Lehmann
pizquie1@jh.edu
Johns Hopkins University and Pontificia Universidad Católica de Chile

Aline Xavier
aline.xavier@usach.cl
Universidad de Santiago de Chile

Marcelo Andia
meandia@uc.cl
Pontificia Universidad Católica de Chile

Carlos A. Sing Long
casinglo@uc.cl
Pontificia Universidad Católica de Chile

Main Result

Let $R_p = \lceil (p-1)/2 \rceil$ and $H_p = 2(1 + 1/2 + \ldots + 1/R_p)$. If there exists $\Delta > 0$ such that

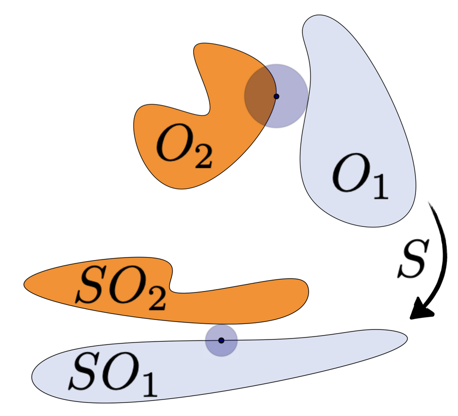
$$|\mathrm{im}(\lambda_k - \lambda_\ell)| \geq \Delta > 2\sqrt{H_p^2 \gamma_+^2 - p^{-2}(\gamma_+ + \gamma_-)^2}$$

for every $\mu \in O_1 - O_2$, then

$$\sqrt{\frac{\eta}{2p}} \leq \frac{\mathrm{dist}_{L^2}(SO_1, SO_2)}{\mathrm{dist}_{\mathrm{TV}}(O_1, O_2)} \leq \sqrt{\frac{1}{2\gamma_+}}$$

where

$$\eta = \frac{1}{2\gamma_+} - \frac{H_p R_p}{\sqrt{(\gamma_+ + \gamma_-)^2 + \Delta^2 R_p}}$$



Numerical Experiments

We used 9.4T [1]H NMR spectroscopy data from the liver of 12 control mice (C) and 12 mice with NASH (N). The high-field data allows us to estimate $p = 9$, $\Delta = 100$ Hz, the average decay rate $\gamma_0 = 3.36$ Hz, and the groups

$$O_C = \{\mu_1, \ldots, \mu_{12}\} \quad \text{and} \quad O_N = \{\mu_{13}, \ldots, \mu_{24}\}.$$

To simulate decay rates, we synthetize the signals

$$t > 0: \quad u_k^\gamma(t) = e^{-\gamma t} S\mu_k(t) \quad i \in \{1, \ldots, 24\}$$

to then define the classes of NMR signals

$$SO_C^\gamma = \{u_1^\gamma, \ldots, u_{12}^\gamma\} \quad \text{and} \quad SO_N^\gamma = \{u_{13}^\gamma, \ldots, u_{24}^\gamma\}.$$

We validate our theoretical findings by computing

$$\mathrm{dist}_{L^2}(SO_C^\gamma, SO_N^\gamma)/\mathrm{dist}_{\mathrm{TV}}(O_C, O_N)$$

and the upper and lower bounds with $\gamma_+ = 1.1\gamma_0 + \gamma$ and $\gamma_- = 0.9\gamma_0 + \gamma$.

Future work

We will analyze the classification of spectra that have an atomic and an absolutely continuous part, modelling spectral leakage, and the effect of distortions due to the push-forward of a measure, modelling spectral shifts.

Scan for paper

Acknowledgements: P.I.L. thanks the The Acheson J. Duncan Fund for their generous support. P.I.L and C.S.L were partially funded by ANID grant Fondecyt 1211643. P.I.L. and M.E.A. were partially funded by ANID grants ICN2021_004 and Fondecyt 1220922. C.S.L. was partially funded by ANID grant Basal FB210017.

JOHNS HOPKINS
WHITING SCHOOL *of* ENGINEERING