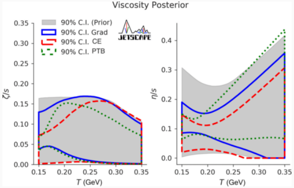
Bayesian inference for relativistic heavy-ion collisions

Exploring nuclear physics across energy scales 2024: intersection between nuclear structure and high-energy nuclear collisions, Peking University & CCAST, Beijing

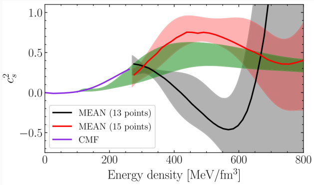
Weiyao Ke
Central China Normal University

April 23, 2024

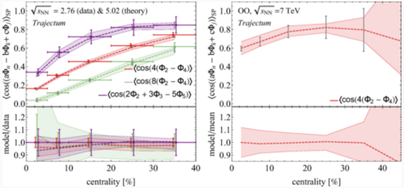
A widely used tool in nuclear physics across energy scales



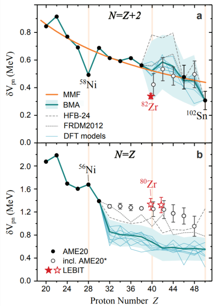
[JETSCAPE, PRC103(2021)054904]



[Kuttan et al, PRL131(2023)202303]



[Trajectum, PRC106(2022)044903]



[BAND, Nat.Phys.17 (2021)1408]

The General Principle is simple: the Bayes' Theorem.

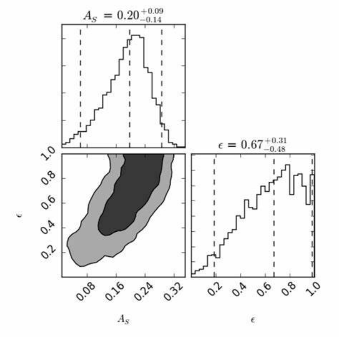
$$\overbrace{P(\mathbf{x}_{\text{true}}|\mathcal{M},\mathbf{y}_{\text{exp}})}^{\text{Posterior}} = \frac{\overbrace{L(\mathbf{y}_{\text{exp}}|\mathcal{M},\mathbf{x}_{\text{true}})}^{\text{Likelihood}}\,\overbrace{P_0(\mathbf{x}_{\text{true}})}^{\text{Prior}}}{\int L(\mathbf{x})P_0(\mathbf{x})d\mathbf{x} \rightarrow \text{Evidence}}$$

How certain we are on a statement given the data and the theory/model? Make predictions and identify anomalies.

Sources of uncertainties

- The exact likelihood function is often unknown. Often molded by multivariate Gaussian:

$$\ln L(x) = -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma| - \frac{1}{2}\Delta\mathbf{y}(x)\Sigma^{-1}\Delta\mathbf{y}^T(x), \quad \Delta\mathbf{y}(x) = \mathbf{y}_{\mathcal{M}}(\mathbf{x}) - \mathbf{y}_{\text{exp}}$$



- Covariance Σ encodes experimental. theory/model uncertainty

$$\Sigma_{ij} = \underbrace{\delta_{ij}[(\delta y^{\text{exp}}_{\text{stat}})^2_i + (\delta y^{\text{exp}}_{\text{sys},0})^2_i]}_{\text{Uncorrelated uncertainty}} + \underbrace{(\delta y^{\text{exp}}_{\text{sys}})_i(\delta y^{\text{exp}}_{\text{sys}})_j c_{ij}}_{\text{Correlated uncertainty}}$$

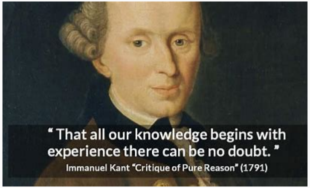
$$+ \underbrace{\Sigma^{\text{emulator}}_{ij}}_{\text{Interpolation}} + \underbrace{\Sigma^{\text{theory}}_{ij}}_{\text{theory}}$$

- Uncertainty of nuisance parameters propagated by marginalization

$$P(p_1) = \int P(p_1, p_2, p_3, \cdots) dp_2 dp_3 \cdots$$

The prior and the uncertainty

Thank Dr. Yi Yin for reminding me of an interesting connection. He asked me "Is the following statement some form of Bayes' theorem? and 04/22 is Kant's 300's birthday".



[http://www.gutenberg.org]

"Kant maintained, knowledge must rest on judgments that are a priori, for it is only as they are separate from the contingencies of experience that they could be necessary and yet also synthetic—i.e., so that the predicate term contains something more than is analytically contained in the subject."

—Period of the three Critiques of Immanuel Kant, Encyclopaedia Britannica

While in Bayes' theorem, the prior can be either analytic (e.g., causality bounds on the transport parameters) or empirical (old data) and even biased. **So the choice of prior also affects uncertainty quantification.**

Outline

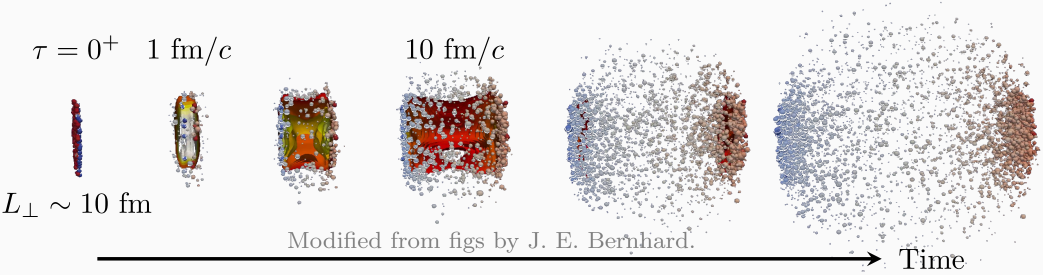
Review HIC modeling and uncertainties

Bayesian analysis of bulk matter properties

Information gain and the prior

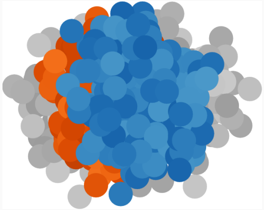
Review HIC modeling and uncertainties

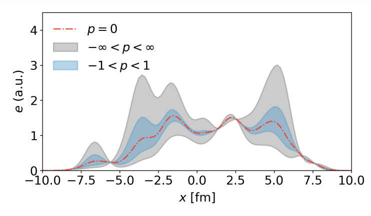
The spirit of a multi-stage modeling



- $\tau = 0^+$: almost instantaneous energy deposition from the two nuclei pancake.
- $\tau < \tau_R$: longitudinal expansion driven, pre-equilibrium stage.
- $\tau \sim L_\perp$: hydrodynamic pressure driven stage.
- $\tau_R \nabla_\perp \sim 1$, $T < T_c$: hadronic transport stage.

From the Glauber model to the initial energy deposition model





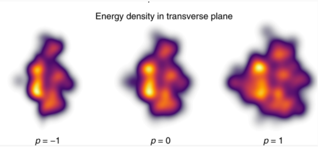
- Glauber model determines the probability for AB to collide at given impact parameter.
- Map participant nucleon density $T_{A.B}$ to $T^{\mu\nu}(\mathbf{x}, \eta_s)$ at $\tau = 0^+$. With boost invariance:

$$\frac{dE_T}{dx_\perp^2\, d\eta_s}(x_\perp, \eta_s = 0) = \mathrm{Norm} \times f(T_A(x_\perp), T_B(x_\perp))$$

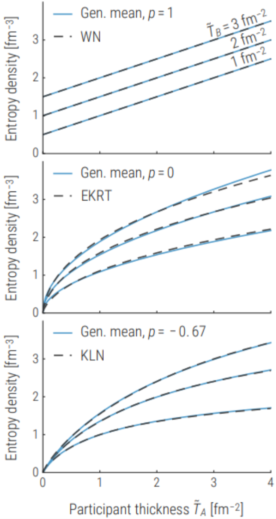
- Direct related to nuclear structure, but often heavily relies on modeling. TRENTo is one of the many ansatz:

$$f(T_A, T_B) = \left(\frac{T_A^p + T_B^p}{2}\right)^{1/p}$$

A motivation for the TRENTo ansatz



To mimic and test different scaling behavior of energy deposition.



- Wounded nucleon model

$$\frac{dS}{dy\, d^2r_\perp} \propto \tilde{T}_A + \tilde{T}_B$$

- EKRT model PRC 93, 024907 (2016)

  after brief free streaming phase

$$\frac{dE_T}{dy\, d^2r_\perp} \sim \frac{K_{sat}}{\pi} p_{sat}^3(K_{sat}, \beta; T_A, T_B)$$

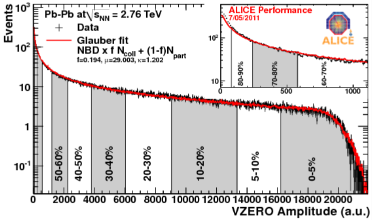
- KLN model PRC 75, 034905 (2007)

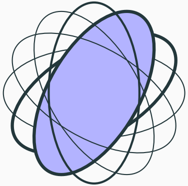
$$\frac{dN_g}{dy\, d^2r_\perp} \sim Q_{s,min}^2 \left[2 + \log\left(\frac{Q_{s,max}^2}{Q_{s,min}^2}\right)\right]$$

[Slide from J. S. Moreland]

Some new worries of initial-condition model

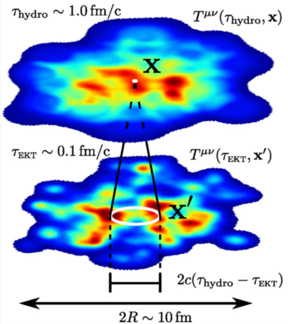
- Centrality (quantiled total inelastic cross-section of AB) is a main handle on the impact parameter $\mathbf{b}$ and geometry!
- Calculation of $\sigma_{AB} = \int d^2\mathbf{b} P_{\rm coll}(\mathbf{b})$ relies on Glauber model.





- In the derivation of Glauber formula (σ_{pA}), one assumes that $\rho_A(r_1, \cdots, r_n) \approx \prod_i \rho_1(r_i)$. Then, $P_{\rm coll}(\mathbf{b})$ can be expressed as functional of nuclear thickness function.
- But the assumption may not work well for a wave function like $|\psi\rangle = \sum_{\vec{\beta}} C(\vec{\beta})|\phi; \vec{\beta}\rangle$.

Pre-equlibirum stage



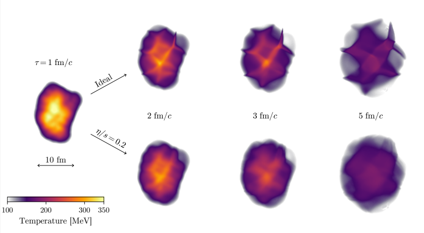
Fast longitudinal expansion dominate over collisions $\frac{\tau_R^{-1}}{\tau^{-1}} \ll 1$

- Zeroth order theory: use collisionless Boltzmann equation to evolve $T^{\tau\tau}(x_T, \tau = 0^+)$ to $T^{\mu\nu}(x_T, \tau_{\rm hydro})$ [PRC91(2015)064906].
- QCD Effective Kinetic Theory. Brings the system closer to local equilibrium. [Kurkela et al, PRL122(2019)122302]
- Anisotropic hydro: gradient expansion around an anisotropic distribution function, extend hydro theory towards early time.

  [Florkowski Ryblewski, Martinez, Strickland, McNelis, Bazow, Heinz et al.]

Main uncertainty: how the system approach to hydrodynamization, and the matching time scale $\tau_{\rm hydro}$.

Applications and developments of the relativistic hydrodynamics



Sensitivity of flow development to viscosity.

The hydrodynamic stage

$$T^{\mu\nu} = eu^\mu u^\nu + (u^\mu u^\nu - g^{\mu\nu})(P(e) + \Pi) + \pi^{\mu\nu}$$

$$\dot{\pi}^{\mu\nu} = -\frac{\pi^{\mu\nu} + 2\eta\partial^{\langle\mu}u^{\nu\rangle}}{\tau_\pi}\cdots,$$

$$\dot{\Pi} = -\frac{\Pi + 2\zeta\nabla\cdot u}{\tau_\Pi}\cdots$$

Main uncertainty: validity of the gradient expansion, higher-order transport coefficients.

Particlization of fluid element

An ill-defined inverse problem: how to reconstruct the distribution function $f_{\rm eq}(p,x) + \delta f_i(p,x)$ for each specie of hadrons from $T^{\mu\nu}(x)$ and conversed currents.

Different models result in different momentum dependence and hadron chemistry.

- Grad 14-moment expansion $\delta f(p) \propto A_\pi \pi^{\mu\nu} p_{\langle\mu} p_{\nu\rangle} + \Pi(A_T m_i^2 + A_E (p\cdot u)^2)$
- Using $1^{\rm st}$-order Chapman-Enskog solution to RTA Boltzmann equation.

$$\delta f \propto \frac{\pi_{\mu\nu} p^{\langle\mu} p^{\nu\rangle}}{2\beta_\pi (p\cdot u) T} + \frac{\Pi}{\beta_\Pi}\left(\frac{\mathcal{F}(p\cdot u)}{T^2} - \frac{\Delta_{\mu\nu} p^\mu p^\nu}{3(p\cdot u)T}\right)$$

- Rotate, stretch, and rescale the equilibrium distribution (Pratt-Torrieri-Bernhard/McNelis)

$$f_{\rm eq} + \delta f = \mathcal{Z} f_{eq}\left(p^i \to [(1+\frac{\Pi}{3\beta_\Pi})\delta_{ij} + \pi^{ij}]p_j,\ T \to T + \beta_\Pi^{-1}\Pi\mathcal{F}\right)$$

How to match a hadron resonance gas EoS (missing states? resonance width? $f_0(500)$?) to lattice QCD EoS near T_c.

Hadronic afterburner

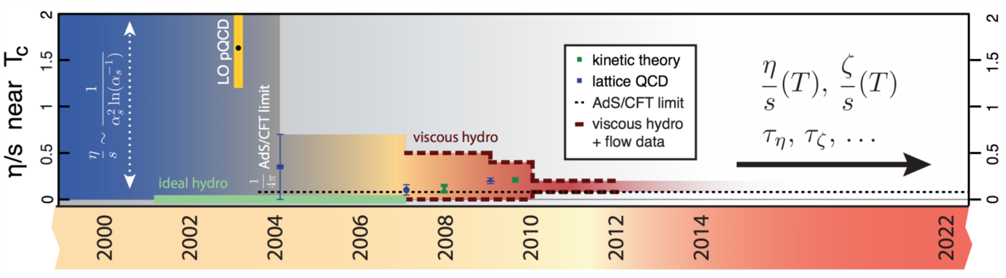
The particlized system is then solved by a hadronic transport equation

$$\frac{\partial}{\partial t} f_i + v \cdot \nabla_x f_i - \nabla_x E_i \cdot \nabla_p f_i = \sum_{i,j,\cdots} C[f_i, f_j, \cdots].$$

- The potential term is often turned-off in the afterburner for ultra-relativistic collisions.
- Parametrization of some hadronic cross-sections.

Despite all these uncertainties, Bayesian analysis allows us to learn something with certainty.

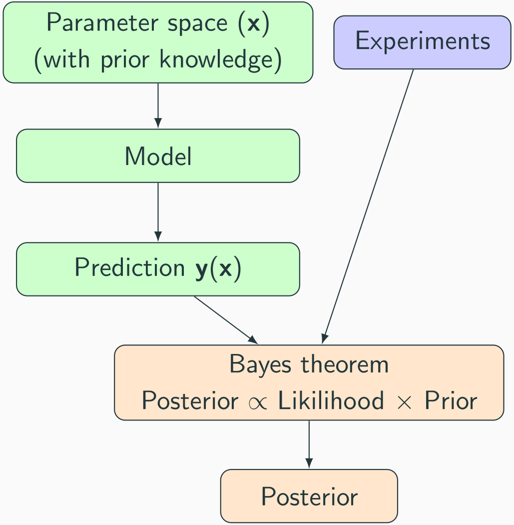
Essential for making quantitative progress



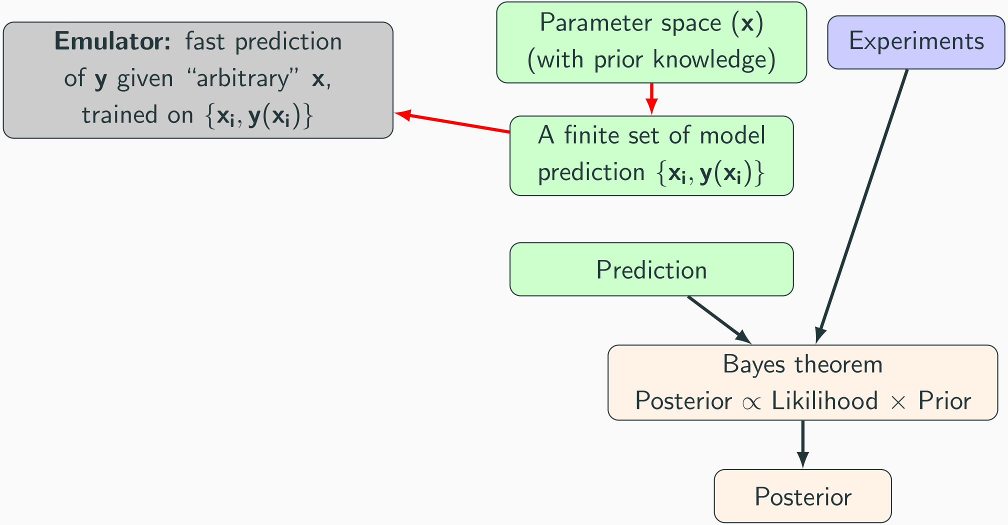
- 2000s: order of magnitude from pQCD.
- 2004: strongly coupled theory $\eta/s = 1/(4\pi) + \cdots$.
- 2006-2013: eyeball fit with viscous hydro $(\eta/s)_{\text{eff}} = 0.1$–$0.2$
- 2013–: Bayesian analysis of EoS and η/s with uncertainty.
- 2016–: Temperature-dependent shear and bulk viscosity $\eta/s(T), \zeta/s(T)$.
- 2021–: Bayesian model averaging, **model improvement, novel observables**.

Bayesian analysis of bulk matter properties

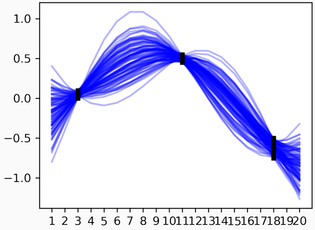
For simple models

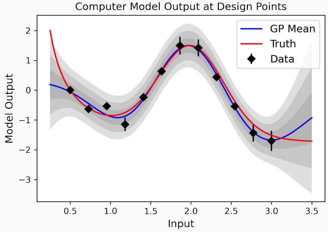


For computationally intensive model



A class of non-parametric emulators: Gaussian emulators





$$y(x) = \mu(x) \pm \sigma(x)$$

Gaussian emulators: to interpolate the value of y at a new input x, it makes use of how $y(x)$ correlates with all the training data $y(x_{T,i}) = y_{T,i}, \quad i = 1, 2, \cdots m$.

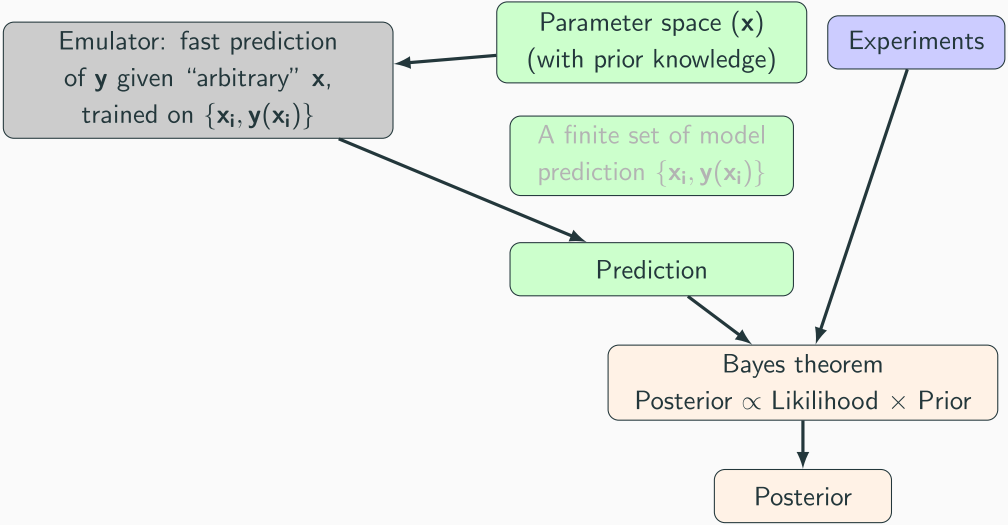
- The distribution of $y(x)$ is also a normal distribution

$$\begin{aligned} P(y(x)) &= \mathcal{N}(\mu, \sigma^2) \\ \mu(x) &= K(x, \vec{x}_T)K^{-1}(\vec{x}_T, \vec{x}_T)\vec{y}_T, \\ \sigma^2(x) &= K(x, x) - K(x, \vec{y}_T)K^{-1}(\vec{y}_T, \vec{y}_T)K(\vec{y}_T, x)) \end{aligned}$$

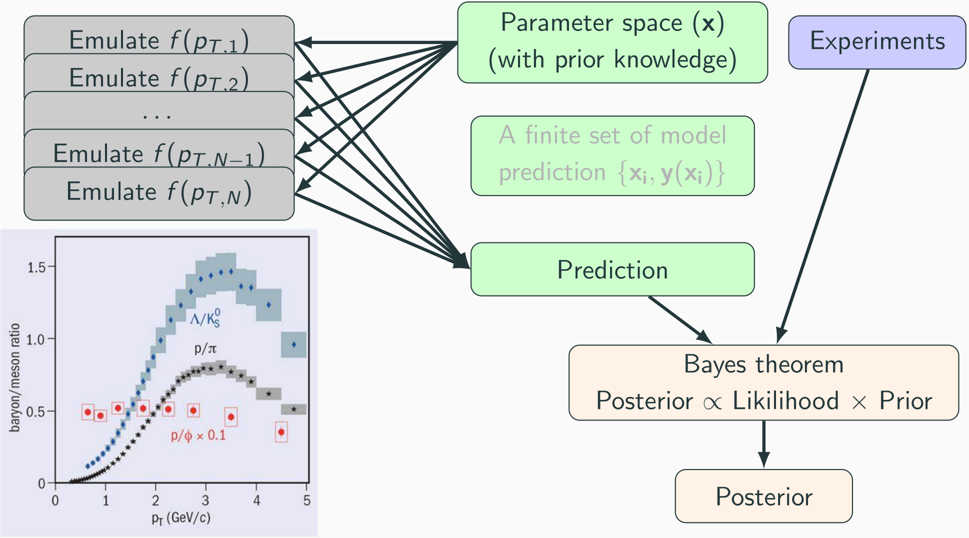
- Interpolation uncertainty

$$\Sigma_{ij} = \delta_{ij}[(\delta y^{\rm exp}_{\rm stat})^2_i + (\delta y^{\rm exp}_{\rm sys,0})^2_i] + (\delta y^{\rm exp}_{\rm sys})_i(\delta y^{\rm exp}_{\rm sys})_j c_{ij} + \Sigma^{\rm emulator}_{ij} + \Sigma^{\rm theory}_{ij}$$

The workflow for analyzing a computationally intensive model



For an array of observables



Dimensional reduction via the Principal-Component Analysis (PCA)

Vector observables at m input points.

$$\vec{y}_T(\vec{x}_{T,1}) = (y_1^{[1]}, y_2^{[1]}, \cdots y_N^{[1]}),$$
$$\cdots$$
$$\vec{y}_T(\vec{x}_{T,m}) = (y_1^{[m]}, y_2^{[m]}, \cdots y_N^{[m]}),$$

Quantify obs variations w.r.t. parameters

$$[\mathrm{Cov}]_{ij} = \frac{1}{m} \sum_{s=0}^{m} \left(y_i^{[s]} - \bar{y}_i\right) \left(y_j^{[s]} - \bar{y}_j\right),$$
$$\bar{y}_j = \frac{1}{m} \sum_{s=0}^{m} y_j^{[s]}$$

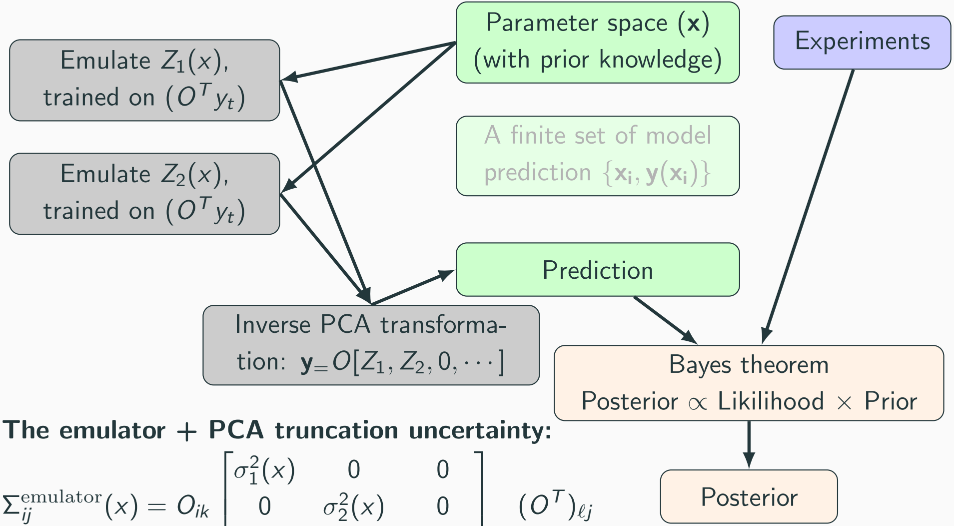
PCA: orthogonal transformation of the basis to an empirical basis

$$\left(O^T\right)_{ik} [\mathrm{Cov}]_{kl} \left(O\right)_{lj} = \begin{bmatrix} \lambda_1 & 0 & \vec{0} \\ 0 & \lambda_2 & \vec{0} \\ \vec{0}^T & \vec{0}^T & \cdots \end{bmatrix} \xrightarrow{\lambda_1 > \lambda_2 \gg \lambda_3 \cdots} \begin{bmatrix} \lambda_1 & 0 & \vec{0} \\ 0 & \lambda_2 & \vec{0} \\ \vec{0}^T & \vec{0}^T & \mathbf{0} \end{bmatrix}$$

- In the new basis, the new observables (features) are $Z_i = \left[O^T\right]_{ij} y_j$.
- Z_i, Z_j's change with x are linearly independent (there may be non-linear correlations).

Dimensional reduction: $Z_i \approx [Z_1(x), Z_2(x), 0, \cdots]$

The workflow of the emulator+PCA-assisted Bayesian analysis

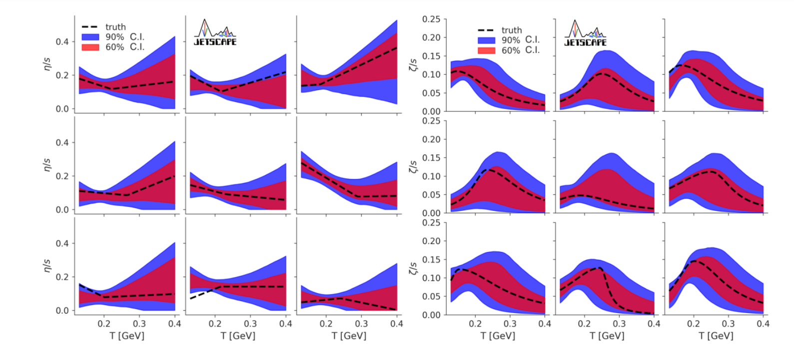


The emulator + PCA truncation uncertainty:

$$\Sigma_{ij}^{\text{emulator}}(x) = O_{ik} \begin{bmatrix} \sigma_1^2(x) & 0 & 0 \\ 0 & \sigma_2^2(x) & 0 \\ 0 & 0 & \lambda_3 \cdots \end{bmatrix}_{k\ell} (O^T)_{\ell j}$$

Diagnosis tools and validation test

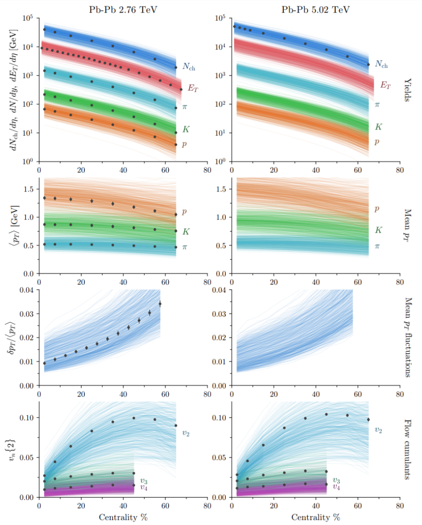
- Modern analysis framework often provides diagnosis tools to access the performance of each module (emulators, PCA, etc).
- A final validation test for the final analysis ∇ [JETSCAPE PRC103(2021)054904].



Calibrate the multi-stage model

Energy deposition model
+ Preequiblirum dynamics
+ Relativistic viscous hydrodynamics
+ Freeze-out to hadron resonance gas
+ Hadronic transport model

Parameter	Description	Range
Norm	Normalization factor	8–20 (2.76 TeV) 10–25 (5.02 TeV)
p	Entropy deposition parameter	$-1/2$ to $+1/2$
σ_{fluct}	Multiplicity fluct. std. dev.	0–2
w	Gaussian nucleon width	0.4–1.0 fm
d^3_{min}	Minimum nucleon volume	0–1.7 fm^3
τ_{fs}	Free streaming time	0–1.5 fm$/c$
η/s hrg	Const. shear viscosity, $T < T_c$	0.1–0.5
η/s min	Shear viscosity at T_c	0–0.2
η/s slope	Slope above T_c	0–8 GeV^{-1}
η/s crv	Curvature above T_c	-1 to $+1$
ζ/s max	Maximum bulk viscosity	0–0.1
ζ/s width	Peak width	0–0.1 GeV
ζ/s T_0	Peak location	150–200 MeV
T_{switch}	Particlization temperature	135–165 MeV



[Jonah E. Bernhard Ph.D. dissertation]

Calibrate the multi-stage model

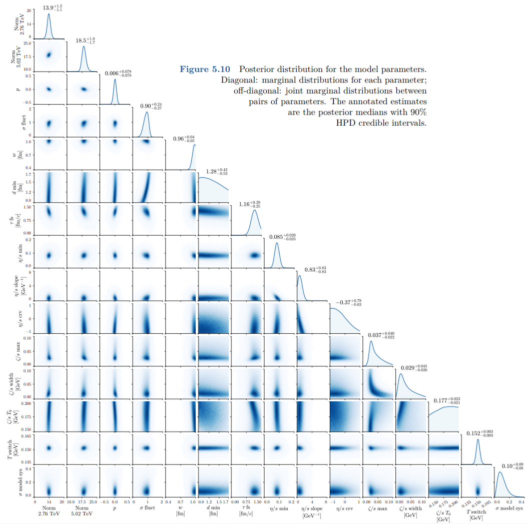
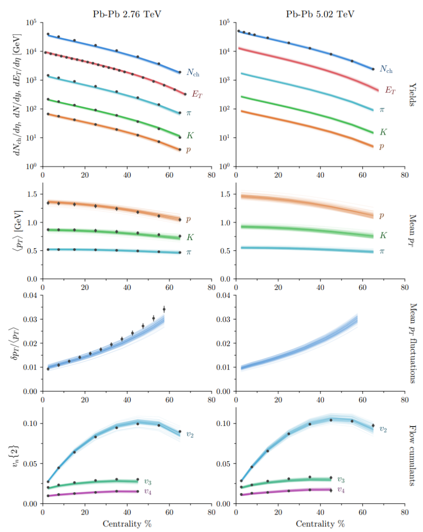
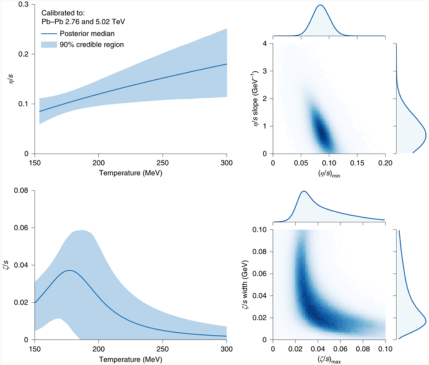


Figure 5.10 Posterior distribution for the model parameters. Diagonal: marginal distributions for each parameter; off-diagonal: joint marginal distributions between pairs of parameters. The annotated estimates are the posterior medians with 90% HPD credible intervals.



[Jonah E. Bernhard Ph.D. dissertation]

Marginalized posterior distribution of shear and bulk viscosity



[JE Bernhard, JS Moreland, SA Bass, Nat. Phys. 15(2019)1113–1117.]

$$\eta/s = (\eta/s)_{\text{min}} + (\eta/s)_{\text{slope}}(T - T_c)\left(\frac{T}{T_c}\right)^{(\eta/s)_{\text{curv}}}$$

$$\zeta/s = \frac{(\zeta/s)_{\text{max}}}{1 + (T - (\zeta/s)_{T_0})^2/(\zeta/s)^2_{\text{width}}}$$

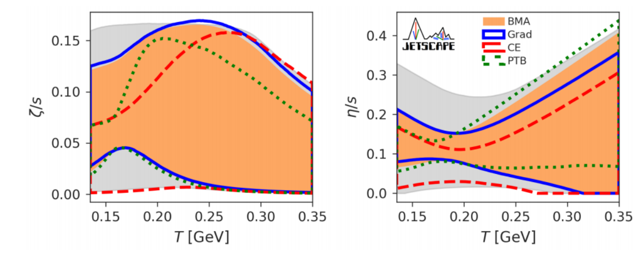
With a high degree of confidence:

- Quark-gluon plasma (QGP) is strongly coupled $\eta/s = (1 \cdots 2)/(4\pi)$.
- QGP has a nonzero bulk viscosity.

Model uncertainty and Bayesian model averaging

Use **Bayesian model averaging (BMA)** to take into account the uncertainty in the fluid-cell particlization procedure:

$$P_{\mathrm{BMA}}(x|y_{\mathrm{exp}}, \{M_i\}) = \sum_i \underbrace{P(x|y_{\mathrm{exp}}, M_i)}_{\text{Posterior for model ``}i\text{''}} \times \underbrace{P(y_{\mathrm{exp}}|M_i)}_{\text{Evidnece of model `}i\text{''}}$$



After model averaging (orange bands), the BMA posterior is dominated by the one with the highest evidence.

[JETSCAPE Collaboration, Phys.Rev.Lett. 126 (2021) 24, 242301]

Information gain and the prior

Quantify the information gain

To quantify the difference between posterior and prior

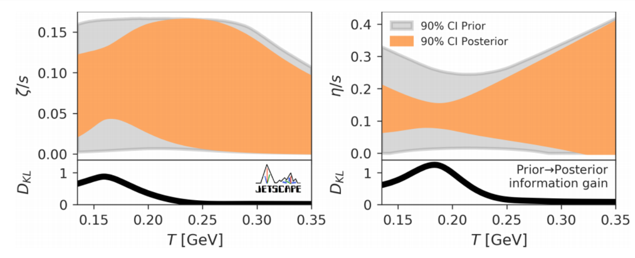
- Use the "Kullback–Leibler divergence" (KL divergence, D_{KL}) to measure the "distance" between two distributions P_1 and P_2

$$D_{\mathrm{KL}}(P_1 \| P_2) \equiv \int dx P_1(x) \ln \frac{P_1(x)}{P_2(x)}, \text{ we take } P_1 = \text{Posterior}, \ P_2 = \text{Prior}.$$

- If $D_{\mathrm{KL}} = 0$, then the posterior is the same as our prior belief, nothing new...
- $D_{\mathrm{KL}} > 0$ signatures information gain from experimental data.

Quantify the information gain

Little sensitivity to viscosity at high T ($D_{KL}(T > 0.25\mathrm{GeV}) \approx 0$). Most information gains in low-temperature regions. **Possible reasons:**

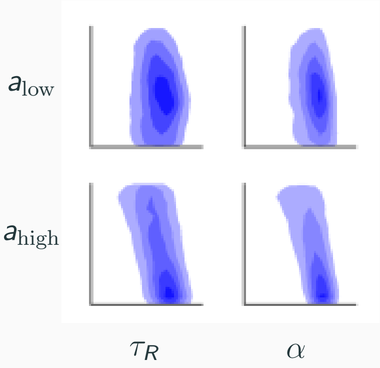


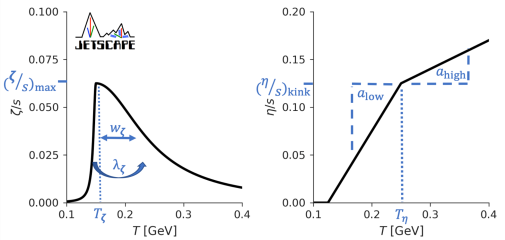
- Some observable only sensitive to "averaged η/s" [J-F Paquet, SA Bass, PRC102(2020)014903]

$$(\eta/s)_{\mathrm{eff}} = \frac{\int_{T_{\mathrm{sw}}}^{T_{\mathrm{max}}} \eta/s(T)/T^{\alpha} dT}{\int_{T_{\mathrm{sw}}}^{T_{\mathrm{max}}} 1/T^{\alpha} dT}$$

- The high-temperature behavior of η and ζ is strongly correlated with other parameters.

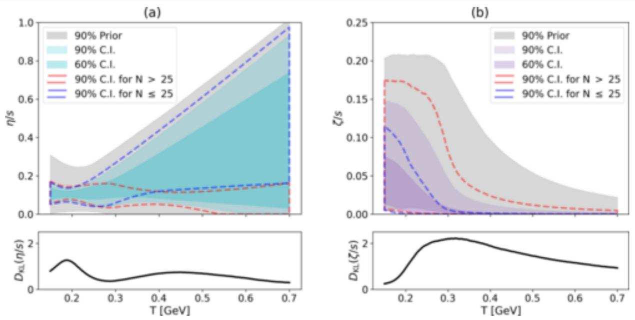
Correlation between high-temperature viscosity with pre-equilibrium parameter





- In JETSCAPE, the free-streaming model in the pre-eq stage does not drive the system close to equilibrium.
- Amount of off-equilibrium effects depend on the matching time between pre-eq dynamics & hydro $\tau_{\rm hydro} = \tau_R \left(\frac{e}{e_0}\right)^{\alpha}$.
- Observe a strong correlation between high-T η/s and the matching time scale in the posterior.

Improvements of the pre-equilibrium models



[Liyanage, Sürer, Plumlee, Wild, Heinz RPC108(2023)054905]

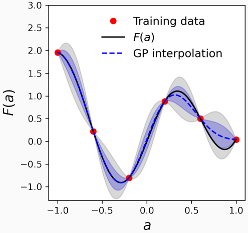
- In viscous anisotropic hydrodynamics (VAH), matching uncertainty between pre-eq dynamics and hydro is reduced. Result in a stronger constrain of η/s and ζ/s at high T.
- Worth checking in other models, such as using the QCD effective kinetic theory to bridge the gap.

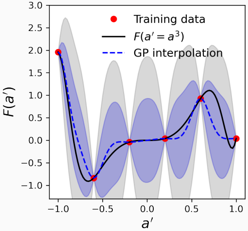
Functional prior

Suppose we parametrize some temperature-dependent quantities:

$$\frac{\eta}{s}(T) = \left[\frac{\eta}{s}\right]_{\mathrm{kink}} + \Theta(T - T_\eta)a_{\mathrm{high}}(T - T_\eta) + \cdots\,, \qquad \frac{\hat{q}(T)}{T^3} = \frac{A}{1 + C(T/T_c)^n}$$

- Contains **long-range correlations**: every parameter affects a large range of temperature.
- Contains a high-degree of **nonlinearity**



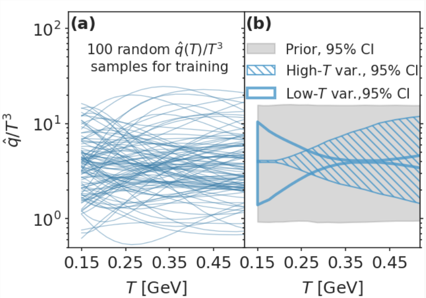


Unnecessary complication for the machine-learning tools

$$a, b, c \xrightarrow[\text{parametrization}]{\text{Nonlinear}} f(a, b, c)$$

$$f(a, b, c) \xrightarrow[\text{models}]{\text{well-behaved}} \text{Observables}$$

Use random field as the prior of an unknown functional



Use a random function/random field to model the prior of an unknown function.

$$\mathrm{Prior}[F(x)] = e^{-\frac{1}{2}\int dxdx'(F(x)-\mu(x))C^{-1}(x,x')(F(x)-\mu(x))}$$

Advantage: suppresses long-range correlation.

Small degree of non-linearity.

$$\mathrm{Posterior}[F(x)] = \mathrm{Prior}[F(x)]e^{-\frac{1}{2}[\mathcal{M}(F)-\mathrm{Exp}]_i\Sigma_{ij}^{-1}[\mathcal{M}(F)-\mathrm{Exp}]_j}$$

Marginalization on the distribution of $F(x)$ at $x = x_0$

$$P[y = F(x_0)] = \int \mathcal{D}F(x)\delta(F(x_0) - y)\mathrm{Posterior}[F(x)]$$

Inference formulated as a field theory problem

- Field-theory approach to infer an 1D probability distribution from N independent draws

[PRL77(1996)4693, hep-ph/9808474v1]

Field Theories for Learning Probability Distributions

William Bialek,[1] Curtis G. Callan,[2] and Steven P. Strong[1]
[1]*NEC Research Institute, 4 Independence Way, Princeton, New Jersey 08540*
[2]*Department of Physics, Princeton University, Princeton, New Jersey 08544*
(Received 25 July 1996)

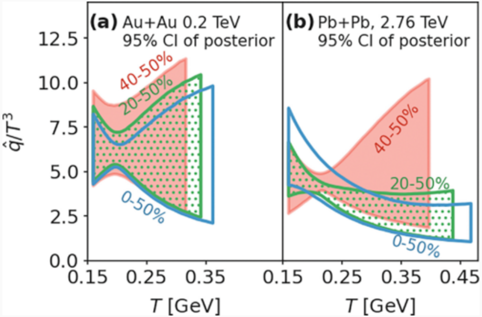
Imagine being shown N samples of random variables drawn independently from the same distribution. What can you say about the distribution? In general, of course, the answer is nothing, unless you have some prior notions about what to expect. From a Bayesian point of view one needs an *a priori* distribution on the space of possible probability distributions, which defines a scalar field theory. In one dimension, free field theory with a normalization constraint provides a tractable formulation of the problem, and we discuss generalizations to higher dimensions. [S0031-9007(96)01804-2]

Functional statistical inference of parton distributions

Vipul Periwal
Department of Physics, Princeton University, Princeton, New Jersey 08544

Bialek, Callan and Strong have recently given a solution of the problem of determining a continuous probability distribution from a finite set of experimental measurements by formulating it as a one-dimensional quantum field theory. This report applies an extension of their formalism to the inference of functional parton distributions from scattering data.

- Extract the temperature dependence of the jet transport parameter:



[M. Xie, WK, H. Zhang, X.-N. Wang PRC108(2023)L011901.]

- For applications in astrophysics ▸ link ▸ link

Summary

- With complex models as used in HIC, the importance of uncertainty quantification is widely accepted. Bayesian analysis has been adopted in many works.
- Understanding the source of uncertainty is the first step of performing a Bayesian analysis.
- Simple machine learn tools is integrated to accelerate Bayesian analysis:
  Gaussian Process-based model emulators + PCA dimensional reduction.
- Uncertainty quantification cannot tell something we don't "know", theory/model improvement is always important.
- The prior choice is nontrivial, especially for functional inference. Random fields may be a reasonable choice.

Questions?