Universidade do Minho
Escola de Ciências

LABORATÓRIO DE INSTRUMENTAÇÃO
E FÍSICA EXPERIMENTAL DE PARTÍCULAS
partículas e tecnologia

DATA SCIENCE *in* (ASTRO)PARTICLE PHYSICS AND BEYOND

Nuno Castro
nfcastro@lip.pt

many thanks to Inês Ochoa, Rute Pedro, Miguel Romão, Maria Gabriela Oliveira, Miguel Caçador Peixoto, Paulo Brás, Ruben Conceição and Joana Sá for the useful discussions and for providing material for this talk

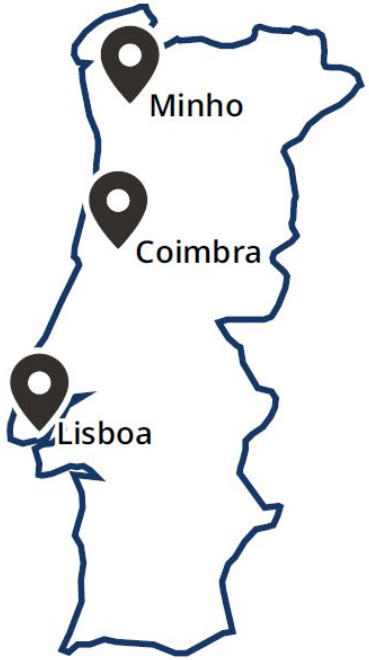
Fundação
para a Ciência
e a Tecnologia

UIDP/50007/2020
LA/P/0016/2020
CERN/FIS-COM/0004/2021

LIP

Laboratório de Instrumentação e Física Experimental de Partículas

- LIP is the reference laboratory for experimental particle physics and associated technologies in Portugal
- LIP exists for the discovery of the fundamental laws of the Universe, ensuring the full participation of the Portuguese scientific community in this endeavour, and to share this knowledge with society
- The laboratory is nation-wide, with nodes in Lisbon, Coimbra and Braga, in close collaboration with the local universities



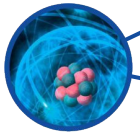
LIP

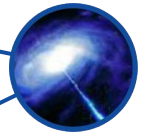
Laboratório de Instrumentação e Física Experimental de Partículas

Experimental particle and **astroparticle Physics**

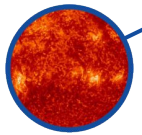
Structure of **matter**
- pQCD
- HADES
- NUC-RIA
- NPstrong

LHC experiments and **phenomenology**
- ATLAS
- CMS
- Phenomenology
- FCC

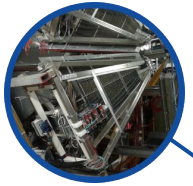




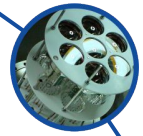
Cosmic Rays
- AMS
- Auger
- SWGO



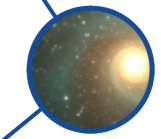
Dark Matter and **neutrinos**
- Dark Matter
- Neutrinos
- SHiP



Development of **new instruments** and **methods**



Detectors for **particles** and **nuclear physics**
- RPC R&D
- Neutron detectors
- Gaseous Detectors R&D
- Liquid Xenon R&D

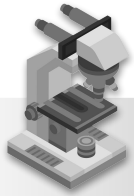


Health and **biomedical applications**
- RPC-PET
- Proton therapy
- OR Imaging
- Dosimetry

Space applications
- Space Rad
- i-Astro

Computing
- GRID - Distributed Computing and Digital Infrastructures
- SPAC - Social Physics and Complexity

R&D



observable universe
8.8 10^{26} m

visible with our own eyes

quarks
< 10^{-19} m

1·000·000·000·000·000·000·000·000·000.000·000·000·000·000·000·01 meter

distance to galactic center

distance light travels in one year

farthest human object from Earth (Voyager 1)

distance Earth-Sun

biological cell

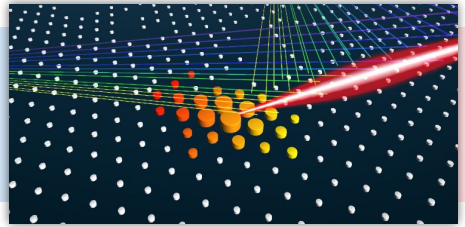
atoms

proton neutron

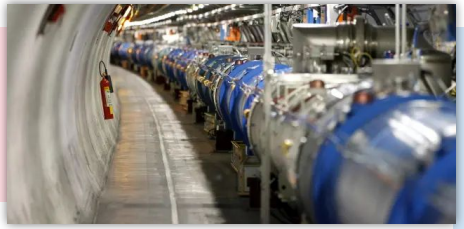
(slide by Jorgen d'Hondt)

age universe
$4.4\ 10^{18}$ s

large surface/volume observatories



powerful accelerators



lifetime top quark
$5\ 10^{-25}$ s

observable universe
$8.8\ 10^{26}$ m

visible with our own eyes

quarks
$< 10^{-19}$ m

1·000·000·000·000·000·000·000·000·000.000·000·000·000·000·000·01 meter

distance to galactic center

distance light travels in one year

farthest human object from Earth (Voyager 1)

distance Earth-Sun

biological cell

atoms

proton neutron

lifetime star
10^{13}-10^{16} s

duration supernova & GRB
0.1-100 s

lifetime proton
$> 3\ 10^{41}$ s

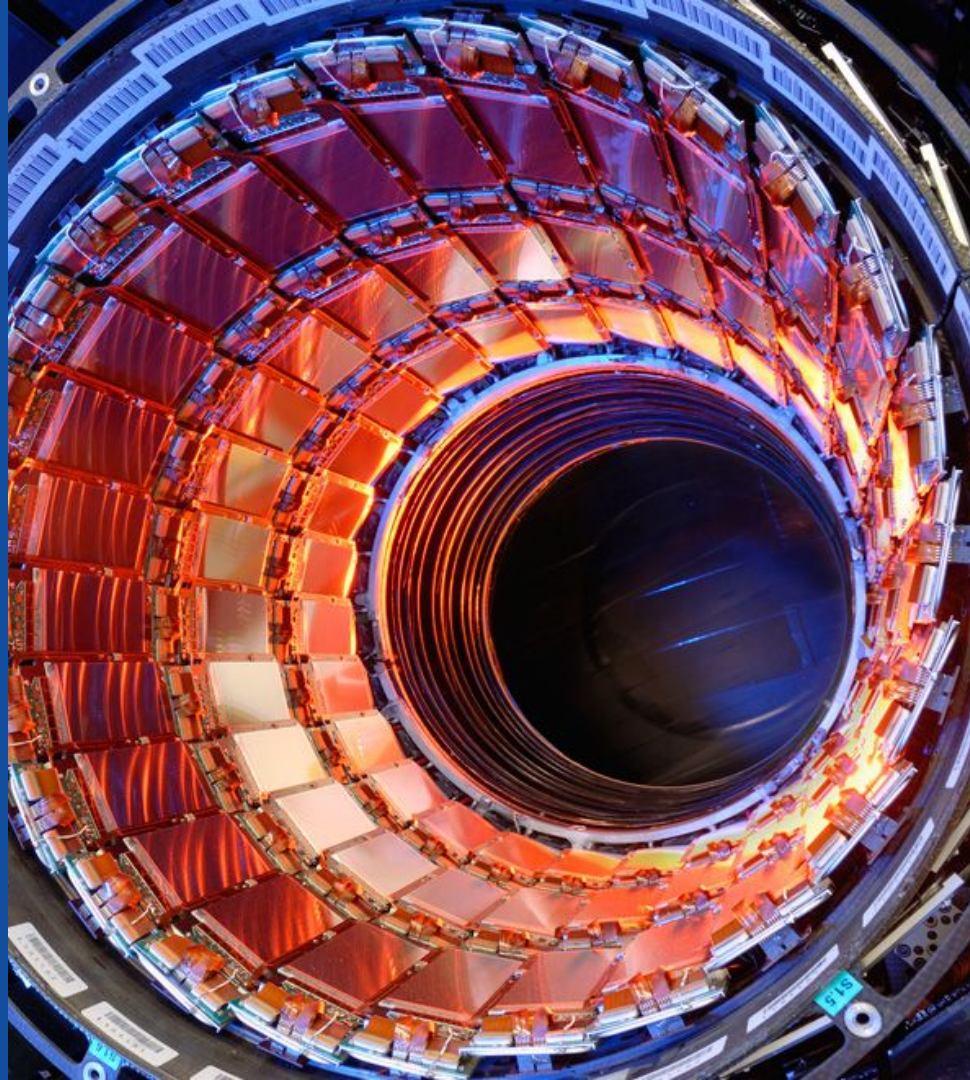
lifetime kaon ($K^{\pm}$)
$1.2\ 10^{-8}$ s

CERN

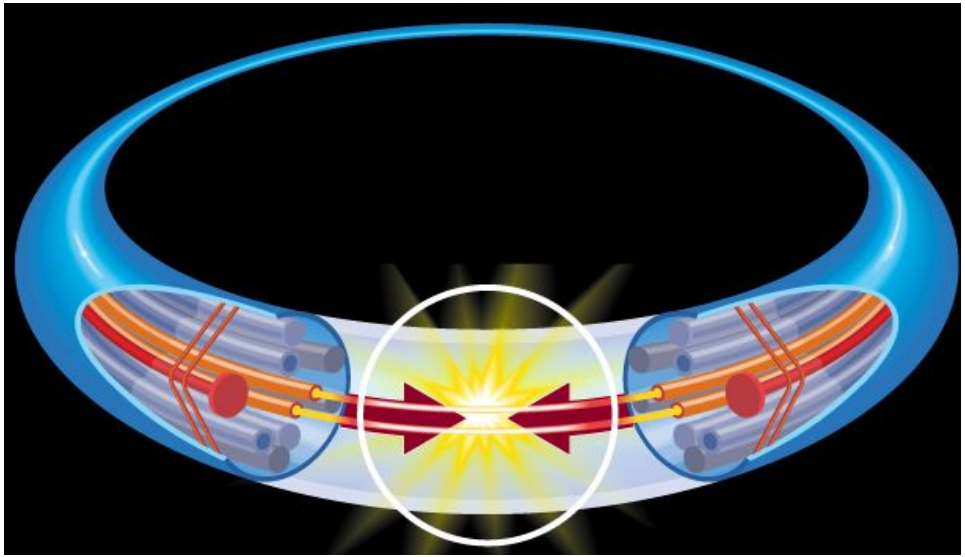
- The European Laboratory for Particle Physics
- Located in the franco-swiss border
- Portugal is a member since 1986, with LIP being the reference portuguese partner

in the frontier of the technology

- Particle physics accelerators and detectors are amongst the most complex devices built by the humankind
- Being on the edge of the technology is required



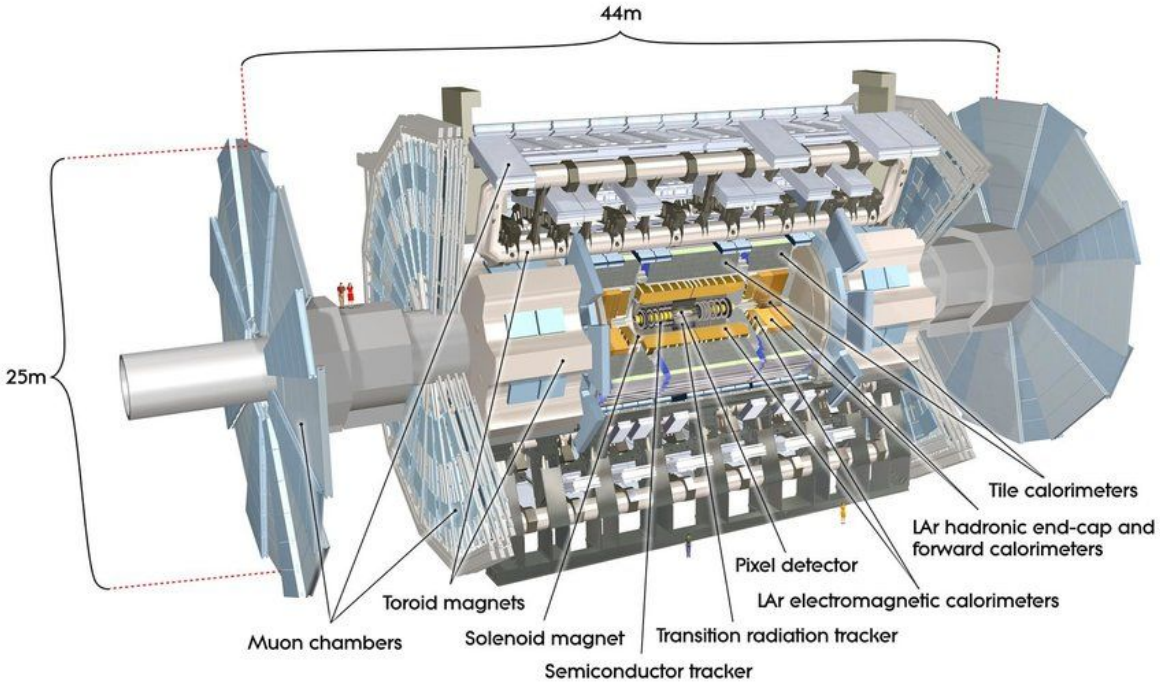
from data to physics at the Large Hadron Collider
a long and complex path



- 40 million proton-proton collisions per second

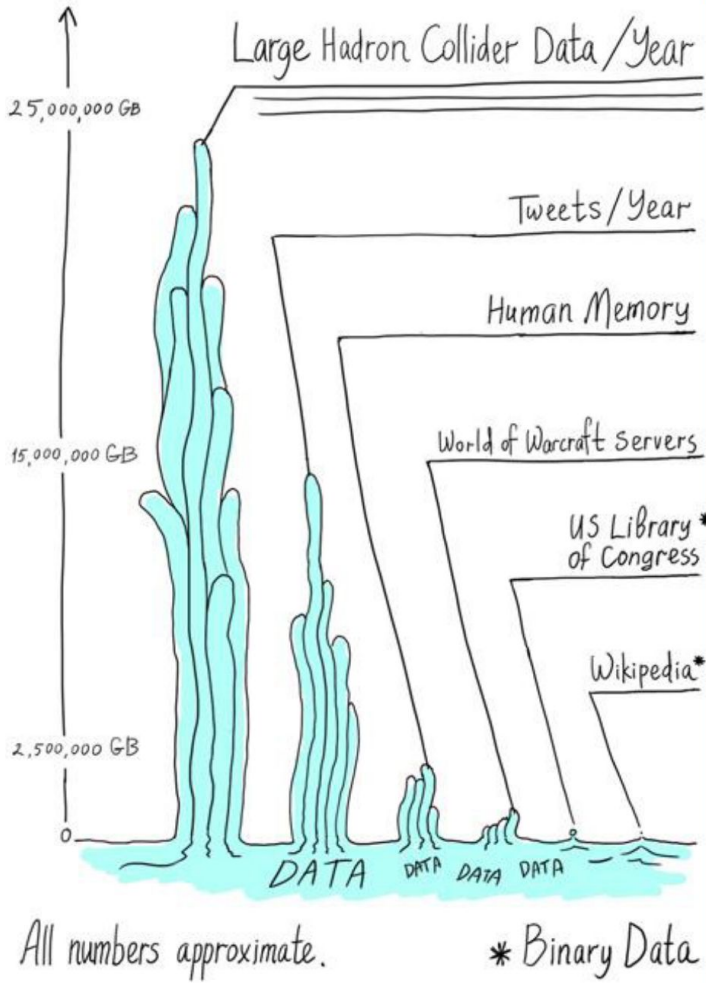
from data to physics at the Large Hadron Collider
a long and complex path

- hundreds of millions of readout channels

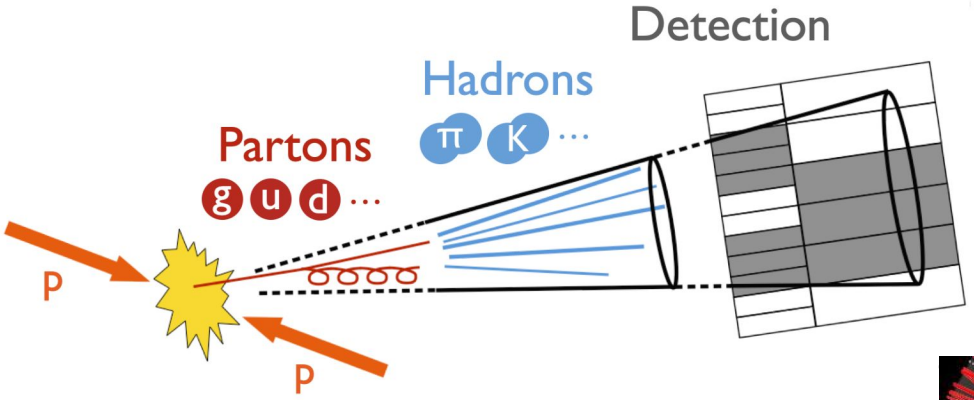


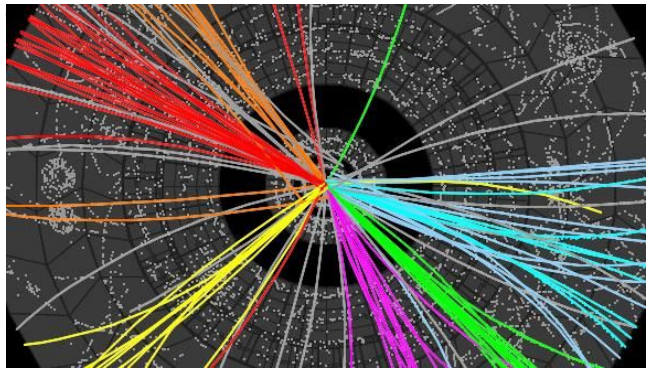
Large Hadron Collider

data, data, data, ...



tagging special objects in collisions

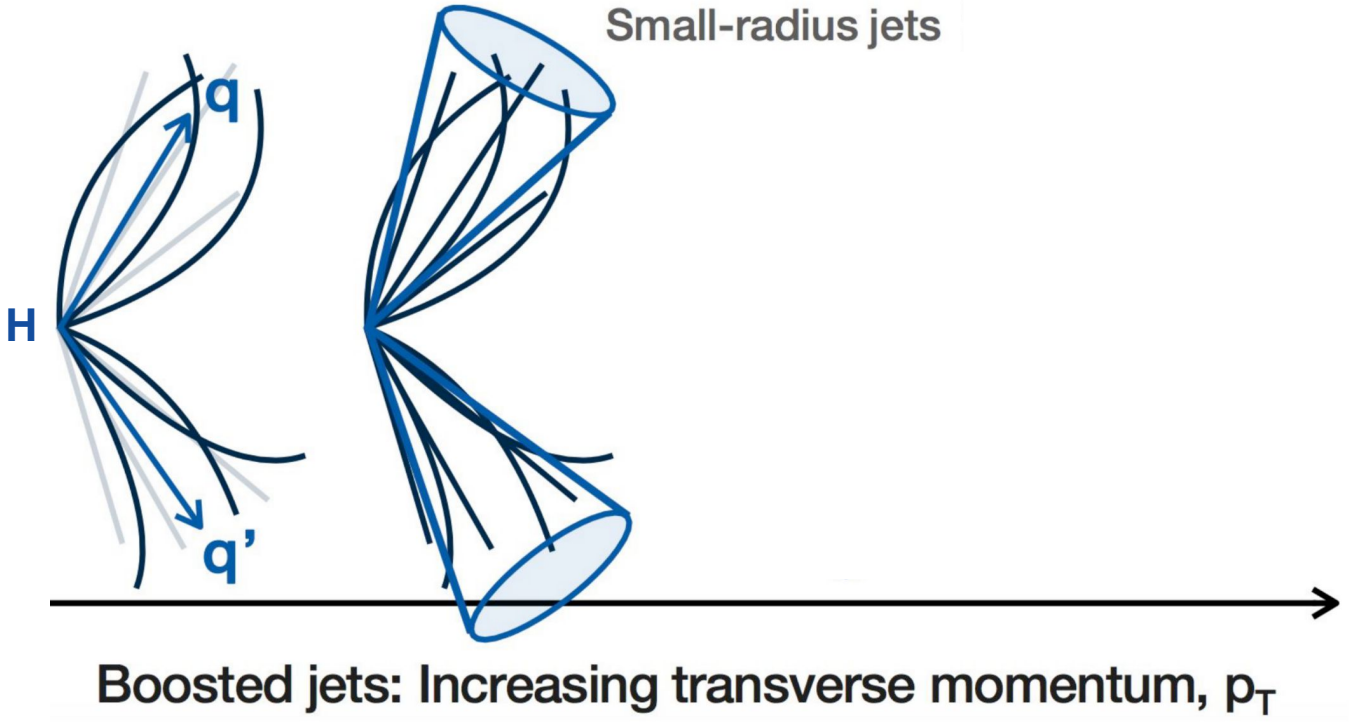




tagging special objects in collisions

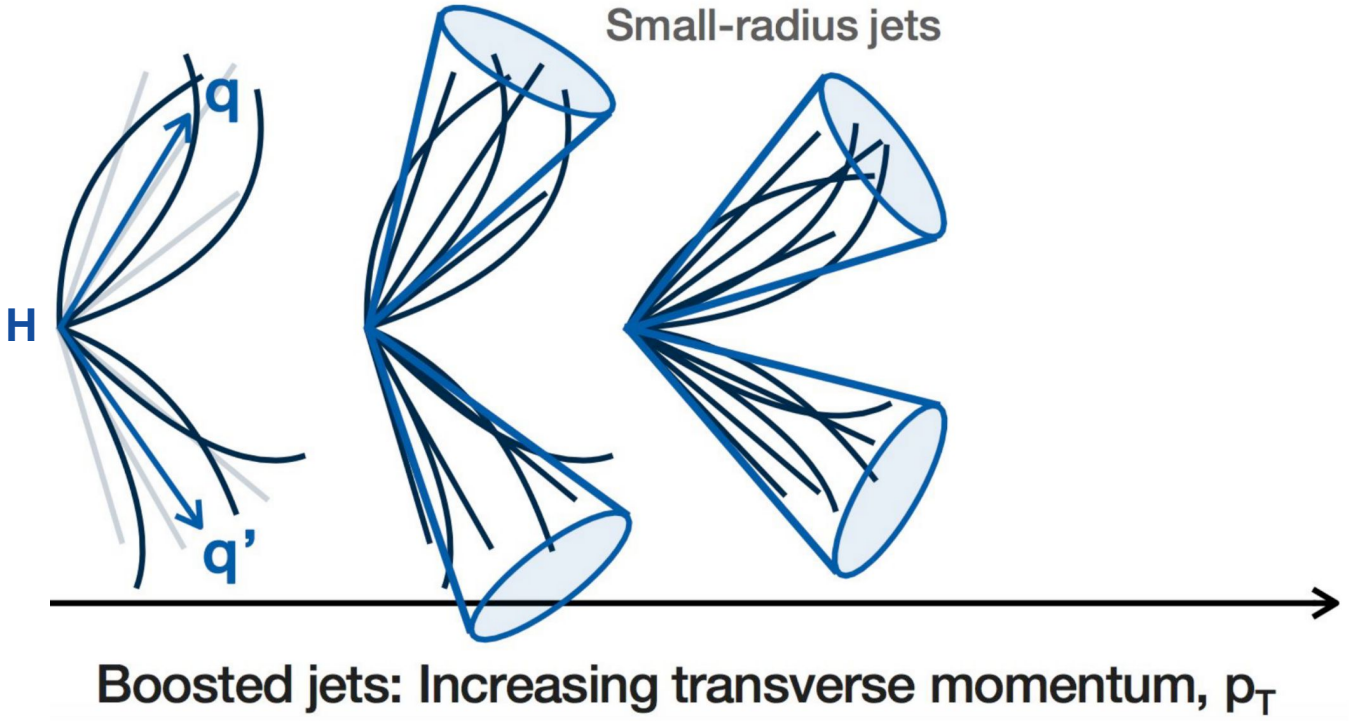
tagging the Higgs boson

tagging special objects in collisions
tagging the Higgs boson



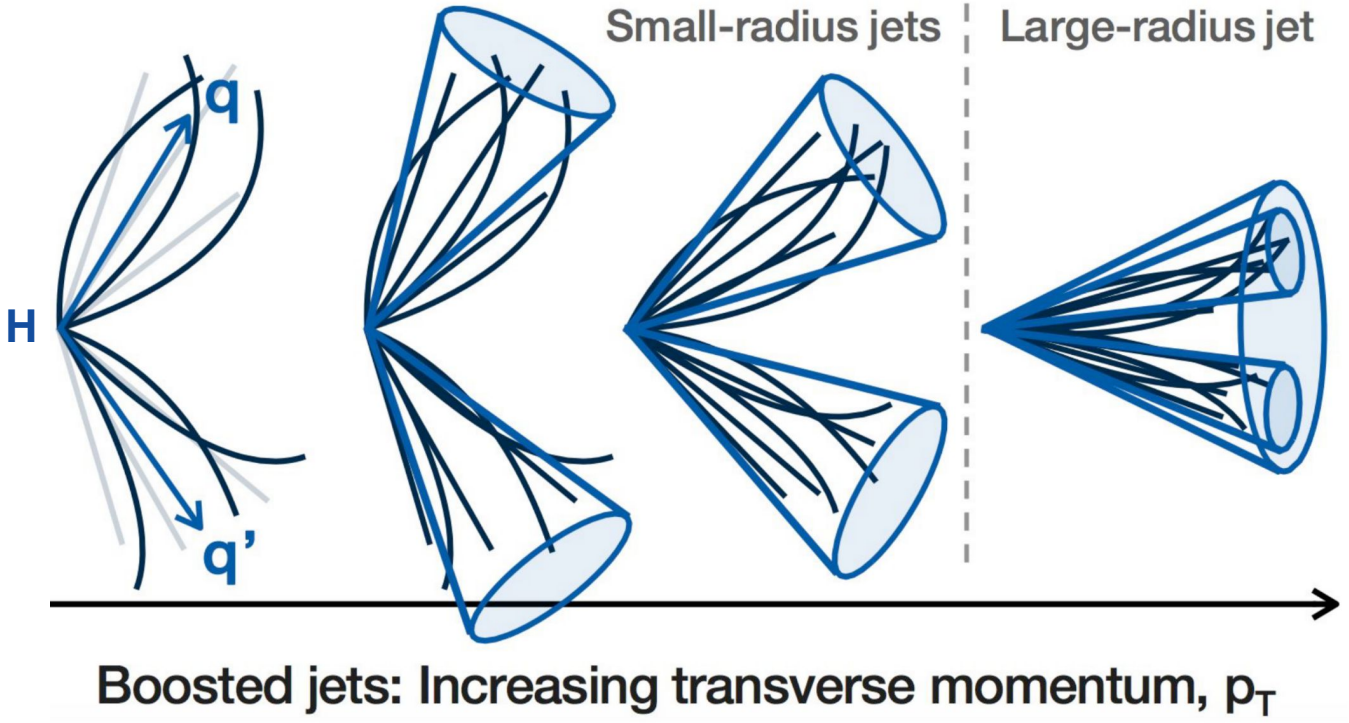
tagging special objects in collisions

tagging the Higgs boson



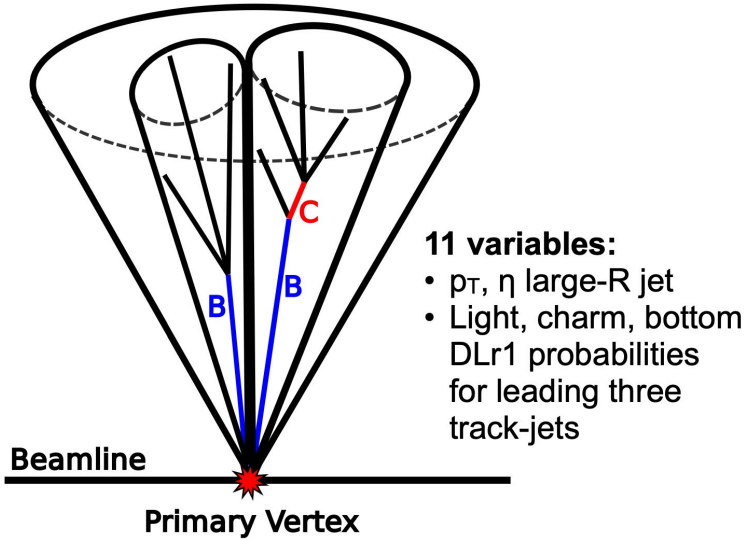
tagging special objects in collisions

tagging the Higgs boson



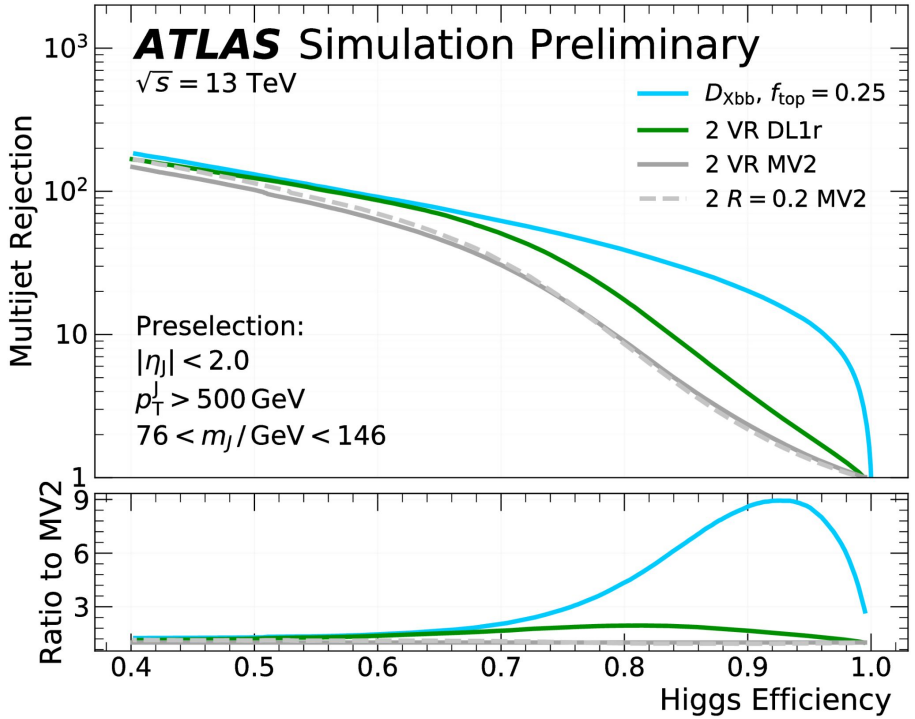
tagging special objects in collisions
tagging the Higgs boson

- Deep neural networks for supervised classification



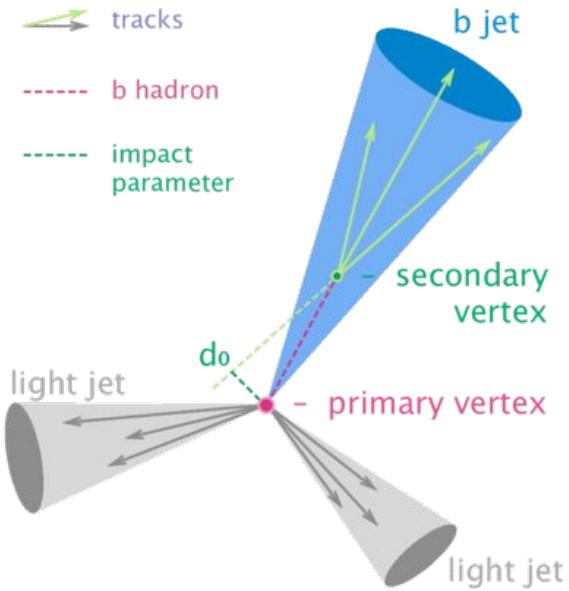
ATL-PHYS-PUB-2020-019

tagging special objects in collisions
tagging the Higgs boson

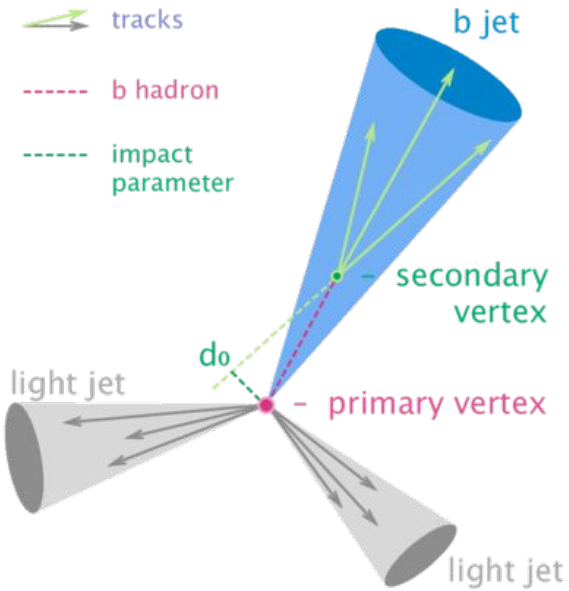


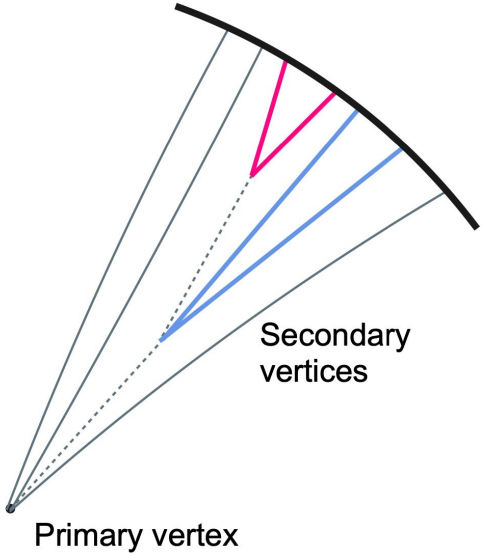
ATL-PHYS-PUB-2020-019

tagging special objects in collisions
tagging long-lived particles



tagging special objects in collisions
tagging long-lived particles

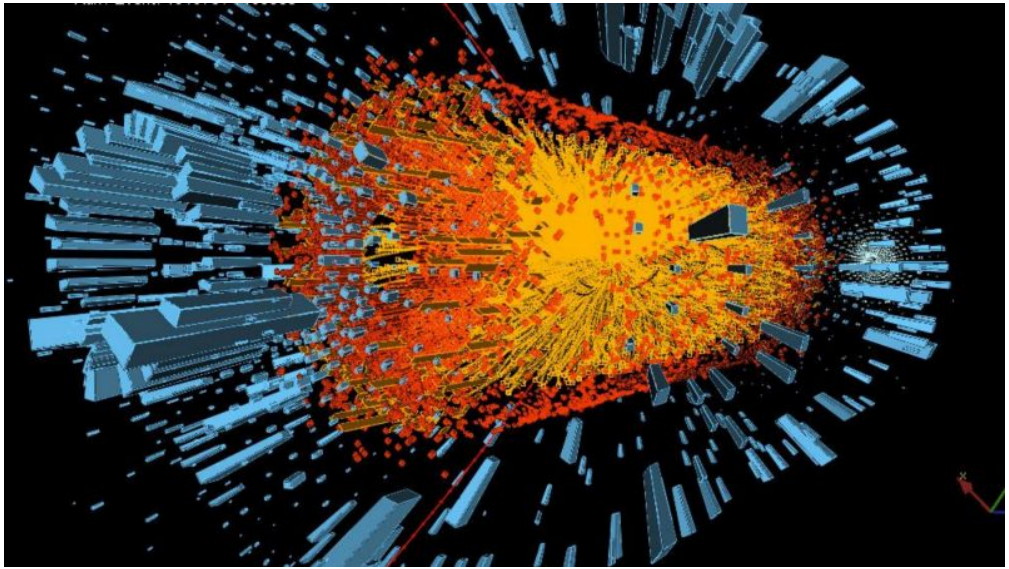




Exploring **graph-neural-networks** for b-tagging, via improved secondary vertexing finding
(very preliminary)

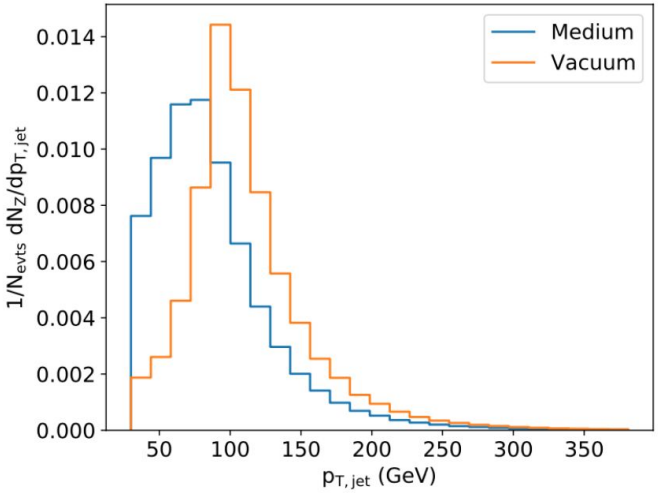
tagging special objects in collisions

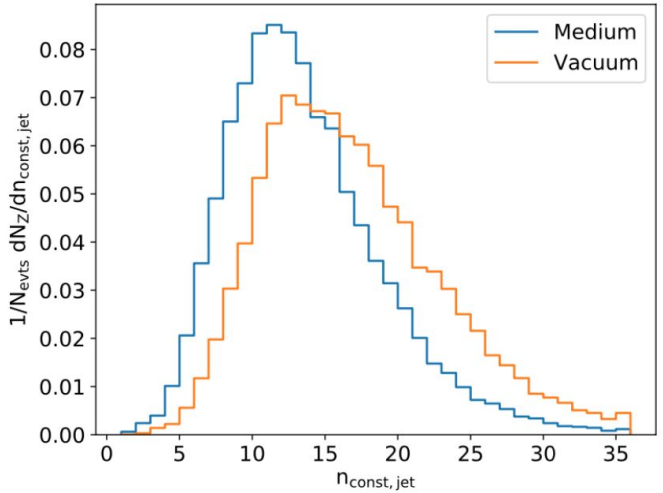
tagging new phenomena: the quark gluon plasma



tagging special objects in collisions

tagging new phenomena: the quark gluon plasma



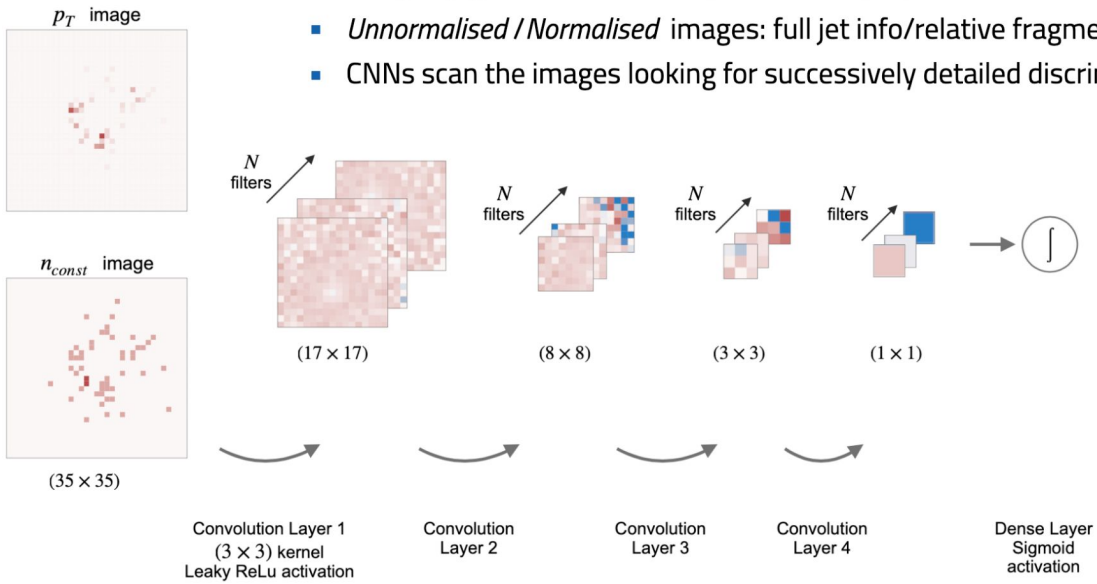


tabular data: use of a supervised dNN

JHEP11 (2021) 219

tagging special objects in collisions
tagging new phenomena: the quark gluon plasma

- 2 $\Delta\eta, \Delta\phi$ grids centred in the jet axis with jet pT and *n* constituents
- *Unnormalised / Normalised* images: full jet info/relative fragmentation pattern
- CNNs scan the images looking for successively detailed discriminant patterns

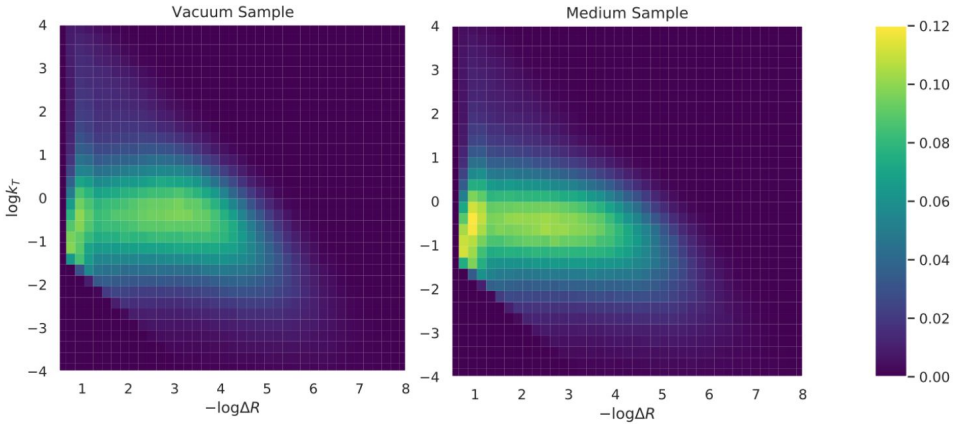


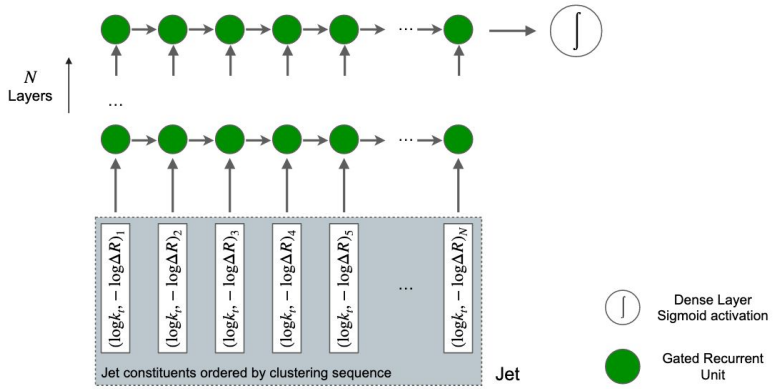
jet images: Convolutional Neural Networks (CNN)

JHEP11 (2021) 219

tagging special objects in collisions

tagging new phenomena: the quark gluon plasma



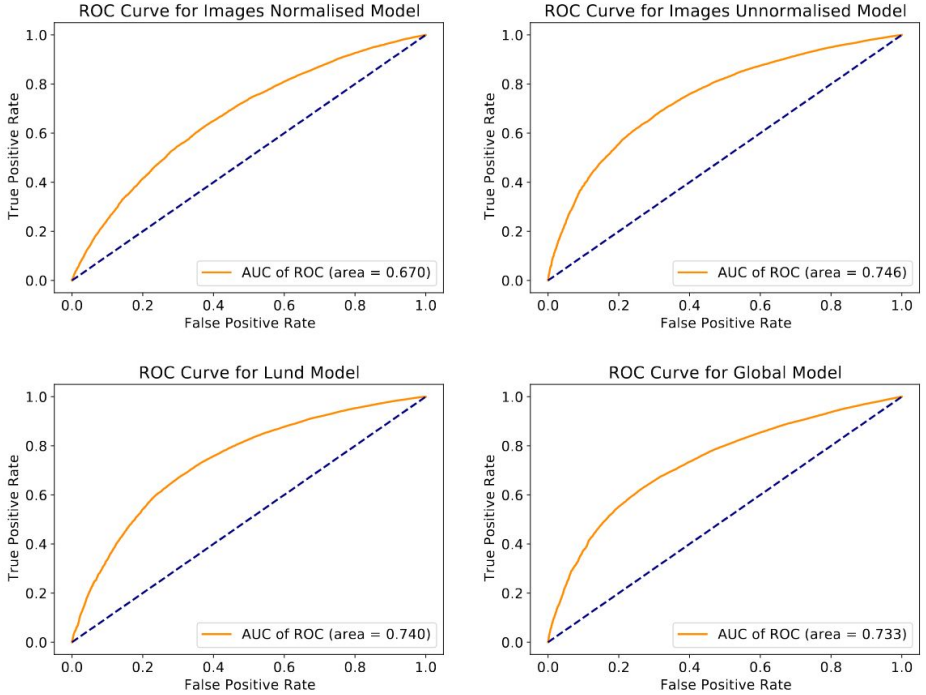


JHEP11 (2021) 219

Lund planes: Recurrent Neural Networks (RNN)

tagging special objects in collisions
tagging new phenomena: the quark gluon plasma

Model Type		Hyperparameter	Value
CNN (Images)	Normalised	Number of Filters	104
		Spatial Dropout Rate	0.3
		Gamma	0.925
	Unnormalised	Number of Filters	88
		Spatial Dropout Rate	0.0
		Gamma	0.970
RNN (Lund)		Number of Layers	2
		Number of Units	15
		Gamma	0.935
DNN (Global)		Number of Layers	6
		Number of Units	116
		Dropout Rate	0.1
		Gamma	0.93

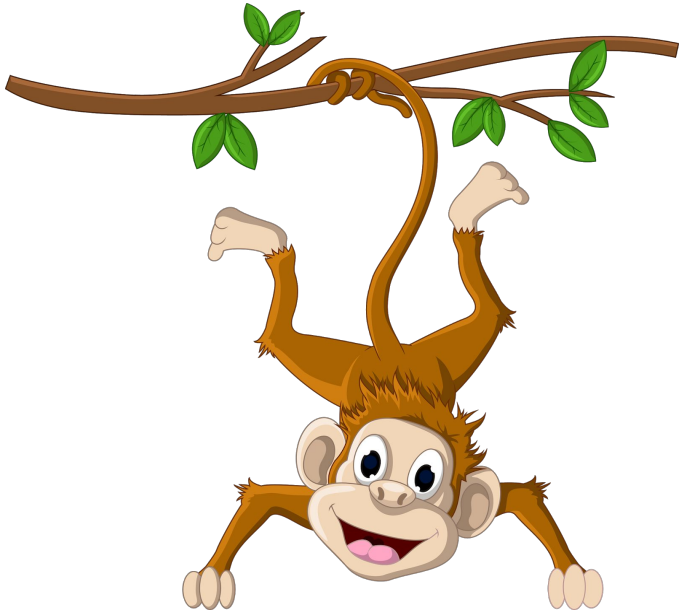


JHEP11 (2021) 219

Searching for new phenomena
supervised learning

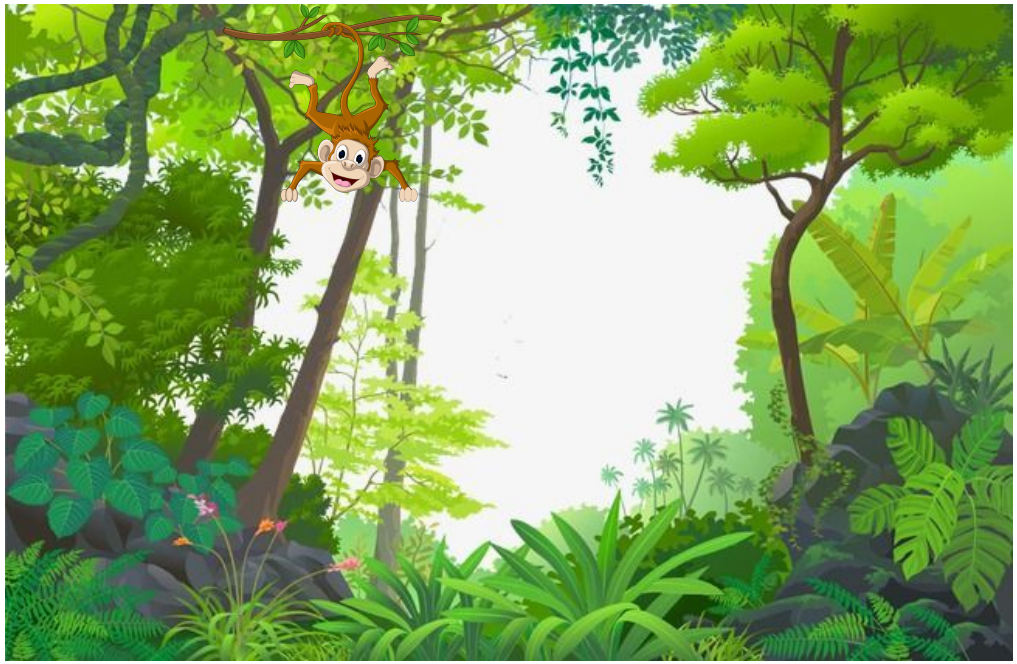
background

signal



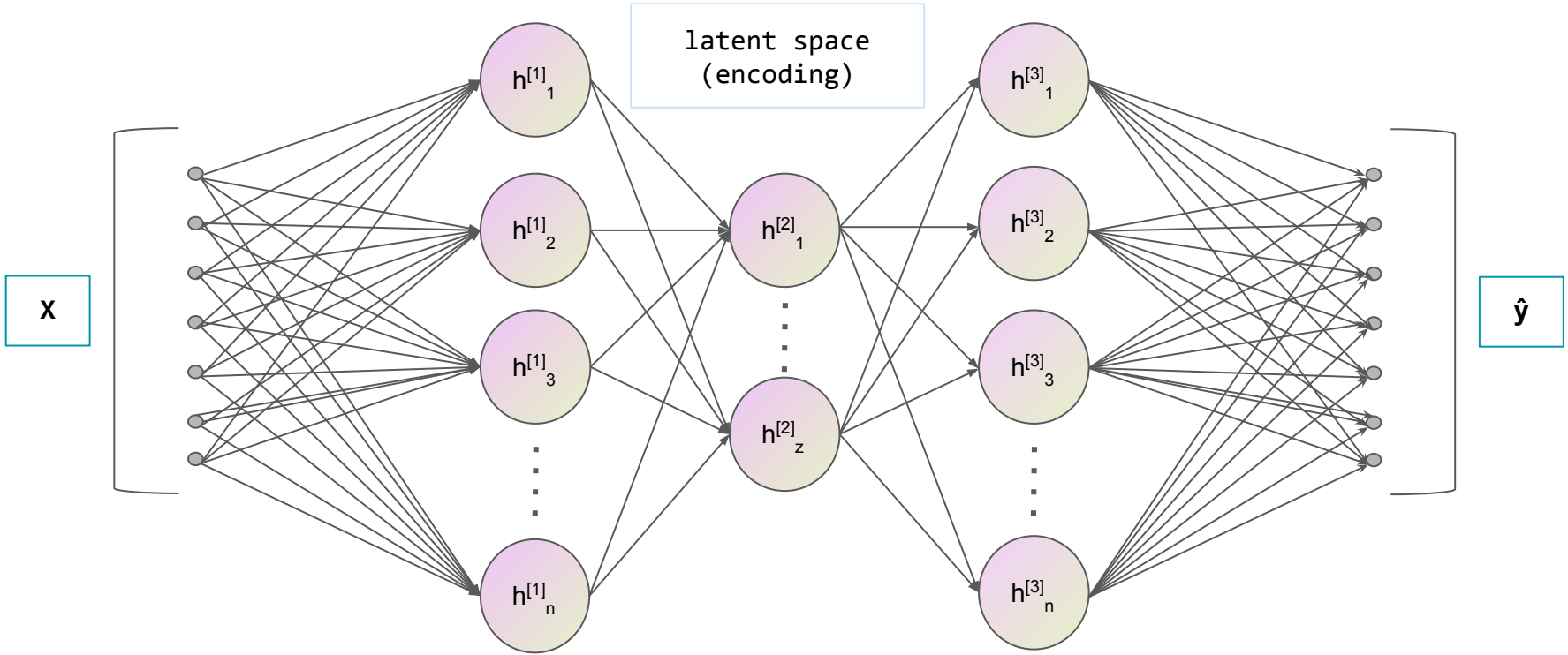
Searching for new phenomena
supervised learning

Searching for new phenomena
supervised learning

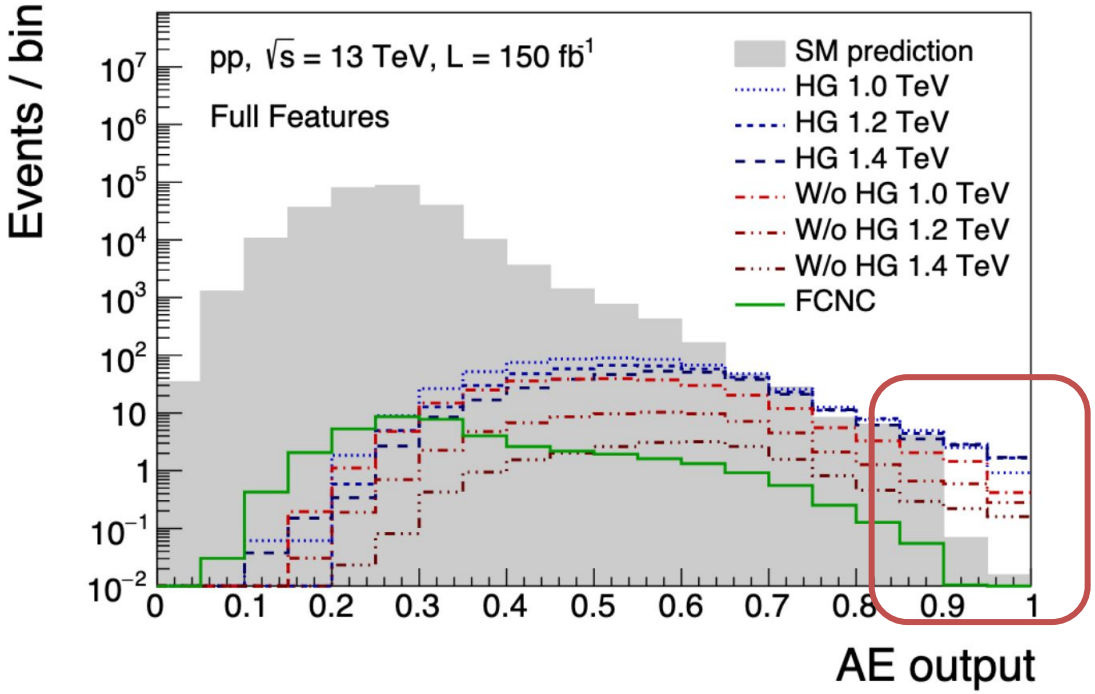


Searching for new phenomena
what if the signal is different from expected?

Searching for new phenomena
autoencoders



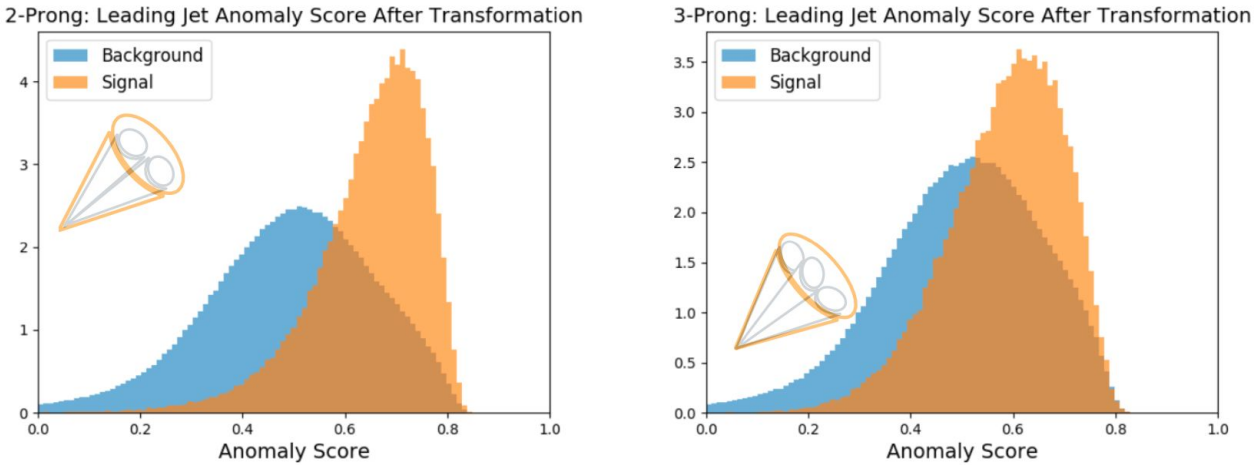
Searching for new phenomena
anomaly detection



Eur.Phys.J.C 81 (2021) 1, 27

Searching for new phenomena
anomaly detection

Detecting “anomalous” jets via sequence modeling

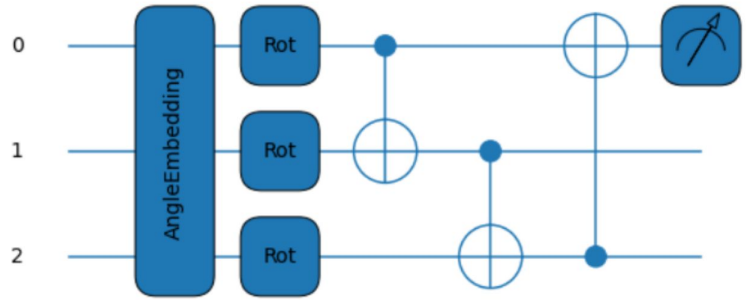


Using a variational recurrent neural network to assign an “anomaly score” to each jet

JINST 16 (2021) P08012

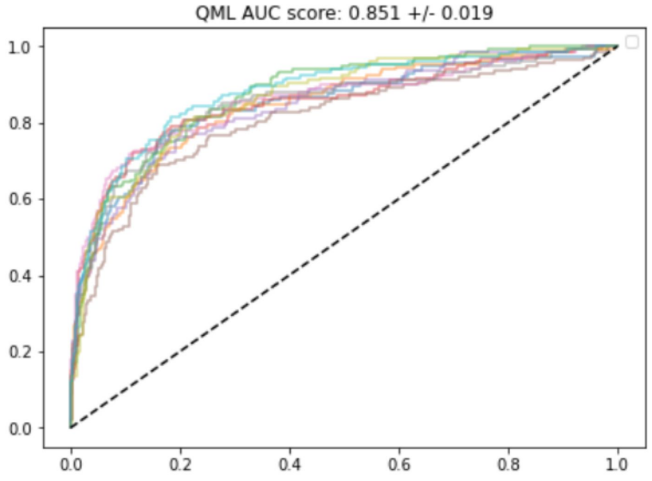
Quantum computing in particle physics: QML

Variational Quantum Classifier



$$|\psi'_x\rangle = U(\theta)\,|\psi_x\rangle$$

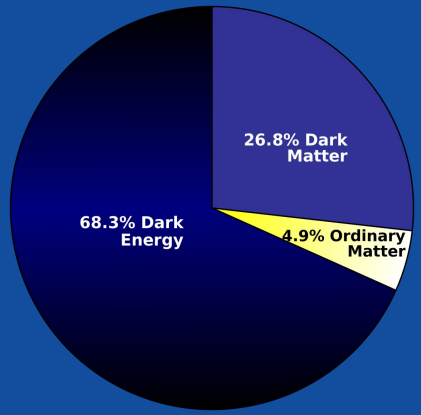
parameterized by a set of learnable parameters

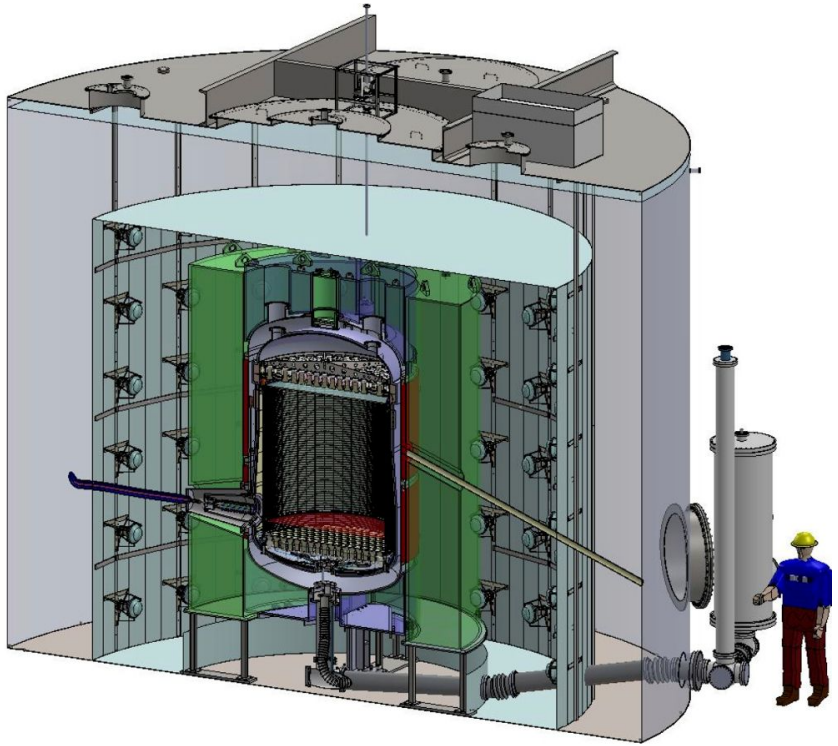


(very preliminary studies)

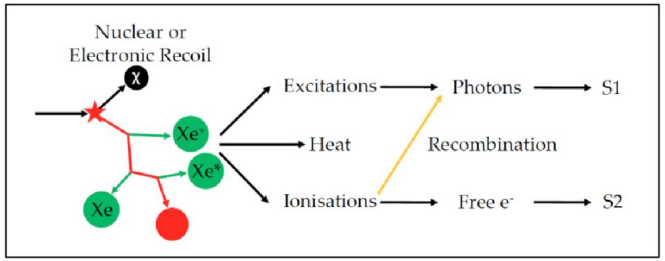
Underground experiments

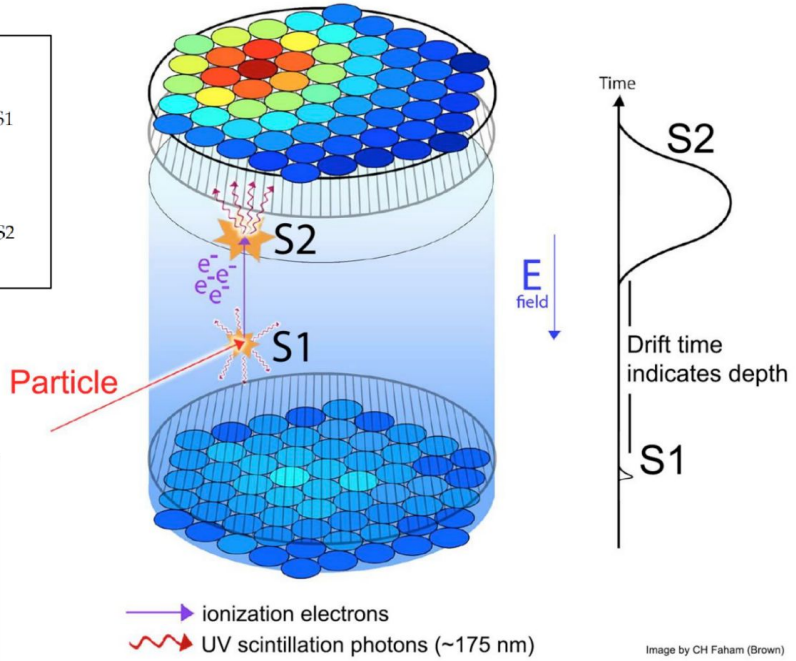
- Searching for the elusive dark matter





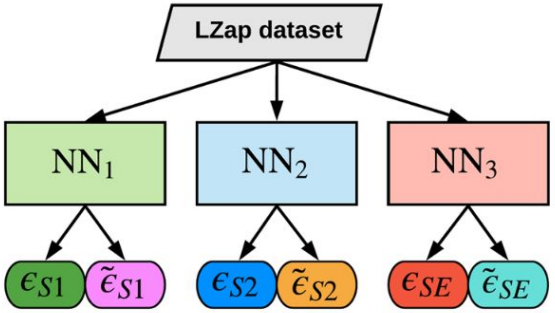
Underground experiments
search for very rare events





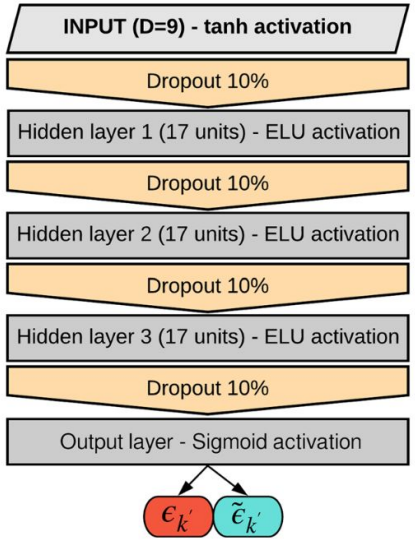
- Single photon and electron sensitivity
- Z position from S1-S2 timing
- X-Y position from S2 signal pattern
- ER/NR discrimination by charge to light ratio (S2/S1)

Underground experiments
search for very rare events



- pulse classification

Name [unit]	Type	Description
pA [phd]	Float	Total integrated area from the start to the end of the pulse
pH [phd/ns]	Float	Pulse maximum amplitude
pHTL	Float	Fraction of pulse length time at which the pulse reaches maximum amplitude
pL90 [ns]	Int	Pulse length time at 90% area, from 5 to 95% integrated area time
pRMSW [ns]	Int	Pulse root mean square (RMS) width
pF50	Float	Fraction of the pulse area integrated in a 50 ns time window starting 10 ns before the 5% integrated area time
pF100	Float	Same as *pF50* but for a 100 ns integration window
pF200	Float	Same as *pF50* but for a 200 ns integration window
pF1k	Float	Same as *pF50* but for a 1 μs integration window
TBA	Float	Top-bottom asymmetry: difference between the top PMT area fraction and bottom PMT area fraction
Coincidence	Int	Number of PMT channels that record signal within pulse boundaries



Eur. Phys. J. C (2022) 82:553

Underground experiments
search for very rare events

GMM class	TriNet predicted class				Total	
	S1	S2	SE	Other		
S1	11,571	0	0	280	11,851	5.9%
S2	0	51,001	444	10	51,455	25.7%
SE	0	380	128,211	8	128,599	64.4%
Other	698	38	28	7331	8095	4.0%
Total	12,269	51,419	128,683	7629	200,000	

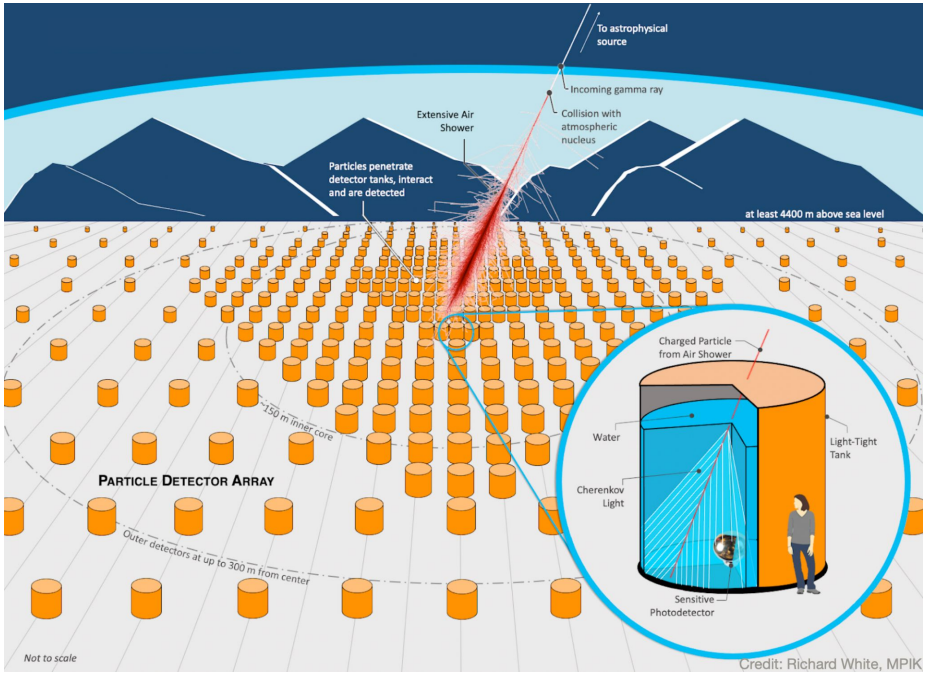
Eur. Phys. J. C (2022) 82:553

Astroparticle physics

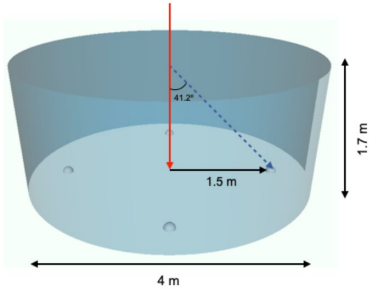
- Detecting extensive air showers produced by gamma rays
- Southern Wide-field Gamma-ray Observatory

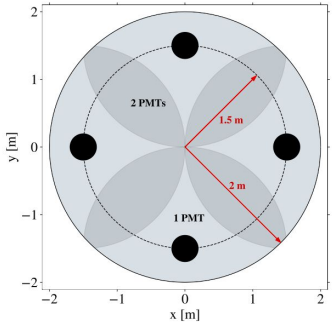
SWGO

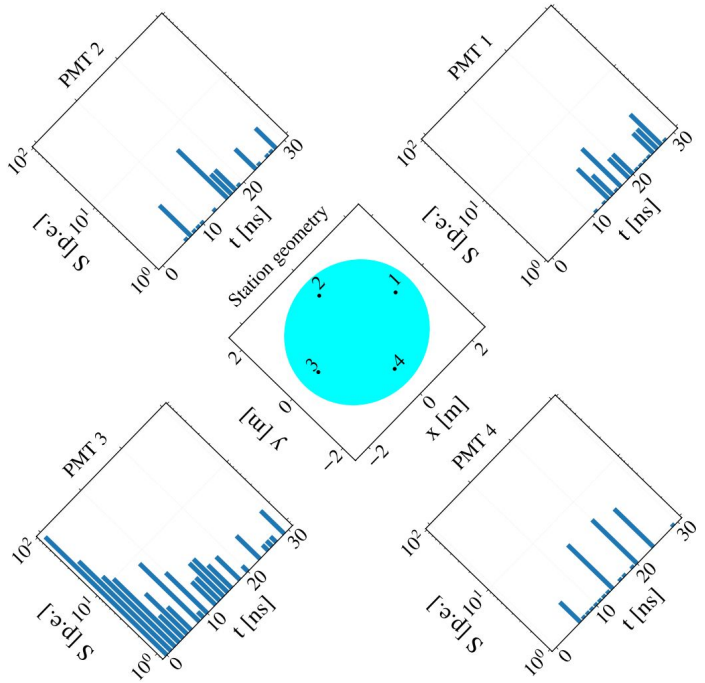
The Southern Wide-field Gamma-ray Observatory



Gamma/hadron discrimination with ML







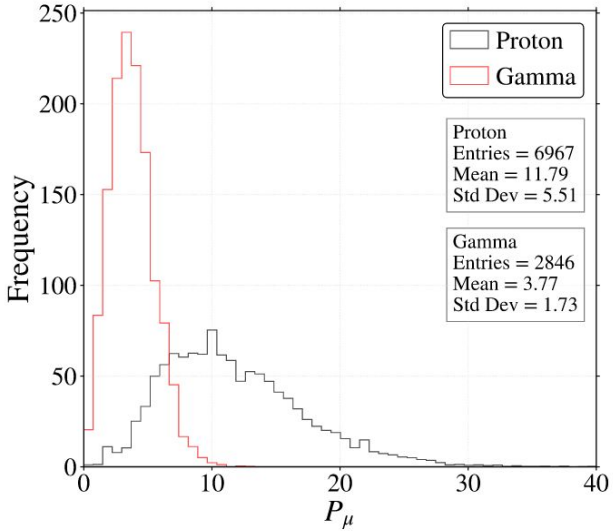
Top view of the WCD station and the signal detected in each PMT

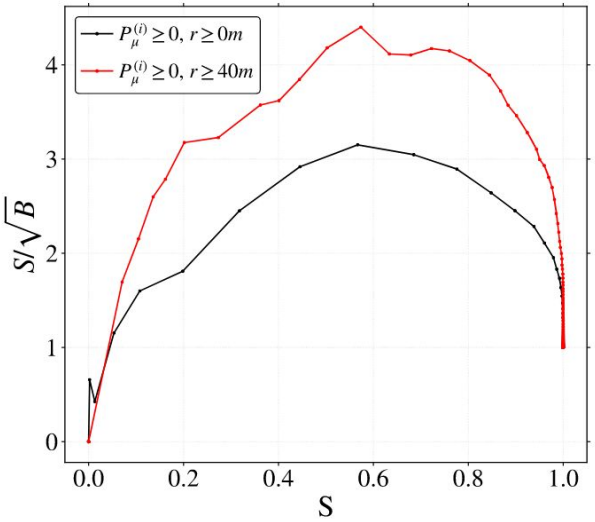
(Single 2 GeV muon injected)

Eur. Phys. J. C (2021) 81:542

Gamma/proton discrimination with ML

- 1-D convolutional neural network





Eur. Phys. J. C (2021) 81:542

Social Physics and Complexity

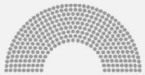
- SPAC uses large scale computational tools to study societal challenges, especially in disease forecasting, human behavior and public policy

Social Physics & Complexity

erc

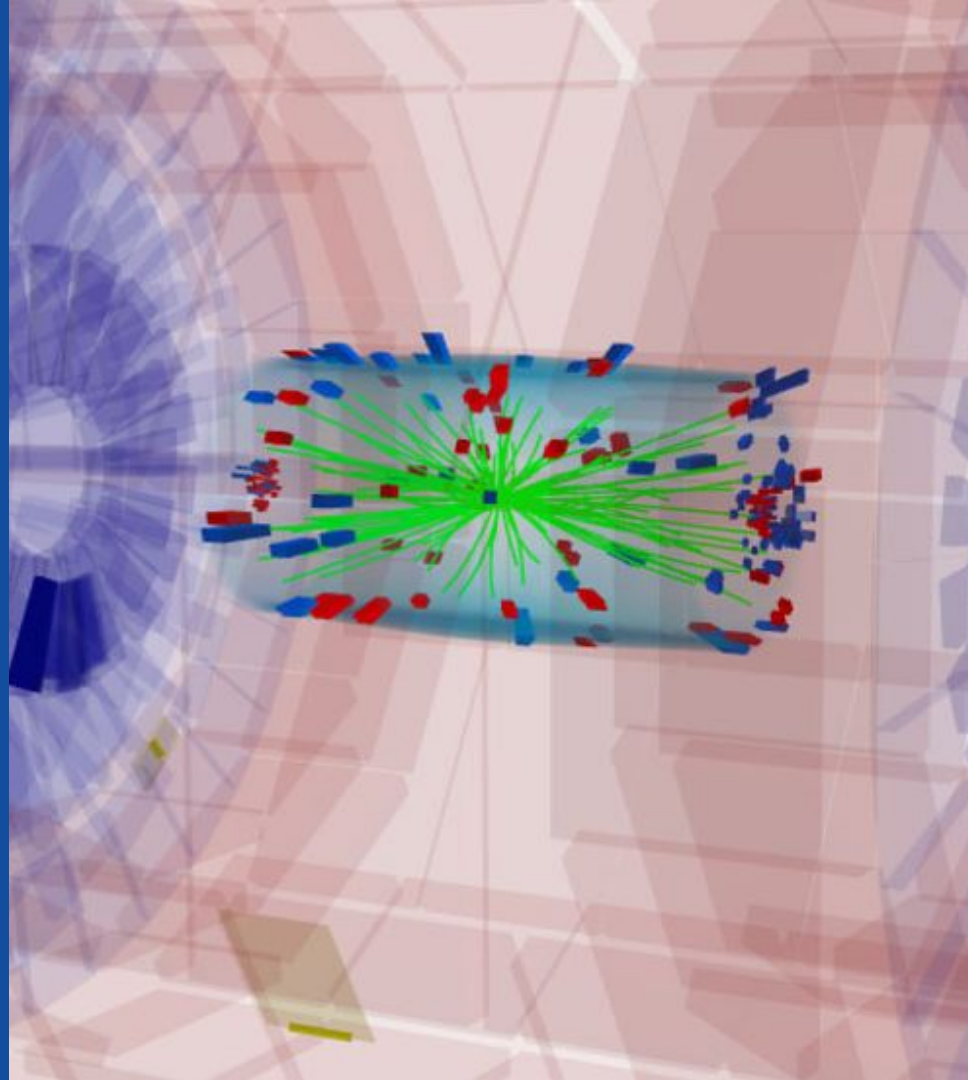
Social Physics and Complexity

Social Physics and Complexity

		QUESTIONS	DATA	TOOLS
HEALTH		Online vs. Offline Patterns Emergency Now-casting Antibiotic Over-prescription	Google Trends SNS24 Twitter ER waiting times E-prescriptions Weather	**Math Modelling** **ML** **Epidemiology**
POLICY		Political Decisions Gender Differences Agenda Setting Voting vs. Discourse	Media records Twitter Parliament data Surveys	**NLP** **Networks** **Math Modelling** **Complex Systems**
BEHAVIOUR		Cognitive Biases Attitudes Towards Science Fake News Sharing	Large scale surveys Behavioral experiments Twitter Economic databases...	**Networks** **Math Modelling** **Psychology** **Information**

summary

- at LIP we have more than 30 years of expertise in the analysis of large and complex data
- the most suitable technique has to be chosen for each problem
  - uncertainties and imperfect datasets
- synergies with other fields and activities



Thanks!

any questions?

you can also find me at nfcastro@lip.pt