DISSERTATION

submitted to the

Combined Faculty of Mathematics, Engineering, and Natural Sciences
of Heidelberg University, Germany

for the degree of

Doctor of Natural Sciences

Put forward by
Johannes Hölck
born in Mannheim, Germany

Oral examination: 23rd November 2022

Dynamical and non-equilibrium–statistical processes in relativistic heavy-ion collisions

Referees:
Prof. Georg Wolschin
Prof. Andreas Mielke

Now it would be very remarkable if any system existing in the real world would be exactly *represented by any simple model. However, cunningly chosen parsimonious models often do provide remarkably useful approximations. For example, the law* $PV = RT$ *relating pressure* P*, volume* V *and temperature* T *of an "ideal" gas via a constant* R *is not exactly true for any real gas, but it frequently provides a useful approximation and furthermore its structure is informative since it springs from a physical view of the behavior of gas molecules.*

For such a model there is no need to ask the question "Is the model true?". If "truth" is to be the "whole truth" the answer must be "No". The only question of interest is "Is the model illuminating and useful?".

— George Box (1979)

Zusammenfassung

Auf der Grundlage stochastischer Differenzialgleichungen werden nichtgleichgewichtsstatistische Modelle für das Abstoppen von Protonen und die Thermalisierung geladener Hadronen in relativistischen Schwerionenkollisionen entwickelt. Hierzu parametrisiere ich die Phasenraumtrajektorien der Teilchen mit einem Satz Koordinaten, welcher als stochastischer Drift-Diffusionsprozess behandelt wird und zu Fokker-Planck-Gleichungen für die Einteilchenverteilungsfunktionen führt. Zur Bestimmung der Form der Koeffizientenfunktionen leite ich Fluktuations-Dissipations-Beziehungen aus den zu erwartenden zeitasymptotischen Teilchenverteilungen ab. Die resultierenden Zeitentwicklungsgleichungen werden numerisch gelöst und mit experimentellen Daten vom Super Proton Synchrotron (SPS), Relativistic Heavy Ion Collider (RHIC) und Large Hadron Collider (LHC) verglichen.

Abstract

On the basis of stochastic differential equations, non-equilibrium–statistical models for the stopping of protons and the thermalization of charged hadrons in relativistic heavy-ion collisions are developed. I parameterize the particles' phase-space trajectories with a set of coordinates that is treated as a stochastic drift–diffusion process, leading to Fokker–Planck equations for the single-particle distribution functions. To deduce the form of the coefficient functions, I derive fluctuation–dissipation relations from the expected time-asymptotic particle distributions. The resulting time-evolution equations are numerically solved and compared with experimental data from the Super Proton Synchrotron (SPS), Relativistic Heavy Ion Collider (RHIC), and Large Hadron Collider (LHC).

Contents

Figures, Tables, and Acronyms

List of Figures

List of Tables

List of Acronyms

CGC	color-glass condensate	46
COM	center of momentum	33
DIS	deep inelastic scattering	43
FDR	fluctuation–dissipation relation	22
FPE	Fokker–Planck equation	20
FPO	Fokker–Planck operator	20
FWHM	full width at half maximum	61
GOF	goodness of fit	39
NDF	number density function	35
ODE	ordinary differential equation	12
PDE	partial differential equation	21
PDF	probability density function	7
PMF	probability mass function	7
QCD	quantum chromodynamics	2
SDE	stochastic differential equation	16
WLS	weighted least squares	38

Particle-physics research facilities and experiments

ALICE	A Large Ion Collider Experiment	57
ATLAS	A Toroidal LHC Apparatus	71
BRAHMS	Broad Range Hadron Magnetic Spectrometers Experiment	47
HERA	Hadron–Electron Ring Accelerator (*Hadron-Elektron-Ringanlage*)	53
LHC	Large Hadron Collider	3
NA49	North Area Experiment 49	47
RHIC	Relativistic Heavy Ion Collider	3
SPS	Super Proton Synchrotron	3

1. Introduction

In 1905, his *annus mirabilis*, Albert Einstein published a series of seminal articles that profoundly shaped the advancement of physics in the 20th century. Besides his famous works on the special theory of relativity[1,2] and the discrete nature of light,[3] which awarded him the 1921 Nobel Prize in Physics, he also provided the first theoretical description of Brownian motion based on statistical mechanics.[4] In the latter, he connected the microscopic particle movements with the macroscopic phenomenon of diffusion and proposed an empirical method to prove the atomic nature of matter.

This sparked lasting interest in probabilistic approaches to the motion of particles (Einstein 1906; Smoluchowski 1906, 1916; Langevin 1908; Uhlenbeck and Ornstein 1930), especially after Perrin (1909) experimentally verified Einstein's claims and thereby the existence of atoms, which was honored with the 1926 Nobel Prize in Physics. Simultaneously, progress was made towards a firm mathematical basis, with notable contributions by Markov (1906), Wiener (1923), Chapman (1928), and Kolmogorov (1931), leading to modern probability theory and the concept of stochastic processes. These objects – effectively time-dependent random variables – formalize the idea of microscopic probabilistic particle trajectories; their associated probability density functions represent the macroscopically observable particle concentrations, *i. e.*, the particle number densities.

Combining Einstein's theories of relativity and diffusion into a description of relativistic particle motion proved difficult, however, as the underlying mathematical structures were found to be irreconcilable (Łopuszański 1953; Dudley 1966; Hakim 1968). One way to overcome this conceptual issue is to move from a purely position-based description of diffusion to a more general formulation in phase space by incorporating the particle's momentum, as proposed by Debbasch, Mallick, and Rivet in 1997, and later, in a slightly different form, by Dunkel and Hänggi (2005a,b). A key advantage of this approach is that the theory remains compatible with Itô's stochastic calculus (1944), which provides a framework for constructing stochastic processes that can be tailored to specific requirements, particularly those of special relativity, by specifying two coefficient functions that

---

[1] "Zur Elektrodynamik bewegter Körper"

[2] "Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?"

[3] "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt"

[4] "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen"

encode the deterministic (“drift”) and probabilistic (“diffusion”) contributions to particle motion. The same two coefficient functions reappear in the time-evolution equation of the associated particle number density, which takes the form of a Fokker–Planck equation for any such drift–diffusion process, so that the determination of these functions is central to phase-space diffusion models of this type.

In particle physics, phenomenological models based on Fokker–Planck equations have been developed by Wolschin (1999) and Biyajima, Ide, Mizoguchi, and Suzuki (2002) to describe the number densities of different kinds of hadrons – subatomic particles comprised of two (“meson”) or three (“baryon”) valence quarks, the fundamental building blocks of matter, and held together by gluons, the carriers of the strong force – during and after the collision of two heavy atomic nuclei at relativistic velocities. Assuming very simple coefficient functions, a linear drift[5] combined with a constant diffusion, good descriptions of experimental net-proton and charged-hadron data in one effective dimension could be achieved. Due to their largely phenomenological nature, however, physical interpretation of the postulated coefficient functions was only possible to a limited extent, and an extension of the models to two or more effective dimensions is still pending.[6]

In this thesis, I attempt to improve the two aforementioned models by providing a firm theoretical basis built upon stochastic drift–diffusion processes in phase space: Starting from a given initial distribution, any particle trajectory is assumed to evolve in time towards an expected asymptotic equilibrium state. Since interactions in relativistic heavy-ion collisions effectively cease after a finite interaction timespan, the time evolution is prematurely halted, so that the system remains in a final non-equilibrium state. To determine the coefficient functions of the associated drift–diffusion processes, I use fluctuation–dissipation relations derived from the particles’ time-asymptotic distribution functions, which allows me to construct coefficients for the two models that are physically motivated from quantum chromodynamics (QCD) and relativistic thermodynamics, respectively.

With this schematic approach, both models can be straightforwardly formulated in two effective dimensions, providing the opportunity for comparison with experimentally measured distributions in two-dimensional transverse-momentum and (pseudo-)rapidity space, which, to the best of my knowledge, has not been done before. Moreover, the models can be easily expressed in different sets of coordinates, since stochastic calculus, and its extensions by Stratonovich (1964), Fisk (1965), Hänggi (1978), and Klimontovich (1990), provide simple transformation rules for all mathematical objects involved. As such,

---

[5] In a later publication (Forndran and Wolschin 2017), a hyperbolic drift function in longitudinal-rapidity space inspired by relativistic thermodynamics was also studied.

[6] It should be noted, however, that Wolschin considered diffusion in a two-dimensional energy space a few years earlier (1996) to describe data measured at the Super Proton Synchrotron (SPS) and Alternating Gradient Synchrotron (AGS).

existing symmetries can be readily exploited by changing coordinates, and model results adapted to the requirements of experimental data sets.

It should be noted that the resulting probabilistic phase-space trajectories, of course, do not provide an accurate representation of the real particle dynamics in relativistic heavy-ion collisions. While a certain degree of information on the fundamental microscopic physics enters via the fluctuation–dissipation relations, the relatively simple mathematical structure of the drift–diffusion process can only rudimentarily reflect complicated interactions. Consequently, the purpose of the two models consists in "finding the 'best' approximation of the 'exact' dynamics" (Dunkel and Hänggi 2009, p. 42) within the relativistic phase-space diffusion formalism, and using it to draw conclusions about the observed macroscopic behavior of the system.

In the subsequent chapter 2, I first give a brief review of the required probability-theoretical concepts. In particular, several properties of the stochastic drift–diffusion process are presented, which will be used extensively throughout the remainder of this thesis. I also provide two simple examples of diffusion models based on stochastic drift–diffusion processes by recreating the historical approaches to Brownian motion by Einstein and Langevin in the modern formalism.

Chapter 3 addresses the adjustments needed to make these types of diffusion models compatible with special relativity, and defines a simple generic relativistic phase-space drift–diffusion process that serves as a starting point for the later models. Afterwards, the general physical setting, relativistic collisions of heavy-ion nuclei, is introduced, along with the set of coordinates in which the majority of calculations is performed. I then outline the common structure of the later models and give some details on the numerical approach.

In chapter 4, the first model on the deceleration of nucleus components in the early stages of the collision ("baryon stopping") is presented. I assume an initial state based on quantum statistics, while the asymptotic distribution, and thereby the coefficient functions, are derived from QCD-inspired phenomenology. The model is then applied to distributions of net protons and compared with data collected at the Super Proton Synchrotron (SPS) and the Relativistic Heavy Ion Collider (RHIC).

Chapter 5 covers the second model, which aims to describe the thermalization of the numerous hadrons produced in the collision. The QCD-based distributions from the preceding chapter are reused to approximate the initial state, and various thermal distributions with collective flow are examined as possible asymptotic-state candidates. With the resulting model, I then attempt to reproduce RHIC pion data and distributions of unidentified charged hadrons from the Large Hadron Collider (LHC).

In the final chapter 6, I summarize the obtained results and give a brief evaluation of the applicability and usefulness of this type of phase-space diffusion models in the context of relativistic heavy-ion collisions. The thesis closes with an overview of open questions, promising improvements, and possible future applications.

2. Modeling diffusion with stochastic processes

The aim of this chapter is to shed some light on the stochastic concepts used throughout this thesis in order to create a common ground for the upcoming discussion. As such, the following statements are not necessarily mathematically rigorous and should merely be regarded as working definitions. Relevant technical terms are *highlighted* the first time they appear in the text for ease of reference.

Readers interested in the mathematical details are encouraged to refer to the relevant standard textbooks – see, for example, Gardiner (2009) and Pavliotis (2014) – as well as Hänggi and Thomas' comprehensive report (1982), which offers a physical perspective on the topic. A lightweight introduction to the measure-theoretic foundations can be found in a series of blog posts by Bernstein (2019, 2020a,b).

2.1. Random variables

For the purpose of this thesis, a *random variable* X can be regarded as a kind of placeholder that can randomly take some value x (*realization*[1]) from a given *state space* S. For X to be well-defined, S has to be *measurable*, which essentially means that forming subsets of S is restricted to a certain class of sets to which some concept of "size" (count, length, area, volume, ...) can be applied in a consistent manner.

The probability that the realization of X is contained in some subset $\mathrm{O} \subseteq \mathrm{S}$ of the state space is given by

$$P_X(\mathrm{O}) = \int_{\mathrm{O}} \mathrm{d}P_X, \tag{2.1}$$

where P_X is the *probability measure* associated with X. A *measure* is a special kind of function that assigns non-negative real numbers to measurable subsets, reflecting their "size". Note that the identity (2.1) implies that the sizes of disjunct sets $\mathrm{O}_1, \mathrm{O}_2 \subseteq \mathrm{S}$ add up

---

[1] In the following, all random variables will be denoted by upper-case letters and their realizations by the corresponding lower-case letters.

to the size of their union,

$$P_X(\mathbf{O}_1) + P_X(\mathbf{O}_2) = P_X(\mathbf{O}_1 \cup \mathbf{O}_2) \quad \text{if} \quad \mathbf{O}_1 \cap \mathbf{O}_2 = \emptyset\,, \tag{2.2}$$

while the empty set is always assigned a size of zero, $P_X(\emptyset) = 0$. As *probabilities*, the function values of P_X are also bounded from above by 1,

$$0 \leq P_X(\mathbf{O}) \leq 1\,, \tag{2.3}$$

and since any realization is contained in $\mathbf{S}$ by definition, P_X fulfills the normalization condition

$$P_X(\mathbf{S}) = 1\,. \tag{2.4}$$

The *law of large numbers* states that, given a function h whose domain includes $\mathbf{S}$, the sequence $\left(s_n\right)_{n\in\mathbb{N}}$ of arithmetic means

$$s_n := \frac{1}{n}\sum_{i=1}^{n} h\left(x_n\right) \tag{2.5}$$

approaches a common limit for any collection of realizations $\{x_1, x_2, ...\}$ of X. This limit is denoted as the *expectation* of $h(X)$ with respect to X, in short $E_X\big(h(X)\big)$, and can be calculated directly via the Lebesgue integral

$$E_X\big(h(X)\big) = \int_{\mathbf{S}} h\,\mathrm{d}P_X\,. \tag{2.6}$$

Depending on the properties of X, this expression may reduce to a sum or a Riemann integral, as will be discussed below.

2.1.1. Probability mass and density functions

While the knowledge of $P_X(\mathbf{O})$ for all viable subsets $\mathbf{O} \subseteq \mathbf{S}$ offers a complete description of a given random variable X, this soon becomes unwieldy when random variables have a large or even infinite number of possible realizations, because the number of subsets can increase tremendously. Fortunately, however, an equivalent probability function can be defined for any random variable that has the same information content but requires only a single element of the state space as an argument, making it easier to handle. In this thesis, only two special kinds of random variables will be of further interest, for which these functions can be derived from P_X as follows:

1. A random variable X whose realizations – neglecting those with probability zero – form a *countable* set is called *discrete*. One can then decompose

$$P_X(\mathbf{O}) = \sum_{x \in \mathbf{O}} P_X(\{x\}) = \sum_{x \in \mathbf{O}} p_X(x) \tag{2.7a}$$

for any $\mathbf{O} \subseteq \mathrm{S}$ with the *probability mass function* (PMF) $p_X(x) := P_X(\{x\})$. The function p_X is normalized by

$$\sum_{x \in \mathrm{S}} p_X(x) = 1\,, \tag{2.7b}$$

which follows directly from eq. (2.4). Similarly, the expectation (2.6) of any function h simplifies to the sum

$$E_X\big(h(X)\big) = \sum_{x \in \mathrm{S}} h(x)\, p_X(x)\,. \tag{2.7c}$$

2. If the probability that the realization of X is contained in some $\mathbf{O} \subseteq \mathrm{S}$ can be written as

$$P_X(\mathbf{O}) = \int_{\mathbf{O}} \mathrm{d}P_X = \int_{\mathbf{O}} f_X(x)\, \mathrm{d}\lambda(x) \tag{2.8a}$$

with the *Lebesgue measure* λ and the associated *Radon–Nikodým derivative* $\mathrm{d}P_X/\mathrm{d}\lambda =: f_X$, the random variable X is called *continuous* with *probability density function* (PDF) f_X. The normalization condition (2.4) and expectation (2.6) then take the form

$$\int_{\mathrm{S}} f_X(x)\, \mathrm{d}\lambda(x) = 1\,, \tag{2.8b}$$

$$E_X\big(h(X)\big) = \int_{\mathrm{S}} h(x)\, f_X(x)\, \mathrm{d}\lambda(x)\,. \tag{2.8c}$$

For an n-dimensional $\boldsymbol{X}$ with $\mathrm{S} \subseteq \mathbb{R}^n$ and $\boldsymbol{x} =: \big(x^1, \ldots, x^n\big)$, $\mathrm{d}\lambda(\boldsymbol{x})$ and $f_{\boldsymbol{X}}(\boldsymbol{x})$ can be simplified to

$$\mathrm{d}\lambda(\boldsymbol{x}) = \prod_{i=1}^{n} \mathrm{d}x^i =: \mathrm{d}^n x, \tag{2.8d}$$

$$f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}^n x \approx P_{\boldsymbol{X}}\left(\bigotimes_{i=1}^{n} \big[x^i, x^i + \mathrm{d}x^i\big]\right). \tag{2.8e}$$

Furthermore, if $\boldsymbol{\varphi}$ is a diffeomorphism with domain S, the object $\boldsymbol{Y} := \boldsymbol{\varphi}(\boldsymbol{X})$ is also an n-dimensional real continuous random variable, and the PDFs of $\boldsymbol{X}$ and $\boldsymbol{Y}$ are connected through the Jacobian determinant of $\boldsymbol{\varphi}$,

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \Big|\det\big(\boldsymbol{J}_{\boldsymbol{\varphi}}(\boldsymbol{x})\big)\Big|\, f_{\boldsymbol{Y}}\big(\boldsymbol{\varphi}(\boldsymbol{x})\big)\,, \tag{2.8f}$$

where $\boldsymbol{J}_{\boldsymbol{\varphi}} := (\nabla \boldsymbol{\varphi}^{\mathsf{T}})^{\mathsf{T}}$ denotes the $(n \times n)$-dimensional Jacobian matrix of $\boldsymbol{\varphi}$.

A random variable is said to follow a given *distribution* if its probability function (PMF or PDF) is part of an associated function family. For example, an n-dimensional real continuous random variable $\boldsymbol{X}$ that has a PDF of the form[2]

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma})}} \exp\Big(-\tfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathsf{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})\Big) \tag{2.9}$$

with $\boldsymbol{\mu} \in \mathbb{R}^n$ and symmetric, positive definite $\boldsymbol{\Sigma} \in \mathbb{R}^{n\times n}$ is called *normally distributed* with *mean* $\boldsymbol{\mu}$ and *covariance* $\boldsymbol{\Sigma}$, or in short,

$$\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) . \tag{2.10}$$

It is easy to show that

$$\boldsymbol{\mu} = E_{\boldsymbol{X}}(\boldsymbol{X}) , \qquad \boldsymbol{\Sigma} = E_{\boldsymbol{X}}\Big((\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^{\mathsf{T}}\Big) \tag{2.11}$$

in this case. The diagonal of $\boldsymbol{\Sigma}$ is known as the *variance* of $\boldsymbol{X}$.

Any linear transformation $\boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}$ of $\boldsymbol{X}$ with $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$ is also normally distributed with

$$\boldsymbol{A}\boldsymbol{X} + \boldsymbol{b} \sim \mathcal{N}(\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{b}, \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^{\mathsf{T}}) . \tag{2.12}$$

Especially, the choice $\boldsymbol{A} = \boldsymbol{\Sigma}^{-1/2}$, $\boldsymbol{b} = -\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}$ yields a *standardized* random variable that follows the *standard normal distribution*

$$\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{X}-\boldsymbol{\mu}) \sim \mathcal{N}(\boldsymbol{0}, \mathbb{1}) , \tag{2.13}$$

with the $(n \times n)$-dimensional identity matrix $\mathbb{1}$.

Normally distributed random variables are frequently encountered in probability theory because many stochastic phenomena asymptotically approach normal distributions; a circumstance described by the *central limit theorem.* As a popular example, the chance of getting k heads from n tosses of a fair coin approaches a normal distribution with $\mu = n/2$ and $\Sigma = n/4$ for large n.

Mean and covariance, as defined in eq. (2.11), can also be calculated for many, but not all, non-normally distributed numerical random variables.

2.1.2. Marginal and conditional probabilities

When dealing with a random variable $\boldsymbol{X}$ that possesses an n-dimensional state space S, *e. g.*, $\mathrm{S} \subseteq \mathbb{R}^n$ as above, $\boldsymbol{X}$ can be split into n one-dimensional random variables $X^1, \dots, X^n$

---

[2] Unless otherwise specified, multidimensional random variables and their realizations will be treated as column vectors in linear-algebra operations.

to examine the behavior of subgroups of these components. For simplicity, this will be illustrated here solely for the subgroup $\boldsymbol{X}' := \left(X^1, \dots, X^k\right)$ with some $k \in \{1, \dots, n-1\}$; however, generalization to arbitrary combinations of components is straightforward.

In case one is only interested in $\boldsymbol{X}'$, the remaining $\left(X^{k+1}, \dots, X^n\right) =: \boldsymbol{X}''$ can be dismissed completely by considering the *marginal* probability measure

$$P_{\boldsymbol{X}'}(\mathrm{O}') := P_{\boldsymbol{X}}(\mathrm{O}' \otimes \mathrm{S}'') \tag{2.14}$$

for subsets $\mathrm{O}' \subseteq \mathrm{S}'$, where $\mathrm{S}' := \{\boldsymbol{x}' \,|\, \boldsymbol{x} \in \mathrm{S}\}$ and $\mathrm{S}'' := \{\boldsymbol{x}'' \,|\, \boldsymbol{x} \in \mathrm{S}\}$ are the slices of S associated with $\boldsymbol{X}'$ and $\boldsymbol{X}''$, respectively. This represents a situation where $\boldsymbol{X}''$ is completely unknown or unobservable, or the state of $\boldsymbol{X}''$ is intentionally left open. If, on the contrary, specific information is available on $\boldsymbol{X}''$ such that it can be restricted to some subset $\mathrm{O}'' \subseteq \mathrm{S}''$, the behavior of $\boldsymbol{X}'$ is governed by the *conditional* probability measure

$$P_{\boldsymbol{X}'|\boldsymbol{X}''}(\mathrm{O}' \,|\, \mathrm{O}'') := \frac{P_{\boldsymbol{X}}(\mathrm{O}' \otimes \mathrm{O}'')}{P_{\boldsymbol{X}''}(\mathrm{O}'')} \,. \tag{2.15}$$

In the limiting case of entirely unspecific information $\mathrm{O}'' = \mathrm{S}''$ on $\boldsymbol{X}''$, eq. (2.15) reproduces eq. (2.14),

$$P_{\boldsymbol{X}'|\boldsymbol{X}''}(\mathrm{O}' \,|\, \mathrm{S}'') = \frac{P_{\boldsymbol{X}}(\mathrm{O}' \otimes \mathrm{S}'')}{P_{\boldsymbol{X}''}(\mathrm{S}'')} = P_{\boldsymbol{X}'}(\mathrm{O}') \,, \tag{2.16}$$

as intuitively expected. When dividing $\boldsymbol{X}$ into three subgroups $\boldsymbol{X}'$, $\boldsymbol{X}''$, and $\boldsymbol{X}'''$, the two concepts can be combined to yield conditional marginal probability measures. The "original" measure $P_{\boldsymbol{X}}$ is then usually referred to as the *joint* probability measure to clearly distinguish it from all the (partial) measures derived from it.

Two subgroups $\boldsymbol{X}'$ and $\boldsymbol{X}''$ of a random variable $\boldsymbol{X}$ are said to be *independent* if (and only if) the product of their marginal probability measures is identical to the joint probability measure,

$$P_{\boldsymbol{X}'}(\mathrm{O}')\, P_{\boldsymbol{X}''}(\mathrm{O}'') = P_{\boldsymbol{X}}(\mathrm{O}' \otimes \mathrm{O}'') \,. \tag{2.17}$$

In this special case, the conditional probability measure of $\boldsymbol{X}'$ matches its marginal probability measure, $P_{\boldsymbol{X}'|\boldsymbol{X}''}(\mathrm{O}' \,|\, \mathrm{O}'') = P_{\boldsymbol{X}'}(\mathrm{O}')$, and the same applies analogously for $\boldsymbol{X}''$. In general, however, $\boldsymbol{X}'$ and $\boldsymbol{X}''$ are *dependent* and the above is *not* the case.

All these considerations also apply to probability mass and density functions if existent. Their joint, marginal, and conditional versions are related through

$$p_{\boldsymbol{X}'}(\boldsymbol{x}') = \sum_{\boldsymbol{x}'' \in \mathrm{S}''} p_{\boldsymbol{X}}(\boldsymbol{x}) \,, \qquad p_{\boldsymbol{X}'|\boldsymbol{X}''}(\boldsymbol{x}' \,|\, \boldsymbol{x}'') = \frac{p_{\boldsymbol{X}}(\boldsymbol{x})}{p_{\boldsymbol{X}''}(\boldsymbol{x}'')} \,, \tag{2.18}$$

$$f_{\boldsymbol{X}'}(\boldsymbol{x}') = \int_{\mathrm{S}''} f_{\boldsymbol{X}}(\boldsymbol{x}) \, \mathrm{d}\lambda(\boldsymbol{x}'') \,, \qquad f_{\boldsymbol{X}'|\boldsymbol{X}''}(\boldsymbol{x}' \,|\, \boldsymbol{x}'') = \frac{f_{\boldsymbol{X}}(\boldsymbol{x})}{f_{\boldsymbol{X}''}(\boldsymbol{x}'')} \,, \tag{2.19}$$

while eq. (2.16) takes the form

$$p_{X'}(x') = \sum_{x''\in S''} p_{X'|X''}(x' \,|\, x'')\, p_{X''}(x'')\,, \tag{2.20}$$

$$f_{X'}(x') = \int_{S''} f_{X'|X''}(x' \,|\, x'')\, f_{X''}(x'')\, \mathrm{d}\lambda(x'')\,. \tag{2.21}$$

Example: Coupled spins

To illustrate some of the concepts introduced so far, consider the decay of a spinless particle into a pair of spin-½ fermions with spins $\vec{s}_1$ and $\vec{s}_2$. For any given axis $\vec{a}$, this system can be described by a two-dimensional discrete random variable $\boldsymbol{X} = \left(X^1, X^2\right)$ with components

$$X^i := \frac{2\vec{a}\cdot\vec{s}_i}{\hbar} \quad \text{for} \quad i = 1, 2\,, \tag{2.22}$$

where $\hbar$ denotes the *reduced Planck constant*, and state space

$$S = \{-1, +1\}^2 = \left\{(-1,-1), (+1,+1), (-1,+1), (+1,-1)\right\} \subseteq \mathbb{Z}^2\,. \tag{2.23}$$

The PMF of $\boldsymbol{X}$ is simply given by

$$p_{\boldsymbol{X}}(\boldsymbol{x}) = \begin{cases} 0 & \text{if} \quad x^1 = +x^2\,, \\ \frac{1}{2} & \text{if} \quad x^1 = -x^2\,, \end{cases} \tag{2.24}$$

which follows immediately from angular momentum conservation and isotropy.

When only one of the fermions is observed and the configuration or existence of the other is unknown, the entanglement of the spins is not visible, and each fermion behaves as if it was an independent particle. This behavior is described by the marginal PMFs[3]

$$p_{X^1}(x^1) = p_{X^1,X^2}(x^1, -1) + p_{X^1,X^2}(x^1, +1) = \tfrac{1}{2} \tag{2.25}$$

and analogously $p_{X^2}(x^2) = \frac{1}{2}$. In contrast, knowing the spin of one fermion with certainty completely predetermines the fate of the other, which then behaves deterministically with conditional PMF

$$p_{X^1|X^2}(x^1 \,|\, x^2) = p_{X^2|X^1}(x^2 \,|\, x^1) = 2p_{\boldsymbol{X}}(\boldsymbol{x}) = \begin{cases} 0 & \text{if} \quad x^1 = +x^2\,, \\ 1 & \text{if} \quad x^1 = -x^2\,, \end{cases} \tag{2.26}$$

respectively.

---

[3] The tuple delimiters in the subscript and argument of the PMFs have been omitted to declutter notation and improve readability.

2.2. Stochastic processes

Stochastic processes are, in essence, parameterized random variables and can be seen as generalizations of the multidimensional random variables discussed above. Mathematically, one can approach this concept from different angles:

First, a stochastic process X is technically a single random variable whose realizations are elements of a *function space*, $\mathrm{S} \subseteq \{x\colon \mathbf{T} \to \mathbf{U}\}$, which is a collection of functions sharing the same domain $\mathbf{T}$ and codomain $\mathbf{U}$, where the latter has to be measurable in this context. The functions $x \in \mathrm{S}$ are referred to as *paths* of the stochastic process X and their domain $\mathbf{T}$ as its *parameter set*. If $\mathbf{T}$ is countable or connected, X is called *parameter-discrete* or *parameter-continuous*, respectively. Note that a stochastic process with finite parameter set $\mathbf{T} = \{1, \dots, n\}$ behaves exactly like an n-dimensional random variable if one identifies $x(i) \leftrightarrow x^i$.

Second, stochastic processes can be considered as collections of random variables: Using its parameter set $\mathbf{T}$ as an index set, a stochastic process X can be decomposed into its *members* $\big(X(t) \mid t \in \mathbf{T}\big)$, a collection of random variables that share the same state space $\mathbf{U}$.[4] This is similar to the decomposition of an n-dimensional random variable into its n components, and the obtained objects are also in general not independent of each other. The number of members, however, can be infinite or even uncountable, depending on the cardinality of $\mathbf{T}$. If all members are discrete (continuous) random variables, the associated stochastic process is called *value-discrete* (*value-continuous*).

These two views offer complementary descriptions of the same object, and hence, one is free to pick the most convenient one for the task at hand. Nevertheless, the mathematical treatment of stochastic processes can be quite difficult when the parameter set is not finite. The further discussion will therefore be limited to stochastic processes possessing the so-called *Markov property* (Markov 1906) or, in short, *Markov processes*.

2.2.1. Markov property

To be eligible as a Markov process, the parameter set $\mathbf{T}$ of a stochastic process X has to be *totally ordered*, which means that statements like $t_1 \leq t_2$ are well-defined for all $t_1, t_2 \in \mathbf{T}$, such that the members of X can be arranged as a sequence $\big(X(t)\big)_{t \in \mathbf{T}}$ of random variables. Furthermore, each member $X(t)$ must be independent of all other members except a *finite* number of *immediate* predecessors. If this number is not greater than some $m \in \mathbb{N}$ for all $t \in \mathbf{T}$, X is said to be a *Markov process of order m*. When the elements of the parameter set correspond to points in time, as will be the case below, the Markov property can be

---

[4] Usually, the members' state space $\mathbf{U}$ and not the function space S is referred to as the state space of the (entire) stochastic process.

paraphrased as a "short-time memory", since the stochastic process seems to "lose track" of everything but its immediate past.

The following discussion will be limited to the case $m = 1$, because any Markov process $\boldsymbol{X}$ of dimension n and order m can be rewritten as an $(n \cdot m)$-dimensional first-order Markov process $\boldsymbol{Y}$: For time-discrete $\boldsymbol{X}$, this can be achieved, *e. g.*, by defining

$$\boldsymbol{Y}(t_k) := \big(\boldsymbol{X}(t_k), \boldsymbol{X}(t_{k-1}), \dots, \boldsymbol{X}(t_{k-m+1})\big)\,, \tag{2.27a}$$

so that all members on which $\boldsymbol{X}(t_k)$ depends, except the most distant one, are included in $\boldsymbol{Y}(t_k)$. For a time-continuous process, where the distance $\mathrm{d}t$ between adjacent elements of $\mathbf{T}$ is infinitesimal, choosing the linear combinations

$$Y^i(t) := \sum_{l=0}^{i} (-1)^l \binom{i}{l} \frac{\boldsymbol{X}(t - l\mathrm{d}t)}{(\mathrm{d}t)^i} =: \frac{\partial^i \boldsymbol{X}(t)}{(\partial t)^i} \tag{2.27b}$$

of the $i = 1, \dots, m-1$ nearest members yields a tuple $\boldsymbol{Y}(t) := \big(Y^1(t), \dots, Y^m(t)\big)$ containing $\boldsymbol{X}(t)$ and its first $m-1$ derivatives. The latter is reminiscent of how an *ordinary differential equation* (ODE) of order m can always be reduced to a system of m coupled first-order ODEs, and as will be seen later, a direct connection exists in the form of *stochastic differential equation*s. Correspondingly, the possibility of obtaining a first-order Markov process $\boldsymbol{Y}$ by rewriting some stochastic process $\boldsymbol{X}$ in the form (2.27b) can also be considered the defining property of time-continuous m^{th}-order Markov processes.

As a consequence of the above, when dividing the parameter set $\mathbf{T}$ of a first-order Markov process X at some breakpoints $t_1 < t_2 < \dots < t_{m-1} \in \mathbf{T}$ into m segments

$$\mathbf{T}_k := \begin{cases} \{t \in \mathbf{T} \,|\, t \le t_1\} & \text{for} \quad k = 1 \\ \{t \in \mathbf{T} \,|\, t_{k-1} < t \le t_k\} & \text{for} \quad k = 2, \dots, m-1 \\ \{t \in \mathbf{T} \,|\, t_{m-1} < t\} & \text{for} \quad k = m \end{cases} \tag{2.28}$$

ensuring that $\biguplus_{k=1}^{m} \mathbf{T}_k = \mathbf{T}$, the probability that X takes a path in $\mathbf{O} = \bigotimes_{k=1}^{m} \mathbf{O}_k$ can be rewritten as

$$P_X(\mathbf{O}) = P_{X_1}(\mathbf{O}_1) \prod_{k=1}^{m-1} P_{X_{k+1}|X_k}(\mathbf{O}_{k+1} \,|\, \mathbf{O}_k) \tag{2.29}$$

with $X_k := \big(X(t) \,|\, t \in \mathbf{T}_k\big)$ by successively applying eq. (2.15). Here, the Markov property of X was used in simplifying the emerging conditional probabilities to marginal conditional probabilities,

$$P_{X_{k+1}|X_1,\dots,X_k}(\mathbf{O}_{k+1} \,|\, \mathbf{O}_1 \otimes \dots \otimes \mathbf{O}_k) = P_{X_{k+1}|X_k}(\mathbf{O}_{k+1} \,|\, \mathbf{O}_k)\,, \tag{2.30}$$

because X_{k+1} does not depend *directly* on members with $t < t_k$. In other words, the problem of describing an entire Markov process can be reduced to the study of its *initial state* X_1 and the *transitions* from one state to the next, $P_{X_{k+1}|X_k}$.

Indeed, the individual transition probabilities (2.30) can be further simplified to

$$P_{X_{k+1}|X_k}\big(\mathrm{O}_{k+1} \,\big|\, \mathrm{O}_k\big) = P_{X_{k+1}|X(t_k)}\big(\mathrm{O}_{k+1} \,\big|\, \mathrm{O}(t_k)\big)\,, \tag{2.31}$$

where $\mathrm{O}(t)$ denotes the slice of O belonging to $X(t)$. Often, one is only interested in the behavior at the breakpoints t_k and leaves the path at the other times unspecified by marginalizing $\mathrm{O}(t) = \mathrm{S}(t)$ for $t \neq t_k$, which yields the two-point transition probability

$$P_{X(t_{k+1})|X(t_k)}\big(\mathrm{O}(t_{k+1}) \,\big|\, \mathrm{O}(t_k)\big)\,. \tag{2.32}$$

Two-point transition probabilities for non-neighboring breakpoints can then be obtained through the *Chapman–Kolmogorov equation* (Chapman 1928; Kolmogorov 1931),

$$\begin{aligned} &P_{X(t_{k+1})|X(t_{k-1})}\big(\mathrm{O}(t_{k+1}) \,\big|\, \mathrm{O}(t_{k-1})\big) \\ &\qquad = P_{X(t_{k+1})|X(t_k)}\big(\mathrm{O}(t_{k+1}) \,\big|\, \mathrm{S}(t_k)\big)\, P_{X(t_k)|X(t_{k-1})}\big(\mathrm{S}(t_k) \,\big|\, \mathrm{O}(t_{k-1})\big)\,. \end{aligned} \tag{2.33}$$

2.2.2. Random walk and Wiener process

To conclude this section, two prominent examples of Markov processes are presented, which will serve as building blocks for further stochastic processes.

A *simple symmetric random walk* is a time- and value-discrete first-order Markov process R that randomly jumps back or forward a constant distance $\Delta r > 0$ with equal probability at each time step. Denoting the elements of the parameter set as $t_0, t_1, t_2, \ldots$, the (two-point) transition PMF for a single jump reads

$$p_{R(t_{k+1})|R(t_k)}(r_{k+1} \,|\, r_k) = \begin{cases} \frac{1}{2} & \text{if} \quad |r_{k+1} - r_k| = \Delta r\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{2.34}$$

Usually, a deterministic initial state $p_{R(t_0)}$ is chosen, *i. e.*, $R(t_0)$ is set to always assume a given value r_0, the starting point of the random walk. The derivation of the transition probability for multiple jumps is best illustrated graphically: Figure 2.1 shows the number of possible ways to arrive at the indicated positions after a certain number of jumps. One can easily see that the graph reproduces Pascal's triangle, because the number of ways to fit l_+ forward jumps and l_- backward jumps into $l_+ + l_- =: l$ total jumps is given by the binomial coefficient $\binom{l}{l_+}$. The corresponding probabilities are obtained by dividing the latter by the total number of possible paths 2^l, and noting that the net distance traveled is

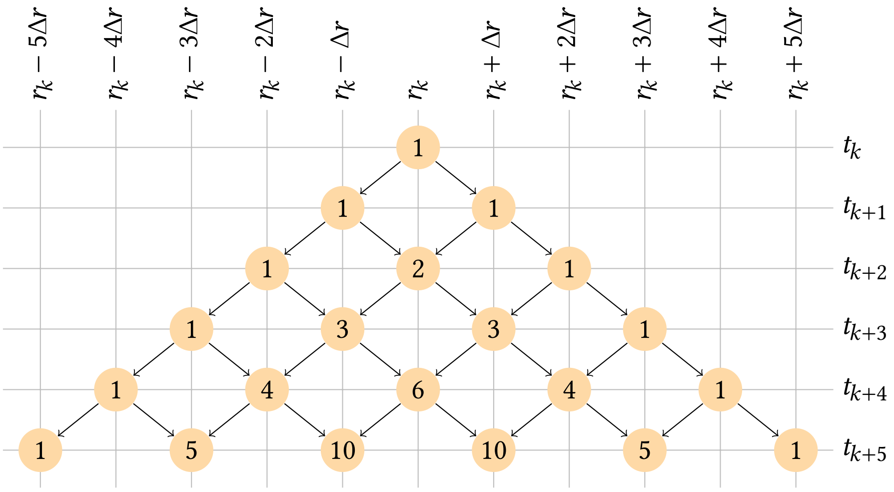


Figure 2.1: All possible paths of a simple symmetric random walk for five time steps. Valid waypoints are marked with circles that indicate the number of possible ways to arrive at said position.

$(2l_+ - l)\,\Delta r$, which finally leads to[5]

$$p_{R(t_{k+l})|R(t_k)}(r_{k+l}\,|\,r_k) = \begin{cases} \frac{1}{2^l}\binom{l}{l_+} & \text{if} \quad l_+ := \frac{1}{2}\left(l + \frac{r_{k+l}-r_k}{\Delta r}\right) \in \mathbb{N} \cup \{0\}\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{2.35}$$

The *standard Wiener process* (Wiener 1923) is a continuous counterpart to the discrete random walk. In physics, it is used to describe phenomena like Brownian motion and Gaussian noise. To qualify as a standard Wiener process, a stochastic process W has to meet the following requirements:

1. $W(0) = 0$
2. $W(t) - W(t')$ is independent of $W(t'')$ for all $t > t' > t''$
3. $W(t + \Delta t) - W(t) \sim \mathcal{N}(0, \Delta t)$ for all $\Delta t > 0$
4. $W(t)$ is continuous[6] in t

The second criterion is basically a stricter form of the Markov property, and consequently, all Wiener processes are necessarily also first-order Markov processes. Together with the third and fourth criteria, it implies that the transition probabilities of W are equal to the

---

[5] Considering backward jumps yields virtually the same result since $\binom{l_+ + l_-}{l_+} = \binom{l_+ + l_-}{l_-}$.

[6] More formally: almost all paths of W are continuous functions

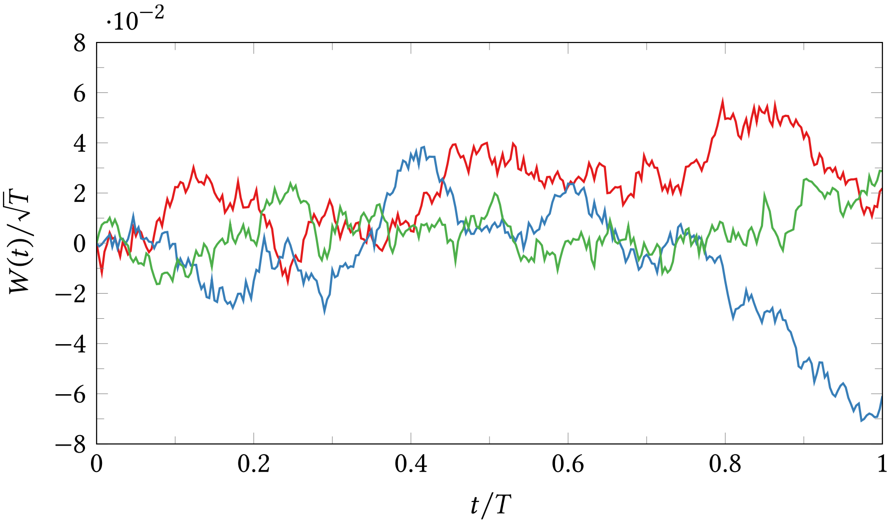


Figure 2.2: Three paths of a standard Wiener process W with parameter set $[0, T]$.

marginal probabilities of its increments,

$$f_{W(t+\Delta t)|W(t)}(w + \Delta w|\, w) = f_{W(t+\Delta t)-W(t)}(\Delta w) = \frac{1}{\sqrt{2\pi\,\Delta t}} \exp\left(-\frac{(\Delta w)^2}{2\,\Delta t}\right) \tag{2.36}$$

for positive time increments $\Delta t > 0$. Moreover, due to the first requirement, all members with $t > 0$ are themselves also normally distributed since

$$W(t) = W(t) - W(0) \sim \mathcal{N}(0, t)\,, \tag{2.37}$$

such that

$$f_{W(t)}(w) = f_{W(t)-W(0)}(w) = f_{W(t)|W(0)}(w|\,0)\,. \tag{2.38}$$

The initial state itself is given by a Dirac function,

$$f_{W(0)}(w) = \delta(w)\,. \tag{2.39}$$

According to *Donsker's theorem* (Donsker 1952, also known as *functional central limit theorem*), a standard Wiener process can be constructed as the scaling limit of a random walk: If R is a simple symmetric random walk with deterministic starting position $R(0) = 0$, jump distance $\Delta r = \sqrt{T/n}$, and evenly-spaced time steps $t_k = kT/n$ for $k = 0, \dots, n$ and $T > 0$, then the limit $W := \lim_{n\to\infty} R$ is a standard Wiener process with parameter set $[0, T]$. Example paths of such a process are shown in fig. 2.2.

2.3. Diffusion as a stochastic process

An n-dimensional *drift–diffusion process* $\boldsymbol{X}$ with constant drift $\boldsymbol{\mu}_X \in \mathbb{R}^n$ and diffusion $\boldsymbol{\sigma}_X \in \mathbb{R}^{n\times m}$ is a time- and value-continuous stochastic process whose increments fulfill the relation

$$\boldsymbol{X}(t+\Delta t) - \boldsymbol{X}(t) = \boldsymbol{\mu}_X \Delta t + \boldsymbol{\sigma}_X \big[\boldsymbol{W}(t+\Delta t) - \boldsymbol{W}(t)\big] , \tag{2.40a}$$

or, component-wise for $i = 1, \dots, n$,

$$X^i(t+\Delta t) - X^i(t) = \mu_X^i \Delta t + \sum_{l=1}^{m} \sigma_X^{il} \big[W^l(t+\Delta t) - W^l(t)\big] , \tag{2.40b}$$

for a collection of m independent standard Wiener processes $\boldsymbol{W} = \big(W^1, \dots, W^m\big)$. It can be shown that $\boldsymbol{X}$ inherits many of the special characteristics of the W^l, especially a normal distribution of the process increments,

$$\boldsymbol{X}(t+\Delta t) - \boldsymbol{X}(t) \sim \mathcal{N}\big(\boldsymbol{\mu}_X \Delta t, 2\boldsymbol{D}_X \Delta t\big) , \tag{2.41}$$

where $\boldsymbol{D}_X \coloneqq \frac{1}{2}\boldsymbol{\sigma}_X \boldsymbol{\sigma}_X^\top$ denotes the *diffusivity* of $\boldsymbol{X}$, and that the PDFs fulfill

$$f_{\boldsymbol{X}(t+\Delta t)|\boldsymbol{X}(t)}(\boldsymbol{x} + \Delta\boldsymbol{x} \,|\, \boldsymbol{x}) = f_{\boldsymbol{X}(t+\Delta t)-\boldsymbol{X}(t)}(\Delta\boldsymbol{x}) \tag{2.42}$$

similar to eq. (2.36). Hence, $\boldsymbol{X}$ is essentially a scaled and shifted Wiener process in multiple dimensions, which is why it is sometimes also called a *generalized Wiener process.* Note that the distribution of $\boldsymbol{X}$ depends on the diffusion coefficient $\boldsymbol{\sigma}_X$ only indirectly via the diffusivity $\boldsymbol{D}_X$, which means that $\boldsymbol{X}$ is invariant under transformations $\boldsymbol{\sigma}_X \mapsto \boldsymbol{\sigma}_X \boldsymbol{Q}$ with orthogonal $(m \times m)$-matrices $\boldsymbol{Q}$. This is even true if $\boldsymbol{Q}$ is a function of $\boldsymbol{X}(t)$.

Although it entails some technical subtleties, extending this concept to non-constant coefficients is principally straightforward: If the coefficients do not vary too much with $\boldsymbol{X}$, they can be considered virtually constant for infinitesimally-small time increments $\Delta t = \mathrm{d}t$, and thus eqs. (2.40) to (2.42) should still hold in this limit. More precisely, an n-dimensional drift–diffusion process $\boldsymbol{X}$ with member state space $\mathrm{U} \subseteq \mathbb{R}^n$ and *Lipschitz-continuous*[7] drift and diffusion coefficient functions $\boldsymbol{\mu}_X \colon \mathrm{U} \to \mathbb{R}^n$ and $\boldsymbol{\sigma}_X \colon \mathrm{U} \to \mathbb{R}^{n\times m}$ fulfills the *stochastic differential equation* (SDE)

$$\mathrm{d}\boldsymbol{X}(t) = \boldsymbol{\mu}_X\big(\boldsymbol{X}(t)\big)\,\mathrm{d}t + \boldsymbol{\sigma}_X\big(\boldsymbol{X}(t)\big)\,\mathrm{d}\boldsymbol{W}(t) \tag{2.43a}$$

---

[7] *I. e.*, there exists a constant $C \geq 0$ such that $\|\boldsymbol{\mu}_X(\boldsymbol{x}_1) - \boldsymbol{\mu}_X(\boldsymbol{x}_2)\| + \|\boldsymbol{\sigma}_X(\boldsymbol{x}_1) - \boldsymbol{\sigma}_X(\boldsymbol{x}_2)\| \leq C\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|$ for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathrm{U}$ in some norm $\|\cdot\|$. This guarantees the existence and uniqueness of $\boldsymbol{X}$ (*cf.* Itô 1946).

with the infinitesimally-small process increments

$$\mathrm{d}\boldsymbol{X}(t) \coloneqq \boldsymbol{X}(t+\mathrm{d}t) - \boldsymbol{X}(t)\,, \qquad \mathrm{d}\boldsymbol{W}(t) \coloneqq \boldsymbol{W}(t+\mathrm{d}t) - \boldsymbol{W}(t)\,. \tag{2.43b}$$

Finite increments are obtained by formally integrating eq. (2.43) over an interval $[t_a, t_b]$, which yields the *stochastic integral*

$$\boldsymbol{X}(t_b) - \boldsymbol{X}(t_a) = \int_{t_a}^{t_b} \mathrm{d}\boldsymbol{X}(t) = \underbrace{\int_{t_a}^{t_b} \boldsymbol{\mu}_X\big(\boldsymbol{X}(t)\big)\,\mathrm{d}t}_{\boldsymbol{I}_1} + \underbrace{\int_{t_a}^{t_b} \boldsymbol{\sigma}_X\big(\boldsymbol{X}(t)\big)\,\mathrm{d}\boldsymbol{W}(t)}_{\boldsymbol{I}_2}\,. \tag{2.44}$$

The expressions $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$ are defined similar to Riemann–Stieltjes integrals,

$$\boldsymbol{I}_1 \coloneqq \lim_{N\to\infty} \sum_{k=1}^{N} \Big[\tfrac{1+\lambda}{2}\,\boldsymbol{\mu}_X\big(\boldsymbol{X}(t_k)\big) + \tfrac{1-\lambda}{2}\,\boldsymbol{\mu}_X\big(\boldsymbol{X}(t_{k-1})\big)\Big]\,\big[t_k - t_{k-1}\big]\,, \tag{2.45a}$$

$$\boldsymbol{I}_2 \coloneqq \lim_{N\to\infty} \sum_{k=1}^{N} \Big[\tfrac{1+\lambda}{2}\,\boldsymbol{\sigma}_X\big(\boldsymbol{X}(t_k)\big) + \tfrac{1-\lambda}{2}\,\boldsymbol{\sigma}_X\big(\boldsymbol{X}(t_{k-1})\big)\Big]\,\big[\boldsymbol{W}(t_k) - \boldsymbol{W}(t_{k-1})\big]\,, \tag{2.45b}$$

by partitioning the interval $[t_a, t_b]$ into N subintervals $[t_k, t_{k-1}]$ with $t_a = t_0 < t_1 < \cdots < t_N = t_b$ and taking the limit $N \to \infty$ for some $\lambda \in [-1, +1]$, which sets the discretization scheme of the integrand. While $\boldsymbol{I}_1$ converges irrespective of the choice of λ, this is not the case for $\boldsymbol{I}_2$, because the stochastic processes $\boldsymbol{X}$ and $\boldsymbol{W}$ depend on each other. As an illustrative example, the stochastic integration of a standard Wiener process with respect to itself yields

$$\int_{t_a}^{t_b} W(t)\,\mathrm{d}W(t) = \tfrac{1}{2}\big[W(t_b)^2 - W(t_a)^2\big] + \tfrac{\lambda}{2}\big(t_b - t_a\big)\,, \tag{2.46}$$

which deviates from the result expected from the fundamental theorem of calculus by an additional, λ-dependent term if $\lambda \neq 0$. See appendix A for the full calculation.

The following discussion will be restricted to the three discretizations $\lambda \in \{-1, \pm 0, +1\}$, which yield the stochastic integrals most frequently encountered in literature: The pre-point rule ($\lambda = -1$), which defines the so-called *Itô integral* (Itô 1944), the mid-point rule ($\lambda = 0$) with the *Stratonovich–Fisk integral* (Stratonovich 1964; Fisk 1965), and the post-point rule ($\lambda = +1$), whose stochastic integral is often referred to as the *Hänggi–Klimontovich integral* (Hänggi 1978; Klimontovich 1990). Table 2.1 summarizes the connection between λ, discretization scheme, and stochastic integral for ease of reference. In the context of drift–diffusion processes, each formulation offers its own advantages and disadvantages, as will become clear below.

Table 2.1: Overview of the most common discretization schemes and the associated integrals in stochastic calculus.

λ	discretization	stochastic integral
-1	pre-point	Itô
± 0	mid-point	Stratonovich–Fisk
$+1$	post-point	Hänggi–Klimontovich

Returning to eq. (2.45b), it can be shown that for $\lambda \in \{-1, \pm 0, +1\}$, the stochastic integral $\boldsymbol{I}_2$ differs by

$$I_2^i\big|_{\lambda=+1} - I_2^i\big|_{\lambda=\pm 0} = \frac{1}{2}\int_{t_a}^{t_b}\left[\sum_{l=1}^{m}\sigma_{\boldsymbol{X}}^{il}(\boldsymbol{x})\sum_{k=1}^{n}\frac{\partial\sigma_{\boldsymbol{X}}^{kl}(\boldsymbol{x})}{\partial x^k}\right]_{\boldsymbol{x}=\boldsymbol{X}(t)}\mathrm{d}t \tag{2.47a}$$

$$I_2^i\big|_{\lambda=+1} - I_2^i\big|_{\lambda=-1} = \int_{t_a}^{t_b}\left[\sum_{k=1}^{n}\frac{\partial D_{\boldsymbol{X}}^{ik}(\boldsymbol{x})}{\partial x^k}\right]_{\boldsymbol{x}=\boldsymbol{X}(t)}\mathrm{d}t \tag{2.47b}$$

$$I_2^i\big|_{\lambda=\pm 0} - I_2^i\big|_{\lambda=-1} = \frac{1}{2}\int_{t_a}^{t_b}\left[\sum_{l=1}^{m}\sum_{k=1}^{n}\frac{\partial\sigma_{\boldsymbol{X}}^{il}(\boldsymbol{x})}{\partial x^k}\sigma_{\boldsymbol{X}}^{kl}(\boldsymbol{x})\right]_{\boldsymbol{x}=\boldsymbol{X}(t)}\mathrm{d}t \tag{2.47c}$$

for $i = 1, \dots, n$, which entails that all increments $\boldsymbol{X}(t_b) - \boldsymbol{X}(t_a)$, and thus the entire drift–diffusion process, implicitly depend on the choice of λ unless $\boldsymbol{\sigma}_{\boldsymbol{X}}$ is constant.

The discretization also affects the substitution of variables in the corresponding stochastic integrals and thereby transformations of stochastic processes: If $\boldsymbol{X}$ is a drift–diffusion process and $\varphi\colon \mathbf{U} \to \mathbb{R}$ a twice-differentiable function, the stochastic process Y defined via $Y(t) \coloneqq \varphi\big(\boldsymbol{X}(t)\big)$ fulfills

$$Y(t_b) - Y(t_a) = \varphi\big(\boldsymbol{X}(t_b)\big) - \varphi\big(\boldsymbol{X}(t_a)\big) = \int_{t_a}^{t_b}\mathrm{d}\varphi\big(\boldsymbol{X}(t)\big)\,. \tag{2.48}$$

According to Itô's lemma (Itô 1944), the infinitesimal increment in the stochastic integral expands to

$$\mathrm{d}\varphi\big(\boldsymbol{X}(t)\big) = \boldsymbol{J}_\varphi\big(\boldsymbol{X}(t)\big)\,\mathrm{d}\boldsymbol{X}(t) + \frac{1}{2}\,\mathrm{d}\boldsymbol{X}(t)^{\mathsf{T}}\boldsymbol{H}_\varphi\big(\boldsymbol{X}(t)\big)\,\mathrm{d}\boldsymbol{X}(t) \tag{2.49}$$

with the $(1 \times n)$-dimensional Jacobian matrix $\boldsymbol{J}_\varphi \coloneqq (\nabla\varphi)^{\mathsf{T}}$ and $(n \times n)$-dimensional Hessian matrix $\boldsymbol{H}_\varphi \coloneqq \boldsymbol{J}_{\nabla\varphi}^{\mathsf{T}}$ of φ. The last, quadratic term vanishes if $\boldsymbol{H}_\varphi$ or λ is zero, but otherwise introduces an additional contribution that is not present in the substitution rule of non-stochastic calculus. For $\lambda = -1$, the above reduces to

$$\begin{aligned}\mathrm{d}\varphi\big(\boldsymbol{X}(t)\big) = \Big\{\boldsymbol{J}_\varphi\big(\boldsymbol{X}(t)\big)\,\boldsymbol{\mu}_{\boldsymbol{X}}\big(\boldsymbol{X}(t)\big) + \frac{1}{2}\,\mathrm{Tr}\Big[\boldsymbol{\sigma}_{\boldsymbol{X}}\big(\boldsymbol{X}(t)\big)^{\mathsf{T}}\boldsymbol{H}_\varphi\big(\boldsymbol{X}(t)\big)\,\boldsymbol{\sigma}_{\boldsymbol{X}}\big(\boldsymbol{X}(t)\big)\Big]\Big\}\,\mathrm{d}t& \\ + \boldsymbol{J}_\varphi\big(\boldsymbol{X}(t)\big)\,\boldsymbol{\sigma}_{\boldsymbol{X}}\big(\boldsymbol{X}(t)\big)\,\mathrm{d}\boldsymbol{W}(t)\,,&\end{aligned} \tag{2.50}$$

which means that Y has the form of a one-dimensional drift–diffusion process. The latter, although not shown here, is also true for other values of λ; especially for $\lambda = 0$, where the term involving the Hessian vanishes. Transformation laws for functions mapping to multiple dimensions are obtained by applying eq. (2.49) or (2.50) component-wise.

The SDE as given in eq. (2.43) is hence ambiguous and not complete until a discretization rule is specified.[8] Usually, the latter is indicated by inserting some special multiplication sign between diffusion coefficient and Wiener-process increments, as in the next section, but the exact notation may differ from author to author.

2.3.1. Discretization-agnostic formulation

Since the terms in eq. (2.47) are merely integrals over t and not $\boldsymbol{W}(t)$, $\boldsymbol{X}$ can be made independent of the discretization scheme, if desired, by using different drift coefficient functions $\boldsymbol{\mu}_{X|-}$, $\boldsymbol{\mu}_{X|\pm}$, $\boldsymbol{\mu}_{X|+}$ for $\lambda = -1, \pm 0, +1$, respectively, and requiring

$$\boldsymbol{I}_1\big|_{\boldsymbol{\mu}_{X|\lambda}} - \boldsymbol{I}_1\big|_{\boldsymbol{\mu}_{X|\lambda'}} \overset{!}{=} \boldsymbol{I}_2\big|_{\lambda'} - \boldsymbol{I}_2\big|_{\lambda} \quad \text{for all} \quad \lambda, \lambda' \in \{-1, \pm 0, +1\}\,, \tag{2.51}$$

which is equivalent to

$$\mu^i_{X|+}(\boldsymbol{x}) \overset{!}{=} \mu^i_{X|\pm}(\boldsymbol{x}) - \tfrac{1}{2}\sum_{l=1}^{m} \sigma^{il}_X(\boldsymbol{x}) \sum_{k=1}^{n} \frac{\partial \sigma^{kl}_X(\boldsymbol{x})}{\partial x^k} \tag{2.52a}$$

$$\overset{!}{=} \mu^i_{X|-}(\boldsymbol{x}) - \sum_{k=1}^{n} \frac{\partial D^{ik}_X(\boldsymbol{x})}{\partial x^k}\,. \tag{2.52b}$$

For such a set of interconnected drift coefficient functions, the sum $\boldsymbol{I}_1 + \boldsymbol{I}_2$ in eq. (2.44) becomes invariant of the choice of λ and defines the very same process $\boldsymbol{X}$ fulfilling the discretization-agnostic SDE

$$\mathrm{d}\boldsymbol{X}(t) = \boldsymbol{\mu}_{X|\lambda}\big(\boldsymbol{X}(t)\big)\,\mathrm{d}t + \boldsymbol{\sigma}_X\big(\boldsymbol{X}(t)\big) *_\lambda \mathrm{d}\boldsymbol{W}(t)\,, \tag{2.53}$$

where the discretization scheme used is explicitly specified by the symbol $*_\lambda$.

With this approach, one benefits from what could be paraphrased as “discretization freedom”: The choice of λ does not alter $\boldsymbol{X}$, but only its representation in terms of coefficient functions; and while it has to be set, one is free to choose the most advantageous discretization for the task at hand. This is roughly comparable to the freedom to choose any coordinate system to describe a multidimensional object, which can be used to exploit given symmetries.

---

[8] It was therefore considered merely a “pre-equation” by van Kampen (1981).

For example, substitution of variables is easiest for $\lambda = \pm 0$ where the anomalous, quadratic term in Itô's lemma (2.49) vanishes: Consequently, for any twice-differentiable diffeomorphism $\boldsymbol{\varphi} = \left(\varphi^1, \dots, \varphi^n\right) \colon \mathrm{U} \to \mathrm{V} \subseteq \mathbb{R}^n$, the transformed coefficient functions of the drift–diffusion process $\boldsymbol{Y} \coloneqq \boldsymbol{\varphi}(\boldsymbol{X})$ simply read

$$\mu^k_{\boldsymbol{Y}|\pm}(\boldsymbol{y}) = \left[\sum_{i=1}^{n} \frac{\partial \varphi^k(\boldsymbol{x})}{\partial x^i} \mu^i_{\boldsymbol{X}|\pm}(\boldsymbol{x})\right]_{\boldsymbol{x}=\boldsymbol{\varphi}^{-1}(\boldsymbol{y})}, \tag{2.54a}$$

$$\sigma^{kl}_{\boldsymbol{Y}}(\boldsymbol{y}) = \left[\sum_{i=1}^{n} \frac{\partial \varphi^k(\boldsymbol{x})}{\partial x^i} \sigma^{il}_{\boldsymbol{X}}(\boldsymbol{x})\right]_{\boldsymbol{x}=\boldsymbol{\varphi}^{-1}(\boldsymbol{y})} \tag{2.54b}$$

for $k = 1, \dots, n$ and $l = 1, \dots, m$. The associated drift coefficients for $\lambda = \pm 1$ can then be easily recovered via eq. (2.52) if needed; alternatively, the PDFs of $\boldsymbol{Y}(t)$ may be determined via eq. (2.8f). This procedure might be faster than applying Itô's lemma directly because the Hessians of $\varphi^1, \dots, \varphi^n$ need not be computed explicitly.

The discretization scheme with $\lambda = -1$, on the other hand, is particularly well suited for the numerical simulation of paths of stochastic processes,[9] while $\lambda = +1$ is beneficial when dealing with the asymptotic behavior of $\boldsymbol{X}(t)$, which will be discussed further below.

2.3.2. Time evolution of the probability distribution

In general, no analytic expression exists for the transition PDF $f_{\boldsymbol{X}(t)|\boldsymbol{X}(t')}$ for finite increments of $\boldsymbol{X}$. However, by performing a *Kramers–Moyal expansion* (Kramers 1940; Moyal 1949) based on eq. (2.40), it can be shown that $f_{\boldsymbol{X}(t)|\boldsymbol{X}(t')}$ fulfills a *Fokker–Planck equation* (FPE)[10] (Fokker 1914; Planck 1917)

$$\frac{\partial}{\partial t} f_{\boldsymbol{X}(t)|\boldsymbol{X}(t')}(\boldsymbol{x}\,|\,\boldsymbol{x}') = L_{\boldsymbol{X}}(\boldsymbol{x})\, f_{\boldsymbol{X}(t)|\boldsymbol{X}(t')}(\boldsymbol{x}\,|\,\boldsymbol{x}') \quad \text{for} \quad t > t' \tag{2.55}$$

with the linear *Fokker–Planck operator* (FPO)

$$L_{\boldsymbol{X}}(\boldsymbol{x}) = \sum_{i=1}^{n} \frac{\partial}{\partial x^i}\left[-\mu^i_{\boldsymbol{X}|-}(\boldsymbol{x}) + \sum_{k=1}^{n} \frac{\partial}{\partial x^k} D^{ik}_{\boldsymbol{X}}(\boldsymbol{x})\right] \tag{2.56a}$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial x^i}\left[-\mu^i_{\boldsymbol{X}|\pm}(\boldsymbol{x}) + \tfrac{1}{2}\sum_{l=1}^{m} \sigma^{il}_{\boldsymbol{X}}(\boldsymbol{x}) \sum_{k=1}^{n} \frac{\partial}{\partial x^k} \sigma^{kl}_{\boldsymbol{X}}(\boldsymbol{x})\right] \tag{2.56b}$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial x^i}\left[-\mu^i_{\boldsymbol{X}|+}(\boldsymbol{x}) + \sum_{k=1}^{n} D^{ik}_{\boldsymbol{X}}(\boldsymbol{x}) \frac{\partial}{\partial x^k}\right] \tag{2.56c}$$

---

[9] Besides, the Itô stochastic integral is the mathematically best understood of the three, and many theorems are valid or available to date solely for $\lambda = -1$.

[10] Also known as *Kolmogorov forward equation.* Its counterpart, the *Kolmogorov backward equation* is a differential equation for $t < t'$ that allows to reconstruct the behavior of $\boldsymbol{X}(t)$ prior to some given *final/target state* $\boldsymbol{X}(t')$.

assuming that eq. (2.52) holds. Otherwise, eqs. (2.56a) to (2.56c) would not match and consequently describe different processes.

Since L_X does not depend on $\boldsymbol{x}'$, multiplying both sides of eq. (2.55) with $f_{X(t')}(\boldsymbol{x}')$ and integrating over all $\boldsymbol{x}'$ yields the same FPE for the marginal PDF $f_{X(t)}$,

$$f_{X(t)}(\boldsymbol{x}) = \int_{\mathrm{U}} \mathrm{d}^n\boldsymbol{x}'\, f_{X(t)|X(t')}(\boldsymbol{x}\,|\,\boldsymbol{x}')\, f_{X(t')}(\boldsymbol{x}')\,, \tag{2.57}$$

$$\frac{\partial}{\partial t} f_{X(t)}(\boldsymbol{x}) = L_X(\boldsymbol{x})\, f_{X(t)}(\boldsymbol{x})\,, \qquad t > t'\,. \tag{2.58}$$

This is a linear parabolic *partial differential equation* (PDE) with the formal solution

$$f_{X(t)}(\boldsymbol{x}) = \exp\big((t-t')\,L_X(\boldsymbol{x})\big)\, f_{X(t')}(\boldsymbol{x})\,, \qquad t > t'\,, \tag{2.59}$$

which can be solved numerically up to some final time t_1 if the coefficient functions, the initial state $f_{X(t_0)}$, and the function values of $f_{X(t)}$ on the codomain boundary $\partial\mathrm{U}$ are known for $t' = t_0 < t \leq t_1$. The FPE (2.58) then completely determines $f_{X(t)}$, which in turn encodes the entire information content of $X(t)$. Thus, FPE and SDE (2.53) offer an equivalent, complementary description of the underlying stochastic phenomenon.

2.3.3. Fluctuation–dissipation relations

Equation (2.58) can be rearranged as a continuity equation

$$\frac{\partial}{\partial t} f_{X(t)}(\boldsymbol{x}) + \nabla^{\mathsf{T}} \boldsymbol{j}_{X(t)}(\boldsymbol{x}) = 0 \tag{2.60}$$

for probability density $f_{X(t)}(\boldsymbol{x})$ and probability current density $\boldsymbol{j}_{X(t)}(\boldsymbol{x})$,

$$j^i_{X(t)}(\boldsymbol{x}) = \mu^i_{X|-}(\boldsymbol{x})\, f_{X(t)}(\boldsymbol{x}) - \sum_{k=1}^{n} \frac{\partial\big[D^{ik}_X(\boldsymbol{x})\, f_{X(t)}(\boldsymbol{x})\big]}{\partial x^k} \tag{2.61a}$$

$$= \mu^i_{X|\pm}(\boldsymbol{x})\, f_{X(t)}(\boldsymbol{x}) - \tfrac{1}{2} \sum_{l=1}^{m} \sigma^{il}_X(\boldsymbol{x}) \sum_{k=1}^{n} \frac{\partial\big[\sigma^{kl}_X(\boldsymbol{x})\, f_{X(t)}(\boldsymbol{x})\big]}{\partial x^k} \tag{2.61b}$$

$$= \mu^i_{X|+}(\boldsymbol{x})\, f_{X(t)}(\boldsymbol{x}) - \sum_{k=1}^{n} D^{ik}_X(\boldsymbol{x})\, \frac{\partial f_{X(t)}(\boldsymbol{x})}{\partial x^k} \tag{2.61c}$$

which mirrors the fact that the total probability is a conserved quantity,

$$\int_{\mathrm{U}} f_{X(t)}(\boldsymbol{x})\,\mathrm{d}^n\boldsymbol{x} = 1 \quad \forall\, t \in \mathbf{T}\,. \tag{2.62}$$

The two summands in eqs. (2.61a) to (2.61c) can be interpreted as advective and diffusive parts of the probability current density, $\boldsymbol{j}_{X(t)} = \boldsymbol{j}_{\mathrm{a},X(t)} + \boldsymbol{j}_{\mathrm{d},X(t)}$, where the former accounts for directed, deterministic and the latter for undirected, probabilistic sources of flux. This decomposition is not necessarily unique for non-constant diffusion coefficients, because whether the flow caused by gradients of $\boldsymbol{\sigma}_X$ is considered "deterministic" or "probabilistic" is a matter of perspective, which is represented on the stochastic level through the choice of a discretization scheme.[11] Notably, the post-point rule yields *Fick's first law*, where the diffusive part originates solely from the gradient of the probability density,

$$\boldsymbol{j}_{\mathrm{a},X(t)} = \boldsymbol{\mu}_{X|+}\, f_{X(t)}\,, \qquad\qquad \boldsymbol{j}_{\mathrm{d},X(t)} = -\boldsymbol{D}_X \nabla f_{X(t)}\,. \tag{2.63}$$

For some processes – marked in the following by the subscript "∞" –, the advective and diffusive probability current densities counterbalance each other, so that the net flux vanishes; a condition denoted as *detailed balance.* From the continuity equation (2.60), it immediately follows that any such X_∞ is *stationary*, which means that its member PDFs do not depend on time,

$$\boldsymbol{j}_{X_\infty(t)}(\boldsymbol{x}) = 0 \overset{(2.60)}{\implies} \frac{\partial}{\partial t} f_{X_\infty(t)}(\boldsymbol{x}) = 0 \quad \forall\, t \in \mathbf{T}\,. \tag{2.64}$$

Furthermore, *any* drift–diffusion process X governed by an identical SDE approaches the same distribution as X_∞ as time t progresses (see, *e. g.*, Hänggi and Thomas 1982, chapter 4),

$$\lim_{t'\to\infty} f_{X(t')}(\boldsymbol{x}) = f_{X_\infty(t)}(\boldsymbol{x}) \quad \forall\, t \in \mathbf{T}, \boldsymbol{x} \in \mathbf{U}\,, \tag{2.65}$$

where for simplicity it was assumed that X and X_∞ share the same parameter set $\mathbf{T}$ and member state space $\mathbf{U}$. This common *asymptotic* PDF will be denoted by $f_{X(\infty)}$ in the following (and similar for derived quantities), regardless of whether ∞ is actually an element of $\mathbf{T}$.

While the initial states of X and its asymptotic X_∞ usually differ, their coefficient functions – which are identical by construction, $\boldsymbol{\mu}_{X|\lambda} = \boldsymbol{\mu}_{X_\infty|\lambda}, \boldsymbol{\sigma}_X = \boldsymbol{\sigma}_{X_\infty}$ – are connected via their shared asymptotic PDF in the form of a *fluctuation–dissipation relation* (FDR),

$$\boldsymbol{j}_{X(\infty)}(\boldsymbol{x}) = 0 \overset{(2.61c)}{\iff} \boldsymbol{\mu}_{X|+}(\boldsymbol{x})\, f_{X(\infty)}(\boldsymbol{x}) = \boldsymbol{D}_X(\boldsymbol{x})\, \nabla f_{X(\infty)}(\boldsymbol{x})\,, \tag{2.66}$$

which takes this simple form only for $\lambda = +1$ and is otherwise a differential equation for the diffusivity $\boldsymbol{D}_X$ or diffusion $\boldsymbol{\sigma}_X$. Note that not all drift–diffusion SDEs allow for processes with detailed balance: For example, constant drift and diffusion coefficients always yield processes that are unbounded and do not converge against stationary states, which can be easily seen from the fact that eq. (2.41) diverges for large time increments.

---

[11] See also the discussion of "external" versus "internal" noise by van Kampen (1981).

The FDR can be used to find pairs of coefficient functions $\boldsymbol{\mu}_{X|+}, \boldsymbol{D}_X$ so that $\boldsymbol{X}$ is guaranteed to approach a given stationary state in detailed balance with PDF $f_{X(\infty)}$. This requires that the latter is non-zero in the region of interest, in which case eq. (2.66) is equivalent to[12]

$$\boldsymbol{\mu}_{X|+}(\boldsymbol{x}) = \boldsymbol{D}_X(\boldsymbol{x})\, \nabla \ln\!\big(f_{X(\infty)}(\boldsymbol{x})\big)\,. \tag{2.67}$$

Inserting the above into the expression for the post-point FPO (2.56c) then allows to express the FPE solely in terms of $\boldsymbol{D}_X$ and $f_{X(\infty)}$,

$$L_X(\boldsymbol{x}) = \sum_{i,k=1}^{n} L_X^{ik}(\boldsymbol{x})\,, \qquad L_X^{ik}(\boldsymbol{x}) := \frac{\partial}{\partial x^i}\, D_X^{ik}(\boldsymbol{x}) \left[-\frac{\partial \ln\!\big(f_{X(\infty)}(\boldsymbol{x})\big)}{\partial x^k} + \frac{\partial}{\partial x^k} \right], \tag{2.68}$$

where the operator naturally decomposes into $n \cdot n$ summands L_X^{ik}.

Conversely, if the coefficient functions are known for $\lambda = +1$ and $\boldsymbol{D}_X$ is invertible, the associated asymptotic PDF can be obtained by applying the *gradient theorem* to eq. (2.67), yielding

$$f_{X(\infty)}(\boldsymbol{x}) = \text{const.} \exp\!\left(\int \mathrm{d}\boldsymbol{x}^{\mathrm{T}} \boldsymbol{D}_X(\boldsymbol{x})^{-1}\, \boldsymbol{\mu}_{X|+}(\boldsymbol{x}) \right), \tag{2.69}$$

where the path of the line integral is chosen such that the integral exists. The integration constant in front of the exponential can be determined via the normalization condition (2.62) for $X(\infty)$.

In any case, the convergence of $f_{X(t)}$ against its asymptote $f_{X(\infty)}$ is exponential in t (see, *e. g.*, Ji, Shen, and Yi 2018), as already indicated by the form of the formal solution (2.59).

2.3.4. Drift–diffusion subprocesses

Sometimes, the components of an n-dimensional drift–diffusion process $\boldsymbol{X}$ can be divided into two subgroups $\boldsymbol{X}'$ and $\boldsymbol{X}''$ such that $\boldsymbol{X}'$ by itself is also a drift–diffusion process, which can be treated separately. For simplicity, this will be demonstrated here solely for $\boldsymbol{X}' := \big(X^1, \dots, X^k\big)$, $\boldsymbol{X}'' := \big(X^{k+1}, \dots, X^n\big)$. In this case, $\boldsymbol{X}'$ is a *drift–diffusion subprocess* of $\boldsymbol{X} = (\boldsymbol{X}', \boldsymbol{X}'')$ if the drift and diffusion coefficient functions associated with $\boldsymbol{X}'$ are constant with respect to $\boldsymbol{x}''$,

$$\frac{\partial \mu^i_{X|\lambda}(\boldsymbol{x})}{\partial x^j} = 0 \quad \text{and} \quad \frac{\partial \sigma^{il}_X(\boldsymbol{x})}{\partial x^j} = 0 \quad \text{for} \quad \begin{cases} i \in \{1, \dots, k\}\,, \\ j \in \{k+1, \dots, n\}\,, \end{cases} \tag{2.70a}$$

---

[12] If $X(t)$ represents a physical quantity, the argument of the logarithm has to be scaled by an appropriate constant to eliminate any physical dimensions.

and $\boldsymbol{X}'$ is not explicitly coupled to $\boldsymbol{X}''$ via the diffusivity,

$$D_{\boldsymbol{X}}^{ij}(\boldsymbol{x}) = 0 \quad \text{for} \quad i \in \{1, \dots, k\}, j \in \{k+1, \dots, n\}. \tag{2.70b}$$

The FPE (2.58) can then be rewritten as

$$\frac{\partial}{\partial t} f_{\boldsymbol{X}(t)}(\boldsymbol{x}) = L_{\boldsymbol{X}'}(\boldsymbol{x}')\, f_{\boldsymbol{X}(t)}(\boldsymbol{x}) + \sum_{i=k+1}^{n} \frac{\partial}{\partial x^i}\, j^i_{\boldsymbol{X}(t)}(\boldsymbol{x}), \tag{2.71}$$

where $L_{\boldsymbol{X}'}$ denotes an FPO defined by the coefficient functions

$$\mu^i_{\boldsymbol{X}'|\lambda}(\boldsymbol{x}') := \mu^i_{\boldsymbol{X}|\lambda}(\boldsymbol{x}), \quad \sigma^{il}_{\boldsymbol{X}'}(\boldsymbol{x}') := \sigma^{il}_{\boldsymbol{X}}(\boldsymbol{x}) \quad \text{for} \quad i = 1, \dots, k \tag{2.72}$$

and $\boldsymbol{j}_{\boldsymbol{X}(t)}$ is the probability current density from eq. (2.60). Integrating out $\boldsymbol{x}''$ on both sides of eq. (2.71) and applying the *divergence theorem* yields an FPE for the marginal PDF of $\boldsymbol{X}'$ with an additional surface-integral term,

$$\frac{\partial}{\partial t} f_{\boldsymbol{X}'(t)}(\boldsymbol{x}') = L_{\boldsymbol{X}'}(\boldsymbol{x}')\, f_{\boldsymbol{X}'(t)}(\boldsymbol{x}') + \underbrace{\int_{\partial \mathrm{S}''} \sum_{i=k+1}^{n} \mathrm{d}\varepsilon_i(\boldsymbol{x}'')\, j^i_{\boldsymbol{X}(t)}(\boldsymbol{x})}_{0}. \tag{2.73}$$

As the probability flux should vanish at the state-space boundary, the integral evaluates to zero, showing that $\boldsymbol{X}'$ is indeed a proper drift–diffusion process with the coefficient functions (2.72).

2.4. Past approaches to non-relativistic diffusion

To conclude this chapter, the mathematical formalism outlined so far will be demonstrated on the basis of two well-known physical models that were devised to describe *Brownian motion*; the jittery movement of small particles suspended in a medium, as observed by botanist Robert Brown in 1827 and published in 1828. In these examples, many calculations can still be performed analytically, which will not be the case for the more complex models presented in the following chapters.

2.4.1. Brownian motion in position space

Consider a drift–diffusion process $\boldsymbol{X}$ with vanishing drift coefficient and a diffusion coefficient equal to a scalar $(d \times d)$-matrix,

$$\boldsymbol{\mu}_{\boldsymbol{X}}(\boldsymbol{x}) = \boldsymbol{0}, \qquad \boldsymbol{\sigma}_{\boldsymbol{X}}(\boldsymbol{x}) = \sqrt{2D}\,\mathbb{1}, \tag{2.74}$$

for some constant $D > 0$, such that the diffusivity is $\boldsymbol{D}_{\boldsymbol{X}}(\boldsymbol{x}) = D\mathbb{1}$.

Despite being conceptually simple, this process is of great historical significance, as it provides a theoretical description of the d-dimensional movement of a Brownian particle in a resting, isotropic, homogeneous medium by identifying $X(t)$ with the particle position expressed in Cartesian coordinates. The theoretical study of this phenomenon by Einstein (1905b, 1906), Smoluchowski (1906, 1916), and Langevin (1908) provided the basis and motivation for the development of stochastic processes in the subsequent years. While mathematically equivalent, the calculations shown here are naturally different from those in the original articles, as the tools of stochastic calculus were not yet available at that time. Nonetheless, Einstein in fact used an approach involving random variables to model the particle displacements and derive a differential equation for the displacements' PDF, which is in essence close to the present-day method.

From eq. (2.41), it is known that the increments of $\boldsymbol{X}$ are normally distributed for arbitrarily large time steps Δt since $\boldsymbol{\mu}_{\boldsymbol{X}}$ and $\boldsymbol{\sigma}_{\boldsymbol{X}}$ are constant,

$$\boldsymbol{X}(t+\Delta t)-\boldsymbol{X}(t)=\sqrt{2D}\left[\boldsymbol{W}(t+\Delta t)-\boldsymbol{W}(t)\right]\sim\mathcal{N}(\boldsymbol{0},2D\Delta t\,\mathbb{1})\,, \tag{2.75}$$

such that eq. (2.42) takes the form[13]

$$f_{\boldsymbol{X}(t+\Delta t)|\boldsymbol{X}(t)}(\boldsymbol{x}+\Delta\boldsymbol{x}\,|\,\boldsymbol{x})=f_{\boldsymbol{X}(t+\Delta t)-\boldsymbol{X}(t)}(\Delta\boldsymbol{x})=\frac{1}{\sqrt{(4\pi D\Delta t)^d}}\exp\!\left(-\frac{(\Delta\boldsymbol{x})^2}{4D\Delta t}\right). \tag{2.76}$$

Notably, the above implies that the mean squared displacement of the Brownian particle is proportional to the elapsed time,

$$E_{\boldsymbol{X}}\!\left(\left[\boldsymbol{X}(t+\Delta t)-\boldsymbol{X}(t)\right]^2\right)=2D\Delta t\,. \tag{2.77}$$

This qualitative prediction is one of the main results of Einstein's[14] and Smoluchowski's work on this topic. It was later experimentally verified by Perrin (1909), which was an important step toward general acceptance of the previously controversial atomic theory of matter.

The FPE (2.58) of $\boldsymbol{X}$ simplifies, as expected, to the *diffusion equation*

$$\frac{\partial}{\partial t}f_{\boldsymbol{X}(t)}(\boldsymbol{x})=D\sum_{i=1}^{d}\frac{\partial^2}{(\partial x^i)^2}f_{\boldsymbol{X}(t)}(\boldsymbol{x}) \tag{2.78}$$

since $L_{\boldsymbol{X}}$ is simply a scaled Laplace operator in Cartesian coordinates. If the coordinate system is chosen such that the Brownian particle is initially located at the origin,

---

[13] Here and in the following, the shorthand $\boldsymbol{a}^2:=\boldsymbol{a}^{\mathsf{T}}\boldsymbol{a}$ is used for the left matrix product of a column vector $\boldsymbol{a}$ with its transpose $\boldsymbol{a}^{\mathsf{T}}$.

[14] Einstein also correctly predicted that D should be proportional to the temperature of the surrounding medium and inversely proportional to the particle radius and medium viscosity.

$f_{X(0)}(\boldsymbol{x}) = \delta(\boldsymbol{x})$, the entire process $\boldsymbol{X}$ is just a scaled standard Wiener process according to eq. (2.75),

$$\boldsymbol{X}(t) = \sqrt{2D}\,\boldsymbol{W}(t) \quad \text{for} \quad \boldsymbol{X}(0) = \boldsymbol{0}\,, \tag{2.79}$$

which explains why the Wiener process is also informally referred to as "Brownian motion". From this point of view, the invariance of $\boldsymbol{X}$ under right multiplications of its diffusion coefficient with normal matrices $\boldsymbol{Q}$, $\boldsymbol{\sigma_X} \mapsto \boldsymbol{\sigma_X Q}$, can intuitively be interpreted as an invariance of Brownian motion under rotations of the coordinate axes due to the isotropy of the system. Hence, if $\boldsymbol{X}$ is a Brownian motion, then $\hat{\boldsymbol{X}} \coloneqq \boldsymbol{QX}$ is also a Brownian motion, and this immediately extends to standard Wiener processes by choosing $D = \frac{1}{2}$.

The same system can also be studied in polar coordinates, which will be done here only for the case of two dimensions ($d = 2$) for reasons of simplicity. These new coordinates are connected to the previous Cartesian ones via the transition map

$$\boldsymbol{\Pi}(x^1, x^2) = \begin{pmatrix} \sqrt{(x^1)^2 + (x^2)^2} \\ \arg\bigl(x^1 + \mathrm{i}x^2\bigr) \end{pmatrix} \quad \Longleftrightarrow \quad \boldsymbol{\Pi}^{-1}(r, \phi) = \begin{pmatrix} r\cos(\phi) \\ r\sin(\phi) \end{pmatrix} \tag{2.80}$$

with the Jacobian matrix

$$\boldsymbol{J_\Pi}\bigl(\boldsymbol{\Pi}^{-1}(r,\phi)\bigr) = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ \frac{-\sin(\phi)}{r} & \frac{+\cos(\phi)}{r} \end{pmatrix} \eqqcolon \frac{\partial(r,\phi)}{\partial\bigl(x^1, x^2\bigr)}\,, \tag{2.81}$$

which, for convenience, is expressed here in polar coordinates. The movement of the Brownian particle is then equally well described by random fluctuations of the new coordinates $(R, \Phi)(t) \coloneqq \boldsymbol{\Pi}\bigl(\boldsymbol{X}(t)\bigr)$, where (R, Φ) is itself a drift–diffusion process with the diffusion and diffusivity coefficient functions

$$\boldsymbol{\sigma}_{R,\Phi}(r,\phi) \overset{(2.54\text{b})}{=} \sqrt{2D}\begin{pmatrix} \cos(\phi) & \sin(\phi) \\ \frac{-\sin(\phi)}{r} & \frac{+\cos(\phi)}{r} \end{pmatrix} \implies \boldsymbol{D}_{R,\Phi}(r,\phi) = D\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{r^2} \end{pmatrix}. \tag{2.82}$$

Note that both coefficients are not constant in the new coordinates, which means that the representation of (R, Φ) is no longer independent of the chosen discretization. As a result, only the mid-point drift coefficient remains zero,

$$\boldsymbol{\mu}_{R,\Phi|\pm}(r,\phi) \overset{(2.54\text{a})}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{2.83a}$$

while the other coefficients receive non-vanishing contributions from derivatives of $\boldsymbol{\sigma}_{R,\Phi}$,

$$\boldsymbol{\mu}_{R,\Phi|+}(r,\phi) \overset{(2.52\text{a})}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix} - D\begin{pmatrix} \cos(\phi) & \sin(\phi) \\ \frac{-\sin(\phi)}{r} & \frac{+\cos(\phi)}{r} \end{pmatrix}\begin{pmatrix} \frac{-\cos(\phi)}{r} \\ \frac{-\sin(\phi)}{r} \end{pmatrix} = D\begin{pmatrix} \frac{1}{r} \\ 0 \end{pmatrix}, \tag{2.83b}$$

$$\boldsymbol{\mu}_{R,\Phi|-}(r,\phi) \overset{(2.52b)}{=} D\begin{pmatrix}\frac{1}{r}\\0\end{pmatrix} + \begin{pmatrix}0\\0\end{pmatrix} = D\begin{pmatrix}\frac{1}{r}\\0\end{pmatrix}. \tag{2.83c}$$

The associated FPO in the pre- and post-point representation is then given by

$$L_{R,\Phi}(r,\phi) = D\left[-\frac{\partial}{\partial r}\frac{1}{r} + \frac{\partial^2}{(\partial r)^2} + \frac{1}{r^2}\frac{\partial^2}{(\partial\phi)^2}\right]. \tag{2.84}$$

Note that the above is *not* a (scaled) Laplace operator expressed in polar coordinates – that would be $L_{\boldsymbol{X}}\big(\boldsymbol{\Pi}^{-1}(r,\phi)\big)$ – but an entirely different object. This is because $L_{R,\Phi}$ is supposed to act on the transformed PDF

$$f_{(R,\Phi)(t)}(r,\phi) = \underbrace{\left|\det\!\big(\boldsymbol{J}_{\boldsymbol{\Pi}^{-1}}(r,\phi)\big)\right|}_{r}\, f_{\boldsymbol{X}(t)}\big(\boldsymbol{\Pi}^{-1}(r,\phi)\big)\,, \tag{2.85}$$

which is not only a re-parameterization of $f_{\boldsymbol{X}(t)}$, but also differs by a Jacobian determinant. In fact, it can be shown by straightforward calculation that

$$L_{R,\Phi}(r,\phi)\left|\det\!\big(\boldsymbol{J}_{\boldsymbol{\Pi}^{-1}}(r,\phi)\big)\right| = \left|\det\!\big(\boldsymbol{J}_{\boldsymbol{\Pi}^{-1}}(r,\phi)\big)\right| L_{\boldsymbol{X}}\big(\boldsymbol{\Pi}^{-1}(r,\phi)\big) \tag{2.86}$$

in the sense of an operator equation, which is similarly true for any other coordinate transformation.

2.4.2. Ornstein–Uhlenbeck process

Consider now a drift–diffusion process $\boldsymbol{V}$ that possesses a diffusion coefficient of the same form as in eq. (2.74) and a simple linear drift,

$$\boldsymbol{\mu}_{\boldsymbol{V}}(\boldsymbol{v}) = \frac{\bar{\boldsymbol{v}} - \boldsymbol{v}}{\tau}\,, \qquad \boldsymbol{\sigma}_{\boldsymbol{V}}(\boldsymbol{v}) = \sqrt{2D'}\,\mathbb{1}\,, \tag{2.87}$$

with constant *mean reversion rate* $1/\tau > 0$, *mean reversion level* $\bar{\boldsymbol{v}} \in \mathbb{R}^d$, and scalar diffusivity $D' > 0$.

Processes of this kind are named *Ornstein–Uhlenbeck processes* in honor of Uhlenbeck and Ornstein (1930), and were originally devised to model the velocity of Brownian particles by identifying $V^1(t), \ldots, V^d(t)$ with the Cartesian vector components of the particle's velocity[15] at time t. Small changes in the velocity are then normally distributed with a linear bias towards positive (negative) values if $\boldsymbol{V}(t)$ is smaller (larger) than $\bar{\boldsymbol{v}}$, which can be motivated as follows: Dividing the SDE of $\boldsymbol{V}$ by $\mathrm{d}t$ and multiplying both sides with the particle mass m

---

[15] The constant D' in eq. (2.87) is then not identical with D from eq. (2.74), as the former is a diffusivity in *velocity* space and the latter a diffusivity in *position* space.

yields a special form of *Newton's second law* for the Brownian particle's acceleration,

$$\begin{gathered} m\,\frac{\mathrm{d}\boldsymbol{V}(t)}{\mathrm{d}t} = \boldsymbol{F}\big(\boldsymbol{V}(t)\big) + \boldsymbol{\eta}(t)\,, \\ \boldsymbol{F}(\boldsymbol{v}) = m\,\frac{\bar{\boldsymbol{v}} - \boldsymbol{v}}{\tau}\,, \qquad \boldsymbol{\eta}(t) = \sqrt{2D'}\,m\,\frac{\mathrm{d}\boldsymbol{W}(t)}{\mathrm{d}t}\,, \end{gathered} \tag{2.88}$$

which is caused by a viscous force $\boldsymbol{F}$ due to *Stokes' law* and a noise term $\boldsymbol{\eta}$ that represents random collisions with the molecules of the medium. Consequently, m/τ is connected to the medium viscosity and $\bar{\boldsymbol{v}}$ is the average velocity of the medium in the observer's reference frame. Note, however, that eq. (2.88) is not mathematically well-defined, since Wiener processes are not differentiable. Therefore, it should rather be taken as a symbolic, more intuitive notation for the associated SDE, which is in turn defined via its stochastic integral.

An equation of this form was proposed by Langevin (1908) as an "infinitely more simple" alternative to Einstein's approach to Brownian motion: While Einstein derived a partial differential equation for the particle's probability distribution, Langevin formulated the above ordinary differential equation for the particle's trajectory, nowadays denoted as *Langevin equation.* The two different historical takes on Brownian motion by Einstein and Langevin already hint at the connection between the SDE (2.53) and its FPE (2.58) that is an integral part of the modern formulation. The analogy is not perfect here, though, as Einstein formulated his model for the particle's position while Langevin's model dealt with its velocity.

The SDE associated with eq. (2.87) can be solved analytically, *e. g.*, by variation of parameters, yielding

$$\boldsymbol{V}(t+\Delta t) - \boldsymbol{V}(t) \sim \mathcal{N}\Big(\big(1 - \mathrm{e}^{-\Delta t/\tau}\big)\,\big(\bar{\boldsymbol{v}} - \boldsymbol{V}(t)\big), D'\tau\big(1 - \mathrm{e}^{-2\Delta t/\tau}\big)\,\mathbb{1}\Big)\,. \tag{2.89}$$

For $D' \to 0$, this reproduces an exponential decay from $\boldsymbol{V}(t)$ to $\bar{\boldsymbol{v}}$ with mean lifetime τ, as one could already expect from the form of the SDE, which is essentially an exponential decay law with an additional disturbance from random noise.

It is obvious from eq. (2.89) that $\boldsymbol{V}(t)$ approaches a stationary limit for $t \to \infty$. For demonstration purposes, the PDF of this state is derived from the FDR of $\boldsymbol{V}$ by using eq. (2.69), resulting in

$$\begin{aligned} f_{\boldsymbol{V}(\infty)}(\boldsymbol{v}) &= \text{const.}\exp\!\left(\frac{1}{D'\tau}\int \mathrm{d}\boldsymbol{v}^{\mathrm{T}}(\bar{\boldsymbol{v}} - \boldsymbol{v})\right) \\ &= \left(\frac{1}{2\pi D'\tau}\right)^{d/2}\exp\!\left(-\frac{(\bar{\boldsymbol{v}} - \boldsymbol{v})^2}{2D'\tau}\right). \end{aligned} \tag{2.90}$$

In the second line, the normalizing constant could be determined immediately from eq. (2.9), because the exponential part corresponds to that of a normal distribution. This agrees with the limit $\Delta t \to \infty$ of eq. (2.89) for any initial state $\boldsymbol{V}(t_0)$ as the influence of the latter decreases exponentially with time.

From a physical point of view, it is to be expected that a prolonged contact with the surrounding medium leads to a thermalization of the Brownian particle, whose PDF would then be given by a *Maxwell–Boltzmann distribution*

$$f_{\boldsymbol{V}_{\mathrm{MB}}}(\boldsymbol{v}) := \left(\frac{m}{2\pi k_{\mathrm{B}} T}\right)^{d/2} \exp\left(-\frac{m\,(\boldsymbol{v}-\bar{\boldsymbol{v}})^2}{2k_{\mathrm{B}}T}\right) \overset{!}{=} f_{\boldsymbol{V}(\infty)}(\boldsymbol{v}) \tag{2.91}$$

with the *Boltzmann constant* k_{B} and medium temperature T. A comparison of eqs. (2.90) and (2.91) suggests that

$$D'\tau = \frac{k_{\mathrm{B}}T}{m}, \tag{2.92}$$

which is known as the *Einstein relation* and was discovered independently by Sutherland (1905), Einstein (1905b), and Smoluchowski (1906).

Conversely, one could also have started by postulating eq. (2.91) as asymptotic stationary PDF of $\boldsymbol{V}$ in combination with a constant scalar diffusivity and leaving the form of the drift coefficient function open. Then, applying eq. (2.67) yields

$$\boldsymbol{\mu}_{\boldsymbol{V}}(\boldsymbol{v}) = D'\,\nabla\left[-\frac{m\,(\boldsymbol{v}-\bar{\boldsymbol{v}})^2}{2k_{\mathrm{B}}T}\right] = -\frac{D'm}{k_{\mathrm{B}}T}\,(\boldsymbol{v}-\bar{\boldsymbol{v}})\,, \tag{2.93}$$

which is identical to eq. (2.87) with τ already expressed in terms of eq. (2.92).

The model can be extended into position space by embedding $\boldsymbol{V}$ into a new $2d$-dimensional stochastic process $(\tilde{\boldsymbol{X}}, \boldsymbol{V})$, where $\tilde{\boldsymbol{X}}$ fulfills the SDE

$$\mathrm{d}\tilde{\boldsymbol{X}}(t) = \boldsymbol{V}(t)\,\mathrm{d}t\,. \tag{2.94}$$

This means that $\tilde{\boldsymbol{X}}$ is the integral of $\boldsymbol{V}$ with respect to time, *i. e.*, the position of the Brownian particle, so that $(\tilde{\boldsymbol{X}}, \boldsymbol{V})$ is essentially[16] its phase-space trajectory. It is easy to see that $(\tilde{\boldsymbol{X}}, \boldsymbol{V})$ is again a drift–diffusion process with the coefficient functions

$$\mu^i_{\tilde{\boldsymbol{X}},\boldsymbol{V}}(\boldsymbol{x},\boldsymbol{v}) = \begin{cases} v^i & \text{if} \quad i \in \{1,\dots,d\}\,, \\ (\bar{v}^i - v^i)/\tau & \text{if} \quad i \in \{d+1,\dots,2d\}\,, \end{cases} \tag{2.95a}$$

$$\sigma^{il}_{\tilde{\boldsymbol{X}},\boldsymbol{V}}(\boldsymbol{x},\boldsymbol{v}) = \begin{cases} \sqrt{2D'} & \text{if} \quad l = i \in \{d+1,\dots,2d\}\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{2.95b}$$

---

[16] The phase space of a particle is usually formulated in terms of its (generalized) position and momentum. In this case, the latter differs from the particle's velocity only by a constant factor, the particle mass m.

Here, the position part $\tilde{\boldsymbol{X}}$ – unlike the velocity part $\boldsymbol{V}$ – does *not* meet the requirements of a drift–diffusion process when considered separately. This is a consequence of the fact that the SDE of $(\tilde{\boldsymbol{X}}, \boldsymbol{V})$ corresponds to a system of $2d$ coupled ODEs of first order, so that eliminating $\boldsymbol{V}$ amounts to rewriting the system as d second-order ODEs for $\tilde{\boldsymbol{X}}$. The latter cannot be represented by an SDE of the form (2.43), and hence, $\tilde{\boldsymbol{X}}$ is a classic example for a second-order Markov process.

A connection to the process $\boldsymbol{X}$ from the previous example can be established by calculating the mean squared displacement of $\tilde{\boldsymbol{X}}$: For $\bar{\boldsymbol{v}} = 0$ and $E_{\boldsymbol{V}}\big(\boldsymbol{V}(0)^2\big) = D'\tau$, it can be shown that (Fürth 1920; Uhlenbeck and Ornstein 1930)

$$E_{\tilde{\boldsymbol{X}}}\Big(\big[\tilde{\boldsymbol{X}}(t+\Delta t) - \tilde{\boldsymbol{X}}(t)\big]^2\Big) = 2D'\tau^3\left(\mathrm{e}^{-\Delta t/\tau} - 1 + \tfrac{\Delta t}{\tau}\right), \tag{2.96}$$

which reproduces eq. (2.77) for $\Delta t/\tau \gg 1$ if one identifies $D = D'\tau^2$. Hence, $\boldsymbol{X}$ can be seen as an approximation of $\tilde{\boldsymbol{X}}$ on long time scales, where the motion is primarily driven by the diffusive forces. On short time scales, however, the movement of the Brownian particle is uniform,

$$E_{\tilde{\boldsymbol{X}}}\Big(\big[\tilde{\boldsymbol{X}}(t+\Delta t) - \tilde{\boldsymbol{X}}(t)\big]^2\Big) \overset{\Delta t/\tau \ll 1}{\approx} D'\tau(\Delta t)^2 = E_{\boldsymbol{V}}\big(\boldsymbol{V}(0)^2\big)\,(\Delta t)^2\,, \tag{2.97}$$

since the influence of the noise is small in this regime.

The two non-relativistic diffusion models discussed above address the rather simple case of particle trajectories in Euclidean position and velocity space. These can be straightforwardly parameterized with Cartesian coordinates, which are then treated as stochastic drift–diffusion processes. In this way, the problem of dealing with probabilistic particle trajectories is reduced to the study of fluctuations of an associated set of coordinates.

Since the member state spaces of drift–diffusion processes do not necessarily have to be vector spaces, this approach can also be applied to more complicated scenarios: Particle trajectories can also be parameterized with non-Cartesian coordinates, as demonstrated above for polar coordinates. By extension, it is also possible to handle cases where the movement of the particles is restricted to some manifold due to conserved quantities or other constraints (van Kampen 1986), as in the case of Fermi surfaces or curved spacetime.

3. Applications to relativistic heavy-ion collisions

The following chapter is intended as a preparation for applications of the previously presented mathematical formalism in the context of high-energy particle physics; the discussion of specific models will be the subject of two separate, subsequent chapters. Here, some general aspects are addressed first, including the intricacies of relativistic diffusion, the choice of appropriate coordinates, and relevant characteristics of relativistic heavy-ion collisions. Furthermore, the common structure of the models to be presented is covered, and some details on the numerical approach are given.

3.1. Preliminary remarks on relativistic diffusion

The requirements of special relativity impose some constraints on the formulation of diffusion models in the high-energy regime; first and foremost that particles always propagate at subluminal speeds. This is clearly violated in the models for Brownian motion by Einstein (1906) and Smoluchowski (1906) presented in section 2.4, where the Brownian particle can travel arbitrarily large distances within finite time intervals due to the normal distribution of its increments (2.75). Moreover, the classical diffusion equation (2.78) contains temporal and spatial derivatives of different orders, which is at odds with Lorentz covariance. Although it provides an adequate description of non-relativistic particle dynamics, applications to relativistic systems are thus bound to fail.

Łopuszański (1953), Dudley (1966), and Hakim (1968) have shown that these problems have their root in a fundamental mathematical principle: Stochastic processes in position space cannot be both Lorentz-invariant and first-order Markovian. Therefore, if one does not want to forgo the mathematical conveniences of the Markov property, at least a second-order process must be used, as Langevin (1908) and Uhlenbeck and Ornstein (1930) did in their treatment of Brownian motion. While the coefficient functions used therein do not put an upper limit on the particle's velocity, this can be fixed quite easily by basing the model on the particle's momentum (Debbasch, Mallick, and Rivet 1997; Dunkel and Hänggi 2005a,b), which remains unbounded in special relativity.[1]

---

[1] If desired, an equivalent velocity-space description can then be obtained through substitution of variables, taking into account the transformation rules of stochastic calculus.

For this thesis, I generalize this "relativized" Ornstein–Uhlenbeck process to model diffusive particle motion as follows: Choosing Cartesian coordinate axes $\{\vec{e}_1, \dots, \vec{e}_d\}$, the particle's trajectory in $2d$-dimensional phase space is approximated by a first-order Markov process $(\boldsymbol{X}, \boldsymbol{P}) = \left(X^1, \dots, X^d, P^1, \dots, P^d\right)$ with the SDE

$$\left.\begin{aligned} \mathrm{d}X^i(t) &= \frac{P^i}{P^0}\, c\, \mathrm{d}t \\ \mathrm{d}P^i(t) &= \mu^i_{\boldsymbol{P}|\lambda}\big(\boldsymbol{P}(t)\big)\, \mathrm{d}t + \sum_{l=1}^{d} \sigma^{il}_{\boldsymbol{P}}\big(\boldsymbol{P}(t)\big) *_\lambda \mathrm{d}W^l(t) \end{aligned}\right\} \quad \text{for} \quad i = 1, \dots, d \tag{3.1}$$

describing the changes of the particle's position $\vec{x} \overset{!}{=} \sum_{i=1}^{d} X^i(t)\, \vec{e}_i$ and momentum $\vec{p} \overset{!}{=} \sum_{i=1}^{d} P^i(t)\, \vec{e}_i$ over time, where c denotes the speed of light. When considered separately, the particle's trajectory $\boldsymbol{X}$ in d-dimensional position space is then second-order Markovian, in accordance with the requirements of relativistic diffusion. Similar to the original Ornstein–Uhlenbeck process, the momentum coefficient functions are assumed to not depend on the position coordinates, so that $\boldsymbol{P}$ is a drift–diffusion subprocess driven by d independent Wiener processes, while the time evolution of the position is completely determined through its relation to relativistic momentum. Note that is assumption does *not* imply that the system is homogenous in position space, as this would require the initial state to be position-independent as well.

Equation (3.1) can be completed to a full $(2 + 2d)$-dimensional description of the particle by setting

$$X^0(t) := ct, \qquad\qquad P^0(t) := \sqrt{(mc)^2 + \sum_{i=1}^{d}\big(P^i(t)\big)^2}\,. \tag{3.2}$$

This relates the time parameter t to the coordinate time x^0 and establishes the mass-shell condition for p^0 with the particle's mass m, but does not introduce additional degrees of freedom. As there is no universal time in relativistic physics, the choice of a time parameter for the stochastic process singles out a preferred class of temporal hypersurfaces in Minkowskian space, which is used implicitly in the definition of the member PDFs $f_{(\boldsymbol{X},\boldsymbol{P})(t)}$ (Dunkel, Hänggi, and Hilbert 2009). For eqs. (3.1) and (3.2), these hypersurfaces are isochronous hyperplanes with $x^0 = ct$, which means that observables involving $f_{(\boldsymbol{X},\boldsymbol{P})(t)}$ (*e. g.*, the average momentum of the particle) are inherently non-local since they may contain information from space-like separated events. Later, the model results will be compared against data that are recorded (more or less) simultaneously in spatially extended detectors, so an isochronous hyperplane should provide a reasonable approximation of the experimental measuring process.

Physically reasonable choices for other use cases include setting t equal to the particle's proper time (see, *e. g.*, Hakim 1968; Dunkel, Hänggi, and Weber 2009) or defining observables with respect to the backward light cone of an observation event, which provides a "photographic measurement" of the system (Dunkel, Hänggi, and Hilbert 2009).

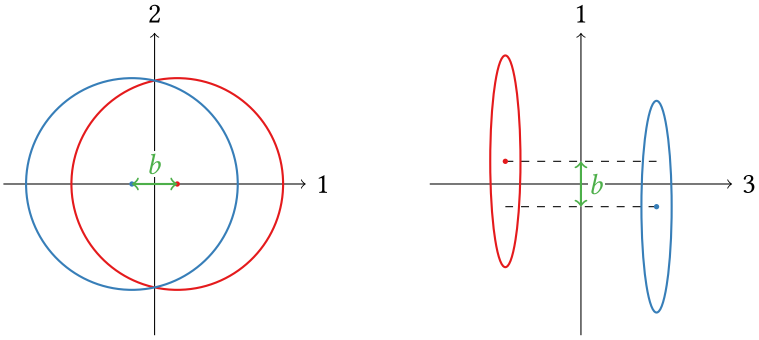
3.2. System and observables

In a relativistic heavy-ion collision, two nuclei moving with velocities close to the speed of light interpenetrate each other, triggering a cascade of interactions that results in the production of numerous subatomic particles. The ions are shattered in the process, separating their constituents each into a group of "spectators", which continue their journey unscathed, and a fragment containing the remnants of the "participants", which made contact with the opposing nucleus. Due to the high energy density, a hot fireball of partonic matter forms between the receding fragments (Bjorken 1983), which cools as it expands and eventually hadronizes in a parton–hadron crossover when its temperature falls below the critical Hagedorn temperature T_H (Hagedorn 1965, pp. 25–30).

The focus of this thesis lies on symmetric heavy-ion collisions in which the two nuclei are identical in terms of their numbers of protons Z and their numbers of neutrons $A - Z$. These collisions are most conveniently described in a center-of-momentum (COM) frame, where the heavy ions move in opposite directions with equal absolute momentum. If all protons and neutrons ("nucleons") participate in the collision, the energy available for particle production will then be equal to the total relativistic energy $\sqrt{s}$ in that frame. The quantity $\sqrt{s}$ is often given in the form $\sqrt{s_\mathrm{NN}} := \sqrt{s}/A$, which represents the maximum collision energy per nucleon pair.

Here and thereafter, the nuclei will be idealized as perfect spheres in their respective rest frames before the collision, neglecting any substructure due to individual nucleons or partons. In the COM frame, these spheres are Lorentz-contracted to oblate spheroids moving parallel to the *beam axis*; the latter is also referred to as the *longitudinal axis* and the plane perpendicular to it as the *transverse plane*. Projecting the geometric centers of the nuclei onto the transverse plane pre-collision yields two stationary points, whose distance is denoted as the *impact parameter b*.

In accordance with the usual convention, the following Cartesian coordinate system is used to describe multidimensional observables before and after the collision: Locating the origin at the geometric center of the system, the third spatial axis is placed parallel to the longitudinal axis so that the first and second spatial axis span the transverse plane. The orientation of the transverse axes is then chosen such that the impact parameter is measured along the first axis and the geometric centers of the nuclei lie in the so-called *reaction plane* spanned by the first and third axis. See fig. 3.1 for a schematic depiction of the system before the initial collision.



(a) transverse plane (b) reaction plane

Figure 3.1: Projection of the pre-collisional system onto the transverse plane (a) and reaction plane (b). Red and blue ellipses denote the forward- and backward-going nuclei, respectively. The impact parameter is indicated by a green arrow.

In the following, the discussion will be limited to “central” collisions, where the impact parameter is small compared to the transverse size of the nuclei. These systems possess an approximate rotational symmetry with respect to the longitudinal axis, which can be exploited by defining polar coordinates for the transverse plane,

$$p_\perp := \sqrt{\left(p^1\right)^2 + \left(p^2\right)^2}\,, \qquad \phi := \arg\left(p^1 + \mathrm{i}p^2\right), \tag{3.3a}$$

since any observable expressed in the new coordinate triple $\left(p_\perp, \phi, p^3\right)$ should be independent of ϕ in good approximation. In particular, this should be true for a particle trajectory in the COM frame that is described by the stochastic drift–diffusion process $\boldsymbol{P} = \left(P^1, P^2, P^3\right)$ defined in the preceding section: Applying the above coordinate transformation (3.3a) to P^1 and P^2 yields a drift–diffusion process $\left(P_\perp, \Phi, P^3\right)$ whose coefficient functions are independent of ϕ, such that $\left(P_\perp, P^3\right)$ is a drift–diffusion subprocess as discussed in section 2.3.4. The coordinate transformation’s Jacobian matrix involved in the substitution rules of stochastic calculus (*cf.* section 2.3.1) reads

$$\frac{\partial\left(p^1, p^2, p^3\right)}{\partial\left(p_\perp, \phi, p^3\right)} = \begin{pmatrix} \cos(\phi) & -p_\perp \sin(\phi) & 0 \\ \sin(\phi) & +p_\perp \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.3b}$$

For convenience, two further coordinate transformations will be performed, which substitute the remaining momentum coordinates with hyperbolic rapidity coordinates (see appendix B for their properties and a physical motivation):

Firstly, the coordinate p^3 is replaced by the longitudinal rapidity

$$y := \operatorname{artanh}\left(\frac{p^3}{p^0}\right), \tag{3.4a}$$

$$\frac{\partial(p_\perp, \phi, p^3)}{\partial(p_\perp, \phi, y)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{\sinh(y)}{\sqrt{1+(mc/p_\perp)^2}} & 0 & \sqrt{(mc)^2 + p_\perp^2}\cosh(y) \end{pmatrix}, \tag{3.4b}$$

to ease Lorentz boosts between the COM frame, the rest frames of the nuclei, and other inertial reference frames moving parallel to the beam axis. The particle trajectory expressed in these coordinates, $(P_\perp, \Phi, Y)$, can then be boosted longitudinally by simply adding the relative boost rapidity Δy to $Y(t)$ or subtracting it from the argument of the PDF,

$$f_{(P_\perp,\Phi,Y)(t)+(0,0,\Delta y)}\big(p_\perp, \varphi, y\big) = f_{(P_\perp,\Phi,Y)(t)}\big(p_\perp, \varphi, y - \Delta y\big)\,. \tag{3.5}$$

Secondly, $p_\perp$ is exchanged for the coordinate h defined via

$$h := \operatorname{arsinh}\left(\frac{p_\perp}{mc}\right), \qquad \frac{\partial(p_\perp, \phi, y)}{\partial(h, \phi, y)} = \begin{pmatrix} mc\cosh(h) & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{3.6}$$

which maps low transverse momenta to a linear and high transverse momenta to a logarithmic scale. This compresses the support of the PDF and thereby reduces the computing time for numerical solutions of the associated FPE.

Note that $(P_\perp, Y)$ and (H, Y) are drift–diffusion subprocesses of $(P_\perp, \Phi, Y)$ and (H, Φ, Y), respectively, since Φ was not involved in the respective coordinate transformations.

3.3. General model structure

In the final stage of a relativistic heavy-ion collision, the produced particles hit an array of surrounding detectors. From the recorded data, it is later possible to reconstruct the distribution of this particle shower in momentum space, rapidity space, or similar motional coordinate spaces, depending on the available detectors and the particle species included in the analysis. The associated particle *number density functions* (NDFS) will be denoted by $\mathrm{d}^nN/\mathrm{d}^nq$, where $(q^1, \dots, q^n) = \boldsymbol{q}$ signify the coordinate space, *e. g.*, $\boldsymbol{q} = (p_\perp, y)$, and the symbol N doubles as a label for the respective particle number, which can be obtained by integrating $\mathrm{d}^nN/\mathrm{d}^nq$ over the entire $\boldsymbol{q}$-space.

In the subsequent chapters, I will devise models for the distributions of certain subsets of produced particles that are closely linked to two characteristic processes in relativistic heavy-ion collisions: The stopping of protons (chapter 4) and the thermalization of charged hadrons (chapter 5). Principally, this is done by setting the respective particle NDFs equal to a PDF scaled by the associated particle number (but see below),

$$\frac{\mathrm{d}^n N}{\mathrm{d}^n q}(\boldsymbol{q}) \overset{!}{=} N f_{\boldsymbol{Q}(t_1)}(\boldsymbol{q}), \tag{3.7}$$

where $\boldsymbol{Q}$ is an n-dimensional drift–diffusion process that describes the particle trajectories in the coordinates $\boldsymbol{q}$, evaluated at some final time t_1. The process $\boldsymbol{Q}$ is constructed by choosing an initial state that is likely to describe the particles well at a given time $t_0 < t_1$ and deriving an FDR from their expected asymptotic state.

Here, it is important to notice that the FDR only fixes the *relationship* between the coefficient functions, so in addition either the drift or the diffusion has to be specified in some other way. For the time being, this gap is bridged by approximating the diffusivity in rapidity space as a constant diagonal matrix,

$$\boldsymbol{D}_{(H,Y)}(h,y) := \begin{pmatrix} D_\perp & 0 \\ 0 & D_\parallel \end{pmatrix}, \tag{3.8}$$

where, in contrast to the previously presented Ornstein–Uhlenbeck process, the diagonal elements may differ from each other. This approach is in line with those used in the previous, one-dimensional iterations of the model, and although it is merely a lowest-order approximation, it will prove sufficient to obtain reasonable results. For the future, it is of course preferable to determine the diffusivity coefficient function from microscopic theories or phenomenological considerations; for example, by fixing the time evolution of macroscopic observables as discussed by Dunkel and Hänggi (2009) for the case of binary elastic collisions in a heat bath. Specifically for the transverse direction, recent work by Caucal and Mehtar-Tani (2021) may be of interest in this regard.

In relativistic heavy-ion collisions, energy and particle number densities are highly inhomogeneous in position space, so it can be expected that particle dynamics differs with position. This is not covered by the Ornstein–Uhlenbeck-like ansatz (3.1), which assumes that the coefficient functions do not depend on any positional coordinates. To address this issue at least partly, the ansatz (3.7) is generalized by splitting the system in multiple disconnected subsystems ("sources") (Wolschin, Biyajima, Mizoguchi, and Suzuki 2006),

$$\frac{\mathrm{d}^n N}{\mathrm{d}^n q}(\boldsymbol{q}) \overset{!}{=} \sum_a N_a f_{\boldsymbol{Q}_a(t_1)}(\boldsymbol{q}) \quad \text{with} \quad \sum_a N_a = N, \tag{3.9}$$

for which different initial and asymptotic states can be chosen, thereby decoupling their time evolutions. Conceptually, each source can be thought to occupy a distinct region in phase space, with overlap between the different regions deemed negligible. Mathematically, eq. (3.9) can be obtained by modeling the coefficient functions as piece-wise constant in position space and calculating the marginal PDF in $\boldsymbol{q}$-space by integrating out the positional coordinates.

3.4. Numerical considerations

The final states of the drift–diffusion processes $\boldsymbol{Q}_a$ are obtained by solving their FPEs from the initial time t_0 to the final time t_1. This is done in transverse- and longitudinal-rapidity space, where the FPOs (2.68) of the individual sources take the particularly simple form

$$L_{(H,Y)_a}(h,y) = L^{11}_{(H,Y)_a}(h,y) + L^{22}_{(H,Y)_a}(h,y) \tag{3.10a}$$

$$L^{11}_{(H,Y)_a}(h,y) = D_\perp \frac{\partial}{\partial h}\left[-\frac{\partial \ln\big(f_{(H,Y)_a(\infty)}(h,y)\big)}{\partial h} + \frac{\partial}{\partial h}\right] \tag{3.10b}$$

$$L^{22}_{(H,Y)_a}(h,y) = D_\parallel \frac{\partial}{\partial y}\left[-\frac{\partial \ln\big(f_{(H,Y)_a(\infty)}(h,y)\big)}{\partial y} + \frac{\partial}{\partial y}\right] \tag{3.10c}$$

due to the above assumption of a constant diagonal diffusivity in rapidity space.

3.4.1. Dimensionless formulation

Since rapidities are physically dimensionless, the only dimensionful quantities remaining in the associated FPEs are the time parameter t and the diffusivity, having a dimension of time and inverse time, respectively. To eliminate these, I define a new, dimensionless evolution parameter δ via $t =: t_0 + \Delta t\,\delta$ with the interaction timespan $\Delta t := t_1 - t_0$, such that the FPEs take the form

$$\frac{\partial}{\partial \delta} f_{(H,Y)_a(t_0+\Delta t\,\delta)}(h,y) = \left[\Delta t\, L_{(H,Y)_a}(h,y)\right] f_{(H,Y)_a(t_0+\Delta t\,\delta)}(h,y)\,, \qquad \delta > 0\,, \tag{3.11}$$

which is then solved up to $\delta = 1$, corresponding to the final time t_1.

Due to the shape of the FPOs, the diffusivity coefficients and interaction timespan can be combined here to form the dimensionless products $D_\perp\,\Delta t$ and $D_\parallel\,\Delta t$, whose numerical values determine how close the final state is to the initial state ($D_\perp\,\Delta t, D_\parallel\,\Delta t \ll 1$) or asymptotic state ($D_\perp\,\Delta t, D_\parallel\,\Delta t \gg 1$). This incidentally exposes an intrinsic symmetry of diffusive systems: changing the interaction timespan while reciprocally adjusting drift and diffusivity leaves the final state unchanged. Mathematically, this is related to a certain self-similarity of the standard Wiener process $W(t)$: for any constant $\alpha > 0$, the process $\widetilde{W}(t) := \frac{1}{\sqrt{\alpha}}\,W(\alpha t)$ is also a standard Wiener process.

In eqs. (3.10) and (3.11), I tacitly assumed that the diffusivity coefficients $D_\perp, D_\parallel$ and the interaction timespan Δt agree for all sources, so that only the initial and asymptotic states differ. While this is certainly not true in general, I expect it to be fulfilled for the physical systems under consideration in this thesis; see chapters 4 and 5 for the model-specific motives behind this assumption. Note that in conjunction with the self-similarity discussed above, these constraints effectively reduce to $D_\perp \Delta t$ and $D_\parallel \Delta t$ being source-independent, which is somewhat less restrictive.

3.4.2. Solution algorithm

For each source, the FPE (3.11) is solved numerically with the method of lines: First, I discretize all derivative operators and functions appearing in eq. (3.10) as well as the initial PDF on an evenly-spaced grid in rapidity space. This yields a matrix representation of the associated FPO, which defines a multidimensional initial-value problem with respect to the evolution parameter δ. The latter is then solved with a Runge–Kutta method (Bogacki and Shampine 1989), interpolated, and transformed from rapidity to $\boldsymbol{q}$-space. Finally, the results from all sources are scaled and summed according to eq. (3.9).

By definition of the coordinates h and y, any PDF vanishes at $h = 0$ and for $h \to \infty$ or $|y| \to \infty$, which provides the boundary conditions for solving the FPEs. Due to the hyperbolic nature of the former, this is also fulfilled in good approximation for finite cut-off values h_{co} and $\pm y_{\text{co}}$, which I choose such that the truncated function values are smaller than 1 ‰ of the PDF's maximum value at all times. The boundary condition is then built directly into the discretized FPOs by modifying the matrix entries associated with $h = 0$, $h = h_{\text{co}}$, or $y = \pm y_{\text{co}}$.

3.4.3. Parameter estimation

Not all model parameters can be fixed via experimental or theoretical input. The remaining free parameters are estimated through the method of *weighted least squares* (WLS), with normalized residuals $r^1, \dots, r^l$ defined as

$$r^i(\boldsymbol{\pi}) := \frac{o^i - M(\mathbf{B}^i; \boldsymbol{\pi})}{\Delta o^i} . \tag{3.12}$$

Here, $o^1, \dots, o^l$ denote the experimentally observed particle number densities, $\Delta o^1, \dots, \Delta o^l$ the associated uncertainties, and $\mathbf{B}^1, \dots, \mathbf{B}^l$ the coordinate domains ("bins") of the measurements, while M is the theoretical model result based on the free parameters $\pi^1, \dots, \pi^k$. Since any observed density o^i is an average value for the coordinate bin $\mathbf{B}^i$, I set the model results equal to bin averages of the theoretical NDF (3.9),

$$M(\mathbf{B}^i;\boldsymbol{\pi}) := \frac{1}{\int_{\mathbf{B}^i} \mathrm{d}^n q} \int_{\mathbf{B}^i} \frac{\mathrm{d}^n N}{\mathrm{d}^n q}(\boldsymbol{q})\, \mathrm{d}^n q\,, \tag{3.13}$$

where the dependence on $\boldsymbol{\pi}$ is implicit in the definition of $\mathrm{d}^n N/\mathrm{d}^n q$.

The WLS method is based on the premise that the normalized residuals (3.12) can be seen as realizations of stochastic processes $R^1, \dots, R^l$ on the free-parameter space if one treats the measuring process as drawing a set of observations $o^1, \dots, o^l$ from a collection of normally-distributed random variables $O^1, \dots, O^l$ with variances $(\Delta o^1)^2, \dots, (\Delta o^l)^2$. Assuming that free parameters $\boldsymbol{\pi}$ exists for which the experimental data are well-described by the model, the associated process members $\boldsymbol{R}(\boldsymbol{\pi})$ are close to zero and roughly standard-normally distributed. Their quadratic sum should then follow a χ^2 distribution,

$$\sum_{i=1}^{l} \left(R^i(\boldsymbol{\pi})\right)^2 \sim \chi^2(\mathit{dof})\,, \tag{3.14}$$

where the number of degrees of freedom $\mathit{dof} \approx l - k$ results from the fact that the $\pi^1, \dots, \pi^k$ introduce dependencies between the $R^1, \dots, R^l$. In this case, the quadratic-sum realization[2] $\varsigma(\boldsymbol{\pi}) := \sum_{i=1}^{l} \left(r^i(\boldsymbol{\pi})\right)^2$ is expected to be close to the mean of the above distribution, *i. e.*, $\varsigma(\boldsymbol{\pi}) \approx \mathit{dof}$.

According to the Gauss–Markov theorem, the most appropriate estimate of $\boldsymbol{\pi}$ is given by the minimizer $\hat{\boldsymbol{\pi}}$ of ς in free-parameter space, which I determine with the Nelder–Mead method (Nelder and Mead 1965). To assess the *goodness of fit* (GOF), the reduced χ^2 statistic $\hat{\varsigma}/\mathit{dof}$ is used, which compares the minimum $\hat{\varsigma} := \varsigma(\hat{\boldsymbol{\pi}})$ with the χ^2-distribution mean. In this context, $\hat{\varsigma}/\mathit{dof} \approx 1$ is taken as an indication that the model provides an adequate description of the experimental data at hand for the parameter set $\hat{\boldsymbol{\pi}}$.

Note, however, that the actual number of degrees of freedom can be substantially less than $l - k$ if the number of observations l is small or the model depends non-linearly on $\boldsymbol{\pi}$ (see, *e. g.*, Andrae, Schulze-Hartung, and Melchior 2010), which will both be the case for the data and model under consideration. Therefore, the GOF assessment described above is more of a guideline than a hard criterion.

By considering small deviations from the minimizing parameter set $\hat{\boldsymbol{\pi}}$, one can derive a covariance matrix $\hat{\Sigma}$ in free-parameter space that provides estimates of the minimizer uncertainties $\Delta\hat{\pi}^1, \dots, \Delta\hat{\pi}^k$,

$$\hat{\Sigma} := \left(\boldsymbol{J}_r(\hat{\boldsymbol{\pi}})^{\mathsf{T}} \boldsymbol{J}_r(\hat{\boldsymbol{\pi}})\right)^{-1}, \qquad \Delta\hat{\pi}^i := \sqrt{\hat{\Sigma}^{ii}} \text{ for } i = 1, \dots, k. \tag{3.15}$$

---

[2] Often, the symbol “χ^2” is used to denote this realization, but I will use a different symbol here to avoid confusion with the distribution of the same name.

Here, a small uncertainty $\Delta\hat{\pi}^i$ implies that the parameter in question has to be close to the minimizer $\hat{\pi}^i$ to obtain a good agreement between observations and model results, while a large uncertainty suggests that the data and model at hand cannot provide a definite statement about the value of π^i.

Being a first order approximation, the uncertainties obtained from eq. (3.15) may deviate from the respective confidence intervals if $\Delta\hat{\pi}^i$ becomes large. However, since only qualitative discussions of fitting results are intended in this thesis, this does not pose a problem.

All numerical routines required for comparison with data are implemented in the Julia programming language (Bezanson, Edelman, Karpinski, and Shah 2017). Operator discretization and the solution of the resulting initial-value problems are performed with the help of the `DifferentialEquations.jl` package suite (Rackauckas and Nie 2017). A performant implementation of the minimization algorithm used is provided by the package `Optim.jl` (Mogensen and Riseth 2018).

4. Baryon stopping as a diffusive process

The participating protons and neutrons contained in the two colliding nuclei are slowed down (“stopped”) as they interpenetrate each other in the early stages of a relativistic heavy-ion collision; a process which is commonly denoted as “baryon stopping”.

I expect that the dynamical behavior of these baryons is dominated by collisions with partons from the opposing nucleus, so that interactions between baryons of the same nucleus can be neglected and the baryon distribution function reduces to a superposition of single-particle PDFs. Moreover, assuming that the density of partons (namely: valence quarks and gluons) is high enough, the opposing nucleus can be approximated as a fluctuating continuum that acts on the baryons via deterministic and stochastic forces, which leads to a diffusive baryon movement.

This ansatz makes use of the so-called *parton–hadron duality*: As the participating baryons break up into their partons during the slow-down, the resulting trajectories are to be interpreted as those of pseudo-particles that are not real baryons most of the time, but merely groups of loosely connected partons. The affiliation of these partons to individual pseudo-particles can change until they recombine into new real baryons in the later stages of the collision.

As will become apparent later, the above reasoning also separates the baryons into two largely independent “sources” associated with the two former nuclei. These will be referred to as the *forward-* (“+”) and *backward-going source* (“−”), respectively, based on the signs of their initial longitudinal rapidities in the COM frame.

An earlier version of this model (Hoelck and Wolschin 2020) mostly neglected the transverse motion of the baryonic pseudo-particles to obtain an effective description in y-space. It has its roots in a series of models by Wolschin (1999, 2004) in which Ornstein–Uhlenbeck processes as well as modified FPEs with thermal FDRs based on the Einstein relation (Forndran and Wolschin 2017) were successfully used towards the same end.

4.1. Initial state: Nuclear Fermi gas

Prior to the collision, the participants are part of an atomic nucleus, which I will model here as a zero-temperature nuclear Fermi gas. In the rest frame $\mathfrak{F}_{\pm}$ of the respective nucleus, the momentum $\vec{p}$ of each participant is then uniformly distributed in a sphere with a radius equal to the Fermi momentum p_{F},

$$f_{P_{\mathrm{NFG}}}(p^1, p^2, p^3) := \frac{3}{4\pi p_{\mathrm{F}}^3}\,\Theta\big(p_{\mathrm{F}} - \|\vec{p}\|\big) \tag{4.1}$$

where Cartesian momentum coordinates are assumed and Θ denotes the Heaviside step function

$$\Theta(x) := \frac{1 + \operatorname{sgn}(x)}{2} = \begin{cases} 1 & \text{if} \quad x > 0 \\ \frac{1}{2} & \text{if} \quad x = 0 \\ 0 & \text{if} \quad x < 0 \end{cases} \tag{4.2}$$

based on the sign function sgn. The Fermi momentum is determined by approximating the nuclear forces through an infinite potential well, yielding

$$p_{\mathrm{F}} := \hbar \sqrt[3]{3\pi^2 \frac{Z}{V}}\,, \qquad V := \frac{4\pi}{3} r^3\,, \tag{4.3}$$

where Z is the nucleus' proton number and V its nuclear charge volume. The latter is calculated from the nuclear charge radius r taken from the table of experimental nuclear ground state charge radii provided by Angeli and Marinova (2013). Here, the nuclides ^{197}Au and ^{208}Pb are of primary interest, since all data used for comparison in this thesis are from (symmetric) collisions of those two nuclear species.

Equation (4.1) can be equivalently formulated in terms of the (relativistic) Fermi energy E_{F},

$$f_{P_{\mathrm{NFG}}}(p^1, p^2, p^3) = \frac{3}{4\pi p_{\mathrm{F}}^3}\,\Theta\big(E_{\mathrm{F}} - cp^0\big)\,, \qquad E_{\mathrm{F}} := \sqrt{m^2c^4 + c^2 p_{\mathrm{F}}^2}\,, \tag{4.4}$$

with the shorthand cp^0 for the participant's energy (*cf.* section 3.1) in $\mathfrak{F}_\pm$. The corresponding PDF in transverse and longitudinal rapidity coordinates reads

$$f_{(H,\Phi,Y)_{\mathrm{NFG}}}(h, \phi, y) = \frac{3}{4\pi}\,\frac{\sinh(h)\cosh(h)^2\cosh(y)}{\sinh(w_{\mathrm{F}})^3}\,\Theta\big(\cosh(w_{\mathrm{F}}) - \cosh(h)\cosh(y)\big) \tag{4.5}$$

with the Fermi rapidity $w_{\mathrm{F}} := \operatorname{arsinh}\big(p_{\mathrm{F}}/mc\big)$. Due to the cylindrical symmetry, integrating out the angular coordinate ϕ merely adds a factor of 2π; the resulting PDF, together with its marginal PDFs in h and y, is shown in fig. 4.1.

Finally, the initial PDFs for the forward- and backward-going participants are obtained by boosting from $\mathfrak{F}_\pm$ to the COM frame,

$$f_{(H,Y)_\pm(t_0)}(h, y) \overset{!}{=} f_{(H,Y)_{\mathrm{NFG}}}(h, y \mp y_{\mathrm{b}})\,, \tag{4.6}$$

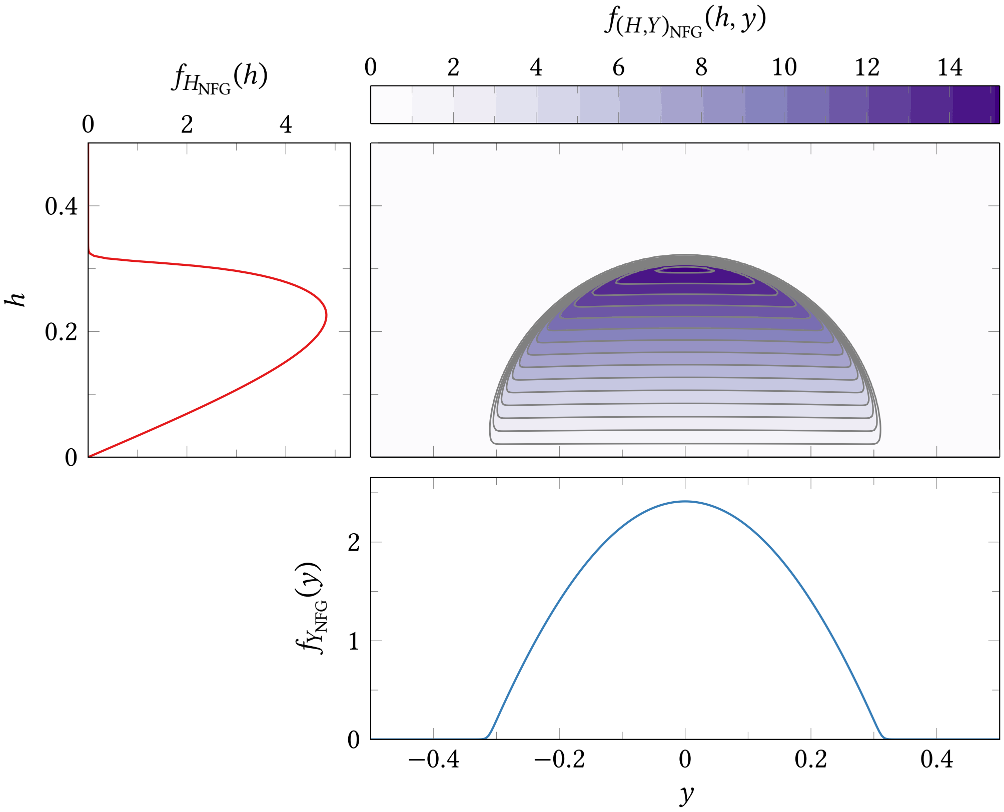


Figure 4.1: Joint (top right) and marginal (top left, bottom right) PDFs of a proton from a nuclear Fermi gas, expressed in the rapidity coordinates $y = \operatorname{artanh}(p^3/p^0)$ and $h = \operatorname{arsinh}(p_\perp/(mc))$, as defined in section 3.2. The joint PDF is depicted as a filled contour plot; refer to the color bar above for the associated function values. The Fermi rapidity used, $w_\mathrm{F} \approx 0.313$, represents a ^{197}Au nucleus; the difference to ^{208}Pb ($w_\mathrm{F} \approx 0.314$) is negligibly small.

where y_b denotes the (absolute) relative longitudinal rapidity between the two frames, which is commonly referred to as the *beam rapidity*. Since each nucleus carries an energy of $\sqrt{s}/2$ in the COM frame, the latter can be obtained from (see appendix B)

$$y_\mathrm{b} = \operatorname{arcosh}\left(\frac{\sqrt{s_\mathrm{NN}}}{2m_\mathrm{N}}\right), \qquad m_\mathrm{N} \approx \frac{Zm_\mathrm{p} + (A-Z)\,m_\mathrm{n}}{A}, \tag{4.7}$$

where the average nucleon mass m_N is approximated as a weighted average of the proton mass $m_\mathrm{p} \approx 938.3\,\mathrm{MeV}/\mathrm{c}^2$ and neutron mass $m_\mathrm{n} \approx 939.6\,\mathrm{MeV}/\mathrm{c}^2$, neglecting the binding energy of the nucleus.

4.2. Asymptotic state: Deep inelastic scattering with gluon saturation

Earlier investigations by Mehtar-Tani and Wolschin (2009a,b) suggest that the participants rapidly escape the collision zone without experiencing a substantial amount of interactions with non-participant hadrons, because their experimentally measured distribution functions seem to be primarily shaped by the initial binary collisions with the opposing nucleons. Therefore, for the time evolution of the participants' trajectories, I will choose an FDR that represents the latter interactions. The associated asymptotic PDF is derived from a QCD-inspired phenomenological formalism based on gluon saturation in *deep inelastic scattering* (DIS), developed with notable contributions from Gribov, Levin, and Ryskin (1983), Mueller and Qiu (1986), Blaizot and Mueller (1987), and McLerran and Venugopalan (1994). In DIS, it is assumed that the nucleon–nucleon interactions are mediated by quark–anti-quark pairs ("dipoles") emitted by the partons of the participating nucleons, which then scatter off the partons of the opposing nucleons. Recombination of the involved partons finally leads to the formation of new hadrons that are emitted from the fragments. Although only the production of new nucleons is of interest for the model at hand, the following discussion will cover hadrons of any kind, because the general case will be needed in the subsequent chapter.

Initially, the transverse motion of the nucleons and their partons can be disregarded in good approximation in the COM frame, so that they move strictly parallel to the beam axis. The momenta of any two forward- and backward-going partons then read

$$\vec{p}_+ \approx x_+ m_\mathrm{N} c \sinh(y_+)\,\vec{e}_3\,, \qquad \vec{p}_- \approx x_- m_\mathrm{N} c \sinh(y_-)\,\vec{e}_3\,, \tag{4.8}$$

where $0 \leq x_+, x_- \leq 1$ are the respective parton's Bjorken-x momentum fractions and y_+, y_- denote the longitudinal rapidities of the containing nucleons. For simplicity, the masses of the latter are approximated here by the average nucleon mass (4.7), which introduces an error of less than 2 ‰. Neglecting the comparatively small parton masses yields

$$p_+^0 \approx x_+ m_\mathrm{N} c \sinh\big(|y_+|\big)\,, \qquad p_-^0 \approx x_- m_\mathrm{N} c \sinh\big(|y_-|\big)\,, \tag{4.9}$$

for the associated energies.

If these two partons combine to form a new hadron with mass m, transverse rapidity h, and longitudinal rapidity y, energy-momentum conservation demands that energy and longitudinal momentum of the produced particle should match the summed initial parton

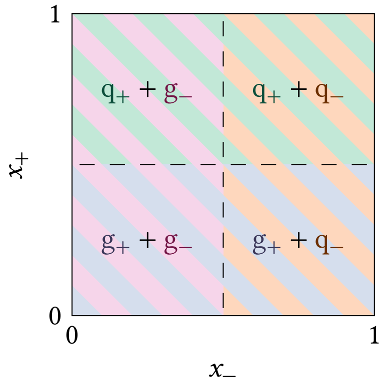


Figure 4.2: Dominant processes in the deep inelastic scattering of partons in relativistic heavy-ion collisions. The symbols "g" and "q" denote gluons and valence quarks, respectively; their affiliations to the forward- and backward-going nucleons are indicated by subscripts.

energies and momenta, respectively,

$$mc\cosh(h)\cosh(y) \overset{!}{\approx} m_{\mathrm{N}}c\left[x_+ \sinh\big(|y_+|\big) + x_- \sinh\big(|y_-|\big)\right] , \tag{4.10a}$$

$$mc\cosh(h)\sinh(y) \overset{!}{\approx} m_{\mathrm{N}}c\left[x_+ \sinh(y_+) + x_- \sinh(y_-)\right] . \tag{4.10b}$$

Note, however, that sources of transverse momentum are ignored here, as are any additional valence quarks that may be needed, making this more of an approximate statement. Adding and subtracting these two equations yields

$$x_\pm \approx \frac{m\cosh(h)\exp(\pm y)}{2m_{\mathrm{N}}\sinh\big(|y_\pm|\big)} , \tag{4.11}$$

where it was used that $\operatorname{sgn}(y_\pm) = \pm 1$ in the COM frame of relativistic heavy-ion collisions. When eq. (4.11) is used to find a Bjorken x, one has to keep in mind that not all combinations of masses and rapidities are valid, since the phase space of the produced hadron is restricted by the maximum available parton energies. If x_+ or x_- exceeds one, the envisaged particle cannot be produced by the nucleons at hand.

At high Bjorken x, the nucleon momentum is mostly carried by its valence quarks, whereas gluons prevail at low values of x. In combination, this yields $2 \cdot 2 = 4$ dominant interaction processes for DIS in the (x_+, x_-)-plane, as shown qualitatively in fig. 4.2. The hadron yield of the valence-quark–valence-quark process is expected to be small in comparison with the other three processes and can thus be safely neglected here. Further, the gluon–gluon process results predominantly in hadronic particle–anti-particle pairs with low to intermediate longitudinal rapidities, which are not of interest in the context of baryon stopping. This leaves two processes, namely the scattering of valence quarks onto gluons and the scattering of gluons onto valence quarks, which is the same process seen from the opposite perspective. As the groups of involved partons are disjunct (see color patterns in fig. 4.2), the hadrons produced from the two processes will be allocated to separate sources (Wolschin 1999). Since their kinematic properties are largely inherited from the valence quarks in the DIS formalism, hadrons produced from scatterings of forward-going valence quarks are assigned to the forward-going source, and *vice versa.*

The PDF of a produced hadron can then be approximated as a product of the distribution function of the scattering valence quarks q_v and their interaction cross-section with the opposing gluons φ (Kharzeev, Kovchegov, and Tuchin 2004; Baier, Mehtar-Tani, and Schiff 2006; Dumitru, Hayashigaki, and Jalilian-Marian 2006),

$$\begin{aligned} f_{(P_\perp,Y)_\pm(\infty)}\big(p_\perp, y\big) \overset{!}{=} f_{(P_\perp,Y)_{\mathrm{DISq}_\pm\mathrm{g}_\mp}}\big(p_\perp, y\big) &\coloneqq \text{const.}\; p_\perp \\ &\times \int_{-\infty}^{+\infty} x_\pm\, q_\mathrm{v}\big(x_\pm\big)\, \Theta\big(1 - x_\pm\big)\, f_{Y_\mathrm{NFG}}\big(y_\pm \mp y_\mathrm{b}\big)\, \mathrm{d}y_\pm \\ &\times \int_{-\infty}^{+\infty} \varphi\big(x_\mp; p_\perp\big)\, \Theta\big(1 - x_\mp\big)\, f_{Y_\mathrm{NFG}}\big(y_\mp \pm y_\mathrm{b}\big)\, \mathrm{d}y_\mp\,, \end{aligned} \tag{4.12}$$

where x_+ and x_- implicitly depend on $p_\perp$, y, y_+, and y_-. Heaviside step functions ensure that only valid production processes with $0 \le x_\pm \le 1$ enter the PDFs (see above). The valence-quark distribution function

$$q_\mathrm{v}(x) \coloneqq u_\mathrm{v}(x) + d_\mathrm{v}(x) \tag{4.13}$$

is taken from Martin, Roberts, Stirling, and Thorne (2002), where the contributions from up and down quarks are weighted according to the number of protons and neutrons in the colliding nuclei. For the scattering amplitude of the emitted quark–anti-quark dipole with the opposing gluons, an approximation by Dumitru, Hayashigaki, and Jalilian-Marian (2006) is employed,

$$\varphi\big(x; p_\perp\big) \coloneqq \frac{4\pi}{\frac{4}{9}Q_\mathrm{s}^2(x)} \exp\left(-\frac{p_\perp^2}{\frac{4}{9}Q_\mathrm{s}^2(x)}\right)(1-x)^4\,, \tag{4.14}$$

which is based on the expectation that the gluon density of the opposing nucleons saturates below a characteristic momentum scale Q_s, such that the gluons form a dense medium called *color-glass condensate* (CGC). For small values of x, the (quadratic) gluon saturation scale can then be parameterized (Golec-Biernat and Wüsthoff 1998) as

$$Q_\mathrm{s}^2(x) = \sqrt[3]{A}\, Q_0^2 \left(\frac{x_0}{x}\right)^\lambda, \tag{4.15}$$

where $Q_0 \coloneqq 1\,\mathrm{GeV}/c$ sets the physical dimension, while x_0 and λ are to be determined experimentally. The leading factor $\sqrt[3]{A}$ explicitly incorporates that the saturation momentum should scale roughly with the cubic root of the mass number A of the nuclei (Kharzeev and Levin 2001). To reflect the fact that the gluon density decreases rapidly with increasing Bjorken x, but also to suppress contributions from the high-x regime where the premises of the CGC formalism are no longer valid, an additional factor $(1-x)^4$ is included in eq. (4.14) (Kharzeev, Kovchegov, and Tuchin 2004).

Due to the Fermi motion in the initial state, the longitudinal rapidities of the nucleons $y_\pm$ do not exactly coincide with the beam rapidities $\pm y_\mathrm{b}$, but may deviate from them up to the Fermi rapidity w_F. This is taken into account in eq. (4.12) by including the initial marginal distribution of the longitudinal rapidities,

$$f_{Y_\mathrm{NFG}}(y) = \int_0^\infty f_{(H,Y)_\mathrm{NFG}}(h,y)\,\mathrm{d}h \tag{4.16a}$$

$$\overset{(4.5)}{=} \frac{1}{2}\left[\left(\frac{\cosh(w_\mathrm{F})}{\cosh(y)}\right)^3 - 1\right]\frac{\cosh(y)}{\sinh(w_\mathrm{F})^3}\,\Theta\big(w_\mathrm{F}-|y|\big)\,, \tag{4.16b}$$

and integrating the entire expression with respect to the nucleon rapidities. Often, the initial Fermi motion is neglected, $f_{Y_\mathrm{NFG}}(y) \approx \delta(y)$, which does not entail a great quantitative change of eq. (4.12). In the context of this thesis, however, the small longitudinal broadening by w_F is crucial, because otherwise the FDR is not defined on the entire support of the initial PDF (4.5). This becomes clear when considering the upper bound of the parton momentum fractions (4.11) for $m \approx m_\mathrm{N}$ and $|y|, |y_\pm| \gg 1$, which implies

$$1 \gtrsim \cosh(h)\exp\big(\pm y - |y_\pm|\big) \quad\Longrightarrow\quad |y_\pm| \gtrsim |y|\,. \tag{4.17}$$

Putting everything together and performing a coordinate transformation from $p_\perp$ to h then yields the distribution shown in fig. 4.3.

4.3. Results

Participant-baryon distributions are not directly accessible experimentally, because the stopped nucleons are indistinguishable from the multitude of other protons and neutrons emitted by the fireball. However, since the latter produces particles and anti-particles in almost equal numbers, this contribution should cancel out when the particle number densities of measured protons and anti-protons are subtracted from each other. The resulting net-proton densities can therefore be expected to provide a reasonable approximation for the distribution of stopped protons.

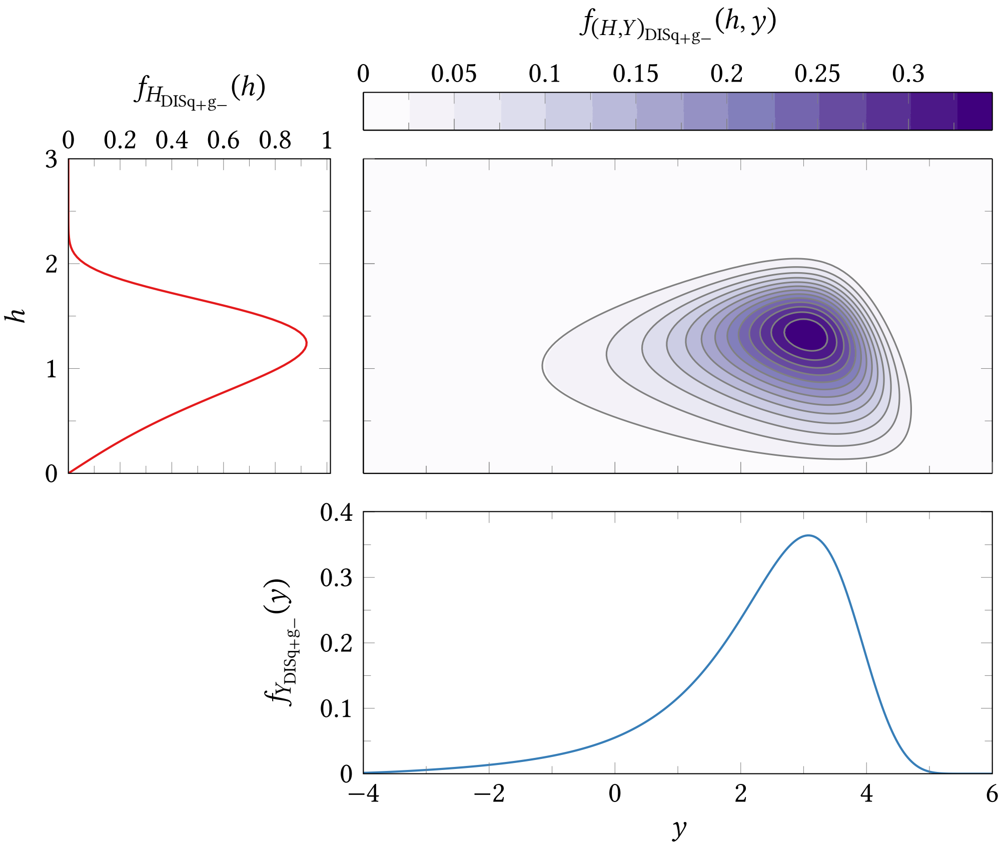


Figure 4.3: Joint (top right) and marginal (top left, bottom right) PDFs of a proton produced in the DIS of forward-going valence quarks and backward-going saturated gluons for a collision of two ^{197}Au nuclei with $\sqrt{s_{\mathrm{NN}}} = 200\,\mathrm{GeV}$ ($y_{\mathrm{b}} \approx 5.36$). For the gluon-saturation parameters, the values $\lambda = 0.2$ and $x_0 = 10^{-3}$ are used (*cf.* section 4.3). A smaller x_0 shifts the PDF towards $h = 0$; if λ is increased, the distribution becomes steeper.

Three net-proton data sets with different collision energies are used in the comparison with model results: One by the NA49[1] collaboration (Appelshäuser *et al.* 1999) obtained from ^{208}Pb collisions with $\sqrt{s_{\mathrm{NN}}} = 17.2\,\mathrm{GeV}$ performed at the SPS and two data sets from ^{197}Au collisions with $\sqrt{s_{\mathrm{NN}}} = 62.4\,\mathrm{GeV}$ and $200\,\mathrm{GeV}$ measured by the BRAHMS[2] collaboration (Arsene *et al.* 2009; Bearden *et al.* 2004) at RHIC. These data contain only contributions from collisions with centralities[3] less than 5 % and 10 %, respectively, so

---

[1] North Area Experiment 49

[2] Broad Range Hadron Magnetic Spectrometers Experiment

[3] The centrality of a heavy-ion collision with impact parameter b is defined as $\sigma_{\mathrm{inel}}(b)/\sigma_{\mathrm{inel}}(\infty)$, where $\sigma_{\mathrm{inel}}(b)$ is its integrated inelastic cross-section for the interval $[0, b]$. As such, its is a number between zero and one that increases monotonically with b.

Table 4.1: Properties of the three net-proton data sets used, complemented by estimates of the average numbers of participating nucleons $\langle A_{\text{part}} \rangle$ and protons $\langle Z_{\text{part}} \rangle$.

data set	$\sqrt{s_{\text{NN}}}$	nuclei	y_{b}	centrality	$\langle A_{\text{part}} \rangle$	$\langle Z_{\text{part}} \rangle$
NA49 (1999)	17.2 GeV	^{208}Pb	2.90	0 – 5 %	352 ± 12	138.8 ± 4.7
BRAHMS (2009)	62.4 GeV	^{197}Au	4.20	0 – 10 %	314 ± 8	125.9 ± 3.2
BRAHMS (2004)	200 GeV	^{197}Au	5.36	0 – 5 %	357 ± 8	143.2 ± 3.2

that cylindrical symmetry holds in good approximation. The aforementioned authors also provide estimates of the average number of participants $\langle A_{\text{part}} \rangle$ for the collisions contained in the data sets, from which the average number of participant protons can be obtained via $\langle Z_{\text{part}} \rangle \coloneqq \langle A_{\text{part}} \rangle \, Z/A$. An overview of all the aforementioned data-set properties is given in table 4.1. Note that $\langle A_{\text{part}} \rangle$ and $\langle Z_{\text{part}} \rangle$ are slightly lower for the BRAHMS (2009) data set because the colliding nuclei overlap less on average here due to the inclusion of larger impact parameters. The value reported by NA49 (1999) only includes particles with $|y| \leq 2.5$ and thus may slightly underestimate the actual number of participants.

All three data sets provide (joint) net-proton number densities in transverse-momentum and rapidity space, as well as associated marginal rapidity densities. Since the detectors do not cover the entire $p_\perp$ range, however, the latter are necessarily biased by assumptions made to extrapolate the low- and high-momentum regions. Therefore, only the joint densities are used in the parameter fitting, and comparisons with the marginal densities are kept at a qualitative level.

In all my numerical computations, the Heaviside step function Θ in the initial-state PDF (4.5) is replaced by the smooth logistic function

$$\ell_s(x) \coloneqq \frac{1 + \tanh\!\left(\frac{x}{2s}\right)}{2} = \frac{1}{1 + \exp\!\left(-\frac{x}{s}\right)}, \qquad s > 0 \tag{4.18}$$

since the discontinuity at the Fermi surface turned out to be detrimental to numerical stability. This function has a finite slope at $x = 0$ that scales with $1/s$, and reproduces Θ for $s \to 0$. For my implementation, it has proven to be a good rule of thumb to set s equal to an eighth of the rapidity grid spacing to ensure stable results. Note that this procedure is mathematically equivalent to introducing a Fermi–Dirac distribution with finite temperature $T \coloneqq mc^2 s/k_{\text{B}}$ for the initial protons. For a typical grid spacing, however, this easily yields temperatures significantly higher than what can be expected physically,[4] so this analogy should not be overrated here.

---

[4] I used a spacing of 2^{-4} in my baryon-stopping computations, which roughly corresponds to $k_{\text{B}}T \approx 8\,\text{MeV}$.

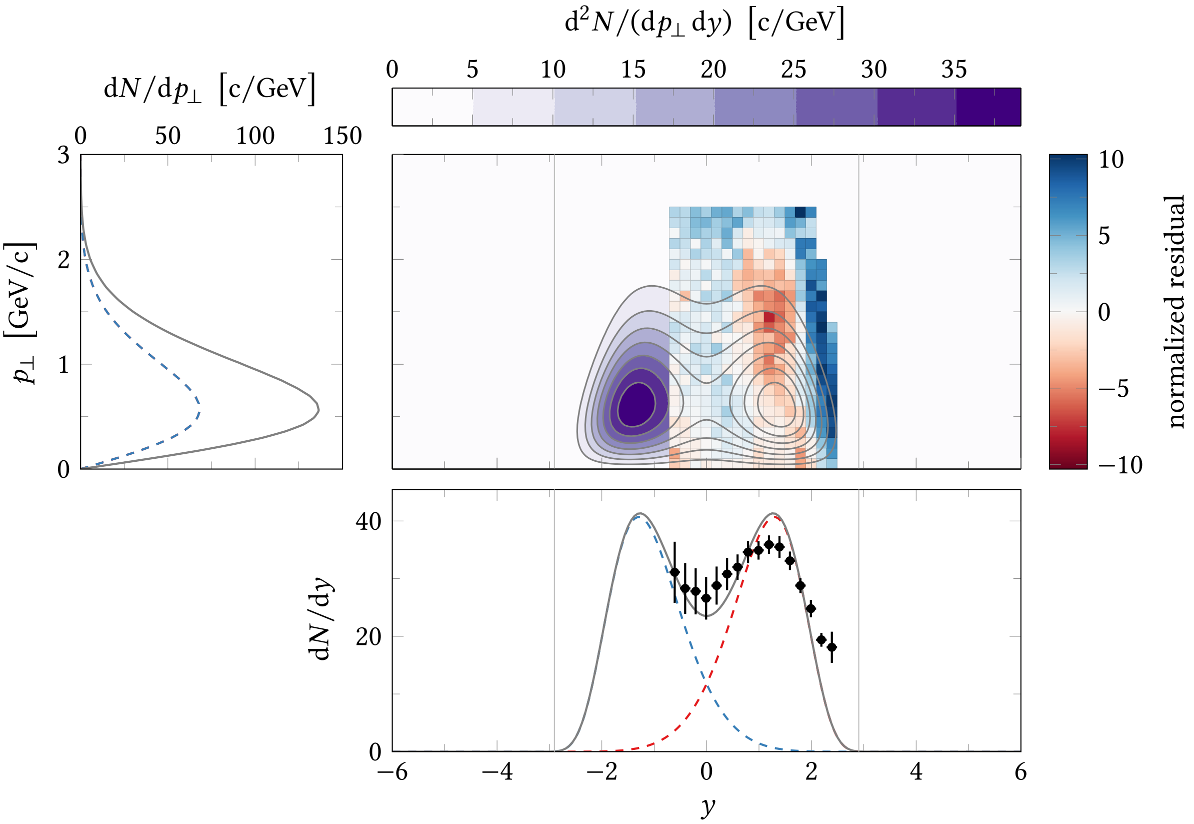


Figure 4.4: Joint (top right) and marginal (top left, bottom right) NDFs of net protons in central collisions of ^{208}Pb nuclei at $\sqrt{s_{\mathrm{NN}}} = 17.2\,\mathrm{GeV}$. The beam rapidity is marked by thin vertical lines.

In the joint plot, filled contour lines depict the model result; the associated normalized residuals for the NA49 (1999) data set are displayed as colored rectangles, see section 3.4.3 for details.

In the marginal plots, the model result is shown as a solid curve, while dashed curves indicate the contributions of the forward- (red) and backward-going (blue) source. Marginal experimental data from the same data set are displayed as circles, with horizontal and vertical bars indicating the bin sizes and total uncertainties, respectively.

Table 4.2: Free model parameters for the stopping of net protons in relativistic heavy-ion collisions, obtained via WLS with respect to the data sets listed in table 4.1. The associated GOF statistic ("χ^2/dof") is given in the last column. Numbers in parentheses indicate the estimated uncertainty of the preceding digit; for parameter values enclosed in brackets or without uncertainties, see text.

data set	$D_\perp\,\Delta t$	$D_\parallel\,\Delta t$	x_0	λ	χ^2/dof
NA49 (1999)	0.242(6)	1.03(1)	0.001	0.2	$\frac{4345}{380}$
BRAHMS (2009)	0.32(9)	2.7(4)	[0.0002]	0.2	$\frac{48.8}{52}$
BRAHMS (2004)	0.13(2)	[7]	0.001	0.2	$\frac{159}{63}$

Figures 4.4 to 4.6 show the resulting joint and marginal theoretical net-proton NDFs as well as the normalized residuals from comparison with experimental data as defined in eq. (3.12). In the chosen convention, a negative residual indicates that the model overestimates the net-proton number in the respective phase-space region, while a positive residual signals that the model result is too low. Since the collisions under consideration are symmetric, all model parameters of the forward- and backward-going source are assumed to be identical. Particularly, the total number of net protons is equally divided between the two sources in each case, and set equal to the estimated number of participating protons, $N = \langle Z_{\text{part}} \rangle$. The remaining free model parameters are obtained via WLS; the results are listed in table 4.2. In both the plots and the parameter estimations, all reported statistical and systematic uncertainties are added in quadrature to obtain overall uncertainties for the data points.

Although there is reasonable agreement between theoretical model and experimental data, the model is not entirely capable of describing the observed net-proton distributions. In the longitudinal direction, the latter appear to be broader than the former, which is well visible for the NA49 data set in fig. 4.4 where the phase-space coverage is most complete. While the trends in the marginal data appear to confirm this observation for the higher energies, it is difficult to make a definite statement here because of the lack of data points at high longitudinal rapidities in the two BRAHMS data sets, *cf.* figs. 4.5 and 4.6. By extension, the good model agreement at $\sqrt{s_{\text{NN}}} = 62.4\,\text{GeV}$ could also merely be a consequence of the sparse data.

A similar effect was already observed in an earlier, purely one-dimensional analysis (Hoelck and Wolschin 2020). Here, however, it became clearly visible only at the higher collision energy of 200 GeV, which is probably related to the fact that the marginal data points gave the fit more freedom in parameter space.

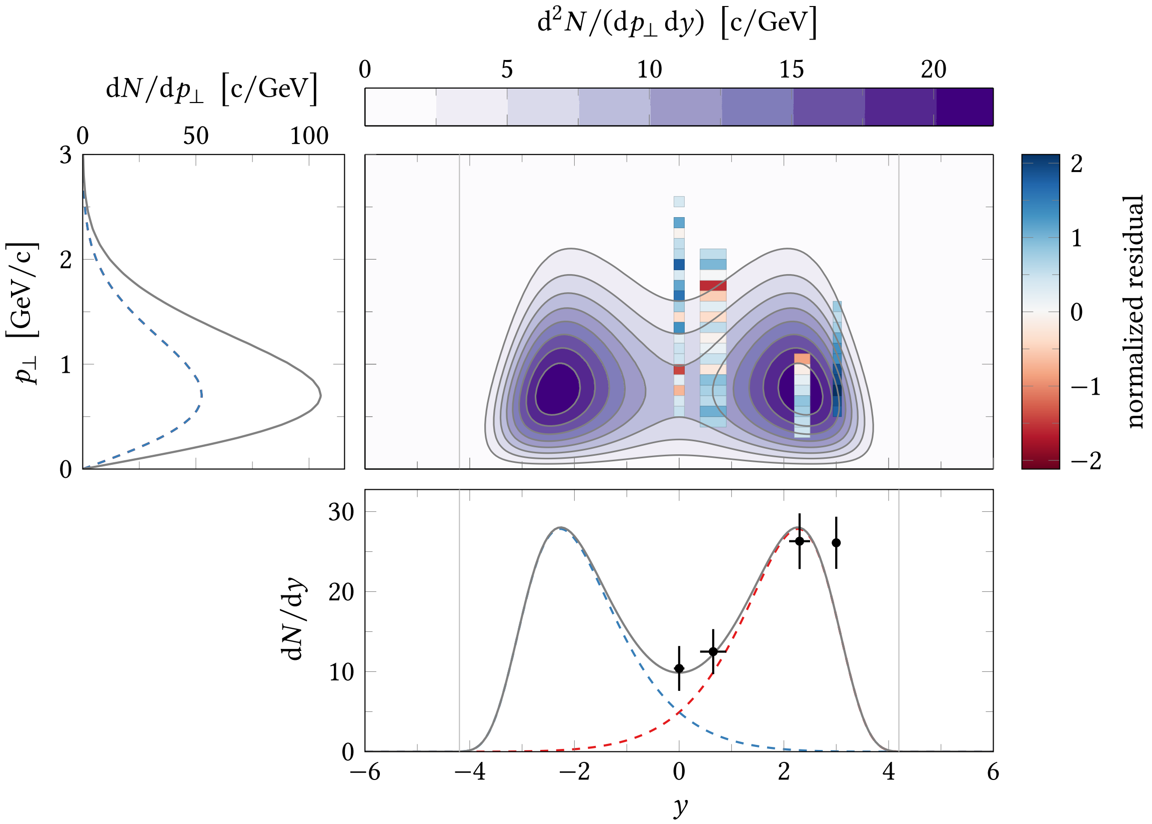


Figure 4.5: Joint (top right) and marginal (top left, bottom right) NDFs of net protons in central collisions of ^{197}Au nuclei at $\sqrt{s_{NN}} = 62.4\,\mathrm{GeV}$. The beam rapidity is marked by thin vertical lines.

In the joint plot, filled contour lines depict the model result; the associated normalized residuals for the BRAHMS (2009) data set are displayed as colored rectangles, see section 3.4.3 for details.

In the marginal plots, the model result is shown as a solid curve, while dashed curves indicate the contributions of the forward- (red) and backward-going (blue) source. Marginal experimental data from the same data set are displayed as circles, with horizontal and vertical bars indicating the bin sizes and total uncertainties, respectively.

This longitudinal broadening might hint at the significance of further physical processes that the current model does not take into account; for example, the presence of non-participant hadrons, which could cause an incipient thermalization. Moreover, interactions between baryons of the same fragment after the initial stopping phase have the potential to redistribute the particles in phase space with similar effect. Alternatively, the discrepancy may simply stem from oversimplification, where the choice of constant diffusivity coefficients first comes to mind. The dependence of the valence-quark distribution (4.13) on the momentum transfer between quark and gluon is also not properly considered in this thesis, but simply fixed to $1\,\mathrm{GeV/c}$, which can be expected to cause distortions at higher collision energies.

In addition, identifying the net-proton number with the number of participating protons does not account for the competing production of other particle species, such as hyperons, in the course of baryon stopping. As a consequence, some partons of the participating protons are bound in other particle species after the initial collision, which presumably leads to a progressive net-proton loss with increasing collision energies. This could explain the predominantly negative residuals at $\sqrt{s_{\mathrm{NN}}} = 200\,\mathrm{GeV}$ and might also contribute to the observed broadening if the conversion rate of protons to other particle species varies accordingly across phase space.

For the gluon-saturation-scale parameters introduced in eq. (4.15), the reference values $\lambda \approx 0.288$ and $x_0 \approx 3.04 \cdot 10^{-4}$ are available from an analysis (Golec-Biernat and Wüsthoff 1998) of lepton–proton DIS data recorded at the Hadron–Electron Ring Accelerator (*Hadron-Elektron-Ringanlage*, HERA). Since the two systems cannot be expected to behave exactly the same, λ and x_0 are treated as free parameters, but constrained to the intervals $0.2 \leq \lambda \leq 0.3$ and $10^{-4} \leq x_0 \leq 10^{-3}$. This is to prevent the WLS method from straying too far from the HERA reference values in an attempt to compensate for the broadening described above, thereby leaving physically reasonable territory. While the λ interval is based on RHIC multiplicity data (Albacete 2007), the stated x_0 range is by no means rigorous and rather intended to reveal a tendency of the parameter towards one direction or the other by allowing the saturation scale to vary by a factor of less than two. The exponent λ consistently hits the lower boundary for all three data sets, which was already observed in a previous analysis (Hoelck and Wolschin 2020), whereas x_0 tends towards its upper limit for the lowest and highest collision energy. For $\sqrt{s_{\mathrm{NN}}} = 62.4\,\mathrm{GeV}$, the WLS favors an intermediate x_0 close to the HERA reference, but also produces an estimated uncertainty that is larger than the parameter value, which means that a conclusive determination of x_0 is not possible here.

The products of diffusivities and interaction timespan provide a measure of how well the idealized, instantaneous redistribution of valence quarks is reproduced in the actual,

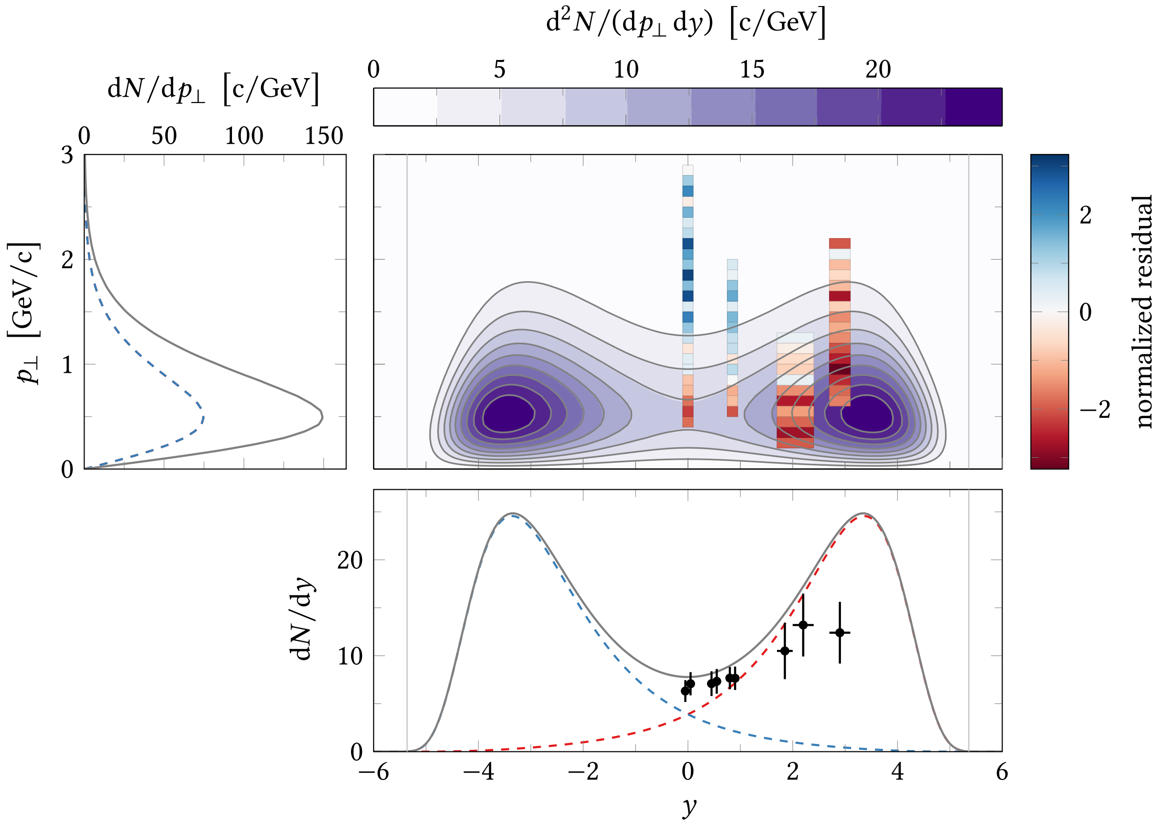


Figure 4.6: Joint (top right) and marginal (top left, bottom right) NDFs of net protons in central collisions of ^{197}Au nuclei at $\sqrt{s_{NN}} = 200\,\text{GeV}$. The beam rapidity is marked by thin vertical lines.

In the joint plot, filled contour lines depict the model result; the associated normalized residuals for the BRAHMS (2004) data set are displayed as colored rectangles, see section 3.4.3 for details.

In the marginal plots, the model result is shown as a solid curve, while dashed curves indicate the contributions of the forward- (red) and backward-going (blue) source. Marginal experimental data from the same data set are displayed as circles, with horizontal and vertical bars indicating the bin sizes and total uncertainties, respectively.

time-dependent quark–gluon interactions. Longitudinally, the results for $D_{\parallel}\,\Delta t$ suggest that this is increasingly well satisfied with growing collision energy. For the highest energy, the final state is even so close to its asymptotic state longitudinally that it is not substantially altered by further increases in $D_{\parallel}\,\Delta t$; as a consequence, the associated uncertainty is large, so that the WLS result can only be interpreted qualitatively rather than quantitatively here. The transverse diffusivities, by contrast, are significantly smaller and show no discernible dependence on the collision energy, implying that the asymptotic DIS distribution is generally too broad to adequately describe the final net-proton state in the transverse direction. However, the difference is not particularly large and might be eliminated by more realistic modeling of the transverse degrees of freedom in the DIS–CGC formalism. The fact that the BRAHMS (2009) data set produces the largest $D_{\perp}\,\Delta t$ is possibly related to its differing centrality cut, since collisions with larger impact parameters induce anisotropic flows in the transverse plane, which could lead to higher effective transverse particle diffusion in baryon stopping.

5. Thermalization of produced hadrons

Due to the large number of particles present in the final stages of a relativistic heavy-ion collision, it is to be expected that mutual interactions lead to an ongoing thermalization of the system. However, due to its rapid expansion, especially in the longitudinal direction, the interaction time is limited: On the one hand, because the time span until the hadrons hit the detectors is of course finite; but on the other hand, because the expansion dilutes the system, so that the mutual interactions effectively cease at some point ("kinetic freeze-out"). It is therefore expected (Wolschin 2016) that the system never reaches thermal equilibrium, but remains in a transient state at the time of measurement.

These are good conditions for the application of (non-equilibrium) statistical models in general and drift–diffusion processes in particular: Due to the multitude of particles and interactions, correlations are expected to be negligible, so that the effect of the particle–hadron collisions should be well reproduced by a stochastic force acting on the hadrons. Simultaneously, the collective expansion of the system and frictional effects with the surroundings should result in an additional deterministic contribution to the hadron motion, which can be accounted for by a suitable drift coefficient.

Indeed, promising results could already be obtained by Wolschin (1999, 2007) and Biyajima, Ide, Mizoguchi, and Suzuki (2002) with a simple one-dimensional Ornstein–Uhlenbeck ansatz, which was later refined through the use of a non-linear drift term based on a thermal FDR (Kellers and Wolschin 2019).

5.1. Initial state: Deep inelastic scattering with gluon saturation

Here, I will model the thermalization process to begin directly after the particle production from partonic interactions of the participating nucleons. The resulting trajectories are therefore again to be interpreted in terms of parton–hadron duality, since many of the produced particles immediately dissolve into their partonic constituents, which only later reassemble into real hadrons.

The associated initial-state PDFs are taken from the DIS–CGC formalism introduced in the previous chapter, which appears to be compatible with the findings of a recent analysis of initial-state signals in ALICE[1] data (Acharya *et al.* 2022). Of course, it would be more accurate to derive the initial state from a dynamical calculation, *e. g.*, by adapting the baryon-stopping model presented in chapter 4; however, it will turn out that the signal of the initial state is not very pronounced in the final state of thermalization, and therefore, this procedure would only lead to a minor correction at the cost of a greatly increased number of free model parameters. For similar reasons, I will ignore here that the kinetic freeze-out does not occur simultaneously throughout the fireball.

In contrast to baryon stopping, the hadrons originating from gluon–gluon interactions are of crucial importance in the context of thermalization, since these account for the bulk of particles produced in the heavy-ion collisions considered in this thesis. Together with the valence-quark–gluon and gluon–valence-quark interactions, whose contributions are expected to remain relevant (Wolschin 2013), this results in three distinct sources of hadron production. Based on the same considerations as in chapter 4, these will be denoted as the *central* (“0”), *forward-* (“+”), and *backward-going hadron source* (“−”), respectively.

The form of the forward- and backward-going initial-state PDFs is already known from the previous chapter and is given by eq. (4.12),

$$f_{(P_\perp,Y)_\pm(t_0)}\big(p_\perp,y\big) \overset{!}{=} f_{(P_\perp,Y)_{\mathrm{DISq}_\pm\mathrm{g}_\mp}}\big(p_\perp,y\big)\,. \tag{5.1}$$

For the central gluon–gluon source, the initial PDF can be approximated by (Kharzeev, Levin, and Nardi 2005)

$$\begin{aligned} f_{(P_\perp,Y)_0(t_0)}\big(p_\perp,y\big) \overset{!}{=} f_{(P_\perp,Y)_{\mathrm{DISg}_\pm\mathrm{g}_\mp}}\big(p_\perp,y\big) &:= \mathrm{const.}\; p_\perp \\ \times \frac{1}{p_\perp^2}\int_{-\infty}^{+\infty} x_\pm\, G\big(x_\pm;p_\perp\big)\,\Theta\big(1-x_\pm\big)\, f_{Y_{\mathrm{NFG}}}\big(y_\pm \mp y_\mathrm{b}\big)\,\mathrm{d}y_\pm& \\ \times \frac{1}{p_\perp^2}\int_{-\infty}^{+\infty} x_\mp\, G\big(x_\mp;p_\perp\big)\,\Theta\big(1-x_\mp\big)\, f_{Y_{\mathrm{NFG}}}\big(y_\mp \pm y_\mathrm{b}\big)\,\mathrm{d}y_\mp&\,, \end{aligned} \tag{5.2}$$

which is completely symmetric under the exchange of the forward- and backward-going gluons, see fig. 5.1. As in the valence-quark–gluon case (4.12), Heaviside functions ensure the validity of the gluons’ Bjorken momentum fractions, while eq. (4.16) is used to describe the longitudinal-rapidity PDFs of the containing nucleons. For the gluon structure function G – the integral of the gluon distribution function – I adopt the simplified form

$$xG\big(x;p_\perp\big) \approx \mathrm{const.}\,(1-x)^4 \begin{cases} p_\perp^2 & \text{if} \quad p_\perp^2 < Q_\mathrm{s}^2(x) \\ Q_\mathrm{s}^2(x) & \text{otherwise} \end{cases} \tag{5.3}$$

proposed by the aforementioned authors, where the factor $(1-x)^4$ regulates the high-x

---

[1] A Large Ion Collider Experiment

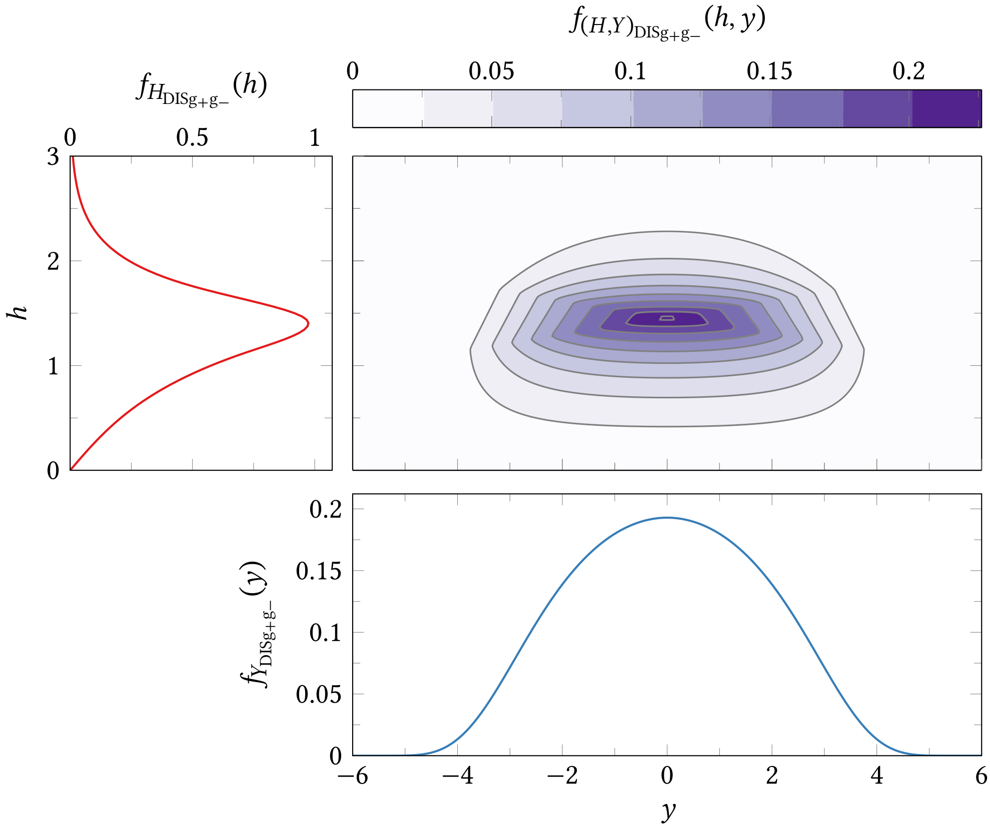


Figure 5.1: Joint (top right) and marginal (top left, bottom right) PDFs of a proton produced in the DIS of saturated gluons for a collision of two ^{197}Au nuclei with $\sqrt{s_{\mathrm{NN}}} = 200\,\mathrm{GeV}$ ($y_b \approx 5.36$). The gluon-saturation parameter values $\lambda = 0.2$ and $x_0 = 10^{-3}$ are identical to those in fig. 4.3; changes in the parameters affect the distribution in a similar way.

behavior as in eq. (4.14). While this is clearly a rather rough approximation, eq. (5.2) will be used exclusively to model initial states for hadron thermalization, where it is sufficient to capture the general shape of the PDF since fine details will fade away during time evolution (see above).

5.2. Asymptotic state: Expanding thermalized gas

Many of the hadrons produced – especially charged pions, which will be of primary interest in later comparison with data – are much lighter and hence less inert than the protons and neutrons previously considered in the process of baryon stopping. I therefore expect that the vast majority of these particles is unable to escape the expanding fireball, and thus subsequently participates in the thermalization process.

With this in mind, it seems reasonable to choose thermal distributions for all three sources to model the asymptotic PDFs, and thereby the FDRs, of produced light hadrons. However, since the system is not spatially confined, it cannot reach a true thermal equilibrium that is independent of its initial configuration (Cubero, Casado-Pascual, Dunkel, Talkner, and Hänggi 2007). I will address this issue by using a family of modified thermal distributions that exhibit a collective expansion of the system, which can act as a remnant of the initial state. The general form of these distributions reads[2]

$$f_{P_{\text{th}}}(p^1, p^2, p^3) := \frac{C_{\text{th}}}{V} \int_\Sigma \frac{\mathrm{d}\bar{\sigma}^* \bar{p}}{p^0} \exp\left(\frac{mc^2 - \bar{u}^* \bar{p}}{k_{\text{B}} T}\right), \tag{5.4}$$

where C_{th} is a normalizing constant and $V := \int_\Sigma \mathrm{d}\bar{\sigma}^* \bar{u}/c$ the volume of the expanding thermal system ("reservoir"), which is contained[3] in some three-dimensional hypersurface $\Sigma \subseteq \mathbb{R}^{1,3}$. Finally, $\bar{u}$ denotes the proper velocity ("proper expansion velocity") of the reservoir's local rest frame, which in general depends on the current position $\bar{\sigma} \in \Sigma$ on the hypersurface, and T the temperature in this frame. For brevity, the expressions above are formulated in terms of Minkowskian vectors, assuming Cartesian coordinates, $\bar{p} := \sum_{\mu=0}^{3} p^\mu \bar{e}_\mu$, and a metric with signature $(1, 3) \mathrel{\hat{=}} (+, -, -, -)$; for other metric conventions, the signs of all Minkowskian inner products must be adjusted accordingly.

The normalizing constant agrees for all distributions and can be calculated analytically (see appendix C) as

$$C_{\text{th}} = \frac{\kappa}{4\pi\,(mc)^3 \exp(\kappa)\, K_2(\kappa)} \tag{5.5}$$

with the dimensionless ratio of the particle's rest and thermal energy

$$\kappa := \frac{mc^2}{k_{\text{B}} T} \tag{5.6}$$

and the modified Bessel function K_a of the second kind and order a. Note that the exponential in the denominator of eq. (5.5) exactly cancels a corresponding term in eq. (5.4), so that it could actually be discarded. The reason why these seemingly redundant exponentials were included here will become clear in section 5.2.4.

In the following, a selection of possible candidates for the three sources will be discussed. Calculations will be largely omitted to keep everything concise; the mathematical details can be found in appendix C.

---

[2] Here, the symbol $\bar{a}^*$ denotes the dual of a Minkowskian vector $\bar{a}$.

[3] The reservoir surface should be regarded as a mathematical idealization rather than a physical entity, as there is no sharp boundary between fireball and freeze-out zone (Schnedermann, Sollfrank, and Heinz 1993).

5.2.1. Non-expanding thermal equilibrium

For completeness, I will first briefly cover the simplest case of a non-expanding reservoir at rest, *i. e.*, $\bar{u} = c\bar{e}_0$ for all $\bar{\sigma} \in \Sigma$. This yields a relativistic generalization of the Maxwell–Boltzmann distribution (2.91), which was proposed by Jüttner in 1911,

$$f_{P_{\mathrm{MJ}}}\left(p^1, p^2, p^3\right) := C_{\mathrm{th}} \exp\left(\frac{mc^2 - cp^0}{k_{\mathrm{B}}T}\right), \tag{5.7}$$

and became known as the Maxwell–Jüttner distribution. The PDF depends only on the relativistic energy of the particle and is thus spherically symmetric. In the limit of small momentum, $cp^0 \approx mc^2 + \|\vec{p}\|^2/(2m)$, and low temperature, $k_{\mathrm{B}}T \ll mc^2$, it reproduces the non-relativistic case as expected.

The Maxwell–Jüttner distribution broadens with increasing temperature and becomes essentially flat in momentum space for infinite temperature. In rapidity coordinates, however, this growth is bounded in the longitudinal direction, which becomes apparent when considering the associated marginal PDFs,

$$f_{H_{\mathrm{MJ}}}(h) = \frac{\kappa}{K_2(\kappa)} \sinh(h) \cosh(h)^2 \, K_1\left(\kappa \cosh(h)\right), \tag{5.8}$$

$$f_{Y_{\mathrm{MJ}}}(y) = \frac{1}{2K_2(\kappa)} \left[1 + \frac{2}{\kappa \cosh(y)} + \frac{2}{[\kappa \cosh(y)]^2}\right] \exp\left(-\kappa \cosh(y)\right). \tag{5.9}$$

While the former diverges for $\kappa \to 0$, the latter approaches the well-defined limit

$$\lim_{\kappa \to 0} f_{Y_{\mathrm{MJ}}}(y) = \frac{1}{2\cosh(y)^2}, \tag{5.10}$$

where the asymptotic behavior $K_2(\kappa) \approx 2/\kappa^2$ of the Bessel function for small κ was used.

Figure 5.2 shows the joint and marginal rapidity PDFs of the Maxwell–Jüttner distribution, including the infinite-temperature limit (5.10). The latter has a finite *full width at half maximum* (FWHM) of $\Delta y_{\mathrm{MJ}} := 2\,\mathrm{arcosh}(\sqrt{2}) \approx 1.76$, which also serves as an upper bound for the width of any Maxwell–Jüttner distribution in longitudinal-rapidity space since the FWHM grows monotonically with temperature.[4] A quick comparison with experimental data reveals, however, that this width is significantly smaller than those observed in measured charged-hadron distributions (*cf.* section 5.3). This indicates that a plain Maxwell–Jüttner distribution is unfit for modeling that observable and confirms the expected necessity to account for collective expansion.

---

[4] A similar statement can be derived for the variance, which is bounded from above by the limiting value $\lim_{\kappa \to 0} E_{Y_{\mathrm{MJ}}}\left(Y_{\mathrm{MJ}}^2\right) = \frac{\pi^2}{12} \approx 0.822$.

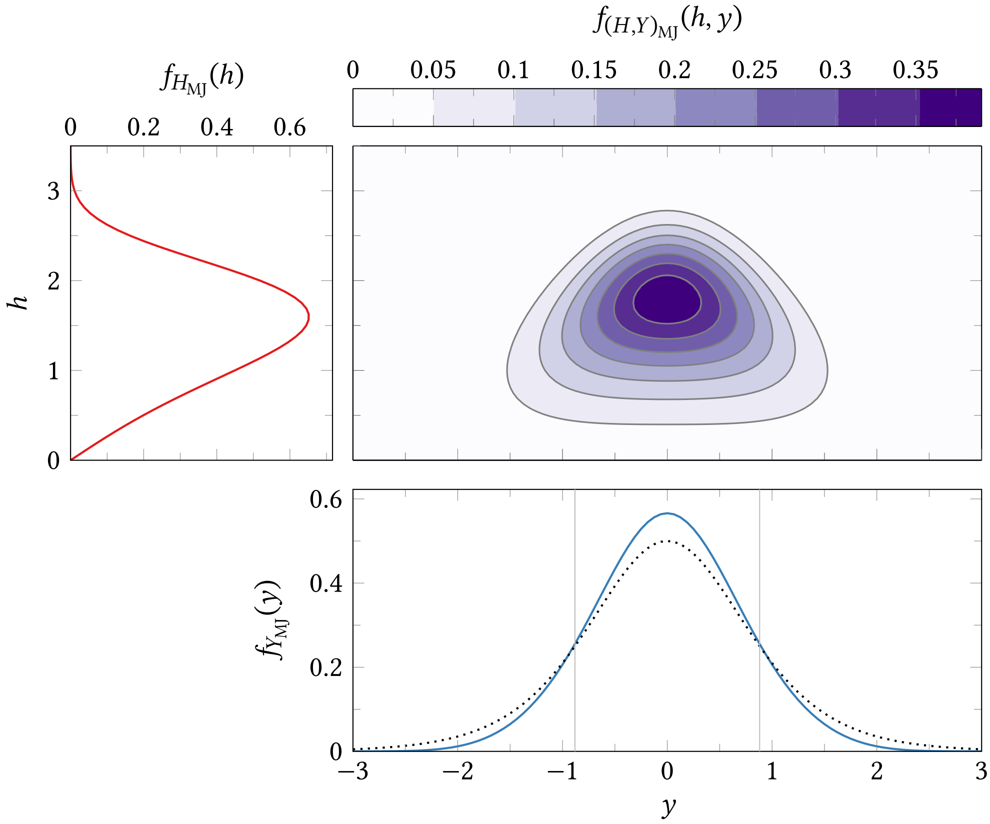


Figure 5.2: Joint (top right) and marginal (top left, bottom right) PDFs of a thermal Maxwell–Jüttner distribution with rest-to-thermal-energy ratio $\kappa = 1$. If κ is increased, the distribution becomes steeper and its maximum moves closer to $h = 0$. In the $f_{Y_{\mathrm{MJ}}}$ plot, the limiting case $\kappa \to 0$ is shown as a dotted curve, and vertical bars mark the upper bound for the FWHM.

5.2.2. Spherical expansion

Bondorf, Garpman, and Zimanyi (1978) and Siemens and Rasmussen (1979) considered the case of an isotropic reservoir expanding with a finite constant radial velocity, $\|\vec{u}\| = \text{const.} > 0$, which became known as the “blast-wave model”. Performing the surface integral in eq. (5.4) then yields

$$f_{\boldsymbol{P}_{\mathrm{BW}}}(p^1, p^2, p^3) := C_{\mathrm{th}} \, \frac{u^0}{c} \, K\!\left(\frac{u^0 p^0}{k_{\mathrm{B}} T}, \frac{\|\vec{u}\| \|\vec{p}\|}{k_{\mathrm{B}} T}\right), \tag{5.11}$$

with $u^0 := \sqrt{c^2 + \|\vec{u}\|^2}$ and the auxiliary functions

$$K(x_1, x_2) := \frac{L(x_1 - x_2) - L(x_1 + x_2)}{2 x_1 x_2}, \qquad L(x) := (1 + x) \exp(\kappa - x). \tag{5.12}$$

Comparisons with experimental data suggest that $\|\vec{u}\|/u^0$ does not exceed 0.7, but give no indication of any lower bound (Schnedermann, Sollfrank, and Heinz 1993). Small increases in the expansion velocity have a qualitatively similar effect as an increased temperature and *vice versa.* Interestingly, the infinite-temperature limit is unaffected by the expansion velocity, which will be examined in more detail in a moment.

The blast-wave model exhibits a singularity at the reservoir center, where the direction of the expansion velocity is undefined, which is physically dubious. Lee and Heinz (1989) addressed this problem by introducing a non-constant radial expansion velocity profile,[5] such that the associated (non-relativistic) expansion velocity fraction $\|\vec{u}\|/u^0$ takes the form of a power law with exponent $\nu > 0$,

$$\frac{\|\vec{u}\|}{u^0} \overset{!}{=} \left(\frac{\|\vec{\sigma}\|}{R}\right)^{\nu} \beta_\mathrm{S} \quad \forall\, \bar{\sigma} \in \Sigma\,. \tag{5.13}$$

With R denoting the radius of the spherical reservoir, the proper expansion velocity vanishes at the center, but increases strictly monotonically with growing distance until it takes its maximum value $\beta_\mathrm{S}/\sqrt{1-\beta_\mathrm{S}^2}$, with $0 < \beta_\mathrm{S} < 1$, at the reservoir surface. The resulting PDF

$$f_{P_\mathrm{SF}}\big(p^1, p^2, p^3\big) := \frac{C_\mathrm{th}}{{}_2F_1\!\left(-\frac{1}{2}, \frac{3}{2\nu}; 1 + \frac{3}{2\nu}; \beta_\mathrm{S}^2\right)} \frac{3}{R^3} \int_0^R \|\vec{\sigma}\|^2\, K\!\left(\frac{u^0 p^0}{k_\mathrm{B} T}, \frac{\|\vec{u}\|\|\vec{p}\|}{k_\mathrm{B} T}\right) \mathrm{d}\|\vec{\sigma}\| \tag{5.14}$$

remains structurally very similar to eq. (5.11), but has an additional integral over the auxiliary function K and a pre-factor involving the ordinary hypergeometric function ${}_2F_1$. Taking the limit $\nu \to 0$ reproduces the blast-wave model, as the expansion velocity becomes constant again.

A finite velocity-profile exponent ν makes the associated marginal rapidity PDFs slightly narrower, which can be understood by taking a look at the reservoir average of eq. (5.13),

$$\left\langle \frac{\|\vec{u}\|}{u^0} \right\rangle := \frac{1}{V} \int_\Sigma \frac{\mathrm{d}\bar{\sigma}^* \bar{u}}{c} \frac{\|\vec{u}\|}{u^0} = \frac{3}{3+\nu} \frac{{}_2F_1\!\left(-\frac{1}{2}, \frac{3+\nu}{2\nu}; 1 + \frac{3+\nu}{2\nu}; \beta_\mathrm{S}^2\right)}{{}_2F_1\!\left(-\frac{1}{2}, \frac{3}{2\nu}; 1 + \frac{3}{2\nu}; \beta_\mathrm{S}^2\right)} \beta_\mathrm{S}\,, \tag{5.15}$$

which is monotonically decreasing in ν for all values of β_S, as depicted in fig. 5.3. Hence, increasing the profile exponent affects the distribution in a roughly similar way to setting a lower maximum expansion velocity.

While the spherically-expanding thermal distributions presented so far allow ample customization in the transverse direction, which makes them promising candidates for modeling transverse hadron distributions, they suffer from a fundamental flaw in longitu-

---

[5] See also the article by Lee, Heinz, and Schnedermann (1990) for more details.

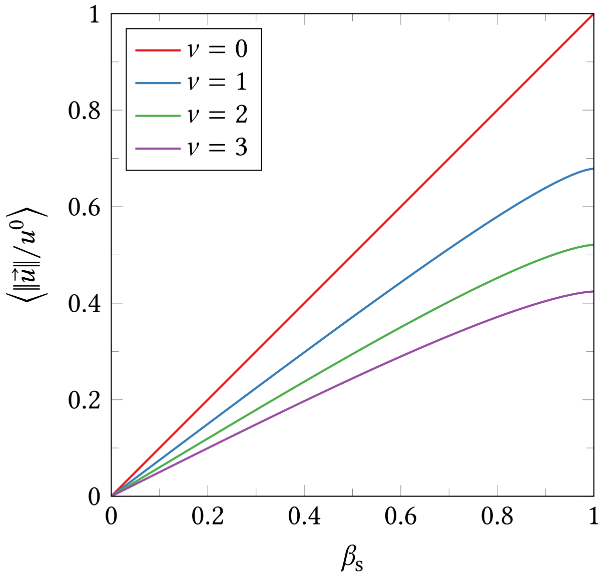


Figure 5.3: Average of the non-relativistic expansion velocity fraction $\|\vec{u}\|/u^0$ as a function of the non-relativistic surface velocity fraction β_s for several velocity-profile exponents ν.

dinal rapidity space: As already indicated, both f_{Y_BW} and f_{Y_SF} approach the same marginal PDF (5.10) for high temperatures as the Maxwell–Jüttner distribution, and thus, their FWHMS are inherently limited by Δy_MJ. It can even be shown (see appendix C.3.2) that this is true for *any* thermal distribution of the form (5.4) with an isotropic radial expansion velocity and local-time hadronization. Consequently, a different kind of collective expansion is needed to successfully model the experimental data under consideration.

5.2.3. Transverse–longitudinal expansion

A possible solution to the aforementioned problem consists in a separation of the collective expansion into independent transverse and longitudinal components, as suggested by Schnedermann, Sollfrank, and Heinz (1993). Here, the expansion-velocity field is modeled as a Bjorken flow (Bjorken 1983) that additionally expands perpendicular to the beam axis, leading to a PDF of the form

$$f_{(H,Y)_\mathrm{TF+LF}}(h, y) := \frac{1}{\zeta_\mathrm{max} - \zeta_\mathrm{min}} \int_{\zeta_\mathrm{min}}^{\zeta_\mathrm{max}} f_{(H,Y)_\mathrm{TF}}(h, y - \zeta)\, \mathrm{d}\zeta, \tag{5.16a}$$

which contains the purely transverse expanding PDF

$$f_{(H,Y)_{\mathrm{TF}}}(h,y) := \frac{2\pi\,(mc)^3\,C_{\mathrm{th}}}{{}_2F_1\!\left(\frac{1}{2},\frac{1}{\nu_\perp};1+\frac{1}{\nu_\perp};\tilde{\beta}_{\mathrm{s}}^2\right)}\sinh(h)\cosh(h)^2\cosh(y) \\ \times\frac{2}{R_\perp^2}\int_0^{R_\perp}\|\vec{\sigma}_\perp\|\,I\!\left(\kappa-\kappa\frac{u_\perp^0}{c}\cosh(h)\cosh(y),\kappa\frac{\|\vec{u}_\perp\|}{c}\sinh(h)\right)\mathrm{d}\|\vec{\sigma}_\perp\|\,. \tag{5.16b}$$

From a mathematical point of view, the first equation (5.16a) is a convolution of a rectangular function and a transverse-flow PDF (5.16b). Physically, this can be understood as an averaging over all Lorentz boosts of $f_{(H,Y)_{\mathrm{TF}}}$ along the beam axis with boost rapidities $\zeta_{\min} < \zeta < \zeta_{\max}$. For symmetric collisions in the COM frame, $\zeta_{\max} = -\zeta_{\min}$ is therefore a natural choice. The auxiliary function

$$I(x_1,x_2) := \exp(x_1)\,I_0(x_2) \tag{5.17}$$

contains an exponential and a zeroth-order modified Bessel function I_0 of the first kind, which encapsulates an angular integral in the transverse plane.

Similar to the spherically-expanding case, a transverse-velocity profile of the form

$$\frac{\|\vec{u}_\perp\|}{u_\perp^0} \overset{!}{=} \left(\frac{\|\vec{\sigma}_\perp\|}{R_\perp}\right)^{\nu_\perp}\tilde{\beta}_{\mathrm{s}} \quad \forall\,\vec{\sigma}\in\Sigma \tag{5.18}$$

is used, where $\vec{\sigma}_\perp$ and $\vec{u}_\perp$ denote the projections of $\vec{\sigma}$ and $\vec{u}$ on the transverse plane, respectively, and $u_\perp^0 := \sqrt{c^2+\|\vec{u}_\perp\|^2}$. Note that the transverse *non-relativistic* expansion velocity fraction $\|\vec{u}_\perp\|/u^0$ is not constant on the reservoir surface, because it also depends on u^3, and is only equal to $\tilde{\beta}_{\mathrm{s}}$ for $\sigma^3 = 0$.

With respect to transverse rapidity, this distribution behaves very similarly to the spherical-flow PDFs discussed earlier, see fig. 5.4, with a slightly more pronounced dent at small h. By construction, the value of $\zeta_{\max}$ does not influence its shape much in the transverse direction, and when computing the marginal PDF, the associated integral drops out completely, such that

$$f_{H_{\mathrm{TF+LF}}}(h) = f_{H_{\mathrm{TF}}}(h) = \frac{\kappa}{K_2(\kappa)}\,\frac{\sinh(h)\cosh(h)^2}{{}_2F_1\!\left(\frac{1}{2},\frac{1}{\nu_\perp};1+\frac{1}{\nu_\perp};\tilde{\beta}_{\mathrm{s}}^2\right)} \\ \times\frac{2}{R_\perp^2}\int_0^{R_\perp}\|\vec{\sigma}_\perp\|\,K_1\!\left(\kappa\frac{u_\perp^0}{c}\cosh(h)\right)I_0\!\left(\kappa\frac{\|\vec{u}_\perp\|}{c}\sinh(h)\right)\mathrm{d}\|\vec{\sigma}_\perp\|\,. \tag{5.19}$$

It is easy to see that this reproduces the marginal Maxwell–Jüttner PDF (5.8) for $\tilde{\beta}_{\mathrm{s}} = 0$. Figure 5.5 shows the longitudinal marginal PDF, whose FWHM is approximately $2\zeta_{\max}$ and

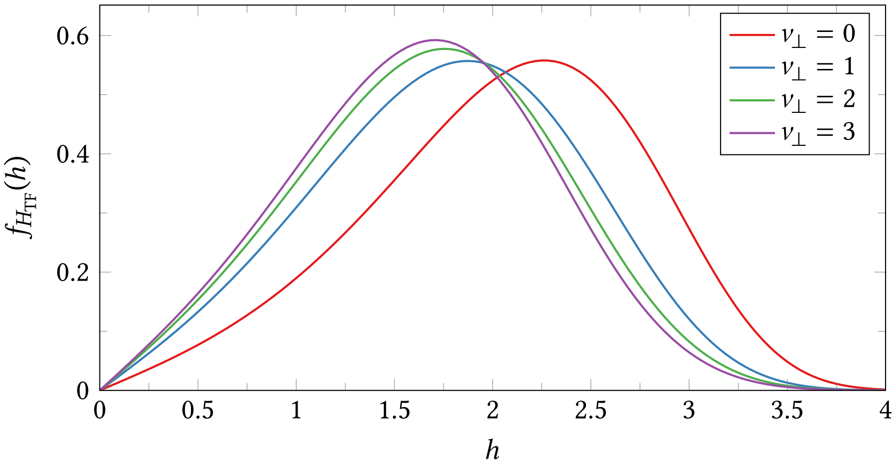


Figure 5.4: Marginal transverse-flow PDF for several velocity-profile exponents at rest-to-thermal-energy ratio $\kappa = 1$ and surface-velocity parameter $\tilde{\beta}_\mathrm{s} = 0.7$.

thus no longer limited by Δy_MJ. For large values of ζ_max, a plateau forms in the central region and the slope of the flanks decreases.

Albeit the treatment of longitudinal expansion is kept rather simple, this distribution appears to be adaptable enough to provide a suitable asymptotic state for the thermalization model. Therefore, I set

$$f_{(H,Y)_i(\infty)}(h,y) \overset{!}{=} f_{(H,Y)_\mathrm{TF+LF}}(h,y) \quad \text{for} \quad i \in \{0,+,-\} \tag{5.20}$$

with the same parameters for all sources (but see below), since the latter are assumed to approach a common thermal state.

5.2.4. Modified high-momentum tails

It has been known for a long time (Hagedorn 1983) that thermal models fail to describe the high-momentum tails of observed transverse particle distributions, where collective expansion is less pronounced than in the longitudinal direction, although they can provide an accurate description of the data at low momenta. Apparently, the distribution of the particles' Cartesian momentum coordinates changes smoothly from being normally distributed at $\|\vec{p}\| \ll mc$ and exponentially distributed for larger momenta – which can be well described by thermal distributions of the form (5.4) – to a Pareto distribution at $\|\vec{p}\| \gg mc$, whose PDF is given by a power law with negative exponent. Physically, this shift can be attributed to the emergence of so-called "hard" processes in this regime, which are presumably governed by perturbative QCD and thus not covered by thermal physics (Biyajima, Mizoguchi, Nakajima, Suzuki, and Wilk 2006).

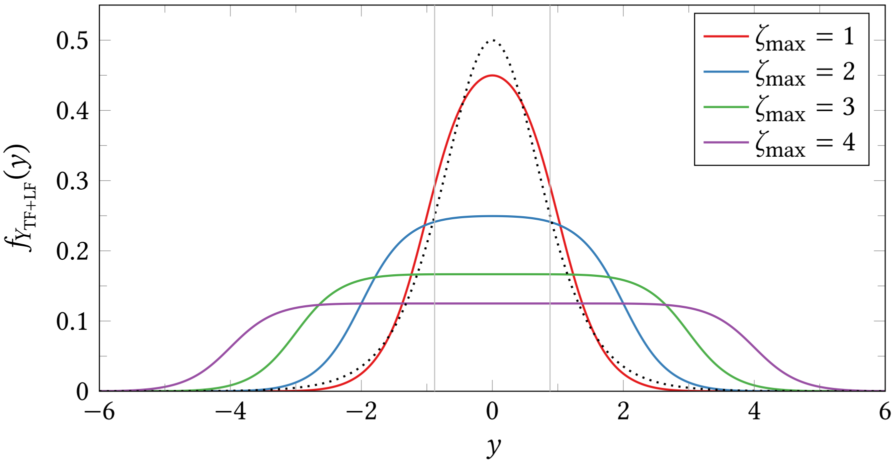


Figure 5.5: Marginal transverse–longitudinal-flow PDF for several maximum boost rapidities $\zeta_{\max} = -\zeta_{\min}$ at $\kappa = 1$, $\tilde{\beta}_{\mathrm{s}} = 0.7$, and $\nu_{\perp} = 1$. For reference, the spherical infinite-temperature PDF and FWHM are shown as dotted curve and vertical bars, respectively.

In this thesis, I will use a phenomenological approach inspired by Michael and Vanryckeghem (1977) and Hagedorn (1983, appendix B.3) to take this transition into account. It is based on the fact that the exponential function can be written as the limit of an infinite sequence,

$$\lim_{n\to\pm\infty}\left(1+\frac{x}{n}\right)^{n} = \exp(x)\,, \tag{5.21}$$

where convergence is faster if the signs of x and n match. Truncating this sequence at a finite number n yields the function

$$\widetilde{\exp}(x;n) \coloneqq \left(1+\frac{x}{n}\right)^{n} \approx \begin{cases} \exp(x) & \text{for} \quad \left|\frac{x}{n}\right| \ll 1\,, \\ \left(\frac{x}{n}\right)^{n} & \text{for} \quad \left|\frac{x}{n}\right| \gg 1\,, \end{cases} \tag{5.22}$$

which is positive and strictly monotonic in x for $x/n > -1$.

Based on this function, the aforementioned authors introduced ad-hoc distributions for the transverse momentum by setting $x \coloneqq \|\vec{p}_{\perp}\|/a$ or $\sqrt{(mc)^2 + \|\vec{p}_{\perp}\|^2}/a$ for some negative constant a and an exponent n in the range -2 to -10. While this adequately reproduces the mid- to high-momentum regime, it does not describe low momenta particularly well,[6] nor does it cover the longitudinal direction. Moreover, the connection to the original thermal distribution is not entirely clear, necessitating the introduction of new, phenomenological parameters.

---

[6] Hagedorn (1983) proposed a piece-wise defined distribution function to circumvent this problem.

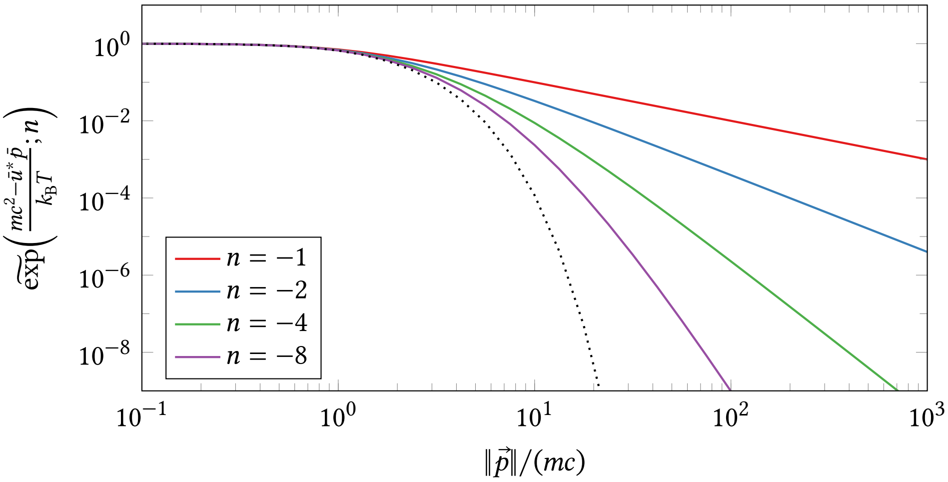


Figure 5.6: High-momentum tails for several power-law exponents at $\kappa = 1$ and $\|\vec{u}\| = 0$. The function shown is identical to the modified Maxwell–Jüttner PDF up to a multiplicative constant. A dotted curve indicates the limit $n \to -\infty$, which reproduces the unmodified exponential.

I address these issues by choosing

$$x := \frac{mc^2 - \bar{u}^* \bar{p}}{k_\mathrm{B} T} \tag{5.23}$$

instead, which is the argument of the exponential in eq. (5.4). This choice therefore amounts to simply replacing $\exp \mapsto \widetilde{\exp}$ in the defining expression of the thermal distribution family discussed above, with the original definition being recovered in the limiting case $|n| \to \infty$. Furthermore, both transverse and longitudinal high-momentum tails are accounted for on a consistent basis.

Equation (5.23) is always non-positive, since $\bar{u}^* \bar{p} \geq mc^2$ due to the (reversed) Cauchy–Schwarz inequality, and approaches zero for small momenta and expansion velocities. Choosing a negative n hence yields the required $\|\vec{p}\|$-dependence over the entire momentum range, as shown in fig. 5.6 for the simple case $\|\vec{u}\| = 0$. The inclusion of the constant mc^2 in eq. (5.23) is crucial here to obtain the correct low-momentum behavior for massive particles.

Starting from the modified definition (5.4), not only the non-expanding Maxwell–Jüttner distribution but also the expanding blast-wave, spherical-flow, and transverse–longitudinal-flow model can be recomputed for finite n. Although the modification of the exponential complicates some calculations, analytical results can still be obtained, see appendix C. Note that $n < -3$ is required for the resulting functions to remain normalizable, which puts an upper limit on the value of n.

A welcome side effect of the power-law tails is that FDRs computed from the thermal distributions do not diverge for large transverse or longitudinal rapidities, *e. g.*,

$$\lim_{h\to\infty} \frac{\partial \ln\Big(f_{(H,Y)_{\mathrm{MJ}}}(h,y)\Big)}{\partial h} = 3+n\,, \tag{5.24}$$

$$\lim_{y\to\pm\infty} \frac{\partial \ln\Big(f_{(H,Y)_{\mathrm{MJ}}}(h,y)\Big)}{\partial h} = \pm(1+n)\,, \tag{5.25}$$

which benefits numerical stability when solving the associated FPEs.

5.3. Results

Experimental data on the distribution of negatively-charged pions (π^-) in transverse-momentum and longitudinal-rapidity space is available from the BRAHMS collaboration for central collisions of ^{197}Au nuclei with $\sqrt{s_{\mathrm{NN}}} = 62.4\,\mathrm{GeV}$ (Arsene *et al.* 2010) and $\sqrt{s_{\mathrm{NN}}} = 200\,\mathrm{GeV}$ (Bearden *et al.* 2005) performed at RHIC. The same authors also provide data on positively-charged pions (π^+), which are practically identical to the π^- data but contain more outliers and are therefore not included in the following analysis.

For the higher collision energies 2.76 TeV and 5.02 TeV reached in ^{208}Pb collisions at the LHC, there is only data for unidentified electrically-charged hadrons ($h^\pm$), because no suitable spectrometer has been installed yet. For the same reasons, the measured distributions are given in terms of the coordinates $(p_\perp, \eta)$, with the so-called *pseudo-rapidity*[7]

$$\eta \coloneqq \operatorname{artanh}\left(\frac{p^3}{\|\vec{p}\|}\right) = \operatorname{arsinh}\left(\sqrt{1+\left(\frac{mc}{p_\perp}\right)^2}\,\sinh(y)\right), \tag{5.26a}$$

$$\frac{\partial(p_\perp,\eta)}{\partial(p_\perp,y)} = \begin{pmatrix} 1 & 0 \\ -\frac{\tanh(\eta)}{p_\perp\left[1+\left(\frac{p_\perp}{mc}\right)^2\right]} & \sqrt{1+\left(\frac{mc}{p_\perp\cosh(\eta)}\right)^2} \end{pmatrix}, \tag{5.26b}$$

since the latter requires merely knowledge of the inclination angle of the particle's trajectory with respect to the beam axis, which is available even without particle identification. The two rapidity coordinates y and η agree for $\|\vec{p}\| \gg mc$, which is well satisfied for many of the experimental data points.

The vast majority of particles produced in relativistic heavy-ion collisions at the LHC are pions, followed by kaons and (anti-)protons (Abelev *et al.* 2013; Acharya *et al.* 2020).

---

[7] Often, a mathematically equivalent definition is used, which makes use of the identity $\operatorname{artanh}(x) = -\ln\big(\tan(\arccos(x)/2)\big)$ and the fact that $\arccos(p^3/\|\vec{p}\|)$ is the inclination angle of the particle's momentum relative to the beam axis.

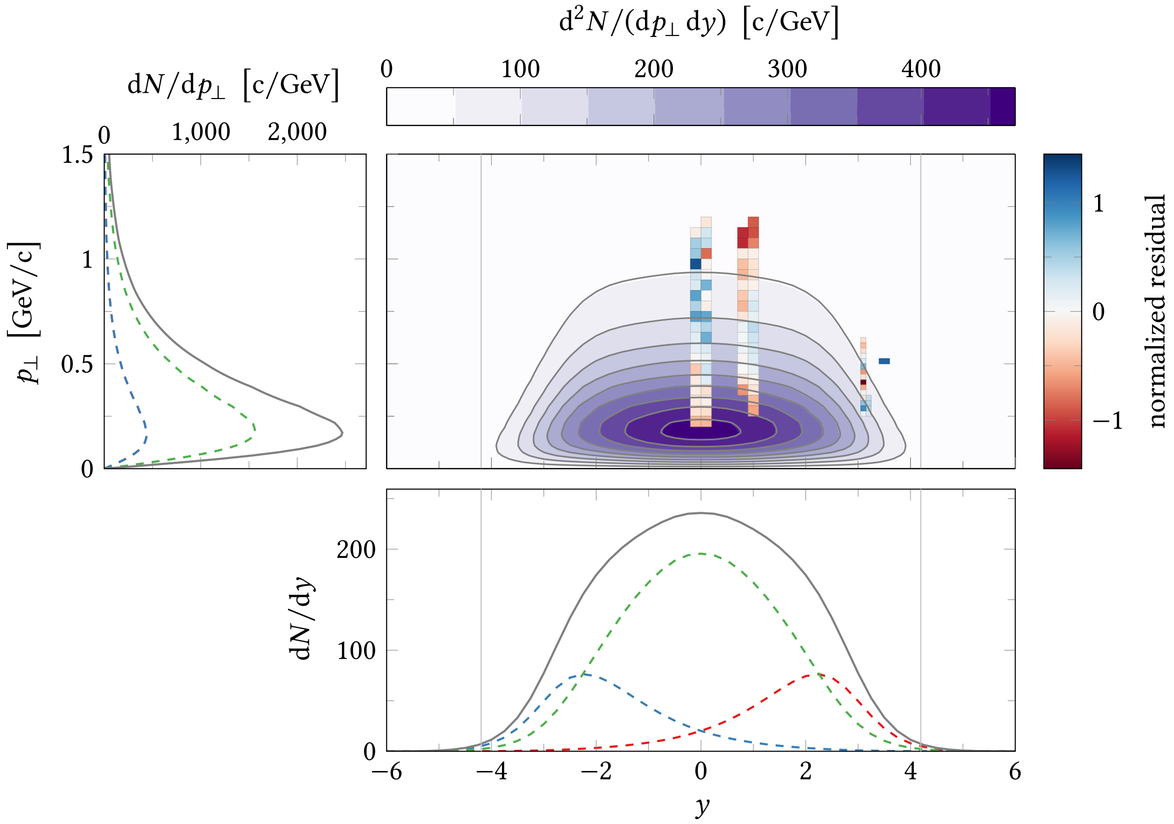


Figure 5.7: Joint (top right) and marginal (top left, bottom right) NDFs of negatively-charged pions in central collisions of ^{197}Au nuclei at $\sqrt{s_{NN}} = 62.4\,\mathrm{GeV}$. The beam rapidity is marked by thin vertical lines.

In the joint plot, filled contour lines depict the model result; the associated normalized residuals for the BRAHMS (2010) data set are displayed as colored rectangles, see section 3.4.3 for details.

In the marginal plots, the model result is shown as a solid curve, while dashed curves indicate the contributions of the central (green), forward- (red), and backward-going (blue) source.

Table 5.1: Shared properties of the data sets used with the hadron-thermalization model. The last column lists the hadron species considered in the comparison with model results.

data set(s)	$\sqrt{s_{\mathrm{NN}}}$	nuclei	y_{b}	centrality	hadron
BRAHMS (2010)	62.4 GeV	^{197}Au	4.20	0 – 10 %	π^-
BRAHMS (2005)	200 GeV	^{197}Au	5.36	0 – 5 %	π^-
ALICE (2018) ALICE (2013) ATLAS (2015)	2.76 TeV	^{208}Pb	7.99	0 – 5 %	$h^{\pm} = \pi^{\pm}, \mathrm{K}^{\pm}, \overset{(-)}{\mathrm{p}}, (\ldots)$
ALICE (2018) ALICE (2017)	5.02 TeV	^{208}Pb	8.58	0 – 5 %	$h^{\pm} = \pi^{\pm}, \mathrm{K}^{\pm}, \overset{(-)}{\mathrm{p}}, (\ldots)$

I therefore address the issue of missing particle identification by calculating separate NDFs for these three particle species, using the same model parameters except for mass,[8] thereby tripling the number of particle sources. Then, the three NDFs are weighted by the measured particle ratios, which are approximately 83 % : 13 % : 4 % for the collision energies in question, and finally summed to obtain an estimate of the total $h^{\pm}$ NDF.

Albeit it would be more accurate, I deliberately refrain from using individual parameters for the different particle species here, since this would inflate the number of free model parameters. As the total NDF is primarily shaped by the contribution of the numerous pions, the error caused by this approximation should be reasonably small.

Unfortunately, none of the available LHC data sets covers a particularly large phase-space region, which is why I use a combination of multiple data sets from the ALICE collaboration for parameter estimation at each collision energy: A first data set covers the transverse direction in the narrow corridor $|\eta| \leq 0.8$ (Acharya *et al.* 2018), while a second data set in marginal pseudo-rapidity space provides information about the longitudinal direction (Abbas *et al.* 2013; Adam *et al.* 2017). For $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$, additional $(p_\perp, \eta)$ data for the region $0 \leq \eta \leq 2$, recorded by the ATLAS[9] collaboration (Aad *et al.* 2015), are also included.

An overview of the selected data sets, including beam rapidities and centrality cuts, is given in table 5.1.

The model results for π^- are shown in figs. 5.7 and 5.8: the calculated theoretical NDF in $(p_\perp, y)$-space, the associated residuals from comparison with BRAHMS data points, and the contributions of the individual sources to the marginal NDFs. The same format is used for the $h^{\pm}$ results in figs. 5.9 and 5.10, but with the (longitudinal) rapidity y replaced by the pseudo-rapidity η. Moreover, the source contributions in the marginal plots are broken

---

[8] The masses of $\pi^{\pm}$, $\mathrm{K}^{\pm}$, and $\overset{(-)}{\mathrm{p}}$ are $139.6\,\mathrm{MeV}/c^2$, $493.7\,\mathrm{MeV}/c^2$, and $938.3\,\mathrm{MeV}/c^2$, respectively.

[9] A Toroidal LHC Apparatus

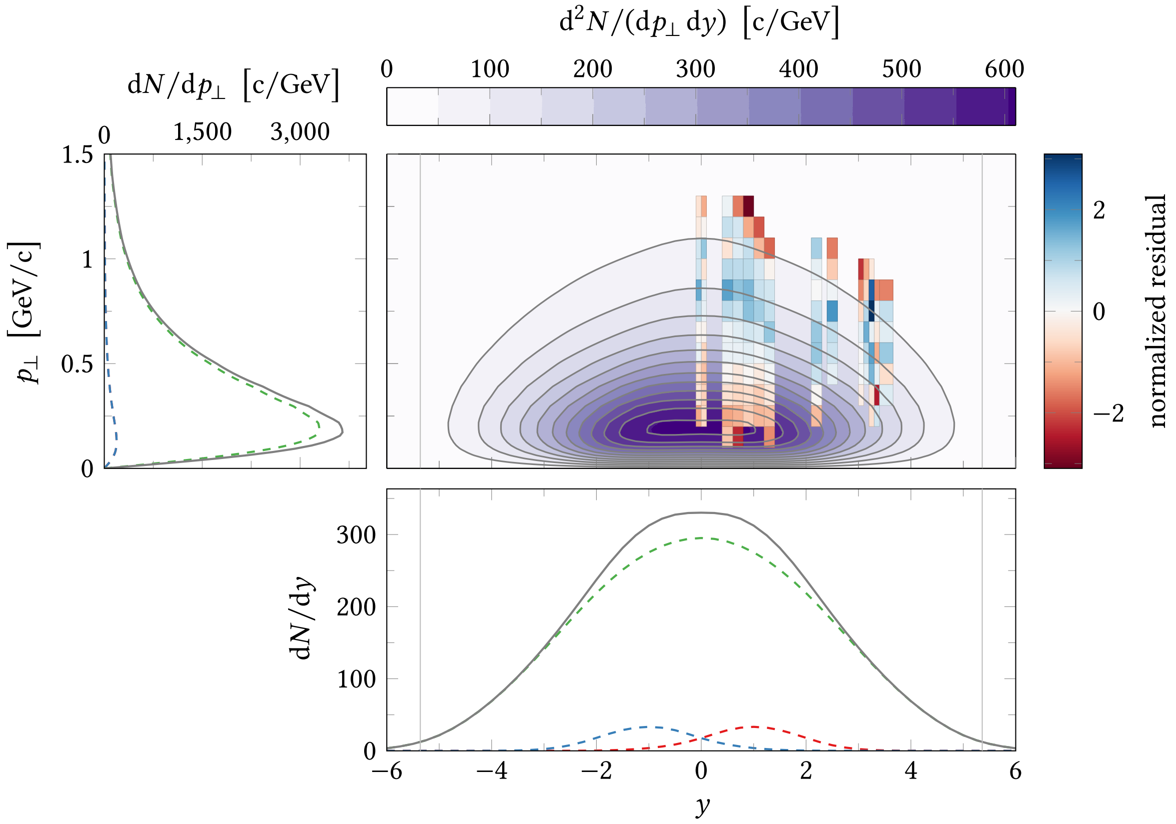


Figure 5.8: Joint (top right) and marginal (top left, bottom right) NDFs of negatively-charged pions in central collisions of ^{197}Au nuclei at $\sqrt{s_{NN}} = 200$ GeV. The beam rapidity is marked by thin vertical lines.

In the joint plot, filled contour lines depict the model result; the associated normalized residuals for the BRAHMS (2005) data set are displayed as colored rectangles, see section 3.4.3 for details.

In the marginal plots, the model result is shown as a solid curve, while dashed curves indicate the contributions of the central (green), forward- (red), and backward-going (blue) source.

down by particle type ($\pi^{\pm}$, $\mathrm{K}^{\pm}$, $\overset{(-)}{\mathrm{p}}$) and shown in comparison with the marginal pseudorapidity ALICE data. Statistical and systematic uncertainties are added in quadrature.

Similar to the baryon-stopping model, all parameters of the forward- and backward-going sources are set to identical values due to the symmetry of the collisions. In addition, most of the parameters of the central sources are assumed to be identical to their forward- and backward-going counterparts:

For the gluon-saturation parameters of the initial DIS–CGC states, this is simply due to the fact that they refer to the very same saturated gluon medium. Since the same medium is also responsible for the stopping of the participants in the previously presented baryon-stopping model, I adopt the values $\lambda = 0.2$ and $x_0 = 10^{-3}$ from section 4.3, which are compatible with all three fitting results obtained therein.

Regarding the asymptotic thermal distribution, the individual sources are expected to interact sufficiently to aim for a common equilibrium. Consequently, the transverse-flow exponent is fixed to $\nu_{\perp} = 0.7$ (Abelev *et al.* 2013; Acharya *et al.* 2020) for all sources, while the surface velocity fraction $\tilde{\beta}_{\mathrm{s}}$, the kinetic freeze-out temperature T, and the products of diffusivities and interaction timespans $D_{\perp}\,\Delta t, D_{\parallel}\,\Delta t$ are left as shared free parameters. The maximum longitudinal boost rapidity ζ_{max} tends to values near the beam rapidity for the central sources, whereas for the forward- and backward-going sources, significantly lower boost rapidities provide substantially better agreement between data and model. For this reason, $\zeta_{\mathrm{max},0}$ and $\zeta_{\mathrm{max},\pm}$ are decoupled and determined separately via WLS here, which is the only exception to the general procedure. Both parameters are restricted to the interval $[0, y_{\mathrm{b}}]$ to prevent eventual divergence, since the available data do not contain measurements close to the beam rapidities.

While the experimental data covers a large transverse-momentum region with particle number densities of many orders of magnitude, only a subset of all available data points is used in the parameter estimation and shown in the plots for technical reasons. Albeit the theoretical NDF is in principle well-defined on the entire phase space, numerical errors accumulate when solving the FPE in practice such that the model results are subject to large relative errors when the number densities become too small. Therefore, I only consider data points no smaller than 5 % of the largest measured density to guarantee that the NDF remains well above numerical accuracy, which effectively limits the discussion to phase-space regions with $p_{\perp} \leq 2\,\mathrm{GeV/c}$. This threshold is by no means fundamental, however, and could be lowered in the future by improving the numerical method.

As a consequence, the high-momentum tails do not play a major role in the comparison between data and model. Nonetheless, I use the finite value $n = -8$ for the associated Hagedorn exponent, based on analyses of mid-rapidity LHC data (Aamodt *et al.* 2011; Adare *et al.* 2011; Michael 1979). Firstly, because it has a minor effect on the marginal longitudinal

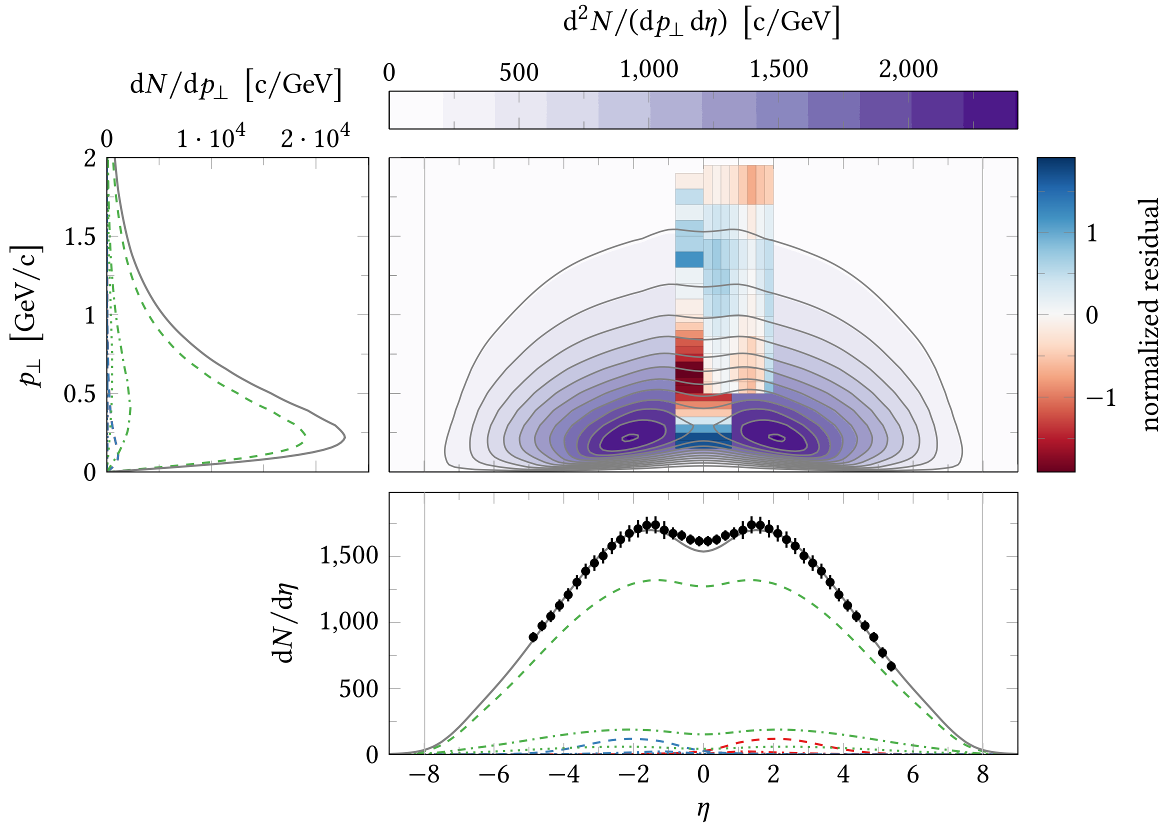


Figure 5.9: Joint (top right) and marginal (top left, bottom right) NDFs of unidentified charged hadrons in central collisions of ^{208}Pb nuclei at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV. The beam rapidity is marked by thin vertical lines.

In the joint plot, filled contour lines depict the model result; the associated normalized residuals for the ALICE (2018) and ATLAS (2015) data sets are displayed as colored rectangles, see section 3.4.3 for details.

In the marginal plots, the model result is shown as a solid curve; the additional curves indicate the contributions of pions (dashed), kaons (dash-dotted), and (anti-)protons (dotted) from the central (green), forward- (red), and backward-going (blue) sources. Marginal experimental data from the ALICE (2013) data set are displayed as circles, with vertical bars indicating the total uncertainties; the bins are smaller than the symbol size.

NDFs that contain high-$p_\perp$ contributions by definition, and secondly because of its positive side effect on numerical stability.

Table 5.2 on page 78 lists the WLS results for the free model parameters. For the two RHIC and the lower-energy LHC data sets, the GOFs indicate a good agreement, while the fit at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$ is less convincing. The latter can be attributed at least in part to the undulating course of the ALICE (2017) data points for $|\eta| \gtrsim 2$, which the model cannot reproduce. Since this pattern is not present in their counterparts at 2.76 TeV, it is not clear whether this is a physical effect or caused by other means. Furthermore, for both 2.76 TeV and 5.02 TeV, the model shows a systematic deviation from the transverse mid-rapidity ALICE (2018) data; being too low for $p_\perp \lesssim 0.4\,\mathrm{GeV/c}$ and too high for $0.4\,\mathrm{GeV/c} \lesssim p_\perp \lesssim 1.1\,\mathrm{GeV/c}$. At the same time, however, there is not much discrepancy with the ATLAS (2015) data set, which also covers a part of this phase-space domain. Whether a similar deviation occurs at RHIC energies is not entirely clear, since the BRAHMS data do not completely cover the maximum of the NDF.

The products of transverse diffusivity and interaction timespan, $D_\perp\,\Delta t$, are pushed towards very large values by the WLS algorithm for the LHC data, which signals a fast thermalization of the transverse degrees of freedom. This divergence is neither problematic nor surprising from a physical point of view, since thermal models with collective expansion or modified high-momentum tails are known to provide an adequate description of transverse charged-hadron distributions (see, *e. g.*, Wolschin 2016, and references therein). However, numerical solution of the FPE is troublesome in this case, because very small time steps are needed to obtain accurate results. To circumvent this problem, an alternative approach detailed in appendix D is taken with respect to $D_\perp\,\Delta t$, where the transverse degrees of freedom are assumed to fully equilibrate in each time step. This mimics the limit $D_\perp\,\Delta t \to \infty$, which hardly affects the resulting NDF quantitatively when $D_\perp\,\Delta t$ is already large, but stabilizes the numerical method. The lack of a finite WLS result for $D_\perp\,\Delta t$ is not a real drawback here since the associated uncertainties also diverge, leaving only a qualitative interpretation of the parameter in any case.

For the lower RHIC energies, the divergence in $D_\perp\,\Delta t$ is less pronounced, but the uncertainties are already relatively large. Therefore, and for consistency, the procedure described above is also applied here. This approximation leads to a moderate increase in $k_{\mathrm{B}}T$ (about 10 MeV for the higher RHIC energy); the remaining free model parameters are hardly affected.

For $D_\parallel\,\Delta t$, on the other hand, relatively small finite values appear to provide the best agreement between model and data. Interestingly, the fit predicts $D_\parallel\,\Delta t$ to vanish for the smallest collision energy, which suggests that almost no thermalization of the longitudinal degrees of freedom takes place here. With rising $\sqrt{s_{\mathrm{NN}}}$, the parameter appears to

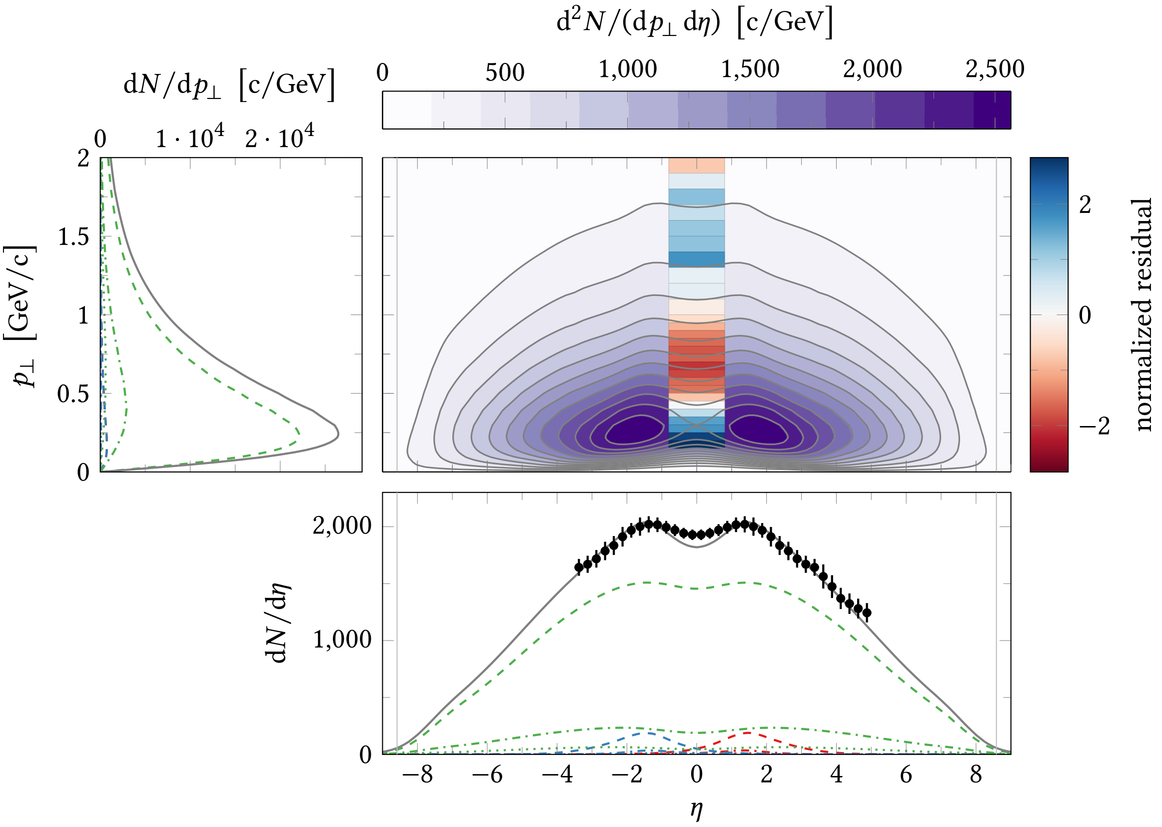


Figure 5.10: Joint (top right) and marginal (top left, bottom right) NDFs of unidentified charged hadrons in central collisions of ^{208}Pb nuclei at $\sqrt{s_{NN}} = 5.02$ TeV. The beam rapidity is marked by thin vertical lines.

In the joint plot, filled contour lines depict the model result; the associated normalized residuals for the ALICE (2018) data set are displayed as colored rectangles, see section 3.4.3 for details.

In the marginal plots, the model result is shown as a solid curve; the additional curves indicate the contributions of pions (dashed), kaons (dash-dotted), and (anti-)protons (dotted) from the central (green), forward- (red), and backward-going (blue) sources. Marginal experimental data from the ALICE (2017) data set are displayed as circles, with vertical bars indicating the total uncertainties; the bins are smaller than the symbol size.

increase monotonically, indicating a growing importance of thermalization, and finally reaches a maximum value close to 1. However, the uncertainties involved are quite large, and longitudinal phase-space coverage is sparse at the lowest collision energy, so these conclusions should be taken with caution.

The total number of particles ($N = N_0 + N_+ + N_-$) grows monotonically with collision energy for pions and unidentified hadrons, which is consistent with physical intuition since more and more energy is available for pair production. However, the allocation to the central (N_0) and the forward- and backward-going sources ($N_+ = N_- =: N_\pm$) cannot be conclusively clarified from the available data since uncertainties are relatively high. This is especially the case for the BRAHMS (2005) data set, where the WLS result for $N_\pm$ is significantly smaller than for all other data sets. In particular, it is less than half of that for the lower RHIC energy, which can hardly be explained physically. Presumably, this is rather a consequence of an insufficient restriction of the free-parameter space by the BRAHMS data sets, which allows the WLS method to approach unphysical minima. This can probably be remedied by adding additional data points, theoretically constraining the free parameters, or revising the model assumptions. For the time being, however, no clear trend can be inferred for $N_\pm$, whereas N_0 apparently grows monotonically with $\sqrt{s_{\mathrm{NN}}}$.

Note that the partitioning of particles between the central and forward/backward-going sources is not directly accessible experimentally and is therefore highly model-dependent. As such, earlier investigations of LHC data with purely longitudinal models yielded significantly larger $N_\pm : N_0$ ratios between $1 : 3$ and $1 : 6$, depending on the drift coefficient functions chosen (Kellers and Wolschin 2019).

A similar picture emerges for the flow parameters: The maximum boost rapidity $\zeta_{\mathrm{max},\pm}$ of the forward- and backward-going sources varies wildly across the data sets with no clear trend. This is again a consequence of the weak signal of these sources in the experimental data, which makes definite statements about their properties difficult. For the central sources, however, $\zeta_{\mathrm{max},0}$ consistently assumes a value close to the beam rapidity. In the case of the BRAHMS data sets, the parameter even reaches the upper limit y_{b}; whether this reflects actual physical behavior is unfortunately not clear due to the sparse phase-space coverage at high longitudinal (pseudo-)rapidities.

Interestingly, the transverse surface velocity fraction $\tilde{\beta}_{\mathrm{s}}$ is compatible with zero for the RHIC pion data, whereas a finite value around 0.6 is found for unidentified hadrons at the two LHC energies. This could point to a distinct collision-energy dependence of transverse flow, implying that high energy densities or particle numbers may be required for its formation. While the $h^\pm$ results compare well with those obtained previously (see below), there are no reference values yet for π^-, however, so further independent research is needed to confirm or refute this hypothesis.

Table 5.2: Free model parameters for the thermalization of charged hadrons in relativistic heavy-ion collisions, obtained via WLS with respect to the data sets listed in table 5.1. The associated GOF statistic ("χ^2/dof") is given in the last column. Numbers in parentheses indicate the estimated uncertainty of the preceding digit; for parameter values enclosed in brackets or without uncertainties, see text.

data set(s)	N_0	$N_\pm$	$D_\parallel \Delta t$	$\tilde{\beta}_\mathrm{s}$	$\zeta_{\mathrm{max},0}$	$\zeta_{\mathrm{max},\pm}$	$k_\mathrm{B}T/\mathrm{MeV}$	χ^2/dof
BRAHMS (2010)	790(60)	210(40)	0	0	4.2	2.2(5)	97(15)	$\frac{21.1}{87}$
BRAHMS (2005)	1750(100)	80(30)	0.6(3)	0	5.4	0	103(9)	$\frac{131}{120}$
ALICE (2018) ALICE (2013) ATLAS (2015)	15 840(250)	460(110)	0.69(5)	0.62(8)	7.1(2)	0.9(2)	102(10)	$\frac{63.5}{141}$
ALICE (2018) ALICE (2017)	19 980(640)	550(150)	1.2(4)	0.62(7)	7.8(5)	1.5(5)	110(10)	$\frac{80.8}{54}$

Albeit uncertainties are high, the kinetic freeze-out temperature shows a mild increase with collision energy, from $k_\mathrm{B}T \approx 100\,\mathrm{MeV}$ at the two RHIC energies to $110\,\mathrm{MeV}$ at the highest LHC energy. This can be well explained by the higher energy density, combined with a more rapid expansion of the fireball, which leaves the hadrons in a higher-temperature state when interactions cease due to large mean free paths. Fits of blast-wave and transverse-flow models to purely transverse mid-rapidity LHC data by Abelev *et al.* (2012, 2013) and Acharya *et al.* (2020) resulted in $85\,\mathrm{MeV} \lesssim k_\mathrm{B}T \lesssim 105\,\mathrm{MeV}$, combined with a higher transverse-flow velocity. As changes in temperature and flow velocity have a similar effect on transverse thermal distributions (*cf.* section 5.2), the different model results are not in conflict. The shift in the $(T, \tilde{\beta}_\mathrm{s})$ parameter space could be a consequence of the additional constraints imposed by the inclusion of longitudinal data points. While this requires further investigation, it might provide a method to obtain more accurate estimates of the freeze-out temperature and transverse-flow parameters in the future.

6. Conclusion

In this thesis, a formalism for relativistic phase-space diffusion on the basis of stochastic drift–diffusion processes was investigated in the context of relativistic heavy-ion collisions. Two specific models were derived to describe the stopping of net protons and the thermalization of charged hadrons, and compared with experimental data from the Super Proton Synchrotron (SPS), Relativistic Heavy Ion Collider (RHIC), and Large Hadron Collider (LHC). The dominant physical processes in different phase-space regions were represented by multiple particle sources with distinct fluctuation–dissipation relations (FDRS). In both cases, measured single-particle number density functions (NDFS) in transverse-momentum and (pseudo-)rapidity space could be reproduced to a good extent, and the free model parameters determined via weighted least squares were mostly compatible with previous findings in one effective dimension.

Major improvements over the previous models by Wolschin and Biyajima *et al.* consist in the systematic construction of the drift and diffusion coefficient functions from the expected FDRS and the extension from one to two effective dimensions by including the transverse degrees of freedom. Since the revised approach is based entirely on the principles of stochastic calculus, coordinate transformations of all mathematical objects involved can be performed without difficulty. Thereby, calculations can always be performed in the most convenient coordinate systems.

Capturing the complex dynamical processes occurring during a relativistic heavy-ion collision in a statistical model is, of course, an ambitious aim, and the present relatively simple treatment may be refined in several ways if a more accurate description is desired. Especially the dependence of the diffusion coefficients on the particle momentum, which has not yet been addressed, offers an opportunity to incorporate further microscopic or phenomenological considerations. In addition, more attention could be paid to the different phases of the collision event by performing multiple successive time evolutions whose coefficient functions reflect the prevailing dynamical processes.

A promising candidate with respect to the latter is the stopping model, where the observed net-proton distributions could not fully be explained by a transition from a Fermi gas to a final state shaped by deep inelastic scattering (DIS) with gluon saturation. This might indicate the significance of further interactions and could be explored in a two-stage model that complements the initial DIS with a subsequent incipient thermalization.

To quantify the net-proton loss, one may devise complementary drift–diffusion models for the production of baryons other than protons or neutrons. The resulting distributions could then be subtracted from the participant-proton NDF for a revised net-proton model or added to obtain an estimate of the net-baryon distribution, whose particle number is conserved throughout the collision and equal to the number of participant nucleons. For the latter variant, comparisons with data would unfortunately be limited to marginal net-baryon distributions in longitudinal-rapidity space, since no experimental net-baryon data in two effective dimensions have been published to date.

In the thermalization model, the role of the forward- and backward-going sources, which account for valence-quark–gluon interactions in the DIS picture, needs more investigation. Assessing the relative importance of the different DIS processes for charged-hadron production from the final NDFs alone proved highly model-dependent, so that it seems advisable to constrain the particle numbers of the three sources by additional criteria. Also, a separate treatment of the dynamics before and after hadronization or the inclusion of further particle-production channels besides gluons and valence quarks should improve the validity of the model.

As for the numerical implementation, it would be interesting to consider alternative techniques for solving the time-evolution equation, *e. g.*, spectral methods. This may enable to include data points with very low particle densities in the analysis, which have been omitted so far, and to investigate the properties of the Fokker–Planck operators.

The structure of phase-space drift–diffusion models is quite generic, and besides the presented models on stopping and thermalization, entirely new applications in the context of relativistic heavy-ion collisions are conceivable.

For example, it might be worthwhile to consider the phenomena of jet quenching or elliptic flow in this formalism. The latter would require consideration of the hitherto largely ignored third spatial dimension, which can be implemented natively by choosing the angle in the transverse plane as an additional coordinate; the experience gained could in turn be incorporated into the two existing models to open up the study of stopping and thermalization in non-central collisions. Besides, the general approach is not limited to single-particle distribution functions, as correlations can be modeled by combining multiple particle trajectories into a single stochastic process.

It is my hope that the above, as well as this thesis as a whole, may serve as an inspiration or starting point for future research in this rich topic area.

Appendix

A. Stochastic integration

Consider the stochastic integral

$$I := \int_{t_a}^{t_b} W(t)\,\mathrm{d}W(t)\,, \tag{A.1}$$

which is defined through the limit

$$I = \lim_{N\to\infty} \sum_{k=1}^{N} \Big[\tfrac{1+\lambda}{2}\,W(t_k) + \tfrac{1-\lambda}{2}\,W(t_{k-1})\Big]\,\big[W(t_k) - W(t_{k-1})\big] \tag{A.2}$$

for a partition $t_a = t_0 < t_1 < \cdots < t_N = t_b$ of the interval $[t_a, t_b]$ with some discretization parameter $\lambda \in [-1, +1]$. Performing the products results in

$$I = \tfrac{1}{2} \lim_{N\to\infty} \sum_{k=1}^{N} \big[W(t_k)^2 - W(t_{k-1})^2\big] + \tfrac{\lambda}{2} \lim_{N\to\infty} \sum_{k=1}^{N} \big[W(t_k) - W(t_{k-1})\big]^2\,, \tag{A.3}$$

where the first term is a telescope sum that is independent of N because $t_0 = t_a$ and $t_N = t_b$, while the second term contains only squares of independent Wiener-process increments. Using the fact that the latter are normally distributed with variance $t_k - t_{k-1}$, the above can be rewritten as

$$I = \tfrac{1}{2} \lim_{N\to\infty} \big[W(t_N)^2 - W(t_0)^2\big] + \tfrac{\lambda}{2} \lim_{N\to\infty} \sum_{k=1}^{N} \big(t_k - t_{k-1}\big)\, Z_k^2 \tag{A.4}$$

with the standardized random variables

$$Z_k := \frac{W(t_k) - W(t_{k-1})}{\sqrt{t_k - t_{k-1}}} \sim \mathcal{N}(0,1) \quad \text{for} \quad k = 1, \dots, N \tag{A.5}$$

whose sum of squares is χ^2-distributed with N degrees of freedom,

$$S := \sum_{k=1}^{N} Z_k^2 \sim \chi^2(N)\,. \tag{A.6}$$

For very large N, the spacing of the partition is approximately $t_k - t_{k-1} \approx (t_b - t_a)/N$, while $S \approx E_S(S) = N$ as a consequence of the law of large numbers, such that

$$I = \frac{1}{2}\left[W(t_b)^2 - W(t_a)^2\right] + \frac{\lambda}{2}\left(t_b - t_a\right). \tag{A.7}$$

B. Rapidity coordinates

Consider a particle with finite mass m and energy-momentum vector $\bar{p}$ in a $(1+d)$-dimensional Minkowskian space. Let $\mathfrak{F}_A$ and $\mathfrak{F}_B$ denote two inertial frames of reference and $\vec{v}_{A\to B}$ the velocity of a resting object in $\mathfrak{F}_B$ as measured by an observer at rest in $\mathfrak{F}_A$.[1] Finally, define Cartesian coordinate systems for both frames such that all corresponding coordinate axes in $\mathfrak{F}_A$ and $\mathfrak{F}_B$ are parallel and the relative motion occurs parallel to the last axis, $\vec{v}_{A\to B} =: v_{A\to B}\,\vec{e}_d$. Then, the Cartesian vector components of $\bar{p}$ in both frames are connected via the *Lorenz boost*

$$\begin{pmatrix} p_A^0 \\ p_A^1 \\ \vdots \\ p_A^{d-1} \\ p_A^d \end{pmatrix} = \begin{pmatrix} \gamma_{A\to B} & 0 & \cdots & 0 & \alpha_{A\to B} \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ \alpha_{A\to B} & 0 & \cdots & 0 & \gamma_{A\to B} \end{pmatrix} \cdot \begin{pmatrix} p_B^0 \\ p_B^1 \\ \vdots \\ p_B^{d-1} \\ p_B^d \end{pmatrix} \tag{B.1}$$

with the shorthands

$$\alpha_{A\to B} := \beta_{A\to B}\,\gamma_{A\to B}\,, \qquad \beta_{A\to B} := \frac{v_{A\to B}}{c}\,, \qquad \gamma_{A\to B} := \frac{1}{\sqrt{1-\beta_{A\to B}^2}}\,. \tag{B.2}$$

By defining the *relative rapidity* of B with respect to A,

$$w_{A\to B} := \operatorname{artanh}\bigl(\beta_{A\to B}\bigr)\,, \tag{B.3}$$

the transformation (B.1) can be rewritten as a “hyperbolic rotation” in the Minkowskian $(0,d)$-plane by a hyperbolic angle $w_{A\to B}$, since

$$\alpha_{A\to B} = \sinh\bigl(w_{A\to B}\bigr)\,, \qquad \gamma_{A\to B} = \cosh\bigl(w_{A\to B}\bigr)\,. \tag{B.4}$$

---

[1] Reversing the roles of $\mathfrak{F}_A$ and $\mathfrak{F}_B$ results in $\vec{v}_{B\to A} = -\vec{v}_{A\to B}$ according to the principle of relativity.

B.1. Parallel particle movement

If the particle moves parallel to $\vec{v}_{A\to B}$ in $\mathfrak{F}_A$ with $\vec{p}_A = p_A\,\vec{e}_d$, it remains parallel to $\vec{e}_d$ in $\mathfrak{F}_B$, $\vec{p}_B = p_B\,\vec{e}_d$, as can easily be seen from B.1. Then, by defining the *particle rapidities*

$$w_A := \operatorname{arsinh}\left(\frac{p_A}{mc}\right) = \operatorname{sgn}(p_A)\operatorname{arcosh}\left(\frac{p_A^0}{mc}\right) = \operatorname{artanh}\left(\frac{p_A}{p_A^0}\right), \tag{B.5a}$$

$$w_B := \operatorname{arsinh}\left(\frac{p_B}{mc}\right) = \operatorname{sgn}(p_B)\operatorname{arcosh}\left(\frac{p_B^0}{mc}\right) = \operatorname{artanh}\left(\frac{p_B}{p_B^0}\right), \tag{B.5b}$$

the matrix–vector equation (B.1) reduces to the single line

$$w_A = w_B + w_{A\to B} \tag{B.6}$$

after applying the hyperbolic addition theorems, which provides a particularly simple representation of this kind of Lorentz transformations.

B.2. Orientation-preserving rapidity coordinates

In case $p_A^1, \dots, p_A^{d-1}$ are non-zero, it is still possible to define rapidities $w_A^1, \dots, w_A^d$ for the individual space-like coordinate axes via

$$w_A^i := \operatorname{arcosh}\left(\frac{p_A^0}{mc}\right)\frac{p_A^i}{\|\vec{p}_A\|} \quad \text{for} \quad i = 1, \dots, d, \tag{B.7}$$

and similarly for $\vec{p}_B$. From these, energy and momentum can be recovered through

$$p_A^0 = mc\cosh\big(\|\vec{w}_A\|\big), \qquad \vec{p}_A = mc\sinh\big(\|\vec{w}_A\|\big)\,\frac{\vec{w}_A}{\|\vec{w}_A\|}, \tag{B.8}$$

where $\vec{w}_A := \sum_{i=1}^{d} w_A^i\,\vec{e}_i$ is a Euclidean vector constructed from the rapidity coordinates. Consequently, the hyperbolic tangent of $\|\vec{w}_A\|$ yields $\|\vec{p}_A\|/p_A^0$, which is the particle's absolute velocity divided by the speed of light.

The definition (B.7) has the advantage that the rapidity vector $\vec{w}_A$ has the same orientation as the particle's momentum $\vec{p}_A$. However, there is no simple relation similar to eq. (B.6) if $\vec{w}_A$ and $\vec{v}_{A\to B}$ are not parallel.

B.3. Transverse–longitudinal rapidity coordinates

A simple transformation law under Lorentz boosts of type (B.1) for arbitrary momenta can be obtained by choosing different rapidity coordinates based on the hyperbolic sine and tangent for the momentum components orthogonal (“transverse”) and parallel (“longitudinal”) to the boost axis $\vec{e}_d$, respectively,

$$h_A^i := \operatorname{arsinh}\left(\frac{\left\|\vec{p}_A - p_A^d\,\vec{e}_d\right\|}{mc}\right)\frac{p_A^i}{\left\|\vec{p}_A - p_A^d\,\vec{e}_d\right\|} \quad \text{for} \quad i = 1, \dots, d-1\,, \tag{B.9a}$$

$$y_A := \operatorname{artanh}\left(\frac{p_A^d}{p_A^0}\right). \tag{B.9b}$$

This yields a relation similar to eq. (B.6),

$$h_A^i = h_B^i \quad \text{for} \quad i = 1, \dots, d-1\,, \qquad\qquad y_A = y_B + w_{A\to B}\,, \tag{B.10}$$

but at the cost of slightly more complicated expressions for energy and momentum,

$$p_A^0 = mc\cosh\left(\left\|\vec{h}_A\right\|\right)\cosh\left(y_A\right), \tag{B.11a}$$

$$\vec{p}_A - p_A^d\,\vec{e}_d = mc\sinh\left(\left\|\vec{h}_A\right\|\right)\frac{\vec{h}_A}{\left\|\vec{h}_A\right\|}\,, \qquad p_A^d = mc\cosh\left(\left\|\vec{h}_A\right\|\right)\sinh\left(y_A\right). \tag{B.11b}$$

Here, the Euclidean vector $\vec{h}_A := \sum_{i=1}^{d-1} h_A^i\,\vec{e}_i$ always points in the same direction as the transverse part of the particle’s momentum.

Note that both eqs. (B.7) and (B.9) reproduce eq. (B.5) if the transverse momentum components vanish.

C. Thermal particle distributions

Section 5.2 introduced a family of thermal distributions whose properties will be considered in more detail here. When expressed in Cartesian momentum coordinates, their PDF has the form

$$f_{P_{\mathrm{th}}}\left(p^1, p^2, p^3\right) = \frac{C_{\mathrm{th}}}{V} \int_{\Sigma} \frac{\mathrm{d}\bar{\sigma}^* \bar{p}}{p^0} \exp\left(\frac{mc^2 - \bar{u}^* \bar{p}}{k_{\mathrm{B}} T}\right), \tag{C.1}$$

using a Minkowski metric with signature $(1, 3) \triangleq (+, -, -, -)$. The shape of the distribution is determined by the temperature T and the local proper velocity $\bar{u}$ of the thermal reservoir, which is enclosed in some hypersurface $\Sigma \subseteq \mathbb{R}^{1,3}$. The energy-like component p^0 of the particle momentum is fixed by the mass-shell condition $mc = \sqrt{\bar{p}^* \bar{p}}$, where $\bar{p}^*$ denotes the dual of the Minkowskian vector $\bar{p}$.

For any coordinate triple $\left(q^1, q^2, q^3\right)$ used to parameterize the spacetime positions $\bar{\sigma} \in \Sigma$ in the reservoir, the hypersurface element $\mathrm{d}\bar{\sigma}$ of Σ follows from

$$\mathrm{d}\sigma_\alpha = \mathrm{d}q^1\, \mathrm{d}q^2\, \mathrm{d}q^3 \sum_{\beta,\gamma,\delta=0}^{3} \epsilon_{\alpha\beta\gamma\delta} \frac{\partial \sigma^\beta}{\partial q^1} \frac{\partial \sigma^\gamma}{\partial q^2} \frac{\partial \sigma^\delta}{\partial q^3} \tag{C.2}$$

with the Levi–Civita symbol ϵ in four dimensions. The PDF is normalized by the thermal-reservoir volume $V := \int_\Sigma \mathrm{d}\bar{\sigma}^* \bar{u}/c$ and an additional constant C_{th} that is independent of the reservoir.

If the reservoir has a rotational symmetry with respect to $\vec{e}_3$ – which will always be the case in the following –, the corresponding PDF for transverse and longitudinal rapidity coordinates (*cf.* appendix B) reads

$$\begin{aligned} f_{(H,Y)_{\mathrm{th}}}\left(h, y\right) = {} & 2\pi\, (mc)^3 \sinh(h) \cosh(h)^2 \cosh(y) \\ & \times f_{P_{\mathrm{th}}}\left(mc \sinh(h), 0, mc \cosh(h) \sinh(y)\right). \end{aligned} \tag{C.3}$$

C.1. Normalizing constant

The value of C_{th} can be obtained by rearranging and then integrating eq. (C.1) over the entire momentum space,

$$C_{\mathrm{th}}^{-1} = \frac{1}{V}\int_{\Sigma}\mathrm{d}\bar{\sigma}^{*}\underbrace{\int_{mc\mathbb{R}^3}\frac{\mathrm{d}^3p}{p^0}\,\bar{p}\exp\!\left(\frac{mc^2-\bar{u}^{*}\bar{p}}{k_{\mathrm{B}}T}\right)}_{=:\bar{I}}\,. \tag{C.4}$$

To solve the integral $\bar{I}$, it is convenient to choose transverse and longitudinal rapidity coordinates h and y with respect to the axis defined by the local reservoir velocity, $\vec{u}/\|\vec{u}\|$. Defining $w := \operatorname{arcosh}(u^0/c)$ and $\kappa := mc^2/(k_{\mathrm{B}}T)$ then yields

$$\bar{I} = (mc)^3\int_0^{\infty}\mathrm{d}h\,\sinh(h)\cosh(h)\int_{-\infty}^{+\infty}\mathrm{d}y\exp\big(\kappa-\kappa\cosh(h)\cosh(y-w)\big)\int_0^{2\pi}\mathrm{d}\phi\,\bar{\alpha} \tag{C.5}$$

with

$$\int_0^{2\pi}\mathrm{d}\phi\,\alpha^0 = 2\pi\cosh(h)\cosh(y) \qquad \text{and} \qquad \int_0^{2\pi}\mathrm{d}\phi\,\vec{\alpha} = 2\pi\cosh(h)\sinh(y)\,\frac{\vec{u}}{\|\vec{u}\|}\,. \tag{C.6}$$

Performing the substitution $y \mapsto y - w =: y'$, expanding the arguments of the hyperbolic functions via the hyperbolic addition theorems, and eliminating the odd integrals in y' results in

$$\begin{aligned}\bar{I} = 2\pi\,(mc)^3\,\frac{\bar{u}}{c}\int_0^{\infty}\mathrm{d}h\,\sinh(h)\cosh(h)^2\int_{-\infty}^{+\infty}\mathrm{d}y'\,\cosh(y')&\\ \times\exp\big(\kappa-\kappa\cosh(h)\cosh(y')\big)\,.&\end{aligned} \tag{C.7}$$

With $\cosh(w') := \cosh(h)\cosh(y)$, the above can then be rewritten as

$$\bar{I} = 4\pi\,(mc)^3\,\frac{\bar{u}}{c}\int_0^{\infty}\mathrm{d}w'\,\sinh(w')^2\cosh(w')\exp\big(\kappa-\kappa\cosh(w')\big) \tag{C.8a}$$

$$= 4\pi\,(mc)^3\,\frac{\bar{u}}{c}\exp(\kappa)\,\frac{K_2(\kappa)}{\kappa}\,, \tag{C.8b}$$

where a hyperbolic addition theorem, a partial integration, and the definition of the modified Bessel function of the second kind and order a,

$$K_a(x) := \int_0^{\infty}\mathrm{d}t\cosh(at)\exp\big(-x\cosh(t)\big) \ \text{ for } \ x>0\,, \tag{C.9}$$

was used. Consequently,

$$C_{\mathrm{th}}^{-1} = \frac{1}{V}\int_{\Sigma}\mathrm{d}\bar{\sigma}^{*}\bar{I} = 4\pi\,(mc)^3\exp(\kappa)\,\frac{K_2(\kappa)}{\kappa}\,. \tag{C.10}$$

C.2. Maxwell–Jüttner distribution

The simplest case (Jüttner 1911) is a reservoir at rest, $\bar{u} := c\,\bar{e}_0$, on an isochronous hypersurface,

$$\Sigma := \left\{ \bar{\sigma} \in \mathbb{R}^{1,3} \,\middle|\, \sigma^0 = \text{const.}, \vec{\sigma} \in \mathbf{V} \right\}, \tag{C.11}$$

for an arbitrary spatial volume $\mathbf{V} \subseteq \mathbb{R}^3$. Choosing Cartesian coordinates $\left(x^1, x^2, x^3\right)$ to parameterize $\mathbf{V}$, the hypersurface element reduces to

$$\mathrm{d}\bar{\sigma} = \mathrm{d}^3 x\, \bar{e}_0\,, \tag{C.12}$$

so that

$$\mathrm{d}\bar{\sigma}^* \bar{u} = \mathrm{d}^3 x\, c \qquad \text{and} \qquad \mathrm{d}\bar{\sigma}^* \bar{p} = \mathrm{d}^3 x\, p^0\,. \tag{C.13}$$

In consequence, the volume integral drops out, leaving

$$f_{P_\text{th}}\left(p^1, p^2, p^3\right) = C_\text{th} \exp\left(\frac{mc^2 - cp^0}{k_\text{B} T}\right) \tag{C.14}$$

since $\bar{u}^* \bar{p} = cp^0$.

C.3. Spherical expansion

Consider an isotropic reservoir,

$$\Sigma := \left\{ \bar{\sigma}(r, \theta, \phi) = c\,t(r)\,\bar{e}_0 + r\,\bar{e}_r(\theta, \phi) \,\middle|\, r \in [0, R], \theta \in [0, \pi], \phi \in [0, 2\pi) \right\} \tag{C.15}$$

with

$$\bar{e}_r(\theta, \phi) := \sin(\theta)\cos(\phi)\,\bar{e}_{1'} + \sin(\theta)\sin(\phi)\,\bar{e}_{2'} + \cos(\theta)\,\bar{e}_{3'}\,, \tag{C.16}$$

that radially expands with proper velocity $\vec{u}(r, \theta, \phi) := \|\vec{u}\|(r)\,\vec{e}_r(\theta, \phi)$. The time-coordinate function $t(r)$ is chosen such that

$$c\,\frac{\partial t}{\partial r} = \frac{\|\vec{u}\|}{u^0}\,, \tag{C.17}$$

which leads to a kinetic freeze-out at constant local time (Lee, Heinz, and Schnedermann 1990). Without loss of generality, one can always choose the zenith direction equal to the direction of particle movement when performing the thermal-reservoir integral, so that

$$\vec{e}_{3'} = \frac{\vec{p}}{\|\vec{p}\|} \quad \Longrightarrow \quad \theta = \arccos\left(\vec{e}_r \cdot \vec{e}_{3'}\right) = \arccos\left(\frac{\vec{\sigma} \cdot \vec{p}}{\|\vec{\sigma}\|\|\vec{p}\|}\right). \tag{C.18}$$

Then,

$$\mathrm{d}\bar{\sigma} = \mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi\, r^2 \sin(\theta)\,\frac{\bar{u}}{u^0}\,, \tag{C.19a}$$

$$\mathrm{d}\bar{\sigma}^*\bar{u} = \mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi\,r^2\sin(\theta)\,\frac{c^2}{u^0}\,, \qquad \mathrm{d}\bar{\sigma}^*\bar{p} = \mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi\,r^2\sin(\theta)\,\frac{\bar{u}^*\bar{p}}{u^0}\,, \tag{C.19b}$$

$$\bar{u}^*\bar{p} = u^0p^0 - \|\vec{u}\|\|\vec{p}\|\cos(\theta)\,. \tag{C.19c}$$

The angular integrations in eq. (C.1) can be easily solved by performing the substitutions

$$\theta \mapsto \cos(\theta) \mapsto \frac{u^0p^0 - \|\vec{u}\|\|\vec{p}\|\cos(\theta)}{k_\mathrm{B}T}\,, \tag{C.20}$$

which leads to the integral

$$\int \mathrm{d}x\,(-x)\exp(\kappa - x) = (1+x)\exp(\kappa - x) + \text{const.} =: L(x)\,. \tag{C.21}$$

Consequently, the spherically-expanding thermal PDF has the form

$$f_{P_\mathrm{th}}\big(p^1, p^2, p^3\big) = 4\pi\,\frac{C_\mathrm{th}}{V}\int_0^1 \mathrm{d}r\,r^2\,K\!\left(\frac{u^0p^0}{k_\mathrm{B}T}, \frac{\|\vec{u}\|\|\vec{p}\|}{k_\mathrm{B}T}\right) \tag{C.22}$$

with the auxiliary function

$$K(x_1, x_2) := \frac{L(x_1 - x_2) - L(x_1 + x_2)}{2x_1x_2}\,. \tag{C.23}$$

C.3.1. Blast wave and power-law flow profile

For a constant expansion velocity, the remaining radial integral can also be solved analytically, yielding the so-called blast-wave model (Bondorf, Garpman, and Zimanyi 1978; Siemens and Rasmussen 1979),

$$V = \frac{4\pi R^3}{3}\,\frac{c}{u^0}\,, \qquad f_{P_\mathrm{th}}\big(p^1, p^2, p^3\big) = C_\mathrm{th}\,\frac{u^0}{c}\,K\!\left(\frac{u^0p^0}{k_\mathrm{B}T}, \frac{\|\vec{u}\|\|\vec{p}\|}{k_\mathrm{B}T}\right). \tag{C.24}$$

Assuming a non-constant expansion velocity of the form (Lee and Heinz 1989)

$$\frac{\|\vec{u}\|}{u^0} = \left(\frac{r}{R}\right)^{\nu}\beta_\mathrm{s} \tag{C.25}$$

with positive exponent $\nu > 0$ and non-relativistic surface velocity fraction $\beta_\mathrm{s} \in (0, 1)$, the reservoir volume is given by

$$V = \frac{4\pi R^3}{3}\,{}_2F_1\!\left(-\tfrac{1}{2}, \tfrac{3}{2\nu}; 1 + \tfrac{3}{2\nu}; \beta_\mathrm{s}^2\right). \tag{C.26}$$

Here, the ordinary hypergeometric function ${}_2F_1$ encapsulates the integral

$$\int_0^1 \mathrm{d}t\,t^a\left(1 - t^bx\right)^c = \tfrac{1}{1+a}\,{}_2F_1\!\left(-c, \tfrac{1+a}{b}; 1 + \tfrac{1+a}{b}; x\right). \tag{C.27}$$

With this, the thermal PDF (C.22) can be written as

$$f_{P_{\mathrm{th}}}\left(p^1, p^2, p^3\right) = \frac{3C_{\mathrm{th}}}{{}_2F_1\left(-\frac{1}{2}, \frac{3}{2\nu}; 1 + \frac{3}{2\nu}; \beta_{\mathrm{s}}^2\right)} \int_0^1 \mathrm{d}\left[\frac{r}{R}\right] \left(\frac{r}{R}\right)^2 K\left(\frac{u^0 p^0}{k_{\mathrm{B}} T}, \frac{\|\vec{u}\|\|\vec{p}\|}{k_{\mathrm{B}} T}\right), \tag{C.28}$$

where the remaining integral is solved numerically.

C.3.2. Infinite-temperature limit

In the limit $\kappa \to 0$, the marginal PDF of eq. (C.22) with respect to the longitudinal rapidity y can be calculated analytically. The key is the substitution

$$h \mapsto c p^0/(k_{\mathrm{B}} T) = \kappa \cosh(h) \cosh(y) =: z\,, \tag{C.29}$$

which then allows to simplify the auxiliary functions K and L:

$$\begin{aligned}
\lim_{\kappa\to 0} f_{Y_{\mathrm{th}}}(y) &= \lim_{\kappa\to 0} \int_0^\infty f_{(H,Y)_{\mathrm{th}}}(h, y)\,\mathrm{d}h && \text{(C.30a)}\\
&= \frac{2\pi\,(mc)^3\,u^0}{c\cosh(y)^2} \lim_{\kappa\to 0} \frac{C_{\mathrm{th}}}{\kappa^3} \int_0^\infty z^2\,K\left(\frac{u^0 z}{c}, \frac{\|\vec{u}\|\sqrt{z^2-\kappa^2}}{c}\right) \mathrm{d}z && \text{(C.30b)}\\
&= \frac{u^0}{4c\cosh(y)^2} \int_0^\infty z^2 \lim_{\kappa\to 0} K\left(\frac{u^0 z}{c}, \frac{\|\vec{u}\| z}{c}\right) \mathrm{d}z && \text{(C.30c)}\\
&= \frac{c}{8\|\vec{u}\|\cosh(y)^2} \int_0^\infty \lim_{\kappa\to 0}\left[L\left(\tfrac{u^0-\|\vec{u}\|}{c} z\right) - L\left(\tfrac{u^0+\|\vec{u}\|}{c} z\right)\right] \mathrm{d}z && \text{(C.30d)}\\
&= \frac{c}{8\|\vec{u}\|\cosh(y)^2}\left[2\tfrac{c}{u^0-\|\vec{u}\|} - 2\tfrac{c}{u^0+\|\vec{u}\|}\right] && \text{(C.30e)}\\
&= \frac{1}{2\cosh(y)^2} && \text{(C.30f)}
\end{aligned}$$

C.4. Transverse–longitudinal expansion

Consider a thermal reservoir of the type (Schnedermann, Sollfrank, and Heinz 1993)

$$\Sigma := \left\{ \bar{\sigma}(r_\perp, \phi, \zeta) = c t(\zeta)\,\bar{e}_0 + r_\perp\,\bar{e}_{r_\perp}(\phi) + z(\zeta)\,\bar{e}_3 \,\middle|\, r_\perp \in [0, R_\perp], \phi \in [0, 2\pi), \zeta \in [\zeta_{\min}, \zeta_{\max}] \right\} \tag{C.31}$$

with the transverse radial unit vector

$$\bar{e}_{r_\perp}(\phi) := \cos(\phi)\,\bar{e}_1 + \sin(\phi)\,\bar{e}_2 \tag{C.32}$$

and the *Milne coordinates*

$$t(\zeta) := \tau \cosh(\zeta)\,, \qquad\qquad z(\zeta) := c\tau \sinh(\zeta) \tag{C.33}$$

for some constant proper freeze-out time τ. The velocity field is chosen to represent a Bjorken flow (Bjorken 1983) that also expands isotropically in the transverse direction,

$$\vec{u}(r_\perp, \phi, \zeta) := \|\vec{u}_\perp\|(r_\perp)\,\vec{e}_{r_\perp}(\phi) + u^0_\perp(r_\perp)\sinh(\zeta)\,\vec{e}_3\,, \tag{C.34}$$

where $\vec{u}_\perp$ is the projection of $\vec{u}$ onto the transverse plane. Analogously to the spherical case, the transverse radial expansion is assumed to follow a power-law,

$$\frac{\|\vec{u}_\perp\|}{u^0_\perp} := \left(\frac{r_\perp}{R_\perp}\right)^{\nu_\perp} \tilde{\beta}_\mathrm{s}\,, \tag{C.35}$$

with exponent $\nu_\perp > 0$ and surface velocity parameter $\tilde{\beta}_\mathrm{s} \in (0,1)$. Then, the surface element is given by

$$\mathrm{d}\bar{\sigma} = \mathrm{d}r_\perp\,\mathrm{d}\phi\,\mathrm{d}\zeta\, r_\perp \left[ct(\zeta)\,\bar{e}_0 + z(\zeta)\,\bar{e}_3\right]\,, \tag{C.36}$$

so that the reservoir volume is

$$V = \pi c\tau R_\perp^2\,(\zeta_\mathrm{max} - \zeta_\mathrm{min})\,{}_2F_1\!\left(\tfrac{1}{2}, \tfrac{1}{\nu_\perp}; 1 + \tfrac{1}{\nu_\perp}; \tilde{\beta}_\mathrm{s}^2\right). \tag{C.37}$$

The integrals in eq. (C.1) are best solved with the help of transverse and longitudinal rapidity coordinates h and y, which allow to compactly express the Minkowskian inner products,

$$\begin{aligned} f_{P_\mathrm{th}}(p^1, p^2, p^3) = {} & \frac{C_\mathrm{th}}{{}_2F_1\!\left(\tfrac{1}{2}, \tfrac{1}{\nu_\perp}; 1 + \tfrac{1}{\nu_\perp}; \tilde{\beta}_\mathrm{s}^2\right)} \frac{2}{\zeta_\mathrm{max} - \zeta_\mathrm{min}} \frac{1}{\cosh(y)} \\ & \times \int_0^1 \mathrm{d}\!\left[\frac{r_\perp}{R_\perp}\right] \int_{\zeta_\mathrm{min}}^{\zeta_\mathrm{max}} \mathrm{d}\zeta\, \frac{r_\perp}{R_\perp} \cosh(y - \zeta)\, I\big(\kappa - k_\mathrm{c}\cosh(y-\zeta), k_\mathrm{s}\big)\,, \end{aligned} \tag{C.38}$$

with the shorthands

$$k_\mathrm{c} := \kappa\,\frac{u^0_\perp}{c}\cosh(h)\,, \qquad\qquad k_\mathrm{s} := \kappa\,\frac{\|\vec{u}_\perp\|}{c}\sinh(h)\,. \tag{C.39}$$

The angular integral is encapsulated in the auxiliary function

$$I(x_1, x_2) := \exp(x_1)\, I_0(x_2)\,, \tag{C.40}$$

which is in turn defined via the zeroth-order modified Bessel function of the first kind

$$I_0(x) := \frac{1}{\pi} \int_0^{\pi} \mathrm{d}\phi \exp\bigl(x \cos(\phi)\bigr) . \tag{C.41}$$

If eq. (C.38) is transformed to rapidity space according to eq. (C.3), the hyperbolic cosine in the first line drops out, so that the entire PDF depends on y only via $y - \zeta$. It can then be rewritten as the average of a longitudinally boosted distribution that expands only in the transverse plane, as shown in eqs. (5.16a) and (5.16b).

C.5. Modified high-momentum tails

The exponential in eq. (C.1) can be replaced by the function $\widetilde{\exp}(x; n) := (1 + x/n)^n$ to account for the observed high-momentum behavior as discussed in section 5.2.4. This leaves the reservoir volume V unchanged, but affects the constant C_{th} and the auxiliary functions used in the spherically and transverse–longitudinally expanding PDFs. Specifically, the function L used in eq. (C.23) changes to

$$\tilde{L}(x; n) := \frac{(n+\kappa)^2 + n\,(n+\kappa)\,x - (n+1)\,x^2}{(n+1)\,(n+2)}\, \widetilde{\exp}(\kappa - x; n) + \text{const.}, \tag{C.42}$$

while eq. (C.40) takes the form

$$\tilde{I}(x_1, x_2; n) := \widetilde{\exp}(x_1; n)\,(1-z)^n\, {}_2F_1\Bigl(-n, \tfrac{1}{2}; 1; \tfrac{2z}{z-1}\Bigr) \quad \text{with} \quad z := \frac{x_2}{n + x_1} . \tag{C.43}$$

D. Fast transverse equilibration

Consider an FPO of the form (2.68),

$$L_{(H,Y)}(h,y) = \sum_{i,j=1}^{2} L^{ij}_{(H,Y)}(h,y)\,, \tag{D.1a}$$

$$L^{ij}_{(H,Y)}(h,y) = \frac{\partial}{\partial(h,y)^i}\, D^{ij}_{(H,Y)}(h,y) \left[-\frac{\partial \ln\bigl(f_{(H,Y)(\infty)}(h,y)\bigr)}{\partial(h,y)^j} + \frac{\partial}{\partial(h,y)^j} \right], \tag{D.1b}$$

for a two-dimensional drift–diffusion process (H, Y) in transverse- and longitudinal-rapidity space. It is then easy to see that the member PDF $f_{(H,Y)(t)}$ fulfills the identity

$$\begin{aligned} L^{ij}_{(H,Y)}(h,y)\, f_{(H,Y)(t)}(h,y) &= \frac{\partial}{\partial(h,y)^i}\, D^{ij}_{(H,Y)}(h,y) \\ &\quad \times \left[-\frac{\partial \ln\bigl(f_{(H,Y)(\infty)}(h,y)\bigr)}{\partial(h,y)^j} + \frac{\partial \ln\bigl(f_{(H,Y)(t)}(h,y)\bigr)}{\partial(h,y)^j} \right] f_{(H,Y)(t)}(h,y)\,. \end{aligned} \tag{D.2}$$

If $D^{11}_{(H,Y)}$ is much larger than the other matrix elements of the diffusivity, the subprocess H rapidly approaches its asymptotic distribution or, put differently, $H(t)$ is always close to $H(\infty)$ for any *given* realization of $Y(t)$, such that $f_{H(t)|Y(t)}(h\,|\,y) \approx f_{H(\infty)|Y(\infty)}(h\,|\,y)$ should hold in good approximation. The conditional member PDF that appears in the previous statement,

$$f_{H(t)|Y(t)}(h\,|\,y) = \frac{f_{(H,Y)(t)}(h,y)}{f_{Y(t)}(y)} = \frac{f_{(H,Y)(t)}(h,y)}{\int_0^\infty f_{(H,Y)(t)}(h,y)\,\mathrm{d}h}\,, \tag{D.3}$$

possesses the logarithmic derivatives

$$\frac{\partial \ln\bigl(f_{H(t)|Y(t)}(h\,|\,y)\bigr)}{\partial h} = \frac{\partial \ln\bigl(f_{(H,Y)(t)}(h,y)\bigr)}{\partial h}\,, \tag{D.4a}$$

$$\frac{\partial \ln\bigl(f_{H(t)|Y(t)}(h\,|\,y)\bigr)}{\partial y} = \frac{\partial \ln\bigl(f_{(H,Y)(t)}(h,y)\bigr)}{\partial y} - \frac{\partial \ln\bigl(f_{Y(t)}(y)\bigr)}{\partial y}\,. \tag{D.4b}$$

Applying the summands of the FPO to $f_{(H,Y)(t)}$ thus results in

$$\begin{aligned} L^{i1}_{(H,Y)}(h,y)\, f_{(H,Y)(t)}(h,y) = {} & \frac{\partial}{\partial(h,y)^i}\, D^{i1}_{(H,Y)}(h,y) \\ & \times \Bigg[\underbrace{-\frac{\partial \ln\big(f_{H(\infty)|Y(\infty)}(h\,|\,y)\big)}{\partial h} + \frac{\partial \ln\big(f_{H(t)|Y(t)}(h\,|\,y)\big)}{\partial h}}_{\approx 0}\Bigg] f_{(H,Y)(t)}(h,y)\,, \end{aligned} \tag{D.5}$$

$$\begin{aligned} L^{i2}_{(H,Y)}(h,y)\, f_{(H,Y)(t)}(h,y) = {} & \frac{\partial}{\partial(h,y)^i}\, D^{i2}_{(H,Y)}(h,y) \\ & \times \Bigg[\underbrace{-\frac{\partial \ln\big(f_{H(\infty)|Y(\infty)}(h\,|\,y)\big)}{\partial y} + \frac{\partial \ln\big(f_{H(t)|Y(t)}(h\,|\,y)\big)}{\partial y}}_{\approx 0} \\ & \qquad - \frac{\partial \ln\big(f_{Y(\infty)}(y)\big)}{\partial y} + \frac{\partial \ln\big(f_{Y(t)}(y)\big)}{\partial y}\Bigg] f_{(H,Y)(t)}(h,y)\,. \end{aligned} \tag{D.6}$$

If $D^{12}_{(H,Y)}(h,y)$ vanishes and $D^{22}_{(H,Y)}(h,y)$ is constant in h, one can then formulate an FPE for the marginal member PDF $f_{Y(t)}$ via

$$\frac{\partial}{\partial t} f_{Y(t)}(y) = \int_0^\infty \frac{\partial}{\partial t} f_{(H,Y)(t)}(h,y)\,\mathrm{d}h \tag{D.7}$$

$$= \int_0^\infty L_{(H,Y)}(h,y)\, f_{(H,Y)(t)}(h,y)\,\mathrm{d}h \tag{D.8}$$

$$\approx \int_0^\infty L^{22}_{(H,Y)}(h,y)\, f_{(H,Y)(t)}(h,y)\,\mathrm{d}h \tag{D.9}$$

$$\begin{aligned} = {} & \frac{\partial}{\partial y}\, D^{22}_{(H,Y)}(h,y) \Bigg[-\frac{\partial \ln\big(f_{Y(\infty)}(y)\big)}{\partial y} + \frac{\partial \ln\big(f_{Y(t)}(y)\big)}{\partial y}\Bigg] \\ & \times \int_0^\infty f_{(H,Y)(t)}(h,y)\,\mathrm{d}h \end{aligned} \tag{D.10}$$

$$= \underbrace{\frac{\partial}{\partial y}\, D^{22}_{(H,Y)}(h,y) \Bigg[-\frac{\partial \ln\big(f_{Y(\infty)}(y)\big)}{\partial y} + \frac{\partial}{\partial y}\Bigg]}_{=:L_Y(y)} f_{Y(t)}(y)\,, \tag{D.11}$$

where $D^{22}_{(H,Y)}(h,y)$ functions as the diffusivity of the approximately independent drift–diffusion process Y. After solving this equation, the joint PDF can be reconstructed through $f_{(H,Y)(t)}(h,y) \approx f_{H(\infty)|Y(\infty)}(h\,|\,y)\, f_{Y(t)}(y)$. Note that this procedure establishes a formal link to the effectively one-dimensional models in longitudinal-rapidity space (Wolschin, Biyajima, Mizoguchi, and Suzuki 2006) and may serve as an *a-posteriori* motivation for the assumptions made therein (Hoelck and Wolschin 2020).

Bibliography

Publications

Hoelck, J. and G. Wolschin (2020), “Baryon stopping as a relativistic Markov process in phase space,” Phys. Rev. Res. **2**, 033409.

In the above article, an earlier iteration of the baryon-stopping model in one effective dimension is discussed. The thesis at hand uses the same underlying mathematical framework and similar physical assumptions, but in an extended, two-dimensional approach.

References

Aad, G. *et al.* [ATLAS Collaboration] (2015), “Measurement of charged-particle spectra in Pb+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$ with the ATLAS detector at the LHC,” J. High Energy Phys. **2015**, 050.

Aamodt, K. *et al.* [ALICE Collaboration] (2011), “Production of pions, kaons and protons in pp collisions at $\sqrt{s} = 900\,\mathrm{GeV}$ with ALICE at the LHC,” Eur. Phys. J. C **71**, 1655.

Abbas, E. *et al.* [ALICE Collaboration] (2013), “Centrality dependence of the pseudorapidity density distribution for charged particles in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$,” Phys. Lett. B **726**, 610–622.

Abelev, B. *et al.* [ALICE Collaboration] (2012), “Pion, kaon, and proton production in central Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$,” Phys. Rev. Lett. **109**, 252301.

Abelev, B. *et al.* [ALICE Collaboration] (2013), “Centrality dependence of π, K, and p production in Pb-Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$,” Phys. Rev. C **88**, 044910.

Acharya, S. *et al.* [ALICE Collaboration] (2018), “Transverse momentum spectra and nuclear modification factors of charged particles in pp, p–Pb and Pb–Pb collisions at the LHC,” J. High Energy Phys. **2018**, 013.

Acharya, S. *et al.* [ALICE Collaboration] (2020), “Production of charged pions, kaons, and (anti-)protons in Pb–Pb and inelastic pp collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$,” Phys. Rev. C **101**, 044907.

Acharya, S. *et al.* [ALICE Collaboration] (2022), "Characterizing the initial conditions of heavy-ion collisions at the LHC with mean transverse momentum and anisotropic flow correlations," Phys. Lett. B **834**, 137393.

Adam, J. *et al.* [ALICE Collaboration] (2017), "Centrality dependence of the pseudorapidity density distribution for charged particles in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$," Phys. Lett. B **772**, 567–577.

Adare, A. *et al.* [PHENIX Collaboration] (2011), "Identified charged hadron production in p+p collisions at $\sqrt{s} = 200$ and $62.4\,\mathrm{GeV}$," Phys. Rev. C **83**, 064903.

Albacete, J. L. (2007), "Particle multiplicities in lead–lead collisions at the CERN large hadron collider from nonlinear evolution with running coupling corrections," Phys. Rev. Lett. **99**, 262301.

Andrae, R., T. Schulze-Hartung, and P. Melchior (2010), "Dos and don'ts of reduced chi-squared," arXiv:`1012.3754 [astro-ph.IM]`.

Angeli, I. and K. P. Marinova (2013), "Table of experimental nuclear ground state charge radii: An update," At. Data Nucl. Data Tables **99**, 69–95.

Appelshäuser, H. *et al.* [NA49 Collaboration] (1999), "Baryon stopping and charged particle distributions in central Pb+Pb collisions at 158 GeV per nucleon," Phys. Rev. Lett. **82**, 2471–2475.

Arsene, I. C. *et al.* [BRAHMS Collaboration] (2009), "Nuclear stopping and rapidity loss in Au+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 62.4\,\mathrm{GeV}$," Phys. Lett. B **677**, 267–271.

Arsene, I. C. *et al.* [BRAHMS Collaboration] (2010), "Kaon and pion production in central Au+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 62.4\,\mathrm{GeV}$," Phys. Lett. B **687**, 36–41.

Baier, R., Y. Mehtar-Tani, and D. Schiff (2006), "Has saturation physics been observed in deuteron–gold collisions at RHIC?" Nucl. Phys. A **764**, 515–527.

Bearden, I. G. *et al.* [BRAHMS Collaboration] (2004), "Nuclear stopping in Au+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200\,\mathrm{GeV}$," Phys. Rev. Lett. **93**, 102301.

Bearden, I. G. *et al.* [BRAHMS Collaboration] (2005), "Charged meson rapidity distributions in central Au+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200\,\mathrm{GeV}$," Phys. Rev. Lett. **94**, 162301.

Bernstein, M. N. (2019), *Demystifying measure-theoretic probability theory (part 1: probability spaces)*, `https://mbernste.github.io/posts/measure_theory_1/`.

Bernstein, M. N. (2020a), *Demystifying measure-theoretic probability theory (part 2: random variables)*, `https://mbernste.github.io/posts/measure_theory_2/`.

Bernstein, M. N. (2020b), *Demystifying measure-theoretic probability theory (part 3: expectation)*, `https://mbernste.github.io/posts/measure_theory_3/`.

Bezanson, J., A. Edelman, S. Karpinski, and V. B. Shah (2017), "Julia: A fresh approach to numerical computing," SIAM Rev. **59**, 65–98.

Biyajima, M., M. Ide, T. Mizoguchi, and N. Suzuki (2002), "Scaling behavior of $(N_{\mathrm{ch}})^{-1}\,\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ at $\sqrt{s_{\mathrm{NN}}} = 130\,\mathrm{GeV}$ by the PHOBOS collaboration and its implication: A possible explanation employing the Ornstein–Uhlenbeck process," Prog. Theor. Phys. **108**, 559–569.

Biyajima, M., T. Mizoguchi, N. Nakajima, N. Suzuki, and G. Wilk (2006), "Modified Hagedorn formula including temperature fluctuation: Estimation of temperatures at RHIC experiments," Eur. Phys. J. C **48**, 597–603.

Bjorken, J. D. (1983), "Highly relativistic nucleus–nucleus collisions: The central rapidity region," Phys. Rev. D **27**, 140–151.

Blaizot, J. P. and A. H. Mueller (1987), "The early stage of ultra-relativistic heavy ion collisions," Nucl. Phys. B **289**, 847–860.

Bogacki, P. and L. F. Shampine (1989), "A 3(2) pair of Runge–Kutta formulas," Appl. Math. Lett. **2**, 321–325.

Bondorf, J. P., S. I. A. Garpman, and J. Zimanyi (1978), "A simple analytic hydrodynamic model for expanding fireballs," Nucl. Phys. A **296**, 320–332.

Box, G. E. P. (1979), "Robustness in the strategy of scientific model building," in *Robustness in statistics*, edited by R. L. Launer and G. N. Wilkinson (Academic Press), pp. 201–236.

Brown, R. (1828), "XXVII. a brief account of microscopical observations made in the months of June, July, and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies," Philos. Mag., 2nd ser. **4**, 161–173.

Caucal, P. and Y. Mehtar-Tani (2021), "Anomalous diffusion in QCD matter," arXiv:`2109.12041 [hep-ph]`.

Chapman, S. (1928), "On the Brownian displacements and thermal diffusion of grains suspended in a non-uniform fluid," Proc. R. Soc. Lond. A **119**, 34–54.

Cubero, D., J. Casado-Pascual, J. Dunkel, P. Talkner, and P. Hänggi (2007), "Thermal equilibrium and statistical thermometers in special relativity," Phys. Rev. Lett. **99**, 170601.

Debbasch, F., K. Mallick, and J. P. Rivet (1997), "Relativistic Ornstein–Uhlenbeck process," J. Stat. Phys. **88**, 945–966.

Donsker, M. D. (1952), "Justification and extension of Doob's heuristic approach to the Kolmogorov–Smirnov theorems," Ann. Math. Stat. **23**, 277–281.

Dudley, R. M. (1966), "Lorentz-invariant Markov processes in relativistic phase space," Ark. Mat. **6**, 241–268.

Dumitru, A., A. Hayashigaki, and J. Jalilian-Marian (2006), "The color glass condensate and hadron production in the forward region," Nucl. Phys. A **765**, 464–482.

Dunkel, J. and P. Hänggi (2005a), "Theory of relativistic Brownian motion: The (1+1)-dimensional case," Phys. Rev. E **71**, 016124.

Dunkel, J. and P. Hänggi (2005b), "Theory of relativistic Brownian motion: The (1+3)-dimensional case," Phys. Rev. E **72**, 036106.

Dunkel, J. and P. Hänggi (2009), "Relativistic Brownian motion," Phys. Rep. **471**, 1–73.

Dunkel, J., P. Hänggi, and S. Hilbert (2009), "Non-local observables and lightcone-averaging in relativistic thermodynamics," Nat. Phys. **5**, 741–747.

Dunkel, J., P. Hänggi, and S. Weber (2009), "Time parameters and Lorentz transformations of relativistic stochastic processes," Phys. Rev. E **79**, 010101.

Einstein, A. (1905a), "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt," Ann. Phys. **322**, 132–148.

Einstein, A. (1905b), "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen," Ann. Phys. **322**, 549–560.

Einstein, A. (1905c), "Zur Elektrodynamik bewegter Körper," Ann. Phys. **322**, 891–921.

Einstein, A. (1905d), "Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?" Ann. Phys. **323**, 639–641.

Einstein, A. (1906), "Zur Theorie der Brownschen Bewegung," Ann. Phys. **324**, 371–381.

Fisk, D. L. (1965), "Quasi-martingales," Trans. Amer. Math. Soc. **120**, 369–369.

Fokker, A. D. (1914), "Die mittlere Energie rotierender elektrischer Dipole im Strahlungsfeld," Ann. Phys. **348**, 810–820.

Forndran, F. and G. Wolschin (2017), "Relativistic diffusion model with nonlinear drift," Eur. Phys. J. A **53**, 37.

Fürth, R. (1920), "Die Brownsche Bewegung bei Berücksichtigung einer Persistenz der Bewegungsrichtung. Mit Anwendungen auf die Bewegung lebender Infusorien," Z. Phys. **2**, 244–256.

Gardiner, C. W. (2009), *Stochastic methods, A handbook for the natural and social sciences*, 4th ed., Springer Series in Synergetics (Springer, Berlin and Heidelberg).

Golec-Biernat, K. and M. Wüsthoff (1998), "Saturation effects in deep inelastic scattering at low Q^2 and its implications on diffraction," Phys. Rev. D **59**, 014017.

Gribov, L., E. Levin, and M. Ryskin (1983), "Semihard processes in QCD," Phys. Rep. **100**, 1–150.

Hagedorn, R. (1965), "Statistical thermodynamics of strong interactions at high energies," Nuovo Cim. Suppl. **3**, 147–186.

Hagedorn, R. (1983), "Multiplicities, p_T distributions and the expected hadron ⟶ quark–gluon phase transition," Riv. Nuovo Cim. **6**, 1–50.

Hakim, R. (1968), "Relativistic stochastic processes," J. Math. Phys. **9**, 1805–1818.

Hänggi, P. (1978), "Stochastic processes. I, Asymptotic behaviour and symmetries," Helv. Phys. Acta **51**, 183–201.

Hänggi, P. and H. Thomas (1982), "Stochastic processes: Time evolution, symmetries and linear response," Phys. Rep. **88**, 207–319.

Hoelck, J. and G. Wolschin (2020), "Baryon stopping as a relativistic Markov process in phase space," Phys. Rev. Res. **2**, 033409.

Itô, K. (1944), "Stochastic integral," Proc. Imp. Acad. **20**, 519–524.

Itô, K. (1946), "On a stochastic integral equation," Proc. Jpn. Acad. **22**, 32–35.

Ji, M., Z. Shen, and Y. Yi (2018), "Convergence to equilibrium in Fokker–Planck equations," J. Dyn. Diff. Equat. **31**, 1591–1615.

Jüttner, F. (1911), "Das Maxwellsche Gesetz der Geschwindigkeitsverteilung in der Relativtheorie," Ann. Phys. **339**, 856–882.

van Kampen, N. G. (1981), "Itô versus Stratonovich," J. Stat. Phys. **24**, 175–187.

van Kampen, N. G. (1986), "Brownian motion on a manifold," J. Stat. Phys. **44**, 1–24.

Kellers, B. and G. Wolschin (2019), "Limiting fragmentation at LHC energies," Prog. Theor. Exp. Phys. **2019**, 053D03.

Kharzeev, D., Y. V. Kovchegov, and K. Tuchin (2004), "Nuclear modification factor in d+Au collisions: onset of suppression in the color glass condensate," Phys. Lett. B **599**, 23–31.

Kharzeev, D. and E. Levin (2001), "Manifestations of high density QCD in the first RHIC data," Phys. Lett. B **523**, 79–87.

Kharzeev, D., E. Levin, and M. Nardi (2005), "Color glass condensate at the LHC: hadron multiplicities in pp, pA and AA collisions," Nucl. Phys. A **747**, 609–629.

Klimontovich, Y. L. (1990), "Ito, Stratonovich and kinetic forms of stochastic equations," Physica A **163**, 515–532.

Kolmogorov, A. N. (1931), "Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung," Math. Ann. **104**, 415–458.

Kramers, H. A. (1940), "Brownian motion in a field of force and the diffusion model of chemical reactions," Physica **7**, 284–304.

Langevin, P. (1908), "Sur la théorie du mouvement brownien," C. R. Acad. Sci. **146**, 530–533.

Lee, K. S. and U. Heinz (1989), “Collective flow model for the pion transverse momentum spectra from relativistic nuclear collisions at CERN,” Z. Phys. C **43**, 425–429.

Lee, K. S., U. Heinz, and E. Schnedermann (1990), “Search for collective transverse flow using particle transverse momentum spectra in relativistic heavy-ion collisions,” Z. Phys. C **48**, 525–541.

Łopuszański, J. (1953), “Relativisierung der Theorie der stochastischen Prozesse,” Acta Phys. Polon. **12**, 87–99.

Markov, A. A. (1906), “Rasprostranenie zakona bolshikh chisel na velichiny zavisyashchie drug ot druga,” Izv. Fiz.-Mat. Obs. Kazan. Univ., 2nd ser. **15**, 135–156.

Martin, A. D., R. G. Roberts, W. J. Stirling, and R. S. Thorne (2002), “NNLO global parton analysis,” Phys. Lett. B **531**, 216–224.

McLerran, L. and R. Venugopalan (1994), “Computing quark and gluon distribution functions for very large nuclei,” Phys. Rev. D **49**, 2233–2241.

Mehtar-Tani, Y. and G. Wolschin (2009a), “Baryon stopping as a new probe of geometric scaling,” Phys. Rev. Lett. **102**, 182301.

Mehtar-Tani, Y. and G. Wolschin (2009b), “Baryon stopping and saturation physics in relativistic collisions,” Phys. Rev. C **80**, 054905.

Michael, C. (1979), “Large transverse momentum and large mass production in hadronic interactions,” Prog. Part. Nucl. Phys. **2**, 1–39.

Michael, C. and L. Vanryckeghem (1977), “Consequences of momentum conservation for particle production at large transverse momentum,” J. Phys. G: Nucl. Phys. **3**, L151–L156.

Mogensen, P. K. and A. N. Riseth (2018), “Optim: A mathematical optimization package for Julia,” J. Open Source Softw. **3**, 615.

Moyal, J. E. (1949), “Stochastic processes and statistical physics,” J. R. Stat. Soc. B **11**, 150–210.

Mueller, A. H. and J. Qiu (1986), “Gluon recombination and shadowing at small values of x,” Nucl. Phys. B **268**, 427–452.

Nelder, J. A. and R. Mead (1965), “A simplex method for function minimization,” Comput. J. **7**, 308–313.

Pavliotis, G. A. (2014), *Stochastic processes and applications, Diffusion processes, the Fokker–Planck and Langevin equations*, 1st ed., Texts in Applied Mathematics (Springer, New York).

Perrin, J.-B. (1909), “Mouvement brownien et réalité moléculaire,” Ann. Chim. Phys., 8th ser. **18**, 5–114.

Planck, M. (1917), “Über einen Satz der statistischen Dynamik und seine Erweiterung in der Quantentheorie,” Sitzungsber. K. Preuß. Akad. Wiss. Berl. **24**, 324–341.

Rackauckas, C. and Q. Nie (2017), "DifferentialEquations.jl – A performant and feature-rich ecosystem for solving differential equations in Julia," J. Open Res. Softw. **5**, 15.

Schnedermann, E., J. Sollfrank, and U. Heinz (1993), "Thermal phenomenology of hadrons from $200A$ GeV S+S collisions," Phys. Rev. C **48**, 2462–2475.

Siemens, P. J. and J. O. Rasmussen (1979), "Evidence for a blast wave from compressed nuclear matter," Phys. Rev. Lett. **42**, 880–883.

Smoluchowski, M. von (1906), "Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen," Ann. Phys. **326**, 756–780.

Smoluchowski, M. von (1916), "Über Brownsche Molekularbewegung unter Einwirkung äußerer Kräfte und deren Zusammenhang mit der verallgemeinerten Diffusionsgleichung," Ann. Phys. **353**, 1103–1112.

Stratonovich, R. L. (1964), "Novaya forma zapisi stokhasticheskikh integralov i uravneniy," Vestn. Mosk. Univ., 1 Mat. Mekh. **1**, 3–12; [SIAM J. Control **4**, 362–371 (1966)].

Sutherland, W. (1905), "LXXV. A dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin," Philos. Mag., 6th ser. **9**, 781–785.

Uhlenbeck, G. E. and L. S. Ornstein (1930), "On the theory of the Brownian motion," Phys. Rev. **36**, 823–841.

Wiener, N. (1923), "Differential-space," J. Math. Phys. **2**, 131–174.

Wolschin, G. (1996), "Diffusion model for relativistic heavy-ion collisions," Z. Phys. A: Hadrons Nucl. **355**, 301–313.

Wolschin, G. (1999), "Relativistic diffusion model," Eur. Phys. J. A **5**, 85–90.

Wolschin, G. (2004), "Diffusion and local deconfinement in relativistic systems," Phys. Rev. C **69**, 024906.

Wolschin, G. (2007), "Diffusion in relativistic systems," Prog. Part. Nucl. Phys. **59**, 374–382.

Wolschin, G. (2013), "Particle production sources at LHC energies," J. Phys. G **40**, 045104.

Wolschin, G. (2016), "Beyond the thermal model in relativistic heavy-ion collisions," Phys. Rev. C **94**, 024911.

Wolschin, G., M. Biyajima, T. Mizoguchi, and N. Suzuki (2006), "Time evolution of relativistic d+Au and Au+Au collisions," Ann. Phys. **518**, 369–378.

An dieser Stelle möchte ich gerne noch all jenen meinen Dank aussprechen, die mich während meiner Promotion begleitet und unterstützt haben. Vor allem Georg Wolschin, der in den vergangenen Jahren stets ein offenes Ohr für meine Fragen hatte, mir immer hilfreich zur Seite stand und dem ich viele spannende Gespräche zu physikalischen und auch nicht-physikalischen Themen verdanke; eine motivierendere Arbeitsumgebung hätte ich mir nicht wünschen können. Bei Andreas Mielke möchte ich mich für seine große Hilfe bei der Suche nach einer Finanzierung bedanken und für seine Bereitschaft, das zweite Gutachten meiner Arbeit zu übernehmen. Ein besonderer Dank geht an das Evangelische Studienwerk, dessen Promotionsstipendium es mir ermöglichte, frei meinen wissenschaftlichen Interessen zu folgen und mit dem ich viele schöne Erinnerungen an die gemeinsame Zeit in Villigst verbinde.