Parameter extractions for RHIC BES using Bayesian statistics

Jussi Auvinen (Duke U.)

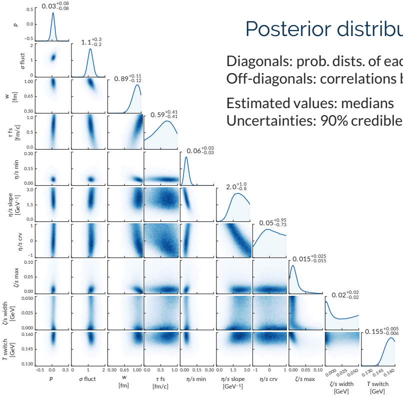
in collaboration with Iu. Karpenko, J. Bernhard and S. A. Bass

CPOD 2017

August 8, 2017

Duke UNIVERSITY

Success of Bayesian approach

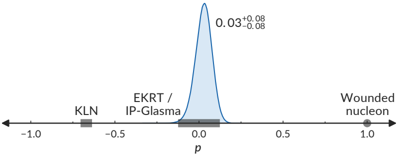


Posterior distribution

Diagonals: prob. dists. of each param.
Off-diagonals: correlations b/w pairs

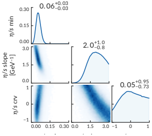
Estimated values: medians
Uncertainties: 90% credible intervals

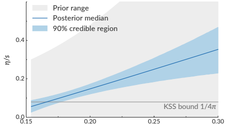
Initial state



Shear viscosity

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

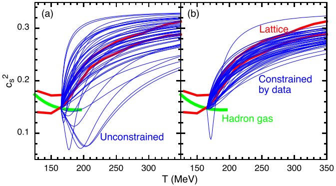


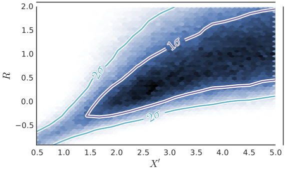


Jonah Bernhard, Quark Matter 2017

Success of Bayesian approach

Equation of state

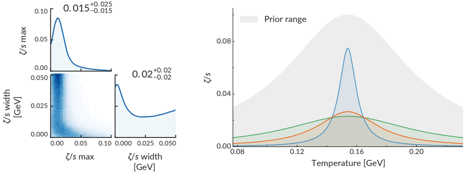




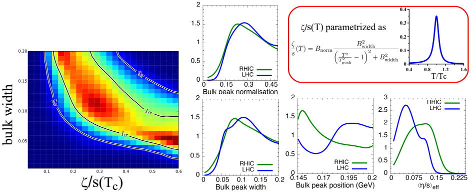
Pratt et al., PRL114, 202301 (2015)

Bulk viscosity

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



Jonah Bernhard, QM2017



Gabriel Denicol, QM2017

RHIC beam energy scan

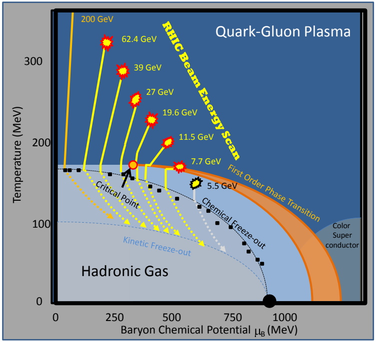
Investigate $\sqrt{s_{NN}}$-dependence of η/s using statistical analysis

- Possible indication of μ_B-dependence in the transport coefficient
- Statistical analysis ⇒ Get both best-fit values and their uncertainties

- Phase A: Find best-fitting model parameters for multiple collision energies independently
- If energy dependence observed in the best-fit parameters

⇓

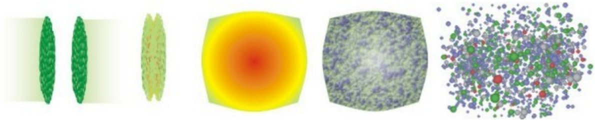
Phase B: parametrize the dependence and find the best fit over all collision energies simultaneously



Picture: G. Odyniec, Acta Phys. Polon. B 43, 627 (2012).

Transport + hydrodynamics hybrid model

Karpenko, Huovinen, Petersen, Bleicher, Phys.Rev.C, 91, 064901 (2015)



- Initial state described by UrQMD[1] hadron transport
- Start the hydrodynamical evolution at time τ_0 when the two nuclei have passed through each other
- Convert energy, momentum and baryon number of each particle into 3D Gaussian distributions with width parameters W_{trans}, W_{long}
- (3+1)D viscous hydrodynamics with constant ratio of shear viscosity over entropy density η/s (bulk viscosity ignored)
- Transition from hydro back to UrQMD ("particlization") when energy density ϵ is smaller than the switching condition ϵ_{SW}

"Cornelius" hypersurface finder, P. Huovinen and H. Petersen, EPJ A48 171 (2012)

[1] S. A. Bass *et al.*, Prog. Part. Nucl. Phys. 41, 255 (1998), M. Bleicher *et al.*, J. Phys. G 25, 1859 (1999).

Bayesian analysis

Model parameters (input): $\vec{x} = (x_1, ..., x_n)$

$(\tau_0,\ W_{\text{trans}},\ W_{\text{long}},\ \eta/s,\ \epsilon_{SW})$

$\Downarrow$

Model output $\vec{y} = (y_1, ..., y_m) \Leftrightarrow$ Experimental values $\vec{y}^{\,\text{exp}}$

$(N_{\text{ch}},\ \langle p_T \rangle,\ v_2, ...)$

Bayes' theorem:

Posterior probability $\propto$ Likelihood $\cdot$ Prior knowledge

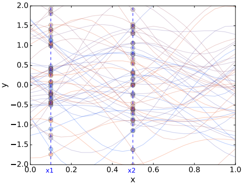
- Prior knowledge: Range of input parameter values to investigate
- Likelihood: $\mathcal{L}(\vec{x}) \propto \exp\left(-\frac{1}{2}(\vec{y}(\vec{x}) - \vec{y}^{\,\text{exp}})\Sigma^{-1}(\vec{y}(\vec{x}) - \vec{y}^{\,\text{exp}})^T\right)$, where Σ is the covariance matrix

Use Markov chain Monte Carlo to sample the posterior probability

- Random walk in input parameter space, constrained by prior
- Use Gaussian process (GP) emulator to estimate model output for likelihood computations

Gaussian process

Stochastic process: A parameterized collection of random variables $\{y_t\}_{t \in T}$ (T possibly infinite). E.g. random walk over time.

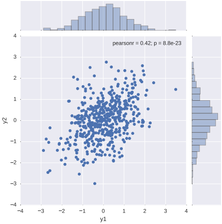


Gaussian process: A stochastic process, in which every finite set $Y = \{y_i\}_{i=1}^{N}$ is a multivariate Gaussian random variable $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = \mu(X) = \{\mu(x_1), ..., \mu(x_N)\}$$

is the mean and

$$\boldsymbol{\Sigma} = \sigma(X, X) = \begin{pmatrix} \sigma(\vec{x}_1, \vec{x}_1) & \cdots & \sigma(\vec{x}_1, \vec{x}_N) \\ \vdots & \ddots & \vdots \\ \sigma(\vec{x}_N, \vec{x}_1) & \cdots & \sigma(\vec{x}_N, \vec{x}_N) \end{pmatrix}$$



is the covariance matrix with covariance function $\sigma(\vec{x}, \vec{x}')$.

Gaussian process

Choice: Squared-exponential covariance function with a noise term

$$\sigma(\vec{x},\vec{x}') = \theta_0 \exp\left(-\sum_{i=1}^{n} \frac{(x_i - x_i')^2}{2\theta_i^2}\right) + \theta_{\text{noise}}\delta_{\vec{x}\vec{x}'}$$

The *hyperparameters* $\vec{\theta} = (\theta_0, \theta_1, ...\theta_n, \theta_{\text{noise}})$ are not known a priori and must be estimated from a set of model outputs ("training points")

$\Rightarrow$ emulator training: Maximise the marginal likelihood (aka “evidence”)

$$\log P(Y|X,\vec{\theta}) = \underbrace{-\frac{1}{2}Y^T\mathbf{\Sigma}^{-1}(X,\vec{\theta})Y}_{\text{model data fit}} \underbrace{-\frac{1}{2}\log|\mathbf{\Sigma}(X,\vec{\theta})|}_{\text{complexity penalty}} \underbrace{-\frac{N}{2}\log(2\pi)}_{\text{normalization}}$$

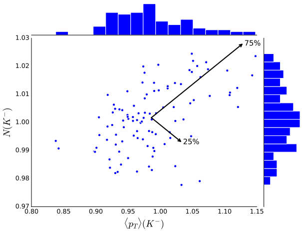
Principal component analysis

m observables $\Rightarrow$ m Gaussian processes needed for model emulation

However, m can be up to $\mathcal{O}(100)$ at top RHIC energies and at the LHC!
Number of emulators can be reduced with principal component analysis

First principal component represents the direction of largest variance in output space, second PC the direction of second largest variance, etc.

- Fraction of variance explained by principal component p_q: $\mathrm{Var}(p_q) = \frac{\lambda_q}{\sum_{i=1}^{m} \lambda_i}$
- Select the number of principal components which together explain desired fraction of total variance; often only a few PCs are needed to explain 99% of the variance



Likelihood function

The likelihood function used in MCMC:

$$\frac{1}{|2\pi(\Sigma_{\rm exp}+\Sigma_{GP})|}\exp\left(-\frac{1}{2}(\vec{z}^{*}_{GP}-\vec{z}_{\rm exp})(\Sigma_{\rm exp}+\Sigma_{GP})^{-1}(\vec{z}^{*}_{GP}-\vec{z}_{\rm exp})^{T}\right)$$

- $\vec{z}^{*}_{GP}$ is the emulator prediction at the input parameter point $\vec{x}^{*}$
- $\vec{z}_{\rm exp}$ is the experimental data transformed to principal component space
- Σ_{GP} is the predictive variance (emulator uncertainty)
- $\Sigma_{\rm exp}$ is the experimental error squared ($\sigma^2_{\rm stat}+\sigma^2_{\rm sys}$), transformed to PC space

Analysis procedure

Produce training data

- run simulations with $\mathcal{O}(100)$ different parameter combinations

Preparation for principal component analysis:

1. Scale with experimental values $\Rightarrow$ Unitless quantities
2. Check that the values are roughly normally distributed; apply a transformation if necessary
3. Center the data; apply possible data weighting

Determine required number of Gaussian processes with PCA

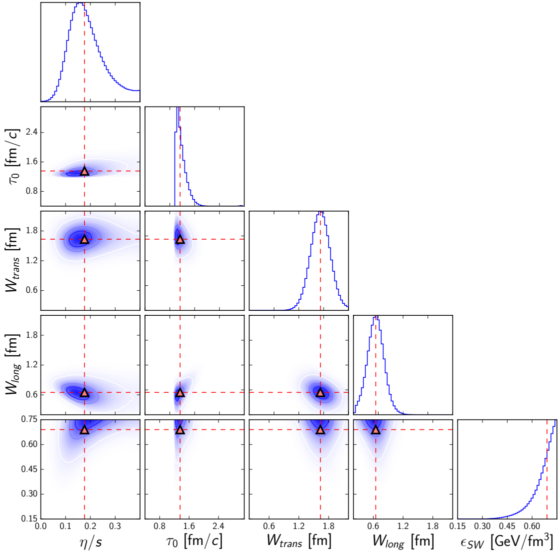
$\Downarrow$

Condition the emulators on training data

$\Downarrow$

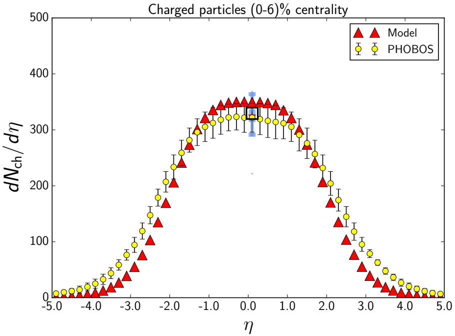
Calibrate on experimental data by running MCMC

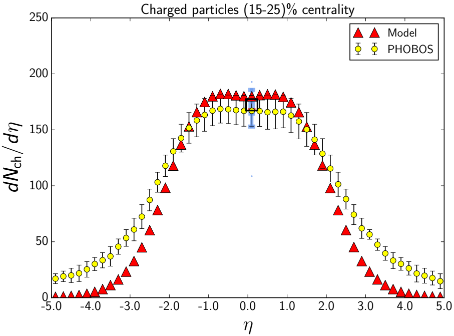
Posterior probabilities at $\sqrt{s_{NN}} = 19.6$ GeV



Observables at $\sqrt{s_{NN}} = 19.6$ GeV

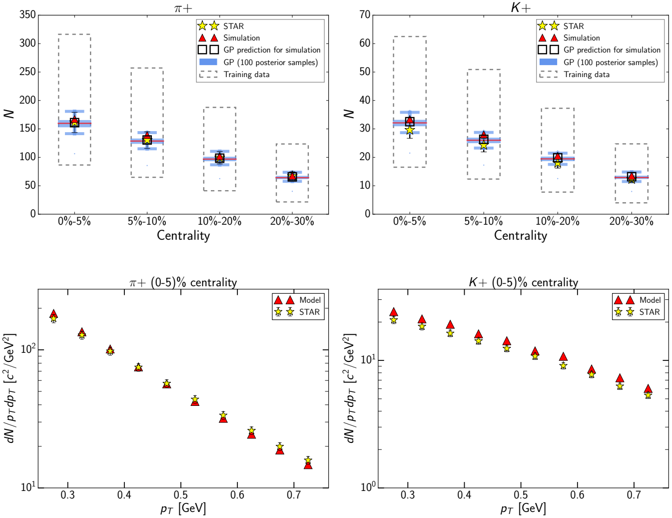
Charged particle pseudorapidity distribution vs. centrality





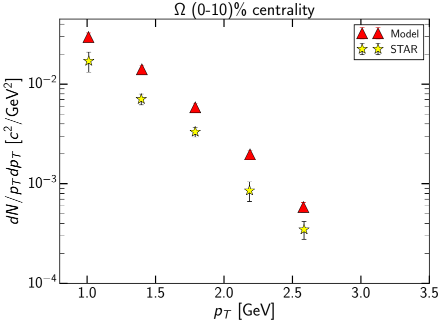
Observables at $\sqrt{s_{NN}} = 19.6$ GeV

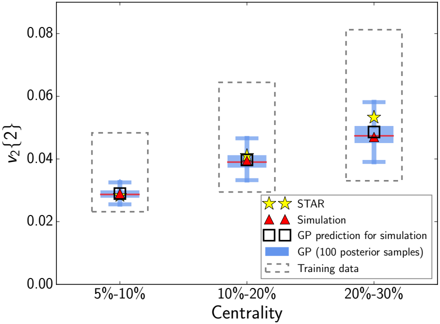
Pion and kaon multiplicities and transverse momentum spectra



Observables at $\sqrt{s_{NN}} = 19.6$ GeV

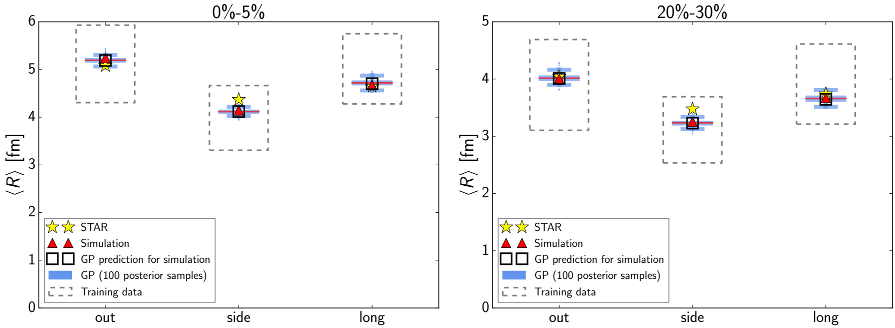
Ω p_T spectrum and elliptic flow $v_2\{2\}$



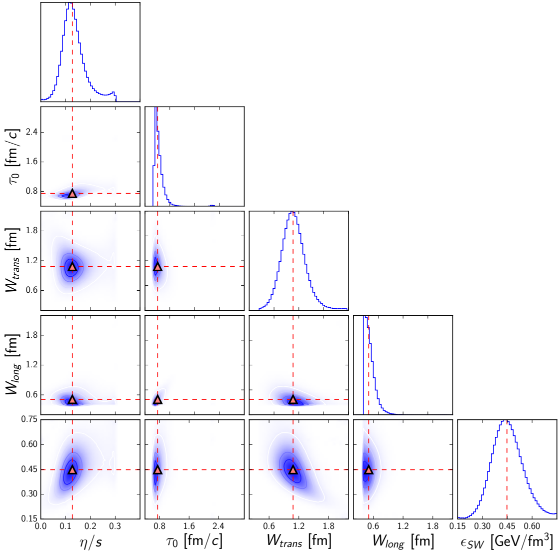


Observables at $\sqrt{s_{NN}} = 19.6$ GeV

HBT radii from two-pion interferometry (averaged over 4 m_T bins)

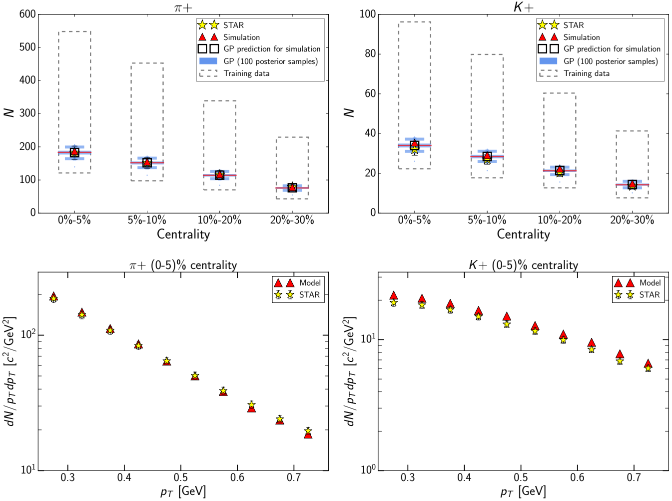


Posterior probabilities at $\sqrt{s_{NN}} = 39$ GeV



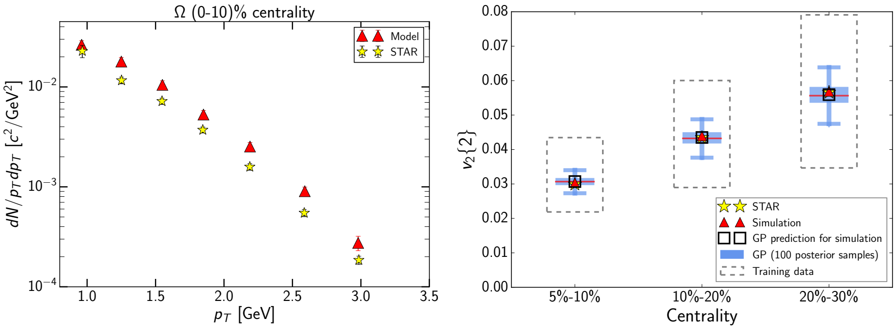
Observables at $\sqrt{s_{NN}} = 39$ GeV

Pion and kaon yields



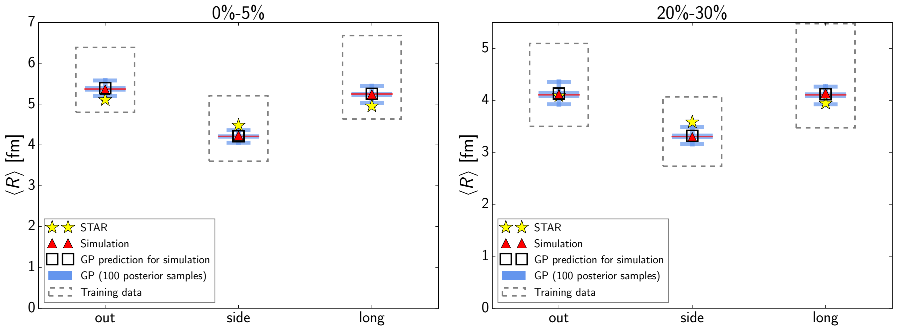
Observables at $\sqrt{s_{NN}} = 39$ GeV

Ω p_T spectrum and elliptic flow $v_2\{2\}$

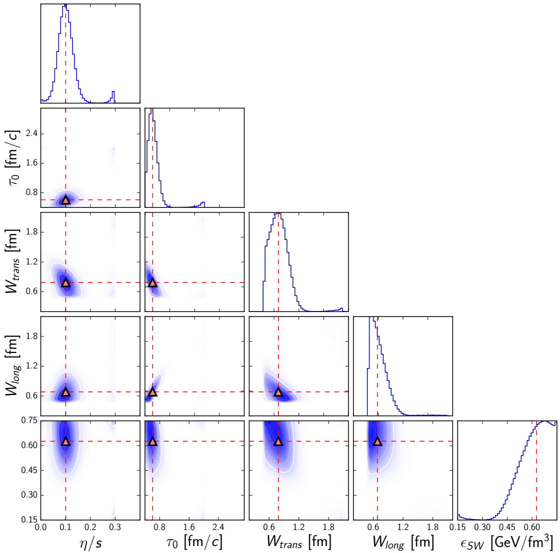


Observables at $\sqrt{s_{NN}} = 39$ GeV

HBT radii (averaged over 4 m_T bins)

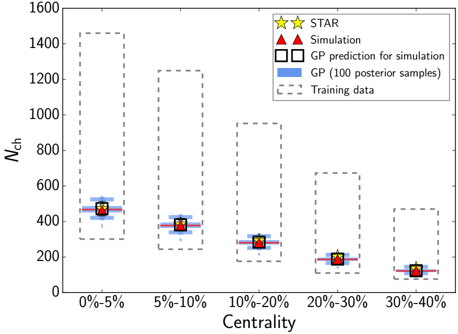


Posterior probabilities at $\sqrt{s_{NN}} = 62.4$ GeV



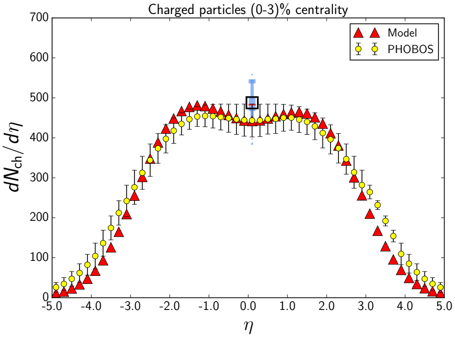
Observables at $\sqrt{s_{NN}} = 62.4$ GeV

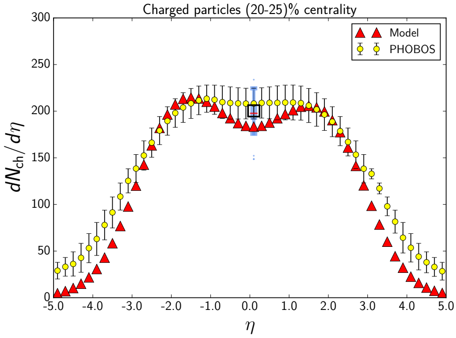
Charged particle yield vs. centrality



Observables at $\sqrt{s_{NN}} = 62.4$ GeV

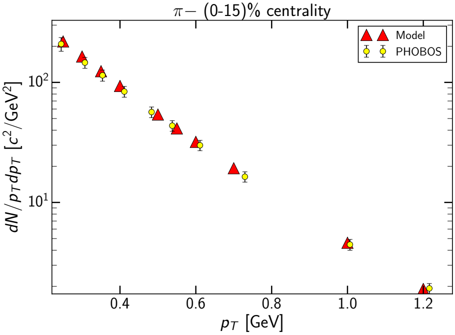
Pseudorapidity distribution vs. centrality

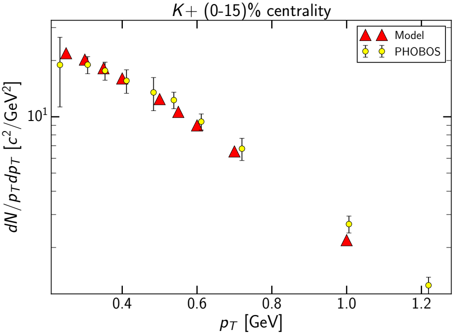




Observables at $\sqrt{s_{NN}} = 62.4$ GeV

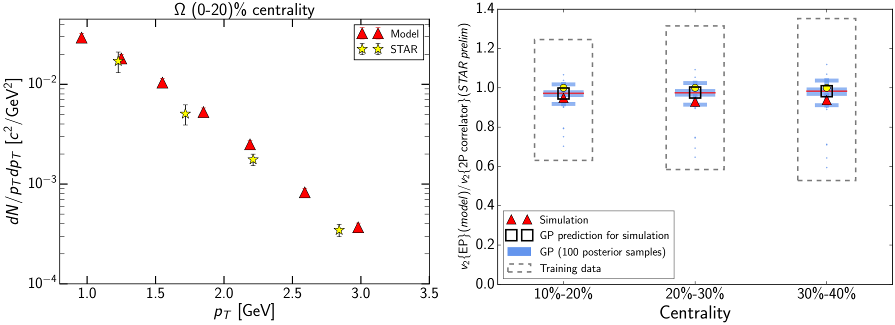
Pion and kaon p_T spectra





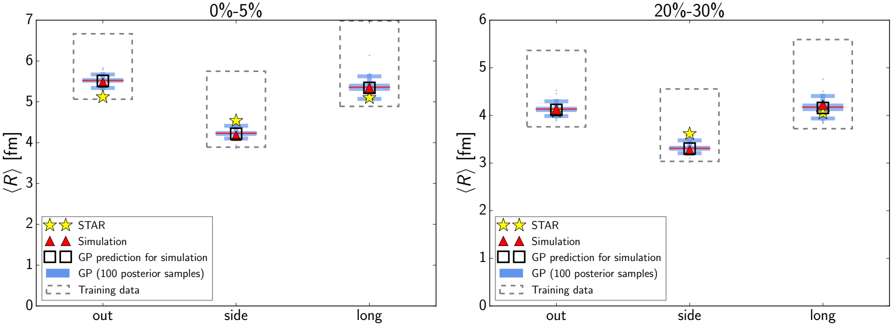
Observables at $\sqrt{s_{NN}} = 62.4$ GeV

Ω p_T spectrum and elliptic flow $v_2\{\mathrm{EP}\}$

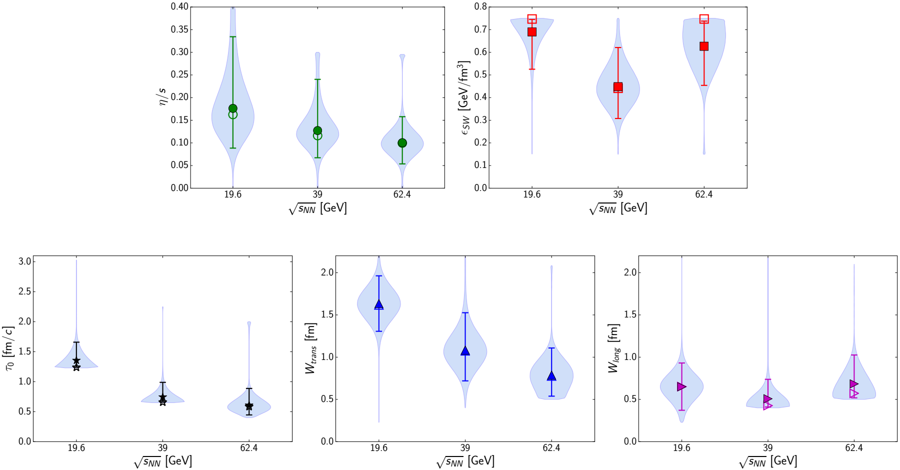


Observables at $\sqrt{s_{NN}} = 62.4$ GeV

HBT radii (averaged over 4 m_T bins)



Parameter dependence on collision energy



Filled symbols: Median values Open symbols: Peak values Error bars: 90% confidence range

Summary

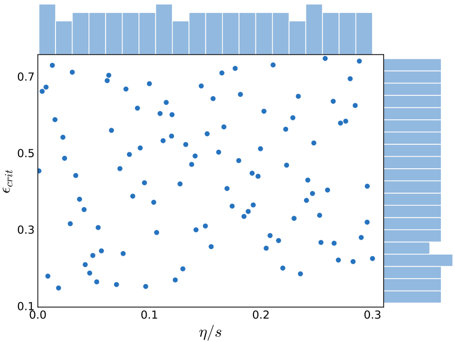
- Bayesian analysis provides a rigorous method for simultaneous estimation of both the best-fit values and the associated uncertainties for the parameters of heavy ion collision models
- Pion p_T spectra and charged particle elliptic flow v_2 well reproduced using median values from the posterior probability distributions
- Charged particle pseudorapidity distribution and R_{side} have systematic deviations from measured values; excess of strange particles at lower beam energies
- Transport-to-hydro switching time τ_0 and longitudinal smoothing W_{long} well constrained by data, while transverse smoothing W_{trans}, hydro-to-transport switching density ϵ_{SW} and shear viscosity η/s have larger uncertainties
- While the most probable value of η/s is higher at lower energies, a constant value of $0.10 - 0.15$ cannot be currently excluded

Extra slides

Investigated parameter ranges

Sample points evenly over whole parameter space using Latin hypercube method

- Shear viscosity over entropy density η/s: 0.001 - 0.4
- Transport-to-hydro transition time τ_0: 0.4 - 3.1 fm
- Transverse Gaussian smearing of particles W_{trans}: 0.2 - 2.2 fm
- Longitudinal Gaussian smearing of particles W_{long}: 0.2 - 2.2 fm
- Hydro-to-transport transition energy density ϵ_{SW}: 0.15 - 0.75 GeV/fm^3



Gaussian process prediction

To predict the model output y_0 in an arbitrary point $\vec{x_0}$, we write the joint distribution

$$\begin{pmatrix} y_0 \\ Y \end{pmatrix} = \mathcal{N}\left(\begin{pmatrix} \mu(\vec{x_0}) \\ \mu(X) \end{pmatrix}, \begin{pmatrix} \Sigma_{0,0} & \Sigma_{0,X} \\ \Sigma_{X,0} & \Sigma_{X,X} \end{pmatrix} \right)$$

and calculate the conditional predictive mean

$$\bar{\mu}(\vec{x_0}) = \mu(\vec{x_0}) + \Sigma_{0,X}\Sigma_{X,X}^{-1}(Y - \mu(X)).$$

Typically we define the mean function $\mu(\vec{x}) \equiv 0$ and the prediction is simply

$$\bar{\mu}(\vec{x_0}) = \Sigma_{0,X}\Sigma_{X,X}^{-1}Y$$

with the associated predictive (co)variance

$$\bar{\Sigma} = \Sigma_{0,0} - \Sigma_{0,X}\Sigma_{X,X}^{-1}\Sigma_{X,0}.$$