Spring 5-13-2023

Numerical Methods for Wave Turbulence: Isotropic 3-Wave Kinetic Equations

Steven Walton
Southern Methodist University, stevenw@smu.edu

NUMERICAL METHODS FOR WAVE TURBULENCE:

ISOTROPIC 3-WAVE KINETIC EQUATIONS

Approved by:

Dr. Alejandro Aceves
Professor of Mathematics, SMU

Dr. Thom Hagstrom
Professor of Mathematics, SMU

Dr. Benno Rumpf
Professor of Mathematics, SMU

Dr. Alain Bensoussan
Lars Magnus Ericsson Chair Professor,
UTD

Dr. Nathaniel Morgan
Scientist 5, E-2: Process Modeling and
Analysis, LANL

Dr. Minh-Binh Tran
Associate Professor of Mathematics,
TAMU

NUMERICAL METHODS FOR WAVE TURBULENCE: ISOTROPIC 3-WAVE KINETIC EQUATIONS

A Dissertation Presented to the Graduate Faculty of the

Dedman College

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Mathematics

by

Steven Walton

B.S., Applied Mathematics, The University of Texas at Dallas
M.S., Applied and Computational Mathematics, Southern Methodist University

May 13th, 2023

Acknowledgments

First, I must acknowledge the RTG program (NSF RTG Grant DMS-1840260) without which I would not have been able to pursue a PhD and supported the first 3 years of my studies in the mathematics department at SMU. The RTG program is headed by Alejandro Aceves who I have to thank for his constant support and encouragement. Next, I must thank Katie Hedrick for her patience, teaching and allowing me to participate in her research group in my first year at SMU and beyond. I'd like to thank Thom Hagstrom for his many useful discussions and advice, especially during the early stages of the deep learning solver for the Smoluchowski equation. I have to express gratitude toward Benno Rumpf for entertaining several meetings to discuss wave turbulence. My understanding of the subject has greatly benefited from the many questions he has posed.

I'm very grateful to Daniel Israel and others for organizing the computational physics workshop at Los Alamos National Laboratory (LANL) in the summer of 2022. It is difficult to exaggerate the benefits of this experience. Which brings me to Nathaniel Morgan and Svetlana Tokareva who were my mentors during the summer workshop but now also as a GRA student at LANL and will continue to be my mentors as a post-doc in T-5. Thank you both for being such supportive mentors and for the wealth of research topics you've introduced me to. Your impact on my early career is inestimable and I am truly grateful. On that tack, I need to express thanks to Angela Herring, my group leader in XCP-4, and Mark Shraad, the division leader, for their kind advice and willingness to listen. I also gratefully acknowledge the funding from the Laboratory Directed Research and Development (LDRD) program at LANL.

To my thesis committee, past and present. I would like to thank Enrique Zuazua for his very useful comments in preparation for my defense and in regards to the present work. To Alain Bensoussan, I want to express gratitude for his interest in my work in the first place, his many helpful suggestions for research directions which played quite a large role in the latter part of this thesis.

To Binh, who taught me first about wave turbulence, the topic of the present thesis, in addition to introducing me to so many other subjects and people. This work would have been impossible without him. I am forever indebted to his mentorship, his constant thoughtfulness and willingness to impart any lesson he has learned so that my time as a fledgling researcher might be successful. Truly, it is his guidance and insight which guided me to pursue the development of numerical methods and is the reason I am on my current path here at LANL.

To my wife, Dusti Chopelas, I want to simply say thank you for your endless support and sacrifice. It's over.

The Los Alamos unlimited release number is: LA-UR-23-24252.

Walton, Steven B.S., Applied Mathematics, The University of Texas at Dallas
M.S., Applied and Computational Mathematics, Southern Methodist University

Numerical Methods for Wave Turbulence:
Isotropic 3-Wave Kinetic Equations

Advisor: Dr. Alejandro Aceves
Doctor of Philosophy degree conferred May 13th, 2023
Dissertation completed April 19th, 2023

Wave turbulence theory has remained an active area of research since its inception in the early part of the last century. In the kinetic regime, the main objects of study are the wave kinetic equations. The breakthrough discovery of constant flux, time independent solutions by Zakharov in the late 1960's has allowed for the theories predictions to be verified both experimentally and computationally in a wide array of physical systems. However, there remain many open questions concerning the time dependent solutions of the wave kinetic equations. In this thesis, we aim to partially address this open area of the wave turbulence theory by providing numerical methods for the time dependent solutions of the isotropic 3-wave kinetic equations. The methods we develop herein are able to confirm previous analysis for time dependent solutions, specifically the behavior of the energy cascade of these solutions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table Page

To Dusti and Ila.

Chapter 1

Some Wave Turbulence Theory

Wave turbulence theory (WTT) has its origins in the early to mid part of the last century with the work of Peierls, Benney-Saffman, Hasselman, Benney-Newell, Zakharov and many others [1–7]. An example we are all familiar with to which wave turbulence theory (WTT) may be applied is that of a wind-driven sea. This scenario was considered by Klaus Hasselmann, the recent Nobel laureate, in 1962-1963 [3–5] wherein one of the first formal derivations of a wave kinetic equation (WKE) is given [3, 4]. The intent was to describe the transfer of energy among waves on the open ocean. For more discussion than what the present chapter provides on wave turbulence the reader is referred to the excellent review article by Newell and Rumpf [8] as well as the recent books [9, 10].

The 3-wave kinetic equation, the protagonist of this thesis and one of the most important classes of wave kinetic equations, reads (see [7, 11–14])

$$\begin{aligned} \partial_t f(t,p) &= \mathcal{Q}[f](t,p), \\ f(0,p) &= f_0(p), \end{aligned} \tag{1.1}$$

in which $f(t,p)$ is the nonnegative wave density at wavenumber $p \in \mathbb{R}^N$, $N \geq 2$; $f_0(p)$ is the initial condition. The quantity $\mathcal{Q}[f]$ describes pure resonance and is of the form

$$\mathcal{Q}[f](p) = \iint_{\mathbb{R}^{2N}} \Big[\mathcal{R}_{p,p_1,p_2}[f] - \mathcal{R}_{p_1,p,p_2}[f] - \mathcal{R}_{p_2,p,p_1}[f]\Big] d^N p_1 d^N p_2 \tag{1.2}$$

with

$$\mathcal{R}_{p,p_1,p_2}[f] := |V_{p,p_1,p_2}|^2 \delta(p - p_1 - p_2)\delta(\omega - \omega_1 - \omega_2)(f_1 f_2 - f f_1 - f f_2)$$

with the short-hand notation $f = f(t,p)$, $\omega = \omega(p)$ and $f_j = f(t,p_j)$, $\omega_j = \omega(p_j)$, for wavenumbers p, p_j, $j \in \{1,2\}$. The function $\omega(p)$ is the dispersion relation of the wave system. The 3-wave kinetic equation has a variety of applications from ocean waves, acoustic waves, gravity capillary waves to Bose-Einstein condensates and many others (see [7,13,15–17] and references therein).

In the isotropic case, we identify $f(t,p)$ with $f(t,\omega)$, the isotropic 3-wave kinetic equation can take the form (see the appendix of [18])

$$\begin{aligned} \partial_t f(t,\omega) &= Q[f](t,\omega), \quad (t,\omega) \in \mathbb{R}^+ \times \mathbb{R}^+, \\ f(0,p) &= f_0(p), \end{aligned} \tag{1.3}$$

where the collision operator is defined by

$$\begin{aligned} Q[f](t,\omega) &= \int_0^\infty \int_0^\infty [R(\omega,\omega_1,\omega_2) - R(\omega_1,\omega,\omega_2) - R(\omega_2,\omega_1,\omega)] d\omega_1 d\omega_2, \\ R(\omega,\omega_1,\omega_2) &:= \delta(\omega-\omega_1-\omega_2)\left[|U(\omega_1,\omega_2)| f_1 f_2 - |U(\omega,\omega_1)| f f_1 - |U(\omega,\omega_2)| f f_2\right], \end{aligned} \tag{1.4}$$

and U satisfies $|U(\omega_1,\omega_2)| = (\omega_1\omega_2)^{\gamma/2}$, in which γ is a non-negative constant which plays an important role in the sequel. In what follows, we identify the variable k with the variable ω as k is simply ω multiplied by a fixed constant. Starting from (4.1), we could rewrite the 3-wave kinetic equation under the following equivalent form

$$\begin{aligned} \partial_t f(t,k) &= \mathbb{Q}[f](t,k), \quad (t,k) \in \mathbb{R}^+ \times \mathbb{R}^+, \\ f(0,k) &= f_0(k), \end{aligned} \tag{1.5}$$

in which $\mathbb{Q}$ is the collision operator defined by

$$\begin{aligned} \mathbb{Q}[f](t,k) \;=\; & \int_0^k [a(k_1,k-k_1) f(k_1) f(k-k_1) - a(k,k_1) f(k) f(k_1) \; - a(k,k-k_1) f(k) f(k-k_1)] \mathrm{d}k_1 \\ & - 2\int_0^\infty [a(k,k_1) f(k) f(k_1) - a(k+k_1,k_1) f(k+k_1) f(k_1) \; - a(k+k_1,k) f(k) f(k_1)] \mathrm{d}k_1, \end{aligned} \tag{1.6}$$

where the collision kernel satisfies $a(k_1, k_2) = (k_1 k_2)^{\gamma/2}$.

To see that this is true, consider

$$\begin{aligned} R(\omega, \omega_1, \omega_2) &:= \delta(\omega - \omega_1 - \omega_2) \left[|U(\omega_1, \omega_2)| f_1 f_2 - |U(\omega, \omega_1)| f f_1 - |U(\omega, \omega_2)| f f_2 \right] , \\ R(\omega_1, \omega, \omega_2) &:= \delta(\omega_1 - \omega - \omega_2) \left[|U(\omega, \omega_2)| f f_2 - |U(\omega_1, \omega)| f_1 f - |U(\omega_1, \omega_2)| f_1 f_2 \right] , \\ R(\omega_2, \omega_1, \omega) &:= \delta(\omega_2 - \omega_1 - \omega) \left[|U(\omega_1, \omega)| f_1 f - |U(\omega_2, \omega_1)| f_2 f_1 - |U(\omega_2, \omega)| f_2 f \right] . \end{aligned} \tag{1.7}$$

In the first and third line, apply the delta function to the ω_2 integral, which sets $\omega_2 = \omega - \omega_1$ and $\omega_2 = \omega + \omega_1$, respectively. In the second line, apply the delta function to the ω_1 integral so that $\omega_1 = \omega + \omega_2$. We then obtain for the first integral, considering that ω_2 is constrained to be positive

$$\int_0^\omega \Big[|U(\omega_1, \omega - \omega_1)| f_1 f_{-1} - |U(\omega, \omega_1)| f f_1 - |U(\omega, \omega - \omega_1)| f f_{-1} \Big] \mathrm{d}\omega_1, \tag{1.8}$$

where we used the notation $f_{-1} = f(\omega - \omega_1)$. The second integral is given by

$$\int_0^\infty \Big[|U(\omega, \omega_2)| f f_2 - |U(\omega + \omega_2, \omega)| f_{+2} f - |U(\omega + \omega_2, \omega_2)| f_{+2} f_2 \Big] \mathrm{d}\omega_2, \tag{1.9}$$

where $f_{+2} = f(\omega + \omega_2)$. Finally, the third integral is given by

$$\int_0^\infty \Big[|U(\omega_1, \omega)| f_1 f - |U(\omega + \omega_1, \omega_1)| f_{+1} f_1 - |U(\omega + \omega_1, \omega)| f_{+1} f \Big] \mathrm{d}\omega_1, \tag{1.10}$$

where $f_{+1} = f(\omega + \omega_1)$. We can change variables in the second integral so that $\omega_2 \to \omega_1$, we then arrive at

$$\begin{aligned} Q[f](t, \omega) = &\int_0^\omega \Big[|U(\omega_1, \omega - \omega_1)| f_1 f_{-1} - |U(\omega, \omega_1)| f f_1 - |U(\omega, \omega - \omega_1)| f f_{-1} \Big] \mathrm{d}\omega_1 \\ &-2 \int_0^\infty \Big[|U(\omega_1, \omega)| f_1 f - |U(\omega + \omega_1, \omega_1)| f_{+1} f_1 - |U(\omega + \omega_1, \omega)| f_{+1} f \Big] \mathrm{d}\omega_1, \end{aligned} \tag{1.11}$$

which after identifying k with ω, as mentioned above, and using the definition of $a(k_1, k_2)$ provided above we arrive at the collision operator (4.3).

Zakharov and Filonenko 1967 [7] were able to derive analytically constant flux solutions to the wave kinetic equation describing capillary wave systems. These solutions are now referred to as the Kolmogorov-Zakharov solutions or K-Z spectra, which is a class of *time-independent solutions* f_∞ of equation (1.1):

$$f_\infty(p) \approx C|p|^{-\kappa}, \quad \kappa > 0,$$

and have been found for many other systems [7,9,19,20]. These solutions are analogous to the Kolmogorov 5/3 law of hydrodynamic turbulence with the important difference that the value of κ in the weak turbulence theory is dependent upon the wave system under consideration. The K-Z spectra describe constant flux of some conserved quantity (here, the energy density) from low to high wavenumbers (direct cascade) and from high to low wavenumbers (inverse cascade) within some inertial interval over the wave numbers $p \in [p_f, p_d]$ where p_f denotes the upper bound for the forcing region and p_d denotes the beginning of the dissipation region. For these constant flux solutions, in many cases, there is some region in the wavenumber domain where the K-Z spectra breaks down [21, 22] and ceases to be valid, attributed to a breakdown in the weak coupling hypothesis, i.e. the assumption that the linear and nonlinear time-scales are well separated is violated at small or large values of wavenumbers (see [23]).

Research in accordance with these stationary solutions has actively continued to the present. Its predictions have been verified numerically and experimentally [24, 25], and applications have been found in weather forecasting, climate modeling, magnetohydrodynamics and gravitational waves [26–29] as well as in local field potential recordings in the hippocampus [30].

However, in absence of forcing and dissipation, the K-Z solution scaling is only expected for infinite capacity systems, e.g. for a forward cascade process such systems that the energy integral $\int \omega(p) f_\infty(p) \mathrm{d}p$ diverges at infinite p. In the opposite case of finite capacity systems, again using the forward cascade process as an example, the energy integral converges at $p = \infty$, indicating only a finite amount of energy can be contained by the system under consideration. It has been argued that once that capacity is met, a sink "opens up" at $p = \infty$ to absorb the extra energy [21] in some finite time t^* which destroys conservation of the energy density. This means that for finite capacity systems, energy conserving solutions are only local in time. In the paragraphs to come discussing time-dependent solutions to (4.1), we call this time, t^*, the blow-up time.

Beyond constant flux solutions, comparatively little has been done on the time-dependent solutions of (1.1). Connaughton and co-authors [18, 31, 32] designed numerical experiments to investigate the isotropic 3-wave equation. In [18], it was pointed out that isotropic 3-wave kinetic equations are equivalent to mean field rate equations for an aggregation-fragmentation problem which possesses an unusual fragmentation mechanism. A numerical method for solving isotropic 3-wave kinetic equations, with forcing and dissipation present, was also introduced in the same work.

The global existence of 3-wave kinetic equations for capillary wave systems in the presence of forcing has been done in [33]. An important work [34] by Gamba et. al. considers the global existence and uniqueness of strong solutions of 3-wave kinetic equations with near resonance conditions, which was inspired by earlier work of the same authors [35] and the related important works [36–41]. An interesting connection to chemical reaction networks was also pointed out in [42].

In [43], Soffer and Tran show that, the energy conserved time-dependent solutions of (4.1), in the finite capacity case (with $\gamma > 1$ in equation (4.1) and, equivalently (4.2)), exhibit the property that the energy is cascaded from small wavenumbers to large wavenumbers. They

show that for a regular initial condition whose energy at infinity, $\omega = \infty$, is initially 0, as time evolves, the energy is gradually accumulated at $\{\omega = \infty\}$. In the long time limit, all the energy of the system is concentrated at $\{\omega = \infty\}$ and the energy function becomes a Dirac function at infinity $\mathcal{E}\delta_{\{\omega=\infty\}}$, where $\mathcal{E}$ is the total energy. To be more precise, let us define the energy of the solution (4.1) as $g(t,\omega) = \omega f(t,\omega)$. It has been proved in [43] that g can be decomposed into two parts

$$g(t,\omega) = \bar{g}(t,\omega) + \tilde{g}(t)\delta_{\{\omega=\infty\}}, \tag{1.12}$$

where $\bar{g}(t,\omega) \geq 0$ is the regular part, which is a function, and $\tilde{g}(t)\delta_{\{\omega=\infty\}}$, is the singular part, which is a measure. The function $\tilde{g}(t)$ is non-negative. Initially, $\bar{g}(0,\omega) = g(0,\omega)$ and $\tilde{g}(0) = 0$. But, there exists a blow-up time t_1^*, such that for all time $t > t_1^*$, the function $\tilde{g}(t)$ is strictly positive. Moreover, starting from time t_1^*, there exists *infinitely many blow-up times*

$$0 < t_1^* < t_2^* < \cdots < t_n^* < \cdots, \tag{1.13}$$

such that

$$\bar{g}(t_1^*,\omega) > \bar{g}(t_2^*,\omega) > \cdots > \bar{g}(t_n^*,\omega) > \cdots \to 0, \tag{1.14}$$

and

$$0 < \tilde{g}(t_1^*) < \tilde{g}(t_2^*) < \cdots < \tilde{g}(t_n^*) < \cdots \tag{1.15}$$

The phenomenon has been explained in [43] that, after the first blow-up time, t_1^*, the energy starts to transfer from the regular part $\bar{g}(t,\omega)$ to the singular part $\tilde{g}(t)\delta_{\omega=\infty}$ at a rate at least like $\mathcal{O}(\frac{1}{\sqrt{t}})$, while the total energy of the two regular and singular parts is still conserved. This decay rate has been obtained in item (iii) of the main theorem of [43] (Theorem 10, pages 22-45), which states as follows. The energy cascade has an explicit rate $\int_{\{|p|=\infty\}} f(t,|p|)\omega_{|p|}|p|^2 d\mu(|p|) \geq \mathfrak{C}_1 - \frac{\mathfrak{C}_2}{\sqrt{t}}$, where $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are explicit constants. Item

(iii) of the main theorem states that the total energy is accumulated at the point $\{\infty\}$ with the rate $\mathcal{O}(\frac{1}{\sqrt{t}})$.

In the limit that $t \to \infty$, all of the energy will be accumulated to the singular part $\tilde{g}(t)\delta_{\omega=\infty}$, while the regular part will vanish, $\bar{g}(t,\omega) \to 0$. This means if we look for a strong (non-measured) solution, whose energy is conserved, it can only exist up to a very short time $t = t_1^*$, at which point it exhibits singular behavior. As a result, we refer to time t_1^* as the first blow-up time. Let $\chi_{[0,R]}(\omega)$ be a cut-off function of ω on the finite domain $[0, R]$, the *multiple blow-up time* phenomenon (1.12)-(1.13)-(1.14)-(1.15), with the decay rate $\mathcal{O}(\frac{1}{\sqrt{t}})$, *can be observed equivalently as the decay of the total energy on any finite interval* $[0, R]$

$$\int_0^R g(t,\omega)d\omega \;=\; \int_{\mathbb{R}_+} \chi_{[0,R]}(\omega)g(t,\omega)d\omega \leq \; \mathcal{O}\Big(\frac{1}{\sqrt{t}}\Big), \tag{1.16}$$

as $t \to \infty$, for all truncation parameters $R > 0$. Inequality (1.16) simply means that the energy of the solution will move away from any truncated finite interval $[0, R]$ as $t \to \infty$ with the rate $\mathcal{O}\Big(\frac{1}{\sqrt{t}}\Big)$. We validate this in chapter 2.

It is interesting to combine the above results of [43] with more physically intuitive arguments. For systems without forcing and dissipation but where an ultra-violet cut-off, $\chi_{[0,R]}(\omega)$, precludes leakage to infinite wavenumbers, the Rayleigh-Jeans spectra is the steady state attractor (see [23], Chapter 1). The wave entropy [44] can be defined

$$S[f](t,p) = \int \log f(t,p)\chi_{[0,R]}(\omega)\mathrm{d}p, \tag{1.17}$$

which if we maximize subject to the constraint that $\dfrac{d\mathcal{E}}{dt} = 0$, with λ a multiplier, we get

$$\max_f S[f](t,p) = \max_f \Big\{ \int \log f(t,p)\chi_{[0,R]}(\omega)\mathrm{d}p - \lambda \int \omega(p) f(t,p)\chi_{[0,R]}(\omega)\mathrm{d}p \Big\}. \tag{1.18}$$

Then taking the variation of (1.17) with the constraint gives

$$\delta S[f](t,p) = \int \left[\frac{\delta f(t,p)}{f(t,p)} - \lambda\omega(p)\delta f(t,p)\right]\chi_{[0,R]}(\omega)\mathrm{d}p \tag{1.19}$$

which, setting the above to zero, gives $f_{\infty}^{RJ}(p) = \frac{1}{\lambda\omega(p)}$ as the maximum entropy solution. The Rayleigh-Jeans spectrum corresponds to an equipartition of the energy, so that as $R \to \infty$, $\mathcal{E}_{\chi_{[0,R]}} \to 0$ [45]. However, unlike the rigorous analysis presented in [43], this formal argument, while adding some physical intuition, we can say nothing about the rate at which the energy goes to zero, only that in a microcananical wave system we expect to reach thermal equilibrium. On the other hand, to be fully understood, the arguments put forth in [43] require a modification of the wave kinetic theory, since it is expected that the wave kinetic equation is only relevant for some finite range of the wavenumber domain (see remark 1.2 page 8 of [43]).

With the above discussion in mind, we are now in a position to state the main contributions of this thesis. First, we wish to consider time-dependent solutions of (4.1) with the only assumption on the solution being that the energy is constant over the extended real line. To this end, in chapter 2 we derive a new identity for the energy density, a conserved quantity for 3-WKEs. This new identity will allow us to use well-known numerical methods for conservation laws to obtain solutions to (4.2). Therefore, we derive a finite volume scheme based on the new identity. The solutions thus obtained are used to validate the theory discussed above, originally put forth in [43]. We provide a CFL type condition for the finite volume scheme in addition to experimental order of convergence tests. The results obtained provide many new lines of inquiry for the analyst of 3-WKEs.

In chapter 3, we provide an alternative numerical method for the Smoluchowski coagulation equation and the 3-WKE. Specifically, we derive a deep learning framework to approximate the solutions to these equations. For the 3-WKE, the obtained solutions also

appear to conform to the theory provided in [43] and are consistent with solutions obtained via the finite volume method of the previous chapter.

Some preliminary results of a forthcoming work studying an inverse problem are given in chapter 4. Here, given an initial condition for the wave-action, data for a single point in time later and some definition of the collision operator, we can first solve an inverse problem to approximate the interaction kernel and from there reconstruct the solution for the wave-action in some time interval $t \in [0, T]$.

In chapter 5, we provide more results of a forthcoming work in which a Residual Distribution method is developed for the 3-wave kinetic equation. It is shown that this method is a generalization of the finite volume method of chapter 2 and indeed zeroth order spatial discretizations reproduce the finite volume solutions.

We conclude in chapter 6 and give some comments on future directions of research and provide some details on ongoing work. A short discussion on transient spectra is provided. The results of the finite volume method in chapter 2 provide a few avenues for the analyst to travel, which we highlight. Finally, we give some comments on open directions provided by the deep learning scheme in chapter 3.

Lastly, we provide an appendix 7 which contains enough background on the deep learning methods employed in chapter 3 to make the thesis self-contained.

Chapter 2

Finite Volume Methods for Isotropic 3-Wave Kinetic Equations

Scientific Product: Walton, Tran. *A numerical scheme for wave turbulence: 3-wave kinetic equations*, SIAM Journal on Scientific Computing (to appear).

2.1. Introduction

The goal of the present chapter is to derive a finite volume scheme that allows us to observe the time evolution of the solutions of (4.2) and to verify the theoretical results of [43] for various values of $\gamma > 1$. In other words, we aim to observe the transferring of energy from the regular to the singular part in (1.12) and to measure precisely the rate of this energy transfer process via the inequality (1.16), that we call the energy cascade rate.

The finite volume scheme derived below relies on the combination of a new energy identity represented in Lemma 2.1 and an adaptation of Filbet and Laurençot's scheme [46] for the Smoluchowski coagulation equation to the 3-wave kinetic equation. Most numerical schemes that approximate integrals on an unbounded domain require the truncation of the unbounded domain to a finite domain. Thanks to the new identity, the number of terms in the collision operator is reduced, reducing the number of truncations needed in the approximation, making the scheme more accurate and reliable. Indeed, we only need to truncate one term in our numerical scheme. Let us comment that the CFL condition is restrictive in certain cases. For other types of equations, the issue of positivity and accurate long-time behavior could be resolved with using the implicit in time discretizations. For instance, a structure preserving scheme has been designed for the Kolmogorov Fokker Planck equation, which is an equation of degenerate parabolic type in the work [47].

The advantage of degenerate parabolic type equations is that those equations are normally local while the 3-wave kinetic equation considered here is highly non-local. Therefore, the previous strategies for parabolic equations, such as the one used for the above Kolmogorov-Fokker-Planck equation, do not immediately carry over to the current 3-wave kinetic equation. We do not attempt to treat an implicit or IMEX time integration method in this thesis.

2.2. A New Identity for the Energy Density

In [46], Filbet and Laurençot derive a finite volume scheme (FVS) for the Smoluchowski coagulation equation (SCE)

$$\begin{aligned} \partial_t f(t,k) &= \mathbb{Q}_{Smo}[f](t,k), \\ f(0,k) &= f_0(k), \end{aligned} \tag{2.1}$$

$$\mathbb{Q}_{Smo}[f](t,k) \;=\; \int_0^k a(k_1,k-k_1)f(k_1)f(k-k_1)\mathrm{d}k_1 \;-2\int_0^\infty a(k,k_1)f(k)f(k_1)\mathrm{d}k_1, \tag{2.2}$$

where $a(\cdot,\cdot)$ is the collision kernel for the 3-wave collision operator (4.3). Let us give a short derivation of the non-conservative form of the SCE (see [48], [49], [46] and the references therein).

Take a test function $\phi(k) = k\chi_{[0,c]}(k)$ and apply it to the SCE

$$\begin{aligned} \int_0^c \partial_t f(t,k)k\mathrm{d}k = \int_0^c\int_0^k a(k,k-k_1)f(k_1)f(k-k_1)k\mathrm{d}k_1\mathrm{d}k \\ -2\int_0^c\int_0^\infty a(k,k_1)f(k)f(k_1)k\mathrm{d}k_1\mathrm{d}k \\ = 2\int_0^c\int_0^{c-k} a(k,k_1)f(k)f(k_1)k\mathrm{d}k_1\mathrm{d}k \\ -2\int_0^c\int_0^\infty a(k,k_1)f(k)f(k_1)k\mathrm{d}k_1\mathrm{d}k. \end{aligned} \tag{2.3}$$

Rearranging the right-hand side, we find

$$\int_0^c \partial_t f(t,k)k\mathrm{d}k \;=\; -2\int_0^c\int_{c-k}^\infty a(k,k_1)f(k_1)f(k)k\mathrm{d}k_1\mathrm{d}k,$$

and upon taking the derivative with respect to c,

$$\partial_t f(t,c)c \;=\; -2\partial_c \int_0^c \int_{c-k}^{\infty} a(k,k_1) f(k_1) f(k) k \mathrm{d}k_1 \mathrm{d}k,$$

and, after truncating the inner integral, we arrive at

$$\partial_t f(t,c)c \;=\; -2\partial_c \int_0^c \int_{c-k}^{R} a(k,k_1) f(k_1) f(k) k \mathrm{d}k_1 \mathrm{d}k, \tag{2.4}$$

where R is a suitable truncation of the volume domain. Then, one can apply any finite volume scheme to solve the truncated problem. We note briefly that the choice of R can effect the accuracy and efficiency of the scheme for the SCE. For example, when considering the case of gelation, R must be quite large in order to avoid a loss of mass before the gelation time. Also, in [50], the authors compare the FVS above with a finite element approximation and find that for smaller truncation values, the finite element scheme is a better choice, while for larger truncation values the FVS should be used.

If we compare equation (4.3) with equation (4.5), we see that the Smoluchowski coagulation equation is a special case of the 3-wave equation. Thus, we would like to adapt the (FVS) of Filbet and Laurençot to derive a numerical scheme for the 3-wave equation. To do this we derive a similar identity to (2.4) for the wave kinetic equation. The role of this identity is to reduce the number of terms in the collision operator. As discussed previously, the truncation of the terms in the collision operator can affect the accuracy of the numerical scheme, thus reducing the number of truncated terms is crucial in the numerical computations of the solutions, as it allows the scheme to be more accurate and reliable.

However, for the 3-wave equation the wave density, $f(t,k)$, is not conserved, but the energy, $g(t,k) = kf(t,k)$, is conserved and so we solve for the energy. What's more, since we do not have an analytic solution to test against, we validate our scheme by verifying

the energy cascade rate for the 3-wave equation found in Soffer and Tran [43] and discussed above.

Lemma 2.1. *The following identity holds true for the energy function* $g(t,k)$

$$\begin{aligned}\partial_t \frac{g(t,c)}{c} &= -2\partial_c \int_0^c \int_0^c a(k,k_1)\frac{g(k)}{k}\frac{g(k_1)}{k_1}\chi\{c<k+k_1\}\mathrm{d}k_1\mathrm{d}k \\ &\quad + \partial_c \int_0^\infty \int_0^\infty a(k,k_1)\frac{g(k)}{k}\frac{g(k_1)}{k_1}\chi\{c<k+k_1\}\mathrm{d}k\mathrm{d}k_1,\end{aligned} \tag{2.5}$$

with $\chi\{\cdot\}$ *the characteristic function and initial condition* $g(0,k)=g_0(k)=kf(0,k)$.

Proof. Let $\phi(k)$ be a test function. From [43], we have the following identity for the 3-wave equation

$$\int_0^\infty \partial_t f(t,k)\phi(k)\mathrm{d}k = \int_0^\infty \int_0^\infty a(k,k_1)f(k)f(k_1)[\phi(k+k_1)+\phi(|k-k_1|)-2\phi(\max\{k,k_1\})]\mathrm{d}k\mathrm{d}k_1. \tag{2.6}$$

If we choose $\phi(k)=\chi_{[0,c]}(k)$ (compare this with $\phi(k)=k\chi_{[0,c]}(k)$ in [46]), we have

$$\begin{aligned}&\int_0^\infty \partial_t f(t,k)\chi_{[0,c]}(k)\mathrm{d}k = \\ &\int_0^\infty \int_0^\infty a(k,k_1)f(k)f(k_1)[\chi_{[0,c]}(k+k_1)+\chi_{[0,c]}(|k-k_1|)-2\chi_{[0,c]}(\max\{k,k_1\})]\mathrm{d}k\mathrm{d}k_1.\end{aligned} \tag{2.7}$$

Set $K(k,k_1)=\chi_{[0,c]}(k+k_1)+\chi_{[0,c]}(|k-k_1|)-2\chi_{[0,c]}(\max\{k,k_1\})$. There are seven cases:

i Assume $k+k_1\le c$ then we have $\chi_{[0,c]}(k+k_1)=1$, $\chi_{[0,c]}(|k-k_1|)=1$ and $\chi_{[0,c]}(\max\{k,k_1\})=1$ which implies that $K(k,k_1)=0$.

ii Assume that $k+k_1>c$ but $k\le c$ and $k_1\le c$ Then we have $\chi_{[0,c]}(k+k_1)=0$, $\chi_{[0,c]}(|k-k_1|)=1$, $\chi_{[0,c]}(\max\{k,k_1\})=1$ and so $K(k,k_1)=-1$.

iii Now let $k+k_1>c$, $k\le c$ but $k_1>c$ then $\chi_{[0,c]}(k+k_1)=0$, $\chi_{[0,c]}(|k-k_1|)=1$, and $\chi_{[0,c]}(\max\{k,k_1\})=0$ giving $K(k,k_1)=1$.

iv Assume $k+k_1 > c$, with $k > c$ and $k_1 \leq c$ and $|k-k_1| > c$, then $\chi_{[0,c]}(k+k_1) = 0$, $\chi_{[0,c]}(|k-k_1|) = 0$, $\chi_{[0,c]}(\max\{k,k_1\}) = 0$ leaving $K(k,k_1) = 0$.

v Assume $k+k_1 > c$, with $k > c$ and $k_1 \leq c$ and $|k-k_1| \leq c$, then $\chi_{[0,c]}(k+k_1) = 0$, $\chi_{[0,c]}(|k-k_1|) = 1$, $\chi_{[0,c]}(\max\{k,k_1\}) = 0$ leaving $K(k,k_1) = 1$.

vi Lastly, set $k+k_1 > c$, $k > c$, $k_1 > c$, and $|k-k_1| > c$, then $\chi_{[0,c]}(k+k_1) = 0$, $\chi_{[0,c]}(|k-k_1|) = 0$ and $\chi_{[0,c]}(\max\{k,k_1\}) = 0$ resulting in $K(k,k_1) = 0$.

vii Lastly, set $k+k_1 > c$, $k > c$, $k_1 > c$, and $|k-k_1| \leq c$, then $\chi_{[0,c]}(k+k_1) = 0$, $\chi_{[0,c]}(|k-k_1|) = 1$ and $\chi_{[0,c]}(\max\{k,k_1\}) = 0$ resulting in $K(k,k_1) = 1$.

Applying these computations to (2.7) and after taking the derivative as done for (2.4), we have

$$\begin{aligned} \partial_t f(t,c) \;=\; & -\partial_c \int_0^\infty \int_0^\infty a(k,k_1) f(k) f(k_1) \chi\{c < k+k_1\} \chi\{k \wedge k_1 \leq c\} \mathrm{d}k_1 \mathrm{d}k \\ & + \partial_c \int_0^\infty \int_0^\infty a(k,k_1) f(k) f(k_1) \chi\{c < k+k_1\} \chi\{k \wedge k_1 > c\} \mathrm{d}k \mathrm{d}k_1, \end{aligned} \tag{2.8}$$

which gives,

$$\begin{aligned} \partial_t f(t,c) \;=\; & -2\partial_c \int_0^c \int_0^c a(k,k_1) f(k) f(k_1) \chi\{c < k+k_1\} \mathrm{d}k_1 \mathrm{d}k \\ & + \partial_c \int_0^\infty \int_0^\infty a(k,k_1) f(k) f(k_1) k_1 \chi\{c < k+k_1\} \mathrm{d}k \mathrm{d}k_1, \end{aligned} \tag{2.9}$$

yielding (2.5). □

With the identity (2.5) in hand, we now derive a finite volume scheme with which to solve it. Note that with (2.5), we only need to truncate a single term, while with the original formula (4.3), we are required to truncate three terms.

2.3. Finite Volume Scheme

We give a discretization for the frequency domain for $k \in [0, R]$. Let $i \in \{1, 2, \dots, M\} = I_h^M$, with the maximum stepsize $h \in (0, 1)$ fixed. We define the set of cells $\mathcal{K} = \bigcup_{i \in I_h^M} K_i = [0, R]$ to be the discretization of the wavenumber domain into M cells. Let

$$K_i = [k_{i-1/2}, k_{i+1/2})_{i \in I_h^M}, \quad \{k_i\}_{i \in I_h^M} = \frac{k_{i+1/2} + k_{i-1/2}}{2}, \quad \{h_i\}_{i \in I_h^M} = k_{i+1/2} - k_{i-1/2},$$

define the cells, pivots and step-size respectively, with the boundary nodes $k_{1/2} = 0$ and $k_{M+1/2} = R$. Note that the discretization does not require a uniform grid, but does restrict the step-size in each cell K_i so that $h_i \leq h$. In practice, it can be useful to drop the restriction that $h \in (0, 1)$, especially when using a non-uniform grid, though we leave it here to simplify the analysis. The set $T_N = \{0, \dots, T\}$ with $N + 1$ nodes, where T is the maximum time, is the discretization of the time domain. We fix the time step to be $\Delta t = \frac{T}{N}$, and denote by $t_n = \Delta t \cdot n$ for $n \in \{0, \dots, N\}$. We approximate equation (2.5) with

$$g^{n+1}(k_i) = g^n(k_i) + \lambda_i \Big(Q^n_{i+1/2}\Big[\frac{g}{k}\Big] - Q^n_{i-1/2}\Big[\frac{g}{k}\Big] \Big), \tag{2.10}$$

where $\lambda_i = \frac{k_i \Delta t}{h_i}$, and

$$Q^n_{i+1/2}\Big[\frac{g}{k}\Big] - Q^n_{i-1/2}\Big[\frac{g}{k}\Big] = -2\Big(Q^n_{1,i+1/2}\Big[\frac{g}{k}\Big] - Q^n_{1,i-1/2}\Big[\frac{g}{k}\Big] \Big) + \Big(Q^n_{2,i+1/2}\Big[\frac{g}{k}\Big] - Q^n_{2,i-1/2}\Big[\frac{g}{k}\Big] \Big),$$

with

$$Q^n_{1,i+1/2}\Big[\frac{g}{k}\Big] = \sum_{m=1}^{i} h_m \frac{g^n(k_m)}{k_m} \left(\sum_{j=1}^{i} h_j \frac{g^n(k_j)}{k_j} a(k_m, k_j) \chi\Big\{ k_{i+1/2} < k_m + k_j \Big\} \right), \tag{2.11}$$

$$Q^n_{2,i+1/2}\left[\frac{g}{k}\right] = \sum_{m=1}^{M} h_m \frac{g^n(k_m)}{k_m}\left(\sum_{j=1}^{M} h_j \frac{g^n(k_j)}{k_j} a(k_m,k_j)\chi\Big\{k_{i+1/2} < k_m + k_j\Big\}\right), \tag{2.12}$$

where we have used the midpoint rule to approximate the integrals in equation (2.5) and we choose an explicit time stepping method.

We will be interested in computing the moments of our solution. We approximate the ℓ-th moment, $\mathcal{M}^\ell(t_n)$, by

$$\mathcal{M}^\ell(t_n) = \sum_{i=1}^{M} h_i g^n(k_i) k_i^\ell, \tag{2.13}$$

with $\ell \in \mathbb{N}$.

The initial condition $g_0(k)$ is approximated by

$$g^0(k_i) = \frac{1}{h_i}\int_{k_{i-1/2}}^{k_{i+1/2}} g_0(k)\mathrm{d}k \approx g_0(k_i),$$

by again employing the midpoint rule. We have that $g^0(k_i) \geq 0$ for all $i \in \{1,\dots,M\}$ since $g_0(k) \geq 0$ for all $k \in [0,R]$ by assumption. To ease notation in the proof of the following proposition, we write g_i^n for $g^n(k_i)$ and the negative flux as

$$\left[Q_i^n\right]^{\mp} = Q^n_{i-1/2}\left[\frac{g}{k}\right] - Q^n_{i+1/2}\left[\frac{g}{k}\right], \tag{2.14}$$

so that the forward Euler scheme becomes $g_i^{n+1} = g_i^n - \lambda_i\left[Q_i^n\right]^{\mp}$. Let us analyze the collision operator (2.14). For the moment, we leave (2.14) in continous form but will apply the same

midpoint approximation when appropriate. Then we have

$$\begin{aligned}\Big[Q_i^n\Big]^{\mp} = &\left(2\int_0^{k_{i+1/2}}\int_0^{k_{i+1/2}} g_1^n g_2^n (k_1k_2)^{\gamma/2-1}\chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1\right.\\ &\qquad\left. - \int_0^R\int_0^R g_1^n g_2^n (k_1k_2)^{\gamma/2-1}\chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1\right)\\ -&\left(2\int_0^{k_{i-1/2}}\int_0^{k_{i-1/2}} g_1^n g_2^n (k_1k_2)^{\gamma/2-1}\chi\Big\{k_{i-1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1\right.\\ &\qquad\left. - \int_0^R\int_0^R g_1^n g_2^n (k_1k_2)^{\gamma/2-1}\chi\Big\{k_{i-1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1\right)\\ &= (\mathrm{I} - \mathrm{II}) - (\mathrm{III} - \mathrm{IV}),\end{aligned} \tag{2.15}$$

with $g_i^n = g^n(k_i)$ for $i = 1, 2$. We may decompose the integrals above so that

$$\begin{aligned}\mathrm{I} = &\,2\int_0^{k_{i-1/2}}\int_0^{k_{i-1/2}} g_1^n g_2^n (k_1k_2)^{\gamma/2-1}\chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1\\ &+2\int_0^{k_{i-1/2}}\int_{k_{i-1/2}}^{k_{i+1/2}} g_1^n g_2^n (k_1k_2)^{\gamma/2-1}\chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1\\ &+2\int_{k_{i-1/2}}^{k_{i+1/2}}\int_0^{k_{i-1/2}} g_1^n g_2^n (k_1k_2)^{\gamma/2-1}\chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1\\ &+2\int_{k_{i-1/2}}^{k_{i+1/2}}\int_{k_{i-1/2}}^{k_{i+1/2}} g_1^n g_2^n (k_1k_2)^{\gamma/2-1}\chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1\\ &= 2\int_0^{k_{i-1/2}}\int_0^{k_{i-1/2}} g_1^n g_2^n (k_1k_2)^{\gamma/2-1}\chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1\\ &+4\int_{k_{i-1/2}}^{k_{i+1/2}}\int_0^{k_{i-1/2}} g_1^n g_2^n (k_1k_2)^{\gamma/2-1}\chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1\\ &+2\int_{k_{i-1/2}}^{k_{i+1/2}}\int_{k_{i-1/2}}^{k_{i+1/2}} g_1^n g_2^n (k_1k_2)^{\gamma/2-1}\chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1,\end{aligned} \tag{2.16}$$

and similar decompositions may be performed on II, III and IV, which, leaving out the $g_1^n g_2^n (k_1 k_2)^{\gamma/2-1}$ integrands to simplify the expressions, gives

$$\begin{aligned}
\Big[Q_i^n\Big]^{\mp} &= 2\int_0^{k_{i-1/2}}\int_0^{k_{i-1/2}} \Big(\chi\Big\{k_{i+1/2} < k_1 + k_1\Big\} - \chi\Big\{k_{i-1/2} < k_1 + k_1\Big\}\Big)\mathrm{d}k_2\mathrm{d}k_1 \\
&\quad + \int_0^{R}\int_0^{R} \Big(\chi\Big\{k_{i-1/2} < k_1 + k_1\Big\} - \chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\Big)\mathrm{d}k_2\mathrm{d}k_1 \\
&\quad + 4\int_{k_{i-1/2}}^{k_{i+1/2}}\int_0^{k_{i-1/2}} \chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1 \\
&\quad + 2\int_{k_{i-1/2}}^{k_{i+1/2}}\int_{k_{i-1/2}}^{k_{i+1/2}} \chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1, \\
&= -2\int_0^{k_{i-1/2}}\int_0^{k_{i-1/2}} \chi\Big\{k_{i-1/2} < k_1 + k_1 \le k_{i+1/2}\Big\}\mathrm{d}k_2\mathrm{d}k_1 \\
&\quad + \int_0^{R}\int_0^{R} \chi\Big\{k_{i-1/2} < k_1 + k_1 \le k_{i+1/2}\Big\}\mathrm{d}k_2\mathrm{d}k_1 \\
&\quad + 4\int_{k_{i-1/2}}^{k_{i+1/2}}\int_0^{k_{i-1/2}} \chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1 \\
&\quad + 2\int_{k_{i-1/2}}^{k_{i+1/2}}\int_{k_{i-1/2}}^{k_{i+1/2}} \chi\Big\{k_{i+1/2} < k_1 + k_1\Big\}\mathrm{d}k_2\mathrm{d}k_1.
\end{aligned} \tag{2.17}$$

At this point, we apply the midpoint rule to approximate the integrals which, after making use the characteristic functions, simplifies the above to

$$\Big[Q_i^n\Big]^{\mp} \approx 2g_i^n k_i^{\gamma/2-1} h_i \Big(\sum_{j=1}^{i} g_j^n k_j^{\gamma/2-1} h_j + \sum_{j=1}^{i-1} g_j^n k_j^{\gamma/2-1} h_j\Big) - \sum_{j=1}^{i-1} g_j^n g_{i-j}^n (k_j k_{i-j})^{\gamma/2-1} h_j h_{i-j}. \tag{2.18}$$

In what follow, we will abuse notation slightly and right $\Big[Q_i^n\Big]^{\mp}$ when we mean its midpoint approximation above.

We are now in a position to state the following sufficient stability condition for the time step.

Theorem 2.2. *If the time step, Δt satisfies*

$$\Delta t R^{\gamma+1} \|g^0\|_{L^\infty(0,R)} \le \frac{\gamma}{16} \min_{i \in I_h^M} h_i, \tag{2.19}$$

then, for all $n \geq 0$ *and* $i \in I_h^M$ *we have* $g_i^n \geq 0$, *and further,*

$$\|g^{n+1}\|_{L^\infty(0,R)} \leq C(\gamma)\|g^n\|_{L^\infty(0,R)},$$

with $C(\gamma) \in [\frac{15}{16}, 1]$ *so that finally,*

$$\|g^n\|_{L^\infty(0,R)} \leq \|g^0\|_{L^\infty(0,R)}, \tag{2.20}$$

for all $n \geq 0$.

Proof. Assume $R >> 1$. We will proceed by induction. Starting from the forward Euler scheme and using $[Q_i^0]^\mp$ as computed above we obtain

$$\begin{aligned} g_i^1 = g_i^0 \Bigg(1/2 - \lambda_i \Big(2k_i^{\gamma/2-1} h_i \Big[\sum_{j=1}^{i} g_j^0 k_j^{\gamma/2-1} h_j + \sum_{j=1}^{i-1} g_j^0 k_j^{\gamma/2-1} h_j \Big] \Big) \Bigg) \\ + \frac{1}{2} g_i^0 + \lambda_i \sum_{j=1}^{i-1} g_j^0 g_{i-j}^0 (k_j k_{i-j})^{\gamma/2-1} h_j h_{i-j} \ = \varphi_i^0(g) + \zeta_i^0(g). \end{aligned} \tag{2.21}$$

Given $g_i^0 \geq 0$, $\zeta_i^0(g)$ is clearly positive. Then, to conclude that $g_i^1 \geq 0$ it is sufficient to have $\varphi_i^0(g) \geq 0$ or equivalently

$$\Delta t \frac{k_i}{h_i} \Big(2k_i^{\gamma/2-1} h_i \Big[\sum_{j=1}^{i} g_j^0 k_j^{\gamma/2-1} h_j + \sum_{j=1}^{i-1} g_j^0 k_j^{\gamma/2-1} h_j \Big] \Big) \leq \frac{1}{2}. \tag{2.22}$$

To this end, note that

$$\begin{aligned} k_i \Big(2k_i^{\gamma/2-1} h_i \Big[\sum_{j=1}^{i} g_j^0 k_j^{\gamma/2-1} h_j + \sum_{j=1}^{i-1} g_j^0 k_j^{\gamma/2-1} h_j \Big] \Big) \ &\leq 4k_i^{\gamma/2} \|g^0\|_{L^\infty(0,R)} \|k^{\gamma/2-1}\|_{L^1(0,R)} \\ &\leq \frac{8}{\gamma} R^{\gamma+1} \|g^0\|_{L^\infty(0,R)}. \end{aligned} \tag{2.23}$$

For now, we provide a provisional stability condition on Δt so that

$$\Delta t R^{\gamma+1} \max_{\ell \le n} \{ \|g^\ell\|_{L^\infty(0,R)} \} \le \frac{\gamma}{16} \min_{i \in I_h^M} h_i, \tag{2.24}$$

which will be seen to be equivalent to the condition (2.19) by the end of the proof. Thus, with (2.24) setting $n = 0$ we have that $\varphi_i^0(g)$ is positive. Now, assuming $g_i^n \ge 0$, we proceed in the same fashion. The steps are identical except that now we have the condition that

$$\Delta t R^{\gamma+1} \|g^n\|_{L^\infty(0,R)} \le \frac{\gamma}{16} \min_{i \in I_h^M} h_i, \tag{2.25}$$

which again is satisfied by the provisional assumption on Δt and we have by induction that $g_i^{n+1} \ge 0$ as desired.

To show stability, let us consider $\zeta_i^n(g)$ and gives estimates for its second term. To this end, note that

$$\begin{aligned} \frac{k_i}{h_i} \sum_{j=1}^{i-1} g_j^n g_{i-j}^n (k_j k_{i-j})^{\gamma/2-1} h_j h_{i-j} \; &\le k_i \|g^n\|_{L^\infty(0,R)}^2 \sum_{j=1}^{M} k_j^{\gamma-1} h_j \\ &\le \frac{k_i}{\gamma} R^\gamma \|g^n\|_{L^\infty(0,R)}^2 \\ &\le \frac{1}{\gamma} R^{\gamma+1} \|g^n\|_{L^\infty(0,R)}^2, \end{aligned} \tag{2.26}$$

where the rearrangement inequality was used in the second line. Then, using (2.24), we obtain

$$\Delta t \frac{k_i}{h_i} \sum_{j=1}^{i-1} g_j^n g_{i-j}^n (k_j k_{i-j})^{\gamma/2-1} h_j h_{i-j} \le \frac{1}{16} \|g^n\|_{L^\infty(0,R)} \tag{2.27}$$

Once again using the assumption (2.24), we see that

$$g_i^{n+1} = \varphi_i^n(g) + \zeta_i^n(g) \le g_i^n \Big(\frac{8-\gamma}{16} \Big) + \frac{1}{2} g_i^n + \frac{1}{16} \|g^n\|_{L^\infty(0,R)}, \tag{2.28}$$

which implies that

$$\|g^{n+1}\|_{L^\infty(0,R)} \leq \left(\frac{17-\gamma}{16}\right)\|g^n\|_{L^\infty(0,R)} = C(\gamma)\|g^n\|_{L^\infty(0,R)}. \tag{2.29}$$

We see that for $\gamma \in [1,2]$ we have $C(\gamma) \in [\frac{15}{16}, 1]$ and thus

$$\|g^{n+1}\|_{L^\infty(0,R)} \leq \|g^0\|_{L^\infty(0,R)}, \tag{2.30}$$

which gives the equivalence of condition (2.24) with (2.19). □

In the following section we provide numerical results for several initial conditions, which give evidence for the dependence of the time steps on the initial condition and truncation parameter as demonstrated above.

2.4. Numerical Tests

As mentioned above, the KZ spectrum is not expected in our equation, in absence of forcing and dissipation. Our numerical tests are therefore only designed to verify the blow up phenomenon first proved in [43] as well as to confirm the energy cascade growth rate bound $\mathcal{O}(1/\sqrt{t})$ obtained in the same paper. To verify this, in any finite interval of the wavenumber domain, we should see the total energy $\mathcal{E}$, or equivalently the zeroth moment of $g(t,k)$, decaying like $\mathcal{O}(\frac{1}{\sqrt{t}})$ in the interval. As (1.12)-(1.13)-(1.14)- (1.15) is equivalent with (1.16), our numerical tests will focus on verifying (1.16).

All tests were performed on a uniform grid, with $h = 0.5$ in the first two test and $h = 0.1$ in the last two. We solve the 3-wave equation via (2.5) for four different initial conditions. For each initial condition, we either vary the truncation parameter, R, in (1.16), and hold the degree of the collision kernel, γ, fixed or we vary γ with R kept at a fixed value. Specifically, for the first two initial conditions, we run tests for $R = 50, 100, 200$ with $\gamma = 2$ fixed or $\gamma = \frac{3}{2}, \frac{9}{5}, 2$ with $R = 100$ fixed. In the last two cases, we do something similar and again run

tests with $\gamma = 2$ fixed but set $R = 25, 50, 80$ and then fix $R = 50$ and let $\gamma = \frac{3}{2}, \frac{9}{5}, 2$. We provide an approximation of the convergence rate for the first initial condition by comparison of approximated solutions over succesively refined grids. In all solutions presented, we use a second-order Runge-Kutta scheme to integrate in time.

The first few moments of the energy, $g(t, k)$, are then computed for each set of parameters. We find our numerical experiments to be in good agreement with [43]. We note that, as with explicit methods for the Smoluchowski equation, the CFL condition can be very restrictive if one wants to maintain positivity (see [51]). Similarly to schemes for other types of kinetic equations, maintaining positivity of the numerical solutions is also an important issue in our scheme.

We provide a condition in Proposition 2.2 that guarantees the positivity of the solutions. Under this condition, the positivity can be preserved when we choose Δt sufficiently small.

We observe the choice of Δt depends on the initial data and the size of the truncated interval. For example, in Test 1 we set $\Delta t = 0.05$ and in Test 2 we set $\Delta t = 0.005$, though we perform computations with the same truncation parameter. Also, in Tests 3 and 4, as we increase the truncation parameter from $R = 50$ to $R = 80$, we must drop the time step from $\Delta t = 0.0004$ to $\Delta t = 0.00025$, respectively.

We believe that common positive preserving techniques like those developed in [52–54] could potentially be a solution to handle this instability issue.

2.4.1. Test 1

Here we choose our initial condition to be

$$g_0(k) = 1.26157 e^{-50(k-1.5)^2} \tag{2.31}$$

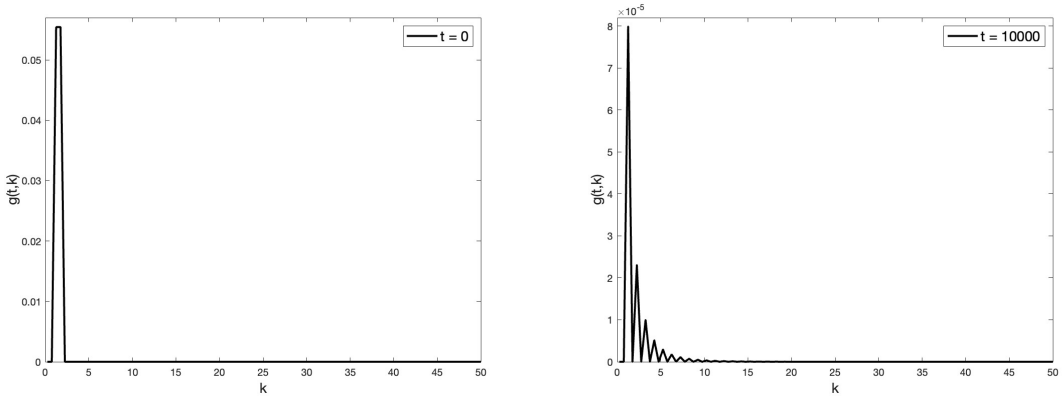
with $\Delta t = 0.05$ for $t \in [0, T]$, $T = 10000$, over a uniform grid, as mentioned previously, with $h = 0.5$. The initial condition and final state, $g(T, k)$, are plotted in Figure 2.1a and 2.1b, respectively, for $\gamma = 2$. The theory developed in [43], in which it has been proven that as t tends to ∞, the energy $g(t, k)$ converges to a delta function $\mathcal{E}\delta_{\{k=\infty\}}$, applies for $\gamma > 1$. Moreover, due to (1.16), the energy on any finite interval also goes to 0

$$\lim_{t\to\infty} g(t, k)\chi_{[0,R]}(k) = 0. \tag{2.32}$$

The two Figures 2.1a and 2.1b indeed give strong evidence for the theoretical result (2.32), proved in [43].

The first four moments of the energy are shown in Figure 2.2 for $\gamma = \frac{3}{2}, \frac{9}{5}, 2$ and $k \in [0, 100]$. The numerical results in Figure 2.2 show that higher moments of g also decay to 0, due to (2.32). We also see that the onset of decay happens later for smaller degree, γ. For the zeroth moment, the curve corresponding to $\gamma = 2$ is below the curve for $\gamma = 3/2$ and the curve for $\gamma = 9/5$. However, for the higher order moments, an interesting phenomenon happens. The $\gamma = 3/2$ curve has a big jump when t is around 10 to 40, and it lies above the other moments from that time on, while the $\gamma = 2$ curve is still below the $\gamma = 9/5$ curve at longer times. While (2.32) can be used to predict what happens in the Figures 2.1, Figure 2.2 indeed needs a different theoretical explanation, which should be an interesting subject of analysis in a follow-up paper.

We then investigate the moments for increasing values of the truncation parameter. The first four moments of $g(t, k)$ are plotted in Figure 2.3 with $\gamma = 2$ fixed. For this initial condition, as might be expected, increasing the truncation parameter seems to have a negligible affect on the higher moments. For the other initial conditions in the test cases that follow, the difference is more distinguishable. When we are able to make a distinction, the decay (2.32) seems to be slower for larger values of the truncation parameter R, though the rate of



(a) Initial Condition (2.31). (b) Final Condition.

Figure 2.1: Test Case 1

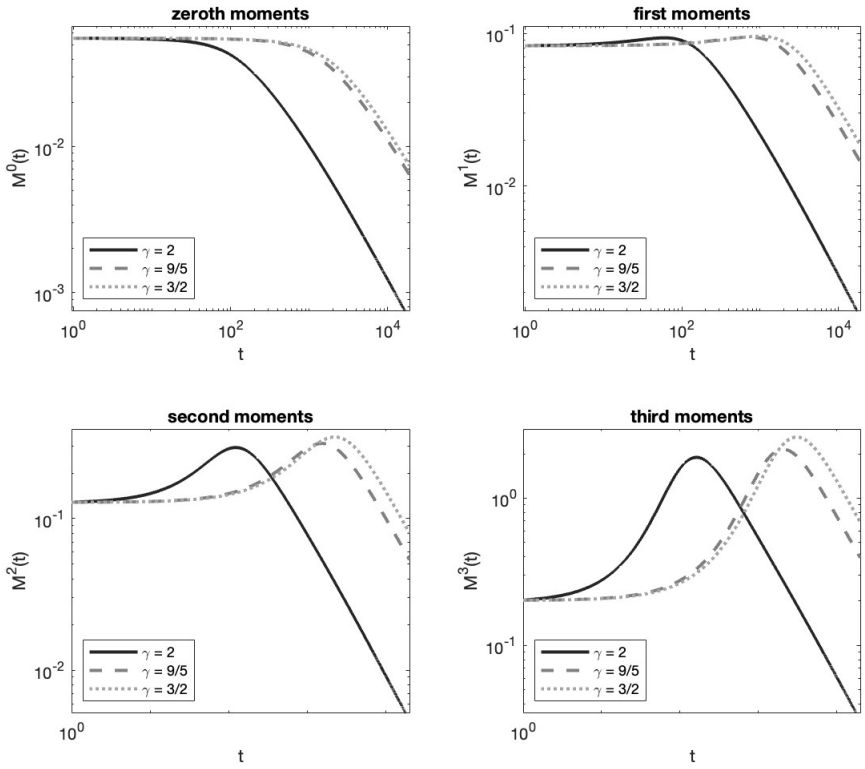


Figure 2.2: Here, we fix $R = 100$ and compute the moments of the energy as a function of time for $\gamma = 3/2, 9/5$ and 2 for initial condition (2.31).

cascade is unaffected. Moreover, when the initial condition spreads the energy across the interval, larger truncation parameters result in larger amounts of energy initially, as one might expect, but again the cascade rate is independent of the truncation value in these cases. This observation is consistent with the findings of [43] and can be understood as follows. As the energy moves away from the zero frequency $k = 0$ and goes to the frequency $k = \infty$ as time evolves, the chance of having some energy in a finite interval $k \in [0, R]$ increases when R increases.

To test convergence without an exact solution, we provide two experimental order of convergence (EOC) tests. First, we compare step sizes h, $h/2$ and $h/4$ for $R = 50$ and $\gamma = 2$. Here we run the simulation for $t \in [0, T_{EOC}]$ with $T_{EOC} = 25$ and $\Delta t = 0.0125$ for every choice of h. The approximate order of convergence, p, is given by [55]

$$p \approx \log_2 \Big(\frac{\|g_h - g_{h/4}\|_{L^1}}{\|g_{h/2} - g_{h/4}\|_{L^1}} - 1 \Big). \tag{2.33}$$

The coarser solutions are interpolated using matlab's built-in piecewise cubic hermite interpolating polynomial (pchip) interpolator. Table 2.1 summarizes the approximation of p for each h at $t_{max} = \arg\max \|g_h - g_{h/4}\|_{L^1}$.

h	0.4	0.3	0.2
p	2.5777	2.6071	2.6392

Table 2.1: Approximation of p with formula 2.33 for various h values.

Next, we approximate p by comparison with a fine grid solution, g_{h^*}, with $h^* = \frac{1}{80}$ and Δt as before. Then we approximate p with the ratio [55]

$$R_h(t) = \Big(\frac{\|g_h - g_{h^*}\|_{L^1(0,R)}}{\|g_{h/2} - g_{h^*}\|_{L^1(0,R)}} \Big). \tag{2.34}$$

Then, the EOC is given by $p \approx \log_2\{R_h(T_{EOC})\}$. The results for $h = 0.2$, $h = 0.1$ are summarized in Table 2.2.

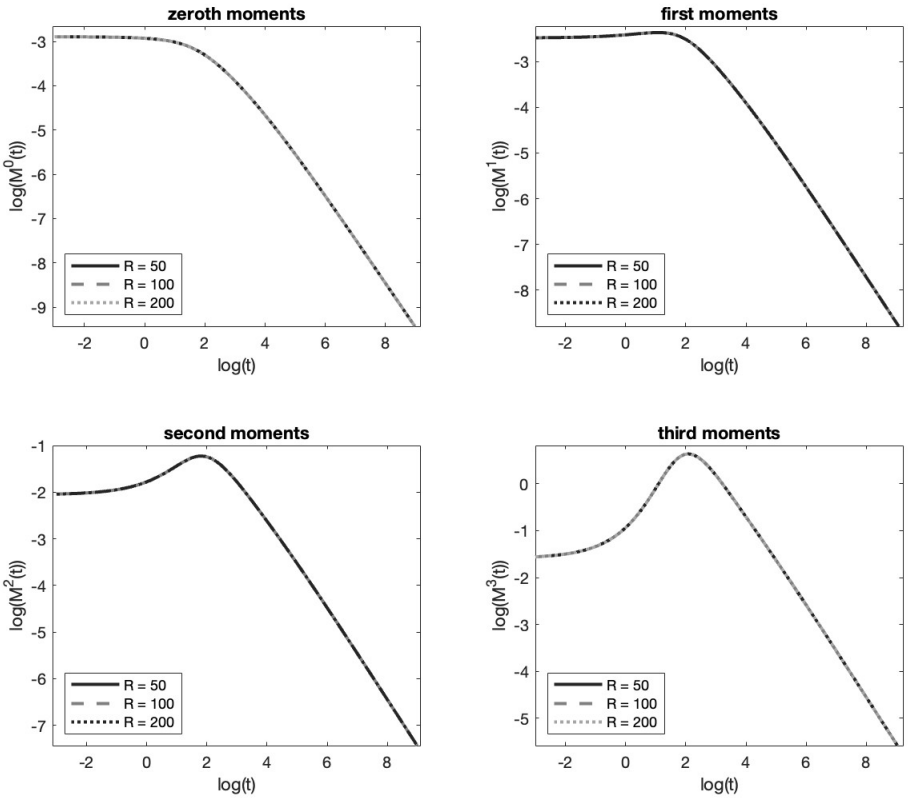


Figure 2.3: For $\gamma = 2$ fixed, we compute the moments of the energy with initial condition (2.31) for truncation parameters $R = 50, 100, 200$.

h	0.2	0.1
p	2.0118	2.5288

Table 2.2: Approximation of p with formula 2.33 for various h values.

A log-log plot of the errors is provided in Figure 2.4. The verification of these experiments requires further analysis, which is the subject of a forthcoming work.

We test the decay rate of the total energy in $[0, \infty)$ obtained in [43] in Figures 2.5a and 2.5b. We give a log-log plot of the zeroth moment of $g(t, k)$ for varying truncation parameter with $\gamma = 2$ and with fixed truncation parameter $R = 200$ with varying degree against the theoretical decay rate of the total energy in $[0, \infty)$. In [43], it has been shown that the decay rate can be bounded by $\mathcal{O}(\frac{1}{\sqrt{t}})$ (see (1.16)), which has very good agreement with the numerical results, where the slope of the decay rate curve in the long time limit are quite below the slope of the line corresponding to $\frac{1}{\sqrt{t}}$. We also compare with a rate like $\mathcal{O}(\frac{1}{t})$. Here, and for the other initial conditions we consider, it would appear that the rate is like $\mathcal{O}(\frac{1}{t^s})$, with $s \in [1/2, 1]$, depending on the degree and the initial condition. Confirmation of this observation requires further analysis.

2.4.2. Test 2 We next choose a Gaussian further away from the origin

$$g_0(k) = (5\pi)^{-1/2} e^{-(k-50/3)^2/2.5} \tag{2.35}$$

as our initial condition. Solutions are computed up to $T = 10000$ with $\Delta t = 0.005$ and $h = 0.5$. The initial condition and final state can be seen in Figure 2.6a and 2.6b, respectively. This test is designed based on Test 1, as we are curious to see what happens if we move the Gaussian away from zero.

In figures 2.6a and 2.6b we see that the energy is pushed slightly toward the origin at some $T_s \in [0, T)$ to $k = 11.75$, away from its initial concentration at $k = 16.667$ at $t = 0$.

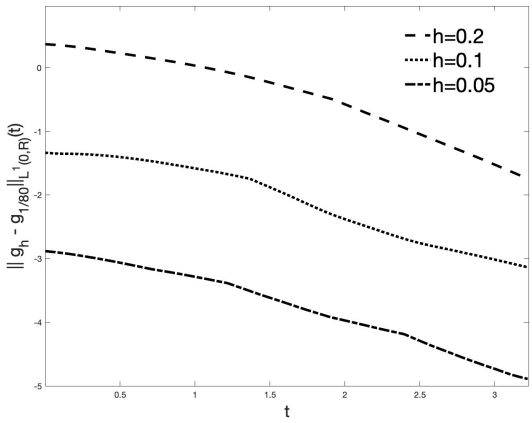
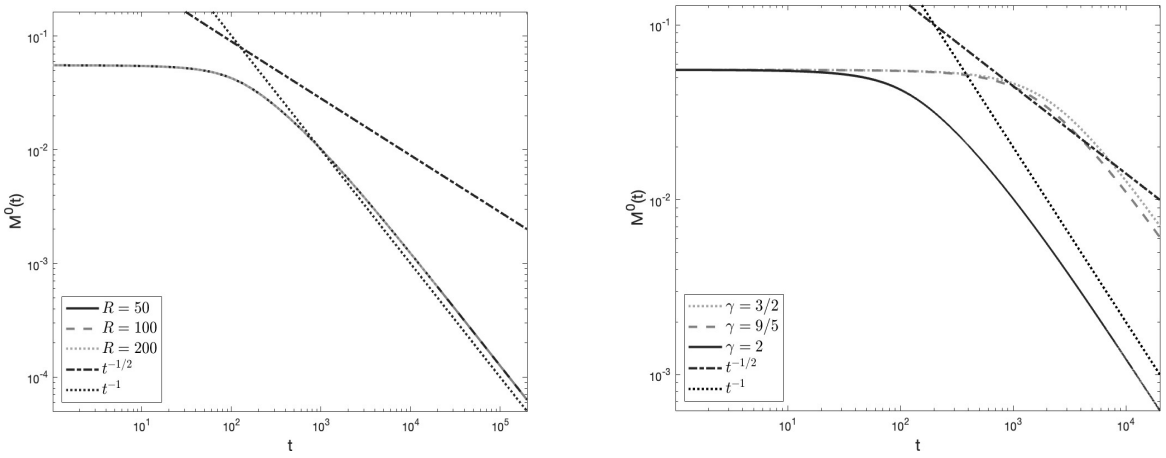
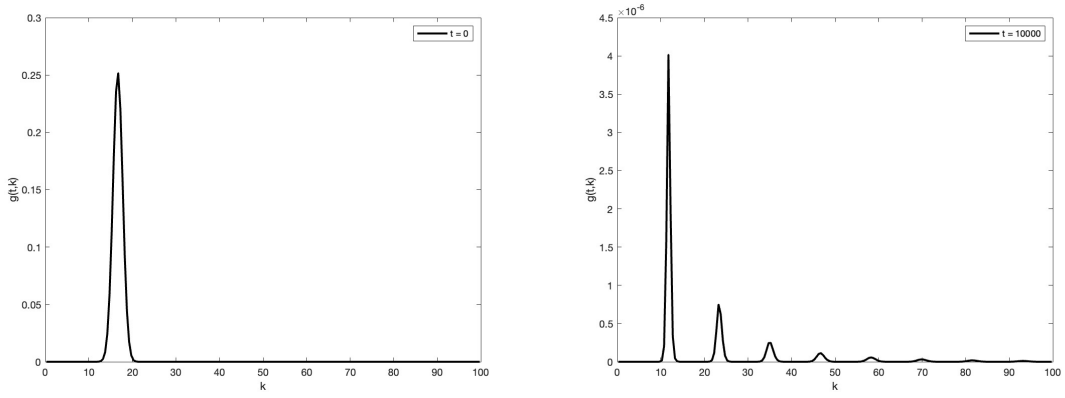


Figure 2.4: Log-log plot of the error between coarse and fine grid solutions.



(a) (b)

Figure 2.5: (a) Rate of decay of the total energy, corresponding to initial condition (2.31) with degree $\gamma = 2$, allowing R to vary. The theoretical cascade rate is shown for comparison. (b) Rate of decay of the total energy corresponding to the initial condition (2.31) with theoretical cascade rate plotted for comparison.



(a) Initial Condition (b) Final Condition

Figure 2.6: Test Case 2

From this time T_s onward, the L^∞ norm of $\|g(t,k)\chi_{[0,R]}(k)\|_{L^\infty}$ decreases to 4.01×10^{-5} for $t = 10000$, maintaining its concentration at $k = 11.75$. This indicates the energy cascade phenomenon does happen and happens in a very special way. An analysis then needs to be done to explain this long time behavior of the solution.

In Figure 2.7, we give results for the computation of the first few moments of the energy when allowing the degree to vary while holding the truncation value fixed. We notice a similar behavior as in the previous test case.

The moment calculations are performed with $\gamma = 2$ fixed and varying truncation parameter. The results are plotted in Figure 2.8. In contrast to Figure 2.3, the difference in moments is more distinguishable. However, as already mentioned, this is consistent with the previous analysis in [43], being that the chance of finding energy in a larger frequency interval is higher.

The theoretical decay rate is compared with the decay of total energy for all considered values of γ and $R = 100$ in Figure 2.9b. As in Test 1, the numerical results have a good agreement with the theoretical findings of [43]. It appears that the decay is more like $\mathcal{O}(\frac{1}{t})$, which is bounded by $\mathcal{O}(\frac{1}{\sqrt{t}})$ as shown in (1.16). We also compare the theoretical cascade

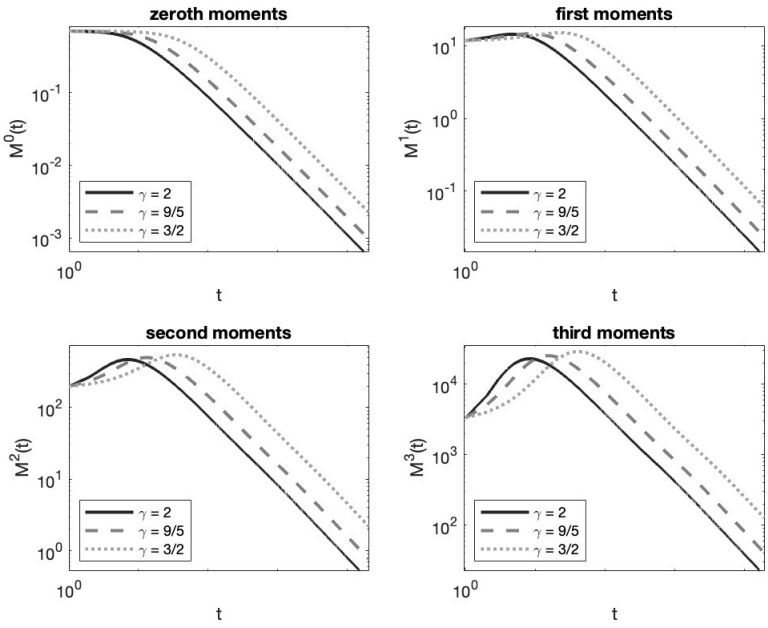


Figure 2.7: Here, we fix $R = 100$ and compute the moments of the energy as a function of time for $\gamma = 3/2, 9/5, 2$ with initial condition (2.31).

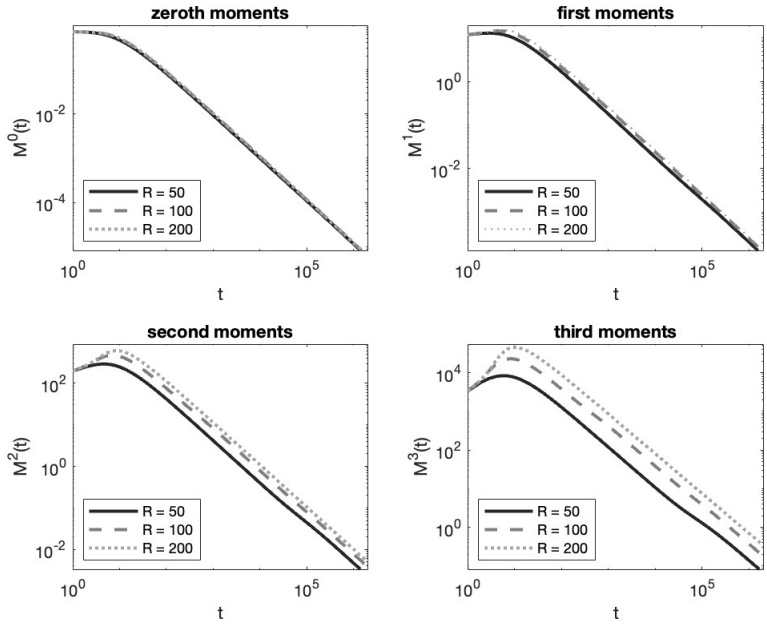
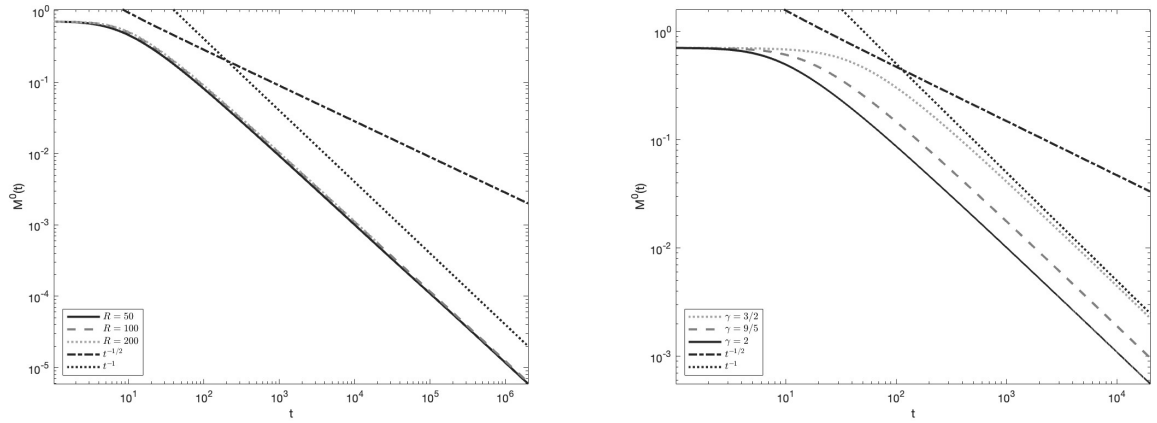


Figure 2.8: Moments of $g(t, k)$ with initial condition (2.35) and fixed degree, $\gamma = 2$, and varying truncation parameter.



(a) (b)

Figure 2.9: (a) Rate of decay of the total energy with varying truncation parameter corresponding to the initial condition shown in figure 2.6a. (b) Rate of decay of the total energy corresponding to the initial condition shown in figure 2.6a with varying degree and theoretical rate plotted for comparison.

rate to the decay of total energy for each interval considered in Figure 2.9a.

We see that increasing the truncation parameter has no influence on the cascade rate and blow-up times, consistent with the theory found in [43]. We notice that here and in the previous test case, that there is also not a distinguishable difference (if any) in the amount of energy contained in the intervals after varying the truncation parameter, which will contrast with the next two test cases, where much more energy is available initially.

2.4.3. Test 3 Here, we consider initial data given by

$$g_0(k) = \begin{cases} 1 & k \in [2n\pi, (2n+1)\pi] \\ 0 & k \in ((2n+1)\pi, 2(n+1)\pi) \end{cases} \quad \text{for } n = 0, 1, 3, 5, \ldots \tag{2.36}$$

and perform test for $t \in [0, T]$ for $T = 100$ and $\Delta t = 0.0004$ when $R = 25$ and $R = 50$ and $\Delta t = 0.00025$ for $R = 80$. The frequency step is $h = 0.1$ for each interval $[0, R]$ considered.

In Figure 2.10a we show the initial condition and in Figure 2.10b the final state at $T = 100$. In the final state, it would appear that the remaining energy in the interval is collected near $k = 0$ with decreasing maximum amplitude at $k = 0.05$. It seems that this profile is maintained as $T \to \infty$, in a similar fashion to Tests 1 and 2.

We then vary the truncation parameter and show the decay of the energy for $R = 25, 50, 80$ in Figure 2.11a. As mentioned previously, we can now see that the amount of energy increases with the interval size, but, importantly, the rate of decay is the same for each truncation value.

As in the previous tests, we explore varying the degree as seen in Figure 2.11b. Here, as in the next case, we begin to see further contrasting behaviour as compared with the previous two test cases. We see that the smaller the value of the degree, the longer the energy is conserved, but now, once the first blow-up time t_1^* is reached, the rate of decay is larger for $\gamma = 3/2, 9/5$ as compared to $\gamma = 2$. To see this, we add a reference line corresponding to a rate of decay like $\mathcal{O}(\frac{1}{t})$ in Figures 2.11a and 2.11b. It would appear that for smaller values of γ, the decay rate is better described by $\mathcal{O}(\frac{1}{t})$, but for $\gamma = 2$, the theoretical bound $\mathcal{O}(\frac{1}{\sqrt{t}})$ (see (1.16)), provides an excellent description of the cascade rate.

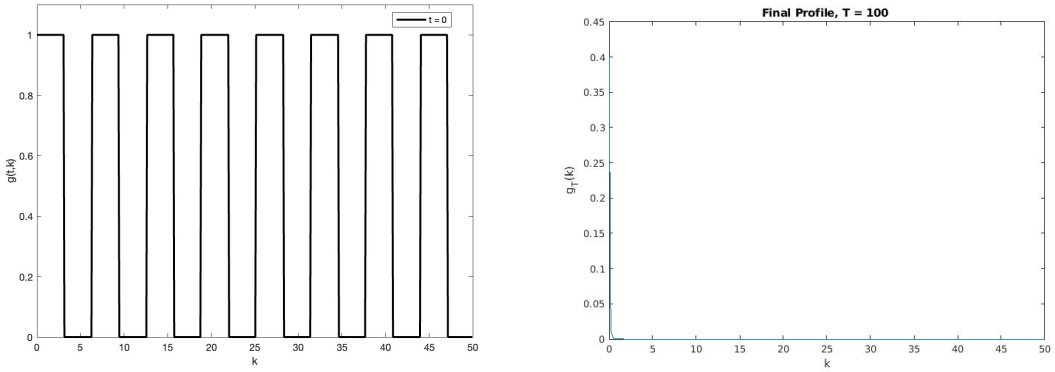
2.4.4. Test 4

Our last test has initial data given by

$$g_0(k) = \frac{k - 2n\pi}{2\pi} \quad k \in [2n\pi, 2(n+1)\pi), \tag{2.37}$$

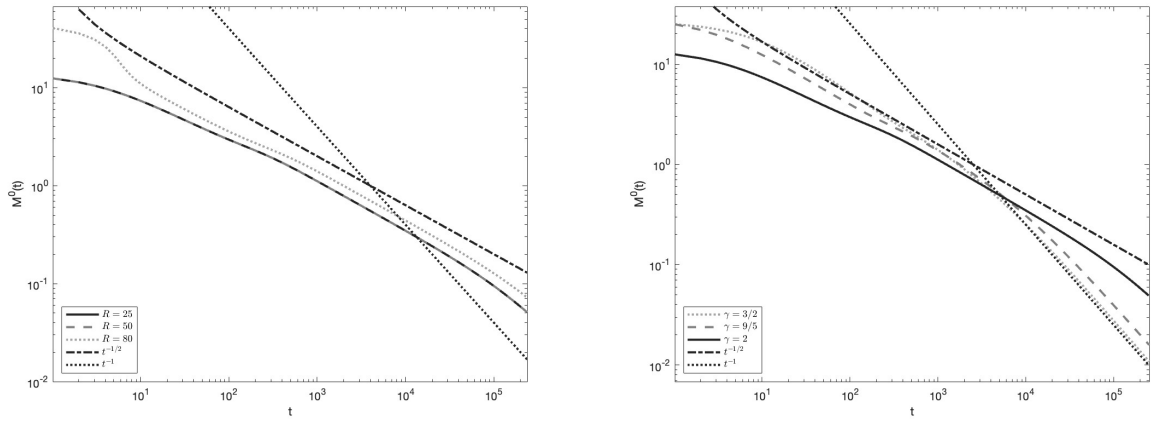
for $n \in \mathbb{N}_0$. As in Test 3, we set $h = 0.1$, $T = 100$ and $\Delta t = 0.0004$ when $R = 25$ and $R = 50$ but $\Delta t = 0.00025$ for $R = 80$.

We again give initial and final conditions in Figures 2.12a and 2.12b, respectively. As in Test 3, we observe that the energy is accumulated to a frequency nearer to $k = 0$ in some finite time T_s, where it remains fixed with decaying L^∞ norm.



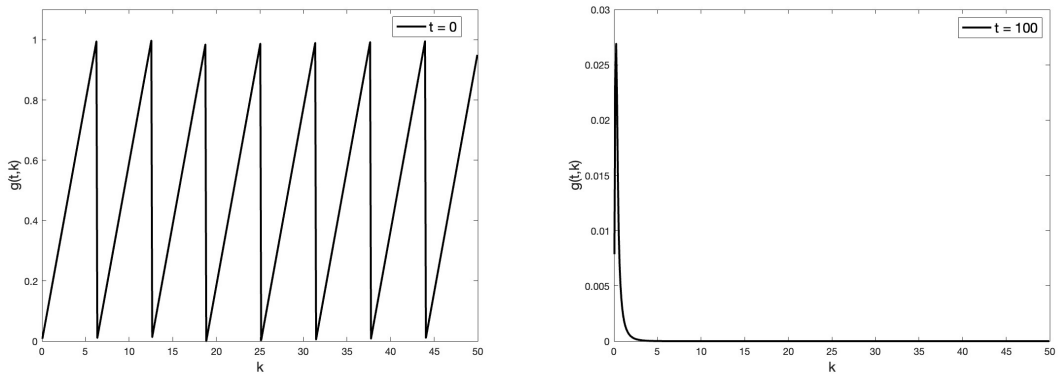
(a) Initial Condition (b) Final Condition

Figure 2.10: Test Case 3



(a) (b)

Figure 2.11: (a) Zeroth moments of solution corresponding to initial condition (2.36), with $\gamma = 2$ and allowing R to vary. (b) Decay rate for initial condition (2.36) plotted against the theoretical rate. The red and black lines are translated to match the intersection of the moments.



(a) Initial Condition (b) Final Condition

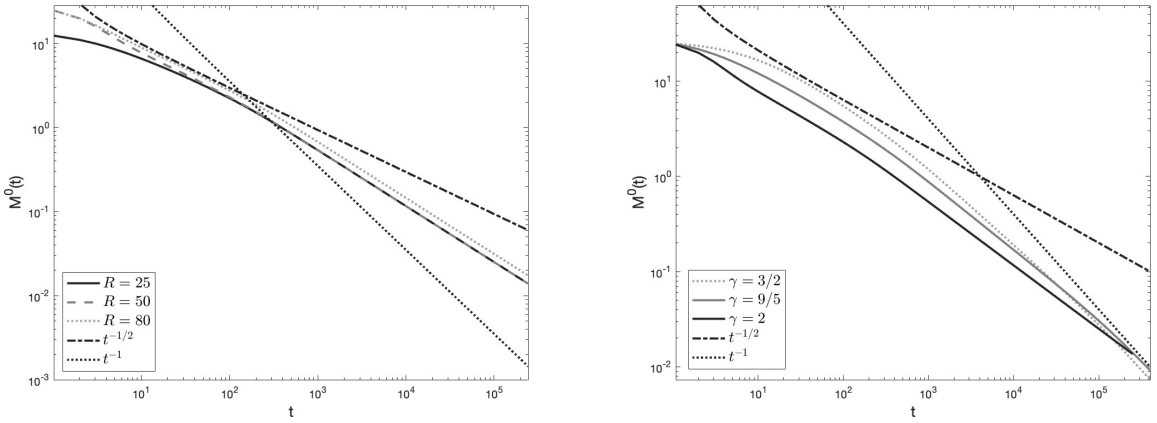
Figure 2.12: Test Case 4

Now fixing $\gamma = 2$ once again, we see in Figure 2.13a that varying the truncation parameter has a similar result on the total energy as in the previous test case. By extending the interval, we see that a larger amount of energy is retained, but that the rate of decay is equal for all truncation values. Further, we see a good fit with the theoretical decay rate.

Next, we let $\gamma = \frac{3}{2}, \frac{9}{5}, 2$ and keep $R = 50$ once again, and show the zeroth moments of these solutions in Figure 2.13b. Here, as for the previous test cases, for $\gamma = 3/2, 9/5$, we see that the energy is conserved for a longer amount of time when compared to $\gamma = 2$. As for the previous initial condition, it would appear that once the decay begins, the rates of decay are faster for $\gamma = 3/2, 9/5$ than for $\gamma = 2$. We again observe that for $\gamma = 2$, the decay rate is more like $\mathcal{O}(\frac{1}{\sqrt{t}})$ and for the smaller values of γ, the decay rate is more like $\mathcal{O}(\frac{1}{t})$.

2.5. Conclusions and Further Discussion

We introduce a finite volume scheme which allows us to observe the long time asymptotics of the solutions of isotropic 3-wave kinetic equations, including the energy cascade behavior proved in [43]. Our numerical algorithm is based on the combination of the identity represented in Lemma 2.1 and Filbet and Laurençot's scheme [46] for the Smoluchowski coagulation equation.



(a) (b)

Figure 2.13: (a) Zeroth moments of solution corresponding to initial condition (2.37), with $\gamma = 2$ and allowing R to vary. The theoretical decay rate is shown for comparison. (b) Zeroth moments of solution corresponding to initial condition (2.37), with $R = 50$ and allowing γ to vary. The theoretical decay rate is shown for comparison.

From the four numerical tests, we can see that the energy cascade behavior happens for $\gamma = \frac{3}{2}, \frac{9}{5}, 2$, and seem to verify the theory found in [43].

From the solutions computed in sections 2.4.1, 2.4.2, 2.4.3, and 2.4.4, One can see that the smaller γ is, the slower is the onset of decay. The energy cascade behavior also seems to occur independently of the smoothness of the initial data for all four cases with $\gamma = \frac{3}{2}, \frac{9}{5}, 2$.

The results in figure 2.3 and 2.8 serve as another verification of the theory since the decay rate of the energy in any finite interval is the same due to the following fact proved in the main theorem of [43]:

$$\int_0^R g(t,\omega)d\omega \;=\; \int_{\mathbb{R}_+} \chi_{[0,R]}(\omega) g(t,\omega)d\omega \leq\; \mathcal{O}\Big(\frac{1}{\sqrt{t}}\Big) \text{ as } t \to \infty,$$

for all truncation parameters R.

Therefore, the rate that the energy leaves any finite interval $[0, R]$ is the same since the convergence does not depend on the truncation parameter. As the theory is only for

the energy cascade at the point $\omega = \infty$, it is a challenging task theoretically to obtain a convergence study pointwise with respect to all of the velocity variables in this ω.

We would like to comment that in contrast to Tests 3 and 4, the amount of energy contained in the interval looks indistinguishable for the various truncation parameters in Tests 1 and 2. This can be explained by comparing $\|g^0\|_{L^1(0,R)}$ of the different test cases. The initial energy serves as a kind of reservoir in the finite intervals [56]. Then, for tests like Tests 3 and 4, we increase the amount of initial energy by increasing the size of the interval. That is for $R_1 < R_2 < R_3$ we have

$$\|g^0\|_{L^1(0,R_1)} < \|g^0\|_{L^1(0,R_2)} < \|g^0\|_{L^1(0,R3)}. \tag{2.38}$$

This is not the case (or is negligible) for Tests 1 and 2, seen in figures 2.3, 2.8, due to the initial conditions selected there. The important thing to notice is that the slopes (decay rates) are independent of how much energy is contained in the interval initially, though the amount of energy can vary depending on the initial condition.

The numerical results confirm the theoretical bound (1.16) and show that the decay rate should be $\mathcal{O}(\frac{1}{t^s})$, with $s \in [\frac{1}{2}, 1]$ for various initial data. In section 2.4.3, the cascade rate is described quite well by $\mathcal{O}(\frac{1}{\sqrt{t}})$ for $\gamma = 2$. We then conclude that the cascade rate bound obtained in [43] is optimal.

Chapter 3

A Deep Learning Approximation

Scientific Product: Walton, Tran, Bensoussan. *Deep learning approximation of non-stationary solutions to wave kinetic equations*, Applied Numerical Mathematics (2022).

3.1. Introduction

It is difficult to exaggerate the impact and scope of machine learning within applied mathematics and scientific computing more broadly. A very active branch of scientific machine learning is the development of neural network approximations of the solutions to partial differential equations. As opposed to data-driven discovery of the dynamics of a physical process in which known governing equations may not be used, a physics-informed neural network (PINN) [57, 58] trains the neural network by minimizing the residual of the governing PDE along with minimization terms for the initial and boundary conditions. For completeness, we describe briefly the general idea.

Let $\mathcal{L}$ be a linear or non-linear (integro-)differential operator. Let $(t,x) \in [0,T] \times \Omega \subset \mathbb{R}^+ \times \mathbb{R}^d$ for $d \geq 1$ and $\Gamma = \partial\Omega$ smooth. Assume for simplicity that $u(t,x) \in L^2((0,T);\Omega)$ is the solution to the evolution equation

$$\begin{aligned} (\partial_t + \mathcal{L})u &= 0, \\ \mathcal{B}u &= u_\Gamma, \\ u(0,x) &= u_0(x), \end{aligned} \tag{3.1}$$

where $\mathcal{B}$ denotes a boundary operator and $u_0 \in L^2(\Omega)$.

If we wish to approximate the solution to (3.1) by a neural network $n(t,x;\theta)$ where $\theta \in \Theta$ the set of weights and biases of the neural network, then a common approach is to first define the residual operator, $\mathcal{R}$, such that

$$n(t,x;\theta) \in \mathcal{H}_{\mathcal{R}} = \{\varphi(t,x) \mid \varphi(t,x) \in L^2((0,T);\Omega),\ \mathcal{R}\varphi(t,x) \in L^2((0,T);\mathcal{R})\},$$

with

$$\|\varphi(t,x)\|_{\mathcal{H}_{\mathcal{R}}} = \|\varphi(t,x)\|_{L^2((0,T);\Omega)} + \|\mathcal{R}\varphi(t,x)\|_{L^2((0,T);\Omega)},$$

for

$$\mathcal{R} = (\partial_t + \mathcal{L}). \tag{3.2}$$

We then define the functional

$$J_\theta[n](t,x) = \|\mathcal{R}n(t,x;\theta)\|^2_{L^2((0,T);\Omega)} + \|\mathcal{B}n(t,x;\theta) - u_\Gamma\|^2_{L^2(\Gamma)} + \|n(0,x;\theta) - u_0\|^2_{L^2(\Omega)}. \tag{3.3}$$

To obtain the neural network approximation to $u(t,x)$, one then solves the following stochastic optimization problem

$$\theta^* = \arg\min_\theta J_\theta[n](t,x), \tag{3.4}$$

for the (probably local and non-unique) minimizer θ^* from which we obtain

$$n(t,x;\theta^*) \approx u(t,x).$$

In solving (3.4), one has many choices even when one excludes the usual hyper-parameter tuning that comes with selecting an architecture. In the absence of empirical measurements, a major choice to be made is how one selects inputs $(t,x) \in (0,T) \times \Omega$ for the residual and initial and boundary conditions. A common approach is to take samples from the uniform

distribution and approximate the integrals in (3.3) by taking expectations ([57, 58] and many, many others). As in [59], one may use knowledge about the underlying distribution of the sample space to select a sampling distribution which more closely mimics empirical measurements and may be more relevant to the approximated solution.

In the present article, we choose neither of the previous approaches and opt for a quasi-Monte Carlo (qMC) approximation of the integrals in (3.3) selecting a low-discrepency sequence as our sample points, the Sobol sequence [60] (see figure 3.4).

We find the qMC approach to be appropriate for WKEs for two reasons. The first is that given the highly non-local nature of the collision term, the approximation of this operator is expensive and when coupled with the number of samples necessary to obtain an accurate approximation of the collision term via MC methods, we find a uniform sampling procedure to be impractical for our needs. The latter method introduces interesting challenges in that the distribution with which we would like to sample the wavenumber domain in order to take expectations is in fact the solution we want to find in the first place.

Thus a qMC approach which requires fewer sample points and being deterministic requires no knowledge of the underlying distribution is a simple and natural choice. For a more detailed analysis of qMC methods in neural network approximations to pdes with applications to fluid dynamics the reader is referred to the works [61, 62].

3.1.0.1. Outline

We now give a short overview of the remainder of the article. To build confidence in our approach, we present a neural network (NN) approximation to the solution of a Smoluchowski coagulation equation (SCE) in section 2. We choose this examples because, 1) as has been discussed elsewhere ([31, 43]) the SCE can be considered as a special case of a 3-WKE and 2) unlike for the 3-WKE the SCE has a known analytic solution which we may

compare our approximation against. In section 3, we present results for a neural network representation of the non-stationary solution to a 3-WKE. The results are compared with previously derived theoretical results for the decay rate of the total energy in any finite interval of the wavenumber domain. As a means of validation for the neural network model, we give solutions for the same 3-WKE obtained via a finite volume scheme (FVS).

3.2. The Smoluchowski Coagulation Equation

Before giving results for wave kinetic equations, we provide a check of the method on a similar type of equation, the Smoluchowski coagulation equation, which has a known analytic solution in contrast to wave kinetic equations. Comparisons of the SCE and 3-WKE can be found in [31, 43, 63] and references therein.

The SCE may be written in the form [46]

$$\begin{aligned} v\partial_t f(t,v) &= -\partial_v \mathcal{Q}_{Smol}[f](t,v) \quad (t,v) \in \mathbb{R}^+ \times \mathbb{R}^+, \\ f(0,v) &= f_0(v) \geq 0 \quad v \in \mathbb{R}^+ := (0,\infty), \end{aligned} \tag{3.5}$$

where the $f(t,v) \geq 0$ gives the density of particles at time t with volume v and

$$\mathcal{Q}_{Smol}[f](t,v) = \int_0^v \int_{v-v_1}^{\infty} a(v_1,v_2) v_1 f(t,v_1) f(t,v_2) \mathrm{d}v_2 \mathrm{d}v_1, \tag{3.6}$$

where the kernel is given by $a(v_1,v_2) = v_1 v_2$ in what follows. Let us define

$$m(t,v) = v f(t,v),$$

then the total volume can be defined as

$$\mathcal{V}(t) = \int_0^{\infty} m(t,v) dv.$$

We will consider the case where

$$f_0(v) = \frac{e^{-v}}{v}, \tag{3.7}$$

i.e.,

$$m_0(v) = e^{-v}, \tag{3.8}$$

and the analytic solution corresponding to this initial condition is given by [46]

$$f(t,v) = e^{-Tv}\frac{I_1(2vt^{1/2})}{v^2 t^{1/2}}, \tag{3.9}$$

where

$$I_1(v) = \frac{1}{\pi}\int_0^\pi e^{v\cos\theta}\cos\theta \mathrm{d}\theta,$$

is the modified Bessel function of the first kind and

$$T = \begin{cases} 1+t & t \le T_{gel} \\ 2t^{1/2} & t > T_{gel} \end{cases},$$

in equation (3.9) above.

This SCE undergoes a gelation phenomenon at time $T_{gel} = 1$, that is,the particle number density is entirely concentrated at $v = \infty$ (again, see [46] and the references therein). This also means that $\mathcal{V}(t) = 1$ if $t \in [0, T_{gel})$ and $\mathcal{V}(t) = t^{-1/2}$ if $t \ge T_{gel}$.

3.2.1. NN Representation of the SCE

To begin, instead of (3.5), we solve the evolution equation for $m(t,v)$. Then,

$$\begin{aligned} &\partial_t m(t,v) = -\partial_v \mathcal{Q}_{Smol}[m](t,v) \quad (t,v) \in \mathbb{R}^+ \times \mathbb{R}^+, \\ &m(0,v) = m_0(v) = v f_0(v) \ge 0 \qquad\qquad\qquad v \in \mathbb{R}^+, \end{aligned} \tag{3.10}$$

and the Smoluchowski collision operator is then written as

$$\mathcal{Q}_{Smol}[m](t,v) = \int_0^v \int_{v-v_1}^{\infty} a(v_1,v_2)m(t,v_1)\frac{m(t,v_2)}{v_2}\mathrm{d}v_2\mathrm{d}v_1. \tag{3.11}$$

Let $m(t,v;\theta)$ be the neural network approximation to the solution of (3.10) with θ denoting the weights and biases of the neural network. The NN approximation is computed for $(t,v) \in [0,T] \times [0,R]$ for some time $T > 0$ and truncation value $R > 0$.
To compute the NN approximation $m(t,v;\theta)$, the problem (3.10) is recast as the following functional minimization problem [**?**, 64]

$$m(t,v;\theta^*) = \min_{\theta\in\Theta} J_{Smol}[m](t,v;\theta), \tag{3.12}$$

with $\theta^* \in \Theta$ a minimizing set of parameters and $J_{Smol}[m](t,v;\theta)$ is

$$J_{Smol}[m](t,v;\theta) = \|\mathcal{R}m(t,v;\theta)\|^2_{L^2((0,T]\times[0,R])} + \|m(0,v;\theta) - m_0(v)\|^2_{L^2([0,R])}, \tag{3.13}$$

where $\mathcal{R}$ is the residual operator defined by

$$\mathcal{R}m(t,v;\theta) = \partial_t m(t,v;\theta) + \partial_v \mathcal{Q}[m](t,v;\theta). \tag{3.14}$$

The functional $J_{Smol}[m]$ and collision term $\mathcal{Q}_{Smol}[m]$ are discretized via a Quasi-Monte Carlo method ([61], [62])). Sample points for the functional are drawn from the Sobol sequence ([65]) in the unit square which are then mapped to the discrete set $S \subset (0,T] \times [0,R]$. For the collision term, for each v, we draw samples from the Sobol sequence in the unit square as before, but map the variables (v_1, v_2) to the discrete intervals $V_1 \subset [0,v]$ and, for each v_1, $V_2 \subset [v - v_1, R]$, respectively.

The discretized collision operator is then given by

$$\hat{\mathcal{Q}}_{Smol}[m](t,v;\theta) = \frac{v}{|V_1|}\sum_{v_1\in V_1} v_1 m(t,v_1;\theta)\left(\frac{|R+v_1-v|}{|V_2|}\sum_{v_2\in V_2} m(t,v_2;\theta)\right), \tag{3.15}$$

which leads to the semi-discretized residual

$$\hat{r}(t,v;\theta) = \partial_t m(t,v;\theta) + \partial_v \hat{\mathcal{Q}}_{Smol}[m](t,v;\theta), \tag{3.16}$$

and so the semi-discrete functional $\hat{J}_{Smol}[m](t,v;\theta)$ to be minimized is

$$\hat{J}_{Smol}[m](t,v;\theta) = \frac{1}{|S|}\sum_{(t,v)\in S}\hat{r}(t,v;\theta)^2 + \frac{1}{|S_0|}\sum_{v\in S_0}(m(0,v;\ \theta) - m_0(v))^2, \tag{3.17}$$

where the set $S_0 \subset S$, $|S_0| \leq |S|$ denotes a subset of the volume samples in S.

3.3. A 3-Wave Kinetic Equation

In [63], a new identity for the energy of the solutions to 3-WKEs [43] is presented which takes the form of a conservation law [46]. This is the equation we shall study in the present paper due to the simplicity provided by this form of the equation. Specifically, there is no need to compute the resonant manifolds of the system, though we emphasize that the method does not depend on this simplification.

Let us write the equation to be solved. The equation is equivalent to (4.1) and is explained in [43, 63]

$$\begin{aligned} \partial_t g(t,p) &= p\partial_p \mathcal{Q}[g](t,p) \quad (t,p)\in\mathbb{R}^+\times\mathbb{R}^+, \\ g(0,p) &= g_0(p) \geq 0 \quad (t,p)\in\{0\}\times\mathbb{R}^+, \end{aligned} \tag{3.18}$$

where the collision term is given by

$$\mathcal{Q}[g](t,p) = -2\int_0^p\int_0^p (p_1,p_2)^{\frac{\gamma}{2}-1} g_1 g_2 \chi_{p,1,2} \mathrm{d}p_{2,1} + \int_0^\infty\int_0^\infty (p_1,p_2)^{\frac{\gamma}{2}-1} g_1 g_2 \chi_{p,1,2} \mathrm{d}p_{2,1}, \tag{3.19}$$

where we have used the notation $g_i = g(t, p_i)$ for $i = 1, 2$, $\mathrm{d}p_{2,1} = \mathrm{d}p_2\mathrm{d}p_1$ and $\chi_{p,1,2} = \chi\{p < p_1 + p_2\}$ where $\chi\{A\}$ is the set characteristic function of some set A. The parameter γ is the degree of the kernel as discussed in the introduction and $\gamma = 2$ in the present work which corresponds to acoustic wave systems.

As in the previous section, we define a neural network approximation, $g(t, p; \theta)$ to be a solution to the optimization problem [66]

$$g(t, p; \theta^*) = \min_{\theta \in \Theta} J[g](t, p; \theta), \tag{3.20}$$

with $\theta^* \in \Theta$ a minimizing set of parameters of the functional

$$J[g](t, p; \theta) = \|r(t, p; \theta)\|^2_{L^2(\mathbb{R}^+ \times \mathbb{R}^+)} + \|g(0, p; \theta) - g_0(p)\|^2_{L^2(\mathbb{R}^+)}, \tag{3.21}$$

where, again, $\mathcal{R}$ denotes the residual operator of the evolution equation defined by

$$\mathcal{R}g(t, p; \theta) = \partial_t g(t, p; \theta) - p\partial_p \mathcal{Q}[g](t, p; \theta). \tag{3.22}$$

The functional (3.21) is approximated via a Quasi-Monte Carlo method with sample points drawn from the Sobol sequence in the unit square and then transformed to some truncated rectangle of the time, wavenumber domain, i.e. we generate the set $W \sim Sobol([0, T] \times [0, R])$, for $T, R > 0$ truncation parameters of the time and wavenumber domain, respectively. The residual (3.22) is approximated similarly where the set of sample points for (p_1, p_2) are given by $P_1 \sim Sobol(0, p)$ for each p, and $P_2 \sim Sobol(p - p_1, p)$ for each p, p_1 in the first term of (3.19) and $\hat{P}_2 \sim Sobol(p - p_1, R)$ for each (p, p_1).

To see how these sample sets are defined, note that the collision term can be rewritten as

$$\mathcal{Q}[g](t,p) = -2\int_0^p \int_{p-p_1}^p (p_1 p_2)^{\frac{\gamma}{2}-1} g_1 g_2 \mathrm{d}p_{2,1} + \int_0^p \int_{p-p_1}^R (p_1 p_2)^{\frac{\gamma}{2}-1} g_1 g_2 \mathrm{d}p_{2,1}, \tag{3.23}$$

where we have applied the truncation parameter of the wavenumber domain to the second expression and enforced the restrictions $p-p_1 < p_2$ in both terms to satisfy the characteristic set function and $p_1 < p$ to guarantee positivity of the variable p_2 in the second term. Thus, using the above collision operator and sample sets defined in the previous paragraph, we can define the discrete collision operator to be

$$\begin{aligned}\hat{\mathcal{Q}}[g](t,p) = \frac{-2p}{|P_1||P_2|}\sum_{p_1\in P_1} p_1^{\frac{\gamma}{2}} g_1 \Bigg(\sum_{p_2\in P_2} p_2^{\frac{\gamma}{2}-1} g_2\Bigg) \\ + \frac{p}{|P_1||\hat{P}_2|}\sum_{p_1\in P_1} p_1^{\frac{\gamma}{2}-1} g_1 \Bigg[(R-p+p_1)\Big(\sum_{p_2\in \hat{P}_2} p_2^{\frac{\gamma}{2}-1} g_2\Big)\Bigg].\end{aligned} \tag{3.24}$$

Using the discrete collision operator, we can define the semi-discrete residual by

$$\hat{\mathcal{R}}g(t,p;\theta) = \partial_t g(t,p;\theta) - p\partial_p \hat{\mathcal{Q}}[g](t,p;\theta), \tag{3.25}$$

and therefore we can minimize the semi-discrete functional

$$\hat{J}[g](t,p);\theta) = \frac{1}{|W|}\sum_{(t,p)\in W} \hat{\mathcal{R}}g(t,p;\theta)^2 + \frac{1}{|W_0|}\sum_{p\in W_0}(g(0,p;\theta) - g_0(p))^2, \tag{3.26}$$

where $W_0 \subset W$, $|W_0| \leq |W|$ contains only sample points for the wavenumber.

3.4. Numerical Results

In this section, we provide numerical results for the SCE with initial condition (3.7) and compare with the anaylitc solution (3.9). A 3-WKE with initial condition (3.27) is then solved and the decay rate of the total energy is compared with the results obtained in [43].

The same initial condition (3.27) is then used to obtain a solution with the finite volume scheme developed in [63]. The results of the two methods are discussed.

All deep learning tests were implemented in TensorFlow [67] with Keras [68]. Training was accelerated using multiple GPUs utilizing Horovod [69]. All computations were performed on SMU's computing cluster MII. The finite volume computations were performed in Matlab.

3.4.1. Test 1

A simple, yet effective, architecture is chosen. Namely, it is enough to employ a feed-forward network with only two hidden layers with 128 hidden units in each layer utilizing sigmoidal activation functions [70]. Given its simplicity, we can write the neural network (NN) out explicitly as

$$m(t, v; \theta) = \mathbf{W}\sigma(\mathbf{W}_2\sigma(\mathbf{W}_1\vec{x} + \mathbf{b}_1) + \mathbf{b}_2) + b,$$

with the entries of $\mathbf{W}_i, \mathbf{b}_i \sim \mathcal{N}(0, 1)$ for $i = 1, 2$, $\mathbf{W}, b \sim \mathcal{N}(0, 1)$ and $\sigma(\cdot)$ denoting the sigmoid function. To fully discretize the functional (3.17), we employ TensorFlow's built-in automatic differentiation method.

The training samples were drawn from the rectangle $(0, T] \times [0, R]$ using the strategy described above for $T = 0.8$ and $R = 8$. To achieve the results shown in Figure 3.1, only the first 32 Sobol points were used to minimize the residual term (16 time samples and 16 volume samples) and only the first 16 Sobol points were needed to train on the initial data.

Figure 3.1 shows the neural network predictions on unseen data within the training interval. The inputs to the neural network to produce the plot were the first 128 Sobol points transformed to the volume training interval and $t = 0.0, 0.2, 0.4$ and 0.62. We see the NN solution gives a good approximation to the analytic solution.

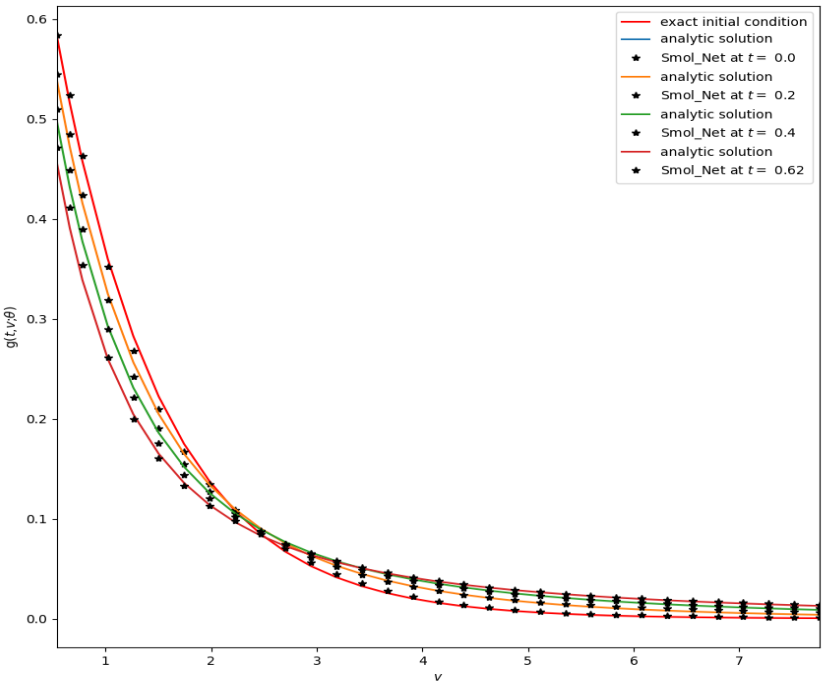


Figure 3.1: Neural Network (NN) approximation to the Smoluchoswki equation (3.10). The solid lines denote the analytic solution (3.9) and $*$ markers denote the NN approximation at equivalent snapshots in time.

In Figure 3.2, we quantify the match of the NN approximation to the analytic solution by giving the sup norm of the error at each snapshot. Here, we supply the first 2^{10} Sobol points within the volume training interval and use the same points in time $t = 0.0, 0.2, 0.4, 0.62$. We note here that while the NN performs well on unseen points within the volume interval $v \in [0, R]$, for $v > R$ we generalization error is larger. This is to be expected given the small training sample size used (16 volume points). We show the absolute error in Figure 3.3 for each snapshot in time over the interval $v \in [0, 1000]$ with 2^{13} Sobol points in this volume interval. This result informed our decision to apply a batch sampling procedure for the wave kinetic equation, the details of which are discussed below.

3.4.2. Test 2

We choose the initial condition

$$g_0(p) = \sqrt{\frac{7}{2\pi}} e^{-\frac{7(p-2)^2}{2}}. \tag{3.27}$$

The neural network was trained using samples from the Sobol sequence in the rectangle $W \sim Sobol((0, T] \times [0, R])$ for $T = 10$ and $R = 10$. As illustrated in figure 3.4, the size of the sample set $|W|$ was 2^{15}, which was broken up into smaller training batches.

The architecture was again chosen to have 2 hidden layers, each with 128 units and sigmoidal activation functions. The loss was again minimized using tensorflow's implementation of ADAM. The collision terms were approximated with the sampling procedure described above with each sample comprised of 32 Sobol points. We consider only the case $\gamma = 2$ here.

We show a few early snapshots of the approximated solution in Figure 3.5. The behaviour seems consistent with predictions in that the L^∞ norm decays as time increases. This is more clearly evidenced by the total energy discussed in the following paragraph. In Figure 3.5,

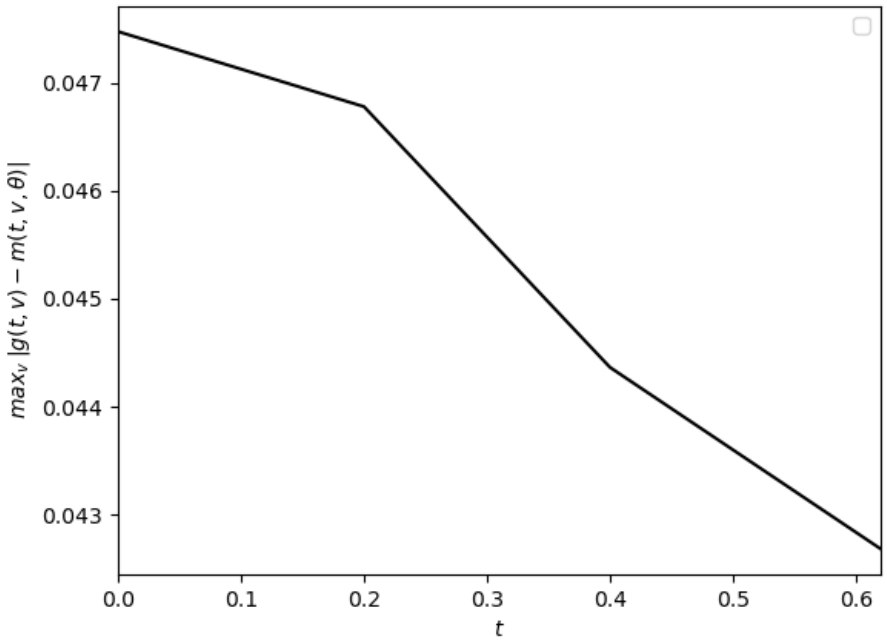


Figure 3.2: Sup norm of the error at $t = 0.0, 0.2, 0.4$ and 0.62 where the maximum is taken over the volume domain $v \in [0, 10]$. This shows the accuracy of the method on unseen data points within the training interval.

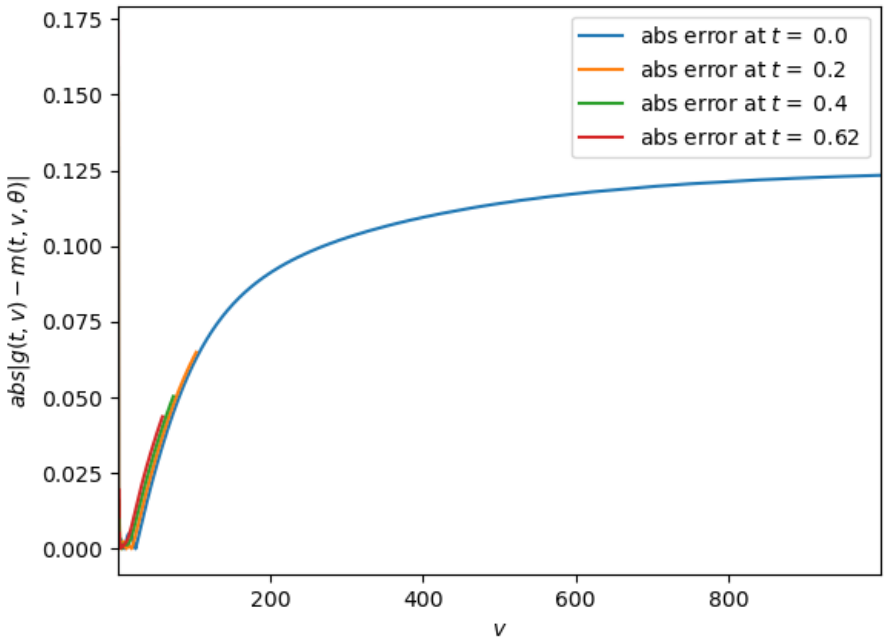


Figure 3.3: Absolute error of the NN approximation versus the analytic solution over the volume interval $v \in [0, 1000]$ with 2^{13} volume samples given as input to the NN. We see a slow growth in the error for large volume numbers.

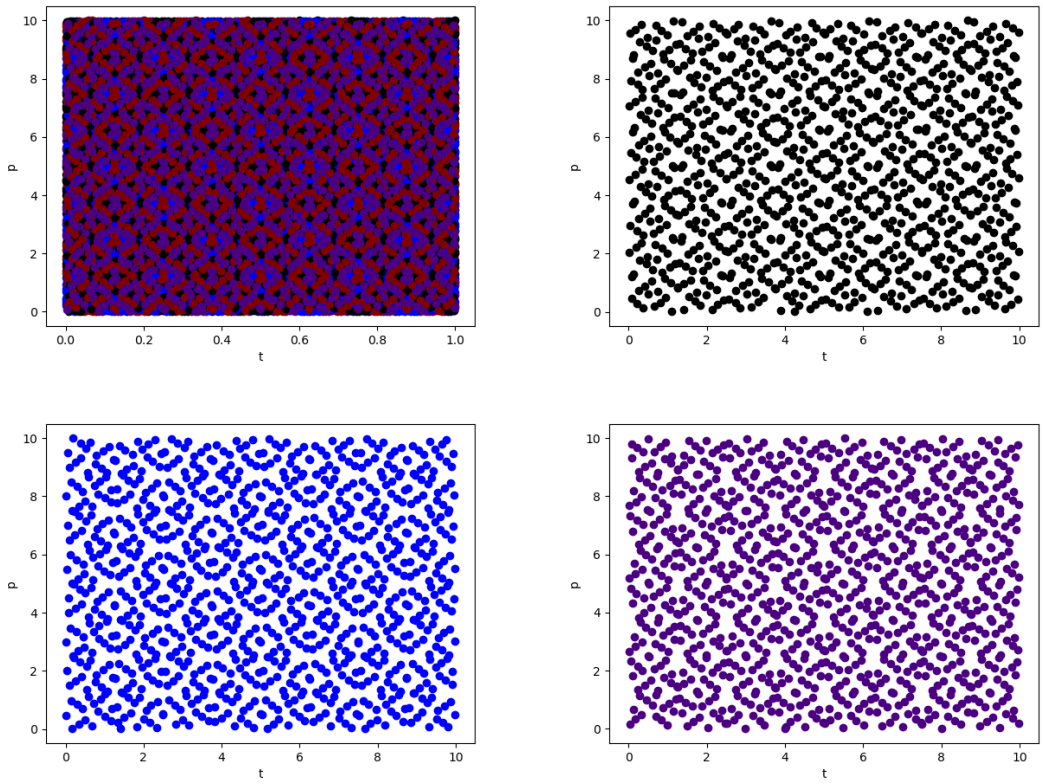


Figure 3.4: Batched samples of Sobol points. (Top Left) All sample points. (Top Right and Bottom Row) Example batches of sampled points.

we also present a comparison with the solution computed the Finite Volume Scheme of [63]. This comparison will be described later in this section, around equation (3.28)-(3.29).

The total energy of the solutions was computed and the log of the total energy is plotted against the log of t. The comparison between the decay of the numerical solution is made with the theoretical decay rate of $t^{-1/2}$, obtained in [43] (see (1.16)), in Figure 3.6. The total energy was predicted up to $t = 148$. From Figure 3.6, the energy decays and the decay is in good agreement with the theoretical rate obtained in [43].

The neural network is able to make reasonable predictions for very large, unseen in training, samples in the wave number domain. In Figure 3.7, we see the predicted solution for wavenumber values up to 1e+06. In contrast, the most we could push for reasonable predicitons in time was $t = 148$ as previously reported in the results shown in figure 3.6. For better predictions in time, we needed to enforce more dense sampling in the time domain. Thus, the neural network was not only trained on the set W for T= 10, but also for $T = 5$ and $T = 2$ with the same number of sample points, 2^{15}.

3.4.3. Comparison with a Finite Volume Scheme

As mentioned earlier, to validate the results, we provide a comparison with another model. Here we use the finite volume scheme presented in [63], which we briefly outline below.

The discretization of the wavenumber domain $p \in [0, R]$, is given as follows. Let $i \in \{0, 1, 2, \ldots, M\} = I_h^M$, with $h \in (0, 1)$ fixed and $M = M(h)$. Define the set $S_d = \{0, \ldots, R\}$ to be the discretization of the interval $[0, R]$. Let

$$S_d = \{p_{i+1/2}\}_{i \in I_h^M}, \ \{p_i\}_{i \in I_h^M \setminus \{0\}} = \frac{p_{i+1/2} + p_{i-1/2}}{2}, \ \{\Delta p_i\}_{i \in I_h^M \setminus \{0\}} = p_{i+1/2} - p_{i-1/2} \leq h, \tag{3.28}$$

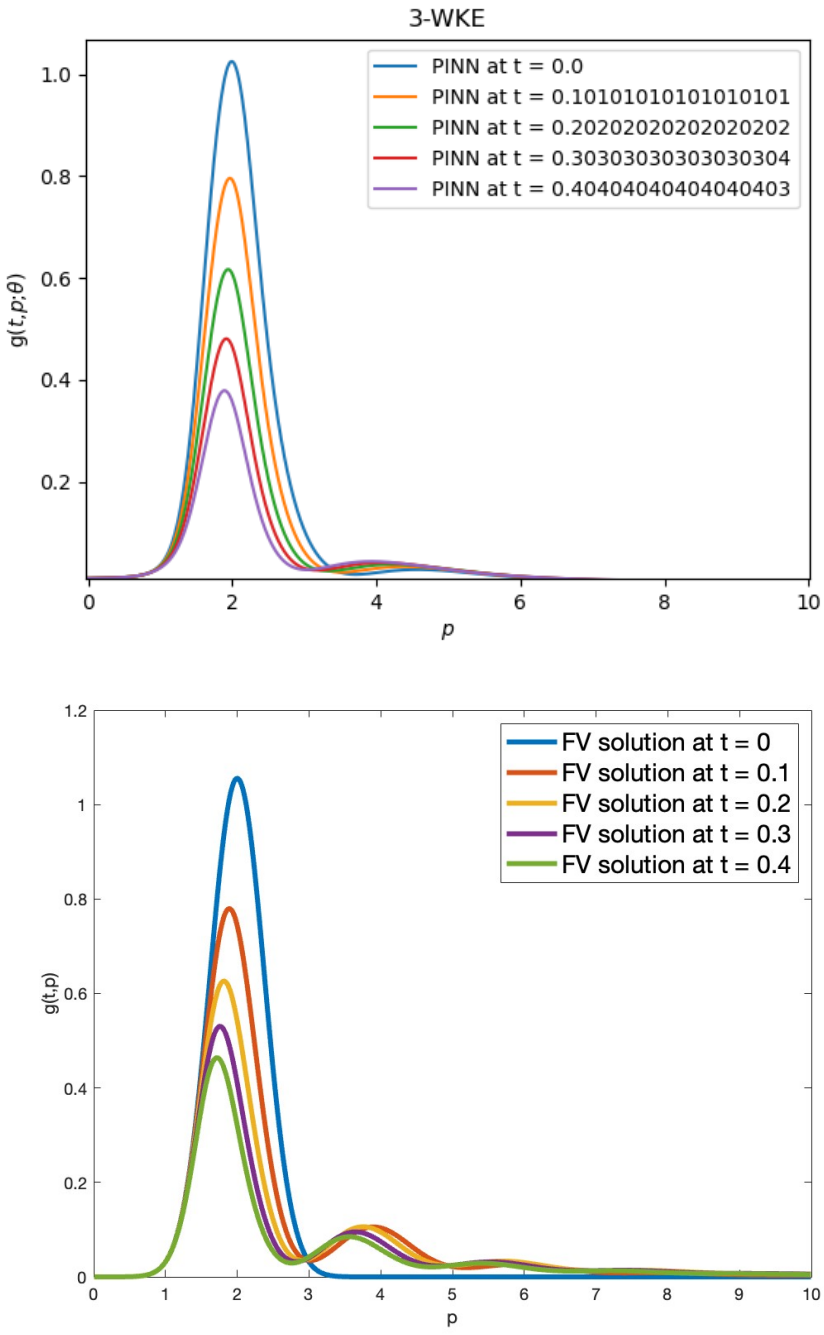


Figure 3.5: Top Picture: A few snapshots of the NN approximation corresponding to initial condition (3.27). Bottom Picture: Comparative snapshots of the FVS solution for the same initial condition (3.27).

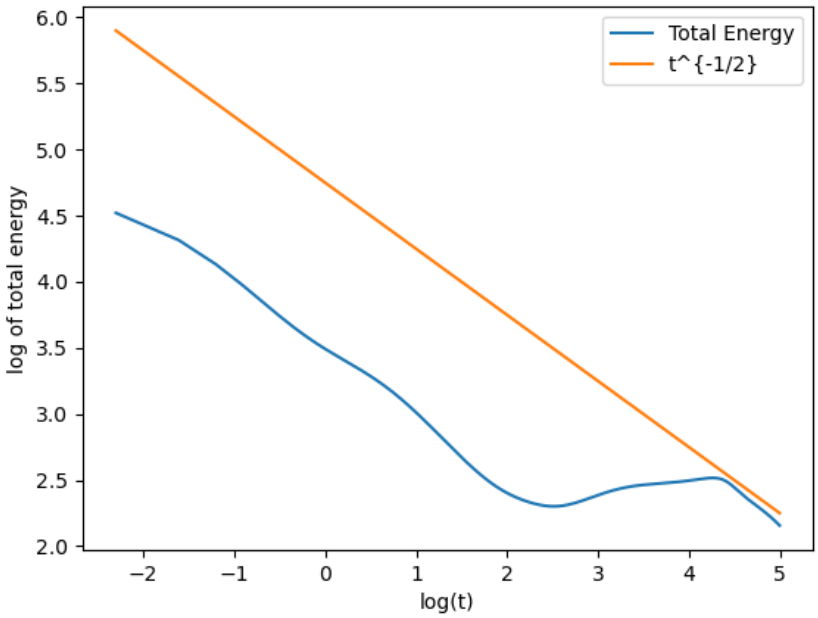


Figure 3.6: Top Picture: Log-Log plot of the total energy corresponding to initial condition (3.27) as predicted up to $t = 148$ the NN. The decay rate is consistent with the decay rates found in 2.

define the faces, pivots and step-size respectively, with $p_{1/2} = 0$ and $p_{M+1/2} = R$. For simplicity, we use a uniform grid and set $h = 0.01$ which leads to

$$S_d = \{ih\}_{i\in I_h^M}, \quad \{p_i\}_{i\in I_h^M\setminus\{0\}} = \frac{h}{2}(2i-1), \quad \{\Delta p_i\}_{i\in I_h^M\setminus\{0\}} = h \in (0,1).$$

The set $T_N = \{0,\dots,T\}$ with $N+1$ nodes and $T = 148$ is the maximum time. We fix the time step to be $\Delta t = \frac{T}{N} = 0.005$, and denote by $t_n = \Delta t \cdot n$ for $n \in \{0,\dots,N\}$. We approximate equation (3.18) by

$$g^{n+1}(p_i) = g^n(p_i) + \lambda_i\Big(Q^n_{i+1/2}\Big[\frac{g}{p}\Big] - Q^n_{i-1/2}\Big[\frac{g}{p}\Big]\Big), \tag{3.29}$$

where $\lambda_i = \frac{p_i\Delta t}{\Delta p_i}$, and

$$Q^n_{i+1/2}\Big[\frac{g}{p}\Big] - Q^n_{i-1/2}\Big[\frac{g}{p}\Big] = -2\Big(Q^n_{1,i+1/2}\Big[\frac{g}{p}\Big] - Q^n_{1,i-1/2}\Big[\frac{g}{p}\Big]\Big) + \Big(Q^n_{2,i+1/2}\Big[\frac{g}{p}\Big] - Q^n_{2,i-1/2}\Big[\frac{g}{p}\Big]\Big),$$

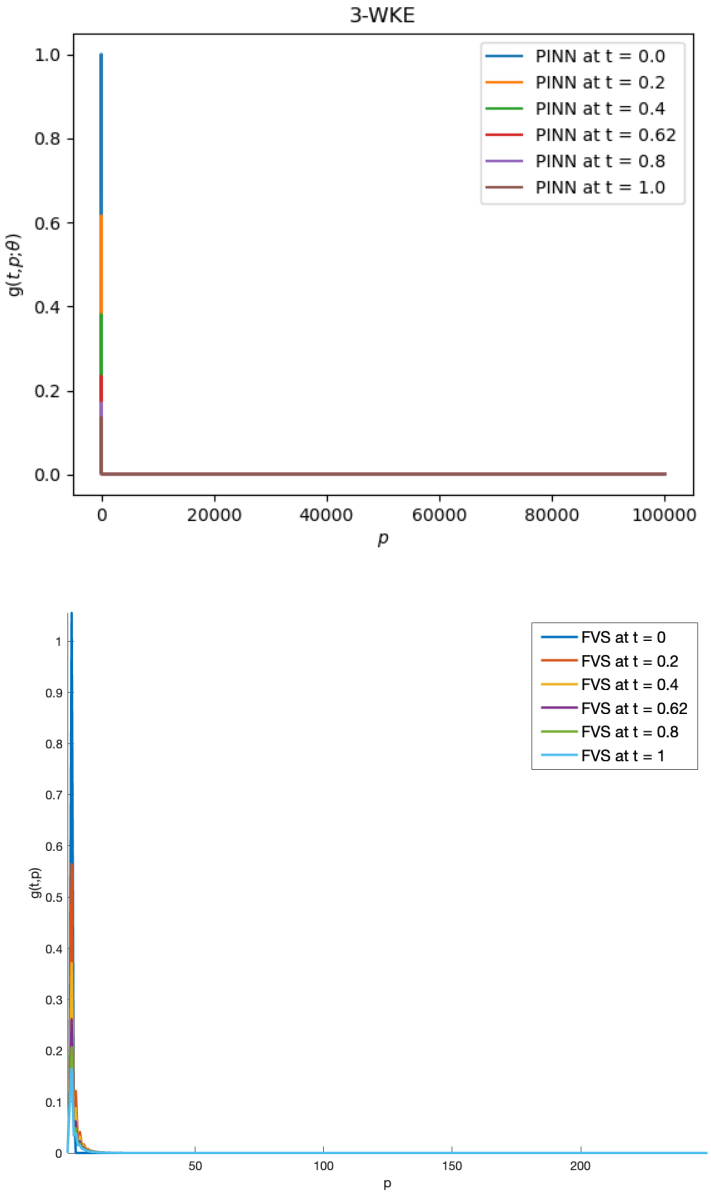


Figure 3.7: Top picture: Neural network prediction corresponding to initial condition (3.27) for wavenumbers up to 1e+06. Bottom picture: The finite volume solution for wavenumbers up to $p = 250$, the largest value for which stability is maintained without further decreasing Δt. The two figures highlight a key strength of the presented method, in that the neural network approximation is consistent with more traditional solvers but is able to produce results for computationally prohibitive values of the wavenumber for traditional methods while maintaining positivity and stability.

with

$$Q^n_{1,i+1/2}\left[\frac{g}{p}\right] = \sum_{m=1}^{i} \Delta p_m \frac{g^n(p_m)}{p_m} \left(\sum_{j=1}^{i} \Delta p_j \frac{g^n(p_j)}{p_j} a(p_m, p_j) \chi\Big\{ p_{i+1/2} < p_m + p_j \Big\} \right), \tag{3.30}$$

$$Q^n_{2,i+1/2}\left[\frac{g}{p}\right] = \sum_{m=1}^{M} \Delta p_m \frac{g^n(p_m)}{p_m} \left(\sum_{j=1}^{M} \Delta p_j \frac{g^n(p_j)}{p_j} a(p_m, p_j) \chi\Big\{ p_{i+1/2} < p_m + p_j \Big\} \right), \tag{3.31}$$

where we have used the midpoint rule to approximate the integrals in equation (3.18) and we choose an explicit time stepping method.

The initial condition (3.27) is approximated by

$$g^0(p_i) = \frac{1}{\Delta p_i} \int_{p_{i-1/2}}^{p_{i+1/2}} g_0(p) \mathrm{d}p \approx g_0(p_i),$$

by again employing the midpoint rule.

We draw the reader's attention again to Figure 3.5 where a few comparative snapshots of the solution are provided. Qualitatively, the two models seem to agree and capture the main features of the theorized behavior of solutions. Namely, an evacuation of the energy within any finite interval. Further, as is typical of simple feedforward architectures, the finer oscillations seen in the FVS solutions appear to be averaged out in the NN solution as expected. For our purposes, the dynamics captured by the simple feed-forward architecture we have employed here are enough to confirm the theory presented in [43]. We leave it to a future work to investigate more complicated architectures. For example, the architecture described in [71] is able to capture highly oscillatory solutions for the stationary Navier-Stokes equations.

In Figure 3.6 (bottom) we see the decay rate of the total energy as provided by the computed solution of the FVS, which is compared with the *predicted* values by the NN (top). Both models are in good agreement with the theorized bound on the rate of decay [43, 63], though the FVS appears to capture a slightly faster rate of decay.

Looking back to Figure 3.7, we make a comparison with the computed solution of the finite volume method for large wavenumber values with predictions provided by the neural network. Here, we have increased h to 0.8 and $R = 250$ while keeping the timestep fixed at $\Delta t = 0.005$ as in the previous figures. The two figures highlight a key strength of the presented deep learning method, in that the neural network approximation is consistent with more traditional solvers but is able to produce results for computationally prohibitive values of the wavenumber for traditional methods, while maintaining positivity and stability. Indeed, constructing positivity preserving schemes for PDEs is a very important direction of research [52–54]. For the FVS (3.29), the positivity and thus stability is lost for wavenumber values larger than 250 with the timestep set at $\Delta t = 0.005$. In order to preserve the positivity of the solutions produced by the FVS (3.29), the CFL condition is restrictive and the time step Δt needs to be chosen sufficiently small as shown in Proposition 3.1 of [63]. As thus, being able to preserve the positivity of the solutions is indeed a very important feature of the presented deep learning approximation.

3.5. Conclusions

We present a deep learning approximation, stochastic optimization based, method for the 3-wave kinetic equation, studied theoretically in [43] and numerically in [63]. We first apply the method to a Smoluchowski coagulation equation with multiplicative kernel for which an analytic solution exists. The deep learning method is proved to give a good approximation of the analytic solution. Next, the learning approach is then used to approximate the solution of the 3-wave kinetic equation. The deep learning approximation is tested and proved to be

as good as the Finite Volume approximation introduced in [63] for the same equation and both approximations are in good agreement with the theoretical results of [43].

Chapter 4

An Inverse Problem and Reconstruction of the Wave-Action for 3-WKEs

Scientific Product: Walton, Tran, Bensoussan. *An inverse problem for wave kinetic equations*, (in preparation).

4.1. Introduction A potentially challenging aspect in determining the behavior of a particular wave system is the derivation of the interaction coefficient. For example, for gravity waves on deep water, Zakharov in [72] derived what is currently the *most compact* form of the interaction coefficient [9] which we provide below. The interaction coefficient for pure gravity waves on deep water is given by

$$\begin{aligned}
W_{12}^{34} = -\frac{1}{16\pi^2}\frac{1}{(p_1p_2p_3p_4)^{\frac{1}{4}}}\Big\{ &-12p_1p_2p_3p_4 \\
&-\frac{2}{g^2}(\omega_1+\omega_2)^2\Big[\omega_3\omega_4((\mathbf{p}_1\cdot\mathbf{p}_2)-p_1p_2)+\omega_1\omega_2((\mathbf{p}_3\cdot\mathbf{p}_2)-p_3p_4)\Big] \\
&-\frac{2}{g^2}(\omega_1-\omega_3)^2\Big[\omega_2\omega_4((\mathbf{p}_1\cdot\mathbf{p}_3)+p_1p_3)+\omega_1\omega_3((\mathbf{p}_2\cdot\mathbf{p}_4)-p_2p_4)\Big] \\
&-\frac{2}{g^2}(\omega_1-\omega_4)^2\Big[\omega_2\omega_3((\mathbf{p}_1\cdot\mathbf{p}_4)+p_1p_4)+\omega_1\omega_4((\mathbf{p}_2\cdot\mathbf{p}_3)-p_2p_3)\Big] \\
&+((\mathbf{p}_1\cdot\mathbf{p}_2)+p_1p_2)((\mathbf{p}_3\cdot\mathbf{p}_4)+p_3p_4) \\
&+(-(\mathbf{p}_1\cdot\mathbf{p}_3)+p_1p_3)(-(\mathbf{p}_2\cdot\mathbf{p}_4)+p_2p_4) \\
&+(-(\mathbf{p}_1\cdot\mathbf{p}_4)+p_1p_4)(-(\mathbf{p}_2\cdot\mathbf{p}_3)+p_2p_3) \\
&+4(\omega_1+\omega_2)^2\frac{((\mathbf{p}_1\cdot\mathbf{p}_2)-p_1p_2)((\mathbf{p}_3\cdot\mathbf{p}_4)-p_3p_4)}{\omega(\mathbf{p}_1+\mathbf{p}_2)^2-(\omega_1+\omega_2)^2} \\
&+4(\omega_1-\omega_3)^2\frac{((\mathbf{p}_1\cdot\mathbf{p}_23-p_1p_3)((\mathbf{p}_2\cdot\mathbf{p}_4)-p_2p_4)}{\omega(\mathbf{p}_1-\mathbf{p}_3)^2-(\omega_1-\omega_3)^2} \\
&+4(\omega_1-\omega_4)^2\frac{((\mathbf{p}_1\cdot\mathbf{p}_4)-p_1p_4)((\mathbf{p}_2\cdot\mathbf{p}_3)-p_2p_3)}{\omega(\mathbf{p}_1-\mathbf{p}_4)^2-(\omega_1-\omega_2)^2},
\end{aligned}$$

where $\omega_i = \omega(\mathbf{p}_i)$ for $i = 1, 2, 3, 4$ is the dispersion relation for a given wavevector $\mathbf{p}$. We denote the wavenumber by $p = |\mathbf{p}|$ throughout.

In this work, we provide a method for the approximation of such interaction coefficients using the method of weighted residuals. The interaction coefficient is approximated by a linear expansion of Legendre basis functions, whose coefficients are the solution to our inverse problem.

As previously mentioned, we consider the isotropic 3-WKE. The dispersion relations we consider are given by $\omega = c|\mathbf{p}|$ which, after absorbing the constant, we simply write $\omega = p$ in the sequel. The the isotropic 3-WKE is given by

$$\begin{aligned}
\partial_t f(t,p) &= Q[f](t,p), \quad \omega \in \mathbb{R}_+,\ f(0,p) = f_0(p), \\
Q[f](t,p) &= \int_0^\infty \int_0^\infty [R(p,p_1,p_2) - R(p_1,p,p_2) - R(p_2,p_1,p)]dp_1 dp_2, \\
&\qquad R(p,p_1,p_2) := \delta(p-p_1-p_2)\left[\mathcal{U}(p_1,p_2)f_1 f_2 - \mathcal{U}(p,p_1)f f_1 - \mathcal{U}(p,p_2)f f_2\right],
\end{aligned} \tag{4.1}$$

where the interaction coefficient, $\mathcal{U}(\cdot,\cdot)$, satisfies $|\mathcal{U}(p_1,p_2)| = (p_1 p_2)^{\gamma/2}$, in which $\gamma \in [1,2]$ is constant. Similar to our notation for the dispersion relation, we us the shorthand for the wave-action $f_i = f(t,p_i)$ and $f = f(t,p)$. Equation (4.1) can be written in the equivalent form

$$\partial_t f(t,p) = \mathbb{Q}[f](t,p), \quad p \in \mathbb{R}_+,\ f(0,p) = f_0(p), \tag{4.2}$$

in which $\mathbb{Q}$ is defined by

$$\begin{aligned}
\mathbb{Q}[f](t,p) = &\int_0^p \Big[|\mathcal{U}(p_1,p-p_1)|f(p_1)f(p-p_1) - |\mathcal{U}(p,p_1)|f(p)f(p_1) - |\mathcal{U}(p,p-p_1)|f(p)f(p-p_1)\Big]\mathrm{d}p_1 \\
&- 2\int_0^\infty \Big[|\mathcal{U}(p,p_1)|f(p)f(p_1) - |\mathcal{U}(p+p_1,p_1)|f(p+p_1)f(p_1) - |\mathcal{U}(p+p_1,p)|f(p+p_1)f(p)\Big]\mathrm{d}p_1.
\end{aligned} \tag{4.3}$$

It is this form of the isotropic 3-WKE which often draws comparison with the Smoluchowski coagulation equation,

$$\partial_t f(t,v) \;=\; \mathbb{Q}_{Smol}[f](t,v),\;\; f(0,v) = f_0(v), \tag{4.4}$$

with the Smoluchowski collision term given by

$$\mathbb{Q}_{Smol}[f](t,v) \;=\; \int_0^v |\mathcal{U}(v_1, v-v_1)| f(v_1) f(v-v_1) \mathrm{d}v_1 \;\; - 2\int_0^\infty |\mathcal{U}(v, v_1)| f(v) f(v_1) \mathrm{d}v_1, \tag{4.5}$$

where v represents the particle volume and $f(t,v)$ is the particle number density.

We do not go into a detailed comparison here but refer the reader to the works [18, 31, 32, 43, 63, 73] and the references therein. However, at this point we would like to mention works on a similar inverse problem involving the Smoluchowski coagulation equation. In the works [74, 75], the authors present a regularized least squares approach to to extract the aggregation kernel (what is our interaction coefficient) with the assumption that the solutions are self-similar and that gelation does not occur, i.e. $\gamma < 1$. In another work [76], the authors drop these assumptions on the solution and do not preclude the possibility of gelation. They reformulate the method of weighted residuals for the forward problem to instead obtain an approximation for the aggregation kernel.

Given that the 3-WKE, for $\gamma > 1$, undergoes a process similar to the gelation behavior observed in the Smoluchowski equation, specifically the energy accumulates at the point $p = \infty$ and undergoes infinitely many blow-up times (see [43, 63], we are better able to draw inspiration from the work [76]. The present article provides an adaptation of the method of weighted residuals applied to the inverse problem for Smoluchowski equation to the 3-WKE, the details of which are the topic of the next section.

However, in the context of wave turbulence, to our knowledge, the interaction coefficients are always time independent. Therefore, in our case, we use only the initial condition and one measurement at some other point in time to solve the inverse problem. From there, we can then reconstruct the solution at every other point in time using our approximated interaction kernel.

4.2. The Inverse Problem

Let us define the residual which will need to be minimized to obtain an approximation of the interaction coefficient. The residual is given by

$$\mathfrak{R}(t,p) = \partial_t f(t,p) - \mathbb{Q}[f](t,p). \tag{4.6}$$

We will require that for any test function $\psi_\ell(p)$ that $\mathfrak{R}$ satisfies

$$\int_{\mathbb{R}^+} \mathfrak{R}(t,p)\psi_\ell(p)\mathrm{d}p = 0, \tag{4.7}$$

as prescribed by the method of weighted residuals. This is quite a general method and, as has been pointed out in the work [77], different choices of the test function $\psi(p)$ and representation of $|\mathcal{U}(p_1,p_2)|$ lead to other well-known methods. For example, if we represent the interaction kernel by a neural network and choose $\psi_\ell(p) = \delta(p-p_\ell)$ we get a physics-informed neural network. However, employing neural networks does not seem to be necessary in the cases we consider. The interaction coefficient will be represented as a linear combination of basis functions

$$P_M|\mathcal{U}(p_1,p_2)| \approx \sum_{i=1}^{M} P_i(p_1,p_2)\mathcal{U}_i, \tag{4.8}$$

where the $P_i(p_1,p_2)$ is a tensor product of 1 dimensional Legendre basis functions defined over the p_1, p_2 domain. The coefficients $\mathcal{U}_i$ are the unknowns we wish to obtain. We will use

the notation $\mathbb{Q}_\varphi[f](t,p)$ to mean the interaction coefficient of (4.2) has been replaced by the expansion (4.8).

Setting $\psi_\ell(p) = \delta(p - p_\ell)$ and splitting the integral (4.7) we arrive at the matrix equation

$$\mathbf{Q}^0\mathbf{u} = \delta\mathbf{f} \tag{4.9}$$

where $\mathbf{Q}^0 \in \mathbb{R}^{L\times M}$, $\delta\mathbf{f} \in \mathbb{R}^L$ and $\mathbf{u} \in \mathbb{R}^M$ where L is the number of nodes in the wavenumber grid and we assume that $M < L$, i.e. the system of equations (4.9) is overdetermined. The entries of $\mathbf{Q}^0$ are

$$\begin{aligned}\mathbf{Q}^0_{\ell,m} = \int_{\mathbb{R}^+} \delta(p-p_\ell)\Bigg(\int_0^p \Big[P_m(p_1, p-p_1) f_1^0 f_{-1}^0 - P_m(p,p_1) f^0 f_1^0 - P_m(p, p-p_1) f^0 f_{-1}^0\Big] \mathrm{d}p_1 \\ -2\int_0^\infty \Big[P_m(p,p_1) f_p^0 f_1^0 - P_m(p+p_1,p_1) f_{+1}^0 f_1^0 - P_m(p+p_1,p) f_{+1}^0 f_1^0\Big] dp_1 \Bigg) \mathrm{d}p\end{aligned} \tag{4.10}$$

where the notation $f^0_{\pm 1} = f^0(p \pm p_1)$ and f^0, f_1^0 denote $f^0(p)$ and $f^0(p_1)$, respectively. Which we find, similarly to [76], has a large condition number. The superscript Specific values will be given in the results section. The entries of $\delta\mathbf{f}$ are given by

$$\delta\mathbf{f} = \frac{1}{\Delta t}\Big(\mathbf{f}^k - \mathbf{f}^0\Big) = \frac{1}{\Delta t}\begin{pmatrix} f_1^k - f_1^0 \\ \vdots \\ f_\ell^k - f_\ell^0 \\ \vdots \\ f_P^k - f_P^0 \end{pmatrix}, \tag{4.11}$$

with $f_\ell^i = f(t^i, p_\ell)$, for $i = 0, k$, where k is some future point in time with $\Delta t = t^k - t^0$. The entries of the vector $\mathbf{u}$ are simply the coefficients in the expansion 4.8,

$$\mathbf{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_m \\ \vdots \\ u_M \end{pmatrix}, \tag{4.12}$$

as expected.

We then have the following optimization problem

$$\arg\min_{\mathbf{u}} \left\{ \|\delta\mathbf{f} - \mathbf{Q}^0\mathbf{u}\|_2^2 + \lambda\|\mathbf{u}\|_2^2 \right\}, \tag{4.13}$$

i.e. a Tikhanov regularized least-squares problem of the overdetermined system (4.9), with λ a regularization parameter.

Once the $\mathbf{u}$ have been computed, we can then reconstruct solutions $\tilde{f}$ for the wave-action with the following composition. Using Δt above, compute

$$\tilde{\mathbf{f}}^1 = \mathbf{f}^0 + \Delta t \mathbf{Q}^0\mathbf{u}. \tag{4.14}$$

Use the approximated solution to compute $\mathbf{Q}^1$ and repeat. The update is then given by

$$\tilde{\mathbf{f}}^n = \tilde{\mathbf{f}}^{n-1} + \Delta t \mathbf{Q}^{n-1}\mathbf{u}. \tag{4.15}$$

4.3. Numerical Results To obtain the data $\mathbf{f}^k$, we use the finite volume method to solve the conservative form of the 3-WKE for the energy density $g(t,p) = pf(t,p)$ as proposed in [63]. We choose the initial condition

$$g_0(p) = pe^{-p} \tag{4.16}$$

which is plotted in Figure 4.1 with the corresponding initial wave-action $f_0(p)$ which is used in the inverse problem.

To test the method presented in this article, we solve the forward problem for various values of the degree of homogeneity, γ, choosing $\gamma = \frac{3}{2}, \frac{15}{8}$ and 2. In each case, we run the forward simulation until $t = 10$ with truncation parameter $R = 10$. We set $\Delta t = 0.05$ and $\Delta p_i = \Delta p = 0.1$, i.e. we use a uniform grid. In the tests that follow, we choose $M = 16$ for the Legendre basis functions. From the forward data, we select only the solution after the first time step as our f^k in (4.11) defined above to perform the reconstruction procedure described in (4.15).

Once the forward solution is obtained, we solve (4.9) to approximate the interaction coefficient in each case. To assemble each $\mathbf{Q}^n$ $n = 0, 1, \ldots, N$, the integrals are computed using SciPy's adaptive quadrature method `quad`. The condition number for $\mathbf{Q}^0$ in each test below is approximately 2.4×10^{10}. The regularized least-squares problem (4.13) is solved using scikit-learn's [78] linear model `Ridge` with L-BFGS solver [79] and additionally enforcing the computed coefficients to be positive. In each test below the penalty parameter $\lambda = 0.01$ throughout. Any subsequent matrix multiplication $\mathbf{Q}^n\mathbf{u}$ is computed with the linear models `predict` method.

Before moving on to the next section, we provide a measure of accuracy for the approximated interaction kernel. Using the single set of coefficients $\mathbf{u}$, we show that finite volume data is reproduced by modifying the reconstruction procedure (4.15) to take in the finite

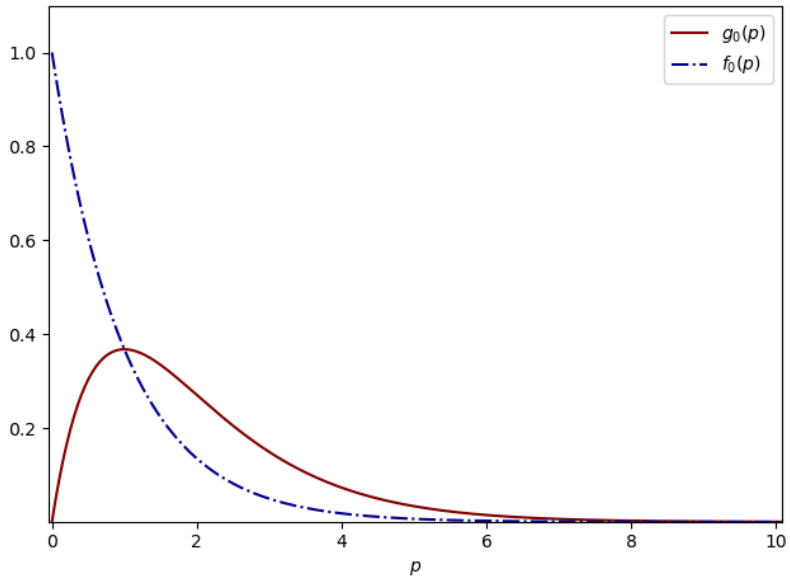


Figure 4.1: Initial condition for the forward problem solved with the finite volume method, $g_0(p)$, and the initial condition for the inverse problem, $f_0(p)$.

volume data $\mathbf{f}$ at each point in time rather than the learned data $\tilde{\mathbf{f}}$. We then compare these results with the finite volume data by computing the L^1 grid norm of the error at each point in time. The results are given below in figure 4.2.

4.3.1. Reconstruction of the Higher Moments

Here the forward solution for the energy density is computed using the finite volume method with $\gamma = 2, \frac{3}{2}$ and $\frac{15}{8}$. We then use the procedure described above to reconstruct this solution using only the initial condition (4.16) and the solution after the first time step, $\mathbf{f}^{0.05}$, provided by the finite volume scheme in each case. That is, we require the initial condition and a single snapshot in time of data. A few snapshots in time are shown for the reconstructed solution in figure 4.3.

The L^∞ grid norm for the difference between the finite volume data and the reconstructed solution given by (4.15) is shown in figure 4.4. It would appear that the error for the reconstructed solution with a single data point is well controlled for about 60 time steps,

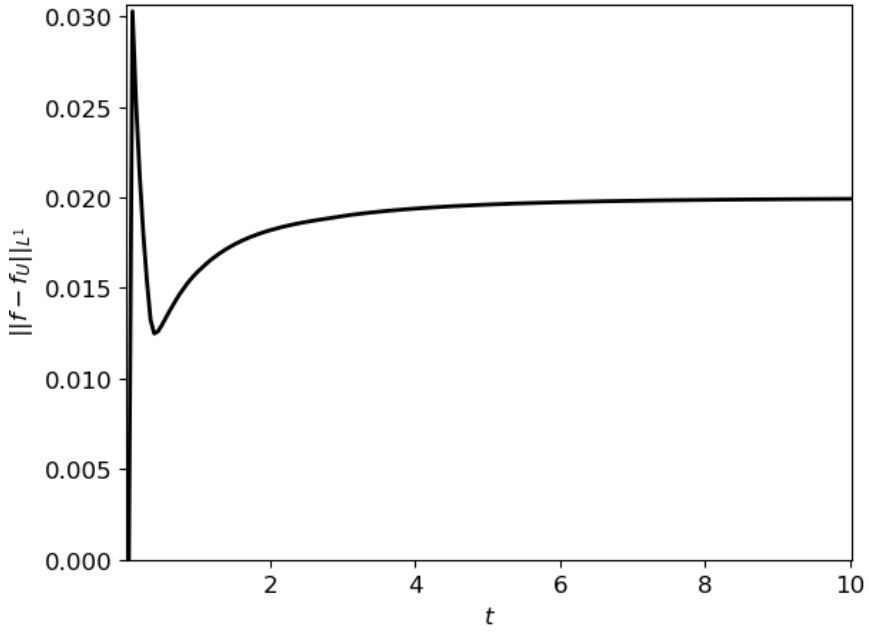


Figure 4.2: We test the accuracy of the approximated interaction kernel by measuring the L^1 grid norm of the error given by measuring the difference $\mathbf{f}^n - \mathbf{d}tQ[\mathbf{f}^{n-1}]\mathbf{u} + \mathbf{f}^{n-1}$ at each point in time, i.e. we check that the approximated interaction kernel can reproduce the data that was used to compute it.

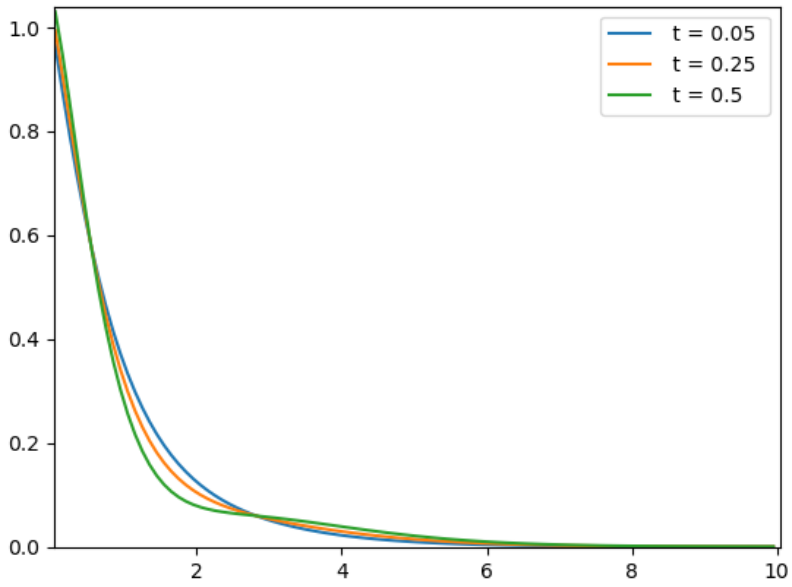


Figure 4.3: Reconstructed solution for the wave-action $f(t, p)$ using the method outlined above.

after which point it begins to grow. The results in figure 4.2 would indicate that adding more measurements from the data would improve this result.

In chapter 2, we observed interesting behavior for the higher moments computed with the finite volume method. Here, we use the computed coefficients for $\gamma_1 = 2, \gamma_2 = \frac{3}{2}, \gamma_3 = \frac{15}{8}$ which will be denoted $\mathbf{u}^{\gamma_1}, \mathbf{u}^{\gamma_2}, \mathbf{u}^{\gamma_3}$ to reconstruct the particle density $\tilde{\mathbf{f}}^{\gamma_i}$, where $i = 1, 2, 3$. The higher moments are then computed using

$$\mathcal{M}^{\ell}_{\gamma_i} = h(\mathbf{p})^{\ell} \cdot \tilde{\mathbf{f}}^{\gamma_i}, \tag{4.17}$$

for $i = 1, 2, 3$ similarly to chapter 2, where the ℓ-th power of the grid data $\mathbf{p}$ is computed element-wise. The results for $\ell = 1, 2, 3, 4$ are given in figure 4.5. It is clear that a longer time simulation is necessary in order to compare with the results in chapter 2, and over a larger wavenumber domain.

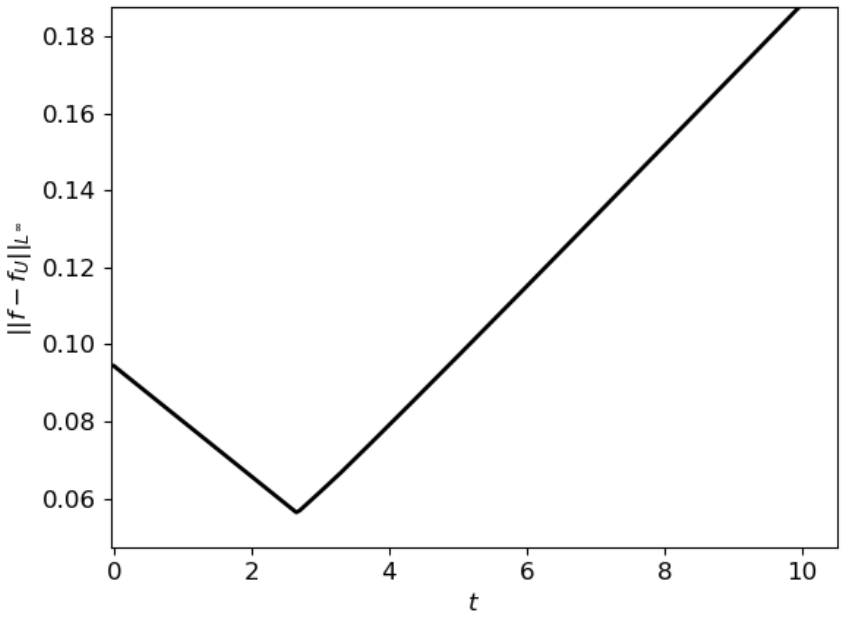


Figure 4.4: The L^∞ grid error is computed to compare the reconstructed solution with the finite volume data at each point in time.

4.4. Conclusion

From the numerical results above in figure 4.2, it would appear that the method is very capable of approximating the unknown interaction kernel with high accuracy. For the reconstruction algorithm, with only an initial condition and a single data vector, There seems to be room for improvement. A possible direction is a higher order time update as is used for the finite volume method. Another possibility is to add additional data vectors at successive points in time. However, we would like to emphasize that using a single inverse solve, we were able to construct the interaction kernel quite accurately, which is probably a more realistic application, since the approximated interaction kernel can be used with any numerical method once obtained. This is still a work in progress.

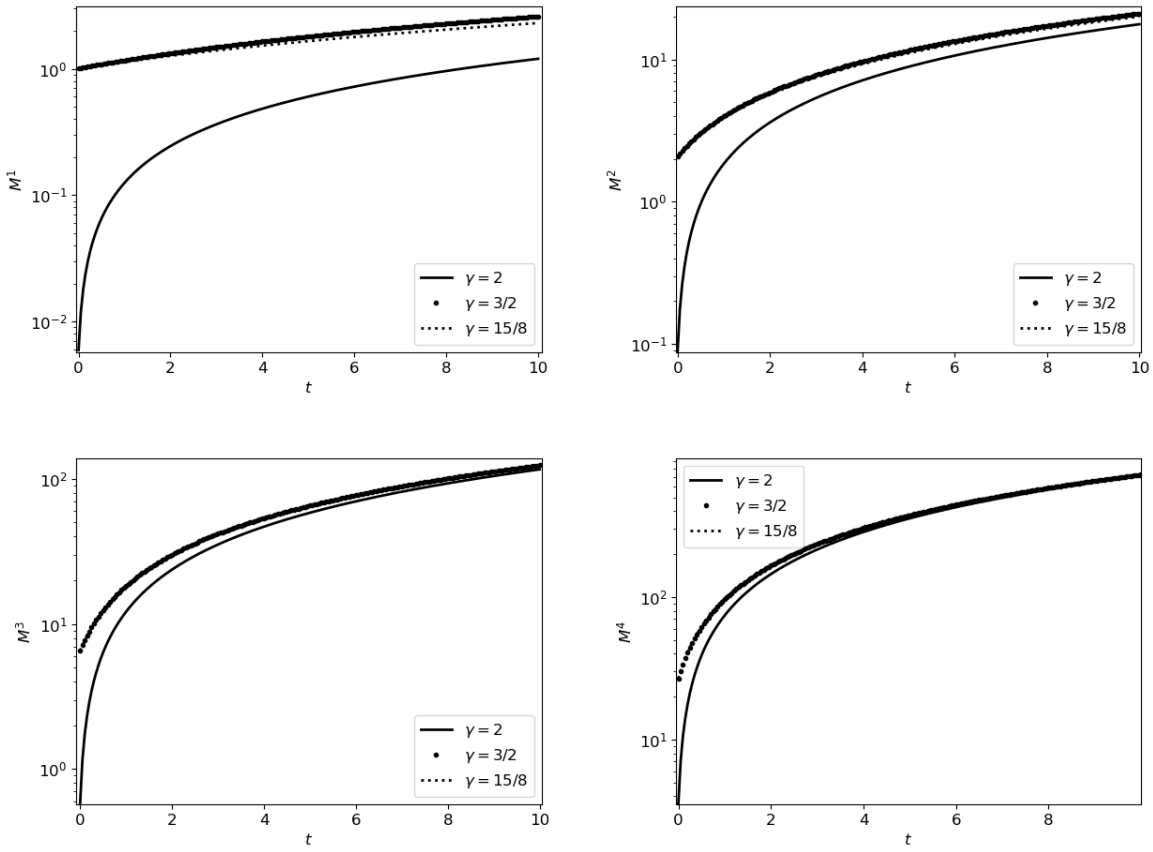


Figure 4.5: Plots (semilogy) of the higher moments, $\mathcal{M}^\ell$ for $\gamma = 2, \frac{3}{2}$ and $\frac{15}{8}$ and $\ell = 1, 2, 3, 4$. There appear to be some similarities with the results in chapter 2, but the short time scale makes a direct comparison difficult.

Chapter 5

A Petrov-Galerkin Residual Distribution Method

Scientific Product: Walton. *A residual distribution approach to 3-wave kinetic equations*, (in preparation).

5.1. Introduction

The present chapter is devoted to deriving a Petrov-Galerkin Residual Distribution (PG RD) [80–83] scheme for equation (2.5) found in chapter 2 (reproduced below). We will see that the PG RD method derived for equation (2.5) generalizes the Finite Volume (FV) scheme derived in 2. In section 5.3, we provide numerical results and compare with the theory. The computations are performed using the MATAR (MAtrix ARray) C++ library [84].

5.2. PG RD Scheme

We will derive the aforementioned PG RD method for the conservation form of the energy density equation (2.5) of chapter 2, which reads

$$\partial_t \frac{g(t,p)}{p} = \partial_p \mathcal{Q}_3[g](t,p), \tag{5.1}$$

where we recall that the collision term is given by

$$\begin{aligned} \partial_p \mathcal{Q}_3[g](t,p) = & -2\partial_p \int_0^p \int_0^p a(p_1,p_2) \frac{g(p_1)}{p_1} \frac{g(p_2)}{p_2} \chi\{p < p_1 + p_2\} \mathrm{d}p_2 \mathrm{d}p_1 \\ & + \partial_p \int_0^\infty \int_0^\infty a(p_1,p_2) \frac{g(p_1)}{p_1} \frac{g(p_2)}{p_2} \chi\{p < p_1 + p_2\} \mathrm{d}p_2 \mathrm{d}p_1. \end{aligned} \tag{5.2}$$

It is easier to begin the derivation for the particle density $f(t,p)$, sometimes called the wave-action, and then solve for the energy density. We define the Petrov-Galerkin Residual at d.o.f. ℓ in the time slab $[t^n, t^k]$ to be

$$R_\ell^{k,E}(f) = \Big(\partial_t f(t,p) - \partial_p \mathcal{Q}_3[f](t,p), \psi_\ell\Big)_{L^2(t^n,t^k;E)} \tag{5.3}$$

with $\psi_\ell = \psi_\ell(p) \in \mathbb{V}_h^0$ a test function, with

$$\mathbb{V}_h^0 = \{\psi \,|\, \psi|_{E_i} \in \mathbb{P}^0(E_i) \subset \{L^2 \cap C_c\}(E_i)\}, \tag{5.4}$$

for all i such that $\cup_i E_i = \Omega \subset \mathbb{R}^+$. The total residual in element E is then defined to be

$$R^{k,E}(f) = \sum_{\ell \in E} R_\ell^{k,E}(f). \tag{5.5}$$

The distribution coefficients are then given by

$$\beta_\ell^{k,E} = \frac{\left(\frac{R_\ell^{k,E}}{R^{k,E}}\right)^+}{\sum_{j\in E}\left(\frac{R_j^{k,E}}{R^{k,E}}\right)^+}, \tag{5.6}$$

where $(\cdot)^+ = \max(0,\cdot)$. The definition of the above distribution coefficients is the one most commonly used in residual distribution schemes, though by no means the only choice. In multidimensional advection dominated problems, the above choice for β imparts a natural upwinding property to the scheme [82]. The simplest case is a steady-state problem, in which the total residual reduces to the flux of some quantity in element E. The sign of the Galerkin residual is then penalized to match that of the flux. The limited residual at ℓ is then

$$\mathcal{R}_\ell^{k,E} = \beta_\ell^{k,E} R^{k,E}, \tag{5.7}$$

and the full update, using the deferred correction method (DeC) [85–88], of the solution is given by

$$f(t^{k+1},p) = f(t^k,p) - \sum_{E\ni\ell} \mathcal{R}_\ell^{k,E}. \tag{5.8}$$

where the notation $\sum_{E\ni\ell}$ denotes the sum of elements, E, containing node ℓ. The details of the DeC scheme are described in subsection 5.2.1 below.

So far, we have only discretized the wavenumber and time domain, and there are still choices to be made for the approximation of $R_\ell^{k,E}(f)$. Returning to (5.3), we have

$$\begin{aligned} R_\ell^{k,E}(f) &= \Big(\partial_t f(t,p) - \partial_p \mathcal{Q}_3[f](t,p), \psi_\ell\Big)_{L^2(t^n,t^k;E)} \\ &= \int_E \{f(t^k,p) - f(t^n,p)\}\psi_\ell \mathrm{d}p - \int_{t^n}^{t^k}\int_E \partial_p \mathcal{Q}_3[f]\psi_\ell \mathrm{d}p\mathrm{d}t \\ &= \int_E \{f(t^k,p) - f(t^n,p)\}\psi_\ell \mathrm{d}p + \int_{t^n}^{t^k}\int_E \partial_p\psi_\ell \mathcal{Q}_3[f]\mathrm{d}p\mathrm{d}t - \int_{t^n}^{t^k}\Big[\mathcal{Q}_3[f]\psi_\ell\Big]_E^{\pm}\mathrm{d}t. \end{aligned} \tag{5.9}$$

Now, let us replace $f(t,p)$ by its basis expansion, $\Phi_{t^k} f = \varphi_\nu f_\nu^k$, where the repeated index implies summation with $\nu \in \{0,1,\ldots,N\}$ and $\varphi \in \mathbb{V}_h^N$, with

$$\mathbb{V}_h^N = \{\varphi \,|\, \varphi|_{E_i} \in \mathbb{P}^N(E_i) \subset L^1 \cap C^\infty(\Omega)\}. \tag{5.10}$$

Actually, we make a specific choice of polynomial functions, the Bernstein polynomials. This choice is motivated by their total positivity. This feature will aid us in the proofs to come and help ensure that the solutions are positive as demanded by the physics.

We may then write, using the notation $\delta^k f = f^k - f^n$,

$$R_\ell^{k,E}(\Phi_{t^k} f) = \int_E \varphi_\nu\psi_\ell \mathrm{d}p\, \delta^k f_\nu + \int_{t^n}^{t^k}\int_E \partial_p\psi_\ell \mathcal{Q}_3[\Phi_{t^k} f]\mathrm{d}p\mathrm{d}t - \int_{t^n}^{t^k}\Big[\mathcal{Q}_3[\Phi_{t^k} f]\psi_\ell\Big]_E^{\pm}\mathrm{d}t, \tag{5.11}$$

choosing our test function to be $\psi_\ell = \frac{1}{\Delta p_\ell}$ and Δp_ℓ as defined in the previous section, we get

$$R_\ell^{k,E}(\Phi_{t^k} f) = \frac{1}{\Delta p_\ell} \int_E \varphi_\nu \mathrm{d}p \, \delta^k f_\nu - \frac{1}{\Delta p_\ell} \int_{t^n}^{t^k} \Big[\mathcal{Q}_3[\Phi_{t^k} f] \Big]_E^{\pm} \mathrm{d}t. \tag{5.12}$$

Lastly, we perform integration over the wavenumber domain by Gaussian quadrature and a transformation to the reference element E_{ref}. The time integral is also performed with an exact Gaussian quadrature upon interpolation of the integrand in time. Specifically, we have

$$R_\ell^{k,E}(\Phi_{t^k} f) = \frac{1}{\Delta p_\ell} \sum_{j=0}^{2N} \varphi_\nu(\xi_j) \omega_j \det(J) \, \delta^k f_\nu - \frac{1}{\Delta p_\ell} \sum_{j=0}^{M-1} \mathcal{I}^M \Big[\mathcal{Q}_3[\Phi_{t^{j,k}} f] \Big]_E^{\pm} \theta_j, \tag{5.13}$$

where ω_m and θ_j and the quadrature weights for the wavenumber and time domain, respectively, and $\mathcal{I}^M$ denotes the time interpolator of order M. Before describing the discretization of the collision operator, let us prove that the scheme given by (5.8) using (5.13) is indeed a generalization of the FV scheme found in [63].

5.2.0.1. Derivation of the FV scheme from the PG RD method

Let us return to equation (5.12) which we repeat here for convenience,

$$R_\ell^{k,E}(\Phi_{t^k} f) = \frac{1}{\Delta p_\ell} \delta^k f_\nu \int_E \varphi_\nu \mathrm{d}p - \frac{1}{\Delta p_\ell} \int_{t^n}^{t^k} \Big[\mathcal{Q}[\Phi_{t^k} f] \Big]_E^{\pm} \mathrm{d}t. \tag{5.14}$$

Now, using the trapezoid rule to approximate the time integral, we could define the PG residual to be

$$R_\ell^{k,E}(\Phi_{t^k} f) = \frac{1}{\Delta p_\ell} \delta^k f_\nu \int_E \varphi_\nu \mathrm{d}p - \frac{\Delta t}{2 \Delta p_\ell} \Big\{ \Big[\mathcal{Q}_3[\Phi_{t^k} f] \Big]_E^{\pm} + \Big[\mathcal{Q}_3[\Phi_{t^n} f] \Big]_E^{\pm} \Big\}. \tag{5.15}$$

Choosing $N = 0$ results in a single point interior to the element in the wavenumber domain so that $\Delta p_\ell = \Delta p = \mu(E)$, $\varphi_\nu = 1$ for all ν and consequently

$$R_\ell^{k,E}(\Phi_{t^k} f) = R_\ell^{k,E}(f^k) = \delta^k f_\ell - \frac{\Delta t}{2\Delta p}\Big\{\Big[\mathcal{Q}_3[f^k]\Big]_E^{\pm} + \Big[\mathcal{Q}_3[f^n]\Big]_E^{\pm}\Big\}. \tag{5.16}$$

A further consequence of $N = 0$ for this choice of ψ is that the distribution coefficients will all be identically one and thus

$$\mathcal{R}_\ell^{k,E} = R^{k,E} = R_\ell^k. \tag{5.17}$$

Applying this to equation (5.8) we get

$$\begin{aligned} f_\ell^{k+1} = f_\ell^k - R_\ell^k &= f_\ell^k - \Big(\delta^k f_\ell - \frac{\Delta t}{2\Delta p}\Big\{\Big[\mathcal{Q}_3[f^k]\Big]^{\pm} + \Big[\mathcal{Q}_3[f^n]\Big]^{\pm}\Big\}\Big) \\ &= f_\ell^n + \frac{\Delta t}{2\Delta p}\Big\{\Big[\mathcal{Q}_3[f^k]\Big]^{\pm} + \Big[\mathcal{Q}_3[f^n]\Big]^{\pm}\Big\}. \end{aligned} \tag{5.18}$$

Lastly, selecting $k = n$, i.e. setting the number of correction steps to one, we arrive at

$$f_\ell^{n+1} = f_\ell^n + \frac{\Delta t}{\Delta p}\Big[\mathcal{Q}_3[f^n]\Big]^{\pm}, \tag{5.19}$$

which gives the finite volume scheme we wished to derive.

5.2.1. Deferred Correction Methods

In this section we give details of the DeC time-stepping method. To begin, we must define the operators $\mathcal{L}^1$ and $\mathcal{L}^2$, which we will call the prediction operator and correction

operator, respectively. Then, we define

$$\mathcal{L}^1(f^k) = \begin{pmatrix} \bar{\varphi}_\ell(f_\ell^{m,k} - f_\ell^{n,0}) + \frac{\Delta t_{m,n}}{\Delta p_\ell}\Big[\mathcal{Q}_3[f^{n,0}]\Big]_E^{\pm} \\ \vdots \\ \bar{\varphi}_\ell(f_\ell^{1,k} - f_\ell^{n,0}) + \frac{\Delta t_{1,n}}{\Delta p_\ell}\Big[\mathcal{Q}_3[f^{n,0}]\Big]_E^{\pm} \end{pmatrix}, \tag{5.20}$$

to be the m-th prediction step, where $\Delta t_{m,n} = \frac{m}{M}\Delta t$, with M the number of sub-time steps. We define

$$\bar{\varphi}_\ell = \frac{1}{\Delta p_\ell}\int_E \varphi_\ell(x)\mathrm{d}x,$$

the element average of the basis function which is positive thanks to the properties of the Bernstein basis. The choice of M plays a direct role in the stability of the method, and should be at least that of the order of the spatial approximation [82, 87, 88]. This is essentially a crude forward Euler approximation [82, 86, 87]. Next, define the correction operator

$$\mathcal{L}^2(f^k) = \begin{pmatrix} \sum_{E\ni\ell}\frac{1}{\Delta p_\ell}\Big\{\int_E \varphi_\nu \mathrm{d}x(f_\nu^{m,k} - f_\nu^{n,0}) + \int_{t^n}^{t^m}\mathcal{I}^M\Big[\mathcal{Q}_3[f^{n,k}]\Big]_E^{\pm}\mathrm{d}t\Big\} \\ \vdots \\ \sum_{E\ni\ell}\frac{1}{\Delta p_\ell}\Big\{\int_E \varphi_\nu \mathrm{d}x(f_\nu^{1,k} - f_\nu^{n,0}) + \int_{t^n}^{t^1}\mathcal{I}^M\Big[\mathcal{Q}_3[f^{n,k}]\Big]_E^{\pm}\mathrm{d}t\Big\} \end{pmatrix}, \tag{5.21}$$

which is an M-th order operator for $\mathcal{L}^2(f^*) = 0$. The full update is then given by

$$\mathcal{L}^1(f^{k+1}) = \mathcal{L}^1(f^k) - \mathcal{L}^2(f^k) \tag{5.22}$$

Moving both $\mathcal{L}^1$ operators to the left-hand side results in equation (5.8). From [85–87], we have the following theorem.

Theorem 5.1. *Assume that there exists an f^* such that $\mathcal{L}^2(f^*) = 0$. Further assume that $\mathcal{L}^1$ is coercive, with constant of coercivity $C_1 > 0$, and that the difference $\mathcal{L}^1 - \mathcal{L}^2$ is Lipschitz*

continuous with Lipschitz constant, $C_2 > 0$, depending on Δt, i.e.

$$C_1\|u - v\| \leq \|\mathcal{L}^1(u) - \mathcal{L}^1(v)\| \tag{5.23}$$

and

$$\|(\mathcal{L}^1(u) - \mathcal{L}^2(u)) - (\mathcal{L}^1(v) - \mathcal{L}^2(v))\| \leq C_2(\Delta t)\|u - v\|. \tag{5.24}$$

Then, if $\frac{C_2(\Delta t)}{C_1} < 1$ the DeC update (5.22) is convergent, and after k corrections, the error is smaller than

$$\Big(\frac{C_2(\Delta t)}{C_1}\Big)^k \|f^0 - f^*\|.$$

Proof. For completeness, we reproduce the proof found in [86]. From the coercivity of $\mathcal{L}^1$ we have

$$\begin{aligned} \|f^k - f^*\| &\leq \frac{1}{C_1}\|\mathcal{L}^1(y^k) - \mathcal{L}^1(y^*)\| \\ &= \|\mathcal{L}^1(f^{k-1}) - \mathcal{L}^2(f^{k-1}) - \mathcal{L}^1(f^*) + \mathcal{L}^2(f^*)\| \\ &\leq \Delta t \frac{C_2}{C_1}\|f^{k-1} - f^*\|, \end{aligned} \tag{5.25}$$

from the Lipschitz property of $\mathcal{L}^1 - \mathcal{L}^2$ and the assumption that $\mathcal{L}^2(f^*) = 0$. The result follows by induction. □

5.2.2. Discretization of the Flux Collision Term

The goal of this section is to derive computationally efficient approximations to the term $\Big[\mathcal{Q}_3[\Phi_{t^k} f]\Big]_E^{\pm}$ in the definition of the residual, equation 5.12, above. Let us first add some detail,

$$\Big[\mathcal{Q}_3[\Phi_{t^k} f]\Big]_E^{\pm} = \mathcal{Q}_3^{i+\frac{1}{2}}[\Phi_{t^k} f] - \mathcal{Q}_3^{i-\frac{1}{2}}[\Phi_{t^k} f], \tag{5.26}$$

where the nodes $i + \frac{1}{2}$ and $i - \frac{1}{2}$ denote the right and left endpoints of the element E, respectively. Recalling the definition of $\mathcal{Q}_3[f](t, p)$, the 3-WKE collision operator and writing

$V_{1,2} = (p_1 p_2)^{\gamma/2}$, we may then write, ignoring the discretization of f for the moment,

$$\mathcal{Q}_3^{i+\frac{1}{2}}[f] = -2\int_0^{p_{i+\frac{1}{2}}}\int_0^{p_{i+\frac{1}{2}}} (p_1p_2)^{\gamma/2} f_1 f_2 \chi\{p_{i+\frac{1}{2}} < p_1 + p_2\}\mathrm{d}p_{21} + \int_0^P\int_0^P (p_1p_2)^{\gamma/2} f_1 f_2 \chi\{p_{i+\frac{1}{2}} < p_1 + p_2\}\mathrm{d}p_{21},$$

and likewise

$$\mathcal{Q}_3^{i-\frac{1}{2}}[f] = -2\int_0^{p_{i-\frac{1}{2}}}\int_0^{p_{i-\frac{1}{2}}} (p_1p_2)^{\gamma/2} f_1 f_2 \chi\{p_{i-\frac{1}{2}} < p_1 + p_2\}\mathrm{d}p_{21} + \int_0^P\int_0^P (p_1p_2)^{\gamma/2} f_1 f_2 \chi\{p_{i-\frac{1}{2}} < p_1 + p_2\}\mathrm{d}p_{21},$$

with $P >> 1$ the truncation parameter. Before further discretization, it is convenient to leave the above two terms as they are and simplify their difference. This simplification not only aids in the analysis to follow but is necessary for a computationally efficient implementation. More specifically, while it is natural to apply a quadrature rule to the above semi-discrete flux collision term as is, this will result in slow computations without the aid of GPUs are parallelism. See [89] for a comparison of a serial implementation of the finite volume scheme (5.19) in Python against a C++ implementation using MATAR.

In what follows, to ease notation, we will only explicitly write the set characteristic functions in the integrands and the term $(p_1p_2)^{\gamma/2} f_1 f_2$ will be implied throughout. We can group the terms in the difference via their bounds of integration so that we have

$$\begin{aligned}\Big[\mathcal{Q}_3[f]\Big]_E^{\pm} &= \int_0^P\int_0^P \Big(\chi\{p_{i+\frac{1}{2}} < p_1 + p_2\} - \chi\{p_{i-\frac{1}{2}} < p_1 + p_2\}\Big)\mathrm{d}p_{21} \\ &\quad + 2\int_0^{p_{i-\frac{1}{2}}}\int_0^{p_{i-\frac{1}{2}}} \chi\{p_{i-\frac{1}{2}} < p_1 + p_2\}\mathrm{d}p_{21} \\ &\quad - 2\int_0^{p_{i+\frac{1}{2}}}\int_0^{p_{i+\frac{1}{2}}} (p_1p_2)^{\gamma/2} f_1 f_2 \chi\{p_{i+\frac{1}{2}} < p_1 + p_2\}\mathrm{d}p_{21} \\ &= (\mathrm{I}) + (\mathrm{II}) + (\mathrm{III}).\end{aligned}$$

We will first analyse (III), which can be written as

$$\begin{aligned}(\text{III}) = -2\Big\{ &\int_0^{p_{i-\frac{1}{2}}} \int_0^{p_{i-\frac{1}{2}}} \chi\{p_{i+\frac{1}{2}} < p_1 + p_2\}\mathrm{d}p_{21} \\ &+ 2\int_0^{p_{i-\frac{1}{2}}} \int_{p_{i-\frac{1}{2}}}^{p_{i+\frac{1}{2}}} \chi\{p_{i+\frac{1}{2}} < p_1 + p_2\}\mathrm{d}p_{21} \\ &+ \int_{p_{i-1/2}}^{p_{i+\frac{1}{2}}} \int_{p_i-\frac{1}{2}}^{p_{i+\frac{1}{2}}} \chi\{p_{i+\frac{1}{2}} < p_1 + p_2\}\mathrm{d}p_{21}\Big\}.\end{aligned} \tag{5.27}$$

Combining this with (II), we get

$$\begin{aligned}(\text{II}) + (\text{III}) = 2\int_0^{p_{i-\frac{1}{2}}} \int_0^{p_{i-\frac{1}{2}}} \Big(\chi\{p_{i-\frac{1}{2}} < p_1 + p_2\}\mathrm{d}p_{21} - \chi\{p_{i+\frac{1}{2}} < p_1 + p_2\}\Big)\mathrm{d}p_{21} \\ -4\int_0^{p_{i-\frac{1}{2}}} \int_{p_{i-\frac{1}{2}}}^{p_{i+\frac{1}{2}}} \chi\{p_{i+\frac{1}{2}} < p_1 + p_2\}\mathrm{d}p_{21} - 2\int_{p_{i-\frac{1}{2}}}^{p_{i+\frac{1}{2}}} \int_{p_{i-\frac{1}{2}}}^{p_{i+\frac{1}{2}}} \chi\{p_{i+\frac{1}{2}} < p_1 + p_2\}\mathrm{d}p_{21}.\end{aligned} \tag{5.28}$$

Taking into account the set characteristic functions, we obtain a further simplification of (I)+(II)+(III) and may write $\Big[\mathcal{Q}_3[f]\Big]_E^{\pm}$ as

$$\begin{aligned}\Big[\mathcal{Q}_3[f]\Big]_E^{\pm} = \int_0^{p_{i-\frac{1}{2}}} \Bigg\{ &\int_{p_{i-\frac{1}{2}}-p_1}^{p_{i+\frac{1}{2}}-p_1} (p_1p_2)^{\frac{\gamma}{2}} f_1 f_2 \mathrm{d}p_2 + 2\int_{p_{i-\frac{1}{2}}-p_1}^{p_{i-\frac{1}{2}}} (p_1p_2)^{\frac{\gamma}{2}} f_1 f_2 \mathrm{d}p_2 \\ -4\int_{p_{i-\frac{1}{2}}\bigvee p_{i+\frac{1}{2}}-p_1}^{p_{i+\frac{1}{2}}} &(p_1p_2)^{\frac{\gamma}{2}} f_1 f_2 \mathrm{d}p_2\Bigg\}\mathrm{d}p_1 - 2\int_{p_{i-\frac{1}{2}}}^{p_{i+\frac{1}{2}}} \int_{p_{i-\frac{1}{2}}\bigvee p_{i+\frac{1}{2}}-p_1}^{p_{i+\frac{1}{2}}} (p_1p_2)^{\frac{\gamma}{2}} f_1 f_2 \mathrm{d}p_2\mathrm{d}p_1,\end{aligned} \tag{5.29}$$

where we use $x \bigvee y$ to indicate taking the maximum of x and y. It is this form which we will fully discretize and approximate with a quadrature rule in the Petrov-Galerkin Residual

$$R_\ell^{k,E}(\Phi_{t^k} f) = \frac{1}{\Delta p_\ell}\sum_{j=0}^{2N} \varphi_\nu(\xi_j)\omega_j \det(J)\, \delta^k f_\nu - \frac{1}{\Delta p_\ell}\sum_{j=0}^{M-1} \mathcal{I}^M\Big[\mathcal{Q}_3[\Phi_{t^{j,k}} f]\Big]_E^{\pm}\theta_j, \tag{5.30}$$

resulting in the update

$$f(t^{k+1}, p) = f(t^k, p) - \sum_{E\ni\ell} \mathcal{R}_\ell^{k,E}, \tag{5.31}$$

already given above.

5.3. Numerical Results

Next, we study the evolution problem with initial condition given by

$$g_0(p) = pe^{-p}, \quad p \in [0, R], \tag{5.32}$$

shown in figure 5.1 with various orders N of the basis functions φ with the number of elements fixed at $2R$, where the figure is shown for $R = 100$, with the x-axis given in log-scale.

To verify numerically the derivation of the FVS from the PGRD method, we take initial condition (5.32), set $\Delta t = 0.05$ and compute solutions the WKE using the finite volume method of chapter 2 and the update given in equation (5.8) choosing zeroth order polynomials and using the trapezoidal rule for the time integration inside the residual. The results shown in figure 5.2 match exactly as predicted.

5.4. Conclusions

Development of the higher-order code leveraging MATAR is currently underway. The results of the zeroth order method are promising. We further plan to treat the issue of convergence the scheme analytically in this forthcoming work. As the PG-RD method generalizes the method of chapter 2, the convergence analysis of the PG-RD method will apply to the finite volume method as well. Similarly, we are constructing a rigorous error analysis of the method, which will also be applicable to the finite volume method. With the higher-order expansion polynomials, we plan to explore the use of non-uniform grids. Further, the use of MATAR allows for easy parallelization, which when enable computations over longer time intervals and larger truncation parameters.

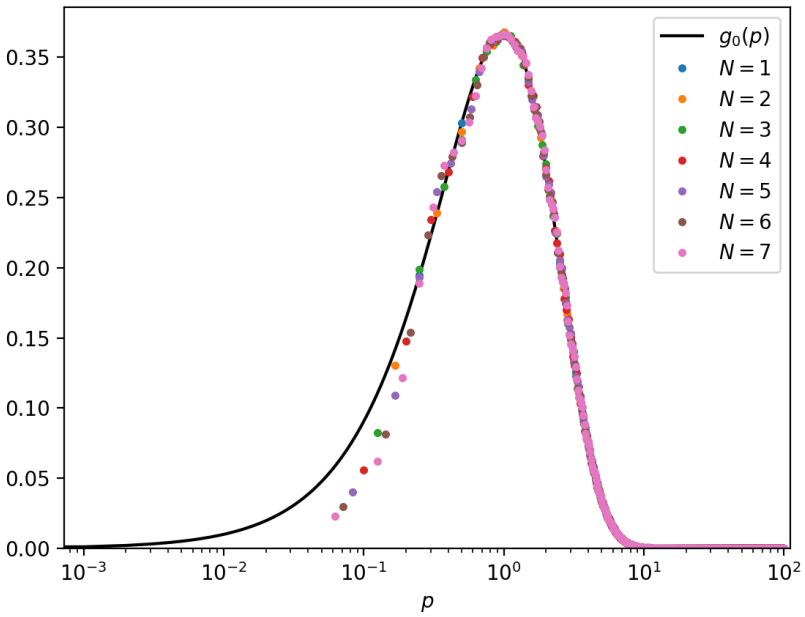


Figure 5.1: The initial condition (5.32) for various orders of interpolation N. The x-axis is given in log-scale for clarity. Here $R = 100$ and the number of elements is $2R$.

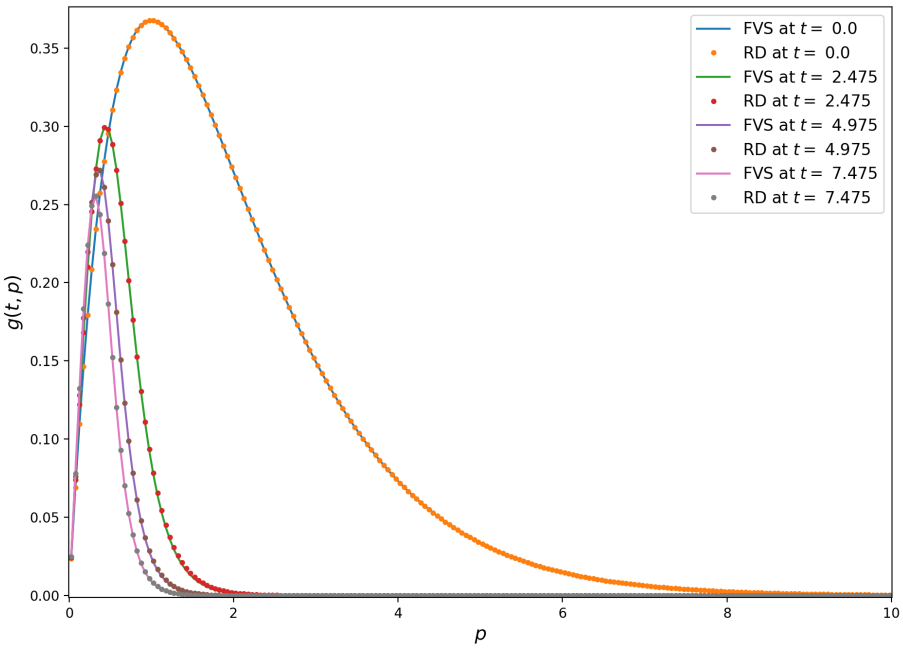


Figure 5.2: Comparison of the finite volume solution with the residual distribution using zeroth order expansion functions and the trapezoidal rule for time integration. The number of elements/control volumes is $10R$ in the figure. The point $p = 0$ is not resolved on the grid as the data corresponds to the cell averages.

Chapter 6

Conclusions

In this final chapter, we would like to provide a discussion on research directions not addressed in the preceding chapters. Some of the topics we address here have been touched on to some extent in the previous chapters, but perhaps it worth collecting them all in one place. Then, the current chapter will be organized as follows. We will address open questions or possible research directions in the context of the work presented in chapter sequentially.

In chapter 2, we are provided with several results which should be studied analytically. The first we will mention is the solution dynamics as the phenomena occurs. We observe that regardless of the initial regularity, after some short time t_s, the remaining energy in the computational domain is concentrated at some small wavenumber value k_s, and for any time $t > t_s$ the energy concentrated at the point decays in a self-similar fashion. For some initial conditions, we observe that in addition to the self-similar decay about k_s, other low amplitude pulses are generated a higher wavenumbers in the computational domain which then also decay self-similarly after their initial generation. Can this behavior be verified analytically?

Can an analytic explanation for the behavior of higher moments of the solutions as γ is varied be given?

The next question is concerned with the so-called transient cascade phenomena. If forcing is present at $k = 0$ and a sink is at $k = \infty$, for finite capacity systems, it has been observed numerically that the constant flux solutions possess exponents $a > \kappa > 0$ where $f_{trans} \approx C_{trans} k^{-a}$ denotes the transient spectra and κ is the K-Z spectra exponent. This transient

flux is carried by a "front" until it reaches $k = \infty$ where a backward propagation of energy then "lifts" f_{trans} to f_∞, the K-Z solution, closing the "energy loop" from $k = 0$ to $k = \infty$) see [21, 23] and the references therein).

In our work, without forcing or dissipation, we found potential evidence for such a transient spectra before the first blow-up time. These transient spectra are different from the KZ spectra as discussed above and cannot be estimated from dimensional arguments or via applying Zakharov transformations in the 3-WKE. Briefly, the transient spectra should occur just before the singular behavior time t_1^*, and the solution should have the form $f_{trans}(t, k) \approx C_{trans} k^{-a}$ for $a > 0$, where, importantly, $a \neq \kappa$ for κ the exponent of the KZ solution. In Figure 6.1, we give a log-log plot of the initial condition and the solution just before the cascade process begins. We draw a comparative line through this snapshot in time of the solution. Lines corresponding to the KZ spectra for capillary and acoustic waves are also shown for comparison. To rigorously compute a, one must solve a nonlinear eigenvalue problem as in [32]. However, as discussed above, solving such a nonlinear eigenvalue problem is a difficult task: in [32], a hypothesis is imposed on the evolution of the solution. The solution is assumed to grow linearly in time and additional hypotheses are also imposed to treat the singularities of the integral. While these hypotheses remain interesting mathematical questions to be verified, we assume the nonlinear interactions to be solely responsible for the transfer of energy in our work. We provide Figure 6.1 as preliminary evidence that the solution seems to give a cascade behavior consistent with the predicted transient spectra, though further rigorous analysis is required for verification. Finding self-similar profiles for the solutions before the n-th blow-up times, similarly to those done in [32, 90, 91], is non-trivial.

In chapter 3, we presented a deep learning algorithm with non-local non-linear terms in the residual of the loss function. However, the solutions were only presented for the radially symmetric case where we can perform computations on $\mathbb{R}^+$. One of the biggest advantages

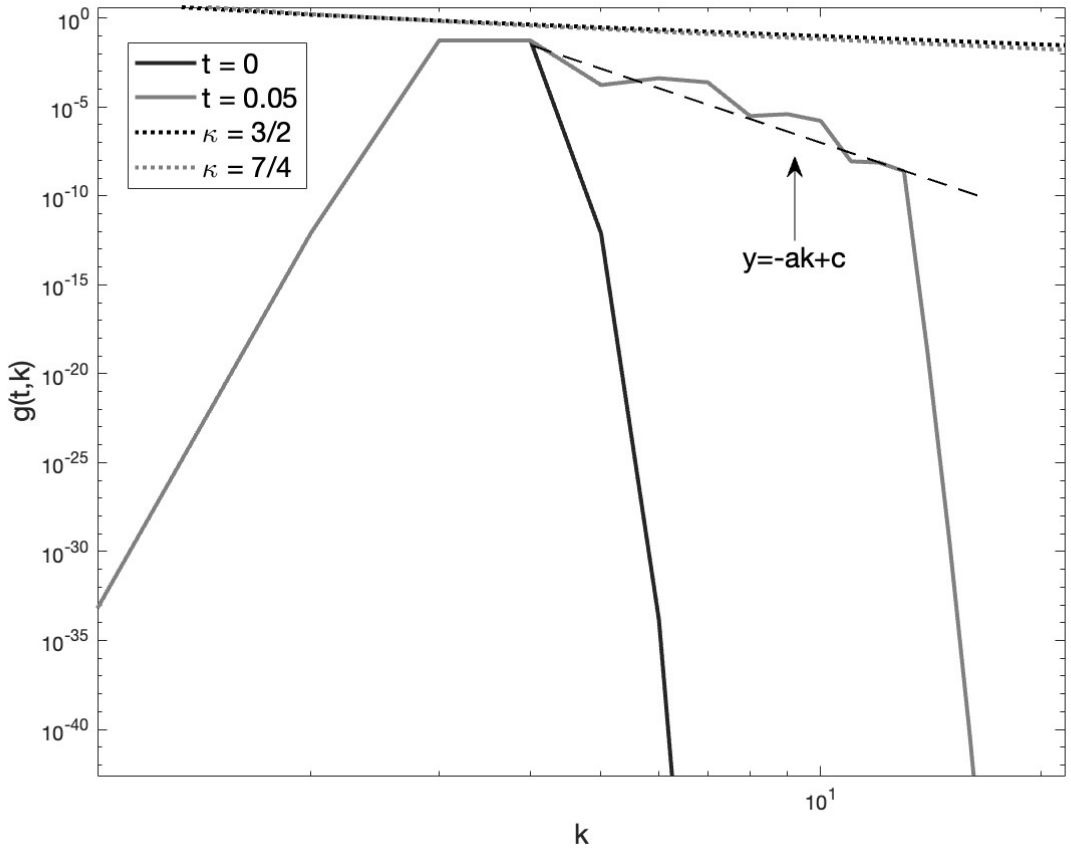


Figure 6.1: Potential evidence for the transient cascade phenomena. The K-Z spectra for acoustic and capillary waves are provided for reference. The above indeed shows a flux with larger exponenet than either of the K-Z exponents.

provided by a deep learning approach to differential equations is the ability to compute solutions for very high dimensional problems. Here there are two challenges. The first is computational. If the computations are not confined to $\mathbb{R}^+$, the one must contend with integration over the resonant manifolds for each wavenumber value of the solution

$$\delta(\omega - \omega_1 - \omega_2)\delta(k - k_1 - k_2), \tag{6.1}$$

for each k in of the solution $f(t, k)$ (see equation (1.2) and [9] for many examples). Then, integration in the collision operators must be restricted to these energy surfaces (resonant manifolds). In some cases, where the energy surface may be cast as an algebraic variety, this can be accomplished via methods developed by computational algebraic geometrician [92] (how efficiently is another question). The second issue is that currently, to the author's knowledge, there is no analysis for the long-time behavior of solutions which are not radially symmetric. This a much more difficult issue.

Chapter 7

Appendix: Background on Methods

For completeness, we give a short account of the methods employed in chapter 3. The literature on deep learning methods is vast and we do not aim to be exhaustive. We will review results on Physics-Informed Neural Networks (PINNs), the Deep Galerkin Method (DGM) and a few applications to fluid mechanics as these are a some of the most relevant results in scientific machine learning to the results in Chapter 3. Each of the previously mentioned machine learning methods are founded upon established universal approximation theorems for neural networks (see for example Hornik, Stithcomba and White, 1989).

7.1. Deep Learning for Partial Differential Equations

The main ideas of the current section can be found in [58] [57], and [59], for example, though we do not follow them exactly. A very nice book chapter which surveys some of the main idea of PINNs can be found in the recent second edition of the book [93] by Brunton and Kutz. Our present goal is to give a brief overview of how one might go about solving partial differential equations (PDEs) with deep neural networks (DNNs) and then to provide a few examples on well known PDEs which, unlike in Chapter 3, only contain local differential operators. While there exist a plethora of architectures one can use to solve PDEs with DNNs, we show that one can get quite far with very simple feed-forward architectures.

For simplicity, we consider equations which are only first order in time. This is not a requirement, but would increase the number of terms in the loss functional below. First, define a partial differential operator, $\mathcal{R}$, the residual operator, to be

$$\mathcal{R} = \partial_t + \mathcal{N}, \tag{7.1}$$

where we do not specify whether $\mathcal{N}$ is linear or nonlinear. We will denote by $\mathcal{B}$ a boundary operator. Let $f(t,x) : \mathbb{R}^+ \times \Omega \to \mathbb{R}$ be the solution to the PDE

$$\begin{aligned} \mathcal{R}f(t,x) &= 0 \quad && (t,x) \in \mathbb{R}^+ \times \Omega \\ \mathcal{B}f(t,x) &= g(t,x) && (t,x) \in \mathbb{R}^+ \times \partial\Omega \\ f(0,x) &= f_0(x) && x \in \Omega, \end{aligned} \tag{7.2}$$

with $\Omega \subseteq \mathbb{R}^d$, $d \geq 1$. Next, we introduce a feed-forward neural network, $n(t,x;\theta)$ for $\theta \in \Theta$ the set of weights and biases of the neural network. Specifically, the neural network is a composition of nonlinear activation functions, σ, with inputs $(t,x) = \mathbf{x} \in \mathbb{R}^{1+d}$. The first layer of the neural network will be denoted $L^1 := \mathbf{W}^1\mathbf{x} + \mathbf{b}^1$ and the ℓth layer $L^\ell = \mathbf{W}^\ell\sigma(L^{\ell-1}) + \mathbf{b}^\ell$, so that we arrive at the following representation for our approximating neural network

$$n(t,x;\theta) = id(L^N) \circ \sigma(L^{N-1}) \circ \cdots \circ \sigma(L^2) \circ \sigma(L^1). \tag{7.3}$$

Note that the activation functions, σ, from composition to composition, are not necessarily the same and that the activation function in the last layer is simply the identity so that $L^N = \mathbf{W}^N\sigma(L^{N-1}) + b^N \in \mathbb{R}$ for our purposes.

Next, we define a loss functional, $J_\theta[n](t,x)$, which we will minimize with respect to the parameter θ to find an approximate solution to (7.2) above. In general, the procedure which we will describe in the sequel produces a solution which is not unique and most likely a global minimum. Let us proceed with the definition of the loss functional. We have

$$J_\theta[n](t,x) = \|\mathcal{R}n(t,x;\theta)\|^2_{L^2(\mathbb{R}^+\times\Omega)} + \|\mathcal{B}n(t,x;\theta) - g(t,x)\|^2_{L^2(\mathbb{R}^+\times\partial\Omega)} + \|f_0(x) - n(0,x;\theta)\|^2_{L^2(\Omega)}, \tag{7.4}$$

and the idea is to find

$$\theta^* = \arg\min_\theta J_\theta[n](t,x), \tag{7.5}$$

via some optimization procedure. There exist a plethora of options to consider when choosing an optimization procedure, but often one uses stochastic gradient descent and specifically the ADAM algorithm [94], which is the approach we will use below. Looking at equation (7.4), we see that we still have choices to make regarding our space-time discretization and just how we intend to approximate the operators $\mathcal{R}$ and $\mathcal{B}$, let alone the L^2 inner products. To this end, let us truncate the time domain by some simulation parameter $T > 0$ so that the integrals above are over $(0, T) \subset \mathbb{R}^+$.

In what follows, we shall use a quasi-Monte Carlo approach to approximate the inner products, which also determines the discretization of $(0, T) \times \Omega$, namely to approximate space-time with some low-discrepancy sequence [65]. The low discrepancy sequence we use is the Sobol sequence [60], which is illustrated above in Chapter 3. The choice of quasi-Monte Carlo is motivated by faster convergence rates and the obviation of the construction of a measure over the samples [59], since it is deterministic. The author is aware of at least two papers [61, 62] which provide comparisons of a quasi-Monte Carlo approach with classical Monte Carlo, as well as a comparison between the various low-discrepancy sequences, in the context of fluid dynamics.

With this in mind, we shall write out the fully discrete loss function, which, with some abuse of notation, we simply denote as $J_\theta[n](t, x)$. Let $\mathcal{S} \subset (0, T) \times \Omega$ denote the set of Sobol space-time samples, $\partial\mathcal{S} \subset (0, T) \times \partial\Omega$ the set of Sobol samples for the boundary condition and $\mathcal{X} \subset \{0\} \times \Omega$ the samples for the initial condition. Then the fully discrete loss function, where the operators below should be understood to be their discrete analogues, is

$$J_\theta[n](t, x) = \frac{1}{|\mathcal{S}|} \sum_{(t,x)\sim\mathcal{S}} \Big(\mathcal{R}n(t, x; \theta)\Big)^2 + \frac{1}{|\partial\mathcal{S}|} \sum_{(t,x)\sim\partial\mathcal{S}} \Big(\mathcal{B}n(t, x; \theta) - g(t, x)\Big)^2 + \frac{1}{|\mathcal{X}|} \sum_{x\sim\mathcal{X}} \Big(f_0(x) - n(0, x; \theta)\Big)^2. \tag{7.6}$$

In what follows, we will consider two examples, a heat equation and an inviscid Burgers' equation. For clarity, and to demonstrate the simplicity of the code, we will provide example

code written in Tensorflow with Keras [67, 68] for the heat equation. We can recommend libraries such as the DeepXDE library [58] and the more recent Modulus toolkit by NVIDIA [95] which one can use to create models rather than writing the code from scratch. To adapt the code for the Burgers' equation, one needs only to modify the residual function and ensure the boundary and initial conditions have been altered correctly. Otherwise, the code is the same.

First, we consider the following heat equation

$$\begin{aligned} (\partial_t - \partial_{xx}) f(t,x) &= 0 \quad (t,x) \in (0,T] \times (0,\pi) \\ f_0(x) &= \sin(x) \qquad\qquad x \in (0,\pi) \\ f(t,0) &= f(t,\pi) = 0, \end{aligned} \tag{7.7}$$

where we take $T = 2$. The exact solution is $f(t,x) = e^{-t}\sin(x)$, as can be seen from inspection. We define our neural network to be

$$n(t,x;\theta) = id(L^3) \circ \sigma(L^2) \circ \sigma(L^1), \tag{7.8}$$

where we select $\sigma = \frac{1}{1+e^{-x}}$, the sigmoid function. Each hidden layer has 128 units [96]. To solve the above heat equation with a neural network, we follow the framework presented above. We take $|\mathcal{S}| = |\partial\mathcal{S}| = 1024$ and $|\mathcal{X}| = 64$. The residual for the heat equation is defined below in the code listing.

```
def heat_res(model, x):
    time = x[:,0]
    space = x[:,1]
    with tf.GradientTape(persistent=True) as tape:
        tape.watch(time)
        tape.watch(space)
```

```
        g = model( tf.stack( [time, space], 1) )
        g_x = tape.gradient(g, space)

    g_t = tape.gradient( g, time )
    g_xx = tape.gradient( g_x, space)
    return tf.subtract( g_t, g_xx)
```

The loss function is defined exactly as above in equation (7.6) and the code for its implementation is provided below.

```
def loss_fcn(model,x):
    # PDE residual
    L1 = tf.reduce_mean( tf.square(  heat_res(model,x)  ) )

    # adherance to initial condition
    t0 = tf.zeros(shape = (2**6, 1))
    x_0 = sobol_transformation(
          tf.math.sobol_sample( dim = 1, num_results = 2**6),
          0 ,L )
    x_ic = tf.concat( [t0, x_0], 1)
    L2 = tf.reduce_mean(tf.square(tf.subtract(heat_IC(x_0),model(x_ic))))

    # boundary conditions
    x_naught = tf.zeros(shape = (loss_samples, ))
    X_bc1 = tf.stack( [ x[:,0], x_naught], 1)
    L3 = tf.reduce_mean( tf.square(  model(X_bc1)  ) )

    x_pi = np.pi*tf.ones(shape = (loss_samples, ))
```

```
    X_bc2 = tf.stack( [x[:,0], x_pi], 1)
    L4 = tf.reduce_mean( tf.square( model(X_bc2) ) )

    return tf.add(L1, tf.add( L2, tf.add( L3, L4 ) ) )
```

The function `sobol_transformation` is a user defined function which tranforms Sobol samples in a unit hypercube to a hypercube with user specified lengths. Lastly, to perform stochastic gradient descent on the loss function to obtain a local minimum we would use the following function.

```
def train(model,x):

    with tf.GradientTape(persistent=True) as tape:
        tape.watch(model.variables)
        loss = loss_fcn(model,x)

    grad = tape.gradient(loss,model.variables)
    optimizer.apply_gradients(zip(grad,model.variables))

    return loss
```

where the optimizer can be set by `optimizer = keras.optimizers.Adam()`. See the online documentation for possible input arguments. The training time depends on the tolerance one sets for the loss function. The results in figure 7.1 were obtained with less than half a minute of training.

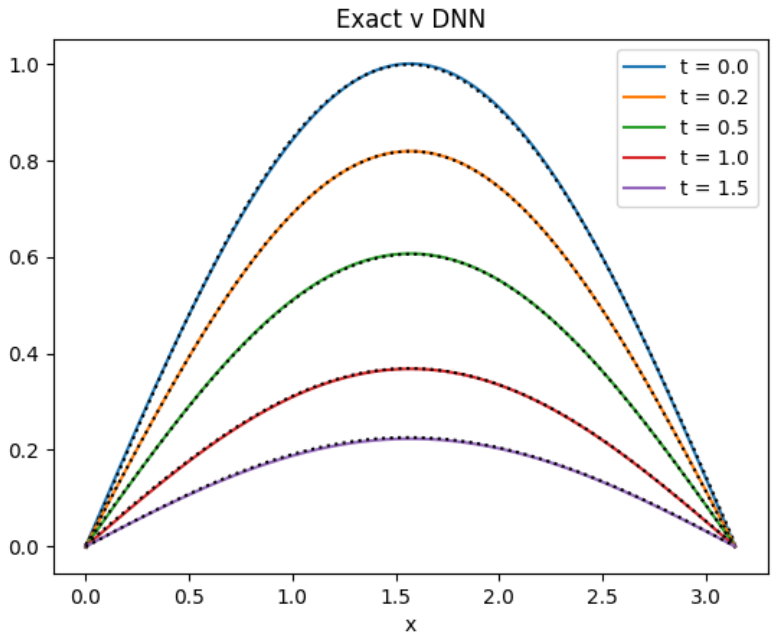


Figure 7.1: Comparison of DNN solution (black dots) to the heat equation against the exact solution (colored lines).

We next consider the inviscid Burgers' equation

$$\begin{aligned} (\partial_t + f(t,x)\partial_x) f(t,x) = 0 \quad & (t,x) \in (0,T] \times (0,2\pi) \\ f_0(x) = 1 - \cos(x) \quad & x \in (0,2\pi) \\ f(t,0) = f(t,2\pi) = 0, & \end{aligned} \tag{7.9}$$

where $T = 1.1$ just after the shock time $t_{shock} = 1$. Here, we take $|\partial\mathcal{S}| = 8192$ and, as before, $|\mathcal{X}| = 64$. For this problem, the architecture of the neural network is modified so that now $\sigma(x) = \tanh(x)$ and we have 8 hidden layers each with 20 units. This architecture increased training speed, though we note that the simpler architecture used for the heat equation also works. We again use ADAM to minimize the loss functional. To better capture the shock, we actually use two space-time samples, $\mathcal{S}^1 \subset (0, 0.8] \times (0, 2\pi)$ and $\mathcal{S}^2 \subset [0.9, T] \times (0, 2\pi)$ where each sample set has size 8192, which adds an extra residual term to the loss function.

Specifically, we have

$$J_\theta[n](t,x) = \frac{1}{|\mathcal{S}^1|}\sum_{(t,x)\sim\mathcal{S}^1}\Big(\mathcal{R}n(t,x;\theta)\Big)^2 + \frac{1}{|\mathcal{S}^2|}\sum_{(t,x)\sim\mathcal{S}^2}\Big(\mathcal{R}n(t,x;\theta)\Big)^2 + \frac{1}{|\partial\mathcal{S}|}\sum_{(t,x)\sim\partial\mathcal{S}}\Big(\mathcal{B}n(t,x;\theta)\Big)^2 + \frac{1}{|\mathcal{X}|}\sum_{x\sim\mathcal{X}}\Big(f_0(x) - n(0,x;\theta)\Big)^2. \tag{7.10}$$

For comparison, we solve the Burgers' equation by applying Fourier transforms, solving the spectral ODE and transforming back to physical space (see [93] for more details). We can see in figure 7.2 that the solutions are in agreement but that after the shock develops only the neural network remains stable.

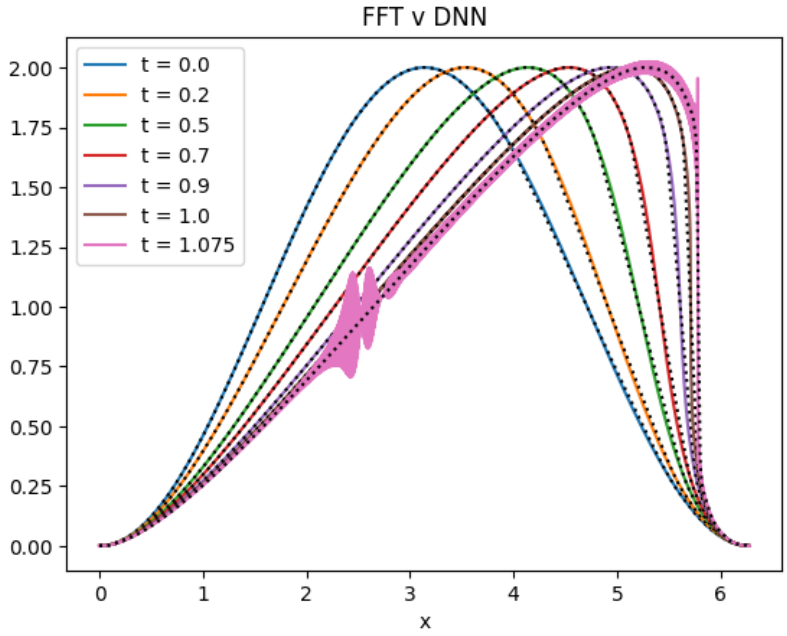


Figure 7.2: Solution of the inviscid Burgers' equation by FFT method (color) and DNN (black dots). The shock time is $t = 1$, notice the breakdown of the FFT solution (pink oscillating line) against the DNN solutions.

BIBLIOGRAPHY

[1] R. Peierls, *Zur kinetischen Theorie der Wärmeleitung in Kristallen*, *Annalen der Physik* **395** (1929) 1055–1101. 1

[2] D. J. Benney and P. G. Saffman, *Nonlinear interactions of random waves in a dispersive medium*, *Proc. R. Soc. Lond. A* **289** (1966) 301–320. 1

[3] K. Hasselmann, *On the non-linear energy transfer in a gravity-wave spectrum part 1. general theory*, *Journal of Fluid Mechanics* **12** (1962) 481–500. 1

[4] K. Hasselmann, *On the non-linear energy transfer in a gravity wave spectrum part 2. conservation theorems; wave-particle analogy; irrevesibility*, *Journal of Fluid Mechanics* **15** (1963) 273–281. 1

[5] K. Hasselmann, *On the non-linear energy transfer in a gravity-wave spectrum. part 3. evaluation of the energy flux and swell-sea interaction for a neumann spectrum*, *Journal of Fluid Mechanics* **15** (1963) 385–398. 1

[6] K. Hasselmann, *Feynman diagrams and interaction rules of wave-wave scattering processes*, *Reviews of Geophysics* **4** (1966) 1–32, [https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/RG004i001p00001]. 1

[7] V. E. Zakharov and N. N. Filonenko, *Weak turbulence of capillary waves*, *Journal of applied mechanics and technical physics* **8** (1967) 37–40. 1, 2, 4

[8] A. C. Newell and B. Rumpf, *Wave turbulence*, *Annual Review of Fluid Mechanics* **43** (2011) 59–78, [https://doi.org/10.1146/annurev-fluid-122109-160807]. 1

[9] S. Nazarenko, *Wave turbulence*, vol. 825 of *Lecture Notes in Physics*. Springer, Heidelberg, 2011. 1, 4, 59, 86

[10] S. Galtier, *Physics of Wave Turbulence*. Cambridge University Press, 2022, 10.1017/9781009275880. 1

[11] A. N. Pushkarev and V. E. Zakharov, *Turbulence of capillary waves*, *Physical review letters* **76** (1996) 3320. 1

[12] A. N. Pushkarev and V. E. Zakharov, *Turbulence of capillary waves: theory and numerical simulation*, *Physica D: Nonlinear Phenomena* **135** (2000) 98–116. 1

[13] V. E. Zakharov, *Stability of periodic waves of finite amplitude on the surface of a deep fluid*, *Journal of Applied Mechanics and Technical Physics* **9** (1968) 190–194. 1, 2

[14] V. E. Zakharov and S. V. Nazarenko, *Dynamics of the Bose-Einstein condensation*, *Phys. D* **201** (2005) 203–211. 1

[15] K. Hasselmann, *On the spectral dissipation of ocean waves due to white capping*, *Boundary-Layer Meteorology* **6** (1974) 107–127. 2

[16] Y. Pomeau and M.-B. Tran, *Statistical physics of non equilibrium quantum phenomena*, *Lecture Notes in Physics, Springer* (2019) . 2

[17] V. E. Zakharov, *Weak turbulence in media with a decay spectrum*, *Journal of Applied Mechanics and Technical Physics* **6** (1965) 22–24. 2

[18] C. Connaughton, *Numerical solutions of the isotropic 3-wave kinetic equation*, *Physica D: Nonlinear Phenomena* **238** (2009) 2282–2297. 2, 5, 61

[19] A. O. Korotkevich, A. I. Dyachenko and V. E. Zakharov, *Numerical simulation of surface waves instability on a homogeneous grid*, *Phys. D* **321/322** (2016) 51–66. 4

[20] V. E. Zakharov, V. S. L'vov and G. Falkovich, *Kolmogorov spectra of turbulence I: Wave turbulence*. Springer Science & Business Media, 2012. 4

[21] L. Biven, S. Nazarenko and A. Newell, *Breakdown of wave turbulence and the onset of intermittency*, *Physics Letters A* **280** (2001) 28–32. 4, 5, 84

[22] A. C. Newell, S. Nazarenko and L. Biven, *Wave turbulence and intermittency*, *Physica D: Nonlinear Phenomena* **152-153** (2001) 520–550. 4

[23] V. Shrira and S. Nazarenko, *Advances in Wave Turbulence*. WORLD SCIENTIFIC, 2013, 10.1142/8269. 4, 7, 84

[24] E. Falcon and N. Mordant, *Experiments in surface gravity–capillary wave turbulence*, *Annual Review of Fluid Mechanics* **54** (2022) 1–25, [https://doi.org/10.1146/annurev-fluid-021021-102043]. 4

[25] E. Kochurin, G. Ricard, N. Zubarev and E. Falcon, *Three-dimensional direct numerical simulation of free-surface magnetohydrodynamic wave turbulence*, *Phys. Rev. E* **105** (Jun, 2022) L063101. 4

[26] S. Galtier, S. V. Nazarenko, A. C. Newell and A. Pouquet, *A weak turbulence theory for incompressible magnetohydrodynamics*, *Journal of Plasma Physics* **63** (2000) 447–488. 4

[27] S. Galtier and S. V. Nazarenko, *Direct evidence of a dual cascade in gravitational wave turbulence*, *Phys. Rev. Lett.* **127** (Sep, 2021) 131101. 4

[28] Y. V. Lvov, K. L. Polzin, E. G. Tabak and N. Yokoyama, *Oceanic internal-wave field: theory of scale-invariant spectra*, *Journal of Physical Oceanography* **40** (2010) 2605–2623. 4

[29] R. Micha and I. I. Tkachev, *Turbulent thermalization*, *Physical Review D* **70** (2004) 043538. 4

[30] A. Sheremet, Y. Qin, J. P. Kennedy, Y. Zhou and A. P. Maurer, *Wave turbulence and energy cascade in the hippocampus*, *Frontiers in Systems Neuroscience* **12** (2019) . 4

[31] C. Connaughton and P. Krapivsky, *Aggregation–fragmentation processes and decaying three-wave turbulence*, *Physical Review E* **81** (2010) 035303. 5, 39, 40, 61

[32] C. Connaughton and A. C. Newell, *Dynamical scaling and the finite-capacity anomaly in three-wave turbulence*, *Physical Review E* **81** (2010) 036303. 5, 61, 84

[33] T. T. Nguyen and M.-B. Tran, *On the kinetic equation in Zakharov's wave turbulence theory for capillary waves*, *arXiv preprint arXiv:1702.03892* (2017) . 5

[34] I. M. Gamba, L. M. Smith and M.-B. Tran, *On the wave turbulence theory for stratified flows in the ocean*, *M3AS: Mathematical Models and Methods in Applied Sciences. Vol. 30, No. 1 105-137* (2020) . 5

[35] R. Alonso, I. M. Gamba and M.-B. Tran, *The cauchy problem and BEC stability for the quantum Boltzmann-Condensation system for bosons at very low temperature*, 2018. 5

[36] P. Germain, A. D. Ionescu and M.-B. Tran, *Optimal local well-posedness theory for the kinetic wave equation*, *Journal of Functional Analysis* **279** (2020) 108570. 5

[37] M. Escobedo and M.-B. Tran, *Convergence to equilibrium of a linearized quantum boltzmann equation for bosons at very low temperature*, *Kinetic and Related Models* **8** (2015) 493–531. 5

[38] A. V. Bobylev and I. M. Gamba, *Boltzmann equations for mixtures of Maxwell gases: exact solutions and power like tails*, *Journal of Statistical Physics* **124** (2006) 497–516. 5

[39] I. Gamba, V. Panferov and C. Villani, *On the Boltzmann equation for diffusively excited granular media*, *Communications in Mathematical Physics* **246** (2004) 503–541. 5

[40] C. Josserand and Y. Pomeau, *Nonlinear aspects of the theory of Bose-Einstein condensates*, *Nonlinearity* **14** (jun, 2001) R25. 5

[41] L. E. Reichl and M.-B. Tran, *A kinetic equation for ultra-low temperature Bose-Einstein condensates*, *Journal of Physics A: Mathematical and Theoretical* **52** (2019) 063001. 5

[42] M.-B. Tran, G. Craciun, L. M. Smith and S. Boldyrev, *A reaction network approach to the theory of acoustic wave turbulence*, *Journal of Differential Equations* **269** (2020) 4332–4352. 5

[43] A. Soffer and M.-B. Tran, *On the energy cascade of 3-wave kinetic equations: beyond Kolmogorov-Zakharov solutions*, *Communications in Mathematical Physics* (2019) 1–48. 5, 6, 7, 8, 9, 10, 13, 21, 22, 23, 25, 27, 29, 31, 34, 35, 36, 39, 40, 43, 45, 52, 56, 57, 58, 61

[44] A. N. Kaufman, *Wave entropy: A derivation by Jaynes' principle*, *The Physics of Fluids* **29** (1986) 2326–2326, [https://aip.scitation.org/doi/pdf/10.1063/1.865570]. 7

[45] B. Rumpf. private communication. 8

[46] F. Filbet and P. Laurençot, *Numerical simulation of the Smoluchowski coagulation equation*, *SIAM Journal on Scientific Computing* **25** (2004) 2004–2028. 10, 11, 13, 34, 40, 41, 43

[47] E. L. Foster, J. Lohéac and M.-B. Tran, *A structure preserving scheme for the Kolmogorov–Fokker–Planck equation*, *Journal of Computational Physics* **330** (2017) 319–339. 10

[48] T. A. Bak and O. Heilmann, *A finite version of Smoluchowski's coagulation equation*, *Journal of Physics A: Mathematical and General* **24** (1991) 4889. 11

[49] F. P. Da Costa, *A finite-dimensional dynamical model for gelation in coagulation processes*, *Journal of nonlinear science* **8** (1998) 619–653. 11

[50] D. D. Keck and D. M. Bortz, *Numerical simulation of solutions and moments of the Smoluchowski coagulation equation*, *arXiv preprint arXiv:1312.7240* (2013) . 12

[51] G. Laibe and M. Lombart, *On the Courant-Friedrichs-Lewy condition for numerical solvers of the coagulation equation*, *arXiv preprint arXiv:2109.11065* (2021) . 22

[52] X. Y. Hu, N. A. Adams and C.-W. Shu, *Positivity-preserving method for high-order conservative schemes solving compressible Euler equations*, *Journal of Computational Physics* **242** (2013) 169–180. 22, 57

[53] J. Huang and C.-W. Shu, *Positivity-preserving time discretizations for production–destruction equations with applications to non-equilibrium flows*, *Journal of Scientific Computing* **78** (2019) 1811–1839. 22, 57

[54] J. Huang, W. Zhao and C.-W. Shu, *A third-order unconditionally positivity-preserving scheme for production–destruction equations with applications to non-equilibrium flows*, *Journal of Scientific Computing* **79** (2019) 1015–1056. 22, 57

[55] R. J. LeVeque, *Finite Difference Methods for Ordinary and Partial Differential Equations*. Society for Industrial and Applied Mathematics, 2007, 10.1137/1.9780898717839. 25

[56] S. Nazarenko, *Wave turbulence*, vol. 825 of *Lecture Notes in Physics*. Springer, Heidelberg, 2011, 10.1007/978-3-642-15942-8. 36

[57] M. Raissi, P. Perdikaris and G. Karniadakis, *Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*, *Journal of Computational Physics* **378** (2019) 686–707. 37, 39, 87

[58] L. Lu, X. Meng, Z. Mao and G. E. Karniadakis, *DeepXDE: A deep learning library for solving differential equations*, *SIAM Review* **63** (2021) 208–228, [https://doi.org/10.1137/19M1274067]. 37, 39, 87, 90

[59] J. Sirignano and K. Spiliopoulos, *DGM: A deep learning algorithm for solving partial differential equations*, *Journal of Computational Physics* **375** (2018) 1339–1364. 39, 87, 89

[60] I. Sobol', *On the distribution of points in a cube and the approximate evaluation of integrals*, *USSR Computational Mathematics and Mathematical Physics* **7** (1967) 86–112. 39, 89

[61] S. Mishra and T. K. Rusch, *Enhancing accuracy of deep learning algorithms by training with low-discrepancy sequences*, *SIAM Journal on Numerical Analysis* **59** (2021) 1811–1834, [https://doi.org/10.1137/20M1344883]. 39, 42, 89

[62] K. O. Lye, S. Mishra and D. Ray, *Deep learning observables in computational fluid dynamics*, *Journal of Computational Physics* **410** (2020) 109339. 39, 42, 89

[63] S. Walton and T. M.-B., *A numerical scheme for wave turbulence: 3-wave kinetic equations*, *SIAM Journal on Scientific Computing (to appear)* (2023) . 40, 43, 46, 52, 57, 58, 61, 65, 74

[64] A. Bensoussan, Y. Li, D. P. C. Nguyen, M.-B. Tran, S. C. P. Yam and X. Zhou, *Machine learning and control theory*, in *Handbook of Numerical Analysis*, pp. 531–558. Elsevier BV, 2022. 42

[65] R. E. Caflisch, *Monte Carlo and quasi-Monte Carlo methods*, *Acta Numerica vol. 7, pp. 1-49.* (1998) . 42, 89

[66] A. Bensoussan, *Stochastic control of partially observable systems*. Cambridge University Press, 1992. 44

[67] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro et al., *TensorFlow: Large-scale machine learning on heterogeneous systems*, 2015. 46, 90

[68] F. Chollet et al., "Keras." https://keras.io, 2015. 46, 90

[69] A. Sergeev and M. D. Balso, *Horovod: fast and easy distributed deep learning in TensorFlow*, *arXiv preprint arXiv:1802.05799* (2018) . 46

[70] S. M. Hornik, K. and H. White, *Multilayer feedforward networks are universal approximators*, *Neural Networks, 2(5), 359–366.* (1989) . 46

[71] L. Liu, B. Wang and W. Cai, *Linearized learning methods with multiscale deep neural networks for stationary Navier-Stokes equations with oscillatory solutions*, 2021. 10.48550/ARXIV.2102.03293. 56

[72] V. Zakharov, *Statistical theory of gravity and capillary waves on the surface of a finite-depth fluid*, *European Journal of Mechanics - B/Fluids* **18** (1999) 327–344. 59

[73] S. Walton, M.-B. Tran and A. Bensoussan, *A deep learning approximation of non-stationary solutions to wave kinetic equations*, *Applied Numerical Mathematics* (2022) . 61

[74] P. P. Jones, R. C. Ball and C. Connaughton, *Nonlinear least-squares method for the inverse droplet coagulation problem.*, *Physical Review.* **88(1)** (2013) . 61

[75] C. Connaughton and P. P. Jones, *Some remarks on the inverse Smoluchowski problem for cluster-cluster aggregation*, *Journal of Physics: Conference Series* **333** (dec, 2011) 012005. 61

[76] J. Chakraborty, J. Kumar, M. Singh, A. Mahoney and D. Ramkrishna, *Inverse problems in population balances. determination of aggregation kernel by weighted residuals*, *Industrial & Engineering Chemistry Research* **54** (2015) 10530–10538, [https://doi.org/10.1021/acs.iecr.5b01368]. 61, 63

[77] A. Vadeboncoeur, D. Akyildiz, I. Kazlauskaite, M. Girolami and F. Cirak, *Deep probabilistic models for forward and inverse problems in parametric PDEs*, 2022. 10.48550/ARXIV.2208.04856. 62

[78] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel et al., *Scikit-learn: Machine learning in Python*, *Journal of Machine Learning Research* **12** (2011) 2825–2830. 65

[79] D. C. Liu and J. Nocedal, *Limited memory BFGS method for large scale optimization*, *Mathematical programming* **45** (1989) 503–528. 65

[80] N. M. R. Abgrall, K. Lipnikov and S. Tokareva, *Multidimensional staggered grid residual distribution scheme for Lagrangian hydrodynamics*, *Journal of Scientific Computing* (2020) . 71

[81] R. Abgrall and S. Tokareva, *Staggered grid residual distribution scheme for Lagrangian hydrodynamics*, *Journal of Scientific Computing* (2017) . 71

[82] R. Abgrall and P. Roe, *High order fluctuation schemes on triangular meshes*, *Journal of Scientific Computing* (2003) . 71, 72, 76

[83] M. Ricchiuto and A. Bollermann, *Stabilized residual distribution for shallow water simulations*, *Journal of Computational Physics* (2009) . 71

[84] J. M. E. N. T. T. R. R. D. Dunning, N. Morgan, *MATAR: A performance portability and productivity implementation of data-oriented design with Kokkos*, *Journal of Parallel and Distributed Computing* (2021) . 71

[85] R. Abgrall, P. Bacigaluppi and S. Tokareva, *High-order residual distribution scheme for the time-dependent Euler equations of fluid dynamics*, *Computers Mathematics with Applications* **78** (2019) 274–297. 73, 76

[86] R. Abgrall, *High order schemes for hyperbolic problems using globally continuous approximation and avoiding mass matrices*, *Journal of Scientific Computing* **73** (2017) . 73, 76, 77

[87] R. Abgrall, E. L. Mélédo, P. Öffner and D. Torlo, *Relaxation deferred correction methods and their applications to residual distribution schemes*, *The SMAI Journal of computational mathematics* **8** (2022) 125–160. 73, 76

[88] B. Ong and R. Spiteri, *Deferred correction methods for ordinary differential equations*, *Journal of Scientific Computing* (2020) . 73, 76

[89] A. D. E. H. C. R. R. M. E. L. S. W. J. M. B. R. D. H. Nathaniel Morgan, Caleb Yenusah and M. Knezevic, *On a simplified approach to achieve parallel performance and portability across CPU and GPU architectures (submitted)*, . 78

[90] N. K. Bell, V. N. Grebenev, S. B. Medvedev and S. V. Nazarenko, *Self-similar evolution of Alfvèn wave turbulence*, *Journal of Physics A: Mathematical and Theoretical* **50** (2017) 435501. 84

[91] B. V. Semisalov, V. N. Grebenev, S. B. Medvedev and S. V. Nazarenko, *Numerical analysis of a self-similar turbulent flow in Bose-Einstein condensates*, *Communications in Nonlinear Science and Numerical Simulation* **102** (2021) 105903. 84

[92] P. Breiding and O. Marigliano, *Random points on an algebraic manifold*, *SIAM Journal on Mathematics of Data Science* **2** (2020) 683–704, [https://doi.org/10.1137/19M1271178]. 86

[93] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, 2 ed., 2022, 10.1017/9781009089517. 87, 94

[94] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, *arXiv preprint arXiv:1412.6980* (2014) . 89

[95] "NVIDIA Modulus Toolkit." https://developer.nvidia.com/modulus. 90

[96] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*. MIT Press, 2016. 90