

KNO-scaling of charged hadron multiplicities within a Machine Learning based approach

Margaret Island Symposium 2023
on Particles & Plasmas
06-09 06 2023

GÁBOR BÍRÓ

biro.gabor@wigner.hu

Gergely Gábor Barnaföldi
Bence Tankó-Bartalis



UNIVERSITY OF
OXFORD

arXiv:2111.15655
arXiv:2210.10548
arXiv:2303.05422

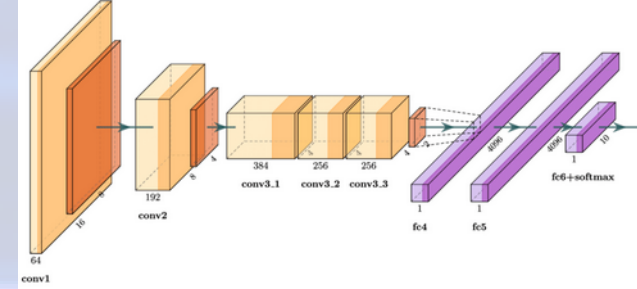


History

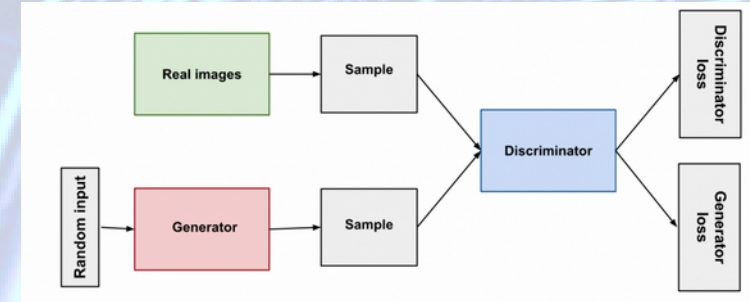
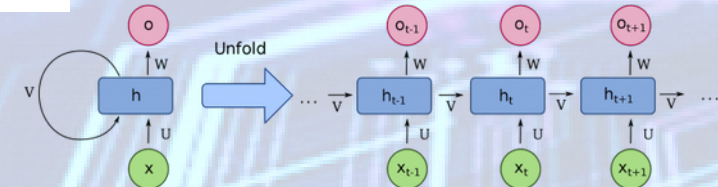
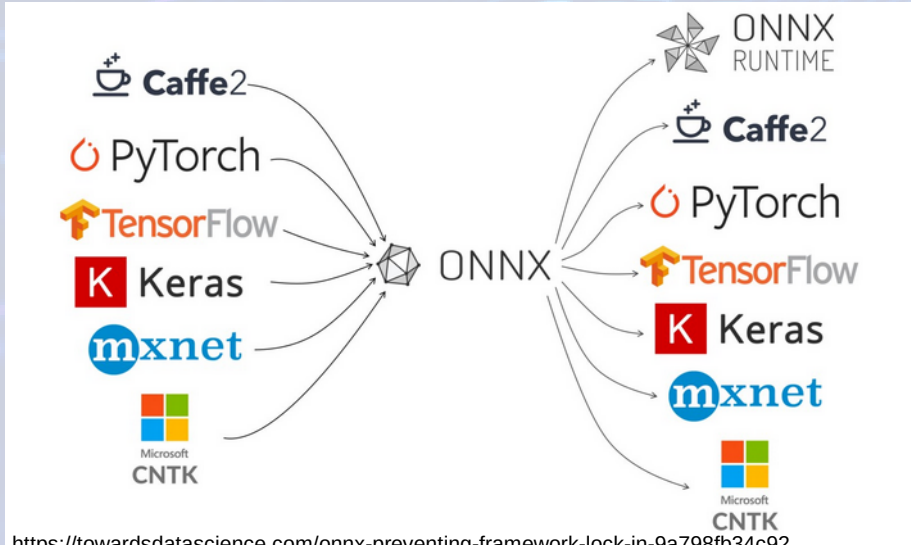
CNN (image classification, object detection, recommender systems)...

Recurrent/recursive neural networks (RNNs): Sequence modeling, next word prediction, translating sounds to words, human language translation...

Generative models: anomaly detection, pattern recognition, reinforced learning

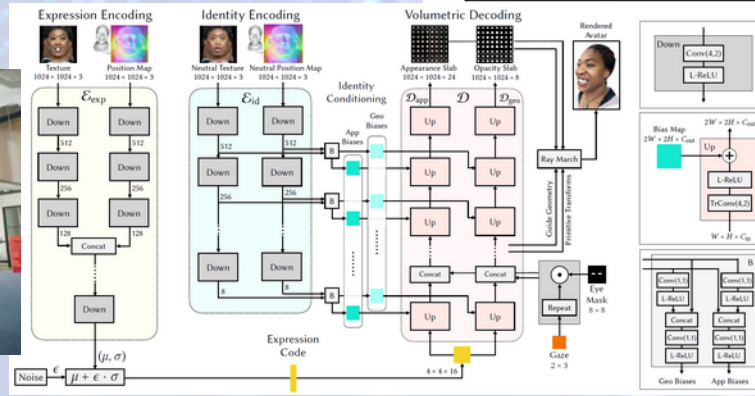
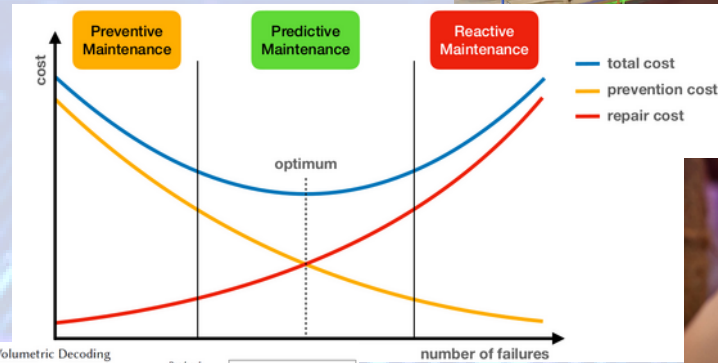
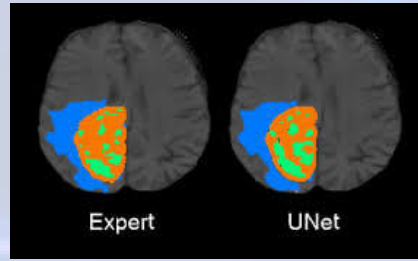


Various frameworks for training and inference:



Motivaton - data, data, more data

- Autonomous driving
- Medical imaging
- Predictive maintenance
- Anomaly detection, fake news detection
- Search of BSM physics
- Stock price prediction
- Natural Language Processing
- Virtual Assistants
- Virtual reality
- Colorization of Black and White Images
- Content generation, examples:
 - <https://infiniteconversation.com/>
 - <https://huggingface.co/spaces/stabilityai/stable-diffusion>
- Robotics



Motivaton - data, data, more data

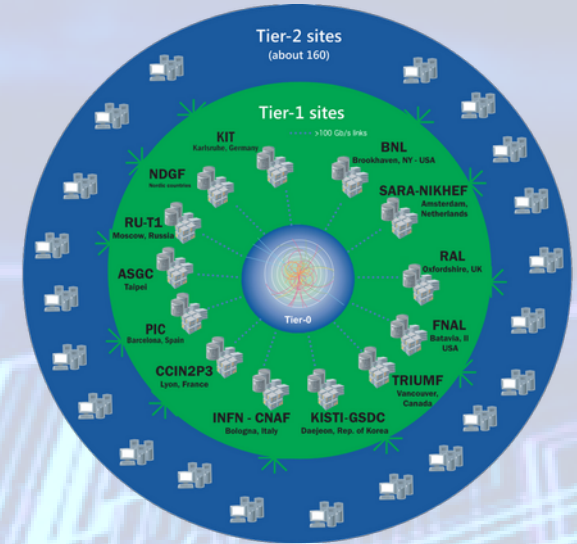


WLCG
Worldwide LHC Computing Grid



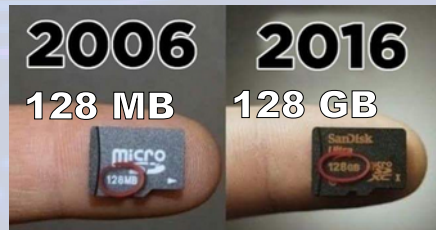
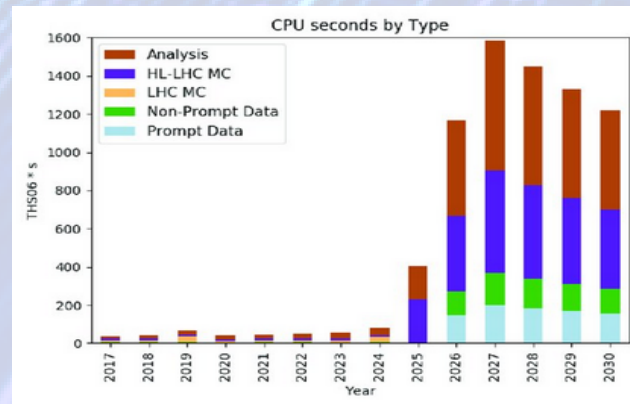
LHC in numbers: **2013** and **now**:

Data:	15 PB/year	VS	200+ PB/year
Tape:	180 PB	VS	740+ PB
Disk:	200 PB	VS	570+ PB
HS06:	2M	VS	100+ B



Storing and distributing the data is only one side of the challenge













→ analysis, simulations

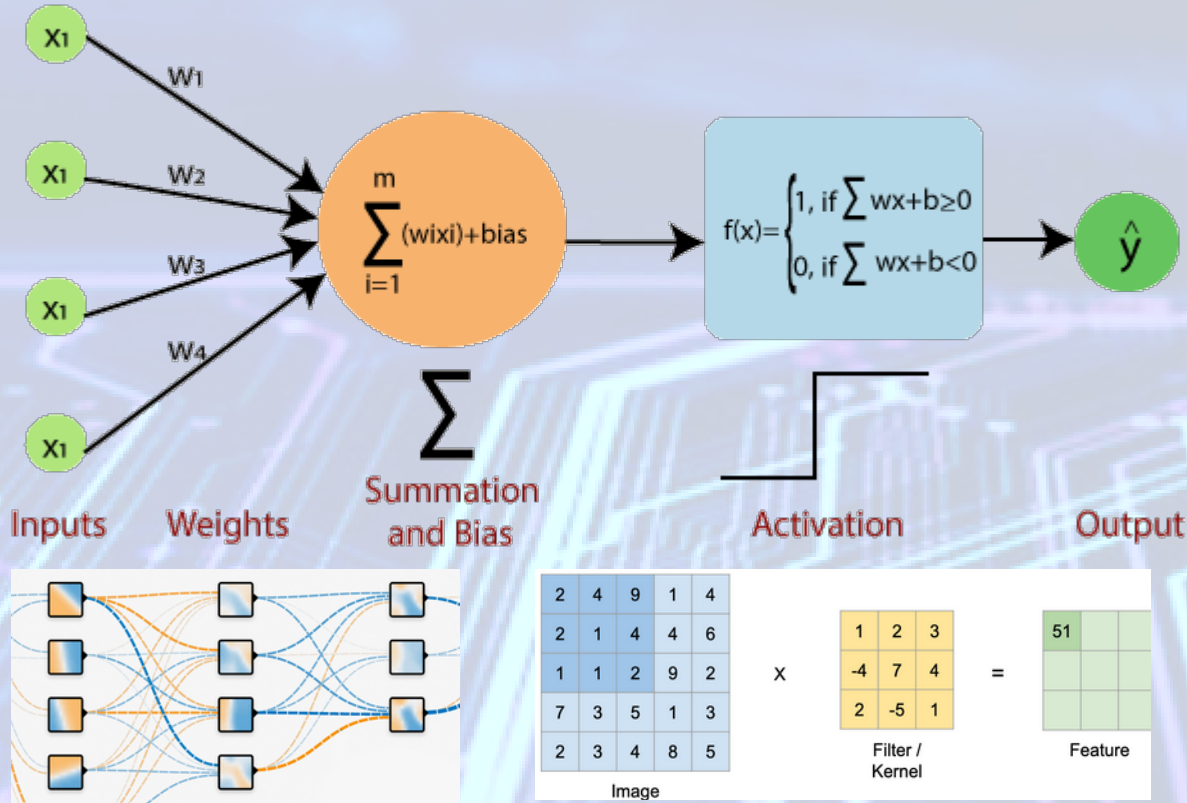


Main ingredients

Perceptrons:

- Input value(s)
- Weight: the connection between the units
- Bias: the intercept added in a linear equation
- Activation Function

<p>Sigmoid</p>  $y = \frac{1}{1+e^{-x}}$	<p>Tanh</p>  $y = \tanh(x)$	<p>Step Function</p>  $y = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$	<p>Softplus</p>  $y = \ln(1+e^x)$
<p>ReLU</p>  $y = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$	<p>Softsign</p>  $y = \frac{x}{1+ x }$	<p>ELU</p>  $y = \begin{cases} \alpha(e^x - 1), & x < 0 \\ x, & x \geq 0 \end{cases}$	<p>Log of Sigmoid</p>  $y = \ln\left(\frac{1}{1+e^{-x}}\right)$
<p>Swish</p>  $y = \frac{x}{1+e^{-x}}$	<p>Sinc</p>  $y = \frac{\sin(x)}{x}$	<p>Leaky ReLU</p>  $y = \max(0.01x, x)$	<p>Mish</p>  $y = x(\tanh(\text{softplus}(x)))$



Other important components: pooling layers, regularization and normalization, recurrent layers...

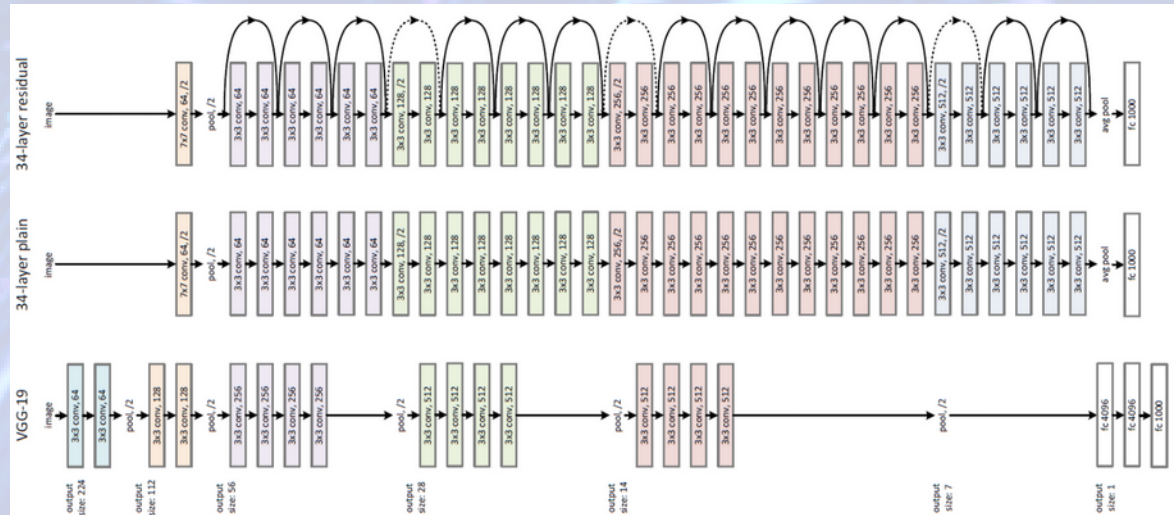
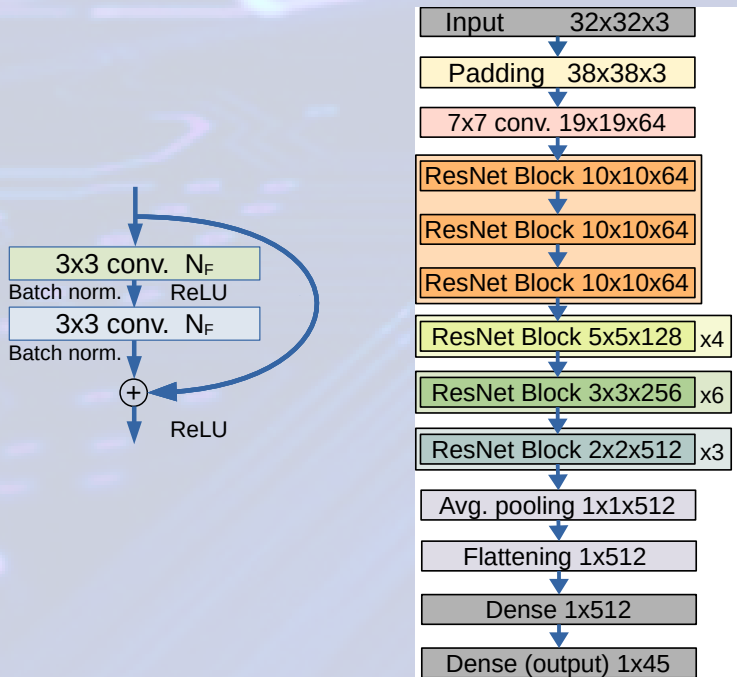
Popular architectures

Stacking more layers: solve complex problems more efficiently, get highly accurate results

BUT:

Vanishing/exploding gradients

ResNet: Residual blocks with “skip connections” (SOTA image classifier of 2015)



Machine Learning in HEP

A Living Review of Machine Learning for Particle Physics

<https://iml-wg.github.io/HEPML-LivingReview/>

Matthew Feickert, Benjamin Nachman, arXiv:2102.02770

2021 May: **417** references

2021 November: **568** references

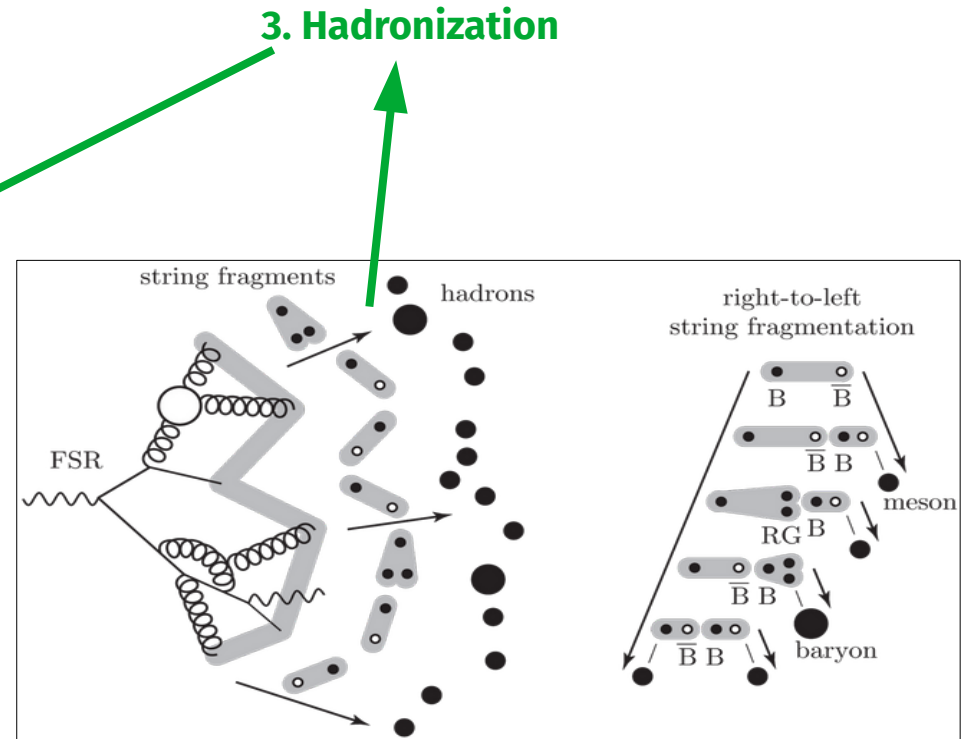
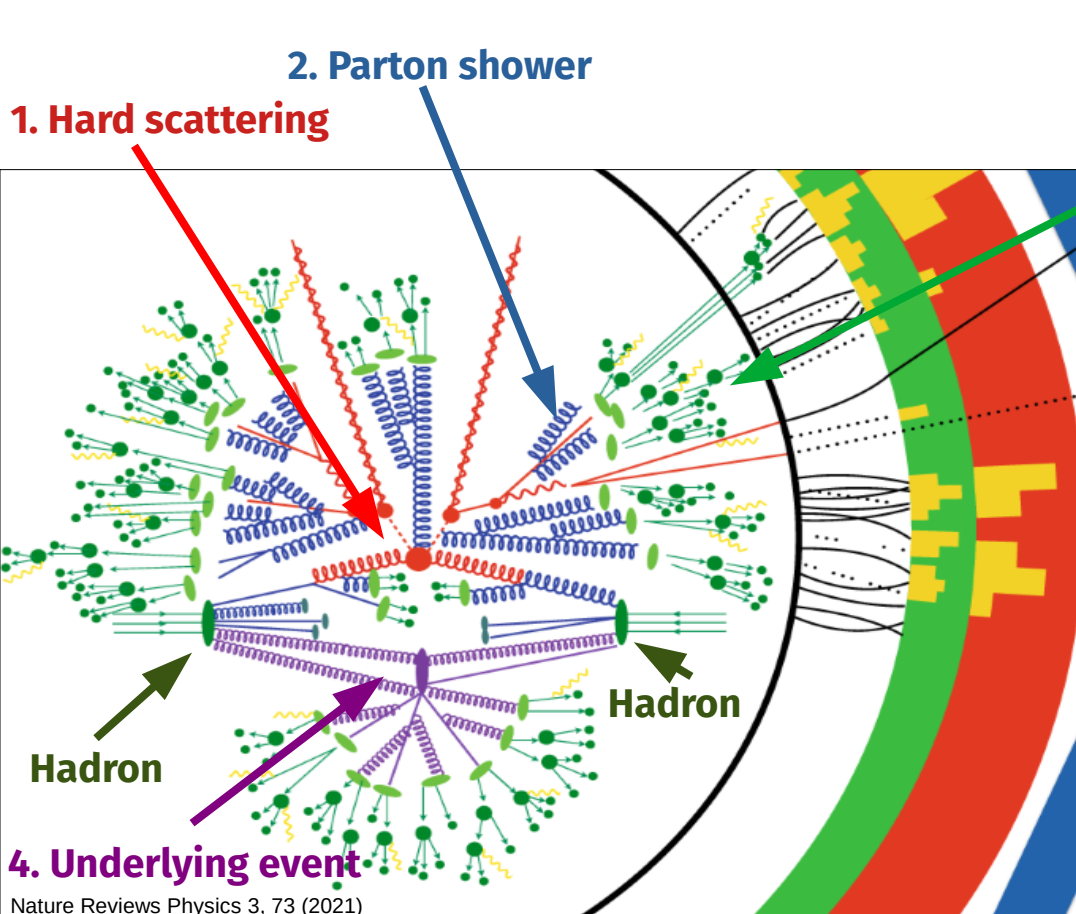
2022 October: **724** references

Today: **849** references

- Track reconstruction
- Quark/gluon jet separation
- Jet reconstruction
- Tuning Monte Carlo event generators
- GAN of detectors

- Accelerated Charged Particle Tracking with Graph Neural Networks on FPGAs
- Particle Track Reconstruction using Geometric Deep Learning
- Jet tagging in the Lund plane with graph networks [DOI]
- Vertex and Energy Reconstruction in JUNO with Machine Learning Methods
- MLPM: Efficient machine-learned particle flow reconstruction using graph neural networks
- 25th International Conference on Computing in High Energy and Nuclear Physics
- 25th International Conference on Computing in High Energy and Nuclear Physics
- Graph Neural Network for Object Reconstruction in Liquid Argon Time Projection Chambers
- Instance Segmentation GNNs for One-Shot Confirmed Tracking at the LHC
- Charged particle tracking via edge-classifying interaction networks
- Jet characterization in Heavy Ion Collisions by QCD-Aware Graph Neural Networks
- Graph Generative Models for Final Detector Simulations in High Energy Physics
- Segmentation of EM showers for neutrino experiments with deep graph neural networks
- Sets (sets (sets))
- Energy Flow Networks: Deep Sets for Particle Jets [DOI]
- ParticleNet: Jet Tagging via Particle Clouds [DOI]
- ARCHER: An attention-based method for particle tagging [DOI]
- Secondary Vertex Finding in Jets with Neural Networks
- Equivalent Energy Flow Networks for Jet Tagging
- Remotely-Supervised Many-Jet Event Reconstruction with Symmetry-Preserving Attention Networks
- Zero-Permutation Jet Partition Assignment using a Self-Attention Network
- Learning to Isolate Muons
- Point Cloud Transformers applied to Collider Physics
- Physics-inspired tasks
- Automating the Construction of Jet Observables with Machine Learning [DOI]
- How Much Information is in a Jet? [DOI]
- Novel Jet Observables from Machine Learning [DOI]
- Energy flow polynomials: A complete linear basis for jet substructure [DOI]
- DeepLearning Top Tagging with a Lorentz Learner [DOI]
- Reconstructing Stripes with kinematic shapes
- SWI2S tagging
- Jet images — deep learning edition [DOI]
- Parton Shower Uncertainties in Jet Substructure Analysis with Deep Neural Networks [DOI]
- QCD-Aware Recursive Neural Networks for Jet Physics [DOI]
- Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques [DOI]
- Boosted SWI and S2S tagging with jet charge and deep learning [DOI]
- Supersymmetry Jet Classification with Graph Neural Networks for Lorentz Boosted Bosons [DOI]
- Jet tagging in the Lund plane with graph networks [DOI]
- A BSM+QCD polarization analyzer from Deep Neural Networks
- Stringnetwise (tree)S
- Automating the Construction of Jet Observables with Machine Learning [DOI]
- Boosting Stripes Inside with Machine Learning [DOI]
- Interaction networks for the identification of boosted Stripes [tree]S decays [DOI]
- Interactable deep learning for two-prong jet classification with jet spectra [DOI]
- Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques [DOI]
- Disentangling Boosted Higgs Boson Production Modes with Machine Learning
- Benchmarking Machine Learning Techniques with Omega Production at the LHC
- The Boosted Higgs Jet Reconstruction via Graph Neural Network
- Enhancing Signals of Higgs Boson From Background Noise Using Deep Neural Networks
- Learning to increase matching efficiency in identifying additional b-jets in the Stripes [tree]S [tree]S [tree]S process
- quarks and gluons
- Quark versus Gluon Jet Tagging Using Jet Images with the ATLAS Detector
- Deep learning in color: towards automated quark/gluon [DOI]
- Recursive Neural Networks in Quark/Gluon Tagging [DOI]
- DeepJet: Generic physics object based jet multiclass classification for LHC experiments
- Probing heavy ion collisions using quark and gluon jet substructure
- JEDI-net: a jet identification algorithm based on interaction networks [DOI]
- Quark-Gluon Tagging: Machine Learning vs Detector [DOI]
- Neutrino Machine Learning Analysis for Jet Substructure [DOI]
- Quark-Gluon Jet Reconstruction with ResNet: Reconstruct 1 primary state
- Classification
- Parametrized neural networks for high-energy physics [DOI]
- Approximating Likelihood Ratios with Calibrated Discriminative Classifiers
- E-Fluxes Unleash Machine Learning from Many Collider Events at Once
- Jet images
- How to tell quark jets from gluon jets
- Jet-Images: Computer Vision Inspired Techniques for Jet Tagging [DOI]
- Playing Top with ANN: Boosted Top Identification with Pattern Recognition [DOI]
- Jet-Images — deep learning edition [DOI]
- Quark versus Gluon Jet Tagging Using Jet Images with the ATLAS Detector
- Boosting Stripes Inside with Machine Learning [DOI]
- Learning to classify from sparse samples with high-dimensional data [DOI]
- Parton Shower Uncertainties in Jet Substructure Analyses with Deep Neural Networks [DOI]
- Deep learning in color: towards automated quark/gluon [DOI]
- Deep-learning Top Taggers or The End of QCD? [DOI]
- Pulling Out All the Stops with Computer Vision and Deep Learning [DOI]
- Reconstructing boosted Higgs jets from event image segmentation
- An Attention Based Neural Network for Jet Tagging
- Quark-Gluon Jet Discrimination Using Convolutional Neural Networks [DOI]
- Learning to Isolate Muons
- Deep learning jet modifications in heavy-ion collisions
- Event images
- Topology classification with deep learning to improve real-time event selection at the LHC [DOI]
- Convolutional Neural Networks with Event Images for Pileup Mitigation with the ATLAS Detector
- Boosting Stripes Inside with Machine Learning [DOI]
- End-to-End Physics Event Classification with the CMS Open Data: Applying Image-based Deep Learning on Detector Data to Directly Classify Collision Events at the LHC [DOI]
- Disentangling Boosted Higgs Boson Production Modes with Machine Learning
- Identifying the nature of the QCD transition in relativistic collision of heavy nuclei with deep learning [DOI]
- Sequences
- Jet Flavor Classification in High-Energy Physics with Deep Neural Networks [DOI]
- Topology classification with deep learning to improve real-time event selection at the LHC [DOI]
- Jet Flavor Classification Using DeepJet [DOI]
- Development of a Vertex Finding Algorithm using Recurrent Neural Network
- Sequence-based Machine Learning Models in Jet Physics
- Tiers
- QCD-Aware Recursive Neural Networks for Jet Physics [DOI]
- Recursive Neural Networks in Quark/Gluon Tagging [DOI]
- Graphs
- Neural Message Passing for Jet Physics
- Graph Neural Networks for Particle Reconstruction in High Energy Physics detectors
- Probing stop pair production at the LHC with graph neural networks [DOI]
- Pileup mitigation at the Large Hadron Collider with graph neural networks [DOI]
- Unraveling CP property of top-Higgs coupling with graph neural networks at the LHC [DOI]
- JEDI-net: a jet identification algorithm based on interaction networks [DOI]
- Learning representations of irregular particle-detector geometry with distance-weighted graph networks [DOI]
- Interpretable deep learning for reconstructing jet classification with jet spectra [DOI]
- Neural Network-based Top Tagger with Two-Point Energy Correlators and Geometry of Soft Emissions [DOI]
- Probing triple Higgs coupling with machine learning at the LHC
- Casting a graph net to catch dark showers [DOI]
- Graph neural networks in particle physics [DOI]
- Distance-Weighted Graph Neural Networks on FPGAs for Real-Time Particle Reconstruction in High Energy Physics [DOI]
- Subervised Jet Classification with Graph Neural Networks for Lorentz Boosted Bosons [DOI]
- Track Seedling and Labeling with Embedding-space Graph Neural Networks
- Graph neural network for 3D classification of antiquities and optical crosslink in scintillator-based neutrino detectors [DOI]

Parton shower and hadronization



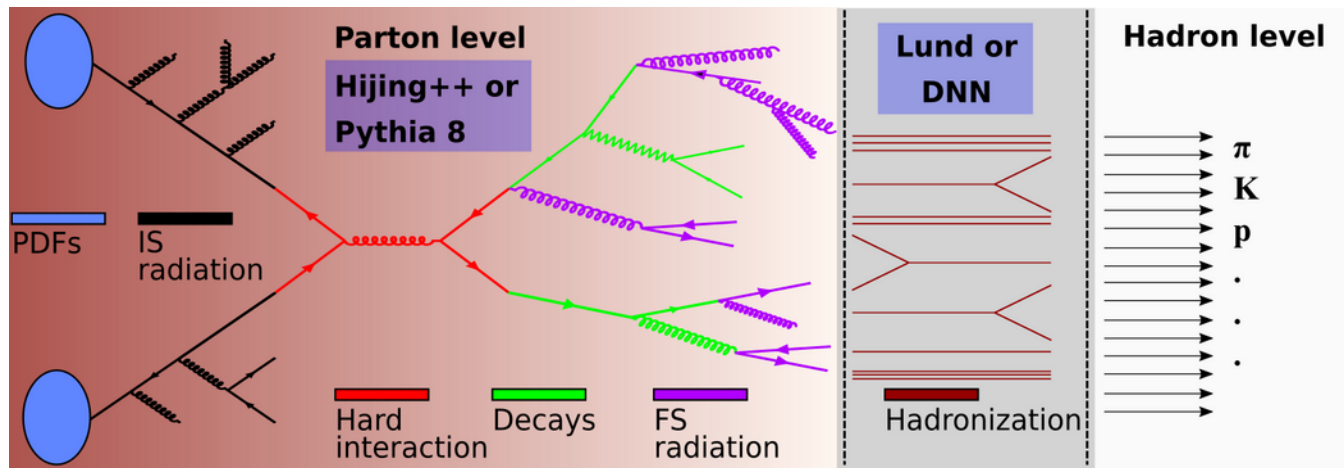
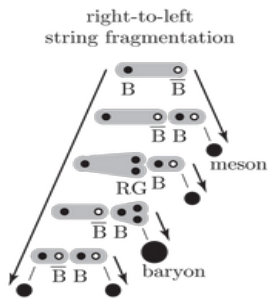
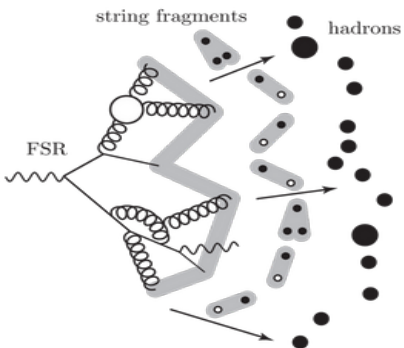
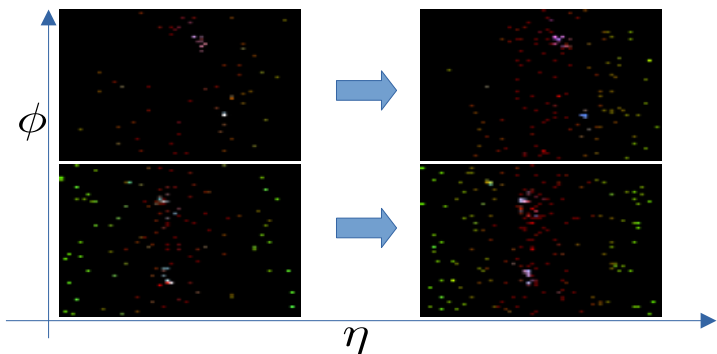
Hadronization

Partons → hadrons

Non-perturbative process

Lund-fragmentation (Comput.Phys.Commun. 27 (1982) 243)

$$f(z) = \frac{1}{z} (1-z)^a e^{-\frac{bm_T^2}{z}}$$



Train and validation sets

Monte Carlo data: Pythia 8.303

Monash tune

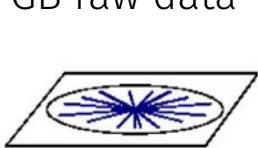
Rescattering and decays turned off
ISR, FSR, MPI: turned on (*)

Selection:

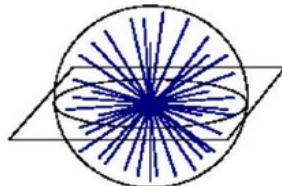
- All final particles with $|y| < \pi$
- At least 2 jets
 - Anti- k_T
 - $R=0.4$
 - $p_T > 40$ GeV

Event number:

- Train: 750 000, $\sqrt{s} = 7$ TeV
- Validation and test: 100 000
- ~20 GB raw data



S=3/4 A=0



S=1 A=1/2

Input:

Parton level

Discretized in the (y, ϕ) plane: p_T , m , multiplicity $\times \sqrt{s}/1\text{GeV}$

$y \in [\pi, \pi]$ 32 bins

$\phi \in [0, 2\pi]$, 32 bins

Hadron level output:

(Charged) event multiplicity, (tr-)sphericity, mean jet p_T , -mass, -width, -multiplicity

$$M_{xyz} = \sum_i \begin{pmatrix} p_{xi}^2 & p_{xi}p_{yi} & p_{xi}p_{zi} \\ p_{yi}p_{xi} & p_{yi}^2 & p_{yi}p_{zi} \\ p_{zi}p_{xi} & p_{zi}p_{yi} & p_{zi}^2 \end{pmatrix}$$

Eigenvalues:

$$\lambda_1 > \lambda_2 > \lambda_3 \quad \sum_i \lambda_i = 1$$

Sphericity:

$$S = \frac{3}{2}(\lambda_2 + \lambda_3)$$

Transverse sphericity:

$$S_{\perp} = \frac{2\lambda_2}{\lambda_1 + \lambda_2}$$

Models

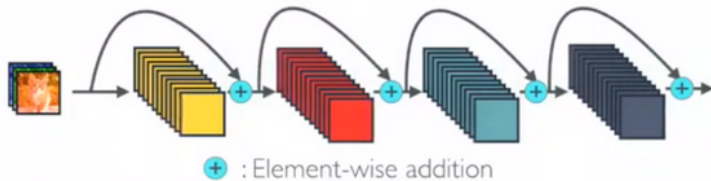
Stacking more layers: solve complex problems more efficiently, get highly accurate results

BUT:

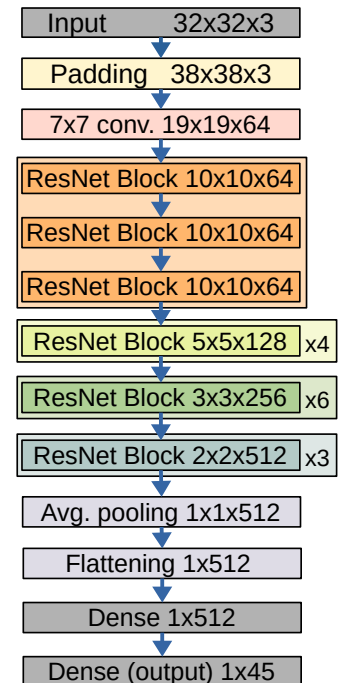
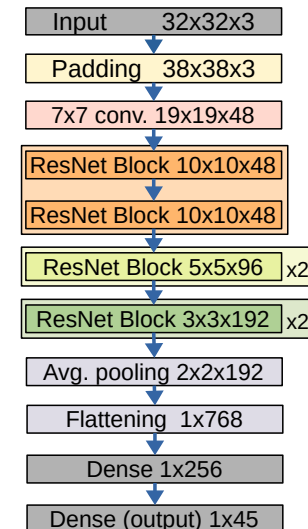
Vanishing/exploding gradients

ResNet:

Residual blocks with “skip connections”



	Model S	Model L
Trainable parameters	1.7 M	20 M

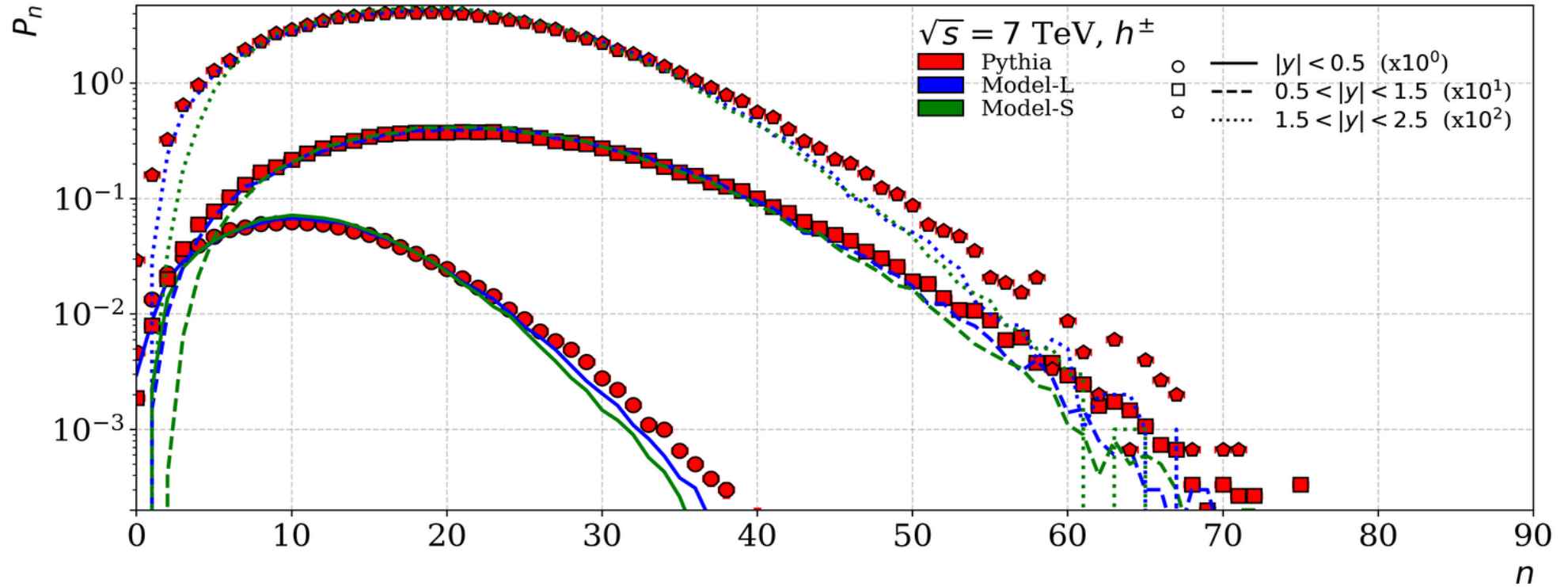


Used hardwares: Nvidia Tesla T4, GeForce GTX 1080
@ Wigner Scientific Computing Laboratory

Framework: Tensorflow 2.4.1, Keras 2.4.0

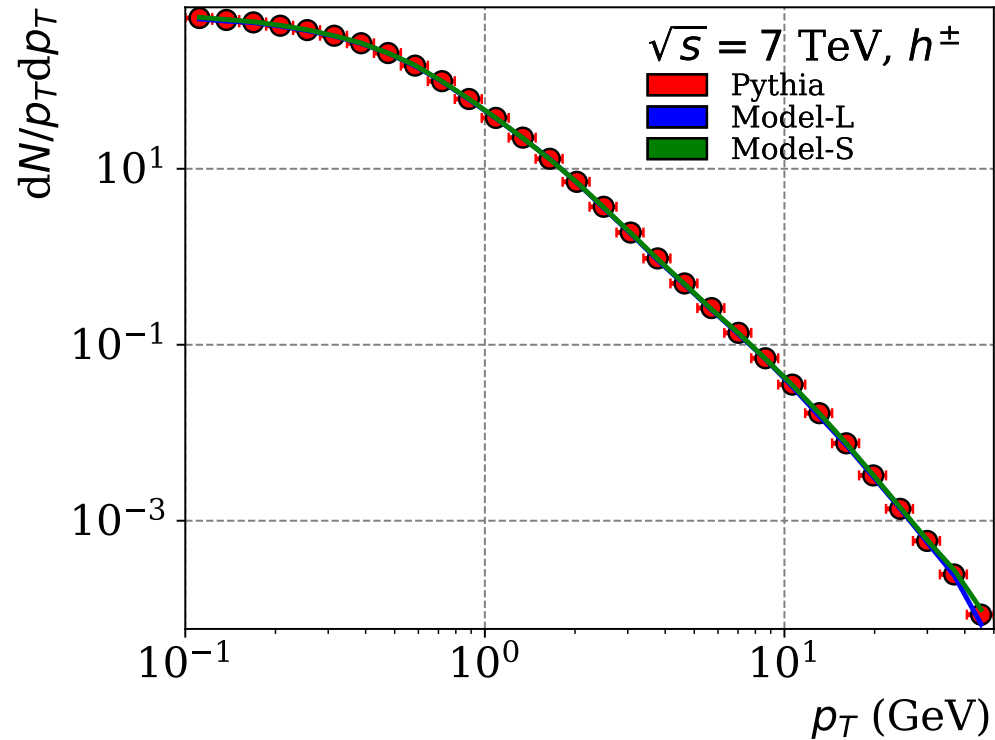
Results

Proton-proton @ 7 TeV, Training + Validation

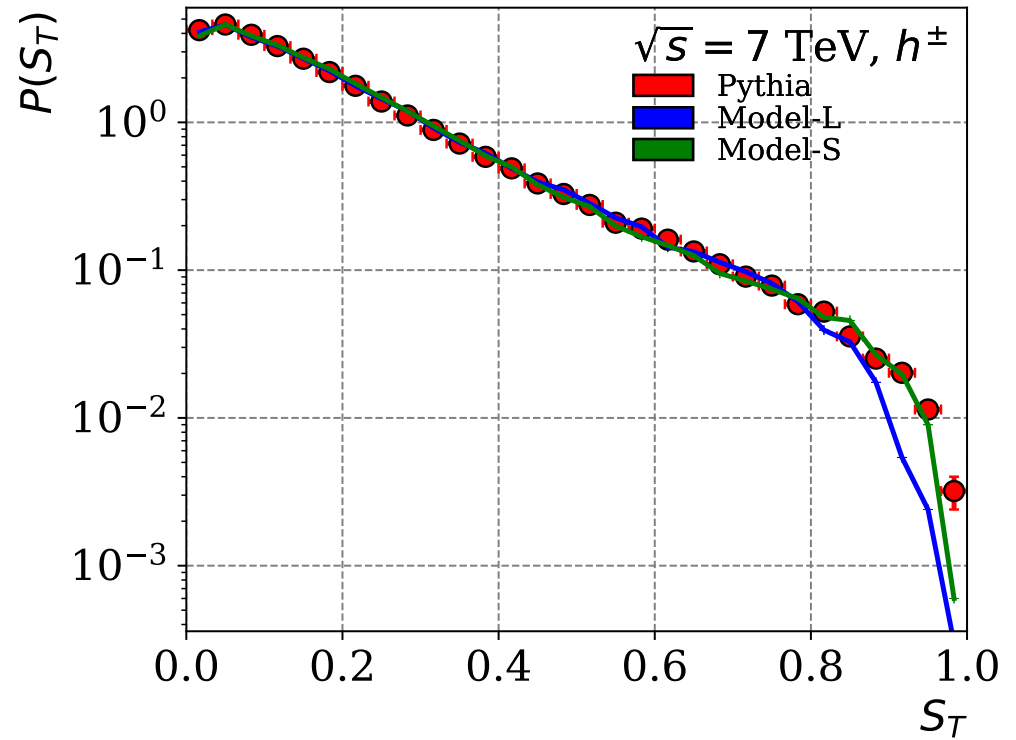


Charged hadron multiplicity at various rapidity windows
Comparison to reference MC model
Good agreement for both models

Proton-proton @ 7 TeV, Training + Validation

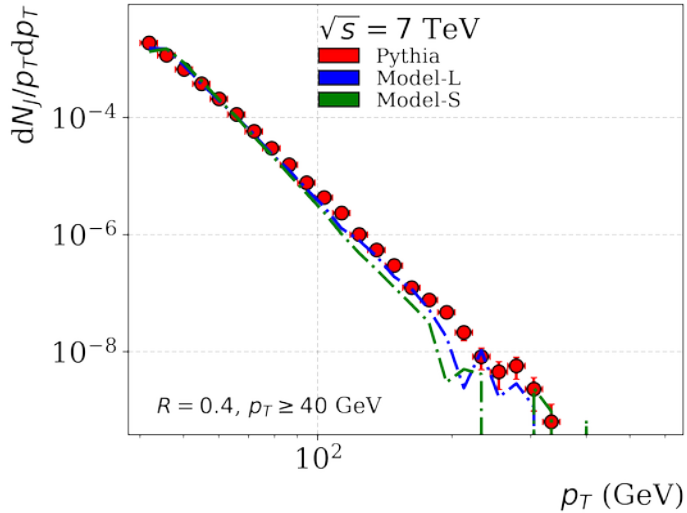


Charged hadron transverse momentum
 $0.1 \text{ GeV} \leq p_T \leq 50 \text{ GeV}$



Event transverse sphericity
The smaller model performs better

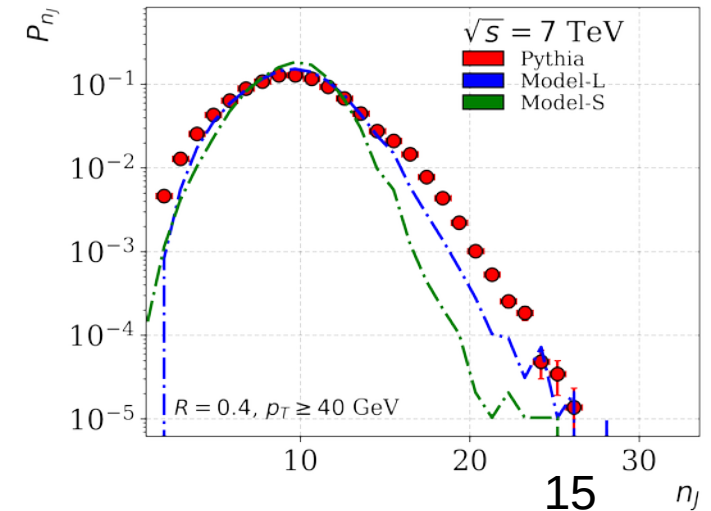
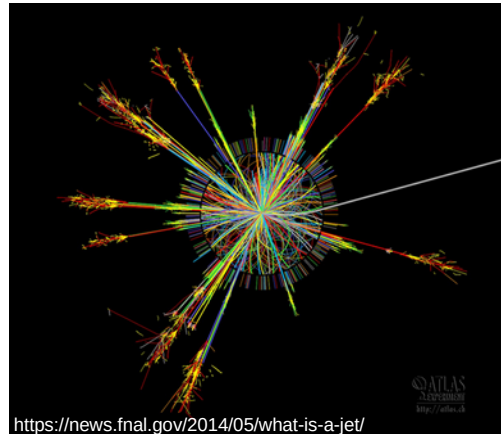
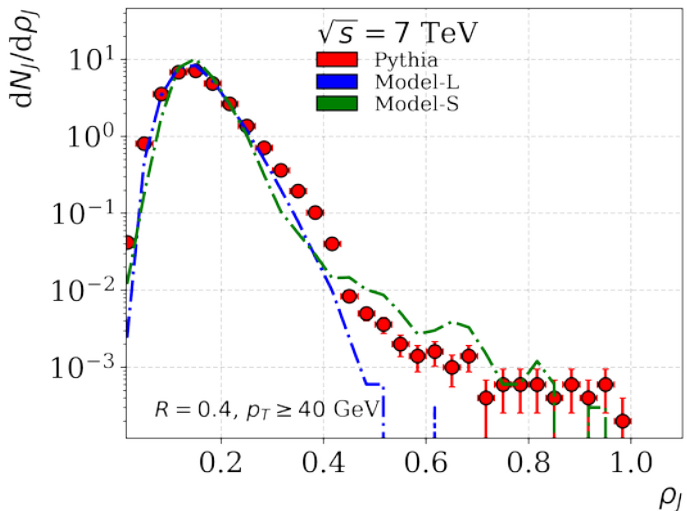
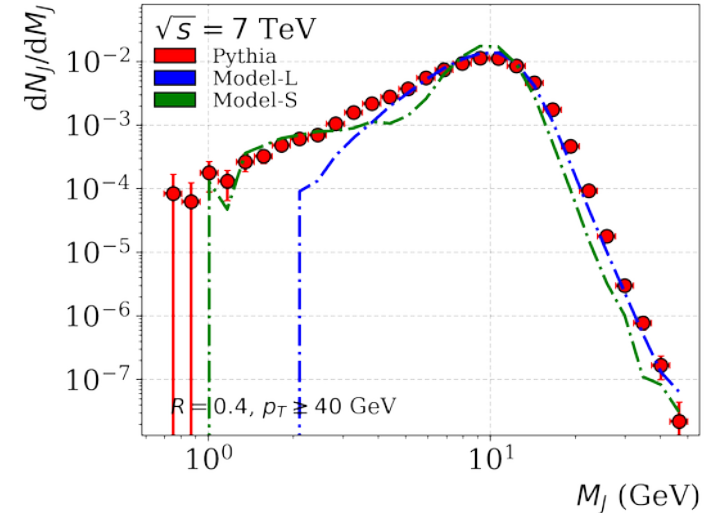
Proton-proton @ 7 TeV, Training + Validation



Jets:

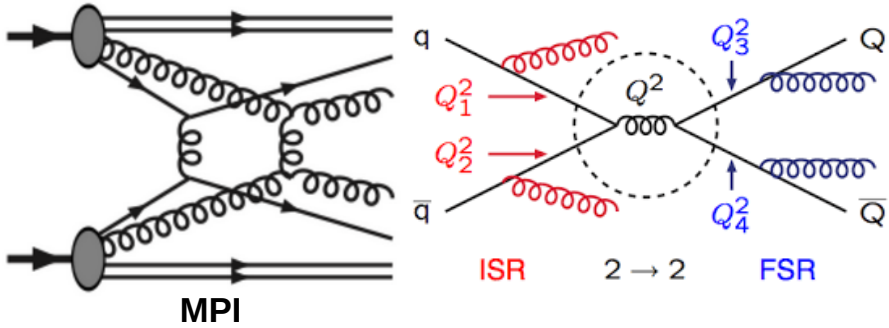
- Mean $p_T \leq 400$ GeV
- Mean mass $p_T \leq 400$ GeV
- Mean multiplicity
- Mean width

The smaller model performs better

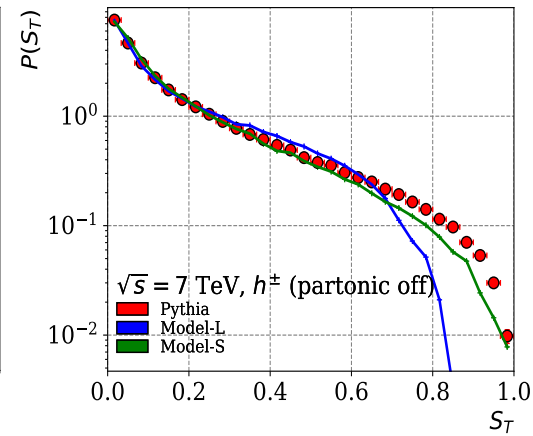
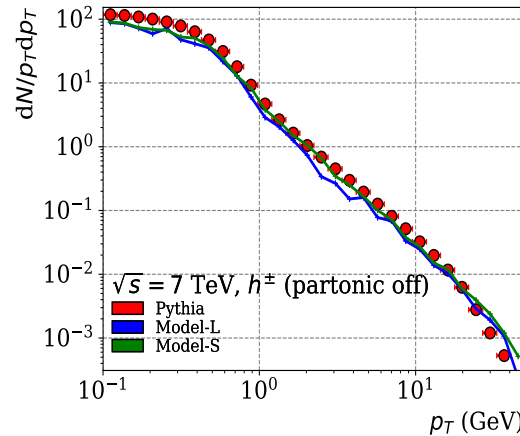
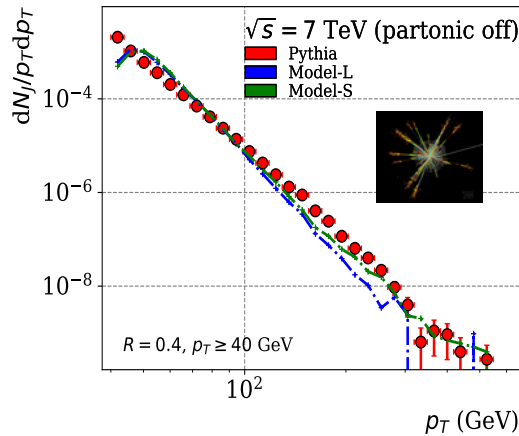
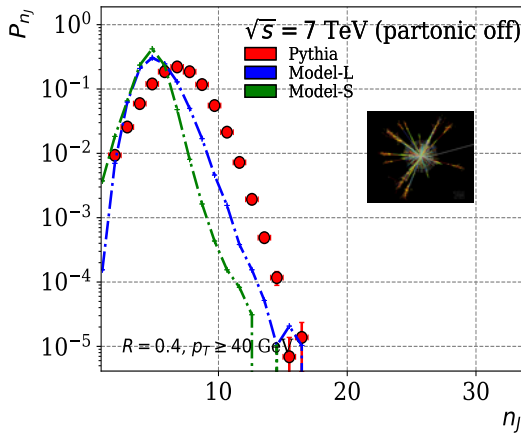
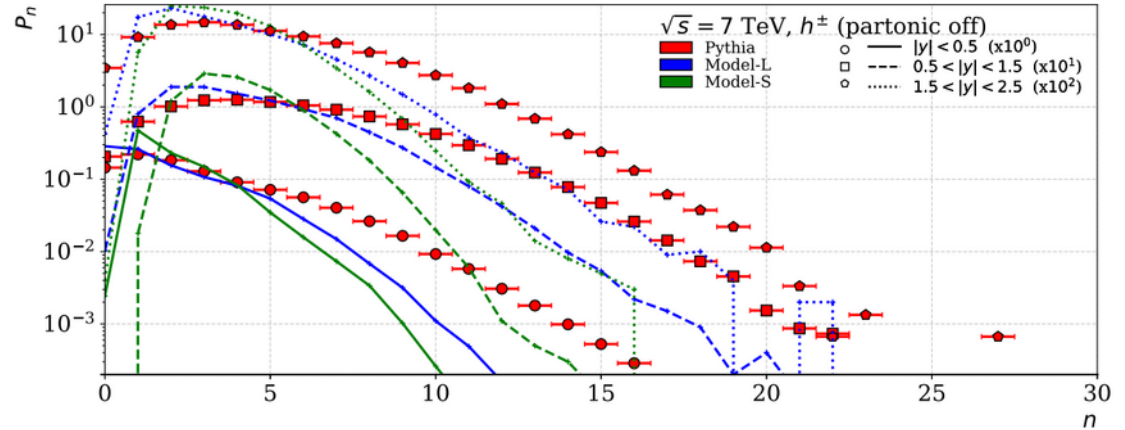


Proton-proton @ 7 TeV, Training + Validation

(* What about the partonic processes?)

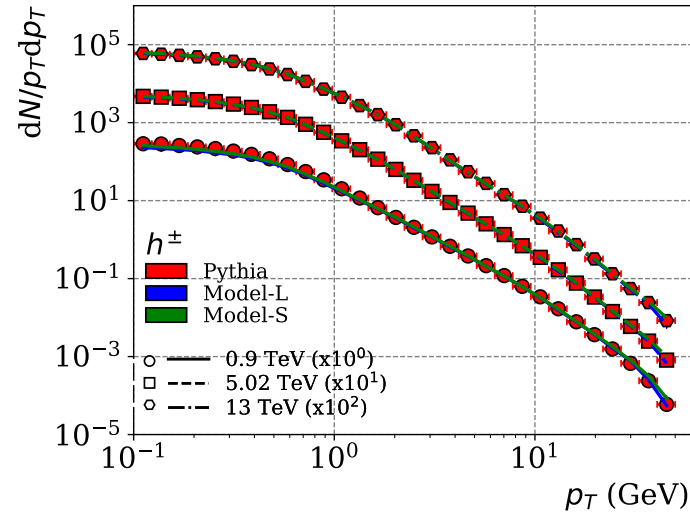
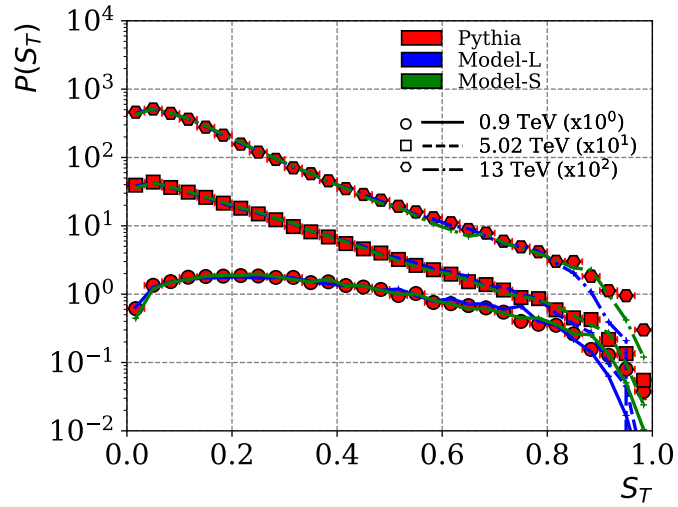
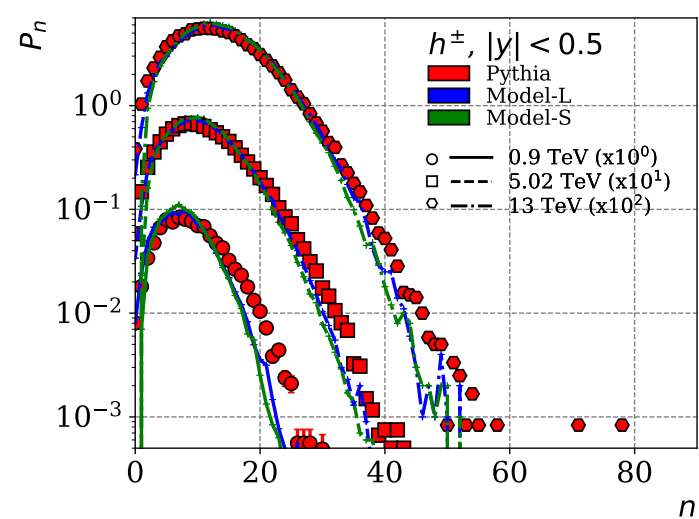


<http://home.thep.lu.se/~torbjorn/talks/cern18cosmic.pdf>

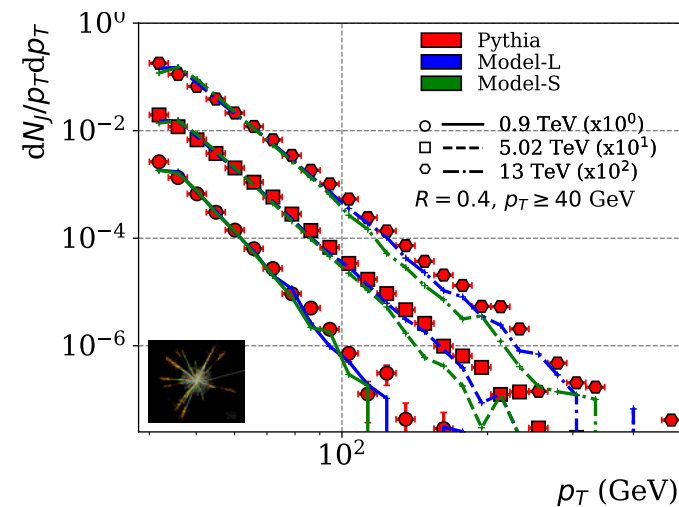


Qualitative agreement \rightarrow the models adopted the hadronization properties

Proton-proton @ 0.9-13 TeV, Predictions



- So far: everything at $\sqrt{s} = 7$ TeV \rightarrow the **ONLY** energy, where the models were trained
- Good agreement for all observable quantities as **predictions** for other LHC energies
- **Multiplicity scaling?**



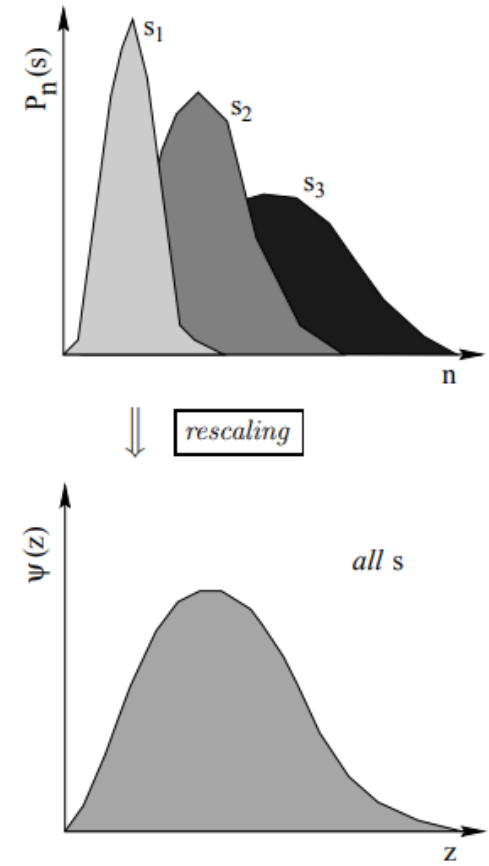
KNO-scaling

The collapse of multiplicity distributions P_n onto a universal scaling curve:

$$P_n = \frac{1}{\langle n \rangle} \Psi \left(\frac{n}{\langle n \rangle} \right)$$

The scale parameters governed by leading particle effects and the growth of average multiplicity

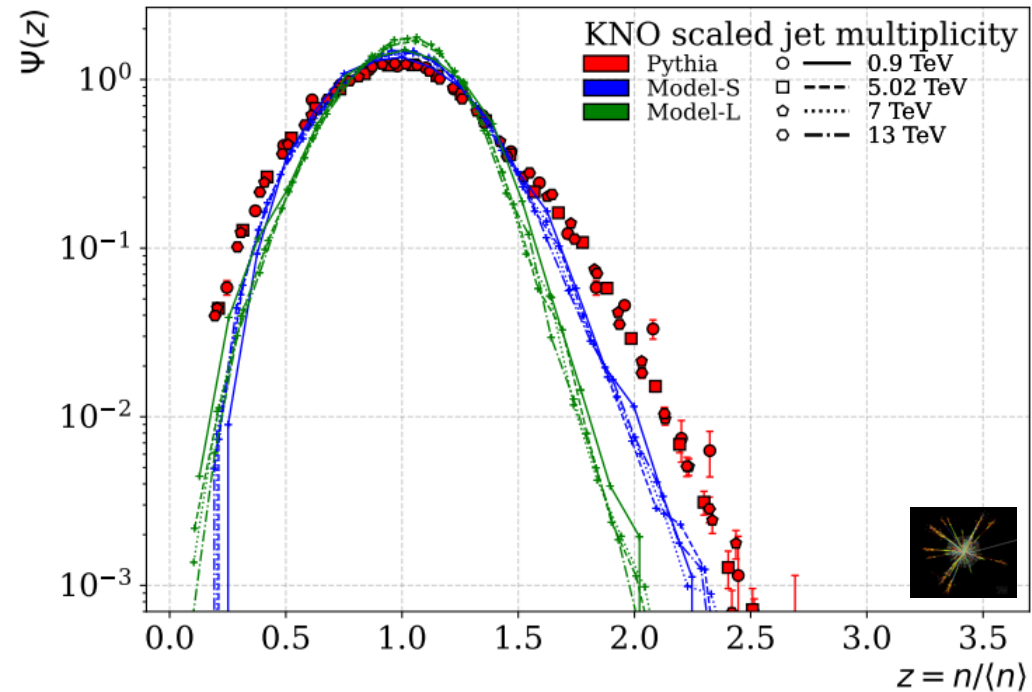
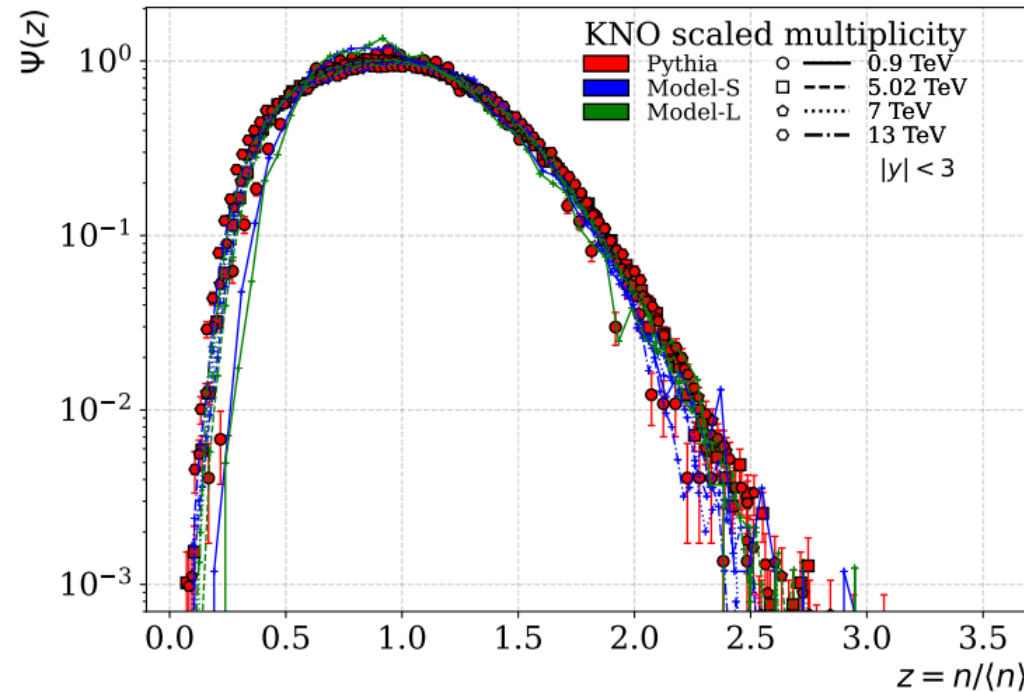
Violation of the scaling at high CM energies: not fully understood (relation to MPI?)



Nuclear Physics B 40 (1972), 317–334.

(Nucl. Phys. B Proc. Suppl. 92 (2001). 122–129)

Test of KNO-scaling for the predictions



Scaling function for multiplicities at various energies: $P_n = \frac{1}{\langle n \rangle} \Psi \left(\frac{n}{\langle n \rangle} \right)$

Charged hadron multiplicities in **jetty** events: good overlap and agreement

Mean jet multiplicities: different scaling for the models

Heavy Ion Jet Interaction Generator (C++ version)

核易经

[Hé -yì -jīng]

A NEW GENERATION OF HEAVY-ION MONTE CARLO

"Nuclear change theory"; Book of Changes, "Originally a divination manual in the Western Zhou period (1000–750 BC)"

First, FORTRAN version: 1991, X.N. Wang, M. Gyulassy, **Phys. Rev. D 44, (1991) 3501.**

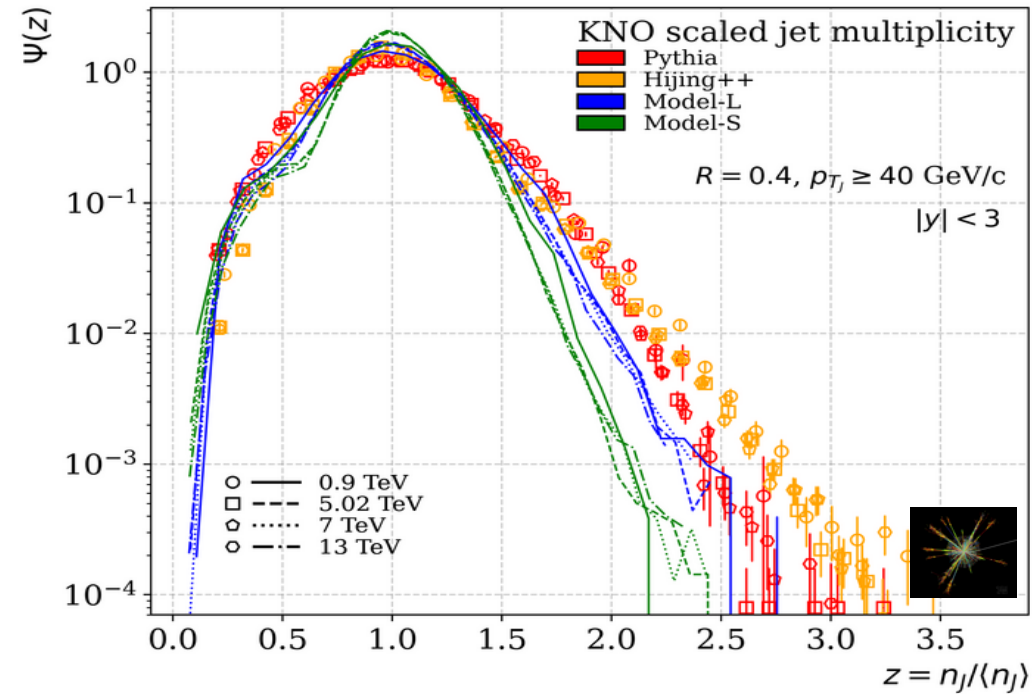
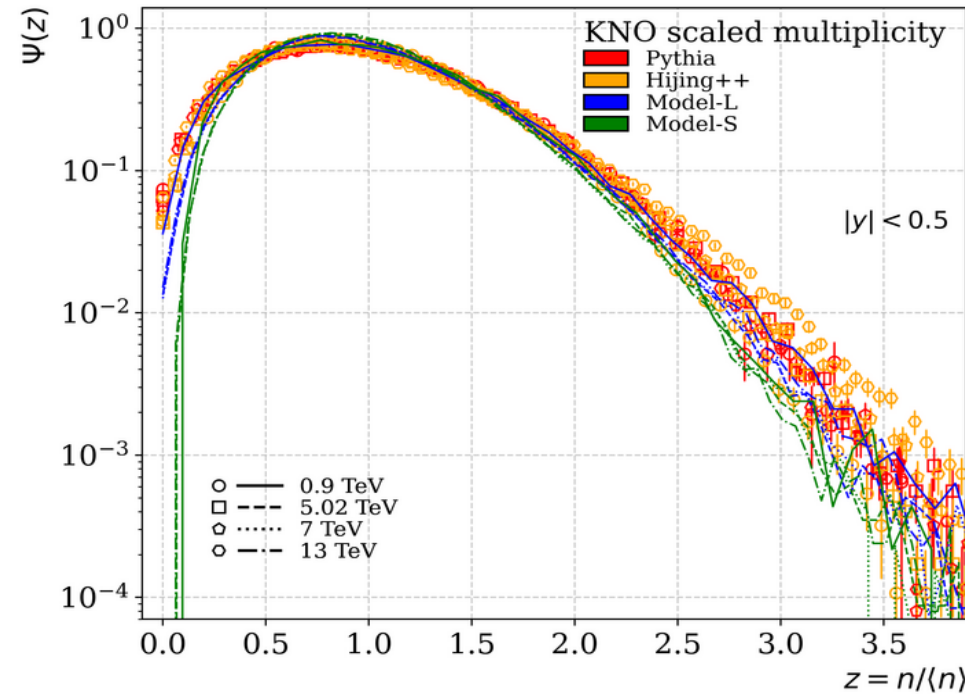
Computational challenge: more than 600 million collision in each second → **HiLumiLHC**: even more

Requirements for a new version: multithreaded mode, maintainability, intuitive usage

	FORTRAN HIJING	HIJING++ v3.0	HIJING++ v3.1
Precision	simple	double	double
Pythia version	5.3	8.2	8.2+
(n)PDF	GRV98lo	LHAPDF6.2	LHAPDF6.2+
Jet quenching	(✓)	(✓)	(✓)
Multithreading	x	x	✓
Analysis interface	x	x	✓
Module management	x	x	✓
Dependencies, build system	Makefile	Makefile	CMake



Test of KNO-scaling for the predictions - **Hijing++**



Scaling function for multiplicities at various energies: $P_n = \frac{1}{\langle n \rangle} \Psi \left(\frac{n}{\langle n \rangle} \right)$

Charged hadron multiplicities in **jetty** events: good overlap and agreement

Mean jet multiplicities: different scaling for the models

Summary

Developed hadronization models with different complexities

Traditional computer vision algorithms capture the main features of high-energy event variables successfully → training only at a **single c.m. energy, predictions at other energies**

Generalization to other CM energies: KNO scaling in jetty events

Valuable input for MC developments

Prospects

Architecture variations (hyperparameter fine-tuning)

Heavy ion (centralities, collective effects)



Thank you for your attention!

The research was supported by OTKA grants K135515, NKFIH 2019-2.1.6-NEMZKI-2019-00011, 2019-2.1.11-TÉT-2019-00078, 2019-2.1.11-TÉT-2019-00050, 2020-2.1.1-ED-2021-00179, the **Wigner Scientific Computing Laboratory** (former Wigner GPU Laboratory) and RRF-2.3.1-21-2022-00004 within the framework of the Artificial Intelligence National Laboratory.

Dimensionality

Input:

Parton level

Discretized in the (y, ϕ) plane: $p_{T,m}$, multiplicity

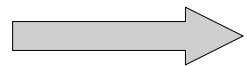
$$\left. \begin{array}{l} y \in [\pi, \pi], \quad 32 \text{ bins} \\ \phi \in [0, 2\pi], \quad 32 \text{ bins} \end{array} \right\} := M$$

Reduction with Singular Value Decomposition:

$$M_{n \times m} = U_{n \times n} \Sigma_{n \times m} V_{m \times m}^T$$

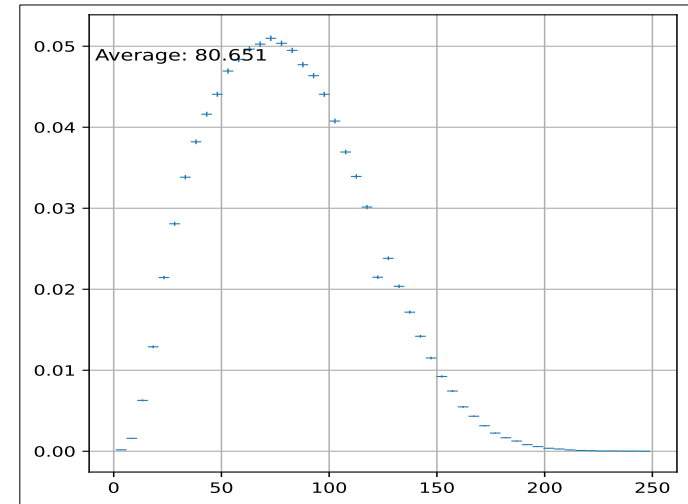
- Unitarity
- Ordered by importance
- Guaranteed to exist, unique

$$M \approx \sum_{i=1}^r \sigma_i u_i v_i^T + \mathcal{O}(\epsilon), \quad r \leq \min\{n, m\}$$



Reduce the input to $\mathcal{O}(10^2)$

$\left. \begin{array}{l} \mathcal{O}(10^3 - 10^4) \text{ Total pixels} \\ \text{vs } \mathcal{O}(10^2) \end{array} \right\}$
Pixels with information



doi:10.1007/BF02288367

Dimensionality (work in progress)

