# Yihong Wu

yihong.wu@umontreal.ca | 514-578-7456|

## EDUCATION

**BEng Chemical Engineering/ Minor Mathematics** — Dalian, Liaoning | Sep 2015 - Jun 2019
Dalian University of Technology

**MSc Engineering Mathematics** — Bristol, England | Oct 2020 - Sep 2021
University of Bristol

**MSc Computer Science GPA: 4.0/4.0** — Montreal, Quebec | Sep 2021 - Jun 2023
Université de Montréal

## RESEARCH EXPERIENCE

### UNIVERSITY OF BRISTOL | Autonomous Protocols

- Designed real-time autonomously switching mechanism for large scale protocols (collaborated with The Defence Science and Technology Laboratory (Dstl), UK) supervised by Alan Champneys.
- Established mathematical model for real world applications; designed and implemented three algorithms.

### UNIVERSITY OF BRISTOL | Training Deep Neural Network by Homotopy Methods

- Master thesis on training deep neural network by homotopy methods supervised by Robert Szalai.
- Used homotopy (continuation) methods to train neural network, namely, Natural Parameter Continuation and Pseudo Arc-length Continuation, to explore the possibility beyond traditional first-order optimization.
- With Application on dynamical system, Shaw-Pierre's model.

### UNIVERSITÉ DE MONTRÉAL | Melodic Phrase Segmentation Using Conditional Random Field

- Implementing core CRF algorithm, $\alpha - \beta$ recursion.
- Using CRF algorithm as backend for symbolic music melodic phrase segmentation. For hand-crafted feature, we achieved f1-score of $61.31\%$ by reimplement while sklearn achieved f1-score of $67.17\%$. The LSTM-derived feature achieved f1-score of $74.33\%$.

### UNIVERSITÉ DE MONTRÉAL | Weight Perturbation Methods to Train Neural Networks

- With the new proposal of forward gradient method, we revisit SPSA (Simultaneous perturbation stochastic approximation) and a bigger category of optimization algorithm, weight perturbation methods.
- However, these methods fails in modern deep learning due to intolerant variance. We provide a matrix perspective to explain this failure and connect compressed sensing for a meaningful lower bound.

### UNIVERSITÉ DE MONTRÉAL | Contrastive Learning for Causal Representation Learning

- We investigate constrastive learning with Siamese network architecture for causal (invariant) feature learning.
- As a feature extractor, the Siamese network fails to learn the causal features (the shape of digits) in the supervised contrastive learning on the CMNIST dataset with label flipping opposite to the no flipping setting.
- Furthermore, we investigate the influence of label flipping in the CMNIST dataset. The low generalization on CMNIST is not due to multiple environments but to the label flipping increasing learning algorithms' perplexity.

## SKILLS

**Languages:** Python, C/C++, MATLAB, Java
**Package:** PyTorch, scikit-learn
**Technology:** LaTeX, Git, Docker

## COURSES

Math: real analysis, complex analysis, ODE, PDE, differential geometry, measure theory, stochastic process
Computer Science: data structure, algorithms, computer system, machine learning, probability graphical model, deep learning, causal inference, learning theory, convex optimization