Heavy quark transport coefficient from a Bayesian analysis

Yingru Xu

Duke University

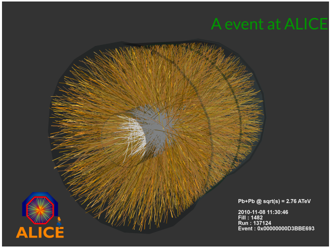
yx59@phy.duke.edu

May 5, 2017

In collaboration with :
Marlene Nahrgang
Jonah E. Bernhard
Shanshan Cao
Steffen A. Bass

This work has been supported by the U.S Department of Energy under grand DE-FG02-05ER41367. Computational resources were provided by the Open Science Grid (OSG).

Heavy quarks in heavy-ion collision – a very complex system



outputs (y)
R_{AA}, v_2

Models
transport model:
Langevin, Boltzmann

calibration parameters (x)
transport coefficients

I: The modeling problem

model eg. HQ transport → **model parameters** α, β, γ eg. $D_s 2\pi T(\alpha, \beta)$ ↔ **experimental obs** eg. R_{AA}, v_2

$$\vec{y} = \mathrm{PM}(\vec{x}) + \delta_{\mathrm{error}} \tag{1}$$

- We have a goal: to learn the unknowns (this case: heavy quark transport coefficients $\vec{x} = \hat{q}, D_s 2\pi T$) from the observables $\vec{y}_{\mathrm{exp}}$
- We have some idea about $\vec{x}$ (or not?), captured in our prior distribution $P(\vec{x})$
- We also have some idea how the outputs depends on the inputs $\vec{x}$, given in the form of likelihood $P(\vec{y}|\vec{x})$(can be calculated by our model)
- Bayesian inference $\Rightarrow$ posterior distribution:

$$P(\vec{x}|\vec{y}) = \frac{P(\vec{y}|\vec{x}) \cdot P(\vec{x})}{\int P(\vec{y}|\vec{x}) \cdot P(\vec{x}) d\vec{x}} \propto P(\vec{y}|\vec{x}) \cdot P(\vec{x}) \tag{2}$$

Overview

Model parameters

$D_s 2\pi T(\alpha, \beta, ...)$

Latin hypercube

Physics model

HQ in-medium evolution

Gussian process emulator

Non-parametric interpolation

fast surrogate

MCMC

random walk through

parameter space

Experiemental data

-LHC + RHIC: R_{AA}, v_2

Bayes' Theorem

$\text{posterior} \propto \text{likelihood} \times \text{prior}$

After equilibrium

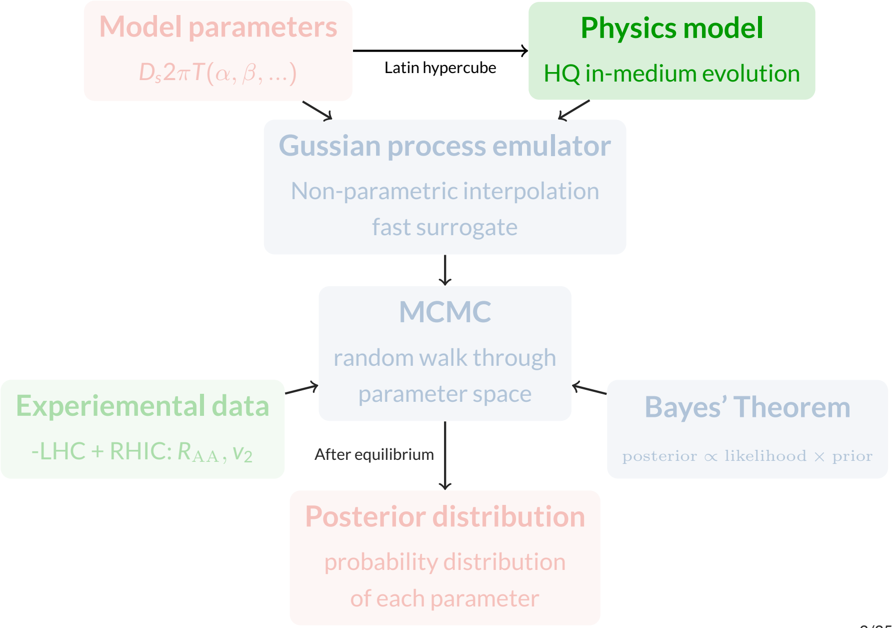
Posterior distribution

probability distribution

of each parameter

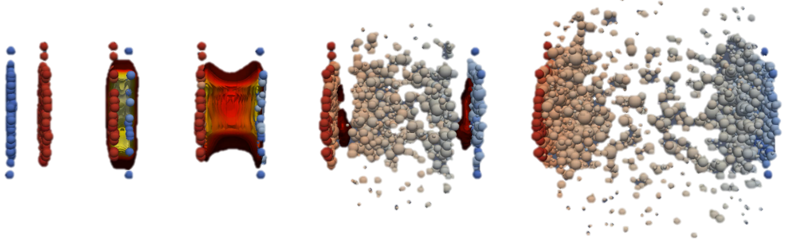
J.E.Bernhard, P.Marcy, C.Coleman-Smith, S.Huzurbazar, R.Wolpert and S.A.Bass, Phys.Rev.C 91, 054910(2015)

Overview


Model parameters
$D_s 2\pi T(\alpha, \beta, ...)$
Latin hypercube
Physics model
HQ in-medium evolution
Gussian process emulator
Non-parametric interpolation
fast surrogate
MCMC
random walk through
parameter space
Experiemental data
-LHC + RHIC: R_{AA}, v_2
Bayes' Theorem
posterior ∝ likelihood × prior
After equilibrium
Posterior distribution
probability distribution
of each parameter

Part II: HQ in heavy-ion collisions

figure credit: Hannah Petersen(Au-Au collisions)



Initial condition:
Spatial IC: T$_R$ENTo
Momentum IC: FONLL

In-medium evolution:
HQ transport: Langevin (col + rad)
Medium: hydrodynamic

Hadronization:
fragmentation + recombination

II: Initial condition

Position space: T_RENTo (A parametric IC model)

- Entropy deposition proportional to eikonal parameterization

$$\left.\frac{ds}{dy}\right|_{\tau=\tau_0} \propto \left(\frac{T_A^p + T_B^p}{2}\right)^{1/p}$$

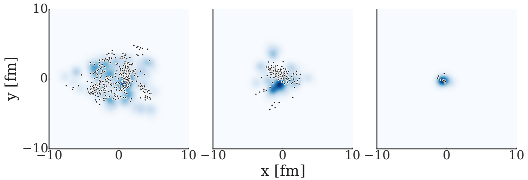
J.S.Moreland,J.Bernhard, and S.A.Bass, Phys.Rev.C 92, 011901(2015)

- $p = 0 \Rightarrow ds/dy \propto \sqrt{T_A T_B}$ (mimic the behavior of IP-Glasma)
- Heavy quark initial production probability: $\left.\frac{dN}{dy}\right|_{\tau=\tau_0} \propto T_{\mathrm{AA}}$

Momentum space: FONLL

- Parton distribution function: CTEQ6
- Nuclear shadowing effect: EPS09 NLO

M.Cacciari,S.Frixione, and P.Nason, arxiv:hep-ph/0102134



II: Calibration of the medium

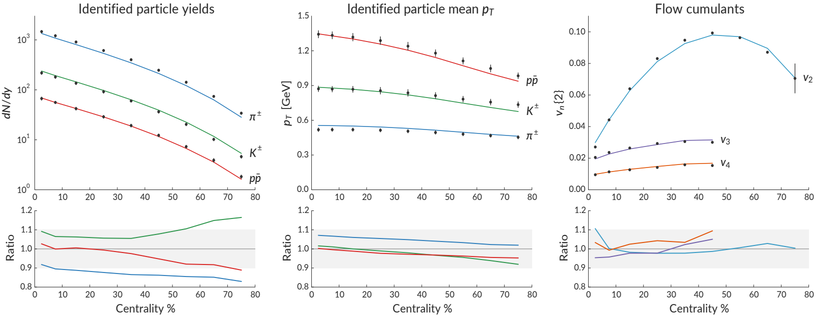
(2+1)D viscous hydro: iEbE-VISHNU

H.Song and U.W.Heinz,
Phys.Rev.C 77, 064901(2008)

- Equation of state from lattice QCD (HotQCD collaboration)
- Temperature-dependent shear + bulk vis correction

$$(\eta/s)(T) = (\eta/s)_{\mathrm{min}} + (\eta/s)_{\mathrm{slope}}(T - T_c), T_c = 154\mathrm{MeV}$$
$$(\zeta/s)(T) = (\zeta/s)_{\mathrm{norm}} \times f(T)$$

- All the initial/medium related parameters ($\mathrm{norm}, p, \eta/s$ etc.) are calibrated by Bayesian analysis with experimental data



J.Bernhard,J.S.Moreland,S.A.Bass,J.Liu, and U.Heinz
Phys.Rev.C 94, 024907(2015)

II: HQ in-medium evolution

Improved Langevin transport model

S.Cao, G.Qin, and S.A.Bass, Phys.Rev.C 92, 024907(2015)

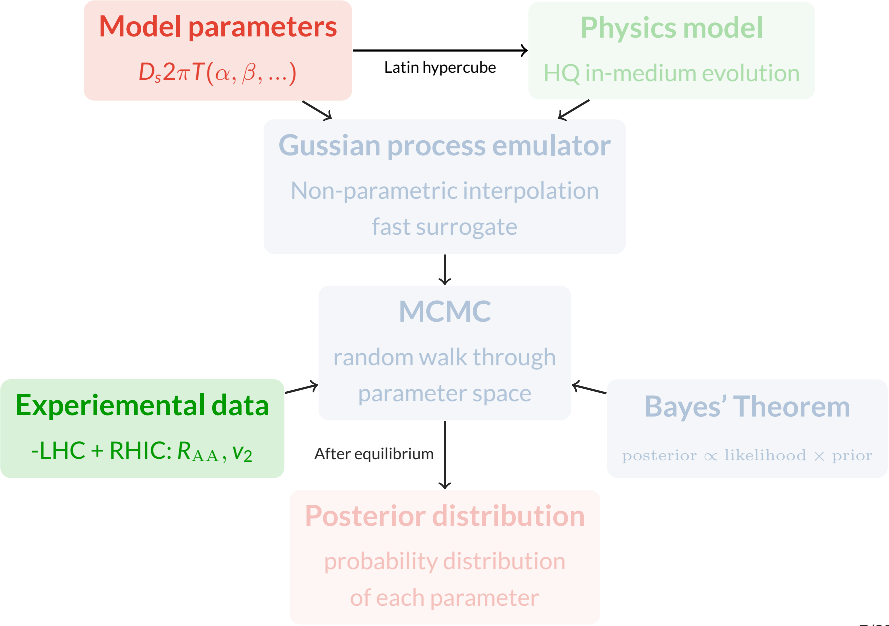
$$\frac{d\vec{p}}{dt} = -\eta_D(p)\vec{p} + \vec{\xi} + \vec{f_g} \tag{3}$$

- Drag force: $\eta_D(p) = \kappa/(2TE)$
- Thermal random force: $\langle \xi^i(t)\xi^j(t') = \kappa\delta^{ij}\delta(t-t')\rangle$
- Recoil force from gluon radiation: $\vec{f_g} = -d\vec{p_g}/dt$
- Gluon emission probability:

$$\frac{dN_g}{dx dk_\perp^2 dt} = \frac{2\alpha_s P(x)\hat{q}_g}{\pi k_\perp^4} \sin^2(\frac{t-t_i}{2\tau_f})(\frac{k_\perp^2}{k_\perp^2 + x^2M^2})^4 \tag{4}$$

- $\hat{q}_g = \hat{q}C_A/C_F = 2\kappa C_A/C_F, D_s = 2T^2/\kappa$
- **Diffusion coefficient D_s=?**

Overview


Model parameters
$D_s 2\pi T(\alpha, \beta, \ldots)$
Latin hypercube
Physics model
HQ in-medium evolution
Gussian process emulator
Non-parametric interpolation
fast surrogate
MCMC
random walk through
parameter space
Experiemental data
-LHC + RHIC: R_{AA}, v_2
Bayes' Theorem
posterior $\propto$ likelihood $\times$ prior
After equilibrium
Posterior distribution
probability distribution
of each parameter

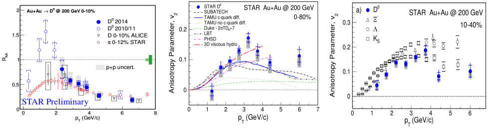
Model parameters and experimental data

HQ diffusion coefficient

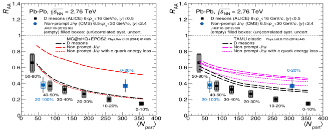
- Linear temperature dependence
  $D_s 2\pi T = A + B \cdot (T - T_c) \propto \eta / s$
- + momentum dependence: $D_s 2\pi T = \frac{A(1+B \cdot T/T_c)}{1+C \cdot \log(E)}$

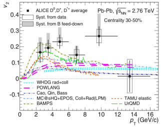
Experiemental observables: $\vec{y}_{40} = (R_{\mathrm{AA}}, v_2)$

- AuAu@ 200 GeV (STAR)
- PbPb@ 2.76 TeV (ALICE)



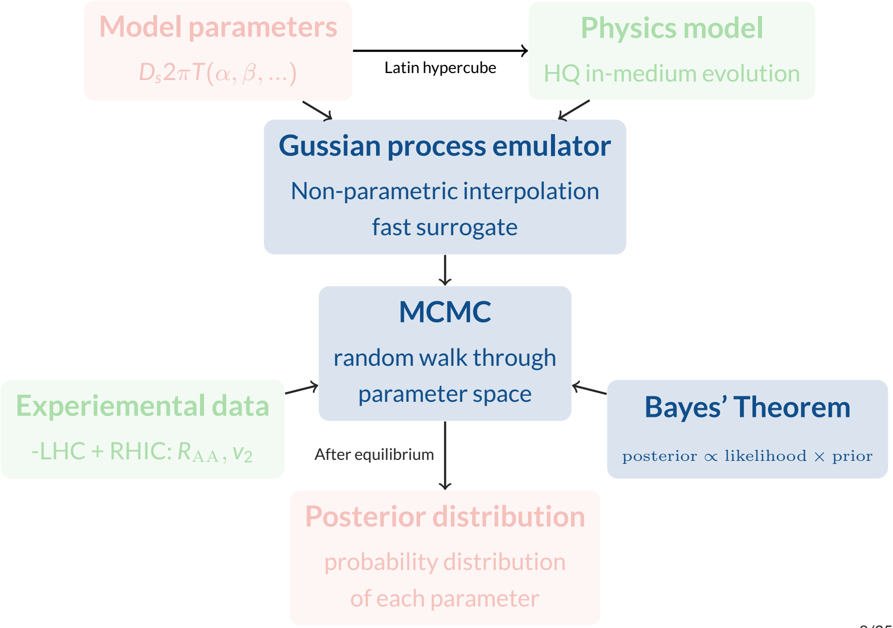
STAR collaboration: arxiv 1701.06060, 1601.00695





ALICE collaboration:
arxiv 1506.06604,
PRC **90**, 034904 (2014)

Overview



Model parameters
$D_s 2\pi T(\alpha, \beta, ...)$
Latin hypercube
Physics model
HQ in-medium evolution
Gussian process emulator
Non-parametric interpolation
fast surrogate
MCMC
random walk through
parameter space
Experiemental data
-LHC + RHIC: R_{AA}, v_2
Bayes' Theorem
posterior ∝ likelihood × prior
After equilibrium
Posterior distribution
probability distribution
of each parameter

Theory DUKE

Bayes' Theorem

$$P(\vec{x}_* | X, Y, \vec{y}_{\rm exp}) \propto P(X, Y, \vec{y}_{\rm exp} | \vec{x}_*) P(\vec{x}_*) \quad (5)$$

- $P(A, B) = P(A)P(B|A) = P(B)P(A|B), P(B|A) \propto P(A|B)P(B)$
- Posterior distribution $P(\vec{x}_* | X, Y, \vec{y}_{\rm exp})$: probability of $\vec{x}_*$ given observing $(X, Y, \vec{y}_{exp})$
- Likelihood $P(X, Y, \vec{y}_{\rm exp} | \vec{x}_*)$: probability of observing $(X, Y, \vec{y}_{\rm exp})$ given $\vec{x}_*$
  $\propto \exp[-\frac{1}{2}(\vec{y} - \vec{y}_{\rm exp})^T \Sigma^{-1} (\vec{y} - \vec{y}_{\rm exp})]$ D.Foreman-Mackey,D.Hogg,D.Lang,J.Goodman arXiv:1202.3665
  - $\vec{y}$ is the model output for input parameter $\vec{x}$
  - covariance matrix $\Sigma = {\rm diag}(\sigma^2_{\rm stat}) + {\rm diag}(\sigma^2_{\rm sys})$
- Prior distribution $P(\vec{x}_*)$: (simplest case) uniformly distributed in parameter space

III: Bayesian model calibration

However...

1. Too many CPU hours

- $O(1000)$ CPU hours for one evaluation of $\vec{y}$ at given $\vec{x}$
- $O(1000)$ nwalkers separate Metropolis-Hastings chains,
- 500 burn-in step, 1000 production step per walker

⇑

Gaussian process emulator

2. Too many observables

- $\vec{y} = R_{\mathrm{AA}}, v_2$
- Independent GP emulators for each output?
- Highly correlated

⇑

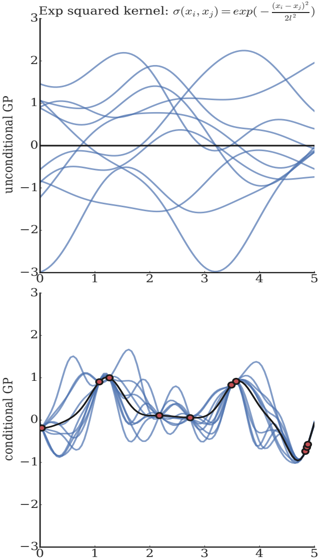
Principal component analysis

Gaussian process

- A collection of random variables, which have a joint Gaussian distribution
- Map inputs to normally-distributed outputs

$$\vec{y} \sim \mathcal{N}(\vec{\mu}, \sigma(X)) \quad (6)$$

- Only need to be specified by mean and covariance functions, for example:
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \mu(\vec{x}_1) \\ \mu(\vec{x}_2) \end{pmatrix}, \begin{pmatrix} \sigma(\vec{x}_1, \vec{x}_1), & \sigma(\vec{x}_1, \vec{x}_2) \\ \sigma(\vec{x}_2, \vec{x}_1), & \sigma(\vec{x}_2, \vec{x}_2) \end{pmatrix}\right]$$
mean $\vec{\mu} = \vec{0}$, $\sigma(\vec{x}_1, \vec{x}_2) = \delta^2 \exp\left(-\frac{(\vec{x}_1 - \vec{x}_2)^2}{2\vec{\beta}^2}\right)$
- Given$(\vec{y}, X)$GP parameters (hyper-parameters) $\delta^2, \vec{\beta}$ can be estimated

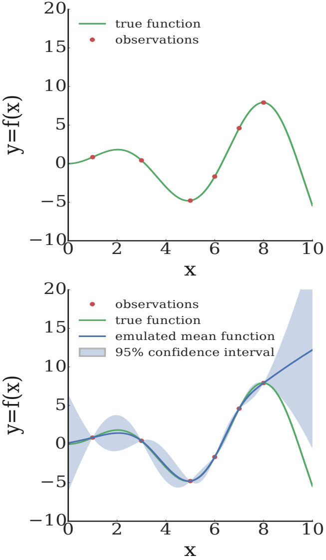


GP as emulator

- Physics model: simulator $y = f(x) + \epsilon$

$$\begin{pmatrix} x_{11} & \dots & x_{1m} \\ \dots & & \dots \\ x_{n1} & \dots & x_{nm} \end{pmatrix} \Rightarrow \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix} \qquad (7)$$

- GP emulator: given a dataset $(X, \vec{y})$, approximation of the simulator + probabilistic prediction
- This work: covariance(include a noise term)
  $\sigma(\vec{x}, \vec{x}') = \sigma_{\rm GP}^2 \exp\left[-\sum_{k=1}^{m} \frac{(x_k - x_k')^2}{2l_k^2}\right] + \sigma_n^2 \delta_{\vec{x},\vec{x}'}$
- Maximize the evidence $\log P(y_* | X, Y, \vec{x}_*) = -\frac{1}{2} Y^T \Sigma^{-1}(X, \vec{x}_*) Y - \frac{1}{2} \log|\Sigma(X, \vec{x}_*)| - N/2 \log(2\pi)$



III: Principal component analysis

Multiple observables, correlated?

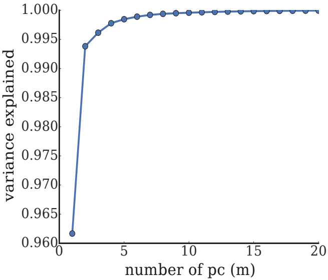
$$\begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \cdots & & \cdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} \Rightarrow \begin{pmatrix} y_{11} & \cdots & y_{1k} \\ \cdots & & \cdots \\ y_{n1} & \cdots & y_{nk} \end{pmatrix} \quad (8)$$

- Decompose into orthogonal linear principal components
- Singular value decomposition: n sets of k-dimension observables $\Rightarrow$ n sets of l-dimensional PCs Z

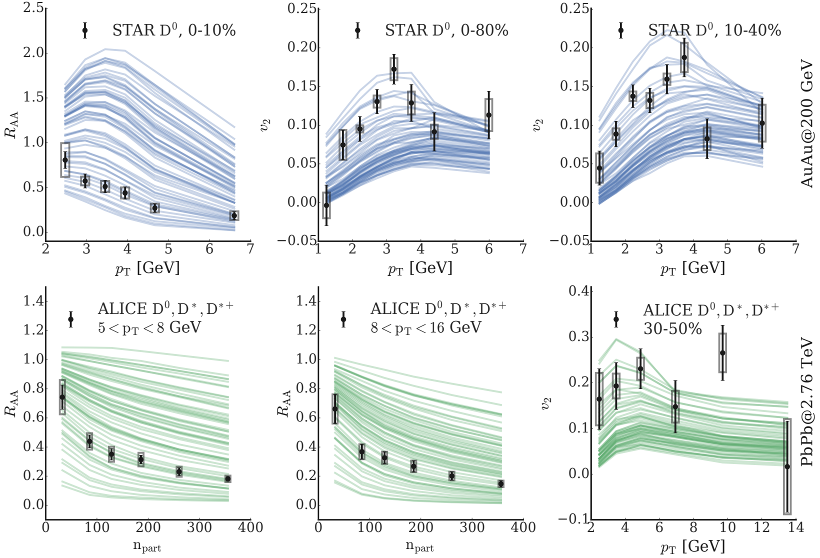
$$Y_{kn} = U_{kl} S_{ll} V^T_{ln} \quad (9)$$

$$Z = \sqrt{n} Y V \quad (10)$$

- Eigenvalue λ_i of Y represents the variance explained by PC $\vec{z}_i$



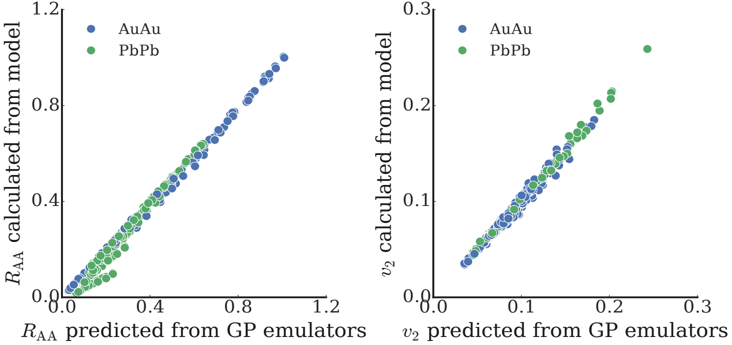
III: Prior (training data)



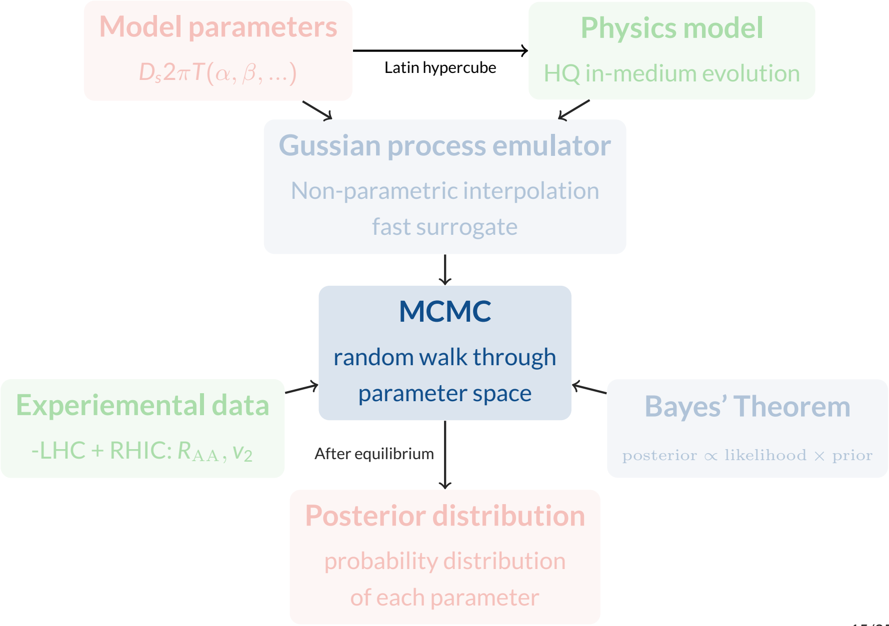
60 sets of inputs ($X = (\vec{x}_1, \ldots, \vec{x}_{60})$) $\Rightarrow$ 60 sets of outputs ($Y = (\vec{y}_1, \ldots, \vec{y}_{60})$)

III: Emulator validation

Another 10 sets of validation inputs $\vec{x}$
Compare between physics model calculation $\vec{y}$ and GP emulator predicted $\vec{y}_{\mathrm{pred}}$



Overview


Model parameters
$D_s 2\pi T(\alpha, \beta, ...)$
Latin hypercube
Physics model
HQ in-medium evolution
Gussian process emulator
Non-parametric interpolation
fast surrogate
MCMC
random walk through
parameter space
Experiemental data
-LHC + RHIC: R_{AA}, v_2
Bayes' Theorem
posterior $\propto$ likelihood $\times$ prior
After equilibrium
Posterior distribution
probability distribution
of each parameter

http://dan.iel.fm/emcee

Affine invariance MCMC ensemble sampler: emcee

- random walk through parameter space, weighted by likelihood
- equilibrium $\Rightarrow$ posterior distribution
- An ensemble $\vec{X}$ (consists of many independent nwalkers)
- Each walker takes random walk by Metropolis-Hasting algorithm, where each step is accepted or rejected based on the likelihood $P(X, Y, \vec{y}_{\mathrm{exp}}|\vec{x})$
- **Acceptance rate:** α
  - $\alpha = 1$ purely random walk
  - $\alpha = 0$ walker stuck
  - Normally optimal proposal has $\alpha = 44\%$ for 1d, 25% larger than 4d
- This work: $\sim 30 - 40\%$

Overview

Model parameters

$D_s 2\pi T(\alpha, \beta, \ldots)$

Latin hypercube

Physics model

HQ in-medium evolution

Gussian process emulator

Non-parametric interpolation

fast surrogate

MCMC

random walk through

parameter space

Experiemental data

-LHC + RHIC: R_{AA}, v_2

Bayes' Theorem

posterior $\propto$ likelihood $\times$ prior

After equilibrium

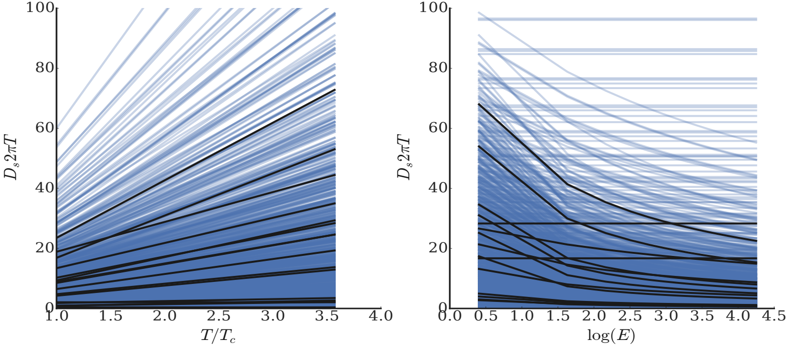
Posterior distribution

probability distribution

of each parameter

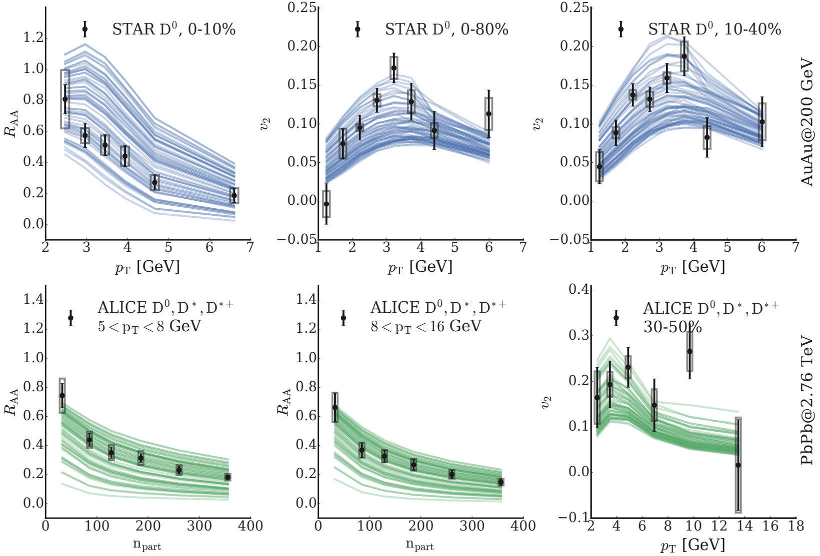
Parametrization

- $D_s 2\pi T = A + B \cdot (T - T_c)$
- $D_s 2\pi T = \frac{A(1+B \cdot T/T_c)}{1+C \cdot \log(E)}$



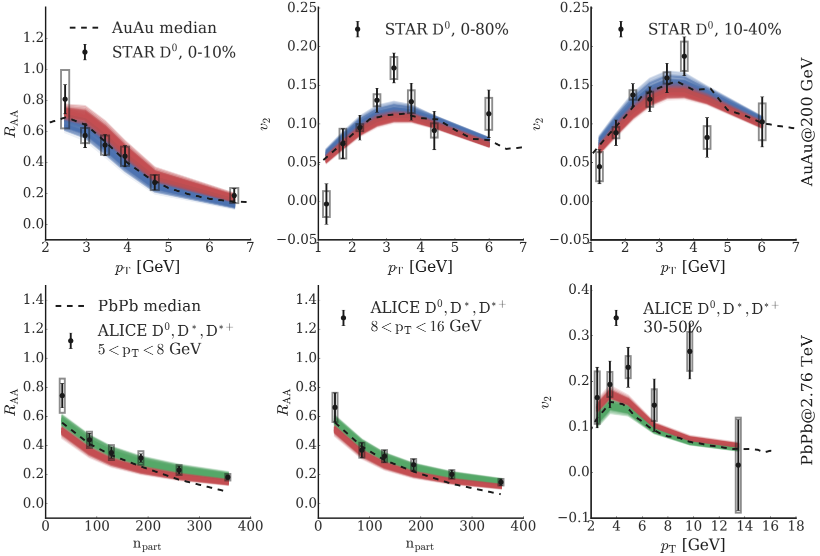
IV: $D_s 2\pi T = A + B \cdot (T - T_c)$

Prior: (60 sets of training data)

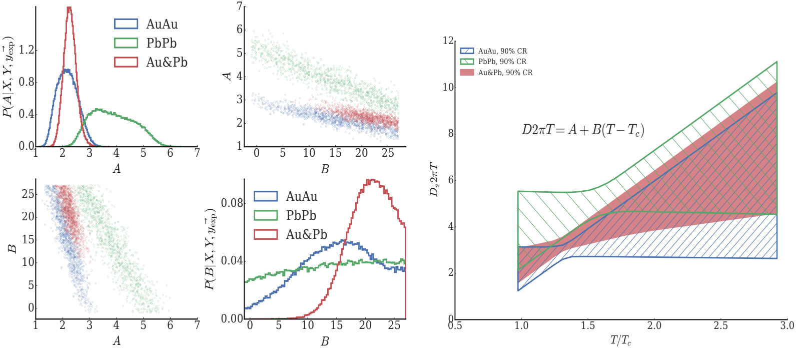


IV: $D_s 2\pi T = A + B \cdot (T - T_c)$

Posterior results: (200 sets of random posterior outputs)

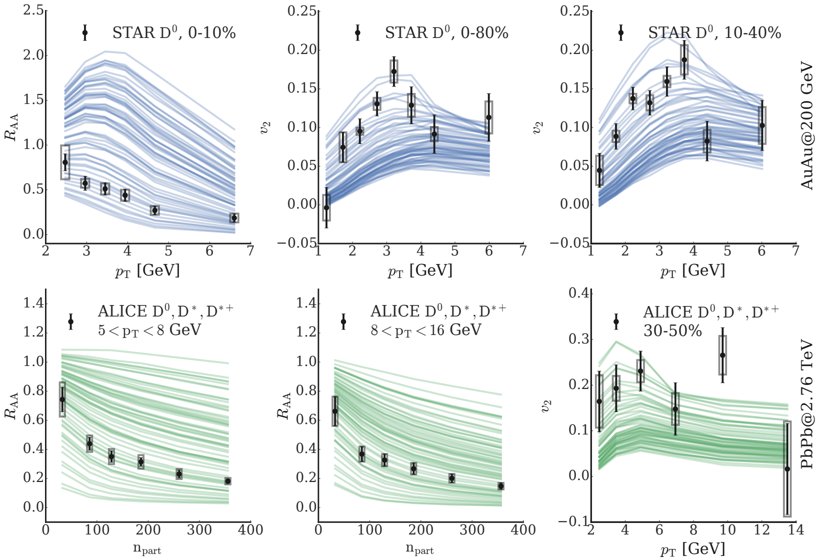


Posterior distributions of (A', B')

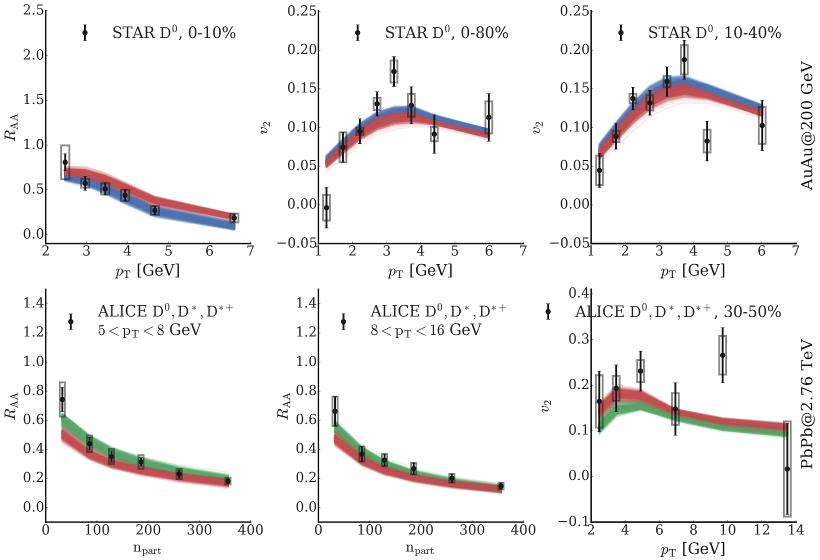


IV: $D_s 2\pi T = \frac{A(1+B\cdot T/T_c)}{1+C\cdot\log(E)}$

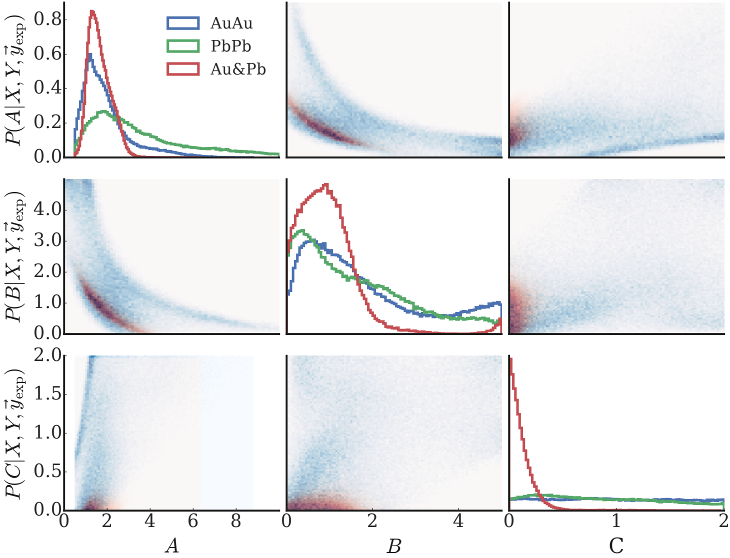
Prior: (60 sets of training data)



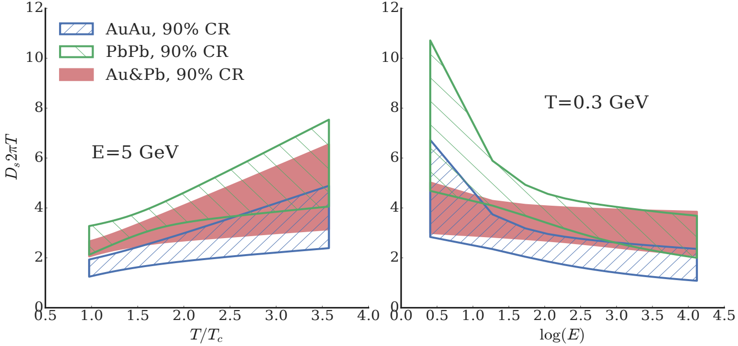
Posterior results: (200 sets of random posterior outputs)



Posterior probability distribution



Posterior probability distribution



Summary

- Bayesian analysis provides a rigorous method to estimate the optimal parameters (with uncertainties) to simultaneously describing the experimental data
- Heavy quark diffusion coefficients can be extracted from the analysis and be constrained with some precision, ($D_s 2\pi T \sim 2 - 6$ near T_c)
- A higher precision of experimental data?
- A more sophisticated momentum dependent inspired by pQCD calculation $D_s = D_{\mathrm{spQCD}} \cdot K(p)$; test on different transport models