Department of Physics and Astronomy

University of Heidelberg

Master Thesis

in Physics

submitted by

Vincent C. Mader

born in Ulm-Söflingen (Germany)

2024

Monte Carlo Coagulation - Utilizing Stochastic Sampling of Colliding Circumstellar Dust Particle Pairs to Lower the Computational Cost of the Smoluchowski Coagulation Equation's Numerical Integration

This Master Thesis has been carried out by

Vincent C. Mader

at the

Heidelberg Institute for Theoretical Astrophysics

under the supervision of

Prof. Dr. Cornelis P. Dullemond

Erklärung:

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 15. Februar 2024

Vincent Christian Mader

Abstract

Context: The Smoluchowski coagulation equation provides a framework that can be used for modeling the temporal evolution of the mass distribution of dust particles in proto-planetary disks. When both coagulation and fragmentation processes are taken into account, numerical integration requires the evaluation of a double integral over the reaction kernel matrix, which makes this approach quite numerically costly.

Aims: The goal of this work is to examine the feasibility of utilizing stochastic Monte Carlo sampling of the kernel matrix in order to lower the integration's computational cost. More specifically, the idea is to sample the most relevant dust particle collisions, where relevance is determined by a given collision's impact on the evolution of the dust particle mass distribution.

Methods: After the construction of a simplified disk model, we define the reaction rates and kernel matrices for both coagulation and fragmentation processes. The Smoluchowski equation is then numerically integrated using a 4th order implicit Radau integration scheme. An appropriate sampling probability distribution is defined, with which an incomplete kernel matrix is constructed. Parameter studies of the sampling density are made for both the simple case of pure hit-and-stick coagulation, as well as with fragmentation included into the model.

Results: The general behavior of the mass distribution's temporal evolution can be reproduced, even if parts of the kernel are neglected. When both coagulation and fragmentation are included, the numerical solution after a few hundred years resembles the expected equilibrium state. It must be noted though, that due to statistical fluctuations arising from the stochastic nature of the sampling method, the distribution's temporal derivative stays well above zero. This, as well as the effectively lower collisions rates, leads to a loss with regards to accuracy. The stability of the algorithm with regards to mass conservation could be ensured down to 10^{-12}.

Conclusions: Whether it makes sense to use this collision sampling approach likely depends on the specific requirements that are imposed on the computational resources, as well as the desired accuracy, since there is a trade-off between the two. If the dust particles are modeled as possessing more than one attribute (e.g. mass and porosity), the utility value of this method could increase. As the amount of "noise" in the kernel can be expected to grow faster with the dimensionality of the problem than the number of relevant kernel entries, for studies of more sophisticated models including additional particle attributes, it may well be worthwhile to consider this approach.

Zusammenfassung

Kontext: Die Smoluchowski-Koagulationsgleichung stellt ein Modell zur Verfügung, das zur Bestimmung der zeitlichen Entwicklung der Verteilung von Staubteilchenmassen in protoplanetaren Scheiben verwendet werden kann. Wenn sowohl Koagulations- als auch Fragmentierungsprozesse berücksichtigt werden, dann erfordert die numerische Integration dieser Gleichung die Auswertung eines Doppelintegrals über die Reaktionskernel-Matrix, wodurch dieser Ansatz aus numerischer Sicht recht aufwendig wird.

Ziele: Das Ziel dieser Arbeit ist es zu untersuchen, ob sich die Nutzung von stochastischem Monte Carlo Sampling anbietet, um die numerischen Kosten der Integration zu senken. Genauer gesagt ist hier die Idee, nur die relevantesten Staubteilchen-Kollisionen bei der Aufsummierung mit zu berücksichtigen. Relevanz wird hierbei durch den Einfluss einer gegebenen Kollision auf die Entwicklung der Massenverteilung der Staubteilchen bestimmt.

Methoden: Nach der Definition eines vereinfachten Scheibenmodells bestimmen wir die Reaktionsrate sowie die Reaktionskernel-Matrizen für sowohl Koagulations- als auch Fragmentierungsprozesse. Die Integration der Smoluchowski-Gleichung wird mithilfe eines impliziten Radau-Integrationsverfahrens vierter Ordnung durchgeführt. Mit einer geeigneten Wahrscheinlichkeitsverteilung wird dann das stochastische Sampling der Kernel-Matrix vollzogen. Parameterstudien der Sampling-Dichte werden sowohl für den einfacheren Fall der reinen Koagulation, als auch unter Einbezug von Fragmentierung gemacht.

Ergebnisse: Das allgemeine Verhalten der zeitlichen Entwicklung der Massenverteilung kann reproduziert werden, auch wenn weite Teile des Kernels vernachlässigt werden. Wenn sowohl Koagulation als auch Fragmentierung einbezogen sind, nähert sich die numerische Lösung nach einigen hundert Jahren dem erwarteten Gleichgewichtszustand an. Es muss jedoch angemerkt werden, dass aufgrund statistischer Schwankungen, die aus der Sampling-Methode resultieren, die zeitliche Ableitung der Verteilung weit über Null bleibt. Dies, sowie die effektiv niedrigeren Kollisionsraten, führt zu einem Verlust an Genauigkeit. Die Stabilität des Algorithmus in Bezug auf Massenerhaltung konnte bis auf 10^{-12} sichergestellt werden.

Schlüsse: Ob es sinnvoll ist, den in dieser Thesis erprobten Ansatz des Kernel Samplings zu verwenden, hängt vermutlich von den spezifischen Anforderungen an die Rechenressourcen sowie der gewünschten Genauigkeit ab, da ein Kompromiss zwischen den beiden gefunden werden muss. Wenn bei der Modellierung der Staubteilchen mehr als nur ein Attribut berücksichtigt wird (e.g. Masse und Porosität), dann könnte die Anwendung dieser Methode an Zweckmäßigkeit gewinnen. Es kann davon ausgegangen werden, dass der Anteil an vernachlässigbaren Einträgen in der Kernel-Matrix mit der Dimensionalität des Problems schneller anwächst als die Zahl der relevanten Einträge. Für detailliertere Studien, die zusätzliche Partikeleigenschaften einschließen, könnte es sich deshalb durchaus als sinnvoll erweisen, diesen Ansatz zu verfolgen.

Acknowledgements

First and foremost, I wish to express my greatest gratitude to Prof. Dr. Cornelis (Kees) Dullemond, for acting as my supervisor while I was writing this thesis, and for providing me with invaluable guidance, advice and feedback.

Second, I want to thank Prof. Dr. Carsten Dominik, who kindly took on the role of second examiner.

Also, since this thesis probably marks the end of my time as a student in Heidelberg, it needs to be said that I am extremely grateful for the last few years that I could spend in this beautiful city, and for all the learning opportunities granted to me by *Ruperto Carola* University of Heidelberg, the Max Planck Institute for Astronomy, and especially the planet formation group and its members at the Heidelberg Institute for Theoretical Astrophysics.

On a more personal note, I wish to thank my mother, father, and grandmother, who kindled my curiosity and interest into the natural sciences from an early age on. Also, in that context I want to mention Kati, Jürgen, and Alexander, as well as my friends and family at DPSG Ulm-Söflingen.

Without the enduring support of the mentioned individuals over many years, this thesis might have never come to fruition. As such, it is dedicated to them.

Contents

1 Introduction **1**
- 1.1 Structure of this Thesis . . . 2
- 1.2 Short Introduction to Planet Formation . . . 2
  - 1.2.1 The Interstellar Medium . . . 3
  - 1.2.2 The Nebular Hypothesis . . . 4
  - 1.2.3 The Road from Dust to Planet . . . 5
- 1.3 Preliminary Notes on Axis Discretization . . . 6
  - 1.3.1 Axis Discretization with Linear Scaling . . . 7
  - 1.3.2 Axis Discretization with Logarithmic Scaling . . . 7
- 1.4 Preliminary Notes on the Dust Model . . . 8
  - 1.4.1 The Individual Dust Particles . . . 8
  - 1.4.2 The Dust Particle Mass Distribution Function . . . 9
  - 1.4.3 Initialization of the Dust Particle Mass Distribution . . . 10

2 The Proto-Planetary Disk Model **11**
- 2.1 Preliminary Definitions and Assumptions . . . 11
  - 2.1.1 Radial Profile of the Midplane Temperature . . . 13
  - 2.1.2 Radial Profile of the Midplane Speed of Sound . . . 14
- 2.2 Gas and Dust Density . . . 15
  - 2.2.1 Radial Profile of the Gas Surface Density . . . 15
  - 2.2.2 Radial Profile of the Gas Volume Density . . . 16
  - 2.2.3 Radial Profile of the Dust Volume Density . . . 17
- 2.3 Gas Particle Kinematics . . . 19
  - 2.3.1 Gas Viscosity . . . 19
  - 2.3.2 Radial Gas Velocity . . . 19
- 2.4 Dust Particle Kinematics . . . 20
  - 2.4.1 Radial Drift . . . 20
  - 2.4.2 Azimuthal Drift . . . 21
  - 2.4.3 Turbulent Motion . . . 21
  - 2.4.4 Brownian Motion . . . 22
  - 2.4.5 Total Relative Velocity . . . 22
- 2.5 Dust Particle Collisions . . . 24
  - 2.5.1 Collision Cross Section . . . 24
  - 2.5.2 Collision Rate . . . 25
  - 2.5.3 Collision Outcomes . . . 25

3 The Dust Coagulation Model **29**
- 3.1 The Smoluchowski Coagulation Equation . . . 29
  - 3.1.1 Discretized Formulation of the Smoluchowski Equation . . . 30
- 3.2 The Kernel . . . 31
  - 3.2.1 Decomposition into Collision Rate and Collision Details . . . 31

3.2.2 Decomposition into Coagulation and Fragmentation Kernel . . . 31
3.2.3 Decomposition into Gain and Loss Term . . . 32
3.2.4 Definition of the Kernel Mass Error . . . 32
3.3 Definition of the Coagulation Kernel . . . 33
3.3.1 Discretization of the Coagulation Kernel . . . 34
3.3.2 The Kovetz-Olund Algorithm . . . 35
3.3.3 Near-Zero-Cancellation Handling . . . 36
3.4 Definition of the Fragmentation Kernel . . . 37
3.4.1 Modeling the Resulting Mass Distribution . . . 37
3.4.2 The MRN Distribution . . . 38
3.5 Visualization of Kernel Matrices . . . 39
3.5.1 Coagulation Kernel on a Linear Mass Axis . . . 39
3.5.2 Coagulation Kernel on a Logarithmic Mass Axis . . . 41
3.5.3 Fragmentation Kernel on a Logarithmic Mass Axis . . . 41
3.5.4 Total Kernel on a Logarithmic Mass Axis . . . 42
3.5.5 Kernel Mass Error . . . 42
3.6 Analytical Solutions . . . 50
3.6.1 Constant Kernel . . . 50
3.6.2 Linear Kernel . . . 51
3.6.3 Quadratic Kernel . . . 51
3.7 Numerical Integration . . . 52
3.7.1 Definition of the Integration Mass Error . . . 52
3.7.2 Explicit Euler Integration . . . 52
3.7.3 Implicit Radau Integration . . . 54
3.8 Visualization of Integration Results . . . 55
3.8.1 Pure Coagulation on Linear Mass Axis . . . 56
3.8.2 Pure Coagulation on Logarithmic Mass Axis . . . 57
3.8.3 Pure Fragmentation on Logarithmic Mass Axis . . . 59
3.8.4 Full Model . . . 61

4 Stochastic Monte Carlo Sampling of the Kernel Matrix 63
4.1 Introduction . . . 63
4.1.1 Motivation . . . 63
4.1.2 Previous Works . . . 65
4.2 Kernel Sampling . . . 65
4.2.1 Basic Idea . . . 65
4.2.2 Definition of the Sampling Probability . . . 67
4.2.3 Sampling of the Pure Coagulation Kernel . . . 69
4.2.4 Sampling of the Total Kernel . . . 73
4.2.5 Accuracy and Stability . . . 77

5 Summary & Discussion 81

Bibliography 85

List of Figures

1.1 Examples of star-forming regions: (a) The Orion Nebula (b) The Eagle Nebula (Zoom in to see the famous *Pillars of Creation* at the center of the image). 3

1.2 Infrared/radio observations performed at *Atacama Large Millimeter Array* (ALMA) of the proto-planetary disks around the stars (a) TW Hydrae and (b) HL Tauri. . . . 4

1.3 Schematic illustration of the discretized mass axis. After having defined a minimum value $m_{\min}$ and a maximum value $m_{\max}$, the interval between these two values is divided evenly into $\mathcal{N}_m$ bins. To do this, we first define the mass values at the bin boundaries and, from that, the mass values at the bin centers. This can be done using either a linear or a logarithmic scaling (for details, see section section 1.3.1 and section 1.3.2, respectively). The definition of the discretized axes for both time t and distance r from the star is done completely analogously. 6

1.4 Schematic drawing of different models for types of dust particle structure. 8

2.1 Radial profile of the midplane gas temperature $T_{\text{mid}}(r)$. 13

2.2 Radial profile of the midplane sound speed $c_s(r)$. 14

2.3 Illustration of Trigonometric relations used for expressing the z-component of the acceleration due to the gravitational influence of the star at the center of the disk. . 16

2.4 *Top left:* Radial profile of the gas surface density $\rho_g(r)$. *Top right:* Radial profile of the gas volume density at the disk midplane $\rho_g^{\text{mid}}(r)$. *Bottom left:* Radial profile of the gas particle number density at the disk midplane $N_g^{\text{mid}}(r)$. *Bottom right:* z-profile of the gas volume density at $r = r_0 = 10$ au. 18

2.5 Total relative dust particle velocity $\Delta v_{ij}^{\text{tot}}$. 22

2.6 Relative dust particle velocities due to (a) Radial drift $\Delta v_{ij}^{\text{RD}}$, (b) Azimuthal drift $\Delta v_{ij}^{\text{AZ}}$, (c) Turbulent motion $\Delta v_{ij}^{\text{TU}}$, and (d) Brownian motion $\Delta v_{ij}^{\text{BR}}$. 23

2.7 Visualization of the dependence of the (a) Dust particle collision cross section $\sigma_{ij}^{\text{coll}}$ and (b) Dust particle collision rate coefficient R_{ij}^{coll} on the mass values of two colliding particles. 24

2.8 Illustration of dust particle collision outcomes (a) Coagulation: Two particles collide and merge into a single new one that carries the combined mass of the initial particles. (b) Fragmentation: The colliding particles are destroyed during the collision. The total mass is subsequently distributed onto a wide range of differently-sized, newly-created particles. (c) Bouncing: Two particles collide, but no mass is exchanged between them. 26

2.9 Dust particle fragmentation probability P_{ij}^{frag}. 27

3.1 Matrix layers of the kernel for pure hit-and-stick coagulation on a linear mass axis with lower mass grid boundary $m_{\min} = 10^{-6}$ kg, upper mass grid boundary $m_{\min} = 10^{-2}$ kg, mass grid resolution $\mathcal{N}_m = 100$, and collision rate coefficent $R_{ij}^{\text{coll}} := 1\ \text{m}^3\ \text{s}^{-1}$. 40

3.2 Matrix layers of the positive kernel component for pure hit-and-stick coagulation on a logarithmic mass axis, with lower mass grid boundary $m_{\min} = 10^{-15}$ kg, upper mass grid boundary $m_{\min} = 10^{-8}$ kg, and mass grid resolution $\mathcal{N}_m = 100$. 43

3.3 Matrix layers of the negative kernel component for pure hit-and-stick coagulation on a logarithmic mass axis, with lower mass grid boundary $m_{\min} = 10^{-15}$ kg, upper mass grid boundary $m_{\min} = 10^{-8}$ kg, and mass grid resolution $\mathcal{N}_m = 100$. 44

3.4 Matrix layers of the positive kernel component for pure fragmentation on a logarithmic mass axis, with lower mass grid boundary $m_{\min} = 10^{-15}$ kg, upper mass grid boundary $m_{\min} = 10^{-8}$ kg, and mass grid resolution $\mathcal{N}_m = 100$. 45

3.5 Matrix layers of the negative kernel component for pure fragmentation on a logarithmic mass axis, with lower mass grid boundary $m_{\min} = 10^{-15}$ kg, upper mass grid boundary $m_{\min} = 10^{-8}$ kg, and mass grid resolution $\mathcal{N}_m = 100$. 46

3.6 Matrix layers of the positive kernel component for both coagulation and fragmentation on a logarithmic mass axis, with lower mass grid boundary $m_{\min} = 10^{-15}$ kg, upper mass grid boundary $m_{\min} = 10^{-8}$ kg, and mass grid resolution $\mathcal{N}_m = 100$. 47

3.7 Matrix layers of the negative kernel component for both coagulation and fragmentation on a logarithmic mass axis, with lower mass grid boundary $m_{\min} = 10^{-15}$ kg, upper mass grid boundary $m_{\min} = 10^{-8}$ kg, and mass grid resolution $\mathcal{N}_m = 100$. 48

3.8 Influence of near-zero cancellation on the mass error of the coagulation kernel. 49

3.9 Dimensionless mass error ΔX_{ij} for (a) Fragmentation kernel, and (b) Total kernel. . 49

3.10 Temporal evolution of the dust particle mass distribution under the influence of only pure hit-and-stick coagulation. Here, the mass axis was discretized making use of a linear scaling. 56

3.11 Relative mass error (stability error) $\Delta_{\text{stab}}(t) = \frac{\rho_d(t)-\rho_d(t=0)}{\rho_d(t=0)}$ of the numerical integration as a function of time. Here, only the process of pure coagulation is included in the model. 57

3.12 Temporal evolution of the dust particle mass distribution under the influence of only pure coagulation. 58

3.13 Root-mean-squared (RMS) temporal derivative of the dust particle mass distribution as a function of time. Here, only the process of pure coagulation is included into the model. 58

3.14 Relative mass error (stability error) $\Delta_{\text{stab}}(t) = \frac{\rho_d(t)-\rho_d(t=0)}{\rho_d(t=0)}$ of the numerical integration as a function of time. Here, only the process of pure fragmentation is included in the model. 59

3.15 Temporal evolution of the dust particle mass distribution under the influence of only pure fragmentation. Here, the initial mass distribution is centered not on the first bin like in the other runs, but on one corresponding a higher mass value. 60

3.16 Root-mean-squared (RMS) temporal derivative of the dust particle mass distribution as a function of time. Here, only the process of fragmentation is included into the model. 60

3.17 Relative mass error (stability error) $\Delta_{\text{stab}}(t) = \frac{\rho_d(t)-\rho_d(t=0)}{\rho_d(t=0)}$ of the numerical integration as a function of time. Here, both coagulation and fragmentation processes are included in the model. 61

3.18 Temporal evolution of the dust particle mass distribution under the influence of both coagulation and fragmentation processes. 62

3.19 Root-mean-squared (RMS) temporal derivative of the dust particle mass distribution as a function of time. Here, both the process of coagulation and fragmentation is included into the model. 62

4.1 Illustration of the basic idea for kernel sampling: Only a fraction of the total number of possible particle collision is taken into account when defining the kernel matrix. This number is determined by the sampling density ρ_{sample}, which is set to 1/2 here. The kernel is plotted for $k = 40$ on a mass grid with resolution $\mathcal{N}_m = 100$. Note that here, we are using a homogenous probability distribution for the collision sampling. Since not all collisions are relevant for the temporal derivative of the dust particle mass distribution to the same extent, we will define a more sensible non-homogenous probability distribution in the next section. 66

4.2 Temporal evolution of the dust particle mass distribution under the influence of (only) pure hit-and-stick coagulation processes for various different values of the sampling density ρ_{sample}. 70

4.3 Illustration of the collision sampling probability distribution P_{ij}^{sample} for $\rho_{\text{sample}} = 0.5$. Here, only the process of pure stick-and-hit coagulation is included into the model. . 71

4.4 Illustration of the amount of times N_{ij}^{sample} that each collision (i, j) is sampled, for $\rho_{\text{sample}} = 0.5$. Here, only the process of pure stick-and-hit coagulation is included into the model. 72

4.5 Temporal evolution of the dust particle mass distribution under the influence of both coagulation and fragmentation processes for various different values of the sampling density ρ_{sample}. 74

4.6 Illustration of the collision sampling probability distribution P_{ij}^{sample} for $\rho_{\text{sample}} = 0.5$. Here, both the process of coagulation as well as fragmentation are included into the model. 75

4.7 Illustration of the amount of times N_{ij}^{sample} that each collision (i, j) is sampled, for $\rho_{\text{sample}} = 0.5$. Here, both the process of coagulation as well as fragmentation are included into the model. 76

4.8 Accuracy error Δ_{acc} and stability error Δ_{stab} for the numerical integration of the Smoluchowski coagulation equation for various sampling densities ρ_{sample}. Stability is assured down to machine precision, accuracy on the other hand is negatively impacted by statistical fluctuations arising from the stochastic nature of the Monte Carlo method. For the stability error, the dashed line signifies a negative value. 78

4.9 Temporal evolution of the root-mean-square (RMS) of the temporal derivative of the mass distribution. Here, both the process of coagulation as well as fragmentation are included into the model. Note that when sampling is enabled, the temporal derivative does not approach zero as quickly as when it is disabled. This is most likely due to the stochastic nature of the Monte Carlo sampling method, leading to statistical fluctuations that prevent the mass distribution from ultimately settling into a true equilibrium state. 79

4.10 Temporal evolution of the root-mean-square (RMS) of the temporal derivative of the mass distribution. Here, only the process of pure hit-and-stick coagulation is included into the model. Interestingly, here the temporal derivatives that were arrived at using a sampled kernel initially approach zero much faster than when using the complete kernel. At some point, also here the temporal derivative of the "complete" solution overtakes the sampled soltions on its way towards zero. 79

4.11 Percentage of non-zero entries in the sampled kernel matrix. Here, both coagulation and fragmentation are included in the model. If $\rho_{\text{sample}} = 1$, i.e. if all collisions are taken into account when defining the kernel, this ratio is approximately 30%. Since a lot of these collisions are neglected when a sampled kernel is used, the number of non-zero kernel entries decreases with lower values of the sampling density. This is especially true in the beginning, where fragmentation processes do not yet happen much. As time progresses, more and more entries of the kernel become relevant. As such, this ratio increases and, in the end, converges towards the value of the sampling density. 80

4.12 Percentage non-zero entries in the sampled kernel matrix. Here, only the process of pure stick-and-hit coagulation is included into the model. The number of non-zero entries increases as time progresses, until there is a steep drop-off when basically the entire mass has been transferred into the highest-mass bin. Overall, the number of non-zero kernel entries is much lower here than in the full model, reflecting the fact that due to the Dirac δ-function in equation (3.21), the definition of the pure coagulation kernel yields a much sparser matrix than one would have if fragmentation processes are included, with only approximately 3% of the kernel being non-zero for $\rho_{\text{sample}} = 1$. 80

Chapter 1

Introduction

As of this writing, the mechanisms involved in planetary formation are still far from being fully understood. Accurate modeling of the underlying processes requires an understanding of a wide variety of different sub-fields of the natural sciences, including (but not limited to) gravitational N-body dynamics, thermodynamics and radiative transfer, magneto-hydro-dynamics, cosmo-chemistry, and coagulation physics.

Bringing all of these different aspects together in a sensible fashion is not easy, thus making this quite a complex field of study. We are not helped by the fact that it is basically impossible to directly run experiments on planet formation in the laboratory. Due to the involved scales with respect to e.g. time, distance, mass, and energy, studies in the laboratory can only be made when focusing on a very small subset of the questions that we would like to have answered. As such, apart from telescope observations, we often have to rely on numerical techniques to gain insight into the respective processes, running simulations based on highly simplified theoretical models.

Various frameworks exist for constructing such models for different aspects of the planetary formation process. One such framework, which is of especially great relevance to the studies done in the context of this thesis, is given by the so-called *Smoluchowski coagulation equation*.

This equation allows the definition of a model for the temporal evolution of a given particle group's distribution of masses under the influence of particle collisions and subsequent merging and/or fragmentation events. As such, the framework provided by this equation can be used to study the agglomeration of particles in young proto-planetary disks. More precisely, we will use this framework to study the evolution of the mass distribution of *circumstellar dust particles*, which are widely believed to play an imporant role in the early phases of planetary core formation.

Previous studies of dust particle agglomeration that made use of the framework provided by the Smoluchowski coagulation equation have often based the definition of a dust model on the simplifying assumption that the individual dust particles can be characterized entirely by the value of their mass alone (see e.g. [DD04], [BDB10]). This is due to the fact that the numerical integration of the Smoluchowski equation from one time-step to the next involves the evaluation of a summation over the multi-dimensional *kernel matrix*. The inclusion of a second attribute like, for instance *particle porosity*, greatly increases the dimensionality of the kernel, which leads to a significant increase in the computational cost required for the integration.

In this thesis, we explore the possibility of lowering this computational cost via the utilization of stochastic Monte Carlo sampling of the kernel matrix. The central idea here is to include only the most relevant sum terms, where relevance is determined by both the rate of a given collision, as well as by the effect that it will have on the particle mass distribution function.

1.1 Structure of this Thesis

The thesis will be structured as follows: To get started, in section 1.2 we will give a short overview of a number of widely accepted ideas which are related to the origin and formation of proto-planetary disks and the processes related to planet formation within them.

Following that, in section 1.3 and section 1.4, we will lay out a few preliminary notes which are related to the discretization of axes, as well as to the method that will be used for modeling the properties of both the individual dust particles, as well as the overall distribution of dust particles masses inside the proto-planetary disk.

Then, we will continue with the definition of a simplified disk model in chapter 2. This model should include the most relevant properties of both the disk itself, as well as its gas and dust contents. In the subsequent sections, it will be used to model the dust particle kinematics, and from that define the dust particle collision rates, which will be an essential part of our dust coagulation model.

In chapter 3 we will then utilize the Smoluchowski coagulation equation to build a simple model of the dust particle mass distribution's temporal evolution under the influence of particle collisions and the following associative/dissociative events. The definition of a dust coagulation model will involve the construction of a kernel matrix for both coagulation and fragmentation processes, as well as the implementation of a method for numerically integrating the Smoluchowski equation using a 4th order implicit Radau integration scheme.

Having constructed the complete kernel matrix, in chapter 4 we will finally focus on the implementation of stochastic kernel sampling using a Monte Carlo method. Here, the task is to identify the most relevant collisions at a given point in time, define a sensible probability distribution from that, and conduct a random sampling of all possible collisions. Including only the chosen collisions, a new (incomplete) kernel matrix is then defined in each time-step, with which the numerical integration is subsequently carried out.

1.2 Short Introduction to Planet Formation

Directly after the Big Bang, the conditions in the young Universe were not favorable for the formation of large-scale structures. Extremely high temperatures prevented the agglomeration of smaller particle constituents into more complex structures.

Current consensus has it that it took approximately 380.000 years for the initial hot quark-gluon plasma to be cooled down sufficiently by the expansion of the Universe to enable the creation of the first hadronic matter via a combination of multiple quarks into one particle. For us, the most relevant of these quark combinations were of course the proton and the neutron, as they provided the building blocks for the nuclei of the first atoms. Subsequently, the capture of electrons by these atomic cores lead to the creation of neutral atoms. As they make up the most basic elements in the periodic table due to their simple atomic structure, the elemental composition of the early Universe was dominated by the contribution of hydrogen present in the form of molecular gas and, to a lesser extent, mono-atomic helium.

Apart from trace amounts of metals that were formed mostly via nuclear fusion inside stellar cores, this statement about the chemical composition of the Universe is still valid today. Though much less abundant, the presence of these metals is assumed to be essential for the formation of planets, as they provide the material for a central rocky core. Initially, they are present in the form of dust particles, which are composed of various materials like silicates, carbon compounds, or ices.

1.2.1 The Interstellar Medium

The term *interstellar medium* (ISM) is used to denote the matter and radiation that can be found in the expanse of space between the stellar systems inside a galaxy. As we will discuss further in the following sections, the ISM is widely considered to serve as the birthplace of proto-planetary disks (PPDs), and subsequently, stars and planets.

At first glance, it might not be easy to imagine how such giant objects like planets and even stars can form in the ISM. After all, one might naively say that it consists of mostly nothing. From the point of view of a human living on Earth, the interstellar regions of space appear to be a near-perfect vacuum. Measurements of the the molcular particle number density in the ISM have been made in previous studies by e.g. [Bur+13], where they showed that the number of particles per unit volume can range from approximately $1 \times 10^{16}\,\mathrm{m}^{-3}$ all the way down to $1 \times 10^{3}\,\mathrm{m}^{-3}$.

The exact value of the particle number density depends of course heavily on the precise location. Nonetheless, a comparison of the mere orders of magnitude involved here demonstrates the huge difference in particle abundances in the ISM when compared to Earth, where the number[1] of gas molecules inside the atmosphere is approximately $1 \times 10^{25}\,\mathrm{m}^{-3}$.

Under the influence of gravity, the material in the ISM can gather into what is appropriately called *giant interstellar clouds*. Examples of such clouds are given by e.g. the Orion Nebula, which can be seen with the naked eye as the "sword" in the Orion constellation, or the Eagle Nebula, which hosts the famous *Pillars of Creation*. Measurements of these two can be seen visualized in figure 1.1.

The fact that giant astronomical objects like PPDs, stars, and planets can form from the low volume densities of the material found in the ISM, may be understood more easily if one brings to mind the absolutely enormous scales of both space and time that are involved here.



(a) Orion Nebula (Messier 42) [Adv06]



(b) Eagle Nebula (Messier 16) [Eur09]

Figure 1.1: Examples of star-forming regions: (a) The Orion Nebula (b) The Eagle Nebula (Zoom in to see the famous *Pillars of Creation* at the center of the image).

[1] Here we assume the validity of the ideal gas law, a sea-level pressure of $p = 1\,\mathrm{bar}$, and a temperature of $T = 25\,^{\circ}\mathrm{C}$.

1.2.2 The Nebular Hypothesis

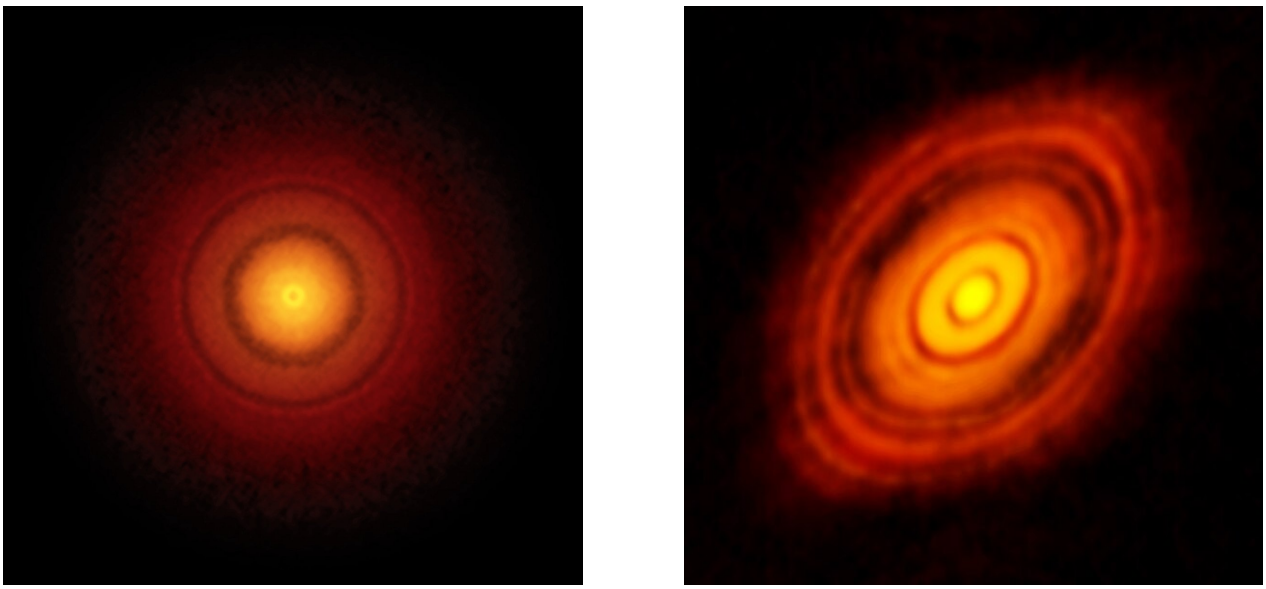
But how *do* stars, and the planetary systems surrounding them, form in these regions? One explanation is given by the so-called *nebular hypothesis*, which provides a widely accepted model for explaining the formation and evolution of not only our own Solar System, but the planetary systems around other stars as well. It was postulated independently by Immanuel Kant and Pierre-Simon Laplace in 1755 and 1796, respectively.

The hypothesis is built upon the idea that the formation of proto-planetary disks, and thus stellar systems, can be explained via the gravitational collapse of interstellar clouds of gas and dust.

Such a collapse may be induced by external influences, like e.g. a shock-wave from a nearby supernova explosion, or just happen due to an irregular distribution of matter in the cloud, leading to a graviational instability.

When such a cloud falls in on itself, one might naively think that this leads to the entire matter collapsing into a single point. But since the cloud possesses a total angular momentum which is most likely $\neq 0$, this does not happen. Instead, the cloud flattens into a disk-like structure, where gravitational influences are balanced out along the radial axis by centrifugal forces. Such a disk will also not continue to flatten into an infinitesimally thin structure. This is due to the non-zero temperatures and the internal gas pressure support along the z-axis, i.e. the axis orthogonal to the disk.

At the center of the disk, densities and temperatures may increase to a point that nuclear fusion becomes possible, and a star can form. The result is a proto-planetary disk, examples of which can be detected all around our own location in the Milky Way. Examples of observations of such objects can be seen visualized in figure 1.2.



(a) Disk around TW Hydrae [S A16] (b) Disk around HL Tauri [ALM14]

Figure 1.2: Infrared/radio observations performed at *Atacama Large Millimeter Array* (ALMA) of the proto-planetary disks around the stars (a) TW Hydrae and (b) HL Tauri.

1.2.3 The Road from Dust to Planet

In the early stages of planet formation, gravity most likely does not play a dominant role in particle agglomeration processes. Instead, dust grains composed of the afore-mentioned materials collide and then stick together mainly via microphysical processes like van der Waals or Coulomb interactions. Over time, larger and larger structures form from initially tiny particles, such that the range of dust particles sizes can range from microns up to millimeter and centimeter sizes, more similar to "sand and pebbles" than to grains. [Nat+06]

This process of dust coagulation is not still entirely understood today. For example, there remains the unanswered question of how exactly it is possible for these aggregates to continue growing without fragmenting into smaller pieces again when collisions with other particles occur. Assuming this so-called fragmentation barrier and other barriers can be overcome though, ever larger structures emerge as particles collide and stick to each other. From the initial dust particles, fractal-shaped "dust bunnies" form, which themselves over long time-scales may evolve into pebbles.

The presence of these pebbles may at some point allow the formation of large enough structures such that one could reasonably speak of *planetesimals*. The gravitational influence of these bodies can then become strong enough to assist in the sticking of smaller grains to the planetesimals, as well as the accretion of gas.

Most modern planet formation models assume the initial buildup of a rocky core, which subsequently may develop into a terrestrial planet, or into a giant planet consisting of a rocky core surrounded by a large gaseous hull. In our own Solar System, the four terrestrial planets Mercury, Venus, Earth, and Mars are situated closest to the Sun, whereas the four gas giants Jupiter, Saturn, Uranus, and Neptune orbit the Sun at much greater distances, in the outer regions of the Solar System.

Most likely, this is no coincidence. As was already mentioned, the greater part of matter in the universe (and therefore also of most matter in newly formed disks), consists of light elements like hydrogen and helium. Because of radiation pressure induced by solar wind, these light elements are pushed outwards away from the disk center, and accumulate in its outer regions. Therefore, an accreting planetoid in the outskirts of the disk will have much more gaseous material available for accretion in its vicinity than one positioned close to the center.

In this thesis, we will focus on the very early stages of the dust coagulation process, and use the Smoluchowski coagulation equation to model the distribution of dust particle masses as well as its temporal evolution.

1.3 Preliminary Notes on Axis Discretization

In order to do this, we will have to deal with several physical quantities that carry values ranging across many orders of magnitude. Most notably, these quantities are time, distance, and mass.

As mentioned earlier, the Smoluchowski coagulation equation serves as the foundation for the coagulation model that we will build. Since this integro-differential equation has analytical solutions in only a limited set of cases, it will be necessary to make use of numerical integration techniques to obtain an approximate solution for the evolution of the dust particle mass distribution in the considered disk region. Before numerical integration, the axes for the above-mentioned physical quantities will have to be discretized in an appropriate manner.

As an example, we will here demonstrate how one might do this for the mass axis m. The construction of the discrete axes for both space (i.e. radial distance r from the central star) and time t can then be done entirely analogously.

The continuous range m of particle masses present in the disk is to be partitioned into a set of $\mathcal{N}_m \in \mathbb{N}$ intervals, which we will refer to as "bins" from this point onwards. Each of these bins can be uniquely characterized by an index $i \in [1, \mathcal{N}_m] \cap \mathbb{N}$, and assigned a corresponding mass value m_i^{c}, which is situated at the *center of the bin* (ergo the superfix "c").

To derive an expression that can be used to calculate the mass values at these bin centers, let us first define the mass values at the *boundaries of the bins*. To do this, consider a collection of $\mathcal{N}_m + 1$ appropriately spaced grid points m_i^{b}. The first and last of these values, i.e. the lower and upper boundaries of the discrete mass grid, shall be labeled $m_{\min}$ and $m_{\max}$, respectively.

The spacing of the grid points of course depends heavily on the utilized scaling. In section 1.3.1 and section 1.3.2, we will go into more detail on how exactly the values for m_i^{b} and m_i^{c} are to be defined when making use of a linear and logarithmic scaling, respectively. If not specified otherwise, throughout this thesis we will always use the logarithmic scaling for the discretization of the axes for mass, distance, and time. Also, we will mostly be considering the grid values at the bin centers, and for thus for reasons of readibility refrain from including the superfix into the equations.

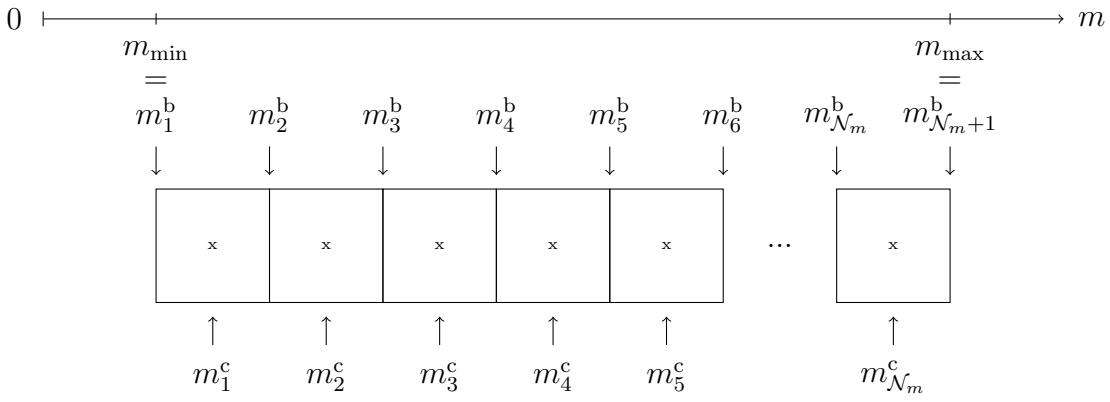


Figure 1.3: Schematic illustration of the discretized mass axis. After having defined a minimum value $m_{\min}$ and a maximum value $m_{\max}$, the interval between these two values is divided evenly into $\mathcal{N}_m$ bins. To do this, we first define the mass values at the bin boundaries and, from that, the mass values at the bin centers. This can be done using either a linear or a logarithmic scaling (for details, see section section 1.3.1 and section 1.3.2, respectively). The definition of the discretized axes for both time t and distance r from the star is done completely analogously.

1.3.1 Axis Discretization with Linear Scaling

For the sake of simplicity, let us assume a linear scaling at the moment. Later on, we will make use of a logarithmic scaling instead, which will help us assure a more appropriate representation of all values along the wide range of dust particle masses that are present in the proto-planetary disk.

For a given bin, which can be characterized by an index i, the mass value corresponding to this bin's lower boundary shall be labeled m_i^{b}. When using a linear scaling, it can be expressed as

$$m_i^{\mathrm{b}} = m_{\min} + (m_{\max} - m_{\min}) \cdot \frac{i}{\mathcal{N}_m} \tag{1.1}$$

Having derived the mass values at the boundaries of each bin, it is now quite easy to calculate the corresponding values at the bin center. To do this, all we need to do is take the arithmetic mean of the two boundary values:

$$m_i^{\mathrm{c}} = \frac{m_i^{\mathrm{b}} + m_{i+1}^{\mathrm{b}}}{2} \tag{1.2}$$

Thus, after having defined only the three numbers $\mathcal{N}_m$, $m_{\min}$, and $m_{\max}$, it is possible to interpolate the values of all mass grid points positioned on both the boundaries as well as the centers of the bins.

The inverse transformation from mass to index can be derived easily by rearranging the above relation, which leads to the following expression:

$$i(m) = \mathcal{N}_m \cdot \frac{m - m_{\min}}{m_{\max} - m_{\min}} \tag{1.3}$$

1.3.2 Axis Discretization with Logarithmic Scaling

The procedure for constructing the discretized axis using a logarithmic scaling works quite analogously to what we just saw in the linear case. The main difference is that we will switch out addition/subtraction by multiplication/division, and multiplication/division by exponentiation.

As before, let $\mathcal{N}_m$ label the total number of bins, each of which can be uniquely identified by an index $i \in [1, \mathcal{N}_m]$. Once again, we will first define an expression for the grid points sitting directly on the lower boundary of each bin. When making use of a logarithmic scaling, these values are given by

$$m_i^{\mathrm{b}} = m_{\min} \cdot \left(\frac{m_{\max}}{m_{\min}} \right)^{i/\mathcal{N}_m} \tag{1.4}$$

To arrive at the mass values at the bin centers, we again take the mean. Contrary to the linear case, here we are not using the arithmetic mean though, but the geometric mean instead:

$$m_i^{\mathrm{c}} = \sqrt{m_i^b \cdot m_{i+1}^b} \tag{1.5}$$

As in the linear case, the inverse transformation can easily be arrived at by rearranging for the bin index i, which leads to the following expression:

$$i(m) = \mathcal{N}_m \cdot \frac{\log(m) - \log(m_{\min})}{\log(m_{\max}) - \log(m_{\min})} \tag{1.6}$$

In contrast to the linear grid, where the "bin width", i.e. the additive offset between bins

$$\Delta m := m_{i+1}^b - m_i^b \tag{1.7}$$

is the same for all bins, in the logarithmic grid this is not the case. Instead, what stays constant here is the *relative* mass increase from one bin to the next:[2]

$$\frac{m_i^{\mathrm{c}}}{m_i^{\mathrm{c}}} = \frac{m_i^{\mathrm{b}}}{m_i^{\mathrm{b}}} = \text{const.} \ \forall i \in [1, \mathcal{N}_m] \tag{1.8}$$

[2] This is true only up to the machine precision of the utilized computer setup.

1.4 Preliminary Notes on the Dust Model

Let us now turn our attention to the question of how one might model the individual dust particles present in the proto-planetary disk. Here, it will be necessary to make simplifying assumptions, that obviously cannot reflect physical reality, but must suffice as an approximate model.

1.4.1 The Individual Dust Particles

In the context of this thesis, we will model the individual dust particles as perfectly spherical bodies, all possessing a shared solid density ρ_s. This allows us to group the particles into different classes, each uniquely identified by the particles' mass value m.

Also, this shape allows us state the following relationship between particle mass m and radius a:

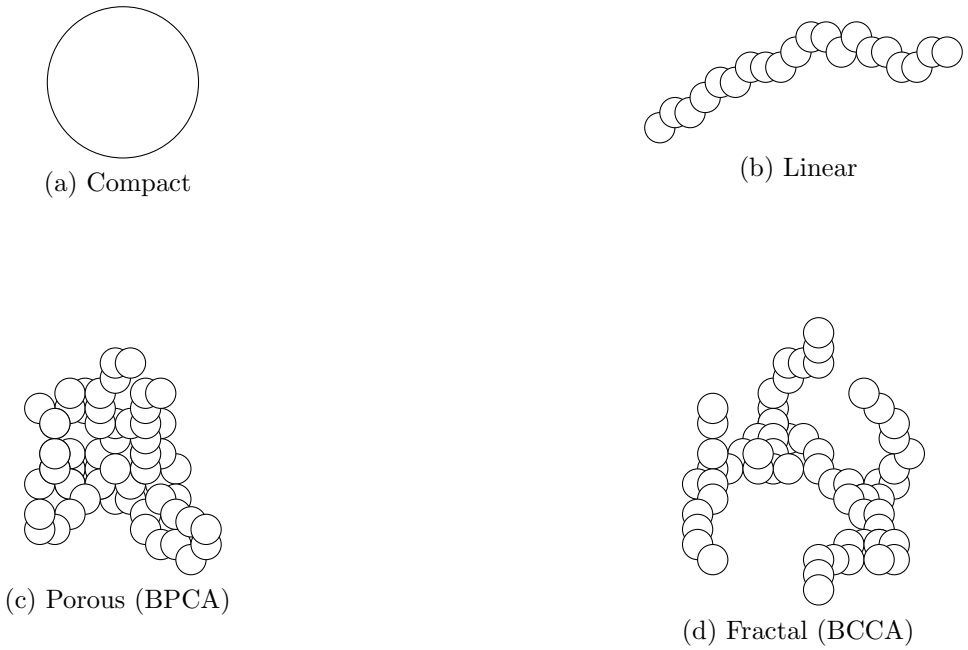
$$m = \rho_s \cdot \frac{4\pi}{3} a^3 \qquad \text{and} \qquad a = \left(\frac{3}{4\pi} \cdot \frac{m}{\rho_s} \right)^{1/3} \tag{1.9}$$

Building upon prior work done in the 2010 paper by Birnstiel, Dullemond & Brauer [BDB10], we will assume that the solid density of the dust particles carries a value of

$$\rho_s = 1600\,\mathrm{kg\,m^{-3}} \tag{1.10}$$

In reality, as a result of collisions leading to sticking, these dust particles can take on very complex, fractal shapes. This complicated structure makes a detailed treatment of large particle numbers basically impossible. Instead, highly simplified assumptions have to be made, visualizations of which are displayed schematically in figure 1.4.

As mentioned earlier, we will base our model on the simplest of these cases, shown in figure 1.4a.



(a) Compact

(b) Linear

(c) Porous (BPCA)

(d) Fractal (BCCA)

Figure 1.4: Schematic drawing of different models for types of dust particle structure.

1.4.2 The Dust Particle Mass Distribution Function

The concept of the *dust particle mass distribution function* $n(m, \vec{r}, t)$ lies at the heart of the dust coagulation model that is built for this thesis. This function is defined in such a way that, at a given time t and position $\vec{r}$ in the disk, the expression

$$N(\vec{r}, t) = \int_0^\infty n(m, \vec{r}, t)\ \mathrm{d}m \tag{1.11}$$

denotes the total number of dust particles N per unit volume, i.e. the dust particle number density, at the considered region in the proto-planetary disk. Its relationship to the mass volume density ρ_d of the dust at a given position $\vec{r}$ in the disk is given by

$$\rho_d(\vec{r}, t) = \int_0^\infty n(m, \vec{r}, t) \cdot m\ \mathrm{d}m \tag{1.12}$$

Before we can use numerical integration techniques to calculate an approximate solution to the Smoluchowski coagulation equation, it will be necessary to define an analogous, but discretized formulation of the expressions we defined above.

Let m_i label the mass values along the discretized mass axis we defined in section 1.3. Specifically, with this we mean the mass values m_i^c at the bin centers, and onwards from here we will do so everywhere in this thesis, if not otherwise specified. As before, the number of bin is labeled $\mathcal{N}_m$, with an lower and upper axis boundary m_{min} and m_{max}, respectively. The mass "width" of a given bin at a position i in the grid is labeled Δm_i.

We will begin with the discretized particle mass distribution function. For this, let us introduce the notation

$$n_i := n(m_i) \tag{1.13}$$

The total number of dust particles per unit unit volume corresponding to a given bin i is then calculated by integrating the particle mass distribution function from the lower to the upper bin boundary:

$$N_i = \int_{m_{i-1/2}}^{m_{1+1/2}} n(m)\ \mathrm{d}m \tag{1.14}$$

$$\approx n_i \cdot \Delta m_i \tag{1.15}$$

The mass density ρ_i, i.e. the dust mass per unit volume belonging to a bin i is then given by

$$\rho_i = \int_{m_{i-1/2}}^{m_{1+1/2}} m \cdot n(m)\ \mathrm{d}m \tag{1.16}$$

$$\approx n_i \cdot m_i \cdot \Delta m_i \tag{1.17}$$

Since each bin in the mass grid contains only a fraction of the total dust mass density $\rho_d(\vec{r})$ at a given point $\vec{r}$ in space, we have to sum over all bins to arrive at the total mass density independent of particle mass:

$$\rho_d = \sum_{i=1}^{\mathcal{N}_m} \rho_i \tag{1.18}$$

Studying the temporal evolution of the functions n_i, N_i, and ρ_i will be one of this thesis' main goals.

1.4.3 Initialization of the Dust Particle Mass Distribution

Before the temporal evolution of the dust mass distribution can be addressed, it is of course necessary to define initial conditions. For this, let us label the initial state of the dust particle mass distribution function as well as the particle number density

$$n_0 := n(m = m_0, \vec{r} = \vec{r}_0, t = 0) \qquad \text{and} \tag{1.19}$$
$$N_0 := N(m = m_0, \vec{r} = \vec{r}_0, t = 0) \tag{1.20}$$

respectively. Here, $\vec{r}_0$ labels the position in the disk where the simulation takes place.

At the beginning of a given simulation, we will assume all dust particles to share the same mass m_0 and the same corresponding radius. The appropriate particle mass distribution function for the this scenario is given by a Dirac δ-distribution:

$$n(m, \vec{r}, t = 0) = \delta_D(m - m_0) \cdot n_0 \tag{1.21}$$

Plugging the above definition back into equation (1.12) leads to

$$\rho_d(\vec{r}, t = 0) = \int_0^\infty n(m, \vec{r}, t = 0) \cdot m \, \mathrm{d}m \tag{1.22}$$
$$= \int_0^\infty \delta_D(m - m_0) \cdot n_0 \cdot m \, \mathrm{d}m \tag{1.23}$$
$$= N_0 \cdot m_0 \tag{1.24}$$

This reflects the fact that we initially want to have N_0 particles per unit volume, each carrying a mass of value m_0.

Now we would like to define a discretized analogon of the above. To do this, we can use the Kronecker δ-function to define list of numbers with length $\mathcal{N}_m$, consisting of $\mathcal{N}_m - 1$ zeros, and a single non-zero value n_0 as the first entry.

$$n_i(t = 0) = \delta_{ij} \cdot n_0 \qquad \text{where} \qquad j = 0 \tag{1.25}$$

Plugging this definition into equation (1.18) yields

$$\rho_d = \sum_{i=1}^{\mathcal{N}_m} n_i \cdot m_i \cdot \Delta m_i \tag{1.26}$$
$$= \sum_{i=1}^{\mathcal{N}_m} \delta_{i0} \cdot n_0 \cdot m_i \cdot \Delta m_i \tag{1.27}$$
$$= \cdot n_0 \cdot m_0 \cdot \Delta m_0 \tag{1.28}$$

If not specified otherwise, the initial particle mass will be set to $m_0 = m_c^1$ for the remainder of this thesis. This is the mass value at the center of the first ("smallest") bin in the discretized mass axis.

Chapter 2

The Proto-Planetary Disk Model

In the following we will concern ourselves with the definition of a simple model for a proto-planetary disk. This model should include descriptions for both the gas and dust contents of the disk as well.

Before we start, it needs to be said here that due to the complexity of the involved processes and the astronomically large number of particles present in the disk, our model can only ever be a highly simplified approximation to reality.

Nonetheless, it is our goal to formulate a model which captures a "typical" PPD's most relevant properties, in order to allow the conduction of studies regarding dust particle coagulation and/or fragmentation in the subsequent chapters.

2.1 Preliminary Definitions and Assumptions

We will base this model on the aforementioned nebular hypothesis: The origin of the modeled proto-planetary disk is assumed to be found in the collapse of a giant interstellar cloud of gas and dust. Having inherited its contents, the disk's elementary and material composition mirror that of the interstellar medium. As such, the disk contains mainly molecular hydrogen H_2, to a lesser extent mono-atomic helium He, and small trace amounts of metals.

Let m_p label the mass of a single proton, and let us use μ to denote the mean molecular weight of the gas. Using these definitions, we can then write

$$m_g := \mu \cdot m_p \qquad \text{with} \qquad \mu := 2.3 \tag{2.1}$$

to express the mean molecular mass m_g of an individual gas particle.

The collapse of the interstellar gas cloud is assumed to have happened sufficiently long ago such that a star could already form from the matter accumulating at the center of the disk. For the studies that will be carried out in the context of this thesis, this star's most relevant properties will be its mass M_* as well as its luminosity L_*. In the following, both of these properties will be assumed to equal those of our own solar system's Sun. Thus, we define

$$M_* := M_\odot = 1.989 \times 10^{30}\,\mathrm{kg}, \text{ and} \tag{2.2}$$
$$L_* := L_\odot \ = 3.828 \times 10^{26}\,\mathrm{J\,s^{-1}}, \tag{2.3}$$

where $M_\odot$ and $L_\odot$ label the solar mass and luminosity, respectively.

Like in our own Solar System, where the Sun is responsible for more than 99.9 % of the total mass, we will assume most of the original gas cloud's mass to be concentrated inside the star in our model as well, while the disk surrounding it holds only a much smaller fraction of the total mass.

Let M_* and M_{disk} label the masses of the star and the disk, respectively. The ratio between the two shall then be defined as

$$q_{\text{disk-to-star}} = \frac{M_{\text{disk}}}{M_*} := 0.01 \tag{2.4}$$

Furthermore, the disk's mass M_{disk} is assumed to be dominated by the contribution of circumstellar gas, with only a comparatively tiny contribution of dust particles. Let M_{dust} and M_{gas} label the total mass present in the form of dust and gas, respectively. Since in the context of our model the disk consists entirely out of gas and dust, they must satisfy the relationship

$$M_{\text{disk}} = M_{\text{gas}} + M_{\text{disk}} \tag{2.5}$$

The relative abundance between the two particle classes can then be encoded into the dust-to-gas ratio, which we will define as

$$q_{\text{dust-to-gas}} = \frac{M_{\text{dust}}}{M_{\text{gas}}} := 0.01 \tag{2.6}$$

To keep things from getting too complicated, we will adopt a relatively straight-forward description of the circumstellar gas particles. We will assume the validity of the ideal gas law, and a perfectly isothermal behavior of the gas. Moreover, let us assume the gas has had enough time to settle into hydrostatic equilibrium, such that the gravitational influence of the star at the disk center is balanced out by an outward force due to the gas pressure gradient.

The isothermal assumption allows us to relate gas pressure P_g, gas volume density ρ_g, and sound speed c_s via the expression

$$P_g = \rho_g \cdot c_s^2 \tag{2.7}$$

The hydrostatic assumption in turn allows us to express the relationship between the gas pressure $P_g(r, z)$ and the gas volume density $\rho_g(r, z)$. If we let g_z label the acceleration due to gravity along the z-axis, then we can write this relationship as

$$\frac{\mathrm{d}P_g}{\mathrm{d}z} = -\rho_g \cdot g_z, \tag{2.8}$$

Both equation (2.7) and equation (2.8) will be useful in the definition of the radial gas density profiles in section 2.2. These profiles, as well as the profiles of other physical quantities, will be defined on a radial axis ranging from a lower to an upper boundary, the value of which is given by

$$r_{\text{min}} = 10^{-2} \text{ au} \quad \text{and} \tag{2.9}$$

$$r_{\text{max}} = 10^{+2} \text{ au} \tag{2.10}$$

respectively. This radial axis is discretized into $\mathcal{N}_r$ bins according to the definition given in section 1.3.

Along that axis, in the next sections we will now define the radial dependency of some of the disk's properties. Most notably, this includes the definition of radial profiles for the gas surface & volume densities. From these, we will then be able to derive the volume density of dust particles, which will be needed for the initialization of the dust particle mass distribution function.

After having defined the spatial distribution of gas and dust in the disk, later on we will focus on only one specific location in the disk for the rest of this thesis. Adopting cylindrical coordinates $\vec{r} = (r, \phi, z)^T$, we will define the position of interest $\vec{r}_0$ as

$$\vec{r}_0 = (r_0, \phi_0, z_0)^T \quad \text{with} \quad r_0 := 10 \text{ au} \quad \text{and} \quad z_0 := 0 \tag{2.11}$$

As such, this position of interest is at the midplane, at a distance of 10 au from the disk center. Assuming a radial symmetry of the disk's properties on macroscopic scales, the definition of a specific azimuthal angle ϕ becomes superfluous here.

2.1.1 Radial Profile of the Midplane Temperature

The stellar luminosity L_* labels the total amount of energy emitted per time by the star at the center of the disk. It is related to the stellar radiation flux B_* via the following surface integral:

$$L_* = \int_A B_* \, \mathrm{d}A \tag{2.12}$$

Assuming the star's radiative power output is directed isotropically in all directions, we can simplify the above and then solve for the radiation flux $B_*(r)$ at a distance r from the star.

$$B_*(r) = \frac{L_*}{4\pi r^2} \tag{2.13}$$

Let f denote the angle of incidence. i.e. the "flaring angle" of the light onto the disk, and assume that one half of the light is scattered into, the other out of the disk. Then, the radiation flux $B_{\mathrm{mid}}(r)$ that actually reaches the midplane can be written as

$$B_{\mathrm{mid}}(r) = \frac{f}{2} \cdot B_*(r) \quad \text{with} \quad f := 0.05 \tag{2.14}$$

We make the assumption that the gas and dust particles in the disk can be modeled as perfect black-body objects. Then, the relationship between the radiation flux B_{mid} and the resulting temperature T_{mid} is given by the Stefan-Boltzmann law, which reads

$$B_{\mathrm{mid}}(T) = \sigma_{\mathrm{SB}} \cdot T_{\mathrm{mid}}^4 \tag{2.15}$$

Plugging this all in and rearranging for the midplane temperature yields

$$\boxed{T_{\mathrm{mid}}(r) = \left(\frac{f}{2} \cdot \frac{L_*}{4\pi r^2 \cdot \sigma_{\mathrm{SB}}} \right)^{1/4}} \tag{2.16}$$

The results can be seen visualized in figure 2.1.

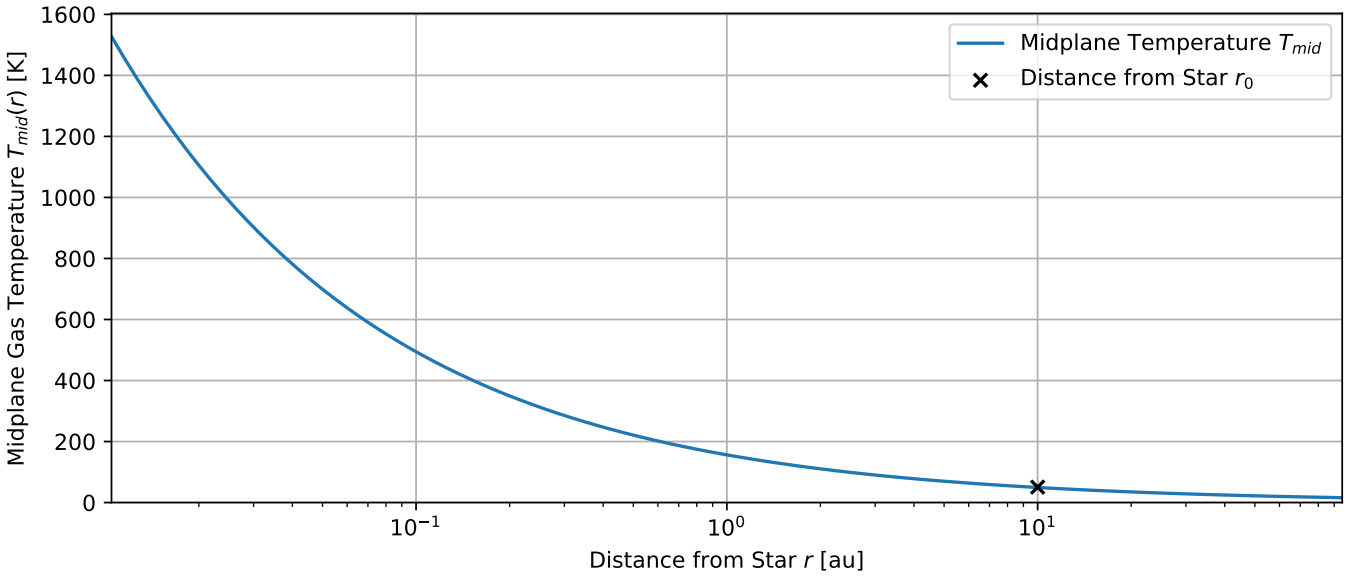


Figure 2.1: Radial profile of the midplane gas temperature $T_{\mathrm{mid}}(r)$.

2.1.2 Radial Profile of the Midplane Speed of Sound

The speed of sound is defined as the speed at which pressure disturbances travel in a given medium. It is given by

$$c_s = \sqrt{\frac{\partial P_g}{\partial \rho_g}} \tag{2.17}$$

where P_g and ρ_g label the gas pressure and the volume density of gas mass, respectively.
As mentioned before in section 2.1, we will assume the validity of the ideal gas equation, which reads

$$P_g \cdot V = N \cdot k_B \cdot T \tag{2.18}$$

Here, the number of particles per volume N/V can be written as the ratio ρ_g/m_g between mass volume density ρ_g and mean particle mass m_g. As such, we can rewrite equation (2.18) into

$$P = \frac{\rho_g \cdot k_B \cdot T_{\text{mid}}}{m_g} \tag{2.19}$$

Note that here we set the temperature $T = T_{\text{mid}}$, constraining our considerations to the disk midplane.

Now we can plug equation (2.19) into equation (2.17), and arrive at the isothermal speed of sound as a function of the distance r from the disk center:

$$c_s(r) = \sqrt{\frac{k_B \cdot T_{\text{mid}}(r)}{m_g}} \tag{2.20}$$

The results can be seen visualized in figure 2.2.

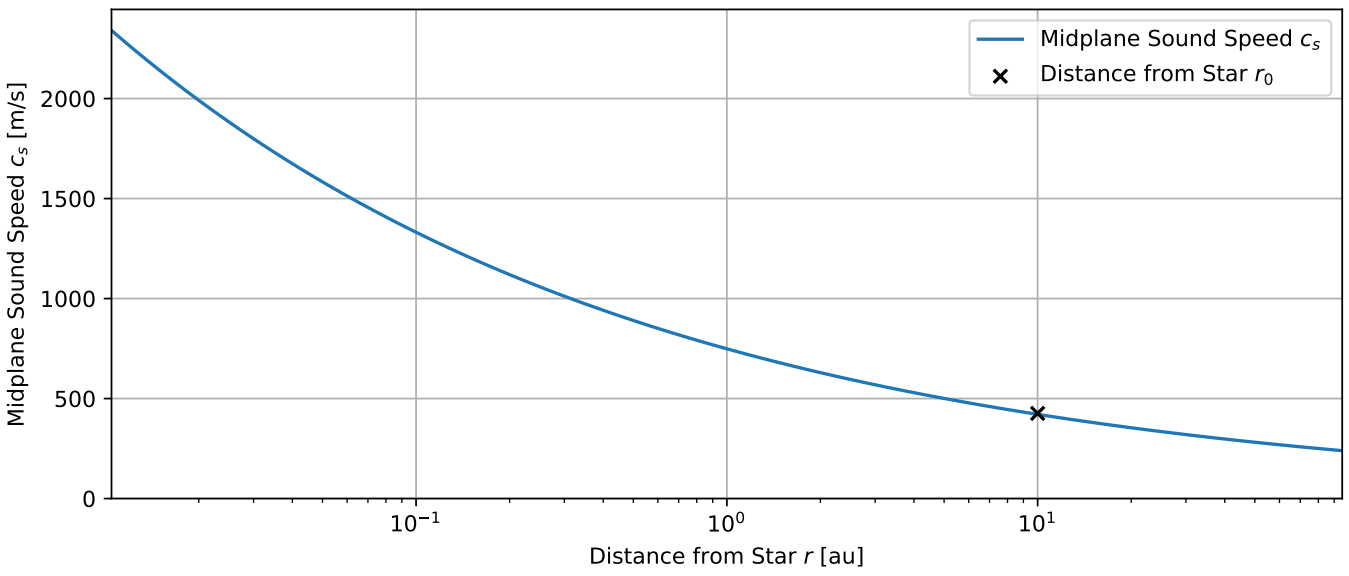


Figure 2.2: Radial profile of the midplane sound speed $c_s(r)$.

2.2 Gas and Dust Density

Our next task will be to formulate a simple model for the spatial distribution of the disk's gaseous contents. In this context, we will adopt cylindrical coordinates $\vec{r} = (r, \phi, z)^T$ and define both the surface and volume density profiles along the radial axis r.

2.2.1 Radial Profile of the Gas Surface Density

The gas surface density $\Sigma_g(r, \phi)$ is directly related to the gas volume density $\rho_g(r, \phi, z)$, and can be derived from it by integrating the volume density from $-\infty$ to $+\infty$ along the z-axis, i.e. along the axis perpendicular to the disk:

$$\Sigma_g(r, \phi) = \int_{-\infty}^{\infty} \rho_g(r, \phi, z) \, \mathrm{d}z \tag{2.21}$$

As noted in the context of the definition given in equation (2.11), we will assume the large-scale structure of the disk to be approximately radially symmetric. This, of course, is quite a substantial simplification that neglects the evolution of the disk structure on macroscopic scales, but will have to suffice as a first approximation in the context of this thesis. For the density distributions, assuming radial symmetry means that

$$\Sigma_g(r, \phi) = \Sigma_g(r) \qquad \text{and} \qquad \rho_g(r, \phi) = \rho_g(r) \tag{2.22}$$

Thus, we can simplify equation (2.21) and write:

$$\Sigma_g(r) = \int_{-\infty}^{\infty} \rho_g(r, z) \, \mathrm{d}z \tag{2.23}$$

Note that we do not yet have an expression for the gas volume density ρ_g so far. Therefore, we cannot derive the surface density Σ_g by simply evaluating the above integral. Instead, we will go the other way round and make an ansatz for the gas surface density. Then we will be able to derive the volume density from that.

Here, we will follow the work of Brauer, Dullemond, & Henning [BDH08] and make the assumption that the radial profile of the surface density can be modeled using an inverse power law of the form

$$\Sigma_g(r) = \Sigma_0 \cdot \frac{1}{r^\sigma} \qquad \text{with} \qquad \sigma := 0.8 \tag{2.24}$$

Before we can use this though, we still need to define a value for Σ_0, which we will do in the following: The total gas mass present in the disk is given by M_{gas}, the exact numerical value of which can be derived from what we defined in section 2.1. Since the integration of the surface density over the entire disk surface must yield this value, we can formulate the following condition:

$$M_{\text{gas}} = \int_{r_{\min}}^{r_{\max}} \Sigma_g(r) \cdot 2\pi r \, \mathrm{d}r \tag{2.25}$$

Due to the discretization of the radial axis into $\mathcal{N}_r$ "bins", we can rewrite this as

$$M_{\text{gas}} = \sum_{i=1}^{\mathcal{N}_r} \Sigma_g(r_i) \cdot 2\pi r_i \cdot \Delta r_i \tag{2.26}$$

Now plug in equation (2.24) and rearrange for Σ_0 to arrive at

$$\Sigma_0 = M_{\text{gas}} \cdot \left[\sum_{i=1}^{\mathcal{N}_m} 2\pi r_i^{1-\sigma} \cdot \Delta r_i \right]^{-1}, \tag{2.27}$$

which is all we need to formulate an expression for the radial profile of the gas surface density.

2.2.2 Radial Profile of the Gas Volume Density

Let us now use the gas surface density that we just defined to derive an expression for the gas volume density. For this, consider again a point $\vec{r} = (r, \phi, z)^T$ in the disk. The acceleration due to Newtonian gravity exerted by the star at the disk center on a massive object at position $\vec{r}$ is given by

$$\vec{g} = -\frac{GM_*}{r^2 + z^2} \cdot \frac{\vec{r}}{|\vec{r}|} \tag{2.28}$$

Now consider the schematic drawing in figure 2.3, which we might imagine as a test mass sitting at a position $\vec{r}$, with a radial distance r from the star, and a height z above the midplane. Making use of basic trigonometry, we can write

$$g_z = |\vec{g}| \cdot \sin\theta \tag{2.29}$$

and

$$\sin\theta = \frac{z}{\sqrt{r^2 + z^2}} \tag{2.30}$$

As was mentioned in section 2.1, the gas in the proto-planetary disk is assumed to be in perfect hydrostatic equilibrium, which lets us make use of equation (2.8). If we then plug in the trigonometric relations from above, we can write

$$\begin{aligned}
\frac{\mathrm{d}P_g}{\mathrm{d}z} &= -\rho_g \cdot g_z && (2.31)\\
&= -\rho_g \cdot |\vec{g}| \cdot \sin\theta && (2.32)\\
&= -\rho_g \cdot \frac{GM_*}{r^2 + z^2} \cdot \frac{z}{\sqrt{r^2 + z^2}} && (2.33)\\
&= -\rho_g \cdot z \cdot \frac{GM_*}{(r^2 + z^2)^{\frac{3}{2}}} && (2.34)
\end{aligned}$$

As defined in equation (2.11), the position of interest lies at a distance of 10 au from the star in the disk midplane. As such, we can safely say that $r \gg z$. This allows us to make the approximation

$$\frac{GM_*}{(r^2 + z^2)^{3/2}} \approx \frac{GM_*}{r^3} = \Omega_K^2 \tag{2.35}$$

and thus, together with equation (2.34), lets us write

$$\frac{\mathrm{d}P_g}{\mathrm{d}z} \approx -\rho_g \cdot z \cdot \Omega_K^2 \tag{2.36}$$

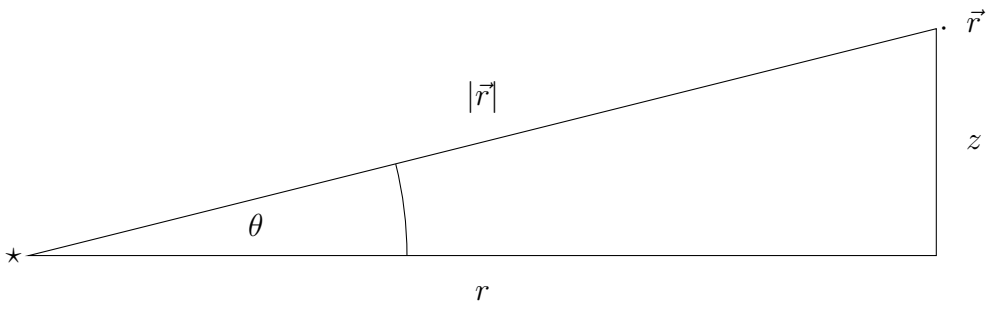


Figure 2.3: Illustration of Trigonometric relations used for expressing the z-component of the acceleration due to the gravitational influence of the star at the center of the disk.

The second assumption we made about the gas in section 2.1 is that its behavior can be modeled as perfectly isothermal. Consider the isothermal condition given in equation (2.7). If we differentiate both sides of this equation with respect to z, we get

$$\frac{\mathrm{d}P_g}{\mathrm{d}z} = \frac{\mathrm{d}\rho_g}{\mathrm{d}z} \cdot c_s^2 \tag{2.37}$$

Thus, we now have two expressions for $\mathrm{d}P_g/\mathrm{d}z$. Equating both of these and then solving for $\mathrm{d}\rho_g/\mathrm{d}z$ leads to the following differential equation:

$$\boxed{\frac{\mathrm{d}\rho_g}{\mathrm{d}z}(r,z) = -\left(\frac{\Omega_K}{c_s}\right)^2 \cdot \rho_g(r,z) \cdot z} \tag{2.38}$$

This differential equation can be solved by making the ansatz of a Gaussian distribution

$$\boxed{\rho_g(r,z) = \rho_g^{\mathrm{mid}} \cdot \exp\left(-\frac{z^2}{2 \cdot H_p^2}\right)} \tag{2.39}$$

where

$$H_p(r) := \frac{c_s}{\Omega_K} \tag{2.40}$$

labels the so-called *pressure scale height* of the disk, Ω_K is the Kepler frequency, and

$$\rho_g^{\mathrm{mid}}(r) := \rho_g(r, z=0) \tag{2.41}$$

labels the gas volume density at the midplane of the disk. Now that we know the behavior of the gas volume density, we can easily derive it from the gas surface density like this:

$$\boxed{\rho_g^{\mathrm{mid}}(r) = \frac{\Sigma_g}{\sqrt{2\pi} \cdot H_p}} \tag{2.42}$$

We can use this to express the number density, i.e. the number of gas particles per unit volume, as

$$N_g = \frac{\rho_g}{m_g} \tag{2.43}$$

2.2.3 Radial Profile of the Dust Volume Density

After having derived an expression for the radial profile of the gas volume density at the disk midplane in equation (2.42), we can use the definition of the dust-to-gas ratio in equation (2.6) to obtain the *dust volume density*, i.e. the amount of dust mass per unit volume:

$$\rho_d(\vec{r},t) = q_{\text{dust-to-gas}} \cdot \rho_g(\vec{r},t) \tag{2.44}$$

This quantity will be relevant to in the context of the studies of dust coagulation and fragmentation that will be done later on, where we will model the temporal evolution of the distribution of dust particle masses. If we were to add up all the particles in the distribution, each mutliplied by the value of its mass, then we would arrive back at this value of ρ_d.

Note that in the highly simplified model used throughout this thesis, we will assume both ρ_g as well as ρ_d to stay constant under translations in time, as we do not include the large-scale movement of matter from one region to another in the disk.

Instead, we position ourselves at a position $\vec{r}_0$ in the disk at a time t_0, and only consider the gas and dust densities at that position, and at that point in time, i.e. we focus on $\rho_d(\vec{r}_0, t_0)$ and $\rho_g(\vec{r}_0, t_0)$.

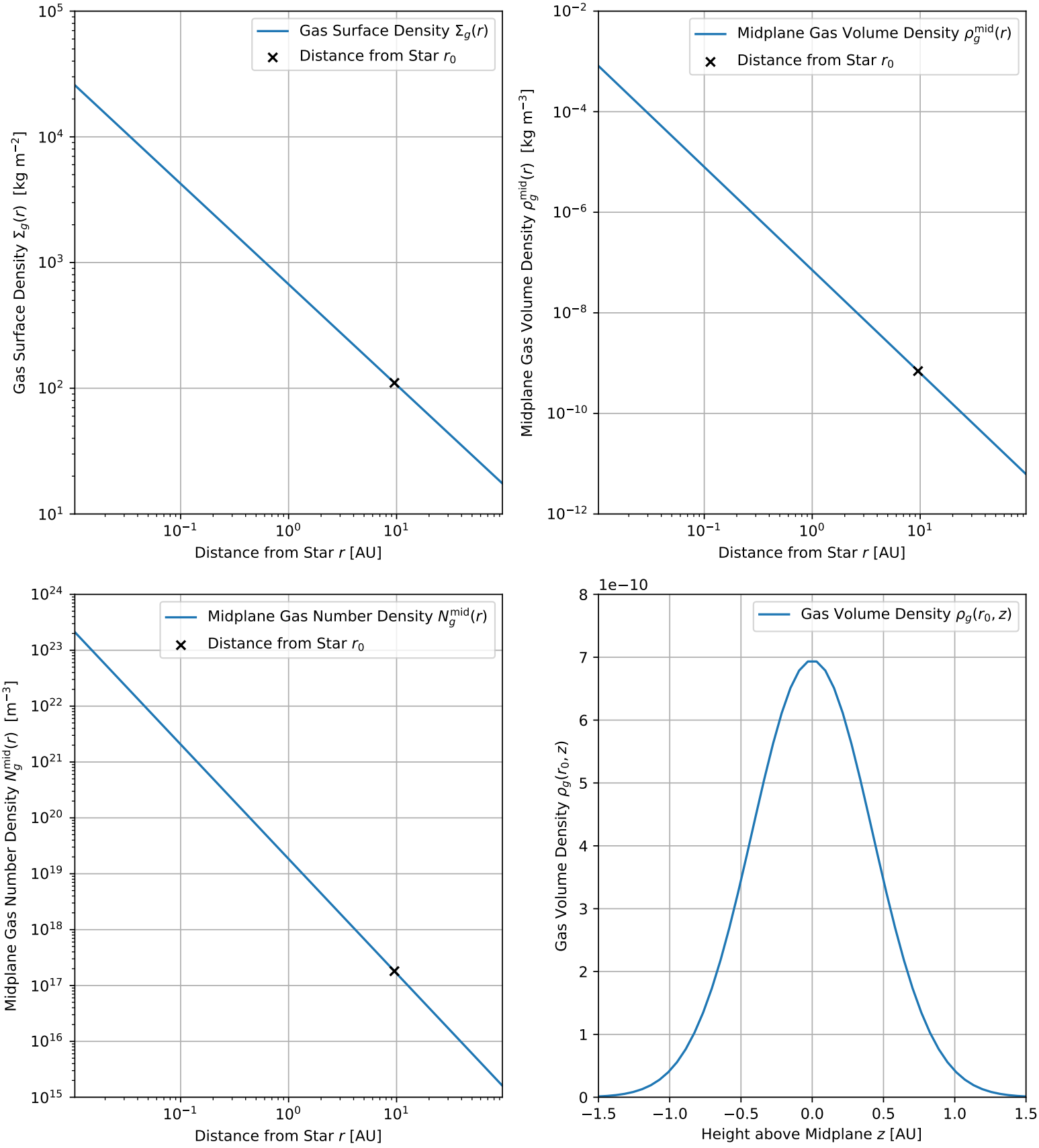


Figure 2.4: *Top left:* Radial profile of the gas surface density $\rho_g(r)$. *Top right:* Radial profile of the gas volume density at the disk midplane $\rho_g^{\mathrm{mid}}(r)$. *Bottom left:* Radial profile of the gas particle number density at the disk midplane $N_g^{\mathrm{mid}}(r)$. *Bottom right:* z-profile of the gas volume density at $r = r_0 = 10$ au.

2.3 Gas Particle Kinematics

Let us now attempt to model the kinematics of circumstellar gas particles along the radial axis. The value of the radial gas velocity u_g impacts the dust kinematics as well, and will later be relevant for the definition of the relative dust velocities due to radial dust particle drift.

2.3.1 Gas Viscosity

Before we can define the radial gas velocity, we first have to focus on the definition of the kinematic gas viscosity. It can be expressed as

$$\nu_g = \frac{1}{2} \cdot u_{\text{th}} \cdot \lambda_{\text{mfp}} \tag{2.45}$$

We define the absolute value of the thermal velocity to be the mean of the magnitude of the molecular velocity. It is given by

$$u_{\text{th}} = \sqrt{\frac{8}{\pi}} \cdot c_s \tag{2.46}$$

A particle's mean free path in a given medium is given by the inverse product between the number of particles per unit volume and the cross section for a collision:

$$\lambda_{\text{mfp}} = \frac{1}{n \cdot \sigma_{\text{H}_2}} \tag{2.47}$$

Here, we follow the 2010 work of Birnstiel, Dullemond, and Brauer [BDB10] and use a value of

$$\sigma_{\text{H}_2} = 2 \times 10^{-19}\,\text{m}^2 \tag{2.48}$$

for the gas particles' collision cross section.

2.3.2 Radial Gas Velocity

According to [BDB10], the viscous evolution of the gas in the proto-planetary disk can be described by the continuity equation

$$\frac{\partial \Sigma_g}{\partial t} - \frac{1}{r}\frac{\partial}{\partial r}\left(\Sigma_g \cdot r \cdot u_g\right) = S_g \tag{2.49}$$

Here, u_g labels the *radial gas velocity*.

A solution to this equation was given in 1974 by Lynden-Bell and Pringle [LP74], who expressed the radial velocity of the gas as

$$u_g = -\frac{3}{\Sigma_g \cdot \sqrt{r}} \cdot \frac{\partial}{\partial r}\left(\Sigma_g \cdot \nu_g \cdot \sqrt{r}\right) \tag{2.50}$$

We will use this expression to model the radial gas velocity as well. It will be needed for the definition of the velocities involved in dust kinematics, which are heavily influenced by the motion of gas in the space surrounding the dust particles.

2.4 Dust Particle Kinematics

The construction of the kernel matrix requires the definition of the rate at which collision occur between pairs of differently sized particles. Therefore, it is essential to model the kinematics of the dust particles in an appropriate manner and arrive at an expression for the relative velocities involved in a given particle collision. These relative velocities will influence both the collision rate as well as their outcomes, since the fragmentation probability is influenced by the relative velocity as well.

We include four processes in the calculation of the relative dust particle velocities. There are the two systematic contributions of radial and azimuthal drift, and the two stochastic contributions of Brownian and turbulent motion.

We will start with the definition of the systematic contributions, then move on to define the stochastic contributions, and finally combine the four different effects into the total relative velocity value

$$\Delta v_{ij} := \Delta v(m_i, m_j) \tag{2.51}$$

for a given pair of particles carrying masses m_i and m_j, respectively.

2.4.1 Radial Drift

Consider a particle suspended in a fluid flow. The ratio between the particle's characteristic time scale and the characteristic time scale of the fluid flow is referred to as *Stokes' number*, named after George Gabriel Stokes. In our model, it is given by

$$\mathrm{St}_i = \frac{\rho_s \cdot a_i}{\Sigma_g \cdot \frac{\pi}{2}} \tag{2.52}$$

Building upon the 2010 paper written by Birnstiel, Dullemond, and Brauer [BDB10], we will model the radial velocity $u_{r,i}$ of particles in the dust population corresponding to a given mass bin i as

$$u_{r,i} = \frac{u_g}{1 + \mathrm{St}_i^2} - \frac{2u_n}{\mathrm{St}_i + \mathrm{St}_i^{-1}} \tag{2.53}$$

Here, the first term in the above equation is the *drag term*. The gas moves along the radial axis with a radial velocity u_g which we defined in equation (2.50). The value of the Stokes number affects the extent to which the dust particles are coupled to the gas. In the case of a low Stokes number, the dust is dragged along with the gas, whereas large Stokes numbers describe particles whose kinematics is less influenced by the motion of the gas.

The second term is related to the fact that due to the disk's own gas pressure gradient, the gas in the disk orbits the disk center on a sub-keplerian trajectory. Heavier particles do not feel this pressure as much and thus move keplerian. As a result, there is a difference in the radial velocity of the dust particle and the gas through which it moves. Slowed down by the gas, the particle loses momentum and starts drifting inwards. The term u_n labels the maximum drift velocity of a particle, which was derived by Weidenschilling [Wei77], and is given by

$$u_n = -\frac{\partial P_g}{\partial r} \cdot \frac{E_d}{2\rho_g \Omega_K} \tag{2.54}$$

In it, the E_d labels the drift efficiency parameter that was introduced by [BDB10]. Here, it will be set to 1 for simplicity.

The relative velocity due to radial drift of dust particles is then given by

$$\Delta v_{ij}^{\mathrm{RD}} = |u_{r,i} - u_{r,j}| \tag{2.55}$$

For a visualization of the results, see figure 2.6a.

2.4.2 Azimuthal Drift

The azimuthal relative velocities of differently-sized dust particles are induced in a similar way to the radial drift. The extent to which dust particles are coupled to the gas depends directly on their Stokes number, and thus on the particle's size and mass.

Building upon the results of earlier studies done by Weidenschilling [Wei77] and Nakagawa [NSH86], we will express the relative azimuthal velocity $\Delta v_{ij}^{\mathrm{AZ}}$ for gas-dominated drag between a pair of particles with mass bin indices (i, j) as

$$\Delta v_{ij}^{\mathrm{AZ}} = \left| u_n \cdot \left(\frac{1}{1 + \mathrm{St}_i^2} + \frac{1}{1 + \mathrm{St}_j^2} \right) \right| \tag{2.56}$$

For a visualization of the resulting distribution of relative velocities, see figure 2.6b.

2.4.3 Turbulent Motion

For the calculation of the relative velocity contributions due to turbulent motion, we adopt the model outlined in the 2007 work by Ormel and Cuzzi [OC07].

There, the relative velocity $\Delta v_{ij}^{\mathrm{TU}}$ due to turbulence is defined as

$$\Delta v_{ij}^{\mathrm{TU}} = \begin{cases} \Delta v_{ij}^{\mathrm{A}} & \text{if } t_{sm} \leq t_s \\ \Delta v_{ij}^{\mathrm{B}} & \text{if } t_s < t_{sm} \leq t_n \\ \Delta v_{ij}^{\mathrm{C}} & \text{if } t_{sm} > t_n \end{cases} \tag{2.57}$$

For this definition, three different cases have to be handled separately.

1. Limit of tightly coupled particles (below the Kolmogorov scale)

$$\Delta v_{ij}^{\mathrm{A}} = u_g \cdot \sqrt{\frac{\mathrm{St}_i - \mathrm{St}_j}{\mathrm{St}_i + \mathrm{St}_j} \cdot \left(\frac{\mathrm{St}_i^2}{\mathrm{St}_i + \mathrm{Re}^{-1/2}} - \frac{\mathrm{St}_j^2}{\mathrm{St}_j + \mathrm{Re}^{-1/2}} \right)} \tag{2.58}$$

   For a detailed explanation, see section 3.4.1 of [OC07], and their equation 26.

2. At least one of the particles is not in the tightly coupled regime. Both particles have $\mathrm{St} < 1$.

$$\Delta v_{ij}^{\mathrm{B}} = u_g \cdot \sqrt{\mathrm{St}_i \cdot \left[2y_a - (1 + \varepsilon) + \frac{2}{1 + \varepsilon} \cdot \left(\frac{1}{1 + y_a} + \frac{\varepsilon^3}{y_a + \varepsilon} \right) \right]} \tag{2.59}$$

   where $\varepsilon < 1$ is the ratio of the stopping times and $y_a := 1.6$. See section 3.4.2 of [OC07], and their equation 28.

3. At least one of the particles is not in the tightly coupled regime. Also, here at least one of the particles has $\mathrm{St} \geq 1$.

$$\Delta v_{ij}^{\mathrm{C}} = u_g \cdot \sqrt{\frac{1}{1 + \mathrm{St}_i} + \frac{1}{1 + \mathrm{St}_j}} \tag{2.60}$$

   For a detailed explanation, see section 3.4.3 of [OC07], their equation 29.

In equation (2.58), the Reynolds number Re is defined as

$$\mathrm{Re} = \frac{\alpha c_s H_p}{\nu_g} \tag{2.61}$$

with the viscosity parameter set to $\alpha := 10^{-3}$. For a visualization of the results, see figure 2.6c.

2.4.4 Brownian Motion

Here, we follow the work published in the 2004 paper written by Dullemond and Dominik [DD04], where the contribution to relative velocity due to Brownian motion is given by

$$\Delta v_{ij}^{\mathrm{BR}} = \sqrt{\frac{8 k_B T}{\pi} \cdot \frac{m_i + m_j}{m_i \cdot m_j}} \tag{2.62}$$

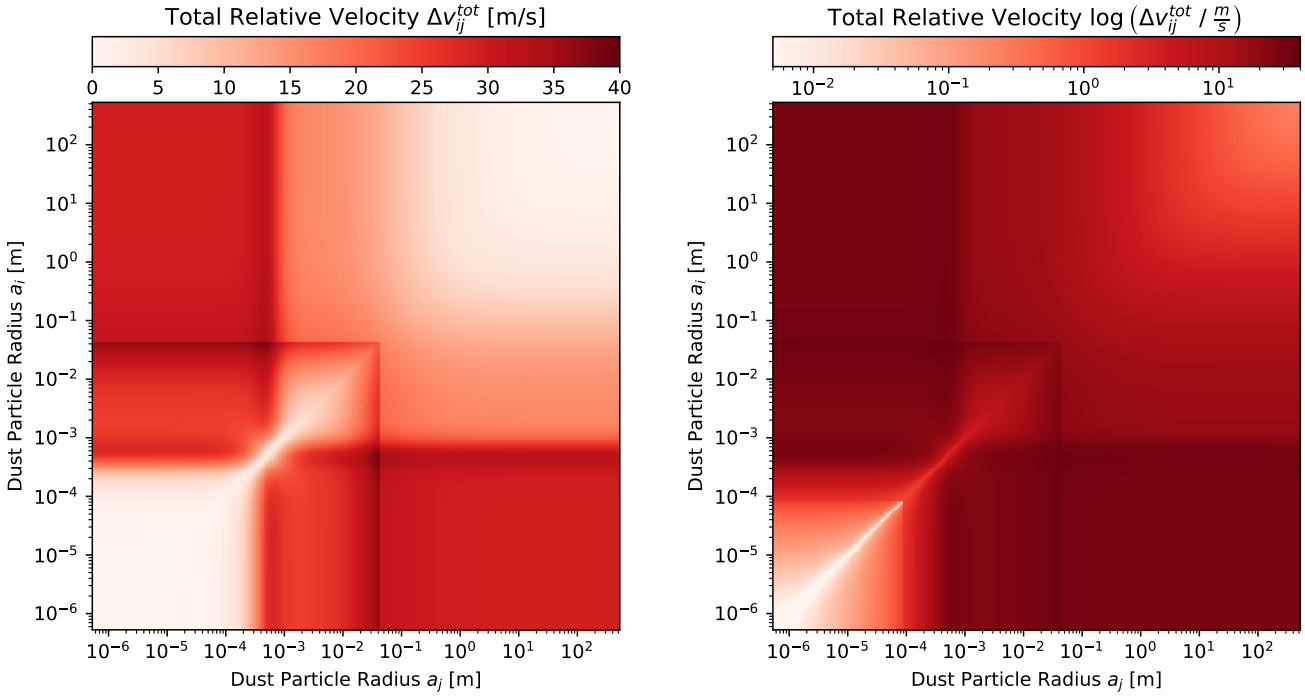
Note that this term will be the largest when both collision partners carry a very low mass. Brownian motion thus favors collisions in which at least one collision partner is a low-mass particles. This can be seen reflected in figure 2.6d.

2.4.5 Total Relative Velocity

The total relative dust particle velocity can be arrived at by calculating the root mean square (RMS) of the individual relative velocity contributions:

$$\Delta v_{ij} = \sqrt{\left(\Delta v_{ij}^{\mathrm{RD}}\right)^2 + \left(\Delta v_{ij}^{\mathrm{AZ}}\right)^2 + \left(\Delta v_{ij}^{\mathrm{BR}}\right)^2 + \left(\Delta v_{ij}^{\mathrm{TU}}\right)^2} \tag{2.63}$$

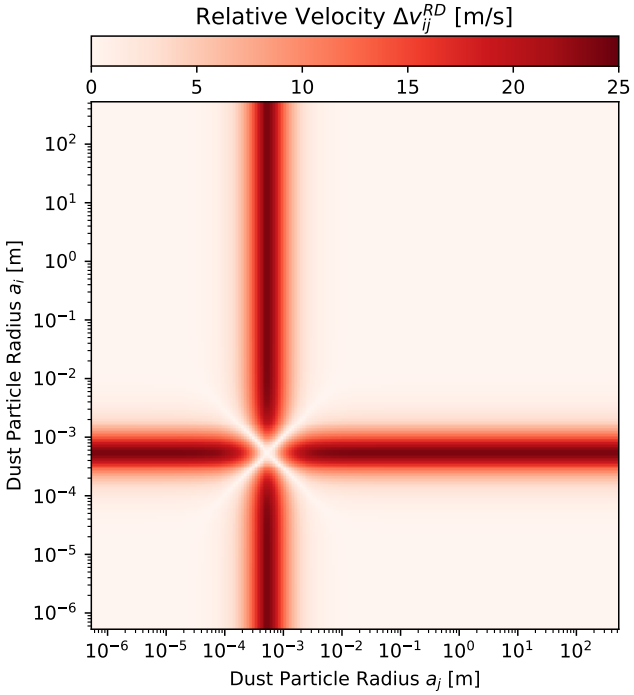
Having arrived at this expression now allows us to use it in the definition of the collision rates between differently-sized particles, which we will do in section 2.5. A visualization of the total relative velocity's dependence on the sizes of the two colliding particles can be seen in figure 2.5. Note that while the relative dust particle velocities are visualized up to a maximum particle radius of $> 10^2$ m, when integrating numerically an equilibrium state will be reached before particles can grow to that size, such that there we have an effective maximum particle radius of 10^{-4} m.



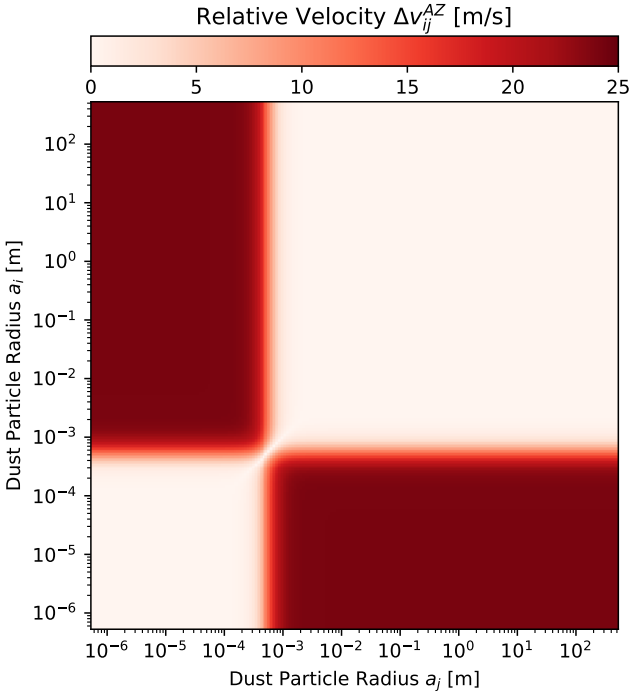
(a) $\Delta v_{ij}^{\mathrm{tot}}$ with Linear Color Scale

(b) $\Delta v_{ij}^{\mathrm{tot}}$ with Logarithmic Color Scale

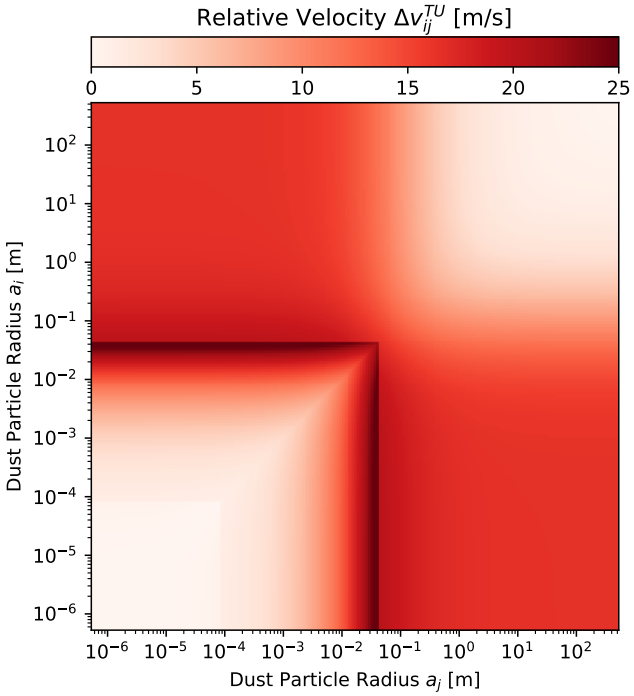
Figure 2.5: Total relative dust particle velocity $\Delta v_{ij}^{\mathrm{tot}}$.



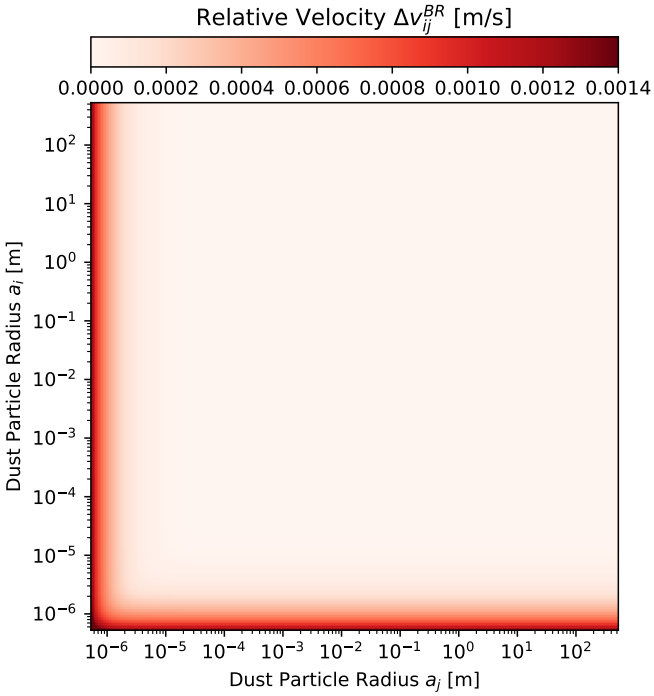
(a) Relative Velocity $\Delta v_{ij}^{\mathrm{RD}}$ due to Radial Drift



(b) Relative Velocity $\Delta v_{ij}^{\mathrm{AZ}}$ due to Azimuthal Drift



(c) Relative Velocity $\Delta v_{ij}^{\mathrm{TU}}$ due to Turbulent Motion



(d) Relative Velocity $\Delta v_{ij}^{\mathrm{BR}}$ due to Brownian Motion

Figure 2.6: Relative dust particle velocities due to (a) Radial drift $\Delta v_{ij}^{\mathrm{RD}}$, (b) Azimuthal drift $\Delta v_{ij}^{\mathrm{AZ}}$, (c) Turbulent motion $\Delta v_{ij}^{\mathrm{TU}}$, and (d) Brownian motion $\Delta v_{ij}^{\mathrm{BR}}$.

2.5 Dust Particle Collisions

As we saw in the last section, the relative velocities between differently-sized dust particls are influenced by partly random, and partly systematic contributions. Due to this, occasional collision events between dust particles are to be expected.

In the context of this thesis, these collisions are assumed to always be two-body interactions. The probability of three or more dust particles colliding at the exact same time is so small, that the occurance rate of three-body interactions can be neglected entirely. Thus, we can safely assume the particle collisions to be interactions of two particles only.

2.5.1 Collision Cross Section

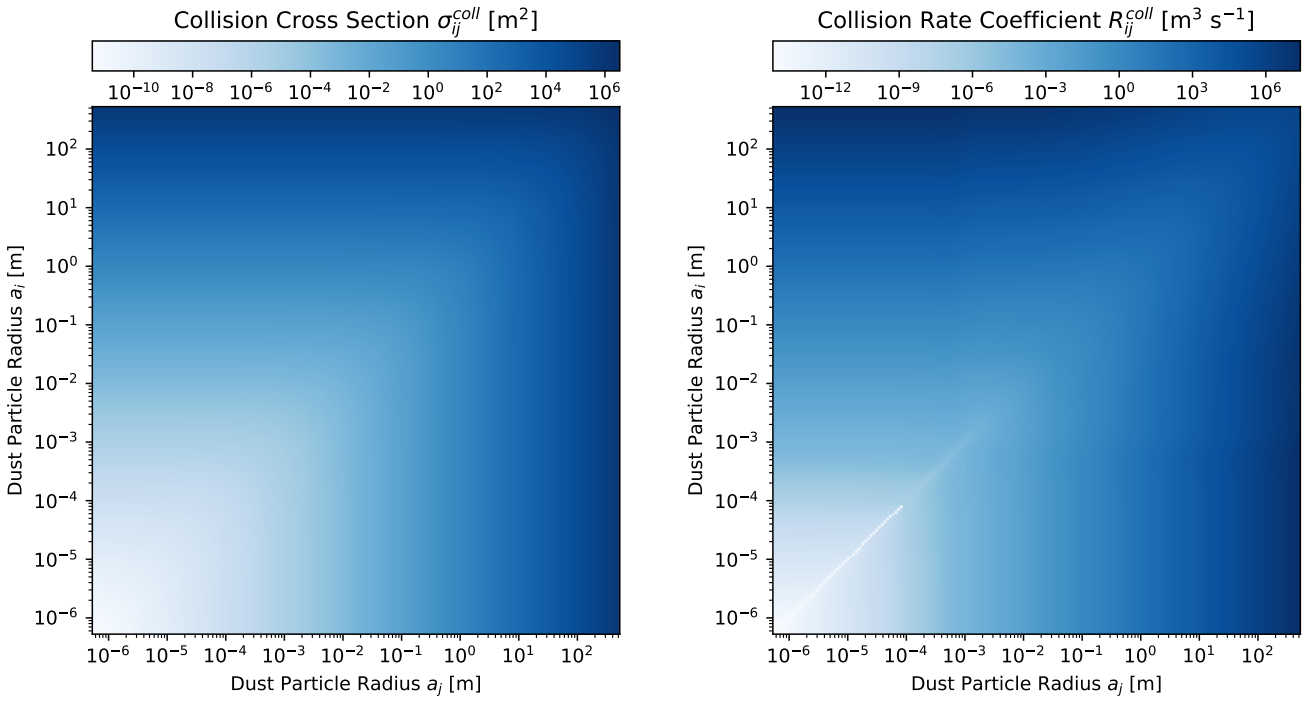
As was mentioned in section 1.4.2, in the context of this thesis we will assume the dust particles to possess a perfectly spherical shape.

Under this assumption, the cross section $\sigma_{ij}^{\text{coll}}$ for a collision of two dust particles with radii a_i and a_j can be approximated by the area of a circle the radius of which equals the sum of the two particles' radii.

Therefore, the collision cross section $\sigma_{ij}^{\text{coll}} := \sigma^{\text{coll}}(m_i, m_j)$ can be written as:

$$\sigma_{ij}^{\text{coll}} = \pi \cdot (a_i + a_j)^2 \tag{2.64}$$

It can be seen visualized in figure 2.7a.



(a) Dust Particle Collision Cross Section $\sigma_{ij}^{\text{coll}}$

(b) Dust Particle Collision Rate Coefficient R_{ij}^{coll}

Figure 2.7: Visualization of the dependence of the (a) Dust particle collision cross section $\sigma_{ij}^{\text{coll}}$ and (b) Dust particle collision rate coefficient R_{ij}^{coll} on the mass values of two colliding particles.

2.5.2 Collision Rate

The rate of collisions between pairs of particles from bins i and j can be described using the definition of the *collision rate coefficient* R_{ij}^{coll}, which is given as the product of the corresponding collision cross section and the two colliding particles' relative velocity:

$$R_{ij}^{\text{coll}} = R^{\text{coll}}(m_i, m_j) \approx \sigma_{ij}^{\text{coll}} \cdot \Delta v_{ij} \tag{2.65}$$

Note: The units of this collision rate coefficent are *not* s^{-1}! Instead, this coefficient is to be interpreted as a collision rate "per particle density", with units of $\text{m}^3\ \text{s}^{-1}$.

The actual rate of collisions per time can be calculated by multiplying the collision coefficient R_{ij}^{coll} with the number density of particles N_i per unit volume at a given position of interest in the disk.

2.5.3 Collision Outcomes

Now that we have a description for *how often* collisions between dust particles will occur in our model, we have to think about how to model *what happens* as a result of a collision.

In the context of this thesis, we will classify the different outcome scenarios into three main categories, which are briefly outlined in the following.

1. *Coagulation:* Here, the two colliding particle merge into a single new particle. Mass is conserved, such that the sum of the two initial particles' masses exactly equals that of newly created one.

2. *Bouncing:* The two particles collide and bounce off of each other. No particles are created or destroyed, and no mass is transferred between the particles. Since this scenario does not lead to a change in the dust particle mass distribution, it will be neglected throughout this work.

3. *Fragmentation:* One or both of the particles involved in the collision breaks apart into a possibly large number of smaller and less massive particles. Figuring out how one might model the distribution of masses among the newly created particles in such a scenario is a problem that we will address in section 3.4.1 and section 3.4.2.

A schematic representation of the collision types is displayed in figure 2.8.[1]

Which of these outcomes actually occurs as a result of a given collision of course heavily depends on the details of said collision. In our highly simplified model, the most notable of these details is the relative dust particle velocity.

To get an intuition for how the relative velocity influences the probability of a certain outcome to occur, one might imagine that collisions at low relative velocities have a high likelyhood of leading to a merge event. Higher velocity values may lead to bouncing, and still higher velocities would eventually have the effect of breaking apart the particles involved in the collision, i.e. fragmentation.

For each of these three possible outcomes, we will define an associated probability distribution, each of which depends on the mass values of the two particles participating in a given collision.

Let P^{coag}, P^{frag}, and P^{bounce} label the probabilities that a collision will lead to the respective outcome.

[1] For studies and visualizations regarding measurements of actual dust particle collisions (here: BPCA) in a laboratory see e.g. [Wad+09].

The rates at which such reactions occur can then described by the *reaction rate coefficients*, which we shall label R_{ij}^{coag}, R_{ij}^{frag}, and R_{ij}^{bounce}. They are defined as

$$R_{ij}^{\text{coag}} = R_{ij}^{\text{coll}} \cdot P_{ij}^{\text{coag}} \tag{2.66}$$

$$R_{ij}^{\text{frag}} = R_{ij}^{\text{coll}} \cdot P_{ij}^{\text{frag}} \tag{2.67}$$

$$R_{ij}^{\text{bounce}} = R_{ij}^{\text{coll}} \cdot P_{ij}^{\text{bounce}} \tag{2.68}$$

Collisions in the scenario of bouncing do not have an effect on the dust mass distribution among particles. But since a high value of P_{ij}^{bounce} will lead to lower reaction rates for the other two collision outcomes, the inclusion of bouncing effectively "slows down" the simulation.

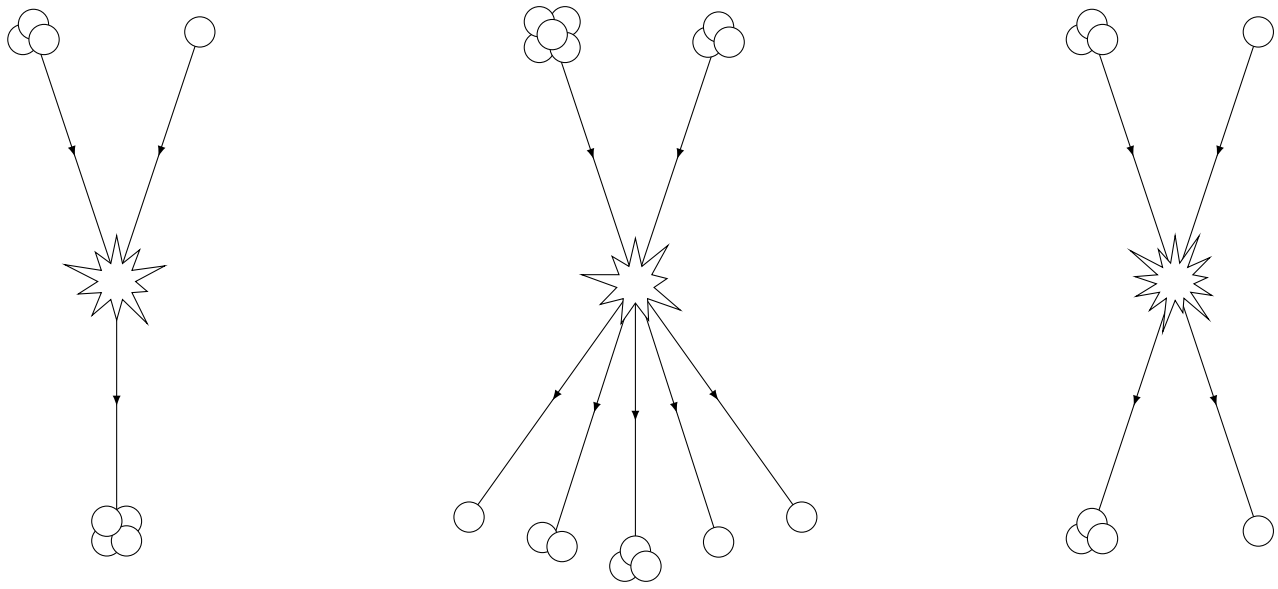
In the context of this thesis, we will not take the effect of bouncing into account, and only model the influence of coagulation and fragmentation processes. As such, we will adopt a bouncing probability

$$P_{ij}^{\text{bounce}} = 0 \tag{2.69}$$

To help with the decision whether a given collision ends up resulting in coagulation or fragmentation, the definition of a *threshold value* for the relative velocity makes sense. Adopting the value used by Dullemond and Dominik in their 2003 paper [DD23], we will define the value of v_{frag} as

$$v_{\text{frag}} := 1\,\text{m}\,\text{s}^{-1} \tag{2.70}$$

The idea is that collisions with slow relative velocities lead to a sticking of the two particles, while high velocities lead to them breaking apart.



(a) Coagulation (b) Fragmentation (c) Bouncing

Figure 2.8: Illustration of dust particle collision outcomes (a) Coagulation: Two particles collide and merge into a single new one that carries the combined mass of the initial particles. (b) Fragmentation: The colliding particles are destroyed during the collision. The total mass is subsequently distributed onto a wide range of differently-sized, newly-created particles. (c) Bouncing: Two particles collide, but no mass is exchanged between them.

Assuming a hard decision boundary, the fragmentation probability P_{ij}^{frag} for a collision involving two particles of masses m_i and m_j could be expressed as

$$P_{ij}^{\text{frag}} = \begin{cases} 1 & \text{if } \Delta v_{ij} \geq v_{\text{frag}} \\ 0 & \text{else} \end{cases} \tag{2.71}$$

A more elegant way though is given by in the 2022 paper by Stammler and Birnstiel [SB22]. Here, a correction to the fragmentation probability is made using the Maxwell-Boltzmann distribution.

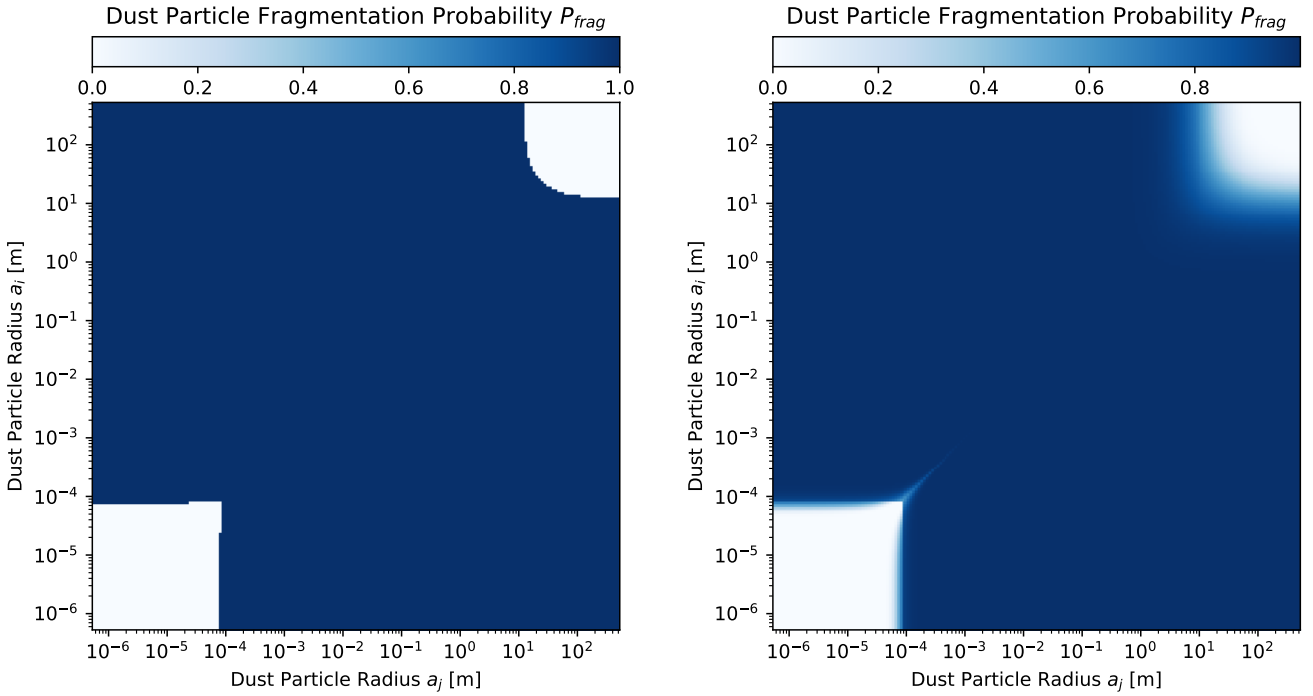
Adopting this method, we define

$$x := \frac{3}{2} \cdot \left(\frac{v_{\text{frag}}}{\Delta v_{ij}} \right)^2 \tag{2.72}$$

and, with this, the fragmentation probability

$$P_{ij}^{\text{frag}} = (1 + x) \cdot e^{-x} \tag{2.73}$$

The resulting probability distributions for the first method and the second, more sophisticated one, are displayed in figure 2.9. For our coagulation studies, we will make use of the second method.



(a) Frag. Prob. Assuming A Hard Decision Boundary (b) Correction Using Maxwell-Boltzmann Distribution

Figure 2.9: Dust particle fragmentation probability P_{ij}^{frag}.

Chapter 3

The Dust Coagulation Model

In the following, we will focus on the definition of a model for describing the processes involved in dust particle coagulation.

Adopting the formalism that was outlined in section 1.4.2, we would like to study the behavior of the particle mass distribution function $n(m)$ under the influence of particle collisions. Both associative and dissociative process (coagulation and fragmentation) are to be taken into account.

3.1 The Smoluchowski Coagulation Equation

The centerpiece of the dust coagulation model that this thesis is built upon is provided by the following integro-differential population balance equation:

$$\boxed{\frac{\partial n}{\partial t}(m) = \int_0^\infty \int_0^\infty K(m, m', m'') \cdot n(m') \cdot n(m'') \, \mathrm{d}m' \, \mathrm{d}m''} \tag{3.1}$$

This it the so-called *Smoluchowski coagulation equation*, which provides a mathematical framework that can be used to model the process of particle aggregation, or *coagulation*, into larger and larger structures.

The equation is named after the Polish physicist Marian Smoluchowski, due to his treatment of the equation in his 1916 paper [Smo16]. It finds applications in a wide range of scientific contexts. An example would be given by the modeling of aerosol particles growth in the atmosphere, but the equation can also be used in biological or chemical contexts, among others.

In the Smoluchowski coagulation equation, the term $K(m, m', m'')$ labels the so-called *coagulation kernel*. This is a function that we will use to encode the information about how the number of particles with mass m changes with time under the influence of collisions between particles with masses m' and m''.

The effects of both coagulation and fragmentation processes can be defined in separate kernel functions, that are subsequently added to acquire the total kernel. We will define these functions separately in section 3.3 and section 3.4, respectively.

3.1.1 Discretized Formulation of the Smoluchowski Equation

To discretize the Smoluchowski coagulation equation, we need to define discretized analogons for both the particle mass distribution function and its temporal derivative, as well as write the integral on the right-hand side of the equation as a sum.

Here we will assume that the particle mass axis from $m_{\min}$ to $m_{\max}$ has been discretized into a grid m_k consisting of $\mathcal{N}_m$ bins, and defined according to what was outlined in section 1.3, where we layed out a method for axis discretization.

For the mass distribution $n(m)$, we use the notation $n_k := n(m_k)$ as defined in equation (1.13). To discretize its derivative, we can switch out the temporal derivative of the particle mass distribution function with its finite difference analogon:

$$\frac{\partial n}{\partial t} \to \frac{\Delta n}{\Delta t} \tag{3.2}$$

Note: Throughout this thesis, we will use the notation $\partial n/\partial t$ interchangeably with the finite difference notation from above, since we will basically always be talking about the numerical solution anyways.

Additionally, we have to replace the integral over the particle mass axis with a discrete sum:

$$\int_0^\infty \mathrm{d}m' \to \sum_{i=1}^{\mathcal{N}_m} \Delta m_i \quad \text{and} \quad \int_0^\infty \mathrm{d}m'' \to \sum_{j=1}^{\mathcal{N}_m} \Delta m_j \tag{3.3}$$

Putting all this together, we arrive at the following expression:

$$\frac{\partial n_k}{\partial t} \approx \sum_{i=1}^{\mathcal{N}_m} \sum_{j=1}^{\mathcal{N}_m} K(m_k, m_i, m_j) \cdot n_i \cdot n_j \cdot \Delta m_i \cdot \Delta m_j \tag{3.4}$$

Let us now make use of the relationship $N_i = n_i \cdot \Delta m_i$ declared in equation (1.14) and define the discrete *kernel matrix*

$$\boxed{K_{kij} := K(m_k, m_i, m_j) \cdot \Delta m_k} \tag{3.5}$$

With these two relationships, we are able to rewrite the equation and arrive at the *discretized Smoluchowski coagulation equation*:

$$\boxed{\frac{\partial N_k}{\partial t} \approx \sum_{i=1}^{\mathcal{N}_m} \sum_{j=1}^{\mathcal{N}_m} K_{kij} \cdot N_i \cdot N_j} \tag{3.6}$$

Here, it makes sense to bring attention to the fact that in the discretized Smoluchowski coagulation equation, we do not make use of the notation $n_k = n(m_k)$ for the dust particle mass distribution function.

Instead, the equation is expressed in terms of the number of particles $N_k = N(m_k)$, which labels the number of particles (per unit volume) in a given bin characterized by an index k.

3.2 The Kernel

In the following sections, we will focus on the definition of the kernel functions for both pure coagulation, as well as pure fragmentation. These two functions can then be added later to arrive at the total kernel function.

For both of these cases, we will first consider the definition on the continuous mass axis, and then implement the analogous definition using a mass axis that was discretized according to section 1.3.

Before we do that though, it will make sense to think about how the kernel function can be decomposed and factorized in several different ways, the first of which is detailed in the following section.

3.2.1 Decomposition into Collision Rate and Collision Details

We can factorize the kernel into two separate terms, given by

$$K(m, m', m'') = R(m', m'') \cdot X(m, m', m'') \tag{3.7}$$

The idea behind this distinction is the following:

1. The first term $R(m', m'')$ is the reaction rate coefficient. It is used to encode the information about *how often collision events occur* per unit time, involving a given pair of particles carrying the masses m' and m''.

2. In contrast to that, the second term $X(m, m', m'')$ helps us understand *how this affects the number of particles* carrying a mass value m under any such collisions, i.e. it gives information about how matter is redistributed onto the range of possible masses as the result of a given collision.

For the detailed definition of the first term $R(m', m'')$, see section 2.5.3. The second term $X(m, m', m'')$ will be defined in the process of kernel construction in section 3.3 and section 3.4. All of these definitions will be made separately for the two cases of coagulation and fragmentation.

3.2.2 Decomposition into Coagulation and Fragmentation Kernel

As was just mentioned, we will define two separate kernels to describe the influence of both collisions leading to coagulation as well as fragmentation events, respectively:

$$K^{\text{coag}}(m, m', m'') = R^{\text{coag}}(m', m'') \cdot X^{\text{coag}}(m, m', m'') \tag{3.8}$$
$$K^{\text{frag}}(m, m', m'') = R^{\text{frag}}(m', m'') \cdot X^{\text{frag}}(m, m', m'') \tag{3.9}$$

After having done this, we can add these two sub-kernels to arrive at the total kernel:

$$X(m, m', m'') = X^{\text{coag}}(m, m', m'') + X^{\text{frag}}(m, m', m'') \tag{3.10}$$

For completeness, it should be said that we could add another sub-kernel K^{bounce} to the sum, for collisions leading to particle bouncing events. We might as well neglect this though, since bouncing does not affect the dust particle mass distribution in any way, and therefore we can set

$$K^{\text{bounce}}(m, m', m'') = R^{\text{bounce}}(m', m'') \cdot \underbrace{X^{\text{bounce}}(m, m', m'')}_{=0} \tag{3.11}$$

3.2.3 Decomposition into Gain and Loss Term

We can also decompose the kernel $K(m, m', m'')$ into a negative and a positive contribution, which in the following we will address as the *loss* and *gain* components L and G of the kernel, respectively:

$$K(m, m', m'') = L(m, m', m'') + G(m, m', m'') \tag{3.12}$$

For both sticking and fragmentation, it is possible to establish the following general definition for these two components:

$$L(m, m', m'') := -R(m', m'') \cdot \frac{1}{2}\Big[\delta_D(m - m') + \delta_D(m - m'')\Big] \tag{3.13}$$

$$G(m, m', m'') := +R(m', m'') \cdot f(m, m', m'') \tag{3.14}$$

This definition (or one that is analogous to it) was also used in previous studies of the dust coagulation process, see e.g. [DD04], [BDB10], [OTS09]. It will hopefully become clear in the following sections why it makes sense to adopt this definition, and why it is presented in the form that is shown here.

3.2.4 Definition of the Kernel Mass Error

The coagulation framework that is given by the Smoluchowski equation and is used throughout this work provides an easy method for expressing the mass conservation error of the discretized kernel function. It can be used to get a first look at the stability properties of the algorithm, without actually running the integration scheme. It makes sense to distinguish between a few different error definitions here:

1. Kernel mass error ΔK_{ij} per second (and per density!), given in kg m^3 s^{-1}:

$$\Delta K_{ij} = \sum_{k=1}^{\mathcal{N}_m} m_k \cdot K_{kij} \tag{3.15}$$

2. Kernel mass error ΔX_{ij} per collision, given in kg:

$$\Delta X_{ij} = \frac{\Delta K_{ij}}{R_{ij}^{\text{coll}}} = \sum_{k=1}^{\mathcal{N}_m} m_k \cdot X_{kij} \tag{3.16}$$

3. Relative kernel mass error $\Delta X_{ij}^{\text{rel}}$ per collision (dimensionless):

$$\Delta X_{ij}^{\text{rel}} = \sum_{k=1}^{\mathcal{N}_m} \frac{m_k}{m_i + m_j} \cdot X_{kij} \tag{3.17}$$

4. Total relative kernel mass error ΔX^{rel} (also dimensionless):

$$\boxed{\Delta X^{\text{rel}} = \sqrt{\sum_{i=1}^{\mathcal{N}_m} \sum_{j=1}^{\mathcal{N}_m} \left(\Delta X_{ij}^{\text{rel}}\right)^2}} \tag{3.18}$$

The last definition is what we will use to classify the properties of our kernel implementations regarding the conservation of mass during collisions.

3.3 Definition of the Coagulation Kernel

Consider a collision between two particles with m' and m''. In the case of pure "hit-and-stick" coagulation, we assume that these particles will merge as a result of the collison.

If the coagulation kernel is defined in the right manner, then the integration of the Smoluchowski equation should, for each collision, lead to the removal of one particle from the mass distribution for each of the masses m' and m''. At the same time, a *new* particle has to be inserted into the distribution at $m = m_{\text{tot}} = m' + m''$.

To implement this, we can define the positive and negative components of the kernel in terms of the gain G and loss L as follows:

$$L^{\text{coag}}(m, m', m'') := -R(m', m'') \cdot \frac{1}{2}\Big[\delta_D(m - m') + \delta_D(m - m'')\Big] \tag{3.19}$$

$$G^{\text{coag}}(m, m', m'') := +R(m', m'') \cdot \frac{1}{2} \cdot \delta_D(m - m' - m'') \tag{3.20}$$

This formulation is equivalent to equation (3.13), only that now we defined

$$f^{\text{coag}}(m, m', m'') = \frac{1}{2} \cdot \delta_D(m - m' - m'') \tag{3.21}$$

The reason for the presence of a factor $1/2$ will become clear at the end of this section.

The loss term L assures that, for each collision, exactly one particle is removed from the mass distribution for each of the collision partners' mass value. The gain term on the other hand makes sure that a single particle is inserted into the mass distribution at $m = m' + m''$.

The total coagulation kernel is given by the sum of the two components G and L:

$$K^{\text{coag}}(m, m', m'') = G(m, m', m'') + L(m, m', m'') \tag{3.22}$$

$$= R(m', m'') \cdot \frac{1}{2}\Big[\delta_D(m - m' - m'') - \delta_D(m - m') - \delta_D(m - m'')\Big] \tag{3.23}$$

If we plug this into equation (3.1), we get the continuous Smoluchowski equation for pure "hit-and-stick" coagulation:

$$\frac{\partial n}{\partial t}(m) = \int_0^\infty \int_0^\infty K_{\text{coag}}(m, m', m'') \cdot n(m') \cdot n(m'') \, \mathrm{d}m' \, \mathrm{d}m'' \tag{3.24}$$

$$= \int_0^\infty \int_0^\infty R(m', m'') \cdot \frac{1}{2}\Big[\delta_D(m - m' - m'') - \delta_D(m - m') - \delta_D(m - m'')\Big] \cdot n(m') \cdot n(m'') \, \mathrm{d}m' \, \mathrm{d}m'' \tag{3.25}$$

To make this rather space-consuming equation a bit more manageable, we can pull apart the expression into a separate integral for each of the three sum terms in the kernel.

Then we can make use of the identity for the Dirac δ-distribution we already discussed in equation (1.21) and collapse one of the two integrals over the mass axis.

This leads to

$$\frac{\partial n}{\partial t}(m) = +\frac{1}{2}\int_0^\infty R(m', m-m')\cdot n(m')\cdot n(m-m')\ \mathrm{d}m' \tag{3.26}$$

$$-\frac{1}{2}\int_0^\infty R(m, m')\cdot n(m)\cdot n(m')\ \mathrm{d}m' \tag{3.27}$$

$$-\frac{1}{2}\int_0^\infty R(m, m')\cdot n(m)\cdot n(m')\ \mathrm{d}m' \tag{3.28}$$

Note that here we performed a renaming of $m'' \to m'$ and $\mathrm{d}m'' \to \mathrm{d}m'$, and made use of the symmetry $R(m, m') = R(m', m)$: It is irrelevant whether particle a hits particle b or particle b hits particle a. We regard these two scenarios as the same event.

We can now further simplify the expression combining the last two terms into a single integral, which allows us to get rid of one of the two factors $1/2$:

$$\boxed{\frac{\partial n}{\partial t}(m) = \frac{1}{2}\cdot\int_0^\infty R(m', m-m')\cdot n(m')\cdot n(m-m')\ \mathrm{d}m' - \int_0^\infty R(m, m')\cdot n(m)\cdot n(m')\ \mathrm{d}m'} \tag{3.29}$$

Due to the symmetry $i \leftrightarrow j$, we can get rid of the other factor $1/2$ as well, if we want to. To achieve this, we will rewrite the integral as

$$\frac{1}{2}\int_0^\infty \mathrm{d}m' \to \int_0^{m/2} \mathrm{d}m' \tag{3.30}$$

We will make use of this in the numerical implementation later on, in order to avoid performing any unneeded computations.

3.3.1 Discretization of the Coagulation Kernel

Having defined the coagulation kernel on a continuous mass axis, we now wish to implement the numerical analogon of the described model to allow numerical simulations of the dust mass distribution's temporal evolution.

Even in a highly simplified scenario, where only hit-and-stick coagulation is included, the definition of the kernel K_{kij} is not at all trivial. To assure both the stability and the accuracy of the algorithm, one has to take care of two separate problems, namely:

1. The conservation of mass *has* to be assured, otherwise the numerical solution can not be assumed to remain stable for long. In the case of hit-and-stick coagulation, this means that for every pair of colliding particles, a single new particle has to be created. At the same time, the two initial particles have to be removed from the distribution. During this process, the total mass should remain unaffected down to machine precision.

2. When using a logarithmically spaced grid for the discretized mass axis, it can not be assumed that after a collision of two dust particles with masses m_i and m_j the resulting particle will carry a mass $m_k = m_i + m_j$ whose value can be mapped trivially onto the grid. In general, the corresponding index will not be an integer, and instead lie between somewhere between the two neighboring grid points with indices k and $k+1$. Therefore, the result of the merging of m_i and m_j has to be divided in some sensible way between these two neighboring bins.

In the following section, we will outline a method for assuring these conditions.

3.3.2 The Kovetz-Olund Algorithm

An elegant way for solving the two problems listed above is given in the 1969 paper by Kovetz & Olund [KO69], where they used the following procedure:

1. The hit-and-stick coagulation kernel is split into two parts, as we did above. The first is the *gain* of particles in bin k due to the collision of particles from the bins i and j. The second is the *loss* of particles from bin k due to collisions of particles in bin k with particles from any other bin j. Using this separation into gain & loss terms (as well as the discrete formalism for the Smoluchowski equation), the dust particle mass distribution's temporal derivative can be expressed in the following form:

$$\frac{\partial N_k}{\partial t} = \sum_{i=0}^{\mathcal{N}_m} \sum_{j=0}^{\mathcal{N}_m} K_{kij}^{\text{gain}} \cdot N_i \cdot N_j - \sum_{j=0}^{\mathcal{N}_m} K_{kj}^{\text{loss}} \cdot N_k \cdot N_j \tag{3.31}$$

   In other words, the total kernel can be written as

$$K_{kij} = K_{kij}^{\text{gain}} - K_{ij}^{\text{loss}} \cdot \delta_{ki} \tag{3.32}$$

   Splitting the kernel like this into a gain & a loss term is a quite general approach, and can be used in more complex scenarios as well (including e.g. particle fragmentation processes). In our case, here the loss component is given by the coagulation rate coefficent R_{ij}^{coag}, and the gain component will shortly be defined from it.

2. For the scenario of pure hit-and-stick coagulation, a unique discretization of the kernel can be defined such that both the number of particles and the conservation of total mass are handled correctly. To do this, consider a pair of colliding particles with indices i and j. Then, let the index $\bar{k}$ be chosen in such a way that the condition

$$m_{\bar{k}} \leq m_i + m_j < m_{\bar{k}+1} \tag{3.33}$$

   is satisfied.

3. As stated before, in hit-and-stick coagulation, a single new particle emerges for each pair of colliding particles. Using the definitions from above, this condition can be expressed as follows:

$$K_{\bar{k},ij}^{\text{gain}} + K_{\bar{k}+1,ij}^{\text{gain}} \stackrel{!}{=} K_{ij}^{\text{loss}} \tag{3.34}$$

4. The second condition is that of mass conservation, which can be written as:

$$m_{\bar{k}} \cdot K_{\bar{k},ij}^{\text{gain}} + m_{\bar{k}+1} \cdot K_{\bar{k}+1,ij}^{\text{gain}} \stackrel{!}{=} (m_i + m_j) \cdot K_{ij}^{\text{loss}} \tag{3.35}$$

5. Now, in order to map the resulting particle's mass onto the two neighboring bins, let us define a parameter ε such that

$$K_{\bar{k},ij}^{\text{gain}} = K_{ij}^{\text{loss}} \cdot (1 - \varepsilon), \text{ and} \tag{3.36}$$

$$K_{\bar{k}+1,ij}^{\text{gain}} = K_{ij}^{\text{loss}} \cdot \varepsilon \tag{3.37}$$

   This assures that equation (3.34) is satisfied. If we now plug these definitions into equation (3.35) and solve for ε, we arrive at

$$\varepsilon := \frac{m_i + m_j - m_{\bar{k}}}{m_{\bar{k}+1} - m_{\bar{k}}} \tag{3.38}$$

This is the Kovetz-Olund algorithm [KO69], which is frequently used in studies of dust coagulation in proto-planetary disks, see e.g. [BDH08] and [BDB10].

3.3.3 Near-Zero-Cancellation Handling

When using floating-point numbers following the representation defined by the IEEE-754 standard, it can occur that

$$a + b = a \qquad \text{for} \qquad b \neq 0 \tag{3.39}$$

Typically, this happens when

$$|b| < \varepsilon_m \cdot |a| \tag{3.40}$$

Here, ε_m labels the *machine precision*, which can be used to give an upper bound on the relative approximation error that may happen in floating point arithmetic due to rounding. For 64-bit floats, this value is approximately given by

$$\varepsilon_{f64} \approx 10^{-16} \tag{3.41}$$

Let i and j once again be the indices used to label two colliding particles. Without loss of generalization, let us assume that particle i is *much smaller* than particle j.

The detailed balance approach from above requires the removal of both the big and the small particle from the mass distribution, followed by the re-insertion of a new particle carrying the initial pair's combined mass. This new particle would then have a mass which is nearly identical to that of the bigger one of the original two particles, it would be only a tiny bit heavier.

In the approach defined above this would mean that $\bar{k} = j$, i.e. the resulting particle will reside in the same bin as the larger original one. Also, it would follow that $\varepsilon \ll 1$.

Let us now take a look at the particle mass distribution in the bin $\bar{k}$ and, more specifically, by how much it changes from one timestep to the next. For this particular pair of i and $\bar{k} = j$, we can write:

$$\frac{\partial n_{\bar{k}}}{\partial t} = K^{\text{gain}}_{\bar{k},i\bar{k}} \cdot n_i \cdot n_{\bar{k}} - K^{\text{loss}}_{\bar{k}i} \cdot n_i \cdot n_{\bar{k}} \tag{3.42}$$

Plugging in equation (3.36) leads to

$$\frac{\partial n_{\bar{k}}}{\partial t} = (1-\varepsilon) K^{\text{loss}}_{\bar{k}i} \cdot n_i \cdot n_{\bar{k}} - K^{\text{loss}}_{\bar{k}i} \cdot n_i \cdot n_{\bar{k}} \tag{3.43}$$

Here, the two terms almost cancel each other out. What remains is a contribution which is proportional to ε. But if ε is small enough, then the double-precision accuracy of the 64-bit floating point representation will lead to a breakdown of the algorithm.

It is relatively easy though to identify the particle pairs (i, j) for which the scenario detailed above will occur. If i is the index of the larger one of the two colliding masses, then cancellation may occur when the resulting k is equal to j.

In that case, we analytically carry out the subtraction in equation (3.43), and write:

$$\frac{\partial n_{\bar{k}}}{\partial t} = -\varepsilon K^{\text{loss}}_{\bar{k}i} \cdot n_i \cdot n_{\bar{k}} \tag{3.44}$$

The effect that this correction has on the kernel mass error $\Delta X^{\text{coag}}_{ij}$ can be seen visualized in figure 3.8.

3.4 Definition of the Fragmentation Kernel

Let us now focus on the definition of the kernel matrix for dust particle fragmentation processes. As before, where we considered the simpler case of pure hit-and-stick coagulation, in the case of fragmentation we also separate the kernel function into a positive gain component G and a negative loss component L.

These two kernel components can be expressed as

$$L^{\text{coag}}(m, m', m'') := -R(m', m'') \cdot \frac{1}{2}\Big[\delta_D(m - m') + \delta_D(m - m'')\Big] \tag{3.45}$$

$$G^{\text{coag}}(m, m', m'') := +R(m', m'') \cdot \frac{1}{2} \cdot f^{\text{frag}}(m, m', m'') \tag{3.46}$$

Here, it should be noted that the loss term is exactly identical to what we defined for pure hit-and-stick coagulation in equation (3.19). If two particles with masses m and m' collide, then these two particles "disappear" from the distribution.

Replacing them, new particles will then have to be added to be distribution, under the strict criterion of mass conservation. How this is implemented can be controlled via the definition of the gain term or, more specifically, via the definition of the function $f^{\text{frag}}(m, m', m'')$. The big distinction between the kernels for coagulation and fragmentation lies in the way that this function is defined.

In the scenario of pure hit-and-stick coagulation, it is given by a Dirac δ-function, which allows us to reduce the double integral in the Smoluchowski equation (or double sum, in the discrete case) into an integral/sum over a single axis. That is very different here, and a reduction of the double integral will not be possible. This is why the numerical cost of modeling fragmentation processes is significantly higher than if one were to only consider pure coagulation.

The reasoning behind the different definitions of the functions f^{coag} and f^{frag} is the following: If two particles merge into one, then (without breaking mass conservation) there is only a single possible mass value that can result from that.

In the case of fragmentation though, this is not true. If two particles collide, they could in principle fragment into a whole range of differently-sized particles, as long as mass is conserved.

In the following, we will attempt to define the function f^{frag} in a sensible fashion. As we did before in other contexts, here as well we will adopt a shorthand index notation, defined by

$$f^{\text{frag}}_{kij} := f^{\text{frag}}(m_k, m_i, m_j) \tag{3.47}$$

There are of course various different approaches for how one might do this.

3.4.1 Modeling the Resulting Mass Distribution

An extremely simplified first approach for modeling the distribution of dust particle masses resulting from a collision event with subsequent particle fragmentation could be derived from the naive assumption that, as a result of a collision, the entire mass involved in the collision is transferred into a single bin with index k.

This is of course not a very sophisticated definition, but was used as a first approach in the process of building and testing the model for this thesis in computer code.

An implementation of this simple approach could be achieved via the definition of

$$f_{kij}^{\text{frag}} = \frac{m_i + m_j}{m_k} \tag{3.48}$$

If, for example, we assume that the entire mass involved in the collision is to be transferred to the lowest-mass bin, i.e. the entire mass is converted into tiny particles, we would simply have to set $k = 1$ (the index of the first bin in the mass grid). This might be imagined as "complete pulverization" of the colliding particles.

3.4.2 The MRN Distribution

A different method, which presents a bit more sophisticated approach, was given by Mathis, Rumpl, and Nordsieck in their 1977 paper [MRN77].

There, the distribution of masses resulting from a collision of two dust particles is modeled making use of an inverse power law with exponent $q := -11/6$.

We will implement it as follows:

$$f_{kij}^{\text{frag}} = \frac{m_i + m_j}{m_k} \cdot \xi_{kij} \tag{3.49}$$

Here, ξ_{kij} labels the fraction of the total mass $m_{\text{tot}} = m_i + m_j$ that is transferred into a given bin k as a result of the collision. This fraction shall be defined as

$$\xi_{kij} := \frac{m_k^q}{S} \tag{3.50}$$

with the normalization factor

$$S := \sum_{k_{\text{min}}}^{k_{\text{max}}} m_k^q \tag{3.51}$$

Here, we define $k_{\text{min}} = 1$, and set k_{max} to the index corresponding to $m_{\text{tot}} = m_i + m_j$. Note that this is a somewhat arbitrary definition, and one could also use $\max(m_i, m_j)$ for the upper mass boundary instead.

3.5 Visualization of Kernel Matrices

On the following pages, we display a few visualizations of the kernel matrices we just defined, plus a few short explanations for why they look the way that they do.

3.5.1 Coagulation Kernel on a Linear Mass Axis

To get started, we will first consider the definition of the kernel matrix for pure hit-and-stick coagulation, defined on a mass axis using linear scaling. Also, for the time being the reaction rate coefficent is set to 1. This is probably the simplest case one might think of in this context.

The kernel matrix is shown in figure 3.1 for various values of k, the value of which we use to index the different particle masses that could result from a given collision. The negative component L_{kij}^{coag} and the positive component G_{kij}^{coag} are drawn in different colors, namely blue and red, respectively.

Let us try to understand why the plot looks the way it does:

Observation 1: The displayed matrix is symmetric along the diagonal, meaning it stays invariant under transformations $i \to j$ and $j \to i$. This is due to the fact that we do not care whether a particle a collides with a particle b, or a particle b collides with a particle a. These two scenarios describe the same event.

Observation 2: The kernel loss (shown in blue) is visualized as two straight lines. For the moment, let us focus on the vertical one: If a collision involving a particle with mass m_k occurs, then we will have to remove one particle from the distribution at that bin k. This explains the vertical line's horizontal position at $j = k$. In addition to the particle characterized by index k, such a collision could involve any other particle mass, which explains why the loss component looks like a line, instead of a point. An analogous reasoning can be made for the horizontal line.

Observation 3: Even though we just mentioned that a given particle could collide with particles from *any other* mass bin, this is only true in principle. As can be seen in the plot, the blue lines do not end at the edges of the kernel matrix, but much earlier instead. This is due to the fact that we have to ensure the conservation of mass over time. Consider two particles with masses m_i and m_j. If their masses do not satisfy the conditions $m_i + m_j < m_{\max}$, then the collision and subsequent merging will require the creation of a particle that can not be represented in the mass grid, which would effectively lead to a loss of mass. In reality, mass can of course move out of the considered range, but we will not allow it to do that in order to be able to better assert the algorithms mass conservation properties. As such, we can expect a "traffic jam" of dust mass accumulating in the highest-mass bin.

Observation 4: The positive component (shown in red) follows a diagonal line. This represents the influence of the mass conservation criterion for newly created particles, which must satisfy the condition $m_{\text{tot}} = m_i + m_j$. Therefore, $m_i = m_{\text{tot}} - m_j$, which can be interpreted as a linear function of the kind "$y = c - x$", and leads to the representation as a diagonal downward-sloped line.

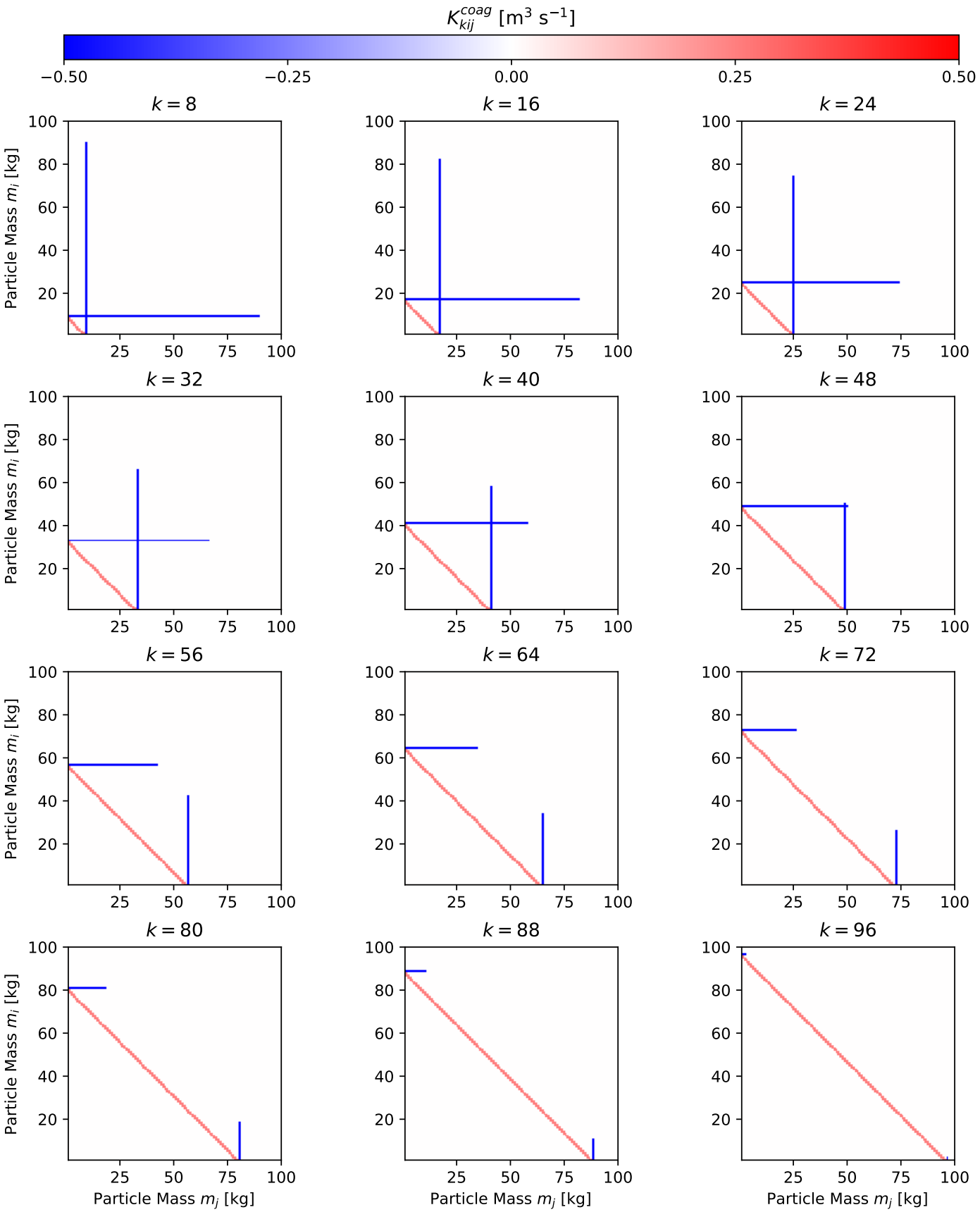


Figure 3.1: Matrix layers of the kernel for pure hit-and-stick coagulation on a linear mass axis with lower mass grid boundary $m_{\min} = 10^{-6}$ kg, upper mass grid boundary $m_{\min} = 10^{-2}$ kg, mass grid resolution $\mathcal{N}_m = 100$, and collision rate coefficent $R_{ij}^{\text{coll}} := 1$ m^3 s^{-1}.

3.5.2 Coagulation Kernel on a Logarithmic Mass Axis

Next, we will define the coagulation kernel again, but making use of a logarithmic scaling for the discretized mass axis instead.

The results can be seen visualized in figure 3.2 and figure 3.3. Note that due to the logarithmic representation, here we need to plot gain and loss components separately.

Observation 1: The negative "loss" component (shown in blue) is quite similar to what we already saw in figure 3.1 for the case of a linearly discretized mass axis. A difference is given by the extension of the blue lines, which reach much further towards the edge of the kernel matrix than they did before. This can be explained by the logarithmic scaling.

Observation 2: In contrast to the linear case, here the positive "gain" component (shown in red) is not represented as a linear diagonal, but an inverse power relation. This can be explained by the logarithmic scaling as well.

Observation 3: The majority of the kernel's entries carry a value of zero, with only about 3% of the entries being non-zero. (Here, a mass grid resolution of $\mathcal{N}_m$ was used.) This will be different if fragmentation is included.

3.5.3 Fragmentation Kernel on a Logarithmic Mass Axis

The positive and negative components of the kernel matrix for modeling pure dust particle fragmentation can be seen in figure 3.4 and figure 3.5, respectively.

Observation 1: The negative "loss" component of the fragmentation kernel looks quite similar to what we saw in the case of the coagulation kernel on a logarithmically discretized mass axis. A difference can be noticed at the edges of the kernel though. While the collisions of high-mass particles had to be ignored in the case of hit-and-stick coagulation, we can include these collisons here. This can be explained by the fact that the particles resulting from a collision+fragmentation event are *always* smaller (and thus less massive) than the original colliding particles. As such, no mass loss can occur here due to a newly created particle's mass not being represented in the mass grid.

Observation 2: The number of non-zero kernel entries is much, much higher here than it was in the case of stick-and-hit coagulation. Here, about 30% of the kernel's entries carry non-zero values. This can be explained by the fact that a merge event always leads to the creation of exactly one new particle from the two original ones, whereas this is not the case in fragmentation. There, possibly very many new particles could be created as a result of a collision, and the mass distribution of newly created particles is modeled via the inverse power law given by the MRN distribution (for details see section 3.4.2).

Observation 3: Another thing one might notice is the existence of a diagonal feature, for the kernel entries with $i = j$. In the plot, it can be seen as a bright line going from top right to top left. This is the same feature we could already see in the relative velocity plot in figure 2.5, as well as the plot of the collision rate coefficent in figure 2.7b. The feature can be explained by the fact that a pair of particles (i, j) has a lower chance of colliding, if $i = j$. Since same-sized particles also experience similar kinematics, here the relative velocity is close to zero, which leads to a low collision rate, and thus smaller values of the kernel entries.

3.5.4 Total Kernel on a Logarithmic Mass Axis

Now, we add the two kernel matrices K_{kij}^{frag} and K_{kij}^{coag} to construct the total kernel, for the full model including both associative and dissociative processes.

Its gain and loss components can be seen visualized in figure 3.6 and figure 3.7, respectively.

3.5.5 Kernel Mass Error

In figure 3.8 and figure 3.9 we display the dimensionless kernel error, i.e. the relative mass deviation resulting from a single collision, for various different scenarios.

First, figure 3.8 shows the influence of near-zero cancellation handling on the error of the kernel for pure coagulation (for implementation see section 3.3.3).

As the visualization shows, the mass error per collision of the coagulation kernel can actually be decreased significantly via the inclusion of a procedure for handling near-zero cancellation in floating point arithmetics.

An additional error is introduced though by the inclusion of fragmentation, as can be seen in figure 3.9a. Here, it is not possible to use the same procedure to get rid of this error, as no obvious cases of cancellation occur.

The error of the total kernel is then shown in figure 3.9b.

All of these (dimensionless/relative) errors could be kept on the order of 10^{-12}.

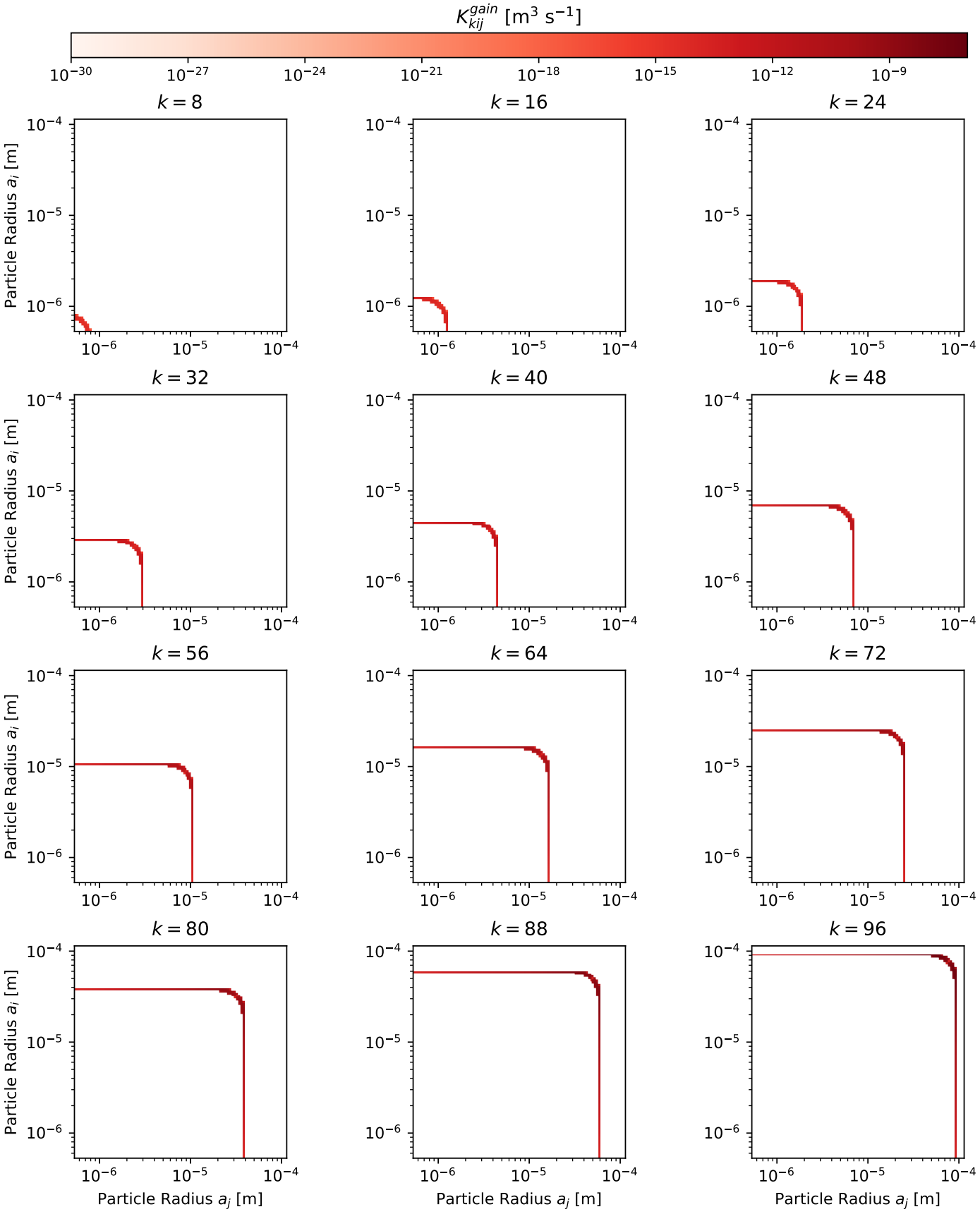


Figure 3.2: Matrix layers of the positive kernel component for pure hit-and-stick coagulation on a logarithmic mass axis, with lower mass grid boundary $m_{\min} = 10^{-15}$ kg, upper mass grid boundary $m_{\min} = 10^{-8}$ kg, and mass grid resolution $\mathcal{N}_m = 100$.

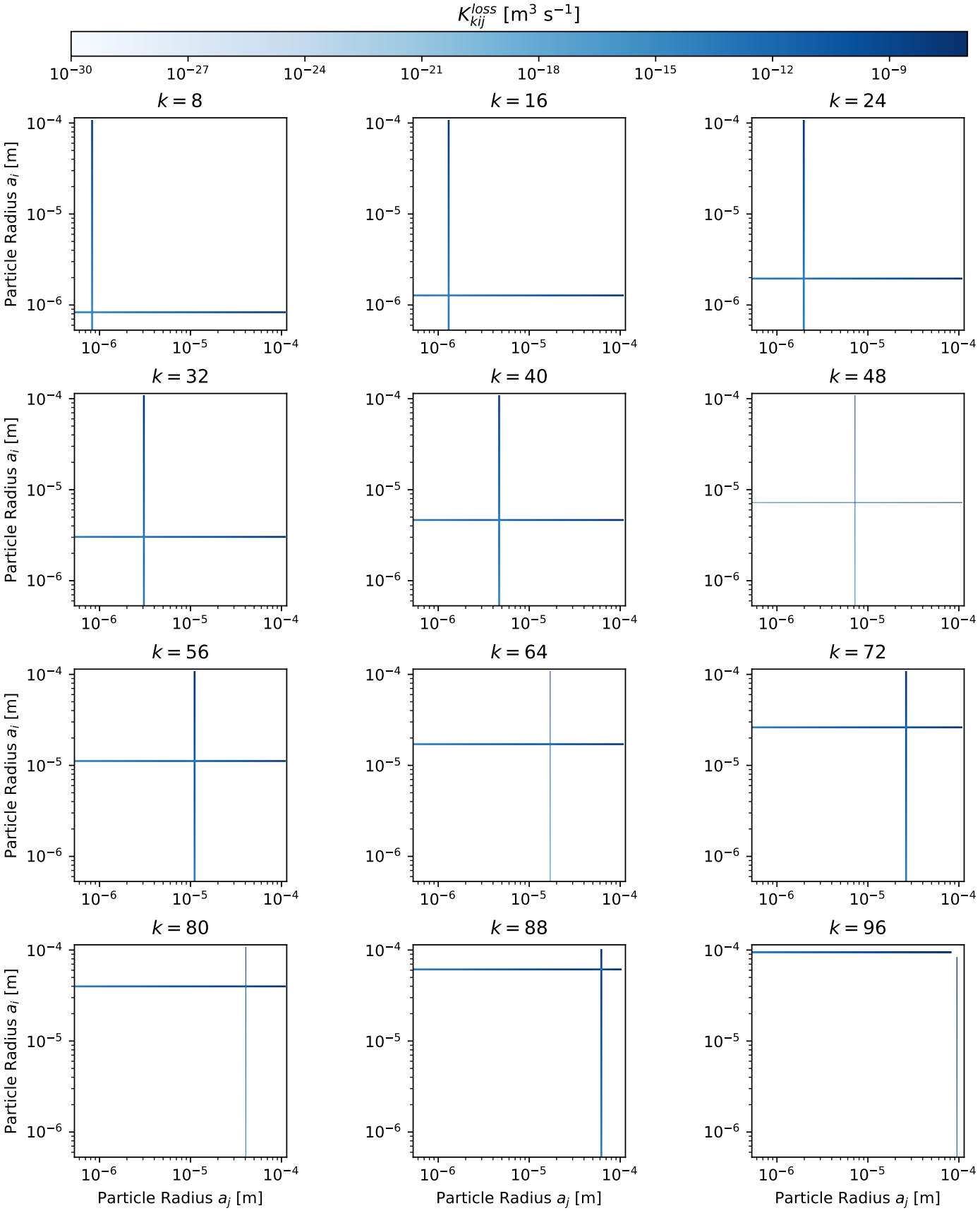


Figure 3.3: Matrix layers of the negative kernel component for pure hit-and-stick coagulation on a logarithmic mass axis, with lower mass grid boundary $m_{\min} = 10^{-15}$ kg, upper mass grid boundary $m_{\min} = 10^{-8}$ kg, and mass grid resolution $\mathcal{N}_m = 100$.

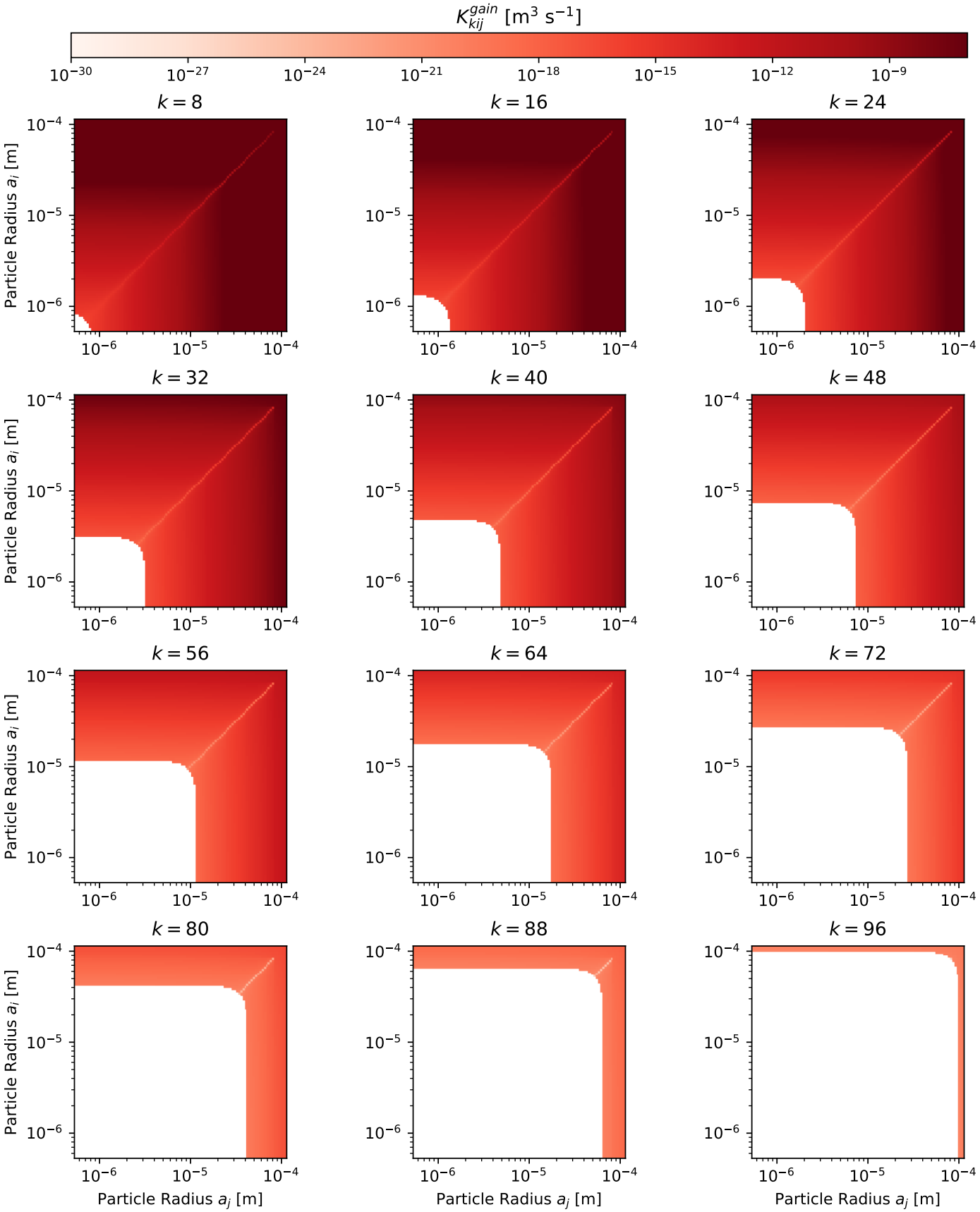


Figure 3.4: Matrix layers of the positive kernel component for pure fragmentation on a logarithmic mass axis, with lower mass grid boundary $m_{\min} = 10^{-15}$ kg, upper mass grid boundary $m_{\min} = 10^{-8}$ kg, and mass grid resolution $\mathcal{N}_m = 100$.

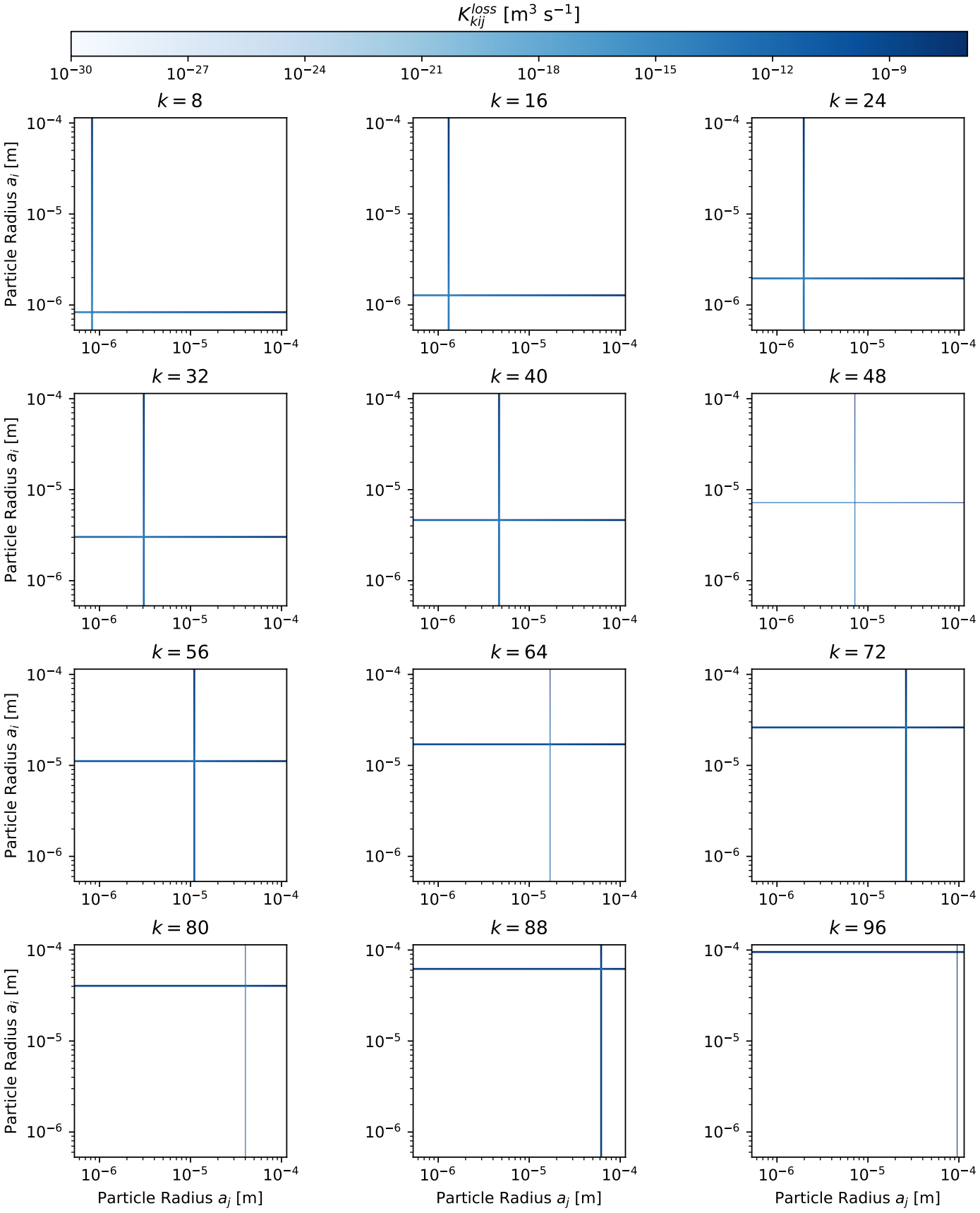


Figure 3.5: Matrix layers of the negative kernel component for pure fragmentation on a logarithmic mass axis, with lower mass grid boundary $m_{\min} = 10^{-15}$ kg, upper mass grid boundary $m_{\min} = 10^{-8}$ kg, and mass grid resolution $\mathcal{N}_m = 100$.

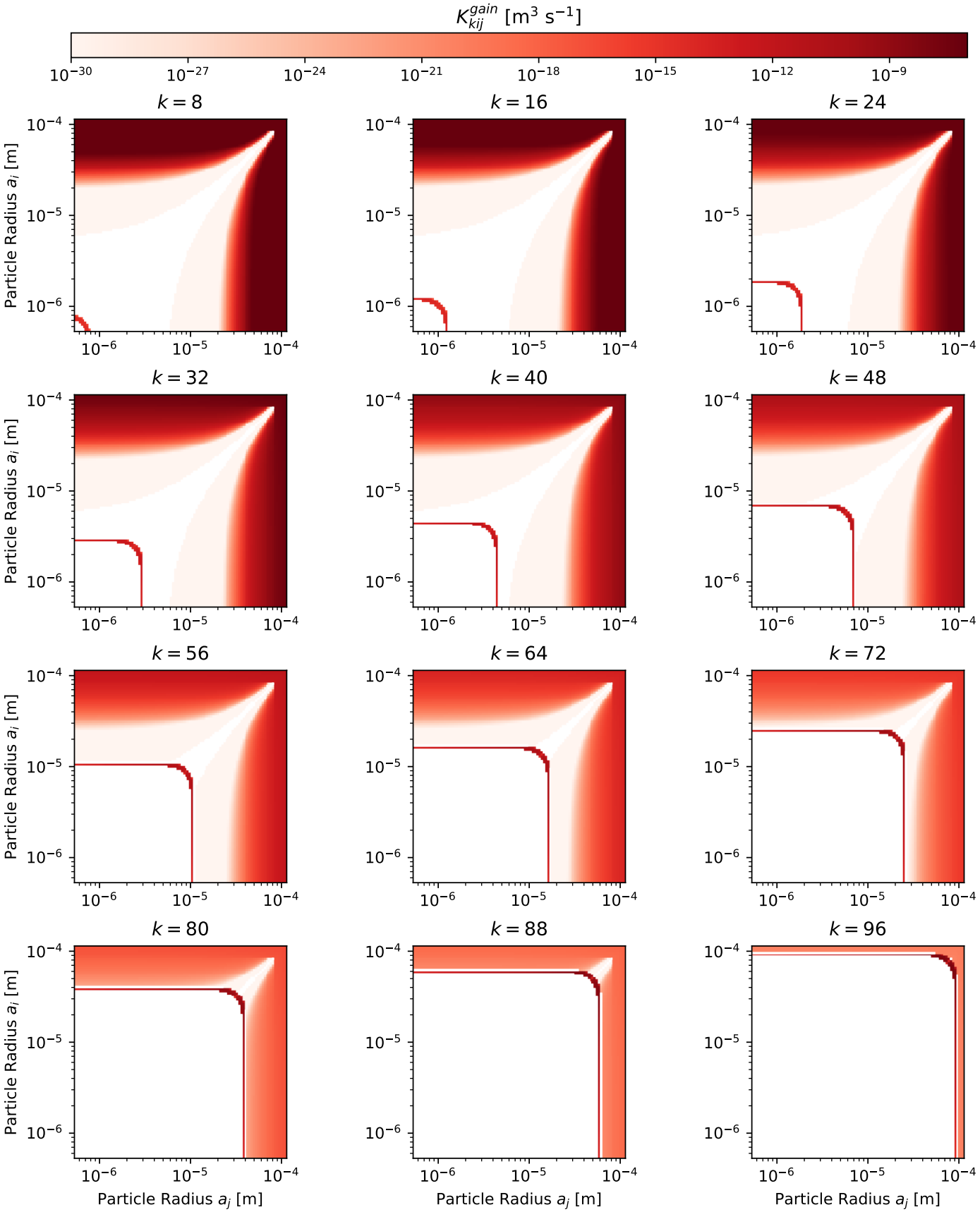


Figure 3.6: Matrix layers of the positive kernel component for both coagulation and fragmentation on a logarithmic mass axis, with lower mass grid boundary $m_{\min} = 10^{-15}$ kg, upper mass grid boundary $m_{\min} = 10^{-8}$ kg, and mass grid resolution $\mathcal{N}_m = 100$.

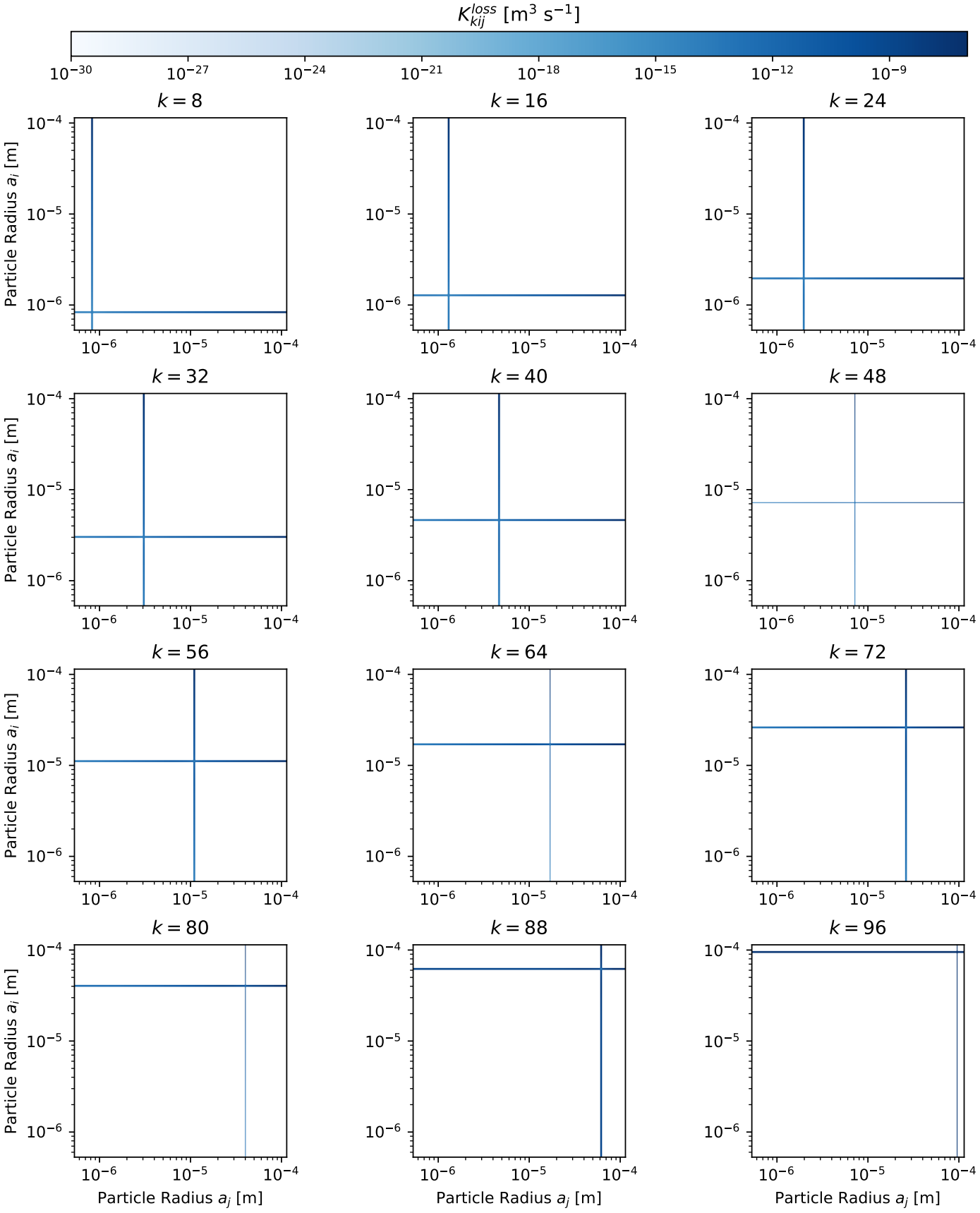
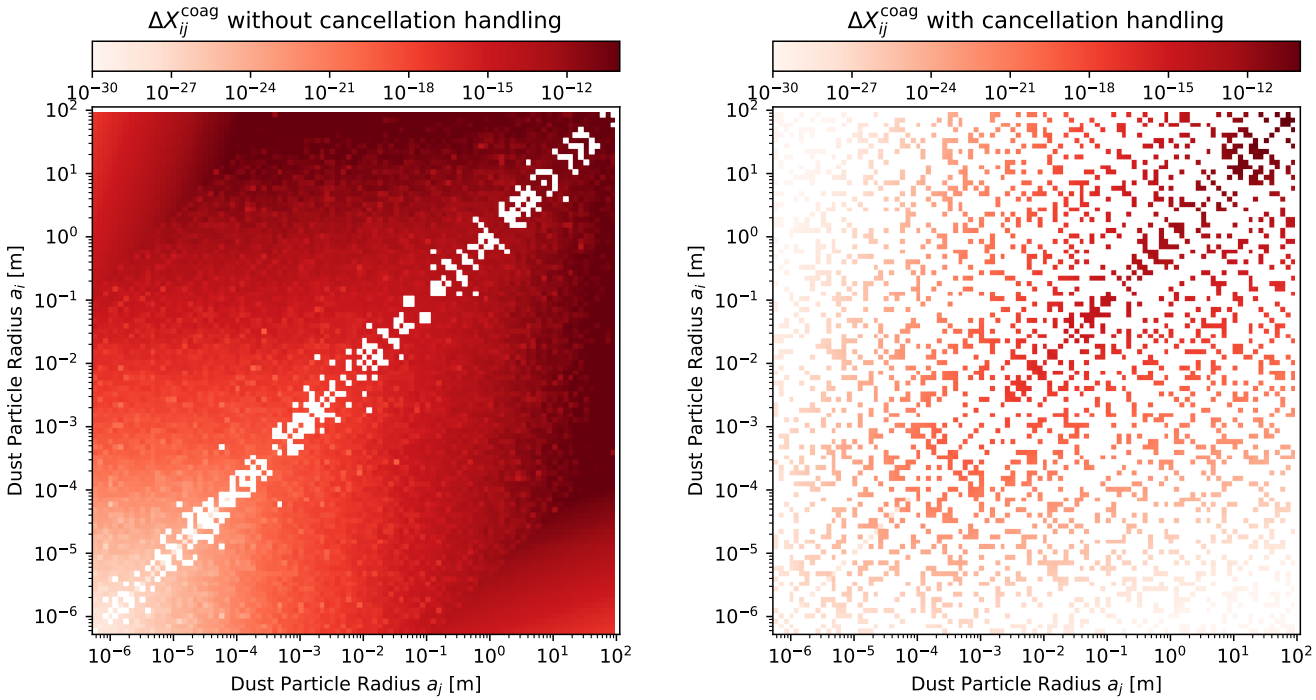


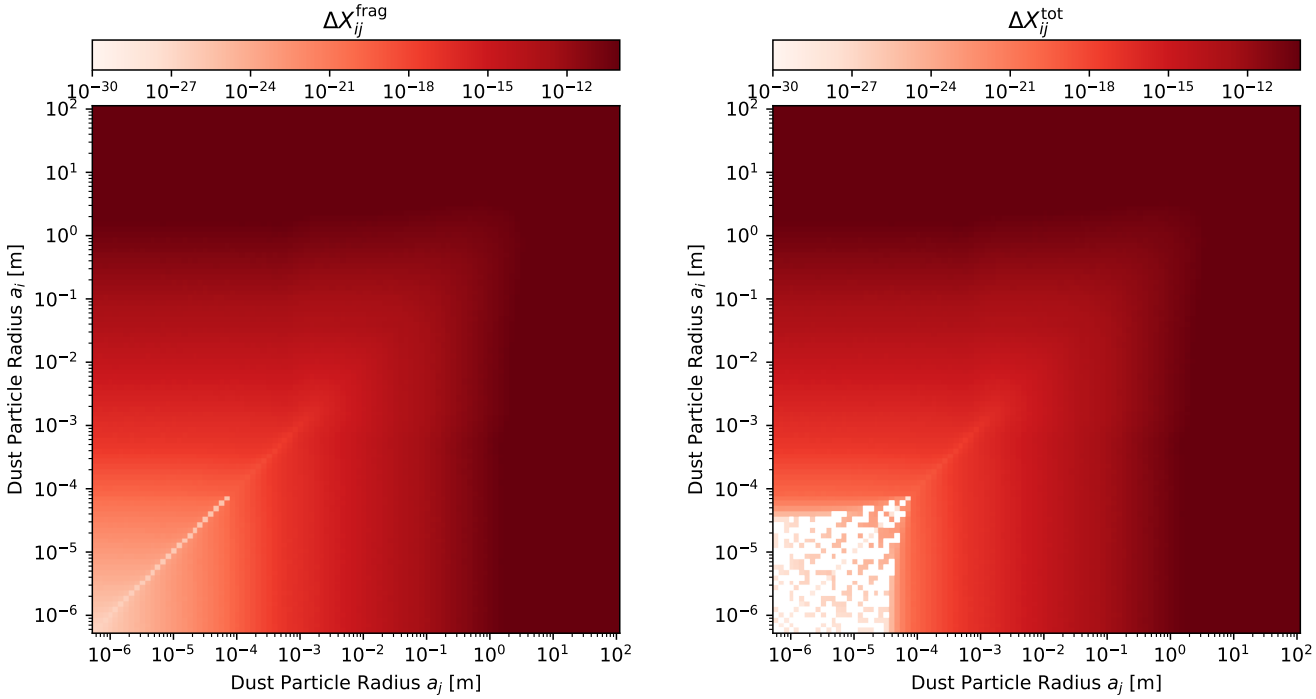
Figure 3.7: Matrix layers of the negative kernel component for both coagulation and fragmentation on a logarithmic mass axis, with lower mass grid boundary $m_{\min} = 10^{-15}$ kg, upper mass grid boundary $m_{\min} = 10^{-8}$ kg, and mass grid resolution $\mathcal{N}_m = 100$.



(a) Coagulation kernel without cancellation handling

(b) Coagulation kernel with cancellation handling

Figure 3.8: Influence of near-zero cancellation on the mass error of the coagulation kernel.



(a) Fragmentation kernel

(b) "Full" kernel: coagulation + fragmentation

Figure 3.9: Dimensionless mass error ΔX_{ij} for (a) Fragmentation kernel, and (b) Total kernel.

3.6 Analytical Solutions

Analytical solutions to the Smoluchowski coagulation equation exist if, and only if the kernel takes on one of the following three forms:

1. Constant kernel:
$$K(m, m', m'') = A_1 \tag{3.52}$$
2. Linear kernel:
$$K(m, m', m'') = A_2 \cdot (m' + m'') \tag{3.53}$$
3. Quadratic kernel:
$$K(m, m', m'') = A_3 \cdot (m' \cdot m'') \tag{3.54}$$

Here, the three numbers A_1, A_2, and A_3 are constants. The analytical solutions in these cases are discussed in detail in the 1990 paper written by K. Ohtsuki [ONN90], as well as the 2006 paper by J. Wattis [Wat06].

3.6.1 Constant Kernel

When making use of the constant kernel, the Smoluchowski reads

$$\frac{\partial n}{\partial t}(m) = \int_0^\infty \int_0^\infty n(m') \cdot n(m'') \, \mathrm{d}m' \, \mathrm{d}m'' \tag{3.55}$$

With the initial condition given by

$$n_1(t = 0) = n_0 \qquad \text{and} \tag{3.56}$$
$$n_k(t = 0) = 0 \qquad \text{for } k > 1 \tag{3.57}$$

the solution to equation (3.55) is given by [Smo16] [MTH17]

$$n_k(\tau) = n_0 \cdot g^2 \cdot (1 - g)^{k-1} \tag{3.58}$$

with

$$g = (1 + \tau/2)^{-1} \tag{3.59}$$

where the normalized time is given by

$$\tau = \frac{A_1 \cdot \rho \cdot t}{m_{\min}} \tag{3.60}$$

3.6.2 Linear Kernel

The Smoluchowski equation with a linear kernel can be written as

$$\frac{\partial n}{\partial t}(m) = \int_0^\infty \int_0^\infty (m' + m'') \cdot n(m') \cdot n(m'') \, \mathrm{d}m' \, \mathrm{d}m'' \tag{3.61}$$

Let $\overline{m}_0$ label the average dust particle mass at time $t = 0$. With the initial condition

$$n(m, t = 0) = C \cdot e^{-\lambda m} \tag{3.62}$$

where

$$C = \frac{N_0}{\overline{m}_0} \qquad \text{and} \qquad \lambda = \frac{1}{\overline{m}_0} \tag{3.63}$$

the solution to equation (3.61) according to [Saf63], [Saf72] is given by

$$n(m, \tau) = \frac{N_0 \cdot g}{m \cdot \sqrt{1-g}} \cdot e^{-\lambda \cdot m(2-g)} \cdot I_1 \cdot 2\lambda m \sqrt{1-g} \tag{3.64}$$

with

$$g = e^{-\tau} \tag{3.65}$$

where I_1 is the modified Bessel function and $\tau = A_2 \rho_g t$ is the normalized time.

3.6.3 Quadratic Kernel

For the quadratic kernel, the Smoluchowski equation reads

$$\frac{\partial n}{\partial t}(m) = \int_0^\infty \int_0^\infty (m' \cdot m'') \cdot n(m') \cdot n(m'') \, \mathrm{d}m' \, \mathrm{d}m'' \tag{3.66}$$

A solution to this equation is derived by in the work by [Wat06], where they give it as

$$n_k(t) \sim \begin{cases} \frac{1}{\sqrt{2\pi} \cdot k^{5/2} \cdot t} \cdot \exp[-k(t - 1 - \log(t))] & \text{for } t < t_c \\ \frac{1}{\sqrt{2\pi} \cdot k^{5/2} \cdot t} & \text{for } t \geq t_c \end{cases} \tag{3.67}$$

Here, t_c labels a critical time value, at which the behavior of the evolution changes. In [Wat06], it is be shown that this is actually the *gelation time*.

3.7 Numerical Integration

Let us now take a look at how to use numerical integration of the discretized Smoluchowski coagulation equation to arrive at an approximate solution for the temporal evolution of the particle mass distribution function under the influence of dust particle collisions.

3.7.1 Definition of the Integration Mass Error

To be able to evaluate a model implementation with regards to its stability properties, it makes sense to define a measure for the total mass error of the integration at a certain point in time.

Let t_i and t_f label the time at the start and end of the integration, respectively. The mass volume density's value at these two points in time can be written as:

$$\rho_{\mathrm{i}}^d := \sum_{k=0}^{\mathcal{N}_m} m_k \cdot n_k(t = t_{\mathrm{i}}) \cdot \Delta m_k \tag{3.68}$$

$$\rho_{\mathrm{f}}^d := \sum_{k=0}^{\mathcal{N}_m} m_k \cdot n_k(t = t_{\mathrm{f}}) \cdot \Delta m_k \tag{3.69}$$

The total mass arror accumulated over the course of the integration is then given by

$$\Delta_{\mathrm{stab}} = \frac{\rho_f^d - \rho_i^d}{\rho_i^d} \tag{3.70}$$

In reality, this mass error should be exactly equal to zero. Otherwise, mass conservation would not be given, since the movement of mass from one point in space to another is not included in our model. The numerical integration introduces an error though, which ideally should be kept at machine precision.

3.7.2 Explicit Euler Integration

A naive approach for solving the Smoluchowski equation numerically is given by the simple explicit Euler integration scheme, which we will briefly address here. Assuming a constant time-step size $\Delta t \in \mathbb{R}_+$, the elapsed time after $m \in \mathbb{N}_0$ integration steps can be expressed as

$$t_m = t_0 + m \cdot \Delta t, \tag{3.71}$$

where t_0 is the time value at the start of the integration.

Now, the goal is to find the corresponding values n_k^m for the particle mass distribution:

$$n_k^m := n(m = m_k,\ t = t_m) \tag{3.72}$$

We define the differences

$$\Delta n_k = n_k^m - n_k^{m-1} \quad \text{and} \tag{3.73}$$

$$\Delta t = t_m - t_{m-1} \tag{3.74}$$

which, in turn, can be used to define the 1st order forward finite difference

$$\mathrm{FFD} = \frac{\Delta n_k}{\Delta t} \tag{3.75}$$

The finite difference can now be used as an approximation for the temporal derivative of n:

$$\frac{\partial n_k}{\partial t} \approx \text{FFD} \tag{3.76}$$

Plugging in equation (3.4) and rearranging for Δn_k leads to

$$\Delta n_k = \Delta t \cdot \sum_{i=1}^{\mathcal{N}_m} \sum_{j=1}^{\mathcal{N}_m} K_{kij} \cdot n_i \cdot n_j \tag{3.77}$$

This can now be used an approximate value for the change in n_k from one time-step to the next, i.e.

$$\begin{aligned} n_k^m &= n_k^{m-1} + \Delta n_k \\ &= n_k^{m-1} + \Delta t \cdot \sum_{i=1}^{\mathcal{N}_m} \sum_{j=1}^{\mathcal{N}_m} K_{kij} \cdot n_i \cdot n_j \end{aligned} \tag{3.78}$$

In order to ensure stability of the numerical integration, the step size needs to respect the Courant-Friedrichs-Lewy (CFL) criterion, which provides an upper limit for the values that Δt can take on [CFL28].

A discussion regarding the formulation of the CFL criterion in the context of the Smoluchowski coagulation equation is given by e.g. [LL21].

3.7.3 Implicit Radau Integration

Since we would like to study the temporal evolution of the dust particle mass distribution with particle masses ranging over multiple orders of magnitude, it makes sense to use a logarithmic scaling for the mass axis discretization.

Similarly, because we want to follow the mass distribution's temporal evolution over very long timespans, the temporal axis will be discretized using a logarithmic scaling as well.

This means that the difference in time Δt from a step to the next will not stay constant over time, and instead satisfies the relationship

$$\Delta t \sim t \tag{3.79}$$

Due to this, we must be careful not to disrespect the CFL criterion, which places an upper limit on the step size Δt. Choosing a larger value would lead to a breakdown of the algorithm, since stability would no longer be assured.

To deal with that problem, it makes sense to make use of an implicit integration scheme.

For the studies made in the context of this thesis, here we chose to utilize a 4th-order Radau integration scheme. This scheme belongs to a class of implicit Runge-Kutta methods that can be used for solving initial value problems.

An implementation of this algorithm is given in the ‘scipy‘ Python library. More specifically, we make use of the ‘scipy.integrate.solve_ivp‘ function.

3.8 Visualization of Integration Results

Having defined the disk model, the kernel matrix, and the integration scheme, now it is the time to integrate the Smoluchowski coagulation equation.

This is done for several different cases, namely:

1. Pure coagulation (on a lin. mass axis)
2. Pure coagulation (on a log. mass axis)
3. Pure fragmentation (on a log. mass axis)
4. Full model of coagulation *and* fragmentation (on a log. mass axis)

The results are displayed in the following sections.

3.8.1 Pure Coagulation on Linear Mass Axis

As an introduction, let us start with a very simple first example: The numerical integration of the Smoluchowski coagulation equation is to be carried out making use of the kernel matrix that we defined in section 3.5.1, i.e. with only pure hit-and-stick coagulation processes being included into the model. For the definition of the discretized mass axis, we make use of a linear scaling here. For the integration, we use the simple forward Euler method that was described in section 3.7.2.

For simplicity, the reaction rate coefficent is set to 1 here, as we're only interesting in getting a qualitative first view on the evolution of the mass distribution, which we initialize with all particles sharing an initial mass value of $m_0 = 1$ kg. The total mass volume density contained in the distribution is set to $\rho^d = 1\ \mathrm{kgm}^{-3}$ as well.

Observation 1: As we would expect in the case of pure coagulation, the distribution of particle masses shifts towards higher mass values as time progress.

Observation 2: The height of the graph stays relatively constant over time. This is due to the fact that here we are not plotting the number of particles per bin against the mass values corresponding to each bin, but instead the fraction of the total mass volume density in each bin. This representation thus gives an intuitive first insight into the mass conservation properties of the algorithm.

Observation 3: Though not shown here (since the plot would not look very interesting), the total mass contained in the distribution is conserved exactly here. The mass error is equal to zero, which stems from the more simple numerical treatment that can be done when using a linearly scaled mass axis. This exact conservation of mass will be lost once the switch to a logarithmic scaling is made.

Observation 4: As described by the CFL criterion, the usage of an explicit Euler algorithm will lead to a breakdown of the method when the time-steps become too large. This can be observed when increasing the size of the time-step Δt, but is not yet the case in the displayed simulation run.

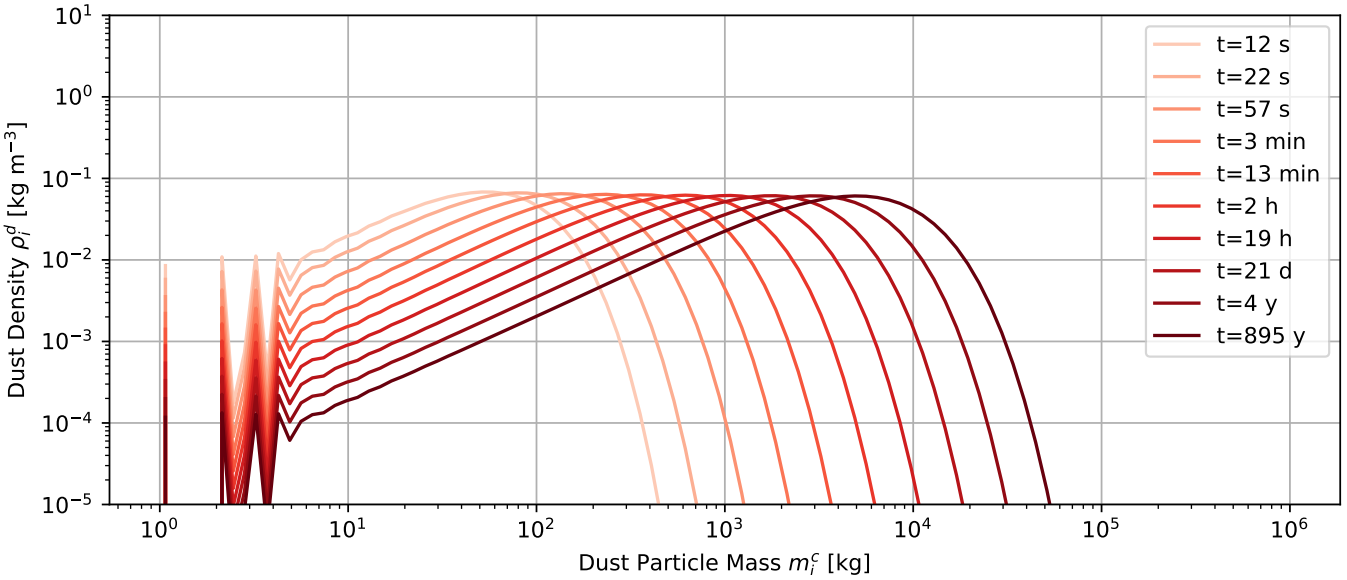


Figure 3.10: Temporal evolution of the dust particle mass distribution under the influence of only pure hit-and-stick coagulation. Here, the mass axis was discretized making use of a linear scaling.

3.8.2 Pure Coagulation on Logarithmic Mass Axis

As before, we start with all mass being concentrated in the lowest bin, i.e. all particles possessing the same small mass value. As the simulation progresses, more and more particle collsions occur. In this model, 100 % of collisions lead to a merging of the two particles.

As such, the particles in the disk grow larger and larger, while the number of particles becomes smaller and smaller.

At some point, the upper boundary of the discretized mass axis is reached. This can be seen on the right side of figure 3.12, where more and more particles gather in the highest-mass bin, and a peak is formed.

The stability properties of numerical integration with only pure coagulation are quite good. As we can see in figure 3.11, the unitless relative mass error stays consistently on the order of 10^{-15}.

As such, we can say that mass conservation is assured down to machine precision here.

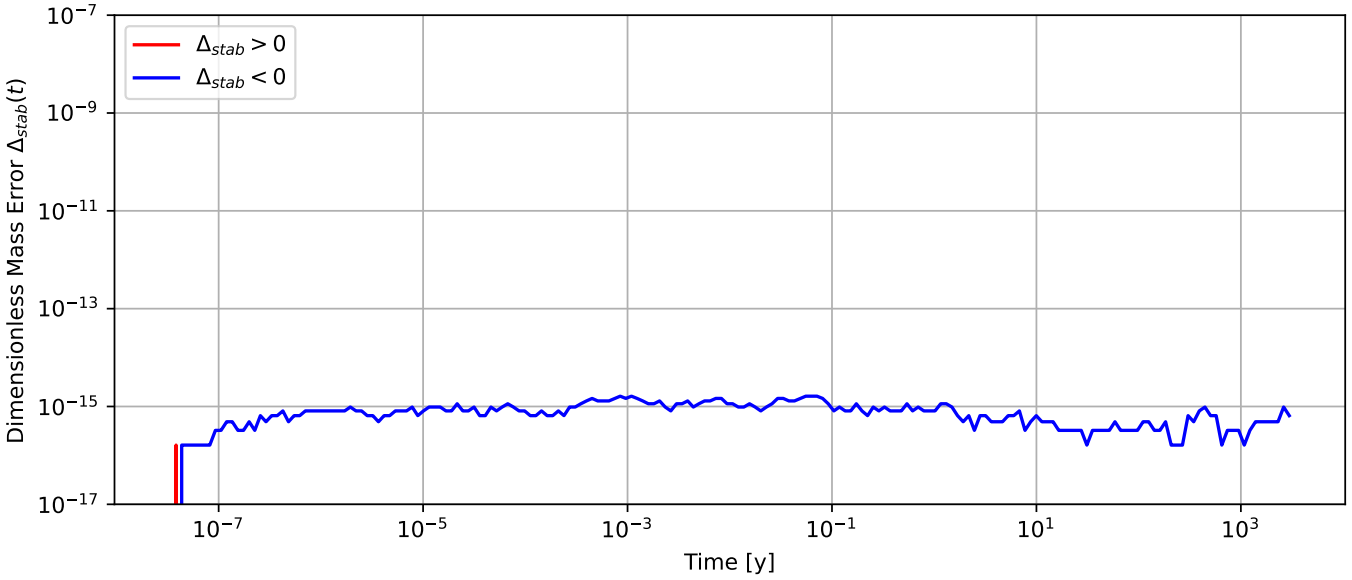


Figure 3.11: Relative mass error (stability error) $\Delta_{\text{stab}}(t) = \frac{\rho_d(t) - \rho_d(t=0)}{\rho_d(t=0)}$ of the numerical integration as a function of time. Here, only the process of pure coagulation is included in the model.

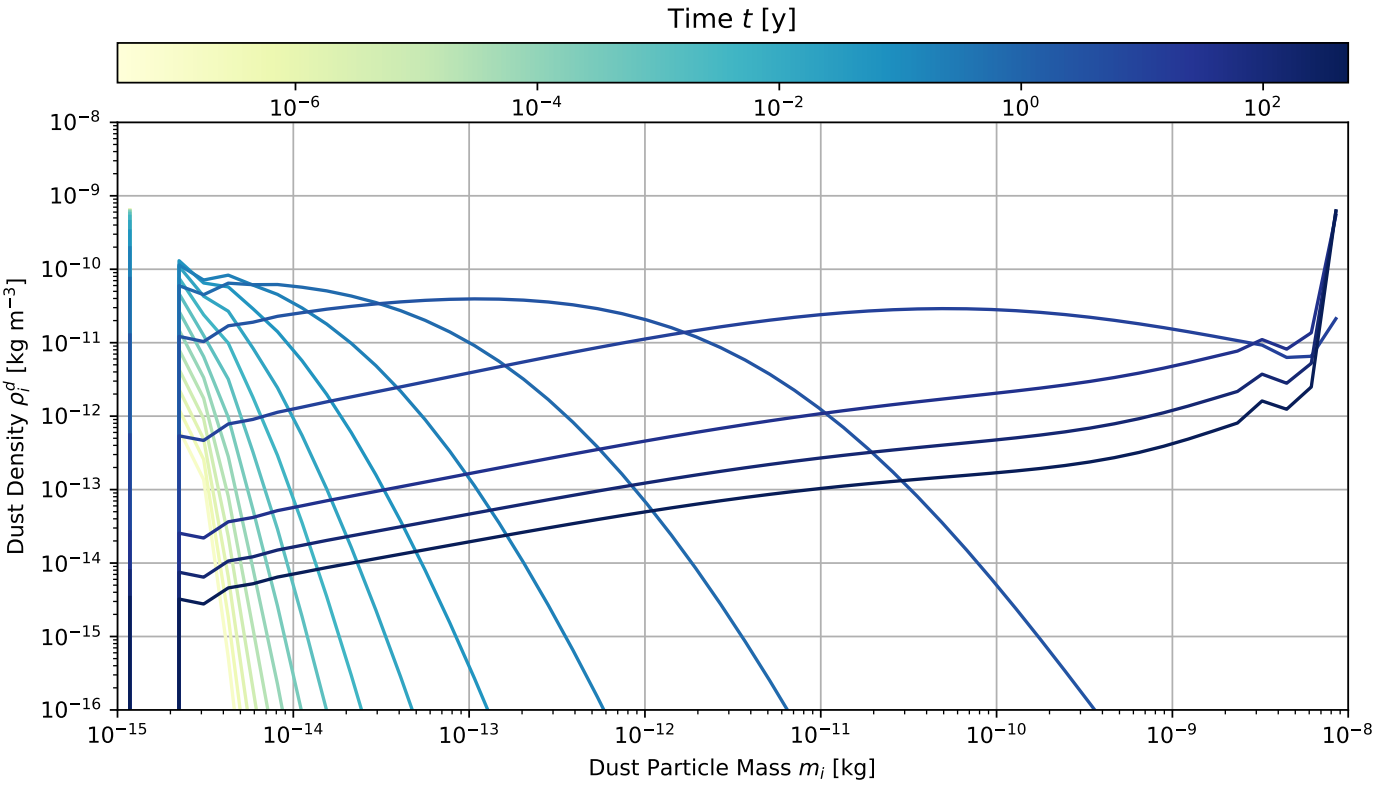


Figure 3.12: Temporal evolution of the dust particle mass distribution under the influence of only pure coagulation.

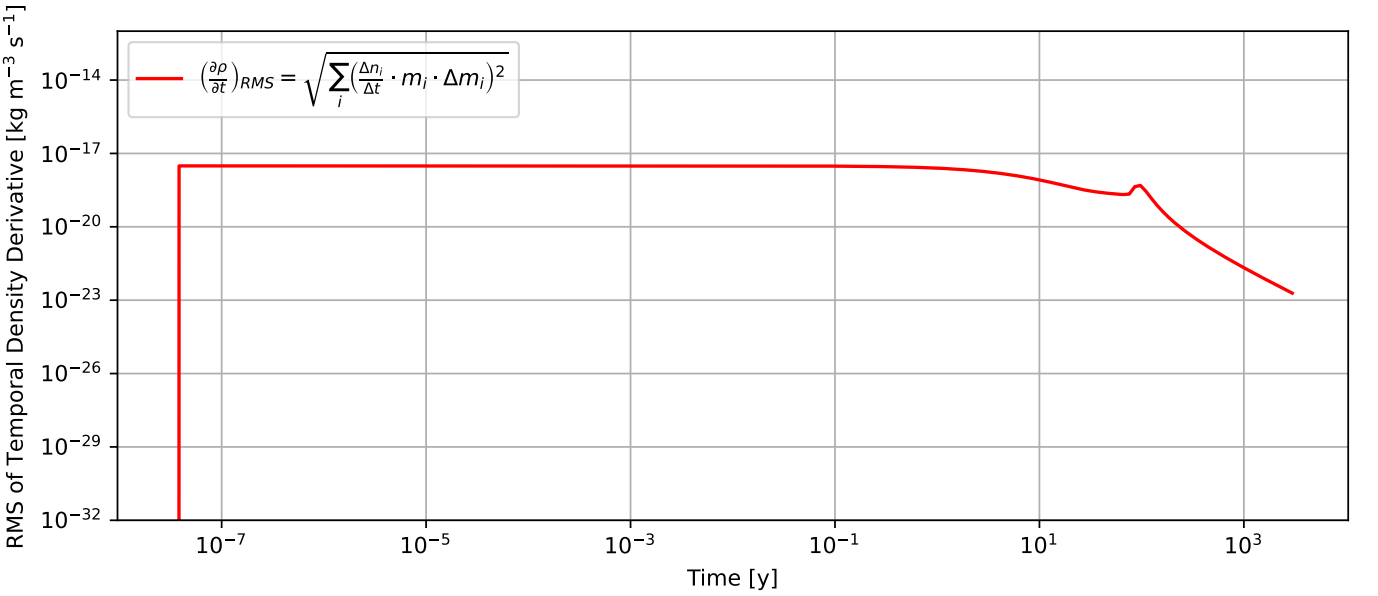


Figure 3.13: Root-mean-squared (RMS) temporal derivative of the dust particle mass distribution as a function of time. Here, only the process of pure coagulation is included into the model.

3.8.3 Pure Fragmentation on Logarithmic Mass Axis

Next, we will consider the scenario in which particle fragmentation is included into the model, but coagulation is not. In this case, it makes sense to not initialize the particle mass distribution as a Dirac δ-function centered on the lowest-mass bin, but choose a larger initial mass for the particles instead. Otherwise, we would not see anything happen at all in figure 3.12, where the particle mass distribution is shown at various points in time.

The behavior is what one might expect. Initially, we have larger particles, which are shattered on collision with another particle, such that the mass is redistributed into the lower-mass bins. Since there is no coagulation here, the mass stays in these low-mass bins and no large particles can be formed.

As we can see in figure 3.14, the inclusion of fragmentation processes into the model leads to the introduction of a mass error, which increases with time.

If we were to run the simulation for longer time-spans, this error would continue to grow, without any boundary in sight. This can be explained by the utilization of a logarithmically scaled mass axis.

Even if the mass error per collision is just a tiny number, it can become arbitrarily large, as it is amplified by the size of the time-step Δt, which itself increases with time.

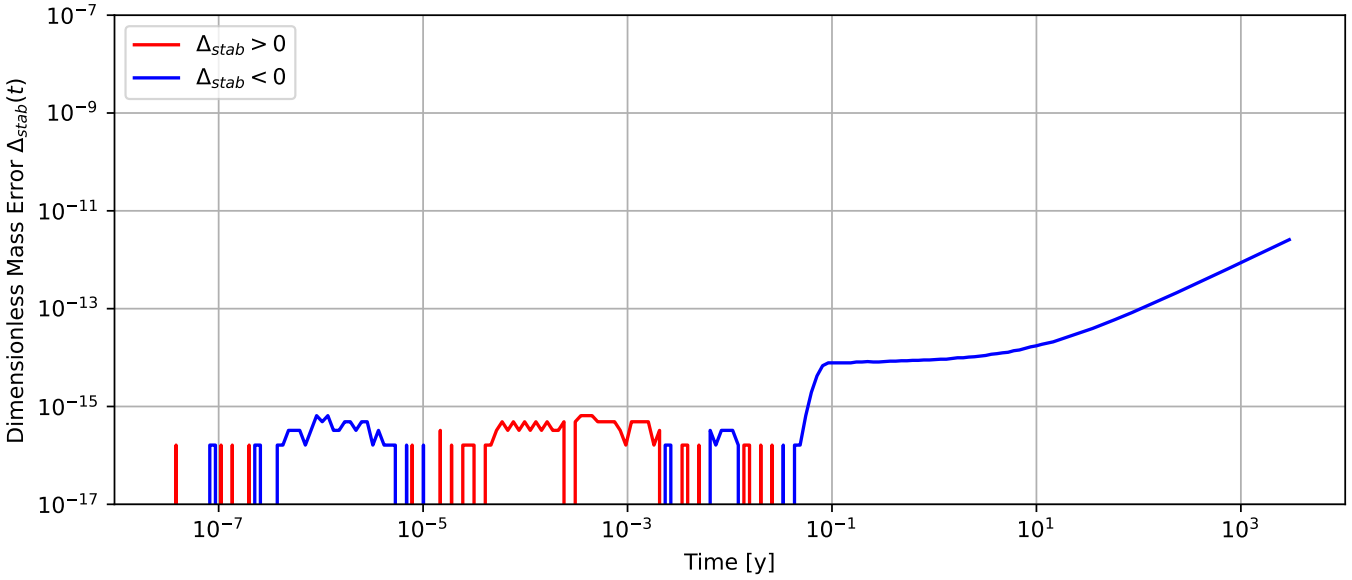


Figure 3.14: Relative mass error (stability error) $\Delta_{\text{stab}}(t) = \frac{\rho_d(t) - \rho_d(t=0)}{\rho_d(t=0)}$ of the numerical integration as a function of time. Here, only the process of pure fragmentation is included in the model.

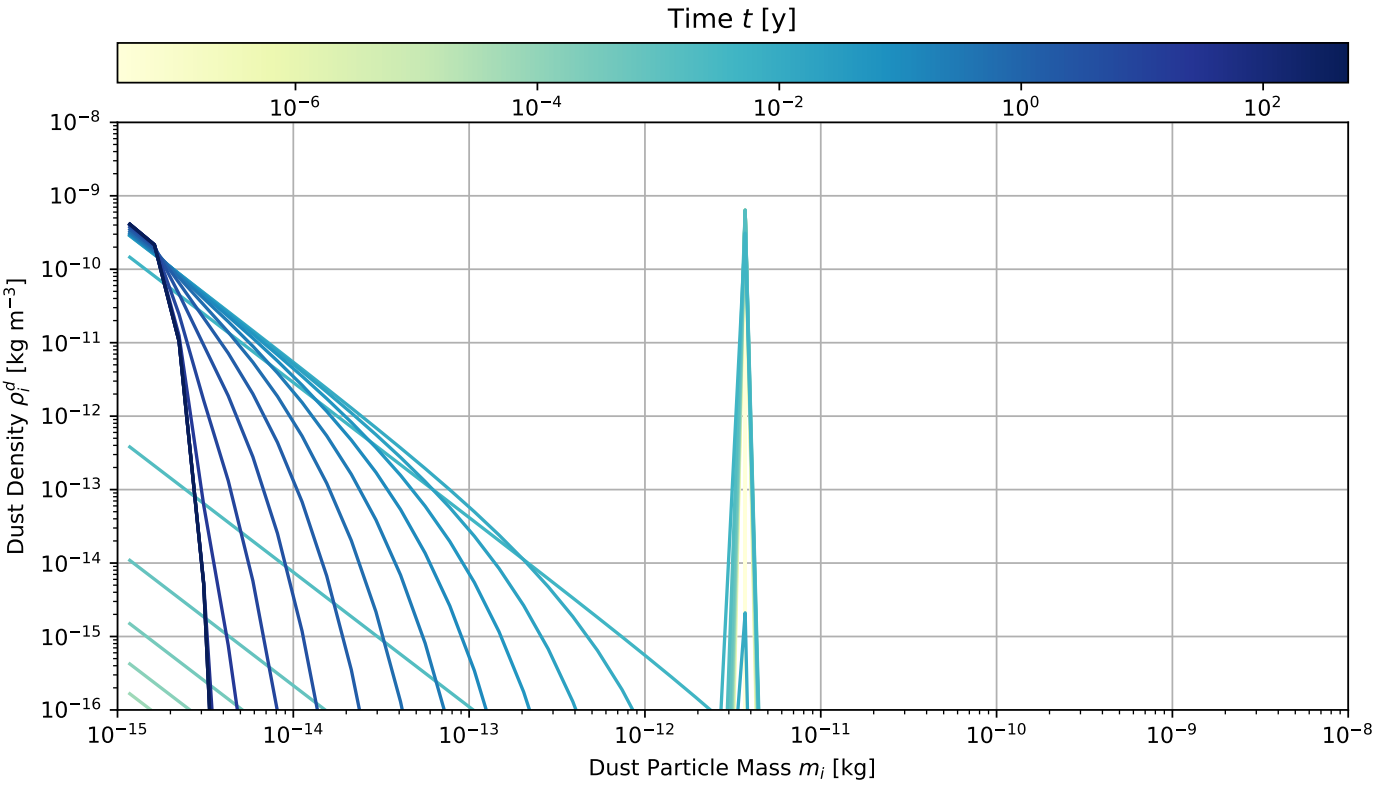


Figure 3.15: Temporal evolution of the dust particle mass distribution under the influence of only pure fragmentation. Here, the initial mass distribution is centered not on the first bin like in the other runs, but on one corresponding a higher mass value.

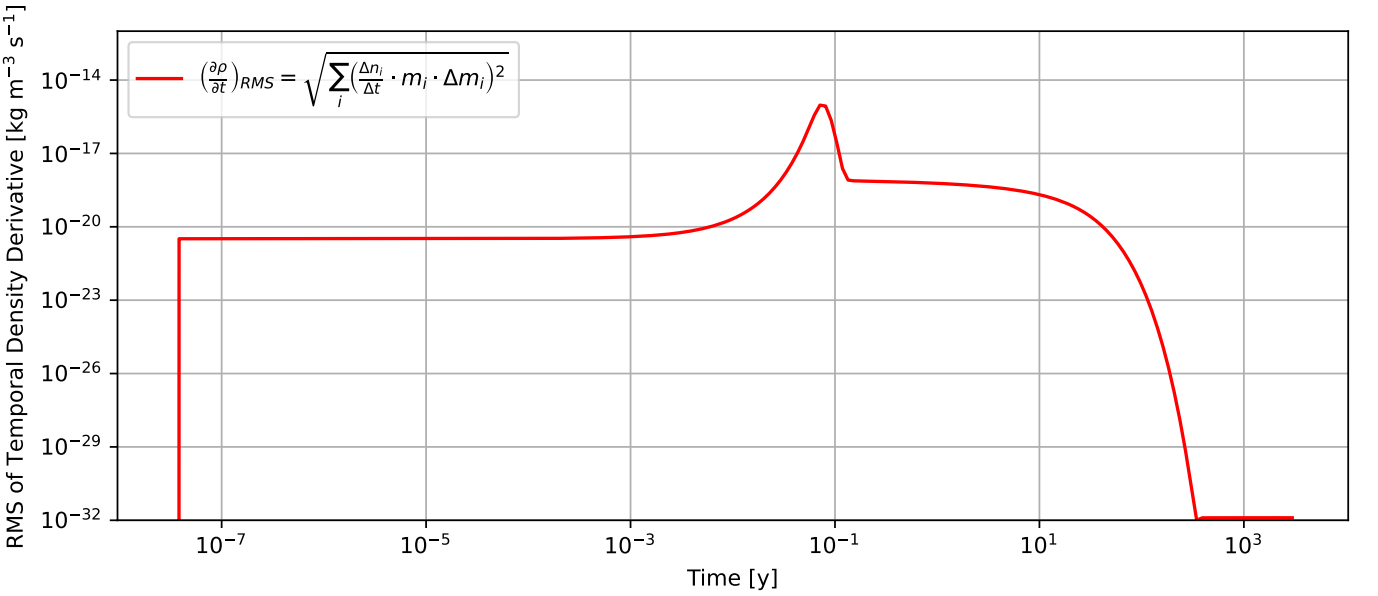


Figure 3.16: Root-mean-squared (RMS) temporal derivative of the dust particle mass distribution as a function of time. Here, only the process of fragmentation is included into the model.

3.8.4 Full Model

Let us now take a look at the full model, where the processes of both coagulation and fragmentation are included into the kernel matrix. The temporal evolution of the dust particle mass distribution is displayed in figure 3.18.

At the beginning, this evolution behaves very similarly to what we saw in the case of pure hit-and-stick coagulation in section 3.8.2. As a consequence of collisions between the initially very low-mass particles, larger bodies are formed via the merging of smaller ones. As such, the low-mass bins are emptied, and the mass contained within them is transferred into bins encoding more massive particles.

As the coagulation process goes on, larger and larger particles form. Eventually, some of them may become large enough that the relative velocity between them and some of the other particles surpasses the fragmentation threshold velocity v_{frag}.

Here, the observed behavior changes. After having emptied by the sticking process, low-mass bins get filled up again. As associative and dissociative processes work against each other, at some point the particle mass distribution function reaches an equilibrium.

When the equilibrium is reached, the temporal derivative of the mass distribution function goes to zero. This can be seen visualized in figure 3.19, where we plot the root-mean-square of the mass distribution's temporal derivative vector as a function of time.

In addition, the relative mass error of the integration that we defined in equation (3.70) is plotted in figure 3.17.

The increase of the error towards the end of the time-span over which we integrate is introduced because of the inclusion of dust particle fragmentation processes into the kernel, as was discussed in section 3.8.3.

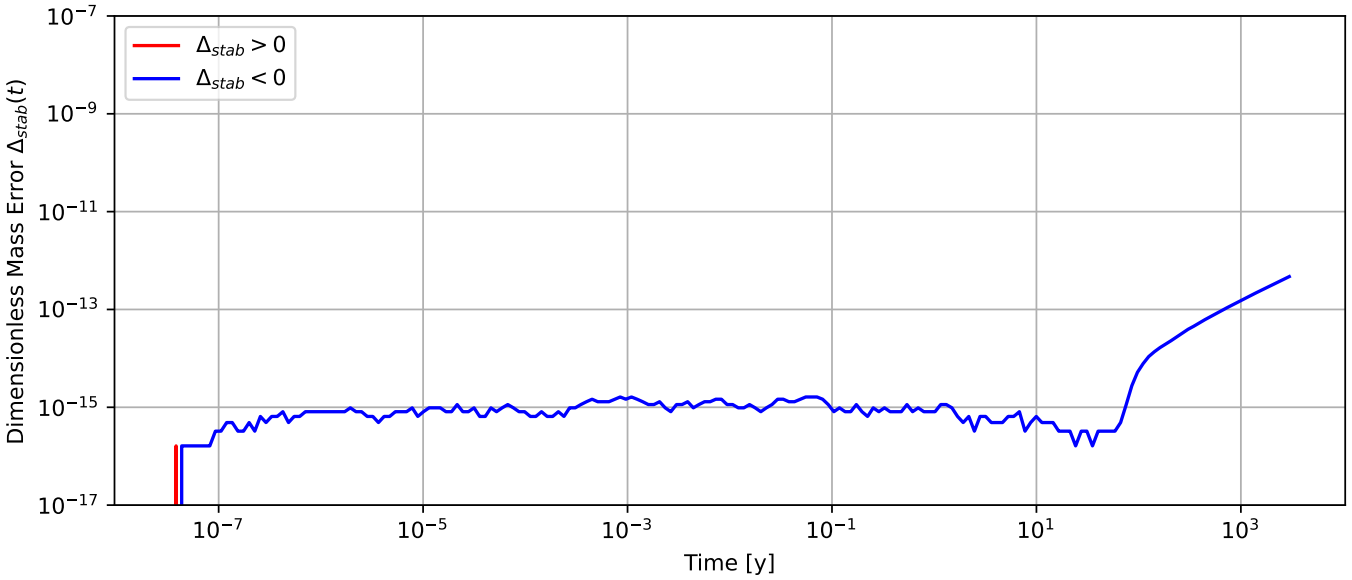


Figure 3.17: Relative mass error (stability error) $\Delta_{\text{stab}}(t) = \frac{\rho_d(t) - \rho_d(t=0)}{\rho_d(t=0)}$ of the numerical integration as a function of time. Here, both coagulation and fragmentation processes are included in the model.

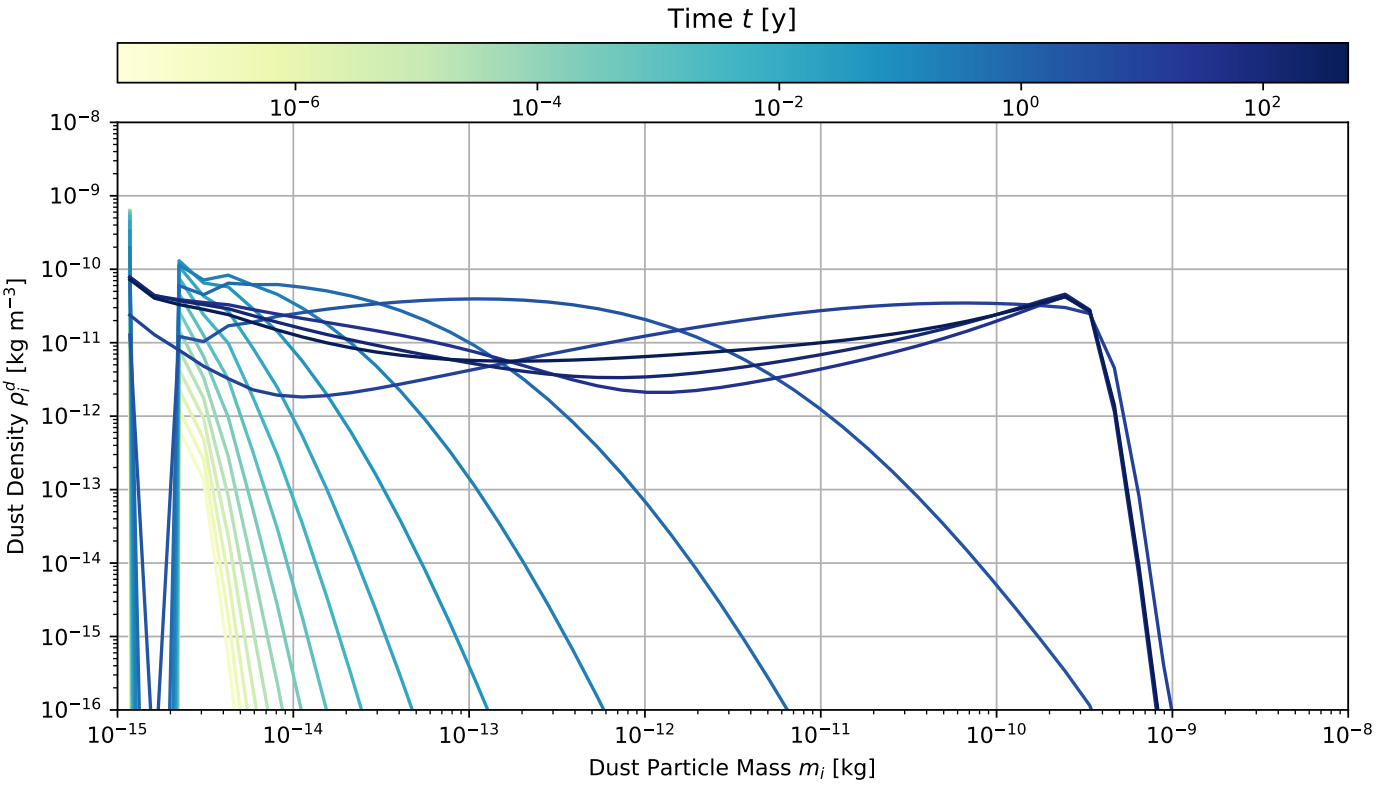


Figure 3.18: Temporal evolution of the dust particle mass distribution under the influence of both coagulation and fragmentation processes.

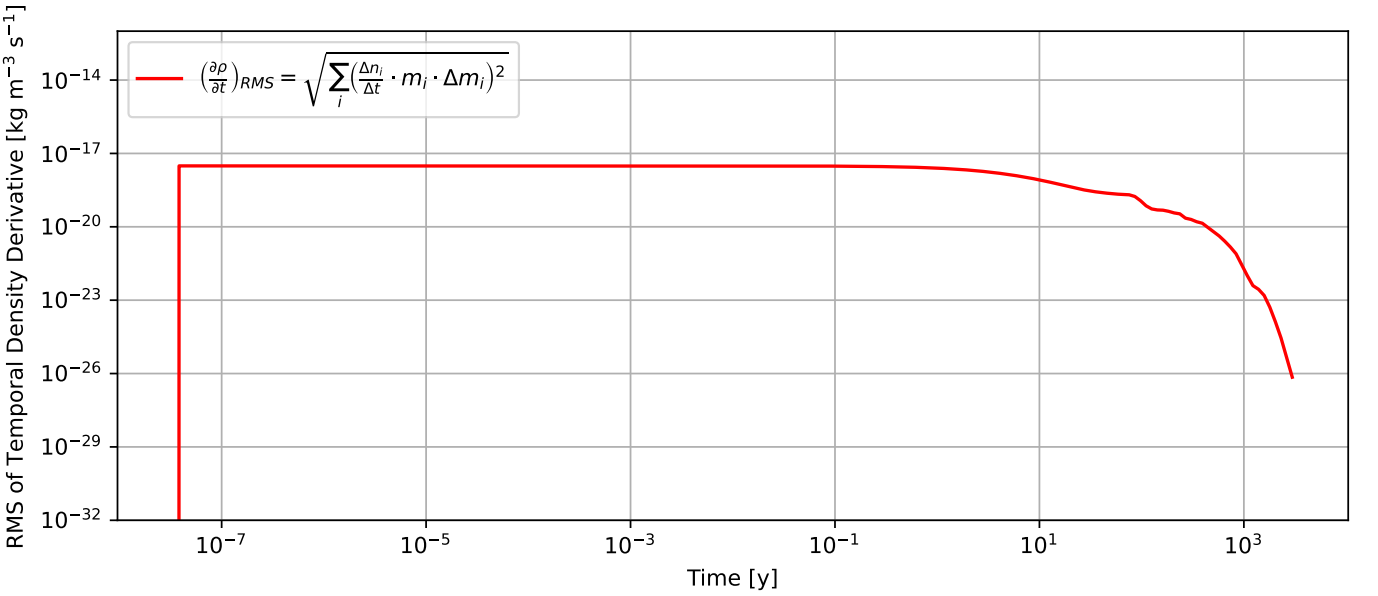


Figure 3.19: Root-mean-squared (RMS) temporal derivative of the dust particle mass distribution as a function of time. Here, both the process of coagulation and fragmentation is included into the model.

Chapter 4

Stochastic Monte Carlo Sampling of the Kernel Matrix

After having defined a simple disk model, and from that calculated the reaction rates for coagulation and fragmentation events, we will now focus on the main goal of this thesis:

We would like to find out whether it makes sense to utilize stochastic Monte Carlo sampling of the kernel matrix, in order to lower the cost of the Smoluchowski coagulation equation's numerical integration.

4.1 Introduction

4.1.1 Motivation

Let us first think about why that would even be necessary. The computational cost of integrating the Smoluchowski equation within the context of our model is affected by a number of the model's parameters:

The first of these is given by the mass grid resolution $\mathcal{N}_m$, i.e. the number of mass "bins". Remember that, in each time-step, we are interested in computing an updated version of the discretized dust particle mass distribution function, which involves updating each of its entries.

For each of these entries k, we need to perform a loop over all collision pairs (i, j) that could potentially lead to a creation or destruction of a particle with mass m_k. Having discretized the mass axis into $\mathcal{N}_m$ bins, the total number of collisions that we need to consider is given by $\mathcal{N}_m^2$.

Since this needs to be done for each value of k, the total numerical cost $\mathcal{C}$ for forwarding the mass distribution at a given location from one time-step to the next follows the relationship

$$\mathcal{C} = \mathcal{O}\big(\mathcal{N}_m^3\big) \tag{4.1}$$

When considering only the simple case of pure hit-and-stick coagulation, this cubic relationship turns into a quadratic one, since we can make use of the criterion of mass conservation: For each particle mass m_k that could emerge from a collision between particles with masses m_i and m_j, there is only a single value m_j that satisfies the mass conservation criterion $m_k = m_i + m_j$, and vice versa. In this case, the kernel matrix does not even have to be cubic, but can be quadratic instead.

The inclusion of fragmentation is what makes this more complicated. Here, a wide range of differently-sized particles could emerge from any given collision, requiring the definition of a cubic kernel matrix, and leading to the cubic dependence of the numerical cost on the mass grid resolution.

Note that up until this point we have only considered the numerical cost of forwarding the mass distribution a single time-step, at a single location, and assuming the individual dust particles can be characterized entirely by their mass value alone.

For the coagulation studies that are done in the context of this thesis, we only consider a single position in the disk, located at $\vec{r} = \vec{r}_0$, defined in equation (2.11). A more sophisticated model would of course require the calculation of the mass distribution's temporal evolution at several different positions in the disk.

If both spatial and temporal discretization are included, the numerical complexity increases by additional factors. Adopting cylindrical coordinates and assuming a radial resolution $\mathcal{N}_r$, azimuthal resolution $\mathcal{N}_\phi$, vertical resolution $\mathcal{N}_z$, and temporal resolution $\mathcal{N}_t$, the complexity becomes

$$\mathcal{C} = \mathcal{O}\big(\mathcal{N}_m^3 \cdot \mathcal{N}_t \cdot \mathcal{N}_r \cdot \mathcal{N}_\phi \cdot \mathcal{N}_z\big) \tag{4.2}$$

In the context of this thesis, the dust particles are modeled as perfectly spherical bodies with a shared solid mass density ρ_s. A possible next step for obtaining a more realistic model is given by the inclusion of a *porosity parameter* p, which allows a slightly more detailed description of the dust particles' internal structure, and that structure's influence on the processes of coagulation, fragmentation, and bouncing. Studies building upon this idea have already been made, for this see e.g. [OST06] [Kat17].

The inclusion of a dust porosity parameter is probably the most significant argument speaking for a Monte Carlo sampling of the kernel. While the cubic dependency of the numerical cost on $\mathcal{N}_m$ is bad enough, the inclusion of porosity makes numerical integration even more costly.

The reason for this is the following: If porosity is included, the individual dust particles can not be characterized by a single number alone anymore, instead two are needed. Thus, we do not have a one-dimensional particle distribution $n(m)$, but a two-dimensional $n(m, p)$. Then, in each time-step and at each location, for each combination of mass and porosity, we need to find out how the interaction between pairs of particles, each carrying an indiviual combination of mass and porosity, influence the particle distribution.

This requires the evaluation of a loop over $\mathcal{N}_m^3 \cdot \mathcal{N}_p^3$ cases. The total cost becomes

$$\mathcal{C} = \mathcal{O}\big(\mathcal{N}_m^3 \cdot \mathcal{N}_p^3 \cdot \mathcal{N}_t \cdot \mathcal{N}_r \cdot \mathcal{N}_\phi \cdot \mathcal{N}_z\big) \tag{4.3}$$

This last equation showcases a behavior that is quite expensive. Consider the following: If we discretize our mass axis into 100 bins (which is not actually that much), and do the same for the porosity axis, evaluating the complete sum on the right-hand side of the Smoluchowski equation would already require the inclusion of

$$\mathcal{N}_m^3 \cdot \mathcal{N}_p^3 = 10^{12} \tag{4.4}$$

terms. Note that this is for a single time-step, and at a single location in the disk!

To allow studies of dust coagulation in a model including both a mass and porosity parameter, this is much too expensive to finish the simulations in a sensible time span. Thus, a method for lowering the computational cost would be desirable.

A first idea for how to do this could be to include only the most relevant collisions into the sum over the kernel, such that fewer calculations have to be made. What exactly we mean by "relevant collision" is of course yet to be defined, see section 4.2.2 for this.

In any case, a Monte Carlo sampling approach for lowering the computational cost seems sensible, and will be implemented in the following sections.

4.1.2 Previous Works

Previous studies have already made use of Monte Carlo methods in an attempt to bring down the numerical cost of the integration of the Smoluchowski equation.

Notable examples for such prior work include e.g. the studies done for the 2006 paper by C. W. Ormel et. al. [OST06], where a direct Monte-Carlo simulation method (DSMC) is used, and the more recent 2023 paper by M. Beutel [BD23], which makes use of a Representative Particle Monte Carlo method (RPMC).

4.2 Kernel Sampling

4.2.1 Basic Idea

As noted above, the goal here is to reduce the cost of numerically integrating the Smoluchowski coagulation equation.

The basic idea is to not include all terms in the summation needed for the evaluation of the Smoluchowski equation's right-hand side. This would lower the cost, as fewer calculations have to be made.

Instead of including *all* collision pairs, we therefore need to identify the "most relevant" ones.[1] Here, relevance is determined by a specific collision's impact on the overall dust particle mass distribution.

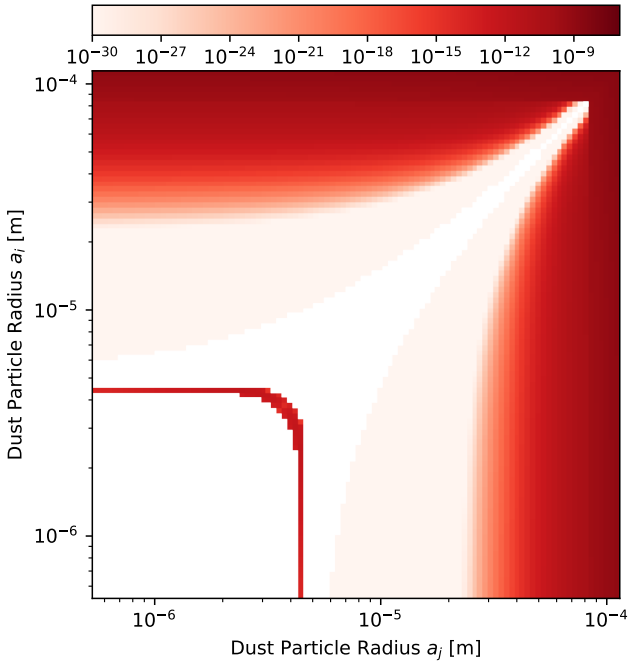
After having defined a sensible sampling probability distribution P_{ij}^{sample} we can, in each time-step, randomly select the most relevant index pairs (i, j) and construct a new kernel from that. This kernel is then used to forward the mass distribution to the next time-step.

If it works, we should be able to decrease the number of terms over which the sum has to be evaluated, while still maintaining acceptable accuracy and stability.

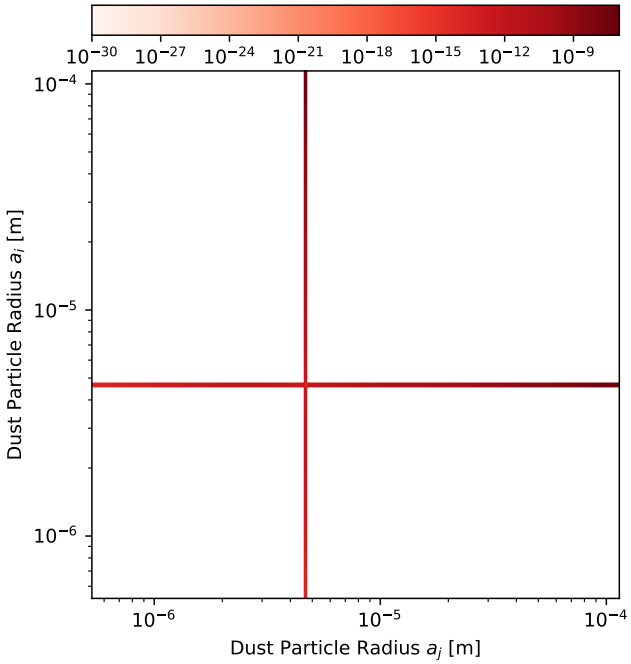
This basic idea is illustrated in figure 4.1, where a comparison between "full" and sampled kernel is displayed. Note that here, as a first example we are using a homogenous sampling probability. Since not all collision are relevant to the same extent for the evolution of the dust particle mass distribution, in the next section we will define a non-homogenous, more sensible sampling density.

It should also be noted that we the displayed kernel matrix is symmetric around the diagonal. This is only relevant for visualization though, and does not mean that the terms on the lower right of the matrix are actually defined as non-zero values. We can safely set them to zero, due to the symmetry $i \leftrightarrow j$ with regards to the particles partaking in a given collision.

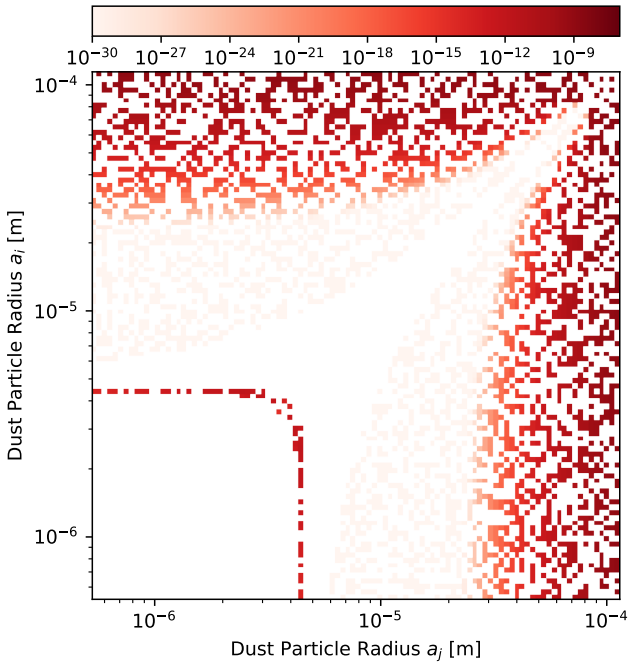
[1] It needs to be emphasized that we do *not* sample i and j separately here! Instead, we sample collision *pairs* (i, j).



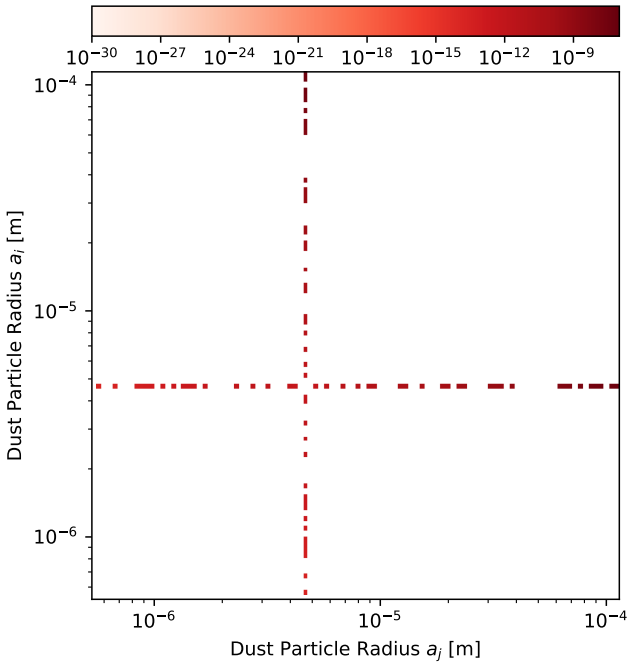
(a) Positive component of total kernel



(b) Negative component of total kernel



(c) Positive component of sampled kernel



(d) Negative component of sampled kernel

Figure 4.1: Illustration of the basic idea for kernel sampling: Only a fraction of the total number of possible particle collision is taken into account when defining the kernel matrix. This number is determined by the sampling density ρ_{sample}, which is set to $1/2$ here. The kernel is plotted for $k = 40$ on a mass grid with resolution $\mathcal{N}_m = 100$. Note that here, we are using a homogenous probability distribution for the collision sampling. Since not all collisions are relevant for the temporal derivative of the dust particle mass distribution to the same extent, we will define a more sensible non-homogenous probability distribution in the next section.

4.2.2 Definition of the Sampling Probability

As noted above, it is our goal to identify the most relevant collisions, i.e. those collisions that lead to the biggest change in the dust particle mass distribution from one time-step to the next. Having done that, we would like to encode this relevance of the collisions into a probability distribution.

It should be chosen in such a way that the most influential collisions are also those that possess highest probability of being selected at random using the probability distribution. For this, we have to take two things into account:

1. First, a certain collision might be deemed "relevant" because a large amount of mass is transferred between bins at each time that such a collision occurs.

2. Secondly, a collision might also be relevant if only a small amount of mass is transferred at each collision event, but these events happens very frequently.

In the following, we will try to encode these two conditions into a sensible probability distribution. For this, consider a collision (i, j) between particles with masses m_i and m_j. Let k be the index of a mass bin representing the population of particles with mass m_k.

When such a collision (i, j) takes place, the amount of mass that is transferred into a bin k is given by the expression

$$\Delta M_{kij} := m_k \cdot K_{kij} \tag{4.5}$$

Also, the total mass transferred between all bins during such a collision is given by the sum

$$\Delta M_{ij} := \sum_{k=1}^{\mathcal{N}_m} \Delta M_{kij} \tag{4.6}$$

$$= \sum_{k=1}^{\mathcal{N}_m} m_k \cdot K_{kij} \tag{4.7}$$

Because no mass is created during collisions, this sum ΔM_{ij} is equal to zero.[2]

To get an idea of the *total net mass* $\Delta M_{ij}^{\text{net}}$ that is being moved between bins, we can slightly alter the expression given in equation (4.6) and write

$$\Delta M_{ij}^{\text{net}} := \sum_{k=1}^{\mathcal{N}_m} \left|\Delta M_{kij}\right| \tag{4.8}$$

$$= \sum_{k=1}^{\mathcal{N}_m} m_k \cdot \left|K_{kij}\right| \tag{4.9}$$

Defined like this, the back-and-forth movement of mass between bins does not lead to canceling, and instead we get a measure for the total net transferred mass resulting from a collision (i, j).

Using this definition to quantify a collision's impact on the dust particle mass distribution, we can recognize the first proportionality relationship for our sampling probability P_{ij}^{sample}:

$$P_{ij}^{\text{sample}} \sim \Delta M_{ij}^{\text{net}} \tag{4.10}$$

[2]This is only really true in physical reality. The model discretization and usage of numerical methods introduces an error, such that this sum will, in fact, *not* be exactly equal to zero. To assure the conservation of mass, and therefore the stability of the algorithm, it is important to keep this error as small as possible. Therefore, in our implementation we need to make sure that mass is conserved down to machine precision at each collision event. This must also be true when only an incomplete (sampled) version of the kernel is used. As such, we must ensure that mass is conserved *in detailed balance*.

Up until now, the sampling probability of a given collision does not depend on the rate at which such collisions occur. It seems plausible though that, all other factors being equal, more frequently occuring collisions should also be sampled more frequently.

To fix this, let us recognize the second proportionality relationship for the sampling probability, which depends directly on the dust particle collision rate coefficient R_{ij}^{coll}:

$$P_{ij}^{\text{sample}} \sim R_{ij}^{\text{coll}} \tag{4.11}$$

Remember that the coefficient R_{ij}^{coll} does not label an actual *collision rate* with units of s^{-1}, but actually a collision rate normalized by dust particle number density with units $\text{m}^3\ \text{s}^{-1}$.

Since more collisions will occur at places where there are a lot of particles, the sampling probability must depend on the number of particles per unit volume as well. Therefore:

$$P_{ij}^{\text{sample}} \sim N_i \quad \text{and} \quad P_{ij}^{\text{sample}} \sim N_j \tag{4.12}$$

Additionally, collisions involving high-mass particles will be weighted more significantly than collisions involving low-mass particles:

$$P_{ij}^{\text{sample}} \sim m_i \quad \text{and} \quad P_{ij}^{\text{sample}} \sim m_j \tag{4.13}$$

The total sampling probability distribution is then directly proportional to the product of the contributions that we discussed so far:

$$P_{ij}^{\text{sample}} \sim \Delta M_{ij}^{\text{net}} \cdot N_i \cdot N_j \cdot m_i \cdot m_j \tag{4.14}$$

To turn this into an actual probability distribution though, we still have to perform a normalization step, to assure that the sum over all probabilities is equal to one:

$$\sum_{i=1}^{\mathcal{N}_m} \sum_{j=1}^{\mathcal{N}_m} P_{ij} \stackrel{!}{=} 1 \tag{4.15}$$

To do this, we define the two preliminary expressions

$$F_{ij}^{\text{sample}} := \Delta M_{ij}^{\text{net}} \cdot N_i \cdot N_j \cdot m_i \cdot m_j \quad \text{and} \tag{4.16}$$

$$F^{\text{sample}} := \sum_{i=1}^{\mathcal{N}_m} \sum_{j=1}^{\mathcal{N}_m} \mathcal{P}_{ij} \tag{4.17}$$

and, with these, arrive at the normalized probability distribution:

$$\boxed{P_{ij}^{\text{sample}} = \frac{F_{ij}^{\text{sample}}}{F^{\text{sample}}} = \frac{\Delta M_{ij}^{\text{net}} \cdot N_i \cdot N_j \cdot m_i \cdot m_j}{\sum_{i'=1}^{\mathcal{N}_m} \sum_{j'=1}^{\mathcal{N}_m} \Delta M_{i'j'}^{\text{net}} \cdot N_{i'} \cdot N_{j'} \cdot m_{i'} \cdot m_{j'}}} \tag{4.18}$$

This probability distribution is what we will use in the following to construct a randomly sampled kernel. Its definition will always depend on the given dust particle mass distribution: Depending on what kind of particles are present at the moment (and the masses they carry), the most relevant collisions can vary significantly.

Note that this sampling will be implemented "without returning", i.e. each collision will only be chosen a single time at most per time-step. An alternative method would of course be given by the implementatin of a method that allows duplicate sampling, i.e. sampling "with returning". This would require the definition of a normalization, e.g. by dividing the kernel matrix K_{kij} by the sampling probability P_{ij}^{sample}. This approach is not pursued in this thesis, but could be in a future study.

To filter out less relevant collisions, we will implement a threshold probability value $P_{\text{min}}^{\text{sample}}$ that needs to be reached before a collision is deemed relevant enough to be sampled. Assuming a mass grid resolution of $\mathcal{N}_m = 100$, we will set this to

$$P_{\text{min}}^{\text{sample}} := 10^{-9} \tag{4.19}$$

4.2.3 Sampling of the Pure Coagulation Kernel

In figure 4.2, the temporal evolution of the dust particle mass distribution is plotted for various different values of the sampling density ρ_{sample}. Here, only the process of pure hit-and-stick coagulation is included into the kernel.

Observation 1: As can be seen when comparing the plots for different values of ρ_{sample}, the accuracy of the integration suffers when the sampling density is set to low values. This is of course not unexpected, since low sampling densities mean that a lot of possible collisions are neglected, leading to an incomplete picture of the particle collision dynamics.

Observation 2: There are statistical fluctuations in the numerical solutions, such that the numerical solution is much less "smooth" as it would be if it were calculated using the complete kernel matrix. This is to be expected as well, as not all collisions are taken into account in each time-step, and there exists a certain under- or over-representation of certain collisions due to the probabilistic nature of the method that is used for selecting kernel the entries that will be included into the summation.

In figure 4.3, the sampling probability P_{ij}^{sample} is plotted for various points in time throughout the numerical integration. figure 4.4 on the other hand attempts to display the number of times N_{ij}^{sample} that each collisions (i, j) was sampled. Since we are implementing a sampling "without returning", this N_{ij}^{sample} can only take on a value of either 0 or 1.

For the sampling density, we adopt a value of $\rho_{\text{sample}} = 0.5$ here.

Observation 3: Initially, i.e. when forwarding the mass distribution from the initial state to the next time-step for the first time, only a single entry of the kernel matrix has to be considered for the summation. This changes as time progresses, and the number of sampled collisions increases as more and more bins in the mass distribution carry a non-zero value.

Observation 4: Towards the end of the simulation, the number of sampled collisions decreases again. In the simple case of hit-and-stick coagulation, on a discretized mass axis with an upper limit m_{max}, at the end almost the entire mass will be contained in the bin corresponding to the largest mass value. As such, only collisions between particles from these high-mass bins need to be considered.

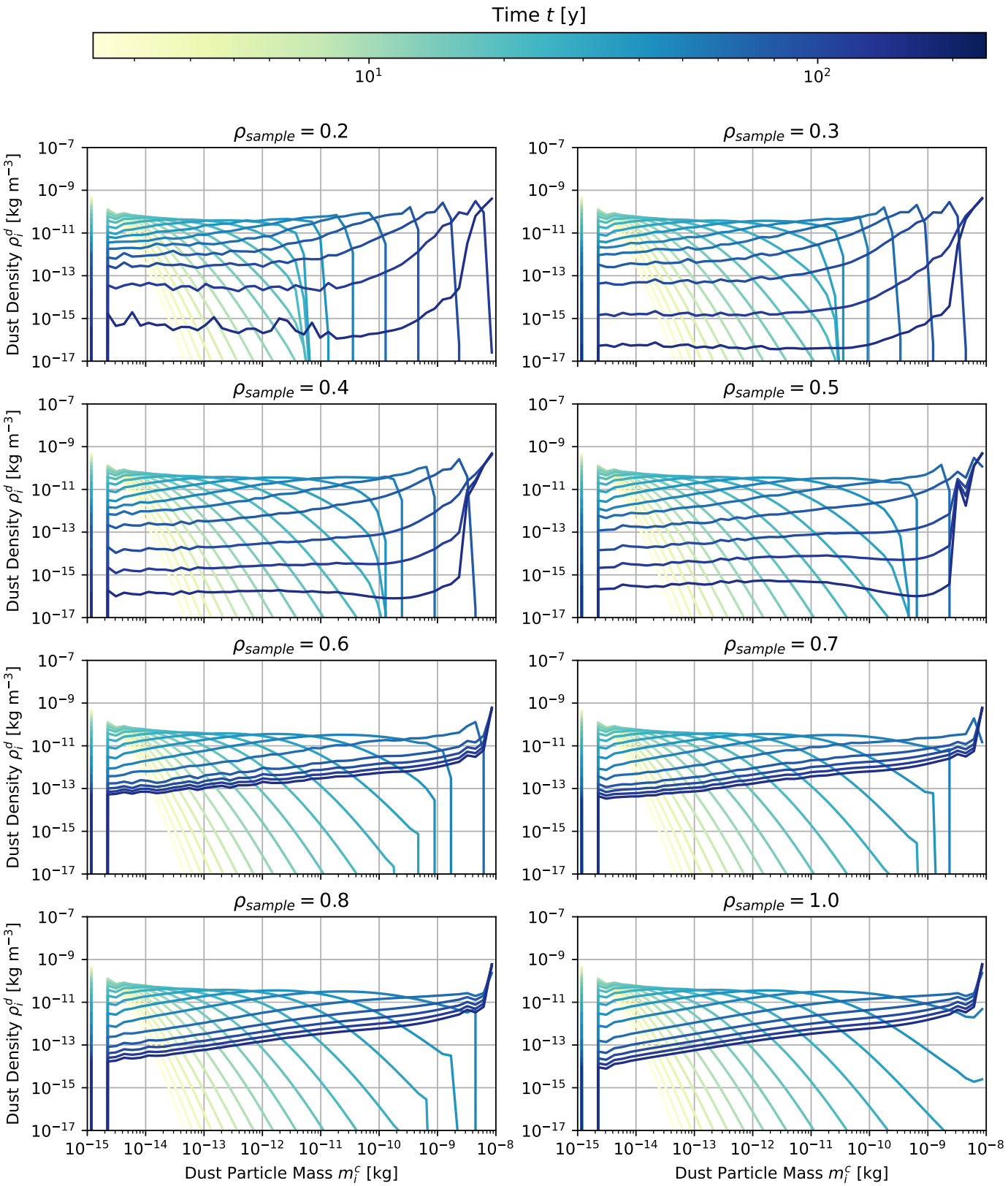


Figure 4.2: Temporal evolution of the dust particle mass distribution under the influence of (only) pure hit-and-stick coagulation processes for various different values of the sampling density ρ_{sample}.

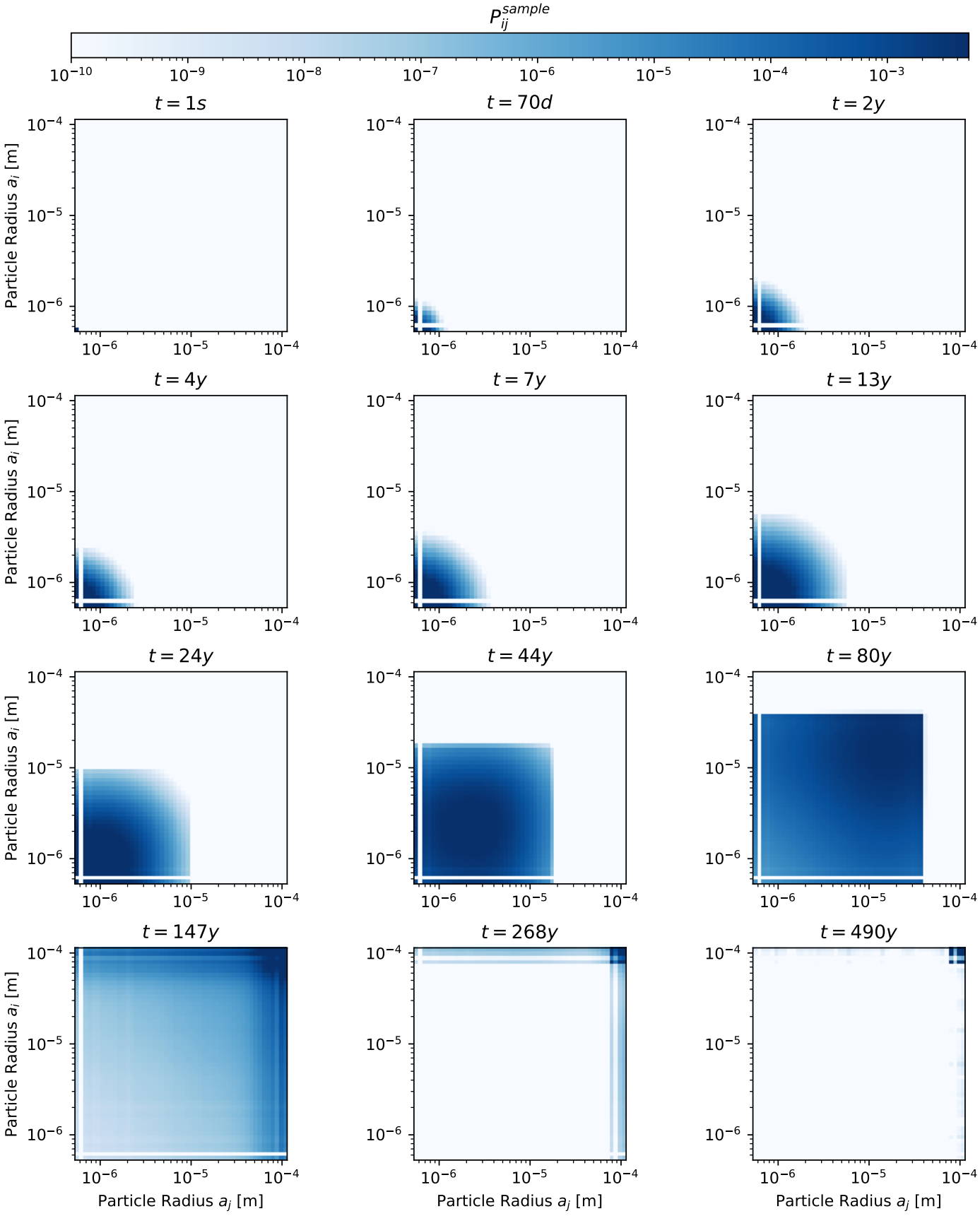


Figure 4.3: Illustration of the collision sampling probability distribution P_{ij}^{sample} for $\rho_{\text{sample}} = 0.5$. Here, only the process of pure stick-and-hit coagulation is included into the model.

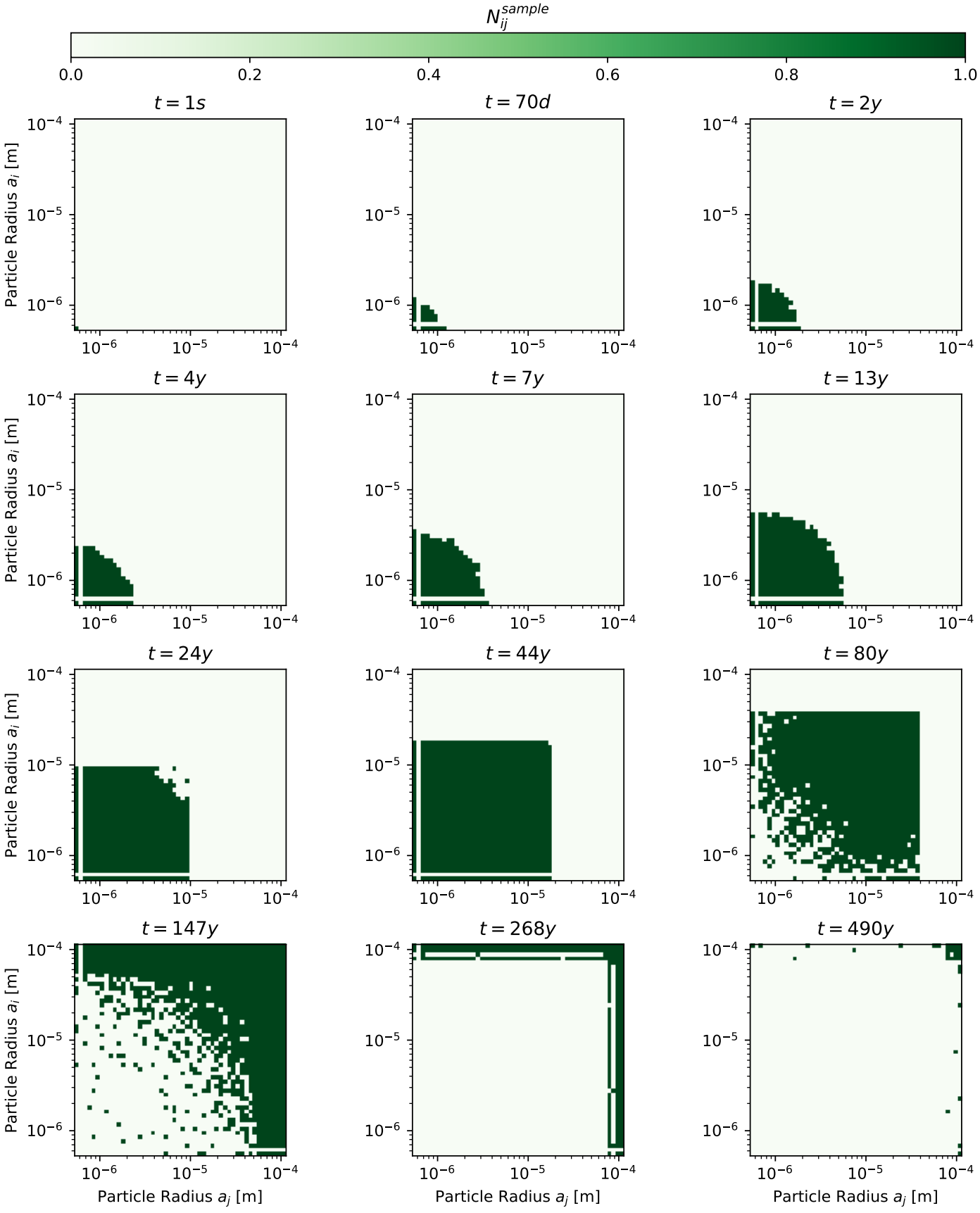


Figure 4.4: Illustration of the amount of times N_{ij}^{sample} that each collision (i, j) is sampled, for $\rho_{\text{sample}} = 0.5$. Here, only the process of pure stick-and-hit coagulation is included into the model.

4.2.4 Sampling of the Total Kernel

The plots in figure 4.5, figure 4.6, and figure 4.7 display analogous results to what we saw in figure 4.2, figure 4.3, and figure 4.4. The difference is that here, the full model is used, i.e. both coagulation and fragmentation processes are taken into account when defining the kernel matrix.

We will focus on the plots displaying the mass distribution's temporal evolution first:

Observation 1: Once again, there are visible fluctuations in the numerical solution of the particle mass distribution. As one might expect, these decrease in magnitude when the sampling density is increased.

Observation 2: Even though the accuracy may suffer from low sampling densities, the results shown here can still be seen as a success, as the mass distribution *does* evolve into a structure that visibily resembles the equilibrium that is reached in the "complete" solution. This is quite satisfactory, as it shows that the sampling method could be implemented in a sensible manner.

Observation 3: It might not be obvious at first sight, but the plots showcase another interesting behavior of the solutions resulting from a sampled kernel: If you look closely, you will notice that the evolution of the mass distribution progresses much slower towards the equilibrium state if the sampling density is low. This of course is nothing unexpected: If not all collisions are taken into account, this effectively leads to a lower collision rate, and thus a slower evolution of the distribution in general.

Observation 4: Another interesting feature one can notice is the existence of peak-like structures at the upper end of the mass axis, which are present even for larger values of the sampling density. This could hint at an under- or over-representation of collisions involving particles whose masses are close to the upper "edge" of the equilibrium state's mass distribution. It might be fixed by adjusting the definition of the sampling probability in a sensible manner.

Let us now take a look at the figures displaying the distribution of the sampling probability and the distribution of sampled collisions. Like before, here we adopt a sampling density value of $\rho_{\text{sample}} = 0.5$.

Observation 5: When comparing the distributions P_{ij}^{sample} and N_{ij}^{sample} with the results from the previous section, we can see that the temporal evolution behaves very similarly, at least during the early parts of the simulation.

Observation 6: Once big enough particles could form, allowing the relative particle velocities to surpass the fragmentation threshold velocity v_{frag}, the similarities come to an end. Because the low-mass bins are "filled up" again by collison events leading to particle fragmentation, they will have to be taken into account here, in contrast to what we saw in the case of pure coagulation.

Observation 7: There exists a "hole" or "gap" in the sampling probability distribution P_{ij}^{sample} for collisions where both particles carry neither a very large nor a very small mass value. Thus, towards the end of the simulation, when the equilibrium state begins to settle in, the most relevant collisions appear to be those containing either at least one large particke, or only small ones. This reflects the fact that the particle populations with those two respective sizes are the most prevalent in the distribution towards the end.

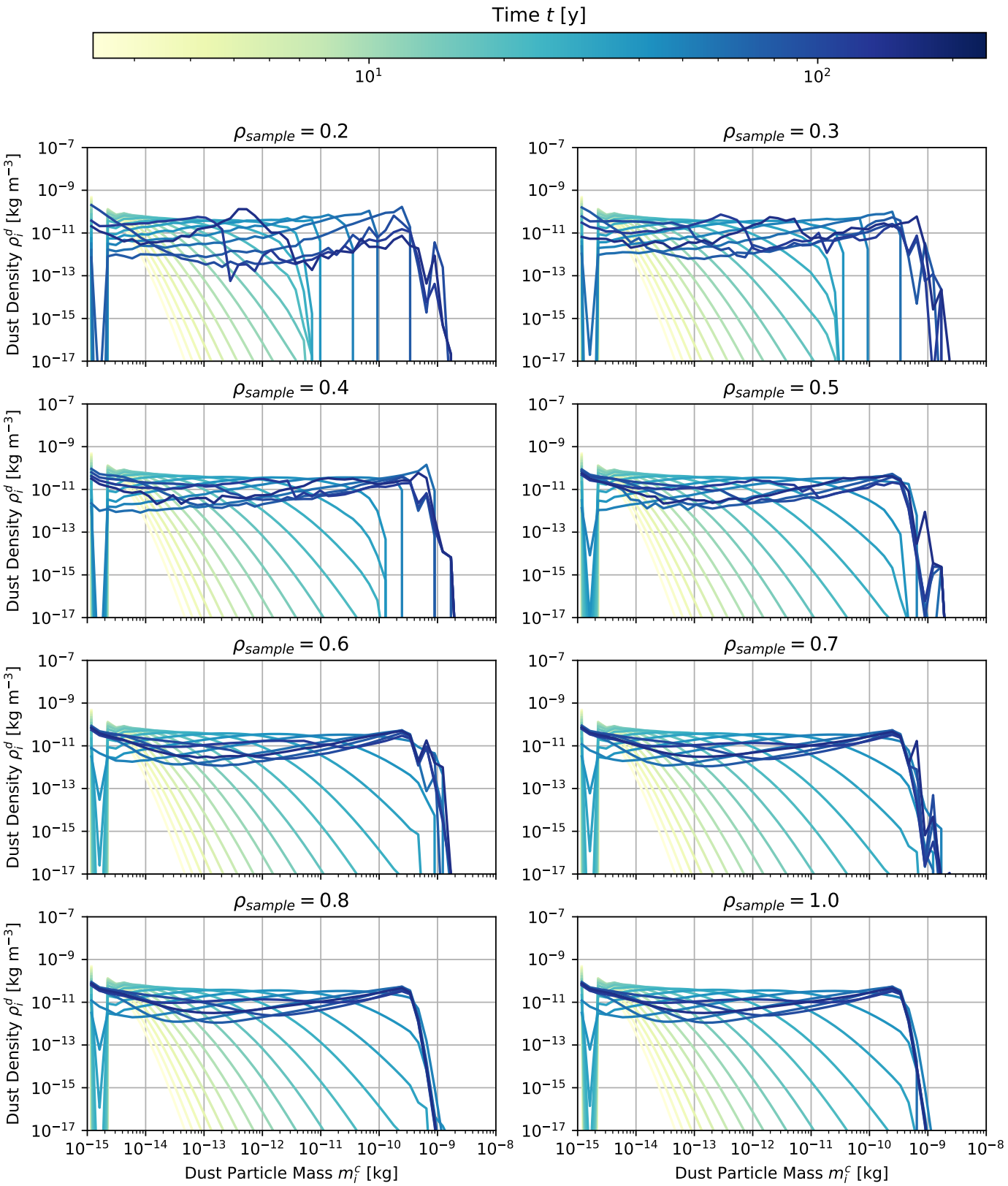


Figure 4.5: Temporal evolution of the dust particle mass distribution under the influence of both coagulation and fragmentation processes for various different values of the sampling density ρ_{sample}.

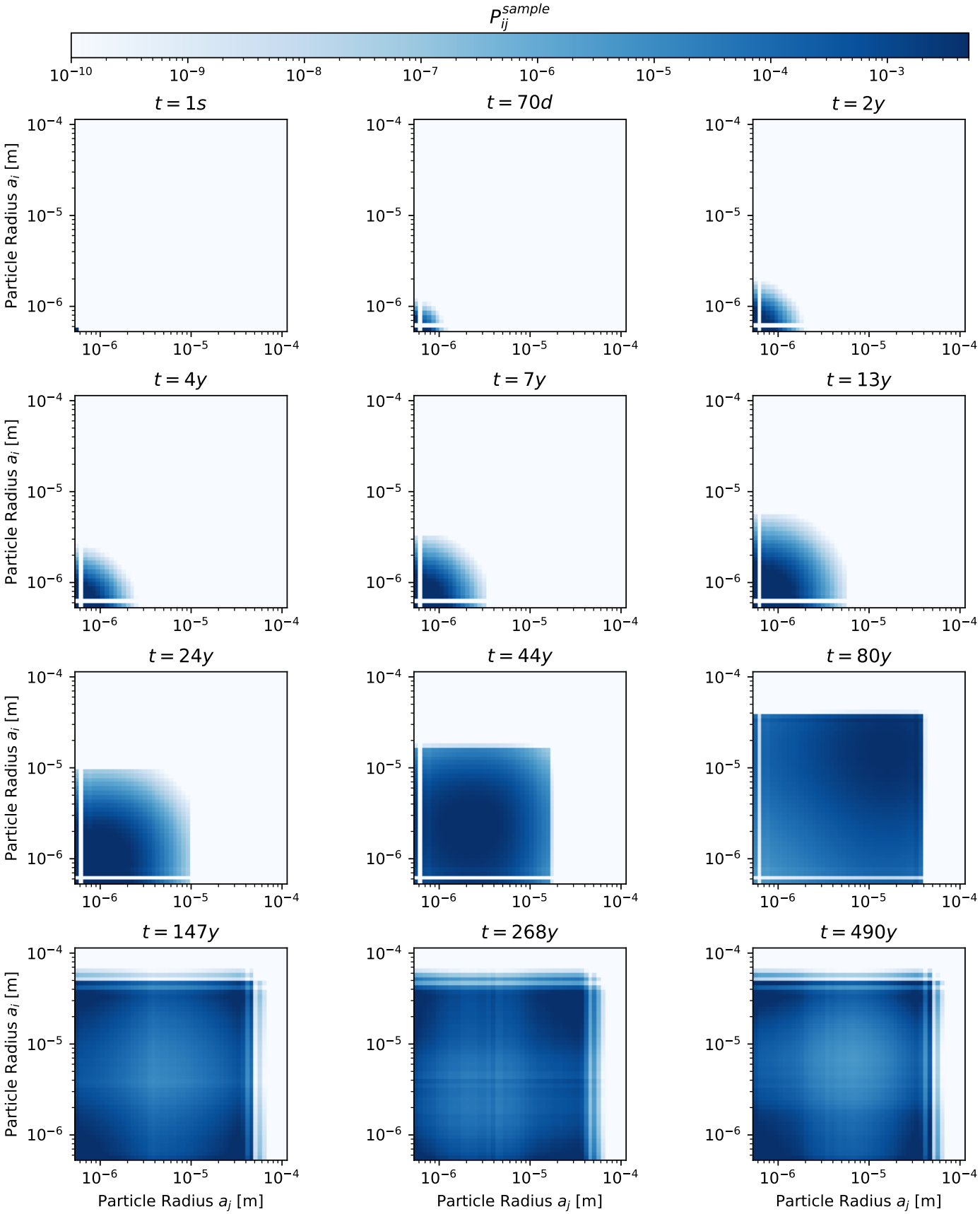


Figure 4.6: Illustration of the collision sampling probability distribution P_{ij}^{sample} for $\rho_{\text{sample}} = 0.5$. Here, both the process of coagulation as well as fragmentation are included into the model.

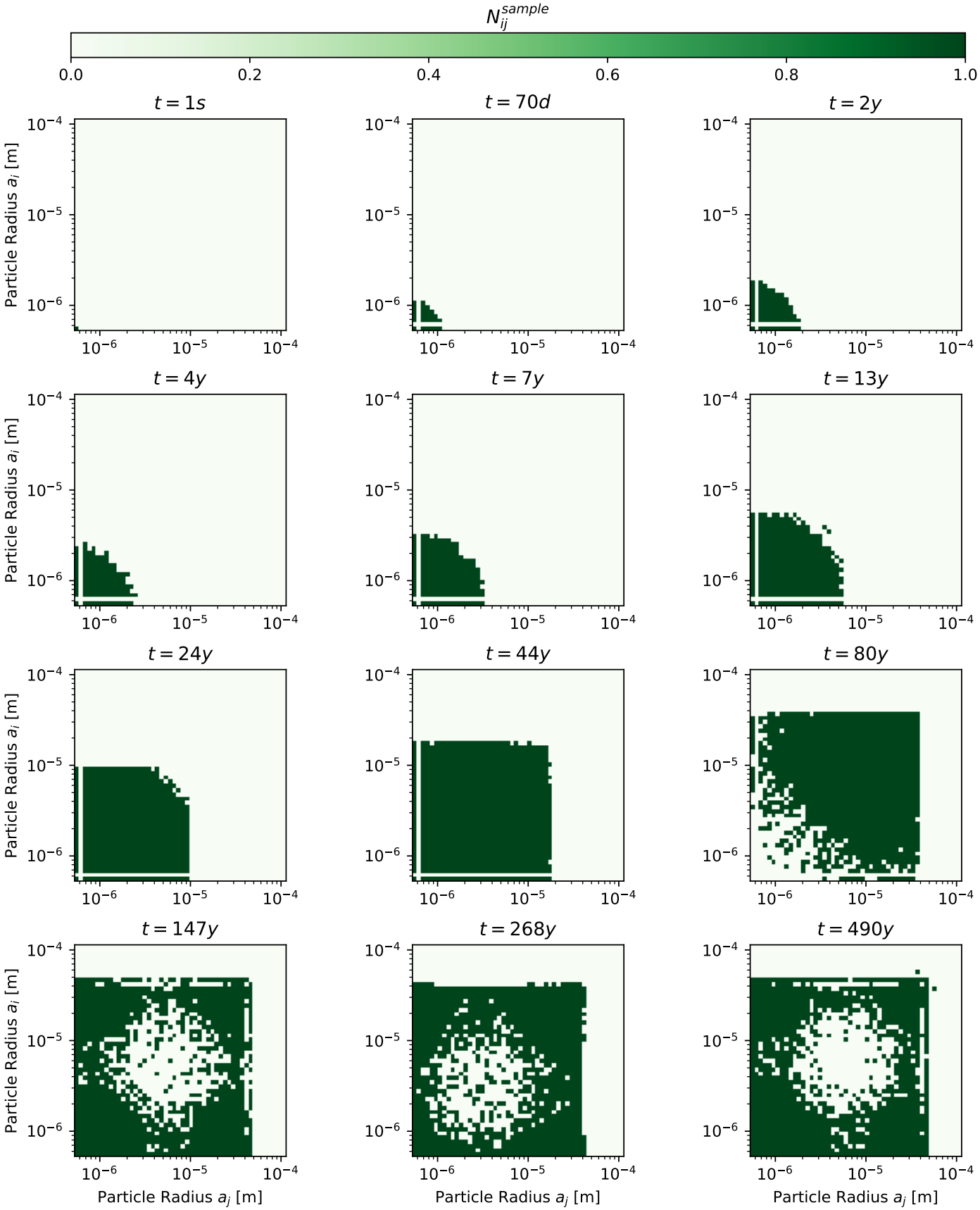


Figure 4.7: Illustration of the amount of times N_{ij}^{sample} that each collision (i, j) is sampled, for $\rho_{\text{sample}} = 0.5$. Here, both the process of coagulation as well as fragmentation are included into the model.

4.2.5 Accuracy and Stability

Additionally to the stability error Δ_{stab} that we defined in equation (3.70), and which is displayed again below, the following definition of Δ_{acc} gives an insight into the accuracy properties of the integration:

$$\Delta_{\text{stab}}(t) = \sum_i \frac{\rho_i^d(t) - \rho_i^d(t=0)}{\rho_i^d(t=0)} \tag{4.20}$$

$$\Delta_{\text{acc}}(t) = \sqrt{\sum_i \left(\frac{\rho_i^d(t) - \rho_i^c(t)}{\rho_i^c(t)} \right)^2} \tag{4.21}$$

Here, $\rho_i^c(t)$ denotes the fraction of the dust mass volume density contained in a given bin i when integrating the model using the *complete* kernel (hence the superfix "c").

Instead of comparing the total mass density value summed over all bins with its initial value at the start of the simulation, in the definition of the accuracy error we calculate the root-mean-square of the relative deviation (at each mass bin) between the numerical solution arrived at using the sampled kernel, and the complete one. The evolution of these two quantities with respect to time is displayed in figure 4.8. In green, we have the stability or mass error (where a dashed line signifies a negative value), and in blue we have the accuracy error that we just defined.[3]

While the stability error could be kept near machine precision for the entirety of the simulation, the accuracy error is much greater than 1. This could be explained partly by statistical fluctuations induced by the stochastic sampling of the kernel, and partly by the effectively lower reaction rates, causing the mass distribution to "lag" behind the solution calculated with the full kernel.

Since fragmentation process only start to be relevant in the latter parts of the integration, the statistical fluctuations can be assumed to effect the accuracy error most significantly at the end as well, whereas the lower reactions rates might well be the more significant contributor to the error in the early stages of the simulation.

This is due to the fact that, apart from a possibly slightly slower evolution, for pure coagulation we can reproduce the results from a complete kernel quite well even for lower sampling densities. This stems of course from the sparsity of the kernel for pure hit-and-stick coagulation.

In any case, while the results from the previous pages show a satisfactory *qualitative* behavior, the accuracy error supplies reasons for doubting the utility of this method when low deviations from the complete kernel solution is to be achieved.

Still, it could be shown that the core idea of the method can be used to reproduce the desired results in at least an approximate fashion.

[3]Note that there a missing values in the evolution of the accuracy error, which occur when the mass distribution in the "complete" solution contains zeros. This is the case in the beginning, before coagulation processes can transfer mass into each bin, and again at near the end, when the lower-mass bins empty before fragmentation fills them up again.

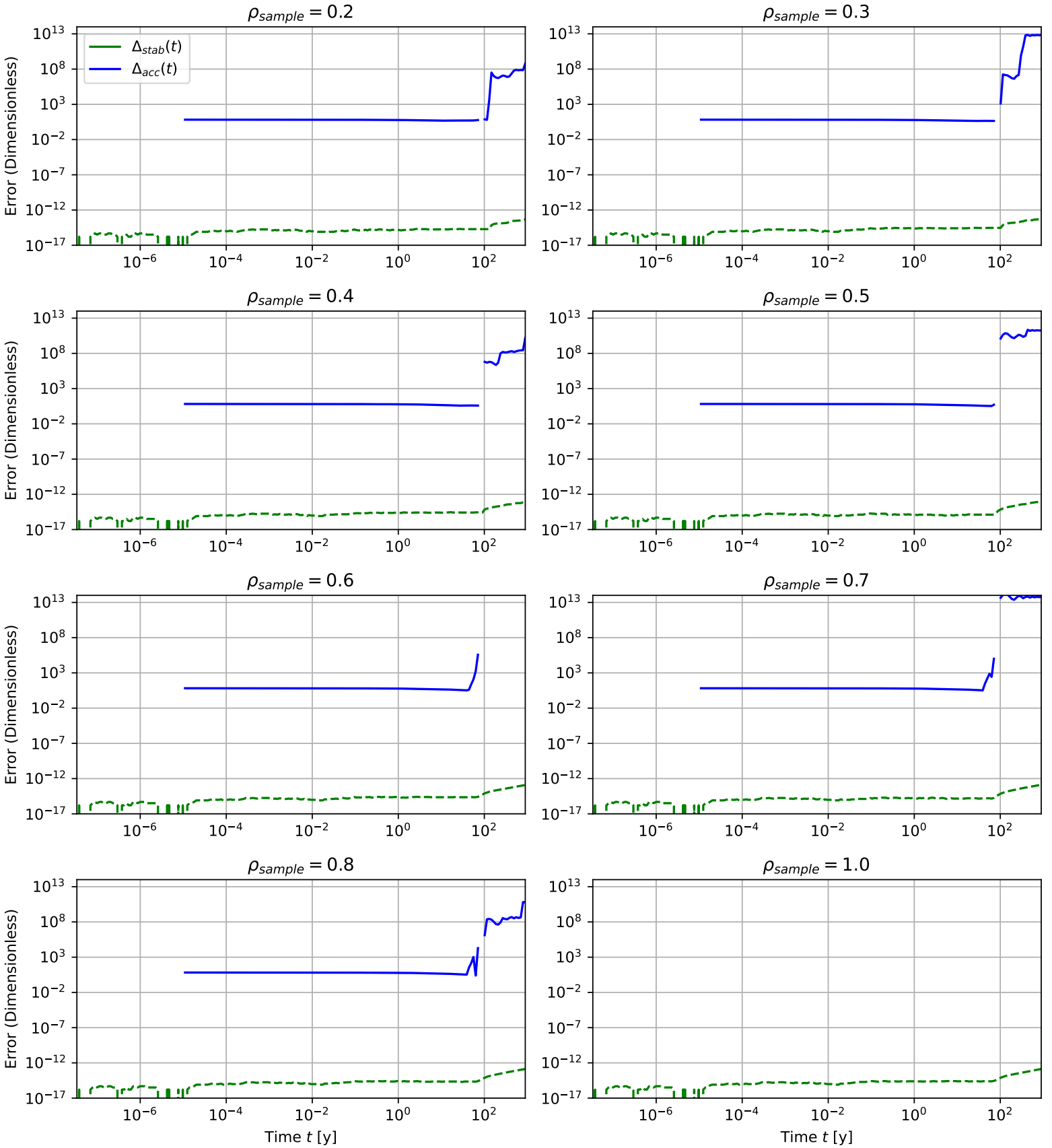


Figure 4.8: Accuracy error Δ_{acc} and stability error Δ_{stab} for the numerical integration of the Smoluchowski coagulation equation for various sampling densities ρ_{sample}. Stability is assured down to machine precision, accuracy on the other hand is negatively impacted by statistical fluctuations arising from the stochastic nature of the Monte Carlo method. For the stability error, the dashed line signifies a negative value.

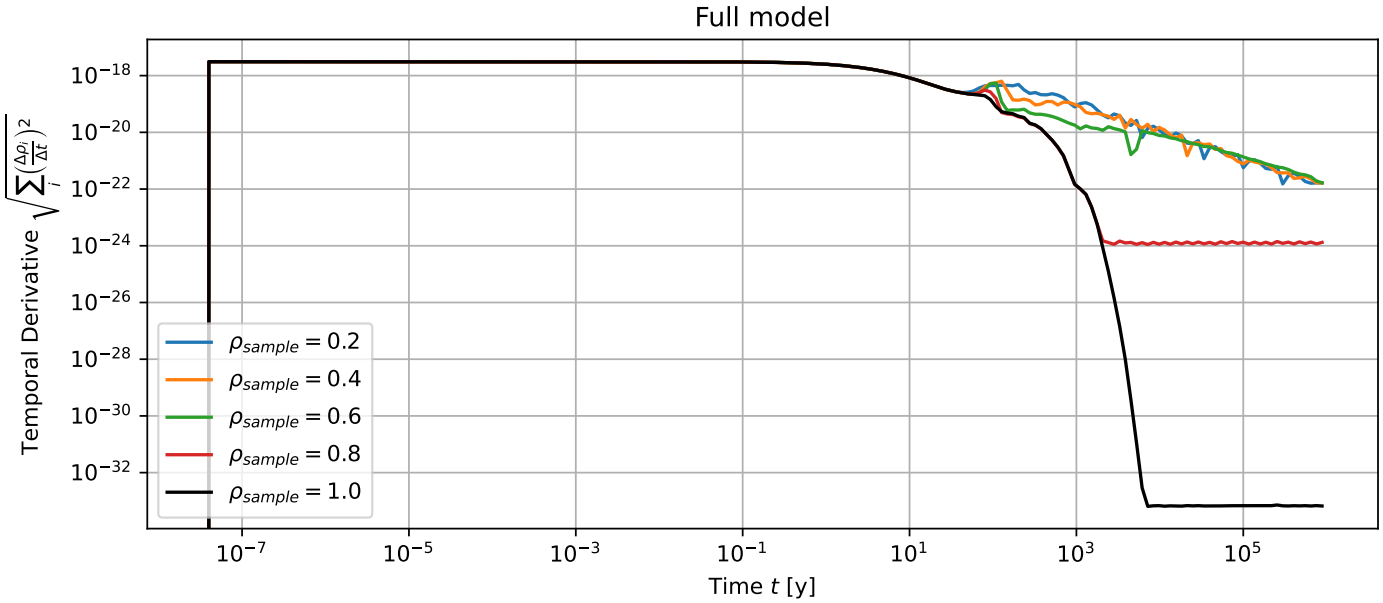


Figure 4.9: Temporal evolution of the root-mean-square (RMS) of the temporal derivative of the mass distribution. Here, both the process of coagulation as well as fragmentation are included into the model. Note that when sampling is enabled, the temporal derivative does not approach zero as quickly as when it is disabled. This is most likely due to the stochastic nature of the Monte Carlo sampling method, leading to statistical fluctuations that prevent the mass distribution from ultimately settling into a true equilibrium state.

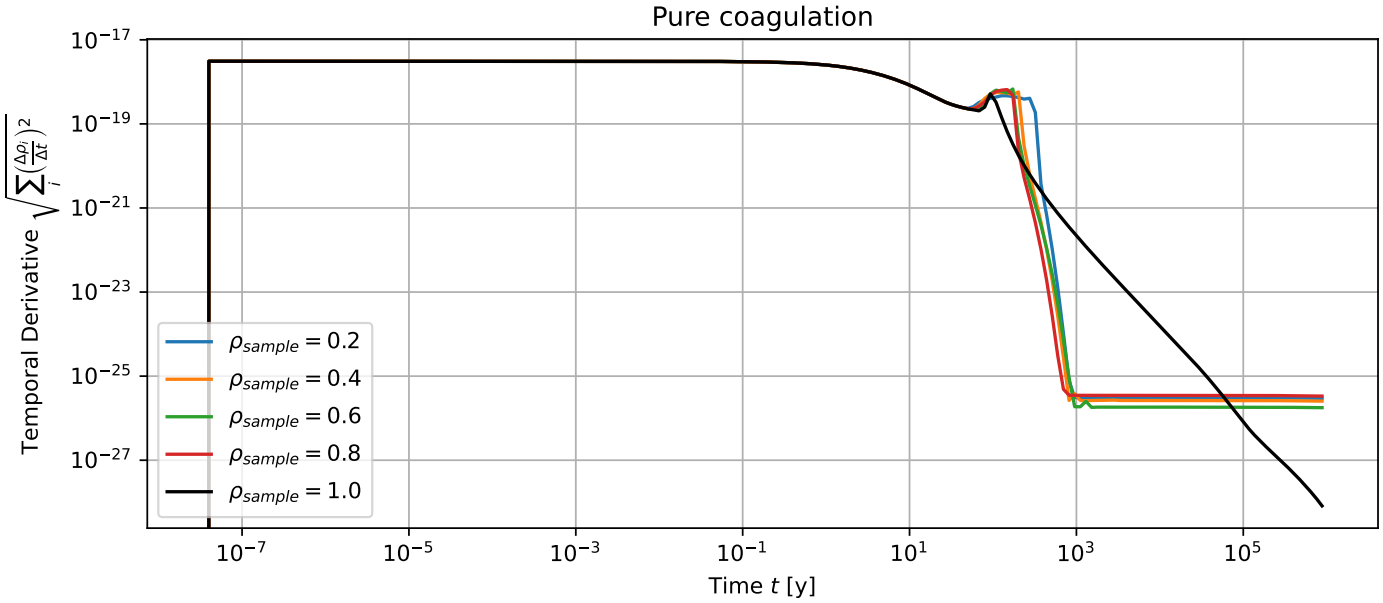


Figure 4.10: Temporal evolution of the root-mean-square (RMS) of the temporal derivative of the mass distribution. Here, only the process of pure hit-and-stick coagulation is included into the model. Interestingly, here the temporal derivatives that were arrived at using a sampled kernel initially approach zero much faster than when using the complete kernel. At some point, also here the temporal derivative of the "complete" solution overtakes the sampled soltions on its way towards zero.

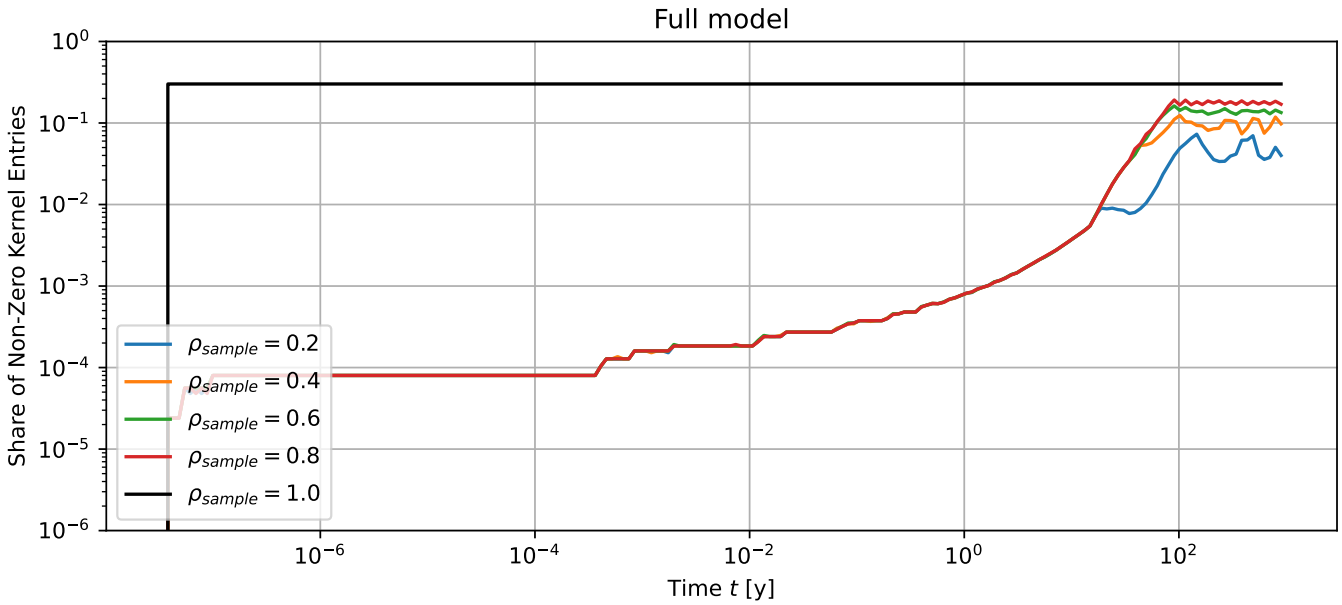


Figure 4.11: Percentage of non-zero entries in the sampled kernel matrix. Here, both coagulation and fragmentation are included in the model. If $\rho_{\text{sample}} = 1$, i.e. if all collisions are taken into account when defining the kernel, this ratio is approximately 30%. Since a lot of these collisions are neglected when a sampled kernel is used, the number of non-zero kernel entries decreases with lower values of the sampling density. This is especially true in the beginning, where fragmentation processes do not yet happen much. As time progresses, more and more entries of the kernel become relevant. As such, this ratio increases and, in the end, converges towards the value of the sampling density.

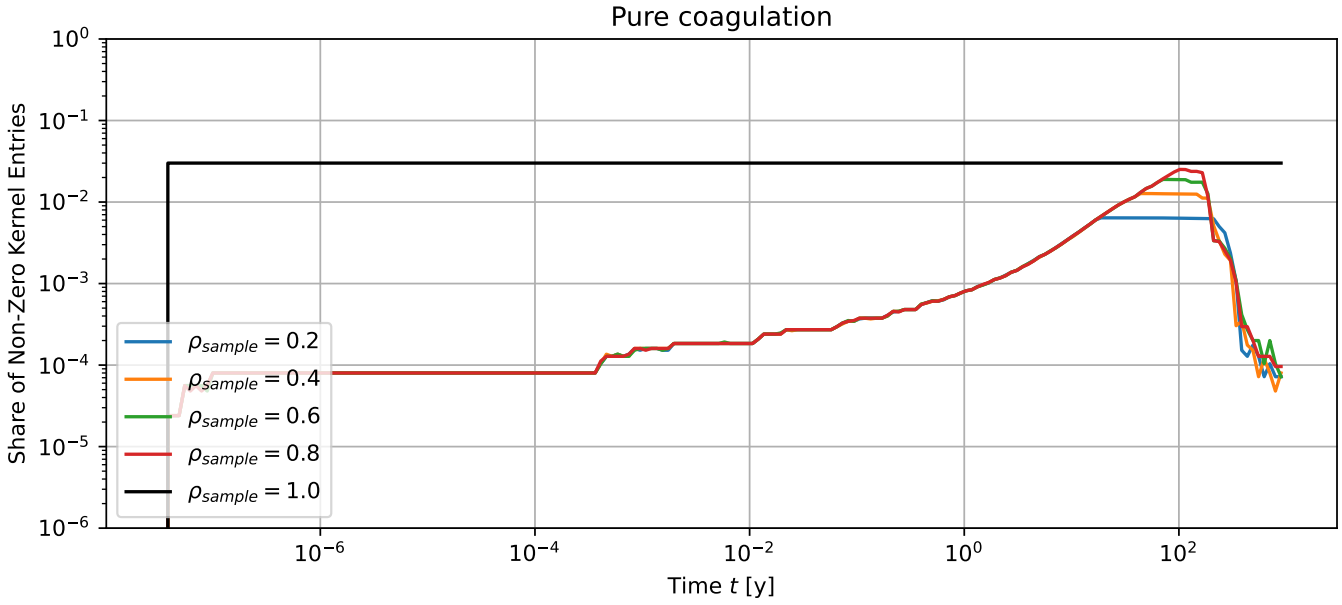


Figure 4.12: Percentage non-zero entries in the sampled kernel matrix. Here, only the process of pure stick-and-hit coagulation is included into the model. The number of non-zero entries increases as time progresses, until there is a steep drop-off when basically the entire mass has been transferred into the highest-mass bin. Overall, the number of non-zero kernel entries is much lower here than in the full model, reflecting the fact that due to the Dirac δ-function in equation (3.21), the definition of the pure coagulation kernel yields a much sparser matrix than one would have if fragmentation processes are included, with only approximately 3% of the kernel being non-zero for $\rho_{\text{sample}} = 1$.

Chapter 5

Summary & Discussion

For this thesis, we set ourselves the goal of attempting to shed light on the feasibility of lowering the computational cost of the Smoluchoswki equation's numerical integration via the utilization of stochastic Monte Carlo sampling of the kernel matrix. Specifically, the goal was to identify the most relevant particle collisions, and only include these into the summation over the kernel matrix.

Before we could start on that, a number of preliminary definitions had to be made. To be able to study the evolution of the distribution of dust particle masses in a proto-planetary disk, it was of course necessary to define a disk model, which we did in chapter 2. Here, we defined the disk's main attributes, which include the spatial distribution of gas and dust within it. Our next point of attention was then given by the task of formulating a model for the gas and dust kinematics, which later on allowed us to define the rate of collisions between differently-sized dust particles, as well as velocity-dependent probabilities for the collision outcomes. This concluded the first part of this thesis.

The second part of the thesis focused on the definition of a model for circumstellar dust particle interaction. In chapter 3, we adopted the formalism given by the Smoluchoswki coagulation equation, which provides a method for calculating the temporal evolution of the distribution of dust particle masses under the influence of collisions with other particles. Before this temporal evolution behavior could be determined via the numerical integration of the Smoluchoswki coagulation equation, it was necessary to build a discretized analogon to the model we discussed so far, which included the construction of the kernel matrices for both dust coagulation and fragmentation processes. Using a 4th order implicit Radau integration scheme, findings from earlier works rearding the temporal evolutions of the dust particle mass distribution could be reproduced.

After having spent a lot of time on preparatory work, in chapter 4 it was finally possible to start with the third part of the thesis, in which studies regarding a sampling of the kernel were made. The idea here was to include only the most "relevant" dust particle collisions into the coagulation model, leading to a reduction of terms that need to be considered in the sum evaluation that is done on the right-hand side of the discretized Smoluchoswki equation. Relevance here is determined by the impact that a given collision has on the particle mass distribution. Only these most relevant collisions were then used to construct a new, incomplete kernel matrix, which was then used to forward the mass distribution by one time-step via the integration of the Smoluchoswki coagulation equation.

For this, a sensible sampling probability distribution had to be defined. The approach used here for the definition is quite straightforward, and could be improved on by future works attempting to use this type of Monte Carlo method. Here, one possible improvement could be given by adjusting the sampling density in such a way that the appearance of peak-like structures at the high-mass end of the equilibrium distribution can be suppressed. Another suggestion for an improvement would be the implementation of a sampling method that allows duplicate selection of collisions, with the definition of an appropriate normalization.

To quantify the quality of the results we arrived at using the Monte Carlo sampling method, it makes sense to focus on three main metrics. These are given by the *stability*, the *accuracy*, as well as the *numerical cost* of the algorithm.

As we discussed earlier, the first of these could be ensured in a satisfactory manner. In the case of pure stick-and-hit coagulation, with the help of the method provided by the Kovetz-Olund algorithm, as well as the implementation of a method for handling near-zero cancellation in floating arithmetics, the total mass error of the integration could be ensured down to machine precision of 64-bit floating point number operations.

When fragmentation processes were included into the model, an additional error was introduced. Still, for the simulations done in the context of this thesis, where the dust particle mass distribution function enters an equilibrium after a few hundred years, the relative mass error of the integration could be kept well below 1, growing no larger than 10^{-12}. This is true both when the complete/unsampled kernel was used, as well as when large parts of the kernel were neglected in the evaluation of the double sum for the numerical integration. As such, we can confidently say that mass conservation in detailed balance could be ensured.

Now, with regards to accuracy of the algorithm: Even when parts of the kernel are left out when evaluating the sum on the right-hand side of the discretized Smoluchowski coagulation equation, the general behavior of the mass distribution under time can approximately be reproduced. When both coagulation and fragmentation are included, the distribution does eventually evolve into the expected equilibrium state.

It must be noted though, that fluctuations arising from the stochastic nature of the method keep the distribution's temporal derivative well above zero. As one might expect, the magnitude of these fluctuations depends on the value of the sampling density. Also, since leaving out some of the collisions effectively leads to a lower total collision rate, the evolution progresses more slowly for lower sampling densities.

Both of these effects negatively impact the accuracy of the numerical solutions derived from a sampled kernel, when compared to the complete one.

Because a lower sampling density leads to a larger accuracy error, the question of whether the sampling approach examined in this thesis should be used likely cannot be answered in a definitive way. Instead, whether it makes sense to use this collision sampling approach depends on the specific requirements that are imposed on the computational resources, as well as the desired accuracy, since there is a trade-off between these two quantities.

When the dust particles are described as entities whose only property is the value of their mass, this sampling approach might only really make sense for very high grid resolutions, or when only an approximate, first look at the results is desired.

If, on the other hand, more attributes (like e.g. particle porosity) are included into the model, the utility value of this method could increase.

As the amount of noise in the kernel may be expected to grow faster with the dimensionality of the problem than the number of relevant kernel entries, for studies of more sophisticated models including additional particle attributes, it may well be worthwhile to consider this approach.

Bibliography

[Adv06] Advanced Camera for Surveys (ACS) aboard NASA/ESA Hubble Space Telescope. *Hubble's sharpest view of the Orion Nebula.* [Online; accessed February 12, 2024]. 2006. URL: https://esahubble.org/images/heic0601a/.

[ALM14] ALMA (ESO/NAOJ/NRAO). *ALMA image of the protoplanetary disc around HL Tauri.* [Online; accessed February 12, 2024]. 2014. URL: https://www.eso.org/public/images/eso1436a/.

[BD23] Beutel, M. and Dullemond, C. P. "An improved Representative Particle Monte Carlo method for the simulation of particle growth". In: *A&A* 670 (2023), A134. DOI: 10.1051/0004-6361/202244955. URL: https://doi.org/10.1051/0004-6361/202244955.

[BDB10] Birnstiel, T., Dullemond, C. P. and Brauer, F. "Gas- and dust evolution in protoplanetary disks". In: *A&A* 513 (2010), A79. DOI: 10.1051/0004-6361/200913731. URL: https://doi.org/10.1051/0004-6361/200913731.

[BDH08] Brauer, F., Dullemond, C. P. and Henning, Th. "Coagulation, fragmentation and radial motion of solid particles in protoplanetary disks". In: *A&A* 480.3 (2008), pp. 859–877. DOI: 10.1051/0004-6361:20077759. URL: https://doi.org/10.1051/0004-6361:20077759.

[Bur+13] Michael Burton et al. *The Interstellar Medium White Paper.* 2013. arXiv: 1307.0712 [astro-ph.GA].

[CFL28] Courant, R., Friedrichs, K. and Lewy, H. "Über die partiellen Differenzengleichungen der mathematischen Physik". In: *Math. Ann.* 100 (1928), pp. 891–921. DOI: https://doi.org/10.1007/BF01448839.

[DD04] Dullemond, C. P. and Dominik, C. "The effect of dust settling on the appearance of protoplanetary disks". In: *A&A* 421.3 (2004), pp. 1075–1086. DOI: 10.1051/0004-6361:20040284. URL: https://doi.org/10.1051/0004-6361:20040284.

[DD23] Carsten Dominik and Cornelis Dullemond. *The bouncing barrier revisited: Impact on key planet formation processes and observational signatures.* 2023. arXiv: 2312.06000 [astro-ph.EP].

[Eur09] European Southern Observatory (ESO). *Three-colour composite mosaic image of the Eagle Nebula (Messier 16, or NGC 6611), based on images obtained with the Wide-Field Imager camera on the MPG/ESO 2.2-metre telescope at the La Silla Observatory.* [Online; accessed February 12, 2024]. 2009. URL: https://www.eso.org/public/images/eso0926a/.

[Kat17] Akimasa Kataoka. "Dust Coagulation with Porosity Evolution". In: *Formation, Evolution, and Dynamics of Young Solar Systems.* Ed. by Martin Pessah and Oliver Gressel. Vol. 445. Astrophysics and Space Science Library. Jan. 2017, p. 143. DOI: 10.1007/978-3-319-60609-5_5.

[KO69] A. Kovetz and B. Olund. "The Effect of Coalescence and Condensation on Rain Formation in a Cloud of Finite Vertical Extent". In: *Journal of Atmospheric Sciences* 26.5 (1969), pp. 1060–1065. DOI: https://doi.org/10.1175/1520-0469(1969)026<1060:TEOCAC>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/atsc/26/5/1520-0469_1969_026_1060_teocac_2_0_co_2.xml.

[LL21] Guillaume Laibe and Maxime Lombart. "On the Courant–Friedrichs–Lewy condition for numerical solvers of the coagulation equation". In: *Monthly Notices of the Royal Astronomical Society* 510.4 (Dec. 2021), pp. 5220–5225. ISSN: 0035-8711. DOI: 10.1093/mnras/stab3499. eprint: https://academic.oup.com/mnras/article-pdf/510/4/5220/42253879/stab3499.pdf. URL: https://doi.org/10.1093/mnras/stab3499.

[LP74] D. Lynden-Bell and J. E. Pringle. "The evolution of viscous discs and the origin of the nebular variables." In: 168 (Sept. 1974), pp. 603–637. DOI: 10.1093/mnras/168.3.603.

[MRN77] J. S. Mathis, W. Rumpl and K. H. Nordsieck. "The size distribution of interstellar grains." In: *Astrophysical Journal* 217 (Oct. 1977), pp. 425–433. DOI: 10.1086/155591.

[MTH17] Sekiyama Makoto, Ohtsuki Toshiya and Yamamoto Hiroshi. "Analytical Solution of Smoluchowski Equations in Aggregation–Fragmentation Processes". In: *Journal of the Physical Society of Japan* 86 (2017), p. 104003. URL: https://api.semanticscholar.org/CorpusID:125747292.

[Nat+06] A. Natta et al. *Dust in Proto-Planetary Disks: Properties and Evolution*. 2006. arXiv: astro-ph/0602041 [astro-ph].

[NSH86] Y. Nakagawa, M. Sekiya and C. Hayashi. "Settling and growth of dust particles in a laminar phase of a low-mass solar nebula". In: 67.3 (Sept. 1986), pp. 375–390. DOI: 10.1016/0019-1035(86)90121-1.

[OC07] C. W. Ormel and J. N. Cuzzi. "Closed-form expressions for particle relative velocities induced by turbulence". In: *Astronomy & Astrophysics* 466.2 (Feb. 2007), pp. 413–420. DOI: 10.1051/0004-6361:20066899. URL: https://doi.org/10.1051%2F0004-6361%3A20066899.

[ONN90] Keiji Ohtsuki, Yoshitsugu Nakagawa and Kiyoshi Nakazawa. "Artificial acceleration in accumulation due to coarse mass-coordinate divisions in numerical simulation". In: *Icarus* 83.1 (1990), pp. 205–215. ISSN: 0019-1035. DOI: https://doi.org/10.1016/0019-1035(90)90015-2. URL: https://www.sciencedirect.com/science/article/pii/0019103590900152.

[OST06] C. W. Ormel, M. Spaans and A. G. G. M. Tielens. "Dust coagulation in protoplanetary disks: porosity matters". In: *Astronomy & Astrophysics* 461.1 (Oct. 2006), pp. 215–232. ISSN: 1432-0746. DOI: 10.1051/0004-6361:20065949. URL: http://dx.doi.org/10.1051/0004-6361:20065949.

[OTS09] Satoshi Okuzumi, Hidekazu Tanaka and Masa-aki Sakagami. "NUMERICAL MODELING OF THE COAGULATION AND POROSITY EVOLUTION OF DUST AGGREGATES". In: *The Astrophysical Journal* 707.2 (Dec. 2009), pp. 1247–1263. ISSN: 1538-4357. DOI: 10.1088/0004-637x/707/2/1247. URL: http://dx.doi.org/10.1088/0004-637X/707/2/1247.

[S A16] S. Andrews (Harvard-Smithsonian CfA); B. Saxton (NRAO/AUI/NSF); ALMA (ESO/NAOJ/NRAO). *ALMA image of the disc around the young star TW Hydrae*. [Online; accessed February 12, 2024]. 2016. URL: https://www.eso.org/public/images/eso1611a/.

[Saf63] VS Safronov. "A restricted solution of the accretion equation". In: *Soviet Physics Doklady*. Vol. 7. 1963, p. 967.

[Saf72] V. S. Safronov. *Evolution of the protoplanetary cloud and formation of the earth and planets.* 1972.

[SB22] Sebastian M. Stammler and Tilman Birnstiel. "DustPy: A Python Package for Dust Evolution in Protoplanetary Disks". In: *The Astrophysical Journal* 935.1 (Aug. 2022), p. 35. DOI: 10.3847/1538-4357/ac7d58. URL: https://dx.doi.org/10.3847/1538-4357/ac7d58.

[Smo16] M. V. Smoluchowski. "Drei Vorträge über Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen". In: *Zeitschrift für Physik* 17 (Jan. 1916), pp. 557–585.

[Wad+09] Koji Wada et al. "COLLISIONAL GROWTH CONDITIONS FOR DUST AGGREGATES". In: *The Astrophysical Journal* 702.2 (Aug. 2009), p. 1490. DOI: 10.1088/0004-637X/702/2/1490. URL: https://dx.doi.org/10.1088/0004-637X/702/2/1490.

[Wat06] J. A. D. Wattis. "An introduction to mathematical models of coagulation–fragmentation processes: A discrete deterministic mean-field approach." In: *Physica D: Nonlinear Phenomena* 322.1–2 (2006), pp. 1–20. DOI: https://doi.org/10.1016/J.PHYSD.2006.07.024.

[Wei77] S. J. Weidenschilling. "Aerodynamics of solid bodies in the solar nebula." In: 180 (July 1977), pp. 57–70. DOI: 10.1093/mnras/180.2.57.