Can You Hear the Shape of A Jet?

An IAIFI Story

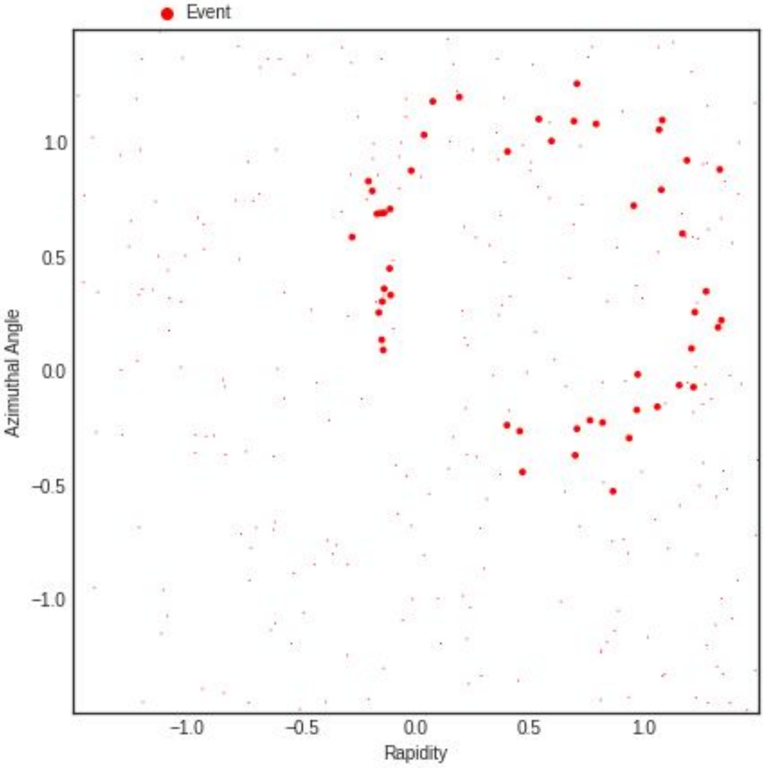
Rikab Gambhir

With Akshunna S. Dogra ( ),

Demba Ba ( ),

& Jesse Thaler ( )

Fundamental Question: What shape is this?

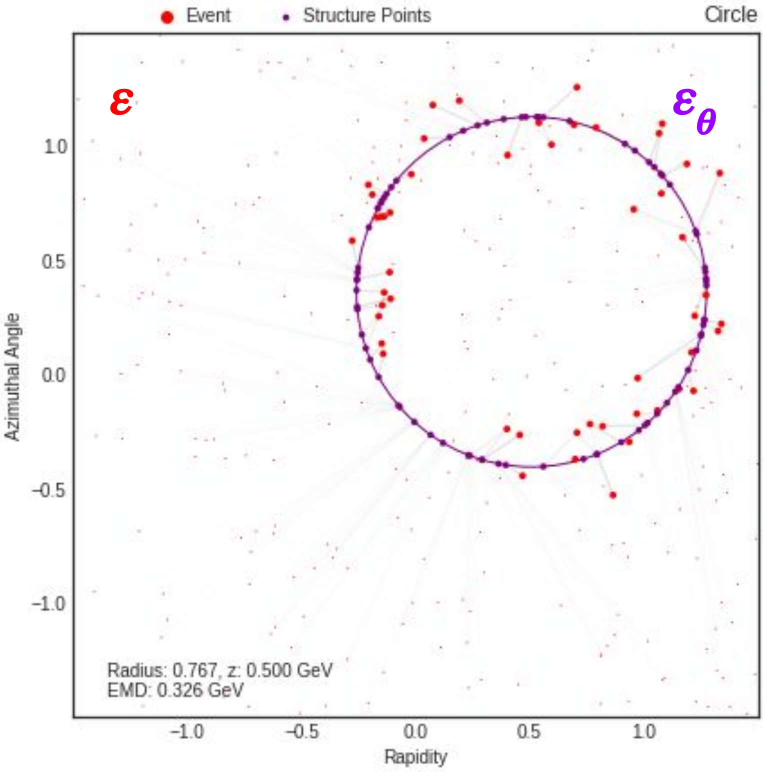


Pictured: (Fake) event that you might have measured at the LHC

Red dots are detector hits on a patch of the LHC cylinder, weighted by energy

Goal: Construct an observable $\mathcal{O}$ that generically answers this question!

Fundamental Question: What shape is this?



Using the **SHAPER** framework and optimal transport ……

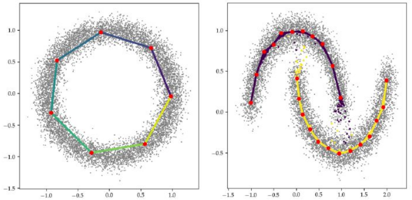
$$\mathcal{O}_{\mathcal{M}}(\mathcal{E}) = \min_{\mathcal{E}'_\theta \in \mathcal{M}} \mathrm{EMD}(\mathcal{E}, \mathcal{E}'_\theta)$$

$$\theta = \underset{\mathcal{E}'_\theta \in \mathcal{M}}{\mathrm{argmin}}\, \mathrm{EMD}(\mathcal{E}, \mathcal{E}'_\theta)$$

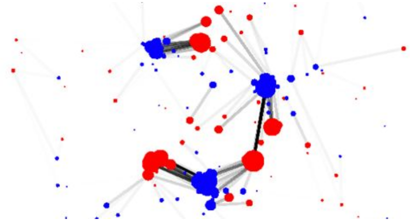
Circle with radius 0.767, center (0.50, 0.36) and a “circle-ness” value of 0.32.

Yes, you *CAN* hear the shape of a jet!

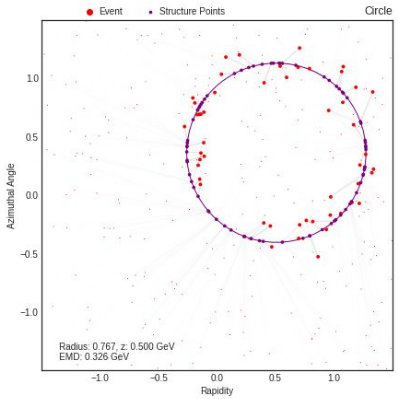
An IAIFI Story



Piecewise-Linear Manifold Approximation with K-Deep Simplices (KDS, 2012.02134)



Well-Defined Metric on Particle Collisions using Energy Mover's Distance (EMD, 2004.04159)



SHAPER: Learning the Shape of Collider Events

$$\mathcal{O}_{\mathcal{M}}(\mathcal{E}) = \min_{\mathcal{E}'_\theta \in \mathcal{M}} \mathrm{EMD}(\mathcal{E}, \mathcal{E}'_\theta)$$

$$\theta = \operatorname*{argmin}_{\mathcal{E}'_\theta \in \mathcal{M}} \mathrm{EMD}(\mathcal{E}, \mathcal{E}'_\theta)$$

Framework for defining and calculating useful observables for collider physics!

Building SHAPER

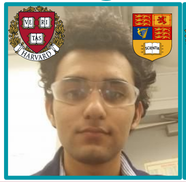
[P. Tankala, A. Tasissa, J. M. Murphy, D. Ba, 2012.02134;
see also F.Dornaika, L.Weng, DOI: 0.1007/s13042-019-01035-z;
see also S. Roweis and L.Saul, DOI: 10.1126/science.290.5500.2323]

Key Component: The Loss function! Step 1: Manifold Learning

$$\mathcal{L}_R(\mathcal{E}, \mathcal{E}') = \min_{\pi_{ij} \geqslant 0} \left[\sum_{i=1}^{M} \sum_{j=1}^{M'} \pi_{ij} \frac{|x_i - x'_j|}{R} \right],$$

$$\text{where } \sum_{i=1}^{M} \pi_{ij} = 1$$

Dogra



Ba

K-Deep Simplices, Dictionary Learning, & Manifold Learning

[P. Komiske, E. Metodiev, J. Thaler, 1902.02346;
see also T. Cai, J. Cheng, K. Craig, N. Craig, 2111.03670;
see also C. Zhang, Y. Cai, G. Lin, C. Shen, 2003.06777;
see also L. Hou, C. Yu, D. Samaras, 1611.05916;
see also M. Arjovsky, S. Chintala, L. Bottou, 1701.07875]

Building SHAPER

Key Component: The Loss function! Step 2: Physical Principles

$$\mathcal{L}_R(\mathcal{E},\mathcal{E}') = \min_{\pi_{ij}\geqslant 0}\left[\sum_{i=1}^{M}\sum_{j=1}^{M'}\pi_{ij}\frac{|x_i - x'_j|}{R}\right] + \left|\sum_{i=1}^{M} z_i - \sum_{j=1}^{M'} z'_j\right|,$$

$$\text{where} \sum_{i=1}^{M}\pi_{ij} \leqslant z'_j, \quad \sum_{j=1}^{M'}\pi_{ij} \leqslant z_i, \quad \sum_{i,j}^{M,M'}\pi_{ij} = \min\left(\sum_{i=1}^{M} z_i, \sum_{j=1}^{M'} z'_j\right)$$

Dogra

Ba

K-Deep Simplices, Dictionary Learning, & Manifold Learning

IRC Safety, Unclustered Radiation, & Wasserstein Geometry

Gambhir

Thaler

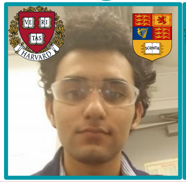
Building SHAPER

Key Component: The Loss function! Step 3: Synthesis

$$\mathcal{L}_R(\mathcal{E},\mathcal{E}') = \min_{\pi_{ij}\geqslant 0}\left[\sum_{i=1}^{M}\sum_{j=1}^{M'}\pi_{ij}\frac{|x_i - x'_j|}{R}\right] + \left|\sum_{i=1}^{M} z_i - \sum_{j=1}^{M'} z'_j\right|,$$

$$\text{where} \quad \sum_{i=1}^{M}\pi_{ij} \leqslant z'_j, \quad \sum_{j=1}^{M'}\pi_{ij} \leqslant z_i, \quad \sum_{i,j}^{M,M'}\pi_{ij} = \min\left(\sum_{i=1}^{M} z_i, \sum_{j=1}^{M'} z'_j\right)$$

Dogra Ba



K-Deep Simplices, Dictionary Learning, & Manifold Learning

+

IRC Safety, Unclustered Radiation, & Wasserstein Geometry

Gambhir Thaler

Building SHAPER

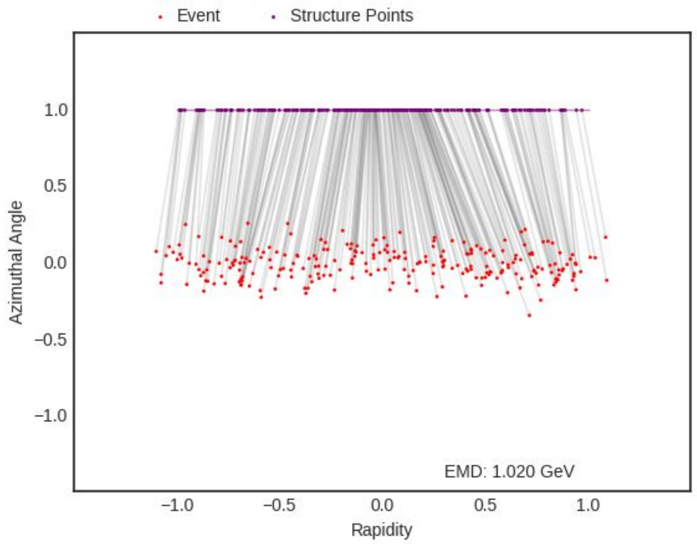
Conversations with:

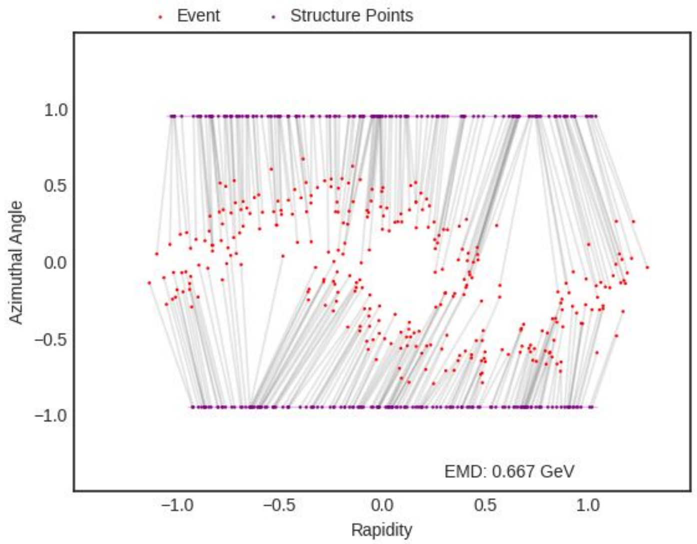
Nolte Williams Kitouni

- Connections to **LHCb Trigger** development
- Potential applications to **future colliders**
- Discussion of **implementation details**

Connections made at the **IAIFI** penthouse and coffee hours!

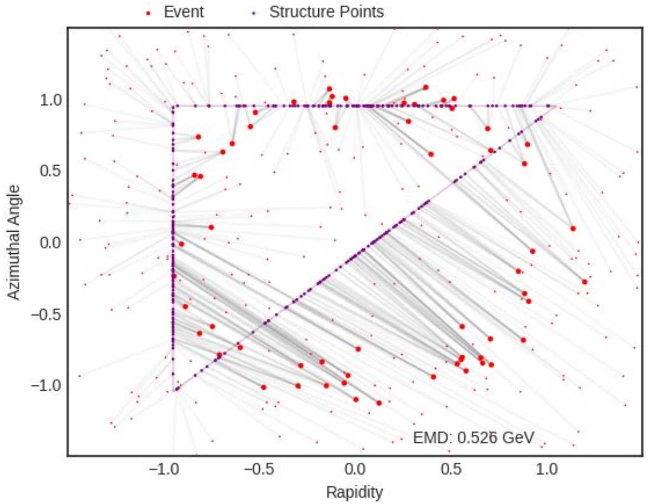
Fun Animations

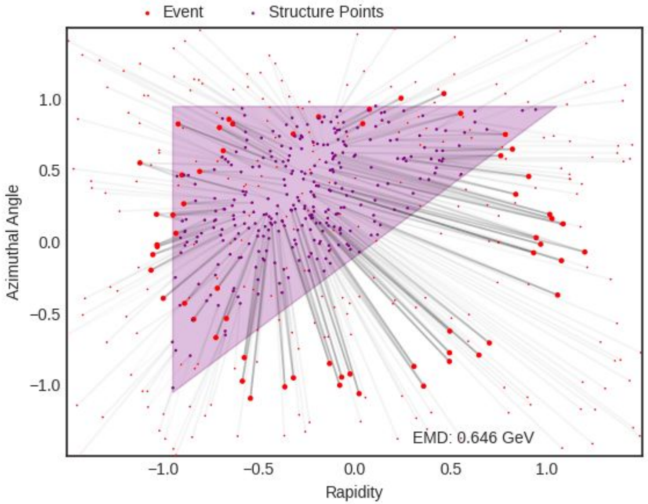




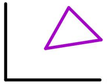
Red: Event Y
Purple: Shape $\mathcal{E}_\theta$ with structure points a_i
Grey: Matrix x_{ij} connecting y's and a_i's

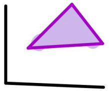
Fun Animations Cont'd





Red: Event Y
Purple: Shape ε_θ with structure points a_i
Grey: Matrix f_{ij} connecting y's and a_i's

Left: ε_θ =  , EMD = 0.245

Right: ε_θ =  EMD = 0.279

New IRC-Safe Observables

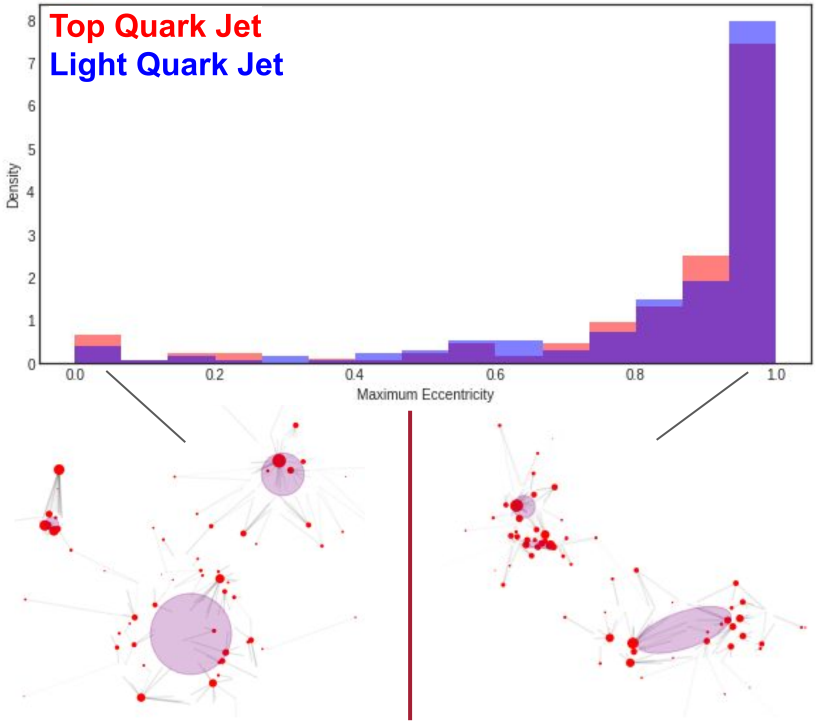
The **SHAPER** framework makes it easy to invent new jet observables!

e.g. ***N*-Ellipsiness+Pileup** as a jet algorithm.

- Learn jet centers
- Dynamic jet radii (no *R* hyperparameter)
- Dynamic **eccentricities** and angles
- Dynamic jet energies
- Uniform Pileup Subtraction
- Learned parameters for discrimination

Can design custom specialized jet algorithms to learn jet substructure!

Max Eccentricity



Low Max Eccentricity (.001)   High Max Eccentricity (.972)

Other Developments: Statistics in Physics

Thaler

Gambhir

Nachman

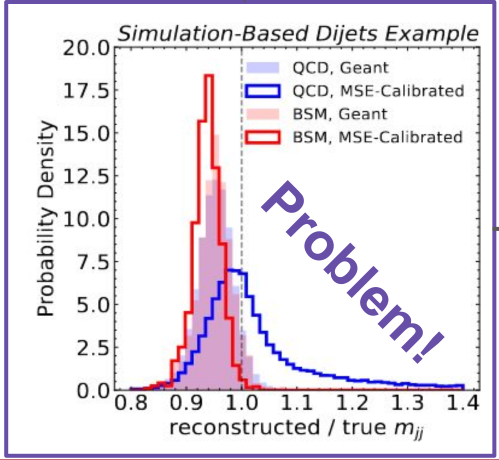
Machine Learning Calibrations (2205.05084)

Bias and Priors in Machine Learning Calibrations for High Energy Physics

Rikab Gambhir,[1,2,*] Benjamin Nachman,[3,4,†] and Jesse Thaler[1,2,‡]

[1] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[2] *The NSF AI Institute for Artificial Intelligence and Fundamental Interactions*
[3] *Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[4] *Berkeley Institute for Data Science, University of California, Berkeley, CA 94720, USA*

Machine learning offers an exciting opportunity to improve the calibration of nearly all reconstructed objects in high-energy physics detectors. However, machine learning approaches often depend on the spectra of examples used during training, an issue known as prior dependence. This is an undesirable property of a calibration, which needs to be applicable in a variety of environments. The purpose of this paper is to explicitly highlight the prior dependence of some machine learning-based calibration strategies. We demonstrate how some recent proposals for both simulation-based and data-based calibrations inherit properties of the sample used for training, which can result in biases for downstream analyses. In the case of simulation-based calibration, we argue that our recently proposed Gaussian Ansatz approach can avoid some of the pitfalls of prior dependence, whereas prior-independent data-based calibration remains an open problem.



Gaussian Ansatz Statistical Framework (2205.03413)

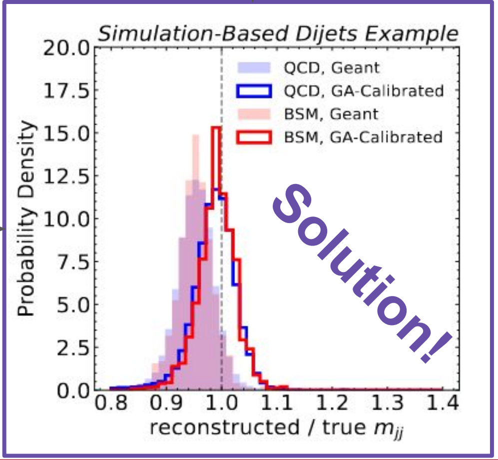
Learning Uncertainties the Frequentist Way: Calibration and Correlation in High Energy Physics

Rikab Gambhir,[1,2,*] Benjamin Nachman,[3,4,†] and Jesse Thaler[1,2,‡]

[1] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[2] *The NSF AI Institute for Artificial Intelligence and Fundamental Interactions*
[3] *Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[4] *Berkeley Institute for Data Science, University of California, Berkeley, CA 94720, USA*

Calibration is a common experimental physics problem, whose goal is to infer the value and uncertainty of an unobservable quantity Z given a measured quantity X. Additionally, one would like to quantify the extent to which X and Z are correlated. In this paper, we present a machine learning framework for performing frequentist maximum likelihood inference with Gaussian uncertainty estimation, which also quantifies the mutual information between the unobservable and measured quantities. This framework uses the Donsker-Varadhan representation of the Kullback-Leibler divergence—parametrized with a novel Gaussian Ansatz—to enable a simultaneous extraction of the maximum likelihood values, uncertainties, and mutual information in a single training. We demonstrate our framework by extracting jet energy corrections and resolution factors from a simulation of the CMS detector at the Large Hadron Collider. By leveraging the high-dimensional feature space inside jets, we improve upon the nominal CMS jet resolution by upwards of 15%.

continues!



Other Developments: MSRP

MIT Summer Research Program

Gaussian Ansatz Statistical Framework (2205.03413)

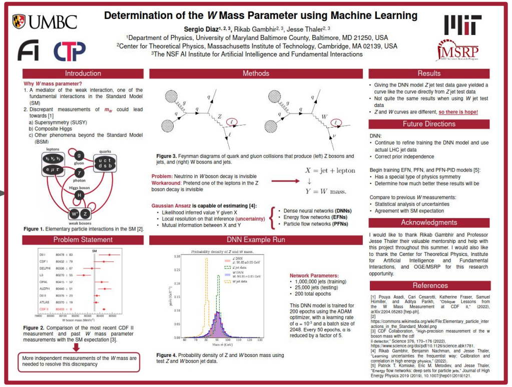
Learning Uncertainties the Frequentist Way:
Calibration and Correlation in High Energy Physics

Rikab Gambhir,[1,2,*] Benjamin Nachman,[3,4,†] and Jesse Thaler[1,2,‡]

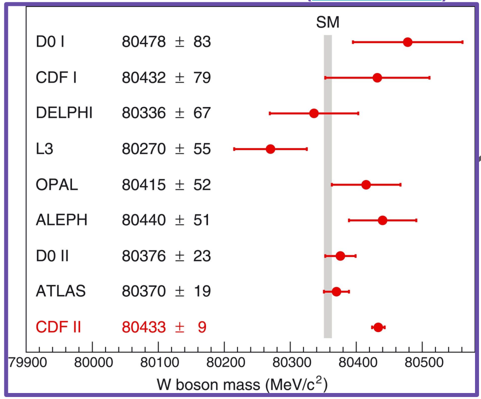
[1] Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2] The NSF AI Institute for Artificial Intelligence and Fundamental Interactions
[3] Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[4] Berkeley Institute for Data Science, University of California, Berkeley, CA 94720, USA

Calibration is a common experimental physics problem, whose goal is to infer the value and uncertainty of an unobservable quantity Z given a measured quantity X. Additionally, one would like to quantify the extent to which X and Z are correlated. In this paper, we present a machine learning framework for performing frequentist maximum likelihood inference with Gaussian uncertainty estimation, which also quantifies the mutual information between the unobservable and measured quantities. This framework uses the Donsker-Varadhan representation of the Kullback-Leibler divergence—parametrized with a novel Gaussian Ansatz—to enable a simultaneous extraction of the maximum likelihood values, uncertainties, and mutual information in a single training. We demonstrate our framework by extracting jet energy corrections and resolution factors from a simulation of the CMS detector at the Large Hadron Collider. By leveraging the high-dimensional feature space inside jets, we improve upon the nominal CMS jet resolution by upwards of 15%.

Sergio Diaz



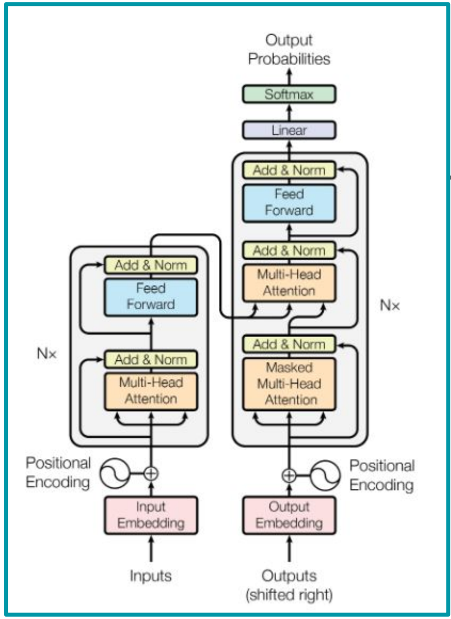
***W* Mass Measurements (DOI: 10.112)**



Exposing students to *both* **particle physics** and **machine learning** to explore new ways to **synthesize** the two!

Other Developments: Summer Students

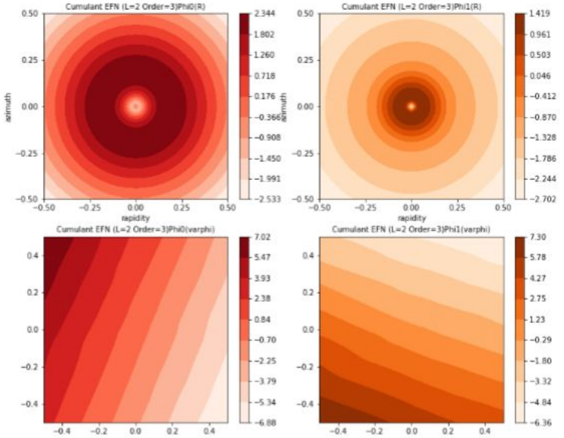
Attention Is All You Need ([1706.03762](https://arxiv.org/abs/1706.03762))



Athis Osathapan

Moment Pooling (WIP)

$$\mathcal{O}_{M_k EFN}\left(\{p_i\}_{i=1}^M\right) = F\left(\left\{\left\langle \prod_{n=1}^{k} \Phi_n^{k_n} \right\rangle\right\}_{\sum k_n \leq k}\right)$$



Translating machine learning language into **physics** language:
What does the attention mechanism look like for a physicist?

Outlook

Exciting research in **physics** and **machine learning** enabled by **IAIFI**!

- Ideas from dictionary and manifold learning to analyze jet data
- Statistical frameworks for precision electroweak measurements
- Efficient machine learning architectures translated to physics language

Made possible by collaborations across fields and institutions!