

5th Inter-experiment Machine Learning Workshop

Monday 9 May 2022 - Friday 13 May 2022

CERN

Book of Abstracts

Contents

An infra-red and collinear safe message passing neural network 1	1
Anomaly detection with SOFIE 2	1
Truncated Marginal Neural Ratio Estimation with swyft 3	1
Learning New Physics aware of systematic uncertainties 4	2
Clustering for interpreting complex high-energy physics models 5	2
CURTAINS for you Sliding Window: Constructing Unobserved Regions by Transporting Adjacent INtervals to improve the reach of bump hunts in the search for new physics 6	3
GNN-based algorithm for full-event filtering and interpretation at the LHCb trigger 7 . .	3
Multi-objective optimization for the CMS High Granularity Calorimeter Level 1 trigger 8	4
Turbo-Sim: a generalised generative model with a physical latent space 9	5
Neural network based primary vertex reconstruction with FPGAs for the upgrade of the CMS level-1 trigger system 10	5
Tracking of Proton Traces in a Digital Tracking Calorimeter using Reinforcement Learning 11	6
MLaaS4HEP: Machine Learning as a Service for HEP 12	7
Information-theoretic stochastic contrastive conditional GAN (InfoSCC-GAN) for physical data generation 13	7
Cosmological Simulation-Based Inference with Truncated Marginal Neural Ratio Estima- tion 14	8
Calibrating stochastic simulations with optimal transport 15	8
ML-based Correction to Accelerate Geant4 Calorimeter Simulations 16	9
Neural network distributed training and optimization library (NNLO) 17	9
Funnels: Exact maximum likelihood with dimensionality reduction 18	9
Quantum Machine Learning algorithms in the latent space of HEP events 19	10
Unsupervised learning for real-time SUEP detection in a High Level Trigger system at the LHC 20	10

Optimized Deep Learning Inference on High Level Trigger at the LHC: Computing time and Resource assessment 21	11
Electron identification in ATLAS using a deep neural network 22	11
Likelihood-Free Frequentist Inference for Calorimetric Muon Energy Measurement 23 . .	12
Using Graph autoencoders to trigger on new physics at the LHC 24	12
IEA-GAN: Intra-Event Aware GAN for the Fast Simulation of PXD Background at Belle II 25	13
ML for SUEP Detection 26	13
SUPA: A Lightweight Diagnostic Simulator for Machine Learning in Particle Physics. 27	14
Particle-based Fast Jet Simulation at the LHC with Variational Autoencoders 28	14
Particle Transformer for Jet Tagging 29	15
Using Machine Learning techniques in phenomenological studies in flavour physics 30 .	15
Targeting Multi-Loop Integrals with Neural Networks 31	16
Summary of Learning To Discover workshop 32	16
Calomplification: The Power of Generative Calorimeter Models 33	16
Invariant Representation Driven Neural Classifier for Anti-QCD Jet Tagging 34	17
Learning Optimal Test Statistics in the Presence of Nuisance Parameters 35	17
Graph Neural Network Track Reconstruction for the ATLAS ITk Detector 36	18
Leveraging universality of jet taggers through transfer learning 37	18
Quarks and gluons in the Lund plane 38	19
Explaining machine-learned particle-flow reconstruction 39	19
Towards a Deep Learning Model for Hadronization 40	20
Ephemeral Learning - Augmenting Triggers with Online-Trained Normalizing Flows 41 .	20
Two-level graphs for muon-tomography inference 42	21
Hadrons, Better, Faster, Stronger 43	21
CBM performance for (multi-)strange hadron measurements using Machine Learning techniques 44	22
FPGA acceleration of the CMS DNN based LLP Jet Algorithm for the LHC High-Luminosity upgrade 45	22
Semi-supervised Graph Neural Networks for Pileup Noise Removal 46	23
Uncertainty Aware Learning for High Energy Physics With A Cautionary Tale 47	23

Point Cloud Deep Learning Methods for Pion Reconstruction in the ATLAS Detector 48	24
Anomaly detection for the quality control of silicon sensor wafers for the CMS HGAL upgrade 49	24
Application of artificial intelligence in the reconstruction of signals from the PADME electromagnetic calorimeter 50	25
Particle identification with machine learning in ALICE Run 3 51	25
Conditional Invertible Network for Neutrino Regression 52	26
Autoencoders for semivisible jet detection 53	27
Measuring a New Class of Jet Observables with Neural Network Based Unfolding 54	27
Object condensation for end-to-end reconstruction in high occupancy calorimeters with graph neural networks 55	28
How to generate all possible simulations with GANs? 56	28
Data-driven machine learning algorithms for the calibration of space-charge distortion fluctuations in the ALICE TPC 57	29
Spatio-Temporal Anomaly Detection for the DQM of the CMS Experiment via Graph Networks 58	29
TBD 59	30
Future applications of ML in HEP 60	30
ML in Cosmology 61	30
Quantum ML 62	30
CERN ML Resources (speaker TBC) 63	30
JAX 64	30
CERN Data Science Seminar 67	31
How to drive scientific progress with community-driven open source projects: the scikit-learn approach 68	31
Hardware and software challenges for massive-scale AI 69	31
CERN Knowledge Transfer (TBC) 70	31
Foundation Models for Accelerated Discovery 71	31
Differentiable Physics Simulations for Deep Learning 72	32
Opening of the workshop 73	32
JAX and Differentiable Programming Tutorial 74	33
Workshop closing 75	33

Workshop / 1**An infra-red and collinear safe message passing neural network****Authors:** Michael Spannowsky¹; Partha Konar²; Vishal Singh Ngairangbam^{None}¹ *IPPP Durham*² *Physical Research Laboratory, Ahmedabad, Gujarat-380 009, INDIA***Corresponding Author:** vishalngairangbam@gmail.com

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 the construction of 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.

Workshop / 2**Anomaly detection with SOFIE****Author:** Juan Antonio Aguilar Saavedra¹¹ *Consejo Superior de Investigaciones Científicas (CSIC) (ES)***Corresponding Author:** juan.antonio.aguilar.saavedra@cern.ch

I will discuss a new concept in anomaly detection based on (1) tagging of objects to detect “exotic” ones (2) search for anomalous features by comparing the untagged and tagged samples. Both (1) and (2) are done with ML tools, either supervised or non-supervised. In the current implementation for multi-pronged jets, tagging is supervised and the search for features is unsupervised.

Workshop / 3**Truncated Marginal Neural Ratio Estimation with swyft****Authors:** Benjamin Kurt Miller¹; Alex Cole¹; Patrick Forré¹; Gilles Louppe^{None}; Christoph Weniger¹¹ *University of Amsterdam***Corresponding Author:** b.k.miller@uva.nl

Parametric stochastic simulators are ubiquitous in science, often featuring high-dimensional input parameters and/or an intractable likelihood. Performing Bayesian parameter inference in this context can be challenging. We present a neural simulation-based inference algorithm which simultaneously offers simulation efficiency and fast empirical posterior testability, which is unique

among modern algorithms. Our approach is simulation efficient by simultaneously estimating low-dimensional marginal posteriors instead of the joint posterior and by proposing simulations targeted to an observation of interest via a prior suitably truncated by an indicator function. Furthermore, by estimating a locally amortized posterior our algorithm enables efficient empirical tests of the robustness of the inference results. Since scientists cannot access the ground truth, these tests are necessary for trusting inference in real-world applications. We perform experiments on a marginalized version of the simulation-based inference benchmark and two complex and narrow posteriors, highlighting the simulator efficiency of our algorithm as well as the quality of the estimated marginal posteriors.

Our implementation of the above algorithm is called *swyft*. It accomplishes the following items: (a) estimates likelihood-to-evidence ratios for arbitrary marginal posteriors; they typically require fewer simulations than the corresponding joint. (b) performs targeted inference by prior truncation, combining simulation efficiency with empirical testability. (c) seamlessly reuses simulations drawn from previous analyses, even with different priors. (d) integrates *dask* and *zarr* to make complex simulation easy.

Relevant code and papers can be found online here:

<https://github.com/undark-lab/swyft>

<https://arxiv.org/abs/2107.01214>

Workshop / 4

Learning New Physics aware of systematic uncertainties

Authors: Andrea Wulzer¹; Gaia Grosso²; Marco Zanetti²; Maurizio Pierini³; Raffaele D’Agnolo^{None}

¹ *CERN and EPFL*

² *Universita e INFN, Padova (IT)*

³ *CERN*

Corresponding Author: gaia.grosso@cern.ch

New Physics Learning Machine (NPLM) is a novel machine-learning based strategy to detect multivariate data departures from the Standard Model predictions, with no prior bias on the nature of the new physics responsible for the discrepancy [1, 2]. The main idea behind the method is to build the log-likelihood-ratio hypothesis test by translating the problem of maximizing the log-likelihood-ratio into the minimization of a loss function. NPLM has been recently extended in order to deal with the uncertainties of the Standard Model predictions [3]. The new formulation directly builds on the specific maximum-likelihood-ratio treatment of uncertainties as nuisance parameters, that is routinely employed in high-energy physics for hypothesis testing. In this talk, after outlining the conceptual foundations of the algorithm, we describe the procedure to account for systematic uncertainties and we show how to implement it in a multivariate setup by studying the impact of two typical sources of experimental uncertainties in two-body final states at the LHC.

Workshop / 5

Clustering for interpreting complex high-energy physics models

Author: Walter Hopkins¹

¹ *Argonne National Laboratory (US)*

Corresponding Author: walter.hopkins@cern.ch

After discovering the last piece of the Standard Model (SM), the Higgs boson, experiments at the Large Hadron Collider (LHC) have been searching for hints of physics Beyond the SM (BSM) to yield insights into these phenomena. These searches have not yet produced any significant deviations from SM predictions. Identifying unexplored regions in experimental observable space (object pTs, MET, etc) is essential in developing future BSM searches. One tool for finding new search regions is the Phenomenological Minimal Supersymmetric Standard Model (pMSSM) scan that is currently ongoing within the ATLAS collaboration. pMSSM models that have not been excluded by current searches can be used to build new search regions. Manually interpreting the high-dimensional space of our observables for each non-excluded model can however be challenging. Unsupervised data exploration algorithms (e.g., clustering) can analyze all of the non-excluded models and identify groups of non-excluded pMSSM models that live in a similar region of observable space. This could reduce thousands of theory models into a significantly smaller number of proto-search regions. These regions can then be developed into new BSM search regions. We present results from applying k-means clustering and dimensional reduction in the form of an autoencoder applied to well-understood simplified SUSY models. These results show that models can be grouped together in an unsupervised manner.

Workshop / 6

CURTAINS for you Sliding Window: Constructing Unobserved Regions by Transporting Adjacent INTERVALS to improve the reach of bump hunts in the search for new physics

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

¹ *Universite de Geneve (CH)*

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

In this talk we present CURTAINS, a new data driven ML technique for constructing a background template on a resonant spectrum, for use in bump hunts in the search for new physics using a sliding window approach. By employing invertible neural networks to parametrise the distribution of side band data as a function of the resonant observable, we learn a transformation to map any data point from its value of the resonant observable to another chosen value. Optimal transport losses are used to learn the transformation between the two sidebands, conditioned on the invariant masses of the input and target data.

CURTAINS constructs a template for the background data in the signal window by transforming the data from the sidebands into the signal region. This conditional transformation can account for changes in their properties due to the correlation with the resonant observable. With this approach we can improve the reach of bump hunts by training a classifier for anomaly detection on observables which may be correlated to the resonant observable, unlike other approaches which are very sensitive to the presence of correlations. Additionally, by transforming the data itself the correct distribution over features and their correlations are preserved.

We demonstrate the robustness and performance improvements over other ML approaches by performing a sliding window scan for various levels of signal contamination in the QCD dijet background, provided by the LHC Olympics R&D dataset. We compare the performance of our model to the leading approaches and demonstrate improved performance, especially when restricting the amount of training data to narrow sidebands. Furthermore, unlike other approaches, thanks to the invertible networks a single model is trained, which can be validated by transforming each sideband to a separate validation region.

Workshop / 7

GNN-based algorithm for full-event filtering and interpretation

at the LHCb trigger

Authors: Andrea Mauri¹; Julian Garcia Pardinias²; Jonas Eschle³; Marta Calvi⁴; Nicola Serra³; Simone Meloni²

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

² *Universita & INFN, Milano-Bicocca (IT)*

³ *University of Zurich (CH)*

⁴ *Univ. degli Studi Milano-Bicocca*

Corresponding Author: julian.garcia.pardinias@cern.ch

The LHCb experiment is currently undergoing its Upgrade I, which will allow it to collect data at a five-times larger instantaneous luminosity. In a decade from now, the Upgrade II of LHCb will prepare the experiment to face another ten-fold increase in instantaneous luminosity. Such an increase in event complexity will pose unprecedented challenges to the online-trigger system, for which a solution needs to be found. On one side, the current algorithms would be too slow to deal with the high level of particle combinatorics. On the other side, the event size will become too large to afford the persistence of all the objects in the event for offline processing. This will oblige to make a very accurate selection of the interesting parts in each event for all the possible channels, which constitutes a gargantuan task. In addition to the challenges for the trigger, the new conditions will also bring a large increase in background levels for many of the offline analyses, due to the enlarged particle combinatorics.

As a combined solution to the previous problems, we propose to evolve from the current signal-based trigger of LHCb towards a Deep-learning based Full Event Interpretation (DFEI) approach, where a new algorithm will process in real time the final-state particles of each event, with two goals: identifying which of them come from the decay of a beauty or charm heavy hadron and reconstructing the hierarchical decay chain through which they were produced. This high-level reconstruction will allow to automatically and accurately identify the part of the event which is interesting for physics analysis, allowing to safely discard the rest of the event. The usage of deep-learning is intended to fight the combinatorics problem, and exploit the complex correlations amongst all the particles in the event.

In this talk, we show the progress in the development of the first DFEI algorithm for LHCb, which is constructed as a sequence of Graph Neural Networks (GNN), where the final-state particles are represented as nodes. The algorithm has evolved from a first version, capable of performing a charged-particle filtering followed by a coarse clustering according to the common beauty-hadron ancestor (https://indico.cern.ch/event/1078058/contributions/4534576/attachments/2322573/3955371/LHCbDFEI_JGPardinias_0) to a more complete one. The new developments include the processing of the neutral-particles in the event, on top of the charged ones, and a new reconstruction step in DFEI, whose goal is to infer the hierarchical structure of the decay chains that led to the production of the particles that were pre-selected in the previous steps. This last addition, also based on a GNN, takes inspiration from the reconstruction of the lowest-common-ancestor matrix, a technique recently proposed for the Belle II experiment. We have adapted the method to the LHCb environment, in which the number of possible different beauty- and charm-hadron decay chains to be reconstructed is much higher.

Workshop / 8

Multi-objective optimization for the CMS High Granularity Calorimeter Level 1 trigger

Author: Alexandre Hakimi¹

Co-author: Jean-Baptiste Sauvan¹

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

Corresponding Author: alexandre.hakimi@cern.ch

The CMS collaboration has chosen a novel High-Granularity Calorimeter (HGCAL) for the endcap regions as part of its planned upgrade for the high luminosity LHC. The high granularity of the detector is crucial for disentangling showers overlapped with high levels of pileup events (140 or more per bunch crossing at HL-LHC). But the reconstruction of the complex events and rejection of background pose significant challenges, particularly for the Level 1 (L1) trigger, where the processing resources and latency are tightly constrained. It is therefore planned to use Machine Learning (ML) models for this task, in particular for the identification of electromagnetic and hadronic showers using the 3D shape of the energy deposits. The 3D shape of a shower is encoded in the form of shape variables computed in the HGCAL trigger primitives generation (TPG) system and sent to the central L1 trigger where they are used as inputs by classification models. The choice of this set of variables is crucial and must take into account their discrimination power, but also the limited bandwidth between the HGCAL TPG and the central L1T and the hardware resources needed to implement the classifiers. In order to find the best compromise, a multi-objective optimization technique based on genetic algorithms is used to optimize together the classification performance, the number of bits required to encode the shape variables, and the classification model complexity. The results of this optimization, and in particular the balance between performance and hardware complexity, will be discussed in this presentation.

Workshop / 9

Turbo-Sim: a generalised generative model with a physical latent space

Author: Guillaume Quétant¹

Co-authors: Mariia Drozdova²; Vitaliy Kinakh¹; Tobias Golling²; Slava Voloshynovskiy¹

¹ *Université de Genève (CH)*

² *Universite de Geneve (CH)*

Corresponding Author: guillaume.quetant@unige.ch

We present Turbo-Sim, a generalised autoencoder framework derived from principles of information theory that can be used as a generative model. By maximising the mutual information between the input and the output of both the encoder and the decoder, we are able to rediscover the loss terms usually found in adversarial autoencoders and generative adversarial networks, as well as various more sophisticated related models. Our generalised framework makes these models mathematically interpretable and allows for a diversity of new ones by setting the weight of each loss term separately. The framework is also independent of the intrinsic architecture of the encoder and the decoder thus leaving a wide choice for the building blocks of the whole network.

We apply Turbo-Sim to a collider physics generation problem: the transformation of the properties of several particles from a theory space, right after the collision, to an observation space, right after the detection in an experiment. We show that our model is able to compete with state-of-the-art method, even outperforming it in critical tasks. Moreover, these results are reached with very basic network building blocks, which is a crucial observation in view of future more expressive implementations.

Another interesting application of such a model is to use it for unfolding tasks. An important aspect of particle physics analysis is also to get back from the observed data to the actual physics. Thanks to the paradigm of using a physically meaningful latent space, i.e. the theoretical distributions of energy and momenta of a given scattering process, our Turbo-Sim model is also trained to achieve this task.

Workshop / 10

Neural network based primary vertex reconstruction with FPGAs for the upgrade of the CMS level-1 trigger system

Authors: Alex Tapper¹; Christopher Edward Brown²; Matthias Komm³; Benjamin Radburn-Smith²; Maurizio Pierini⁴; Sioni Paris Summers⁴

¹ *Imperial College London*

² *Imperial College (GB)*

³ *Deutsches Elektronen-Synchrotron (DE)*

⁴ *CERN*

Corresponding Authors: christopher.eward.brown@cern.ch, sioni.paris.summers@cern.ch, matthias.komm@cern.ch

The CMS experiment will be upgraded to maintain physics sensitivity and exploit the higher luminosity of the High Luminosity LHC. Part of this upgrade will see the first level (Level-1) trigger use charged particle tracks within the full outer silicon tracker volume as an input for the first time and new algorithms are being designed to make use of these tracks. One such algorithm is primary vertex finding which is used to identify the hard scatter in an event and separate the primary interaction from additional simultaneous interactions. This work presents a novel approach to regress the primary vertex position and to reject tracks from additional soft interactions, which uses an end-to-end neural network. This neural network possesses simultaneous knowledge of all stages in the reconstruction chain, which allows for end-to-end optimisation. The improved performance of this network versus a baseline approach in the primary vertex regression and track-to-vertex classification is shown. A quantised and pruned version of the neural network is deployed on an FPGA to match the stringent timing and computing requirements of the Level-1 Trigger.

Workshop / 11

Tracking of Proton Traces in a Digital Tracking Calorimeter using Reinforcement Learning

Author: Tobias Kortus¹

Co-authors: Ralf Keidel¹; Nicolas R. Gauger²; pCT collaboration Bergen ; research training group SIVERT

¹ *University of Applied Sciences Worms*

² *TU Kaiserslautern*

Corresponding Author: kortus@ztt.hs-worms.de

Particle therapy using protons or heavy ions is a relatively new cancer treatment modality which has acquired increasing popularity in the last decade, due to its potential in reducing undesired dose to the nearby healthy tissues, with respect to conventional radiotherapy. However, current clinical treatment planning based on computed tomography suffers from modest range uncertainties due to inaccurate conversions of Hounsfield units (HU) to relative stopping power (RSP). Proton computed tomography (pCT) poses an alternative imaging technique promising accurate pre-imaging of patients for treatment planning reducing uncertainties in dose distribution calculations. In contrast to X-rays used in conventional CT, protons do not travel on a straight line throughout patient and detector due to interactions with the traversed matter and thus require a reconstruction of the taken path prior to image reconstruction.

We propose a novel track following technique based on deep reinforcement learning (RL) for recovering proton traces inside the DTC where we formalize the task at hand as a Markov decision process (MDP) on a graph. Here we aim to learn a deterministic policy parametrized by a deep neural network optimizing the physical plausibility of sequential transitions between nodes, describing proton hit centroids.

In a proof of principle study on Monte Carlo simulated data, we show that modeling of elastic nuclear interactions is a sufficient metric for a dense reward function allowing the optimization of proton

traces in homogeneous detector configurations without knowledge of the ground truth. Moreover, with reinforcement learning we can reconstruct at the current stage trajectories originating from a variety of phantoms and particle densities with accuracies in the 50-98% range while being able to relocate the optimization steps to an initial training phase and thus avoid performing recursive or iterative optimization of proton tracks during inference. Currently this approach is limited to homogeneous detectors lacking the ability to efficiently trace protons over tracking layers. Finally, at the moment we rely on ground-truth seeding for finding initial track seeds in order to avoid unwanted behavior on the reinforcement learning approach.

Workshop / 12

MLaaS4HEP: Machine Learning as a Service for HEP

Author: Luca Giommi¹

Co-authors: Daniele Spiga²; Valentin Y Kuznetsov³; Daniele Bonacorsi⁴; Mattia Paladino⁵

¹ *Universita e INFN, Bologna (IT)*

² *Universita e INFN, Perugia (IT)*

³ *Cornell University (US)*

⁴ *University of Bologna / INFN*

⁵ *University of Bologna*

Corresponding Author: luca.giommi@cern.ch

Nowadays Machine Learning (ML) techniques are widely adopted in many areas of HEP and certainly will play a significant role also in the upcoming High-Luminosity LHC (HL-LHC) upgrade foreseen at CERN, when a huge amount of data will be produced by LHC and collected by the experiments, facing challenges at the exascale.

Here, we present a Machine Learning as a Service solution for HEP (MLaaS4HEP) to perform an entire ML pipeline (in terms of reading data, processing data, training ML models, serving predictions) in a completely model-agnostic fashion, directly using ROOT files of arbitrary size from local or distributed data sources.

With the new version of MLaaS4HEP code based on uproot4, we provide new features to improve users' experience with the framework and their workflows, e.g. users can define preprocessing operations to be applied on ROOT data before starting the ML pipeline. Then our approach is extended to use local and cloud resources via HTTP proxy which allows physicists to submit their workflows using the HTTP protocol.

Workshop / 13

Information-theoretic stochastic contrastive conditional GAN (InfoSCC-GAN) for physical data generation

Authors: Guillaume Quétant¹; Mariia Drozdova¹; Vitaliy Kinakh²; Slava Voloshynovskiy²; Tobias Golling¹; Stefan Hackstein³

¹ *Universite de Geneve (CH)*

² *Université de Genève (CH)*

³ *University of Applied Sciences and Arts Northwestern Switzerland FHNW*

Corresponding Author: vitaliy.kinakh@unige.ch

Cosmological simulations use generative deep learning models to generate galaxy images that are indiscernible from real images. Such simulations allow for a precise modeling of competing cosmological models as well as realistic propagation effects that affect observations. We present a new

stochastic contrastive conditional generative adversarial network (InfoSCC-GAN) with explorable latent space that can be used for generation of natural images as well as images of galaxies. The InfoSCC-GAN architecture is based on an unsupervised contrastive encoder built on the InfoNCE paradigm, attributes' classifier, and stochastic EigenGAN generator. We propose two approaches for selecting the class attributes: external attributes from the dataset annotations and internal attributes from the clustered latent space of the encoder. We propose a novel training method based on a generator regularization using external or internal attributes every n -th iteration using the pre-trained contrastive encoder and pre-trained attributes' classifier. The proposed InfoSCC-GAN is derived from an information-theoretic formulation of mutual information maximization between the input data and latent space representation for the encoder and the latent space and generated data for the decoder. Thus, we demonstrate a link between the training objective functions and the above information-theoretic formulation. The experimental results show that InfoSCC-GAN outperforms vanilla EigenGAN in image generation on several popular datasets, yet providing an interpretable latent space. In addition, we investigate the impact of regularization techniques and each part of the system by performing an ablation study. Finally, we demonstrate that thanks to the stochastic EigenGAN generator, the proposed framework enjoys a truly stochastic generation of natural images and galaxy images in contrast to vanilla deterministic GANs yet with the independent training of an encoder, a classifier, and a generator.

Workshop / 14

Cosmological Simulation-Based Inference with Truncated Marginal Neural Ratio Estimation

Author: Alex Cole^{None}

Co-authors: Benjamin Miller ; Christoph Weniger¹

¹ *University of Amsterdam*

Corresponding Author: a.e.cole@uva.nl

I will describe some applications of Truncated Marginal Neural Ratio Estimation (TMNRE) to cosmological simulation-based inference. In particular, I will report on using SBI for CMB power spectra (based on <https://arxiv.org/abs/2111.08030>) and realistic 21cm simulations (work in progress). Along the way, I plan to discuss some thoughts on how to incorporate active learning scenarios with high-dimensional nuisance parameter spaces, as well as criteria we need to trust results generated via simulation-based inference.

Workshop / 15

Calibrating stochastic simulations with optimal transport

Authors: Chris Pollard¹; Philipp Windischhofer¹

¹ *University of Oxford (GB)*

Corresponding Authors: chris.pollard@cern.ch, philipp.windischhofer@cern.ch

Stochastic simulators are an indispensable tool in many branches of science. Often based on first principles, they deliver a series of samples whose distribution implicitly defines a probability measure to describe the phenomena of interest. However, the fidelity of these simulators is not always sufficient for all scientific purposes, necessitating the construction of ad-hoc corrections to “calibrate” the simulation and ensure that its output is a faithful representation of reality. In this recently-published work, we leverage methods from transportation theory to construct such corrections in a systematic way. We use a neural network to compute minimal modifications to the individual samples produced by the simulator such that the resulting distribution becomes properly calibrated.

We illustrate the method and its benefits in the context of experimental particle physics, where the need for calibrated stochastic simulators is particularly pronounced.

Workshop / 16

ML-based Correction to Accelerate Geant4 Calorimeter Simulations

Authors: Evangelos Kourlitis¹; Walter Hopkins¹

¹ *Argonne National Laboratory (US)*

Corresponding Author: evangelos.kourlitis@cern.ch

The Geant4 detector simulation, using full particle tracking (FullSim), is usually the most accurate detector simulation used in HEP but it is computationally expensive. The cost of FullSim is amplified in highly segmented calorimeters where large fraction of the computations are performed to track the shower's low-energy photons through the complex geometry. A method to limit the amount of these photons is in the form of Geant4's production energy thresholds. Increased computational speed can be achieved by high values of these thresholds, however reduction of the simulation accuracy occurs beyond a geometry specific value. We propose a post-hoc machine learning (ML) correction method for calorimeter cell energy depositions. The method is based on learning the density ratio between the reduced accuracy simulation and the nominal one to extract multi-dimensional weights using a binary classifier. We explore the method using an example calorimeter geometry from the International Large Detector project and showcase initial results. The use of ML to correct calorimeter cells allows for more efficient use of heterogeneous computing resources with FullSim running on the CPU while the ML algorithm applies the correction in an event-parallel fashion on GPUs.

Workshop / 17

Neural network distributed training and optimization library (NNLO)

Authors: Irena Veljanovic¹; Jean-Roch Vlimant²; Maurizio Pierini¹; Vladimir Loncar¹

¹ *CERN*

² *California Institute of Technology (US)*

Corresponding Author: irena.veljanovic@cern.ch

With deep learning becoming very popular among LHC experiments, it is expected that speeding up the network training and optimization will soon be an issue. To this purpose, we are developing a dedicated tool at CMS, Neural Network Learning and Optimization library (NNLO). NNLO aims to support both widely known deep learning libraries Tensorflow and PyTorch. It should help engineers and scientists to easier scale neural network learning and hyperparameter optimization. Supported training configurations are a single GPU, a single node with multiple GPUs and multiple nodes with multiple GPUs. One of the advantages of the NNLO library is the seamless transition between resources, enabling researchers to quickly scale up from workstations to HPC and cloud resources. Compared to manual distributed training, NNLO facilitates the transition from a single to multiple GPUs without losing performance. With this contribution, we will discuss the status of the project and perspectives for the future.

Workshop / 18

Funnels: Exact maximum likelihood with dimensionality reduction

Authors: Johnny Raine¹; Samuel Byrne Klein¹; Sebastian Pina-Otey^{None}; Slava Voloshynovskiy²; Tobias Golling¹

¹ *Universite de Geneve (CH)*

² *University of geneva*

Corresponding Author: samuel.byrne.klein@cern.ch

Normalizing flows are exact likelihood models that have been useful in several applications in HEP. The use of these models is hampered by the dimension preserving nature of the transformations, which results in many parameters and makes the models unusable for some techniques. In this talk we introduce funnels, a new family of dimension reducing exact likelihood models.

Funnel models allow existing normalizing flows to be extended to dimension reducing transformations, and also to calculate the likelihood using existing non-invertible transformations such as convolutions. We can show that these models outperform standard flows on several downstream tasks, such as generation and anomaly detection, on standard image datasets.

We apply funnels to high energy physics datasets for generative modelling and demonstrate the advantages with respect to using standard flows or other generative models.

Workshop / 19

Quantum Machine Learning algorithms in the latent space of HEP events

Authors: Kinga Anna Wozniak¹; Ema Puljak²; Vasilis Belis³; Guenther Dissertori³; Michele Grossi⁴; Maurizio Pierini⁴; Sofia Vallecorsa⁴

¹ *University of Vienna (AT)*

² *Universitat Autònoma de Barcelona (ES)*

³ *ETH Zurich (CH)*

⁴ *CERN*

Corresponding Author: kinga.anna.wozniak@cern.ch

We present a study, based on supervised and unsupervised quantum machine learning algorithms, with the goal of proposing a new strategy for anomaly detection at the LHC. This study focuses on designing an algorithm capable of finding hidden patterns in the jet data. The algorithm is structured as a sequence of a classic and a quantum machine learning algorithm: the classic algorithm is the encoder of an autoencoder, used to reduce the dimensionality of the problem to a manageable level. The reduced latent space representation is then given as input to a Quantum Support Vector Machine (QSVM) and unsupervised quantum clustering algorithm, trained to learn a metric of the distance between jets which can be used to isolate anomalous jets. We experiment with a supervised approach and different quantum clustering algorithms, benchmarking their performance against their classical counterparts on several new physics scenarios. We also study the dependence of the algorithm performance on the number of latent-space dimensions.

Workshop / 20

Unsupervised learning for real-time SUEP detection in a High Level Trigger system at the LHC

Authors: Benedikt Maier¹; Simranjit Singh Chhibra¹; Maurizio Pierini¹; Syed Anwar Ul Hasan²

¹ CERN

² Università & INFN Pisa (IT)

Corresponding Author: simranjit.singh.chhibra@cern.ch

We propose a signal-agnostic strategy to reject QCD jets and identify anomalous signatures in a High Level Trigger (HLT) system at the LHC. Soft unclustered energy patterns (SUEP) could be such a signal —predicted in models with strongly-coupled hidden valleys —primarily characterized by a nearly spherically-symmetric signature of an anomalously large number of soft charged particles, in contrast with a comparatively collimated spray-of-hadrons signature of QCD jets. We target the experimental nightmare scenario, i.e., SUEP in exotic Higgs decays, where all dark hadrons decay promptly to standard model hadrons. We design a three-channel convolutional autoencoder (reconstructed energy deposits at the HLT in the eta-phi plane in inner-tracker, electromagnetic calorimeter, and hadron calorimeter). By processing raw-event information, this application would be ideal for central online or offline computing workflows. Our study focuses on detecting a SUEP signal; however, the technique can be applied to any new physics model that predicts signatures anomalous to QCD jets.

Workshop / 21

Optimized Deep Learning Inference on High Level Trigger at the LHC: Computing time and Resource assessment

Authors: Kinga Anna Wozniak¹; Maurizio Pierini²; Pratik Jawahar³; Syed Anwar Ul Hasan⁴; Nadezda Chernyavskaya²

¹ CERN and University of Vienna

² CERN

³ Worcester Polytechnic Institute and CERN

⁴ Università & INFN Pisa (IT)

Corresponding Authors: syed.hasan@sns.it, maurizio.pierini@cern.ch

We present a study on latency and resource requirements for deep learning algorithms to run on a typical High Level Trigger computing farm at a high-pT LHC experiment at CERN. As a benchmark, we consider convolutional and graph autoencoders, developed to perform real-time anomaly detection on all the events entering the High Level Trigger (HLT) stage. The benchmark dataset consists of synthetic multijet events, simulated at a center-of-mass energy 13 TeV. Having in mind a next-generation heterogeneous computing farm powered with GPUs, we consider both optimized CPU and GPU inference, using hardware-specific optimization tools to meet the constraints of real-time processing at the LHC: ONNX runtime for CPU and NVIDIA TensorRT for GPUs. We observe O(msec) latency with different event batch sizes for both CPU- and GPU-based model inference with maximal gain seen at batch size of 1 (corresponding to the typical use case of event-parallelized HLT farms). We show that these optimized workflows offer significant savings with respect to native solutions (Tensorflow 2 and Keras) both in terms of time and computing resources.

Workshop / 22

Electron identification in ATLAS using a deep neural network

Authors: Johnny Raine¹; Lukas Ehrke¹; Knut Zoch¹; Tobias Golling¹

¹ *Universite de Geneve (CH)*

Corresponding Author: lukas.fabian.ehrke@cern.ch

The currently used identification of electrons in ATLAS uses a likelihood approach without considering correlations between the input variables. In this talk we introduce the next generation identification algorithm using the same input information but with a deep neural network in order to extract more information from the input variables and substantially improve the rejection of fake electrons. In simulated data, the network improves the rejection of background over the current likelihood approach by a factor of up to four while maintaining the same signal efficiency.

Furthermore, instead of using a binary classifier a multiclass classifier is used allowing for more flexibility in rejecting specific backgrounds by giving more weight to the output of the network of that background. This will also allow to have two different identification scores, one accepting and one rejecting electrons with misidentified charge, without the need of an additional algorithm.

Workshop / 23

Likelihood-Free Frequentist Inference for Calorimetric Muon Energy Measurement

Author: Luca Masserano¹

Co-authors: Ann Lee ¹; Mikael Kuusela ¹; Rafael Izbicki ²; Tommaso Dorigo ³

¹ *Carnegie Mellon University*

² *Federal University of São Carlos*

³ *INFN, Padova*

Corresponding Author: masserano.luca@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 interactions at different configurations of the 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 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 intervals within an LFI setting. Waldo reframes the well-known Wald test to convert parameter point estimates from any prediction algorithm to confidence sets that are guaranteed to have the nominal coverage even in finite samples. We exploit an existing LFI framework, which 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) that has proposed this new avenue as an alternative to curvature-based measurements in a magnetic field.

Workshop / 24

Using Graph autoencoders to trigger on new physics at the LHC

Authors: Muhammad-Hassan Shahid^{None}; Katya Govorkova¹; Maurizio Pierini¹

¹ *CERN*

Corresponding Author: hassan.shahid@cern.ch

We investigate the potential of graph neural networks in unsupervised search for new physics signatures in the extremely challenging environment at the L1 at the Large Hadron Collider (LHC). On a dataset mimicking the hardware-level trigger input, we demonstrate that graph autoencoders can significantly enhance new physics contributions. Moreover, we implement the graph autoencoder on FPGA to check if the strict constraints from the L1 are satisfied.

Workshop / 25

IEA-GAN: Intra-Event Aware GAN for the Fast Simulation of PXD Background at Belle II

Author: Hosein Hashemi¹

Co-authors: Nikolai Hartmann²; Sahand Sharifzadeh²; Matej Srebre ; Thomas Kuhr²

¹ *LMU Munich*

² *Ludwig Maximilian University of Munich*

Corresponding Author: gh.hashemi@physik.uni-muenchen.de

A realistic detector simulation is extremely important in particle physics. However, the current methods are very inefficient computationally since large amounts of resources are required for the readout, storage and distribution of simulation data. Deep generative models allow for more effective fast simulation of this information. 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), in order to generate sensor-dependent images for the pixel vertex detector (PXD) which is the most sensitive sub-detector at the Belle II Experiment. First, we show that using the domain-specific relational inductive bias by introducing a Relational Reasoning Module, one can approximate the concept of an “event” in the detector simulation. Second, we incorporate a Uniformity loss in order to maximize the information entropy of the discriminator’s knowledge. Lastly, we develop the IEA-loss for the generator in order to imitate the class-to-class knowledge of the discriminator. As a result we show that the IEA-GAN not only captures fine-grained semantic and statistical similarity among the images but also it leads to a significant enhancement in the image fidelity and diversity in comparison to previous state of the art models.

Workshop / 26

ML for SUEP Detection

Authors: Chad Wells Freer¹; Luca Marco Lavezzo²; Benedikt Maier³; Adrian Alan Pol⁴; Isobel Ojalvo⁴; Christoph Paus¹; Maurizio Pierini³

¹ *Massachusetts Inst. of Technology (US)*

² *MIT*

³ *CERN*

⁴ *Princeton University (US)*

Corresponding Authors: lavezzo@mit.edu, adrian.pol@cern.ch, benedikt.maier@cern.ch, freerc@mit.edu

We explore a possible avenue for detecting Dark Showers that manifest as Soft Unclustered Energy Patterns (SUEP) in the detector with the use of supervised machine learning techniques and transfer learning. We employ a ResNet model based on Convolutional Neural Networks (CNNs) to classify events. Additionally, a robust, data-driven background estimation technique is embedded into the

model architecture through a Distance Correlation (DiSco) term in the loss function of the network; this achieves decorrelation between the classifier output and another physics-motivated discriminator in order to estimate background in the signal region through the ABCD method.

Workshop / 27

SUPA: A Lightweight Diagnostic Simulator for Machine Learning in Particle Physics.

Authors: Atul Kumar Sinha¹; Daniele Paliotta²; Bálint Máté¹; François Fleuret¹; John A. Raine¹; Sebastian Pina-Otey¹; Tobias Golling¹

¹ *University of Geneva*

² *Universite de Geneve (CH)*

Deep learning methods have gained popularity in high energy physics for fast modeling of particle showers in detectors. Detailed simulation frameworks such as the gold standard GEANT4 are computationally intensive, and current deep generative architectures work on discretized, lower resolution versions of the detailed simulation.

The development of models that work at higher spatial resolutions is currently hindered by the complexity of the full simulation data, and by the lack of simpler, more interpretable benchmarks.

Our contribution is SUPA, the SURrogate PArTicle propagation simulator, an algorithm and software package for generating data by simulating simplified particle propagation, scattering and shower development in matter. The generation is extremely fast and easy to use compared to GEANT4, but still exhibits the key characteristics and challenges of the detailed simulation. We support this claim experimentally by showing that performance of generative models on data from our simulator reflects the performance on a dataset generated with GEANT4. The proposed simulator generates thousands of particle showers per second on a desktop machine, a speed up of up to 6 orders of magnitudes over GEANT4, and stores detailed geometric information about the shower propagation. SUPA provides much greater flexibility for setting initial conditions and defining multiple benchmarks for the development of models. Moreover, interpreting particle showers as point clouds creates a connection to geometric machine learning and provides challenging and fundamentally new datasets for the field.

Workshop / 28

Particle-based Fast Jet Simulation at the LHC with Variational Autoencoders

Authors: Mary Touranakou¹; Nadya Chernyavskaya²; Javier Mauricio Duarte³; Dimitrios Gunopoulos⁴; Raghav Kansal³; Breno Orzari⁵; Maurizio Pierini²; Thiago Tomei Fernandez⁵; Jean-Roch Vlimant⁶

¹ *National and Kapodistrian University of Athens (GR)*

² *CERN*

³ *Univ. of California San Diego (US)*

⁴ *National and Kapodistrian University of Athens*

⁵ *UNESP - Universidade Estadual Paulista (BR)*

⁶ *California Institute of Technology (US)*

Corresponding Author: mary.touranakou@cern.ch

At the LHC, the full-simulation workflow requires a large fraction of the computing resources available for experiments. With the planned High Luminosity upgrade of the LHC, the amount of needed

simulated datasets would even increase. Speeding up the simulation workflow is of crucial importance for the success of the HL-LHC program and Deep Learning is considered as a promising approach to achieve this goal. In this study, we employ Deep Variational Autoencoders to train a fast simulation of jets of particles at the LHC. Starting from a generator-level view of a jet, we train a variational autoencoder to model the detector response and get the corresponding jet at reconstruction-level bypassing the time-consuming detector simulation and reconstruction steps. This approach achieves inference time comparable to that of a rule-based fast simulation and an accurate description of the momenta of the jet and its constituents.

Workshop / 29

Particle Transformer for Jet Tagging

Authors: Congqiao Li¹; Sitian Qian¹; Huilin Qu²

¹ Peking University (CN)

² CERN

Corresponding Author: stjian@pku.edu.cn

Jet tagging is a critical yet challenging classification task in particle physics. While deep learning has transformed jet tagging and significantly improved performance, the lack of a large-scale public dataset impedes further enhancement. In this work, we present JetClass, a new comprehensive dataset for jet tagging. The JetClass dataset consists of 100 M jets, about two orders of magnitude larger than existing public datasets. A total of 10 types of jets are simulated, including several types unexplored for tagging so far. Based on the large dataset, we propose a new Transformer-based architecture for jet tagging, called Particle Transformer (ParT). By incorporating pairwise particle interactions in the attention mechanism, ParT achieves higher tagging performance than a plain Transformer and surpasses the previous state-of-the-art, ParticleNet, by a large margin. The pre-trained ParT models, once fine-tuned, also substantially enhance the performance on two widely adopted jet tagging benchmarks.

<https://arxiv.org/abs/2202.03772>

Workshop / 30

Using Machine Learning techniques in phenomenological studies in flavour physics

Authors: Jorge Alda Gallo¹; Jaume Guasch²; Siannah Penaranda-Rivas^{None}

¹ Universidad de Zaragoza

² Barcelona

Corresponding Author: jalda@unizar.es

In the recent years, a series of measurements in the observables $R_{K^{(*)}}$ and $R_{D^{(*)}}$ concerning the semileptonic decays of the B mesons have shown hints of violations of Lepton Flavour Universality (LFU). An updated model-independent analysis of New Physics violating LFU, by using the Standard Model Effective Field Theory (SMEFT) Lagrangian with semileptonic dimension six operators at $\Lambda = 1$ TeV is presented. We perform a global fit, in order to assess the impact of the New Physics in a broad range of observables including B -physics, electroweak precision test, Higgs physics and nuclear β decays. We discuss the relevance of the mixing in the first generation for the observables with heavier lepton flavours. We use for the first time in this context a Montecarlo analysis of the likelihood function to extract the confidence intervals and correlations between observables. Our

results show that a suitable strategy is to use a Gradient Boosting predictor as a proxy of the real likelihood function, and to analyze the SHAP values as a measure of the impact of each parameter of SMEFT Lagrangian in the fit.

Based on arXiv:2109.07405 [hep-ph], submitted to JHEP. Recent talks: IFT Seminar, Universidad Autónoma de Madrid (Spain) 27th January 2022 & XIII CPAN Days, Huelva (Spain) 22th March 2022.

Workshop / 31

Targeting Multi-Loop Integrals with Neural Networks

Authors: Ramon Winterhalder¹; Vitalii Maheria^{None}; Emilio Villa²; Stephen Philip Jones³; Matthias Kerner⁴; Anja Butter^{None}; Gudrun Heinrich⁵; Tilman Plehn^{None}

¹ *UC Louvain*

² *Karlsruhe Institute of Technology*

³ *University of Durham (GB)*

⁴ *University of Zürich*

⁵ *KIT*

Corresponding Author: ramon.winterhalder@uclouvain.be

Numerical evaluations of Feynman integrals often proceed via a deformation of the integration contour into the complex plane. While valid contours are easy to construct, the numerical precision for a multi-loop integral can depend critically on the chosen contour. We present methods to optimize this contour using a combination of optimized, global complex shifts and a normalizing flow. They can lead to a significant gain in precision.

Workshop / 32

Summary of Learning To Discover workshop

Author: David Rousseau¹

¹ *IJCLab-Orsay*

Corresponding Author: rousseau@lal.in2p3.fr

Learning To Discover event has taken place 19th to 29th April 2022.

Three themes have been selected based on the one hand, their interest for HEP, and the fact there is already a number of HEP teams working on it, on the other hand, their importance in the Machine Learning field : Representation Learning over Heterogeneous/Graph Data, Dealing with Uncertainties and Generative Models. A final HEP and AI 3 days conference concluded the event.

This talk is a summary of Learning To Discover event, specially tailored to IML workshop in order to avoid overlaps.

All slides and records : <https://indico.ijclab.in2p3.fr/event/5999/timetable/#20220419.detailed>

Workshop / 33

Calomplification: The Power of Generative Calorimeter Models

Author: Sebastian Guido Bieringer¹

Co-authors: Anja Butter ; Ben Nachman²; Daniel Hundhausen³; Engin Eren ; Frank-Dieter Gaede⁴; Gregor Kasieczka⁵; Mathias Trabs⁶; Sascha Daniel Diefenbacher⁵; Tilman Plehn

¹ *Hamburg University*

² *Lawrence Berkeley National Lab. (US)*

³ *Deutsches Elektronen-Synchrotron DESY*

⁴ *Deutsches Elektronen-Synchrotron (DE)*

⁵ *Hamburg University (DE)*

⁶ *Department of Mathematics, Karlsruhe Institute of Technology, Germany*

Corresponding Author: sebastian.bieringer@desy.de

Motivated by the high computational costs of classical simulations, machine-learned generative models can be extremely useful in particle physics and elsewhere. They become especially attractive when surrogate models can efficiently learn the underlying distribution, such that a generated sample outperforms a training sample of limited size. This kind of GANplification has been observed for simple Gaussian models¹ and large ranges of training sample sizes. In this talk, we extend this histogram based method to show the same effect for a physics simulation, specifically photon showers in an electromagnetic calorimeter².

¹ <https://arxiv.org/abs/2008.06545>

² <https://arxiv.org/abs/2202.07352>

Workshop / 34

Invariant Representation Driven Neural Classifier for Anti-QCD Jet Tagging

Authors: Aaron Courville¹; Taoli Cheng¹

¹ *University of Montreal*

Corresponding Author: chengtaoli.1990@gmail.com

We leverage representation learning and the inductive bias in neural-net-based Standard Model jet classification tasks, to detect non-QCD signal jets. In establishing the framework for classification-based anomaly detection in jet physics, we demonstrate that with a *well-calibrated* and *powerful enough feature extractor*, a well-trained *mass-decorrelated* supervised Standard Model neural jet classifier can serve as a strong generic anti-QCD jet tagger for effectively reducing the QCD background. Imposing *data-augmented* mass-invariance (decoupling the dominant factor) not only facilitates background estimation, but also induces more substructure-aware representation learning. We are able to reach excellent tagging efficiencies for all the test signals considered. This study indicates that supervised Standard Model jet classifiers have great potential in general new physics searches.
(<https://arxiv.org/abs/2201.07199>)

Workshop / 35

Learning Optimal Test Statistics in the Presence of Nuisance Parameters

Author: Lukas Alexander Heinrich¹

¹ CERN

Corresponding Author: lukas.heinrich@cern.ch

The design of optimal test statistics is a key task in frequentist statistics and for a number of scenarios optimal test statistics such as the profile-likelihood ratio are known. By turning this argument around we can find the profile likelihood ratio even in likelihood-free cases, where only samples from a simulator are available, by optimizing a test statistic within those scenarios. We propose a likelihood-free training algorithm that produces test statistics that are equivalent to the profile likelihood ratios in cases where the latter is known to be optimal.

Workshop / 36

Graph Neural Network Track Reconstruction for the ATLAS ITk Detector

Authors: Alexis Vallier¹; Charline Rougier¹; Daniel Thomas Murnane²; Jan Stark¹; Paolo Calafiura²; Steven Farrell³; Sylvain CAILLOU⁴; Xiangyang Ju²

¹ *Laboratoire des 2 Infinis - Toulouse, CNRS / Univ. Paul Sabatier (FR)*

² *Lawrence Berkeley National Lab. (US)*

³ *Lawrence Berkeley National Laboratory*

⁴ *CNRS*

Corresponding Author: dtmurnane@lbl.gov

Graph Neural Networks (GNNs) have been shown to produce high accuracy performance on a variety of HEP tasks, including track reconstruction in the TrackML challenge, and tagging in jet physics. However, GNNs are less explored in applications with noisy, heterogeneous or ambiguous data. These elements are expected from ATLAS Inner Tracker (ITk) detector data, when it is reformulated as a graph. We present the first comprehensive studies of a GNN-based track reconstruction pipeline on ATLAS-generated ITk data.

Significant challenges exist in translating graph methods to this dataset. We analyze several approaches to low-latency and high-efficiency graph construction, including heuristics-based construction, discrete mappings of spacepoints to detector modules, and neural network learned mappings. We also extend these ideas to mappings of spacepoint doublets for more performant graph construction. Innovations in GNN training are required for ITk, and we discuss memory management for the very large ITk point clouds, and novel constructions of loss for noisy spacepoints and background tracks.

Track candidates constructed from GNN link prediction may always suffer some inefficiency, particularly on noisy point clouds. We present several methods for post-processing GNN output for either very fast triplet seeding on GPU, or for recovering efficiency with learned embeddings of tracklets and with Kalman Filters. Finally, the performance of several configurations of GNN architecture based on the Interaction Network are considered, for various hardware and latency constraints.

Workshop / 37

Leveraging universality of jet taggers through transfer learning

Authors: Frederic Alexandre Dreyer¹; Pier Francesco Monni²; Radoslaw Piotr Grabarczyk^{None}

¹ *University of Oxford*

² CERN**Corresponding Authors:** pier.francesco.monni@cern.ch, frederic.dreyer@physics.ox.ac.uk, radoslaw.piotr.grabarczyk@cern.ch

A significant challenge in the tagging of boosted objects via machine-learning technology is the prohibitive computational cost associated with training sophisticated models. Nevertheless, the universality of QCD suggests that a large amount of the information learnt in the training is common to different physical signals and experimental setups. In this article, we explore the use of transfer learning techniques to develop fast and data-efficient jet taggers that leverage such universality. We consider the graph neural networks LundNet and ParticleNet, and introduce two prescriptions to transfer an existing tagger into a new signal based either on fine-tuning all the weights of a model or alternatively on freezing a fraction of them. In the case of W-boson and top-quark tagging, we find that one can obtain reliable taggers using an order of magnitude less data with a corresponding speed-up of the training process. Moreover, while keeping the size of the training data set fixed, we observe a speed-up of the training by up to a factor of three. This offers a promising avenue to facilitate the use of such tools in collider physics experiments.

Workshop / 38

Quarks and gluons in the Lund plane

Authors: Adam Takacs¹; Frederic Alexandre Dreyer²; Gregory Soyez³¹ University of Bergen² University of Oxford³ IPhT, CEA Saclay**Corresponding Authors:** gregory.soyez@cern.ch, frederic.dreyer@physics.ox.ac.uk, adam.takacs@uib.no

Discriminating quark and gluon jets is a long-standing topic in collider phenomenology. In this paper, we address this question using the Lund jet plane substructure technique introduced in recent years. We present two complementary approaches: one where the quark/gluon likelihood ratio is computed analytically, to single-logarithmic accuracy, in perturbative QCD, and one where the Lund declusterings are used to train a neural network. For both approaches, we either consider only the primary Lund plane or the full clustering tree. The analytic and machine-learning discriminants are shown to be equivalent on a toy event sample resumming exactly leading collinear single logarithms, where the analytic calculation corresponds to the exact likelihood ratio. On a full Monte Carlo event sample, both approaches show a good discriminating power, with the machine-learning models usually being superior. We carry on a study in the asymptotic limit of large logarithm, allowing us to gain confidence that this superior performance comes from effects that are subleading in our analytic approach. We then compare our approach to other quark-gluon discriminants in the literature. Finally, we study the resilience of our quark-gluon discriminants against the details of the event sample and observe that the analytic and machine-learning approaches show similar behaviour.

Workshop / 39

Explaining machine-learned particle-flow reconstruction

Author: Farouk Mokhtar¹**Co-authors:** Raghav Kansal¹; Daniel Diaz¹; Javier Mauricio Duarte¹; Joosep Pata²; Maurizio Pierini³; Jean-Roch Vlimant⁴¹ Univ. of California San Diego (US)² National Institute of Chemical Physics and Biophysics (EE)

³ CERN

⁴ California Institute of Technology (US)

Corresponding Author: fmokhtar@ucsd.edu

The particle-flow (PF) algorithm is used in general-purpose particle detectors to reconstruct a comprehensive particle-level view of the collision by combining information from different subdetectors. A graph neural network (GNN) model, known as the machine-learned particle-flow (MLPF) algorithm, has been developed to substitute the rule-based PF algorithm (<https://arxiv.org/abs/2101.08578>), and have shown comparable performance.

Understanding the model's decision making is not straightforward, especially given the complexity of the set-to-set prediction task, dynamic graph building, and message-passing steps. In this talk, we explore the application of an explainable AI technique, called the layerwise-relevance propagation, for GNNs and apply it to the MLPF algorithm to gauge the relevant nodes and features for its predictions. Through this process, we gain insight into the model's decision-making. Results can be found in our paper: <https://arxiv.org/abs/2111.12840>

A sneak peak of a public talk given about the topic can be found here: <https://indico.cern.ch/event/1136420/contribution/Learned%20Particle%20Flow.pdf>

Workshop / 40

Towards a Deep Learning Model for Hadronization

Authors: Aishik Ghosh¹; Andrzej Konrad Siodmok²; Ben Nachman³; Xiangyang Ju³

¹ University of California Irvine (US)

² Jagiellonian University (PL)

³ Lawrence Berkeley National Lab. (US)

Corresponding Authors: xiangyang.ju@cern.ch, a.siodmok@cern.ch, aishik.ghosh@cern.ch, benjamin.philip.nachman@cern.ch

Hadronization is a complex quantum process whereby quarks and gluons become hadrons. The widely-used models of hadronization in event generators are based on physically-inspired phenomenological models with many free parameters. We propose an alternative approach whereby neural networks are used instead. Deep generative models are highly flexible, differentiable, and compatible with Graphical Processing Unit (GPUs). We make the first step towards a data-driven machine learning-based hadronization model by replacing a component of the hadronization model within the Herwig event generator (cluster model) with a Generative Adversarial Network (GAN). We show that a GAN is capable of reproducing the kinematic properties of cluster decays. Furthermore, we integrate this model into Herwig to generate entire events that can be compared with the output of the public Herwig simulator as well as with e+e- data.

Based on: <https://arxiv.org/abs/2203.12660>

Workshop / 41

Ephemeral Learning - Augmenting Triggers with Online-Trained Normalizing Flows

Authors: Anja Butter^{None}; Sascha Daniel Diefenbacher¹; Ben Nachman²; David Shih^{None}; Gregor Kasieczka¹; Ramon Winterhalder³; Tilman Plehn^{None}

¹ Hamburg University (DE)

² *Lawrence Berkeley National Lab. (US)*

³ *UC Louvain*

Corresponding Author: sascha.diefenbacher@desy.de

The large data rates at the LHC make it impossible to store every single observed interaction. Therefore we require an online trigger system to select relevant collisions. We propose an additional approach, where rather than compressing individual events, we compress the entire data set at once. We use a normalizing flow as a deep generative model to learn the probability density of the data online. The events are then represented by the generative neural network and can be inspected offline for anomalies or used for other analysis purposes.

We demonstrate our new approach for a toy model and a correlation-enhanced bump hunt.

Workshop / 42

Two-level graphs for muon-tomography inference

Author: Giles Chatham Strong¹

¹ *Universita e INFN, Padova (IT)*

Corresponding Author: giles.chatham.strong@cern.ch

Muon tomography is a useful imaging technique for studying volumes of interest by examining the scattering and absorption of cosmic muons which pass through them. Inferring properties of the volumes, however, is challenging since muons will scatter many times within the volume, the detectors involved have finite resolution, and each muon only ever traverses a sub-portion of the whole volume.

Traditional inference approaches either function by extrapolating incoming and outgoing muon trajectories inside the volume and assign the entire scattering to a single “point of closest approach”, or use a maximum likelihood fit to infer material densities. The PoCA approach being inherently biased, and the fit being challenging to implement. In both cases the volume can be discretised into voxels.

As part of ongoing work to develop a fully-differentiable pipeline for optimising muon tomography detectors (TomOpt - see Ref*), we have studied how graph neural networks are applicable to inferring properties of the volumes of interest by learning representations of the available data in two stages: a representation of the muons for each voxel, and a representation of the surrounding voxels for each voxel. Not only does such a setup allow the full population of muons to be adequately exploited in all parts of the volume, but the resulting per-voxel representations can be easily adapted for predictions at both the voxel and volume level.

In this presentation we present results on several benchmark examples, and discuss the pros and cons of such an approach.

*MODE et al. (2022) Toward the End-to-End Optimization of Particle Physics Instruments with Differentiable Programming: a White Paper, arXiv:2203.13818 [physics.ins-det]

Workshop / 43

Hadrons, Better, Faster, Stronger

Authors: Daniel Hundhausen¹; Engin Eren^{None}; Erik Buhmann²; Frank-Dieter Gaede³; Gregor Kasieczka²; Katja Kruger³; Lennart Rustige⁴; Peter McKeown^{None}; Sascha Daniel Diefenbacher²; William Korcar²

¹ *Deutsches Elektronen-Synchrotron DESY*

² *Hamburg University (DE)*

³ *Deutsches Elektronen-Synchrotron (DE)*

⁴ *Université Clermont Auvergne (FR)*

Corresponding Author: engin.eren@desy.de

Precise modeling of physical processes is a crucial part of modern particle physics. However simulation of particle showers within a calorimeter requires significant computational resources. Fast and exact machine-learning-based shower simulators offer a promising way of alleviate this problem.

This work reports progress on two important fronts. First, the WGAN and BIB-AE generative models, previously investigated for photon showers, are improved and demonstrated to successfully learn hadronic showers initiated by charged pions in a segment of the hadronic calorimeter of the International Large Detector (ILD). Second, we consider how state-of-the-art reconstruction software applied to generated shower energies affects the obtainable energy response and resolution.

Workshop / 44

CBM performance for (multi-)strange hadron measurements using Machine Learning techniques

Authors: Andrea Dubla¹; Shahid Khan^{None}; Ilya Selyuzhenkov²; Oleksii Lubynets³; Viktor Klochkov¹

¹ *GSI*

² *GSI, Darmstadt*

³ *Taras Shevchenko National University of Kyiv (UA)*

Corresponding Author: shahidzafarkhan@gmail.com

The Compressed Baryonic Matter (CBM) experiment at FAIR will investigate the QCD phase diagram at high net-baryon density ($B > 400$ MeV) in the energy range of $\sqrt{s_{NN}} = 2.7-4.9$ GeV. Precise determination of dense baryonic matter properties requires multi-differential measurements of strange hadron yields, both for most copiously produced kaons and Λ as well as for rare (multi-)strange hyperons and their anti-particles.

In this presentation, the CBM performance for the multi-differential yield measurements of strange hadrons (K_s^0 , Λ , and Ξ^-) will be reported. The strange hadrons are reconstructed via their weak decay topology using the Kalman Filter algorithm. Machine Learning techniques, such as XGBoost, are used for non-linear multi-parameter selection of weak decay topology, resulting in high signal purity and efficient rejection of the combinatorial background. Yield extraction and extrapolation to unmeasured phase space is implemented as a multi-step fitting procedure, differentially in centrality, transverse momentum, and rapidity at different collision energies. Variation of the analysis parameters allows to estimate systematic uncertainties. A novel approach to study feed-down contribution to the primary strange hadrons using Machine Learning algorithms will also be discussed.

Workshop / 45

FPGA acceleration of the CMS DNN based LLP Jet Algorithm for the LHC High-Luminosity upgrade

Author: Tarik Ourida^{None}

Co-authors: Wayne Luk ; Niranjana Bhat ¹; Marco Barbone ²; Robert John Bainbridge ²; Alex Tapper ¹

¹ *Imperial College London*

² *Imperial College (GB)*

Corresponding Author: tarik.ourida17@imperial.ac.uk

The CMS experiment at the Large Hadron Collider (LHC) at CERN adopts the LLP (Long-Lived Particle) Jet Algorithm, to search for new physics by tagging hadronic jets which stem from exotic long-lived particles. The LLP Jet Algorithm is a multiclass classifier based on a state-of-the-art Deep Neural Network (DNN). The jet tagging model's forward inference stage employs 12 convolutional layers and 5 dense layers to produce predictions from over 600 input parameters. Its DNN based architecture is highly computationally intensive, thus not meeting real-time latency constraints for data selection systems when implemented on a CPU; while a hardware accelerated FPGA implementation would meet real-time requirements. The motivation for exploring hardware acceleration for the LLP Jet Algorithm is further justified by the projected upscale in particle collisions and hence in data collection and processing requirements due to the planned High-Luminosity LHC upgrade. This work presents an FPGA acceleration of the LLP Jet Algorithm, exploring the use of kernelization to divide the algorithm into self-contained units, allowing simple orchestration of dataflow within the network and reuse of multiplication units enabling reduced hardware resource utilization, thus yielding performance enhancements by over an order of magnitude in the most computationally expensive sections of the network.

Workshop / 46

Semi-supervised Graph Neural Networks for Pileup Noise Removal

Authors: Shikun Liu^{None}; Tianchun Li^{None}; Yongbin Feng¹; Garyfallia Paspalaki²; Miaoyuan Liu²; Pan Li^{None}; Nhan Tran¹

¹ *Fermi National Accelerator Lab. (US)*

² *Purdue University (US)*

Corresponding Author: miaoyuan.liu@cern.ch

The high instantaneous luminosity of the CERN Large Hadron Collider leads to multiple proton-proton interactions in the same or nearby bunch crossings (pileup). Advanced pileup mitigation algorithms are designed to remove this noise from pileup particles and improve the performance of crucial physics observables. This study implements a semi-supervised graph neural network for particle-level pileup noise removal, by identifying individual particles produced from pileup. The graph neural network is firstly trained on charged particles with known labels, which can be obtained from detector measurements on data or simulation, and then inferred on neutral particles for which such labels are missing. This semi-supervised approach does not depend on the ground truth information from simulation and thus allows us to perform training directly on experimental data. The performance of this approach is found to be consistently better than widely-used domain algorithms and comparable to the fully-supervised training using simulation truth information. The study serves as the first attempt at applying semi-supervised learning techniques to pileup mitigation, and opens up a new direction of fully data-driven machine learning pileup mitigation studies.

Workshop / 47

Uncertainty Aware Learning for High Energy Physics With A Cautionary Tale

Authors: Aishik Ghosh¹; Ben Nachman²; Daniel Whiteson¹

¹ *University of California Irvine (US)*

² *Lawrence Berkeley National Lab. (US)*

Corresponding Author: aishik.ghosh@cern.ch

Machine learning tools provide a significant improvement in sensitivity over traditional analyses by exploiting subtle patterns in high-dimensional feature spaces. These subtle patterns may not be well-modeled by the simulations used for training machine learning methods, resulting in an enhanced sensitivity to systematic uncertainties. Contrary to the traditional wisdom of constructing an analysis strategy that is invariant to systematic uncertainties, we study the use of a classifier that is fully aware of uncertainties and their corresponding nuisance parameters. We show on two datasets that this dependence can actually enhance the sensitivity to parameters of interest compared to baseline approaches. Finally, we provide a cautionary example for situations where uncertainty mitigating techniques may serve only to hide the true uncertainties.

Workshop / 48

Point Cloud Deep Learning Methods for Pion Reconstruction in the ATLAS Detector

Authors: Aaron Angerami¹; Russell Bate²; Wojtek Fedorko³; Ming Fong⁴; Sanmay Ganguly⁵; Piyush Karande^{None}; Alison Lister²; Ben Nachman⁶; Jan Tuzlic Offermann⁷; Mariel Pettee⁶; Dilia Maria Portillo Quintero⁸; Maximilian J Swiatkowski⁸

¹ *Lawrence Livermore Nat. Laboratory (US)*

² *University of British Columbia (CA)*

³ *TRIUMF*

⁴ *University of California, Berkeley*

⁵ *University of Tokyo (JP)*

⁶ *Lawrence Berkeley National Lab. (US)*

⁷ *University of Chicago (US)*

⁸ *TRIUMF (CA)*

Corresponding Author: mariel.pettee@gmail.com

Reconstructing the type and energy of isolated pions from the ATLAS calorimeters is a key step in the hadronic reconstruction. The baseline methods for local hadronic calibration were optimized early in the lifetime of the ATLAS experiment. Recently, image-based deep learning techniques demonstrated significant improvements over the performance over these traditional techniques. We present an extension of that work using point cloud methods that do not require calorimeter clusters or particle tracks 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. We demonstrate the performance of these new approaches as an important step towards a full deep learning-based low-level hadronic reconstruction.

Workshop / 49

Anomaly detection for the quality control of silicon sensor wafers for the CMS HGCal upgrade

Authors: Sonja Grönroos¹; Thorben Quast²

¹ *University of Helsinki (FI)*

² *CERN*

Corresponding Author: sonja.gronroos@cern.ch

With the approaching HL-LHC upgrade, the current endcap calorimeters of the CMS are to be replaced with the High Granularity Calorimeter (HGCAL). Most of the sensitive part of HGCAL will consist of approximately 25,000 silicon pad sensor wafers, each approximately 20 cm in diameter, covering a total area of more than 600 m² of silicon sensors. Electrical breakdowns have been observed during prototype testing. Those could often be attributed to the presence of various anomalies on the sensor surface, such as scratches and dust.

Therefore, visual inspection of the sensor surface might become an imperative step in the quality control program of the HGCAL sensors. A visual inspection system that is in use for this purpose consists of a programmable xy-table, microscope and a camera, which takes approximately 500 images per sensor. While the photo taking is automatized in this way, a human still has to inspect the images and look for anomalies, which is a subjective and laborious process. We are considering the application of deep learning-based tools, specifically convolutional neural networks, for the task of image classification and anomaly detection. Our goal is to use them to develop a trigger-like model that preselects images containing potential anomalies for a human to subsequently validate.

Our strategy is the implementation of an anomaly detector as an ensemble of two independent neural networks. First, an autoencoder is trained to encode and decode normal images so that the reconstruction error is minimized. Thus, the error will increase in the event of anomalous input. Second, the pixel-wise reconstruction error will be given as input to a convolutional neural network for classification. In this talk, we present the task of anomaly detection in the context of HGCAL silicon sensor qualification, we present the proof-of-concept of our approach and complement this with preliminary results.

¹ N. Akchurin et al., “Deep learning applications for quality control in particle detector construction”, arXiv:2203.08969 [hep-ex], 2022.

Workshop / 50

Application of artificial intelligence in the reconstruction of signals from the PADME electromagnetic calorimeter

Author: Paola Gianotti¹

¹ *INFN Laboratori Nazionali di Frascati (IT)*

Corresponding Authors: kallystoimenova@gmail.com, paola.gianotti@cern.ch

PADME experiment at LNF-INFN is devoted to the search for the associate production of new light particles using accelerated positrons which annihilate in a thin active diamond target.

The core of the experiment is an electromagnetic calorimeter made of 616 BGO crystals which is dedicated to the measurement of the energy and the position of the final state photons.

The high beam particle multiplicity over a short bunch duration requires reliable identification and measurement of overlapping signals. A regression machine learning based algorithm was developed to disentangle close-in-time events with high efficiency and precisely reconstruct the amplitude of the hits and their time with a sub-nanosecond resolution.

The performance of the algorithm and the sequence of improvements leading to the achieved results will be presented and discussed.

Workshop / 51

Particle identification with machine learning in ALICE Run 3

Authors: Kamil Rafal Deja¹; Lukasz Kamil Graczykowski¹; Maja Kabus¹; Lukasz Sawicki¹; Monika Joanna Jakubowska¹

¹ *Warsaw University of Technology (PL)*

Corresponding Author: maja.kabus@cern.ch

Particle identification (PID) is an essential ingredient of many measurements performed by the ALICE Collaboration. The ALICE detectors provide PID information via complementary experimental techniques, allowing for the identification of particles over a broad momentum interval ranging from about 100 MeV/c up to 20 GeV/c. The biggest challenge lies in combining the information from the different detectors. Up to now, in ALICE, particles were identified by hand-crafted selections based on the detector signals and by the Bayesian method.

We propose to use machine learning methods to classify particle species, aiming to better exploit the detector information and to improve identification in the regions where different particle species give overlapping signals. During LHC Run 2, preliminary studies were pursued by using Random Forests 1 with the tree generation based on the Gini index. This method resulted in much higher efficiencies and purities for selected particles than standard techniques.

For the coming Run 3, a more advanced approach based on Domain Adaptation Neural Networks is under investigation (<https://indico.bnl.gov/event/10699/contributions/53933/>, proceedings accepted in JINST). The new approach accounts for the discrepancies between the Monte Carlo simulations and the experimental data. The algorithm consists of a feature mapping network, whose outputs – domain-invariant features – are inputs to a particle classifier and to a domain classifier. The particle classifier outputs the predicted particle species, while the domain classifier discriminates between the real and simulated data. The classifiers are trained independently. Preliminary studies show that domain adaptation improves particle classification. This solution will be integrated in the ALICE Run 3 Analysis Framework. Preliminary results for the PID of selected particle species will be discussed as well as the possible optimizations and further developments.

¹ Tomasz Trzeciński, Łukasz Graczykowski, Michał Glinka, ALICE Collaboration, et al. Using Random Forest classifier for particle identification in the ALICE experiment. In Conference on Information Technology, Systems Research and Computational Physics, pages 3–17. Springer, 2018.

Workshop / 52

Conditional Invertible Network for Neutrino Regression

Authors: John A. Raine¹; Matthew Leigh¹; Tobias Golling¹

¹ *University of Geneva*

Corresponding Author: matthew.leigh@cern.ch

The mass of the top quark is of paramount importance as it is a highly sensitive probe of the structure and stability of the Standard Model. The final state of a leptonically decaying top quark contains a neutrino. In collider physics the neutrino escapes detection, leaving only experimental proxies for its momentum in the plane transverse to the beam pipe and no information about its longitudinal momentum. This hinders the full reconstruction of the final state and thus the invariant mass of the top quark. Conventional methods for deriving the neutrino momentum from kinematic constraints sometimes yield no real solutions. We propose a novel method for the estimation of the neutrino kinematics in the single lepton $t\bar{b}$ decay channel using a conditional invertible neural network. This is achieved by viewing the reconstruction process as an inverse problem and thus approach it as a task of conditional inference using the flow. The flow is trained to estimate the conditional probability distribution of the neutrino's 3-momentum conditioned on observed event variables. We present the performance of this method in comparison to standard reconstruction techniques used in analyses.

Workshop / 53**Autoencoders for semivisible jet detection****Author:** Jeremi Niedziela¹¹ *ETH Zurich (CH)***Corresponding Author:** jeremi.niedziela@cern.ch

The production of dark matter particles from confining dark sectors may lead to many novel experimental signatures. Depending on the details of the theory, dark quark production in proton-proton collisions could result in semivisible jets of particles: collimated sprays of dark hadrons of which only some are detectable by particle collider experiments. The experimental signature is characterised by the presence of reconstructed missing momentum collinear with the visible components of the jets. This complex topology is sensitive to detector inefficiencies and mis-reconstruction that generate artificial missing momentum.

We propose a signal-agnostic strategy to reject ordinary jets and identify semivisible jets via anomaly detection techniques. A deep neural autoencoder network with jet substructure variables as input proves highly useful for analyzing anomalous jets. The study focuses on the semivisible jet signature; however, the technique can apply to any new physics model that predicts signatures with anomalous jets from non-SM particles.

Related publication: [https://link.springer.com/article/10.1007/JHEP02\(2022\)074](https://link.springer.com/article/10.1007/JHEP02(2022)074)

Workshop / 54**Measuring a New Class of Jet Observables with Neural Network Based Unfolding****Author:** Jingjing Pan¹**Co-authors:** Jennifer Kathryn Roloff²; Matt LeBlanc³; Ben Nachman⁴¹ *Yale University (US)*² *Brookhaven National Laboratory (US)*³ *CERN*⁴ *Lawrence Berkeley National Lab. (US)***Corresponding Author:** jingjing.pan@cern.ch

Jets are collimated sprays of hadrons produced in high energy collisions. Jets play an important role in many searches for new physics, and provide an experimental window into the real time dynamics of hadronization, namely the confinement of asymptotically free quarks and gluons into hadrons. With the advent of a new class of theoretical observables, so called energy correlators, that probe the energy flow within jets, it has recently becomes possible, at least in theory, to directly “image” the confinement transition. In this talk, we will discuss the major challenges of realizing this measurement experimentally, and introduce a number of new techniques that will be required to achieve this goal. We first show that the angular resolution necessary can only be achieved using tracking information, and we perform the first measurement of “track functions” that describe the conversion of quarks and gluons into the electrically charged hadrons seen by trackers. We then show that radically new ways of unfolding data are required for these new observables, and show how this can be achieved using novel Neural Network based approaches.

More precisely, observables in high energy particle physics are computed on the distribution of energy imprinted on the detectors. To compare with theoretical calculations, the detector effects must be removed. Remarkably, this energy distribution has many similarities with point clouds,

allowing us to leverage data representations used in this more well studied context. We discuss an architecture “Energy flow networks” based on an application of the deep sets theorem to the particle physics context. We then show how one can develop a novel unfolding approach that generalizes iterative Bayesian unfolding, call “Omnifold”, and simultaneously unfolds all observables at the level of the energy distribution. We present preliminary studies, and discuss their power for particle physics measurements.

Workshop / 55

Object condensation for end-to-end reconstruction in high occupancy calorimeters with graph neural networks

Authors: Jan Kieseler¹; Shah Rukh Qasim²; Kenneth Long³; Maurizio Pierini¹; Nadya Chernyavskaya¹; Oleksandr Viazlo⁴; Raheel Nawaz⁵

¹ CERN

² Manchester Metropolitan University (GB)

³ Massachusetts Inst. of Technology (US)

⁴ Florida State University (US)

⁵ Manchester Metropolitan University

Corresponding Author: shah.rukh.qasim@cern.ch

We present an end-to-end reconstruction algorithm to build particle candidates from detector hits in next-generation granular calorimeters similar to that foreseen for the high-luminosity upgrade of the CMS detector. The algorithm exploits a distance-weighted graph neural network 2, trained with object condensation 1, a graph segmentation technique. Through a single-shot approach, the reconstruction task is paired with energy regression. We describe the reconstruction performance in terms of reconstructed-to-truth matching as well as in terms of energy resolution. In addition, we show the jet reconstruction performance of our method and discuss its inference computational cost. This work is the first-ever example of machine-learning-based single-shot calorimetric reconstruction in high-luminosity conditions with 200 pileup to the best of our knowledge.

1 <https://arxiv.org/abs/2002.03605>

2 <https://arxiv.org/abs/1902.07987>

Workshop / 56

How to generate all possible simulations with GANs?

Authors: Jan Michal Dubinski¹; Kamil Rafal Deja¹; Sandro Christian Wenzel²; Tomasz Piotr Trzcinski¹; Przemyslaw Rokita¹

¹ Warsaw University of Technology (PL)

² CERN

Corresponding Author: jan.michal.dubinski@cern.ch

The development of faster simulation methods is one of the crucial tasks currently undertaken at CERN. A part of this process in the ALICE experiment is the deep learning-based simulation tool for the Zero Degree Calorimeters (ZDC).

Generative models such as GANs that are currently used in the fast simulation framework successfully replicate the results for input particles that produce consistent calorimeter responses. However, those methods struggle to reflect the variety of possible outcomes for highly non-deterministic input

particles. Existing techniques for increasing the diversity of GAN results are able to mitigate this shortcoming but at the price of producing unrealistic and less precise results for consistent particle data.

To address this problem, we propose a novel method for the selective increase of diversity in GAN-generated samples. Our approach encourages the model to generate diverse results for input particles that allow for many possible simulation outcomes by penalizing an insufficient variety of generated results. At the same time, we allow the model to generate low-diversity results for input particles that produce consistent responses. This improvement successfully increases the diversity of generated samples for a selected subset of input data. Our method leads to higher simulation fidelity, decreasing the differences between the original and fast simulation and smoothing the distribution of the generated results.

Workshop / 57

Data-driven machine learning algorithms for the calibration of space-charge distortion fluctuations in the ALICE TPC

Author: Marian Ivanov¹

¹ *GSI - Helmholtzzentrum für Schwerionenforschung GmbH (DE)*

Corresponding Author: marian.ivanov@cern.ch

The Time Projection Chamber (TPC) of the ALICE experiment at CERN LHC was upgraded for Run 3 and Run 4. Readout chambers based on Gas Electron Multiplier (GEM) technology and a new readout scheme allow continuous data acquisition at the highest interaction rates expected in Pb-Pb collisions. In the absence of a gating grid system, a significant amount of ions generated in the multiplication region are expected to enter the TPC drift volume and distort the uniform electric field. The fluctuation of the ion space-charge density leads to a corresponding fluctuation of the space-point distortions. In order to achieve the intrinsic resolution of the detector system of $O(100 \mu\text{m})$, the distortions of $O(5 \text{ cm})$ need to be corrected in time intervals of the order of 10 ms.

To account for unknown detector parameters such as ion drift velocity or local ion transparency, data-driven methods are considered. A combination of a physical model approximation and machine learning techniques will be used to correct for distortion fluctuation. The results of preliminary studies are shown and the prospects for further development and optimization are also discussed.

Workshop / 58

Spatio-Temporal Anomaly Detection for the DQM of the CMS Experiment via Graph Networks

Authors: David Yu¹; Long Wang²; Mulugeta Asres³

¹ *Brown University (US)*

² *University of Maryland (US)*

³ *University of Agder (NO)*

Corresponding Author: mulugeta.asres@cern.ch

The Data Quality Monitoring (DQM) is in place to spot and diagnose particle physics data problems as promptly as possible to avoid data loss in the CMS experiment of CERN. Several studies have proposed to leverage the DQM automation using machine learning algorithms. However, only a few efforts explored temporal characteristics to underpin system monitoring automation of the CMS detectors via anomaly detection models thus far. Moreover, the DQM for the HCAL detector of the

CMS experiment poses multidimensional challenges, yet it is relatively unexplored with machine learning models. In this study, we propose a time-aware deep learning model for anomaly detection on the multidimensional spatial quantity of the DQM for the HCAL detector. The model employs convolutional, recurrent and graph networks to learn three-dimensional spatial characteristics, exploit physical connections of the QIE channels into RBX of the detector systems, and temporal evolution, respectively. Performance evaluation on artificially generated anomalies such as dead, hot and degrading channels has demonstrated the efficacy of the proposed model in detecting and robustly localizing anomalies in temporal and spatial contexts on digioccupancy histograms. Finally, we have carried-out comparison among several models that shows the substantial gain of the proposed model architecture and temporal modeling.

59

TBD

HEP Plenary / 60

Future applications of ML in HEP

Corresponding Author: tommaso.dorigo@cern.ch

HEP Plenary / 61

ML in Cosmology

Corresponding Author: c.weniger@uva.nl

HEP Plenary / 62

Quantum ML

Corresponding Author: sofia.vallecora@cern.ch

63

CERN ML Resources (speaker TBC)

64

JAX

67

CERN Data Science Seminar

Agenda and zoom connection are dedicated to this seminar (different from the zoom room of the workshop)

Industry Session / 68

How to drive scientific progress with community-driven open source projects: the scikit-learn approach

From the first images of a black hole by Katie Bouman using Matplotlib to neuroscience research that motivated the development of the scikit-learn library, open source has now revolutionized the way we do science. The scikit-learn software has now been cited more than 50000 times in 10 years. It's the most used software by machine learning experts on kaggle. One considers that about 2 millions of data scientists are using it every month. Yet this was made possible with limited resources and mostly by researchers and engineers in academia. In this talk I will try to convince you that this immense success is not pure luck and that it can be explained by conscious decisions.

Then I will review some recent efforts in the team in order to keep scikit-learn a leading software in the field.

Industry Session / 69

Hardware and software challenges for massive-scale AI

OpenAI's GPT-3 language model has triggered a new generation of Machine Learning models. Leveraging transformers architectures at billion-size parameters trained on massive unlabeled datasets, these language models achieve new capabilities such as text generation, question answering, or even zero-shot learning - tasks the model has not been explicitly trained for. However, training these models represent massive computing tasks, sometimes performed on dedicated supercomputers. Scaling up these models will require new hardware and optimized training algorithms.

At LightOn - a spinoff of university research -, we develop a set of hardware and software technologies to address such massive-scale computing challenges. The Optical Processing Unit (OPU) technology makes some matrix-vector multiplications in a massively parallel fashion, at record-low power consumption. Now accessible on-premises or through the cloud, the OPU technology has been used by engineers and researchers worldwide in a variety of applications, for Machine Learning and scientific computing. We also train in an efficient manner large language models that can be used for various research and business applications.

Industry Session / 70

CERN Knowledge Transfer (TBC)

Industry Session / 71

Foundation Models for Accelerated Discovery

AI is having a transformational impact for accelerating discovery. Massive volumes of scientific data, which are continuously growing due to tireless efforts from many scientific communities of discovery, are enabling data-driven AI methods to be developed at ever increasing scale and applied in novel ways to breakthrough bottlenecks in the scientific method, and to speed up the discovery process. Examples include using AI models to assist in knowledge extraction and reasoning over large repositories of scientific publications, creating AI surrogate models to predict the output of simulations and speed up complex simulation campaigns, training AI generative models to create novel hypotheses – such as to enable de-novo design of molecules by leveraging data about known chemicals and their properties, and developing AI models that can predict chemical reactions and automate synthesis and experimentation. At the same time, foundation models have emerged as a powerful new development in AI that will make further impact on accelerating scientific discovery. Foundation models learn “universal representations” from massive-scale data, typically using unsupervised or self-supervised training methods, with the goal to enable and simplify a diversity of downstream tasks. Prominent examples of foundation models are the large-language models trained from massive corpora of text that have been driving the state-of-the-art for natural language processing. In this talk, we review how foundation models work and discuss how they can learn effective representations for scientific discovery. We show examples how foundation models apply for challenges such as materials discovery and drug development. We discuss the potential for foundation models to have a key role for a broader set of scientific challenges and drive further impact of AI for accelerating discovery.

Invited talk / 72

Differentiable Physics Simulations for Deep Learning

In this talk I will focus on the possibilities that arise from recent advances in the area of deep learning for physical simulations. In this context, especially the Navier-Stokes equations represent an interesting and challenging advection-diffusion PDE that poses a variety of challenges for deep learning methods.

In particular, I will focus on differentiable physics solvers from the larger field of differentiable programming. Differentiable solvers are very powerful tools to integrate into deep learning processes. The existing numerical methods for efficient solvers can be leveraged within learning tasks to provide crucial information in the form of reliable gradients to update the weights of a neural networks. Interestingly, it turns out to be beneficial to combine supervised and physics-based approaches. The former poses a much simpler learning task by providing explicit reference data that is typically pre-computed. Physics-based learning on the other hand can provide gradients for a larger space of states that are only encountered at training time. Here, differentiable solvers are particularly powerful to, e.g., provide neural networks with feedback about how inferred solutions influence the long-term behavior of a physical model.

I will demonstrate this concept with several examples from learning to reduce numerical errors, over long-term planning and control, to generalization. I will conclude by discussing current limitations and by giving an outlook about promising future directions.

Opening session / 73

Opening of the workshop

Corresponding Authors: simon.akar@cern.ch, riccardo.torre@cern.ch, pietro.vischia@cern.ch, michael.aaron.kagan@cern.ch, lorenzo.moneta@cern.ch, gian.michele.innocenti@cern.ch, fabio.catalano@cern.ch, rousseau@lal.in2p3.fr,

butter@thphys.uni-heidelberg.de, andrea.wulzer@cern.ch

Tutorials / 74

JAX and Differentiable Programming Tutorial

Corresponding Authors: nathan.simpson@hep.lu.se, lukas.heinrich@cern.ch

Closing session / 75

Workshop closing

Corresponding Authors: simon.akar@cern.ch, riccardo.torre@cern.ch, pietro.vischia@cern.ch, michael.aaron.kagan@cern.ch, lorenzo.moneta@cern.ch, gian.michele.innocenti@cern.ch, fabio.catalano@cern.ch, rousseau@lal.in2p3.fr, butter@thphys.uni-heidelberg.de, andrea.wulzer@cern.ch