# LA-UR-22-31711

**Approved for public release; distribution is unlimited.**


**Title:** Complex-Valued Signal Denoising and Bayesian Optimization for Detection of Synthetic Opioids


**Author(s):** Day, Amber Joule


**Intended for:** PhD Candidacy presentation at University of Texas at Austin


**Issued:** 2022-11-04

# Complex-Valued Signal Denoising and Bayesian Optimization for Detection of Synthetic Opioids

Candidacy Talk (November 2022)

Presented By: Amber Day

UT Advisor - Sinead Williamson

LANL Mentor - Natalie Klein

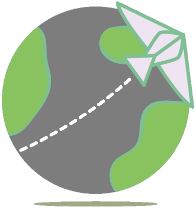
The University of Texas at Austin
Department of Statistics and Data Sciences
College of Natural Sciences

# Table of Contents

# Overseas manufacturers ship synthetic opioids into the United States through international mail

This is due to the:

- high potency of synthetic opioids
- ease and low cost to manufacture synthetic opioids
- persistent demand in the United States
- low risk of detection when sent in small quantities[1]

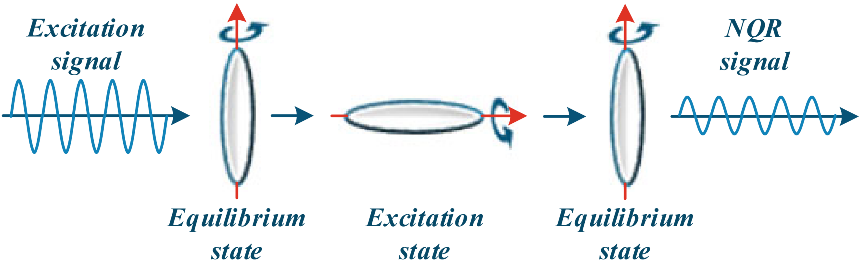


# Synthetic opioids are largely responsible for the overdose crisis in the United States

From October 2020 to October 2021 there were 105,000 drug overdose deaths. This was the highest number ever recorded in a 12-month period in the United States.[1]

[1] Thaivalappil, M. (2022, June 21). *Addressing the Overdose Crisis.* United States Department of State. https://www.state.gov/addressing-the-overdose-crisis/

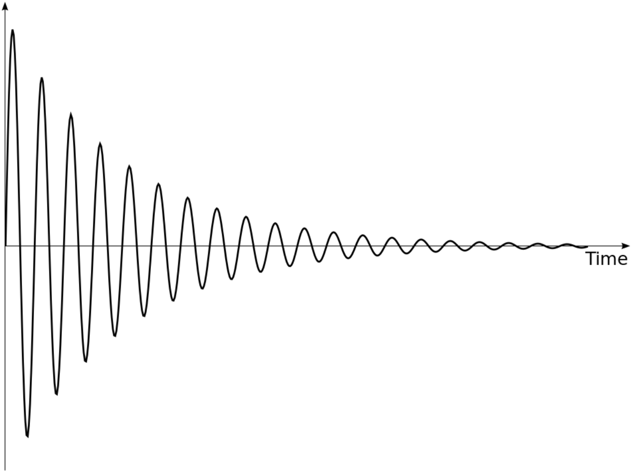
# Nuclear Quadrupole Resonance (NQR) spectroscopy is a chemical analysis technique used for detection



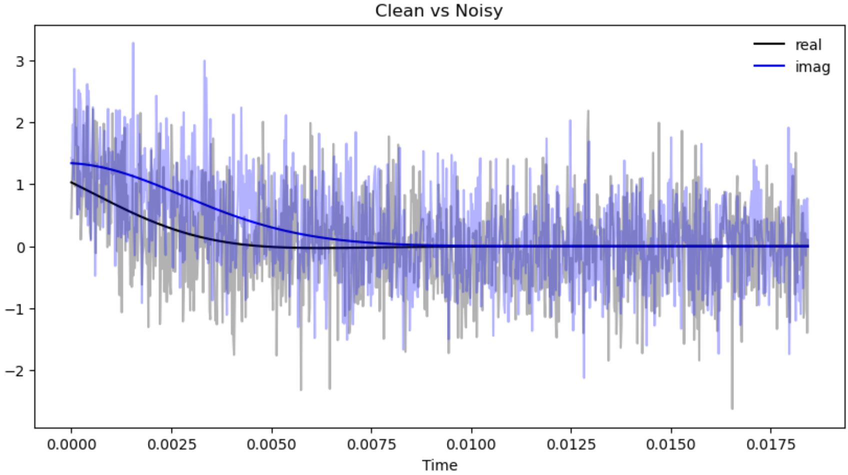

When the atoms return to equilibrium, after the excitation has ended, the excess energy released takes the form of a measurable radio frequency signal, called free induction decay (FID).

[2] Monea, C., Bizon, N. (2021). *Signal Processing and Analysis Techniques for Nuclear Quadrupole Resonance Spectroscopy.* Springer Publishing.

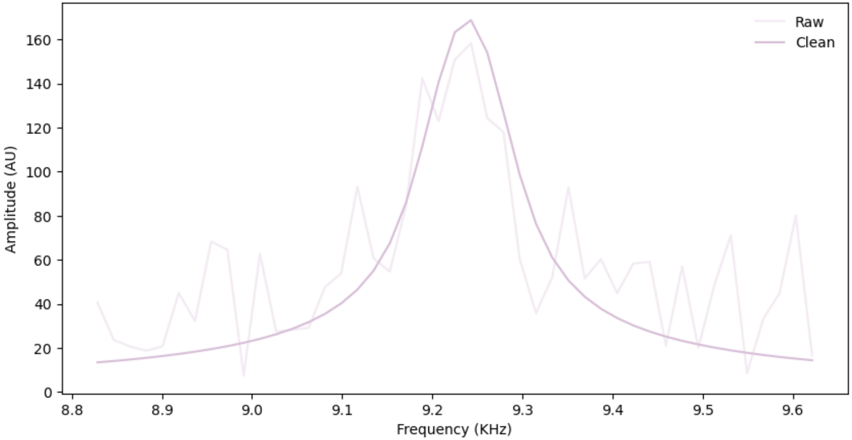
# Example of measured radio frequency signal



# Complex demodulation brings the component of interest near zero frequency.

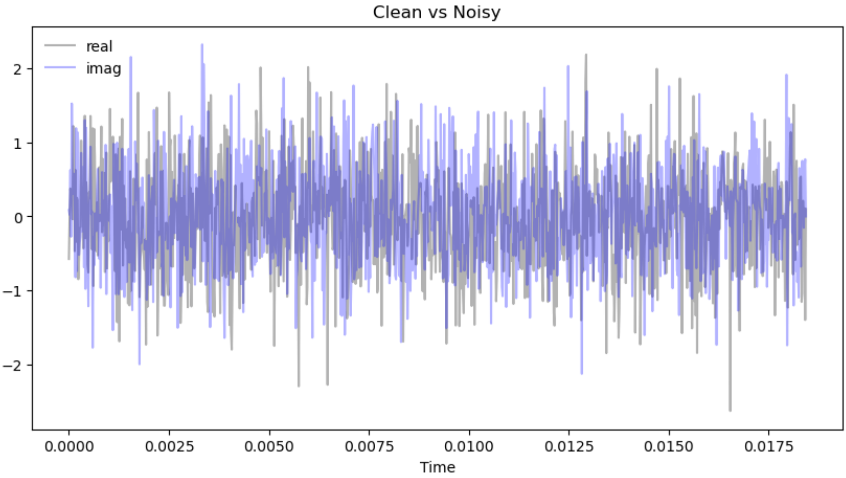


# Complex demodulation brings the component of interest near zero frequency.

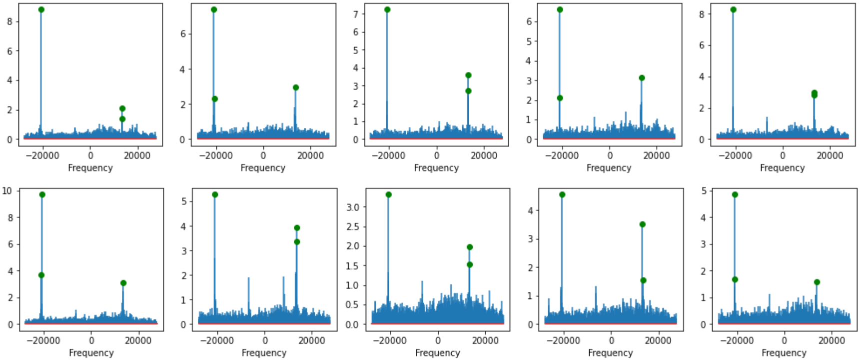


# Table of Contents

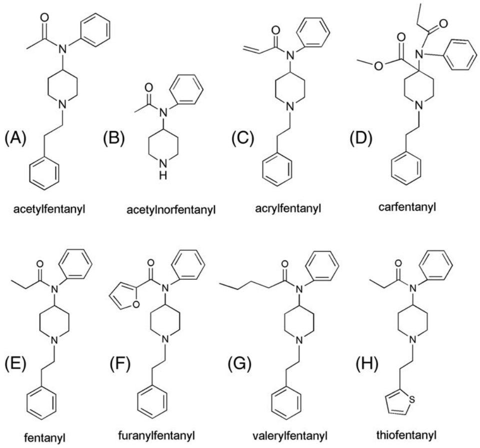
# Commonly occurring radio frequency interference.



# Commonly occurring radio frequency interference.



# Increasing number of new compounds of fentanyl analogs.




Acquiring the $^{1}HT_{1}$ dispersion curve for each fentanyl analog can take months of conducting experiments.

# Table of Contents

# Two Statistics Goals

1. Denoise low signal-to-noise ratio NQR measurements.
2. Speed up acquisition of the $^1$H$T_1$ dispersion curves for fentanyl analogs.

# Ultimate Goal

Develop a technology capable of detecting the presence of synthetic opioids in unopened packages using NQR spectroscopy.

# Ultimate Goal

Develop a technology capable of detecting the presence of synthetic opioids in unopened packages using NQR spectroscopy.

This research is being conducted by a group of physicists and chemists at Los Alamos National Laboratories led by principal investigators Natalie Klein, Michael Malone, and Harris Mason.

# Table of Contents

# Data Generation

We utilize synthetically generated time-varying signals $y(t)$ following a noisy decaying complex exponential (Voigt profile[3]),

$$y(t) = A \exp\left(-\frac{t^2}{2\sigma^2} - \frac{t}{T_2} + i[2\pi wt + \phi]\right) + \epsilon(t),$$

where $i$ is the imaginary unit, $A$ is the initial amplitude, $\sigma$ and $T_2$ control the rate of decay, $\phi$ is the phase offset, $w$ is the frequency, and $\epsilon$ is additive noise.

[3] K Ramani, S Ganapathy, and R Srinivasan, "A Fourier transform approach to Voigt profile analysis and its ap- plication to nuclear magnetic resonance," *Journal of Magnetic Resonance (1969)*, vol. 24, no. 2, pp. 231– 237, 1976.

# Datasets

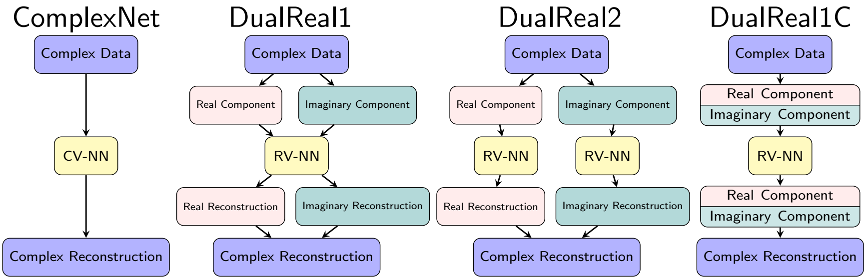
We generate eight datasets by selecting from two noise distributions, two signal frequency bands centered at $f_0 = 0$Hz and $f_0 = F = 1600$Hz, and two signal-to-noise ratio regimes (low and high).

# Comparison between two different types of Neural Networks, Autoencoder (AE) and ConvTasNet(CVN)

We used two different kinds (autoencoders and ConvTasNet[4]) of real-valued neural networks and complex-valued neural networks for signal denoising.
ConvTasNet is a fully-convolutional network which utilizes an encoder and decoer, similar to an autoencoder, but explicitly learns to separate multiple sources (e.g., signal and noise) with a masking module.

[4]Yi Luo and Nima Mesgarani, "Conv-TasNet: Surpass- ing ideal time–frequency magnitude masking for speech separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 8, pp. 1256– 1266, 2019.

# Complex-valued (CV) and real-valued (RV) Neural Network (NN) Architectures

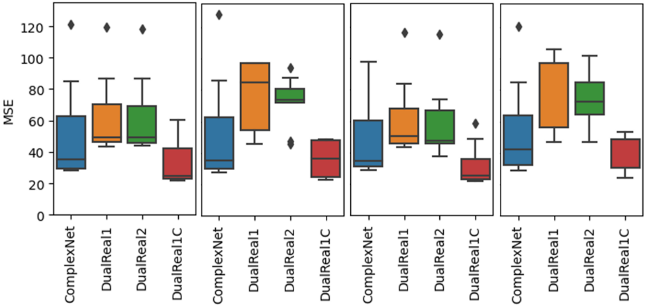



ComplexNet natively processes complex-valued signals[5].

[5] C Trabelsi, O Bilaniuk, Y Zhang, D Serdyuk, S Subra- manian, JF Santos, S Mehri, N Rostamzadeh, Y Bengio, and CJ Pal, "Deep complex networks," *arXiv preprint arXiv:1705.09792, 2017.*

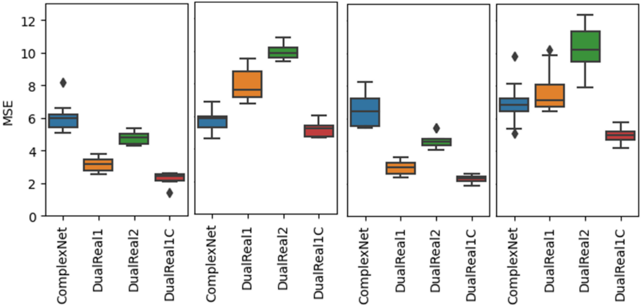
# Complex- and Real-valued autoencoder MSE comparisons

| | White noise | | | | Measured noise | | | |
|---|---|---|---|---|---|---|---|---|
| | $f \sim 0$ | | $f \sim F$ | | $f \sim 0$ | | $f \sim F$ | |
| Model | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| ComplexNet | 50.6 | 29.8 | 51.4 | 31.7 | 48.1 | 23.6 | 53.3 | 28.4 |
| DualReal1 | 61.8 | 23.5 | 77.2 | 21.3 | 60.9 | 22.3 | 81.4 | 22.9 |
| DualReal2 | 61.4 | 23.3 | 72.7 | 14.7 | 58.0 | 22.2 | 74.2 | 17.6 |
| DualReal1C | 33.0 | 13.0 | 36.1 | 10.8 | 32.1 | 12.1 | 41.9 | 11.5 |



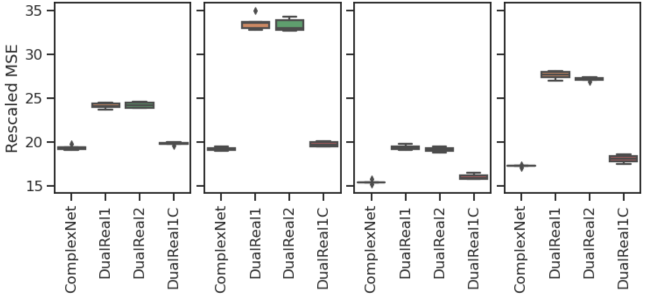
# Complex- and Real-valued autoencoder MSE comparisons

| | White noise | | | | Measured noise | | | |
|---|---|---|---|---|---|---|---|---|
| | $f \sim 0$ | | $f \sim F$ | | $f \sim 0$ | | $f \sim F$ | |
| Model | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| ComplexNet | 6.03 | 0.87 | 5.80 | 0.67 | 6.51 | 0.96 | 6.94 | 1.28 |
| DualReal1 | 3.16 | 0.44 | 7.95 | 0.92 | 2.99 | 0.42 | 7.70 | 1.28 |
| DualReal2 | 4.78 | 0.37 | 9.96 | 0.48 | 4.65 | 0.44 | 10.4 | 1.34 |
| DualReal1C | 4.96 | 0.46 | 2.30 | 0.34 | 5.19 | 0.43 | 2.30 | 0.23 |



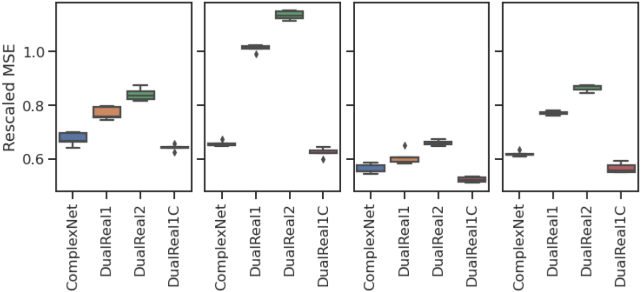
# Complex- and Real-valued ConvTasNet MSE comparisons

| | White noise | | | | Measured noise | | | |
|---|---|---|---|---|---|---|---|---|
| | $f \sim 0$ | | $f \sim F$ | | $f \sim 0$ | | $f \sim F$ | |
| Model | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| ComplexNet | 19.38 | 0.23 | 19.24 | 0.16 | 15.43 | 0.21 | 17.25 | 0.10 |
| DualReal1 | 24.15 | 0.28 | 33.59 | 0.74 | 19.34 | 0.24 | 27.62 | 0.38 |
| DualReal2 | 24.19 | 0.30 | 33.29 | 0.65 | 19.14 | 0.23 | 27.18 | 0.18 |
| DualReal1C | 19.80 | 0.13 | 19.74 | 0.27 | 16.04 | 0.30 | 18.07 | 0.40 |

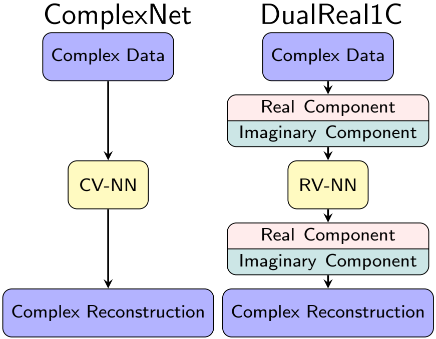


# Complex- and Real-valued ConvTasNet MSE comparisons

| | White noise | | | | Measured noise | | | |
|---|---|---|---|---|---|---|---|---|
| | $f \sim 0$ | | $f \sim F$ | | $f \sim 0$ | | $f \sim F$ | |
| Model | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| ComplexNet | 0.67 | 0.02 | 0.66 | 0.01 | 0.56 | 0.02 | 0.62 | 0.01 |
| DualReal1 | 0.77 | 0.02 | 1.01 | 0.01 | 0.61 | 0.02 | 0.77 | 0.01 |
| DualReal2 | 0.84 | 0.02 | 1.13 | 0.01 | 0.66 | 0.01 | 0.86 | 0.01 |
| DualReal1C | 0.64 | 0.01 | 0.62 | 0.01 | 0.52 | 0.01 | 0.57 | 0.02 |



# Complex-valued (CV) and real-valued (RV) Neural Network (NN) Architectures



# Comparison of test set ensemble performance across the eight datasets

- Compare test set ensemble performance across the eight datasets.
- Compare two model architectures: autoencoder (AE) and ConvTasNet (CTN).
- Compare two complex-valued architectures: CV-NN and RV-NN.

We use a scaled $R^2$ metric (best possible $R^2 = 100$)

$$R^2 = 100 \times \left(1 - \frac{\sum_i \sum_t (y_i(t) - \hat{y}_i(t))^2}{\sum_i \sum_t y_i(t)^2}\right)$$

- $y_i(t)$ is the true clean signal instance $i$ at time $t$
- $\hat{y}_i(t)$ is the corresponding denoised signal prediction

# Comparisons of results from RV-NNs and CV-NNs

Table: High SNR regime NN scaled $R^2$ values.

| | White noise | | | | Measured noise | | | |
|---|---|---|---|---|---|---|---|---|
| | $f \sim 0$ | | $f \sim F$ | | $f \sim 0$ | | $f \sim F$ | |
| Model | CV | RV | CV | RV | CV | RV | CV | RV |
| AE | 96.0 | 98.1 | 96.1 | 95.5 | 95.5 | 98.1 | 95.2 | 95.8 |
| CTN | **99.4** | **99.4** | **99.5** | **99.5** | 99.5 | **99.6** | **99.5** | **99.5** |

Table: Low SNR regime NN scaled $R^2$ values.

| | White noise | | | | Measured noise | | | |
|---|---|---|---|---|---|---|---|---|
| | $f \sim 0$ | | $f \sim F$ | | $f \sim 0$ | | $f \sim F$ | |
| Model | CV | RV | CV | RV | CV | RV | CV | RV |
| AE | 67.1 | 76.9 | 65.9 | 75.8 | 72.2 | 77.6 | 65.5 | 70.0 |
| CTN | **81.7** | 81.1 | **81.8** | 81.2 | **85.5** | 84.9 | **83.7** | 82.9 |

# Comparison of NNs to baseline methods

Comparison of RV-AE and CV-CTN to three baseline methods: time-domain deconvolution (DC; in which we fit the signal model of Eq. 1 to noisy data via nonlinear least squares), singular spectrum analysis (SSA; a subspace-based method), and wavelet decomposition (W).

# Comparison of NNs to baseline methods

Table: High SNR regime scaled $R^2$ values.

| | White noise | | Measured noise | |
|---|---|---|---|---|
| Method | $f \sim 0$ | $f \sim F$ | $f \sim 0$ | $f \sim F$ |
| RV-AE | 98.1 | 95.5 | 98.1 | 95.8 |
| CV-CTN | **99.4** | **99.5** | **99.5** | **99.5** |
| DC | 99.2 | 99.1 | 99.0 | **99.5** |
| SSA | 98.6 | 98.6 | 52.4 | 52.6 |
| W | 97.8 | 87.0 | 98.4 | 89.0 |

# Comparison of NNs to baseline methods

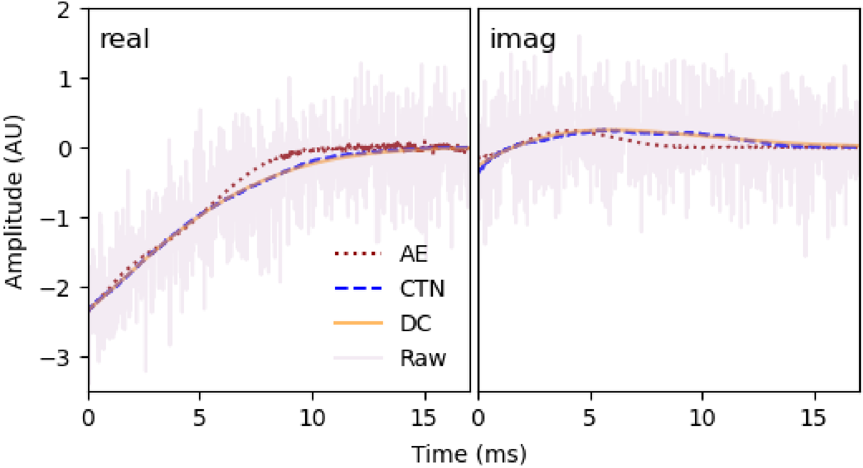
Table: Low SNR regime scaled $R^2$ values.

| | White noise | | Measured noise | |
|---|---|---|---|---|
| Method | $f \sim 0$ | $f \sim F$ | $f \sim 0$ | $f \sim F$ |
| RV-AE | 76.9 | 75.8 | 77.6 | 70.0 |
| CV-CTN | **81.7** | **81.8** | **85.5** | **83.7** |
| DC | 57.7 | 65.5 | 68.0 | 62.4 |
| SSA | -42.6 | -41.7 | -1970.2 | -1970.2 |
| W | 21.5 | -38.1 | 51.5 | -10.9 |

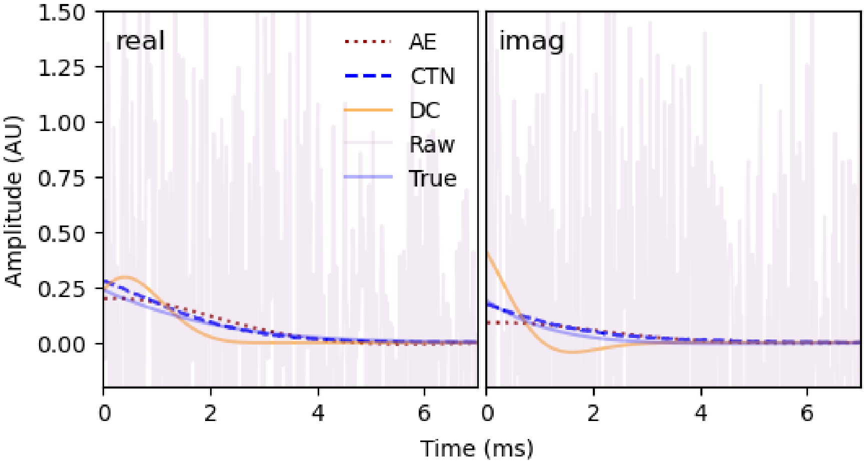
# Application to experimental data

We apply the best-performing NNs (CV-CTN and RV-AE) and the best high signal-to-noise ratio (SNR) baseline method (DC) to experimental data that exhibits signal frequency near zero and high SNR, suggesting NNs trained on $f \sim 0$, high SNR data are suitable.

# Application to experimental data



# Application to realistic low SNR data



# Application to out of distribution datasets

| SNR | $f$ | w $\rightarrow$ w | w $\rightarrow$ m | m $\rightarrow$ m | m $\rightarrow$ w |
|---|---|---|---|---|---|
| high | 0 | 0.54(0.04) | 0.44(0.03) | 0.44(0.03) | 0.56(0.03) |
| high | $F$ | 0.50(0.02) | 0.48(0.02) | 0.45(0.02) | 0.51(0.02) |
| low | 0 | 17.9(0.23) | 14.6(0.19) | 14.1(0.21) | 18.3(0.28) |
| low | $F$ | 17.8(0.16) | 17.1(0.26) | 15.8(0.10) | 18.2(0.17) |

Table: Ensemble MSE(SD) $\times 10^{-3}$ for CTN (CV) models.

| SNR | $f$ | w $\rightarrow$ w | w $\rightarrow$ m | m $\rightarrow$ m | m $\rightarrow$ w |
|---|---|---|---|---|---|
| high | 0 | 1.86(0.25) | 1.81(0.16) | 1.89(0.13) | 1.95(0.12) |
| high | $F$ | 4.36(0.22) | 4.28(0.23) | 4.10(0.35) | 4.17(0.34) |
| low | 0 | 22.5(5.68) | 22.3(0.06) | 21.8(0.87) | 22.4(8.44) |
| low | $F$ | 23.6(4.83) | 23.4(5.28) | 29.3(6.34) | 29.5(5.75) |

Table: Ensemble MSE(SD) $\times 10^{-3}$ for AE (DualReal1C) models.

# Application to out of distribution datasets

| | White noise | | | | Measured noise | | | |
|---|---|---|---|---|---|---|---|---|
| Trained | $f \sim 0$ | | $f \sim 1600$ | | $f \sim 0$ | | $f \sim 1600$ | |
| Tested | $f \sim 0$ | $f \sim 300$ | $f \sim 1600$ | $f \sim 1900$ | $f \sim 0$ | $f \sim 300$ | $f \sim 1600$ | $f \sim 1900$ |
| AE | 1.86(0.25) | 29.5(0.32) | 4.36(0.22) | 35.0(0.18) | 1.89(0.13) | 29.7(0.09) | 4.10(0.35) | 34.7(0.15) |
| CTN | 0.54(0.04) | 16.3(1.26) | 0.50(0.02) | 21.5(3.44) | 0.44(0.03) | 14.7(1.42) | 0.45(0.02) | 18.1(4.09) |

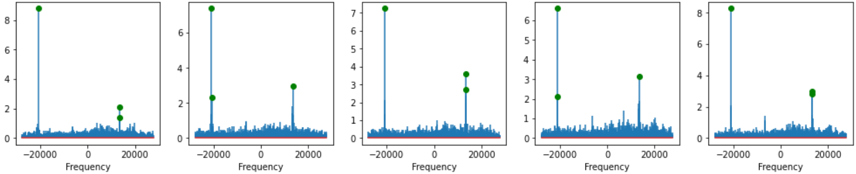
(a) Relative MSE for out of distribution high-SNR data.

| | White noise | | | | Measured noise | | | |
|---|---|---|---|---|---|---|---|---|
| Trained | $f \sim 0$ | | $f \sim 1600$ | | $f \sim 0$ | | $f \sim 1600$ | |
| Tested | $f \sim 0$ | $f \sim 300$ | $f \sim 1600$ | $f \sim 1900$ | $f \sim 0$ | $f \sim 300$ | $f \sim 1600$ | $f \sim 1900$ |
| AE | 22.5(5.68) | 47.1(1.41) | 23.6(4.83) | 48.6(1.13) | 21.8(0.87) | 47.1(1.32) | 29.3(6.34) | 48.4(1.16) |
| CTN | 17.9(0.23) | 65.2(0.24) | 17.8(0.16) | 65.9(0.45) | 14.1(0.21) | 59.5(0.26) | 15.8(0.10) | 63.2(0.67) |

(b) Relative MSE for out of distribution low-SNR data.

# Classification

We developed a binary classifier using complex-valued feed-forward neural networks, that predicts the presence of a substance of interest based on the denoised signal. The accuracy of the classifier is highly dependent on the SNR. With moderate noise levels the classifier obtains $92.5\%$ accuracy on the test set while more extreme levels of noise bring this down to $47\%$.



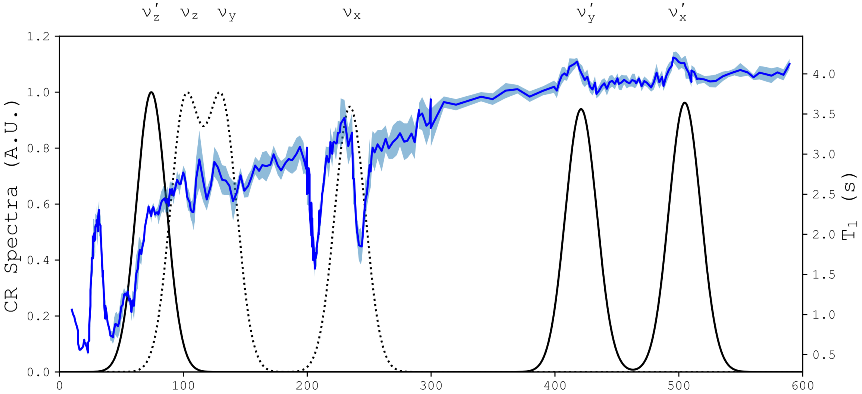
# Future work

- Create classifiers which instead use frequency-domain signal as input.
- Explore the relationship between real and imaginary components.
- Denoise Experimental Data.

# Table of Contents

# Acquiring the $^{1}HT_1$ dispersion curve for each fentanyl analog can take months of conducting experiments.

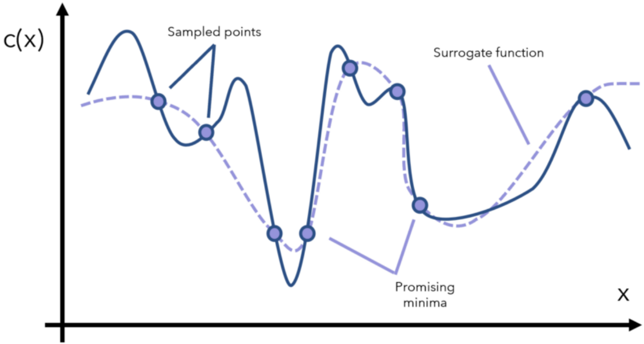
[6]




[6] Malone MW, Espy MA, He S, Janicke MT, Williams RF. (2020). The $^{1}H$ $T_1$ dispersion curve of fentanyl citrate to identify NQR parameters. *Solid state nuclear magnetic resonance, 110,* 101697. https://doi.org/10.1016/j.ssnmr.2020.101697

# Speed up acquisition of the $^{1}HT_{1}$ dispersion curves for fentanyl analogs.

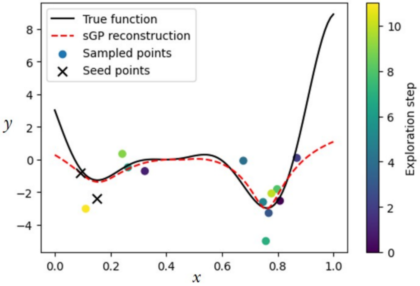
[7]




[7] Ye, A. (2020, October 31). *The Beauty of Bayesian Optimization, Explained in Simple Terms.* Medium. https://towardsdatascience.com/the-beauty-of-bayesian-optimization-explained-in-simple-terms-81f3ee13b10f

# Bayesian optimization and active learning via augmented Gaussian process

We plan to explore the use of Bayesian optimization (BO) and active learning via augmented Gaussian process to make statistically informed decisions on which subset of frequencies to conduct experiments on. [8]




[8] Ziatdinov, Maxim A., Ghosh, Ayana, Kalinin, Sergei V. *Physics makes the difference: Bayesian optimization and active learning via augmented Gaussian process.* United Kingdom. https://doi.org/10.1088/2632-2153/ac4baa

# Speed up acquisition of the $^{1}\mathrm{H}T_1$ dispersion curves for fentanyl analogs.

- Need new techniques capable of locating local optima amongst high levels of noise while exploiting prior knowledge of a physical model.
- This will increase experimental efficiency, making the scanners adaptable and capable of detecting new synthetic opioids as they are discovered.

# Special Thanks To:


- Natalie Klein
- Sinead Williamson

# Questions