# QUANTIFICATION OF THE QUARK-GLUON PLASMA WITH STATISTICAL LEARNING

MATTHEW RAMIN HAMEDANI HEFFERNAN

Department of Physics
McGill University, Montréal

August 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACRONYMS

ALICE A Large Ion Collider Experiment

ATLAS A Toroidal LHC ApparatuS

BMA Bayes Model Average

C.-E. Chapman-Enskog

C.I. Credible Interval

FS Freestreaming

HICs Heavy Ion Collisions

LHC Large Hadron Collider

MAP Maximum a Posteriori

MCMC Markov Chain Monte Carlo

ML Maximum Likelihood

MUSIC MUScl for Ion Collisions

QGP Quark-Gluon Plasma

SMASH Simulating Many Accelerated Strongly-interacting Hadrons

T$_R$ENTo Thickness (reduced) Event-by-event Nuclear Topology

UrQMD Ultra relativistic Quantum Molecular Dynamics

iSS iSpectra Sampler

iS$_3$D iSpectra 3D

# ABSTRACT

Heavy ion collisions performed at facilities such as the Large Hadron Collider (LHC) and the Relativistic Heavy Ion Collider (RHIC) produce the hottest matter in the universe at $\sim 10^{12}$ K. This generates an energetic state of matter in which quarks and gluons become deconfined, known as the Quark Gluon Plasma. However, this material only survives for approximately $10^{-23}$ seconds, presenting many challenges for precise study.

This thesis uses Bayesian methods for systematic model-to-data comparison to quantify the properties of this material with a fully state-of-the-art hybrid model of heavy ion collisions. The hybrid model consists of IP-Glasma for the initial collision and pre-equilibrium evolution, 2+1D MUSIC viscous hydrodynamics, iS3D particlization, and SMASH for the hadronic cascade. This work has produced the most accurate and precise constraint in the literature on the understanding of the shear and bulk viscosity, energy scale, hydrodynamic onset time, and particlization temperature of the Quark Gluon Plasma. By using a pre-equilibrium model with physically-motivated microscopic physics for the first time, the sensitivity of experimental results to material properties of the plasma has been revealed. This highlights the importance of accurately incorporating pre-equilibrium physics in the hybrid model to describe the hydrodynamic evolution.

Sensitivity of inferred parameters to modeling choices is investigated with model averaging and computational improvements are realized through computer experiment design (an ordered Maximum Projection Latin Hypercube) and transfer learning, both of which are applied to the study of heavy ion collisions for the first time. Importantly, this work also lays the foundations for further extensions using transfer learning.

# RESUMÉ

Les collisions d'ions lourds menées au Grand collisionneur de hadrons (LHC) et au Collisionneur d'ions lourds relativistes (RHIC) produisent la matière la plus chaude de l'univers, à $\sim 10^{12}$ K. Ces conditions sont propices à la création d'un état de la matière où quarks et gluons se déconfinent, soit le plasma de quarks-gluons. Cet état est toutefois éphémère puisqu'il ne survit qu'environ $10^{-23}$ secondes, compliquant ainsi toute analyse expérimentale précise.

Cette thèse de doctorat emploie les méthodes bayésiennes dans la comparaison systématique de modèle avec des données expérimentales pour quantifier les propriétés de cet état en s'appuyant sur un modèle hybride des collisions d'ions lourds. Les composantes du modèle sont IP-Glasma pour les conditions initiales et l'évolution pré-équilibre, MUSIC pour l'évolution hydrodynamique visqueuse en 2+1D et SMASH pour la cascade hadronique finale. Cette recherche établie les contraintes les plus précises sur les viscosités de cisaillement et de volume, l'échelle énergétique, le temps de commencement de la phase hydrodynamique et la température de fragmentation particulaire du plasma de quarks-gluons. En se basant sur un modèle de la physique microscopique pré-équilibre (soit IP-Glasma) pour la toute première fois, la sensibilité des résultats expérimentaux aux propriétés matérielles du plasma sont révélées. Ces découvertes soulignent l'importance de modèles physiques de pré-équilibre pour l'évolution hydrodynamique.

La sensibilité des paramètres liés au choix du modèle est étudiée en faisant une moyenne statistique de modèles. De plus, des améliorations de calculs informatiques ont été réalisées grâce à l'utilisation d'une procédure informatique expérimentale (dite "ordered Maximum Projection Latin Hypercube") et l'apprentissage par transfert, dont l'usage combiné est une première dans l'étude des collisions d'ions lourds. Ce travail servira aussi de fondation pour de futures extensions employant l'apprentissage par transfert.

# ACKNOWLEDGMENTS

I gratefully acknowledge scholarships from the Natural Sciences and Engineering Research Council of Canada (NSERC), first a Postgraduate Scholarship - Doctoral (2019-2021) and then a Canada Graduate Scholarship - Doctoral (2021-2022). I also wish to acknowledge the high performance computing resources of the Compute Canada Federation and its regional partners, Calcul Québec, Compute Ontario, SciNet, SHARCNET, and WestGrid. Calculations took place on the following Compute Canada systems: Narval, Béluga, Niagara, Graham, and Cedar. Additional resources were supplied by the McGill Department of Physics with support from NSERC.

I would like to thank Charles Gale, my supervisor, for his mentorship, guidance, support, and advice throughout my graduate studies. His support and supervision were key to my research progress and development as a physicist. I would also like to thank Jean-François Paquet for his support and guidance throughout my doctoral studies and for always pushing me to be certain beyond a reasonable doubt that details were exactly and precisely correct. Sangyong Jeon's advice and guidance have been deeply appreciate and I have benefited greatly from his physics expertise and perspective throughout my graduate work. I also wish to thank Nikolas Provatas for taking me under his wing as an unofficial student and for his boundless support.

This research was enabled in part by discussions as a member of the JETSCAPE Collaboration, especially with Derek Everett and Dananjaya Liyanage on Bayesian methods. I also wish to thank my fellow graduate students in the high energy nuclear theory group at McGill, whose expertise and companionship were of great scientific and personal value. In particular, I would like to thank Scott McDonald, Mayank Singh, Nicolas Fortier, and Rouzbeh Modarresi Yazdi. Thanks are also due to Matthew Frick, for his camaraderie and soundness.

I wish to thank Sophie Huddart for her unending love and support that have made my life better in every way. Last but not least, I thank my family for their care, love, and support of my interests that has enabled me to come this far.

# CONTRIBUTION OF AUTHOR TO ORIGINAL KNOWLEDGE

Chapters 1-3 These chapters are a general introduction to heavy ion collisions, modeling the soft sector, and statistical methods with a review of relevant literature. In chapter 3, I use the simple pendulum as a tractable example to demonstrate Bayesian inference and Bayesian modeling workflow. The idea and implementation are novel and performed by me, first appearing in pre-print in [1], but adapted for this thesis.

Chapter 4 This chapter produces novel comparisons and analysis between the existing state-of-the-art models and the models used in this work to motivate the main study in this thesis. The comparisons were performed by me as well as a substantial number of the underlying computations. Some comparisons previously appeared in [2]. The chapter then motivates and discusses design considerations for a new Bayesian study, with some comparisons previously appearing in [3]. These motivations and design considerations are novel to this thesis. The investigation of design space sampling strategies is also novel to this thesis and this thesis employs an ordered Maximum Projection Latin Hypercube design for the first time in heavy ion collisions.

Chapter 5 This chapter investigates the pre-equilibrium stage in isolation. I developed significant automation to produce the underlying calculations on high performance computing infrastructure. I then implemented the inference, analysis, and discussion and they are novel to this thesis. I also implemented non-uniform priors for the first time in a Bayesian study of heavy ion collisions and motivated a larger-scale Bayesian study of a hybrid model with an IP-Glasma pre-equilibrium stage.

Chapter 6 I designed, adapted software, implemented, and executed a global Bayesian analysis of heavy ion collisions using a hybrid model with an IP-Glasma pre-equilibrium stage. I performed calculations requiring approximately 500 core-

years of computing resources and 100 TB of storage on high performance computing infrastructure. I implemented a linear transfer learning model, first proposed with proof of concept by coauthors and myself in [4], but never before implemented or used in practice for Bayesian inference in heavy ion collisions. I inferred the properties of matter produced in heavy-ion collisions and used Bayesian model comparison to reveal physical features required by the data. I used Bayesian model averaging to produce best estimates with uncertainty of the properties of strongly-interacting matter. I calculated physical observables with maximum a posteriori configurations of the model and make predictions.

# Part I

# PHYSICAL FOUNDATIONS

This part provides an introduction to heavy ion physics as well as the phenomenology investigated in this thesis, with a review of relevant literature. The multi-stage physics model used for simulation is described and important considerations are discussed.

# 1

# INTRODUCTION TO HEAVY ION COLLISIONS

On 15 September 1896, two unmanned railway locomotives were crashed into each other at high speed in a publicity stunt known as "The Crash at Crush". Imagine yourself to be an enthusiastic student of thermodynamics who arrived at the site of the crash, which had remained untouched for a century, knowing only that two locomotives collided and wishing to learn about heat engines such as the steam engine in the locomotives. If this student's only way of studying heat engines were to collide many different models of period trains until the scattered parts resembled those at the Crash at Crush, they would be on their way to being a heavy-ion physicist attempting to study the quark-gluon plasma produced in heavy ion collisions.

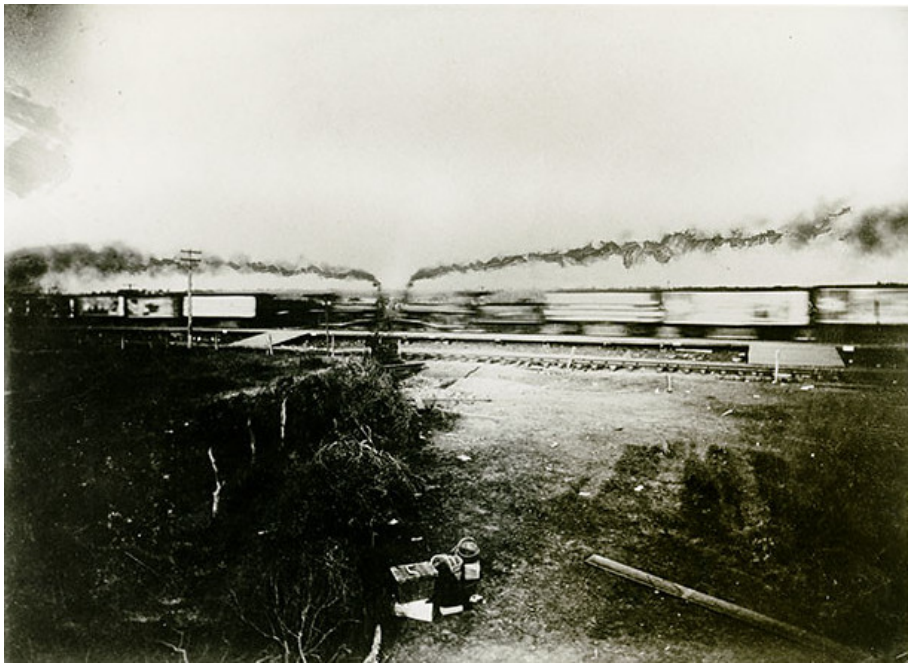

Figure 1.1: The moment of collision of the Crash at Crush [5].

Modeling heavy-ion collisions is a challenging exercise: the system is extraordinarily dynamic, but occurs so quickly that it is difficult to gain direct insight. Instead, the material produced must be studied using the particles that reach the detectors. While the thermodynamics enthusiast arrived armed with a knowledge of thermodynamics,

those curious about quark-gluon plasma must arm themselves with the fundamental theory of the strong nuclear force.

## 1.1 QUANTUM CHROMODYNAMICS

Four broadly-independent forces describe the present understanding of nature: electromagnetism, gravitation, and the strong and weak nuclear forces. Electromagnetism and the weak nuclear force are unified at high energy scales and become the electroweak interaction. This thesis is concerned with studying the strong nuclear force, described by Quantum ChromoDynamics (QCD). This theory governs the dynamics of quarks and gluons and has three charges. These charges are described in analogy to the colors of light, hence the term "chromo", and composite particles are always color-neutral. The color-neutral state is called "white" and the charges are red, green, and blue.

The QCD Lagrangian,

$$\mathcal{L}_{QCD} = \bar{\psi}_{q,a}\left(i\gamma^{\mu}\partial_{\mu}\delta_{ab} - g_s\gamma^{\mu}t^{C}_{ab}\mathcal{A}^{C}_{\mu} - m_q\delta_{ab}\right)\psi_{q,b} - \frac{1}{4}F^{A}_{\mu\nu}F^{A\mu\nu}, \tag{1.1}$$

$$F^{A}_{\mu\nu} = \partial_{\mu}\mathcal{A}^{A}_{\nu} - \partial_{\nu}\mathcal{A}^{A}_{\mu} - g_s f_{ABC}\mathcal{A}^{B}_{\mu}\mathcal{A}^{C}_{\nu}, \tag{1.2}$$

characterizes the strong interaction. $\mathcal{L}_{QCD}$ is the QCD Lagrangian and $g_s$ is the QCD coupling constant. Color indices are $(a, b)$, the gluon indices $(A, B, C)$, the SU(3) generators $(t^{C}_{ab})$, and the structure constants $f_{ABC}$ arise by the commutation relation $\left[t^{A}, t^{B}\right] = if_{ABC}t^{C}$ [6]. Throughout this thesis, repeated indices such as in the QCD Lagrangian, are summed over.

The strength of the QCD coupling poses a unique challenge. In Quantum Electrodynamics (QED) and the weak nuclear force, the coupling is small at low energies and increases with increasing energy scale. The strong force behaves differently; it is comparatively strong at low energies and decreases in strength at progressively higher energies. This is revealed by the sign of the QCD $\beta$ function,

$$\beta(g) = -\frac{g^3}{(4\pi)^2}\left(\frac{11N_c - 2n_f}{3}\right), \tag{1.3}$$

where $g$ is the strong coupling constant, $N_c$ is the number of colors, and $n_f$ is the number of approximately massless quark flavors (2 in QCD). The negative sign of the

QCD β function arises because of the gluon self-coupling. The β function describes the rate of change of the renormalized coupling $\alpha_s(Q)$ as the renormalization scale (*i.e.* energy scale) Q is varied,

$$\begin{aligned}\alpha_s(Q) &= -\frac{g^3}{8\pi\beta(g)\log(Q/\Lambda)} \\ &= \frac{6\pi}{(11N_c - 2n_f)\log(Q/\Lambda)}\end{aligned} \tag{1.4}$$

where $\Lambda$ is a scale parameter, usually taken to be $\approx 200$ MeV and $N_c = 3$ for QCD [7]. The renormalized coupling $\alpha_s$ thus clearly decreases with increasing energy.

A standard mathematical tool for investigating the behavior of forces is perturbation theory, which assumes a weak coupling. This is not the case for the strong force at low energies. As a result, its low energy behavior is difficult to calculate. Non-perturbative numerical techniques on the lattice (Lattice QCD) excel above the deconfinement transition or crossover temperature [8] where the degrees of freedom are quarks and gluons but do not smoothly connect to the hadronic phase. However, at high energies, QCD is weakly interacting and perturbative calculations are possible. The running of the strong coupling $\alpha_s$ with energy scale Q is shown in Fig. 1.2.

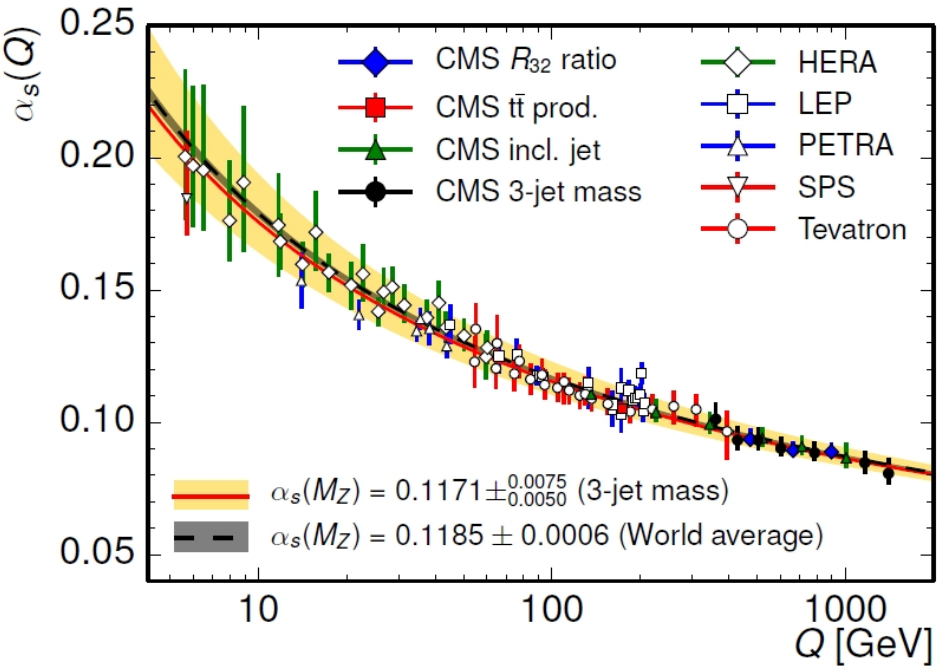


Figure 1.2: The running of the QCD coupling constant with energy scale Q, adapted from [9].

Exotic phases of nuclear matter are theorized to exist in different regions of the QCD phase diagram (Fig. 1.3). Of these phases, the high-temperature plasma of quarks and gluons, transitioning to a hadron gas, is the primary focus of this work.

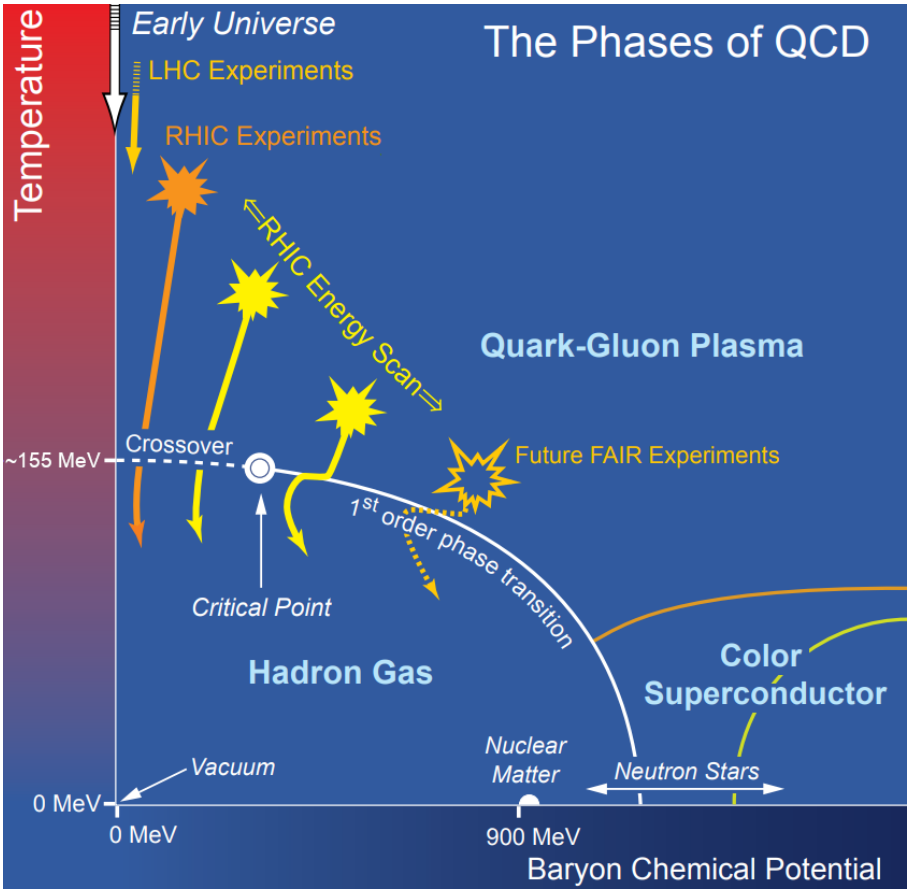


Figure 1.3: A schematic QCD phase diagram, adapted from [10]. The estimated crossover temperature is from [11].

Significant efforts are underway to understand the collective behavior of quarks and gluons and involves constructing models of relativistic hydrodynamics. This is used to explore the physics and reproduce observable quantities from the collision of heavy ions at particle accelerators [12], such as the Relativistic Heavy-Ion Collider (RHIC) at Brookhaven National Laboratory in New York and the Large Hadron Collider (LHC) in Geneva. RHIC and the LHC determined that heavy-ion collisions create such a macroscopic state of quarks and gluons that behaves as a strongly-interacting viscous fluid [13]. As it is a viscous fluid, the dynamics of this macroscopic state can be described using transport coefficients such as shear viscosity, bulk viscosity, and charge conductivity. Transport coefficients determine how the medium reacts to small deviations from equilibrium. These transport coefficients can either

be determined through microscopic calculations [14–18] or through inference using model-to-data comparison. Determining the value and temperature dependence of transport coefficients and other properties of the medium using statistical learning is the central focus of this work.

In the low-temperature region, QCD matter forms a gas of hadrons. In the high-temperature region, QCD matter takes the form of a plasma of quarks and gluons. The transition between the hadron gas and quark-gluon plasma occurs in a crossover region [19, 20], shown in Fig. 1.3, and estimated to be at $155 \pm 1.5$ MeV [11]. At higher net baryon density, a first-order phase transition has been predicted [20, 21]. This implies the existence of a critical point where there ceases to be a crossover from the hadron gas to a quark-gluon plasma. The investigation of the critical point is a subject of intense effort, but is beyond the scope of this work.

## 1.2 HEAVY ION COLLISIONS

When heavy ions such as lead or gold collide, there are initial large gradients and high energies in a phase strongly out of equilibrium. Then, the system approaches local thermal equilibrium and enters an intermediary region. In this intermediate energy region, the droplets of strongly coupled matter are well-described by hydrodynamics. Like a fireball, the strongly coupled matter expands and cools, eventually forming a gas of hadrons. Finally, the mean free path becomes long and the system is no longer well-described by hydrodynamics, but is better described by a lower-energy hadron gas. This hadronic ensemble interacts until it becomes sufficiently dilute and travels to the detectors. These stages are shown in Fig. 1.4. This description is the "standard model" of the soft sector of heavy ion physics and has been constrained by years of data-taking and theoretical modeling. For recent reviews, including experimental evidence for collective behavior, see [22, 23]. The "soft" sector indicates that jet physics is not considered at this level although interactions between jets and the hydrodynamic medium are the subject of active research [24–26]. This thesis is focused exclusively on studying the soft sector of heavy ion collisions.

As described previously, the stages shown in Fig. 1.4 can be condensed into three epochs: pre-equilibrium and thermalization, hydrodynamics, and hadronization. Ad-

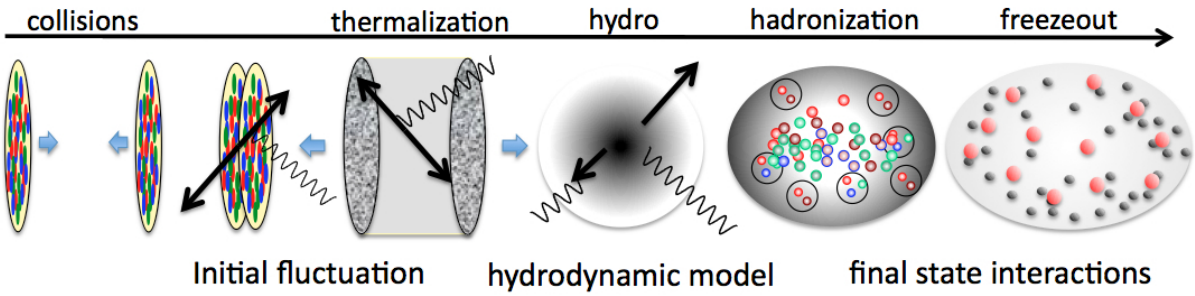


Figure 1.4: Stages of heavy ion collisions, adapted from [27].

vanced hybrid model simulations of heavy ion collisions with a hydrodynamic stage have been successful in reproducing hadronic observables [28–30]. This supports the notion that hydrodynamics is an accurate effective field theory of the long wavelength modes and is applicable when the mean free path is short relative to the size of the system. As the system cools and expands, the mean free path grows to the point where hydrodynamics is no longer applicable. This is known as particlization. After this, a "hadronic afterburner" can be used to evolve particles from the end of hydrodynamics to the particle shower observed in detectors. Two examples of contemporary hadronic afterburners are UrQMD [31] and SMASH [32].

The evolution from collision to freezeout all takes place within approximately 10 fm/c, or $1 \times 10^{-22}$ seconds. As a result, it is not possible to directly observe each stage as one would with a traditional microscope or camera. Instead, the only way to image these dynamics is via particles that reach the detectors. The traditional coordinate system used is with the $z$-axis along the beam and with $x$ and $y$ denoting the transverse plane. The momenta of the particles is recorded, typically the transverse momentum $p_T = \sqrt{p_x^2 + p_y^2}$ and the azimuthal angle in momentum space $\phi_p = \arctan\left(\frac{p_x}{p_y}\right)$. Additional kinematic variables that will be used throughout this thesis are the rapidity

$$y = \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right) \tag{1.5}$$

where $E$ is the energy $E = \sqrt{\mathbf{p}^2 + m^2}$. Pseudorapidity $\eta$ is also often used and is identical to the rapidity for massless particles,

$$\eta = \frac{1}{2} \ln\left(\frac{|\mathbf{p}| + p_z}{|\mathbf{p}| - p_z}\right) \tag{1.6}$$

where bold has been used to denote the momentum three-vector. This slightly over-subscribes the notation, with $\eta$ representing both the pseudorapidity and the shear viscosity. The shear viscosity is always referred to as the specific shear viscosity $\eta/s$ and the difference will always be made clear in context. This thesis specifically considers models that assume boost-invariance, *i.e.* without longitudinal structure. This approximation is best at mid-rapidity, and so comparisons are often made within a small rapidity (or pseudo-rapidity) window such as $|\eta| < 0.5$, chosen to match experimental measurement. This is the case for the charged hadron multiplicity $dN_{ch}/d\eta$ or identified particle yields, *e.g.* $dN_{\pi^+}/dy$. A more complete reference for particle kinematics, as well as particle listings, can be found in [6].

While such overall yields and multiplicity provide information about the overall size of the material produced, the geometry is revealed by the azimuthal angle of the particles $\phi_p$. Each collision produces many particles, whose spectra can be decomposed with a Fourier series in azimuthal angle, *e.g.*

$$\frac{dN_{ch}}{p_T dp_T dy d\phi_p} = \frac{dN_{ch}}{2\pi p_T dp_T dy}\left\{1 + 2\sum_{n=1}^{\infty} v_n(y, p_T)\cos\left[n(\phi_p - \Psi_n(y, p_T))\right]\right\} \quad (1.7)$$

where $\Psi_n$ denote the event plane angle, or the offset of the overall angular orientation from 0. The expansion coefficients of this series $v_n$ yield valuable geometric insights and are an important observable used throughout this thesis. Because these expansion coefficients $v_n$ reveal information about the momentum-space geometry, they are a key probe of collective behavior and are thus an ideal way to learn about the hydrodynamic stage.

This begins to yield a glance at how heavy ion collisions can be imaged, with the size and geometry characterized through detected particles. More quantities can be carefully constructed, yielding insight into different aspects of the collision. Examples of how different particle species and quantities provide multi-messenger insight into the evolution of heavy ion collisions may be found in [33]. Despite such success in reconstructing the collision, the hadrons detected in experiments are all produced at relatively late stages of the collision. Understanding the early stages of the collision must use additional probes, such as particles that only interact electromagnetically as they traverse the medium, or an extensive campaign of model building and systematic comparison to data.

## 1.3 STATUS OF BAYESIAN STUDIES IN HICS

Theoretical study of heavy ion collisions has produced many physical insights but has only recently begun to incorporate more advanced statistical methods. These statistical methods are a key step in gaining quantitative insight, but also in quantifying uncertainty. This quantification with uncertainty is beginning to yield guidance for theoretical efforts.

The first efforts in heavy ion collisions to gain quantitative insights using advanced statistical techniques for model-to-data comparison were [34, 35] followed by the MADAI collaboration [36]. These used simpler hybrid models without any pre-equilibrium dynamics and demonstrated the wealth of information that could be gained through systematic comparison.

These analyses were improved on by the DukeQCD research efforts [37–44], which also introduced a more flexible parametric model to allow the data to speak more for itself rather than priming the extractions with constraining prior assumptions. These analyses used a more sophisticated hybrid model, including pre-equilibrium dynamics in the form of free streaming. They also considered p+A and A+A collisions simultaneously for the first time and introduced more systematic treatments about questions such as surrogate models and principal component analysis. The work of the DukeQCD group set the field on a firm foundation.

These efforts have inspired numerous others, most prominently the JETSCAPE Collaboration's study of the soft sector [45, 46], with which I was involved. This effort, along with the thesis of D. Everett [47], represents the current state-of-the-art in heavy ion collisions. This analysis used Bayesian model selection and Bayesian model averaging to assess previously unquantified sources of uncertainty for the first time. This analysis also resolved several outstanding or previously-unexplored issues present in previous analyses, such as the treatment of the $\sigma$ meson. The application of Bayesian tools to heavy ion collisions beyond the soft sector is also becoming increasingly common, c.f. [48, 49].

Other soft-sector Bayesian efforts have been published recently, such as [50–53], and have been primarily focused on increasing statistical power in model calculations or incorporating additional observables. Only one of these studies has explicitly consid-

ered the use of more advanced observables, such as symmetric cumulants, in constraining the transport properties of QGP [53]. Despite different levels of complexity, all of these have used simplified parametric models of the initial state without microscopic physics. A parametric model of the initial state, T$_R$ENTo [37, 42], has been used to extract physically relevant and meaningful quantities with apparent success in all of these studies. This is typically followed by a parametrized set of dynamics, such as freestreaming, resulting in sharp discontinuities at the transition to hydrodynamics in addition to other systematic, unquantified uncertainties [2].

## 1.4 PRESENT INVESTIGATION

This thesis is in support of the ongoing campaign of model building and systematic model-to-data comparison. It introduces a structured workflow for Bayesian modeling to heavy ion collisions for the first time. I will further extend the literature by using a physically-motivated microscopic physics model in a comprehensive Bayesian study – IP-Glasma [54]. IP-Glasma uses information from deep inelastic scattering (DIS) experiments to constrain the gluon saturation scale before evolving gluon gauge fields with the classical Yang-Mills (CYM) equations. Further details of the IP-Glasma model are provided in the following chapter. I will use the IP-Glasma model for the pre-equilibrium stage, allowing for a consistent theoretical treatment of both initial energy deposition and pre-equilibrium evolution to the onset of hydrodynamics. I also demonstrate novel contributions to computer experiment design and statistical learning in heavy ion collisions to more efficiently and systematically explore the parameter space of this study.

With these extensions, I proceed to embed the hybrid model into a Bayesian study of the quark-gluon plasma to systematically test a CGC-based model for the first time. I do so with a focus on Pb-Pb collisions at the LHC with $\sqrt{s_{NN}} = 2.76$ TeV. I first demonstrate constraint, and a new total covariance, in the parameter space of the IP-Glasma model. I demonstrate otherwise that the hybrid model with an IP-Glasma initial state can be constrained by final-state observables and is self-consistent, *i.e.* it can recover known inputs. I then proceed to perform an 11-dimensional study of the parameter space of the full hybrid model using final state observables believed

to have constraining power on the temperature-dependence of the shear and bulk viscosity. With this study, I will constrain the overall energy scale as well as the onset of hydrodynamics, the particlization temperature, and the specific shear and bulk viscosities. Finally, I incorporate a transfer learning approach [4] to investigate similar models in a computationally-efficient way.

The goal of this work is to more accurately describe the state of knowledge of the properties of strongly-interacting matter in heavy ion collisions by using the best tools and information available. To do this, realistic physical models and advanced statistical techniques are applied in order to precisely, but accurately, describe physical constraint and modeling uncertainty. Critically, this goes beyond a model fitting exercise where a specific parameter "tune" is the result: the goal of this thesis is to rigorously improve the understanding of the strongly-interacting matter produced in heavy ion collisions through systematic comparison of physical models to experimental measurements. The result of this work is the state of knowledge *posterior* to comparison of model calculations to experimental results. This result may also be used to improve other studies that utilize a hybrid model of the soft sector of heavy ion collisions and the best estimates of its properties to study other phenomena. For example, this will provide a new background for both jet physics and photon/dilepton studies that rely on a calibrated soft sector.

The thesis is organized as follows: Part I introduces the physical foundations of heavy ion physics and the models used for the soft sector (Chapter 2). In Part II, the statistical framework is introduced in Chapter 3 and demonstrated in a pedagogical example. Detailed numerical comparisons between the current state of the art Bayesian and IP-Glasma studies, followed by design considerations and advances are shown in Chapter 4. Chapter 5 investigates the self-consistency and sensitivity of IP-Glasma as the pre-equilibrium stage. Chapter 6 performs the full 11-dimensional study. In the final chapter, I discuss the results from the totality of the thesis, conclude, and describe the outlook for future work.

### 1.4.1 *Note on conventions*

The mostly negative Minkowski metric signature $(+,-,-,-)$ is used as well as units where $\hbar = c = k_B = 1$ unless otherwise noted. Here, $c$ is the speed of light, $\hbar$ is the reduced Planck's constant, and $k_B$ is the Boltzmann constant. Reference values for these constants are $c = 299792458\ \mathrm{ms}^{-1}$, $\hbar = (6.62607015 \times 10^{-34})/2\pi\ \mathrm{Js}$, and $k_B = 1.380649 \times 10^{-23}\ \mathrm{JK}^{-1}$ and they are a subset of the defining constants of the International System of Units (SI) [55]. These units allow for straightforward conversion between time, space, energy, and temperature. For example, $\hbar c = 0.19733\ \mathrm{GeV\ fm} = 1$ in these units. The standard notation $\beta = 1/T$ is also used throughout. Oversubscription of notation is somewhat unavoidable, but is disambiguated in the body of the text.

# 2

# MODELING THE SOFT SECTOR

Now that the broad physics of heavy ion physics has been introduced, it is time to introduce each model in turn. As seen in Fig. 1.4, there are three broad stages of heavy ion collisions:

1. Initial energy deposition and pre-equilibrium dynamics,
2. Relativistic viscous hydrodynamics, and
3. Hadronic interactions.

In the first stage, the nuclei collide, deposit energy, and are driven toward (local) equilibrium by some dynamics in this extremely energetic material. In the following stage, the material interacts hydrodynamically as it approaches local equilibrium. Next, the mean free path of individual particles becomes long enough that the material ceases to be hydrodynamic, but interacts more like a gas of particles. The hydrodynamic medium transitions into particles via "particlization", which remains an area of active research and a source of theoretical uncertainty. The resulting hadron gas is evolved to late times where they can be compared to experimental measurements. In this chapter, I will elucidate details of each stage as well as the model I use to describe it. However, this space is insufficient to provide a detailed technical report on each model. For such insight, I refer to relevant literature.

## 2.1 PRE-EQUILIBRIUM DYNAMICS – IP-GLASMA

IP-Glasma is a QCD-based pre-equilibrium (initial state) model for heavy ion collisions [28, 54, 56, 57]. It incorporates small momentum fraction gluon saturation using the Impact Parameter SATuration (IPSAT) model [58] and then evolves the resuling semi-classical gluon fields [59–62]. IP-Glasma was first implemented in [54] and in an improved computational formulation by [28, 57].

The theoretical underpinning of IP-Glasma is the Color Glass Condensate (CGC) framework. This framework is an effective theory of QCD in the limit where gluons dominate a system. Color refers to the three color charges in QCD. In the CGC framework, the boosted partons that source the gluon field experience time dilation and are "frozen" during the interactions between the two nuclei, but are of course still dynamic. The epithet "condensate" arises as the phase space density of gluons is high and these gluons behave coherently over length scales proportional to their saturation scale, $Q_s$, discussed further below.

CGC's connection to QCD becomes transparent when comparing the CGC and QCD actions. The QCD action is

$$S_{QCD} = \int d^4x \left( -\frac{1}{4} F^a_{\mu\nu} F^{\mu\nu a} + \bar{\psi}_f \left( i \not{D} - m_f \right) \psi_f \right), \tag{2.1}$$

where $f$ is an index over quark flavors, $a$ is an index 1-8 over colors, and Feynman's slash notation has been used. Once again, repeated indices denote summation by convention. $F_{\mu\nu}$ is the field strength tensor,

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_{strong} f^{abc} A^b_\mu A^c_\nu, \tag{2.2}$$

and the covariant derivative is

$$D_\mu = \partial_\mu - i g_{strong} A^a_\mu t^a \tag{2.3}$$

where $g_{strong}$ is the strong coupling constant. The CGC action is the QCD action without dynamical quarks [59],

$$S_{CGC} = \int d^4x \left( -\frac{1}{4} F^a_{\mu\nu} F^{\mu\nu a} + J^{\mu a} A^a_\mu \right). \tag{2.4}$$

The first term is identical between QCD and the CGC, but differences arise in the second term. The CGC separates fast and slow degrees of freedom: the fast partons (quarks) act as sources $J^{\mu a}$ for the slow parton (gluon) gauge fields $A^a_\mu$. The separation between fast and slow partons is determined using the momentum fraction $x = \frac{k}{P}$ where $k$ is the momentum of an individual parton while $P$ is the total momentum of the nucleon.

The CGC is an appropriate effective theory for QCD in heavy ion collisions because at the relativistic velocities at which the collisions occur, the system is *very nearly* only

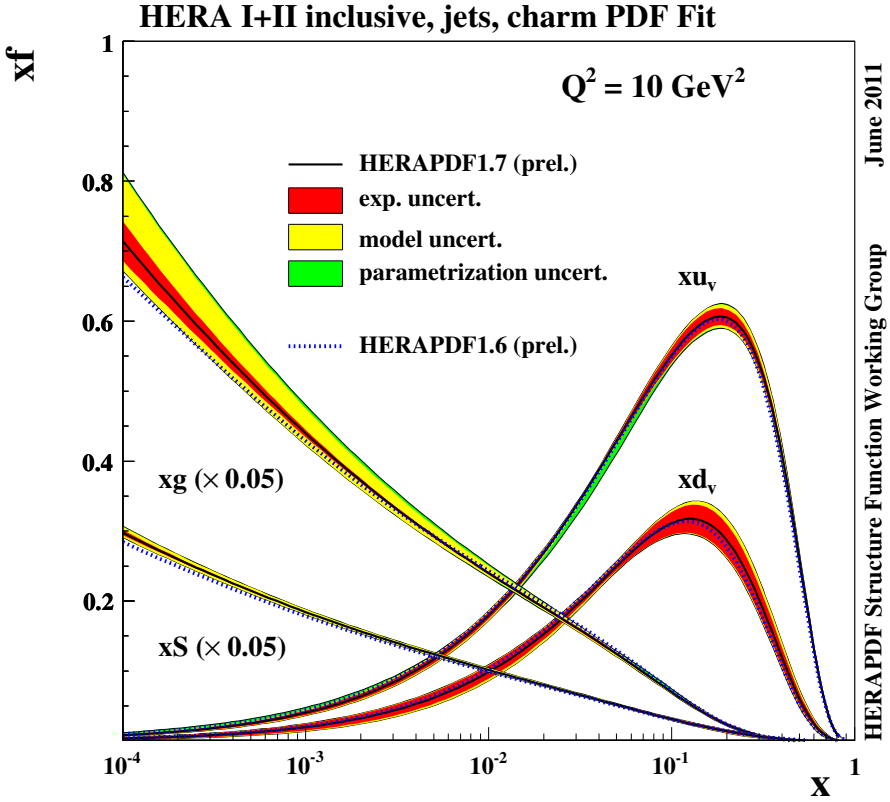


Figure 2.1: Parton distribution functions (PDFs) from HERA, adapted from [64]. $xu_v$ denotes the up quark PDF, $xd_v$ the down quark PDF, $xS$ the sea quark PDF, and $xg$ the gluon PDF. Note that the gluon and sea quark PDFs have been scaled dramatically.

gluons. Using the CMS fit [63], $x \approx 10^{-4}$ at LHC energies, which corresponds to the very lefthand margin of Fig. 2.1, where gluons dominate the momentum fraction.

At the LHC, Pb+Pb collisions currently take place at either $\sqrt{s_{NN}} = 2.76$ or 5.02 TeV. To determine the Lorentz factor,

$$\gamma = \frac{E}{m} = \frac{\sqrt{s}}{2m}. \tag{2.5}$$

For 2.76 TeV collisions, this yields $\gamma \approx 1470$ for protons and neutrons while for 5.02 TeV, this yields $\gamma \approx 2510$. A lead nucleus has an approximate radius of 7 fm, thus the length contraction of the nuclei is

$$R_{\text{contracted}} = \frac{2R_{Pb}}{\gamma} = \frac{14\text{fm}}{1470} \approx 0.009\text{fm}. \tag{2.6}$$

It follows from this that it is reasonable to approximate the nucleus as a 2-dimensional pancake. The CGC framework assumes that the colliding nuclei are perfectly frozen,

*i.e.* moving with the speed of light. For the purposes of heavy ion collisions with such a large $\gamma$, it is clear that treating the fermions in the nucleus as frozen sources is not far from the truth because of time dilation with the same $\gamma$.

Now that the use of the CGC framework has been motivated, it is time to turn to the details of IP-Glasma. This section condenses, but broadly follows, the discussion in [57] and references therein.

As the nuclei are "frozen" and moving at the speed of light, they are on the light cone. As a result, the color currents can only be delta functions moving on the light cone,

$$J^{\mu a}(x) = \rho^a_{(A)}(\mathbf{x}_\perp)\,\delta^{\mu+}\delta\left(x^-\right) + \rho^a_{(B)}(\mathbf{x}_\perp)\,\delta^{\mu-}\delta\left(x^+\right) \tag{2.7}$$

where A and B denote the colliding nuclei. The charge density distribution is sampled stochastically and cannot be determined analytically in event-by-event collisions.

The next piece of the puzzle comes in the form of the Classical Yang-Mills (CYM) field and equations of motion. In electromagnetism, there is a scalar charge. The Yang-Mills equations generalize Maxwell's Equations from a scalar charge to a vector of charges, *i.e.* multiple charges. QCD has three color charges and the equations of motion can be derived by varying the action, resulting in

$$[D_\mu, F^{\mu\nu}]_a = J^\nu_a. \tag{2.8}$$

This source is the same as the source of the color currents along the light cone. This is familiar from electromagnetism (c.f. [65]), but with an additional color index $a$. For details of the derivation as well as technical implications, the interested reader is referred to the seminal McLerran-Venugopalan model [66–68], early IP-Glasma work [54], and relevant theses [57, 69].

The gauge fields ($A^a_\mu$ in the action and field strength tensor) satisfy a two-dimensional Poisson equation where gauge fields go to zero as $\mathbf{x}_\perp$ goes to infinity.

$$-\nabla^2_\perp A^\pm_C = J^\pm = g_{strong}\rho(x^\mp, \mathbf{x}_\perp) \tag{2.9}$$

where subscript C denotes that the field is in the covariant gauge. Thus, the gauge fields of a nucleus prior to collision may be constructed. Of course, in heavy ion collisions, the quantity of interest is the gauge field of two 2D colliding nuclei. In this case, two sources moving on the $x^\pm$ axes are summed,

$$J = g_{strong}\rho_A\left(x^-, \mathbf{x}_\perp\right) + g_{strong}\rho_B\left(x^+, \mathbf{x}_\perp\right) \tag{2.10}$$

and $A^{\pm} = 0$,

$$A^i = \theta\left(x^-\right)\theta\left(-x^+\right)A^i_A\left(\mathbf{x}_\perp\right) + \theta\left(x^+\right)\theta\left(x^-\right)A^i_B\left(\mathbf{x}_\perp\right) \tag{2.11}$$

satisfy the equations of motion prior to the collision where $A_A$ and $A_B$ are the non-Abelian gauge fields of nuclei A and B prior to the collision. The discontinuity of the fields across the nuclei gives the charge density, in which the physical information is embedded. This can be written as

$$\nabla_i A^i_{A(B)} = \rho_{A(B)}(\mathbf{x}_\perp), \tag{2.12}$$

which is simply Gauss' Law. Next, the initial condition is uniquely determined by the gauge fields before the collision in an initial value problem. This relies explicitly on the charge distributions. Once the initial condition has been determined, the gauge fields can be evolved using the equations of motion.

Solving the Poisson equation to determine the pre-collision fields is therefore all that stands in the way of a theoretical solution to implement computationally.

Beginning with the Poisson equation, but suppressing $x^{\pm}$ and the color index yields

$$-\nabla^2_\perp A^{+(-)}_a = g_{strong}\rho_a(x^{-(+)}) \tag{2.13}$$

and the Fourier transform may be written

$$A(\mathbf{k}_\perp) = \frac{g_{strong}\rho(\mathbf{k}_\perp)}{\mathbf{k}^2_\perp}. \tag{2.14}$$

This equation can then be solved and the inverse Fourier transform taken to find the gauge field at any position in the transverse plane. These gauge fields in position space are the initial conditions for the pre-equilibrium evolution. The path-ordered exponential of this gauge field gives the Wilson Line [70]

$$V\left(\mathbf{x}_\perp\right) = Pe^{i\int dx^- A^+\left(x^-,\mathbf{x}_\perp\right)}. \tag{2.15}$$

Discretizing this equation and connecting it to a gauge transformation yields the initial gauge fields. These are the fields that are then evolved using the CYM equations as the pre-equilibrium phase of the collision. Implementing this process on a computational lattice is a technically-involved process which can only be given a detailed treatment in a dedicated work. However, this brief introduction should serve as

a conceptual introduction to CGC-based pre-equilibrium models and demonstrates the degree of first-principles theory involved compared to parametric *ansätze* used in other works.

A basic recipe for IP-Glasma is given below, adapted from [57] for a (2+1)D study:

1. Determine nuclei positions. This can be done by sampling from a Woods-Saxon distribution, but this work exploits pre-calculated nucleon configurations that include nucleon-nucleon correlations [71].

2. Sample impact parameter from

$$P(b)db = \frac{2b}{b_{max}^2 - b_{min}^2} db \tag{2.16}$$

and shift nucleus $A(B)$ by $+(-)b/2$ in the $x$-direction.

3. Sum contributions from individual nucleons to construct the nuclear thickness function

$$T_{nucleus}(\mathbf{x}) = \sum_{i=1}^{A} \frac{e^{(\mathbf{x}-\mathbf{x}_i)^2/2B_G}}{2\pi B_G} \tag{2.17}$$

where $B_G$ denotes the size of the nucleon hot spots.

4. Compute the saturation scale $Q_s^2$ at midrapidity using IPSAT,

$$Q_s^2 = \frac{1}{r_s^2} = \frac{\pi^2}{N_c} \alpha_s(\mu^2(r^2)) x g(x, \mu^2(r^2)) T(b) \tag{2.18}$$

where $\mu$ is the energy scale of the interaction at which deep inelastic scattering takes place, $xg$ is the gluon PDF, $N_c$ is the number of colors and $Q_s \propto g_{strong}^2 \mu$; or, perhaps more intuitively, $\frac{Q_s}{g_{strong}^2 \mu} = const$. This proportionality must be estimated from data and is the parameter $\mu_{Q_s}$ throughout this thesis. Throughout this work, $N_c = 3$, corresponding to QCD. Constraining $\mu_{Q_s}$ using experimental data is one of the goals of this thesis, as is investigating simultaneous constraint on $g_{strong}$.

5. Sample the color charge density using

$$\langle \rho_{A(B)}^a(\mathbf{x}_\perp) \rho_{A(B)}^b(\mathbf{x}_\perp) \rangle = g_{strong}^2 \mu_{A(B)}^2(x, \mathbf{x}_\perp) \delta^{ab} \delta^2(\mathbf{x}_\perp - \mathbf{y}_\perp) \tag{2.19}$$

6. Solve the Poisson equation

$$\nabla_{\perp}^{2} A^{a} = \rho^{a} \tag{2.20}$$

This is simplified in 2D, since $A^{\eta}$ is zero at $\tau = 0$ fm [57]. As each spatial component results in 8 new colour gauge fields, the lack of an $\eta$ direction reduces the complexity of the problem.

7. Construct Wilson lines at midrapidity ($\eta = 0$) for a 2+1D simulation, as the boost-invariant approximation is most accurate at midrapidity.

8. Construct pure gauge fields for each nucleus

$$A_{\mu}^{A(B)} = -\frac{i}{g_{strong}} V^{A(B)} \partial_{\mu} V^{A(B)\dagger} \tag{2.21}$$

9. Solve Gauss' Law. In 2D, this becomes trivial because derivatives in the $\eta$ direction vanish, leaving $[D_i, E_i] = 0$, whose solution is simply $E_i = 0$ and the initial transverse fields start at 0.

10. Evolve using the sourceless classical Yang-Mills equations of motion

$$[D_{\mu}, F^{\mu\nu}] = 0. \tag{2.22}$$

Note that as the sources for the initial gauge fields travel at the speed of light, they are now causally disconnected from the evolution and thus cannot source new fields.

11. Construct the energy-momentum tensor $T^{\mu\nu}$

$$T^{\mu\nu} = -g^{\mu\alpha} g^{\nu\beta} g^{\gamma\delta} F_{\alpha\gamma} F_{\beta\delta} + \frac{1}{4} g^{\mu\nu} g^{\alpha\gamma} g^{\beta\delta} F_{\alpha\beta} F_{\gamma\delta} \tag{2.23}$$

12. Diagonalize $T^{\mu\nu}$

$$T_{\nu}^{\mu} u^{\nu} = \epsilon u^{\mu} \tag{2.24}$$

13. Separate $T^{\mu\nu}$ into ideal and viscous components to initialize the hydrodynamic evolution

## 2.2 RELATIVISTIC VISCOUS HYDRODYNAMICS – MUSIC

Hydrodynamics is an effective theory of long-wavelength modes. Practically, this means that the evolution of a collection of differential elements can be described not by tracking microscopic particles, but considering their long-wavelength (or spatially-coarse) collective motion. Analogously, to model a hurricane or storm front, it is not necessary (or even relevant) to model the behavior of every constituent water droplet or molecule of air. Instead, the collective dynamics at a much larger scale reveal the physics of interest.

In this section, I will introduce the basic equations of relativistic viscous hydrodynamics broadly following a similar discussion in [30]. However, the equations of relativistic viscous hydrodynamics can describe any hydrodynamic evolution. To connect this to heavy ion physics explicitly, I will also discuss details of the equation of state. With this connection, I motivate the functional form of the shear and bulk viscosity used in this thesis.

It is traditional to proceed by introducing the energy-momentum tensor order-by-order in the deviation from equilibrium. I will begin by introducing the ideal energy-momentum tensor and proceed to add dissipative (viscous) elements.

The ideal symmetric energy-momentum tensor is

$$T^{\mu\nu}_{ideal} = -Pg^{\mu\nu} + wu^{\mu}u^{\nu} \tag{2.25}$$

where $P$ is the temperature-dependent pressure (the functional dependence on temperature $T$ is suppressed for clarity), $s = dP/dT$ is the entropy density, $w = Ts$ is the enthalpy density, and $g^{\mu\nu}$ is the metric tensor. This is also commonly written in the equivalent form

$$T^{\mu\nu}_{ideal} = (\epsilon + P)u^{\mu}u^{\nu} - Pg^{\mu\nu} \tag{2.26}$$

using the relation $Ts = \epsilon + P$. In this thesis, I use the Landau frame[1] where $u^\mu(x)$ is the velocity of energy transport. This is also known as enforcing the eigenvalue equation

$$u_\mu T^{\mu\nu} = \epsilon u^\nu. \tag{2.27}$$

The energy-momentum tensor is next extended to first order in deviation from equilibrium, where shear and bulk viscosity are encountered for the first time. This is called "Navier-Stokes" hydrodynamics. Because of the choice of the Landau frame, $u_\mu \Delta T^{\mu\nu} = 0$, where $\Delta T^{\mu\nu}$ is the first order deviation from the ideal energy-momentum tensor. The full energy-momentum tensor thus becomes

$$T^{\mu\nu}_{\text{N.S.}} = T^{\mu\nu}_{\text{ideal}} + \Delta T^{\mu\nu} \tag{2.28}$$

$$\Delta T^{\mu\nu} = \eta \left( D^\mu u^\nu + D^\nu u^u - \frac{2}{3} \Delta^{\mu\nu} \partial_\rho u^\rho \right) + \zeta \Delta^{\mu\nu} \partial_\rho u^\rho \tag{2.29}$$

with the projection tensor

$$\Delta^{\mu\nu} = g^{\mu\nu} - u^\mu u^\nu \tag{2.30}$$

and the covariant derivative

$$D_\mu = \partial_\mu - u_\mu u^\beta \partial_\beta, \tag{2.31}$$

both of which are normal to the flow $u^\mu$.

In the local rest frame, the change in the entropy current $s^\mu = s u^\mu$ is given by

$$\partial_\mu s^\mu = \frac{\eta}{2T} \left( \partial^i u^j + \partial^j u^i - \frac{2}{3} \delta^{ij} \nabla \cdot \mathbf{u} \right)^2 + \frac{\zeta}{T} (\nabla \cdot \mathbf{u})^2 \tag{2.32}$$

The non-decrease of entropy must be respected. As a result, the shear viscosity $\eta$ and the bulk viscosity $\zeta$ must be non-negative.[2]

Thus far, the focus has been on the gradient expansion to the first order. However, the first order expansion is insufficient for relativistic fluids because it is possible for acausal propagation of information via superluminal flow [72], although recent

1 An alternate frame, the Eckart frame, considers the velocity of baryon number flow. In this thesis, the system's conserved quantum numbers are set to be identically 0 and the Eckart frame cannot be applied.

2 As $s$ is an overall multiplicative factor in the entropy current, the same requirement of non-negative specific shear ($\eta/s$) and bulk ($\zeta/s$) viscosity naturally follows.

progress has been made in the search for causal theories [73–76]. In this thesis, second-order viscous hydrodynamics is used.

The first derivation of second-order viscous hydrodynamics was by Müller, Israel, and Stewart [77, 78]. Müller-Israel-Stewart adds transport coefficients in the form of shear and bulk relaxation times, as well as two additional hydrodynamic equations,

$$\begin{aligned} \tau_\pi \dot{\pi}^{\langle\mu\nu\rangle} + \pi^{\mu\nu} &= 2\eta\sigma^{\mu\nu} - \tfrac{4}{3}\pi^{\mu\nu}\theta, \\ \tau_\Pi \dot{\Pi} + \Pi &= -\zeta\theta - \tfrac{2}{3\tau_\Pi}\Pi\theta. \end{aligned} \tag{2.33}$$

As before, $\eta$ and $\zeta$ are the shear and bulk viscosities, respectively. $\theta = \partial_\mu u^\mu$ is the expansion rate and $\sigma^{\mu\nu} = \partial^{\langle\mu} u^{\nu\rangle}$ is the Navier-Stokes tensor. $\Pi$ is the bulk pressure and $\pi^{\mu\nu} = 2\eta\sigma^{\mu\nu}$ is the shear tensor in the Navier-Stokes limit. Finally, the shear and bulk relaxation times are denoted by $\tau_\pi$ and $\tau_\Pi$. The dot above a quantity indicates a proper time derivative $\dot{\Pi} = u^\mu \partial_\mu \Pi = d\Pi/d\tau$ while the brackets $\langle\rangle$ refer to the traceless part of the symmetrized tensor, $\dot{\pi}^{\langle\mu\nu\rangle} = \frac{1}{2}\left(\dot{\pi}^{\mu\nu} + \dot{\pi}^{\nu\mu} - \frac{2}{3}\Delta^{\mu\nu}\dot{\pi}^{\alpha}_{\alpha}\right)$. These shear and bulk relaxation times characterize the timescale on which shear stress tensor and bulk pressure approach their first-order solutions, ensuring causal solutions [79]. These ensure that superluminal propagation of information does not occur as, for example, in the absence of shear viscosity, the Navier-Stokes shear tensor decays instantaneously while the shear relaxation time ensure it decays proportional to $\exp(-\tau/\tau_\pi)$. Recent works have further investigated and developed second-order hydrodynamic formulations specifically to constrain when hydrodynamics becomes acausal [80, 81].

In this thesis, a variant of Müller-Israel-Stewart hydrodynamics is chosen for the hydrodynamic simulation. This kinetic-theory-based formulation by Denicol, Niemi, Molnar, and Rischke – referred to by their initials as DNMR hydrodynamics [82] – is more commensurate with how modern physical theories are constructed by consideration of all allowable symmetries. It is derived from the relativistic Boltzmann equation,

$$p^\mu \partial_\mu f_{\mathbf{p}} = C[f]. \tag{2.34}$$

From left to right, $p^\mu$ is the four-momentum, $f_p$ is the distribution function of a single particle, and finally $C[f]$ is the classic Boltzmann collision kernel, where DNMR hydrodynamics includes $2 \to 2$ collisions. Terms up to the second order in Knudsen number $k_n = \frac{\lambda}{L}$ and shear/bulk inverse Reynolds numbers are included. Small

Knudsen number quantifies the ratio between microscopic ($\lambda$) and macroscopic (L) length scales, ensuring that a long-range effective theory like hydrodynamics applies. Small inverse Reynolds number quantifies the ratio of viscous to equilibrium forces and guarantees the system is not far from equilibrium, so a gradient expansion is still appropriate.

Specifically, DNMR second-order relativistic viscous hydrodynamics adds the following equations to those of Müller, Israel, and Stewart:

$$\tau_\pi \dot{\pi}^{\langle\mu\nu\rangle} + \pi^{\mu\nu} = 2\eta\sigma^{\mu\nu} - \frac{4}{3}\tau_\pi \pi^{\mu\nu}\theta + \frac{9}{70P}\pi_\alpha^{\langle\mu}\pi^{\nu\rangle\alpha} - \frac{10}{7}\tau_\pi \pi_\alpha^{\langle\mu}\sigma^{\nu\rangle\alpha} + \frac{6}{5}\tau_\pi \Pi\sigma^{\mu\nu} \quad (2.35)$$

and

$$\tau_\Pi \dot{\Pi} + \Pi = -\zeta\theta - \frac{2}{3}\tau_\Pi \Pi\theta + \frac{8}{5}\left(\frac{1}{3} - c_s^2\right)\tau_\Pi \pi^{\mu\nu}\sigma_{\mu\nu} \quad (2.36)$$

where the relaxation times are from the collision kernel and depend on underlying interactions, providing the timescale by which higher-order dissipative modes decay toward the Navier-Stokes limit. In this thesis, the relaxation times are fixed by ratios to viscosity

$$\tau_\pi = \frac{5\eta}{\epsilon + P} \quad (2.37)$$

$$\tau_\Pi = \frac{\zeta}{\epsilon + P}\frac{1}{14.55\,(1/3 - c_s^2)^2}, \quad (2.38)$$

where the constants are obtained through an expansion in the relaxation time approximation [83]. DNMR hydrodynamics is implemented numerically in the code known as MUSIC (*MUSCl for Ion Collisions*). MUSIC uses the Kuganov-Tadmor (KT) algorithm for calculating spatial derivatives and Heun's method for temporal derivatives. It is well-suited to application in heavy ion collisions as the numerical viscosities are small and the KT algorithm is robust to discontinuities and shock waves. For further details of MUSIC, including a simple example of the algorithm, see [12, 30].

The transport coefficients of interest in this study, notably $\eta/s$ and $\zeta/s$, characterize the first-order deviation from equilibrium in the hydrodynamic medium. This begs the question, "deviation from *what* equilibrium?" This is characterized by the Equation of State (EoS), which is necessary to close the system of hydrodynamic equations above. Equations of state relate the thermodynamic properties (state variables) such as pressure, temperature, or entropy density under given conditions. This is where

the specific material properties of QCD matter inform the hydrodynamic stage and as a result, must be constructed to be consistent with the model choices. The EoS at high temperature is matched to lattice calculations [84]. At low temperature, the EoS matches that of the particle list used in the hadronic cascade, which ensures that the EoS is continuous across the transition between the two stages. The matching between the high and low temperature results must be done in a manner consistent with the QCD crossover, *i.e.* it must have a smooth crossover rather than a sharp phase transition at vanishing baryochemical potential.[3] Attempts to constrain the equation of state directly from hadronic observables have shown promise, but as of yet still have significant remaining uncertainty [91, 92]. The difficulty in constraining the EoS largely arises from the relative insensitivity of hadronic observables. Active learning techniques are also being applied to the efforts to characterize the equation of state [93]. The EoS used in this work smoothly connects the HotQCD calculation [94] at high temperatures to a list of stable resonances at low temperatures and matches that of [45] as shown explicitly in Fig. A.1 and the code that produced it is publicly available with the default parameters [95]. More on constructing a QCD equation of state as well as detailed comparisons between various candidates may be found in [96–98] with a Bayesian study focused on constraining the EoS in [36].

## 2.3 PARTICLIZATION – IS3D

As the fireball of quark-gluon plasma expands, it cools and the mean free path between particles becomes large. The relevant degrees of freedom in hydrodynamics transitions from being quarks and gluons to being hadrons and, as the medium continues to cool, it ceases to interact hydrodynamically. Once hydrodynamics is no longer an effective theory of long-wavelength modes, it is necessary to transition to a discrete particle description of the strongly-interacting matter. This transition is known as "particlization." This is a change in the degrees of freedom from those of hydrodynamics to particles. In practice, it becomes a change of language in how

3 Discussions of non-zero baryochemical potential are beyond the scope of this thesis, but are a vibrant field which features a search for a hypothesized QCD critical point [22, 85–90].

the material is described and occurs in a regime where both hydrodynamics and a discrete particle kinetic theory are applicable.

To particlize a hydrodynamic medium, the current best practice is to draw a surface at constant temperature, energy density, or entropy - these choices are equivalent in the case of zero baryochemical potential, which this thesis strictly respects. This temperature is called the switching, or particlization, temperature.[4] Once this surface has been drawn, particles can be sampled stochastically, respecting energy and momentum conservation *on ensemble average*. This means that to ensure that the sampled distribution converges to the true distribution of particles, momenta, etc.. it is necessary to *over*sample this surface. This is either done a fixed number of times (often 100 to 300 times), or until a sufficient number of particles has been sampled. This ensures that for large surfaces, where individual fluctuations balance, fewer events are oversampled, reducing computational cost, while for small surfaces where individual samples from the surface have greater variation, more samples are taken and conservation laws are respected on ensemble average.

The way the sampling is performed is via the Cooper-Frye prescription [101], in this thesis implemented in iS3D [102]. Given an isothermal (or isentropic, etc.) hypersurface $\Sigma$ with normal vector $\sigma_\mu(x)$, the invariant momentum spectra of a particle species $i$ with degeneracy $g_i$ is

$$E\frac{dN_i}{d^3p} = \frac{g_i}{(2\pi)^3}\int_\Sigma f_i(x,p)p_\mu d\sigma^\mu(x) \tag{2.39}$$

where $f_i(x,p)$ is the phase-space distribution. This distribution function reproduces the energy-momentum tensor of hydrodynamics at the particlization surface,

$$T^{\mu\nu}(x) = \sum_i \frac{g_i}{(2\pi)^3}\int \frac{p^\mu p^\nu f_i(x,p)}{E} d^3p. \tag{2.40}$$

Here, $f_i(x,p)$ is species-specific, representing Bose-Einstein statistics for particles with integer spin and Fermi-Dirac statistics for particles with half-integer spin.

The out-of-equilibrium nature of the material generates interesting physics, but presents significant mathematical challenges. If at the time of particlization the hydrodynamic medium was in equilibrium, the choice of the distribution function would

4 Interesting ongoing work is investigating chemical freezeout. For more on chemical freezeout, see [99, 100].

simply be the equilibrium form. The rest frame velocity and temperature would be fixed by the hydrodynamic velocity and the energy density in the local rest frame. The medium is not in equilibrium and consistency between the kinetic description of particles and viscous hydrodynamics must be attempted. The shear and bulk stress – $\pi^{\mu\nu}$ and $\Pi$ – equate to deviations of the microscopic distributions and yields from the equilibrium ones. While many different methods to capture these dynamics exist, there are two main classes: linear and exponentiated viscous corrections. A recent work has even extended the maximum entropy ideas of statistics and information theory to the problem [103]. This thesis will exclusively consider the more established linear viscous corrections as these are systematically-improvable. These are Grad's 14-moment approximation [104] and the linear Chapman-Enskog expansion in the relaxation time approximation [105]. The following discussion closely follows [102] and summarizes key parts for completeness.

The distribution function for a fluid out of local equilibrium may be separated as

$$f_i(x,p) = f_{eq,i}(x,p) + \delta f_i(x,p) \tag{2.41}$$

where $f_{eq,i}(x,p)$ is the equilibrium distribution function (Bose-Einstein or Fermi-Dirac for different particle species) and $\delta f_i(x,p)$ is the non-equilibrium correction.[5] An advantage of mapping between a hydrodynamic description and a kinetic theory (particle) description is that information from hydrodynamics can be used to provide some information for the non-equilibrium correction. The net baryon current $J^\mu_B$ and energy-momentum tensor $T^{\mu\nu}$ from viscous hydrodynamics translate to the first and second moments of the distribution,

$$J^\mu_B(x) = \sum_i b_i \int_p p^\mu f_i(x,p) \tag{2.42}$$

$$T^{\mu\nu}(x) = \sum_i \int_p p^\mu p^\nu f_i(x,p) \tag{2.43}$$

where the shorthand notation $\int_p \equiv \int \frac{d^3p}{(2\pi)^3 E_p}$ has been used.

5 An explicit assumption of viscous corrections is that they are small. However, they still modify the momentum dependence of observables as shown in [102]. Studies explicitly using momentum-dependence for constraint should exercise caution about overly-certain conclusions, especially as corrections grow larger.

Unfortunately, the separation of the distribution function into the equilibrium contribution and a viscous correction, despite the constraints from matching to hydrodynamics, does not fully specify the momentum-dependence of $\delta f_i(x,p)$. This means that the choice of correction remains a modeling choice with inherent ambiguity that can have a notable impact on hadronic observables [106]. To constrain further, the reasonable assumption is made that hydrodynamics and relativistic kinetic theory are simultaneously applicable at the transition between them.

Linearized viscous corrections linearize the correction $\delta f_i$ in the shear stress tensor ($\pi^{\mu\nu}$), bulk viscous pressure $\Pi$, and baryon diffusion current $V_B^\mu$,

$$\begin{aligned}\delta f_i(x,p) \approx c_{\pi,i}(x,p)p_{\langle\mu}p_{\nu\rangle}\pi^{\mu\nu}(x) + c_{\Pi,i}(x,p)\Pi(x) \\ + c_{V,i}(x,p)p_{\langle\mu\rangle}V_B^\mu(x)\end{aligned} \tag{2.44}$$

where $c_{\pi,i}$, $c_{\Pi,i}$, and $c_{V,i}$ are expansion coefficients with both spatial and momentum dependence. In this thesis, baryon diffusion not considered, but is included in this discussion for completeness. In linearized viscous corrections, the expansion coefficients are adjusted to exactly reproduce $J_B^\mu$ and $T^{\mu\nu}$.

Grad's 14-moment approximation expands the correction $\delta f_i(x,p)$ in momentum moments of the distribution function [104], only truncating at the level with terms involving $p^\mu$ and $p^\mu p^\nu$, *i.e.* at hydrodynamic order, which after some manipulation [78, 82, 107, 108] becomes

$$\begin{aligned}\delta f_i = f_{eq,i}\bar{f}_{eq,i}(c_T m_i^2 + b_n(c_B(u_\mu p^\mu) + c_V^{\langle\mu\rangle}p_{\langle\mu\rangle}) \\ + c_E(u_\mu p^\mu)^2 + c_Q^{\langle\mu\rangle}(u_\mu p^\mu)p_{\langle\mu\rangle} + c_\pi^{\langle\mu\nu\rangle}p_{\langle\mu}p_{\nu\rangle})\end{aligned} \tag{2.45}$$

where $\bar{f}_{eq,i} \equiv 1 - g_n^{-1}\Theta_n f_{eq,n}$ and the brackets $\langle\cdot\rangle$ again denote the traceless part of the symmetrized tensor. Here, $g_n$ is the spin degeneracy $g_n = 2s_n + 1$ and $\Theta_n = (-1, 1)$ to account for quantum statistics of a particle (Bose-Einstein and Fermi-Dirac). The coefficients $c$ are adjusted to ensure that the viscous correction to the distribution does not contribute to the energy or net baryon density and reproduces the shear stress tensor, bulk viscous pressure, and baryon diffusion current that were linearized. Their forms are lengthy and may be found in [102, 104].

The Chapman-Enskog expansion is a gradient expansion around $f_{eq,i}$. The relaxation time approximation (RTA) is used for the collision term of the Boltzmann equation [109], which thus reduces to

$$p^\mu \partial_\mu f_i = -\frac{u_\mu p^\mu}{\tau_r} \delta f_i \tag{2.46}$$

where the relaxation time $\tau_r$ is momentum and species independent. Then expanding $f_i$ into its equilibrium component and correction and assuming hydrodynamic gradients are small in comparison to the relaxation time, a first order gradient correction for the thermal distribution may be derived [110],[6]

$$\delta f_i = -\frac{p^\mu \partial_\mu f_{eq,i}}{u_\mu p^\mu / \tau_r}. \tag{2.47}$$

This may then be expanded using conservation laws and Navier-Stokes relations to produce the final form of the viscous correction,

$$\begin{aligned} \delta f_i = & f_{eq,i} \bar{f}_{eq,i} \left[ \frac{\Pi}{\beta_\Pi} \left( b_i \mathcal{G} + \frac{(u_\mu p^\mu)\mathcal{F}}{T^2} + \frac{(-p_\mu \Delta^{\mu\nu} p_\nu)}{3(u_\mu p^\mu) T} \right) \right. \\ & \left. + \frac{V_B^\mu p_{\langle\mu\rangle}}{\beta_V} \left( \frac{n_B}{\mathcal{E} + \mathcal{P}_{eq}} - \frac{b_i}{(u_\mu p^\mu)} \right) + \frac{\pi_{\mu\nu} p^{\langle\mu} p^{\nu\rangle}}{2\beta_\pi u_\mu p^\mu T} \right]. \end{aligned} \tag{2.48}$$

Here, $V_B^\mu = \kappa_B \Delta^{\mu\nu} \partial_\nu \alpha_B$ is the baryon diffusion current, $\kappa_B$ is the baryon diffusion coefficient, and $\alpha_B = \mu_B / T$ is the ratio of baryon chemical potential and temperature. The conservation of net baryon number is $\dot{\alpha}_B = \mathcal{G}\theta$, conservation of energy is $\dot{T} = \mathcal{F}\theta$, and conservation of momentum is $\dot{u}^\mu = \Delta^{\mu\nu} \partial_n u \ln T$. The ratios of transport coefficients to relaxation times are denoted $\beta_\pi = \eta/\tau_r$, $\beta_\Pi = \zeta/\tau_r$ and $\beta_V = \kappa_B/\tau_r$. The remaining quantities, $\mathcal{F}$ and $\mathcal{G}$ are defined by thermal integrals and may be found in [102, 110].

The uncertainty in the treatment of the $\sigma$ meson presents challenges to sampling it in the particlization stage. In particular, the uncertainty of the mass and decay width can significantly impact hadronic observables as explored in [45]. Additionally, the contribution of the $\sigma$ to isospin-averaged observables is minimal due to repulsion from the isotensor-scalar channel and including it spuriously increases pion yields [111]. The preferred practice is to omit it from sampling at particlization but allow it to be dynamically formed in the hadron gas phase so the $\pi - \pi$ cross-section remains well-described.

6 This equation demonstrates the competition between global expansion and local scattering in determining $\delta f_i$ [102].

## 2.4 HADRON GAS – SMASH

Once a list of particles has been sampled from a hydrodynamic hypersurface, they can be evolved using kinetic theory via the SMASH transport code [32]. The particles interact each other, scattering, decaying, and forming resonances. These are computed in SMASH using measured particle properties and channels [6] via a tower of coupled Boltzmann equations.

$$p^\mu \partial_\mu f_i(x,p) = C[f_i] \tag{2.49}$$

where $i$ is an index over species. Once again, $f_i(x,p)$ is species-specific, representing Bose-Einstein statistics for particles with integer spin and Fermi-Dirac statistics for particles with half-integer spin. This list of species is given in the SMASH documentation.[7] This thesis uses SMASH Version 1.8.

Inclusion of this stage makes physical sense: the particles sampled directly from the medium must travel to the detectors and do so in the presence of the rest of the hadron gas. Consequently, they must be allowed to interact and decay dynamically; not doing so would introduce a sharp discontinuity in the dynamics inconsistent with the assumption that this is a continuous system one is forced to model disjointedly. By using a dynamic hadron cascade, SMASH allows for a dynamical decoupling of the strongly-interacting medium dictated by the mean free path and density. This is consistent with physical intuition that the fireball should gradually decouple as the density drops and particles are less likely to collide.

This work will vary none of the parameters of SMASH, similar to [45, 46], which validated its usage. As this thesis does not vary parameters of this stage, the interested reader is referred to more comprehensive references for specific implementation of details, including determining collisions, implementing mean fields, and particle properties [32].

7 Available at smash-transport.github.io

## 2.5 MODELING AND CENTRALITY SELECTION

To calculate observables from particle lists, a large number of collisions must be simulated. Each event's properties can be calculated and then event-averaged quantities may be compared to experiment. Given a set of input parameters of interest to the model, the following procedure is followed:

1. Generate 2,500 IP-Glasma events with impact parameter restricted to between 0 and 13 fm,
2. Propagate each event through viscous hydrodynamics using MUSIC, saving a hypersurface at a specified switching temperature,
3. Particlize the information from the hydrodynamic hypersurface into a hadron gas using a specified model for the viscous corrections,
4. Propagate the hadron gas dynamically with scatterings and decays using SMASH.

This produces lists of particles and their properties that can be analyzed to reproduce calculations of observables from experiments. The analysis follows the centrality selection of [28], which reproduces Glauber centrality classes and cross-sections within experimental error. This method produces a cross-section of $\sigma_{PbPb} \sim 8.1b$, compatible with the ALICE result of $\sigma_{PbPb} = 7.7b \pm 0.1b(stat) \pm 0.5b(sys)$ [112]. The 100% centrality cutoff - a poorly-constrained idea in simulation - is treated as a parameter use to ensure the model is properly calibrated to the data and experimental effects. To bins events by centrality, the following procedure is followed:

1. Calculate the single-event properties such as the number of charged hadrons,
2. Take the ratio between two centrality bins' charged hadron multiplicities and compare to experiment,
3. Order all events by charged hadron multiplicity and use percentiles to define centrality classes,
4. If the ratio between the bins of the experimental charged hadron multiplicity is greater than the simulated ratio, add empty events (zeros) to the list of simulated charged hadron multiplicities and re-bin until the ratios match.

5. If the ratio between the bins of the experimental charged hadron multiplicity is less than the simulated ratio, drop events from the list until the ratio matches.

This is analogous to varying an energy cut-off that defines 100% centrality. Its ability to reproduce the Glauber cross section is remarkable as there are no nucleons in IP-Glasma: to reproduce the Glauber result, central nucleon positions must be added and participants counted. The range $0-13$ fm typically maps onto a range of $\sim 0-75\%$ centrality for Pb-Pb collisions at 2.76 TeV. The advantage of this method of centrality selection is that hydrodynamics is generally considered to break down in highly peripheral collisions. In minimum bias calculations, this corresponds to discarding a significant number of events at great computational cost.

## 2.6 OBSERVABLES

A precise understanding of various candidate observables is critical for extracting physics insight. It is not merely sufficient to treat them as black-box quantities, but rather the reader should appreciate the details so that it is clear how properties of the quark-gluon plasma may be revealed through careful, simultaneous study. This section describes the observables used in this thesis. I will refer to two classes of observable: *first generation observables* and *next generation observables*. This distinction is wholly my own, but will be a useful shorthand. This division is broadly made as the first generation observables are those first reported by experimental collaborations and those first used in Bayesian studies. The next generation observables are dominated by observables reported by experimental collaborations after the first generation or those included in a Bayesian study for the first time in this thesis. I by no means list all possible soft sector observables here, merely a selection that are used in this thesis and with which the interested reader must be conversant.

### 2.6.1 *First generation observables*

These broadly describe large-scale features of the fireball and add four-particle radial Fourier coefficients to the set of observables used in [45, 46] with the exception of

correlated momentum fluctuations. Charged hadron multiplicity, transverse energy, and mean transverse momenta all probe the total energy contained in the collision as well as the total entropy produced and the chemical composition of the fireball.

- $dN_{ch}/d\eta$: The number of charged hadrons per unit pseudorapidity. Measurements are from the ALICE Collaboration [113].
- $dN_i/dy$, $i \in \{\pi, p, K, etc.\}$: Identified charged hadrons per unit rapidity. Measurements are from the ALICE Collaboration [114].
- $dE_T/d\eta$: Transverse energy, defined as $E_T = \sqrt{m^2 + p_T^2}$, per unit rapidity. Measurements are from the ALICE Collaboration [115].
- $\langle p_T \rangle_i$, $i \in \{\pi, K, p\}$: Mean transverse momenta of identified hadrons. Measurements are from the ALICE Collaboration [114].
- $v_n\{2\}$: Two-particle radial Fourier coefficients. Measurements are from the ALICE Collaboration [116].
- $v_n\{4\}$: Four-particle radial Fourier coefficients. Measurements are from the ALICE Collaboration [117].

The momentum-space geometry of the fireball is a traditional sign of hydrodynamic behavior and is quantified by the radial Fourier coefficients. In an almond-shaped medium, the pressure gradient is larger along the narrower axis; this means that the expansion rate in this axis is faster due to the higher gradient. As event-by-event fluctuations of partons grow more complex, it is possible to probe more of the momentum space geometry by including more advanced, next generation observables.[8]

### 2.6.2 *Next generation observables*

These observables explore more sophisticated effects using correlations between geometric features or momentum fluctuations and decompositions of observables into a

8 While this thesis considers only the two and four-particle Fourier coefficients, coefficients with higher numbers of particles are also measured.

linear and non-linear response of the medium. These have been shown in [2, 118] and others to be sensitive to details of the hydrodynamic evolution such as the viscosity or the initial state geometry. As these are key details arising from the CGC initial state or viscous hydrodynamic evolution, they will be key to providing new insights. Near this study's conclusion, it was shown that other next-generation observables can also be effective in constraining the transport coefficients of QCD matter [53].

- Two- and three-plane Scalar Product Event Plane Correlators: Correlations between expansion coefficients $v_n$ reveal patterns of fluctuations in the initial state and non-linear effects in hydrodynamics. Measurements, as well as detailed definitions, are from the ATLAS Collaboration [119]. These patterns are coupled and reproduction of them in parametric models has been shown to be highly model-dependent [118]. The ALICE Collaboration measures similar quantities, which are also used, statistics allowing [120].

- $\chi_{n,mk}$: Nonlinear response coefficients that quantify mixing between higher and lower-order modes. These decompose higher order $v_n$ into a linear component from the corresponding position space energy density Fourier coefficients ($\epsilon_n$) and a non-linear component from lower modes. For example, $v_5 = v_5^L + \chi_{5,32} v_3 v_2$. Measurements and more details may be found in [120].

- Linear and non-linear flow modes: these quantify the linear and non-linear response of the flow to collision geometry, similar to the event plane correlators and $\chi_{n,mk}$ above [120].

- $\delta p_T / \langle p_T \rangle$: Correlated transverse momentum fluctuations, alternately denoted as $\sqrt{C_m}/M$. This quantifies the correlations between deviations from the mean transverse momentum. If the deviations are uncorrelated over all events, this quantity is 0 [121].

The purpose of using these carefully chosen observables is to efficiently constrain the properties of strongly-interacting matter. For example, the multiplicities constrain the overall energy of the system, the radial Fourier coefficients constrain the momentum-space geometry of the hydrodynamic stage, and next-generation observables couple

various aspects of the medium evolution. More specifically, details such as the linear and non-linear relationships between $v_n$ can specifically target the hydrodynamic phase of the evolution. This will be coupled with state-of-the-art physical simulation and Bayesian inference to improve the understanding of strongly interacting matter.

# Part II

# STATISTICAL FOUNDATIONS

Bayesian inference is introduced with historical overview and some details of probability theory. Relevant considerations for numerical implementation are discussed, including computer experiment design for training a surrogate model on expensive numerical calculations. Finally, a simple example is used to demonstrate how to use Bayesian methods to provide quantitative guidance in a simple physics problem.

# 3

# INTRODUCTION TO BAYESIAN INFERENCE

When doing science with simulations, it is necessary to engage in plausible reasoning. Especially in numerically-expensive pursuits such as heavy ion collisions,[1] scientists attempt to make conclusions by considering the results of computation with a fixed set of parameter values. This discards the inherent uncertainty on the model inputs by construction. It then becomes difficult to convincingly argue that an effect is (or is not) present: it may simply be a choice of having arbitrary confidence in input parameters.

The choice of particular parameters is, of course, not random. It is often chosen by what is winkingly referred to as "chi-by-eye", referring to the $\chi^2$ metric. "Chi-by-eye" calibrations are simply varying model inputs by hand and heuristically determining a combination of parameters that fits "well-enough".

The issue that arises with such a method is that it is misleadingly precise: numerical simulations always have some uncertainty and experimental data is always measured with uncertainties arrived at through painstaking modeling. It is a disservice to these efforts to ignore this uncertainty when performing quantitative comparisons. Instead, scientific efforts connecting rigorous theoretical modeling to experimental results should be honest in their comparisons and account for both theoretical and experimental uncertainties.

To account for these uncertainties consistently when comparing to data, this thesis employs the tools of Bayesian inference. By doing so, it is possible to make conclusive general statements rather than attempting to hide the fact that conclusions were made with a naive exploration of the parameter space.

This is comparable to the state of science in the 18th century. James Bernoulli (1713) highlighted the difference between deductive logic in games of chance and the induc-

1 The primary culprit behind the computational expense of heavy ion collisions is the need to evolve differential equations on a spatial grid with sufficient resolution to faithfully reproduce the underlying physics.

tive logic required for more general reasoning and sought a way to make plausible statements about everyday outcomes. The solution to this came from a Presbyterian minister, Reverend Thomas Bayes, in a letter published after his death in 1763 and formulated into its now-familiar expression by Laplace in 1812. In this approach to statistics, probability denoted a degree of belief (and consequently, uncertainty) in a given proposition. For example, Laplace used the inductive logic proposed by Reverend Bayes to reason plausibly about the possible values of the mass of Saturn given available observations and came to a result where "... it is a bet of 11,000 to 1 that the error of this result is not 1/100th of its value." All these years after Laplace's death, the estimate for the mass of Saturn has only changed by approximately 0.63%, demonstrating the power of reasoning plausibly and with quantified uncertainty [122].

However, Bayes' Theorem includes a quantity to explicitly account for prior beliefs, or rather beliefs informed by observations other than those under consideration. This led to the denouncement of this school of statistical analysis as "subjective" and the rise of so-called Frequentist statistics. In this new school of statistics, probability is not interpreted as a "degree of belief" a la Bayes, but rather "the long-run limiting frequency of an event occurring were it to be identically repeated an infinite number of times". This less-intuitive definition of probability is what is often taught in schools, but in this author's experience, is rarely how physicists actually interpret probabilities in practice. While Bayesians attempt to answer statistical problems by reasoning plausibly, Frequentists calculate limiting frequencies of events that may only be able to occur once. This split in the fundamental interpretation of probability resulted in a quip by Louis Lyons: *Bayesians address the question everyone is interested in by using assumptions no-one believes, while frequentists use impeccable logic to deal with an issue of no interest to anyone* [123, 124]. It is worth noting that, often enough, the numerical results between the two schools of statistics are equivalent, merely the interpretation differs. Bayesians additionally tend to treat the entire final probability distribution as the result of interest as it quantifies the full degree of belief, while Frequentist results

are often summary statistics or similar estimates drawn from the final probability distribution.[2]

Such was the state of science in the 19th and the first part of the 20th centuries. In 1946, Richard Cox resurrected the Bayesian perspective by establishing the rules for logical and consistent quantitative reasoning via the rules of Boolean logic and ordinary algebra in a move recognized as "the most important advance in the conceptual ... formulation of probability theory since Laplace" [125]. In the late 20th and early 21st centuries, the rapid increase in available computational resources has allowed for the computing of Bayes' Theorem in more general, less analytically-solvable cases. This has proven valuable as uncertainty quantification is often as valuable as parameter estimation, which has been the focus of more traditional statistical approaches. It is with Cox's rules that I will derive Bayes' Theorem anew.

## 3.1 BAYES' THEOREM

To begin a discussion of probabilities, it is necessary to introduce a notation. In this work, I will strive to stick to that of Gelman's Bayesian Data Analysis [126]. First, $p(A)$ denotes the probability density $p(\cdot)$ of a proposition $A$. Next, $p(A|B)$ denotes the probability density of proposition $A$ conditional on proposition $B$, *i.e.* the probability density of $A$ *given* $B$. There may be multiple statements to which the proposition of interest is conditional; these are all contained to the right of the vertical bar, *e.g.* $p(A|B, C, D, \dots)$. In general, the comma in the probability density denotes "and", thus $p(A, B|C, D)$ may be read as the probability density of $A$ and $B$ given $C$ and $D$. The overall probability of a proposition, rather than the probability density, is denoted $\Pr(A)$. Negation is indicated with a bar over a quantity, thus *not* $A$ is denoted $\bar{A}$. Throughout this thesis, the shorthand proposition $I$ is used to represent all other information one may have about a particular probability statement. For example, one may be interested in the value of a physical quantity, but are fundamentally

2 It is worth noting that one can use Bayes' Theorem while remaining a Frequentist. The division between Bayesians and Frequentists is in the interpretation of probability, not in the simple use of a foundational theorem.

restricted to investigating it in the context of a particular model. The particular model is absorbed into the additional information I.

Next, some fundamental statements of probability theory are introduced. The first is the *sum rule*,

$$\Pr(X|I) + \Pr(\bar{X}|I) = 1 \tag{3.1}$$

that states that the sum of the probability of proposition X being true plus the probability of proposition X being false must be 1. In applications where a proposition is not Boolean (*i.e.* strictly true or false) but may be continuous, this is simply the normalization requirement of probability. The second fundamental statement is called the *product rule*,

$$\Pr(X, Y|I) = \Pr(X|Y, I) \times \Pr(Y|I). \tag{3.2}$$

The product rule reveals that the probability that both X and Y are true can be expressed as the probability of X being true given that Y is true times the probability that Y is true. Both of these are broadly intuitive statements, but have a firm foundation in Boolean logic.[3]

Relaxing the Boolean nature of the above statements results in dealing with *probability densities*. Revisiting the product rule yields

$$p(X, Y|I) = p(X|Y, I)p(Y|I). \tag{3.3}$$

If the goal is to make probability statements about X given Y, this rule may be rearranged to provide a joint probability distribution for X and Y,

$$p(X|Y) = \frac{p(X, Y)}{p(Y)} \tag{3.4}$$

where the additional information I is suppressed for conciseness. It is always assumed to present, but left implicit. The product rule may then be used again since X and Y obey the principle of exchangeability: the joint probability of X and Y $p(X, Y)$ is indistinguishable from the joint probability of Y and X, $p(Y, X)$.[4]

$$p(X|Y) = \frac{p(Y, X)}{p(Y)} = \frac{p(Y|X)p(X)}{p(Y)} \tag{3.5}$$

3 For the reader interested in Boolean logic, the proof of the sum and product rules may be found in chapter 2 of [125] and alternatively, appendix B of [122].

4 By connecting $p(X|Y)$ and $p(Y|X)$ via a joint distribution and exchangeability, inference and plausible reasoning are made possible.

This is *Bayes' Theorem*. Each piece of this theorem is of critical importance to the remainder of this thesis and considerable space will be devoted to each. To connect Bayes' Theorem to physical modeling questions, I will rewrite it in a more convenient notation,

$$p(H|d, I) = \frac{p(d|H, I)p(H, I)}{p(d, I)}. \tag{3.6}$$

In this notation, $H$ represents a particular *hypothesis*, such as the proposed values of a set of parameters and $d$ represents data to which the hypothesis is compared. $p(H|I)$ is the *prior* and is the belief in the hypothesis $H$ *prior* to comparison with measurements. $p(d|H, I)$ is the *likelihood* and is the likelihood for the data $d$ given the hypothesis $H$. $p(d|I)$ is the *Bayes* or *Bayesian evidence* and is typically treated as a normalization constant such that $\int p(H|d, I)dH = 1$. Older names for the evidence are the *prior predictive* or the *marginal likelihood*, but these are poor descriptions of the meaning of this quantity. The Bayes evidence quantifies a balance between quality of fit via the likelihood and predictive power, by penalizing increasing dimensionality. It can therefore be used in model selection: the "best" model is the one that fits the data best with the fewest number of free parameters. Finally, the quantity of interest is $p(H|d, I)$, the *posterior*. It quantifies the belief in a given hypothesis $H$ *posterior* to comparison with measured data $d$. Bayes' theorem formalizes statistical learning by making a prior belief explicit and then comparing it to data, after which the prior belief is determined to be relatively more or less likely. The result *posterior* to comparison with data is the new state of understanding.

Specifics of the prior, likelihood, and data are application-specific and are addressed explicitly for each study. Bayes' theorem may be used to solve inverse problems, calculating "how likely is the data given a set of model parameters" (the likelihood) in order to answer a primary question of interest, "how likely is the set of model parameters given the data" (the posterior). Consequently, a hypothesis for which the observed data is unlikely is itself an unlikely hypothesis. This allows for the translation of a forward problem more typical in physics phenomenology – given some input, what is the outcome? – to an inverse problem – given some outcome, what were the inputs? – through which one hopes to reason plausibly about the physical world.

The next piece of the puzzle that must be introduced is known as *marginalization.* Marginalization follows directly from the sum and product rules. First, the sum r)

$$\Pr(X, Y) = \Pr(Y, X) = \Pr(Y|X)\Pr(X). \tag{3.7}$$

Next, the same expression for $\Pr(X, \bar{Y})$ may be added to both sides,

$$\Pr(X, Y) + \Pr(X, \bar{Y}) = [\Pr(Y|X) + \Pr(\bar{Y}|X)]\Pr(X). \tag{3.8}$$

Of course, the term in the square brackets must be 1 from the sum rule and the expression simplifies,

$$\Pr(X) = \Pr(X, Y) + \Pr(X, \bar{Y}). \tag{3.9}$$

In this analysis, $Y$ and $\bar{Y}$ form a mutually exclusive and exhaustive set of outcomes. In most applications, there are more than two binary outcomes and thus the mutually exclusive and exhaustive set of outcomes can be represented by, for example, $\{Y_n\} = \{Y_1, Y_2, \ldots, Y_M\}$. Thus, each of these must be accounted for and results in the marginalization rule

$$\Pr(X) = \sum_{n=1}^{M} \Pr(X, Y_n). \tag{3.10}$$

It is worth noting that this *only* holds if the sum rule may be used as it was above. For the sum rule to hold, the set $\{Y_n\}$ must obey the normalization condition

$$\sum_{n=1}^{M} \Pr(Y_n) = 1 \tag{3.11}$$

revealing it to be a mutually exclusive and exhaustive set. In the continuum limit (equivalently, $M \to \infty$), the marginalization rule becomes an integral

$$p(X) = \int_{-\infty}^{\infty} p(X, Y) dY. \tag{3.12}$$

The use of marginalization is important because it allows one to "integrate out" nuisance parameters that aren't relevant to the problem but must be included in the inference and it allows for the visualization of complex high-dimensional distributions.[5]

5 Marginalization does not only go from two dimensions to one, but may go from 10 to 4 or any other combination so long as the result is a lower-dimensional projection of the original distribution of interest.

## 3.2 MARKOV CHAIN MONTE CARLO

All this is very well and good, but how does one actually go about putting this into practice? Not all posteriors have an analytically-tractable form and, due to practical limitations, one cannot query the posterior at every point in the parameter space. Similarly, simply querying the posterior on a fixed grid over a certain range is inefficient as much of the space may be uninteresting and unlikely, meaning that the region of interest is poorly explored.

In this thesis, this problem is addressed using Markov Chain Monte Carlo (MCMC). In MCMC, samples are drawn from the true posterior and the set of samples may be used to explore the posterior distribution numerically. To explore how this is done, I will introduce Markov Chain Monte Carlo in general and will proceed to discuss the details of one of the first MCMC algorithms, the Metropolis algorithm. Many different MCMC algorithms exist, including those used in this work [127, 128], but the technical details prohibit them from being an accessible introduction to the unfamiliar. They are often purposefully chosen to efficiently sample posteriors with particular features.

To begin in the beginning, it is necessary to introduce the Markov Property. This is sometimes described as memorylessness, *i.e.* a process which obeys the Markov Property is independent of history of the process. A Markov Chain is a collection of the states of a Markov process. Each state in the collection (or chain) is only determined by the previous state. In this way, it is often introduced in textbooks as a "drunken walk" as it follows no directed pattern and the drunkard's next step only depends on where they are at present. To put this in more formal statistical language, a Markov chain is a sequence of random variables $\theta^1, \theta^2, \ldots$ for which the distribution of $\theta^n$ given all previous $\theta$ depends only on $\theta^{n-1}$.

But how is this useful? A Markov Chain can be constructed so that it has a unique stationary distribution. Then, one can show that the stationary distribution is equal to the target distribution of interest. For the Markov Chain to have a unique stationary distribution, the chain must be irreducible, aperiodic, and not transient. The chain is irreducible if the random walk has some positive, finite probability of reaching any state from any other state - *i.e.* it can jump to anywhere from anywhere. Aperiodicity

and non-trasience of the chain hold for random walks on any proper distribution. Thus, what remains is to prove that the stationary distribution of the chain is the target distribution. To do this, I first introduce the Metropolis algorithm, and then prove explicitly that the distributions are equal.

The Metropolis algorithm is a random walk with a rule for accepting or rejecting a proposed step such that the walk converges to a specified target.

Step 1: Choose a set of starting points. While these can be chosen intelligently, the common practice is to randomly disperse points throughout the sample space, allow the chains to walk toward the target distribution, and then throw out the first N steps of the chain so that the chain has an opportunity to explore the potentially arbitrarily-complex space.

Step 2: For $n = 1, 2, \ldots$:

Step 2.1: Sample a proposed next step $\theta'$ from a proposal (or jumping) distribution at step $n$, $J_n(\theta'|\theta^{n-1})$. In the Metropolis algorithm, this proposal distribution must be symmetric – $J_n(\theta_a|\theta_b) = J_n(\theta_b|\theta_a)\ \forall\ \theta_a, \theta_b, n$.

Step 2.2: Calculate the *Metropolis ratio*:

$$r = \frac{p(\theta'|y)}{p(\theta^{n-1}|y)} \tag{3.13}$$

where $y$ are the data. This is immediately recognizable as the ratio of the posteriors for parameter hypotheses $\theta$.

Step 2.3: Add a state to the Markov chain as follows

$$\theta^n = \begin{cases} \theta' & \text{with probability } \min(r, 1) \\ \theta^{n-1} & \text{otherwise} \end{cases} \tag{3.14}$$

until the chain is a sufficiently close approximation to the target distribution for its intended purpose.

It is important to note two details: first, that even if the distribution does not jump to the next state, this still constitutes an iteration in the chain. Second, it is possible to denote a transition distribution $T_n(\theta^n|\theta^{n-1})$ as a point mass at $\theta^n = \theta^{n-1}$ and the proposal distribution $J_n(\theta^n|\theta^{n-1})$ weighted by the acceptance rate.

With the Metropolis algorithm in hand, it is possible to return to the proof of MCMC convergence with a digestible example. Next, consider a target distribution $p(\theta|y)$ at step $n-1$ in order to demonstrate that the target distribution and the Metropolis algorithm Markov chain's stationary distribution are one and the same. Two draws from $p(\theta|y)$ are hereon labeled as $\theta_a$ and $\theta_b$.

The unconditional probability density of transition from $\theta_a$ to $\theta_b$ is

$$p(\theta^{n-1} = \theta_a, \theta^n = \theta_b) = p(\theta_a|y) J_n(\theta_b, \theta_a) r. \tag{3.15}$$

The Metropolis ratio $r$ weights the right hand side, but because of the labeling of $a$ and $b$ the acceptance probability is 1, so it has only been included here for clarity. The unconditional probability density of transition from $\theta_b$ to $\theta_a$ is

$$p(\theta^n = \theta_a, \theta^{n-1} = \theta_b) = p(\theta_b|y) J_n(\theta_a, \theta_b) \left( \frac{p(\theta_a|y)}{p(\theta_b|y)} \right) \tag{3.16}$$

$$= p(\theta_a|y) J_n(\theta_a|\theta_b), \tag{3.17}$$

where the transition probability density from $\theta_a$ to $\theta_b$ is recovered due to the symmetry of the proposal distribution. As the joint distributions are symmetric, $\theta^n$ and $\theta^{n-1}$ have the same marginal distributions and so $p(\theta|y)$ is the stationary distribution of the Markov chain of $\theta$. While this example specifically exploits the Metropolis algorithm, similar proofs may be found for other algorithms.

Of course, it is all very well and good to claim it is a random walk, but it is important to check this explicitly. One of the first diagnostics of a Markov chain is the *autocorrelation*. This quantifies if the chain is taking a directed walk – meaning that the chain is not truly in a Markov state – or if the chain has converged to a target distribution, where steps are uncorrelated and are thus draws from the target distribution. The autocorrelation is calculated by defining some lag, $l$. A chain is then shifted by this number of steps, and then the dot product of the chain is taken, normalized by the product of the magnitudes of the two chains. Thus, the 0 lag autocorrelation is 1 by construction and two tightly-correlated chains have autocorrelation close to 1 for many lag steps. A decorrelated chain's autocorrelation drops to 0 quickly, perhaps even after a single step. Additional diagnostics include the Gelman-Rubin diagnostic [129], and the between- and within-sequence variance [126], although a technical discussion on various MCMC diagnostics is beyond the scope of this thesis.

In this thesis, the parallel-tempering Markov Chain Monte Carlo software *ptemcee* is used [128]. This defines a ladder of inverse temperatures $\beta_i$ and evolves an ensemble of walkers on "tempered" distributions $p(y|x)^{\beta_i} p(x)$ where in the limit $\beta \to 0$, the prior distribution $p(x)$ is recovered while at $\beta = 1$, the prior is multiplied by the likelihood. The walkers are given the chance to exchange with walkers at neighboring temperatures. The advantage of this technique is that walkers at high temperatures are progressively less influenced by peaks in the likelihood as these peaks are tempered. Conversely, walkers at $\beta = 1$ sample from the target posterior. Thus, only samples from the $\beta = 1$ chain are used in analysis and gain all the advantage from the higher chains in efficiently sampling multimodal distributions.

Another convenient feature of using parallel-tempering MCMC is in the calculation of the Bayes Evidence. If one defines the Bayes Evidence as a function of inverse temperature,

$$Z(\beta) = \int dx \, [p(y|x)]^{\beta} \, p(x), \tag{3.18}$$

then it satisfies a differential equation

$$\begin{aligned} \frac{d \ln Z(\beta)}{d\beta} &= \frac{1}{Z(\beta)} \int dx p(x) \ln [p(y|x)] \, [p(y|x)]^{\beta} \\ &\equiv \langle \ln [p(y|x)] \rangle_{\beta} \end{aligned} \tag{3.19}$$

and the log Bayes Evidence $\ln Z(\beta = 1)$ is calculated by integrating the average at each temperature with error determined by the number of rungs in the temperature ladder. The convergence behavior of *ptemcee* is also greatly superior to other MCMC algorithms and rapidly converges to the target distribution of interest with autocorrelation dropping to 0 more than an order of magnitude more quickly than in similar algorithms [47].

## 3.3 BAYESIAN MODEL COMPARISON

Bayesian model preference is evaluated using the ratio of two models' Bayes evidences (the denominator of Eq. 3.6), called the Bayes factor

$$B_{01} = \frac{p(d|M_0)}{p(d|M_1)} \tag{3.20}$$

where models $M_i$ are subscripted 0 and 1. The Bayes factor $B_{01}$ is used to quantify model preference. A Bayes factor greater than one represents an increase of support for $M_0$ relative to $M_1$ and directly corresponds to the odds of $M_0$:$M_1$. The logarithm of the Bayes factor, $\ln B$, is often used both for ease of computation as in *ptemcee* but also to quantify the relative probability in a more concise manner. The Bayes factor allows for determination of a preferred model when the result is not immediately obvious. In sophisticated experiments, multiple observables are often considered simultaneously and the comparison becomes more complex. While tools such as the $\chi^2$-per-degree-of-freedom exist, these often make implicit assumptions about the shape of the posterior distribution.

The Bayes factor makes no additional distributional assumptions beyond those explicitly chosen by the priors and the likelihood. It also penalizes the incorporation of additional parameters. John von Neumann reportedly said that "with four parameters I can fit an elephant, and with five I can make him wiggle his trunk" [130]. The Bayes factor implements Occam's Razor, which says that it is futile to do with more things that which can be done with fewer [131]. A model with many parameters should be disfavored in comparison with an equally-successful model with fewer parameters. A further discussion of Bayesian model selection may be found in [132].

Empirical scales are used to determine when there is weak, moderate, and strong evidence for $M_0$ vs. $M_1$. The Bayes factor is easily interpretable as it gives the direct odds ratio of one model to the other. In this work, the Jeffreys' Scale is used (Table 3.1) [133]. The Jeffreys' Scale is a standard scale for model selection in Bayesian studies.

| $\lvert\ln B_{01}\rvert$ | Odds | Probability | Strength of evidence |
|---|---|---|---|
| $< 1.0$ | $\leqslant 3:1$ | $< 0.750$ | Inconclusive |
| 1.0 | $\sim 3:1$ | 0.750 | Weak evidence |
| 2.5 | $\sim 12:1$ | 0.923 | Moderate evidence |
| 5.0 | $\sim 150:1$ | 0.993 | Strong evidence |

Table 3.1: The Jeffreys' Scale, reproduced from [132].

## 3.4 PRINCIPAL COMPONENT ANALYSIS

In many situations, modeling is designed to predict a single outcome. In contrast, heavy-ion collisions is such a rich physics laboratory precisely because of the multi-messenger nature of the QGP, with different particle species providing different insights into the chemistry and evolution of the fireball. In turn, these different probes allow for the construction of many observable quantities as summary statistics of a shower of particles. These summary statistics can each be reconstructed from theoretical simulations and compared, constraining the simulation.

Doing this directly is the naive, but inefficient approach. For example, the number of protons produced in a particular collision shares mutual information with the total number of charged hadrons (including protons) produced. Instead of considering each observable totally independently, it should be possible to determine this mutual information and determine the minimum amount of information required to describe all the various observables.

Enter principal component analysis (PCA). In a space defined by the observables, where each dimension corresponds to a particular observed quantity, it is possible to identify correlations. Principal component analysis is a simple technique to "rotate" in observable space into a linear combination of the original axes such that every dimension of the data is linearly independent.

This rotation is also invertible, meaning that predictions can be made for the transformed space and inverted back to the observable space. This is useful as it is no longer necessary to interpolate between hundreds of dimensions in the observable space, but rather only interpolate in a $\mathcal{O}(10)$ dimensional space, which is much more feasible.

Another way to think of this rotation is by a decomposition of the data in question to its eigenvalues and eigenvectors. The eigenvalues are the fraction of the total variance in the data described by each eigenvector.

More formally, PCA is a dimensionality-reduction tool and is sometimes described as an elementary type of unsupervised learning. This is useful for analyses with a large amount of data, such as ours, because much of the information contained in the data is redundant due to features such as correlations. One motivation for this is to

speed up calculations, but another one is simple: surrogate modeling techniques (or "emulators") often perform poorly with multidimensional outputs and this remains an area of active research. As such, one has to train an emulator for each feature one wishes to predict. With a long list of observables, this becomes many independent surrogate models that need to be trained and deployed in MCMC. This is computationally expensive and redundant.

A much more intelligent way of doing this is to construct new characteristics that adequately summarize the data and allows for reconstruction of the high-dimensional data with fewer-dimensional inputs. The way this is done is by finding linear combinations of the original inputs (correlated) that project the inputs into a space of these new characteristics (uncorrelated), which are called principal components. This also has the useful feature of discarding features in the data that do not explain much variation while focusing on the features that do explain the variation.

To put this another way, the different measured quantities are not totally independent of each other. However, there exists a vector space into which these different quantities may be projected. This transformation orders the components of the new, independent vector space so that the component that describes most of the variation is the first component; the component that describes the next largest fraction of the variation is the second component; and so on and so forth. The discussion that follows broadly reproduces that of [134, 135] to make this more mathematically explicit, borrowing a simple example.

First, the data has to be encoded as a matrix. Suppose that the width $w$, height $h$, and length $l$ of some large number of boxes are known. Box $i$'s measurements are recorded in a column vector $x_i$ of length $d = 3 - x_i^T = (w_i, h_i, l_i)$. Now there are some large number of data vectors to encode in a data matrix. This is done by simply stacking the vectors,

$$X = \begin{pmatrix} w_1 & h_1 & l_1 \\ w_2 & h_2 & l_2 \\ & \vdots & \\ w_n & h_n & l_n \end{pmatrix} \tag{3.21}$$

Each column thus refers to a single coordinate and the matrix is $n \times d$ dimensional. This matrix is also standardized in practice, meaning the mean has been subtracted:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \left(\frac{1}{n}\sum_{i=1}^{n} x_{i1}, \ldots, \frac{1}{n}\sum_{i=1}^{n} x_{id}\right)^T \tag{3.22}$$

The zero-centered vectors are used as rows of the matrix X:

$$X = \begin{pmatrix} x_1^T - \mu^T \\ x_2^T - \mu^T \\ \vdots \\ x_n^T - \mu^T \end{pmatrix} \tag{3.23}$$

We then write the sample covariance around the mean S:

$$S = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)(x_i - \mu)^T = \frac{1}{n-1}X^T X \tag{3.24}$$

Note that dividing by $\frac{1}{n-1}$ rather than $1/n$ is simply Bessel's correction to correct bias in the estimation of the population variance from a sample.

With this notation in hand, it becomes possible to formally tackle PCA. The goal is the find a collection of $k \leqslant d$ unit vectors $v_i \in \mathbb{R}^d$ for $i \in (1, \ldots, k)$ such that the variance of the dataset projected onto the direction determined by $v_i$ is *maximized* and the $v_i$ are all mutually orthogonal. These unit vectors $v_i$ are the principal components.

The projection of $x \in \mathbb{R}^d$ onto the line determined by any $v_i$ is the dot product $v_i^T x$, so the variance of the dataset projected onto the first principal component $v_!$ is just

$$\frac{1}{n-1}\sum_{i=1}^{n}(v_1^T x_i - v_1^T \mu)^2 = v_1^T S v_1. \tag{3.25}$$

The problem is now how to find $v_1$ such that it maximizes Eq. 3.25 subject to the constraint that $\|v_1\| = 1$. This is done with Lagrange multipliers and results in

$$Sv_1 = \lambda_1 v_1 \tag{3.26}$$

thereby implying that $v_1$ is an eigenvector of S. Since $\|v_1\| = v_1^T v_1 = 1$, the eigenvalue is just the variance of the dataset along $v_1$, i.e. $v_1^T S v_1 = \lambda_1$. This process can be repeated for a new direction $v_2$ with the additional constraint that $v_1 \perp v_2$ and for a further new direction $v_3$ with the additional constraint that $v_3 \perp v_1, v_2$ and so on and so forth.

The first $k$ principal components of $X$ are the $k$ eigenvectors of the covariance matrix $S$ ordered by their eigenvalues. The following discussion closely follows [134].

Since $S$ is a real symmetric matrix, i.e. $S = S^T$, there exists a full set of orthonormal eigenvectors for $S$ over $\mathbb{R}$. Consider the case $k = d < n$. Taking $k = d$ PCs, one can construct a $d \times d$ matrix $V$ whose columns are the eigenvectors of $S$ and that diagonalizes $S$.

$$S = V\Lambda V^T = \sum_{i=1}^{r} \lambda_i v_i v_i^T \tag{3.27}$$

where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_d)$ and $r = \text{rank}(X)$. This means that the PCs are the columns of a rotation matrix and forms a new basis for the data aligning the dataset $X$ so that the dimensions are uncorrelated.

A similar method exists, called Singular Value Decomposition, which is a matrix factorization method used in linear algebra that is generalized and efficient for non-square matrices. This is usually used in practice since it is more general, although the outcome is similar to the simpler process outlined above but the process is less transparent.

## 3.5 INFERENCE WITH EXPENSIVE LIKELIHOODS

All the proceeding text is true for general inference, but what if it is not practical to call the model tens or hundreds of thousands of times to generate an MCMC chain? The computationally expensive piece of Bayes' Theorem is the likelihood function, which requires the model to be evaluated. As a result, this problem is known as the problem of an expensive likelihood function.[6]

While specifics of the model implementation are addressed later in this thesis, suffice it to say that a single model run requires 0.5-1 core-years.[7] The resources required to run a single MCMC run would thus take an entire high performance computing cluster for a year (and even then, with relatively few steps) and the results would all

6 Likelihood-free methods exist, such as Approximate Bayesian Computing. However, these have disadvantages that make them poorly-suited to this thesis. They are often also special cases of more general likelihoods.

7 A core-year is the equivalent computing of a single CPU core running continuously for 1 year.

be dependent upon specific choices made in the study at hand – hardly an efficient way to proceed.

A more efficient and general strategy is obviously required. In this section, I outline the idea of surrogate modeling as well as its implementation in this study using regression with Gaussian Processes (GPs). Following this, strategies for more efficient surrogate modeling are pursued via improvements to computer experiment design, *i.e.* how training data is chosen for the surrogate models.

### 3.5.1 *Surrogate Modeling*

Surrogate modeling is an effective strategy for expensive likelihood functions. A cheaper model (the "surrogate") is trained to emulate the expensive model by using calculations from the more expensive model. This less computationally expensive model can be considered a low-fidelity model, or a model-of-a-model, and compromises a degree of accuracy for computational time. It does this by simply mapping inputs to outputs and learning the functional relationship between them rather than attempting to produce a "coarse" version of all the intermediary physics.

Many different surrogate modeling strategies exist, varying from simple interpolation through to techniques such as deep convolutional neural networks. In this study, I use Gaussian Process Emulators as a nonparametric surrogate model to learn the mapping between parameter inputs and model outputs.

Gaussian processes are a regression technique to determine the relationship between variables and data and are used in heavy ion collisions to make predictions as surrogate models. Given a set of training points, there are infinitely many functions that can describe the points. Gaussian processes assign a probability to each of these functions, meaning that the output is a probability distribution of the characterization of the data. Conveniently, this also allows one to determine the relative confidence in the prediction. The mathematical details of the following discussion are reproduced from [126, 136]. The only assumptions by the GP are that it assumes the function is continuous and smoothly varying with respect to the length scales of the observations.

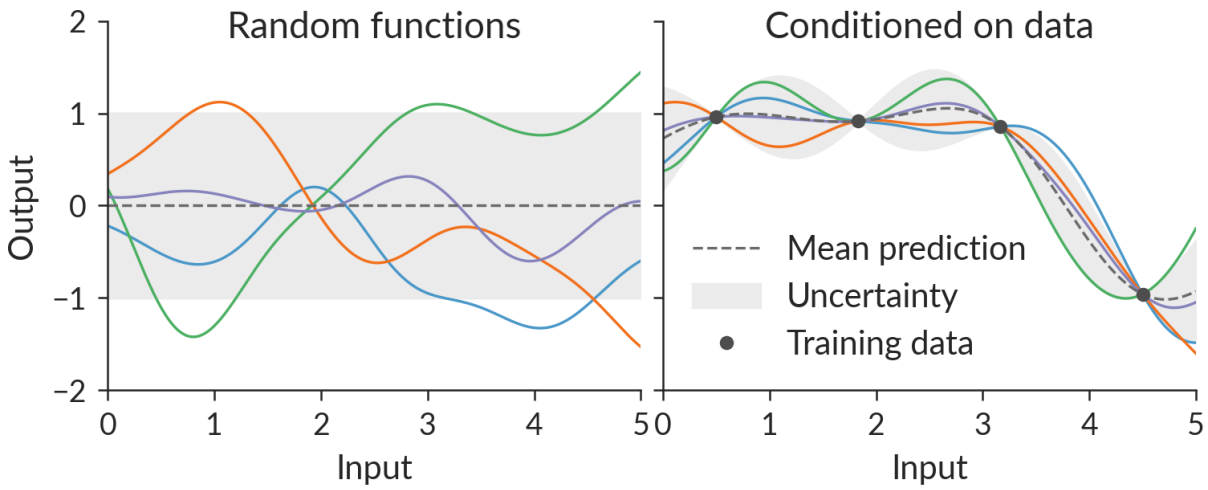


Figure 3.1: Draws from a Gaussian Process emulator prior to conditioning on data (left) and posterior to conditioning (right).

A GP can be thought of as putting a Gaussian prior on an unknown regression function $\mu(x)$, where individual draws from the Gaussian process are random functions whose outputs are normally-distributed around the mean of the process.[8] Due to the flexibility of an interpolation through function space, many features of interest can be reproduced without extensive modeling choices and careful tuning.

A Gaussian process $\mu \sim \mathrm{GP}(m, )$ is parametrized by a mean function $m$ and covariance function $k$, where $\mu$ and its prior define the GP as a random function. The values at specified points $x_1, \ldots, x_n$ are a draw from the $n$-dimensional Gaussian (normal) distribution with mean $m$ and covariance $k$

$$\mu(x_1), \ldots, \mu(x_n) \sim \mathrm{N}\Big( (m(x_1), \ldots, m(x_n)), \; k(x_1, \ldots, x_n) \Big). \tag{3.28}$$

This is a nonparametric model. The mean function is an initial guess at the regression function and is the starting point when conditioning the GP on training data.

$k$ specifies the covariance between any two points of the process, with $k$ as an $n \times n$ matrix in which $k_{pq}$ corresponds to the covariance between two points $k(x_p, x_q)$, or alternately $k(x, x')$. The smoothness of the GP is controlled by $k$ and also controls the

8 A GP is also the limit of a neural network given certain specific choices [137].

degree of shrinkage toward the mean. The most common choice for the covariance function is the Gaussian,

$$k(x, x') = \tau^2 \exp\left(-\frac{|x - x'|^2}{2l^2}\right) \tag{3.29}$$

where $\tau$ and $l$ are unknown parameters in the covariance and $|x - x'|^2$ is the Euclidean distance between two points $x$ and $x'$. $\tau$ determines the magnitude and $l$ the smoothness. In practice, the hyperparameters are independent in each dimension,

$$k(x, x') = \mathrm{cov}(\mu(x), \mu(x')) = \tau^2 \exp\left(-\sum_{j=1}^{p} \frac{(x_j - x'_j)^2}{2l_j^2}\right). \tag{3.30}$$

The hyperparameters can then be varied and a posterior or maximum likelihood may be found, allowing the GP to be used as a regression model.

But how does one actually arrive at the trained GP? Bayesian inference allows for a convenient explanation.

$$p(X|Y) \propto p(Y|X)p(X) = p(X, Y) \tag{3.31}$$

Given a prior $p(X)$ on the test points $X$ and a set of training points $Y$, one can obtain a posterior. First, the joint distribution $p(X, Y)$ is formed between the test points $X$ and training points $Y$. The result is a multivariate Gaussian with dimensions $|Y| + |X|$. Conditioning allows one to find $p(X|Y)$ from $p(X, Y)$. Formally, when conditioning two distributions $X \sim N(\mu_x, \Sigma_{XX})$ and $Y \sim N(\mu_y, \Sigma_{YY})$, then conditioning is defined by

$$X|Y \sim N\left(\mu_X + \Sigma_{XY}\Sigma_{YY}^{-1}(Y - \mu_Y), \Sigma_{XX} - \Sigma XY\Sigma_{YY}^{-1}\Sigma YX\right) \tag{3.32}$$

and

$$Y|X \sim N\left(\mu_Y + \Sigma_{YX}\Sigma_{XX}^{-1}(X - \mu_X), \Sigma_{YY} - \Sigma YX\Sigma_{XX}^{-1}\Sigma XY\right). \tag{3.33}$$

This should be read as, for example, $X|Y$ is conditional on $Y$ and one can then find the conditioned outputs from the GP given training points $Y$ for a set of input hyperparameters. Specific methods of tuning the hyperparameters and exploring the hyperparameter space is beyond the scope of this discussion.

Because this is important to understand and is still comparatively new to physics audiences, I will provide one last explanation of how a GP is trained. Priors on GP

hyperparameters are also conditionally conjugate such that for a given Gaussian process, the posterior when compared to data is itself a GP, but with updated inputs.

It is impossible to evaluate the GP everywhere, so the focus is on data points $x = (x_1, \ldots, x_n)$ and additional locations $\tilde{x}$ where predictions may be of interest.

A Gaussian process $\mathrm{GP}(0, k)$ prior to conditioning has a multivariate Gaussian joint density for observations $y$ and $\mu$ at additional locations $\tilde{x}$,

$$\begin{pmatrix} y \\ \tilde{\mu} \end{pmatrix} \sim \mathrm{N}\left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} k(x,x) + \sigma^2 I & k(\tilde{x},x) \\ k(x,\tilde{x}) & k(\tilde{x},\tilde{x}) \end{pmatrix} \right). \tag{3.34}$$

In this case noise variance $\sigma^2$ has been added to the diagonal of the covariance of $\mu$ to get the covariance for $y$. This corresponds to how the GP is used in practice, with a multivariate Gaussian kernel and an additional variance that accounts for uncertainty. It is then possible to calculate other properties such as the expected value

$$\mathrm{E}(\tilde{\mu}) = k(\tilde{x},x)(k(x,x) + \sigma^2 I)^{-1} y \tag{3.35}$$

and covariance

$$\mathrm{cov}(\tilde{\mu}) = k(\tilde{x},\tilde{x}) - k(\tilde{x},x)(k(x,x) + \sigma^2 I)^{-1} k(x,\tilde{x}). \tag{3.36}$$

The clear next question is how does one choose the relevant hyperparameters, since everything else follows relatively straightforwardly. This is done by maximizing the log marginal likelihood (LML), which identical to the Bayes factor for the hyperparameters. The LML is given in closed form by

$$\begin{aligned} p(\mathbf{y} \mid \mathbf{X}, \theta) = &-\frac{1}{2}\mathbf{y}^{\mathsf{T}} \left( k(x,x) + \sigma^2 I \right)^{-1} \mathbf{y} \\ &-\frac{1}{2}\log \left| k(x,x) + \sigma^2 I \right| - \frac{n}{2}\log 2\pi \end{aligned} \tag{3.37}$$

where the first term is for model fitting and the second term penalizes complexity, balancing the potential for overfitting.

In this thesis, GPs are used as surrogates for the much more computationally-intensive hybrid model of heavy ion collisions. The purpose of the surrogate model is to learn the true dependence of the observables on the parameters of the hybrid model so that MCMC can be performed on a reasonable timescale and with realistic uncertainties. One can think of it in analogy to game of telephone: the full physics

model with its uncertainties is the quantity of interest, but the only practical way to access the information in it is via a third party, the GP, whose reliability must be verified. The MCMC queries the GP, which refers to its training to give the best guess – with uncertainty – of what the truth is. In this thesis, a Gaussian (or Radial Basis Function) kernel and a white noise kernel are used for the underlying Gaussian processes. The Gaussian kernel describes the underlying functional behavior while the white noise kernel describes the inherent variability from finite Monte Carlo model evaluations.

#### 3.5.1.1 *Transfer learning*

A surrogate modeling technique only recently applied to heavy ion collisions is transfer learning [4, 138]. This is a method used to learn about a task of interest (the target task) by using information from related tasks (source tasks). In heavy ion collisions, inductive transfer learning, where the source and target have the same input domain, can be readily deployed. This allows for straightforward transfer learning between models of viscous corrections at particlization that do not introduce additional parametric flexibility.

Transfer learning is performed by first having a trained surrogate model for a source task. Then, the discrepancy between the source and the target is found and is encapsulated in a discrepancy function. Finally, the task in transfer learning is to use comparably little new training information about the target to learn the discrepancy function.

More formally, if $f_S(x)$ is the source and $f_T(x)$ the target, then a simple relationship can be identified,

$$f_T(x) = \rho f_S(x) + \delta(x) \tag{3.38}$$

where $\rho$ is a linear correlation between the source and target estimated via maximum likelihood methods and $\delta(x)$ is the discrepancy between the source and target models. $f_S(x)$ and $\delta(x)$ are considered to be independent Gaussian processes. This model for transfer learning was developed from multifidelity emulation, where the source is a computationally-inexpensive low-fidelity model and the target is a computationally-expensive high-fidelity model [139].

Practically, this vastly reduces the computational cost of training new models that are similar to an already-trained target. In the case of viscous corrections, particularly those between Grad's 14-moment viscous corrections and the Chapman-Enskog viscous corrections, an order of magnitude fewer training points are required to reach a specified accuracy when using transfer learning as opposed to training a new surrogate model [4]. This thesis is the first time transfer learning methods will be used in heavy ion collisions for Bayesian inference. I will use transfer learning to implement a second viscous correction model with Grad's 14-moment viscous corrections as the source $f_S$ and Chapman-Enskog RTA viscous corrections as the target $f_T$.

### 3.5.2 *Computer Experiment Design*

Previous studies using uniform priors have sampled the allowed parameter space using maximin latin hypercube sampling (LHS) techniques, which maximize the minimum Euclidean distance between points. Latin hypercubes are designed to provide uniform coverage when projected into 1 dimension while the maximin algorithm helps select points that give a fairly reasonable coverage of the volume.

This is a sub-optimal space-filling sampling technique. The purpose of the sampling is not to give even coverage when projected into 1 dimension or have good coverage only in the full d-dimensional space but rather to provide a uniform coverage of the volume – and of the active subspace of that volume – for the training of Gaussian Process emulators (GPs). Choosing a maximin Latin hypercube requires the specification of a number of samples and is inherently tied to this number of samples; adding additional points is thus an open question.

One approach to rectify these issues is through designs which have criteria that make them straightforwardly extensible, including the use of low-discrepancy sequences such as the Sobol Low-Discrepancy Sequence (LDS) [140]. Low-discrepancy sequences are deterministic sequences of points that rapidly converge to a uniform coverage of an N-dimensional volume. In fact, it can be shown that a Sobol LDS converges to uniform coverage of the space $\sim \mathcal{O}(1/N)$ while the Latin hypercube converges $\sim \mathcal{O}(1/\sqrt{N})$. Further comparisons can be found in a comprehensive study comparing the two [141].

Use of extensible designs produces a variety of advantages. First, it ensures convergence with increasing points and guarantees that points are being placed in such a way that the volume is given uniform coverage. Second, it is efficient and sequentially improvable, meaning that future studies can simply continue to add points to improve the resolution of the data and thus represents an investment in research infrastructure. Third, it ensures that there is always a uniform coverage of the volume, not simply a 1-d projection of this volume and thus is more suited to the task at hand.

However, using the Sobol LDS has a limitation. In relatively few dimensions, converged Sobol sequences can be constructed with relatively few points. The Sobol Low Discrepancy Sequence has two convergence properties, called Property A and Property A'. Property A is met *if for any binary segment (not an arbitrary subset) of the d-dimensional sequence of length* $2^d$ *there is exactly one draw in each* $2^d$ *hypercube that results from subdividing the unit hypercube along each of its length extensions into half.* Property A' is satisfied *if for any binary segment (not an arbitrary subset) of the d-dimensional sequence of length* $4^d$ *there is exactly one draw in each* $4^d$ *hypercube that results from subdividing the unit hypercube along each of its length extensions into four equal parts [140].* These define two convergence thresholds, dependent on the dimensionality of the sample space. In a few-dimensional space, these properties are easy to meet; for example, in 3 dimensions, Property A requires $2^3 = 8$ samples while Property A' requires $4^3 = 64$ samples. This scales poorly to higher dimensions, *e.g.* in 11 dimensions, Property A requires $2^{11} = 2048$ samples. This rapidly becomes impractical. With the available computational budget for this thesis, a design size of 350 training points is the largest possible. I compute the centered discrepancy [142] from uniform sampling for a variety of sampling strategies in Fig. 3.2.

In few dimensions where uniform exploration of the parameter space is the goal, this thesis uses a Sobol low-discrepancy sequence, a first for the field. However, when facing the curse of dimensionality and when maximizing the performance of a surrogate model is of primary interest, the Sobol sequence is poorly suited to the goal and another design must be used.

Recall that the desire was for a space-filling, extensible approach for placing points for maximal surrogate model performance – not necessarily uniformity of coverage. One issue that arises in surrogate model performance is projection behavior: not all

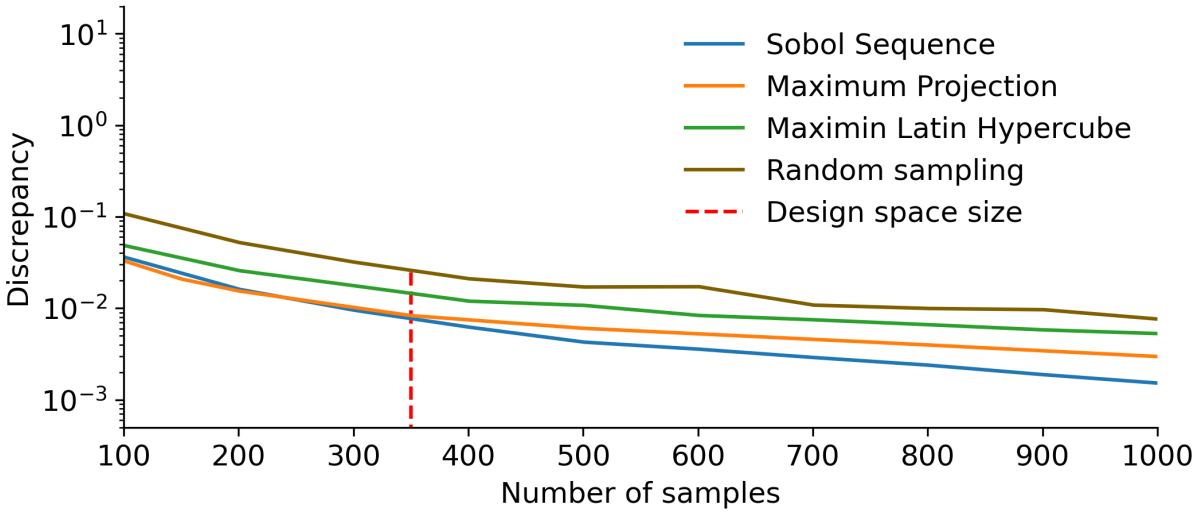


Figure 3.2: Centered Discrepancy from uniformity in a number of different sampling techniques on the 11-dimensional unit hypercube. The Sobol sequence only converges after 2048 samples. At the design space size in this thesis, the Sobol sequence and the Maximum Projection design have comparable discrepancy, but the Maximum Projection is designed for interpolator performance in high dimensions and covers the space.

parameters are equally impactful – some may even have little impact on the final result. As a result, there is a smaller space that impacts the outputs, called the "active subspace". It is not possible to know the active subspace ahead of time, but it is possible to construct a space filling design that maximizes *all* arbitrary projections of the space to lower dimensions. This is the idea behind the Maximum Projection (MaxPro) design strategy [143].

Specifically, this study will utilize a MaxPro Latin Hypercube Design. By focusing on all arbitrary projections rather than only the 1-D projection, a major critique of Latin Hypercubes is resolved in a way more commensurate with the desired properties than naive Maximin designs.[9] Rather than maximizing the minimum distance

9 An additional benefit is that the points can be ordered so that the points that maximize projection behavior can be calculated first, meaning intermediate results are more rapidly useful [3].

between points as in a Maximin design, the Maximum Projection design criteria (or measure) is

$$\min_{D} \psi(D) = \left\{ \frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{1}{\prod_{l=1}^{p} \left(x_{il} - x_{jl}\right)^2} \right\}^{1/p} . \tag{3.39}$$

$p$ is the number of dimensions of the space and $n$ are the number of points in the sample space. Overlapping points are infinitely penalized, as $\psi(D) = \infty$ for $x_{il} = x_{jl}$, $i \neq j$ and so points that maximize distance to another in all projections are preferred. Rather than consider the simple Euclidean distance, the denominator has the product of the squared distances, discouraging two points from approaching each other in any projection. The design that minimizes $\psi(D)$ maximizes the projection behavior in all subspaces for sampled points $x$ [143].[10]

The sampling must also be made commensurate with the parameter ranges and priors used. This is accomplished by sampling designs on a unit hypercube with the relevant number of dimensions. The priors are chosen and the percent point function (or quantile) can be straightforwardly calculated. The sample location on each dimension of the unit hypercube corresponds to a percentile of the prior range in each dimension. This ensures uniform coverage of the probability volume by weighting by the prior density. This deformation technique is shown for a simple 2D example in Fig. 3.3

An additional feature of the Maximum Projection design is that it provides a measure to be minimized. This allows for choosing points such that the first N points minimize a design chosen from the larger subset. This is in contrast to other, unordered techniques and allows for intermediary results as the first N points have maximized the coverage of the space. A comparison for the coverage of a single parameter dimension is shown in Fig. 3.4. As can clearly be seen, the ordered MaxPro design spans the 1-dimensional projection much more quickly than the design which only maximizes the minimum d-dimensional Euclidean distance between design points [3].

10 All designs are generated on the p-dimensional unit hypercube, $[0, 1]^p$ and then scaled and shifted to be commensurate with priors.

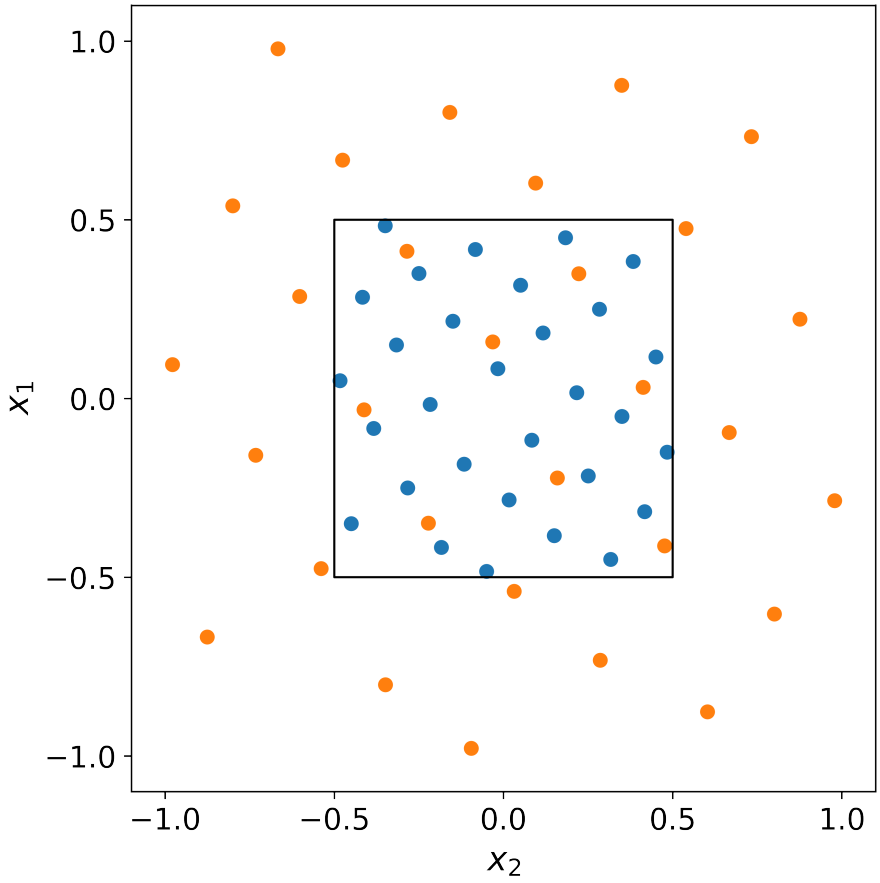


Figure 3.3: Deformation of a 2 dimensional Maximum Projection design on the unit hypercube centred at 0 according to a standard Generalized Normal distribution (Version 1) with $\beta = 10$. The centred unit hypercube is highlighted with a square box.

The ordered Maximum Projection design compares comparably well to designs at each number of points and outperforms other design methods also chosen for the chosen number of points. A comparison between purpose-built MaxPro designs and subsequent subsets from a larger design is shown in Fig. 3.5. This choice of sequential subset allows for intermediary analysis to be run reliably before the full design has been computed, providing substantial speedup in producing physics results.

## 3.6 BAYESIAN MODELING WORKFLOW

The final plank in the statistical foundation of this study is the idea of a Bayesian modeling workflow. In addition to supporting trustworthy inference, it ensures an

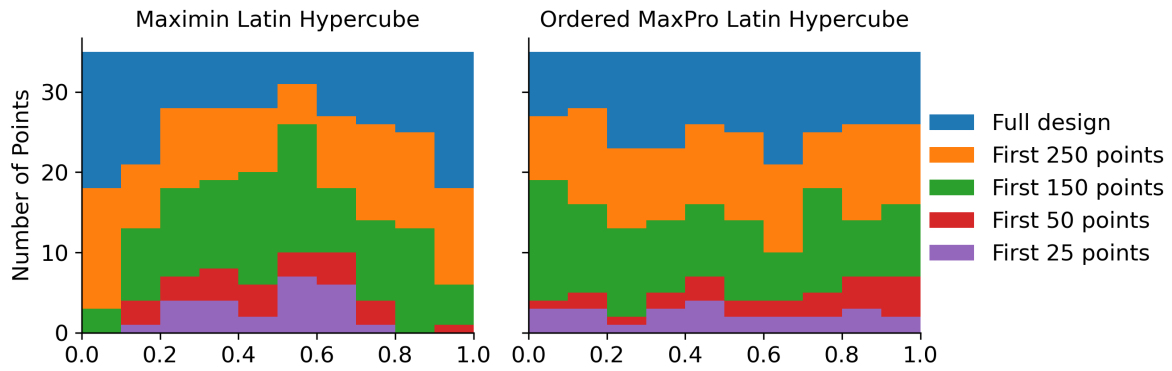


Figure 3.4: Subsequent subsets from a one-dimensional projection of the unit 11-dimensional Maximum Projection design with 350 design points compared to similar subsets for a maximin latin hypercube algorithm [144] used by nearly all other studies of heavy ion collisions. Adapted from [3].

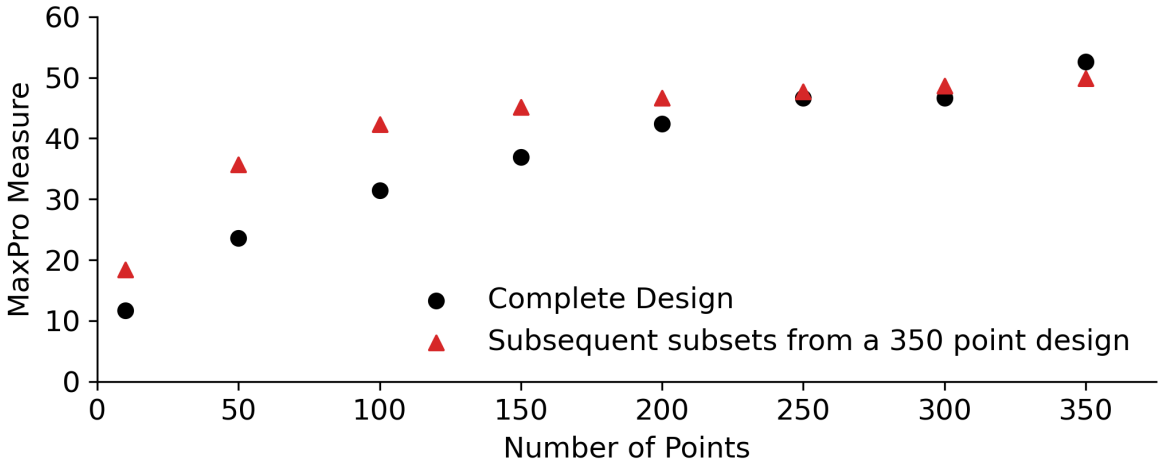


Figure 3.5: Design measure with subsets from am 11-dimensional, 350-point Maximum Projection design compared to Maximum Projection designs chosen for various increasing design sizes.

awareness of the decisions made at each step of the analysis. It also highlights that producing a Bayesian study is not the end of a task - rather, it sets in motion a guided process of refinement and improvement.

In the next section, I deploy a Bayesian modeling workflow to produce quantitative guidance for the simple pendulum in the introductory laboratory. The steps are listed below and details are expanded upon in the next chapter.

Step 1: Define the prior state of knowledge.

Choose or build the model that will be investigated and define principled expectations for prior knowledge of the parameter values.

Step 2: Prior Predictive Checks.

Evaluate the model given the prior knowledge of the parameters defined in the preceeding step. Are these outputs reasonable or is there pathological behavior that suggests mis-specified prior knowledge?

Step 3: Model Validation.

Generate pseudodata with the model in question given known inputs. Can the known inputs be recovered to show that the model is self-consistent?

Step 4: Inference with Data.

Replace the pseudodata with real experimental data. Analyze the results, in particular the differences from expectations gathered with the pseudodata comparison.

Step 5: Posterior Predictive Checks.

Evaluate the model given the posterior knowledge gathered in the preceding step. Are these outputs consistent with the data? If not, are discrepancies interpretable?

Step 6: Identify Future Directions.

Identify avenues for model improvement or physics study. Return to step 1, improve the model and begin the process anew.

# A SIMPLE EXAMPLE: THE PENDULUM

*Note: A reader only interested in heavy ion collisions can skip this chapter without loss of relevant physics; the reader unfamiliar with Bayesian techniques is encouraged to review this example as a simple demonstration before a higher-dimensional application.* This example is lightly adapted from [1]. Preference should be given to citing that work whenever possible. At time of writing, the preprint is in the process of peer review.

Introductory physics courses commonly teach that the period of a simple pendulum displaced by a small angle $\theta$ may be found by applying the approximation $\sin(\theta) \approx \theta$, which holds when $\theta$ is small. However, the period of a pendulum also depends on the initial angular displacement. Attention has so far been focused on more closely approximating the full non-analytic integral expression for the period so students in undergraduate laboratories can investigate the dependence of the period on initial angular displacement [145]. Recently, model comparison and complexity has been applied to the pendulum problem or the teaching laboratory [146].

Thus far, guidance for the use of the small angle approximation in simple pendula has not been rigorously quantified using realistic uncertainties measured in teaching laboratories. Textbooks often state that after $\sim 15^\circ$, the difference between $\sin(\theta)$ and $\theta$ in radians exceeds 1% and the small angle approximation should no longer be used [147]. This does not take into account the variety of measurement uncertainties in practice or if the data demands additional complexity. In this demonstration, I show that with both an exact and approximate analytical formula for the period of a pendulum can successfully establish a quantitative preference at moderate displacements using Bayesian tools provided sufficiently accurate timing methods are available. This demonstration produces quantitatively-motivated guidance for restricting the initial angular displacement or employing more complicated formula for the first time.

## 3.7 THE PERIOD OF A SIMPLE PENDULUM

An ideal simple pendulum is comprised of a bob with mass $m$ suspended from a frictionless attachment point by a massless rod of length $L$. The motion of the pendulum is localized in a single angular plane $\theta$. In the remainder of this example, non-conservative forces such as friction and drag are explicitly excluded. The equation of motion of the ideal simple pendulum is

$$\ddot{\theta} = -\frac{g \sin(\theta)}{L} \tag{3.40}$$

where $g$ is the gravitational acceleration. For small displacements, $\sin\theta \approx \theta$ and Eq. 3.40 is a simple harmonic oscillator. The period of a simple harmonic oscillator is $T = 2\pi\sqrt{\frac{m}{k}}$ and, for the simple pendulum, $k = \frac{mg}{L}$. This results in

$$T_0 = 2\pi\sqrt{\frac{L}{g}}, \tag{3.41}$$

in which there is no angular dependence. Thus, when propagating uncertainty, only the uncertainty on the length is propagated through to the final period.

To derive an exact expression for the period of a pendulum, one can begin by considering the conservation of energy. An exact expression for the period of a simple pendulum is readily derived using the conservation of energy. By choosing the zero of potential energy when $\theta(t) = 0$ and assuming the pendulum to be initially stationary and at its initial angular displacement $\theta_0$ makes this process straightforward. The energy conservation equation is $mgL(1-\cos\theta_0) = \frac{1}{2}mL^2\dot{\theta}^2 + mgL(1-\cos\theta)$. It is then straightforward to solve for $\dot{\theta}$ and integrate $\theta$ from $0$ to $\theta_0$, corresponding to one-quarter of the period. $T = 2\sqrt{2}\sqrt{\frac{L}{g}}\int_0^{\theta_0} \frac{1}{\sqrt{\cos\theta - \cos\theta_0}} d\theta$. This integral is improper when $\theta = \theta_0$, but the substitution $\cos\theta = 1 - 2\sin^2(\theta/2)$ may be combined with a change of variables $\sin\phi = \frac{\sin(\theta/2)}{\sin(\theta_0/2)}$ to yield

$$T = 4\sqrt{\frac{L}{g}}\int_0^{\pi/2} \frac{1}{\sqrt{1 - \sin^2(\theta_0/2)\sin^2(\phi)}} d\phi \tag{3.42}$$

This can then be evaluated numerically [145, 148]. Uncertainty on the initial angular displacement $\theta_0$ is propagated through the integral expression, assuming that the uncertainties are Gaussian and uncorrelated. The exact expression contains both the

length and angular displacement, so must account for the measurement uncertainty on both. The small angle approximation (Eq. 3.41) and the exact expression (Eq. 3.42) are the two models that will be considered in this example.

### 3.7.1 *Generating realistic measurements*

In order to determine when the small angle approximation becomes insufficient using realistic measurements, realistic choices must be made. Using real measurements in order to determine model preference thresholds demands a sufficiently large number of measurements as to be impractical. For 5 measurements of 10 periods at 20 displacements each with 3 timers, a single maximum angular displacement can require 300 measurements, so some of the data in this example is generated using Eq. 3.42 and assigned realistic uncertainties determined in this section.

Generated data used in model validation is called "pseudodata" and real measurements are performed for the final inference. Generated data is used again to calculate model preference thresholds as this is a particularly data hungry process. Pseudodata is generated for model validation by making a set of simple choices informed by standard guidance in undergraduate teaching laboratories. In these laboratories, students are told that a rule of thumb for measurement uncertainty is equal to half of the smallest increment of the measurement device. A common measuring device in teaching laboratories is a meter stick with millimetre gradations, corresponding to an uncertainty of $\pm 0.0005$ m. In this example, the length of the pendulum was measured to be $L = 0.807 \pm 0.0005$m. The uncertainty on initial angular displacement corresponds to the use of digital protractor with $0.01^\circ$ gradations and thus all angular model inputs have uncertainty $\pm 0.005^\circ$.Uncertainty is propagated from the measured inputs assuming independent (or uncorrelated) variables, where for a function $f$ of variables $a$, $b$, etc., the variance of the model prediction is

$$\sigma_f^2 = \left(\frac{\partial f}{\partial a}\right)^2 \sigma_a^2 + \left(\frac{\partial f}{\partial b}\right)^2 \sigma_b^2 + \dots . \tag{3.43}$$

More precise timing mechanisms such as photodiodes may also be available and the use of smartphone accelerometers in teaching laboratories has become more common. These measurements are significantly more precise than the precision of a stop-

watch and bypass human reaction times. Measurements with smartphone accelerometers in recent courses at McGill had uncertainty of $\pm 0.02$ s [149], while timing uncertainties with the photodiode timer I used to measure data for this example are smaller still at approximately $\pm 0.005$ s. Finally, for timing uncertainties corresponding to the use of a stopwatch in model selection is taken to be $0.250\sqrt{2}$ s, corresponding to a reaction time of 0.250 s [150] to start and stop the timer added in quadrature. Ten periods were timed to reduce the relative error of the measurements.

It is possible to specify the initial angular displacement and produce pseudodata that reflects the real precision of data, which can in turn be used to to assess self-consistency. Once the models are shown to be self consistent, they are compared to data. If a model describes the data sufficiently well, then it can be used to generate realistic-enough pseudodata for general guidance. While a simple calculation is sufficient to show that the models become increasingly differentiable as $\theta_0$ increases, this is a tractable example to demonstrate Bayesian model selection.

## 3.8 A WORKFLOW FOR REPRODUCIBLE INFERENCE

Following a step-by-step process acts as a safeguard that supports the validity of the analysis – or straightforward troubleshooting – and a rich understanding of the models. It also establishes practices for reproducibility in analysis. In both research and pedagogical settings, clearly-defined steps are used to guide analysis and allow for careful implementation and refinement. This simple example is an ideal testbed for demonstrating this workflow before applying it to more involved research problems. At the end of the workflow, when both models have been validated and calibrated to data, Bayesian model comparison is used to establish rigorous criteria for when the data demands one model over the other.

### 3.8.1 *Step 1: Defining the prior state of knowledge*

The first step in Bayesian data analysis is to clearly motivate the choice of prior given the models in hand. The prior should reflect realistic understanding of the parameters, with a goal of accurately reflecting a physical understanding blind to the data in

question. This then allows for posterior constraint to be determined by the systematic comparison of the model to the data via the likelihood [151].

The only free model parameter in both the equations using the small angle approximation and the exact expression for the period of a pendulum is the gravitational acceleration $g$. It must be positive definite and nonzero, otherwise no oscillation will occur. Physical intuition may also be applied to the prior. Gravitational acceleration on the surface of a body scales with the mass of the body and gravitational acceleration on the Moon is approximately 1.625 $m/s^2$ [152] while the surface gravity of Jupiter is 24.79 $m/s^2$ [153]. Proceeding with confidence that the mass of the Earth is likely between that of the Moon and Jupiter, a weakly-informative prior can be constructed that allows for some probability of surprise. A common choice for a weakly-informative prior on positive definite quantities is the inverse gamma distribution. It has support on the interval $(0, \infty)$, matching the positive definite constraint and is easily tuned so that only 1% of the probability falls below 1.625 $m/s^2$ and above 24.79 $m/s^2$. This prior is shown in Fig. 3.6. The location of the mode of the distribution is not necessarily relevant – by constructing a weakly-informative prior, the data is allowed to speak for itself. Other choices of prior that have similar support are possible, but the inverse gamma distribution has the nice feature of smoothly approaching 0 on the bounds of its support, which is consistent with physical expectations.

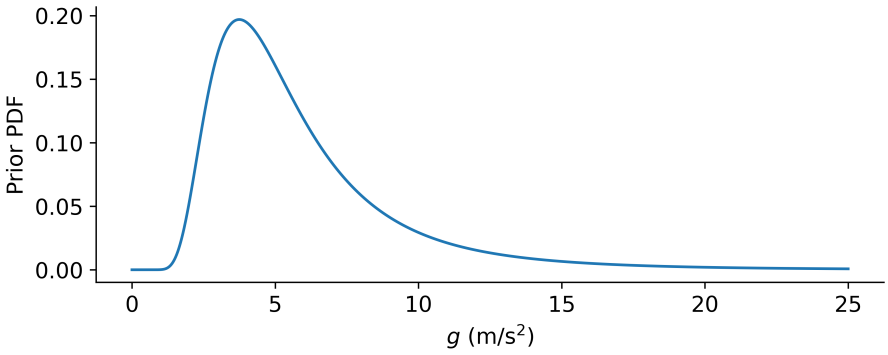


Figure 3.6: The prior probability density function (PDF) of the gravitational acceleration $g$. This PDF is $p(H|I)$ in Eq. 3.6 where each value of $g$ is a hypothesis $H$.

### 3.8.2 *Step 2: Prior predictive checks*

Exploring the parameter space is an important step in data analysis to understand any system. Prior predictive checking provides an understanding of model behavior because it can reveal interesting or improper features of the model [154]. The model is evaluated with draws from the prior distribution to form the prior predictive distribution, which is is used to assess the model's behavior. A non-trivial feature could be an interaction between two components of a multi-component model that yields unphysical results and should be excluded by a more carefully constructed prior. In this example, few such features exist, but the step is included for completeness of the demonstration. Predictions for the period with $g$ sampled from the prior distribution are shown in Fig. 3.7. It is expected that measured periods of this pendulum will be between approximately 1 and 5 seconds, while longer and shorter periods and corresponding values of $g$ are still theoretically possible, but implausible.

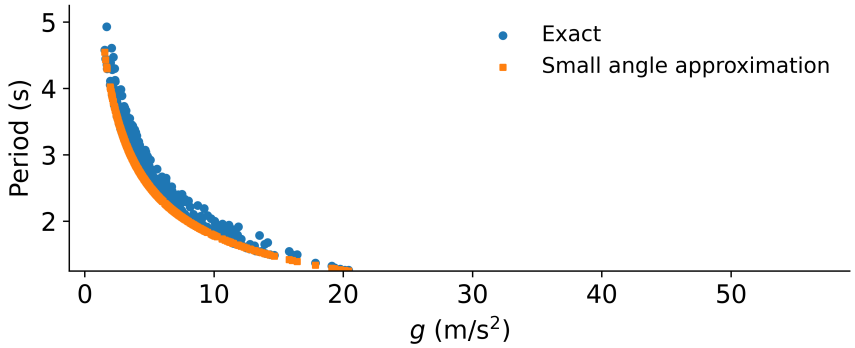


Figure 3.7: 100 predictions for the period for both models, with $g$ sampled from the prior and initial angular displacement $\theta_0$ sampled from a uniform distribution $U(0, \pi/4)$.

### 3.8.3 *Step 3: Model validation*

Testing for self-consistency is a means of establishing model validity. This could also be called "inverse validation", as it requires solving the inverse problem and comparing the result to known inputs. This is an important step, but also represents the best-case scenario: the model is known to contain all the features of the pseudodata. While this is straightforward in a simple example, it is a critical feature in more

intensive numerical applications with sophisticated observables such as heavy ion collisions.

Pseudodata is generated using the NIST reference value $g = 9.80665\ m/s^2$ [155] and a timing uncertainty corresponding to the use of a photodiode timer over 10 periods. Initial displacements are chosen with arbitrary spacing, primarily due to the difficulty in regular spacing on the experimental apparatus.

$$\theta_0 = [2.86°, 11.46°, 20.05°, 22.92°, 35.98°]$$

Calculations of the corresponding periods are shown in Fig. 3.8. The arbitrarily-spaced displacements exaggerate features of real measurements rather than provide an unrealistically-even test case.

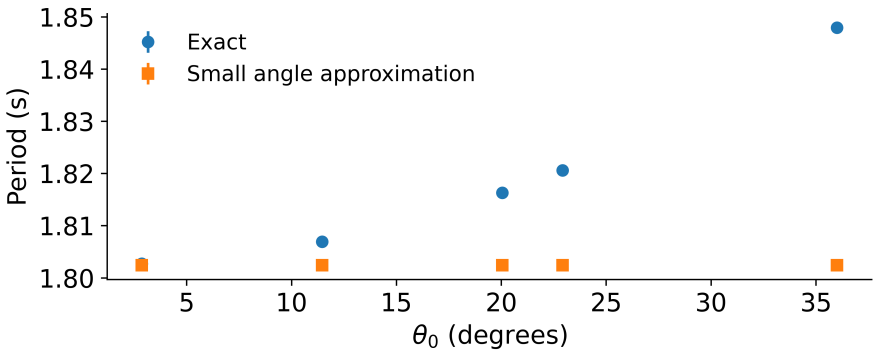


Figure 3.8: Pseudodata used for model validation. In model validation, models are compared to data generated by the model under consideration.

The models are then both calibrated to their respective validation pseudodata (Fig. 3.8) and the posterior distributions are shown in Fig. 3.9. Both models are found to be self-consistent. The small angle approximation model's posterior is slightly more peaked than that of the exact model because the predictions are more precise. This increased precision arises as the small angle approximation does not account for uncertainty on $\theta_0$. For each hypothesized parameter value, the Gaussian likelihood is calculated and the posterior is trivially found. While this is simple in a small number of dimensions, this is impractical for high-dimensional applications with expensive models due to computational cost. To overcome this, the posterior is evaluated numerically using Markov Chain Monte Carlo. This is used here here for consistency with later analysis.

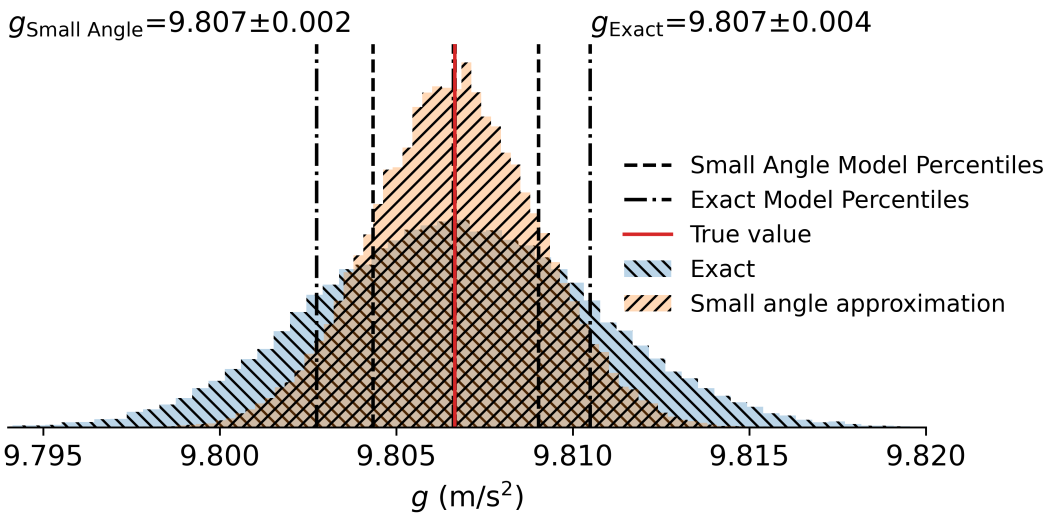


Figure 3.9: Posterior distributions for g in model validation using the exact formula for the period and the small angle approximation. The true value is shown in red and vertical dashed lines denote the 16th, 50th, and 84th percentile of the samples in increasing g with values provided above the figure. When the models contain the full information present in the data to which they are compared, both are able to recover the true value.

Posterior predictive checks help ensure that the posterior is consistent with the data. Predictions are made by taking draws from the posterior distribution and evaluating the model to form the posterior predictive distribution, shown in Fig. 3.10. Both models clearly produce predictions consistent with their respective validation data. A mismatched posterior predictive distribution can indicate an issue in the inference, the MCMC, or may reveal underlying structure contrary to expectations but useful for improving the model. For example, the posterior predictive distributions may reproduce some – but not all – data, indicating that there is missing physics and what that missing physics may be. Where the models are calibrated to data generated by themselves, both are able to recover the known input (Fig. 3.9) and reproduce pseudodata (Fig. 3.10). The distribution of model predictions are shown with violin plots. In violin plots, the width of the regions corresponds to the probability density of the data.

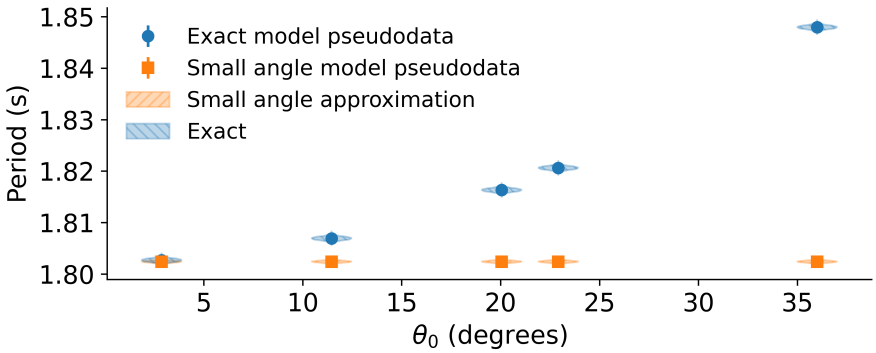


Figure 3.10: Posterior predictive distributions for g from model validation. The posterior distributions are those shown in Fig. 3.9.

### 3.8.4 *Step 4: Inference with data*

Only after self-consistency has been established is it appropriate to move on to comparison with real data. Measurements were made with a pendulum of length $L = 0.807 \pm 0.0005$ m and a cylindrical bob of mass $10 \pm 0.05$g with times recorded using a photodiode connected to a Pasco 850 universal interface Model UI-5000 for data acquisition. The bob was released from the same location using a stand. Five measurements of ten periods were made at each arbitrary angular displacement, consistent with the typical number of measurements performed by students in introductory teaching laboratories. The means of these measurements and associated uncertainty are provided in Table 3.2.

| $\theta_0$ (deg) $\pm 0.005°$ | Period (s) $\pm 0.0001$ s |
|---|---|
| 6.40 | 1.8028 |
| 11.73 | 1.8056 |
| 16.41 | 1.8113 |
| 19.99 | 1.8152 |
| 23.18 | 1.8196 |
| 32.67 | 1.8339 |

Table 3.2: Experimental results for the period of the simple pendulum.

The posterior for the gravitational acceleration g is shown in Fig. 3.11. The exact model is able to infer the gravitational acceleration with a relatively accurate and precise 68% credible region – $9.818 \pm 0.003$ m/s$^2$ – while the small angle approximation infers a result with higher precision – $9.715 \pm 0.001$ m/s$^2$ – but sacrifices accuracy with a remarkable $\sim 0.1$ m/s$^2$ bias from the standard value 9.80665 m/s$^2$ [155] and from 9.80636 m/s$^2$, the best-available model calculation of g for the experimental apparatus' location [156]. While in Fig. 3.9, both models were self consistent and recovered results within expectations, one can clearly see in Fig. 3.11 that the small angle approximation's posterior is less consistent with expectations than that of the exact expression for the period.

Increased precision at the cost of bias is commonly known as the bias-variance tradeoff. Addressing this tradeoff is a focus of model development and many techniques in statistical learning [157]. While the reference value is not within the 68% credible region of the exact model, the bias is greatly reduced and predictions with the posterior are investigated further in the next subsection and match the data well.

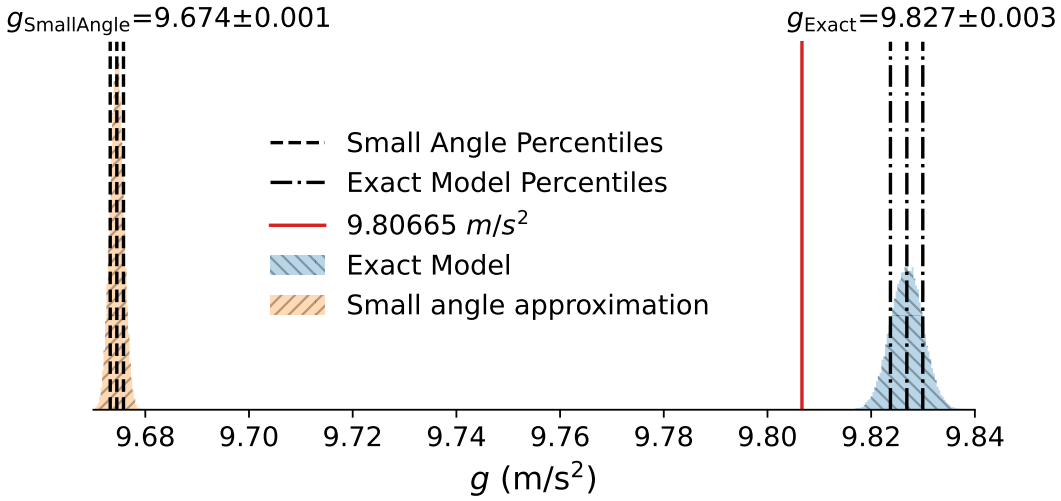


Figure 3.11: Independently normalized posterior distributions for g using the exact formula for the period and the small angle approximation. The NIST value [155] is shown in red and vertical dashed lines denote the 16th, 50th, and 84th percentile of the samples in increasing g with exact values provided above the corresponding distribution. Because the two posteriors are independently normalized, it is important to consider each posterior independently and not interpret the relative height of the peaks.

### 3.8.5 *Step 5: Posterior predictive checks*

Posterior predictive distributions can elucidate features of the predictions and provide insights into why one model may fail why the other succeeds. The posterior predictive distributions are shown in Fig. 3.12. The failure of the small angle approximation to reproduce the measured data is striking. As it was shown in Fig. 3.10 that the small angle approximation is self-consistent, the only conclusion can be that important physics is absent, *e.g.* angular dependence. The predictions from the exact model posterior are constrained around the measurements and are broadly consistent with the measured uncertainty, while the small angle model attempts to match the data by undershooting at large angular displacement and overshooting small displacements. It thus fails to reproduce the data outside a small intermediate window. As a result, the posterior predictive distributions make the incompleteness of the small angle approximation immediately apparent and interpretable, while the failure of the exact model requires more attention. Features not yet accounted for include the flattening of the data at large displacement in Fig. 3.12, likely due to non-conservative forces such as drag. Accounting for such effects is beyond the scope of this example. This establishes that it is possible to detect the bias in calculations of gravitational

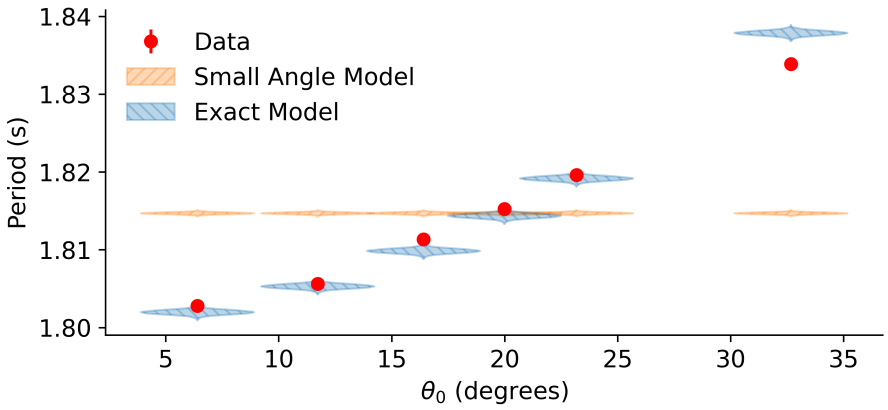


Figure 3.12: Violin plot of the posterior predictive distributions for g. The posterior distributions are those shown in Fig. 3.11.

acceleration introduced by using the small angle approximation of the period of a pendulum while including large initial angular displacements and realistic measurement uncertainties. Through the use of a modeling workflow, it is straightforward

to identify and interpret this bias through visualization of the posterior and posterior predictive distributions. It is also clear that the exact formula for the period of a simple pendulum contains much of the relevant physics necessary to describe real data.

## 3.9 MODEL COMPARISON

It is time to translate this into quantitative guidance using the tools of Bayesian model comparison. In the remainder of this chapter, the exact formula and the small angle approximation for the period are subscripted 0 and 1, respectively. Taking sufficient accurate measurements by hand involves thousands of measurements of the period and it has already been demonstrated that the exact expression of the period describes the motion of a pendulum well. Rather than perform an unreasonable number of measurements for a simple demonstration, the model selection study will use generated data. It is unreasonable to assume that a student in an introductory laboratory would be able to make reliable measurements at increments finer than $1^\circ$ with a meter-scale pendulum or at more than 12 increments in initial angular displacement $\theta_0$. This represents an attempt to find an idealized, maximally-restrictive constraint on initial angular displacement $\theta_0$.

12 proxy data points linearly spaced in $\theta_0$ are generated if the maximum $\theta_0$ is larger than $12^\circ$. If the maximum $\theta_0$ is less than $12^\circ$, data is generated in $1^\circ$ increments. The Bayes factor is calculated and interpolated to determine at what $\text{Max}[\theta_0]$ the evidence for the exact model becomes weak, moderate, or strong. Including these results demonstrates the importance of accounting for uncertainty when offering data-driven guidance and will allow for appropriate implementation.

The ln Bayes factors are shown in Fig. 3.13. As the maximum initial angular displacement $\theta_0$ increases, there is progressively stronger preference for the exact expression. The constraint diminishes as the relative magnitude of the angular displacement is reduced and the displacement is increasingly small. Preference thresholds for the exact calculation at maximum initial angular displacement $\theta_0$ is given in Table 3.3. The small angle approximation is appropriate below these model preference thresholds.

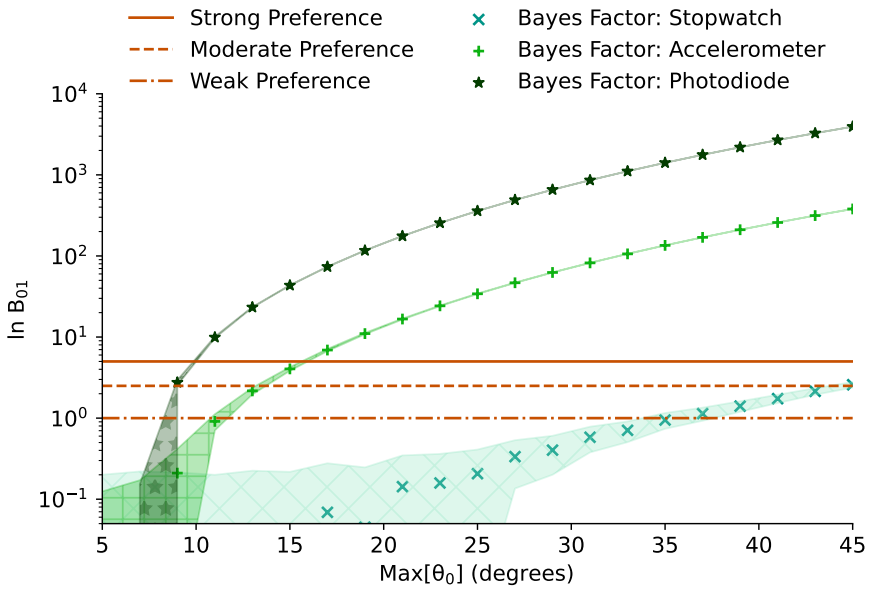


Figure 3.13: $\ln B_{01}$ demonstrating preference for the exact expression for the period of the pendulum over the small angle approximation for a variety of timing methods. Positive values denote preference for the exact model. Horizontal lines denote thresholds for weak, moderate, and strong preference and the filled areas represent the integration uncertainty on the Bayes factors arising from numerical calculation of the Bayes evidence. When timing is imprecise, it becomes difficult to establish model preference.

| Strength of evidence | $\lvert \ln B_{01} \rvert$ | $\theta_0$ (rad) | $\theta_0$ (deg) |
|---|---|---|---|
| Photodiode timing | | | |
| Weak | 1.0 | $0.140 \pm 0.003$ | $8.0^{+0.1}_{-0.2}$ |
| Moderate | 2.5 | $0.155 \pm 0.002$ | $8.89^{+0.09}_{-0.1}$ |
| Strong | 5.0 | $0.171 \pm 0.001$ | $9.81 \pm 0.06$ |
| Accelerometer timing | | | |
| Weak | 1.0 | $0.195 \pm 0.007$ | $11.2 \pm 0.4$ |
| Moderate | 2.5 | $0.234 \pm 0.004$ | $13.4^{+0.2}_{-0.3}$ |
| Strong | 5.0 | $0.275 \pm 0.003$ | $15.8 \pm 0.2$ |
| Stopwatch timing | | | |
| Weak | 1.0 | $0.62^{+0.04}_{-0.03}$ | $35 \pm 2$ |
| Moderate | 2.5 | $0.78 \pm 0.02$ | $44.6 \pm 0.9$ |

Table 3.3: Model preference thresholds in which the exact formula is preferred over the small angle approximation. Subscript 0 denotes the exact formula and subscript 1 denotes the small angle approximation. Strong preference was not found for stopwatch timing when maximum angular displacement was limited to $\pi/4$ rad.

| Timing Precision | Recommended Restriction | |
|---|---|---|
| | (rad) | (deg) |
| Photodiode ($\pm \sim 0.005$ s) | 0.17 | 10 |
| Accelerometer ($\pm \sim 0.02$ s) | 0.26 | 15 |
| Stopwatch ($\pm \sim \sqrt{2}(0.250)$ s) | $\pi/4$ | 45 |

Table 3.4: Guidance for restricting initial angular displacement for simple pendula when using the small angle approximation.

## 3.10 CONCLUSIONS

This example demonstrates how to approach statistical modeling through a workflow and how to set data-driven constraints on the use of the small angle approximation using Bayesian inference.

The heuristic guidance of constraining simple pendula to 15° is excessively restrictive when timing with the typical precision found in stopwatches, but is reasonable when using a smartphone accelerometer and overly-permissive when using more precise timing methods. The data does not exhibit strong preference for the exact expression for the period with photodiode timing until $9.81 \pm 0.06°$ and with accelerometer timing until $15.8 \pm 0.2°$, while stopwatch timing is insufficiently precise to establish strong preference between the two models at maximum displacements up to $\pi/4$ radians (45°). This suggests that current guidance of restriction to $\sim 15°$ is appropriate for timing of precision $\pm\mathcal{O}(0.01)$ s, but displacements with more precise timing $\pm\mathcal{O}(0.001)$ s should be restricted to 10° while displacements with less precise timing commensurate with stopwatches and human reaction times should not be restricted below 45°.

This analysis represents the current state-of-the-art guidance for laboratory manuals and textbooks as it has used realistic best-case uncertainties and advanced statistical techniques to determine when the breakdown of the approximation is actually detectable. Future works should continue with the present guidance, but modify their justification to be data-driven rather than being motivated by a 1% discrepancy in the small angle approximation. This would both be more directly applicable to instructors and more faithfully communicate the derivation of scientific guidance to students. Additionally, this has demonstrated a statistical modeling workflow in which more advanced modeling and guidance questions may be addressed.

# Part III

# QUANTIFYING THE QUARK-GLUON PLASMA

Bayesian tools are applied to modeling heavy ion collisions; first in a limited case and then in a global Bayesian analysis. Details of the modeling process are discussed before results are provided and predictions are made.

# 4

# MOTIVATION AND DESIGN

One question at the outset of any study is to ask what one hopes to learn from performing it. In this case, Bayesian analyses with T$_R$ENTo have been performed many times, often with relatively small changes between them. A new analysis, replacing the parametric T$_R$ENTo + freestreaming pre-equilibrium stage with IP-Glasma's microscopic physics, is justified by the clear successes of IP-Glasma-based models in the literature [28, 29, 57].

I will also motivate a new analysis with reference to the current state of the field and by detailed quantitative comparison to the hybrid model used in most Bayesian studies of heavy ion collisions. I first demonstrate important physical differences between the current standard hybrid model with a T$_R$ENTo + freestreaming initial state and a more physically-derived initial state model – notably, IP-Glasma. I show that these differences are of physical importance to modeling heavy ion collisions and outline that the current modeling standard is insufficient for accurate inference of the properties of the QGP. With these differences established through detailed comparison, I design a new, IP-Glasma based study to infer the properties of the QGP through more physically interpretable simulations and model-to-data comparison.

A major reason to motivate these differences is not only to demonstrate increased interpretability of physically-derived models, but rather the adage that "hydrodynamics is agnostic." In other words, hydrodynamics is a deterministic process.[1] and thus constitutes an initial value problem. Assuming two models can both reproduce the same initial values, hydrodynamics will evolve them in the same way and it will be impossible to distinguish between them. Thus, it is important to unambiguously differentiate the CGC-based, QCD-inspired IP-Glasma from the standard T$_R$ENTo + freestreaming implementation to physically justify a new study.

1 Hydrodynamic fluctuations are beyond the scope of this thesis. For more on hydrodynamic fluctuations in heavy ion collisions, see [30].

First, I demonstrate that a TRENTo + freestreaming initial state does not reproduce important details of an IP-Glasma initial state at hydrodynamic starting time. Then, I argue that in order to have confidence that physical quantities are extracted rather than merely useful model parameters, a physical initial state model must be used. I then highlight that known sources of theoretical uncertainty arise from the coupling of the initial state and the hydrodynamic stage.

In this chapter, I begin with a brief overview of TRENTo to establish the current standard of modeling in Bayesian inference in heavy ion collisions in Sec. 4.1. I then compare TRENTo and IP-Glasma energy deposition in Sec. 4.2. Next, I compare various aspects of TRENTo + freestreaming initial states and IP-Glasma directly using initial state pseudo-observables in Sec. 4.2.1. In the same section, I highlight systematic impacts of these differences on the subsequent evolution and the implications for Bayesian studies. I also list some physically-motivated and methodological concerns and expand to issues in previous Bayesian works. I show comparisons using final-state observables with differentiable features of IP-Glasma and TRENTo + freestreaming in Sec. 4.3 and highlight proper comparisons as well as particularly constraining observables that have so far been unused. I include some brief notes on the implementation before describing what will be learned from this study. I also describe improvements compared to previous analyses as well as a few potential extensions. Finally, to conclude this chapter, I explicitly describe parameters and parametrizations as well as their physical meaning in Sec. 4.5.4. These are the quantities which will be constrained through systematic Bayesian model-to-data comparison.

## 4.1 INITIAL STATE MODELS

The goal of every Bayesian study in HICs is to accurately infer properties of the strongly-interacting QCD medium. To do this, the initial stage must be described. However, there are many models of this pre-equilibrium stage. Physically-inspired parametric models have achieved wide use and include Glauber models [158] and TRENTo [37]; transport models such as AMPT [159] used to produce initial states for hydrodynamics; and saturation-based models such as EKRT [160], KLN [161], and IP-Glasma [54].

Due to both availability and flexibility, the model nearly-exclusively used in modern Bayesian analyses in heavy-ion phenomenology is T$_R$ENTo. To allow for completeness of discussion, I provide a brief introduction of the parametric T$_R$ENTo model before continuing to comparisons with IP-Glasma.

T$_R$ENTo, or Thickness (Reduced) Event-by-event Nuclear Topology, is a parametric model of the initial stage of a heavy-ion collision.[2] This model seeks to use a continuously-varying parameter $p$ to modulate the energy deposition into the initial state and introduces other parameters for flexible modeling of the initial energy deposition. It has also been used to "interpolate" between various features – such as geometry – of commonly-used models, including IP-Glasma, MC-KLN, EKRT, and the wounded nucleon model, although these comparisons must be made cautiously [2].

The motivation for a simplified model is that microscopically-derived models of the initial state are computationally intensive and therefore take a significant amount of resources to employ. They are also inflexible, meaning that they effectively "bake in" prior beliefs.[3] T$_R$ENTo is capable of running orders of magnitude more events than IP-Glasma in a comparable time while also enjoying a degree of quantitative success and generating event-by-event energy density profiles.

To provide some mathematical detail, imagine that two colliding heavy ions each can be assigned some "thicknesses" $T_A$ and $T_B$ denoting the amount of nuclear material at a particular position. When these two thickness functions collide (or "participate" in a collision), they deposit energy into a reaction plane proportional to some function of the participant thickness functions. This deposited energy is called the reduced thickness function as the two participant thickness functions are "reduced" to a third thickness. T$_R$ENTo *angesetzt* that this can be modeled by a function $T_R$ of the form

$$T_R\left(p; T_A, T_B\right) \equiv \left(\frac{T_A^p + T_B^p}{2}\right)^{1/p} \tag{4.1}$$

2 As T$_R$ENTo's usage has evolved, model interpolation has been de-emphasized and much effort has gone into differentiating it from models it was designed to interpolate between [162] and must be understood as highly conditional on other choices [2].

3 Flexibility is a double edged sword: because parametric models are not necessarily physical by construction, they can produce unphysical outcomes.

which may easily be recognized as the generalized mean. Depending on the value of $p$, one produces $\max(T_A, T_B)$ ($p \to \infty$), the arithmetic mean $(T_A + T_B)/2$ at $p = 1$, the geometric mean $\sqrt{T_A T_B}$ at $p = 0$, the harmonic mean $2T_A T_B/(T_A + T_B)$ at $p = -1$, and finally $\min(T_A, T_B)$ at $p \to -\infty$. The impact of the variation of $p$ is shown in fig. 4.1.

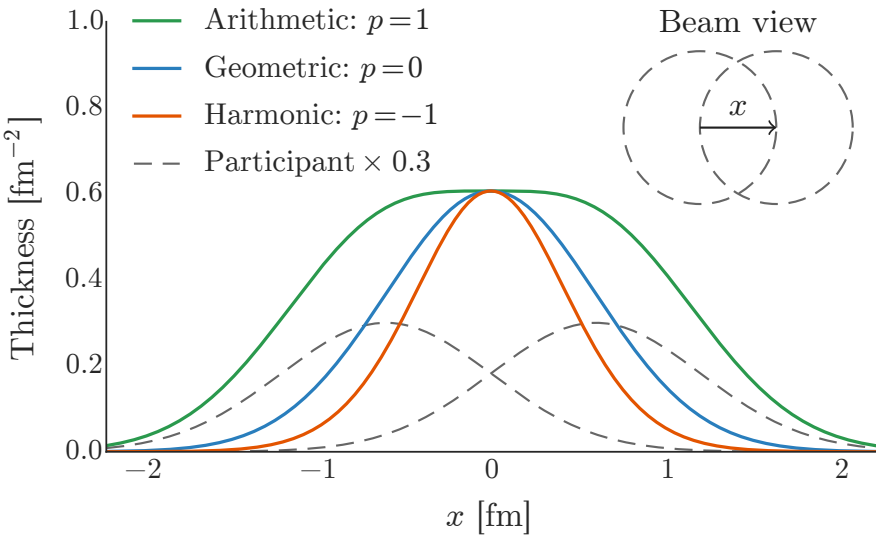


Figure 4.1: TRENTo thickness functions for different values of the parameter $p$, reproduced from [37].

The individual thickness functions $T_A$ and $T_B$ also deserve some attention for the upcoming discussion. At least in 2D TRENTo, these are functions of only the transverse coordinates $x$ and $y$. The nuclear density of two colliding protons A and B centred along the $x$ axis is given by

$$\rho_{A,B} = \rho_{\text{proton}}(x \pm b/2, y, z) \tag{4.2}$$

where the protons collide with some impact parameter $b$. A collection of nucleons colliding similar to these protons may be complied into a total thickness function for nucleus A or B where each proton is also given a fluctuating weight, $w$, itself an element of the set of weights $w_{A,B}$.

$$T_{A,B}(x, y) = w_{A,B} \int dz \rho_{A,B}(x, y, z). \tag{4.3}$$

These weights $w_{A,B}$ are independent random weights sampled from a unit mean gamma distribution,

$$P_k(w) = \frac{k^k}{\Gamma(k)} w^{k-1} e^{-kw}, \tag{4.4}$$

and are introduced to reproduce large multiplicity fluctuations in proton-proton collisions. The shape parameter $k$ is tuned by fitting to data. Additional physical parameters include the nucleon width, the minimum nucleon-nucleon distance for Woods-Saxon sampling of nuclei positions, and the overall normalization of the energy (or entropy) deposition.

Before continuing to comparisons with IP-Glasma, it is worth issuing some words of caution when using parametric models. In the case of T$_R$ENTo, parameters such as nucleon size have strong degeneracy with follow-up stages which diffuse hotpots. These can thus compensate for each other and T$_R$ENTo's relatively diffuse energy (entropy) density deposition results in less constraint about the knowledge of the length of a more physical pre-equilibrium stage. In fact, substantial covariance with hydrodynamic quantities has been shown [42, 45].

T$_R$ENTo does not provide an initial full $T^{\mu\nu}$ nor does it generate flow, so an additional step is often used between the initial energy density and the onset of hydrodynamics. Typically, a freestreaming model [163, 164] is used. This introduces another opportunity for unquantified uncertainties as freestreaming is known to be an incomplete approximation to pre-equilibrium dynamics. In freestreaming, the energy distribution is initialized with an isotropic momentum distribution and allowed to evolve without interactions. In heavy ion modeling, freestreaming is often initialized using the T$_R$ENTo energy (or entropy) density.

A physical problem with connecting freestreaming to hydrodynamics directly is that freestreaming implies an infinite mean free path (no interactions). Hydrodynamics evolves a strongly-coupled medium whose mean free path is very short. This is a source of unquantified theoretical uncertainty in hybrid models with this discontinuity. A method to more smoothly connect initial states to hydrodynamics can be found in [165], although more detailed considerations are beyond the scope of this thesis.[4]

A phenomenological concern is that, when followed by freestreaming, the hot spots in T$_R$ENTo diffuse further. This has a clear degeneracy with the nucleon size parameter and, if evolved long enough, can significantly impact the geometry at the onset of hydrodynamics. This suggests an additional compensating effect between the

4 A new model, KTiso [166], has been proposed that is more physically-inspired than freestreaming and appears promising, but has yet to be used at scale.

freestreaming time and various TRENTo parameters and will be explored further in Sec. 4.2.1. TRENTo attempts to provide a sufficiently flexible model to match data by compiling features and expecting that the parametric flexibility is able to partially quantify the uncertainty on the initial stage and its impact on later stages of the model. The success of this will be investigated in Secs. 4.2 and 4.2.1.

## 4.2 INITIAL-STATE COMPARISON

Parametric, geometric models sacrifice microscopic physics in order to capture broad features in an efficient, intuitive way. The goal of the entire program of statistical analysis in heavy ion collisions is to accurately and precisely quantify the present state of knowledge of the properties of strongly-interacting matter. As such, striving for infinitely precise inference without making substantial effort to understand and reduce unquantified bias only serves to harm the broader theoretical program. In contrast, IP-Glasma's foundation in the QCD-based CGC framework includes microscopic physics and allows for a direct connection of pen and paper theory to phenomenology. IP-Glasma presents an opportunity to quantitatively constrain physical and theoretical understanding of strongly-interacting matter in the initial stage of heavy ion collisions with an already-successful model, attempt to identify weaknesses, and allow for them to be improved.

In this and the following section, I demonstrate differences between TRENTo and IP-Glasma that directly motivate a Bayesian analysis with an IP-Glasma initial state. I show a series of comparisons at switching time to hydrodynamics with important differences between the two models. In the direct comparisons in this chapter, I use maximum a posteriori (MAP) results for simplicity and due to computational expense. The Grad MAP parameter set from [45, 46] is used for TRENTo + freestreaming, while the calculations with IP-Glasma are performed with the calibration from [28] for Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV.

Comparisons are made at hydrodynamic switching time. Some researchers would prefer comparisons between the models at the same proper time $\tau$ [2]. A comparison at the same time would be informative if the relevant comparison were to understand synchronous differences between the evolutions. Instead, the goal of this work

is to understand what differences are presented to the remaining parts of the hybrid model and the impact of these differences. The remaining evolution is sensitive to the way the model is used in practice, rather than all possible configurations. Finally, many such comparisons of T$_R$ENTo and other models already exist while the comparison at the onset of hydrodynamics is novel and of more relevance to this discussion. All subsequent T$_R$ENTo quantities are calculated at $\tau \sim 1.45$ fm, subject to the energy-dependent switching time prescription in [45], while IP-Glasma calculations are performed at 0.4 fm. The JETSCAPE workflow is reproduced and validated, with brief details shown in Appendix A.

### 4.2.1 *Comparisons at the onset of hydrodynamics*

As noted earlier, "hydrodynamics is agnostic". If T$_R$ENTo and freestreaming were to somehow successfully reproduce the features of an IP-Glasma initial state at hydrodynamic switching time, then the subsequent analysis would not be impacted by the choice of initial state and differences in the evolution to that point would be immaterial. I show conclusively that differences do exist in the initial state produced by these different modeling choices and therefore hydrodynamics does not see indistinguishable initial conditions.

I will make two broad classes of comparison at the onset of hydrodynamics as the models are used in practice: first, I will compare the geometry and second, I will compare the dynamics. In the geometric comparisons, I will compare the shapes of the deposited energy densities in the reaction plane. This will reveal both whether event-averaged geometries are similarly distributed, but also if the centrality dependence of the geometries are comparable. In the comparisons with dynamics, I investigate differences in pre-equilibrium flow to determine whether to expect comparable response from the hydrodynamic stage in both models as well as if the hydrodynamic stage is forced to compensate for an incomplete physical picture from the pre-equilbrium phase.

To explore the geometry of the medium at hydrodynamic switching time, a pseudo-observable called $\epsilon_n$ is used and are the coefficients of a radial Fourier decomposition of the energy density in position space. In T$_R$ENTo, these are defined as

$$\varepsilon_n e^{in\phi} = -\frac{\int dx dy r^n e^{in\phi} T_R}{\int dx dy r^n T_R} \tag{4.5}$$

where $\epsilon_n$ are the coefficients of order $n$, $\phi$ is the azimuthal angle, and $r$ is the radius $r = \sqrt{x^2 + y^2}$. When using models other than T$_R$ENTo, $T_R$ is replaced by the energy density of the model under consideration. These are similar to $v_n$ – simply the spatial energy density rather than the momentum-space energy density – and are often considered a proxy for $v_n$ as hydrodynamics converts initial state spatial gradients into final state momentum space gradients. Thus, $\epsilon_2$ quantifies how elliptical the energy density is, $\epsilon_3$ quantifies how triangular the energy density is, and so on and so forth. This therefore allows for a meaningful geometric comparison between the two models as it quantifies differences in the shape of the energy densities. The event-averaged $\epsilon_n$ for central to peripheral collisions are shown in Fig. 4.2. It is immediately apparent that the shapes are strikingly different and do not have comparable centrality dependence or in the case of $\epsilon_{n\neq 2}$, magnitude. This suggests that the energy density supplied by T$_R$ENTo + freestreaming is more diffuse than the more structure-rich energy density of IP-Glasma.

The observed suppression in the higher components can be understood as the hotspots diffusing and overlapping each other to produce a remarkably different geometry from the strikingly non-flat IP-Glasma $\epsilon_n$. This is the first sign that even the basic geometry of the events at the onset of hydrodynamics is radically different.

Although the geometry supplied to hydrodynamics is quite different, it is important to consider the dynamics. This will allow for details such as medium response to be put into their proper context. As previously described, the dynamics typically implemented with T$_R$ENTo are freestreaming, as opposed to the microscopic CGC-based evolution used by IP-Glasma. The purpose of freestreaming is to provide pre-equilibrium dynamics to the hydrodynamic phase. The first comparison of the dynamics is the root mean square (RMS) transverse velocity of T$_R$ENTo + freestreaming and IP-Glasma, in Fig. 4.3. It is clear that both the shape and values of the RMS initial transverse flow velocities are strikingly different. An important difference between the two models is the centrality dependence of the RMS transverse flows. Fig. 4.3

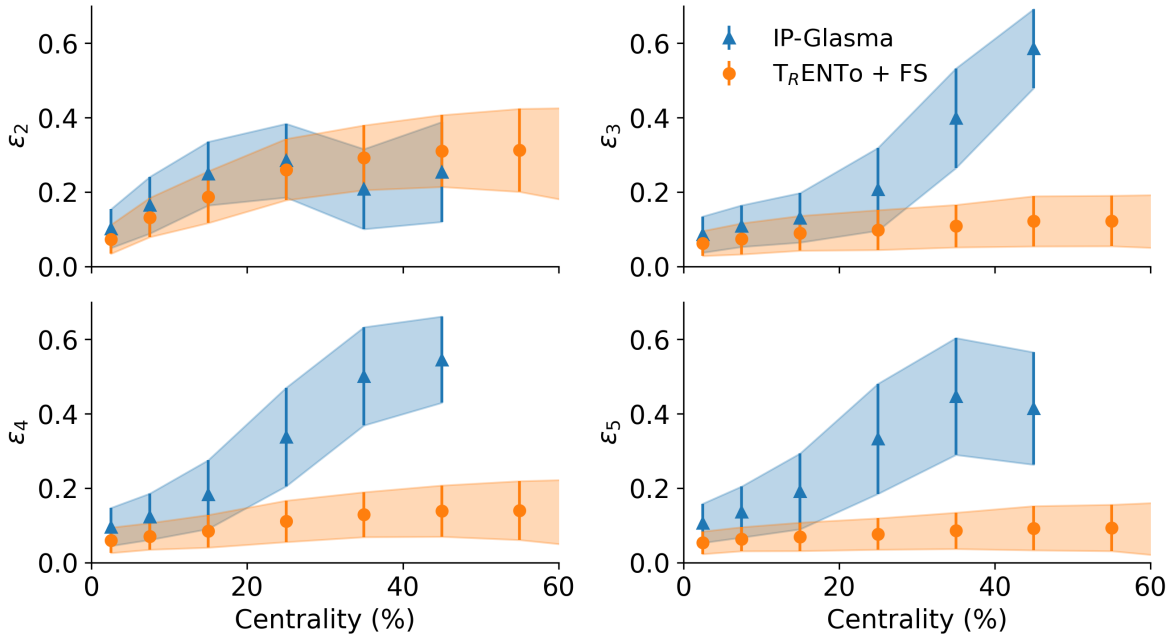


Figure 4.2: $\epsilon_2$ through $\epsilon_5$ at the onset of hydrodynamics

clearly shows a linear centrality dependence for T$_R$ENTo + freestreaming, while IP-Glasma's transverse velocity centrality dependence is clearly non-linear. This is certain to impact the hydrodynamic response and its centrality dependence.

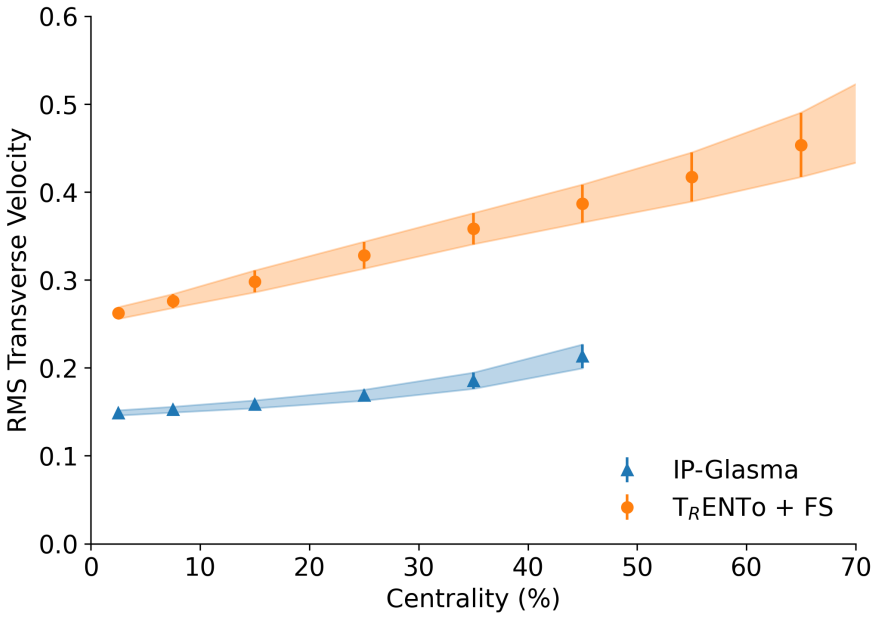


Figure 4.3: Transverse flow in T$_R$ENTo + freestreaming and IP-Glasma at the onset of hydrodynamics. The freestreaming time used is the prescription of [45] and IP-Glasma is evolved to 0.4 fm in accordance with [28].

Another differentiating feature between the standard T$_R$ENTo + freestreaming pre-equilibrium model and IP-Glasma is the momentum anisotropy. Because there are no interactions in freestreaming, the initial isotropic momentum distribution is frozen-in and the anisotropy is identically 0 as freestreaming is presently initialized with an isotropic momentum distribution. In IP-Glasma, there is no such constraint and a clear nonzero momentum anisotropy is observed.

In [167], a significant component (35-45%) of the momentum anisotropy comes from the initial state, particularly in pA collisions. The bulk viscosity also impacts the evolution of momentum anistropy, which suggests that freestreaming's isotropic momentum distribution is certain to have an impact on the evolution of the medium and influence estimates of the bulk viscosity.

The momentum anisotropy quantity is defined so that it is comparable to the final state momentum anisotropy $v_n$ coefficients, which must include the imaginary term. This definition is

$$\epsilon_p = \frac{T^{xx} - T^{yy} + 2iT^{xy}}{T^{xx} + T^{yy}} \tag{4.6}$$

and its magnitude is

$$|\epsilon_p| = \sqrt{\frac{(T^{xx} - T^{yy})^2 + (2T^{xy})^2}{(T^{xx} + T^{yy})^2}}. \tag{4.7}$$

This recovers the expected physical result for freestreaming and IP-Glasma, shown in Fig. 4.4. This is another clear difference between a T$_R$ENTo + freestreaming initial state and an IP-Glasma initial state and will have implications for the final momentum distributions observed in particle spectra.

This is another clear signal of how freestreaming does not successfully mimic the microscopic physics of the classical Yang-Mills evolution and why quantitative results generated with a T$_R$ENTo + freestreaming pre-equilibrium phase are likely unreliable reflections of the physical dynamics of heavy ion collisions. Both the geometry and the dynamics show that there is a clear need for a Bayesian study using an IP-Glasma initial state in order to better constrain the understanding of the QGP.

A difficulty in constraining the current state of knowledge lies in significant sources of theoretical uncertainty arising from a variety of available models. In the initial state, a variety of models are available and were discussed in Sec. 4.1. If a model

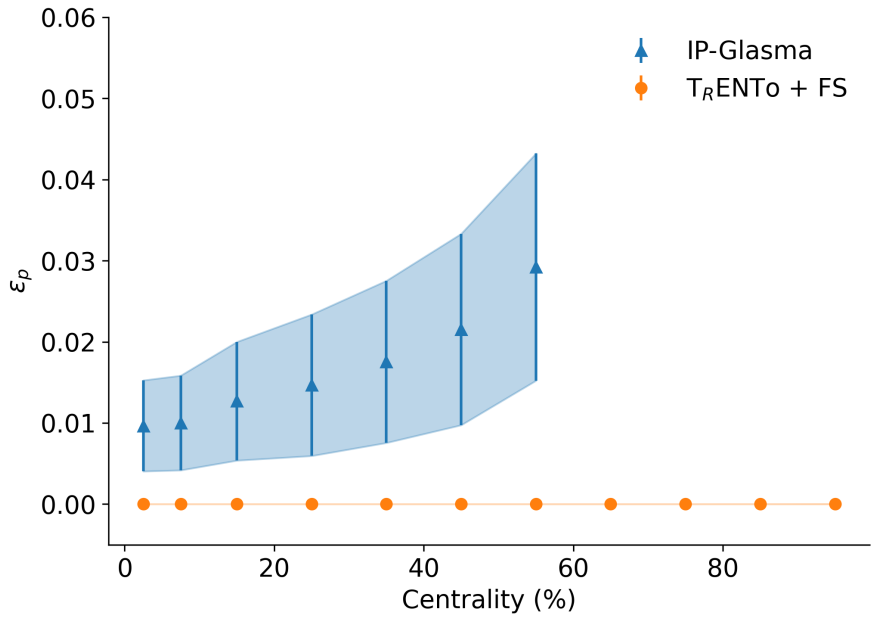


Figure 4.4: Momentum anisotropy in T$_R$ENTo + freestreaming and IP-Glasma. Freestreaming explicitly assumes an isotropic momentum distribution.

existed which was able to interpolate between broad features of these models, it would be able to constrain the model space and highlight important features using systematic model-to-data comparison. Such a claim was made using the T$_R$ENTo model, reproduced in Fig. 4.5, and has resulted in the need for repeated emphasis that $p = 0$ T$_R$ENTo and IP-Glasma are distinguishable using a variety of methods, including those shown in this chapter [2, 162].

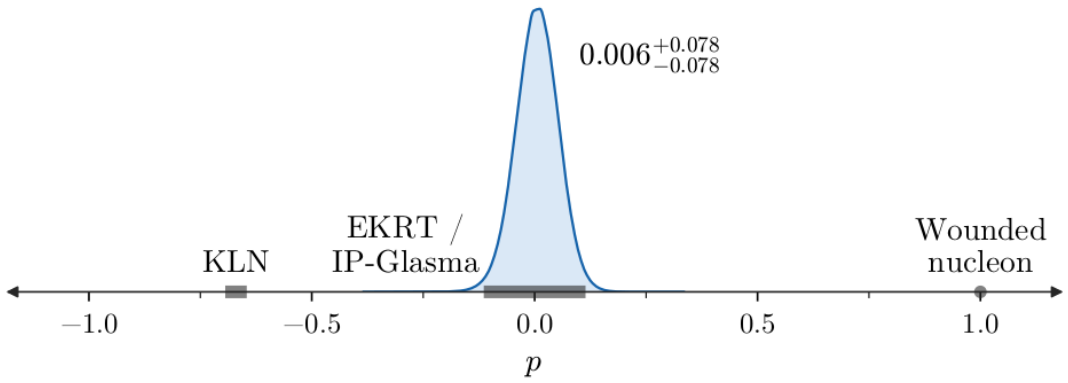


Figure 4.5: Model selection in T$_R$ENTo. Reproduced from Fig 5.13 in [42].

At the outset, it was emphasized that T$_R$ENTo + freestreaming *entropy* deposition, conditional on certain parameter choices, is able to reproduce some geometric fea-

tures of IP-Glasma and EKRT [42]. This value of $p$ is repeatedly favored by experimental data [42, 45, 51]. The comparison to other models arises solely from comparison to $\epsilon_2$, $\epsilon_3$, and $\sqrt{\langle\epsilon_2^2\rangle}/\sqrt{\langle\epsilon_3^2\rangle}^{0.6}$ for IP-Glasma and to thermal distribution profiles with special values of $p$ for KLN, EKRT, and the Wounded Nucleon model [42]. Note that these comparisons are only geometric and do not investigate dynamics, which of course modify the geometry. Additionally, these are with entropy density, not energy density as T<sub>R</sub>ENTo is now used. Insofar as $p = 0$ T<sub>R</sub>ENTo have been compared to IP-Glasma, EKRT, and other models, it is within this limited context.

Since the initial states are distinguishable using geometry, flow, and momentum pseudo-observables at hydrodynamic matching as used in practice, it can be conclusively stated that $p = 0$ T<sub>R</sub>ENTo and IP-Glasma are both qualitatively and quantitatively different. The strongest statement that can be made is that it is possible for a particular configuration of T<sub>R</sub>ENTo + freestreaming to mimic a particular configuration of IP-Glasma. However, this is not a guarantee that $p = 0$ is sufficient to do so, although it seems to suggest IP-Glasma as a candidate for further study.

Now that the two initial state models have been differentiated, it is important to investigate how those differences are expected to impact final state observables and, consequently, inference of strongly-interacting matter in heavy ion collisions. Critically, I do not seek to prove which model is "right" or "wrong", but how each of these two models is most *useful*.[5]

It is critical to ask what implications the identified differences will have for Bayesian inference – were these differences immaterial and nothing new would be learned, a new study could not be justified. In [160], the authors highlight that event plane correlators are particularly sensitive both to the value and to the temperature dependence of $\eta/s$. To date, no Bayesian study has used this to constrain the temperature dependence of shear viscosity, which is a critical aspect of study as [45, 46] find that it is difficult to constrain $\eta/s(T)$.

The evolution from the initial participant-plane correlations to final state event-plane correlations is highlighted in [118]. There is a compensating effect in this ob-

5 The famous adage is "all models are wrong, but some are more wrong than others". Thinking of the usefulness of models for various problems is a much more productive way to approach phenomenology than in stark binary categories and is generally ascribed to George Box.

servable between initial state model choice and shear viscosity required to match data. I have established significant differences between the predictions for these $\epsilon_n$ between T$_R$ENTo and IP-Glasma, suggesting that the geometric differences will impact the inferred shear viscosity. This coupling of initial state model and shear viscosity is not isolated to this observable; the effect is seen again in Fig. 5.4 of [42] where Glauber and KLN initial states produce nearly-non-overlapping distributions of $\eta/s$. In a study on nonlinear response coefficients, the ALICE collaboration [120] found that simply changing the shear viscosity in a single initial state model can change the prediction for certain coefficients by up to a factor of 10 in more central collisions. Changing the initial state model and holding the viscosity fixed can also produce a difference up to a factor of 3, again demonstrating the coupled nature of the initial state and hydrodynamic medium.

Viscosities modify the mapping of initial states to final states [168, 169], making it apparent that geometrically-motivated models without underlying physical constraints or dynamics are insufficient and significantly impact the results of statistical inference.

As previously discussed, $\epsilon_p$ is sensitive to the bulk viscosity and, when properly constructed, directly corresponds to $v_2$ in the final state. When 35-45% of $\epsilon_p$ in hybrid models of pA collisions can be traced to the initial state [56] and there are nontrivial correlations between shear and bulk viscosity [42, 45, 50, 51], it becomes imperative to use a physically motivated, interpretable, and systematically-improvable model to accurately constrain transport coefficients. To explicitly test determine if the hydrodynamic phase masks differences in initial state behaviors, the initial states must be evolved through hybrid models and compare final state observables.

## 4.3 FINAL STATE COMPARISONS

In this section, I compare T$_R$ENTo + freestreaming + MUSIC + iS3D + SMASH events using the Grad MAP parameters of JETSCAPE SIMS to the work of [28] in order to thoroughly explore differences in final-state predictions. Both of these choices presently represent the state-of-the-art for each hybrid model.

While I have comprehensively demonstrated differences at switching time to hydrodynamics, final state differences between the two hybrid models can suggest where the models experience tension in describing the data and what features may be lacking. In particular, the differences between these hybrid models are in the pre-equilibrium geometry and dynamics. The pre-equilibrium geometry and dynamics are mapped to final-state quantities via hydrodynamics, particlization, and the hadronic cascade.

Event plane correlators and hydrodynamic response coefficients are examples of observables that couple shear viscosity and the initial geometry [118, 120, 160]. Additionally, the initial state can be the source of significant momentum anisotropy. This momentum anisotropy is modified by the bulk viscosity [56]. Differences in initial geometry and momentum anisotropy will have a clear signal in the final state and corresponding implications for Bayesian inference. No Bayesian analysis in heavy ion collisions has yet considered such observables that couple these quantities in such an explicit manner, nor exploited them to constrain the temperature-dependence of the viscosity. This increases the value of choosing observables with such interpretable sensitivity.

T$_R$ENTo as a successful proxy for IP-Glasma was partially identified because of eccentricity scaling and binary scaling of the thickness function [42]. However, the parameters that result in this matched scaling are not chosen in practice. Thanks to sensitivity analyses, the underlying cause for this is revealed. The final state geometry is sensitive to both initial state geometry and $\eta/s$ [118, 160]. The momentum-space anisotropy, characterized by $v_n$, is also sensitive to shear viscosity. In a sensitivity analysis of a T$_R$ENTo + freestreaming + MUSIC + iS3D + SMASH hybrid model, the primary observables sensitive to changes in the freestreaming time are $v_n$, followed by $\langle p_T \rangle$ [45].

Thus, the momentum space anisotropy and correlated momentum fluctuations are connected to overly-simplistic dynamics. This has a knock-on effect as the initial $\epsilon_p$ and $\zeta/s$ have a demonstrable impact on the evolution of the hydrodynamic momentum anisotropy [56, 167]. Additionally, bulk viscosity is necessary to reproduce $\langle p_T \rangle$ and a range of other hadronic observables [170, 171]. As a result, the initial state produced by T$_R$ENTo that has the same geometric properties as a more physical model

is undermined by the heuristic dynamics that follow. The deficiencies' impact of the freestreaming stage on the extracted hydrodynamization time is clear. As a result, there are interpretable and identifiable deficiencies in the pre-equilbrium stage produced by T$_R$ENTo + freestreaming that impacts inference of the hydrodynamic stage. The impacts of these effects will be demonstrated in this section.

A disclaimer is worth including for the following comparisons. I compare calculations at single sets of parameter values, IP-Glasma in a hybrid model as used by [28] and T$_R$ENTo + freestreaming as specified by the MAP parameters of [45, 46]. The posterior predictive distribution from Bayesian model-to-data comparison would be a more informative quantity to differentiate the models and the true uncertainty in predictions from varying the parameters. However, this is not easily available, is prohibitively computationally expensive for exploratory work, and is beyond the scope of this analysis.

While the T$_R$ENTo + freestreaming hybrid model was systematically tuned to experimental observables, the work of [28] uses the following procedure to tune to data by hand: first, normalize IP-Glasma to match the charged hadron multiplicity and second, adjust a constant shear viscosity and a parametrized bulk viscosity to visually match flow coefficients $v_n$ and mean transverse momentum $\langle p_T \rangle$. Small adjustments are then made to normalization to ensure the best visual matching. Event plane correlators and nonlinear response coefficients were not used as inputs in either calibration and therefore represent a prediction by both models.

### 4.3.1 *First generation observables*

In this section, the first generation observables discussed in Ch. 2 are considered. Here, T$_R$ENTo + freestreaming is expected to produce satisfactory calculations as these are broadly observables to which the T$_R$ENTo + freestreaming hybrid model was systematically compared. However, there is tension in the JETSCAPE-SIMS calibration and this results in an underestimation of the charged hadron multiplicity in Fig. 4.6. A hand-calibrated IP-Glasma produces a better description of this observable.

Next, consider $v_n$. In Fig. 4.7, it is apparent that T$_R$ENTo + freestreaming undershoots the data systematically at MAP, while IP-Glasma better reproduces the data

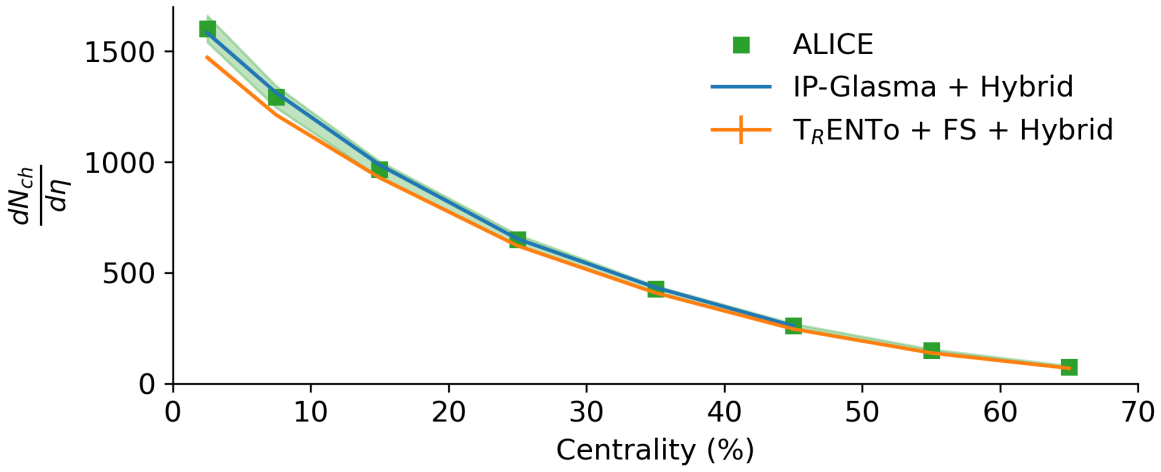


Figure 4.6: Charged hadron multiplicity from a hybrid model with IP-Glasma compared to that of TRENTo + freestreaming and data measured by ALICE. Shaded bands throughout this thesis denote uncertainty.

but does not fully capture the centrality dependence as it is currently configured. These two results are both worthy of attention, as it is important to emphasize that two imperfect models are being compared rather than a "perfect" model and a challenger. Sufficient differences between data and MAP TRENTo + freestreaming calculations suggests that IP-Glasma may be worthwhile to pursue systematically.

Another interesting feature is the relatively better matching to data from IP-Glasma compared to TRENTo + freestreaming with $\langle p_T \rangle$ in Fig. 4.8 despite the fact that these observables were included in the systematic tune. This suggests that sufficiently large tension existed elsewhere that it was not able to simultaneously match this quantity and other data.

### 4.3.2 *Next generation observables*

While both hybrid models produce reasonable, albeit different, reproductions of experimental results for first generation, simpler observables, it is time to investigate next generation observables beginning with the ATLAS Collaboration's measurement of event plane correlators [172]. The impact of hydrodynamics is key to a proper understanding of next-generation observables and provides a unique way to simultane-

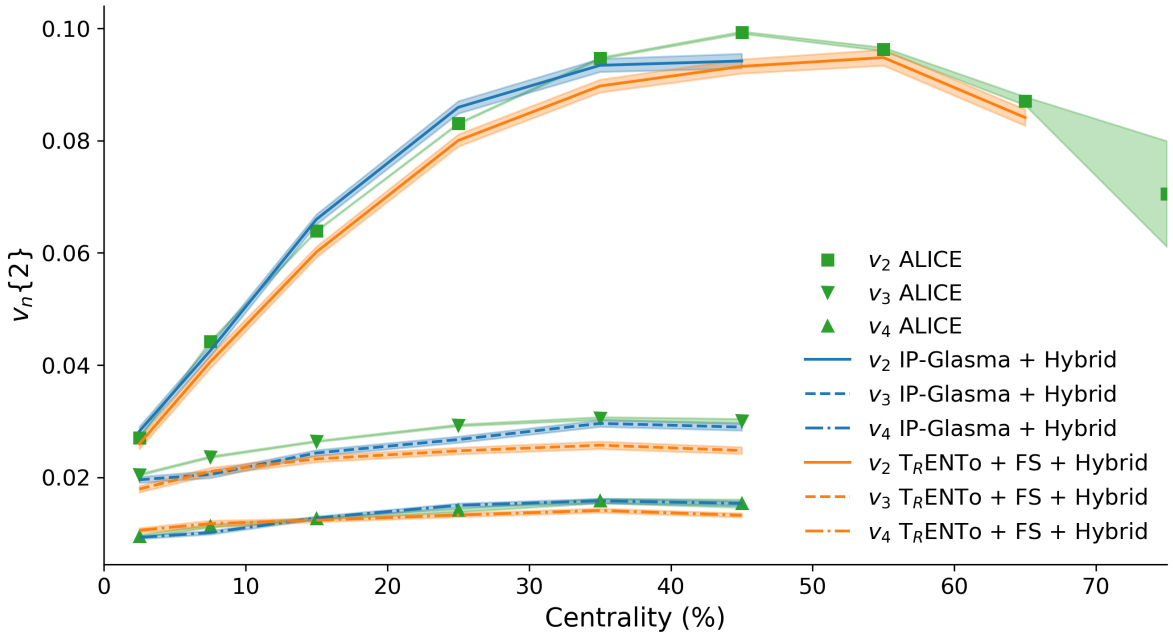


Figure 4.7: $v_n\{2\}$ from a hybrid model with IP-Glasma compared to that of T$_R$ENTo + freestreaming and data measured by ALICE. IP-Glasma comparisons are restricted to a conservative $0-50\%$ centrality range in keeping with [28].

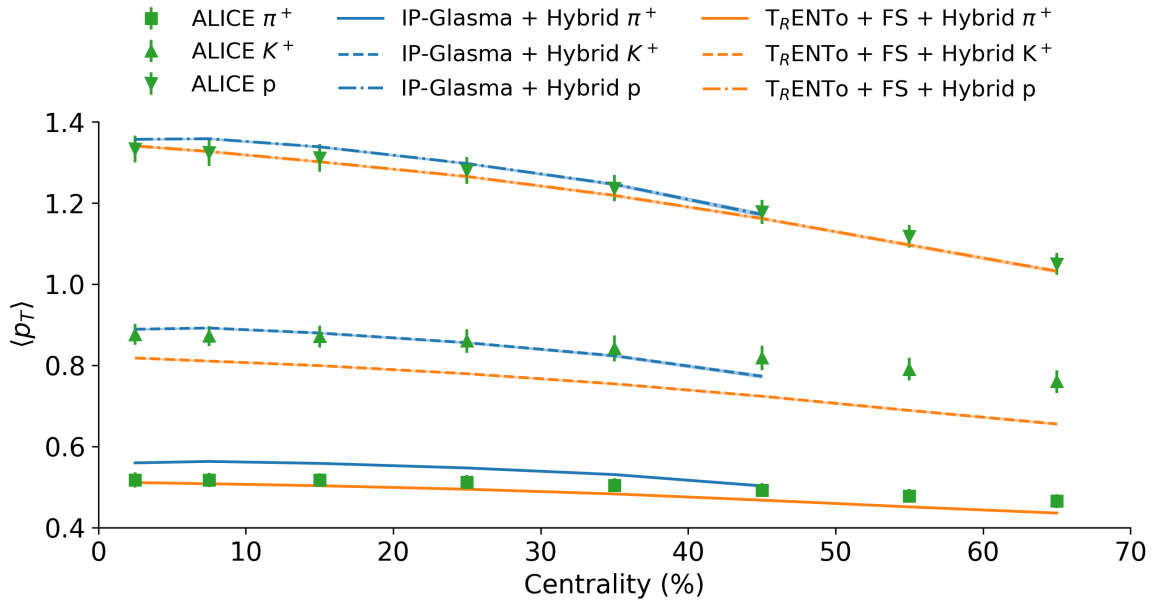


Figure 4.8: Identified particle mean transverse momentum from a hybrid model with IP-Glasma compared to that of T$_R$ENTo + freestreaming and data measured by ALICE.

ously constrain initial geometry and dynamics in addition to details of the hydrodynamic phase.

We show event plane correlations in Figs. 4.9– 4.15. In these plots, the number of participants at different centrality bins are estimated using the MC-Glauber model according to Table 1 in Ref. [172]. In some cases, there is little statistically-significant difference between T$_R$ENTo and IP-Glasma event plane correlators (Fig. 4.9), while in other cases IP-Glasma predicts these quantities more accurately (such as in Fig. 4.10). As a result, for these next generation observables, IP-Glasma has a clear advantage in predictive power. This demonstrates that there is physics in IP-Glasma that is not yet captured by T$_R$ENTo + freestreaming that impacts final-state results. This cannot be ascribed to the fact that IP-Glasma uses pre-generated nuclear configurations as T$_R$ENTo uses the same configurations, therefore the correlation is due to additional geometric and dynamic differences in the pre-equilibrium stage.

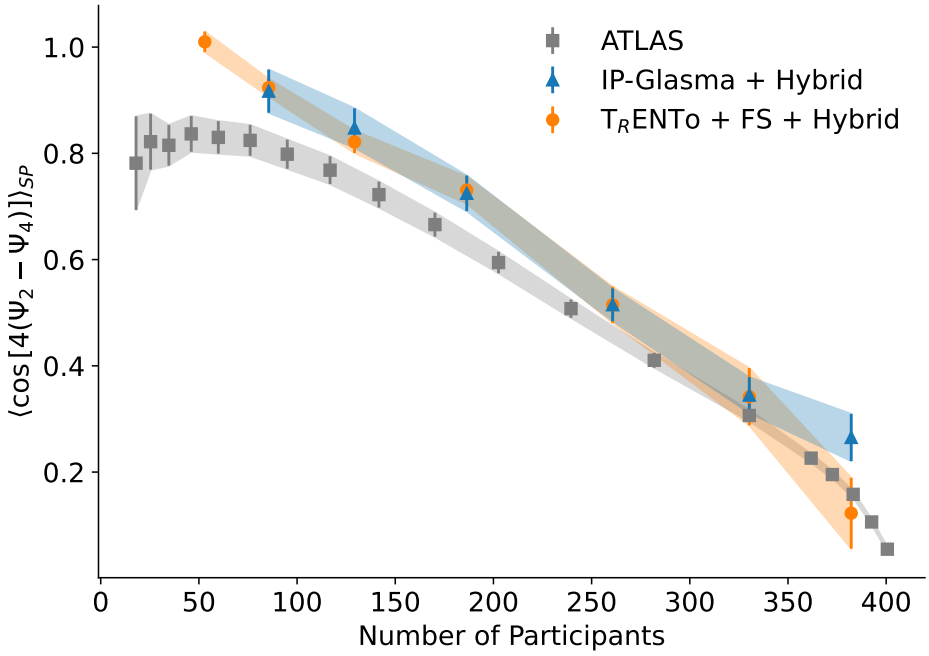


Figure 4.9: The $\langle\cos(4(\Psi_2-\Psi_4))\rangle$ charged hadron event-plane correlation as measured by the ATLAS experiment compared to predictions from IP-Glasma and T$_R$ENTo + freestreaming-based hybrid models.

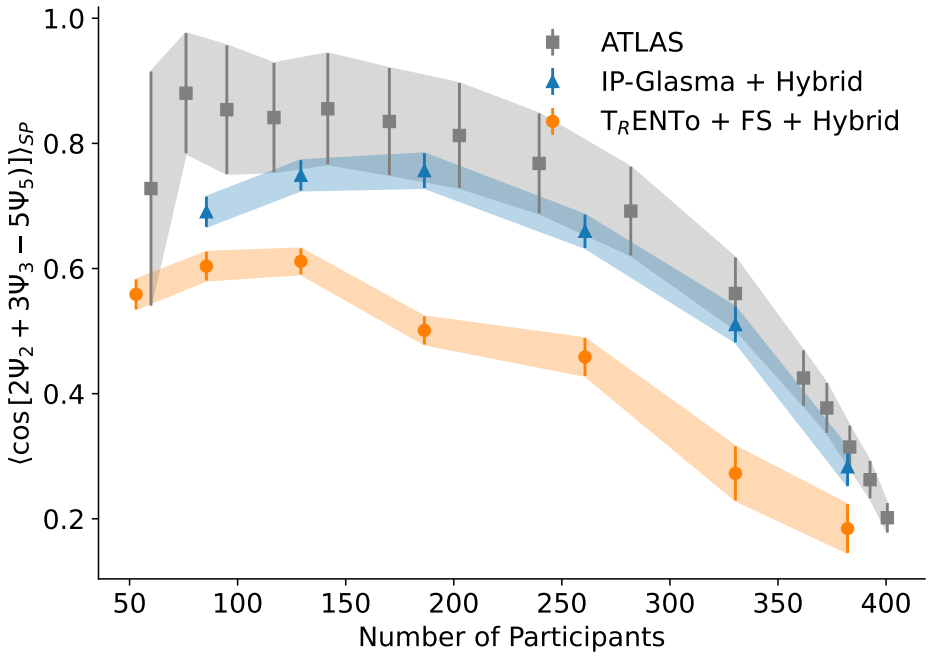


Figure 4.10: The $\langle\cos(2\Psi_2 + 3\Psi_3 - 5\Psi_5)\rangle$ charged hadron event-plane correlation as measured by the ATLAS experiment compared to predictions from IP-Glasma and T$_R$ENTo + freestreaming-based hybrid models.

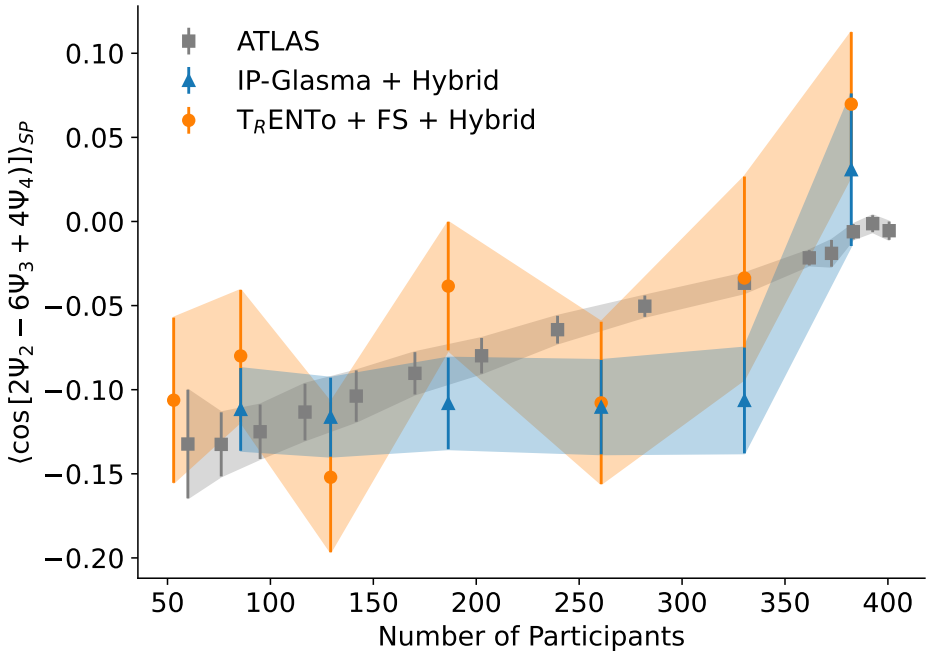


Figure 4.11: The $\langle\cos(2\Psi_2 - 6\Psi_3 + 4\Psi_4)\rangle$ charged hadron event-plane correlation as measured by the ATLAS experiment compared to predictions from IP-Glasma and T$_R$ENTo + freestreaming-based hybrid models.

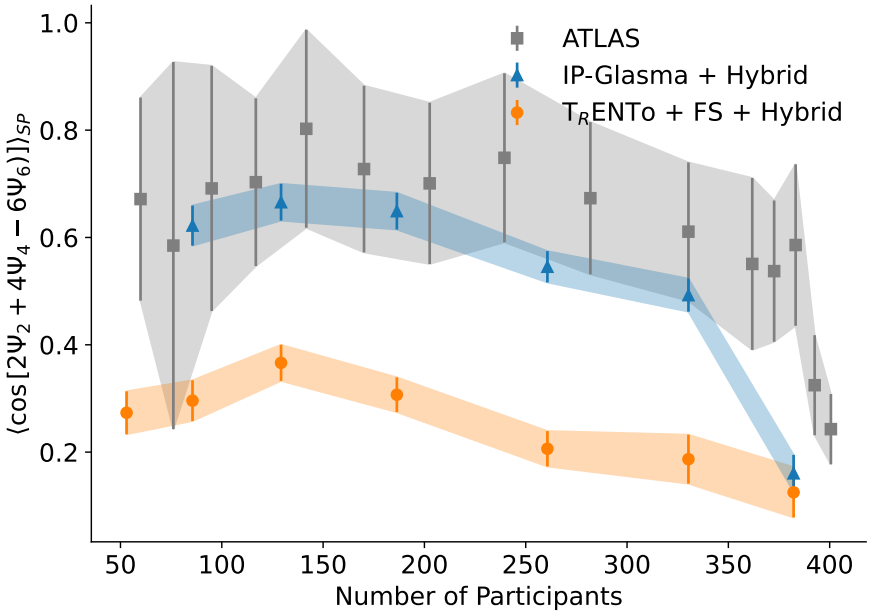


Figure 4.12: The $\langle\cos(2\Psi_2 + 4\Psi_4 - 6\Psi_6)\rangle$ charged hadron event-plane correlation as measured by the ATLAS experiment compared to predictions from IP-Glasma and T$_R$ENTo + freestreaming-based hybrid models.

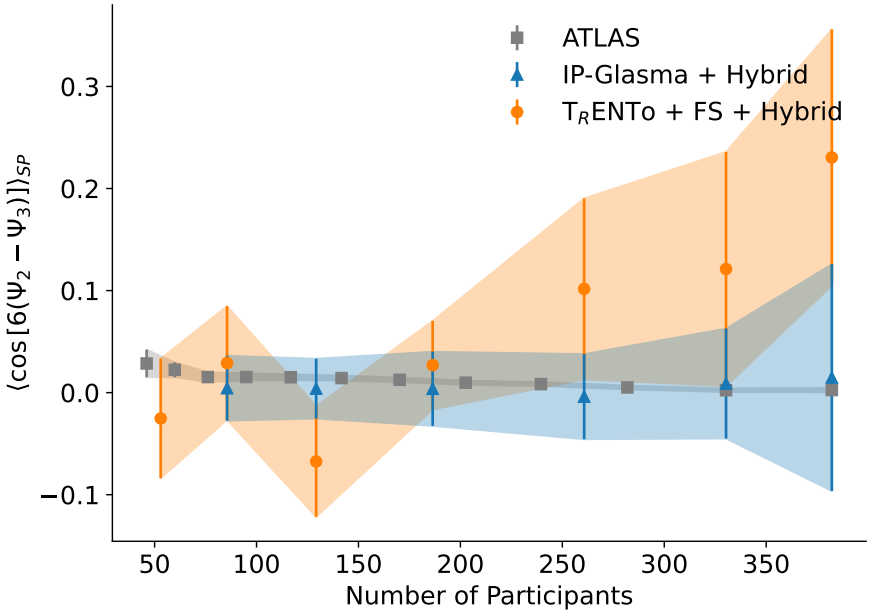


Figure 4.13: The $\langle\cos(6(\Psi_2 - \Psi_3))\rangle$ charged hadron event-plane correlation as measured by the ATLAS experiment compared to predictions from IP-Glasma and T$_R$ENTo + freestreaming-based hybrid models.

### 4.3.3 *Nonlinear response coefficients*

Nonlinear response is another observable sensitive to both the viscosity and the initial state model [120]. While second and third order anisotropic flow – $v_2\{2\}$ and $v_3\{2\}$, for

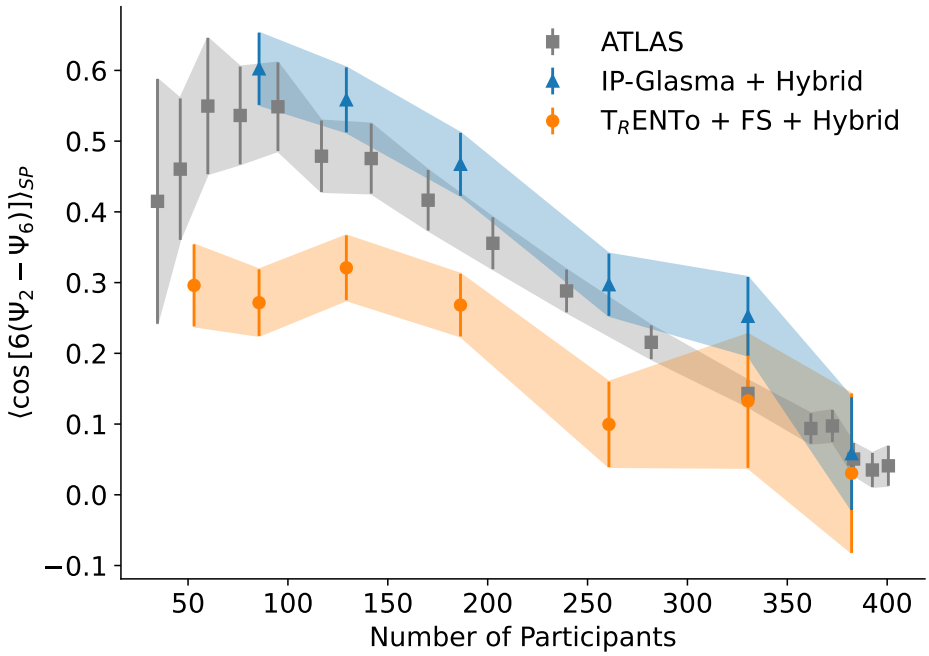


Figure 4.14: The $\langle\cos(6(\Psi_2 - \Psi_6))\rangle$ charged hadron event-plane correlation as measured by the ATLAS experiment compared to predictions from IP-Glasma and T$_R$ENTo + freestreaming-based hybrid models.

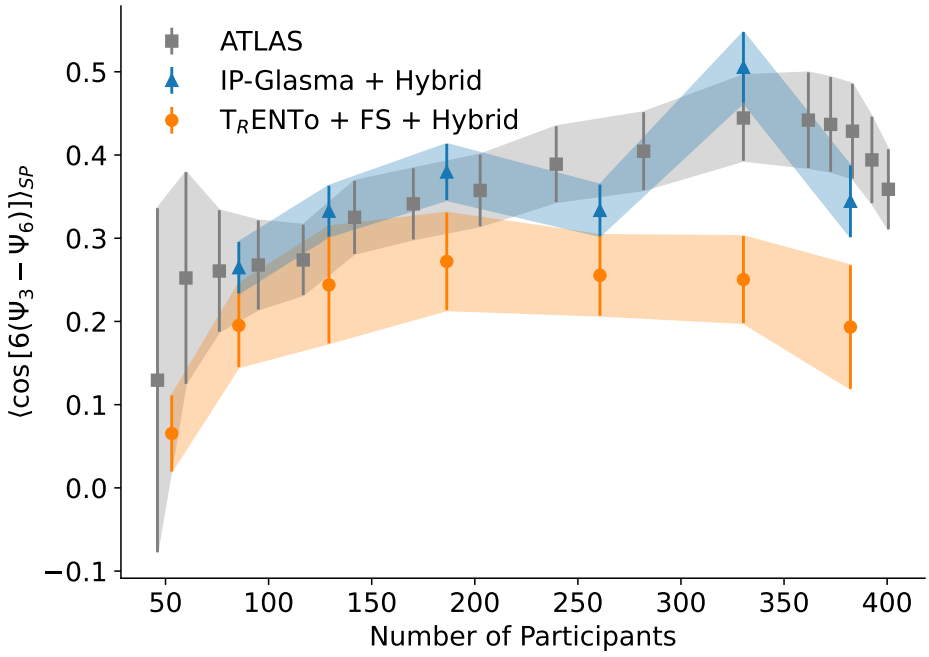


Figure 4.15: The $\langle\cos(6(\Psi_3 - \Psi_6))\rangle$ charged hadron event-plane correlation as measured by the ATLAS experiment compared to predictions from IP-Glasma and T$_R$ENTo + freestreaming-based hybrid models.

example – are mostly determined by the initial spatial anisotropy $\epsilon_2$ and $\epsilon_3$, higher order flows ($v_n\{m\}, n > 3$) can have significant mode coupling to lower order flows. This mode coupling is described by the nonlinear response coefficient as it breaks the otherwise dominantly linear relationship between $\epsilon_n$ and $v_n$. This nonlinear mode coupling is sensitive to the hydrodynamic phase, which maps $\epsilon_n$ to $v_n$ as well as to the various $\epsilon_n$ themselves. This makes it a useful and sensitive discriminating observable for constraining both the initial geometry and the viscosity of the fireball in heavy ion collisions.

In these calculations, T$_R$ENTo + freestreaming and IP-Glasma initial states produce comparable predictions; rarely are the distributions not overlapping. IP-Glasma's ability to simultaneously describe the event plane correlators provides further evidence that it contains important physics not present in simple parametric models with more freely-tunable parameters. Additionally, the temperature dependence of the viscosity's impact on $\chi_{n,mk}$ in IP-Glasma-based hybrid models, which was not explored in [120] or [28], is of significant interest.

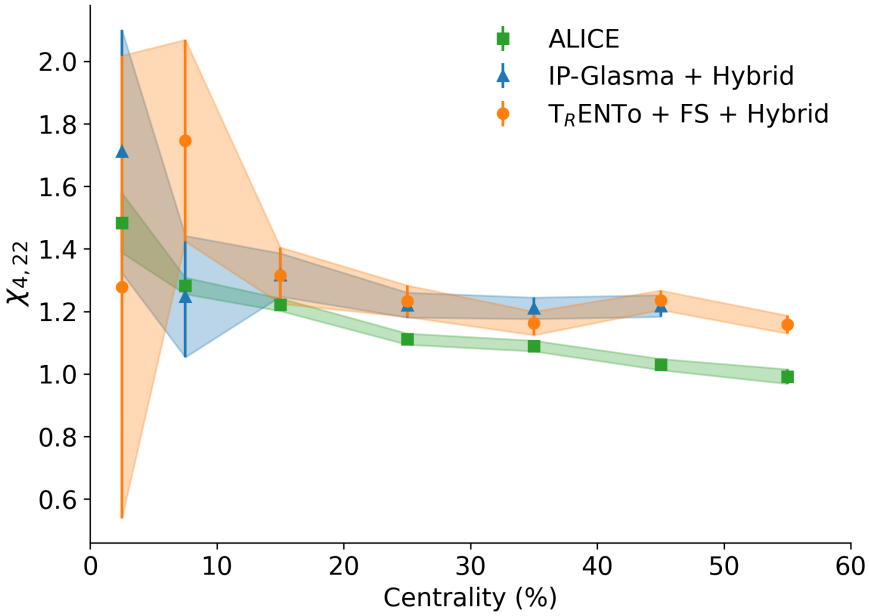


Figure 4.16: Predictions of the $\chi_{4,22}$ nonlinear response coefficient and measurements in Pb+Pb collisions at 2.76 TeV.

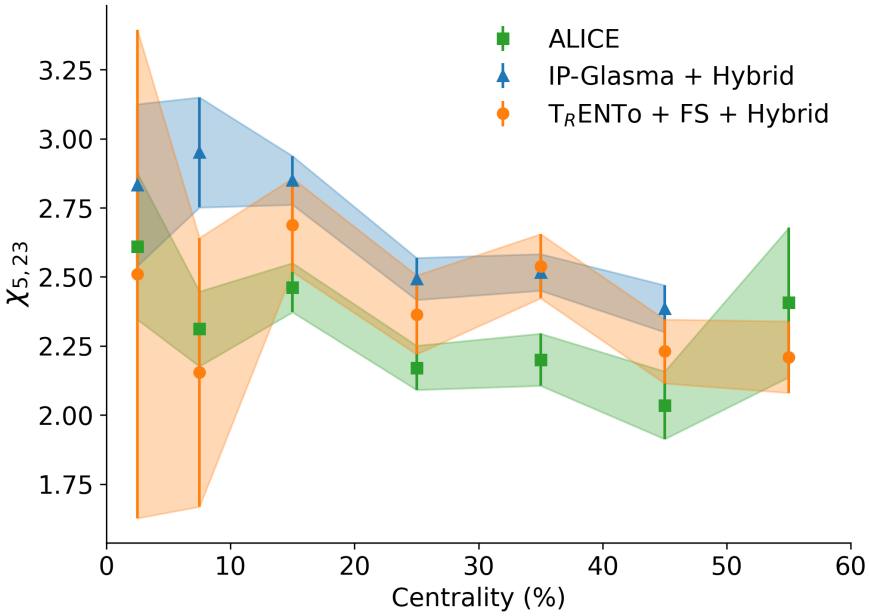


Figure 4.17: Predictions of the $\chi_{5,23}$ nonlinear response coefficient and measurements in Pb+Pb collisions at 2.76 TeV.

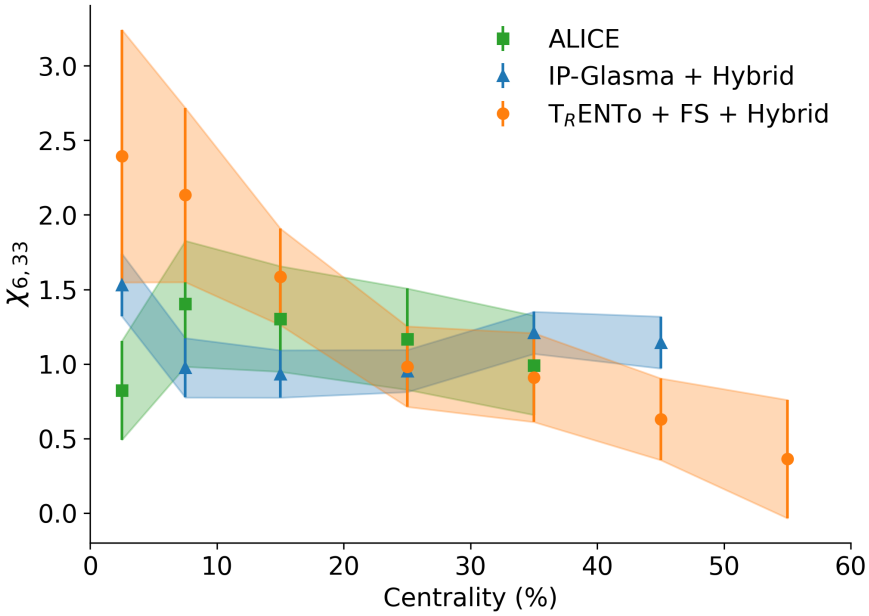


Figure 4.18: Predictions of the $\chi_{6,33}$ nonlinear response coefficient and measurements in Pb+Pb collisions at 2.76 TeV.

## 4.4 EXPLICIT MOTIVATION FOR A NEW BAYESIAN STUDY

It has now been demonstrated that IP-Glasma's predictive power is superior, which suggests that it is advantageous for accurate inference of physical properties as it is a

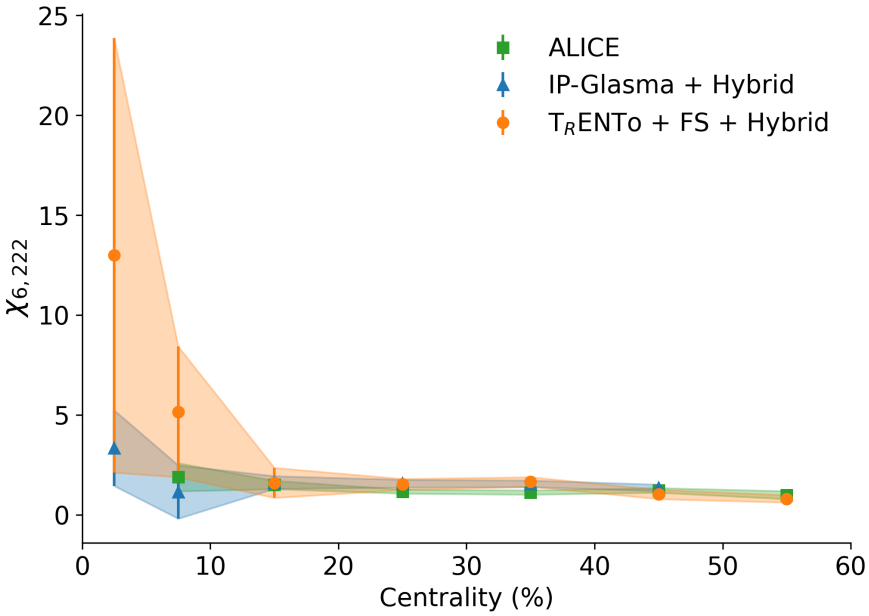


Figure 4.19: Predictions of the $\chi_{6,222}$ nonlinear response coefficient and measurements in Pb+Pb collisions at 2.76 TeV.

more complete description of the dynamics. This further supports the goal of motivating a new Bayesian study with an IP-Glasma pre-equilibrium stage. Methodological improvements have also been made that will allow this new study to leverage the latest techniques for computational design, efficiency, and accuracy. Throughout this chapter, it has been shown that T$_R$ENTo + freestreaming initial states do not successfully mimic IP-Glasma's geometry or dynamics and deliver a quantitatively different and inferior initial state to the hydrodynamic phase. It has also been shown that this impacts the predictive power of the hybrid model and, in Chapter 2, the physical basis of the IP-Glasma model was thoroughly expounded. The implications of the deficiencies of a parametric pre-equilbrium stage are expected to have an impact on parameter inference and as a result, introduce bias to describing the underlying QCD medium.

Another motivation behind a new study is that a hybrid model of heavy ion collisions with an IP-Glasma pre-equilibrium stage has never been studied in a systematic way. This provides a unique opportunity to test a physically-derived microscopic theory for the initial stage of HICs and identify aspects for future study and improvement. Bayesian inference should be treated as a precision tool to quantify the parameters of physical models, therefore if reliable inference with minimal bias is

the goal, state-of-the-art models at each stage are a requirement. Advancing the field from Bayesian techniques with T$_R$ENTo+freestreaming to IP-Glasma is a clear improvement to the theoretical study of heavy ion collisions.

Through using IP-Glasma as the pre-equilibrium stage, this thesis will address these points. Not only is IP-Glasma based on microscopic physics, but it is already tuned to data from deep inelastic scattering experiments.[6] It also has the advantage of reducing the curse of dimensionality as there are fewer free parameters in the initial state, making exploration of the space more tractable. Since the whole collision system will be informed and constrained by microscopic dynamics, this analysis is able to extend to specific observables that couple the pre-equilibrium and hydrodynamic stages of heavy ion collisions. The expectation is that this will constrain the temperature dependence of the shear and bulk viscosity of the QGP to unprecedented accuracy and precision.

Finally, from a philosophical perspective, science should be concerned with rigorously testing and comparing theoretical models to data. The state-of-the-art previous to this study is a good-faith attempt to do this, albeit with a parametric estimate of the pre-equilibrium stage guided by sound physical reasoning, but by no means a consistent and detailed description of microscopic physics. This consequently results in knock-on effects in parameter inference of transport coefficients in an uncontrollable way that is more systematically improvable through physics progress. Rather than attempt to make estimates of what physics is primarily relevant for quantities of interest and include it by hand or determine what constitutes a "good enough" description of the physics, it is clearly preferable to employ the leading physical model. IP-Glasma is the leading physical model of the initial stage of heavy ion collisions and has not yet been systematically been compared to data with Bayesian methods. This thesis will rectify this.

6 In principle, a future study of IP-Glasma could perform parameter inference on IP-SAT and use this as an empirical prior for additional parameters in IP-Glasma.

## 4.5 DESIGN CONSIDERATIONS

I now lay out details of how the study will be performed such as various physics choices in the model – such as the treatment of the σ meson – and details of the Bayesian inference. In Sec. 4.5.3, I describe the choice of likelihood function as well as which parameters will be varied and the relevant parameterizations.

The design discussion begins with physics choices which must be made prior to further decisions. In Sec. 4.5.1, the choice of viscous corrections is described as well as how to account for this source of uncertainty. In Sec. 4.5.2, the treatment of the sigma meson and how to construct the equation of state are described. Finally, in Sec. 4.5.3, the choice of likelihood function is justified, specifics priors are motivated, and the varied parameters in the model are discussed.

### 4.5.1 *Viscous correction*

Viscous corrections at particlization are an important source of uncontrolled theoretical uncertainty at switching from a hydrodynamic description of the strongly-interacting matter in heavy ion collisions to a particle-based kinetic theory description. The two leading candidates, Grad's 14-moment viscous corrections and Chapman-Enskog relaxation time approximation viscous corrections were introduced in Ch. 2. Recent work by the JETSCAPE Collaboration's Simulations and Distributed Computing working group (to which I contributed) tested several forms of the viscous corrections and found somewhat differing inferred transport coefficients and particlization temperatures. In addition, through Bayesian model comparison, the JETSCAPE Collaboration's hybrid description of HICs strongly favored Grad's 14-moment expansion [45, 46].

This preference was due to the failure of the Chapman-Enskog RTA viscous corrections to simultaneously capture proton and pion yields in a hybrid model with a T<sub>R</sub>ENTo + freestreaming initial state. If one discounts proton yield in the calculation of the Bayes factors in [45], the relative preference for Grad viscous corrections over Chapman-Enskog decreases from $\ln B_{A/B} = 8.2 \pm 2.3$ to $\ln B_{A/B} = 2.7 \pm 2.8$. In other words, a preference of 4000 : 1 decreases to a preference of $\approx 5 : 1$ and con-

sistent with no preference at all. The work also has a posterior whose specific bulk viscosity is large at particlization, suggesting that the influence of viscous corrections may be exaggerated and pushes the assumption that the corrections are small compared to the equilibrium distribution. This demonstrates that this debate is not settled and is worth revisiting and accounting for such uncertainties in this work. In addition, hybrid models with an IP-Glasma pre-equilibrium stage show agreement within statistical error bars with the proton yield [28]. This suggests that this discrepancy in the proton yield may be an effect deserving further study. In this thesis, I will investigate only linearized viscous correction models. In addition to their longer history in the literature, an important feature of these viscous corrections is that they are systematically-improvable by expanding to higher order in the expansion. For example, the 14-moment viscous corrections contain all terms in an expansion to hydrodynamic order – *i.e.* terms with $p^\mu$ and $p^\mu p^\nu$ – while Chapman-Enskog viscous corrections are a gradient expansion around the equilibrium distribution function.

Previous work in selecting models of viscous corrections at particlization informs the choice in this thesis to primarily use Grad's 14-moment viscous corrections. Because of the focus on linear viscous corrections, this is also well-suited to exploit transfer learning for the first time in a large scale study [4]. This will then allow for a rigorous treatment of the systematic uncertainties arising from the choice of viscous correction at particlization that directly impacts matching to observable quantities.

This thesis will primarily consider Grad's 14-moment viscous corrections with a transfer learning extension to the Chapman-Enskog relaxation time approximation viscous corrections in order to test another model and better account for theoretical uncertainty. Potential extension to exponentiated viscous corrections or Maximum Entropy viscous corrections is left for future work and will require significantly less investment due to this study already computing the pre-equilibrium and hydrodynamic stages needed for such a study, assuming transfer learning.

### 4.5.2 $\sigma$ *meson and the equation of state*

The treatment of the $\sigma$ meson varies between studies of heavy ion collisions. The $\sigma$ meson was proposed in the linear $\sigma$ model to describe low-energy pion dynamics $\sigma$

model [173] and can be used for exact calculations of transport coefficients in low-temperature hadron gases [18]. There is significant uncertainty surrounding its pole mass and width [6]. As such, it is necessary to make informed choices of how to treat the σ meson as it impacts both the hadron resonance gas phase as well as the equation of state matching to high-temperature QCD, as discussed in Ch. 2. The impact of including the σ meson at particlization has also been studied and a variety of choices that are consistent with the experimental understanding have been shown influence final observed results [45]. The contribution by σ to the partition function is nearly completely cancelled, further supporting its lack of inclusion [111]. As such, the σ is excluded from the SMASH list of resonances for sampling at particlization. However, the σ is important in the hadron resonance gas phase to reproduce the $\pi - \pi$ cross-section and is allowed to form dynamically.

In this thesis, the equilibrium equation of state is matched to the SMASH box list of particles that excludes extremely massive resonances.[7] The inclusion of only the lighter species in the construction of the equation of state is the correct practice recommended by the developers of SMASH.

### 4.5.3 *Bayesian analysis*

Bayes theorem derives the posterior via the priors, the likelihood, and the Bayes Evidence. The Bayes Evidence is a normalization and is calculated by integration as described in Ch. 3. Thus, what remains are the likelihood and priors. In this section, specific choices are justified. This thesis focuses on Pb-Pb collisions at the LHC with $\sqrt{s_{NN}} = 2.76$ TeV; extending to other collision systems is beyond the scope of this thesis but is a logical next step.

#### 4.5.3.1 *Likelihood function*

The least informative, Maximum Entropy (MaxEnt) choice for data with a mean and a variance is a Gaussian likelihood [122]. When quantities are typically reported in HICs, the result is given as a mean and two variances, one "statistical" – arising from

7 The list of heavy resonances is incomplete in SMASH and work is ongoing to identify and characterize important contributions [174–178].

a finite number of collisions – and another systematic, arising from detailed understanding and modeling of the experimental apparatus $\mu \pm \sigma_{\text{stat.}} \pm \sigma_{\text{sys.}}$. However, no further information is provided. As a result, the least informative course of action is to assume that the two variances are independent and uncorrelated.[8] Without further information, a mean and a variance characterize a Gaussian distribution and thus a Gaussian likelihood is the only realistic option for a non-informative study. Interesting work has been done on proposing theoretical covariances between quantities, but without detailed information from experiment, these remain hypotheses.

#### 4.5.3.2 *Priors*

The prior distribution, or state of knowledge, for each parameter must be justified for each study. However, the general form of the prior deserves some attention for the first time in heavy ion collisions. In all Bayesian inference in heavy ion collisions to date, uniform priors are used. However, uniform priors are unrealistic and unfaithful representations of the true state of knowledge due to sharp cutoffs. Sharp cutoffs can, of course, be physical: a quantity restricted to be positive definite may be reasonably likely to be zero, but infinitely unlikely to be even slightly negative. However, this is not always the case. For example, if a uniform prior is put on the boiling point of water at sea level $\mathcal{U}(99.0, 101.0)$ Celsius, this is tantamount to suggesting that it is equally likely for the boiling point to be anywhere between 99 and 101 Celsius but forbidden to be $101.0 + 10^{-9}$ Celsius. This kind of sharp change in prior is an improper representation of the prior state of knowledge without explicit justification for such seeming ambivalence and absolutism. From a more practical perspective, sharp cutoffs in high dimensions can also impact the performance of MCMC algorithms.

Existing guidance from Bayesian practitioners in the statistics community suggests using uniform priors with sharp cutoffs if that is an accurate reflection of the underlying constraint and not as a general non-informative choice. Additionally, priors may be chosen with features such as boundary-avoidance or invariance under reparametrization [179]. Another important consideration is to interrogate what

8 Recently, experimentalists have commented that error bars do not necessarily denote a mean and variance but rather characterize a range in which there may be a uniform probability and outside which there is none. Without additional information, no progress can be made from theory alone.

"weakly-" or "non-informative" means in the absence of explicit reference to the likelihood. If the dominant constraint comes from the prior, then the prior is informative. Conversely, if the likelihood is the dominant source of constraint, then the prior is less informative.[9]

It is high time for the priors used in HICs to be brought in line with more standard statistical practice. In order to bias the priors as little as possible, this study uses a Generalized Normal distribution (Version 1) with varying mean $\mu$, location $\alpha$, and shape parameter $\beta$. The shape parameter $\beta$ which controls the tails of the distribution. When $\beta = \infty$, the distribution becomes the uniform distribution. When $\beta = 2$, the distribution is Gaussian, while when $\beta = 1$, the distribution is Laplacian. When $\beta \gtrsim 8$, the distribution resembles a flat plateau with power law tails, smoothly interpolating between the current practice (effectively $\beta = \infty$) and priors more reflective of the underlying physics. This is an important step towards including yet-more physics information in the priors and I demonstrate in this thesis that such priors are both weakly-informative and can be consistently incorporated into study design. Future studies should use this infrastructure to further incorporate theoretical expectations in systematic model-to-data comparisons.

The Generalized Normal distribution has support on the whole real line and can be shifted. Additionally, the Half Generalized Normal distribution exists for instances where a sharp cutoff is reasonable, e.g. positive specific viscosity for non-decrease of entropy. Quantities such as the probability density function (PDF) and cumulative distribution function (CDF) are well defined, as is the entropy of the distribution, and thus it is well-suited for continuing, flexible study.

The probability density function of the Generalized Normal distribution is

$$p(x, \mu, \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{(-|x-\mu|/\alpha)^\beta} \tag{4.8}$$

where $\Gamma$ is the Gamma function. I generate a variety of sample PDFs for the standard Generalized Normal distribution (i.e. mean 0 and variance 1) with different $\beta$ in Fig. 4.20.

9 An interesting way one could check whether the prior or the likelihood is dominantly informative is to calculate a metric such as Kullback-Leibler Divergence between the likelihood and the posterior vs the prior and the posterior.

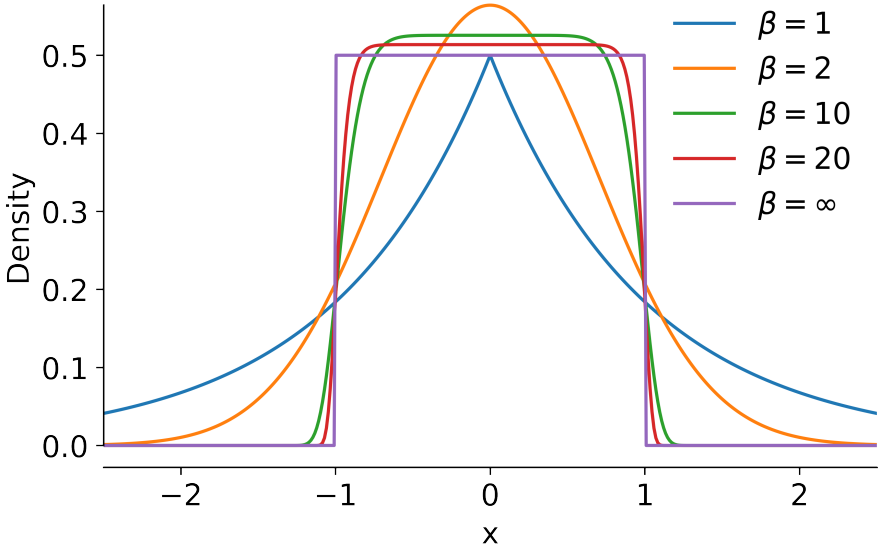


Figure 4.20: The standard Generalized Normal Distribution (Version 1) probability density function for several sample values of the shape parameter β.

### 4.5.4 *Parameters*

In this section, the physical meaning of the free parameters investigated in this thesis is described. The specific choices for individual priors are given in each study in subsequent chapters. However, the general form of the prior remains the same: a Generalized Normal distribution with a specified shape parameter β and a central 99% interval. Rather than specify values of the location and scale, the central 99% interval is chosen and the parameters that produce this interval are found through numerical optimization. This is more interpretable as it specifies a 99% degree of belief that the parameters are within a certain range and is directly comparable to the 100% central interval used to characterize the uniform distribution.

In this thesis, only parameters in IP-Glasma and MUSIC are varied. The choice of viscous correction is a parameter in iS3D, but is fixed for each calculation. The impact of this choice on parameter inference is investigated in Ch. 6.

#### 4.5.4.1 *IP-Glasma*

The parameters in IP-Glasma are mostly fixed via the IP-SAT model's comparison to deep inelastic scattering experiments, as described in Ch. 2. Two parameters are

poorly constrained: the proportionality between the saturation scale and color charge densities, and the onset of hydrodynamics while the strong coupling `g_strong` has previously been fixed to reproduce the low-energy $\alpha_s$. The strong coupling and the proportionality between the saturation scale and color charge density are totally degenerate (as confirmed in Ch. 5) and thus the strong coupling is fixed to 2. Suppressing the running of the coupling,

$$\alpha_s = \frac{g^2}{4\pi}. \tag{4.9}$$

In (2+1)D IP-Glasma, the coupling is fixed during the Classical Yang-Mills evolution, thus fixing it to 2 results in $\alpha_s = 1/\pi$. This is commensurate with experimental measurements and model predictions for $\alpha_s$ at low energy scale $Q < 0.5$ GeV [180].

Each parameter is now described in more detail and is given a shorthand notation.

1. `Multiplier_for_mu_to_Qs` ($\mu_{Q_s}$): Multiplier from the saturation scale to the color charge density profile ($Q_s \propto g^2\mu$). In the CGC, these quantities are proportional but a constraint on this proportionality is not known from theory. Phenomenologically, it is important for determining the energy density in IP-Glasma and the variance of color charge fluctuations. The value used in 3D is typically higher because the 2D simulation does not evolve the gluon densities and thus the boost invariant assumption modifies the behavior of the gluon fields.

2. `g_strong` ($g_{strong}$): Strong coupling constant. Believed to be completely degenerate with $\mu_{Q_s}$, but worth investigating in an initial state closure test to check systematically. The value currently used in IP-Glasma is 2.0, but a wide variation will be allowed as this has not been systematically studied before.

3. `Switching_time_to_MUSIC` ($\tau_0$): Proper time in fm of the transition between IP-Glasma and hydrodynamics. In IP-Glasma, the Glasma phase stabilizes within approximately 0.2 fm while flow continues to build as shown in Fig. 4.21 reproduced from [69]. The onset of hydrodynamics is additionally expected on order $1/Q_s$, typically around 0.2 fm. However, recent studies with freestreaming [42, 45, 51] extract a longer time to the onset of hydrodynamics than the 0.4 fm typically used with IP-Glasma [28]. This thesis will allow for switching times up

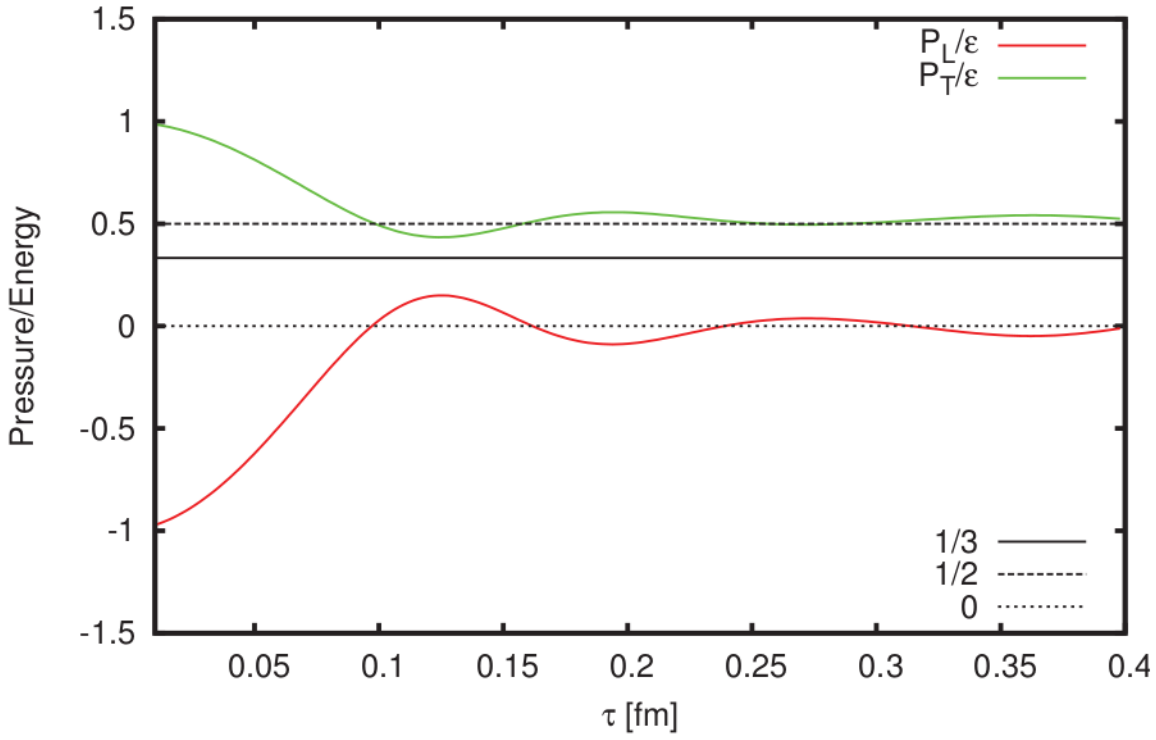


Figure 4.21: Pressure to energy ratio in (2+1)D IP-Glasma, adapted from [69].

to ∼ 1.2 fm, informed by longer hydrodynamic onset time and the approach to hydrodynamics [181], to determine if these long onset times are plausible in a model with more physically-inspired dynamics.

#### 4.5.4.2 *MUSIC*

The MUSIC parameters included are the shear and bulk viscosity parametrizations and the freezeout (or particlization) temperature. Due to the parametric flexibility in the viscosity, these parameters dominate the analysis.

It has been proposed to avoid using a parametrization, which correlates values of the viscosity at different temperatures [47]. However, extrapolating outside of the range of chosen temperatures is uncontrolled and complete lack of correlation between points is unphysical as the approximate shape of the shear viscosity is constrained to have a variety of well-known features. The advantage of the decorrelation is straightforward constraint to ensure viscosity is small at freezeout, ensuring small viscous corrections. However, this is insufficient to overcome the unphysical nature

of such decorrelation. As a result, this work uses the viscosities as parametrized in [45, 46] but widens the prior ranges.

The specific shear viscosity is parametrized as

$$\begin{aligned} \eta/s(T) &= (\eta/s)_{kink} + \alpha_{\eta,low} * (T - T_{\eta,kink})\Theta(T_{\eta,kink} - T) \\ &\quad + \alpha_{\eta,high} * (T - T_{\eta,kink})\Theta(T - T_{\eta,kink}) \end{aligned} \tag{4.10}$$

where the function has four parameters: $(\eta/s)_{kink}$, $\alpha_{\eta,low}$, $\alpha_{\eta,high}$ and $T_{\eta,kink}$. These control the value of $\eta/s$ at some kink, the slope below and above the kink, and the temperature of the kink. In practice, this can be less than 0, so the value used is $max(0, \eta/s)$. This parametrization is justified by first principles calculations which show such a kinked behavior, if any temperature dependence at all [182, 183]. IP-Glasma has so far experienced success without any temperature dependence in $\eta/s$.

The specific bulk viscosity is parametrized as the pdf of a skewed Cauchy distribution,

$$\begin{aligned} \zeta/s(T) &= \frac{(\zeta/s)_{max}\Lambda^2}{\Lambda^2 + (T - T_{\zeta,c})^2} \\ \Lambda &= w_\zeta \left[1 + \lambda_\zeta \, \text{sign}\,(T - T_{\zeta,c})\right] \end{aligned} \tag{4.11}$$

where the function again has four parameters: the maximum of the bulk viscosity $(\zeta/s)_{max}$, the temperature at which the bulk viscosity is maximum $T_{\zeta,c}$, the width of the bulk viscosity $w_\zeta$ and the skewness $\lambda_\zeta$. This parametrization, as highlighted in [45], is consistent with the expectation that specific bulk viscosity for QCD matter reaches a peak near the deconfinement transition and is related to the trace anomaly of QCD or a corresponding dip in the speed of sound in-medium while at high temperature, QCD becomes increasingly conformal and the specific bulk viscosity is expected to smoothly approach zero [18, 184–188].

The full list of the parameters varied in MUSIC is as follows:

1. $(\eta/s)_{kink}$: The value of $\eta/s$ at the kink temperature.

2. $T_{\eta,kink}$: The temperature at which $\eta/s$ changes slope.

3. $\alpha_{\eta,low}$: The slope of the $\eta/s$ below the kink temperature. This is broadly expected to be negative or 0, but has not yet been constrained conclusively by model-to-data comparison.

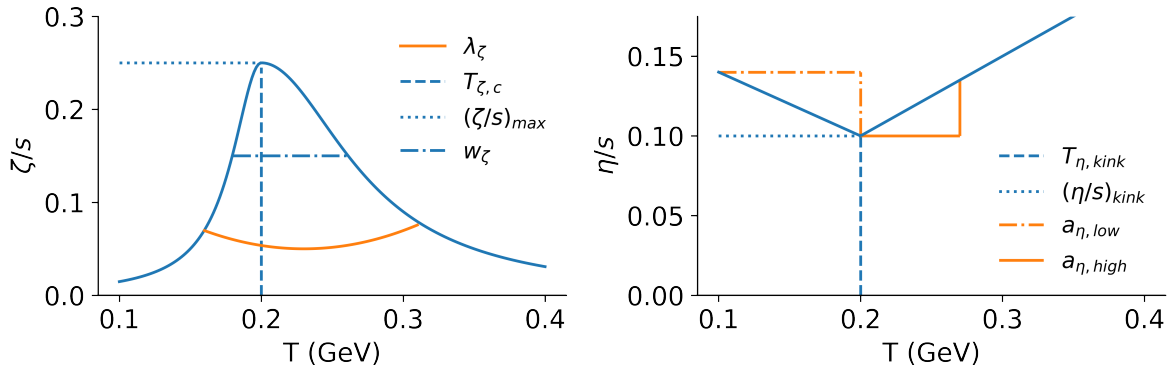


Figure 4.22: The parametrization of the viscosities.

4. $a_{\eta,high}$: The slope of $\eta/s$ above the kink temperature. This is anticipated to be positive or 0, but has not yet been constrained conclusively by model-to-data comparison. A theoretical exception to this expectation can be found in the NJL model for SU(3) [189].

5. $(\zeta/s)_{max}$: The maximum of $\zeta/s$.

6. $T_{\zeta,c}$: The temperature of the maximum of $\zeta/s$.

7. $w_\zeta$: The width of the peak in $\zeta/s$.

8. $\lambda_\zeta$: The asymmetry of the peak in $\zeta/s$.

9. `T_freeze`, also known as the particlization or switching temperature ($T_{sw}$): At this temperature, the hydrodynamic medium ceases to behave as a fluid and is converted to particles. A surface at constant temperature is drawn (assuming no baryochemical potential) from which particles are sampled by iS3D.

# 5

# BAYESIAN INVESTIGATION OF THE INITIAL STAGE

It is now time to turn to an application of the methods and models introduced thus far. The specific choices made in the hydrodynamic stage have already been thoroughly tested in other analyses, leaving the pre-equilibrium stage comparatively unexplored. In this exploratory study, this will be rectified by first performing tests of self-consistency, *i.e.* if known inputs can be recovered before proceeding to a preliminary comparison to data. This will systematically explore the free parameters in IP-Glasma for the first time and establish the sensitivity of final-state observables to variation in the pre-equilibrium stage.

First, the priors and observables used in this analysis are explicitly motivated and specified in Sec. 5.1. Then, tests of self-consistency are performed in Sec. 5.2 to validate the model and statistical workflow. A discussion of the results concludes this chapter in Sec. 5.3.

## 5.1 IMPLEMENTATION AND OBSERVABLES

The test of self-consistency, also known as a closure test, is performed with the physical parameters for IP-Glasma not already constrained by deep inelastic scattering experiments. This leaves three parameters: the strong coupling ($g_{\mathrm{strong}}$), the multiplier from the color charge density profile to the saturation scale ($\mu_{Q_s}$), and the switching time between IP-Glasma and MUSIC ($\tau_0$).

Two IP-Glasma parameters, $g_{\mathrm{strong}}$ and $\mu_{Q_s}$, should have a strong covariance as they usually appear together (see Sec. 2.1). However, it would be of consequence if this degeneracy is broken and this has yet to be systematically investigated. This would in turn allow for the first constraint on the strong coupling from heavy-ion phenomenology. As there has been no systematic study of observable sensitivity to these parameters, it is important to assess in a limited study what parameters may

or may not be possible to constrain. If a parameter does not impact final observables, there is no sense needlessly expanding the dimensionality of the parameter space in a larger study.

For all of the parameters in this study, the prior distribution is a Generalized Normal Distribution with shape parameter $\beta = 10$. Rather than specify the remaining parameters independently, which can be difficult to interpret, the central interval that contains 99% of the probability density is specified by identifying the values of the $0.5^{\text{th}}$ and $99.5^{\text{th}}$ percentiles of the distribution. The explicit parameters for the distributions are then found using numerical optimization.

The prior for the strong coupling is chosen with some knowledge of Fig. 1.2 in mind. 2+1D IP-Glasma uses a coupling that does not run with the energy scale and the form of the QCD coupling is $\alpha_s = \frac{g^2_{strong}}{4\pi}$. At $g_{strong} = 2.0$, the low-energy value of $\alpha_s$ is recovered ($\sim 0.3$). A simple test of the impact of parameter variation can be performed with an empirically-derived proxy between IP-Glasma energy and final-state multiplicity [28],

$$\frac{dN_{ch}}{d\eta} \approx 0.839 \left(\frac{dE}{\tau d\eta}\right)^{0.833} . \tag{5.1}$$

Varying $g_{strong}$ and holding all other parameters constant yields the proxy multiplicity dependence in Fig. 5.1. Allowing for an approximate factor of 2 variation from the value used by [28] (marked as "standard value") yields a 99% prior range of 1.30 to 3.25.

This same procedure with the proxy multiplicity is repeated to determine a broad prior range for the multiplier from the color charge density profile to the saturation scale, demonstrated in Fig. 5.2. Again allowing for an approximate factor of 2 variation in the prior yields a 99% prior range from 0.45 to 0.80. An interesting feature of both of these variations is that the dependence of the proxy multiplicity on both $g_{strong}$ and $\mu_{Q_s}$ is parametrically $1/x$. Thus, these take the form of an overall inverse normalization.

The final parameter to motivate a prior range for is the switching time between IP-Glasma and MUSIC, $\tau_0$. As seen in Fig. 4.21, the pressures do not start to come to a steady-state until approximately 0.2 fm. The current practice with IP-Glasma at in $\sqrt{s_{NN}} = 2.76$ TeV collisions is to wait until 0.4 fm, but recent works with varying pre-equilibrium models have favored increasingly large times to transition between the

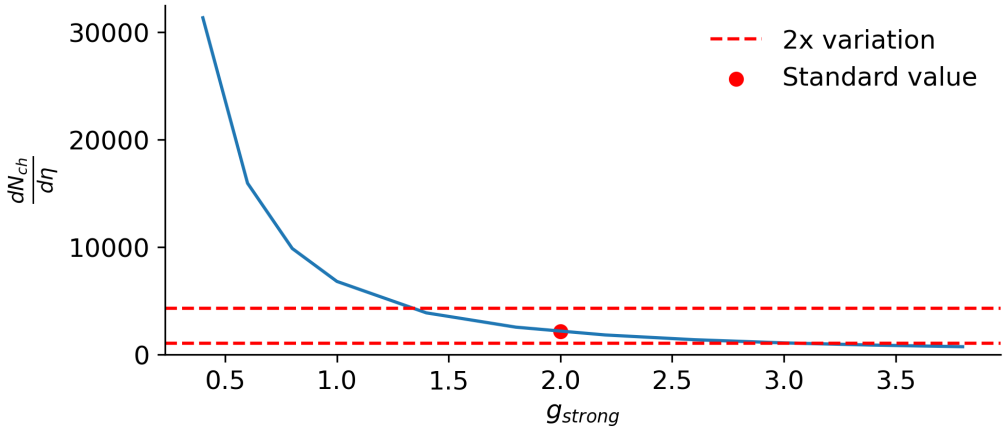


Figure 5.1: Multiplicity dependence of a central event on variation of $g_{strong}$ with all other parameters held fixed.

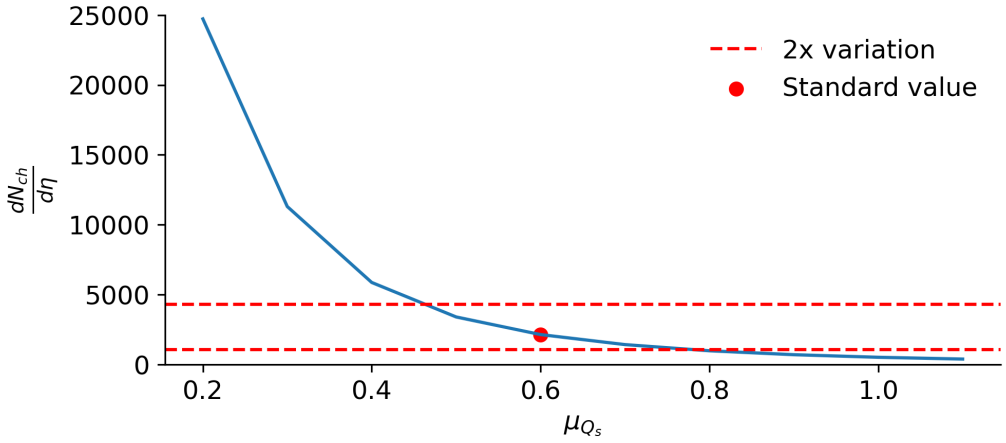


Figure 5.2: Multiplicity dependence in a central event on variation of $\mu_{Q_s}$ with all other parameters held fixed.

pre-equilibrium and hydrodynamic stages [33, 45]. As this has yet to be systematically explored and is of clear interest, $\tau_0$ is allowed to vary between 0.2 and 1.0 fm and its impact on observables will be investigated in this chapter. The final prior ranges are collected in Table 5.1. Not varied are the viscous parameters or the particlization temperature, which are fixed to the MAP values for Grad viscous corrections from [45]. In this chapter, only the Grad viscous corrections are investigated.

With the priors in hand, a design can be chosen. As there are only 3 dimensions and the primary goal of this study is to explore the parameter space, a Sobol Low Discrepancy sequence is used for the design. Sobol convergence property A is met

| Parameter | $0.5^{\text{th}}$ percentile | $99.5^{\text{th}}$ percentile |
|---|---|---|
| $g_{\text{strong}}$ | 1.30 | 3.25 |
| $\mu_{Q_s}$ | 0.45 | 0.80 |
| $\tau_0$ | 0.2 | 1.0 |

Table 5.1: Prior ranges for testing self-consistency in IP-Glasma.

with $2^3$ points and the curse of dimensionality does not impact this limited study. The GP emulators with a Gaussian + white noise kernel are trained with 30 main design points (10 points per dimension). The Gaussian (or Radial Basis Function) kernel characterizes the smoothness of the function draws and the white noise component accounts for the fluctuations in the training data. This combined kernel will also be used in the global Bayesian analysis in the following chapter.

In preliminary testing, it was found that 2500 events per design point are sufficient to calculate first generation and most next generation observables of interest described in Sec. 2.6. However, calculation of an observable does not necessarily guarantee successful surrogate modeling. While next generation observables were demonstrated to be a discriminating observable in Ch. 4, the choice of the best set of observables for a full comparison remains unclear and must be decided by both physical intuition and practical considerations.

The set of observables that are reliably calculated and distinguishable from statistical fluctuation are all of the first generation observables; the nonlinear response coefficients $\chi_{4,22}$, $\chi_{5,23}$, $\chi_{6,222}$, and $\chi_{6,33}$; the linear and nonlinear flow modes $v_4^{\text{L}}$, $v_4(\Psi_2)$, $v_5(\Psi_{23})$, $v_6(\Psi_2)$, $v_6(\Psi_3)$; and the event plane correlations $\rho_{422}$, $\langle\cos(4(\Phi_2-\Phi_4))\rangle$, $\langle\cos(6(\Phi_2-\Phi_3))\rangle$, $\langle\cos(6(\Phi_2-\Phi_6))\rangle$, $\langle\cos(4(\Phi_3-\Phi_6))\rangle$, $\langle\cos(2\Phi_2+3\Phi_3-5\Phi_5)\rangle$, $\langle\cos(2\Phi_2+4\Phi_4-6\Phi_6)\rangle$, and $\langle\cos(2\Phi_2-6\Phi_3+4\Phi_4)\rangle$. The calculations at each design point for these observables are shown in Fig. 5.3. Note that even though the hydrodynamic phase has not been explicitly calibrated with an IP-Glasma initial state, most of the data fall within the prior predictive range for this exploratory study. This stands in contrast to Ch. 4, where many of these observables were shown to not be reproduced with a T<sub>R</sub>ENTo + freestreaming initial state with the same hydrodynamic parameters. These correlations are thus isolated to the difference between pre-equilibrium models and

IP-Glasma clearly incorporates important physics by construction. However, this does not account for correlations between the observables.

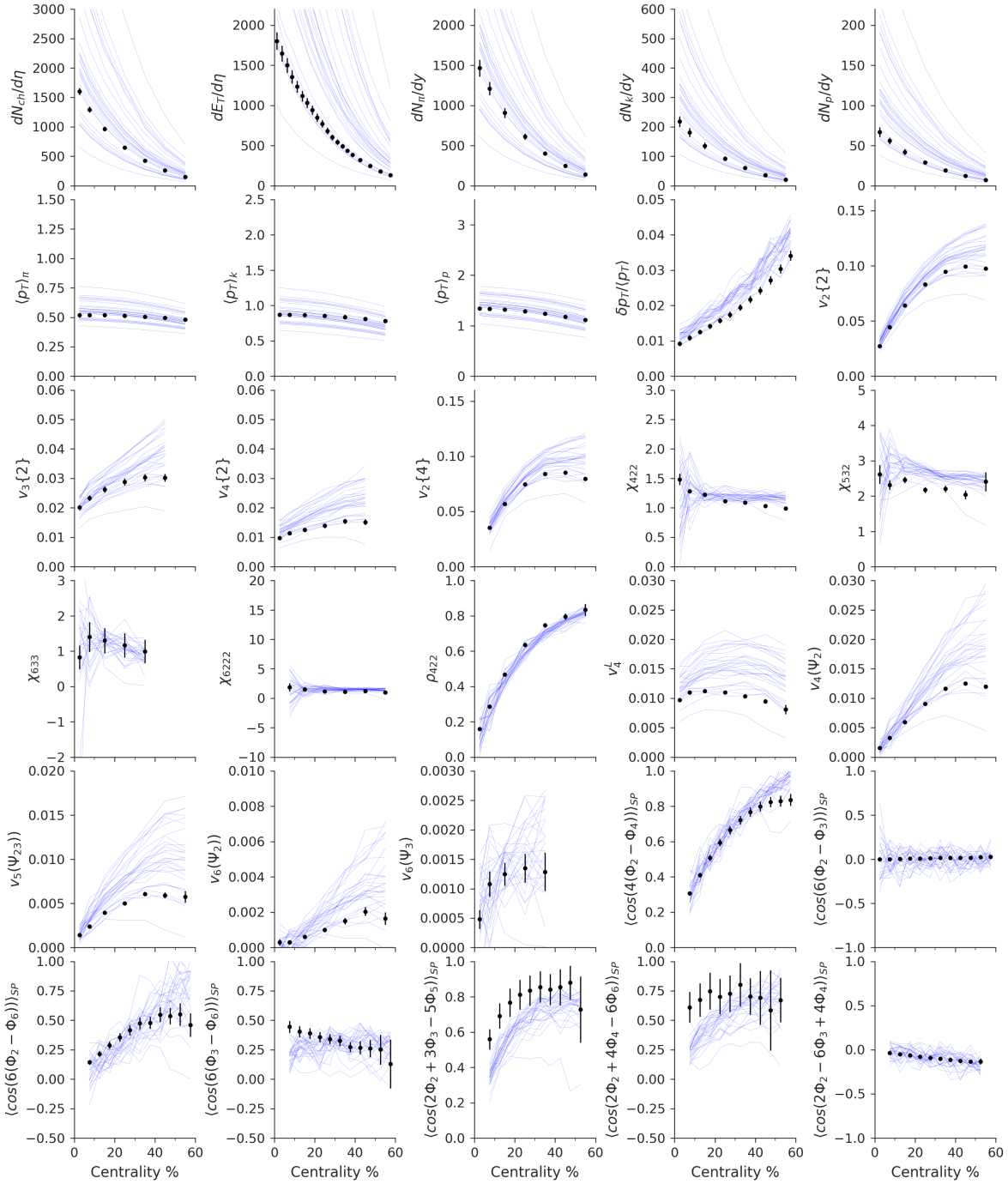


Figure 5.3: Calculations at each design point forming the prior predictive distribution for each observable. Points are experimental data.

To reduce the dimensionality of this data with minimal loss, PCA is used. First, it must be explicitly checked that the correlations between observables are linear as PCA is a linear transformation. Once this has been verified, PCA is performed. It was found that 20 principal components capture 96.205% percent of the variation in the data. Finally, a Gaussian Process emulator is trained on the principal components of the observables vs. the parameters. The predictions of this GP are then compared to validation points that were not included in the training set for each observable of interest, shown in Fig. 5.4.

In Fig. 5.4, a successful prediction of the validation point's observables is consistent with the line $y = x$, shown by a dashed line. Clearly, the surrogate model is not successful at making predictions for all these observables. This is often because the underlying model calculations are statistically under-powered or fluctuate about zero. For reliable outcomes from this analysis, the observable set must be restricted to only those observables that are successfully emulated. This set is comprised of the first generation observables; the flow modes $v_4^L$, $v_4(\Psi_2)$, $v_5(\Psi_{23})$, $v_6(\Psi_2)$, $v_6(\Psi_3)$; and the plane correlations $\rho_{422}$, $\langle\cos(4(\Phi_2 - \Phi_4))\rangle$, $\langle\cos(2\Phi_2 + 3\Phi_3 - 5\Phi_5)\rangle$, and $\langle\cos(2\Phi_2 + 4\Phi_4 - 6\Phi_6)\rangle$. The remaining observables then become excellent candidates for predictions with higher-statistics calculations. Once these observables have been selected, the PCA and GP emulation is repeated and are sufficiently reliable for performing self-consistency tests.

## 5.2 SELF-CONSISTENCY

Tests for self-consistency, also known as closure tests or empirical coverage tests, are important to assess whether a model is trustworthy. After all, a model that cannot recover its own inputs – *i.e.* is not self-consistent – cannot be trusted to reveal the truth in experimental data. The self-consistency tests are also less of a challenge for the model as there is no tension in the data as the model in testing generated the data exactly. In an ideal scenario for a single-valued posterior, the parameters map one-to-one to values of the observables. As a result, the posterior on the parameters would become a delta function at each known truth. However, this is smeared in practice by

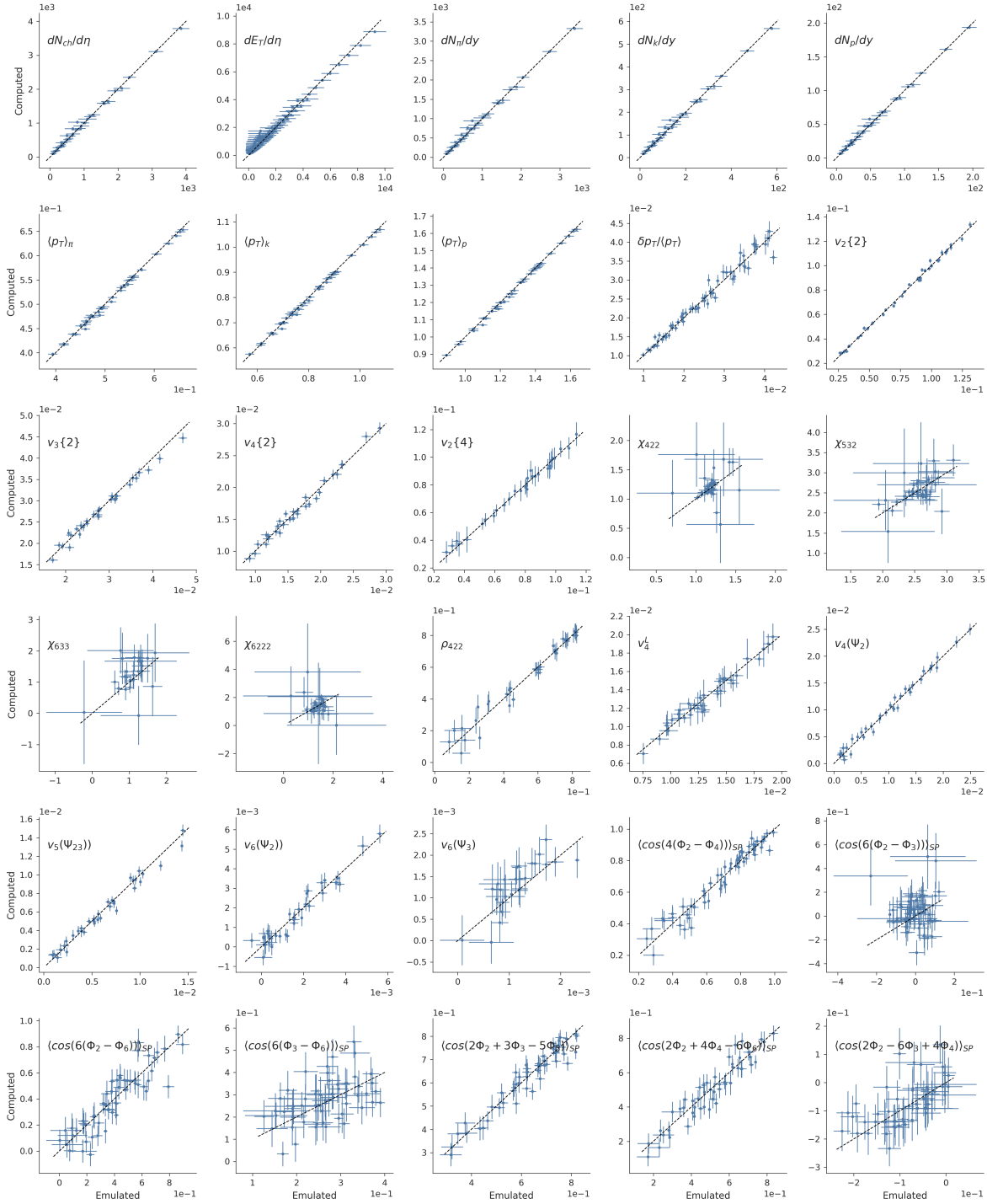


Figure 5.4: Emulated vs. computed for all observables considered. Successful emulation is clustered around $y = x$, shown as a dashed line. Error in the x direction is emulator uncertainty while error in the y direction is uncertainty from the hybrid model.

both uncertainties in the parameter estimation and the underlying data as well as by any degeneracy in the parameter space.

In these tests of self-consistency, Bayesian inference is performed with a validation point not included in the training set for the GP. Then, the resulting posterior can be compared to the known true values and the degree of constraint can be assessed. Finally, this exploratory study will conclude with a preliminary calibration to data given the MAP parameter values for the hydrodynamic stage from [45]. For conciseness, not all the closure test results are shown.

The first closure test is performed on a validation point with true parameter values $(g_{\text{strong}}, \mu_{Q_s}, \tau_0) = (2.11587, 0.49354, 0.624)$. A total of 10,000 MCMC steps are taken and the chains have clearly come to the stationary distribution with autocorrelation dropping almost immediately to zero, shown in Fig. 5.5. The posterior distribution for this test is shown in Fig. 5.6 and the posterior predictive distribution is shown in Fig. 5.7. Constraint tighter than experimental uncertainty is due to the fact that some of the observables appear exclusively sensitive to the hydrodynamic phase, which is not varied in this study. While the data is successfully reproduced by the posterior predictive distribution, a strong covariance can be seen between $g_{\text{strong}}$ and $\mu_{Q_s}$. The bimodality in this covariance is not of significance and is likely due to a small number of design points.

This strong covariance is a feature seen in all validation points and with a relatively sparse 30-point design, any breaking of this covariance is just as likely to be a feature of the emulation as it is to be physical. This suggests that the theoretical understanding was correct and $g_{\text{strong}}$ and $\mu_{Q_s}$ are totally degenerate and one of the two parameters should be fixed. Because $g_{\text{strong}}$ is well-constrained by experiment while there is no experimental guidance for $\mu_{Q_s}$, the natural choice is to fix $g_{\text{strong}}$.

With $g_{\text{strong}}$ fixed to the known truth, Bayesian inference can be performed again and the results can be investigated further. The MCMC behaviour is comparable to that of Fig. 5.5 and is omitted for conciseness. The posterior distribution is shown in Fig. 5.8 and the posterior predictive distribution as well as the validation data are shown in Fig. 5.9. With the covariance between $g_{\text{strong}}$ and $\mu_{Q_s}$ resolved by fixing the strong coupling, good constraint is found and the known truths are well constrained within the posterior. This performance in constraining the posterior is comparable

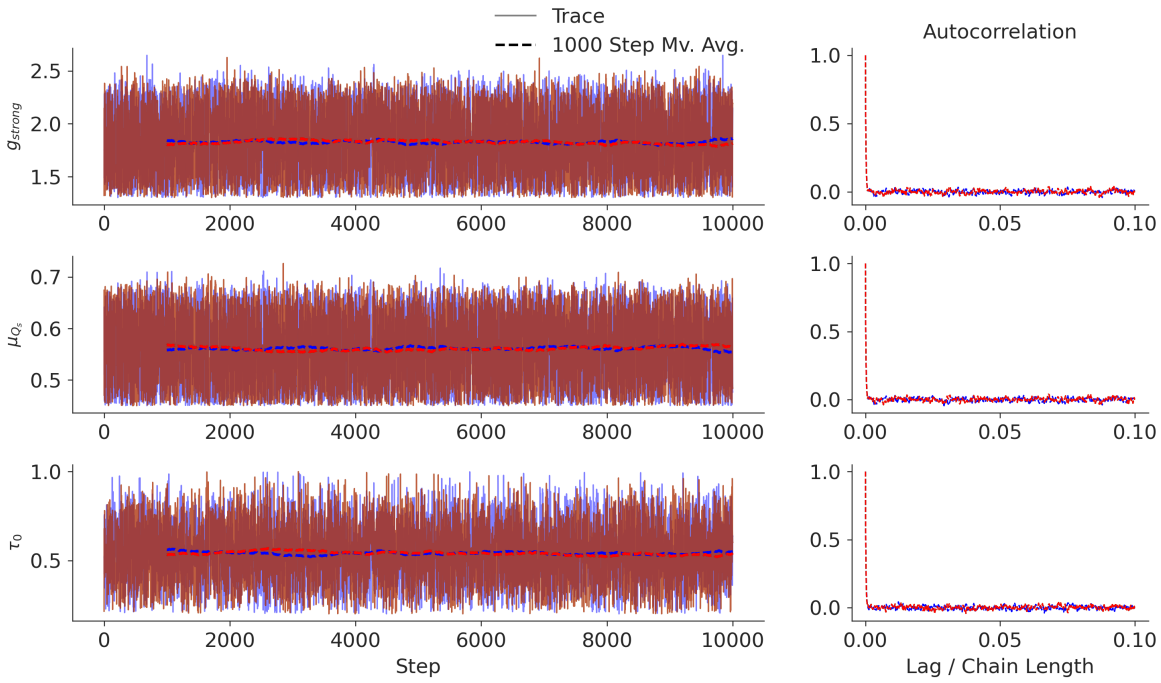


Figure 5.5: MCMC chain trace and autocorrelation after calibration to a validation point with true parameter values $(g_{strong}, \mu_{Q_s}, \tau_0) = (2.11587, 0.49354, 0.624)$.

across all validation points, suggesting that once the $g_{strong} - \mu_{Q_s}$ covariance is removed, IP-Glasma's free parameters are excellent candidates for study in a Bayesian inference framework.

## 5.3 DISCUSSION

This exploratory study has been the first application of Bayesian methods to a hybrid model with an IP-Glasma initial state and the first systematic model-to-data comparison of the free parameters in IP-Glasma. Observable selection was demonstrated and self-consistency of the IP-Glasma initial state was established with confirmation that the strong coupling and the proportionality between the color charge density and saturation scales are degenerate. The closure tests demonstrate that the true values are well-contained within the inferred posterior regions

The primary goals of this study were to investigate potential covariances in IP-Glasma and to establish if sufficient constraint were possible on the parameters of IP-

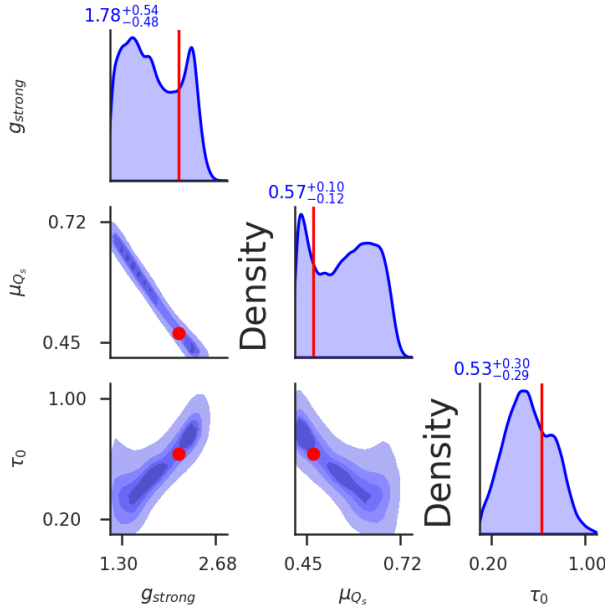


Figure 5.6: Posterior distribution after calibration to a validation point with true parameter values ($g_{strong}$, $\mu_{Q_s}$, $\tau_0$) = (2.11587, 0.49354, 0.624). 1-dimensional marginal distributions are along the diagonal and off-diagonal elements are the 2-dimensional marginal distributions. True values are shown in red and the median and central 95% interval are given. The quoted values are the median and 95% credible interval.

Glasma not fixed by comparison to DIS experiments. These goals were successfully met and have been demonstrated by the tests for self-consistency.

This exploratory analysis also aimed to develop an automated workflow for producing a design space, running the hybrid model at design points, performing analysis on each design point, and collecting the results into a format compatible with existing Bayesian analysis infrastructure. These goals were met and successfully prepared the IP-Glasma hybrid model workflow for execution at scale on high performance computing systems.

Finally, a brief word must be included on the value of closure tests. In addition to demonstrating the self-consistency of a model – and thus the reliability of comparisons to data – it must be emphasized that closure tests are an idealized comparison in which the model contains all the information in the data. This is not the case when comparing to experimental data. A healthy degree of caution must be exercised regarding the comparison between the degree of constraint achieved in closure tests to

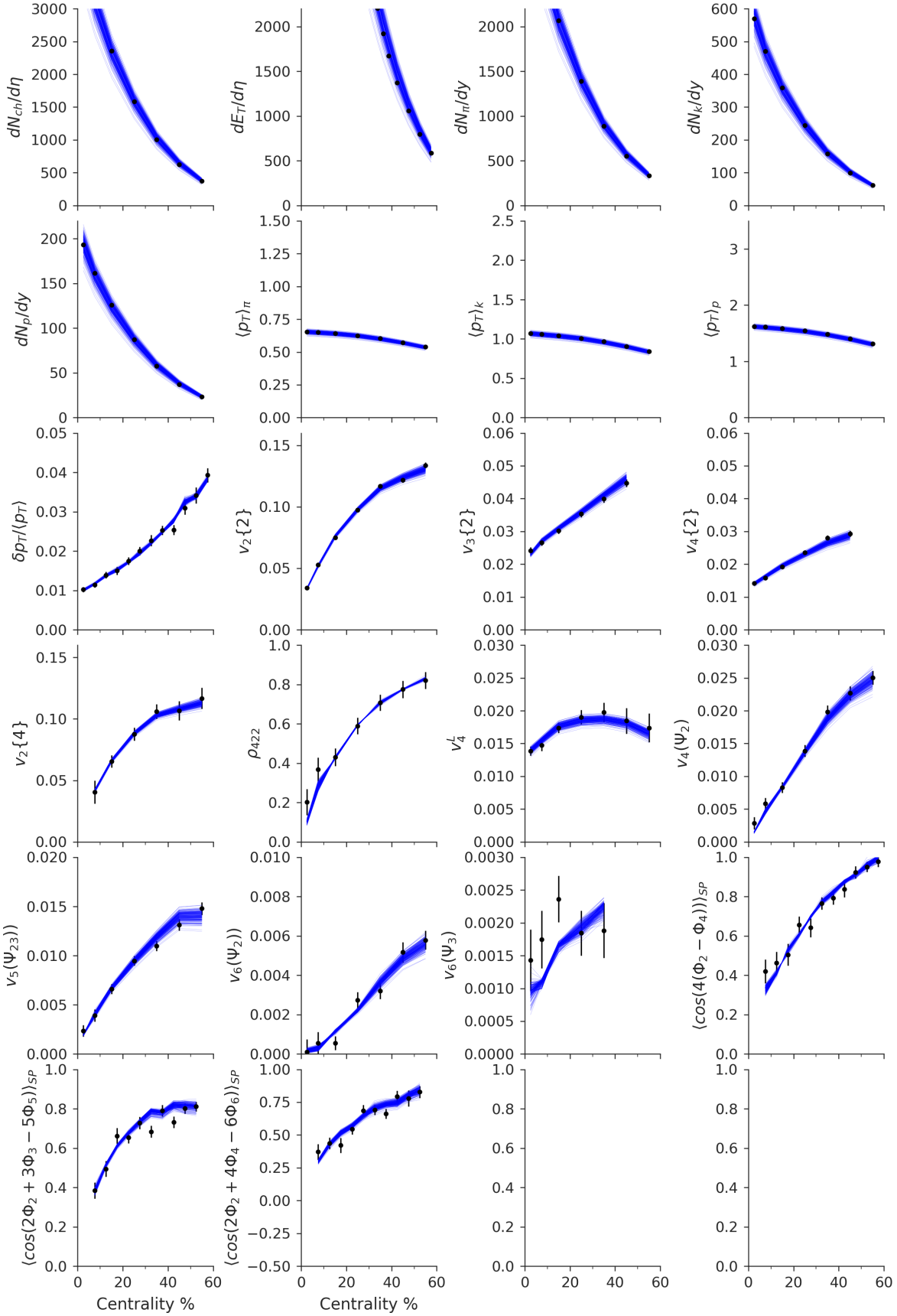


Figure 5.7: Posterior predictive distribution after comparison to a validation point with true parameter values $(g_{strong}, \mu_{Q_s}, \tau_0) = (2.11587, 0.49354, 0.624)$. Validation data are shown as black points and the posterior predictive distribution is shown in blue.

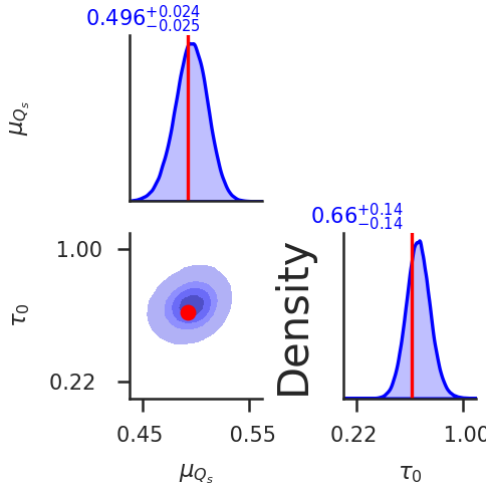


Figure 5.8: Posterior distribution after comparison to a validation point with true parameter values ($g_{strong}$, $\mu_{Q_s}$, $\tau_0$) = (2.11587, 0.49354, 0.624) and holding $g_{strong}$ fixed to its true value. 1-dimensional marginal distributions are along the diagonal and off-diagonal elements are the 2-dimensional marginal distributions. True values are shown in red and the median and central 95% interval are given.

the degree of constraint when comparing to data. Nonetheless, closure tests remain a valuable laboratory for ensuring the validity of analysis before comparing to experimental measurements and are an indispensable way to understand the limitations of a model before deployment.

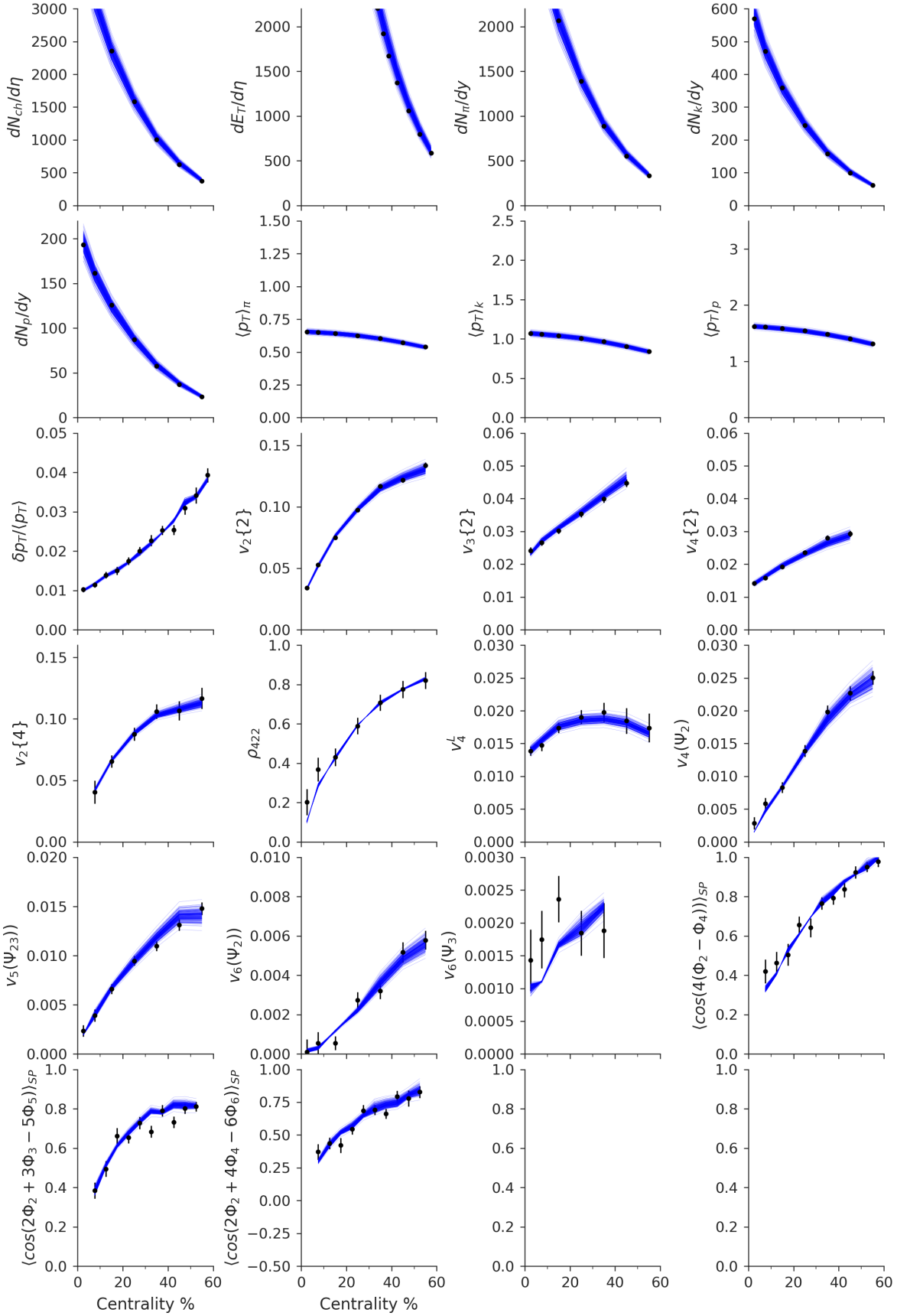


Figure 5.9: Posterior predictive distribution after comparison to a validation point with true parameter values ($g_{strong}$, $\mu_{Q_s}$, $\tau_0$) = (2.11587, 0.49354, 0.624), holding $g_{strong}$ fixed to its true value.

# 6

# A MODERN BAYESIAN INFERENCE FOR HEAVY ION COLLISIONS

With pertinent motivation in hand and self consistency established, all the pieces are in place for a Bayesian study of an IP-Glasma hybrid model with simultaneous variation of pre-equilibrium and hydrodynamic quantities. In this chapter, I begin by explicitly motivating priors and choice of design space sampling method. The model selection process from Ch. 5 is repeated to determine what physically-motivated observables are can be practically implemented. Once this is complete, tests for self-consistency are performed for the full parameter space. Then, comparison to data is performed with Grad's 14-moment viscous corrections. Finally, transfer learning is used to perform closure tests and comparison to data using the Chapman-Enskog RTA viscous corrections. A discussion of the results concludes this chapter.

## 6.1 PRIORS

The first step is to again specify prior knowledge. The motivation behind specifying the IP-Glasma parameters varied in this study, $\mu_{Q_s}$ and $\tau_0$ are unchanged from Ch. 5 although the prior for $\tau_0$ is extended to even later times to ensure late times are properly explored.

The remaining parameters whose priors must be motivated are those of the hydrodynamic stage: the 8 parameters of the specific shear and bulk viscosity as well as the particlization temperature. For the parameters of the specific shear viscosity, the parametrization in Eq. 4.10 is used and some inspiration was taken from a lack of ability to constrain the temperature dependence in previous studies [45]. As a result, the priors for the parametrization of the specific shear viscosity were widened, but the same belief was kept that it is more likely for the slope below the kink to be negative and the slope above the kink to be positive, although the opposite is also ex-

| Parameter | $0.5^{\text{th}}$ percentile | $99.5^{\text{th}}$ percentile | $\beta$ | Distribution |
|---|---|---|---|---|
| $\mu_{Q_s}$ | 0.55 | 0.90 | 10 | Generalized Normal |
| $\tau_0$ [fm] | 0.20 | 1.20 | 20 | Generalized Normal |
| $T_{\eta,kink}$ [GeV] | 0.120 | 0.320 | 20 | Generalized Normal |
| $a_{\eta,low}$ [GeV$^{-1}$] | -2.10 | 1.20 | 20 | Generalized Normal |
| $a_{\eta,high}$ [GeV$^{-1}$] | -1.20 | 2.10 | 20 | Generalized Normal |
| $(\eta/s)_{kink}$ | 0.00 | 0.30 | 10 | Half Generalized Normal |
| $(\zeta/s)_{max}$ | 0.00 | 0.30 | 10 | Half Generalized Normal |
| $T_{\zeta,c}$ [GeV] | 0.100 | 0.350 | 10 | Generalized Normal |
| $w_\zeta$ [GeV] | 0.02 | 0.18 | 30 | Generalized Normal |
| $\lambda_\zeta$ | -1.0 | 1.0 | 20 | Generalized Normal |
| $T_{sw}$ [GeV] | 0.135 | 0.180 | 10 | Generalized Normal |

Table 6.1: Prior hyperparameters and distributions for each parameter varied.

plored. For the specific bulk viscosity, the parametrization in Eq. 4.11 is used and the same procedure was applied, but also incorporated knowledge of the prior success of peaked bulk viscosities and the physical motivation for their form [190]. Finally, the particlization temperature is believed to be around the crossover temperature, but recent studies have found success in lower particlization temperatures. Both will be explored in this study. The form of the priors is again the Generalized Normal Distribution (Version 1) or the half Generalized Normal Distribution (Version 1) if the quantity is commensurate with a sharp cutoff (*e.g.* is required to be positive definite). Unlike in the exploratory study, the value of $\beta$ is not kept uniform, but is instead modified on a case-by-case basis to depending on how quickly the tails of the prior drop off. The full set of parameter priors, with the central 99% range and the Generalized Normal distribution shape parameter $\beta$ is collected in Table 6.1. The viscous priors with respect to temperature are shown in Fig. 6.1.

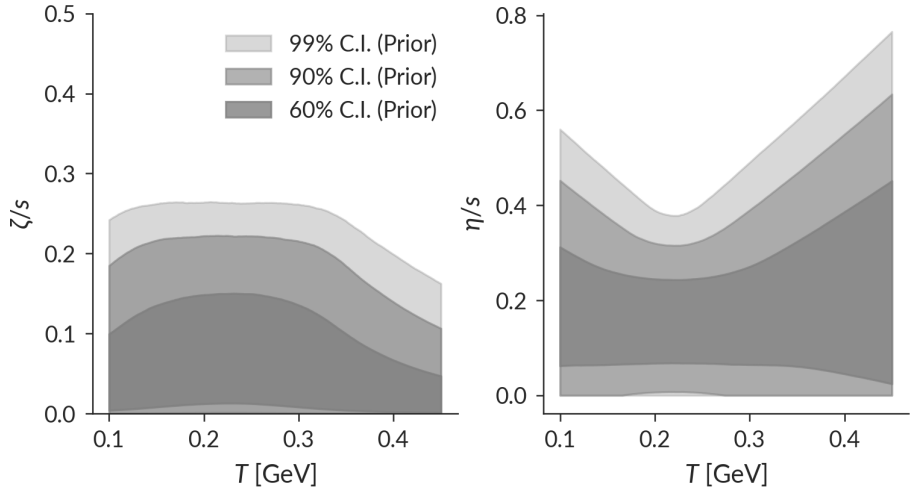


Figure 6.1: Viscous priors with respect to temperature with credible intervals (C.I.). The specific bulk viscosity is shown on the left and the specific shear viscosity is on the right.

## 6.2 SAMPLING THE DESIGN SPACE

Unlike in Ch. 5, the goal of this design is to maximize the emulator performance. Additionally, in an 11-dimensional study, the convergence properties of the Sobol LDS make it an untenable choice by requiring nearly 200 points per parameter dimension. As the goal is a computationally-efficient choice of design maximizing surrogate model performance, the clear choice is the Maximum Projection design. This design is sampled on the 11-dimensional unit hypercube $[0, 1]^{11}$ and deformed according to the specified priors.

The number of design points must maximize the underlying model performance given the constraint of a finite computational budget. The initial state closure test demonstrated that next generation observables can be reliably calculated with 2500 events per design point. With a single design point requiring approximately 1.25 core-years, the budget allows for 350 design points in the training set as well as an additional 50 from an independent design to form the validation set. This translates to approximately 32 points per parameter dimension, well above the rule-of-thumb of 10 points per parameter dimension, suggesting that emulator performance will outperform that in the exploratory study. Additionally, these design points will be

ordered to maximize the Maximum Projection criteria for the design containing the 0 to $n^{th}$ point, allowing for intermediary results [3].

Additional computational resources are reserved for repeating the particlization and hadronic cascade for the 50 most important training design points as well as 20 validation design points with Chapman-Enskog viscous corrections. This will allow for transfer learning to account for this source of modeling uncertainty. Further computing resources are reserved higher-statistics runs at the MAP points for both viscous corrections found in inference.

## 6.3 OBSERVABLE SELECTION AND EMULATOR FORWARD VALIDATION

Once again, it is important to investigate which parameters are both reliably calculated with the underlying hybrid model and are reliably emulated by the Gaussian process surrogate model. This step, revisited at the outset of each study, must be performed to ensure that predictions made by the GP are sensible and will provide physical – rather than spurious – constraint. The set of observables that are reliably calculated and distinguishable from statistical fluctuations are again all of the first generation observables; the nonlinear response coefficients $\chi_{4,22}$, $\chi_{5,23}$, $\chi_{6,222}$, and $\chi_{6,33}$; the linear and nonlinear flow modes $v_4^L$, $v_4(\Psi_2)$, $v_5(\Psi_{23})$, $v_6(\Psi_2)$, $v_6(\Psi_3)$; and the event plane correlations $\rho_{422}$, $\langle\cos(4(\Phi_2-\Phi_4))\rangle$, $\langle\cos(6(\Phi_2-\Phi_3))\rangle$, $\langle\cos(6(\Phi_2-\Phi_6))\rangle$, $\langle\cos(4(\Phi_3-\Phi_6))\rangle$, $\langle\cos(2\Phi_2+3\Phi_3-5\Phi_5)\rangle$, $\langle\cos(2\Phi_2+4\Phi_4-6\Phi_6)\rangle$, and $\langle\cos(2\Phi_2-6\Phi_3+4\Phi_4)\rangle$. The calculation of these observables at each design point are shown in Fig. 6.2. PCA is once again performed and 30 principal components explain 90.642% of the variance in the calculations. Finally, emulators are trained on the PCs vs. the parameters and predictions for observables can be made. The emulator predictions at validation points vs. the computed results are shown in Fig. 6.3.

Observables that are loosely clustered along the $y = x$ lines in Fig. 6.3 are not kept for the final analysis and observables that are extremely uncertain are also not included. This constitutes "forward model validation." Given a known set of inputs, the predictions are compared to model calculations and observables the surrogate model predicts poorly are inappropriate for inclusion in a physics study. Finally, the $\langle\cos(4(\Phi_2-\Phi_4))\rangle$ event plane correlator is not included as it quantifies the same

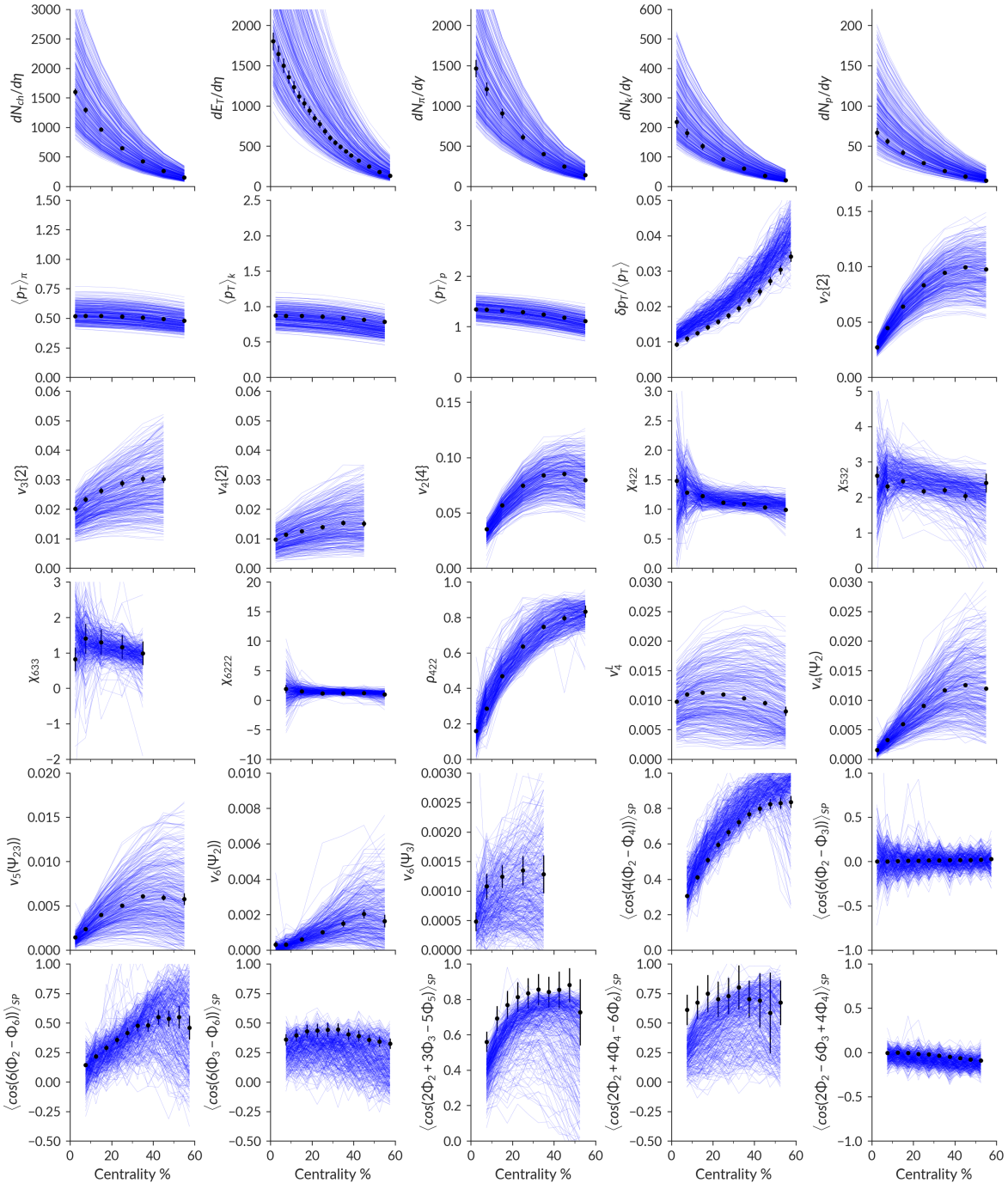


Figure 6.2: Calculations at each design point forming the prior predictive distribution for each observable. Points are experimental data.

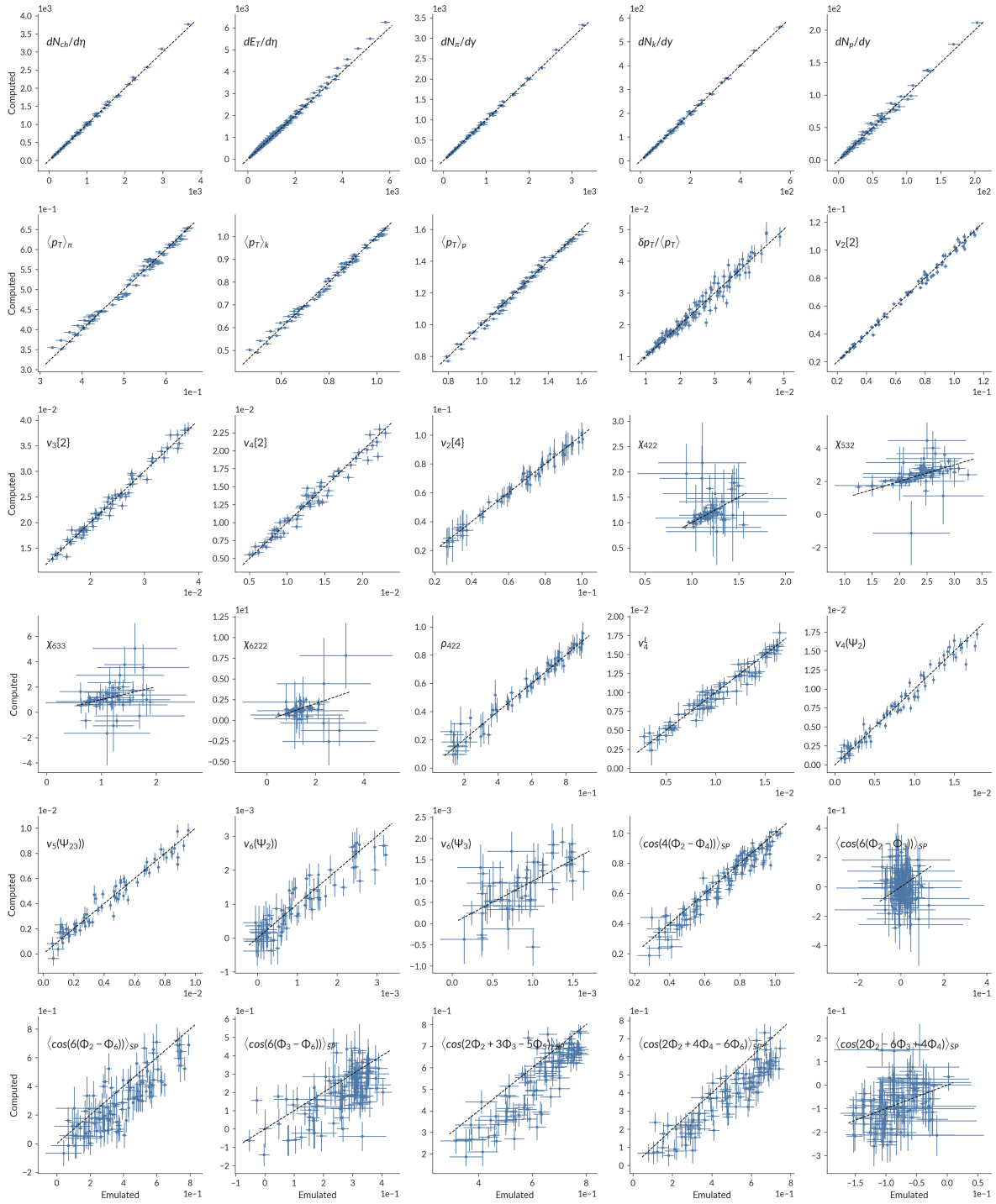


Figure 6.3: Emulated vs. computed for all observables considered. Successful emulation is clustered around $y = x$, shown as a dashed line. Error in the x direction is emulator uncertainty while error in the y direction is uncertainty from the hybrid model.

correlation as the $\rho_{422}$ correlator and has an overall bias. The finalized set of observables for testing self-consistency and comparison to data is comprised of the first generation observables; the flow modes $v_4^L$, $v_4(\Psi_2)$, $v_5(\Psi_{23})$, $v_6(\Psi_2)$; and the plane correlations $\rho_{422}$, $\langle\cos(2\Phi_2 + 3\Phi_3 - 5\Phi_5)\rangle$, and $\langle\cos(2\Phi_2 + 4\Phi_4 - 6\Phi_6)\rangle$. The remaining observables are again excellent candidates for predictions with higher-statistics calculations to test the posterior state of knowledge. Once these observables have been selected, the PCA and GP emulation is repeated and are found to be sufficiently reliable for performing self-consistency tests and comparisons to data. Further details of the principal component analysis for the final observable set are shown in Fig. 6.4, where the relationship between the first three principal components (PCs) are shown as well as the cumulative explained variance fraction. The first few principal components contain the majority of the variance of the data and it can be clearly seen that the first three PCs relate clearly to the observables, further supporting the idea that they are successfully reducing the dimensionality of the data with minimal loss of underlying signal. With the final observable set, 30 principal components explain 97.94% of the variance in the data. The full set of principal components to explain the total variance in the data consists of 161 PCs, meaning that the remaining 131 principal components represent 2.06% of the variance in the data, which is almost certainly dominated by noise in the underlying calculations. Note that the presence of exclusively linear correlations between observables must be investigated for the final set of chosen observables. This investigation was performed and correlations were found to be sufficiently linear for PCA to be a valid transformation, but the resulting figure is sufficiently large (a 334 x 334 matrix of plots to show pairwise combinations of every observable in every centrality) as to not fit in this thesis.

### 6.3.1 *Transfer learning for Chapman-Enskog $\delta$f*

The viscous corrections are an important source of uncontrolled theoretical uncertainty to quantify. The most computationally-efficient way to do so is by using transfer learning, where a linear transfer learning model was introduced in Ch. 3.5.1.1. This uses information learned from a source system – in this study, the already validated Grad viscous correction – to learn about a similar target system, the Chapman-

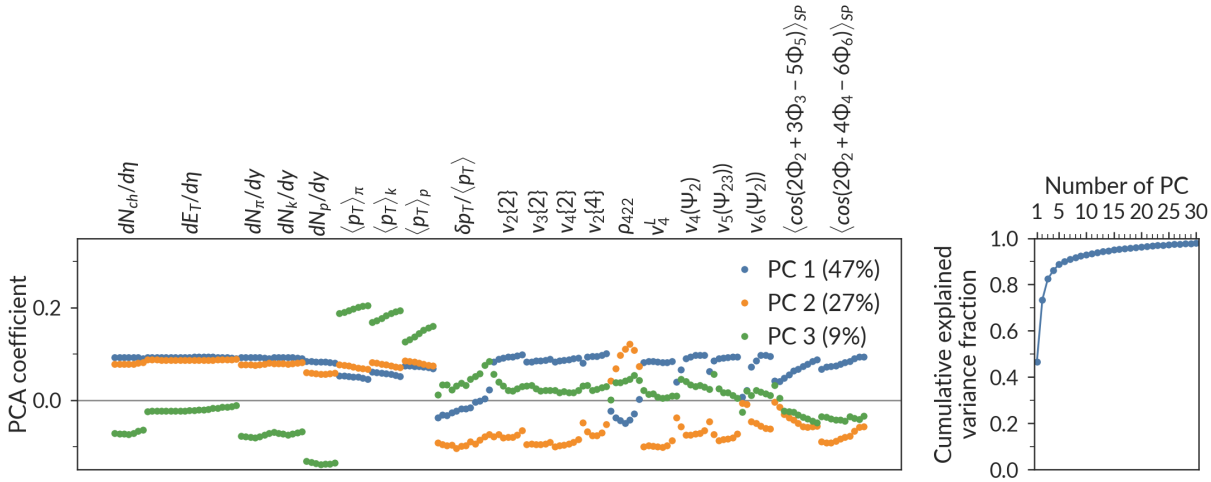


Figure 6.4: Observable relation to the first three principal components (left) and cumulative explained variance fraction (right).

Enskog RTA $\delta f$. By construction, these are both linearized viscous corrections and are designed to be small corrections to the equilibrium distribution function. This is a prime opportunity to use transfer learning to enable Bayesian inference for the first time in heavy ion collisions.

Transfer learning is implemented using *emukit* and *GPy*'s [191, 192] multifidelity emulation framework and builds off work introduced as proof-of-concept in [4]. Nonetheless, the proof-of-concept was not yet ready for deployment in a full-scale study. The proof-of-concept did not yet incorporate principal component analysis nor did it return the covariance matrix necessary for evaluating the likelihood function.

The information contained in the principal component analysis for the Grad viscous corrections (see Fig. 6.4) is exploited so that the transfer learning can take place on the principal components. The Grad PCA, trained on a large number of design points, can be understood to perform a critical covariance-revealing and noise-filtering function. By acting as a rotation in the observable space, the true underlying signal is contained in the first N PCs and noise fluctuations are reduced. This reveals mutual information between observables, *e.g.* that one can be fairly confident of $dN_{ch}/d\eta$ in the 30-40% bin given its value in the 0-5% bin. It also means that observables that require higher statistics to calculate reliably, such as $\delta p_T/\langle p_T \rangle$, become correlated with observables that do not, resulting in noise reduction and more successful surrogate modeling. Additionally, by training the transfer learning emulator

on the same principal components as the source emulator, the comparison between the two is put on an even footing.

A second improvement to the transfer learning is using "transformed parameters", introduced and used in [45, 46, 193] and the above work. Although the parametrization of the specific shear and bulk viscosity may appear intuitive and concise, it can present challenges to nonparametric models such as Gaussian processes as the relationship between the observables and these parameters can be highly non-linear and non-uniform. However, observables are often more straightforwardly-dependent on the value of the specific shear and bulk at a given temperature. Because for any one set of parameters in Eqs. 4.10 and 4.11 there exists one and only one set of values of $\eta/s$ and $\zeta/s$ at a set of temperatures, a one-to-one mapping takes place. Thus, no information is gained or lost by performing this transformation. By using the transformed observables, the transfer learning emulator's mean squared error was reduced by a factor between 2 and 20 for every observable considered as well as corresponding improvement in the distance between the coefficient of determination $R^2$ and its maximum value of one. Finally, software changes were made to make it indistinguishable from the original Emulator object and therefore compatible with existing MCMC software and ready for use. This software, as well as the general improvements to the heavy-ion collisions Bayesian software implemented for this thesis, will be made available on GitHub [194].

The transfer learning emulator validation begins with comparing emulated predictions to computed values at validation points not used in training, shown in Fig. 6.5. All the observables considered for the study with Grad viscous corrections are well predicted by the transfer learning model, in some cases even better than the source emulator trained on the full design. Uncertainties are often larger in the transfer learning model than in the Grad emulator, but this does not interfere significantly with the quality of predictions and is consistent with having two Gaussian Processes, each with their own variance, rather than just one. Predictions by the transfer learning emulator are broadly consistent with the true values and the emulator uncertainty is well-balanced with the computed uncertainty in the most statistics-hungry observables. Were one source of uncertainty systematically larger than the other, this would suggest imbalance between the number of design points and the number of model

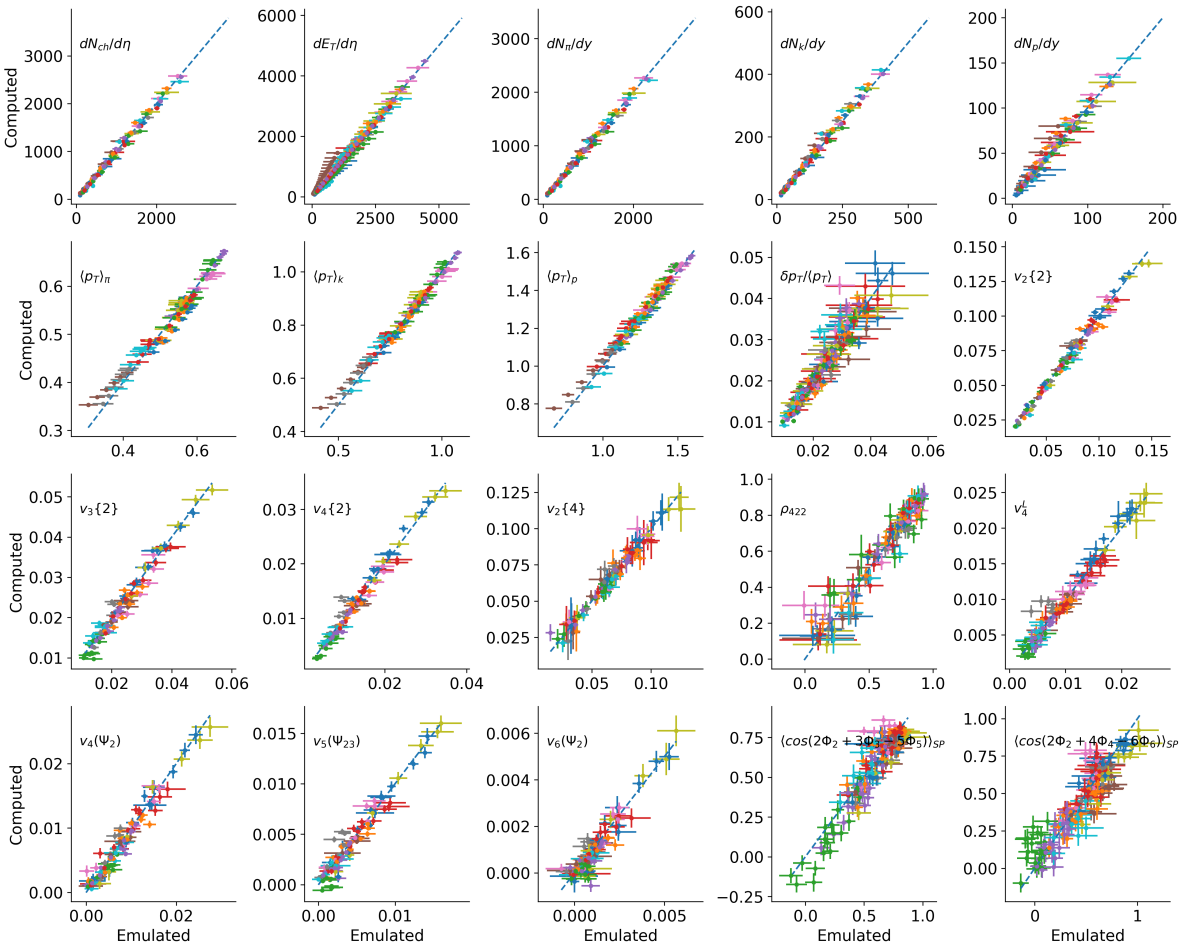


Figure 6.5: Transfer learning emulated vs. computed for all observables considered. Validation points are shown with a consistent color to identify correlations between points. The diagonal dashed line is located at $y = x$, denoting perfect prediction.

runs at each design point, which must be judged by the most statistics-hungry calculations. In this case, there are the correlated momentum fluctuations and event plane correlators: $\delta p_T/\langle p_T\rangle$, $\rho_{422}$, and $\langle\cos(2\Phi_2 + 4\Phi_4 + 6\Phi_6)\rangle_{SP}$. Further worth highlighting is what appears to be slight emulator bias in the three-plane correlators in Fig. 6.3 is resolved in the transfer learning emulation, suggesting yet-more accurate predictions.

### 6.3.2 *Observable sensitivity*

The first physics results of this chapter can be derived now that a surrogate model has been trained and validated with a reliable set of observables. Using the Gaussian process emulator as a surrogate, the global sensitivity of the model to variation

of input parameters can be calculated. This analysis is an example of "analysis of variance" (ANOVA) that decompose the total variance of a model into variance of model parameters (at first order), pairs of model parameters (second order), and so on. The first-order Sobol indices quantify the global variance in model observable due to variance in model parameters [195] and are readily available [196, 197].

For a given observable output $y$, suppose it can be represented as a function of model parameters $\mathbf{x}$, $y = f(\mathbf{x})$. A prior predictive distribution $p(y)$ is produced for each output by marginalization,

$$p(y) = \int d\mathbf{x} p(y|\mathbf{x}) p(\mathbf{x}). \tag{6.1}$$

The quantity of interest is, however, the variance associated with a single parameter. In this case, suppose one fixed a single parameter $x_i$ to take a particular value $a$. The variance of the resulting distribution of outputs can be readily computed,

$$p(y|x_i = a) = \int dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n p(y|\mathbf{x}) p(\mathbf{x}) \tag{6.2}$$

where $\mathbf{x}$ consists of $n$ elements. The variance of this distribution is $\mathrm{Var}(y)|_a \equiv \mathrm{Var}(p(y|x_i = a))$ and is the variance of the observable $y$ due to varying all parameters except $x_i$, *i.e.* conditional on $x_i = a$. By marginalizing over possible values of $x_i$, determined in turn by the prior, the variance due to variation of $x_i$ is found,

$$\mathrm{Var}(y)|_{x_i} = \int da \mathrm{Var}(y)|_a p(a). \tag{6.3}$$

The first-order Sobol sensitivity index $S_1$ for a parameter $x_j$ and observable $y$ is then

$$S_1[x_j] \equiv \frac{\mathrm{Var}(y) - \mathrm{Var}(y)|_{x_j}}{\mathrm{Var}(y)}, \tag{6.4}$$

the *fractional* variance in the observable from variation of parameter $x_j$ alone. Therefore, if $S_1[x_j] = 0.7$, this is interpreted as 70% of the global variation in this observable being ascribed to variation of $x_j$ alone.

The first-order Sobol sensitivities of the observables in the most central centrality bin are shown in Figs. 6.6-6.10. Due to length, they are divided into the following groups of observables: multiplicities and transverse energy in Fig. 6.6; identified hadron mean transverse momentum and correlated transverse momentum fluctuations in Fig. 6.7; anisotropic flow in Fig. 6.8; flow modes $v_4^L$, $v_4(\Psi_2)$, $v_5(\Psi_{23})$, $v_6(\Psi_2)$

in Fig. 6.9; and the plane correlations $\rho_{422}$, $\langle\cos(2\Phi_2 + 3\Phi_3 - 5\Phi_5)\rangle$, and $\langle\cos(2\Phi_2 + 4\Phi_4 - 6\Phi_6)\rangle$ in Fig. 6.10. The similar sensitivity of the two viscous correction models, almost always overlapping, demonstrates that the models are similar, but not identical, and that the viscous corrections are small compared to the overall effect of parameter variation.

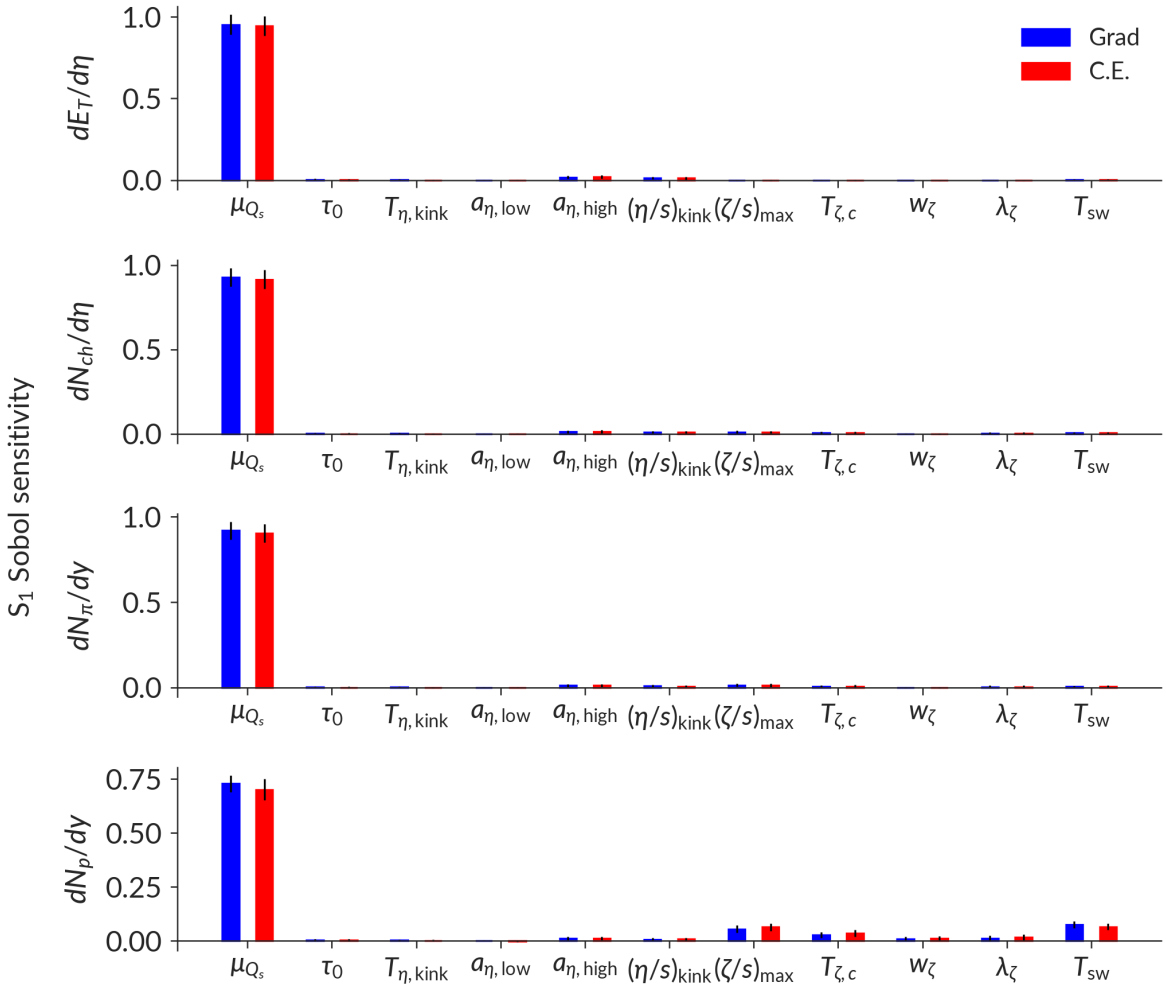


Figure 6.6: First-order Sobol sensitivity of charged hadron multiplicity, identified particle multiplicity, and transverse energy to input parameters.

In these Sobol sensitivities, physical intuition is confirmed - an important step in further verifying that the model behaves as expected. The multiplicities and transverse energy (Fig. 6.6) are dominantly sensitive to the overall normalization ($\mu_{Q_s}$) and the viscosity presents a small correction. The proton multiplicity, one of the most

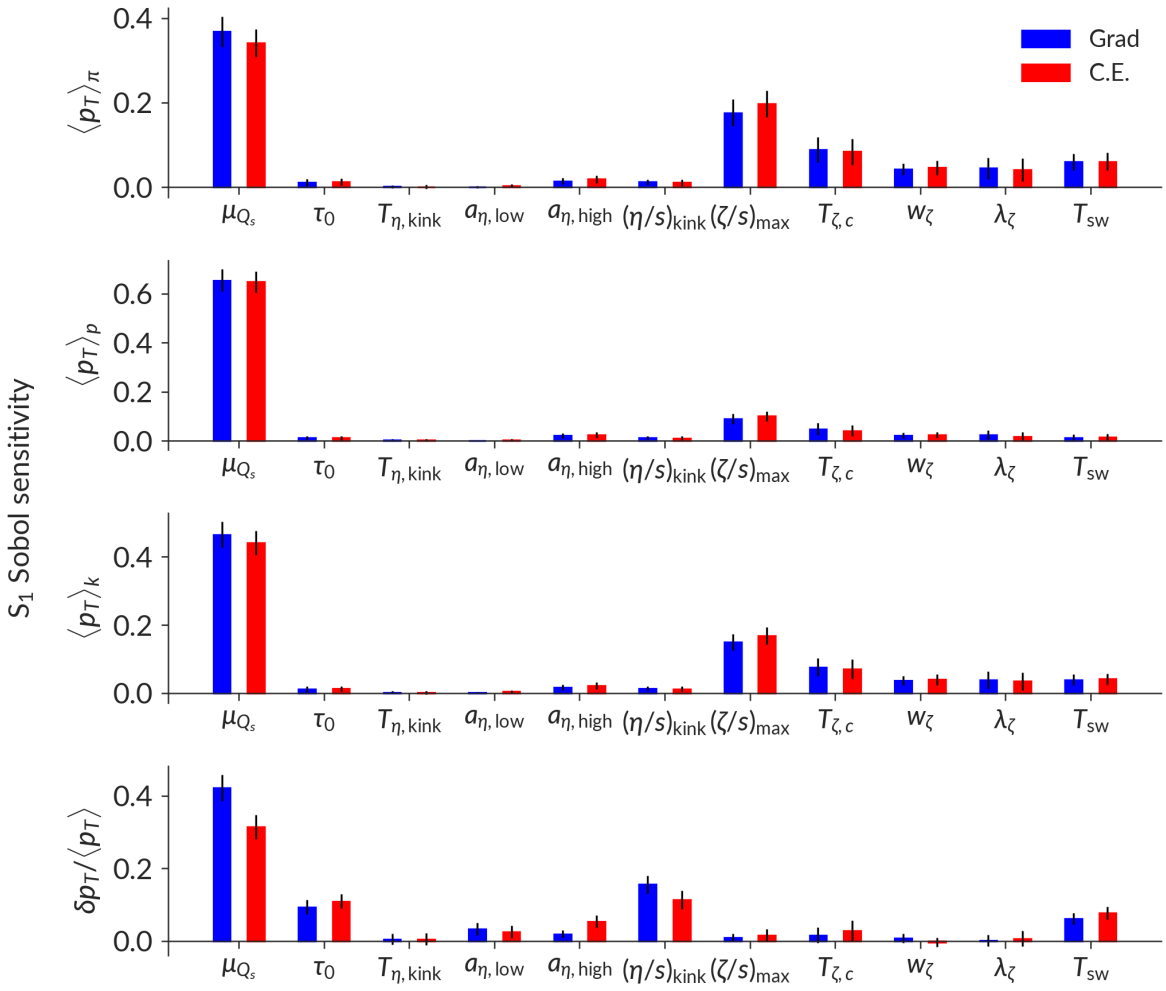


Figure 6.7: First-order Sobol sensitivity of identified particle mean transverse energy and correlated momentum fluctuations to input parameters.

sensitive observables to the chemistry of the system, is the more sensitive than the rest to the bulk viscosity and the switching temperature.

The mean transverse momentum (Fig. 6.7) again confirm the prior expectation that the dominant sensitivity is to the overall normalization and the bulk viscosity. To-date, however, no study using an IP-Glasma initial state has been able to reproduce experimental results for $\delta p_T/\langle p_T \rangle$. This sensitivity analysis reveals that while the dominant sensitivity is to the overall normalization, the fluctuations are primarily sensitive to the shear viscosity, onset of hydrodynamics, and the switching temperature. This study is the first with IP-Glasma to consider simultaneous variation of these parame-

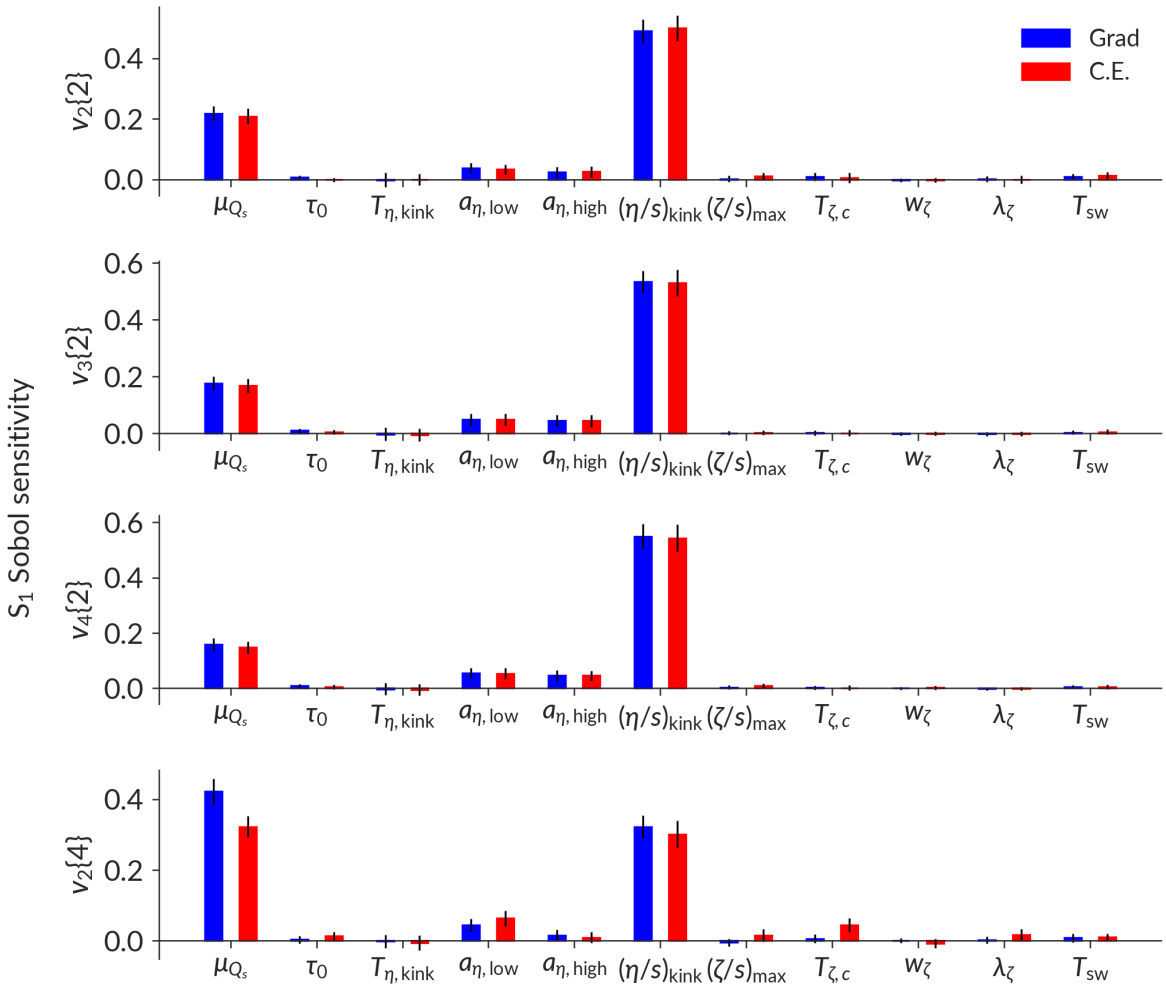


Figure 6.8: First-order Sobol sensitivity of charged hadron anisotropic flow coefficients to input parameters.

ters – in concert with all observables – to learn what can be learned about and from this observable.

The anisotropic flow coefficients (Fig. 6.8) reveal, as expected, that the dominant sensitivity is to shear viscosity and overall normalization. The overall normalization is related to the lifetime of the hydrodynamic phase, in turn allowing for more time for the shear viscosity to act. Interestingly, the difference in sensitivity between the two-particle and four-particle anisotropic flow is isolated to an increased sensitivity to overall normalization and a decreased sensitivity to the high-temperature slow of the shear viscosity. However, the $v_n\{m\}$ are broadly insensitive to the bulk viscosity and particlization temperature, as anticipated.

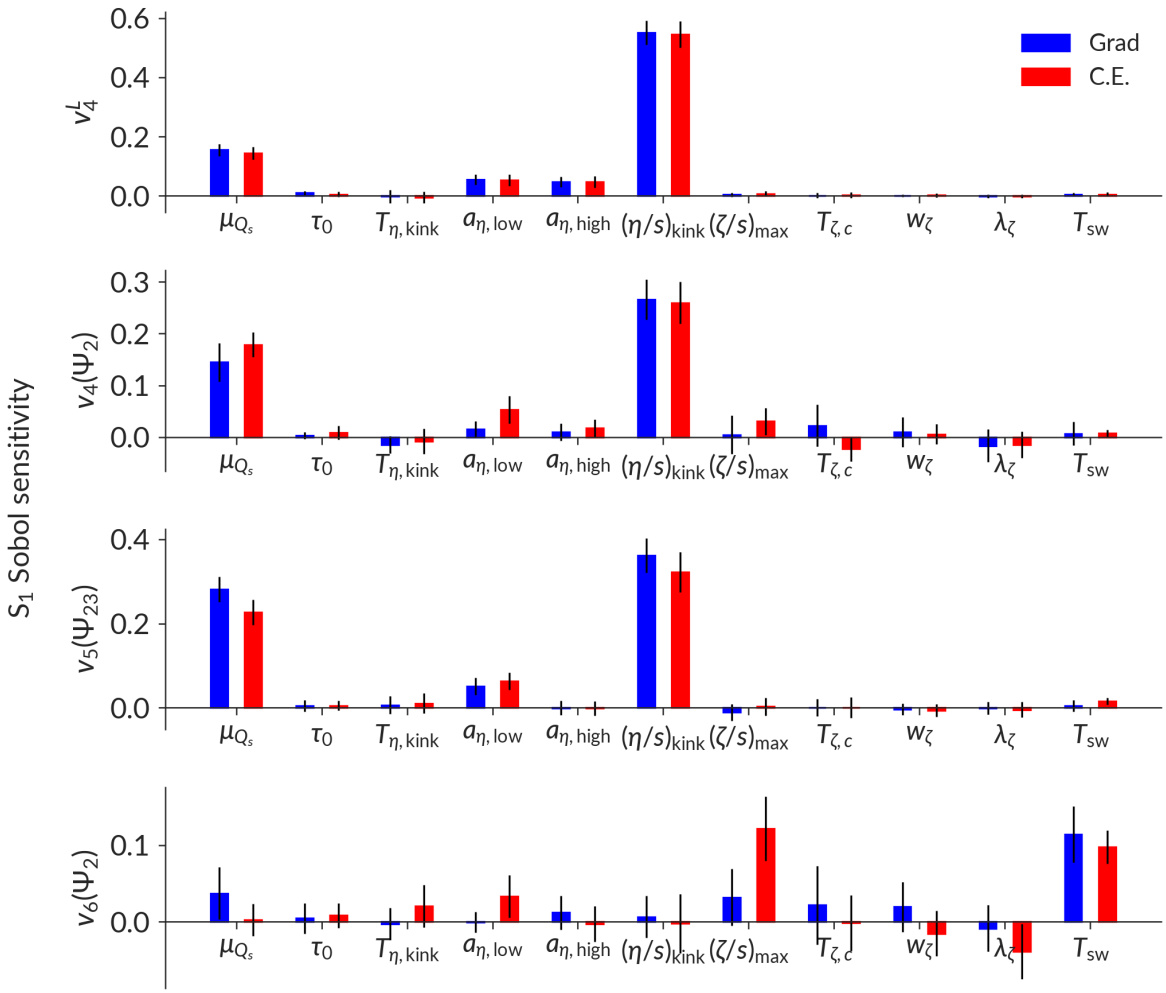


Figure 6.9: First-order Sobol sensitivity of charged hadron linear and nonlinear flow modes to input parameters.

In summary of the first generation observables, the expectations set by both hand-tuning and previous studies is confirmed: the normalization, shear viscosity, and bulk viscosity are broadly sensitive to separate parameter families, but yield constraint across the input parameter space. Disappointingly perhaps, the switching time between the pre-equilibrium and hydrodynamic stage does not appear to be a dominant factor in the variance of any of these observables.

Consideration of the global sensitivity of next generation observables begins with linear and nonlinear flow modes in Fig. 6.9. These observables have less defined and smaller sensitivity to the input parameters, alternately suggesting that they are insensitive to the parameters or that the IP-Glasma initial state by construction con-

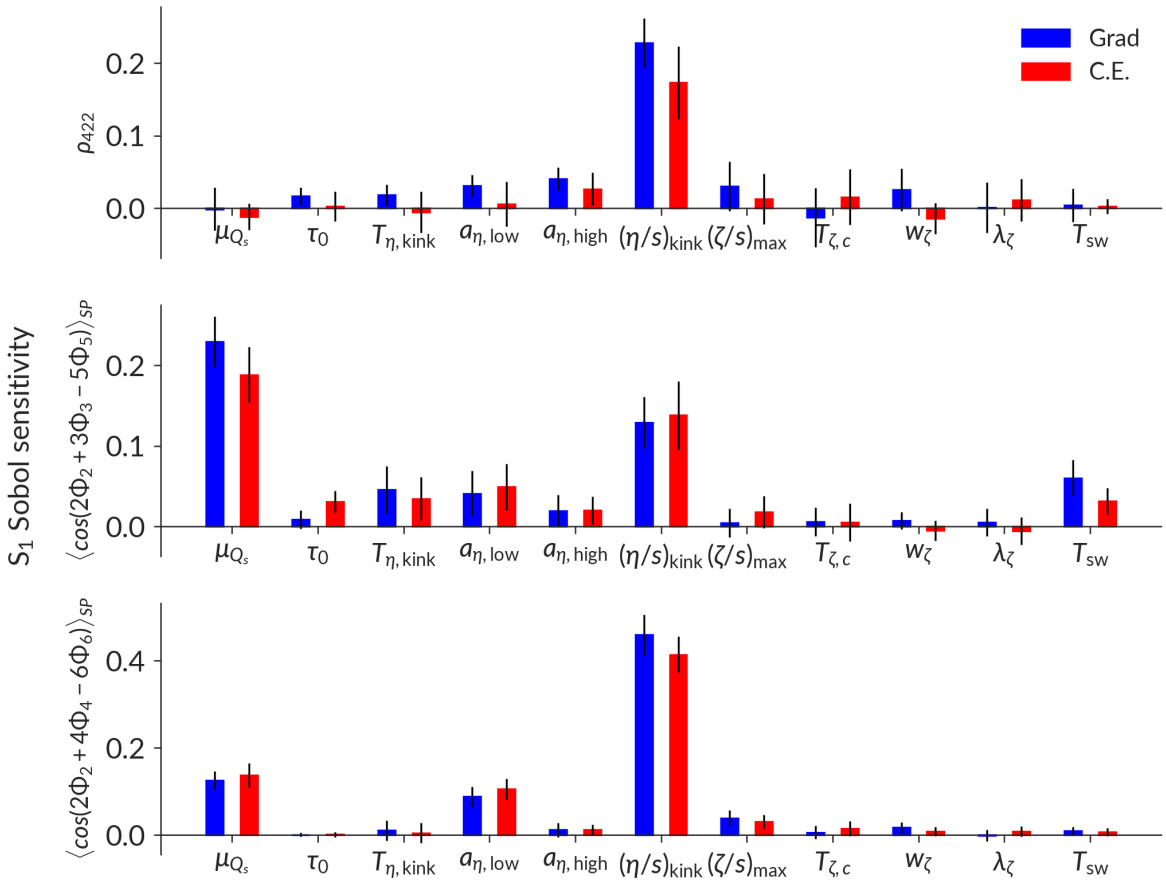


Figure 6.10: First-order Sobol sensitivity of charged hadron event plane correlators to input parameters.

tains the information needed to reproduce these quantities. The exception in this case is to the shear viscosity, whose kink value dominates the constraint of these quantities. Nonetheless, physical expectations suggest that these quantities couple the pre-equilibrium and hydrodynamic stages in a way that similar first generation observables do not – a feature that can be readily seen by their different relations to the principal components in Fig. 6.4.

The correlators (Fig. 6.10) are also less sensitive to variation of the parameters than the first generation observables. Additionally, the large uncertainties on the Sobol indices suggests that the sensitivity is less uniform across the space than in the first generation observables. Nonetheless, these observables further couple the initial state geometry to the hydrodynamic phase in a way poorly quantified by the first generation observables alone.

The sensitivity analysis provides the first glimpse into the response of an IP-Glasma + MUSIC + iS3D + SMASH hybrid model to parameter variation. This result, with a validated surrogate model, yields the first physics insights by calculating the global sensitivity of these observables with the leading physics model of the pre-equilibrium stage and how a realistic hydrodynamic medium responds. The PCA and sensitivity analysis confirm, before any inference has taken place, that the model outputs will yield constraint and that there is information in the next generation observables not contained in the first generation observables that have been the focus of most previous studies. Combined with a physical understanding of the observables themselves, this suggests strongly that exciting opportunities lie ahead for learning through model-to-data comparison.

## 6.4 BAYESIAN PARAMETER ESTIMATION WITH PB-PB COLLISIONS AT 2.76 TEV

As must be becoming familiar, once the observables have been selected and the surrogate model is validated, the next steps are tests of self-consistency to verify the model achieves closure and – at long last – inference with data. These final steps will once again ensure that the model is capable of recovering known inputs before comparing with data to extract a physical result.

### 6.4.1 *Establishing self-consistency*

Once again, the model is tested for self-consistency with pseudodata generated by the underlying hybrid model at known points in the parameter space that were not used in training the surrogate model. The surrogate model is then used for inference with pseudodata and the resulting posterior is investigated to determine how well it recovers underlying truth. Due to the fact that a particular parametrization has been chosen for the specific shear and bulk viscosity, the test for self-consistency is best compared as, for example, $\eta/s$ vs. temperature. After all, despite the motivation for the parametrization, the physics is contained in the temperature dependence of the viscosity, not a particular representation.

It is cumbersome to show this result for all validation points, but care is taken to show a representative sample of validation points in this section. The viscous and non-viscous parts of the posterior are shown separately as the viscous posterior is shown as $\eta/s$ or $\zeta/s$ vs. temperature and no discernible covariance is seen between the viscous and non-viscous parameters. While no covariances are seen, when the model is pushed to the edges of the prior region, the distribution can become bimodal. Examples are shown in Figs. 6.11-6.13.

What is important to inspect is if the posterior consistently contains the known truth. For example, does the true value fall within the 90% credible interval approximately 90% of the time? If so, then it is plausible that, provided with a 90% credible interval, a gambler would break exactly even assuming they were presented with fair odds by the bookmaker. This is clearly the case for the viscous posteriors and the non-viscous posteriors shown in Figs. 6.11 and 6.13. The sample validation points chosen for these figures additionally demonstrate the resolution of a large, relatively flat bulk viscosity (Fig. 6.13) and a bulk viscosity with a comparatively sudden peak at high temperature (Fig. 6.11) in addition to a variety of $\eta/s$. All are recovered well and within the 90% C.I.

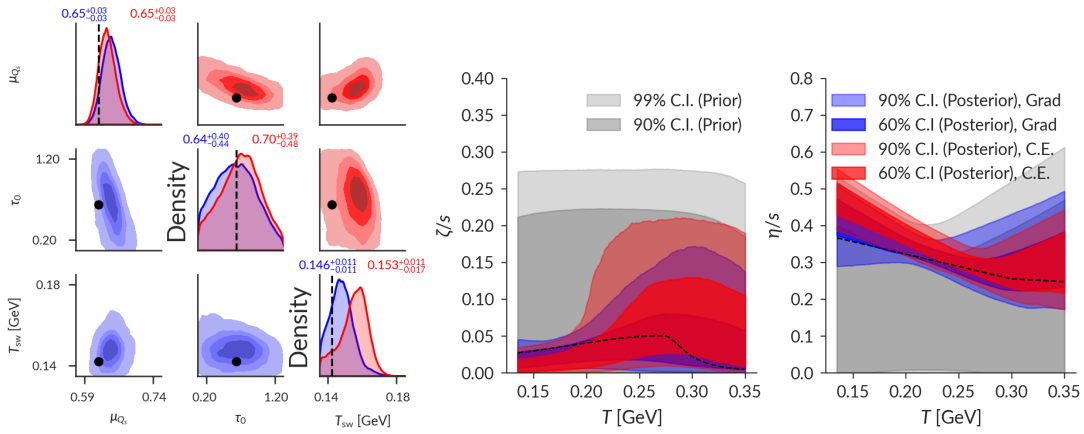


Figure 6.11: Non-viscous (left) and viscous (right) posteriors for a sample validation point. The true values are highlighted in black. The quoted values are the median and 95% C.I.

The posteriors in Fig. 6.12 demonstrate a strong bimodality for Grad viscous corrections and bias in C.-E. $\delta f$ and, while the truths are partially recovered, the posteriors seem at odd with physical intuition and are not in particularly good agreement with each other, such as in $\tau_0$. This occurs because the true value of the bulk viscosity peak

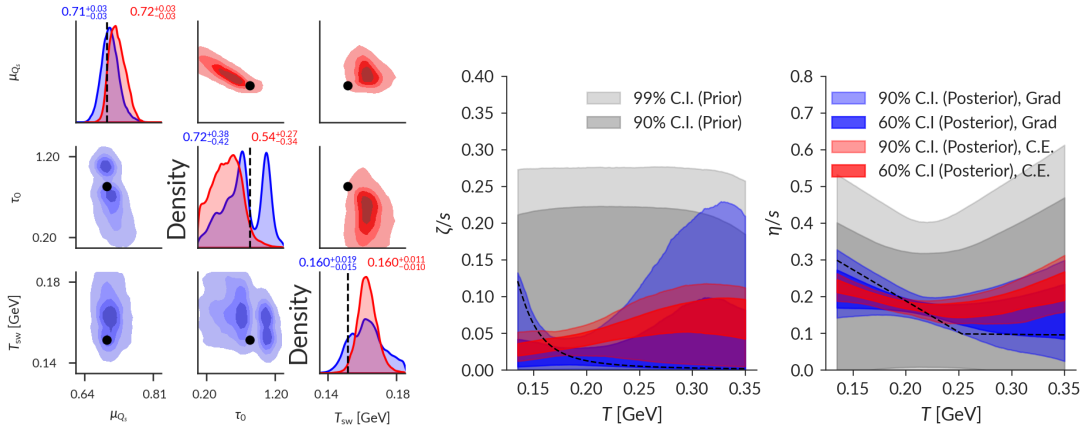


Figure 6.12: Non-viscous (left) and viscous (right) posteriors for a sample validation point. The true values are highlighted in black. This is an important example of interpretable failure. The quoted values are the median and 95% C.I.

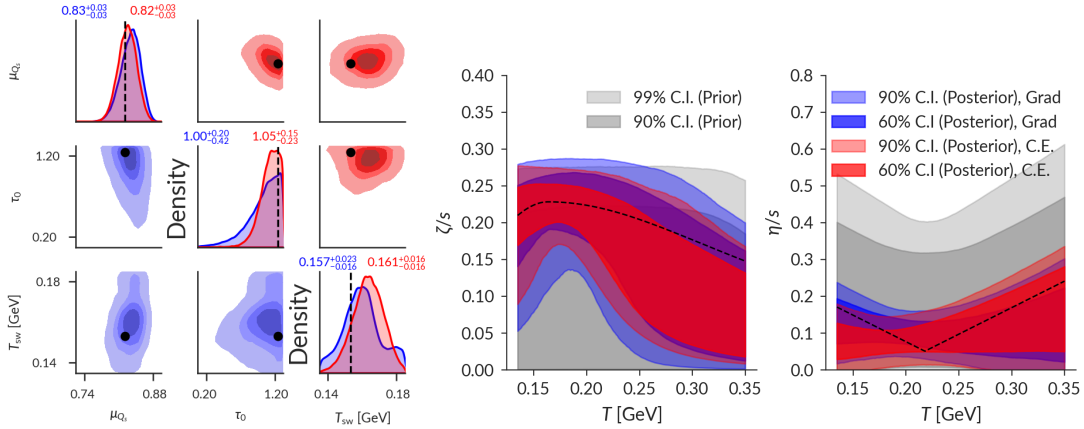


Figure 6.13: Non-viscous (left) and viscous (right) posteriors for a sample validation point. The true values are highlighted in black. The quoted values are the median and 95% C.I.

is *below* the particlization temperature and a bimodality develops in $\zeta/s$ for Grad $\delta f$, while the C.-E. $\delta f$ attempts to compensate and does not resolve the second $\zeta/s$ mode and poorly resolves $\eta/s$. For each peak of $\zeta/s$, a different value of the switching time between IP-Glasma and MUSIC is preferred as the model is pushed into a corner, causing bimodality in the non-viscous posterior. The observable that couples these quantities is $\delta p_T/\langle p_T \rangle$, whose pseudodata is noisier than the experimental data, further exacerbating the issue. This is an example of an interpretable failure is an edge case in the parameter space. It is intuitive that the model struggles to reproduce true values of hydrodynamic quantities that are located outside the hydrodynamic evolution. Joint priors (*i.e.* requiring the bulk peak temperature to be greater than

particlization) have not yet been developed for heavy ion collision studies and doing so is beyond the scope of this work. Note as well that this is a particular feature of the multi-modal bulk viscosity as the true value of $T_{sw}$ in Fig. 6.11 is close to the edge but can still be well-constrained. Nonetheless, the ability to interpret these failures of the modeling workflow further strengthens the results derived from this study.

A reassuring feature of the inferential framework is that all of the closure points reproduce the pseudodata well, as exemplified in Fig. 6.14. As can be seen, the emulator is not overfitting by going through every potentially noisy data point, but is instead robust to statistical fluctuations in the underlying data. This further suggests that the model is behaving well and is well-conditioned for the problem at hand while also not exhibiting strong bias. An example of low-bias can also be seen in the marginal distributions for $\mu_{Q_s}$ in Figs. 6.11-6.13 – the truth is not always exactly located at the peak of the marginal distribution, but instead the peaks are distributed around the true value. Additionally, the two $\delta f$ models are differentiable and reveal that the transfer learning model is not simply reproducing the source model's results.

An exciting feature in these closure tests in comparison to previous studies is the constraint on $\eta/s$ and $\zeta/s$ at higher temperatures. In previous studies, constraint was limited to the low temperature regions and the model was insensitive to the high-temperature (or early-time) behavior of the fireball evolution unless the temperature-dependence was explicitly specified by the parametrization [42, 45, 46, 50, 51]. In these closure test, for the first time, constraint on the viscosity can be achieved even at high temperature. This raises the exciting prospect that the bulk viscosity of strongly-interacting matter in heavy ion collisions may be constrained to an unprecedented precision without sacrificing accuracy.

### 6.4.2 *Parameter estimation from data*

Now that the model is known to behave in accordance with expectations for test points and failures are interpretable, the validation pseudodata is exchanged for real experimental data. The previous section has confidently established that the Bayesian parameter estimation produces reasonable results for known inputs, leading to the belief that this should plausibly reveal the underlying properties of experimentally-

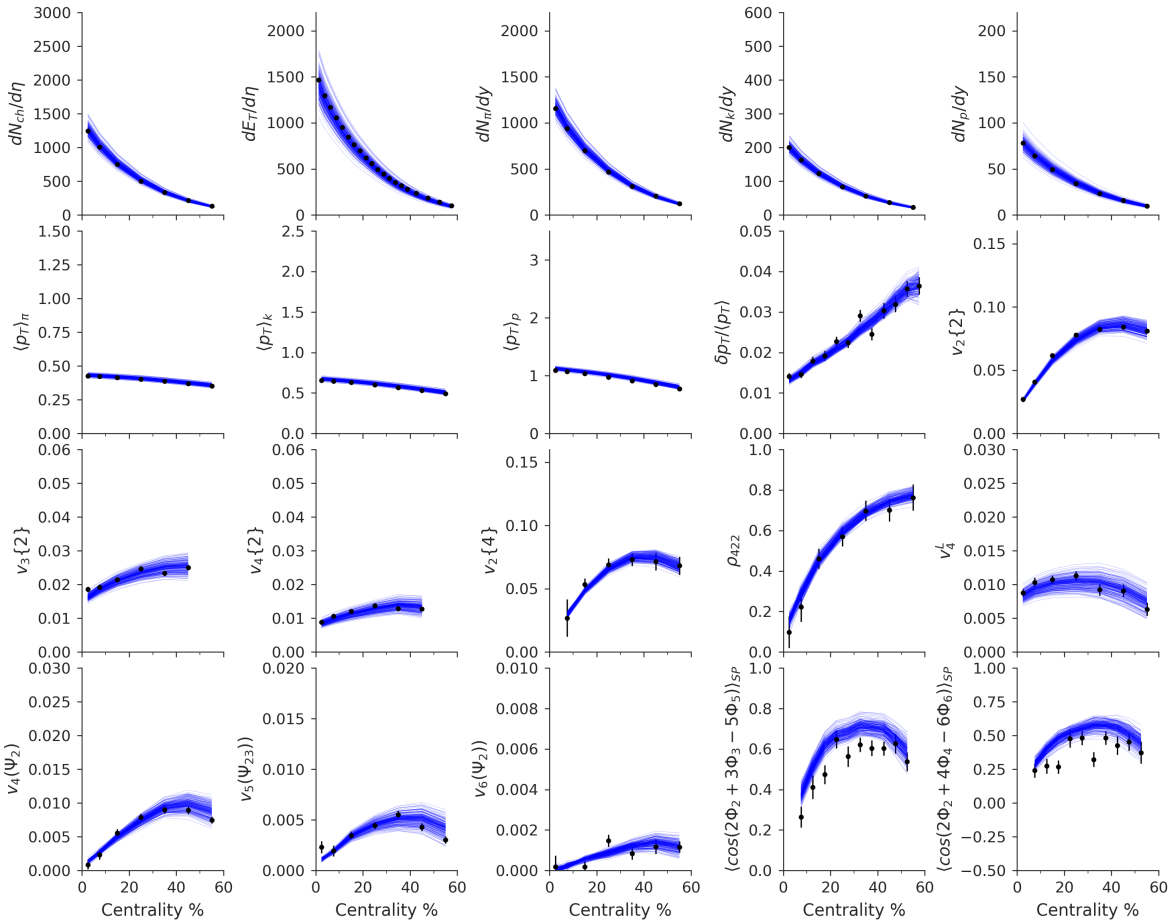


Figure 6.14: Posterior predictive distributions with Grad viscous corrections for the posterior shown in Fig. 6.13 with pseudodata used for comparison shown as data points.

produced quark-gluon plasma in heavy ion collisions. The repeated validation, observable selection, closure testing, and sanity checks of the surrogate modeling and inference have established beyond a reasonable doubt that the models are reliable and well-conditioned for the problem at hand.

The calculations at the design points form the prior predictive distribution and were shown in Fig. 6.2 for a superset of observables. These calculations cover the experimental results well, although correlations between calculations are difficult to discern and likely introduce some tension. The MCMC is again performed using a parallel tempering algorithm. The above closure test and the below comparison to data are performed using Grad's 14-moment viscous corrections and while the above closure tests were performed with 10,000 MCMC steps with 10 walkers per dimension and 10 rungs in the parallel tempering temperature ladder, the below comparison to

data is performed with 20,000 MCMC steps with 50 walkers per dimension and 20 rungs in the parallel tempering ladder for improved sampling resolution. The trace, moving average, and autocorrelation of the final MCMC chain is shown in Fig. 6.15 for three sample walkers. These walkers have clearly thermalized as the trace exhibits no discernible autocorrelation and are thus sampling from the target distribution.

With confidence in the MCMC, it is finally time to look at the posterior distribution after comparison with data. The non-viscous parameter posterior for both viscous corrections is shown in Fig. 6.16, the viscous posterior for both viscous corrections is shown in Fig. 6.17, and the marginal and joint marginal distributions of the 11-dimensional posterior are shown in Fig. 6.18. The non-viscous parameters demonstrate clear constraint, particularly in the case of the normalization $\mu_{Q_s}$. This constraint is consistent with results derived entirely from $J/\Psi$ production in e+p collisions at HERA [52]. The switching time between IP-Glasma and MUSIC is well-localized to early times $\tau_0 \lesssim 0.7$ fm, which is in accordance with physical expectations and appears to rule out very late hydrodynamic onset times seen in models with more simplistic pre-equilibrium dynamics.

The particlization temperature $T_{sw}$ is also well-constrained within the prior region. A recent estimate of the crossover temperature from lattice QCD places it at $T_c = 155 \pm 1.5$ MeV [11], precisely in the region of highest posterior density for the particlization temperature. The constraint of the particlization temperature is particularly interesting as the chemistry of the hydrodynamic medium is identical to that of [45], which required a much lower particlization temperature with the same viscous correction. This also provides a limit on the lifetime over which the viscosity can act by reducing the lifetime of the fireball, placing limits on the viscous contribution. As demonstrated in the closure tests, the viscosity is not required to be small in this region and so has the potential to be influenced by large viscous corrections. However, this is seen to not be the case: in the temperature region probed by particlization – approximately bounded by 0.14 and 0.18 GeV – the data itself prefers the specific bulk viscosity to be small. In testing for self-consistency, it was found that the model can recover large viscosity at particlization (Fig. 6.13), meaning that the demand for small viscous corrections is an authentic feature of the data.

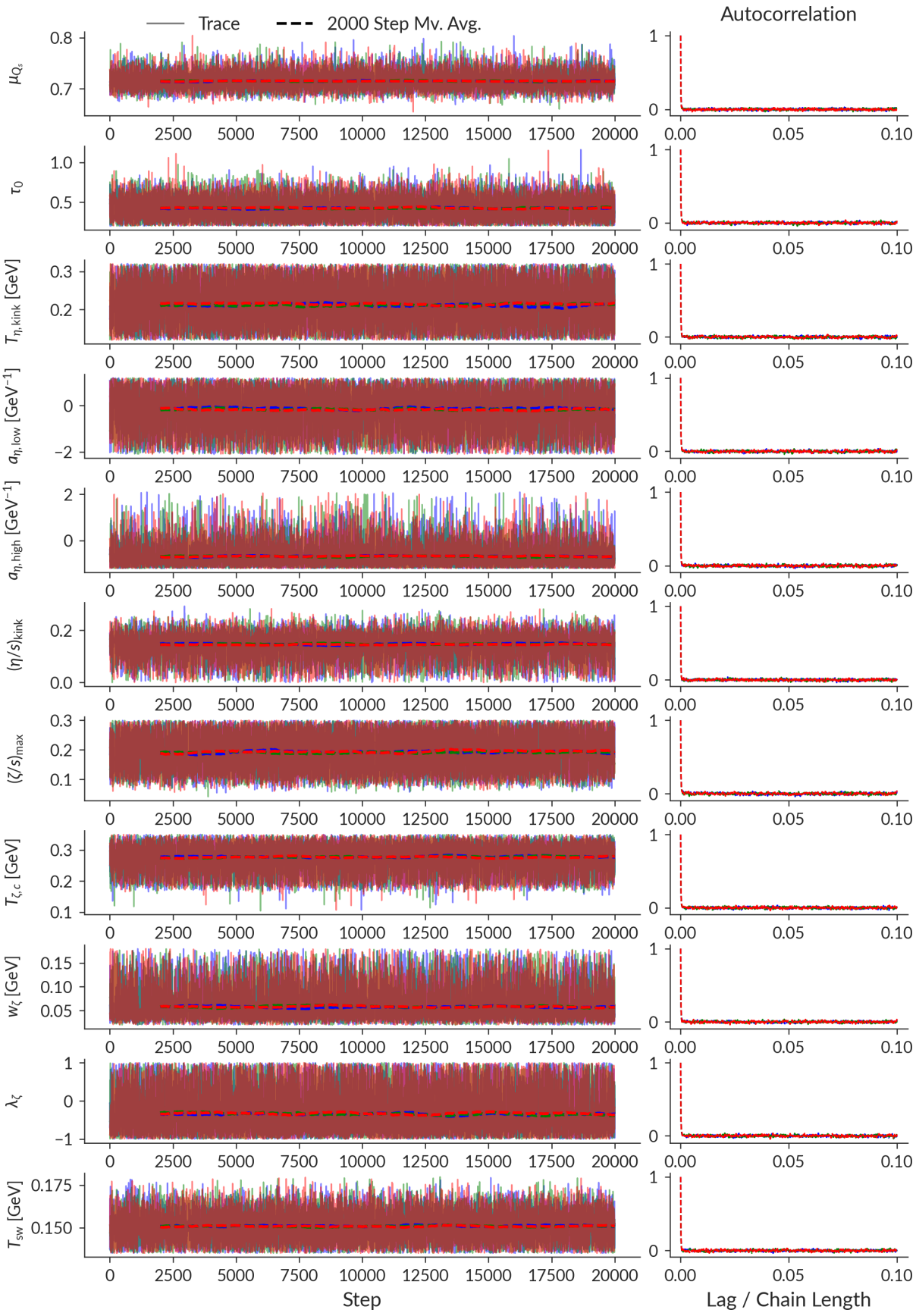


Figure 6.15: MCMC trace, moving average, and autocorrelation from comparison to experimental data with Grad viscous corrections. The C.-E. MCMC behavior is comparable.

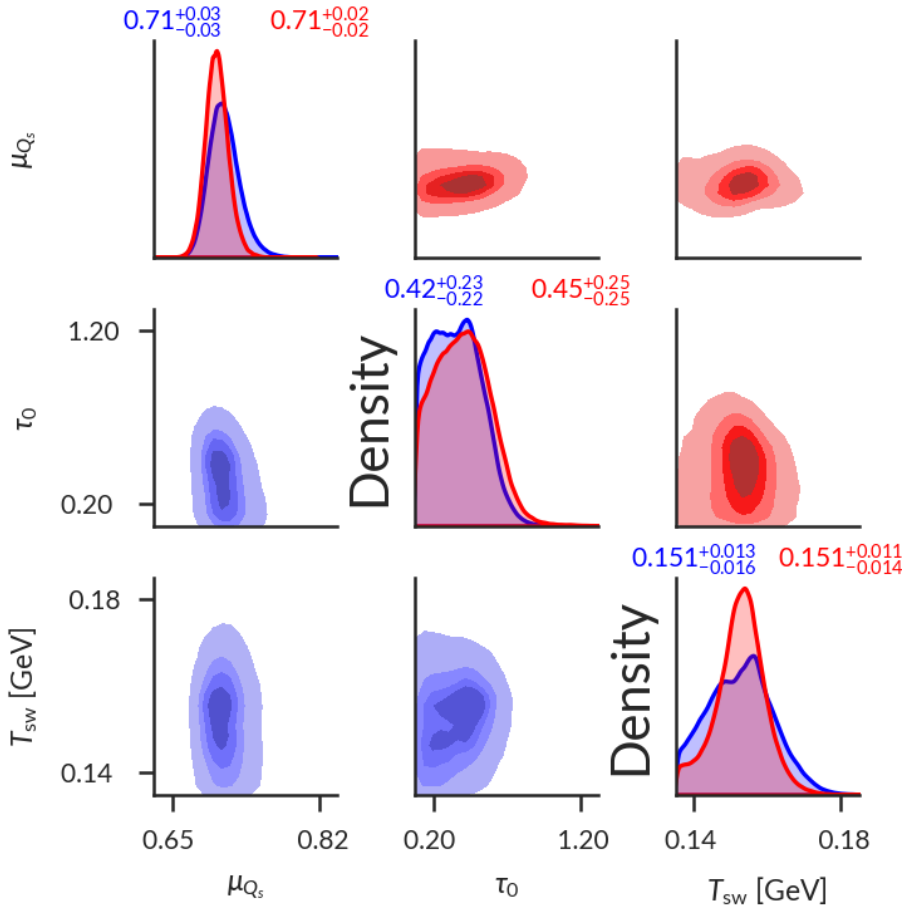


Figure 6.16: Non-viscous posterior from comparison to experimental data with Grad $\delta f$ (blue, lower triangle) and Chapman-Enskog $\delta f$ (red, upper triangle). The quoted values along the diagonal are the median and 95% C.I. of the 1-dimensional marginal distribution.

In Fig. 6.17, the temperature-dependent specific bulk viscosity $\zeta/s$ demonstrates a clear peak and the 99% C.I. is inconsistent with 0 below $T \approx 0.34$ GeV for Grad viscous corrections while for Chapman-Enskog, it is inconsistent with zero over the entire range shown. Randomly drawn example samples from the Grad posterior are shown in Fig. 6.19, demonstrating the diversity of choices that are compatible with data. The constraint certainly weakens at high temperature, but the peaked specific bulk viscosity is well-constrained at low and intermediate temperatures. This is the first time a large, nonzero, peaked specific bulk viscosity has been recovered from data. A peaked result is consistent with expectations from previously-used bulk viscosity motivated by purely physical considerations, demonstrating phenomenological self-consistency between purely theoretical considerations and model-to-data comparison,

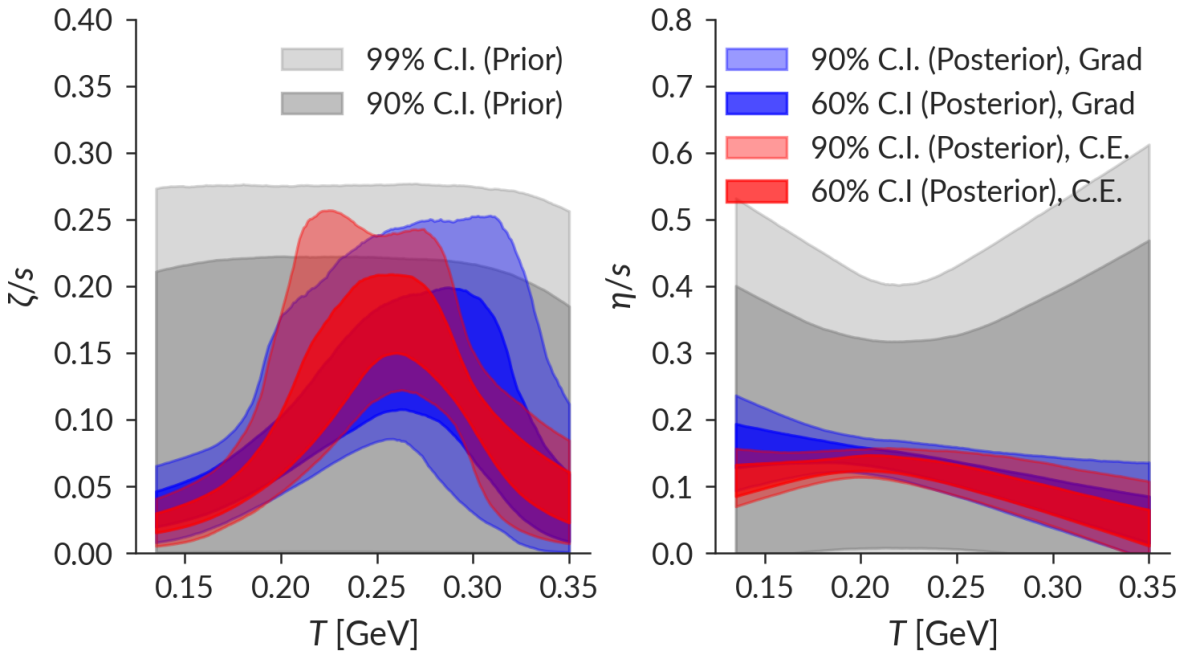


Figure 6.17: Viscous posterior with Grad viscous corrections (blue) and Chapman-Enskog viscous corrections (red) from comparison to experimental data.

although the extracted peak is (mostly) constrained to be at higher temperatures. The peak of the specific bulk viscosity shifts slightly between the two viscous correction models, but the posteriors are broadly consistent with each other, particularly the 60% credible intervals. Both viscous correction models strongly indicate a peaked nonzero bulk viscosity throughout the hydrodynamic evolution.

An unexpected feature of the viscous posterior is a slight preference for a negatively-sloped specific shear viscosity at higher temperatures. This is driven in part by peripheral $v_3\{2\}$, a fluctuation-driven quantity, and central $v_4\{2\}$. As higher temperatures correspond to earlier times in the fireball evolution, this decreasing high-temperature $\eta/s$ dissipates initial-state fluctuations more slowly. Worth noting, however, is that the high-temperature $\eta/s$ posterior is still compatible with both a flat line through the 99% credible interval and the AdS/CFT-derived bound of $1/4\pi$ [198]. This is a consideration worth investigating in more depth. Recent theoretical work has estimated effective viscosity from first-principles calculations; the shear and bulk posteriors in this study are consistent with those derived from first principles theory based on

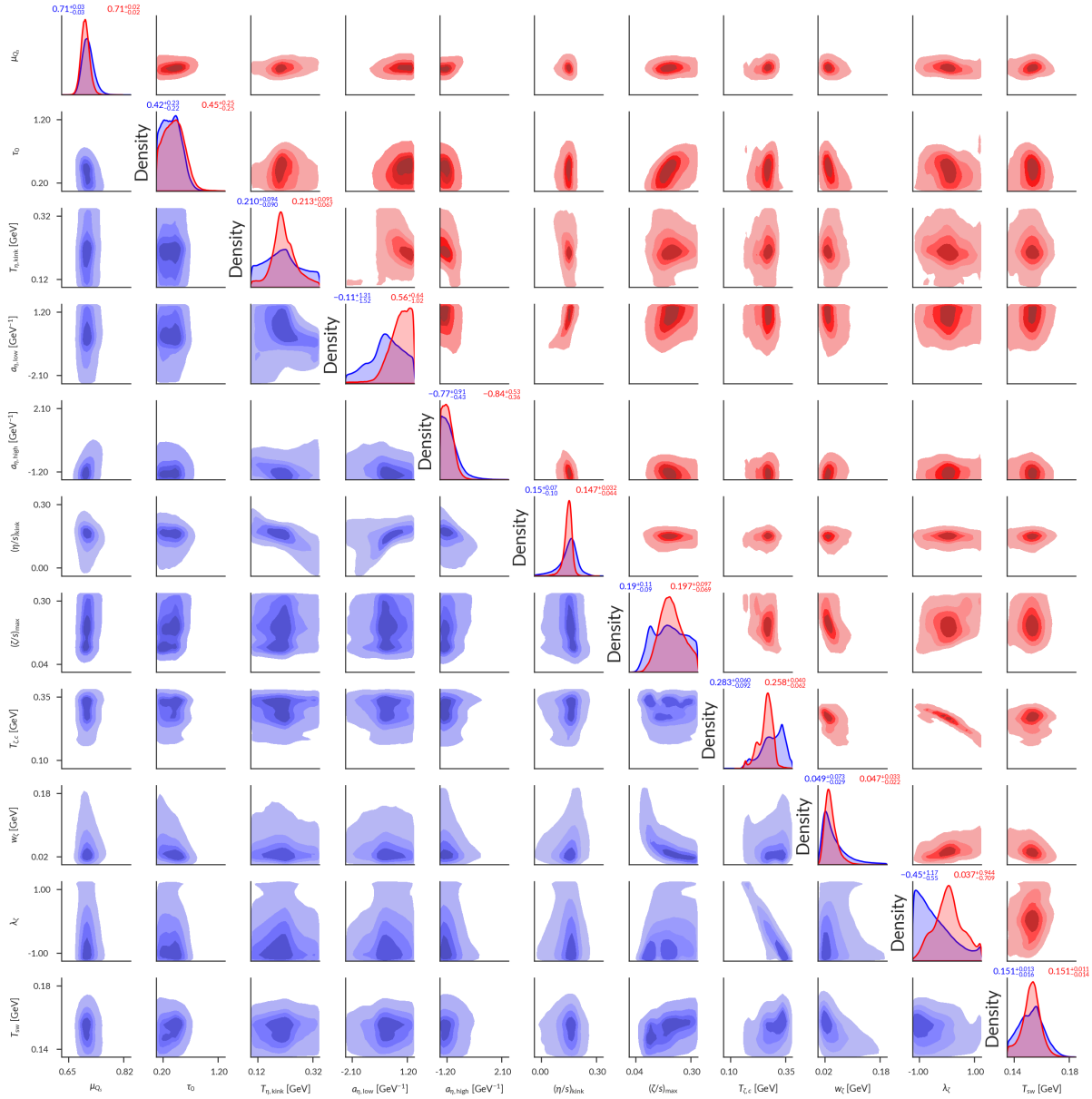


Figure 6.18: 11-dimensional posterior showing marginal and joint marginal distributions with Grad viscous corrections (blue, lower triangle) and Chapman-Enskog viscous corrections (red, upper triangle) from comparison to experimental data. Values along the diagonal are the median and 95% C.I. of the 1-dimensional marginal distribution.

comparisons to other work at a different collision energy [199]. This is also worth investigating in more depth.

To ensure the quality of the fit and to identify tension in the model, one can inspect the posterior predictive distribution (Fig. 6.20) and the ratio of the posterior predictive distribution to experimental data (Fig. 6.21). It is clear from the posterior

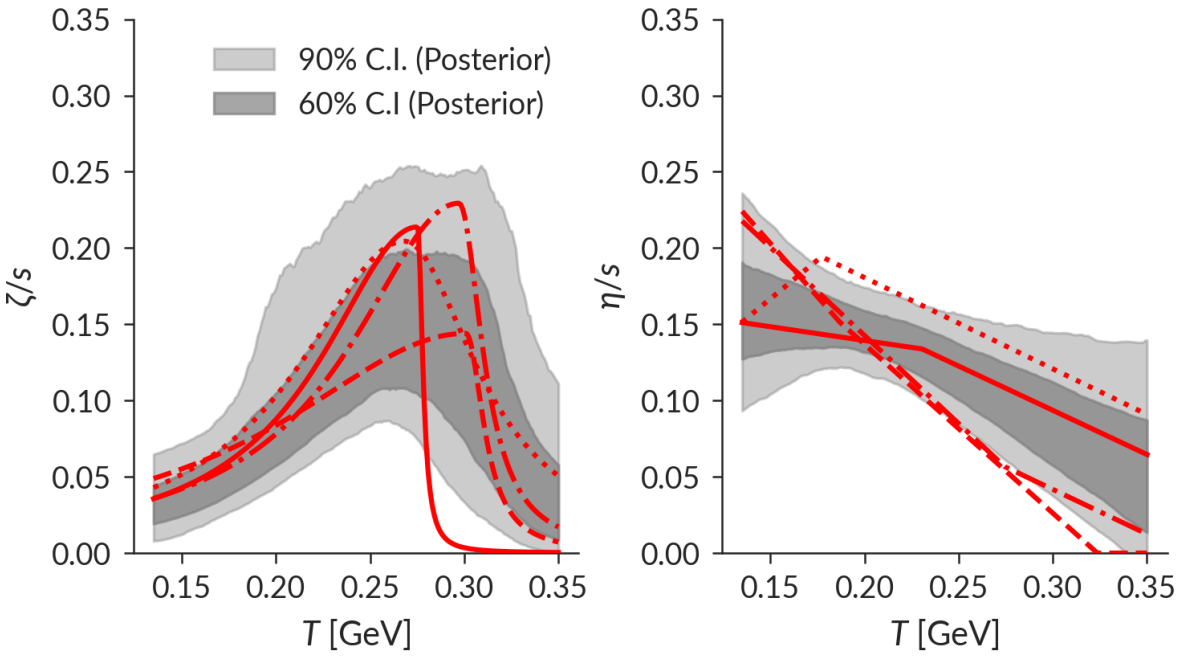


Figure 6.19: Samples from the viscous posterior for Grad viscous corrections after comparison to experimental data.

predictive distributions that the model fits the data well, but exhibits tension, seen in the transverse energy and the three-plane correlators.

The tension between $dE_T/d\eta$ is not new to this work and was also seen in [47]. This suggests that it is a feature independent of the pre-equilibrium stage and potentially a feature of either the specifics of the hadronic chemistry or – less likely – a feature of 2+1D hydrodynamics. As the transverse energy captures correlations between particle multiplicity and transverse momentum, the chemical explanation is more plausible and is consistent with the differences in reproducing experimental results first shown in Fig. 4.8 and present in both models.

The difficulty reproducing the three-plane correlators is also not new, but the postdictions shown in Fig. 6.20 are consistent with the results previously shown in Fig. 4.10. Of note is that the C.-E. $\delta f$ is closer to reproducing these correlations than the Grad viscous corrections. Insight can be gained by investigating the sensitivity of these observables to various parameters in Fig. 6.10. The dominant sensitivities are to normalization and the shear viscosity kink temperature, similar to the anisotropic flow that naturally influence the correlations. The other potential underlying cause

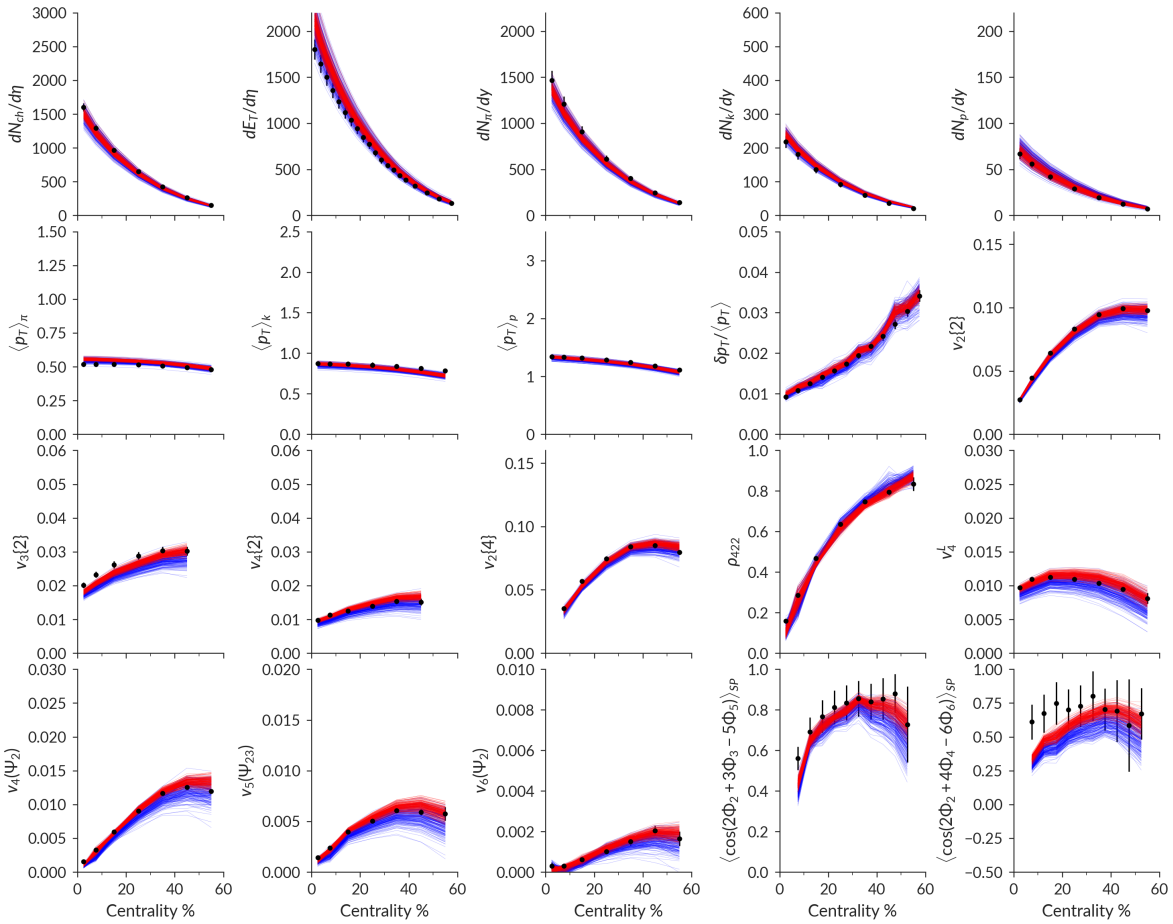


Figure 6.20: Posterior predictive distribution with Grad viscous corrections (blue) and C.-E. viscous corrections (red) after comparison to data.

of difficulty in matching these observables is geometric - the observables match as well as they can, but the prior predictive distributions do not cover the data. With the geometry in IP-Glasma fixed by nuclear configurations and deep inelastic scattering, insufficient freedom remains. Before leaving this to future analysis, it must be noted that $\delta p_T/\langle p_T \rangle$ is also at the edge of the prior predictive region. If the three-plane correlators and the $p_T$ fluctuations are correlated, this has potential to reveal further insight. The correlation between these observables at mid-centrality is shown in Fig. 6.22 and reveals that these observables are uncorrelated, suggesting that their tension is independent. A future analysis should attempt to address this by revisiting the constraint from deep inelastic scattering simultaneously with observables from heavy ion collisions.

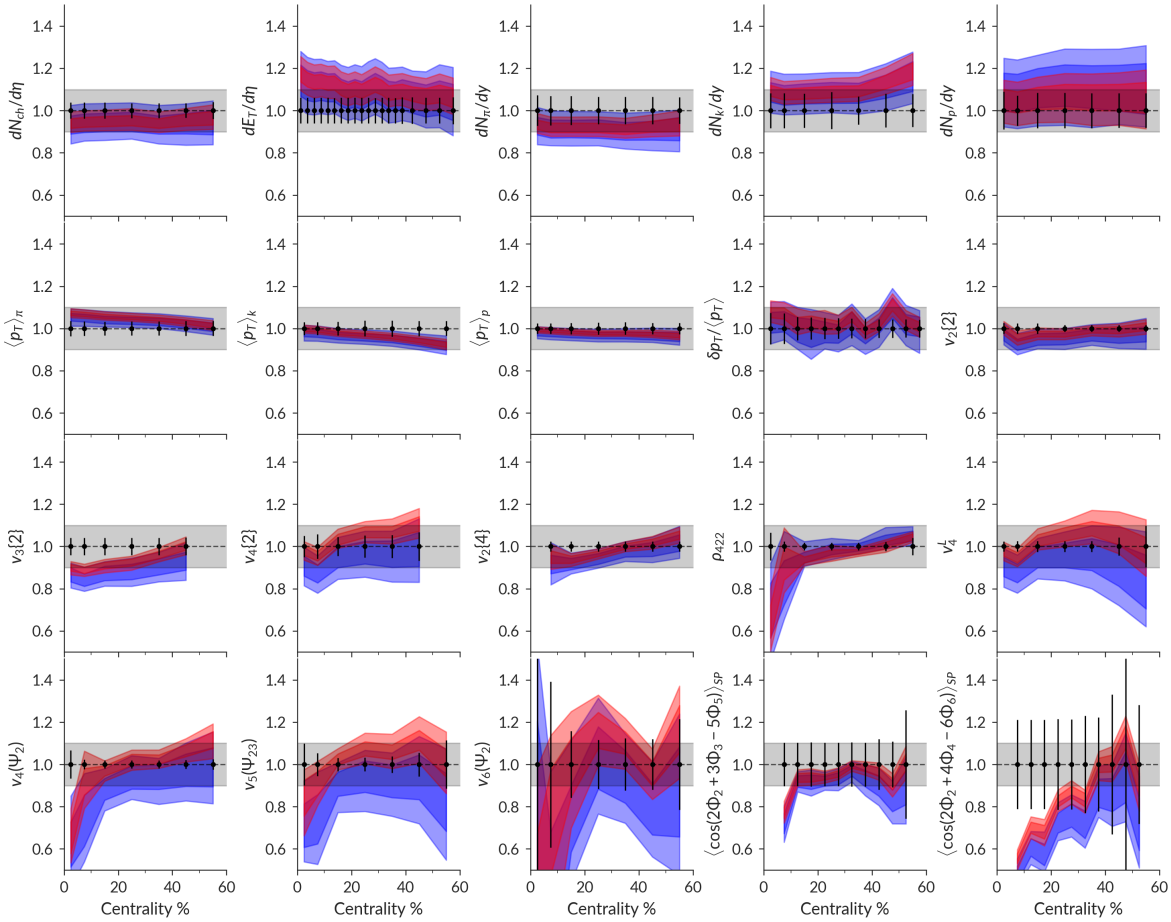


Figure 6.21: Posterior predictive ratio with Grad (blue) and C.-E. (red) $\delta f$ after comparison to data.

The posterior predictive distribution for the correlated $p_T$ fluctuations produces the most accurate postdiction of any IP-Glasma calculation and yields the correct centrality dependence, a feature not seen in other models. Investigating this sensitivity, the overall magnitude is reduced by a larger $(\zeta/s)_{max}$ and constraints $\tau_0$ to early times. This suggests yet further that the bulk viscosity must be further investigated for a narrower, taller peak to better reproduce experimental results. This is, unfortunately, beyond the scope of this work.[1]

The success of the model with respect to every other observable must be highlighted: nearly every experimental measurement in nearly every observable is consistent with the posterior predictive distribution shown in Fig. 6.20. This is nothing short

1 This narrower, taller peak is difficult to resolve without reparametrization of the width of the bulk viscosity or carefully constructing a scale-invariant prior. This too is beyond the scope of this work, but should be strongly considered in future studies.

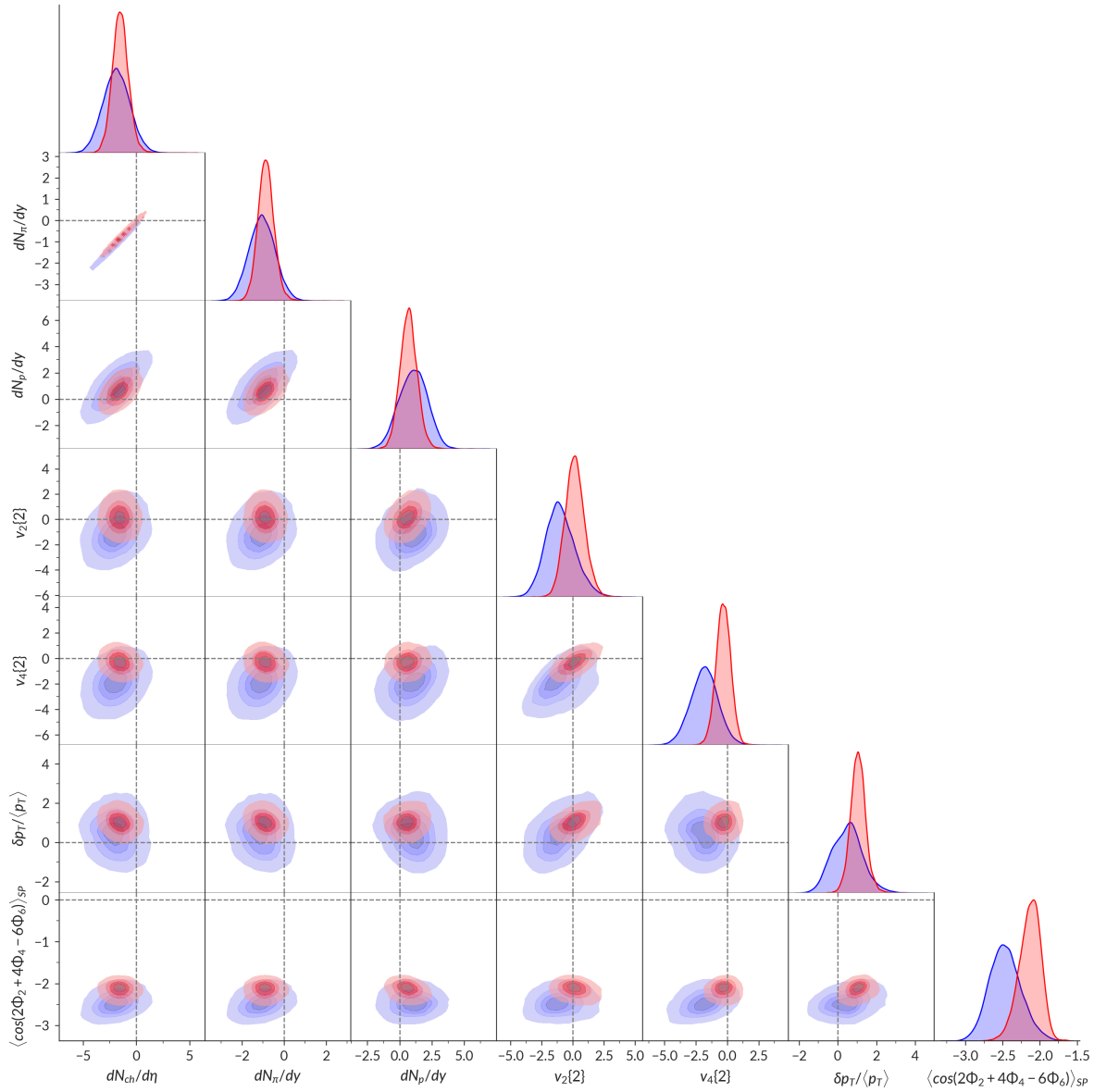


Figure 6.22: Correlations between posterior predictive distributions for selected observables for central collisions. Dashed lines denote the central experimental result and x- and y-axis units are the experimental uncertainty for the respective observables. Grad viscous corrections are in blue while Chapman-Enskog viscous corrections are shown in red.

of remarkable and is by no means guaranteed. Bayesian studies in heavy ion physics have broadly exhibited success, but previous studies have used flexible parametric models and fewer observables. To have a pre-equilibrium stage with microscopic physics produce such a resounding success is a thorough and non-trivial validation of the theory and implementation of IP-Glasma. This represents a milestone in rig-

orously constructing a hybrid model with each stage containing the best-available microscopic physics and testing it via comparison to data. Certainly, future work remains in theoretical development, but this establishes the physical models on a firm scientific foundation.

### 6.4.3 *Bayesian Model Comparison*

Bayesian model comparison, introduced in Ch. 3.3, can be used to determine if the data exhibits a preference for one model or another, if additional complexity is justified by the model, or even if the model can differentiate between pseudodata and experimental data. This is extremely valuable as it does not attempt to falsify a model, but rather puts it to a binary test to determine which model is the most useful in describing the data.

To test, as always, with self-consistency, the first use of Bayesian Model Comparison is to determine if the model can differentiate between pseudodata used for the previous self-consistency testing and experimental data. This hypothesizes the following scenario: a "true" model underlies the experimental data just as a known model underlies the pseudodata generated to test self-consistency. A distinct model is never expected to systematically defeat the true underlying model and, if it did, would be a sign of systematic bias. As a result, the Bayes evidence for the pseudodata is expected to be greater than the Bayes evidence for the experimental data and strong preference is expected from the Bayes factor. This is found when comparing the model estimate of the Bayes evidence for pseudodata and true data using Grad viscous corrections: the ln Bayes Factor determining which data the model is best suited to ranges from $118.7 \pm 3.1$ to $147.9 \pm 2.3$ in favor of the pseudodata, corresponding to odds of around $2 \times 10^{56} : 1$ to $10^{62} : 1$ differentiating the two data sources. Comparable, albeit slightly reduced preference is found using Chapman-Enskog viscous corrections ($\ln B \sim 90 \pm 5$). This is an overwhelming validation of the model's ability to differentiate the data and demonstrates the self-consistency of the Bayesian model selection. It remains a humbling revelation of just how much information is not yet captured even by this state-of-the-art model.

The self-consistent Bayesian model comparison can now be used to determine if the model exhibits a preference for a variety of features. For example, it can be used to test if the model demands a temperature-dependent shear viscosity by fixing the high and low temperature slopes to 0 and fixing the kink temperature to any value in the prior range as it is meaningless with no change in slope. Performing this comparison yields $\ln B = 0.2 \pm 2.4$ in favor of temperature-dependent shear viscosity with Grad viscous corrections and $\ln B = 1 \pm 4$ for Chapman-Enskog viscous corrections. On the Jeffreys' scale (Table 3.1), this is consistent with no preference between the two models, suggesting that the evidence is still inconclusive *in favor or against* temperature-dependent $\eta/s$ given the data considered in this study. This suggests that further, and higher-resolution, studies are required to conclusively demonstrate the temperature dependence (or lack thereof) of $\eta/s$ in heavy ion collisions.

The lack of such preference for or against $\eta/s(T)$ is not surprising. Hybrid models with IP-Glasma have demonstrated considerable success in describing experimental results using a constant specific shear viscosity and the viscous posteriors in this study are themselves consistent with a constant value. In the study requiring a constant $\eta/s$, the result is well-constrained – $\eta/s = 0.137^{+0.025}_{-0.028}$ for Grad viscous corrections and $\eta/s = 0.125^{+0.021}_{-0.022}$ for Chapman-Enskog viscous corrections, where the uncertainty denotes the 95% C.I. – and with minimal covariance. By inspection, it is apparent that this is entirely consistent with the $\eta/s(T)$ posteriors in Fig. 6.16 and nearly spans the full width at the narrowest point.

As many Bayesian works require $\eta/s(T)$ to strictly increase or be constant above a fixed kink temperature, this is also a useful comparison and is performed with only Grad viscous corrections as both models are consistently in agreement. To do this, $a_{\eta,low}$ is fixed to zero as it is in those studies and $T_{\eta,kink}$ is fixed to 0.154 GeV. Finally, $a_{\eta,high}$'s prior range is reduced to require it to be positive definite. Comparing the evidence for this configuration to the full study produces $\ln B = 3.8 \pm 2.6$ in favor of the full study allowing for a negatively-sloped $\eta/s(T)$. This corresponds to moderate-to-strong evidence on the Jeffreys' Scale in Table 3.1. Comparing the requirement of a positive-definite slope to $\eta/s(T)$ to a constant $\eta/s$, the Bayes factor is $\ln B = 3.6 \pm 2.6$ in favor of the constant specific shear viscosity. Because the Bayes factor penalizes complexity, the additional complexity is not justified by the data.

Next generation observables are employed in this study in the hope of determining the features of $\eta/s$ and $\zeta/s$ with greater accuracy and precision. Some studies use next generation correlations that require much greater computational expenditure to attempt to find this constraint, but suffer from parametric initial conditions [53]. It is clear from these Bayesian model comparisons that success in learning the physical specific viscosity of strongly-interacting matter will only come from combining realistic initial conditions and well-chosen observables. A promising candidate for increased constraint are $v_n - p_T$ correlations, which are not readily calculable at the precision of this study, but further couple pre-equilibrium geometry to the hydrodynamic evolution [200, 201]. This is investigated later in this work as a prediction made at Maximum a Posteriori.

Recent Bayesian works with a T$_R$ENTo + freestreaming initial state have been finding success with smaller and smaller specific bulk viscosity [50, 51, 53], contrasting with prior studies demonstrating the need for $\zeta/s$ to reproduce hadronic observables. By fixing $(\zeta/s)_{max}$ to zero and holding the other parameters fixed to arbitrary values as they no longer have any impact, it is straightforward to assess the demand for nonzero $\zeta/s$. This comparison results in $\ln B = 34.4 \pm 2.4$ in favor of non-zero $\zeta/s$ when using Grad viscous corrections, corresponding to odds of $\sim 8 \times 10^{14} : 1$. With Chapman-Enskog viscous corrections, this preference for the inclusion of bulk viscosity increases to $\ln B = 61 \pm 5$, conclusively demonstrating that bulk viscosity is strongly justified when using a physically-motivated pre-equilibrium stage no matter the viscous corrections at particlization. The physical impacts of the lack of bulk viscosity arise in enhancement of the identified particle $\langle p_T \rangle$ and the momentum fluctuations $\delta p_T/\langle p_T \rangle$ with simultaneous suppression of $v_3\{2\}$ and the three-plane correlators. The particlization temperature is also forced to the highest possible temperature allowed in the prior while the switching time to MUSIC is required to be as short as possible. This arises from a need to preserve as many initial-state fluctuations as possible as they must reproduce fluctuation-driven final-state observables. The high particlization temperature additionally preserves fluctuations by allowing for less viscous dissipation in the hydrodynamic phase.

Comparing the relative likelihood of the viscous correction models is a useful way to assess model applicability and begin to quantify the uncertainty introduced by

the choice of viscous correction. Comparing Grad and Chapman-Enskog viscous corrections to data with none of the parameters held fixed, the relative preference for the Grad over the Chapman-Enskog RTA viscous corrections is $\ln B = -0.1 \pm 3.1$ in imperceptibly-slight favor of Grad viscous correction, although this should be interpreted as the models being indistinguishable in this analysis.

This indistinguishable nature of the viscous correction models deserves further study. The posteriors, as shown previously, are quite similar but not identical, but are equally well-suited to experimental measurements. As a result, the viscous corrections chosen in a study are an important source of theoretical uncertainty to quantify and not doing so results in an artificially precise posterior. Progress in adding additional constraining observables must not neglect quantification of uncertainty as a parallel goal lest analyses fall into the trap of the bias-variance tradeoff. The goal is not to constrain these quantities the most precisely, but to do so both accurately *and* precisely. By not including sources of theoretical uncertainties, an analysis focuses on the latter and sacrifices the former.

A natural question to ask is why is the model preference between the Grad and Chapman-Enskog viscous corrections indeterminate in this study where, in the only other application of Bayesian model comparison, it was strongly in favor of the Grad $\delta f$ [45, 46]? The answer is essentially two-fold: first, the viscosity in the previous study were larger at particlization and as a result, enhanced the effect of the corrections; and second, the inclusion of more realistic dynamics in the pre-equilibrium stage means that observables are less sensitive to the hydrodynamic viscosity. As the viscosity is not wholly responsible for introducing momentum-state anisotropy, for example, the viscous corrections at particlization impact the observables less, in turn resulting in less model preference between modeling choices which should be small effects by construction. The lack of preference between the two models of viscous corrections is in accordance with prior theoretical expectations.

The most likely value in the 11-dimensional parameter space is the Maximum a Posteriori estimate , determined by numerical optimization on the MCMC chain. As the Bayesian model comparison exhibits no preference for or against temperature-dependent specific shear viscosity, estimates of the MAP are provided for both temperature dependence and a lack thereof in Table 6.2.

| Parameter | Grad $\delta f$, $\eta/s$ | Grad $\delta f$, $\eta/s(T)$ | C.-E. $\delta f$, $\eta/s$ | C.-E. $\delta f$, $\eta/s(T)$ |
|---|---|---|---|---|
| $\mu_{Q_s}$ | 0.72341 | 0.70808 | 0.72654 | 0.70858 |
| $\tau_0$ [fm] | 0.52127 | 0.51291 | 0.40142 | 0.55159 |
| $T_{\eta,kink}$ [GeV] | 0.150 | 0.22333 | 0.150 | 0.21123 |
| $a_{\eta,low}$ [GeV$^{-1}$] | 0.000 | -0.16259 | 0.000 | 0.65272 |
| $a_{\eta,high}$ [GeV$^{-1}$] | 0.000 | -0.80217 | 0.000 | -0.89472 |
| $(\eta/s)_{kink}$ | 0.13577 | 0.13944 | 0.12504 | 0.14888 |
| $(\zeta/s)_{max}$ | 0.28158 | 0.22085 | 0.17391 | 0.20117 |
| $T_{\zeta,c}$ [GeV] | 0.31111 | 0.29198 | 0.2706 | 0.25455 |
| $w_\zeta$ [GeV] | 0.02878 | 0.03625 | 0.05255 | 0.04506 |
| $\lambda_\zeta$ | -0.96971 | -0.56235 | -0.14178 | 0.06408 |
| $T_{sw}$ [GeV] | 0.15552 | 0.15429 | 0.15069 | 0.1513 |

Table 6.2: Maximum a Posteriori estimates with Grad's 14-moment and Chapman-Enskog RTA viscous corrections. Estimates with (denoted $\eta/s(T)$) and without (denoted $\eta/s$) temperature-dependent specific shear viscosity are reported.

A variety of interesting features arise in Table 6.2. First, the lattice QCD estimate of crossover temperature – $T_c = 155 \pm 1.5$ MeV – is consistent with both Grad MAP estimates of the particlization temperature using Grad viscous corrections, with the MAP estimate from constant $\eta/s$ nearly identical to the central lattice estimate. Using Chapman-Enskog viscous corrections results in a slightly lower estimate of the particlization temperature, but still close to the estimated crossover temperature, suggesting that the hadrons may behave hydrodynamically for a brief period after recombination. Next, the switching time $\tau_0$ is consistent with IP-Glasma's pressures having come to a steady state (see Fig. 4.21) and with sufficient time for the build-up of pre-equilibrium dynamics that was hypothesized to be of critical importance in describing the strongly-interacting medium (see Ch. 4). The value of the specific shear viscosity is broadly consistent with other Bayesian results and the constant $\eta/s$ is very close to past chi-by-eye fits of 0.13. The bulk viscosity maximum and width are consistent with a large, peaked bulk viscosity, further supporting a consistent pic-

ture between theoretical expectations and prior modeling success. The asymmetry of the bulk viscosity is of interest as it suggests a bulk viscosity peaked at high temperature and slowly decreasing as it approaches the particlization temperature, where it is well-constrained by the data to be small. While the MAP estimates for the bulk viscosity differ in their parameters between $\eta/s$ and $\eta/s(T)$, the actual value at any temperature differs by a maximum of ~10% below the region where it nears the lower peak location at $T \approx 0.28$ GeV.

The MAP estimates are used to make predictions of observables not used in the model-to-data comparison. Strictly speaking, this is due to computational limitation: the most appropriate comparison is a full posterior predictive distribution with perhaps a surrogate model trained on a reasonable quantity of high-statistics calculations. At the same time, the MAP estimates are the recommended parameters for use in other studies, such as hard sector studies of jet-medium interactions or photon/dilepton calculations, and therefore represent a faithful picture of how the model will be used in practice.

### 6.4.4 *Bayesian Model Averaging*

In Bayesian model comparison, the question under investigation is "which model is best suited to the data?" This informs which model to use and how best to use it. A related question is "given two models, how does one best estimate the truth?" For this, Bayesian model averaging (BMA) is employed. In Bayesian model averaging, two posteriors are combined using a weighted average in which the weights are the Bayes evidence [202]. In a simplified example, if two models are equally likely, then the truth is most likely to be in the region where the model posteriors overlap. This is formalized as

$$p_{BMA}(x|y) \propto \sum_i p_i(y) p_i(x|y) \tag{6.5}$$

for models indexed $i$.

BMA was first used in heavy ion collisions to perform model averaging of the transport coefficients and later for model averaging of non-viscous parameters [46, 47]. The BMA viscous posteriors are shown in Fig. 6.23 along with the Kullback-

Leibler divergence, which quantifies the distance between two distributions and is used here to calculate the information gained from the prior to the BMA posterior [203]. The BMA posterior for non-viscous parameters is shown in Fig. 6.24.

The BMA viscous posterior clearly demonstrates the value of accounting for the uncertainty due to viscous corrections at particlization by showing the state of knowledge by considering both simultaneously. The two models contribute their constraint throughout the temperature evolution of both $\zeta/s$ and $\eta/s$, although the impact is clearer in the specific bulk viscosity due to the differences in constraint between the two models. Particularly of interest is that BMA leverages the information content of both models to produce a more-constrained 60% C.I. than either model independently, demonstrating how to address the bias-variance tradeoff with multiple models in a rigorous way.

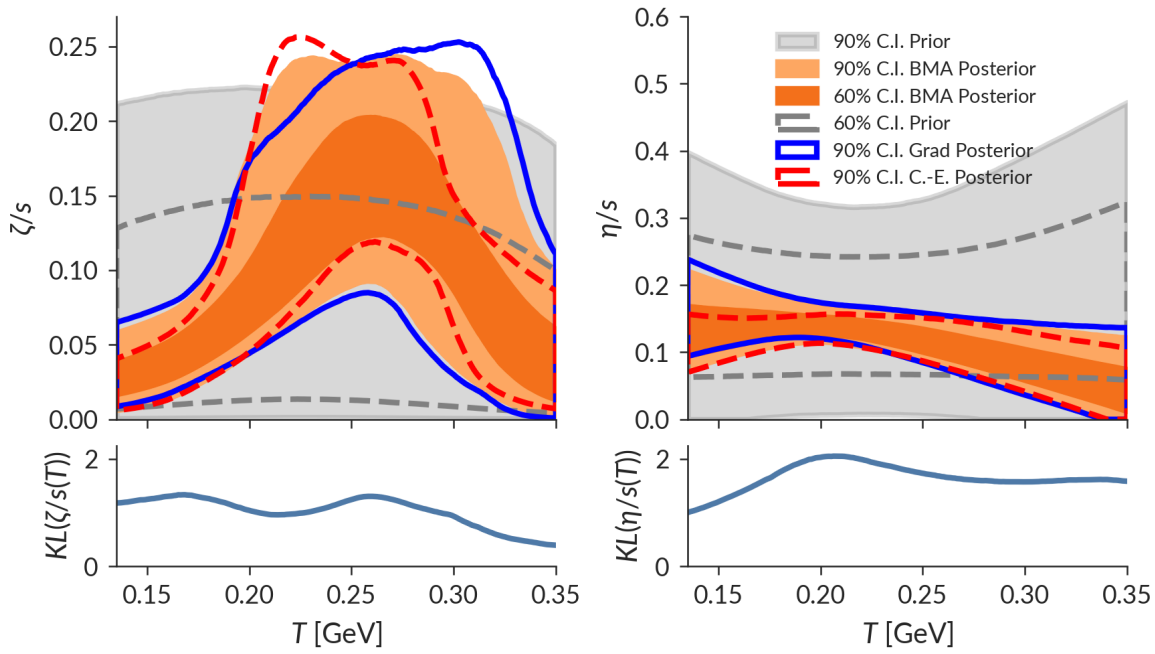


Figure 6.23: Bayes Model Averaged viscous posterior shown with with Grad 90% C.I. (blue) and Chapman-Enskog 90% C.I. (red) and the Kullback-Leibler Divergence quantifying information gain from the priors to the BMA posterior in bits (bottom panels).

The KL Divergence in Fig. 6.23 is also of note: information on the viscosity is gained by comparing to data over the entire temperature region considered, decreasing at higher temperatures that are probed more briefly and earlier in the collision evolu-

tion. This is consistent with the only other study to investigate this, but has substantially increased the amount of learning from prior to posterior. As the hydrodynamic, partilcization, and hadronic cascade stages were intentionally chosen to be identical, this difference can be ascribed to microscopic physics in the pre-equilibrium evolution of the plasma. This difference in dynamics is most pronounced at early times in the evolution, roughly corresponding to higher temperatures, where the increased constraint is found. The KL Divergence is also non-monotonic for the BMA posterior, which is commensurate with increased constraint around the peak of $\zeta/s$ and the kink of $\eta/s$ as well as constraint of $\zeta/s$ near particlization. This demonstrates model sensitivity to key phenomenological features and less constraint otherwise. As the field attempts to better constrain the understanding of heavy ion collisions by constraining the temperature-dependent transport coefficients, using state-of-the-art physics models rather than parametric substitutes must be a critical plank in the theoretical program.

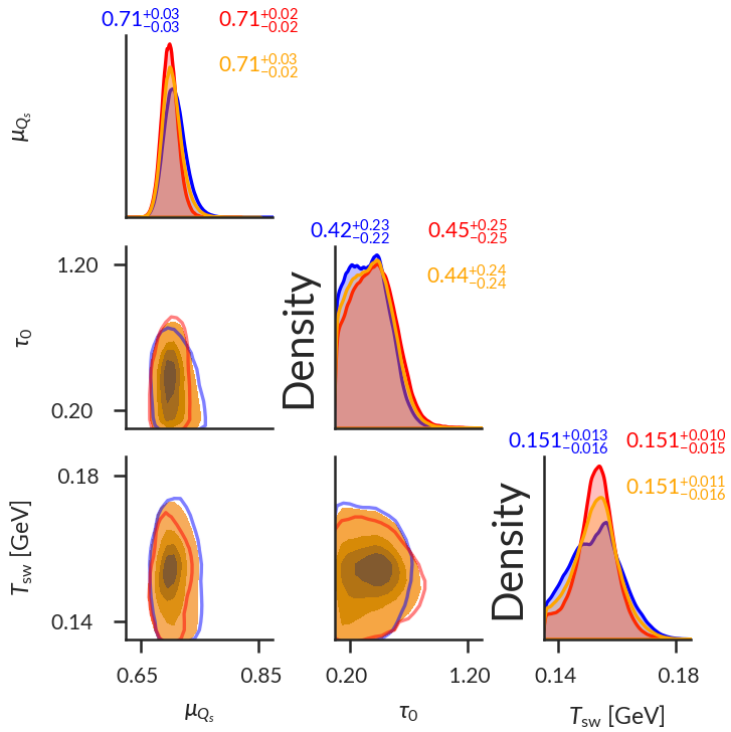


Figure 6.24: Bayes Model Averaged posterior for non-viscous parameters (orange) shown with with Grad (blue) and Chapman-Enskog (red). The lowest contour shown is the $5^{th}$ percentile.

The non-viscous BMA posterior demonstrates this as well and it can be seen in the joint marginal distributions that the BMA posterior is in between those of the two underlying models. This incorporates this source of modeling uncertainty and is the most precise and accurate physical understanding of these quantities yet. The non-viscous posteriors for the Grad and C.-E. viscous correction models are quite similar, demonstrating that these are robust to a modeling choice intended to be a small correction. The largest difference between the two models is in the particlization temperature, which is robustly accounted for in the BMA posterior, and the median value remains consistent between the models.

Bayesian model averaging remains the current state-of-the-art in heavy-ion collisions for leveraging the information in multiple models to best constrain the physical understanding of strongly-interacting matter without over-fitting. This is only the second study in this field, following [46] and elaborated in [47], to utilize BMA for improving uncertainty quantification and has further demonstrated its importance. Further sources of unquantified uncertainty still exist in heavy ion collisions, usually at the interface between models at each stage in the evolution of the fireball, but how to incorporate such interface effects in BMA is not yet clear. The majority of Bayesian effort in studying the strongly-interacting matter produced in heavy ion collisions is in improving the precision of the models; it cannot be emphasized enough that studies that pursue arbitrary precision without accounting for sources of uncertainty using techniques such as BMA do not fully leverage the information available and are falling prey to bias. Simultaneous consideration of observables and uncertainty quantification, as performed in this chapter, are required for reliable inference of the physical properties of strongly-interacting matter.

### 6.4.5 *Calculations at Maximum a Posteriori*

Scientific models can be evaluated by how well they can describe experimental measurements in systematic model-to-data comparison, as performed up to this point, but also by how well they predict quantities to which they were not explicitly tuned. A model that can only describe quantities to which it is systematically compared is less useful than a model that, once compared to a carefully-selected set, makes ac-

curate predictions. The Bayesian inference performed in this chapter was performed using a surrogate model trained at a large number of design points, not the underlying model itself. As a result, before moving on to predictions, it is important to explore the veracity of the MAP points in Table 6.2. To do this, the model is run as before, but with 6000 collision events from $0-13$ fm rather than 2500. This increase in statistics allows for higher precision results.

First, the veracity of the MAP points is determined via postdiction, in which the underlying computationally-expensive hybrid model is compared to quantities used in the inference above. In the following figures, the Grad and C.-E. MAP are shown in blue and red, respectively. The MAP with temperature-dependent $\eta/s$ is shown as a solid line while constant $\eta/s$ is shown as a dashed line. Shaded regions denote uncertainty. Where available, a MAP calculation from [45] is shown for comparison to the previous state-of-the-art Bayesian study in HICs. The charged hadron multiplicity, Fig. 6.25, compares very favorably with the MAP calculations within the experimental uncertainty for all viscous correction models. A variety of identified particle multiplicities and transverse energy per rapidity slice, Fig. 6.26, also compare very well, albeit the proton and kaons are overestimated while the pions are underestimated. This balancing act combined with the overall charged hadron multiplicity shows that aspects of the hadron chemistry are imbalanced. As discussed previously, details such as chemical freezeout (c.f. [100, 204]) are not included in this study and are likely to particularly influence higher-mass particles, particularly kaons. The overestimation of the number of higher-mass particles in turn results in an overestimation of transverse energy. Nonetheless, the differences between the MAP calculations imply an influence of viscous corrections on the hadronic chemistry.

The mean transverse momentum of identified particles, Fig. 6.27, further reveals the success of the model-to-data comparison while demonstrating how overestimation of multiplicity combined with good estimation of the transverse momentum results in overestimation of transverse energy. The $\langle p_T \rangle$ shows less tension in the chemical makeup than previous results with the same hydrodynamic equation of state, revealing the role of bulk viscosity and a physically-motivated pre-equilibrium model with microscopic dynamics. The primary difference between Grad and C.-E. MAP calcu-

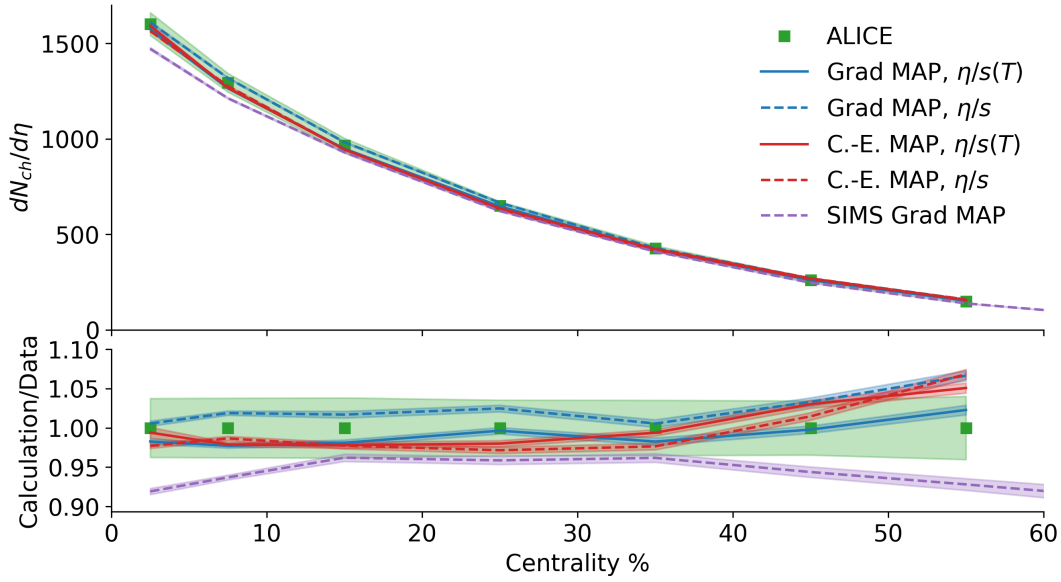


Figure 6.25: Postdictions of charged hadron multiplicity at Maximum a Posteriori.

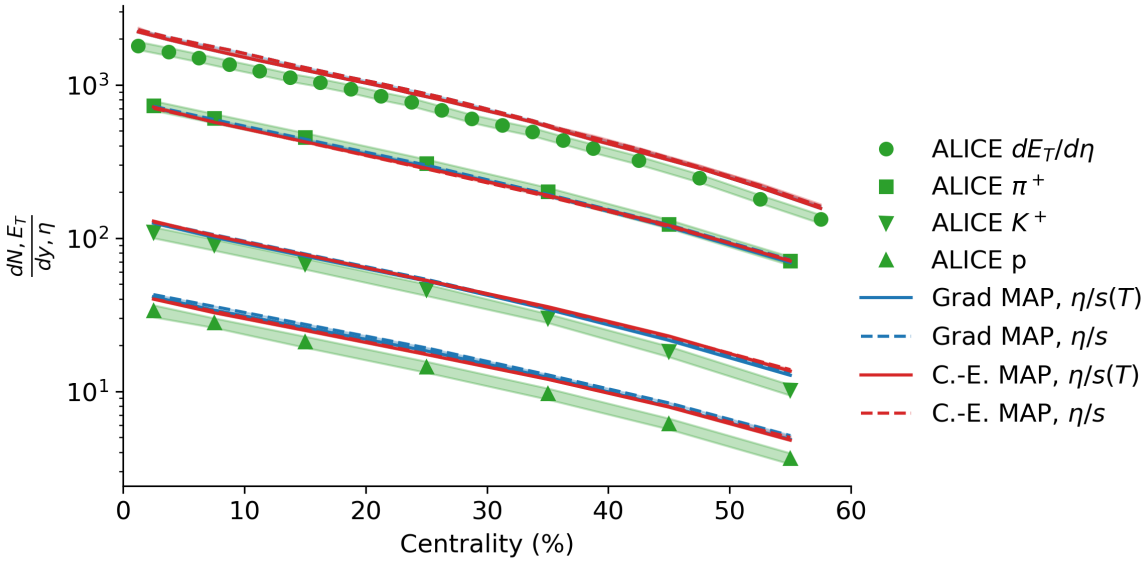


Figure 6.26: Postdictions of identified hadron multiplicity at Maximum a Posteriori.

lations is in enhanced proton $\langle p_T \rangle$, in which the C.-E. MAP better reproduces the experimental results.

The two-particle integrated $v_n$ further reveal good, albeit imperfect, reproduction of experimental results in Fig. 6.28. Notably, $v_2\{2\}$ and $v_4\{2\}$ are well described, particularly in central collisions, while $v_3\{2\}$ is underestimated. The underprediction of $v_3\{2\}$ is a feature of nearly every study and remains an object of continuing study. Peripheral $v_2\{2\}$ reveal that the MAP temperature-dependence of the Grad shear viscosity

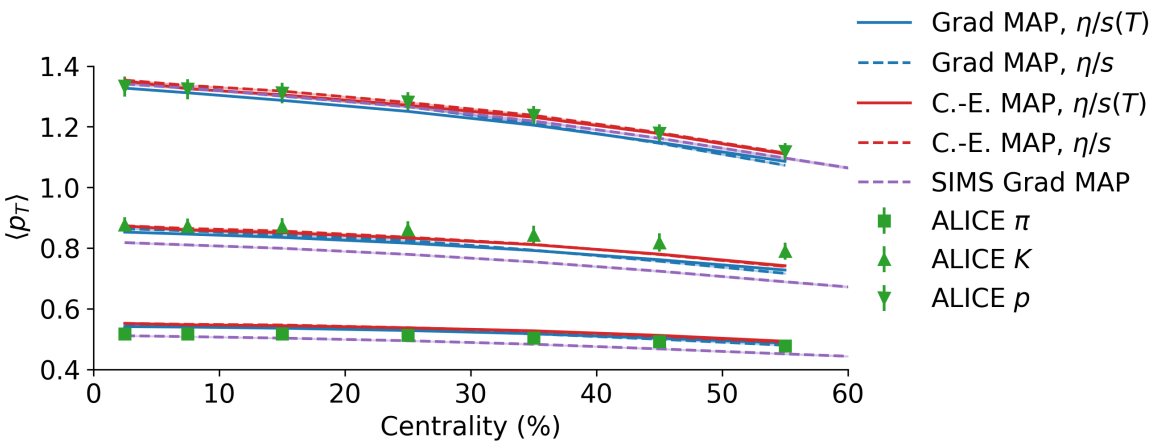


Figure 6.27: Postdiction of identified particle $\langle p_T \rangle$ at Maximum a Posteriori.

results in an overestimate, while the constant shear more closely reproduces the experimental centrality dependence as do the C.-E. MAP calculations. For all $v_n\{2\}$, the MAP prediction of this study performs better than the previous state-of-the-art and the tension revealed here produces useful insight both into temperature-dependent $\eta/s$ and remaining progress required in describing the geometric fluctuations that drive $v_3\{2\}$. The four-particle integrated $v_2$ is shown in Fig. 6.29, showing agreement with data until the most peripheral bin where it is overestimated, consistent with the two-particle $v_2$ in Fig. 6.28, suggesting that these observables capture broadly similar physics and are similarly-well described by the model, although less tension is observed in $v_2\{4\}$ compared to $v_2\{2\}$.

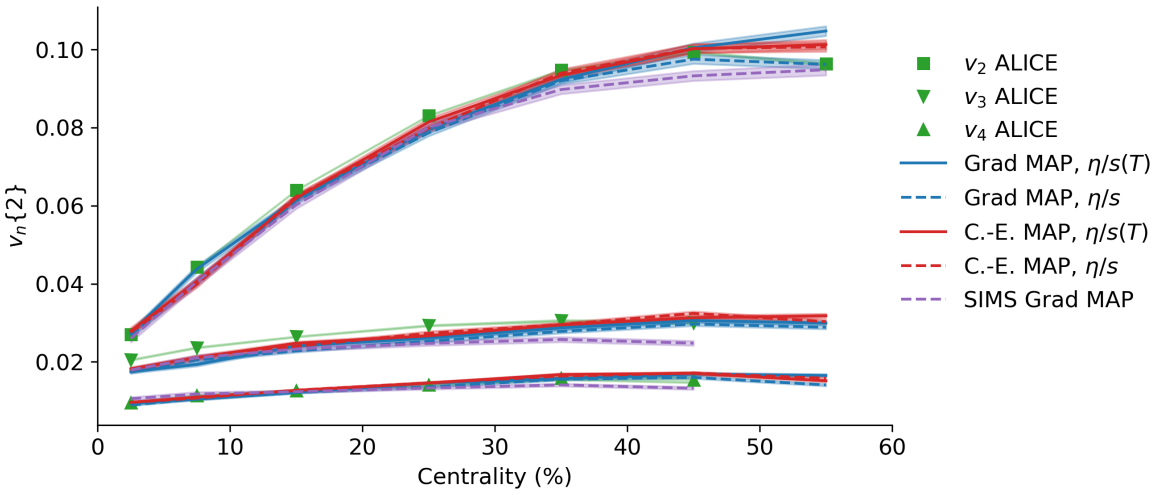


Figure 6.28: Postdiction of $v_n\{2\}$ at Maximum a Posteriori.

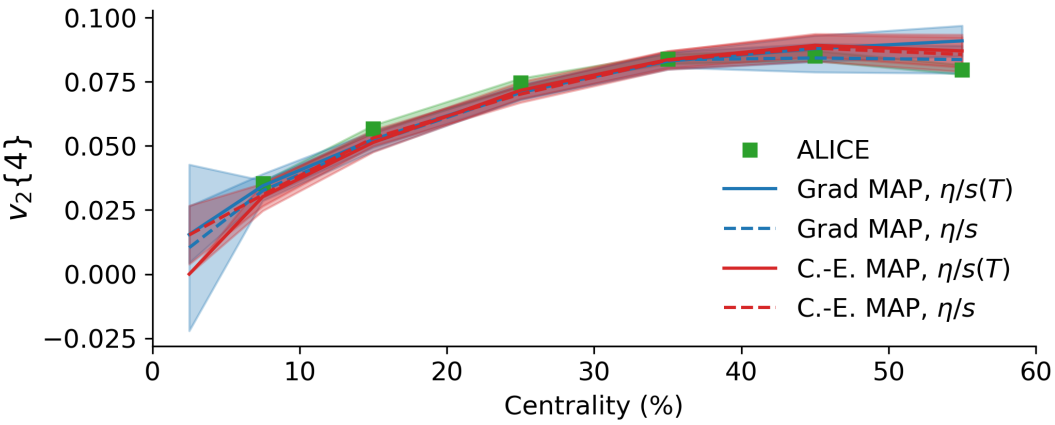


Figure 6.29: Postdiction of $v_2\{4\}$ at Maximum a Posteriori.

The correlated momentum fluctuations $\delta p_T/\langle p_T\rangle$, also denoted $\sqrt{C_m}/M$ or $\sqrt{C_m}/\langle p_T\rangle$, in Fig. 6.30 are the first calculations to successfully describe this observable from a model with an IP-Glasma pre-equilibrium state and this description is consistent. The only prior work showing this calculation matches less well and does so by underestimating charged hadron multiplicity [205], while this study is able to simultaneously describe both quantities with a variety of different viscosities and viscous corrections. These fluctuations are also sensitive to the temperature-dependence of the specific shear viscosity, where the constant $\eta/s$ systematically overestimates the data while correctly reproducing the centrality dependence (itself not seen in either other calculations with IP-Glasma or in the previous state-of-the-art), while the temperature-dependent $\eta/s$ better reproduces the data beginning in mid-central collisions. The C.-E. MAP reproduces the fluctuations more closely, save for the $\eta/s(T)$ calculation in the most central bin, which is likely the impact of statistical fluctuations.

The decomposition of higher order $v_n$ further reveals the ability to simultaneously describe flow observables in Fig. 6.31. For every quantity other than central $v_4^L$, both models produce successful predictions of the experimental data, with the temperature-dependent $\eta/s$ again overpredicting peripheral flow as seen in $v_4(\Psi_2)$. Although is often consistent with the data within uncertainty, $v_4^L$ is overpredicted by the C.-E. MAP calculations. Nonetheless, this broad reproduction of the experimental flow decomposition suggests that the momentum-space geometry of the hybrid model successfully reproduces the physical picture in heavy ion collisions.

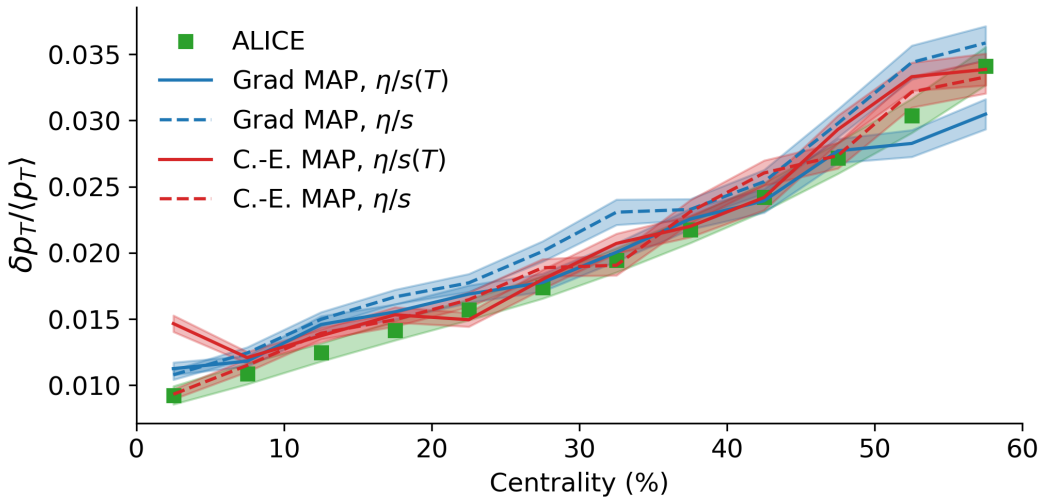


Figure 6.30: Postdiction of $\delta p_T/\langle p_T \rangle$ at Maximum a Posteriori.

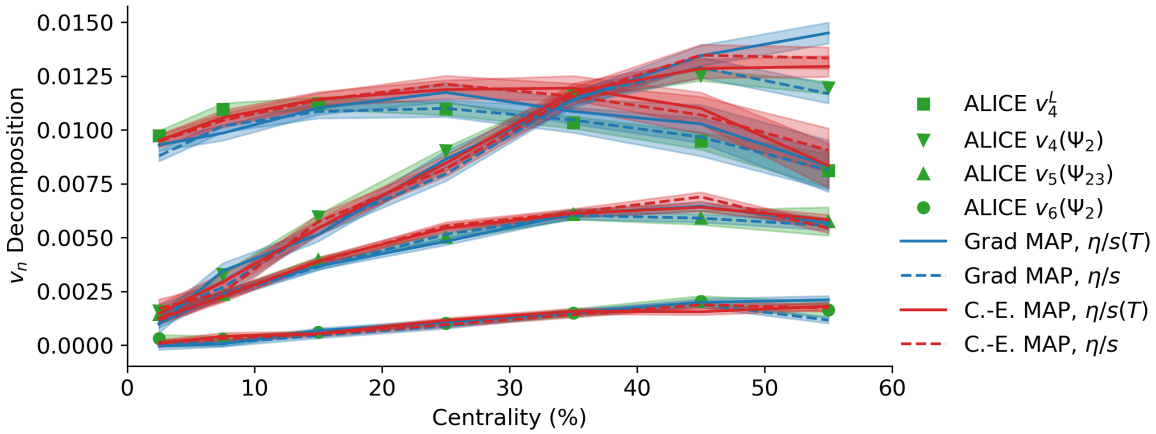


Figure 6.31: Postdiction of the decomposition of $v_n$ at Maximum a Posteriori.

As highlighted in Ch. 4, the simultaneous reproduction of flow decomposition and event plane correlation constrains both the initial state geometry and the hydrodynamic evolution. In Fig. 6.32, the correlators are also well-described by the postdictions and are consistent with experimental uncertainty, save for central $\langle \cos(2\Phi_2 + 4\Phi_4 - 6\Phi_6) \rangle$. The purely-even correlations are particularly well described and primarily relate the conversion of event planes of initial state geometry to momentum space via hydrodynamics. The mixed even-odd plane correlations reveal that the fluctuation structure is well described and correlates properly with even planes. These calculations can be compared to Figs. 4.9, 4.10, and 4.12 where the previous state-of-the-art systematically underpredicted correlations and the previous IP-Glasma results pre-

dicted less well. This postdiction is also well in line with the posterior predictive distributions, further supporting the accuracy of the surrogate modeling.

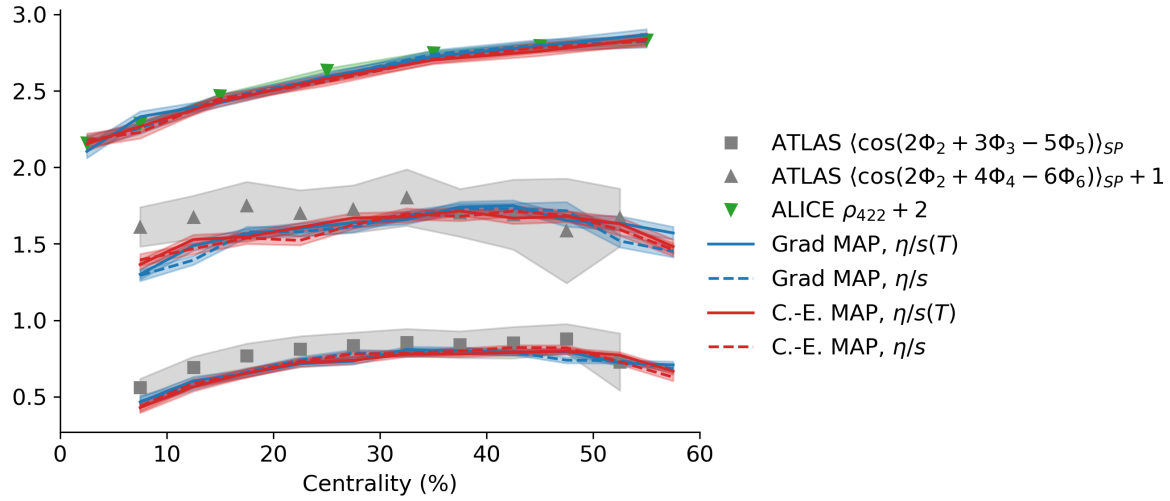


Figure 6.32: Postdiction of event plane correlations at Maximum a Posteriori. Data and calculations are shifted for clarity.

The postdictions show that MAP parameter values are able to successfully describe the observables used in inference with never-before-seen accuracy for a hybrid model with an IP-Glasma pre-equilibrium stage. This alone is a resounding success of the Bayesian inference in this study and conclusively demonstrates the performance of the Gaussian Process emulators as well as the study design. Tension is seen in the hadron chemistry, impacting the transverse energy, as well as in some of the description of the flow harmonics, notably $v_2\{2\}$ and $v_3\{2\}$. However, the decomposition of higher order flow is successful and the overwhelming majority of observables are well-described while the same tension is seen in $v_2\{4\}$, ensuring this is effect is not a result of two-particle correlations. In the case of $\delta p_T/\langle p_T\rangle$, successful description is shown for the first time. The impact of viscous corrections is minimal, showing that the different posteriors are accurately accounting for differences in the underlying model calculations.

#### 6.4.5.1 *Predictions*

Having established the success of the surrogate modeling and demonstrated unprecedented description of a wide range of observables, it is time to turn to predictions of

quantities not included in the calibration. Here, "predictions" is used to highlight that these observables were not used in systematic comparisons. As a result, the model is blind to these observables beyond information contained in other quantities. If models are differentiable at this stage, perhaps it can shed light on model quality not revealed in the more limited model-to-data comparison. In the following comparisons, centrality bins are chosen to match experimental results and predictions for bins not shown are simply due to dominance by theoretical uncertainty from a small number of events per bin.

The comparisons begin with measures of event plane correlation from ALICE in Fig. 6.33 and ATLAS in Fig. 6.34. In both cases, the model predictions are very well-aligned with experimental results. Both $\rho_{532}$ and $\rho_{633}$ are accurately predicted within experimental uncertainty, while $\rho_{6222}$ is accurately predicted below 30% centrality. With respect to the ALICE measurements, the MAP calculations are broadly indistinguishable. A similarly indistinguishable picture is painted by comparison to ATLAS measurements, where $\langle\cos(2\Phi_2+3\Phi_3+4\Phi_4)\rangle_{SP}$, $\langle\cos 6(\Phi_2-\Phi_3)\rangle_{SP}$, and $\langle\cos 6(\Phi_3-\Phi_3)\rangle_{SP}$ are very well predicted by the model. Two predictions that perform less successfully above 30% centrality are $\langle\cos 4(\Phi_2-\Phi_4)\rangle_{SP}$ and $\langle\cos 6(\Phi_2-\Phi_6)\rangle_{SP}$. The consistent picture drawn from comparison to both experiments suggests that the event planes produced by an IP-Glasma initial state are better suited to collisions below 30% despite successful comparison to observables across the whole centrality range. These predictions outperform previous predictions made by a hybrid model with IP-Glasma, shown in Ch. 4.

A motivation for the use of IP-Glasma as a pre-equilibrium model was its success in simultaneous description of next generation observables, particularly both the event plane correlations and nonlinear response coefficients. With demonstrated success in prediction of event plane correlations not used in model-to-data comparison, predictions for nonlinear response coefficients are shown in Figs. 6.35 and 6.36. These broadly describe the experimental results within experimental uncertainty, with slight overestimation in $\chi_{5,23}$ between 20 and 40% centrality and peripheral $\chi_{4,22}$. In this case, the model with $\eta/s(T)$ slightly outperforms predictions with constant $\eta/s$, although they are often consistent within standard error. This demonstrates that a hybrid model with an IP-Glasma pre-equilibrium stage is able to produce simultaneous

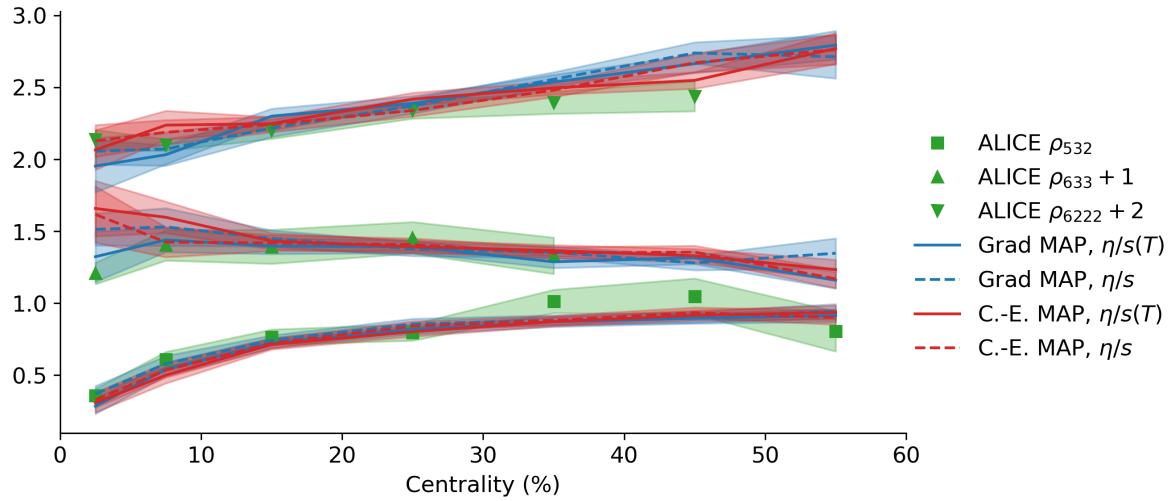


Figure 6.33: Prediction of ALICE event plane correlations at Maximum a Posteriori. Data and calculations are shifted for clarity.

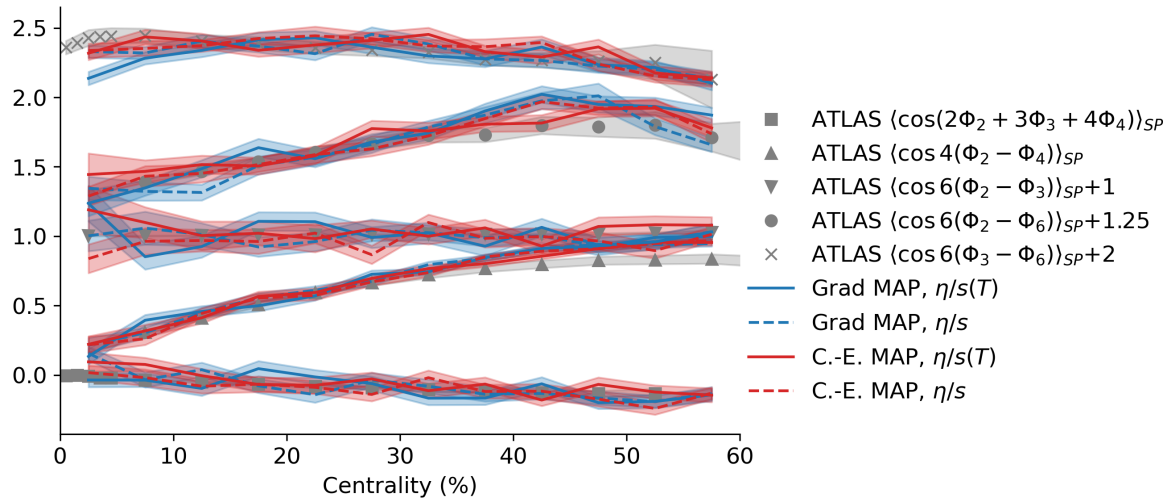


Figure 6.34: Prediction of ATLAS event plane correlations at Maximum a Posteriori. Data and calculations are shifted for clarity.

accurate predictions of the event plane correlations and hydrodynamic response with a initial geometry broadly fixed by low-energy nuclear correlations. This strongly suggests that the hydrodynamic phase is accurately described as there is no geometric flexibility to exploit and the hydrodynamic response to geometry matches that seen in experiment. The centrality dependence is also often accurately captured, such as in $\chi_{5,23}$, which was not the case in previous calculations.

Predictions for the final category of observables used in the analysis are shown in Fig. 6.37 for the linear and nonlinear flow decomposition. These predictions are ac-

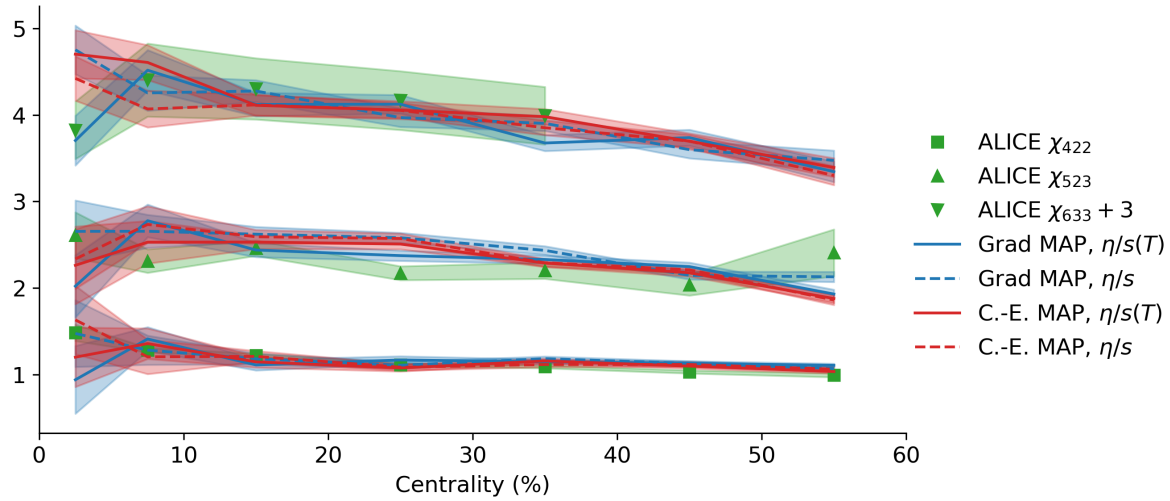


Figure 6.35: Prediction of ALICE nonlinear response coefficients at Maximum a Posteriori. Data and calculations are shifted for clarity.

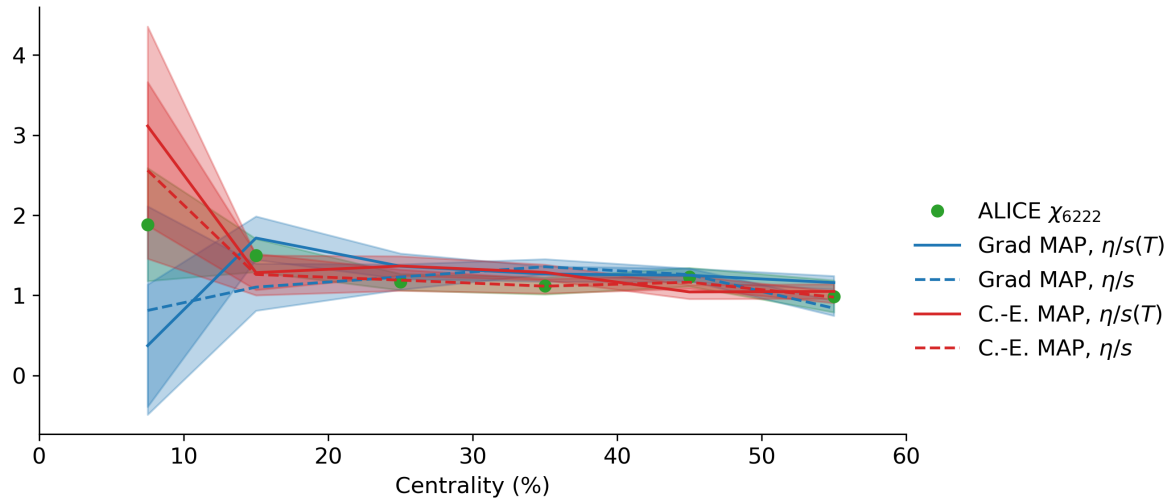


Figure 6.36: Prediction of the ALICE $\chi_{6222}$ nonlinear response coefficients at Maximum a Posteriori. Data and calculations are shifted for clarity.

curate and are clearly consistent with experimental results within uncertainties, save for $30-40\%$ $v_5^L$. This demonstrates the continuing success of the hybrid model with IP-Glasma as it is able to both describe and predict a wide range of observables. In the $v_5^L$ predictions, the constant $\eta/s$ prediction is more consistent with the experimental measurement, further supporting an inconclusive preference for one model over the other as the quality of predictions depends on which observable is considered.

The final $p_T$-integrated prediction is made for the modified Pearson correlation between $v_2^2$ and $p_T$, shown in Fig. 6.38. As no experimental results at this energy are

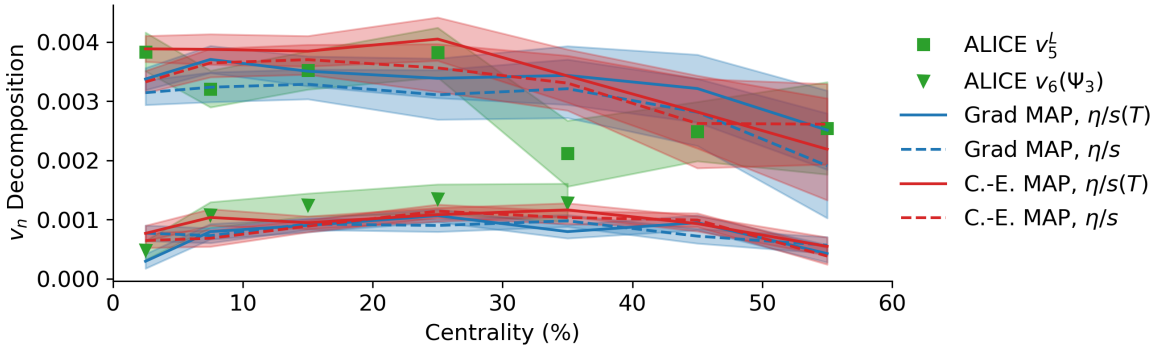


Figure 6.37: Prediction of ALICE linear and nonlinear flow at Maximum a Posteriori.

available, preliminary results for a higher Pb-Pb collision energy system ($\sqrt{s_{NN}} = 5.02$ TeV) are used [206] for comparison. The predictions made at $\sqrt{s_{NN}} = 2.76$ TeV describe the higher-energy data and its centrality dependence well, which is not seen in T$_R$ENTo-based hybrid model predictions and has been shown to be sensitive to nucleon size [201]. This study has not utilized sub-nucleonic degrees of freedom and has used a nucleon size of 4 GeV$^{-2}$. There is no significant difference seen between predictions with different viscous corrections or between $\eta/s$ and $\eta/s(T)$. Even with variation of the nucleon width in previous calculations of this quantity with IP-Glasma and T$_R$ENTo-based hybrid models, successful prediction of the value and centrality dependence has proved elusive. Hybrid models with T$_R$ENTo + freestreaming initial states, as well as previous calculations with IP-Glasma, have sign changes as they become increasingly peripheral. This feature is not seen in the data, nor in this prediction. Based on the prediction in Fig. 6.38, there is no anticipated collision-energy dependence of this correlation and the IP-Glasma initial state at maximum a posteriori is able to successfully describe this observable. The lack of collision-energy dependence is supported by the comparison of Pb-Pb at $\sqrt{s_{NN}} = 5.02$ TeV data compared to Xe-Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV data from ALICE [206]. Of note is that it appears to not yield further constraint on the temperature-dependence of $\eta/s$. Nonetheless, comparing it directly to the previous state-of-the-art Bayesian study using a T$_R$ENTo + freestreaming initial state, one immediately sees that the microscopic physics of the IP-Glasma pre-equilibrium stage plays an important role in capturing the relevant physics. This represents a true prediction as data at $\sqrt{s_{NN}} = 2.76$ TeV has yet to be published.

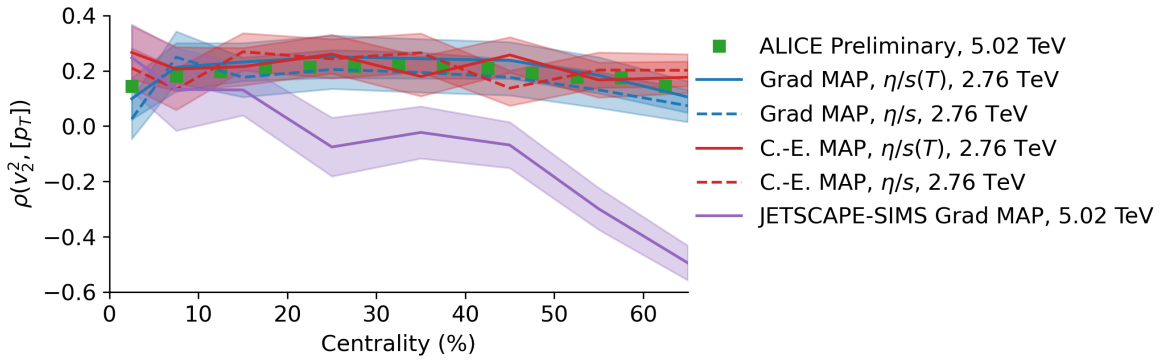


Figure 6.38: Prediction of correlation between $v_2^2$ and $p_T$ at Maximum a Posteriori, compared to data from a higher-energy collision and a prediction using [45]. Note that data and the JETSCAPE prediction are at $\sqrt{s_{NN}} = 5.02$ TeV while the MAP predictions are at $\sqrt{s_{NN}} = 2.76$ TeV.

Up to this point, only $p_T$-integrated observables have been considered. Differential observables also exist and provide interesting and discriminating probes of the soft sector. However, the boundary between the soft sector and the hard sector (such as jets and jet-medium interactions) is unclear. By considering the integrated quantities up to now, the sensitivity of the inference to the precise location of this boundary is reduced and predictions can be made. This sensitivity is reduced because integrated observables are weighted by the multiplicity, which drops exponentially. By considering each differential $p_T$ bin, this exponentially-decreasing weighting would be removed and each bin would be treated on an equal footing, in turn giving the bins on the boundary of the soft and hard sectors a higher proportional weighting.

The first differential observable investigated is the differential charged hadron $v_n\{2\}$, with predictions shown in Fig. 6.39 compared to experimental measurements from ALICE [116]. Tension is clearly present in reproducing the spectra, with predictions from integrated observables often undershooting at lower transverse momentum and overshooting at higher momenta. Nonetheless, the majority of predictions are consistent with experimental measurements for the first time or the distance from the prediction to measurement has been greatly reduced from the previous IP-Glasma state-of-the-art [28]. The greatest tension is observed in the differential $v_2\{2\}$ in the $0-5\%$ and $30-40\%$ centrality bins and low-$p_T$ $v_3\{2\}$ in more peripheral collisions. This low-momentum region is expected to be the region best described by hydrody-

namics, suggesting that relevant physics remains missing from the hybrid model. As $v_3\{2\}$ is primarily fluctuation driven, this suggests that fluctuation structure is missing. The underestimate of $v_2\{2\}$ in contrast suggests that a geometric aspect is not included or an aspect of the conversion between position-space and momentum-space geometry remains incomplete. This is not necessarily a concern for the validity of the hydrodynamic description, as the higher-order differential $v_n$ are well-described, but instead suggests that additional physics may be at play. Recent works including the differential momentum spectra suggest that their inclusion in systematic model-to-data comparison can yield insight, but various analysis errors and inclusion of momentum bins in regions where unincluded physics is relevant hinders the interpretation of results [50, 51]. The posterior predictive distribution, rather than single MAP predictions, may provide more insight into the present apparent mismatch of the model predictions and data.

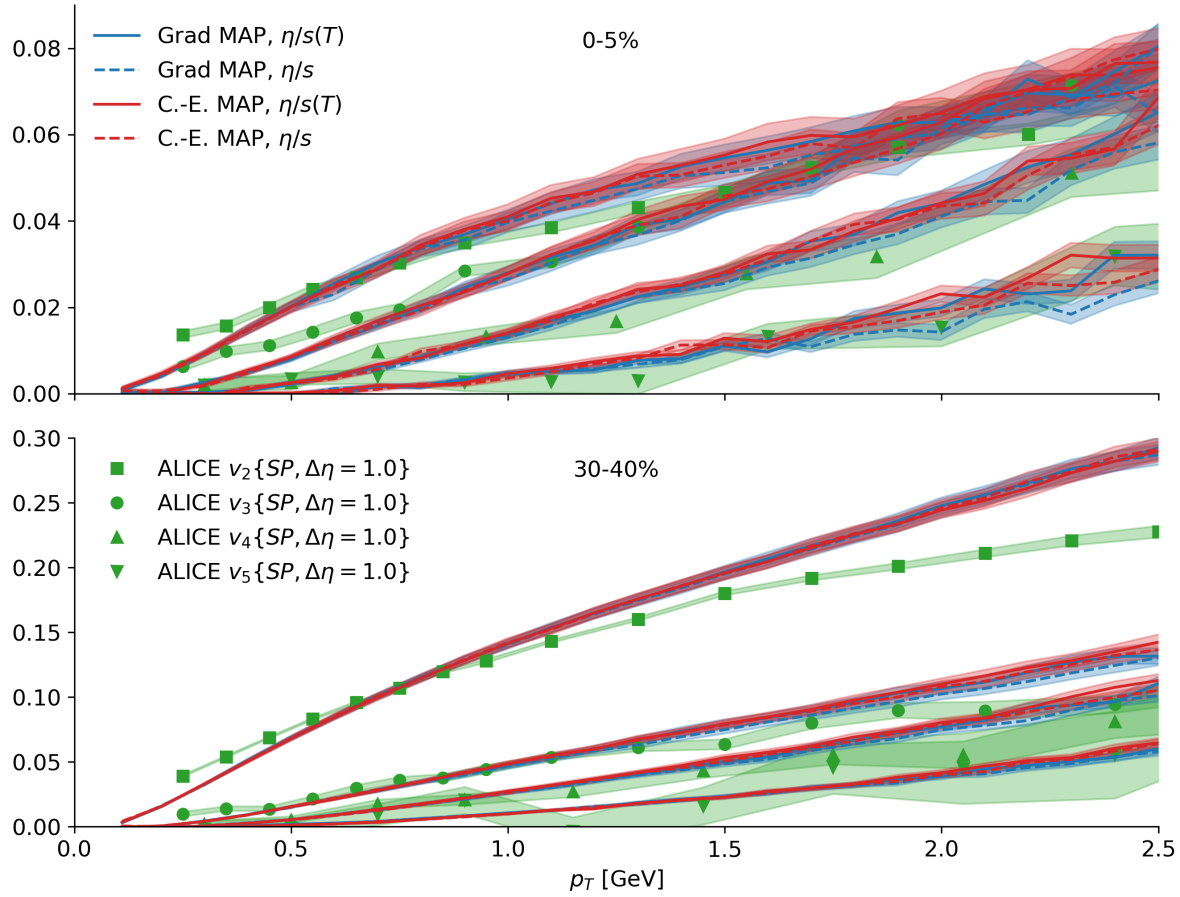


Figure 6.39: Prediction of differential $v_n\{2\}$ at Maximum a Posteriori for the $0-5\%$ centrality bin (upper panel) and the $30-40\%$ centrality bin (lower panel).

The light hadron multiplicity spectra, shown in Fig. 6.40 for selected central and mid-central centrality bins, paints a complimentary picture to the integrated multiplicity in Fig. 6.26. The integrated proton and kaon multiplicity were overestimated, while the pion multiplicity was slightly underestimated; the same is found here. The momentum dependence of the spectra, however, remains well-predicted until the higher momentum region ($p_T > 1.75$ GeV), where mini-jets, jet showers, and other hard-sector considerations begin to gain relevance. Beginning in this region, all the identified light hadrons are underpredicted. To include these additional effects is a matter of ongoing theoretical effort and is beyond the scope of this thesis.

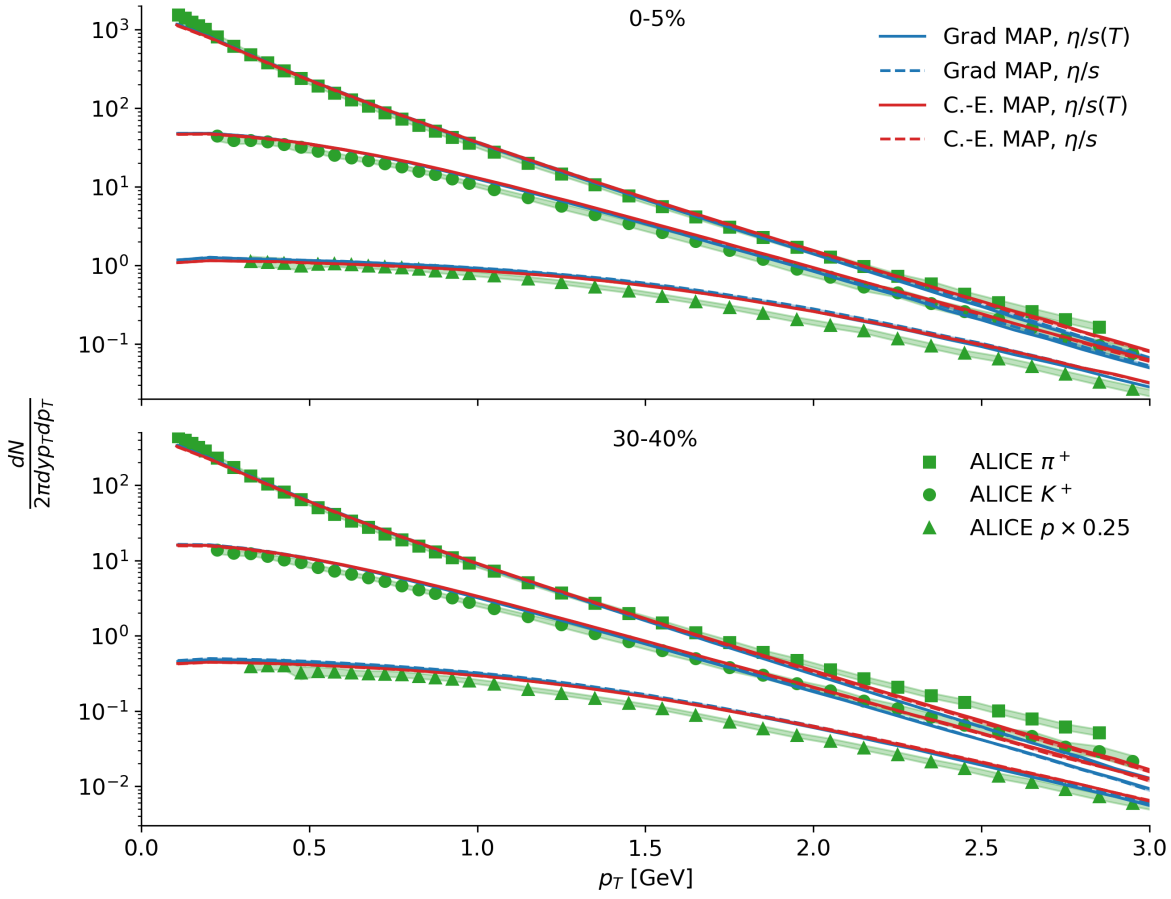


Figure 6.40: Prediction of differential light hadron multiplicity spectra at Maximum a Posteriori for the $0-5\%$ centrality bin (upper panel) and the $30-40\%$ centrality bin (lower panel).

Postdictions and predictions using four MAP calculations have been shown, comparing Grad and Chapman-Enskog viscous corrections with and without temperature-dependent shear viscosity. The inconclusive preference between viscous correction

models is consistent when comparing MAP parameter sets, as is the inconclusive preference for or against temperature-dependent $\eta/s$, in keeping with the Bayesian model comparison.

This study has implemented rigorous Bayesian model-to-data comparison with IP-Glasma for the first time, incorporated transfer learning for the first time, and has demonstrated ongoing inconclusive preference between viscous correction models and between temperature-dependent and temperature-independent shear viscosity. A large number of postdictions and predictions are shown at maximum a posteriori and should be considered the new state-of-the-art theoretical result to which future measurements and calculations should be compared. The posterior distributions are the main results of this chapter and are the current best estimate of the properties of strongly-interacting matter in ultra-relativistic heavy ion collisions.

# 7

# CONCLUSIONS, OUTLOOK, AND FINAL REMARKS

## CONCLUSIONS

In this thesis, the tools of Bayesian inference were used to study the strongly-interacting matter produced in heavy ion collisions, employing a hybrid model with an IP-Glasma pre-equilibrium state in such a framework for the first time. This study was motivated by extensive comparison to the previous state-of-the-art and advances were made to the computer experiment design, incorporating non-uniform priors and design space sampling for the first time in heavy ion collisions (Ch. 4). The self-consistency of IP-Glasma was demonstrated using a limited closure test and it was shown that a hydrodynamic stage calibrated using a parametric initial-state model was unable to successfully reproduce experimental measurements (Ch. 5). A large-scale Bayesian study was then performed, simultaneously varying the remaining free parameters of IP-Glasma, the transport coefficients of the hydrodynamic phase, and the particlization temperature (Ch. 6).

The motivations for this study were collected in Ch. 4 and demonstrated that systematically motivating a study yields useful insight into how to constrain physical properties. Certainly, a prior belief existed that a physically-inspired model with microscopic physics was a higher-fidelity description than a parametric model with parametric dynamics. By demonstrating specifically where a systematically-constrained parametric model fails, it was possible to identify discriminating quantities and interpret them physically. With this in hand, it was possible to design a study that exploited the information contained in these quantities. The design, also described in Ch. 4, was another source for methodological progress. The ordered Maximum Projection Latin Hypercube has distinct advantages to the maximin Latin Hypercube used by previous studies. In particular, a fair sampling of the design space is reached much earlier and the active space of the model is well-covered rather than only the full de-

sign space and its 1-dimensional projections. This produced clear improvements in surrogate model performance. Non-uniform priors were also motivated and implemented, beginning the process of more accurately describing physical understanding of these quantities. The previous standard – a uniform prior – has unphysical features, such as sharp cutoffs without explicit justification. By developing and demonstrating non-uniform priors, a more faithful representation of physical priors can be used and design points can be placed in regions with higher prior probability.

A Bayesian modeling workflow was also introduced and implemented throughout this work. This formalizes the process of model development, comparison, and criticism. It was shown with a pedagogical, but nonetheless physical, example – the simple pendulum. In this example, adapted from [1], a common piece of introductory laboratory guidance was investigated using the tools of Bayesian model comparison. It was found that the guidance currently standard in laboratories is overly-constraining in the absence of relatively precise timing.

The self-consistency of IP-Glasma had not been established before Ch. 5, which demonstrated the total covariance of the strong coupling $g_s$ with the multiplier between the color charge density profile and the saturation scale, $\mu_{Q_s}$. Once this total covariance was broken by fixing the strong coupling to a value commensurate with the generated simulations, a hybrid model with an IP-Glasma initial state was able to recover the input parameters of IP-Glasma. This step demonstrated that the sensitivity of chosen final-state hadronic observables to the pre-equilibrium phase was sufficient to reliably extract accurate information. The self-consistency of the subsequent phases had been demonstrated in other studies, but this demonstration served to show that IP-Glasma was well-suited for a large-scale statistical study.

The final study in this thesis, Ch. 6, simultaneously varied physical parameters of the hybrid model in the pre-equilibrium and hydrodynamic phases. The ranges over which these parameters were varied are given in Tab. 6.1 and the maximum a posteriori estimates for their values are given in Tab. 6.2. It implemented transfer learning for the first time in heavy-ion collisions and used Bayesian model comparison and Bayesian model averaging to estimate the most likely true values of the physical quantities considered. Of particular focus were the the specific shear and bulk viscosity of strongly-interacting matter, the understanding of which are a long-standing

goal of the field. By using multiple viscous correction models, an important source of otherwise-unquantified theoretical uncertainty was studied for only the second time in heavy-ion collisions and was found to have less impact than in previous studies.

The temperature dependence of the specific shear viscosity remains indeterminate, but is substantially constrained in the high-temperature/early-time regime of the collision. This is the result of much reduced parametric flexibility in the initial state and in the inclusion of microscopic physics in the pre-equilibrium dynamics. The specific bulk viscosity $\zeta/s$ was found to be large and peaked during the hydrodynamic phase and is strongly inconsistent with zero.

The proportionality between the color charge density and the saturation scale was successfully constrained and found to be only weakly sensitive to the impact of different viscous corrections. To my knowledge, this is the first such systematic constraint derived for a Color Glass Condensate Model. The switching time between the pre-equilibrium stage and hydrodynamics was also well-constrained and was found to be incompatible with a long pre-equilibrium evolution. In fact, pre-equilibrium evolution longer than 0.6 fm was strongly disfavored, in contrast to many studies with a freestreaming pre-equilibrium stage which sometimes evolve for up to 1.5 fm. This short time to the onset of hydrodynamics is the first Bayesian study, to my knowledge, to yield a constraint consistent with the theoretical expectations of approximately $1/Q_s$. This is another advantage of using a physically-inspired model with microscopic physics: with realistic hybrid model components, realistic constraints are derived. It is my hope that with future rigorous study and inclusion of more discriminating observables, these constraints can be improved yet further.

The particlization temperature was also well-constrained by the hybrid model for both viscous correction models. Excitingly, these constraints are consistent with the latest lattice QCD estimates of the crossover in the QCD phase diagram. Details of viscous particlization models were investigated, but chemical freezeout remains an exciting avenue for future research not included in this thesis. Nonetheless, the constraint on particlization temperature's consistent identification of the change of the degrees of freedom from the quark gluon plasma to a hadron gas is another success of this work.

Calculation of observables not used in the Bayesian model-to-data comparison were also shown, demonstrating the striking predictive power of the model in both non-linear response coefficients and in $p_T$-differential quantities. Comparisons are shown with both constant and temperature-dependent $\eta/s$, revealing little systematic preference for one over the other in either postdictions or predictions. Similarly, Grad and Chapman-Enskog viscous corrections are investigated and no significant preference is observed.

A true prediction for $\rho(v_2^2, [p_T])$ was shown for $\sqrt{s_{NN}} = 2.76$ TeV and compared to data at 5.02 TeV in Fig. 6.38. The maximum a posteriori predictions are consistent with the higher-collision-energy data and with its centrality dependence, again a first to my knowledge. This suggests that the most likely outcome of such a measurement is that there is little $\sqrt{s_{NN}}$ dependence in this observable and that a hybrid model with an IP-Glasma pre-equilibrium stage is able to successfully describe both underlying quantities as well as their correlation.

The hybrid model is – of course – incomplete, but useful. By systematically comparing the model to data while including sources of uncertainty, it is possible to guide theoretical effort on a quantitative foundation. For example, a tension between multiplicity and transverse energy first identified in parametric studies remains in this one, pointing to a need for increased study in the chemical construction of the equation of state. The models considered in this thesis are also wholly (2+1)-dimensional. (3+1)-dimensional studies are currently prohibitively computationally expensive, but remain a rich avenue for the study of both longitudinal dynamics and potential further constraint of the transport coefficients of strongly-interacting matter. A final limitation is that only one collision system was considered, Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. Other studies have found comparatively little additional benefit from including multiple collision systems, but with the increased constraint demonstrated by this study it is plausible that with multiple systems, the temperature dependence – or lack thereof – of the specific shear viscosity could be determined.

## OUTLOOK

The future of Bayesian techniques in physics is bright, with the potential for application ranging from the gamut from introductory physics laboratory (Ch. 3.6) through to heavy-ion collisions. I briefly outline a few avenues for contribution to this burgeoning sub-discipline in the hope that it proves useful to future researchers.

First, recent studies are increasingly expending ever-greater computational resources to drive down statistical uncertainty from Monte Carlo models and thus better constrain the properties of strongly-interacting matter. This must be balanced with increased efforts for uncertainty quantification. An arbitrarily precise but potentially inaccurate result is of no scientific value. The choice of observables must be motivated by a physical understanding of what must be learned and care must be taken to ensure that the model is best suited to the data at hand. For example, care must be taken to ensure that the model is compared to data in the realm of its applicability; attempting to describe a regime in which jet effects become prominent with a purely hydrodynamic model will not produce useful results. The inclusion of more new observables must be guided by what can be learned by the comparison.

Another method of reducing expenditures is by developing and employing novel surrogate modeling strategies, such as the transfer learning employed in this thesis. By choosing an intentionally weakly-informative prior for the overall normalization, transfer learning and multifidelity emulation can be exploited to extend this study to other collision energies and systems. It is my hope and expectation that this will improve the understanding of the material produced in heavy-ion collisions and that it sheds further light on the still-indeterminate temperature dependence of $\eta/s$.

A still-unquantified source of theoretical uncertainty is the onset of hydrodynamics. KøMPøST [207, 208] seeks to model pre-equilibrium dynamics with non-equilibrium linear response to more smoothly approach hydrodynamics while other approaches seek to remove the constraint that the transition between the pre-equilibrium and hydrodynamic phases is isochronous throughout the medium [209–214]. These approaches have not yet been included in a Bayesian study, but have potential to improve this transition and would be of great value.

A final important avenue for progress is in dialogue with experimental collaborations. Bayesian studies with efficient surrogate models can be used to determine the potential constraint yielded by new observables or better measurement of existing observables by artificially reducing errors to determine optimal targets for experimental effort. The correlation between experimental uncertainties is still rarely reported and better elucidation of the shape of these uncertainties would be of great value. This thesis took the maximum entropy approach and assumed the uncertainties were Gaussian and uncorrelated, but this is unlikely to be the case of systematic errors that undergo rigorous experimental study. More information about the correlation and distribution of these uncertainties will yield further information via priors placed on the covariance matrix.

## FINAL REMARKS

The novel results and methodological improvements presented in this thesis represent the state-of-the-art of prior elicitation, computer experiment design, parameter estimation, surrogate modeling with transfer learning, and uncertainty quantification in heavy ion collisions. The Bayes model averaged transport coefficients are the best estimate of the properties of hydrodynamic strongly-interacting matter. By establishing a clear workflow for Bayesian inference, the progress made in this thesis can clearly inform future study design and progress in understanding the most perfect fluid. By further developing the statistical tools and incorporating the leading physical models, this thesis was able to advance the knowledge of the strongly-interacting matter produced in heavy ion collisions.

It is astounding that strongly-interacting matter is produced on Earth. When it was first produced in colliders, it was likely the first time that hot and dense QCD matter was produced in the aftermath of the Big Bang. The inference of the properties of this material is no less stunning than if, with only the rusted remains of the Crash at Crush, a dedicated global effort was able to build a working steam locomotive. The theoretical program will continue, incorporate additional physics, and the understanding of hot QCD matter will improve. This thesis is another tie in building that railroad.

Part IV

# APPENDIX

# A

# WORKFLOW VALIDATION

In order to do a systematic comparison of full events from T$_R$ENTo vs. IP-Glasma, it is necessary to actually produce full T$_R$ENTo hybrid model events. The work of the JETSCAPE Collaboration's Simulations and Distributed Computing (SIMS) working group (of which I am a member) serves as the reference model and calculation. The goal is to reproduce the Grad MAP results, within the limits of Monte Carlo simulations, in order to demonstrate a validated model that can then be used compare other final-state observables. The parametrizations and modeling choices of [45] are reproduced exactly. Brief comparisons are shown here to confirm that the workflow was validated before the pre-equilibrium model was replaced with IP-Glasma.

## A.1 EQUATIONS OF STATE

In Fig. A.1, it is apparent that the EoS used in [28] differs by a maximum of 2% with that used by JETSCAPE. With large bulk viscous corrections, this can introduced unquantified differences. As a result, software[1] was adapted to properly reproduce equations of state with the HotQCD result [94] matched to both UrQMD and SMASH particle lists at the crossover. The results produced by this software are shown in Fig. A.1, where it can be explicitly seen the the equations of state are identical.

## A.2 FINAL STATE COMPARISON

The results of several hundred events were used to perform a statistically-noisy validation to ensure that model predictions with a hybrid workflow matching that of [45] was able to reproduce the results of that work. Single-event tests were performed, but

1 available at https://github.com/mrhheffernan/eos_maker

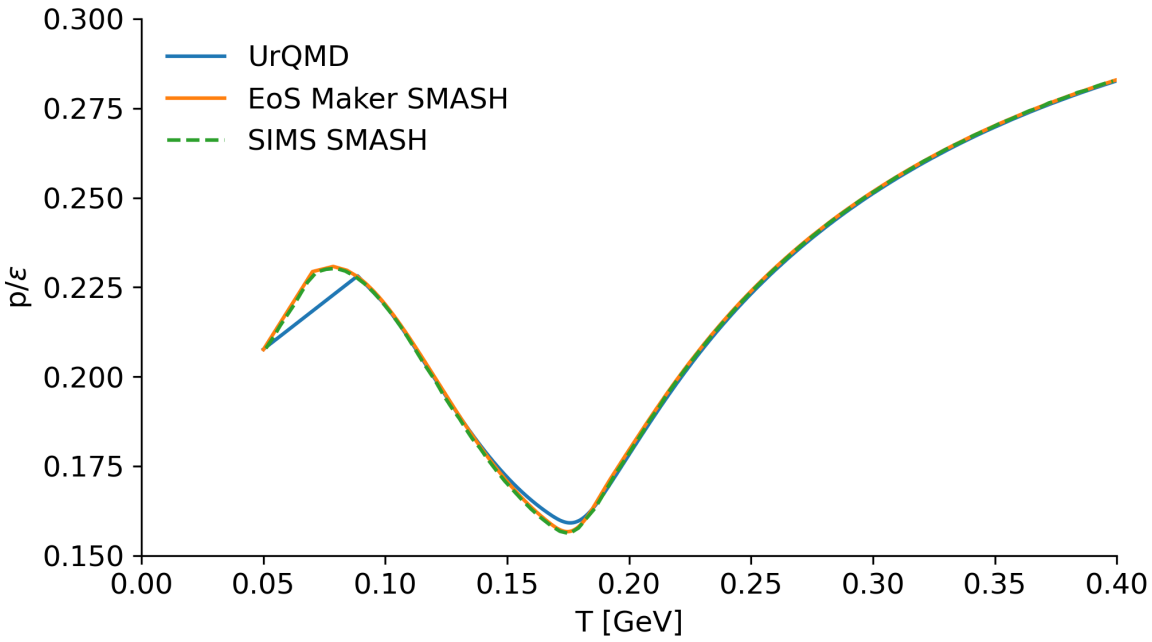


Figure A.1: Pressure to energy density ratio from the SMASH equation of state produced by `mrhheffernan/eos_maker` compared to that of SIMS and the EoS used by [28], labeled UrQMD.

are omitted here for conciseness. These tests found reproduction of physical quantities within statistical uncertainties.

In Fig. A.2, the result is in statistical agreement with that of SIMS for the majority of points and differences are consistent with expectations. The ratio between the new workflow and SIMS show that these results are consistent within statistical error. This confirms the reproduction of the geometry.

The results for $\langle p_T \rangle$ are sufficiently close that plotting the SIMS calculation and my result on the same plot is nearly unproductive. The ratio of all calculations to those of JETSCAPE-SIMS is shown in Fig. A.3. The maximal difference is ~2%, which is a feat in comparisons of statistical codes. For kaons and pions, the ratio is largely statistically consistent with 1 and does not differ much more than 0.5% with the reference calculation.

In Fig. A.4, the $\pi^0$ spectra from a fixed seed test is shown. In this case, the entire evolution prior to the freezeout (or particlization) surface is identical due to the fixing of the seed in TRENTo (see Fig. A.5 for a close-up of the energy density comparison between particlization surfaces). To ensure that the final read-in and Cooper-Frye

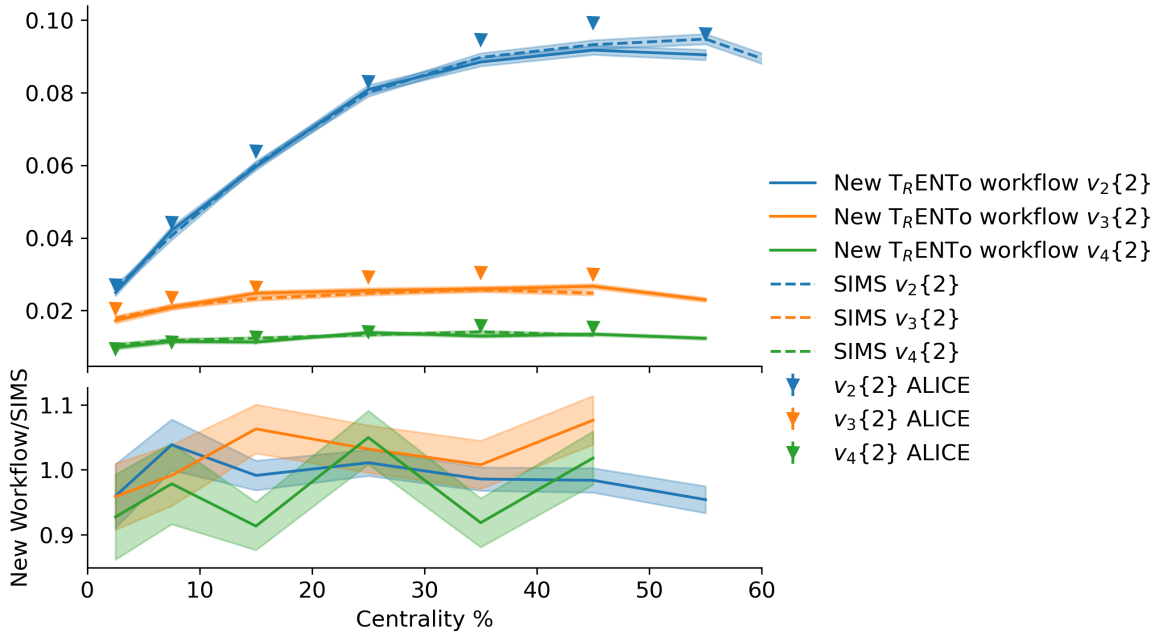


Figure A.2: New workflow $v_n\{2\}$ compared to that of JETSCAPE-SIMS and data measured by ALICE, ratios between the new workflow and SIMS are shown in the bottom panel.

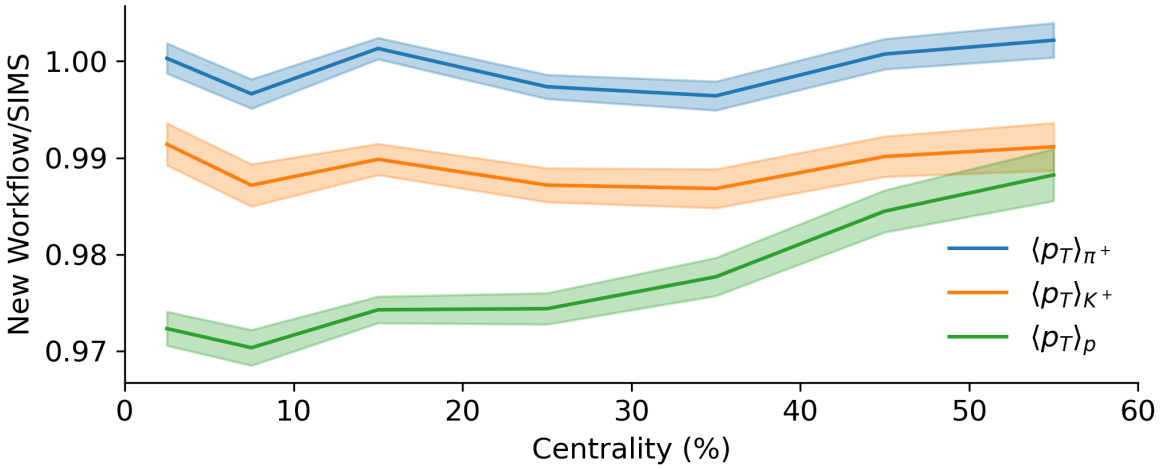


Figure A.3: Average transverse momentum compared to the work of JETSCAPE-SIMS. Note that this scale has a maximum difference less than +1/-2%.

sampling is identical, 57 samplings of the surface are taken and individual particle spectra can be determined. As can be seen, the sampled spectra are identical within statistical fluctuations with shaded bars denoting standard error in each transverse momentum bin.

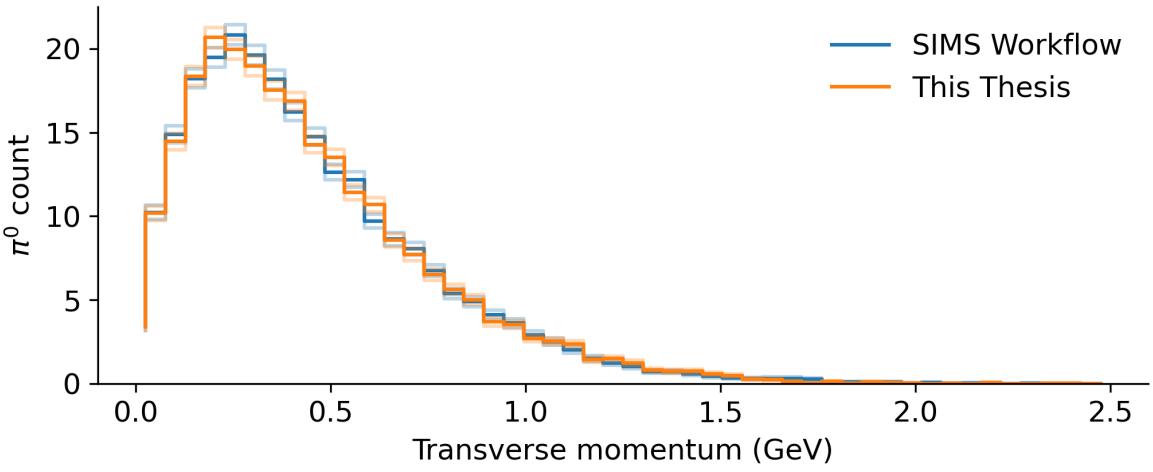


Figure A.4: $\pi^0$ spectra in a single event fixed seed comparison between the JETSCAPE SIMS workflow and the workflow used in this thesis.

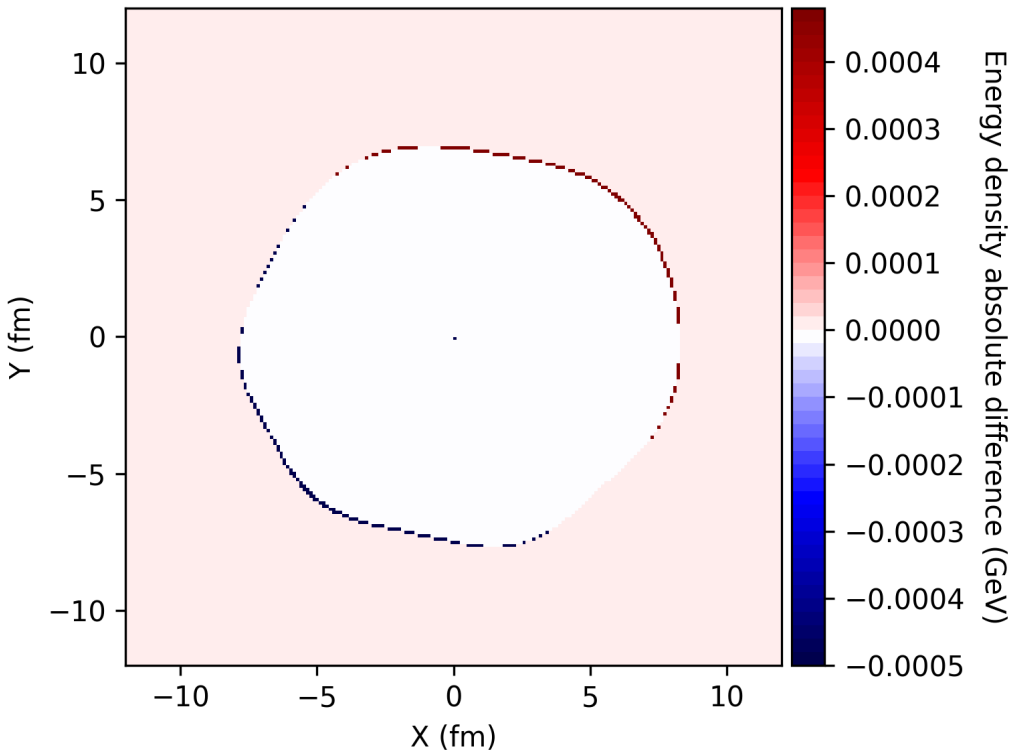


Figure A.5: Absolute difference in energy density in a single event fixed seed comparison between the JETSCAPE SIMS workflow and the workflow used in this thesis, albeit with a T$_R$ENTo + freestreaming initial state. The difference is large on edges due to numerical effects from interpolation to make this comparison. In the body of the comparison, the differences are substantially less than 0.1 MeV.

The workflow successfully reproduces the predictions of JETSCAPE-SIMS up to differences consistent with statistical fluctuations and reproduction of large scale Monte Carlo simulations. This successfully validates the model and ensures it can be used in comparisons of other observables not reported by that analysis. Additional extensive comparisons were made, but are not shown here for sake of brevity and clarity of discussion. In each comparison made, the differences between available calculations from [45] and those of the T$_R$ENTo workflow underconsideration here were immaterial. This comparison ensures that differences in calculations are isolated to the models in question and are not impacted by details such as the equation of state, specifics of the particlization or hadronic cascade, or parametrization. It was also useful in developing a workflow as it allowed for quick identification of software issues at the boundary between stages, which were readily resolved. The validated model is ready for comparison to other model calculations (see Ch. 4).

# OPEN SOURCE PROGRAMMING LIBRARIES

The computational models used in this thesis are research works and have been cited appropriately in the text. This work has further utilized a number of open-source programming libraries for the Python language. These are listed alphabetically below:

- Emukit [191]
- GPy [192]
- h5py [215]
- HSLuv [216]
- matplotlib [217]
- NumPy [218]
- pandas [219, 220]
- ptemcee [128]
- SALib [196, 197]
- scikit-learn [221]
- SciPy [222]
- seaborn [223]

# BIBLIOGRAPHY


[1] Matthew Heffernan. "How about that Bayes: Bayesian techniques and the simple pendulum." In: (Apr. 2021). arXiv: 2104.08621 [physics.ed-ph] (cit. on pp. xxiii, 67, 190).

[2] Matthew Heffernan. *Constraining the initial state through many-body observables*. https://indico.cern.ch/event/854124/contributions/4145022/ (cit. on pp. xxiii, 11, 35, 87, 90, 95).

[3] Matthew Heffernan, Charles Gale, Sangyong Jeon, and Jean-Francois Paquet. "Bayesian quantification of the Quark-Gluon Plasma: Improved design and closure demonstration." In: July 2022. arXiv: 2207.14751 [hep-ph] (cit. on pp. xxiii, 61, 62, 64, 138).

[4] Dananjaya Liyanage, Yi Ji, Derek Everett, Matthew Heffernan, Ulrich Heinz, Simon Mak, and Jean-Francois Paquet. "Efficient emulation of relativistic heavy ion collisions with transfer learning." In: *Phys. Rev. C* 105.3 (2022), p. 034910. DOI: 10.1103/PhysRevC.105.034910. arXiv: 2201.07302 [nucl-th] (cit. on pp. xxiv, 12, 58, 59, 111, 142).

[5] Jervis C. Deane. *The Trains Just as They Struck, Views of the Head End Collision at Crush, Texas*. 10 September 1896. URL: https://commons.wikimedia.org/wiki/File:The_Trains_Just_as_They_Struck,_Views_of_the_Head_End_Collision_at_Crush,_Texas_(cropped).jpg (cit. on p. 3).

[6] R.L. Workman et al. *Review of Particle Physics*. To be published in Prog. Theor. Exp. Phys. 2022, 083C01 (2022) (cit. on pp. 4, 9, 31, 112).

[7] Michael E Peskin and Daniel V Schroeder. *An introduction to quantum field theory*. Includes exercises. Boulder, CO: Westview, 1995 (cit. on p. 5).

[8] Claudia Ratti and Rene Bellwied. *The Deconfinement Transition of QCD: Theory Meets Experiment*. Vol. 981. Lecture Notes in Physics. June 2021. ISBN: 978-3-

030-67234-8, 978-3-030-67235-5. DOI: 10.1007/978-3-030-67235-5 (cit. on p. 5).

[9] V. Khachatryan et al. "Measurement of the inclusive 3-jet production differential cross section in proton–proton collisions at 7 TeV and determination of the strong coupling constant in the TeV range." In: *The European Physical Journal C* 75.5 (2015), p. 186 (cit. on p. 5).

[10] "The Frontiers of Nuclear Science, A Long Range Plan." In: (Sept. 2008). arXiv: 0809.3137 [nucl-ex] (cit. on p. 6).

[11] Jana N. Guenther. "Overview of the QCD phase diagram: Recent progress from the lattice." In: *Eur. Phys. J. A* 57.4 (2021), p. 136. DOI: 10.1140/epja/s10050-021-00354-6. arXiv: 2010.15503 [hep-lat] (cit. on pp. 6, 7, 156).

[12] Bjoern Schenke, Sangyong Jeon, and Charles Gale. "(3+1)D hydrodynamic simulation of relativistic heavy-ion collisions." In: *Phys. Rev.* C82 (2010), p. 014903 (cit. on pp. 6, 25).

[13] Charles Gale, Sangyong Jeon, and Bjoern Schenke. "Hydrodynamic Modeling of Heavy-Ion Collisions." In: *Int. J. Mod. Phys.* A28 (2013), p. 1340011 (cit. on p. 6).

[14] D. Fernandez-Fraile and A. Gomez Nicola. "Transport coefficients and resonances for a meson gas in Chiral Perturbation Theory." In: *Eur. Phys. J. C* 62 (2009). Ed. by Steffen Bass, Helen Caines, Manuel Calderon de la Barca Sanchez, Alessandro De Falco, Christian Kuhn, James Nagle, Marzia Nardi, Carlos Salgado, and Julia Velkovska, pp. 37–54. DOI: 10.1140/epjc/s10052-009-0935-0. arXiv: 0902.4829 [hep-ph] (cit. on p. 7).

[15] Antonio Dobado and Juan M. Torres-Rincon. "Bulk viscosity and the phase transition of the linear sigma model." In: *Phys. Rev. D* 86 (2012), p. 074021. DOI: 10.1103/PhysRevD.86.074021. arXiv: 1206.1261 [hep-ph] (cit. on p. 7).

[16] Rudy Marty, Elena Bratkovskaya, Wolfgang Cassing, Jörg Aichelin, and Hamza Berrehrah. "Transport coefficients from the Nambu-Jona-Lasinio model for $SU(3)_f$." In: *Phys. Rev. C* 88 (2013), p. 045204. DOI: 10.1103/PhysRevC.88.045204. arXiv: 1305.7180 [hep-ph] (cit. on p. 7).

[17] A. Puglisi, S. Plumari, and V. Greco. "Electric Conductivity from the solution of the Relativistic Boltzmann Equation." In: *Phys. Rev. D* 90 (2014), p. 114009. DOI: 10.1103/PhysRevD.90.114009. arXiv: 1408.7043 [hep-ph] (cit. on p. 7).

[18] Matthew Heffernan, Sangyong Jeon, and Charles Gale. "Hadronic transport coefficients from the linear σ model at finite temperature." In: *Phys. Rev. C* 102.3 (2020), p. 034906. DOI: 10.1103/PhysRevC.102.034906. arXiv: 2005.12793 [hep-ph] (cit. on pp. 7, 112, 118).

[19] Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz, and K. K. Szabo. "The Order of the quantum chromodynamics transition predicted by the standard model of particle physics." In: *Nature* 443 (2006), pp. 675–678. DOI: 10.1038/nature05120. arXiv: hep-lat/0611014 (cit. on p. 7).

[20] Owe Philipsen. "Lattice Constraints on the QCD Chiral Phase Transition at Finite Temperature and Baryon Density." In: *Symmetry* 13.11 (2021), p. 2079. DOI: 10.3390/sym13112079. arXiv: 2111.03590 [hep-lat] (cit. on p. 7).

[21] Tanmoy Bhattacharya et al. "QCD Phase Transition with Chiral Quarks and Physical Quark Masses." In: *Phys. Rev. Lett.* 113 (2014), p. 082001 (cit. on p. 7).

[22] Wit Busza, Krishna Rajagopal, and Wilke van der Schee. "Heavy Ion Collisions: The Big Picture, and the Big Questions." In: *Ann. Rev. Nucl. Part. Sci.* 68 (2018), pp. 339–376. DOI: 10.1146/annurev-nucl-101917-020852. arXiv: 1802.04801 [hep-ph] (cit. on pp. 7, 26).

[23] Björn Schenke. "The smallest fluid on Earth." In: *Rept. Prog. Phys.* 84.8 (2021), p. 082301. DOI: 10.1088/1361-6633/ac14c9. arXiv: 2102.11189 [nucl-th] (cit. on p. 7).

[24] Karen M. Burke et al. "Extracting the jet transport coefficient from jet quenching in high-energy heavy-ion collisions." In: *Phys. Rev. C* 90 (1 2014), p. 014909. DOI: 10.1103/PhysRevC.90.014909. URL: https://link.aps.org/doi/10.1103/PhysRevC.90.014909 (cit. on p. 7).

[25] Megan Connors, Christine Nattrass, Rosi Reed, and Sevil Salur. "Jet measurements in heavy ion physics." In: *Rev. Mod. Phys.* 90 (2018), p. 025005. DOI: 10.1103/RevModPhys.90.025005. arXiv: 1705.01974 [nucl-ex] (cit. on p. 7).

[26] S. Cao et al. "Determining the jet transport coefficient $\hat{q}$ from inclusive hadron suppression measurements using Bayesian parameter estimation." In: *Phys. Rev. C* 104.2 (2021), p. 024905. DOI: 10.1103/PhysRevC.104.024905. arXiv: 2102.11337 [nucl-th] (cit. on p. 7).

[27] Chiho Nonaka and Masayuki Asakawa. "Modeling a Realistic Dynamical Model for High Energy Heavy Ion Collisions." In: *PTEP* 2012 (2012), 01A208 (cit. on p. 8).

[28] Scott McDonald, Chun Shen, Francois Fillion-Gourdeau, Sangyong Jeon, and Charles Gale. "Hydrodynamic predictions for Pb+Pb collisions at 5.02 TeV." In: *Phys. Rev. C* 95.6 (2017), p. 064913. DOI: 10.1103/PhysRevC.95.064913. arXiv: 1609.02958 [hep-ph] (cit. on pp. 8, 15, 32, 85, 90, 93, 97, 99, 101, 106, 111, 116, 122, 184, 197, 198).

[29] Scott McDonald, Sangyong Jeon, and Charles Gale. "Exploring Longitudinal Observables with 3+1D IP-Glasma." In: *Nucl. Phys. A* 1005 (2021). Ed. by Feng Liu, Enke Wang, Xin-Nian Wang, Nu Xu, and Ben-Wei Zhang, p. 121771. DOI: 10.1016/j.nuclphysa.2020.121771. arXiv: 2001.08636 [nucl-th] (cit. on pp. 8, 85).

[30] Mayank Singh. "Characterizing the quark gluon plasma using soft thermal fluctuations and hard parton interactions." PhD thesis. McGill U., 2020 (cit. on pp. 8, 22, 25, 85).

[31] M Bleicher et al. "Relativistic hadron-hadron collisions in the ultra-relativistic quantum molecular dynamics model." In: *Journal of Physics G: Nuclear and Particle Physics* 25.9 (1999), p. 1859 (cit. on p. 8).

[32] J. Weil et al. "Particle production and equilibrium properties within a new hadron transport approach for heavy-ion collisions." In: *Phys. Rev. C* 94 (2016), p. 054905 (cit. on pp. 8, 31).

[33] Charles Gale, Jean-François Paquet, Björn Schenke, and Chun Shen. "Multi-messenger heavy-ion collision physics." In: *Phys. Rev. C* 105.1 (2022), p. 014909. DOI: 10.1103/PhysRevC.105.014909. arXiv: 2106.11216 [nucl-th] (cit. on pp. 9, 123).

[34] Hannah Petersen, Christopher Coleman-Smith, Steffen A. Bass, and Robert Wolpert. "Constraining the initial state granularity with bulk observables in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV." In: *J. Phys. G* 38 (2011), p. 045102. DOI: 10.1088/0954-3899/38/4/045102. arXiv: 1012.4629 [nucl-th] (cit. on p. 10).

[35] John Novak, Kevin Novak, Scott Pratt, Joshua Vredevoogd, Chris Coleman-Smith, and Robert Wolpert. "Determining Fundamental Properties of Matter Created in Ultrarelativistic Heavy-Ion Collisions." In: *Phys. Rev. C* 89.3 (2014), p. 034917. DOI: 10.1103/PhysRevC.89.034917. arXiv: 1303.5769 [nucl-th] (cit. on p. 10).

[36] Scott Pratt, Evan Sangaline, Paul Sorensen, and Hui Wang. "Constraining the Equation of State of Superhadronic Matter from Heavy-Ion Collisions." In: *Phys. Rev. Lett.* 114 (20 2015), p. 202301. DOI: 10.1103/PhysRevLett.114.202301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.114.202301 (cit. on pp. 10, 26).

[37] J. Scott Moreland, Jonah E. Bernhard, and Steffen A. Bass. "Alternative ansatz to wounded nucleon and binary collision scaling in high-energy nuclear collisions." In: *Phys. Rev.* C92.1 (2015), p. 011901. DOI: 10.1103/PhysRevC.92.011901. arXiv: 1412.4708 [nucl-th] (cit. on pp. 10, 11, 86, 88).

[38] Jonah E. Bernhard, Peter W. Marcy, Christopher E. Coleman-Smith, Snehalata Huzurbazar, Robert L. Wolpert, and Steffen A. Bass. "Quantifying properties of hot and dense QCD matter through systematic model-to-data comparison." In: *Phys. Rev. C* 91.5 (2015), p. 054910. DOI: 10.1103/PhysRevC.91.054910. arXiv: 1502.00339 [nucl-th] (cit. on p. 10).

[39] Jonah E. Bernhard, J. Scott Moreland, Steffen A. Bass, Jia Liu, and Ulrich Heinz. "Applying Bayesian parameter estimation to relativistic heavy-ion collisions: simultaneous characterization of the initial state and quark-gluon plasma medium." In: *Phys. Rev. C* 94.2 (2016), p. 024907. DOI: 10.1103/PhysRevC.94.024907. arXiv: 1605.03954 [nucl-th] (cit. on p. 10).

[40] Jonah E. Bernhard, J. Scott Moreland, and Steffen A. Bass. "Characterization of the initial state and QGP medium from a combined Bayesian analysis of LHC data at 2.76 and 5.02 TeV." In: *Nucl. Phys. A* 967 (2017). Ed. by Ulrich Heinz,

Olga Evdokimov, and Peter Jacobs, pp. 293–296. DOI: 10.1016/j.nuclphysa.2017.05.037. arXiv: 1704.04462 [nucl-th] (cit. on p. 10).

[41] Jonah E. Bernhard, J. Scott Moreland, and Steffen A. Bass. "Bayesian estimation of the specific shear and bulk viscosity of quark–gluon plasma." In: *Nature Phys.* 15.11 (2019), pp. 1113–1117. DOI: 10.1038/s41567-019-0611-8 (cit. on p. 10).

[42] Jonah E. Bernhard. "Bayesian parameter estimation for relativistic heavy-ion collisions." PhD thesis. Duke U., 2018-04-19. arXiv: 1804.06469 [nucl-th] (cit. on pp. 10, 11, 89, 95–98, 116, 154).

[43] J. Scott Moreland, Jonah E. Bernhard, and Steffen A. Bass. "Estimating nucleon substructure properties in a unified model of p-Pb and Pb-Pb collisions." In: *Nucl. Phys.* A982 (2019), pp. 503–506. DOI: 10.1016/j.nuclphysa.2018.07.011. arXiv: 1806.04802 [nucl-th] (cit. on p. 10).

[44] J. Scott Moreland, Jonah E. Bernhard, and Steffen A. Bass. "Bayesian calibration of a hybrid nuclear collision model using p-Pb and Pb-Pb data at energies available at the CERN Large Hadron Collider." In: *Phys. Rev. C* 101.2 (2020), p. 024911. DOI: 10.1103/PhysRevC.101.024911. arXiv: 1808.02106 [nucl-th] (cit. on p. 10).

[45] D. Everett et al. "Multisystem Bayesian constraints on the transport coefficients of QCD matter." In: *Phys. Rev. C* 103.5 (2021), p. 054904. DOI: 10.1103/PhysRevC.103.054904. arXiv: 2011.01430 [hep-ph] (cit. on pp. 10, 26, 30, 31, 33, 89–91, 93, 96–99, 110, 112, 116, 118, 123, 128, 135, 143, 154, 156, 168, 174, 184, 197, 201).

[46] D. Everett et al. "Phenomenological constraints on the transport properties of QCD matter with data-driven model averaging." In: *Phys. Rev. Lett.* 126.24 (2021), p. 242301. DOI: 10.1103/PhysRevLett.126.242301. arXiv: 2010.03928 [hep-ph] (cit. on pp. 10, 31, 33, 90, 96, 99, 110, 118, 143, 154, 168, 170, 173).

[47] Derek Everett. "Quantifying the Quark Gluon Plasma." In: (July 2021). arXiv: 2107.11362 [hep-ph] (cit. on pp. 10, 48, 117, 161, 170, 173).

[48] J. Mulligan et al. "Determining the jet transport coefficient $\hat{q}$ of the quark-gluon plasma using Bayesian parameter estimation." In: *55th Rencontres de Moriond on QCD and High Energy Interactions*. June 2021. arXiv: 2106.11348 [nucl-th] (cit. on p. 10).

[49] Kianusch Vahid Yousefnia, Atharva Kotibhaskar, Rajeev Bhalerao, and Jean-Yves Ollitrault. "Bayesian approach to long-range correlations and multiplicity fluctuations in nucleus-nucleus collisions." In: *Phys. Rev. C* 105.1 (2022), p. 014907. DOI: 10.1103/PhysRevC.105.014907. arXiv: 2108.03471 [nucl-th] (cit. on p. 10).

[50] Govert Nijs, Wilke van der Schee, Umut Gürsoy, and Raimond Snellings. "Transverse Momentum Differential Global Analysis of Heavy-Ion Collisions." In: *Phys. Rev. Lett.* 126.20 (2021), p. 202301. DOI: 10.1103/PhysRevLett.126.202301. arXiv: 2010.15130 [nucl-th] (cit. on pp. 10, 97, 154, 167, 185).

[51] Govert Nijs, Wilke van der Schee, Umut Gürsoy, and Raimond Snellings. "Bayesian analysis of heavy ion collisions with the heavy ion computational framework Trajectum." In: *Phys. Rev. C* 103.5 (2021), p. 054909. DOI: 10.1103/PhysRevC.103.054909. arXiv: 2010.15134 [nucl-th] (cit. on pp. 10, 96, 97, 116, 154, 167, 185).

[52] Heikki Mäntysaari, Björn Schenke, Chun Shen, and Wenbin Zhao. "Bayesian inference of the fluctuating proton shape." In: (Feb. 2022). arXiv: 2202.01998 [hep-ph] (cit. on pp. 10, 156).

[53] J. E. Parkkila, A. Onnerstad, F. Taghavi, C. Mordasini, A. Bilandzic, and D. J. Kim. "New constraints for QCD matter from improved Bayesian parameter estimation in heavy-ion collisions at LHC." In: (Nov. 2021). arXiv: 2111.08145 [hep-ph] (cit. on pp. 10, 11, 35, 167).

[54] Björn Schenke, Prithwish Tribedy, and Raju Venugopalan. "Event-by-event gluon multiplicity, energy density, and eccentricities in ultrarelativistic heavy-ion collisions." In: *Phys. Rev. C* 86 (3 2012), p. 034908. DOI: 10.1103/PhysRevC.86.034908. URL: https://link.aps.org/doi/10.1103/PhysRevC.86.034908 (cit. on pp. 11, 15, 18, 86).

[55] Eite Tiesinga, Peter Mohr, David Newell, and Barry Taylor. *2018 CODATA recommended values of the fundamental constants of physics and chemistry*. 2019. URL: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=928211 (cit. on p. 13).

[56] Björn Schenke, Chun Shen, and Prithwish Tribedy. "Features of the IP-Glasma." In: *Nucl. Phys. A* 982 (2019). Ed. by Federico Antinori, Andrea Dainese, Paolo Giubellino, Vincenzo Greco, Maria Paola Lombardo, and Enrico Scomparin, pp. 435–438. DOI: 10.1016/j.nuclphysa.2018.08.015. arXiv: 1807.05205 [nucl-th] (cit. on pp. 15, 97, 98).

[57] Scott McDonald. "Towards fully 3-dimensional simulations of heavy ion collisions in the IP-Glasma initial state framework." PhD thesis. McGill U., 2020 (cit. on pp. 15, 18, 20, 21, 85).

[58] Henri Kowalski and Derek Teaney. "Impact parameter dipole saturation model." In: *Phys. Rev. D* 68 (11 2003), p. 114005. DOI: 10.1103/PhysRevD.68.114005. URL: https://link.aps.org/doi/10.1103/PhysRevD.68.114005 (cit. on p. 15).

[59] Alex Krasnitz and Raju Venugopalan. "Non-perturbative computation of gluon mini-jet production in nuclear collisions at very high energies." In: *Nuclear Physics B* 557.1 (1999), pp. 237–270. ISSN: 0550-3213. DOI: https://doi.org/10.1016/S0550-3213(99)00366-1. URL: https://www.sciencedirect.com/science/article/pii/S0550321399003661 (cit. on pp. 15, 16).

[60] Alex Krasnitz and Raju Venugopalan. "Initial Energy Density of Gluons Produced in Very-High-Energy Nuclear Collisions." In: *Phys. Rev. Lett.* 84 (19 2000), pp. 4309–4312. DOI: 10.1103/PhysRevLett.84.4309. URL: https://link.aps.org/doi/10.1103/PhysRevLett.84.4309 (cit. on p. 15).

[61] Alex Krasnitz, Yasushi Nara, and Raju Venugopalan. "Coherent Gluon Production in Very-High-Energy Heavy-Ion Collisions." In: *Phys. Rev. Lett.* 87 (19 2001), p. 192302. DOI: 10.1103/PhysRevLett.87.192302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.192302 (cit. on p. 15).

[62] Alex Krasnitz and Raju Venugopalan. "Initial Gluon Multiplicity in Heavy-Ion Collisions." In: *Phys. Rev. Lett.* 86 (9 2001), pp. 1717–1720. DOI: 10.1103/

PhysRevLett.86.1717. URL: https://link.aps.org/doi/10.1103/PhysRevLett.86.1717 (cit. on p. 15).

[63] Vardan Khachatryan et al. "Transverse-momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s} = 7$ TeV." In: *Phys. Rev. Lett.* 105 (2010), p. 022002. DOI: 10.1103/PhysRevLett.105.022002. arXiv: 1005.3299 [hep-ex] (cit. on p. 17).

[64] Ringaile Placakyte. "Parton Distribution Functions." In: *31st International Symposium on Physics In Collision*. Nov. 2011. arXiv: 1111.5452 [hep-ph] (cit. on p. 17).

[65] D.J. Griffiths. *Introduction to Electrodynamics*. v. 2. Cambridge University Press, 2017. ISBN: 9781108420419. URL: https://books.google.ca/books?id=ndAoDwAAQBAJ (cit. on p. 18).

[66] Larry D. McLerran and Raju Venugopalan. "Computing quark and gluon distribution functions for very large nuclei." In: *Phys. Rev. D* 49 (1994), pp. 2233–2241. DOI: 10.1103/PhysRevD.49.2233. arXiv: hep-ph/9309289 (cit. on p. 18).

[67] Larry D. McLerran and Raju Venugopalan. "Gluon distribution functions for very large nuclei at small transverse momentum." In: *Phys. Rev. D* 49 (1994), pp. 3352–3355. DOI: 10.1103/PhysRevD.49.3352. arXiv: hep-ph/9311205 (cit. on p. 18).

[68] Larry D. McLerran and Raju Venugopalan. "Green's functions in the color field of a large nucleus." In: *Phys. Rev. D* 50 (1994), pp. 2225–2233. DOI: 10.1103/PhysRevD.50.2225. arXiv: hep-ph/9402335 (cit. on p. 18).

[69] Scott McDonald. *The initial State of heavy ion collisions in the IP-glasma framework*. Masters Thesis, McGill University. 2017 (cit. on pp. 18, 116, 117).

[70] T. Lappi. "Wilson line correlator in the MV model: Relating the glasma to deep inelastic scattering." In: *Eur. Phys. J. C* 55 (2008), pp. 285–292. DOI: 10.1140/epjc/s10052-008-0588-4. arXiv: 0711.3039 [hep-ph] (cit. on p. 19).

[71] M. Alvioli, H.J. Drescher, and M. Strikman. "A Monte Carlo generator of nucleon configurations in complex nuclei including nucleon–nucleon correlations." In: *Physics Letters B* 680.3 (2009), pp. 225–230. ISSN: 0370-2693. DOI:

https://doi.org/10.1016/j.physletb.2009.08.067. URL: https://www.sciencedirect.com/science/article/pii/S0370269309010442 (cit. on p. 20).

[72] William A. Hiscock and Lee Lindblom. "Generic instabilities in first-order dissipative relativistic fluid theories." In: *Phys. Rev. D* 31 (4 1985), pp. 725–733. DOI: 10.1103/PhysRevD.31.725. URL: https://link.aps.org/doi/10.1103/PhysRevD.31.725 (cit. on p. 23).

[73] Pavel Kovtun. "First-order relativistic hydrodynamics is stable." In: *JHEP* 10 (2019), p. 034. DOI: 10.1007/JHEP10(2019)034. arXiv: 1907.08191 [hep-th] (cit. on p. 24).

[74] Arpan Das, Wojciech Florkowski, and Radoslaw Ryblewski. "Correspondence between Israel-Stewart and first-order casual and stable hydrodynamics for the boost-invariant massive case with zero baryon density." In: *Phys. Rev. D* 102 (3 2020), p. 031501. DOI: 10.1103/PhysRevD.102.031501. URL: https://link.aps.org/doi/10.1103/PhysRevD.102.031501 (cit. on p. 24).

[75] Arpan Das, Wojciech Florkowski, and Radoslaw Ryblewski. "Correspondence between Israel-Stewart and first-order causal and stable hydrodynamics for Bjorken-expanding baryon-rich systems with vanishing particle masses." In: *Phys. Rev. D* 103.1 (2021), p. 014011. DOI: 10.1103/PhysRevD.103.014011. arXiv: 2010.02596 [nucl-th] (cit. on p. 24).

[76] Fabio S. Bemfica, Marcelo M. Disconzi, and Jorge Noronha. "First-Order General-Relativistic Viscous Fluid Dynamics." In: *Phys. Rev. X* 12.2 (2022), p. 021044. DOI: 10.1103/PhysRevX.12.021044. arXiv: 2009.11388 [gr-qc] (cit. on p. 24).

[77] Ingo Müller. "Zum Paradoxon der Wärmeleitungstheorie." In: *Zeitschrift für Physik* 198.4 (1967), pp. 329–344. ISSN: 0044-3328. DOI: 10.1007/BF01326412. URL: https://doi.org/10.1007/BF01326412 (cit. on p. 24).

[78] W. Israel and J.M. Stewart. "Transient relativistic thermodynamics and kinetic theory." In: *Annals of Physics* 118.2 (1979), pp. 341–372. ISSN: 0003-4916. DOI: https://doi.org/10.1016/0003-4916(79)90130-1. URL: https://www.sciencedirect.com/science/article/pii/0003491679901301 (cit. on pp. 24, 29).

[79] Alina Czajka and Sangyong Jeon. "Kubo formulas for the shear and bulk viscosity relaxation times and the scalar field theory shear $\tau_\pi$ calculation." In: *Phys. Rev. C* 95.6 (2017), p. 064906. DOI: 10.1103/PhysRevC.95.064906. arXiv: 1701.07580 [nucl-th] (cit. on p. 24).

[80] Christopher Plumberg, Dekrayat Almaalol, Travis Dore, Jorge Noronha, and Jacquelyn Noronha-Hostler. "Causality violations in realistic simulations of heavy-ion collisions." In: (Mar. 2021). DOI: 10.1103/PhysRevC.105.L061901. arXiv: 2103.15889 [nucl-th] (cit. on p. 24).

[81] Jorge Noronha, Michał Spaliński, and Enrico Speranza. "Transient Relativistic Fluid Dynamics in a General Hydrodynamic Frame." In: *Phys. Rev. Lett.* 128.25 (2022), p. 252302. DOI: 10.1103/PhysRevLett.128.252302. arXiv: 2105.01034 [nucl-th] (cit. on p. 24).

[82] G. S. Denicol, H. Niemi, E. Molnar, and D. H. Rischke. "Derivation of transient relativistic fluid dynamics from the Boltzmann equation." In: *Phys. Rev. D* 85 (2012). [Erratum: Phys.Rev.D 91, 039902 (2015)], p. 114047. DOI: 10.1103/PhysRevD.85.114047. arXiv: 1202.4551 [nucl-th] (cit. on pp. 24, 29).

[83] G. S. Denicol, S. Jeon, and C. Gale. "Transport coefficients of bulk viscous pressure in the 14-moment approximation." In: *Phys. Rev. C* 90 (2 2014), p. 024912. DOI: 10.1103/PhysRevC.90.024912. URL: https://link.aps.org/doi/10.1103/PhysRevC.90.024912 (cit. on p. 25).

[84] A. Bazavov et al. "Equation of state in $(2+1)$-flavor QCD." In: *Phys. Rev. D* 90 (9 2014), p. 094503. DOI: 10.1103/PhysRevD.90.094503. URL: https://link.aps.org/doi/10.1103/PhysRevD.90.094503 (cit. on p. 26).

[85] J. Noronha-Hostler, P. Parotto, C. Ratti, and J. M. Stafford. "Lattice-based equation of state at finite baryon number, electric charge and strangeness chemical potentials." In: *Phys. Rev. C* 100.6 (2019), p. 064910. DOI: 10.1103/PhysRevC.100.064910. arXiv: 1902.06723 [hep-ph] (cit. on p. 26).

[86] Jacquelyn Noronha-Hostler. "Theory Summary at Strangeness in Quark Matter 2019." In: *Springer Proc. Phys.* 250 (2020). Ed. by Domenico Elia, Giuseppe E. Bruno, Pietro Colangelo, and Leonardo Cosmai, pp. 519–525. DOI: 10.1007/978-3-030-53448-6_81. arXiv: 1911.01328 [nucl-th] (cit. on p. 26).

[87] Lipei Du. "Hydrodynamic description of the baryon-charged quark-gluon plasma." PhD thesis. Ohio State U., 2021. arXiv: 2107.08368 [nucl-th] (cit. on p. 26).

[88] Xiaofeng Luo and Nu Xu. "Search for the QCD Critical Point with Fluctuations of Conserved Quantities in Relativistic Heavy-Ion Collisions at RHIC : An Overview." In: *Nucl. Sci. Tech.* 28.8 (2017), p. 112. DOI: 10.1007/s41365-017-0257-0. arXiv: 1701.02105 [nucl-ex] (cit. on p. 26).

[89] J. M. Karthein, D. Mroczek, A. R. Nava Acuna, J. Noronha-Hostler, P. Parotto, D. R. P. Price, and C. Ratti. "Strangeness-neutral equation of state for QCD with a critical point." In: *Eur. Phys. J. Plus* 136.6 (2021), p. 621. DOI: 10.1140/epjp/s13360-021-01615-5. arXiv: 2103.08146 [hep-ph] (cit. on p. 26).

[90] Travis Dore, Jacquelyn Noronha-Hostler, and Emma McLaughlin. "Far-from-equilibrium search for the QCD critical point." In: *Phys. Rev. D* 102.7 (2020), p. 074017. DOI: 10.1103/PhysRevD.102.074017. arXiv: 2007.15083 [nucl-th] (cit. on p. 26).

[91] Scott Pratt, Evan Sangaline, Paul Sorensen, and Hui Wang. "Constraining the Eq. of State of Super-Hadronic Matter from Heavy-Ion Collisions." In: *Phys. Rev. Lett.* 114 (2015), p. 202301. DOI: 10.1103/PhysRevLett.114.202301. arXiv: 1501.04042 [nucl-th] (cit. on p. 26).

[92] Agnieszka Malgorzata Sorensen. "Density Functional Equation of State and Its Application to the Phenomenology of Heavy-Ion Collisions." PhD thesis. UCLA, Los Angeles (main), UCLA, 2021. arXiv: 2109.08105 [nucl-th] (cit. on p. 26).

[93] D. Mroczek, M. Hjorth-Jensen, J. Noronha-Hostler, P. Parotto, C. Ratti, and R. Vilalta. "Mapping out the thermodynamic stability of a QCD equation of state with a critical point using active learning." In: (Mar. 2022). arXiv: 2203.13876 [nucl-th] (cit. on p. 26).

[94] A. Bazavov et al. "Equation of state in ( 2+1 )-flavor QCD." In: *Phys. Rev. D* 90 (2014), p. 094503. DOI: 10.1103/PhysRevD.90.094503. arXiv: 1407.6387 [hep-lat] (cit. on pp. 26, 197).

[95] URL: https://github.com/mrhheffernan/eos_maker/tree/smash/ (cit. on p. 26).

[96] Pasi Huovinen and Pter Petreczky. "QCD Equation of State and Hadron Resonance Gas." In: *Nucl. Phys. A* 837 (2010), pp. 26–53. DOI: 10.1016/j.nuclphysa.2010.02.015. arXiv: 0912.2541 [hep-ph] (cit. on p. 26).

[97] Paolo Parotto, Marcus Bluhm, Debora Mroczek, Marlene Nahrgang, Jacquelyn Noronha-Hostler, Krishna Rajagopal, Claudia Ratti, Thomas Schäfer, and Mikhail Stephanov. "QCD equation of state matched to lattice data and exhibiting a critical point singularity." In: *Phys. Rev. C* 101.3 (2020), p. 034901. DOI: 10.1103/PhysRevC.101.034901. arXiv: 1805.05249 [hep-ph] (cit. on p. 26).

[98] Claudia Ratti. "Lattice QCD and heavy ion collisions: a review of recent progress." In: *Rept. Prog. Phys.* 81.8 (2018), p. 084301. DOI: 10.1088/1361-6633/aabb97. arXiv: 1804.07810 [hep-lat] (cit. on p. 26).

[99] Paolo Alba, Wanda Alberico, Rene Bellwied, Marcus Bluhm, Valentina Mantovani Sarti, Marlene Nahrgang, and Claudia Ratti. "Freeze-out conditions from net-proton and net-charge fluctuations at RHIC." In: *Phys. Lett. B* 738 (2014), pp. 305–310. DOI: 10.1016/j.physletb.2014.09.052. arXiv: 1403.4903 [hep-ph] (cit. on p. 27).

[100] Jamie M. Karthein, Paolo Alba, Valentina Mantovani-Sarti, Jacquelyn Noronha-Hostler, Paolo Parotto, Israel Portillo-Vazquez, Volodymyr Vovchenko, Volker Koch, and Claudia Ratti. "Thermal-model-based characterization of heavy-ion-collision systems at chemical freeze-out." In: *EPJ Web Conf.* 259 (2022), p. 11010. DOI: 10.1051/epjconf/202225911010. arXiv: 2201.03645 [hep-ph] (cit. on pp. 27, 174).

[101] Fred Cooper and Graham Frye. "Single-particle distribution in the hydrodynamic and statistical thermodynamic models of multiparticle production." In: *Phys. Rev. D* 10 (1 1974), pp. 186–189. DOI: 10.1103/PhysRevD.10.186. URL: https://link.aps.org/doi/10.1103/PhysRevD.10.186 (cit. on p. 27).

[102] Mike McNelis, Derek Everett, and Ulrich Heinz. "Particlization in fluid dynamical simulations of heavy-ion collisions: The iS3D module." In: *Comput. Phys. Commun.* 258 (2021), p. 107604. DOI: 10.1016/j.cpc.2020.107604. arXiv: 1912.08271 [nucl-th] (cit. on pp. 27–30).

[103] Derek Everett, Chandrodoy Chattopadhyay, and Ulrich Heinz. "Maximum entropy kinetic matching conditions for heavy-ion collisions." In: *Phys. Rev. C* 103.6 (2021), p. 064902. DOI: 10.1103/PhysRevC.103.064902. arXiv: 2101.01130 [hep-ph] (cit. on p. 28).

[104] Harold Grad. "On the kinetic theory of rarefied gases." In: *Communications on Pure and Applied Mathematics* 2.4 (1949), pp. 331–407. DOI: https://doi.org/10.1002/cpa.3160020403. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.3160020403. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.3160020403 (cit. on pp. 28, 29).

[105] Sydney Chapman and Thomas George Cowling. *The mathematical theory of non-uniform gases: an account of the kinetic theory of viscosity, thermal conduction and diffusion in gases*. Cambridge university press, 1990 (cit. on p. 28).

[106] Kevin Dusling, Guy D. Moore, and Derek Teaney. "Radiative energy loss and v(2) spectra for viscous hydrodynamics." In: *Phys. Rev. C* 81 (2010), p. 034907. DOI: 10.1103/PhysRevC.81.034907. arXiv: 0909.0754 [nucl-th] (cit. on p. 29).

[107] Werner Israel. "Nonstationary irreversible thermodynamics: A causal relativistic theory." In: *Annals of Physics* 100.1 (1976), pp. 310–331. ISSN: 0003-4916. DOI: https://doi.org/10.1016/0003-4916(76)90064-6. URL: https://www.sciencedirect.com/science/article/pii/0003491676900646 (cit. on p. 29).

[108] Akihiko Monnai and Tetsufumi Hirano. "Effects of bulk viscosity at freeze-out." In: *Phys. Rev. C* 80 (5 2009), p. 054906. DOI: 10.1103/PhysRevC.80.054906. URL: https://link.aps.org/doi/10.1103/PhysRevC.80.054906 (cit. on p. 29).

[109] James L Anderson and HR Witting. "A relativistic relaxation-time model for the Boltzmann equation." In: *Physica* 74.3 (1974), pp. 466–488 (cit. on p. 30).

[110] Amaresh Jaiswal, Radoslaw Ryblewski, and Michael Strickland. "Transport coefficients for bulk viscous evolution in the relaxation-time approximation." In: *Phys. Rev. C* 90 (4 2014), p. 044908. DOI: 10.1103/PhysRevC.90.044908. URL: https://link.aps.org/doi/10.1103/PhysRevC.90.044908 (cit. on p. 30).

[111] Wojciech Broniowski, Francesco Giacosa, and Viktor Begun. "Cancellation of the σ meson in thermal models." In: *Phys. Rev. C* 92.3 (2015), p. 034905. DOI:

10.1103/PhysRevC.92.034905. arXiv: 1506.01260 [nucl-th] (cit. on pp. 30, 112).

[112] B. Abelev et al. "Centrality determination of Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV with ALICE." In: *Phys. Rev. C* 88 (4 2013), p. 044909. DOI: 10.1103/PhysRevC.88.044909. URL: https://link.aps.org/doi/10.1103/PhysRevC.88.044909 (cit. on p. 32).

[113] Kenneth Aamodt et al. "Centrality dependence of the charged-particle multiplicity density at mid-rapidity in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV." In: *Phys. Rev. Lett.* 106 (2011), p. 032301. DOI: 10.1103/PhysRevLett.106.032301. arXiv: 1012.1657 [nucl-ex] (cit. on p. 34).

[114] Betty Abelev et al. "Centrality dependence of $\pi$, K, p production in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV." In: *Phys. Rev. C* 88 (2013), p. 044910. DOI: 10.1103/PhysRevC.88.044910. arXiv: 1303.0737 [hep-ex] (cit. on p. 34).

[115] Jaroslav Adam et al. "Measurement of transverse energy at midrapidity in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV." In: *Phys. Rev. C* 94.3 (2016), p. 034903. DOI: 10.1103/PhysRevC.94.034903. arXiv: 1603.04775 [nucl-ex] (cit. on p. 34).

[116] K. Aamodt et al. "Higher harmonic anisotropic flow measurements of charged particles in Pb-Pb collisions at $\sqrt{s_{NN}}$=2.76 TeV." In: *Phys. Rev. Lett.* 107 (2011), p. 032301. DOI: 10.1103/PhysRevLett.107.032301. arXiv: 1105.3865 [nucl-ex] (cit. on pp. 34, 184).

[117] Jaroslav Adam et al. "Anisotropic flow of charged particles in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV." In: *Phys. Rev. Lett.* 116.13 (2016), p. 132302. DOI: 10.1103/PhysRevLett.116.132302. arXiv: 1602.01119 [nucl-ex] (cit. on p. 34).

[118] Zhi Qiu and Ulrich Heinz. "Hydrodynamic event-plane correlations in Pb+Pb collisions at $\sqrt{s} = 2.76$ATeV." In: *Phys. Lett. B* 717 (2012), pp. 261–265. DOI: 10.1016/j.physletb.2012.09.030. arXiv: 1208.1200 [nucl-th] (cit. on pp. 35, 96, 98).

[119] Georges Aad et al. "Measurement of event-plane correlations in $\sqrt{s_{NN}} = 2.76$ TeV lead-lead collisions with the ATLAS detector." In: *Phys. Rev. C* 90.2 (2014), p. 024905. DOI: 10.1103/PhysRevC.90.024905. arXiv: 1403.0489 [hep-ex] (cit. on p. 35).

[120] Shreyasi Acharya et al. "Linear and non-linear flow modes in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV." In: *Phys. Lett. B* 773 (2017), pp. 68–80. DOI: 10.1016/j.physletb.2017.07.060. arXiv: 1705.04377 [nucl-ex] (cit. on pp. 35, 97, 98, 104, 106).

[121] Betty Bezverkhny Abelev et al. "Event-by-event mean $p_T$ fluctuations in pp and Pb-Pb collisions at the LHC." In: *Eur. Phys. J. C* 74.10 (2014), p. 3077. DOI: 10.1140/epjc/s10052-014-3077-y. arXiv: 1407.5530 [nucl-ex] (cit. on p. 35).

[122] Devinderjit Sivia and John Skilling. *Data analysis: a Bayesian tutorial*. OUP Oxford, 2006 (cit. on pp. 40, 42, 112).

[123] Roberto Trotta. "Bayes in the sky: Bayesian inference and model selection in cosmology." In: *Contemp. Phys.* 49 (2008), pp. 71–104. DOI: 10.1080/00107510802066753. arXiv: 0803.4089 [astro-ph] (cit. on p. 40).

[124] L Lyons. "A particle physicist's perspective on astrostatistics." In: *Statistical Challenges in Modern Astronomy IV*. Vol. 371. 2007, p. 361 (cit. on p. 40).

[125] Edwin T Jaynes. *Probability theory: The logic of science*. Cambridge university press, 2003 (cit. on pp. 41, 42).

[126] Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian data analysis*. CRC press, 2013 (cit. on pp. 41, 47, 54).

[127] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. "emcee: The MCMC Hammer." In: *Publications of the Astronomical Society of the Pacific* 125.925 (2013), pp. 306–312. DOI: 10.1086/670067. URL: https://doi.org/10.1086/670067 (cit. on p. 45).

[128] W. D. Vousden, W. M. Farr, and I. Mandel. "Dynamic temperature selection for parallel tempering in Markov chain Monte Carlo simulations." In: *Monthly Notices of the Royal Astronomical Society* 455.2 (Nov. 2015). https://github.com/willvousden/ptemcee, pp. 1919–1937. ISSN: 0035-8711. DOI: 10.1093/mnras/stv2422. URL: https://doi.org/10.1093/mnras/stv2422 (cit. on pp. 45, 48, 203).

[129] Andrew Gelman and Donald B Rubin. "Inference from iterative simulation using multiple sequences." In: *Statistical science* 7.4 (1992), pp. 457–472 (cit. on p. 47).

[130] Freeman Dyson. "A meeting with Enrico Fermi." In: *Nature* 427.6972 (2004), pp. 297–297 (cit. on p. 49).

[131] W. M. Thorburn. "The myth of Occam's razor." In: *Mind* XXVII.3 (Jan. 1918). https://academic.oup.com/mind/article-pdf/XXVII/3/345/9877777/345.pdf, pp. 345–353. ISSN: 0026-4423. DOI: 10.1093/mind/XXVII.3.345. URL: https://doi.org/10.1093/mind/XXVII.3.345 (cit. on p. 49).

[132] Roberto Trotta. "Bayes in the sky: Bayesian inference and model selection in cosmology." In: *Contemporary Physics* 49.2 (2008). https://doi.org/10.1080/00107510802066753, pp. 71–104. DOI: 10.1080/00107510802066753. URL: https://doi.org/10.1080/00107510802066753 (cit. on p. 49).

[133] Bradley Efron, Alan Gous, R. E. Kass, G. S. Datta, and P. Lahiri. "Scales of Evidence for Model Selection: Fisher versus Jeffreys." In: *Lecture Notes-Monograph Series* 38 (2001), pp. 208–256. ISSN: 07492170. URL: http://www.jstor.org/stable/4356166 (cit. on p. 49).

[134] Andre Perunicic. *How are principal component analysis and singular value decomposition related?* August 23, 2017. URL: https://intoli.com/blog/pca-and-svd/ (cit. on pp. 51, 53).

[135] thanks_in_advance. *Making sense of principal component analysis, eigenvectors & eigenvalues*. CrossValidated. 2021. URL: https://stats.stackexchange.com/questions/2691/making-sense-of-principal-component-analysis-eigenvectors-eigenvalues (cit. on p. 51).

[136] Oliver Deussen Jochen Görtler Rebecca Kehlbeck. *A Visual Exploration of Gaussian Processes*. 2019. URL: https://distill.pub/2019/visual-exploration-gaussian-processes/ (cit. on p. 54).

[137] Mengwu Guo. *A brief note on understanding neural networks as Gaussian processes*. 2021. DOI: 10.48550/ARXIV.2107.11892. URL: https://arxiv.org/abs/2107.11892 (cit. on p. 55).

[138] Sinno Jialin Pan and Qiang Yang. "A Survey on Transfer Learning." In: *IEEE Transactions on Knowledge and Data Engineering* 22.10 (2010), pp. 1345–1359. DOI: 10.1109/TKDE.2009.191 (cit. on p. 58).

[139] Marc C Kennedy and Anthony O'Hagan. "Predicting the output from a complex computer code when fast approximations are available." In: *Biometrika* 87.1 (2000), pp. 1–13 (cit. on p. 58).

[140] I.M. Sobol. "Uniformly distributed sequences with an additional uniform property." In: *USSR Computational Mathematics and Mathematical Physics* 16.5 (1976), pp. 236–242. ISSN: 0041-5553. DOI: https://doi.org/10.1016/0041-5553(76)90154-3. URL: https://www.sciencedirect.com/science/article/pii/0041555376901543 (cit. on pp. 59, 60).

[141] Sergei Kucherenko, Daniel Albrecht, and Andrea Saltelli. *Exploring multi-dimensional spaces: a Comparison of Latin Hypercube and Quasi Monte Carlo Sampling Techniques*. 2015. arXiv: 1505.02350 [stat.AP] (cit. on p. 59).

[142] Yong-Dao Zhou, Kai-Tai Fang, and Jian-Hui Ning. "Mixture discrepancy for quasi-random point sets." In: *Journal of Complexity* 29.3-4 (2013), pp. 283–301 (cit. on p. 60).

[143] V. Roshan Joseph, Evren Gul, and Shan Ba. "Maximum projection designs for computer experiments." In: *Biometrika* 102.2 (2015), pp. 371–380. ISSN: 00063444. URL: http://www.jstor.org/stable/43908541 (cit. on pp. 61, 62).

[144] Furong Sun, Robert B. Gramacy, Benjamin Haaland, Siyuan Lu, and Youngdeok Hwang. "Synthesizing simulation and field data of solar irradiance." In: *Statistical Analysis and Data Mining: The ASA Data Science Journal* 12.4 (2019), pp. 311–324. DOI: https://doi.org/10.1002/sam.11414. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sam.11414. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/sam.11414 (cit. on p. 64).

[145] F. M. S. Lima and P. Arun. "An accurate formula for the period of a simple pendulum oscillating beyond the small angle regime." In: *American Journal of Physics* 74.10 (2006). https://doi.org/10.1119/1.2215616, pp. 892–895. DOI: 10.1119/1.2215616. URL: https://doi.org/10.1119/1.2215616 (cit. on pp. 67, 68).

[146] Brian S. Blais. "Model Comparison in the Introductory Physics Laboratory." In: *The Physics Teacher* 58.3 (2020). https://doi.org/10.1119/1.5145420, pp. 209–213. DOI: 10.1119/1.5145420. URL: https://doi.org/10.1119/1.5145420 (cit. on p. 67).

[147] Douglas C Giancoli. *Physics: principles with applications*. Boston: Pearson, 2016 (cit. on p. 67).

[148] Stephen D. Schery. "Design of an inexpensive pendulum for study of large-angle motion." In: *American Journal of Physics* 44.7 (1976). https://doi.org/10.1119/1.10352, pp. 666–670. DOI: 10.1119/1.10352. URL: https://doi.org/10.1119/1.10352 (cit. on p. 68).

[149] Private communication with instructor. *McGill University Physics 101 Lab Manual*. 2020 (cit. on p. 70).

[150] Robert T. Wilkinson and Sue Allison. "Age and Simple Reaction Time: Decade Differences for 5,325 Subjects." In: *Journal of Gerontology* 44.2 (Mar. 1989), P29–P35. ISSN: 0022-1422. DOI: 10.1093/geronj/44.2.P29. URL: https://doi.org/10.1093/geronj/44.2.P29 (cit. on p. 70).

[151] Stan Development team. *Prior Choice Recommendations*. https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations. 2020. URL: https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations (cit. on p. 71).

[152] C. Hirt and W.E. Featherstone. "A 1.5km-resolution gravity field model of the Moon." In: *Earth and Planetary Science Letters* 329-330 (2012), pp. 22–30. ISSN: 0012-821X. DOI: https://doi.org/10.1016/j.epsl.2012.02.012. URL: https://www.sciencedirect.com/science/article/pii/S0012821X12000969 (cit. on p. 71).

[153] NASA. *By the Numbers: Jupiter*. https://solarsystem.nasa.gov/planets/jupiter/by-the-numbers/. 2021. URL: https://solarsystem.nasa.gov/planets/jupiter/by-the-numbers/ (cit. on p. 71).

[154] George EP Box. "Sampling and Bayes' inference in scientific modelling and robustness." In: *Journal of the Royal Statistical Society: Series A (General)* 143.4 (1980), pp. 383–404 (cit. on p. 72).

[155] Ambler Thompson and Barry N. Taylor. "Guide for the Use of the International System of Units (SI)." In: *NIST Special Publication* 811 (2008). https://physics.nist.gov/cuu/pdf/sp811.pdf. URL: https://physics.nist.gov/cuu/pdf/sp811.pdf (cit. on pp. 73, 76).

[156] Nikolaos K. Pavlis, Simon A. Holmes, Steve C. Kenyon, and John K. Factor. "The development and evaluation of the Earth Gravitational Model 2008 (EGM2008)." In: *Journal of Geophysical Research: Solid Earth* 117.B4 (2012). DOI: https://doi.org/10.1029/2011JB008916. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JB008916 (cit. on p. 76).

[157] Brady Neal. *On the Bias-Variance Tradeoff: Textbooks Need an Update*. 2019. arXiv: 1912.08286 [cs.LG] (cit. on p. 76).

[158] Michael L. Miller, Klaus Reygers, Stephen J. Sanders, and Peter Steinberg. "Glauber modeling in high energy nuclear collisions." In: *Ann. Rev. Nucl. Part. Sci.* 57 (2007), pp. 205–243. DOI: 10.1146/annurev.nucl.57.090506.123020. arXiv: nucl-ex/0701025 (cit. on p. 86).

[159] Zi-Wei Lin, Che Ming Ko, Bao-An Li, Bin Zhang, and Subrata Pal. "A Multi-phase transport model for relativistic heavy ion collisions." In: *Phys. Rev. C* 72 (2005), p. 064901. DOI: 10.1103/PhysRevC.72.064901. arXiv: nucl-th/0411110 (cit. on p. 86).

[160] H. Niemi, K. J. Eskola, and R. Paatelainen. "Event-by-event fluctuations in a perturbative QCD + saturation + hydrodynamics model: Determining QCD matter shear viscosity in ultrarelativistic heavy-ion collisions." In: *Phys. Rev. C* 93.2 (2016), p. 024907. DOI: 10.1103/PhysRevC.93.024907. arXiv: 1505.02677 [hep-ph] (cit. on pp. 86, 96, 98).

[161] Dmitri Kharzeev and Marzia Nardi. "Hadron production in nuclear collisions at RHIC and high density QCD." In: *Phys. Lett. B* 507 (2001), pp. 121–128. DOI: 10.1016/S0370-2693(01)00457-9. arXiv: nucl-th/0012025 (cit. on p. 86).

[162] Steffen Bass. *Hydrodynamic simulations of relativistic nuclear collisions with nucleon substructure: combined analysis of p+Pb and Pb+Pb collision systems at 5.02 TeV*. https://indico.bnl.gov/event/5391/contributions/29252/ (cit. on pp. 87, 95).

[163] Jia Liu, Chun Shen, and Ulrich Heinz. "Pre-equilibrium evolution effects on heavy-ion collision observables." In: *Phys. Rev. C* 91 (6 2015), p. 064906. DOI: 10.1103/PhysRevC.91.064906. URL: https://link.aps.org/doi/10.1103/PhysRevC.91.064906 (cit. on p. 89).

[164] Wojciech Broniowski, Wojciech Florkowski, Mikolaj Chojnacki, and Adam Kisiel. "Free-streaming approximation in early dynamics of relativistic heavy-ion collisions." In: *Phys. Rev. C* 80 (3 2009), p. 034902. DOI: 10.1103/PhysRevC.80.034902. URL: https://link.aps.org/doi/10.1103/PhysRevC.80.034902 (cit. on p. 89).

[165] Aleksi Kurkela, Aleksas Mazeliauskas, Jean-François Paquet, Sören Schlichting, and Derek Teaney. "Matching the Nonequilibrium Initial Stage of Heavy Ion Collisions to Hydrodynamics with QCD Kinetic Theory." In: *Phys. Rev. Lett.* 122.12 (2019), p. 122302. DOI: 10.1103/PhysRevLett.122.122302. arXiv: 1805.01604 [hep-ph] (cit. on p. 89).

[166] Dananjaya Liyanage, Derek Everett, Chandrodoy Chattopadhyay, and Ulrich Heinz. "Pre-hydrodynamic evolution and its impact on quark-gluon plasma signatures." In: (May 2022). arXiv: 2205.00964 [nucl-th] (cit. on p. 89).

[167] Bjoern Schenke. *Elucidating the properties of hot nuclear matter with a comprehensive description of ultra-relativistic heavy-ion collisions*. https://indico.cern.ch/event/656452/contributions/2869794/ (cit. on pp. 94, 98).

[168] Fernando G. Gardim, Jacquelyn Noronha-Hostler, Matthew Luzum, and Frédérique Grassi. "Effects of viscosity on the mapping of initial to final state in heavy ion collisions." In: *Phys. Rev. C* 91.3 (2015), p. 034902. DOI: 10.1103/PhysRevC.91.034902. arXiv: 1411.2574 [nucl-th] (cit. on p. 97).

[169] Jacquelyn Noronha-Hostler, Li Yan, Fernando G. Gardim, and Jean-Yves Ollitrault. "Linear and cubic response to the initial eccentricity in heavy-ion collisions." In: *Phys. Rev. C* 93.1 (2016), p. 014909. DOI: 10.1103/PhysRevC.93.014909. arXiv: 1511.03896 [nucl-th] (cit. on p. 97).

[170] Sangwook Ryu. "Integrated description of heavy ion collisions at RHIC and the LHC." PhD thesis. McGill U., 2016 (cit. on p. 98).

[171] S. Ryu, J. F. Paquet, C. Shen, G. S. Denicol, B. Schenke, S. Jeon, and C. Gale. "Importance of the Bulk Viscosity of QCD in Ultrarelativistic Heavy-Ion Collisions." In: *Phys. Rev. Lett.* 115.13 (2015), p. 132301. DOI: 10.1103/PhysRevLett.115.132301. arXiv: 1502.01675 [nucl-th] (cit. on p. 98).

[172] Georges Aad et al. "Measurement of event-plane correlations in $\sqrt{s_{NN}} = 2.76$ TeV lead-lead collisions with the ATLAS detector." In: *Phys. Rev. C* 90.2 (2014), p. 024905. DOI: 10.1103/PhysRevC.90.024905. arXiv: 1403.0489 [hep-ex] (cit. on pp. 100, 102).

[173] Murray Gell-Mann and M Levy. "The axial vector current in beta decay." In: *Nuovo Cim.* 16 (1960), p. 705 (cit. on p. 112).

[174] Jacquelyn Noronha-Hostler. "Implications of Missing Resonances in Heavy Ions Collisions." In: *Excited Hyperons in QCD Thermodynamics at Freeze-Out*. Dec. 2016, pp. 118–127. arXiv: 1612.07765 [nucl-th] (cit. on p. 112).

[175] Jamie M. Stafford, Paolo Alba, Rene Bellwied, Valentina Mantovani-Sarti, Jacquelyn Noronha-Hostler, Paolo Parotto, Israel Portillo-Vazquez, and Claudia Ratti. "Determination of Chemical Freeze-out Parameters from Net-kaon Fluctuations at RHIC." In: *Springer Proc. Phys.* 250 (2020). Ed. by Domenico Elia, Giuseppe E. Bruno, Pietro Colangelo, and Leonardo Cosmai, pp. 367–371. DOI: 10.1007/978-3-030-53448-6_58. arXiv: 1912.12968 [hep-ph] (cit. on p. 112).

[176] Jishnu Goswami, Frithjof Karsch, Christian Schmidt, Swagato Mukherjee, and Peter Petreczky. "Comparing conserved charge fluctuations from lattice QCD to HRG model calculations." In: *Acta Phys. Polon. Supp.* 14 (2021), p. 251. arXiv: 2011.02812 [hep-lat] (cit. on p. 112).

[177] P. Alba, V. Mantovani Sarti, J. Noronha-Hostler, P. Parotto, I. Portillo-Vazquez, C. Ratti, and J. M. Stafford. "Influence of hadronic resonances on the chemical freeze-out in heavy-ion collisions." In: *Phys. Rev. C* 101.5 (2020), p. 054905. DOI: 10.1103/PhysRevC.101.054905. arXiv: 2002.12395 [hep-ph] (cit. on p. 112).

[178] J. Salinas San Martin, J. Noronha-Hostler, H. Elfner, J. Hammelmann, and R. Hirayama. "Influence of heavy resonances in SMASH." In: July 2022. arXiv: 2207.09607 [hep-ph] (cit. on p. 112).

[179] Andrew Gelman and the Stan Development Team. *Prior choice recommendations*. 2020. URL: https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations (cit. on p. 113).

[180] A. Deur, V. Burkert, J.P. Chen, and W. Korsch. "Experimental determination of the effective strong coupling constant." In: *Physics Letters B* 650.4 (2007), pp. 244–248. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2007.05.015. URL: https://www.sciencedirect.com/science/article/pii/S037026930700562X (cit. on p. 116).

[181] Soeren Schlichting and Derek Teaney. "The First fm/c of Heavy-Ion Collisions." In: *Ann. Rev. Nucl. Part. Sci.* 69 (2019), pp. 447–476. DOI: 10.1146/annurev-nucl-101918-023825. arXiv: 1908.02113 [nucl-th] (cit. on p. 117).

[182] Olga Soloveva, David Fuseau, Jörg Aichelin, and Elena Bratkovskaya. "Shear viscosity and electric conductivity of a hot and dense QGP with a chiral phase transition." In: *Phys. Rev. C* 103.5 (2021), p. 054901. DOI: 10.1103/PhysRevC.103.054901. arXiv: 2011.03505 [nucl-th] (cit. on p. 118).

[183] Laszlo P. Csernai, Joseph I. Kapusta, and Larry D. McLerran. "Strongly Interacting Low-Viscosity Matter Created in Relativistic Nuclear Collisions." In: *Phys. Rev. Lett.* 97 (15 2006), p. 152303. DOI: 10.1103/PhysRevLett.97.152303. URL: https://link.aps.org/doi/10.1103/PhysRevLett.97.152303 (cit. on p. 118).

[184] Dmitri Kharzeev and Kirill Tuchin. "Bulk viscosity of QCD matter near the critical temperature." In: *Journal of High Energy Physics* 2008.09 (2008), pp. 093–093. DOI: 10.1088/1126-6708/2008/09/093. URL: https://doi.org/10.1088/1126-6708/2008/09/093 (cit. on p. 118).

[185] J-B Rose, J M Torres-Rincon, and H Elfner. "Inclusive and effective bulk viscosities in the hadron gas." In: *Journal of Physics G: Nuclear and Particle Physics* 48.1 (2020), p. 015005. DOI: 10.1088/1361-6471/abbc5c. URL: https://doi.org/10.1088/1361-6471/abbc5c (cit. on p. 118).

[186] Frithjof Karsch, Dmitri Kharzeev, and Kirill Tuchin. "Universal properties of bulk viscosity near the QCD phase transition." In: *Physics Letters B* 663.3 (2008), pp. 217–221. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.

2008.01.080. URL: https://www.sciencedirect.com/science/article/pii/S0370269308003997 (cit. on p. 118).

[187] Jacquelyn Noronha-Hostler, Jorge Noronha, and Carsten Greiner. "Transport Coefficients of Hadronic Matter Near $T_c$." In: *Phys. Rev. Lett.* 103 (17 2009), p. 172302. DOI: 10.1103/PhysRevLett.103.172302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.103.172302 (cit. on p. 118).

[188] Peter Arnold, and Guy D. Moore. "Bulk viscosity of high-temperature QCD." In: *Phys. Rev. D* 74 (2006), p. 085021 (cit. on p. 118).

[189] R. Marty, E. Bratkovskaya, W. Cassing, J. Aichelin, and H. Berrehrah. "Transport coefficients from the Nambu–Jona-Lasinio model for $SU(3)_f$." In: *Phys. Rev. C* 88 (4 2013), p. 045204. DOI: 10.1103/PhysRevC.88.045204. URL: https://link.aps.org/doi/10.1103/PhysRevC.88.045204 (cit. on p. 119).

[190] Sangwook Ryu, Jean-François Paquet, Chun Shen, Gabriel Denicol, Björn Schenke, Sangyong Jeon, and Charles Gale. "Effects of bulk viscosity and hadronic rescattering in heavy ion collisions at energies available at the BNL Relativistic Heavy Ion Collider and at the CERN Large Hadron Collider." In: *Phys. Rev. C* 97.3 (2018), p. 034910. DOI: 10.1103/PhysRevC.97.034910. arXiv: 1704.04216 [nucl-th] (cit. on p. 136).

[191] Andrei Paleyes, Mark Pullin, Maren Mahsereci, Neil Lawrence, and Javier González. "Emulation of physical processes with Emukit." In: *Second Workshop on Machine Learning and the Physical Sciences, NeurIPS*. Vol. 74. American Physical Society, 2019, p. 085021 (cit. on pp. 142, 203).

[192] GPy. *GPy: A Gaussian process framework in python*. http://github.com/SheffieldML/GPy. since 2012 (cit. on pp. 142, 203).

[193] Weiyao Ke. "Training emulators with transformed \supplemental parameters." Unpublished JETSCAPE note, private communication. 2019 (cit. on p. 143).

[194] Matthew Heffernan. *GitHub: mrhheffernan*. https://github.com/mrhheffernan, year = 2022. 2006 (cit. on p. 143).

[195] I.M Sobol. "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates." In: *Mathematics and Computers in Simulation* 55.1 (2001). The Second IMACS Seminar on Monte Carlo Methods, pp. 271–280. ISSN: 0378-4754. DOI: https://doi.org/10.1016/S0378-4754(00)00270-6. URL: https://www.sciencedirect.com/science/article/pii/S0378475400002706 (cit. on p. 145).

[196] Takuya Iwanaga, William Usher, and Jonathan Herman. "Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses." In: *Socio-Environmental Systems Modelling* 4 (May 2022), p. 18155. DOI: 10.18174/sesmo.18155. URL: https://sesmo.org/article/view/18155 (cit. on pp. 145, 203).

[197] Jon Herman and Will Usher. "SALib: An open-source Python library for Sensitivity Analysis." In: *The Journal of Open Source Software* 2.9 (2017), p. 085021. DOI: 10.21105/joss.00097. URL: https://doi.org/10.21105/joss.00097 (cit. on pp. 145, 203).

[198] Pavel Kovtun, Dam T. Son, and Andrei O. Starinets. "Holography and hydrodynamics: diffusion on stretched horizons." In: *Journal of High Energy Physics* 2003.10 (2003), p. 064 (cit. on p. 159).

[199] Fernando G. Gardim and Jean-Yves Ollitrault. "Effective shear and bulk viscosities of the quark-gluon plasma: QCD versus heavy-ion data." In: *Phys. Rev. D* 74 (July 2022), p. 085021. arXiv: 2207.08692 [nucl-th] (cit. on p. 160).

[200] Giuliano Giacalone, Fernando G. Gardim, Jacquelyn Noronha-Hostler, and Jean-Yves Ollitrault. "Correlation between mean transverse momentum and anisotropic flow in heavy-ion collisions." In: *Phys. Rev. C* 103.2 (2021), p. 024909. DOI: 10.1103/PhysRevC.103.024909. arXiv: 2004.01765 [nucl-th] (cit. on p. 167).

[201] Giuliano Giacalone, Björn Schenke, and Chun Shen. "Constraining the Nucleon Size with Relativistic Nuclear Collisions." In: *Phys. Rev. Lett.* 128.4 (2022), p. 042301. DOI: 10.1103/PhysRevLett.128.042301. arXiv: 2111.02908 [nucl-th] (cit. on pp. 167, 183).

[202] Tiago M Fragoso, Wesley Bertoli, and Francisco Louzada. "Bayesian model averaging: A systematic review and conceptual classification." In: *International Statistical Review* 86.1 (2018), pp. 1–28 (cit. on p. 170).

[203] S. Kullback and R. A. Leibler. "On Information and Sufficiency." In: *The Annals of Mathematical Statistics* 22.1 (1951), pp. 79 –86. DOI: 10.1214/aoms/1177729694. URL: https://doi.org/10.1214/aoms/1177729694 (cit. on p. 171).

[204] Jamie M. Karthein, Debora Mroczek, Angel R. Nava Acuna, Jacquelyn Noronha-Hostler, Paolo Parotto, Damien R. P. Price, and Claudia Ratti. "Strangeness neutral equation of state with a critical point." In: *International Conference on Critical Point and Onset of Deconfinement*. Vol. 74. American Physical Society, Oct. 2021, p. 085021. arXiv: 2110.00622 [nucl-th] (cit. on p. 174).

[205] Björn Schenke, Chun Shen, and Derek Teaney. "Transverse momentum fluctuations and their correlation with elliptic flow in nuclear collision." In: *Phys. Rev. C* 102.3 (2020), p. 034905. DOI: 10.1103/PhysRevC.102.034905. arXiv: 2004.00690 [nucl-th] (cit. on p. 177).

[206] E.G. Nielsen. "Studies of anisotropic flow with event-shape engineering and mean transverse momentum - flow correlations in Pb-Pb, Xe-Xe and pp collisions with ALICE." In: *The European PhysicalSociety Conference on High Energy Physics 2021 (EPSHEP2021)*. Vol. 74. American Physical Society, 2021, p. 085021. URL: \url{https://indico.desy.de/event/28202/contributions/105721/} (cit. on p. 183).

[207] Aleksi Kurkela, Aleksas Mazeliauskas, Jean-François Paquet, Sören Schlichting, and Derek Teaney. "Matching the Nonequilibrium Initial Stage of Heavy Ion Collisions to Hydrodynamics with QCD Kinetic Theory." In: *Phys. Rev. Lett.* 122.12 (2019), p. 122302. DOI: 10.1103/PhysRevLett.122.122302. arXiv: 1805.01604 [hep-ph] (cit. on p. 193).

[208] Aleksi Kurkela, Aleksas Mazeliauskas, Jean-François Paquet, Sören Schlichting, and Derek Teaney. "Effective kinetic description of event-by-event pre-equilibrium dynamics in high-energy heavy-ion collisions." In: *Phys. Rev. C* 99.3 (2019), p. 034910. DOI: 10.1103/PhysRevC.99.034910. arXiv: 1805.00961 [hep-ph] (cit. on p. 193).

[209] Michito Okai, Koji Kawaguchi, Yasuki Tachibana, and Tetsufumi Hirano. "New approach to initializing hydrodynamic fields and mini-jet propagation in quark-gluon fluids." In: *Phys. Rev. C* 95.5 (2017), p. 054914. DOI: 10.1103/PhysRevC.95.054914. arXiv: 1702.07541 [nucl-th] (cit. on p. 193).

[210] Chun Shen, Gabriel Denicol, Charles Gale, Sangyong Jeon, Akihiko Monnai, and Bjoern Schenke. "A hybrid approach to relativistic heavy-ion collisions at the RHIC BES energies." In: *Nucl. Phys. A* 967 (2017). Ed. by Ulrich Heinz, Olga Evdokimov, and Peter Jacobs, pp. 796–799. DOI: 10.1016/j.nuclphysa.2017.06.008. arXiv: 1704.04109 [nucl-th] (cit. on p. 193).

[211] Chun Shen and Björn Schenke. "Dynamical initial state model for relativistic heavy-ion collisions." In: *Phys. Rev. C* 97.2 (2018), p. 024907. DOI: 10.1103/PhysRevC.97.024907. arXiv: 1710.00881 [nucl-th] (cit. on p. 193).

[212] Yukinao Akamatsu, Masayuki Asakawa, Tetsufumi Hirano, Masakiyo Kitazawa, Kenji Morita, Koichi Murase, Yasushi Nara, Chiho Nonaka, and Akira Ohnishi. "Dynamically integrated transport approach for heavy-ion collisions at high baryon density." In: *Phys. Rev. C* 98.2 (2018), p. 024909. DOI: 10.1103/PhysRevC.98.024909. arXiv: 1805.09024 [nucl-th] (cit. on p. 193).

[213] Lipei Du, Ulrich Heinz, and Gojko Vujanovic. "Hybrid model with dynamical sources for heavy-ion collisions at BES energies." In: *Nucl. Phys. A* 982 (2019). Ed. by Federico Antinori, Andrea Dainese, Paolo Giubellino, Vincenzo Greco, Maria Paola Lombardo, and Enrico Scomparin, pp. 407–410. DOI: 10.1016/j.nuclphysa.2018.09.015. arXiv: 1807.04721 [nucl-th] (cit. on p. 193).

[214] Yuuka Kanakubo, Yasuki Tachibana, and Tetsufumi Hirano. "Unified description of hadron yield ratios from dynamical core-corona initialization." In: *Phys. Rev. C* 101.2 (2020), p. 024912. DOI: 10.1103/PhysRevC.101.024912. arXiv: 1910.10556 [nucl-th] (cit. on p. 193).

[215] Andrew Collette. *Python and HDF5*. Vol. 74. O'Reilly, 2013, p. 085021 (cit. on p. 203).

[216] Alexei Boronine et al. *HSluv*. https://github.com/hsluv/hsluv. 2006 (cit. on p. 203).

[217] J. D. Hunter. "Matplotlib: A 2D graphics environment." In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/MCSE.2007.55 (cit. on p. 203).

[218] Charles R. Harris et al. "Array programming with NumPy." In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2. URL: https://doi.org/10.1038/s41586-020-2649-2 (cit. on p. 203).

[219] The pandas development team. *pandas-dev/pandas: Pandas*. Version latest. Feb. 2020. DOI: 10.5281/zenodo.3509134. URL: https://doi.org/10.5281/zenodo.3509134 (cit. on p. 203).

[220] Wes McKinney. "Data Structures for Statistical Computing in Python." In: *Proceedings of the 9th Python in Science Conference*. Ed. by Stéfan van der Walt and Jarrod Millman. Vol. 74. American Physical Society, 2010, pp. 56 –61. DOI: 10.25080/Majora-92bf1922-00a (cit. on p. 203).

[221] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python." In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830 (cit. on p. 203).

[222] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python." In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2 (cit. on p. 203).

[223] Michael L. Waskom. "seaborn: statistical data visualization." In: *Journal of Open Source Software* 6.60 (2021), p. 3021. DOI: 10.21105/joss.03021. URL: https://doi.org/10.21105/joss.03021 (cit. on p. 203).

## COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/