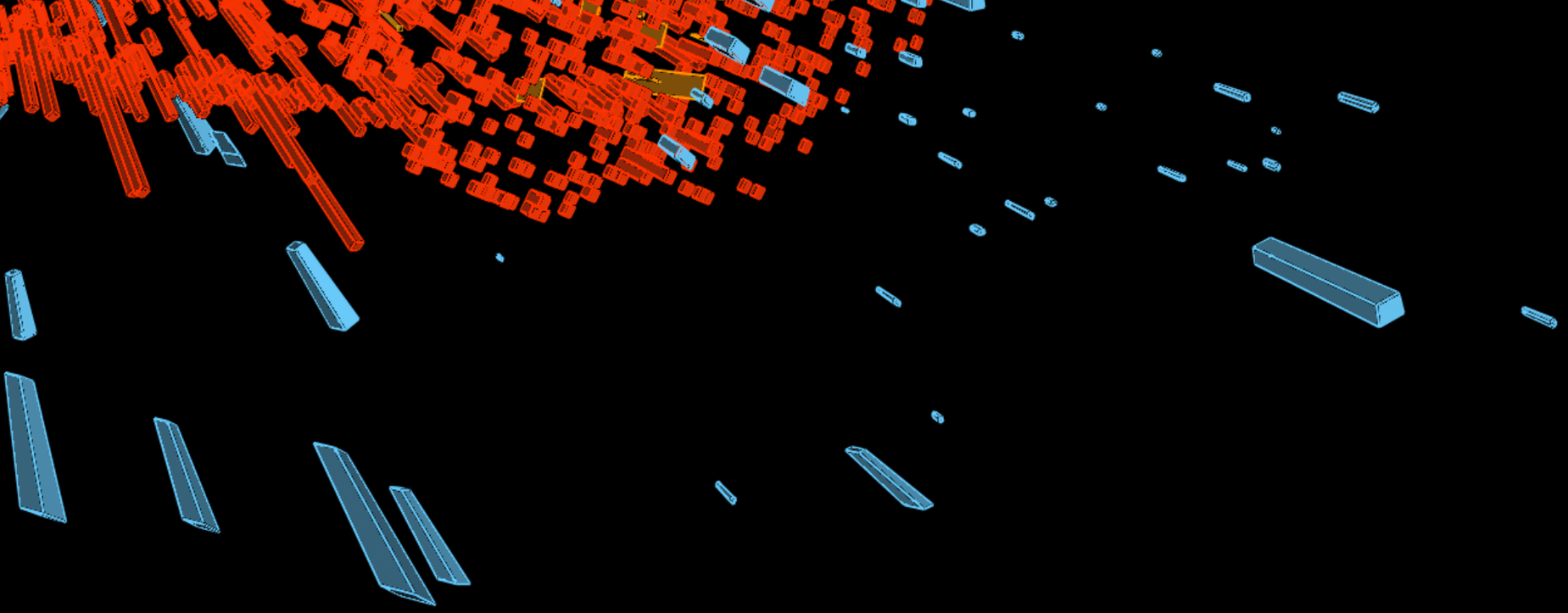
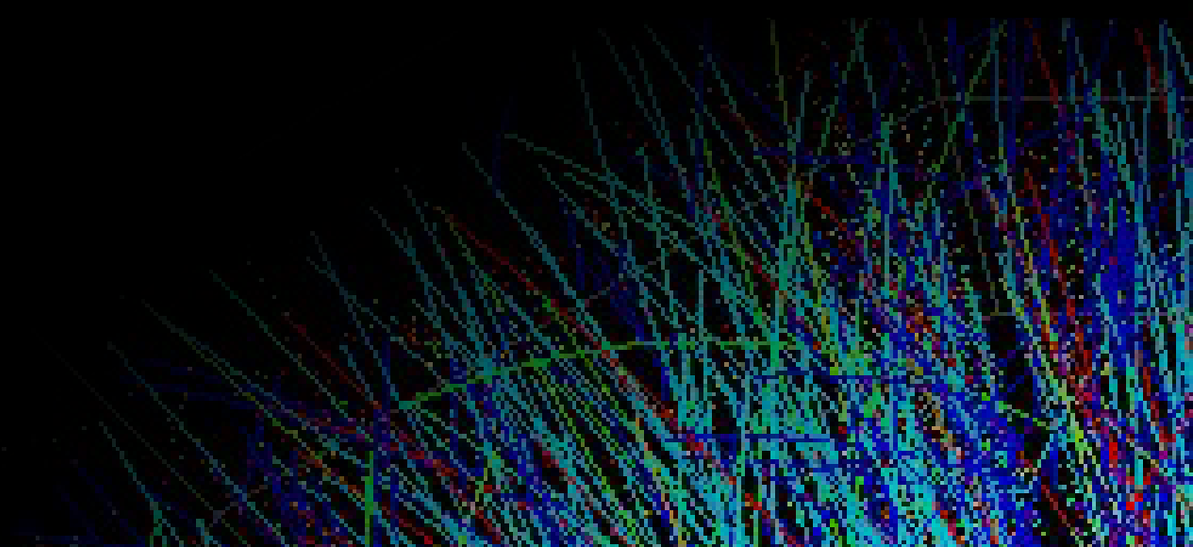


Duke
UNIVERSITY

A data-driven approach to quantifying the shear viscosity of nature's most ideal liquid

Steffen A. Bass



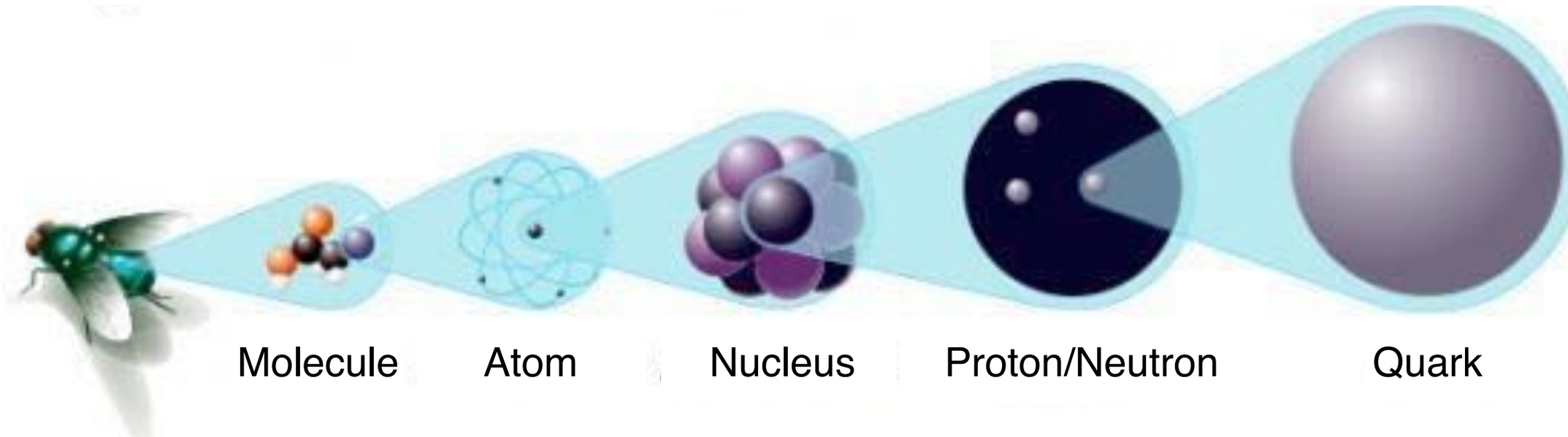
http://www.facebook.com/DukeQCD

@Steffen_Bass

Introduction:

- Phase diagram of matter
- Quark-Gluon-Plasma

Quarks & Gluons: Elementary Building-Blocks of Matter



Elementary Particles:

- 12 elementary building blocks of nature (plus anti-particles)
- only need three for creation of ordinary matter **(u, d, e)**
- strong force mediates the interaction between quarks via exchange of gluons

Phases of Matter

by adding/removing heat, phase of matter can be changed between solid, liquid and gaseous

liquid

solid

Pressure plays an important role for the value of the transition temperature between the phases

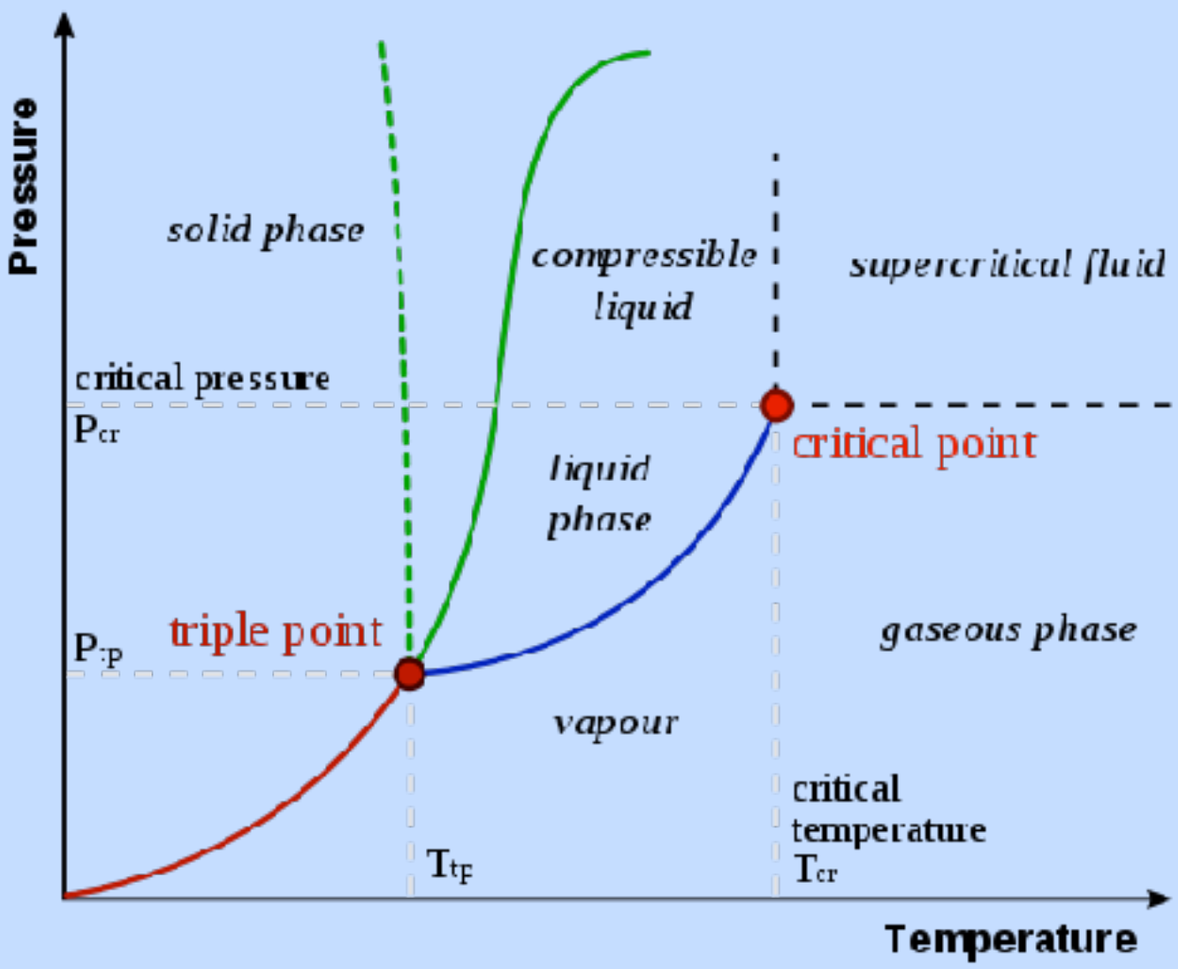
boiling temperature:

- sea level: 100 °C
- Mt. Everest: 71 °C

Phase Diagram

Ordinary Matter:

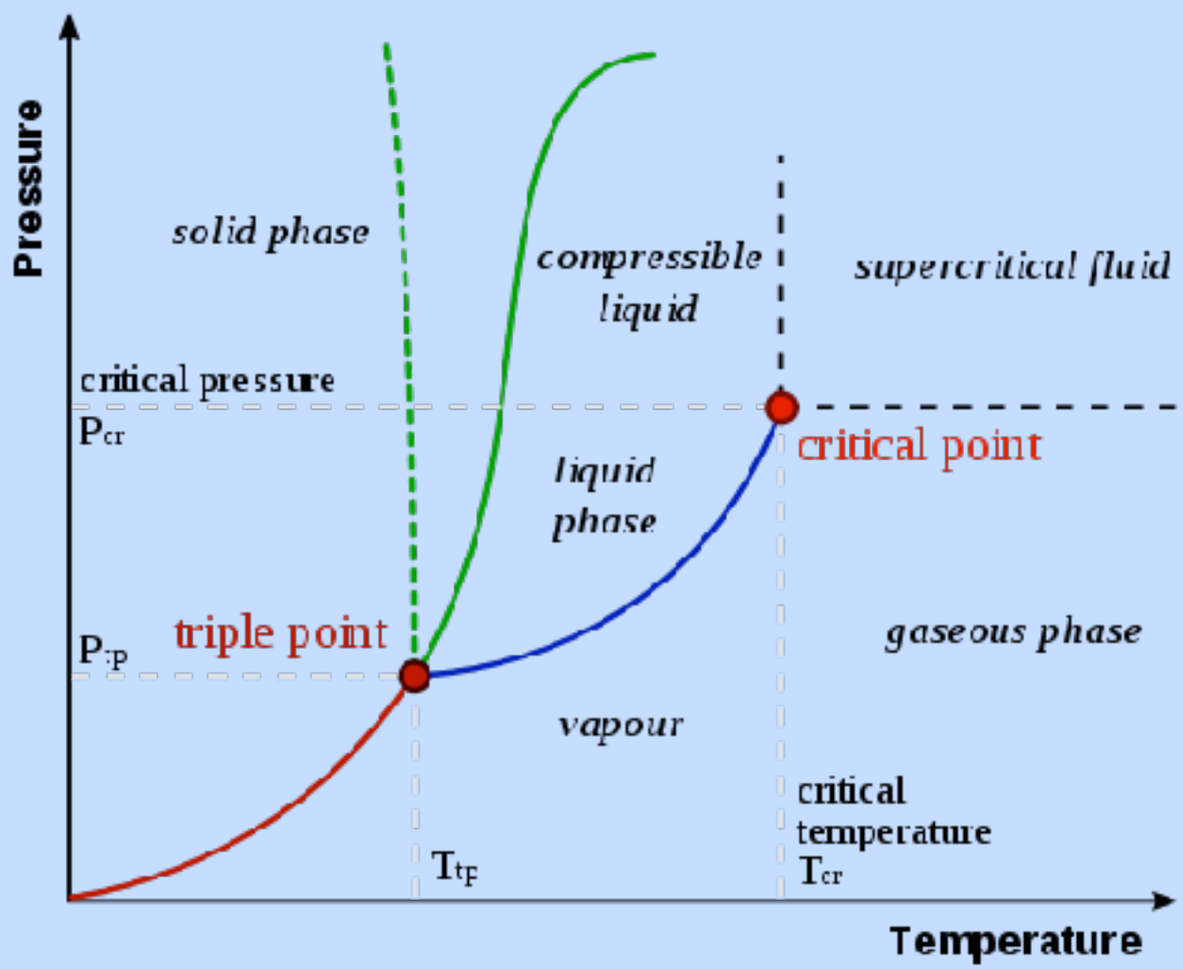
- phases determined by (electro-magnetic) interaction
- apply heat & pressure to study phase-diagram



Phase Diagram

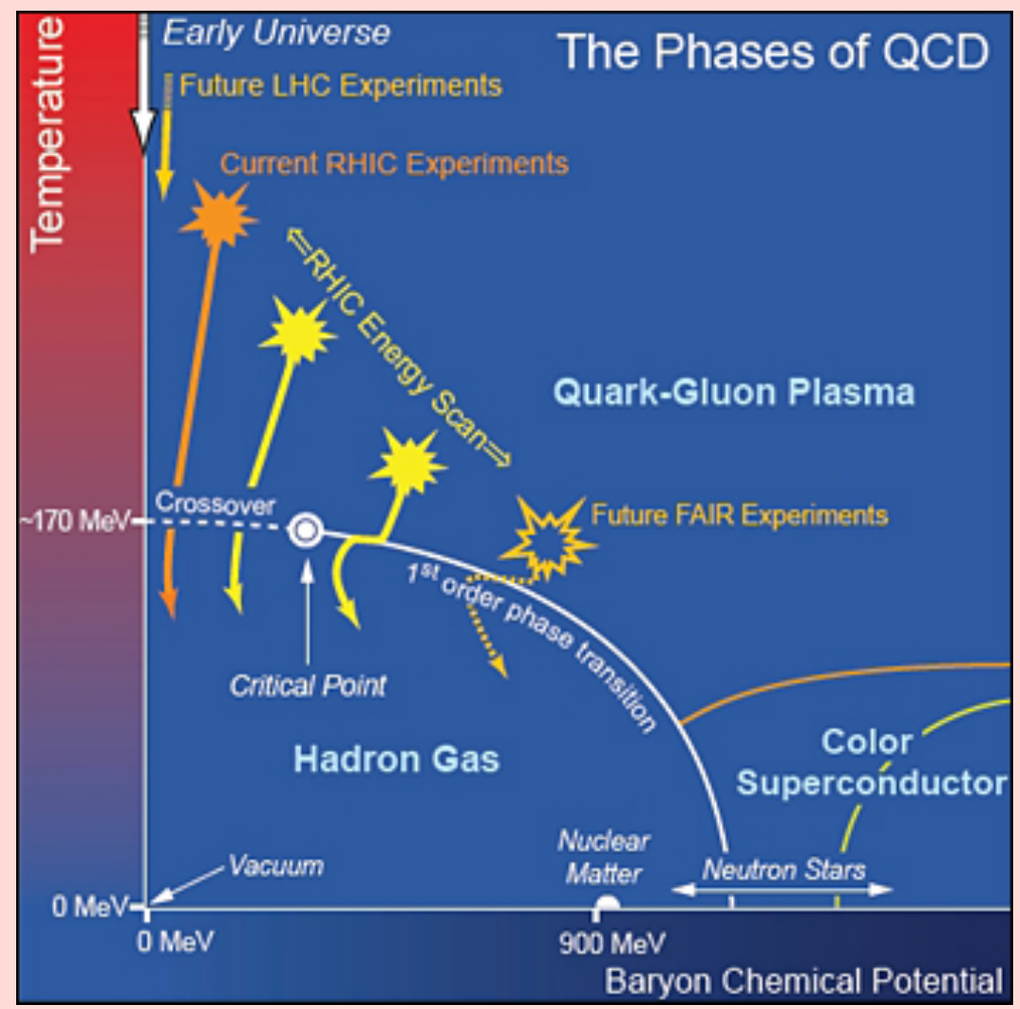
Ordinary Matter:

- phases determined by (electro-magnetic) interaction
- apply heat & pressure to study phase-diagram

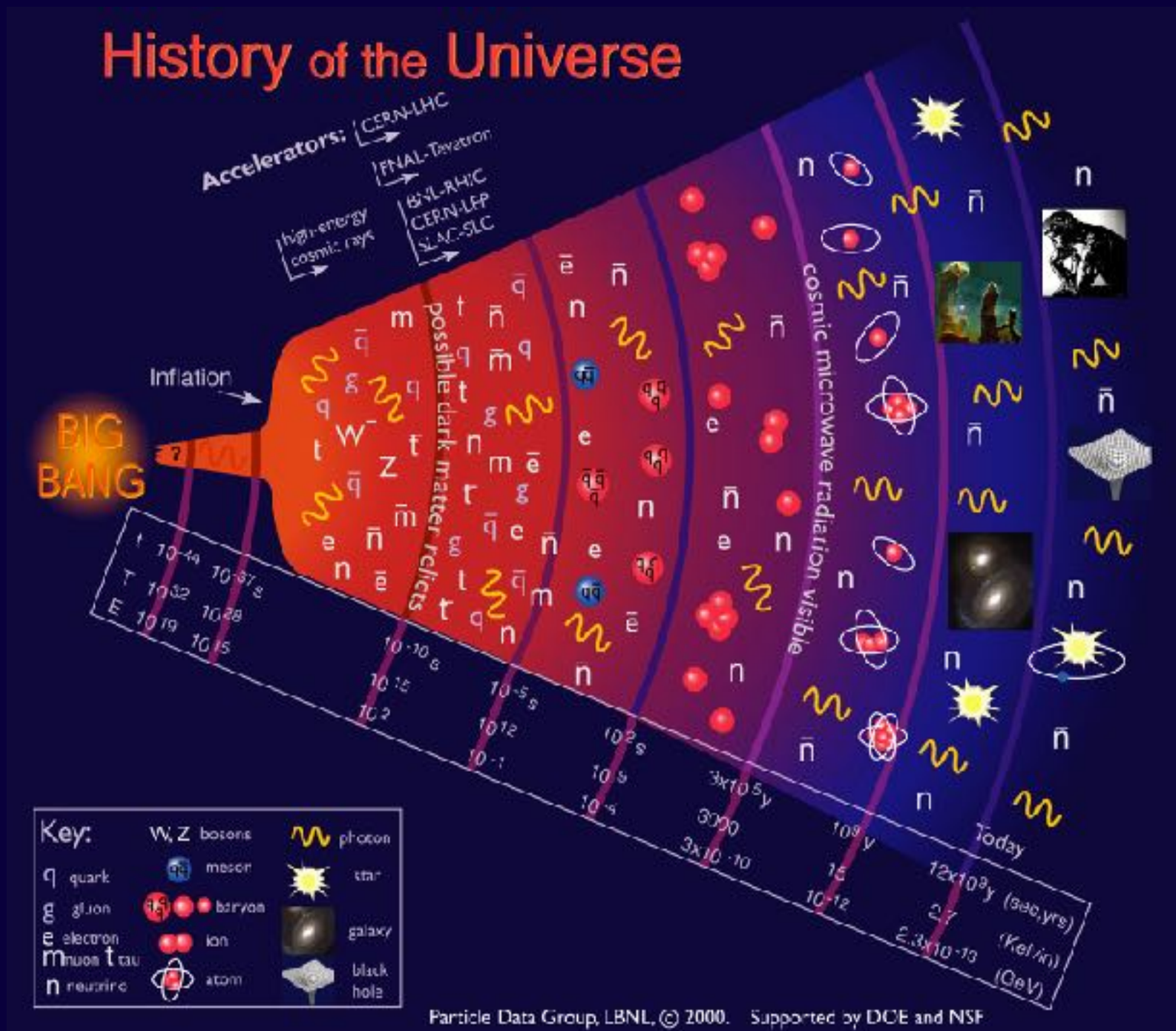


Phases of QCD matter:

- heat & compress QCD matter:
  - collide heavy atomic nuclei
- numerical simulations:
  - solve partition function (Lattice)



QGP and the Early Universe

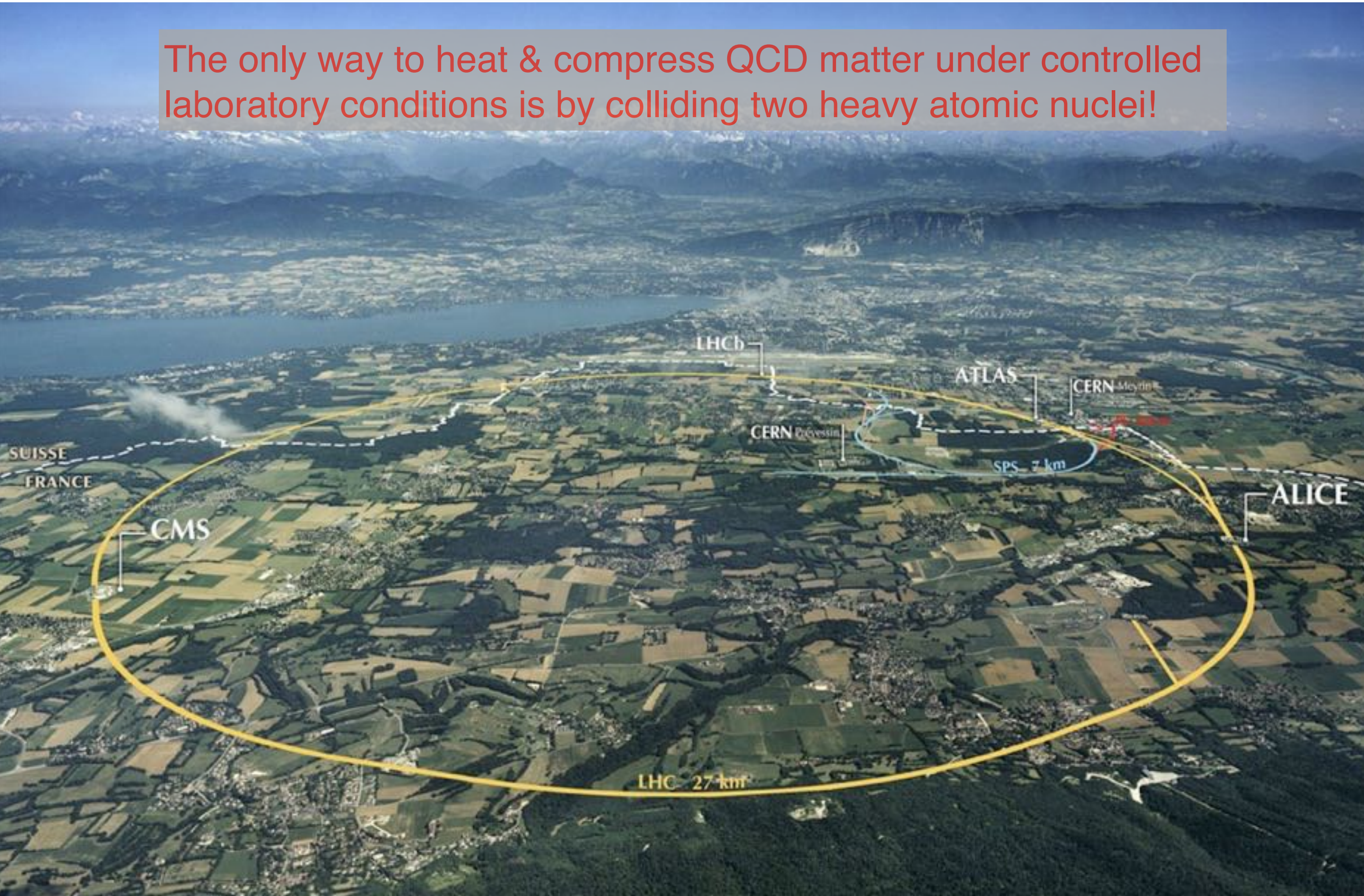


- a few microseconds after the Big Bang the entire Universe was in a QGP state
- compressing & heating nuclear matter allows to investigate the history of the Universe
- the only means of recreating temperatures and densities of the early Universe is by colliding beams of ultra-relativistic heavy-ions

Telescopes for the Early Universe: Heavy-Ion Collider Facilities

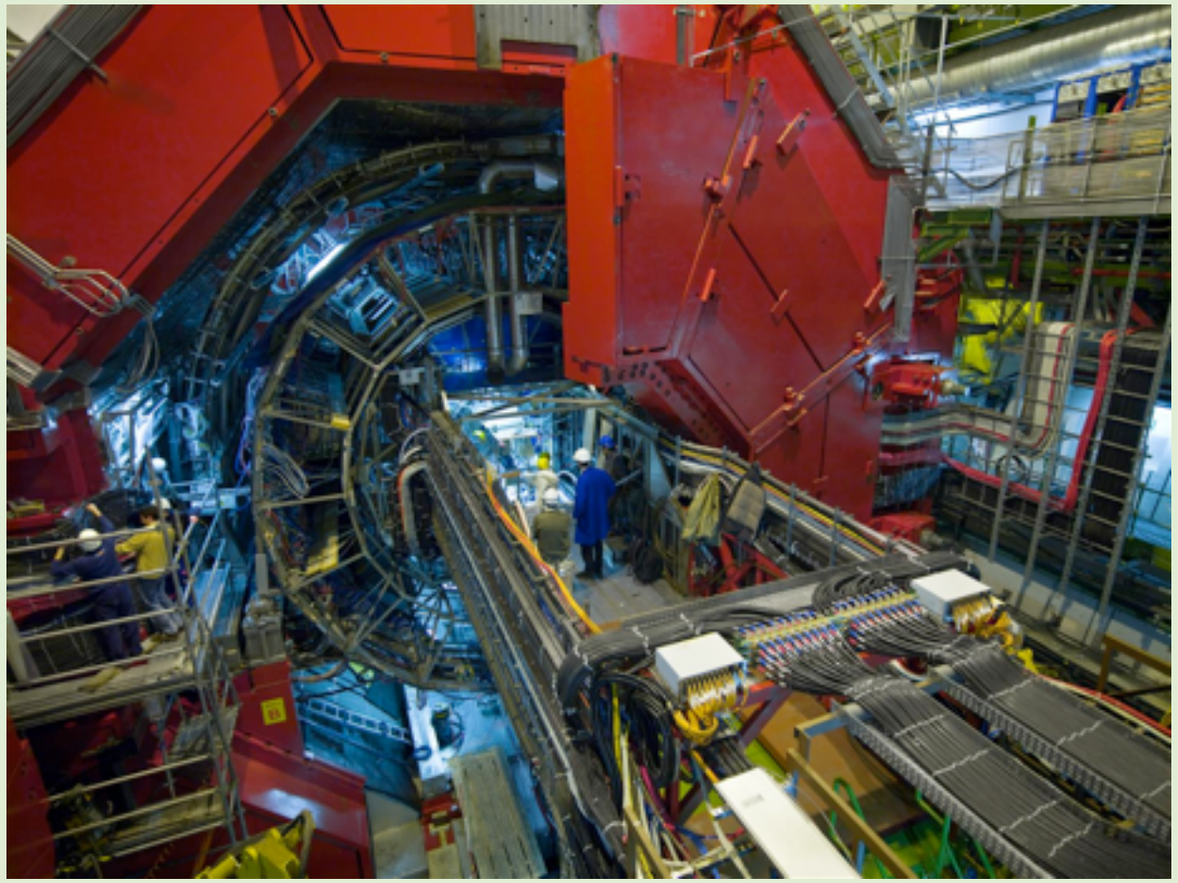
Heating & Compressing QCD Matter

The only way to heat & compress QCD matter under controlled laboratory conditions is by colliding two heavy atomic nuclei!



Heating & Compressing QCD Matter

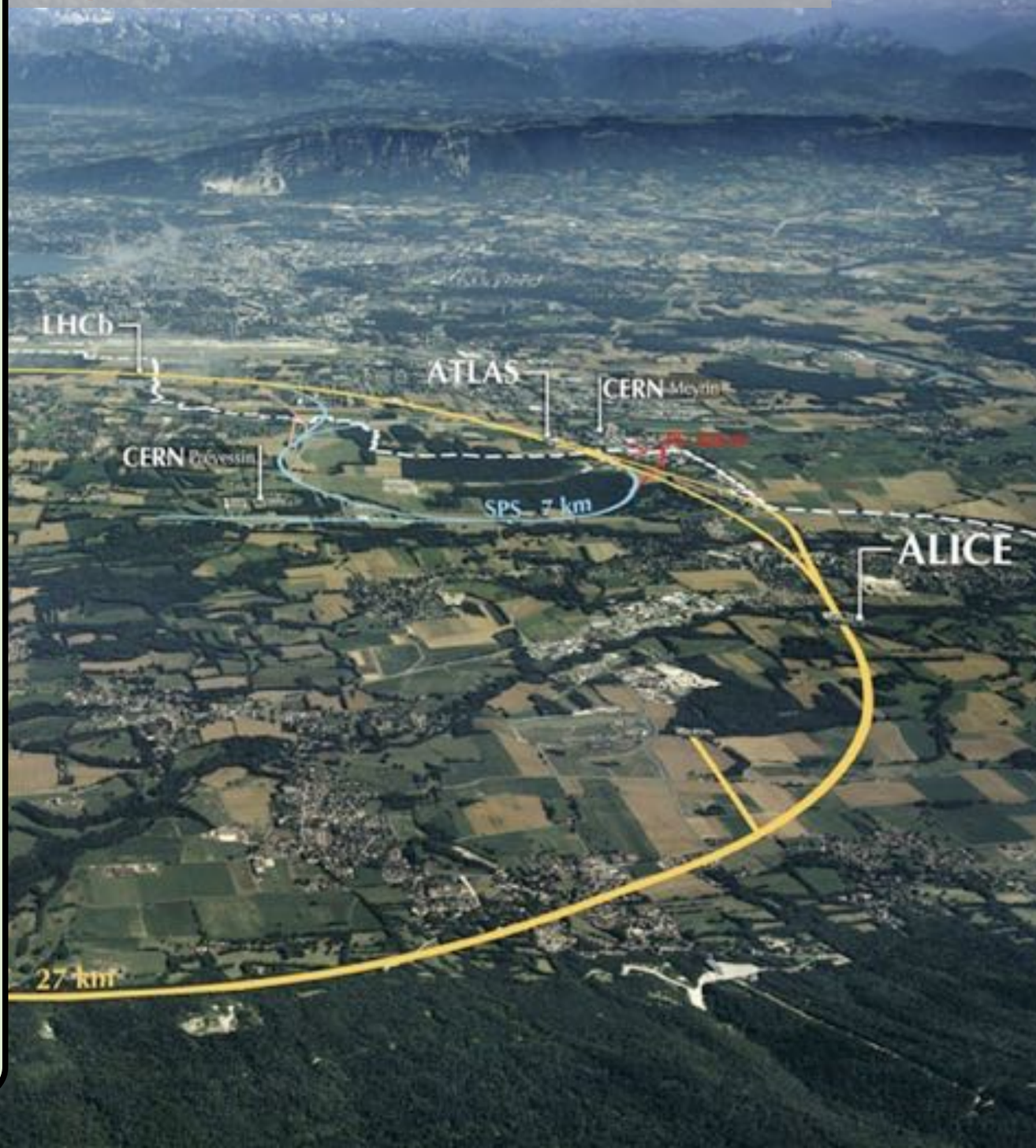
ALICE experiment @ CERN:



- 1000+ scientists from 105+ institutions
- dimensions: 26m long, 16m high, 16m wide
- weight: 10,000 tons

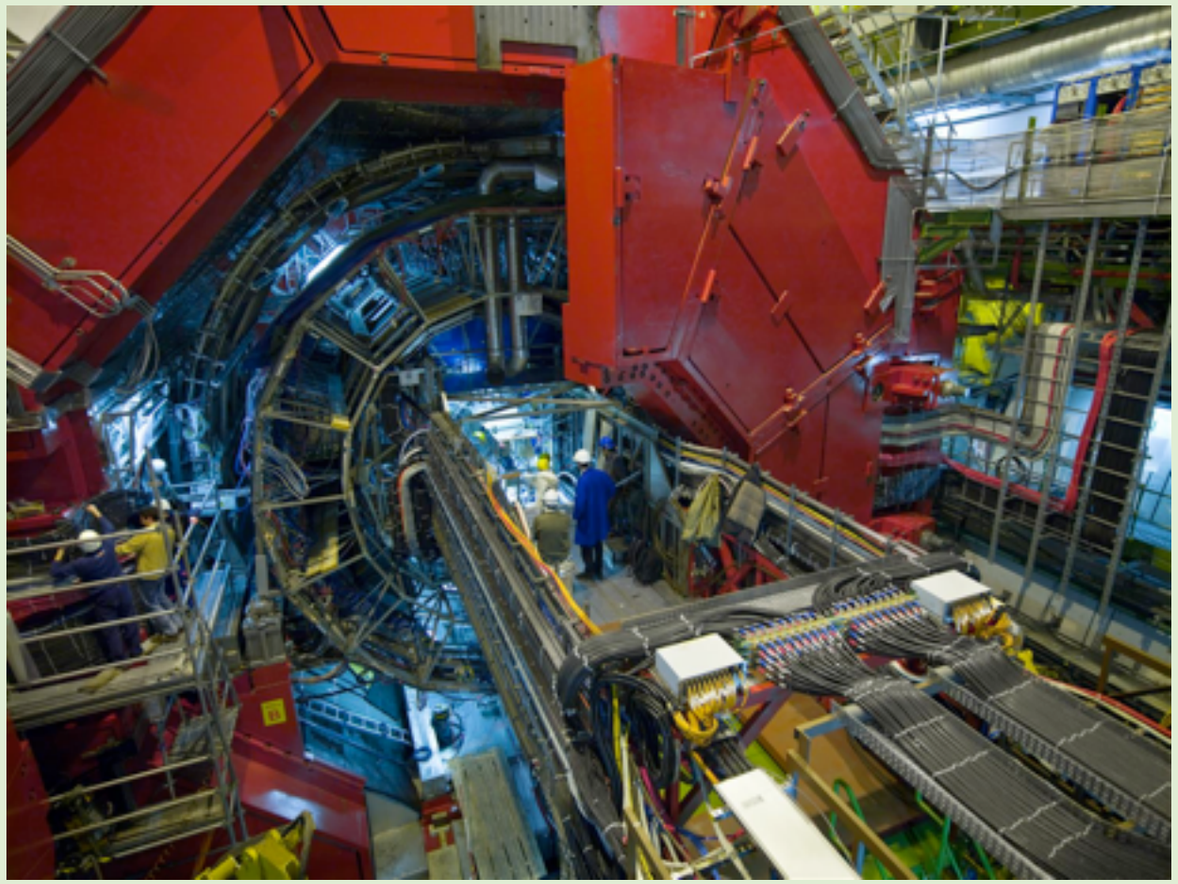
two more experiments w/ Heavy-Ions:
- CMS, ATLAS

ess QCD matter under controlled
ding two heavy atomic nuclei!



Heating & Compressing QCD Matter

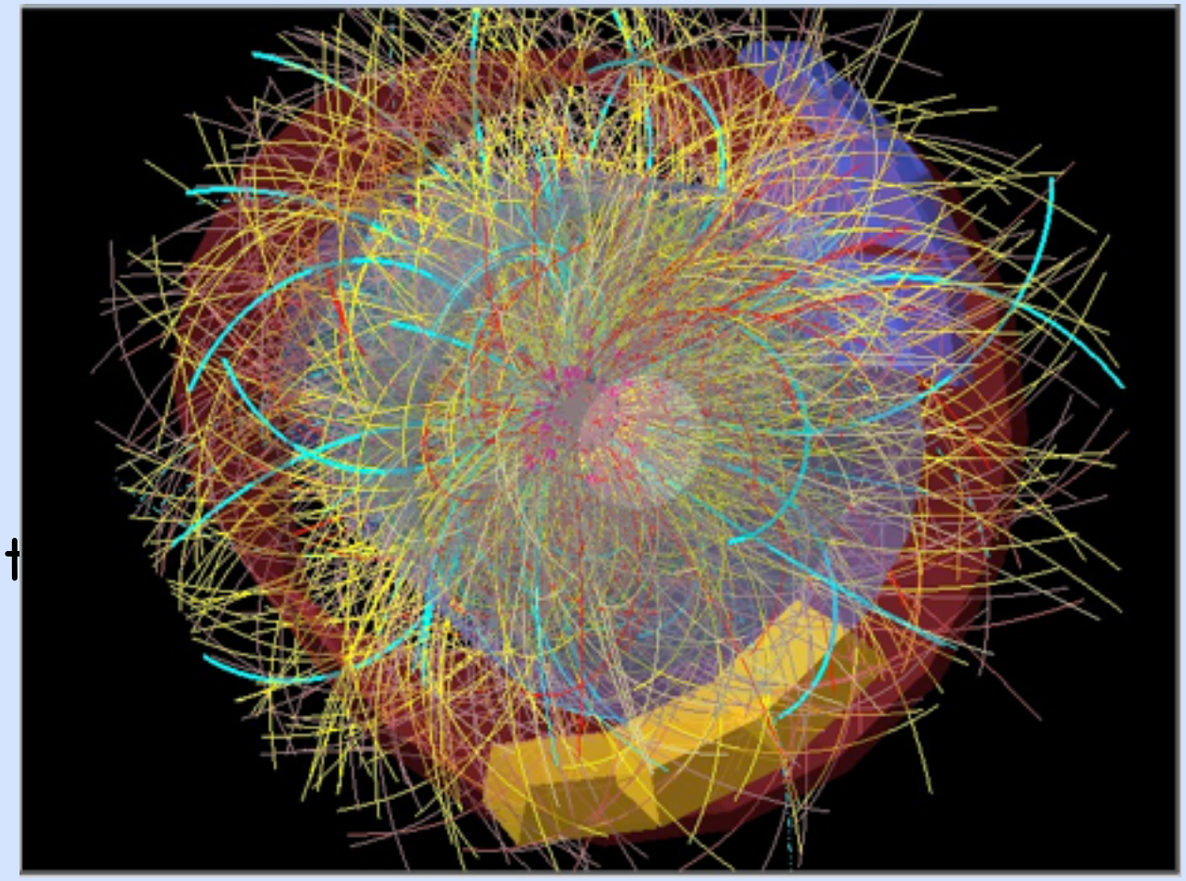
ALICE experiment @ CERN:



- 1000+ scientists from 105+ institutions
- dimensions: 26m long, 16m high, 16m wide
- weight: 10,000 tons

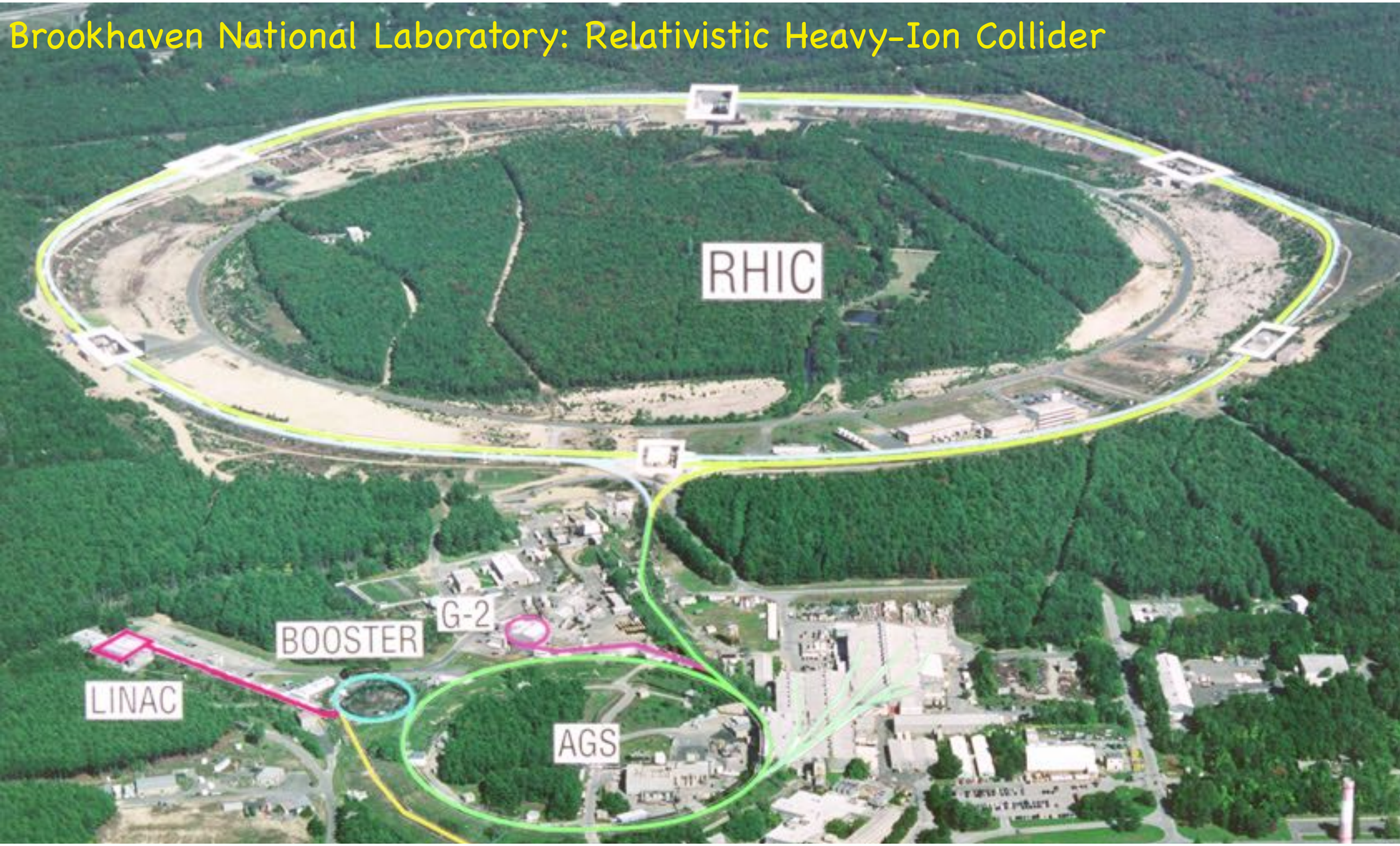
two more experiments w/ Heavy-Ions:
- CMS, ATLAS

typical Pb+Pb collision @ LHC:



- 1000s of tracks
- task: reconstruction of final state to characterize matter created in collision

RHIC: A dedicated QGP Machine



RHIC: A dedicated QGP Machine

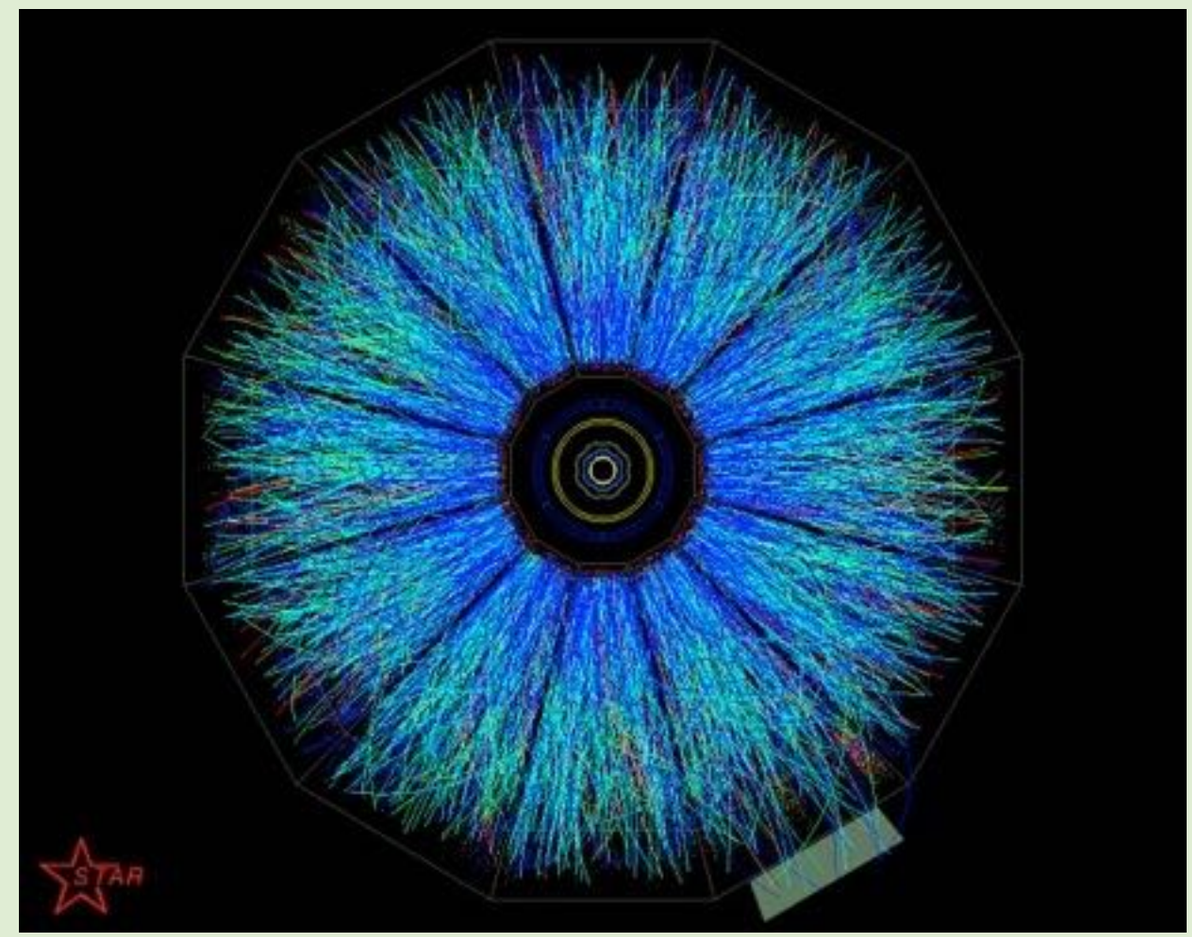
Brookhaven National Laboratory: Relativistic Heavy-Ion Collider

- 2 large experiments (STAR, PHENIX)
- 2 small experiments (PHOBOS, BRAHMS)

RHIC: A dedicated QGP Machine

Brookhaven National Laboratory: Relativistic Heavy-Ion Collider

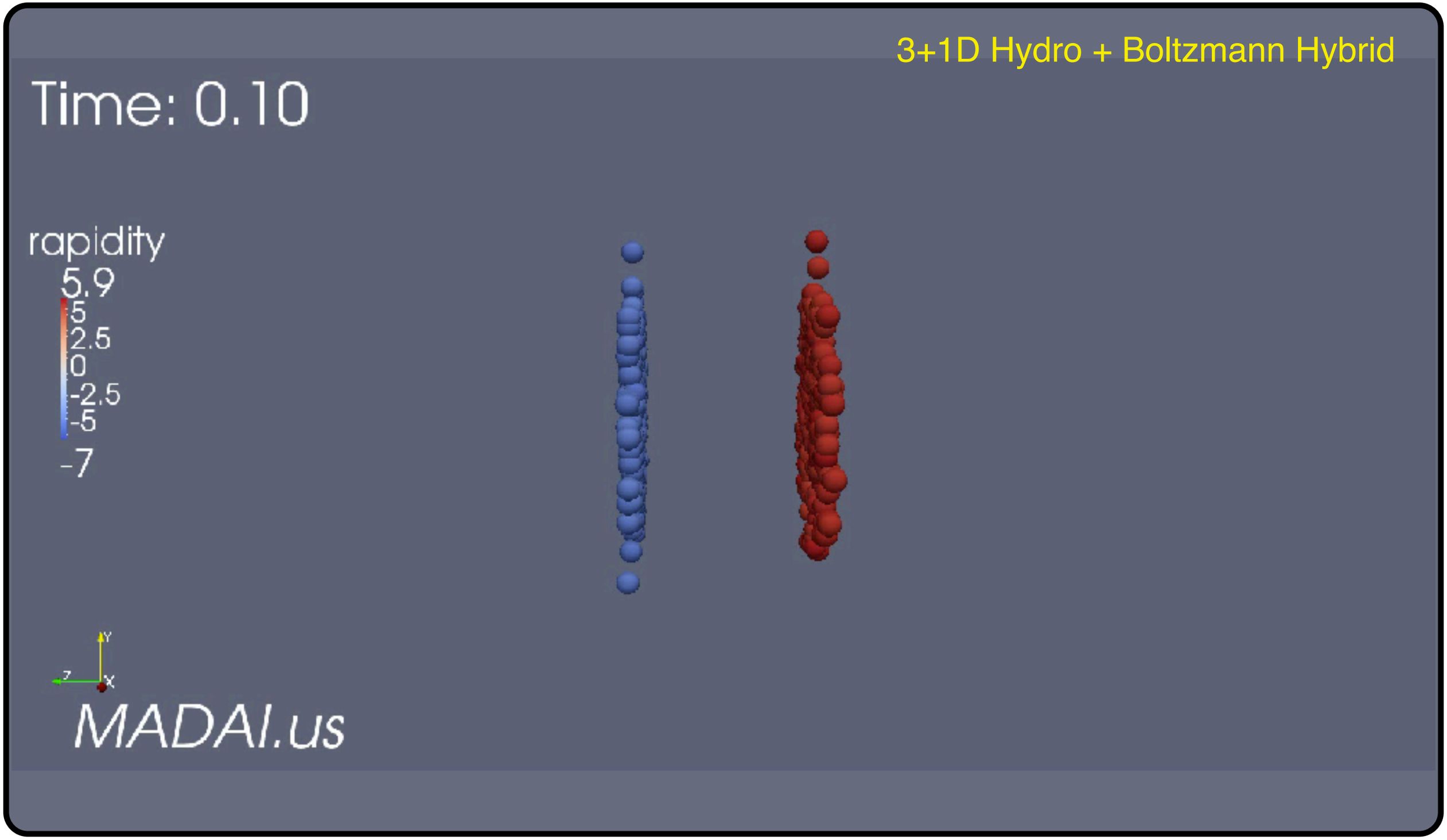
- 2 large experiments (STAR, PHENIX)
- 2 small experiments (PHOBOS, BRAHMS)



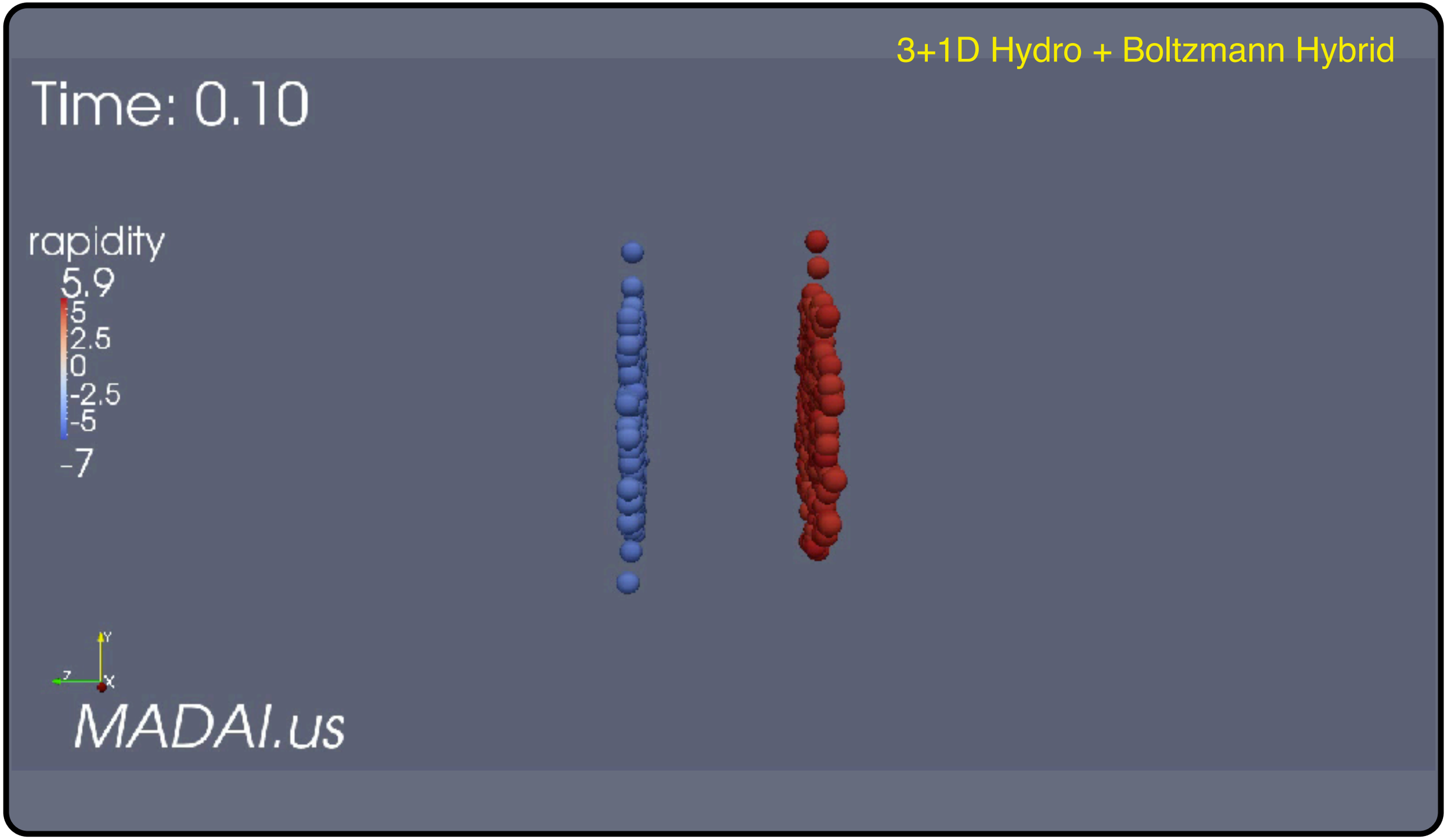
- typical collision @ RHIC: 1000s of tracks
- task: reconstruction of final state to characterize matter created in collision

Transport Theory:
Connecting Data to Knowledge

Computational Modeling

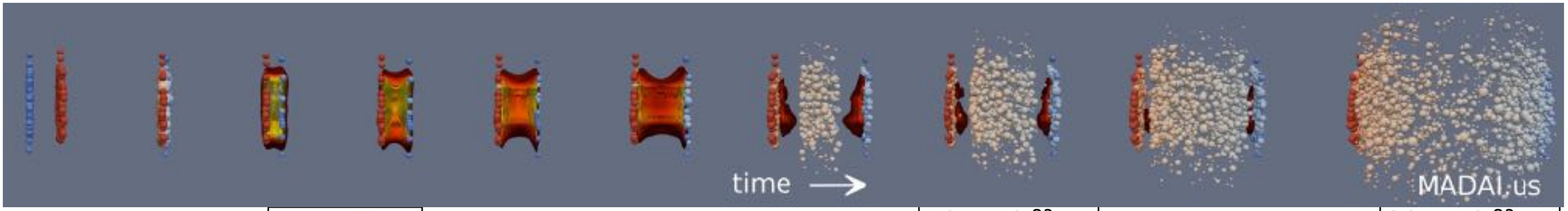


Computational Modeling



Time-Evolution of a Heavy-Ion Collision

Time-Evolution of a Heavy-Ion Collision



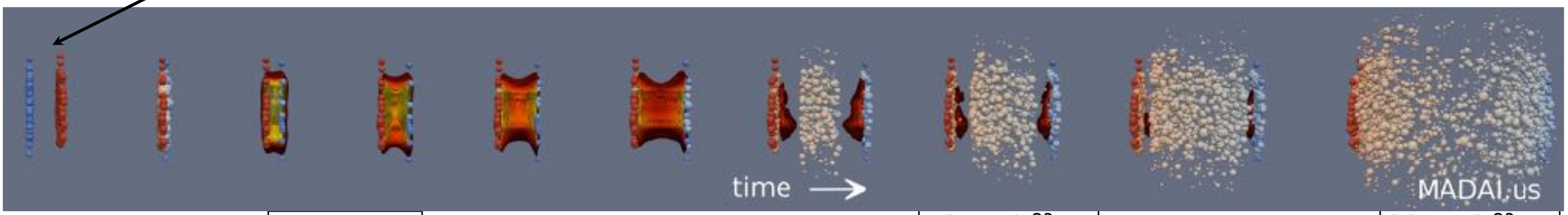
1×10^{-23} s

10×10^{-23} s

30×10^{-23} s

Time-Evolution of a Heavy-Ion Collision

nuclei at 99.99% speed of light



1×10^{-23} s

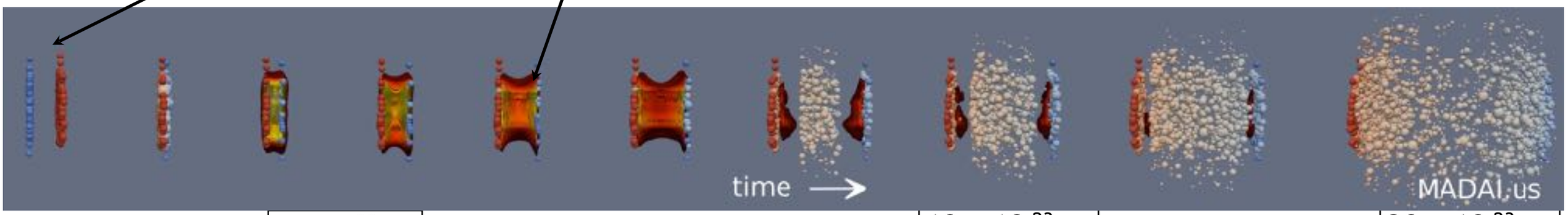
10×10^{-23} s

30×10^{-23} s

Time-Evolution of a Heavy-Ion Collision

nuclei at 99.99% speed of light

Quark-Gluon-Plasma

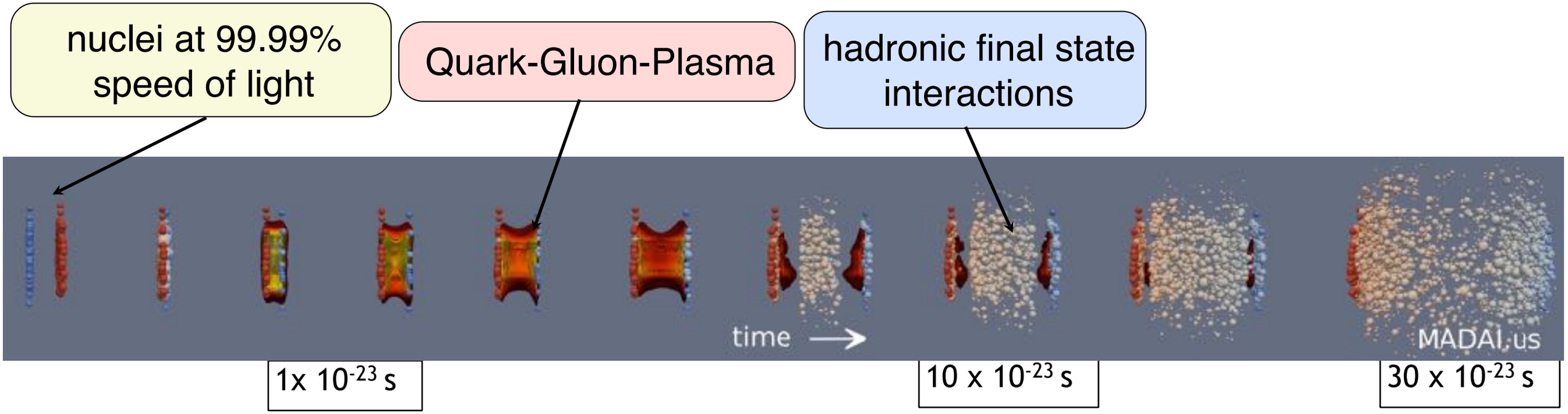


1×10^{-23} s

10×10^{-23} s

30×10^{-23} s

Time-Evolution of a Heavy-Ion Collision



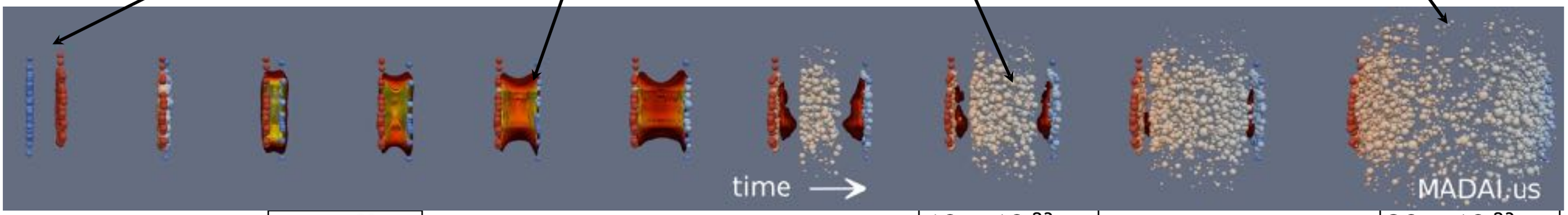
Time-Evolution of a Heavy-Ion Collision

nuclei at 99.99% speed of light

Quark-Gluon-Plasma

hadronic final state interactions

measurable (stable) particles in detector

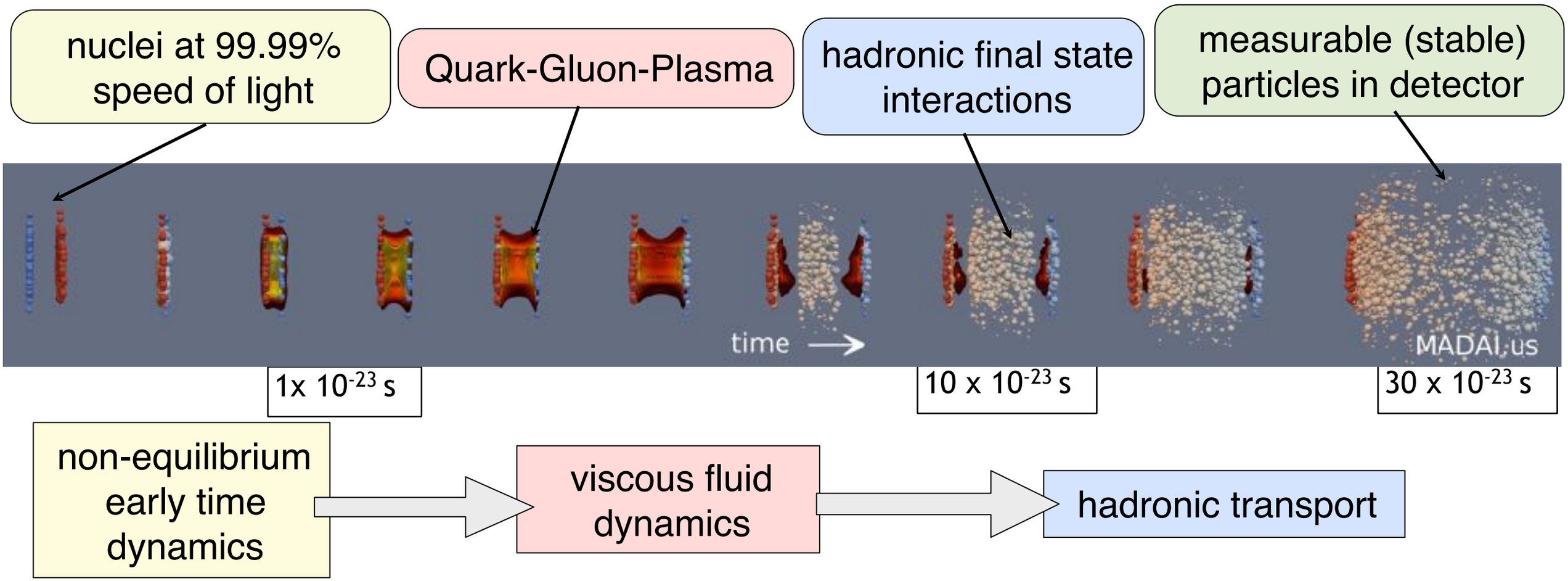


1×10^{-23} s

10×10^{-23} s

30×10^{-23} s

Time-Evolution of a Heavy-Ion Collision



Principal Challenges of Probing the QGP with Heavy-Ion Collisions:

- time-scale of the collision process: 10^{-24} seconds! [too short to resolve]
- characteristic length scale: 10^{-15} meters! [too small to resolve]
- confinement: quarks & gluons form bound states, experiments don't observe them directly

‣computational models are need to connect the experiments to QGP properties!

Transport Theory

microscopic transport models based on the **Boltzmann Equation**:

- transport of a system of microscopic particles
- all interactions are based on binary scattering

$$\left[\frac{\partial}{\partial t} + \frac{\vec{p}}{E} \times \frac{\partial}{\partial \vec{r}}\right] f_1(\vec{p}, \vec{r}, t) = \sum_{processes} C(\vec{p}, \vec{r}, t)$$

diffusive transport models based on the **Langevin Equation**:

- transport of a system of microscopic particles in a thermal medium
- interactions contain a drag term related to the properties of the medium and a noise term representing random collisions

$$\vec{p}(t + \Delta t) = \vec{p}(t) - \frac{\kappa}{2T}\vec{v} \cdot \Delta t + \vec{\xi}(t)\Delta t$$

(viscous) relativistic fluid dynamics:

- transport of macroscopic degrees of freedom
- based on conservation laws:

$$\partial_\mu T^{\mu\nu} = 0$$

$$\begin{aligned} T_{ik} &= \varepsilon u_i u_k + P\left(\delta_{ik} + u_i u_k\right) \\ &- \eta\left(\nabla_i u_k + \nabla_k u_i - \frac{2}{3}\delta_{ik}\nabla \cdot u\right) \\ &+ \varsigma\, \delta_{ik}\nabla \cdot u \end{aligned}$$

(plus an additional 9 eqns. for dissipative flows)

hybrid transport models:

- combine microscopic & macroscopic degrees of freedom
- current state of the art for RHIC modeling

Transport Theory

microscopic transport models based on the **Boltzmann Equation**:

- transport of a system of microscopic particles
- all interactions are based on binary scattering

$$\left[\frac{\partial}{\partial t} + \frac{\vec{p}}{E} \times \frac{\partial}{\partial \vec{r}}\right] f_1(\vec{p}, \vec{r}, t) = \sum_{processes} C(\vec{p}, \vec{r}, t)$$

diffusive transport models based on the **Langevin Equation**:

- transport of a system of microscopic particles in a thermal medium
- interactions contain a drag term related to the properties of the medium and a noise term representing random collisions

$$\vec{p}(t + \Delta t) = \vec{p}(t) - \frac{\kappa}{2T}\vec{v} \cdot \Delta t + \vec{\xi}(t)\Delta t$$

(viscous) relativistic fluid dynamics:

- transport of macroscopic degrees of freedom
- based on conservation laws:

$$\partial_\mu T^{\mu\nu} = 0$$

$$\begin{aligned} T_{ik} &= \varepsilon u_i u_k + P\left(\delta_{ik} + u_i u_k\right) \\ &- \eta\left(\nabla_i u_k + \nabla_k u_i - \frac{2}{3}\delta_{ik}\nabla \cdot u\right) \\ &+ \varsigma\, \delta_{ik}\nabla \cdot u \end{aligned}$$

(plus an additional 9 eqns. for dissipative flows)

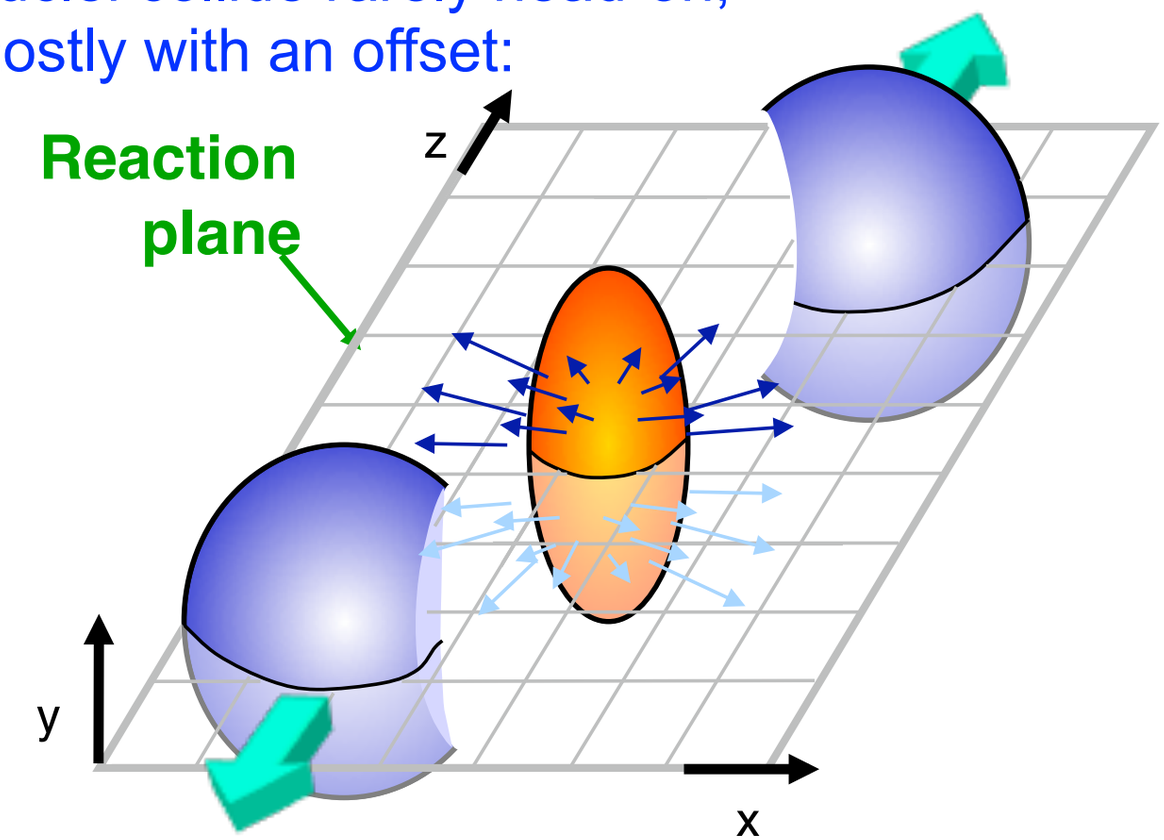
hybrid transport models:

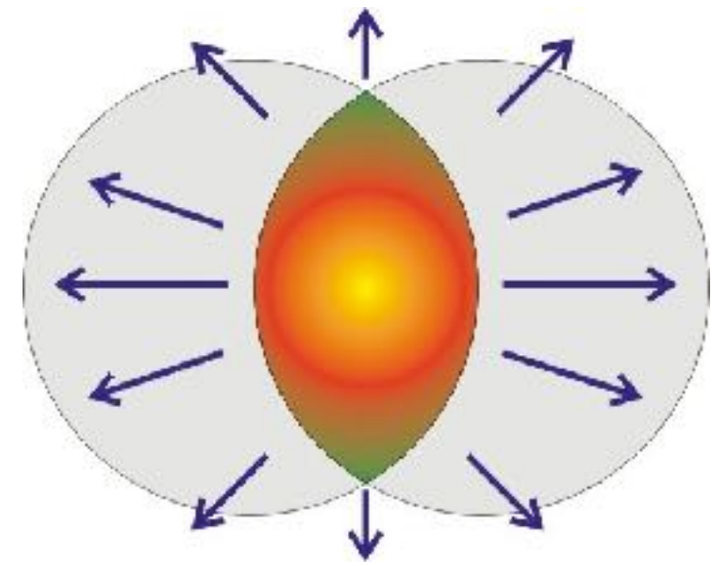
- combine microscopic & macroscopic degrees of freedom
- current state of the art for RHIC modeling

Each transport model relies on roughly a dozen physics parameters to describe the time-evolution of the collision and its final state. These physics parameters act as a representation of the information we wish to extract from RHIC & LHC.

Collision Geometry: Elliptic Flow

- two nuclei collide rarely head-on, but mostly with an offset:

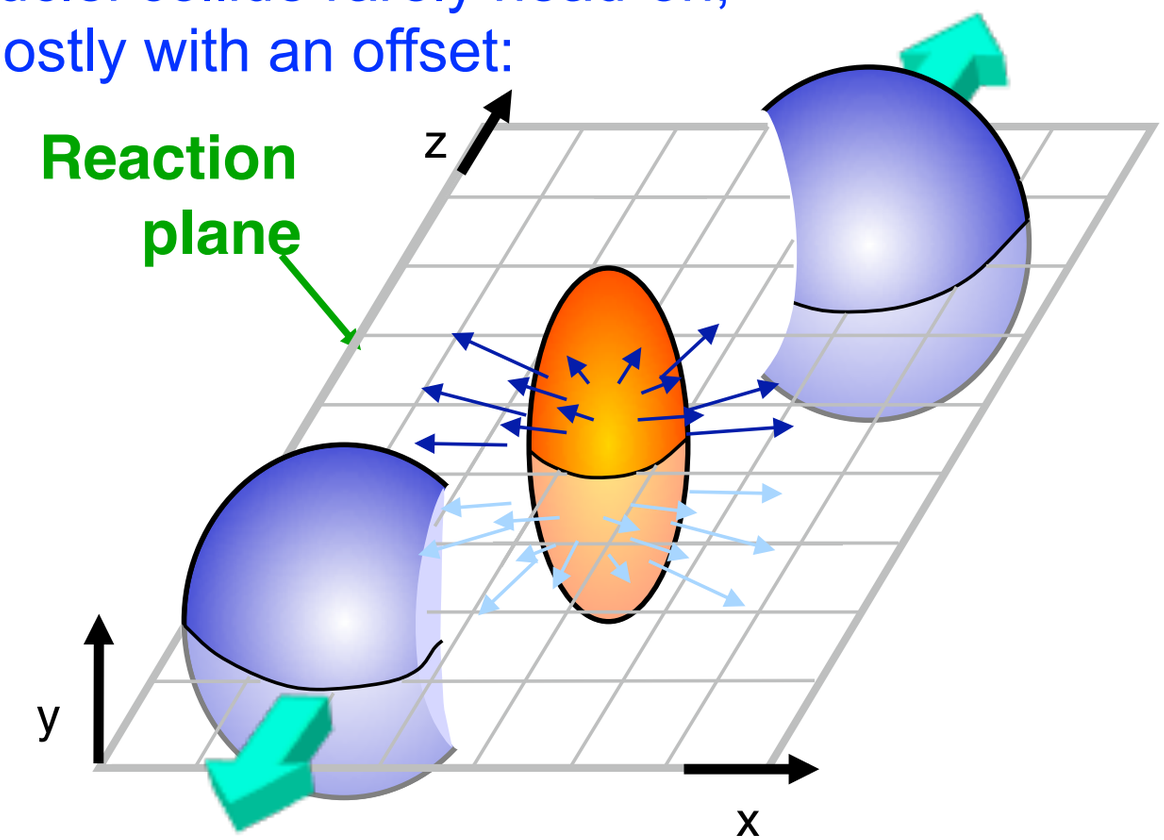


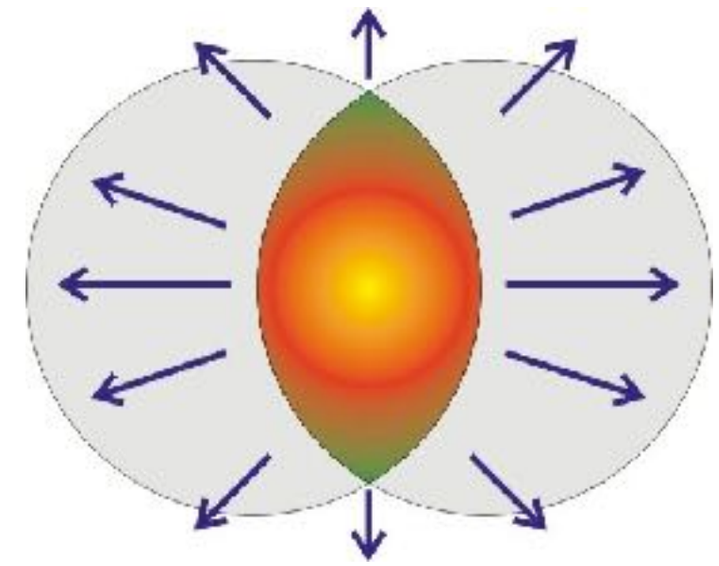


only matter in the overlap area gets compressed and heated up

Collision Geometry: Elliptic Flow

- two nuclei collide rarely head-on, but mostly with an offset:



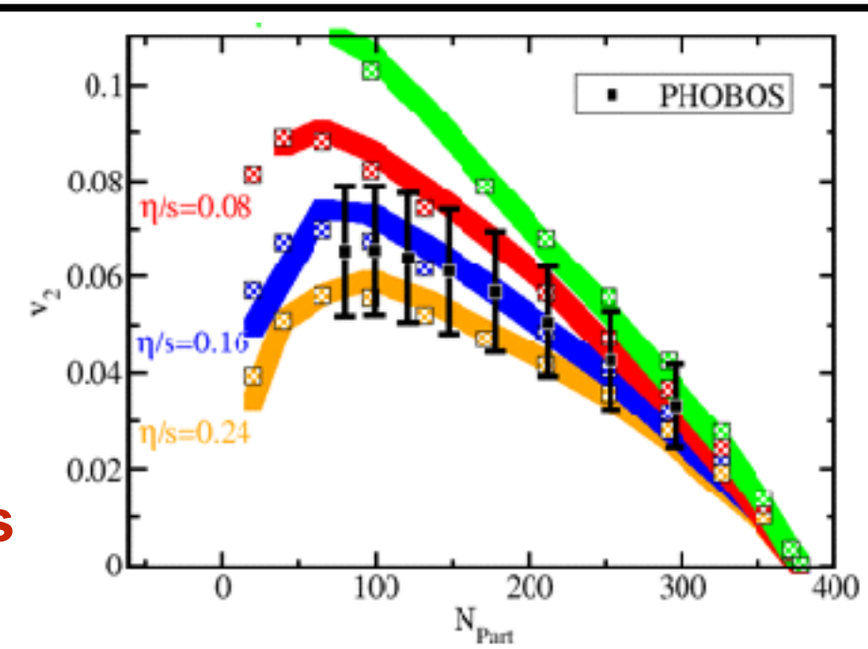


only matter in the overlap area gets compressed and heated up

elliptic flow (v_2):

- gradients of almond-shape surface will lead to preferential emission in the reaction plane
- asymmetry out- vs. in-plane emission is quantified by 2nd Fourier coefficient of angular distribution: v_2

➢ **vRFD: good agreement with data for very small η/s**



How to determine the the shear viscosity of the QGP?

Determining the QGP Shear Viscosity via a Model to Data Comparison

Model Parameter:

eqn. of state

shear viscosity

initial state

pre-equilibrium dynamics

thermalization time

quark/hadron chemistry

particlization/freeze-out

experimental data:

π/K/P spectra

yields vs. centrality & beam

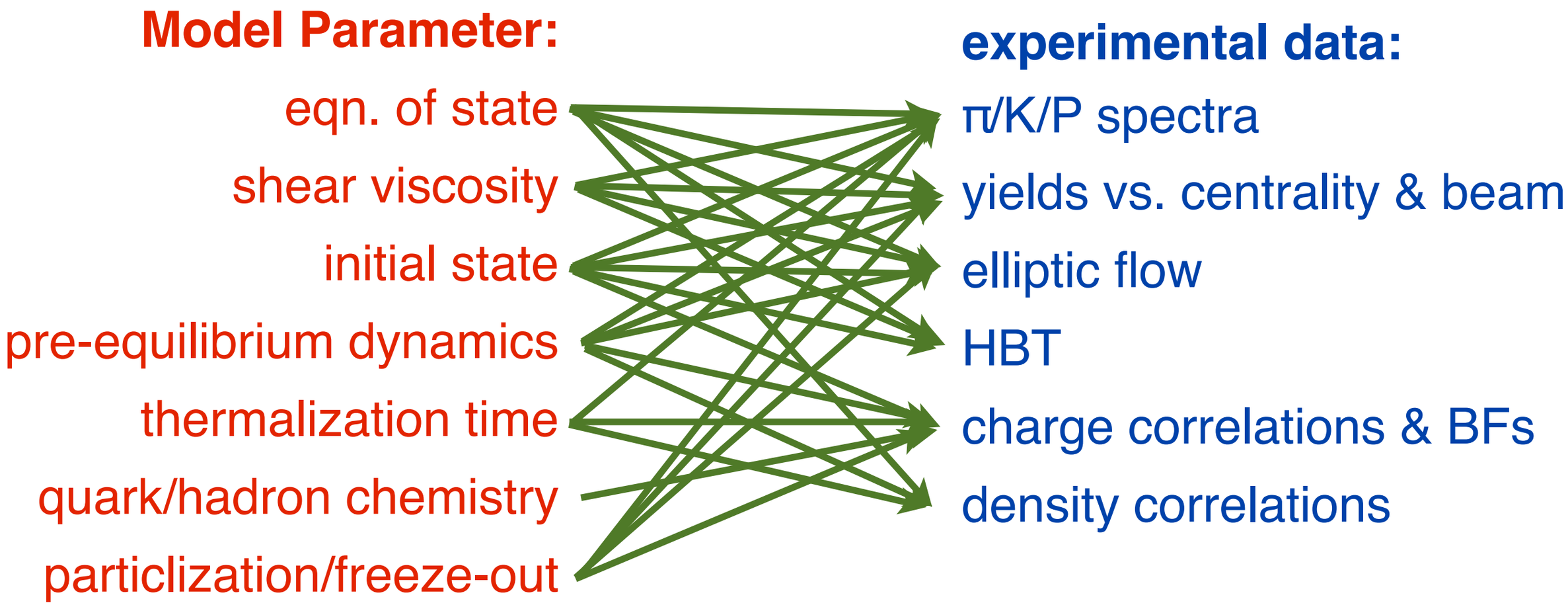
elliptic flow

HBT

charge correlations & BFs

density correlations

Determining the QGP Shear Viscosity via a Model to Data Comparison



Determining the QGP Shear Viscosity via a Model to Data Comparison

Model Parameter:

- eqn. of state
- shear viscosity
- initial state
- pre-equilibrium dynamics
- thermalization time
- quark/hadron chemistry
- particlization/freeze-out

experimental data:

- π/K/P spectra
- yields vs. centrality & beam
- elliptic flow
- HBT
- charge correlations & BFs
- density correlations

- large number of interconnected parameters w/ non-factorizable data dependencies
- data have correlated uncertainties
- develop novel optimization techniques: Bayesian Statistics and MCMC methods
- transport models require too much CPU: need new techniques based on emulators
- general problem, not restricted to RHIC Physics

→collaboration with Statistical Sciences

Bayesian Analysis

Each computational model relies on a set of physics parameters to describe the dynamics and properties of the system. These physics parameters act as a representation of the information we wish to extract from RHIC & LHC.

Model Parameters - System Properties
- initial state
- temperature-dependent viscosities
- hydro to micro switching temperature

estimate or calculate parameters

Physics Model:
- Trento
- iEbE-VISHNU

calculate observables & compare to data

Experimental Data
- ALICE flow & spectra

Bayesian Analysis

Each computational model relies on a set of physics parameters to describe the dynamics and properties of the system. These physics parameters act as a representation of the information we wish to extract from RHIC & LHC.

Model Parameters - System Properties
- initial state
- temperature-dependent viscosities
- hydro to micro switching temperature

Physics Model:
- Trento
- iEbE-VISHNU

Bayesian analysis

Experimental Data
- ALICE flow & spectra

- Bayesian analysis allows us to simultaneously calibrate all model parameters via a model-to-data comparison
- determine parameter values such that the model best describes experimental observables
- extract the probability distributions of all parameters

Setup of a Bayesian Statistical Analysis

Model Parameters - System Properties
- initial state
- temperature-dependent viscosities
- hydro to micro switching temperature

calculate events on Latin hypercube

Physics Model:
- Trento
- iEbE-VISHNU

Gaussian Process Emulator
- non-parametric interpolation
- fast surrogate to full Physics Model

Experimental Data
- ALICE flow & spectra

MCMC
(Markov-Chain Monte-Carlo)
- random walk through parameter space weighted by posterior probability

after many steps, MCMC equilibrates to

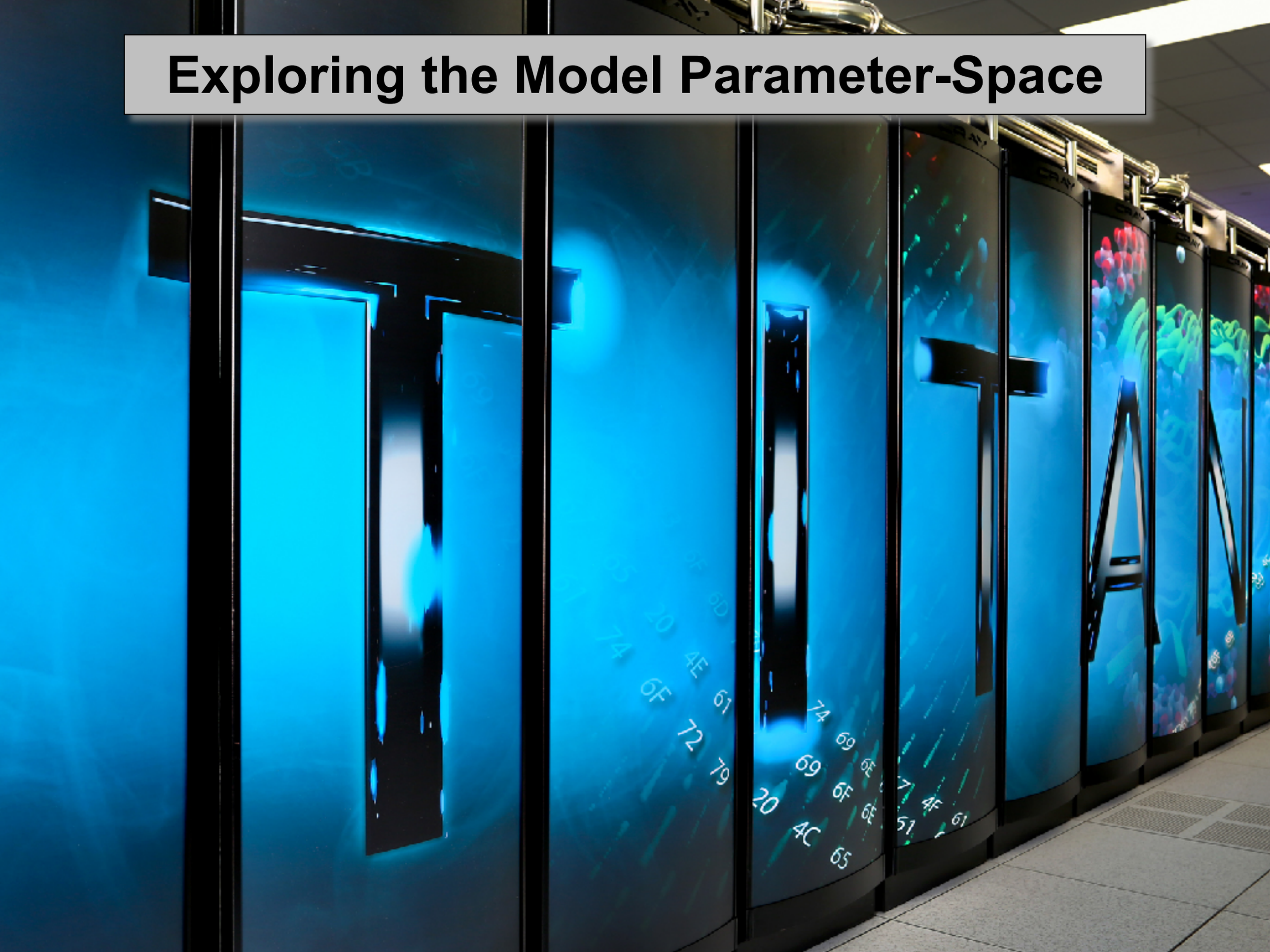
Posterior Distribution
- **diagonals**: probability distribution of each parameter, integrating out all others
- **off-diagonals**: pairwise distributions showing dependence between parameters

Bayes' Theorem

posterior $\propto$ likelihood $\times$ prior

- **prior**: initial knowledge of parameters
- **likelihood:** probability of observing exp. data, given proposed parameters

Exploring the Model Parameter-Space



Exploring the Model Parameter-Space

brute force analysis:

- 14 model parameters
- 9 centrality bins
- 20 bins per parameter
- need to evaluate model at 9×20^{14} points
- fluctuating initial conditions: $\mathcal{O}(10^4)$ events per point $\rightarrow 10^{18}$ events
- assume 1 cpu hour per event: 10^{18} cpu-hours!
- **2 billion years 100% use of TITAN @ ORNL (Cray XK7 w/ 560,640 cores)**
- then start MCMC to find point that optimally describes data…

Exploring the Model Parameter-Space

brute force analysis:

- 14 model parameters
- 9 centrality bins
- 20 bins per parameter
- need to evaluate model at 9×20^{14} points
- fluctuating initial conditions: $\mathcal{O}(10^4)$ events per point $\rightarrow 10^{18}$ events
- assume 1 cpu hour per event: 10^{18} cpu-hours!
- **2 billion years 100% use of TITAN @ ORNL (Cray XK7 w/ 560,640 cores)**
- then start MCMC to find point that optimally describes data…

Need to find techniques that cut down the cpu needed by at least a factor of 10^{10}: **Gaussian Process Emulators**

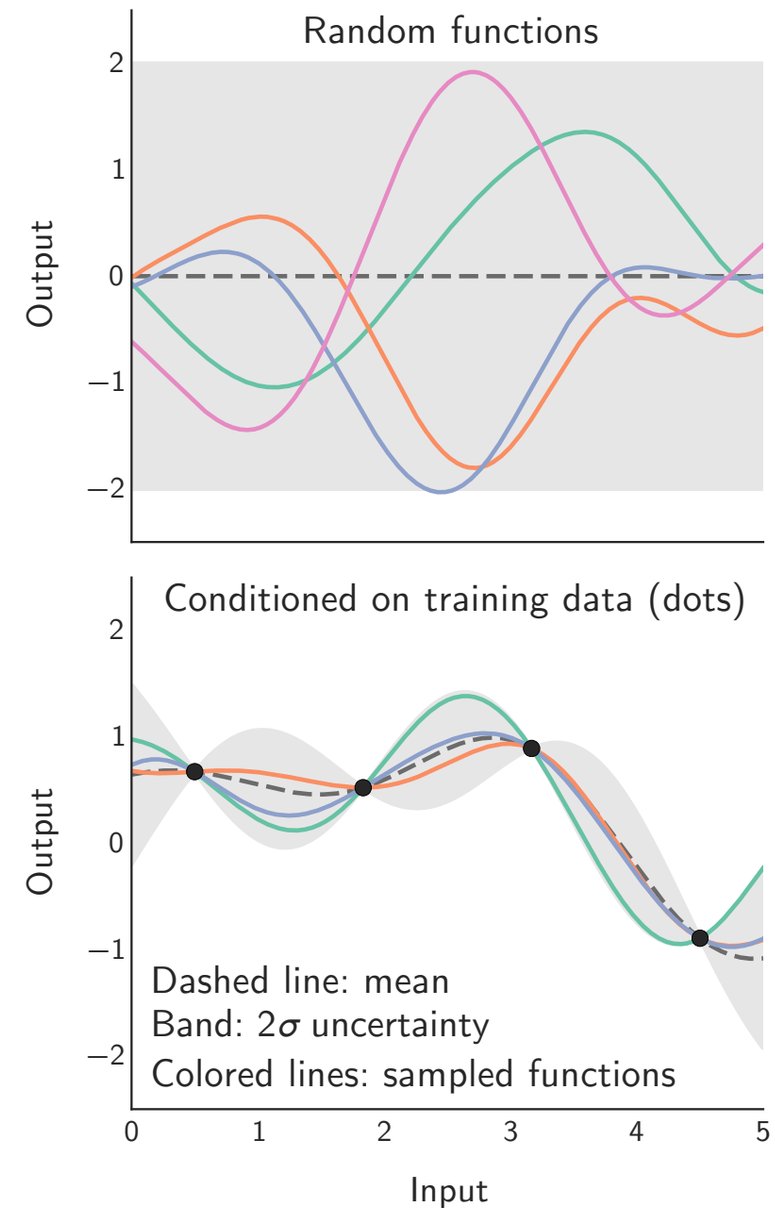
Exploring the Model Parameter-Space

Gaussian process:

- stochastic function:
  maps inputs to normally distributed outputs
- specified by mean and covariance functions

GP as a model emulator:

- non-parametric interpolation of physics model
- predicts probability distributions for model output at any given input value
  - narrow near training points, wide in gaps
- needs to be conditioned on training data (Latin hypercube points)
- fast *surrogate* to actual model



Computer Experiment Design

Latin hypercube:

- algorithm for generating semi-randomized, space-filling points (here: maximin Latin hypercube)
- avoids large gaps and tight clusters
- all parameters varied simultaneously
- needs only $m \geq 10n$ points, with n: number of model parameters

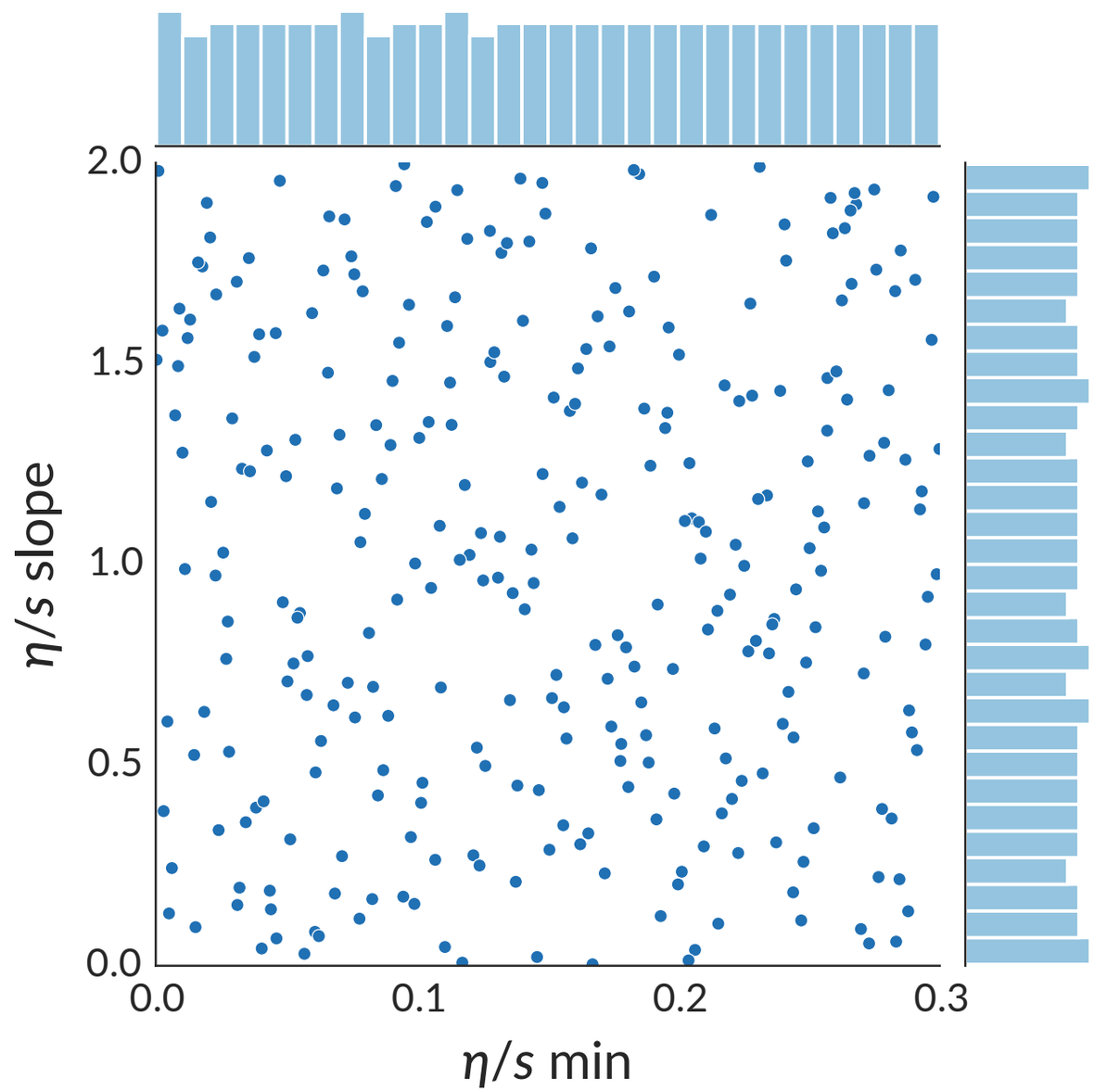
Computer Experiment Design

Latin hypercube:

- algorithm for generating semi-randomized, space-filling points (here: maximin Latin hypercube)
- avoids large gaps and tight clusters
- all parameters varied simultaneously
- needs only *m≥10n* points, with n: number of model parameters

Example:

- Latin-hypercube projection for η/s parameters



Computer Experiment Design

Latin hypercube:

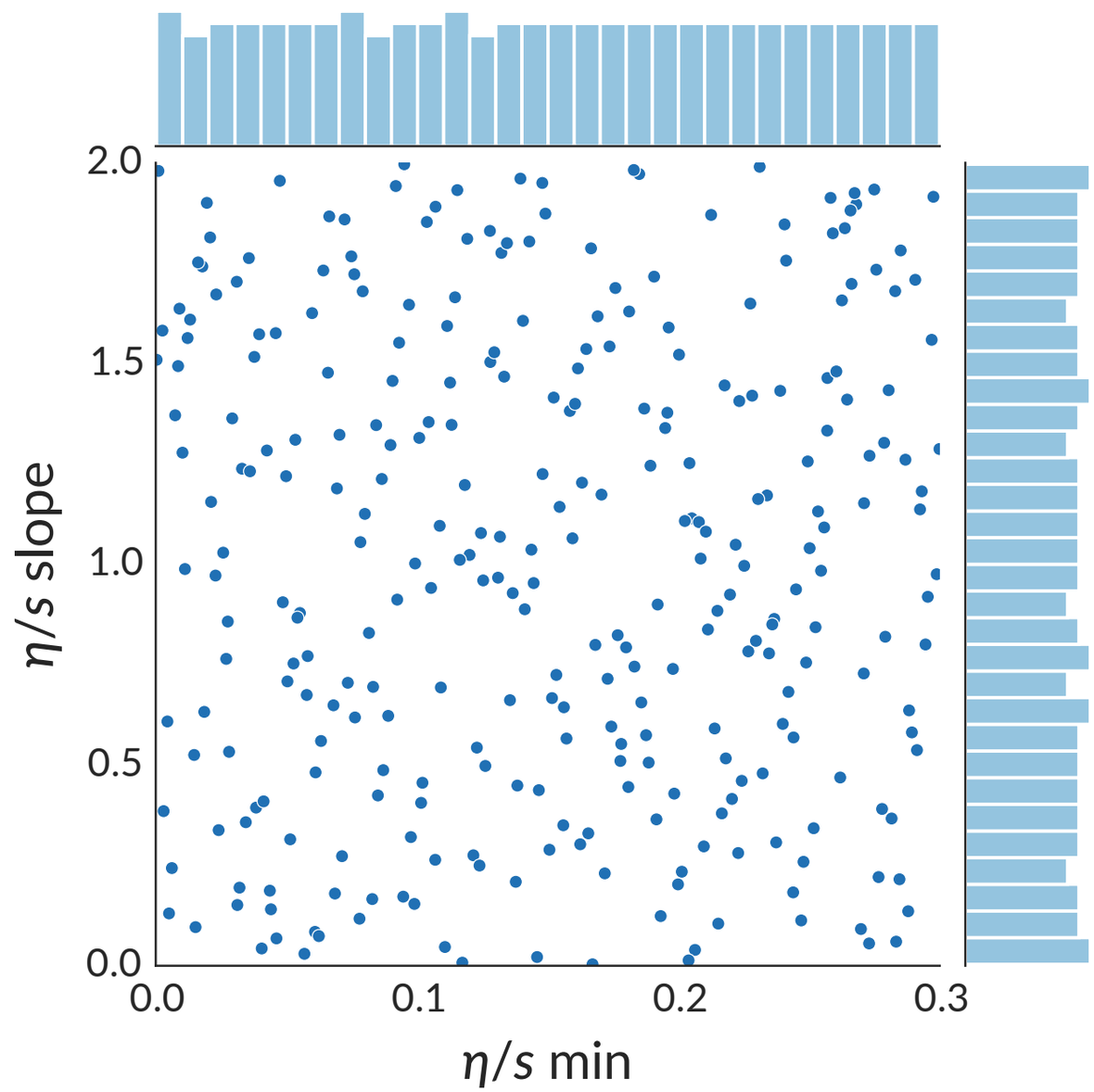
- algorithm for generating semi-randomized, space-filling points (here: maximin Latin hypercube)
- avoids large gaps and tight clusters
- all parameters varied simultaneously
- needs only *m≥10n* points, with n: number of model parameters

this design:

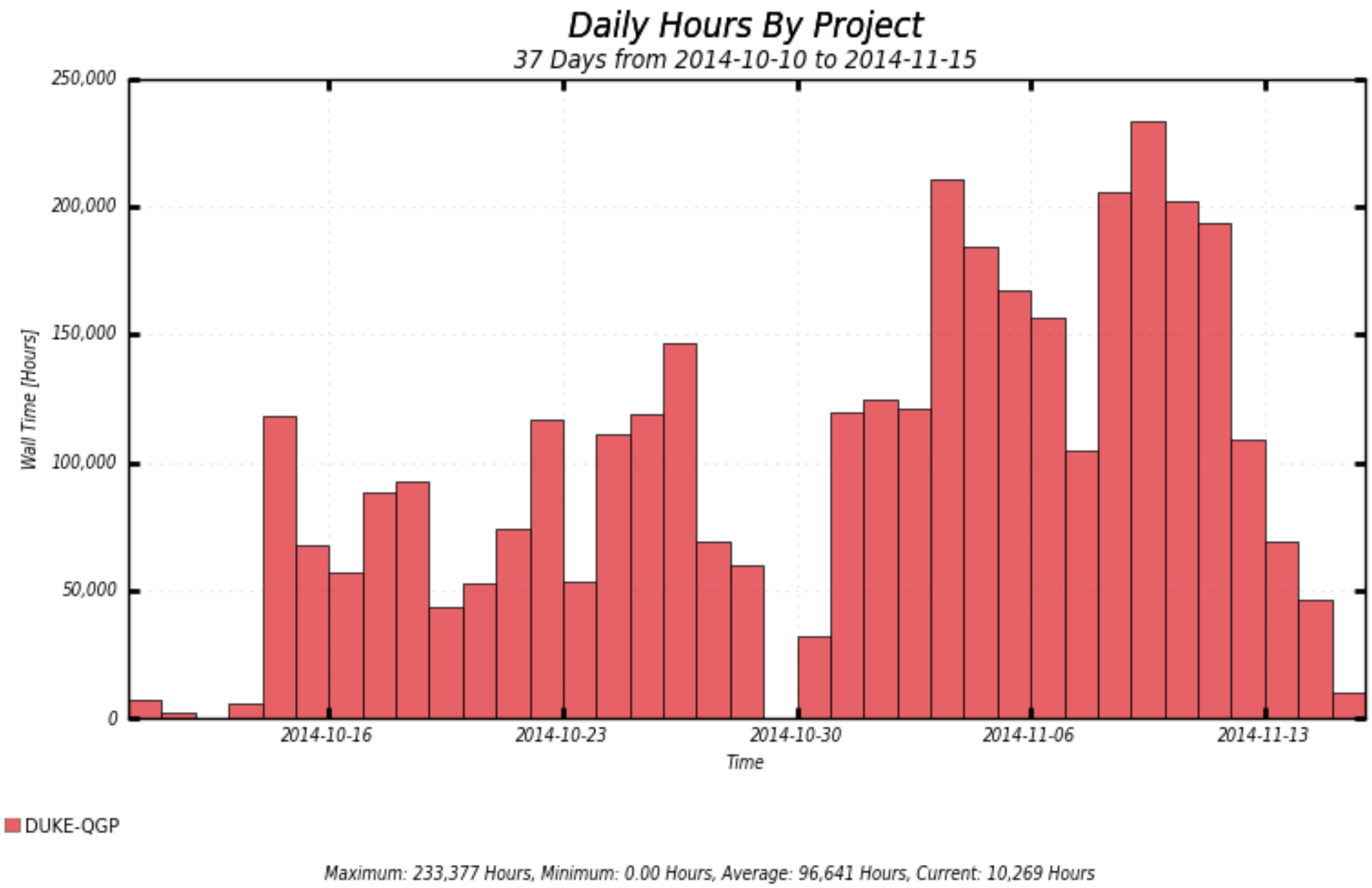
- n=14 model parameters
- 9 centrality bins, 2 energies
- Latin hypercube with m=500 points
- $\mathcal{O}(10^4)$ events per point, for a total of approx. 35,000,000 events
- use Gaussian Process Emulators to interpolate between points

Example:

- Latin-hypercube projection for η/s parameters



OSG Throughput: Duke QCD Group



- up to 233,000 cpu hours per day from OSG resources!
  - computational projects previously thought unfeasible are becoming doable
  - still need to utilize statistical tools such as Gaussian Process Emulators to reduce the computational footprint

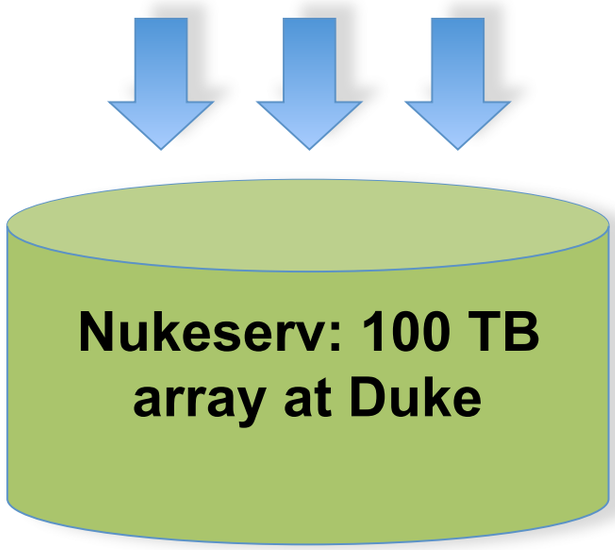
OSG Workflow @ Duke

XSEDE OSG submit node:
- CONDOR script transmits job to OSG nodes
- job may run on 1-100,000 nodes independently

local desktop at Duke:
- prepare executable & input files
- configure job for 10-20 cpu-hours

Open Science Grid is used for:
- trivially parallelizable MC-based simulations
- tasks which can be completed within 10-30 cpu hours

GridFTP protocol



compute cluster @ Duke:
- combine individual job outputs
- run analysis on output files
- perform visualization tasks

Stepping up the Game: NERSC

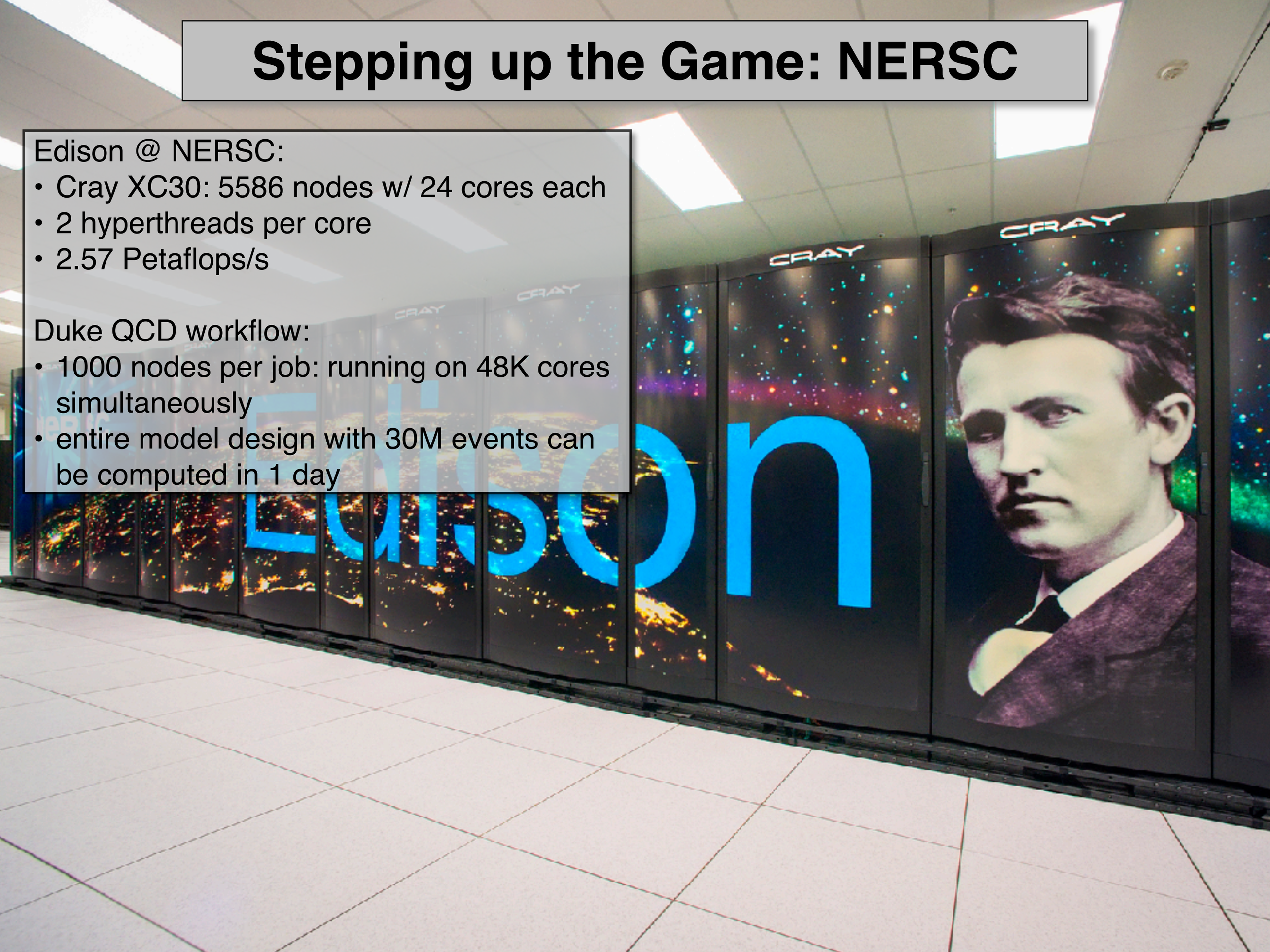
Stepping up the Game: NERSC

Edison @ NERSC:

- Cray XC30: 5586 nodes w/ 24 cores each
- 2 hyperthreads per core
- 2.57 Petaflops/s

Duke QCD workflow:

- 1000 nodes per job: running on 48K cores simultaneously
- entire model design with 30M events can be computed in 1 day



Stepping up the Game: NERSC

Edison @ NERSC:

- Cray XC30: 5586 nodes w/ 24 cores each
- 2 hyperthreads per core
- 2.57 Petaflops/s

Duke QCD workflow:

- 1000 nodes per job: running c
  simultaneously
- entire model design with 30M
  be computed in 1 day

NOW COMPUTING

A small sample of massively parallel scientific computing jobs running right now at NERSC.

PROJECT	MACHINE	NODES	NERSC HOURS USED
NERSC Staff Accounts PI: Sudip S. Dosanjh, Lawrence Berkeley National Lab - NERSC	Cori KNL	1,008	115,874.8
NERSC Staff Accounts PI: Sudip S. Dosanjh, Lawrence Berkeley National Lab - NERSC	Cori KNL	1,008	77,866.5
Extraction of QCD transport coefficients from ultra-relativistic heavy-ion collisions through a Bayesian model to data analysis PI: Steffen A. Bass, Duke University	Edison	1,000	443,890.9
Extraction of QCD transport coefficients from ultra-relativistic heavy-ion collisions through a Bayesian model to data analysis PI: Steffen A. Bass, Duke University	Edison	1,000	399,224.3
Extraction of QCD transport coefficients from ultra-relativistic heavy-ion collisions through a Bayesian model to data analysis PI: Steffen A. Bass, Duke University	Edison	750	229,928.2
NERSC Staff Accounts PI: Sudip S. Dosanjh, Lawrence Berkeley National Lab - NERSC	Cori KNL	512	282,594.2

Calibration

Vector of input parameters: $\mathbf{x}$=[p,k,w,$(\eta/s)_{min}$,$(\eta/s)_{slope}$,$(\zeta/s)_{norm}$,T_{sw}]

- assume true parameters $\mathbf{x}_\star$ exist ⇒ find probability distribution for $\mathbf{x}_\star$

Bayes' Theorem: $P(\mathbf{x}_\star \mid X,Y,\mathbf{y}_{exp}) \propto P(X,Y,\mathbf{y}_{exp} \mid \mathbf{x}_\star)P(\mathbf{x}_\star)$

- X: training data design points
- Y: model output on X
- $P(\mathbf{x}_\star)$ = prior
  ⇒ initial knowledge of $\mathbf{x}_\star$

- $P(X,Y,\mathbf{y}_{exp} \mid \mathbf{x}_\star)$ = likelihood
  ⇒ probability of observing $(X,Y,\mathbf{y}_{exp})$ given proposed $\mathbf{x}_\star$
- $P(\mathbf{x}_\star \mid X,Y,\mathbf{y}_{exp})$ = posterior
  ⇒ probability of $\mathbf{x}_\star$ given observations $(X,Y,\mathbf{y}_{exp})$

Calibration

Vector of input parameters: $\mathbf{x}$=[p,k,w,$(\eta/s)_{min}$,$(\eta/s)_{slope}$,$(\zeta/s)_{norm}$,T_{sw}]

- assume true parameters $\mathbf{x}_\star$ exist ⇒ find probability distribution for $\mathbf{x}_\star$

Bayes' Theorem: $P(\mathbf{x}_\star \mid X,Y,\mathbf{y}_{exp}) \propto P(X,Y,\mathbf{y}_{exp} \mid \mathbf{x}_\star)P(\mathbf{x}_\star)$

- X: training data design points
- Y: model output on X
- $P(\mathbf{x}_\star)$ = prior
  ⇒ initial knowledge of $\mathbf{x}_\star$

- $P(X,Y,\mathbf{y}_{exp} \mid \mathbf{x}_\star)$ = likelihood
  ⇒ probability of observing $(X,Y,\mathbf{y}_{exp})$ given proposed $\mathbf{x}_\star$
- $P(\mathbf{x}_\star \mid X,Y,\mathbf{y}_{exp})$ = posterior
  ⇒ probability of $\mathbf{x}_\star$ given observations $(X,Y,\mathbf{y}_{exp})$

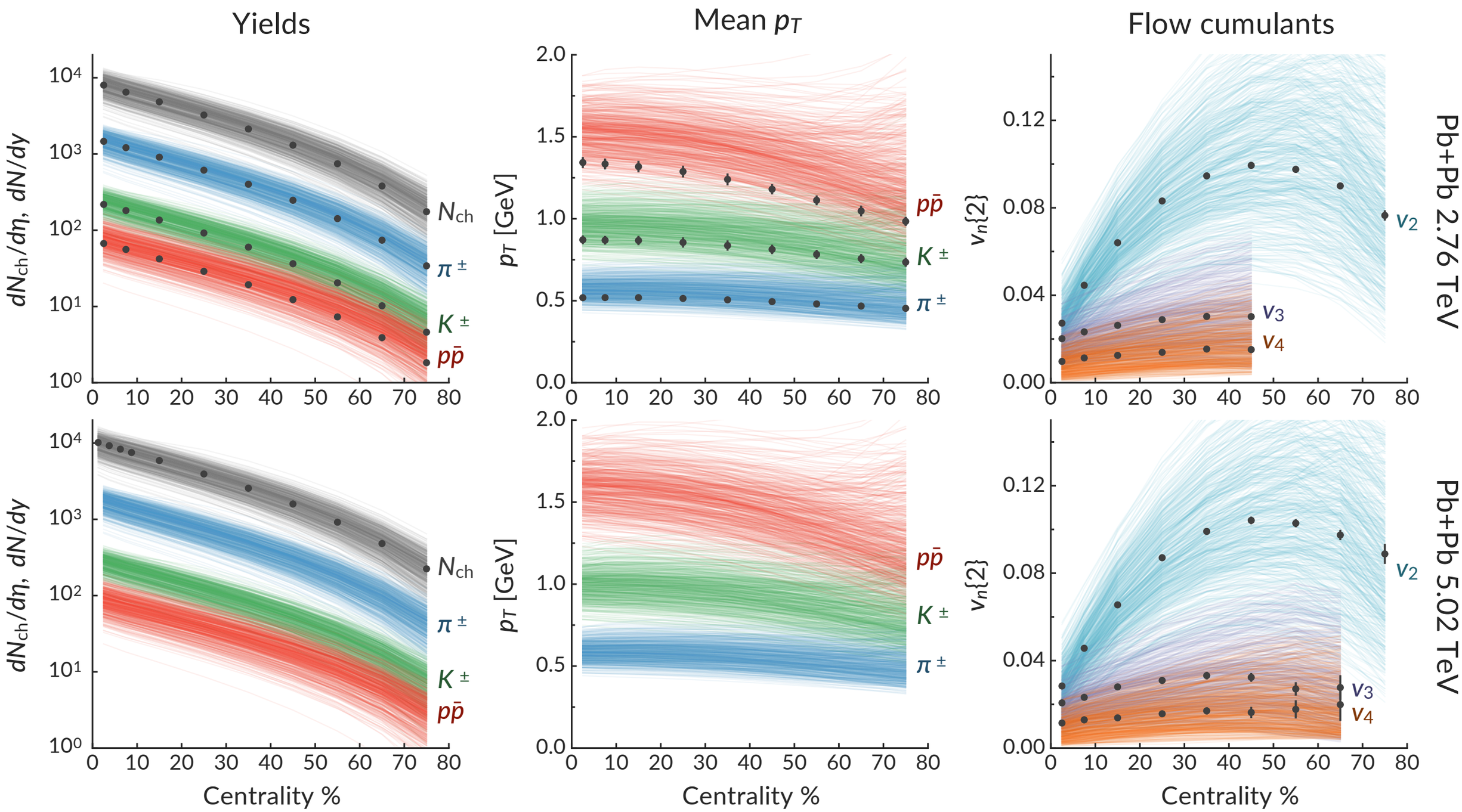
Markov-Chain Monte-Carlo:

- random walk through parameter space weighted by posterior
- large number of samples ⇒ chain equilibrates to posterior distribution

- flat prior within design range, zero outside
- likelihood: $\log[P(X,Y,\mathbf{y}_{exp} \mid \mathbf{x}_\star)] \sim -(\mathbf{y}(\mathbf{x}_\star) - \mathbf{y}_{exp})^2/(2\sigma^2)$
  - σ=0.1 on principal components (includes correlations)
- posterior ~ likelihood within design range, zero outside

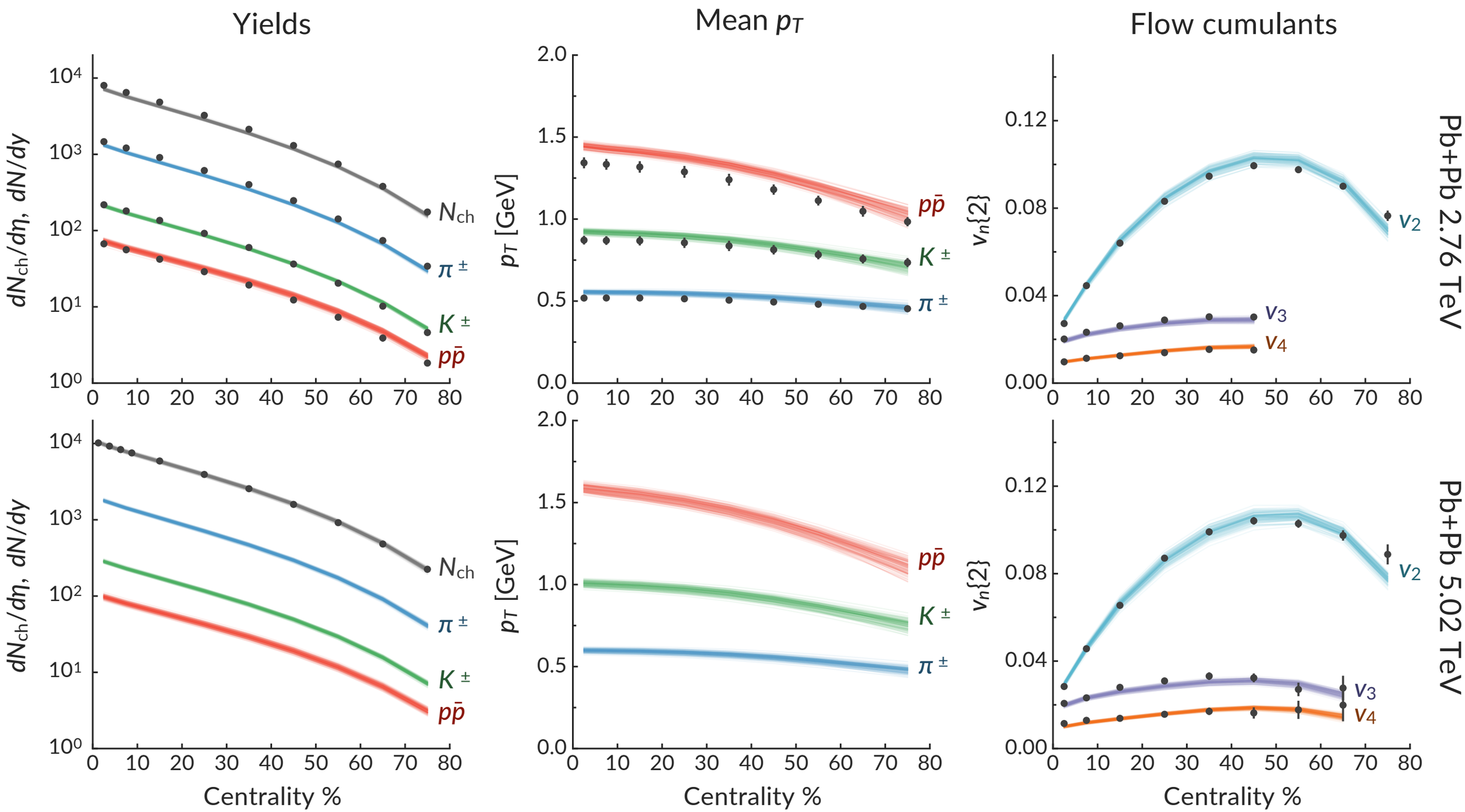
Prior vs. Posterior

Prior: model calculations evenly distributed over full design space

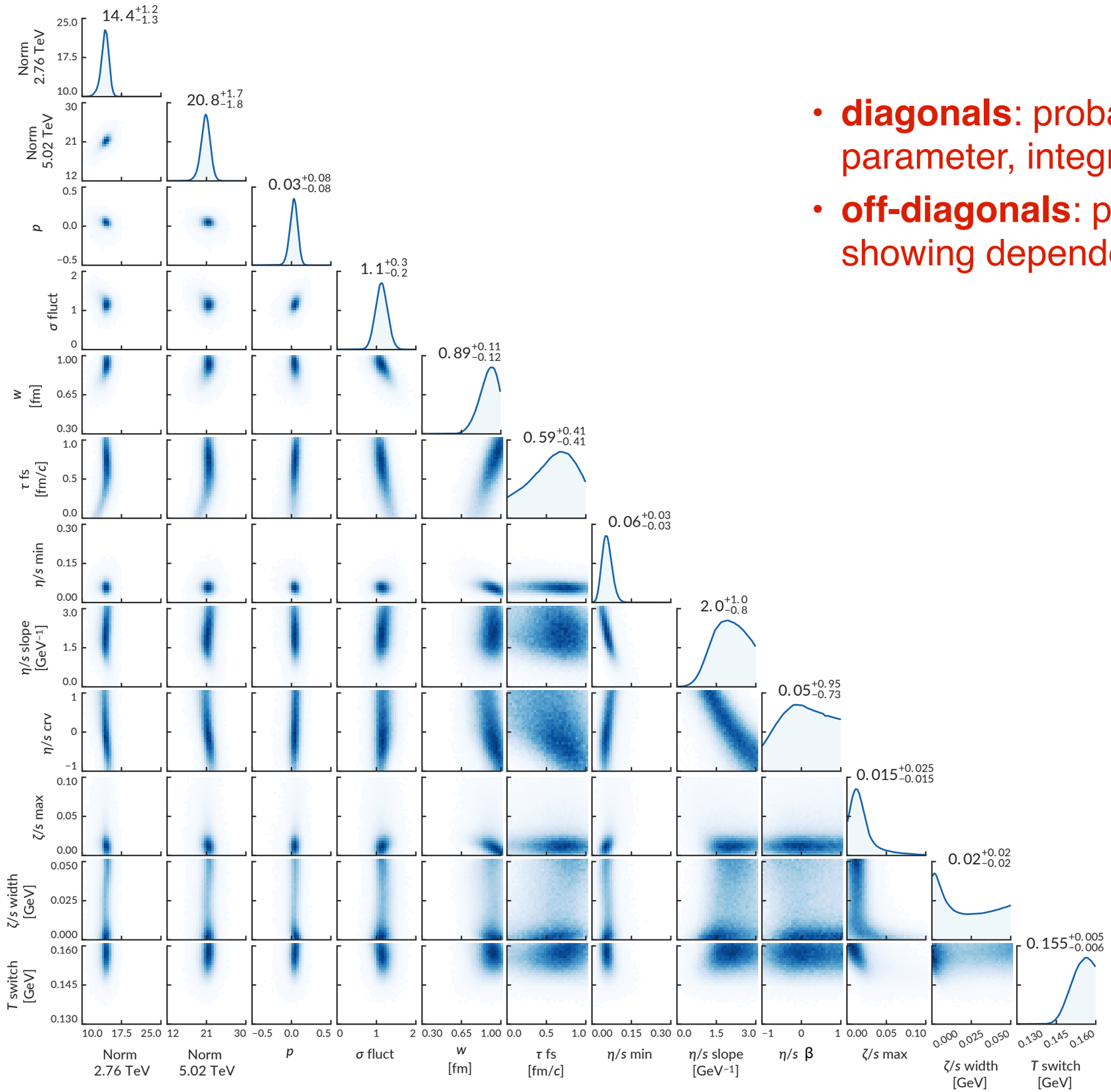


Prior vs. Posterior

Posterior: emulator predictions for highest likelihood parameter values

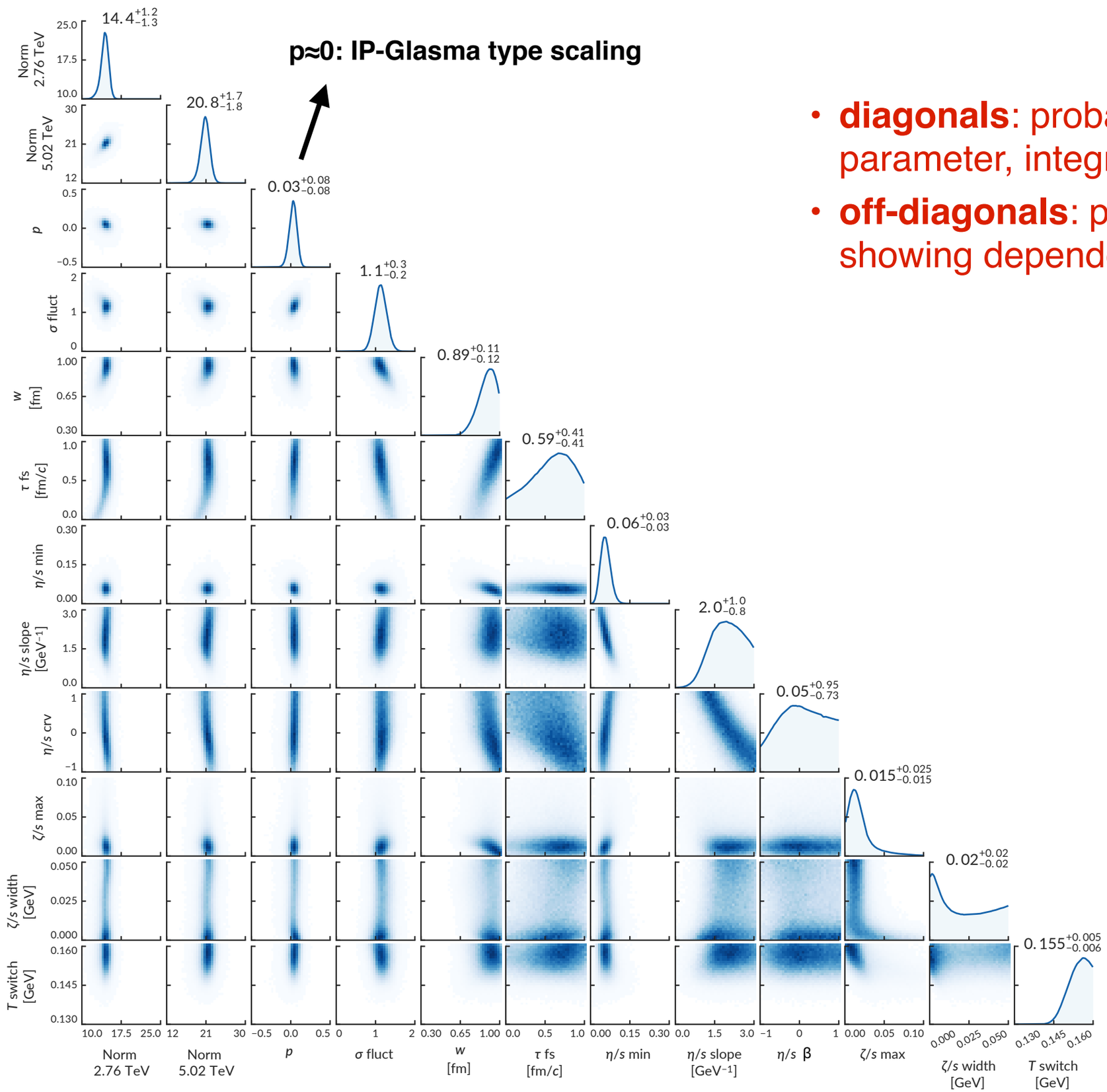


Calibrated Posterior Distribution



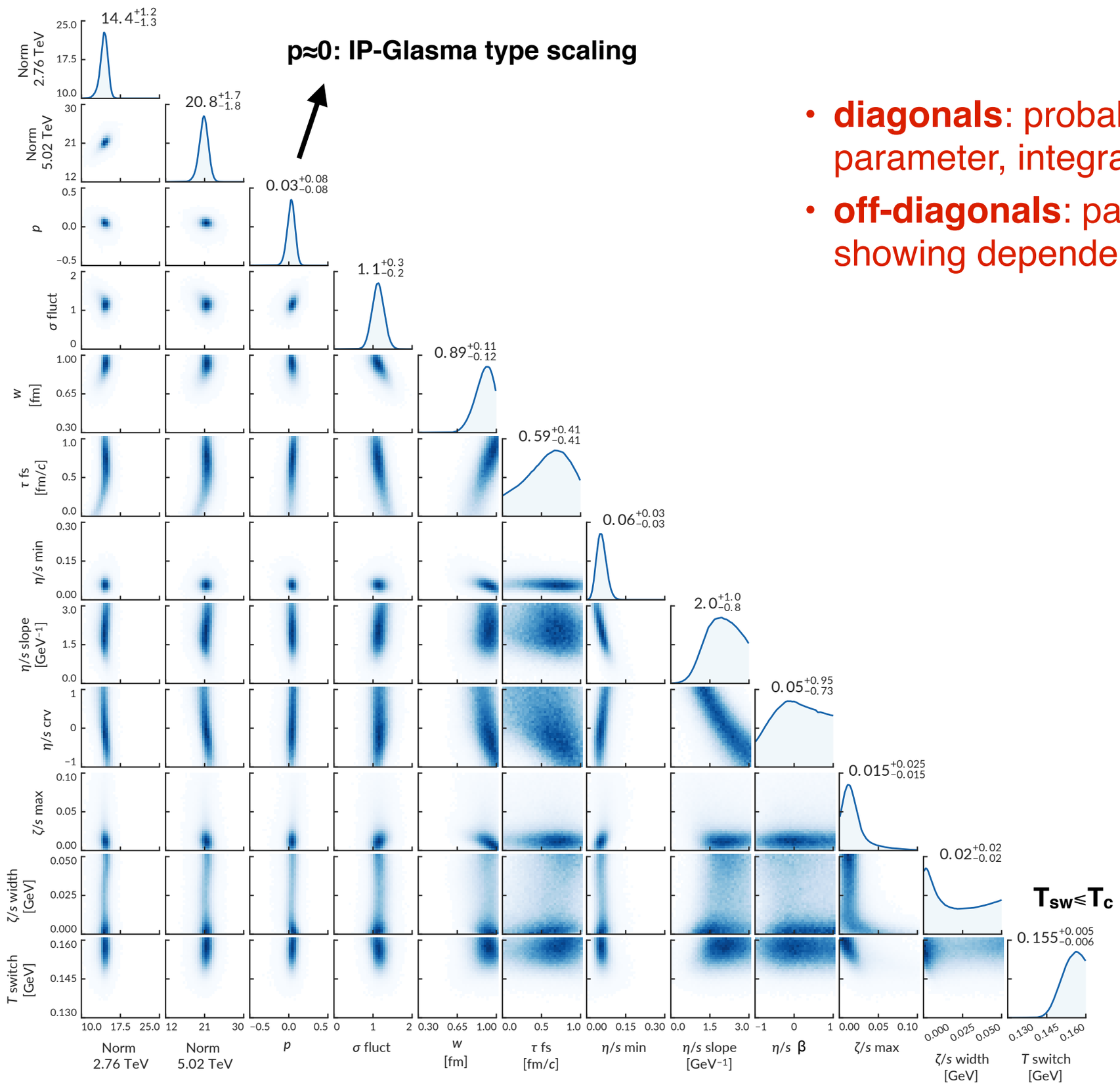
- **diagonals**: probability distribution of each parameter, integrating out all others
- **off-diagonals**: pairwise distributions showing dependence between parameters

Calibrated Posterior Distribution



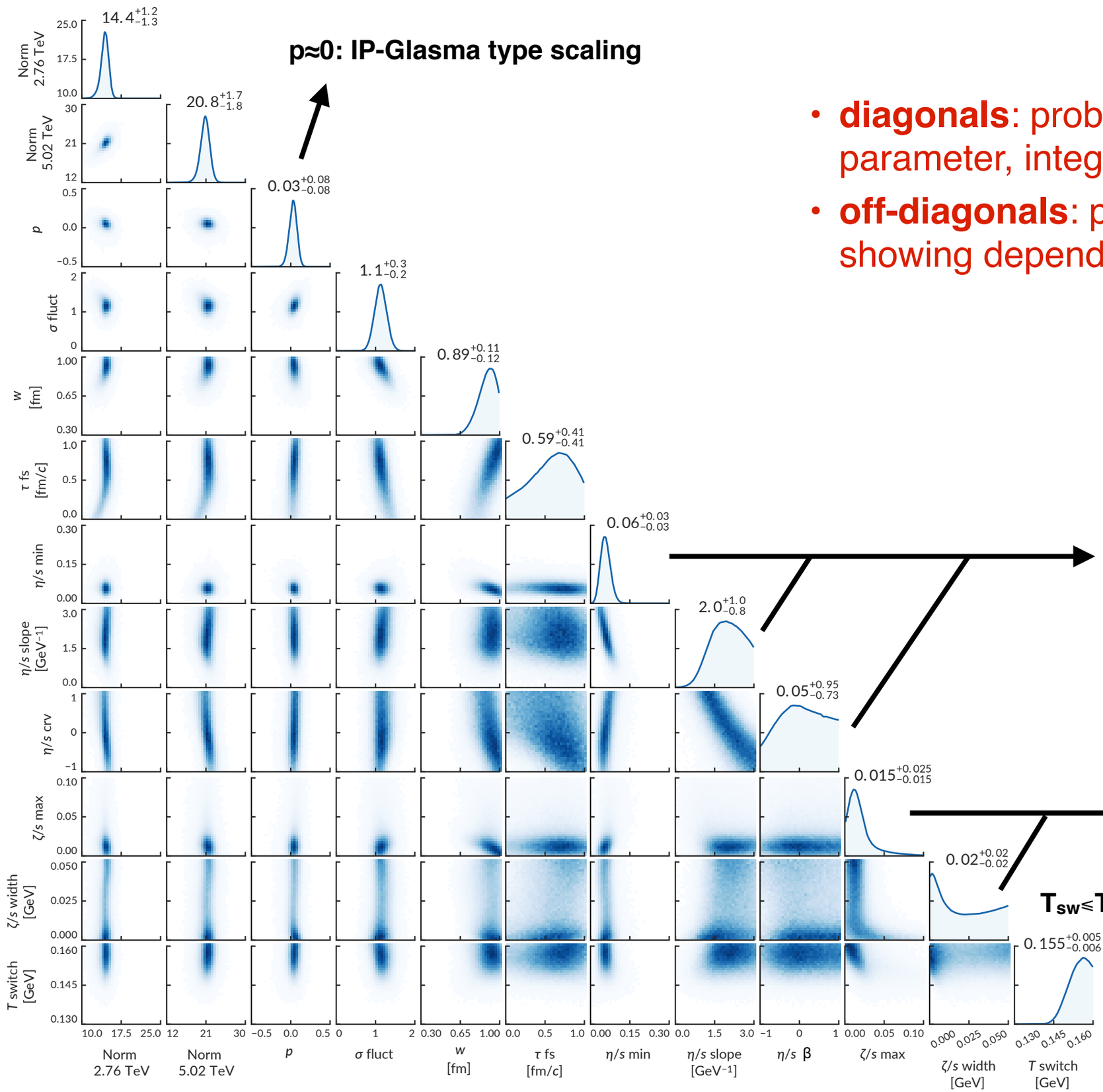
- **diagonals**: probability distribution of each parameter, integrating out all others
- **off-diagonals**: pairwise distributions showing dependence between parameters

Calibrated Posterior Distribution



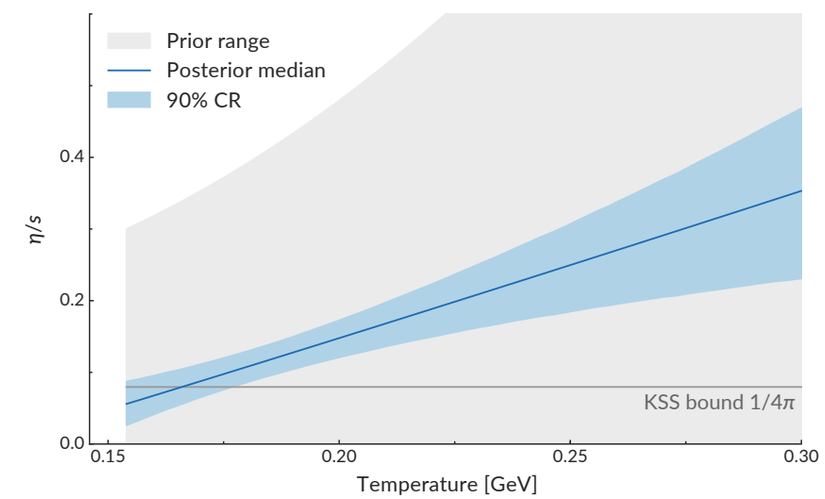
- **diagonals**: probability distribution of each parameter, integrating out all others
- **off-diagonals**: pairwise distributions showing dependence between parameters

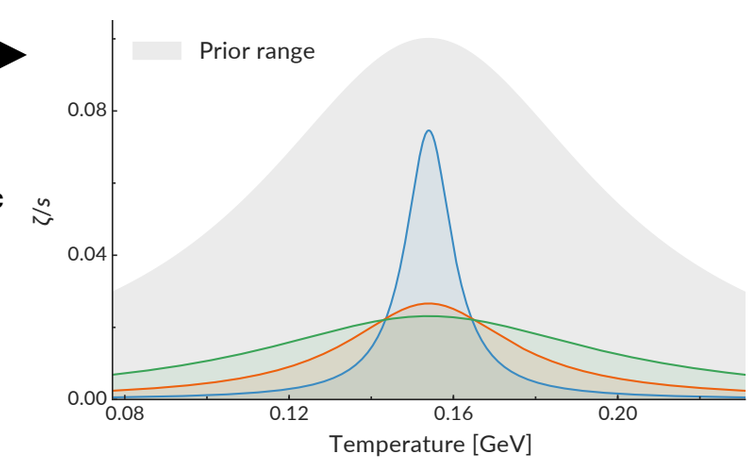
Calibrated Posterior Distribution



- **diagonals**: probability distribution of each parameter, integrating out all others
- **off-diagonals**: pairwise distributions showing dependence between parameters

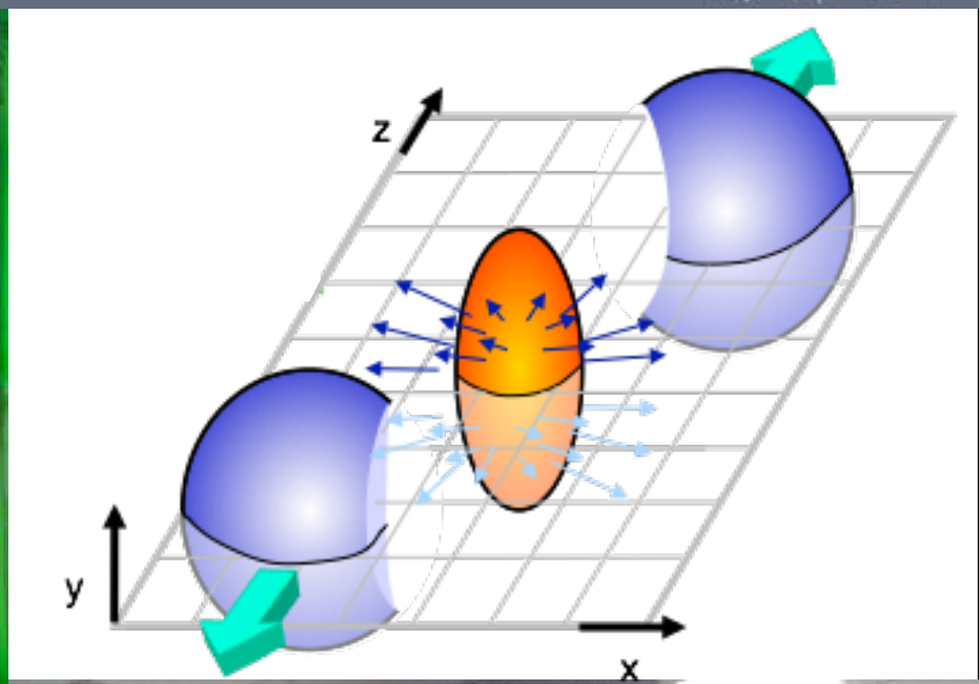
temperature-dependent viscosities:

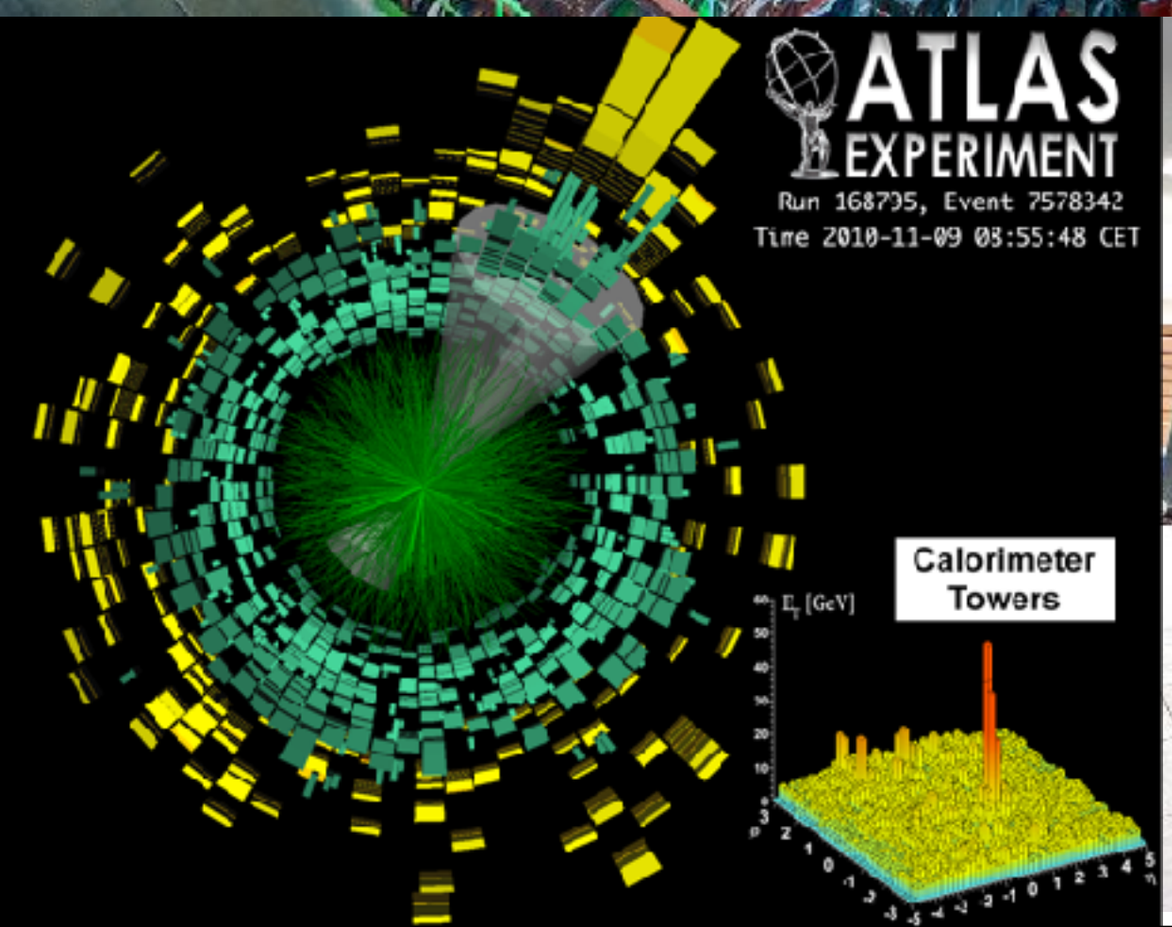


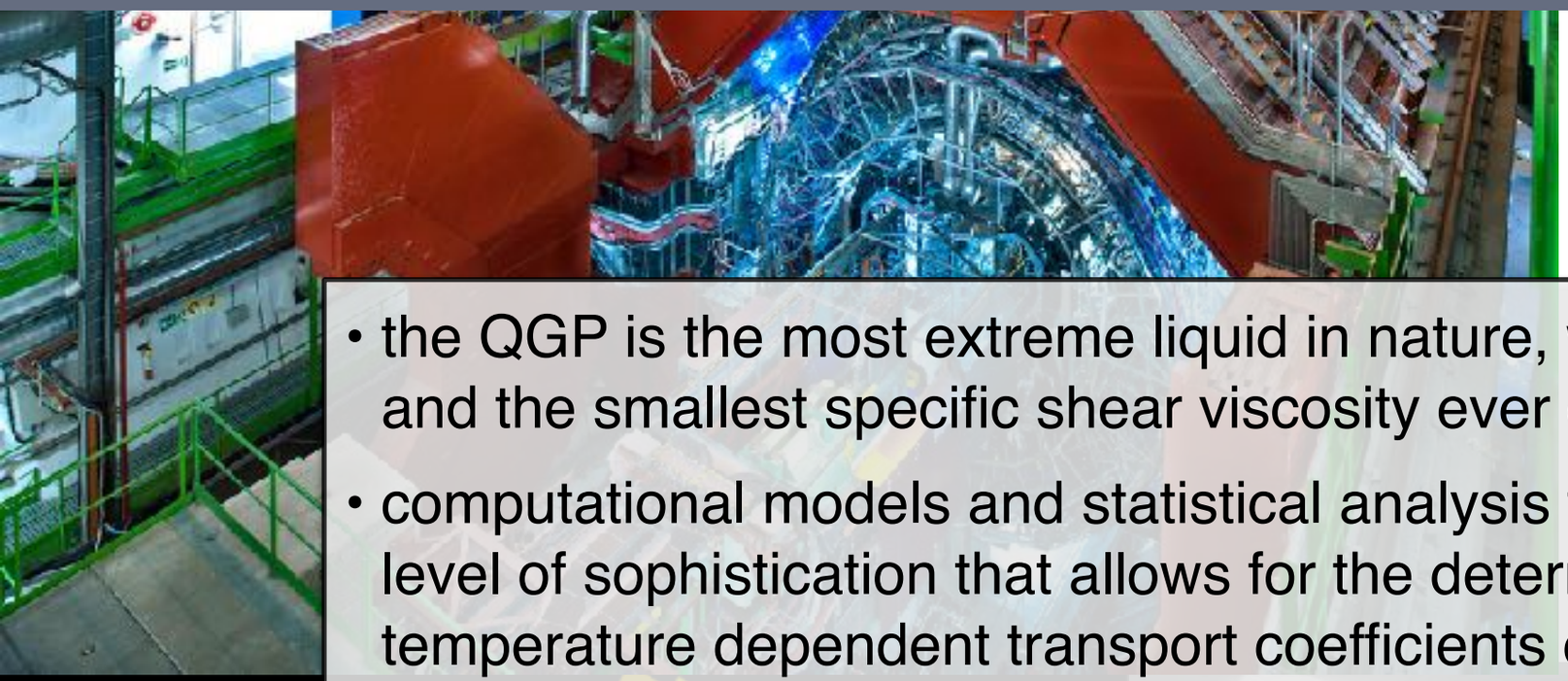


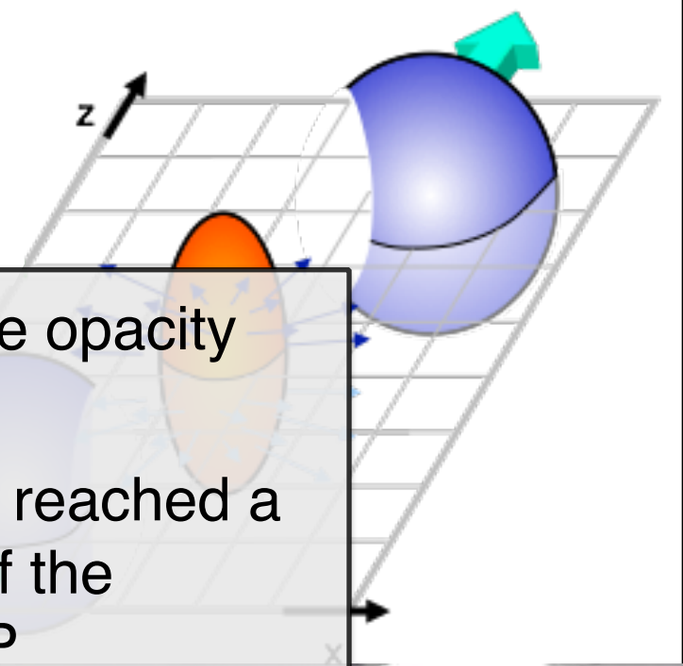
Summary & Outlook

time
MADAI.us

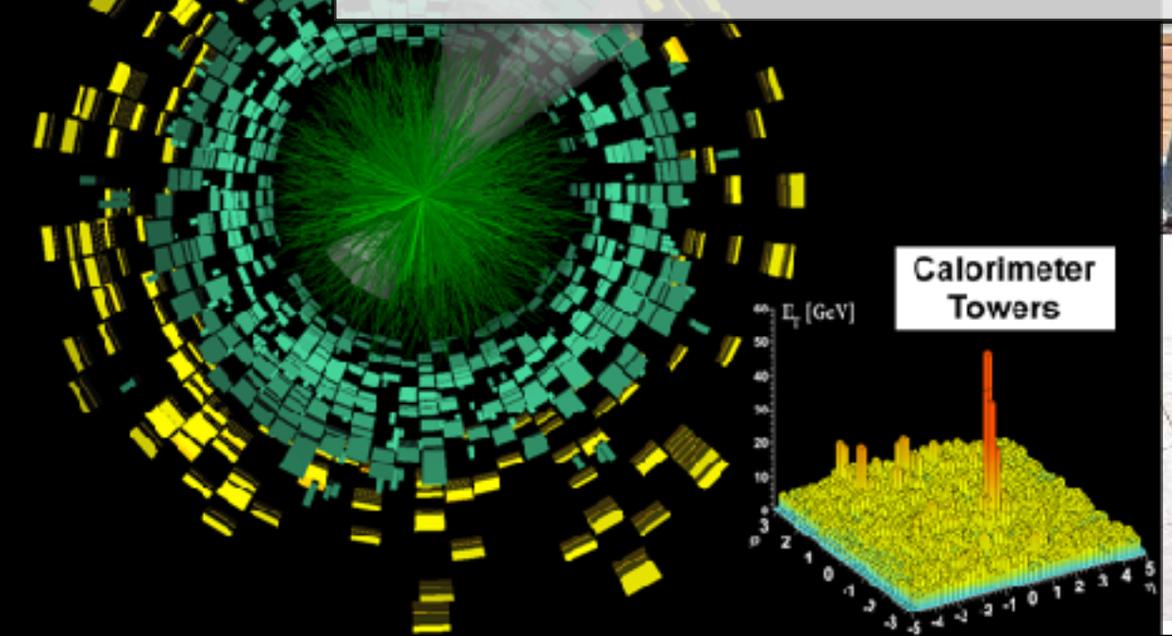

z
y
x


ATLAS
EXPERIMENT
Run 168795, Event 7578342
Time 2010-11-09 08:55:48 CET
Calorimeter Towers
E_T [GeV]





- the QGP is the most extreme liquid in nature, with a large opacity and the smallest specific shear viscosity ever observed
- computational models and statistical analysis tools have reached a level of sophistication that allows for the determination of the temperature dependent transport coefficients of the QGP
- capacity and high-performance computing have been essential for the feasibility and the success of the science!
- frontier science entering the precision era!



Resources

Trento:

- J. Scott Moreland, Jonah E. Bernhard & Steffen A. Bass: Phys. Rev. C 92, 011901(R)
- https://github.com/Duke-QCD/trento

iEbE-VISHNU:

- Chun Shen, Zhi Qiu, Huichao Song, Jonah Bernhard, Steffen A. Bass & Ulrich Heinz: Computer Physics Communications in print, arXiv:1409.8164
- http://u.osu.edu/vishnu/

UrQMD:

- Steffen A. Bass et al. Prog. Part. Nucl. Phys. 41 (1998) 225-370 , arXiv:nucl-th/9803035
- Marcus Bleicher et al. J.Phys. G25 (1999) 1859-1896 , arXiv:hep-ph/9909407
- http://urqmd.org

MADAI Collaboration:

- Visualization and Bayesian Analysis packages
- https://madai-public.cs.unc.edu

Duke Bayesian Analysis Package:

- https://github.com/jbernhard/mtd

this work has been made possible through support from

End

Backup Slides

Data Storage: HPSS at NERSC

HPSS Capabilities:

- theoretical capacity: 200 Petabytes
- buffer (disk) cache: 288 Terabytes
- theoretical maximum throughput: 6.4 GB/sec

Problem:

- network bandwidth/capacity to transfer data to/from NERSC
- not well-interfaced with OSG

Multivariate Output

Scaling of analysis with # of observables:

- independent emulators for each output?
  - neglects correlations among outputs
  - what if # of outputs scales to 100?
  - ▸ training of individual GPE's may become unfeasible and unnecessary in case of strong correlations

Multivariate Output

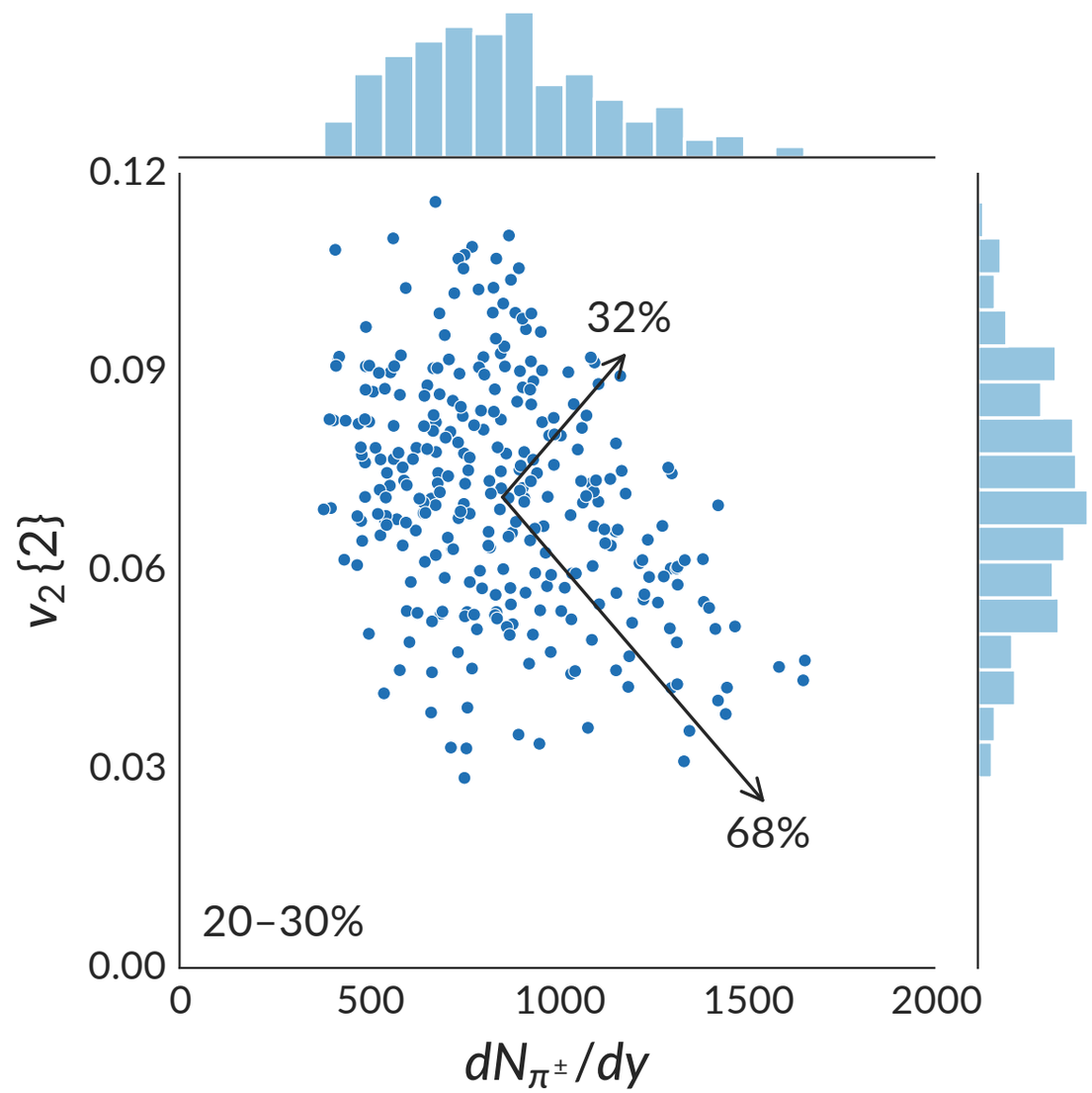
Scaling of analysis with # of observables:

- independent emulators for each output?
  - neglects correlations among outputs
  - what if # of outputs scales to 100?
  - ▸ training of individual GPE's may become unfeasible and unnecessary in case of strong correlations

Principal Components:

- linear combinations of model output
- orthogonal and uncorrelated

⇒ emulate each PC



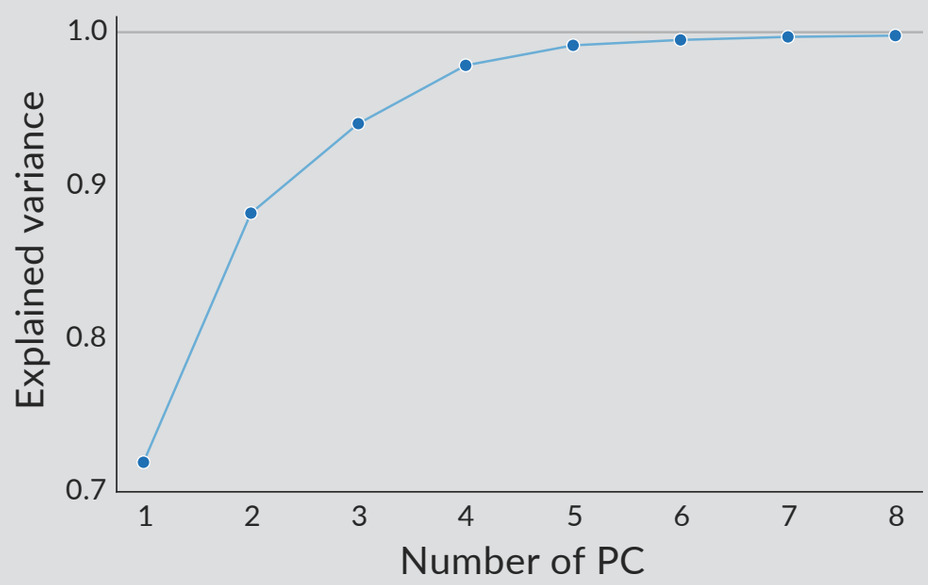
Multivariate Output

Scaling of analysis with # of observables:

- independent emulators for each output?
  - neglects correlations among outputs
  - what if # of outputs scales to 100?
  - ▸ training of individual GPE's may become unfeasible and unnecessary in case of strong correlations

this analysis:

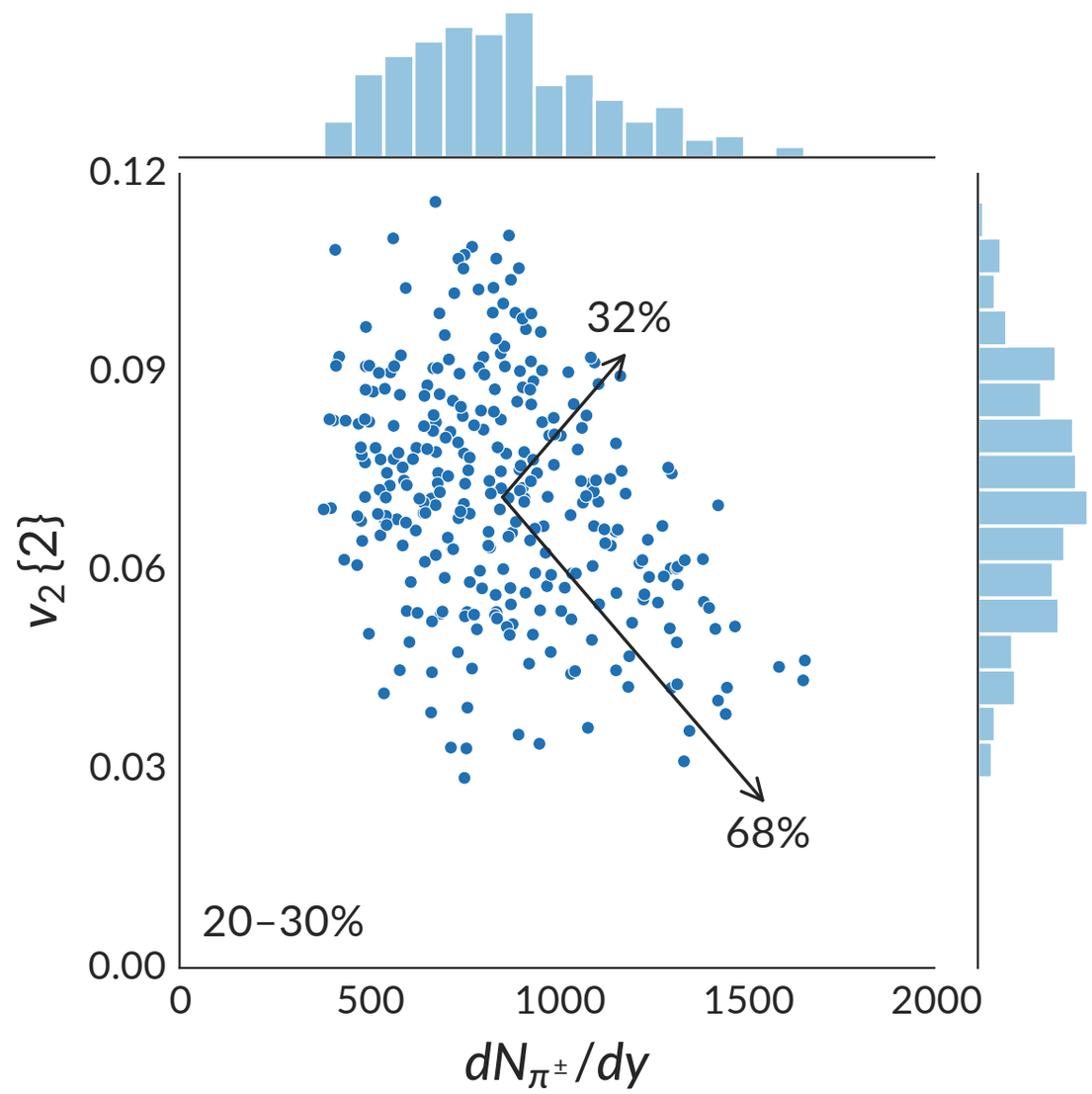
- model outputs are yields, $\langle p_T \rangle$, v_2, v_3 and v_4
- 68 original output dimensions
- 8 principal components used



Principal Components:

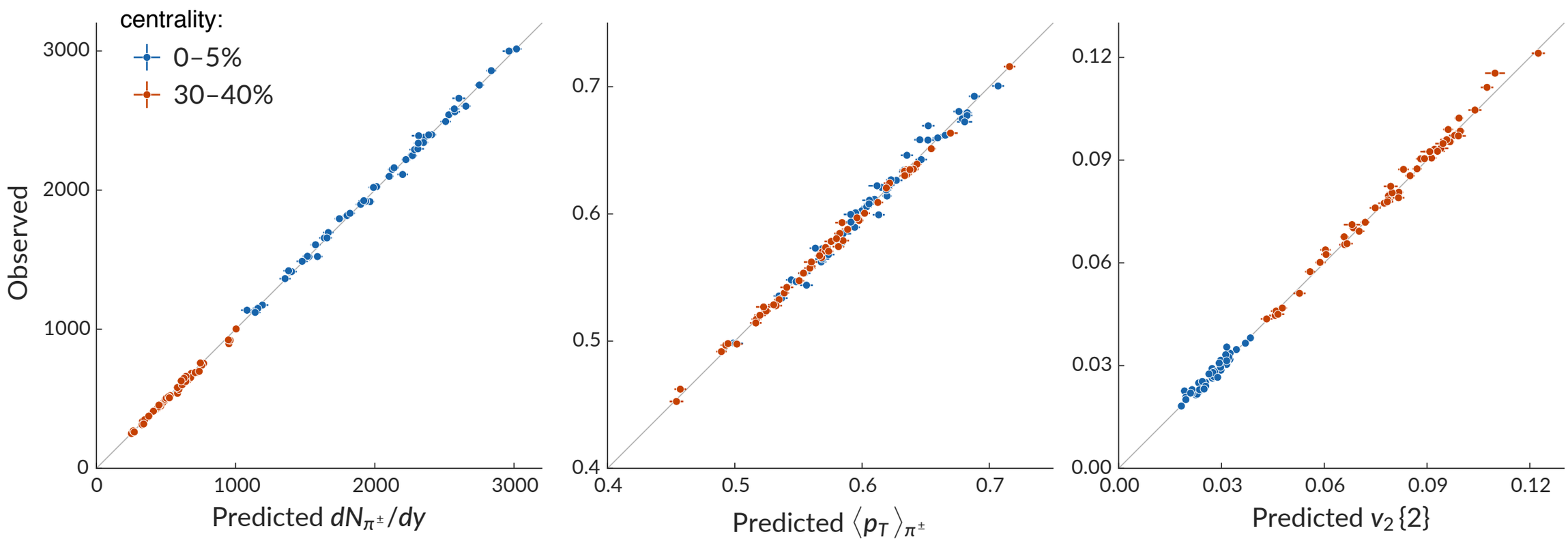
- linear combinations of model output
- orthogonal and uncorrelated

⇒ emulate each PC



Validation

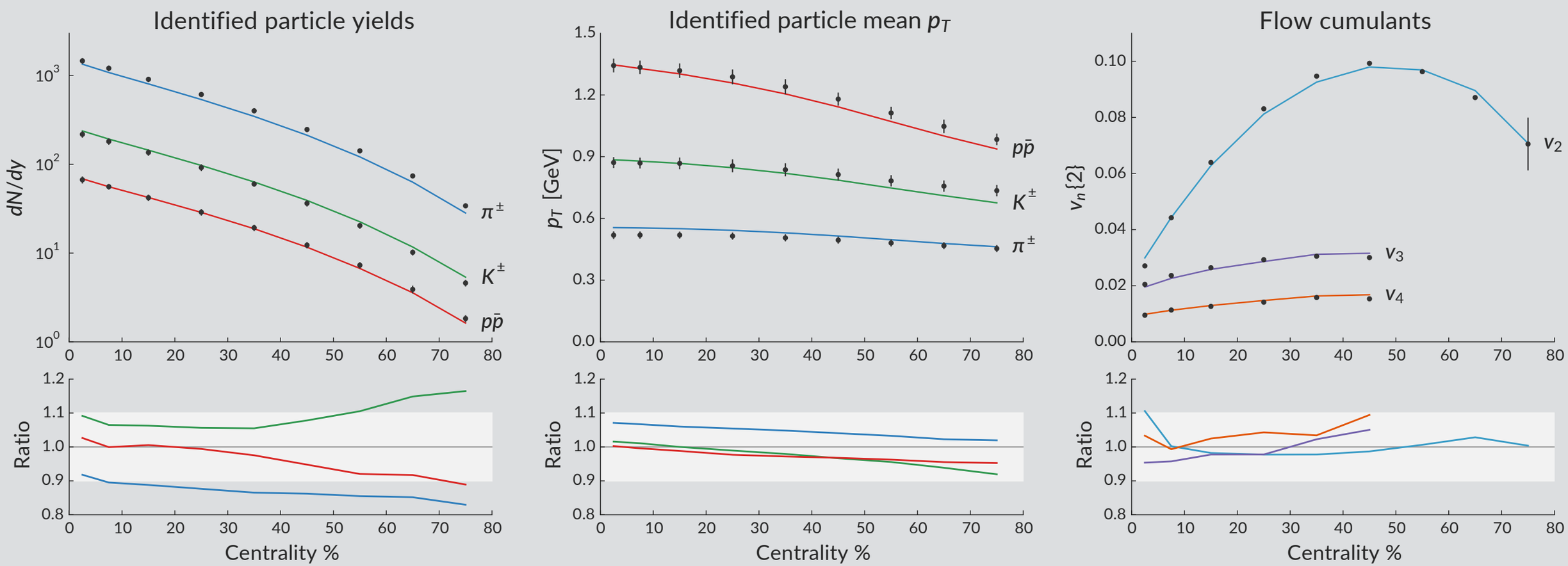
- generate a separate Latin hypercube validation design with 50 points
- evaluate the full physics model at each validation point
- compare physics model output to that of the previously conditioned GP emulators:



- note that since GPEs are stochastic functions, only ~68% of predictions need to fall within 1 standard deviation

Verification: Explicit Model Calculation

- explicit physics model calculations (no emulator) with parameter values set to the maximum of the posterior probability distributions yield excellent agreement with data:



- description of identified particle mean p_T and flow cumulants to within ±10%
- from 0% to 60% centrality, identified particle yields are described within ±15%