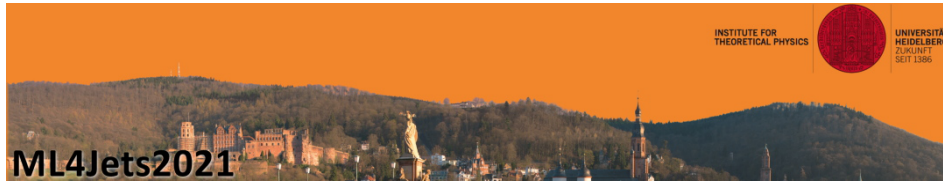


# ML4Jets2021

Tuesday, 6 July 2021 - Thursday, 8 July 2021



## Book of Abstracts



# Contents

Spectral Clustering for Jet Formation . . . . .	1
A $W^\pm$ polarization analyzer from Deep Neural Networks . . . . .	1
Detecting hidden patterns in jet substructure with probabilistic models . . . . .	2
How to GAN Event Unweighting . . . . .	2
Bump Hunting in Latent Space . . . . .	3
Deep learning jet modifications in heavy-ion collisions . . . . .	3
Better latent spaces for better autoencoders . . . . .	4
Explainable AI for ML Jet Taggers . . . . .	4
Higgs tagging with the Lund jet plane . . . . .	5
Measuring QCD Splittings with Invertible Networks . . . . .	6
Uncertainties Associated with GAN Generated Datasets . . . . .	6
Linearized Optimal Transport for Jet Physics . . . . .	7
Amplifying Statistics with Generative Models . . . . .	7
OnlineFlow: Trigger Free Analysis Using Online Learned Generative Models . . . . .	8
Parameter Inference from Event Ensembles and the Top-Quark Mass . . . . .	8
Decoding Photons: Physics in the Latent Space of a BIB-AE Generative Network . . . . .	9
Anomaly Detection in the Copula Space . . . . .	9
Jet Identification in L1 Trigger at HL-LHC based on DNN implementation on FPGA . . . . .	10
The Blessing of Dimensionality: Dimensionality Estimation for Event Clustering . . . . .	10
Supervised Attention for Jet Classification . . . . .	11
Latent Space Refinement for Deep Generative Models . . . . .	11
Uncertainty Aware Learning for High Energy Physics . . . . .	12
Learning Symmetries and Conserved Quantities of Physical Systems . . . . .	13

CaloFlow: Fast and Accurate Generation of Calorimeter Showers with Normalizing Flows	13
Synergies between Quantum Computing and Machine Learning	14
Machine learning based Particle Flow algorithm and application of super-resolution techniques	14
Invertible Networks or Partons to Detector and Back Again	15
Detecting Anomalous jets with Graph Neural Networks	15
Measurement of Muon Energy From Radiative Losses in a Granular Calorimeter	16
Particle Cloud Generation with Message Passing GANs	16
Review of the Dark Machine Anomaly Score Challenge I	17
Jet Metrics and Autoencoders	17
Jet tagging in the Lund plane with graph networks	18
Testing Universality in various Monte Carlo Generators in Deep Learning with Application in Higgs-boson pair searches in 2HDM	18
Compressing PDF sets using Generative Adversarial Networks	19
Fast and Accurate Electromagnetic and Hadronic Showers from Generative Models	20
A new approach to unsupervised learning in jet physics	20
Deep-Learned Event Variables	21
Computing the exact optimal classifier for Ginkgo jets	21
Thoughts on the expressive power and inductive bias of DeepSets and Tree-Based models	22
Tuning the parton shower parameters with the marginal likelihood	22
Emerging techniques for sampling, searching, and summing over the combinatorially large space of shower histories	23
Autoencoders for unsupervised anomaly detection in high energy physics	23
Exploring phase space with Neural Importance Sampling	24
Lightweight Jet Reconstruction as an Object Detection Task	25
SPANet: Generalized Permutationless Set Assignment for Particle Physics using Symmetry Preserving Attention	25
Symmetry Discovery with Deep Learning	26
Jet Topology	26
Conditional invertible neural networks to probe cosmic-ray sources	27
Combine and Conquer: Event Reconstruction with Bayesian Ensemble Neural Networks	27

Matrix Element Calculations on the GPU . . . . .	28
CATHODE part 1: introducing a new model-agnostic search strategy for resonant new physics at the LHC . . . . .	29
CATHODE part 2: robustness and comparison to other methods . . . . .	29
High-dimensional Anomaly Detection with Radiative Return in e+e- Collisions . . . . .	30
Implementation of Jupyter Notebooks into The Reproducible Open Benchmarks for Data Analysis Platform (ROB) . . . . .	30
Jet-based TMD measurements with H1 data, unfolded using machine-learning techniques	31
Angular Conditioning of Deep Generative Models for Fast Simulation of High Granularity Calorimeters . . . . .	32
Safety of Quark/Gluon Jet Classification . . . . .	32
Boosting new physics sensitivity with Variational Autoencoders . . . . .	33
Learning Partially Known Stochastic Dynamics with Empirical PAC Bayes . . . . .	34
Towards a new generation of PDFs using ML . . . . .	34
Using Machine Learning for Heavy-Ion Jet $p_T$ Reconstruction in ALICE . . . . .	34
Shared Data and Algorithms for Deep Learning in Fundamental Physics . . . . .	35
Foundations of a Fast, Data-Driven, Machine-Learned Simulator . . . . .	36
Lorentz Group Equivariant Autoencoder . . . . .	36
Equivariant energy flow networks for jet tagging . . . . .	37
Learning Uncertainties the Frequentist Way: Calibration and Correlation in High Energy Physics . . . . .	37
Combining Neural Network predictions with Hypothesis Testing for discovery in the LHC . . . . .	38
ML in jet physics beyond classification . . . . .	39
AtlFast3: The next generation of fast simulation in ATLAS . . . . .	39
Boosted jet tagging in CMS . . . . .	40
Pileup mitigation in CMS . . . . .	40
Morphology for Jet Classification . . . . .	40
Parametrized classifiers for optimal EFT sensitivity . . . . .	41
Sparse Data Generation with Convolutional VAE . . . . .	42
Via Machinae . . . . .	42
Fast Simulation of Jets with VAEs . . . . .	43

Attention and Dynamic Graph Convolution Neural Network in the context of classifying $t\bar{t}H(bb)$ vs. $t\bar{t}(bb)$ in the semi-leptonic top quark pair decay channel . . . . .	43
Identifying the Quantum Properties of Hadronic Resonances using Machine Learning . . . . .	44
White Box AI for parton shower development . . . . .	45
The information content of quenched jets . . . . .	45
Identifying Heavy-Flavor Jets Using Vectors of Locally Aggregated Descriptors . . . . .	46
SO(3)-equivariant Neural Network for b-tagging . . . . .	46
Pushing the limit of jet tagging with graph neural networks . . . . .	47
Classifier-based Anomalous Jet Tagging . . . . .	47
Particle Convolution for Jets . . . . .	48
Multi-detector geometry modeling and Geant4 Integration . . . . .	48
Introduction to Anomaly Detection Challenge . . . . .	49
Generative Adversarial Networks for Anomaly Detection at the LHC . . . . .	49
ML in Cosmological Simulations . . . . .	50
Generative Networks with Uncertainties . . . . .	50
Invertible Neural Networks beyond Particle Physics . . . . .	50
Recognizing hadronic SUEP at the LHC with Unsupervised Machine Learning . . . . .	51
Super-Resolution for QCD and Top Jets . . . . .	51
Review of the Dark Machine Anomaly Score Challenge II . . . . .	52
Calorimeter Simulation Challenge Proposal . . . . .	52
Bayesian Inference in for Four Tops at the LHC . . . . .	52
ML4Jets2022 . . . . .	53
Closeout and ML4Jets2022 . . . . .	53

**Interpretability, Robustness, and Uncertainties / 2****Spectral Clustering for Jet Formation**

**Author:** Henry Day-Hall<sup>1</sup>

**Co-authors:** Stefano Moretti<sup>2</sup>; Billy Ford<sup>1</sup>; Claire Shepherd-Themistocleous<sup>3</sup>; Srinandan Dasmahapatra<sup>1</sup>

<sup>1</sup> *University of Southampton*

<sup>2</sup> *Science and Technology Facilities Council STFC (GB)*

<sup>3</sup> *Rutherford Appleton Laboratory (GB)*

**Corresponding Authors:** h.a.day-hall@soton.ac.uk, stefano.moretti@cern.ch, claire.shepherd-themistocleous@cern.ch, b.ford@soton.ac.uk, sd@ecs.soton.ac.uk

Machine learning (ML) is pushing through boundaries in computational physics. Jet physics, with its large and detailed dataset, is particularly well suited. In this talk I will discuss the application of an unusual ML technique, Spectral Clustering, to jet formation.

Spectral clustering differs from much of ML as it has no “black-box” elements. Instead, it is based on a simple, elegant algebraic manipulation. This allows us to inspect the way the algorithm is interpreting the data, and apply physical intuition.

Infrared-collinear (IRC) safety is of critical importance to jet physics. IRC safety requires that jets formed are insensitive to collinear splitting and soft emissions. Spectral clustering is shown to be possible to apply in an IRC safe way, and the conditions for this are noted.

Finally, the capacity of spectral clustering to handle different datasets is shown. Its excellent performance, both in terms of multiplicity and mass peaks is demonstrated. In particular we show great performance on two datasets from the extended Higgs sector, alongside the semileptonic top. The reasons for its flexibility are discussed, and potential developments offered.

**Academic Rank:**

PhD student

**Affiliation:**

University of Southampton

**Classification / 3****A  $W^\pm$  polarization analyzer from Deep Neural Networks**

**Authors:** Taegyun Kim<sup>1</sup>; Adam Orion Martin<sup>2</sup>

<sup>1</sup> *University of Notre Dame*

<sup>2</sup> *University of Notre Dame (US)*

**Corresponding Authors:** amarti41@nd.edu, tkim12@nd.edu

We train a Convolutional Neural Network to classify longitudinally and transversely polarized hadronic  $W^\pm$  using the images of boosted  $W^\pm$  jets as input. The images capture angular and energy information from the jet constituents that is faithful to the properties of the original quark/anti-quark  $W^\pm$  decay products without the need for invasive substructure cuts. We find that the difference between

the polarizations is too subtle for the network to be used as an event-by-event tagger. However, given an ensemble of  $W^\pm$  events with unknown polarization, the average network output from that ensemble can be used to extract the longitudinal fraction  $f_L$ . We test the network on Standard Model  $pp \rightarrow W^\pm Z$  events and on  $pp \rightarrow W^\pm Z$  in the presence of dimension-6 operators that perturb the polarization composition.

**Academic Rank:**

PhD student

**Affiliation:**

University of Notre Dame

**Exploring the Latent Structure of Data / 5**

## Detecting hidden patterns in jet substructure with probabilistic models

**Author:** Darius Faroughy<sup>1</sup>

<sup>1</sup> *University of Zurich*

**Corresponding Author:** faroughy@physik.uzh.ch

We build a simple probabilistic model for collider events represented by a pattern of points in a space of high-level observables. The model is based on three assumptions for the point data: the measurements in individual events are discrete, exchangeable, and generated from a mixture of latent distributions, or ‘themes’. The result is a mixed-membership model known as Latent Dirichlet Allocation (LDA), extensively used in natural language processing, biology and many unsupervised machine learning applications. By training on point patterns in the Lund jet plane, we demonstrate that a two-theme LDA model can be used for fully unsupervised event classification. As an example, we show that the LDA classifier can detect a BSM heavy resonance hidden in dijet data.

**Academic Rank:**

**Affiliation:**

Zurich U.

**Simulation and Generative Models / 6**

## How to GAN Event Unweighting

**Authors:** Ramon Winterhalder<sup>1</sup>; Mathias Backes<sup>2</sup>; Anja Butter<sup>None</sup>; Tilman Plehn<sup>None</sup>

<sup>1</sup> *ITP Heidelberg*

<sup>2</sup> *Uni Heidelberg*

**Corresponding Authors:** r.winterhalder@thphys.uni-heidelberg.de, butter@thphys.uni-heidelberg.de, plehn@uni-heidelberg.de

Event generation with neural networks has seen significant progress recently. The big open question is still how such new methods will accelerate LHC simulations to the level required by upcoming LHC runs. We target a known bottleneck of standard simulations and show how their unweighting



procedure can be improved by generative networks. This can, potentially, lead to a very significant gain in simulation speed.

**Academic Rank:**

Master Student

**Affiliation:**

ITP Heidelberg

**Exploring the Latent Structure of Data / 7**

## Bump Hunting in Latent Space

**Authors:** Blaz Bortolato<sup>1</sup>; Barry Dillon<sup>1</sup>; Jernej F. Kamenik<sup>1</sup>; Aleks Smolkovic<sup>2</sup>

<sup>1</sup> *Jozef Stefan Institute*

<sup>2</sup> *Jozef Stefan Institute Ljubljana*

**Corresponding Authors:** jernej.kamenik@ijs.si, aleks.smolkovic@ijs.si, barry.dillon@ijs.si

Unsupervised anomaly detection could be crucial in future analyses searching for rare phenomena in large datasets, as for example collected at the LHC. To this end, we introduce a physics inspired variational autoencoder (VAE) architecture which performs competitively and robustly on the LHC Olympics Machine Learning Challenge datasets. We demonstrate how embedding some physical observables directly into the VAE latent space, while at the same time keeping the classifier manifestly agnostic to them, can help to identify and characterise features in measured spectra as caused by the presence of anomalies in a dataset.

**Academic Rank:**

PhD student

**Affiliation:**

Jozef Stefan Institute, Ljubljana

**Regression, Calibration, and Fast Inference / 8**

## Deep learning jet modifications in heavy-ion collisions

**Authors:** Yilun Du<sup>1</sup>; Daniel Pablos<sup>2</sup>; Konrad Tywoniuk<sup>3</sup>

<sup>1</sup> *University of Bergen*

<sup>2</sup> *INFN, Sezione di Torino*

<sup>3</sup> *University of Bergen (NO)*

**Corresponding Authors:** konrad.tywoniuk@uib.no, daniel.pablosalfonso@gmail.com, yldu@fias.uni-frankfurt.de

Jet interactions in a hot QCD medium created in heavy-ion collisions are conventionally assessed by measuring the modification of the distributions of jet observables with respect to the proton-proton baseline. However, the steeply falling production spectrum introduces a strong bias toward small energy losses that obfuscates a direct interpretation of the impact of medium effects in the measured

jet ensemble. In this talk, we will explore the power of deep learning techniques to tackle this issue on a jet-by-jet basis.

Toward this goal, we employ a convolutional neural network (CNN) to diagnose such modifications from jet images where the training and validation is performed using the hybrid strong/weak coupling model. By analyzing measured jets in heavy-ion collisions, we extract the original jet transverse momentum, i.e., the transverse momentum of an identical jet that did not pass through a medium, in terms of an energy loss ratio. Despite many sources of fluctuations, we achieve good performance and put emphasis on the interpretability of our results. We observe that the angular distribution of soft particles in the jet cone and their relative contribution to the total jet energy contain significant discriminating power, which can be exploited to tailor observables that provide a good estimate of the energy loss ratio.

With a well-predicted energy loss ratio, we study a set of jet observables to estimate their sensitivity to bias effects and reveal their medium modifications when compared to a more equivalent jet population, i.e., a set of jets with similar initial energy. Then, we show how this new technique provides unique access to the initial configuration of jets over the transverse plane of the nuclear collision, both with respect to their production point and initial orientation. Finally, we demonstrate the capability of our new method to locate with unprecedented precision the production point of a dijet pair in the nuclear overlap region, in what constitutes an important step forward towards the long term quest of using jets as tomographic probes of the quark-gluon plasma.

[1] Yi-Lun Du, Daniel Pablos, Konrad Tywoniuk, Deep learning jet modifications in heavy-ion collisions, arXiv:2012.07797 [hep-ph], JHEP. 2021, 206 (2021)

**Academic Rank:**

Postdoctoral researcher

**Affiliation:**

University of Bergen

**Exploring the Latent Structure of Data / 9**

## **Better latent spaces for better autoencoders**

**Author:** Barry Dillon<sup>1</sup>

<sup>1</sup> *University of Heidelberg*

**Corresponding Author:** dillon@thphys.uni-heidelberg.de

Autoencoders as tools behind anomaly searches at the LHC have the structural problem that they only work in one direction, extracting jets with higher complexity but not the other way around. To address this, we derive classifiers from the latent space of (variational) autoencoders, specifically in Gaussian mixture and Dirichlet latent spaces. In particular, the Dirichlet setup solves the problem and improves both the performance and the interpretability of the networks.

**Academic Rank:**

Postdoctoral researcher

**Affiliation:**

University of Heidelberg

**Interpretability, Robustness, and Uncertainties / 10****Explainable AI for ML Jet Taggers**

**Authors:** Christine Angela McLean<sup>1</sup>; Garvita Agarwal<sup>2</sup>; Lauren Meryl Hay<sup>1</sup>; Salvatore Rappoccio<sup>2</sup>; Margaret Morris<sup>3</sup>; Ia Iashvili<sup>2</sup>

<sup>1</sup> *SUNY Buffalo*

<sup>2</sup> *The State University of New York SUNY (US)*

<sup>3</sup> *University at Buffalo*

**Corresponding Authors:** lmhay@buffalo.edu, garvita.agarwal@cern.ch, rappoccio@gmail.com, ch.mclean@cern.ch, iashvili@buffalo.edu, morris35@buffalo.edu

A framework is presented to extract and understand decision-making information from a deep neural network classifier of jet substructure tagging techniques. The general method studied is to provide expert variables that augment inputs (“eXpert AUGmented” variables, or XAUG variables), then apply layerwise relevance propagation (LRP) to networks that have been provided XAUG variables and those that have not. The XAUG variables are concatenated to the classifier’s intermediate input to the final layer. The results show that XAUG variables can be used to interpret classifier behavior, increase discrimination ability when combined with low-level features, and in some cases capture the behavior of the classifier completely. The LRP technique can be used to find relevant information the network is using, and when combined with the XAUG variables, can be used to rank features, allowing one to find a reduced set of features that capture a majority of network performance. These identified XAUG variables can also be added to low-level networks as a guide to improve performance.

**Academic Rank:**

Postdoctoral researcher

**Affiliation:**

University at Buffalo

**Classification / 11****Higgs tagging with the Lund jet plane**

**Authors:** Charanjit Kaur Khosa<sup>1</sup>; Simone Marzani<sup>1</sup>

<sup>1</sup> *University of Genova and INFN Genova*

**Corresponding Authors:** simone.marzani@ge.infn.it, khosacharanjit@gmail.com

In this talk we will present a procedure to separate boosted Higgs bosons decaying into hadrons, from the background due to strong interactions. We employ the Lund jet plane to obtain a theoretically well-motivated representation of the jets of interest and we use the resulting images as the input to a convolutional neural network. In particular, we consider two different decay modes of the Higgs boson, namely into a pair of bottom quarks or into light jets, against the respective backgrounds. The performance of the tagger is compared to what is achieved using a traditional single-variable analysis which exploits a QCD inspired color-singlet tagger, namely the jet color ring observable. Furthermore, we study the dependence of the tagger’s performance on the requirement that the invariant mass of the selected jets should be close to the Higgs mass.

**Academic Rank:**

Postdoctoral researcher

**Affiliation:**

University of Genova and INFN Genova, Italy

**ML-Assisted Measurements and Searches / 12****Measuring QCD Splittings with Invertible Networks**

**Authors:** Anja Butter<sup>None</sup>; Sebastian Bieringer<sup>None</sup>; Stefan Hoche<sup>1</sup>; Stefan Radev<sup>None</sup>; Theo Heigel<sup>2</sup>; Tilman Plehn<sup>None</sup>; Ullrich Köthe<sup>None</sup>

<sup>1</sup> *Fermi National Accelerator Lab. (US)*

<sup>2</sup> *Universität Heidelberg*

**Corresponding Authors:** butter@thphys.uni-heidelberg.de, t.heigel@stud.uni-heidelberg.de, plehn@uni-heidelberg.de, stefan.hoeche@cern.ch

QCD splittings are among the most fundamental theory concepts at the LHC. In this talk, I will show how they can be studied systematically with the help of invertible neural networks. These networks work with sub-jet information to extract fundamental parameters from jet samples. Our approach expands the LEP measurements of QCD Casimirs to a systematic test of QCD properties based on low-level jet observables. Starting with a toy example, I will present the effect of the full shower, hadronization, and detector effects.

**Academic Rank:**

PhD student

**Affiliation:**

Heidelberg University

**Interpretability, Robustness, and Uncertainties / 13****Uncertainties Associated with GAN Generated Datasets**

**Authors:** Alex Roman<sup>1</sup>; Konstantin Matchev<sup>2</sup>; Prasanth Shyamsundar<sup>3</sup>

<sup>1</sup> *University of Florida*

<sup>2</sup> *University of Florida (US)*

<sup>3</sup> *Fermi National Accelerator Laboratory*

**Corresponding Authors:** alexroman@ufl.edu, prasanth@fnal.gov, matchev@phys.ufl.edu

Recently, Generative Adversarial Networks (GANs) trained on samples of traditionally simulated collider events have been proposed as a way of generating larger simulated datasets at a reduced computational cost. In this talk we will present an argument cautioning against the usage of this method to meet the simulation requirements of an experiment, namely that data generated by a GAN cannot statistically be better than the data it was trained on.

We will also state and prove a theorem that limits the ability of GANs to replace traditional simulators in collider physics.

**Affiliation:**

Fermi National Accelerator Laboratory

**Academic Rank:**

Postdoctoral researcher

**New architectures / 14**

## Linearized Optimal Transport for Jet Physics

**Author:** Tianji Cai<sup>1</sup>

**Co-authors:** Junyi Cheng<sup>2</sup>; Nathaniel Craig<sup>1</sup>; Katy Craig<sup>3</sup>; Bernhard Schmitzer<sup>4</sup>; Matthew Thorpe<sup>5</sup>

<sup>1</sup> *Department of Physics, University of California, Santa Barbara*

<sup>2</sup> *University of California, Santa Barbara*

<sup>3</sup> *Department of Mathematics, University of California, Santa Barbara*

<sup>4</sup> *CAMPUS INSTITUTE DATA SCIENCE, UNIVERSITÄT GÖTTINGEN, GÖTTINGEN, GERMANY*

<sup>5</sup> *DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MANCHESTER, MANCHESTER, UK*

**Corresponding Authors:** tianji\_cai@ucsb.edu, schmitzer@cs.uni-goettingen.de, junyi\_cheng@ucsb.edu, matthew.thorpe-2@manchester.ac.uk, ncraig@ucsb.edu, kcraig@math.ucsb.edu

Optimal Transport has been applied to jet physics for the computation of distance between collider events. Here we generalize the Energy Mover's Distance to include both the balanced Wasserstein-2 (W2) distance and the unbalanced Hellinger-Kantorovich (HK) distance. Whereas the W2 distance only allows for mass to be transported, the HK distance allows mass to be transported, created and destroyed, therefore naturally incorporating the total pt difference of the jets. Both distances enjoy a weak Riemannian structure and thus admit linear approximation. Such a linear framework significantly reduces the computational cost and in addition provides a Euclidean embedding amenable to simple machine learning algorithms and visualization techniques downstream. Here we demonstrate the benefit of this linear approach for jet classification and study its behavior in the presence of pileup.

**Affiliation:**

Department of Physics, University of California, Santa Barbara

**Academic Rank:**

PhD student

**Interpretability, Robustness, and Uncertainties / 15**

## Amplifying Statistics with Generative Models

**Authors:** Anja Butter<sup>None</sup>; Sascha Daniel Diefenbacher<sup>1</sup>; Ben Nachman<sup>2</sup>; Gregor Kasieczka<sup>1</sup>; Tilman Plehn<sup>None</sup>; Sebastian Bieringer<sup>None</sup>

<sup>1</sup> *Hamburg University (DE)*

<sup>2</sup> *Lawrence Berkeley National Lab. (US)*

**Corresponding Authors:** gregor.kasieczka@cern.ch, sascha.diefenbacher@desy.de, benjamin.philip.nachman@cern.ch, butter@thphys.uni-heidelberg.de, plehn@uni-heidelberg.de

Monte Carlo simulations are a vital part of modern particle physics. However classical approaches to these simulations require a vast amount of computational resources. Generative Machine Learning models offer a chance to reduce this strain on computing capabilities by allowing us to generate simulated data at a significantly greater speed. The applicability of such generative models has been demonstrated for many problems in particle physics, ranging from event generation to fast calorimeter simulation to many more.

However, one question that needs to be addressed before we can fully utilise generative models is whether a generative model can achieve a more precise description of a given underlying distribution than the data the model was originally trained on. We explore this using a simple toy example and show that a generative model can indeed be used to amplify a data set.

**Affiliation:**

Universität Hamburg

**Academic Rank:**

**Regression, Calibration, and Fast Inference / 16**

## OnlineFlow: Trigger Free Analysis Using Online Learned Generative Models

**Authors:** Anja Butter<sup>None</sup>; Sascha Daniel Diefenbacher<sup>1</sup>; Ben Nachman<sup>2</sup>; David Shih<sup>3</sup>; Gregor Kasieczka<sup>1</sup>; Ramon Winterhalder<sup>4</sup>; Tilman Plehn<sup>None</sup>

<sup>1</sup> *Hamburg University (DE)*

<sup>2</sup> *Lawrence Berkeley National Lab. (US)*

<sup>3</sup> *Rutgers University*

<sup>4</sup> *ITP Heidelberg*

**Corresponding Authors:** benjamin.philip.nachman@cern.ch, r.winterhalder@thphys.uni-heidelberg.de, dshih@physics.rutgers.edu, plehn@uni-heidelberg.de, butter@thphys.uni-heidelberg.de, sascha.diefenbacher@desy.de, gregor.kasieczka@cern.ch

The high collision rates at the Large Hadron Collider (LHC) make it impossible to store every single observed interaction. For this reason, only a small subset that passes so-called triggers —which select potentially interesting events—are saved while the remainder is discarded. This makes it difficult to perform searches in regions that are usually ignored by trigger setups, for example at low energies. However a sufficiently efficient data compression method could help these searches by storing information about more events than can be saved offline.

We investigate the use of a generative machine learning model (specifically a normalizing flow) for the purpose of this compression. The model is trained online on the collision data, essentially encoding the underlying data structure in the network weights. The data generated by the trained model can then be analysed and for example probed for anomalies offline.

We demonstrate this method for a simple bump hunt, showing that we can detect resonances that would have been missed under regular trigger setups.

**Affiliation:**

Universität Hamburg

**Academic Rank:**

**ML-Assisted Measurements and Searches / 17**

## Parameter Inference from Event Ensembles and the Top-Quark Mass

**Author:** Katherine Fraser<sup>1</sup>

**Co-authors:** Matthew Schwartz ; Bryan Ostdiek <sup>1</sup>

<sup>1</sup> *Harvard University*

**Corresponding Authors:** kfraser@g.harvard.edu, bostdiek@gmail.com, schwartz@physics.harvard.edu

Measurements at colliders are often done by fitting data to simulations, which depend on many physical and unphysical parameters. One example is the top-quark mass, where parameters in simulation must be profiled when fitting the top-quark mass parameter. In particular, the dependence of top-quark mass fits on simulation parameters contributes to the error in the best measurements of the top-quark mass. In this talk, I discuss a simple new fitting method to reduce this error, where regression is done directly on ensembles of events. This method is superior at reducing the top-quark mass uncertainty when compared to both traditional histogram fitting methods as well as the modern ML DCTR method. More generally, machine learning from ensembles for parameter estimation has broad potential for collider physics measurements.

**Affiliation:**

Harvard University

**Academic Rank:**

PhD Student

Exploring the Latent Structure of Data / 18

## Decoding Photons: Physics in the Latent Space of a BIB-AE Generative Network

**Authors:** Anatolii Korol<sup>1</sup>; Engin Eren<sup>2</sup>; Erik Buhmann<sup>3</sup>; Frank-Dieter Gaede<sup>4</sup>; Gregor Kasieczka<sup>3</sup>; Katja Kruger<sup>4</sup>; Sascha Daniel Diefenbacher<sup>3</sup>

<sup>1</sup> *Taras Shevchenko National University of Kyiv*

<sup>2</sup> *Deutsches Elektronen-Synchrotron DESY*

<sup>3</sup> *Hamburg University (DE)*

<sup>4</sup> *Deutsches Elektronen-Synchrotron (DE)*

**Corresponding Authors:** katja.kruger@cern.ch, gregor.kasieczka@cern.ch, anatolii.korol@desy.de, frank-dieter.gaede@cern.ch, sascha.diefenbacher@desy.de, engin.eren@desy.de, erik.buhmann@desy.de

Given the increasing data collection capabilities and limited computing resources of future collider experiments, interest in using generative neural networks for the fast simulation of collider events is growing. In our previous study, the Bounded Information Bottleneck Autoencoder (BIB-AE) architecture for generating photon showers in a high-granularity calorimeter showed a high accuracy modeling of various global differential shower distributions. In this work, we investigate how the BIB-AE encodes this physics information in its latent space. Our understanding of this encoding allows us to propose methods to optimize the generation performance further, for example, by altering latent space sampling or by suggesting specific changes to hyperparameters. In particular, we improve the modeling of the shower shape along the particle incident axis.

**Affiliation:**

University of Hamburg

**Academic Rank:**

BSM / 19

## Anomaly Detection in the Copula Space

**Author:** Tommaso Dorigo<sup>1</sup>

<sup>1</sup> *Universita e INFN, Padova (IT)*

**Corresponding Author:** tommaso.dorigo@cern.ch

A unsupervised learning tool that searches for localized, overdense regions of the copula space of a multidimensional feature space is discussed. The algorithm, named RanBox, exists in two versions - one which searches multiple times in random subspaces (typically of 8 to 12 dimensions) of the feature space, and a second one (RanBoxIter) which iteratively adds dimensions to the searched space. Gradient descent is used to localize the multi-dimensional interval which maximizes a suitable test statistic proportional to the significance of the observed data in the box. Applications to UCI datasets from fundamental physics and from fraud detection are discussed.

**Affiliation:**

INFN, Padova

**Academic Rank:**

professor

**Regression, Calibration, and Fast Inference / 20**

## Jet Identification in L1 Trigger at HL-LHC based on DNN implementation on FPGA

**Author:** Andre Sznajder<sup>1</sup>

<sup>1</sup> *UERJ (Brazil)*

**Corresponding Author:** andre.sznajder@cern.ch

We investigate the possibility of using Deep Learning algorithms for jet identification in the L1 trigger at HL-LHC. We perform a survey of architectures (MLP, CNN, Graph Networks) and benchmark their performance and resource consumption on FPGAs using a QKeras+hls4ml compression-aware training procedure. We use the HLS4ML jet dataset to compare the results obtained in this study to previous literature on Fast Machine Learning applications on FPGAs.

**Affiliation:**

UERJ(Brazil)

**Academic Rank:**

Professor

**Exploring the Latent Structure of Data / 21**

## The Blessing of Dimensionality: Dimensionality Estimation for Event Clustering



**Authors:** Erik Buhmann<sup>1</sup>; Gregor Kasieczka<sup>1</sup>; Malte Jacobsen<sup>1</sup>; Tilman Plehn<sup>2</sup>

<sup>1</sup> *Hamburg University (DE)*

<sup>2</sup> *Heidelberg University*

**Corresponding Authors:** erik.buhmann@desy.de, tilman.plehn@cern.ch, malte.jacobsen@desy.de, gregor.kasieczka@cern.ch

Fundamental laws of physics introduce specific topological features in the phase-space of n-body processes in collider events. We introduce a new analysis approach relying on analyzing such global topological properties of the manifold over the distribution of events. One specific property of potential interest is the dimensionality of the phase space. It can, for example, be used for clustering events and discovering anomalies in an unsupervised way.

Focusing on the Drell-Yan process with and without Z-resonance, we show that the dimensionality can be accurately estimated using the minimal neighborhood information. As a resonance reduces the dimensionality by one, we can use this to separate the two processes. Our approach can be extended to more complicated processes and generally has a potentially wide range of applications in particle physics.

**Affiliation:**

University of Hamburg

**Academic Rank:**

**New architectures / 22**

## Supervised Attention for Jet Classification

**Authors:** Eilam Gross<sup>1</sup>; Jonathan Shlomi<sup>1</sup>; Kyle Stuart Cranmer<sup>2</sup>; Sanmay Ganguly<sup>1</sup>

<sup>1</sup> *Weizmann Institute of Science (IL)*

<sup>2</sup> *New York University (US)*

**Corresponding Authors:** eilam.work@gmail.com, jonathan.shlomi@weizmann.ac.il, kyle.cranmer@cern.ch, sanmay.ganguly@cern.ch

Secondary vertex reconstruction is a key intermediate step in building powerful jet classifiers. We use a neural network to perform vertex finding inside jets in order to improve classification performance. This can be thought of as a supervised attention mechanism - directing the classifier towards the relevant information inside the jet. We show supervised attention outperforms an identical network with standard unsupervised attention.

<https://arxiv.org/abs/2008.02831>

**Affiliation:**

Weizmann Institute

**Academic Rank:**

PhD student

**Compression / 24**

## Latent Space Refinement for Deep Generative Models

**Authors:** Ramon Winterhalder<sup>1</sup>; Marco Bellagente<sup>2</sup>; Ben Nachman<sup>3</sup>

<sup>1</sup> *ITP Heidelberg*

<sup>2</sup> *Universität Heidelberg*

<sup>3</sup> *Lawrence Berkeley National Lab. (US)*

**Corresponding Authors:** r.winterhalder@thphys.uni-heidelberg.de, benjamin.philip.nachman@cern.ch, md239@uni-heidelberg.de

Deep generative models are becoming widely used across science and industry for a variety of purposes. A common challenge is achieving a precise implicit or explicit representation of the data probability density. Recent proposals have suggested using classifier weights to refine the learned density of deep generative models. We extend this idea to all types of generative models and show how latent space refinement via iterated generative modeling can circumvent topological obstructions and improve precision. This methodology also applies to cases where the target model is non-differentiable and has many internal latent dimensions which must be marginalized over before refinement. We demonstrate our Latent Space Refinement (LaSeR) protocol on a variety of examples, focusing on the combinations of Normalizing Flows and Generative Adversarial Networks.

**Affiliation:**

ITP Heidelberg

**Academic Rank:**

Postdoctoral researcher

**Interpretability, Robustness, and Uncertainties / 25**

## Uncertainty Aware Learning for High Energy Physics

**Authors:** Aishik Ghosh<sup>1</sup>; Ben Nachman<sup>2</sup>; Daniel Whiteson<sup>1</sup>

<sup>1</sup> *University of California Irvine (US)*

<sup>2</sup> *Lawrence Berkeley National Lab. (US)*

**Corresponding Authors:** aishik.ghosh@cern.ch, daniel.whiteson@cern.ch, benjamin.philip.nachman@cern.ch

Machine learning techniques are becoming an integral component of data analysis in High Energy Physics (HEP). These tools provide a significant improvement in sensitivity over traditional analyses by exploiting subtle patterns in high-dimensional feature spaces. These subtle patterns may not be well-modeled by the simulations used for training machine learning methods, resulting in an enhanced sensitivity to systematic uncertainties.

Contrary to the traditional wisdom of constructing an analysis strategy that is invariant to systematic uncertainties, we study the use of a classifier that is fully aware of uncertainties and their corresponding nuisance parameters. We show that this dependence can actually enhance the sensitivity to parameters of interest. Studies are performed using a synthetic Gaussian dataset as well as a more realistic HEP dataset based on Higgs boson decays to tau leptons. For both cases, we show that the uncertainty aware approach can achieve a better sensitivity than alternative machine learning strategies.

<https://arxiv.org/abs/2105.08742>

**Affiliation:**

University of California, Irvine, Lawrence Berkeley National Laboratory

**Academic Rank:**

Postdoctoral researcher

## Exploring the Latent Structure of Data / 26

### Learning Symmetries and Conserved Quantities of Physical Systems

**Author:** Sven Krippendorf<sup>None</sup>

**Corresponding Author:** sl.krippendorf@googlemail.com

This talk is about how we can use ML to identify symmetries (conserved quantities) of physical systems. I report on three different strategies to find symmetries:

- 1) By examining the embedding a (deep) neural network adapts on a simple supervised task (2003.13679).
- 2) By imposing a modification to Hamiltonian Neural Networks such that a coordinate transformation ensures the emergence of conserved quantities (symmetry control neural networks, 2104.14444).
- 3) By searching for a Lax pair/connection to identify whether a system is integrable (2103.07475), i.e. it has as many conserved quantities as degrees of freedom.

I comment on how strategies 1) and 3) enable us to search for new mathematical structures and how 2) can be used to accelerate simulations.

**Affiliation:**

LMU Munich

**Academic Rank:**

## Simulation and Generative Models / 27

### CaloFlow: Fast and Accurate Generation of Calorimeter Showers with Normalizing Flows

**Authors:** Claudius Krause<sup>1</sup>; David Shih<sup>1</sup>

<sup>1</sup> *Rutgers University*

**Corresponding Authors:** dshih@physics.rutgers.edu, claudius.krause@rutgers.edu

We introduce CaloFlow, a fast detector simulation framework based on normalizing flows. For the first time, we demonstrate that normalizing flows can reproduce high-granularity calorimeter simulations with extremely high fidelity, providing a fresh alternative to computationally expensive GEANT4 simulations, as well as other state-of-the-art fast simulation frameworks based on GANs and VAEs. Besides the usual histograms of physical features and images of calorimeter showers, we introduce a new metric for judging the quality of generative modeling: the performance of a classifier trained to differentiate real from generated images. We show that GAN-generated images can be identified by the classifier with 100% accuracy, while images generated from CaloFlow are able to fool the classifier much of the time. More broadly, normalizing flows offer several advantages compared to other state-of-the-art approaches (GANs and VAEs), including: tractable likelihoods; stable and convergent training; and principled model selection. Normalizing flows also provide a bijective mapping between data and the latent space, which could have other applications beyond simulation, for example, to detector unfolding.

**Affiliation:**

Rutgers University

**Academic Rank:**

Postdoc

**New Horizons / 28**

## **Synergies between Quantum Computing and Machine Learning**

**Author:** Michael Spannowsky<sup>1</sup>

<sup>1</sup> *University of Durham (GB)*

**Corresponding Author:** michael.spannowsky@cern.ch

I will give a very brief (and incomplete) review on quantum machine learning techniques and focus then on novel quantum computing approaches for the task of finding a solution to an optimisation problem. I will then give explicit examples how quantum machine learning techniques can be used for classification tasks and to calculate solutions to nonperturbative problems in quantum field theory.

**Affiliation:**

IPPP Durham

**Academic Rank:**

**Regression, Calibration, and Fast Inference / 29**

## **Machine learning based Particle Flow algorithm and application of super-resolution techniques**

**Author:** Sanmay Ganguly<sup>1</sup>

**Co-authors:** Jonathan Shlomi<sup>1</sup>; Eilam Gross<sup>1</sup>; Marumi Kado<sup>2</sup>; Lorenzo Santi<sup>2</sup>; Francesco Armando Di Bello<sup>2</sup>

<sup>1</sup> *Weizmann Institute of Science (IL)*

<sup>2</sup> *Sapienza Universita e INFN, Roma I (IT)*

**Corresponding Authors:** jonathan.shlomi@weizmann.ac.il, francesco.armando.di.bello@cern.ch, eilam.work@gmail.com, l.santi@cern.ch, sanmay.ganguly@cern.ch, marumi.kado@cern.ch

In High Energy Physics experiments Particle Flow (PFlow) algorithms are designed to provide an optimal reconstruction of the nature and kinematic properties of the particles produced within the detector acceptance during collisions. At the heart of PFlow algorithms is the ability to distinguish the calorimeter energy deposits of neutral particles from those of charged particles, using the complementary measurements of charged particle tracking devices, to provide a superior measurement of the particle content and kinematics. In this presentation, a computer vision approach to this fundamental aspect of PFlow algorithms, based on calorimeter images, is proposed. A comparative study of the state of the art deep learning techniques is performed. A significantly improved reconstruction of the neutral particle calorimeter energy deposits is obtained in a context of large overlaps with the deposits from charged particles. Calorimeter images with augmented finer granularity are also obtained using super-resolution techniques.

**Affiliation:**

Weizmann Institute of Science

**Academic Rank:**

Postdoctoral researcher

**ML-Assisted Measurements and Searches / 30****Invertible Networks or Partons to Detector and Back Again****Authors:** Anja Butter<sup>None</sup>; Gregor Kasieczka<sup>1</sup>; Marco Bellagente<sup>2</sup>; Ramon Winterhalder<sup>3</sup>; Tilman Plehn<sup>None</sup>; Ullrich Köthe<sup>None</sup><sup>1</sup> *Hamburg University (DE)*<sup>2</sup> *Universität Heidelberg*<sup>3</sup> *ITP Heidelberg***Corresponding Authors:** gregor.kasieczka@cern.ch, butter@thphys.uni-heidelberg.de, md239@uni-heidelberg.de, plehn@uni-heidelberg.de, r.winterhalder@thphys.uni-heidelberg.de

For simulations where the forward and the inverse directions have a physics meaning, invertible neural networks are especially useful. A conditional INN can invert a detector simulation in terms of high-level observables, specifically for ZW production at the LHC. It allows for a per-event statistical interpretation. Next, we allow for a variable number of QCD jets. We unfold detector effects and QCD radiation to a pre-defined hard process, again with a per-event probabilistic interpretation over parton-level phase space.

**Affiliation:**

ITP Heidelberg

**Academic Rank:**

Postdoc

**BSM / 31****Detecting Anomalous jets with Graph Neural Networks****Authors:** Akanksha Bhardwaj<sup>1</sup>; Christoph Englert<sup>1</sup>; Michael Spannowsky<sup>2</sup>; Oliver Atkinson<sup>1</sup>; Vishal Singh Ngairangbam<sup>3</sup><sup>1</sup> *University of Glasgow*<sup>2</sup> *University of Durham (GB)*<sup>3</sup> *Physical Research Laboratory***Corresponding Authors:** o.atkinson.1@research.gla.ac.uk, michael.spannowsky@cern.ch, bhardwajakanksha22@gmail.com, vishalngairangbam@gmail.com, christoph.englert@googlemail.com

We devise an autoencoder based strategy to facilitate anomaly detection for boosted jets, employing Graph Neural Networks (GNNs) to do so. To overcome known limitations of GNN autoencoders, we design a symmetric decoder capable of simultaneously reconstructing edge features and node features. Focusing on latent space based discriminators, we find that such setups provide a promising avenue to isolate new physics and competing SM signatures from sensitivity-limiting QCD jet contributions. We demonstrate the flexibility and broad applicability of this approach using examples of W bosons, top quarks, and exotic hadronically-decaying exotic scalar boson

**Affiliation:**

Physical Research Laboratory, Ahmedabad

**Academic Rank:**

PhD student

**Regression, Calibration, and Fast Inference / 32****Measurement of Muon Energy From Radiative Losses in a Granular Calorimeter**

**Authors:** Filippo Chiandotto<sup>None</sup>; Giles Chatham Strong<sup>1</sup>; Jan Kieseler<sup>2</sup>; Lukas Layer<sup>1</sup>; Tommaso Dorigo<sup>1</sup>

<sup>1</sup> *Universita e INFN, Padova (IT)*

<sup>2</sup> *CERN*

**Corresponding Authors:** jan.kieseler@cern.ch, llayer@cern.ch, giles.chatham.strong@cern.ch, tommaso.dorigo@cern.ch

The performance demands of future particle-physics experiments investigating the high-energy frontier pose a number of new challenges, forcing us to find new solutions for the detection, identification, and measurement of final-state particles in subnuclear collisions. One such challenge is the precise measurement of muon momenta at very high energy, where the curvature provided by conceivable magnetic fields in realistic detectors proves insufficient to achieve the desired resolution. In this work we show the feasibility of an entirely new avenue for the measurement of the energy of muons based on their radiative losses in a dense, finely segmented calorimeter.

Using a task-specific 3D convolutional neural network, the raw energy deposits in the calorimeter cells may be used as inputs to regress to the energy of the originating muon. We demonstrate that this approach provides superior resolution for high energy muons. Additionally, due to the differing energy dependence, we show that the regression is entirely complementary to traditional tracker-based measurements, which degrades with energy, together allowing one to achieve good resolution across the energy spectrum.

**Affiliation:**

CERN, CMS, Padova

**Academic Rank:**

1 PhD student, 1 Postdoctoral researcher (speaker), 1 Professor, 1 research staff, 1 masters student

**Simulation and Generative Models / 33****Particle Cloud Generation with Message Passing GANs**

**Authors:** Raghav Kansal<sup>1</sup>; Javier Mauricio Duarte<sup>1</sup>; Hao Su<sup>None</sup>; Maurizio Pierini<sup>2</sup>; Mary Touranakou<sup>3</sup>; Breno Orzari<sup>4</sup>; Thiago Tomei Fernandez<sup>4</sup>; Jean-Roch Vlimant<sup>5</sup>; Dimitrios Gunopoulos<sup>None</sup>

<sup>1</sup> *Univ. of California San Diego (US)*

<sup>2</sup> *CERN*

<sup>3</sup> *National and Kapodistrian University of Athens (GR)*

<sup>4</sup> *UNESP - Universidade Estadual Paulista (BR)*

<sup>5</sup> *California Institute of Technology (US)*

**Corresponding Authors:** mary.touranakou@cern.ch, raghav.kansal@cern.ch, javier.mauricio.duarte@cern.ch, jean-roch.vlimant@cern.ch, thiago.tomei@cern.ch, breno.orzari@cern.ch, maurizio.pierini@cern.ch

There has been significant development recently in generative models for accelerating LHC simulations. Work on simulating jets has primarily used image-based representations, which tend to be sparse and of limited resolution. We advocate for the more natural ‘particle cloud’ representation of jets, i.e. as a set of particles in momentum space, and discuss four physics- and computer-vision-inspired metrics: (1) the 1-Wasserstein distance between high- and low-level feature distributions; (2) a new Fréchet ParticleNet Distance; (3) the coverage; and (4) the minimum matching distance as means of quantitatively and holistically evaluating generated particle clouds. We then present our new message-passing generative adversarial network (MPGAN), which has excellent performance on gluon, top quark, and lighter quark jets on all metrics, evaluated against real samples via bootstrapping as well as existing point cloud GANs, and shows promise for use in HEP.

**Affiliation:**

UC San Diego

**Academic Rank:**

PhD student

BSM / 34

## Review of the Dark Machine Anomaly Score Challenge I

**Authors:** Bryan Ostdiek<sup>1</sup>; Joe Davies<sup>None</sup>

<sup>1</sup> *Harvard University*

**Corresponding Authors:** j.m.m.davies@qmul.ac.uk, bostdiek@gmail.com

We describe the outcome of a data challenge conducted as part of the Dark Machines Initiative and the Les Houches 2019 workshop on Physics at TeV colliders. The challenge aims at detecting signals of new physics at the LHC using unsupervised learning algorithms. We define and describe a large benchmark dataset, consisting of > 1 Billion simulated LHC events. We then review a wide range of anomaly detection algorithms and measure their performance on the data challenge. We then assess the best-performing models on a still blinded dataset. Similarities between the best-performing models are observed and discussed.

**Affiliation:**

Harvard University

**Academic Rank:**

Postdoc

BSM / 35

## Jet Metrics and Autoencoders

**Authors:** Katherine Fraser<sup>1</sup>; Samuel Homiller<sup>2</sup>; Rashmish Mishra<sup>1</sup>; Bryan Ostdiek<sup>1</sup>; Matthew Schwartz<sup>None</sup>

<sup>1</sup> *Harvard University*

<sup>2</sup> *Harvard*

**Corresponding Authors:** kfraser@g.harvard.edu, schwartz@physics.harvard.edu, bostdiek@gmail.com, rashmish-mishra@fas.harvard.edu, shomiller@g.harvard.edu

The Energy Movers Distance was recently proposed as an advantageous metric to distinguish certain types of signals at the LHC. We explore generalizations of this distance to multiple families of signals and find similar performance anomaly detection through variational autoencoders. We investigate this connection by exploring the correlation of event distances with distances in the latent space of the autoencoder.

**Affiliation:**

Harvard University

**Academic Rank:**

Postdc

**Classification / 36**

## Jet tagging in the Lund plane with graph networks

**Author:** Frederic Alexandre Dreyer<sup>1</sup>

<sup>1</sup> *University of Oxford*

**Corresponding Author:** frederic.dreyer@physics.ox.ac.uk

The identification of boosted heavy particles such as top quarks or vector bosons is one of the key problems arising in experimental studies at the Large Hadron Collider. In this article, we introduce LundNet, a novel jet tagging method which relies on graph neural networks and an efficient description of the radiation patterns within a jet to optimally disentangle signatures of boosted objects from background events. We apply this framework to a number of different benchmarks, showing significantly improved performance for top tagging compared to existing state-of-the-art algorithms. We study the robustness of the LundNet taggers to non-perturbative and detector effects, and show how kinematic cuts in the Lund plane can mitigate overfitting of the neural network to model-dependent contributions. Finally, we consider the computational complexity of this method and its scaling as a function of kinematic Lund plane cuts, showing an order of magnitude improvement in speed over previous graph-based taggers.

**Affiliation:**

Oxford

**Academic Rank:**

**Classification / 37**

## Testing Universality in various Monte Carlo Generators in Deep Learning with Application in Higgs-boson pair searches in 2HDM

**Authors:** Kingman Cheung<sup>1</sup>; Yi-Lun Chung<sup>1</sup>; Shih-Chieh Hsu<sup>2</sup>; Ben Nachman<sup>3</sup>



<sup>1</sup> *National Tsing Hua University (TW)*

<sup>2</sup> *University of Washington Seattle (US)*

<sup>3</sup> *Lawrence Berkeley National Lab. (US)*

**Corresponding Authors:** benjamin.philip.nachman@cern.ch, s107022801@m107.nthu.edu.tw, schsu@uw.edu, kingman.cheung@cern.ch

It is widely known that predictions for jet substructure features vary significantly between Monte Carlo generators. This is especially true for the output of deep neural networks (NN) trained with high-dimensional feature spaces to tag the origin of a jet. However, even though the spectra of a given NN varies between generators, it could be that the function learned by different generators is the same. We investigate the universality of jet substructure information by training a NN with a variety of generators and testing these NNs on the same generator. By fixing the testing generator, we can see if the NNs have learned to use the same information, even if the extent to which that information is expressed varied between training datasets. Our target physics process is boosted Higgs bosons and we explore the implications of universality on uncertainties for searches for new particles at the Large Hadron Collider and beyond.

**Affiliation:**

National Tsing Hua University

**Academic Rank:**

PhD student

**Compression / 38**

## Compressing PDF sets using Generative Adversarial Networks

**Authors:** Tanjona Radonirina Rabemananjara<sup>1</sup>; Stefano Carrazza<sup>2</sup>; Juan M. Cruz Martínez<sup>3</sup>

<sup>1</sup> *INFN - National Institute for Nuclear Physics*

<sup>2</sup> *CERN*

<sup>3</sup> *University of Milan*

**Corresponding Authors:** juacumar@gmail.com, tanjona.rabemananjara@mi.infn.it, stefano.carrazza@cern.ch

Data compression plays a major role in the field of Machine Learning and recent works based on generative models such as Generative Adversarial Networks (GANs) have shown that deep-learning-based compression can outperform state-of-the-art classical compression methodologies. Such techniques can be adapted and applied to various areas in high energy physics, in particular to the study of the Parton Distribution Functions (PDFs) in which large Monte Carlo replicas samples are required in order to get accurate results. In this talk, we present a compression algorithm for parton densities in which the statistics of a given input PDF set is further enhanced by the generation of synthetic replicas using GAN prior to compression. This results in a compression methodology that is able to provide a compressed set with smaller number of replicas and a more adequate representation of the original probability distribution.

**Affiliation:**

University of Milan

**Academic Rank:**

PhD Student

**Simulation and Generative Models / 39****Fast and Accurate Electromagnetic and Hadronic Showers from Generative Models**

**Authors:** Anatolii Korol<sup>1</sup>; Daniel Hundhausen<sup>2</sup>; Engin Eren<sup>2</sup>; Erik Buhmann<sup>3</sup>; Frank-Dieter Gaede<sup>4</sup>; Gregor Kasieczka<sup>3</sup>; Katja Kruger<sup>4</sup>; Lennart Rustige<sup>5</sup>; Peter McKeown<sup>2</sup>; Sascha Daniel Diefenbacher<sup>3</sup>; William Korcari<sup>3</sup>

<sup>1</sup> *Taras Shevchenko National University of Kyiv*

<sup>2</sup> *Deutsches Elektronen-Synchrotron DESY*

<sup>3</sup> *Hamburg University (DE)*

<sup>4</sup> *Deutsches Elektronen-Synchrotron (DE)*

<sup>5</sup> *CDCS / DESY*

**Corresponding Authors:** anatolii.korol@desy.de, peter.mckeown@desy.de, lennart.rustige@desy.de, gregor.kasieczka@cern.ch, daniel.hundhausen@desy.de, katja.kruger@cern.ch, frank-dieter.gaede@cern.ch, engin.eren@desy.de, sascha.diefenbacher@desy.de, erik.buhmann@desy.de, william.korcari@cern.ch

Generative machine learning models are a promising way to efficiently amplify classical Monte Carlo generators' statistics for event simulation and generation in particle physics. The high computational cost of the simulation and the expected increase in data in the high-precision era of the LHC and at future colliders indicate that we urgently need such fast surrogate simulators. We present a status update on simulating particle showers in high granularity calorimeters for future colliders. Building on prior work using Generative Adversarial Networks (GANs), Wasserstein-GANs, and the information-theoretically motivated Bounded Information Bottleneck Autoencoder (BIB-AE), we achieve further improvements of the fidelity of generated photon showers. The key to this improvement is a detailed understanding and optimization of the latent space. The richer structure of hadronic showers compared to electromagnetic ones makes their precise modeling an important yet challenging problem. We present initial progress towards accurately simulating the core of hadronic showers in a highly granular scintillator calorimeter

**Affiliation:**

DESY, University of Hamburg

**Academic Rank:****New architectures / 40****A new approach to unsupervised learning in jet physics**

**Authors:** Peter Rangi Sorrenson<sup>1</sup>; Barry Dillon<sup>2</sup>; Tilman Plehn<sup>3</sup>; Hans Olschläger<sup>2</sup>; Lorenz Vogel<sup>2</sup>

<sup>1</sup> *Universität Heidelberg*

<sup>2</sup> *University of Heidelberg*

<sup>3</sup> *Heidelberg University*

**Corresponding Authors:** tilman.plehn@cern.ch, wq173@uni-heidelberg.de, dillon@thphys.uni-heidelberg.de, hans@olischlaeger.com, mail@lorenz-vogel.de

TBC

**Affiliation:**

University of Heidelberg

**Academic Rank:**

PhD student

**Interpretability, Robustness, and Uncertainties / 41****Deep-Learned Event Variables****Author:** Doojin Kim<sup>1</sup><sup>1</sup> *Texas A & M University (US)***Corresponding Author:** doojin.kim@cern.ch

The choice of optimal event variables is crucial for achieving the maximal sensitivity of experimental analyses, and suitable kinematic variables for many well-motivated event topologies have been developed in collider physics. Here we propose a deep-learning-based algorithm to design good event variables that are sensitive to a wide range of the unknown model parameter values. We demonstrate that the neural networks trained with our algorithm on some simple event topologies are able to reproduce standard variables like invariant mass, transverse mass, and stransverse mass. These simple exercises can address two issues: 1) what have machines learned (explainability)? and 2) are human-engineered features best (optimality)? The method is automatable, completely general, and can be used to derive sensitive, previously unknown, event variables for other, more complex event topologies.

**Affiliation:**

Texas A&amp;M University

**Academic Rank:**

Postdoctoral researcher

**ML-Assisted Measurements and Searches / 42****Computing the exact optimal classifier for Ginkgo jets****Authors:** Lauren Greenspan<sup>1</sup>; Matthew Drnevich<sup>2</sup>; Sebastian Macaluso<sup>3</sup>; Kyle Stuart Cranmer<sup>2</sup><sup>1</sup> *NYU*<sup>2</sup> *New York University (US)*<sup>3</sup> *New York University***Corresponding Authors:** sm4511@nyu.edu, matthew.drnevich@cern.ch, kyle.cranmer@cern.ch, greenspan.lauren@gmail.com

In the last several years, the ML4Jets community has worked to improve performance for jet tagging and performed a number of comparisons of different architectures for jet tagging and other tasks. We have seen that combining multiple classifiers together into a meta-tagger or an ensemble improves performance. But is there still room for improvement? In other words, are we approaching the performance of the optimal tagger? Formally, the optimal classifier is defined by a likelihood ratio (Neyman-Pearson lemma), but the likelihood for the observed jet is typically intractable as it involves marginalizing over the enormous number of showering histories. Additionally, the likelihood for a particular shower is, in general, not easily accessible. We consider new datasets with signal and background generated with the Ginkgo model and use the cluster trellis to exactly compute the marginal likelihood under each hypothesis in order to calculate the exact optimal likelihood ratio. As a result, we can compare the performance of ML-based taggers to this optimal classifier.

**Affiliation:**

NYU

**Academic Rank:****Interpretability, Robustness, and Uncertainties / 43**

## Thoughts on the expressive power and inductive bias of DeepSets and Tree-Based models

**Authors:** Kyle Stuart Cranmer<sup>1</sup>; Sebastian Macaluso<sup>2</sup>; Lauren Greenspan<sup>2</sup><sup>1</sup> *New York University (US)*<sup>2</sup> *New York University***Corresponding Authors:** kyle.cranmer@cern.ch, greenspan.lauren@gmail.com, sm4511@nyu.edu

Nearly five years ago we introduced tree-based recursive NN models for jet physics, which intuitively reflected the sequence of 1-to-2 splittings found in a parton shower. Subsequently, tree-based models like JUNIPR were developed as (probabilistic) generative models that could be used for classification and reweighting. One result that somewhat undermined the narrative of the connection between the inductive bias of the architectures and the underlying physics for this class of tree-based models was that they continued to perform well even if the jet algorithm that was used did not reflect the underlying physics of the parton shower (eg. anti-kT, a simple pT ordering, or a 2d-printer). Later, even simpler models based on DeepSets were introduced that focused on permutation invariance and performed without reference to the underlying showering picture at all. In this talk I'd like to revisit these two forms of inductive bias for models for jets from a new perspective. In particular, I will discuss the expressive power of these network architectures, their connection to the jet clustering algorithm, and make some predictions for experiments that will be conducted over the coming months.

**Affiliation:**

New York University

**Academic Rank:****ML-Assisted Measurements and Searches / 44**

## Tuning the parton shower parameters with the marginal likelihood

**Authors:** Matthew Drnevich<sup>1</sup>; Kyle Stuart Cranmer<sup>1</sup>; Sebastian Macaluso<sup>2</sup><sup>1</sup> *New York University (US)*<sup>2</sup> *New York University***Corresponding Authors:** kyle.cranmer@cern.ch, matthew.drnevich@cern.ch, sm4511@nyu.edu

Tuning parton shower models to data is an important task for HEP experiments. We are performing exploratory research for what tuning the parton shower might look like if the parton shower were described by a generative model with a tractable likelihood, which might be implemented with a hybrid of theoretically-motivated components or generic neural network components. For this work we consider the Ginkgo model, which is a simplified parton shower with 1-to-2 splittings and a tractable likelihood that has been designed to facilitate this research. While the parton shower is traditionally tuned by matching one dimensional projections for various observables, ideally we would tune the it with a maximum likelihood fit. The challenge is that the likelihood for the data

given the shower parameters must marginalize over the  $(2N-3)!!$  possible showering histories, where  $N$  is the number of jet constituents. We demonstrate that with the hierarchical cluster trellis we can exactly marginalize over this enormous space of showering histories and fit the parameters of the Ginkgo model.

**Affiliation:**

New York University

**Academic Rank:**

PhD student

**ML-Assisted Measurements and Searches / 45**

## Emerging techniques for sampling, searching, and summing over the combinatorially large space of shower histories

**Authors:** Sebastian Macaluso<sup>1</sup>; Johann Brehmer<sup>2</sup>; Duccio Pappadopulo<sup>3</sup>; Kyle Stuart Cranmer<sup>4</sup>; Matthew Drnevich<sup>4</sup>

<sup>1</sup> *New York University*

<sup>2</sup> *NYU*

<sup>3</sup> *EPFL*

<sup>4</sup> *New York University (US)*

**Corresponding Authors:** johann.brehmer@nyu.edu, matthew.drnevich@cern.ch, kyle.cranmer@cern.ch, pap-pakus@gmail.com, sm4511@nyu.edu

A central challenge in jet physics is that the evolution of the jet is an unobserved, latent process. In a semi-classical parton shower, this corresponds to a sequence of 1-to-2 splittings that form a tree-like showering history. Framing jet physics in probabilistic terms is attractive as it provides a principled framework to think about tasks as diverse as clustering, classification, parton shower tuning, matrix element —parton shower matching, and event generation of jets in complex, signal-like regions of phase space. Unfortunately, this usually involves either marginalizing (summing) or maximizing (searching) over the enormous space of clustering histories, which is typically intractable. We review three recently published works that address these challenges by building on techniques from statistics, machine learning, and combinatorial optimization. Each of these works are enabled by Ginkgo, a simplified, generative model for jets, designed to facilitate research in this area. We show how probabilistic programming can be used to efficiently sample the showering process, how a novel trellis algorithm can be used to efficiently marginalize over the enormous number of showering histories for the same observed particles, and how dynamic programming, A\* search, and reinforcement learning can be used to find the maximum likelihood clustering in this enormous search space. (edited)

**Affiliation:**

New York University

**Academic Rank:**

postdoc

**BSM / 46**

## Autoencoders for unsupervised anomaly detection in high energy physics

**Authors:** Thorben Finke<sup>None</sup>; Michael Krämer<sup>1</sup>; Alessandro Morandini<sup>1</sup>; Alexander Mück<sup>1</sup>; Ivan Oleksiyuk<sup>1</sup>

<sup>1</sup> RWTH Aachen

**Corresponding Author:** finke@physik.rwth-aachen.de

Autoencoders have been introduced in high energy physics as a promising tool for model-independent new physics searches. As a benchmark scenario, we study the tagging of top jet images in a background of QCD jet images. Although we reproduce the positive results from the literature, we show that the standard autoencoder setup cannot be considered as a model-independent anomaly tagger by inverting the task: the autoencoder fails to tag QCD jets if it is trained on top jets. We suggest improved performance measures for the task of model-independent anomaly detection. We also improve the capability of the autoencoder to learn non-trivial features of the jet images, such that it is able to achieve both top jet tagging and QCD jet tagging with the same setup. However, we want to stress that a truly model-independent and powerful autoencoder-based unsupervised jet tagger still needs to be developed.

**Affiliation:**

RWTH Aachen

**Academic Rank:**

**Compression / 47**

## Exploring phase space with Neural Importance Sampling

**Author:** Timo Janßen<sup>1</sup>

**Co-authors:** Enrico Bothmann<sup>2</sup>; Max Knobbe<sup>1</sup>; Steffen Schumann<sup>3</sup>; Tobias Schmale

<sup>1</sup> Georg-August-Universität Göttingen

<sup>2</sup> University of Göttingen

<sup>3</sup> Georg-August-Universität Göttingen

**Corresponding Authors:** steffen.schumann@phys.uni-goettingen.de, max.knobbe@gmx.net, timo.janssen@theorie.physik.uni-goettingen.de, enrico.bothmann@uni-goettingen.de

Due to the expected increase in LHC data from the HL upgrade it is important to work on the efficiency of MC Event Generators in order to make theoretical predictions with the necessary precision accessible. One part of the calculation that could benefit from improvements is the generation of unweighted parton-level events. While adaptive multi-channel importance sampling combined with the Vegas algorithm is a very effective method for a wide range of scattering processes, it can become inefficient for challenging examples. Normalizing Flows are a recent machine learning development based on neural networks that provide trainable bijective mappings. We propose to use Normalizing Flows as a direct replacement for Vegas. The method guarantees full phase space coverage and the exact reproduction of the desired target distribution. We study the performance of the algorithm for a few representative examples, including top-quark pair production and gluon scattering into three- and four-gluon final states. We show that our method is able to achieve higher sampling performance than the traditional method for the simpler examples. Furthermore, we discuss the computational challenges and propose possible improvements that could boost the performance of the method also for more complex examples.

**Affiliation:**

Georg-August-Universität Göttingen

**Academic Rank:**

PhD student

**Regression, Calibration, and Fast Inference / 48**

## Lightweight Jet Reconstruction as an Object Detection Task

**Author:** Adrian Alan Pol<sup>1</sup>

**Co-authors:** Thea Aarrestad<sup>1</sup>; Katya Govorkova<sup>1</sup>; Roi Halily<sup>2</sup>; Vladimir Loncar<sup>1</sup>; Anat Klempner<sup>2</sup>; Tal Kopetz<sup>2</sup>; Jennifer Ngadiuba<sup>3</sup>; Maurizio Pierini<sup>1</sup>; Olya Sirkin<sup>2</sup>; Sioni Paris Summers<sup>1</sup>

<sup>1</sup> CERN

<sup>2</sup> CEVA

<sup>3</sup> FNAL

**Corresponding Authors:** thea.aarrestad@cern.ch, ekaterina.govorkova@cern.ch, jennifer.ngadiuba@cern.ch, adrian.pol@cern.ch, sioni.paris.summers@cern.ch, maurizio.pierini@cern.ch, vladimir.loncar@cern.ch

We apply object detection techniques based on convolutional blocks to jet reconstruction and identification at the CERN Large Hadron Collider. We use particles reconstructed through a Particle Flow algorithm to represent each event as an image composed of a calorimeter and tracker cells as input and a Single Shot Detection network, called PFJet-SSD. The network performs simultaneous localization, classification and auxiliary regression tasks to measure jet features. We investigate Ternary Weight Networks with weights quantized to  $\{-1, 0, 1\}$  set, times a layer- and channel-dependent scaling factors for reducing memory and latency constraints. We show that the quantized version of the network closely matches the performance of its full-precision equivalent while both outperform the physics baseline. Finally, we report the inference latency on Nvidia Tesla T4.

**Affiliation:**

CERN

**Academic Rank:**

**New architectures / 49**

## SPANet: Generalized Permutationless Set Assignment for Particle Physics using Symmetry Preserving Attention

**Authors:** Alex Shmakov<sup>1</sup>; Michael James Fenton<sup>2</sup>; Pierre Baldi<sup>1</sup>; Ta-Wei Ho<sup>3</sup>

<sup>1</sup> UCI

<sup>2</sup> University of California Irvine (US)

<sup>3</sup> National Tsing Hua University (TW)

**Corresponding Authors:** michael.james.fenton@cern.ch, ta-wei.ho@cern.ch, ashmakov@uci.edu, pfbaldi@ics.uci.edu

One of the most ubiquitous challenges in analyses at the LHC is event reconstruction, whereby heavy resonance particles (such as top quarks, Higgs bosons, or vector bosons) must be reconstructed from the detector signatures left behind by their decay products. This is particularly challenging when all decay products have similar or identical signatures, such as all-jet events. Existing methods typically require the evaluation of every possible permutation of these events to find the “best” assignment. In this work, we present a novel neural network architecture “SPANet”, or Symmetry Preserving Attention networks. By casting this problem as a set assignment problem on a variable size set, and embedding our knowledge of the symmetries in the problem into the neural network

architecture, we demonstrate that these problems can be solved efficiently even in cases for which existing methods are intractable. We demonstrate the approach using a suite of progressively more complex benchmarks, going from all-hadronic ttbar, via ttH, to 4top final states, and provide an easy to use and flexible software package to design and train networks for arbitrary final states.

**Affiliation:**

UCI

**Academic Rank:**

Postdoc

**Exploring the Latent Structure of Data / 50**

## Symmetry Discovery with Deep Learning

**Authors:** Ben Nachman<sup>1</sup>; Krish Desai<sup>2</sup>; Jesse Thaler<sup>3</sup>

<sup>1</sup> *Lawrence Berkeley National Lab. (US)*

<sup>2</sup> *University of California, Berkeley*

<sup>3</sup> *MIT*

**Corresponding Authors:** benjamin.philip.nachman@cern.ch, krish.desai@berkeley.edu, jthaler@mit.edu

Symmetries are a fundamental property of functions applied to datasets. A key function for any dataset is the probability density, and the corresponding symmetries are often referred to as the symmetries of the dataset itself. We provide a rigorous statistical notion of symmetry for a dataset, which involves reference datasets that we call inertial in analogy to inertial frames in classical mechanics. Then, we construct a novel approach to automatically discover symmetries from a dataset using a deep learning method based on an adversarial neural network. We show how this model performs on simple examples and provide a corresponding analytic description of the loss landscape. Symmetry discovery may lead to new insights and can reduce the effective dimensionality of a dataset to increase its effective statistics.

**Affiliation:**

Department of Physics, University of California, Berkeley, Physics Division, Lawrence Berkeley National Laboratory,

**Academic Rank:**

PhD Student

**Exploring the Latent Structure of Data / 51**

## Jet Topology

**Authors:** Sijun Xu<sup>1</sup>; LINGFENG LI<sup>2</sup>; Tao Liu<sup>3</sup>

<sup>1</sup> *Hong Kong University of Science and Technology*

<sup>2</sup> *HKUST*

<sup>3</sup> *The Hong Kong University of Science and Technology*



**Corresponding Authors:** iaslfli@ust.hk, sijunxu.phy@gmail.com, taoliu@ust.hk

We introduce persistent Betti numbers to characterize topological structure of jets. These topological invariants measure multiplicity and connectivity of jet branches at a given scale threshold, while their persistence records evolution of each topological feature as this threshold varies. With this knowledge, in particular, we are able to reconstruct branch phylogenetic tree of each jet. These points are demonstrated in the benchmark scenario of light-quark versus gluon jets. This study provides a topological tool to develop jet taggers, and opens a new angle to look into jet physics.

**Affiliation:**

Hong Kong University of Science and Technology

**Academic Rank:**

PhD student

**New Horizons / 52**

## Conditional invertible neural networks to probe cosmic-ray sources

**Authors:** Josina Schulte<sup>1</sup>; Martin Erdmann<sup>2</sup>; Teresa Bister<sup>1</sup>; Ullrich Köthe<sup>None</sup>

<sup>1</sup> RWTH Aachen University

<sup>2</sup> Rheinisch Westfaelische Tech. Hoch. (DE)

**Corresponding Authors:** josina.schulte@rwth-aachen.de, martin.erdmann@cern.ch, teresa.bister@rwth-aachen.de

To obtain information on the still unknown sources of ultra-high-energy cosmic rays (UHECRs), a combined fit of the observed energy spectrum and depths of the shower maximum can be used, which constrains characteristic parameters of the sources. During propagation from the sources to Earth, UHECRs can experience numerous stochastic processes such that no explicit inverse function, which would describe the source parameters as a function of the measured quantities, can be formulated.

Previously, the high-dimensional space of possible combinations of source parameters has been investigated with Bayesian sampler methods like the Markov Chain Monte Carlo (MCMC), which in general is very computationally expensive. Here, we introduce the application of a new method using deep-learning techniques, the so-called *conditional Invertible Neural Network* (cINN). The network implicitly learns a mapping between the source parameters and the observables from a large set of training data. The backward pass through the trained invertible network then provides the full posterior probabilities including correlations. It needs far less computational resources than the MCMC and can be evaluated in seconds, and therefore the cINN enables extensive rapid tests of the method's accuracy. In this work, we compare the results of the two methods, MCMC and cINN, applied on a simulated scenario inspired by current UHECR measurements of the energy spectrum and the depths of shower maximum distributions. For the cINN, we also evaluate the performance on many test data sets and show that the mean estimates of the source parameters as well as the widths of the posterior distributions can be described accurately.

**Affiliation:**

RWTH Aachen University

**Academic Rank:**

PhD student

**Classification / 53****Combine and Conquer: Event Reconstruction with Bayesian Ensemble Neural Networks****Authors:** Jack Araz<sup>1</sup>; Michael Spannowsky<sup>2</sup><sup>1</sup> *IPPP - Durham University*<sup>2</sup> *University of Durham (GB)***Corresponding Authors:** jack.araz@durham.ac.uk, michael.spannowsky@cern.ch

Ensemble learning is a technique where multiple component learners are combined through a protocol. In this talk, we will present an Ensemble Neural Network (ENN) that uses the combined latent-feature space of multiple neural network classifiers to improve the representation of the network hypothesis. We apply this approach to construct an ENN from Convolutional and Recurrent Neural Networks to discriminate top-quark jets from QCD jets. Such ENN provides the flexibility to improve the classification beyond simple prediction combining methods by linking different sources of error correlations, hence improving the representation between data and hypothesis. In combination with Bayesian techniques, we show that it can reduce epistemic uncertainties and the entropy of the hypothesis by simultaneously exploiting various kinematic correlations of the system, which also makes the network less susceptible to a limitation in training sample size.

**Affiliation:**

IPPP - Durham University

**Academic Rank:**

Postdoctoral researcher

**Regression, Calibration, and Fast Inference / 54****Matrix Element Calculations on the GPU****Author:** Joshua Isaacson<sup>1</sup>**Co-authors:** Enrico Bothmann<sup>2</sup>; Max Knobbe<sup>3</sup>; Stefan Hoeche<sup>1</sup>; Walter Giele<sup>1</sup> *Fermilab*<sup>2</sup> *University of Göttingen*<sup>3</sup> *Georg-August-Universität Göttingen***Corresponding Authors:** max.knobbe@gmx.net, shoeche@fnal.gov, enrico.bothmann@uni-goettingen.de, giele@fnal.gov, isaacson@fnal.gov

Generating large numbers of events efficiently is a major bottleneck for ML projects. As a first step towards a full-fledged event generator for modern GPUs, we investigated different recursive strategies. The GPU implementations are compared to the state-of-the-art CPU codes, showing promise for using these in other pipelines. Finally, we propose baseline implementations for the development of a future full scale event generator on GPUs.

**Affiliation:**

Fermilab

**Academic Rank:**

BSM / 55

## CATHODE part 1: introducing a new model-agnostic search strategy for resonant new physics at the LHC

**Authors:** Anna Hallin<sup>1</sup>; Ben Nachman<sup>2</sup>; Claudius Krause<sup>3</sup>; David Shih<sup>3</sup>; Gregor Kasieczka<sup>4</sup>; Joshua Isaacson<sup>5</sup>; Manuel Sommerhalder<sup>4</sup>; Matthias Schlaffer<sup>None</sup>; Tobias Loesche<sup>4</sup>

<sup>1</sup> *Test IDP - Rutgers, The State University of New Jersey*

<sup>2</sup> *Lawrence Berkeley National Lab. (US)*

<sup>3</sup> *Rutgers University*

<sup>4</sup> *Hamburg University (DE)*

<sup>5</sup> *Fermilab*

**Corresponding Authors:** isaacson@fnal.gov, anna.hallin@rutgers.edu, benjamin.philip.nachman@cern.ch, matthias.j.schlaffer@gmail.com, dshih@physics.rutgers.edu, tobias.loesche@desy.de, claudius.krause@rutgers.edu, gregor.kasieczka@cern.ch, manuel.sommerhalder@cern.ch

We propose Classifier-based Anomaly detection THrough Outer Density Estimation (CATHODE), a new approach to search for resonant new physics at the LHC in a model-agnostic way. In CATHODE, we train a conditional density estimator on additional features in the sideband region, interpolate it into the signal region, and sample from it. This produces in a data-driven way events that follow the SM background model without any reliance on simulation. Then we train a classifier to distinguish background events and data events in the signal region to find anomalies. Using the LHCO R&D dataset, we show that CATHODE can discover new physics that would otherwise be hiding in the data, improving the nominal statistical significance in a specific example from  $\sim 1 \sigma$  to as much as  $\sim 15 \sigma$ .

### Affiliation:

Rutgers University

### Academic Rank:

PhD student

BSM / 56

## CATHODE part 2: robustness and comparison to other methods

**Authors:** Anna Hallin<sup>1</sup>; Manuel Sommerhalder<sup>2</sup>; Ben Nachman<sup>3</sup>; Claudius Krause<sup>4</sup>; David Shih<sup>4</sup>; Gregor Kasieczka<sup>2</sup>; Joshua Isaacson<sup>5</sup>; Matthias Schlaffer<sup>None</sup>; Tobias Loesche<sup>2</sup>

<sup>1</sup> *Test IDP - Rutgers, The State University of New Jersey*

<sup>2</sup> *Hamburg University (DE)*

<sup>3</sup> *Lawrence Berkeley National Lab. (US)*

<sup>4</sup> *Rutgers University*

<sup>5</sup> *Fermilab*

**Corresponding Authors:** claudius.krause@rutgers.edu, gregor.kasieczka@cern.ch, tobias.loesche@desy.de, anna.hallin@rutgers.edu, manuel.sommerhalder@cern.ch, dshih@physics.rutgers.edu, benjamin.philip.nachman@cern.ch, matthias.j.schlaffer@gmail.com, isaacson@fnal.gov

We explore the robustness of the CATHODE (Classifier-based Anomaly detection THrough Outer Density Estimation) method against correlation in the input features. We also compare CATHODE to other related approaches, specifically ANODE and CWoLa Hunting. Using the LHCO R&D dataset,

we will demonstrate that in the absence of feature correlations, CATHODE outperforms both ANODE and CWoLa Hunting, and even approaches the performance of a supervised classifier trained to distinguish data from background. Meanwhile, in the presence of feature correlations, CWoLa Hunting breaks down, while ANODE is robust. Here we demonstrate that CATHODE is also robust against correlations, maintaining its spectacular performance.

**Affiliation:**

Hamburg University

**Academic Rank:**

PhD Student

BSM / 57

## High-dimensional Anomaly Detection with Radiative Return in $e^+e^-$ Collisions

**Authors:** Julia Lynne Gonski<sup>1</sup>; Jerry Lai<sup>2</sup>; Ben Nachman<sup>3</sup>; Ines Ochoa<sup>4</sup>

<sup>1</sup> *Columbia University (US)*

<sup>2</sup> *University of California, Berkeley*

<sup>3</sup> *Lawrence Berkeley National Lab. (US)*

<sup>4</sup> *LIP Laboratoriao de Instrumentacao e Fisica Experimental de Part*

**Corresponding Authors:** ines.ochoa@cern.ch, benjamin.philip.nachman@cern.ch, julia.gonski@cern.ch, thejerrylai@berkeley.edu

Experiments at a future  $e^+e^-$  collider will be able to search for new particles with masses below the nominal centre-of-mass energy by analyzing collisions with initial-state radiation (radiative return). We show that machine learning methods based on semisupervised and weakly supervised learning can achieve model-independent sensitivity to the production of new particles in radiative return events. In addition to a first application of these methods in  $e^+e^-$  collisions, our study is the first to combine weak supervision with high-dimensional information by deploying a deep sets neural network architecture. We have also investigated some of the experimental aspects of anomaly detection in radiative return events and discuss these in the context of future detector design.

**Affiliation:**

Columbia University, University of California Berkeley, Lawrence Berkeley National Laboratory, Laboratory of Instrumentation and Experimental Particle Physics (Lisbon)

**Academic Rank:**

Datasets / 58

## Implementation of Jupyter Notebooks into The Reproducible Open Benchmarks for Data Analysis Platform (ROB)

**Authors:** Aaron Wang<sup>None</sup>; Ajay Rawat<sup>1</sup>; Heiko Mueller<sup>None</sup>; Shih-Chieh Hsu<sup>2</sup>

<sup>1</sup> *University of Washington (US)*

<sup>2</sup> *University of Washington Seattle (US)*

**Corresponding Authors:** anrunw@uw.edu, ajay.rawat@cern.ch, hm74@nyu.edu, schsu@uw.edu

The Reproducible Open Benchmarks for Data Analysis Platform (ROB) is a platform that allows for the evaluation of different data analysis algorithms in a controlled competition-style format [1]. One example for such a comparison and evaluation of different algorithms is the “The Machine Learning Landscape of Top Taggers” paper, which compiled and compared multiple different top tagger neural networks [2]. Motivated by the significant amount of time required to organize and evaluate such benchmarks, ROB provides a platform that automates the collection, execution, and comparison of participant submissions in a benchmark. Although convenient, the ROB currently requires participants to package their submissions into docker containers, which can pose an additional burden due to the steep learning curve.

To increase ease of use, we implement support for the commonly used Jupyter Notebooks [3] in ROB. Jupyter Notebooks are a popular tool that many physicists are already familiar with. Using Jupyter notebooks, physicists are able to combine live code, comments, and documentation inside one document. By utilizing the PaperMill package [4], we allow ROB users to submit their implementations directly as Jupyter Notebooks in order to evaluate different data analysis algorithms without the need to package the code into Docker containers. To demonstrate functionality and spur usage of the ROB, we provide demos using bottom and top tagging neural networks that display the application of the ROB within particle physics as a way of providing a competition style platform for algorithm evaluation [5].

References:

- [1] “Reproducible and Reusable Data Analysis Workflow Server”, <https://github.com/scaifin/flowserv-core>
- [2] Kasieczka, Gregor, Plehn, Tilman, Butter, Anja, Cranmer, Kyle, Debnath, Dipsikha, Dillon, Barry M, . . . Varma, Sreedevi. (2019). The Machine Learning landscape of top taggers. *SciPost Physics*, 7(1), 014.
- [3] “Jupyter Notebooks”, <https://jupyter.org/>
- [4] “Papermill”, <https://papermill.readthedocs.io/en/latest/>
- [5] “Particle Physics”, <https://github.com/anrunw/ROB>

**Affiliation:**

University of Washington

**Academic Rank:**

Undergraduate

**ML-Assisted Measurements and Searches / 59**

## Jet-based TMD measurements with H1 data, unfolded using machine-learning techniques

**Authors:** Ben Nachman<sup>1</sup>; Miguel Ignacio Arratia Munoz<sup>1</sup>; Stefan Schmitt<sup>2</sup>; Miguel Ignacio Arratia Munoz<sup>None</sup>

<sup>1</sup> *Lawrence Berkeley National Lab. (US)*

<sup>2</sup> *Deutsches Elektronen-Synchrotron (DE)*

**Corresponding Authors:** miguel.arratia@gmail.com, miguel.arratia@cern.ch, stefan.schmitt@cern.ch, benjamin.philip.nachman@cern.ch

Recently, jet measurements in DIS events close to Born kinematics have been proposed as a new probe to study transverse-momentum-dependent (TMD) PDFs, TMD fragmentation functions, and TMD evolution. We report measurements of lepton-jet momentum imbalance and hadron-in-jet correlations in high- $Q^2$  DIS events collected with the H1 detector at HERA. The jets are reconstructed with the kT algorithm in the laboratory frame. These are two examples of a new type of TMD studies in DIS, which will serve as pathfinder for the Electron-Ion Collider program.

**Affiliation:**

Lawrence Berkeley National Lab

**Academic Rank:****Simulation and Generative Models / 60****Angular Conditioning of Deep Generative Models for Fast Simulation of High Granularity Calorimeters**

**Authors:** Anatolii Korol<sup>1</sup>; Erik Buhmann<sup>2</sup>; Katja Kruger<sup>3</sup>; Lennart Rustige<sup>4</sup>; Peter McKeown<sup>5</sup>; Daniel Hundhausen<sup>5</sup>; Engin Eren<sup>5</sup>; Frank-Dieter Gaede<sup>3</sup>; Gregor Kasieczka<sup>2</sup>; Sascha Daniel Diefenbacher<sup>2</sup>; William Korcari<sup>2</sup>

<sup>1</sup> *Taras Shevchenko National University of Kyiv*

<sup>2</sup> *Hamburg University (DE)*

<sup>3</sup> *Deutsches Elektronen-Synchrotron (DE)*

<sup>4</sup> *CDCS / DESY*

<sup>5</sup> *Deutsches Elektronen-Synchrotron DESY*

**Corresponding Authors:** engin.eren@desy.de, lennart.rustige@desy.de, sascha.diefenbacher@desy.de, anatolii.korol@desy.de, william.korcari@cern.ch, frank-dieter.gaede@cern.ch, daniel.hundhausen@desy.de, katja.kruger@cern.ch, peter.mckeown@desy.de, erik.buhmann@desy.de, gregor.kasieczka@cern.ch

Modern high energy physics crucially relies on simulation to connect experimental observations to underlying theory. While traditional methods relying on Monte Carlo techniques produce powerful simulation tools, they prove to be computationally expensive. This is particularly true when they are applied to calorimeter shower simulation, where many particle interactions occur. The strain on computing resources due to simulation is projected to be so large as to be a major bottleneck at the high luminosity stage of the LHC and for future colliders.

Deep generative models have attracted significant attention as an approach which promises to drastically reduce the computing time required for simulation. Recent work in our group has demonstrated the capability of various generative models to accurately reproduce showers displaying key physics properties in a highly granular calorimeter. This initial work focused on the specific case of a particle incident perpendicular to the calorimeter face, however a practical simulator must incorporate arbitrary angles of incidence and simulate them correctly. This talk will address ongoing efforts to add conditioning on the incident angle of the particle. In particular, we demonstrate the crucial importance of modifying an existing loss function via the addition of an auxiliary constrainer network, in order to improve the angular performance of a generator.

**Affiliation:**

DESY

**Academic Rank:**

PhD student

**Interpretability, Robustness, and Uncertainties / 61****Safety of Quark/Gluon Jet Classification**

**Author:** Alexis Romero<sup>None</sup>

**Co-authors:** Pierre Baldi<sup>1</sup>; Julian Collado Umana<sup>2</sup>; Michael James Fenton<sup>2</sup>; Daniel Whiteson<sup>2</sup>

<sup>1</sup> UCI

<sup>2</sup> University of California Irvine (US)

**Corresponding Authors:** alexir2@uci.edu, pfbaldi@ics.uci.edu, julius506@gmail.com, daniel.whiteson@cern.ch, michael.james.fenton@cern.ch

The classification of jets as quark- versus gluon-initiated is an important yet challenging task in the analysis of data from high-energy particle collisions and in the search for physics beyond the Standard Model. The recent integration of deep neural networks operating on low-level detector information has resulted in significant improvements in the classification power of quark/gluon jet tagging models. However, the improved power of such models trained on simulated samples has come at the cost of reduced interpretability, raising concerns about their reliability. We elucidate the physics behind quark/gluon jet classification decisions by comparing the performance of networks with and without constraints of infrared and collinear safety, and identify the nature of the unsafe information by revealing the energy and angular dependence of the learned models. This in turn allows us to approximate the performance of the low-level networks (by 99% or higher) using equivalent sets of interpretable high-level observables, which can be used to probe the fidelity of the simulated samples and define systematic uncertainties.

**Affiliation:**

University of California, Irvine

**Academic Rank:**

BSM / 62

## Boosting new physics sensitivity with Variational Autoencoders

**Authors:** Kinga Anna Wozniak<sup>1</sup>; Maurizio Pierini<sup>2</sup>; Jennifer Ngadiuba<sup>3</sup>; Thea Aarrestad<sup>2</sup>; Thong Nguyen<sup>4</sup>; Javier Mauricio Duarte<sup>5</sup>

<sup>1</sup> University of Vienna (AT)

<sup>2</sup> CERN

<sup>3</sup> FNAL

<sup>4</sup> California Institute of Technology (US)

<sup>5</sup> Univ. of California San Diego (US)

**Corresponding Authors:** javier.mauricio.duarte@cern.ch, maurizio.pierini@cern.ch, kinga.anna.wozniak@cern.ch, jennifer.ngadiuba@cern.ch, thong@caltech.edu, thea.aarrestad@cern.ch

We show how an anomaly detection algorithm could be integrated in a typical search for new physics in events with jets at the CERN Large Hadron Collider (LHC). We assume that an anomaly detection algorithm is given, trained to identify rare jet types, such as jets originating from the decay of a highly boosted massive particle. We demonstrate how this algorithm could be integrated in a search without disrupting the background-estimate strategy while enhancing the sensitivity to new physics. As an example, we consider convolutional variational autoencoders (VAEs) applied to dijet events. The proposed procedure can be generalized to any final state with jets. Once applied to real data, it could contribute to extend the sensitivity of the LHC experiments to previously uncovered new physics scenarios, e.g. broad-resonance and non-resonant jet production from new physics processes.

**Affiliation:**

CERN, University of Vienna

**Academic Rank:**

**Interpretability, Robustness, and Uncertainties / 63****Learning Partially Known Stochastic Dynamics with Empirical PAC Bayes****Author:** Manuel Haußmann<sup>1</sup>**Co-authors:** Sebastian Gerwinn<sup>2</sup>; Andreas Look<sup>2</sup>; Barbara Rakitsch<sup>2</sup>; Melih Kandemir<sup>2</sup><sup>1</sup> *Universität Heidelberg*<sup>2</sup> *Bosch Center for Artificial Intelligence***Corresponding Author:** manuel.haussmann@iwr.uni-heidelberg.de

Neural Stochastic Differential Equations model a dynamical environment with neural nets assigned to their drift and diffusion terms. The high expressive power of their nonlinearity comes at the expense of instability in the identification of the large set of free parameters. This work presents a recipe to improve the prediction accuracy of such models in three steps: i) accounting for epistemic uncertainty by assuming probabilistic weights, ii) incorporation of partial knowledge on the state dynamics, and iii) training the resultant hybrid model by an objective derived from a PAC-Bayesian generalization bound. We observe in our experiments that this recipe effectively translates partial and noisy prior knowledge into an improved model fit.

**Affiliation:**

Heidelberg University

**Academic Rank:**

PhD student

**ML-Assisted Measurements and Searches / 64****Towards a new generation of PDFs using ML****Author:** Roy Stegeman<sup>1</sup><sup>1</sup> *University of Milan***Corresponding Author:** roystegeman@live.nl

We present the machine learning methodology that is the backbone of the new release of the NNPDF family of parton distribution functions. The new methodology introduces state of the art machine learning techniques such as stochastic gradient descent for neural network training which results in a major reduction in computational costs, and an automated optimization of the hyperparameters which reduces a source of bias. We further show how correlations between PDF sets can be used to assess the efficiency of the methodology, and why the use of correlations for the combination of different PDFs into a joint set could lead to severely distorted results. We discuss the “future test”, a recently developed method of validating the generalization power of the methodology, which checks whether the uncertainty on PDFs, in regions in which they are not constrained by current data, are compatible with future data.

**Affiliation:**

University of Milan

**Academic Rank:**



**Regression, Calibration, and Fast Inference / 65****Using Machine Learning for Heavy-Ion Jet  $p_T$  Reconstruction in ALICE****Author:** Hannah Bossi<sup>1</sup><sup>1</sup> *Yale University (US)***Corresponding Author:** hannah.bossi@cern.ch

Reconstructing the jet transverse momentum ( $p_T$ ) is a challenging task, particularly in heavy-ion collisions due to the large fluctuating background from the underlying event. In the recent years, ALICE has developed a novel method to correct jets for this large background using machine learning techniques. This analysis intentionally does not utilize deep learning methods and instead utilizes a shallow neural network for simplicity. This approach uses jet properties, including the constituents of the jet, to create a mapping between the corrected and uncorrected jet  $p_T$ . In comparison to the standard ALICE method, this machine learning based estimator demonstrates a significantly improved performance. The ML estimator was then applied to data in order to perform a measurement of full jets (jets containing both charged and neutral constituents) to lower transverse momenta than previously possible in ALICE. An ongoing challenge facing ML for jet physics is the interpretation of results and understanding potential biases. In this particular result, including constituent information in training introduces a bias towards PYTHIA-like fragmentation patterns, which has been shown to differ from the fragmentation measured in Pb–Pb data. Recent studies focusing on this bias in an attempt to further investigate and quantify its impact will be shown.

**Affiliation:**

Yale University

**Academic Rank:**

PhD student

**Datasets / 66****Shared Data and Algorithms for Deep Learning in Fundamental Physics**

**Authors:** William Korcari<sup>1</sup>; Lisa Benato<sup>1</sup>; Erik Buhmann<sup>1</sup>; Martin Erdmann<sup>2</sup>; Peter Fackelday<sup>None</sup>; Jonas Glombitza<sup>3</sup>; Nikolai Hartmann<sup>4</sup>; Gregor Kasieczka<sup>1</sup>; Thomas Kuhr<sup>4</sup>; Tilman Plehn<sup>5</sup>; Jan Steinheimer<sup>None</sup>; Horst Stoecker<sup>6</sup>; Kai Zhou<sup>7</sup>

<sup>1</sup> *Hamburg University (DE)*<sup>2</sup> *Rheinisch Westfaelische Tech. Hoch. (DE)*<sup>3</sup> *RWTH AACHEN UNIVERSITY*<sup>4</sup> *Ludwig Maximilians Universitat (DE)*<sup>5</sup> *Heidelberg University*<sup>6</sup> *GSI*<sup>7</sup> *FIAS, Goethe-University Frankfurt am Main*

**Corresponding Authors:** steinheimer@fias.uni-frankfurt.de, tilman.plehn@cern.ch, zhou@fias.uni-frankfurt.de, william.korcari@cern.ch, glombitza@physik.rwth-aachen.de, gregor.kasieczka@cern.ch, nikolai.marcel.hartmann@cern.ch, peter.fackelday@cern.ch, stoecker@uni-frankfurt.de, martin.erdmann@cern.ch, erik.buhmann@desy.de, thomas.kuhr@cern.ch, lisa.benato@cern.ch

We introduce a collection of datasets from fundamental physics research including particle physics, astroparticle physics, hadron, and nuclear physics for supervised machine learning studies. These

datasets, containing hadronic top quarks, cosmic air showers, phase transitions in the hadronic matter, and generator-level histories, are combined and made public to simplify future work on cross-disciplinary machine learning and transfer learning in fundamental physics. Based on these samples, we present two simple and yet flexible models: a fully connected neural network and a graph-based neural network architecture that can easily be applied to a wide range of supervised learning tasks in these domains. Furthermore, we show that our approaches reach performance close to state-of-the-art dedicated methods on all datasets.

**Affiliation:**

Hamburg University

**Academic Rank:**

PhD student

**Simulation and Generative Models / 67****Foundations of a Fast, Data-Driven, Machine-Learned Simulator****Author:** Jessica N. Howard<sup>1</sup>**Co-authors:** Stephan Mandt<sup>2</sup>; Daniel Whiteson<sup>1</sup>; Yibo Yang<sup>2</sup><sup>1</sup> *Department of Physics & Astronomy, UC Irvine*<sup>2</sup> *Department of Computer Science, UC Irvine***Corresponding Authors:** jnhoward@uci.edu, jessica.nicole.howard@cern.ch, daniel.whiteson@cern.ch, mandt@uci.edu, yibo.yang@uci.edu

We introduce a novel strategy for machine-learning-based fast simulators, which is the first that can be trained in an unsupervised manner using observed data samples to learn a predictive model of detector response and other difficult-to-model transformations. Across the physical sciences, a barrier to interpreting observed data is the lack of knowledge of a detector's imperfect resolution, which transforms and obscures the unobserved latent data. Modeling this detector response is an essential step for statistical inference, but closed-form models are often unavailable or intractable, requiring the use of computationally expensive, ad-hoc numerical simulations. Using particle physics detectors as an illustrative example, we describe a novel strategy for a fast, predictive simulator called Optimal Transport based Unfolding and Simulation (OTUS), which uses a probabilistic autoencoder to learn this transformation directly from observed data, rather than from existing simulations. Unusually, the probabilistic autoencoder's latent space is physically meaningful, such that the decoder becomes a fast, predictive simulator for a new latent sample, and has the potential to replace Monte Carlo simulators. We provide proof-of-principle results for  $Z$ -boson and top-quark decays, but stress that our approach can be widely applied to other physical science fields.

**Affiliation:**

Department of Physics &amp; Astronomy, UC Irvine

**Academic Rank:**

PhD student

**Compression / 68****Lorentz Group Equivariant Autoencoder**

**Authors:** Zichun Hao<sup>1</sup>; Daniel Diaz<sup>1</sup>; Javier Mauricio Duarte<sup>1</sup>; Raghav Kansal<sup>1</sup>; Farouk Mokhtar<sup>1</sup>

<sup>1</sup> *Univ. of California San Diego (US)*

**Corresponding Authors:** zihao@ucsd.edu, daniel.cipriano.diaz@cern.ch, fmokhtar@ucsd.edu, raghav.kansal@cern.ch, javier.mauricio.duarte@cern.ch

Symmetries are ubiquitous and essential in physics, and the framework to describe symmetries is group theory. The symmetry described by the Lorentz group is essential in the dynamics of all particle physics experiments. A Lorentz-group-equivariant deep neural network framework, called the Lorentz group network (LGN), has been introduced by Bogatskiy et al. and tested for performance in classifying jets. The model uses irreducible representations of the Lorentz group to achieve equivariance with respect to Lorentz transformations. However, the architecture has not yet been extended to generative, compression, or anomaly detection tasks yet. We develop an autoencoder based on the architecture of the LGN for jet compression and reconstruction tasks, using a complex permutation invariant loss function. The model is tested to be fully equivariant (within numerical precision) and is trained on a dataset of high momentum jets simulated at the LHC. We analyze the latent space after training, and explore how the choices of hyperparameters, such as the multiplicities of scalars and vectors in the latent spaces and the number of basis functions for the edge features, can influence the model's performance.

**Affiliation:**

Department of Physics, University of California, San Diego

**Academic Rank:**

Undergraduate student, Postdoctoral researcher, Professor, PhD student, PhD student

**New architectures / 69**

## **Equivariant energy flow networks for jet tagging**

**Authors:** Ayodele Ore<sup>1</sup>; Matthew Dolan<sup>2</sup>

<sup>1</sup> *The University of Melbourne*

<sup>2</sup> *University of Melbourne*

**Corresponding Authors:** ayodeleo@student.unimelb.edu.au, maitiu.o.dolain@gmail.com

The Energy Flow Network (EFN) is a neural network architecture that represents jets as point clouds and enforces infrared and collinear (IRC) safety on its outputs. In this talk, I will introduce a new variant of the EFN architecture based on the Deep Sets formalism, incorporating permutation-equivariant layers. I will discuss the conditions under which IRC safety can be maintained in the new architecture and showcase the performance of these networks on the canonical example of W-boson tagging. The equivariant EFNs have similar performance to Particle Flow Networks, which are superior to standard EFNs. Finally I will comment on how the equivariant networks sculpt the jet mass compared to unaugmented EFNs.

**Affiliation:**

The University of Melbourne

**Academic Rank:**

PhD Student

**Regression, Calibration, and Fast Inference / 70****Learning Uncertainties the Frequentist Way: Calibration and Correlation in High Energy Physics****Author:** Rikab Gambhir<sup>1</sup>**Co-authors:** Jesse Thaler<sup>1</sup>; Ben Nachman<sup>2</sup><sup>1</sup> MIT<sup>2</sup> Lawrence Berkeley National Lab. (US)**Corresponding Authors:** rikab@mit.edu, jthaler@mit.edu, benjamin.philip.nachman@cern.ch

A common problem that appears in collider physics is the inference of a random variable  $Y$  given a measurement of another random variable  $X$ , and the estimation of the uncertainty on  $Y$ . Additionally, one would like to quantify the extent to which  $X$  and  $Y$  are related. We present a machine learning framework for performing frequentist maximum likelihood inference with uncertainty estimation and measuring the mutual information between random variables. By using the Donsker-Varadhan representation of the KL divergence, the framework learns the likelihood ratio  $p(x|y)/p(x)$ . This can be used to calculate the mutual information between  $X$  and  $Y$ . The framework is parameterized using a Gaussian ansatz, which enables a manifest extraction of the maximum likelihood values and uncertainties. All of this can be accomplished in a single training of the model. We then demonstrate our framework for a simple Gaussian example, apply it to a realistic calibration task by calculating jet energy correction (JEC) and jet energy resolution (JER) factors for CMS open data.

**Affiliation:**

Massachusetts Institute of Technology

**Academic Rank:**

PhD Student

**ML-Assisted Measurements and Searches / 71****Combining Neural Network predictions with Hypothesis Testing for discovery in the LHC****Author:** Michael Soughton<sup>1</sup>**Co-authors:** Charanjit Kaur Khosa<sup>2</sup>; Veronica Sanz<sup>3</sup><sup>1</sup> University of Sussex<sup>2</sup> Universit' a di Genova<sup>3</sup> Universitat de Valencia**Corresponding Authors:** ms731@sussex.ac.uk, veronica.sanz@uv.es, khosacharanjit@gmail.com

As the use of Machine Learning techniques become more widespread within High Energy Physics it is important to consider how the results from Neural Networks can be applied within hypothesis testing. We show how a Log-Likelihood Ratio test can be performed using the the output of Neural Network classifiers trained on different physical datasets to yield a detection significance between two simple hypotheses, which we find to be superior to a common naive result. We also show how a generalised Log-Likelihood Ratio test can be performed using the output of a Variational Autoencoder when one hypothesis is not fully known beforehand, thereby providing a discovery significance that is useful in anomaly detection scenarios, which we showcase in the example of a Heavy Higgs EFT search.

**Affiliation:**

University of Sussex

**Academic Rank:**

PhD student

**Regression, Calibration, and Fast Inference / 72****ML in jet physics beyond classification****Author:** Loukas Gouskos<sup>1</sup><sup>1</sup> CERN**Corresponding Author:** loukas.gouskos@cern.ch

Advanced machine-learning techniques started recently to be explored by the CMS collaboration in various areas of jet physics, beyond jet classification. We present the most recent developments for the jet energy calibration and the jet mass reconstruction. In both cases novel algorithms using state-of-the-art machine-learning techniques have been developed. Significant improvement compared to traditional methods is observed which translates to improved physics reach.

**Affiliation:**

CERN

**Academic Rank:****Simulation and Generative Models / 73****AtlFast3: The next generation of fast simulation in ATLAS****Authors:** Douglas Michael Schaefer<sup>1</sup>; Michele Fauci Giannelli<sup>2</sup>; Hasib Ahmed<sup>3</sup>; Joshua Falco Beirer<sup>4</sup><sup>1</sup> University of Chicago (US)<sup>2</sup> INFN e Universita Roma Tor Vergata (IT)<sup>3</sup> The University of Edinburgh (GB)<sup>4</sup> CERN, Georg-August-Universitaet Goettingen (DE)**Corresponding Authors:** joshua.beirer@cern.ch, michele.fauci.giannelli@cern.ch, a.hasib@cern.ch, schaed@uchicago.edu

AtlFast3 is the next generation of high precision fast simulation in ATLAS that is being deployed by the collaboration and will replace AtlFastII, the fast simulation tool that was successfully used until now. AtlFast3 combines a parametrization-based Fast Calorimeter Simulation and a new machine-learning based Fast Calorimeter Simulation based on Generative Adversarial Networks (GANs). The new fast simulation improves the accuracy of simulating objects used in analyses when compared to Geant4, with a focus on those that were poorly modelled in AtlFastII. In particular, the simulation of jets of particles reconstructed with large radii and the detailed description of their substructure, are significantly improved in Atlfast3. Additionally the agreement between AtlFast3 and Geant4 is improved for high momentum  $\tau$ -leptons. The modelling and performance are evaluated on events produced at 13 TeV centre-of-mass energy in the Run-2 data-taking conditions.

**Affiliation:**

ATLAS

**Academic Rank:****Classification / 74**

## Boosted jet tagging in CMS

**Author:** Congqiao Li<sup>1</sup><sup>1</sup> *Peking University (CN)***Corresponding Author:** congqiao.li@cern.ch

Identification of hadronic decays of highly Lorentz-boosted W/Z/Higgs bosons and top quarks provides powerful handles to a wide range of new physics searches and Standard Model measurements at the LHC. This talk presents recent advances in boosted jet tagging algorithms in CMS. The application of novel machine-learning techniques has substantially improved the tagging performance and led to a significant increase in the physics reach.

**Affiliation:**

CERN

**Academic Rank:****Regression, Calibration, and Fast Inference / 75**

## Pileup mitigation in CMS

**Author:** Benedikt Maier<sup>None</sup>

At the LHC, each bunch crossings is able to create thousands of particles per collisions. Identifying a collision of interest from additional “pileup” collisions is a difficult task, requiring the development of dedicated methods. Commonly used methods are however not scalable to future LHC upgrades, where the average number of interactions will increase by almost an order of magnitude. To tackle this challenge, machine learning methods for pileup mitigation are currently been developed to improve and replace standard algorithms. In this talk, an overview of pileup mitigation methods using machine learning in CMS are described.

**Affiliation:**

CERN

**Academic Rank:****Classification / 76**

## Morphology for Jet Classification

**Authors:** Mihoko Nojiri<sup>1</sup>; Sung Hak Lim<sup>2</sup><sup>1</sup> *Theory Center, IPNS, KEK*<sup>2</sup> *Rutgers University*

**Corresponding Authors:** [nojiri@post.kek.jp](mailto:nojiri@post.kek.jp), [sunghak.lim@rutgers.edu](mailto:sunghak.lim@rutgers.edu)

We introduce a morphological analysis based on a neural network analyzing the Minkowski Functionals (MFs) of pixellated jet images. The MFs describe the geometric measures of binary images, and their changes by dilation encode the jet constituents' geometric structures that appear at various angular scales. We explicitly show that this morphological analysis can be considered a constrained convolutional neural network (CNN). Conversely, CNN could model the MFs, and we show their correlation in the example of tagging semi-visible jets emerging from the strong interaction of a hidden valley scenario. The MFs are independent of the IRC-safe observables commonly used in jet physics. We combine this morphological analysis with an IRC-safe relation network, which models two-point energy correlations. While the resulting network uses constrained input parameters, it shows comparable dark jet and top jet tagging performances to the CNN. The architecture has a significant advantage when the available data is limited, and we show that its tagging performance is much better than that of the CNN with a small number of training samples. We also qualitatively discuss their parton-shower model dependency. The results suggest that the MFs can be an efficient parameterization of the IRC-unsafe feature space of jets.

**Affiliation:**

Rutgers, the State University of New Jersey

**Academic Rank:**

Postdoctoral researcher

**ML-Assisted Measurements and Searches / 77**

## Parametrized classifiers for optimal EFT sensitivity

**Author:** Alfredo Glioti<sup>1</sup>

**Co-authors:** Siyu Chen<sup>1</sup>; Giuliano Panico<sup>2</sup>; Andrea Wulzer<sup>3</sup>

<sup>1</sup> EPFL - Ecole Polytechnique Federale Lausanne (CH)

<sup>2</sup> University of Florence and INFN Florence

<sup>3</sup> CERN and EPFL

**Corresponding Authors:** [giuliano.panico@cern.ch](mailto:giuliano.panico@cern.ch), [andrea.wulzer@cern.ch](mailto:andrea.wulzer@cern.ch), [alfredo.glioti@epfl.ch](mailto:alfredo.glioti@epfl.ch), [siyu.chen@cern.ch](mailto:siyu.chen@cern.ch)

We study unbinned multivariate analysis techniques, based on Statistical Learning, for indirect new physics searches at the LHC in the Effective Field Theory framework. We focus in particular on high-energy ZW production with fully leptonic decays, modeled at different degrees of refinement up to NLO in QCD. We show that a considerable gain in sensitivity is possible compared with current projections based on binned analyses. As expected, the gain is particularly significant for those operators that display a complex pattern of interference with the Standard Model amplitude. The most effective method is found to be the “Quadratic Classifier” approach, an improvement of the standard Statistical Learning classifier where the quadratic dependence of the differential cross section on the EFT Wilson coefficients is built-in and incorporated in the loss function. We argue that the Quadratic Classifier performances are nearly statistically optimal, based on a rigorous notion of optimality that we can establish for an approximate analytic description of the ZW process.

**Affiliation:**

EPFL

**Academic Rank:**

PhD student

**Simulation and Generative Models / 78****Sparse Data Generation with Convolutional VAE**

**Authors:** Breno Orzari<sup>1</sup>; Dimitrios Gunopulos<sup>None</sup>; Javier Mauricio Duarte<sup>2</sup>; Jean-Roch Vlimant<sup>3</sup>; Mary Touranakou<sup>4</sup>; Maurizio Pierini<sup>5</sup>; Raghav Kansal<sup>2</sup>; Thiago Tomei Fernandez<sup>1</sup>

<sup>1</sup> *UNESP - Universidade Estadual Paulista (BR)*

<sup>2</sup> *Univ. of California San Diego (US)*

<sup>3</sup> *California Institute of Technology (US)*

<sup>4</sup> *National and Kapodistrian University of Athens (GR)*

<sup>5</sup> *CERN*

**Corresponding Authors:** maurizio.pierini@cern.ch, mary.touranakou@cern.ch, breno.orzari@cern.ch, javier.mauricio.duarte@cern.ch, jean-roch.vlimant@cern.ch, raghav.kansal@cern.ch, thiago.tomei@cern.ch

A key aspect for the study of particle collisions is the comparison of the experiments data with those resulting from computer simulations, mainly obtained using Monte Carlo-based generators. However the amount of data required in simulations makes this task very time consuming. One approach to avoid this issue is by using machine learning techniques to speed up this process.

In this work, we focus on the simulation of one of the final-state objects of particle collisions that are the hadronic jets. For this study the input dataset consists of the particle constituents of the jets due to its sparsity and the possibility to assay the network's capacity to describe the jets and particles properties. The generative neural network architecture chosen is a variational autoencoder consisting of convolutional layers. For the reconstruction error term we choose a permutation-invariant loss on the particles' properties along with mean-squared error terms measuring the distinction between input and output jets transverse momentum and mass, which imposes physics constraints, allowing the model to learn the kinematics of the jets.

**Affiliation:**

SPRACE-UNESP, Brazil

**Academic Rank:**

PhD student

**New Horizons / 79****Via Machinae**

**Author:** Matthew Buckley<sup>1</sup>

**Co-authors:** David Shih<sup>1</sup>; Lina Necib<sup>2</sup>; John Tamanas<sup>3</sup>

<sup>1</sup> *Rutgers University*

<sup>2</sup> *California Institute of Technology*

<sup>3</sup> *University of California, Santa Cruz*

**Corresponding Authors:** dshih@physics.rutgers.edu, jtamanas@ucsc.edu, lnecib@mit.edu, mbuckley@physics.rutgers.edu

I describe a new machine learning algorithm, Via Machinae, to identify cold stellar streams in data from the Gaia telescope. Via Machinae is based on ANODE, a general method that uses conditional density estimation and sideband interpolation to detect local overdensities in the data in a model agnostic way. By applying ANODE to the positions, proper motions, and photometry of stars observed by Gaia, Via Machinae obtains a collection of those stars deemed most likely to belong to a stellar stream. In this talk, I will provide an overview of the Via Machinae algorithm, using the



known stream GD-1 as a worked example, and show preliminary results of our analysis across the full sky.

**Affiliation:**

Rutgers University

**Academic Rank:**

Professor

**Simulation and Generative Models / 80**

## Fast Simulation of Jets with VAEs

**Authors:** Breno Orzari<sup>1</sup>; Dimitrios Gunopoulos<sup>None</sup>; Mary Touranakou<sup>2</sup>; Javier Mauricio Duarte<sup>3</sup>; Jean-Roch Vlimant<sup>4</sup>; Maurizio Pierini<sup>5</sup>; Raghav Kansal<sup>3</sup>; Thiago Tomei Fernandez<sup>1</sup>

<sup>1</sup> UNESP - Universidade Estadual Paulista (BR)

<sup>2</sup> National and Kapodistrian University of Athens (GR)

<sup>3</sup> Univ. of California San Diego (US)

<sup>4</sup> California Institute of Technology (US)

<sup>5</sup> CERN

**Corresponding Authors:** jean-roch.vlimant@cern.ch, thiago.tomei@cern.ch, mary.touranakou@cern.ch, javier.mauricio.duarte@cern.ch, raghav.kansal@cern.ch, maurizio.pierini@cern.ch, breno.orzari@cern.ch

Typically, high-energy physics (HEP) data analysis heavily relies on the production and the storage of large datasets of simulated events. At the LHC, the end-to-end simulation workflow can require up to 50% of the available computing resources of an experiment. Speeding up the simulation process would be crucial to save resources that could be otherwise utilized.

In our study, we investigate the use of Deep Generative Models, and more specifically, of Variational Autoencoders for fast simulation of jets at the LHC. We represent jets as lists of jet constituents (particles) characterized by their momenta. Starting from a simulation of the jet before detector effects, we train a Deep Variational Autoencoder (VAE) to learn a parametric description of the detector response and produce the corresponding list of jet constituents. Doing so, we bypass both the detector simulation and the reconstruction, potentially speeding up significantly the events generation workflow.

For our domain-specific application, we use a permutation-invariant loss function, the Chamfer distance, as the reconstruction loss of the Variational Autoencoder based on the jet constituents. We further modify the reconstruction loss by adding extra penalty terms for the jet mass and the jet transverse momentum  $p_T$  to impose physics constraints to the model to learn the jet kinematics. Preliminary results show that jet features like jet pseudorapidity  $\eta$ , jet polar angle in the transverse plane  $\phi$ , or jet cartesian momenta  $p_x, p_y, p_z$  are modelled within 10% of accepted accuracy, whereas it is more complicated for the jet mass to be learned.

**Affiliation:**

National and Kapodistrian University of Athens

**Academic Rank:**

PhD student

**Exploring the Latent Structure of Data / 81**

## Attention and Dynamic Graph Convolution Neural Network in the context of classifying $t\bar{t}H(bb)$ vs. $t\bar{t}(bb)$ in the semi-leptonic top quark pair decay channel

**Authors:** Christina Reissel<sup>1</sup>; Gregor Kasieczka<sup>2</sup>; Pasquale Musella<sup>1</sup>; Tobias Loesche<sup>2</sup>

<sup>1</sup> *ETH Zurich (CH)*

<sup>2</sup> *Hamburg University (DE)*

**Corresponding Authors:** christina.reissel@cern.ch, gregor.kasieczka@cern.ch, tobias.loesche@desy.de, pasquale.musella@cern.ch

Deep neural networks (DNNs) are essential tools in particle physics targeting various use cases ranging from reconstruction of particles up to event classification and anomaly detection. Whereas DNNs for event classification are primarily trained on quantities deduced from the kinematic properties of the particles in the final state (high-level observables), we present an alternative approach of using exclusively basic kinematic object information (low-level observables) in combination with attention and dynamic graph convolution neural networks. Their performance is evaluated in terms of the potential of discriminating events in which the Higgs boson is produced in association with top quarks and decays into a pair of bottom quarks from the overwhelming background, which is top quark pair production in association with b jets. Studying the Higgs boson in such a complex final state with high object multiplicity is simultaneously challenging and crucially important for precision tests of the standard model.

Furthermore, DNNs are often seen as black boxes due to the difficulty of understanding what information they learn. We present detailed studies of the latent spaces of both the attention and graph networks in order to provide insights into what the neural networks learn about the physics and event topology as well as to demonstrate the potential of such networks for alternative analysis approaches, e.g., mass peak searches.

### Affiliation:

ETH Zurich

### Academic Rank:

PhD Student

### Classification / 82

## Identifying the Quantum Properties of Hadronic Resonances using Machine Learning

**Authors:** Ben Nachman<sup>1</sup>; Jakub Filippek<sup>2</sup>; John Kruper<sup>2</sup>; Kirtimaan Mohan<sup>None</sup>; Shih-Chieh Hsu<sup>3</sup>

<sup>1</sup> *Lawrence Berkeley National Lab. (US)*

<sup>2</sup> *University of Washington*

<sup>3</sup> *University of Washington Seattle (US)*

**Corresponding Authors:** benjamin.philip.nachman@cern.ch, kamohan@msu.edu, jk232@uw.edu, jakubflpk@gmail.com, schsu@uw.edu

With the great promise of deep learning, discoveries of new particles at the Large Hadron Collider (LHC) may be imminent. Following the discovery of a new Beyond the Standard model particle in an all-hadronic channel, deep learning can also be used to identify its quantum numbers. Convolutional neural networks (CNNs) using jet-images can significantly improve upon existing techniques to identify the quantum chromodynamic (QCD) ('color') as well as the spin of a two-prong resonance using its substructure. Additionally, jet-images are useful in determining what information

in the jet radiation pattern is useful for classification, which could inspire future taggers. These techniques improve the categorization of new particles and are an important addition to the growing jet substructure toolkit, for searches and measurements at the LHC now and in the future.

**Affiliation:**

University of Washington

**Academic Rank:**

Master's Student

**Simulation and Generative Models / 83**

## White Box AI for parton shower development

**Author:** Felix Ringer<sup>1</sup>

**Co-authors:** Mateusz Ploskon<sup>2</sup>; Yue Shi Lai ; Duff Neill

<sup>1</sup> *Lawrence Berkeley National Laboratory*

<sup>2</sup> *Lawrence Berkeley National Lab. (US)*

**Corresponding Authors:** f.m.ringer@gmail.com, mploskon@lbl.gov, duff.neill@gmail.com, ylai@phys.columbia.edu

We present an implementation of an explainable and physics-aware machine learning model capable of inferring the underlying physics of high-energy particle collisions using the information encoded in the energy-momentum four-vectors of the final state particles. We demonstrate the proof-of-concept of our White Box AI approach using a Generative Adversarial Network (GAN) which learns from a DGLAP-based parton shower Monte Carlo event generator. Our approach leads to a network that is able to learn not only the final distribution of particles, but also the underlying parton branching mechanism, i.e. the Altarelli-Parisi splitting function, the ordering variable of the shower, and the scaling behavior. While the current work is focused on perturbative physics of the parton shower, we foresee a broad range of applications of our framework to areas that are currently difficult to address from first principles in QCD. Examples include nonperturbative and collective effects, factorization breaking and the modification of the parton shower in heavy-ion, and electron-nucleus collisions.

**Affiliation:**

LBNL

**Academic Rank:****New architectures / 84**

## The information content of quenched jets

**Authors:** Felix Ringer<sup>1</sup>; James Mulligan<sup>2</sup>; Mateusz Ploskon<sup>3</sup>; Yue Shi Lai<sup>3</sup>

<sup>1</sup> *Lawrence Berkeley National Laboratory*

<sup>2</sup> *University of California, Berkeley (US)*

<sup>3</sup> *Lawrence Berkeley National Lab. (US)*

**Corresponding Authors:** f.m.ringer@gmail.com, yue.shi.lai@cern.ch, james.mulligan@berkeley.edu, mploskon@lbl.gov

In high energy heavy-ion collisions the substructure of jets is modified compared to that in proton-proton collisions due to the presence of the quark-gluon plasma (QGP). This modification of jets in the QGP is called “jet quenching”. We employ machine learning techniques to quantify how much information about this process is within the substructure observables. We formulate the question as a binary classification problem where the machine is trained to learn information that distinguishes jets in proton-proton and heavy-ion collisions. We perform the classification task using *i*) deep sets which directly includes Infrared-Collinear (IRC) unsafe information, and *ii*) a complete basis of IRC safe jet substructure observables which is passed to a Dense Neural Network (DNN). From the trained DNN, we identify optimal observables using symbolic regression. We perform our analysis using parton shower event generator models, and outline possible future directions to apply these methods directly to the raw data. We expect that the automated design of observables for heavy-ion collisions can provide new guidance for inferring properties of QGP from jet substructure data. In addition, the proposed framework for jets can be extended to study event-wide properties of any nuclear collisions - in particular, to study electron-ion collisions at the future Electron-Ion Collider.

**Affiliation:**

LBNL

**Academic Rank:**

Postdoc

**New architectures / 85**

## Identifying Heavy-Flavor Jets Using Vectors of Locally Aggregated Descriptors

**Author:** Jitka Mrazkova<sup>None</sup>**Co-authors:** Jana Bielcikova ; Raghav Kunnawalkam Elayavalli ; Georgy Ponimatkin**Corresponding Author:** jitka.mrazk@gmail.com

Jets of collimated particles originating from hard scattered partons are utilized in a wide range of analyses in high energy physics. Our study is focused on identifying jets originating from heavy quarks. We introduce a novel approach to tagging heavy-flavor jets at collider experiments utilizing the information contained within jet constituents via the JetVLAD model architecture. This model is based on the concept of Vectors of Locally Aggregated Descriptors, which takes a set of feature descriptors as an input and returns a fixed-length feature vector that characterizes each set. We show the performance of this model as characterized by common metrics and showcase its ability to extract high purity heavy-flavor jet sample at various realistic jet momenta and production cross-sections. The method was demonstrated on PYTHIA generated proton-proton collisions at center-of-mass energies 200 and 510 GeV.

**Affiliation:**

NPI CAS

**Academic Rank:****New architectures / 86**

## SO(3)-equivariant Neural Network for b-tagging

**Authors:** Chase Owen Shimmin<sup>1</sup>; Ema Catalina Smith<sup>None</sup><sup>1</sup> Yale University (US)

**Corresponding Authors:** ema.smith@yale.edu, chase.shimmin@cern.ch

Jets originating from bottom quarks, b-jets, are of particular interest in high energy physics. While b-jets are similar to other jets, they have certain qualities that present unique challenges in the context of machine learning. Generally, there is an underlying rotational symmetry of the particles about a jet's axis. However, in the case of b-jets, some of the most discriminating observables, the transverse and longitudinal impact parameters, are not compatible with this symmetry. To address this, we instead consider the full 3-dimensional vector-valued impact parameter as the relevant observable for b-tagging. We then compare the b-tagging performance of a novel SO(3)-equivariant neural network against a standard baseline architecture of similar complexity. A preliminary study using DELPHES-simulated events with track smearing shows that the equivariant network significantly outperformed the baseline model. In particular, the equivariant model was less prone to over-training and improved background rejection overall and by over 1.5x at the 70% working point.

**Affiliation:**

Yale University

**Academic Rank:**

Undergraduate

**Classification / 87**

## Pushing the limit of jet tagging with graph neural networks

**Author:** Huilin Qu<sup>1</sup>

<sup>1</sup> CERN

**Corresponding Author:** huilin.qu@cern.ch

Graph neural networks (GNNs) have shown a lot of potential for jet tagging. Recent GNN algorithms such as ParticleNet, ABCNet, and LundNet represent the state-of-the-art in various jet tagging tasks. In this talk, we present some new progress on GNN design for jet tagging. With the incorporation of edge features and optimized network architecture, the new algorithm achieves a significant performance improvement on the top tagging benchmark.

**Affiliation:**

CERN

**Academic Rank:**

Postdoctoral researcher

**BSM / 88**

## Classifier-based Anomalous Jet Tagging

**Authors:** Taoli Cheng<sup>1</sup>; Aaron Courville<sup>2</sup>

<sup>1</sup> University of Montreal

<sup>2</sup> Mila, University of Montreal

**Corresponding Author:** chengtaoli.1990@gmail.com

As an alternative approach (w.r.t. deep generative models) for detecting out-of-distribution samples, we explore the possibility of employing jet classifiers as anomalous jet taggers. We also discuss the advantages and limitations of different approaches.

**Affiliation:**

Mila, University of Montreal

**Academic Rank:**

**New architectures / 89**

## Particle Convolution for Jets

**Author:** Chase Owen Shimmin<sup>1</sup>

<sup>1</sup> *Yale University (US)*

**Corresponding Author:** chase.shimmin@cern.ch

We introduce the Particle Convolution Network (PCN), a new type of equivariant neural network layer suitable for many tasks in jet physics. The particle convolution layer can be viewed as an extension of Deep Sets and Energy Flow network architectures, in which the permutation-invariant operator is promoted to a group convolution. While the PCN can be implemented for various kinds of symmetries, we consider the specific case of rotation in the  $\eta - \phi$  plane. In two standard benchmark tasks, q/g tagging and top tagging, we show that the rotational PCN (rPCN) achieves performance comparable to graph networks such as ParticleNet. Moreover, we show that it is possible to implement an IRC-safe rPCN, which significantly outperforms existing IRC-safe tagging methods on both tasks. We speculate that by generalizing the PCN to include additional convolutional symmetries relevant to jet physics, it may outperform the current state-of-the-art set by graph networks, while allowing some control over physically-motivated inductive biases.

**Affiliation:**

Yale University

**Academic Rank:**

Postdoc

**Simulation and Generative Models / 90**

## Multi-detector geometery modeling and Geant4 Integration

**Author:** Dalila Salamani<sup>1</sup>

**Co-authors:** Anna Zaborowska<sup>1</sup>; Witold Pokorski<sup>1</sup>

<sup>1</sup> *CERN*

**Corresponding Authors:** witold.pokorski@cern.ch, dalila.salamani@cern.ch, anna.zaborowska@cern.ch

The extensive physics program of HEP experiments relies on simulated Monte Carlo events. This simulation provides a highly detailed detector response modeling. However, this simulation dominated by the calorimeter showers becomes very slow in the context of high luminosity LHC. Collecting

order of magnitude more data remains necessary to lower the statistical uncertainties. Several research directions investigated the use of Machine Learning (ML) based models to for fast simulation of one specific detector. Each models tries to mimic the subtle and complex detector response resulting in a very finely tuned simulation. In this study, we explore the use of a ML multi-detector geometry model for fast simulation.

**Affiliation:**

CERN

**Academic Rank:**

**Datasets / 91**

## Introduction to Anomaly Detection Challenge

**Authors:** Ema Puljak<sup>1</sup>; Jennifer Ngadiuba<sup>2</sup>; Katya Govorkova<sup>3</sup>; Thea Aarrestad<sup>3</sup>; Maurizio Pierini<sup>3</sup>

<sup>1</sup> *Rudjer Boskovic Institute (HR)*

<sup>2</sup> *FNAL*

<sup>3</sup> *CERN*

**Corresponding Authors:** ema.puljak@cern.ch, maurizio.pierini@cern.ch, thea.aarrestad@cern.ch, ekaterina.govorkova@cern.ch, jennifer.ngadiuba@cern.ch

The data challenge is “anomaly detection @ 40 MHz” for which the biggest concern is to fit an algorithm in the tight constraints, which are presented in the talk. Considering as a benchmark an inclusive data stream, which has been pre-filtered by requiring the presence of one lepton, we discuss different possible strategies to detect new physics events as anomalies. The main goal of the challenge is to then seek for new ideas on how to do anomaly detection.

**Affiliation:**

**Academic Rank:**

**BSM / 92**

## Generative Adversarial Networks for Anomaly Detection at the LHC

**Authors:** Ben Nachman<sup>1</sup>; Daniel Sun<sup>2</sup>; David Shih<sup>3</sup>; Dukaixuan Ling<sup>2</sup>; Elham E Khoda<sup>4</sup>; Htet Aung Myin<sup>4</sup>; Ines Ochoa<sup>5</sup>; Shih-Chieh Hsu<sup>6</sup>

<sup>1</sup> *Lawrence Berkeley National Lab. (US)*

<sup>2</sup> *University of Washington*

<sup>3</sup> *Rutgers University*

<sup>4</sup> *University of Washington (US)*

<sup>5</sup> *LIP Laboratorio de Instrumentacao e Fisica Experimental de Part*

<sup>6</sup> *University of Washington Seattle (US)*

**Corresponding Authors:** schsu@uw.edu, danielcsun@hotmail.com, benjamin.philip.nachman@cern.ch, elham.e.khoda@cern.ch, ldkxnb@gmail.com, dshih@physics.rutgers.edu, ines.ochoa@cern.ch, htet.aung.myin@cern.ch

Anomaly detection techniques offer exciting possibilities to significantly extend the search for new physics at the Large Hadron Collider (LHC) in a model-agnostic approach. We study how Generative Adversarial Networks could be used for this purpose, using the LHC Olympics 2020 dataset as an example.

**Affiliation:**

**Academic Rank:**

**New Horizons / 93**

## **ML in Cosmological Simulations**

**Interpretability, Robustness, and Uncertainties / 94**

## **Generative Networks with Uncertainties**

**Authors:** Manuel Haußmann<sup>1</sup>; Marco Bellagente<sup>1</sup>; Michel Luchmann<sup>None</sup>; Tilman Plehn<sup>None</sup>

<sup>1</sup> *Universität Heidelberg*

**Corresponding Author:** plehn@uni-heidelberg.de

We show how Bayesian neural networks can be used to estimate uncertainties associated with regression, classification, and now also generative networks. For generative INNs, the combination of the learned density and uncertainty maps also provide insights into how these networks learn. These results show that criticizing the use of neural networks in LHC physics as black boxes is a sociological rather than scientific statement.

**Academic Rank:**

**Affiliation:**

**ML-Assisted Measurements and Searches / 95**

## **Invertible Neural Networks beyond Particle Physics**

Invertible Neural Networks (INNs) are an extremely versatile class of generative models. Their invertibility allows for exact modelling of probability densities, computation of information-theoretic quantities, interpretable and disentangled features, among other things. Due to these properties, INNs have seen growing adoption in recent years, especially in natural sciences and engineering disciplines. In this talk, we present a number of examples for successful applications of INN-specific methods to real-world problems, covering various scientific fields beyond particle physics.

**Academic Rank:**

**Affiliation:**



BSM / 96

## Recognizing hadronic SUEP at the LHC with Unsupervised Machine Learning

**Corresponding Author:** jared.barron@mail.utoronto.ca

Models with dark showers represent one of the most challenging possibilities for new physics at the LHC. One of the most difficult examples is a novel collider signature called a Soft Unclustered Energy Pattern (SUEP), which can arise in certain BSM models with a hidden valley sector that is both pseudo-conformal and strongly coupled over a large range of energy scales. Large-angle emissions are unsuppressed during the showering process, and if the hidden sector hadrons decay hadronically and promptly back into the Standard Model, the result is a high-multiplicity shower of SM final state particles that possess more democratically distributed energies and a much higher degree of isotropy than typically seen in QCD jets. This signature presents significant challenges to model, trigger on and search for, due to high theoretical uncertainties and the lack of isolated hard objects to identify in the detector. We outline an analysis strategy to look for SUEP produced by exotic decays of the Higgs boson, using conventional cuts on event-level observables and employing an autoencoder neural network trained on QCD background as an anomaly detector. We compare this unsupervised approach to a simple cut-and-count strategy as well as supervised machine learning models. We find that our strategy could allow the HL-LHC to exclude branching ratios of Higgs decay to SUEP down to a few percent.

**Academic Rank:**

**Affiliation:**

**Simulation and Generative Models / 97**

## Super-Resolution for QCD and Top Jets

**Authors:** Anja Butter<sup>None</sup>; Daniel Whiteson<sup>1</sup>; Gregor Kasieczka<sup>2</sup>; Jessica Nicole Howard<sup>1</sup>; Lukas Blecher<sup>3</sup>; Pierre Baldi<sup>4</sup>; Tilman Plehn<sup>None</sup>

<sup>1</sup> *University of California Irvine (US)*

<sup>2</sup> *Hamburg University (DE)*

<sup>3</sup> *Universität Heidelberg*

<sup>4</sup> *UCI*

**Corresponding Author:** plehn@uni-heidelberg.de

QCD-jets at the LHC are described by simple physics principles. We show how super-resolution generative networks can learn the underlying structures and use them to improve the resolution of jet images. We test this approach on massless QCD-jets and on fat top-jets and find that the network reproduces their main features even without training on pure samples. In addition, we show how a slim network architecture can be constructed once we have control of the full network performance.

**Academic Rank:**

Master Student

**Affiliation:**

Heidelberg University

BSM / 98

## Review of the Dark Machine Anomaly Score Challenge II

**Author:** Bryan Ostdiek<sup>1</sup>

<sup>1</sup> *Harvard University*

**Corresponding Author:** bostdiek@gmail.com

We describe the outcome of a data challenge conducted as part of the Dark Machines Initiative and the Les Houches 2019 workshop on Physics at TeV colliders. The challenge aims at detecting signals of new physics at the LHC using unsupervised learning algorithms. We define and describe a large benchmark dataset, consisting of > 1 Billion simulated LHC events. We then review a wide range of anomaly detection algorithms and measure their performance on the data challenge. We then assess the best-performing models on a still blinded dataset. Similarities between the best-performing models are observed and discussed.

**Academic Rank:**

Postdoc

**Affiliation:**

Harvard University

**Datasets / 99**

## Calorimeter Simulation Challenge Proposal

**Authors:** Ben Nachman<sup>1</sup>; Dalila Salamani<sup>2</sup>; David Shih<sup>3</sup>; Gregor Kasieczka<sup>4</sup>; Michele Fauci Giannelli<sup>5</sup>

<sup>1</sup> *Lawrence Berkeley National Lab. (US)*

<sup>2</sup> *CERN*

<sup>3</sup> *Rutgers University*

<sup>4</sup> *Hamburg University (DE)*

<sup>5</sup> *INFN e Universita Roma Tor Vergata (IT)*

**Corresponding Authors:** michele.fauci.giannelli@cern.ch, gregor.kasieczka@cern.ch, dshih@physics.rutgers.edu, dalila.salamani@cern.ch, benjamin.philip.nachman@cern.ch

**Interpretability, Robustness, and Uncertainties / 100**

## Bayesian Inference in for Four Tops at the LHC

**Author:** Ezequiel Alvarez de los Alvarez de San Luis<sup>None</sup>

**Corresponding Author:** sequi@df.uba.ar

Four-tops (and its backgrounds) is very hard to model at the LHC, it represents a unique window for detecting top-philic NP, and its current measurements have some tension with theory and predictions. We find that simple, clean and powerful Bayesian Inference can be applied on the data to infer signal and background true distributions. We propose that these results could be used in a novel way to test for SM agreement and/or NP effects in four-top final state at the LHC.

**Academic Rank:**

**Affiliation:**

**New Horizons / 101**

**ML4Jets2022**

**Corresponding Author:** [dshih@physics.rutgers.edu](mailto:dshih@physics.rutgers.edu)

**102**

**Closeout and ML4Jets2022**

**Corresponding Authors:** [plehn@uni-heidelberg.de](mailto:plehn@uni-heidelberg.de), [dshih@physics.rutgers.edu](mailto:dshih@physics.rutgers.edu), [benjamin.philip.nachman@cern.ch](mailto:benjamin.philip.nachman@cern.ch)