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

Universidad Nacional Autónoma de México

Instituto de Ciencias Nucleares - Seminar

05 05 2023

GÁBOR BÍRÓ
biro.gabor@wigner.hu

Gergely Gábor Barnaföldi
Bence Tankó-Bartalis

arXiv:2111.15655
arXiv:2210.10548
arXiv:2303.05422

History

1950: Alan Turing creates the "Turing Test"

1957: Frank Rosenblatt: the first neural network for computers (the **perceptron**), which simulate the thought processes of the human brain.

1959: Arthur Samuel, IBM: *Machine Learning*

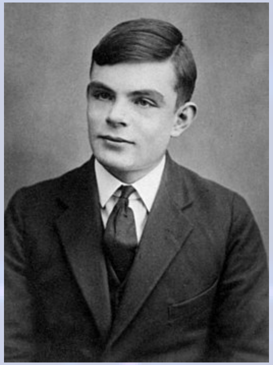
1967: The first general, working learning algorithm for supervised, deep, feedforward, multilayer perceptrons by A. G. Ivakhnenko and V. G. Lapa

1986: First mention of *Deep Learning* by Rina Dechter (*Learning While Searching in Constraint-Satisfaction Problems*)

1989: Yann LeCun et al: standard backpropagation algorithm for recognizing handwritten ZIP codes on mail

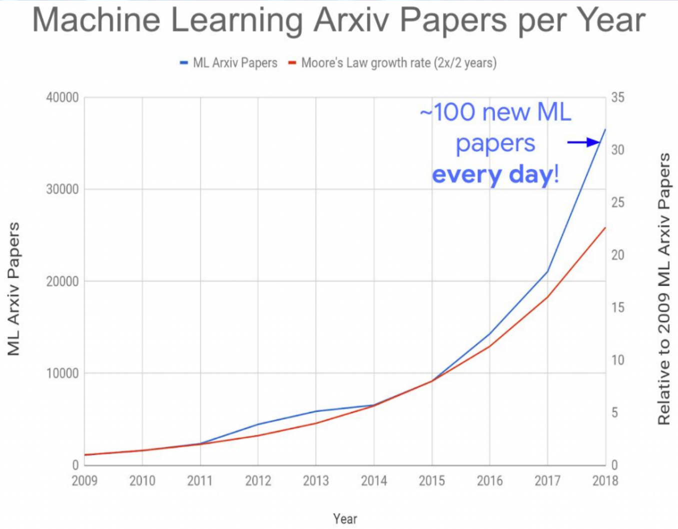
1997: "*A computer program is said to learn from experience **E** with respect to some class of tasks **T** and performance measure **P** if its performance at tasks in **T**, as measured by **P**, improves with experience **E**.*" - Tom M. Mitchell: Machine Learning

1997: IBM's Deep Blue beats Garri Kaszparov (the world champion at chess). Computing capacity: 11.38 GFLOPS, TOP500: 259th (comparison: Nvidia RTX 4090: 82.6 TFLOPS)



Wikipedia

https://researcher.watson.ibm.com/researcher/view_page.php?id=6814



arXiv:1911.05289

History

2009: ImageNet by prof. Fei-Fei Li a database of 14 million labeled images in 2009

2011: IBM's Watson: winner of game show Jeopardy!

2011: Google Brain: cats in Youtube videos

2012: AlexNet by Alex Krizhevsky: first CNN

2013: Word2vec algorithms: foundations for language models

2014: DeepFace by Facebook

2014: Generative adversarial networks (GAN) by Ian Goodfellow

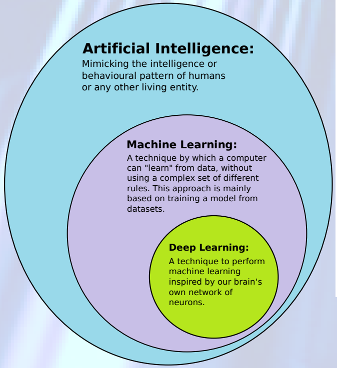
2016: AlphaGo by Deepmind

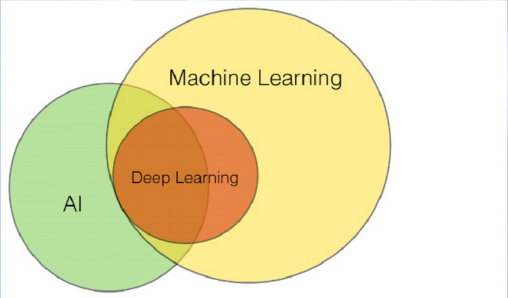
2016: Face2Face (baseline for 'DeepFake')

2017: Waymo: first self-driving car company to operate without human intervention

2018: AlphaFold by Deepmind

2020: GPT-3 by OpenAI to generate human-like text. Trainable parameters: **175 billion**





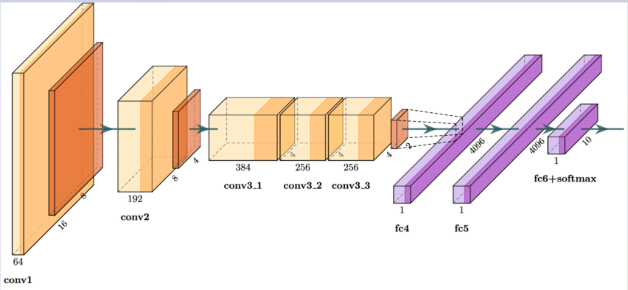
Wikipedia

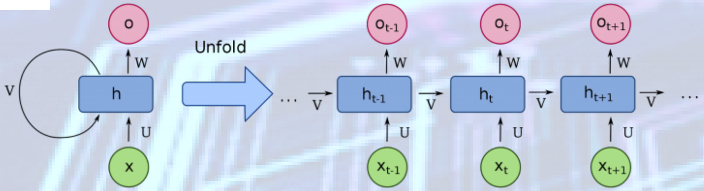
History

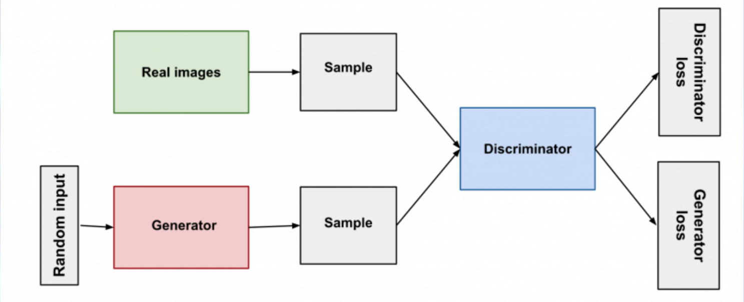
CNN (image classification, object detection, recommender systems)...

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

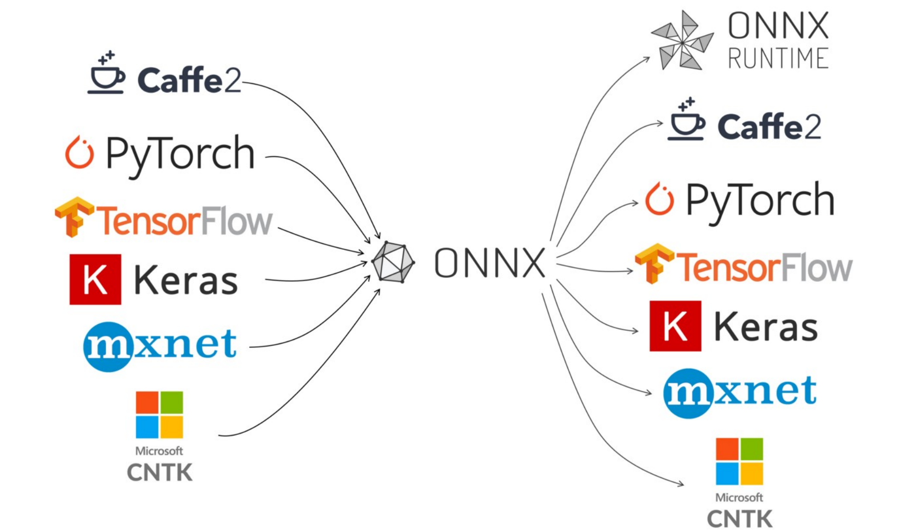
Generative models: anomaly detection, pattern recognition, reinforced learning







Various frameworks for training and inference:



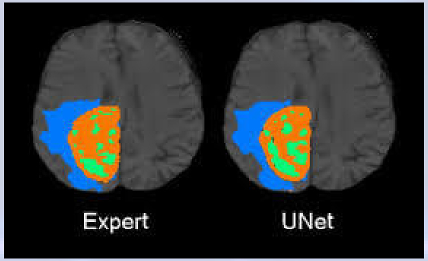
https://towardsdatascience.com/onnx-preventing-framework-lock-in-9a798fb34c92

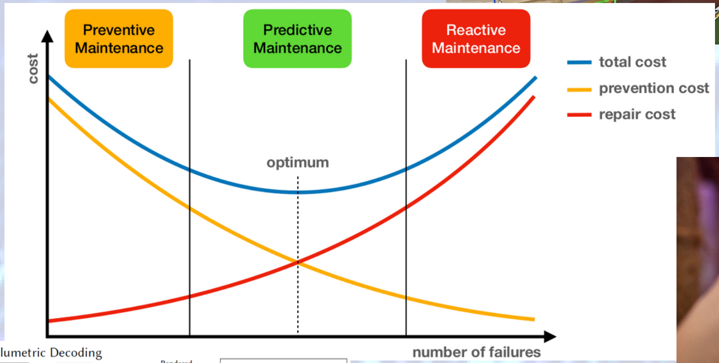
Motivaton – data, data, more data

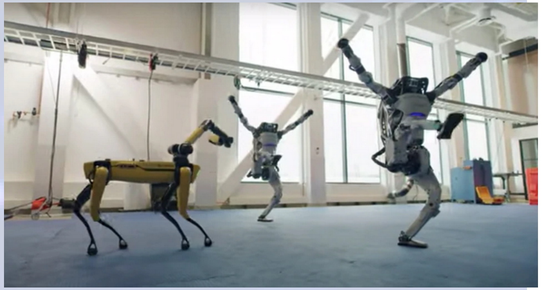
Autonomous driving
Medical imaging
Predictive maintenance
Anomaly detection, fake news detection
Search of BSM physics
Stock price prediction
Natural Language Processing
Virtual Assistants
Virtual reality
Colorization of Black and White Images
Content generation, examples:

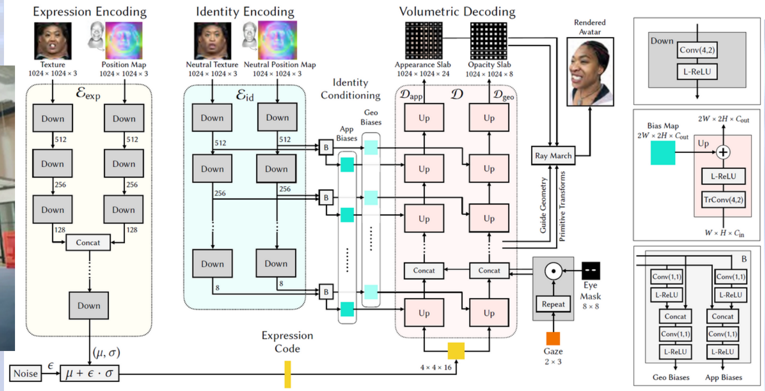
https://infiniteconversation.com/
https://huggingface.co/spaces/stabilityai/stable-diffusion

Robotics
…

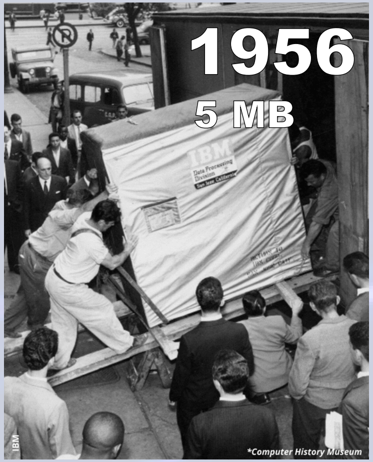


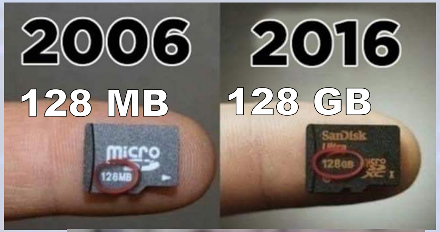






Motivaton – data, data, more data



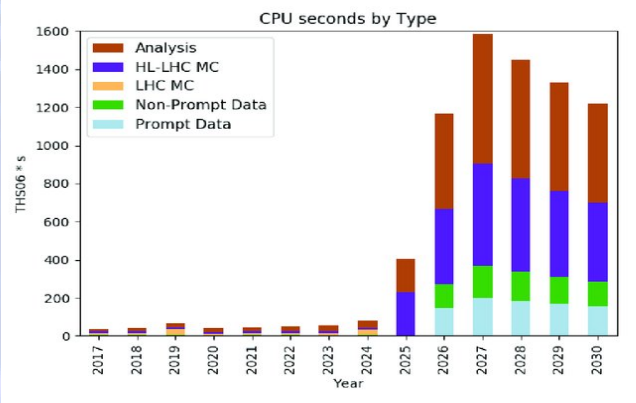


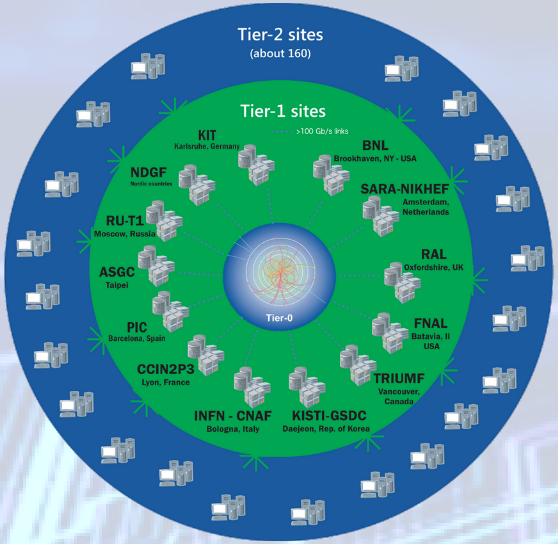
LHC in numbers: **2013** and **now:**

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**

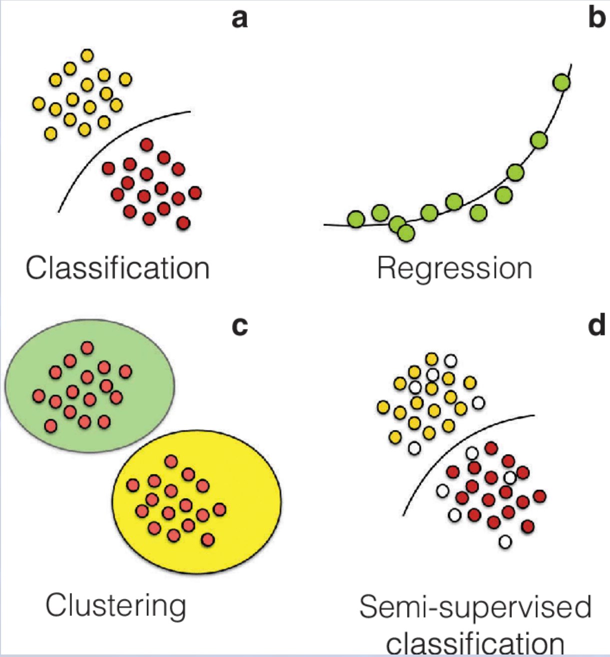
Storing and distributing the data is only one side of the challange

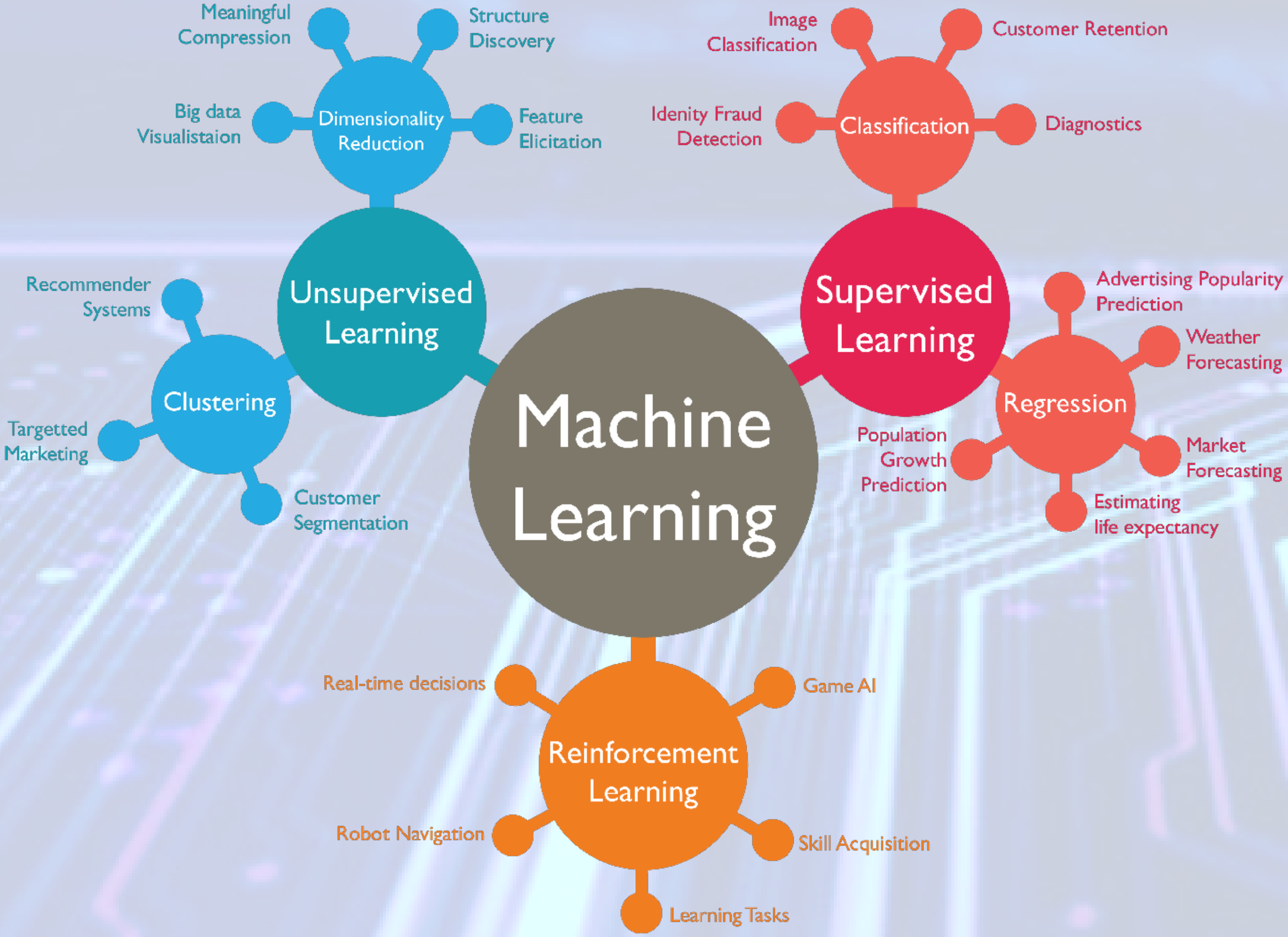
→ **analysis, simulations**





Approaches

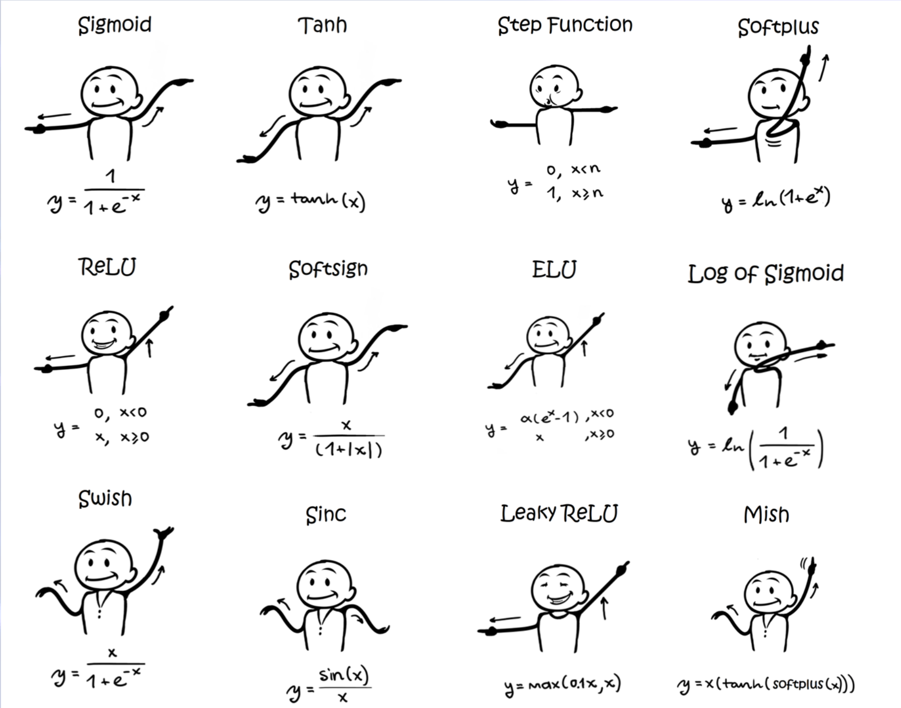

a
b
Classification
Regression
c
d
Clustering
Semi-supervised classification

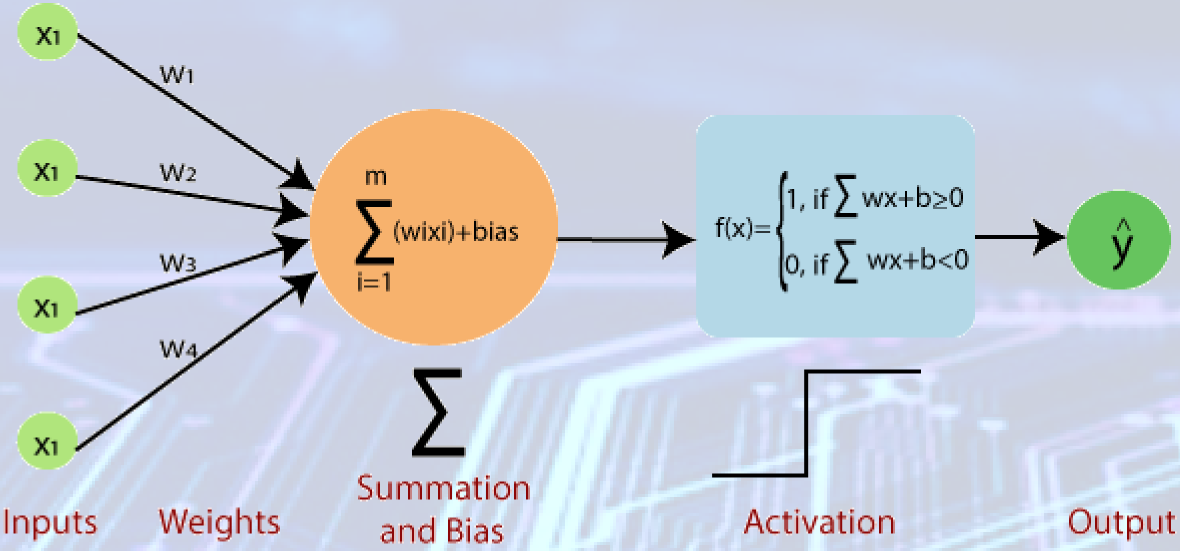

Machine Learning
Unsupervised Learning
Dimensionality Reduction
Meaningful Compression
Structure Discovery
Big data Visualistaion
Feature Elicitation
Clustering
Recommender Systems
Targetted Marketing
Customer Segmentation
Supervised Learning
Classification
Image Classification
Customer Retention
Idenity Fraud Detection
Diagnostics
Regression
Advertising Popularity Prediction
Weather Forecasting
Population Growth Prediction
Market Forecasting
Estimating life expectancy
Reinforcement Learning
Real-time decisions
Game AI
Robot Navigation
Skill Acquisition
Learning Tasks

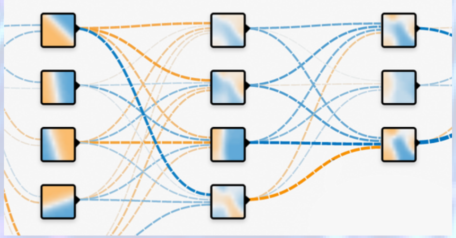
Main ingredients

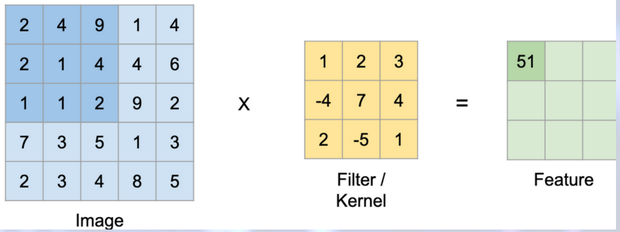
Perceptrons:

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





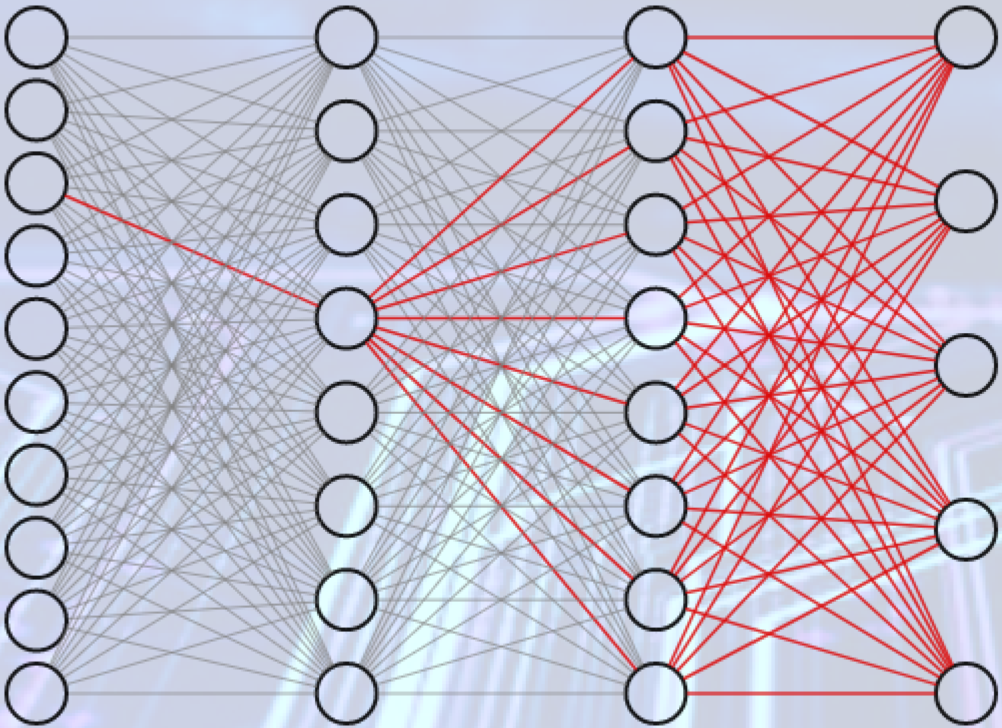




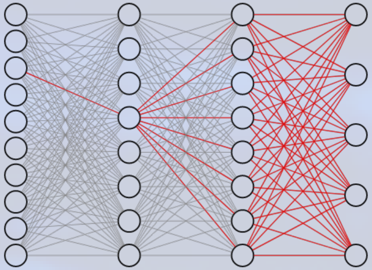
Other important components: pooling layers, regularization and normalization, recurrent layers...

https://sefiks.com/2020/02/02/dance-moves-of-deep-learning-activation-functions/

Main ingredients



Main ingredients

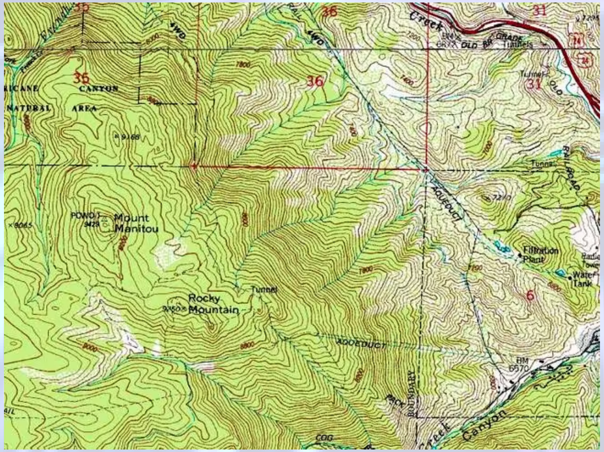


The *learning* part: optimizing “somehow” the weights (Curse of dimensionality)

Loss function: $\mathcal{L} = \frac{1}{n}\sum_i (y_i - f(x_i))^2 := \frac{1}{n}\sum_i (y_i - (mx_i + b))^2$

The gradient descent method:

1) Start with random weights
2) Evaluate the loss
3) Figure out which direction the loss function steeps downward the most (with respect to changing the parameters)
4) Repeat from 2)



Gradient of the loss function with respect to all of the parameters

$$\frac{\partial \mathcal{L}}{\partial m} = \frac{2}{n}\sum_i x_i \cdot (y_i - (mx_i + b))$$

$$\frac{\partial \mathcal{L}}{\partial b} = \frac{2}{n}\sum_i (y_i - (mx_i + b))$$

$$\Rightarrow$$

$$m := m - \alpha \cdot \frac{\partial \mathcal{L}}{\partial m}$$

$$b := b - \alpha \cdot \frac{\partial \mathcal{L}}{\partial b}$$

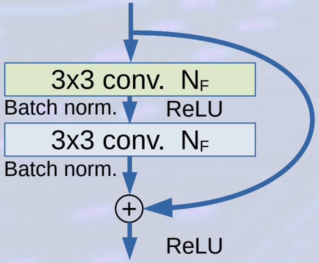
Popular architectures

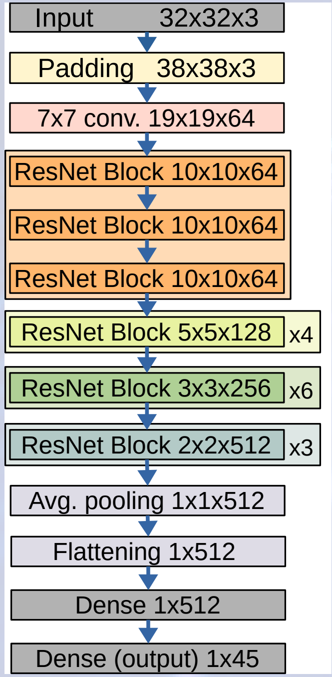
Stacking more layers: solve complex problems more efficiently, get highly accurate results

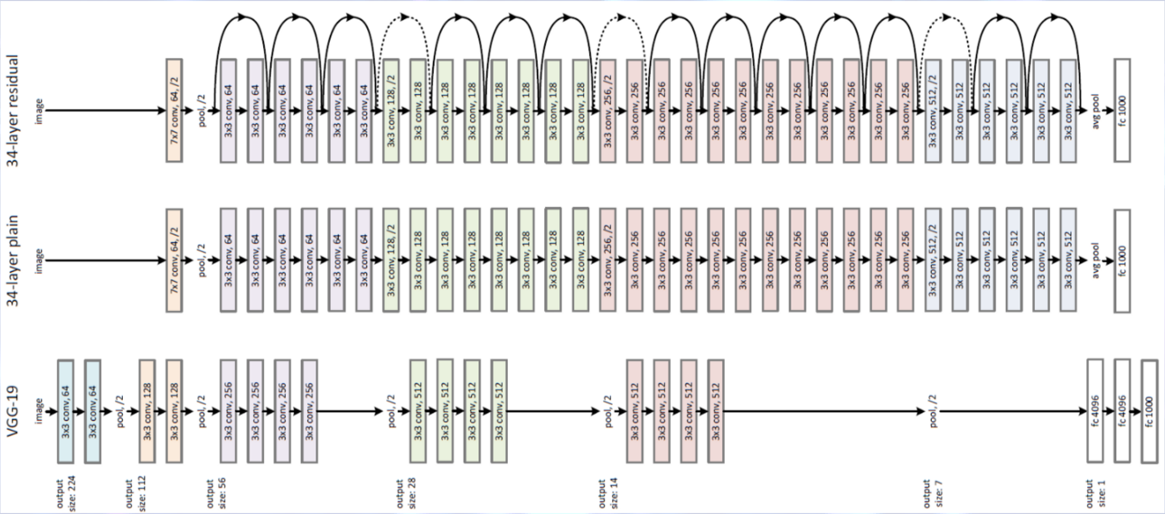
BUT:

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: **759** 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
- MLPF: 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 Conformal 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 Fast Detector Simulations in High Energy Physics
- Segmentation of EM showers for neutrino experiments with deep graph neural networks

- Sets (point clouds)
  - Energy Flow Networks: Deep Sets for Particle Jets [DOI]
  - ParticleNet: Jet Tagging via Particle Clouds [DOI]
  - ABCNet: An attention-based method for particle tagging [DOI]
  - Secondary Vertex Finding in Jets with Neural Networks
  - Equivariant Energy Flow Networks for Jet Tagging
  - Permutationless Many-Jet Event Reconstruction with Symmetry Preserving Attention Networks
  - Zero-Permutation Jet-Parton Assignment using a Self-Attention Network
  - Learning to Isolate Muons
  - Point Cloud Transformers applied to Collider Physics
- Physics-inspired basis
  - 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]
  - Deep-learned Top Tagging with a Lorentz Layer [DOI]
  - Resurrecting $b\bar{b}h$ with kinematic shapes
- W/Z tagging
  - Jet-images — deep learning edition [DOI]
  - Parton Shower Uncertainties in Jet Substructure Analyses 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 W and Z tagging with jet charge and deep learning [DOI]
  - Supervised Jet Clustering with Graph Neural Networks for Lorentz Boosted Bosons [DOI]
  - Jet tagging in the Lund plane with graph networks [DOI]
  - A $W^\pm$ polarization analyzer from Deep Neural Networks
- $H\rightarrow b\bar{b$}
  - Automating the Construction of Jet Observables with Machine Learning [DOI]
  - Boosting $H\to b\bar b$ with Machine Learning [DOI]
  - Interaction networks for the identification of boosted $H \rightarrow b\overline{b}$ decays [DOI]
  - Interpretable 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 Di-Higgs Production at the LHC
  - The Boosted Higgs Jet Reconstruction via Graph Neural Network
  - Extracting Signals of Higgs Boson From Background Noise Using Deep Neural Networks
  - Learning to increase matching efficiency in identifying additional b-jets in the $\text{t}\bar{\text{t}}\text{b}\bar{\text{b}}$ 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]
  - Towards Machine Learning Analytics for Jet Substructure [DOI]
  - Quark Gluon Jet Discrimination with Weakly Supervised Learning [DOI]

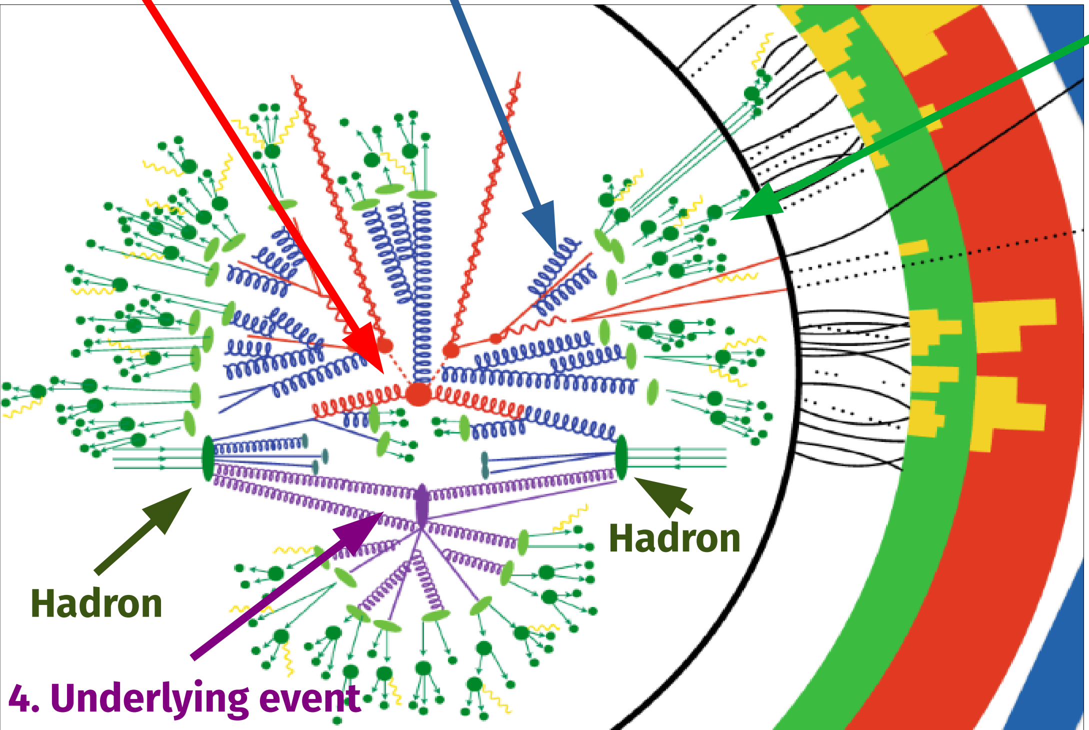
- Classification
  - Parameterized classifiers
    - Parameterized neural networks for high-energy physics [DOI]
    - Approximating Likelihood Ratios with Calibrated Discriminative Classifiers
    - E Pluribus Unum Ex Machina: Learning from Many Collider Events at Once
    - Jet images
      - How to tell quark jets from gluon jets
      - Jet-Images: Computer Vision Inspired Techniques for Jet Tagging [DOI]
      - Playing Tag 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 $H\to b\bar b$ with Machine Learning [DOI]
      - Learning to classify from impure 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 Tops 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 $H\to b\bar b$ 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 Flavour Classification Using DeepJet [DOI]
      - Development of a Vertex Finding Algorithm using Recurrent Neural Network
      - Sequence-based Machine Learning Models in Jet Physics
    - Trees
      - 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]
      - Unveiling 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 two-prong jet classification with jet spectra [DOI]
      - Neural Network-based Top Tagger with Two-Point Energy Correlations 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]
      - Supervised Jet Clustering with Graph Neural Networks for Lorentz Boosted Bosons [DOI]
      - Track Seeding and Labelling with Embedded-space Graph Neural Networks
      - Graph neural network for 3D classification of ambiguities and optical crosstalk in scintillator-based neutrino detectors [DOI]

Parton shower and hadronization

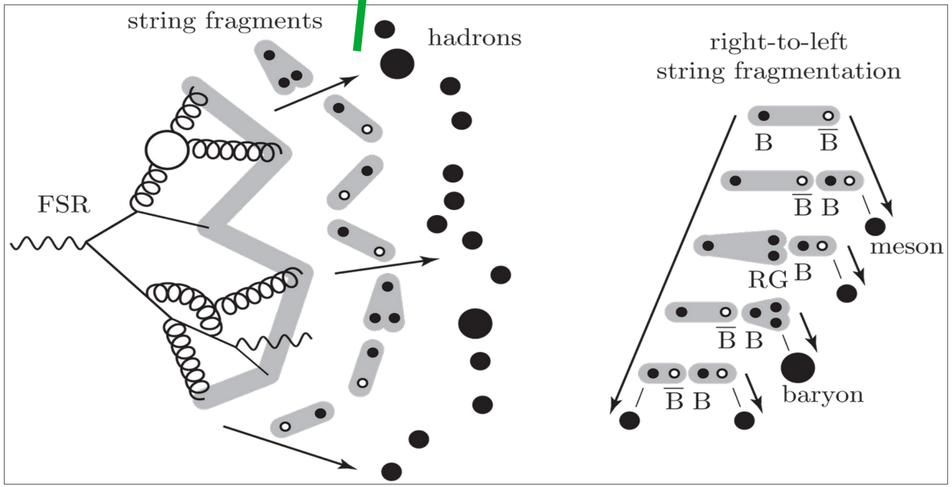
1. Hard scattering

2. Parton shower

3. Hadronization



Nature Reviews Physics 3, 73 (2021)



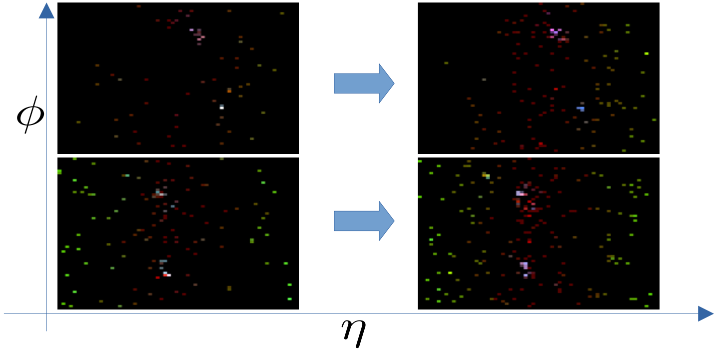
Hadronization

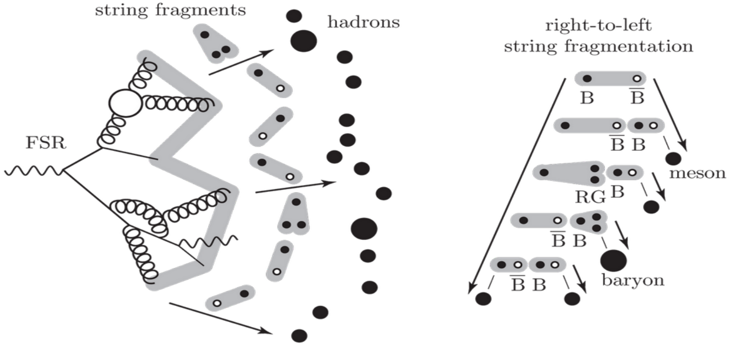
Partons → hadrons

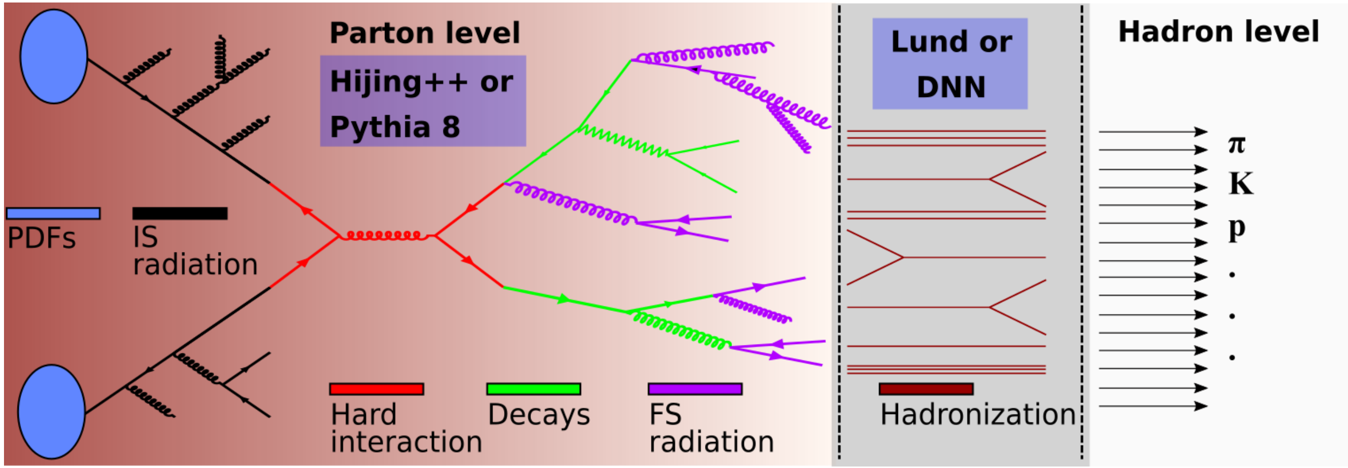
Non-perturbative process

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

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





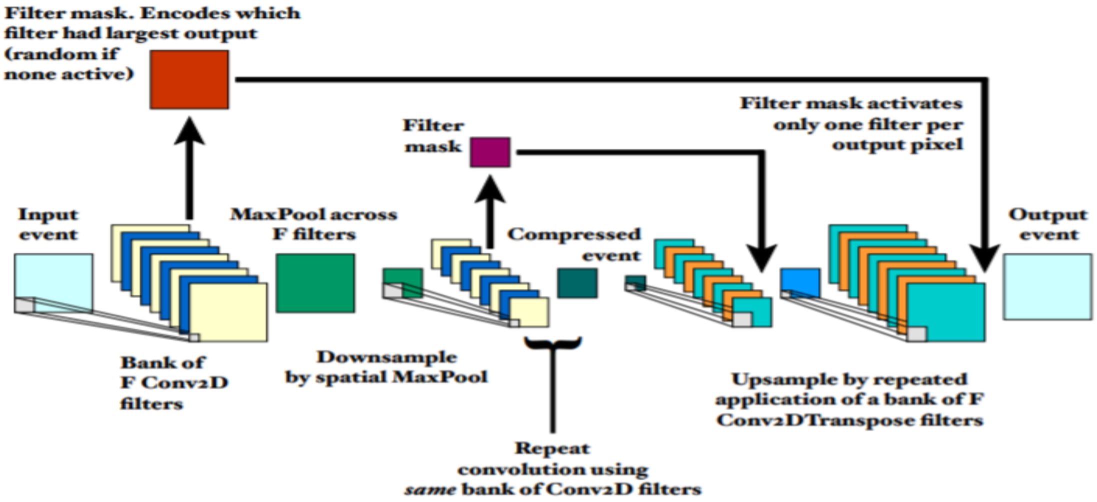


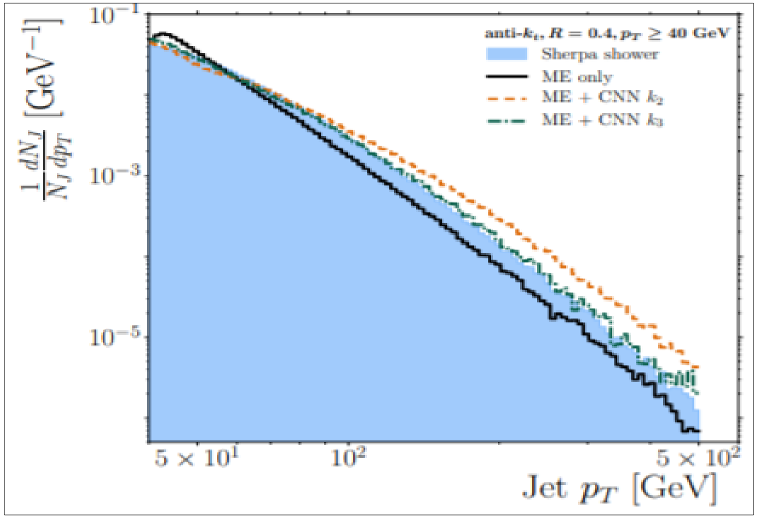
Inspiration

J.W. Monk: Deep Learning as a Parton Shower (arXiv:1807.03685)

Dataset: 500 000 QCD pp event @ 7 TeV, generated by Sherpa

parameter	model k_2	model k_3
Kernel size, k	2	3
Input image size, N	64	81
Size of filter bank, F	9	7
Levels of decomposition	5	3
Regularisation, λ	500	300
Learning rate	5×10^{-5}	1×10^{-5}
Loss weight w_1	5	4
Loss weight w_2	2	2
Loss weight w_3	1	1
Total number of trained weights	72	126





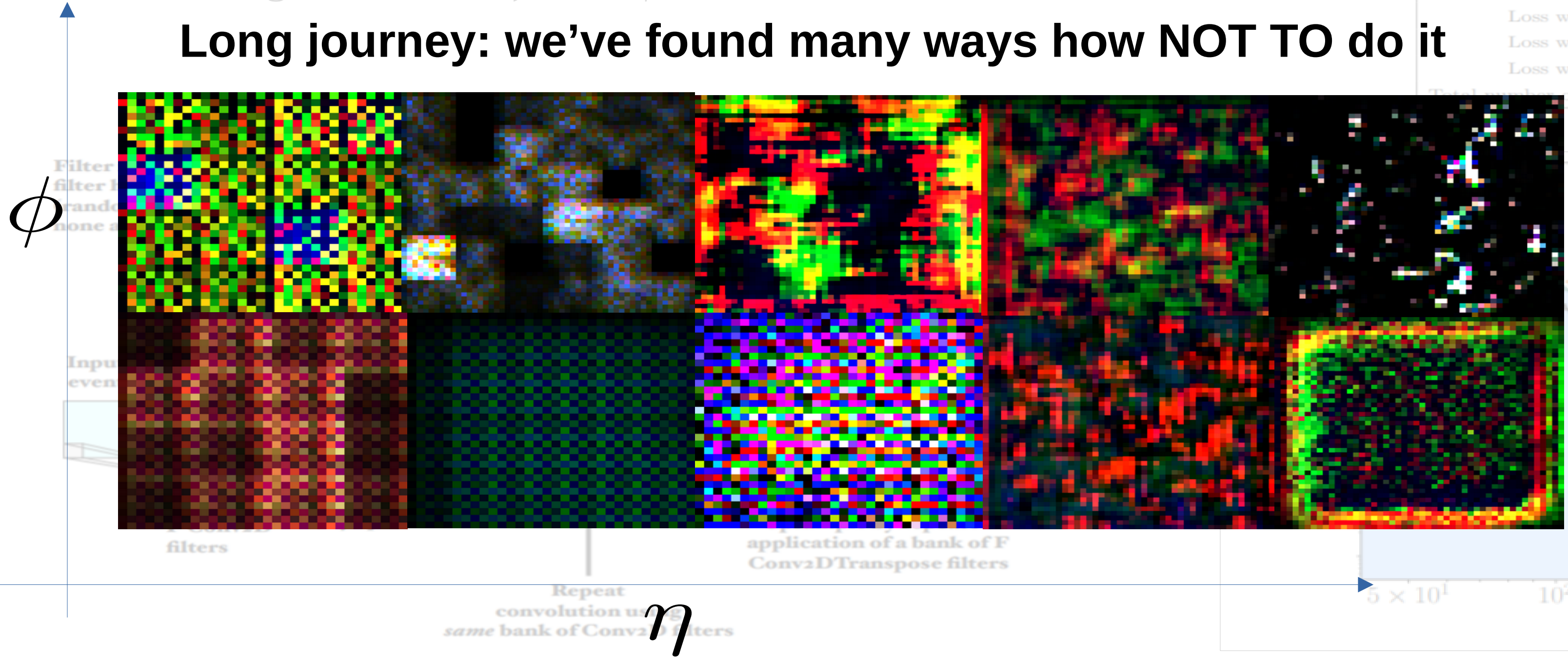
Inspiration

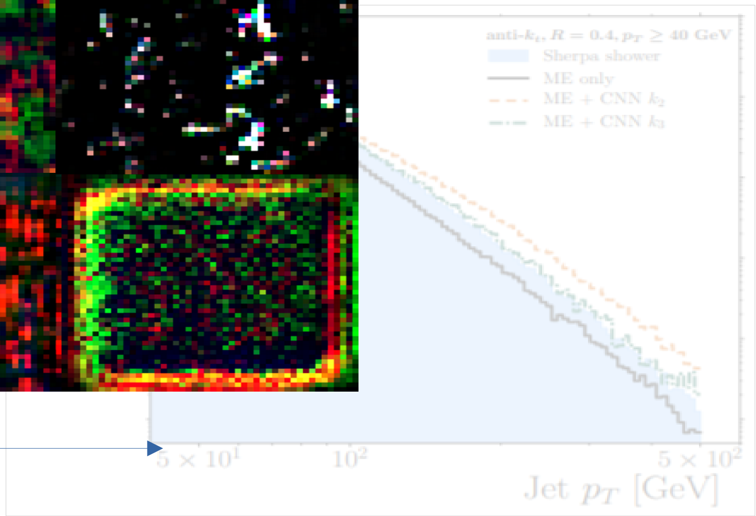
J.W. Monk: Deep Learning as a Parton Shower (arXiv:1807.03685)

Dataset: 500 000 QCD pp event @ 7 TeV, generated by Sherpa

parameter	model k_2	model k_3
Kernel size, k	2	3
Input image size, N	64	81
Size of filter bank, F	9	7
Levels of decomposition	5	3
Regularisation, λ	500	300
Learning rate	5×10^{-5}	1×10^{-5}
Loss weight w_1	5	4
Loss weight w_2	2	2
Loss weight w_3	1	1
Total [illegible] of trained weights	72	126

Long journey: we’ve found many ways how NOT TO do it



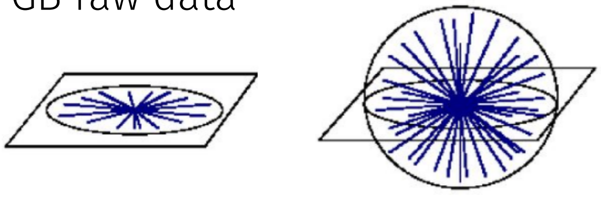


Train and validation sets

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, **√s = 7 TeV**
- Validation and test: 100 000
- ~20 GB raw data



Input:

Parton level

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

$y \in [\pi, \pi]$ 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: $\lambda_1 > \lambda_2 > \lambda_3$ $\sum_i \lambda_i = 1$

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

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

Models

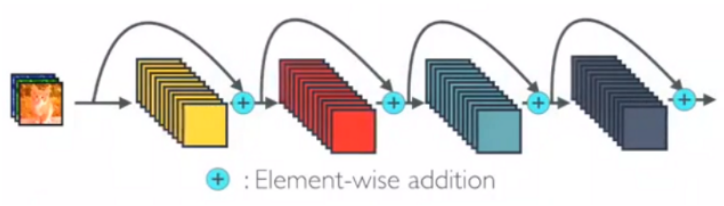
Stacking more layers: solve complex problems more efficiently, get highly accurate results

BUT:

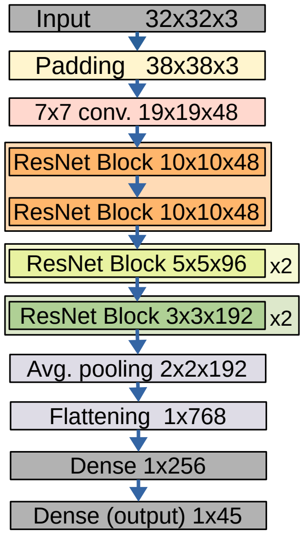
Vanishing/exploding gradients

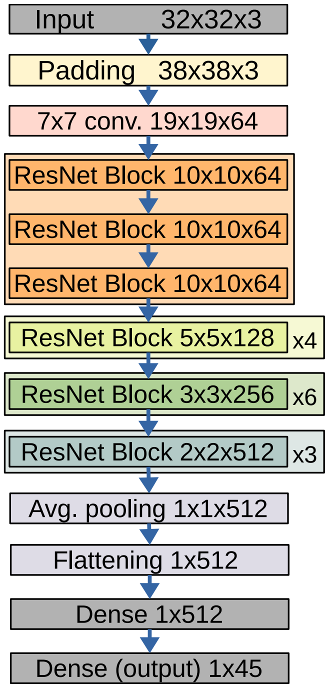
ResNet:

Residual blocks with "skip connections"



	Model S	Model L
Trainable parameters	1.7 M	20 M



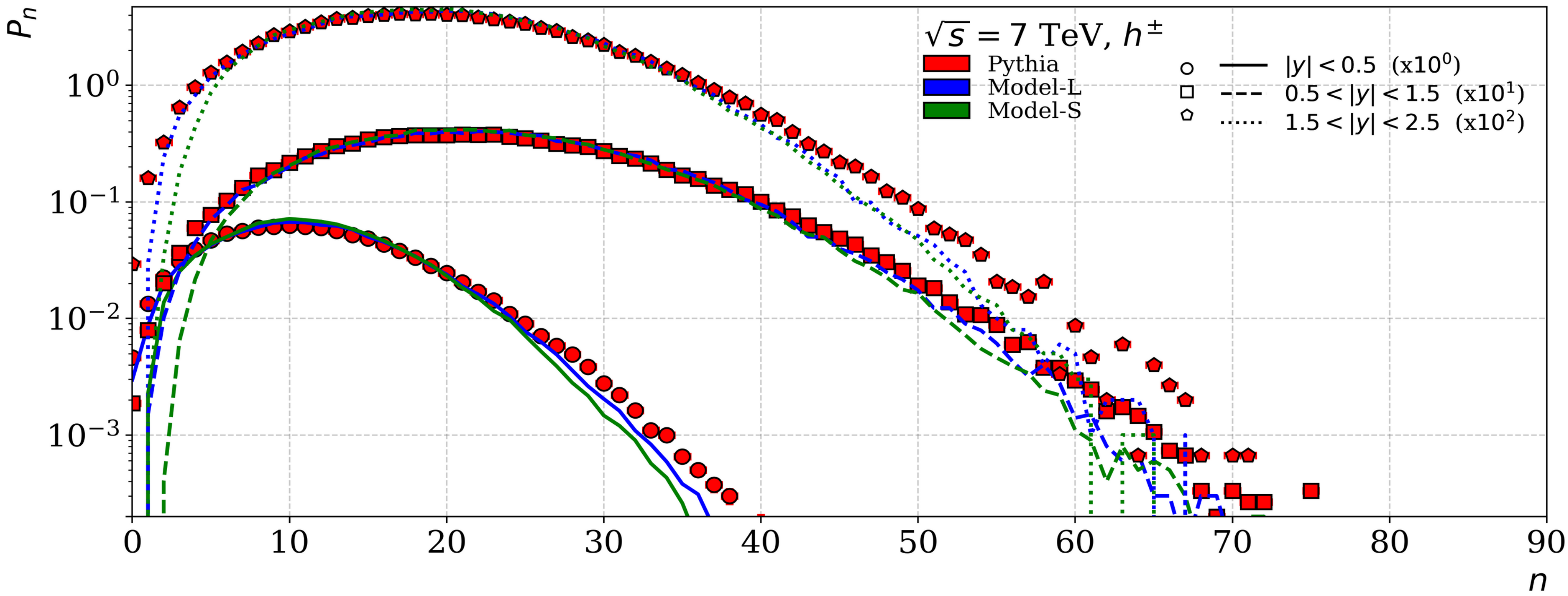


Used hardwares: Nvidia Tesla T4, GeForce GTX 1080
@ Wigner Scientific Computing Laboratory

Framework: Tensorflow 2.4.1, Keras 2.4.0

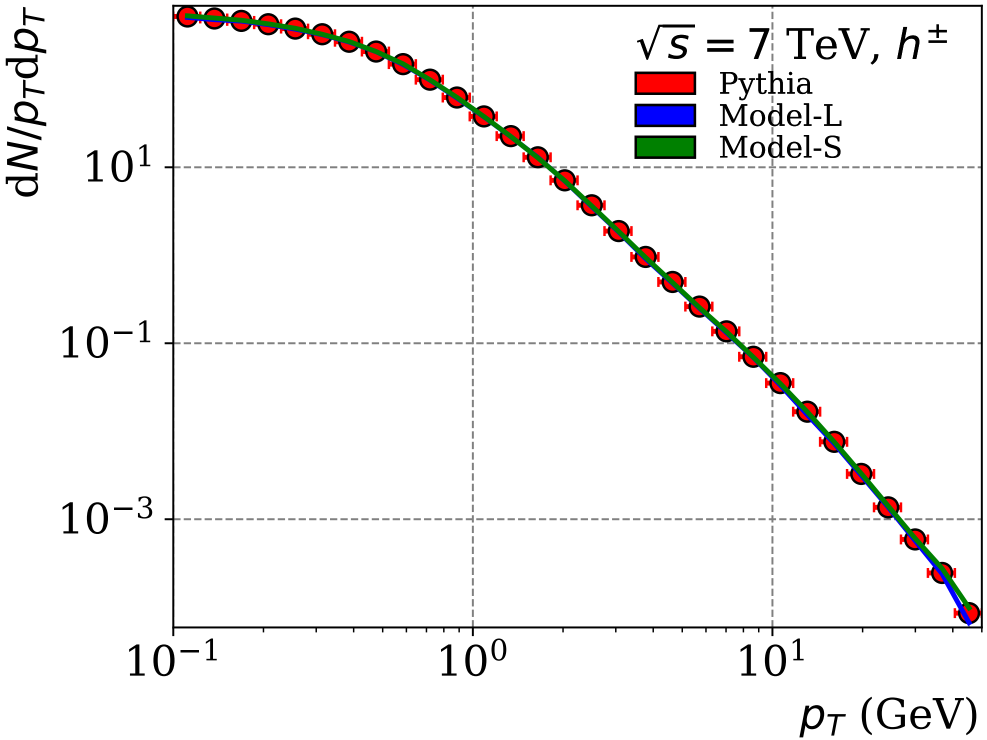
Results

Proton-proton @ 7 TeV, Training + Validation



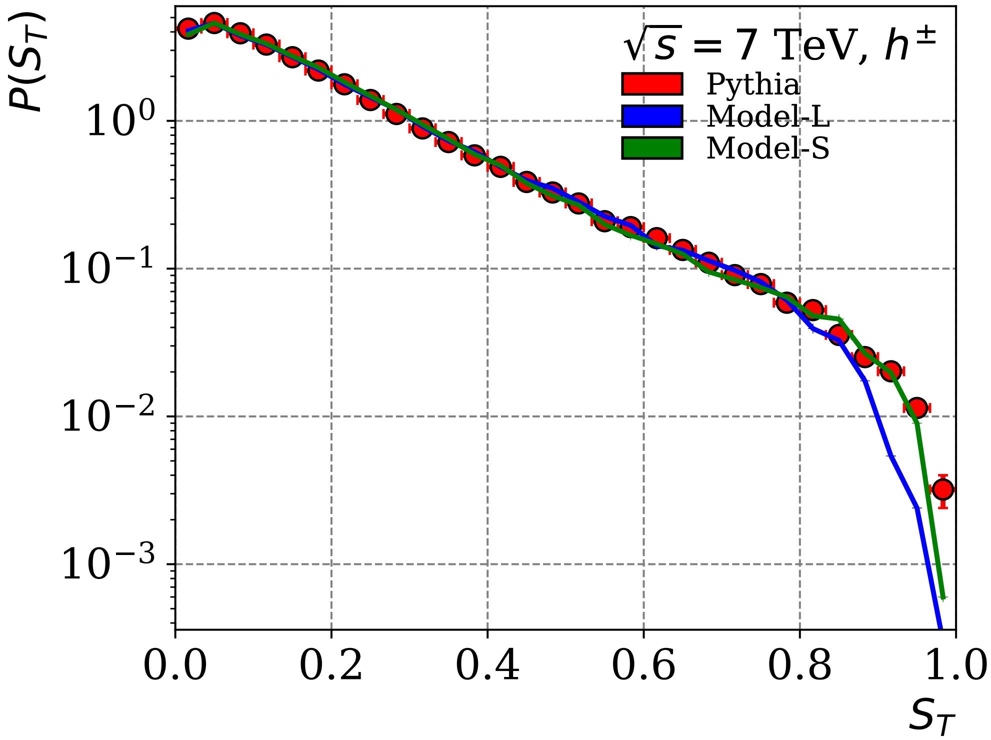
Charged hadron multiplicity at various rapidity windows
Comparison to reference MC model
Good agreement for both models

Proton-proton @ 7 TeV, Training + Validation



Charged hadron transverse momentum

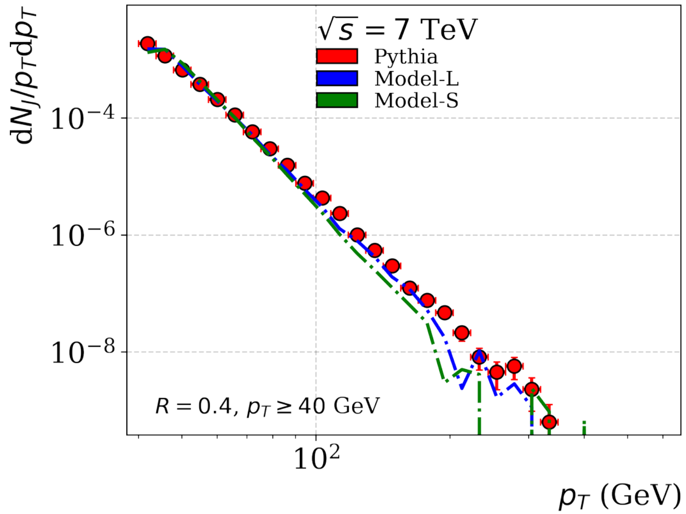
0.1 GeV ≤ p_T ≤ 50 GeV



Event transverse sphericity

The smaller model performs better

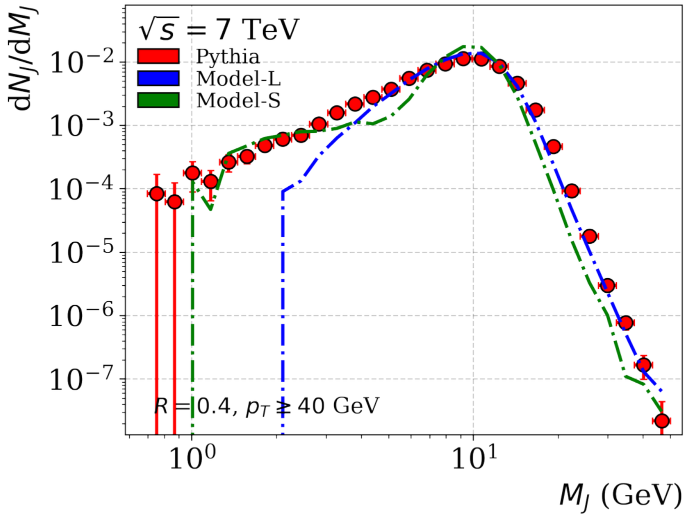
Proton-proton @ 7 TeV, Training + Validation

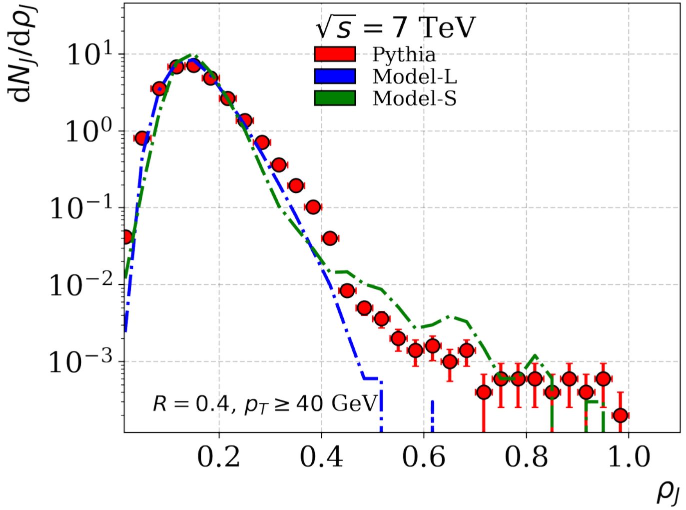


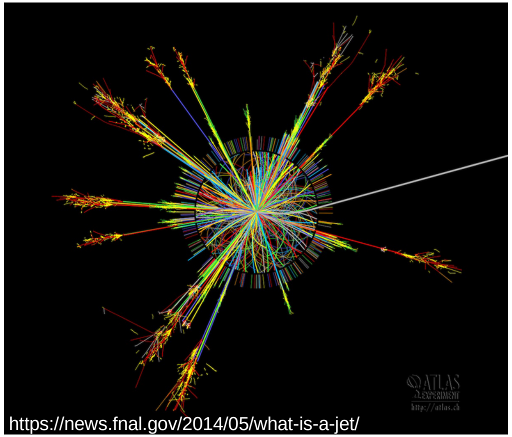
Jets:

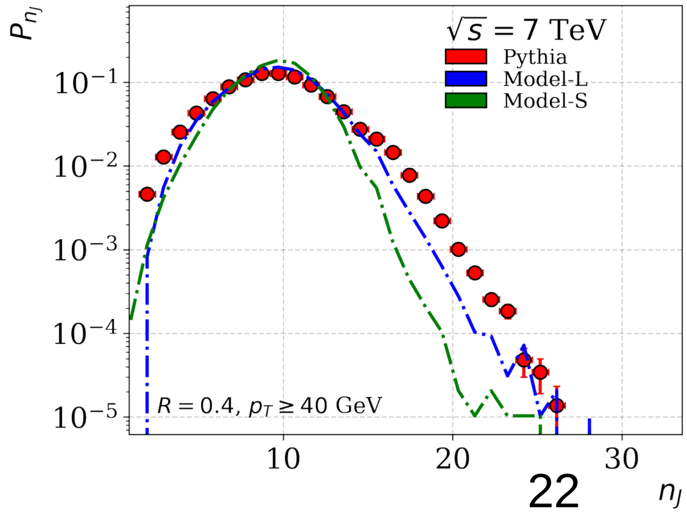
- Mean $p_T \leq 400$ GeV
- Mean mass $p_T \leq 400$ GeV
- Mean multiplicity
- Mean width

The smaller model performs better



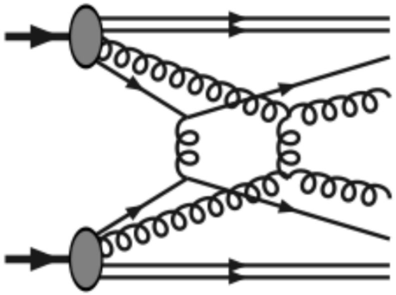




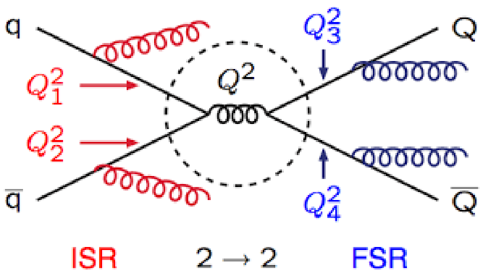


Proton-proton @ 7 TeV, Training + Validation

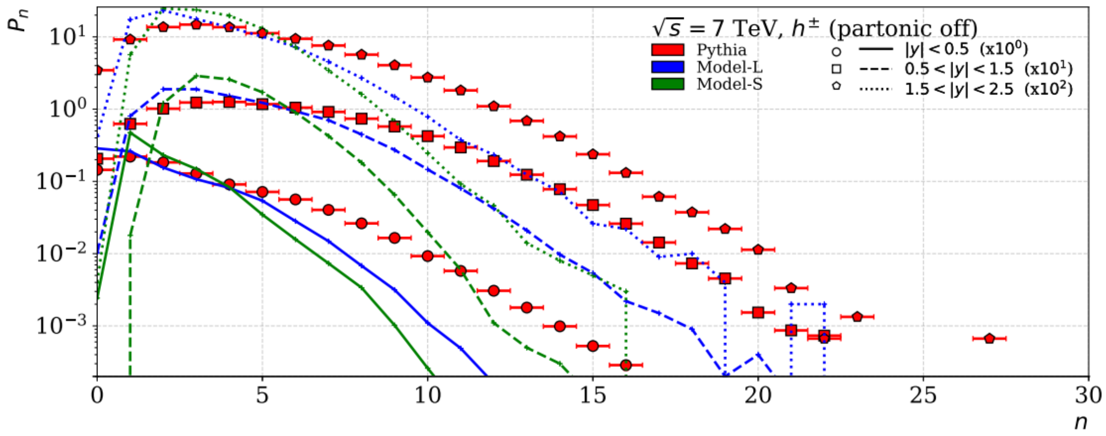
(*) What about the partonic processes?

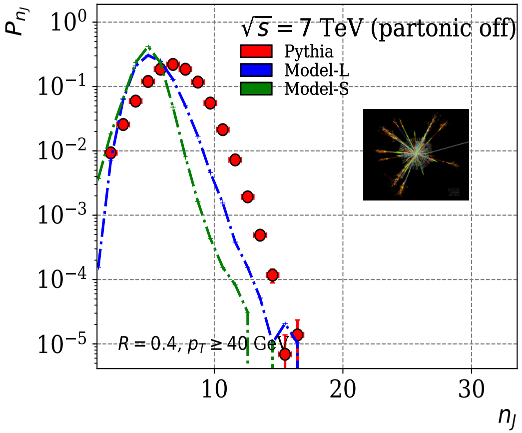


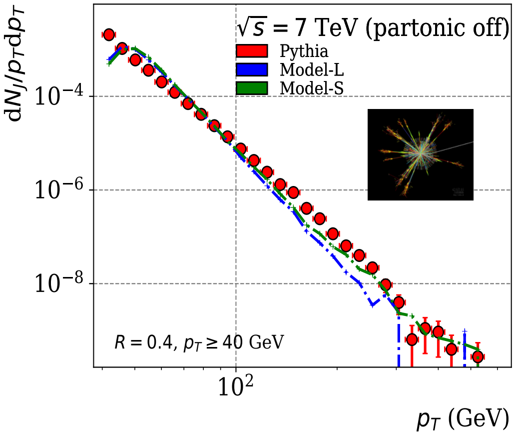
MPI

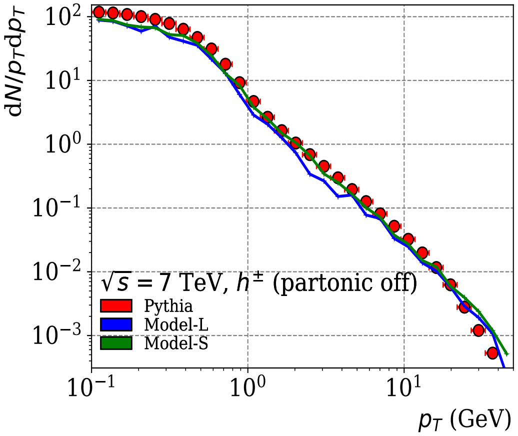


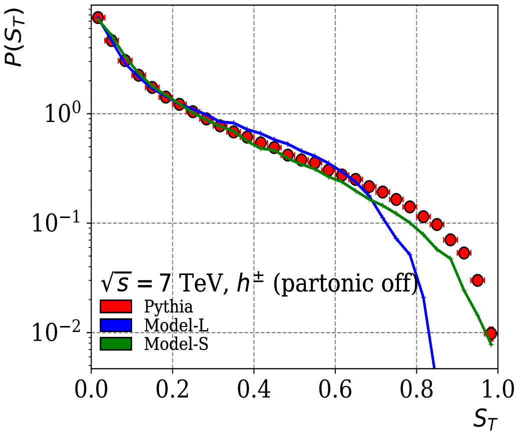
http://home.thep.lu.se/~torbjorn/talks/cern18cosmic.pdf





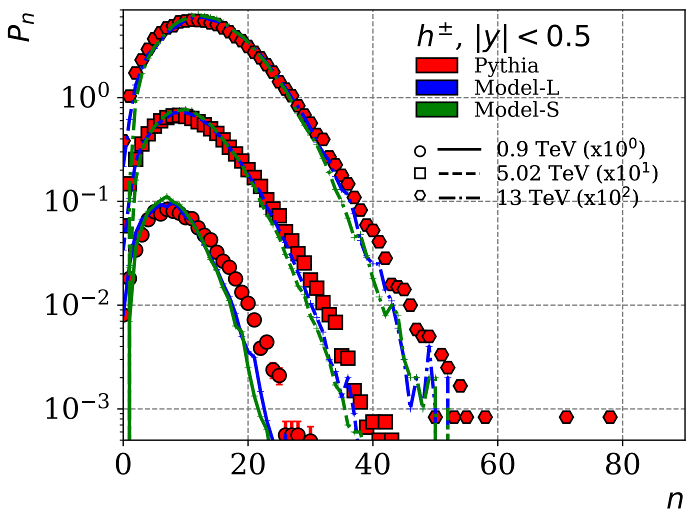


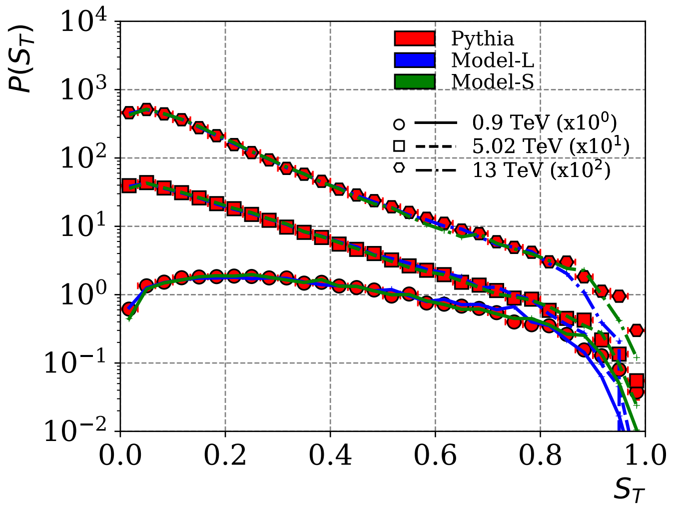


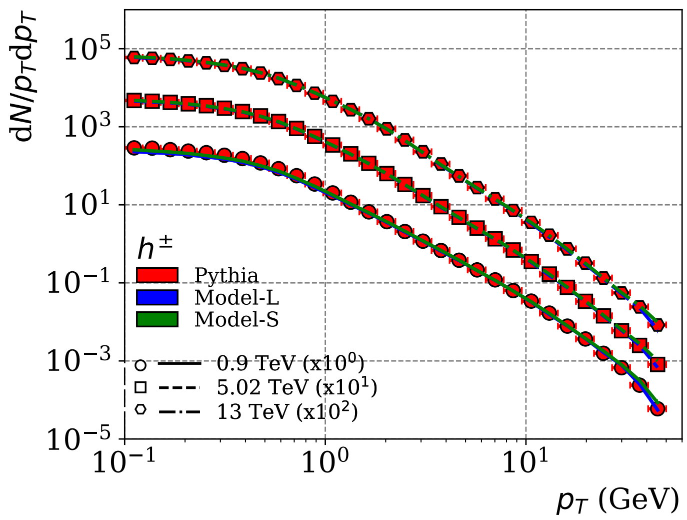


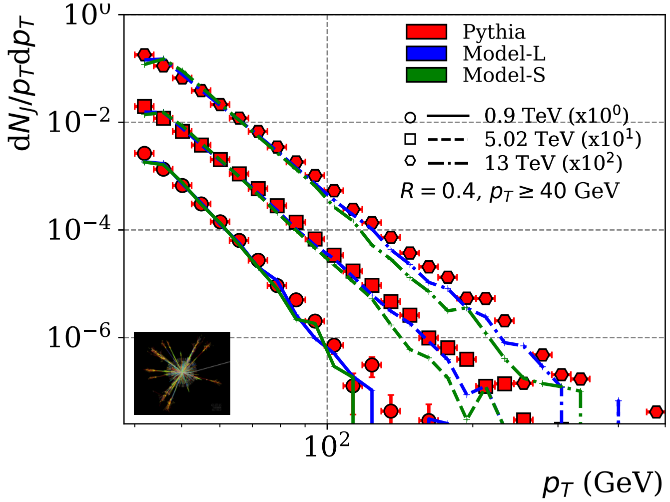
Qualitative agreement → the models adopted the hadronization properties

Proton-proton @ 0.9-13 TeV, Predictions









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

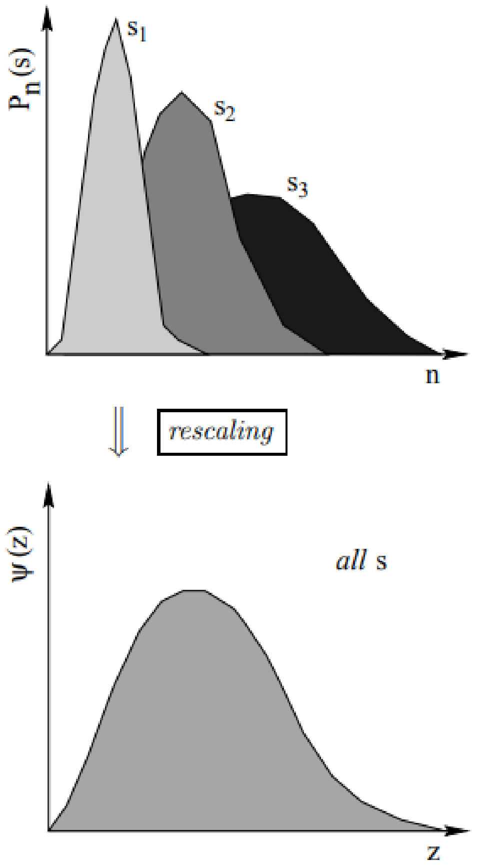
KNO-scaling

The collapse of multiplicity distributions $\mathbf{P_n}$ 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

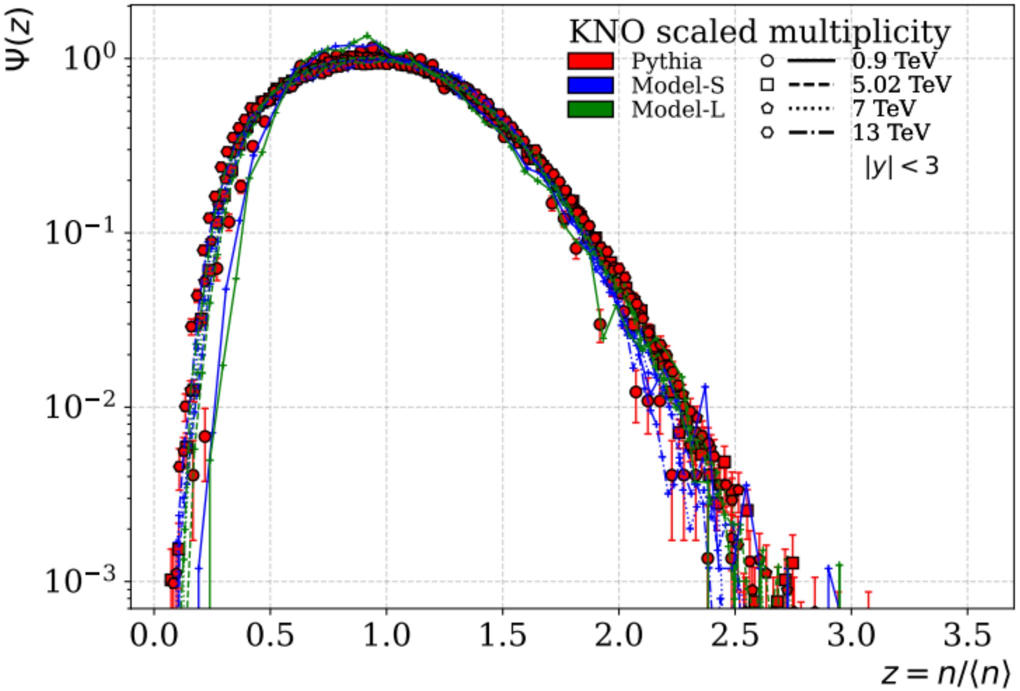
Violation of the scaling at high CM energies: not fully understood (relation to MPI?)

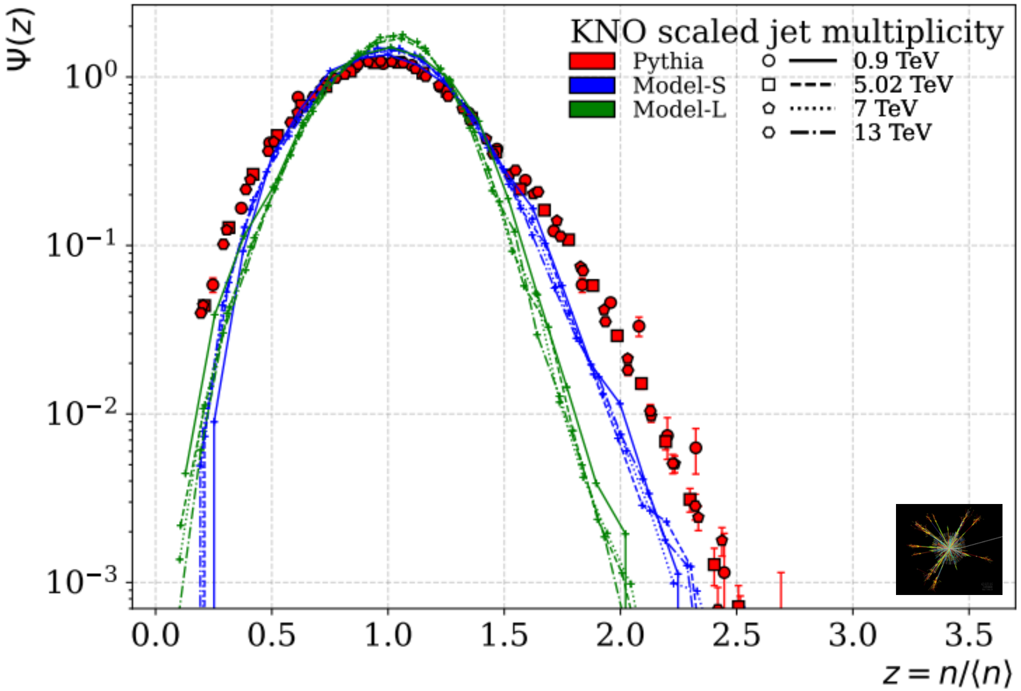


Nuclear Physics B 40 (1972), 317–334.

(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

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

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)”

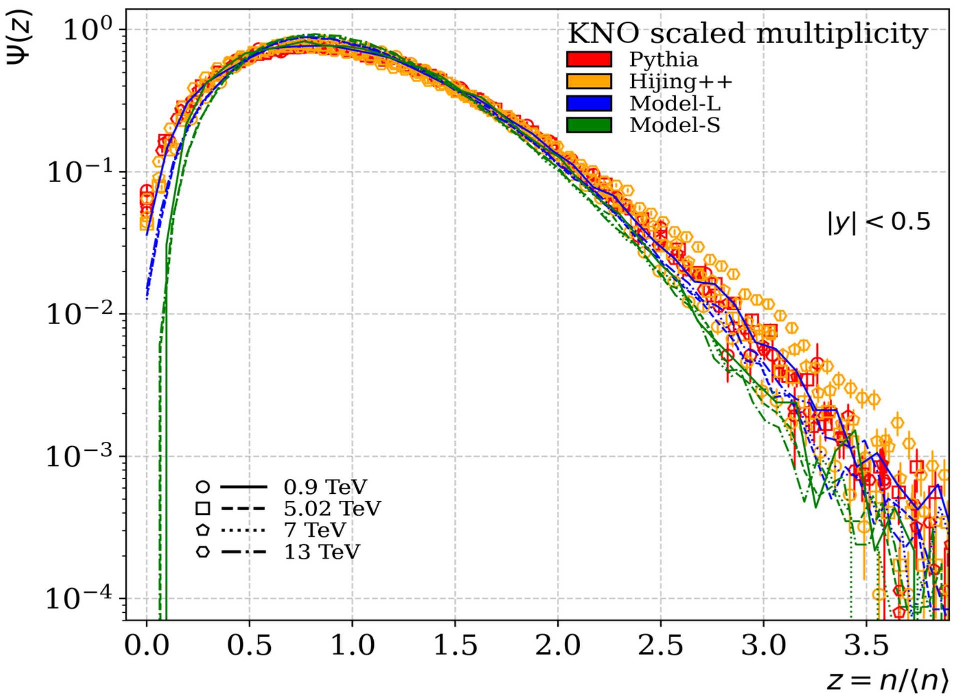
First, FORTRAN version: 1991, X.N. Wang, M. Gyulassy, **Phys. Rev. D 44, (1991) 3501.**

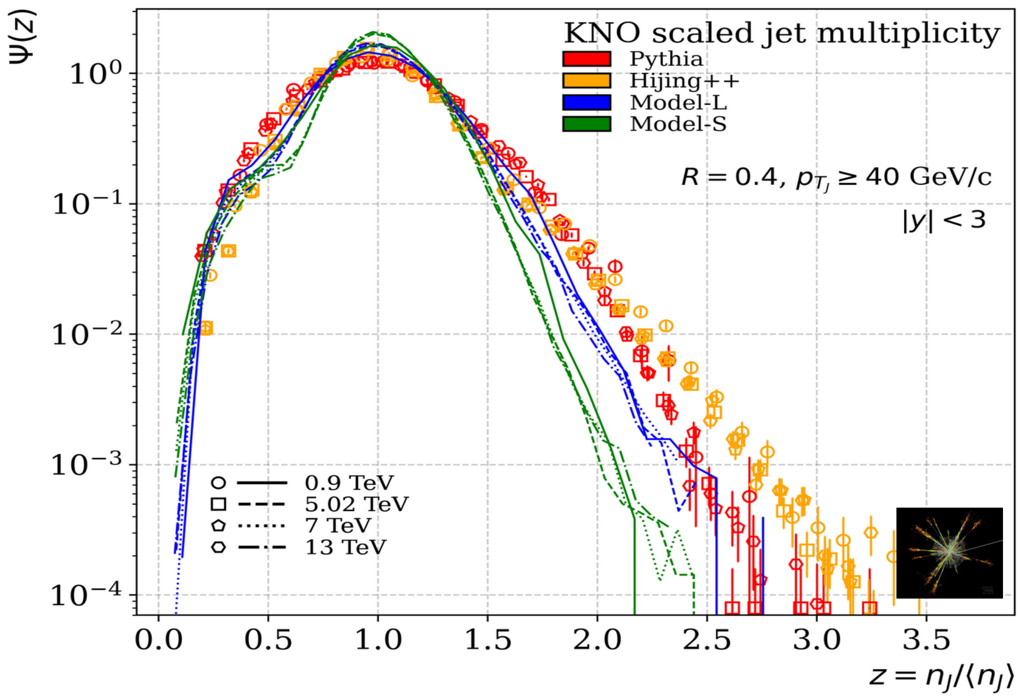
Computational challenge: more than 600 million collision in each second → HiLumiLHC: even more

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	✗	✗	✓
Analysis interface	✗	✗	✓
Module management	✗	✗	✓
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

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

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)

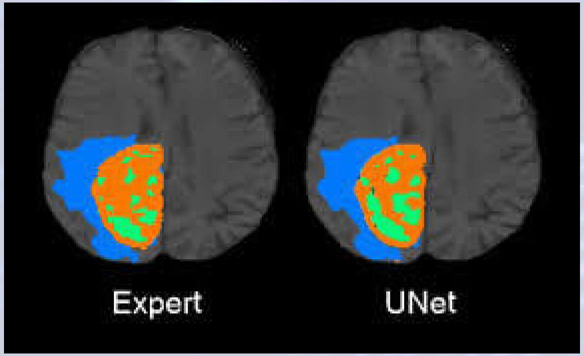
Heavy ion (centralities, collective effects)

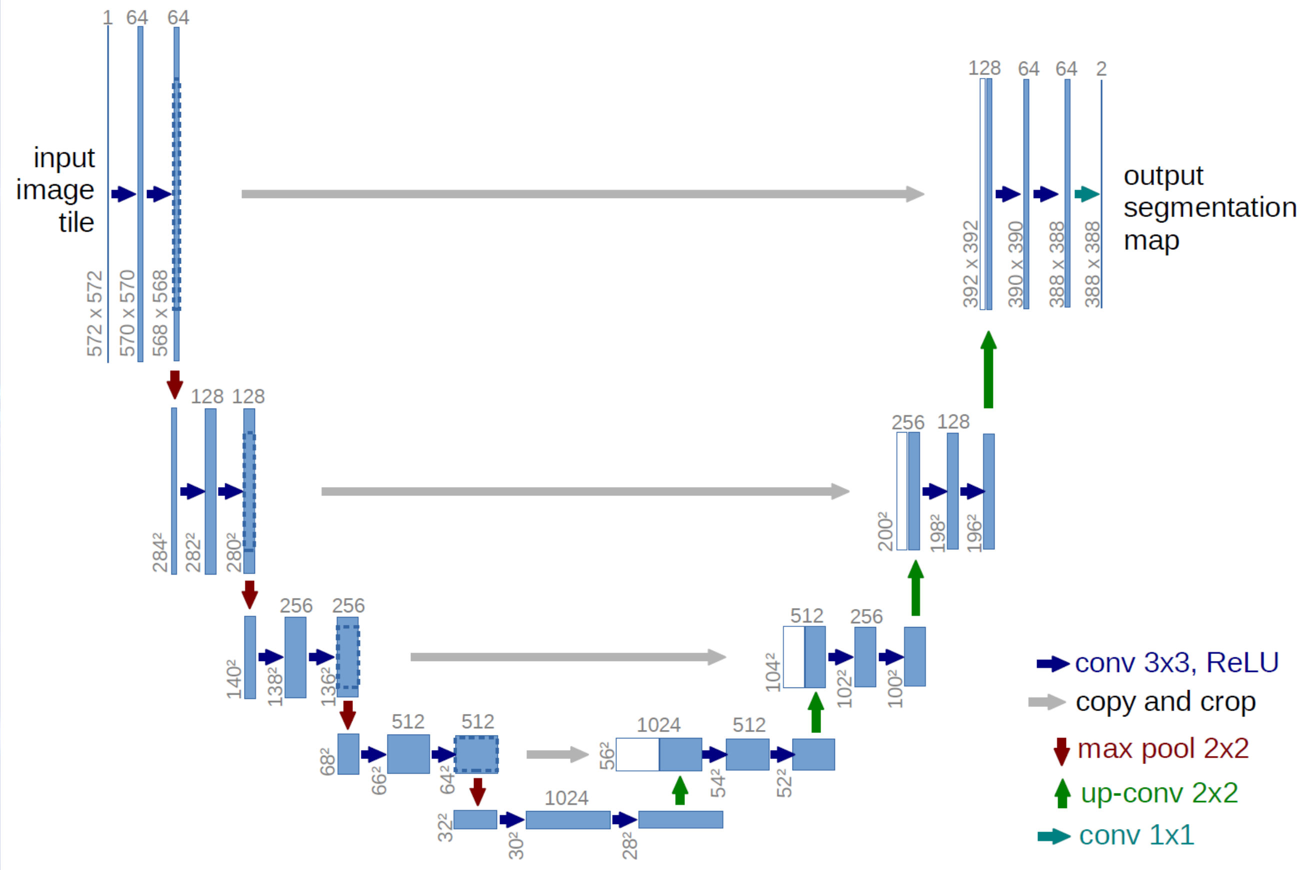
Thank you for your attention!

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 Computating Laboratory** (former Wigner GPU Laboratory) and RRF-2.3.1-21-2022-00004 within the framework of the Artificial Intelligence National Laboratory.

Popular architectures

U-Net: biomedical image segmentation





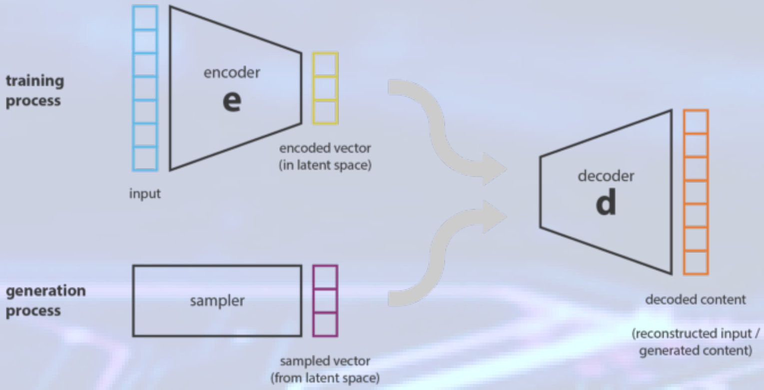
Popular architectures

(Conditional (Variational)) autoencoders

Dimension reduction

Denoising data

Latent space conditioning



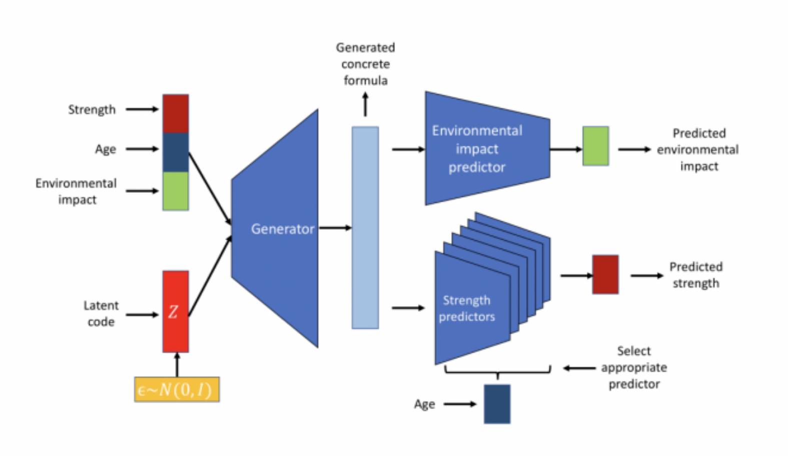
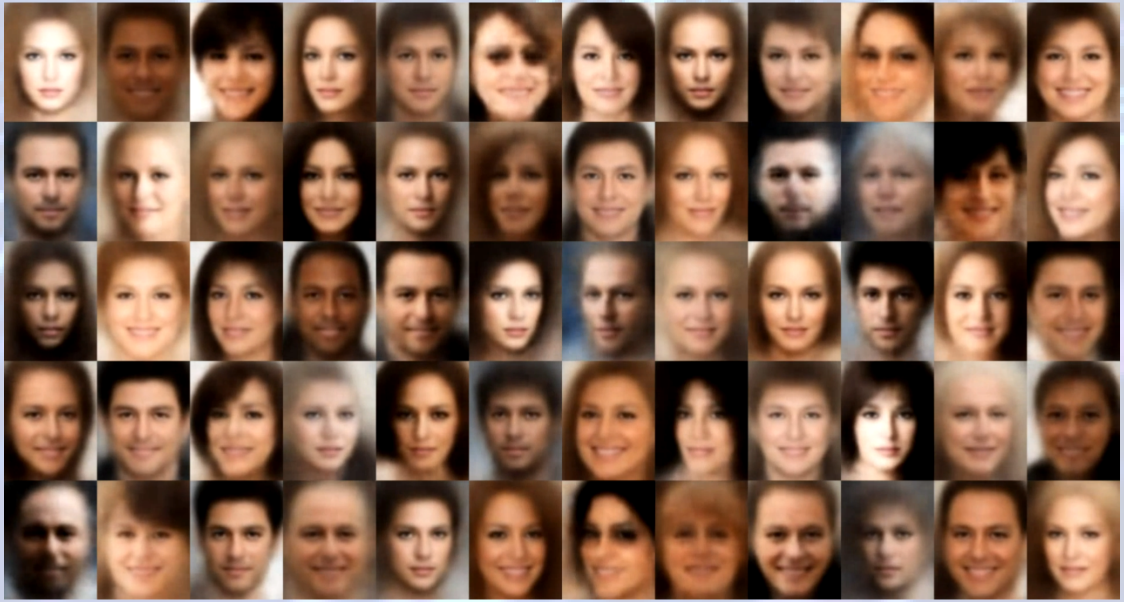


Fig. 4. Generating new concrete formulas and evaluating their properties

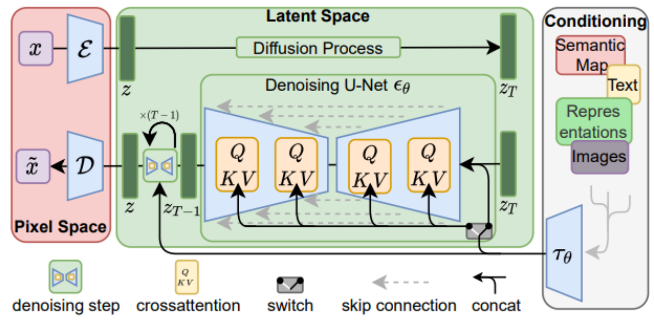
https://arxiv.org/pdf/2204.05397.pdf

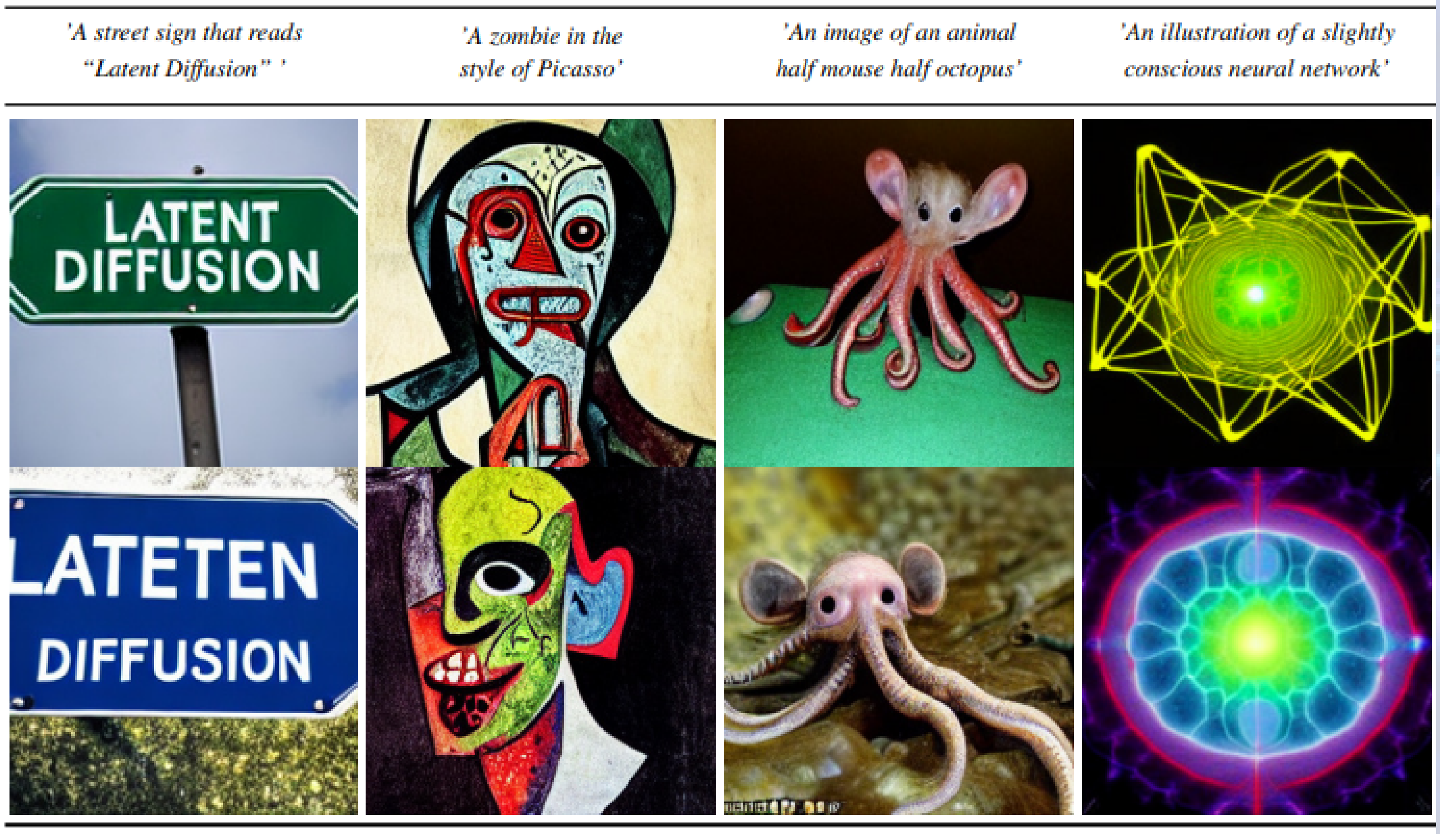


Popular architectures

Diffusion models: https://huggingface.co/spaces/stabilityai/stable-diffusion

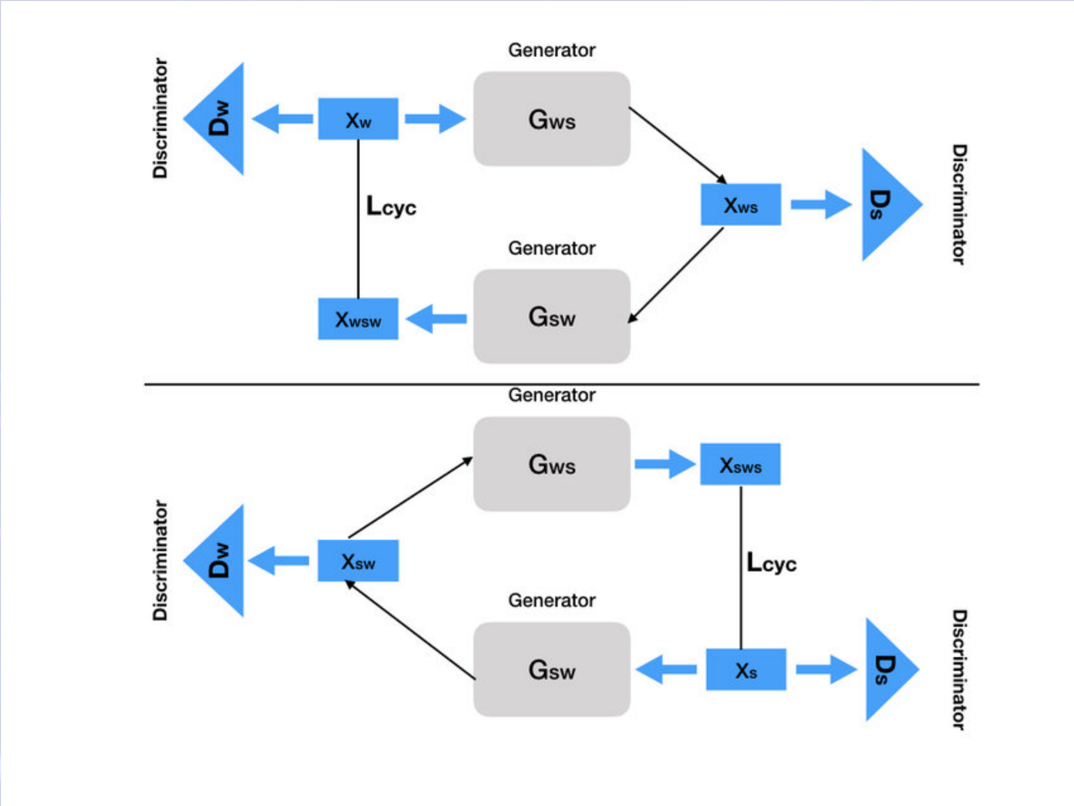
Gradually perturbate he input data over several steps by adding Gaussian noise





Popular architectures

GAN: data generation via competing generator-discriminator



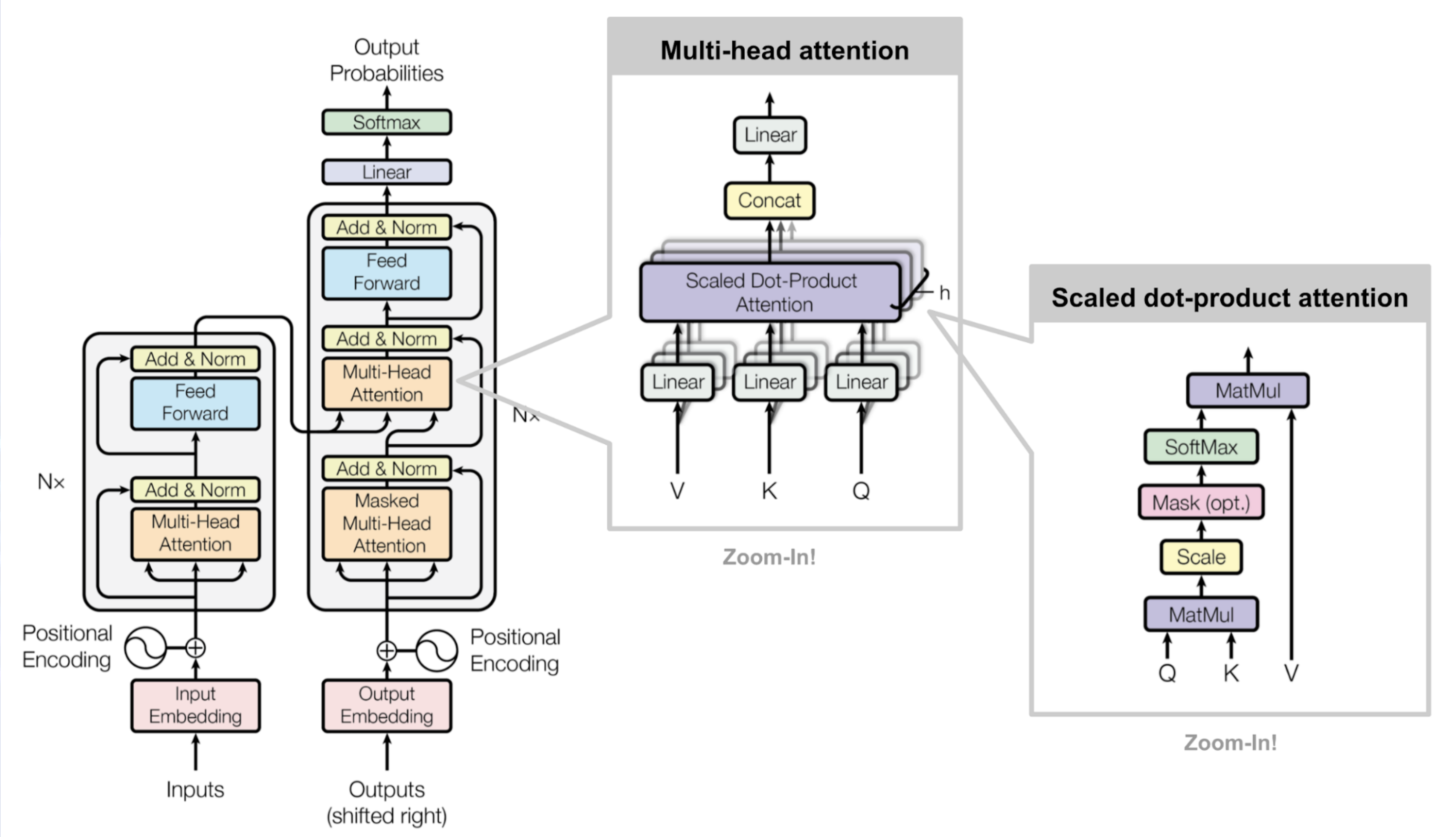
Popular architectures

GAN: data generation via competing generator-discriminator

Popular architectures

Attention and Transformers :

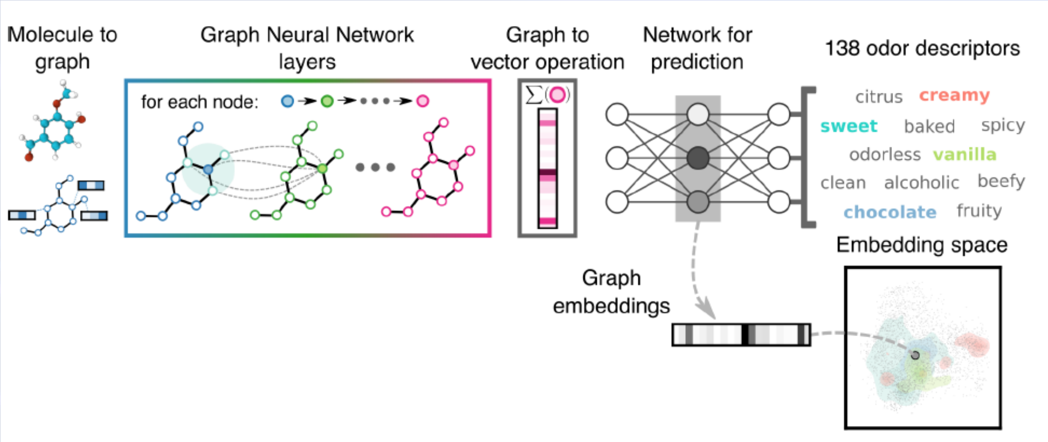
A revolution in natural language processing

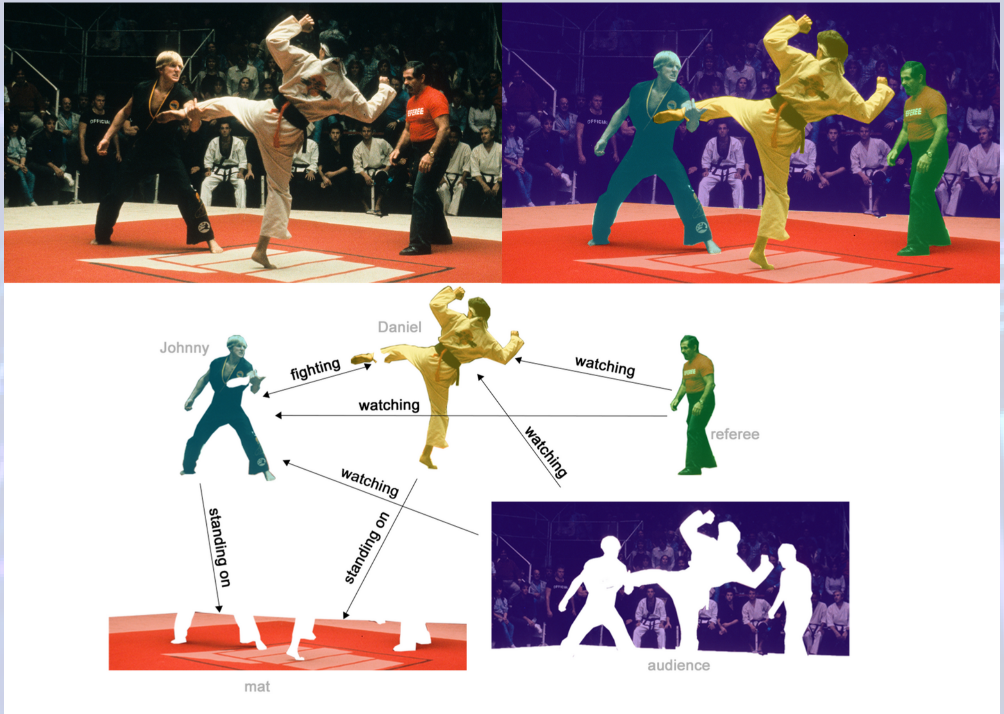


https://arxiv.org/abs/1706.03762

Popular architectures

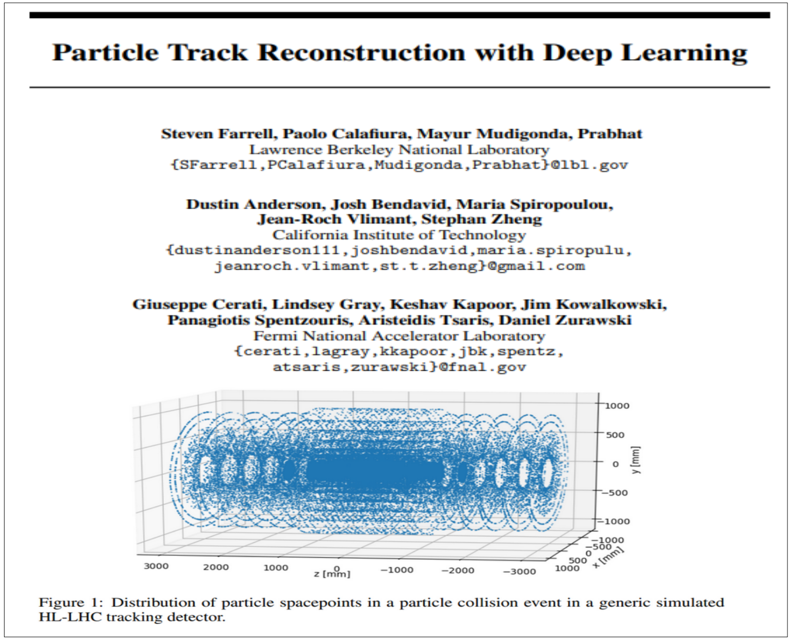
Graph Neural Networks





Machine Learning in HEP

Track reconstruction



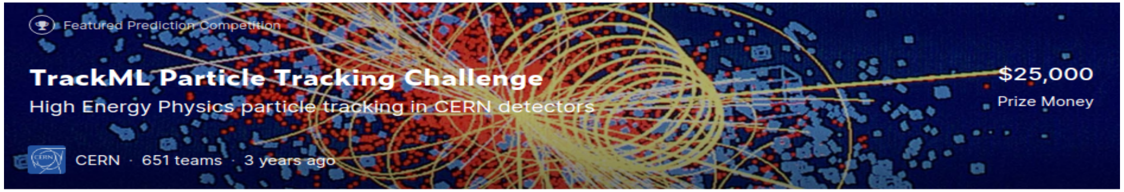
Particle Track Reconstruction with Deep Learning

Steven Farrell, Paolo Calafiura, Mayur Mudigonda, Prabhat
Lawrence Berkeley National Laboratory
{SFarrell,PCalafiura,Mudigonda,Prabhat}@lbl.gov

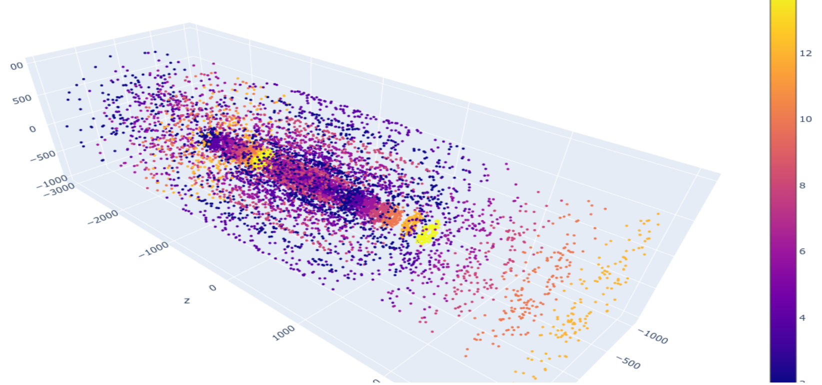
Dustin Anderson, Josh Bendavid, Maria Spiropoulou, Jean-Roch Vlimant, Stephan Zheng
California Institute of Technology
{dustinanderson111,joshbendavid,maria.spiropulu, jeanroch.vlimant,st.t.zheng}@gmail.com

Giuseppe Cerati, Lindsey Gray, Keshav Kapoor, Jim Kowalkowski, Panagiotis Spentzouris, Aristeidis Tsaris, Daniel Zurawski
Fermi National Accelerator Laboratory
{cerati,lagray,kkapoor,jbk,spentz, atsaris,zurawski}@fnal.gov

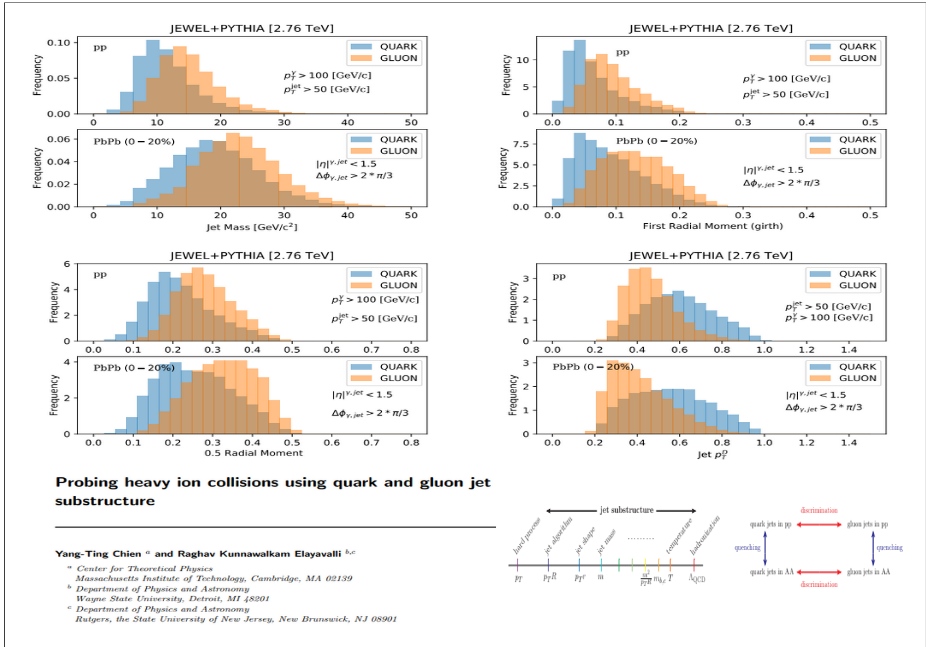
Figure 1: Distribution of particle spacepoints in a particle collision event in a generic simulated HL-LHC tracking detector.



https://www.kaggle.com/c/trackml-particle-identification

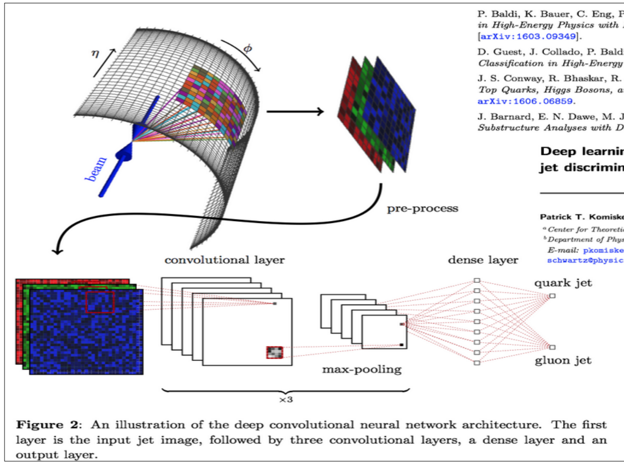


Machine Learning in HEP



arXiv:1803.03589

Quark/gluon jet separation



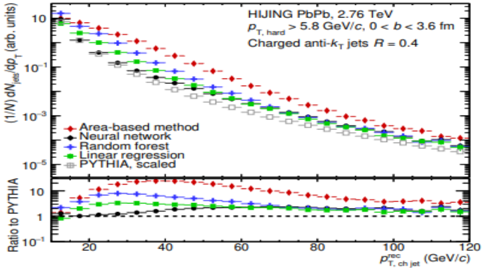
https://doi.org/10.1007/JHEP01(2017)110

Deep CNN match or outperform traditional jet observables.

Machine Learning in HEP

Machine Learning based jet momentum reconstruction in heavy-ion collisions

Rüdiger Haake[1] and Constantin Loizides[2]
[1]*Yale University, Wright Laboratory, New Haven, CT, USA*
[2]*ORNL, Physics Division, Oak Ridge, TN, USA*
(Dated: June 24, 2019)



Feature	Score	Feature	Score
Jet p_T (no corr.)	**0.1355**	$p^1_{T,\,const}$	0.0012
Jet mass	0.0007	$p^2_{T,\,const}$	**0.0039**
Jet area	0.0005	$p^3_{T,\,const}$	0.0015
Jet p_T (area-based corr.)	**0.7876**	$p^4_{T,\,const}$	0.0011
LeSub	0.0004	$p^5_{T,\,const}$	0.0009
Radial moment	0.0005	$p^6_{T,\,const}$	0.0009
Momentum dispersion	0.0007	$p^7_{T,\,const}$	0.0008
Number of constituents	0.0008	$p^8_{T,\,const}$	0.0007
Mean of const. p_T	**0.0585**	$p^9_{T,\,const}$	0.0006
Median of const. p_T	0.0023	$p^{10}_{T,\,const}$	0.0007

FIG. 9. Reconstructed charged jet spectra in HIJING events and the ratio to (N_{coll}-scaled) PYTHIA jet spectra.

https://doi.org/10.1103/PhysRevC.99.064904

Jet reconstruction

Machine Learning based jet momentum reconstruction in Pb–Pb collisions measured with the ALICE detector

Rüdiger Haake* **for the ALICE Collaboration**
Yale University, Wright Laboratory, New Haven, CT, USA
E-mail: ruediger.haake@cern.ch

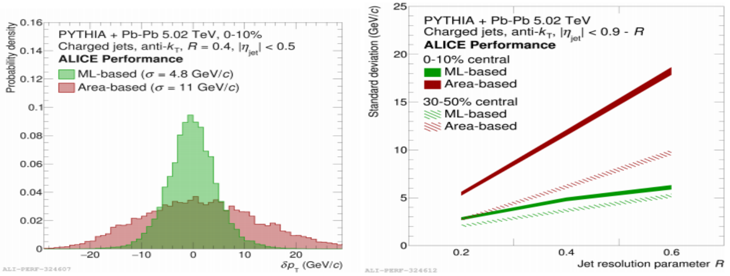
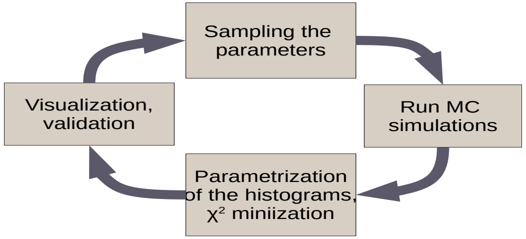


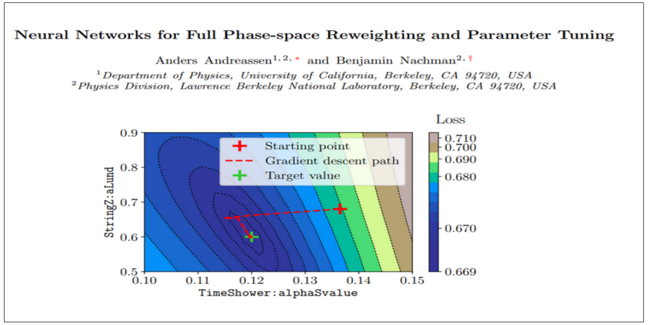
Figure 1: Residual p_T-distributions of embedded jet probes of known transverse momentum.

https://doi.org/10.22323/1.364.0312

Machine Learning in HEP

Tuning Monte Carlo event generators




Neural Networks for Full Phase-space Reweighting and Parameter Tuning

Anders Andreassen[1,2,*] and Benjamin Nachman[2,†]
[1]Department of Physics, University of California, Berkeley, CA 94720, USA
[2]Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

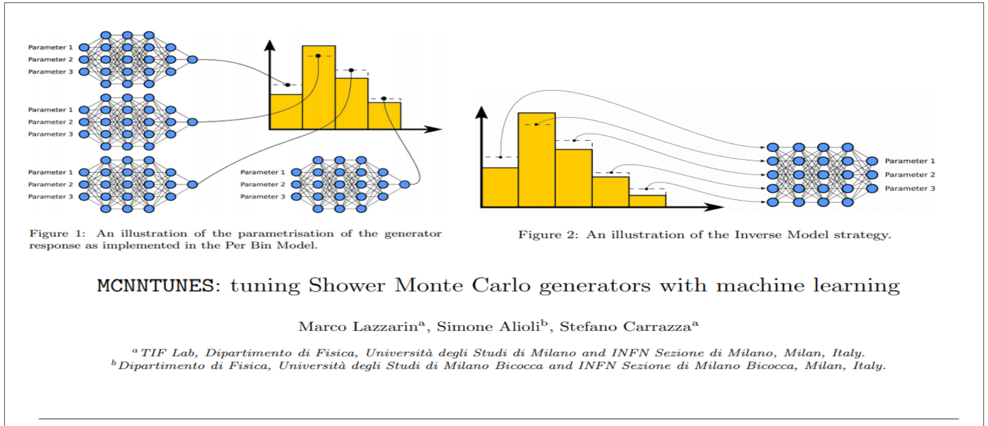


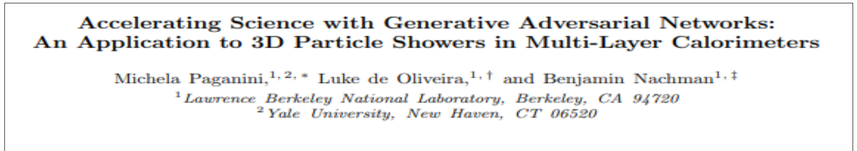
Figure 1: An illustration of the parametrisation of the generator response as implemented in the Per Bin Model.

Figure 2: An illustration of the Inverse Model strategy.

MCNNTUNES: tuning Shower Monte Carlo generators with machine learning

Marco Lazzarin[a], Simone Alioli[b], Stefano Carrazza[a]
[a]TIF Lab, Dipartimento di Fisica, Università degli Studi di Milano and INFN Sezione di Milano, Milan, Italy.
[b]Dipartimento di Fisica, Università degli Studi di Milano Bicocca and INFN Sezione di Milano Bicocca, Milan, Italy.

https://doi.org/10.1016/j.cpc.2021.107908


Accelerating Science with Generative Adversarial Networks: An Application to 3D Particle Showers in Multi-Layer Calorimeters

Michela Paganini,[1,2,*] Luke de Oliveira,[1,†] and Benjamin Nachman[1,‡]
[1]Lawrence Berkeley National Laboratory, Berkeley, CA 94720
[2]Yale University, New Haven, CT 06520

https://doi.org/10.1103/PhysRevLett.120.042003

Dimensionality

Input:
Parton level

Discretized in the (y, ϕ) plane: p_T,m, multiplicity

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

$\mathcal{O}(10^3 - 10^4)$ Total pixels

vs $\mathcal{O}(10^2)$

Pixels with information

Reduction with Singular Value Decomposition:

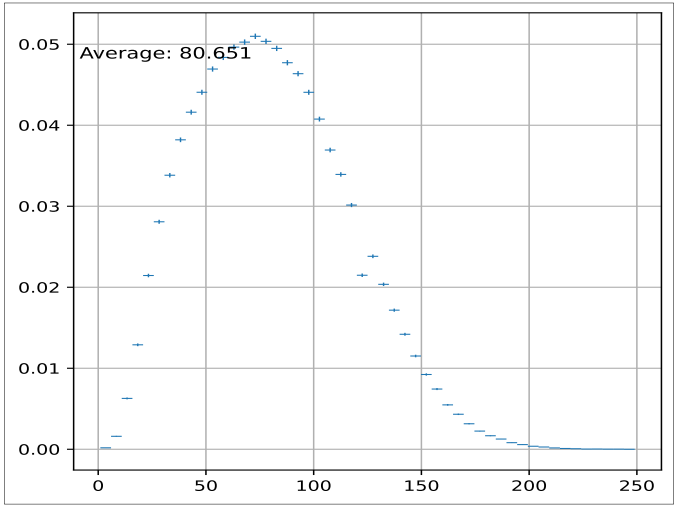
$$M_{n\times m} = U_{n\times n}\Sigma_{n\times m}V^T_{m\times m}$$

- 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 $\mathcal{O}(10^2)$

doi:10.1007/BF02288367



Data-Driven Science and Engineering (S. L. Brunton, J. N. Kutz)

Dimensionality (work in progress)