



**HAL**  
open science

# Models and numerical methods for multiscale transport problems

Laurent Navoret

► **To cite this version:**

Laurent Navoret. Models and numerical methods for multiscale transport problems. Mathematics [math]. Université de Strasbourg, 2024. tel-04826632

**HAL Id: tel-04826632**

**<https://theses.hal.science/tel-04826632v1>**

Submitted on 9 Dec 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Habilitation à diriger des recherches**

**INSTITUT DE  
RECHERCHE  
MATHÉMATIQUE  
AVANCÉE**

**UMR 7501**

**Strasbourg**

Université de Strasbourg  
Spécialité MATHÉMATIQUES APPLIQUÉES

**Laurent Navoret**

**Models and numerical methods  
for multiscale transport problems**

Soutenu le 20 décembre 2024  
devant la commission d'examen

Vincent Calvez, examinateur  
Pierre Degond, examinateur  
Philippe Helluy, garant  
Pauline Lafitte, examinatrice  
Raphaël Loubère, examinateur  
Lorenzo Pareschi, rapporteur  
Nicolas Seguin, rapporteur  
Éric Sonnendrücker, rapporteur

<https://irma.math.unistra.fr>



**Université**

**de Strasbourg**



---

## Remerciements

Les quatorze dernières années qui se sont écoulées depuis ma soutenance de thèse ont été véritablement passionnantes, scientifiquement et humainement. C'est comme une belle balade, parfois avec des détours sinueux et des forêts bien sombres, mais que c'est beau quand les paysages se découvrent ! Et, surtout, que de belles rencontres en chemin ! Ce manuscrit d'Habilitation à Diriger les Recherches s'en veut le reflet et il est temps de remercier toutes les personnes qui ont contribué de près ou de loin.

En cette toute fin d'année 2024, la balade est ponctuée par la tant attendue soutenance. Je tiens donc à remercier très chaleureusement tous les membres de mon jury et en premier lieu Lorenzo Pareschi, Nicolas Seguin et Éric Sonnendrücker, qui ont rapporté ce manuscrit. C'est un très grand honneur pour moi. J'exprime également toute ma gratitude envers Pauline Lafitte d'avoir accepté de présider ce jury. Merci également très sincèrement à vous : Vincent Calvez, Pierre Degond et Raphaël Loubère. Votre présence me réjouit. Enfin, je tiens à remercier tout particulièrement Philippe Helluy de m'avoir accompagné scientifiquement depuis mon arrivée à Strasbourg. Merci pour ta bienveillance, merci pour ta relaxation à toute épreuve !

Les contributions présentées dans ce manuscrit sont le fruit de collaborations, qui me sont toutes très chères. Je suis extrêmement reconnaissant envers tous les collègues pour leur gentillesse et leur passion communicative. Que c'est agréable de travailler avec vous ! Je remercie donc (par ordre à peu près chronologique) : Bertrand M., Bérénice G., Nicolas M., Philippe H. (encore !), Anaïs C., Pierre D., Piotr M., Ewelina Z., Matthieu B., Cédric F., Michel M., Mehdi B., Ali E., Emmanuel F., David C., Florence D., François B., Charlotte P., Nilasis C., Thomas B., Vincent V., Clémentine C., Yannick P., Kilian L., Raphaël C., Victor M.-D. (<3), Emmanuel O, Léopold T., Youssouf N., et tous ceux que j'ai rencontrés au cours des différents projets Cemracs ! J'en profite pour remercier mes co-organisateurs du Cemracs 2022, le CIRM et tous les participants, pour ce très bel été. Je remercie également de manière plus spécifique Marcela S., Daniel R., Simon L., Romain L., He L. pour la belle aventure scientifique à la frontière entre biologie, physique et mathématique : que c'est enrichissant de discuter avec vous ! Merci également à mes aventuriers de l'apprentissage : merci Vincent V., Emmanuel F. et Victor M.-D. pour le chemin parcouru ensemble. Durant ces années, j'ai eu la joie de collaborer avec plusieurs doctorants : Nhung Pham, Pierre Gerhard, Romane Hélie, Léo Bois, Mickaël Bestard et actuellement Guillaume Steimer, Claire Schnoebelen, Roxana Sublet, Nicolas Pailliez. Quel plaisir d'interagir et de grandir avec vous ! Ce manuscrit vous doit énormément.

Je souhaiterais également remercier chaleureusement tous les collègues de l'IRMA. C'est vraiment une chance d'être aussi bien entouré. Commençons par le voisinage rapproché : Vincent, mon délicieux collègue de bureau, qui m'a fait découvrir tant de choses, dont le jazz, le handball et les réseaux de neurones ! Je continue avec les membres du deuxième étage de l'UFR que je croise régulièrement, avec un remerciement spécial pour Jean B. et Pierre B. et les discussions sur les politiques françaises et américaines. Je continue par remercier tous les membres et ex-membres de l'équipe Moco et de l'équipe INRIA : Michaël G., Pierre N., Christophe S., Sever H., Vilmos K., Christophe P., Bopeng R., Joubine A., Rémi I., Michel D., Antoine D., Andréa T., Hung T.. Quel dynamisme, quel tonus, quels gourmands ! Je pense bien sûr également

à tous les autres membres du laboratoire avec qui j'ai pu interagir : Nicolas J., Xiaolin Z., Jacques F., Rutger N., Christine H., Ana R., Yann B., Frédéric C., Thomas D., Charles F., Laurent G., Christophe Po., Étienne B. Je ne vous citerai malheureusement pas tous tant vous êtes nombreux. Je suis très enfin heureux de partager avec vous, Brieuc F. et Thomas S., cette journée de soutenances !

Depuis quatre ans déjà, je suis embarqué avec Nalini A. puis Raphaël C. et Laurène P. dans l'aventure ITI (Institut Thématique Interdisciplinaire). Je suis très heureux de mener avec vous ce projet, dans une atmosphère toujours chaleureuse. Merci ! Je tiens plus particulièrement à remercier Laurène pour son travail dévoué et admirable d'efficacité : c'est précieux. Je remercie tous les collègues astrophysiciens (Hubert B., Benoît F., Jonathan F., Rodrigo I., Katarina K., Yassin R.) et informaticiens (Étienne B., Jean-Michel D., Franck H.W., Pierre K., Nicolas M., Stella M.Z.) que j'ai pu ainsi rencontrer : c'est une joie d'échanger avec vous. Je tiens à remercier également vivement tous les soutiens du service administratif de l'IRMA, Delphine S., Delphine K., Jessica M.S., du service informatique, Matthieu B., Alexis P., Alain S., et du service des bibliothèques, Christine C., Christine D., Marounia K., Grégory T.. Je remercie également vivement Ouiza H. côté INRIA.

Bien que ce manuscrit soit un rapport de recherche, l'enseignement tient une place très importante dans mon quotidien. Je remercie donc tous les personnels administratifs de l'UFR, actuels et passés, pour leur soutien apporté dans cette mission, notamment Claudine B., Marie-Line V. et Marjorie S.. C'est vraiment très agréable de pouvoir compter sur vous. Merci également à Vincent B., Benjamin E., Michaël G. et Olivier G. pour leurs conseils et soutiens lors de mes différentes prises de responsabilités. Je remercie aussi tous les étudiants et étudiantes que j'ai pu croiser. Je pense plus spécialement à ceux du Master Agrégation, du Diplôme d'Université MARI et du master de Physique Cellulaire : c'est très agréable de vous accompagner.

Pour finir, je remercie vivement mes amours : Perrine et (les grandes) Esther et Daphné. Merci pour tous les sourires, toutes les émotions, toute l'énergie et pour toute la force que vous me procurez ! Je n'oublierai pas ces magnifiques panoramas sur les falaises de craies, ces crabes qui se cachaient sous les galets et ces quelques heures passées à (presque) finaliser ce manuscrit loin de vous. Merci pour tout. Je finis ces remerciements par une chanson :

C'est quoi l'essentiel ?  
C'est de toujours y croire  
Et un souvenir ?  
Un dessin sur la mémoire  
C'est quoi un sourire ?  
C'est du vent dans les voiles  
Et la poésie ?  
Une épuisette à étoiles.

[...]

Mais dis, papa, la vie, c'est quoi ?  
Petite, tu vois  
La vie, c'est un peu de tout ça,  
mais surtout c'est toi

Aldebert, "La vie c'est quoi ?" - Les enfantillages, 2017

## Abstract

This manuscript presents different contributions in the mathematical modeling and the development of numerical methods for multiscale transport problems, arising in particular in plasma physics and collective cell dynamics.

In the first chapter, we present a few reduction methods for kinetic transport equations. Performing a finite element semi-discretization in velocity leads to a linear hyperbolic model with good conservation properties. Fluid models with entropic or neural network closures have been studied to obtain even smaller reduced models with only physical quantities as reduced macroscopic variables. Reduced order modeling techniques are also proposed to construct reduced macroscopic variables specifically adapted to a given set of parametric solutions.

The second chapter deals with two numerical issues concerning the discontinuous Galerkin and its semi-Lagrangian version, namely the recurrence numerical artifact and the approximate well-balanced property by basis enrichment.

The third chapter presents two implicit numerical methods for solving hyperbolic systems with low-computational cost thanks to relaxation techniques. The first method relies on the vectorial kinetic relaxation scheme, whose transport steps are solved using an implicit solver with explicit cost. The second one proposes a splitting method applied to the two-speed relaxation system for the Euler system, whose acoustic part involves only constant coefficients. This leads to implicit steps with controlled conditioning.

Finally, the fourth chapter regards the development of models for collective dynamics of cells. We present three different models with specific interaction rules: layer interactions, back-to-back anti-alignment and interactions with surrounding elastic cables. We study their large-scale dynamics either theoretically, by deriving the associated macroscopic models, or using numerical simulations. Comparisons with experimental results have also been performed.



# Contents

<b>Introduction</b>	<b>1</b>
<b>I Reduced methods for kinetic models</b>	<b>9</b>
1 Introduction . . . . .	9
2 Reduction with nodal values . . . . .	11
3 Fluid models with an entropic closure . . . . .	17
4 Fluid models with a data-driven closure . . . . .	20
5 Fluid models with a data-driven velocity basis . . . . .	22
6 Conclusion and perspectives . . . . .	25
References . . . . .	26
<b>II Numerical methods for kinetic and hyperbolic systems</b>	<b>31</b>
1 Introduction . . . . .	31
2 Recurrence phenomenon on a regular finite element mesh . . . . .	31
3 Prior-basis enrichment for the DG scheme with PINNs . . . . .	35
4 Conclusion . . . . .	37
References . . . . .	38
<b>III Multi-scale hyperbolic problems and implicit schemes</b>	<b>41</b>
1 Introduction . . . . .	41
2 Implicit vectorial kinetic relaxation schemes . . . . .	42
2.1 Reformulation . . . . .	42
2.2 Implicit splitting scheme . . . . .	44
2.3 High order time splitting method . . . . .	46
2.4 Equivalent equations and stability issues . . . . .	47
2.5 Applications to plasma physics simulations . . . . .	50
3 Implicit Suliciu relaxation scheme . . . . .	52
4 Conclusion and perspectives . . . . .	57
References . . . . .	58
<b>IV Models for cells collective dynamics</b>	<b>61</b>
1 Introduction . . . . .	61
2 A multi-layer model for interacting disks . . . . .	63
3 Asymmetric interactions and phase transitions . . . . .	65
4 Effect of self-generated active boundaries and hard-repulsion . . . . .	70
5 Conclusion and perspectives . . . . .	74
References . . . . .	75



# Introduction

---

This manuscript presents different contributions on the mathematical modeling and the development of numerical methods for multiscale transport problems, arising in particular in plasma physics and collective cell dynamics.

When considering a large number of interacting and moving particles, we are generally interested in the resulting dynamics at large time and space scales of macroscopic quantities, rather than in the precise individual trajectories of each particle. Examples of particle systems are (a) ions-electrons with long-range electromagnetic interactions in plasma, (b) cells with local polarity alignment in biological tissues or (c) molecules with hard-sphere interactions in a neutral gas. To derive the macroscopic dynamics from the underlying particle models, an intermediate step consists in looking for the dynamics of the statistical distribution function of the particles in phase space. Assuming some independence between particles when their number is large, this distribution function satisfies a so-called kinetic transport equation with a collision operator, which makes the distribution relax towards a local thermodynamical equilibrium. Then, in the large collisional regime, the dynamics can be reduced to the fluid equations satisfied by the macroscopic quantities that parameterize the local thermodynamical equilibria. In a very schematic way, this is how the equations of fluid mechanics, e.g. the Euler or the Navier-Stokes equations, are derived from the particle dynamics.

The obtained fluid models are very attractive from a computational point of view as they are far less computationally demanding than their kinetic counterparts. However, their domain of validity is limited to collisional regimes. In weakly- or non-collisional regimes, we must rely on the statistical kinetic description of the particles. This is for instance the case for the dynamics of plasma in a fusion device. Solving the corresponding kinetic equation on a supercomputer is, however, not feasible at the moment, as it would demand too much computational and storage resources. This is why a number of **reduced kinetic or fluid models**, with specific domains of validity, have been developed: they require to first identify appropriate reduced kinetic or macroscopic quantities, and then to derive their dynamics. Some of them are based on asymptotic analysis and require a good knowledge of the physical system. Alternatively, reduced quantities can be constructed using standard discretization methods or directly from data coming from experiments or numerical simulations. **A central question in this program are how mathematical structures, such as the conservation of physical quantities, are preserved at the reduced level and more fundamentally how to ensure good stability properties for the reduced models.**

Whether for kinetic or fluid models, numerical methods need to be carefully optimized to tackle simulations in high dimensions of space or phase space. This is why high-order accurate methods are generally considered. For instance, when solving fluid models with a hyperbolic structure, the Discontinuous Galerkin (DG) method is a method of choice, as it can handle unstructured meshes and is quite appropriate for code parallelization. However specific algorithms are required to run the code on supercomputers or to take advantage of Graphics Processing Units

(GPUs). Another way of improving performance would be to rely on the recent development of machine learning tools. **One current question is how the numerical methods could be optimized by using data from simulations to increase the accuracy or improve the stability of the methods.**

Another difficulty comes from the fact that explicit numerical methods are limited by the Courant-Friedrichs-Lewy (CFL) stability condition: the time step has to be smaller than the maximal characteristic speed of the problem multiplied by the minimal mesh element size. This condition may be too stringent in possibly two different situations: either the maximal characteristic speed of the problem corresponds to the propagation of some low amplitude perturbations, which is not of major interest for the physical problem under consideration, or the size of the minimal mesh element is not determined by the target dynamics but by some mesh generation constraints. To bypass this stability condition, implicit methods are generally employed. However, the difficulty is now transferred to the resolution of large linear systems, possibly with bad condition numbers. **One challenge is the development of implicit methods with low computational cost.**

Returning to the initial hierarchy of models (particle, kinetic, macroscopic), one fundamental modeling issue concerns replacing inert particles with active ones, like cells. For instance, cells have the ability to self-propel and have polarity or velocity alignment interactions. At a large scale, this leads to the emergence of cell collective motions, which can be observed for instance in embryo development or tissue remodeling. One key feature of these models is that there is a priori no conservation of momentum and energy, which complexifies the development of macroscopic models. Another difficulty results from the appearance of phase transitions, like liquid-gas or disorder-order. From a practical point of view, estimating the parameters included in such models is crucial to assess the prediction ability of the mathematical models against experiments. **This calls for the development of new mathematical models as well as close collaborations with biologists and biophysicists.**

This manuscript presents several contributions introducing new mathematical models and numerical methods to address these questions. They have been possible thanks to the stimulating collaborations with colleagues from mathematics, computer science, physics and biology. We provide below a more precise overview, insisting on the contributions of the PhD students, whom we supervised. Note that the affiliations are the current ones.

In Chapter I, we present a few **reduction methods for kinetic equations** developed in different physical contexts:

- (i) The first consists in performing a finite element semi-discretisation in velocity, which leads to a linear hyperbolic reduced model. The model has good conservation properties and can be solved using a high-order DG method. The method was tested on the Curie supercomputer in a Cartesian geometry for the Vlasov-Poisson system in 4 dimensions (2 space, 2 velocity). These results were obtained during the PhD thesis of Nhung Pham, co-advised with Philippe Helluy (Univ. Strasbourg), and in collaboration with Anaïs Crestetto, Sébastien Guisset, Michaël Gutnic and Michel Massaro. She defended her PhD thesis in December 2016. The corresponding publications are [Red7, Red5, Red4].
- (ii) Another reduction technique consists in looking at the dynamics of moments of the distribution function. In the context of the study of room acoustics, we apply the entropic closure method to the kinetic transport equation satisfied by sound particles. This study has led to the development of a code in the parallel `pyopenc1` language (GPU). Comparisons on room acoustics test cases have shown that accurate sound energy decays can be achieved.

These developments have been achieved during the PhD of Pierre Gerhard, co-advised with Philippe Helluy (Univ. Strasbourg) and Cédric Foy (Cerema, Strasbourg). He defended his PhD in January 2020.

- (iii) Considering the Vlasov-Poisson dynamics with a BGK collision operator, we explored the ability to extend the validity of the classical fluid models to weakly collisional regimes by constructing a neural network closure. It consists in learning a nonlocal heat flux as a function of the density, velocity and temperature profiles. The learning process is carried out using data from numerical simulations of the kinetic model, and the network has a multi-scale convolutional structure. This work has been obtained during the PhD of Léo Bois, co-advised with Emmanuel Franck (INRIA, Univ. Strasbourg) and Vincent Vigon (Univ. Strasbourg). He defended his PhD in December 2023. The corresponding publication is [Red3].
- (iv) Finally, we present a reduction method which provides the space-time dynamics of a few macroscopic quantities constructed from data following the Reduced Order Method (ROM) in the velocity variable. The method is applied to the semi-discretized Vlasov-Poisson system with a non-linear Fokker-Planck collision operator. This work was initiated during a CEMRACS 2022 project, co-supervised with Emmanuel Franck (INRIA, Univ. Strasbourg), and involving Ibtissem Lannabi (Univ. Pau), Youssouf Nasseri (Univ. Strasbourg), Guiseppe Parasiliti Rantone (Sorbonne Univ.), and Guillaume Steimer (Univ. Strasbourg). The corresponding publication is [Red6].

In Chapter II, two different works related to Discontinuous Galerkin methods and its semi-Lagrangian version are presented.

- (i) The recurrence phenomenon is a numerical artifact, appearing in the numerical simulations of the Vlasov-Poisson system, that results from phase-space filamentation and aliasing effects. We have proposed an analysis of this phenomenon when a finite element grid is used in velocity, like in the Semi-Lagrangian Discontinuous Galerkin method. We show that the first recurrence time depends on the quadrature rule used to compute the density. This work was conducted in collaboration with Michel Mehrenberger (Aix-Marseille Univ.) and Nhung Pham. The associated publication is [Num4].
- (ii) When considering hyperbolic systems with source terms, a main issue is to preserve equilibria at the discrete level, at least approximately. This would enable us to capture perturbation dynamics around these equilibria. To this aim, we propose a DG method with local Taylor bases enriched with a prior estimate of the equilibrium. This prior can be obtained using Physics-Informed Neural Networks (PINNs), to capture a whole family of equilibria. This work was carried out with Emmanuel Franck (INRIA, Univ. Strasbourg) and Victor Michel-Dansac (INRIA, Univ. Strasbourg). The associated publication is [Num3].

Chapter III concerns **implicit numerical methods for multiscale hyperbolic problems** using relaxation techniques. They are based on reformulations of the hyperbolic systems as singular limits of larger systems, with more variables, but with a simpler mathematical structure.

- (i) Based on the vectorial kinetic relaxation technique, we propose a high-order implicit method with explicit cost. Each additional unknown satisfies kinetic transport equations at constant velocities that are solved using an implicit solver. A composition method based on symmetric second-order splitting is used to obtain fourth- or sixth-order time accuracy. An equivalent equation for both the variables of interest and the additional variables have been obtained. Finally, we applied this method to the numerical simulation of a three-dimensional version of the guiding center model for plasma physics: the method pre-

vents refined poloidal geometries from constraining the resolution of the toroidal dynamics. This last result has been obtained during the PhD thesis of Romane H elie, co-advised with Philippe Helluy (Univ. Strasbourg) and Emmanuel Franck (INRIA, Univ. Strasbourg). She defended her PhD in March 2023. All these developments were also obtained thanks to the collaborations with Thomas Bellotti (Univ. Strasbourg), Mathieu Boileau (CNRS, Univ. Strasbourg), B erenger Bramas (INRIA, Univ. Strasbourg), David Coulette (ENS Lyon), Cl ementine Court es (Univ. Strasbourg), Florence Drui (Univ. Strasbourg), Michel Mehrenberger (Aix-Marseille Univ.). The corresponding publications are [Imp6, Imp7, Imp3, Imp1].

- (ii) A specific implicit relaxation method for the compressible Euler equations is designed to overcome the CFL stability condition in the low Mach limit while still capturing the material wave with good accuracy. The method is based on a two-speed Suliciu relaxation formulation solved using a semi-implicit splitting strategy. The key point is that, due to the specific relaxation system, the acoustic part leads to a constant-coefficient elliptic problem whose conditioning can be controlled. This work has been carried out in collaboration with F. Bouchut (CNRS, Univ. Gustave Eiffel) and E. Franck (INRIA, Univ. Strasbourg). The corresponding publication is [Imp4].

In Chapter IV, several **models for cell collective dynamics** are proposed and studied.

- (i) Inspired by the layered movements of spermatozoa observed experimentally, we present a model of disc-shaped particles that exhibits both velocity-aligned interactions (Vicsek-type interactions) within each layer and nematic-type interactions between layers. The derivation of the associated macroscopic model required the extension of the method of generalized collisional invariants. Numerical simulations also show the appearance of interesting metastable states. This work was carried out in collaboration with Pierre Degond (CNRS, Univ. Toulouse) and the associated publication is [Cell6].
- (ii) Phase transitions are studied in the context of an original two-dimensional particle model with either opposite alignment or alignment of the velocities of neighboring particles depending on whether they are back to back or not. The model exhibit three kinds of phases: disordered phase, nematic phase and polar-nematic phase. Macroscopic equations have been obtained for the first two phases. This ongoing work is in collaboration with Nicolas Meunier (Univ. Cergy) and Rapha el Voituriez (CNRS, Sorbonne Univ.).
- (iii) We finally introduced a Vicsek-type model, with contact forces and specific boundary conditions: cells have a ring configuration, confined between two elastic cables. They model actomyosin cables that form at the periphery of the cell cluster. A parameter calibration has been performed and we were able to reproduce the emergence of the collective rotation observed in experiments. This study shows for the first time the decisive impact of these cables in the emergence of the collective movement. This work was carried out with Simon Lo Vecchio and Daniel Riveline (IGBMC) and Marcela Szopos (Sorbonne Univ.). The associated publication is [Cell1].

Each chapter will conclude with a number of future research directions.

The following is a list of our publications in peer-reviewed journals or in conference proceedings, arranged by theme and in reverse chronological order. Publications directly following the PhD work have been grouped together at the end. The publications that have not been previously cited will not be described in this manuscript or will be mentioned very briefly.

## Reduction methods

- [Red1] R. Côte, E. Franck, L. Navoret, G. Steimer, and V. Vigon. “Hamiltonian reduction using a convolutional auto-encoder coupled to an Hamiltonian neural network”. In: *preprint* (2024). ([hal-04237799](#)).
- [Red2] C. Courtès, E. Franck, K. Lutz, L. Navoret, and Y. Privat. “Reduced modelling and optimal control of epidemiological individual-based models with contact heterogeneity”. In: *Optimal Control Appl. Methods* 45.2 (2024), pp. 459–493. ([hal-03664271](#)).
- [Red3] L. Bois, E. Franck, L. Navoret, and V. Vigon. “A neural network closure for the Euler-Poisson system based on kinetic simulations”. In: *Kinetic and Related Models* 15.1 (2022), pp. 49–89. ([hal-02965954](#)).
- [Red4] S. Guisset, M. Gutnic, P. Helluy, M. Massaro, L. Navoret, N. Pham, and M. Roberts. “Lagrangian/Eulerian solvers and simulations for Vlasov-Poisson”. In: *ESAIM: Proc.* 53 (2016), pp. 120–132. ([hal-01239673](#)).
- [Red5] P. Helluy, L. Navoret, N. Pham, and A. Crestetto. “Reduced Vlasov–Maxwell simulations”. In: *Comptes Rendus. Mécanique* 342.10-11 (2014), pp. 619–635. ([hal-00957045](#)).

### Conference papers

- [Red6] E. Franck, I. Lannabi, Y. Nasser, L. Navoret, G. Parasiliti Rantone, and G. Steimer. “Hyperbolic reduced model for Vlasov-Poisson equation with Fokker-Planck collision”. In: *ESAIM: Proc.* (2024). ([hal-04099697](#)).
- [Red7] N. Pham, P. Helluy, and L. Navoret. “Hyperbolic approximation of the Fourier transformed Vlasov equation”. In: *ESAIM: Proceedings and Surveys* 45 (2014), pp. 379–389. ([hal-00872972v1](#)).

## Numerical schemes for kinetic and hyperbolic equations

- [Num1] M. Bestard, E. Franck, L. Navoret, and Y. Privat. “Optimal scenario for road evacuation in an urban environment”. In: *Z. Angew. Math. Phys.* 75.4 (2024), p. 146. ([hal-04253010v1](#)).
- [Num2] L. Bois, E. Franck, L. L. Navoret, and V. Vigon. “Optimal control deep learning approach for viscosity design in DG schemes”. In: *preprint* (2024). ([hal-04213057](#)).
- [Num3] E. Franck, V. Michel-Dansac, and L. Navoret. “Approximately well-balanced Discontinuous Galerkin methods using bases enriched with Physics-Informed Neural Networks”. In: *J. Comp. Phys.* (2024), p. 113144. ([hal-04246991v2](#)).
- [Num4] M. Mehrenberger and L. Navoret. “Recurrence Phenomenon for Vlasov-Poisson Simulations on Regular Finite Element Mesh”. In: *Commun. in Comput. Physics* 28.3 (2020). ([hal-01942708](#)), pp. 877–901.

### Conference papers

- [Num5] M. Mehrenberger, L. Navoret, and A.-T. Vu. “Composition schemes for the guiding-center model”. In: *International Conference on Finite Volumes for Complex Applications*. Springer. 2023, pp. 239–247. ([hal-04035921](#)).
- [Num6] M. Badsì, M. Mehrenberger, and L. Navoret. “Numerical stability of a plasma sheath”. In: *ESAIM: Proceedings and Surveys* 64 (2018), pp. 17–36. ([hal-01676656](#)).
- [Num7] M. Billaud-Friess, B. Boutin, F. Caetano, G. Faccanoni, S. Kokh, F. Lagoutière, and L. Navoret. “A second order anti-diffusive Lagrange-remap scheme for two-component flows”. In: *ESAIM: Proceedings*. Vol. 32. EDP Sciences. 2011, pp. 149–162. ([hal-00708605v1](#)).

- [Num8] F. Charles, N. Vauchelet, C. Besse, T. Goudon, I. Lacroix-Violet, J.-P. Dudon, and L. Navoret. “Numerical approximation of Knudsen layer for the Euler-Poisson system”. In: *ESAIM: Proceedings*. Vol. 32. EDP Sciences. 2011, pp. 177–194. ([hal-00708605v1](#)).

## Implicit schemes for multi-scale hyperbolic problems

- [Imp1] T. Bellotti, P. Helluy, and L. Navoret. “Fourth-order entropy-stable lattice Boltzmann schemes for hyperbolic systems”. In: *preprint* (2024). ([hal-04510582](#)).
- [Imp2] N. Chaudhuri, L. Navoret, C. Perrin, and E. Zatorska. “Hard congestion limit of the dissipative Aw–Rascle system”. In: *Nonlinearity* 37.4 (2024), p. 045018. ([hal-03786853](#)).
- [Imp3] M. Boileau, B. Bramas, E. Franck, R. Hélie, P. Helluy, and L. Navoret. “Parallel kinetic scheme for transport equations in complex toroidal geometry”. In: *SMAI J. Comput. Mathematics* 8 (2022), pp. 249–271. ([hal-02404082](#)).
- [Imp4] F. Bouchut, E. Franck, and L. Navoret. “A low cost semi-implicit low-Mach relaxation scheme for the full Euler equations”. In: *J. Scientific Comput.* 83 (2020), pp. 1–47. ([hal-02420859v3](#)).
- [Imp5] C. Courtès, D. Coulette, E. Franck, and L. Navoret. “Vectorial kinetic relaxation model with central velocity. application to implicit relaxations schemes”. In: *Commun. Comput. Physics* 27.4 (2020). ([hal-01942317](#)).
- [Imp6] D. Coulette, E. Franck, P. Helluy, M. Mehrenberger, and L. Navoret. “High-order implicit palindromic discontinuous Galerkin method for kinetic-relaxation approximation”. In: *Comput. Fluids* 190 (2019), pp. 485–502. ([hal-0170661](#)).
- [Imp7] F. Drui, E. Franck, P. Helluy, and L. Navoret. “An analysis of over-relaxation in a kinetic approximation of systems of conservation laws”. In: *Comptes Rendus Mécanique* 347.3 (2019), pp. 259–269. ([hal-01839092](#)).
- [Imp8] P. Degond, P. Minakowski, L. Navoret, and E. Zatorska. “Finite volume approximations of the Euler system with variable congestion”. In: *Comput. Fluids* 169 (2018), pp. 23–39. ([hal-01651479](#)).

### Conference papers

- [Imp9] E. Franck and L. Navoret. “Semi-implicit two-speed well-balanced relaxation scheme for Ripa model”. In: *Finite Volumes for Complex Applications IX-Methods, Theoretical Aspects, Examples: FVCA 9, Bergen, Norway, June 2020 IX*. Springer. 2020, pp. 735–743. ([hal-02407820](#)).
- [Imp10] R. Hélie, P. Helluy, E. Franck, and L. Navoret. “Kinetic over-relaxation method for the convection equation with Fourier solver”. In: *Finite Volumes for Complex Applications IX-Methods, Theoretical Aspects, Examples: FVCA 9, Bergen, Norway, June 2020 IX*. Springer. 2020, pp. 745–753. ([hal-02427044](#)).
- [Imp11] D. Coulette, E. Franck, P. Helluy, M. Mehrenberger, and L. Navoret. “Palindromic discontinuous Galerkin method”. In: *Finite Volumes for Complex Applications VIII-Hyperbolic, Elliptic and Parabolic Problems: FVCA 8, Lille, France, June 2017 8*. Springer. 2017, pp. 171–178. ([hal-01653049](#)).

## Modelling for cell dynamics

- [Cell1] S. Lo Vecchio, O. Pertz, M. Szopos, L. Navoret, and D. Riveline. “Spontaneous rotations in epithelia as an interplay between cell polarity and boundaries”. In: *Nat. Phys.* 20.2 (2024), pp. 322–331. ([hal-03427805](#)).

- [Cell2] N. Meunier, L. Navoret, and R. Voituriez. “Phase transitions in cell assemblies with asymmetric aligning interactions”. In: *in preparation* (2024).
- [Cell3] L. Navoret, R. Sublet, and M. Szopos. “From jamming to flocking in collective cell dynamics: a Vicsek-like model including contact forces”. In: *in preparation* (2024).
- [Cell4] S. Lo Vecchio, R. Thiagarajan, D. Caballero, V. Vigon, L. Navoret, R. Voituriez, and D. Riveline. “Collective dynamics of focal adhesions regulate direction of cell motion”. In: *Cell Systems* 10.6 (2020), pp. 535–542. DOI: [10.1101/750331](https://doi.org/10.1101/750331).
- [Cell5] B. Grec, B. Maury, N. Meunier, and L. Navoret. “A 1D model of leukocyte adhesion coupling bond dynamics with blood velocity”. In: *J. Theo. Biol.* 452 (2018), pp. 35–46. ([hal-01566770](https://hal.archives-ouvertes.fr/hal-01566770)).
- [Cell6] P. Degond and L. Navoret. “A multi-layer model for self-propelled disks interacting through alignment and volume exclusion”. In: *Math. Models Methods Appl. Sci.* 25.13 (2015), pp. 2439–2475. ([hal-01118377](https://hal.archives-ouvertes.fr/hal-01118377)).

#### Conference papers

- [Cell7] A. Cucchi, C. Etchegaray, N. Meunier, L. Navoret, and L. Sabbagh. “Cell migration in complex environments: chemotaxis and topographical obstacles”. In: *ESAIM Proc. Surveys* 67 (2020), pp. 191–209. ([hal-02126709v2](https://hal.archives-ouvertes.fr/hal-02126709v2)).
- [Cell8] C. Etchegaray, B. Grec, B. Maury, N. Meunier, and L. Navoret. “An integro-differential equation for 1D cell migration”. In: *Integral Methods in Science and Engineering: Theoretical and Computational Advances*. Springer, 2015, pp. 195–207. ([hal-01083510](https://hal.archives-ouvertes.fr/hal-01083510)).

## Articles based on my PhD work

- [Phd1] L. Navoret. “A two-species hydrodynamic model of particles interacting through self-alignment”. In: *Mathematical Models and Methods in Applied Sciences* 23.06 (2013), pp. 1067–1098. ([arxiv:1401.1379](https://arxiv.org/abs/1401.1379)).
- [Phd2] P. Degond, J. Hua, and L. Navoret. “Numerical simulations of the Euler system with congestion constraint”. In: *J. Comput. Phys.* 230.22 (2011), pp. 8057–8088. ([hal-01651479v1](https://hal.archives-ouvertes.fr/hal-01651479v1)).
- [Phd3] S. Motsch and L. Navoret. “Numerical simulations of a nonconservative hyperbolic system with geometric constraints describing swarming behavior”. In: *Multiscale Model. Simul.* 9.3 (2011), pp. 1253–1275. ([arxiv:0910.2951](https://arxiv.org/abs/0910.2951)).
- [Phd4] P. Degond, F. Deluzet, L. Navoret, A.-B. Sun, and M.-H. Vignal. “Asymptotic-preserving particle-in-cell method for the Vlasov–Poisson system near quasineutrality”. In: *J. Comp. Phys.* 229.16 (2010), pp. 5630–5652. ([hal-00629561](https://hal.archives-ouvertes.fr/hal-00629561)).
- [Phd5] P. Degond, L. Navoret, R. Bon, and D. Sanchez. “Congestion in a macroscopic model of self-driven particles modeling gregariousness”. In: *J. Stat. Phys.* 138 (2010), pp. 85–125. ([arxiv:0908.1817](https://arxiv.org/abs/0908.1817)).

#### Conference papers

- [Phd6] P. Degond, A. Frouvelle, J.-G. Liu, S. Motsch, and L. Navoret. “Macroscopic models of collective motion and self-organization”. In: *Séminaire Laurent Schwartz—EDP et Applications* (2012), pp. 1–27. ([hal-00816752v1](https://hal.archives-ouvertes.fr/hal-00816752v1)).
- [Phd7] L. Navoret, R. Bon, P. Degond, J. Gautrais, D. Sanchez, and G. Theraulaz. “Analogies Between Social Interaction Models and Supply Chains”. In: *Progress in Industrial Mathematics at ECMI 2008* (2010), pp. 535–540. ([hal-01872081](https://hal.archives-ouvertes.fr/hal-01872081)).

- [Phd8] P. Degond, F. Deluzet, and L. Navoret. “An asymptotically stable Particle-in-Cell (PIC) scheme for collisionless plasma simulations near quasineutrality”. In: *Comptes rendus. Mathématique* 343.9 (2006), pp. 613–618. DOI: [10.1016/j.crma.2006.09.033](https://doi.org/10.1016/j.crma.2006.09.033).

---

## Chapter I

# Reduction methods of kinetic models

---

## 1 Introduction

We are here interested in the dynamics of a large system of interacting particles. In many physical applications, there are so many particles that a microscopic description of the individual motions of each particle would be numerically intractable. Instead, we turn to kinetic models, that describes the dynamics of the statistical distribution function of particles in position-velocity phase space. The distribution function thus writes  $f(x, v, t)$  with  $x \in \Omega_x \subset \mathbb{R}^d, v \in \Omega_v \subset \mathbb{R}^d$  and  $t \geq 0$  and satisfies a so-called kinetic equation. Kinetic equations are, however, still difficult to solve numerically due to the large dimension of the phase space: considering  $N$  discretization points in each dimension would lead to  $N^{2d}$  unknowns. Thus updating the distribution function requires many computational resources and just storing the distribution function also requires large resources.

Therefore, we are interested in devising so-called reduced models, which aims at providing approximate solutions valid at least in some range of parameters. In our work, we focus on **reduced fluid models**, where the reduction is made in the velocity variable only. In the present context, reduced variables can be nodal values of the distribution function at some velocity grid points  $(v_k)$ ,

$$U_k(x, t) = f(x, v_k, t), \quad \forall k \in \llbracket 1, K \rrbracket, \quad (1)$$

or integral quantities in velocity

$$U_k(x, t) = \int_{\Omega_v} m_k(v) f(x, v, t) dv, \quad \forall k \in \llbracket 1, K \rrbracket, \quad (2)$$

where  $(m_k(v))_k$  are linearly independent functions. We then would like to derive the space-time dynamics of the reduced quantities  $U(x, t) = (U_1(x, t), \dots, U_K(x, t))$ . Without any other assumptions, they satisfy a system like

$$\frac{dU}{dt}(x, t) = \mathcal{F}(U(x, t), f(x, v, t)),$$

where  $\mathcal{F}$  is a differential operator. This model still depends on the distribution function  $f$ . To close the system, i.e. remove the dependency on the distribution function, this latter should be supposed to be entirely reconstructible from the  $K$  reduced variables, with a formula like

$$f(x, v, t) = M_{U(x, t)}(v),$$

where  $M_U(v)$  denotes a parametrized function with linear or non-linear dependency with respect to the reduced vector  $U(x, t)$  such that relations (1) or (2) are satisfied. However, the range of

validity can be limited with small values of  $K$ . Extending the range of validity of these reduced models are thus of interest.

Let us now mention some of the existing methods proposed in the literature. First, reduced fluid models using spectral discretization methods in velocity have been obtained using either Fourier transform [34, 20] or Hermite functions [32, 35, 14, 2] when  $\Omega_v = \mathbb{R}^d$  or spherical harmonics when  $\Omega_v = \mathbb{S}^{d-1}$  and a linear reconstruction to get the reconstructed distribution function. In these methods, the number of reduced quantity to be taken into account is dictated by the regularity in velocity of the target solutions. In contrast, moment methods consider possibly only a few reduced quantities and apply a non-linear reconstruction [38, 5, 49]. For the radiative transfer equation, where  $\Omega_v = \mathbb{S}^{d-1}$  is compact, this leads to well-posed reduced fluid models [17, 18, 25]. However, for gas dynamics equations where  $\Omega_v = \mathbb{R}^d$ , so-called realizability issues may arise: the reduced vector  $U(x, t)$  may no longer correspond to a non-negative distribution function in the course of the dynamics [33, 45].

In this chapter, we explore four different approaches to construct reduced fluid models using (i) values at velocity grid points and a **finite-element vector space closure**, (ii) moments and **the entropic minimization closure**, (iii) moments and a **data-driven closure**, (iv) **data-driven reduced variables** made of linear combinations of values at velocity grid points. A specific attention has been paid to the efficient implementation of the reduced models. The main application will be the construction of reduced models for the **Vlasov equation and plasma simulations**, except for (ii), which is dedicated to the **free transport kinetic equation** with application in acoustics. We refer to Box 1 for more information on this physical context.

#### Box 1: Plasma dynamics in Tokamak

Plasma is a gas made of charged particles, i.e. ions and electrons. Ions and electrons interact with each other through the self-consistent electromagnetic fields  $E(x, t)$  and  $B(x, t)$ , which satisfies the Maxwell equations.

Fusion reactor aims at extracting energy present in the kernel of ions, in the strong interactions between nucleons. Indeed, when two ions fuse, part of the energy is transformed into kinetic energy of the reaction output and this energy can be captured by heating water. To carry out this fusion reaction, the plasma has to be at a very high temperature during some time so that a few ions can pass through the Coulombic repulsion barrier and fuse. From a practical point of view, due to the large temperature, the plasma has to be confined without contact with any material. A solution is to trap it into closed magnetic field lines. Tokamak is such a fusion device, with a toroidal geometry: a strong toroidal magnetic field is created using superconducting coils. To avoid plasma leaks due to the combination of the  $B \times \nabla_x B$  and  $E \times B$  drifts, a toroidal current is generated inside the plasma to produce a self-consistent magnetic field that bends the magnetic field lines. The resulting dynamics of the plasma is highly complex, with the appearance of strong density and temperature gradients and turbulence. Numerical simulations are thus required to predict and then control the quality of the confinement.

The most complete physical model is the **kinetic Vlasov-Maxwell** system, which describes the dynamics of the particles under mean-field approximation. The Vlasov equation describes the transport in position-velocity phase space of the ions and electrons, each described by their kinetic distribution functions. The Maxwell equations provide the dynamics of the self-consistent electromagnetic fields. Collision operators can be added to take into account short-range interactions. However, this model raises three main issues from a computational point of view:

1. large dimension: simulations of the dynamics in six-dimensional position-velocity phase space is at the moment out of reach because of computation and storage limitations.
2. multi-scale dynamics: the problem involves many different time and space scales.
3. complex geometry: despite the toroidal symmetry of the fusion device, poloidal sections may have complex geometries.

To overcome these difficulties, **kinetic reduced models** have been developed. Indeed, in the strong magnetic field regime, ions and electrons oscillate rapidly around the magnetic field lines. By somehow averaging the fast rotations, the so-called gyrokinetic models only describe the slow dynamics of the rotation centers and the reduced velocity variable has only two components, one of which is constant over time. The reduced kinetic phase space is thus of dimension 5 instead of 6.

We refer to [8, 28] for more details on plasma physics and gyrokinetic models.

## 2 Reduction with nodal values

This first reduction method we studied is based on a semi-discretization of the problem in velocity: it consists into projecting the distribution function  $f(x, v, t)$  in a finite-element vector space in velocity. The method has first been proposed in dimension 1 in [42] and then extended in [Red5, Red7, Red4, 41] during the PhD thesis of Nhung Pham, whom I co-supervised with Philippe Helluy.

We first consider a bounded cubic velocity domain  $\Omega_v = [-v_{\max}, v_{\max}]^d$  as well as a regular Cartesian mesh on this domain. Then we choose the space of continuous piece-wise polynomial functions of degree at most  $q$  on each mesh element, based on the tensorial Lagrange  $\mathbb{Q}_q$  finite-element space. Denoting  $N_v$  the total dimension of this space and  $(\varphi_i(v))_{i=1, \dots, N_v}$  an interpolation basis of this vector space, the distribution function writes:

$$M_{U(x,t)}(v) = \sum_{i=1}^{N_v} U_i(x, t) \varphi_i(v),$$

where  $(U_i(x, t))_{i=1, \dots, N_v} \in \mathbb{R}^{N_v}$  denote the  $N_v$  coefficients in this basis. We note that the distribution function is defined such as to satisfy an interpolation property:  $U_i(x, t) = M_{U(x,t)}(v_i)$  for all  $i \in \llbracket 1, N_v \rrbracket$ .

**Vlasov model.** Instead of considering at the Vlasov-Maxwell system mentioned in Box 1, let us consider the simplified Vlasov equation

$$\begin{aligned} \partial_t f(x, v, t) + v \cdot \nabla_x f(x, v, t) + E(x, t) \cdot \nabla_v f(x, v, t) &= 0, & \text{on } \Omega_x \times \Omega_v \times \mathbb{R}^+, \\ (E(x, t) \cdot n_v)^- f(x, v, t) &= 0, & \text{on } \Omega_x \times \partial\Omega_v \times \mathbb{R}^+, \end{aligned}$$

with a given electric field  $E(x, t)$ , on the bounded domain  $\Omega_x \times \Omega_v = [0, L]^d \times [-v_{\max}, v_{\max}]^d$ , with periodic boundary conditions in space and zero inflow boundary conditions in velocity: when the electric field is incoming ( $E(x, t) \cdot n_v < 0$  with  $n_v$  the outward normal to the velocity domain), then the distribution function is set to zero. Then the weak formulation of this transport equation writes as follows: find  $f(x, \cdot, t) \in V = \{f \in L^2(\Omega_v), E \cdot \nabla_v f \in L^2(\Omega_v)\}$  such that

$$\partial_t \int_{\Omega_v} f \varphi + \nabla_x \cdot \int_{\Omega_v} v f \varphi + E \cdot \int_{\Omega_v} \nabla_v f \varphi - \int_{\partial\Omega_v} (E \cdot n_v)^- f \varphi = 0, \quad (3)$$

for all  $\varphi \in V$ , where the outflow boundary conditions are weakly imposed.

**Reduced model.** The semi-discrete formulation can then be obtained by considering the following problem: find  $f(x, \cdot, t) \in V_h = \text{span}(\{\varphi_i\}_i)$  such that

$$\partial_t \int_{\Omega_v} f \varphi + \nabla_x \cdot \int_{\Omega_v} v f \varphi + E \cdot \int_{\Omega_v} \nabla_v f \varphi - \int_{\partial\Omega_v} (E \cdot n_v)^- f \varphi = 0, \quad (4)$$

for all  $\varphi \in V_h$ . Once expressed into the basis  $(\varphi_i(v))$ , we get the following reduced model:

**Proposition 1** (Reduced Vlasov model [41]). *The vector  $U(x, t) = (U_1(x, t), \dots, U_{N_v}(x, t))^T$  satisfies the following system*

$$M \partial_t U + \sum_{k=1}^d A_k \partial_{x_k} U + B(E) U = 0, \quad (5)$$

where  $M$ ,  $A_k$  and  $B(E)$  are matrices of dimension  $N_v \times N_v$ , whose elements are given by

$$M_{ij} = \int_{\Omega_v} \varphi_i \varphi_j, \quad (A_k)_{ij} = \int_{\Omega_v} v_k \varphi_i \varphi_j, \quad k = 1 \dots 3,$$

$$B(E)_{ij} = \int_{\Omega_v} \varphi_i (E \cdot \nabla_v) \varphi_j - \int_{\partial\Omega_v} (E \cdot n_v)^- \varphi_j \varphi_i.$$

$M$  is symmetric positive definite and  $(A_k)$  are symmetric: the system is thus **hyperbolic**.

In addition to its hyperbolicity, the reduced model ensures  $L^2$  stability, meaning that  $f = \sum_{i=1}^{N_v} U_i \varphi_i$  satisfies

$$\frac{1}{2} \frac{d}{dt} \left( \int_{\Omega_x \times \Omega_v} f^2 \right) \leq - \frac{1}{2} \int_{\Omega_x \times \partial\Omega_v} |E \cdot n_v| f^2.$$

Moreover, as soon as  $p \geq 2$ , the polynomials  $1$ ,  $v$  and  $v_i v_j$ , for  $1 \leq i, j \leq d$  belong to the approximation space  $V_h = \text{span}(\{\varphi_i\}_i)$  and the following conservation properties hold:

$$\begin{aligned} \frac{d}{dt} \left( \int_{\Omega_x \times \Omega_v} f \right) &= - \int_{\Omega_x \times \partial\Omega_v} (E \cdot n_v)^+ f, \\ \frac{d}{dt} \left( \int_{\Omega_x \times \Omega_v} v^2 f \right) &= - \int_{\Omega_x \times \partial\Omega_v} (E \cdot n_v)^+ f \left( \frac{|v|^2}{2} + \nabla_x E \right), \\ \frac{d}{dt} \left( \int_{\Omega_x \times \Omega_v} v f \right) &= - \int_{\Omega_x \times \partial\Omega_v} (E \cdot n_v)^+ f v, \end{aligned}$$

supposing periodic boundary conditions in space.

**Reduction with quadrature approximations.** The system is further simplified by considering a specific choice of basis and quadrature points to compute the coefficients of the matrices  $M$  and  $A_k$ . In each square element  $L$ , Gauss-Lobatto points  $(v_{L,j})_{0 \leq j < (q+1)^d}$  are chosen and integrals are approximated by Gauss quadrature rules with weights  $(\omega_{L,j})_{0 \leq j < (q+1)^d}$ . The associated Lagrange basis  $(\varphi_i(v))$  is made of Gauss-Lobatto polynomials on each element and satisfies the interpolation property:

$$\varphi_i(v_k) = \delta_{ik},$$

where we use the notation  $v_i = v_{L,j}$  with a local-global index correspondence. When using the Gauss quadrature, the approximated mass matrix  $\tilde{M}$  and the approximated advection matrices  $\tilde{A}_k$  become diagonal with coefficients:

$$\tilde{M}_{i,i} = \sum_{v_i=v_{L,j}} \omega_{L,j}, \quad (\tilde{A}_k)_{i,i} = M_{i,i} (v_i)_k,$$

and the approximated system writes as a simple **diagonal transport** system:

$$\partial_t U + \sum_{k=1}^d \text{Diag}((v_i)_k) \partial_{x_k} U + \tilde{M}^{-1} B(E) U = 0. \quad (6)$$

Note that this is reminiscent of mass lumping technique in finite element methods and is classical in nodal Discontinuous Galerkin methods [31].

**Different spatial and time discretizations.** Regarding only the transport part of the equation, different discretization have been investigated. A **finite-volume discretization** with the upwind flux<sup>1</sup> has been proposed in [42] in the one-dimensional case for System (5). It has also been observed that high-order Runge-Kutta time discretizations allow us to dispense with artificial viscosity or at least to adjust it to as low a level as desired. In [41], a detailed stability analysis has been carried out, in the one-dimensional case ( $d = 1$ ) on a regular Cartesian grid with space step  $\Delta x$ , which we present here. Considering the standard decentered numerical flux, the space semi-discretized scheme writes:

$$\partial_t U_i + (M^{-1}A) \frac{U_{i+1} - U_{i-1}}{2\Delta x} - \frac{\kappa}{2\Delta x} (U_{i+1} - 2U_i + U_{i-1}) = 0,$$

with a viscosity parameter  $\kappa \geq 0$ . After noting that the spectral radius of the matrix  $M^{-1}A$  is actually smaller than  $v_{\max}$ , the following proposition gathers stability results for the first four Runge-Kutta schemes:

**Proposition 2** (Stability without the source term [41]). *The Runge-Kutta finite-volume schemes with decentered numerical flux are stable under the condition  $\Delta t = \alpha \Delta x / v_{\max}$  with*

$$\begin{array}{ll} (RK1) & \alpha \in (0, 1), & \alpha^{-1} > \kappa / v_{\max} > \alpha, \\ (RK2) & \alpha \in (0, \sqrt{3}), & \alpha^{-1} > \kappa / v_{\max} > \alpha^3 / 4, \\ (RK3) & \alpha \in (0, 1.5), & (3/4)\alpha^{-1} > \kappa / v_{\max}, \\ (RK4) & \alpha \in (0, 2), & (3/4)\alpha^{-1} > \kappa / v_{\max}. \end{array}$$

*The second case has been only numerically checked and the last two conditions are non optimal.*

This proposition shows that the first and second order numerical schemes always require some viscosity to be stable, while the third and fourth order in time schemes can be used with a tunable viscosity parameter. We refer to [48] for recent results on this topic.

Taking benefit from the diagonal formulation (6), a **semi-Lagrangian scheme** can be used to solve the  $N_v$  resulting transport equations. This has been investigated in [Red4]. We refer to Box 2 for a general presentation of semi-Lagrangian methods. Using a dimensional splitting, the method consists into the resolution of a succession of one-dimensional transport equations. High order spatial accuracy can be obtained using either cubic spline or Lagrange interpolation. The main advantage of the semi-Lagrangian methods is to be exempt from stability conditions: the time step can be considered as large as desired. Strang splitting can then be used to gain accuracy in time. However, conservation properties are only ensured when considering a regular Cartesian grid: this limits us to using cubic spatial domains  $\Omega_x$ .

To hold general domains and still be high order accurate in space, a **discontinuous Galerkin scheme (DG)** can also be considered [Red5, Red7], as presented in Box 3. Runge-Kutta time

<sup>1</sup>The upwind flux coincides with the Godunov flux for linear hyperbolic systems.

discretizations can be used to achieve high order accuracy in time. Note, however, that this method has to satisfy a CFL stability constraint.

The reduced Vlasov model also contains a source term corresponding to the advection in the velocity direction with the electric field. The whole system can be either solve directly or using an operator splitting strategy, meaning that the two following equations are solved successively

$$\begin{aligned}\partial_t U + \sum_{k=1}^d (M^{-1} A_k) \partial_{x_k} U &= 0, \\ \partial_t U + (M^{-1} B(\mathbf{E})) U &= 0.\end{aligned}$$

In both cases, the source term induces some constraints on the time step: to capture the dynamics induced by the corresponding CFL like condition should be considered:  $\Delta t \leq \Delta v / \|E\|_\infty$ .

**Implementation issues.** In [Red4], we have carried out performance comparisons between the finite volume (FV), the semi-Lagrangian (SL) and the DG solvers on a two-dimensional square spatial domain. The FV and the SL version have been implemented in the `SeLaLib`<sup>2</sup> library, whose parallelism is based on MPI subdomain decomposition for the former and on MPI-based grid transposition for the latter. The DG version has been implemented in the `schnaps`<sup>3</sup> library based on the OpenCL framework to perform simulations on GPUs. Figure 1a shows that the three methods have roughly similar computation time per second-order time step when considering  $32 \times 32$  points in velocity, degree  $p = 2$  finite element in velocity, between  $32 \times 32$  to  $256 \times 256$  points in space and with deactivated source term. A first comparison between the SL (order 3 - cubic spline, order 9 - Lagrange) and the FV schemes have been done on a classical 1D Landau damping test case with 65 grid points in velocity and 32 spatial points. As expected, Figure 1d shows that both methods preserve mass up to some loss due to the boundary conditions in velocity. On Figure 1c, we also observe that the SL methods order 3 and 9 better preserve the  $L^2$  norm of the distribution function than the FV method as they are less diffusive. However, as they hardly preserve the positivity of the distribution function, the SL lack of preservation of the  $L^1$  norm of the distribution function (see Fig. 1b).

### Box 2: Semi-Lagrangian method

Considering a transport equation

$$\partial_t U + a \cdot \nabla_x U = 0,$$

with velocity field  $a(t, x) \in \mathbb{R}^d$ , the semi-Lagrangian method relies on the fact that the solution is constant along the characteristics curves  $(s, X(s; t, x))$ , solutions to the differential equations:

$$\frac{dX}{ds}(s; t, x) = a(X(s; t, x), t), \quad X(t; t, x) = x.$$

In particular, the solution at time  $t^{n+1}$  at point  $x$  equals the solution at time  $t^n$  evaluated at the root of the characteristic:

$$U(t^{n+1}, x) = U(t^n, X(t^n; t^{n+1}, x)).$$

<sup>2</sup><https://selalib.github.io/>

<sup>3</sup><https://github.com/phelluy/schnaps>

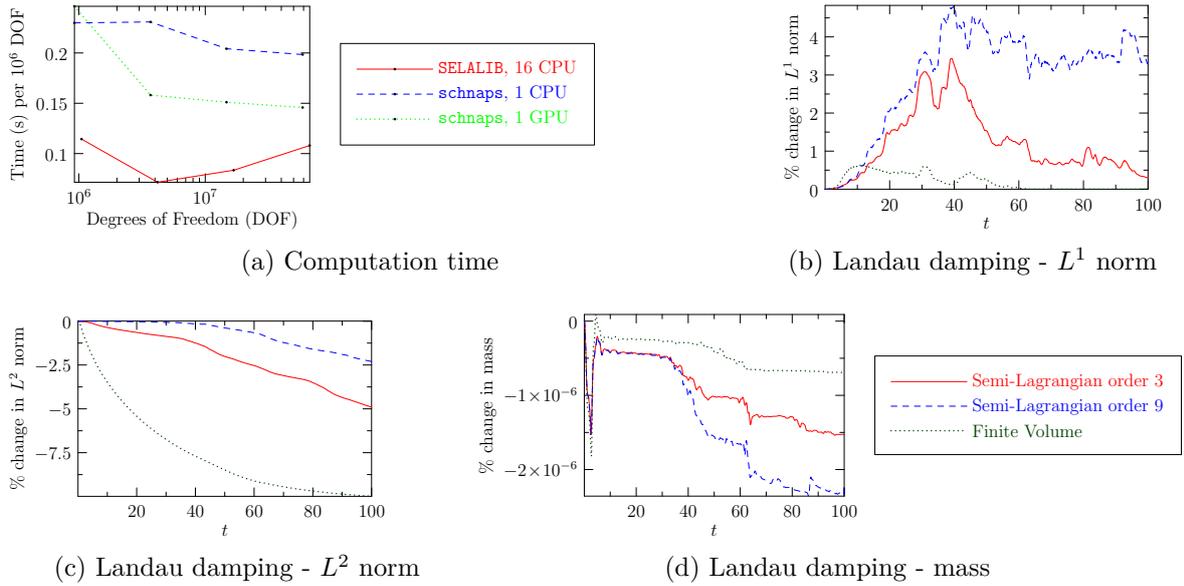


Figure 1: Top: Computation time per second-order time step per  $10^6$  degrees of freedom for SeLaLib with 16 CPUs and schnaps with 1 CPU or 1 GPU. Bottom: Time evolution of the change of the  $L^2$  norm of the distribution function (left) and of the change in mass for the strong Landau damping test-case [Red4].

Now, given a discrete solution ( $U_i^n$ ) at time  $t^n$  and points ( $x_i$ ), the solution at time  $t^{n+1}$  can be defined similarly :

$$U_i^{n+1} = [\Pi U^n](X(t^n; t^{n+1}, x_i)),$$

where  $\Pi$  denotes a reconstruction operator that defines a function of the whole space  $\Pi U^n$  from the nodal values ( $x_i, U_i^n$ ).

The method requires the resolution of the characteristic equations. If the velocity field  $a$  is constant (independent of  $t$  and  $x$ ), then they are given explicitly:  $X(t^n; t^{n+1}, x_i) = x_i - a(t^{n+1} - t^n)$ . If the velocity field  $a$  is nonconstant, this requires to integrate the characteristic equations with a numerical solver.

Regarding the reconstruction operator  $\Pi$ , several methods have been studied in dimension 1 (see e.g. [13]): it can be either a global interpolation operator using cubic splines (with accuracy  $\mathcal{O}(\Delta x^3)$ ) or a local Lagrange polynomial interpolation with  $2r + 1$  neighboring points (with accuracy  $\mathcal{O}(\Delta x^{2r+1})$ ):

$$[\Pi U]_{|[x_i, x_{i+1}]} = \pi((x_j, U_j)_{i-r \leq j \leq i+r}).$$

Convergence has been proven on uniform meshes [7, 21]. The main advantage of the semi-Lagrangian method is that the time step is not subject to a CFL stability condition. Consequently, the choice of the time step is dictated solely by the precision required. They are also conservative for constant advection on a uniform mesh [13]. We refer to [22] for comparisons.

In higher dimensions, the method becomes computationally more expensive. This is why, semi-Lagrangian methods are often used with dimensional splitting on uniform Cartesian meshes. Splitting errors can, however, accentuate the problem due to the non-conservative nature of the method [47]. For non-uniform meshes (in dimension 1)

and unstructured meshes (in higher dimensions), stability issues also arise when using Lagrange interpolation. Several variants of the semi-Lagrangian scheme have been proposed to ensure either the stability and the conservativity: e.g. conservative semi-Lagrangian (PFC) method [23] or semi-Lagrangian Discontinuous Galerkin (SLDG) methods (see Box 4).

We refer to [46, 16] for a more detailed presentation of the method.

### Box 3: Discontinuous Galerkin method

Let us consider a hyperbolic system:

$$\partial_t U + \nabla \cdot F(U) = 0,$$

with  $F(U) = [F_1(U), \dots, F_d(U)]$ . We recall that it is hyperbolic if for any  $n \in \mathbb{R}^d$ ,  $\sum_{k=1}^d n_k \nabla F_k(U)$  is diagonalizable with real eigenvalues. Let us consider a mesh of the domain  $\Omega_x$ . In each cell  $L$  of the mesh, the solution is approximated by polynomial basis functions

$$U_L(x, t) = \sum_{j=1}^q U_{L,j}(t) \psi_{L,j}(x), \quad x \in L.$$

where  $q$  is the number of degrees of freedom per cell. The approximated solution satisfies the discontinuous Galerkin (DG) scheme

$$\int_L \partial_t U_L \psi_{L,i} - \int_L F(U_L) \nabla \psi_{L,i} + \int_{\partial L} \tilde{F}(U_L, U_R, n_{LR}) \psi_{L,i} = 0,$$

where  $R$  denotes cells adjacent to  $L$ ,  $n_{LR}$  is the unit normal vector on  $\partial L$  oriented from  $L$  to  $R$ ,  $\tilde{F}(U_L, U_R, n)$  is a standard finite-volume numerical flux like the local Lax-Friedrichs flux:

$$\tilde{F}(U_L, U_R, n) = \left( \frac{F(U_L) + F(U_R)}{2} \right) n - \frac{\alpha}{2} (U_R - U_L),$$

where  $\alpha = \max_k (|\lambda_k(U_L, n)|, |\lambda_k(U_R, n)|)$  is a local upper bound of the characteristic speeds of the system.

The **nodal Discontinuous Galerkin method** consists into considering a specific polynomial basis made of interpolant polynomials at some given points  $(x_{L,j})$ :  $\psi_{L,j}(x_{L,i}) = \delta_{ij}$ . Therefore, the degrees of freedom  $U_{L,j}$  coincides with the nodal values of the approximate solutions at these points. In addition, the integrals are approximated by Gauss quadrature and then  $(x_{L,j})$  are Gauss points.

For linear systems, the semidiscrete nodal DG scheme is  $L^2$  dissipative as soon as the interface flux is  $L^2$  dissipative too. This also holds for scalar equations if the volume integrals are computed exactly. When using quadrature methods for the integrals, the method is no longer stable a priori. Entropic stability can however be recovered using a split formulation of the method [26]. This has also been extended to a few systems with general entropic functions. Note that obtaining general entropy stability results for the fully discrete DG scheme is still an open problem. We refer to [31] for more details.

### 3 Fluid models with an entropic closure

Another possible reduction is made by considering the dynamics of some specific moments of the distribution function  $f(x, v, t)$ :

$$U(x, t) = \int_{\Omega_v} m(v) f(x, v, t) dv, \quad \forall k \in \llbracket 1, K \rrbracket,$$

where  $m(v) = (m_j(v))_j$  denotes polynomial functions. In order to close the model satisfied by this parameter, we use a reduced form of the distribution function. Here we explore the entropy minimization method to define the distribution function as functions of its moments:

$$M_U(v) = \underset{\substack{g \text{ s.t. } g(v) \geq 0, \\ \int_{\Omega_v} m(v) g(v) dv = U.}}{\operatorname{argmin}} \int_{\Omega_v} g(v) \ln g(v) dv.$$

Using this ansatz, we can obtain the reduced dynamics on the moments  $U$  and this naturally leads to a hyperbolic system, as introduced by Levermore in [38]. This method has been first developed for radiative transport applications [17]. During the PhD thesis of Pierre Gerhard, co-supervised with Cedric Foy (Cerema) and Philippe Helluy, we apply the methodology to construct reduced models for room acoustics [27].

**Acoustic kinetic model.** Acoustics in the audible domain and at the building scale is based on the propagation of pressure waves at the speed of sound  $c = 343 \text{ms}^{-1}$  with frequencies between 20 Hz and 20 kHz. In this regime, acoustic waves can be approximated by the transport of acoustic particles, also called phonons [44]. Each particle carries a quantity of acoustic energy and travels in straight lines at the constant sound speed  $c$ . They also have reflections on the boundary. Here the variable  $v$  will refer to their directions of propagation: they belong to the unit sphere  $\Omega_v = \mathbb{S}^{d-1} = \{v \in \mathbb{R}^d, |v| = 1\}$ . The distribution function  $f(x, v, t)$  thus satisfies the free transport equation:

$$\partial_t f(x, v, t) + c v \cdot \nabla_x f(x, v, t) = 0, \quad \text{in } \Omega_x \times \mathbb{S}^{d-1} \times \mathbb{R}^+. \quad (7)$$

complemented with the following boundary conditions:

$$f(x, v, t) = (1 - \alpha) \left[ \beta f^s(x, v, t) + (1 - \beta) f^d(x, v, t) \right], \quad \forall (v, t) \in \mathbb{S}^-(x) \times \mathbb{R}^+, \quad (8)$$

where  $\alpha \in [0, 1]$  is the absorption coefficient,  $\beta \in [0, 1]$  the so-called accommodation coefficient and where  $f^s$  and  $f^d$  denote the specular and diffusive distribution function of incoming velocities:

$$f^s(x, v, t) = f(x, v - 2(v \cdot n(x))n(x), t), \quad (9)$$

$$f^d(x, v, t) |v \cdot n(x)| = \int_{v' \in \mathbb{S}^+(x)} P_x(v', v) f(x, v', t) |v' \cdot n(x)| dv', \quad (10)$$

and  $P_x(v', v)$  the reflection probability. For any boundary points  $x \in \partial\Omega_x$ , we denote by  $\mathbb{S}^+(x)$  (resp.  $\mathbb{S}^-(x)$ ) the half sphere of outgoing velocities (resp. incoming velocities):  $\mathbb{S}^\pm(x) = \{v \in \mathbb{S}^{d-1} \mid \pm v \cdot n(x) \geq 0\}$ , where  $n(x)$  denotes the outward normal.

**Entropic reduced model.** When considering  $m(v) = (1, v_1, \dots, v_d)^T$ , the macroscopic unknown are denoted  $U(x, t) = (w(x, t), I_1(x, t), \dots, I_d(x, t))^T \in \mathbb{R}^{d+1}$ : it is composed of the energy density  $w(x, t)$  and the intensity direction  $I(x, t)$ :

$$w(x, t) = \int_{\mathbb{S}^{d-1}} f(x, v, t) dv, \quad I(x, t) = \int_{\mathbb{S}^{d-1}} v f(x, v, t) dv. \quad (11)$$

Multiplying Equation (7) by  $m(v)$  and integrating leads to the following fluid model:

$$\partial_t w(x, t) + c \nabla_x \cdot I(x, t) = 0, \quad (12)$$

$$\partial_t I(x, t) + c \nabla_x \cdot P(x, t) = 0, \quad (13)$$

where  $P(x, t)$  is the so-called pressure tensor:

$$P(x, t) = \int_{\mathbb{S}^{d-1}} (v \otimes v) f(x, v, t) dv.$$

To close this system, we define the distribution function  $f_{(w, I)} = M_U$  that solves the optimization problem (3). Then we define the approximate pressure tensor as the one associated with this distribution function:

$$\bar{P}(w, I) = \int_{\mathbb{S}^{d-1}} v \otimes v f_{(w, I)}(x, v, t) dv. \quad (14)$$

Since the velocity domain is compact, the optimization problem has actually a unique solution [SCHNEIDER2004] as soon as the moment vector  $U$  belongs to the so-called realizable set

$$\mathcal{U} = \{U = (w, I) \in \mathbb{R}_+ \times \mathbb{R}^d, |I| < cw\}.$$

Thus we can state the following proposition:

**Proposition 3.** *Let  $(w, I) \in \mathcal{U}$ . Then problem (3) has a unique solution given by:*

$$f_{(U, I)}(v) = w \frac{\exp\left(\kappa \frac{I \cdot v}{|I|}\right)}{Z_\kappa}, \quad (15)$$

where  $Z_\kappa$  is a normalization constant and  $\kappa \geq 0$  is given by:  $\kappa = G^{-1}(r)$  with  $r = |I|/(cw)$ . The two quantities  $Z_\kappa$  and  $G(\kappa)$  are given by:

$$\begin{aligned} (d=2) \quad Z_\kappa &= I_0(\kappa), & G(\kappa) &= \frac{I_1(\kappa)}{I_0(\kappa)}, \\ (d=3) \quad Z_\kappa &= \frac{4\pi \sinh \kappa}{\kappa}, & G(\kappa) &= \coth \kappa - \frac{1}{\kappa}, \end{aligned}$$

where  $I_n(\kappa)$  denotes the Bessel functions of the first kind:  $I_n(\kappa) = \int_0^{2\pi} \cos(n\theta) \exp(\kappa \cos \theta) d\theta / (2\pi)$ .

**Proposition 4** ( $M_1$  model). *The reduced model writes*

$$\partial_t w + c \nabla_x \cdot I = 0, \quad (16)$$

$$\partial_t I + c \nabla_x \cdot \bar{P}(w, I) = 0. \quad (17)$$

where the pressure tensor (14) associated to  $f_{(w, I)}(v)$ :

$$\bar{P}(w, I) = w \left[ \frac{(1 - \chi(r))}{d-1} \mathbf{Id} + \frac{(d\chi(r) - 1)}{d-1} \frac{I \otimes I}{I^2} \right], \quad (18)$$

where  $\mathbf{Id}$  is the identity matrix and  $\chi(r)$  the Eddington factor,  $\chi(r) = 1 - (d-1)r/\kappa$ , with  $r = |I|/(cw)$ . By construction, this is a hyperbolic system, equipped with a convex entropy  $H(f_{w, I})$  [38].

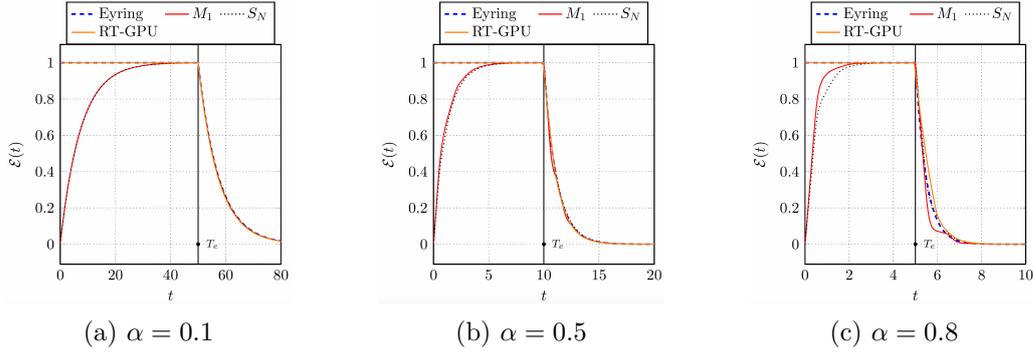


Figure 2: Time evolution of the total energy in a two-dimensional room  $[0, 5] \times [0, 1]$  for accommodation parameter  $\beta = 0.7$  and different absorption parameters  $\alpha$ . Parameters for the  $M_1$ ,  $S_N$  ( $N = 64$  velocities):  $\Delta x = \Delta y = 0.002$ , CFL = 0.5,  $t = 0.00025$ . Parameter RT:  $10^7$  particles. [27].

The one-dimensional version of the model has been studied in [11, 10] and showing that the Riemann problem can be solved globally. The extension of this result in the two-dimensional case has been studied in [27]. Higher order models, using more moments, can be derived following the same strategy. However, the optimization problem (3) has no longer explicit solutions and, in practice, optimization algorithms have to be used to solve the entropy minimization problem. Moreover, the realizable set is not fully identified in these cases. We refer to [1, 43] for recent developments to overcome these difficulties.

**Numerical discretization.** In [27], the reduced model has been solved using a standard finite volume scheme with a Lax-Friedrich flux for internal edges and a kinetic flux for boundary edges.

$$F^{*,b}(U_L, I_L, n) = \int_{\substack{v \in \mathbb{S}^1 \\ v \cdot n > 0}} f_{U_L, I_L}(v) m(v) (v \cdot n) dv \quad (19)$$

$$+ \int_{\substack{v \in \mathbb{S}^1 \\ v \cdot n < 0}} (1 - \alpha) \left[ \beta f_{U_L, I_L}^s(x, v, t) + (1 - \beta) f_{U_L, I_L}^d(x, v, t) \right] (v) m(v) (v \cdot n) dv.$$

The implementation has been done using a PyOpenCL python code to perform simulations on GPUs. Following [18], the Eddington factor has been approximated by a rational fraction to gain in performance.

**Numerical results.** The  $M_1$  reduced model has been compared with two other methods: a ray-tracing algorithm (RT) and the discrete-ordinate method ( $S_N$ ). The latter consists into solving  $N$  independent transport equations corresponding to  $N$  discrete velocities, taken regularly spaced on  $\mathbb{S}^1$  and as Lebedev points on  $\mathbb{S}^2$  [36]. A comparison between the models have been performed to assess their ability to recover the right acoustic energy decay after removing a sound source. The total acoustic energy is defined as the integral of  $w$  over the spatial domain:  $\mathcal{E}(t) = \int_{\Omega_x} w(t, x) dx$ . The results can be compared with the Eyring law, which is valid for cubic rooms. Figure 2 shows that both  $M_1$  and  $S_N$  model (with  $N = 64$  velocities) capture quite well this energy decay, the  $M_1$  model being less computationally demanding than the  $S_N$  one.

## 4 Fluid models with a data-driven closure

In this section, we investigate another way to design fluid models. We directly look for the dynamics of the moments of the distribution function:

$$U(x, t) = \int_{\Omega_v} m(v) f(x, v, t) dv,$$

by using a data-driven closure. Hence, instead of relying on a reduced form of the distribution function  $f(x, v, t)$ , the reduced model is learned from data coming from numerical simulations. This would help into extending the regime of validity of the reduced models. The following work has been carried out during the PhD thesis of Léo Bois, whom I co-supervised with Emmanuel Franck and Vincent Vigon.

**Vlasov-Poisson-BGK model.** In this work, we were interested in the Vlasov-Poisson-BGK equation:

$$\begin{aligned} \partial_t f(t, x, v) + v \cdot \nabla_x f(t, x, v) - E(t, x) \cdot \nabla_v f(t, x, v) &= \frac{1}{\varepsilon} (\mathcal{M}(f)(t, x, v) - f(t, x, v)), \\ \nabla_x \cdot E(t, x) &= \rho(t, x), \end{aligned}$$

where  $\varepsilon > 0$  denotes the Knudsen number, that is the mean free path between two collisions divided by the characteristic length under consideration, and the right-hand side denotes the simplified BGK (Bathnagar-Gross-Krook) collisional operator. This is a relaxation operator towards the local Maxwellian distribution:

$$\mathcal{M}(f) = \frac{\rho(x, t)}{(2\pi T(x, t))^{d/2}} e^{-\frac{|v-u(x,t)|^2}{2T(x,t)}},$$

where the density  $\rho(x, t)$ , the fluid velocity  $u(x, t)$  and the temperature  $T(x, t)$  are defined by:

$$\rho(x, t) = \int_{\mathbb{R}^d} f(x, v, t) dv, \quad (20)$$

$$\rho(x, t)u(x, t) = \int_{\mathbb{R}^d} v f(x, v, t) dv, \quad (21)$$

$$\rho(x, t)T(x, t) = \frac{1}{d} \int_{\mathbb{R}^d} |v - u(x, t)|^2 f(x, v, t) dv. \quad (22)$$

**Fluid model.** Classical fluid models look at the dynamics of the first three moments in velocity of the distribution function are the density  $\rho(x, t)$ , the momentum  $\rho(x, t)u(x, t)$  and the energy  $w(x, t)$ :

$$w(x, t) = \int_{\mathbb{R}^d} \frac{|v|^2}{2} f(x, v, t) dv,$$

which can be related to the temperature through the relation:  $w = \rho|u|^2/2 + (d/2) \rho T$ . The dynamics of these first three moments are expected to be enough to capture the dynamics at least in the strong collisional regime when the distribution function is close to the manifold of local Maxwellian distribution. We are therefore interested in deriving their dynamics. The moment method consists into multiplying the Vlasov equation against 1,  $v$ ,  $|v|^2/2$  and then integrating in velocity. This results into the following system:

$$\begin{aligned} \partial_t \rho + \nabla_x \cdot (\rho u) &= 0, \\ \partial_t \rho u + \nabla_x \cdot (\rho u \otimes u) + \nabla_x p + \nabla_x \cdot \Pi &= -\rho E, \\ \partial_t w + \nabla_x \cdot [w u + p u + \Pi u + q] &= -E \rho u, \end{aligned}$$

where  $p(x, t) = \rho(x, t)T(x, t)$  denotes the pressure and  $\Pi(x, t) \in M_d(\mathbb{R})$  is the traceless part of the pressure tensor and  $q(x, t) \in \mathbb{R}^d$  is the heat flux, given by:

$$\begin{aligned}\Pi(x, t) &= \left( \int_{\mathbb{R}^d} (v - u(x, t)) \otimes (v - u(x, t)) f(x, v, t) dv \right) - p(x, t) \text{Id}, \\ q(x, t) &= \int_{\mathbb{R}^d} v |v - u(x, t)|^2 f(x, v, t) dv.\end{aligned}$$

We thus get a system for the three moments  $\rho$ ,  $\rho u$  and  $w$  but this system is not closed as  $\Pi$  and  $q$  still depend on the underlying kinetic distribution  $f$ .

The first possibility to close the system would be to provide an expression for the distribution function as a function of its first three moments. Using a Chapman-Enskog analysis, we can formally determine the following approximation:

$$f = \mathcal{M}(f) - \varepsilon (\mathcal{A} : \sigma(u) + \mathcal{B} \cdot \nabla T) M(f) + \mathcal{O}(\varepsilon^2),$$

and then deduce the following Navier-Stokes closure (in dimension  $d = 3$ )

$$\begin{aligned}\hat{\Pi} &= -\varepsilon p \sigma(u) + \mathcal{O}(\varepsilon^2), \quad \text{with } \sigma(u) = \left( \nabla_x u + \nabla_x u^T \right) - (\nabla \cdot u) \text{Id}. \\ \hat{q} &= -\varepsilon \frac{3}{2} p \nabla T + \mathcal{O}(\varepsilon^2).\end{aligned}$$

However, these relations are only valid for small Knudsen numbers,  $\varepsilon \leq 10^{-2}$ . We refer to [15] for a review on this topic.

**Data-driven reduced model.** Another possible way to determine the dynamics is to learn a closure relation from data, coming from physical measures or numerical simulations. In [Red3], we have proposed a supervised strategy to learn a nonlocal closure relation of the form:

$$(\Pi(\cdot, t), q(\cdot, t)) = \mathcal{C}_{\hat{\theta}}(\varepsilon, \rho(\cdot, t), u(\cdot, t), T(\cdot, t)),$$

where the closure function is defined by a neural network with parameters  $\hat{\theta} \in \Theta$ . Kinetic numerical simulations for various initial conditions and Knudsen parameters  $\varepsilon \in [10^{-2}, 1]$  have been performed and have provided a collection of data composed of tuples,  $(\varepsilon, \rho_i, u_i, T_i, \Pi_i, q_i)_{i=1, \dots, N}$ . Then the neural network has been trained such as to reach a local minima of the following function:

$$\hat{\theta} \in \underset{\theta \in \Theta}{\text{argmin}} \sum_{i=1}^N |C_{\theta}(\varepsilon_i, \rho_i, u_i, T_i) - (\Pi_i, q_i)|^2$$

Regarding the architecture of the neural network, a V-net architecture has been employed: it makes multi-scale analysis of the signals thanks to a succession of convolution layers and up and down-sampling layers [39]. Specific data post-processing has been carried out to avoid the prediction of two small values, using a normalization thanks to the Navier-Stokes closure.

**Stability issues.** While the learning process can be made with success, troubles can arise while using the obtained closure into a fluid solver. Indeed, there is no stability guarantee and numerical instabilities develop. In the one-dimensional case [Red3], a spatial smoothing of the neural network output, with a convolution with a Gaussian kernel, proves to be sufficient to stabilize the numerical methods, while still preserving the prediction accuracy. In the two-dimensional case, however, this process is no longer enough. Several possible numerical treatments have been tested (smoothing, increased numerical viscosity, restricted range of Knudsen number, pressure thresholding) but none stabilizes without deteriorating significantly the results. We refer to [3] for more details.

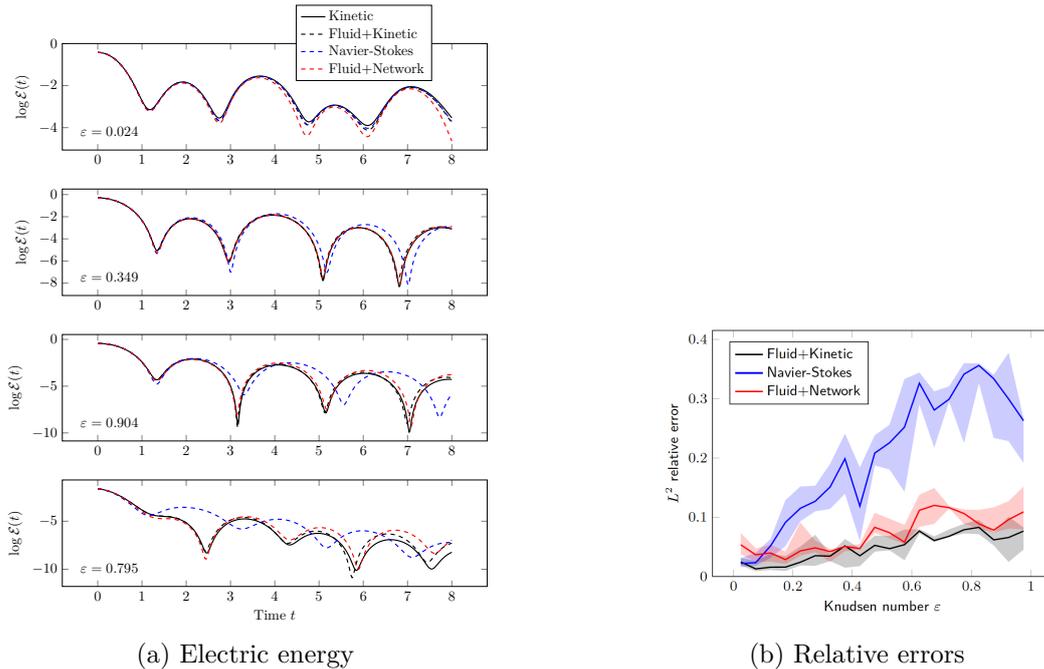


Figure 3: Left: Examples of the evolution of the electric energy with different initial conditions and different Knudsen numbers. Right:  $L^2$  relative errors of the fluid models over the kinetic model depending on the Knudsen number. It shows the median and interquartile interval for 20 uniform classes of Knudsen numbers between 0.01 and 1. Taken from [Red3].

**Implementation issues.** The neural network and the numerical simulations have been implemented using TensorFlow. While the use of the neural network in a fluid solver is not competitive with the kinetic numerical simulations in the one-dimensional case, an acceleration factor of order 7 is obtained in the two-dimensional case. This is mostly due to the fact the neural network approximation capabilities is not greatly affected by the dimension of the problem: neural networks have approximately the same number of parameters in the two- and three-dimensional problem.

**Numerical results.** Figure 3a shows electric energy time evolution obtained with the model with the neural network closure (Fluid+Network), for different initial conditions and different Knudsen number, while Figure 3b presents statistical results. The proposed model is compared with the Navier-Stokes model and the Fluid+Kinetic model, which uses the true heat flux obtained from the kinetic simulation. The latter is exact at the continuous level, but already leads to some error once discretized. The main observation is that the Fluid+Network model is actually able to capture the right dynamics on the considered Knudsen interval, with uniform relative error below 10%.

## 5 Fluid models with a data-driven velocity basis

Reduced models have also been constructed by extracting information from the solutions themselves. More precisely, in [Red6], we adopt a Reduced Order Modeling (ROM) approach, where the reduced variables are linear combinations of the nodal values:

$$\hat{U}(x, t) = \Phi^T [f(x, v_j, t)]_{j=1, \dots, N_v}$$

where  $(v_j)_{j=1,\dots,N_v}$  are velocity grid points and  $\Phi^T \in M_{N_v, K}(\mathbb{R})$  is a linear encoder operator and  $\Phi \in M_{K, N_v}(\mathbb{R})$  is the associated linear decoder operator: they satisfy the decoding-encoding property,  $\Phi^T \Phi = \text{Id}$ . This linear operator is obtained as the solution of the following problem:

$$\min_{\substack{\Phi \in M_{N_v, K}(\mathbb{R}) \\ \Phi^T \Phi = \text{Id}}} \sum_i |f_i - \Phi \Phi^T f_i|^2,$$

where  $f_i = (f_i(x_i, v_j, t_i))_{j=1,\dots,N_v}$  are  $N_s$  samples of the discrete distribution functions on the velocity grid. The optimization problem thus consists into looking for the decoding operator  $\Phi$  so as to minimize the encoding-decoding error on the samples. The solution to this minimization problem is obtained by performing a Proper Orthogonal Decomposition of the matrix  $(f_{i,j})$ . With this technique, the goal is to determine a reduced model valid around a reduced set of solutions but with the smallest possible number of parameters  $K$ .

Note that such a method has been applied for particle discretization of the Vlasov-Poisson system [30] with a reduction in both space and velocity. In order to efficiently tackle Eulerian discretization in the six-dimensional phase space, low-rank tensor bases have been proposed [6, 19], as well as their dynamic version [29]. Here we propose a mixed method where the ROM approach only compresses the dynamics in the velocity variable. Like moment models, the resulting reduced model provides the spatial dynamics of the reduced quantities. Therefore, the method provides a model, independent of the spatial discretization of the computational domain.

**Vlasov-Poisson Fokker-Planck model.** We have considered the one-dimensional Vlasov-Poisson Fokker-Planck model:

$$\partial_t f(t, x, v) + v \partial_x f(t, x, v) + E(t, x) \partial_v f(t, x, v) = \frac{1}{\varepsilon} Q(f)(t, x, v), \quad (23)$$

where  $\varepsilon > 0$  still denotes the Knudsen number and  $Q(f)$  denotes the non-linear Fokker-Planck operator:

$$Q(f) = \partial_v \left( (v - u_f) f + T_f \partial_v f \right),$$

where  $\rho(t, x)$ ,  $u(t, x)$  and  $T(t, x)$  denote the particle density (in space), the velocity, and the temperature, as defined in (20)-(21)-(22).

**Reduced model.** Starting from the semi-discrete model (6) with finite element approximation of order  $p = 1$ , the nodal values  $U(x, t) = (f(x, v_j, t))$  satisfies the following equation:

$$\partial_t U(t, x) + \text{Diag}(v_j) \partial_x U(t, x) + E[U](t, x) D U(t, x) = \frac{1}{\varepsilon} Q_{\Delta v}(U(t, x)), \quad (24)$$

where  $D$  refers to the matrix associated with the centered finite difference approximation and  $E[U]$  refers to the resolution of the Poisson equation from the discrete solution  $U$ . The operator  $Q_{\Delta v}(U(t, x))$  is a discretization of the Fokker-Planck operator, like the one proposed in [4], which is explicit, density, momentum and energy preserving and entropy dissipative. Then to obtain the dynamics of the reduced variables  $\hat{U}$ , we perform a Galerkin projection. We insert the ansatz  $U = \Phi \hat{U}$  into the equation and apply the projection  $\Phi^T$ . Using the identity  $\Phi^T \Phi = \text{Id}$ , we get the following reduced model:

$$\partial_t \hat{U}(t, x) + \widehat{\text{Diag}}(v) \partial_x \hat{U}(t, x) + E[\Phi \hat{U}] \hat{D} \hat{U}(t, x) = \frac{1}{\varepsilon} \hat{Q}_{\Delta v}(\hat{U}), \quad (25)$$

where matrices  $\widehat{\text{Diag}}(v)$ ,  $\hat{D}$  are respectively of size  $K \times K$ ,  $K \times K$  and  $1 \times K$  and given by:

$$\widehat{\text{Diag}}(v) = \Phi^T \text{Diag}(v) \Phi, \quad \hat{D} = \Phi^T D \Phi,$$

and the reduced collision operator is defined by:

$$\widehat{Q}_{\Delta v}(\hat{U}) = \Phi^T Q_{\Delta v}(\Phi \hat{U}(t, x)). \quad (26)$$

**Preservation of moments.** Defining the reduced moment operators:

$$\hat{\mathbf{m}} = \begin{bmatrix} \hat{\mathbf{m}}_\rho \\ \hat{\mathbf{m}}_{\rho u} \\ \hat{\mathbf{m}}_w \end{bmatrix} = \begin{bmatrix} \Delta v \mathbf{1}^T \\ \Delta v \mathbf{1}^T \Phi (\Phi^T \text{Diag}(\mathbf{v}) \Phi) \\ \frac{1}{2} \Delta v \mathbf{1}^T \Phi (\Phi^T \text{Diag}(\mathbf{v}) \Phi)^2 \end{bmatrix} \in M_{3,K}(\mathbb{R}). \quad (27)$$

the following proposition holds

**Proposition 5.** *The reduced moments  $(\hat{\rho}, \hat{\rho} \hat{u}, \hat{w})^T = \hat{\mathbf{m}} \hat{U}$  satisfy the following system:*

$$\begin{cases} \partial_t \hat{\rho} + \partial_x(\hat{\rho} \hat{u}) = E[\Phi \hat{U}] \hat{\mathbf{m}}_\rho \hat{D} \hat{U} + \frac{1}{\varepsilon} \hat{\mathbf{m}}_\rho \widehat{Q}_{\Delta v}(\hat{U}), \\ \partial_t(\hat{\rho} \hat{u}) + \partial_x(\hat{\rho} \hat{u}^2 + \hat{p}) = E[\Phi \hat{U}] \hat{\mathbf{m}}_{\rho u} \hat{D} \hat{U} + \frac{1}{\varepsilon} \hat{\mathbf{m}}_{\rho u} \widehat{Q}_{\Delta v}(\hat{U}), \\ \partial_t \hat{w} + \partial_x(\hat{w} \hat{u} + \hat{p} \hat{u} + \hat{q}) = E[\Phi \hat{U}] \hat{\mathbf{m}}_w \hat{D} \hat{U} + \frac{1}{\varepsilon} \hat{\mathbf{m}}_w \widehat{Q}_{\Delta v}(\hat{U}), \end{cases} \quad (28)$$

where  $\hat{p}$  is such that  $\hat{w} = \hat{\rho} \hat{u}^2 / 2 + \hat{p} / 2$  and  $\hat{q} = \frac{1}{2} \Delta v \mathbf{1}^T \Phi (\hat{A} - \hat{u} \text{Id})^3 \hat{U}$ .

From this proposition, we see that if  $\widehat{Q}_{\Delta v}$  preserved the reduced moments, then the singular terms would vanish and the reduced model would lead to a close fluid model. To enforce this property, a classical method consists into correcting the collision operator in the following way:

$$\widehat{Q}_{\Delta v}^c(\Phi \hat{U}) = \underset{\hat{Q} \in \mathbb{R}^K \text{ s.t. } \hat{\mathbf{m}} \hat{Q} = 0}{\text{argmin}} \quad \|\Phi^T \hat{Q} - Q_{\Delta v}(\Phi \hat{U})\|^2.$$

The solution to this minimization problem is given by:

$$\widehat{Q}_{\Delta v}^c(\Phi \hat{U}) = \Phi^T Q_{\Delta v}(\Phi \hat{U}) - \hat{\mathbf{m}}^T (\hat{\mathbf{m}} \hat{\mathbf{m}}^T)^{-1} \hat{\mathbf{m}} \Phi^T Q_{\Delta v}(\Phi \hat{U}). \quad (29)$$

**Numerical results.** We consider the Landau damping test case with initial distribution

$$f(0, x, v) = \frac{1}{\sqrt{2\pi}} \exp(-v^2/2) (1 + \alpha \cos(kx)), \quad x \in [0, 2\pi/k], \quad v \in [-v_{\max}, v_{\max}],$$

with  $k = 0.5$ ,  $v_{\max} = 6$  and the aim is to build a reduced model valid for all values in  $(\varepsilon, \alpha)$  in the domain  $\mathcal{D} = [1, 10] \times [0.01, 0.1]$  on the time interval  $[0, 20]$ . Therefore, we consider 25 pairs  $(\varepsilon, \alpha)$ , randomly chosen in  $\mathcal{D}$  and the reduced models is built using 30 samples of the distribution function per pairs  $(\varepsilon, \alpha)$ , resulting into  $N_s = 750$  samples. These samples are obtained by solving (24), using a Finite Volume spatial discretization with  $N_x = N_v = 128$ . In Figure 4, we can observe the electric energy damping for three sets of parameters. The first set of parameters belongs to  $\mathcal{D}$  and the approximation results are quite good, with a reduced dimension equal to  $K = 10$ . The reduced dimension  $K$  have been chosen by computing the singular values associated with the samples (red curve on Fig. 4a). When considering parameters outside of the learning domain  $\mathcal{D}$ , the correction introduced in (29) enables the model to capture the right dynamics at least for  $\varepsilon$  smaller than 1 (see Fig. 4c).

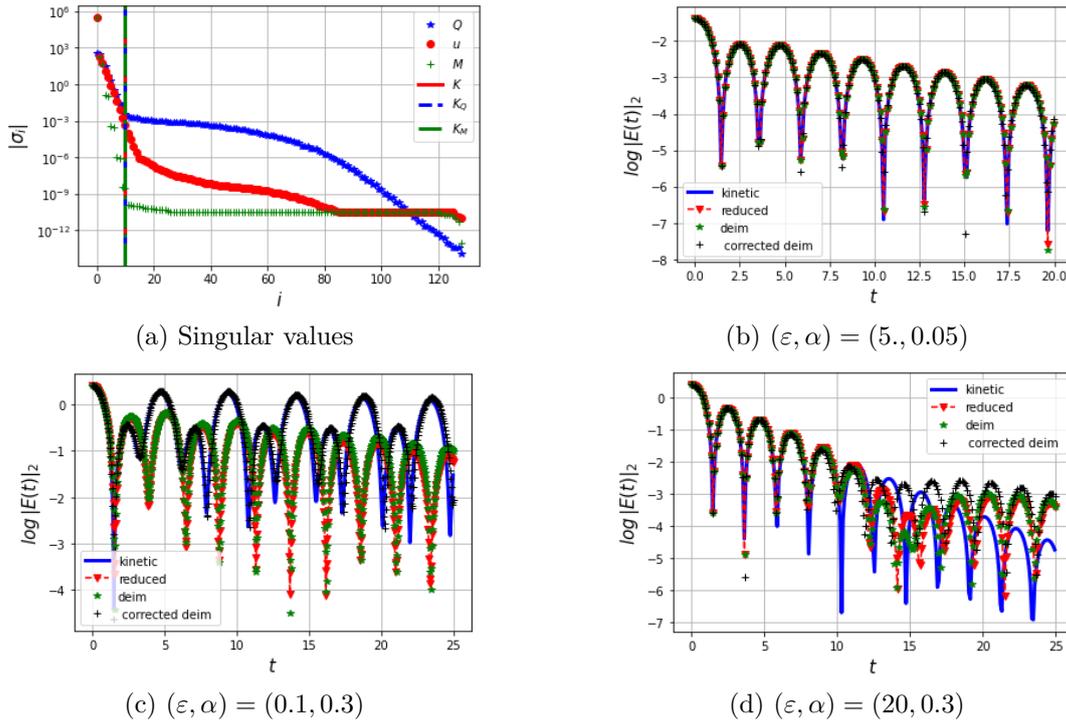


Figure 4: Logarithm of the norm of the electric field. Taken from [Red6].

## 6 Conclusion and perspectives

In this chapter, we have presented several reduction methods for kinetic equations. Reduced models based on a semi-discretization techniques or entropic closures enjoy hyperbolicity properties by construction, but may have a limited range of validity unless considering a large number of reduced variables or restrict ourselves to collisional physical regimes. On the other hand, the proposed data-driven methods enable us to construct reduced models valid in weakly-collisional regimes, with a limited number of reduced variables, but without stability guarantees.

To ensure stability, data-driven reduced models should be constrained to preserve some mathematical structure. This is what we investigated in the recent work [Red1], which proposes a data-driven reduction method that preserves Hamiltonian structure. The reduced variables and the reduced dynamics are learned simultaneously by coupling a convolutional auto-encoder and a Hamiltonian neural network. The work consisted in implementing this approach on large differential systems obtained by semi-discretization in the space of non-linear wave equations. This has been obtained during the ongoing PhD thesis of Guillaume Steimer, co-advised with Raphaël Côte (Univ. Strasbourg), Emmanuel Franck (INRIA, Univ. Strasbourg) and Vincent Vigon (Univ. Strasbourg).

In continuation of these works, we would like to further investigate the construction of reduced models using neural networks. We can list several directions of research:

1. **Reduction of particle dynamics.** Particle systems arising in plasma physics are described by large systems of differential equations with Hamiltonian or Poisson structure. In the reduction program, the first difficulty concerns the identification of reduced vari-

ables. As there is no underlying spatial structure, data-driven reduction technique would require dense auto-encoders that are much more difficult to train. We propose to use a first algebraic reduction by proper orthogonal decomposition, adapted to the symplectic case [40]. Another technique to be studied could be the use of permutation-invariant neural networks [37]. The second issue regards the learning of the reduced dynamics. Specific work needs to be done to extend the Hamiltonian technique to the Poisson case. Finally, in order to reduce the algorithmic complexity of the reduced models obtained by neural networks, SINDY-type models could be considered [9].

2. **Moment models.** Regarding the reduction of kinetic models, the advantage of velocity moment models is that they give the dynamics of variables of physical interest such as density, momentum or energy. In Section 4, we only constructed closures with the minimum number of moments appearing in the classical fluid equations. An interesting extension would be to construct the dynamics of higher-order moments. One approach could be to learn the dynamics of these additional moments using recurrent neural networks. The method could be tested on the Vlasov-Poisson dynamics and could subsequently be extended to the closure of the gyrofluid model [24].
3. **Construction of reduced variables in a strongly oscillating regime.** In the strong magnetic regime, with non-periodic magnetic fields, it is difficult, if not impossible, to analytically figure out a reduced asymptotic model, which would provide the main dynamics of the particles deprived of fast rotation. Neural networks could also be used in this regime to learn appropriate reduced variables. In addition, the possible reconstruction of the fast rotation dynamics would make it possible to add relevant corrections to the slow dynamics model. This would allow us to build valid schemes in different magnitude regimes of the magnetic field [12].
4. **Adaptive models** Reduced models provide a good approximation to solutions, but only in a necessarily restricted regime of parameters. It would therefore be interesting to determine locally whether the reduced model can be used or whether it is necessary to resort to the initial model. To do this, one method would be to build classifiers using neural networks, for which the labels - reduced model or complete model - will be determined using the error between the reduced model and the complete model. We can also consider unsupervised reinforcement learning strategies to ensure better stability. Once this classifier has been trained, it will be necessary to implement model coupling methods, while ensuring the stability and convergence of the methods. In the longer term, it would be interesting to design a method in which the reduced model is built on the fly from the simulations carried out by the full model and where the full model is gradually replaced by the reduced model when this proves relevant.

## References

- [1] G. W. Alldredge, M. Frank, and C. D. Hauck. “A regularized entropy-based moment method for kinetic equations”. In: *SIAM Journal on Applied Mathematics* 79.5 (2019), pp. 1627–1653.
- [2] M. Bessemoulin-Chatard and F. Filbet. “On the convergence of discontinuous Galerkin/Hermite spectral methods for the Vlasov–Poisson system”. In: *SIAM J. Numer. Anal.* 61.4 (2023), pp. 1664–1688.
- [3] L. Bois. “Méthodes numériques basées sur l’apprentissage pour les EDP hyperboliques et cinétiques”. PhD thesis. Université de Strasbourg, 2023. ([tel-04589397v1](tel:04589397v1)).

- [4] Christophe Buet, Stéphane Dellacherie, and Rémi Sentis. “Numerical Solution of an Ionic Fokker–Planck Equation with Electronic Temperature”. In: *SIAM Journal on Numerical Analysis* 39.4 (2001), pp. 1219–1253.
- [5] Z. Cai, Y. Fan, and R. Li. “Globally hyperbolic regularization of Grad’s moment system”. In: *Comm. Pure Appl. Math.* 67.3 (2014), pp. 464–518.
- [6] F. Cassini and L. Einkemmer. “Efficient 6D Vlasov simulation using the dynamical low-rank framework Ensign”. In: *Comput. Phys. Commun.* 280 (2022), p. 108489.
- [7] F. Charles, B. Després, and M. Mehrenberger. “Enhanced convergence estimates for semi-Lagrangian schemes application to the Vlasov–Poisson equation”. In: *SIAM J. Numer. Anal.* 51.2 (2013), pp. 840–863.
- [8] F.F. Chen. *Introduction to plasma physics*. Springer Science & Business Media, 2012.
- [9] P. Conti, G. Gobat, S. Fresca, A. Manzoni, and A. Frangi. “Reduced order modeling of parametrized systems through autoencoders and SINDy approach: continuation of periodic solutions”. In: *Comput. Methods Appl. Mech. Engrg.* 411 (2023), p. 116072.
- [10] J.-F. Coulombel, F. Golse, and T. Goudon. “Diffusion approximation and entropy-based moment closure for kinetic equations”. In: *Asympt. Anal.* 45.1-2 (2005), pp. 1–39.
- [11] Jean-François Coulombel and Thierry Goudon. “Entropy-based moment closure for kinetic equations: Riemann problem and invariant regions”. In: *J. Hyperbolic Differ. Equations* 3.04 (2006), pp. 649–671.
- [12] N. Crouseilles, M. Lemou, F. Méhats, and X. Zhao. “Uniformly accurate Particle-in-Cell method for the long time solution of the two-dimensional Vlasov–Poisson equation with uniform strong magnetic field”. In: *J. Comput. Phys.* 346 (2017), pp. 172–190.
- [13] N. Crouseilles, M. Mehrenberger, and E. Sonnendrücker. “Conservative semi-Lagrangian schemes for Vlasov equations”. In: *J. Comput. Phys.* 229.6 (2010), pp. 1927–1953.
- [14] R. Dai and B. Després. “On the quadratic stability of asymmetric Hermite basis and application to plasma physics”. In: *arXiv preprint arXiv:2405.07811* (2024).
- [15] P. Degond. “Macroscopic limits of the Boltzmann equation: a review”. In: *Modeling and computational methods for kinetic equations* (2004), pp. 3–57.
- [16] G. Dimarco and L. Pareschi. “Numerical methods for kinetic equations”. In: *Acta Numer.* 23 (2014), pp. 369–520.
- [17] Bruno Dubroca and Jean-Luc Feugeas. “Etude théorique et numérique d’une hiérarchie de modèles aux moments pour le transfert radiatif”. In: *Comptes Rendus de l’Académie des Sciences-Series I-Mathematics* 329.10 (1999), pp. 915–920.
- [18] R. Ducloux, B. Dubroca, and M. Frank. “A deterministic partial differential equation model for dose calculation in electron radiotherapy”. In: *Phys. Med. Biol.* 55.13 (2010), p. 3843.
- [19] V. Ehrlacher and D. Lombardi. “A dynamical adaptive tensor method for the Vlasov–Poisson system”. In: *J. Comput. Phys.* 339 (2017), pp. 285–306.
- [20] B. Eliasson. “Outflow boundary conditions for the Fourier transformed two-dimensional Vlasov equation”. In: *J. Comput. Phys.* 181.1 (2002), pp. 98–125.
- [21] R. Ferretti and M. Mehrenberger. “Stability of semi-Lagrangian schemes of arbitrary odd degree under constant and variable advection speed”. In: *Math. Comput.* 89.324 (2020), pp. 1783–1805.
- [22] F. Filbet and E. Sonnendrücker. “Comparison of eulerian vlasov solvers”. In: *Comput. Phys. Commun.* 150.3 (2003), pp. 247–266.
- [23] F. Filbet, E. Sonnendrücker, and P. Bertrand. “Conservative numerical schemes for the Vlasov equation”. In: *J. Comput. Phys.* 172.1 (2001), pp. 166–187.

- [24] B. Frei, A. Hoffmann, and P. Ricci. “A Gyrokinetic moment-based method to model the plasma boundary of fusion devices at arbitrary collisionality”. In: *Bull. Am. Phys. Soc.* (2022).
- [25] C.K. Garrett and C.D. Hauck. “A comparison of moment closures for linear kinetic transport equations: The line source benchmark”. In: *Transp. Theory Stat. Phys.* 42.6-7 (2013), pp. 203–235.
- [26] G.J. Gassner and A.R. Winters. “A novel robust strategy for discontinuous Galerkin methods in computational fluid mechanics: Why? When? What? Where?” In: *Front. Phys.* 8 (2021), p. 500690.
- [27] P. Gerhard. “Réduction de modèles cinétiques et applications à l’acoustique du bâtiment”. PhD thesis. Université de Strasbourg, 2020. ([tel-02445322v3](tel:02445322v3)).
- [28] V. Grandgirard. “The GYSELA project: A semi-Lagrangian code addressing gyrokinetic full-f global simulations of flux driven tokamak plasmas”. PhD thesis. Université de Strasbourg, IRMA, 2016.
- [29] W. Guo, J. Ferdous Ema, and J.-M. Qiu. “A Local Macroscopic Conservative (LoMaC) low rank tensor method with the discontinuous Galerkin method for the Vlasov dynamics”. In: *Commun. Appl Math. Comput.* (2023), pp. 1–26.
- [30] J. Hesthaven, C. Pagliantini, and N. Ripamonti. “Adaptive symplectic model order reduction of parametric particle-based Vlasov–Poisson equation”. In: *Math. Comput.* 93.347 (2024), pp. 1153–1202.
- [31] J.S. Hesthaven and T. Warburton. *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*. Springer Science & Business Media, 2007.
- [32] J.P. Holloway. “Spectral velocity discretizations for the Vlasov–Maxwell equations”. In: *Transp. Theory Stat. Phys.* 25.1 (1996), pp. 1–32.
- [33] M. Junk. “Domain of definition of Levermore’s five-moment system”. In: *J. Stat. Phys.* 93.5-6 (1998), pp. 1143–1167.
- [34] A.J. Klimas and W.M. Farrell. “A splitting algorithm for Vlasov simulation with filamentation filtration”. In: *J. Comp. Phys.* 110.1 (1994), pp. 150–163.
- [35] S. Le Bourdiec, F. De Vuyst, and L. Jacquet. “Numerical solution of the Vlasov–Poisson system using generalized Hermite functions”. In: *Comput. Phys. Comm.* 175.8 (2006), pp. 528–544.
- [36] V.I. Lebedev. “Quadratures on a sphere”. In: *USSR Comput. Math. Math. Phys.* 16.2 (1976), pp. 10–24.
- [37] J. Lee, Y. Lee, J. Kim, A. Kosiorek, S. Choi, and Y.W. Teh. “Set transformer: A framework for attention-based permutation-invariant neural networks”. In: *International conference on machine learning*. PMLR, 2019, pp. 3744–3753.
- [38] C.D. Levermore. “Moment closure hierarchies for kinetic theories”. In: *J. Stat. Phys.* 83 (5-6 1996), pp. 1021–1065.
- [39] F. Milletari, N. Navab, and S. Ahmadi. “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation”. In: *2016 Fourth International Conference on 3D Vision (3DV)*. 2016, pp. 565–571.
- [40] L. Peng and K. Mohseni. “Symplectic model reduction of Hamiltonian systems”. In: *SIAM J. Sci. Comput.* 38.1 (2016), A1–A27.
- [41] N. Pham. “Méthodes numériques pour l’équation de Vlasov réduite”. PhD thesis. Université de Strasbourg, 2016. ([tel-01412750](tel:01412750)).
- [42] N. Pham, P. Helluy, and A. Crestetto. “Space-only hyperbolic approximation of the Vlasov equation”. In: *ESAIM: Proceedings*. Vol. 43. EDP Sciences, 2013, pp. 17–36.

- 
- [43] W. A. Porteous, M. T. P. Laiu, and Cory D Hauck. “Data-driven, structure-preserving approximations to entropy-based moment closures for kinetic equations”. In: *Communications in Mathematical Sciences* 21.4 (2023), pp. 885–913.
  - [44] O. Runborg. “Mathematical models and numerical methods for high frequency waves”. In: *Commun. Comput. Phys* 2.5 (2007), pp. 827–880.
  - [45] J. Schneider. “Entropic approximation in kinetic theory”. In: *Esaim Math Model Numer Anal.* 38.3 (2004), pp. 541–561.
  - [46] E. Sonnendrücker and K. Kormann. “Numerical methods for Vlasov equations”. In: *Lecture notes* 107 (2013), p. 108.
  - [47] E. Sonnendrücker, J. Roche, P. Bertrand, and A. Ghizzo. “The semi-Lagrangian method for the numerical resolution of the Vlasov equation”. In: *J. Comput. Phys.* 149.2 (1999), pp. 201–220.
  - [48] Eitan Tadmor. “Runge-Kutta methods are stable”. In: *arXiv preprint arXiv:2312.15546* (2023).
  - [49] M. Torrilhon. “Modeling nonequilibrium gas flow based on moment equations.” In: *Annual review of fluid mechanics* 48 (2016), pp. 429–458.



---

## Chapter II

# Numerical methods for kinetic and hyperbolic systems

---

## 1 Introduction

Among the large variety of method for hyperbolic systems, the discontinuous Galerkin method, already presented in Box 3, has the advantage to be high-order accurate, compatible with unstructured meshes and have stability properties in the linear case. To overcome the CFL stability condition, a semi-Lagrangian version, called Semi-Lagrangian Discontinuous Galerkin (SLDG), have been proposed for transport equations, which inherit its stability properties. We refer to Box 4 for a short presentation of the method.

In this short chapter, we will explore two very different kinds of numerical issues. The first one is how to analyze the recurrence phenomenon appearing in the simulations of kinetic transport equations with the SLDG scheme. The second one is how to improve the accuracy of the DG method on stationary solutions.

## 2 Recurrence phenomenon on a regular finite element mesh

Recurrence phenomenon is a numerical artifact appearing when solving kinetic transport equations on a regular phase-space grid with periodic boundary conditions, which suddenly destroys the accuracy of the numerical solution. It has been first mentionned in the seminal paper of Cheng and Knorr [8]. When considering periodic boundary conditions, the Vlasov dynamics exhibit phase mixing: small initial spatial perturbations result in oscillations in the velocity distribution, whose frequency grows in time. Consequently, at some time, the velocity mesh cannot capture these oscillations anymore. This leads to an aliasing effect: when charge density is computed, the large frequency oscillations are numerically interpreted as small frequency ones. As the charge density and the electric field are related through the Poisson equation, the recurrence phenomenon is rather observed on the electric energy time evolution.

Several works have considered methods to suppress or at least reduce the recurrence phenomenon. This can be done by adding an artificial diffusion [16], outflow boundary conditions in the Fourier velocity variable [13], using cutoff procedure in Fourier space [12, 11], Filtering techniques for Hermite-spectral method [7]. In another direction, randomization of the velocity points in each velocity cells have been considered in [1].

When using a **uniform finite-element mesh in velocity**, as it can be the case for the SLDG scheme, it has been observed surprisingly that the recurrence phenomenon behaves unfavourably

when increasing the degree of approximation [9, 18, 19, 11]: the recurrence time does not depend on the degree of approximation.

This section briefly presents the analysis of this phenomenon performed and how trigonometric quadrature weights optimally defers the occurrence of the recurrence phenomenon. This was a work made in collaboration with M. Mehrenberger (Aix-Marseille Univ.) and N. Pham (Univ. Strasbourg) and published in [Num4].

**Recurrence phenomenon on a regular mesh.** Let us consider the following kinetic transport problem:

$$\begin{aligned}\partial_t f + v \partial_x f &= 0, \\ f_{\text{init}}(x, v) &= A \cos(kx) \frac{1}{\sqrt{2\pi}} \exp(-v^2/2),\end{aligned}$$

on the domain  $\Omega_x \times \Omega_v = [0, 2\pi/k] \times [-v_{\max}, v_{\max}]$ , with spatial periodic boundary conditions. The solution is given by:

$$f(x, v, t) = f_{\text{init}}(x - vt, v) = \cos(k(x - vt)) \frac{1}{\sqrt{2\pi}} e^{-v^2/2}.$$

Then computing the charge density with rectangle quadrature on a regular mesh  $(v_\ell)_\ell$  with  $\ell \in \llbracket -N_v/2, N_v/2 \rrbracket$ ,  $v_\ell = \ell \Delta v$ ,  $\Delta v = v_{\max}/N_v > 0$ , then the approximated density writes:

$$\tilde{\rho}(x, t) = \sum_{\ell=-N_v/2}^{N_v/2-1} f(t, x, v_\ell) \Delta v = \sum_{\ell=-N_v/2}^{N_v/2-1} \cos(k(x - v_\ell t)) \frac{1}{\sqrt{2\pi}} e^{-v_\ell^2/2} \Delta v.$$

Thus this approximate charge density is **periodic in time**, with period  $T_R = 2\pi/(k\Delta v)$ . This numerical artifact also appears in the numerical simulation of the Vlasov-Poisson system.

**Recurrence phenomenon on a regular finite element mesh.** Considering the following finite element velocity mesh

$$v_{(\ell,j)} = (\ell + \bar{v}_j) \Delta v, \quad \forall (\ell, j) \in \llbracket -N_v/2, N_v/2 - 1 \rrbracket \times \llbracket 0, q \rrbracket$$

with  $\bar{v}_0 < \dots < \bar{v}_q \in [0, 1]$  and  $q \in \mathbb{N}$ . Then, solving the kinetic transport equation consists in solving  $d \times N_v$  transport equations:

$$\begin{aligned}\partial_t f_{(\ell,j)} + v_{(\ell,j)} \partial_x f_{(\ell,j)} &= 0, \\ (f_{\text{init}})_{(\ell,j)}(x) &= \cos(kx) \frac{1}{\sqrt{2\pi}} e^{-\frac{v_{(\ell,j)}^2}{2}},\end{aligned}$$

and the approximate density writes:

$$\tilde{\rho}(x, t) = \sum_{\ell=-N_v/2}^{N_v/2-1} \sum_{j=0}^q \Delta v \omega_j (f_{\text{init}})_{(\ell,j)}(x - v_{(\ell,j)} t) = \frac{\pi}{k} \cos(kx) h(t)$$

with

$$h(t) = \sum_{\ell=-N_v/2}^{N_v/2-1} \sum_{j=0}^q \Delta v \omega_j \cos(kv_{(\ell,j)} t) e^{-\frac{v_{(\ell,j)}^2}{2}},$$

and where  $(\omega_\ell)$  are quadrature weights. To analyze the recurrence effect, we introduce the correlation function:

$$C(t) = \int_{\Omega_x} \tilde{\rho}(x, t) \tilde{\rho}(x, 0) dx = \frac{\pi}{k} h(t) h(0).$$

We then perform a rescaling in time  $t = (2\pi\tilde{t})/(k\Delta v)$  and  $\tilde{C}(\tilde{t}) = C(t)$  and we get the following result:

**Proposition 6.** *We suppose that  $N_v = +\infty$  and the  $(\bar{v}_j)_{j=0,\dots,q}$  are symmetrically distributed with respect to 0.5. For all  $m \in \mathbb{Z}$  and any  $\tau \in \mathbb{R}$ , we have:*

$$\tilde{C}\left(m + \frac{k\Delta v}{2\pi}\tau\right) = \frac{\pi}{k} h(0) \left( \sum_{j=0}^q \omega_j \cos(2\pi m \bar{v}_j) \right) e^{-\frac{k^2\tau^2}{2}} + O(\Delta v), \quad (30)$$

and for  $t \notin \mathbb{Z}$ , we have:  $\tilde{C}(t) = O(\Delta v)$ .

Consequently, the correlation at time  $m(2\pi)/(k\Delta v)$ , with  $m \in \mathbb{N}^*$  equals the numerical error of the quadrature rule on the trigonometric function  $v \mapsto \cos(2\pi mv)$  (since  $\int_0^1 \cos(2\pi vm) dv = 0$ ) up to a  $O(\Delta v)$  term. Note that this correlation does not vanish as  $\Delta v$  goes to 0. From this analysis, we define the recurrence phenomenon as the appearance of local extrema of the semi-discrete correlation function at positive time (in the limit  $\Delta v$  goes to 0). Without additional assumptions on the quadrature rule, this may hold at time  $\tilde{t}_{\text{rec}} = 1$  that is  $T_{\text{rec}} = 2\pi/(k\Delta v)$ . **The first recurrence time thus does not depend on approximation degree  $q$  and thus on the accuracy of the scheme.**

One way to make the recurrence time as large as possible is to consider the trigonometric quadrature rule, which is exact on trigonometric polynomials. Weights  $(\omega_j)$  are then defined such that

$$\forall m \in \llbracket 1, n \rrbracket, \quad \sum_{j=0}^q \omega_j \cos(2\pi m \bar{v}_j) = \int_0^1 \cos(2\pi mv) dv,$$

with  $n = \lfloor q/2 \rfloor$ . Thus using these trigonometric weights allows us to defer the first recurrence time to  $T_{\text{rec}} = (\lfloor q/2 \rfloor + 1)2\pi/(k\Delta v)$ . On regularly distributed points, the trigonometric quadrature is even exact on trigonometric polynomials up to degree  $d$ , which results in a recurrence time equal to  $T_{\text{rec}} = (q + 1)2\pi/(k\Delta v)$ .

**Applications.** Figure (5) shows the time evolution of the electric energy for a Landau damping test case and compares the results obtained with Gauss (left) and trigonometric (right) quadratures. The sudden rises of the electric energy correspond to recurrence events and the black vertical lines indicate the multiple of the recurrence time  $T_{\text{rec}} = 2\pi/(k\Delta v)$ . We first consider a linear Landau damping test case with a perturbation of size  $\alpha = 10^{-6}$  and the SLDG (see Box 4) method with Gauss-Legendre points in velocity. As expected from the analysis, although performed in the transport case only, we observe that the trigonometric weight indeed allows us to defer the recurrence time. Note that, for  $q \geq 4$ , the first recurrence event is almost not visible because of the already very good precision of the Gauss quadrature. For  $q = 2$ , the recurrence effect is actually not totally shifted in time due to non-linear effects.

As experimentally observed for non-linear test cases in [Num4], trigonometric weights have more impact when considering low-order of accuracy. Consequently, rather than a new numerical method, this analysis provide a validation tool for the concerned codes. It can also lead to further developments by combining the use of trigonometric weights with filtering techniques.

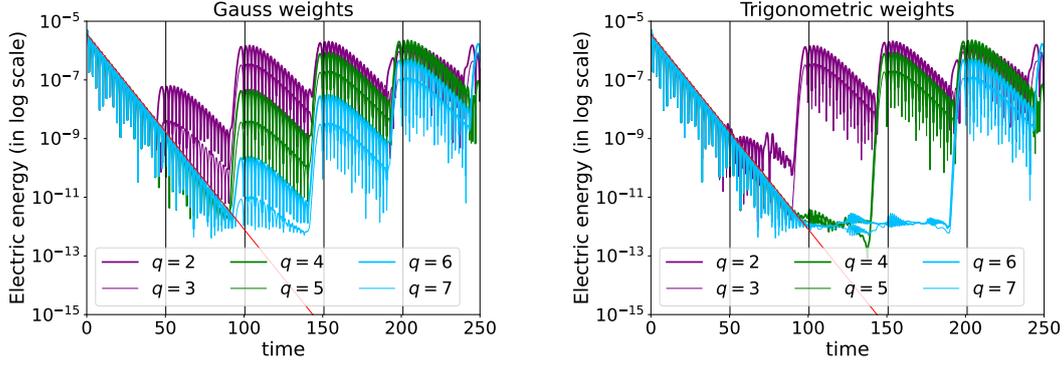


Figure 5: Time evolution of the electric energy  $\mathcal{E}(t)$  (in log scale) when varying  $q$ . The black vertical lines indicate the multiples of the recurrence time  $T_{\text{rec}} = 2\pi/(k\Delta v)$ . The red line is the theoretical damping of the linearized Vlasov-Poisson system. Parameters:  $N_x = N_v = 64$ ,  $v_{\text{max}} = 8$ , Gauss-Legendre points in space and velocity.

#### Box 4: Semi-Lagrangian Discontinuous Galerkin method

In this box, we briefly present a Semi-Lagrangian Discontinuous Galerkin (SLDG) method introduced in [10, 20, 22]. We restrict ourselves to the one-dimensional constant advection equation:

$$\partial_t U + a \partial_x U = 0.$$

with  $a \in \mathbb{R}$  on the domain  $[x_{\min}, x_{\max}]$ . Considering a subdivision  $(x_\ell)_\ell$  of the domain and a time step  $\Delta t > 0$ , the approximate solution at time  $t^n = n\Delta t$  is expressed as a piecewise polynomial function:

$$U^n(x) = \sum_{j=0}^q U_{\ell,j}^n \psi_{\ell,j}(x), \quad \forall x \in [x_\ell, x_{\ell+1}]$$

where  $(\psi_{\ell,j})_{0 \leq j \leq q}$  is a basis of polynomial of order  $q$  in the  $\ell$ -th cell. To compute the approximate solution at time  $t^{n+1}$ , the method consists in first advecting exactly the approximate solution and then taking the  $L^2$  projection on the vector space of piecewise polynomials. Supposing that  $(\psi_{\ell,j})_{\ell,j}$  form an  $L^2$  orthonormal basis, we have:

$$U^{n+1}(x) = \sum_{\ell,j} U_{\ell,j}^{n+1} \psi_{\ell,j}(x),$$

with coefficients given by:

$$U_{\ell,j}^{n+1} = \int \left[ \sum_{\ell',j'} U_{\ell',j'}^n \psi_{\ell',j'}(y - a\Delta t) \right] \psi_{\ell,j}(y) dy.$$

The method is  $L^2$  stable by construction. Moreover, since we have:

$$\int U^{n+1}(x) \psi_{\ell,j}(x) dx = \int U^n(x - a\Delta t) \psi_{\ell,j}(x) dx, \quad \forall \psi_{\ell,j},$$

and the constant 1 belongs to the vector space spanned by  $(\psi_{\ell,j})$ , the scheme is mass conservative. We also refer to [6] for recent developments.

### 3 Prior-basis enrichment for the Discontinuous Galerkin scheme with Physics-Informed Neural Networks

In this section, we present a method to increase the accuracy of the DG scheme on stationary solutions to hyperbolic systems with source terms:

$$\partial_t U + \nabla_x \cdot F(U) = S(x, U). \quad (31)$$

on the domain  $\Omega \subset \mathbb{R}^d$ . The scheme is said **approximately well-balanced**. This is particularly important to capture perturbation dynamics around these equilibria. There is an extensive literature on high order exact or approximately well-balanced schemes (e.g. [25, 2, 15, 14] and references therein): they generally require an additional computational cost compared with traditional schemes. Instead, the method proposed here is quite generic, with almost no additional cost. It is based on a basis enrichment using the knowledge of a prior  $U_\theta(x)$ , which is supposed to be a good approximation of the stationary solution. This is reminiscent of Trefftz methods [17, 4].

This work has been done in collaboration with Emmanuel Franck (INRIA, Univ. Strasbourg) and Victor Michel-Dansac (INRIA, Univ. Strasbourg) and is published in [Num3].

**Basis enrichment.** For the sake of simplicity, let us expose the methodology in the one-dimensional case ( $d = 1$ ). We decompose the domain into cells  $\Omega_k = (x_{k-1/2}, x_{k+1/2})$  of size  $\Delta x_k$  and of centers  $x_k$ . Instead of the nodal DG scheme used in the other sections of this manuscript, we here consider the modal Taylor basis. On each cell  $L_k$ , the polynomial basis writes:

$$V_h = \text{Span}(\psi_{k,0}, \psi_{k,1}, \psi_{k,2}, \dots, \psi_{k,q}) = \text{Span}\left(1, (x - x_k), \frac{1}{2}(x - x_k)^2, \dots, \frac{1}{q!}(x - x_k)^q\right).$$

The main idea is to insert the prior into this approximation basis by considering either

$$V_h^+ = \text{Span}(\psi_{k,0}^+, \psi_{k,1}^+, \psi_{k,2}^+, \dots, \psi_{k,q}^+) = \text{Span}\left(U_\theta(x), (x - x_k), \dots, \frac{1}{q!}(x - x_k)^q\right). \quad (32)$$

or

$$V_h^* = \text{Span}(\psi_{k,0}^*, \psi_{k,1}^*, \psi_{k,2}^*, \dots, \psi_{k,q}^*) = \text{Span}\left(U_\theta(x), (x - x_k)U_\theta(x), \dots, \frac{1}{q!}(x - x_k)^q U_\theta(x)\right). \quad (33)$$

In [24], convergence criteria have been exhibited for a non-polynomial basis and we show in [Num3] that they are satisfied by the two previous bases providing that the prior  $U_\theta$  has differentiable class  $C^{q+1}(\mathbb{R})$ . Then the following proposition provides refined error estimates for the second basis:

**Proposition 7.** *Assume that the prior  $U_\theta$  satisfies  $U_\theta(x)^2 > m^2 > 0$ ,  $\forall x \in \Omega$ . For a given cell  $L_k$ , for any function  $U \in H^{q+1}(\Omega_k)$ , the  $L^2$  projector onto  $V_h^*$ , denoted by  $P_h$  and such that  $P_h(U) \in V_h^*$ , satisfies the inequality*

$$\|U - P_h(U)\|_{L^\infty(\Omega_k)} \lesssim \left| \frac{U}{U_\theta} \right|_{H^{q+1}(\Omega_k)} (\Delta x_k)^{q+\frac{1}{2}} \left( 1 + \frac{\|U_\theta^2\|_{L^\infty(\Omega_k)}}{m^2} \right) \|U_\theta\|_{L^\infty}.$$

We deduce that for any function  $U \in H^{q+1}(\Omega)$ ,

$$\|U - P_h(U)\|_{L^2(\Omega)} \lesssim \left| \frac{U}{U_\theta} \right|_{H^{q+1}(\Omega)} (\Delta x_k)^{q+1} \|U_\theta\|_{L^\infty(\Omega)}.$$

This proposition shows that instead of variations of  $U$ , it is variations of  $U/U_\theta$  that control the error. Therefore, if  $U$  is close to  $U_\theta$ , then the error is expected to be small.

**Prior construction using Physics Informed Neural Networks (PINNs).** The prior  $U_\theta$  should be an approximate solution to the stationary equation:

$$\nabla_x \cdot F(U) = S(x, U)$$

To get this approximation, PINNs look for a solution as a fully connected neural network, with input  $x \in \mathbb{R}$  and output  $U_\theta(x) \in \mathbb{R}^m$ . We refer to [21, 5]. Thus, the solution is a parameterized function, whose parameters  $\theta$  are actually the weights of the neural network. In order to strongly enforce the boundary and initial conditions, the solution  $U_\theta$  is actually written as

$$U_\theta = B(\tilde{U}_\theta),$$

where  $B$  is a boundary operator such as  $U_\theta$  satisfies the boundary and initial conditions. Then the weights are obtained by minimizing the loss function, which is the  $L^2$  norm of the residual:

$$J(\theta) = \|\nabla_x \cdot F(U_\theta) - S(x, U_\theta)\|_{L^2(\Omega)}^2.$$

In practice this  $L^2$  norm is evaluated by a Monte Carlo method. A data loss can be added if samples of the solution are available. As the loss function involves the derivatives of  $U_\theta$ , the PINNs method is expected to provide a good approximation of the stationary solution and its derivatives. This is a key point to benefit from the error estimate of Proposition 7.

The main advantage of PINNs is that they are mesh free and less sensitive to dimension than standard numerical methods. In particular, instead of computing one single equilibrium, we can rather **approximate a family of equilibria indexed by some parameters  $\mu$** .

**Application.** We present some results obtained on the two-dimensional shallow-water equation on the domain  $\Omega = [-3, 3]^2$ . The system writes:

$$\begin{aligned} \partial_t h + \nabla_x \cdot Q &= 0, \\ \partial_t Q + \nabla_x \cdot \left( \frac{Q \otimes Q}{h} + \frac{1}{2} g h^2 \text{Id} \right) &= -g h \nabla_x Z, \end{aligned}$$

with  $h(x, t)$  the water height,  $Q(x, t) = (Q_1(x, t), Q_2(x, t))^T$  the water discharge and  $Z(x, \mu)$  is a parameterized topography, here chosen equal to:

$$Z(x; \mu) = \Gamma \exp(\alpha(r_0^2 - \|x\|^2)).$$

with  $\mu = (\alpha, \Gamma, r_0) \in [0.25, 0.75] \times [0.1, 0.4] \times [0.5, 1.25]$ . A family of radially symmetric stationary solution, denoted  $(h_{\text{eq}}(x; \mu), Q_{\text{eq}}(x, \mu))$ , can be determined for this system. Inspired by its expression, the prior is constructed using the following relations

$$h_\theta(x; \mu) = 2 - Z(x; \mu) \tilde{h}_\theta(x; \mu)^2 \quad \text{and} \quad Q_\theta(x; \mu) = -Z(x, \mu) h_\theta(x; \mu) \tilde{u}_\theta(x; \mu) x^\perp,$$

where  $\tilde{h}_\theta$  and  $\tilde{u}_\theta$  are two neural networks with about 2500 parameters each. These expressions involve the function  $Z$ , that almost vanishes at the boundary of the domain, in order to approximately impose the boundary conditions compatible with the targeted stationary solutions  $(h_{\text{eq}}(\cdot; \mu), Q_{\text{eq}}(\cdot, \mu))$ . The loss data, using the analytical expression, is used to prevent the neural network from learning the lake at rest stationary solutions ( $h + Z = \text{cst}$ ,  $Q = 0$ ). The training takes about 10 minutes on an NVIDIA V100 GPU, until the loss function reaches about  $4 \times 10^{-7}$ .

$K$	$h$ , basis $V_h$		$Q_1$ , basis $V_h$			$Q_2$ , basis $V_h$		
	error	order	error	order	gain	error	order	gain
20	$2.17 \cdot 10^{-2}$	—	$2.58 \cdot 10^{-2}$	—	18.21	$2.58 \cdot 10^{-2}$	—	18.21
40	$5.46 \cdot 10^{-3}$	1.99	$8.88 \cdot 10^{-3}$	1.54	23.99	$8.88 \cdot 10^{-3}$	1.54	23.99
80	$1.37 \cdot 10^{-3}$	2.00	$2.50 \cdot 10^{-3}$	1.83	26.74	$2.50 \cdot 10^{-3}$	1.83	26.74
160	$3.42 \cdot 10^{-4}$	2.00	$6.46 \cdot 10^{-4}$	1.95	27.54	$6.46 \cdot 10^{-4}$	1.95	27.54
320	$8.56 \cdot 10^{-5}$	2.00	$1.62 \cdot 10^{-4}$	2.00	27.55	$1.62 \cdot 10^{-4}$	2.00	27.55

$K$	$h$ , basis $V_h^+$			$Q_1$ , basis $V_h^+$			$Q_2$ , basis $V_h^+$		
	error	order	gain	error	order	gain	error	order	gain
20	$8.51 \cdot 10^{-5}$	—	254.60	$1.42 \cdot 10^{-3}$	—	18.21	$1.42 \cdot 10^{-3}$	—	18.21
40	$3.23 \cdot 10^{-5}$	1.40	169.11	$3.70 \cdot 10^{-4}$	1.94	23.99	$3.70 \cdot 10^{-4}$	1.94	23.99
80	$9.43 \cdot 10^{-6}$	1.78	145.10	$9.35 \cdot 10^{-5}$	1.98	26.74	$9.35 \cdot 10^{-5}$	1.98	26.74
160	$2.47 \cdot 10^{-6}$	1.94	138.89	$2.35 \cdot 10^{-5}$	2.00	27.54	$2.35 \cdot 10^{-5}$	2.00	27.54
320	$6.19 \cdot 10^{-7}$	1.99	138.25	$5.87 \cdot 10^{-6}$	2.00	27.55	$5.87 \cdot 10^{-6}$	2.00	27.55

Table II.1: (2D shallow water equations) Errors, orders of accuracy, and gain obtained when approximating a steady solution for bases with and without prior. Parameters: DG order  $q = 1$ ,  $\mu = (\alpha, \mu, r_0) = (0.5, 0.25, 0.875)$ ,  $T = 0.01$ .

Table II.1 compares the numerical errors obtained using the DG method of degree  $q = 1$  with the standard basis  $V_h$  and the prior-enriched basis  $V_h^+$ , on a given set of parameters  $\mu = (\alpha, \mu, r_0) = (0.5, 0.25, 0.875)$ . As expected, a significant accuracy gain is observed, especially on the water height  $h$ . Figure 6 shows the deviation from the equilibrium obtained at time  $T = 1.2$ , with a numerical simulation with  $16^2$  discretization cells, after initializing with a perturbed water height:

$$h_{\text{init}}(x) = h_{\text{eq}}(x; (0.5, 0.25, 0.875)) - 0.02 \exp\left(-2\|x - (-2, -2)\|^2\right),$$

$$Q_{\text{init}}(x) = Q_{\text{eq}}(x; (0.5, 0.25, 0.875)).$$

We clearly observe that the prior-enriched basis  $V_h^+$  better captures the perturbation.

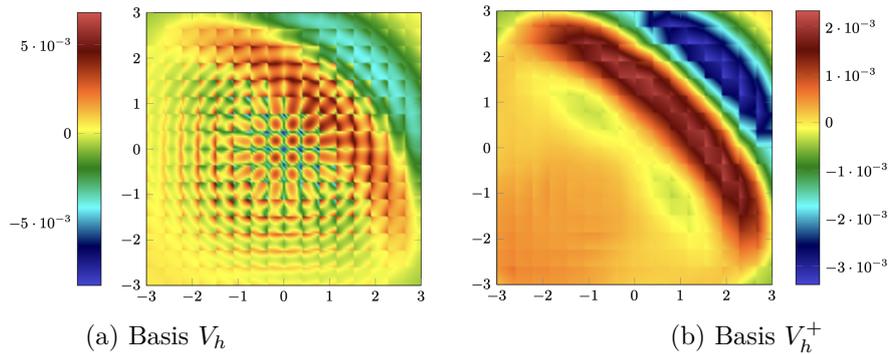


Figure 6: (2D shallow water equations) Error between the perturbed solution and the underlying steady solution  $h(x, T) - h_{\text{eq}}(x, T)$  at time  $T = 1.2$ . Parameters:  $K = 16$ .

## 4 Conclusion

In this chapter, we presented two works related to high-order methods for kinetic and hyperbolic equations. The first work provided an analysis of the recurrence effect appearing when solving

the kinetic transport equation with a finite element grid in velocity and shows that it can be related to the quadrature rule used for computing the density. The second work proposes an approximate well-balanced DG method for hyperbolic problems with source terms, based on a prior enrichment of the approximation basis. The prior can be constructed using the PINNs method, which is able to capture a whole family of stationary solutions.

Regarding the future developments on these topics, we can mention two of them:

1. **Construction of bases adapted to the problems.** A second way of proceeding would be to determine the optimal basis functions, represented by a neural network, directly by minimizing the approximation error. This requires that the entire scheme can be written as a differentiable function. The same methodology could also be applied to particle methods, using a local prior of the solution, in order to reduce the variance.
2. **Semi-Lagrangian schemes.** Semi-Lagrangian schemes are a very interesting class of numerical method for solving transport equations and therefore kinetic equations. However, the stability of the method is not always guaranteed, particularly when it is used with unstructured meshes. But the use of unstructured meshes is particularly useful for capturing solutions locally and taking complex geometries into account. Work has already been done with Hermite-type (Hsieh-Clough-Tocher) interpolators - of order 3 - for the Vlasov-Poisson equation [3]. In an ongoing work with Michel Mehrenberger, we would like to explore the use of a higher degree interpolator. It would also be interesting to exploit the use of enriched bases in the interpolation operator in order to improve the results. Finally, an interesting approach would be to couple these semi-Lagrangian methods with adaptive meshing methods, using strong gradient detection by neural networks on graphs [23].

## References

- [1] Abbasi, H. and Jenab, M.H. and Pajouh, H. Hakimi. “Preventing the recurrence effect in the Vlasov simulation by randomizing phase-point velocities in phase space”. In: *Phys. Rev. E* 84.3 (2011), p. 036702.
- [2] C. Berthon, S. Bulteau, F. Foucher, M. M’Baye, and V. Michel-Dansac. “A Very Easy High-Order Well-Balanced Reconstruction for Hyperbolic Systems with Source Terms”. In: *SIAM J. Sci. Comput.* 44.4 (2022), A2506–A2535.
- [3] N. Besse, J. Segré, and E. Sonnendrücker. “Semi-Lagrangian schemes for the two-dimensional Vlasov-Poisson system on unstructured meshes”. In: *Transport Theory and Statistical Physics* 34.3-5 (2005), pp. 311–332.
- [4] C. Buet, B. Despres, and G. Morel. “Trefftz discontinuous Galerkin basis functions for a class of Friedrichs systems coming from linear transport”. In: *Adv. Comput. Math.* 46.3 (2020).
- [5] S. Cai, Z. Mao, Z. Wang, Y. Yin, and G. E. Karniadakis. “Physics-informed neural networks (PINNs) for fluid mechanics: a review”. In: *Acta Mech. Sin.* 37.12 (2021), pp. 1727–1738.
- [6] X. Cai, W. Guo, and J.-M. Qiu. “A high order conservative semi-Lagrangian discontinuous Galerkin method for two-dimensional transport simulations”. In: *J. Sci. Comput.* 73 (2017), pp. 514–542.
- [7] Z. Cai and Y. Wang. “Suppression of recurrence in the Hermite-spectral method for transport equations”. In: *arXiv preprint arXiv:1709.08194* (2017).

- [8] C.-Z. Cheng and G. Knorr. “The integration of the Vlasov equation in configuration space”. In: *J. Comput. Phys.* 22.3 (1976), pp. 330–351.
- [9] Y. Cheng, I.M. Gamba, and P.J. Morrison. “Study of conservation and recurrence of Runge–Kutta discontinuous Galerkin schemes for Vlasov–Poisson systems”. In: *J. Sci. Comput* 56.2 (2013), pp. 319–349.
- [10] N. Crouseilles, M. Mehrenberger, and F. Vecil. “Discontinuous Galerkin semi-Lagrangian method for Vlasov–Poisson”. In: *ESAIM: Proceedings*. Vol. 32. EDP Sciences. 2011, pp. 211–230.
- [11] L. Einkemmer. “A study on conserving invariants of the Vlasov equation in semi-Lagrangian computer simulations”. In: *J. Plasma Phys.* 83.2 (2017).
- [12] L. Einkemmer and A. Ostermann. “A strategy to suppress recurrence in grid-based Vlasov solvers”. In: *Eur. Phys. J. D.* 68 (2014), p. 197.
- [13] B. Eliasson. “Outflow boundary conditions for the Fourier transformed two-dimensional Vlasov equation”. In: *J. Comput. Phys.* 181.1 (2002), pp. 98–125.
- [14] I. Gómez-Bueno, M. J. Castro, and C. Parés. “High-order well-balanced methods for systems of balance laws: a control-based approach”. In: *Appl. Math. Comput.* 394 (2021), p. 125820.
- [15] I. Gómez-Bueno, M. J. Castro Díaz, C. Parés, and G. Russo. “Collocation Methods for High-Order Well-Balanced Methods for Systems of Balance Laws”. In: *Mathematics* 9.15 (2021), p. 1799.
- [16] A.J. Klimas and W.M. Farrell. “A splitting algorithm for Vlasov simulation with filamentation filtration”. In: *J. Comp. Phys.* 110.1 (1994), pp. 150–163.
- [17] F. Kretschmar, S. M. Schnepp, I. Tsukerman, and T. Weiland. “Discontinuous Galerkin methods with Trefftz approximations”. In: *J. Comput. Appl. Math.* 270 (2014), pp. 211–222.
- [18] E. Madaule, M. Restelli, and Eric Sonnendrücker. “Energy conserving discontinuous Galerkin spectral element method for the Vlasov–Poisson system”. In: *J. Comput. Phys.* 279 (2014), pp. 261–288.
- [19] N. Pham. “Méthodes numériques pour l’équation de Vlasov réduite”. PhD thesis. Université de Strasbourg, 2016. (tel-01412750).
- [20] Jing-Mei Qiu and Chi-Wang Shu. “Positivity preserving semi-Lagrangian discontinuous Galerkin formulation: theoretical analysis and application to the Vlasov–Poisson system”. In: *J. Comput. Phys.* 230.23 (2011), pp. 8386–8409.
- [21] M. Raissi, P. Perdikaris, and G. E. Karniadakis. “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations”. In: *J. Comput. Phys.* 378 (2019), pp. 686–707.
- [22] James A Rossmanith and David C Seal. “A positivity-preserving high-order semi-Lagrangian discontinuous Galerkin scheme for the Vlasov–Poisson equations”. In: *J. Comput. Phys.* 230.16 (2011), pp. 6203–6232.
- [23] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S.Y. Philip. “A comprehensive survey on graph neural networks”. In: *IEEE Trans. Neural Netw. Learn. Syst.* 32.1 (2020), pp. 4–24.
- [24] Y. Xing and C.-W. Shu. “High order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms”. In: *J. Comput. Phys.* 214.2 (2006), pp. 567–598.
- [25] Y. Xing, X. Zhang, and C.-W. Shu. “Positivity-preserving high order well-balanced discontinuous Galerkin methods for the shallow water equations”. In: *Adv. Water Resour.* 33.12 (2010), pp. 1476–1493. ISSN: 0309-1708.



---

## Chapter III

# Multi-scale hyperbolic problems and implicit schemes

---

## 1 Introduction

This chapter focuses on numerical schemes dedicated to the solution of **non-linear hyperbolic systems** incorporating multi-scale phenomena. Indeed, the characteristic velocities of the waves present in the solutions of this type of model can differ by several orders of magnitude. This poses a real challenge for numerical simulations because fast wave velocities impose very restrictive CFL-type stability conditions, even though these same waves may be of low magnitude and therefore represent only a small perturbation of the global dynamics. For instance, such challenges arise in the simulation of low-Mach number flows, for which acoustic waves propagate small pressure perturbations and are very fast compared to the so-called material wave.

To tackle such difficulties and bypass the stringent CFL stability condition, fast waves have to be discretized using **implicit methods**. However, implicit methods generally raise difficult issues in terms of matrix assembly and storage, matrix inversion and parallelization when the size of the system to be solved is large.

We thus develop and study numerical schemes, whose implicit steps are as simple as possible in order to improve their efficiency. These methods are based on a **relaxation techniques**, which consists in expressing the non-linear hyperbolic system as a singular limit of an extended system with more unknowns but with a simpler mathematical structure. More precisely, we explore two kinds of relaxation:

1. **vectorial kinetic relaxation**, where the extended system contains only transport equations with constant velocities; this relaxation can be applied to any hyperbolic systems.
2. **Suliciu relaxation**, where the extended system has only linearly degenerate characteristic fields; this relaxation is specific to the Euler equation and enables us to separate acoustic and material transport with a splitting method. Thanks to a modified Suliciu relaxation, the implicit treatment of the acoustic part can be made very simple.

It is important to note that there is an extensive literature on relaxation schemes, in particular to develop explicit schemes with controlled mathematical properties. Our aim is clearly different here.

## 2 Implicit vectorial kinetic relaxation schemes

To solve a hyperbolic system of conservation law:

$$\partial_t U + \nabla \cdot F(U) = 0, \quad (34)$$

where  $U(x, t) \in \mathbb{R}^m$  is the vector of  $m$  conserved quantities depending on  $x \in \mathbb{R}^d$  and  $t \geq 0$  and  $F(U) \in M_{m,d}(\mathbb{R})$  denotes the flux, kinetic relaxation schemes consist into introducing a distribution function  $f(x, \lambda, t)$ , with  $\lambda \in \Omega_\lambda \subset \mathbb{R}^d$ , satisfying a **kinetic transport equation with BGK relaxation** term. The BGK relaxation is determined such that, formally, in the limit of large relaxation frequency, some velocity moments associated with  $f$  actually satisfies (34). As a result, the kinetic formulation is an ad hoc mesoscopic equation, whose macroscopic version is Equation (34). As the relaxation dynamics is not relevant for our purpose, the relaxation operator is approximated by a simple projection at the discrete level. Therefore, a **transport-projection** discretization is usually considered for the kinetic equation and it results in the so-called kinetic relaxation schemes for Equation (34). They have been introduced in [29]. Using the same approach, Lattice Boltzmann Method (LBM) considers a discrete velocity set  $\Omega_v$ . Discrete kinetic relaxation schemes have then been proposed for any scalar conservation laws in [27] and **vectorial kinetic relaxation** have been proposed to tackle any hyperbolic system [25, 2, 4]. Based on this reformulation, we have proposed an implicit kinetic relaxation scheme in a series of works [Imp11, Imp7, Imp6, Imp10, Imp3, Imp1].

### 2.1 Reformulation

**Kinetic model.** The **vectorial kinetic relaxation** method consists in introducing  $p$  kinetic velocities

$$\lambda_1, \dots, \lambda_p \in \mathbb{R}^d,$$

and a discrete kinetic distribution

$$f(x, t) = (f_1(x, t), \dots, f_p(x, t)) \in (\mathbb{R}^m)^p.$$

According to the naming convention, we will refer to this vectorial kinetic representation as  $[Dd, Qp]^m$ , the power referring to the fact that there is a discrete kinetic distribution for each  $m$  conservative variables. The kinetic distribution is then chosen to satisfy the following kinetic BGK equation:

$$\partial_t f_i + \lambda_i \cdot \nabla f_i(x, t) = \frac{1}{\varepsilon} (\mathcal{M}(U_f)_i - f_i), \quad \forall i \in \{1, \dots, p\}, \quad (35)$$

$$U_f(x, t) = \sum_{i=1}^p f_i(x, t), \quad (36)$$

where  $U_f$  is the density associated with  $f$ . Equation (35) is a linear transport equation with a non-linear source term. Each component  $f_i$  of the kinetic distribution is transported at constant velocity  $\lambda_i$ , while their coupling results from the BGK source term: it makes the distribution  $f$  relax to an equilibria kinetic distribution  $\mathcal{M}(U_f) \in (\mathbb{R}^m)^N$ , which will be defined below. Summing the  $N$  equations, the density equation writes

$$\partial_t U_f + \nabla \cdot Q_f = \frac{1}{\varepsilon} (U_{\mathcal{M}(U_f)} - U_f), \quad \forall i \in \{1, \dots, N\}, \quad (37)$$

where  $Q_f$  denotes the momentum

$$(Q_f)_k(x, t) = \sum_{i=1}^p (\lambda_i)_k f_i(x, t), \quad \forall k \in \{1, \dots, d\}.$$

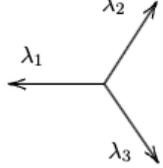
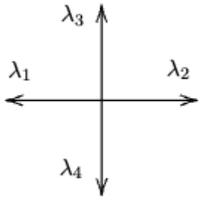
Name	Velocities	Equilibria	Moments
[D1Q2]		$\mathcal{M}(U)_i = \frac{U}{2} + \frac{F(U)\lambda_i}{2\lambda^2}$	$(1, x_1)$
[D2Q3]		$\mathcal{M}(U)_i = \frac{U}{3} + \frac{F(U)\lambda_i}{3\lambda^2}$	$(1, x_1, x_2)$
[D2Q4]		$\mathcal{M}(U)_i = \frac{U}{4} + \frac{F(U)\lambda_i}{2\lambda^2}$	$(1, x_1, x_2, x_1^2 - x_2^2)$

Table III.1: Discrete Velocities

**Compatibility conditions and equilibrium distribution.** In order to recover the initial hyperbolic system in the limit  $\varepsilon \rightarrow 0$ , the equilibria kinetic distribution  $\mathcal{M}(U)$  is chosen such that the so-called compatibility conditions hold:

$$U_{\mathcal{M}(U)} = U, \quad Q_{\mathcal{M}(U)} = F(U). \quad (38)$$

The first condition makes the right-hand side of Equation (37) vanish. Then, in the limit  $\varepsilon \rightarrow 0$ , assuming that  $U_{f^\varepsilon}$  converges to  $U$ ,  $f^\varepsilon$  formally converges to  $\mathcal{M}(U)$  and  $Q_{f^\varepsilon}$  converges to  $Q_{\mathcal{M}(U)} = F(U)$ . Therefore, the limit function  $U$  satisfies the original hyperbolic system (34) and the kinetic model does provide an approximation to this equation.

To define the kinetic approximation completely, we need to identify the equilibrium distribution function  $\mathcal{M}$ . As it is composed of  $n$  vectors of  $\mathbb{R}^m$ , this requires  $n$  relations in  $\mathbb{R}^m$ . The compatibility conditions actually provide  $(1 + d)$  linear relations in  $\mathbb{R}^m$ . Consequently, if  $p$  is strictly larger than  $(1 + d)$ , additional moments  $Z_f$  of the equilibrium should be prescribed to define it. Denoting the moment operator as:

$$\mathbf{m}f = (U_f, (Q_f)_1, \dots, (Q_f)_d, Z_f) \in (\mathbb{R}^m)^n,$$

the equilibrium is then fully defined by a relation like:

$$\mathbf{m}\mathcal{M}(U) = (U, (F(U))_1, \dots, (F(U))_d, 0),$$

as soon as  $\mathbf{m}$  is invertible. Table III.1 lists the three different choices of kinetic approximation we use in practice: all kinetic velocities have the same magnitude, denoted  $\lambda > 0$ . While [D1Q2] and [D2Q3] are fully defined by the compatibility conditions, [D2Q4] requires an additional moment, defined by:

$$Z_f = \sum_{i=1}^p \left( (\lambda_i)_{x_1}^2 - (\lambda_i)_{x_2}^2 \right) f_i,$$

**Box 5: Jin-Xin relaxation**

The Jin-Xin relaxation method has been introduced in [25]. It corresponds to the [D1Q2]<sup>m</sup> method. Starting from the following one-dimensional conservation equation:

$$\partial_t U + \partial_x F(U) = 0,$$

the method consists into introducing an additional variable  $q$  and considering the following system:

$$\begin{aligned} \partial_t U + \partial_x Q &= 0, \\ \partial_t Q + \lambda^2 \partial_x U &= \frac{1}{\varepsilon}(F(U) - Q). \end{aligned}$$

The non-linear equation has been approximated by a linear hyperbolic system with a stiff source term. Writing the system in the characteristic variables,  $f_1 = U/2 - Q/(2\lambda)$  and  $f_2 = U/2 + Q/(2\lambda)$ , we get the diagonal hyperbolic system:

$$\begin{aligned} \partial_t f_1 - \lambda \partial_x f_1 &= \frac{1}{\varepsilon} \left( \left( \frac{U}{2} - \frac{F(U)}{2\lambda} \right) - f_1 \right), \\ \partial_t f_2 + \lambda \partial_x f_2 &= \frac{1}{\varepsilon} \left( \left( \frac{U}{2} + \frac{F(U)}{2\lambda} \right) - f_2 \right), \end{aligned}$$

which coincides with Equation (35) with kinetic velocities  $\lambda_1 = -\lambda$  and  $\lambda_2 = \lambda$ . We also refer to [24].

**2.2 Implicit splitting scheme**

The vectorial kinetic formulation (35)-(36) naturally leads us to consider a numerical method based on a splitting strategy, where the transport part and the relaxation part are solved successively at each time step. This is a common strategy to solve such kinetic equations involving a stiff source term modeling collisions (e.g. [10, 13]). Note, however, that, in this part, this collision term has no physical significance: we are only interested in designing a numerical method for the limit model (34) and there is no point in capturing the collisional dynamics precisely. Thus, the above numerical scheme is not, strictly speaking, a discretization of (35)-(36) but rather a discretization of (34) directly.

The time semi-discretization writes as follows. Denoting  $f^n$  the distribution function at discrete time  $t^n = n\Delta t$ , then each time step is decomposed into:

- (i) a **transport step**:  $f^* = T_h(\Delta t)f^n$ , which consists of solving the  $p$  transport equations at constant velocities over the time step  $\Delta t$ :

$$(T_h(\Delta t)f^n)_i \approx f_i^n(\cdot - \lambda_i \Delta t).$$

- (ii) a **relaxation step**:  $f^{n+1} = R_\omega f^*$ , defined by:

$$f^{n+1} = f^* + \omega \left( \mathcal{M}(U_{f^*}) - f^* \right), \quad (39)$$

where  $\omega \in ]0, 2]$  is a numerical parameter. This relaxation operator  $R_\omega$  can be obtained as the limit  $\varepsilon \rightarrow 0$  in a  $\theta$ -scheme applied to the relaxation differential equation. We refer to [Imp5] for more details.

Gathering both steps, the discrete operator writes:

$$\Psi_\omega(\Delta t) = R_\omega \circ T_h(\Delta t) \quad (40)$$

We emphasize once more that the relaxation operator is not aimed at capturing the stiff relaxation dynamics. In both LBM and kinetic schemes, the classical method is indeed the transport-projection scheme obtained with  $\omega = 1$ . Taking  $\omega = 2$  yields to the so-called over-relaxation operator, introduced in [12], which is a symmetric operator with respect to the equilibria manifold.

The kinetic scheme can be equivalently written on the moments of  $f^n$ :

$$(U^n, (Q^n)_1, \dots, (Q^n)_d, Z^n) = \mathbf{m}f^n,$$

and  $(U^n)$  is actually an approximated solution to the initial hyperbolic equation (34), as it will be discussed in Section 2.4. This method presents several advantages. First, if the transport solver is conservative, then the full method is conservative. Secondly, the scheme concentrates all the non-linear dynamics in the local relaxation operator, while the non-local coupling between data is ensured by the linear transport step. Because this transport step is at constant velocities, it can be solved using very efficient **implicit solvers**.

**Implicit transport  $T_h(\Delta t)$ .** The mere advantage of this method is to simplify the advection step, which is the more time and CPU demanding. The main solver we have considered is an **implicit Discontinuous Galerkin** method with upwind fluxes (see Box 3). This method naturally handles **unstructured meshes** and has high-order accuracy in space. Since each of the  $N$  linear transport equations are at constant velocities, the scheme formally results into  $N$  triangular systems to invert. Indeed, data in one cell only depends on the data of the upwind ones and as the transport velocities are constant, there is no dependency loop. More precisely, with the notation introduced in Box 3, the DG discretization of the  $i$ -th component writes:

$$\int_L \frac{(f_i)_L^n - (f_i)_L^{n-1}}{\Delta t} \psi_j^L - \int_L \lambda_i \cdot \nabla \psi_j^L (f_i)_L^n + \int_{\partial L} \left( (\lambda_i \cdot n_{LR})^+ (f_i)_L^n + (\lambda_i \cdot n_{LR})^- (f_i)_R^n \right) \psi_j^L = 0. \quad (41)$$

We can rewrite it in the matrix form

$$M_L (f_i)_L^n = M_L (f_i)_L^{n-1} + \Delta t \left( A_L (f_i)_L^n - \sum_{\substack{LR \in \partial L, \\ \lambda_i \cdot n_{LR} > 0}} B_{LR} (f_i)_L^n - \sum_{\substack{LR \in \partial L, \\ \lambda_i \cdot n_{LR} < 0}} C_{LR} (f_i)_R^n \right), \quad (42)$$

where  $M_L$  denotes the mass matrix,  $A_L$  the volume advection matrix,  $B_{LR}$  and  $C_{LR}$  the flux advection matrices between cell  $L$  and downwind or upwind cells  $R$ : cell  $R$  is said to be upwind with respect to cell  $L$  if  $\lambda_i \cdot n_{LR} < 0$  and downwind otherwise. Their coefficients are given by

$$(M_L)_{j,k} = \int_L \psi_j^L \psi_k^L, \quad (A_L)_{j,k} = \int_L (\lambda_i \cdot \nabla \psi_j^L) \psi_k^L, \\ (B_{LR})_{j,k} = \int_{\partial L} (\lambda_i \cdot n_{LR}) \psi_j^L \psi_k^L \quad \text{and} \quad (C_{LR})_{j,k} = \int_{\partial L} (\lambda_i \cdot n_{LR}) \psi_j^L \psi_k^R.$$

In [Imp6], we have proposed a resolution of these triangular systems with an **explicit cost**. Indeed, as the system is triangular, a **lifting-like algorithm** can be used. To each kinetic velocity  $\lambda_i$ , we first compute and store the associated dependency graph between cells. Then, one step of the implicit solver simply consists of going through the dependency graph. Consequently, there is no need of an iterative process and although requiring the storage of the

dependency graphs, the overall algorithm has an explicit cost. In practice, a **second-order Crank-Nicolson** time discretization is used, following the same method.

In order to tackle large problems, a parallel implementation of the algorithm has been carried out in the `chuckrut` code, based on a task graph approach [Imp11, Imp6]. It is based on the `StarPu` library developed at INRIA Bordeaux, dedicated to **task graph parallelism** [3] on GPU-CPU hybrid architecture. One main issue regards the granularity of the tasks: too large tasks reduce the level of parallelism and too small tasks require a lot of task management and data movements. A specialized Direct Acyclic Graph (DAG) clustering algorithm has been proposed in [8, 7] to control the granularity of the tasks.

### 2.3 High order time splitting method

The first order scheme (40), for  $\omega \in [1, 2[$ , is very diffusive due to the  $O(\Delta t)$  time error coming from the splitting method. Even with the second-order version  $\omega = 2$ , the scheme is not accurate enough for tackling realistic plasma simulations since too small-time steps would be required. We thus developed a higher order splitting method based on **composition** techniques.

**Time symmetric second order splitting.** High order composition methods require a second-order time-symmetric splitting [20]. As soon as the transport operator is second-order accurate, the classical Strang splitting

$$\Psi_{\text{Strang}}(\Delta t) = T\left(\frac{\Delta t}{2}\right) \circ R_2 \circ T\left(\frac{\Delta t}{2}\right)$$

is second order accurate and satisfies the symmetric property:  $\Psi_{\text{Strang}}^{-1}(-\Delta t) = \Psi_{\text{Strang}}(\Delta t)$ . However,  $\Psi_{\text{Strang}}(0) = R_2$  preserves the macroscopic quantity  $U$  but does not equal the identity operator on the kinetic distribution: by definition, this prevents the scheme to be a time-symmetric splitting. Therefore we consider the following splitting:

$$\begin{aligned} \Psi_{\text{sym}}(\Delta t) &= \Psi_{\text{Strang}}\left(\frac{\Delta t}{2}\right) \circ \Psi_{\text{Strang}}\left(\frac{\Delta t}{2}\right) \\ &= T\left(\frac{\Delta t}{4}\right) \circ R_2 \circ T\left(\frac{\Delta t}{2}\right) \circ R_2 \circ T\left(\frac{\Delta t}{4}\right), \end{aligned}$$

which satisfies the condition:  $\Psi_{\text{Strang}}(0) = \text{Id}$ . This splitting is now time-symmetric. Note that, while the dynamics of  $U$  is Eq. (34) by construction, the dynamics of  $f$  is not. We refer to Section 2.4 for further discussion.

**Palindromic composition method.** Equipped with such a time-symmetric splitting, higher-order method can be considered just by composing this building blocks:

$$\Psi_{\text{compo}}(\Delta t) = \Psi_{\text{sym}}(\gamma_0 \Delta t) \circ \Psi_{\text{sym}}(\gamma_1 \Delta t) \circ \cdots \circ \Psi_{\text{sym}}(\gamma_n \Delta t), \quad (43)$$

where  $\gamma_0, \gamma_1, \dots, \gamma_n$  are the coefficients of the method. To ensure the symmetric property, they have to be a palindromic sequence:  $\gamma_p = \gamma_{n-p}$ , for all  $0 \leq p \leq n$ . Suzuki and Kahan-Li coefficients lead to respectively fourth and six order schemes, with respectively  $n + 1 = 5$  steps and 9 steps [20].

Figure 7 shows a convergence study performed on the one-dimensional isothermal Euler equation,  $U = (\rho, \rho u)^T$ ,  $F(U) = (\rho u, \rho u^2 + c^2 \rho)^T$  with sound speed  $c = 0.6$ . We use the  $[D1Q2]^2$

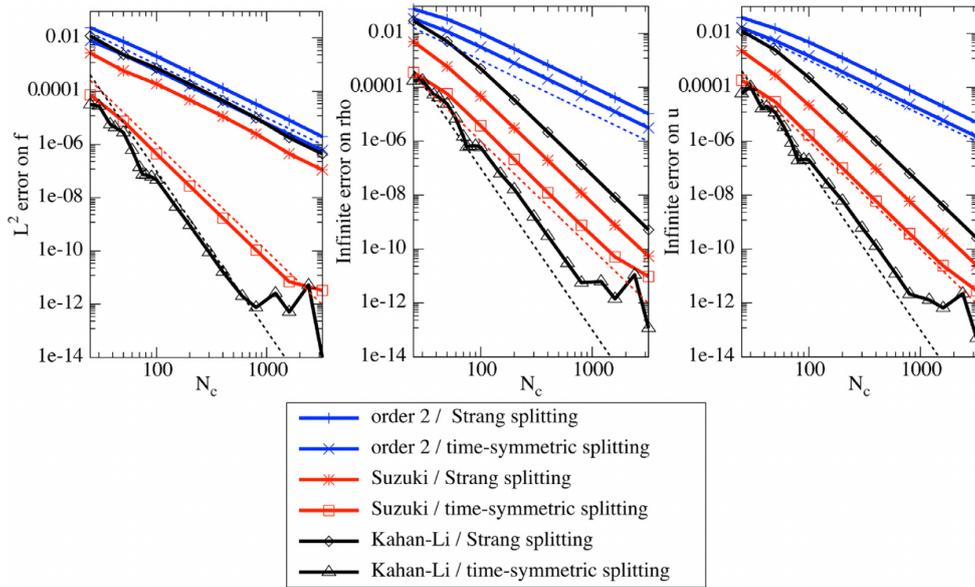


Figure 7: Convergence study (left:  $f$ , middle:  $\rho$ , right:  $\rho u$ ) for several palindromic methods and the second order building blocks  $\Psi_{\text{Strang}}$  and  $\Psi_{\text{sym}}$ . The dotted lines are reference lines with slopes 2, 4 and 6 respectively. CFL number  $\beta = 5$ . Taken from [Imp6].

kinetic vectorial scheme, with speed  $\lambda = 2$ , and a DG scheme of order 5 for the implicit transport steps. The initial condition is given by:

$$\rho(x, 0) = 1 + e^{-30x^2}, \quad u(x, 0) = 0,$$

and the solution is computed on the interval on the domain  $[-2, 2]$  up to time  $t_{\text{max}} = 0.4$ . A reference solution is computed with a very fine grid ( $N_x = 3200$ ). As expected, we can observe that fourth and sixth orders are obtained on the distribution function  $f$  only when using the time-symmetric splitting in the composition (44). Regarding the conserved quantities  $U = (\rho, \rho u)^T$ , keeping the Strang splitting in the composition method seems sufficient for the Suzuki composition to be fourth-order accuracy, but the time-symmetric splitting is required to reach six-order accuracy with Kahan-Li.

**Composition method with rational coefficients.** Let us consider the subclass of palindromic composition methods with even  $n$  and which writes:

$$\Psi_{\text{compo}}(\Delta t) = (\Psi_{\text{sym}}(\gamma_0 \Delta t))^{n/2} \circ \Psi_{\text{sym}}(\gamma_{n/2} \Delta t) \circ (\Psi_{\text{sym}}(\gamma_0 \Delta t))^{n/2}, \quad (44)$$

Then, as stated in [26], fourth order accuracy is reached when  $\gamma_0$  and  $\gamma_{n/2}$  satisfies the relations:

$$n\gamma_0 + \gamma_{n/2} = 1, \quad n\gamma_0^3 + \gamma_{n/2}^3 = 0.$$

For  $n = 8$ , these equations have actually a rational solution:  $\gamma_0 = 1/6$ ,  $\gamma_4 = -1/3$ . On a regular Cartesian mesh, this enables us to replace the DG scheme by a LBM, with exact transport [Imp1].

## 2.4 Equivalent equations and stability issues

**Equivalent equation.** To analyze the time stability of the scheme and derive an equivalent equation for the approximated solution ( $U^n$ ), a Chapman-Enskog expansion in  $\Delta t$  can

be formally performed on the transport-relaxation scheme [14, 17, Imp5], supposing that the transport is exact. For the  $[D1Q2]^m$  method, the equivalent equation thus writes:

$$\partial_t U + \partial_x F(U) = \Delta t \left( \frac{1}{\omega} - \frac{1}{2} \right) \partial_x \left( (\lambda^2 \text{Id} - F'(U)^2) \partial_x U \right) + O(\Delta t^2), \quad (45)$$

where  $\text{Id}$  and  $F'(U)$  are matrices of size  $m \times m$ . The equivalent equation for the  $[D2Q4]^m$  transport-relaxation scheme can be found in [3]. A formal stability condition would be to enforce the diffusion matrix to be positive definite. For the  $[D1Q2]^m$  method, this leads to the so-called subcharacteristic or **diffusive stability** condition:

$$\sigma(F'(U)) \in ] -\lambda, \lambda[,$$

where, for any matrix  $A$ ,  $\sigma(A)$  denotes its spectrum. The kinetic speed  $\lambda$  should be larger than the characteristic speed of the hyperbolic system. Note, however, that this condition may not be sufficient to ensure the stability.

**Equivalent system.** Instead of focusing on the quantities  $U^n$  only, an equivalent equation can be obtained for the dynamics of the  $p$  variables  $(U^n, Q^n, Z^n) = M f^n$  or equivalently on the following moments:

$$(U^n, Y^n, Z^n) = (U^n, (Q^n - F(U^n))_1, \dots, (Q^n - F(U^n))_d, Z^n) \in (\mathbb{R}^m)^p.$$

where the momentum has been replaced by the flux error. In these new variables, the relaxation operator takes the form:

$$\bar{R}_\omega(U, Y, Z) = (U, (1 - \omega)Y, Z)$$

where the bar on the operator indicates that the operator is expressed in the moment variables  $(U, Y, Z)$ . Hence the flux error is relaxed to zero for  $\omega = 1$  or changes sign for  $\omega = 2$ . In order to obtain the equivalent equation while filtering the oscillation due to the relaxation, a finite difference involving two time steps is expanded:

$$\frac{\bar{\Psi}_{\text{sym}}(\Delta t) - \bar{\Psi}_{\text{sym}}(-\Delta t)}{2\Delta t} \begin{bmatrix} U(\cdot, t) \\ Y(\cdot, t) \\ Z(\cdot, t) \end{bmatrix} = \partial_t \begin{bmatrix} U(\cdot, t) \\ Y(\cdot, t) \\ Z(\cdot, t) \end{bmatrix} + O(\Delta t^2).$$

Performing a Taylor expansion of the first term, we have obtained:

**Proposition 8** ([Imp7]). *The equivalent system of the  $[D1Q2]^m$  transport - over-relaxation scheme ( $\omega = 2$ ) is given by:*

$$\partial_t \begin{bmatrix} U \\ Y \end{bmatrix} + \begin{bmatrix} F'(U) & 0 \\ 0 & -F'(U) \end{bmatrix} \partial_x \begin{bmatrix} U \\ Y \end{bmatrix} = O(\Delta t^2).$$

Considering the scalar transport equation, with  $F(U) = aU$  with  $U \in \mathbb{R}$  and  $a \in \mathbb{R}$ , the equivalent system of the  $[D1Q2]^m$  transport - over-relaxation scheme ( $\omega = 2$ ) writes:

$$\partial_t \begin{bmatrix} U \\ Y \end{bmatrix} + \begin{bmatrix} a & 0 \\ 0 & -a \end{bmatrix} \partial_x \begin{bmatrix} U \\ Y \end{bmatrix} + B_2 \Delta t^2 \partial_{x^2} \begin{bmatrix} U \\ Y \end{bmatrix} = O(\Delta t^3).$$

with  $B_2 = \begin{bmatrix} (\lambda^2 - a^2) & 3a \\ 3a(\lambda^2 - a^2) & -(\lambda^2 - a^2) \end{bmatrix}$ . Under periodicity boundary conditions on the domain  $[0, L]$ , this equivalent system preserves the total energy

$$E(t) = \int_{[0, L]} (\lambda^2 - a^2) \frac{U(x, t)^2}{2} + \frac{Y(x, t)^2}{2} dx,$$

which is convex under the subcharacteristic condition  $\lambda < |a|$ .

This result shows that, at first order, the conserved quantity  $U$  and the flux error  $Y$  propagate in opposite directions. Still for the scalar transport equation, this has been further generalized to  $\omega \in ]1, 2[$  in [16] using a symbolic computing library. Then the equivalent equation reads:

$$\partial_t \begin{bmatrix} U \\ Y \end{bmatrix} - \frac{A_\omega}{\Delta t} \begin{bmatrix} U \\ Y \end{bmatrix} + B_\omega \partial_x \begin{bmatrix} U \\ Y \end{bmatrix} = O(\Delta t^2),$$

with

$$A_\omega = \begin{bmatrix} 0 & 0 \\ 0 & \alpha_\omega \end{bmatrix}, \quad B_\omega = \begin{bmatrix} a & \gamma_\omega \\ \gamma_\omega(\lambda^2 - a^2) & -\beta_\omega a \end{bmatrix},$$

with  $\alpha_\omega = \omega(\omega - 2)(\omega^2 - 2\omega + 2)/(2(\omega - 1)^2)$ ,  $\beta_\omega = (\omega^4 - 4\omega^3 + 6\omega^2 - 4\omega + 2)/(2(\omega - 1)^2)$  and  $\gamma_\omega = (\omega - 2)^2(\omega^2 - 2\omega + 2)/(8(\omega - 1)^2)$ . The equivalent equation (45) can then be recovered by supposing that  $Y = O(\Delta t)$ . It has been shown that the system is hyperbolic under the subcharacteristic condition: **hyperbolic stability condition** for the equivalent system and diffusive stability condition for the equivalent equation match. Note, however, that it is not always the case: in [16], the authors show that the hyperbolicity of the equivalent system of the [D2Q4]<sup>m</sup> is more restrictive than the diffusive stability condition.

**Remark.** The information given by the equivalent system has been used in [Imp7] to devise **boundary conditions** that ensure second order accuracy for the scalar transport equation. Indeed, the kinetic scheme involves the additional variables  $Y$ , whose boundary conditions are lacking. Numerical experiments show that second accuracy is reached when imposing a Neumann boundary condition on  $Y$  at the outgoing boundary in addition to the Dirichlet boundary condition on  $U$  at the incoming boundary. Generalizations to non-linear systems have been investigated in [22].

**Entropic stability conditions.** Conditions are not sufficient and entropic conditions have been introduced in [9, 4, 15]. Starting from an entropy-entropy flux pair for the original hyperbolic system, it consists into constructing a kinetic entropy  $H(f)$  for the relaxation system. We refer to Box 6 on this construction. Then transport-projection scheme (with  $\omega = 1$ ) is entropy decreasing, since the projection is so and the transport preserves it. As explored in [Imp1], entropy conserving method can be designed by defining a non-linear relaxation parameter  $\omega(f)$  such that the kinetic entropy is conserved during the relaxation phase.

### Box 6: Entropic stability

In this box, we briefly review the entropic stability developed in [4, 15]. Let us consider a strictly convex entropy  $\eta(U)$  and the associated entropy flux  $G(U)$ , which satisfies:  $\eta'(U)F'(U) = G'(U)$ , where the primes denote the Jacobian of the functions in this box. Then we define entropic variables  $V$  through the following mapping:

$$V = \eta'(U) \quad \Leftrightarrow \quad U = \eta^{*'}(V)$$

where  $\eta^*(V) = \inf\{U \cdot V - \eta(U)\}$  denotes the Legendre transform of  $\eta$ . We denote by  $G^{\star}(V) = \inf\{V \cdot F(U) - G(U)\}$ , which satisfies the following relation:  $G^{\star}'(V) = F(U)$ . Let us consider functions  $h_i^*(V)$  which satisfies the following compatibility conditions:

$$\eta^*(V) = \sum_{i=1}^p h_i^*(V), \quad G^{\star}(V) = \sum_{i=1}^p \lambda_i h_i^*(V). \quad (46)$$

Supposing that they are differentiable, we can define an equilibrium distribution  $\mathcal{M}(U)$  as the gradient of these functions

$$\mathcal{M}_i(U) = (h_i^*(V))'{}^T.$$

These are indeed an equilibrium distribution, since taking the derivatives of Eq. (46) with respect to  $V$  results in the compatibility conditions (38). In particular, when there exists a unique solution to these compatibility conditions, like for the [D1Q2]<sup>m</sup> or the [D2Q3]<sup>m</sup>, we recover the same standard equilibrium distributions. Supposing that the functions  $h_i^*$  are convex (which is equivalent to  $(\mathcal{M}_i)'(U)$  has nonnegative real eigenvalues, since  $(\mathcal{M}_i)'(U) = h_i^{*''}(V)\eta''(U)$ ), we can introduce their Legendre transform  $(h_i)$ . Then we have the following relations:

$$\eta(U) = \min_{f, \sum_{i=1}^p f_i = U} \sum_{i=1}^p h_i(f) = \sum_{i=1}^p h_i(\mathcal{M}_i(U)).$$

where the second one results from the Lagrange multiplier relation. Consequently, in that case, we have just defined a kinetic entropy

$$H(f) = \sum_{i=1}^p h_i(f_i),$$

minimal on the equilibria distribution with given  $U$  and compatible with the macroscopic entropy  $\eta$ . The main point of this construction is to ensure that the  $h_i^*$  are convex. For the [D1Q2]<sup>m</sup> applied to the transport equation, it can be shown that we recover the subcharacteristic condition.

## 2.5 Applications to plasma physics simulations

In [Imp3], we have adapted the vectorial kinetic scheme to the toroidal geometry with the aim to perform numerical simulations of plasma in Tokamak. As already presented in Box 1, plasma dynamics involves several multiscale transport phenomena, which implies stringent constraints on the discretization parameters. Indeed, the large toroidal magnetic field results in a scale separation between toroidal and poloidal dynamics: particles exhibit fast transport in the toroidal direction, while the poloidal dynamics follow nearly incompressible dynamics. On the other hand, numerical constraints could result only from the available mesh: indeed, for explicit numerical schemes, the time step could be severely constrained by the CFL conditions in presence of small mesh cells.

We aim at solving non-linear transport equations of a charge density  $\rho(x, t)$  in a cylindrical domain:  $(r, z, \varphi) \in \Omega \times [0, L]$ . The numerical methods must naturally be adapted to such geometry both in their structure and parallelism. Very simple numerical methods can be envisaged in the  $\varphi$  direction (e.g. Fourier or semi-Lagrangian method), while more sophisticated solvers have to be considered to deal with the transport in unstructured meshes in  $(r, z)$  planes. Here we consider the [D3Q6] kinetic relaxation schemes with different kinetic speeds in the toroidal and poloidal directions,  $\lambda_p$ ,  $\lambda_t$ , and different spatial discretization of the transport step in the poloidal and toroidal directions:

- (i) in the toroidal direction  $[0, L]$ , the mesh is supposed to be uniform with a space step  $\Delta\varphi$ . Thus we can consider an exact transport solver: the time step is then chosen such as to

verify the relation

$$\Delta t = \Delta\varphi/\lambda_t,$$

as in Lattice Boltzmann method. Then the transport step boils down to a simple shift. Each data associated with one poloidal plane is associated with one MPI process and the shift consists in an MPI send/receive operation.

- (ii) in the poloidal direction  $\Omega$ , we consider the implicit DG Galerkin scheme as described in Section 2.2, implemented using the task-based parallelism.

We note that the full CFL number is defined by:

$$n_{\text{CFL}} = \frac{u_{\text{max}}\Delta t}{\min(\delta_p, \Delta\varphi)}, \quad (47)$$

where  $u_{\text{max}}$  the maximal transport speed over the simulation and  $\delta_p > 0$  denotes the minimal distance between two interpolation points in the poloidal plane. Unlike explicit methods, **this CFL number is not constrained by the scheme and, as a consequence,  $\delta_p$  can be set without stability considerations.**

As an application, we consider the following three-dimensional version of the 2D guiding center model:

$$\begin{aligned} \partial_t \rho + \nabla_{(r,z)} \cdot \left( \rho \begin{bmatrix} u_r \\ u_z \end{bmatrix} \right) + \partial_\varphi(\rho u_\varphi) &= 0, \\ u &= E \times e_\varphi + B, \\ E &= -\nabla V, \quad -\Delta_{(r,z)} V = \rho, \\ B &= (-b_\theta \cos(\theta), b_\theta \sin(\theta), b_\varphi)^T. \end{aligned}$$

with  $b_\theta = 0.1$  and  $b_\varphi = 200$ , on the domain  $\{(r, z) \mid \alpha_{\min} \leq \sqrt{(r - r_t)^2 + z^2} \leq \alpha_{\max}\} \times [0, L]$ . We consider periodic boundary conditions in  $\varphi$  and homogeneous Dirichlet boundary conditions for the electric potential  $V$ . The density is initialized as the perturbation of a stationary solution:

$$\rho_{\text{init}}(r, z, \varphi) = \rho_{\text{init}}(\alpha, \theta, \varphi) = \left( 1 + \varepsilon \cos \left( k\theta + \ell\varphi \frac{2\pi}{L} \right) \right) e^{-\frac{(\alpha - \alpha_0)^2}{2\sigma^2}}.$$

where  $(\alpha, \theta)$  are polar coordinates in the poloidal plane, with  $\sigma > 0$ ,  $\alpha_0 \in (\alpha_{\min}, \alpha_{\max})$  and where  $\varepsilon > 0$  and  $k, \ell \in \mathbb{N}^*$  parametrize the perturbation. This model, made of the transport equation along  $B$  and the poloidal  $E \times B$  drift, can be viewed as a very simplified drift-kinetic model with a unique parallel velocity [18].

Figure 8 show the results of one simulation, whose parameters are summarized in its caption. Let us emphasize the main points. Each poloidal mesh is composed of  $4 \times N_\alpha \times N_\theta = 4 \times 80 \times 50$  elements. In the toroidal direction, we use  $n_p = 128$  poloidal planes and so 128 MPI processes. Consistently with the involved advection velocities, we choose kinetic speeds of norm  $\lambda_p = 7$  in the poloidal plane and  $\lambda_t = 600$  in the toroidal direction. As we are using an exact transport solver, the time step is set to  $\Delta t = L/(n_p \lambda_t) = 0.002604$ . The relaxation parameter is chosen equal to  $\omega = 1.99$ . As expected, we can observe that the density in the poloidal planes are identical up to a rotation. As time goes on, two vortices develop and lead to fine structures in the poloidal planes as in the two-dimensional case. The computation takes 26 hours.

We would like to emphasize that the full CFL number, defined in (47), takes the value 33.17. This is due to the very fast speed in the toroidal direction which is not resolved by the poloidal

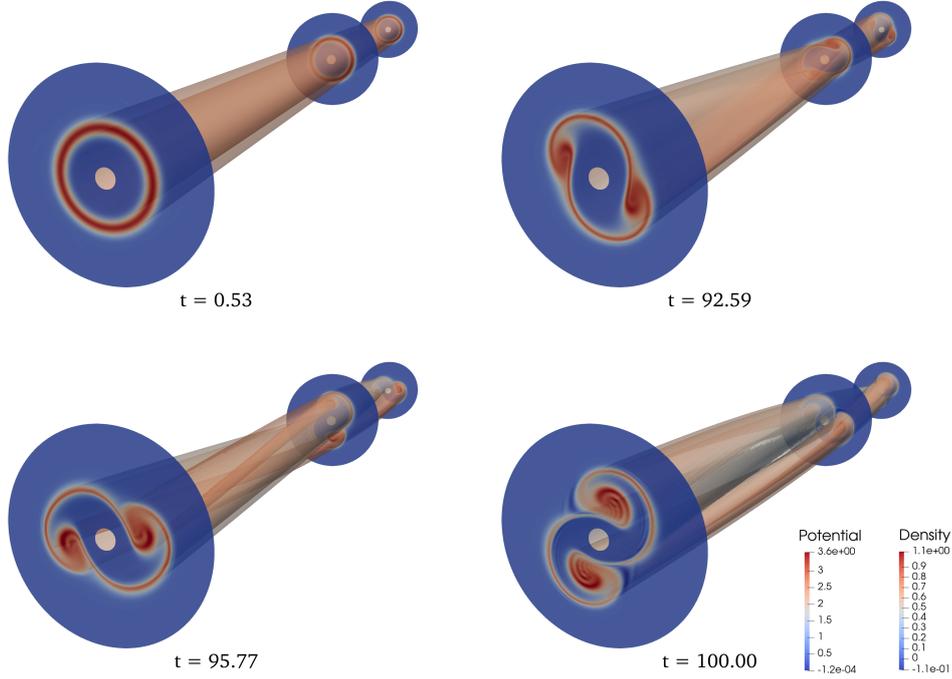


Figure 8: (Three-dimensional diocotron test case) Four successive instants of the 3D solution showing cross sections of density colormaps and density isosurface colored by the potential. Domain parameters:  $r_t =$ ,  $\alpha_{\min} = 1$ ,  $\alpha_{\max} = 100$ ,  $L = 200$ . Initial condition parameters:  $\sigma = 0.5$ ,  $\varepsilon = 10^{-6}$ ,  $k = 2$  and  $\ell = 1$ . Numerical parameters:  $\lambda_p = 7$ ,  $\lambda_t = 600$ , DG order  $p = 2$ ,  $\omega = 1.99$ ,  $n_p = 128$ . Taken from [Imp3].

plane mesh. Indeed, the minimal distance between interpolation points in the poloidal planes equals  $\delta_p = 0.01570$  and the time step is not small enough to capture a speed of order 200 for this fine resolution. An explicit scheme would require either decreasing the time step and thus increasing the toroidal discretization, or a coarser poloidal mesh at the cost of a loss of accuracy. Consequently, explicit schemes would not be able to cope with both the large velocity in the toroidal direction and a very fine grid in the poloidal plane.

### 3 Implicit Suliciu relaxation scheme

In this section, we focus on the Euler system and propose a specific relaxation method to capture the low-Mach regime.

**The Euler system and the low-Mach regime.** The Euler system describes the dynamics of a compressible gas, described by its density  $\rho(x, t) \in \mathbb{R}$ , its velocity  $u(x, t) \in \mathbb{R}^d$  and its total energy  $E(x, t) \in \mathbb{R}$ . As already introduced in Section I.4, they satisfied the following system:

$$\begin{aligned} \partial_t \rho + \nabla_x \cdot (\rho u) &= 0, \\ \partial_t (\rho u) + \nabla_x \cdot (\rho u \otimes u + p \text{Id}) &= 0, \\ \partial_t E + \nabla_x \cdot (Eu + pu) &= 0. \end{aligned}$$

For a polytropic ideal gaz, the energy is related to the pressure with the following state law:

$$E = \rho \frac{|u|^2}{2} + \frac{p}{\gamma - 1}$$

where  $\gamma > 1$  denotes the adiabatic exponent<sup>1</sup>. The Euler system is hyperbolic with characteristic speeds given by:

$$\lambda_1(n) = u \cdot n - c, \quad \lambda_{2:d+2}(n) = u \cdot n, \quad \lambda_{d+2}(n) = u \cdot n + c,$$

where  $c > 0$  denotes the sound speed:

$$c = \sqrt{\frac{\gamma p}{\rho}}.$$

The two extreme characteristic speeds are associated with the acoustic waves, i.e. the propagation of pressure variations, while the central speeds are those of material waves, i.e. the propagation of the fluid itself. The Mach number is defined as the ratio between the fluid velocity and the sound speed:

$$\text{Ma} = |u|/c.$$

In flow regime where the Mach number is very small,  $\text{Ma} \ll 1$ , the Euler system involves two-scale dynamics: acoustic waves propagates very quickly compared to material waves and the dynamics is asymptotically incompressible. Box 7 presents the its formal derivation.

#### Box 7: Low-Mach asymptotics of the Euler system

In this section, we recall the formal derivation of the low-Mach asymptotic dynamics (see also [19]). After space-time rescaling and nondimensionalization,

$$\tilde{x} = x/x_0, \quad \tilde{u} = u/u_0, \quad \tilde{t} = (u_0/x_0)t, \quad \tilde{\rho} = \rho/\rho_0, \quad \tilde{p} = p/p_0, \quad \tilde{E} = E/p_0,$$

the Euler system writes in nondimensional form (after removing the tildes):

$$\begin{aligned} \partial_t \rho + \nabla_x \cdot (\rho u) &= 0, \\ \partial_t (\rho u) + \nabla_x \cdot (\rho u \otimes u + \frac{p}{\text{Ma}^2} \text{Id}) &= 0, \\ \partial_t E + \nabla_x \cdot (E u + p u) &= 0, \\ E &= \text{Ma}^2 \rho \frac{|u|^2}{2} + \frac{p}{\gamma - 1}, \end{aligned}$$

with  $\text{Ma}^2 = u_0/(p_0/\rho_0)$ . Then plugging the following expansions,

$$\begin{aligned} \rho &= \rho_0 + \text{Ma} \rho_1 + O(\text{Ma}^2), & u &= u_0 + \text{Ma} u_1 + O(\text{Ma}^2), \\ E &= E_0 + \text{Ma} E_1 + O(\text{Ma}^2), & p &= p_0 + \text{Ma} p_1 + \text{Ma}^2 p_2 + O(\text{Ma}^2), \end{aligned}$$

we get, at order  $O(1/\text{Ma}^2)$  and  $O(1/\text{Ma})$ :  $\nabla_x p_0 = 0$  and  $\nabla_x p_1 = 0$ . Thus  $p_0, p_1$  are uniform in space. We can further assume that  $p_1$  equals zero (up to renaming  $p_0 + \text{Ma} p_1$  to  $p_0$ ) and we will further assume that boundary conditions impose  $p_0$  to be constant

<sup>1</sup>The adiabatic exponent for a monoatomic gaz equals  $\gamma = 1 + 2/3 = 5/3$ , while for a gaz made of diatomic molecules (like dinitrogen  $N_2$  and dioxygen  $O_2$ )  $\gamma = 1 + 2/5 = 7/5 = 1.4$ . The three-dimensionnal Boltzmann-BGK equation naturally leads to pressure law with  $\gamma = 5/3$ .

also in time. At  $O(1)$ , we get the following system

$$\begin{aligned}\partial_t \rho_0 + \nabla_x \cdot (\rho_0 u_0) &= 0, \\ \partial_t (\rho_0 u_0) + \nabla_x \cdot (\rho_0 u_0 \otimes u_0 + p_2 \text{Id}) &= 0, \\ \partial_t E_0 + \nabla_x \cdot (E_0 u_0 + p_0 u_0) &= 0, \\ p_0 &= (\gamma - 1)E_0.\end{aligned}$$

Using that  $p_0$  and thus  $E_0$  are constants, the energy equation shows that  $u_0$  is divergence free. The system thus simplifies into

$$\begin{aligned}\partial_t \rho_0 + u_0 \cdot \nabla_x \rho_0 &= 0, \\ \partial_t (\rho_0 u_0) + \nabla_x \cdot (\rho_0 u_0 \otimes u_0 + p_2 \text{Id}) &= 0, \\ \nabla_x \cdot u_0 &= 0, \\ p_0 &= (\gamma - 1)E_0.\end{aligned}$$

Therefore, the density is advected by the velocity  $u_0$ . The pressure  $p_2$  is the Lagrange multiplier associated with this constraint.

**Suliciu relaxation system.** The Suliciu relaxation system is a specific relaxation method for the Euler system, where there is only one additional variable: an auxiliary pressure  $\pi(x, t)$ . The extended system writes:

$$\begin{aligned}\partial_t \rho + \nabla_x \cdot (\rho u) &= 0, \\ \partial_t (\rho u) + \nabla_x \cdot (\rho u \otimes u + \pi \text{Id}) &= 0, \\ \partial_t E + \nabla_x \cdot (E u + \pi u) &= 0, \\ \partial_t (\rho \pi) + \nabla_x \cdot (\rho \pi u) + \lambda^2 \nabla \cdot u &= \frac{\rho}{\varepsilon} (p - \pi),\end{aligned}$$

where the pressure  $p$  has been replaced by  $\pi$  in the momentum and energy equations and the last equation mimics the pressure equation:  $\lambda > 0$  denotes the relaxation speed and  $\varepsilon > 0$  the relaxation rate. The left-hand side system is hyperbolic with all fields linearly degenerate. Hence, Riemann problems have explicit solutions. This results in an approximate Riemann solver for the Euler system [5]: at each time step and at each cell interface, the solution to the Riemann problems in the  $(\rho, \rho u, E, \rho \Pi)$  coordinates is solved after setting  $\Pi = p$ . The stability of the approximation (positivity and entropy preserving) is ensured by the subcharacteristic condition  $\lambda \geq \rho c$  and the CFL condition:

$$\Delta t \leq \Delta x / \max_x \max_i \{|\lambda_i(n(x))|\}$$

In the low-Mach regime  $\text{Ma} \ll 1$ , this stability condition becomes very stringent and seems all the more restrictive as the acoustic waves are small perturbations.

**Implicit Suliciu relaxation systems.** To overcome the CFL stability conditions, implicit solvers have to be considered. In [23], a classical splitting strategy has been proposed: the system is split into a convective part and an acoustic part, with a splitting parameter  $\alpha \in [0, 1]$  depending on the effective Mach number. Then the convective part is solved explicitly and the acoustic part implicitly. However, this second step requires to solve a non-homogeneous elliptic equation, where the diffusion coefficient depends on the density. Assembling a new matrix at

each iteration could have important CPU cost for 2D and 3D simulations. Moreover, the matrix may have a large condition number when the density has large spatial variations.

In [Imp4], we have proposed to introduce a **two-speed relaxation system** for the Euler equation following the strategy considered in the barotropic case [6]. Introducing both an auxiliary pressure  $\pi(x, t)$  and an auxiliary velocity  $v(x, t)$ , the relaxation system writes:

$$\begin{aligned}\partial_t \rho + \nabla_x \cdot (\rho v) &= 0, \\ \partial_t(\rho u) + \nabla_x \cdot (\rho u \otimes v + \pi \text{Id}) &= 0, \\ \partial_t E + \nabla_x \cdot (Eu + \pi v) &= 0, \\ \partial_t(\rho v) + \nabla_x \cdot (\rho v \otimes v) + \frac{a}{b} \nabla \pi &= \frac{\rho}{\varepsilon}(u - v), \\ \partial_t(\rho \pi) + \nabla_x \cdot (\rho \pi v) + ab \nabla \cdot v &= \frac{\rho}{\varepsilon}(p - \pi), \\ \partial_t a + v \cdot \nabla_x a &= 0, \\ \partial_t b + v \cdot \nabla_x b &= 0,\end{aligned}$$

where  $a(x, t) > 0$  and  $b(x, t) > 0$  are speed relaxation variables that are advected by the flow. When  $a = b$  and  $v = u$ , we recover the Suliciu relaxation model. This system has two main advantages: first, the diffusion matrix of the associated equivalent equation has a structure compatible with the low-Mach flows, since it involves only terms depending on the gradient of the pressure and the divergence of the velocity, which are small in this regime. Secondly, a **splitting strategy** can be applied between a convective, an acoustic and a relaxation part:

[Convective part]	[Acoustic part]	[Relaxation part]
$\partial_t \rho + \nabla_x \cdot (\rho v) = 0,$	$\partial_t \rho = 0,$	$\partial_t \rho = 0,$
$\partial_t(\rho u) + \nabla_x \cdot (\rho u \otimes v) + \alpha^2 \nabla_x \pi = 0,$	$\partial_t(\rho u) + (1 - \alpha^2) \nabla_x \pi = 0,$	$\partial_t(\rho u) = 0,$
$\partial_t E + \nabla_x \cdot (Eu) + \alpha^2 \nabla \cdot (\pi v) = 0,$	$\partial_t E + (1 - \alpha^2) \nabla \cdot (\pi v) = 0,$	$\partial_t E = 0,$
$\partial_t v + v \cdot \nabla_x v + \alpha^2 \frac{a}{b\rho} \nabla \pi = 0,$	$\partial_t v + (1 - \alpha^2) \frac{a}{b\rho} \nabla \pi = 0,$	$\partial_t v = \frac{\rho}{\varepsilon}(u - v),$
$\partial_t \pi + v \cdot \nabla_x \pi + \chi_c \frac{ab}{\rho} \nabla \cdot v = 0,$	$\partial_t \pi + (1 - \alpha^2) \chi_a \frac{ab}{\rho} \nabla \cdot v = 0,$	$\partial_t \pi = \frac{1}{\varepsilon}(p - \pi),$

with  $\alpha = \text{proj}_{[M_{\min}, 1]}(\text{Ma}) = \max(M_{\min}, \min(\text{Ma}, 1))$  and  $\chi_c + (1 - \alpha^2)\chi_a = 1$ . The splitting of the  $\pi$  equation is slightly different from the splitting of the other equations. The main point here is that, as auxiliary variables, the relaxation speeds  $a$  and  $b$  can be reinitialized at each time step, such that  $a/(b\rho)$  and  $(ab/\rho)$  are constant in space. The resolution of the acoustic part can then be made **using an implicit numerical scheme**:

$$\pi^{n+1} - \Delta t^2 (1 - \alpha^2)^2 \chi_a \left( \frac{a}{b\rho} \right) \left( \frac{ab}{\rho} \right) \text{Div}_h (\nabla_h \pi^{n+1}) = \pi^n$$

where  $\text{Div}_h$  and  $\nabla_h$  refers to finite difference approximation of the divergence and gradient operators. This implicit linear system has **constant coefficients**. Regarding the convective part, it can be solved using an explicit Finite-Volume method with the local Lax-Friedrichs (Rusanov) numerical flux for the conservative parts and the upwind flux for the non-conservative parts. Finally, the relaxation part can be approximated by a simple projection.

As exposed in [Imp4], entropic stability criteria on the convective and acoustic parts impose conditions on the relaxation parameters:

$$a \geq b, \quad \chi_c \frac{ab}{\rho^2} \geq (1 + \alpha^2(\gamma - 1)) \frac{p}{\rho}, \quad \chi_a \frac{ab}{\rho^2} \geq (\gamma - 1) \frac{p}{\rho}.$$

Supposing that  $a/\rho$  and  $b$  are constant and using the change of variables  $a = \rho\lambda$  and  $b = \phi\lambda$ , with  $\lambda, \phi > 0$ , the inequalities become:

$$\rho \geq \phi, \quad \phi(\chi_c \lambda^2) \geq \frac{(1 + \alpha^2(\gamma - 1))}{\gamma} \rho c^2, \quad \phi(\chi_a \lambda^2) \geq \frac{(\gamma - 1)}{\gamma} \rho c^2. \quad (48)$$

Thus  $\varphi$  is chosen as a lower bound of the density and  $\lambda$  large enough so as to satisfy the two other equations. In practice,  $\lambda_c^2 = \chi_c \lambda^2$  and  $\lambda_a^2 = \chi_a \lambda^2$  are rather set separately so as to satisfy (48). This is the reason why parameters  $\chi_c$  and  $\chi_a$  have been introduced. Then  $\lambda^2$  is determined such that  $\lambda^2 = \lambda_c^2 + (1 - \alpha^2)\lambda_a^2$  and  $\chi_c$  and  $\chi_a$  are determined by the relation We note that, defined in this way,  $a$  and  $b$  are of the order  $O(1/\text{Ma})$ .

**Numerical results.** We first consider a moving smooth contact wave with constant velocity. The initial data are given by

$$\rho_{\text{init}}(x) = 0.1 + \frac{1}{5\sigma\sqrt{\pi}} e^{-\|x-x_0\|^2/2\sigma^2}, \quad u_{\text{init}}(x) = (0.01, 0.01), \quad p_{\text{init}}(x) = 1,$$

on the domain  $\Omega = [0, 1]^2$  with  $\sigma = 0.05$  and the perfect gas law ( $\gamma = 1.4$ ). The solution to this problem consists in the advection of the density at constant velocity  $u$ . The numerical solution is computed up to time  $T = 1$ . For the relaxation parameters, we take

$$\lambda_c = 3.2, \quad \lambda_a = 2.5, \quad \phi = 0.099.$$

As expected, it is shown that for a similar accuracy the proposed method, called ‘‘SI two-speed scheme’’ hereafter, enables us to take a 40 times larger timestep than the one of the explicit scheme. The time step is also 1.5 times larger than the one used by the semi-implicit Suliciu scheme, denoted ‘‘SI Suliciu’’. Moreover, Table III.3 shows that by removing the spatial dependency on the density in the elliptic solver, the SI two-speed relaxation method enables a division by 3.3 of the condition number (at a given time step) compared with the SI Suliciu.

Figure 9 presents different results for the Gresho vortex test case. The initial condition is given by

$$\rho_{\text{init}}(x) = 1, \quad (u_{\text{init}}(x), p_{\text{init}}(x)) = \begin{cases} (5r \mathbf{e}_\phi, & p_c + \frac{25}{2}r^2), & \text{for } r \leq 0.2, \\ ((2 - 5r) \mathbf{e}_\phi, & p_c + 4 \ln(5r) + 4 - 20r + \frac{25}{2}r^2), & \text{for } 0.2 < r \leq 0.4, \\ (0, & p_c + 4 \ln 2 - 2), & \text{for } r > 0.4, \end{cases}$$

on the domain  $\Omega = [0, 1]^2$  with  $r = \|x - x_0\|$ ,  $x_0 = (0.5, 0.5)$  and  $p_c = 1/(\gamma \text{Ma}^2) - 1/2$  with a perfect gas law with  $\gamma = 1.4$ . The Mach number  $\text{Ma}$  will take the different values 0.5, 0.1, 0.01 or 0.001. Then, since the initial velocity is divergence free, the initial data is well-prepared. Indeed the above initial data is a steady solution to the incompressible Euler system, as well as a steady solution to the compressible one. Figure 9 shows 2D plots of the norm of the velocity at final time  $T = 1$  as well as 1D cuts of the final density and initial density obtained with the explicit scheme (top) and the SI two-speed scheme (bottom). We observe that with

a larger time-step independent of the Mach number, the SI two-speed scheme preserves the structure of the velocity field, contrary to the explicit scheme. More precisely, the accuracy of the classical scheme decreases with the Mach number and the CPU cost increases since the time step decreases. Concerning the SI two-speed scheme, the accuracy on the density is almost the same whatever the Mach number is, and the CPU cost remains almost constant as the time step is not modified. The CPU cost can be still impacted by the condition number of the implicit operator, which could be mitigated using preconditioning technique.

	$\lambda_{\max}$	$\Delta t$
Explicit	$\max( (u \cdot n) \pm c )$	0.0021
SI Suliciu	$\max( (u \cdot n) \pm \alpha \lambda_c / \rho )$	0.056
SI two-speed	$\max( (u \cdot n) \pm \alpha \lambda_c )$	0.088

Table III.2: (Smooth contact wave, constant velocity, 2d) Maximal characteristic speed  $\lambda_{\max}$  of the explicit part and timestep  $\Delta t$  used for the different numerical schemes, for  $N = 200$  cells.

	$\Delta t$	conditioning
Explicit	0.0575	2400
SI Suliciu	0.088	1750
SI two-speed	0.057	720

Table III.3: (Smooth contact wave, constant velocity, 2d) Condition numbers for the two semi-implicit relaxation scheme for  $N = 200$  cells.

## 4 Conclusion and perspectives

This chapter presents two implicit schemes for hyperbolic equations. The first one is based on a vectorial kinetic approach and can be applied to any hyperbolic system. The second one is a two-speed Suliciu relaxation system and is dedicated to the Euler system.

We identify two main perspectives of these works:

- **Extension to more complex plasma test cases.** In order to deal with real physical test cases, the code already developed for the three-dimensional guiding-center model could be extended to simulate full gyrokinetic models, where there is an additional velocity dimension. Additionally, the fourth-order Lattice Boltzmann scheme, developed [Imp1], could be used to increase the accuracy of the toroidal transport.
- **Data-driven optimization of the relaxation schemes.** An interesting avenue concerns the improvement of the relaxation methods using neural networks. The method requires the relaxation parameter to be chosen optimally in order to obtain the greatest possible accuracy while remaining stable. The choice of this parameter could be learnt by neural networks. An initial study was carried out during the PhD thesis of Léo Bois [3] on the transport, the Burgers and the linear wave equations. It will also be interesting to take a more detailed account of the non-linear stability criteria [21].

In this conclusion, we also mention the related semi-implicit methods for fluid models with congestion. These models involve a singular pressure that maintains the density below a given maximal value. By increasing the stiffness of the pressure, the model leads to the coexistence of incompressible dynamics, in regions of maximal density, and compressible dynamics in regions of lower density. Based on the methodology exposed in [Phd5]. In [Imp8], we develop a second order (RK2-Crank Nicolson) scheme for the pedestrian model introduced in [11] and performed two-dimensional evacuation simulations. In [Imp2], in addition to the theoretical analysis of the hard congestion limit, we propose a numerical scheme for the one-dimensional dissipative Aw-Rascle model and performed numerical comparisons with the asymptotic congested Euler model, whose exact solutions can be obtained thanks to an optimal transport formulation [28].

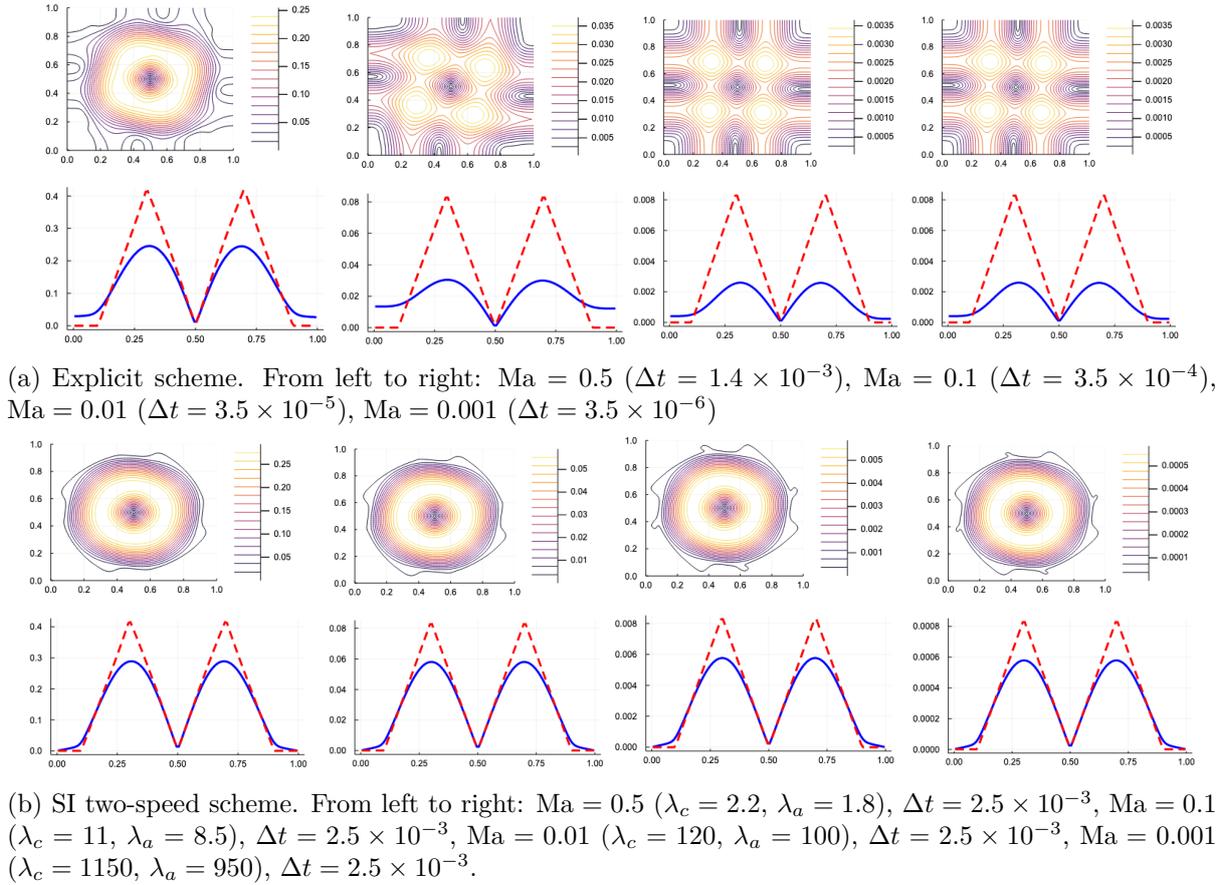


Figure 9: (Gresho vortex) Norm of the velocity (2D plot) at final time  $T = 1$  and initial (red) and final (blue) density profiles for the explicit (top) and the SI two-speed scheme (bottom) for different values of the Mach number.

The same numerical method has been introduced independently in [1] and have been extended in two dimensions there. My collaborators on this topic are Nilasis Chaudhuri (Warsaw Univ.), Pierre Degond (CNRS, Univ. Toulouse), Piotr Minakowski (Heidelberg Univ.), Charlotte Perrin (Aix-Marseille Univ.), Ewelina Zatorska (Imperial College, London)

## References

- [1] P. Aceves-Sánchez, R. Bailo, P. Degond, and Z. Mercier. “Pedestrian models with congestion effects”. In: *ath. Models Methods Appl. Sci.* (2024).
- [2] D. Aregba-Driollet and R. Natalini. “Discrete Kinetic Schemes for Systems of Conservation Laws”. In: *Hyperbolic Problems: Theory, Numerics, Applications*. Basel: Birkhäuser Basel, 1999, pp. 1–10.
- [3] C. Augonnet, S. Thibault, R. Namyst, and P.-A. Wacrenier. “StarPU: A Unified Platform for Task Scheduling on Heterogeneous Multicore Architectures”. In: *CCPE - Concurrency and Computation: Practice and Experience, Special Issue: Euro-Par 2009 23.2* (2011), pp. 187–198.
- [3] L. Bois. “Méthodes numériques basées sur l’apprentissage pour les EDP hyperboliques et cinétiques”. PhD thesis. Université de Strasbourg, 2023. ([tel-04589397v1](tel:04589397v1)).

- [4] F. Bouchut. “Construction of BGK models with a family of kinetic entropies for a given system of conservation laws”. In: *J. Stat. Phys.* 95.1-2 (1999), pp. 113–170.
- [5] F. Bouchut. *Nonlinear Stability of Finite Volume Methods for Hyperbolic Conservation Laws and Well-Balanced Schemes for Sources*. Frontiers in Mathematics. Birkhäuser Basel, 2004.
- [6] F. Bouchut, C. Chalons, and S. Guisset. “An entropy satisfying two-speed relaxation system for the barotropic Euler equations: application to the numerical approximation of low Mach number flows”. In: *Numer. Math.* 145 (2020), pp. 35–76.
- [7] B. Bramas, P. Helluy, L. Mendoza, and B. Weber. “Optimization of a discontinuous Galerkin solver with OpenCL and StarPU”. In: *nt. J. Finite Vol.* 15.1 (2020), pp. 1–19.
- [8] B. Bramas and A. Ketterlin. “Improving parallel executions by increasing task granularity in task-based runtime systems using acyclic DAG clustering”. In: *PeerJ Comput. Sci.* 6 (2020), e247.
- [9] G.-Q. Chen, C.D. Levermore, and T.-P. Liu. “Hyperbolic conservation laws with stiff relaxation terms and entropy”. In: *Commun. Pure Appl. Math.* 47.6 (1994), pp. 787–830.
- [10] F. Coron and B. Perthame. “Numerical passage from kinetic to fluid equations”. In: *SIAM J. Numer. Anal.* 28.1 (1991), pp. 26–42.
- [11] P. Degond, P. Minakowski, and E. Zatorska. “Transport of congestion in two-phase compressible/incompressible flows”. In: *Nonlinear Anal. Real World Appl.* 42 (2018), pp. 485–510.
- [12] P.J. Dellar. “An interpretation and derivation of the lattice Boltzmann method using Strang splitting”. In: *Comput. Math. Appl.* 65.2 (2013), pp. 129–141.
- [13] G. Dimarco and L. Pareschi. “Asymptotic preserving implicit-explicit Runge–Kutta methods for nonlinear kinetic equations”. In: *SIAM J. Numer. Anal.* 51.2 (2013), pp. 1064–1087.
- [14] F. Dubois. “Equivalent partial differential equations of a lattice Boltzmann scheme”. In: *Comput. Math. Appl.* 55.7 (2008), pp. 1441–1449.
- [15] F. Dubois. “Stable lattice Boltzmann schemes with a dual entropy approach for monodimensional nonlinear waves”. In: *Comput. Math. Appl.* 65.2 (2013), pp. 142–159.
- [16] C. Flint, K. Guillon, R. Hélie, and P. Helluy. “Schémas cinétiques: aspects théoriques et optimisations informatiques”. In: *{hal-03986533v2}* (2022).
- [17] B. Graille. “Approximation of mono-dimensional hyperbolic systems: A lattice Boltzmann scheme as a relaxation method”. In: *J. Comput. Phys.* 266 (2014), pp. 74–88.
- [18] V. Grandgirard, J. Abiteboul, J. Bigot, T. Cartier-Michaud, N. Crouseilles, G. Dif-Pradalier, C. Ehrlacher, D. Esteve, X. Garbet, P. Ghendrih, et al. “A 5D gyrokinetic full-f global semi-Lagrangian code for flux-driven ion turbulence simulations”. In: *Comput. Phys. Commun.* 207 (2016), pp. 35–68.
- [19] H. Guillard and C. Viozat. “On the behaviour of upwind schemes in the low Mach number limit”. In: *Comput. Fluids* 28.1 (1999), pp. 63–86.
- [20] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*. Vol. 31. Springer Science & Business Media, 2006.
- [21] R. Hélie and P. Helluy. “Stable Second Order Boundary Conditions for Kinetic Approximations”. In: *Finite Volumes for Complex Applications X—Volume 2, Hyperbolic and Related Problems*. Ed. by Emmanuel Franck, Jürgen Fuhrmann, Victor Michel-

- Dansac, and Laurent Navoret. Springer Nature Switzerland, 2023, pp. 161–170. ISBN: 978-3-031-40860-1.
- [22] R. Hélié and P. Helluy. “Stable second order boundary conditions for kinetic approximations”. In: *International Conference on Finite Volumes for Complex Applications*. Springer. 2023, pp. 161–170.
- [23] D. Iampietro, F. Daude, P. Galon, and J.-M. Hérard. “A Mach-sensitive implicit–explicit scheme adapted to compressible multi-scale flows”. In: *J. Comput. Appl. Math.* 340 (2018), pp. 122–150.
- [24] S. Jin. “Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: a review”. In: *Lecture notes for summer school on methods and models of kinetic theory (M<sup>3</sup>MKT), Porto Ercole (Grosseto, Italy)* (2010), pp. 177–216.
- [25] S. Jin and Z. Xin. “The relaxation schemes for systems of conservation laws in arbitrary space dimensions”. In: *Commun. Pure Appl. Math.* 48.3 (1995), pp. 235–276.
- [26] R.I. McLachlan and G.R.W. Quispel. “Splitting methods”. In: *Acta Numer.* 11 (2002), pp. 341–434.
- [27] R. Natalini. “A Discrete Kinetic Approximation of Entropy Solutions to Multidimensional Scalar Conservation Laws”. In: *J. Differ. Equat.* 148.2 (1998), pp. 292–317.
- [28] C. Perrin and M. Westdickenberg. “One-dimensional granular system with memory effects”. In: *SIAM J. Math. Anal.* 50.6 (2018), pp. 5921–5946.
- [29] B. Perthame. “Boltzmann type schemes for gas dynamics and the entropy property”. In: *SIAM J. Numer. Anal.* 27.6 (1990), pp. 1405–1421.

---

## Chapter IV

# Models for cells collective dynamics

---

## 1 Introduction

In this chapter, we present some contributions to the modeling of cell collective dynamics. They range from the development of relevant new particle models to the derivation of associated large-scale models and comparisons with experimental results.

Collective cell movements are observed in several biological processes, like morphogenesis, tissue remodeling, bacterial migration or collective swimming of spermatozoa. These collective movements spontaneously emerge at large scale without any leader: they result from the numerous interactions between cells.

Cells are complex entities involving complex internal physical and chemical processes. In particular, they are capable of extracting energy from their environment and converting it into a mechanical rearrangement of their cytoskeletons. As a result, many cells are active particles that can exhibit persistent movement: this is self-propulsion.

One possible description of their dynamics is provided by particle models: each particle is represented by its position  $X_i(t) \in \mathbb{R}^d$ , its velocity  $V_i(t) \in \mathbb{R}^d$ , and possibly its polarity vector  $P_i(t) \in \mathbb{R}^d$  and other geometric quantities. Self-propulsion of the cells generally leads us to assume that the velocity norms remain constant over time. Interactions between cells can be of the attraction-repulsion type, but also of the velocity or polarity alignment type. These alignment interactions can be viewed as a mesoscopic description of the internal activity of the cytoskeletons. They are generally modeled through Vicsek-type alignment interactions. We refer to Box 8 for a short presentation of this model, which will be central to our work.

Particle models have the advantage of being quite flexible: modelling assumptions, like specific cell geometry or specific interactions rules, can be inserted fairly easily. They also has the advantage to allow the rigorous derivation of the large scale hydrodynamic models in some regimes. They can thus serves in finding the key dynamical ingredients to explain the observed collective motions [12, 10, 9]. Let us mention that other modeling approaches for tissues also exist, either at the cellular scale with vertex models [16, 5] or directly at the macroscopic scale [18].

In what follows, we present three models with three distinct features: (i) disc-shaped cells organised in layers, (ii) anti-alignment interaction rules, (iii) self-confinement with an actin cable. For the first two models, the associated macroscopic models were derived. For the third, extensive comparisons with experimental results were carried out.

**Box 8: Vicsek Model**

The Vicsek model, introduced in the seminal paper [26], consider particles moving at constant speed interacting with velocity alignment interactions in competition with noise. The continuous version of the model, introduced in [12], reads:

$$\frac{d}{dt}X_i = cV_i, \quad dV_i = \text{Proj}_{V_i^\perp} \circ \left( \nu (\bar{V}_i - V_i)dt + D (dB_t)_i \right), \quad \bar{V}_i = \frac{\sum_{j, |X_j - X_i| \leq R} V_j}{\left| \sum_{j, |X_j - X_i| \leq R} V_j \right|},$$

where  $\bar{V}_i$  is the normalized local averaged velocity of the neighboring particles in a neighborhood of radius  $R$ . In this model, the velocities belongs to the sphere  $\mathbb{S}^{d-1} = \{V \in \mathbb{R}^d \mid \|V\| = 1\}$ . This property is preserved over time thanks to the projection operator onto  $V_i^\perp$ .  $(dB_t)$  refers to a Gaussian white noise and the operator  $\circ$  indicates that the stochastic differential equation has to be taken in the Stratonovitch sens.<sup>a</sup> This model involves three parameters: the particle speed  $c > 0$ , the alignment rate between velocities  $\nu > 0$ , and the angular diffusion of velocities  $D > 0$ .

The statistical properties of this model have been intensively studied in the physical community to characterize the observed phase transition between an ordered-liquid phase, a disordered-gas phase and a coexistence liquid-gas phase with bands like dynamics. We refer to [2] for a review.

To analyze the associated large-scale dynamics, a mean-field version of the dynamics have been considered in [12]. It results in a Fokker-Planck equation satisfied by the phase-space distribution function of the particles  $f(t, x, v)$ . This mean-field version of the model is obtained in the limit of large number of particles [1]. From this mean-field model, a hydrodynamic model has been derived: this is the so-called Self-Organized Hydrodynamics (SOH) model. We refer to Box 9.

<sup>a</sup>Integrating the equation in the Stratonovitch sens consists into considering the limit of the following discretization scheme:  $V_i^{n+1} - V_i^n = \text{Proj}_{((V_i^n + V_i^{n+1})/2)^\perp} (\nu(\bar{V}_i^n - V_i^n)\Delta t + (B_{t^{n+1}} - B_{t^n})_i)$ . Multiplying by  $(V_i^{n+1} + V_i^n)$ , we get that the norm is preserved:  $|V_i^{n+1}| = |V_i^n|$ .

**Box 9: SOH Model**

The Self-Organized Hydrodynamic (SOH) model has been first derived in [12]. It describes the dynamics of the density  $\rho(x, t) \in \mathbb{R}$  and the mean direction  $u(x, t) \in \mathbb{S}^{d-1}$  and writes

$$\begin{aligned} \partial_t \rho + c_1 \nabla_x \cdot (\rho u) &= 0, \\ \rho (\partial_t u + c_2 u \cdot \nabla_x u) + \kappa^{-1} \text{Proj}_{u^\perp} \nabla \rho &= 0. \end{aligned}$$

where  $\kappa = \nu/D$  and the constants  $c_1, c_2$  are given by:

$$c_1 = \frac{\int_0^\pi \cos \theta \exp(\kappa \cos \theta) d\theta}{\int_0^\pi \exp(\kappa \cos \theta) d\theta} \geq c_2 = \frac{\int_0^\pi \cos \theta (\sin \theta)^{d-1} h_\kappa(\cos \theta) \exp(\kappa \cos \theta) d\theta}{\int_0^\pi (\sin \theta)^{d-1} h_\kappa(\cos \theta) \exp(\kappa \cos \theta) d\theta} > 0.$$

with  $h_\kappa(\cos \theta)$  is a given function, which is a solution to a second order differential equation. This system is hyperbolic. However, in contrast with the Euler equation for classical fluids, it is **non-Galilean invariant** as soon as  $c_1 \neq c_2$ , i.e.  $(\rho(x + Vt, t), u(x + Vt, t) - V)$  does not solve the system for non-zero constant velocity  $V \in \mathbb{R}^d$ . This also reflects the fact that the associated sound speed is not isotropic. Due to the projection

term  $\text{Proj}_{u^\perp}$ , this model is also **non-conservative**. In particular, there is no clear definition for shock speeds for discontinuous solutions. In [Phd3], a numerical scheme based on a relaxation method has been proposed. Although the approximated system has a restricted hyperbolic region, the designed numerical scheme is able to capture the dynamics associated with the underlying particle system. We refer to [Phd6] for a review of this model.

## 2 A multi-layer model for interacting disks

In collaboration with P. Degond (CNRS - Univ. Toulouse), we have proposed a modeling framework for the three-dimensional motion of spermatozoa [Cell6]. At a large scale, spermatozoa show collective movements: dark waves are observed and these motions are currently used as a measure of fertility. These dark waves are thought to be linked with the inclination of the spermatozoa heads relative to the observer. Indeed, spermatozoa have flat ellipsoidal heads, which is modeled as **infinitely thin flat discs**.

There is some experimental evidence that sperm-cell motion in the most common experiments is mostly planar. We thus propose a **multi-layer model** where cells remain in a given layer and interact with the cells of the same layer and the cells of the two neighboring ones.

Although cells exhibit hydrodynamic interactions with the surrounding liquid, the predominant induction result from the **volume exclusion interactions** between cells. Indeed, sperm cell concentration is very high (as large as  $5 \times 10^9 \text{ cm}^{-3}$ ) and they occupy a volume ratio of about 50%. As shown in [21], volume exclusion interactions result in **local alignment** for self-propelled elongated particles. This has been also experimentally demonstrated for spermatozoa dynamics in [4].

**Particle model.** We consider horizontal layers with inter-distance  $d$ . Each spermatozoon is described by its altitude index  $h_k \in \mathbb{N}$  (constant over time), its position  $X_k(t) \in \mathbb{R}^2$ , its angle velocity  $\varphi_k(t) \in [0, 2\pi)$  and its inclination with respect to the vertical axis  $\theta_k(t) \in [0, \pi)$ . Then the particle model reads:

$$\begin{aligned} \frac{dX_k}{dt} &= V(\varphi_k), \\ d\varphi_k &= -\nu \sin(\varphi_k - \bar{\varphi}_k^{\text{tot}}) dt + \sqrt{2D} dB_t^{\varphi,k} \quad \text{modulo } 2\pi, \\ d\theta_k &= (-K \sin(2(\theta_k - \bar{\theta}_k)) + \mu T_k) dt + \sqrt{2\delta} dB_t^{\theta,k} \quad \text{modulo } \pi. \end{aligned}$$

where  $V(\varphi_k) = (\cos \varphi_k, \sin \varphi_k)^T$ . The dynamics of positions and velocity angles are very similar to the original Vicsek model (see Box 8), expressed in the variable  $(x, \varphi)$ , except that the averaged velocity direction  $\bar{\varphi}_k^{\text{tot}}$  consider the discs of the same layer and the two neighboring layers only. The dynamics of the inclination angles result from the torque balance equation of the discs in the overdamped regime. There are two main ingredients: there is a local alignment for the discs of the same layer, with rotational stiffness  $K$ , and particles of the two neighboring layers exert a torque  $T_k$  defined as follows:

$$T_k = \sum_{\substack{h_j = h_k \pm 1 \\ |X_k - X_j| \leq R}} T_{kj}, \quad T_{kj} = \text{sign}(h_k - h_j) (V(\varphi_k) \times V(\varphi_j)) \cdot \hat{z}.$$

Angular diffusion of parameter  $\delta > 0$  is also considered to the inclination dynamics. To summarize, this model combines a two-dimensional **Vicsek alignment** dynamics in layers and

**nematic alignment dynamics** for inclinations. The latter has been further studied in [11]. Let mention that other three-dimensional interactions could be considered like the body-attitude model [13, 6].

**Mean-field kinetic equation.** In order to study the large-scale dynamics of the proposed model, we are interested in the dynamics of the phase-space distribution  $f(x, \varphi, \theta, h, t)$  with  $x \in \mathbb{R}^2, \varphi \in [0, 2\pi), \theta \in [0, \pi)$  and  $h \in \mathbb{N}$ . Formally, this distribution satisfies the following equation:

$$\partial_t f + \nabla_x \cdot (V(\varphi)f) = Q_{\text{nb}}(f) + Q_{\text{in}}(f)$$

where we have split the interaction of the neighboring layers,

$$Q_{\text{nb}}(f) = -\partial_\theta(\mu T_f f) - \partial_\varphi(S_f f)$$

from the interactions that take place inside layers:

$$Q_{\text{in}}(f) = -\partial_\varphi(-\nu \sin(\varphi - \bar{\varphi}_f)f) + D \partial_\varphi^2 f - \partial_\theta(-K \sin(2(\theta - \bar{\theta}_f))f) + \delta \partial_\theta^2 f.$$

where  $S_f, T_f \in \mathbb{R}, \bar{\varphi}_f \in [0, 2\pi), \bar{\theta}_f \in [0, \pi)$  are mean-field approximations of the associated quantities at the particle scale. Then, looking at the large time dynamics and space scale, of size  $O(1/\varepsilon)$  with  $\varepsilon \ll 1$ , and supposing that the interaction frequency inside layers are much larger interaction frequency between neighboring layers ( $\mu, \beta$  are  $O(\varepsilon)$ ), the distribution function  $f^\varepsilon(x, \varphi, \theta, h, t) = f(x/\varepsilon, \varphi, \theta, h, t/\varepsilon)$  satisfies the equation:

$$\partial_t f^\varepsilon + \nabla_x \cdot (V(\varphi)f^\varepsilon) = Q_{\text{nb}}(f^\varepsilon) + \frac{1}{\varepsilon} Q_{\text{in}}(f^\varepsilon).$$

**Equilibria and macroscopic equations.** The set of local equilibria associated with the non-linear tensorial Fokker-Planck operator  $Q_{\text{in}}$  is given by:

$$\mathcal{E} = \{f, Q_{\text{in}}(f) = 0\} = \left\{ \rho M_{\bar{\varphi}, \bar{\theta}} \mid \rho \in \mathbb{R}^+, \bar{\varphi} \in [0, 2\pi], \bar{\theta} \in [0, \pi] \right\}$$

with  $M_{\bar{\varphi}, \bar{\theta}}(\varphi, \theta)$  the following function

$$M_{\bar{\varphi}, \bar{\theta}}(\varphi, \theta) = \frac{1}{Z_{\kappa, \lambda}} \exp(\kappa \cos(\varphi - \bar{\varphi})) \exp(\lambda \cos(2(\theta - \bar{\theta}))), \quad (49)$$

where  $\kappa = \nu/D$  and  $\lambda = K/\delta$  and  $Z_{\kappa, \lambda}$  is a normalization constant.

Supposing that the distribution function  $f^\varepsilon$  tends to  $f$  as  $\varepsilon \rightarrow 0$ , then  $f$  belongs to the kernel of  $Q_{\text{in}}$ . This means that there exists macroscopic quantities  $\rho(x, h, t) \geq 0, \bar{\varphi}(x, t) \in [0, 2\pi)$  and  $\bar{\theta}(x, t) \in [0, \pi)$  such that

$$f(x, \varphi, \theta, h, t) = \rho(x, h, t) M_{\bar{\varphi}(x, h, t), \bar{\theta}(x, h, t)}(\varphi, \theta).$$

To determine the dynamics of these quantities, we consider the so-called generalized collisional invariants associated with  $Q_{\text{in}}$  as introduced in [12]: these are functions  $\psi_{(\bar{\varphi}, \bar{\theta})}(\varphi, \theta)$  orthogonal to  $Q_{\text{in}}(f)$ , for any distribution  $f$  with mean velocity and mean inclination  $(\bar{\varphi}, \bar{\theta})$ . We refer to [Cell6] for a detailed description. After integrating the equation against these generalized collisional invariants, we get the following equations:

$$\partial_t \rho + c_1 \nabla_x \cdot (\rho V(\bar{\varphi})) = 0, \quad (50)$$

$$\rho \left( \partial_t \bar{\varphi} + c_2 (V(\bar{\varphi}) \cdot \nabla_x) \bar{\varphi} \right) + \frac{1}{\kappa} V(\bar{\varphi})^\perp \cdot \nabla_x \rho = \frac{\nu \beta'}{c_3} V(\bar{\varphi})^\perp \cdot \mathcal{S}, \quad (51)$$

$$\rho \left( \partial_t \bar{\theta} + c_1 (V(\bar{\varphi}) \cdot \nabla_x) \bar{\theta} \right) = \frac{\mu'}{c_4} (c_1 \rho V(\bar{\varphi}))^\perp \cdot \mathcal{T}, \quad (52)$$

where the constants  $c_1, c_2, c_3, c_4$  are weighted averages of the equilibria. On the left-hand side, this system combines a SOH model (Eqs.(50)-(51), see also Box 9) and a transport equation for the mean inclination (Eq.(52)). Interactions between layers appear on the right-hand side.

**Numerical results: metastable states.** Figure 10 presents numerical comparisons between the initial particle model and the macroscopic one for a homogeneous simulation. We consider 3 layers of size  $[0, L_x] \times [0, L_y]$  with 25000 particles per layer, uniformly distributed. Particle velocity and inclination angles are randomly distributed according to their respective von Mises distribution (49), with mean velocity angles and mean inclination angles:

$$\bar{\varphi}^1 = -1, \quad \bar{\varphi}^2 = 1, \quad \bar{\varphi}^3 = -2, \quad (53)$$

$$\bar{\theta}^1 = 0, \quad \bar{\theta}^2 = -0.9, \quad \bar{\theta}^3 = -0.8. \quad (54)$$

For the macroscopic model, the initial density is thus constant in each layer given by  $\rho^\ell = (25000)/(L_x L_y)$ , for  $\ell \in \{1, 2, 3\}$ , and the uniform initial values of  $\bar{\varphi}^\ell$  and  $\bar{\theta}^\ell$  given by (53) - (54). A time rescaling is performed in the microscopic model to match the macroscopic time scale:  $\tilde{t} = t/\varepsilon$  with  $\varepsilon = 0.1$ .

Fig. 10 depicts the time evolution of the mean velocity and mean inclination for both microscopic simulations (dotted line) and macroscopic simulations (continuous line), with different interlayer distances  $h$ . We first consider strong interactions between layers by choosing  $h = R$  (Fig. 10a). The dynamics can be divided into two phases: a first relaxation of the mean velocities of layers 1 and 2 and then a second relaxation dynamics between layers 1 and 3 that leads to the alignment of the three layers. We note that microscopic and macroscopic simulations coincide during the first 1.5 time unit (including the first relaxation mechanism), which confirms that the macroscopic model captures the right interaction time scale. However, we see that, due to the finite number of particles, stochastic fluctuations make the second relaxation occur earlier around time  $t = 3$  (micro) instead of  $t = 3.5$  (macro). We also see that the time of the second relaxation depends on the simulation, always occurs before the macroscopic relaxation and strongly determines the final inclinations. When increasing the interlayer distance (Fig. 10b), some inclination configurations prevent layers from interacting. We observe that, in the macroscopic simulation, this slight increase of the interlayer distance results in large-time translation of the second relaxation step going from  $t = 3.5$  to  $t = 7.5$ . This highlights the metastability of the system between the two relaxation steps. This is in contrast with the microscopic simulations where this increase does not result in so much change of the second relaxation time. This second relaxation time strongly depends on the stochastic fluctuations, which make particle interacting instead of remaining in a non-interacting configuration.

### 3 Asymmetric interactions and phase transitions

In an ongoing work with N. Meunier (Univ. Évry) and R. Voituriez (CNRS - Sorbonne Univ.), we are interested in the dynamics of a Vicsek-type model with asymmetric alignment interactions. The effect of the visual field has been already studied in [15]. In this work, we further investigate this by adding an anti-alignment when the particles are back to back. This is motivated by the so-called Contact Inhibition of Locomotion (CIL) reported in the literature which is expected to disadvantage the formation of largescale cell clusters with coherent motion (see for instance [24]).

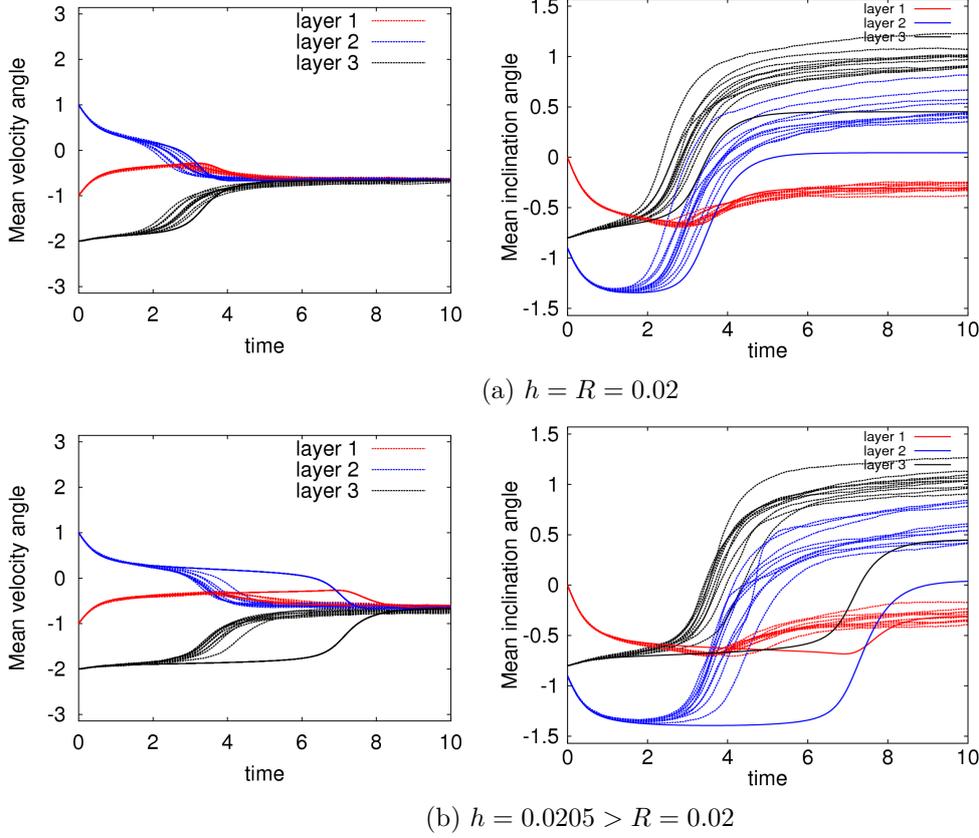


Figure 10: (Homogeneous case, interacting layers,  $h = R$ ) Comparison of macroscopic (continuous line) and microscopic (dashed lines) simulations. Left: Mean velocity angles  $\bar{\varphi}^\ell$  as a function of time. Right: Mean inclination angles  $\bar{\theta}^\ell$  as a function of time. Alignment interaction parameters:  $\nu = 4$ ,  $D = 0.1$ ,  $K = 4$ ,  $\delta = 0.1$ ,  $R = 0.02$ . Layer-interaction parameters:  $R = 0.02$ ,  $\beta = 0.5$ . Domain parameters:  $L_x = L_y = 1$ . Microscopic parameters: 25000 particles per layer,  $\mu = 3$ ,  $\Delta t = 1 \times 10^{-2}$ . Macroscopic parameters:  $\Delta x = \Delta y = 0.5$ ,  $\Delta t = 1 \times 10^{-2}$ . Taken from [Cell6].

**Particle model.** We study the two-dimensional dynamics of  $N$  particles, described by their positions  $X_k(t) \in \mathbb{R}^2$  and velocities  $V_k(t) \in \mathbb{S}^1$ . They satisfy the Vicsek-type model

$$dX_k = c V_k dt, \quad (55)$$

$$dV_k = P_{V_k^\perp} \circ (\nu (W_k - V_k) dt + \sqrt{2D} dB_t^k). \quad (56)$$

where the vector  $W_k \in \mathbb{R}^2$  is defined by:

$$W_k = \frac{\alpha}{N} \sum_{\substack{j, |X_j - X_k| < R \\ (X_j - X_k) \cdot V_k > 0 \text{ and/or} \\ (X_k - X_j) \cdot V_j > 0}} V_j + \frac{(1 - \alpha)}{N} \sum_{\substack{j, |X_j - X_k| < R \\ (X_j - X_k) \cdot V_k < 0 \text{ and} \\ (X_k - X_j) \cdot V_j < 0}} -V_j. \quad (57)$$

In this expression, the first-term models the fact that two neighboring particles tend to align their velocities if one is in the field of vision of the other, while the second term models the fact that they tend to anti-align their velocities otherwise.

**Mean-field kinetic model.** The mean-field kinetic model associated with this particle model writes:

$$\partial_t f + \nabla_x \cdot (c v f) = -\nabla_v \cdot \left( P_{v^\perp} F[f] f \right) + D \Delta_v f, \quad (58)$$

where  $F[f]$  denotes the mean-field interaction force:

$$F[f](x, v, t) = \nu \int_{y \in \mathbb{R}^2, w \in \mathbb{S}^1} K(|x - y|, (y - x) \cdot v, (x - y) \cdot w) w f(y, w, t) dy dw.$$

with interaction kernel  $K$  given by:

$$K(r, \delta_1, \delta_2) = \mathbf{1}_{[r < R]}(r) \left( \alpha \mathbf{1}_{[(\delta_1 > 0) \text{ and/or } (\delta_2 > 0)]}(\delta_1, \delta_2) - (1 - \alpha) \mathbf{1}_{[(\delta_1 < 0) \text{ and } (\delta_2 < 0)]}(\delta_1, \delta_2) \right).$$

Then considering macroscopic variables  $x' = \varepsilon x, t' = \varepsilon^s t$ , with  $s \in \{1, 2\}$ , the new kinetic distribution,  $f^\varepsilon(x', v, t') = f(x, v, t)$ , satisfies the following kinetic equation:

$$\varepsilon^s \partial_{t'} f^\varepsilon + \varepsilon \nabla_{x'} \cdot (c v f^\varepsilon) = -\nabla_v \cdot \left( P_{v^\perp} F^\varepsilon[f^\varepsilon] f^\varepsilon \right) + D \Delta_v f^\varepsilon. \quad (59)$$

where the force field is the same as previously expressed in the macroscopic variable. After applying some linear approximation of the interaction kernel, the expansion of the force field in the powers of  $\varepsilon$  leads to the following kinetic equation:

$$\varepsilon^s \partial_t f^\varepsilon + \varepsilon \nabla_x \cdot (c v f^\varepsilon) = \Gamma(f^\varepsilon) + \varepsilon \Phi(f^\varepsilon) + O(\varepsilon^2), \quad (60)$$

where  $\Gamma(f^\varepsilon)$  and  $\Phi(f^\varepsilon)$  are the  $\mathcal{O}(1)$  and  $\mathcal{O}(\varepsilon)$  interaction operators defined by:

$$\Gamma(f^\varepsilon) = -D \nabla_v \cdot \left( P_{v^\perp} \left[ a \mu J_{f^\varepsilon} + \mu (Q_{f^\varepsilon} v) \right] f^\varepsilon \right) + D \Delta_v f^\varepsilon, \quad (61)$$

$$\Phi(f^\varepsilon) = -\mu' \nabla_v \cdot \left( P_{v^\perp} \left[ \nabla_x \cdot Q_{f^\varepsilon} + \nabla_x \frac{\rho_{f^\varepsilon}}{2} \right] f^\varepsilon \right). \quad (62)$$

with  $a = 4\alpha - 1$ ,  $\mu = \nu\pi R^2/4$  and  $\mu' = (\nu R^2)(\alpha - 1/2)/2$  and where the density  $\rho_f(x, t)$ , the mean flux  $J_f(x, t)$  and the Q-tensor  $Q_f(x, t)$  are defined by

$$\begin{aligned} \rho_f(x, t) &= \int_{v \in \mathbb{S}^1} f(x, v, t) dv, \\ J_f(x, t) &= \int_{v \in \mathbb{S}^1} v f(x, v, t) dv, \\ Q_f(x, t) &= \int_{v \in \mathbb{S}^1} \left( v \otimes v - \frac{1}{2} \text{Id} \right) f(x, v, t) dv. \end{aligned}$$

We note that the interaction operator  $\Gamma$  both involves a polar alignment term and a nematic alignment one. As both terms are acting at the same time scale, they constrain the dynamics.

**Equilibria and phase transition.** We first note that the interaction can be written:

$$\Gamma(f) = D \nabla_v \cdot \left( \mathcal{M}_{a\mu J_f, \mu Q_f} \nabla_v \left( \frac{f}{\mathcal{M}_{a\mu J_f, \mu Q_f}} \right) \right).$$

The set of equilibria of the interaction operator  $\Gamma$  is

$$\mathcal{E} = \left\{ \rho \mathcal{M}_{u, \kappa, \lambda}^s \text{ with } \rho \geq 0, \kappa \geq 0, u \in \mathbb{S}^1, \lambda \in \mathbb{R} \text{ and compatibility conditions} \right\},$$

where the distribution  $\mathcal{M}_{u,\kappa,\lambda}^s$  is defined by:

$$\mathcal{M}_{u,\kappa,\lambda}^s(v) = \frac{1}{\mathcal{Z}_{\kappa,\lambda}^s} \exp(\kappa(u \cdot v)) \exp(\lambda(u \cdot v)^2),$$

and the compatibility conditions are:

$$a\mu J_{\rho} \mathcal{M}_{u,\kappa,\lambda}^s = \kappa u, \quad \mu Q_{\rho} \mathcal{M}_{u,\kappa,\lambda}^s = \frac{\lambda}{2}(u \otimes u) - \frac{\lambda}{2}(u^\perp \otimes u^\perp),$$

which are equivalent to the following relations:

$$a\mu \rho c(\kappa, \lambda) = \kappa, \quad \mu \rho \frac{\ell(\kappa, \lambda)}{2} = \lambda, \quad (63)$$

where the function  $c(\kappa, \lambda)$  and  $\ell(\kappa, \lambda)$  are defined by:

$$c(\kappa, \lambda) = \int_0^{2\pi} M_{\kappa,\lambda}^s(\theta) \cos \theta \frac{d\theta}{2\pi}, \quad \ell(\kappa, \lambda) = \int_0^{2\pi} M_{\kappa,\lambda}^s(\theta) \cos(2\theta) \frac{d\theta}{2\pi}. \quad (64)$$

We note that like in the previous work the equilibria are the product of a polar and a nematic distribution in velocity (see Eq. (49)). However, here, both distributions are forced to share the same axis of symmetry in velocity. Moreover, the compatibility conditions show that the two order parameters  $\kappa$  and  $\lambda$  cannot be chosen arbitrarily and their values depend on the density  $\rho$ . This is exactly the same kind of conditions appearing in the Vicsek-type model, without normalization of the local averaged velocity, studied [8, 9]: this can be viewed as a direct extension with additional nematic interactions.

We then can classify the equilibria into three different phases depending on the values of polar and nematic parameters  $\kappa$  and  $\lambda$ :

- a) The **disordered phase** corresponds to equilibria with both  $\kappa = 0$  and  $\lambda = 0$ . This is a uniform distribution in velocity orientation. In that case, the polar and nematic order parameters,  $c(\kappa, \lambda)$  and  $\ell(\kappa, \lambda)$ , both equal 0.
- b) The **nematic phase** corresponds to equilibria with  $\kappa = 0$ . In that case, the polar order parameter  $c(\kappa, \lambda)$  equals 0, which also means that there is no flux.
- c) The **polar-nematic phase** corresponds to equilibria with positive  $\kappa > 0$  and  $\lambda \neq 0$ .

A bifurcation analysis shows that there exists two bifurcation points on the curves  $F(\rho, 0, 0) = 0$ , with  $F(\rho, \kappa, \lambda) = (a\mu \rho c(\kappa, \lambda) - \kappa, \mu \rho / 2 \ell(\kappa, \lambda) - \lambda)^T$ . Indeed, we have the following

**Proposition 9.**  $\nabla_{(\kappa,\lambda)} F(\rho, 0, 0)$  is singular if and only if  $\rho = 8/\mu$  or  $\rho = 2/(\mu a)$ . These are two bifurcation points.

On Figure 11, we numerically observed three different kinds of phase transitions: (I) the disorder / nematic phase transition at density  $\rho_{d|n}^* = 8/\mu$  (in purple), (II) the disorder / polar-nematic phase transition at density  $\rho_{d|pn}^* = 2/(a\mu)$  (in green and red) and (III) the nematic / polar-nematic phase transition at density  $\rho_{n|pn}^*(a)$  which is an unknown function of  $a$  (in black and blue). Note that the second phase transition is the same as in [9].

**Macroscopic model.** Following the method developed in [9], we have been able to derive the dynamics of the density  $\rho(x, t) \in \mathbb{R}$  and the mean orientation  $u(x, t) \in \mathbb{S}^1$  in the disordered phase and in the polar-nematic phase.

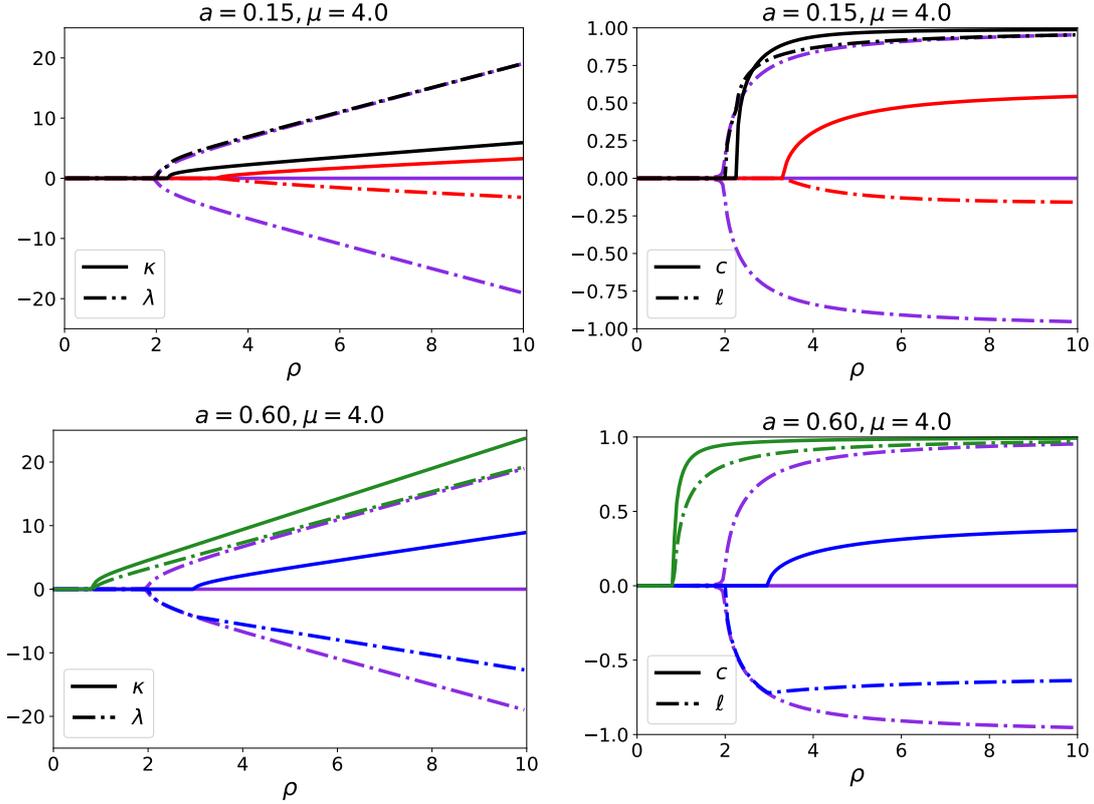


Figure 11: Left plots: The solutions  $(\kappa, \lambda)$  to the compatibility conditions (63) are plotted as functions of the density  $\rho$  for  $a \in \{0.15, 0.6\}$  and  $\mu = 4$ . Right plots: the associated order parameter  $(c(\kappa, \lambda), \ell(\kappa, \lambda))$ . We observe five different phase transitions. Purple curves: disorder / nematic phase transition. Green and red curves: disorder / nematic-polar phase transition, Black and blue curves: nematic / polar-nematic phase transition.

**Proposition 10.** 1. In the disordered phase, the density  $\rho^\varepsilon$  associated with  $f^\varepsilon$  satisfies the following equation:

$$\partial_t \rho^\varepsilon - \varepsilon \nabla_x \cdot \left( \frac{c/(2D) - (\mu'/(4D))\rho^\varepsilon}{(1 - \rho^\varepsilon/\rho_{d|pn})} \nabla_x \rho^\varepsilon \right) = \mathcal{O}(\varepsilon^2).$$

2. In the polar-nematic phase, if  $f^\varepsilon \rightarrow f = \rho \mathcal{M}_{\Omega_p, \kappa, \lambda}^s$ , then  $\rho(x, t)$ ,  $\theta_p(x, t)$  satisfy the following system:

$$\begin{aligned} \partial_t \rho + c \nabla_x \cdot (\rho c_1 \Omega_p) &= 0, \\ \rho \left( \partial_t \theta_p + (c_2^{(1)} + 2\lambda \rho c_5^{(1)}) \Omega_p \cdot \nabla_x \theta_p \right) + \left( -c_3^{(1)} + (-c_4^{(1)} + c_5^{(1)} \left( \frac{\mu}{2} - \partial_\rho \lambda \right)) \rho \right) \Omega_p^\perp \cdot \nabla_x \rho &= 0, \\ \rho \left( \partial_t \theta_p + (c_2^{(2)} + 2\lambda \rho c_5^{(2)}) \Omega_p \cdot \nabla_x \theta_p \right) + \left( -c_3^{(2)} + (-c_4^{(2)} + c_5^{(2)} \left( \frac{\mu}{2} - \partial_\rho \lambda \right)) \rho \right) \Omega_p^\perp \cdot \nabla_x \rho &= 0, \end{aligned}$$

with  $\Omega_p = (\cos \theta_p, \sin \theta_p)^T$  and where functions  $c_i$  are functions of  $(\kappa(\rho), \lambda(\rho))$  and are defined as weighted averages of the equilibria.

Let us make a few comments on this result. The density equation obtained in the nematic phase is similar to the one obtained for the Vicsek model in [15], except that there is an extra

term involving parameter  $\mu'$  due to the interaction term  $\Phi$  of order  $O(1)$  of Equation (60). The stability of the disorder state is, however, modified since this parameter can be positive (when  $\alpha < 1.2$ ): in that case, the disorder state is stable under the condition that  $(\rho^\varepsilon - \rho_{d|pn})$  and  $(2c/\mu' - \rho^\varepsilon)$  have the same sign. Regarding the macroscopic model in the polar-nematic phase, the main observation is that there are two different propagation speeds for the polarity angle  $\theta_p$ , which makes the system very constrained. One non-trivial solution, however, consists of considering a constant velocity  $\Omega_p(x, t) = \text{cst}$  and traveling densities,  $\rho(x, t) = g((x \cdot \Omega_p) - c_1 t)$  with a given function  $g : \mathbb{R} \rightarrow \mathbb{R}^+$ . The dynamics of the system in the nematic phase is still under investigation.

## 4 Effect of self-generated active boundaries and hard-repulsion

In collaboration with Simon Lo Vecchio, Daniel Riveline (IGBMC, Univ. Strasbourg), Olivier Pertz (Univ. Berne) Marcela Szopos (MAP5, Univ. Paris Cité), we propose a particle model with self-generated surrounding actomyosin cables [Cell1]. Indeed, as observed in experiments, actomyosin cables appear at the boundary of multicellular rings: they are intracellular and form a continuous line that encloses cells and acts as an active physical boundary. These cables actually play a decisive role in the emergence of the collective movement.

**Particle model and hard repulsion.** Cells are considered as two-dimensional hard disks and the vectors of positions, velocities, polarities are denoted  $\mathbf{X}(t)$ ,  $\mathbf{V}(t) \in (\mathbb{R}^2)^N$ ,  $\mathbf{P}(t) \in (\mathbb{S}^1)^N$ , with  $N \in \mathbb{N}$ . In an overdamped regime with large friction, the velocities are supposed to be aligned with the polarities direction up to contact forces that prevent cells from overlapping. Using the framework introduced in [19, 20] for contact forces, the dynamics thus write:

$$\frac{d\mathbf{X}}{dt} = \mathbf{V}, \quad (65)$$

$$\mathbf{V} = \mathbf{Proj}_{\mathcal{C}_\mathbf{X}}(c\mathbf{P}), \quad (66)$$

where  $\mathbf{Proj}_{\mathcal{C}_\mathbf{X}}$  denotes the projection operator onto the set of admissible velocities defined by:

$$\mathcal{C}_\mathbf{X} = \left\{ \mathbf{V} \in (\mathbb{R}^2)^N \mid \forall i < j, \quad D_{i,j}(\mathbf{X}) = 0 \quad \Rightarrow \quad \nabla D_{i,j}(\mathbf{X}) \cdot \mathbf{V} \geq 0 \right\},$$

where  $D_{i,j}(\mathbf{X}) = \|\mathbf{X}_i - \mathbf{X}_j\| - 2r$  corresponds to the distance between the  $i$ -th and  $j$ -th particles of radius  $r$ . Regarding the polarities dynamics, we consider a Vicsek-type model to account for the local alignment of the polarities in competition with noise:

$$dP_k = \text{Proj}_{P_k^\perp} \circ \left( \mu \left( \frac{V_k}{\|V_k\|} - P_k \right) dt + \nu \left( \bar{P}_k - P_k \right) dt + \sqrt{2D} (dB_t)_k \right). \quad (67)$$

The only additional ingredient with respect to the standard Vicsek model (see Box 8) is the addition of the relaxation of the polarity onto the velocity direction at a rate  $\mu > 0$ . Note that a variant of this model has already been proposed and studied in [7], where repulsion was, however, treated by using a soft interaction potential.

**Self-generated cables and interactions.** In the following study, the  $N$  cells are initially distributed in a ring domain. As long as the cells keep a ring-like topological configuration, we can define inner and outer cables as piece-wise affine curves linking the cells at the boundary. These two cables are then considered as ensembles of small springs connecting each particle at the outer and inner boundaries and each particle lying at the boundary will be sensitive to the

elastic force exerted by the cables, with bending and stretching components. However, thanks to a scaling analysis, this direct effect is shown to be negligible. Instead, the cables are supposed to indirectly act on the cell by changing the polarity dynamics. The polarity equations then read:

$$\begin{aligned} \frac{d\mathbf{X}}{dt} &= \mathbf{V}, \\ \mathbf{V} &= \mathbf{Proj}_{\mathcal{C}_X} \left( c\mathbf{P} + \gamma (F_{\text{passive}}^{\text{cable}})_k \right), \\ dP_k &= \text{Proj}_{P_k^\perp} \circ \left( \mu \left( \frac{V_k}{\|V_k\|} - P_k \right) dt + \nu (\bar{P}_k - P_k) dt + \sqrt{2D} (dB_t)_k \right. \\ &\quad \left. + \alpha_p (F_{\text{passive}}^{\text{cable}})_k dt + \alpha_a (F_{\text{active}}^{\text{cable}})_k dt \right), \end{aligned}$$

where  $(F_{\text{passive}}^{\text{cable}})_k$  is the passive elastic force and  $(F_{\text{active}}^{\text{cable}})_k$  is an additional active force, which makes the polarity align to the cables. Note that in [25], a similar action of boundary cables have been proposed.

**Numerical methods and simulation set up.** The polarity equation is solved using the semi-implicit method proposed in [Phd3] and the positions and velocities are then updated using the scheme introduced in [19], which writes:

$$\frac{\mathbf{X}^{n+1} - \mathbf{X}^n}{\Delta t} = \mathbf{V}^{n+1}, \quad \mathbf{V}^{n+1} = \mathbf{Proj}_{\mathcal{C}_{\mathbf{X}^n}^{\Delta t}} \left( c\mathbf{P}^{n+1} \right),$$

where the set of admissible velocities has been replaced by a first-order approximation:

$$\mathcal{C}_{\mathbf{X}^n}^{\Delta t} = \left\{ \mathbf{V} \in (\mathbb{R}^2)^N \mid \forall i < j, \quad D_{i,j}(\mathbf{X}^n) + \Delta t \nabla D_{i,j}(\mathbf{X}^n) \cdot \mathbf{V} \geq 0 \right\}.$$

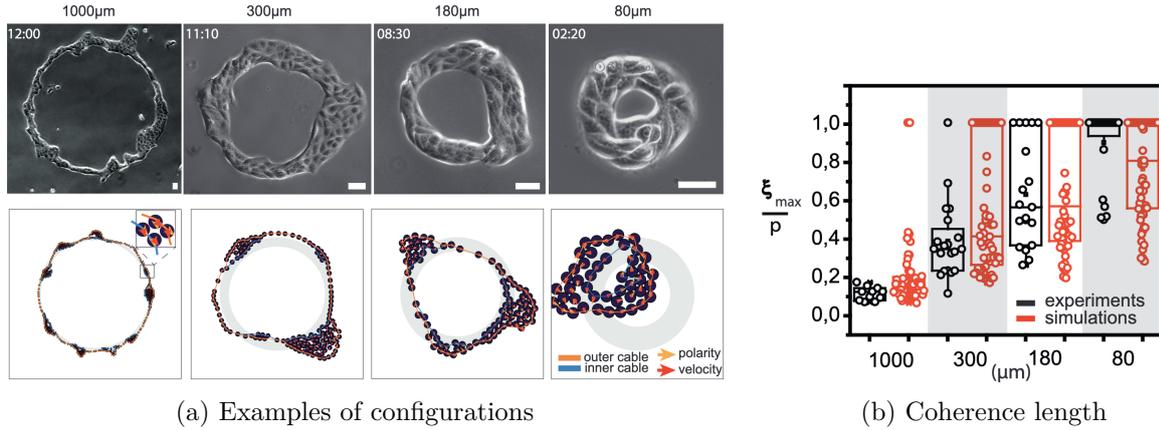
Then this discrete problem is reformulated as an optimization problem with linear constraints, which is solved using the Uzawa algorithm [3]. We refer to [Cell1] for more details.

Regarding the simulation set up, we consider cells initialized on a ring domain,  $\Omega = \{(r, \theta) \in [R_{\min}, R_{\max}] \times [0, 2\pi[ \}$ , with  $R_{\min} = 40, 90, 150$  or  $500 \mu\text{m}$  and  $R_{\max} = R_{\min} + 30 \mu\text{m}$ . As cells have a diameter equal to  $15 \mu\text{m}$ , two cells can be present on each radial section of the ring at initial condition. Accordingly, by taking into account the annular radius, the number of cells is set to 46, 89, 140 or 439. In all simulations, the time step is chosen equal to  $\Delta t = 10^{-3}\text{h}$  and the final time  $T = 24\text{h}$ .

**Parameter calibrations.** The model involves 12 parameters gathered in Table IV.1. Whenever possible, we used physical values obtained from measurements or derived from references. Let us focus on the relaxation parameters  $\nu$ . While maximal speed and angular diffusion can be measured on individual cell dynamics, this is not the case of this parameter, as it parameterizes cell interactions. To extract its value, typical examples of polarity relaxation have been isolated in experiments: cells with a polarity opposite to the direction of rotation, assumed to be  $\bar{P}_k$ , have a shift of angle  $\pi$  after an average time  $\tau$  and relax to  $\bar{P}_k$ .

It is noteworthy that only two parameters were calibrated *a posteriori*:  $\alpha_a$  was set to  $10^{-2}h^{-1}\text{pN}^{-1}$  and  $\alpha_p$  was determined by running numerical simulations, with ring radius  $R_{\min} = 80$ , while aiming at reaching a closure time set by experiments (see Fig.12). We then select the value  $6.5 \times 10^{-5}h^{-1}\text{pN}^{-1}$  in order to retrieve the experimentally reported closing time.

We refer to [Cell1] for a detailed discussion on the other parameters.



(a) Examples of configurations

(b) Coherence length

Figure 13: Taken from [Cell1].

**Comparisons with experiments.** Using the above parameters, we are now able to investigate the dynamics of rings of internal diameter  $80 \mu\text{m}$ ,  $180 \mu\text{m}$ ,  $300 \mu\text{m}$  and  $1000 \mu\text{m}$  respectively. Remarkably, we are able to recover typical multicellular patterns like fingers and clusters 13a. We then compare the coherence length, which is defined by:

$$\frac{\xi}{p} = \min \left\{ -\frac{1}{2\pi} \frac{C_v(0)}{C'_v(0)}, 1 \right\},$$

where  $C_v$  denotes the correlation of the tangential velocity  $v(\theta)$  in polar coordinates relatively to the barycenter of the cell assembly:

$$C_v(\theta) = \int_0^{2\pi} v(t + \theta)v(t)dt.$$

computed in practice using a discrete Fourier transform. This quantity is introduced as a collective motion indicator in biological works [14]. Remarkably, as shown in Figure 13b, we observe a good agreement between the coherence lengths measured in experiments and in numerical simulations for the different sizes of rings, although the parameters have been calibrated on rings of diameter  $80 \mu\text{m}$  only.

Figure 14 gathers results from two numerical test cases to assess the importance of the cables. First, deactivating the active component of the cables induces a clear reduction of the correlation length (Fig. 14a): the active component is thus required for the onset of coherence. Secondly, we perform a comparison with dynamics in a ring-shape domain with a static wall-like confinement as experimentally studied in [17]. To this aim, we change the properties of the actomyosin cables to make them very stiff ( $k_{st}^{\text{stiff}} = 10^5 \times k_{st}$ ,  $k_b^{\text{stiff}} = 10^5 \times k_b$ ). Note that the contribution in the velocity dynamics is not negligible anymore. We also deactivate the dynamic update of the cells involved in the cables to prevent the ring from contracting when the number of cells per cables decreases, and thus to preserve the geometry of the ring over time. In Fig. 14b, we performed a similar numerical experience as in Fig. 13b: in contrast to the case of elastic cables, the coherent motion occurs systematically for confined rings of diameter  $80 \mu\text{m}$ . Note, however, that these simulations are not strictly equivalent to a confined environment dynamics as collective translation of the ring is possible.

Parameters extracted from measures		
Cells diameter	$2r$	$15 \mu\text{m}$
Cells maximal velocity amplitude	$c$	$21.6 \mu\text{m h}^{-1}$ (mean)
Angular diffusion	$D$	$0.96 \text{ rad}^2 \text{ h}^{-1}$
Relaxation parameter: polarity to velocity	$\mu$	$6.2 \text{ rad h}^{-1}$
Relaxation parameter: polarity to mean polarity	$\nu$	$6.2 \text{ rad h}^{-1}$
Parameters estimated from references		
Cable stretching modulus	$k_{\text{st}}$	$5 \times 10^3 \text{ pN } \mu\text{m}^{-1}$
Cable bending modulus	$k_{\text{b}}$	$10^3 \text{ pN } \mu\text{m}$
Inverse friction coefficient	$\gamma$	$10^{-6} \text{ pN}^{-1} \text{ h}^{-1} \mu\text{m}$
Symmetric (wall-type) force amplitude	$\beta_{\text{sym}}$	$10^4 \text{ pN}$
Asymmetric (nematic-type) force amplitude	$\beta_{\text{asym}}$	$10^4 \text{ pN}$
Parameters obtained by numerical calibration		
Polarization per Newton (passive)	$\alpha_{\text{p}}$	$6.5 \times 10^{-5} \text{ h}^{-1} \text{ pN}^{-1}$
Polarization per Newton (active)	$\alpha_{\text{a}}$	$10^{-2} \text{ h}^{-1} \text{ pN}^{-1}$

Table IV.1: Parameters of the model.

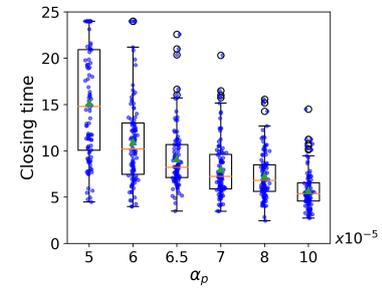


Figure 12: Effect of the variation of the parameter  $\alpha_p$  on the closing time. Each blue point corresponds to an independent numerical simulation and we perform  $n = 100$  simulations per parameter value. The median value is represented as an orange horizontal line, and the mean depicted as a green triangle. Taken from [Cell1].

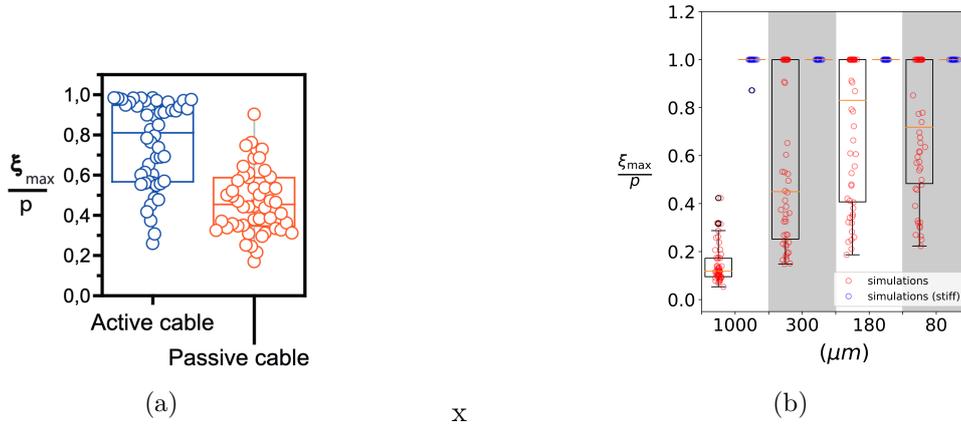


Figure 14: Comparison of the maximal coherence length obtained for simulations with physical cables (in red) and with stiff cables (in blue) for different ring diameters. Each red and blue point corresponds to an independent numerical simulation and we perform  $n = 50$  simulations and  $n_{\text{stiff}} = 50$  simulations per diameter. Taken from [Cell1].

## 5 Conclusion and perspectives

We summarized in this chapter three models for cell collective dynamics, with different modeling ingredients. The first model is a three-dimensional model of disc structured in layers, the second one a two-dimensional model with asymmetric interactions rules and the third two-dimensional hard-sphere model with an active boundary. Of course, each model raises its own mathematical challenges: well-posedness, rigorous derivation, analysis of the phase transitions. However, in this conclusion, we rather focus on new modeling objectives:

1. **Three-dimensional dynamics of discs.** The model presented in Section 2 is based on the assumption that cells are arranged in layers. It would be interesting to remove this modeling assumption, while retaining the same kind of anisotropic character of the cells interactions.
2. **Neural network closure of fluid models.** Macroscopic models have been derived in strong interactions regimes and the viscous corrections are derived under specific assumptions [10], which may not be satisfied. Extending the validity of such models with appropriate closures could therefore be investigated, following the methodology developed in [Red3].
3. **Free-boundary active fluid models.** A macroscopic version of the model proposed in Section 4 could be developed and analyzed, starting from the model proposed in [23]. A comparison with viscoelastic active fluid models could also be carried out. The inclusion of steep congestion terms in Vicsek models could also be studied.
4. **Including cellular events.** Our aim is to explore in a systematic and generic way how cell apoptosis and mitosis influence the rheological properties of tissue. This can be done by incorporating these cellular events into a Vicsek-type model, as presented in Section 4. Cell contractility could also be considered. The model will then be refined by comparison with both in vitro and in vivo experiments. In order to carry out the comparisons, common indicators of fluidity will be used, such as those proposed in [22]. Programming work will have to be carried out in order to generate simulations involving a large number of particles in a relatively short time so that a parametric study of the model can be

envisaged. This may prove essential, in particular, to identify certain parameters that are not accessible by experiment. A comparison with vertex-type models, commonly used in biology, could be carried out. This is part of the ANR MAPEFLU project (2023-2026), involving Marcela Szopos (Univ. Paris Cité), Daniel Riveline (CNRS, Univ. Strasbourg) and Romain Levayer (Institut Pasteur). Roxana Sublet is currently doing her PhD on this topic, under my supervision and that of Marcela Szopos.

## References

- [1] F. Bolley, J.A. Cañizo, and J.A. Carrillo. “Mean-field limit for the stochastic Vicsek model”. In: *Appl. Math. Lett.* 25.3 (2012), pp. 339–343.
- [2] H. Chaté. “Dry aligning dilute active matter”. In: *Annu. Rev. Condens. Matter Phys.* 11.1 (2020), pp. 189–212.
- [3] P.G. Ciarlet. “Introduction à l’analyse numérique matricielle et à l’optimisation”. In: (2007).
- [4] A. Creppy, F. Plouraboué, O. Praud, X. Druart, S. Cazin, Hui Yu, and P. Degond. “Symmetry-breaking phase transitions in highly concentrated semen”. In: *J. R. Soc. Interface* 13.123 (2016), p. 20160575.
- [5] M. Czajkowski, D.M. Sussman, M.C. Marchetti, and M.L. Manning. “Glassy dynamics in models of confluent tissue with mitosis and apoptosis”. In: *Soft Matter* 15.44 (2019), pp. 9133–9149.
- [6] P. Degond, A. Diez, A. Frouvelle, and S. Merino-Aceituno. “Phase transitions and macroscopic limits in a BGK model of body-attitude coordination”. In: *J. Nonlin. Sci.* 30.6 (2020), pp. 2671–2736.
- [7] P. Degond, G. Dimarco, T. B. N. Mac, and N. Wang. “Macroscopic models of collective motion with repulsion”. In: *Comm. in Math. Sci.* 13.6 (2015), pp. 1615–1638.
- [8] P. Degond, A. Frouvelle, and J.-G. Liu. “Macroscopic limits and phase transition in a system of self-propelled particles”. In: *J. Nonlinear Sci.* 23 (2013), pp. 427–456.
- [9] P. Degond, A. Frouvelle, and J.-G. Liu. “Phase transitions, hysteresis, and hyperbolicity for self-organized alignment dynamics”. In: *Arch. Ration. Mech. Anal.* 216 (2015), pp. 63–115.
- [10] P. Degond, J.-G. Liu, S. Motsch, and V. Panferov. “Hydrodynamic models of self-organized dynamics: Derivation and existence theory”. In: *Methods Appl. Anal.* 20.2 (2013), pp. 89–114.
- [11] P. Degond, A. Manhart, and H. Yu. “A continuum model for nematic alignment of self-propelled particles”. In: *Discrete Contin. Dyn. Syst. - Ser. B* 22 (2017).
- [12] P. Degond and S. Motsch. “Continuum limit of self-driven particles with orientation interaction”. In: *Math. Models Methods Appl. Sci.* 18.supp01 (2008), pp. 1193–1215.
- [13] Pierre Degond, Amic Frouvelle, and Sara Merino-Aceituno. “A new flocking model through body attitude coordination”. In: *Mathematical Models and Methods in Applied Sciences* 27.06 (2017), pp. 1005–1049.
- [14] K. Doxzen, S.R.K. Vedula, M.C. Leong, H. Hirata, N.S. Gov, A.J. Kabla, B. Ladoux, and C.T. Lim. “Guidance of collective cell migration by substrate geometry”. In: *Integr. Biol.* 5.8 (2013), pp. 1026–1035.
- [15] A. Frouvelle. “A continuum model for alignment of self-propelled particles with anisotropy and density-dependent parameters”. In: *Math. Models Methods Appl. Sci.* 22.07 (2012), p. 1250011.
- [16] V. Hakim and P. Silberzan. “Collective cell migration: a physics perspective”. In: *Rep. Progr. Phys.* 80.7 (2017), p. 076601.

- [17] S. Jain, V.M.L. Cachoux, G.H.N.S. Narayana, S. de Beco, J. D'alessandro, V. Cellerin, T. Chen, M.L. Heuzé, P. Marcq, R.-M. Mège, A.J. Kabla, C.T. Lim, and B. Ladoux. “The role of single-cell mechanical behaviour and polarity in driving collective cell migration”. In: *Nat. Phy.* 16.7 (2020), pp. 802–809.
- [18] M.C. Marchetti, J.-F. Joanny, S. Ramaswamy, T.B. Liverpool, J. Prost, M. Rao, and R.A. Simha. “Hydrodynamics of soft active matter”. In: *Rev. Mod. Phys.* 85.3 (2013), pp. 1143–1189.
- [19] B. Maury. “A time-stepping scheme for inelastic collisions”. In: *Numer. Math.* 102 (2006), pp. 649–679.
- [20] B. Maury and J. Venel. “A mathematical framework for a crowd motion model”. In: *C. R. Acad. Sci. Paris, Ser. I* 346 (2008), pp. 1245–1250.
- [21] F. Peruani, A. Deutsch, and M. Bär. “Nonequilibrium clustering of self-propelled rods”. In: *Phys. Rev. E* 74.3 (2006), p. 030904.
- [22] J. Ranft, M. Basan, J. Elgeti, J.-F. Joanny, J. Prost, and F. Jülicher. “Fluidization of tissues by cell division and apoptosis”. In: *Proc. Natl. Acad. Sci.* 107.49 (2010), pp. 20863–20868.
- [23] M. Saadaoui, D. Rocancourt, J. Roussel, F. Corson, and J. Gros. “A tensile ring drives tissue flows to shape the gastrulating amniote embryo”. In: *Science* 367.6476 (2020), pp. 453–458.
- [24] B. Stramer and R. Mayor. “Mechanisms and in vivo functions of contact inhibition of locomotion”. In: *Nat. Rev. Mol. Cell Biol.* 18 (Sept. 2016).
- [25] V. Tarle, A. Ravasio, V. Hakim, and N.S. Gov. “Modeling the finger instability in an expanding cell monolayer”. In: *Integr. Biol.* 7.10 (2015), pp. 1218–1227.
- [26] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet. “Novel type of phase transition in a system of self-driven particles”. In: *Phys. Rev. Lett.* 75.6 (1995), p. 1226.

This manuscript presents different contributions in the mathematical modeling and the development of numerical methods for multiscale transport problems, arising in particular in plasma physics and collective cell dynamics.

In the first chapter, we present a few reduction methods for kinetic transport equations. Performing a finite element semi-discretization in velocity leads to a linear hyperbolic model with good conservation properties. Fluid models with entropic or neural network closures have been studied to obtain even smaller reduced models with only physical quantities as reduced macroscopic variables. Reduced order modeling techniques are also proposed to construct reduced macroscopic variables specifically adapted to a given set of parametric solutions.

The second chapter deals with two numerical issues concerning the discontinuous Galerkin and its semi-Lagrangian version, namely the recurrence numerical artifact and the approximate well-balanced property by basis enrichment.

The third chapter presents two implicit numerical methods for solving hyperbolic systems with low-computational cost thanks to relaxation techniques. The first method relies on the vectorial kinetic relaxation scheme, whose transport steps are solved using an implicit solver with explicit cost. The second one proposes a splitting method applied to the two-speed relaxation system for the Euler system, whose acoustic part involves only constant coefficients. This leads to implicit steps with controlled conditioning.

Finally, the fourth chapter regards the development of models for collective dynamics of cells. We present three different models with specific interaction rules: layer interactions, back-to-back anti-alignment and interactions with surrounding elastic cables. We study their large-scale dynamics either theoretically, by deriving the associated macroscopic models, or using numerical simulations. Comparisons with experimental results have also been performed.

**INSTITUT DE RECHERCHE MATHÉMATIQUE AVANCÉE**  
UMR 7501  
Université de Strasbourg et CNRS  
7 Rue René Descartes  
67 084 STRASBOURG CEDEX

**cnrs**

Université de Strasbourg

**IRMA**  
Institut de Recherche Mathématique Avancée

ISSN 0755-3390

Tél. 03 68 85 01 29  
Fax 03 68 85 03 28  
<https://irma.math.unistra.fr>  
[irma@math.unistra.fr](mailto:irma@math.unistra.fr)

IRMA 2024/014  
<https://tel.archives-ouvertes.fr/>