

ML4Jets2022

Tuesday, 1 November 2022 - Friday, 4 November 2022

Rutgers University

Book of Abstracts

Contents

Boost-Invariant Polynomials: an efficient and interpretable approach to jet tagging . . .	1
Score-based Generative Models for Calorimeter Shower Simulation	1
Multi-differential Jet Substructure Measurement in High Q^2 Deep-Inelastic Scattering with the H1 Detector	1
Modeling Hadronization with Machine Learning	2
Heterogeneous Graph Representation for Identifying Hadronically Decayed Tau Leptons at the High Luminosity LHC	2
Equivariant Point Cloud Generation for Particle Jets	2
Solving Combinatorial Problems in Multijet Signatures Using Machine Learning	3
Learning to Identify Semi-Visible Jets	3
Moment Unfolding using Deep Learning	4
Results from Unsupervised Machine Learning in an ATLAS Dijet Resonance Search . . .	4
Truth tagging for efficiency parametrization of b-jets using Graph Neural Networks . . .	5
Challenges for unsupervised anomaly detection in particle physics	5
Estimating Uncertainties for Trained Neural Networks	5
Invertible Networks for the Matrix Element Method	6
MadNIS: Neural networks for multi-channel integration	6
Optimal Mass Variables for Semivisible Jets	7
IEA-GAN: Intra-Event Aware GAN with Relational Reasoning for the Fast Detector Simulation	7
A boosted kNN regressor with 66 million parameters	7
Blueprints for Training Information Bottlenecks for Collider Analyses	8
Topological Data Analysis for Collider Events	8
Adversarial training for b-tagging algorithms in CMS	9
Optimal transport solutions for pileup mitigation at hadron colliders	9

Graph Neural Networks for a Deep-learning based Full Event Interpretation (DFEI) at the LHCb trigger	10
Neural Embedding: Learning the Embedding of the Manifold of Physics Data	10
Does Lorentz-symmetric design boost network performance in jet physics?	11
Constituent-Based Top-Quark Tagging with the ATLAS Detector	11
Generative Models for Fast Simulation of Electromagnetic and Hadronic Showers in Highly Granular Calorimeters	11
Transformer models for heavy flavor jet identification in CMS	12
HEP-Sim2Real: creating background templates with normalizing flows	12
Machine learning based jet and event classification at the Electron-Ion Collider	13
CURTAINS for your Sliding Window: Constructing Unobserved Regions by Transporting Adjacent Intervals	13
Robust Signal Detection using a Classifier Decorrelated through Optimal Transport (CDOT)	14
Neural Estimation of Energy Movers Distance	14
VBF vs. GGF Higgs with Full-Event Deep Learning: Towards a Decay-Agnostic Tagger	15
CaloMan: Fast generation of calorimeter showers with density estimation on learned manifolds	15
Recent ML-usage in searches with boosted jets in CMS	16
Feature selection with Distance Correlation	16
Can You Hear the Shape of a Jet?	16
Search for dimuon events in IceCube using decision trees	17
Jet tagging with deep sets of subsets	17
Point Cloud Deep Learning Methods for Pion Reconstruction in the ATLAS Experiment	17
Conditional generative networks for pure quark and gluon jets	18
Constraining quark and gluon jet energy loss distributions in quark-gluon plasma using Bayesian inference	18
Transformer Architectures for Quenched Jet Tagging	19
Resilience of Quark-Gluon Tagging	19
v-flows: Conditional neutrino momentum regression	20
Particle reconstruction in jets with set transformer and hypergraph prediction architectures	20
Weakly Supervised Learning for Muon Discrimination in Unlabeled Collider Data	20

A Normalized Autoencoder for LHC triggers	21
Point Cloud Generation using Transformer Encoders and Normalising Flows	21
Machine learning for top physics in CMS	22
Overview of ML for Gravitational Waves	22
Fast calorimeter simulation with VQVAE	23
GalaxyFlow: Upsampling Hydrodynamical Simulations for Realistic Gaia Mock Catalogs	23
Equivariant Neural Networks for Particle Physics: PELICAN	24
Likelihood-Free Frequentist Inference for Calorimetric Muon Energy Measurement . . .	24
Jet SIFT-ing	25
Weakly supervised methods for LHC analyses	25
Robust anomaly detection using NuRD	26
Measuring Galactic Dark Matter through Unsupervised Machine Learning	26
Hunting for signals using Gaussian Process regression	26
A Holistic Approach to Predicting Top Quark Kinematic Properties with the Covariant Particle Transformer	27
Generative Models for Resonant Anomaly Detection	27
Infra-red and collinear safe Graph Neural Networks	28
Using Machine Learning to Improve our Understanding of the Jet Background in Nucleus- Nucleus Collisions.	28
CaloFlow for CaloChallenge	29
AtlFast3, the new ATLAS fast simulation tool	29
Particle Cloud Generation	30
ML Unfolding based on conditional Invertible Neural Networks using iterative training .	30
Advances in developing deep neural networks for finding primary vertices in proton-proton collisions at the LHC	30
Registration	31
Welcome from Local Organizers	31
Theory Opening	31
Experimental Opening I	31
Experimental Opening II	31
Closing	32

Loop Amplitudes from Precision Networks	32
ML Keynote Talk – Generative models, manifolds and symmetries: From QFT to molecules	32
Resonant anomaly detection without background sculpting	33
Introduction to Generative Models for Fast Detector Simulation	33
Introduction to Anomaly Detection	33
Discussion	33
Overview of ML for Astro/Cosmo	33
Overview of ML for Neutrinos	34
Panel Discussion	34
Overview of ML for Gaia	34
Symmetries, Safety, and Self-Supervision	35
Identification of hadronic tau decays using a deep neural network with the CMS experiment at LHC	35
Machine learning for particle flow at CMS	35
How can Bayesian networks be used for uncertainty quantification in particle physics?	36
Anomaly detection in a perspective of interdisciplinary research	36

Interpretability / 1**Boost-Invariant Polynomials: an efficient and interpretable approach to jet tagging**

Authors: Christoph Ortner¹; Ilyes Batatia²; Jose Miguel Munoz Arias³

¹ *UBC Math Department*

² *Engineering Laboratory, University of Cambridge*

³ *ELA University*

Corresponding Authors: jose.miguel.munoz.arias@cern.ch, christophortner0@gmail.com, ilyes.batatia@ens-paris-saclay.fr

Besides modern architectures designed via geometric deep learning achieving high accuracies via Lorentz group invariance, this process involves high amounts of computation. Moreover, the framework is restricted to a particular classification scheme and lacks interpretability.

To tackle this issue, we present BIP, an efficient and computationally cheap framework to build rotational, permutation, and boost in the jet mean axis invariances. Moreover, we show the versatility of our approach to obtaining state-of-the-art range accuracies in both supervised and unsupervised jet tagging by using several out-of-the-box classifiers.

Generative Models – Detector Level / 2**Score-based Generative Models for Calorimeter Shower Simulation**

Authors: Ben Nachman¹; Vinicius Massami Mikuni¹

¹ *Lawrence Berkeley National Lab. (US)*

Corresponding Authors: vinicius.massami.mikuni@cern.ch, benjamin.philip.nachman@cern.ch

Score-based generative models are a new class of generative algorithms that have been shown to produce realistic images even in high dimensional spaces, currently surpassing other state-of-the-art models for different benchmark categories and applications. In this work we introduce CaloScore, a score-based generative model for collider physics applied to calorimeter shower generation. Three different diffusion models are investigated using the Fast Calorimeter Simulation Challenge 2022 dataset. CaloScore is the first application of a score-based generative model in collider physics and is able to produce high-fidelity calorimeter images for all datasets, providing an alternative paradigm for calorimeter shower simulation.

Measurement / 3**Multi-differential Jet Substructure Measurement in High Q^2 Deep-Inelastic Scattering with the H1 Detector**

Authors: Ben Nachman¹; Stefan Schmitt²; Vinicius Massami Mikuni¹

¹ *Lawrence Berkeley National Lab. (US)*

² *Deutsches Elektronen-Synchrotron (DE)*

Corresponding Authors: benjamin.philip.nachman@cern.ch, vinicius.massami.mikuni@cern.ch, stefan.schmitt@cern.ch

A study of different jet observables in high Q^2 Deep-Inelastic Scattering events close to the Born kinematics is presented. Differential and multi-differential cross-sections are presented as a function of the jet's charged constituent multiplicity, momentum dispersion, jet charge, as well as three values of jet angularities. Results are split into multiple Q^2 intervals, probing the evolution of jet observables with energy scale. These measurements probe the description of parton showers and provide insight into non-perturbative QCD. Unfolded results are derived without binning using the machine learning-based method Omnifold. All observables are unfolded simultaneously by using reconstructed particles inside jets as inputs to a graph neural network. Results are compared with a variety of predictions.

H1prelim-22-03

Generative Models – Particle Level / 4

Modeling Hadronization with Machine Learning

Author: Manuel Szwec^{None}

Corresponding Author: manuelsing24@gmail.com

A fundamental part of event generation, hadronization is currently simulated with the help of fine-tuned empirical models. In this talk, I'll present MLHAD, a proposed alternative for hadronization where the empirical model is replaced by a surrogate Machine Learning-based model to be ultimately data-trainable. I'll detail the current stage of development and discuss possible ways forward.

Classification / 5

Heterogeneous Graph Representation for Identifying Hadronically Decayed Tau Leptons at the High Luminosity LHC

Authors: Andris Huang¹; Xiangyang Ju²

¹ *University of California-Berkeley*

² *Lawrence Berkeley National Lab. (US)*

Corresponding Authors: xiangyang.ju@cern.ch, andrewhz@berkeley.edu

We present a new algorithm that identifies reconstructed jets originating from hadronic decays of tau leptons against those from quarks or gluons. No tau lepton reconstruction algorithm is used. Instead, the algorithm represents jets as heterogeneous graphs using the associated low-level objects such as tracks and energy clusters and trains a Graph Neural Network (GNN) to identify hadronically decayed tau leptons from other jets. Simulated events are generated to emulate the dense environment at the High Luminosity Large Hadron Collider (HL-LHC). We compare the physics performance and the computational effectiveness for different graph representations of jets and for different GNNs (homogeneous vs heterogeneous). In addition, we compare the GNNs with the RNN that is used in ATLAS.

Equivariance and New Architectures / 6

Equivariant Point Cloud Generation for Particle Jets

Authors: Erik Buhmann¹; Gregor Kasieczka¹; Jesse Thaler²

¹ *Hamburg University (DE)*

² *MIT*

Corresponding Authors: jthaler@mit.edu, gregor.kasieczka@cern.ch, erik.buhmann@desy.de

With current and future high-energy collider experiments' vast data collecting capabilities comes an increasing demand for computationally efficient simulations. Generative machine learning models allow fast event generation, yet so far are largely constrained to fixed data and detector geometries.

We introduce a Deep Sets based permutation equivariant generative adversarial network (GAN) for generation of permutation invariant point clouds with variable cardinality - a flexible data structure optimal for collider events such as jets. The generator utilizes an interpretable global latent vector and does not rely on pairwise information sharing between particles, leading to a significant speed-up over graph-based approaches. The model can be fine-tuned for minimal information sharing between particles and model complexity. We show that our GAN scales well to large particle multiplicities and achieves high generation fidelity for quark jets.

Equivariance and New Architectures / 7

Solving Combinatorial Problems in Multijet Signatures Using Machine Learning

Author: Lawrence Lee Jr¹

Co-author: Anthony Badea²

¹ *University of Tennessee (US)*

² *Harvard University (US)*

Corresponding Authors: abadea@g.harvard.edu, lawrence.lee.jr@cern.ch

High-multiplicity signatures at particle colliders can arise in Standard Model processes and beyond. With such signatures, difficulties often arise from the large dimensionality of the kinematic space. For final states containing a single type of particle signature, this results in a combinatorial problem that hides underlying kinematic information. We explore using a neural network that includes a Lorentz Layer to extract high-dimensional correlations. We use the case of squark decays to jets in R-Parity-violating Supersymmetry as a benchmark, comparing the performance to that of classical methods. With this approach, we demonstrate significant improvement over traditional methods. Based on arXiv:2201.02205.

Interpretability / 8

Learning to Identify Semi-Visible Jets

Authors: Daniel Whiteson¹; Shih-Chieh Hsu²; Taylor James Faucett³

¹ *University of California Irvine (US)*

² *University of Washington Seattle (US)*

³ *University of California, Irvine*

Corresponding Authors: schsu@uw.edu, daniel.whiteson@cern.ch, tfaucett@uci.edu

We train a network to identify jets with fractional dark decay (semi-visible jets) using the pattern of their low-level jet constituents, and explore the nature of the information used by the network by mapping it to a space of jet substructure observables. Semi-visible jets arise from dark matter particles which decay into a mixture of dark sector (invisible) and Standard Model (visible) particles. Such objects are challenging to identify due to the complex nature of jets and the alignment of the momentum imbalance from the dark particles with the jet axis, but such jets do not yet benefit from the construction of dedicated theoretically-motivated jet substructure observables. A deep network operating on jet constituents is used as a probe of the available information and indicates that classification power not captured by current high-level observables arises primarily from low- p_T jet constituents

Measurement / 9

Moment Unfolding using Deep Learning

Author: Krish Desai^{None}

Co-authors: Ben Nachman¹; Jesse Thaler²

¹ *Lawrence Berkeley National Lab. (US)*

² *MIT*

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

Deconvolving ('unfolding') detector distortions is a critical step in the comparison of cross section measurements with theoretical predictions. However, most of these approaches require binning while many predictions are at the level of moments. We develop a new approach to directly unfold distribution moments as a function of any other observables without having to first discretize. Our Moment Unfolding technique uses machine learning and is inspired by Generative Adversarial Networks (GANs). We demonstrate the performance of this approach using jet substructure measurements in collider physics. We also discuss challenges with unfolding all moments simultaneously, drawing connections to the renormalization of the partition function.

Anomaly Detection / 10

Results from Unsupervised Machine Learning in an ATLAS Dijet Resonance Search

Author: Julia Lynne Gonski¹

¹ *Columbia University (US)*

Corresponding Author: julia.gonski@cern.ch

An application of unsupervised machine learning-based anomaly detection to a generic dijet resonance is presented using the full LHC Run 2 dataset collected by ATLAS. A novel variational recurrent neural network (VRNN) is trained over data, specifically large-radius jets that are modeled using a sequence of constituent four-vectors and substructure variables, to identify anomalous jets based on their energy deposition pattern. The VRNN produces a per-jet anomaly score, whose performance is evaluated across a wide variety of hadronic topologies. This score is used to define a model-independent signal region in a search for new particles Y and X in association with a Higgs boson, representing the first application of unsupervised machine learning to an ATLAS analysis. A selection on the anomaly score of the X jet is shown to yield between 5-30% increase in significance across a variety of potential decays, and a comparison of the cross section upper limit on a variety of X hypotheses shows that the anomaly score provides competitive and broad sensitivity compared to traditional high-level variables.

Classification / 11**Truth tagging for efficiency parametrization of b-jets using Graph Neural Networks****Author:** Krunal Bipin Gedia¹¹ *ETH Zurich (CH)***Corresponding Author:** krunal.gedia@cern.ch

In high-energy physics experiments, estimating the efficiency of a process using selection cuts is a widely used technique. However, this method is limited by the number of events that could be simulated in the required analysis phase space. A way to improve this sensitivity is to use efficiency weights instead of selecting events by selection cuts. This method of efficiency measurements is called Truth tagging. In this talk, we propose a GNN-based approach for Truth-tagging which provides efficiency estimates parameterized in the multi-dimensional phase for b-tagging classifiers in CMS as firstly studied in arXiv:2004.02665.

Anomaly Detection / 12**Challenges for unsupervised anomaly detection in particle physics****Author:** Katherine Fraser¹¹ *Harvard University***Corresponding Author:** kfraser@g.harvard.edu

I discuss several approaches to anomaly detection in collider physics, including using variational autoencoders, which rely on the ability to reconstruct certain types of data (background) but not others (signals), and optimal transport distances, which which measures how easily one pT distribution can be changed into another. I discuss advantages and challenges associated with each approach. I also discuss a connection we uncovered between the latent space of a variational autoencoder trained using mean squared error and using optimal transport distances within the dataset.

Measurement / 13**Estimating Uncertainties for Trained Neural Networks****Author:** Sebastian Guido Bieringer¹**Co-authors:** Gregor Kasieczka²; Mathias Trabs³; Maximilian Steffen⁴¹ *Hamburg University*² *Hamburg University (DE)*³ *Karlsruhe Institute of Technology*⁴ *Karlsruhe Institute of Technology***Corresponding Authors:** trabs@kit.edu, maximilian.steffen@kit.edu, sebastian.bieringer@desy.de, gregor.kasieczka@cern.ch

Uncertainty estimation is a crucial issue when considering the application of deep neural network to problems in high energy physics such as jet energy calibrations.

We introduce and benchmark a novel algorithm that quantifies uncertainties by Monte Carlo sampling from the models Gibbs posterior distribution. Unlike the established ‘Bayes By Backpropagation’ training regime, it does not rely on any approximations of the network weight posterior, is flexible to most training regimes, and can be applied after training to any network. For a one-dimensional regression task, we show that this novel algorithm describes epistemic uncertainties well, including large errors for extrapolation.

Measurement / 16

Invertible Networks for the Matrix Element Method

Authors: Anja Butter^{None}; Theo Heimel¹; Till Martini²; Sascha Peitzsch³; Tilman Plehn^{None}

¹ *Heidelberg University*

² *HU Berlin*

³ *Humboldt-Universität zu Berlin*

Corresponding Authors: martinit@physik.hu-berlin.de, heimel@thphys.uni-heidelberg.de, butter@thphys.uni-heidelberg.de, sascha.peitzsch@gmail.com, plehn@uni-heidelberg.de

The matrix element method is widely considered the perfect approach to LHC inference, but computationally expensive. We show how a combination of two conditional Invertible Neural Networks can be used to learn the transfer function between parton level and reconstructed objects, and to make integrating out the partonic phase space numerically tractable. We illustrate our approach for the CP-violating phase of the top Yukawa coupling in associated Higgs and single-top production.

Generative Models – Particle Level / 17

MadNIS: Neural networks for multi-channel integration

Authors: Anja Butter^{None}; Claudius Krause¹; Fabio Maltoni²; Joshua Isaacson^{None}; Olivier Mattelaer³; Ramon Winterhalder⁴; Theo Heimel⁵; Tilman Plehn^{None}

¹ *Rutgers University*

² *Universite Catholique de Louvain (UCL) (BE) and Università di Bologna*

³ *UCLouvain*

⁴ *UC Louvain*

⁵ *Heidelberg University*

Corresponding Authors: butter@thphys.uni-heidelberg.de, olivier.mattelaer@uclouvain.be, ramon.winterhalder@uclouvain.be, fabio.maltoni@cern.ch, claudius.krause@rutgers.edu, isaacson@fnal.gov, heimel@thphys.uni-heidelberg.de, plehn@uni-heidelberg.de

High-precision theory predictions require the numerical integration of high-dimensional phase-space integrals and the simultaneous generation of unweighted events to feed the full simulation chain and subsequent analyses. While current methods are based on first principles and are mathematically guaranteed to converge to the correct answer, the computational cost to decrease the numerical error to a sub-percent level is enormous. Therefore, we combine current methods with fast and flexible machine-learning algorithms. In detail, we use a conditional normalizing flow that extends and generalizes the idea of i-flow, as well as machine-learned multi-channel weights to reduce the Monte Carlo error. Additionally, we employ a two-stage training procedure that reuses previously generated samples to reduce the number of potentially expensive integrand evaluations.

Anomaly Detection / 18**Optimal Mass Variables for Semivisible Jets****Author:** Kevin Pedro¹¹ *Fermi National Accelerator Lab. (US)***Corresponding Author:** pedrok@cern.ch

We apply the artificial event variable technique, a deep neural network with an information bottleneck, to strongly coupled hidden sector models. These models of physics beyond the standard model predict collider production of invisible, composite dark matter candidates mixed with regular hadrons in the form of semivisible jets. We explore different resonant production mechanisms to determine in which cases the machine learning approach provides an advantage over classical mass reconstruction. The results show that this technique is quite general and can be successfully applied even to very complicated physical models. We further demonstrate the viability of conducting an actual search for new physics using this method.

Generative Models – Detector Level / 19**IEA-GAN: Intra-Event Aware GAN with Relational Reasoning for the Fast Detector Simulation****Author:** Hosein Hashemi¹**Co-authors:** James Kahn²; Sahand Sharifzadeh³; Thomas Kuhr⁴¹ *LMU Munich*² *Helmholtz AI, Karlsruhe Institute of Technology (KIT)*³ *DeepMind*⁴ *Ludwig Maximilians Universitat (DE)***Corresponding Authors:** nikolai.hartmann@physik.uni-muenchen.de, thomas.kuhr@cern.ch, james.kahn@kit.edu, sharifzadeh@deepmind.com, gh.hashemi@physik.uni-muenchen.de

A realistic detector simulation is an essential component of experimental particle physics. However, it is currently very inefficient computationally since large amounts of resources are required to produce, store, and distribute simulation data. Deep generative models allow for more cost-efficient and faster simulations. Nevertheless, generating detector responses is a highly non-trivial task as they carry fine-grained information and have correlated mutual properties within an “event”, a single readout window after the collision of particles. Thus, we propose the Intra-Event Aware GAN (IEA-GAN) and demonstrate its use in generating sensor-dependent images for the pixel vertex detector (PXD), the sub-detector with the highest spatial resolution at the Belle II Experiment. First, we show that using the domain-specific relational inductive bias introduced by our novel Relational Reasoning Module; one can approximate the concept of an event in the detector simulation. Second, we propose a Uniformity loss in order to maximize the information entropy of the discriminator’s knowledge. Lastly, we develop an Intra-Event Aware loss for the generator to imitate the discriminator’s dyadic class-to-class knowledge. As a result, we show that the IEA-GAN not only captures fine-grained semantic and statistical similarity among the images but also finds correlations among them. Consequently, It leads to a significant enhancement in image fidelity and diversity compared to the previous state-of-the-art models.

Reconstruction / 20**A boosted kNN regressor with 66 million parameters**

Author: Tommaso Dorigo¹

¹ *Universita e INFN, Padova (IT)*

Corresponding Author: tommaso.dorigo@cern.ch

We develop a nearest neighbor algorithm for regressor for the problem of estimating the energy of multi-TeV muons in a high-granularity calorimeter, exploiting the pattern of soft photon deposits around the muon track. The algorithm is heavily overparametrized by assigning weights and biases to the training events. Parameters are learnt by batch gradient descent. The performance compares favourably with that of xgboost and a neural network, although CPU consumption is orders of magnitude larger.

Interpretability / 21

Blueprints for Training Information Bottlenecks for Collider Analyses

Authors: Kevin Pedro¹; Prasanth Shyamsundar²

¹ *Fermi National Accelerator Lab. (US)*

² *Fermi National Accelerator Laboratory*

Corresponding Authors: pedrok@cern.ch, prasanth@fnal.gov

Dimensionality reduction is a crucial aspect of data analysis in high energy physics, even if accompanied by information loss. Several methods, including histogram- and kernel-based analyses, are only computationally feasible for low-dimensional data. Furthermore, simulation models used in HEP can often only be validated for low-dimensional data. We provide several blueprints for using machine learning to create low-dimensional data representations (continuous event variables and discrete classification labels) for use in signal discovery and parameter estimation tasks. We also describe how to design the learned representation to facilitate a) searches with unknown model parameters and b) validation of simulation models in data control regions.

Equivariance and New Architectures / 22

Topological Data Analysis for Collider Events

Author: Tianji Cai¹

Co-authors: Ben Nachman²; Ian Dyckes²; Junyi Cheng

¹ *University of California, Santa Barbara*

² *Lawrence Berkeley National Lab. (US)*

Corresponding Authors: junyi_cheng@ucsb.edu, tianji_cai@ucsb.edu, benjamin.philip.nachman@cern.ch, george.dyckes@cern.ch

We introduce a novel framework to capture the inherent topological structure of collider events. Using persistence homology, the evolution of various topological features across scales is recorded graphically in a persistence diagram, and further encoded as scalars and vectors amenable to machine learning classifiers, showing excellent performance on both jet tagging and event classification tasks. We further propose a way to metricize the space of persistence diagrams by means of linearized optimal transport, which offers a new representation especially suited for topologically more challenging datasets. These topological taggers are inherently invariant to certain transformations of the underlying datasets, thus eliminating the need to preprocess jets and events in an *ad hoc* fashion. This constitutes another major advantage of the Topological Data Analysis framework applied to collider physics.

Classification / 23

Adversarial training for b-tagging algorithms in CMS

Authors: Annika Stein¹; CMS Collaboration²¹ *Rheinisch Westfaelische Tech. Hoch. (DE)*² *CMS Experiment, CERN***Corresponding Author:** annika.stein@cern.ch

Deep learning is a standard tool in high-energy physics, facilitating identification of physics objects. In particular, complex neural network architectures play a major role for jet flavor tagging. However, these methods are reliant on accurate simulations and a calibration is required to treat non-negligible performance differences with respect to data. In order to reduce residual disagreement between these two domains, adversarial methods are applied to close the generalization gap and to improve a classifier's robustness. Extensive studies have been carried out on a publicly accessible dataset with a fully connected neural network. Studying the impact of adversarial attacks on the inputs mimics the effect of systematic uncertainties. To mitigate this effect, an enhanced algorithm that adapts adversarial training is presented. We show that this strategy can also be applied to the DeepJet algorithm, a commonly used tagger at the CMS experiment. Due to the large number of inputs, small differences can add up to a considerable impact on performance. Utilizing the interplay of frameworks for sample creation, training, evaluation and scale factor derivation, we show that this mitigation strategy successfully improves agreement between data and simulated samples while maintaining a high performance. Thus, the introduction of the adversarial module is envisaged to become a useful ingredient for the upcoming generation of flavor tagging algorithms developed for Run3.

Reconstruction / 24

Optimal transport solutions for pileup mitigation at hadron colliders

Authors: Benedikt Maier¹; Fabio Iemmi²; Huilin Qu¹; Loukas Gouskos¹; Sascha Pascal Liechti³; Vinicius Massami Mikuni⁴¹ *CERN*² *Chinese Academy of Sciences (CN)*³ *University of Zurich (CH)*⁴ *Lawrence Berkeley National Lab. (US)***Corresponding Authors:** vinicius.massami.mikuni@cern.ch, sascha.liechti@cern.ch, fabio.iemmi@cern.ch, loukas.gouskos@cern.ch, huilin.qu@cern.ch, benedikt.maier@cern.ch

Hadronic jets and missing transverse energy are key experimental probes when searching for new physics or performing standard model precision measurements in collision events at the LHC. In this work, we propose a graph neural network algorithm for obtaining a global event description that demonstrates greatly improved resolution in the aforementioned objects obtained with a fast simulation of the CMS detector and reconstruction. This is achieved through a novel approach employing metrics inspired by optimal transport problems as the cost function of the neural network. By learning the difference between two particle collections, one containing only the event from the hard scattering and one containing additional products from secondary proton collisions, our network is able to reject contributions from pileup more effectively than other algorithms, which translates into a better resolution for observables related to jets or missing transverse energy. The implementation of such an algorithm would lead to a quasi-global improvement for analyses performed on proton collision data and would become crucial in the high-pileup scenario expected at the High Luminosity LHC.

Reconstruction / 25**Graph Neural Networks for a Deep-learning based Full Event Interpretation (DFEI) at the LHCb trigger**

Authors: Andrea Mauri¹; Jonas Eschle²; Julian Garcia Pardinas³; Marta Calvi⁴; Nicola Serra²; Simone Meloni^{None}

¹ *Nikhef National institute for subatomic physics (NL)*

² *University of Zurich (CH)*

³ *Universita & INFN, Milano-Bicocca (IT)*

⁴ *Univ. degli Studi Milano-Bicocca*

Corresponding Authors: sim.meloni@gmail.com, nicola.serra@cern.ch, julian.garcia.pardinas@cern.ch, marta.calvi@cern.ch, jonas.eschle@cern.ch, a.mauri@cern.ch

In a decade from now, the Upgrade II of LHCb experiment will face an instantaneous luminosity ten times higher than in the current Run 3 conditions. This will bring LHCb to a new era, with huge event sizes and typically several signal heavy-hadron decays per event. The trigger scope will shift from deciding ‘which events are interesting?’ to ‘which parts of the event are interesting?’. To allow for an inclusive, automatic and accurate multi-signal selection per event, we propose evolving from the current signal-based trigger to a Deep-learning based Full Event Interpretation (DFEI) approach. We have designed the first prototype for the DFEI algorithm, leveraging the power of Graph Neural Networks (GNN). The algorithm takes as input the final-state particles and has a two-folded goal: select the sub-set of particles originated in heavy-hadron decays, and reconstruct the decay chains in which they were produced. In this talk, we describe the design and development of this prototype, and discuss the latest performance studies on simulation, as well as the requirements for an eventual integration in the Real Time Analysis (RTA) system of LHCb.

Anomaly Detection / 26**Neural Embedding: Learning the Embedding of the Manifold of Physics Data**

Author: Sang Eon Park¹

Co-authors: Bryan Ostdiek²; Philip Coleman Harris¹

¹ *Massachusetts Inst. of Technology (US)*

² *Harvard University*

Corresponding Authors: bostdiek@gmail.com, sang.eon.park@cern.ch, philip.coleman.harris@cern.ch

There is a growing recent interest in endowing the space of collider events with a metric structure calculated directly in the space of its inputs. For quarks and gluons, the recently developed energy mover’s distance has allowed for a quantification of what is different between physical events. However, the large number of particles within jets makes using metrics and interpreting these metrics particularly difficult. In this work, we introduce a flexible framework based on neural embedding to embed a manifold from a jet to lower-dimensional spaces using a defined metric. We demonstrate a low distortion and robust embedding can be achieved with Energy movers distance in two dimensions. Furthermore, we show that we can construct a self-organized space that captures the core physical features of a jet, including the splitting angularity and the number of prongs. Using the notion of volume in the embedded space, we propose the volume-adjusted roc-curve to measure the energy mover’s volume that a dedicated jet selection has on the total phase space of jets. Finally, we equate the volume to the inclusivity of a jet kinematic selection and show how this approach can quantify the effectiveness of anomaly searches and measurements in performing unbiased, inclusive measurements.

Equivariance and New Architectures / 27**Does Lorentz-symmetric design boost network performance in jet physics?**

Authors: Congqiao Li¹; Huilin Qu²; Sitian Qian¹; Qi Meng³; Shiqi Gong⁴; Jue Zhang³; Tie-Yan Liu³; Qiang Li¹

¹ *Peking University (CN)*

² *CERN*

³ *Microsoft Research Asia*

⁴ *Academy of Mathematics and Systems Science, Chinese Academy of Sciences*

Corresponding Authors: stjian@pku.edu.cn, jue.zhang@microsoft.com, gongshiqi15@mails.ucas.ac.cn, qi.meng@microsoft.com, qiang.li@cern.ch, congqiao.li@cern.ch, huilin.qu@cern.ch

In the deep learning era, improving the neural network performance in jet physics is a rewarding task as it directly contributes to more accurate physics measurements at the LHC. Recent research has proposed various network designs in consideration of the full Lorentz symmetry, but its benefit is still not systematically asserted, given that there remain many successful networks without taking it into account. We conduct a detailed study on the Lorentz-symmetric design. We propose two generalized approaches for modifying a network - these methods are experimented on PFN, ParticleNet, and LorentzNet, and exhibit a general performance gain. We also reveal for the first time that the underlying reason behind the known improvement the “pairwise mass” feature brings to the network is that it introduces a structure that adheres to full Lorentz symmetry. We confirm that Lorentz-symmetry preservation serves as a strong inductive bias of jet physics, hence calling for attention to such general recipes in future network designs.

This talk is based on <https://arxiv.org/abs/2208.07814> and includes new relevant studies.

Classification / 28**Constituent-Based Top-Quark Tagging with the ATLAS Detector**

Author: Kevin Thomas Greif¹

¹ *University of California Irvine (US)*

Corresponding Author: kgreif@uci.edu

This talk will present the performance of constituent-based jet taggers on large radius boosted top quark jets reconstructed from optimized jet input objects in simulated collisions at $\sqrt{s}=13$ TeV. Several taggers which consider all the information contained in the kinematic information of the jet constituents are tested, and compared to a tagger which relies on high-level summary quantities similar to the taggers used by ATLAS in Runs 1 and 2. Several constituent based taggers are found to out-perform the high level quantity based tagger, with the best achieving a factor of two increase in background rejection across the kinematic range. To enable further development and study, the data set described in this note is made publicly available.

Generative Models – Detector Level / 29**Generative Models for Fast Simulation of Electromagnetic and Hadronic Showers in Highly Granular Calorimeters**

Authors: Anatolii Korol¹; Engin Eren^{None}; Erik Buhmann²; Frank-Dieter Gaede³; Gregor Kasieczka²; Katja Kruger³; Lennart Rustige⁴; Peter McKeown^{None}; Sascha Daniel Diefenbacher²; William Korcari²

¹ *Centre National de la Recherche Scientifique (FR)*

² *Hamburg University (DE)*

³ *Deutsches Elektronen-Synchrotron (DE)*

⁴ *Université Clermont Auvergne (FR)*

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

Simulation in High Energy Physics (HEP) places a heavy burden on the available computing resources and is expected to become a major bottleneck for the upcoming high luminosity phase of the LHC and for future Higgs factories, motivating a concerted effort to develop computationally efficient solutions. Methods based on generative machine learning methods hold promise to alleviate the computational strain produced by simulation while providing the physical accuracy required of a surrogate simulator.

In this contribution, an overview of a growing body of work focused on simulating showers in highly granular calorimeters will be reported, which is making significant steps towards realistic fast simulation tools based on deep generative models. Progress on the simulation of both electromagnetic and hadronic showers will be presented, with a focus on the high degree of physical fidelity and computational performance achieved. Additional steps taken to address the challenges faced when broadening the scope of these simulators, such as those posed by multi-parameter conditioning, will also be discussed.

Equivariance and New Architectures / 30

Transformer models for heavy flavor jet identification in CMS

Authors: Alexandre De Moor¹; Congqiao Li²; Denise Muller¹; Huilin Qu³; Sitian Qian²

¹ *Vrije Universiteit Brussel (BE)*

² *Peking University (CN)*

³ *CERN*

Corresponding Authors: alexandre.de.moor@vub.be, huilin.qu@cern.ch, stjian@pku.edu.cn, congqiao.li@cern.ch, denise.muller@cern.ch

During Run2 of the Large Hadron Collider (LHC), deep-learning-based algorithms were established and led to a significantly improved heavy flavor (b and c) jet tagging performance. In the scope of large-radius boosted jets like top-quark jets, Graph Neural Network (GNN) based models, e.g. ParticleNet, have reached state-of-the-art performance. As a step further, we present Particle Transformer (ParT), a new algorithm that incorporates physics-inspired interactions in an augmented self-attention mechanism. We show that ParT substantially improves the heavy flavor jet tagging performance compared to the state-of-the-art DeepJet algorithm. ParT is therefore a promising algorithm to be used for heavy flavor jet identification during Run3 of LHC.

Anomaly Detection / 31

HEP-Sim2Real: creating background templates with normalizing flows

Authors: Ben Nachman¹; Radha Mastandrea²; Samuel Byrne Klein³; Tobias Golling³

¹ *Lawrence Berkeley National Lab. (US)*

² *University of California, Berkeley*

³ *Universite de Geneve (CH)*

Corresponding Authors: benjamin.philip.nachman@cern.ch, tobias.golling@unige.ch, rmastand@berkeley.edu, samuel.byrne.klein@cern.ch

Resonant anomaly detection is a promising framework for model-independent searches for new particles. Weakly supervised resonant anomaly detection methods compare data with a potential signal against a template of the Standard Model (SM) background inferred from sideband regions. We propose a means to generate this background template that uses a normalizing flow to create a mapping between high-fidelity SM simulations and the data. The flow is trained in sideband regions with the signal region blinded, and the flow is conditioned on the resonant feature (mass) such that it can be interpolated into the signal region. To illustrate this approach, we use simulated collisions from the Large Hadron Collider (LHC) Olympics Dataset. We find that our flow-constructed background method has competitive sensitivity with other recent proposals and can therefore provide complementary information to improve future searches.

Classification / 32

Machine learning based jet and event classification at the Electron-Ion Collider

Author: James Mulligan¹

¹ *University of California, Berkeley (US)*

Corresponding Author: james.mulligan@berkeley.edu

In this talk, we explore machine learning-based event and jet identification at the future Electron-Ion Collider (EIC). We study the effectiveness of machine learning-based classifiers at the relatively low EIC energies, focusing on (i) identifying the flavor of the jet, in terms of both quark flavor tagging and quark vs. gluon tagging, and (ii) identifying the hard-scattering process, using full event information instead of using only information associated with the identified jet. We establish first benchmarks and contrast the performance of flavor tagging at the EIC with that at the LHC. We will discuss applications of these machine learning-based taggers in the key research areas at the future EIC, including the extraction of (transverse momentum dependent) parton distribution functions, studies of hadronization, and quantifying the modification of hadrons and jets in the cold nuclear matter environment in electron-nucleus collisions. Moreover, we outline how machine learning techniques can help to improve experimental access to transverse spin asymmetries in current experiments at the Relativistic Heavy Ion Collider (RHIC) and the future EIC.

Anomaly Detection / 33

CURTAINS for your Sliding Window: Constructing Unobserved Regions by Transporting Adjacent Intervals

Authors: Debajyoti Sengupta¹; Johnny Raine¹; Samuel Byrne Klein¹; Tobias Golling¹

¹ *Universite de Geneve (CH)*

Corresponding Authors: samuel.byrne.klein@cern.ch, johnny.raine@cern.ch, debajyoti.sengupta@cern.ch, tobias.golling@unige.ch

We introduce a new model independent technique for constructing background data templates for use in searches for new physics processes at the LHC.

This method, called CURTAINS, uses invertible neural networks to parametrise the distribution of side band data as a function of the resonant observable. The network learns a transformation to map any data point from its value of the resonant observable to another chosen value.

We demonstrate the performance on CURTAINS at anomaly detection on the LHC Olympics R&D dataset, hunting for a dijet resonance, by transforming data in sidebands into the signal region.

We will also present the latest developments and improvements to the method.

Preprint: <https://arxiv.org/abs/2203.09470>

Classification / 34

Robust Signal Detection using a Classifier Decorrelated through Optimal Transport (CDOT)

Authors: Purvasha Chakravarti¹; Mikael Kuusela²; Larry Wasserman³

¹ *University College London*

² *Carnegie Mellon University (US)*

³ *Carnegie Mellon University*

Corresponding Authors: mikael.kuusela@cern.ch, larrywasserman.cool@gmail.com, purvashac@gmail.com

New physics searches are usually done by training a supervised classifier to separate a signal model from a background model. However, even when the signal model is correct, systematic errors in the background model can influence supervised classifiers and might adversely affect the signal detection procedure. To tackle this problem, one approach is to find a classifier constrained to be decorrelated with one or more protected variables, e.g. the invariant mass. We do this by considering an optimal transport map of the classifier output that makes it independent of the invariant mass for the background. We then fit a semi-parametric mixture model to the invariant mass for different cuts on the transformed classifier to detect the presence of signal. We compare and contrast this decorrelation method with previous approaches, show that the decorrelation procedure is robust to background misspecification, and analyze the power of the test that simultaneously fits multiple classifier output bins.

Interpretability / 35

Neural Estimation of Energy Movers Distance

Authors: J Michael Williams¹; Ouail Kitouni¹

¹ *Massachusetts Inst. of Technology (US)*

Corresponding Authors: ouail.kitouni@cern.ch, mwill@mit.edu

We propose a novel neural architecture that enforces an exact upper bound on the Lipschitz constant of the model by constraining the norm of its weights. This architecture was useful in developing new algorithms for the LHCb trigger which have robustness guarantees as well as powerful inductive biases leveraging the neural network's ability to be monotonic in any subset of features. A new and interesting direction for this architecture is that it can also be used in the estimation of the Wasserstein metric (or the Earth Mover's Distance) in optimal transport using the Kantorovich-Rubinstein duality. We describe how such architectures can be leveraged for developing new clustering algorithms

using the Energy Mover's Distance. Clustering using optimal transport generalizes all previous well-known clustering algorithms in HEP (anti-kt, Cambridge-Aachen, *etc.*) to arbitrary geometries and offers new flexibility in dealing with effects such as pile-up and unconventional topologies.

Classification / 36

VBF vs. GGF Higgs with Full-Event Deep Learning: Towards a Decay-Agnostic Tagger

Author: Cheng-Wei Chiang¹

Co-authors: David Shih ; Shang-Fu Wei ¹

¹ *National Taiwan University*

Corresponding Authors: b06202061@ntu.edu.tw, dshih@physics.rutgers.edu, chengwei@phys.ntu.edu.tw

We study the benefits of jet- and event-level deep learning methods in distinguishing vector boson fusion (VBF) from gluon-gluon fusion (GGF) Higgs production at the LHC. We show that a variety of classifiers (CNNs, attention-based networks) trained on the complete low-level inputs of the full event achieve significant performance gains over shallow machine learning methods (BDTs) trained on jet kinematics and jet shapes, and we elucidate the reasons for these performance gains. Finally, we take initial steps towards the possibility of a VBF vs. GGF tagger that is agnostic to the Higgs decay mode, by demonstrating that the performance of our event-level CNN does not change when the Higgs decay products are removed. These results highlight the potentially powerful benefits of event-level deep learning at the LHC.

Generative Models – Detector Level / 37

CaloMan: Fast generation of calorimeter showers with density estimation on learned manifolds

Authors: Anthony Caterini¹; Brendan Ross¹; Gabriel Loaiza-Ganem¹; Humberto Reyes-González²; Jesse Cresswell¹; Marco Letizia^{None}

¹ *Layer 6 AI*

² *University of Genoa*

Corresponding Authors: jesse@layer6.ai, gabriel@layer6.ai, anthony@layer6.ai, marco.letizia@edu.unige.it, brenndan@layer6.ai, harg.zepelin@gmail.com

The efficient simulation of particle propagation and interaction within the detectors of the Large Hadron Collider is of primary importance for precision measurements and new physics searches. The most computationally expensive simulations involve calorimeter showers, which will become ever more costly and high-dimensional as the Large Hadron Collider moves into its High Luminosity era. Advances in deep generative modelling have opened the possibility of creating models that can generate realistic calorimeter showers orders of magnitude more quickly than physics-based simulation. Deep generative models have recently made stunning advances in modelling high-dimensional data like images and audio, however, the high-dimensional nature of calorimeter data belies the relative simplicity of the underlying physical laws which govern shower processes. In machine learning this relates to the manifold hypothesis which states that high-dimensional data is supported on low dimensional manifolds. We propose modelling calorimeter showers by first learning their manifold structure, then estimating the distribution of data on the manifold. Learning manifold structure reduces the dimensionality of the data, which enables fast training and generation. For our proof of concept, we model the datasets provided by the Fast Calorimeter Simulation Challenge 2022.

Classification / 38

Recent ML-usage in searches with boosted jets in CMS

Author: Oz Amram¹

¹ *Johns Hopkins University (US)*

Corresponding Author: oz.amram@cern.ch

CMS has a wide search program making use of ML for jet tagging and event reconstruction. This talk will report recent usage of ML in searches for heavy resonances involving boosted W, Z, H and top quark jets.

Interpretability / 39

Feature selection with Distance Correlation

Authors: David Shih^{None}; Gregor Kasieczka¹; RANIT DAS^{None}

¹ *Hamburg University (DE)*

Corresponding Authors: dshih@physics.rutgers.edu, gregor.kasieczka@cern.ch, rd804@scarletmail.rutgers.edu

Feature selection algorithms can be an important tool for AI explainability. If the performance of neural networks trained on low-level data can be reproduced by a small set of high-level features, we can hope to understand “what the machine learned”. We present a new algorithm that selects features by ranking their Distance Correlation (DisCo) values with truth labels. We apply this algorithm to the classification of boosted top quarks and use a set of 7,000 Energy Flow Polynomials (EFPs) as our feature space. We show that our method is able to select a small set of high-level features, with a classification performance comparable to the state-of-the-art top taggers.

Interpretability / 40

Can You Hear the Shape of a Jet?

Author: Rikab Gambhir¹

Co-authors: Akshunna Dogra²; Demba Ba³; Jesse Thaler¹

¹ *MIT*

² *Imperial College*

³ *Harvard*

Corresponding Authors: jthaler@mit.edu, adogra@nyu.edu, demba@seas.harvard.edu, rikab@mit.edu

The identification of interesting substructures within jets is an important tool to search for new physics and probe the Standard Model. In this talk, we present SHAPER, a general framework for defining computing shape-based observables, which generalizes the N -jettiness from point clusters to any extended shape. This is accomplished by minimizing the p -Wasserstein metric between events and parameterized manifolds of energy flows representing idealized shapes, implemented using the dual-potential Sinkhorn approximation for efficient minimization. We show how this geometric language of observables as manifolds can be used to easily define novel event and jet-substructure observables with built-in IRC safety that are useful for physics analyses, generalizing the notion of an event shape. We then demonstrate the SHAPER framework by performing example jet substructure analyses using these new shape-based observables.

Classification / 41**Search for dimuon events in IceCube using decision trees****Author:** Nakul Aggarwal¹¹ *University of Alberta***Corresponding Author:** naggarw@ualberta.ca

The dominant neutrino-nucleon interaction above 100 GeV is Deep Inelastic Scattering (DIS) in which an incoming neutrino scatters off a quark in the nucleon by exchanging a weak boson, producing an outgoing lepton accompanied by a hadron shower. Two sub-dominant processes are expected to produce two high energy charged leptons in the final state. The first one is a subset of DIS where a charmed meson is produced, which can decay into a charged lepton. The second one involves the exchange of a weak boson and a photon, resulting in a final state with two charged leptons and a neutrino, in a process known as neutrino trident production (NTP).

If an excess of these events is observed above the Standard Model (SM) prediction, it can serve as an indicator of Beyond Standard Model (BSM) physics. Since IceCube Neutrino Observatory has detected thousands of high energy neutrinos above 100 GeV and has collected over 10 years of data taking, it is an excellent candidate for their search. For the purposes of this work, we consider the channel where the outgoing leptons are muons. Since muons leave a track-like Cherenkov signature in IceCube, our central goal is to search for double-track events (from two muons or dimuons) and separate them from single track events (from a single muon). In this work, we perform this classification using decision trees.

Interpretability / 42**Jet tagging with deep sets of subsets****Authors:** Dimitrios Athanasakos^{None}; Felix Ringer¹; James Mulligan²; Andrew Larkoski³; Mateusz Ploskon⁴¹ *Stony Brook University*² *University of California, Berkeley (US)*³ *SLAC National Accelerator Laboratory*⁴ *Lawrence Berkeley National Lab. (US)***Corresponding Authors:** felix.ringer@stonybrook.edu, dimitrios.athanasakos@stonybrook.edu, mploskon@lbl.gov, james.mulligan@berkeley.edu, larkoski@slac.stanford.edu

We introduce a complete basis of subsets for machine learning-based jet tagging. The subsets are obtained with (i) a fixed radius or (ii) the clustering is performed until a fixed number of subsets is obtained.

For nonzero values of the subset radius, the resulting classifier is Infrared-Collinear (IRC) safe. By lowering the subset radius, we can increase the sensitivity to nonperturbative physics. In the limit of a vanishing subset radius, the exclusive subset basis approximates deep sets/particle flow networks (IRC unsafe). The basis introduced here is thus ideally suited to quantify the information content of jets at the boundary of perturbative vs. nonperturbative physics.

Reconstruction / 44**Point Cloud Deep Learning Methods for Pion Reconstruction in the ATLAS Experiment**

Author: Piyush Karande¹

¹ *Lawrence Livermore National Laboratory*

Corresponding Author: karande1@llnl.gov

The reconstruction and calibration of hadronic final states in the ATLAS detector present complex experimental challenges. For isolated pions in particular, classifying π^0 versus π^\pm and calibrating pion energy deposits in the ATLAS calorimeters are key steps in the hadronic reconstruction process. The baseline methods for local hadronic calibration were optimized early in the lifetime of the ATLAS experiment. Here we present a significant improvement over existing techniques using machine learning methods that do not require the input variables to be projected onto a fixed and regular grid. Instead, Transformer, Deep Sets, and Graph Neural Network architectures are used to process calorimeter clusters and particle tracks as point clouds, or a collection of data points representing a three-dimensional object in space. We demonstrate the performance of these new approaches as an important step towards a low-level hadronic reconstruction scheme that fully takes advantage of deep learning to improve its performance.

Generative Models – Particle Level / 45

Conditional generative networks for pure quark and gluon jets

Authors: Ayodele Ore^{None}; Matthew Dolan¹

¹ *University of Melbourne*

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

The separation of quarks and gluons is of key interest at hadron colliders. While it is only possible to obtain mixed samples of quark and gluon jets from experimental data, some recent works have proposed methods for disentangling the underlying distributions in an unsupervised manner. However, these approaches typically lack a generative model for the separated distributions. In this work we provide a framework based on conditional generative networks that is able to separate mixed samples of quark and gluon jets. We present results using normalising flows and generative adversarial networks and discuss how the models could be used to enhance quark/gluon classification at colliders.

Measurement / 46

Constraining quark and gluon jet energy loss distributions in quark-gluon plasma using Bayesian inference

Authors: Alexandre Falcão¹; Konrad Tywoniuk²; Mathias Nilsen¹

¹ *University of Bergen*

² *University of Bergen (NO)*

Corresponding Authors: konrad.tywoniuk@uib.no, alexandre.falcao@uib.no

QCD factorization allows us to model the jet energy-loss in A-A collisions as a convolution between the jet cross section in p-p collisions and an energy loss distribution. Meanwhile, Bayesian inference provides a data-driven way of constraining the energy loss distribution parameterization. Only a few efforts have been made in this direction, and solely using untagged jets. However, gluon and quark jets are known to lose energy differently. By discriminating them, we distinguish the energy loss distributions of each parton-jet and arrive at a different set of parameters for each. This allows for a

more universal model that can be used for prediction in other jet measurements where quark/gluon ratio is different. A form for the energy loss distribution is chosen in the soft scattering approximation and the Markov Chain Monte Carlo method is then employed to estimate the parameters. The jet suppression obtained from the extracted energy loss distribution for inclusive jets show good agreement with measured one. However, it is sensitive to the collision energy. This might be caused by a poor constraining power of only relying on inclusive jets. We show the improvement by including photon tagged jets to the analysis.

With this study, we hope to achieve a better and more constrained modeling of the jet energy loss distribution, as well as to retrieve insights on how current theoretical models can improve by adding more insight from measurements.

Equivariance and New Architectures / 47

Transformer Architectures for Quenched Jet Tagging

Authors: João Pedro de Arruda Gonçalves¹; Miguel Crispim Romao²; Guilherme Milhano³

¹ *LIP*

² *University of Southampton*

³ *LIP-Lisbon & CERN TH*

Corresponding Authors: guilherme.milhano@cern.ch, jgoncalves@lip.pt, m.crispim-romao@soton.ac.uk

In high-energy heavy-ion collisions, the unconfined state of partons known as the Quark Gluon Plasma (QGP), is known to suppress the yield of jets with respect to proton-proton collision, as well as modify the structure of jets that transverse it. Nonetheless, samples of heavy-ion jets, even at the highest centralities, will contain a significant fraction of jets that, for one reason or the other, were not significantly modified by the QGP. Our community is therefore in need of a jet by jet tagger of the quenching phenomena. Gearing towards this end, we propose the transformer architecture to tackle the problem. We have obtained excellent discrimination with respect to the current state of the art, showing that this architecture can indeed be a viable solution. Further studies should consider dealing with the underlying event and some approach close to CWoLa with jet topics, for ROC curves directly pertaining to actually quenched jets.

Interpretability / 48

Resilience of Quark-Gluon Tagging

Authors: Anja Butter¹; Barry Dillon¹; Lukas Fabrizio Klassen¹; Tilman Plehn¹; Lorenz Vogel¹

¹ *ITP, Heidelberg University*

Corresponding Authors: plehn@uni-heidelberg.de, l.f.klassen@thphys.uni-heidelberg.de, butter@thphys.uni-heidelberg.de, vogel@thphys.uni-heidelberg.de, dillon@thphys.uni-heidelberg.de

Discriminating quark-initiated from gluon-initiated jets is an extremely challenging yet important task in high-energy physics. Recent studies have shown that the discriminating features between quark and gluon jets produced by the Monte Carlo generator Pythia differ significantly from the features produced by Herwig. To understand this simulation-dependent discrepancy, we propose a Bayesian version of ParticleNet (a state-of-the-art graph neural network that treats jets as particle clouds). Our Bayesian ParticleNet (BPN) shows similar performance to the deterministic ParticleNet, while providing additional information about uncertainties. We use the uncertainty estimates provided by our Bayesian ParticleNet to study the resilience/robustness of quark-gluon tagging and assess the differences between Pythia and Herwig jets.

Reconstruction / 49**v-flows: Conditional neutrino momentum regression****Authors:** Johnny Raine¹; Matthew Leigh²; Tobias Golling¹¹ *Universite de Geneve (CH)*² *University of Geneva***Corresponding Authors:** tobias.golling@unige.ch, johnny.raine@cern.ch, matthew.leigh@cern.ch

We present v-Flows, a novel method for restricting the likelihood space of neutrino kinematics in high energy collider experiments using conditional normalizing flows and deep invertible neural networks.

This method allows the recovery of the full neutrino momentum, which is usually left as a free parameter, and permits one to sample neutrino values under a learned conditional likelihood given event observations.

We demonstrate the success of v-Flows in a case study by applying it to simulated semileptonic $t\bar{b}$ events and show that it can lead to more accurate momentum reconstruction, particularly of the longitudinal coordinate. We also show that this has direct benefits in a downstream task of jet association, leading to an improvement of up to a factor of 1.41 compared to conventional methods.

Reconstruction / 50**Particle reconstruction in jets with set transformer and hypergraph prediction architectures****Authors:** Francesco Armando Di Bello¹; Etienne Dreyer²; Nilotpal Kakati²; Sanmay Ganguly³; Eilam Gross²; Lukas Alexander Heinrich⁴; Anna Ivina²; Marumi Kado⁴; Lorenzo Santi⁵; Matteo Tusoni⁵¹ *INFN e Universita Genova (IT)*² *Weizmann Institute of Science (IL)*³ *University of Tokyo (JP)*⁴ *Max Planck Society (DE)*⁵ *Sapienza Universita e INFN, Roma I (IT)***Corresponding Authors:** marumi.kado@cern.ch, matteo.tusoni@cern.ch, l.santi@cern.ch, lukas.heinrich@cern.ch, anna.ivina@cern.ch, eilam.work@gmail.com, nilotpal.kakati@cern.ch, etienne.dreyer@cern.ch, sanmay.ganguly@cern.ch, francesco.armando.di.bello@cern.ch

Particle reconstruction is a task underlying virtually all analyses of collider-detector data. Recently, the application of deep learning algorithms on graph-structured low-level features has suggested new possibilities beyond the scope of traditional parametric approaches. In particular, we explore the possibility to reconstruct and classify individual neutral particles in a collimated environment by studying single-jet events in a realistic calorimeter simulation. We develop two novel algorithms which approach reconstruction as a set-to-set task between tracks and calorimeter clusters as input and final-state particles as output. Notably, an algorithm designed to predict hypergraph structure shows superior performance on particle and jet-level metrics –surpassing a parametric particle-flow baseline –and provides a high degree of interpretability.

Interpretability / 51

Weakly Supervised Learning for Muon Discrimination in Unlabeled Collider Data

Authors: Daniel Whiteson¹; Edmund Witkowski²

¹ *University of California Irvine (US)*

² *UCI*

Corresponding Authors: daniel.whiteson@cern.ch, witkowse@uci.edu

We use unlabeled collision data from CMS and weakly-supervised learning to train models which can distinguish prompt muons from non-prompt muons using patterns of low-level particle activity in vicinity of the muon, and interpret the models in the space of energy flow polynomials. Particle activity associated with muons is a valuable tool for identifying prompt muons, those due to heavy boson decay, from muons produced in the decay of heavy flavor jets. The high-dimensional information it typically reduced to a single scalar quantity, isolation, but previous work in simulated samples suggests that valuable discriminating information is lost in this reduction. We extend these studies in LHC collisions recorded by the CMS experiment, where true class labels are not available, requiring the use of the invariant mass spectrum to obtain macroscopic sample information. This allows us to employ Classification Without Labels (CWoLa), a weakly supervised learning technique, to train models. Our results confirm that isolation does not describe events as well as the full low-level calorimeter information, and allows us to interpret the resulting network in terms of energy flow polynomials.

Anomaly Detection / 52

A Normalized Autoencoder for LHC triggers

Authors: Barry Dillon¹; Friedrich Feiden^{None}; Luigi Favaro^{None}; Tanmoy Modak^{None}; Tilman Plehn^{None}

¹ *University of Heidelberg*

Corresponding Authors: luigi.favaro@stud.uni-heidelberg.de, plehn@uni-heidelberg.de, tanmoy@thphys.uni-heidelberg.de, dillon@thphys.uni-heidelberg.de

The main goal for the upcoming LHC runs is still to discover BSM physics. It will require analyses able to probe regions not linked to specific models but generally identified as beyond the Standard Model. Autoencoders are the typical choice for fast anomaly detection models. However, they have shown to misidentify anomalies of low complexity signals over background events. I will present an energy-based Autoencoder called Normalized AE, a density-based high-performance anomaly search algorithm. I will show NAE applications on jet tagging and on reconstructed events. In particular, I will discuss how NAE is able to symmetrically tag QCD and top jet images as well as the BSM events proposed for the Anomaly Detection Challenge 2021.

Generative Models – Particle Level / 53

Point Cloud Generation using Transformer Encoders and Normalising Flows

Author: Benno Kach¹

Co-authors: Dirk Kruecker²; Isabell Melzer-Pellmann¹

¹ *Deutsches Elektronen-Synchrotron (DE)*

² *Deutsches Elektronen-Synchrotron (DESY)-Unknown-Unknown*

Corresponding Authors: dirk.bernhard.kruecker@cern.ch, isabell.melzer-pellmann@cern.ch, benno.kaech@desy.de

Machine-learning-based data generation has become a major topic in particle physics, as the current Monte Carlo simulation approach is computationally challenging for future colliders, which will have a significantly higher luminosity. The generation of particles poses difficult problems similar as is the case for point clouds. We propose that a transformer setup is well fitted to this task. In this study, a novel refinement model is presented, which uses normalizing flows as a prior and then enhances the generated points using an adversarial setup with two Transformer encoder networks. Different training architectures and procedures were tested and compared on the jetnet datasets.

Measurement / 54

Machine learning for top physics in CMS

Authors: Jan Kieseler¹; Philip Daniel Keicher²

¹ CERN

² Hamburg University (DE)

Corresponding Authors: philip.daniel.keicher@cern.ch, jan.kieseler@cern.ch

Machine learning (ML) plays a significant role in the physics analyses at the CMS experiment. Many different techniques and strategies have been deployed to a wide range of applications. In this presentation we will illustrate the most advanced techniques used in top quark physics measurements, such as using ML algorithms to improve the extraction of effective field theory contributions, and to predict background shapes in the region that are hard to be covered by conventional methods. Potential future developments will be discussed too.

Beyond Jets / 55

Overview of ML for Gravitational Waves

Authors: Alec Gunny¹; Chia-Jui Chou²; Deep Chatterjee¹; Dylan Sheldon Rankin³; Eric Anton Moreno¹; Erik Katsavounidis¹; Ethan Jacob Marx¹; Hong-Yin Chen²; Li-Cheng Yang²; Michael Coughlin⁴; Muhammed Saleem Cholayil⁴; Philip Coleman Harris³; Ryan Raikman⁵; Will Benoit⁴

¹ Massachusetts Institute of Technology (US)

² National Yang Ming Chiao Tung University

³ Massachusetts Inst. of Technology (US)

⁴ University of Minnesota

⁵ Carnegie Mellon University

Corresponding Authors: emarx@mit.edu, mcholayi@umn.edu, dylan.sheldon.rankin@cern.ch, rraikman@mit.edu, philip.coleman.harris@cern.ch, cough052@umn.edu, alec.gunny@gmail.com, benoi090@umn.edu, deep1018@mit.edu, kats@ligo.mit.edu, eric.anton.moreno@cern.ch

At an increasing number of interferometer sites with constantly-changing detector conditions, AI can play an important role in real-time and offline data processing. In this talk, we develop novel algorithms and training schemes that sift through noise and instrumental glitches to detect gravitational waves (GW) from compact binary coalescences (CBCs). For real-time processing, we create custom low-latency pipelines and packages for time-series analysis including parallel processing of hardware accelerators, inference as a service (iaas), and non-linear noise regression using autoencoders. We improve the efficiency of ML idea-to-deployment using end-to-end model iteration, optimization, and analysis which can be trained/tested with full observation runs and mock data. Beyond CBCs, we establish source-agnostic anomaly detection algorithms using Transformers and

LSTMs to build embedded spaces that identify glitches and search for a variety of hypothesized astrophysical sources that may emit GWs in the LIGO frequency band including supernovae, neutron star glitches, and cosmic strings from the early universe. In presenting at ML4Jets, we hope to establish a bridge between the high energy and gravitational-wave communities, introducing our open data and frameworks under the A3D3/ML4GW organization that make time-series generation and analysis simple.

Generative Models – Detector Level / 56

Fast calorimeter simulation with VQVAE

Authors: Qibin Liu¹; Chase Owen Shimmin²; Xiulong Liu³; Eli Shlizerman³; Shih-Chieh Hsu⁴

¹ *Tsung-Dao Lee Institute (CN) & Shanghai Jiao Tong University (CN)*

² *Yale University (US)*

³ *University of Washington, Seattle*

⁴ *University of Washington Seattle (US)*

Corresponding Authors: liuxiulong1995@gmail.com, shlizee@gmail.com, qibin.liu@cern.ch, schsu@uw.edu, chase.shimmin@cern.ch

Simulation of calorimeter response is important for modern high energy physics experiments. With the increasingly large and high granularity design of calorimeters, the computational cost of conventional MC-based simulation of each particle-material interaction is becoming a major bottleneck. We propose a new generative model based on a Vector-Quantized Variational Autoencoder (VQ-VAE) to generate the calorimeter response. This model achieved a speedup of more than 5×10^4 times over GEANT4 on the CaloGAN dataset and the comparable performance of energy deposition and shower shape as existing ML-models such as CaloGAN and CaloFlow, with substantially fewer parameters and factor of 2 more speedup. We also demonstrate that the VQVAE approach can be adapted to a variety of encoder/decoder architectures, ranging from fully-connected to convolutional networks. The former is more suited to smaller, or irregular geometries, while the latter can perform well on very high granularity datasets with regular structure.

Beyond Jets / 57

GalaxyFlow: Upsampling Hydrodynamical Simulations for Realistic Gaia Mock Catalogs

Authors: Matthew Buckley^{None}; Sung Hak Lim¹; Kailash Anirudh Raman²; David Shih^{None}

¹ *Rutgers University*

² *Rutgers State Univ. of New Jersey (US)*

Corresponding Authors: sunghak.lim@rutgers.edu, kar352@physics.rutgers.edu, mbuckley@physics.rutgers.edu, dshih@physics.rutgers.edu

The Gaia DR3 catalog provides high-quality measurements of stars in the Milky Way, but the current N -body simulation-based mock Gaia catalogs have larger resolutions compared to those of the original Gaia dataset.

Because of that, using the mock catalogs to aid statistical analyses on the Gaia dataset in a small position/velocity resolution scale is very limited.

To solve this issue, we introduce an unsupervised deep-learning method for upsampling the star particles representing clusters of stars in N -body simulations.

Our method is based on a density estimation technique called normalizing flows for smoothly modeling the coarse star particle distribution so that it can be used for generating a more smooth synthetic star distribution.

We demonstrate our upsampling algorithm on a sub-dataset of two simulated galaxies, h277 and Auriga 6, and we compare the upsampling performance of our approach to conventionally used kernel smoothing algorithms such as EnBiD.

Also, using a machine-learned likelihood-ratio-based test, we show that the distribution of our synthetic stars is not only smooth but also kinematically consistent with that of the original star particles.

Equivariance and New Architectures / 58

Equivariant Neural Networks for Particle Physics: PELICAN

Authors: Alexander Bogatskiy¹; Timothy Hoffman^{None}; Xiaoyang Liu²; David Miller³; Jan Tuzlic Offermann³

¹ Flatiron Institute, Simons Foundation

² University of Chicago

³ University of Chicago (US)

Corresponding Authors: blackmail1807@gmail.com, xiaoyang1@uchicago.edu, hoffmant@uchicago.edu, jan.offermann@cern.ch, david.w.miller@cern.ch

A lot of attention has been paid to the applications of common machine learning methods in physics experiments and theory. However, much less attention is paid to the methods themselves and their viability as physics modeling tools. One of the most fundamental aspects of modeling physical phenomena is the identification of the symmetries that govern them. Incorporating symmetries into a model can make it more physically satisfactory, reduce the risk of over-parameterization, and consequently improve robustness and predictive power. As usage of neural networks continues to grow in the field of particle physics, more effort will need to be invested in narrowing the gap between the black-box models of ML and the analytic models of physics.

Building off of previous work, we demonstrate how careful choices in the details of network design – creating a model both simpler and more grounded in physics than the traditional approaches – can yield state-of-the-art performance within the context of problems including jet tagging and four-momentum reconstruction. We present the Permutation-Equivariant and Lorentz-Invariant or Covariant Aggregator Network (PELICAN), which is based on three key ideas: symmetry under permutations of particles, Lorentz symmetry, and the ambiguity of the aggregation process in Graph Neural Networks. For the first, we use the most general permutation-equivariant layer acting on square arrays, which can be viewed as a powerful generalization of Message Passing. For the second, we use classical theorems of Invariants Theory to reduce the 4-vector inputs to an array of Lorentz-invariant quantities. Finally, the flexibility of the aggregation process commonly used in Graph Networks can be leveraged for improved accuracy, in particular to allow variable scaling with the size of the input. We demonstrate the performance of this architecture on two problems: top tagging, and W momentum reconstruction.

Reconstruction / 59

Likelihood-Free Frequentist Inference for Calorimetric Muon Energy Measurement

Author: Luca Masserano¹

Co-authors: Ann Lee¹; Mikael Kuusela²; Rafael Izbicki³; Tommaso Dorigo⁴

¹ Carnegie Mellon University

² Carnegie Mellon University (US)

³ Federal University of São Carlos

⁴ *Universita e INFN, Padova (IT)*

Corresponding Authors: annlee@andrew.cmu.edu, tommaso.dorigo@cern.ch, masserano.luca@gmail.com, mikael.kuusela@cern.ch, rafaelizbicki@gmail.com

Calorimetric muon energy estimation in high-energy physics is an example of a likelihood-free inference (LFI) problem, where simulators that implicitly encode the likelihood function are used to mimic complex particle interactions at different values of the physical parameters. Recently, Kieseler et al. (2022) exploited simulated measurements from a dense, finely segmented calorimeter to infer the true energy of incoming muons and improve the resolution at high energies using a prediction approach based on a custom neural network architecture. Nonetheless, it remains an open question whether these tools produce reliable measures of uncertainty. In this work we present Waldo, a novel method to construct frequentist confidence sets within an LFI setting. WALDO reframes the well-known Wald test and uses Neyman inversion to convert point predictions or posterior distributions from any prediction algorithm or posterior estimator to confidence sets with correct conditional coverage, even for finite observed sample sizes. The LFI framework we exploit also allows to check empirical coverage across the entire parameter space. Finally, we demonstrate the effectiveness of Waldo by applying it to the muon energy estimation problem. Our results further support the work of Kieseler et al. (2022) who proposed this new avenue as an alternative to curvature-based measurements in a magnetic field.

Reconstruction / 60

Jet SIFT-ing

Author: Joel Walker¹

Co-authors: Andrew Larkoski²; Denis Rathjens³; Jason Robert Veatch⁴

¹ *Sam Houston State University*

² *SLAC National Accelerator Laboratory*

³ *Texas A & M University (US)*

⁴ *California State University (US)*

Corresponding Authors: denis.rathjens@cern.ch, larkoski@slac.stanford.edu, jwalker@shsu.edu, jason.veatch@cern.ch

We describe a new scale-invariant jet clustering algorithm which does not impose a fixed cone size on the event. The proposed construction unifies fat-jet finding, substructure axis-finding, and recursive filtering of soft wide-angle radiation into a single procedure. The sequential clustering measure history facilitates high-performance substructure tagging with a boosted decision tree. Excellent object discrimination is maintained for highly-boosted partonic systems, while asymptotically recovering favorable behaviors of both the standard KT anti-KT algorithms.

Anomaly Detection / 61

Weakly supervised methods for LHC analyses

Author: Thorben Finke^{None}

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

We investigate how weakly supervised methods like CWoLa and CATHODE can be used to enhance the sensitivity of searches at the LHC. These methods do not rely on truth level labels and are thus applicable in a model agnostic setting. In particular, we examine how these methods generalize to low level features, i.e. to higher dimensional inputs. As one example, we show how CWoLa can enhance the sensitivity of a monojet search at the LHC for models with modified jet dynamics.

Anomaly Detection / 62**Robust anomaly detection using NuRD****Authors:** Aahlad Puli^{None}; Abhijith Gandrakota¹; Lily Zhang²**Co-authors:** Jennifer Ngadiuba³; Nhan Tran¹¹ *Fermi National Accelerator Lab. (US)*² *New York University*³ *FNAL***Corresponding Authors:** jennifer.ngadiuba@cern.ch, abhijith.gandrakota@cern.ch, nhan.viet.tran@cern.ch, zhang, aahlad@nyu.edu

Anomaly Detection algorithms are crucial tools for identifying unusual decays from proton collisions at the LHC and are efficient methods for seeking out the possibility of new physics. These detection algorithms should be robust against nuisance kinematic variables and detector conditions. To achieve this robustness, popular detection models built via autoencoders, for example, have to go through a decorrelation stage, where the anomaly thresholds for the scores are decorrelated with the nuisances; this post-training procedure sacrifices detection accuracy. We propose a class of robust anomaly detection technique that accounts for nuisances in the prediction, called Nuisance-Randomized Distillation (NuRD). Our nuisance-aware anomaly detection methods we build with NuRD do not require the extra decorrelation step (and therefore do not suffer the associated accuracy loss).

Beyond Jets / 63**Measuring Galactic Dark Matter through Unsupervised Machine Learning****Author:** Matthew Buckley^{None}**Co-authors:** David Shih ; Eric Putney¹; Sung Hak Lim²¹ *Rutgers, The State University of New Jersey*² *Rutgers University***Corresponding Authors:** sunghak.lim@rutgers.edu, eputney@physics.rutgers.edu, mbuckley@physics.rutgers.edu, dshih@physics.rutgers.edu

Measuring the density profile of dark matter in the Solar neighborhood has important implications for both dark matter theory and experiment. In this work, we apply masked autoregressive flows to stars from a realistic simulation of a Milky Way-type galaxy to learn – in an unsupervised way – the stellar phase space density and its derivatives. With these as inputs we calculate the gravitational acceleration field and mass density directly from the Boltzmann Equation only assuming dynamic equilibrium. We demonstrate our approach can accurately reconstruct the mass density and acceleration profiles of the simulated galaxy, even in the presence of Gaia-like errors in the kinematic measurements. Finally, we discuss applications of this analysis to real astrometric data from Gaia Data Release 3.

Anomaly Detection / 64**Hunting for signals using Gaussian Process regression****Author:** Abhijith Gandrakota¹

Co-authors: Alexandre Morozov ; Amitabh Lath ²; Sindhu Murthy ³

¹ *Fermi National Accelerator Lab. (US)*

² *Rutgers, The State University of New Jersey*

³ *Carnegie-Mellon University (US)*

Corresponding Authors: sindhum@andrew.cmu.edu, morozova@physics.rutgers.edu, lath@physics.rutgers.edu, abhijith.gandrakota@cern.ch

We present a novel computational approach for extracting weak signals, whose exact location and width may be unknown, from complex background distributions with an arbitrary functional form. We focus on datasets that can be naturally presented as binned integer counts, demonstrating our approach on the datasets from the Large Hadron Collider. Our approach is based on Gaussian Process (GP) regression - a powerful and flexible machine learning technique that allowed us to model the background without specifying its functional form explicitly, and to separate the background and signal contributions in a robust and reproducible manner. Unlike functional fits, our GP-regression-based approach does not need to be constantly updated as more data becomes available. We discuss how to select the GP kernel type, considering trade-offs between kernel complexity and its ability to capture the features of the background distribution. We show that our GP framework can be used to detect the Higgs boson resonance in the data with more statistical significance than a polynomial fit specifically tailored to the dataset. Finally, we use Markov Chain Monte Carlo (MCMC) sampling to confirm the statistical significance of the extracted Higgs signature.

Equivariance and New Architectures / 65

A Holistic Approach to Predicting Top Quark Kinematic Properties with the Covariant Particle Transformer

Authors: Ben Nachman¹; Haichen Wang¹; Shikai Qiu¹; Shuo Han¹; Xiangyang Ju¹

¹ *Lawrence Berkeley National Lab. (US)*

Corresponding Authors: xiangyang.ju@cern.ch, haichen.wang@cern.ch, benjamin.philip.nachman@cern.ch, shikai.qiu@cern.ch, shuo.han@cern.ch

Precise reconstruction of top quark properties is a challenging task at the Large Hadron Collider due to combinatorial backgrounds and missing information. We introduce a physics-informed neural network architecture called the Covariant Particle Transformer (CPT) for directly predicting the top quark kinematic properties from reconstructed final state objects. This approach is permutation invariant and partially Lorentz covariant and can account for a variable number of input objects. In contrast to previous machine learning-based reconstruction methods, CPT is able to predict top quark four-momenta regardless of the jet multiplicity in the event. Using simulations, we show that the CPT performs favorably compared with other machine learning top quark reconstruction approaches.

Anomaly Detection / 66

Generative Models for Resonant Anomaly Detection

Authors: Ali Garabaglu¹; Ben Nachman²; Daniel Sun^{None}; David Shih^{None}; Elham E Khoda¹; Htet Aung Myin¹; Hui-Chi Lin³; Ines Ochoa⁴; Shih-Chieh Hsu⁵

¹ *University of Washington (US)*

² *Lawrence Berkeley National Lab. (US)*

³ *University of Michigan (US)*

⁴ *LIP - Laboratorio de Instrumentação e Física Experimental de Partículas (PT)*

⁵ *University of Washington Seattle (US)*

Corresponding Authors: ali.garabaglu@cern.ch, danielcsun@hotmail.com, dshih@physics.rutgers.edu, htet.aung.myin@cern.ch, elham.e.khoda@cern.ch, schsu@uw.edu, hui-chi.lin@cern.ch, ines.ochoa@cern.ch, benjamin.philip.nachman@cern.ch

Machine learning-based anomaly detection techniques offer exciting possibilities to significantly extend the search for new physics at the Large Hadron Collider (LHC) and elsewhere by reducing the model dependence. In this work, we focus on resonant anomaly detection, where generative models can be trained in sideband regions and interpolated into a signal region to provide an estimate of the Standard Model background. This estimate can then be compared with data using a machine learning classifier. This approach has been studied for normalizing flows and we explore the possibility of also using Variational Autoencoders (VAEs). We demonstrate this idea by conducting a di-jet resonance search using the LHC Olympics 2020 challenge dataset. The generative algorithm is trained by conditioning on the dijet invariant mass in the mass side-band, and so it can be evaluated in the signal region using a number of kinematic distributions for new resonance classification. The preliminary results are promising and open a new opportunity to address anomaly detection challenges at the LHC.

Interpretability / 68

Infra-red and collinear safe Graph Neural Networks

Authors: Michael Spannowsky¹; Partha Konar²; Vishal Singh Ngairangbam^{None}

¹ *University of Durham (GB)*

² *Physical Research Laboratory, Ahmedabad, Gujarat-380 009, INDIA*

Corresponding Authors: konar@prl.res.in, vishalngairangbam@gmail.com, michael.spannowsky@cern.ch

Hadronic signals of new-physics origin at the Large Hadron Collider can remain hidden within the copiously produced hadronic jets. Unveiling such signatures require highly performant deep-learning algorithms. We construct a class of Graph Neural Networks (GNN) in the message-passing formalism that makes the network output infra-red and collinear (IRC) safe, an important criterion satisfied within perturbative QCD calculations. Including IRC safety of the network output as a requirement in constructing the GNN improves its explainability and robustness against theoretical uncertainties in the data. We generalise Energy Flow Networks (EFN), an IRC-safe deep-learning algorithm on a point cloud, defining energy-weighted local and global readouts on GNNs. Applying the simplest of such networks to identify top quarks, W bosons and quark/gluon jets, we find that it outperforms state-of-the-art EFNs. Additionally, we obtain a general class of graph construction algorithms that give structurally invariant graphs in the IRC limit, a necessary criterion for the IRC safety of the GNN output.

Measurement / 69

Using Machine Learning to Improve our Understanding of the Jet Background in Nucleus-Nucleus Collisions.

Author: Tanner Mengel¹

¹ *University of Tennessee*

Corresponding Author: tmengel@vols.utk.edu

Jets in heavy ion collisions contain contributions from a background of soft-particles. The kinematic reach into low jet momentum is largely driven by the precision of the method used to subtract this

background. This precision is also a significant contribution to uncertainties of jet measurements. Previous studies have suggested that deep neural networks can improve momentum resolution at LHC energies when compared to the traditional area-based subtraction method. Applying deep neural networks to subtract background in Au+Au collisions at 200 GeV yields similar performance in low momentum jet resolution. This talk presents investigations into the relationship between corrected jet momentum and input jet feature-space which provide insight into the improved performance achieved by a deep neural network. These insights are used to develop a simplified neural network architecture and a background subtraction method based on jet track multiplicity. Both achieve similar performance to the deep neural network.

Generative Models – Detector Level / 70

CaloFlow for CaloChallenge

Authors: Claudius Krause¹; David Shih^{None}; Ian Pang²; Matthew Buckley^{None}; Yi En Ian Pang^{None}

¹ *Rutgers University*

² *Rutgers*

Corresponding Authors: ian.pang@physics.rutgers.edu, mbuckley@physics.rutgers.edu, dshih@physics.rutgers.edu, yp325@rutgers.edu, claudius.krause@rutgers.edu

Simulating particle detector response is the single most computationally expensive step in the Large Hadron Collider computational pipeline. Recently it was shown that normalizing flows can accelerate this process while achieving unprecedented levels of accuracy (CaloFlow).

Applying CaloFlow to the photon and charged pion GEANT4 showers of Dataset 1 of the Fast Calorimeter Simulation Challenge 2022, we are able to produce samples of high-fidelity with a sampling time less than 0.1ms per shower. We demonstrate the fidelity of the samples using calorimeter shower images, histograms of high level features, and aggregate metrics such as a classifier trained to distinguish generated from GEANT4 samples.

Scaling this approach up to higher resolutions relevant for future detector upgrades introduces prohibitive memory constraints. We introduce a fast detector simulation based on an inductive series of normalizing flows which overcomes this problem. By training the flow on the pattern of energy deposition in both the current and previous layer of a GEANT event, Inductive CaloFlow is capable of efficiently generating new events even for large calorimeter geometries. We demonstrate our architecture using CaloChallenge Datasets 2 and 3, and demonstrate they reproduce GEANT-like events at higher fidelity than previously possible.

Generative Models – Detector Level / 71

AtlFast3, the new ATLAS fast simulation tool

Author: Michele Faucci Giannelli¹

¹ *INFN e Universita Roma Tor Vergata (IT)*

Corresponding Author: michele.faucci.giannelli@cern.ch

AtlFast3 is the new, high-precision fast simulation in ATLAS that was deployed by the collaboration to replace AtlFastII, the fast simulation tool that was successfully used for most of Run2. 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. During the next data-taking period, Run3, the fast simulation will become the main simulation used by analyses; for this reason, the collaboration is further improving the tool expanding the use of ML Preliminary results will be presented on the performance of the new AtlFast3.

Generative Models – Particle Level / 72

Particle Cloud Generation

Author: Raghav Kansal¹

¹ *Univ. of California San Diego (US)*

Corresponding Author: raghav.kansal@cern.ch

Particle Cloud Generation

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 evaluation metrics for the generation of such data. We then introduce our new graph network and attention-based generative models, which have excellent qualitative and quantitative performance in generating sparse jets.

Measurement / 73

ML Unfolding based on conditional Invertible Neural Networks using iterative training

Authors: Anja Butter^{None}; Mathias Josef Backes¹

¹ *Universität Heidelberg*

Corresponding Authors: butter@thphys.uni-heidelberg.de, mathias.backes@stud.uni-heidelberg.de

The unfolding of detector effects impacting experimental measurements is crucial for the comparison of data to theory predictions. While traditional methods were limited to low dimensional data, machine learning has enabled new techniques to unfold high-dimensional data. Generative networks like conditional Invertible Neural Networks (cINN) enable a probabilistic unfolding, which map individual events to their corresponding unfolded probability distribution. The precision of this method is however limited by the similarity between simulated training data and the measurement we want to unfold. We therefore introduce an improved version of the cINN Unfolding by combining it with an iterative reweighting which adjusts for deviations between simulation and data. We validate the performance on toy data and an EFT-dependent example.

Reconstruction / 75

Advances in developing deep neural networks for finding primary vertices in proton-proton collisions at the LHC

Authors: Elliott Kauffman¹; Henry Fredrick Schreiner²; Michael David Sokoloff³; Michael Peters^{None}; Rocky Bala Garg⁴; Simon Akar³; William James Tepe^{None}

¹ *Princeton University (US)*

² *Princeton University*

³ *University of Cincinnati (US)*

⁴ *Stanford University (US)*

Corresponding Authors: michael.david.sokoloff@cern.ch, simon.akar@cern.ch, henry.fredrick.schreiner@cern.ch, elliot.kauffman@duke.edu, rocky.bala.garg@cern.ch, will.tepe@cern.ch, peter2mj@mail.uc.edu

We have been studying the use of deep neural networks (DNNs) to identify and locate primary vertices (PVs) in proton-proton collisions at the LHC. Earlier work focused on finding primary vertices in simulated LHCb data using a hybrid approach that started with kernel density estimators (KDEs) derived from the ensemble of charged track parameters and predicted “target histograms” from which the PV positions are extracted. We have recently demonstrated that using a UNet architecture performs indistinguishably from a “flat” convolutional neural network model and that “quantization”, using FP16 rather than FP32 arithmetic, degrades its performance minimally. We have demonstrated that the KDE-to-hists algorithm developed for LHCb data can be adapted to ATLAS data. Finally, we have developed an “end-to-end” tracks-to-hists DNN that predicts target histograms directly from track parameters using simulated LHCb data that provides better performance (a lower false positive rate for the same high efficiency) than the best KDE-to-hists model studied.

76

Registration

Welcome and Introduction / 77

Welcome from Local Organizers

Authors: David Shih^{None}; Dean Thu Nguyen¹

¹ *Rutgers University*

Corresponding Authors: mpsdean@sas.rutgers.edu, dshih@physics.rutgers.edu

Welcome and Introduction / 78

Theory Opening

Author: Ian James Moulton^{None}

Corresponding Author: ianmoulton@lbl.gov

Welcome and Introduction / 79

Experimental Opening I

Corresponding Author: petar.maksimovic@cern.ch

Welcome and Introduction / 80

Experimental Opening II

Corresponding Author: tobias.golling@unige.ch

81

Closing

Authors: David Shih^{None}; Manuel Sommerhalder¹

¹ *Hamburg University (DE)*

Corresponding Authors: manuel.sommerhalder@cern.ch, dshih@physics.rutgers.edu

Measurement / 82

Loop Amplitudes from Precision Networks

Authors: Anja Butter^{None}; Peter Rangi Sorrenson¹; Tilman Plehn^{None}

¹ *Universität Heidelberg*

Corresponding Authors: wq173@uni-heidelberg.de, plehn@uni-heidelberg.de, butter@thphys.uni-heidelberg.de

Evaluating loop amplitudes is a time-consuming part of LHC event generation. I will shown for di-photon production with jets how simple, Bayesian networks can learn such amplitudes and model their uncertainties reliably. A boosted training of the Bayesian network further improves the uncertainty estimate and the network precision in critical phase space regions. In general, boosted network training of Bayesian networks allows us to move between fit-like and interpolation-like regimes of network training.

ML Keynote Talk / 83

ML Keynote Talk – Generative models, manifolds and symmetries: From QFT to molecules

Corresponding Author: rezende.danilo@gmail.com

The study of symmetries in physics has revolutionized our understanding of the world. Inspired by this, the development of methods to incorporate internal (Gauge) and external (space-time) symmetries into machine learning models is a very active field of research. We will introduce some of the latest work in the field. We will then present our latest work on invariant generative models and its applications to lattice-QCD and molecular dynamics simulations. In the molecular dynamics front, we'll talk about how we constructed permutation and translation-invariant normalizing flows on a torus for free-energy estimation. In lattice-QCD, we'll present our work that introduced the first U(N) and SU(N) Gauge-equivariant normalizing flows for pure Gauge simulations and its extensions to incorporate fermions.

Anomaly Detection / 84**Resonant anomaly detection without background sculpting**

Authors: Anna Hallin^{None}; David Shih^{None}; Gregor Kasieczka¹; Manuel Sommerhalder¹; Tobias Quadfasel^{None}

¹ *Hamburg University (DE)*

Corresponding Authors: tobias.quadfasel@uni-hamburg.de, dshih@physics.rutgers.edu, gregor.kasieczka@cern.ch, manuel.sommerhalder@cern.ch, anna.hallin@rutgers.edu

We introduce a new technique named Latent CATHODE (LaCATHODE) for performing “enhanced bump hunts”, a type of resonant anomaly search that combines conventional one-dimensional bump hunts with a model-agnostic anomaly score in an auxiliary feature space where potential signals could also be localized. The main advantage of LaCATHODE over existing methods is that it provides an anomaly score that is well behaved when evaluating it beyond the signal region, which is essential to prevent the sculpting of background distributions in the bump hunt. LaCATHODE accomplishes this by constructing the anomaly score directly in the latent space learned by a conditional normalizing flow trained on sideband regions. We demonstrate the superior stability and comparable performance of LaCATHODE for enhanced bump hunting on the LHC Olympics R&D dataset.

Generative Models – Detector Level / 85**Introduction to Generative Models for Fast Detector Simulation**

Author: Claudius Krause¹

¹ *Rutgers University*

Corresponding Author: claudius.krause@rutgers.edu

Anomaly Detection / 86**Introduction to Anomaly Detection**

Author: Barry Dillon¹

¹ *University of Heidelberg*

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

I will give an overview of recent progress in less-than-supervised methods for new physics searches at the LHC.

Generative Models – Detector Level / 87**Discussion**

Beyond Jets / 88**Overview of ML for Astro/Cosmo**

Authors: Miles Cranmer¹; Shirley Ho^{None}

¹ *Princeton*

Corresponding Authors: miles.cranmer@gmail.com, shirleyho@flatironinstitute.org

I will give an overview of recent progress in ML applications to Astro/Cosmo.

Beyond Jets / 89**Overview of ML for Neutrinos**

Author: Fernanda Psihas¹

¹ *Fermi National Accelerator Laboratory*

Corresponding Author: psihas@fnal.gov

I will give an overview of ML applications to Neutrino Physics.

ML Keynote Talk / 90**Panel Discussion**

Authors: Danilo Rezende^{None}; David Shih^{None}; Jesse Thaler¹; Nick Dunn²; Savannah Jennifer Thais³; Tilman Plehn^{None}

¹ *MIT*

² *Two Sigma*

³ *Princeton University (US)*

Corresponding Authors: plehn@uni-heidelberg.de, savannah.jennifer.thais@cern.ch, nick.dunn@twosigma.com, jthaler@mit.edu, dshih@physics.rutgers.edu, rezende.danilo@gmail.com

“ML connections between industry and HEP”

Beyond Jets / 91**Overview of ML for Gaia**

Corresponding Author: mbuckley@physics.rutgers.edu

The Gaia space telescope measures the position and proper motion of a billion stars in the neighborhood of the Sun. This dataset contains stellar streams, tidal debris, and other structures that can cast light on the structure of the Galaxy, its merger history, and its dark matter component. I review the machine learning approaches – including classifiers, normalizing flows, and anomaly detection

– which have been used to understand the Gaia dataset, and possible directions of future work of interest to the machine learning and particle physics community.

Equivariance and New Architectures / 92

Symmetries, Safety, and Self-Supervision

Authors: Barry Dillon¹; Gregor Kasieczka²; Lorenz Vogel³; Peter Rangi Sorrenson⁴; Tilman Plehn^{None}

¹ *University of Heidelberg*

² *Hamburg University (DE)*

³ *Heidelberg University*

⁴ *Universität Heidelberg*

Corresponding Authors: dillon@thphys.uni-heidelberg.de, plehn@uni-heidelberg.de, vogel@thphys.uni-heidelberg.de, wq173@uni-heidelberg.de, gregor.kasieczka@cern.ch

Collider searches face the challenge of defining a representation of high-dimensional data such that physical symmetries are manifest, the discriminating features are retained, and the choice of representation is new-physics agnostic. We introduce JetCLR to solve the mapping from low-level data to optimized observables through self-supervised contrastive learning. As an example, we construct a data representation for top and QCD jets using a permutation-invariant transformer-encoder network and visualize its symmetry properties. We compare the JetCLR representations with alternative representations using linear classifier tests and demonstrate its performance on an anomaly detection task.

Classification / 93

Identification of hadronic tau decays using a deep neural network with the CMS experiment at LHC

Author: Mykyta Shchedrolosiev¹

¹ *Deutsches Elektronen-Synchrotron (DE)*

Corresponding Author: mykyta.shchedrolosiev@cern.ch

Tau leptons are a key ingredient to perform many Standard Model measurements and searches for new physics at LHC. The CMS experiment has released a new algorithm to discriminate hadronic tau lepton decays against jets, electrons, and muons. The algorithm is based on a deep neural network and combines fully connected and convolutional layers. It combines information from all individual reconstructed particles near the tau axis with information about the reconstructed tau candidate and other high-level variables. Many CMS Run 2 analyses have already benefitted from the improvement brought in performance. The algorithm is presented together with its measured performance in CMS Run 2 data.

Reconstruction / 94

Machine learning for particle flow at CMS

Authors: Dylan Sheldon Rankin¹; Eric Anton Moreno²

¹ *Massachusetts Inst. of Technology (US)*

² *Massachusetts Institute of Technology (US)*

Corresponding Authors: dylan.sheldon.rankin@cern.ch, eric.anton.moreno@cern.ch

The particle-flow (PF) algorithm is of central importance to event reconstruction at the CMS detector, and has been a focus of developments in light of planned Phase-2 running conditions with an increased pileup and detector granularity. Current rule-based implementations rely on extrapolating tracks to the calorimeters, correlating them with calorimeter clusters, subtracting charged energy and creating neutral particles from significant energy deposits. Such rule-based algorithms can be difficult to extend and may be computationally inefficient under high detector occupancy, while also being challenging to port to heterogeneous architectures in full detail.

In recent years, end-to-end machine learning approaches for event reconstruction have been proposed, including for PF at CMS, with the possible advantage of directly optimising for the physical quantities of interest, being highly reconfigurable to new conditions, while also being a natural fit for deployment on heterogeneous hardware accelerators.

One of the proposed approaches for machine-learned particle-flow (MLPF) reconstruction relies on graph neural networks to infer the full particle content of an event from the tracks and calorimeter clusters based on a training on simulated samples, and has been recently implemented in CMS as a possible future reconstruction R&D direction to fully map out the characteristics of such an approach in a realistic setting.

We discuss progress in CMS towards an MLPF implementation, thus paving the way to potentially improving the detector response in terms of physical quantities of interest while also allowing for native deployment on heterogeneous architectures.

Measurement / 95

How can Bayesian networks be used for uncertainty quantification in particle physics?

Author: Christina Peters¹

¹ *University of Delaware*

Corresponding Author: petersc@udel.edu

Uncertainty quantification is crucial for data analysis and hypothesis testing. Many machine learning algorithms were not designed to provide information about the reliability of their predictions, and the methods for estimating uncertainties from these algorithms can lack transparency. In this talk we demonstrate the Bayesian network framework, which was developed using a rigorous formalism for probabilistic reasoning including both representation and inference (Pearl 1988). This framework uses a graph-based representation of the joint probability distribution as the basis for compactly encoding a high-dimensional distribution, resulting in a simple, interpretable model that is designed for uncertainty quantification. Bayesian networks are well suited to problems where uncertainty quantification is paramount, the scientist would like to constrain the model based on domain knowledge, and the set of variables is small. We constructed a Bayesian network for reconstruction of interaction position in dark matter direct-detection experiments and found that it yielded highly informative per-interaction uncertainties, while also demonstrating precision on reconstructed positions comparable to existing methodologies. This framework can be applied similarly to other inverse problems in particle physics, such as jet classification.

Anomaly Detection / 96

Anomaly detection in a perspective of interdisciplinary research

Author: Taoli Cheng¹

¹ *University of Montreal*

Corresponding Author: chengtaoli.1990@gmail.com

Following the previous work of leveraging Standard Model jet classifiers as generic anomalous jet taggers (<https://arxiv.org/abs/2201.07199>), we present an analysis of regularized SM jet classifiers serving as anti-QCD taggers. In the second part of the presentation, from the perspective of interdisciplinary research, we initiate a discussion on the opportunities and challenges involved in the pipeline of applying deep learning techniques in scientific discovery.