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

arXiv:2111.15655
arXiv:2210.10548
arXiv:2303.05422

UNIVERSITY OF OXFORD

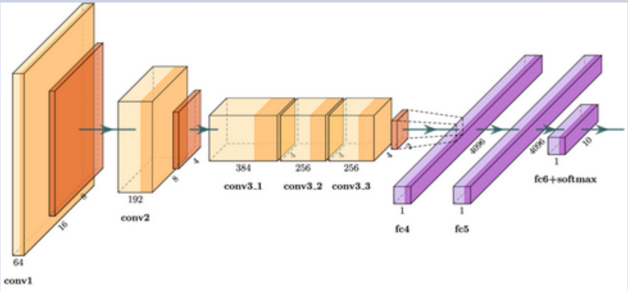
25 YEARS OF EPJ.org
your physics journal

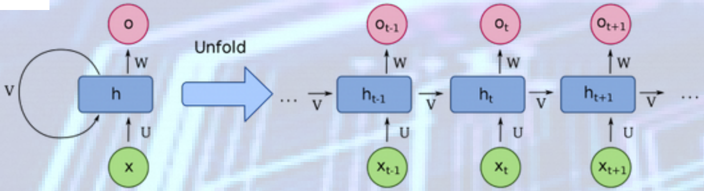
History

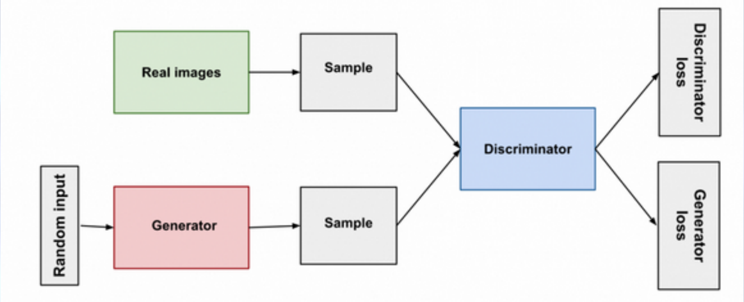
CNN (image classification, object detection, recommender systems)...

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

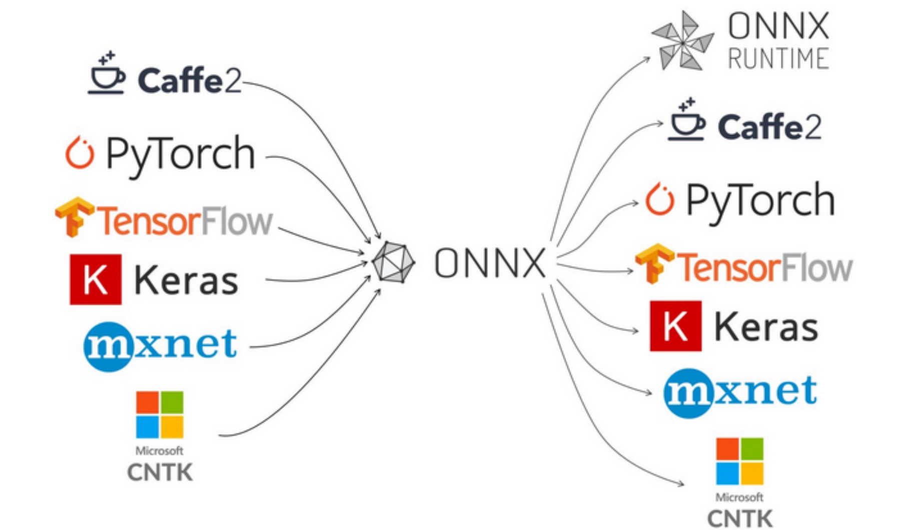
Generative models: anomaly detection, pattern recognition, reinforced learning







Various frameworks for training and inference:



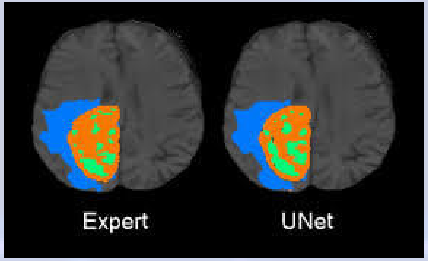
https://towardsdatascience.com/onnx-preventing-framework-lock-in-9a798fb34c92

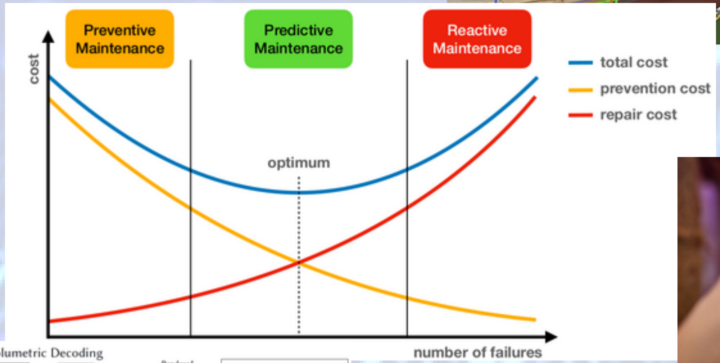
Motivaton – data, data, more data

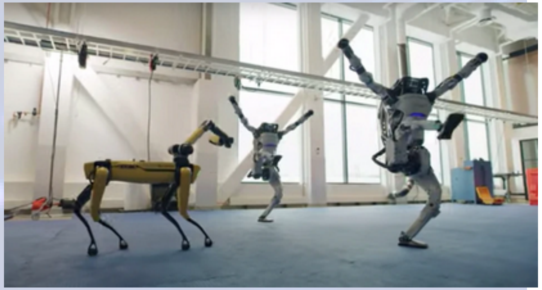
Autonomous driving
Medical imaging
Predictive maintenance
Anomaly detection, fake news detection
Search of BSM physics
Stock price prediction
Natural Language Processing
Virtual Assistants
Virtual reality
Colorization of Black and White Images
Content generation, examples:

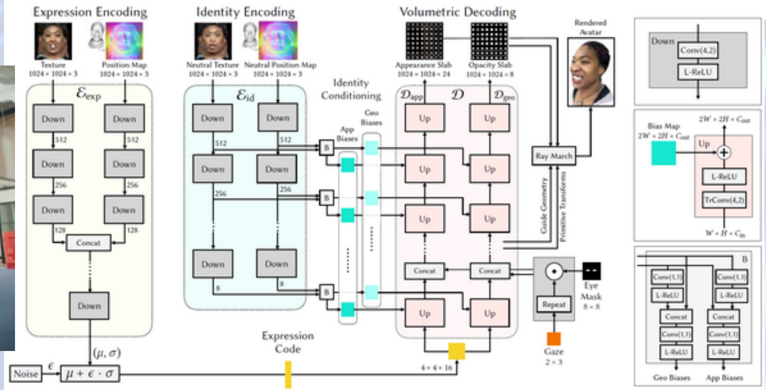
- https://infiniteconversation.com/
- https://huggingface.co/spaces/stabilityai/stable-diffusion

Robotics
…



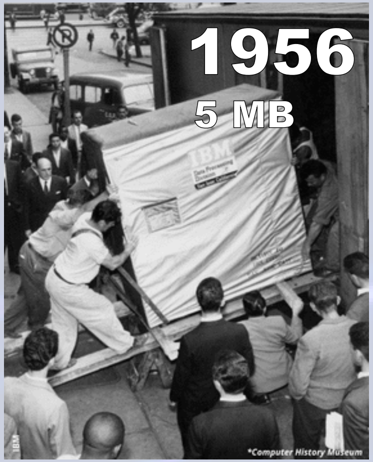


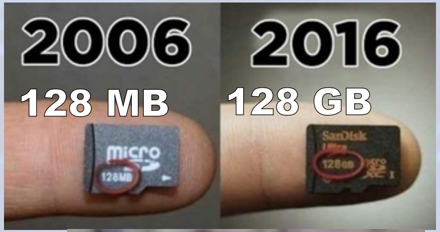


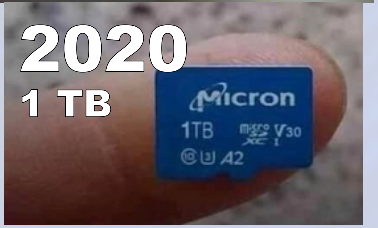


Motivaton – data, data, more data

WLCG Worldwide LHC Computing Grid





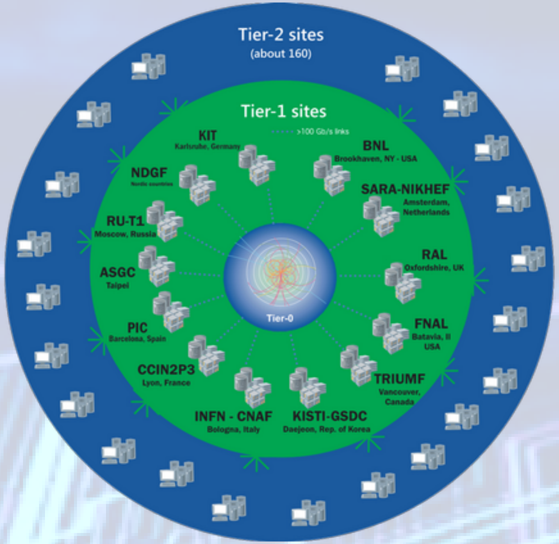


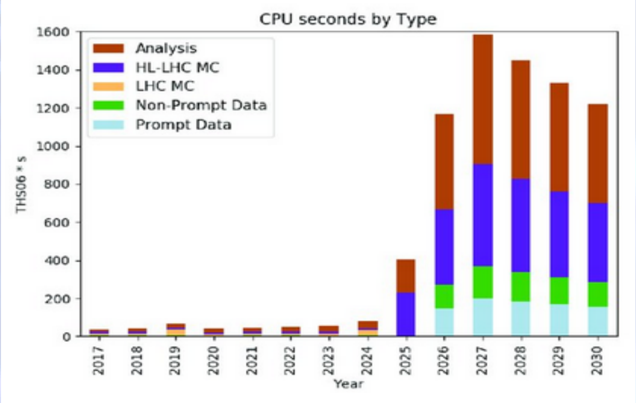
LHC in numbers: **2013** and **now:**

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 challange

→ **analysis, simulations**

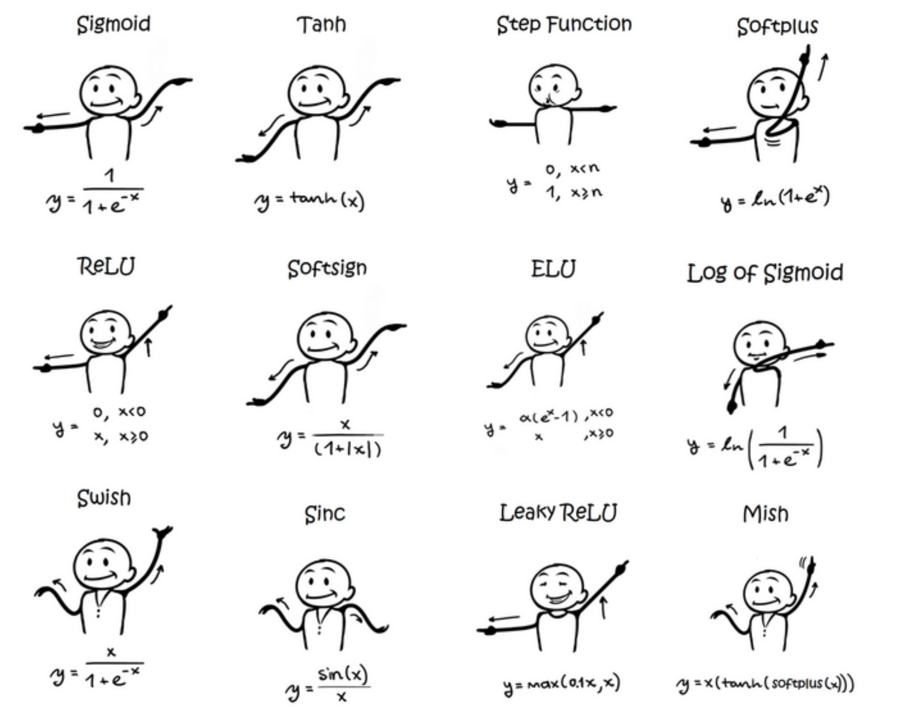


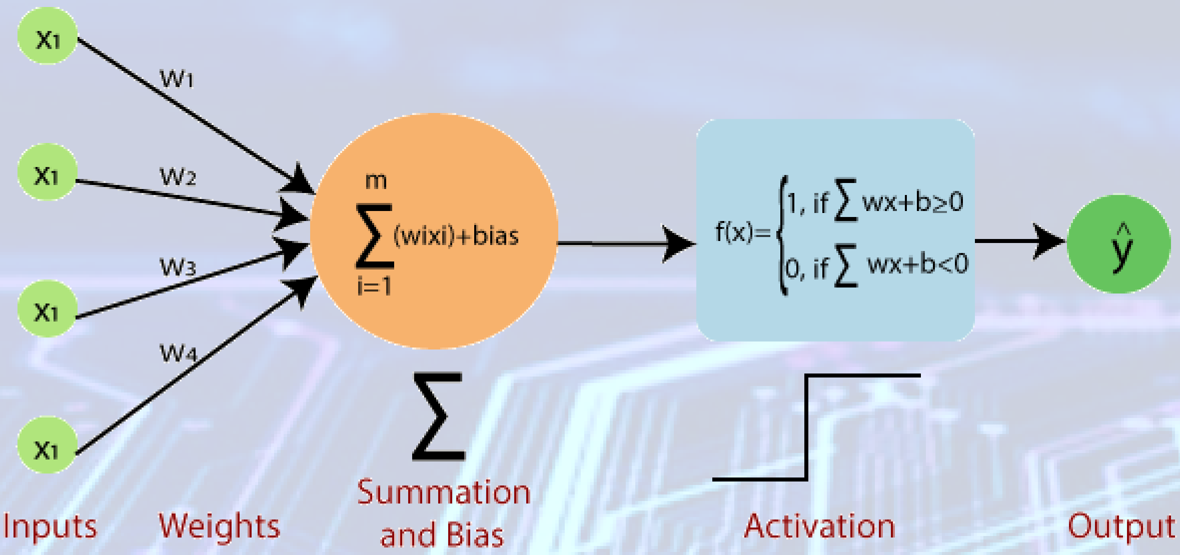


Main ingredients

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/

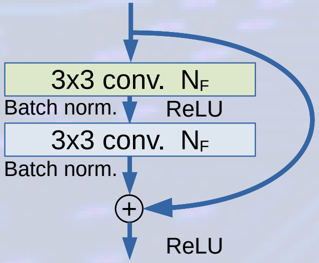
Popular architectures

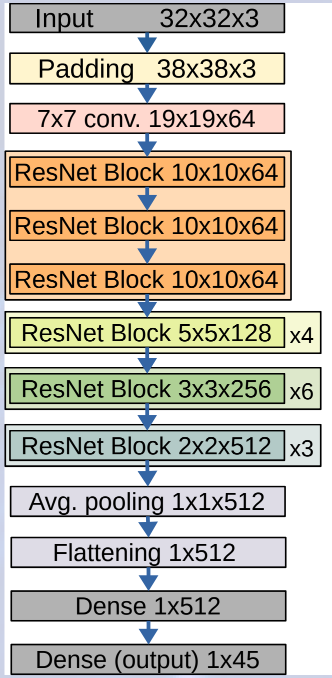
Stacking more layers: solve complex problems more efficiently, get highly accurate results

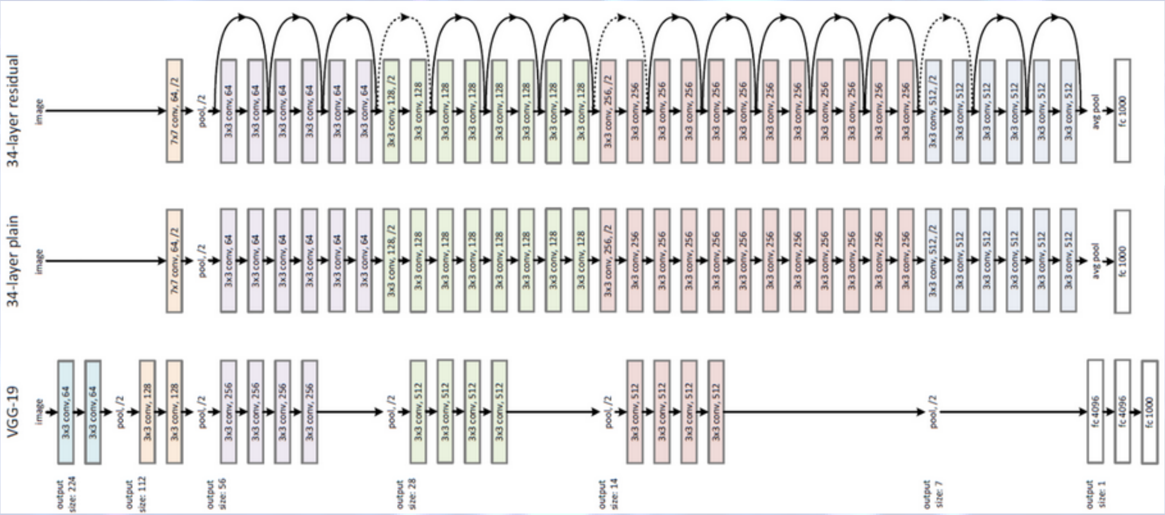
BUT:

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

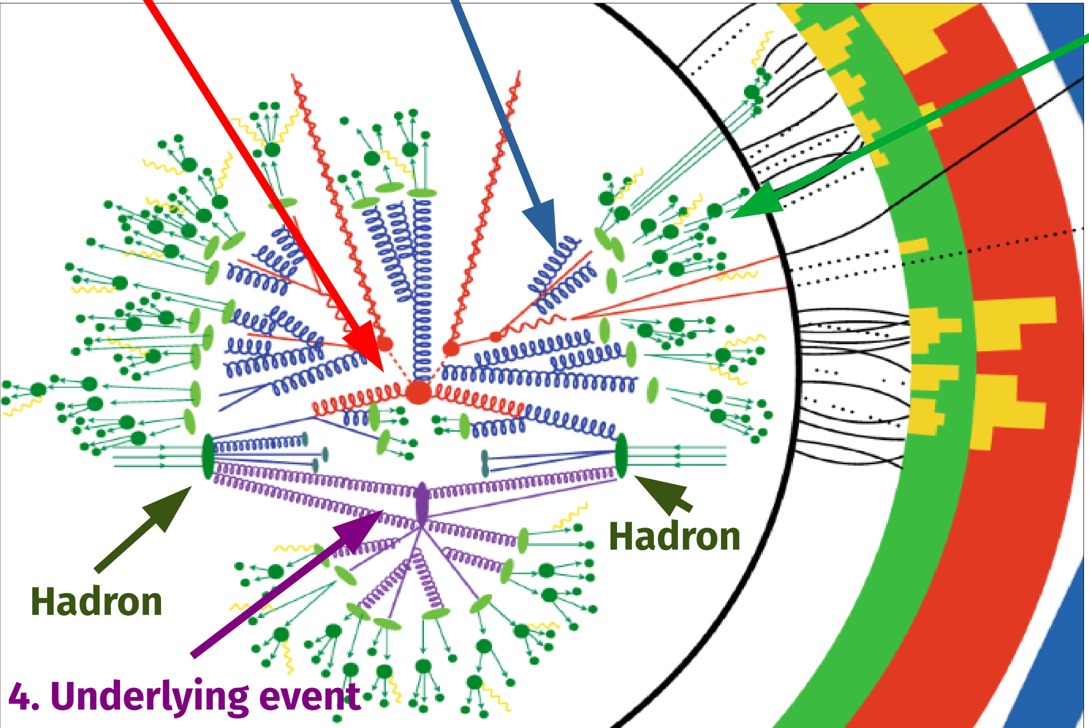
- Track reconstruction
- Quark/gluon jet separation
- Jet reconstruction
- Tuning Monte Carlo event generators
- GAN of detectors

Parton shower and hadronization

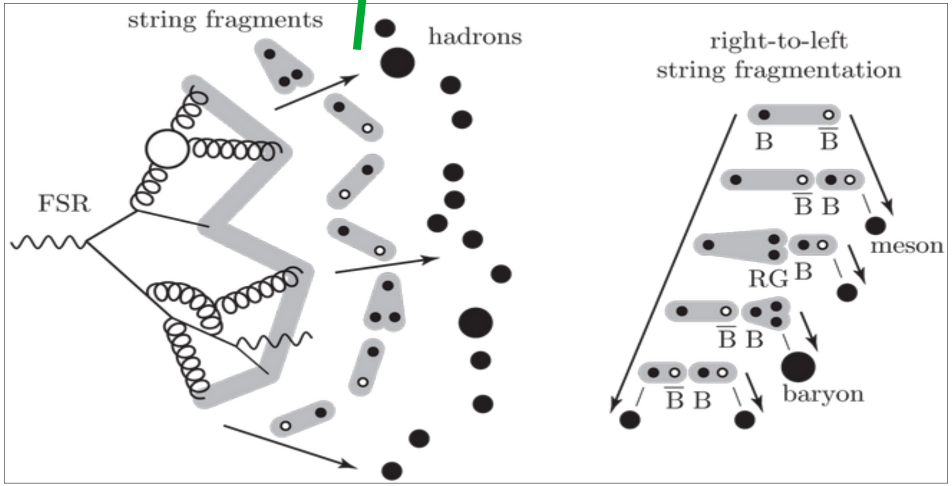
3. Hadronization

2. Parton shower

1. Hard scattering



Nature Reviews Physics 3, 73 (2021)



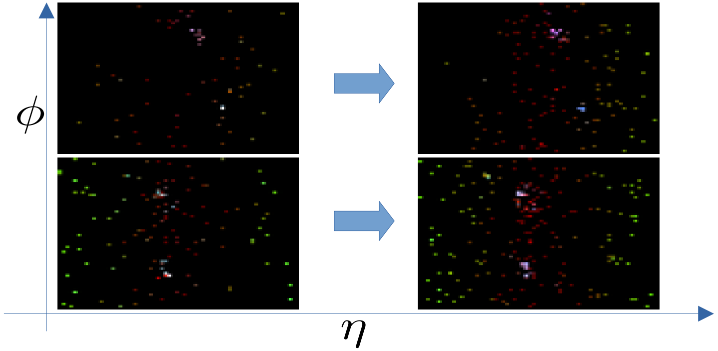
Hadronization

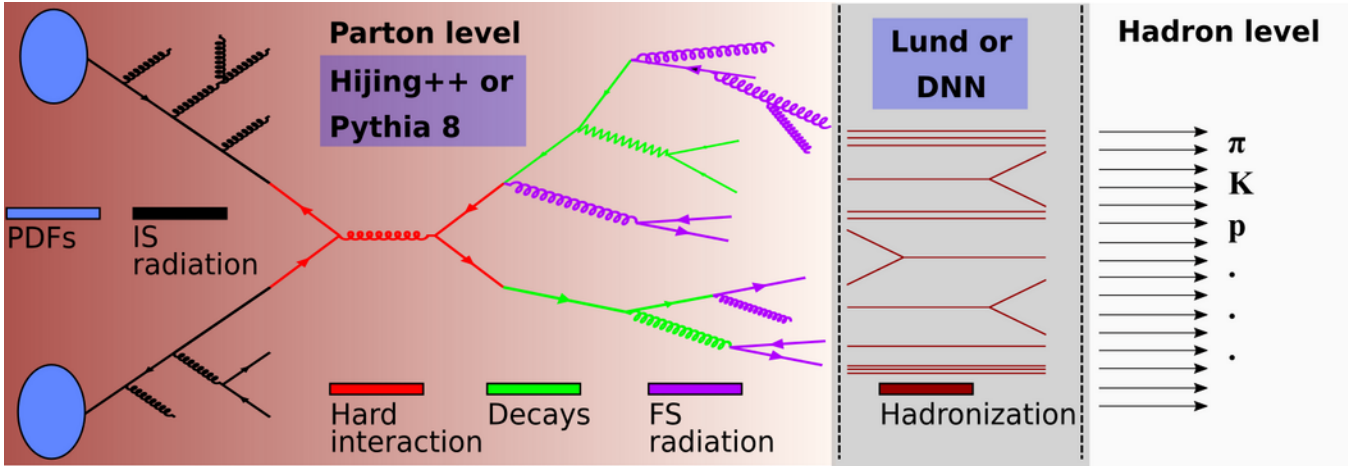
Partons → hadrons

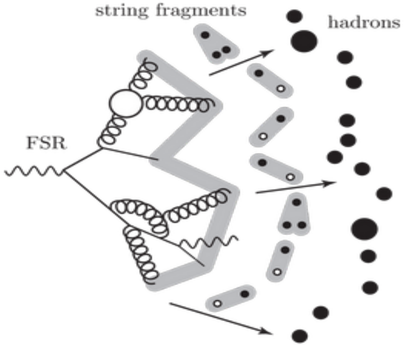
Non-perturbative process

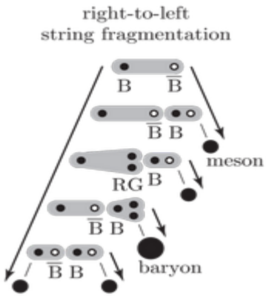
Lund-fragmentation (Comput.Phys.Commun. 27 (1982) 243)

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









Train and validation sets

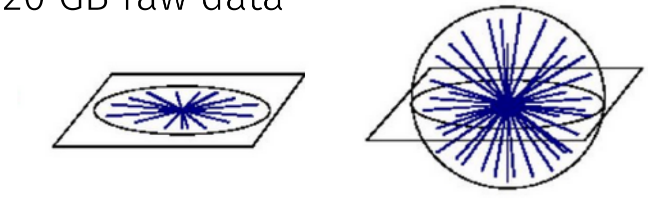
Monte Carlo data: Pythia 8.303

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



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

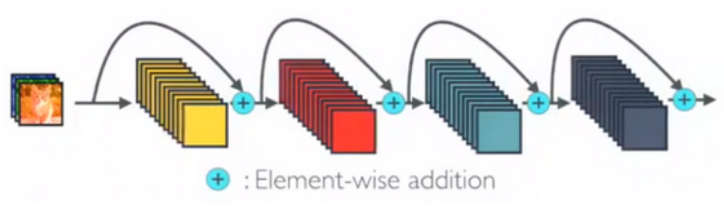
Stacking more layers: solve complex problems more efficiently, get highly accurate results

BUT:

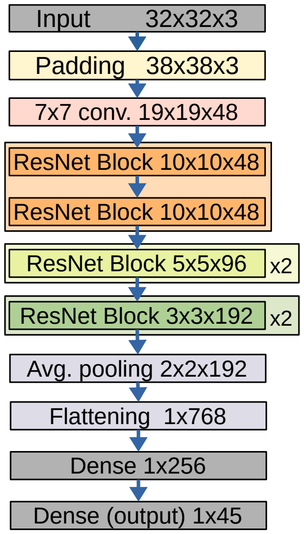
Vanishing/exploding gradients

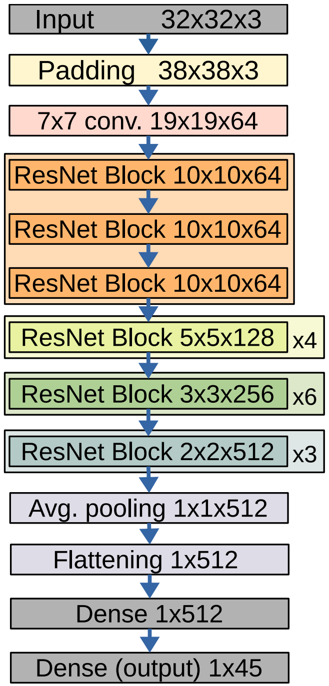
ResNet:

Residual blocks with "skip connections"



	Model S	Model L
Trainable parameters	1.7 M	20 M



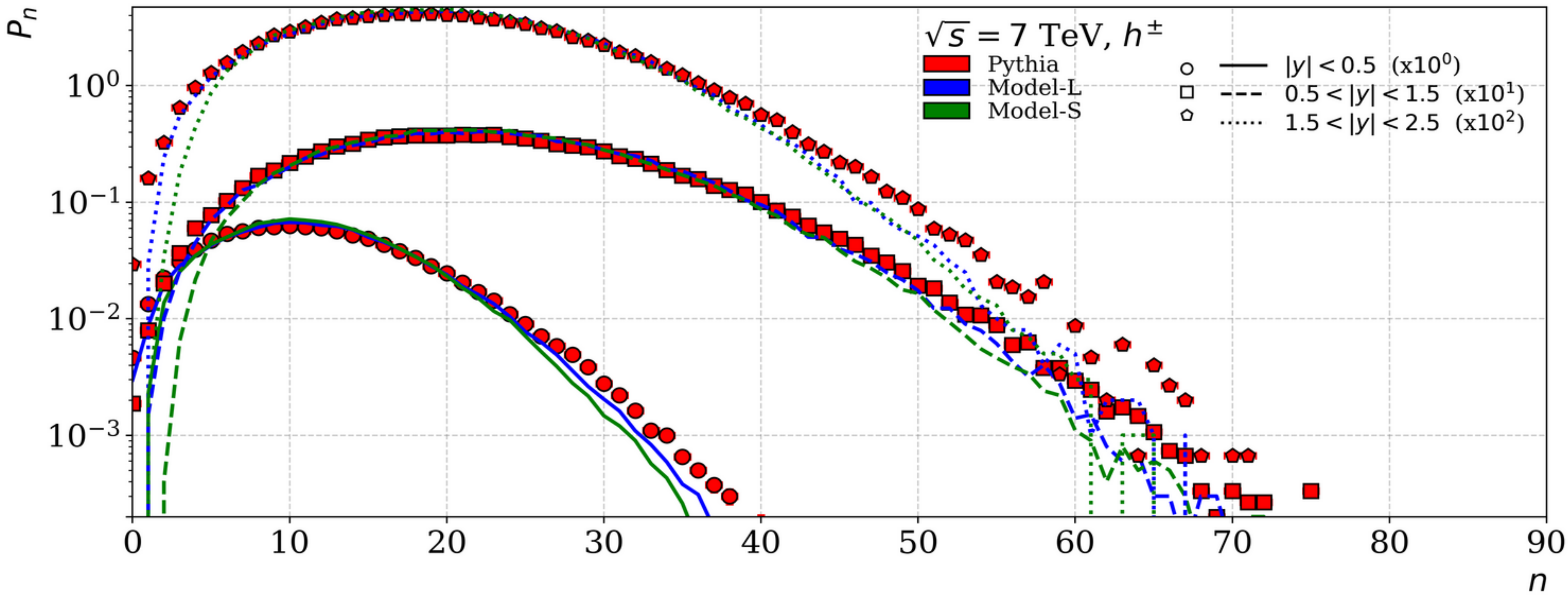


Used hardwares: Nvidia Tesla T4, GeForce GTX 1080
@ Wigner Scientific Computing Laboratory

Framework: Tensorflow 2.4.1, Keras 2.4.0

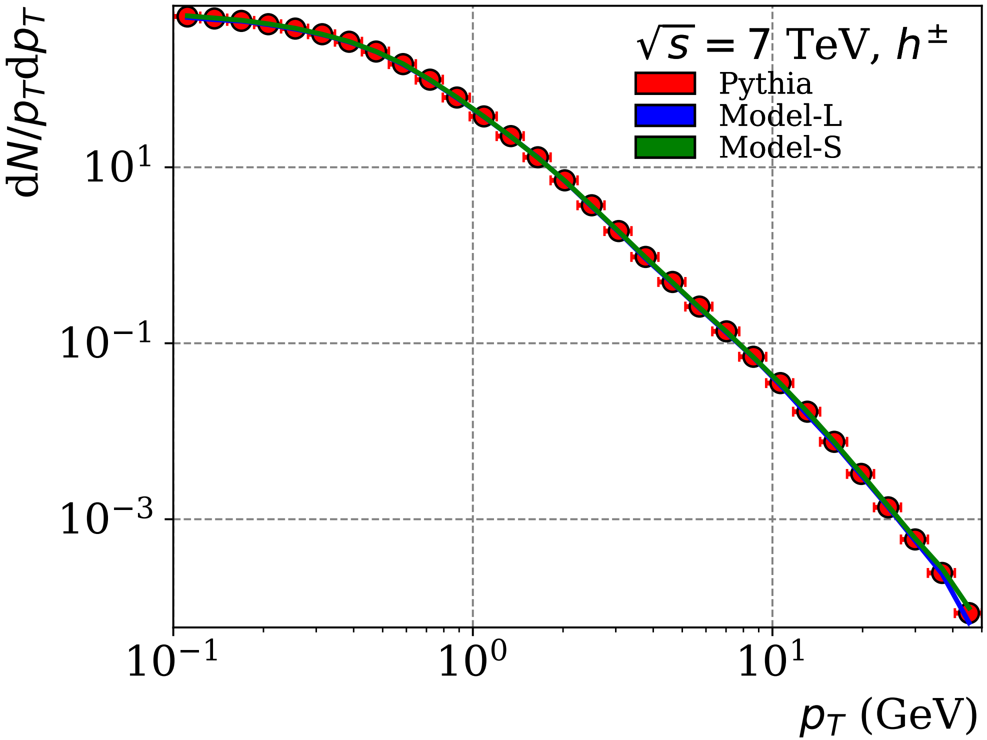
Results

Proton-proton @ 7 TeV, Training + Validation



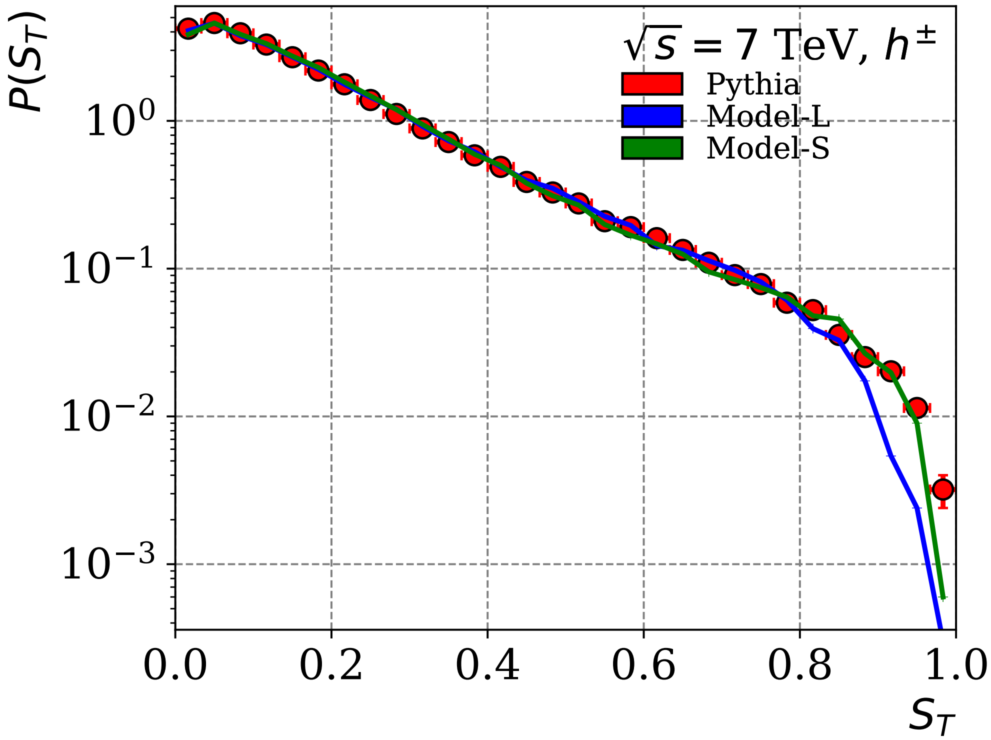
Charged hadron multiplicity at various rapidity windows
Comparison to reference MC model
Good agreement for both models

Proton-proton @ 7 TeV, Training + Validation



Charged hadron transverse momentum

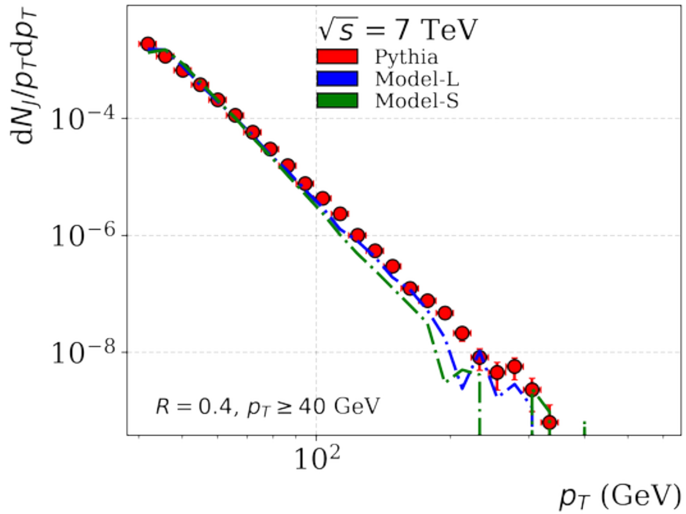
0.1 GeV ≤ p_T ≤ 50 GeV



Event transverse sphericity

The smaller model performs better

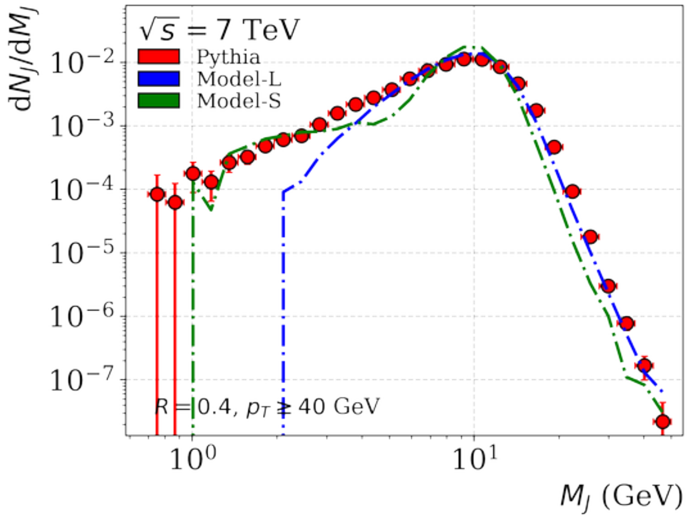
Proton-proton @ 7 TeV, Training + Validation

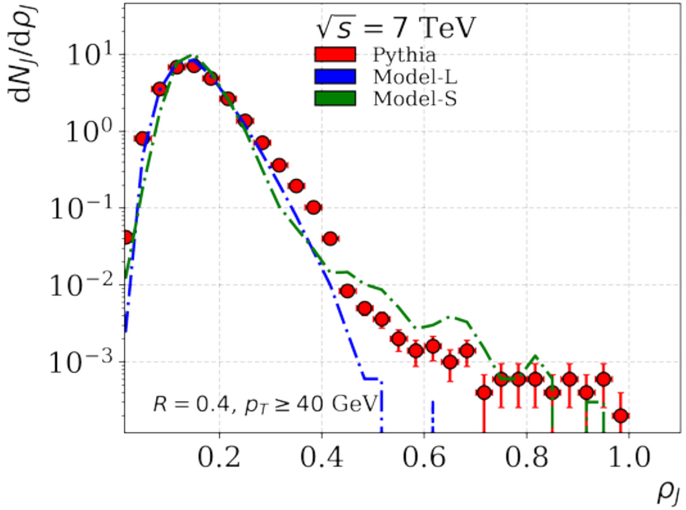


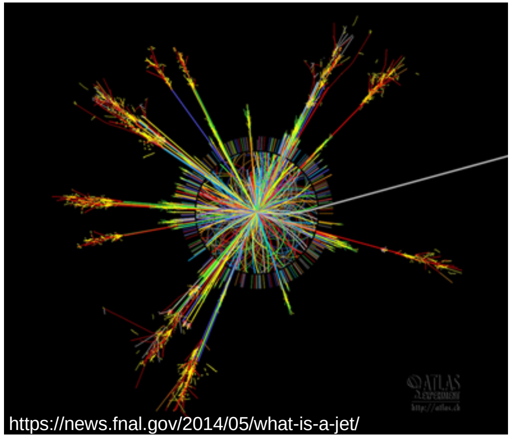
Jets:

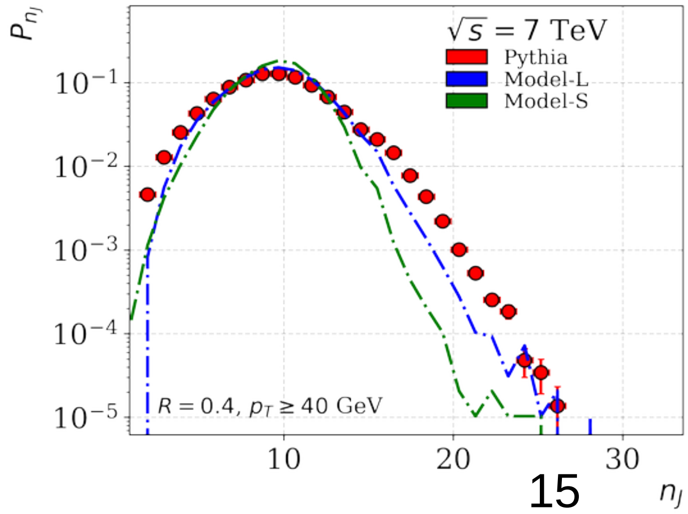
- Mean $p_T \leq 400$ GeV
- Mean mass $p_T \leq 400$ GeV
- Mean multiplicity
- Mean width

The smaller model performs better



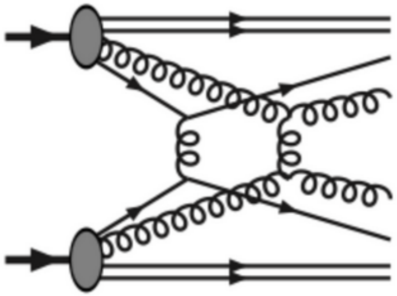




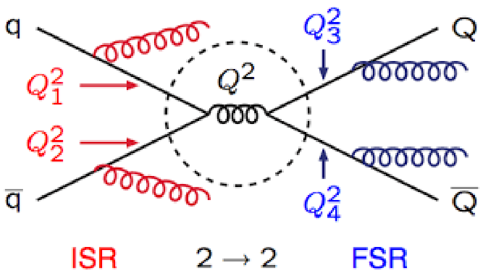


Proton-proton @ 7 TeV, Training + Validation

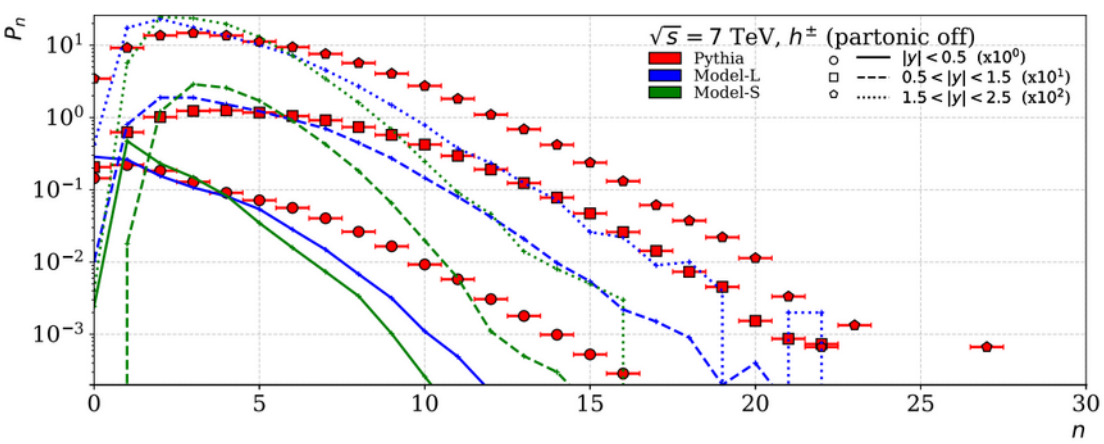
(*) What about the partonic processes?

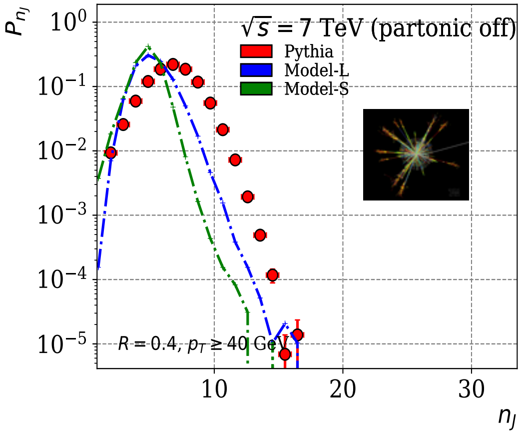


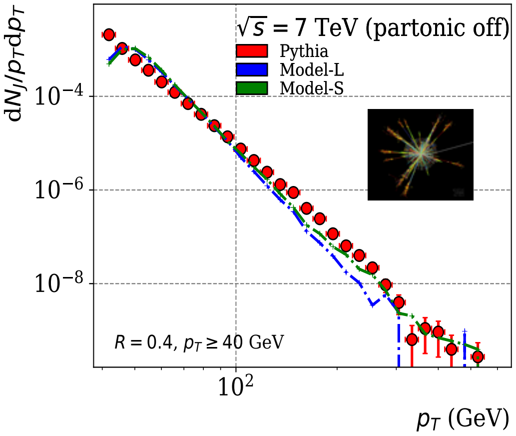
MPI

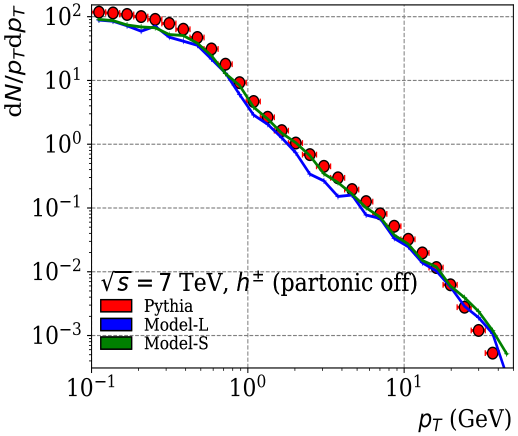


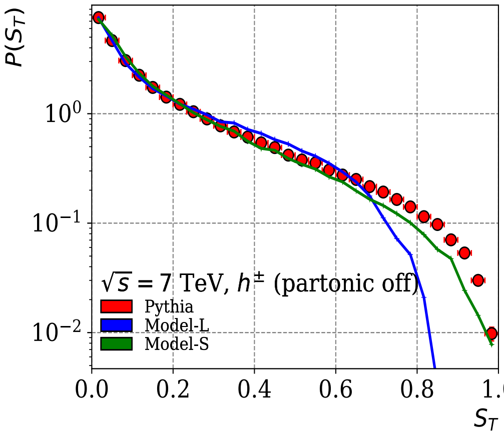
http://home.thep.lu.se/~torbjorn/talks/cern18cosmic.pdf





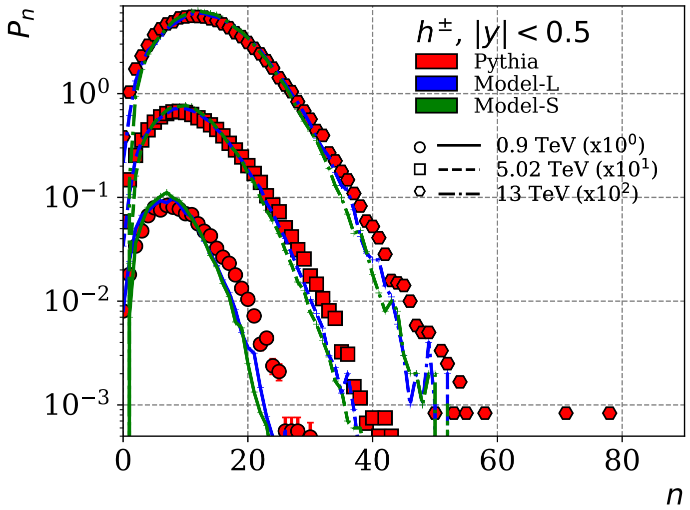


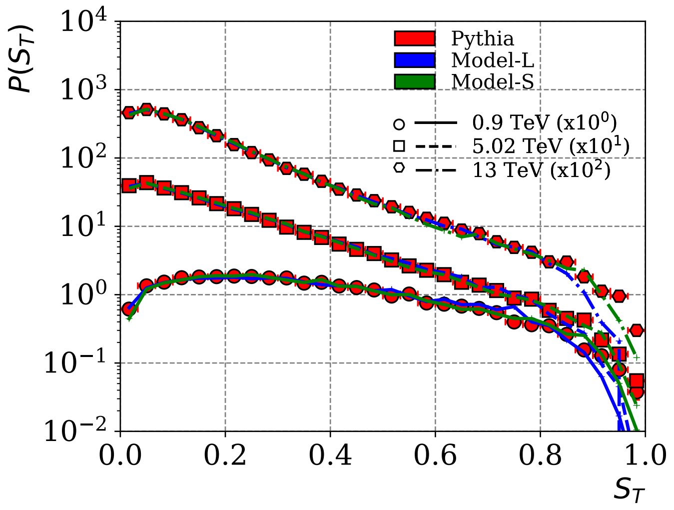


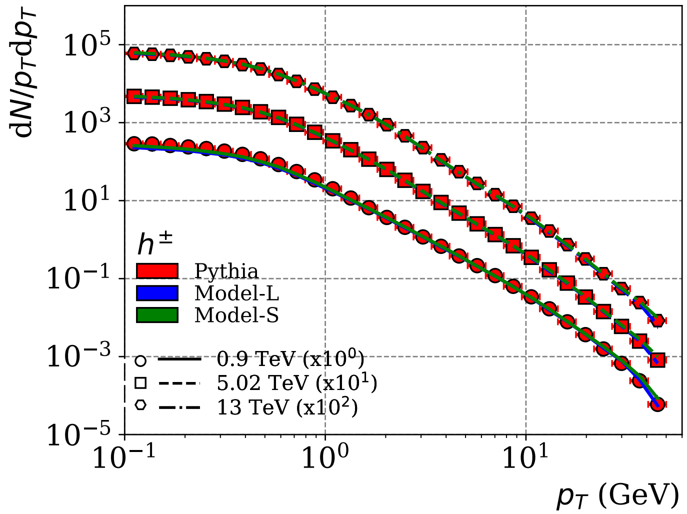


Qualitative agreement → the models adopted the hadronization properties

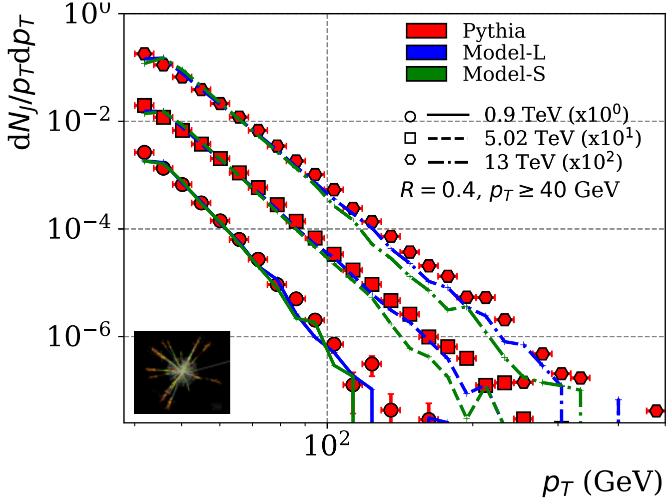
Proton-proton @ 0.9-13 TeV, Predictions







- 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?**



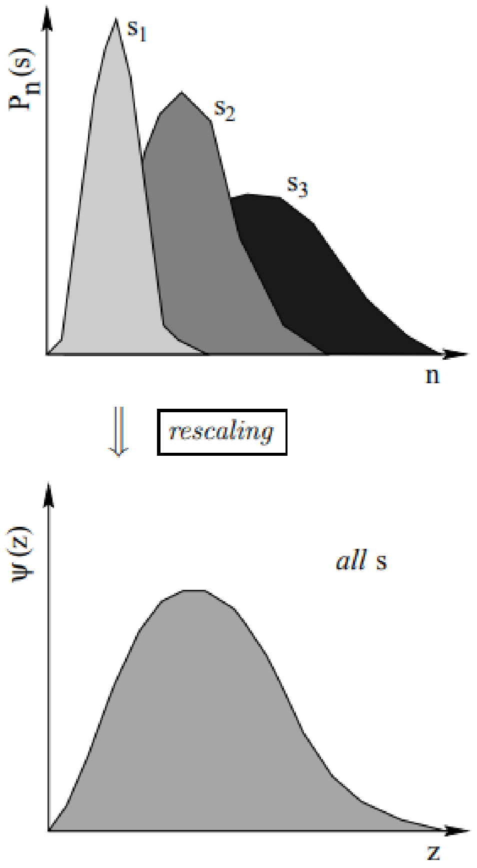
KNO-scaling

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

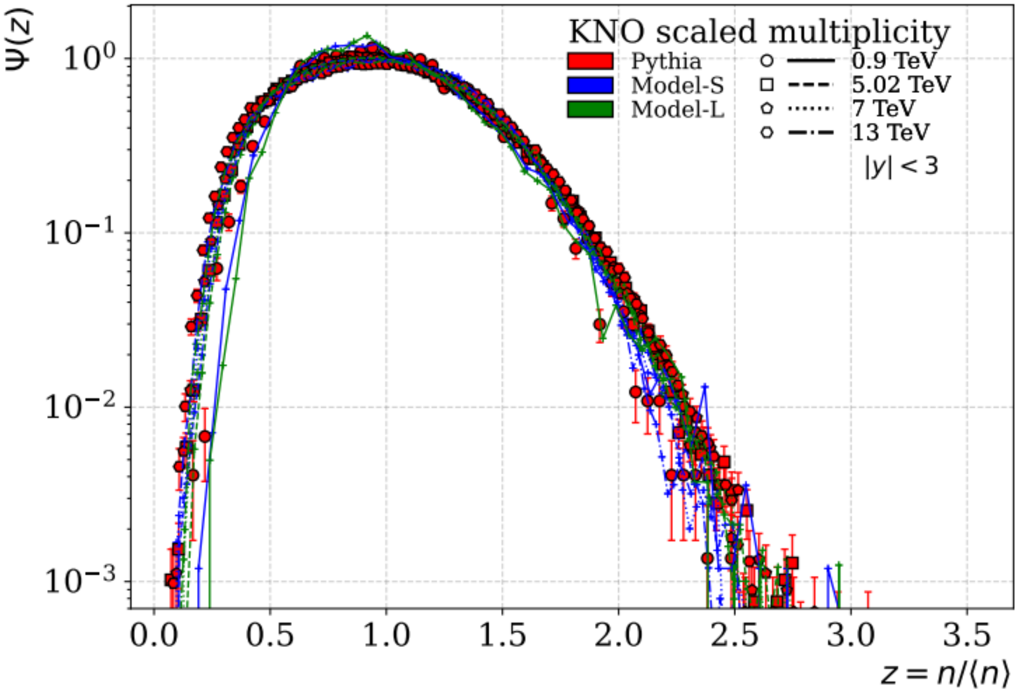
Violation of the scaling at high CM energies: not fully understood (relation to MPI?)

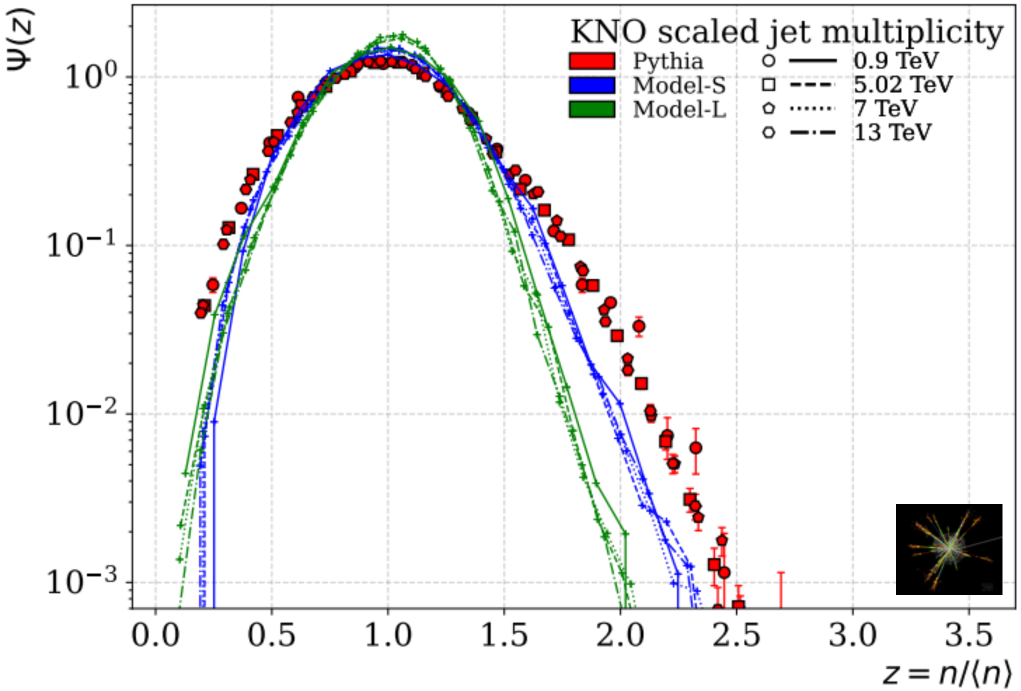


Nuclear Physics B 40 (1972), 317–334.

(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)”

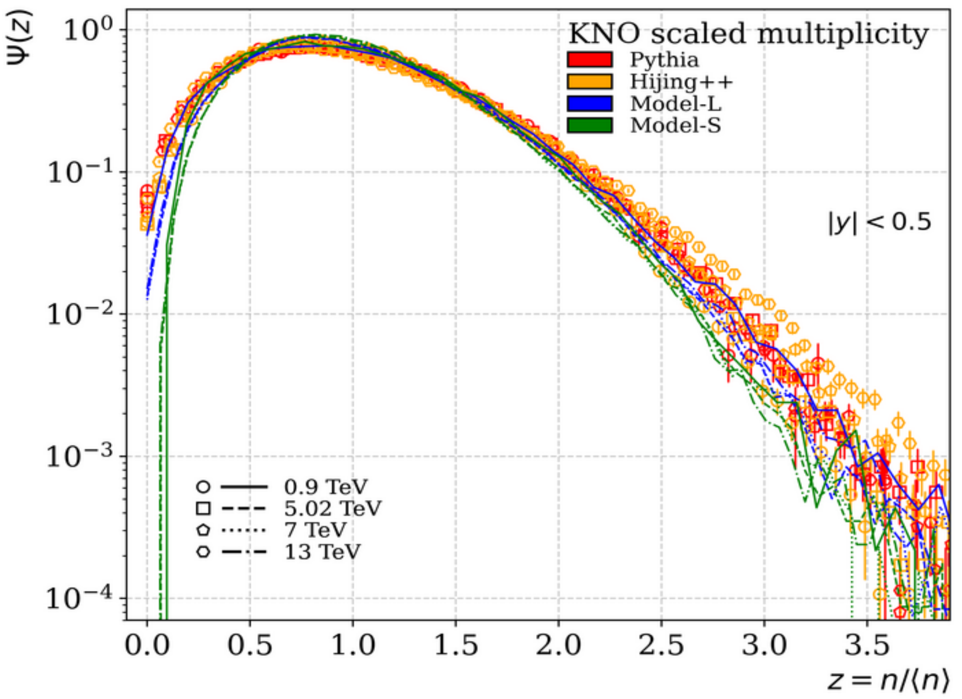
First, FORTRAN version: 1991, X.N. Wang, M. Gyulassy, **Phys. Rev. D 44, (1991) 3501.**

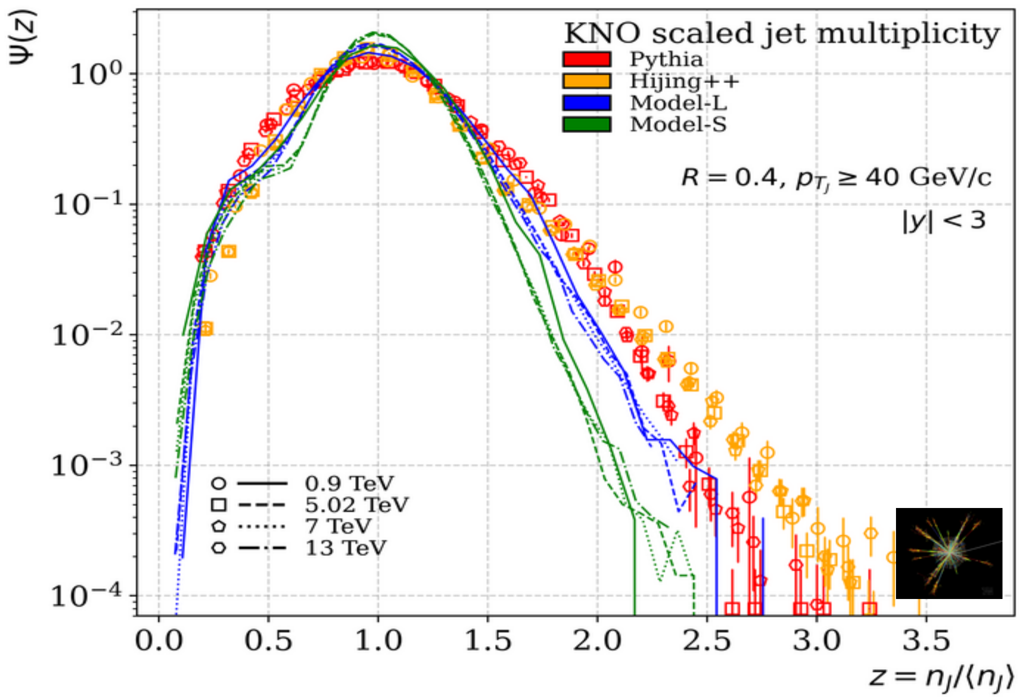
Computational challenge: more than 600 million collision in each second → HiLumiLHC: even more

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)

25 YEARS OF EPJ.org your physics journal

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.

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

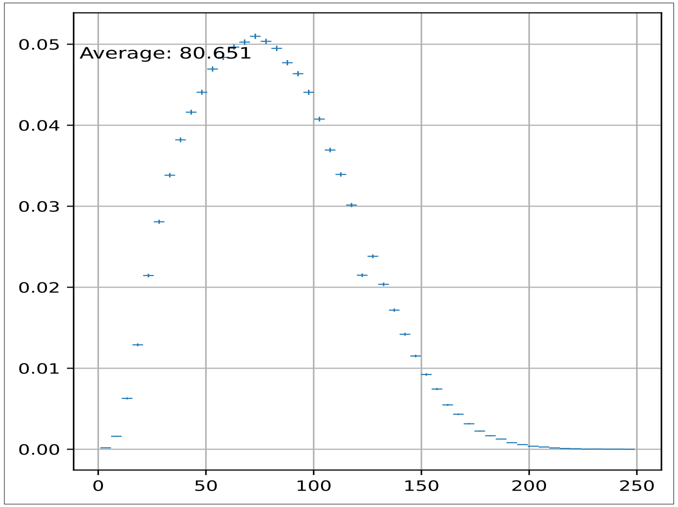
Reduction with Singular Value Decomposition:

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

- 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)