Probing Quark-Gluon Plasma at the Large Hadron Collider in the Charmonia sector using Machine Learning

Raghunath Sahoo

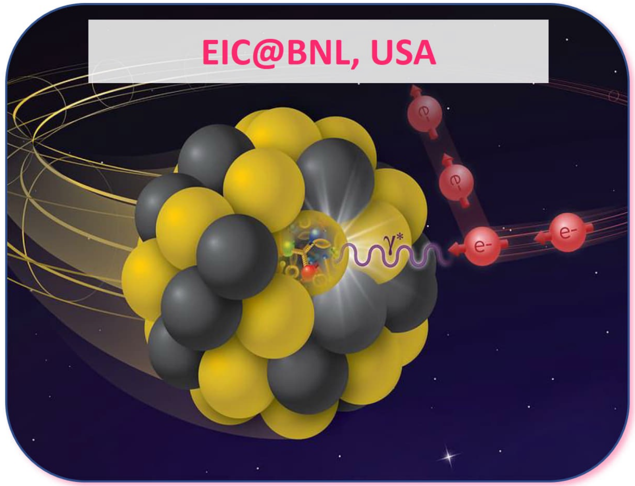
Indian Institute of Technology Indore, India

CBM Juniors' Meeting

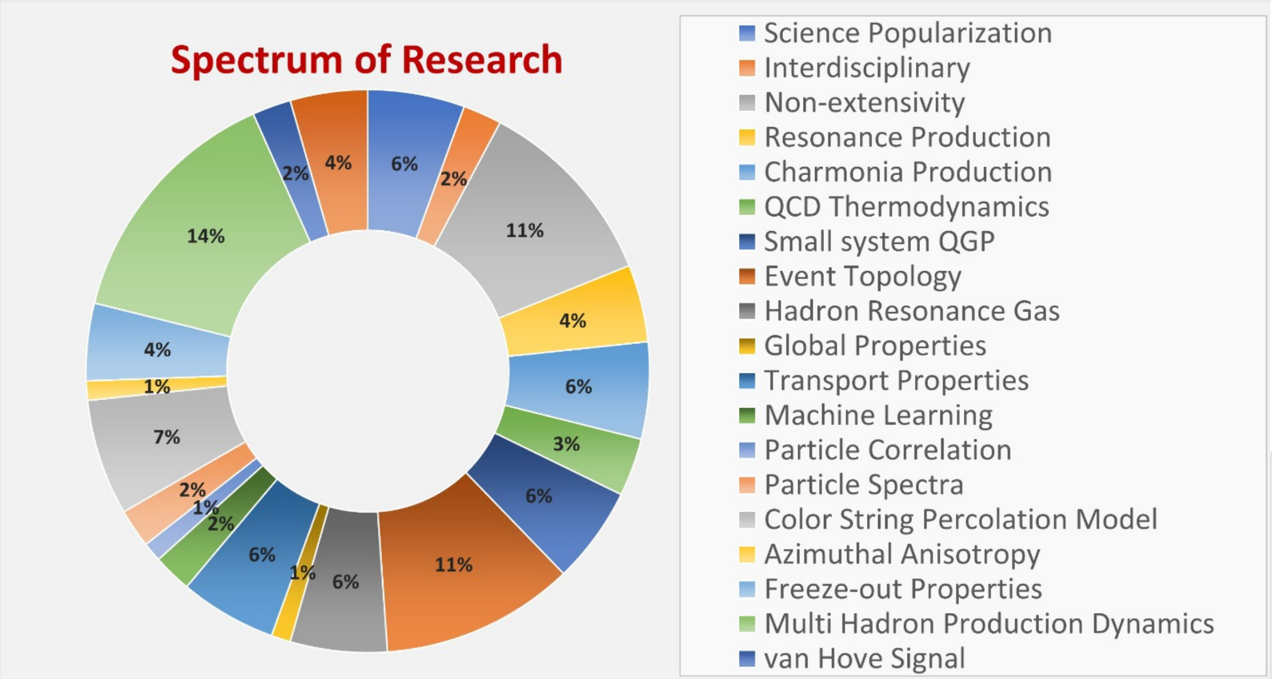
Raghunath.Sahoo@cern.ch, raghunath@iiti.ac.in

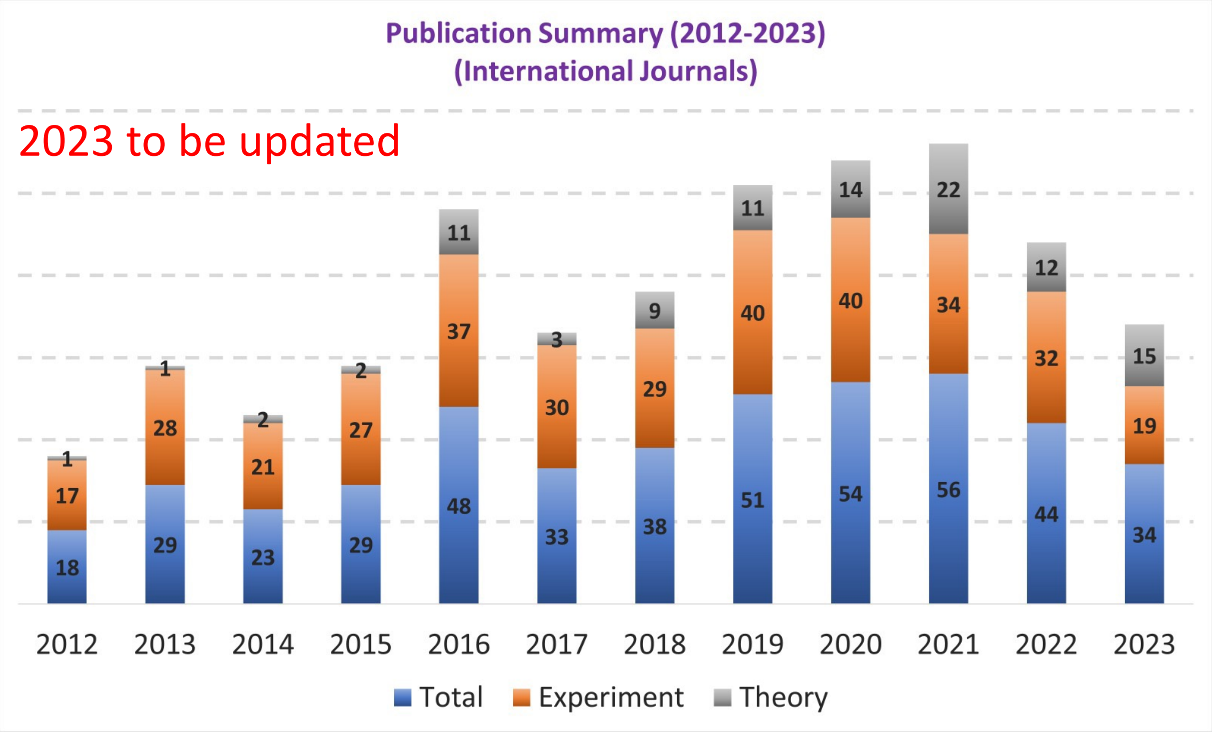
Experimental High-Energy Physics Group

Contact: Prof. Raghunath Sahoo, FInstP
e-mail: raghunath@iiti.ac.in
Web: www.iiti.ac.in/~raghunath



Research Spectrum





Details: www.iiti.ac.in/~raghunath

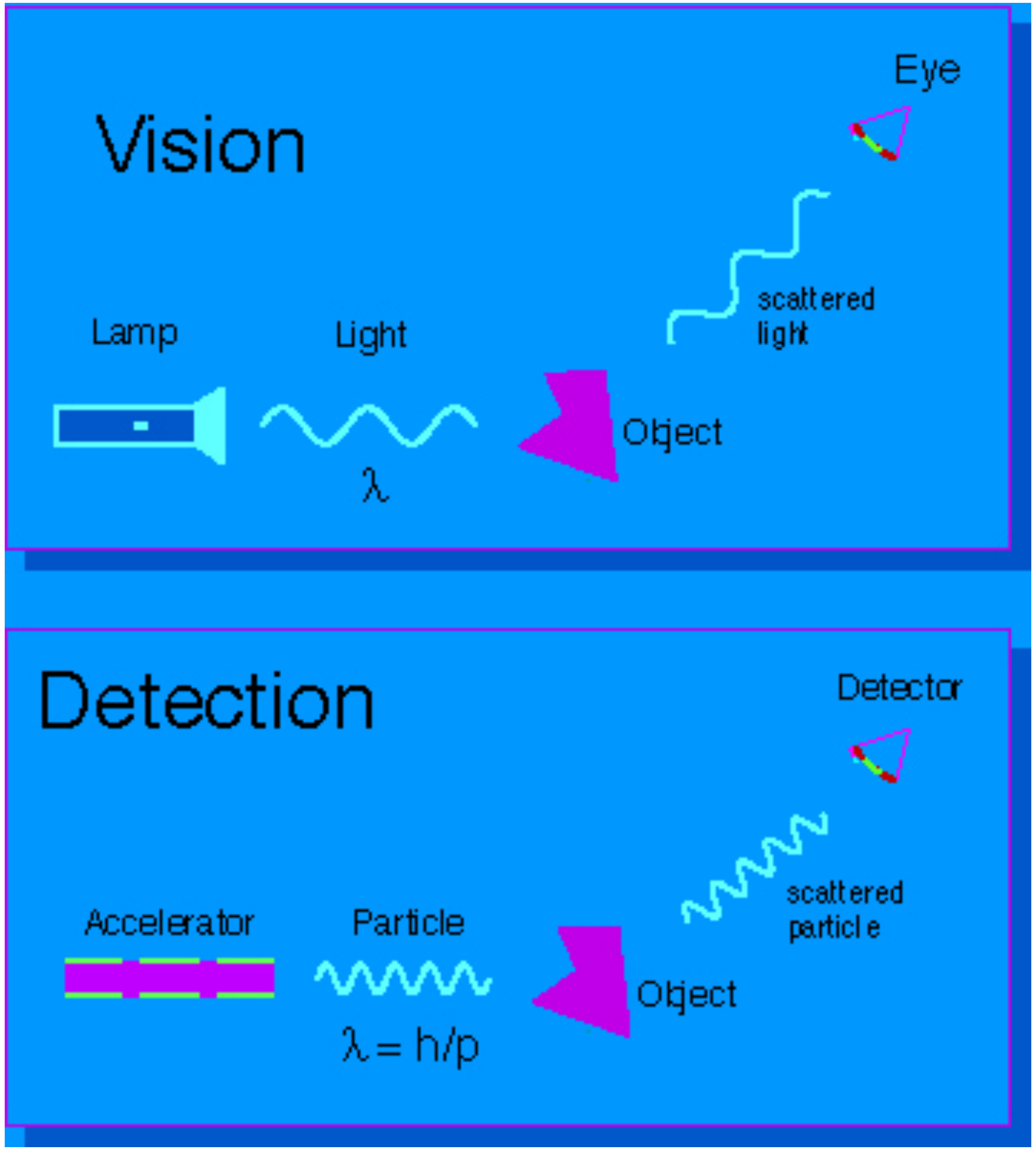
The Right Light to Look Inside

Vision works by scattering of 'visible' light

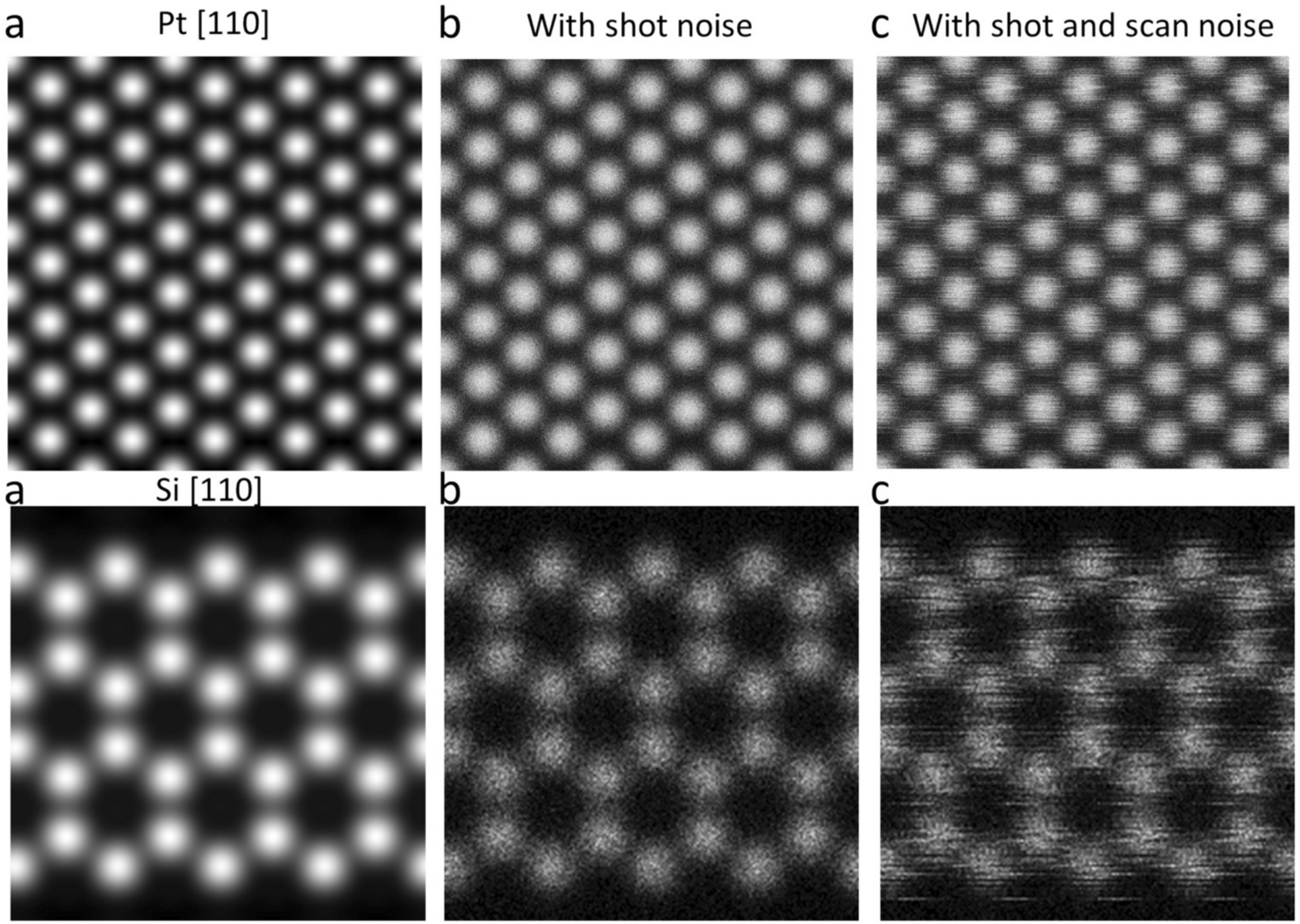
$\lambda = 400\text{-}700$ nm

"Vision" of even smaller structures via scattering of particles

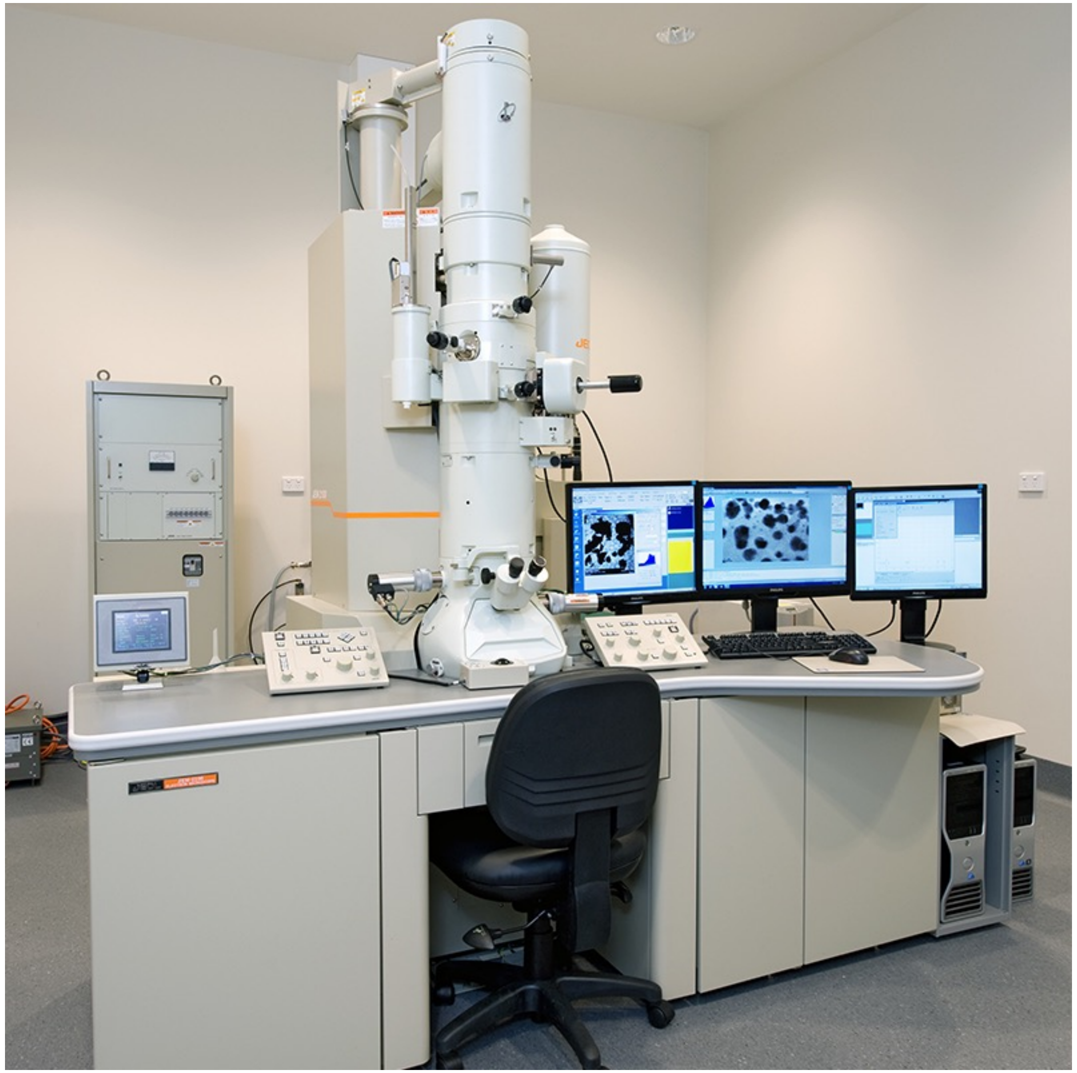
$\lambda = h/p$



Seeing the Atoms



(for demonstration purpose)



Transmission Electron Microscope

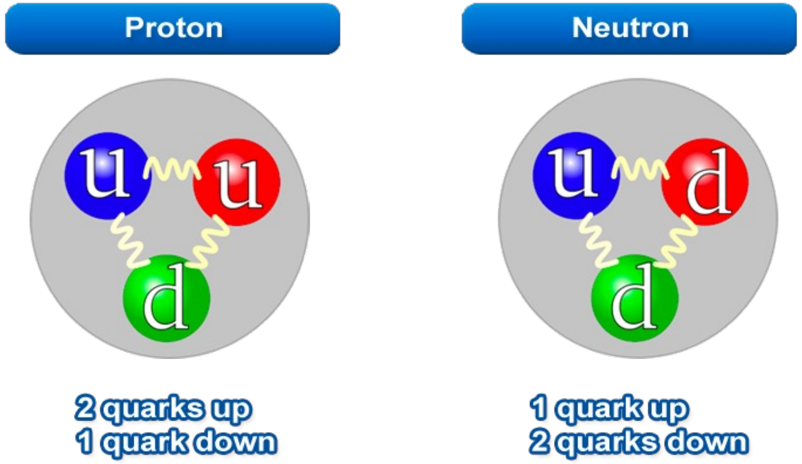
A Journey from Electrons to Quarks

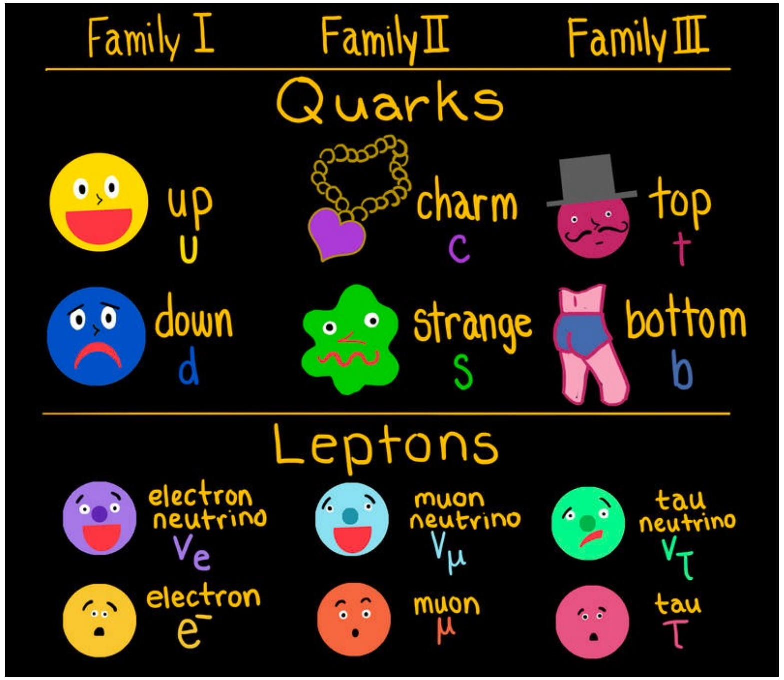
SLAC: Stanford Linear Accelerator Center (3.2 Kms)
Electron-positron energy ~ 50 GeV

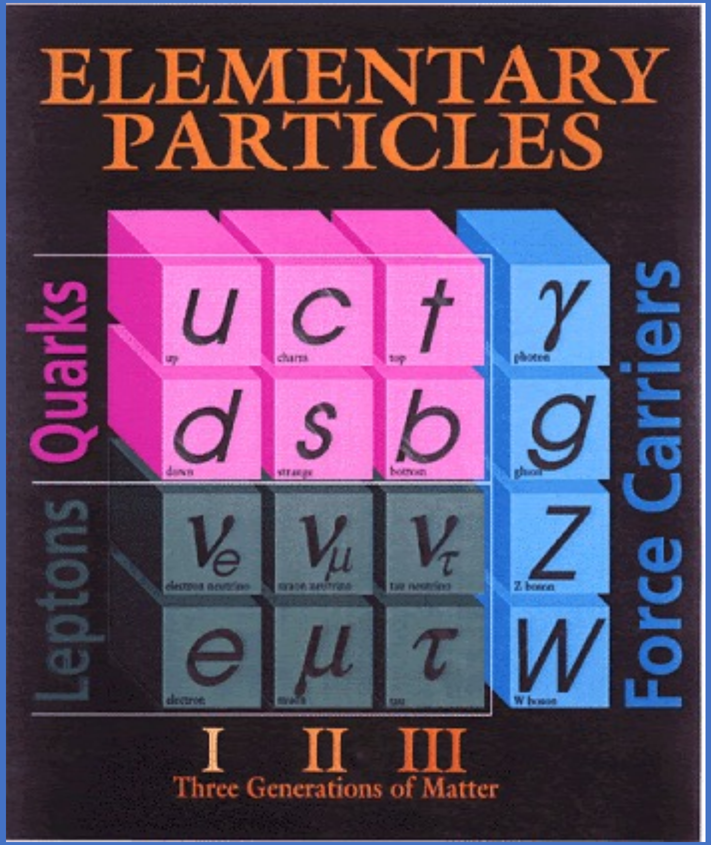
$$\lambda = h/p : \alpha\, 1/E$$

This universal formula of de Broglie helps in deciding the energy of the probe.

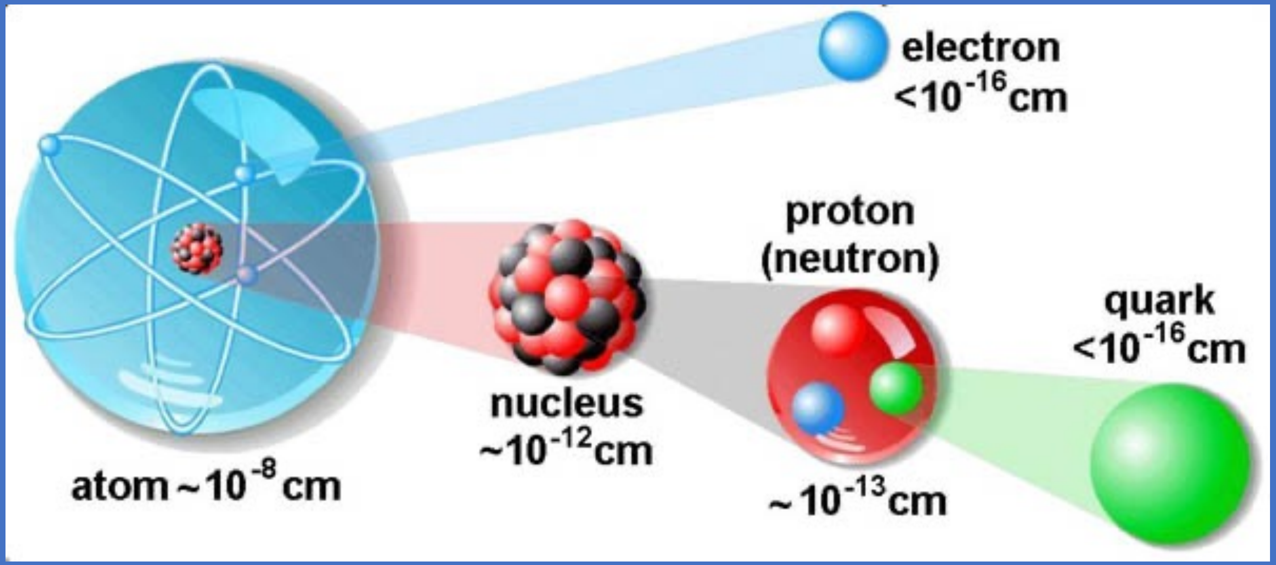
Proton charge radius ~ 0.843 fermi.
(1 fermi = 10^{-15} meter)





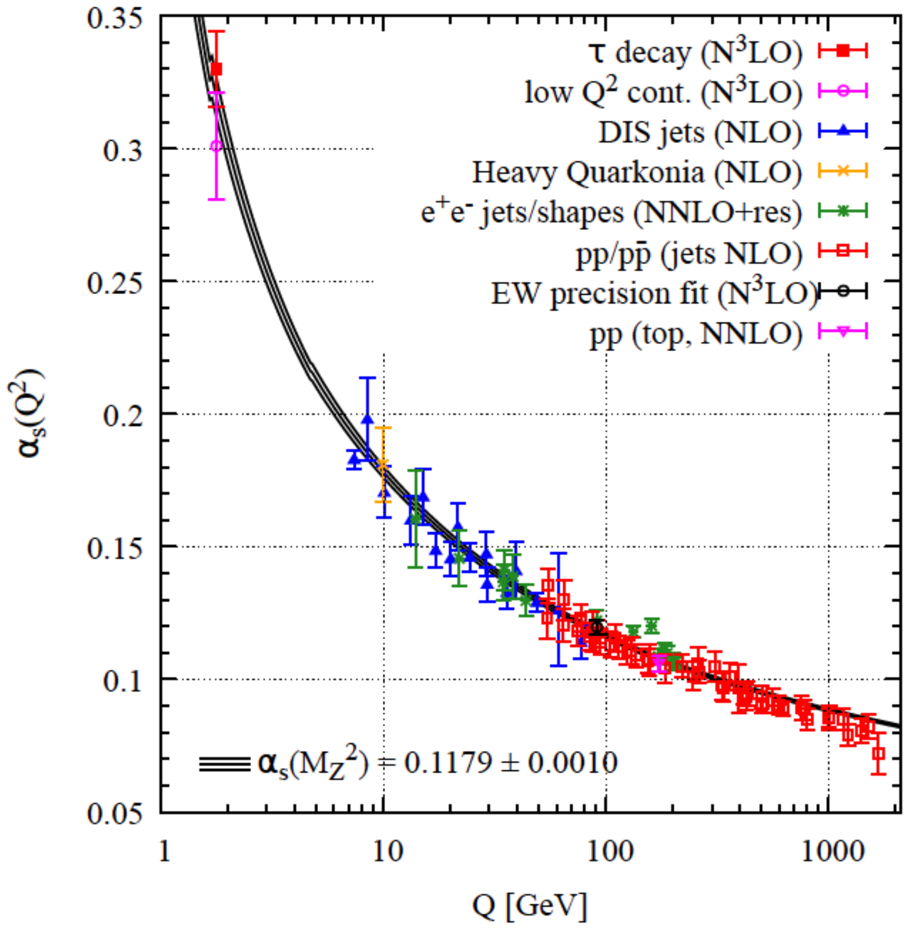


ELEMENTARY PARTICLES
Quarks
Leptons
Force Carriers
u c t γ
d s b g
ν_e ν_μ ν_τ Z
e μ τ W
I II III
Three Generations of Matter



electron
$<10^{-16}$cm
proton
(neutron)
quark
$<10^{-16}$cm
nucleus
$\sim10^{-12}$cm
atom $\sim10^{-8}$cm
$\sim10^{-13}$cm

Can we find free quarks?



To the leading order:

$$\alpha_s(Q^2) = \frac{4\pi}{(33-2n_f)ln(Q^2/\Lambda^2)}$$

n_f is the no. of quark flavors, with a mass less than Q/2,
Λ is the QCD scale parameter, obtained experimentally ~ 200 MeV

👉Coupling becomes weaker as the momentum transfer increases,

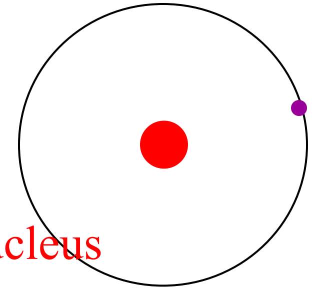
Or as we go smaller in the length-scale ➔ small distance, partons move freely (asymptotic freedom of QCD)

👉As the inter-quark distance increases, strong coupling grows faster making the quarks confined inside the cage of the hadrons → Quark confinement

Asymptotic freedom and infrared slavery is an inbuilt property of QCD (Strong interaction)→ No free quarks in nature!

Analogies and differences between QED and QCD

to study structure of an atom…



electron

…separate constituents

neutral atom

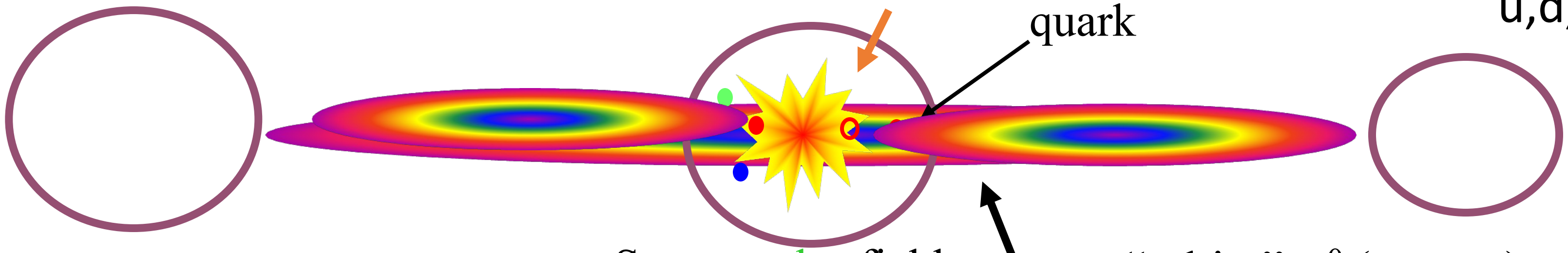
QED Quantum Electro Dynamics

Confinement: fundamental & crucial (but *not* understood!) feature of strong force
- colored objects (quarks) have ∞ energy in normal vacuum

•

quark-antiquark pair
created from vacuum

quarks
u,d, (s,c,t,b)



“white” proton (baryon)
(confined quarks)

Strong color field
Force *grows* with separation !!!
“white” proton

“white” π^0 (meson)
(confined quarks)

QCD: Quantum Chromo Dynamics

Can we find free quarks?

1973: Asymptotic freedom
D.J. Gross, F. Wilczek, H.D. Politzer

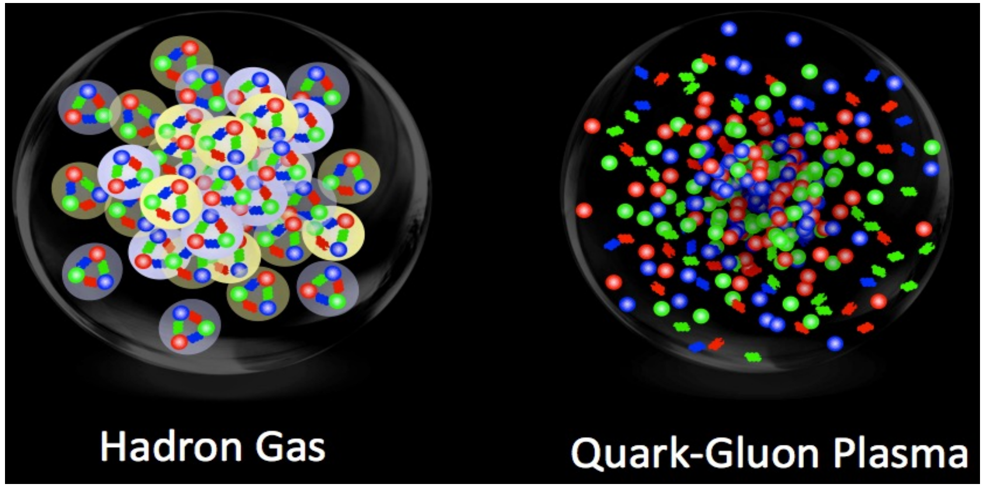
1975: Asymptotic QCD & deconfinement:
N. Cabibbo and G. Parisi;
J. Collins and M. Perry

2004: Nobel Prize

David J. Gross   H. David Politzer   Frank Wilczek

QCD predicts that normal nuclear matter undergoes a phase transition to quark-gluon plasma (QGP) under extreme temperatures and energy densities.

Quark Gluon Plasma

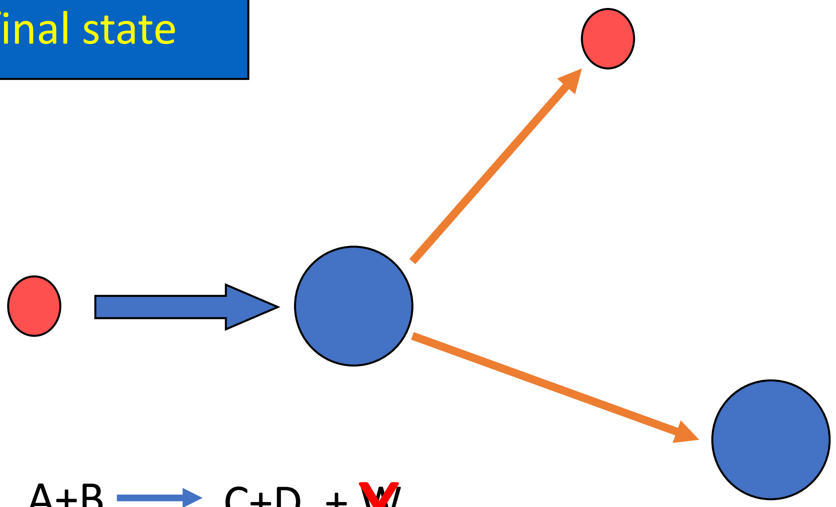


Quark Gluon Plasma (QGP): (locally) thermally equilibrated state of matter in which quarks and gluons are deconfined from hadrons, so that color degrees of freedom become manifest over nuclear, rather than merely nucleonic, volumes.

A journey to the beginning of the universe.....

Collisions: Classical vs Relativistic

Two-particle final state



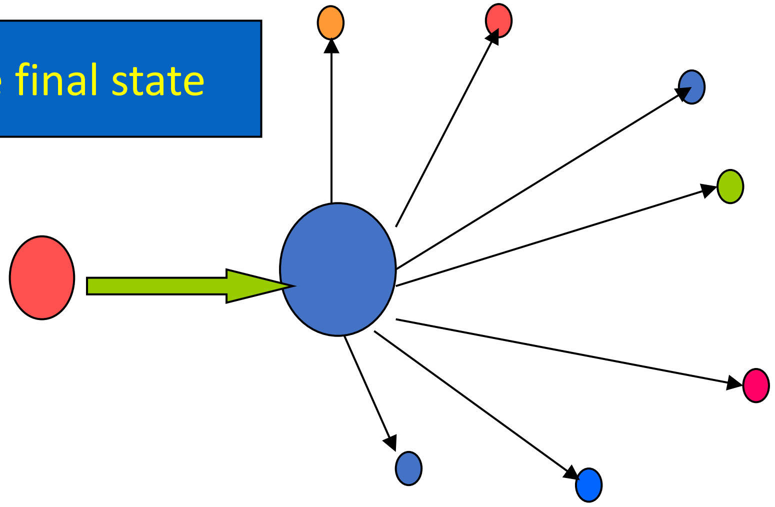
$A+B \longrightarrow C+D + \cancel{W}$

Classical collisions:

a) Mass is conserved: $m_A+m_B = m_C+m_D$
b) Momentum is conserved: $p_A+p_B = p_C+p_D$
c) Kinetic energy may(not) be conserved

i. Sticky- K.E. decreases: $T_A+T_B > T_C+T_D$
ii. Explosive- K.E. increases: $T_A+T_B < T_C+T_D$
iii. Elastic- K.E. is conserved: $T_A+T_B = T_C+T_D$

Multi-particle final state



Relativistic collisions:

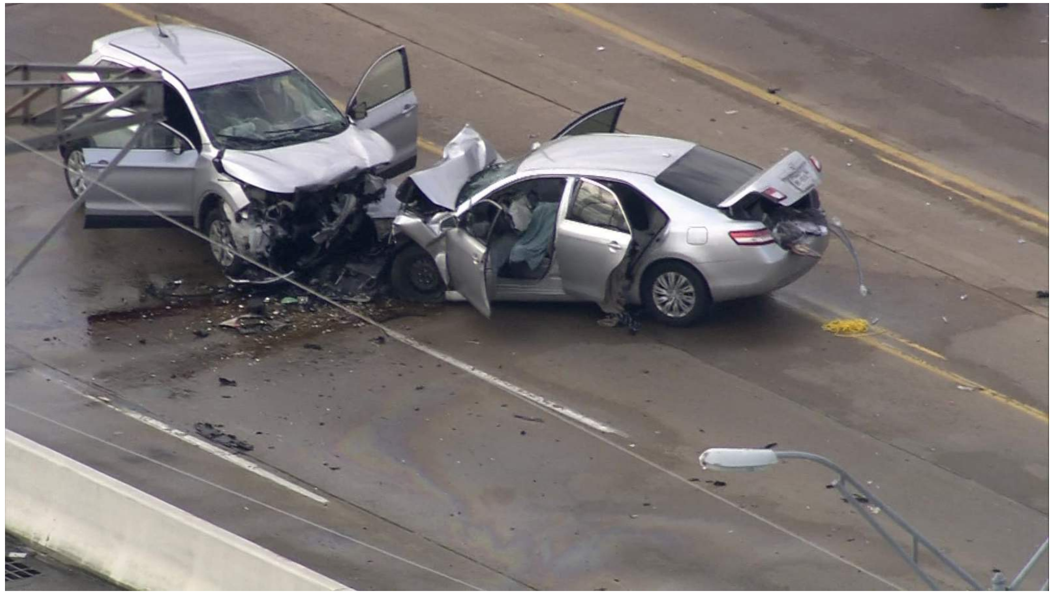
a) Energy is conserved: $E_A+E_B = E_C+E_D$
b) Momentum is conserved: $p_A+p_B = p_C+p_D$
c) Kinetic energy may(not) be conserved

Mass-energy conversions in "explosive collisions" !

If mass is conserved: collision is elastic.

Why Collider Experiments?

The severity of damage is less.



The damage is much higher !

Energy available for particle production:

In laboratory frame (fixed target):

$E_{cm} = \sqrt{(2m_t E_{beam})}$

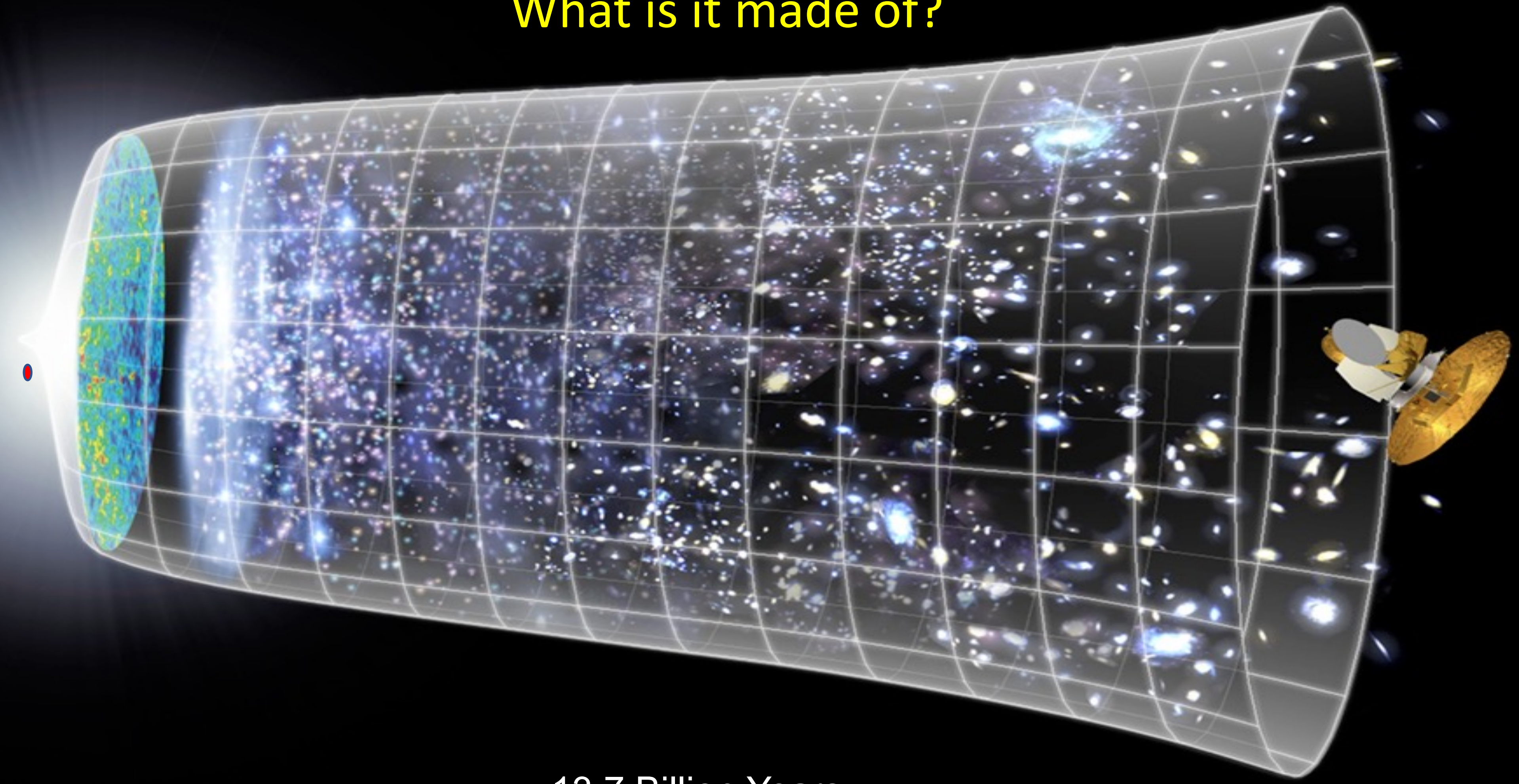
In CM frame (collider):

$E_{cm} = (2E_{beam})$

Our Universe ….. How did it start?

What is it made of?

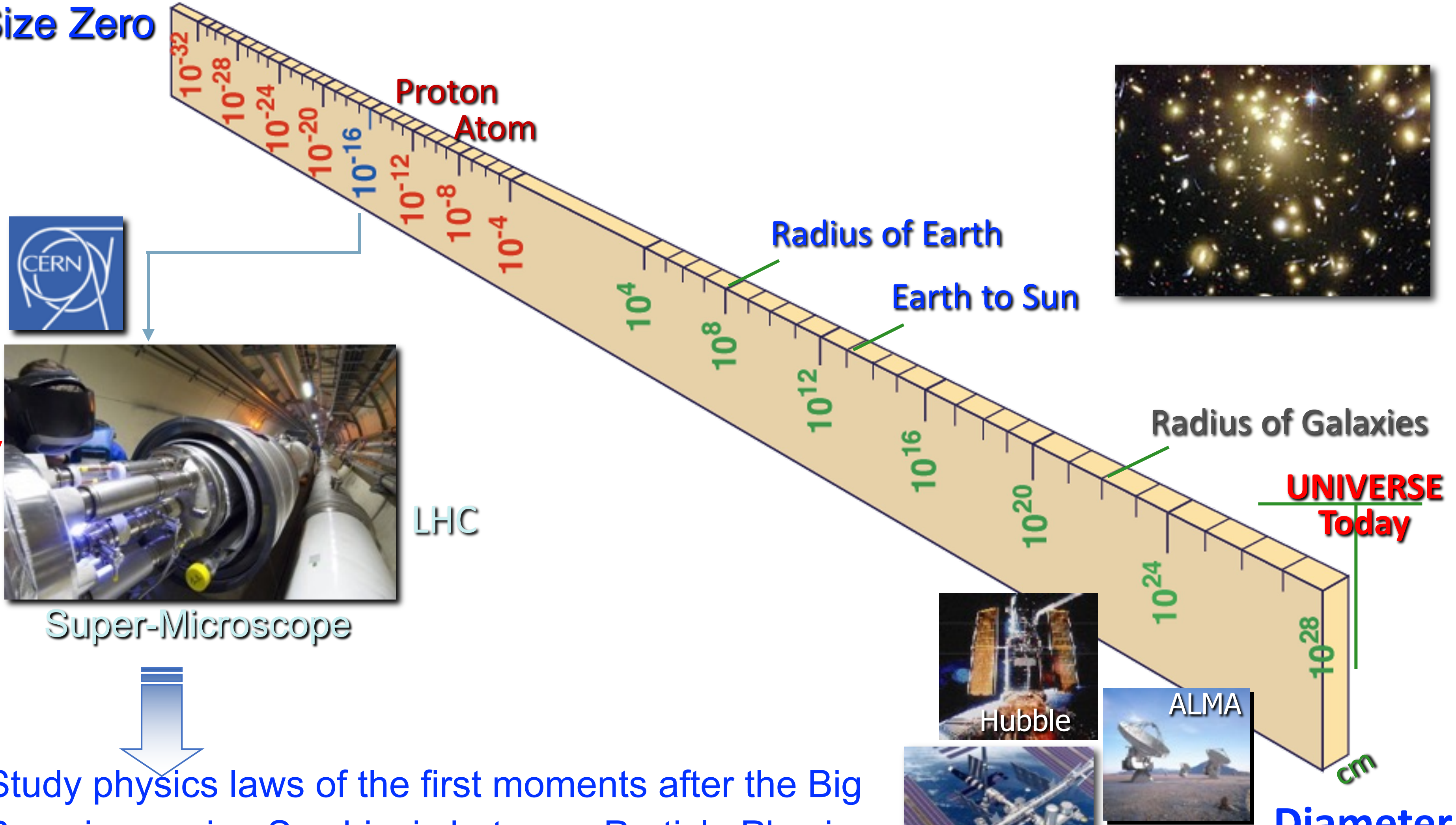
Big Bang



13.7 Billion Years

Today

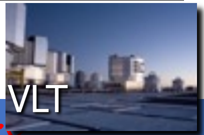
Accessing Big Bang conditions



High Energy Accelerator

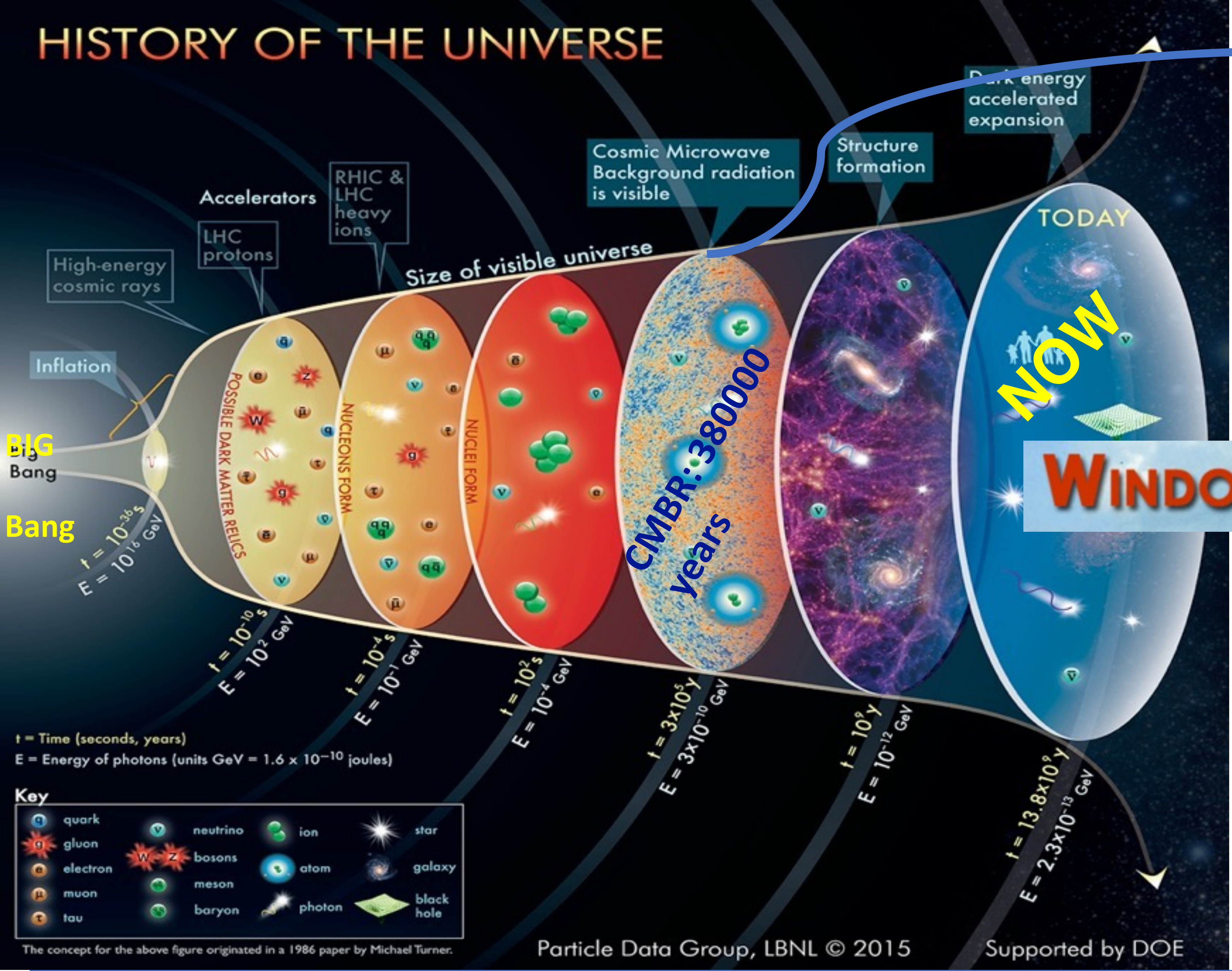
LHC

Super-Microscope



Study physics laws of the first moments after the Big Bang increasing Symbiosis between Particle Physics, Astrophysics and Cosmology

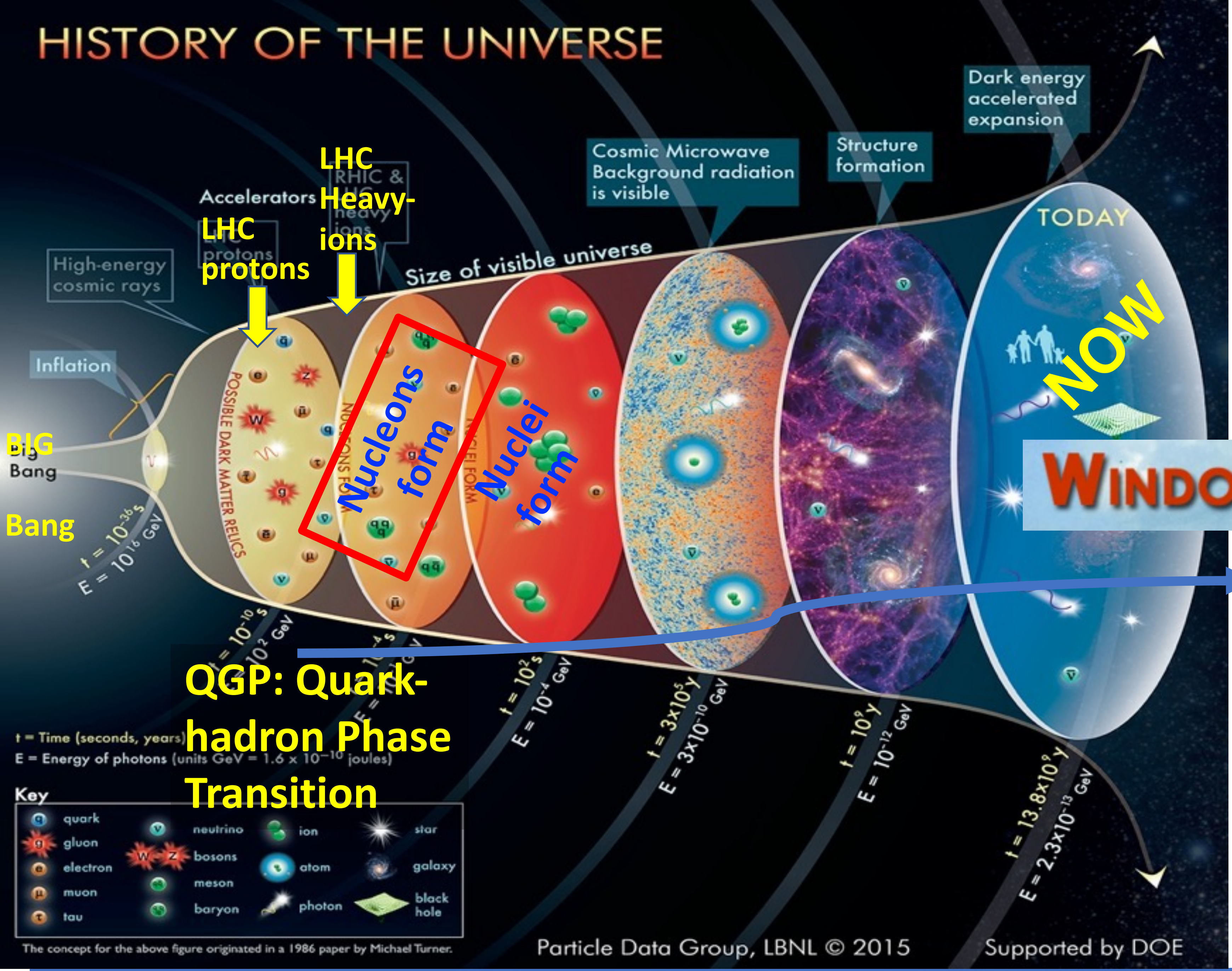
Diameter: 8.8×10^{26} m



Astrophysical Probes

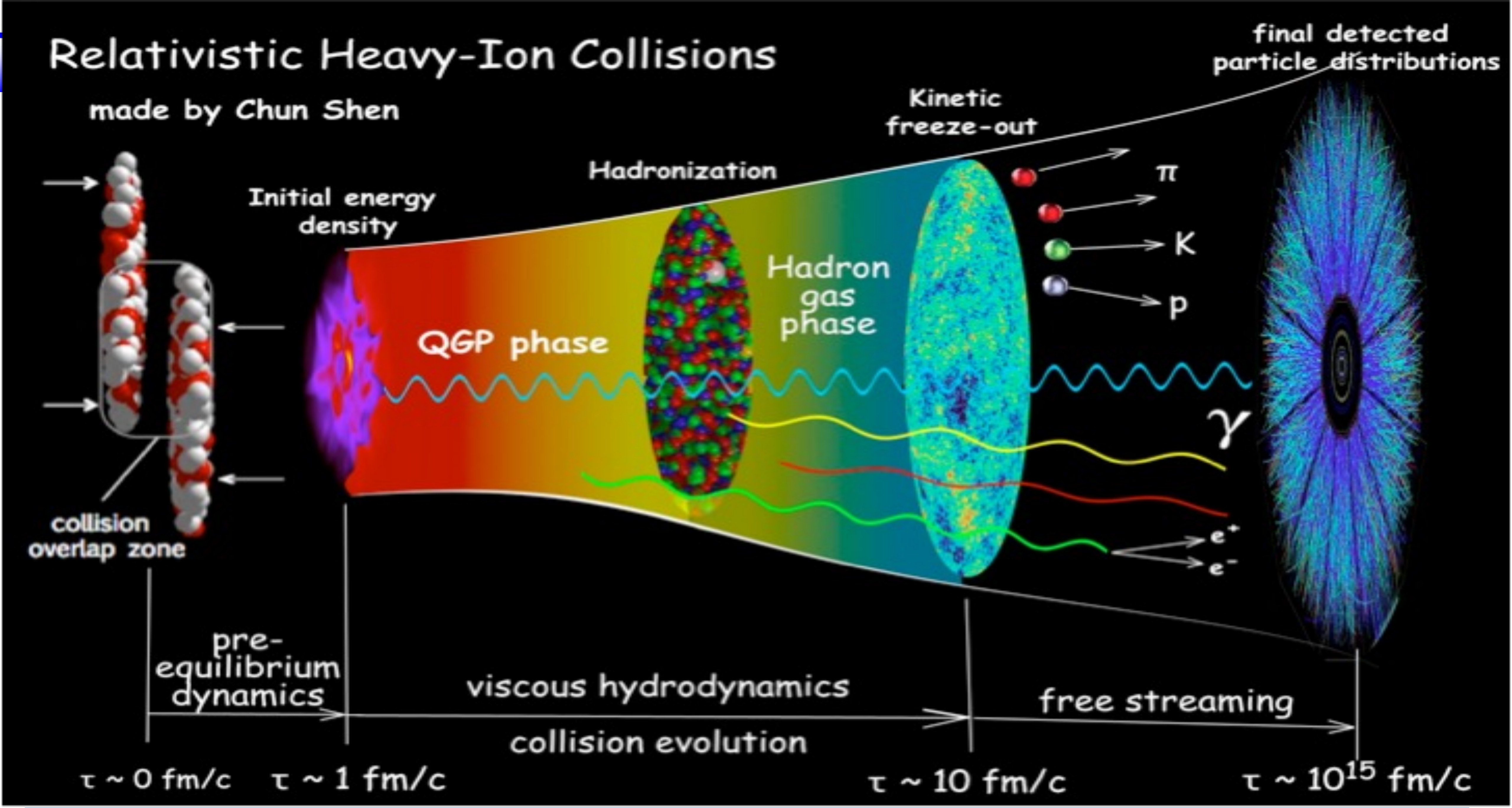
Takes us back to 380,000 years after the Big Bang

WINDOWS ON THE UNIVERSE



Accelerators (LHC)

Takes us back to within few Microseconds of the Big Bang


Relativistic Heavy-Ion Collisions
made by Chun Shen
final detected
particle distributions
Kinetic
freeze-out
Hadronization
Initial energy
density
Hadron
gas
phase
QGP phase
π
K
p
γ
e^+
e^-
collision
overlap zone
pre-
equilibrium
dynamics
viscous hydrodynamics
free streaming
collision evolution
τ ~ 0 fm/c
τ ~ 1 fm/c
τ ~ 10 fm/c
τ ~ 10^{15} fm/c

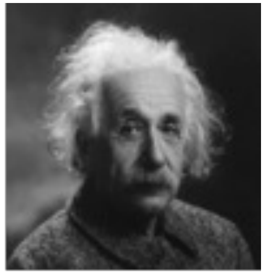
How much high is high-energy?

High energies allow us:

To look deeper into Nature (E α 1/size),
("powerful microscopes")

de Broglie

To discover new particles with high(er) mass ($E = mc^2$)



Einstein

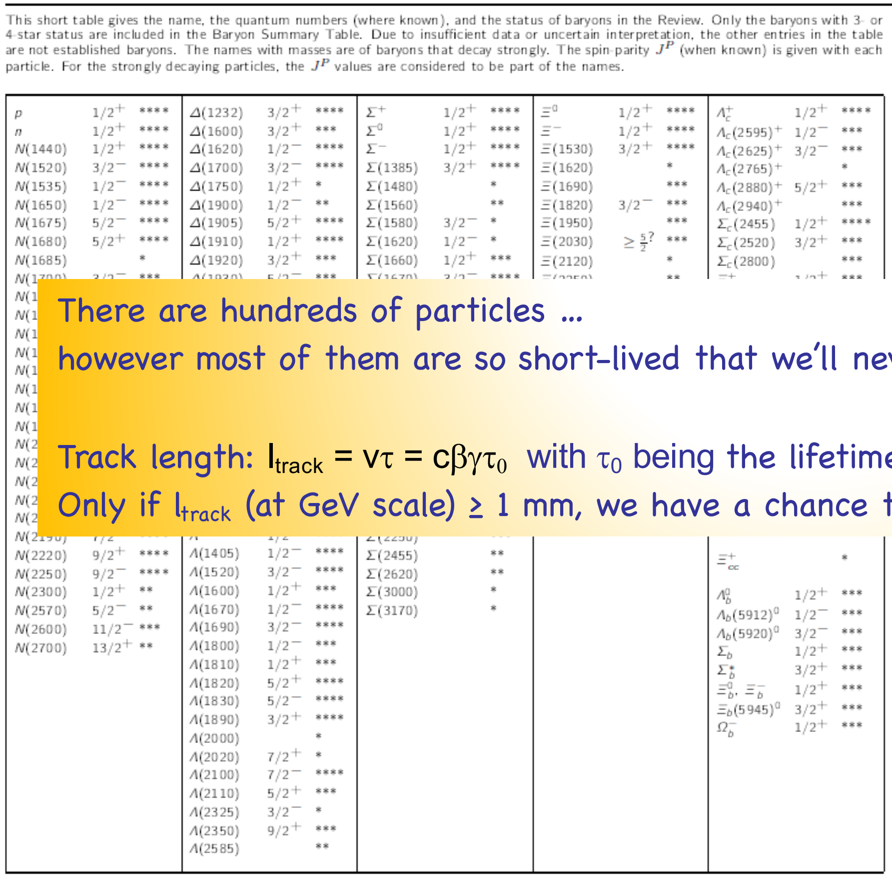
To study the early universe ($E= kT$)

Boltzmann

Lesson: To probe the subatomic universe, we need very high energies, which could produce many particles in the final state with very high temperature system.

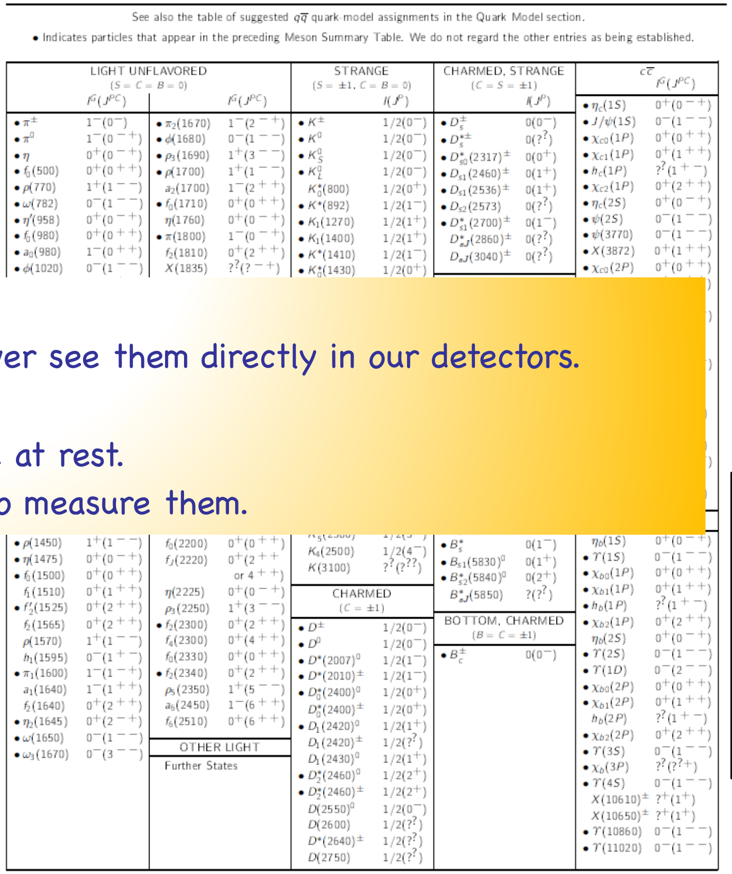
How do we do that?

Which particles do we detect?



Baryon Summary Table

This short table gives the name, the quantum numbers (where known), and the status of baryons in the Review. Only the baryons with 3- or 4-star status are included in the Baryon Summary Table. Due to insufficient data or uncertain interpretation, the other entries in the table are not established baryons. The names with masses are of baryons that decay strongly. The spin-parity J^P (when known) is given with each particle. For the strongly decaying particles, the J^P values are considered to be part of the names.



Meson Summary Table

See also the table of suggested $q\bar{q}$ quark model assignments in the Quark Model section.

• Indicates particles that appear in the preceding Meson Summary Table. We do not regard the other entries as being established.

Leptons

- e
- μ
- τ
- ν_e
- ν_μ
- ν_τ

There are hundreds of particles …

however most of them are so short-lived that we'll never see them directly in our detectors.

Track length: $l_{track} = v\tau = c\beta\gamma\tau_0$ with τ_0 being the lifetime at rest.

Only if l_{track} (at GeV scale) ≥ 1 mm, we have a chance to measure them.

Gauge bosons

- γ
- $W^{+/-}$
- Z
- g
- H

Which are left then? These 8 particles (and their antiparticles).

	γ	p	n	$e^\pm$	$\mu^\pm$	$\pi^\pm$	$K^\pm$	K_0 (K_S/K_L)
τ_0	∞	∞	∞	∞	2.2μs	26 ns	12 ns	89 ps / 51 ns
$l_{track\ (p=1GeV)}$	∞	∞	∞	∞	6.1 km	5.5 m	6.4 m	5 cm / 27.5 m

Accelerator facilities for heavy-ions

USA
1984

Bevalac

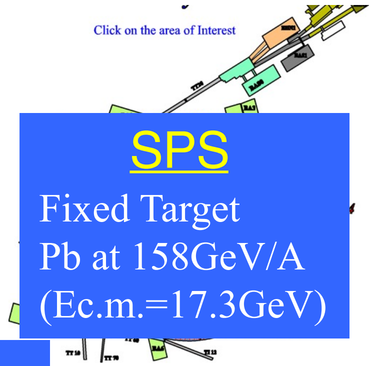
Bevalac
Fixed Target
Au at 1GeV/A

USA
1991

AGS Layout

AGS
Fixed Target Au at 11.7 GeV/A
(Ec.m.=4.86GeV)

CERN
1993 -



SPS
Fixed Target
Pb at 158GeV/A
(Ec.m.=17.3GeV)

SPS Layout

1988:
WA98
NA49
CERES

USA
2001 -

RHIC
Collider
Au+Au @ 200GeV/A

2001:
STAR
PHENIX
PHOBOS
BRAHMS

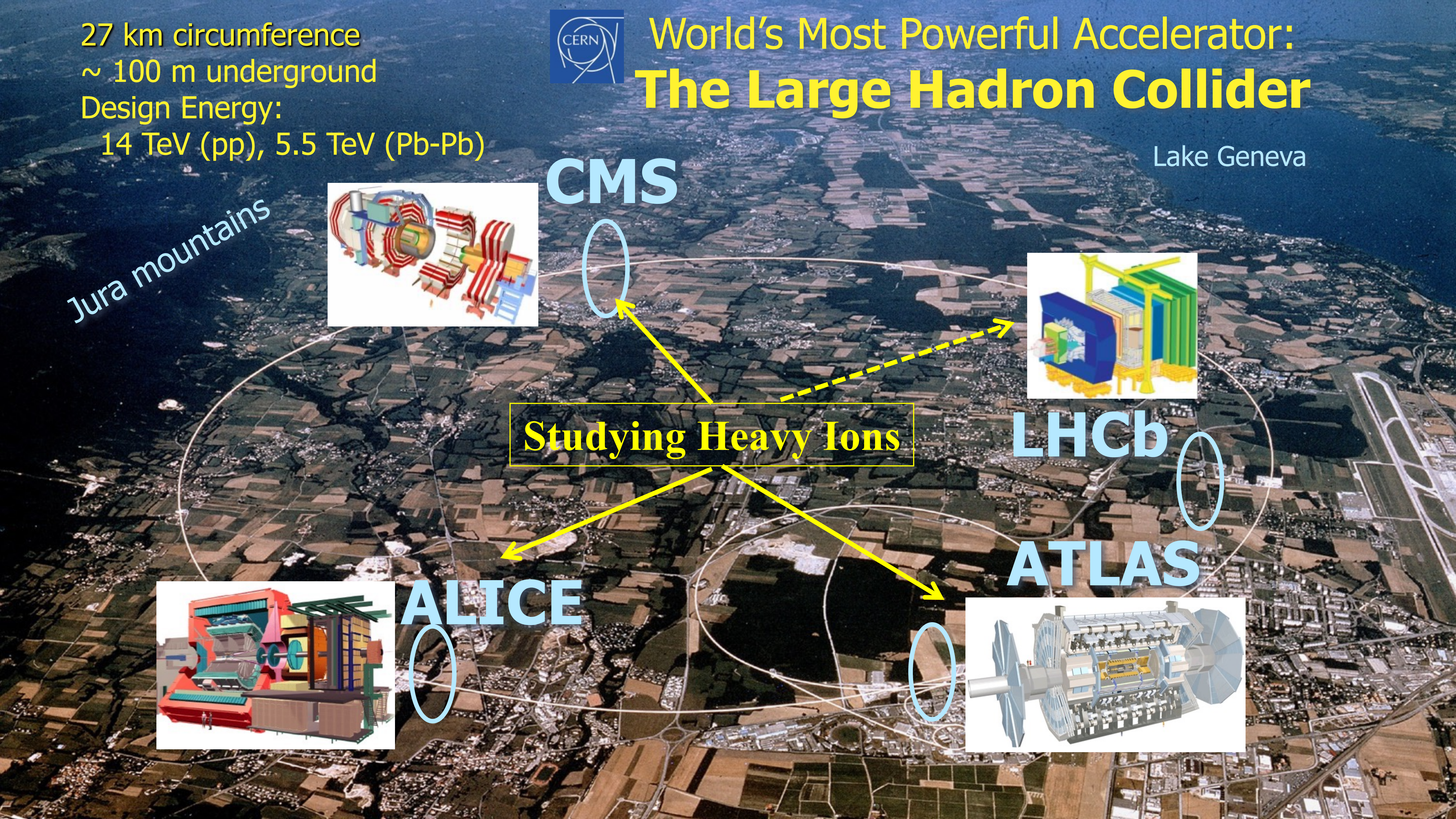
CERN
2009 -

LHC
Collider
Pb+Pb @5.5TeV/A

2009:
ALICE
CMS
ATLAS
LHCB


Geneva, Switzerland

CERN

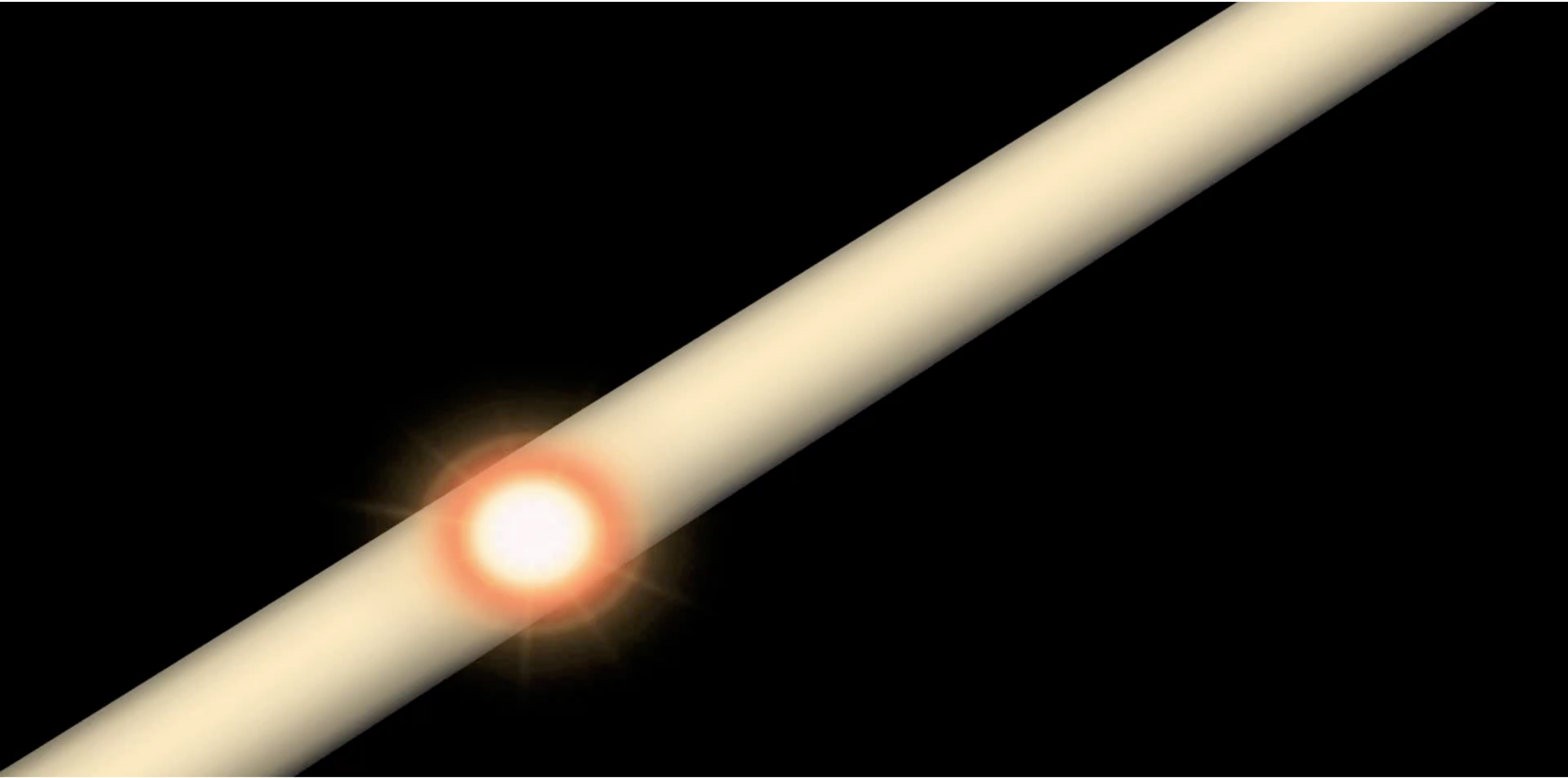

CERN
World's Most Powerful Accelerator:
The Large Hadron Collider
27 km circumference
~ 100 m underground
Design Energy:
14 TeV (pp), 5.5 TeV (Pb-Pb)
Lake Geneva
Jura mountains
CMS
LHCb
Studying Heavy Ions
ATLAS
ALICE

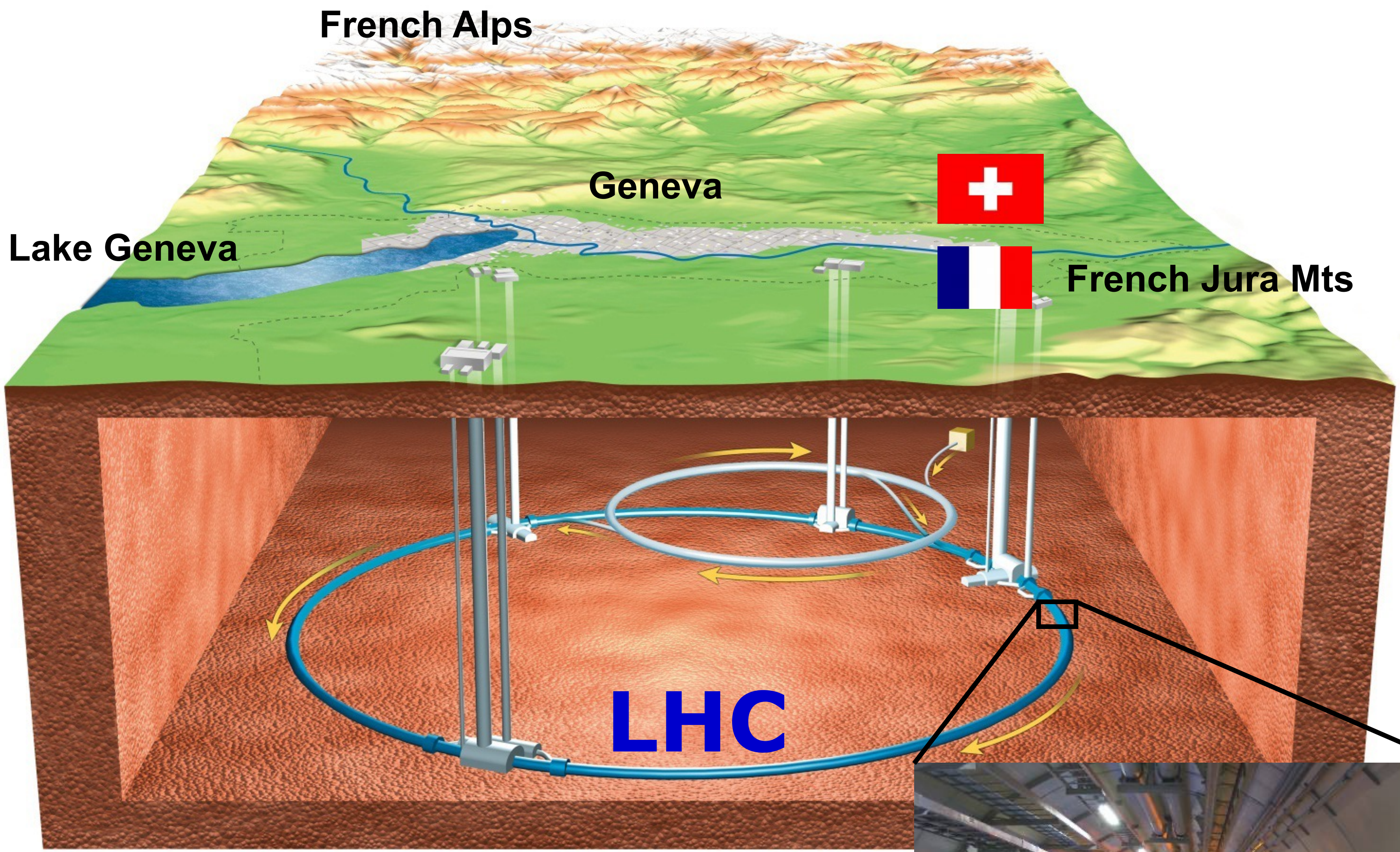
Large Hadron Collider (LHC)

27km tunnel:

- 50-150m below ground
- Two beams circulating in opposite directions
- Total of 9300 magnets: beams controlled by 1800 superconducting magnets (up to 8T)

The Large Hadron Collider

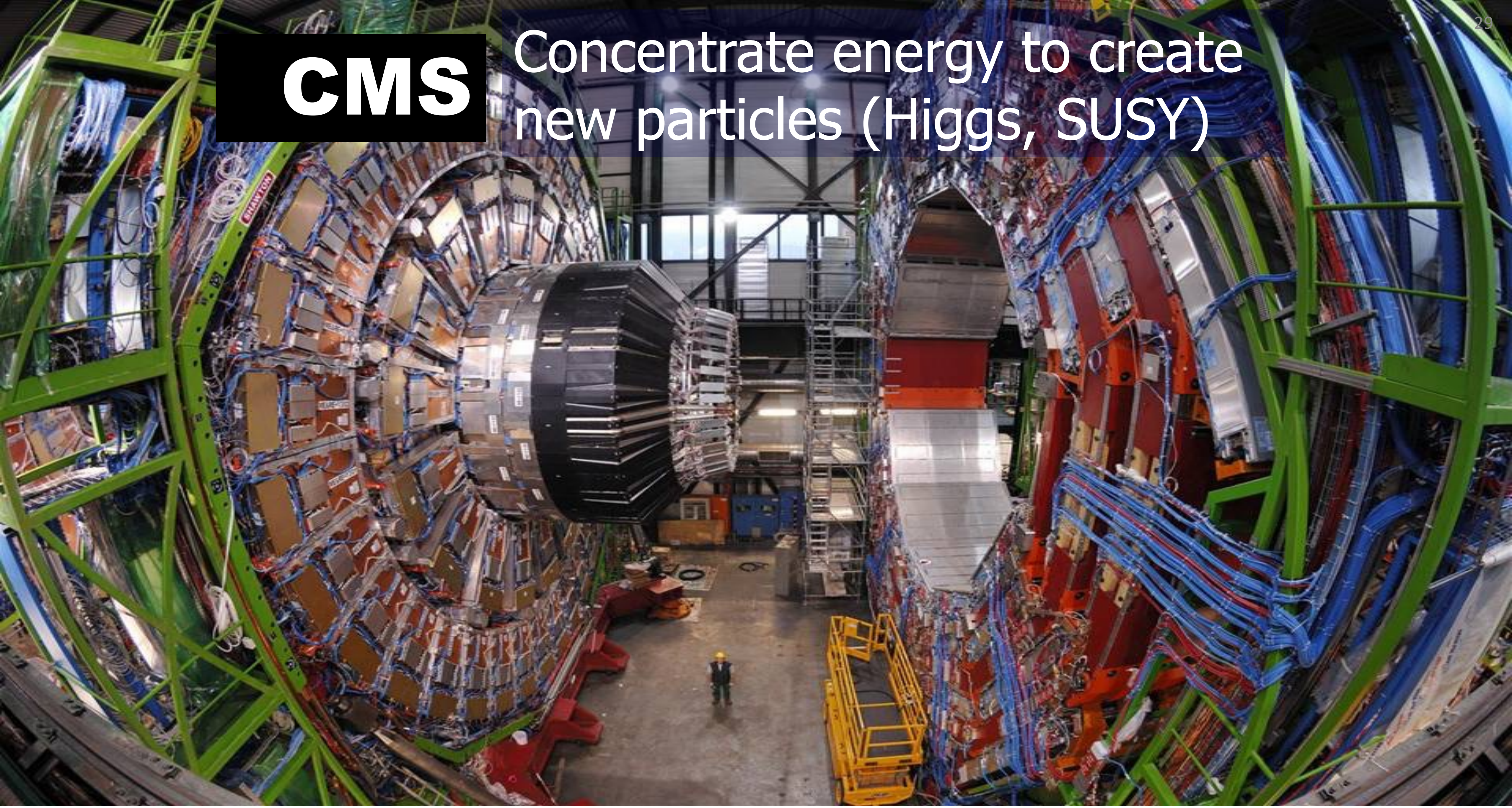




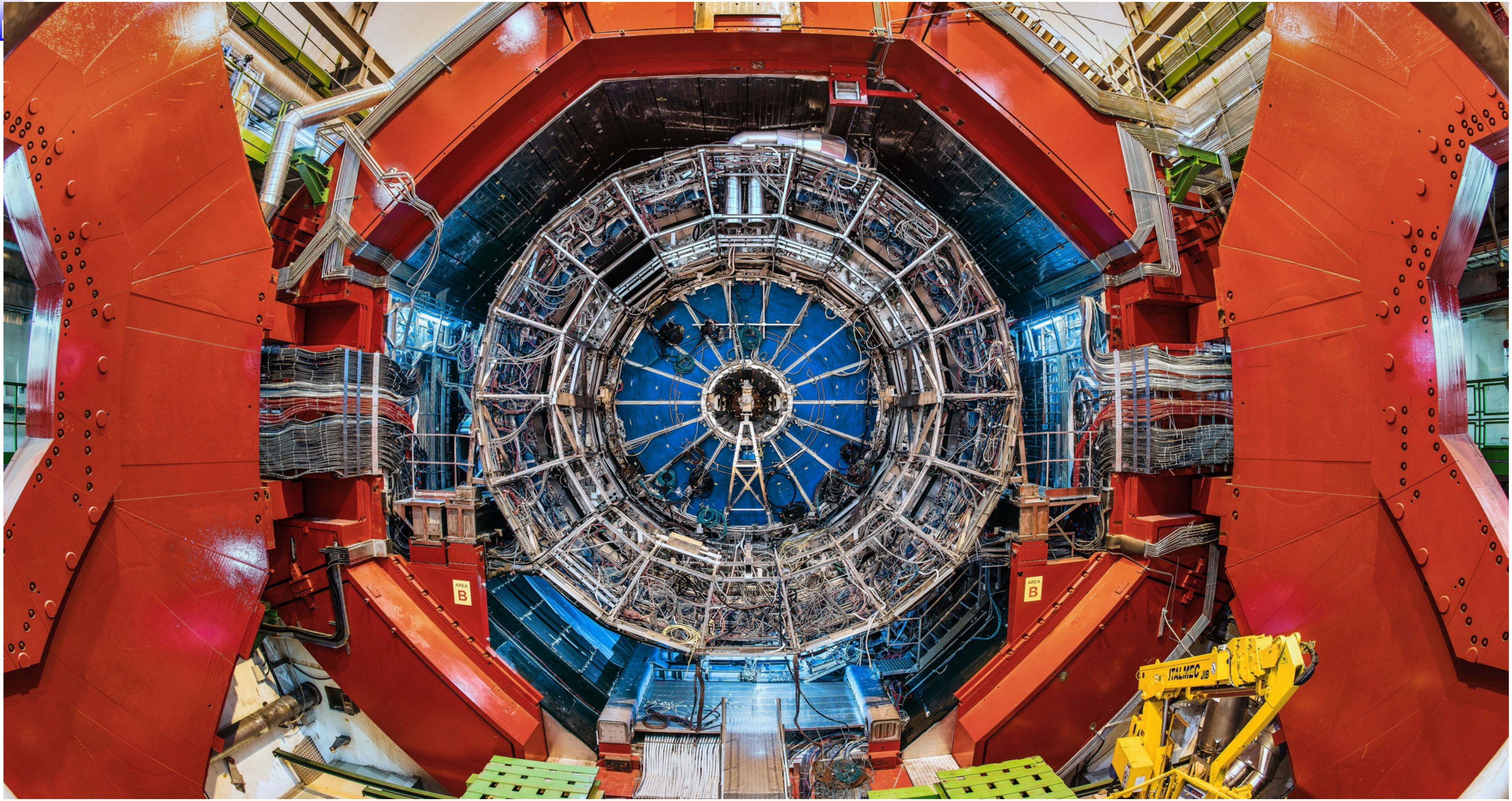
- Electric waves speed particles up
- Magnets bend them in a circle

CMS

Concentrate energy to create new particles (Higgs, SUSY)



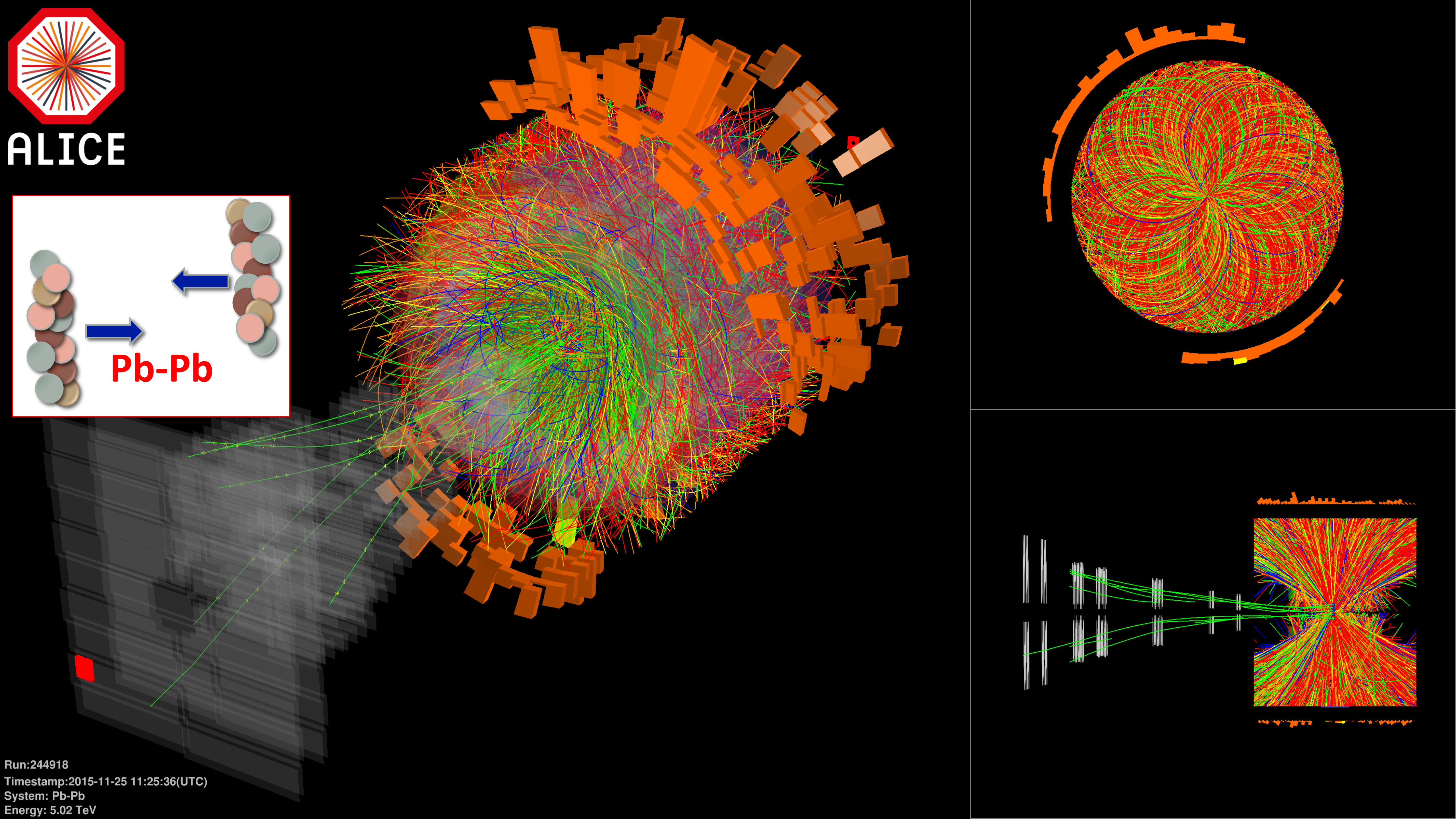
ALICE at Point-2 of the LHC



ALICE

Dedicated experiment to study QGP Matter

Size: 16 x 26 meters
Weight: 10,000 tonnes

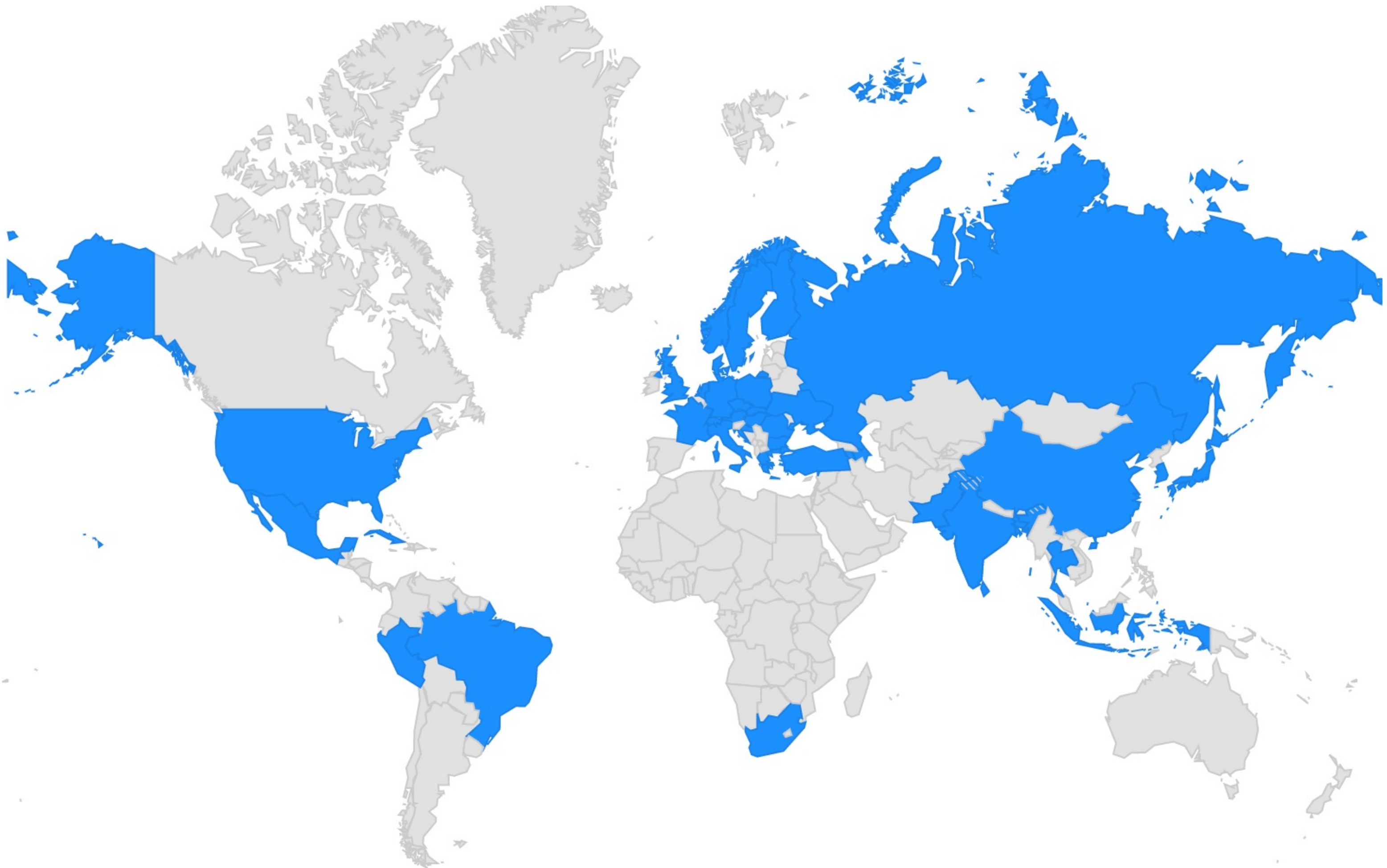

ALICE
Pb-Pb
Run:244918
Timestamp:2015-11-25 11:25:36(UTC)
System: Pb-Pb
Energy: 5.02 TeV

ALICE Collaboration

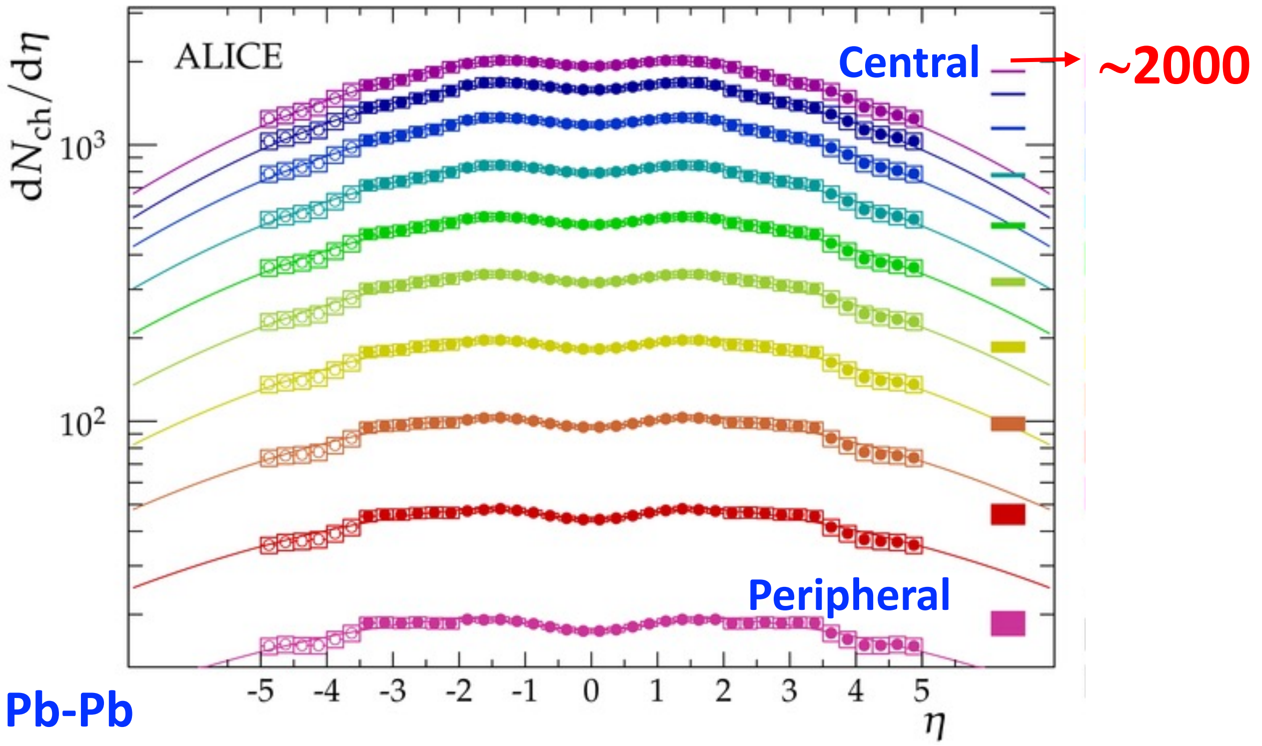
alice.cern

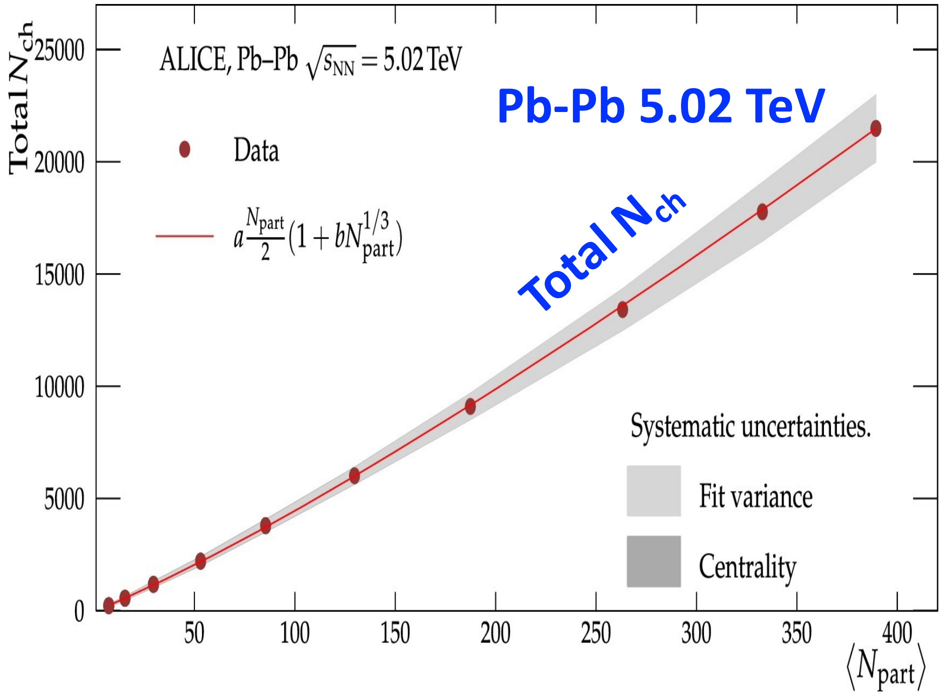
ALICE

40 countries, 171 institutes, 1996 members



Charged particle multiplicity





Number of charged particles in one collision:

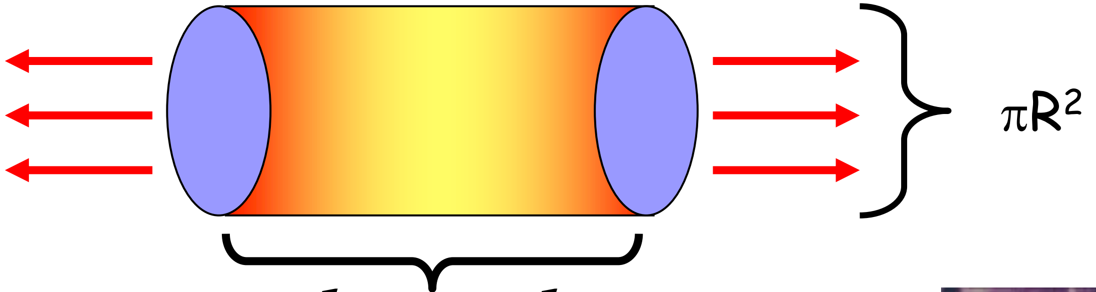
- Central collisions: 21400 ± 1300
- Peripheral collisions: 230 ± 38

Phys.Lett. B 772 (2017) 567577
Phys. Rev. Lett. 116 (2016) 222302

VERY LARGE NUMBER OF PRODUCED PARTICLES

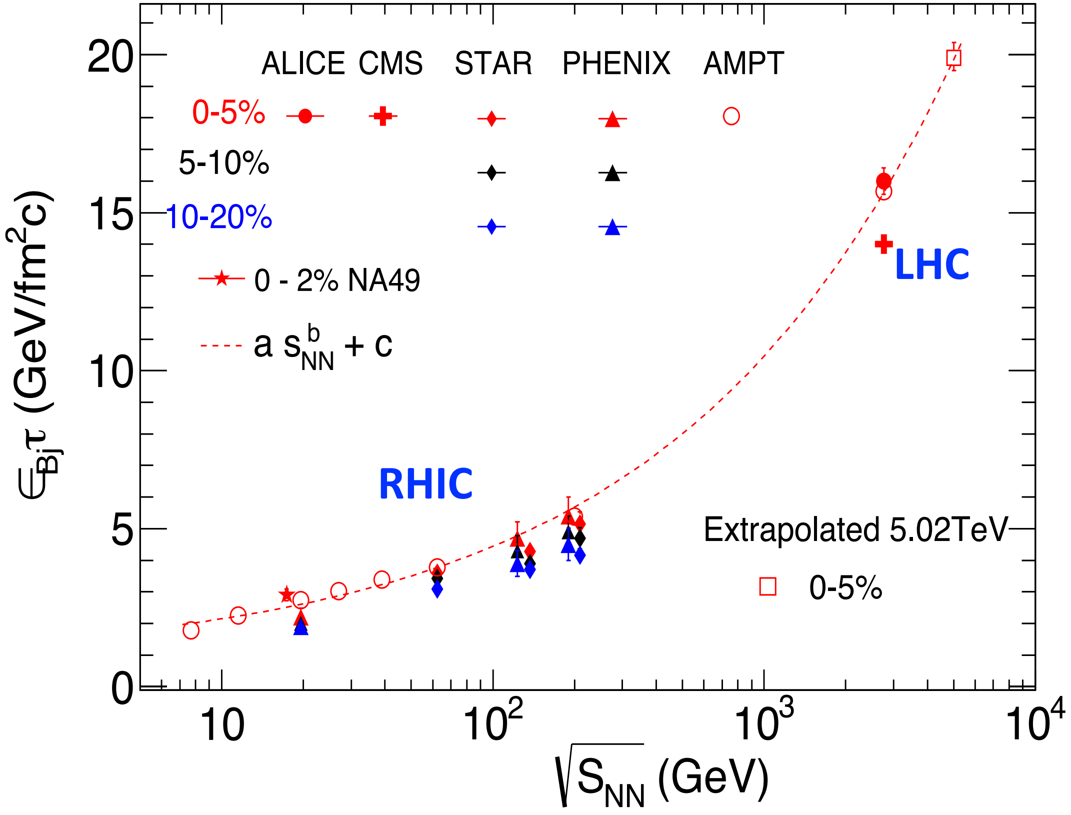
Particle density & Energy density

J. D. Bjorken, Phys. Rev. D 27, 140 (1983).



$$dz = \tau_0 dy$$

$$\varepsilon_{Bj}(\tau) = \frac{1}{\pi R^2 \tau}\frac{dE_T}{dy}$$
$$\approx \frac{1}{\pi R^2 \tau} < m_T > \frac{3}{2}\frac{dN_{ch}}{d\eta}$$



$\varepsilon.\tau \sim 16$ GeV/fm^{2}c

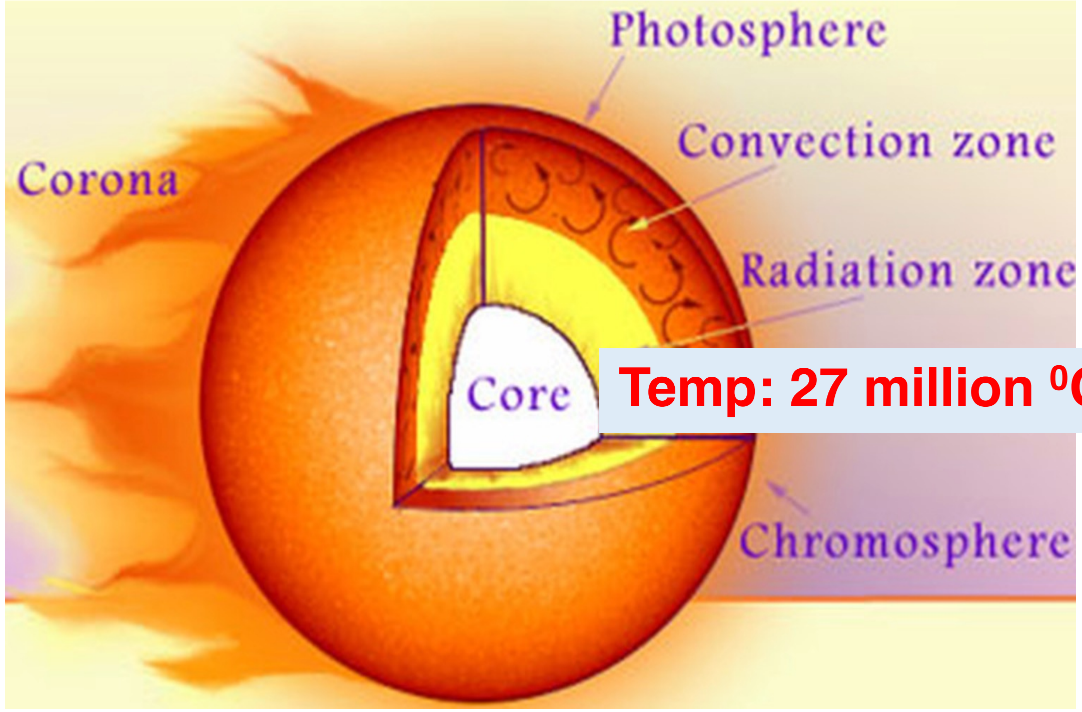
LARGEST ENERGY DENSITIES EVER ACHIEVED ....

S. Basu et al. PRC 93 (2016) 064902
R. Sahoo et al. Adv. in HEP, Vol. 2015

Photon Spectra and QGP temperature

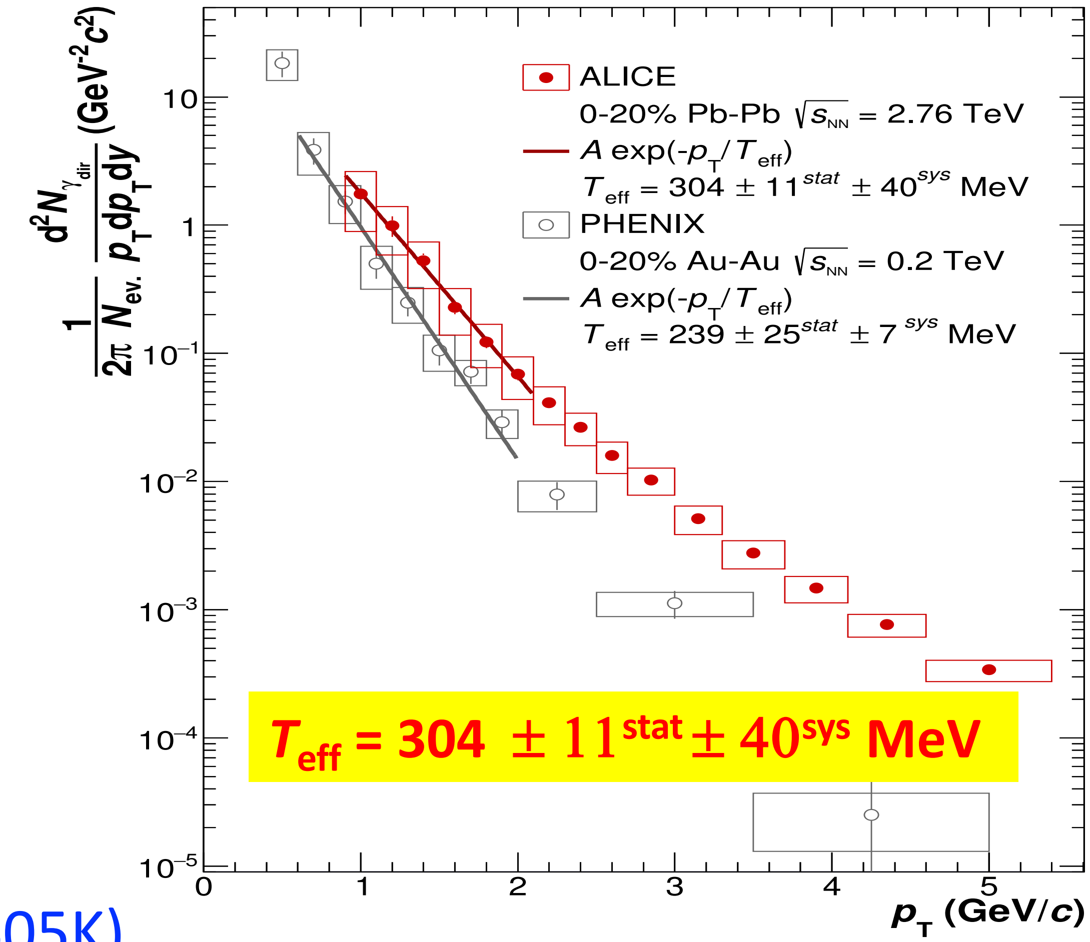
- Photons do not interact via the nuclear force → transparent to the medium
- Photons are emitted in all stages and are unaffected by the medium.

Core of the Sun:



Temp: 27 million ^{0}C

Phys. Lett. B 754 (2016) 235-248



T_{eff} = 304 ± 11^{stat} ± 40^{sys} MeV

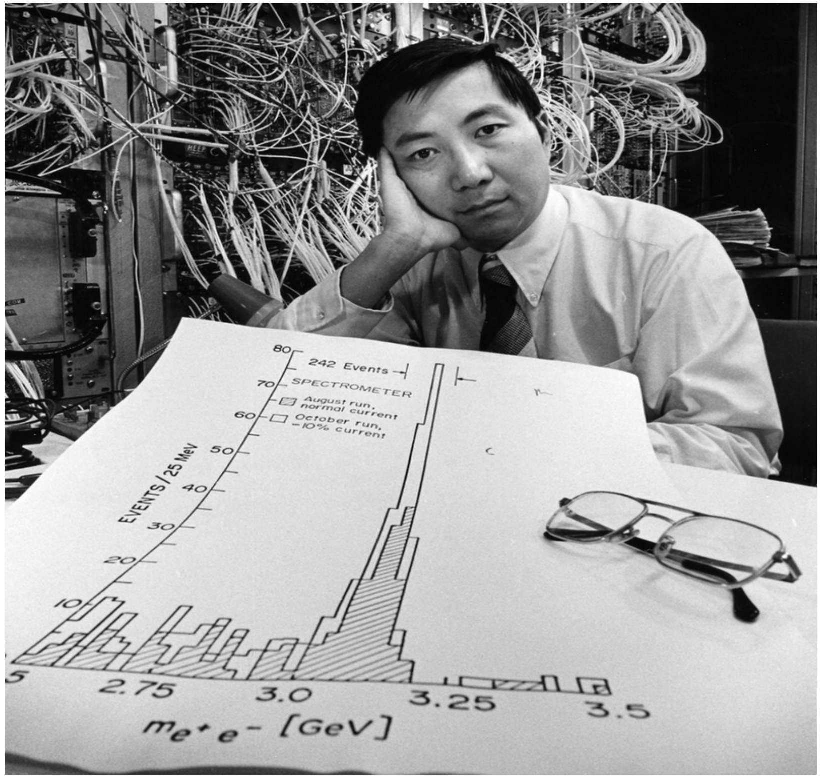
(1eV=11605K)

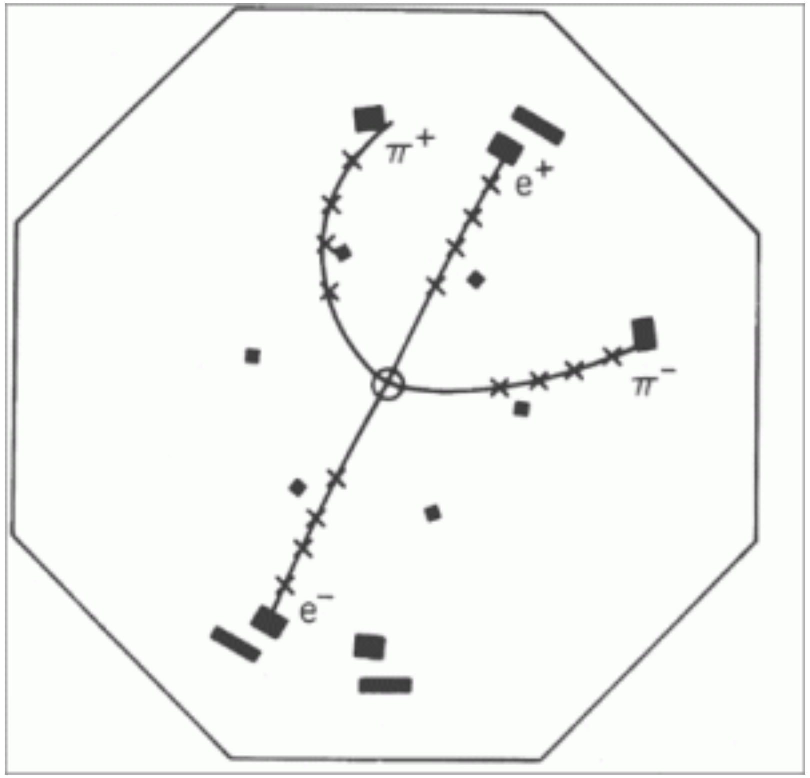
T_{eff} = 3,527,920 million deg

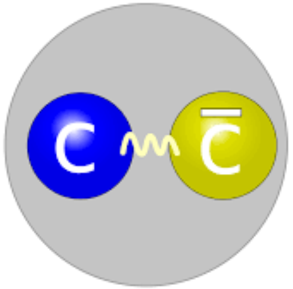
LARGEST EVER TEMPERATURE REACHED IN THE LAB ….

Charmonia, Machine Learning and QGP

The November Revolution: 1974







Crosses: spark chamber hits
Dark recatangles: ToF counters

$$\psi' \rightarrow \psi\pi^+\pi^-$$

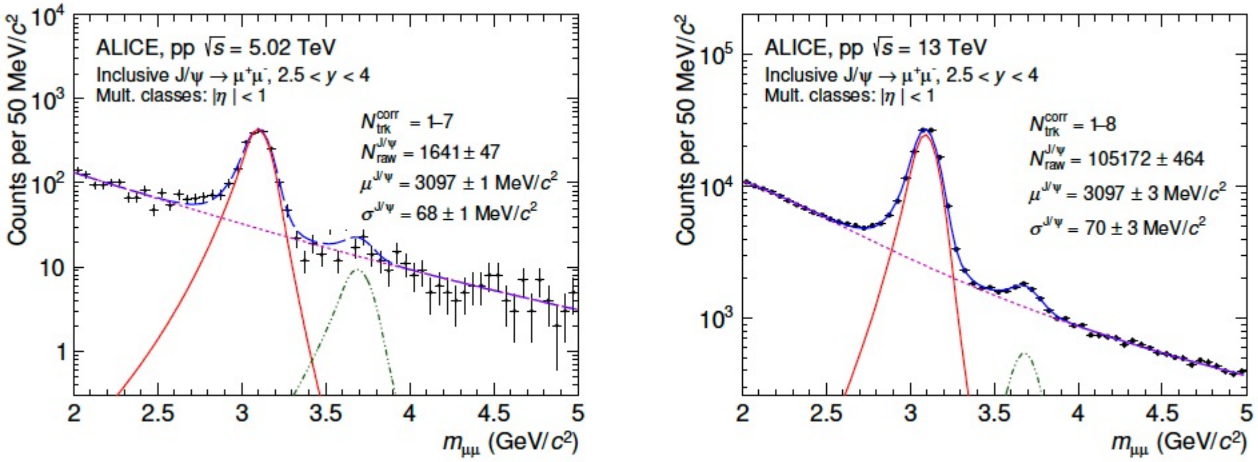
Griffiths

1.14. Using four quarks (u, d, s, and c), construct a table of all the possible baryon species. How many combinations carry a charm of +1? How many carry charm +2, and +3?

The Nobel Prize in Physics 1976

Burton Richter@SLAC: e^+e^-
Sam Ting@BNL: pp

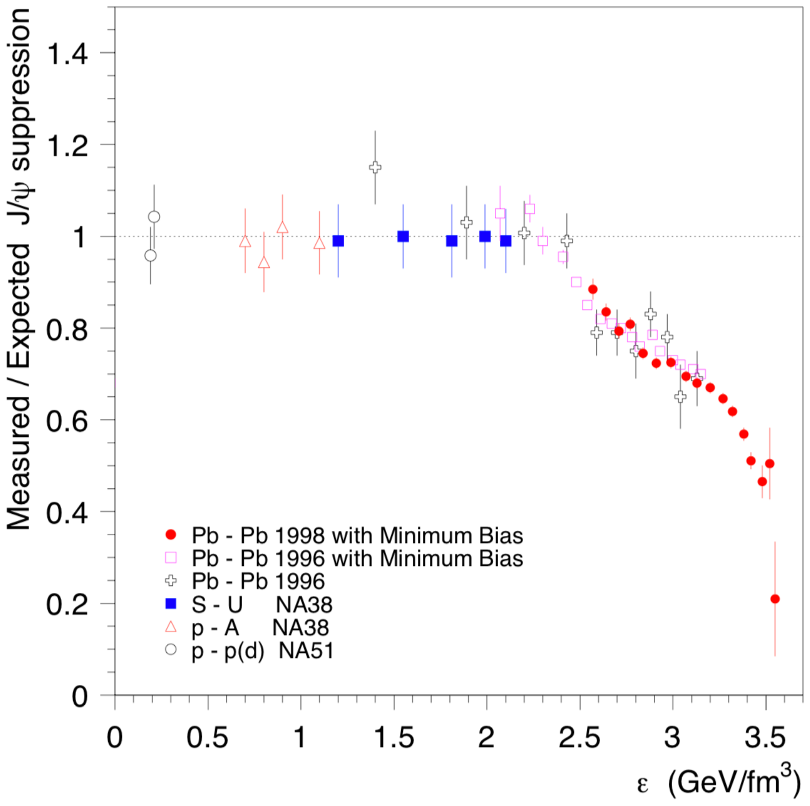
Our ~~discovery~~ (measurement) of J/ψ



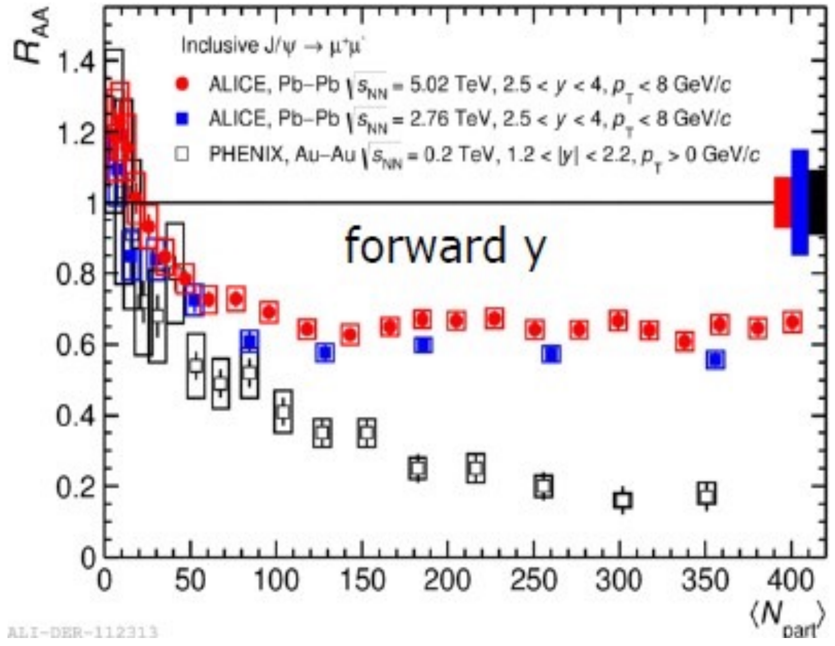
D. Thakur, R. Sahoo et al. ALICE Collaboration JHEP06(2022)015

https://alice-notes.web.cern.ch/node/734, ANA-734, D. Thakur and R. Sahoo
https://alice-publications.web.cern.ch/node/5122, D. Thakur, R. Sahoo et al.

Experimental observation of J/ψ Suppression

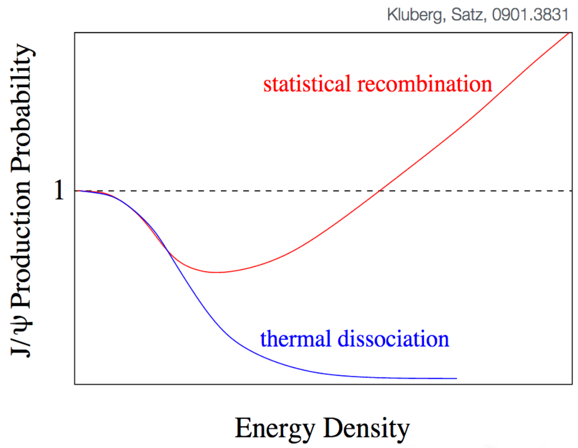


CERN SPS (NA50): observed J/ψ suppression as a function of energy density for various collision species. Note that the critical energy density for a partonic medium is 1 GeV/fm³.



(Physics Letters B 766 (2017) 212–224)

http://alice.web.cern.ch/content/mystery-jpsi



$$R_{AA}(p_T) = \frac{d N_{AA}/dp_T}{\langle T_{AA}\rangle \times d\sigma_{pp}/dp_T}$$

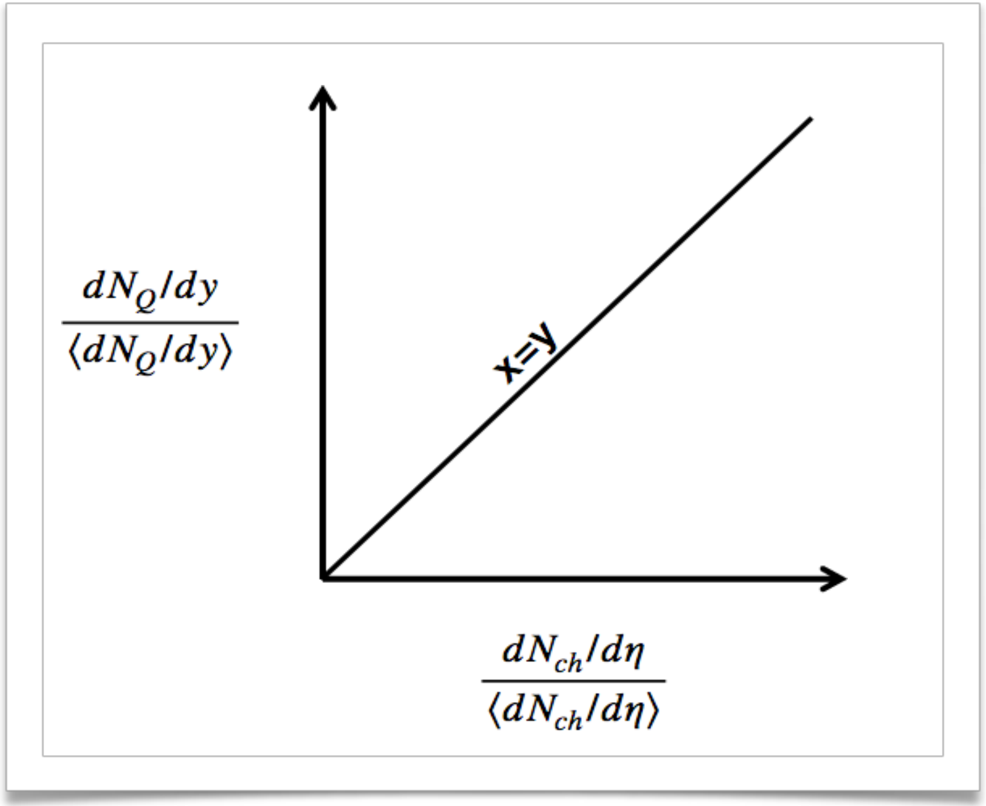
RAA < 1 : Suppression of yield due to presence of medium

RAA > 1 : No suppression and hence no medium

- **Big question:** use of pp as reference at LHC
- **Need :** pp should be investigated properly

Study of J/ψ production in pp collisions

- Investigate charmonium production in pp as a function of multiplicity



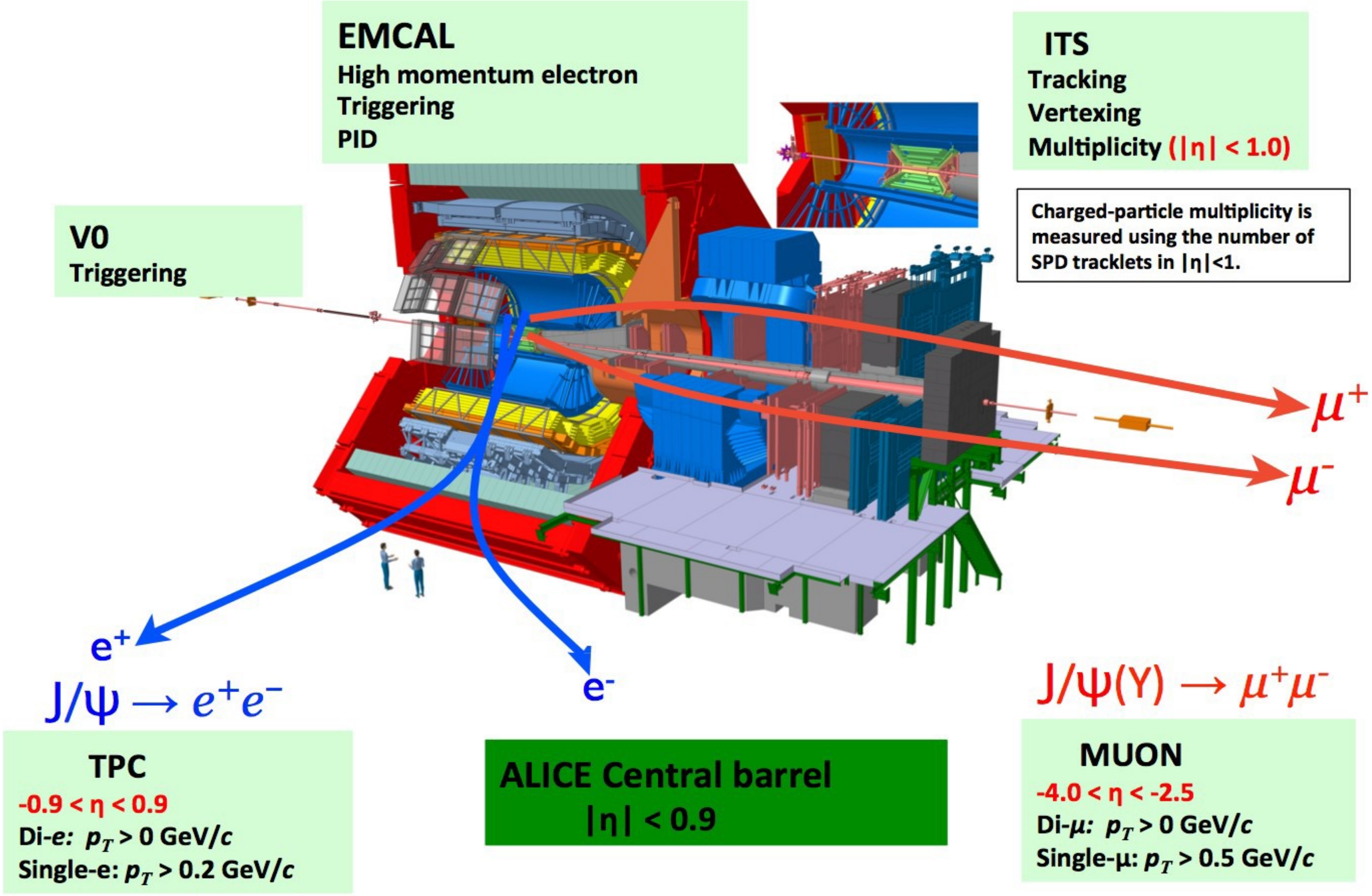
Observables:

- X axis: relative charged-particle density
- Y axis: relative quarkonium yield (w.r.t. minimum bias)

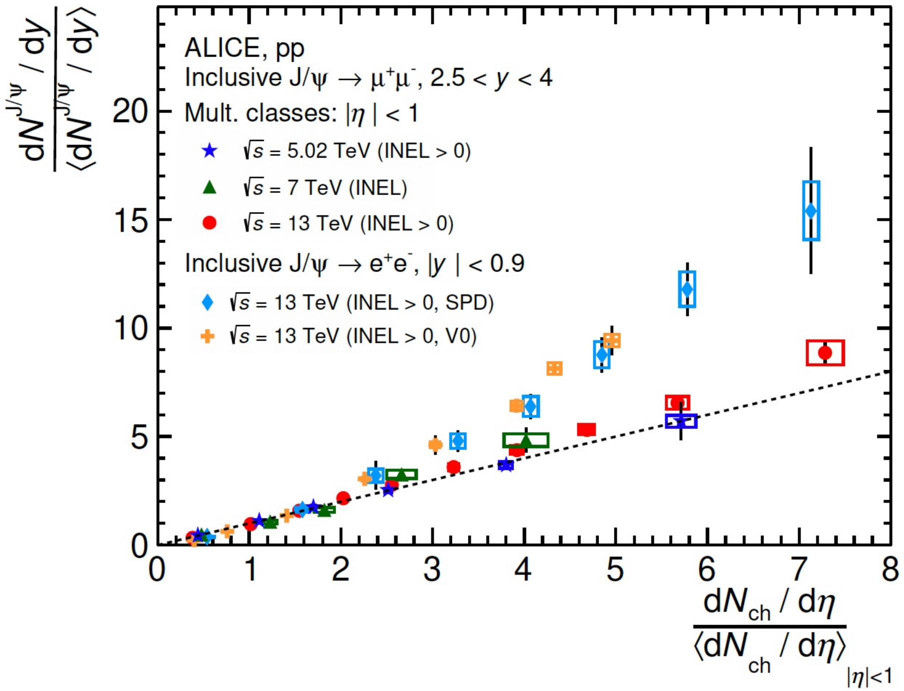
Advantages:

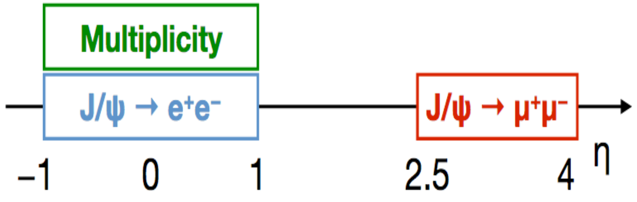
- Several correction factors cancel out in relative yields
- Easy to compare the results across collision systems and energies

Study of J/ψ production in ALICE



Study of J/ψ production in ALICE pp collisions



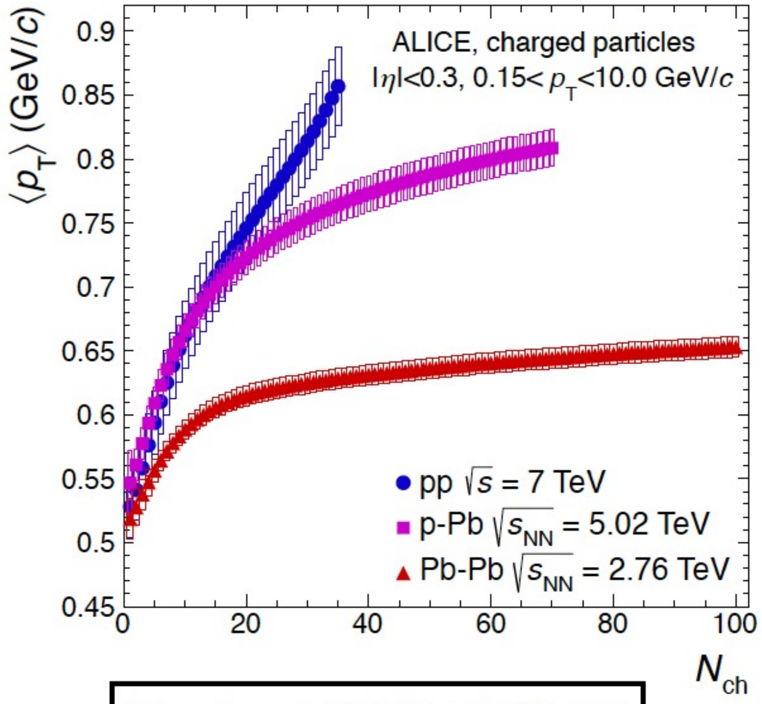


ALICE: JHEP06(2022)015

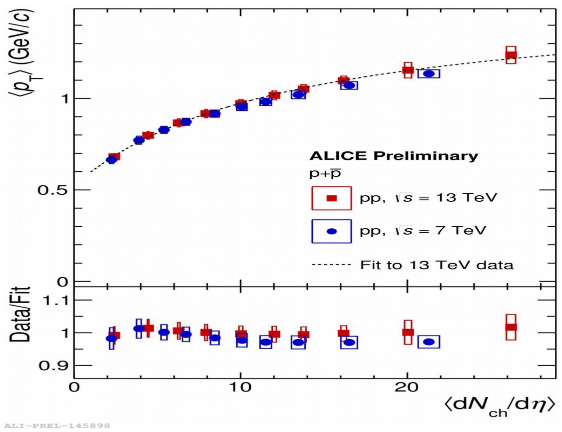
https://alice-notes.web.cern.ch/node/734, ANA-734, D. Thakur and R. Sahoo
https://alice-publications.web.cern.ch/node/5122, D. Thakur, R. Sahoo et al.

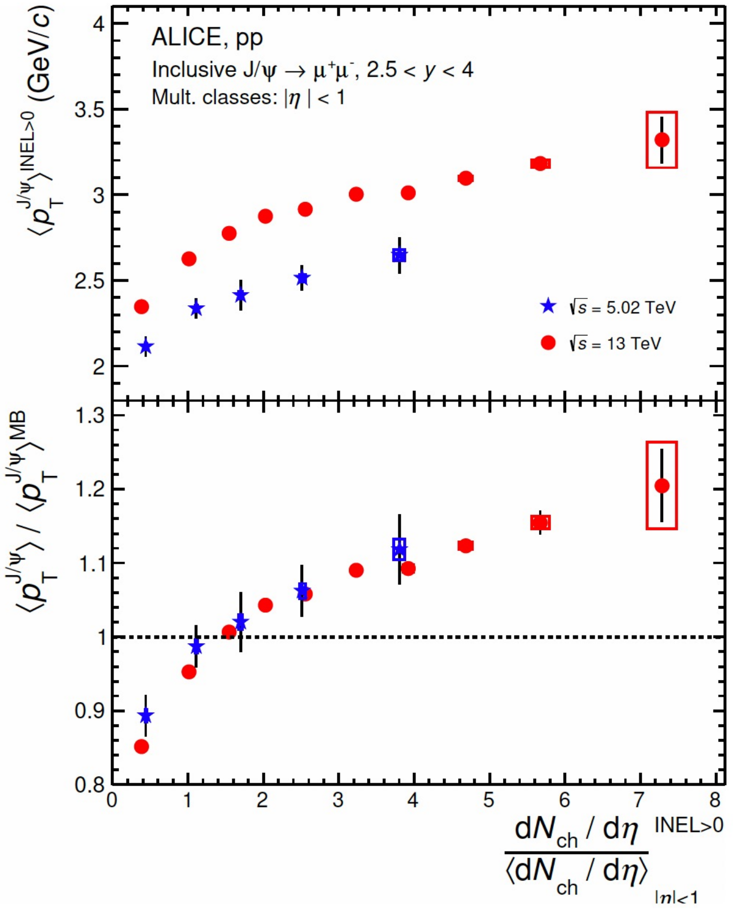
- Relative J/Ψ yield at midrapidity is compared to the forward rapidity yield as a function of midrapidity relative charged-particle multiplicity
- Midrapidity yields exhibit faster than linear increase
- The results using midrapidity multiplicity selection based on the SPD detector ($|\eta| < 1$) and forward-rapidity multiplicity selection based on the V0 detector ($-3.7 < \eta < -1.7$ and $2.8 < \eta < 5.1$) are found to be compatible within the uncertainties
- Therefore, the different trends in the multiplicity dependence of the J/Ψ production observed at midrapidity and forward rapidity are not due to a possible auto-correlation bias

⟨ p_T ⟩ of J/ψ vs. multiplicity in pp collisions



Phys. Lett. B 727 (2013) 371-380





- ❖⟨ p_T ⟩ of J/ψ increases with increasing multiplicity with a little saturation towards higher multiplicities
  - collectivity? (Phys. Lett. B 727 (2013) 371-380)

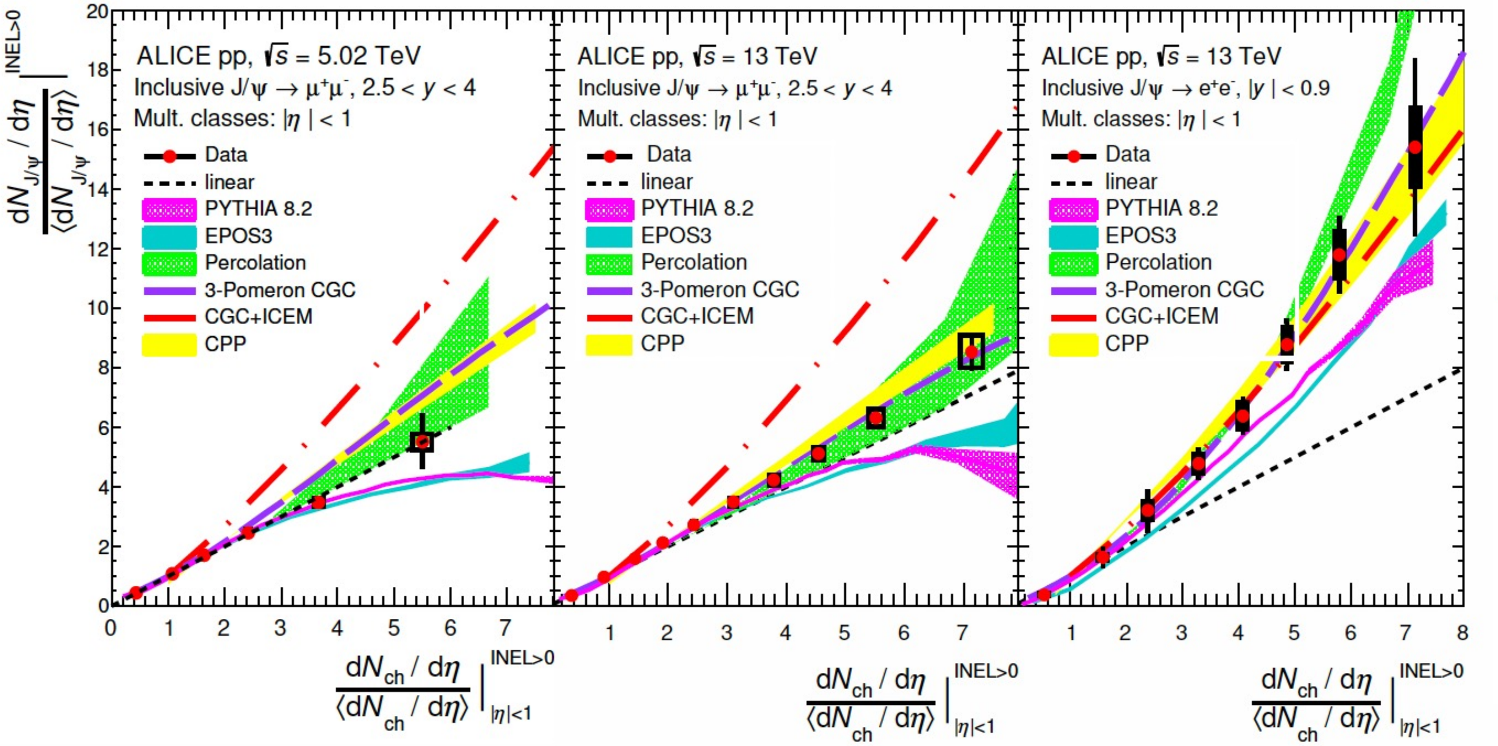
- ❖Possible explanation of MPI prospect: the high multiplicity events are produced by MPIs and in the absence of CR, incoherent superposition of MPIs would lead to constant ⟨ p_T ⟩ at high multiplicity.

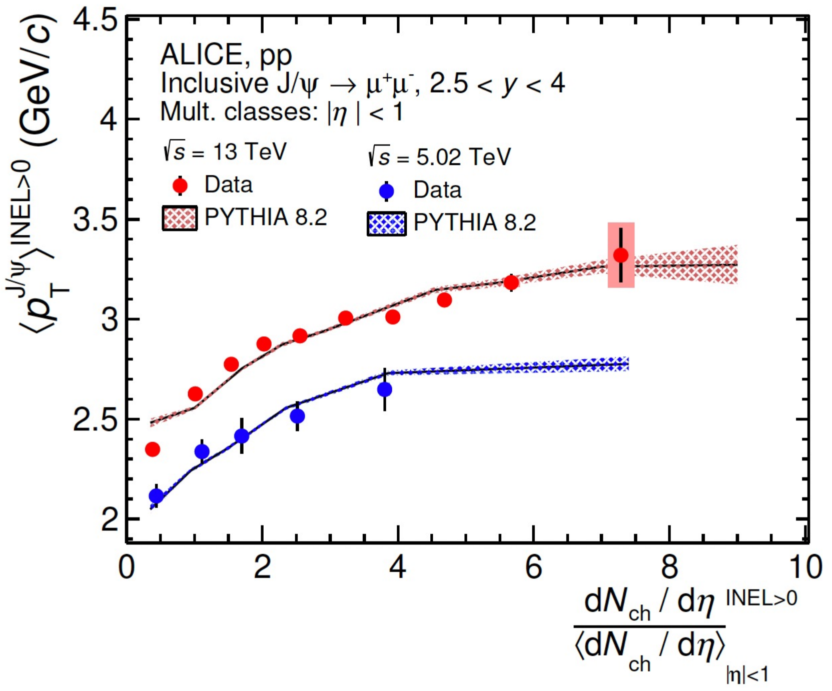
- ❖ <p_T> of J/ψ vs multiplicity trend is energy-independent

- ❖ The multiplicity dependent trend is the same for light-flavor and heavy-flavor particles

https://alice-notes.web.cern.ch/node/734, ANA-734, D. Thakur and R. Sahoo
https://alice-publications.web.cern.ch/node/5122, D. Tahkur, R. Sahoo et al.

Theoretical model comparison: J/ψ production





None of the theoretical models give a complete picture of the J/ψ prodution dynamics in pp collisions

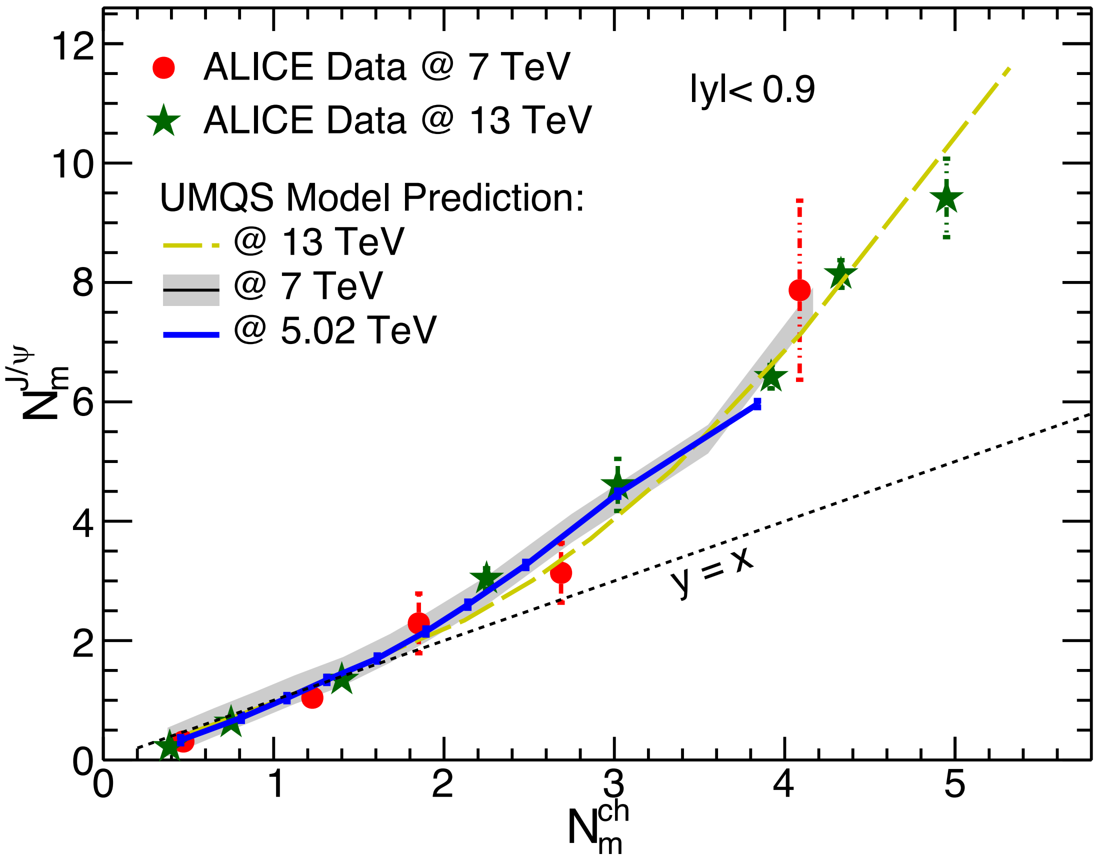
PYTHIA8 with CR gives a reasonable description of the multiplicity dependence of J/ψ $\langle p_T \rangle$ in pp collisions

https://alice-notes.web.cern.ch/node/734, ANA-734, D. Thakur and R. Sahoo

https://alice-publications.web.cern.ch/node/5122, D. Tahkur, R. Sahoo et al.

A lot of room for theory/phenomenology!

UMQS model explains ALICE J/ψ production

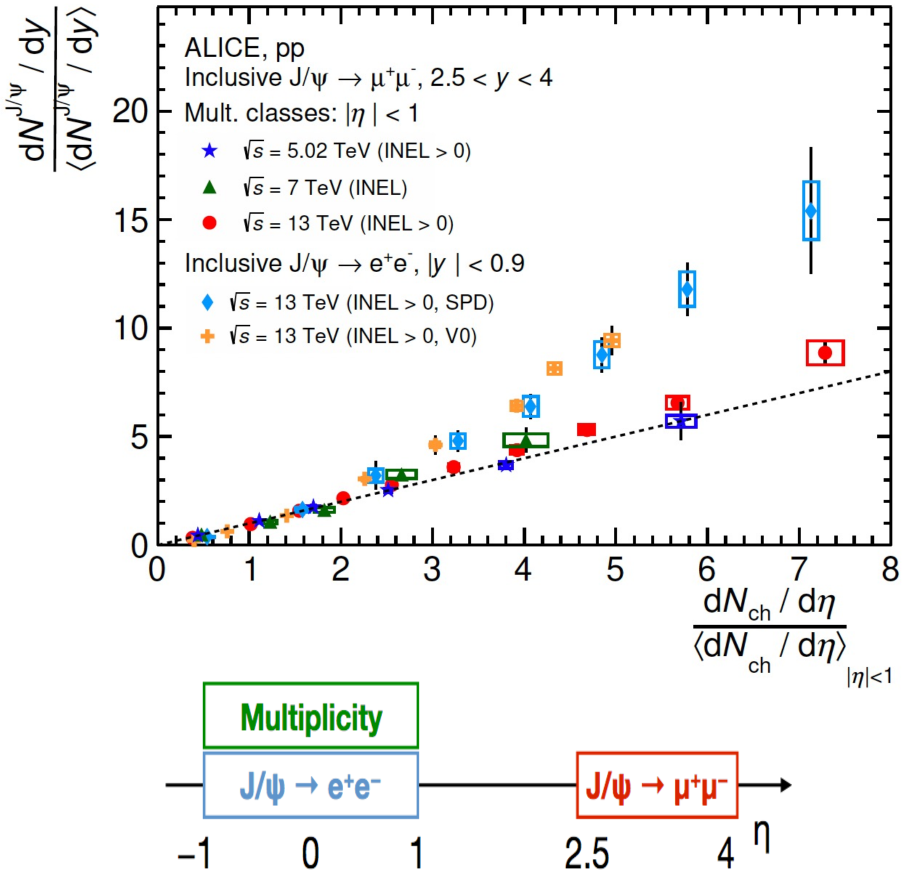


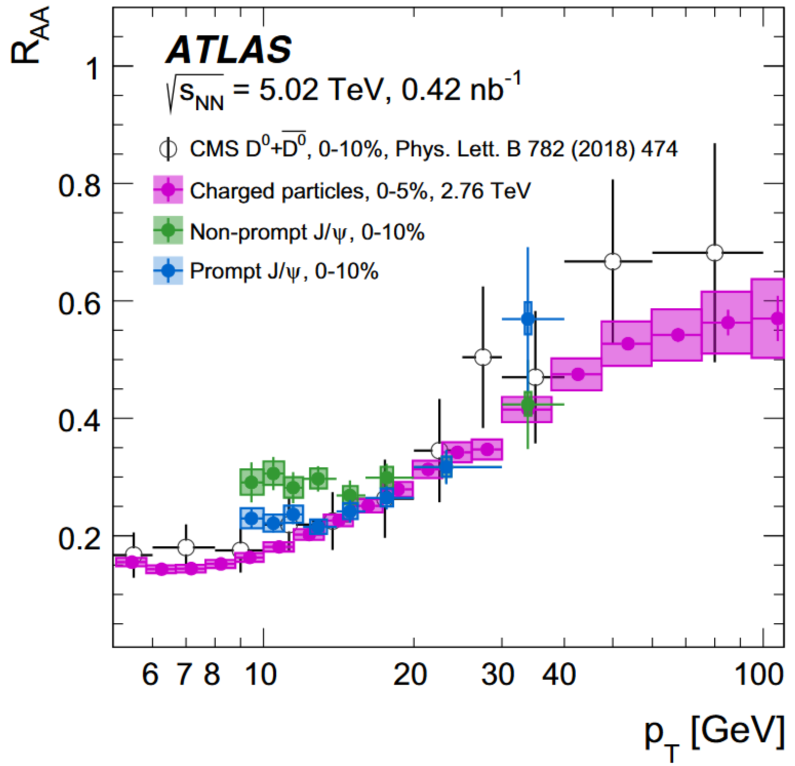
- ✓ J/ψ self-normalized yield as a function of self-normalized multiplicity follows a scaling across collision energies.
- ✓ Unified Model of Quarkonia Suppression (UMQS) model which incorporates the suppression of J/ψ through color screening, gluonic dissociation, and collision damping and regeneration of charmonium due to correlated c – cbar pairs.

C.R. Singh, S. Deb, R. Sahoo, J. Alam, Eur. Phys. J. C, 82, 542 (2022)

Need of ML: Separating prompt from nonprompt

Inclusive



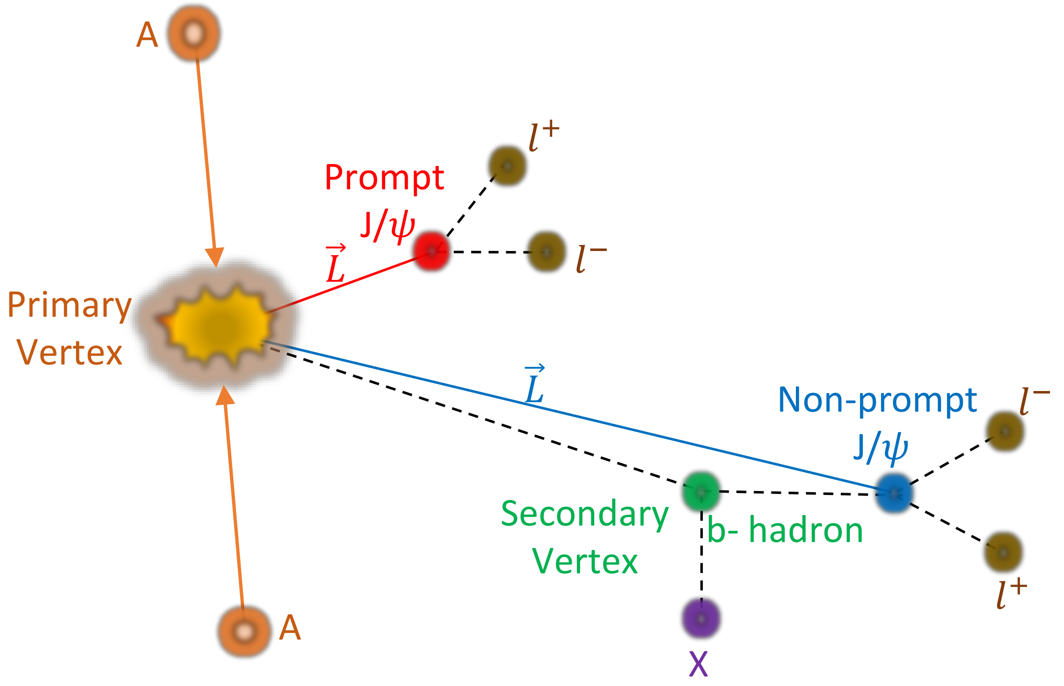


Nuclear modification factor for prompt and non-prompt J/ψ are different.

D. Thakur, R. Sahoo et al (ALICE), JHEP (2022)

ML: Separating prompt and non-prompt J/ψ @LHC

- J/ψ (3.096 GeV/c^2)
- In experiments, $J/\psi \rightarrow \mu^+ + \mu^-$ or $J/\psi \rightarrow e^+ + e^-$
- Prompt Production: Direct production/ decay of heavier charmonia states
- Non-prompt Production: Products of beauty hadron weak decays
- Prompt and non-prompt J/ψ are topologically different

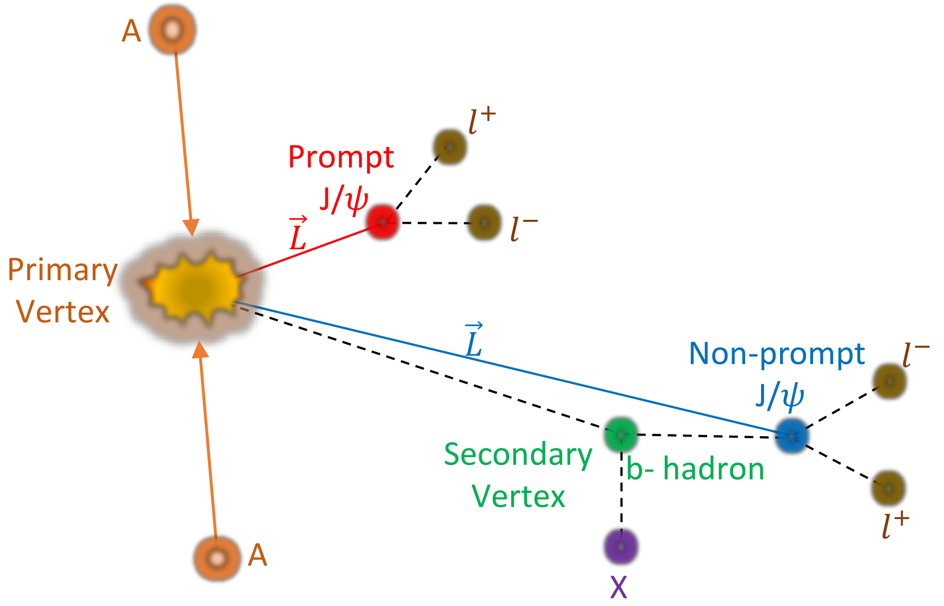


S. Prasad, N. Mallick and R. Sahoo, *Phys. Rev. D 109, 014005 (2024)*

ML to separate prompt and non-prompt J/ψ @LHC

Simulating pp collisions at $\sqrt{s}$ =13 TeV using PYTHIA8 → Training machine-learning models → Using the machine-learning models to predict prompt and non-prompt yields at different energies → Results

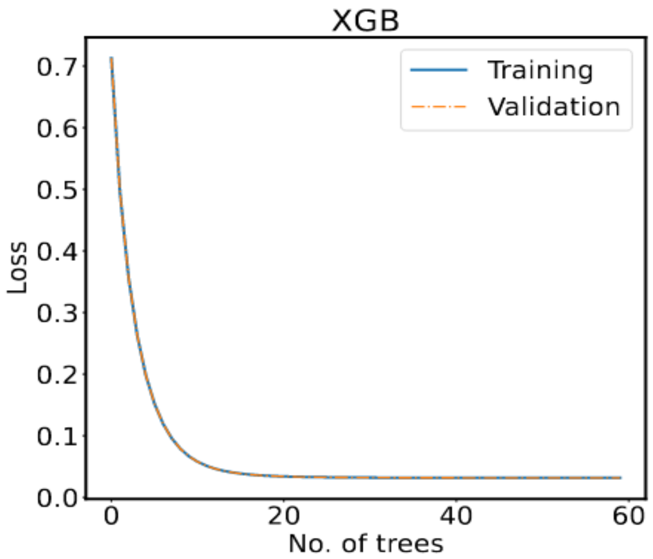
- PYTHIA8 (4C-tune) 20 billion minimum bias events for pp $\sqrt{s} = 13$ TeV
- The coordinates of the primary vertex are randomised following a Gaussian distribution: experiment-like scenario
- $J/\psi \rightarrow \mu^+ + \mu^-$ channel is used to reconstruct invariant mass ($m_{\mu\mu}$), transverse momentum ($p_{T,\mu\mu}$), pseudorapidity ($\eta_{\mu\mu}$) and rapidity ($y_{\mu\mu}$) of the dimuons
- Pseudoproper decay length ($c\tau$) of the reconstructed dimuon pairs along with $m_{\mu\mu}$, $p_{T,\mu\mu}$, and $\eta_{\mu\mu}$ are taken as inputs



$$c\tau = \frac{c\, m_{J/\psi}\, \vec{L}\, .\, \vec{p_T}}{|\vec{p_T}|^2}$$

ML parameters to separate prompt and non-prompt J/ψ @LHC

- Background : Prompt : Non-prompt = 20 : 10 : 1
- Classification models required to be trained on similar number of training instances → oversampling of data is done
- Dataset for Training : Testing : Validation = 81 : 10 : 9
- Parameters are chosen through a grid search method (Making an array of all possible parameters and training to find the parameter values for minimum loss)

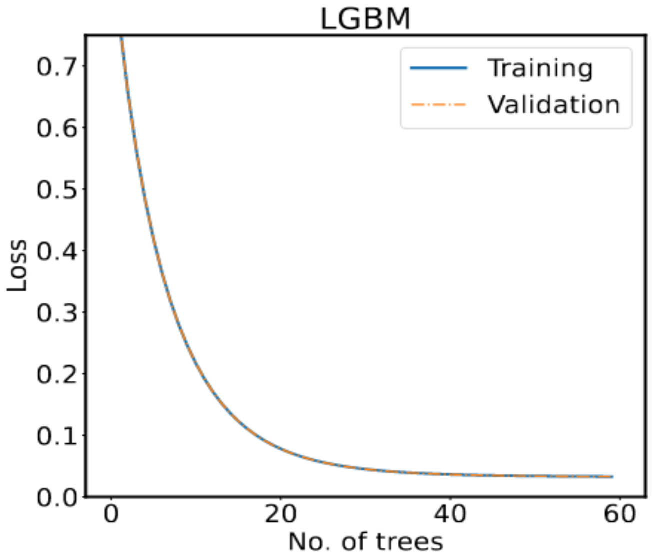


- Grid Search Method to find best parameters

	XGB	LGBM
Learning rate	0.3	0.1
Sub-sample	1.0	1.0
No. of trees	60	60
Maximum depth	3	3
Objective	*softmax*	*softmax*
Metric	*mlogloss*	*multilogloss*

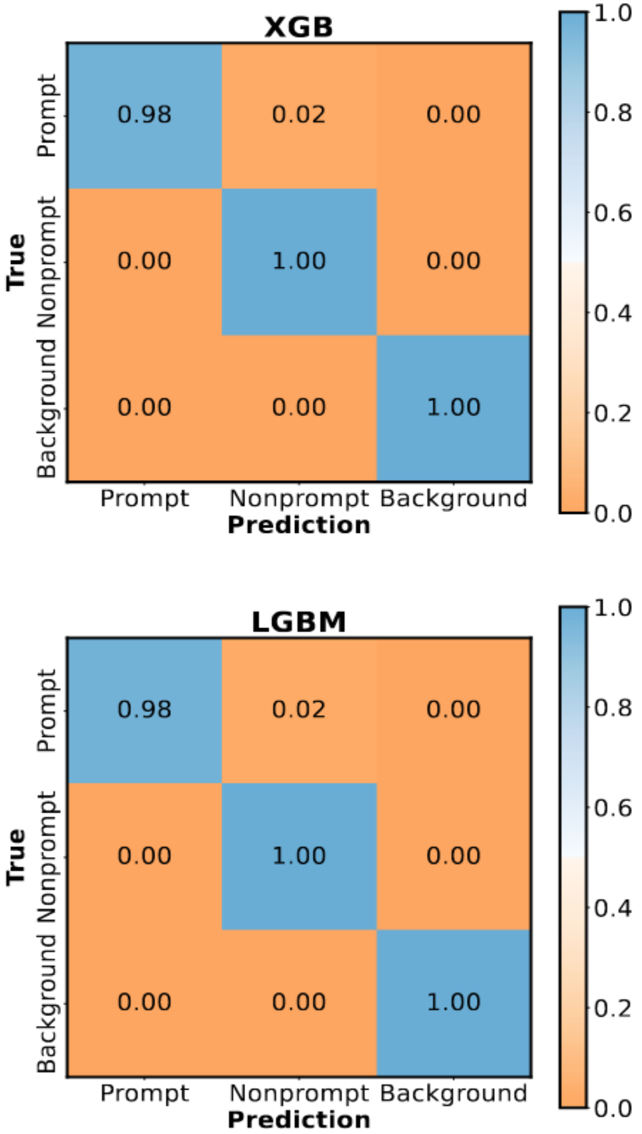
S. Prasad, N. Mallick and R. Sahoo, *Phys. Rev. D 109, 014005 (2024)*

- Loss saturates around 25 and 45 trees for XGB and LGBM
- Training and validation curves are on top of each other → No overfitting/underfitting

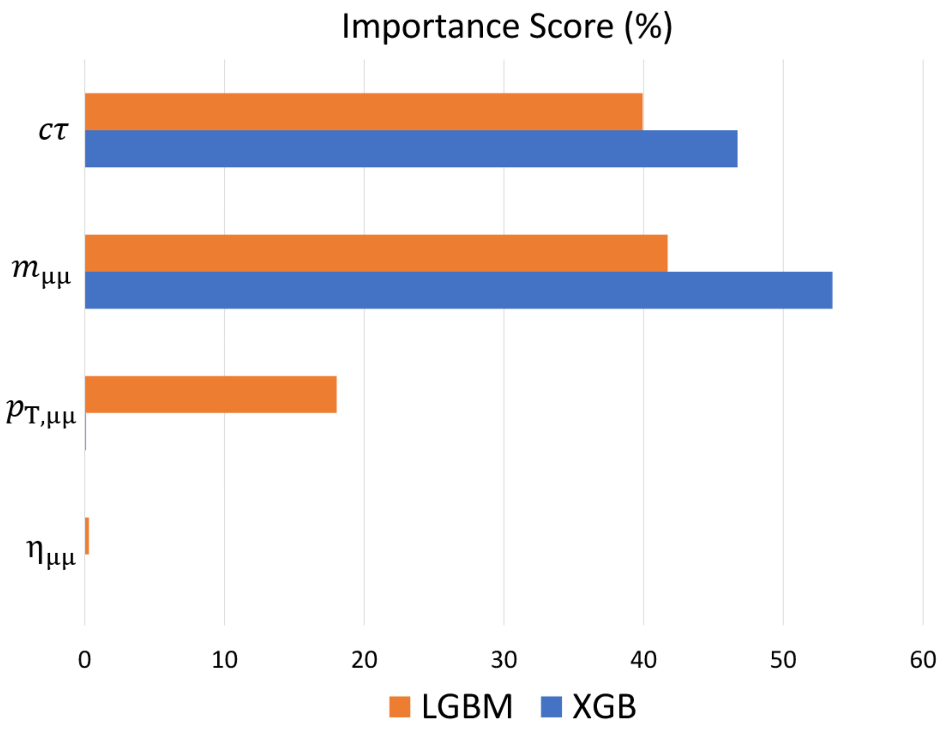


LGBM: Light Gradient Boosting Machine
XGB: Extreme Gradient Boosting

ML Model performance to separate prompt and non-prompt J/ψ @LHC

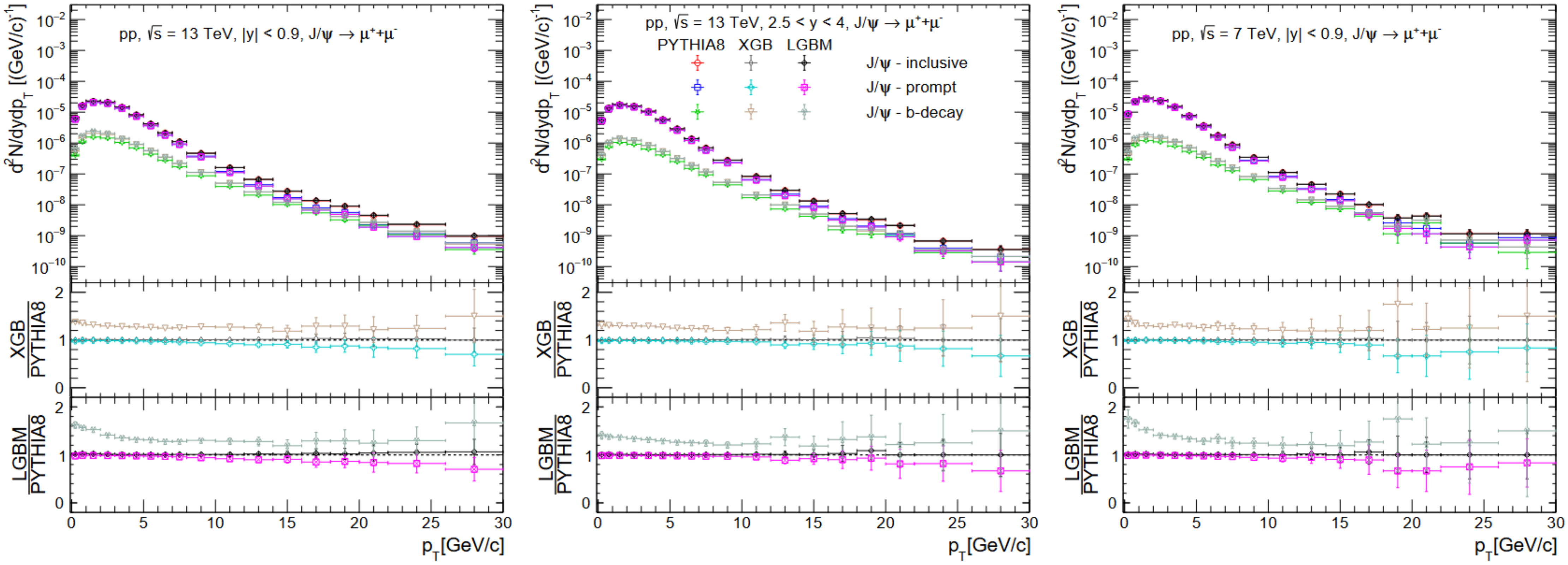


- Confusion Matrix talks about the mispredictions given by the model for each class
- Both XGB and LGBM perfectly separate the inclusive J/ψ from the uncorrelated background pairs
- Both models mispredict 2% of prompt J/ψ as the non-prompt → Raises non-prompt yield



- The importance score tells how important a feature is for the decision-making of the models
- The importance score of the invariant mass of dimuons is highest for both models
- $c\tau$ contributes to decision making of the models significantly

ML Results: Transverse momentum spectra

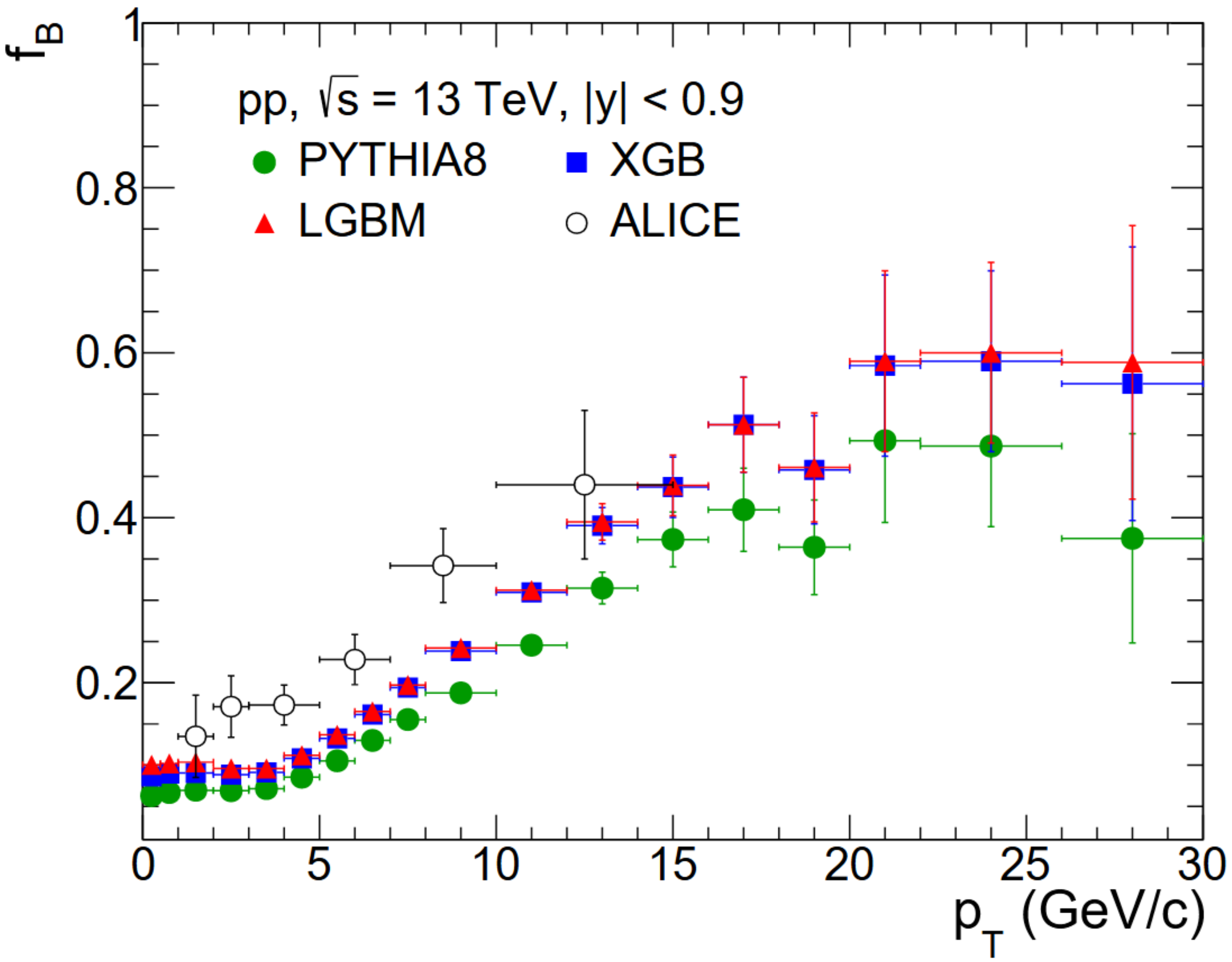


- Both XGB and LGBM give accurate predictions for p_{T}-spectra for inclusive and prompt-J/ψ both in mid and forward rapidity in pp collisions at $\sqrt{s} = 13$ TeV and 7 TeV
- The ML models overpredict the non-prompt J/ψ throughout the p_{T} spectra for both the collision energy and rapidity

  → Expected from the confusion matrix

S. Prasad, N. Mallick and R. Sahoo, *Phys. Rev. D 109, 014005 (2024)*

ML Results: Fraction of non-prompt J/ψ yield



$$\sigma_{\text{nonprompt J}/\psi} = f_B \cdot \sigma_{\text{J}/\psi},$$
$$\sigma_{\text{prompt J}/\psi} = (1 - f_B) \cdot \sigma_{\text{J}/\psi}.$$

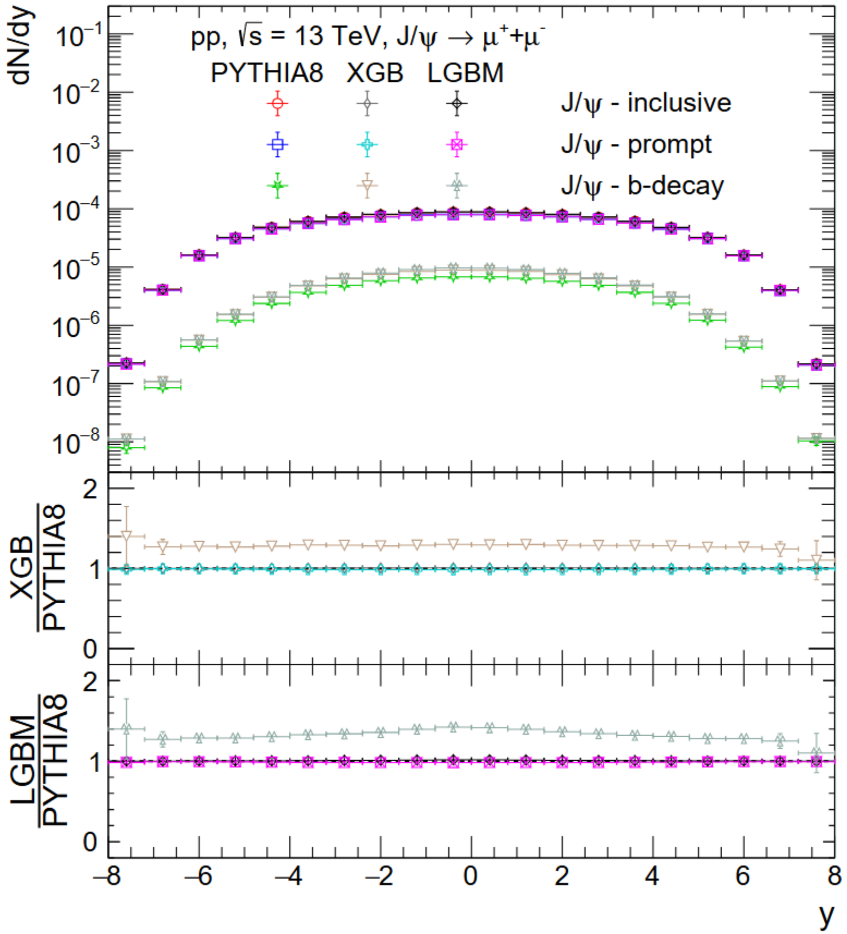
- f_B is the fraction of the non-prompt production (b-hadron decays)
- f_B increases with increase in p_T → The b-hadron production is favoured towards higher p_T
- PYTHIA8 underestimates the experimental data following a similar trend
- Both XGB and LGBM overestimate PYTHIA8
- As this method does not require fitting, it can be used in both low and high statistics without affecting its efficiency

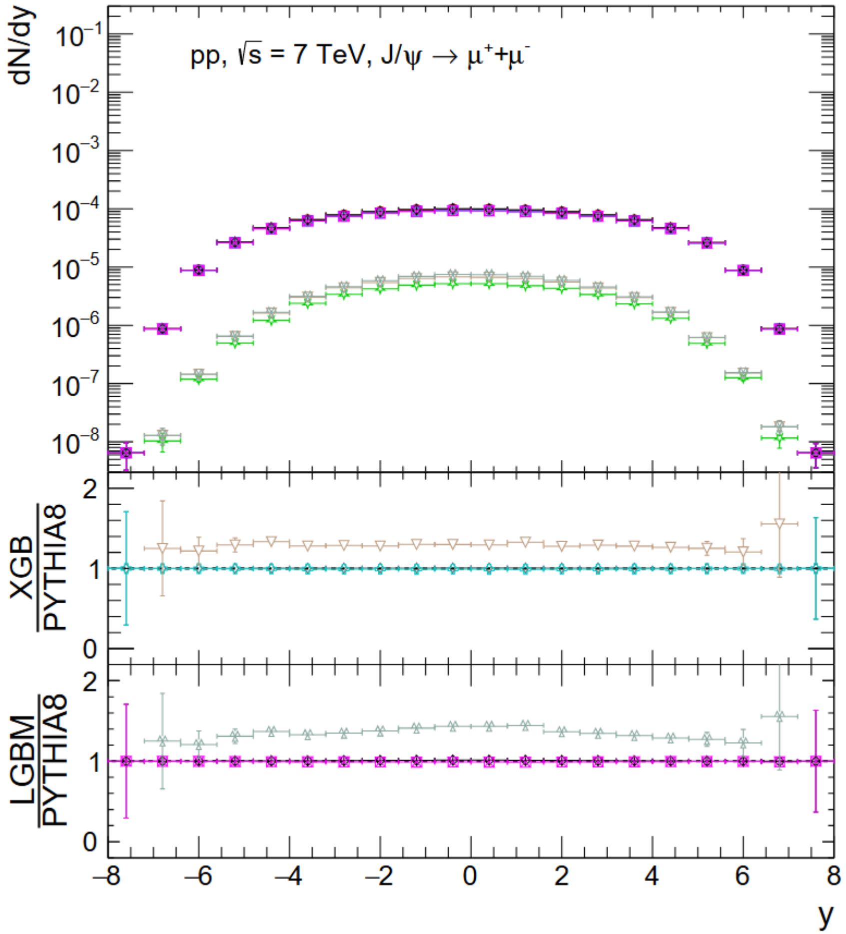
S. Prasad, N. Mallick and R. Sahoo, *Phys. Rev. D 109, 014005 (2024)*

ML Results: Rapidity Spectra J/ψ

- Both XGB and LGBM give accurate predictions for rapidity spectra for inclusive and prompt-J/ψ in pp collisions at $\sqrt{s} = 13$ TeV and 7 TeV

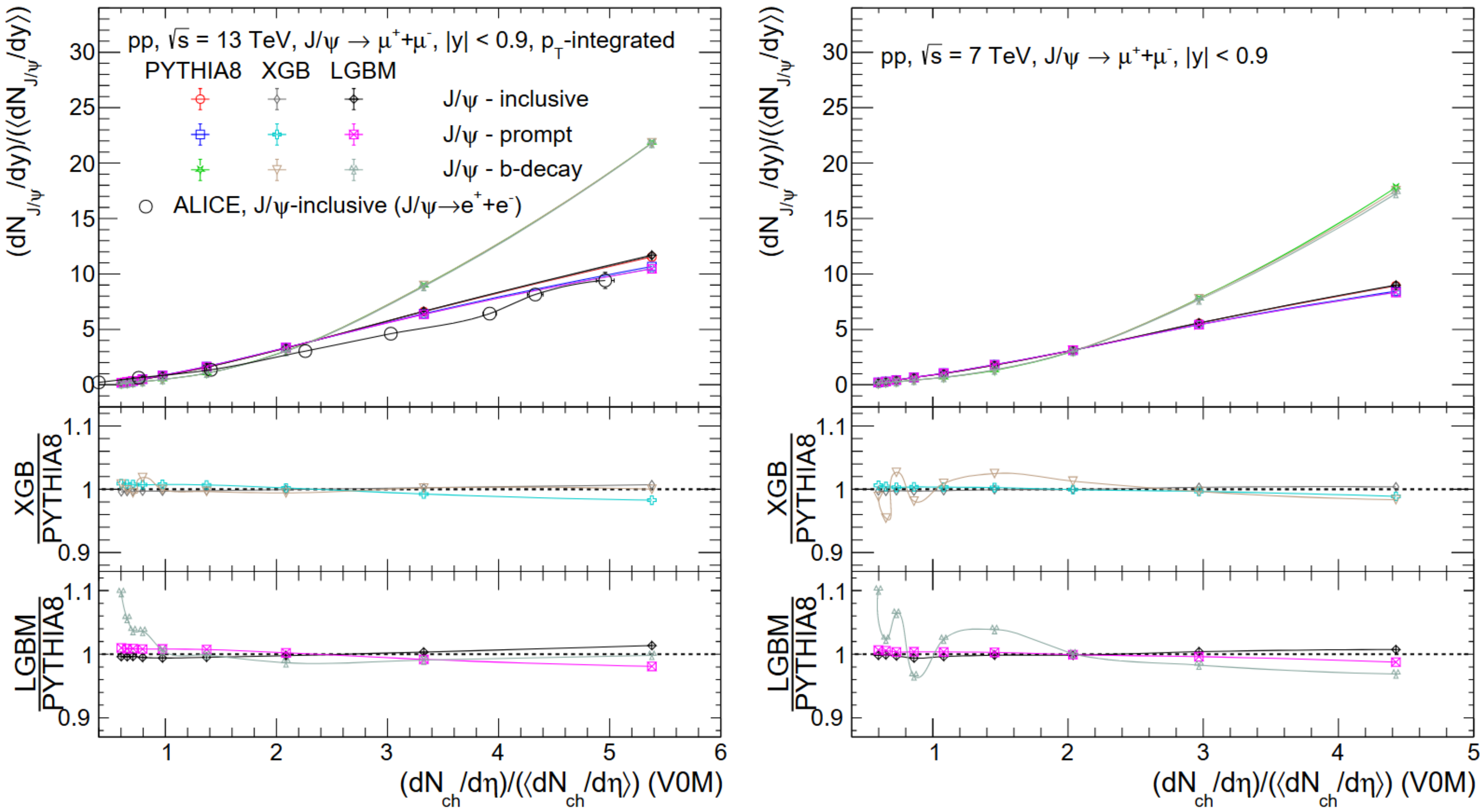
- The ML models overpredict the non-prompt J/ψ throughout rapidity region for both the collision energies





S. Prasad, N. Mallick and R. Sahoo, *Phys. Rev. D 109, 014005 (2024)*

ML Results: Normalized J/ψ Yields



- The normalised yield for inclusive J/ψ from PYTHIA8 matches qualitatively with the ALICE results
- Both XGB and LGBM reproduce the PYTHIA8 results very precisely for inclusive and prompt J/ψ
- The predictions for non-prompt J/ψ from both XGB and LGBM matches PYTHIA8 findings within 10% uncertainty

S. Prasad, N. Mallick and R. Sahoo, *Phys. Rev. D 109, 014005 (2024)*

ML Results J/ψ : Summary

- We have used BDT based ML models such as XGBoost and LGBM to segregate the prompt, non-prompt and inclusive J/ψ production in pp collisions at $\sqrt{s} = 13$ TeV
- The models the parameters such as, pseudo-proper decay length ($c\tau$), invariant mass ($m_{\mu\mu}$), transverse momentum ($p_{T,\mu\mu}$), pseudorapidity ($\eta_{\mu\mu}$) of the dimuons as the input, which are accessible in the experiments
- The model almost achieves 99% overall accuracy
- The estimations for the prompt and inclusive J/ψ from the ML models match with the PYTHIA8 for the inclusive and non-prompt J/ψ
- Using these models, track label identification is possible, and it avoids the necessity of fit for the identification

Open Charms: D^0

PHYSICAL REVIEW D **VOL..XX,** 000000 (XXXX)

Machine learning-based study of open-charm hadrons in proton-proton collisions at the Large Hadron Collider

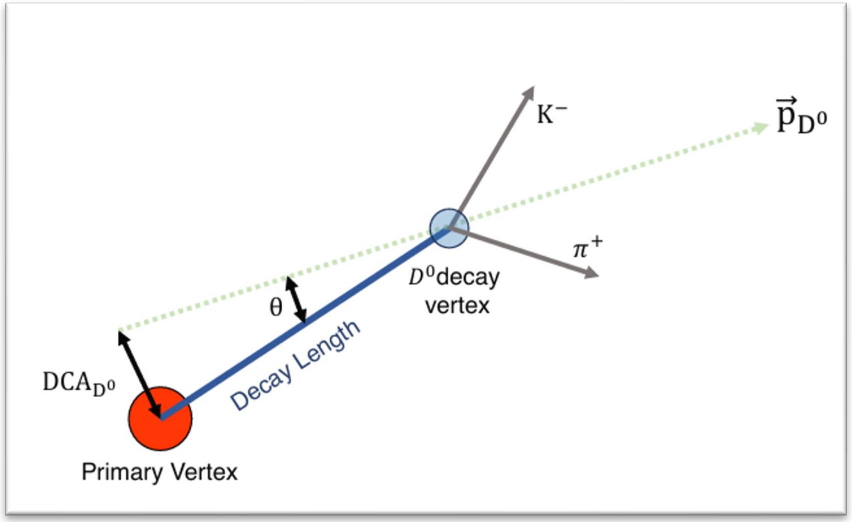
Kangkan Goswami, Suraj Prasad, Neelkamal Mallick, and Raghunath Sahoo*
Department of Physics, Indian Institute of Technology Indore, Simrol, Indore 453552, India

Gagan B. Mohanty
Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India

(Received 16 April 2024; accepted 2 July 2024)

In proton-proton and heavy-ion collisions, the study of charm hadrons plays a pivotal role in understanding the QCD medium and provides an undisputed testing ground for the theory of strong interaction, as they are mostly produced in the early stages of collisions via hard partonic interactions. The lightest open charm, D^0 meson ($c\bar{u}$), can originate from two separate sources. The prompt D^0 originates from either direct charm production or the decay of excited open charm states, while the nonprompt stems from the decay of beauty hadrons. In this paper, using different machine learning (ML) algorithms such as XGBoost, CatBoost, and Random Forest, an attempt has been made to segregate the prompt and nonprompt production modes of the D^0 meson signal from its background. The ML models are trained using the invariant mass through its hadronic decay channel, i.e., $D^0 \rightarrow \pi^+ K^-$, pseudoproper time, pseudoproper decay length, and distance of closest approach of D^0 meson, using PYTHIA8 simulated pp collisions at $\sqrt{s} = 13$ TeV. The ML models used in this analysis are found to retain the pseudorapidity, transverse momentum, and collision energy dependence. In addition, we report the ratio of nonprompt to prompt D^0 yield, the self-normalized yield of prompt and nonprompt D^0, and explore the charmonium, J/ψ to open charm, D^0 yield ratio as a function of transverse momenta and normalized multiplicity. The observables studied in this paper are well predicted by all the ML models compared to the simulation.

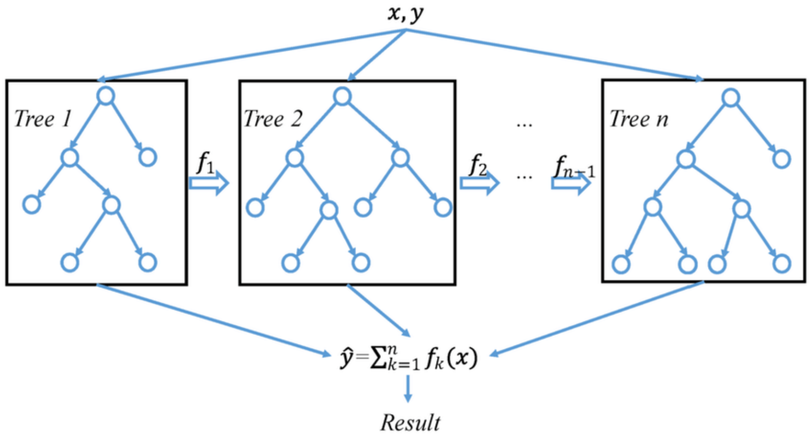
DOI:



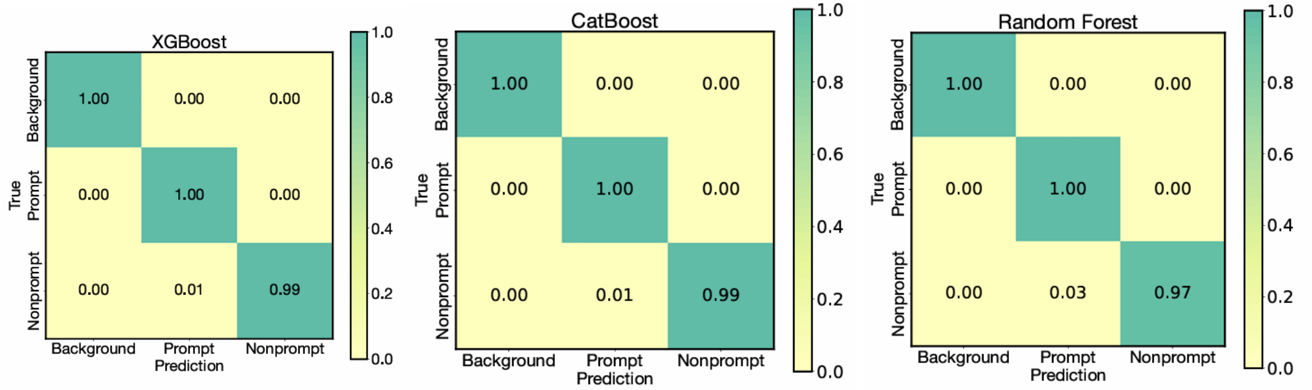
Phys. Rev. D (2024: In Press)
arXiv: 2404.09839

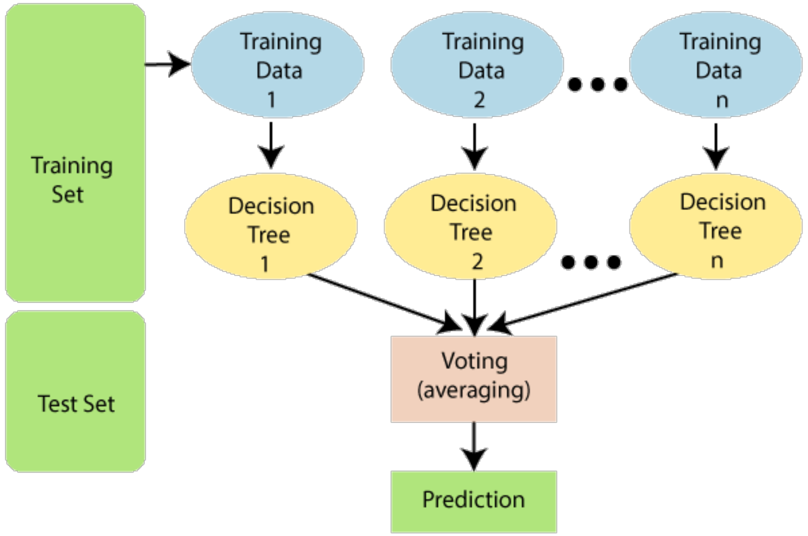
ML Techniques

- Extreme Gradient Boost (XGBoost): Combines the predictions of multiple weak models to produce a stronger model.
- Categorical Boosting (CatBoost): Similar working principle as XGBoost but faster and more efficient when working with categorical data.
- Random Forest: In a Random Forest classifier, multiple decision trees are created, each on a different subset of the data. Each tree gets a vote on the class label for a new instance. The class that gets the most votes is chosen as the final prediction.



XGBoost and CatBoost Architecture



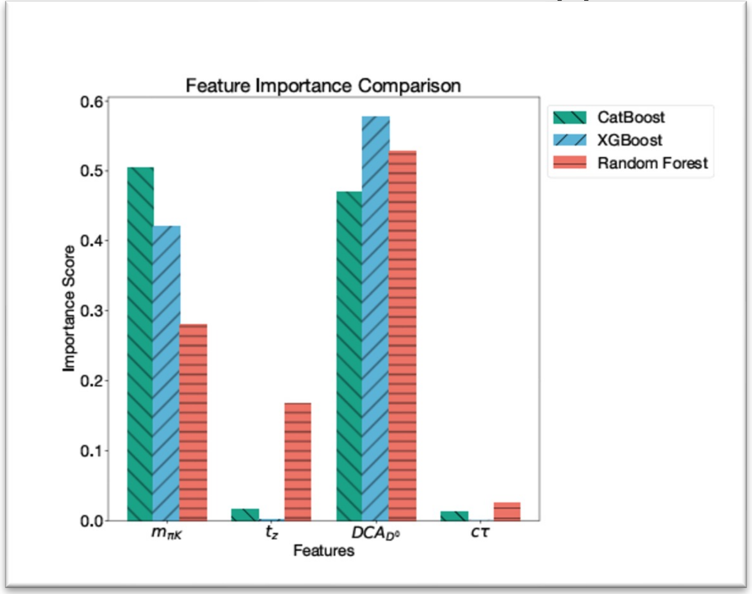


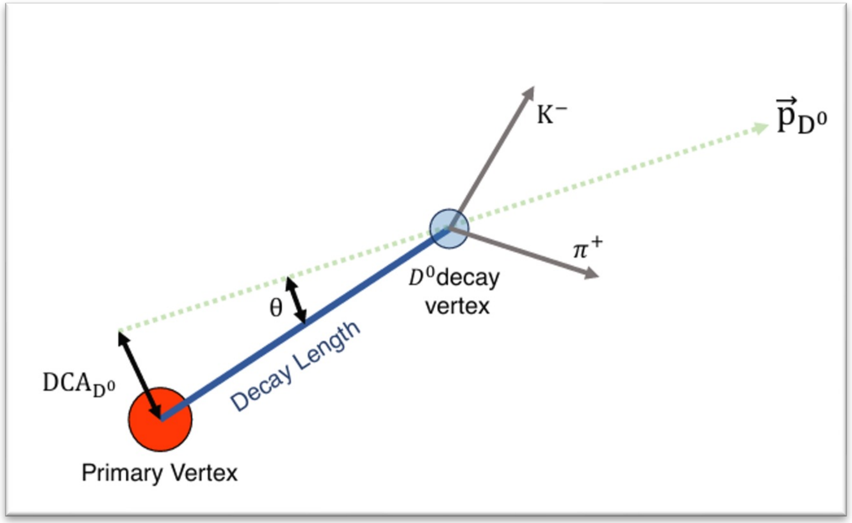
Random Forest Classifier Architecture

Training ML Models

Input Variables

1. Invariant Mass
2. The pseudo-proper time: $$t_z = \frac{(z_{D^0} - z_{PV}) \times m_{D^0}}{p_z}$$
3. Pseudo-proper decay length: $$c\tau = \frac{c m_{D^0} \vec{L}.\overrightarrow{p_T}}{p_T^2}$$
4. Distance of closest approach: $$DCA_{D^0} = L \times \sin\theta$$



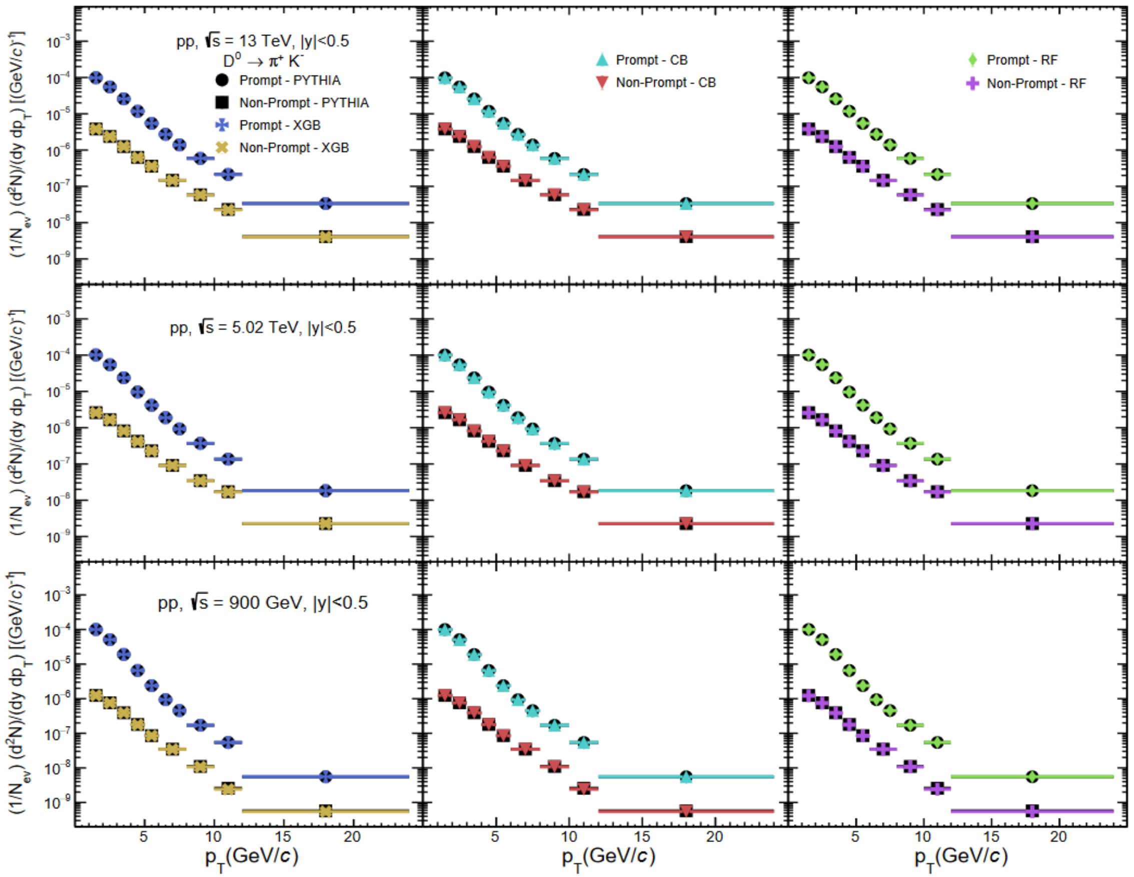


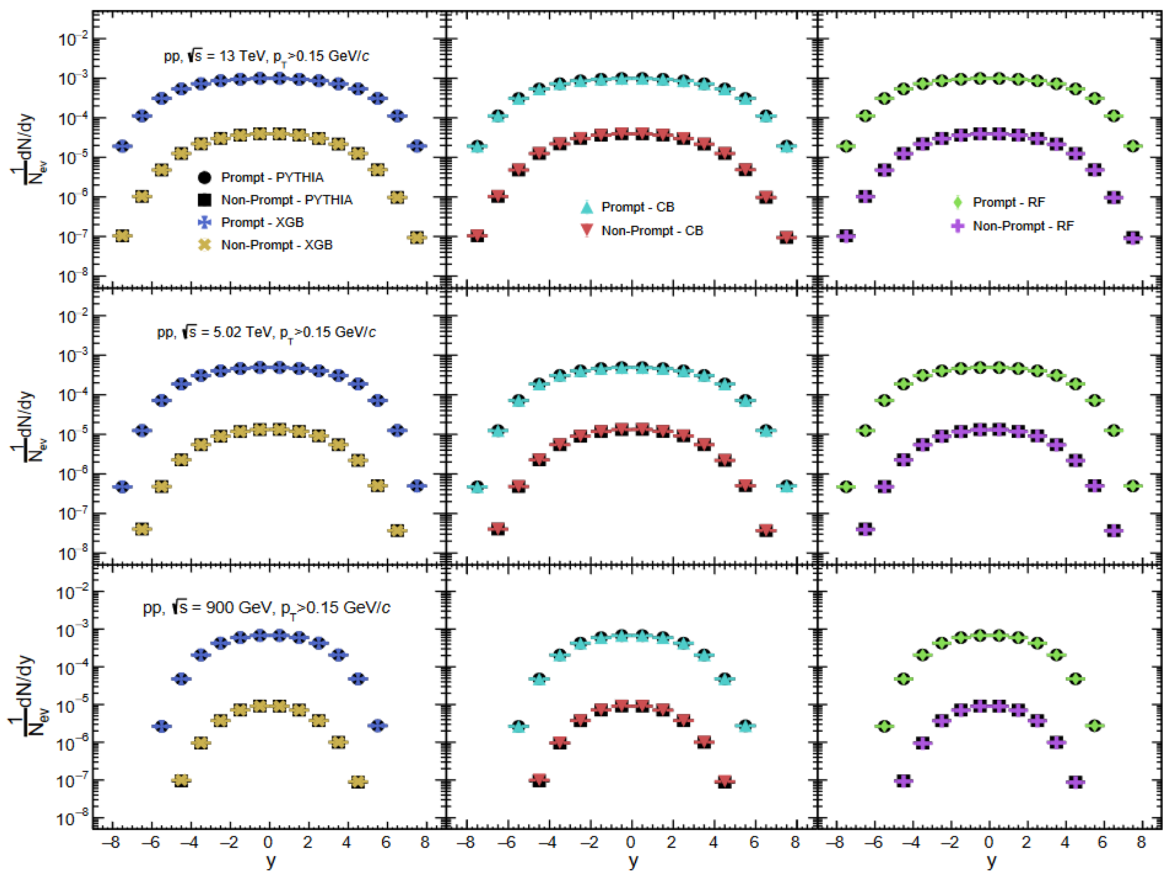
$\vec{L}$ is the vector pointing from the primary vertex towards D^0 decay vertex, i.e. $\vec{L} = \vec{V} - \vec{S}$. $\vec{V}$ is the position of the primary vertex and $\vec{S}$ is the position of the D^0 decay vertex given by,

$$S_i = \frac{(t_1 + d_{i,1}m_1/p_{i,1}) - (t_2 + d_{i,2}m_2/p_{i,2})}{m_1/p_{i,1} - m_2/p_{i,2}}$$

Results

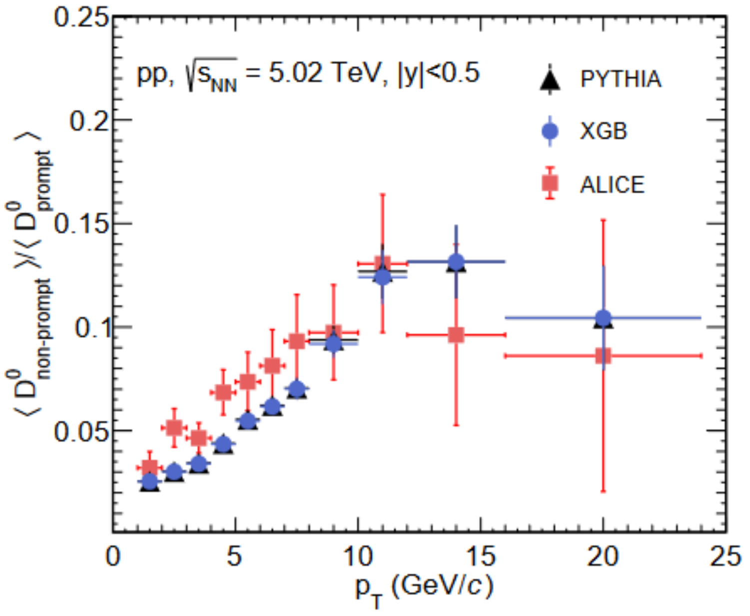
- Transverse momenta and rapidity spectra
- Trained the ML models only at one centre-of-mass energy: $\sqrt{s} = 13$ TeV
- Predicted normalized D^0 meson yield at $\sqrt{s} = 13$ TeV, 5.02 TeV and 900 GeV

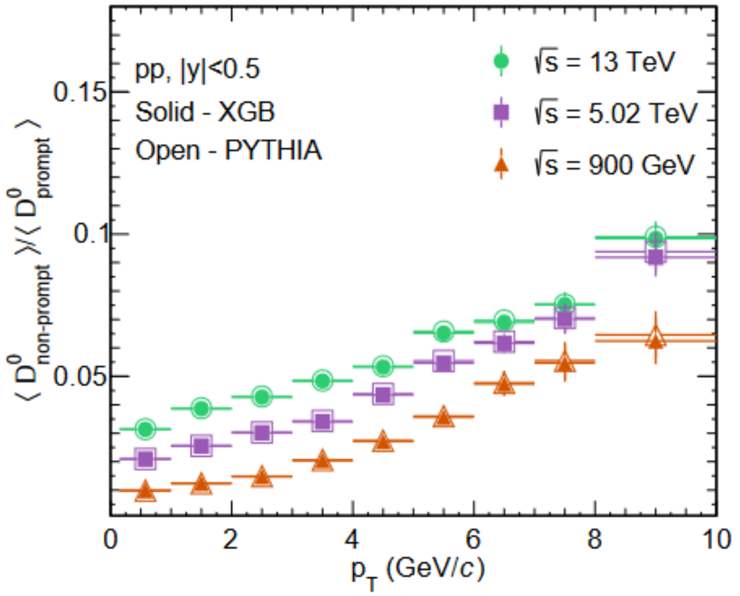




Results

- ❖ Predicted the non-prompt to prompt ratio at 5.02 TeV and compared with ALICE data
- ❖ We observe a linear increase in the ratio with p_T. However, the linear trend holds true only up to a certain p_T range.





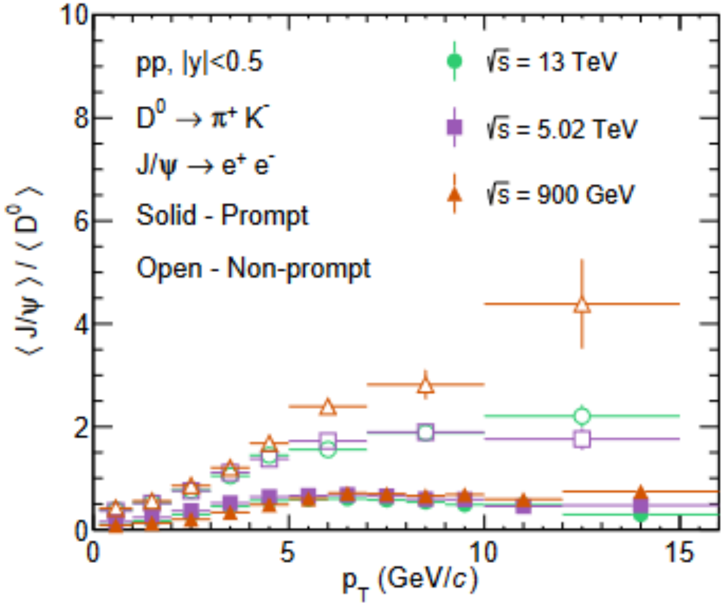
- ➢ One can clearly notice the increase in the ratio with increasing p_T across all the collision energies.
- ➢ We observe an energy-dependent hierarchy in the ratio.

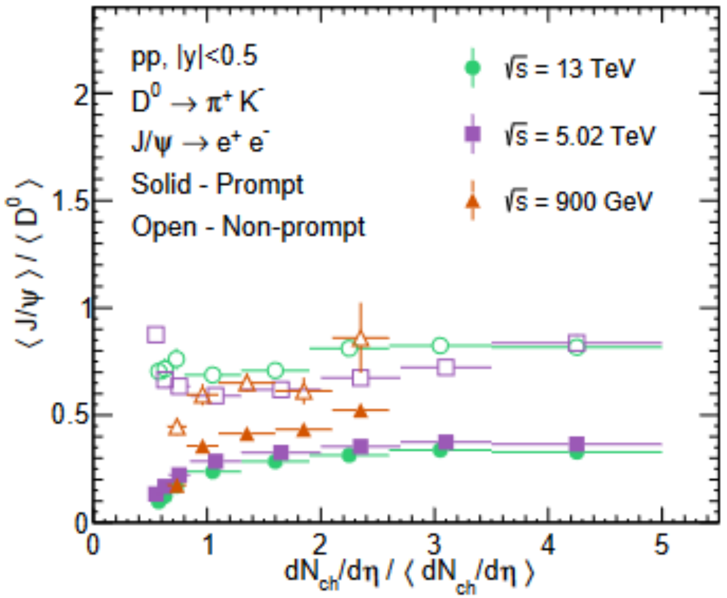
Results

This ratio tells us about the production dynamics of charmonium state wrt open-charmed hadrons.

➢ The ratio shows a similar rising trend as a function of p_T, upto 5 GeV.

➢ The prompt J/ψ to D^0 ratio is always less than 1. However, the non-prompt ratio rises above 1, indicating a higher contribution of beauty hadrons towards J/ψ states.





➢ For the prompt ratio, there is a slight increase and then it follows a flat trend.

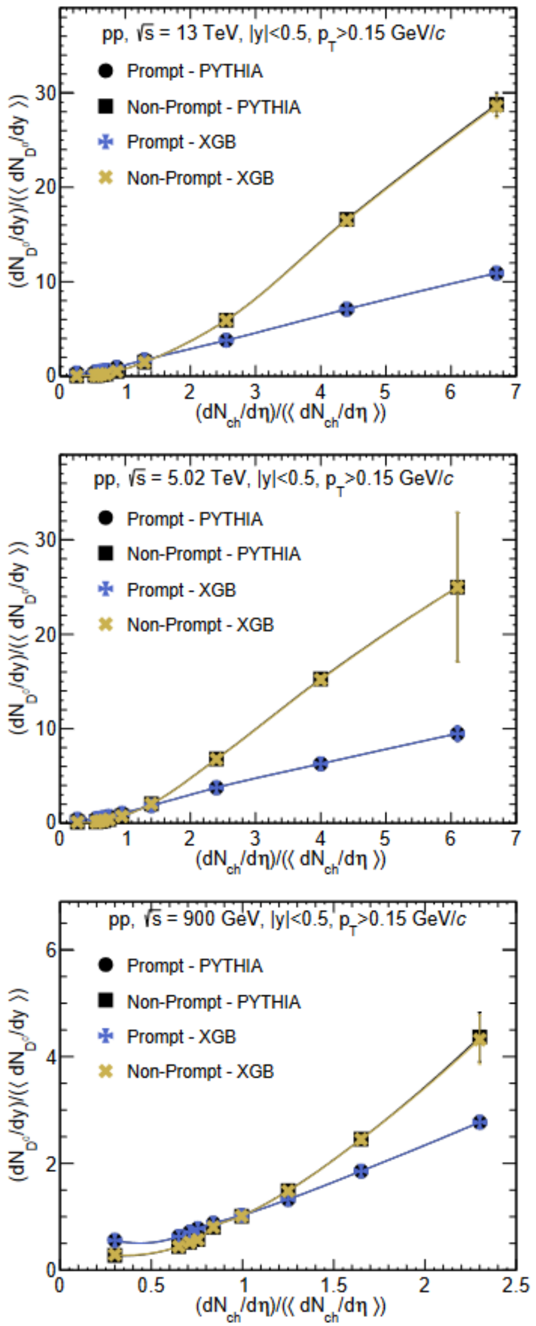
➢ However, the non-prompt ratio is independent of the charged particle multiplicity.

Results

- Self-normalized yield of prompt and non-prompt D^0 meson as a function of charged particle multiplicity.

- A linear rise can be seen for the prompt D^0 meson.

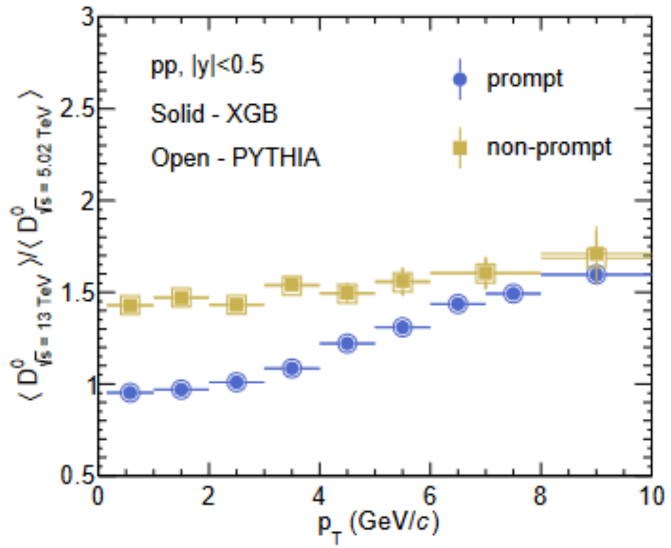
- A non-linear rise is observed for non-prompt D^0 meson.

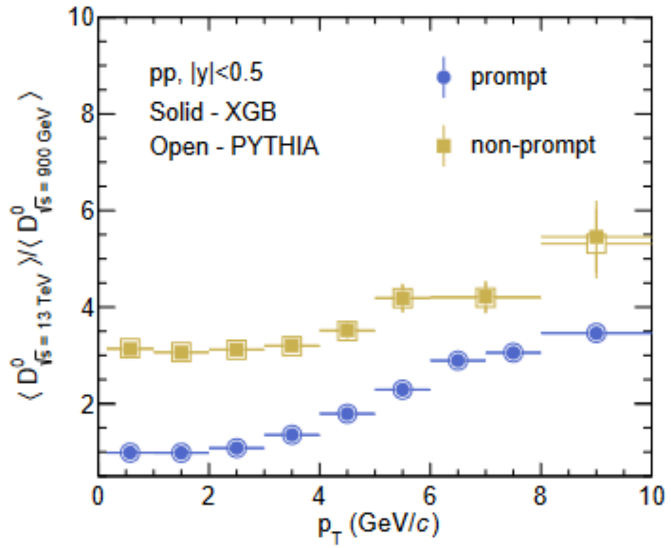


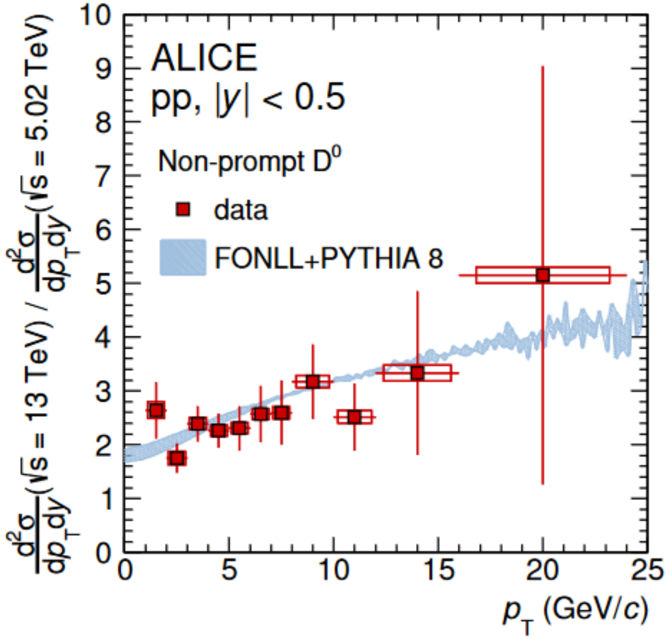
Results

- Role of centre-of-mass energy in D^0 meson production.
- In both plots, a clear increase in the ratio for prompt D^0 meson can be observed.

- In the top panel, a flat trend can be seen for the non-prompt case. Similar results have been observed in ALICE.
- In the bottom panel, the same ratio tends to increase slightly due to the significant difference in energies.
- A higher value of non-prompt ratio indicates more abundant production of beauty hadrons at higher $\sqrt{s}$.

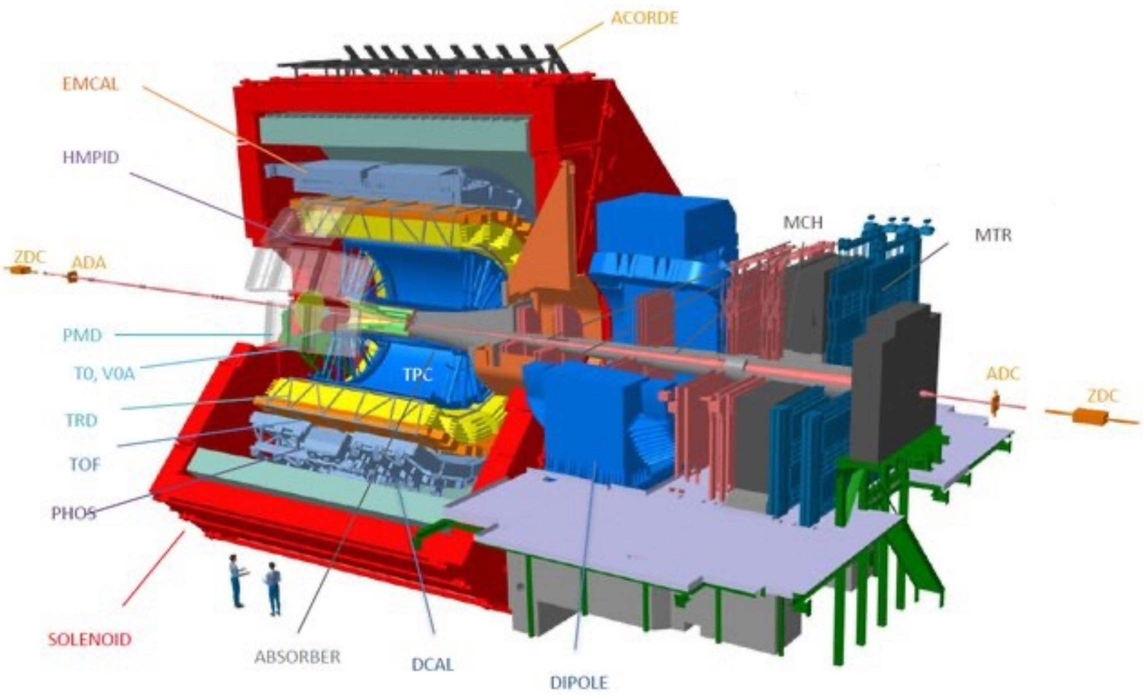




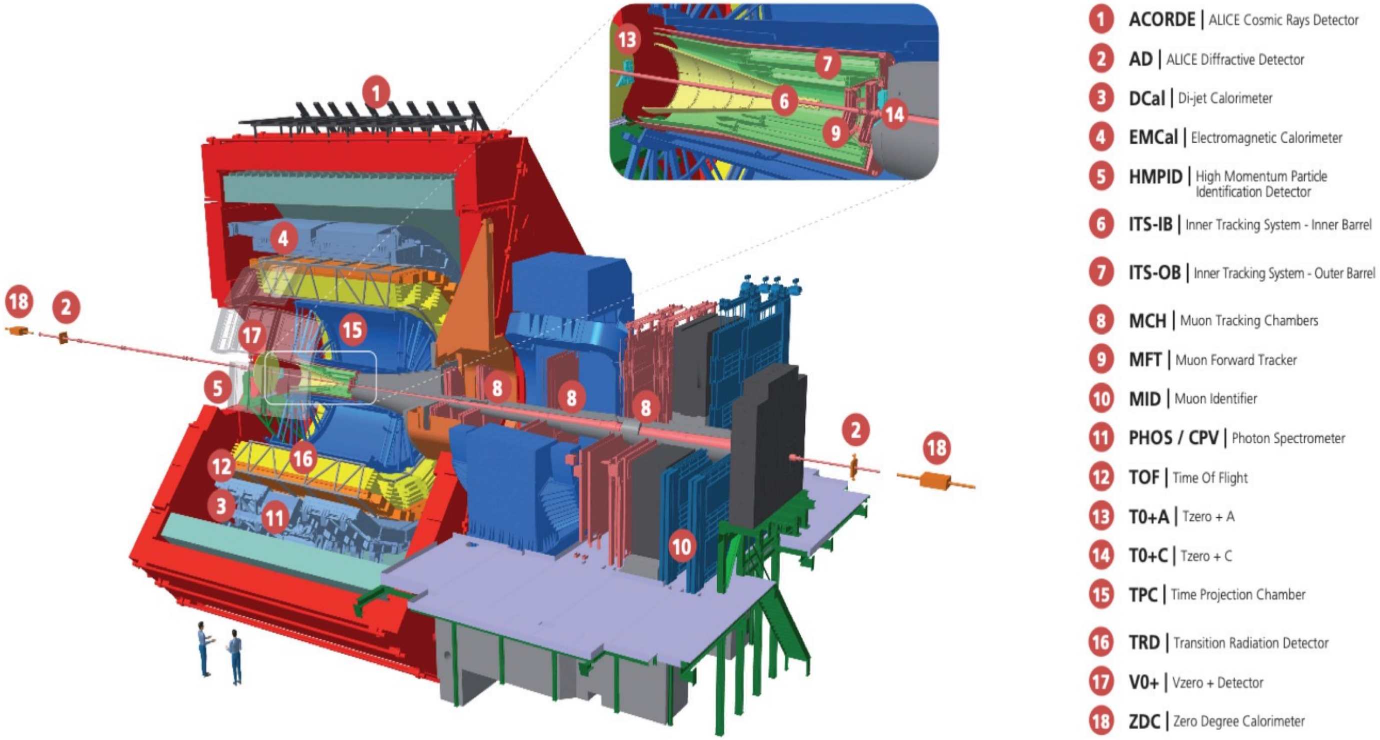


Source: ALICE Collaboration, arXiv:2402.16417

ALICE in RUN 3 with Muon Forward Tracker (MFT) will enhance the charmonia measurement Capability while distinguishing prompt vs non-prompt J/ψ.

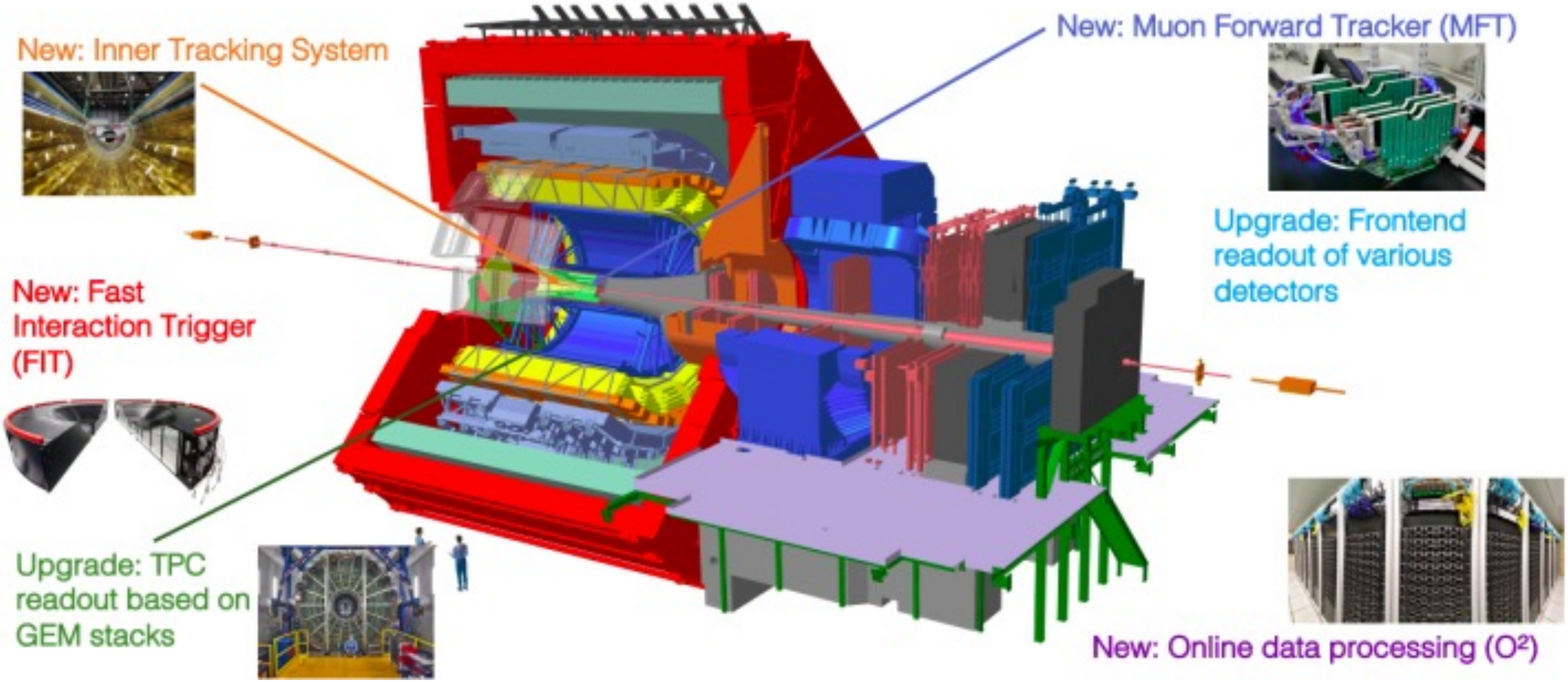


ALICE RUN 2 Setup

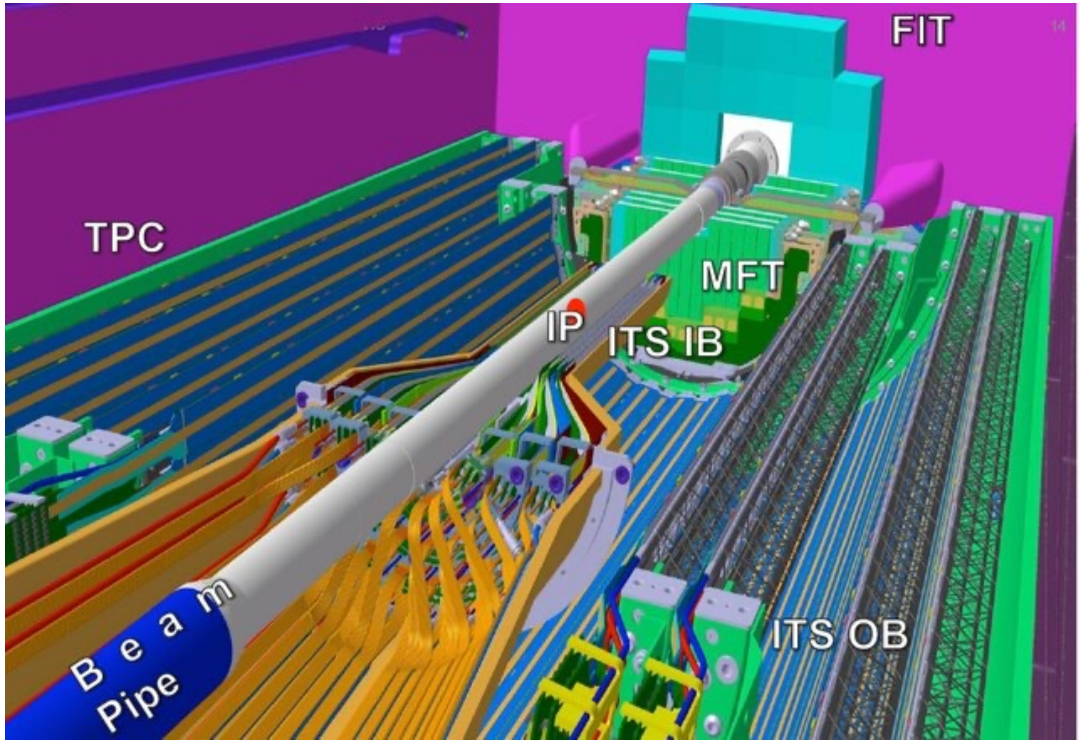


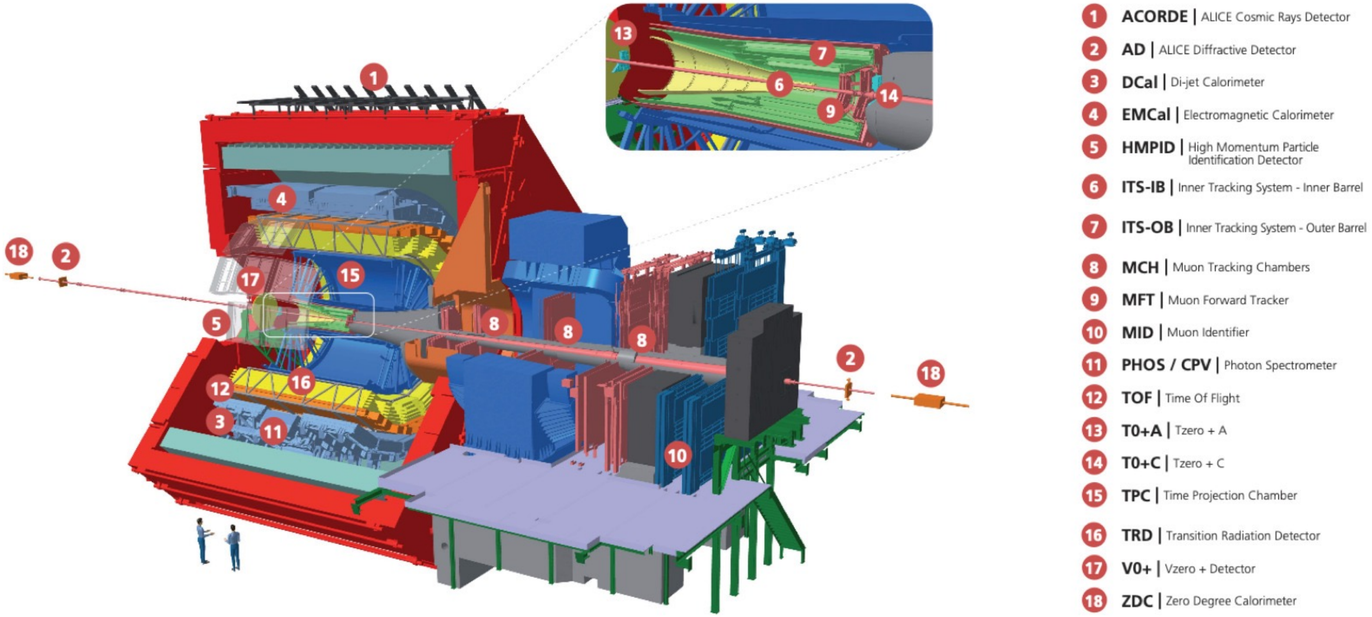
ALICE RUN 3 Setup

ALICE RUN 3 Upgrades



ALICE in RUN 3 with Muon Forward Tracker (MFT) will enhance the charmonia measurement Capability while distinguishing prompt vs non-prompt J/ψ.





ALICE RUN 3 Setup showing MFT

ALICE RUN 3 and beyond stays interesting given HF measurements at the LHC!

Bon Voyage

Thank you