On the Uncertainty Calibration of Equivariant Functions

Edward Berman

Northeastern University

Boston, MA, March 3, 2025

Collaborators


Figure: Jake Ginesin (left) and Robin Walters (right)

Outline

Motivation

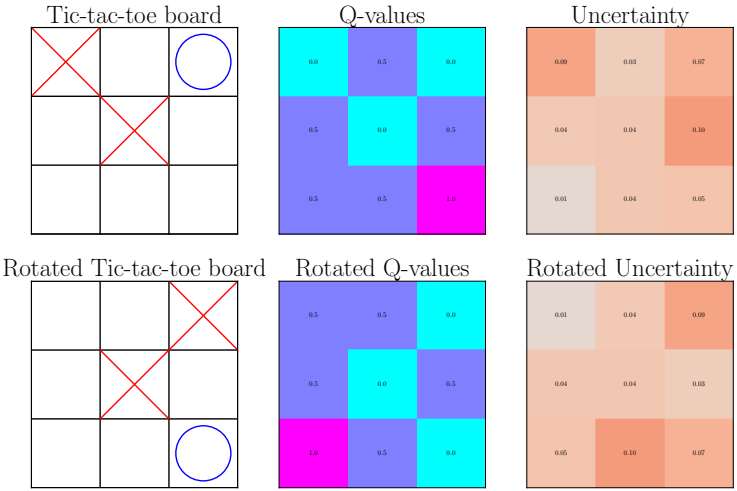


Figure: Equivariance with Uncertainty!

Motivation

Equivariance is a property of a function that allows for us to reason about group symmetries in neural networks and beyond.

- More sample efficient (auto-generalization across symmetry)
-
-

Motivation

Equivariance is a property of a function that allows for us to reason about group symmetries in neural networks and beyond.

- More sample efficient (auto-generalization across symmetry)
- More parameter efficient
-

Motivation

Equivariance is a property of a function that allows for us to reason about group symmetries in neural networks and beyond.

- More sample efficient (auto-generalization across symmetry)
- More parameter efficient
- Are they better **calibrated?**

Motivation

Preliminary works argue yes, but more work required

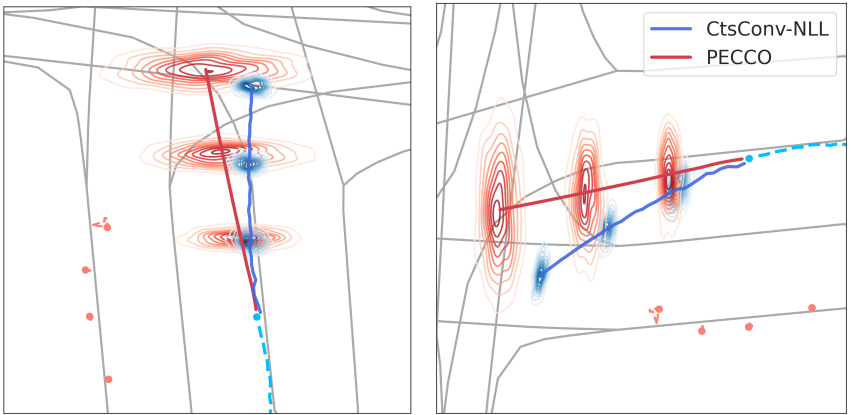


Figure: Equivariant Countour from `https://arxiv.org/pdf/2205.01927`

Outline

Equivariance

Equivariance is a property of an operator $\phi : \mathcal{X} \to \mathcal{Y}$ that maps between input and output vector spaces $\mathcal{X}$ and $\mathcal{Y}$. Given a group G and its representations $\rho^{\mathcal{X}}$ and $\rho^{\mathcal{Y}}$ which transform vectors in $\mathcal{X}$ and $\mathcal{Y}$ respectively, an *operator $\phi : \mathcal{X} \to \mathcal{Y}$ is said to be equivariant if it satisfies the following constraint*

$$\rho^{\mathcal{Y}}(g)[\phi(x)] = \phi(\rho^{\mathcal{X}}(g)[x]) \ , \ \text{for all } g \in G, x \in \mathcal{X} \ . \tag{1}$$

Invariance is a special case of equivariance in which $\rho^{\mathcal{Y}} = \mathcal{I}^{\mathcal{Y}}$ for all $g \in G$. I.e., *an operator $\phi : \mathcal{X} \to \mathcal{Y}$ is said to be invariant if it satisfies the following constraint*

$$\phi(x) = \phi(\rho^{\mathcal{X}}(g)[x]) \ , \ \text{for all } g \in G, x \in \mathcal{X} \ . \tag{2}$$

Equivariance

Examples

1. Galaxy Morphology Classification
2. Chemical Compound Spectral Lines
3. Pick and Place Robotics

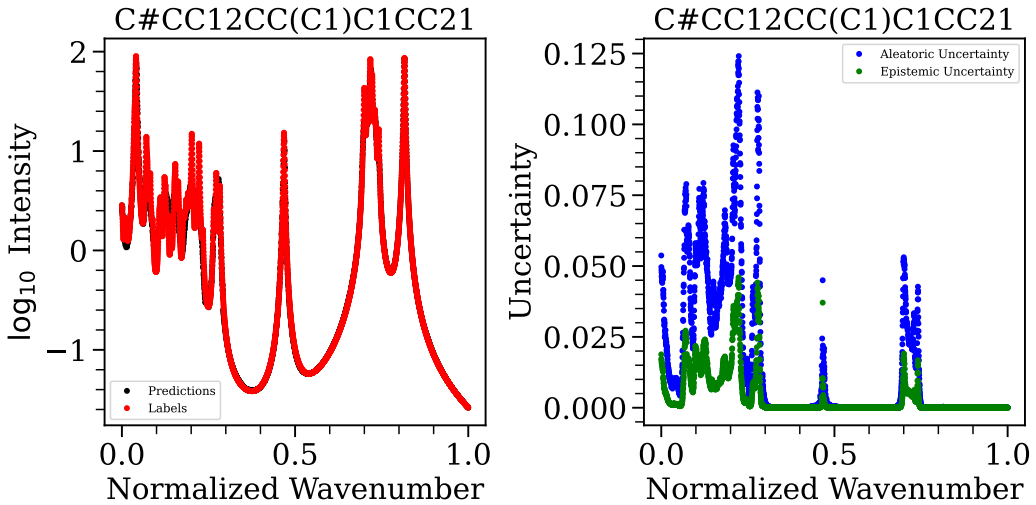
Equivariance

Galaxy Morphology Classification: $E(2)$ Classification Invariance

Figure: Galaxy Zoo Galaxies

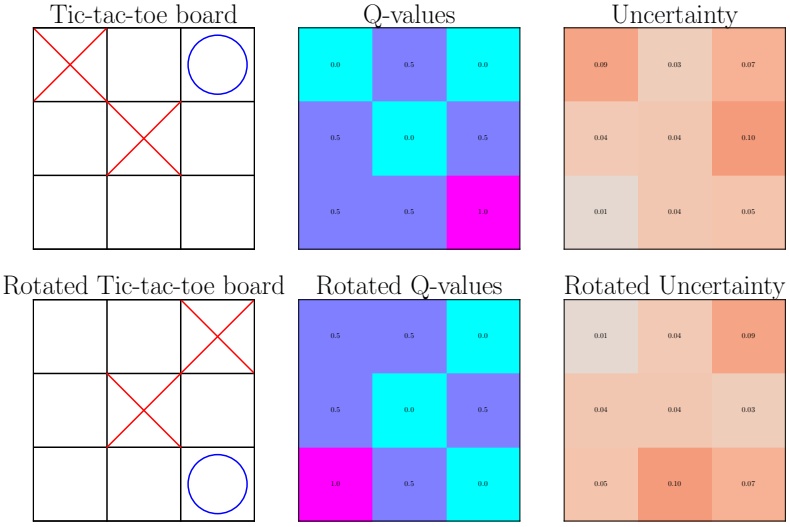
Equivariance

Chemical Compound Spectral Lines: $E(3)$ Regression Invariance



Equivariance

Pick and Place Robotics: $E(3)$ Regression Equivariance



Outline

Uncertainty

As outlined by our last slide, we seek to make rigorous the idea of calibration. We have the ground truth classification given by $f : X \to Y$ and a model class of function $\{h : X \to Y \times P\}$ used to approximate f.

$$\underset{\text{Classification}}{ECE}(h) = \mathbb{E}_{h_P}\left[\left| \mathbb{P}\left(f_Y = h_Y | h_P = p\right) - \underset{\text{Confidence}}{p} \right|\right], \tag{3}$$

$\underset{\text{Accuracy}}{\mathbb{P}\left(f_Y = h_Y | h_P = p\right)}$

we are well calibrated if a model's predicted **confidence** (as given by the logit) matches a model's **accuracy**.

- "If I am 80% confident, I should be correct 80% of the time"

Uncertainty

We can study regression analogously. Now, we have a model class $\{h : X \to \mu \times \sigma^2\}$ for μ, $\sigma^2 \in \mathbb{R}^n$ and $\mathbb{R}^n_+$ respectively. We are well calibrated if

$$\mathbb{E}[\|f - h_\mu\|_2^2 \;\Big|\; h_{\sigma^2} = \sigma^2] = \sigma^2 \qquad \forall \sigma^2 \in [0, \infty). \tag{4}$$

Uncertainty

Equipped with these notions of uncertainty, we wish to understand how equivariance (or invariance) on the model class $\{h\}$ can effect our calibration.

Outline

Invariant Regression — A Case Study on Chemical Spectra

Our first goal is see if we can upper bound the calibration error for a class of functions $\{h : X \to \mu \times \sigma^2\}$ that are arbitrarily expressive except for being constrained to be invariant with respect to a group G.

Intuition: Let $e(x)$ be the error vector defined by taking the squared error between $f_Y(x)$ and $h_\mu(x)$ on each coordinate, $e(x)_i = (f_Y(x)_i - h_\mu(x)_i)^2$.

$$\mathbb{E}_{x,y}\left[\left\|\sigma^2 - e(x)\right\|_2^2\right] \leq \mathbb{E}_{x,y}\left[\left\|\sigma^2\right\|_2^2 + \left\|e(x)\right\|_2^2\right],$$

the miscalibration is in some sense upper bounded by the error.

Invariant Regression — A Case Study on Chemical Spectra

Now, we consider the error on individual orbits in the input domain X. Specifically, by assumption of $G-$invariance, $h_{\sigma^2}(x) = \sigma^2 \implies h_{\sigma^2}(gx) = \sigma^2$.
In other words, every variance predicted by the model corresponds to (at least) one orbit in the domain X.

At a high level, we bound the calibration error by considering the regression error on individual orbits induced by σ^2.

Indeed, the upper bound comes out to be

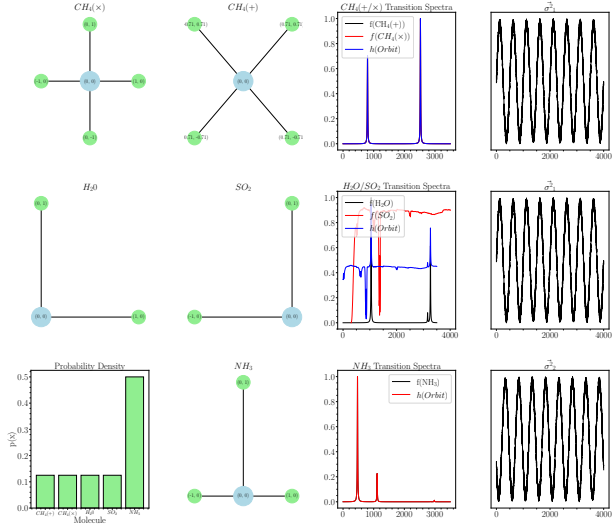
$$ENCE \leq 1 + \mathbb{E}\left[\frac{\mathrm{err}_{\mathrm{reg}}(h, \sigma^2)}{\|\sigma\|_2^2}\right],$$

the average regression error on orbits induced by σ^2!

Invariant Regression — A Case Study on Chemical Spectra

We now explore this on a fabricated example of chemical compounds. For the size and shapes of the compounds,..., hold a suspension of disbelief!

Invariant Regression — A Case Study on Chemical Spectra



Invariant Regression — A Case Study on Chemical Spectra

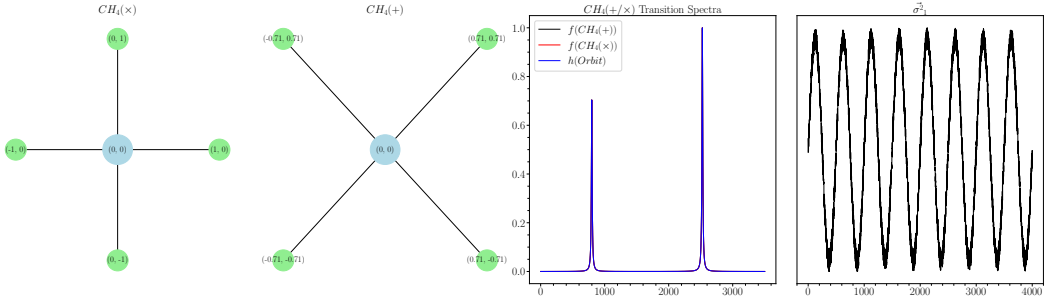


Figure: Methane Spectra

Invariant Regression — A Case Study on Chemical Spectra

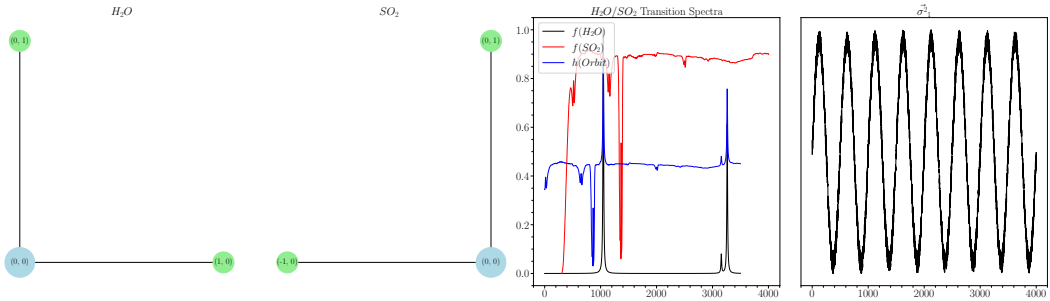


Figure: Polarized Molecule

Invariant Regression — A Case Study on Chemical Spectra

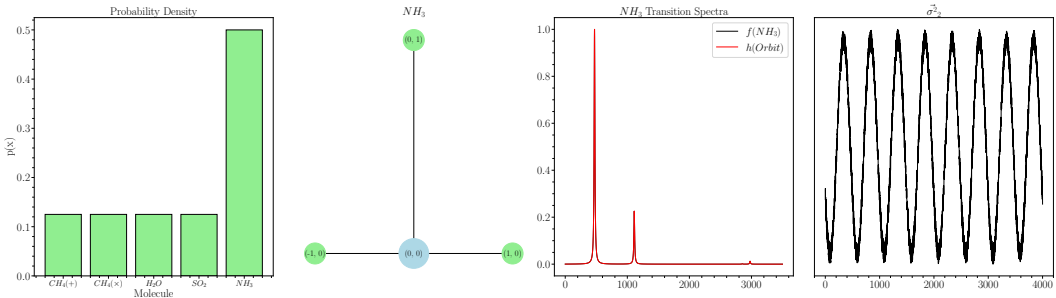


Figure: Ammonia Molecule

Invariant Regression — A Case Study on Chemical Spectra

From theory to practice! We train an $E(3)$ invariant model to predict both a spectral line and an uncertainty estimate on QM9s molecules.

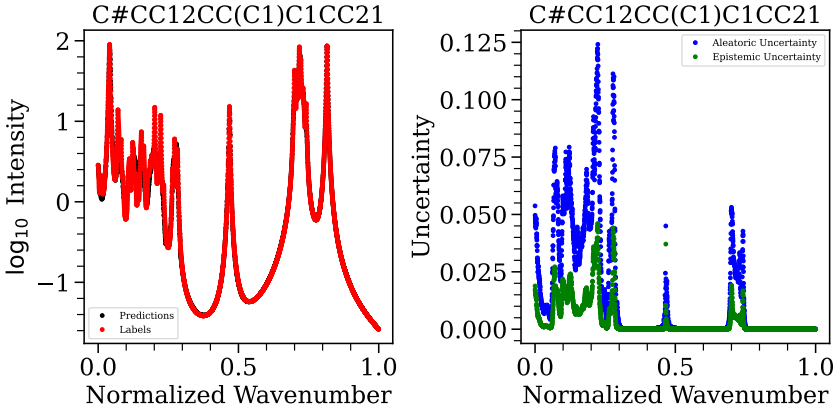


Figure: Spectral Lines

Due to binning approximations, it can be hard to estimate the true calibration error from data. However, regardless of binning approximation, we find that our bound is satisfied!

Invariant Regression — A Case Study on Chemical Spectra

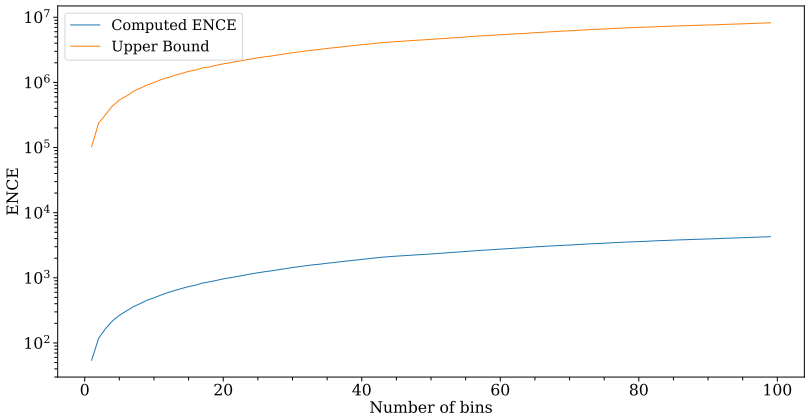
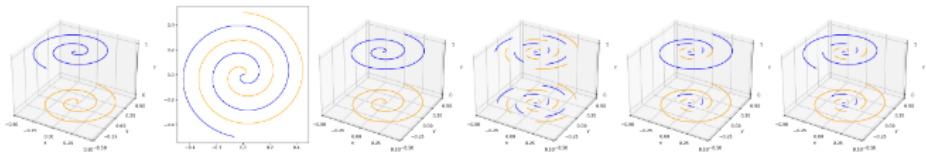


Figure: Upper Bound vs Computation

Outline

Equivariance Type on Swiss Roll Classification



(a) Swiss Roll Data | (b) C_2-Invariant View | (c) Incorrect | (d) Correct | (e) $c = .5, i = .5, e = 0$ | (f) $c = .5, i = 0, e = .5$

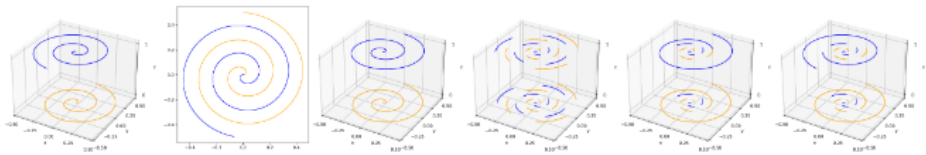
Figure 7: (a) (b) The Swiss Roll data distribution that leads to harmful extrinsic equivariance. (c) (d) The correct and incorrect data distribution in the Swiss Roll experiment. Here the spirals overlap with mismatched and matched labels respectively. (e) (f) Data distribution example with different correct ratio (c), incorrect ratio (i), and extrinsic ratio (e) values.

Figure: Swiss Roll Distributions from `https://arxiv.org/pdf/2303.04745`

Equivariance Type on Swiss Roll Classification

The Goal is to predict the color (blue/yellow) from the position (x,y,z). We sample in different proportions from the Correct and Incorrect Swiss Roll Distributions.

Equivariance Type on Swiss Roll Classification



(a) Swiss Roll Data | (b) C_2-Invariant View | (c) Incorrect | (d) Correct | (e) $c = .5, i = .5, e = 0$ | (f) $c = .5, i = 0, e = .5$

Figure 7: (a) (b) The Swiss Roll data distribution that leads to harmful extrinsic equivariance. (c) (d) The correct and incorrect data distribution in the Swiss Roll experiment. Here the spirals overlap with mismatched and matched labels respectively. (e) (f) Data distribution example with different correct ratio (c), incorrect ratio (i), and extrinsic ratio (e) values.

Figure: Swiss Roll Distributions from `https://arxiv.org/pdf/2303.04745`

Equivariance Type on Swiss Roll Classification

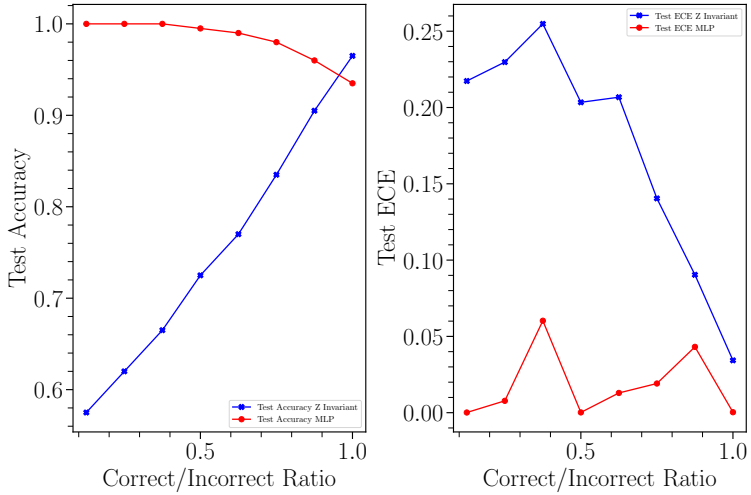


Figure: Accuracy and ECE as a function of correct equivariance in the dataset

Group Order on Galaxy Morphology Classification

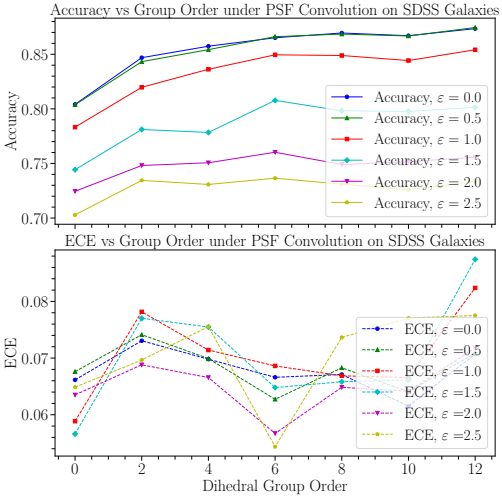


Figure: Galaxy Morphological Accuracy and Calibration vs. Group Order

Questions? Contact berman.ed@northeastern.edu or visit https://ebrmn.space/

Applying to PhD programs next fall :D