Bayesian analysis of heavy ion collisions with the heavy ion computational framework Trajectum

Govert Nijs [1,2] Wilke van der Schee,[3] Umut Gürsoy,[2] and Raimond Snellings[4,5]

[1]*Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2]*Institute for Theoretical Physics and Center for Extreme Matter and Emergent Phenomena, Utrecht University,*
3584 CC Utrecht, The Netherlands
[3]*Theoretical Physics Department, CERN, CH-1211 Genève 23, Switzerland*
[4]*Nikhef, 1098 XG Amsterdam, The Netherlands*
[5]*Institute for Gravitational and Subatomic Physics (GRASP), Utrecht University, 3584 CC Utrecht, The Netherlands*

We introduce a model for heavy ion collisions named TRAJECTUM, which includes an expanded initial stage with a variable free streaming velocity v_{fs} and a hydrodynamic stage with three varying second order transport coefficients. We describe how to obtain a Gaussian emulator for this 20-parameter model and show results for key observables. This emulator can be used to obtain Bayesian posterior estimates on the parameters, which we test by an elaborate closure test as well as a convergence study. Lastly, we employ the optimal values of the parameters found in Nijs *et al.* [G. Nijs, W. van der Schee, U. Gürsoy, and R. Snellings, Phys. Rev. Lett. **126**, 202301 (2021)] to perform a detailed comparison to experimental data from PbPb and *p*Pb collisions. This includes both observables that have been used to obtain these values as well as wider transverse momentum ranges and new observables such as correlations of event-plane angles.

I. INTRODUCTION

The collisions of heavy ions at the Relativistic Heavy Ion Collider (RHIC) at Brookhaven and the Large Hadron Collider (LHC) at CERN have led to an accepted picture of a short prehydrodynamic phase followed by a relativistic fluid composed of quark-gluon plasma (QGP) with remarkably small shear viscosity [1].

The most convincing evidence for an almost perfect fluid comes from the anisotropies of the particle spectra at low momentum in combination with elaborate hydrodynamical modeling that can explain these nontrivial correlations [2]. This started with ideal hydrodynamics (effectively having shear viscosity $\eta = 0$) [3], after which simulations including viscosity pointed to a surprisingly small value of the specific viscosity of $\eta/s \approx 0.08$, with s the entropy density [4]. This value was particularly exciting, since results from string theory using a holographic dual black hole indicate that quantum matter that is infinitely strongly interacting has a universal specific viscosity equaling $\eta/s = 1/4\pi \approx 0.08$ [5].

This hydrodynamic modeling relied on two pillars. First, it was assumed early on that the exploding debris left after the collision interacts so strongly that starting at a time as early as $1\,\mathrm{fm}/c$ (as in [4]) a fluid would be formed that can be described by viscous hydrodynamics. This assumption is nontrivial since the full stress-energy tensor has many more degrees of freedom (nine for conformal matter) as compared to hydrodynamics (the energy density plus fluid velocity gives four degrees of freedom). This assumption of fast "hydrodynamization" has been substantiated by studies in holography which showed that far-from-equilibrium matter can always be described by hydrodynamics within a time of the inverse temperature $1/T$ [6–8]. Similarly fast times have been found utilizing kinetic theory at a realistic value of the coupling constant [9,10].

Second, the hydrodynamic simulations need initial profiles for the energy density and fluid velocity. A Monte Carlo Glauber model [11] can provide a realistic estimate of the distribution of nucleons inside a heavy ion, but *a priori* it is not clear how this translates into an energy density for the plasma. Two naive pictures are to add up all left- and right-moving colliding pairs (binary collisions), as would be appropriate for a completely transparent collision where all interactions are independent (this is indeed applicable for the production of highly energetic quarks or gluons). For the low energy dynamics the collisions are not independent and it is more realistic to add up colliding nucleons, without doubly counting them if they collide several times (wounded nucleon approximation). In Sec. II A we review the more elaborate phenomenological TRENTO model [12], but here we note that none of these descriptions are based on microscopic insights. Exceptions to this are the color glass condensate (CGC) initial model [13], Eskola-Kajantie-Ruuskanen-Tuominen (EKRT) model [14], or models from holography [15].

Especially the uncertainty in the initial prehydrodynamic stage has led to considerable uncertainty for the hydrodynamic transport coefficients. One example is the difference in

the estimate for $\eta/s \approx 0.08$ of [4] (using binary collisions) or $\eta/s \approx 0.20$ when using CGC initial conditions [16]. Furthermore, in these references the bulk viscosity was assumed to vanish and the comparison with experiment focused on the anisotropic flow coefficients. When also demanding a simultaneous fit of the mean transverse momentum of all identified particle species it turns out that [16] needed a sizable bulk viscosity, which in turn meant that the shear viscosity was refitted to a value of $\eta/s \approx 0.095$ [17].

All of this is to stress the significant uncertainties in the precision analysis of the QGP and furthermore the importance of having a model that describes all experimental results simultaneously, preferably among different colliding systems as well as across vastly different colliding energy scales. This can be done by a sufficiently flexible model setup, where all free parameters are obtained from a global analysis, comparing with as much experimental data as possible. References [18–22] pioneered such studies using large computing resources combined with emulation and Markov chain tools inspired from similar analyses in cosmology. This culminated in the most precise temperature dependent estimate for the QGP shear viscosity to date [23].

In the current paper we will build on this expertise, which was generously made available on GitHub together with an excellent documentation [24,25]. First, we enlarge the scope of the prehydrodynamic phase by allowing a free streaming parameter that can interpolate between noninteracting free streaming with the speed of light to no streaming at all. On the one hand this is motivated by possible interactions during this phase, which can slow down the free streaming. On the other hand, by enlarging the scope of the model we can gauge the sensitivity of (physical) QGP parameters such as the viscosities to the specific parametrization used for the initial condition. Second, we enlarge the set of parameters with three second order hydrodynamic transport coefficients. This is again motivated to study the robustness of previous results on the specifics of the hydrodynamic code, but we also hope that after the relatively successful determination of η/s constraints can be obtained on second order transport coefficients. Third, an important addition is the addition of more experimental observables, such as the transverse momentum differential observables for pions, kaons, and protons, both for the yield as well as the anisotropic flow coefficients.

Of special recent interest is the debated question whether a QGP can also form in (high multiplicity) proton-ion (pA) collisions [26]. It is for this reason that we include pPb collisions in our analysis (also done in [27]). First, by obtaining results both with and without pPb collisions it is possible to at least give a partial answer if hydrodynamic modeling can describe low momentum experimental data in pPb. Second, if this is indeed possible it is the expectation that these collisions can in principle be more sensitive to several observables influenced by large gradients, which would average out for the larger and longer lived PbPb system.

In Sec. II we will describe this enlarged model which includes ten parameters for the initial stage, six first order and three second order transport coefficients, and one final parameter at which temperature to particlize our hydrodynamic fluid into hadrons for a subsequent hadronic cascade code. In
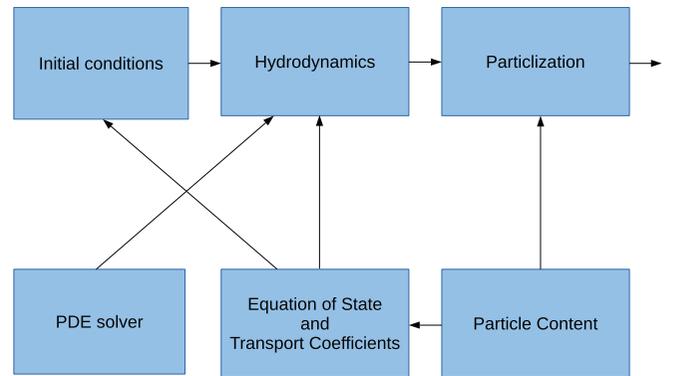
FIG. 1. Components implemented in TRAJECTUM. Also shown are the dependencies between the various components, where the information flow is indicated by the arrows. Note the (perhaps nonobvious) arrow between the equation of state/transport coefficients and the initial conditions. This allows the initial conditions to access information which is for example needed for Gubser flow.

Sec. III we describe its results, including a detailed study of how several observables depend on any of our twenty parameters. We furthermore present several computational aspects, including elaborate closure tests.

Importantly, global analyses have never been performed with transverse momentum dependent spectra or (identified) transverse momentum dependent anisotropic flow coefficients. In our companion paper [28] we attempt to fill this lacuna by comparing to such experimental data, both for PbPb and pPb collisions. A detailed analysis of the resulting QGP properties is presented in Sec. IV.

II. THE TRAJECTUM FRAMEWORK

The results in this paper have been generated using the new heavy ion code TRAJECTUM. TRAJECTUM is named after the old Roman name for the city of Utrecht, where the code was developed. The code is written in C++ and incorporates the computation of the initial conditions, prehydrodynamic phase, hydrodynamic phase, and the particlization inside a single executable. Furthermore, for each of these components a base class specifies the interface with which the component communicates with the other components. In this way, it becomes possible to have several versions for each component, which the user can swap out as desired. The common interface then guarantees that whichever choice the user makes the component will consistently interact with the others. Additionally, TRAJECTUM will also query of each chosen component which parameters it requires to function properly, and will check a user-specified parameter file to read these parameters.

In Fig. 1 we show the various components, as well as the interactions between them. The particle content can serve as an example of the statement that components automatically interact correctly with the others. In the current implementation, it is possible to choose from the particle content of URQMD [29,30] or SMASH [31–33]. Upon the user choosing one of these options, the choice is automatically communicated to the class handling particlization, which

needs this to work out which particles to produce, but also to compute viscous corrections and amounts of particles of each type to be produced. Upon the user choosing one of these options, the choice is automatically communicated to the class handling particlization, which uses this to compute a list of particles coming out of the hydrodynamic phase. In addition, the choice is automatically communicated to the class handling the equation of state (EOS), where it is needed for the hadron resonance gas (HRG) part of the equation of state.

As another example, consider the choice between $(2 + 1)$-dimensional $(2 + 1D)$ and $(3 + 1)$-dimensional $(3 + 1D)$ hydrodynamics. These models require different behaviors from the initial conditions, the partial differential equation (PDE) solvers, and the particlization algorithms. In particular, components like the initial conditions need extra parameters to set up a $3 + 1D$ simulation [34]. This does not lead to incompatibility between different sets of components though, as each component is guaranteed to work well together with whichever other components the user chooses. Instead, the components communicate, and change their behavior even according to which other components they are combined with. As a result, in case the user specifies $3 + 1D$ hydrodynamics, the component responsible for the initial conditions will make sure that the user also specifies the extra needed parameters.

A. Initial conditions

For the initial conditions, one can choose from the following options: (1) Monte Carlo Glauber [35–38]; (2) Ohio State University [39,40]; (3) TRENTO, both with and without nucleon substructure [12,27]; and (4) Gubser flow [41–43]. Although conceptually separate, in TRAJECTUM the prehydrodynamic phase is part of the initial conditions. Therefore the prehydrodynamic phase presented in the next subsection is only available when using TRENTO initial conditions. The Gubser initial conditions initialize the fluid according to the Gubser solution, but the time dependent Gubser flow will only be recovered if the ideal gas equation of state is used for the hydrodynamical evolution (see Sec. II D), as this is a conformal equation of state.

In the analysis presented in this paper, we use the TRENTO model with nuclear substructure [27]. The TRENTO model describes nuclei using a Woods-Saxon distribution, where a minimal internucleon distance $d_{\min}$ is imposed. Each nucleon subsequently is made up of n_c constituents, each with width given by

$$v_{\min} + \chi_{\text{struct}}(w - v_{\min}),$$

with $v_{\min} = 0.2$ fm, and w and χ_{struct} parameters. The constituents are sampled from a Gaussian distribution, such that the width of the matter distribution in each nucleon is w on average. After sampling the configuration of each nucleus, every pair of nucleons is labeled *wounded* with a probability based on their overlap, in such a way that the nucleon-nucleon cross section σ_{NN} equals the proton-proton cross section for that particular collision energy (63 mb for 2.76 TeV and 70 mb for 5.02 TeV). The constituents in each wounded nucleon each then source a thickness function with a Gaussian distribution,

where for each individual constituent the norm of the Gaussian is given by $N\gamma/n_c$. Here γ is sampled from a gamma distribution with width $\sigma_{\text{fluct}}\sqrt{n_c}$, and N and σ_{fluct} are parameters. In this way, for each nucleus a thickness function is computed, $\mathcal{T}_A$ and $\mathcal{T}_B$. These functions are combined together as

$$\mathcal{T} = \left(\frac{\mathcal{T}_A^p + \mathcal{T}_B^p}{2} \right)^{1/p}, \tag{1}$$

with p a parameter. This parametrizes a wide range of initial conditions, whereby in particular for $p = 1$ we have $\mathcal{T} = \mathcal{T}_A + \mathcal{T}_B$ (also called wounded nucleon scaling) and for $p = 0$ (1) reduces to $\mathcal{T} = \sqrt{\mathcal{T}_A \mathcal{T}_B}$, which is qualitatively similar to EKRT [14,21] or holography [15]. This thickness function $\mathcal{T}$ is then assumed to be boost invariant and proportional to the energy density, whereby the proportionality factor is given by $1/\tau$, as is appropriated for a pressureless fluid. TRAJECTUM is capable of extending this to the $3 + 1D$ case [34], but this was not used in this paper.

B. Prehydrodynamic phase

After obtaining the initial conditions, the matter is free streamed for a time τ_{fs} with velocity v_{fs}. This is a physically significant modification of the implementation [44–46] used in [23], where $v_{\text{fs}} = 1$. Since the free streaming evolution starts at proper time $\tau = 0^+$, the initial velocity in the longitudinal direction is zero and we have $T^{\mu\eta} = 0$, with η here referring to spacetime rapidity. This results in the following stress-energy tensor:

$$T^{\mu\nu}(x, y) = \frac{1}{2\pi\tau_{\text{fs}}} \int_0^{2\pi} d\phi \, \hat{p}^\mu \hat{p}^\nu$$
$$\times \mathcal{T}(x - v_{\text{fs}}\tau_{\text{fs}}\cos\phi, y - v_{\text{fs}}\tau_{\text{fs}}\sin\phi) \tag{2}$$

with

$$\hat{p}^\mu \hat{p}^\nu = \begin{pmatrix} 1 & v_{\text{fs}}\cos\phi & v_{\text{fs}}\sin\phi \\ v_{\text{fs}}\cos\phi & v_{\text{fs}}^2\cos^2\phi & v_{\text{fs}}^2\cos\phi\sin\phi \\ v_{\text{fs}}\sin\phi & v_{\text{fs}}^2\cos\phi\sin\phi & v_{\text{fs}}^2\sin^2\phi \end{pmatrix}.$$

It is interesting to examine the small τ_{fs} behavior of the prehydrodynamic stress-energy tensor. For this purpose, we expand (2) for small τ_{fs}. In this case, the integral can be performed analytically, resulting in the following expression:

$$T^{\mu\nu} = \begin{pmatrix} \frac{\mathcal{T}}{\tau_{\text{fs}}} & -\frac{\mathcal{T}v_{\text{fs}}^2}{2}\partial_x \log\mathcal{T} & -\frac{\mathcal{T}v_{\text{fs}}^2}{2}\partial_y \log\mathcal{T} \\ -\frac{\mathcal{T}v_{\text{fs}}^2}{2}\partial_x \log\mathcal{T} & \frac{\mathcal{T}v_{\text{fs}}^2}{2\tau_{\text{fs}}} & 0 \\ -\frac{\mathcal{T}v_{\text{fs}}^2}{2}\partial_y \log\mathcal{T} & 0 & \frac{\mathcal{T}v_{\text{fs}}^2}{2\tau_{\text{fs}}} \end{pmatrix},$$

where we omit the spatial arguments for brevity. By solving the eigenvalue problem

$$T^{\mu\nu}u_\nu = \rho u^\mu,$$

we can then extract the fluid velocity:

$$u^\mu = \left(1, -\frac{v_{\text{fs}}^2\tau_{\text{fs}}}{2 + v_{\text{fs}}^2}\partial_x \log\mathcal{T}, -\frac{v_{\text{fs}}^2\tau_{\text{fs}}}{2 + v_{\text{fs}}^2}\partial_y \log\mathcal{T}\right) + \mathcal{O}(\tau_{\text{fs}}^2). \tag{3}$$

Using holographic simulations [47–49] it was found in [50] that at strong coupling in a conformal theory an appropriate initial velocity can be initialized with $u^\perp \approx -\frac{\tau}{3.0}\nabla_\perp \log \mathcal{T}$. Quite curiously this exactly agrees with (3) for $v_{\rm fs} = 1$. This is somewhat coincidental, since the holographic evolution is quite unlike free streaming. Free streaming has zero longitudinal pressure, whereas the holographic evolution has a large negative longitudinal pressure at early times, but quickly hydrodynamizes towards a positive longitudinal pressure (see also [51,52]). Depending on the starting time of hydrodynamics this pressure can average to the zero free streaming result. Note also that this does not mean that the entire stress tensor agrees. In particular, the holographic result features fast hydrodynamization, whereas we will see in Sec. III F that in our model the prehydrodynamic phase takes the fluid away from hydrodynamics, and that a free streaming velocity of around 0.93 minimizes deviations from hydrodynamics for the bulk viscous pressure. For the shear tensor, minimization of deviations from hydrodynamics is achieved for $v_{\rm fs} \approx 0.6$.

C. Hydrodynamic phase

For the hydrodynamics model, one can choose between one including only first order transport coefficients, and one including also the second order transport coefficients from the 14-moment approximation [53]. Both of these choices are available in both $2 + 1$D (assuming boost invariance) and $3 + 1$D versions. In the analysis presented in this paper, we used the second order $2 + 1$D version. This solves the conservation equations for the stress-energy equation

$$\partial_\mu T^{\mu\nu} = 0, \qquad (4)$$

with ∂_μ the covariant derivative. We use the constitutive relation

$$T^{\mu\nu} = \rho u^\mu u^\nu - (P + \Pi)\Delta^{\mu\nu} + \pi^{\mu\nu}, \qquad (5)$$

with $\Delta^{\mu\nu} = g^{\mu\nu} - u^\mu u^\nu$, u^μ the local fluid velocity, and $g_{\mu\nu}$ a mostly minus metric. The relation between the energy density ρ and the pressure P is given by the equation of state, which is discussed in Sec. II D. There are also two viscous corrections, $\pi^{\mu\nu}$ and Π. The shear tensor $\pi^{\mu\nu}$ is a symmetric traceless tensor which is transverse to u^μ. It and the bulk pressure Π obey the second order 14-moment approximation equations of motion [53], where we keep only the transport coefficients that have been computed explicitly in [53]:

$$D\Pi = -\frac{1}{\tau_\Pi}[\Pi + \zeta\nabla\cdot u + \delta_{\Pi\Pi}\nabla\cdot u\Pi$$
$$- \lambda_{\Pi\pi}\pi^{\mu\nu}\sigma_{\mu\nu}], \qquad (6)$$

$$\Delta^\mu_\alpha\Delta^\nu_\beta D\pi^{\alpha\beta} = -\frac{1}{\tau_\pi}\big[\pi^{\mu\nu} - 2\eta\sigma^{\mu\nu}$$
$$+ \delta_{\pi\pi}\pi^{\mu\nu}\nabla\cdot u - \phi_7\pi^{\langle\mu}_\alpha\pi^{\nu\rangle\alpha}$$
$$+ \tau_{\pi\pi}\pi^{\langle\mu}_\alpha\sigma^{\nu\rangle\alpha} - \lambda_{\pi\Pi}\Pi\sigma^{\mu\nu}\big]. \qquad (7)$$

Here $D \equiv u^\mu\partial_\mu$, $\nabla^\mu \equiv \Delta^{\mu\nu}\partial_\nu$, and $\sigma^{\mu\nu} \equiv \frac{1}{2}(\nabla^\mu u^\nu + \nabla^\nu u^\mu) - \frac{1}{3}(\nabla\cdot u)\Delta^{\mu\nu}$. These constitutive equations do not include vortical transport as well as some second order transport coefficients involving the bulk viscous pressure Π.

Both of these are expected to be small in boost invariant settings with approximate conformal symmetry.

D. Equation of state and transport coefficients

For the equation of state and transport coefficients, there are two possible choices available in TRAJECTUM.

1. Ideal gas equation of state $P = \alpha T^4$, with α a constant specifiable by the user

This equation of state comes with constant user-specifiable values for the dimensionless ratios of transport coefficients:

$$\frac{\eta}{s}, \quad \frac{\zeta}{s}, \quad \frac{\tau_\pi s T}{\eta}, \quad \frac{\tau_\Pi s T}{\zeta}, \quad \frac{\delta_{\pi\pi}}{\tau_\pi},$$

$$\phi_7 P, \quad \frac{\tau_{\pi\pi}}{\tau_\pi}, \quad \frac{\lambda_{\pi\Pi}}{\tau_\pi}, \quad \frac{\delta_{\Pi\Pi}}{\tau_\Pi}, \quad \frac{\lambda_{\Pi\pi}}{\tau_\Pi}.$$

Note that the dimensionless ratios for τ_Π and $\lambda_{\Pi\pi}$ differ from that used by the other equation of state and transport coefficients model, by factors $(1/3 - c_s^2)^2$ and $(1/3 - c_s^2)^{-1}$, respectively. The reason for this is that for this particular equation of state the speed of sound c_s equals $\sqrt{1/3}$, which would cause τ_Π to be infinite, with similar issues for $\lambda_{\Pi\pi}$.

2. Lattice QCD/HRG hybrid equation of state with temperature dependent shear and bulk viscosities

In this model, the equation of state is a hybrid constructed from the HotQCD lattice equation of state, together with the HRG constructed from the chosen particle content [24,54,55]. Here the analytical fit described in [55] to the lattice equation of state is used instead of a tabulated form, and the transition between the two parts of the equation of state lies in the region 165–200 MeV. A consistent switch from hydrodynamics to a HRG is therefore possible for temperatures below 165 MeV, as in this region the EOS is constructed from the HRG (see also Sec. II F). The user can also specify mostly the same dimensionless ratios of transport coefficients as in the ideal gas model, except for the bulk relaxation time τ_Π and the coefficient $\lambda_{\Pi\pi}$, for which we instead specify

$$\frac{\tau_\Pi s T (1/3 - c_s^2)^2}{\zeta}, \quad \frac{\lambda_{\Pi\pi}}{\tau_\Pi (1/3 - c_s^2)}.$$

Another difference from the previous model is that we now allow for temperature dependence in the dimensionless ratios for both the shear and bulk viscosities, which are given by the following expressions:

$$\frac{\eta}{s} = \begin{cases} a + b(T - T_c)\big(\frac{T}{T_c}\big)^c & T \geqslant T_c, \\ 0.06 & T < T_c, \end{cases}$$

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

with $T_c = 154$ MeV, and $a = (\eta/s)_{\rm min}$, $b = (\eta/s)_{\rm slope}$, and $c = (\eta/s)_{\rm crv}$, $(\zeta/s)_{\rm max}$, $(\zeta/s)_{\rm width}$, and $(\zeta/s)_{T_0}$ user-specifiable parameters. In the analysis presented in this paper, the lattice QCD/HRG hybrid equation of state with temperature dependent viscosities was used.

In the Bayesian analysis, we varied all parameters governing the shear and bulk viscosities η/s and ζ/s, as well as the following three second order transport coefficients:

$$\frac{\tau_\Pi sT \left(1/3 - c_s^2\right)^2}{\zeta}, \quad \frac{\tau_\pi sT}{\eta}, \quad \frac{\tau_{\pi\pi}}{\tau_\pi}.$$

We keep the remaining coefficients

$$\frac{\delta_{\Pi\Pi}}{\tau_\Pi} = \frac{2}{3}, \quad \frac{\lambda_{\Pi\pi}}{\tau_\Pi \left(1/3 - c_s^2\right)} = \frac{8}{5}, \quad \frac{\delta_{\pi\pi}}{\tau_\pi} = \frac{4}{3},$$

$$\phi_7 P = \frac{9}{70}, \quad \frac{\lambda_{\pi\Pi}}{\tau_\pi} = \frac{6}{5}$$

fixed to the kinetic theory values [53] in the limit of small particle masses, which is a common choice from earlier analyses. In this paper we made the choice to vary the two relaxation times, which most likely have the largest influence on the hydrodynamical evolution, as well as one additional coefficient $\tau_{\pi\pi}$ that does not vanish in the conformal limit. In future work it would be important to vary also other second order coefficients, but we note that varying more parameters can come at a significant computational cost.

E. PDE solvers

There are two choices available for the PDE solvers that solve the equations of motion (4), (6), and (7).

1. Finite difference [56]

This method discretizes spatial derivatives by simply taking a symmetric finite difference. An advantage of this method is that it is fast since it is simple and does not need a lot of computations. A disadvantage, though, is that this method is not always stable, and can diverge in the presence of shocks or other large spatial gradients, especially if the shear viscosity is small. Of course, in the setting of a Bayesian analysis, this is not a good tradeoff, as we need to be able to generate theoretical predictions even for very "exotic" choices of parameters.

2. MUSCL [57,58]

This method separates the equations of motion into a part describing conserved quantities, plus a source term. The conserved part is then solved by examining the fluxes between each pair of neighboring grid points, where the left- and right-moving fluxes are combined using the Kurganov-Tadmor algorithm [57]. Crucially, the fluxes are subsequently limited using the minmod flux limiter, which guarantees the stability under all circumstances. For small gradients the flux limiting vanishes and the result of MUSCL agrees exactly. Stability makes this method well suited for our purposes, as it will produce an accurate result under all circumstances. The price to pay for stability, though, is speed, as the MUSCL algorithm is more complicated. The most computationally expensive part of using MUSCL is the computation of the velocity and energy density from the primary variables ($T^{\tau\tau}$, $T^{\tau x}$, $T^{\tau y}$, $\pi^{\mu\nu}$, and Π). This step is required five times more often in MUSCL as compared to finite difference in order to evaluate its flux limiting feature that makes the code stable in presence of

shocks. For this reason, we wrote a "fast" version of MUSCL that applies the same flux limiting procedure directly to the velocities instead of computing the velocities from the flux limited stress tensors. In a regime in which the velocity can be approximated as a linear function of the stress tensor, this modified flux limiting procedure is identical to the original MUSCL. We have also verified that this fast version is just as stable as the unmodified version, and that the difference between both results is negligible for our simulations. In this paper, we used the fast version of MUSCL.

F. Particlization and hadronic phase

For the particlization, only the Cooper-Frye procedure with viscous corrections as presented in [24,59] is available. This model finds the isotemperature surface of the switching temperature T_{switch}, and subsequently generates particles at that surface from a thermal distribution in the local restframe with temperature T_{switch}. In the Bayesian analysis we vary this particlization temperature. For a continuous evolution the equation of state must match the one derived from the particle content used in particlization, which in our implementation is only the case for the lattice QCD/HRG hybrid equation of state. The other equation of state is implemented mostly for testing purposes and is not used in the results presented.

For the viscous corrections at each location, the spatial momenta p_i of the sampled particles are subsequently rescaled by

$$p_i \mapsto p_i + \sum_j [(\lambda_{\text{shear}})_{ij} + \lambda_{\text{bulk}}\delta_{ij}]p_j.$$

Here, $(\lambda_{\text{shear}})_{ij} \propto \pi_{ij}$, where the proportionality constant, which is species independent, is determined by demanding that to linear order the resulting HRG has the same shear tensor as the fluid from which π_{ij} was taken. Likewise, λ_{bulk} is computed from a similar HRG computation, however the dependence of λ_{bulk} on Π is nonlinear. Lastly, the density of sampled particles is adjusted as a function of Π in order to match the energy density of the fluid [60]. In this way, the entire stress-energy tensor is by construction on average continuous across the particlization surface. The nonlinearity of the dependence of λ_{bulk} on Π guarantees that large negative values for Π cannot flip the sign of the momenta of the sampled particles. It is important to note that this method is only one way to generate a continuous stress-energy tensor and for future work it would be interesting to verify that different methods do not strongly affect the results, such as, e.g., studied in [61,62].

The sampled particles subsequently undergo a hadronic cascade. TRAJECTUM does not incorporate this hadronic cascade itself, but it does offer the flexibility to work with more than one possible code for this cascade, namely, URQMD [29,30] and SMASH [31–33]. These two codes require their inputs to be in slightly different formats, but importantly for the whole simulation their resonance content is different. This means that in order for TRAJECTUM to consistently interact with the hadronic cascade, it must use the exact same resonance content, to indeed ensure continuity across the particlization surface. In this paper, we have solely used
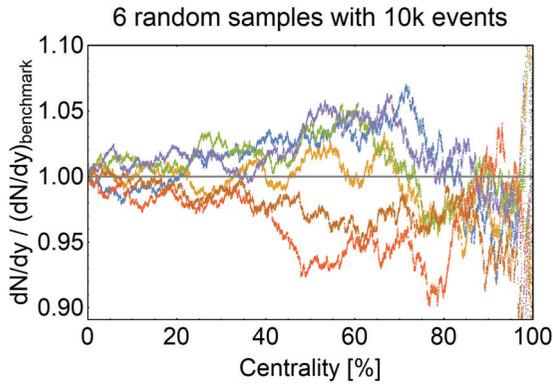
FIG. 2. Particle multiplicity where centrality has been determined from 10 000 events, relative to the true distribution from which the samples were obtained.

SMASH, as that contains a few more resonances as compared to URQMD.

G. Validation of TRAJECTUM

Because TRAJECTUM is a new code, it is important to validate its results. We performed several checks to verify the correct functioning of the code.

(1) To test the equations of motion, we compared the numerical computation of Gubser flow to the analytically known result [41–43]. This leads to good agreement, but in Gubser flow there is no bulk viscosity, and there is also little dependence on second order coefficients. For this reason we performed an additional check, by comparing the implementation of the equations of motion for generic events to an independently performed computation in MATHEMATICA. The two computations agree to machine precision.

(2) The particlization procedure we use is designed so that on average the stress-energy tensor is continuous across the particlization surface. We checked that this is indeed true to good precision, by oversampling the number of produced particles, producing a statistically significant sample.

(3) We compared various quantities such as the initial state eccentricities ϵ_n between our implementation of TRENTO and the original version [63], finding good agreement.

(4) We computed the observables from [23] using the maximum *a posteriori* (MAP) values for the parameters specified there. The results from TRAJECTUM agree well with theirs.

H. Centrality classes

A common but relatively large problem is the determination of centrality classes. When using a finite number of minimum bias events [$\mathcal{O}(10^4)$] the boundary between centrality classes can easily be the dominant error, especially for observables that depend strongly on centrality. This is illustrated in Fig. 2, where we show the ratio of the multiplicity

of six random samples to the "true" distribution (taken from Sec. IV, we also used this distribution to draw the random samples). For a realistic simulation using 10 000 minimum bias events this results in errors up to 10%, which is especially unfortunate as this particular observable has low experimental uncertainties.

For the theoretical centrality determination it is however possible to use a trick to resolve this statistical problem without running extra hydrodynamic simulations. First, it is known that for PbPb the initial entropy correlates very well with final state multiplicity. In turn, the initial entropy correlates very well with the spatial integral of the initial thickness function in TRENTO, for which one does not need to perform the computationally expensive free streaming stage. The way in which one can use these facts is to generate a large number $\mathcal{O}(10^5)$ of TRENTO simulations, and in this way generate an accurate distribution of initial thickness functions. We then sort the simulations into bins, and we save the upper and lower initial thickness functions for each bin. Subsequently, for each bin we run the entire hydrodynamical evolution for some fixed number of events with initial thickness function within the bin boundaries. In this way, we guarantee that for these events the initial thickness functions are distributed the same way as for the large set described above. Since the initial thickness function correlates well with the final multiplicity, we generate final multiplicities with the correct distribution as well, and hence get a determination of centrality as if we had used $\mathcal{O}(10^5)$ events. In the limit of a large number of events, the distribution still converges to the true distribution as long as the number of initial condition simulations is larger than the number of full hydrodynamics computations. The only difference is that it converges to the true distribution faster than without this optimization, so the resulting events are still minimal bias. In particular, since our centrality determination is based on final state multiplicity this means that even if the final state multiplicity does not correlate well with the initial thickness function the final distribution will still converge to the true distribution, albeit at a slower rate than with a strong correlation and our improved algorithm.

One can also use this method to generate weighted samples, by simply generating more events for the desired bins. For our pPb simulations, we bias the selection towards higher multiplicities, and subsequently weigh events in the analysis accordingly. In this way we obtain better statistics for (relatively rare) higher multiplicity events.

A related, but separate issue is the experimental centrality selection. In our boost invariant simulations our only choice is to bin events according to their multiplicity at midrapidity. Experimentally, however, centrality classes are often selected by binning at relatively high rapidity (ALICE often uses the V0 detectors, located at pseudorapidities -3.7 up to -1.7 and at 2.8 up to 5.1). For PbPb collisions this bias is not so severe, since the V0 amplitude correlates strongly with multiplicity at midrapidity (see Fig. 15 in [64]), but for pPb this can introduce a serious bias [65]. Ideally we would compare to experimental data that also determine centrality at midrapidity, but unfortunately this is rarely available. This is part of the reason why we chose to independently fit the normalizations in PbPb and pPb collisions.
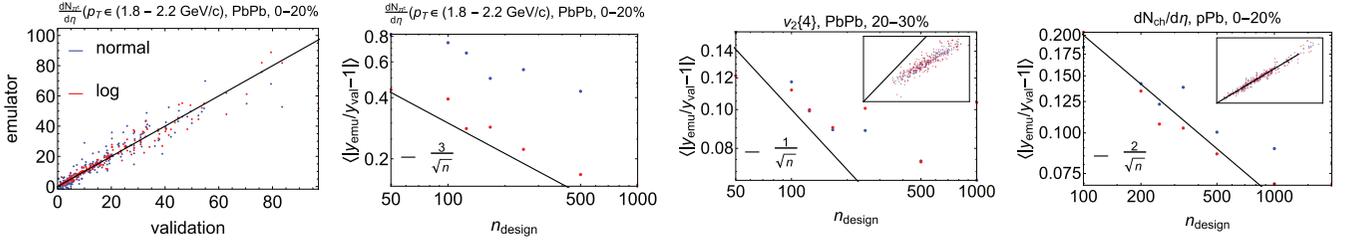
FIG. 3. Left: Emulator prediction shown against model output (validation) for charged pion multiplicity in the 1.8–2.2-GeV p_T bin for the 0–20% centrality class using 1000 design points. When emulating the logarithm of the yield (labeled log, red) the accuracy of the emulator is much improved. The right three plots compare the emulator error, defined as $\langle |y_{\text{emu}}/y_{\text{val}} - 1| \rangle$ vs the number of design points for the left plot, $v_2\{4\}$ at 20–30% centrality and $dN_{\text{ch}}/d\eta$ in pPb at 0–20% centrality, respectively (validation vs emulator results are shown as insets). For $v_2\{4\}$ the statistical error becomes dominant at about 250 design points and the log scaling does not improve. For pPb the emulator error is relatively large, which is the reason why this paper uses 2000 design points for pPb.

III. RESULTS

A. Emulator

As our aim is to perform a global analysis of heavy ion collisions depending on all parameters specified in Sec. II it is necessary to obtain a large sample of events for many points in this parameter space. Even for a single parameter point it is however computationally expensive to obtain sufficient statistics for interesting observables, and it is therefore standard practice to evaluate the model at a number n_{design} of "design" points in the parameter space and employ an emulator to interpolate those results to any point in parameter space. These design points are typically distributed on a latin hypercube with reasonable (physical) limits on the parameter range [24,66], that in our case will equate with the range of prior probabilities.

To further reduce computational time we decided to leave the nucleon-nucleon cross section σ_{NN} as a parameter. The major advantage of doing so is that a single scan over the design points can be used to compare to experimental data from several colliding energies. The addition of one extra parameter is much more efficient than performing two separate scans. Each energy will have its own normalization and cross section, whereby either we can choose to fix the cross section to its independently measured proton-proton value, or we can leave it up to the Bayesian analysis to estimate the cross section. In the current paper we fix $\sigma_{\text{NN}} = 63\,(70)$ mb for collisions at $\sqrt{s_{\text{NN}}} = 2.76\,(5.02)$ TeV.

For the emulator we train a Gaussian process emulator with a squared exponential (a Gaussian) plus a noise kernel [24,25,67], as implemented in the SCIKIT-LEARN library for PYTHON. Even though the emulator itself estimates its own uncertainty it is important to validate its output, which is shown in Fig. 3 for three representative observables. The uncertainty of the emulator is not shown for clarity, but corresponds well with the actual validation error (equal to the deviation from the black diagonal line). For some observables such as transverse momentum spectra the spread of points is such that for points with small values a small emulator error can make a large relative change. In these situations it is often better to train the emulator on the logarithm of the datapoints, and afterwards exponentiate the emulator re-
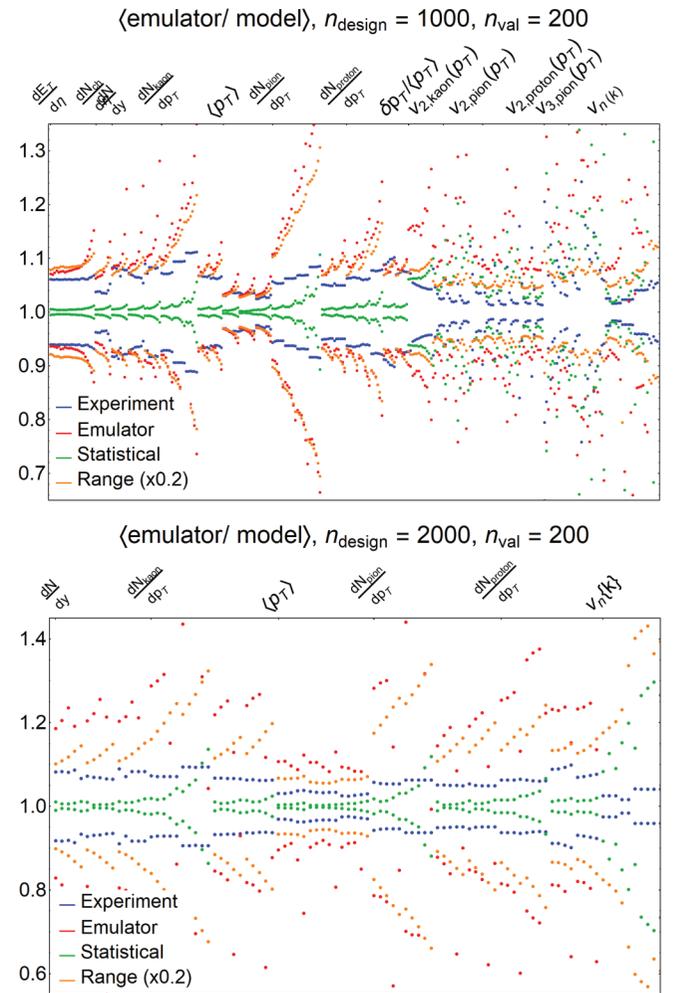
FIG. 4. We show the validation of our emulator for PbPb (top) and pPb (bottom) systems in red, whereby we show the average of the emulator discrepancy with the true model result for 200 validation points (red). Experimental uncertainties are shown and often of the same order as the emulator uncertainty, which in turn is much smaller than the parameter range probed (orange). Statistical model uncertainty is dominant for a few statistically difficult observables (green). The pPb system is much harder to emulate, as also found in [25].
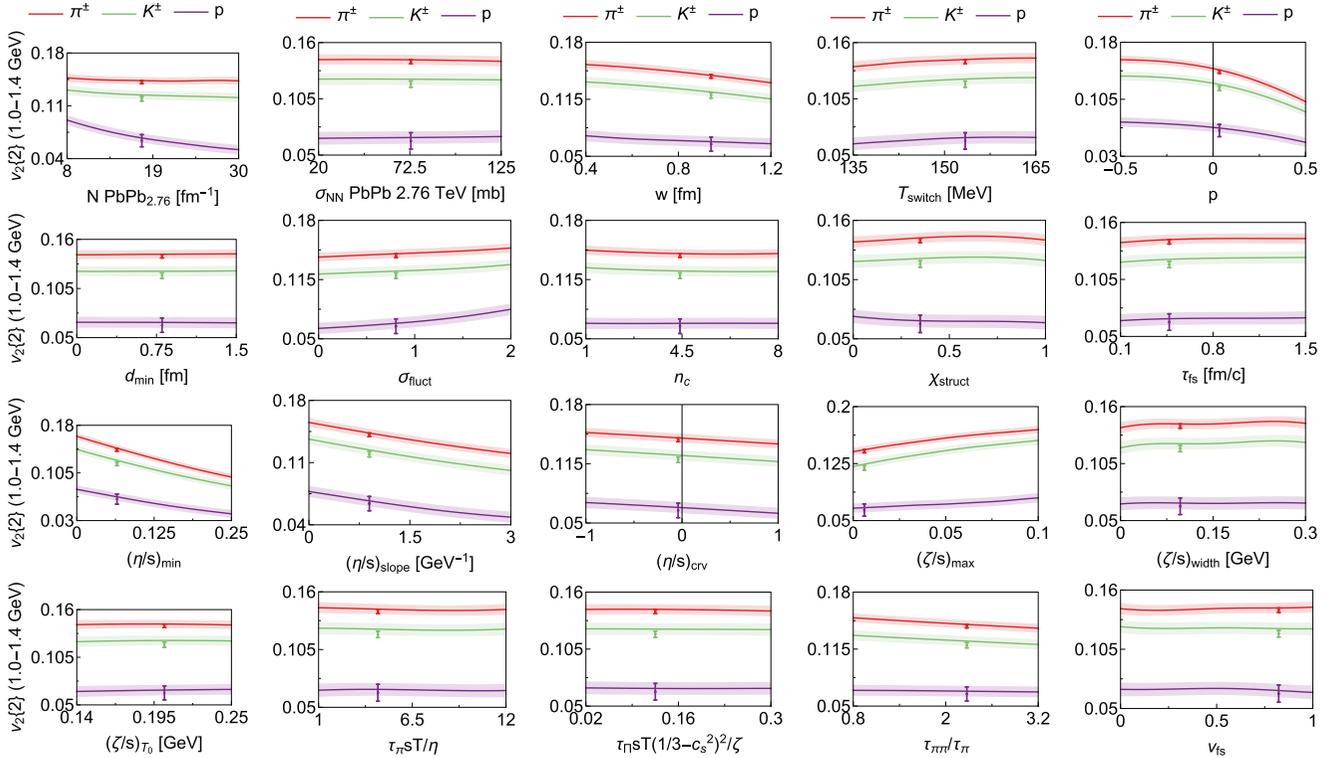
FIG. 5. We show $v_2\{2\}$ for the 1.0–1.4-GeV p_T bin for pions, kaons, and protons in the 20–30% centrality class as a function of our model parameters. All parameters are fixed to the optimal point found in [28], with the exception of the varying parameter displayed. The relevant datapoints [68] are displayed at its optimal point. As expected the elliptic flow decreases as the shear viscosity is increased. Similarly strong dependencies are seen for the TRENTO parameter p [affecting the initial geometric anisotropy, see (1)] and for protons the norm and fluctuations σ_{fluct} are important.

sults. This is shown as the "normal" versus "log" result in Fig. 3, and gains can be as large as a factor 3 reduction in emulator error. In this paper we decided to designate $dN/d\eta$, $dE_T/d\eta$, and the transverse momentum spectra of pions, kaons, and protons as observables on which to perform the log transformation.

For an accurate emulation with limited computing resources it is important to balance the number of design points versus the number of events per design point. For some observables, such as the $v_2\{4\}$ shown in Fig. 3 (the observable $v_n\{k\}$ is defined in Sec. IV C), statistics is often the limiting factor and increasing the number of design points does not improve the emulator uncertainty. For statistically easier observables, such as the multiplicity, the emulator error decreases as approximately $1/\sqrt{n_{\text{design}}}$ up to larger n_{design}. Second, we noticed that pPb observables are more difficult to emulate. For these reasons we chose to model PbPb and pPb with 1000 and 2000 design points, using 6000 and 40 000 events per design point, respectively.

To emulate the full dataset it is important to perform a linear transformation on the datapoints, as these are often highly correlated. This is done using a principal component analysis (PCA), which extracts the few principal components (PCs) that contain independent information. In our dataset we found that the first 25 PCs per colliding system suffice, capturing over 97% of the variation in the data. An important criterion on the number of PCAs to use is the estimated

(statistical) noise found by the emulator training. For our case only from the ninth PC onward the noise levels can surpass 10%, and only at the 22nd and 24th the noise was dominant (over 40%), which led us to conclude that 25 PCs is sufficient to capture all nontrivial information. Performing the PCA also automatically gives a good estimate of theoretical correlations among the datapoints, which will be important later (the log transformation above also needs to be applied to the correlation matrix).

After performing the PCA it is again necessary to validate the emulator, which is shown in Fig. 4 for PbPb (top) and pPb (bottom) in red. This is estimated by taking the model values m and emulator predictions e for the same 200 validation points for all our observables, where the red points then correspond to $\langle e/m \rangle \pm s$, with s the standard deviation of e/m. For comparison we also show the experimental error of the datapoint (blue), the average statistical error of our model (green), and the range of model predictions given our prior range, defined as the standard deviation over the mean (orange). Ideally the emulator uncertainty should be smaller than the experimental uncertainty, but this is only the case for a few observables and adding further design points is hence likely to improve our results. The statistical uncertainty is only dominant for statistically difficult observables, such as p_T-differential v_2, and can be improved by running design points with more events. Lastly, we find that the range of model output is often roughly five times the emulator uncertainty,
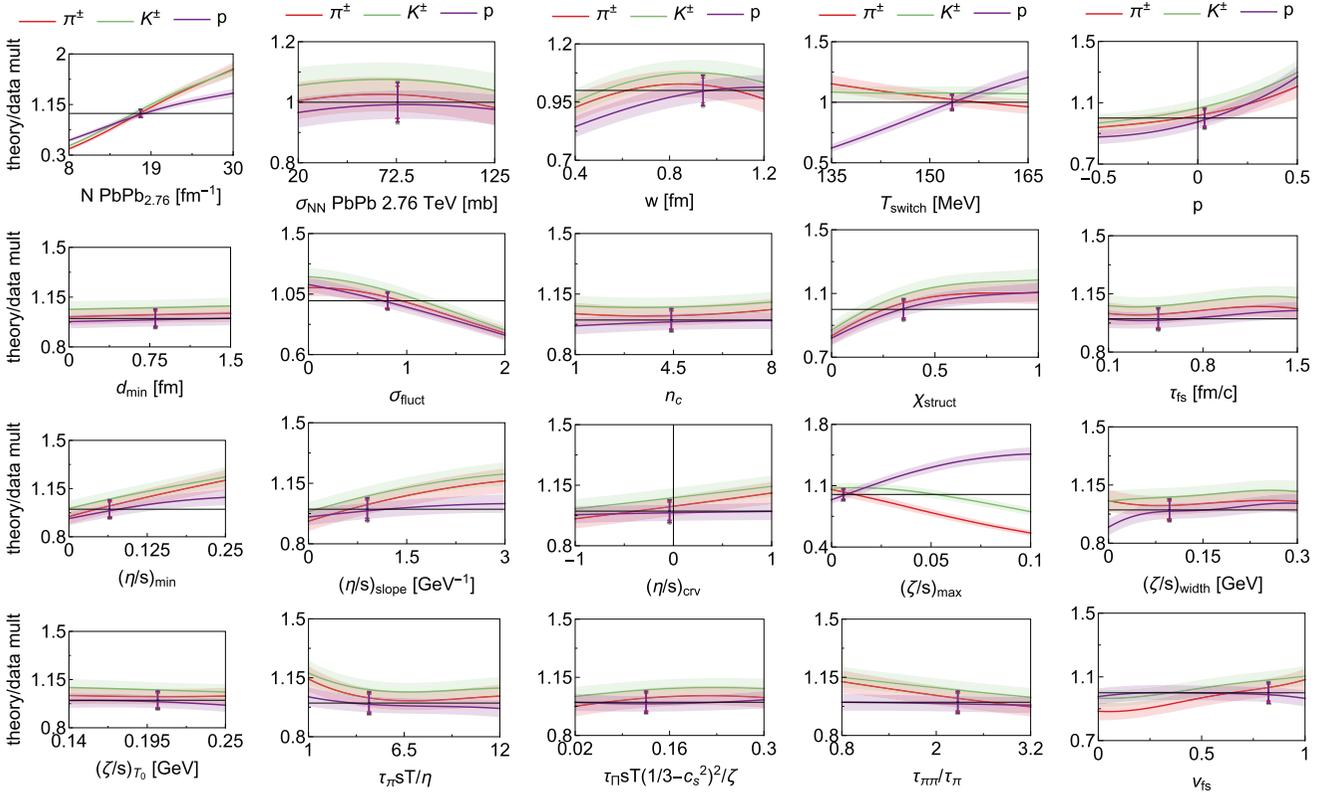
FIG. 6. Similar to Fig. 5 we show the multiplicity for the 1.0–1.4-GeV p_T bin for pions, kaons, and protons in the 20–30% centrality class as a function of our model parameters normalized to the relevant datapoints [69]. As expected the multiplicity depends strongly on the norm. Strong dependencies for the proton to pion ratio are seen for the switching temperature (as expected) as well as the bulk viscosity.

which gives a good discriminating power. As is clear from the plots the emulator performs worse for pPb collisions than for PbPb collisions (consistent with [25]), which is the reason why we ran pPb with 2000 design points and used only 1000 design points for PbPb.

B. Parameter dependence of observables

Once we have a fully trained emulator we can evaluate the emulator on a set of parameters and immediately obtain all observables, reducing the computational time from days to fractions of a second. As we have 514 observables as a function of a 20-dimensional parameter space it is impossible to display all this information here, but in Figs. 5 and 6 we, respectively, show the representative $v_2\{2\}$ and multiplicity in the intermediate p_T bin of 1.0–1.4 GeV for pions, kaons, and protons in the 20–30% centrality class. All panels use the optimal parameters found in [28], with the exception of the parameter being displayed. The experimental datapoint is shown at the optimal value for the parameter itself, such that at this point all curves agree exactly with each other and are indeed also close to the actual datapoint.

A few things are as qualitatively expected: $v_2\{2\}$ decreases as we increase either $(\eta/s)_{\mathrm{min}}$ or $(\eta/s)_{\mathrm{slope}}$. $v_2\{2\}$ depends sensitively on the TRENTO parameter p. The multiplicity depends strongly on the norm (but perhaps surprisingly less so for protons) and the proton to pion ratio increases for higher T_{switch}. As also expected the dependence on second order transport

coefficients is much less pronounced, and it is not possible to extract a preferred value by just looking at these plots; we will come back to a global analysis shortly. We also see a
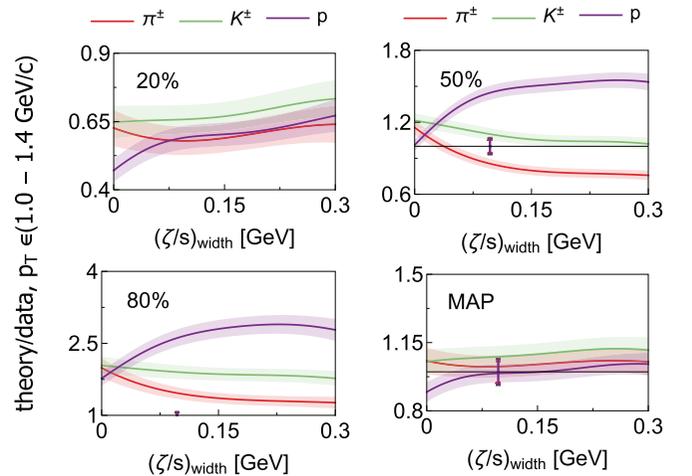
FIG. 7. To illustrate the interaction of the norm and width of the bulk viscosity we fix all parameters at, respectively, 20, 50, and 80% of their range and at the MAP value. Clearly, when the norm is at 80% there is a much stronger dependence on the width. Only if either the norm or the width is close to zero it is possible to simultaneously fit the pion, kaon, and proton multiplicities in this intermediate p_T bin.
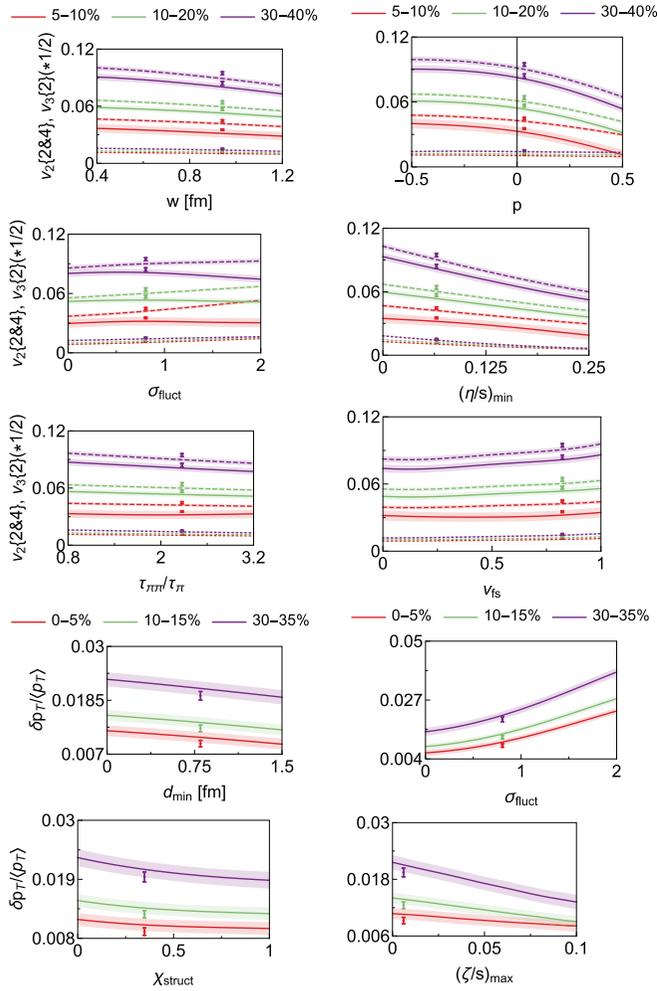
FIG. 8. To illustrate the centrality dependence more clearly we show $v_2\{4\}$ (solid), $v_2\{2\}$ (dashed), and $v_3\{2\}$ (dotted) coefficients (top) and δp_T fluctuations (bottom) for several centrality classes as a function of representative parameters for 2.76-TeV PbPb collisions. There is a clear increase in both δp_T as well as the difference between $v_2\{4\}$ and $v_2\{2\}$ when increasing fluctuations through σ_{fluct}.

FIG. 9. We show the p_T-differential $v_2\{2\}$ for five p_T bins as a function of the shear and bulk viscosity for the 20–30% centrality class (top), as well as the integrated $v_2\{2\}$ for several centrality classes (middle). Quite interestingly the bulk viscosity causes an increase of $v_2\{2\}$ in all p_T bins, even though the integrated $v_2\{2\}$ decreases. This can be understood by the fact that the mean p_T decreases as a function of the bulk viscosity (bottom left). We also note an increase in mean p_T as a function of v_{fs} (bottom right).

cosity for the optimal parameters, which in turn (physically) means there is little to no effect of $(\zeta/s)_{\text{width}}$. Note though that this (physical) correlation is not explicitly put in for the emulator, and even for $(\zeta/s)_{\text{max}} = 0$ some dependence on the width is still expected, at least within the error band of the emulator. All this can be clearly seen in Fig. 7, which shows the sensitivity of the number of pions, kaons, and protons with $p_T \in 1.0$–1.4 GeV to $(\zeta/s)_{\text{width}}$ where all parameters are chosen to lie at 20, 50, and 80% of the prior range or at the MAP point, respectively. In this case we hence have that $(\zeta/s)_{\text{max}} = (0.06, 0.15, 0.24, 0.0060)$. Clearly the largest variation is present for $(\zeta/s)_{\text{max}} = 0.24$, which incidentally also has the highest overall yield since the norm N is then also fixed at 80%. The figure also shows that unless either $(\zeta/s)_{\text{max}}$ or $(\zeta/s)_{\text{width}}$ is close to zero it is hard to obtain simultaneous agreement with the experimental datapoints for all identified particles.

Some observables have a strong centrality dependence. For this reason we show in Fig. 8 the anisotropic flow coefficients $v_2\{2\}$, $v_2\{4\}$, and $v_3\{2\}$ as well as the transverse momentum fluctuations for three different centrality classes as a function of a few representative parameters. As can be expected the transverse momentum fluctuations depend sensitively on the fluctuating TRENTO parameter σ_{fluct} (see Sec. II A). A large bulk viscosity on the other hand reduces the p_T fluctuations. The difference between $v_2\{4\}$ and $v_2\{2\}$ increases strongly as σ_{fluct} increases, especially for central collisions. This is in

strong dependence on the bulk viscosity through $(\zeta/s)_{\text{max}}$ on both the particle ratios as well as $v_2\{2\}$. This is partly due to the influence of the bulk viscosity on hydrodynamics, and partly due to its influence on the stress tensor at the moment of particlization (see [61] for a more detailed study on the influence of particlization; we use what they call the PTB prescription). In the Appendix we show similar results for pPb collisions, this time showing the elliptic flow coefficient and the mean kaon transverse momentum for several centrality classes. As already mentioned emulator uncertainty is larger for this smaller system, but it is still clear that dependence on our parameters is increasingly nontrivial as we go to these smaller systems.

Of course the sensitivity of observables on a parameter may (strongly) depend on the location in the parameter space. From Figs. 5 and 6 it may for instance seem that observables are rather insensitive to the width of the bulk viscosity. This, however, is due to the low value of the norm for the bulk vis-

FIG. 10. Mean (solid) and median (dashed) combined with 68 and 95% confidence intervals (shaded dark and light, respectively), computed over all 600 walkers as a function of the step number. After about 9000 steps the distributions are converged (apart from some statistical fluctuations) and the remainder of the chain points can be used as an approximation for the true posterior distribution.

agreement with the statistical interpretation for the difference between these two estimates of elliptic flow [70].

Lastly, it can be surprising that the $v_2\{2\}$ in the 1.0–1.4-GeV bin in Fig. 5 increases as a function of the bulk viscosity, whereas from the dissipative character of transport coefficients the viscosity is expected to isotropize the plasma. In Fig. 9 we show this in more detail, both for the shear and bulk viscosity, and indeed for all p_T bins the $v_2\{2\}$ increases, but nevertheless the integrated $v_2\{2\}$ decreases. This can be understood by the fact that the $v_2\{2\}$ is larger for larger transverse momentum and the fact that the bulk viscosity reduces the mean transverse momentum (Fig. 9 bottom).

C. Bayesian posterior estimate

After having obtained an emulator and hence model predictions as a function of our input parameters the next step is to estimate the probability distributions for our parameter space $\boldsymbol{x}$. Using Bayes's theorem this is done by first assuming a prior probability distribution, which is typically flat within a given (physically motivated) range and zero outside this range, and by second reweighing this distribution taking into account new evidence from comparing the model output with some set of experimental data. Mathematically, this amounts to a posterior distribution given by

$$\mathcal{P}(\boldsymbol{x}|\mathbf{y}_{\mathrm{exp}}) = \frac{e^{-\Delta^2/2}}{\sqrt{(2\pi)^n \det{[\Sigma(\boldsymbol{x})]}}} \mathcal{P}(\boldsymbol{x}) \qquad (8)$$

with $\mathcal{P}(\boldsymbol{x})$ the (flat) prior probability density and where

$$\Delta^2 = [\mathbf{y}(\boldsymbol{x}) - \mathbf{y}_{\mathrm{exp}}]\Sigma(\boldsymbol{x})^{-1}[\mathbf{y}(\boldsymbol{x}) - \mathbf{y}_{\mathrm{exp}}], \qquad (9)$$

with $\mathbf{y}(\boldsymbol{x})$ the predicted data for parameters $\boldsymbol{x}$, $\mathbf{y}_{\mathrm{exp}}$ the n experimental datapoints, and $\Sigma(\mathbf{x}) = \Sigma_{\mathrm{emu}}(\mathbf{x}) + \Sigma_{\mathrm{exp}}$ the sum of the emulator and experimental covariance uncertainty matrices. For the emulator this follows from the correlations in the theoretical model, as determined by a PCA. In principle the experimental matrix should be provided by the relevant experiment, but as this is rarely available we follow the simple

prescription from [24]. Numerically it is often more convenient to work with the logarithm of (8), which we will refer to as the log likelihood (LL), and we note that (8) should still be normalized.

For the experimental datapoints we use the same set of 514 observables as in [28], which built upon [24,25]. These include PbPb charged particle multiplicity $dN_{ch}/d\eta$ at 2.76 [71] and 5.02 TeV [72], transverse energy $dE_T/d\eta$ at 2.76 TeV [73], identified yields dN/dy and mean p_T for pions, kaons, and protons at 2.76 TeV [69], integrated anisotropic flow $v_n\{k\}$ for both 2.76 and 5.02 TeV [74], and p_T fluctuations [75] at 2.76 TeV. All of these use centrality classes of width 2.5% (for central transverse energy) up to 20% (for peripheral $v_n\{k\}$). On top of this we added identified transverse momentum spectra using six coarse grained p_T bins separated at (0.5, 0.75, 1.0, 1.4, 1.8, 2.2, 3.0) GeV both for PbPb at 2.76 [69] and pPb at 5.02 TeV [76], anisotropic identified flow coefficients using the same p_T bins (statistics allowing) at 2.76 [68] and 5.02 TeV [77]. As in [25] we use $\tilde{v}_n\{k\}$ anisotropic flow coefficients for pPb at 5.02 TeV [78]. Since in pPb $v_n\{k\}$ can become imaginary we here use $\tilde{v}_n\{k\} \equiv \mathrm{sgn}(v_n\{k\}^k)|v_n\{k\}|$, which equals $v_n\{k\}$ when it is real. For pPb we furthermore compare to the mean p_T for pions, kaons, and protons at 5.02 TeV [79]. Here we typically use the 0–5, 5–10, 10–20, up to 50–60% centrality classes, except for the pPb flow coefficients, where high multiplicity events are specifically interesting. In that case we follow [25] and use five multiplicity classes, separated according to charged particle multiplicity over the mean particle multiplicity. With the bias introduced in Sec. II H this was possible with reasonable statistics up to the bin with multiplicity four to five times the average particle multiplicity. In total this leads to 418 and 96 datapoints for PbPb and pPb collisions, respectively.

To compute the actual posterior distributions we use a Markov chain Monte Carlo, which produces a chain of parameters with a distribution that converges to the true distribution. For this we use the EMCEE code [80], with a chain of 600 walkers initialized using a uniform distribution and subsequently
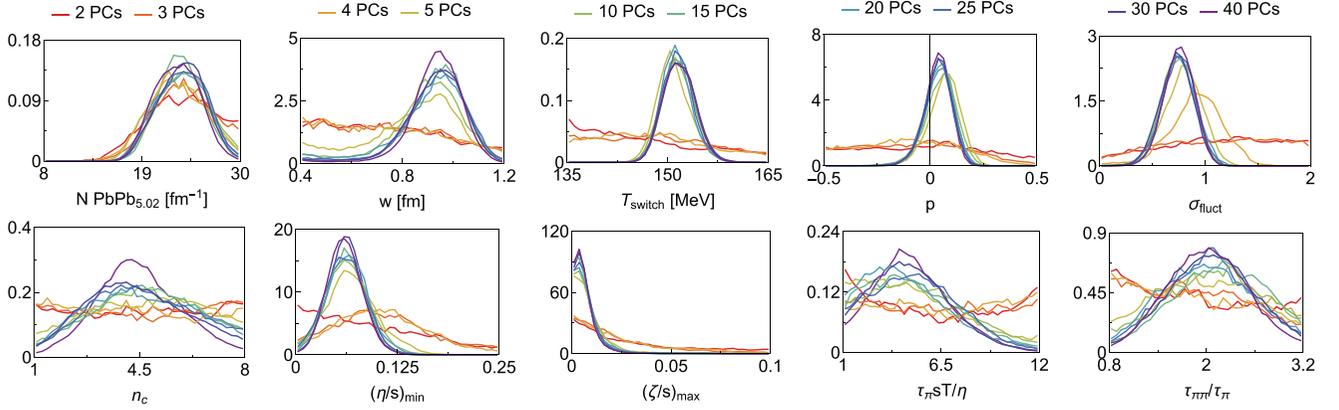
FIG. 11. For ten representative posteriors we show the influence of including a varying number of principal components (PCs) for the posteriors of PbPb at 2.76 and 5.02 TeV. Some parameters converge more smoothly than others, but from ten PCs onward the precision gains are modest.

evolved for 15 000 steps. It is important to run the computation for a sufficient number of steps to allow the distribution to converge to the true distribution. Figure 10 shows the mean and median over all walkers for all parameters as a function of step number. The 68 and 95% confidence intervals are shaded, whereby we note that we take these intervals symmetrically with for the second region 2.5% probability left on either side of the distribution (see [24] for an alternative definition that minimizes the width of the distribution). For the distribution to have converged it is necessary that not only the mean but also the entire distribution has converged, apart from statistical fluctuations. As can be seen in Fig. 10, some parameters satisfy this condition faster than others. For example, σ_{fluct} converges quite fast, while w only ends its downward trend after around 9000 steps, so that for all analyses only the part of the chain after this point should be used.

As another test on the convergence of the chains we varied the number of PCs in Fig. 11, shown for PbPb at 2.76 and 5.02 TeV only. For some parameters the posterior changes quite suddenly, for instance the switching temperature is clearly most constrained by the fifth PC, after which not much precision is gained anymore. Throughout this paper we chose to show results using 25 PCs.

D. Closure tests

For any model including emulators and Bayesian estimates it is crucial to implement a closure test: given a set of parameters one determines the model output ("experimental data") and from this output one obtains the posterior estimates on the parameters, which naturally should be consistent with the input parameters (see also [81]). This test is useful for two reasons: first it gives a strong test that all steps from emulation to the MCMC are working correctly, and importantly that the uncertainties obtained are realistic. Second, since the closure test uses the same dataset as the "real" model it shows how constraining this dataset is on each of the varying parameters. It could for instance be that the model is not sensitive to one of the input parameters, and the closure test should then result in a flat distribution regardless of the input parameter. This latter effect can also depend on where we are in the parameter

space; on physical grounds it is for instance clear that the width of the bulk viscosity is unconstrained if the norm of the bulk viscosity vanishes. It is therefore important to perform the closure test at several points in the parameter space.

Figure 12 shows the posterior distribution for six such points; the dashed lines indicate the (random) input parameters for each closure test using different colors. Note that both the norm as well as the nucleon-nucleon cross section are chosen randomly for each system separately, whereby we did not constrain the norm to be larger for 5.02 TeV as compared to 2.76-TeV collisions. For reference we include the real posterior distributions as obtained in [28] (black). A few parameters are clearly well constrained, including the norms, the width, the switching temperature, the TRENTO parameter p, the amount of fluctuations, the free streaming velocity as well as the minimum of η/s, and the norm of ζ/s. Other parameter distributions are more similar to the priors (being flat over the range), or depend more sensitively on the point in parameter space. $\tau_\pi sT/\eta$ is for instance quite well constrained for points 1, 2, and 5, but less so for the others.

Most posterior distributions align well with the input parameter. This can be quantified by looking at the percentile in the posterior distribution of the input parameter, of which a histogram is shown for all 6×24 parameters in Fig. 13. Clearly the distribution of the percentiles is approximately flat, as expected. There is a small excess around 100%, which can indicate that we should not have put (random) input parameters close to the edge of the parameter range. In that case the prior probabilities, which are zero outside the parameter range, guarantee a percentile that is close to zero or 100%. For the real dataset this issue is not present, as all prior probabilities have been chosen to include a parameter range that should reasonably include the physical point. Lastly, in rare cases the posterior is clearly off, such as for instance for v_{fs} for point 4. In this case there is a strong correlation with χ_{struct} (see also [28]), which is also off, thereby at least providing part of the explanation.

A more detailed and refined look can be obtained by showing the average posterior deviations of the actual 514 datapoints for each closure point, as shown in blue in Fig. 14 as a ratio with respect to the uncertainty that includes both
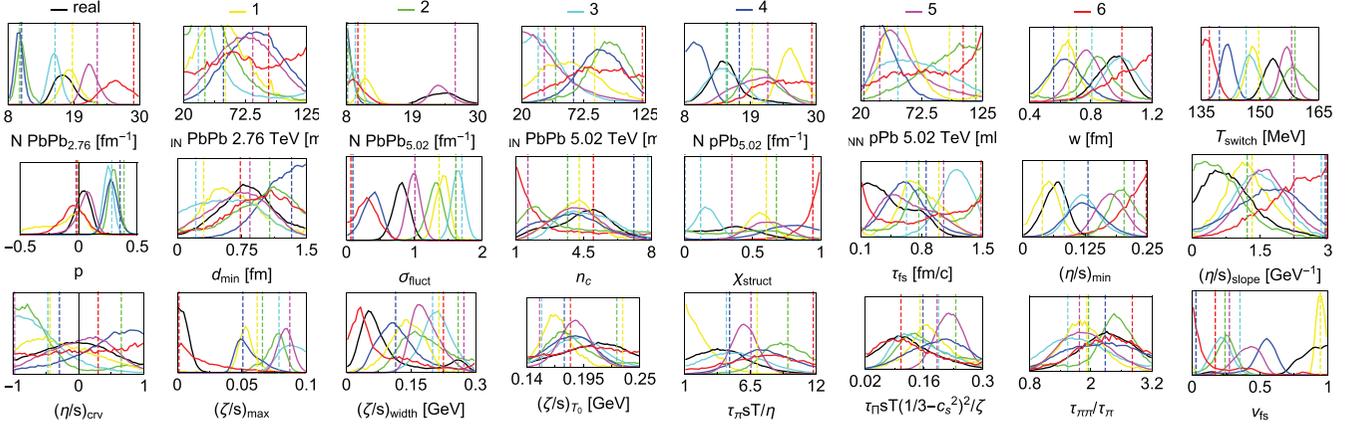
FIG. 12. For all parameters that we vary we show the posterior distributions obtained from model output at six random locations in parameter space (shown as dashed lines). For comparison we also show the posterior comparing to real data from [28] (black). The posterior distributions match well with the input parameter, and for many parameters strong constraints are possible using the experimental data as presented in Sec. III C.

emulator uncertainty as well as statistical uncertainty in computing the "data" point. Systems are numbered 1, 2, and 3 for PbPb at 2.76 and 5.02 and pPb at 5.02 TeV, respectively. The red points show the deviation when using the emulator on the input parameters directly, without using the posterior distribution. Ideally all these points would agree within their normal errors, whereby the red points test solely the emulator and the blue points also test the posterior distributions. In the text we also display the (average) values of the deviation computed by (9) as well as the LL of (8). The red validation is supposed to give the best estimates of the datapoints and naively should have the highest LL. Nevertheless, the posterior distributions are optimized to maximize this likelihood, so it is no surprise that these often give a higher LL. Nevertheless, their relative closeness gives further credibility that all works well. Lastly, it is worthwhile to study a few outliers more specifically. Point 2 has a four (blue) or six (red) standard deviation difference

for a pion v_2 in system 2. Looking more carefully, this is a high p_T off-central datapoint whereby the norm for this closure point is furthermore small (8.21 in this case). This discrepancy is hence a result of a lack of statistics for this particular somewhat unlucky combination. Point 5 has a rather strong deviation in red for the centrality dependence in system 1 already at the emulator level (as seen from red points becoming consistently worse for $dN_{ch}/d\eta$ and $dE_T/d\eta$), which would be interesting to investigate further.

Finally, we can compare to the equivalent result using the real experimental data, as shown in Fig. 15. Here we do not have knowledge of the true physical parameters, so there are no validation points. Instead we do show the part of σ_y due to experimental error (in black). We see that the mean of all posterior distributions nicely follows a normal distribution with no clear outliers. The pPb results have larger deviations, especially when compared to the experimental error. This is because it is much harder to simulate this smaller system, as already noticed in Fig. 4.

E. Discretization error

An important source of potential error is the discretization error in the simulation. Our simulations are performed on a square lattice with lengths of 50 fm, containing $N_{sites} \times N_{sites}$ lattice sites. A straightforward way to examine the effect of discretization error is to perform the hydrodynamic simulation on the same event with multiple values for N_{sites} and comparing the results. We performed this check for $N_{sites} = 200, 300, 400$, and as shown in Fig. 16 we find no significant differences between the results.

In addition to this, we also employed a novel technique, by which we vary N_{sites} in the Bayesian analysis as if it were a physical parameter. In particular, one looks for two important indications that the error is indeed under control, i.e., that the smallest value for N_{sites} in the chosen range is large enough. First, it is possible to utilize the emulator to verify that our observables do not depend on N_{sites}. This is shown for a few representative observables in Fig. 17 (top left, top right,
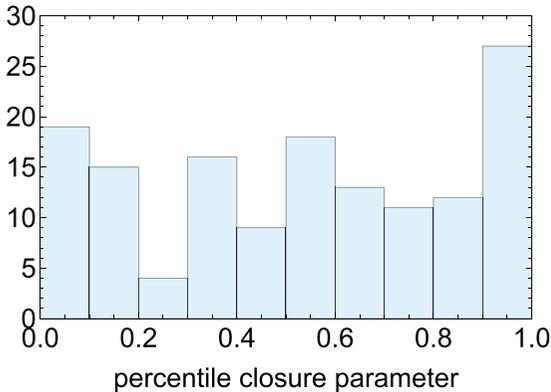
FIG. 13. We show the distribution of percentiles of the input parameter in the posterior probability distribution for all 6×24 parameters displayed in Fig. 12. Ideally this should be a flat distribution, with each bin containing 14.4 ± 3.8 points. The highest bin is overrepresented, which is partly due to our choice of random input parameters that can lie at the edge of the prior probability distribution.
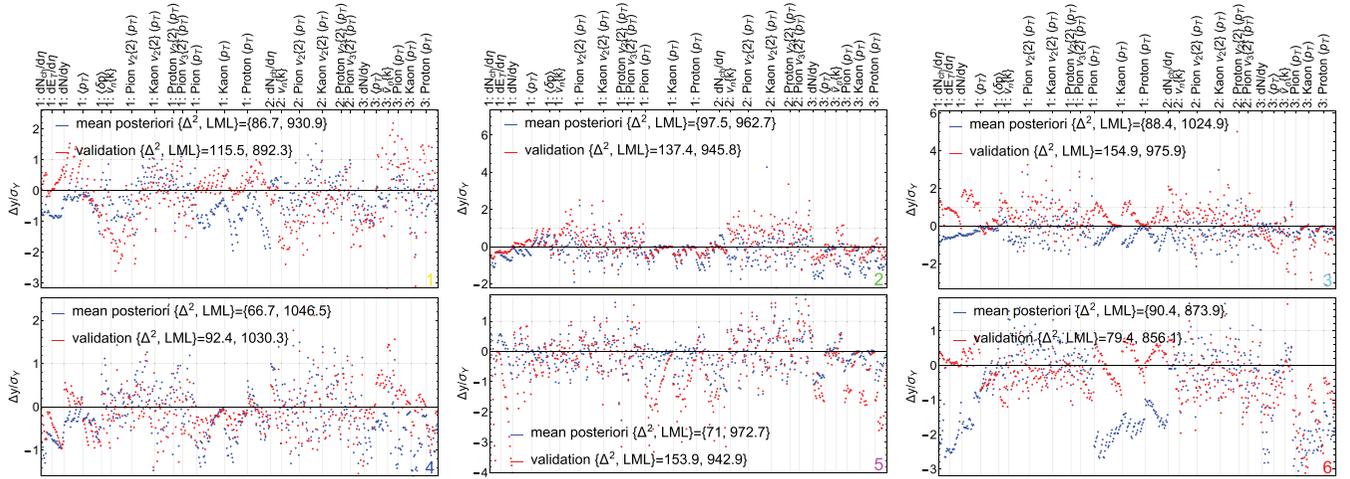
FIG. 14. For the closure points presented in Fig. 12 we present the average of 100 posterior datasets, as compared to the true datapoint with the estimated uncertainty σ_y (including both emulator uncertainty and statistical uncertainty from obtaining the datapoint in the model). The validation dataset is obtained from the emulator prediction at the input parameters. Data are labeled by class, where the number refers to the system (PbPb at 2.76 and 5.02 TeV and pPb at 5.02 TeV, respectively) and subclasses or centralities are not displayed. In total we included 514 datapoints.

bottom left). All observables shown are indeed consistent with being independent of N_{sites} when taking into account the uncertainty associated with the emulator (shown as a band). $\tilde{v}_2\{2\}$ for pPb shows the strongest dependence, whereby the relatively wide uncertainty band is due to both the statistical difficulty obtaining accurate results for (low multiplicity) pPb flow coefficients, as well as the more difficult emulation (see also Sec. III A). Second, the posterior distribution treating N_{sites} as a physical parameter should not show any preference or correlation, which is implied if indeed observables do not depend on N_{sites}. Figure 17 (bottom right) shows the result of the six closure tests that we performed for N_{sites}, together with the posterior from the comparison with real data. One can see that while none of the distributions are completely flat, the
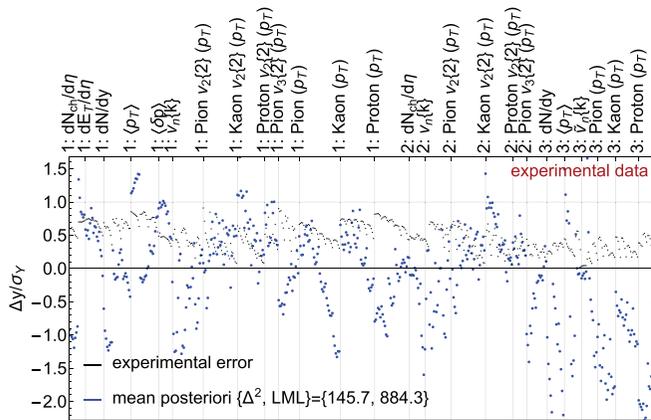
FIG. 15. We show the equivalent plot of Fig. 14 for the real experimental data. Black points represent the part of σ_y due to the experimental error, where for pPb collisions (labeled 3) the emulator error is clearly dominant. All posterior averages fit our 514 experimental datapoints nicely according to a normal distribution.

locations of the maxima do not correlate with the known input values used in the closure tests. This means that the nonflat distributions are unlikely to reflect any true correlations found through the posterior estimates.

We also checked the correlations of all 24 parameters with N_{sites} for the six closure tests and the real posterior distributions. For all parameters the average correlation never exceeded 10%, from which we conclude that also the correlations with N_{sites} are insignificant. Lastly, we note that this technique could also be used to estimate a minimum number of required grid points, or point to observables and parameter settings that need particularly accurate simulations.

Another source of discretization error can come from the fact that τ_π or τ_Π can become of the same order as the time step Δt. For the simulations used in this paper and in [28], $\Delta t = 0.08 \Delta x$, where Δx varies between 0.17 and 0.25 fm (for pPb more refined simulations are necessary, whereby we used a lattice of 12 fm in size with at least 70 lattice sites such that $0.11 < \Delta x < 0.17$ fm). This leads to $\Delta t \leqslant 0.02$ fm/c for our simulations. For this reason we require that $\tau_\pi > 2\Delta t$ and $\tau_\Pi > 0.1$ fm/c throughout the simulation. Due to the temperature dependence in both η/s and ζ/s, whether this rule is invoked often depends in a relatively complicated way on the parameters.

We verified that in less than 0.2% of our posterior distribution we needed to invoke this requirement for τ_π at any temperature below 400 MeV and hence we can trust the resulting distribution not to be affected by the minimum time step. For τ_Π this is however different, as our analysis keeps the ratio $\tau_\Pi s T (1/3 - c_s^2)^2/\zeta$ fixed. As our posterior distribution obtains a small specific bulk viscosity ζ/s this implies that often τ_Π is small as well. We hence verified that in most of our simulations the constraint $\tau_\Pi > 0.1$ fm/c is dominant. This means we should not put much trust in our posterior distributions for τ_Π, which in our case did not give strong constraints anyway.
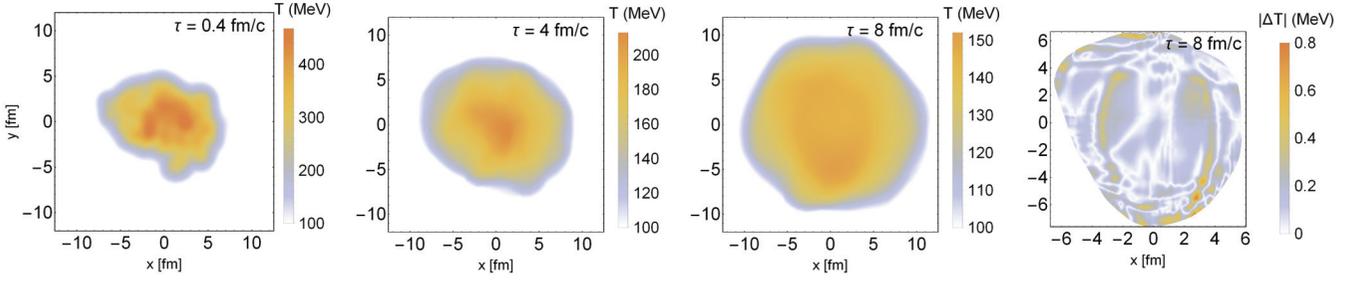
FIG. 16. We show the temperature of a typical event at the initial time $\tau = 0.4 \, \text{fm}/c$ and at 4 and 8 fm/c, together with the difference in temperature at the end of the evolution comparing resolutions of $N_\text{sites} = 200$ or 300. The difference is shown only where there is plasma (defined by $T > 140 \, \text{MeV}$) and shows that the discretization error is small.

F. Prehydrodynamic phase

This section attempts a more refined understanding of the prehydrodynamic phase and more in particular the influence of the free streaming velocity v_fs. First of all we note that the free streaming velocity directly affects the pressure of the free streaming stress tensor (2), whereby $v_\text{fs} = 0$ leads to a pressureless fluid and $v_\text{fs} = 1$ corresponds to a conformal stress tensor for which the EOS equals $P = \rho/3$. This pressure is in turn matched to the viscous shear tensor $\pi^{\mu\nu}$ and the bulk viscous pressure Π in order to guarantee a continuous stress tensor. Second, it is important to realize that the EOS at the relevant temperatures is not yet fully conformal, and in fact at $T = 0.4 \, \text{GeV}$ we have $P \approx 0.85\rho/3$ [54]. From the equation of state one may therefore suspect that physically a free streaming velocity around 0.85 may be the most realistic.

To quantify this effect we are interested in the effects of changing the free streaming velocity v_fs on the initialization

FIG. 17. Top left: $v_2\{2\}$ (1.0–1.4 GeV) for PbPb at 2.76 TeV, shown for pions, kaons, and protons in the 20–30% centrality class. Top right: Charged particle multiplicity, normalized by the experimental data, for PbPb at 2.76 TeV, shown for pions, kaons, and protons in the same centrality class. Bottom left: $\tilde{v}_2\{2\}$ for pPb at 5.02 TeV, in the multiplicity classes $N_\text{ch}/\langle N_\text{ch}\rangle \in [0, 1]$, [2, 3], and [4,5]. Bottom right: Posterior distribution for N_sites, shown together with six posterior distributions for the closure chains. Colors are as in Fig. 12.

of the shear stress $\pi^{\mu\nu}$ and the bulk viscous pressure Π. These are shown in, respectively, the left and right panel of Fig. 18, in a cross section through the plasma, relative to the energy density ρ. Since the shear tensor is traceless, we define the absolute value of the shear tensor as the simplest nonvanishing scalar:

$$|\pi| \equiv \sqrt{g_{\mu\nu}g_{\rho\sigma}\pi^{\mu\rho}\pi^{\nu\sigma}}. \tag{10}$$

The event shown is generated using the MAP values for the parameters presented in [28], which in particular means that these figures are shown at $\tau_\text{fs} = 0.47 \, \text{fm}/c$. Furthermore, this one event is computed multiple times, each time using a different value for v_fs. In addition, the "quasiequilibrium" values of $|\pi|/\rho$ and Π/ρ are also shown. These are obtained by, respectively, setting the left hand sides of (7) and (6) to zero. In other words, these are the values that the shear stress and bulk viscous pressure would take if the fluid was initialized close to the value that (7) and (6) would relax to given enough time.

In the left panel of Fig. 18, one can see that the value closest to the quasiequilibrium value for the shear stress is $v_\text{fs} \approx 0.6$. However, because $\pi^{\mu\nu}$ is a tensor it is not clear that these configurations resemble each other. To investigate this further we furthermore computed the absolute value of the difference, defined analogously as $|\Delta\pi| \equiv \sqrt{g_{\mu\nu}g_{\rho\sigma}\Delta\pi^{\mu\rho}\Delta\pi^{\nu\sigma}}$, with $\Delta\pi \equiv \pi^{\mu\nu}_\text{fs} - \pi^{\mu\nu}_\text{quasi-eq}$. This shows that $v_\text{fs} \approx 0.6$ indeed minimizes the deviation from the quasiequilibrium value, whereby $|\Delta\pi|/\rho \approx 0.05$.

In the right panel of Fig. 18, one can see that the quasiequilibrium value for the bulk viscous pressure is slightly negative, in agreement with the small value found for the bulk viscosity ζ. However, when initializing the plasma with $v_\text{fs} = 1$, the bulk viscous pressure is much larger and positive, as is indeed expected for an EOS that has a pressure below the conformal value. As one lowers v_fs, the bulk viscous pressure decreases, and eventually changes sign around $v_\text{fs} \approx 0.93$. The posterior distribution for v_fs is peaked around $v_\text{fs} \approx 0.9$, suggesting that data perhaps show a preference for initialization near quasiequilibrium. Such an initialization is also what is suggested by holography, where the system quickly *hydrodynamizes* after the initial stage [48].

The fact that the v_fs for a quasiequilibrium initialization of $\pi^{\mu\nu}$ and Π are different means that within the free streaming model it is impossible to initialize both the shear stress and bulk pressure near quasiequilibrium. This implies that while

FIG. 18. Left: Ratio of the absolute shear viscous pressure $|\pi|$ and the energy density ρ, where we define $|\pi| = \sqrt{g_{\mu\nu}g_{\rho\sigma}\pi^{\mu\rho}\pi^{\nu\sigma}}$. Right: Ratio of the bulk viscous pressure Π and the energy density ρ, computed at the time of initialization in a cross section through the plasma. The free streaming has been performed on the same event with three different free streaming velocities (different colors). We also show the "quasiequilibrium" value, defined as the zeros of the right hand sides of (7) and (6).

the stress-energy tensor itself is continuous at the switch to hydrodynamics, its derivative will not be continuous. This is in marked contrast with, e.g., holographic models [48], which would hence be interesting to study further in future work.

The behavior of Π as a function of v_{fs} can be understood as follows: For $v_{\mathrm{fs}} = 0$, the stress tensor has the form $T^{\mu\nu} = \rho\delta_0^\mu\delta_0^\nu$. Using the constitutive relation (5), one can decompose this stress tensor, yielding $\pi^{\mu\nu} = 0$ and $\Pi = -P$. This therefore leads to Π being negative. At early times we can use (2), which for $v_{\mathrm{fs}} = 1$ yields a traceless stress tensor, i.e., $\rho - 3(P + \Pi) = 0$. For our equation of state the speed of sound is below the conformal bound, which implies $3P < \rho$ and hence that Π is positive. For intermediate values of v_{fs}, the behavior of Π interpolates between these two cases.

Another interesting dependence of the prehydrodynamic phase is that on the free streaming time τ_{fs}. In Fig. 19, the deviation from quasiequilibrium is shown for the shear stress (left) and the bulk viscous pressure (right). As in Fig. 18, the same event is computed, but here we used $v_{\mathrm{fs}} = 1$ together with

various values of τ_{fs}. It can clearly be seen that the shear tensor moves away from quasiequilibrium during the prehydrodynamic phase, even though the τ_{fs} dependence is mild. For the bulk viscous pressure, the interior of the plasma moves away from quasiequilibrium during the prehydrodynamic phase, whereas the edges move towards quasiequilibrium. These edges contain only a small part of the total energy though, so we can conclude that the prehydrodynamic stage moves the fluid away from quasiequilibrium.

G. Correlation between $\tau_{\pi\pi}/\tau_\pi$ and $(\eta/s)_{\mathrm{min}}$

As shown in [28], $\tau_{\pi\pi}/\tau_\pi$ and $(\eta/s)_{\mathrm{min}}$ are negatively correlated. The reason for this is that η/s is mostly determined by the measurement of $v_2\{2\}$, and that both η/s and $\tau_{\pi\pi}$ tend to lower this observable, making the effects of these two transport coefficients more or less interchangeable. The mechanism ultimately causing this is that both of these transport coefficients are dissipative, and hence increase the entropy by making the fluid less anisotropic. For this reason, we expect
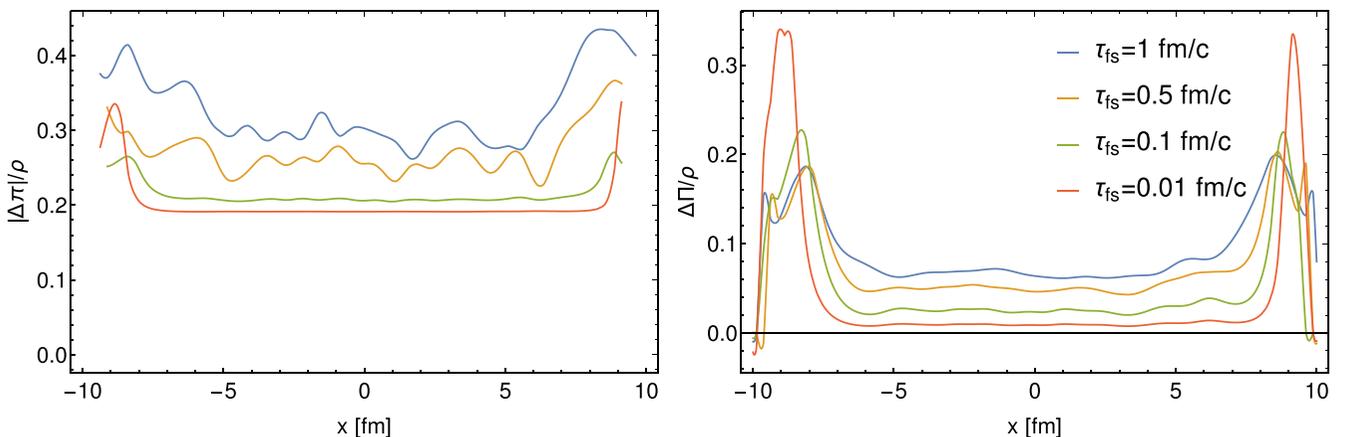
FIG. 19. Difference from quasiequilibrium relative to the energy density ρ for different values of the free streaming time τ_{fs}, shown for the shear stress (left) and the bulk viscous pressure (right). Here we define $|\Delta\pi| \equiv \sqrt{g_{\mu\nu}g_{\rho\sigma}\Delta\pi^{\mu\rho}\Delta\pi^{\nu\sigma}}$, with $\Delta\pi^{\mu\nu} \equiv \pi_{\mathrm{fs}}^{\mu\nu} - \pi_{\mathrm{quasi\text{-}eq}}^{\mu\nu}$ and $\Delta\Pi = \Pi_{\mathrm{fs}} - \Pi_{\mathrm{quasi\text{-}eq}}$.
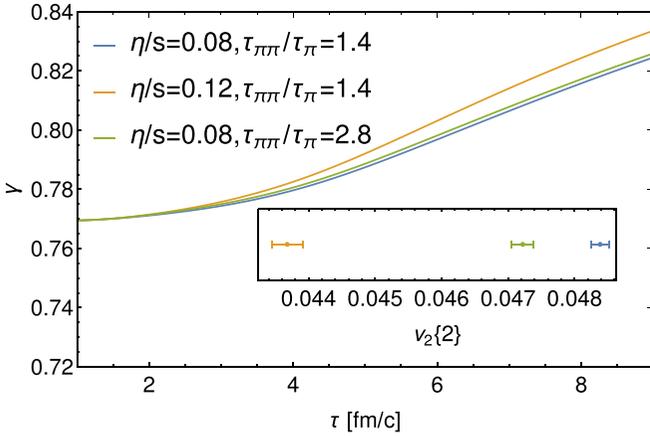
FIG. 20. γ, defined by (12), as a function of proper time τ for the idealized plasma initialized according to (11), for different values of η/s and $\tau_{\pi\pi}/\tau_\pi$. Inset: The resulting values for the elliptic flow $v_2\{2\}$.

this correlation to be generically present, but in the following we will look at a specific configuration, to examine whether the mechanism by which the anisotropy decreases is similar in both cases.

In the case of the shear viscosity, the decrease in anisotropy is well understood to work through the mechanism that in the "short" direction pressure gradients are larger, driving the buildup of momentum in that direction, while the shear viscosity counters this momentum buildup [82]. In a similar spirit, we look at an idealized plasma, where we initialize $\pi^{\mu\nu}$ and Π to zero, set $u^\mu = (1, 0, 0, 0)$, and take the energy density to be

$$\rho(x, y) = \frac{\alpha}{1 + \exp\left(\frac{\sqrt{x^2+(1.3y)^2}-R}{\theta}\right)}, \tag{11}$$

with $R = 5$ fm, $\theta = 1$ fm, and $\alpha = 50$ fm^{-4}. Initializing the plasma as such at $\tau = 0.98$ fm/c, we let it evolve with η/s independent of temperature, and the bulk viscosity equal to that found in [24]. We then define the quantity γ as follows:

$$\gamma(\tau) = \frac{\int_0^\infty dx \int_0^\infty dy\, T^{01}(x, y, \tau)}{\int_0^\infty dx \int_0^\infty dy\, T^{02}(x, y, \tau)}, \tag{12}$$

which measures the ratio of the total x momentum and y momentum present in one quadrant. In an isotropic setting γ equals unity, and by using the departure from unity we can analyze how the dynamics of the fluid convert the spatial anisotropy into an anisotropic momentum distribution. In Fig. 20, the result of this analysis is shown for three different choices of η/s and $\tau_{\pi\pi}/\tau_\pi$ (note that increasing η/s by 50% also increases $\tau_{\pi\pi}$ by 50% since we fix our ratios such that $\tau_\pi \propto \eta/s$). One can clearly see that increasing η/s decreases the amount of anisotropy, thereby decreasing $v_2\{2\}$. One can also see that indeed $\tau_{\pi\pi}/\tau_\pi$ has a similar, albeit smaller, effect. Hence we can indeed conclude that both of these dissipative corrections tend to lower the amount of momentum anisotropy.

In the inset of Fig. 20, we see the associated $v_2\{2\}$ of these three simulations. It can clearly be seen that the observed

momentum anisotropy translates into $v_2\{2\}$, explaining the observed correlation.

IV. MAXIMUM *A POSTERIORI*

With the emulator and MCMC together with 514 datapoints we obtained optimal (MAP) values for our varying parameters [28]. Strictly speaking these values are just the mean expectation value for each parameter and nontrivial correlations could mean that such a combination is not maximizing the LL of (8). We verified, however, that this method led to a LL that is comparable with the maximum LL obtained otherwise. In this section we analyze a high statistics run (400 000 and 4×10^6 events for PbPb and pPb, respectively) using these parameters, which allows us both to verify the emulator at this point specifically and importantly to compare to experimental data that were statistically not feasible to include in the Bayesian optimization.

A. Transverse momentum spectra

In Fig. 21 we show the transverse momentum spectra for central and very peripheral PbPb collisions as well as central pPb collisions including ALICE data [76,83] for the entire range where particle identification is possible. These spectra have been included in the posterior estimate using p_T bins separated at $(0.5, 0.75, 1.0, 1.4, 1.8, 2.2, 3.0)$ GeV for centralities up to the 40–50% class. At high p_T the spectra are dominated by hard processes described by perturbative QCD, which have a characteristic polynomial falloff as a function of p_T. Our hydrodynamic thermal model falls off exponentially due to the Boltzmann factor with the switching temperature, and hence it is expected that around $p_T \approx 3$ GeV the hydrodynamic prediction starts to deviate from the full experimental result. This effect is more pronounced for very peripheral collisions as well as pPb collisions. At low p_T below 500 MeV (not included in the posterior) there is a small surplus of pions, especially for very peripheral collisions where this deviation exceeds two standard deviations (see also [22] for a similar result).

From the spectra it is possible to obtain the average transverse momentum, as shown in Fig. 22. Overall this provides an excellent fit, but the experimental mean transverse momentum for kaons in peripheral pPb collisions is surprisingly high (we show the neutral kaons here, as their experimental uncertainty is much smaller than that of the charged kaons [76]). The hydrodynamic results seem consistent with [25] and given the significantly lower mean kaon momentum in PbPb collisions it is possible that this requires a nonhydrodynamic explanation, which is also apparent from the deviation of the kaon spectrum at intermediate p_T in Fig. 21 (bottom). Also the spectra of pions and protons deviate from the experimental data at intermediate p_T, but this is compensated at lower p_T, which then leads to the correct mean transverse momentum. This stresses once more why it is important to include the full identified transverse spectra. Since the average kaon transverse momentum has perhaps the most apparent deviation we chose this observable when showing its dependence on our parameters in Fig. 30 in the Appendix.
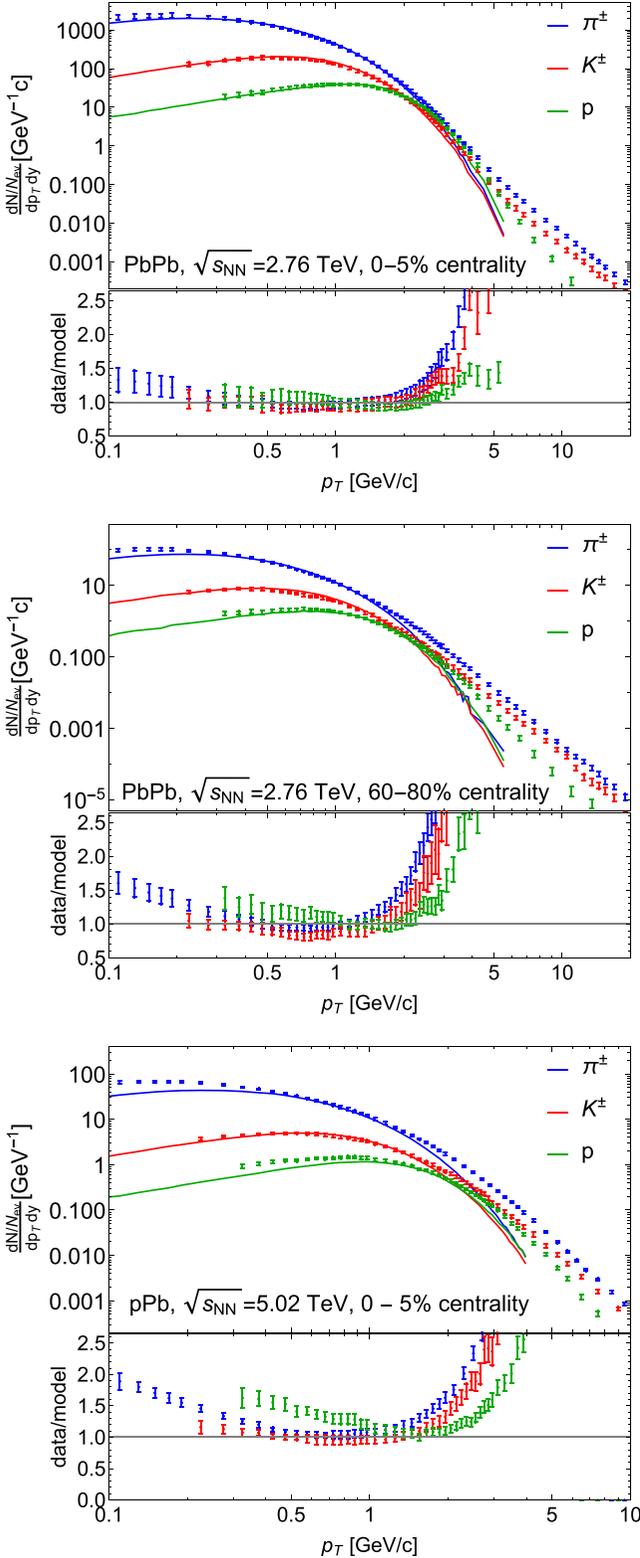
FIG. 21. We compare the MAP results for the transverse momentum spectra for charged pions, kaons, and protons for central (top), peripheral (middle, data from ALICE [83]), and central pPb (bottom, data in [76]) collisions. The model somewhat underpredicts the low p_T pions (also seen in [22]), as well as the spectra at high p_T. The latter is more pronounced in peripheral and pPb collisions and can possibly be explained by polynomial processes in pQCD.
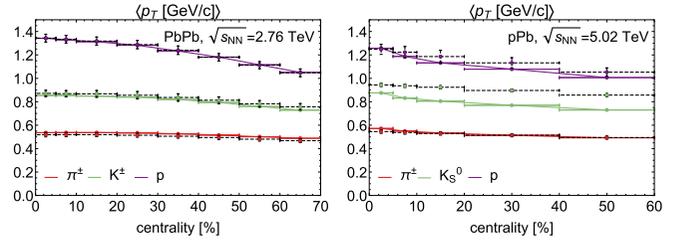
FIG. 22. We compare the mean of the transverse momentum for pions, kaons, and protons as a function of centrality (colored) to experimental data from ALICE (black, [76]). The hydrodynamic results match perfectly, with the possible interesting exception of the mean kaon transverse momentum in peripheral pPb collisions that possibly requires a nonhydrodynamic explanation.

We also stress that the MAP results are obtained using the expectation values found in [28], but importantly they do not include the posterior uncertainties. Especially for pPb this is relevant, as the emulator has a significant uncertainty, and indeed the posterior mean kaon transverse momentum has a wide band that may just about include the experimental result (see Fig. 1 in [28]). It could hence be that certain parameters especially relevant for pPb can still be found that provide a better fit to the kaon mean p_T. One likely option here would be to have a larger free streaming velocity $v_{\rm fs}$ (see Fig. 30).

Lastly, we show the transverse momentum fluctuations in Fig. 23, defined as $\delta p_T^2 = \langle\langle (p_{T,i} - \langle p_T \rangle)(p_{T,j} - \langle p_T \rangle) \rangle\rangle$, where the double brackets denote an average over all pairs in the same event as well as averaging over the centrality class. Up to 60% in centrality this observable is also used to obtain the posteriors (as also done in [24]), but even then it turns out to be difficult to capture the peripheral bins (see [24]). It is therefore satisfactory that we manage to obtain a good agreement, with perhaps the exceptions of the most central and most peripheral bins.

Transverse momentum fluctuations have not been measured for pPb collisions and it is therefore interesting to make a prediction, which is shown as red points in Fig. 23 (see also [27] for a similar prediction). The fluctuations do not depend strongly on centrality, but a relatively strong decrease is seen for ultraperipheral collisions (see inset). Remarkably we see a similar decrease for very peripheral collisions.

B. Strangeness enhancement

The fraction of strange quarks in the final particle spectra is enhanced as one goes to higher and higher multiplicities [84], which was originally proposed as a signature of QGP formation. The naive explanation for this is that in nucleon-nucleon collisions the fraction of strange quarks is canonically suppressed as compared to the fraction that would exist in thermal equilibrium. In Fig. 24 we show the ratio of multiplicities of strange hadrons with respect to pions versus event activity including experimental data for both pp and pPb collisions [84,85]. The theoretical curves contain results from our pPb (solid) and PbPb (dashed) collisions. To stress the importance of resonance decays as well as hadronic interactions
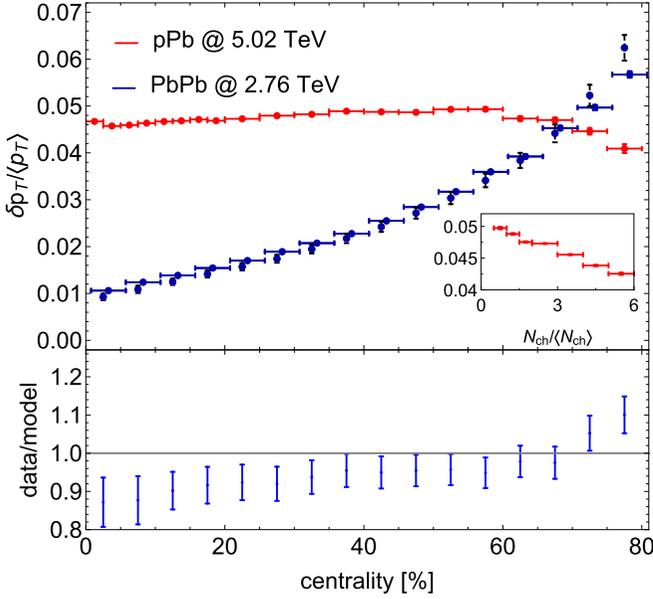
FIG. 23. We compare the transverse momentum fluctuations of charged particles as a function of centrality (blue) to experimental data from ALICE (black, [75]). Especially for peripheral collisions this is a difficult observable [24] and even though we did not use the centrality bins beyond 60% to obtain the MAP values there is still a satisfactory match with the model. The red points are predictions for pPb collisions at $\sqrt{s_{NN}} = 5.02\,$TeV for the same $0.15 < p_T[\mathrm{GeV}/c] < 2.0$ range. The inset shows the pPb fluctuations for ultracentral events with multiplicities up to six times the average multiplicity.

we also include the corresponding results without using the SMASH afterburner as dotted lines.

In general the hydrodynamic results align well with the high-multiplicity pp and pPb results, indeed confirming that
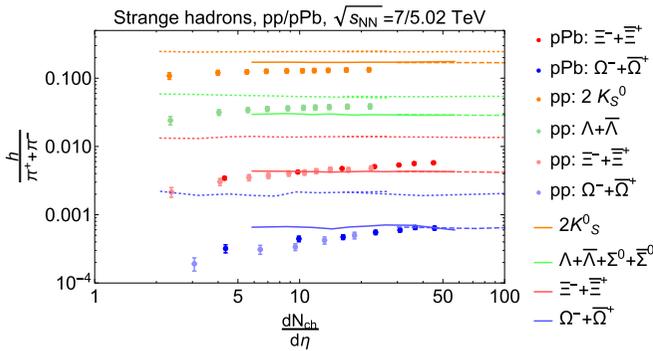
FIG. 24. We show the ratio of hadrons containing strange quarks with respect to the pion yield as a function of multiplicity for pp (experiment only [84]), pPb [experiment [85] and model (solid)], and PbPb collisions (model only, dashed). To show the importance of the resonances and the afterburner we include results without applying the hadronic afterburner (dotted). Experimental strangeness increases with multiplicity, while the theoretical curves do not depend on multiplicity. Small but significant deviations exist even for high multiplicity events for both kaons and lambda hadrons.

a thermal model like ours can explain the saturation of the strangeness fraction. There is no significant multiplicity dependence in our hydrodynamic model. There is also a small but significant discrepancy even at high multiplicities: the kaons are overpredicted and the lambdas are underpredicted. In [86,87] fits have been performed in the canonical ensemble, which can become important as the system size becomes smaller; see also [88] for thermal fits performed by the ALICE collaboration.

C. Anisotropic flow

Anisotropic flow coefficients are an interesting family of observables that encode the anisotropy of the particle distributions in the final state. Examples are integrated ($v_n\{k\}$) over a certain p_T and η range, and p_T differential [$v_n\{k\}(p_T)$]. The integrated flow coefficients are computed by first computing the Q_n vector for each event [89,90]:

$$Q_n = \sum_{i=1}^{M} e^{in\phi_i}, \qquad (13)$$

where ϕ_i is the azimuthal angle of particle i and the sum runs over a certain set of M particles. Each different set of chosen particles defines a different observable. For example, one can choose to use only charged particles, or to also include neutral particles. Also, one can choose to incorporate only particles of a certain species, leading to identified flow, or to include all species, leading to unidentified flow. In this paper, we chose to compute unidentified charged particle flow coefficients. For this, we used all charged particles with $0.2 \leqslant p_T \leqslant 5\,$GeV and $|\eta| \leqslant 0.8$ in accordance with [74]. After having computed the Q_n vector, one can compute the two- and four-particle correlations for that particular event as

$$\langle 2 \rangle_n = \frac{|Q_n|^2 - M}{M(M-1)},$$

$$\langle 4 \rangle_n = \frac{|Q_n|^4 + |Q_{2n}|^2 - 2\mathrm{Re}[Q_{2n}Q_n^*Q_n^*]}{M(M-1)(M-2)(M-3)}$$
$$- 2\frac{2(M-2)|Q_n|^2 - M(M-3)}{M(M-1)(M-2)(M-3)}.$$

These two- and four-particle correlations are then averaged over all events in a given centrality class, with per-event weights given by $M(M-1)$ for the two-particle correlations, and $M(M-1)(M-2)(M-3)$ for the four-particle correlations. This yields the averaged two- and four-particle correlations $\langle\langle 2 \rangle\rangle_n$ and $\langle\langle 4 \rangle\rangle_n$. The integrated flow coefficients are then given by

$$v_n\{2\} = \sqrt{\langle\langle 2 \rangle\rangle_n},$$

$$v_n\{4\} = \sqrt[4]{2\langle\langle 2 \rangle\rangle_n^2 - \langle\langle 4 \rangle\rangle_n}.$$

The differential flow is computed in a similar way, but now we define two groups of particles for each observable: the reference flow particles (RFPs) and the particles of interest (POI). In this way, one can reconstruct the p_T dependence of the flow. First, one defines the Q_n vector in the same way as for the integrated flow, where the sum now runs over all RFPs.
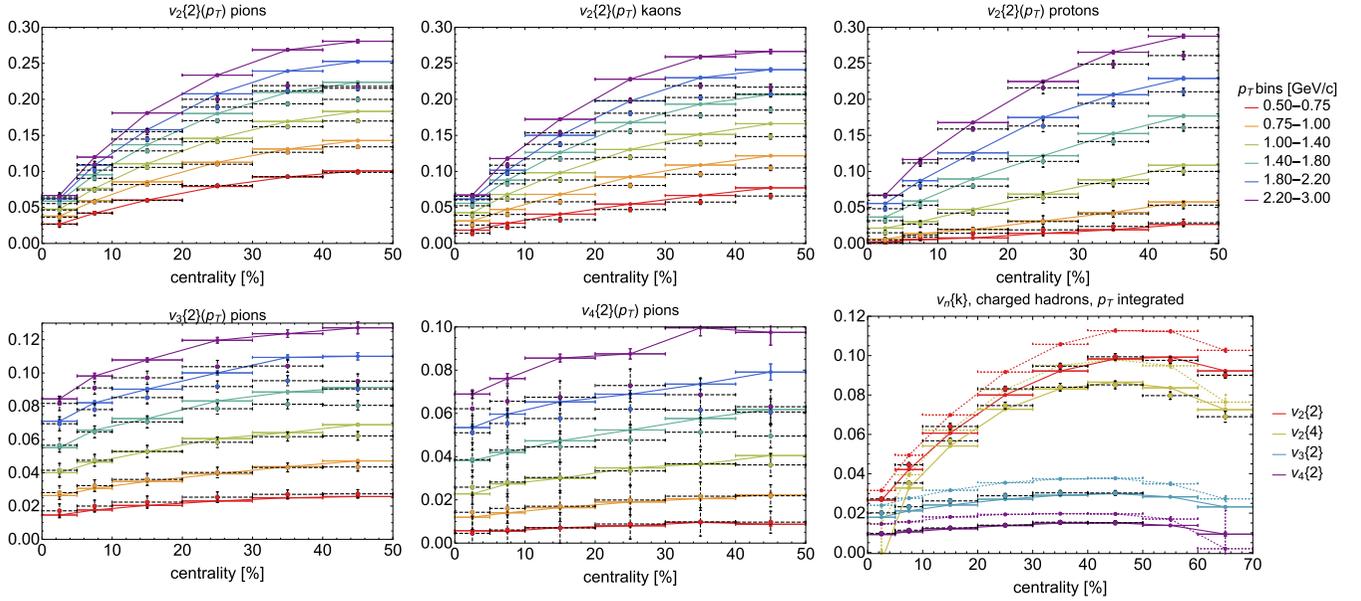
FIG. 25. We show the anisotropic flow coefficients for PbPb collisions at 2.76 TeV using our MAP values. The top row shows for six p_T bins $v_2\{2\}$ for pions (left), kaons (middle), and protons (right) as a function of centrality. On the bottom row, the left (middle) panel shows $v_3\{2\}$ ($v_4\{2\}$) for pions, and the rightmost panel shows integrated flow for $v_2\{2\}$, $v_2\{4\}$, $v_3\{2\}$, and $v_4\{2\}$. The solid colored lines indicate the full model result, and the dotted lines are the result without using the afterburner. The black datapoints are from [68,74].

In addition, one defines

$$q_n = \sum_{i=1}^{m_q} e^{in\phi_i},$$

where the sum runs over only the m_q POI. One then defines the two- and four-particle correlations:

$$\langle 2' \rangle_n = \frac{q_n Q_n^* - m_q}{m_q M - m_q},$$

$$\langle 4' \rangle_n = [q_n Q_n Q_n^* Q_n^* - q_{2n} Q_n^* Q_n^* - q_n Q_n Q_{2n}^* - 2M q_n Q_n^*$$
$$- 2m_q |Q_n|^2 + 7q_n Q_n^* - Q_n q_n^* + q_{2n} Q_{2n}^*$$
$$+ 2q_n Q_n^* + 2m_q M - 6m_q]/$$
$$[(m_q M - 3m_q)(M-1)(M-2)].$$

In similar fashion to the integrated flow coefficients, these are averaged over all events, with weights $m_q M - m_q$ and $(m_q M - 3m_q)(M-1)(M-2)$ for the two- and four-particle correlations, respectively. The resulting averages over all events, $\langle\langle 2' \rangle\rangle_n$ and $\langle\langle 4' \rangle\rangle_n$, can then be used to obtain the differential flow coefficients for the p_T bin the POI were taken from:

$$v_n\{2\}(p_T) = \frac{\langle\langle 2' \rangle\rangle_n}{\sqrt{\langle\langle 2 \rangle\rangle_n}},$$

$$v_n\{4\}(p_T) = \frac{2\langle\langle 2' \rangle\rangle_n \langle\langle 2 \rangle\rangle_n - \langle\langle 4' \rangle\rangle_n}{\left(2\langle\langle 2 \rangle\rangle_n^2 - \langle\langle 4 \rangle\rangle_n\right)^{3/4}}.$$

In Fig. 25 we show the anisotropic flow coefficients for pions, kaons, and protons in different p_T bins as a function of centrality, as well as the integrated $v_n\{k\}$ for a simulation both including (solid) and without (dotted) the hadronic afterburner

SMASH. By showing the flow coefficients before applying the afterburner we clearly show how the afterburner can in this case significantly decrease anisotropic flow. The figure shows an impressive agreement with the data, especially considering how both the p_T bin as well as the centrality affect these different hadrons differently. The only significant deviations are found in the three highest p_T bins of pions and the highest p_T bin of kaons. It is interesting that for protons all curves agree, even though in the Bayesian analysis only the (1.0, 1.4) and (1.4, 1.8) bins were used due to the more limited statistics for the protons.

Figure 26 shows the equivalent figure for the flow coefficients in pPb collisions. Recall that we defined $\tilde{v}_n\{k\} \equiv$ sgn$(v_n\{k\}^k)|v_n\{k\}|$, for which $\tilde{v}_n\{k\}$ is negative when $v_n\{k\}$ is imaginary. Only the second and third harmonic are used for the posterior estimates. As also mentioned in Sec. IV A the emulator error for pPb is large for this observable, which is not included in the theoretical errors shown. It is however comforting to see a reasonable description of the data, including statistically significantly negative values for $\tilde{v}_3\{2\}$ and $\tilde{v}_4\{2\}$ (for $\tilde{v}_4\{2\}$ this is opposite from both ATLAS [78] and ALICE [91] experimental data), which in particular shows that even within a purely hydrodynamic model with hadronic cascade $v_n\{2\}$ is not necessarily real.

D. Event-plane correlations

Event-plane angle correlations are another interesting family of observables that we did not use in the Bayesian analysis. It is hence interesting to compare the results of TRAJECTUM to experimental data. We used the definition from the ATLAS collaboration [92]. For two-plane correlations, this starts by splitting each event into two subevents. Subevent N contains
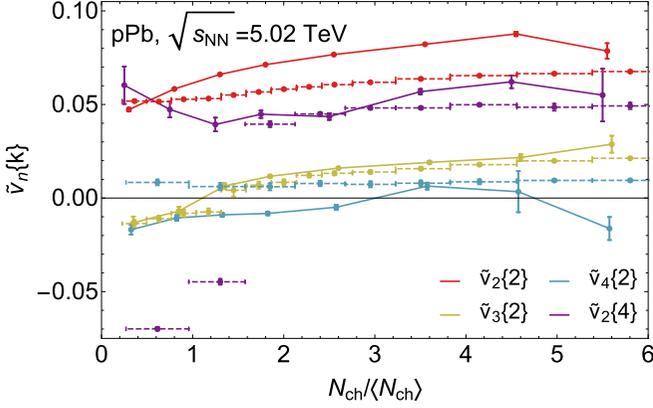
FIG. 26. We show the anisotropic flow coefficients for pPb collisions at 5.02 TeV obtained for our MAP values. Given the difficulty obtaining accurate pPb results and hence posterior distributions the model compares well with the data (dashed, [78]). Interestingly our method also reproduces the flow coefficients for negative $\tilde{v}_n\{k\}$, which corresponds to imaginary values for the standard $v_n\{k\}$.

all charged particles with pseudorapidity $-2.5 \leqslant \eta \leqslant -0.5$, while subevent P contains all charged particles with pseudorapidity $0.5 \leqslant \eta \leqslant 2.5$. In addition to the pseudorapidity requirement, both subevents require the transverse momentum to satisfy $0.5 \leqslant p_T \leqslant 5$ GeV. For each subevent, we then compute the Q vectors

$$Q_n = \sum_{j=1}^{M} e^{in\phi_j},$$

where the sum runs over all M particles in the subevent, n is an integer, and ϕ_i is the azimuthal angle of particle i. We then decompose each Q vector into a magnitude v_n and event-plane angle Ψ_n:

$$Q_n = M v_n e^{in\Psi_n}. \qquad (14)$$

The event-plane correlations are then defined as correlations between the Ψ_n as follows:

$$\langle \cos[k(\Phi_n - \Phi_m)] \rangle_{\text{SP}}$$
$$= \frac{\left\langle \left(v_n^P\right)^{k/n}\left(v_m^N\right)^{k/m} \cos\left[k\left(\Psi_n^P - \Psi_m^N\right)\right]\right\rangle + P \leftrightarrow N}{2\text{Res}\{k\Psi_n\}\text{Res}\{k\Psi_m\}}, \qquad (15)$$

where the average is over all events in the centrality class. For the observable to be properly defined k must be a multiple of both n and m. This correlation aims to measure the correlation between the true event-plane angles Φ_n, as defined from a (hypothetical) smooth distribution emitting a large number of particles. However, due to finite statistics the correlation between the measured event-plane angles Ψ_n is different. To compensate for this difference, the definition includes the resolution factors

$$\text{Res}\{k\Psi_n\} = \sqrt{\left\langle \left(v_n^N v_n^P\right)^{k/n} \cos\left[k\left(\Psi_n^P - \Psi_n^N\right)\right]\right\rangle}.$$

The three-plane correlations are defined in a similar fashion. In this case, each event is subdivided into three subevents. Subevent A contains all charged particles with $-2.5 \leqslant \eta \leqslant$

-1.5, subevent B contains charged particles with $-1 \leqslant \eta \leqslant 1$, and subevent C contains the charged particles with $1.5 \leqslant \eta \leqslant 2.5$. Transverse momenta satisfy $0.5 \leqslant p_T \leqslant 5$ GeV for each subevent. The three-plane correlations are then defined by

$$\langle \cos(i\Phi_l + j\Phi_m + k\Phi_n) \rangle_{\text{SP}}$$
$$= \frac{\left(c_{lmn}^{ijk}\right)^{BAC} + \left(c_{lmn}^{ijk}\right)^{BCA}}{\text{Res}\{i\Phi_l^B\}\text{Res}\{j\Phi_m^A\}\text{Res}\{k\Phi_n^C\} + A \leftrightarrow C}, \qquad (16)$$

where

$$\left(c_{lmn}^{ijk}\right)^{ABC}$$
$$= \left\langle \left(v_l^A\right)^{l/i}\left(v_m^B\right)^{m/j}\left(v_n^C\right)^{n/k} \cos\left(i\Psi_l^A + j\Psi_m^B + k\Psi_n^C\right)\right\rangle$$

and

$$\left(\text{Res}\{i\Phi_l^A\}\right)^2 = \left\langle \left(v_l^A v_l^B\right)^{i/l} \cos\left[i\left(\Psi_l^A - \Psi_l^B\right)\right]\right\rangle$$
$$\times \left\langle \left(v_l^A v_l^C\right)^{i/l} \cos\left[i\left(\Psi_l^A - \Psi_l^C\right)\right]\right\rangle /$$
$$\left\langle \left(v_l^B v_l^C\right)^{i/l} \cos\left[i\left(\Psi_l^B - \Psi_l^C\right)\right]\right\rangle.$$

For the observable to be properly defined, i/l, j/m, and k/n must be integers, and i, j, and k must add up to zero. Again, resolution factors are employed to compensate for the difference between the idealized true event-plane angles Φ_n and the measured event-plane angles Ψ_n.

Because the cosines always evaluate between -1 and 1, both the numerator and the denominator of (15) and (16) become quite accurate rather fast. However, for many observables both the numerator and the denominator are also close to zero, leading to large uncertainties in the final result. Figure 27 shows several event-plane correlations that have reasonably small final uncertainties for 200 000 events generated with TRAJECTUM using the MAP values.

The experimental agreement with ATLAS is excellent for the two-plane correlations, which is nontrivial since this observable is in a different class than the ones used to obtain the MAP values. We note that [14] obtained a similarly good agreement, but their agreement was only possible when specifically choosing $\eta/s = 0.20$. Reference [14] also obtained the event-plane angles directly from the hydrodynamic profile, whereas it is clear from (15) that significant modifications can come from both the finite statistic sample as well as the hadronic afterburner used in the current paper. For the three-plane correlations the statistics is more limited, but for those observables that are feasible to compute we also obtain good agreement.

E. Eccentricities and anisotropic flow

It is well known that hydrodynamics translates initial state spatial anisotropies into final state momentum anisotropies [95–98]. In particular, there is an approximately linear relation between initial state ellipticity ϵ_2 and the final state elliptic flow v_2 and similarly for the triangular flow v_3. The proportionality constant depends on quantities such as the shear viscosity. Here, v_n are defined on an event-by-event basis through (14), and ϵ_n can be defined from the initial state
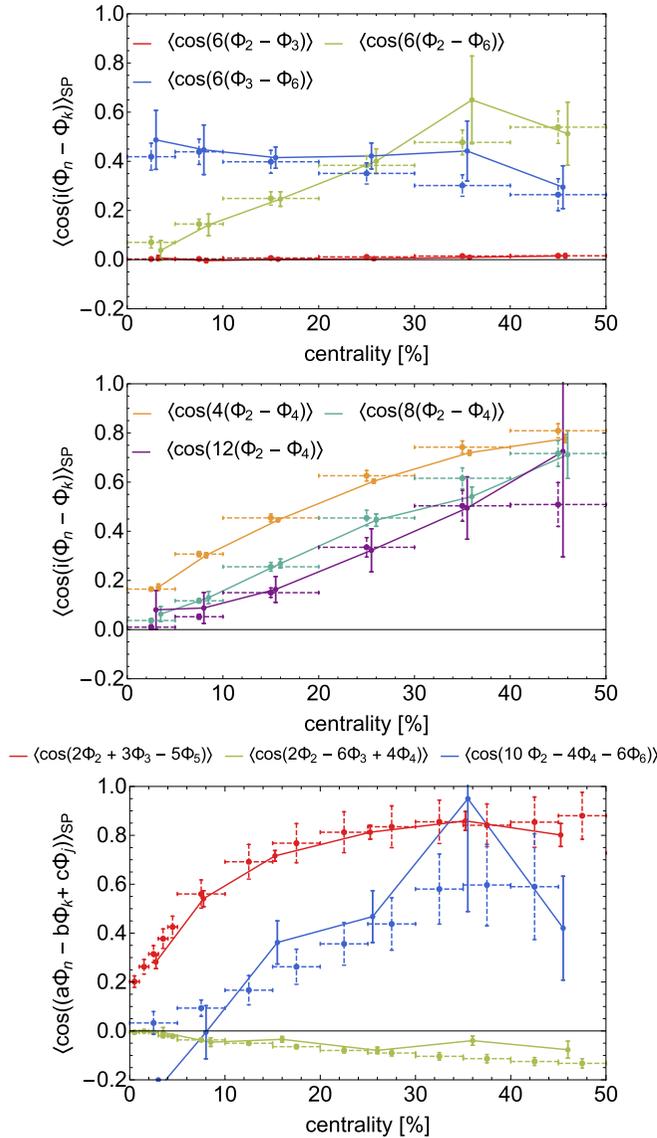
FIG. 27. Event-plane correlations using the MAP parameters (solid), for PbPb at 2.76 TeV. ATLAS data (dashed) are shown for comparison [92].

through [14]

$$\epsilon_n = \left| \frac{\int dx \, dy \, r^n e^{in\phi} \rho(x, y)}{\int dx \, dy \, r^n \rho(x, y)} \right|,$$

where $r^2 = x^2 + y^2$, $\phi = \arctan(y/x)$, and $\rho(x, y)$ is the energy density in the transverse plane. From a computational point of view, this is interesting, because it opens up the possibility of obtaining v_2 by running a small number of full hydrodynamic simulations to determine the proportionality constant, after which it is sufficient to obtain the initial state ϵ_2 to determine v_2.

In principle there is then the initial anisotropy ϵ_2, the anisotropy of the freeze-out surface which leads to anisotropic flow, and finally the anisotropy of the final particle spectra v_n. *A priori* it is not clear how to relate the final particle spectra to the hydrodynamic freeze-out surface, first of all because

of the influence of the afterburner and second since the finite multiplicity will lead to statistical fluctuations. In [93] these effects were unfolded using a Bayesian method, leading to a v_n distribution that is constructed to be similar to the freeze-out surface. In our code we have direct access to the initial eccentricity and the final particle spectra, and with a trick by generating a factor 100 more particles (oversampling) we can obtain an accurate estimate of the freeze-out surface anisotropy.

These results are illustrated in Fig. 28, where we show v_n as a function of ϵ_n for $n = 2$, 3, and 4 for three different cases. In the left panels we perform a full hydrodynamics simulation with afterburner, in the center left we perform the same calculation without the afterburner, and in the center right panels we perform a computation without afterburner, but increasing the number of sampled particles in the output by a factor 100. Different from the rest of this section, the results for "without afterburner" and "100x oversampled" were performed using 50 000 events. All panels also show linear fits through the distribution. One can clearly see that the correlation for $n = 2$ and 3 is much better when using oversampling. For $n = 4$, the correlation is not more pronounced with oversampling, indicating that the absence of the correlation is not due to statistics. The reason for this is actually physical, as v_4 is determined by a nonlinear combination of ϵ_2 and ϵ_4 [99].

A more detailed view can be obtained from the v_n and ϵ_n distributions, normalized to their mean values, as shown in the right column of Fig. 28. For ϵ_2 and v_2 the oversampled distribution differs significantly from the final v_n distributions, both with and without the hadronic afterburner, and is also different from the ϵ_2 distribution. For the third and fourth harmonics all distributions are similar, which reflects the fact that these harmonics originate from fluctuations, whereas v_2 has a large mean contribution from the ellipticity from the geometry. It is hence no surprise that the statistical fluctuations from the finite number of particles in the afterburner widen the v_2 distribution. Experimentally it is possible to unfold this statistical effect, and after this unfolding the data indeed agree with the oversampled result [93]. From a theoretical point of view this has the disadvantage that a direct comparison with the output of the hydrodynamic code with afterburner is not possible, but on the other hand it has the advantage that without using an afterburner the hydrodynamic profiles themselves can be directly compared to the experimental distributions, such as for instance done in [14].

V. DISCUSSION

We performed a detailed closure test, including carefully comparing the posterior distributions with the true parameters, which gives convincing evidence that the emulation and Monte Carlo work well. Nevertheless, we have to stress that a closure test does not tell if our model is reasonable as a physical description of heavy ion collisions. After all, the closure test only compares to output from the model itself and hence has to work by construction. It is hence always important to study weaknesses in the physical model in higher detail, especially where part of the physics may be missing. Clear examples can be intermediate p_T particles, which are not
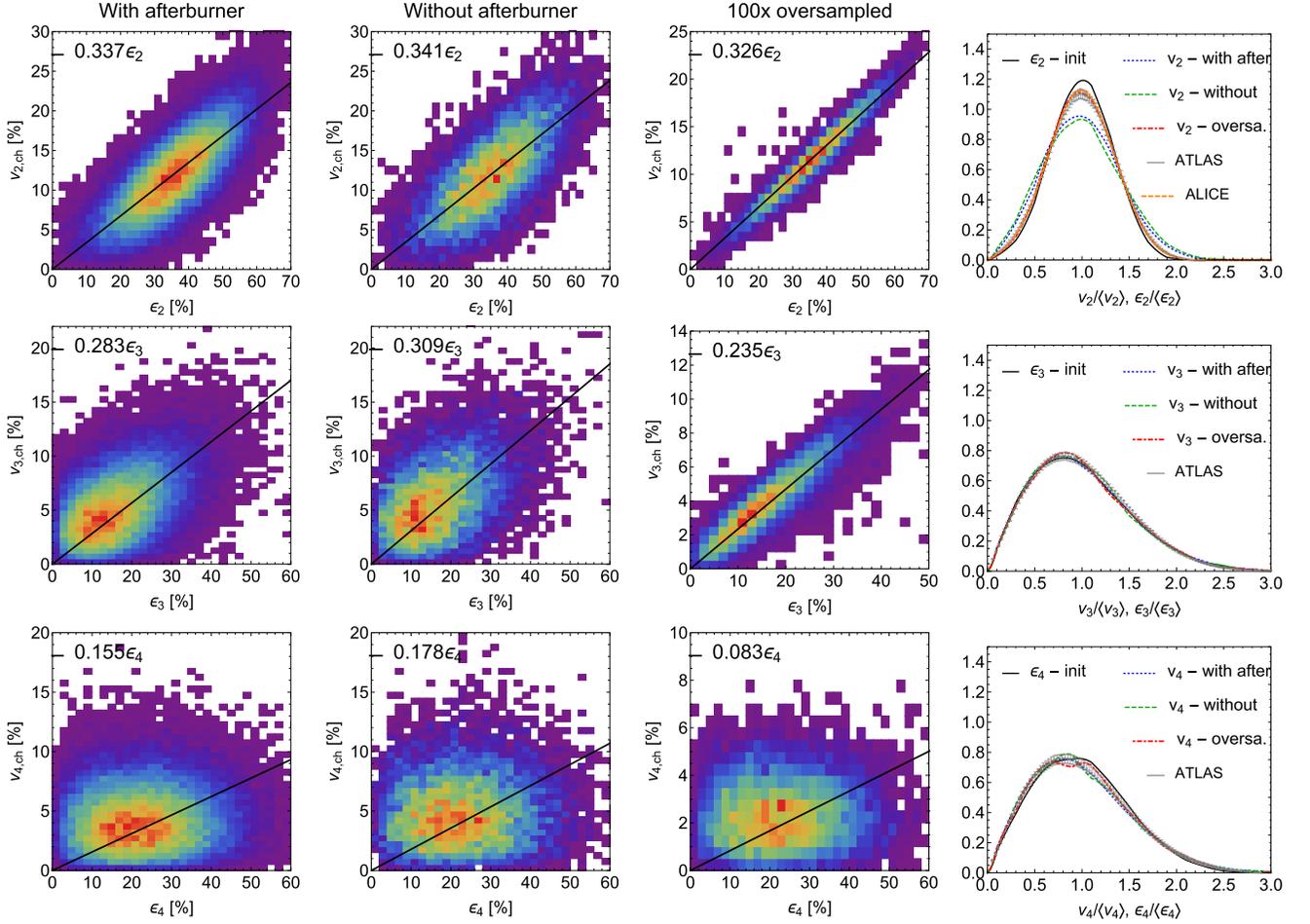
FIG. 28. Event-by-event anisotropic flow v_n [defined by (14)] as a function of initial state eccentricities ϵ_n, for $n = 2$ (top row), $n = 3$ (middle row), and $n = 4$ (bottom row), at 20–40% centrality for PbPb collisions at 2.76 TeV. The left column shows the result with the hadronic afterburner, and the center left column shows the result where the particles produced at particlization do not have any further interactions. The center right column is similar to the middle column, except that the number of produced particles at particlization has been artificially multiplied by a factor 100. A linear fit through the origin is also shown for each of these three panels. The rightmost panel shows projections on the ϵ_n and v_n axes of the three other columns together with experimental data from ATLAS (gray, [93]). The data are unfolded to match the oversampled distribution, which indeed compares well with these MAP results. ALICE has a similar measurement (orange, dashed, [94]), which uses a different method and agrees with [93].

necessarily described by a hydrodynamic or thermal model, especially in pPb systems. Nevertheless, as shown in Fig. 15 and verified in Sec. IV, our model comfortably fits almost all data, thereby giving confidence that our (phenomenological) model scope is wide enough to capture most of the physics presented.

As far as we are aware we are the first to obtain imaginary values for $v_3\{2\}$ and $v_4\{2\}$ using a purely hydrodynamic model (see Fig. 26). This is interesting, since without statistical fluctuations both $v_n\{4\}$ and especially $v_n\{2\}$ are expected to be real in hydrodynamics. Even though our posterior probabilities for pPb collisions are relatively uncertain, our MAP parameters give a significantly imaginary value for $v_3\{2\}$, which is in agreement with ATLAS data [78].

There are several avenues for improvement. First of all this can include adding more experimental data. Especially important to further constrain the temperature dependence

of the transport coefficients will be the addition of data at lower energies from RHIC (also done in [61,100]). One could also add additional system sizes, such as XeXe at the LHC or ^{3}HeAu at RHIC, or one could even study what would be gained by performing OO or ArAr collisions. It is computationally relatively difficult to add more PbPb and pPb observables at LHC energies, as many of these would require a much larger number of events per design point. One way forward in this scheme is to use simpler observables, such as initial state eccentricities, that correlate well with (more difficult) final state flow correlations. Initial investigations into such methods were presented in Sec. IV E (see also [14,93]). Another opportunity for such observables could be to perform a smaller run with better statistics.

Another potential improvement lies in the particlization procedure. Here, thermal particle production is modified by the viscous corrections to ensure continuity of the stress-
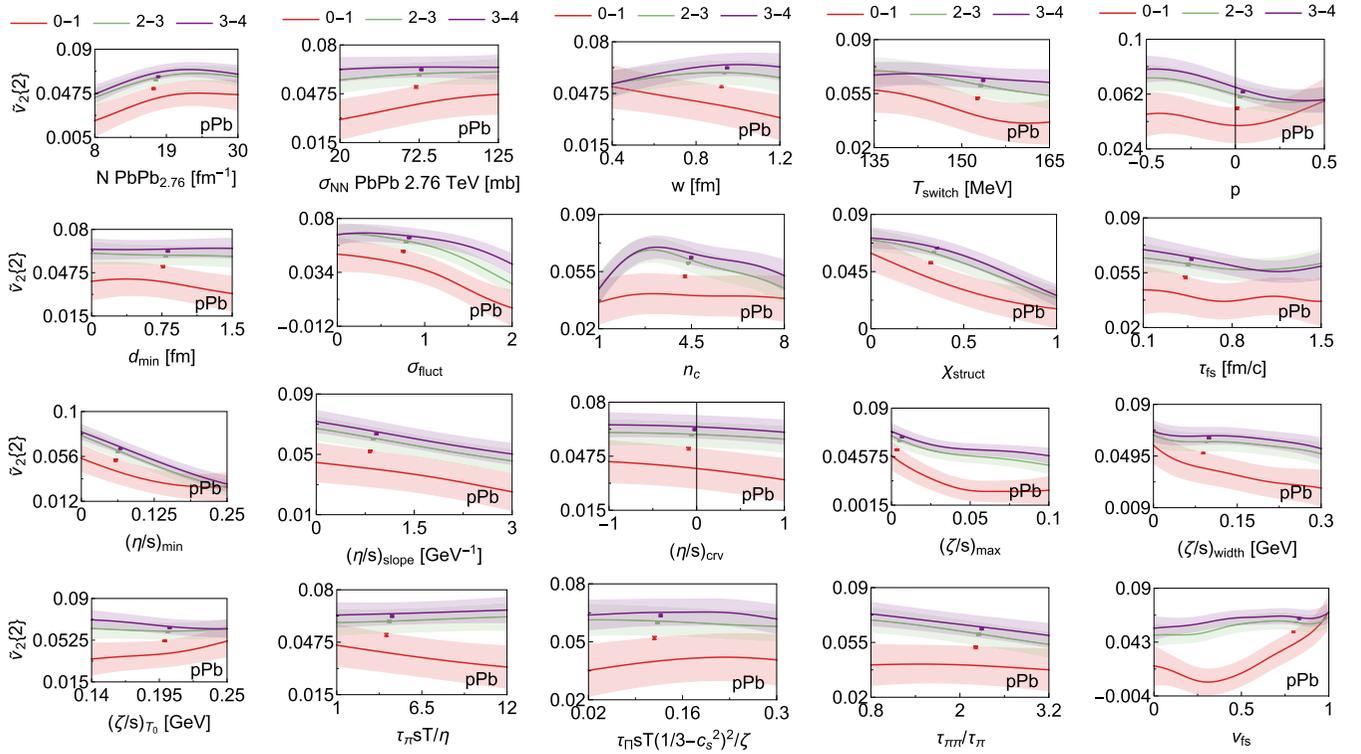
FIG. 29. Similar to Figs. 5 and 26 we show the $\tilde{v}_2\{2\}$ for several multiplicity classes. What is clear is that there is a much more refined dependence on all parameters, as opposed to the PbPb case, where the viscosity is mostly dominant. For these small systems also the subnuclear structure, parametrized by χ_{struct} and n_c, is more pronounced.

energy tensor. However, the precise manner in which this should be done is unknown. Several different prescriptions exist, and the choice of prescription is known to affect

the posterior distributions [61]. In the future, it would be interesting to improve this part of the simulation, either by data-driven methods such as those presented in [61],
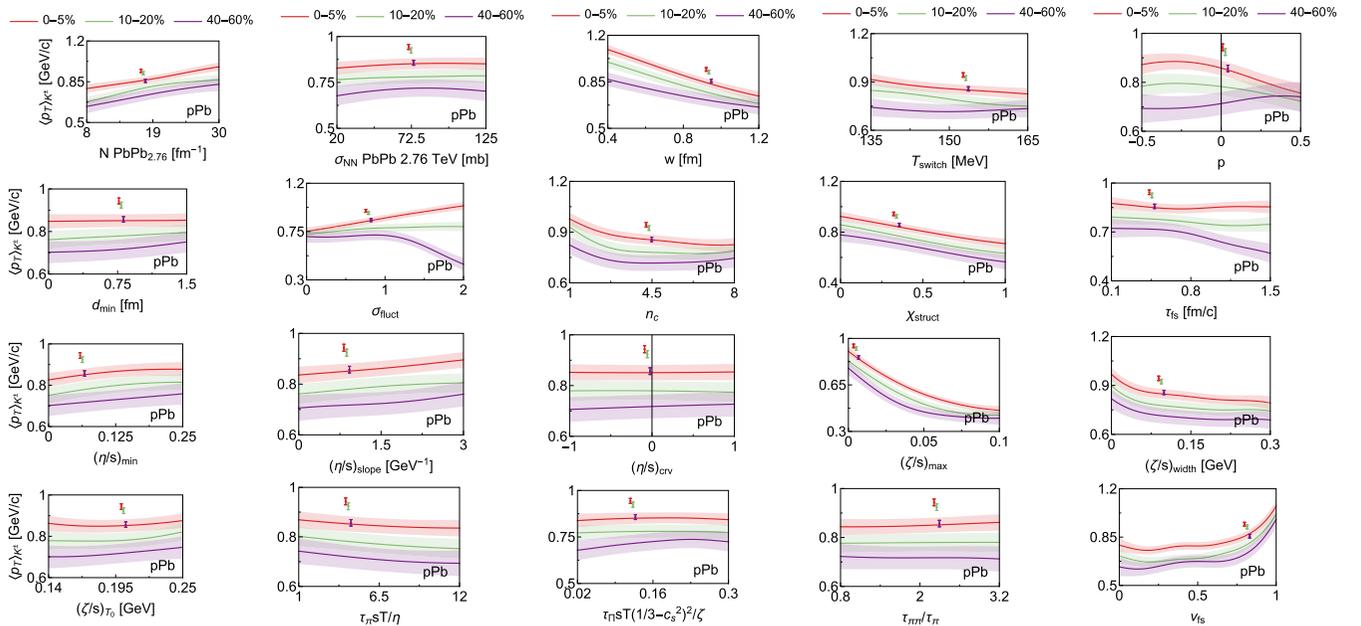
FIG. 30. Similar to Fig. 5 we show the mean transverse momentum for kaons for three centrality classes as a function of our model parameters. Datapoints [79] are shown at the MAP values and as discussed in Sec. IV A this particular datapoint does not match well. Note, however, that for this case a different and not excluded value of the width w or free streaming velocity v_{fs} could lead to a better fit.

or by microscopic insights guiding towards the correct prescription.

Currently the difficult emulation of pPb prevents the pPb data from providing a precise constraint in the Bayesian analysis. Our high statistics MAP results for $\tilde{v}_4\{2\}$ (see Fig. 26) do not agree well with ATLAS data [78]. This is an indication that perhaps tighter constraints can be obtained by improving the emulation of pPb further. Furthermore, we observe a deficit of kaon transverse momentum in pPb (see Fig. 22). One hint for resolving this deficit is the strong dependence on $v_{\rm fs}$ in Fig. 30. Higher values of $v_{\rm fs}$ would increase the mean kaon p_T, while these values are not excluded from the Bayesian analysis. It would be interesting to see whether an improved emulation and global fit will be able to resolve this discrepancy, or whether the discrepancy will require a nonhydrodynamic explanation.

Lastly, we wish to go beyond the more phenomenological model for the initial stage as presented here. Both the TRENTO initial energy deposition as well as the free streaming phase are not motivated by microscopic models, even though they may be wide enough in scope to describe some of these models. Perhaps most pressing is the fact that the switch from free streaming to hydrodynamics is not smooth, as shown in Fig. 19, whereas this transition is smooth in holographic [47,48] or weakly coupled models [9,101].

APPENDIX: PARAMETER DEPENDENCE OF pPb OBSERVABLES

We show the emulator results for pPb collisions in Figs. 29 and 30 for both anisotropic flow coefficients as well as the mean transverse momentum for kaons. The latter is of special interest, as for peripheral pPb collisions there is currently a clear tension between the MAP result and the experimental data (see Fig. 22).

[1] W. Busza, K. Rajagopal, and W. van der Schee, Annu. Rev. Nucl. Part. Sci. **68**, 339 (2018).

[2] U. Heinz and R. Snellings, Annu. Rev. Nucl. Part. Sci. **63**, 123 (2013).

[3] D. Teaney, J. Lauret, and E. V. Shuryak, Phys. Rev. Lett. **86**, 4783 (2001).

[4] P. Romatschke and U. Romatschke, Phys. Rev. Lett. **99**, 172301 (2007).

[5] P. K. Kovtun, D. T. Son, and A. O. Starinets, Phys. Rev. Lett. **94**, 111601 (2005).

[6] P. M. Chesler and L. G. Yaffe, Phys. Rev. D **82**, 026006 (2010).

[7] M. P. Heller, R. A. Janik, and P. Witaszczyk, Phys. Rev. Lett. **108**, 201602 (2012).

[8] M. P. Heller, D. Mateos, W. van der Schee, and D. Trancanelli, Phys. Rev. Lett. **108**, 191601 (2012).

[9] A. Kurkela and Y. Zhu, Phys. Rev. Lett. **115**, 182301 (2015).

[10] L. Keegan, A. Kurkela, P. Romatschke, W. van der Schee, and Y. Zhu, J. High Energy Phys. 04 (2016) 031.

[11] B. Alver, M. Baker, C. Loizides, and P. Steinberg, arXiv:0805.4411.

[12] J. S. Moreland, J. E. Bernhard, and S. A. Bass, Phys. Rev. C **92**, 011901(R) (2015).

[13] F. Gelis, E. Iancu, J. Jalilian-Marian, and R. Venugopalan, Annu. Rev. Nucl. Part. Sci. **60**, 463 (2010).

[14] H. Niemi, K. J. Eskola, and R. Paatelainen, Phys. Rev. C **93**, 024907 (2016).

[15] W. van der Schee and B. Schenke, Phys. Rev. C **92**, 064907 (2015).

[16] C. Gale, S. Jeon, B. Schenke, P. Tribedy, and R. Venugopalan, Phys. Rev. Lett. **110**, 012302 (2013).

[17] S. Ryu, J. F. Paquet, C. Shen, G. S. Denicol, B. Schenke, S. Jeon, and C. Gale, Phys. Rev. Lett. **115**, 132301 (2015).

[18] J. Novak, K. Novak, S. Pratt, J. Vredevoogd, C. E. Coleman-Smith, and R. L. Wolpert, Phys. Rev. C **89**, 034917 (2014).

[19] S. Pratt, E. Sangaline, P. Sorensen, and H. Wang, Phys. Rev. Lett. **114**, 202301 (2015).

[20] E. Sangaline and S. Pratt, Phys. Rev. C **93**, 024908 (2016).

[21] J. E. Bernhard, J. S. Moreland, S. A. Bass, J. Liu, and U. Heinz, Phys. Rev. C **94**, 024907 (2016).

[22] D. Devetak, A. Dubla, S. Floerchinger, E. Grossi, S. Masciocchi, A. Mazeliauskas, and I. Selyuzhenkov, J. High Energy Phys. 06 (2020) 044.

[23] J. E. Bernhard, J. S. Moreland, and S. A. Bass, Nat. Phys. **15**, 1113 (2019).

[24] J. E. Bernhard, Bayesian parameter estimation for relativistic heavy-ion collisions, Ph.D. thesis, Duke University, 2018.

[25] J. S. Moreland, Initial conditions of bulk matter in ultrarelativistic nuclear collisions, Ph.D. thesis, Duke University, 2019.

[26] J. L. Nagle and W. A. Zajc, Annu. Rev. Nucl. Part. Sci. **68**, 211 (2018).

[27] J. S. Moreland, J. E. Bernhard, and S. A. Bass, Phys. Rev. C **101**, 024911 (2020).

[28] G. Nijs, W. van der Schee, U. Gürsoy, and R. Snellings, Phys. Rev. Lett. **126**, 202301 (2021).

[29] S. Bass *et al.*, Prog. Part. Nucl. Phys. **41**, 255 (1998).

[30] M. Bleicher *et al.*, J. Phys. G **25**, 1859 (1999).

[31] J. Weil *et al.*, Phys. Rev. C **94**, 054905 (2016).

[32] D. Oliinychenko, V. Steinberg, J. Weil, M. Kretz, J. Staudenmaier, S. Ryu, A. Schäfer, J. Rothermel, J. Mohs, F. Li, H. E. Petersen, L. Pang, D. Mitrovic, A. Goldschmidt, L. Geiger, J.-B. Rose, J. Hammelmann, and L. Prinz, smash-transport/smash: Smash-1.8 (2020), doi:10.5281/zenodo.3742965.

[33] T. Sjostrand, S. Mrenna, and P. Z. Skands, Comput. Phys. Commun. **178**, 852 (2008).

[34] W. Ke, J. S. Moreland, J. E. Bernhard, and S. A. Bass, Phys. Rev. C **96**, 044912 (2017).

[35] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg, Annu. Rev. Nucl. Part. Sci. **57**, 205 (2007).

[36] M. Barej, A. Bzdak, and P. Gutowski, Phys. Rev. C **97**, 034901 (2018).

[37] C. Loizides, Phys. Rev. C **94**, 024914 (2016).

[38] P. Bożek, W. Broniowski, and M. Rybczyński, Phys. Rev. C **94**, 014902 (2016).

[39] K. Welsh, J. Singer, and U. W. Heinz, Phys. Rev. C **94**, 024919 (2016).

[40] R. D. Weller and P. Romatschke, Phys. Lett. B **774**, 351 (2017).

[41] S. S. Gubser, Phys. Rev. D **82**, 085027 (2010).

[42] S. S. Gubser and A. Yarom, Nucl. Phys. B **846**, 469 (2011).

[43] H. Marrochio, J. Noronha, G. S. Denicol, M. Luzum, S. Jeon, and C. Gale, Phys. Rev. C **91**, 014903 (2015).

[44] W. Broniowski, W. Florkowski, M. Chojnacki, and A. Kisiel, Phys. Rev. C **80**, 034902 (2009).

[45] U. W. Heinz and J. Liu, Nucl. Phys. A **956**, 549 (2016).

[46] https://github.com/Duke-QCD/freestream.

[47] W. van der Schee, Phys. Rev. D **87**, 061901(R) (2013).

[48] W. van der Schee, P. Romatschke, and S. Pratt, Phys. Rev. Lett. **111**, 222302 (2013).

[49] P. Romatschke and J. Hogg, J. High Energy Phys. 04 (2013) 048.

[50] M. Habich, J. Nagle, and P. Romatschke, Eur. Phys. J. C **75**, 15 (2015).

[51] D. Grumiller and P. Romatschke, J. High Energy Phys. 08 (2008) 027.

[52] J. Casalderrey-Solana, M. P. Heller, D. Mateos, and W. van der Schee, Phys. Rev. Lett. **111**, 181601 (2013).

[53] G. S. Denicol, S. Jeon, and C. Gale, Phys. Rev. C **90**, 024912 (2014).

[54] P. Huovinen and P. Petreczky, Nucl. Phys. A **837**, 26 (2010).

[55] A. Bazavov *et al.* (HotQCD Collaboration), Phys. Rev. D **90**, 094503 (2014).

[56] M. Luzum and P. Romatschke, Phys. Rev. C **78**, 034915 (2008); **79**, 039903 (2009).

[57] L. Du and U. Heinz, Comput. Phys. Commun. **251**, 107090 (2020).

[58] D. Bazow, U. W. Heinz, and M. Strickland, Comput. Phys. Commun. **225**, 92 (2018).

[59] S. Pratt and G. Torrieri, Phys. Rev. C **82**, 044901 (2010).

[60] Without bulk viscous corrections, the energy density would be automatically satisfied by sampling from a thermal distribution, but the λ_{bulk} rescaling spoils this.

[61] D. Everett *et al.* (JETSCAPE), arXiv:2010.03928.

[62] D. Everett, C. Chattopadhyay, and U. Heinz, arXiv:2101.01130.

[63] https://github.com/Duke-QCD/trento.

[64] B. Abelev *et al.* (ALICE Collaboration), Phys. Rev. C **88**, 044909 (2013).

[65] J. Adam *et al.* (ALICE Collaboration), Phys. Rev. C **91**, 064905 (2015).

[66] M. Stein, Technometrics **29**, 143 (1987).

[67] C. K. Williams and C. E. Rasmussen, *Gaussian Processes for Machine Learning* (MIT, Cambridge, MA, 2006), Vol. 2.

[68] J. Adam *et al.* (ALICE Collaboration), J. High Energy Phys. 09 (2016) 164.

[69] B. Abelev *et al.* (ALICE Collaboration), Phys. Rev. C **88**, 044910 (2013).

[70] M. Miller and R. Snellings, arXiv:nucl-ex/0312008.

[71] K. Aamodt *et al.* (ALICE Collaboration), Phys. Rev. Lett. **106**, 032301 (2011).

[72] J. Adam *et al.* (ALICE Collaboration), Phys. Rev. Lett. **116**, 222302 (2016).

[73] J. Adam *et al.* (ALICE Collaboration), Phys. Rev. C **94**, 034903 (2016).

[74] J. Adam *et al.* (ALICE Collaboration), Phys. Rev. Lett. **116**, 132302 (2016).

[75] B. B. Abelev *et al.* (ALICE Collaboration), Eur. Phys. J. C **74**, 3077 (2014).

[76] J. Adam *et al.* (ALICE Collaboration), Phys. Lett. B **760**, 720 (2016).

[77] S. Acharya *et al.* (ALICE Collaboration), J. High Energy Phys. 09 (2018) 006.

[78] M. Aaboud *et al.* (ATLAS Collaboration), Eur. Phys. J. C **77**, 428 (2017).

[79] B. B. Abelev *et al.* (ALICE Collaboration), Phys. Lett. B **728**, 25 (2014).

[80] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, Publ. Astron. Soc. Pac. **125**, 306 (2013).

[81] J.-F. Paquet *et al.* (JETSCAPE Collaboration), Nucl. Phys. A **1005**, 121749 (2021).

[82] D. Teaney, Phys. Rev. C **68**, 034913 (2003).

[83] B. B. Abelev *et al.* (ALICE Collaboration), Phys. Lett. B **736**, 196 (2014).

[84] J. Adam *et al.* (ALICE Collaboration), Nat. Phys. **13**, 535 (2017).

[85] J. Adam *et al.* (ALICE Collaboration), Phys. Lett. B **758**, 389 (2016).

[86] V. Vovchenko, B. Dönigus, and H. Stoecker, Phys. Rev. C **100**, 054906 (2019).

[87] J. Cleymans, P. M. Lo, K. Redlich, and N. Sharma, Phys. Rev. C **103**, 014904 (2021).

[88] F. Bellini (ALICE Collaboration), Nucl. Phys. A **982**, 427 (2019).

[89] A. Bilandzic, R. Snellings, and S. Voloshin, Phys. Rev. C **83**, 044913 (2011).

[90] A. Bilandzic, Anisotropic flow measurements in ALICE at the large hadron collider, Ph.D. thesis, Utrecht University, 2012 .

[91] S. Acharya *et al.* (ALICE Collaboration), Phys. Rev. Lett. **123**, 142301 (2019).

[92] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. C **90**, 024905 (2014).

[93] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. 11 (2013) 183.

[94] S. Acharya *et al.* (ALICE Collaboration), J. High Energy Phys. 07 (2018) 103.

[95] J.-Y. Ollitrault, Phys. Rev. D **46**, 229 (1992).

[96] H. Niemi, G. S. Denicol, H. Holopainen, and P. Huovinen, Phys. Rev. C **87**, 054901 (2013).

[97] G. Giacalone, J. Noronha-Hostler, and J.-Y. Ollitrault, Phys. Rev. C **95**, 054910 (2017).

[98] R. S. Bhalerao, G. Giacalone, P. Guerrero-Rodríguez, M. Luzum, C. Marquet, and J.-Y. Ollitrault, Acta Phys. Pol. B **50**, 1165 (2019).

[99] L. Yan and J.-Y. Ollitrault, Phys. Lett. B **744**, 82 (2015).

[100] J. Auvinen, J. E. Bernhard, S. A. Bass, and I. Karpenko, Phys. Rev. C **97**, 044905 (2018).

[101] A. Kurkela, A. Mazeliauskas, J.-F. Paquet, S. Schlichting, and D. Teaney, Phys. Rev. C **99**, 034910 (2019).