THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*
Cotutelle internationale avec *Universität Potsdam*

Présentée et soutenue le *11/07/2017* **par :**

Laure PÉDÈCHES

Stochastic models for collective motions of populations
Modèles stochastiques pour des mouvements collectifs de populations

JURY

Gilles BLANCHARD	Universität Potsdam	Examinateur
François BOLLEY	Université Pierre et Marie Curie	Rapporteur
Patrick CATTIAUX	Université Paul Sabatier	Directeur de Thèse
Laure COUTIN	Université Paul Sabatier	Examinatrice
Nicolas FOURNIER	Université Pierre et Marie Curie	Rapporteur
Sylvie ROELLY	Universität Potsdam	Directrice de Thèse

École doctorale et spécialité :
MITT : Domaine Mathématiques : Mathématiques appliquées
Unité de Recherche :
Institut de Mathématiques de Toulouse (UMR 5219)
Directeur(s) de Thèse :
Patrick CATTIAUX et *Sylvie ROELLY*
Rapporteurs :
François BOLLEY et *Nicolas FOURNIER*

À mes parents,

Remerciements

Devant la menace du syndrome de la page blanche – non par défaut de personnes à mentionner ici, commençons par le commencement...

Merci à Patrick Cattiaux et à Sylvie Roelly d'avoir, à leurs risques et périls, encadré cette thèse, entre Toulouse et Potsdam. Merci de m'avoir initiée au mystérieux monde de la recherche mathématique, et pour leur patiente aide dans ma quête de compréhension de quelques-uns de ses microscopiques fragments. Et pour les coups de mains avec la paperasse, française comme allemande. Merci, Sylvie, pour la recherche perpétuelle de petites bêtes, souvent au fin fond d'un *cluster* ou d'une *slide* ; merci, Patrick, pour une certaine cécité devant de tels nuisibles. Merci, surtout, pour votre disponibilité, compréhension, discrétion et empathie.

Merci à François Bolley et à Nicolas Fournier pour avoir rapporté ce manuscrit, et pour les remarques et corrections qui en ont découlé. Merci à Gilles Blanchard pour un petit crochet toulousain entre Potsdam et Barcelone pour venir participer au jury ; merci à Laure Coutin, pour un crochet un peu plus court et une excursion à St-Sernin.

Merci à l'Université Franco-Allemande pour l'attribution d'une allocation pour la mise en oeuvre d'une cotutelle de thèse franco-allemande, à l'Ecole Doctorale MITT, en particulier Agnès Requis et Martine Labruyère pour leur patience et leur disponibilité, aux membres de l'Institut de Mathématiques de Toulouse et de l'*Institut für Mathematik* de l'*Universität Potsdam.*

J'ai passé les (presque) trois dernières années à faire quelque chose que je m'étais jurée que jamais, ô grand jamais, je ne ferais... Préparer un doctorat ! En maths ! Quelle étrange idée...! Trois ans : un certain nombre de feuilles de brouillon, de montées et descentes d'escaliers (et une chute dans les escalators), d'invectives diverses et variées à l'encontre de LateX ou de Matlab, de cascades sur un terrain de basket, de trajets en train et en métro, d'aller-retours entre le 1R1 et l'Upsidum, entre mille autres choses.

Trois ans, aussi, remplis de rencontres et de retrouvailles, de gens qui filent un coup de main ou égaient les journées (n'est ce pas, après tout, la même chose ?). Combien de fois, depuis septembre 2014, me suis-je dit "si un jour, j'écris les remerciements de ma thèse, il faudra que je remercie M. Y ou Mme X" ? J'espère ci-dessous avoir limité les oublis à un minimum...

Et pourquoi ne pas, pour une fois, commencer par ces personnes que le sang vous impose ? Rémi, Maguelone, Antoine, merci pour votre volonté d'aller de l'avant, qui ne m'a guère laissé le choix que d'essayer de faire la même chose. Entre nous, si vous ne vouliez plus me voir, vous auriez pu, d'une, me le dire, de deux, choisir des lieux de villégiature moins éloignés que le Chili, l'Inde, et le terrifiant monde des classes prépas scientifiques ! Merci à ma grand-mère, pour son exemple, à mes oncles et tantes, en particulier Paul et Reine pour leur soutien indéfectible – et très utile – et à Anne et Jörg pour ces quelques week-ends nord-allemands et la traduction du résumé.

J'ai eu la chance de pouvoir un peu découvrir le système allemand, et surtout la très cosmopolite – et très peu allemande – équipe de probabilités de l'*Institut für Mathematik* de l'*Universität Potsdam*. Un merci tout particulier à Sara, demi-soeur de thèse, guide et interprête, à Giovanni, pour, entre autres, un match de beach volley en plein centre de Berlin, ainsi qu'au gang des italiennes de Potsdam. *Tania, Yura, vielen Danke für eure Hilfe ! Sorry, my Ukrainian is not up to scratch...*

En dehors de quelques excursions potsdamiennes, c'est à Toulouse que j'ai passé l'essentiel de mes journées, et que se trouvent donc les gens qui ont dû le plus me supporter... à commencer par mes co-bureaux ! Merci donc à Anna, pour avoir apporté un peu de calme, à Pierre, pour avoir répondu patiemment et sans sourciller à mes questions plus foireuses et triviales les unes que les autres, à Mickaël et Maël pour un peu de diversité, culinaire comme disciplinaire. Un merci tout particulier à Guillaume, à qui j'ai tourné le dos pendant trois ans, pour ne pas m'avoir trop fait culpabiliser, et à Hugo, le bizut de service, dont le poisson rouge aura animé la discussion (sauf les vendredis), pour pas mal de trucs, pas tous de mesure nulle.

Dans les bureaux environnants, la plupart au-dessus, dans un ordre complètement stochastique, merci à Fanny Delebecque, pour ses simulations et sa disponibilité, merci aux "vieux" – les Claire B. et C., Benoît, Malika, Mélanie, Kévin, Sofiane, Nil, Stéphane, mes trois "collègues de groupe de travail" Antoine P., Laurent et Maria –, aux "petits jeunes" – Maylis, Valentin, Pierre, Antoine B., William –, et à ceux avec qui j'ai découvert peu à peu les joies – et les déboires – de la thèse – Ioana, Fanny, Sylvain – et tous ceux que j'oublie honteusement. Un grand grand merci à Anton et Claire D., qui, je l'espère, comprendront pourquoi (*hint* : pas seulement pour les mots fléchés ; ni les horoscopes).

Parce que tout ne commence pas avec la thèse, merci aux cachano-orcéennes, notamment Lucile, Nolwenn, Manon et Lucie.

Dans la "vraie vie", en dehors des maths, il y a des gens aussi. A commencer par des gens un peu virtuels, pseudo-psychopathes rencontrés sur la toile il y maintenant plusieurs années... merci à la bande des jyhèmes, spécialement à Juliette et son ornithorynque de compagnie, à Estelle (one día, wir saura alemán properly sprechen !) pour quelques spots de stop inoubliables, à Thibaut pour deux-trois trucs, à Robin pour deux-trois trucs aussi, et à Léo, bah comme ça.

Parfois, on croise des gens, une ou deux heures par semaine... merci à toute l'équipe du Salin du jeudi soir, en tête Pierre-Henri, Marino et Jean-Mi, et désolée pour mes absences chroniques... la faute de mes genoux, dira-t-on.

Tant qu'à être repassé par Toulouse, autant mentionner Françoise, pour son hospitalité et ses réponses à mes coups de fil... douteux, et Patrice, parce qu'en vrai, je t'aime bien, et que l'autodérision est la chose la plus cool qui soit. Et Jean-François, pour ses coups de pouce.

Pour finir, merci à la "famille du basket gersois" pour son soutien dans des moments qu'il releverait d'un léger euphémisme de qualifier de "peu évidents". Merci aux gens de l'ABC, aux personnes dont le soutien a été essentiel sur l'année écoulée, Sév' en particulier, Pauline et Noémie, également, Noémy, en partie. Merci à Annick, pour m'avoir parfois donné l'impression de servir à quelque chose, à Pascale pour d'innombrables passages au cabinet, et m'avoir laissé jouer dans des situations un peu limites, à Yannick pour des discussions où nous avons refait le basket, à Patrick, grâce à qui je sais compter jusqu'à 24, à l'envers.

C'est dans des situations critiques qu'on découvre la force des liens qu'on entretient avec les gens, et pas forcément dans la fréquence des interactions. Floris, Anaëlle [1], merci pour votre présence... je savais qu'il y avait une raison pour laquelle j'étais particulièrement fan de vous deux !

Il y a des personnes sans qui je ne peux absolument pas dire où j'en serais aujourd'hui (peut-être à l'autre bout du monde ? encore plus à l'ouest ?), et quelques mots dérisoires ne peuvent rendre justice à la gratitude que je ressens pour elles... Marie-Amandine, Marie-Claude, Bastien, Julie... merci, pour tout.

Vingt-cinq “merci” jusque là. Il en manque un.

Papa, maman, merci.

[1] et pas Alice.

Modèles stochastiques pour des mouvements collectifs de populations

Dans cette thèse, nous nous intéressons à des systèmes stochastiques modélisant un des phénomènes biologiques les plus mystérieux, les *mouvements collectifs de populations*. On les observe lors de vols d'oiseaux et au sein de bancs de poissons, mais aussi chez certaines populations de bactéries, troupeaux de bétail ou encore pour des populations humaines. Ce genre de comportements apparaît également dans de nombreux autres domaines tels que la finance, la linguistisque ou encore la robotique.

Nous étudions la dynamique d'un *groupe de N individus*, et plus particulièrement deux types de *comportements asymptotiques*. D'une part, nous nous intéressons aux *propriétés d'ergodicité* en *temps long* : existence d'une probabilité invariante via des fonctionnelles de Lyapunov, vitesse de convergence du semi-groupe de transition vers cette probabilité. Egalement au centre de nos recherches la notion de *flocking* : on la définit comme le fait qu'un ensemble d'individus atteigne un consensus en l'absence d'une structure hiérarchique ; d'un point de vue mathématique, cela correspond à l'alignement des vitesses et au regroupement des individus en essaim. D'autre part, nous étudions le phénomène de *propagation du chaos* quand le nombre de particules N tend vers l'infini : les dynamiques des différents individus deviennent asymptotiquement indépendantes.

Le *modèle de Cucker-Smale*, un modèle déterministe cinétique de champ moyen pour une population sans structure hiérarchique, est notre point de départ. L'interaction entre deux particules varie selon leur "taux de communication", qui dépend de leur distance relative et décroît polynomialement.

Dans le premier chapitre, nous étudions les comportements asymptotiques d'un modèle de Cucker-Smale avec perturbation stochastique et de certaines de ces variantes.

Le chapitre 2 présente plusieurs définitions du flocking dans un cadre aléatoire : diverses dynamiques stochastiques, correspondant à différentes formes de bruit – évoquant par exemple un environnement perturbé, le "libre-arbitre" de chaque individu ou une transmission brouillée – sont reprises et étudiées en conjonction avec ces notions.

Le troisième chapitre est basé sur la *méthode de développement en amas*, outil issu de la mécanique statistique. Nous prouvons l'ergodicité exponentielle de certains processus non-markoviens à dérive non-régulière, et nous appliquons ces résultats à des perturbations du processus d'Ornstein-Uhlenbeck.

Dans la dernière partie, nous nous intéressons à l'*équation parabolique-elliptique en dimension 2 de Keller-Segel*, et en particulier au système de particules en champ moyen que l'on peut en dériver. Nous démontrons l'existence d'une solution, unique dans un certain sens, en déterminant les types de collisions possibles entre les particules, grâce à des comparaisons avec des processsus de Bessel et à la théorie des formes de Dirichlet.

Stochastisches Modell für kollektive Bewegung von Populationen

In dieser Doktorarbeit befassen wir uns mit stochastischen Systemen, die eines der mysteriösesten biologischen Phänomene als Modell darstellen: die *kollektive Bewegung von Gemeinschaften*. Diese werden bei Vögel- und Fischschwärmen, aber auch bei manchen Bakterien, Viehherden oder gar bei Menschen beobachtet. Dieser Verhaltenstyp spielt ebenfalls in anderen Bereichen wie Finanzwesen, Linguistik oder auch Robotik eine Rolle.

Wir nehmen uns der Dynamik einer *Gruppe von N Individuen*, insbesondere *zweier asymptotischen Verhaltenstypen* an. Einerseits befassen wir uns mit den Eigenschaften der *Ergodizität in Langzeit*: Existenz einer invarianten Wahrscheinlichkeitsverteilung durch Ljapunow-Funktionen, und Konvergenzrate der Übergangshalbgruppe gegen diese Wahrscheinlichkeit. Eine ebenfalls zentrale Thematik unserer Forschung ist der Begriff *Flocking*: es wird damit definiert, dass eine Gruppe von Individuen einen dynamischen Konsens ohne hierarchische Struktur erreichen kann; mathematisch gesehen entspricht dies der Aneinanderreihung der Geschwindigkeiten und dem Zusammenkommen des Schwarmes. Andererseits gehen wir das Phänomen der *"Propagation of Chaos"* an, wenn die Anzahl N der Teilchen ins Unendliche tendiert: die Bewegungen der jeweiligen Individuen werden asymptotisch unabhängig.

Unser Ausgangspunkt ist das *Cucker-Smale-Modell*, ein deterministisches kinetisches Molekular-Modell für eine Gruppe ohne hierarchische Struktur. Die Wechselwirkung zwischen zwei Teilchen variiert gemäß deren "Kommunikationsrate", die wiederum von deren relativen Entfernung abhängt und polynomisch abnimmt.
Im ersten Kapitel adressieren wir das asymptotische Verhalten eines Cucker-Smale-Modells mit Rauschstörung und dessen Varianten.
Kapitel 2 stellt mehrere Definitionen des Flockings in einem Zufallsrahmen dar: diverse stochastische Systeme, die verschiedenen Rauschformen entsprechen (die eine gestörte Umgebung, den "freien Willen" des jeweiligen Individuums oder eine unterbrochene Übertragung suggerieren) werden im Zusammenhang mit diesen Begriffen unter die Lupe genommen.
Das dritte Kapitel basiert auf der *"Cluster Expansion"-Methode* aus der statistischen Mechanik. Wir beweisen die exponentielle Ergodizität von gewissen nicht-Markow-Prozessen mit nicht-glattem Drift und wenden diese Ergebnisse auf Störungen des Ornstein-Uhlenbeck-Prozesses an.
Im letzten Teil, nehmen wir uns der *zweidimensionalen parabolisch-elliptischen Gleichung von Keller-Segel* an. Wir beweisen die Existenz einer Lösung, welche in gewisser Hinsicht einzig ist, indem wir, mittels Vergleich mit Bessel-Prozessen und der Dirichlet Formtheorie, mögliche Stoßtypen zwischen den Teilchen ermitteln.

Contents

A general introduction

0.1 About the modelling of collective behaviour

As if a cast of grain leapt back to the hand,
A landscapeful of small black birds, intent
On the far south, convene at some command
At once in the middle of the air, at once are gone
With headlong and unanimous consent
From the pale trees and fields they settled on.

What is an individual thing? They roll
Like a drunken fingerprint across the sky!

Richard Wilbur

A flock of birds crosses the sky, and disappears on the horizon, between the dark, quiet blue of the sea and the fading, once explosive, lights of the sunset. The closing seconds of a nice, random, wildlife documentary about the mating rituals of an underrated flying species. Probably about to be extinct, thanks to global warming, deforestation and plastic bags.

A school of fish, roaming in the cold, inhospitable, unsettling abysses of the ocean, hidden from the scrutiny of all but the least casual of observers. Hundreds, thousands of them, entrenched in the most fluid of choreographies, without any jostling or bumping. How are they doing it? This is a rather good question. Why are they doing it? This is a rather good question. Because of an inherent, visceral belief in the "safety in numbers" theory? Because of the twinge of hunger, and the conviction that if being part of such a large entity does not bring easier food, at least, cannibalism will not be so far-fetched an idea?

0.1.1 Motivations: collective behaviour, a near-ubiquitous phenomenon

Those are but two instances of "collective motion of population", a rather loosely defined expression encompassing the "group behaviour" of a population of members of a single (or closely related) species. The word "flocking" is frequently used as a synonym. Once specifically employed about birds, as in a "flock of birds", it will take on both a wider and more specific meaning in a mathematical context. For now, let it be, from the words of [67], a "collective, coherent motion of large numbers of organisms". Note that the term "swarming" – derived from insect behaviour – can also, at times, describe similar antics: according to Wikipedia it is "a collective behaviour exhibited by entities, particularly animals, of similar size which aggregate together, perhaps milling about the same spot or perhaps moving en masse or migrating in some direction". Later, we shall distinguish them.

"One of the most familiar and ubiquitous biological phenomena" ([67] again), flocking can be found, for a large enough cluster of individuals, among many living species with few in common between each other: from the flights of starlings in the German sky to amoeba and bacteria involved in chemotaxis ([39], [40]) phenomena (say hello to the biologists' darling, *Escherichia Coli*), from marching locusts to pods of dolphins, from insect clouds to wolf packs, from herds of rhinoceros or elephants to loose platypus...

Beyond those examples, we can also mention the ant colonies, and their incredible ability to find – collectively – the best source of food in the surroundings to then implement a pheromone-based foraging trail between their colony and their future meal (see for instance [53]).
And humans. We are not exempt from this sort of behaviour, far from it. Off chance (!), let us consider a crowd exiting a basketball arena at the end of a game: people tends to avoid walls, each other (in most cases, at least) and any other hurdles coming their way, while trying to find the quickest way outside (usually, following the person in front of them as closely as (humanly) possible). Imagine now that the team star player makes an impromptu appearance in the middle of the corridor. Instinctively, most people will converge towards him, or her.

As Lawren puts it in [42], flocking behaviour is "one of nature's oldest and most confounding mysteries", one that has fascinated and bewildered onlookers in equal measures for millennia. Well, at least two millennia: the earliest recorded observations of such phenomena are widely considered to be those of Pliny the Elder, the famous Roman naturalist, in its *Naturalis Historia*, written between the years 77 and 79. In particular, in Book X, *The Natural History of Birds*, he describes different flying patterns: for instance, not long after discussing the mythical phoenix, one can learn in chapter XXXI that stocks are never seen either to depart or to arrive, and he writes in chapter XXXV:

"It is a peculiarity of the starling to fly in troops, as it were, and then to wheel around in a globular mass like a ball, the central troop acting as a pivot for the rest."

This underlines that even amongst a family of species, birds in this case, different kinds of collective motions can be observed: the "V-formation" (or "follow-the-leader" formation), and other line formations, but also cluster formations, for example for large groups of pigeons.

In front of such behaviour, one's mind is assaulted by two main questions: "how?" and "why?". The "how does it work" issue is too wide to consider; we will ignore the foundation of biological motion, and the "tiny motor proteins operating at the molecular scale" evoked by Vicsek in [68] (himself quoting [27] and [3]).

While not irrelevant, as we are considering emergent phenomena, we will not delve into it, as our interest will be in the collective movements observed. How are these patterns possible, what are the decision-making processes involved? Is there a hierarchical structure, or are decisions reached by consensus? Pliny raises the problem of leadership, concerning cranes:

"These birds agree by common consent at what moment they shall set out, fly aloft to look out afar, select a leader for them to follow, and have sentinels duly posted in the rear, which relieve each other by turns, utter loud cries, and with their voice help the whole flight in proper array."

Another additional question follows: in a group of hundreds or thousands individuals in what seems like a coordinated motion, how is the information about the movement direction transmitted to each member? The seemingly near-instantaneous changes of direction that can be observed in flocks of birds can be mind-boggling: how to explain that seagulls, for instance, when suddenly disturbed (by a strolling passer-by or a sharp noise) take flight in a what looks like a perfectly coordinated sweep, without visual interaction between every individual?

It led Selous, in 1931, after decades of studies of birds behaviours, to emit the rather far-fetched hypothesis that birds have some psychic abilities that enable them to communicate by telepathy, through some sort of group mind, in a book aptly titled *Thought-transference (or what?) in birds* [59]. Around this time, Nichols [48] was also mesmerized by this collective precision: "flocks at times fly holding a close ranked formation, and the seemingly instantaneous precision with which they wheel in unison, as though each individual were motivated by a common impulse, rather than adjusting itself to the movements of its companions". While noticing that the flocks seem to proceed without leadership, he made a conjecture.

"A simple explanation of mechanism would be that the faster [birds] finding themselves isolated in the van turn back and in so doing provide a single visual impulse on which the remainder of the flock may swerve almost instantaneously."

A detailed review of the study of bird behaviour, both for line formations and cluster flocks can be found in [5], as well as a survey of the proposed explanations for the different types of flocking formation they can adopt.

Hypotheses are more readily available for the "why?", that is the reasons for flocking behaviours. Chief among them, when one thinks of a migrating species covering (very) long distances, is the lessening of the fluid resistance. Let us quote Pliny, again, who gives an aerodynamic reason for the "V-formation", about geese and swans, in chapter XXXII:

"The flocks, forming a point, move along with great impetus, much, indeed, after the manner of our Liburnian beaked galleys, and it is by doing so that they are enabled to cleave the air more easily than if they presented to it a broad front. The flight gradually enlarges in the rear, much in the form of a wedge, presenting a vast surface to the breeze as it impels them onward; those that follow place their necks on those that go before, while the leading birds, as they become wary, fall to the rear."

Beyond birds, one cannot help but think of middle-distance runners or, especially, cyclists in a peloton, looking to benefit from a sensible drag reduction (and in case of strong sideways winds, trying to trap opponents with the "elastic bang" effect !). One should also mention migration amongst aquatic species, notably eels: the endangered European eel lives in fresh water rivers or lakes of mainland Europe, but spawns in the Sargasso sea, near North America, thousands of kilometres from there! The better the hydrodynamic (and so the higher the locomotion) efficiency, the better their chances to make it out of this trek alive...

Which leads us to a second clue: the "security in numbers" concept, which can take two (and even three) aspects. On the one hand (the dissuasion effect), one can imagine that a predator will have more qualms about attacking a large group of animals, maybe doubting its ability to prevail uninjured, and will rather go for an easier victim, typically an isolated prey. On the other hand (the anonymity effect), if the predator does attack, then the risk for a given individual to be the actual target decreases as the size of the group augments. According to Radakov [54], in a school (different of a shoal, a more loosely defined structure) of fish – defined as "a temporary group of individuals, usually of the same species, all or most of which are in the same phase of the life cycle, actively maintains mutual contact, and manifest, or may manifest at any moment, organized actions which are as a rule biologically useful for all the members of the group" – predators can be avoided if each fish exactly copies the relative movements of its nearest neighbours, even though most of them cannot see the incoming danger. Staying with fish, Milinski and Heller, in [45], study the "confusion effect", that is the difficulty encountered by a predator when aiming for an individual target in "high prey densities".

Looking the other way at the prey-predator relationship, one would expect an increased foraging efficiency in a large population: either – for a carnivorous species – in killing preys or in localizing (we will talk about intelligence gathering in a few sentences) sources of food or water for the group to share (such as for the previously mentioned ant colonies). However, in an environment with limited resources, being part of too large a gathering can be a hindrance: too few to split between too many can lead to rather nefarious consequences...
Millions of Mormon crickets – not real crickets, but members of a flightless grasshopper-like insect species from North America – swarm over long distances in destructive marches, in search for food... and to avoid being eaten by their congeners. Cannibalism, indeed, becomes a frequent answer, especially as the Mormon crickets themselves are a major source of salt and protein, the two nutritional resources they are the most desperate for, according to [61]. As a stationary individual is more likely to be attacked, this lead to a "forced march" (see [8]) and more devastation.

Going back to the positives, being surrounded by members of the same species makes it easier to mate, and probably to raise the offspring as well, so ensuring the survival of the species. As touched upon above when mentioning food gathering, the decision making, and thus the continued existence, of the group is improved by the information and intelligence provided by its members, which share the knowledge at their disposal (presence of preys, predators, water, unidentified specimen or (flying) object). Unfortunately, they can also spread around infections with a greater likelihood (see [71]).

As complex, mysterious and fascinating as it is on its own, collective motion is only a part of a much larger field, collective behaviour, as is explained in Vicsek and Zafeiris comprehensive survey on collective motion [70]. According to [58], in a sociological sense, collective behaviour is

"Potentially a very wide-ranging field of study which deals with the ways in which collective behaviours emerge as responses to problematic circumstances and situations. At one extreme this can mean the study of coordinated and organized social movements; at the other, it refers to the seemingly spontaneous eruption of common behavioural patterns, as for example in episodes of mass hysteria. Between these are responses to natural disasters, riots, lynchings, crazes, fads, fashions, rumours, booms, panics, and even rebellions or revolutions."

For us, in a much wider context, collective behaviour will concern a set of "agents" that interact between each other, at times exhibiting patterns at group-level not always predictable from the individual behaviour.

The diversity of the fields impacted is huge: linguistic in [23], with the emergence of languages (and vowel systems) in primitive societies an example mentioned in Cucker and Smale first paper [22] and abundantly rehearsed since; another classic is the emergence (this word again) of a common belief in a price system (for instance the correlation structure of a creditworthiness index in [37]), thus involving finance, or economy [64]. Nowadays, robotics – see [62] or [43] for systems of multiple mobile robots – is also affected, with systems of mobile autonomous agents, such as vehicles, mobile sensors or satellites. One of the most striking examples was the involvement in the control mechasnisms of the proposed Darwin space mission [52], finally abandoned, involving between four and nine spacecrafts, searching for life on hospitable planets, and requiring for the spacecrafts to adopt, and keep, a group structure. One can also mention artificial life research, animation movies or urban design (based on the behaviour of pedestrian human crowds, as in Section 6.2 of [68], and cars in the streets). Even social networks are concerned, with a study of the network of scientific cooperations [7], as are cell populations (and thus medicine) and granular media. Many other references on the spectrum of collective behaviour and emergence can be found in [49].

The existence or formation of collective behaviour is closely linked to the idea of emergence, one that is not clearly defined. According to the Online Oxford Dictionary, emergence, from the latin *emergere*, "to bring to light" is either "the process of becoming visible after being concealed" or "the process of coming into existence or prominence". On the internet, another definition one can stumble upon is "order arising out of chaos", which can be a rather apt description of some of the phenomena described above, when nothing about the intrinsic properties of each agent suggests such a collective pattern and self-organization. One can also say that it is the idea that the collective is more than the sum of its parts, a faithful motto for most team sport coaches.

0.1.2 Modelling: from wildebeests and ants to self-propelled particles

The big question, now that we have mentioned some of the phenomena that we are interested in, is how to model them, and first, how to model an "agent". In everything that follows, unless stated otherwise, whether representing a wildebeest fleeing a lion, a cancerous cell or a car on a highway, we will considered a point-like, weightless, self-propelled particle, which will interact with other point-like, weightless, self-propelled particles.

Onto the models themselves now: as one would expect, there are loads of them, simple or complex, discrete (step by step) or continuous, deterministic or stochastic, made by computer scientists, biologists, physicists or mathematicians. Most of the earliest, often computer-friendly, models, however, have in common some basic "behavioural rules", general guidelines for each particle to follow: in Reynolds' words: separation, alignment and cohesion (see Figure 1):

- separation: there will be no collisions, since the particles will avoid each other, as they "steer to avoid crowding local flockmates";
- alignment: they move in the same direction, as they "steer towards the average heading of local flockmates";
- cohesion: formation of a group, as the particles "steer to move toward the average position of local flockmates".

These rules can vary, but most models are built around (some variations of) these three principles.

Reynolds' model, published in [56], was, in 1987, one the firsts to appear: it is a purely deterministic model, in which the particles, self-propelled, called "boids" (from "bird-oid objects") follow the three aforementioned rules. An artificial life program, it has been used in various video games and animation films (e.g. the "bat swarms and armies of penguins marching through the streets of Gotham City" in *Batman returns*, in 1992, according to [5], and the same theory was used for the wildebeest stampede ("animators working with computers can figure out what the behavior of the animal is and replicate it" according to Scott Johnson, Computer Generated Imagery supervisor) in *The Lion King*, in 1994).

Three years later, another computer-simulated bird flock model was introduced, by Heppner and Grenander in [38]. Much more biologically-oriented, with 15 parameters that can be altered (such as the maximum distance repulsion between birds or the strength of the attractiveness of

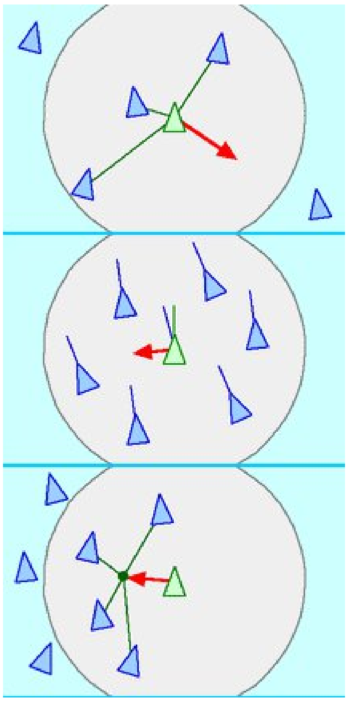

Figure 1: Image taken from Craig Reynolds' website, http://www.red3d.com/cwr/boids/. From top to bottom: separation, alignment and cohesion. The red arrow corresponds to the movement of the green particle.

the neighbours), this is a stochastic model, with randomness given by a Poisson process, assumed to represent "wind gusts and random local disturbances", which can reproduce "flock-like behaviour". This model was inspired by Conway's zero-player Game of Life, a cellular automaton on a square-grid, with each square representing a cell either living or dead, governed by very simple rules, step by step: if a live cell has two or three live neighbours, it lives; if not, it dies, either by isolation or over-population. If a "dead" cell has three neighbours, it becomes alive. Depending on the starting position, the subsequent evolution can be fascinating; it is a classical example of emergence. One example can be seen in Figure 2.

One of the reference models in the field is the so-called Vicsek model [69], stochastic, time-

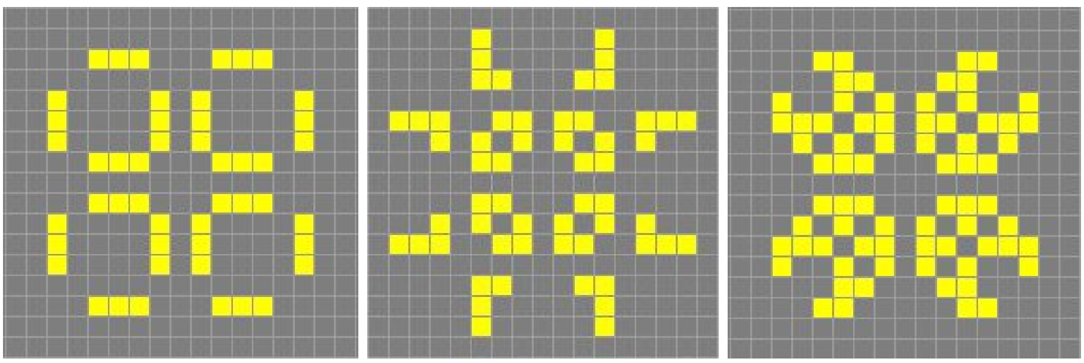

Figure 2: From left to right, a 3-stage loop for Conway's Game of Life. Screenshots from https://bitstorm.org/gameoflife/

discrete, with particles evolving at constant absolute speed such that:

"The only rule of the model is: at each time step a given particle driven with a constant absolute velocity assumes the average direction of motion of the particles in its neighborhood of radius r with some random perturbation added."

Let us consider N particles, with x_i (resp. v_i) the position (resp. velocity) of the i-th particle. They follow the Vicsek model if, for given parameters r, η and v, they satisfy the system, for every $t \in \mathbb{N}$,

$$\begin{cases} x_i(t+1) & = & x_i(t) + v_i(t)\ \Delta t, \quad i \in \{1, ..., N\} \\ \\ \theta_i(t+1) & = & < \theta_j(t) >_r \ + \delta_i(t), \end{cases}$$

where $\Delta t = 1$ is the time step, $< \theta_j(t) >_r$ is the average direction of the velocities of the particles j within range r of particle i, $\delta_i(t)$ is identically distributed between $-\dfrac{\eta}{2}$ and $\dfrac{\eta}{2}$, and $v_i(t+1)$ is such that $|v_i(t+1)| = v$ and $\arg v_i(t+1) = \theta_i(t+1)$.

This model has been the basis for many subsequent studies, with many variations introduced down the years, both discrete and continuous (for instance [41] or [28]), with or without an explicit alignment rule (which is not necessary, in some settings). Its main characteristic is that when the "density of particles" increases, one obtains an ordered collective motion.

Amongst the simplest organisms exhibiting collective behaviour, bacterial colonies have not been forgotten by researchers in the field: notably, a biological model, giving numerical results, was proposed for the hydrodynamics and colony evolution of swimming bacteria, *Bacillus subtilis*, by [24] which appears to coincide with experimental data. A general "phenomenological model", "based on a ferromagneticlike coupling of the velocities of self-propelled particles" and taking in at least five microscopic interactions. In the same vein, we can also mention [25] and [17].

Different types of models have been studied, such as the "nonequilibrium continuum dynamical model" of Toner and Tu [66], for biological organisms such as birds and slime molds, starting from the continuum equations of motion. In this kind of models, with infinitely many agents modelled as a fluid, it is the density, or the concentration, of the organism involved which is studied, like in the Keller-Segel model [39] for chemotaxis, an attractive chemical phenomenon between bacteria.

And then, there is Cucker-Smale model. Introduced in 2007 by Cucker and Smale in [21] and [22], it is a **time-continuous**, **mean-field kinetic deterministic** model describing a N-particle system in $\mathbb{R}^d$:

$$
\begin{cases}
x_i'(t) &= v_i(t), \qquad\qquad i \in \{1, ..., N\} \\
v_i'(t) &= -\dfrac{1}{N} \displaystyle\sum_{j=1}^{N} \psi(x_i(t), x_j(t)) \, (v_i(t) - v_j(t)),
\end{cases}
\tag{0.1.1}
$$

where ψ is a positive symmetric function called *communication rate.*

This model, the starting point from chapters 1 and 2, is a good test subject to discuss the interactions taken into account. Here, there is only one deterministic interaction, an attractive force depending on the relative velocities between the particles, weighted by the "communication rate".

There are at least three items from the previous sentence that are open to debate: first, how to choose this communication rate (and more generally, the shape of the interactions)? Typically, in Cucker and Smale works, it is of the form

$$
\psi(x, y) = \bar{\psi}(x - y) \text{ with } \bar{\psi}(u) = \frac{\lambda}{(1 + |u|^2)^\gamma},
$$

where λ is a positive constant, representing the intensity of this interaction. Thus, it vanishes when the distance between the two agents – let them be two birds – goes to infinity. It sounds pretty sensible (even though one could also imagine a straight cut-off): two birds too far away are not going to be able to "see" – or detect – each other. However, when two birds are very close, the rate goes to λ, thus potentially leading to collisions, or at least very close proximity (remember that the birds are modelled by point-like particles). One could imagine that the interaction would become strongly repulsive between two birds in each other vital space... after all, one does not often see bird collisions in full flight! Of course, as we will see, results will be hard enough to come by with the Cucker-Smale communication rate, let alone with such a 3-regime one as described above... and one should recall that the aim of a model is not always to describe the reality as closely as possible, and can be to give a better understanding of a part of this reality.
This said, this 3-phase model is not without reminding of Aoki [4] computer model, as early as 1982, for schools of fish: two fish interact between each other if they are at a distance smaller than the "radius of extent of near-field interactions", denoted by RC. If the distance between them is larger than a given D_2, the "approach distance", they are attracted to each other; it it is smaller than the "avoidance distance" D_1, then a repulsive force comes into play. The same ideas can be found in [18], with zones of repulsion, orientation and attraction.

Second, and strongly linked to what precedes, one can see that, in the dynamics (0.1.1), the sum is over every particle of the system, with a strength depending on the distance between them. However, Ballerini et al. [6] argue that one should consider the topological distance rather than the metric distance. More precisely, it seems that birds only take into account their six or seven – depending on the species – closest neighbours, whatever their relative distances. In [12], the authors agree, as they found the models based on this topological distance to be more stable than those based on the metric distance. Though, as noticed in [70], other studies, with other kinds of agents do tend in the opposite direction: for instance in [11], when studying marching locusts, the authors observe that "insects adjust their direction to align with neighbors within an interaction range". The problem was also tested in [31]. In the case of the Cucker-Smale model, the authors in [47] try to offset this problem by giving more relative weight to the interactions due to the nearest neighbours, the drawback being a loss of symmetry in the system.

Third, the deterministic nature of the interaction, and, subsequently, of the dynamics. Indeed, why should the trajectories of a group of herrings be entirely predetermined? What about the glorious uncertainty of life? And, as put in [14], the craziness of each individual? Hence the need for a stochastic touch. But then, under what guise? From a practical point of view, a random noise, from a mathematical one, a Brownian motion. If one wishes to represent the wind, or a stream, global enough to concern all individuals, then a common noise, affecting similarly all particles (e.g. [65]). If one wish to represent localized wind gusts, or oceanic currents, or the free will of each agent, its right to (a slice of) self-determination, then a different noise for each individual (e.g. [35]). Stochastic dynamics are what we will be interested in (mind you, maybe the title had given that away).

Numerous deterministic variations of the Cucker-Smale model have been proposed: hierarchical leadership is presented by Shen ([60]), a collision-avoiding model is introduced by Cucker and Dong ([19]), the idea of a "vision cone" for the agents (birds in this case) is studied by Agueh, Illner and Richardson ([1]), amongst many others. As well as a few stochastic ones: in addition to those previously quoted, one should mention [2], where is considered multiplicative noise, and [20], where is shown "nearly-alignment" with a certain probability.

0.1.3 Mathematical aims: asymptotic behaviours

To sum it up, models exist. Loads of them. From very different horizons and of very different kinds. And we will introduce a few others in the following two chapters of this work. That is great. What for, though? What are we looking for? The answer takes two words. Asymptotic behaviour.

On the one hand, time-asymptotic behaviour. What does the system "look like", after a long time? Are the particles clustered, or scattered? And how to quantify the proximity of the individuals? This is where we need a few useful notions.
Time to define what flocking and swarming are, for us, in a more precise way. *Flocking*, phenomena in which a large number of agents reaches a consensus without a hierarchical structure, shall correspond to both an alignment of the velocities and the formation of a group structure:

it happens if, for all $i \in \{1, ..., N\}$,

$$\lim_{t \to \infty} |v_i(t) - v_c(t)|^2 = 0 \quad \text{and} \quad \sup_{0 \leqslant t < \infty} |x_i(t) - x_c(t)|^2 < \infty \tag{0.1.2}$$

with $x_c = \dfrac{1}{N} \displaystyle\sum_{j=1}^{N} x_j$ and $v_c = \dfrac{1}{N} \displaystyle\sum_{j=1}^{N} v_j$ the centres of mass for the positions and the velocities (see figure 3).

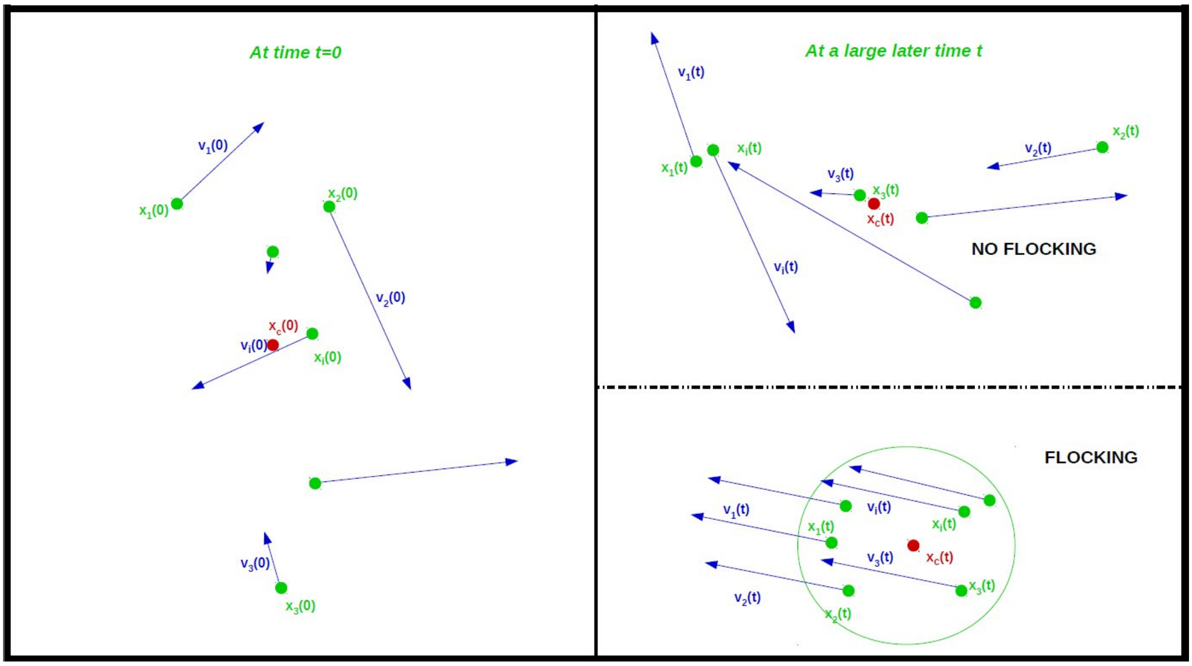

Figure 3: Flocking for a group of particles.

For instance, the main results concerning the Cucker-Smale model can be summed up by the following assertion (see [22] and [36]): when considering the dynamics (0.1.1), for a communication rate

$$\psi(x, y) = \frac{\lambda}{(1 + |x - y|^2)^{\gamma}},$$

there is always flocking if $\gamma \leqslant 1/2$; otherwise, there is "conditional flocking", depending on the initial configuration (positions and velocities) of the system.

A less restrictive notion, *swarming* will only require "cohesion preserving" of the group, that is

$$\sup_{0 \leqslant t < \infty} |v_i(t) - v_c(t)|^2 < \infty \quad \text{and} \quad \sup_{0 \leqslant t < \infty} |x_i(t) - x_c(t)|^2 < \infty.$$

Of course, both definitions are for a deterministic setting. To extend them to a stochastic framework is quite a challenge, to find which definition(s) make(s) the most sense... do we ask for similar results almost surely? Or simply in average, as in [35]? Or in L^p? Or by involving concentration theory? Attempts were made in [14], that is Chapter 3 of this work: we give more details later in this introduction (see paragraph 0.2.2), and for now we only define what probably are the two more relevant notions.

First, we say that there is *almost sure flocking* for the stochastic system (x_i, v_i) if almost surely the property (0.1.2) is satisfied.
Second, L^2-*flocking* happens if

$$\lim_{t\to\infty} \mathbb{E}[|v_i(t) - v_c(t)|^2] = 0 \quad \text{and} \quad \sup_{0\leqslant t<\infty} \mathbb{E}[\,|x_i(t) - x_c(t)|] < \infty.$$

Armed with those definitions, and others, we shall thus determine whether the long-time behaviour can be classified as such, or whether we have an invariant probability measure, and so ergodicity of the system, or whether an explosion or a scattering of the particles...

On the other hand, the behaviour when the number N of particles goes to infinity (as was done for example for an instance of stochastic Vicsek model in [10]). In particular, we shall hope for propagation of chaos results (in the spirit of those first obtained by Sznitman [63]). Take, for any integer k, a group of k particles. If those k particles become independent and identically distributed when N goes to infinity, then there is propagation of chaos. In mathematical terms, one says that a sequence $(Q_N)_{N\geqslant 1}$ of probability measure on E^N is Q-chaotic for a probability measure Q on a Polish space E, if for any fixed integer k and any continuous bounded functions $f_1, ..., f_k$ on E,

$$\lim_{N\to\infty} \int f_1(x_1) \ldots f_k(x_k)\, dQ_N(x_1, ..., x_N) = \prod_{i=1}^{k} \int f_i(x_i)\, dQ(x_i). \tag{0.1.3}$$

In this context, *chaos* simply means some independence of the particles.

One can notice that both those asymptotic aspects, one corresponding to self-organization, the other to propagation of chaos are not exactly two sides of the same coin!

The main initial objective of this thesis was to extend the results obtained in [35] to more general stochastic Cucker-Smale models (see Chapter 2), using both Lyapunov functions and propagation of chaos techniques (see Chapter 1), and also to focus on collision avoidance, with the introduction of hard spheres, and dynamics involving local times, in the spirit of [34] and [15]. This second part was rather quickly forsaken, but should be the aim of future works. Instead the focus shifted to other avenues: one was the 2-D Keller-Segel equation (see Chapter 4) and the mean-field particle system one can associate with it. While the propagation of chaos for this dynamics is still a big open challenge, we were able to prove existence and weak uniqueness. Another lead was to consider the Cucker-Smale model as a system for the velocities only, with infinite delay and to use the cluster expansion method (see Chapter 3), which also turned out to be not fully satisfactory, because of assumptions too restrictive.

0.2 Contents of this work

Each of the four chapters of this thesis corresponds to an article, either published or a preprint. We now give an overview of each of them, briefly presenting their respective framework, the methods and techniques used, and underlining the main results.

0.2.1 About Chapter 1: Asymptotic properties of various stochastic Cucker-Smale dynamics

The starting point of the chapter ([51], submitted in April 2017) is the stochastic Cucker-Smale model introduced by Ha, Lee and Levy in [35], with a random Brownian noise independent for each of the N particles in $\mathbb{R}^d$,

$$\begin{cases} dx_i(t) &=& v_i(t)\ dt \\ dv_i(t) &=& -\dfrac{1}{N}\displaystyle\sum_{j=1}^{N}\psi(x_i(t),x_j(t))\ (v_i(t)-v_j(t))\ dt+\sqrt{D}\ dW_i(t) \end{cases}$$

We give an overview of the asymptotic results for this model, and a variation of it, mainly with a constant communication rate, as well as for a more general rate, in particular settings.

The dynamics is decomposed in two parts: on one side, a "macroscopic" part, $x_c = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} x_i$ and $v_c = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} v_i$, that satisfies

$$\begin{cases} x_c(t) &=& x_c(0)+t\ v_c(0)+\sqrt{D}\displaystyle\int_0^t W_c(s)\ ds \\ v_c(t) &=& v_c(0)+\sqrt{D}\ W_c(t) \end{cases}$$

with $W_c(t) = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} W_i(t)$, and follows a Gaussian dynamics.
On the other side, a "microscopic" part, the relative fluctuations, $\hat{x}_i = x_i - x_c$ and $\hat{v}_i = v_i - v_c$ for all $i \in \{1, ...N\}$.

For a constant communication rate $\psi = \lambda$, the global system is

$$\begin{cases} dx_i(t) &=& v_i(t)\ dt \\ dv_i(t) &=& -\lambda\ (v_i(t)-v_c(t))\ dt+\sqrt{D}\ dW_i(t), \end{cases} \tag{0.2.1}$$

and with $\widehat{W}_i = W_i - W_c$, the microscopic part can be given explicitly

$$\begin{cases} \hat{x}_i(t) &= \displaystyle\int_0^t \hat{v}_i(t)\, dt \\ \\ \hat{v}_i(t) &= e^{-\lambda t}\, \hat{v}_i(0) + \displaystyle\int_0^t e^{-\lambda(t-s)}\, d\widehat{W}_i(s). \end{cases}$$

It is an Ornstein-Uhlenbeck type process, and thus a Gaussian process as well, associated with a covariance matrix of the form $\dfrac{1}{2\lambda}\,(1 - e^{-2\lambda t})\,\Pi_{N,d}$, where $\Pi_{N,d}$ is a certain $Nd \times Nd$ matrix.

While the process $\hat{v}$ of the relative velocities admits the law $\mathcal{N}\left(0, \dfrac{1}{2\lambda}\,\Pi_{N,d}\right)$ as its unique reversible probability measure, the behaviour of the relative positions $\hat{x}$ is less to our taste, as they satisfy the central limit theorem, inspired by [13],

$$\frac{1}{\sqrt{t}}\,\hat{x}(t) \xrightarrow[t\to\infty]{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{\lambda^2}\,\Pi_{N,d}\right).$$

Hence the absence of formation of a group structure as the particles scatter, and of a recognizable collective behaviour. Furthermore, in this scenario, there is only a weak form of flocking (see the next subsection), no almost sure or L^p-flocking of any kind.

To obtain a more aggregating behaviour, we add an attractive interaction in $-\beta\,(x_i - x_j)$; the microscopic part of the stochastic differential system becomes

$$\begin{cases} d\hat{x}_i(t) &= \hat{v}_i(t)\, dt \\ \\ d\hat{v}_i(t) &= -\lambda\hat{v}_i(t)\, dt - \beta\hat{x}_i(t)\, dt + d\widehat{W}_i(t), \end{cases}$$

with the macroscopic part unchanged. With this new interaction, without a true biological meaning (but then, taking a constant communication rate is already something of an aberration from this point of view!), one can find an invariant probability measure for the microscopic system $(\hat{x}, \hat{v})$, restricted to the subspace $\{(x, v) \in \mathbb{R}^{Nd} | x_1 + ... + x_N = 0 \text{ and } v_1 + ... + v_N = 0\}$. It is the measure with density

$$\frac{1}{Z}\,exp\,(-\lambda\,\;(\phi(\hat{x}_1, ..., \hat{x}_{N-1}) + \beta\,\phi(\hat{v}_1, ..., \hat{v}_{N-1})))\,,$$

with respect to the Lebesgue measure, where $\phi(z_1, ..., z_{N-1}) = \displaystyle\sum_{i=1}^{N-1} |z_i|^2 + \left|\sum_{i=1}^{N-1} z_i\right|^2$ and Z is a renormalization constant.
Furthermore, thanks to a Down-Meyn-Tweedie (see [29]) type argument, resting on the existence of a well-chosen Lyapunov function, one can conclude to the exponential ergodicity of the dynamics.

In the last part, for a non-constant communication rate ψ (but still under rather restrictive assumptions: ψ must be bounded above and below by positive numbers, and Lipschitz continuous), we prove that chaos propagates (in the sense of (0.1.3)), following standard proofs in this field, notably by Sznitman [63] and Méléard [44]. This is done in three stages: first, we show the tightness of the sequence of probability measures, then we link the accumulation points of this sequence with a martingale problem, and finally we prove the uniqueness of this martingale problem.

With η_N, defined by $\eta_N(\omega) = \frac{1}{N}\sum_{i=1}^{N} \delta_{(\hat{x}_i^N,\hat{v}_i^N)(\omega)}$, the empirical measure associated with the N-particle system

$$\begin{cases} d\hat{x}_i(t) &= \hat{v}_i(t)dt \\ d\hat{v}_i(t) &= -\frac{1}{N}\sum_{i=1}^{N} \psi(\hat{x}_i(t), \hat{x}_j(t))\,(\hat{v}_i(t) - \hat{v}_j(t))\,dt + \sqrt{D}\,d\widehat{W}_i(t), \end{cases} :$$

one can show that the sequence $(\eta_N)_N$ is η-chaotic, where η, a probability measure on $\mathcal{C}([0,T],\mathbb{R}^{2Nd})$, is the unique solution of the martingale problem associated with the non-linear dynamics

$$\begin{cases} \mathbf{x}_t &= \mathbf{x}_0 + \int_0^t \mathbf{v}_s\,ds \\ \mathbf{v}_t &= \mathbf{v}_0 + W_t - \int_0^t \int \psi(\mathbf{x}_s, x)(\mathbf{v}_s - v)\,\mathbf{Q}_s(dx,dv)\,ds \\ \mathbf{Q}_t &= \mathcal{L}(\mathbf{x}_t, \mathbf{v}_t), \end{cases}$$

which itself admits a unique solution, as we will prove. This propagation of chaos result was previously proven, using a different method, in [9].

0.2.2 About Chapter 2: Stochastic Cucker-Smale models: old and new

This chapter corresponds to [14], and was submitted in April 2017.

Different stochastic Cucker-Smale models, and in particular different kinds of random noises, lead to different asymptotic behaviours, and to different forms of flocking. One of the main goals here is to look at the effect of the random noise on the dynamics: what happens when one disturbs a certain equilibrium? Conversely, is it possible to obtain flocking when the corresponding deterministic model was deprived of it?

To quantify self-organization behaviour, different sorts of stochastic flocking (as well as swarming) are introduced:

- *mean-flocking*(the weakest form of flocking) : flocking property (0.1.2) is satisfied in mean:

for all $i \in \{1, ..., N\}$,

$$\lim_{t\to\infty} |\mathbb{E}[v_i(t)] - \mathbb{E}[v_c(t)]|^2 = 0 \quad \text{and} \quad \sup_{0\leqslant t<\infty} |\mathbb{E}[x_i(t)] - \mathbb{E}[x_c(t)]|^2 < \infty\,;$$

- *weak-flocking*: it more or less corresponds to the convergence in probability towards the center of mass. There is weak flocking with rate $\varepsilon(R)$, with $\lim_{R\to+\infty} \varepsilon(R) = 0$, if for all $R > 0$ and all $i \in \{1, ..., N\}$,

$$\limsup_{t\to+\infty} \mathbb{P}(|v_i(t) - v_c(t)| > R) \leqslant \varepsilon(R)\,;$$

- *almost-sure-flocking*: as previously mentioned, it means that, almost surely (0.1.2) holds:

$$\forall i \in \{1, ..., N\}, \quad \lim_{t\to\infty} |v_i(t) - v_c(t)| = 0 \;\; a.s. \quad \text{and} \quad \sup_{0\leqslant t<\infty} |x_i(t) - x_c(t)| < \infty \;\; a.s.\,;$$

- *$L^{p,q}$-flocking*: a generalized, quite self-explanatory, version of the L^2-flocking defined earlier. There is convergence in L^p of the velocities towards the center of mass, and the difference between the positions and their center of mass is bounded in L^q:

$$\forall i \in \{1, ..., N\}, \quad \lim_{t\to\infty} \mathbb{E}[|v_i(t) - v_c(t)|^p] = 0 \quad \text{and} \quad \sup_{0\leqslant t<\infty} \mathbb{E}[\,|x_i(t) - x_c(t)|^q] < \infty.$$

If $q = 1$, we simply say that there is L^p-flocking.

Then, we take a tour of various stochastic Cucker-Smale models proposed in the literature, to see what kind of flocking behaviour they exhibit, for a constant communication rate.
In [35], stochastic flocking is defined as mean-flocking. We prove in Chapter 1 that there is weak flocking with a rate of convergence given by a χ^2-tail for the dynamics (0.2.1):

$$dv_i(t) = -\lambda\,(v_i(t) - v_c(t))\,dt \;+ \sqrt{D}\,dw_i(t) \tag{0.2.2}$$

For the system inspired by [2], for a common noise w and some constant velocity v_e,

$$dv_i(t) = -\lambda\,(v_i(t) - v_c(t))\,dt \;+\; D\,(v_i(t) - v_e)\,dw(t)\,, \tag{0.2.3}$$

we show that there are always almost-sure-flocking and L^1-flocking, and, if and only if $2\lambda > D^2$, $L^{2,2}$-flocking as well.
The parameters play an even bigger role for a version of the model introduced in [30], with an individual-dependent noise, and a constant σ,

$$dv_i(t) = -\lambda\,(v_i(t) - v_c(t))\,dt \;+\; \sigma\,(v_i(t) - v_c(t))\,dw_i(t)\,, \tag{0.2.4}$$

Indeed, the value $\alpha := (1 - 1/N)\,\sigma^2 - 2\lambda$ determines the behaviour of the system: almost-sure and $L^{2,2}$-flocking if $\alpha < 0$ but no L^2-flocking, and even the norm in L^2 of v_i going to infinity if $\alpha \geqslant 0$!

These three simple equations, only distinguished by their diffusion coefficient, have quite different flocking behaviours; this is not surprising, considering that the last two admits some kind of "dynamical equilibrium" for $v_i(t) = v_c(0)$ and $v_i(t) = v_e$ respectively, and not the first one, whose flocking properties are much weaker.

We also establish results for general communication rates, first in a noisy environment, that is the same random noise $w(t)$ affecting all particles,

$$dv_i(t) = -\frac{\lambda}{N} \sum_{j=1}^{N} \psi_{ij}(t)(v_i(t) - v_j(t))\, dt \,+\, \sigma(v_i(t))\, dw(t)\,, \tag{0.2.5}$$

with $\psi_{i,j} = \psi_{i,j}(v(.), x(.))$ locally Lipschitz, as a function of t, non-negative and symmetric, and σ globally K-Lipschitz continuous.

Theorem 1. *(i) If $2\lambda \inf_{i,j,x,v} \psi_{i,j}(v,x) > 4K^2d^2$, there is almost-sure and $L^{2,2}$-flocking.*

(ii) Moreover, if $\sigma(v_i)$ is linear in v_i, then, the system always flocks almost surely *(this does not hold in the deterministic framework!).*

(iii) Besides, with this same dynamics, assuming that $\psi_{i,j} = \bar{\psi}(|x_i - x_j|^2)$, we can affirm that there is conditional flocking *(which is not without recalling the deterministic case), that is for a subset of initial conditions, with a* positive probability*.*

We also prove that similar results hold with noisy communication rates,

$$dv_i(t) = -\frac{\lambda}{N} \sum_{j=1}^{N} \psi_{ij}(t)(v_i(t) - v_j(t))\, dt \,+\, \frac{1}{N} \sum_{j=1}^{N} \sigma_{ij}(t)\,(v_i(t) - v_j(t))\, dw_{i,j}(t).$$

From this few elements, the most demanding type of flocking appears to be the L^2-flocking, rather than the almost-sure-flocking, which can be found (contrary to L^2-flocking) even in cases (see (ii) of Theorem 1 just above) where the corresponding deterministic system does not flock. Thus, in some particular scenarios, adding noise can foster flocking. Conversely, in other settings, it can create an environment averse to collective behaviour, and leads to a scattering of the particles (see the dynamics (0.2.1), studied in Chapter 1).

0.2.3 About Chapter 3: Exponential ergodicity for a class of non-Markovian stochastic processes

Chapter 3, [50], was submitted in March 2017.

One can see the system (0.2.1) as an autonomous stochastic differential equation in v_i, for every $i \in \{1, ..., N\}$, with infinite delay:

$$dv_i(t) = -\frac{1}{N} \sum_{j=1}^{N} \psi((v_j)_0^t, (v_i)_0^t)\; (v_i(t) - v_j(t))\; dt + \sqrt{D}\; dW_i(t),$$

where $(v)_0^t = \{v(s) \mid 0 \leqslant s \leqslant t\}$ is the trajectory between the times 0 and t.

Even though it was ultimately unsuccessful, as can be seen in Section 4.2 of Chapter 1, with an unconvincing result for finite delay (even though one can very well imagine that the impact of the far way past is negligible, thus justifying a finite delay) and non-explicit communication rates, this was the starting point of this chapter. The tools used here are very different from those of the previous two chapters, and quite new for the study of the ergodicity of stochastic processes.

Indeed, this work is based on techniques coming from statistical mechanics, and in particular Gibbs theory: using the *cluster expansion method*, as for instance in [46] and [26], we prove the exponential ergodicity of a class of non-Markovian stochastic processes, with non-regular drift and finite delay.

Consider stochastic differential dynamics of the form

$$dX_t = \Big(-\frac{1}{2}\,\nabla V(X_t) + \beta\, b((X)_{t-t_0}^t)\Big)\, dt + \, dW_t, \tag{0.2.6}$$

seen as a perturbation of a stochastic differential equation, the "reference process",

$$dX_t = -\frac{1}{2}\,\nabla V(X_t)\, dt + \, dW_t. \tag{0.2.7}$$

Our main result is the theorem:

Theorem 2. *Assume that there exists a reversible probability measure for the dynamics (0.2.7) admitting a Poincaré inequality, and a transition density with a finite moment of order* 8*. If* β *is small enough and the perturbation drift* b *bounded, measurable, then,*

- *the system (0.2.6) admits a unique stationary weak solution (this was already known, see [57]).*
- *moreover, there is exponential ergodicity, and the property of exponential decorrelation holds.*

The proof rests on the convergence of certain sequence of probability measures (Q_N) – law on the finite-time windows $[-N, N]$ of an approximation of (0.2.6) – and the finding of a "cluster representation", and then "cluster estimates", for their renormalization constants Z_N called "partition functions".

Then, we apply this result taking as reference process the Ornstein-Uhlenbeck process, in dimension 1, thus considering perturbed equations of the form

$$dX_t = \Big(-\lambda X_t + \beta\, b((X)_{t-t_0}^t)\Big)\, dt + \, dW_t.$$

As the existence of a unique stationary solution was already known since works from Scheutzow (see [57]), the true novelty here is the explicit rate of convergence, which is exponential.

This result has, however, two main drawbacks: first, because of the nature of the cluster expansion method, there is no way to truly quantify what "small enough" means for the parameter β, thus ruling out explicit drifts and applications; second, verifying that the hypotheses needed for Theorem 2 really hold is, in practice, not an easy thing to do (with the notable exception of the Ornstein-Uhlenbeck case), especially the existence of a moment of order 8 for the density, whose explicit expression is seldom known.

0.2.4 About Chapter 4: The 2-D stochastic Keller-Segel particle model: existence and uniqueness.

Chapter 4, [16], was published in ALEA in 2016.

The Keller-Segel model for chemotaxis was introduced by... Keller and Segel in [39] and [40]. Originally, it represents the interaction of micro-organisms, amoebae, which can lead to aggregation, through the mediation of acrasin, or chemo-attractant, a chemical substance which is then degraded by an enzyme, acrasinase, both produced by the amoebae.

The modelling of the chemical reaction between acrasin and acrasinase, of Fick's law of diffusion and of the variations of the concentration of amoeba due to "an oriented chemotactic motion in the direction of a positive gradient of acrasin and a random motion analogous to diffusion" brings a number of partial differential equations, one each for the evolution of the different concentrations of the substances involved.

A simplified version of this system, known as the 2-D parabolic-elliptic Keller-Segel model, is the following single non-linear partial differential equation in $\mathbb{R}^2$:

$$\partial_t \rho_t(x) = \Delta_x \rho_t(x) + \chi \nabla_x.((K * \rho_t)\rho_t)(x)$$

where $\rho_t : \mathbb{R}^2 \mapsto \mathbb{R}$ is the density at time t of the organisms (e.g. the amoebae), χ is a positive constant – linked to the chemotactic sensitivity, the rate of production of acrasin, the total mass and the product of diffusivities – and $K : x \mapsto \nabla \log(\|x\|) = \dfrac{x}{\|x\|^2}$ is the gradient of the harmonic kernel in dimension 2.

One of the most striking characteristics of this equation is the blow-up phenomenon it exhibits: basically, if $\chi > 4$, there is explosion of the solution in a known finite time; else, there exists a global solution.

Our interest here lies in the mean-field stochastic particle system one would imagine to admit this non-linear equation as the limit law, when N goes to infinity, that is, for $i \in \{1, ..., N\}$:

$$dX_t^i = \sqrt{2}\, dB_t^i - \frac{\chi}{N} \sum_{j \neq i} \frac{X_t^i - X_t^j}{\|X_t^i - X_t^j\|^2}\, dt.$$

Long before focusing on the original ultimate goal, the propagation of chaos – which has, as of yet, not been reached –, one has to deal with the matter of the existence and uniqueness for these dynamics.
Indeed, because of the obvious singularity of the drift $K(.)$ in 0, the system is not defined, as the potential explodes when two particles i and j "collide". The most pressing question thus becomes: do collisions happen? If so, what kind of collisions? And what happens then?

Using comparison theorems for one-dimensional diffusion processes and well-known properties of squared Bessel processes, we show that when $N \geqslant 4$ and $\chi < 4\left(1 - \dfrac{1}{N-1}\right)$, the only possible k-collisions (that is k particles colliding simultaneously) are for $k = 2$, and that, then, the separation is instantaneous (as squared Bessel processes are immediately reflected when reaching 0).

This leads to the following result, with help from Dirichlet forms theory and a uniqueness result from [33] (where very similar results are obtained, using a different method).

Theorem 3. *Set $M = \{x \in \mathbb{R}^{2N}$ such that there exists at most one pair $i \neq j$ such that $x_i = x_j\}$. Then,*

- *for $N \geqslant 4$ and $\chi < 4\left(1 - \dfrac{1}{N-1}\right)$, there exists a unique (in distribution) non explosive solution, starting from any $x \in M$. Moreover, the process is strong Markov, lives in M and admits a symmetric, σ-finite, invariant measure.*
- *if $N \geqslant 2$ and $\chi > 4$, there is no global solution.*
- *if $N \geqslant 2$ and $\chi = 4$, in finite time, either there is explosion or the N particles are "glued".*

Notice that the value 4 remains the pivotal threshold for the parameter χ and that the blow-up phenomenon is also present, as with the partial differential equation. The next step should be to prove that system (0.2.4) "converges" towards the stochastic differential equation

$$\begin{cases} dX_t & = \sqrt{2}\, dB_t - \chi\,(K * \rho_t)(X_t)\, dt \\ \\ \rho_t(x)\, dx & = \mathcal{L}(X_t). \end{cases}$$

However, it has, for now, eluded us.

0.3 Mathematical stuff: notations, framework and stochastic calculus

We recall here a few (mainly) standard notations, and a few results from probability theory for stochastic processes.

0.3.1 Notations

Better be safe than sorry. Or lost in an ocean of unknown letters and symbols.
We give here a few basic and well-spread notations that will be used throughout this work.

Given a random variable X, we will denote by $\mathbb{E}[X]$ its expectation and $\mathrm{var}(X)$ its variance. For a probability measure μ, $\mathbb{P}_\mu$ will be a law on the pathspace with initial distribution μ, and $\mathbb{E}_\mu$ (resp. $\mathrm{var}_\mu(X)$) the associated expectation (resp. variance).

$\mathcal{L}(X)$ will indicate the law of X and, if X follows the distribution μ, we will write $\mathcal{L}(X) \sim \mu$; $\mathcal{N}(m, \Sigma)$ is the Gaussian law in $\mathbb{R}^d$, for a certain integer d, with mean m and covariance matrix Σ. In the same way, $X_n \longrightarrow Z$, with $\mathcal{L}(Z) = \mu$, means that X converges in law (or in distribution) towards the probability measure μ.

$\mathcal{P}(\Omega)$ is the set of probability measures on any given set Ω. Let μ_1 and μ_2 be two measures from $\mathcal{P}(\Omega)$; then $\mu_1 \otimes \mu_2$ is the product probability measure on $\Omega \times \Omega$. The norm in $L^p(\Omega)$, $p \in (0, \infty]$, for any space Ω is denoted by $\|.\|_p$ and the total variation norm by $\|.\|_{TV}$.

0.3.2 Stochastic calculus: tools and basic notions

Consider $(X_t)_{t\in\mathbb{R}_+}$ a $\mathbb{R}^d$-valued process solution of the stochastic differential equation

$$dX_t = b(X_t)\, dt + \Sigma(X_t)\, dB_t \tag{0.3.1}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$, the drift function, and $\Sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$, the diffusion coefficient, are measurable functions and $(B_t)_{t\in\mathbb{R}_+}$ is a standard d-dimensional Brownian motion.

The transition semi-group associated to the Markovian process $(X_t)_{t\geqslant 0}$, denoted by $(P_t)_{t\geqslant 0}$, is the family of operators defined, for f regular enough (a priori, measurable and bounded), by

$$P_t f(x) = \mathbb{E}_x[f(X_t)].$$

It has a number of fundamental properties, two of which we mention here.

Proposition 1. *(Properties of the semi-group)*

- $(P_t)_{t\geqslant 0}$ *satisfies the "semi-group" property (also called the Chapman-Kolmogorov equation), that is for every non-negative* t *and* s, $P_{t+s} = P_t P_s$.
- *A contraction property holds: for every bounded function* f, $\|P_t f\|_\infty \leqslant \|f\|_\infty$.

The infinitesimal generator L is given, for any function f for which it makes sense, by

$$Lf = \lim_{t\to 0} \frac{P_t f - f}{t}.$$

In our case, if $(X_t)_{t\geqslant 0}$ satisfies equation (0.3.1), L can be explicitly stated:

$$L = \frac{1}{2} \sum_{i,j} a_{ij} \partial^2_{ij} + \sum_i b_i \partial i,$$

where $a_{ij} = (\Sigma\Sigma^*)_{ij}$ and b_i is the i-th component of b. The generator L associated with a diffusion process is a second-order, elliptic, operator, defined at least on the set of $\mathcal{C}^2$-functions with a compact support.

A measure μ is said to be invariant for the process, and the process (X_t) stationary, if for every f measurable and bounded,

$$\int P_t f \; d\mu = \int f \; d\mu.$$

Invariant measures can also be characterized with the help of the infinitesimal generator.

Proposition 2. *(Invariance and generator, see [32])*
The measure μ is invariant for the process associated with the generator L if, and only if, for every function f in the domain of definition of L,

$$\int Lf \; d\mu = 0$$

Notice that, if μ is an invariant measure which happens to be a probability, it means that if μ is the initial law of the process, then μ is the law of the process at all time – that is $\mathcal{L}(X_0) = \mu$ implies that for every non-negative t, $\mathcal{L}(X_t) = \mu$. It is also equivalent to the fact that for every integer k, every positive $t_1 < ... < t_k$, and time translation τ, $(X_{t_1}, ..., X_{t_k})$ and $(X_{t_1+\tau}, ..., X_{t_k+\tau})$ have the same law under $\mathbb{P}_\mu$.

Now, let us give similar properties about reversible measure: one says that a measure μ is symmetric, and that the process $(X_t)_t$ is reversible, if for every t, every f and g, measurable and bounded,

$$\int f \; P_t g \; d\mu = \int g \; P_t f \; d\mu.$$

There also exists a characterization of such objects involving the infinitesimal generator:

Proposition 3. *(Reversibility and generator, see [32])*
The measure μ is symmetric for the process $(X_t)_t$ if, and only if, for every f and g smooth enough for the generator to make sense,

$$\int f \; Lg \; d\mu = \int g \; Lf \; d\mu$$

Consider a probability measure μ: it being symmetric means that for every integer k, every positive τ, the processes $(X_{t_1}, ..., X_{t_k})_{t_1,...,t_k}$ and $(X_{\tau-t_1}, ..., X_{\tau-t_k})_{t_1,...,t_k}$ have the same law under $\mathbb{P}_\mu$.

From what is above, one can easily see that if a measure is symmetric for a certain process, then it is also invariant. The converse is not true.

In this report, we will be very (very) interested in finding symmetric or invariant measures, especially symmetric or invariant probability measures. This is not just some weird obsession. There is a (real) (mathematical) reason for it. Here it comes.

Proposition 4. *(Probability measure: uniqueness and ergodicity)*
Under some assumptions of irreducibility and strong recurrence, there is uniqueness of the in-

variant probability measure. Furthermore, for every starting point x, $f \mapsto P_t f(x)$ converges towards μ, for the total variation distance, when t goes to infinity.

Thus, knowing that a probability measure is invariant for a process gives us a feeling about the long-time behaviour of the dynamics considered.

The process is then said to be ergodic, and depending on the speed of this convergence, one will have “geometric ergodicity”, or “exponential ergodicity”, amongst other possibilities. Thus, the existence of an invariant probability measure brings valuable information as to the time-asymptotic behaviour of the considered process.

For readers not overly familiar with stochastic calculus, we finish this introduction by recalling one of the paramount results in the field. Due to Itô, it can be seen as the basis of stochastic differential calculus, or as a stochastic generalization of the chain rule. It will be used throughout this thesis. Here, we present it in one of its multi-dimensional versions, taken from [55].

Theorem 4. *(Itô's formula)*
Let X^i be the i-th component of the process X and $\phi : \mathbb{R}^d \to \mathbb{R}$ a smooth enough function. Then,

$$\begin{aligned}\phi(X_t) &= \phi(X_0) + \sum_{i=1}^{d} \int_0^t \partial_i \phi(X_s)\, dX_s^i + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \partial_{i,j}^2 \phi(X_s)\, d < X^i, X^j >_s \\ &= \phi(X_0) \;+\; \sum_{i=1}^{d} \int_0^t \partial_i \phi(X_s)\, b_i(X_s)\, ds \;+\; \sum_{i,j=1}^{d} \int_0^t \partial_i \phi(X_s)\, \Sigma_{i,j}(X_s)\, dB_s^j \\ &\qquad + \frac{1}{2} \sum_{i,j,k=1}^{d} \Sigma_{i,k}(X_s)\, \Sigma_{j,k}(X_s)\, \partial_{i,j}^2 \phi(X_s)\, ds,\end{aligned}$$

the second equality holding for the process X defined by (0.3.1).

Bibliography

[1] M. Agueh, R. Illner, and A Richardson. Analysis and simulations of a refined flocking and swarming model of Cucker-Smale type. *Kinet. Relat. Models*, 48(1):1–16, 2011.

[2] S.M. Ahn and Y. Ha. Stochastic flocking dynamics of the Cucker-Smale model with multiplicative white noises. *J. Math. Phys.*, 51, 2010.

[3] B. Alberts, D. Bray, M. Raff, K. Roberts, and J. Watson. *Molecular Biology of the Cell.* New York : Garland, 1994.

[4] I. Aoki. A simulation study on the schooling mechanism in fish. *Nippon Suisan Gakkaishi*, 48(8):1081–1088, 1982.

[5] I.L. Bajec and F.H. Heppner. Organized flight in birds. *Animal Behaviour*, 78(4):777 – 789, 2009.

[6] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the National Academy of Sciences*, 105(4):1232–1237, 2008.

[7] A.L. Barabasi, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek. Evolution of the social network of scientific collaborations. *Physica A : Statistical Mechanics and its Applications*, 311(3):590–614, 2002.

[8] S. Bazazi, J. Buhl, J.J. Hale, M.L. Anstey, G.A. Sword, S.J. Simpson, and I.D. Couzin. Collective motion and cannibalism in locust migratory bands. *Current Biology*, 18(10):735 – 739, 2008.

[9] F. Bolley, J.A. Cañizo, and J.A. Carrillo. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Math. Models Methods Appl. Sci.*, 21(11):2179–210, 2011.

[10] F. Bolley, J.A. Cañizo, and J.A. Carrillo. Mean-field limit for the stochastic Vicsek model. *Math. Models Methods Appl. Sci.*, 25(3):339–343, 2012.

[11] J. Buhl, D. J. T. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, and S. J. Simpson. From disorder to order in marching locusts. *Science*, 312(5778):1402–1406, 2006.

[12] M. Camperi, A. Cavagna, I. Giardina, G. Parisi, and E. Silvestri. Spatially balanced topological interaction grants optimal cohesion in flocking models. *Interface Focus*, 2(6):715–725, 2012.

[13] P. Cattiaux, D. Chafaï, and A. Guillin. Central limit theorems for additive functionals of ergodic Markov diffusions processes. *ALEA Lat. Am. J. Probab. Math. Stat.*, 9(2):337–82, 2012.

[14] P. Cattiaux, F. Delebecque, and L. Pédèches. Stochastic Cucker-Smale models : old and new. *Preprint*, 2016.

[15] P. Cattiaux, M. Fradon, A.M. Kulik, and S. Roelly. Long time behavior of stochastic hard ball systems. *Bernoulli*, 22(2):681–710, 05 2016.

[16] P. Cattiaux and L. Pédèches. The 2-d stochastic keller-segel particle model : existence and uniqueness. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 13(1):447–463, 2016.

[17] L.H. Cisneros, R. Cortez, C. Dombrowski, R.E. Goldstein, and J.O. Kessler. Fluid dynamics of self-propelled microorganisms, from individuals to concentrated populations. *Experiments in Fluids*, 43(5):737–753, 2007.

[18] I.D. Couzin, J. Krause, R. James, G.D. Ruxton, and N.R. Franks. Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, 218(1):1 – 11, 2002.

[19] F. Cucker and J-G. Dong. Avoiding collisions in flocks. *IEEE Trans. Automat. Control* ,, 55(5):1238–43, 2010.

[20] F. Cucker and E. Mordecki. Flocking in noisy environments. *J. Math. Pures Appl.*, 89:278–96, 2008.

[21] F. Cucker and S. Smale. Emergent Behavior in Flocks. *IEEE Trans. Automat. Control*, 52(5):852–61, 2007.

[22] F. Cucker and S. Smale. On the mathematics of emergence. *Jpn. J. Math.*, 2:197–227, 2007.

[23] F. Cucker, S. Smale, and D. Zhou. Modeling language evolution. *Foundations of Computational Mathematics*, 4(3):315–343, 2004.

[24] A. Czirók, E. Ben-Jacob, I. Cohen, and T. Vicsek. Formation of complex bacterial colonies via self-generated vortices. *Physical Review E*, 54(2):1791–1801, 1996.

[25] A. Czirók, M. Matsushita, and T. Vicsek. Theory of periodic swarming of bacteria: Application to Proteus mirabilis. *Physical Review E*, 63(3), 2001.

[26] P. Dai Pra and S. Roelly. An existence result for infinite-dimensional Brownian diffusions with non-regular and non-Markovian drift. *Markov Processes Relat. Fields*, 10:113–136, 2004.

[27] J. Darnell, H. Lodish, and D. Baltimore. *Molecular Cell Biology*. New York : Scientific American Books, 1990.

[28] P. Degond and S Motsch. Continuum limit of self-driven particles with orientation interaction. *Mathematical Models and Methods in Applied Sciences*, 18:1193–1215, 2008.

[29] D. Down, S. Meyn, and R. Tweedie. Exponential and uniform ergodicity of Markov processes. *Ann. Probab.*, 23(4):1671–91, 1995.

[30] R. Erban, J. Haskovec, and Y. Sun. A Cucker-Smale model with noise and delay. *SIAM J. Appl. Math.*, 76(4):1535–1557, 2016.

[31] A. Eriksson, M. Nilsson Jacobi, J. Nystrom, and K. Tunstrom. Determining interaction rules in animal swarms. *Behavioral Ecology*, 21(5):1106, 2010.

[32] S.N. Ethier and T.G. Kurtz. *Markov processes : characterization and convergence*. John Wiley and Sons, Inc., New York, 1986.

[33] N. Fournier and B. Jourdain. Stochastic particle approximation of the Keller-Segel equation and two-dimensional generalization of Bessel processes. Preprint, 2015.

[34] M. Fradon and S. Roelly. Infinite dimensional diffusion processes with singular interaction. *Bulletin des Sciences Mathématiques*, 124(4):287 – 318, 2000.

[35] S-Y. Ha, K. Lee, and D. Levy. Emergence of time-asymptotic flocking in a stochastic Cucker-Smale system. *Commun. Math. Sci.*, 7(2):453–69, 2009.

[36] S-Y. Ha and E. Tadmor. From particle to kinetic and hydrodynamic descriptions of flocking. *Kinet. Relat. Models*, 1(3):415–35, 2008.

[37] S.Y. Ha, K.K. Kim, and K. Lee. A mathematical model for multi-name credit based on community flocking. *Quantitative Finance*, 15(5):841–851, 2015.

[38] F. Heppner and U. Grenander. *A stochastic nonlinear model for coodinated bird flocks*, pages 233–238. AAAS publications, 1990.

[39] E.F. Keller and L.A. Segel. Initiation of slime mold aggregation viewed as an instability. *J. Theor. Biol.*, 26:399–415, 1970.

[40] E.F. Keller and L.A. Segel. Model for chemotaxis. *J. Theor. Biol.*, 30:225–234, 1971.

[41] V. L. Kulinskii, V. I. Ratushnaya, A. V. Zvelindovsky, and D. Bedeaux. Hydrodynamic model for a system of self-propelling particles with conservative kinematic constraints. *EPL (Europhysics Letters)*, 71(2):207, 2005.

[42] B. Lawren. Going with the crowd. *National Wildlife*, 1992.

[43] K. Lerman, A. Martinoli, and A. Galstyan. *A Review of Probabilistic Macroscopic Models for Swarm Robotic Systems*, pages 143–152. Springer Berlin Heidelberg, 2005.

[44] S. Méléard. Asymptotic behaviour of some interacting particle systems ; McKean-Vlasov and Boltzmann models. *Lectures Notes in Math.*, 1627:42–95, 1995.

[45] M. Milinski and R. Heller. Influence of a predator on the optimal foraging behaviour of sticklebacks (Gasterosteus aculeatus l.). *Nature*, 275:642 – 644, 1978.

[46] R.A. Minlos, S. Roelly, and H. Zessin. Gibbs states on space-time. *Potential Annal.*, 13(4):367–408, 2000.

[47] S. Motsch and E. Tadmor. A new model for self-organized dynamics and its flocking behavior. *Journal of Statistical Physics*, 144(5):923, 2011.

[48] J.T. Nichols. Notes on the flocking of shore birds. *The Auk*, 48(2):181–185, 1931.

[49] D. Pais. *Emergent Collective Behavior in Multi-Agent Systems: An Evolutionary Perspective*. PhD thesis, 2012.

[50] L. Pédèches. Exponential ergodicity for a class of non-Markovian stochastic processes. *Preprint*, 2016.

[51] L. Pédèches. Asymptotic properties of various stochastic Cucker-Smale dynamics. *Preprint*, 2017.

[52] L. Perea, P. Elosequi, and G. Gómez. Extension of the Cucker-Smale control law to space flight formation. *Journal of Guidance, Control, and Dynamics*, 32(2):527–537, 2009.

[53] B. Prabhakar, K.N. Dektar, and D.M. Gordon. The regulation of ant colony foraging activity without spatial information. *PLOS Computational Biology*, 8(8):1–7, 08 2012.

[54] D.V. Radakov. *Schooling in the ecology of fish*. New York : J. Wiley, 1973.

[55] D. Revuz and M. Yor. *Continuous martingales and Brownian motion.* Springer, 2001.

[56] C.W. Reynolds. Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics*, 21(4):25–34, 1987.

[57] M. Scheutzow. Qualitative behaviour of stochastic delay equations with a bounded memory. *Stochastics*, 12:41–80, 1984.

[58] J. Scott and G. Marshall. *A dictionary of sociology.* Oxford University Press, 2009.

[59] E. Selous. *Thought-transference (or what ?) in birds.* London Constanble and Company Ltd, 1931.

[60] J. Shen. Cucker-Smale flocking under hierarchical leadership. *SIAM J. Appl. Math.*, 68(3):694–719, 2007.

[61] S.J. Simpson, G.A. Sword, P.D. Lorch, and I.D. Couzin. Cannibal crickets on a forced march for protein and salt. *Proceedings of the National Academy of Sciences of the United States of America*, 103(11):4152–4156, 2006.

[62] I. Suzuki and M. Yamashita. Distributed anonymous mobile robots: Formation of geometric patterns. *SIAM Journal on Computing*, 28(4):1347–1363, 1999.

[63] A. Sznitman. Topics in propagation of chaos. *Lectures Notes in Math.*, 1464:167–252, 1989.

[64] L. Testfatsion. Economic agents and markets as emergent phenomena. *Proceedings of the National Academy of Sciences*, 99:7191–7192, 2002.

[65] T.V. Ton, N.T.H. Linh, and A. Yagi. Flocking and non-flocking behavior in a stochastic Cucker-Smale system. *Anal. Appl. (Singap.)*, 12(1):63–73, 2014.

[66] J. Toner and Y. Tu. Long-range order in a two-dimensional dynamical XY model: How birds fly together. *Phys. Rev. Lett.*, 75(23):4326–4329, 1995.

[67] J. Toner, Y. Tu, and S. Ramaswamy. Hydrodynamics and phases of flocks. *Annals of Physics*, 318(1):170 – 244, 2005.

[68] T. Vicsek. *Fluctuations and scaling in biology.* OUP Oxford, 2001.

[69] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel Type of Phase Transition in a System of Self-Driven Particles. *Physical Review Letters*, 75:1226–1229, 1995.

[70] T. Vicsek and A. Zafeiris. Collective motion. *Physics Reports*, 517(3-4):71–140, 2012.

[71] E.O. Wilson. *Sociobiology : the new synthesis.* Cambridge, Mass. :Belknap Press of Harvard University Press, 1975.

Chapter 1

Asymptotic properties of various stochastic Cucker-Smale dynamics

This chapter was submitted, under the same title, in March 2017.

Abstract : Starting from the stochastic Cucker-Smale model introduced by Ha, Lee, and Levy, we look into its asymptotic behaviours. First in term of ergodicity, when t goes to infinity, seeking invariant probability measures and using Lyapunov functionals. Second, when the number N of particles becomes large, leading to results about propagation of chaos.

Keywords : Cucker-Smale dynamics, ergodicity, propagation of chaos.

Contents

1.1 Introduction

Phenomena in which a large number of agents reaches a consensus without a hierarchical structure have been studied abundantly in recent years, as they occur in numerous scientific fields. Indeed, can be considered as such, events as diverse as the emergence of a new language in a primitive society, the belief in a price system in an active market or the collective motions of a population. This last instance encompasses itself very different situations : amongst others, the behaviours of school of fish, flights of birds, bacterial populations or even human groups.

The so-called Cucker-Smale model is one of many attempts at representing such phenomena. It was introduced in 2007 by Cucker and Smale in [11] and [12]. It is a mean-field kinetic deterministic model describing a N-particle system evolving in $\mathbb{R}^d$ through the position x_i and the velocity v_i of particle number i for $i \in \{1, ..., N\}$:

$$\begin{cases} x_i'(t) &= v_i(t) \\ v_i'(t) &= -\dfrac{1}{N}\displaystyle\sum_{j=1}^{N} \psi(x_j(t), x_i(t))\ (v_i(t) - v_j(t)) \end{cases} \tag{1.1.1}$$

where ψ a positive, symmetric function called *communication rate*. Typically, in Cucker and Smale works, it is of the form

$$\psi(x, y) = \bar{\psi}(x - y) \text{ with } \bar{\psi}(u) = \frac{\lambda}{(1 + |u|^2)^\gamma}$$

where λ is a positive constant, representing the intensity of this interaction.

A fundamental property of this model, due to the symmetry of the communication rate, is that the center of mass of the velocities, $v_c = \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} v_j$, is constant at all times : that is, for every t, $v_c(t) = v_c(0)$. Thus, if the initial velocities $v_i(0)$ are all equal, then the velocities are constant, and so equal, at all times : for every i and t, $v_i(t) = v_c(0)$. This is a kind of equilibrium situation, towards which tends the system.

Indeed, the main result of Cucker and Smale is related to flocking, a phenomenon in which self-propelled individuals or particles organize themselves to reach a motion with global coherence, characterized in a mathematical sense by both velocity alignment and formation of a group structure. More precisely, here is the definition for (deterministic) flocking :

Definition 1. Flocking happens for a set of N particles if, for all $i \in \{1, ...N\}$,

- $\displaystyle\lim_{t\to\infty} |v_i(t) - v_c(t)|^2 = 0$, with $v_c(t) = \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} v_j(t)$;
- $\displaystyle\sup_{0\leqslant t<\infty} |x_i(t) - x_c(t)|^2 < \infty$, where $x_c(t) = \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} x_j(t)$.

Cucker and Smale main statement is that there is flocking, whatever the initial conditions, when $\gamma < 1/2$ and that it still occurs under some condition depending only on the initial configuration $(x_i(0), v_i(0))_i$ otherwise ; it was later shown (see [16]) that there is also always flocking if $\gamma = 1/2$.

This result is a major reason why, since then, various authors have studied properties of such models, given alternative proofs of the results (for instance, [15]) and proposed refined, more "reality-compliant" versions of the model : hierarchical leadership is presented by Shen ([22]), a collision-avoiding model is introduced by Cucker and Dong ([9]), the idea of a "vision cone" for the agents (birds in this case) is studied by Agueh, Illner and Richardson ([1]), amongst many others.

There have been a fair number of attempts (including by Cucker and Mordecki in [10], where is added smooth noise, Ahn and Ha in [2] or Ton, Linh and Yagi in [24]) to introduce a random component in this model. Indeed, in the above system, the effects of the environment are neglected : what of the effects of some (very) localized ocean currents or wind gusts, for fishes or birds respectively ? What of the free will of each individual ? And why should the trajectory of a particle be totally predetermined by its initial configuration ?

In this paper, we will focus mainly on the model presented in [14] by Ha, Lee and Levy in 2009, the main difference with the system (1.1.1) being the addition of a stochastic noise, which takes the form of a Brownian motion :

$$\begin{cases} dx_i(t) &= v_i(t)\, dt \\ dv_i(t) &= -\dfrac{1}{N}\displaystyle\sum_{j=1}^{N} \psi(x_j(t), x_i(t))\, (v_i(t) - v_j(t))\, dt + \sqrt{D}\, dW_i(t) \end{cases} \tag{1.1.2}$$

for every $i \in \{1, ..., N\}$, where D is a non-negative number, representing the intensity of the noise, $\psi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+^*$ is positive symmetric function, and $W_1, ..., W_N$ are d-dimensional independent standard Brownian motions.

Here the choice of stochastic noise is such that one can see the W_i, random and independent from each other, as a way of representing the degree of freedom of each individual. In order to model the wind, for example, one should consider the same Brownian motion for all particles : the behaviour of such models is studied in [8]. Notice that, here, because of this choice of noise, and contrary to the deterministic model, the equality of all initial velocities $v_i(0)$ does not imply the equality of all velocities at any given time. That is, having for every i $v_i(0) = v_c(0)$ does not mean that $v_i(t) = v_c(t)$ at time $t > 0$. Thus, the model is no more a perturbation of some equilibrium.

The first question that comes to mind is whether it is possible to obtain results similar to those proven by Cucker and Smale. However, one must first determine what should be the stochastic equivalent of flocking, especially concerning the velocities.

This is a complex problem : in [14], stochastic flocking is defined as deterministic flocking for the expectations of the velocities and positions of the particles. This is a weak condition, one that does not require much in terms of behaviour of the model. Alternatively, one can imagine an almost sure type of flocking, as in [2], if definition 1 holds almost surely, a $\mathcal{L}^p$-flocking (see [24]) if the difference between the velocity of a particle and the center of mass goes to 0 in $\mathcal{L}^p$ or even some kind of weak flocking. For more detailed explanations about the different types of flocking and their appearances in conjunction with different types of stochastic noise, one can read [8].
Here, the independence of the W_i rules out almost sure- and $\mathcal{L}^p$-flockings. We thus focus on the asymptotic behaviour of the system (1.1.2).

In order to study this stochastic dynamics, we will decompose it in two different parts, as is done in [14], corresponding to two different scales :

- On the one hand, we consider a macroscopic (or coarse-scale) system represented by the center of mass x_c of the positions x_i, and its velocity v_c (which, incidentally, is also the center of mass of the velocities v_i) : $x_c = \frac{1}{N}\sum_{i=1}^{N} x_i$ and $v_c = \frac{1}{N}\sum_{i=1}^{N} v_i$.
  From (1.1.2), we deduce the stochastic differential equations satisfied by x_c and v_c :

$$\left\{\begin{array}{lll} dx_c(t) & = & v_c(t)\, dt \\ \\ dv_c(t) & = & \sqrt{D}\, dW_c(t) \end{array}\right.$$

  where $W_c(t) = \frac{1}{N}\sum_{i=1}^{N} W_i(t)$ is a $\mathbb{R}^d$-valued Gaussian process, with expectation 0 and covariance matrix $\frac{D}{N}\, t\, I_d$, for every $t \geqslant 0$.
  This system can therefore be explicitly solved :

$$\left\{\begin{array}{lll} x_c(t) & = & x_c(0) + t\, v_c(0) + \sqrt{D}\int_0^t W_c(s)\, ds \\ \\ v_c(t) & = & v_c(0) + \sqrt{D}\, W_c(t) \end{array}\right. \tag{1.1.3}$$

- On the other hand, we will be interested in a microscopic (or fine-scale) system described by the relative fluctuations of the positions and velocities, around the center of mass and its velocity, $\hat{x}_i = x_i - x_c$ and $\hat{v}_i = v_i - v_c$. Notice that for every positive t,

$$\sum_{i=1}^{N} \hat{x}_i(t) = \sum_{i=1}^{N} \hat{v}_i(t) = 0. \tag{1.1.4}$$

  Assume that ψ is of the form $\psi(x, y) = \bar{\psi}(x - y)$, with $\bar{\psi} : \mathbb{R}^d \to \mathbb{R}_+^*$ a positive, even, function. Therefore, the relative fluctuations satisfy, for all $i \in \{1, ...N\}$,

$$\begin{cases} d\hat{x}_i(t) &= \hat{v}_i(t)dt \\ d\hat{v}_i(t) &= -\dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \psi(\hat{x}_i(t), \hat{x}_j(t)) \, (\hat{v}_i(t) - \hat{v}_j(t)) \, dt + \sqrt{D} \, d\widehat{W}_i(t), \end{cases} \tag{1.1.5}$$

setting $\widehat{W}_i = W_i - W_c$.

Studying these equations will prove to be much more challenging, especially with some nondescript communication rate ψ. We will mainly focus on this relative (or microscopic) system in this work.

We first turn our attention to the particular – nonsensical from a biological point of view but computation-friendly – case of a constant communication rate. We study the time-asymptotic behaviour of system (1.1.5) in this setting and then try to improve it by slightly modifying it in the following section, with the addition of an attractive, linear, input of the positions in the velocity equations, in the same vein as [9]. In this case, as proven in section 3, there exists an invariant probability measure for the couple position-velocity and the system is exponentially ergodic.

Then, in two particular settings, we obtain the existence of an invariant probability measure and ergodicity for variations from the constant communication rate case, in section 4. On the one hand we prove polynomial ergodicity for a class of explicit non-constant communication rates for two particles. On the other hand, using the cluster expansion method presented in [20], we obtain exponential ergodicity for non-explicit drifts with finite delay.

In section 5, we give further results on stationarity, based on Itô-Nisio celebrated result (see [17]) : in particular, we obtain stationary solutions for a larger class of communication rates. This approach requires moment controls as does the final part, section 6, where is investigated system (1.1.2) when the number N of agents goes to infinity to obtain propagation of chaos results, after proving the uniqueness of the associated non-linear stochastic differential system. The results we present in this section can be seen as a particular case of those obtained in [6] ; however our proof, purely based on probability theory, is very different from [6], where transport equations methods are adapted to this framework.

1.2 The basic stochastic Cucker-Smale model with a constant communication rate

From here to the end of Section 4, we place ourselves on $\Omega = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^{2d})$, the canonical continuous $\mathbb{R}^{2d}$-valued path space, with $\mathcal{F}$ the canonical Borel σ-field on Ω.

As it shall not have any impact on any of the results presented in this paper, we shall take in everything, except subsection 1.4.1, that follows $D = 1$, for the sake of simplicity.

Suppose that the communication rate is constant, and more specifically that $\psi = \lambda$, for a certain positive constant λ. This means that whatever the distance between two particles, the interaction between them will be of the same intensity. This does not really seem credible : one would likely imagine that if two individuals are close, their interaction will become strongly repulsive, so as to avoid collisions, and on the contrary, if they are far from each other, they will have difficulties to be influenced mutually, thus resulting in a weak interaction. Unfortunately, such a 3-regime communication rate is beyond what we can do, and we deal first with the easiest possible rate.

1.2.1 Explicit expression and distribution for the relative velocities

From observation (1.1.4), the microscopic system (1.1.5) becomes, in this case, for every positive t, for every $i \in \{1, ...N\}$,

$$\begin{cases} d\hat{x}_i(t) &= \hat{v}_i(t)\ dt \\ d\hat{v}_i(t) &= -\lambda \hat{v}_i(t)\ dt + d\widehat{W}_i(t) \end{cases} \tag{1.2.1}$$

Remark that the second equation is autonomous in $\hat{v}_i$; moreover it is the equation satisfied by what is called an Ornstein-Uhlenbeck type process. Stochastic calculus, and in particular Itô's formula, gives us an explicit expression for $\hat{v}_i$.

Proposition 5. *For every $t \geqslant 0$, and for every $i \in \{1, ..., N\}$,*

$$\hat{v}_i(t) = e^{-\lambda t}\ \hat{v}_i(0) + \int_0^t e^{-\lambda(t-s)}\ d\widehat{W}_i(s).$$

Proof. Apply Itô's formula to $t \mapsto f(t, \hat{v}_i(t)) = e^{\lambda t}\ \hat{v}_i(t)$. Then,

$$e^{\lambda t}\hat{v}_i(t) = \hat{v}_i(0) + \int_0^t e^{\lambda s}\ d\widehat{W}_i(s) + \int_0^t \Big(\lambda\ e^{\lambda s}\ \hat{v}_i(s) - e^{\lambda s}\ \lambda\ \hat{v}_i(s)\Big)\ ds.$$

□

Furthermore, the $\hat{v}_i$ are Gaussian processes : as such, their distribution is entirely determined by their expectation and their variance. Thus, setting $\hat{v}(t)$ the element of $\mathbb{R}^{Nd}$ defined by

$$\hat{v}(t) = (\hat{v}_1(t), ..., \hat{v}_N(t)),$$

and $\Pi_{N,d}$ the following square block matrix of size Nd,

$$\Pi_{N,d} = \begin{pmatrix} (1-\frac{1}{N})I_d & -\frac{1}{N}I_d & \cdots & -\frac{1}{N}I_d \\ -\frac{1}{N}I_d & (1-\frac{1}{N})I_d & \cdots & -\frac{1}{N}I_d \\ \vdots & \ddots & \ddots & \vdots \\ -\frac{1}{N}I_d & -\frac{1}{N}I_d & \cdots & (1-\frac{1}{N})I_d \end{pmatrix}, \tag{1.2.2}$$

the proposition below holds.

Proposition 6. *For every $t \geqslant 0$,*

$$\hat{v}(t) \sim \mathcal{N}\left(e^{-\lambda t}\,\mathbb{E}[\hat{v}(0)], \Lambda_{N,d}(t)\right)$$

where $\Lambda_{N,d}(t) = \dfrac{1}{2\lambda}\,(1 - e^{-2\lambda t})\,\Pi_{N,d}$.

Remark 1. Notice that the eigenvalues of $\Pi_{N,d}$ are 0 with multiplicity d and 1 with multiplicity $(N-1)d$. Thus, the matrix $\Pi_{N,d}$ is not invertible and the law of $\hat{v}$ is degenerate at all time. In particular, it is not absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^{Nd}$, which will prove problematic in later stages.

Remark 2. One can notice from the form of $\Pi_{N,d}$ that many results in dimension d are going to be straightforward generalizations of results in dimension 1.

1.2.2 Invariant probability measure for $\hat{v}$

Let μ be the Gaussian distribution $\mathcal{N}\left(0, \dfrac{1}{2\lambda}\,\Pi_{N,d}\right)$.

Firstly, we prove that μ is a reversible (and thus, invariant) probability measure for the vector $\hat{v}$ of the relative velocities with respect to the center of mass v_c.

Proposition 7. *The process $\hat{v}(.)$ admits $\mu = \mathcal{N}\left(0, \dfrac{1}{2\lambda}\,\Pi_{N,d}\right)$ as its unique reversible probability measure.*

Remark 3. As previously noticed in Remark 1, μ does not admit a probability density with respect to the Lebesgue measure and, so, we cannot use the characterization of reversible measures involving the infinitesimal generator associated with the process in $\mathcal{L}^2(\mu)$.

For every $i \in \{1, ..., N\}$, we write $\hat{x}_i^\alpha(t)$ (resp. $\hat{v}_i^\alpha(t)$) for $\alpha \in \{1, ..., d\}$ the α^{th}-component of the $\mathbb{R}^d$ vector $\hat{x}_i(t)$ (resp. $\hat{v}_i(t)$).

Proof. We prove the invariance of μ. One can obtain its reversibility in a very similar way.

μ is invariant if and only if for every n, $0 < t_1 < ... < t_n$, $\tau > 0$, $(\hat{v}(t_1), ..., \hat{v}(t_n))$ and $(\hat{v}(t_1 + \tau), ..., \hat{v}(t_n + \tau))$ have the same distribution under $\mathbb{P}_\mu$, the law with initial distribution μ. As both are Gaussian processes, it is sufficient to show that they have same expectation and covariance matrix.
For every $i \in \{1, ..., N\}$, $t \geqslant 0$,

$$\mathbb{E}_\mu[\hat{v}_i(t)] = 0$$

and for every $\alpha, \beta \in \{1, ..., d\}$, $i, j \in \{1, ..., N\}$, $t \geqslant 0$,

$$cov_\mu(\hat{v}_i^\alpha(t), \hat{v}_j^\beta(t)) = \delta_{\alpha,\beta} \, \frac{1}{2\lambda} \, (\delta_{i,j} - \frac{1}{N}).$$

Thus, for every non negative t and τ, $v(t)$ and $v(t+\tau)$ have the same distribution under $\mathbb{P}_\mu$.

Furthermore, for $t_1 < t_2$,

$$\begin{aligned}
\mathbb{E}_\mu[\hat{v}_i^\alpha(t_1 + \tau)\hat{v}_j^\beta(t_2 + \tau)] &= e^{-\lambda(t_1+t_2+2\tau)} \, \mathbb{E}_\mu[\hat{v}_i^\alpha(0) \, \hat{v}_j^\beta(0)] \\
&+ \delta_{\alpha,\beta} \, (\delta_{i,j} - \frac{1}{N}) \, \mathbb{E}_\mu \left[\int_0^{t_1+\tau} e^{-\lambda(t_1+\tau-s)} \, dW_i(s) \int_0^{t_2+\tau} e^{-\lambda(t_2+\tau-s)} \, dW_i(s) \right] \\
&= e^{-\lambda(t_1+t_2+2\tau)} \, \delta_{\alpha,\beta} \, \frac{1}{2\lambda}(\delta_{i,j} - \frac{1}{N}) + \delta_{\alpha,\beta} \, (\delta_{i,j} - \frac{1}{N}) \int_0^{t_1+\tau} e^{-\lambda(t_1+t_2+2\tau-s)} \, ds \\
&= \delta_{\alpha,\beta} \, \frac{1}{2\lambda} \, (\delta_{i,j} - \frac{1}{N}) \, \left(e^{-\lambda(t_1+t_2+2\tau)} + e^{-\lambda(t_1+t_2+2\tau)} \, (e^{2\lambda(t_1+\tau)} - 1) \right) \\
&= \delta_{\alpha,\beta} \, \frac{1}{2\lambda} \, (\delta_{i,j} - \frac{1}{N}) \, e^{-\lambda(t_2-t_1)} = \mathbb{E}_\mu[\hat{v}_i^\alpha(t_1) \, \hat{v}_j^\beta(t_2)].
\end{aligned}$$

Hence, $cov_\mu(\hat{v}_i^\alpha(t_1 + T), \hat{v}_j^\beta(t_2 + T)) = cov_\mu(\hat{v}_i^\alpha(t_1), \hat{v}_j^\beta(t_2))$.

$(\hat{v}(t_1), ..., \hat{v}(t_n))$ and $(\hat{v}(t_1 + \tau), ..., \hat{v}(t_n + \tau))$ have same expectation and covariance matrix under $\mathbb{P}_\mu$, and subsequently the same distribution.

In the same fashion, to prove the reversibility of μ, one has to show that for every $n \in \mathbb{N}$, $0 < t_1 < ... < t_n < \tau$, $(\hat{v}(t_1), ..., \hat{v}(t_n))$ and $(\hat{v}(\tau - t_1), ..., \hat{v}(\tau - t_n))$ have the same distribution under $\mathbb{P}_\mu$.

□

Secondly, setting $\hat{x} = (\hat{x}_1, ..., \hat{x}_N)$, we show that the pair $(\hat{x}, \hat{v})$ does not admit an invariant probability measure but an invariant σ-finite measure.

Proposition 8. *The measure $dx \otimes \mu$ is invariant for $(\hat{x}, \hat{v})$, with dx the Lebesgue measure on $\mathbb{R}^{Nd}$.*

Proof. Using Itô's formula, we obtain the infinitesimal generator $\hat{L}$ associated with the system (1.2.1) : for a function f regular enough, for $x, v \in \mathbb{R}^{Nd}$,

$$\hat{L}f(x, v) = \hat{L}_x f(x, v) + \hat{L}_v f(x, v)$$

where

$$\hat{L}_x f(x,v) = \sum_{i=1}^{N} v_i \nabla_{x_i} f$$

and

$$\hat{L}_v f(v) = -\lambda \sum_{i=1}^{N} v_i \nabla_{v_i} f + \frac{1}{2} \sum_{i=1}^{N} \left(\Delta_{v_i} f - \frac{1}{N} \sum_{j=1}^{N} \sum_{\alpha=1}^{d} \partial^2_{v_i^\alpha v_j^\alpha} f \right).$$

As μ is invariant for $\hat{v}$,

$$\int \left(\int \hat{L}_v f(v) \, d\mu(v) \right) dx = \int 0 \, dx = 0.$$

Besides, denoting by $x_i^{\bar{\alpha}}$ the vector x missing its (i,α)-component,

$$\begin{aligned} \int \hat{L}_x f \, dx \otimes \mu &= \sum_{i,\alpha} \int \left(\int v_i^\alpha \, \partial_{x_i^\alpha} f(x,v) \, dx_i^\alpha \right) dx_i^{\bar{\alpha}} \, d\mu(v) \\ &= -\sum_{i,\alpha} \int \left(\int f(x,v) \, \partial_{x_i^\alpha} v_i^\alpha \right) dx \, d\mu(v) \quad = \; 0 \end{aligned}$$

because $\partial_{x_i^\alpha} v_i^\alpha = 0$.
It follows that $\int \hat{L} f \, dx \otimes \mu = 0$.

□

As $dx \otimes \mu$ is a measure with an infinite mass, there is no invariant probability measure for the random system $(\hat{x}(.), \hat{v}(.))$.

1.2.3 Behaviour of $\hat{x}$ and central limit theorem

In this subsection, we will shed light on the asymptotic behaviour of the relative positions $\hat{x}_1(t)$, ..., $\hat{x}_N(t)$, when t goes to infinity. This is indeed of interest to us as we try to assert whether the flocking definitions mentioned in the introduction are really making sense in this context.

Recall that

$$\hat{x}(t) = \hat{x}(0) + \int_0^t \hat{v}(s) \, ds.$$

Using the ergodic theorem,

$$\frac{1}{t} \, \hat{x}(t) = \frac{1}{t} \left(\int_0^t \hat{v}(s) \, ds + \hat{x}(0) \right) \longrightarrow \int \hat{v} \, \mu(d\hat{v}) = 0.$$

Looking for a more precise result, we prove the following convergence result :

Proposition 9. *The central limit theorem below holds :*

$$\frac{1}{\sqrt{t}} \, \hat{x}(t) \xrightarrow[t\to\infty]{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{\lambda^2} \, \Pi_{N,d}\right).$$

Proof. The proof is based on results and methods presented by Cattiaux, Chafaï and Guillin in [7]. We start with a lemma about the variance of the components of $\hat{x} = (\hat{x}_i^\alpha)_{\alpha \in \{1,\dots,d\}, i \in \{1,\dots,N\}}$.

Lemma 1. For every $\alpha \in \{1, ..., d\}$ and $i \in \{1, ..., N\}$,

$$\frac{1}{t}\ var_\mu\left(\int_0^t \hat{v}_i^\alpha(s)\ ds\right) \underset{t\to+\infty}{\longrightarrow} \frac{1}{\lambda^2}\left(1-\frac{1}{N}\right).$$

Proof. There are different ways to prove this lemma (a direct computation works fairly well for instance). We do it by using the method, based on the invariance of the probability measure μ, used in the proof of Lemma 2.3 of [7].

$$\begin{aligned} var_\mu\left(\int_0^t \hat{v}_i^\alpha(s)\ ds\right) = \mathbb{E}_\mu\left[\left(\int_0^t \hat{v}_i^\alpha(s)\ ds\right)^2\right] &= 2\ \mathbb{E}_\mu\left[\int_0^t\int_0^s \hat{v}_i^\alpha(s)\hat{v}_i^\alpha(u)\ du\ ds\right] \\ &= 2\int_0^t\int_0^s \mathbb{E}_\mu[\hat{v}_i^\alpha(s-u)\hat{v}_i^\alpha(0)]\ du\ ds \ \text{ by invariance of } \mu. \end{aligned}$$

Moreover, as the initial conditions and the Brownian motions are independent, it follows that,

$$\mathbb{E}_\mu[\hat{v}_i^\alpha(s-u)\ \hat{v}_i^\alpha(0)] = \mathbb{E}_\mu[e^{-\lambda(s-u)}\ \hat{v}_i^\alpha(0)^2] = e^{-\lambda(s-u)}\frac{1}{2\lambda}\left(1-\frac{1}{N}\right),$$

and

$$\begin{aligned} \frac{1}{t}\ var_\mu\left(\int_0^t \hat{v}_i^\alpha(s)\ ds\right) &= \frac{1}{\lambda t}\left(1-\frac{1}{N}\right)\int_0^t e^{-\lambda s}\left(\int_0^s e^{\lambda u}\ du\right) ds \\ &= \frac{1}{\lambda^2}\left(1-\frac{1}{N}\right)\left(1-\frac{1}{t}\ (1-e^{-\lambda t})\right) \\ &\underset{t\to+\infty}{\longrightarrow} \frac{1}{\lambda^2}\left(1-\frac{1}{N}\right). \end{aligned}$$

□

Therefore, according to Theorem 3.3 of [7], under $\mathbb{P}_\mu$, for all $i \in \{1, ..., N\}$ and $\alpha \in \{1, ..., d\}$,

$$\frac{1}{\sqrt{t}}\int_0^t \hat{v}_i^\alpha(s)ds \xrightarrow[t\to\infty]{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{\lambda^2}(1-\frac{1}{N})\right). \tag{1.2.3}$$

Most of the work towards obtaining the proposition has been done : the only thing left to prove is that for $i \neq j$, α, β,

$$\frac{1}{t}\ cov\left(\int_0^t \hat{v}_i^\alpha(s)\ ds, \int_0^t \hat{v}_j^\beta(s)\ ds\right) \underset{t\to\infty}{\longrightarrow} -\ \delta_{\alpha,\beta}\ \frac{1}{\lambda^2 N}.$$

A direct computation leads to :

$$\begin{aligned} \frac{1}{t}\ cov\left(\int_0^t \hat{v}_i^\alpha(s)\ ds, \int_0^t \hat{v}_j^\beta(s)\ ds\right) &= \frac{1}{t}\ (1-e^{-\lambda t})^2\ \mathbb{E}[\hat{v}_i^\alpha(0)\ \hat{v}_j^\beta(0)] \\ &+ \frac{1}{t}\ \mathbb{E}\left[\int_0^t\int_0^s e^{-\lambda(s-u)}\ d\widehat{W}_i^\alpha(u)\ ds\ \int_0^t\int_0^s e^{-\lambda(s-u)}\ d\widehat{W}_j^\beta(u)\ ds\right] \\ &= -\ \delta_{\alpha,\beta}\ \frac{1}{2\lambda N t}\ (1-e^{-\lambda t})^2 - \delta_{\alpha,\beta}\ \frac{1}{Nt}\ \mathbb{E}\left[\left(\int_0^t\int_0^s e^{-\lambda(s-u)} dW_i^\alpha(u)\ ds\right)\right] \\ &= -\ \delta_{\alpha,\beta}\left(\frac{1}{2\lambda N t}\ (1-e^{-\lambda t})^2 + \frac{1}{Nt}\int_0^t\left(\int_u^t e^{-\lambda(s-u)}\ ds\right)^2 ds\right) \\ &\underset{t\to\infty}{\longrightarrow} -\ \delta_{\alpha,\beta}\ \frac{1}{N}\ \left(\int_u^\infty e^{-\lambda(s-u)}\ ds\right)^2 = -\delta_{\alpha,\beta}\ \frac{1}{\lambda^2 N}. \end{aligned}$$

□

This asymptotic behaviour of $\hat{x}$ confirms the result of the previous subsection : there is no invariant probability measure for the relative system $(\hat{x}, \hat{v})$ associated with the model (1.1.2) introduced in [14] : indeed the particles do not particularly tend to come closer from each other, and thus there is not formation of a group structure.

We will come back to this later, but first we briefly turn our attention to the global process (x, v) and what can be said about its behaviour when t is large.

1.2.4 Back to the original system of absolute motion

In this setting, that is with a constant communication rate, the stochastic differential equations (1.1.2) become :

$$\begin{cases} dx_i(t) &= v_i(t)\, dt \\ dv_i(t) &= -\lambda\, (v_i(t) - v_c(t))\, dt + dW_i(t). \end{cases} \tag{1.2.4}$$

As was shown previously, $x = \hat{x} + x_c$ and $v = \hat{v} + v_c$ have a known explicit expression. v_c is a Brownian motion (and thus admits the Lebesgue measure as an invariant measure) ; therefore v cannot admits an invariant measure with finite mass. Nevertheless, we can find an invariant (and even symmetric) measure for the vector of the global velocity v. Indeed,

Proposition 10. *The measure ν with infinite mass given by*

$$d\nu(v) = exp\left(-\lambda \sum_{\alpha=1}^{d} \sum_{i=1}^{N} (v_i^\alpha - v_c^\alpha)^2\right) dv$$

is reversible for v defined in (1.2.4).

The proof of this proposition follows from a classical result on gradient diffusions :

Lemma 2. If X is solution in $\mathbb{R}^n$ of

$$dX_t = \eta\, dW_t - \nabla F(X_t)\, dt$$

where W_t is a n-dimensional standard Brownian motion, F a smooth function and η a real number, then X admits $\rho(dx) = e^{-\frac{2}{\eta^2}F(x)}\, dx$ as reversible measure.

Proof. The associated infinitesimal generator is defined by

$$Lf = -\sum_{i=1}^{n} \partial_i f\, \partial_i F + \frac{\eta^2}{2} \sum_{i=1}^{n} \partial_i^2 f.$$

To prove the reversibility of ρ, we have to show that for every smooth f and g, $\int gLf\ d\rho = \int fLg\ d\rho$.

$$\begin{aligned}
&\int f(x)Lg(x)\ d\rho(x) - \int g(x)Lf(x)\ d\rho(x) \\
&\quad = -\sum_{i=1}^{n} \int (f\ \partial_i g - g\ \partial_i f)\, \partial_i F\ e^{-\frac{2}{\eta^2}F(x)}\ dx\ +\ \frac{\eta^2}{2}\sum_{i=1}^{n}\int \left(f\ \partial_i^2 g - g\ \partial_i^2 f\right) e^{-\frac{2}{\eta^2}F(x)}\ dx \\
&\quad = -\sum_{i=1}^{n} \int \left(\left(\partial_i g\ \partial_i f + g\ \partial_i^2 f - \partial_i f\ \partial_i g - f\ \partial_i^2 g\right)\frac{\eta^2}{2} + \frac{\eta^2}{2}\left(f\ \partial_i^2 g - g\ \partial_i^2 f\right)\right) d\rho(x) = 0.
\end{aligned}$$

□

Proposition 10 follows by applying this lemma, with $n = Nd$, for

$$F(v) = -\frac{\lambda}{2}\sum_{\alpha=1}^{d}\sum_{i=1}^{N}\left(v_i^\alpha - v_c^\alpha\right)^2.$$

In the next section we will present a slightly modified version of (1.1.2), and, again, study its asymptotic properties.

1.3 Introducing x in the v-equation

As mentioned at the beginning of the previous section, one can imagine that the correlation between the velocities of two particles i and j will be more complex that a linear interaction in $v_i - v_j$. A simple idea can be to add a linear attractive term in $x_i - x_j$, as is done for instance with kinetic models. Indeed it seems reasonable to imagine that the distance between the particles will play a part in the evolution of the direction and the amplitude of the movement of the particle.

For $i \in \{1, ..., N\}$, we now consider

$$\begin{cases} dx_i(t) &= v_i(t)\ dt \\ dv_i(t) &= -\lambda\left(v_i(t) - v_c(t)\right) dt - \beta\left(x_i(t) - x_c(t)\right) dt + dW_i(t) \end{cases} \tag{1.3.1}$$

where β is a positive parameter coding the intensity of this new interaction.

As in the previous section, we divide this system in two parts, a "macroscopic" one and a "microscopic" one. The center of mass (x_c, v_c) is subject to exactly the same dynamics and the expressions of x_c and v_c are still given by (1.1.3).
Changes appear for the relative fluctuations, however, and instead of (1.2.1), one now obtains,

for every i in $\{1, ..., N\}$,

$$\begin{cases} d\hat{x}_i(t) &= \hat{v}_i(t)\, dt \\ \\ d\hat{v}_i(t) &= -\lambda \hat{v}_i(t)\, dt - \beta \hat{x}_i(t)\, dt + d\widehat{W}_i(t). \end{cases} \tag{1.3.2}$$

We now focus on finding an invariant probability measure for stochastic differential processes satisfying equations (1.3.2).

1.3.1 Invariant probability measure on a "d-hyperplane" for the relative fluctuations

We hope that the introduction of this new interaction will bring an invariant probability measure for the microscopic system. Such an occurrence is impossible for the global system (x, v) ; however, one can easily check the validity of the following proposition as well as the fact that the measure involved has an infinite mass.

Proposition 11. *The measure μ_β defined on $(\mathbb{R}^d \times \mathbb{R}^d)^N$ by*

$$d\mu_\beta(x, v) = exp\left(-\lambda \left[\sum_{i=1}^{N} |v_i - v_c|^2 + \beta \sum_{i=1}^{N} |x_i - x_c|^2\right]\right) dx\, dv$$

is invariant for the system of stochastic differential equations (1.3.1).

This result is nevertheless helpful : by substituting variables and projecting on the subspace

$$\mathcal{H} = \{(x, v) \in \mathbb{R}^{Nd} | x_1 + ... + x_N = 0 \text{ and } v_1 + ... + v_N = 0\}$$

of codimension $2d$, we obtain what we were seeking :

Proposition 12. *Define $\phi(z_1, ..., z_{N-1}) = \sum_{i=1}^{N-1} |z_i|^2 + \left|\sum_{i=1}^{N-1} z_i\right|^2$.*
The probability measure $\hat{\mu}_\beta$ on $(\mathbb{R}^d \times \mathbb{R}^d)^{N-1}$ whose density is

$$\hat{f}_\beta(\hat{x}_1, ..., \hat{x}_{N-1}, \hat{v}_1, ..., \hat{v}_{N-1}) \;=\; \frac{1}{Z}\, exp\left(-\lambda\ \left(\phi(\hat{x}_1, ..., \hat{x}_{N-1}) + \beta\ \phi(\hat{v}_1, ..., \hat{v}_{N-1})\right)\right),$$

where Z is a renormalisation constant, is invariant for the projection on $\mathcal{H}$ of the stochastic dynamics defined in (1.3.2).

Having found this elusive invariant probability measure, we wish to determine its rate of convergence towards the associated semi-group. This is where Lyapunov function theory comes in.

1.3.2 Lyapunov functions and ergodicity

Even restricting oneself to the field of stochastic analysis, one can find a plethora of definitions for Lyapunov functions, however similar they usually are. Thus, we first precise in exactly what sense we mean it here.

Definition 2. A positive, continuous, smooth enough, function V is called Lyapunov function for the system associated with the infinitesimal generator L if there exists $K \geqslant 0$ such that, outside of a certain compact set U,

$$LV \leqslant -K\,V.$$

Remark 4. Given a compact U, one can always find a ball $B(0,R)$, with $R>0$ such that $U \subset B(0,R)$.

The existence of Lyapunov functions associated with an infinitesimal generator is strongly linked with the ergodicity of the system. This can be seen in the following theorem, the main result of this section.

Theorem 5. *Let P_t^β be the semi-group associated with the system (1.3.2). Assume that $\lambda^2 > 2\beta$. For all $(\tilde{x},\tilde{v}) \in (\mathbb{R}^d \times \mathbb{R}^d)^{N-1}$, $P_t^\beta((\tilde{x},\tilde{v}),.)$ converges exponentially towards μ_β for the total variation distance : there exists $\rho > 0$ and $C > 0$ such that for all t :*

$$\|P_t^\beta((\tilde{x},\tilde{v}),.) - \mu_\beta\|_{TV} \leqslant C\,V(\tilde{x},\tilde{v})\,e^{-\rho\,t}$$

where V is the Lyapunov function defined in (1.3.3), associated with the stochastic system (1.3.2).
In this case, we shall say that μ_β is exponentially ergodic.

Proof. The proof rest on the following theorem, due to Down, Meyn and Tweedie (see [13]) ; it also appears in other papers, such as [4]. It explicits the link between Lyapunov functions and ergodicity.

Theorem 6 (Down, Meyn, Tweedie, 1995)**.** *Let a process be irreducible, in the sense defined in [13]. Let L be its infinitesimal generator and $(P_t)_{t\geqslant 0}$ its semi-group. Assume that it admits an invariant probability measure μ. Assume, in addition, that there exists some Lyapunov function V for the generator L.*
Then, μ is exponentially ergodic.

What is left to prove is that, if $\lambda^2 > 2\beta$, the function V defined by

$$V(x,v) = exp\left(\sum_i (\frac{1}{2}\beta\lambda\,|x_i|^2 + \beta\,x_i v_i + \frac{1}{2}\lambda\,|v_i|^2)\right) \tag{1.3.3}$$

is a Lyapunov function for the system (1.3.2).
The infinitesimal generator associated with system (1.3.2) is, for all f regular enough,

$$L_\beta f(x,v) = \frac{1}{2}\sum_{i,\alpha}\left(\partial^2_{v_i^\alpha} f(x,v) - \frac{1}{N}\sum_j \partial^2_{v_i^\alpha v_j^\alpha} f(x,v)\right) + \sum_{i,\alpha} v_i^\alpha \partial_{x_i^\alpha} f(x,v) - \sum_{i,\alpha}(\lambda v_i^\alpha + \beta x_i^\alpha)\partial_{v_i^\alpha} f(x,v),$$

where $i \in \{1,...,N\}$ and $\alpha \in \{1,...,d\}$.

We start by computing $L_\beta V$ for every $(x,v)\in\mathbb{R}^{2Nd}$:

$$\begin{aligned}
L_\beta V(x,v) &= \frac{1}{2}\sum_{i,\alpha}\left(\lambda+(\beta\, x_i^\alpha+\lambda\, v_i^\alpha)^2-\frac{1}{N}\sum_j(\beta\, x_i^\alpha+\lambda\, v_i^\alpha)\left(\beta\, x_j^\alpha+\lambda\, v_j^\alpha\right)\right)\, V(x,v)\\
&\quad +\sum_{i,\alpha}(\beta(v_i^\alpha)^2+\lambda v_i^\alpha x_i^\alpha)\, V(x,v)-\sum_{i,\alpha}(\lambda v_i^\alpha+\beta x_i^\alpha)^2\, V(x,v)\\
&= \left[\frac{1}{2}\lambda Nd-\frac{1}{2}\sum_{i,\alpha}(\lambda v_i^\alpha+\beta x_i^\alpha)^2-\frac{1}{N}\sum_\alpha\left(\sum_i(\beta x_i^\alpha+\lambda v_i^\alpha)\right)^2\right.\\
&\qquad\left.+\sum_{i,\alpha}(\beta(v_i^\alpha)^2+\lambda v_i^\alpha x_i^\alpha)\right]V(x,v)\\
&\leqslant \left[\frac{1}{2}\lambda Nd-\frac{1}{2}\sum_{i,\alpha}\left(-\lambda^2(v_i^\alpha)^2-\beta^2(x_i^\alpha)^2+2\beta(v_i^\alpha)^2\right)\right]V(x,v)\\
&= -\frac{1}{2}\left((\lambda^2-2\beta)|v|^2+\beta^2|x|^2-\lambda Nd\right)V(x,v).
\end{aligned}$$

Thus, if $\lambda^2>2\beta$, setting $K=\min((\lambda^2-2\beta),\beta^2)$, when $|x|^2+|v|^2$ is large enough,

$$L_\beta V(x,v)\leqslant -K\, V(x,v).$$

□

1.4 Non-constant communication rate : two particular cases

We go back to the stochastic Cucker-Smale model (1.1.2) : we now turn our attention to non-constant, and, one hopes, more realistic communication rates.

First, we try to adapt the method of the previous paragraph, using some Lyapunov functions to obtain ergodicity, when considering only two particles. Even in this reduced setting, we are only able to obtain a polynomial convergence of the semi-group towards its invariant probability measure, which we are not able to explicit.

Second, we consider as communication rate a small perturbation of a constant one. Applying results from [20], based on the cluster expansion method, from statistical physics, we obtain some exponential ergodicity for general drifts with finite delay.

1.4.1 One (or two) particle(s) along the real line

Consider (x_1,v_1) and (x_2,v_2) satisfying a generalized version of (1.3.1) for $N=2$ and $d=1$, for a communication rate $\psi(x,y)=\dfrac{\lambda}{(1+(x-y)^2)^\gamma}$, similar to those introduced by Cucker and Smale in their original model (see [12] and [11]). It can also be seen as equation 1.1.2 in which is added the term in β introduced in the previous section.

If we set $x = x_1 - x_2$ and $v = v_1 - v_2$, then x and v form a solution of the stochastic differential system :

$$\begin{cases} dx_t &= v_t\, dt \\ \\ dv_t &= -\dfrac{\lambda\, v_t}{(1+x_t^2)^\gamma}\, dt - \beta x_t\, dt + \sqrt{D} dW_t. \end{cases} \tag{1.4.1}$$

One can also consider these equations as the modelization of a single particle moving along the real line, according to some version of the modified stochastic Cucker-Smale model (1.3.1), studied in the previous section.

The main result of this paragraph is the following convergence theorem :

Theorem 7. *We define the function* ϕ_γ *by*

$$\phi_\gamma(t) = \begin{cases} (\gamma\, t + 1)^{-\frac{1-\gamma}{\gamma}} & \text{for } \gamma \leqslant \dfrac{1}{2} \\ \left(\dfrac{4\gamma - 1}{4\gamma}\, t + 1\right)^{-\frac{1}{4\gamma-1}} & \text{for } \gamma \geqslant \dfrac{1}{2}. \end{cases} .$$

The process (x_t, v_t) *defined by (1.4.1) admits an invariant probability measure, called* μ_γ*. Moreover, the semi-group converges towards* μ_γ *for the total variation distance and the convergence rate is at least* ϕ_γ*.*

To prove this result, we first establish a criterion for the existence of an invariant probability measure, before applying it to our system (1.4.1). Then, we prove the polynomial ergodicity.

A criterion for the existence of an invariant probability measure

We first give in the theorem below a sufficient condition for the existence of an invariant probability measure. Though this result is part of the folklore, we did not find a proof of it in the literature. The one we propose is adapted from Theorem 12.3.4 in Meyn and Tweedie's book [19], from discrete time to continuous time.

Theorem 8. *Let* $(U_t)_{t\geqslant 0}$ *be a Feller Markov process on* Ω*, whose infinitesimal generator* L *is such that there exists a non-negative function* V*, a real* b *and a compact set* C *satisfying, for every* u*,*

$$LV(u) \leqslant -1 + b\, \mathbf{1}_C(u) \tag{1.4.2}$$

Then, the process $(U_t)_{t\geqslant 0}$ *admits a unique invariant probability measure on* Ω*.*

Proof. We begin by showing the following proposition :

Proposition 13. *If a Feller Markov process does not admit an invariant probability measure, then for any* u *and any compact set* C*,*

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t P_s(u, C)\, ds = 0.$$

Proof. We prove this result by contradiction : we suppose that the aforementioned limit is not 0 and that there is no invariant probability measure for the process considered.

Thereby, we can find a continuous function f with compact support, a positive number δ and two sequences $(t_i)_{i\in\mathbb{N}}$ and $(y_i)_{i\in\mathbb{N}}$ satisfying, for every i, $\rho_i(f) = \frac{1}{t_i}\int_0^{t_i} P_s f(y_i)\, ds > \delta$. There exists a subsequence $(n_i)_{i\in\mathbb{N}}$, and a subprobability ρ_∞ such that $(\rho_{n_i})_{i\in\mathbb{N}}$ converges vaguely towards ρ_∞ (that is, for every continuous f which goes to 0 at infinity, $\rho_{n_i}(f) \longrightarrow \rho_\infty(f)$), thanks to Proposition D.5.6(i) of [19]).

First, remark that $\rho_\infty(f) \geqslant \liminf_{i\to\infty} \rho_{n_i}(f) \geqslant \delta > 0$, which implies $\rho_\infty \neq 0$.

Next, for every non-negative r, for every g continuous with compact support,

$$\begin{aligned}\rho_{n_i}(P_r g) &= \frac{1}{t_i}\int_0^{t_i} P_{s+r} g(y_i)\, ds = \frac{1}{t_i}\int_r^{t_i+r} P_u g(y_i)\, du \\ &= \frac{1}{t_i}\int_0^{t_i+r} P_u g(y_i)\, du - \frac{1}{t_i}\int_0^{r} P_u g(y_i)\, du \longrightarrow \rho_\infty(g),\end{aligned}$$

as the first part goes towards $\rho_\infty(g)$ and the second to 0 when i goes to infinity.

Furthermore, as we deal here with a Feller process, for all r, $P_r g$ is continuous ; thus, $\rho_\infty(P_r g) \leqslant \liminf_{i\to\infty} \rho_{n_i}(P_r g)$, by Lemma D.5.5 of [19], and $\rho_\infty(P_r g) \leqslant \rho_\infty(g)$.

As $\rho_\infty \circ P_r(1) = \rho_\infty(1)$, $\bar{\rho_\infty} \circ P_r(g) \leqslant \bar{\rho}_\infty(g)$ where $\bar{\rho}_\infty$ is the probability $\frac{\rho_\infty}{\rho_\infty(1)}$.

Finally, for every non-negative r, $\bar{\rho}_\infty(P_r g) = \bar{\rho}_\infty(g)$.

As $\bar{\rho}_\infty$ is not the trivial measure, it is an invariant probability measure, which contradicts the hypothesis and concludes the proof of the proposition. □

We now prove the theorem. Thanks to Itô's formula,

$$\begin{aligned}\mathbb{E}V(U_t) &= \mathbb{E}V(U_0) + \int_0^t \mathbb{E}[LV(U_s)]\, ds \\ &= V(u) + \int_0^t \int V(y) P_s(u, dy)\, ds \\ &\leqslant V(u) - t + b\int_0^t P_s(u, C)\, ds \text{ , with the help of (1.4.2).}\end{aligned}$$

Hence,

$$\begin{aligned}\frac{1}{t}\int_0^t P_s(u, C)\, ds &\geqslant \frac{1}{b}\left(1 + \frac{1}{t}(\mathbb{E}V(U_t) - V(u))\right) \\ &\geqslant \frac{1}{b}\left(1 - \frac{1}{t}V(u)\right) \xrightarrow[t\to\infty]{} \frac{1}{b}.\end{aligned}$$

Thus,

$$\liminf_{t\to\infty} \frac{1}{t}\int_0^t P_s(u,C)\, ds \geqslant \frac{1}{b} > 0.$$

We then apply the proposition, which gives us the existence of an invariant probability measure.

□

Application to our system (1.4.1)

The infinitesimal generator associated with (1.4.1) is the differential operator defined by

$$L_{x,v} = D\,\partial_v^2 + v\,\partial_x - \frac{\lambda v}{(1+x^2)^\gamma}\,\partial_v - \beta x\,\partial_v.$$

Set

$$V_\gamma : (x,v) \mapsto \beta x^2 + \lambda\, f_\gamma(x)\, v + v^2,$$

where f_γ is the primitive of $\psi_\gamma : x \mapsto \dfrac{1}{(1+x^2)^\gamma}$ that vanishes at 0 (which exists by continuity of ψ_γ).

Applying the generator L to V_γ, we obtain

$$LV_\gamma(x,v) = D - \frac{\lambda v^2}{(1+x^2)^\gamma} - \frac{\lambda^2 v}{(1+x^2)^\gamma}\, f_\gamma(x) - \lambda\beta\, x f_\gamma(x). \tag{1.4.3}$$

The proposition below guarantees that Theorem 8 can be applied.

Proposition 14. *For a positive large enough R , if $\max(|x|,|v|) > R$, then*

$$LV_\gamma(x,v) \leqslant -1.$$

Proof. We begin by giving a few properties of f_γ : as ψ_γ is positive, f_γ is increasing, with $f_\gamma(x)$ positive (resp. negative) if x is positive (resp. negative). In particular, for every x, $xf_\gamma(x)$ is non-negative. Furthermore, ψ_γ is an even function, making f_γ an odd one.

ψ_γ is integrable on $\mathbb{R}$ if, and only if, $\gamma > \dfrac{1}{2}$; in this case f_γ tends towards the finite number $\displaystyle\int_0^\infty \frac{1}{(1+u^2)^\gamma}\, du$ when x goes to infinity. When $\gamma < \dfrac{1}{2}$, there exists a positive C_γ such that for x large, $f_\gamma(x) \sim C_\gamma x^{1-2\gamma}$.

Suppose that $\max(|x|,|v|) > R$.

- If $|v| < R$, from (1.4.3),

$$\begin{aligned} LV_\gamma(x,v) &\leqslant 2D - \lambda\, f_\gamma(x)\left(\beta\, x + \frac{\lambda\, v}{(1+x^2)^\gamma}\right) \\ &\leqslant 2D - \lambda\,|f_\gamma(x)|\left(\beta\,|x| - \frac{\lambda\,|v|}{(1+x^2)^\gamma}\right) \\ &\leqslant 2D - \lambda\,|f_\gamma(x)|\left(\beta\, R - \frac{\lambda\, R}{(1+R^2)^\gamma}\right). \end{aligned}$$

We briefly notice that $\dfrac{\lambda\, R}{(1+R^2)^\gamma} \sim_{R\to\infty} \lambda\, R^{1-2\gamma}$ and $1-2\gamma<1$ if and only if $\gamma<0$.

Whatsoever,

$$\beta\, R \geqslant \frac{\lambda\, R}{(1+R^2)^\gamma} \Leftrightarrow (1+R^2)^\gamma \geqslant \frac{\lambda}{\beta} \Leftrightarrow R \geqslant \sqrt{1+\left(\frac{\lambda}{\beta})\right)^\gamma}.$$

For $R \geqslant \sqrt{1+\left(\frac{\lambda}{\beta})\right)^\gamma}$, $LV_\gamma(x,v) \leqslant 2D - \lambda\,|f_\gamma(R)|\left(\beta\, R - \dfrac{\lambda\, R}{(1+R^2)^\gamma}\right)$.

Hence, if R is such that $R \geqslant \sqrt{1+(\lambda/\beta)^\gamma}$ and $2D - \lambda\,|f_\gamma(R)|\,(\beta\, R - \dfrac{\lambda\, R}{(1+R^2)^\gamma}) \leqslant -1$, we have

$$LV_\gamma(x,v) \leqslant -1.$$

- If $|x|<R$, for every positive γ, $|f_\gamma(x)| \in o(x)$; therefore, for R sufficiently large,

$$|v| > \lambda\,|f_\gamma(x)| \text{ and } \lambda\,|f_\gamma(x)| \in o(|v|). \tag{1.4.4}$$

Accordingly, on the one hand, if $\gamma<1$,

$$\begin{aligned} LV_\gamma(x,v) &\leqslant 2D - \frac{\lambda\, v}{(1+x^2)^\gamma}\,(v+\lambda\, f_\gamma(x)) \\ &\leqslant 2D - \frac{\lambda\,|v|}{(1+x^2)^\gamma}\,(|v|-\lambda\,|f_\gamma(x)|) \\ &\leqslant 2D - \frac{\lambda\, R}{(1+R^2)^\gamma}\,(R-\lambda\, f_\gamma(R)) \\ &\sim -\lambda\, R^{2(1-\gamma)} \text{ for } 2(1-\gamma)>0, \text{ that is } \gamma<1 \end{aligned}$$

We can thus find R such that $LV_\gamma(x,v) \leqslant -1$.

On the other hand, when $\gamma \geqslant 1$, keeping (1.4.4) in mind,

$$LV_\gamma(x,v) \leqslant 2D - \frac{\lambda\,|v|}{(1+x^2)^\gamma}\,(|v|-\lambda\,|f_\gamma(x)|) - \lambda\beta\, x f_\gamma(x).$$

(i) If $v \leqslant x^{2\gamma}$, $LV_\gamma(x,v) \leqslant 2D - \lambda\beta\, f_\gamma(R^{\frac{1}{2\gamma}})R^{\frac{1}{2\gamma}} \leqslant -1$ for R large enough.

(ii) Otherwise, if $v > x^{2\gamma}$, $LV_\gamma(x,v) \leqslant 2D - \dfrac{\lambda\,|v|\,(|v|-\lambda\,|f_\gamma(x)|)}{(1+v^{\gamma^{-1}})^\gamma}$.
For $|v|$ sufficiently large, the fraction is equivalent to λv, which allows us to conclude.

- If $|x| > R$ and $|v| > R$, we will use Young's inequality : for every positive real numbers p and q satisfying $\frac{1}{p} + \frac{1}{q} = 1$,

$$\frac{\lambda^2 \, |v|}{(1+x^2)^\gamma} \, |f_\gamma(x)| \leqslant \frac{1}{p} \frac{\lambda^p \, |v|^p}{(1+x^2)^{\gamma p}} + \frac{1}{q} \lambda^q \, |f_\gamma(x)|^q.$$

Subsequently,

$$\begin{aligned} LV_\gamma(x,v) &\leqslant 2D - \left(\frac{\lambda \, v^2}{(1+x^2)^\gamma} - \frac{1}{p} \frac{\lambda^p \, |v|^p}{(1+x^2)^{\gamma p}} \right) - \left(\lambda\beta \, x f_\gamma(x) - \frac{1}{q} \lambda^q \, |f_\gamma(x)|^q \right) \\ &= 2D - a_1(x,v) - a_2(x,v). \end{aligned}$$

Roughly, $a_1(x,v) \approx \frac{\lambda \, v^2}{|x|^{2\gamma}} - \frac{\lambda^p}{p} \frac{|v|^p}{|x|^{2\gamma p}}$ and $a_2(x,v) \approx C_\gamma \lambda\beta \, |x|^{1+\max(0,1-2\gamma)} - C_\gamma^q \frac{\lambda^q}{q} \, |x|^{q\max(0,1-2\gamma)}$ where C_γ is a positive real number depending only on γ (α_γ if $\gamma < \frac{1}{2}$, else $f_\gamma(\infty)$).

We would like to find p and q such that a_1 and a_2 tends to infinity when x and v do, regardless of the ratio x/v.

We will obtain such a thing if each of the following assumptions is satisfied :

(a) $\frac{1}{p} + \frac{1}{q} = 1$

(b) $p < 2$

(c) $2\gamma < 2\gamma p$

(d) $q\max(0,1-2\gamma) < 1 + \max(0,1-2\gamma)$.

Conditions (b) and (c) come from the expression of a_1, (d) from the expression of a_2.

As $\gamma \neq 0$, (b) and (c) can be summed up by (e) : $1 < p < 2$.

If $1 - 2\gamma > 0$, ie $\gamma < \frac{1}{2}$,

$$q < \frac{2(1-\gamma)}{1-2\gamma} \Leftrightarrow 1 - \frac{1}{p} > \frac{1-2\gamma}{2(1-\gamma)} \Leftrightarrow \frac{1}{p} < \frac{1}{2(1-\gamma)} \Leftrightarrow p > 2(1-\gamma).$$

Thus, for every $\gamma > 0$, $p \in (\max(1, 2(1-\gamma)), 2)$.

When $\gamma < \frac{1}{2}$, we set $p = 2 - \gamma$, hence $q = \frac{2-\gamma}{1-\gamma}$; otherwise, we choose $p = \frac{3}{2}$ and $q = 3$.

Next, we check that we indeed observe the behaviour we were looking for :

If $\gamma > \dfrac{1}{2}$, $a_1(x,v) \leqslant \dfrac{\lambda v^2}{(1+x^2)^\gamma} - \dfrac{3\lambda^{\frac{3}{2}}}{2}\dfrac{|v|^{\frac{3}{2}}}{(1+x^2)^{\frac{3\gamma}{2}}}$ is non-negative for R sufficiently large and $a_2(x,v) = \lambda\beta\, x f_\gamma(x) - \dfrac{\lambda^3}{3} f_\gamma(\infty)^3 \geqslant 2D+1$ when x is large enough, and we are therefore able to conclude in this situation.

A similar verification can be done when $\gamma < \dfrac{1}{2}$.

Finally, for every positive real number γ, there exists some positive real number R such that

$$LV_\gamma \leqslant -1 + \delta \mathbf{1}_{B_\gamma}$$

where $B_\gamma = \{(x,v) \in \mathbb{R}^2 |\, \max(|x|,|v|) \leqslant R\}$ and δ a real number.

□

B_γ being a compact set of $\mathbb{R}^2$, and V_γ being bounded on $\bar{B}(0,R)$, the stochastic dynamical system (1.4.1) admits an invariant probability measure, thanks to Theorem 8.

Now, we know there exists an invariant measure, μ_γ ; we would like to go one step further and find a convergence rate for the semi-group towards this probability measure.

Polynomial ergodicity

Recall that $\phi_\gamma(t) = \begin{cases} (\gamma\, t + 1)^{-\frac{1-\gamma}{\gamma}} & \text{for } \gamma \leqslant \dfrac{1}{2} \\ \left(\dfrac{4\gamma - 1}{4\gamma}\, t + 1\right)^{-\frac{1}{4\gamma - 1}} & \text{for } \gamma \geqslant \dfrac{1}{2}. \end{cases}$

The proof of the second part of Theorem 7 follows from Theorem 1.2 of [4] and the following proposition.

Proposition 15. *Let H be the function defined on $\mathbb{R}^+$ by $H(u) = |u|^{1-\gamma}$ if $\gamma \leqslant \dfrac{1}{2}$ and $H(u) = |u|^{\frac{1}{4\gamma}}$ if $\gamma \geqslant \dfrac{1}{2}$.*
For a positive, large enough, R , if $\max(|x|,|v|) > R$, then

$$\forall (x,v) \in \mathbb{R}^2, \quad LV_\gamma(x,v) \leqslant -K_\gamma\, H(V_\gamma(x,v))) \tag{1.4.5}$$

where K_γ is a positive constant, depending only on γ.

Proof. We prove the proposition. It will follow a pattern similar to the one of proposition 14.

Notice that H is a non-negative, increasing and concave map.
Suppose that $\max(|x|,|v|) > R$.

- If $\gamma \leqslant \frac{1}{2}$, we would like to prove that there exists K_γ such that for R large enough,

$$K_\gamma(\beta x^2 + \lambda f_\gamma(x)v + v^2)^{1-\gamma} \leqslant -D + \frac{\lambda v^2}{(1+x^2)^\gamma} + \frac{\lambda^2 v}{(1+x^2)^\gamma} f_\gamma(x) + \lambda\beta x f_\gamma(x). \quad (1.4.6)$$

  – Suppose that $|x| < |v|$ and $|v| > R$.
  On the one hand, for R such that $\frac{\lambda f_\gamma(R)}{R} < 1$,

$$(\beta x^2 + \lambda f_\gamma(x)v + v^2)^{1-\gamma} = |v|^{2(1-\gamma)} \left(1 + \frac{\beta x^2}{v^2} + \frac{\lambda f_\gamma(x)}{v}\right)^{1-\gamma} \leqslant (2+\beta)^{1-\gamma}|v|^{2(1-\gamma)}.$$

  On the other hand, if R is sufficiently large,

$$-D + \frac{\lambda v^2}{(1+x^2)^\gamma} + \frac{\lambda^2 v}{(1+x^2)^\gamma} f_\gamma(x) + \lambda\beta x f_\gamma(x) \geqslant -D + \frac{\lambda v^2}{2(1+x^2)^\gamma} \geqslant \frac{\lambda v^2}{4(1+x^2)^\gamma}.$$

  Furthermore, if $R \geqslant 1$,

$$\frac{\lambda v^2}{4(1+x^2)^\gamma} \geqslant \frac{\lambda v^2}{4 \times 2^\gamma \max(1, x^2)^\gamma} \geqslant \frac{1}{8}\lambda v^2 \min(1, x^{-2\gamma}) \geqslant \frac{1}{8}\lambda v^2 \min(1, |v|^{-2\gamma}) \geqslant \frac{1}{8}\lambda |v|^{2(1-\gamma)}.$$

  Thus, with, for instance, $K_\gamma = \frac{1}{16(2+\beta)^{1-\gamma}}$, (1.4.6) is satisfied.

  – Suppose that $|v| \leqslant |x|$ and $|x| > R$.
  We proceed in exactly the same way, swapping x and v, to obtain the inequality

$$(\beta x^2 + \lambda f_\gamma(x)v + v^2)^{1-\gamma} \leqslant (2+\beta)^{1-\gamma}|x|^{2(1-\gamma)}.$$

  Besides, with similar arguments as those previously used,

$$-D + \frac{\lambda v^2}{(1+x^2)^\gamma} + \frac{\lambda^2 v}{(1+x^2)^\gamma} f_\gamma(x) + \lambda\beta x f_\gamma(x) \geqslant -D + \frac{1}{2}\lambda\beta|x| \times C_\gamma |x|^{1-2\gamma}$$

  where C_γ is a positive constant such that $f_\gamma(|x|) \sim C_\gamma |x|^{1-2\gamma}$ when $|x|$ is quite large.

  Thus,

$$-D + \frac{\lambda v^2}{(1+x^2)^\gamma} + \frac{\lambda^2 v}{(1+x^2)^\gamma} f_\gamma(x) + \lambda\beta x f_\gamma(x) \geqslant \frac{1}{4}\lambda\beta C_\gamma |x|^{2(1-\gamma)} \geqslant K_\gamma (2+\beta)^{1-\gamma}|x|^{2(1-\gamma)}$$

  for R big enough and K_γ below $\frac{\lambda\beta C_\gamma}{4(2+\beta)^{1-\gamma}}$, which implies inequality (1.4.6).

- If $\gamma \geqslant \frac{1}{2}$, we aim to show that we can find a positive constant K_γ such that for R large enough,

$$K_\gamma(\beta x^2 + \lambda f_\gamma(x)v + v^2)^{\frac{1}{4\gamma}} \leqslant -D + \frac{\lambda v^2}{(1+x^2)^\gamma} + \frac{\lambda^2 v}{(1+x^2)^\gamma} f_\gamma(x) + \lambda\beta x f_\gamma(x). \quad (1.4.7)$$

– Suppose that $|v| \leqslant |x|$ and $|x| > R$.

Then, for a large enough R,

$$(\beta x^2 + \lambda f_\gamma(x)v + v^2)^{\frac{1}{4\gamma}} \leqslant (2+\beta)^{\frac{1}{4\gamma}} |x|^{\frac{1}{2\gamma}}.$$

Moreover,

$$-D + \frac{\lambda v^2}{(1+x^2)^\gamma} + \frac{\lambda^2 v}{(1+x^2)^\gamma} f_\gamma(x) + \lambda\beta x f_\gamma(x) \geqslant \frac{1}{4}\lambda\beta C_\gamma |x|$$

when R is large enough, with $C_\gamma = \lim_{|x|\to\infty} f_\gamma(|x|)$.

Hence (1.4.7), with $K_\gamma \leqslant \dfrac{\lambda\beta C_\gamma}{4(2+\beta)^{\frac{1}{\gamma}}}$.

– Suppose that $|x| < |v|$ and $|v| > R$. We have, by analogy with previous assumptions,

$$(\beta x^2 + \lambda f_\gamma(x)v + v^2)^{\frac{1}{4\gamma}} \leqslant (2+\beta)^{\frac{1}{4\gamma}} |v|^{\frac{1}{2\gamma}}.$$

Furthermore,

$$-D + \frac{\lambda v^2}{(1+x^2)^\gamma} + \frac{\lambda^2 v}{(1+x^2)^\gamma} f_\gamma(x) + \lambda\beta x f_\gamma(x) \geqslant \frac{\lambda v^2}{2(1+x^2)^\gamma} + \frac{1}{2}\lambda\beta C_\gamma |x|.$$

(i) If $|x^{2\gamma} \leqslant |v|$, then

$$\frac{\lambda v^2}{2(1+x^2)^\gamma} + \frac{1}{2}\lambda\beta C_\gamma |x| \geqslant \frac{\lambda v^2}{2(1+v^{\frac{1}{\gamma}})^\gamma} \geqslant \frac{\lambda v^2}{2(2v^{\frac{1}{\gamma}})^\gamma} \geqslant \frac{1}{8}|v| \geqslant K_\gamma (2+\beta)^{\frac{1}{4\gamma}} |v|^{\frac{1}{2\gamma}}$$

with $K_\gamma = \dfrac{1}{8(2+\beta)^{\frac{1}{4\gamma}}}$, as $2\gamma \leqslant 1$ and (1.4.7) is satisfied.

(ii) If $|x^{2\gamma} > |v|$, we have

$$\frac{\lambda v^2}{2(1+x^2)^\gamma} + \frac{1}{2}\lambda\beta C_\gamma |x| \geqslant \frac{1}{2}\lambda\beta C_\gamma |x| \geqslant \frac{1}{2}\lambda\beta C_\gamma |v|^{\frac{1}{2\gamma}} \geqslant K_\gamma (2+\beta)^{\frac{1}{4\gamma}} |v|^{\frac{1}{2\gamma}}$$

for $K_\gamma = \dfrac{\lambda\beta C_\gamma}{2((2+\beta)^{\frac{1}{4\gamma}})}$ and R large enough, ensuring us of the validity of inequality (1.4.7).

□

For two different values of γ, we illustrate the veracity of the proposition we have just proven. These graphics were realised with Matlab ; in blue/dark is LV_γ, in green/light is $-K_\gamma H(V_\gamma)$.

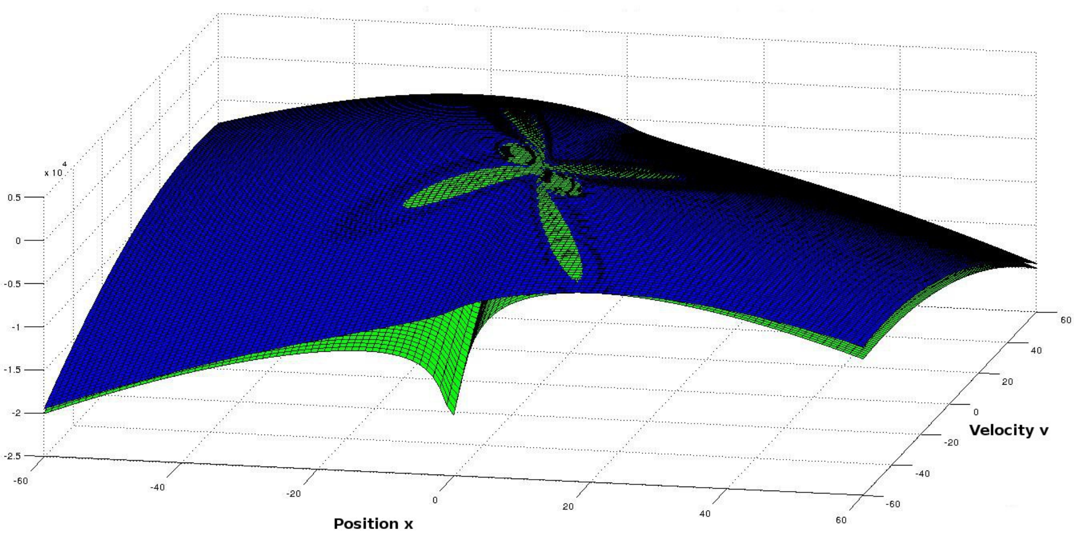

Figure 1.1: Case $\gamma = 0.2$, with $D = 0.1$, $\lambda = 5$, $\beta = 2$ and $K_\gamma = 8.3$.

We indeed observe that, in both situations, when we are far enough from the origin, the green/lighter surface is under the blue/darker one, which illustrates the drift condition shown in this section.

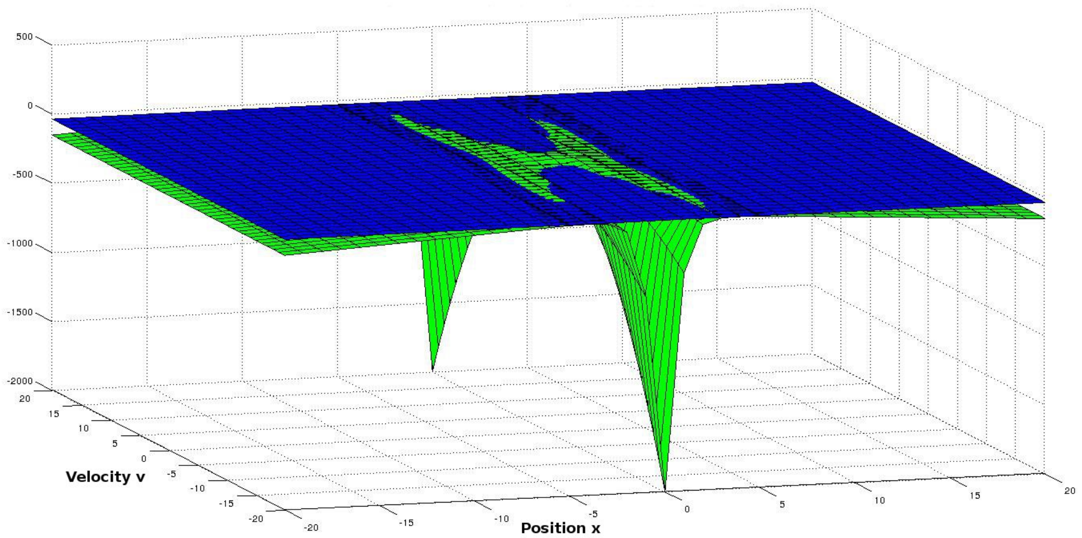

Figure 1.2: Case $\gamma = 2$, with $D = 0.1$, $\lambda = 5$, $\beta = 2$ and $K_\gamma = 17$.

Thanks to Theorem 1.2 of [4], we know that the semi-group associated with the process converges towards the invariant measure ; furthermore we can obtain an indication about the convergence rate, hence ϕ_γ.

Remark 5. It should be noticed that other communication rates ψ give similar results as those obtained here, as long as they satisfy the following hypotheses, where Ψ is the primitive of ψ vanishing at 0 :

- ψ is bounded
- Ψ is non-decreasing
- for all $x \in \mathbb{R}$, $x\ \Psi(x) \geqslant 0$.

This is for instance the case of $\psi(x) = \dfrac{1}{1+\ln(1+x^2)^b}$ for a non-negative b.

Remark 6. Results of this subsection are only valid for $d = 1$. Indeed, we are not able to find Lyapunov functions in higher dimensions.

1.4.2 Ergodicity for small perturbations : the cluster expansion method

In this section, we look to apply the cluster expansion method to our problem : we start from a well-known symmetric diffusion, the case of the constant communication rate. Our aim is to disrupt it through a small perturbation with finite delay t_0, to obtain a perturbation of a stochastic Cucker-Smale model whose drift has a finite delay t_0.

For the sake of simplicity, computations are done here in the case $d = 1$. The symmetry of the system ensures it is possible without loss of generality on the final result, even though some constants depend on d.

We saw in paragraph 1.2.2 that the system corresponding to the communication rate $\psi = \lambda$,

$$d\hat{v}_i(t) = -\lambda \hat{v}_i(t)dt + d\widehat{W}_i(t), \quad i \in \{1, ..., N\} \tag{1.4.8}$$

admits a reversible probability measure, $\mu = \mathcal{N}\left(0, \dfrac{1}{2\lambda}\,\Pi_N\right)$, setting $\Pi_N = \Pi_{N,1}$, defined in (1.2.2).

In this subsection, we wish to apply results using the cluster expansion method established in [20] to obtain ergodicity for slight perturbations of the drift in this model.

Consider the dynamics :

$$d\hat{v}(t) = -\lambda \hat{v}(t)dt + \Pi_N\ dW(t), \quad t \in \mathbb{R}_+ \tag{1.4.9}$$

where $\hat{v} \in \mathbb{R}^N$ and W is a N-dimensional standard Brownian motion.

The law of the Ornstein-Uhlenbeck process studied in section 1.2 is degenerate - in particular, the $N \times N$ matrix Π_N is not invertible (see Remark 1), nor is $\Pi_N \Pi_N^*$ (and a projection on the first $N-1$ coordinates of Π_N satisfies neither of these requirements). Thus, in order to apply results from [20], we have to project our system on an ad hoc subspace, where the process will be elliptic, once we have established an adequate orthonormal basis on it. We then introduce the perturbation on this subspace.

As previously mentioned, the vector of the microscopic velocities $(\hat{v}_1, ..., \hat{v}_N)$ is living on the hyperplane $H = \{ v \in \mathbb{R}^N \mid v_1 + ... + v_N = 0 \}$. One can check that the set $(e_i)_{i\in\{1,...,N-1\}}$ defined by

$$e_i^j = \sqrt{\frac{i}{i+1}} \left(\frac{1}{i} \delta_{j \leqslant i} - \delta_{j=i+1} \right)$$

for every $j \in \{1, ..., N\}$, is an orthonormal basis of H.

In what follows, we set $\alpha_i = \sqrt{\frac{i}{i+1}}$. Let e_N be the element of $\mathbb{R}^N$ such that $e_N^j = \frac{1}{\sqrt{N}}$ for every j. Then $(e_i)_{i\in\{1,...,N\}}$ is an orthonormal basis of $\mathbb{R}^N$.

The microscopic system $\hat{v}(t)$ has in the basis $(e_i)_{i\in\{1,...,N\}}$ the coordinates $u_i(t) = e_i^* \hat{v}(t)$, $i = 1, ..., N$.

Thus,

$$u_N(t) = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \hat{v}_k(t)$$

and for $i \in \{1, ...N-1\}$,

$$u_i(t) = \alpha_i \left(\frac{1}{i} \sum_{k=1}^{i} \hat{v}_k - \hat{v}_{i+1} \right).$$

This means that, on the one hand,

$$u_N(t) = \sqrt{\frac{1}{N}} \sum_{k=1}^{N} \widehat{W}_k(t) = 0$$

and on the other hand,

$$du_i(t) = \alpha_i \left(\frac{1}{i} \sum_{k=1}^{i} \left(-\lambda \hat{v}_k(t)\, dt + (dW_k(t) - \frac{1}{N} \sum_{j=1}^{N} dW_j(t)) \right) \right.$$
$$\left. + \lambda \hat{v}_{i+1}(t)\, dt - (dW_{i+1}(t) - \frac{1}{N} \sum_{j=1}^{N} dW_j(t)) \right)$$

so that

$$du_i(t) = -\lambda u_i(t)\, dt + \alpha_i \left(\frac{1}{i} \sum_{k=1}^{i} dW_k(t) - dW_{i+1}(t) \right).$$

Setting $U = (u_1, ..., u_{N-1})$, it satisfies in $\mathbb{R}^{N-1}$

$$dU(t) = -\lambda U(t)\, dt + \sigma\, dW(t) \tag{1.4.10}$$

with σ the $(N-1) \times N$ matrix whose j-th row is $\alpha_j\ e_j^*$. The system is now non degenerate : $\sigma\sigma^*$ is invertible.

U is another Ornstein-Uhlenbeck type process (different from (1.4.9)) : as was done in section 1.2, one can give the explicit expression of U, its expectation and its covariance matrix :

Proposition 16. *For every $t \geqslant 0$,*

$$U(t) = e^{-\lambda t}\, U(0) + \int_0^t e^{-\lambda(t-s)}\, \sigma\, dW(s),$$

$$U(t) \sim \mathcal{N}\left(e^{-\lambda t}\, \mathbb{E}[U(0)]\ ,\ \frac{1}{2\lambda}\,(1 - e^{-2\lambda t})\, I_{N-1}\right).$$

Finally, we find a reversible probability measure for U. This proposition can be proven using the same tools as before.

Proposition 17. $\rho = \mathcal{N}(0, \frac{1}{2\lambda}\, I_{N-1})$ *is a reversible probability measure for the process U defined in (1.4.10).*

One can check that all the hypotheses required for Theorem 2 in [20] are satisfied. As a consequence, theorem 9 holds :

Theorem 9. *Assume that $b : \mathcal{C}(\mathbb{R}_+, \mathbb{R}^{N-1}) \to \mathbb{R}^{N-1}$ is a measurable function, bounded by 1, and local, in the sense that there exists $t_0 > 0$ such that, for any $u \in \Omega$, $b(u) = b((u)_{t-t_0}^t)$. Then, when β is small enough, the system with delay*

$$dZ(t) = (-\lambda Z(t) + \beta\, b((Z)_{t-t_0}^t)\, dt + \sigma\, dW(t),$$

where $(Z)_{t-t_0}^t$ is the trajectory of Z between times $t - t_0$ and t, admits a unique weak stationary solution Q, probability measure on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^{N-1})$.
Moreover, there is exponential ergodicity : there exist $\theta > 0$ and $C : \mathbb{R}^{N-1} \to \mathbb{R}_+$ such that for t and t' large enough, for every $z \in \mathbb{R}^{N-1}$, for every bounded measurable function f,

$$|\mathbb{E}_Q[f(Z(t))|Z(0) = z] - \mathbb{E}_Q[f(Z_{t'})|Z(0) = z]| \leqslant C(z)\, e^{-\theta\, |t-t'|}.$$

Finally, we go back to the canonical basis, where this result will hold for a certain class of perturbation drift.

Let, for $b = (b_1, ..., b_{N-1}) : \mathcal{C}(\mathbb{R}_+, \mathbb{R}^{N-1}) \to \mathbb{R}^{N-1}$, the function $B = (B_1, ..., B_{N-1}) : \mathcal{C}(\mathbb{R}_+, \mathbb{R}^{N-1}) \to \mathbb{R}^{N-1}$ be given by

$$B_i(.) = \sum_{j=i}^{N-1} \frac{1}{\sqrt{j(j-1)}}\, b_j(P\, .) - \sqrt{\frac{i-1}{i}}\, b_{i-1}(P\, .)$$

with $P = (P_{ij})_{i,j \in \{1,\dots,N\}}$ the square matrix of size N such that, for all $j \in \{1, ..., N\}$,

$$P_{ij} = \sqrt{\frac{i}{i+1}}\left(\frac{1}{i}\,\delta_{j \leqslant i} - \delta_{i=j+1}\right) \text{ if } i < N \quad \text{ and } \quad P_{Nj} = \frac{1}{\sqrt{N}}.$$

Corollary 1. *Assume that b is as in Theorem 9 and B as defined just above. Then, if β is small enough, the dynamics*

$$d\hat{v}(t) = \big(-\lambda \hat{v}(t) + \beta\, B\big((\hat{v})_{t-t_0}^t\big)\big)\, dt + \Pi_N\, dW(t)$$

admits a weak stationary solution and there is exponential ergodicity.

Remark 7. Following Scheutzow in [21] (see Theorem 3), we knew already that there exists a unique invariant probability measure for such dynamics. The only novelty here is the fact that the convergence happens at an exponential rate.

Remark 8. That we consider a finite delay instead of an unbounded one, as in the original Cucker-Smale model, is not that much of a stretch : indeed, one can imagine that the behaviour of a particle at time t depends on the difference of the positions at times t and $t - t_0$, for a certain t_0.

1.5 Stationarity solutions and moment controls

In the previous section, we have obtained very partial results in two particular situations ; here our goal is to obtain a more general result about the existence of an invariant probability measure, in a certain sense. To do this, we apply results from the work of Itô and Nisio ([17]) to prove the existence of stationary solutions, and thus of a certain form of invariant probability measures.

First, however, we introduce a few hypotheses :

- (H1) : There exists a even, positive, function $\bar{\psi} : \mathbb{R}^d \to \mathbb{R}$ such that, for all x and y, $\psi(x,y) = \bar{\psi}(x-y)$.

- (H2) : There exists two constants ψ_1 and ψ_2 such that, for all $s \in \mathbb{R}^d$, $0 < \psi_1 \leqslant \bar{\psi}(s) \leqslant \psi_2$.

- (H3) : $\bar{\psi}$ is bounded and Lipschitz continuous.

1.5.1 Stationarity results

We place ourselves in the general case of the microscopic velocities of the stochastic Cucker-Smale system (1.1.5), seen as a delayed equation with unbounded delay :

$$d\hat{v}_i(t) = -\frac{1}{N}\sum_{j=1}^{N} \widetilde{\psi}\left((\hat{v}_i)_{-\infty}^t, (\hat{v}_j)_{-\infty}^t\right)(\hat{v}_i(t) - \hat{v}_j(t))\,dt + d\widehat{W}_i(t), \quad i \in \{1, ..., N\} \tag{1.5.1}$$

where $(X)_{-\infty}^t = (X_s)_{s\in(-\infty,t]}$ and $\widetilde{\psi}\left((v)_{-\infty}^t, (v')_{-\infty}^t\right) := \psi(x_t, x'_t)$, and setting $\hat{v}(t) = \hat{v}(0)$ for every $t \leqslant 0$.

Theorem 10. *Assume (H1) and (H2). Then the delayed equation (1.5.1) admits at least one stationary solution.*

Proof. The key ingredient of the proof is Theorem 3 of [17] ; we give it here in the version that we will be interested in :

Theorem 11 (Itô-Nisio, 1964)**.** *Consider the stochastic differential equation*

$$dX_t = a((X)_{-\infty}^t)\,dt + b((X)_{-\infty}^t)\,dB_t, \quad t \in \mathbb{R}. \tag{1.5.2}$$

Let the following hypotheses be satisfied :

- *(H4) : $a(f)$ and $b(f)$ are continuous on the space of the continuous functions on $\mathbb{R}_-$.*
- *(H5) : there exist $M > 0$ and a bounded measure K with compact support on $\mathbb{R}$ such that for every continuous f,*

$$|a(f)|^2 + |b(f)|^2 \leqslant M + \int_{-\infty}^{0} |f(t)|^2\,dK_t.$$

- *(H6) : there is a uniform control of the second-order moments :*

$$\sup_{t\in\mathbb{R}} \mathbb{E}[X_t^2] < +\infty.$$

Then, equation (1.5.2) admits a stationary solution, that is, a solution that is invariant under the time shift.

An argument of weak compactness is at the center of its proof.

In our case, one can easily verify that (H4) is satisfied. Furthermore, so is (H5), when ψ is bounded by ψ_2, with $M = |(\delta_{i=j} - 1/N)_{i,j}|_\infty = 1 - 1/N$ and $dK_t = 4N^2\psi_2\,\delta_0(dt)$.

The crucial point to apply this result is the hypothesis (H6) : we will show in the next paragraph that it holds under more restrictive assumptions on ψ. Indeed, from Proposition 18, we will be able to state that

$$\mathbb{E}\left[|\hat{v}(t)|^2\right] \leqslant \mathbb{E}\left[|\hat{v}(0)|^2\right] + \frac{dN}{2\psi_1}, \tag{1.5.3}$$

if there exists a positive constant ψ_1 such that for all non-negative s, $0 < \psi_1 \leqslant \bar{\psi}(s)$, that is, if assumption (H3) holds.

Thus, we have the existence of a stationary solution for this particular class of communication rates. □

Remark 9. To the best of our knowledge, we cannot conclude anything about the uniqueness of such stationary solutions.

We now prove the necessary results to obtain the upper bound (1.5.3), as well as other moment controls that will be useful in the last section of this paper.

1.5.2 Various controls of first and second order moments

Lemma 3, proposition 18 and corollary 2 can all be found in [14].

Assume (H1) and (H2). First we truly conclude the proof of Theorem 10 with the crucial result that brings about the inequality (1.5.3).

Proposition 18. *Suppose that (H1) and (H2) are satisfied, and that the initial law has a finite second order moment.*
For all $t \geqslant 0$,

$$\sum_{i=1}^{N} \mathbb{E}[|\hat{v}_i(t)|^2] \leqslant \sum_{i=1}^{N} \mathbb{E}[|\hat{v}_i(0)|^2]\, e^{-2\psi_1 t} + \frac{d(N-1)}{2\psi_1}\,(1-e^{-2\psi_1 t}).$$

Proof. We start with two almost sure inequalities.

Lemma 3. Suppose that (H1) and (H2) are satisfied, and that the initial law has a finite second order moment. Let t be any (stopping) time. Then, almost surely,

a) $|\hat{x}(t)|^2 \leqslant |\hat{x}(0)|^2 + 2\displaystyle\int_0^t \sqrt{|\hat{x}(s)|^2\,|\hat{v}(s)|^2}\,ds.$

b) $|\hat{v}(t)|^2 \leqslant |\hat{v}(0)|^2 - 2\psi_1 \displaystyle\int_0^t |\hat{v}(s)|^2\,ds + d(N-1)\,t + 2\sum_{i=1}^{N}\int_0^t \hat{v}_i(s)\,d\widehat{W}_i(s).$

Proof. a) We apply Itô's formula to the function $t \mapsto |\hat{x}_i(t)|^2$:

$$|\hat{x}_i(t)|^2 = |\hat{x}_i(0)|^2 + 2\int_0^t \sum_{\alpha=1}^{d} \hat{x}_i^{\alpha}(s)\hat{v}_i^{\alpha}(s)\,ds.$$

Then, as $\sum_{i=1}^{N} ||\hat{x}_i(t)|^2 = |\hat{x}(t)|^2$, we have $|\hat{x}(t)|^2 = |\hat{x}(0)|^2 + 2\sum_{i=1}^{N}\int_0^t \hat{x}_i(s)\hat{v}_i(s)\,ds$. Furthermore, by using twice Cauchy-Schwarz inequality, almost surely,

$$\begin{aligned}\sum_{i=1}^{N}\hat{x}_i(s)\,\hat{v}_i(s) &\leqslant \sum_{i=1}^{N}|\hat{x}_i(s)|\,|\hat{v}_i(s)| \\ &\leqslant \sqrt{\sum_{i=1}^{N}|\hat{x}_i(s)|^2}\sqrt{\sum_{i=1}^{N}|\hat{v}_i(s)|^2}.\end{aligned}$$

b) Using once more Itô's formula :

$$\begin{aligned}|\hat{v}_i(t)|^2 &= |\hat{v}_i(0)|^2 + \int_0^t \sum_{j=1}^{N}\frac{2}{N}\,\bar{\psi}(\hat{x}_i(s)-\hat{x}_j(s))\hat{v}_i(s)\,(\hat{v}_j(s)-\hat{v}_i(s))\,ds \\ &\quad + 2\int_0^t \hat{v}_i(s)\,d\widehat{W}_i(s) + \int_0^t \sum_{\alpha=1}^{d}(\sigma\sigma^T)_{\alpha,\alpha}\,ds.\end{aligned}$$

with $\sigma_{\alpha,j} = \left(\delta_{i,j} - \dfrac{1}{N}\right)$ for $j \in \{1,...,N\}$ and $\alpha \in \{1,...,d\}$.

$$\text{Thus, } |\hat{v}(t)|^2 = |\hat{v}(0)|^2 + \frac{2}{N} \sum_{i,j=1}^{N} \int_0^t \bar{\psi}(\hat{x}_i(s) - \hat{x}_j(s)) \, \hat{v}_i(s) \, (\hat{v}_j(s) - \hat{v}_i(s)) \, ds$$
$$+2 \int_0^t \hat{v}_i(s) \, d\widehat{W}_i(s) + d(N-1) \, t.$$

Besides,

$$\begin{aligned} 2 \sum_{i,j=1}^{N} \bar{\psi}(\hat{x}_i - \hat{x}_j) \, \hat{v}_i \, (\hat{v}_j - \hat{v}_i) &= \sum_{i,j=1}^{N} \bar{\psi}(\hat{x}_i - \hat{x}_j) \, \hat{v}_i \, (\hat{v}_j - \hat{v}_i) \\ &\quad + \sum_{i,j=1}^{N} \bar{\psi}(\hat{x}_i - \hat{x}_j) \, \hat{v}_j \, (\hat{v}_i - \hat{v}_j) \\ &= \sum_{i,j=1}^{N} \bar{\psi}(\hat{x}_i - \hat{x}_j) \, (\hat{v}_i - \hat{v}_j) \, (\hat{v}_j - \hat{v}_i) \\ &= - \sum_{i,j=1}^{N} \bar{\psi}(\hat{x}_i - \hat{x}_j) \, |\hat{v}_i - \hat{v}_j|^2, \end{aligned}$$

and $\bar{\psi}(\hat{x}_i - \hat{x}_j) \geqslant \psi_1$.

Finally, as

$$\begin{aligned} \sum_{i,j=1}^{N} |\hat{v}_i - \hat{v}_j|^2 &= \sum_{i,j=1}^{N} \left[|\hat{v}_i|^2 + |\hat{v}_j|^2 - 2 \, \hat{v}_i \hat{v}_j \right] \\ &= 2N|\hat{v}|^2 - 2 \sum_{i=1}^{N} \left(\hat{v}_i \sum_{j=1}^{N} \hat{v}_j \right) = 2N \, |\hat{v}|^2, \end{aligned}$$

we have a.s.

$$2 \sum_{i,j=1}^{N} \bar{\psi}(\hat{x}_i - \hat{x}_j) \, \hat{v}_i \, (\hat{v}_j - \hat{v}_i) \leqslant -2N\psi_1 \, |\hat{v}|^2,$$

which leads to b).

□

We introduce $T_k = \inf \{ \, u \geqslant 0 \mid |\hat{v}(u)|^2 \geqslant k \, \}$. Then, with part b) of lemma 3,

$$\begin{aligned} \mathbb{E}[|\hat{v}(T_k)|^2] &\leqslant \mathbb{E}[|\hat{v}(0)|^2] - 2 \, \psi_1 \, \mathbb{E}\left[\int_0^{T_k} |\hat{v}(s)|^2 \, ds \right] + d(N-1) \, \mathbb{E}[T_k] \\ &\leqslant |\hat{v}(0)|^2 + d(N-1) \, t. \end{aligned}$$

Hence, when k goes to infinity, we obtain the finiteness of $\mathbb{E}[|\hat{v}(t)|^2]$.

Furthermore,

$$\mathbb{E}[|\hat{v}(t)|^2] = \mathbb{E}[|\hat{v}(0)|^2] - \frac{1}{N} \int_0^t \mathbb{E}\left[\sum_{i,j=1}^{N} \bar{\psi}(\hat{x}_i - \hat{x}_j) \, |\hat{v}_i - \hat{v}_j|^2 \right] \, ds + d(N-1) \, t,$$

so that by differentiation,

$$\partial_t \left(\mathbb{E}[|\hat{v}(t)|^2] - \frac{d(N-1)}{2\psi_1} \right) = -\frac{1}{N} \, \mathbb{E}\left[\sum_{i,j=1}^{N} \bar{\psi}(\hat{x}_i - \hat{x}_j) \, |\hat{v}_i - \hat{v}_j|^2 \right] + d(N-1)$$
$$\leqslant d(N-1) - 2\psi_1 \, \mathbb{E}[|\hat{v}(t)|^2]$$

Thus, by Gronwall's lemma,

$$\mathbb{E}[|\hat{v}(t)|^2] \leqslant \mathbb{E}[|\hat{v}(0)|^2] \, e^{-2\psi_1 t} + \frac{d(N-1)}{2\psi_1} \, (1 - e^{-2\psi_1 \, t}).$$

□

As $var(\hat{v}_i^{\alpha}(t)) \leqslant \mathbb{E}[|\hat{v}(t)|^2]$, a uniform bound for the variance can be produced :

Corollary 2. *Suppose that (H1) and (H2) are satisfied, and that the initial law has a finite second order moment.*

$$\forall i \in \{1, ..., N\}, \, \alpha \in \{1, ..., d\}, \, \lim_{t \to \infty} var(\hat{v}_i^{\alpha}(t)) \leqslant \frac{d(N-1)}{2\psi_1}.$$

We now focus on results that will be needed in the next section, dealing with propagation of chaos, adding exchangeability to our assumptions.
We recall (see for instance [3]) that particles are said to be exchangeable if every permutation of these particles has the same law : that is $(X_1, ..., X_n)$ are exchangeable if for any permutation σ of $\{1, ..., n\}$, $(X_1, ..., X_n)$ and $(X_{\sigma(1)}, ..., X_{\sigma(n)})$ have same law.

Proposition 19. *Suppose that (H1) and (H2) are satisfied and that the particles are exchangeable at time $t = 0$; in particular particles have the same initial law. Assume also that this initial law has a finite second order moment.*
For all $i \in \{1, ...N\}$,

$$\sup_{t \geqslant 0} \mathbb{E}\left[|\hat{v}_i(t)|^2 \right] \leqslant \mathbb{E}\left[|\hat{v}_i(0)|^2 \right] + \frac{d}{2\psi_1}.$$

Proof. We have previously seen that :

$$\mathbb{E}\left[\sum_{i=1}^{N} |\hat{v}_i(t)|^2 \right] \leqslant \mathbb{E}\left[\sum_{i=1}^{N} |\hat{v}_i(0)|^2 \right] \, e^{-2\psi_1 \, t} + \frac{d(N-1)}{2\psi_1} \, (1 - e^{-2\psi_1 \, t}).$$

Exchangeability leads to :

$$\mathbb{E}\left[|\hat{v}_i(t)|^2 \right] \leqslant \mathbb{E}\left[|\hat{v}_i(0)|^2 \right] \, e^{-2\psi_1 \, t} + \frac{d(N-1)}{2\psi_1 N} \, (1 - e^{-2\psi_1 \, t}),$$

which brings the conclusion of the proof.

□

Corollary 3. *Suppose that (H1) and (H2) are satisfied and that the particles are exchangeable at time $t = 0$. Assume also that the common initial law has a finite second order moment.*
For all non-negative t, there exists a positive constant M_t, such that

$$\sup_{i \in \{1, ..., N\}} \mathbb{E}\left[|\hat{x}_i(t)), \hat{v}_i(t)| \right] \leqslant M_t.$$

Proof. Let $C_1 = \mathbb{E}\left[|\hat{v}_i(0)|^2\right] + \dfrac{d}{2\psi_1}$. Then,

$$|\hat{x}_i(t)|^2 = |\hat{x}_i(0)|^2 + \left|\int_0^t \hat{v}_i(s)\,ds\right|^2 + 2\,\hat{x}_i(0)\,\int_0^t \hat{v}_i(s)\,ds.$$

Thus, by multiple uses of Cauchy-Schwarz inequality,

$$\mathbb{E}\left[|\hat{x}_i(t)|^2\right] \leqslant \mathbb{E}\left[|\hat{x}_i(0)|^2\right] + C_1\,t^2 + 2\,\sqrt{C_1\,\mathbb{E}\left[|\hat{x}_i(0)|^2\right]}\;t$$

Then, choosing $M_t = \sqrt{C_1 + \mathbb{E}\left[|\hat{x}_i(0)|^2\right] + C_1\,t^2 + 2\,\sqrt{C_1\,\mathbb{E}\left[|\hat{x}_i(0)||^2\right]}\;t}$, by Proposition 19,

$$\mathbb{E}\left[|(\hat{x}_i(t), \hat{v}_i(t))|\right] \leqslant \sqrt{\mathbb{E}\left[|(\hat{x}_i(t), \hat{v}_i(t))|^2\right]} \leqslant M_t.$$

□

We now seek a different kind of moment control, one involving a single particle and a stopping time. This will be useful to apply Aldous criterion in section 1.6.

As we now delve into controls that are not uniform in time, we restrain the trajectories to a finite time interval : we place ourselves on $\Omega_T = \mathcal{C}([0,T], \mathbb{R}^{2d})$, the canonical continuous $\mathbb{R}^{2d}$-valued path space, where T is a fixed positive constant. This proposition will be helpful to prove a tightness result via Aldous lemma.

Proposition 20. *Suppose that (H1) and (H2) are satisfied and that the particles are exchangeable at time $t = 0$. Assume also that the common initial law has a finite second order moment. There exist two constants, C and K, independent of N, such that, for two stopping times τ_1 and τ_2 on Ω_T satisfying $\tau_1 \leqslant \tau_2 \leqslant (\tau_1 + \theta) \wedge T$,*

$$\sup_{i\in\{1,\dots,N\}} \mathbb{E}\left[|(\hat{x}_i(\tau_2) - \hat{x}_i(\tau_1), \hat{v}_i(\tau_2) - \hat{v}_i(\tau_1))|^2\right] \leqslant K\,\theta + C\,\theta^2.$$

Proof. We apply Itô's formula : :

$$\begin{aligned}|\hat{v}_i(\tau_2) - \hat{v}_i(\tau_1)|^2 = M_{\tau_1}^{\tau_2} + \int_{\tau_1}^{\tau_2} d\left(1 - \frac{1}{N}\right)\,du \\ - \frac{2}{N}\int_{\tau_1}^{\tau_2} (\hat{v}_i(u) - \hat{v}_i(\tau_1))\,\sum_{j=1}^{N} \psi(\hat{x}_i(u), \hat{x}_j(u))\,(\hat{v}_i(u) - \hat{v}_j(u))\,du\end{aligned}$$

where $M_\tau^{\tau+u}$ is a martingale and satisfies $\mathbb{E}[M_\tau^{\tau+u}] = 0$ for every u.

This leads to

$$
\begin{aligned}
&\mathbb{E}[|\hat{v}_i(\tau_2) - \hat{v}_i(\tau_1)|^2] \\
&= d\left(1 - \frac{1}{N}\right)\ \mathbb{E}[\tau_2 - \tau_1] - \frac{2}{N}\ \mathbb{E}\left[\int_{\tau_1}^{\tau_2} (\hat{v}_i(u) - \hat{v}_i(\tau_1))\ \sum_{j=1}^{N} \psi(\hat{x}_i(u), \hat{x}_j(u))\ (\hat{v}_i(u) - \hat{v}_j(u))\ du\right] \\
&\qquad \leqslant d\left(1 - \frac{1}{N}\right)\ \theta + \frac{2\psi_2}{N}\ \sum_{j=1}^{N} \left(\mathbb{E}\left[\int_{\tau_1}^{\tau_2} |\hat{v}_i(u) - \hat{v}_i(\tau_1)|\ |\hat{v}_i(u) - \hat{v}_j(u)|\ du\right]\right) \\
&\qquad\qquad \leqslant d\left(1 - \frac{1}{N}\right) \theta + \frac{2\psi_2}{N}\ \sum_{j=1}^{N}\ \int_0^{\theta} \sqrt{\mathbb{E}[|\hat{v}_i(\tau_1 + u) - \hat{v}_i(\tau_1)|^2]} \\
&\qquad\qquad\qquad\qquad \times \sqrt{\mathbb{E}[|\hat{v}_i(\tau_1 + u) - \hat{v}_j(\tau_1 + u)|^2]}\ du,
\end{aligned}
$$

thanks to Cauchy-Schwarz inequality.
According to Lemma 3, for all τ (stopping) time smaller than $T + \theta$,

$$
\mathbb{E}[|\hat{v}(\tau)|^2] \leqslant \mathbb{E}[|\hat{v}(0)|^2] - 2\psi_1\ \mathbb{E}\left[\int_0^{\tau}\ |\hat{v}(s)|^2\ ds\right] + d(N-1)\ \mathbb{E}[\tau] \leqslant \mathbb{E}[|\hat{v}(0)|^2] + 2d(N-1)T
$$

Using the exchangeability, for all i,

$$
\mathbb{E}[|\hat{v}_i(\tau)|^2] \leqslant \mathbb{E}[|\hat{v}_i(0)|^2] + 2d\left(1 - \frac{1}{N}\right) T \leqslant \mathbb{E}[|\hat{v}_i(0)|^2] + 2dT =: C.
$$

It means that

$$
\mathbb{E}[|\hat{v}_i(\tau_2) - \hat{v}_i(\tau_1)|^2] \leqslant d\left(1 - \frac{1}{N}\right)\ \theta + 2\psi_2\ \int_0^{\theta} 4\ C\ du \leqslant K\,\theta,
$$

with

$$
K = d + 8\psi_2\, C.
$$

Besides,

$$
|\hat{x}_i(\tau_2) - \hat{x}_i(\tau_1)|^2 = \left|\int_{\tau_1}^{\tau_2} \hat{v}_i(u)\ du\right|^2 \leqslant \left(\int_{\tau_1}^{\tau_2} |\hat{v}_i(u)|\ du\right)^2 \leqslant (\tau_2 - \tau_1)\ \int_{\tau_1}^{\tau_2} |\hat{v}_i(u)|^2\ du,
$$

using once again Cauchy-Schwarz inequality.

Thus, the proof is concluded, as

$$
\mathbb{E}[|\hat{x}_i(\tau_2) - \hat{x}_i(\tau_1)|^2] \leqslant \theta\ \int_0^{\theta} \mathbb{E}\left[|\hat{v}_i(\tau_1 + u)|^2\right]\ du \leqslant C\ \theta^2.
$$

□

1.6 Propagation of chaos

We now switch our perspective : we look into the behaviour of the system when considering a very large number N of particles. As we are considering mean-field systems, one would expect propagation of chaos properties, as introduced by Sznitman [23] in the late 1980s.

The results presented here have been previously obtained, in a more general case by Bolley, Cañizo and Carrillo in [6]. One should note however, that the two proofs are very different.

In this section we assume that hypotheses (H1), (H2) and (H3), introduced at the beginning of Section 1.5, are satisfied.

Let T be a fixed positive constant. Recall that $\Omega_T = \mathcal{C}([0,T], \mathbb{R}^{2d})$ is the canonical continuous $\mathbb{R}^{2d}$-valued path space, with $\mathcal{F}$ the canonical Borel σ-field on Ω_T.

First, we recall the definition of chaoticity.

Definition 3. We consider E a Polish space, Q a probability measure on E and for $N \in \mathbb{N}$, Q_N a probability measure on E^N. The sequence $(Q_N)_{N\geqslant 1}$ is Q-chaotic if for any fixed integer k and any continuous bounded functions $f_1, ..., f_k$ on E,

$$\lim_{N\to\infty} \int f_1(x_1) \dots f_k(x_k)\, dQ_N(x_1, ..., x_N) = \prod_{i=1}^{k} \int f_i(x_i)\, dQ(x_i).$$

In other words, it means that when N goes towards infinity, any fixed finite number of coordinates become independent with the same distribution Q.

The objective here is on the one hand to show the convergence, in law and in probability, of the empirical measure, in N, associated with the N-particle system (1.1.5) towards a limit η, and one the other hand to prove that we have a chaotic behaviour.

Remember that the system (1.1.5) is, for every $i \in \{1, ...N\}$,

$$\begin{cases} d\hat{x}_i(t) &= \hat{v}_i(t)\, dt \\ d\hat{v}_i(t) &= -\dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \psi(\hat{x}_i(t), \hat{x}_j(t))\, (\hat{v}_i(t) - \hat{v}_j(t))\, dt + d\widehat{W}_i(t) \end{cases}$$

If there is chaoticity, the "natural" limit would be the non linear system :

$$\begin{cases} \mathbf{x}_t &= \mathbf{x}_0 + \displaystyle\int_0^t \mathbf{v}_s\, ds \\ \mathbf{v}_t &= \mathbf{v}_0 + W_t - \displaystyle\int_0^t \int \psi(\mathbf{x}_s, x)(\mathbf{v}_s - v)\, \mathbf{Q}_s(dx, dv)\, ds \\ \mathbf{Q}_t &= \mathcal{L}(\mathbf{x}_t, \mathbf{v}_t) \end{cases} \tag{1.6.1}$$

At this point, we need to introduce a few notations.

Let, for every integer N larger than 1, $\eta_N(\omega) = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \delta_{(\hat{x}_i^N, \hat{v}_i^N)(\omega)}$ be the empirical measure on Ω_T associated with the N-particle system defined by (1.1.5), and π_N its law on $\mathcal{P}(\Omega_T)$.

We then introduce the martingale problems, associated respectively with systems (1.1.5) and (1.6.1) :

1. A probability measure Q^N on $\mathcal{C}([0,T],\mathbb{R}^{2dN})$ is a solution of the martingale problem $(\mathcal{P}_N)$ if for all Φ in $\mathcal{C}_b^2(\mathbb{R}^{2dN})$, $M_t^N(\Phi)$ defined by

$$M_t^N(\Phi) = \Phi(\hat{x}(t),\hat{v}(t)) - \Phi(\hat{x}(0),\hat{v}(0)) - \int_0^t \widehat{L}_N\phi(\hat{x}(s),\hat{v}(s))\,ds \tag{1.6.2}$$

is a Q^N-martingale such that,

$$< M_t^N(\Phi) >= \sum_{i=1}^N \int_0^t |\nabla_{v_i}\Phi(\hat{x}(s),\hat{v}(s)) - \frac{1}{N}\sum_{j=1}^N \nabla_{v_j}\Phi(\hat{x}_s,\hat{v}_s)|^2\,ds$$

where $\widehat{L}_N$ is the infinitesimal generator associated with (1.1.5), that is

$$\widehat{L}_N\Phi(\hat{x},\hat{v}) = \sum_{i=1}^N \hat{v}_i\,\nabla_{\hat{x}_i}\Phi - \frac{1}{N}\sum_{i,j=1}^N \psi(\hat{x}_i,\hat{x}_j)\,(\hat{v}_i-\hat{v}_j)\,\nabla_{\hat{v}_i}\Phi + \frac{1}{2}\sum_{i=1}^N\left(\Delta_{\hat{v}_i}\Phi - \frac{1}{N}\sum_{j=1}^N\sum_{\alpha=1}^d \partial^2_{\hat{v}_i^\alpha\,\hat{v}_j^\alpha}\Phi\right).$$

When $\Phi(\hat{x},\hat{v}) = \phi(\hat{x}_i,\hat{v}_i)$ with ϕ in $\mathcal{C}_b^2(\mathbb{R}^{2d})$, we set $M_t^{N,i}(\phi) := M_t^N(\Phi)$.

2. A probability measure Q on $\Omega_T = \mathcal{C}([0,T],\mathbb{R}^{2d})$ is a solution of the martingale problem $(\mathcal{P}_\infty)$ if for all ϕ in $\mathcal{C}_b^2(\mathbb{R}^{2d})$,

$$\begin{aligned} M_t^\phi(Q) &= \phi(\mathbf{x}_t,\mathbf{v}_t) - \phi(\mathbf{x}_0,\mathbf{v}_0) - \int_0^t \nabla_x\phi(\mathbf{x}_s,\mathbf{v}_s)\,\mathbf{v}_s\,ds \\ &+ \int_0^t\int \psi(\mathbf{x}_s,x)\,\nabla_v\phi(\mathbf{x}_s,\mathbf{v}_s)\,(\mathbf{v}_s - v)\,Q_s(dx,dv)\,ds - \frac{1}{2}\int_0^t \Delta_v\phi(\mathbf{x}_s,\mathbf{v}_s)\,ds, \end{aligned} \tag{1.6.3}$$

where Q_s is defined by $Q_s = Q\circ(\mathbf{x}_s,\mathbf{v}_s)^{-1}$, is a Q-martingale such that

$$< M_t^\phi >= \int_0^t |\nabla_v\phi(\mathbf{x}_s,\mathbf{v}_s)|^2\,ds.$$

The main result is the following theorem :

Theorem 12. *Assume (H1), (H2) and (H3).*
Suppose that the particles are exchangeable at time $t=0$. Assume also that the initial law η_0 on $\mathbb{R}^d\times\mathbb{R}^d$ has a finite second order moment and that for all $a>0$, $\mathbb{E}[e^{a|v_0|}]<\infty$.
The sequence of the empirical measures $(\eta_N)_{N\geqslant 1}$ converges in law and in probability to η, the unique solution of (1.6.3, if $\eta_N(0)$ converges in probability towards η_0 when N goes to infinity

Remark 10. Notice that, while the uniqueness of the solution of (1.6.3) will be established as we prove the theorem, its existence derives from the convergence and will be a consequence of the proof.

From this, to obtain the chaoticity of the system, we need the following proposition, whose proof can be found in [23] or in [18].

Proposition 21. *If $(Q_N)_{N\geqslant 1}$ is a sequence of exchangeable probability measures on E^N, it is Q-chaotic if and only if the associated empirical measure converges in law - and in probability - as $\mathcal{P}(E)$-valued variables under Q_N, towards the probability measure Q.*

And thus, we deduce that the system is indeed chaotic, and that there is propagation of chaos.

Corollary 4. *Under the hypotheses of Theorem 12, the sequence $(\eta_N)_{N\geqslant 1}$ is η-chaotic.*

Remark 11. If we consider a small number of particles among a large amount, they behave independently from each other, which seems quite far from the concept of flocking.

To prove Theorem 12, we will follow a classical procedure and proceed in three steps, in the last three subsections of this work :

1. first, the tightness of $(\pi_N)_{N\geqslant 1}$ in $\mathcal{P}(\mathcal{P}(\Omega_T))$;
2. then, the link between the accumulation points of $(\pi_N)_{N\geqslant 1}$ and a martingale problem ;
3. finally, the uniqueness of the solution of (1.6.3), coming from the uniqueness of the solution of the limit process (1.6.1).

We actually start with the third one.

1.6.1 Uniqueness of the non-linear equation, and of the associated martingale problem

We thank Andrey Pilipenko for the discussion in Potsdam in December 2015 which helped us to solve this uniqueness problem.

Consider the non-linear stochastic differential system, on $[0,T]$, :

$$(\mathcal{S}_W)\left\{\begin{array}{lcl} \mathbf{x}_t & = & \mathbf{x}_0 + \int_0^t \mathbf{v}_s\, ds \\ \mathbf{v}_t & = & \mathbf{v}_0 + W_t - \int_0^t \int \psi(\mathbf{x}_s, x)(\mathbf{v}_s - v)\ \mathbf{Q}_s(dx,dv)\ ds \\ \mathbf{Q}_t & = & \mathcal{L}(\mathbf{x}_t, \mathbf{v}_t) \end{array}\right.$$

Recall in particular that $\psi(x,y) = \bar{\psi}(x-y)$ with $\bar{\psi}$ an even, k-Lipschitz continuous function such that $0 < \bar{\psi}(x) \leqslant \psi_2$ for all $x \in \mathbb{R}^d$.

Theorem 13. *Assume (H1) and (H3).*
For a fixed initial condition (x_0, v_0) *with a finite second order moment and such that for all* $a > 0$, $\mathbb{E}[e^{a|v_0|}] < \infty$, *the non-linear stochastic system* $(\mathcal{S}_W)$ *admits at most one strong solution.*

Corollary 5. *Assume (H1) and (H3).*
For a fixed initial condition (x_0, v_0) *with a finite second order moment and such that for all* $a > 0$, $\mathbb{E}[e^{a|v_0|}] < \infty$, *the martingale problem (1.6.3) associated with* $(\mathcal{S}_W)$ *admits at most one solution.*

The proof that we give here is based on manipulations of expectations and multiple uses of Gronwall lemma.

Let $(W_t)_{t\in[0,T]}$ and $(\widetilde{W}_t)_{t\in[0,T]}$ be two independent, standard $\mathbb{R}^d$-valued Brownian motions. We will sometimes construct W (resp. $(\widetilde{W})$) on the first (resp. second) component of the product space $\Omega_T \times \Omega_T$, and denote, in this subsection alone, by $\mathbb{E}$ (resp. $\widetilde{\mathbb{E}}$) the expectation with respect to the first coordinate (resp. the second coordinate) of this product space.

Reformulation of the problem

System $(\mathcal{S}_W)$ can also be seen as :

$$\left\{\begin{array}{rcl} \mathbf{x}_t &=& \mathbf{x}_0 + \int_0^t \mathbf{v}_s \, ds \\ \mathbf{v}_t &=& \mathbf{v}_0 + W_t - \int_0^t \widetilde{\mathbb{E}}[\psi(\mathbf{x}_s, \widetilde{\mathbf{x}}_s)(\mathbf{v}_s - \widetilde{\mathbf{v}}_s)] \, ds \end{array}\right. \tag{1.6.4}$$

where $(\widetilde{\mathbf{x}}_t, \widetilde{\mathbf{v}}_t)_{t\in[0,T]}$ is an independent copy of $(\mathbf{x}_t, \mathbf{v}_t)_{t\in[0,T]}$ and satisfies system $(\mathcal{S}_{\widetilde{W}})$, that is

$$\left\{\begin{array}{rcl} \widetilde{\mathbf{x}}_t &=& \widetilde{\mathbf{x}}_0 + \int_0^t \widetilde{\mathbf{v}}_s \, ds \\ \widetilde{\mathbf{v}}_t &=& \widetilde{\mathbf{v}}_0 + \widetilde{W}_t - \int_0^t \int \psi(\widetilde{\mathbf{x}}_s, x)(\widetilde{\mathbf{v}}_s - v) \, \mathbf{Q}_s(dx, dv) \, ds \\ \mathbf{Q}_t &=& \mathcal{L}(\widetilde{\mathbf{v}}_t, \widetilde{\mathbf{v}}_t) \end{array}\right.$$

Suppose now that there exist two processes, $(\mathbf{x}, \mathbf{v})$ and $(\mathbf{x}', \mathbf{v}')$, strong solutions of $\mathcal{S}_W$ on Ω_T, with the same initial condition $(\mathbf{x}_0, \mathbf{v}_0)$ and respective laws $\mathbf{Q}$ and $\mathbf{Q}'$. Considering the processes, in $\Omega_T \times \Omega_T$, $((\mathbf{x}, \mathbf{v}), (\widetilde{\mathbf{x}}, \widetilde{\mathbf{v}}))$ and $((\mathbf{x}', \mathbf{v}'), (\widetilde{\mathbf{x}}', \widetilde{\mathbf{v}}'))$ – defined as in equation (1.6.4), we will show that they are almost surely equal, hence the strong uniqueness.

We can write :

$$\mathbf{v}_t = \mathbf{v}_0 + W_t - \int_0^t \widetilde{\mathbb{E}}[\psi(\mathbf{x}_s, \widetilde{\mathbf{x}}_s)(\mathbf{v}_s - \widetilde{\mathbf{v}}_s)] \, ds, \tag{1.6.5}$$

$$\mathbf{v}'_t = \mathbf{v}_0 + W_t - \int_0^t \widetilde{\mathbb{E}}[\psi(\mathbf{x}'_s, \widetilde{\mathbf{x}}'_s)(\mathbf{v}'_s - \widetilde{\mathbf{v}}'_s)]\ ds.$$

Control of the trajectories

First we track an upper bound for $\sup_{t\in[0,T]} |\mathbf{v}_t|$ and $\sup_{t\in[0,T]} |\mathbf{v}'_t|$.

$$\mathbb{E}[|\mathbf{v}_t|] = |\mathbf{v}_0| + \mathbb{E}[|W_t|] + \int_0^t \mathbb{E}\left[\widetilde{\mathbb{E}}[\psi(\mathbf{x}_s, \widetilde{\mathbf{x}}_s)(\mathbf{v}_s - \widetilde{\mathbf{v}}_s)]\right]\ ds \leqslant |\mathbf{v}_0| + \sqrt{\frac{2T}{\pi}} + 2\psi_2 \int_0^t \mathbb{E}[|\mathbf{v}_s|]\ ds.$$

We can then apply Gronwall lemma : for every $t < T$,

$$\mathbb{E}[|\mathbf{v}_t|] \leqslant \left(|\mathbf{v}_0| + \sqrt{\frac{2T}{\pi}}\right) e^{2\psi_2 T}.$$

From there, and keeping in mind that $\mathbb{E}[|\mathbf{v}_t|] = \widetilde{\mathbb{E}}[|\widetilde{\mathbf{v}}_t|]$

$$\begin{aligned} |\mathbf{v}_t| &\leqslant |\mathbf{v}_0| + \psi_2 \int_0^t |\mathbf{v}_s|\ ds + \psi_2 \int_0^t \widetilde{\mathbb{E}}[|\widetilde{\mathbf{v}}_s|]\ ds + \sup_{t\in[0,T]} |W_t| \\ &\leqslant |\mathbf{v}_0| + \psi_2 T \left(|\mathbf{v}_0| + \sqrt{\frac{2T}{\pi}}\right) e^{2\psi_2 T} + \sup_{t\in[0,T]} |W(t)| + \psi_2 \int_0^t |\mathbf{v}_s| ds. \end{aligned}$$

Thus, thanks again to Gronwall lemma,

$$\sup_{t\in[0,T]} |\mathbf{v}_t| \leqslant C_W \tag{1.6.6}$$

where the random variable C_W satisfies

$$C_W = \left(|\mathbf{v}_0| + \psi_2 T \left(|\mathbf{v}_0| + \sqrt{\frac{2T}{\pi}}\right) e^{2\psi_2 T} + \sup_{t\in[0,T]} |W(t)|\right) e^{\psi_2 T}.$$

Besides, as $\mathbf{x}_t = \mathbf{x}_0 + \int_0^t \mathbf{v}_s\ ds$,

$$\sup_{t\in[0,T]} |\mathbf{x}_t| \leqslant |\mathbf{x}_0| + T\ C_W.$$

In a similar way, with

$$C_{\widetilde{W}} = \left(|\mathbf{v}_0| + \psi_2 T \left(|\mathbf{v}_0| + \sqrt{\frac{2T}{\pi}}\right) e^{2\psi_2 T} + \sup_{t\in[0,T]} |\widetilde{W}(t)|\right) e^{\psi_2 T},$$

one has

$$\sup_{t\in[0,T]} |\widetilde{\mathbf{v}}_t| \leqslant C_{\widetilde{W}}. \tag{1.6.7}$$

Note that we also have $\sup_{t\in[0,T]} |\mathbf{v}'_t| \leqslant C_W$ and $\sup_{t\in[0,T]} |\widetilde{\mathbf{v}}'_t| \leqslant C_{\widetilde{W}}$.

Constants and Lipschitz continuity

Let x, x', v and v' be in $\mathbb{R}^d$.
If we suppose that $|v| \leqslant M$, then,

$$|\bar{\psi}(x)\ v - \bar{\psi}(x')\ v'| \leqslant |\bar{\psi}(x) - \bar{\psi}(x')|\ |v| + \bar{\psi}(x')\ |v - v'| \leqslant kM\ |x - x'| + \psi_2 |v - v'|$$
$$\leqslant (kM + \psi_2)\ (|x - x'| + |v - v'|).$$

In particular, using (1.6.6) and (1.6.7), as $\bar{\psi}$ is bounded by ψ_2 and k-Lipschitz continuous,

$$\begin{aligned}
&|\bar{\psi}(\mathbf{x}_s - \widetilde{\mathbf{x}}_s)\ (\mathbf{v}_s - \widetilde{\mathbf{v}}_s) - \bar{\psi}(\mathbf{x}'_s - \widetilde{\mathbf{x}}'_s)\ (\mathbf{v}'_s - \widetilde{\mathbf{v}}'_s)| \\
&\qquad \leqslant (k\ (C_W + C_{\widetilde{W}}) + \psi_2)\ (|\mathbf{x}_s - \widetilde{\mathbf{x}}_s - \mathbf{x}'_s + \widetilde{\mathbf{x}}'_s| + |\mathbf{v}_s - \widetilde{\mathbf{v}}_s - \mathbf{v}'_s + \widetilde{\mathbf{v}}'_s|) \\
&\qquad \leqslant (K_W + K_{\widetilde{W}})\ (|\mathbf{x}_s - \mathbf{x}'_s| + |\mathbf{v}_s - \mathbf{v}'_s|) + (K_W + K_{\widetilde{W}})\ (|\widetilde{\mathbf{x}}_s - \widetilde{\mathbf{x}}'_s| + |\widetilde{\mathbf{v}}_s - \widetilde{\mathbf{v}}'_s|), \quad (1.6.8)
\end{aligned}$$

setting the random variables $K_W = k\ C_W + \dfrac{\psi_2}{2}$ and $K_{\widetilde{W}} = k\ C_{\widetilde{W}} + \dfrac{\psi_2}{2}$.

Computations towards the uniqueness

Using (1.6.5) and (1.6.8),

$$\begin{aligned}
&|\mathbf{v}_t - \mathbf{v}'_t| \leqslant - \int_0^t \widetilde{\mathbb{E}}[|\psi(\mathbf{x}_s, \widetilde{\mathbf{x}}_s)(\mathbf{v}_s - \widetilde{\mathbf{v}}_s) - \psi(\mathbf{x}'_s, \widetilde{\mathbf{x}}'_s)(\mathbf{v}'_s - \widetilde{\mathbf{v}}'_s)]\ ds \\
&\leqslant \int_0^t \widetilde{\mathbb{E}}[(K_W + K_{\widetilde{W}})\ (|\mathbf{x}_s - \mathbf{x}'_s| + |\mathbf{v}_s - \mathbf{v}'_s|)\ ds + \int_0^t \widetilde{\mathbb{E}}[(K_W + K_{\widetilde{W}})\ (|\widetilde{\mathbf{x}}_s - \widetilde{\mathbf{x}}'_s| + |\widetilde{\mathbf{v}}_s - \widetilde{\mathbf{v}}'_s|)\ ds \\
&= \widetilde{\mathbb{E}}[K_W + K_{\widetilde{W}}]\ \int_0^t (|\mathbf{x}_s - \mathbf{x}'_s| + |\mathbf{v}_s - \mathbf{v}'_s|)\ ds + \int_0^t \widetilde{\mathbb{E}}[(K_W + K_{\widetilde{W}})\ (|\widetilde{\mathbf{x}}_s - \widetilde{\mathbf{x}}'_s| + |\widetilde{\mathbf{v}}_s - \widetilde{\mathbf{v}}'_s|)\ ds.
\end{aligned}$$

Thus, setting

$$S_W(s) = \sup_{u \in [0,s]} |\mathbf{x}_u - \mathbf{x}'_u| + \sup_{u \in [0,s]} |\mathbf{v}_u - \mathbf{v}'_u|,$$
$$S_{\widetilde{W}}(s) = \sup_{u \in [0,s]} |\widetilde{\mathbf{x}}_u - \widetilde{\mathbf{x}}'_u| + \sup_{u \in [0,s]} |\widetilde{\mathbf{v}}_u - \widetilde{\mathbf{v}}'_u|,$$

we can affirm that

$$\sup_{u \in [0,t]} |\mathbf{v}_u - \mathbf{v}'_u| \leqslant \widetilde{\mathbb{E}}[K_W + K_{\widetilde{W}}] \int_0^t S_W(s)\ ds\ + \int_0^t \widetilde{\mathbb{E}}[(K_W + K_{\widetilde{W}})\ S_{\widetilde{W}}(s)]\ ds.$$

As $\sup_{u \in [0,t]} |\mathbf{x}_u - \mathbf{x}'_u| \leqslant T \sup_{u \in [0,t]} |\mathbf{v}_u - \mathbf{v}'_u|$, we have

$$S_W(t) \leqslant (1 + T)\ \widetilde{\mathbb{E}}[K_W + K_{\widetilde{W}}] \int_0^t S_W(s)\ ds + (1 + T) \int_0^t \widetilde{\mathbb{E}}[(K_W + K_{\widetilde{W}})\ S_{\widetilde{W}}(s)]\ ds.$$

Applying a generalized version of Gronwall inequality,

$$S_W(t) \leqslant c_W \int_0^t \widetilde{\mathbb{E}}[(K_W + K_{\widetilde{W}})\ S_{\widetilde{W}}(s)]\ ds,$$

with $c_W = (1+T)\, e^{\,T(1+T)\,\widetilde{\mathbb{E}}[K_W + K_{\widetilde{W}}]}$.

In order to bound $t \mapsto \mathbb{E}[(K_W + K_{\widetilde{W}})\ S_W(t)]$, we will apply yet again Gronwall lemma. One can notice that both the following equalities are true :

$$\mathbb{E}[(K_W + K_{\widetilde{W}})\ S_W(t)] = K_{\widetilde{W}}\ \mathbb{E}[S_W(t)] + \mathbb{E}[K_W\ S_W(t)],$$

$$\widetilde{\mathbb{E}}[(K_W + K_{\widetilde{W}})\ S_{\widetilde{W}}(t)] = K_W\ \mathbb{E}[S_W(t)] + \mathbb{E}[K_W\ S_W(t)].$$

From there,

$$\begin{aligned}
\mathbb{E}[(K_W + K_{\widetilde{W}})\ S_W(t)] &\leqslant \mathbb{E}\left[(K_W + K_{\widetilde{W}})\ c_W \int_0^t \widetilde{\mathbb{E}}[(K_W + K_{\widetilde{W}})\ S_{\widetilde{W}}(s)]\ ds\right] \\
&\leqslant \mathbb{E}[c_W(K_W + K_{\widetilde{W}})\ K_W\ ] \int_0^t \mathbb{E}[S_W(s)]\ ds + \mathbb{E}[c_W(K_W + K_{\widetilde{W}})] \int_0^t \mathbb{E}[K_W\ S_W(s)]\ ds \\
&\leqslant \mathbb{E}[c_W(K_W + K_{\widetilde{W}})^2\ ] \int_0^t \mathbb{E}[S_W(s)]\ ds + \mathbb{E}\left[c_W(K_W + K_{\widetilde{W}})\left(1 + \frac{K_W}{K_{\widetilde{W}}}\right)\right] \int_0^t \mathbb{E}[K_W\ S_W(s)]\ ds,
\end{aligned}$$

which leads, finally, to

$$\mathbb{E}[(K_W + K_{\widetilde{W}})\ S_W(t)] \leqslant \mathbb{E}\left[c_W\ (K_W + K_{\widetilde{W}})\left(1 + \frac{K_W}{K_{\widetilde{W}}}\right)\right] \int_0^t \mathbb{E}[(K_W + K_{\widetilde{W}})\ S_W(s)]\ ds.$$

Thus, by Gronwall inequality,

$$\mathbb{E}[(K_W + K_{\widetilde{W}})\ S_W(t)] = 0 \quad a.s.$$

for all $t \in [0,T]$, which implies, as all terms are non-negative, that

$$(K_W + K_{\widetilde{W}})\ S_W(T) = 0 \quad a.s.$$

By definition of S_W, it means that, for all $t \in [0,T]$, $(\mathbf{x}_t, \mathbf{v}_t) = (\mathbf{x}'_t, \mathbf{v}'_t)$ a.s., which ends the proof.

1.6.2 Tightness of the $(\pi_N)_N$

The following lemma can be found in [18], and we will admit it :

Lemma 4. The tightness of $(\pi_N)_{N\geqslant 1}$ in $\mathcal{P}(\mathcal{P}(\Omega_T))$ is equivalent to the tightness of the law of $(\hat{x}_1^N, \hat{v}_1^N)_{N\geqslant 1}$ in $\mathcal{P}(\Omega_T)$.

In order to prove the tightness of the law of $(\hat{x}_1^N, \hat{v}_1^N)$, we will use Aldous criterion.
We start by proving the tightness of the law of $(\hat{x}_1^N(t), \hat{v}_1^N(t))$ for a.e. t. Take $\epsilon > 0$.

$$\mathbb{P}(|(\hat{x}_1^N(t), \hat{v}_1^N(t))| > \alpha) \leqslant \frac{1}{\alpha}\ \mathbb{E}[|(\hat{x}_1^N(t), \hat{v}_1^N(t))|] \leqslant \frac{M_t}{\alpha},$$

according respectively to Markov's inequality and Corollary 3.
Thus, for $\alpha = \dfrac{M_t}{\epsilon}$, $\mathbb{P}((\hat{x}_1^N(t), \hat{v}_1^N(t)) \in \bar{B}(0,\alpha)) > 1 - \epsilon$.

We take $\varepsilon, \eta > 0$. According to Aldous criterion (see [5]) what we need to show is that there exist $\delta > 0$ and an integer N_0 such that

$$\sup_{N \geqslant N_0} \sup_{\substack{\tau_1, \tau_2 \in \mathbf{T} \\ \tau_1 \leqslant \tau_2 \leqslant (\tau_1 + \delta) \wedge T}} \mathbb{P}(|(\hat{x}_1^N(\tau_2) - \hat{x}_1^N(\tau_1), \hat{v}_1^N(\tau_2) - \hat{v}_1^N(\tau_1))| > \varepsilon) \leqslant \eta,$$

where $\mathbf{T}$ is the set of stopping times on Ω_T.
Again thanks to Markov's inequality, and Proposition 20, we have :

$$\begin{aligned} &\mathbb{P}(|(\hat{x}_1^N(\tau_2) - \hat{x}_1^N(\tau_1), \hat{v}_1^N(\tau_2) - \hat{v}_1^N(\tau_1))| > \varepsilon) \\ &\qquad \leqslant \frac{1}{\varepsilon^2} \mathbb{E}[|(\hat{x}_1^N(\tau_2) - \hat{x}_1^N(\tau_1), \hat{v}_1^N(\tau_2) - \hat{v}_1^N(\tau_1))|^2] \leqslant \frac{K\delta + C\delta^2}{\varepsilon^2}. \end{aligned}$$

Thus, δ such that $K\delta + C\delta^2 = \eta\, \varepsilon^2$, which is $\delta = \dfrac{-K + \sqrt{K^2 + 4C\eta\, \varepsilon^2}}{2C}$, provides the solution, and allows us to conclude to the tightness of $(\hat{x}_1^N, \hat{v}_1^N)$, as K and C are independent from N (but depend on T).

1.6.3 The accumulation points of $(\pi_N)_N$

We now know that the sequence $(\pi_N)_{N \geqslant 1}$ is tight ; hence its relative compactness, thanks to Prohorov's theorem.

Let π_∞ be one of its accumulation points ; we still denote by $(\pi_N)_{N \geqslant 1}$ the subsequence that converges towards it. Our goal is to show that under π_∞, for almost every Q in $\mathcal{P}(\Omega_T)$,

$$\mathbb{E}_Q[M_t^\phi(Q) - M_s^\phi(Q) | \mathcal{F}_s] = 0,$$

with M_t^ϕ defined in (1.6.3), this shall mean that Q is a solution of the martingale problem $(\mathcal{P}_\infty)$.

For $q \in \mathbb{N}^*$, $0 \leqslant s_1 < ... < s_q \leqslant s \leqslant t \leqslant T$ and $g_1, ..., g_q \in \mathcal{C}_b(\mathbb{R}^{2d})$, we define

$$F_{s,t}(Q) = \int_{\Omega_T} (M_t^\phi(Q) - M_s^\phi(Q))\, g_1(x_{s_1}, v_{s_1}) \ldots g_q(x_{s_q}, v_{s_q})\, dQ(x, v).$$

.

Lemma 5. For every $q \in \mathbb{N}$, $0 \leqslant s_1 < ... < s_q \leqslant s \leqslant t \leqslant T$ and $g_1, ..., g_q \in \mathcal{C}_b(\mathbb{R}^{2d})$,

$$\int_{\mathcal{P}(\Omega_T)} |F_{s,t}(Q)| \pi_\infty(dQ) = 0.$$

Proof. For the sake of simplicity, we forego here the ever-present exponent N for the $\hat{x}_i$ and the $\hat{v}_i$.
Recall that π_N is the law on $\mathcal{P}(\Omega_T)$ of $\eta_N = \dfrac{1}{N}\displaystyle\sum_{i=1}^N \delta_{(\hat{x}_i, \hat{v}_i)}$, the empirical measure on Ω_T associated

with the N-particle system defined by (1.1.5). It immediately follows that

$$\int F_{s,t}^2(Q)\ \pi_N(dQ) = \mathbb{E}[F_{s,t}(\eta_N)^2].$$

As $F_{s,t}(\eta_N) = \dfrac{1}{N}\sum_{i=1}^{N}(M_t^{N,i}(\phi) - M_s^{N,i}(\phi))g_1(\hat{x}_i(s_1), \hat{v}_i(s_1))...g_q(\hat{x}_i(s_q), \hat{v}_i(s_q))$,

$$\begin{aligned}\int F_{s,t}^2(Q)\pi_N(dQ) = &\frac{1}{N}\ \mathbb{E}\left[(M_t^{N,1}(\phi) - M_s^{N,1}(\phi))^2(g_1(\hat{x}_1(s_1), \hat{v}_1(s_1))...g_q(\hat{x}_1(s_q), \hat{v}_1(s_q)))^2\right] \\ &+ \frac{N(N-1)}{N^2}\ \mathbb{E}\left[(M_t^{N,1}(\phi) - M_s^{N,1}(\phi))(M_t^{N,2}(\phi) - M_s^{N,2}(\phi))\right. \\ &\left.\times g_1(\hat{x}_1(s_1), \hat{v}_1(s_1))...g_q(\hat{x}_1(s_q), \hat{v}_1(s_q))g_1(\hat{x}_2(s_1), \hat{v}_2(s_1))...g_q(\hat{x}_2(s_q), \hat{v}_2(s_q))\right].\end{aligned}$$

The first part goes to zero when N tends towards infinity because $g_1, ..., g_q$ are bounded, and for $t \in [0, T]$, the expectation of $M_t^{N,1}(\phi)^2$ is uniformly bounded in N.

As for the second term,

$$< M^{N,1}(\phi), M^{N,2}(\phi) >= \frac{1}{2}(< M^{N,1}(\phi) + M^{N,2}(\phi) > - < M^{N,1}(\phi) > - < M^{N,2}(\phi) >) = 0.$$

Thus, we have $\lim_{N\to\infty}\int F_{s,t}^2(Q)\pi_N(dQ) = 0$ which implies $\lim_{N\to\infty}\int |F_{s,t}(Q)|\pi_N(dQ) = 0$.

$(\pi_N)_{N\geqslant 1}$ is a sequence of probability measures converging towards π_∞, thus the uniform integrability of $(F_{s,t}(\eta_N))$ (by virtue of being bounded in L^2) allows us to affirm that

$$\int |F_{s,t}(Q)|\pi_\infty(dQ) = 0$$

by inverting limit and integral. □

Then, for every $q \in \mathbb{N}$, $0 \leqslant s_1 < ... < s_q \leqslant s \leqslant t$ and $g_1, ..., g_q \in \mathcal{C}_b(\mathbb{R}^{2d})$, for π_∞-a.e. Q in $\mathcal{P}(\Omega_T)$, $F_{s,t}(Q) = 0$. Using the pathwise continuity, we conclude that for π_∞-a.e. Q, $(M_t^\phi(Q))_{t\geqslant 0}$ is a Q-martingale.

This means that if π_∞ is some limiting point of $(\pi_N)_{N\geqslant 1}$, then every Q in $\mathcal{P}(\Omega_T)$ which is in the support of π_∞ is solution of (1.6.3).
Thanks to corollary 5, we know that there exists a unique probability measure η on Ω_T such that $\pi_\infty = \delta_\eta$; furthermore, π_∞ is entirely determined, and subsequently, unique.

As δ_η is a Dirac measure, this convergence in law implies the convergence in probability. And so, Theorem 12 holds.

Bibliography

[1] M. Agueh, R. Illner, and A Richardson. Analysis and simulations of a refined flocking and swarming model of Cucker-Smale type. *Kinet. Relat. Models*, 48(1):1–16, 2011.

[2] S.M. Ahn and Y. Ha. Stochastic flocking dynamics of the Cucker-Smale model with multiplicative white noises. *J. Math. Phys.*, 51, 2010.

[3] D.J. Aldous. Exhangeability and related topics. *Lectures Notes in Math.*, 1117:1–198, 1983.

[4] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic Markov processes : Lyapunov versus Poincaré. *J. Funct. Anal.*, 254:727–59, 2008.

[5] P. Billingsley. *Convergence of probability measures.* Wiley, 1968.

[6] F. Bolley, J.A. Cañizo, and J.A. Carrillo. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Math. Models Methods Appl. Sci.*, 21(11):2179–210, 2011.

[7] P. Cattiaux, D. Chafaï, and A. Guillin. Central limit theorems for additive functionals of ergodic Markov diffusions processes. *ALEA Lat. Am. J. Probab. Math. Stat.*, 9(2):337–82, 2012.

[8] P. Cattiaux, F. Delebecque, and L. Pédèches. Stochastic Cucker-Smale models : old and new. *Preprint*, 2016.

[9] F. Cucker and J-G. Dong. Avoiding collisions in flocks. *IEEE Trans. Automat. Control ,*, 55(5):1238–43, 2010.

[10] F. Cucker and E. Mordecki. Flocking in noisy environments. *J. Math. Pures Appl.*, 89:278–96, 2008.

[11] F. Cucker and S. Smale. Emergent Behavior in Flocks. *IEEE Trans. Automat. Control*, 52(5):852–61, 2007.

[12] F. Cucker and S. Smale. On the mathematics of emergence. *Jpn. J. Math.*, 2:197–227, 2007.

[13] D. Down, S. Meyn, and R. Tweedie. Exponential and uniform ergodicity of Markov processes. *Ann. Probab.*, 23(4):1671–91, 1995.

[14] S-Y. Ha, K. Lee, and D. Levy. Emergence of time-asymptotic flocking in a stochastic Cucker-Smale system. *Commun. Math. Sci.*, 7(2):453–69, 2009.

[15] S-Y. Ha and J-G. Liu. A simple proof of the Cucker-Smale flocking dynamics and mean-field limit. *Commun Math. Sci.*, 7(2):297–325, 2009.

[16] S-Y. Ha and E. Tadmor. From particle to kinetic and hydrodynamic descriptions of flocking. *Kinet. Relat. Models*, 1(3):415–35, 2008.

[17] K. Itô and M. Nisio. On stationary solutions of a stochastic differential equation. *J. Math. Kyoto Univ.*, 4(1):1–75, 1964.

[18] S. Méléard. Asymptotic behaviour of some interacting particle systems ; McKean-Vlasov and Boltzmann models. *Lectures Notes in Math.*, 1627:42–95, 1995.

[19] S. Meyn and R. Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, 1993.

[20] L. Pédèches. Exponential ergodicity for a class of non-Markovian stochastic processes. *Preprint*, 2016.

[21] M. Scheutzow. Qualitative behaviour of stochastic delay equations with a bounded memory. *Stochastics*, 12:41–80, 1984.

[22] J. Shen. Cucker-Smale flocking under hierarchical leadership. *SIAM J. Appl. Math.*, 68(3):694–719, 2007.

[23] A. Sznitman. Topics in propagation of chaos. *Lectures Notes in Math.*, 1464:167–252, 1989.

[24] T.V. Ton, N.T.H. Linh, and A. Yagi. Flocking and non-flocking behavior in a stochastic Cucker-Smale system. *Anal. Appl. (Singap.)*, 12(1):63–73, 2014.

Chapter 2

Stochastic Cucker-Smale models : old and new.

This paper was a joint work with Fanny Delebecque and Patrick Cattiaux, and was submitted in May 2017.

Abstract : In this paper we revisit and generalize various stochastic models extending the deterministic Cucker-Smale model for self organization. We study flocking and swarming properties. We show how these properties strongly depend on the structure and on the variance of the noise.

Keywords : Cucker-Smale dynamics, stochastic interacting particles, flocking.

Contents

2.1 Introduction, motivations and existing models.

In recent years the observation, the description and the modeling of collective motions deserved a lot of attention and consequently produced a huge literature. These kinds of collective behaviors have been observed for several types of populations: humans, fishes, birds, insects, bacteria, macromolecules, cells ... We refer to the beautiful survey [15] for a nice description of various models introduced during the last fifteen years. Despite its fundamental importance, the validation of such models will be ignored in the present work, where we will focus on mathematical properties. However we shall make some small comments on the structure of the models under study throughout the whole paper, and summarize them (with some additional comments) in the final section.
If we read a lot of interesting papers on the subject, it turns out that we do not always completely understand all the mathematical arguments contained in some of them, in particular those dealing with stochastic models. That is why, instead of pointing out these misunderstandings, we decided to make this paper self-contained, at least for the potential readers a little bit familiar with stochastic calculus.
Finally we shall only look at stochastic models where the noise comes from some Brownian motion (or some continuous Ito process). Of course one should also look at jump processes (P.D.M.P. for instance) or fractional processes whose local behavior could introduce other interesting properties.
Let us come to the subject of this work.

The so-called Cucker-Smale model introduced in [4, 5] is a mean-field kinetic deterministic model that intends to describe self organization of individuals in a population. Originally it is written as

$$\begin{aligned} \frac{d}{dt}\, x_i(t) &= v_i(t) \\ \frac{d}{dt}\, v_i(t) &= -\frac{\lambda}{N} \sum_{j=1}^{N} \psi_{ij}(t)\, (v_i(t) - v_j(t))\,. \end{aligned} \tag{2.1.1}$$

Here the pair $(x_i(t), v_i(t)) \in \mathbb{R}^d \otimes \mathbb{R}^d$ denotes the pair position/velocity of the "particle" $i \in \{1, ..., N\}$ at time t, λ is some positive parameter and $\psi_{ij}(.)$ is for all (i, j) a non-negative function called the *communication rate*.
In the original model

$$\psi_{ij}(t) = \psi(|x_i(t) - x_j(t)|^2) \quad \text{with} \quad \psi(u) = \frac{1}{(1+u)^r} \ \text{ for } r > 0. \tag{2.1.2}$$

The goal was to propose a model for *flocking*. In the deterministic context, flocking means the following. Introduce the center of mass of the system

$$\bar{x}(t) = \frac{1}{N} \sum_{j=1}^{N} x_j(t) \quad , \quad \bar{v}(t) = \frac{1}{N} \sum_{j=1}^{N} v_j(t)\,, \tag{2.1.3}$$

the system (2.1.1) is said to flock if

$$\text{for all } i, \quad \lim_{t\to\infty} |v_i(t) - \bar{v}(t)| = 0 \ \text{ and } \ \sup_{t\geqslant 0} |x_i(t) - \bar{x}(t)| < +\infty \,. \tag{2.1.4}$$

It is known that in the situation of (2.1.2), flocking occurs for all initial conditions (unconditional flocking) provided $r \leqslant \frac{1}{2}$, and for some initial conditions otherwise (see [4, 5, 8, 9]). Of course this is nothing else than convergence to some "equilibrium". Indeed if all initial velocities are the same (hence all equals to $\bar{v}$), they do not evolve in time and the motion of the positions block is simply a translation. This is some equilibrium for the model and flocking is thus some kind of convergence to this equilibrium.

A lot of modified models have then been studied in the deterministic context, including delays, no collisions and many other features. Some of them have introduced some randomness in the model, in various ways. The goal of the present paper is to revisit, extend and study these stochastic Cucker-Smale models.

The first question to ask is: where (and why) does randomness enter the game ?

The first idea is to consider that each individual has a degree of freedom (or craziness) represented by some random noise independent of the behavior of all other individuals in the population. This leads to the following system for the velocities

$$dv_i(t) = -\frac{\lambda}{N}\sum_{j=1}^{N} \psi_{ij}(t)\,(v_i(t) - v_j(t))\,dt \;+\; \sigma_i(t)\,dw_i(t)\,, \tag{2.1.5}$$

where σ_i only depends on (x_i, v_i) and the w_i's are independent $\mathbb{R}^d$ valued *noises.* This kind of model has been studied in [3] for "smooth" noises (actually smooth regularizations of Brownian motions) and in [7] for independent d-dimensional standard Brownian motions w_i and a constant diffusion matrix σ_i (actually $\sigma_i = \sqrt{D}\, Id_d$). The latter case has been revisited and completed by one of us in the recent [12]. Here and in what follows the meaning of dw is the Ito differential (we shall come back later to this).

The second idea is to consider that the dynamics of the velocities is perturbed by a noisy environment. This yields the following model

$$dv_i(t) = -\frac{\lambda}{N}\sum_{j=1}^{N} \psi_{ij}(t)\,(v_i(t) - v_j(t))\,dt \;+\; \sigma_i(t)\,dw(t)\,, \tag{2.1.6}$$

where this time the noise w is the same for all particles.
A very peculiar form of this model is studied in [1]. The authors consider therein a noise $w = (W, W, ..., W)$ i.e. the same Brownian motion in all the directions of $\mathbb{R}^d$ and a diagonal diffusion matrix $\sigma_i(t)$ whose diagonal entries are given by the vector $\sigma(v_i(t))$ where

$$\sigma(v) = D(v - v_e)$$

for some constant state v_e, telling us that the "noise intensity" depends (in a simple way) on the localization of the velocity.

Another idea is to consider that the "infinitesimal" communication rate is perturbed by some noise. This leads to the following model

$$dv_i(t) = -\frac{\lambda}{N} \sum_{j=1}^{N} (v_i(t) - v_j(t)) \; (\psi_{ij}(t)\, dt \, + \, \sigma_{ij}(t)\, dw_{i,j}(t)) \; , \tag{2.1.7}$$

where the w_{ij} are again one dimensional noises.
This is done in [14] with $w_{i,j} = w$ for all i, j and with a constant $\sigma_{i,j} = \sigma$, i.e. for some new constant $\bar{\sigma}$,

$$dv_i(t) = -\frac{\lambda}{N} \sum_{j=1}^{N} \psi_{ij}(t)\, (v_i(t) - v_j(t))\, dt \, + \, \bar{\sigma}\, (v_i(t) - \bar{v}(t))\, dw(t)\, . \tag{2.1.8}$$

Actually the authors replaced the Ito differential by a Stratonovitch differential. This choice is not really natural since it introduces some repulsive modification on the drift due to the Ito-Stratonovitch correction.
In the recent [6] the authors consider instead N independent one dimensional Brownian noises w_i and the following system

$$dv_i(t) = -\frac{\lambda}{N} \sum_{j=1}^{N} \psi_{ij}(t)\, (v_i(t) - v_j(t))\, dt \, + \, \frac{\sigma_i}{N} \sum_{j=1}^{N} \psi_{ij}(t)\, (v_i(t) - v_j(t))\, dw_i(t)\, , \tag{2.1.9}$$

for constant σ_i. Actually these authors also introduce some delay in the coefficients. A similar model to (2.1.9) is also discussed in [13].

One immediately sees an important difference in nature between all these models. In (2.1.8) or (2.1.9), the dynamics $v_i(t) = v_i(0) = \bar{v}(0)$ for all i is still a solution, hence as for the deterministic system we have some "dynamical" equilibrium. Similarly, if we assume $v_i(0) = v_e$ for all i, $v_i(t) = v_e$ furnishes again some dynamical equilibrium for (2.1.6). In the general case of (2.1.5) such a trivial solution does no more exist. This shows that the asymptotic behavior of these stochastic systems may be (and actually is) very different.

The second point is to understand what kind of asymptotic flocking is expected. Indeed since the solutions are random processes, one can look at various behaviors: almost sure behavior, moments behavior, distribution behavior. We will thus introduce three different notions of *stochastic* flocking

Definition 2.1.10. Let $(x_i(t), v_i(t))_{i=1,...,N}$ be a $\mathbb{R}^d \otimes \mathbb{R}^d$ valued stochastic process such that $dx_i(t) = v_i(t)\, dt$ for all $i = 1, ..., N$. Denote by $\bar{v}$ and $\bar{x}$ the centers of masses defined in (2.1.3). We shall say that:

1) The system is almost surely flocking if (2.1.4) holds almost surely.

2) The system is flocking in $\mathbb{L}^{p,q}$ $(p, q \geqslant 1)$ if for all i,

$$\mathbb{E}(|v_i(t) - \bar{v}(t)|^p) \to 0 \quad \text{as} \quad t \to +\infty$$

and

$$\sup_{t\geqslant 0} \mathbb{E}(|x_i(t)-\bar{x}(t)|^q) \;<\; +\infty\,.$$

Actually we will only look at the cases $(p,q)=(1,1),(2,1),(2,2)$. When $q=1$ we simply write $\mathbb{L}^p$ flocking.

3) The system is weakly flocking with rate $\varepsilon(R)$ if for all $R>0$ and all i,

$$\limsup_{t\to+\infty} \mathbb{P}(|v_i(t)-\bar{v}(t)|>R) \;\leqslant\; \varepsilon(R)\,.$$

Remark 2.1.11. Of course quick enough convergence to 0 for the "centered" velocities is enough to ensure boundedness for the "centered" positions.
For instance,

$$\mathbb{E}(\sup_{t\geqslant 0}|x_i(t)-\bar{x}(t)|) \leqslant \mathbb{E}(|x_i(0)-\bar{x}(0)|) + \int_0^{+\infty} \mathbb{E}(|v_i(s)-\bar{v}(s)|)\,ds < +\infty\,.$$

Similarly if for some function η,

$$\int_0^{+\infty} \mathbb{E}(|v_i(t)-\bar{v}(t)|^2)\,\eta(t)\,dt \;<\; +\infty$$

we have

$$\begin{aligned}
\mathbb{E}(\sup_{t\geqslant 0}|x_i(t)-\bar{x}(t)|^2) &\leqslant 2\,\mathbb{E}(|x_i(0)-\bar{x}(0)|^2) + 2\,\mathbb{E}\left[\left(\int_0^{+\infty}|v_i(s)-\bar{v}(s)|\,ds\right)^2\right]\\
&\leqslant 2\,\mathbb{E}(|x_i(0)-\bar{x}(0)|^2) + \\
&\quad + 2\left(\int_0^{+\infty}\mathbb{E}(|v_i(s)-\bar{v}(s)|^2)\,\eta(s)\,ds\right)\left(\int_0^{+\infty}\eta^{-1}(s)\,ds\right)
\end{aligned}$$

so that if in addition $\int_0^{+\infty}\eta^{-1}(s)\,ds<+\infty$, $\mathbb{E}(\sup_{t\geqslant 0}|x_i(t)-\bar{x}(t)|^2)<+\infty$ too. ♢

Results in [1] concern almost sure flocking, results in [14] concern $\mathbb{L}^2$ flocking and those in [12] concern weak flocking.
Another very weak form of stochastic flocking is sometimes discussed: mean-flocking, i.e.

$$\lim_{t\to+\infty}|\mathbb{E}(v_i(t)-\bar{v}(t))|=0 \quad\text{ and }\quad \sup_{t\geqslant 0}|\mathbb{E}(x_i(t)-\bar{x}(t))|<+\infty\,. \tag{2.1.12}$$

Actually it is this type of flocking which is studied in [7, 6]. We confess that we are not really convinced that this kind of property really describes some "collective" behavior, though (2.1.12) can be seen at a first glance as the immediate generalization of deterministic flocking.

In all cases the same strategy of study is used: first look at the motion of the center of mass $\bar{v}(t)$ (the macroscopic level), then look at the fluctuations $\hat{v}_i(t)=v_i(t)-\bar{v}(t)$ (the microscopic level). As in all the previous works we shall assume in the whole paper that

$$\text{for all } i,j\,,\quad \psi_{ij}=\psi_{ji}\,. \tag{2.1.13}$$

Under this assumption,

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \psi_{ij}(t)\,(v_i(t) - v_j(t)) \;=\; 0$$

so that the motion of $\bar{v}(t)$ is only driven by the noise.

In order to understand the difference in nature of all these models we shall first look at the simplest case i.e. with a constant communication rate and a constant diffusion coefficient. This is done in the next section 2.2. In the following section we introduce the notion of *swarming* and look at its connection with flocking, as it is the case in the deterministic situation. In the two following sections, we still look at constant communication rates but with more general diffusion coefficients for (2.1.6) and (2.1.7). This will be the opportunity to introduce the methods that will be mainly used in the general case. In addition, as we shall see in section 2.6, many results for a non constant communication rate can be deduced from the ones obtained in the constant case. Up to section 2.6 what is obtained is "unconditional" flocking, that is, without restriction on the initial condition.
Section 2.6 studies the case of non constant communication rate for the latter two models (2.1.6) and (2.1.7). Actually if the communication rate is bounded from below, one can reduce the study to the one with constant communication rate. If it is not bounded from below, we prove some "conditional" flocking results, that is we extend for the first time the corresponding deterministic results to the stochastic situation. The final section deals with comments and simulations.

In order to keep the paper into a reasonable size, we will not discuss here other models of Cucker-Smale type, introducing a mean field term depending on the positions too, or a local mean field dependence as in [11]. This will be the aim of future work(s). However, some aspects are already contained in [12] for the model (2.1.5).

For the sake of simplicity we will assume throughout the paper that the initial conditions $(v(0), x(0))$ are deterministic. All the results can be extended to random initial conditions such that $v(0) - \bar{v}(0)$ and $x(0) - \bar{x}(0)$ are almost surely bounded. We shall also denote by $|y|$ the euclidean norm of a vector $y \in \mathbb{R}^m$ whatever m is.

2.2 Constant communication rate. A new visit of the existing models.

In this section we assume that, for all t,

$$\psi_{ij}(t) = \psi_{ji}(t) = \psi \;>\; 0\,. \tag{2.2.1}$$

Notice that in this situation, under mild assumptions on the diffusion coefficients (ensuring that the stochastic integral is a true martingale) the expectations $(\mathbb{E}(v_i(t)), \mathbb{E}(x_i(t)))$ satisfy (2.1.1) with a constant communication rate, so that one always has mean-flocking.

First we will revisit (and extend) the known results we recalled in the introduction, hence we assume that

(H1) in (2.1.5) we consider $\sigma_i(t) = \sqrt{D}\, Id_d$, (H2) in (2.1.6) as in [1] we consider $\sigma_i(t) = D\,(v_i(t) - v_e)$, but here we assume that w is a d-dimensional process $w = (w^1, ..., w^d)$ such that each w^k is a standard linear Brownian motion (we do not make any assumption on the correlations),

(H3) in (2.1.8) the same assumption for w is made as in (H2),

(H4) in (2.1.9) the same assumption is made for each $w_i = (w_i^1, ..., w_i^d)$ (the w_i's being independent) and in addition $\sigma_i = \sigma$ for all i.

We will prove the following

Theorem 2.2.2. Consider the previous models assuming (2.2.1). Then,

1. If (H1) is satisfied, the system (2.1.5) is weakly flocking with a rate $\varepsilon(R)$ given by some χ^2 tail.

2. If (H2) is satisfied, the system (2.1.6) is always almost surely flocking and $\mathbb{L}^1$ flocking, but is $\mathbb{L}^2$ flocking if and only if $2\lambda\psi > D^2$ (or $v_i(0) = \bar{v}(0)$ for all i). In this case it is also $\mathbb{L}^{2,2}$ flocking.
   In addition $\bar{v}(t)$ goes almost surely to v_e as t goes to infinity and $\bar{x}(t) - tv_e$ is almost surely bounded.

3. If (H3) is satisfied, the system (2.1.8) is always almost surely flocking and $\mathbb{L}^1$ flocking, but is $\mathbb{L}^2$ flocking if and only if $2\lambda\psi > \bar{\sigma}^2$ (or $v_i(0) = \bar{v}(0)$ for all i). In this case it is also $\mathbb{L}^{2,2}$ flocking.
   In addition $\bar{v}$ is constant hence $\bar{x}(t)$ is linear in t.

4. Assume that the system (2.1.9) is not at equilibrium i.e. does not satisfy $v_i(0) = \bar{v}(0)$ for all i. If (H4) is satisfied for (2.1.9), we have the following situation: define

$$\alpha = (1 - 1/N)(\sigma\psi)^2 - 2\lambda\psi\,,$$

   then

   (a) if $\alpha < 0$ the system is almost surely and $\mathbb{L}^{2,2}$ flocking. In addition the center of mass $\bar{v}(t)$ converges almost surely and in $\mathbb{L}^1$ to some given random variable, while $\bar{x}(t)$ has some asymptotic linear behavior.

   (b) If $0 \leqslant \alpha$ the system is not $\mathbb{L}^2$ flocking, moreover when $\alpha > 0$, the $\mathbb{L}^2$ norm of all the $\hat{v}_i(t)$ are going to infinity,

   (c) if $\dfrac{2(\sigma\,\psi)^2}{N} > \alpha \geqslant 0$ the system is almost surely flocking (but not $\mathbb{L}^2$). This condition implies $N \leqslant 2$, which is not really interesting.

Remark 2.2.3. The previous Theorem clearly shows the importance of defining the type of stochastic flocking one wants to get, since on the same elementary model one can have one flocking property and not another one. It also seems that $\mathbb{L}^2$ flocking is more demanding. $\diamondsuit$

Proof. In the first three cases one can find an explicit solution for the involved stochastic differential equations using that

$$\frac{1}{N}\sum_{j=1}^{N}(v_i - v_j) = v_i - \bar{v}\,.$$

Let start with (2.1.8) assuming (H3). It can be rewritten for all $i = 1, ..., N$ and all $k = 1, ..., d$,

$$dv_i^k(t) = -\,\lambda\,\psi\,(v_i^k(t) - \bar{v}^k(t))\,dt \,+\, \bar{\sigma}\,(v_i^k(t) - \bar{v}^k(t))\,dw^k(t)\,. \tag{2.2.4}$$

In particular $d\bar{v}^k(t) = 0$ so that $\bar{v}^k(t) = \bar{v}^k(0) = v_e^k$ and (2.2.4) becomes a particular case of (2.1.6) with $v_e = \bar{v}(0)$. This yields the following explicit solution

$$v_i^k(t) = \bar{v}^k(0) \,+\, (v_i^k(0) - \bar{v}^k(0))\, e^{\bar{\sigma}\,w_t^k \,-\, (\frac{1}{2}\,\bar{\sigma}^2 + \lambda\,\psi)\,t}\,. \tag{2.2.5}$$

Since

$$\frac{w_t^k}{t} \;\to\; 0 \text{ almost surely as } t \to +\infty,$$

there is almost sure convergence to the constant center of mass for the velocities. But if $B_.$ is a linear standard Brownian motion,

$$\int_0^{+\infty} e^{aB_t - bt}\,dt \text{ is almost surely bounded for any } a \in \mathbb{R} \text{ and } b > 0 \tag{2.2.6}$$

thanks to the previous remark on the asymptotic behavior of B_t/t. Thus we have shown almost sure flocking for the model (2.2.4). Notice that the center of mass of the positions is here simply given by $\bar{x}(t) = \bar{x}(0) + t\bar{v}(0)$.

In addition, on one hand

$$\begin{aligned}\mathbb{E}(|v_i^k(t) - \bar{v}_i^k(t)|) &= |v_i^k(0) - \bar{v}^k(0)|\,\mathbb{E}(e^{\bar{\sigma}\,w_t^k \,-\, \frac{1}{2}\bar{\sigma}^2\,t})\,e^{-2\lambda\,\psi\,t} \\ &= |v_i^k(0) - \bar{v}^k(0)|\,e^{-2\lambda\,\psi\,t}\end{aligned}$$

while

$$\begin{aligned}\mathbb{E}((v_i^k(t) - \bar{v}_i^k(t))^2) &= (v_i^k(0) - \bar{v}^k(0))^2\,\mathbb{E}(e^{2\bar{\sigma}\,w_t^k \,-\, (\bar{\sigma}^2 + 2\lambda\,\psi)\,t}) \\ &= (v_i^k(0) - \bar{v}^k(0))^2\,e^{(-2\lambda\psi + \bar{\sigma}^2)t}\,\mathbb{E}(e^{2\bar{\sigma}\,w_t^k \,-\, 2\bar{\sigma}^2\,t}) \\ &= (v_i^k(0) - \bar{v}^k(0))^2\,e^{(-2\lambda\psi + \bar{\sigma}^2)t}\,.\end{aligned}$$

Hence if $-2\lambda\psi + \bar{\sigma}^2 \geqslant 0$ there is no $\mathbb{L}^2$ flocking. For the positions we may use the Remark (2.1.11) to get $\mathbb{L}^1$ flocking. For $\mathbb{L}^{2,2}$ flocking, assuming $-2\lambda\psi + \bar{\sigma}^2 < 0$, since

$$e^{\theta t}\,\mathbb{E}(|\hat{v}_i^k(t)|^2) \to 0 \qquad \text{as } t \to +\infty$$

for some $\theta > 0$, we also have,

$$\int_0^{+\infty} e^{\theta t/2}\,\mathbb{E}(|\hat{v}_i^k(t)|^2)\,dt \,<\, +\infty\,,$$

so that we are again in the situation of the remark.

Remark 2.2.7. This result differs from [14] since almost sure flocking occurs in all cases while small noise is required in [14] (that actually does not really study almost sure flocking). This is only due to the fact that, as we said before, the Ito-Stratonovitch correction introduces some repulsive part in the drift in [14]. $\diamondsuit$

(2.1.6) assuming (H2) is thus a little bit more general if $\bar{v}(0) \neq v_e$. In this case

$$\bar{v}^k(t) = v_e^k + (\bar{v}^k(0) - v_e^k)\, e^{Dw_t^k - \frac{D^2 t}{2}} \tag{2.2.8}$$

converges almost surely to v_e^k, and at the microscopic level $\hat{v}_i^k(t) = v_i^k(t) - \bar{v}^k(t)$ satisfies

$$d\hat{v}_i^k(t) = -\lambda\, \psi\, \hat{v}_i^k(t)\, dt \,+\, D\, \hat{v}_i^k(t)\, dw^k(t)$$

so that

$$\hat{v}_i^k(t) = \hat{v}_i^k(0)\, e^{D\, w_t^k - (\frac{1}{2}\, D^2 + \lambda\, \psi)\, t}$$

and we get almost sure flocking as before. This time the center of mass $\bar{v}(t)$ goes almost surely to v_e as t goes to infinity and $\bar{x}(t) - t\, v_e$ is almost surely bounded.

The key point here is that, summing up the equations over i, we obtain an autonomous S.D.E. for the motion of $\bar{v}$.

Actually the same occurs under (H1) in (2.1.5). Thanks to the independence of the w_i's, $\bar{v}$ is simply a Brownian motion with covariance matrix $\dfrac{D}{N}\, Id_d$. We can then get an explicit solution for the motion of $\hat{v}$ which becomes some degenerate dN-dimensional Ornstein-Uhlenbeck process (see [12] section 1),

$$d\hat{v}_i(t) = -\lambda \hat{v}_i(t)\, dt \,+\, \sqrt{D}\, \left(1 - \frac{1}{N}\right)\, dw_i(t) \,-\, \frac{\sqrt{D}}{N}\, \sum_{j \neq i} dw_j(t)\,, \tag{2.2.9}$$

degenerate means that since $\sum_i \hat{v}_i = 0$ the process is an O-U process on this subspace. It is then easy to show that $\hat{v}$ is ergodic with a (degenerate but explicit) gaussian invariant distribution so that it is weakly flocking with a rate $\varepsilon(R)$ corresponding to some χ^2 tail. However using a Central Limit Theorem one can see that $\hat{x}(t)$ behaves like $\sqrt{t}$ times a gaussian vector (in distribution) so that the Probability for $\hat{x}(t)$ to belong to some bounded set goes to 0 as $t \to +\infty$ for all bounded sets, that is, weak flocking really only concerns the velocities. We refer to [12] for the details and the explicit computations.

Finally let us look at (2.1.9) assuming (H4). We first get

$$d\bar{v}(t) \;=\; \frac{\sigma\, \psi}{N}\, \left(\sum_{i=1}^{N} v_i(t)\, dw_i(t) \,-\, \bar{v}(t)\, \sum_{i=1}^{N} dw_i(t)\right) \tag{2.2.10}$$

and then

$$d\hat{v}_i^k(t) \;=\; -\lambda\,\psi\,\hat{v}_i^k(t)\,dt \;+\; \sigma\,\psi\,\left(1-\frac{1}{N}\right)\,\hat{v}_i^k(t)\,dw_i(t) \;-\; \frac{\sigma\,\psi}{N}\,\sum_{j\neq i}\,\hat{v}_j^k(t)\,dw_j(t)\,. \qquad (2.2.11)$$

Of course since the coefficients are global Lipschitz, (2.2.11) admits a unique strong solution. Thus using Ito's formula we get,

$$\begin{aligned} z^k(t) := \sum_{i=1}^N (\hat{v}_i^k)^2(t) \;&=\; z^k(0) \,-\, 2\lambda\psi \int_0^t z^k(s)\,ds \,+\, 2\sigma\psi\,\sum_{i=1}^N \left(\int_0^t (\hat{v}_i^k)^2(s)\,dw_i(s)\right) \\ &\quad - \frac{2\sigma\psi}{N}\,\sum_{i,j=1}^N \left(\int_0^t (\hat{v}_i^k\,\hat{v}_j^k)(s)\,dw_j(s)\right) + (\sigma\psi)^2\,\left(1-\frac{1}{N}\right)^2\,\int_0^t z^k(s)\,ds \\ &\quad + \frac{(\sigma\psi)^2}{N^2}\,\sum_{i\neq j=1}^N \int_0^t (\hat{v}_j^k)^2(s)\,ds \qquad (2.2.12) \\ &=\; z^k(0) + \left(-2\lambda\psi + (\sigma\psi)^2\,\left(1-\frac{1}{N}\right)\right)\,\int_0^t z^k(s)\,ds \\ &\quad + 2\sigma\psi\,\sum_{i=1}^N \left(\int_0^t (\hat{v}_i^k)^2(s)\,dw_i(s)\right)\,, \end{aligned}$$

since $\sum_j \hat{v}_j^k(s) = 0$.

It follows

$$u^k(t) := \mathbb{E}(z^k(t)) \;=\; u^k(0) + \left(-2\lambda\psi + (\sigma\psi)^2\,\left(1-\frac{1}{N}\right)\right)\,\int_0^t u^k(s)\,ds\,. \qquad (2.2.13)$$

A rigorous proof of (2.2.13) is straightforward: it is enough to stop the process at the exit time of open balls of radius R (to be sure that the stochastic integrals are true martingales), to take the expectation and then to use the monotone convergence theorem for letting R go to infinity. (2.2.13) is exactly solved by

$$u^k(t) = u^k(0)\,e^{\alpha t} \qquad \text{where} \qquad \alpha = (1-1/N)(\sigma\psi)^2 - 2\lambda\psi\,. \qquad (2.2.14)$$

We thus have to distinguish three cases: when $\alpha > 0$ $u^k(t)$ grows to infinity and there is no $\mathbb{L}^2$ flocking, when $\alpha < 0$ we may have $\mathbb{L}^2$ flocking, when $\alpha = 0$ there is no $\mathbb{L}^2$ flocking.

We can be more precise. First we have (with an obvious new notation)

$$\begin{aligned} u_i^k(t) \;&=\; u_i^k(0) + \left(-2\lambda\psi + (\sigma\psi)^2\,\left(1-\frac{1}{N}\right)^2\right)\,\int_0^t u_i^k(s)\,ds \,+\, \frac{(\sigma\psi)^2}{N^2}\,\sum_{j\neq i}\int_0^t u_j^k(s)\,ds \\ &=\; u_i^k(0) + \left(-2\lambda\psi + (\sigma\psi)^2\,\left(1-\frac{2}{N}\right)\right)\,\int_0^t u_i^k(s)\,ds \,+\, \frac{(\sigma\psi)^2}{N^2}\,\int_0^t u^k(s)\,ds \\ &=\; u_i^k(0) + \left(\alpha - \frac{(\sigma\psi)^2}{N}\right)\,\int_0^t u_i^k(s)\,ds \,+\, \frac{(\sigma\psi)^2\,u^k(0)}{\alpha\,N^2}\,(e^{\alpha t}-1)\,, \end{aligned}$$

so that it is easily seen (by contradiction for instance) that when $\alpha > 0$, $u_i^k(.)$ cannot be bounded by some $C\,e^{\beta t}$ for $\beta < \alpha$. Hence all the $u_i^k(t)$ are growing to infinity at an exponential rate.

Of course for $\alpha < 0$, $\mathbb{E}((\hat{v}_i^k)^2(s)))$ decays exponentially fast; hence we get $\mathbb{L}^{2,2}$ flocking as before.

Notice in this case that

$$\mathbb{E}(|\bar{v}^k(t)-\bar{v}^k(0)|^2) = \frac{(\sigma\psi)^2}{N^2}\int_0^t u^k(s)\,ds \;=\; \frac{(\sigma\psi)^2}{N^2}\,\frac{u^k(0)}{|\alpha|}\,(1-e^{\alpha t})$$

is bounded. According to (2.2.10), $(\bar{v}^k(t)-\bar{v}^k(0))_{t\geqslant 0}$ is thus a martingale which is bounded in $\mathbb{L}^2$. According to Doob's convergence of martingale theorem we know that there exists a random variable a^k such that

$$(\bar{v}^k(t)-\bar{v}^k(0)) \to a^k \quad \text{ a.s. as } \quad t\to+\infty\,.$$

Since the convergence also holds in $\mathbb{L}^1$ we get in addition that $\bar{x}(t)-\bar{x}(0)-t(\bar{v}(0)+a)$ is bounded in $\mathbb{L}^1$.

What can be said about the almost sure behavior ? Using Ito formula we get that for all $t<T_0$, where T_0 is the hitting time of 0 for $z^k(.)$ (notice that for $t\geqslant T_0$ one has $z_t^k=0$ almost surely),

$$\ln(z_t^k) = \ln(z^k(0)) + \alpha t + 2\sigma\psi \sum_{i=1}^N \int_0^t \frac{(\hat{v}_i^k)^2(s)}{z_s^k}\,dw_i(s) - 2(\sigma\,\psi)^2 \sum_{i=1}^N \int_0^t \frac{(\hat{v}_i^k)^4(s)}{(z_s^k)^2}\,ds\,. \quad (2.2.15)$$

Since $\sum_i \beta_i^4 \leqslant (\sum_i \beta_i^2)^2$, the martingale term

$$M^k(t) = \sum_{i=1}^N \int_0^t \frac{(\hat{v}_i^k)^2(s)}{z_s^k}\,dw_i(s)$$

whose bracket is given by

$$\langle M^k\rangle(t) = \int_0^t \frac{(\sum_{i=1}^N \hat{v}_i^k)^4(s)}{(\sum_{i=1}^N (\hat{v}_i^k)^2(s))^2}\,ds \;\leqslant\; t$$

satisfies the two following properties

1. $t^{-1/2}\,M_t^k$ is bounded in $\mathbb{L}^2$ for $t\in[1,+\infty[$,
2. $t^{-1}\,M_t^k \to 0$ almost surely as $t\to+\infty$.

The second point is the standard law of large numbers for martingales.

Using in addition that $\sum_{i=1}^N \beta_i^4 \geqslant \frac{1}{N}(\sum_{i=1}^N \beta_i^2)^2$, we immediately deduce that $z^k(t)$ converges almost surely to 0, hence that we have almost sure flocking, as soon as $\alpha < \frac{2(\sigma\,\psi)^2}{N}$ and that $z^k(t)$ goes to infinity (hence no almost sure flocking) if $\alpha > 2(\sigma\,\psi)^2$. But the latter cannot occur due to the value of α. □

2.3 Some general properties.

In this section we introduce some general properties (holding true for any of the model we are considering) that we will use in the sequel.

We start with some simple algebraic remarks:

$$\sum_{1\leqslant i,j\leqslant N}|v_i-v_j|^2 = 2\sum_{1\leqslant i,j\leqslant N}\langle v_i, v_i-v_j\rangle = 2N\sum_{i=1}^N |v_i|^2 - 2N^2|\bar v|^2 \;=\; 2N\sum_{i=1}^N |v_i-\bar v|^2\,, \quad (2.3.1)$$

$$\sum_{1\leqslant i,j\leqslant N}|v_i-v_j|^2 = \sum_{1\leqslant i,j\leqslant N}|\hat v_i-\hat v_j|^2 = 2N\sum_{i=1}^N |\hat v_i|^2\,. \quad (2.3.2)$$

and similarly, if $\psi_{ij}=\psi_{ji}$,

$$\sum_{1\leqslant i,j\leqslant N}\psi_{ij}\langle v_i, v_i-v_j\rangle = \frac{1}{2}\sum_{1\leqslant i,j\leqslant N}\psi_{ij}|v_i-v_j|^2\,, \quad (2.3.3)$$

and more generally

$$\sum_{1\leqslant i,j\leqslant N}\psi_{ij}\langle u_i, v_i-v_j\rangle = \frac{1}{2}\sum_{1\leqslant i,j\leqslant N}\psi_{ij}\langle u_i-u_j, v_i-v_j\rangle\,. \quad (2.3.4)$$

The final (2.3.4) will allow us, to control in some cases, flocking by a weaker notion called *swarming* we will define now.

Definition 2.3.5. Let $(x_i(t), v_i(t))_{i=1,\ldots,N}$ be a $\mathbb{R}^d\otimes\mathbb{R}^d$ valued stochastic process such that $dx_i(t) = v_i(t)\,dt$ for all $i=1,...,N$. Denote by $\bar v$ and $\bar x$ the centers of masses defined in (2.1.3). We shall say that:

1) The system is almost surely (resp. $\mathbb{L}^p$) weakly swarming if

$$\text{for all } i, \quad \sup_{t\geqslant 0}|v_i(t)-\bar v(t)| < +\infty \;\text{ almost surely} \quad (2.3.6)$$

respectively

$$\text{for all } i, \quad \sup_{t\geqslant 0}\mathbb{E}(|v_i(t)-\bar v(t)|^p) < +\infty\,. \quad (2.3.7)$$

2) The system is almost surely (resp. $\mathbb{L}^p$, resp. $\mathbb{L}^{p,q}$) strongly swarming if in addition, for all i,

$$\sup_{t\geqslant 0}|x_i(t)-\bar x(t)| < +\infty$$

almost surely, respectively

$$\sup_{t\geqslant 0}\mathbb{E}(|x_i(t)-\bar x(t)|) < +\infty\,,$$

respectively

$$\sup_{t\geqslant 0}\mathbb{E}(|x_i(t)-\bar x(t)|^q) < +\infty\,.$$

When $\mathbb{E}(\sup_{t\geqslant 0}|x_i(t)-\bar x(t)|^q) < +\infty$ we shall say that the swarming property is uniform (in time).

In some situations, proving swarming is enough to get flocking. Indeed, assume that

$$\psi_{ij}(v,x) = \psi(|x_i-x_j|^2) \;\text{ and define }\; \Psi(b) = \int_0^b \psi(a)\,da\,. \quad (2.3.8)$$

We thus have

$$\Psi(|x_i(t)-x_j(t)|^2) - \Psi(|x_i(0)-x_j(0)|^2) \;=\; 2\int_0^t \psi_{ij}(s)\,\langle x_i(s)-x_j(s), v_i(s)-v_j(s)\rangle\, ds\,.$$

Hence, denoting $x_{ij} = x_i - x_j$ and $v_{ij} = v_i - v_j$;

$$\begin{aligned}\langle x_{ij}(t), v_{ij}(t)\rangle \;&=\; \langle x_{ij}(0), v_{ij}(0)\rangle - \frac{\lambda}{N}\int_0^t \sum_{l=1}^N \psi_{il}(s)\,\langle x_{ij}(s), v_{il}(s)\rangle\, ds \\ &\quad + \frac{\lambda}{N}\int_0^t \sum_{l=1}^N \psi_{jl}(s)\,\langle x_{ij}(s), v_{jl}(s)\rangle\, ds \;+\; \int_0^t |v_i(s)-v_j(s)|^2 ds + M_{ij}(t)\end{aligned}$$

where $M_{ij}(.)$ is a local martingale term. Let sum up in i, j. The following term appears

$$A = -\sum_{i,j,l} \psi_{il}\langle x_i - x_j, v_i - v_l\rangle + \sum_{i,j,l} \psi_{jl}\langle x_i - x_j, v_j - v_l\rangle\,.$$

Let us calculate A, first exchanging the role of i and j in the second term,

$$\begin{aligned}A \;&=\; -2\sum_{i,j,l} \psi_{il}\langle x_i - x_j, v_i - v_l\rangle \\ &=\; -2N\sum_{i,l} \psi_{il}\langle x_i, v_i - v_l\rangle \;+\; 2\sum_j \left\langle x_j, \sum_{i,l} \psi_{il}\,(v_i - v_l)\right\rangle \\ &=\; -2N\sum_{i,l} \psi_{il}\langle x_i, v_i - v_l\rangle \;=\; -N\sum_{i,l} \psi_{il}\langle x_i - x_l, v_i - v_l\rangle\end{aligned}$$

thanks to (2.3.4) and since $\sum_{i,l} \psi_{il}\,(v_i - v_l) = 0$.
As usual using some exhausting sequence of stopping times (if it exists) we may integrate up to these random times, for which we get true martingales, take the expectation and then pass to the limit. So we may assume that we have true martingales if we can check that the brackets of the M_{ij} have finite expectation. We shall come back to this point later.
Hence we sum up over all indices and take the expectation, in order to get

$$\begin{aligned}\sum_{i,j=1}^N \mathbb{E}(\langle x_{ij}(t), v_{ij}(t)\rangle) \;&=\; \sum_{i,j=1}^N \mathbb{E}(\langle x_{ij}(0), v_{ij}(0)\rangle) + \int_0^t \mathbb{E}(\sum_{i,j=1}^N |v_i(s)-v_j(s)|^2)\, ds \\ &\quad + \lambda\int_0^t \sum_{i,j=1}^N \mathbb{E}(\psi_{ij}(s)\,\langle x_i(s)-x_j(s), v_i(s)-v_j(s)\rangle)\, ds\end{aligned}$$

and finally

$$
\begin{aligned}
\int_0^t \mathbb{E}(\sum_{i,j=1}^N |v_i(s)-v_j(s)|^2)\, ds &= \sum_{i,j=1}^N (\mathbb{E}(\langle x_{ij}(t), v_{ij}(t)\rangle - \langle x_{ij}(0), v_{ij}(0)\rangle)) \\
&\quad - \frac{\lambda}{2}\sum_{i,j=1}^N \mathbb{E}(\Psi(|x_i(t)-x_j(t)|^2) - \Psi(|x_i(0)-x_j(0)|^2)) \\
&\leqslant 2N^2 \max_{i,j}\left((\sup_{s\geqslant 0} \mathbb{E}(|x_i(s)-x_j(s)|^2))^{\frac{1}{2}} (\sup_{s\geqslant 0} \mathbb{E}(|v_i(s)-v_j(s)|^2))^{\frac{1}{2}}\right) \\
&\quad - \frac{\lambda}{2}\sum_{i,j=1}^N \mathbb{E}(\Psi(|x_i(t)-x_j(t)|^2) - \Psi(|x_i(0)-x_j(0)|^2))\,.
\end{aligned}
\tag{2.3.9}
$$

We shall thus use the following elementary Lemma

Lemma 2.3.10. Let $h : \mathbb{R} \to \mathbb{R}^+$ be a C^1 function with a bounded derivative.
If $\int_0^{+\infty} h(s)\, ds < +\infty$, then $h(t) \to 0$ as $t \to +\infty$.

We can thus easily deduce that $\mathbb{L}^{2,2}$ swarming implies $\mathbb{L}^{2,2}$ flocking, as soon as Ψ is at most linear, in order to control the second term in the previous sum. Let us state a general result that will be completed in the situations we are looking at later

Lemma 2.3.11. Consider any of our models. Assume that (2.3.8) is fulfilled for some bounded function ψ. Assume in addition that,

(1) a unique solution $(v(.), x(.))$ exists and is such that for all i, j, $\langle v_i(.) - v_j(.), x_i(.) - x_j(.)\rangle$ is a $\mathbb{L}^2$ semi martingale,

(2) for all i, j, $s \mapsto \mathbb{E}(|v_i(s) - v_j(s)|^2)$ is well defined and differentiable with a bounded derivative,

(3) the system is $\mathbb{L}^{2,2}$ strongly swarming.

Then the system is $\mathbb{L}^{2,2}$ flocking.

We shall check the required assumptions for each model.

2.4 Relaxing (H2) in (2.1.6) for constant communication rates.

In this section we shall study the model (2.1.6), still assuming that (2.2.1) is satisfied, but relaxing the assumption (H2). Namely we will consider the following general model

$$
dv_i(t) = -\frac{\lambda\psi}{N}\sum_{j=1}^N (v_i(t)-v_j(t))\, dt \,+\, \sigma(v_i(t), x_i(t))\, dw(t)\,, \tag{2.4.1}
$$

where w is a d-dimensional Brownian motion (the same for all the particles). That is, we consider that the dynamics of a particle is perturbed by a noisy environment depending on the position and the velocity of this particle.

Once again the dynamics of the center of mass is given by a (at least local) martingale

$$
d\bar{v}(t) \,=\, \frac{1}{N}\left(\sum_{i=1}^N \sigma(v_i(t), x_i(t))\right) dw(t) \,:=\, s(v(t), x(t))\, dw(t)\,. \tag{2.4.2}
$$

It follows

$$d\hat{v}_i^k(t) = -\lambda\psi\, \hat{v}_i^k(t)\, dt + \sum_{l=1}^{d} \theta_i^{k,l}(v(t), x(t))\, dw^l(t)$$

where

$$\theta_i^{k,l}(v,x) = \sigma^{k,l}(v_i, x_i) - s^{k,l}(v,x) \;=\; \frac{1}{N} \sum_{j=1}^{N} (\sigma^{k,l}(v_i,x_i) - \sigma^{k,l}(v_j,x_j))\,.$$

Of course we will assume enough regularity on σ for (2.4.1) to admit a unique solution. Notice that if

$$v_i(0) = v_0 \quad \text{and} \quad x_i(0) = x_0 \quad \text{for all } i,$$

then the unique solution of (2.4.1) is given by a dynamic equilibrium $v_i(t) = \bar{v}(t)$ and $x_i(t) = \bar{x}(t)$ for all i, where $(\bar{v}, \bar{x})$ solves

$$\begin{aligned} d\bar{v}(t) &= \sigma(\bar{v}(t), \bar{x}(t))\, dw(t) \\ d\bar{x}(t) &= \bar{v}(t)\, dt\,. \end{aligned}$$

There is however a difference with the deterministic model (or the model assuming (H2)): this time one has in general to fix the initial positions to get some equilibrium.

As we did in the first section we define

$$\begin{aligned} z^k(t) := \sum_{i=1}^{N} (\hat{v}_i^k)^2(t) \;&=\; z^k(0) \;-\; 2\lambda\psi \int_0^t z^k(s)\, ds \;+\; 2 \sum_{i=1}^{N} \sum_{l=1}^{d} \int_0^t \hat{v}_i^k(s)\, \theta_i^{k,l}(v(s), x(s))\, dw^l(s) \\ &\quad + \int_0^t \left(\sum_{i=1}^{N} \sum_{l=1}^{d} (\theta_i^{k,l})^2(v(s), x(s)) \right) ds\,, \end{aligned} \tag{2.4.3}$$

so that

$$\begin{aligned} z(t) := \sum_{k=1}^{d} z^k(t) \;&=\; z(0) \;-\; 2\lambda\psi \int_0^t z(s)\, ds \;+\; 2 \sum_{i=1}^{N} \sum_{l=1}^{d} \int_0^t \left(\sum_{k=1}^{d} \hat{v}_i^k(s)\, \theta_i^{k,l}(v(s), x(s)) \right) dw^l(s) \\ &\quad + \int_0^t \left(\sum_{k=1}^{d} \sum_{i=1}^{N} \sum_{l=1}^{d} (\theta_i^{k,l})^2(v(s), x(s)) \right) ds\,. \end{aligned} \tag{2.4.4}$$

Hence

$$u(t) := \mathbb{E}(z(t)) = u(0) \;-\; 2\lambda\psi \int_0^t u(s)\, ds \;+\; \int_0^t U(v(s), x(s))\, ds$$

where

$$U(v(s), x(s)) = \sum_{k=1}^{d} \sum_{i=1}^{N} \sum_{l=1}^{d} \mathbb{E}\left[(\theta_i^{k,l})^2(v(s), x(s)) \right]\,. \tag{2.4.5}$$

Finally, at least formally (and rigorously up to the first time $z(.)$ hits 0)

$$\begin{aligned}
\ln z(t) \;&=\; \ln z(0) \,+\, 2 \sum_{i=1}^{N} \sum_{l=1}^{d} \int_0^t \left(\sum_{k=1}^{d} \frac{\hat{v}_i^k(s)\, \theta_i^{k,l}(v(s),x(s))}{z(s)}\right) dw^l(s) \,-\, 2\lambda\psi\, t \\
&\quad + \int_0^t \left(\sum_{i=1}^{N}\sum_{l=1}^{d}\sum_{k=1}^{d} \frac{(\theta_i^{k,l})^2(v(s),x(s))}{z(s)}\right) ds \qquad (2.4.6)\\
&\quad - 2 \int_0^t \sum_{l=1}^{d} \left(\frac{\left[\sum_{i=1}^{N}\sum_{k=1}^{d} \hat{v}_i^k(s)\, \theta_i^{k,l}(s)\right]^2}{(z(s))^2}\right) ds\,.
\end{aligned}$$

2.4.1 A first natural generalization of (H2).

Introduce the following assumption

(H2-1) σ only depends on v and is Lipschitz continuous, i.e. there exists K such that for all k,l, all (v,v') ,

$$|\sigma^{k,l}(v) - \sigma^{k,l}(v')| \leqslant K\,|v-v'|\,.$$

In this situation we have

$$\begin{aligned}
|\theta_i^{k,l}(v,x)| \;&\leqslant\; \frac{1}{N}\sum_{j\neq i} |\sigma^{k,l}(v_i) - \sigma^{k,l}(\bar{v}) + \sigma^{k,l}(\bar{v}) - \sigma^{k,l}(v_j)| \\
&\leqslant\; \frac{N-1}{N}\,|\sigma^{k,l}(v_i) - \sigma^{k,l}(\bar{v})| + \frac{1}{N}\sum_{j\neq i} |\sigma^{k,l}(v_j) - \sigma^{k,l}(\bar{v})| \\
&\leqslant\; K\,|\hat{v}_i| + \frac{K}{N}\sum_{j=1}^{N} |\hat{v}_j|\,.
\end{aligned}$$

Hence

$$\sum_{i=1}^{N} |\theta_i^{k,l}(v,x)|^2 \leqslant K^2 \left(\sum_{i=1}^{N} |\hat{v}_i|^2 + \frac{3}{N}\left(\sum_{j=1}^{N} |\hat{v}_j|\right)^2\right) \leqslant 4K^2 \sum_{i=1}^{N} |\hat{v}_i|^2$$

and finally

$$\sum_{k=1}^{d}\sum_{i=1}^{N}\sum_{l=1}^{d} (\theta_i^{k,l})^2(v(t),x(t)) \,\leqslant\, 4d^2K^2\, z(t)\,. \qquad (2.4.7)$$

Of course if σ is diagonal, we may replace d^2 by d, and if in addition $\sigma^{k,k}$ only depends on v^k we may replace d by 1 (as for (H2)).
Similarly, using Cauchy-Schwartz inequality it is easily seen that

$$\left(\frac{\left[\sum_{i=1}^{N}\sum_{k=1}^{d} \hat{v}_i^k(t)\, \theta_i^{k,l}(t)\right]^2}{(z(t))^2}\right)$$

is uniformly bounded above.
We may thus use the same arguments as for the end of the previous proof of Theorem 2.2.2,

except that we do no more have any better lower bound for

$$\left(\frac{\left[\sum_{i=1}^N \sum_{k=1}^d \hat{v}_i^k(t)\, \theta_i^{k,l}(t)\right]^2}{(z(t))^2}\right)$$

than 0. We have thus obtained

Theorem 2.4.8. Assume that (H2-1) is satisfied in (2.4.1). Then if $2\lambda\psi > 4K^2d^2$, the system (2.4.1) is almost surely and $\mathbb{L}^{2,2}$ flocking.
However, contrary to what happens when (H2) is satisfied, the center of mass $\bar{v}(t)$ does not necessarily converge as $t \to +\infty$.

Let us look at a very particular case: the case when σ is diagonal and $\sigma^{k,k}(v) = \sigma^{k,k}(v^k)$. We can thus rewrite (2.4.1)

$$dv_i^k(t) = -\frac{\lambda\psi}{N}\sum_{j=1}^N (v_i^k(t) - v_j^k(t))\, dt \,+\, \sigma^{k,k}(v_i^k(t))\, dw^k(t)\,, \tag{2.4.9}$$

i.e. we can look at the system independently for each coordinate k, or if one prefers, reduce the problem to the case of one dimensional particles i.e. $d = 1$. In the sequel we thus suppress the superscript k.
The first elementary remark is that, if $v_i(0) = v_j(0)$ for some pair $i \neq j$, the uniqueness of the solution shows that $v_i(t) = v_j(t)$ for all t. Using the Markov property, the same holds for $t \geqslant T$ for any stopping time T such that $v_i(T) = v_j(T)$. Reordering the indices if necessary we may assume that $v_1(0) \leqslant v_2(0) \leqslant ... \leqslant v_N(0)$ so that the dynamics preserves the order of the velocities of the particles. The best quantity to look at is thus $D_{1N}(t) = v_N(t) - v_1(t)$ instead of $z(t)$, since $D_{1N} \geqslant \hat{v}_i$ for all i. The dynamics of D_{1N} is given by

$$dD_{1N}(t) = -\lambda\psi\, D_{1N}(t)\, dt + (\sigma(v_N(t)) - \sigma(v_1(t)))\, dw(t)\,.$$

Since D_{1N} is non negative we have

$$\ln D_{1N}(t) = \ln D_{1N}(0) - \lambda\psi\, t - \int_0^t \frac{(\sigma(v_N(s)) - \sigma(v_1(s)))^2}{2D_{1,N}^2(s)}\, ds + \int_0^t \frac{(\sigma(v_N(s)) - \sigma(v_1(s)))}{D_{1N}(s)}\, dw(s)\,,$$

and

$$\begin{aligned} D_{1N}^2(t) \;&=\; D_{1N}^2(0) - \int_0^t (2\lambda\psi\, D_{1N}^2(s) \,-\, (\sigma(v_N(s)) - \sigma(v_1(s)))^2)\, ds \\ &\quad + \int_0^t 2D_{1N}(s)\, (\sigma(v_N(s)) - \sigma(v_1(s)))\, dw(s)\,. \end{aligned}$$

Using the same arguments as before we thus have

Theorem 2.4.10. Assume that (H2-1) is satisfied in (2.4.9). Then the system (2.4.9) is always almost surely flocking. If in addition $2\lambda\psi > K^2$, then it is also $\mathbb{L}^{2,2}$ flocking.

2.4.2 More general environments.

One may ask about the physical meaning of a random environment acting on the velocities only. It can be the case for some aerodynamical perturbations for instance. But of course, it is more natural (or at least as natural) to add some random perturbation that depends on the position (and possibly the velocity too) of each particle. We shall now discuss briefly this situation.

Assume for instance that all $\sigma^{k,l}$ are bounded, say by M. We thus have

$$\mathbb{E}(z_t) := u(t) \leqslant u(0) - 2\lambda\psi \int_0^t u(s)\, ds \,+\, 4M^2\, N\, d^2\, t\,, \tag{2.4.11}$$

so that

$$\limsup_{t\to+\infty} u(t) \,\leqslant\, \frac{2M^2 N d^2}{\lambda\psi}\,. \tag{2.4.12}$$

In particular $u(.)$ is bounded on $\mathbb{R}^+$. Hence if all the σ's are bounded in (2.4.1), the system is $\mathbb{L}^2$ weakly swarming, while in general it is hard to say anything about strong swarming. Concerning this last point let us look at some particular case, namely

(H2-2) $\sigma(v,x) = \sigma(x)$ is C^1 with partial derivatives bounded by K (but the $\sigma^{k,l}$ are not necessarily bounded themselves).

As for (2.4.9) we may look at each coordinate k individually, i.e. consider a system of Nd 1-dimensional particles governed by

$$dv_i^k(t) = -\,\frac{\lambda\psi}{N}\,\sum_{j=1}^{N}\,(v_i^k(t) - v_j^k(t))\, dt \,+\, \sum_{l=1}^{d}\,\sigma^{k,l}(x_i(t))\, dw^l(t)\,. \tag{2.4.13}$$

Contrary to the situation of Theorem 2.4.10, in general the order of the velocities v_i^k is not preserved by the dynamics, and the only trivial equilibrium is given by $v_i(t) = \bar{v}(0)$ and $x_i(t) = \bar{x}(0) + t\bar{v}(0)$ for all t.
We shall nevertheless look at

$$v_{i,j}^k(t) = v_i^k(t) - v_j^k(t) \quad \text{ and } \quad x_{i,j}(t) = x_i(t) - x_j(t)$$

which solves

$$dv_{i,j}^k(t) = -\lambda\psi\, v_{i,j}^k(t)\, dt + \sum_{l=1}^{d}\,\sigma_{i,j}^{k,l}(t)\, dw_t^l$$

where $\sigma_{i,j}^{k,l}(t) = \sigma^{k,l}(x_i(t)) - \sigma^{k,l}(x_j(t))$. We already know that, if σ is bounded, the system is $\mathbb{L}^2$ weakly swarming. Here we assume that

$$\sup_{t>0}\, \mathbb{E}(|x_i(t) - x_j(t)|^2) \leqslant M_{i,j}^2 < +\infty. \tag{2.4.14}$$

Of course when σ is bounded (2.4.14) implies that the system is $\mathbb{L}^2$ strongly swarming. We shall first show that it is still the case when (H2-2) is satisfied.

Let us make some computations: first if T_R denotes the first time $|v_{i,j}|(.)$ exceeds R, we have

$$\begin{aligned}
\mathbb{E}((v^k_{i,j})^2(t\wedge T_R)) &= \mathbb{E}((v^k_{i,j})^2(0)) - 2\lambda\psi\,\mathbb{E}\left(\int_0^{t\wedge T_R} (v^k_{i,j})^2(s)\,ds\right) + \mathbb{E}\left(\int_0^{t\wedge T_R} \sum_{l=1}^d (\sigma^{k,l}_{i,j})^2(s)\,ds\right)\\
&\leqslant \mathbb{E}((v^k_{i,j})^2(0)) + d\,K^2\,M^2_{i,j}\,t
\end{aligned}$$

so that $u^k_{i,j}(t) = \mathbb{E}((v^k_{i,j})^2(t))$ is well defined and satisfies

$$\mathbb{E}((v^k_{i,j})^2(t)) := u^k_{i,j}(t) \leqslant u^k_{i,j}(0) - 2\lambda\psi \int_0^t u^k_{i,j}(s)\,ds + d\,K^2\,M^2_{i,j}\,t$$

and finally

$$\limsup_{t\to+\infty} u^k_{i,j}(t) \leqslant \frac{d\,K^2\,M^2_{i,j}}{2\lambda\psi}\,. \tag{2.4.15}$$

It follows that

$$\sup_{t\geqslant 0} \mathbb{E}((v^k_{i,j})^2(t)) \leqslant N^2_{i,j} < +\infty\,.$$

Using what precedes we also see that $s \mapsto \mathbb{E}((v^k_{i,j})^2(s))$ is differentiable with

$$\frac{d}{ds}\,\mathbb{E}((v^k_{i,j})^2(s)) = -\,2\lambda\psi\,\mathbb{E}((v^k_{i,j})^2(s)) + \mathbb{E}(\sum_{l=1}^d (\sigma^{k,l}_{i,j})^2(s))$$

which is bounded below by $-\,2\lambda\psi\,N^2_{i,j}$ and bounded above by $d\,M^2_{i,j}K^2$. Hence we may use Lemma 2.3.11 in order to get

Lemma 2.4.16. Consider the system (2.4.1) under the assumption (H2-2). If for all pair (i,j),

$$\sup_{t\geqslant 0} \mathbb{E}(|x_i(t) - x_j(t)|^2) \leqslant M < +\infty$$

then the system (2.4.1) is $\mathbb{L}^{2,2}$ flocking.

But we can go further. Indeed, in the situation of the previous lemma, we first of all have

$$\mathbb{E}((v^k_{i,j})^2(0)) + \int_0^{+\infty} \mathbb{E}(\sum_{l=1}^d (\sigma^{k,l}_{i,j})^2(s))\,ds = 2\lambda\psi \int_0^{+\infty} \mathbb{E}((v^k_{i,j})^2(s))\,ds < +\infty\,. \tag{2.4.17}$$

On one hand, using lemma 2.3.10 again (it is easily seen that the assumptions are satisfied) we thus obtain

$$\lim_{t\to+\infty} \mathbb{E}((\sigma^{k,l}_{i,j})^2(t)) = \lim_{t\to+\infty} \mathbb{E}((\sigma^{k,l}(x_i(t)) - \sigma^{k,l}(x_j(t)))^2) = 0\,. \tag{2.4.18}$$

On the other hand, as before the martingale $m^k_{i,j}(t) = \sum_{l=1}^d \int_0^t \sigma^{k,l}_{i,j}(s)\,dw^l(s)$ converges (as $t \to +\infty$) almost surely and in $\mathbb{L}^2$ to a random variable $m^k_{i,j}$ such that

$$\mathbb{E}[m^k_{i,j}|\mathcal{F}^k(t)] = m^k_{i,j}(t)\,,$$

$\mathcal{F}^k(.)$ being the filtration of the Brownian motion $w(.)$. Notice that

$$m^k_{i,j}(t) = (\lambda\psi\, x^k_{i,j}(t) + v^k_{i,j}(t)) - (\lambda\psi\, x^k_{i,j}(0) + v^k_{i,j}(0))\,. \tag{2.4.19}$$

We deduce that $x^k_i(t) - x^k_j(t)$ converges in Probability as $t \to +\infty$ to

$$\frac{1}{\lambda\psi}\, m^k_{i,j} + ((x^k_i(0) - x^k_j(0)) + \frac{1}{\lambda\psi}(v^k_i(0) - v^k_j(0)))$$

(since $v^k_{i,j}(t)$ goes to 0 in $\mathbb{L}^2$ hence in Probability). In addition

$$\lim_{t\to+\infty} \mathbb{E}((x^k_{i,j})^2(t)) = \mathbb{E}((x^k_{i,j}(0) + \frac{1}{\lambda\psi}\, v^k_{i,j}(0))^2) + \frac{1}{(\lambda\psi)^2}\,\mathbb{E}((m^k_{i,j})^2)\,. \tag{2.4.20}$$

It follows that the above convergence in Probability also holds in $\mathbb{L}^p$ for all $p < 2$.
Hence

Proposition 2.4.21. Consider the system (2.4.1) under the assumption (H2-2). If for all pair (i, j),

$$\sup_{t\geqslant 0} \mathbb{E}(|x_i(t) - x_j(t)|^2) \leqslant M < +\infty$$

then the system (2.4.1) satisfies the following

(1) it is $\mathbb{L}^{2,2}$ flocking,

(2) there exists some random vector $\hat{x}(\infty)$ such that $\hat{x}(t)$ converges in $\mathbb{L}^p$ $(p < 2)$ towards $\hat{x}(\infty)$ as $t \to +\infty$.

Remark 2.4.22. Notice that if $\lim_{t\to+\infty} \mathbb{E}((x^k_{i,j})^2(t)) = 0$, then $0 = \lambda\psi\, x^k_{i,j}(0) + v^k_{i,j}(0)$ and $m^k_{i,j} = 0$.
If $m^k_{i,j} = 0$, then $m^k_{i,j}(t) = 0$ for all $t \geqslant 0$, so that

$$v^k_{i,j}(t) = v^k_{i,j}(0)\, e^{-\lambda\psi\, t} \quad ; \quad x^k_{i,j}(t) = x^k_{i,j}(0) + \frac{v^k_{i,j}(0)}{\lambda\psi}\,(1 - e^{-\lambda\psi\, t})\,.$$

So

$$0 = \lambda\psi\, x^k_{i,j}(0) + v^k_{i,j}(0) = \lambda\psi\, x^k_{i,j}(t) + v^k_{i,j}(t)\,.$$

But, since $m^k_{i,j}(t) = 0$ for all t, we also have for all l,

$$\sigma^{k,l}(x_i(t)) - \sigma^{k,l}(x_j(t)) = 0 \quad \text{for all } t \geqslant 0\,.$$

In particular if $\sigma^{k,\cdot} : \mathbb{R}^d \mapsto \mathbb{R}^d$ is one to one, we get $x_{i,j}(t) = 0$ for all t, hence $v_{i,j}(t) = 0$ for all t. ♢

Let us illustrate the previous remark with a simple example

Example 2.4.23. *Almost affine diffusion coefficient.*
Assume that for some k,

$$\sigma^k(x) = Ax + B$$

for some constant invertible matrix A and constant vector B. Then, if (2.4.14) is satisfied for all pair (i,j), (2.4.18) yields

$$\lim_{t\to+\infty} \mathbb{E}(|x_i(t)-x_j(t)|^2) = \lim_{t\to+\infty} \mathbb{E}(|A^{-1}(\sigma^k(x_i(t))-\sigma^k(x_j(t))|^2) = 0\,,$$

for all pair (i,j). According to the previous remark, the system is thus at equilibrium. Hence

Proposition 2.4.24. In addition to (H2-2), if for some k, $\sigma^k(x) = Ax+B$ for some constant invertible matrix A and constant vector B, the system (2.4.1) cannot be strongly $\mathbb{L}^{2,2}$ swarming, except if it is at equilibrium (all coordinates are equal).

More generally (almost) the same occurs if one of the σ^k (k-th row of the matrix σ), in addition to be one to one, satisfies the following property: for a sequence $(x(n),y(n))$, $\sigma^k(x(n))-\sigma^k(y(n))\to 0$ implies $x(n)-y(n)\to 0$.
To see it, recall that (2.4.18) implies that $\sigma^k(x_i(t))-\sigma^k(x_j(t))\to 0$ in Probability. Hence up to a subsequence t_n we may assume that it converges almost surely, so that $x_i(t_n)-x_j(t_n)\to 0$ almost surely. But since $x_i(t)-x_j(t)$ goes to some $x_{i,j}(\infty)$ as $t\to+\infty$ in Probability, we deduce that $x_{i,j}(\infty)=0$. Using Lebesgue's bounded convergence theorem we can thus deduce

Proposition 2.4.25. In addition to (H2-2), assume that for some k, σ^k is one to one and satisfies (H2-21): for a sequence $(x(n),y(n))$, $\sigma^k(x(n))-\sigma^k(y(n))\to 0$ implies $x(n)-y(n)\to 0$.
Then if the system (2.4.1) is uniformly $\mathbb{L}^{2,2}$ swarming (i.e. $\max_{i,j}\sup_{t\geqslant 0}|x_i(t)-x_j(t)| = M\in\mathbb{L}^2$), the system (2.4.1) is at equilibrium (all coordinates are equal).

The latter statement can be extended: if for instance $\sigma^k(x(n))-\sigma^k(y(n))\to 0$ only implies $x^k(n)-y^k(n)\to 0$, then the conclusion of the proposition is still true provided the previous property is satisfied for all k. ♢

The previous assumptions on σ imply in a sense that it cannot be bounded. Indeed for $d=1$, a smooth one to one function from $\mathbb{R}$ to $\mathbb{R}$ which is bounded, admits a limit at infinity and thus cannot satisfy (H2-21). The typical example of smooth bounded (and presumably interesting from a physical point of view) function is the case of periodic functions we shall look at now.

Example 2.4.26. *Periodic diffusion coefficient.*
Assume now that σ is T-periodic. For $x\in\mathbb{R}^d$ we denote $\tilde{x}$ the unique vector in $[0,T[^d$ such that $x^k-\tilde{x}^k$ belongs to $T\mathbb{Z}$ for all $k=1,...,d$. By T-periodic we mean that $\sigma(x)=\sigma(\tilde{x})$.
We shall introduce a new "one to one" assumption:
(H2-22). The set

$$N = \{\tilde{z} = \tilde{x}-\tilde{y} \text{ such that for all } (k,l)\,,\ \sigma^{k,l}(x)-\sigma^{k,l}(y)\,\} \text{ is reduced to } \{0\}\,.$$

For instance if $d=2$, the matrix

$$\sigma(x^1,x^2) = \begin{pmatrix} \sin(x^1) & \cos(x^2) \\ \cos(x^1) & \sin(x^2) \end{pmatrix}$$

satisfies (H2-22) with $T = 2\pi$. The matrix

$$\sigma(x^1, x^2) = \begin{pmatrix} \sin(x^1) & \cos(x^1) \\ \\ \cos(x^2) & \sin(x^2) \end{pmatrix}$$

also does, but this case reduces after an immediate change of Brownian motion, to the case of a constant diffusion coefficient.

If the system is strongly $\mathbb{L}^{2,2}$ swarming, we can as in the previous example, find some sequence t_n such that for all (k,l), $\sigma^{k,l}(\tilde{x}_i(t_n)) - \sigma^{k,l}(\tilde{x}_j(t_n)) \to 0$ almost surely. According to proposition 2.4.21, $x_i(t) - x_j(t)$ goes to $x_{i,j}(\infty)$ in probability, so that taking a subsequence of t_n if necessary (we still denote by t_n), we may assume that the convergence is almost sure. It follows that $\tilde{x}_i(t_n) - \tilde{x}_j(t_n)$ goes almost surely to $\tilde{x}_{i,j}(\infty)$.
Thanks to compactness, we have that for each ω for which both previous convergences hold, extracting another subsequence if necessary both $\tilde{x}_i(t'_n, \omega)$ and $\tilde{x}_j(t'_n, \omega)$ converge to limits $\tilde{x}_i(\infty, \omega)$ and $\tilde{x}_j(\infty, \omega)$, for which, using the continuity of σ, it holds that $\sigma^{k,l}(\tilde{x}_i(\infty, \omega)) - \sigma^{k,l}(\tilde{x}_j(\infty, \omega)) = 0$. If (H2-22) is satisfied, we deduce that $\tilde{x}_i(\infty, \omega) - \tilde{x}_j(\infty, \omega) = 0$, i.e $\tilde{x}_{i,j}(\infty, \omega) = 0$ for almost all ω. It means that $x_{i,j}(\infty)$ is a random variable taking its values in $(T\mathbb{Z})^d$.
The key point now is the following: if we add to $x(0)$ any $\mathbb{L}^2$ random vector whose coordinates belong to $(T\mathbb{Z})^d$, we do not change the dynamics of the $v(.)$. Hence, replacing all $x_i(0)$ by $x'_i(0) = x_i(0) + x_{1,i}(\infty)$ (for $i > 1$), we do not change the dynamics of the v_i, we do not change the strong swarming property, nor the uniform swarming property, and we get in the limit $x'_{1,i}(\infty) = 0$, hence for all (i,j), $x'_{i,j}(\infty) = 0$. But now we may use remark (2.4.22), periodicity and (H2-22) to conclude that all $x'_{i,j}(.)$ and all $v_{i,j}(.)$ are equal to 0 as soon as $\lim_{t \to +\infty} \mathbb{E}(((x')^k_{i,j}(t))^2) = 0$, which is satisfied, thanks to Lebesgue bounded convergence theorem as soon as the system is uniformly $\mathbb{L}^{2,2}$ swarming. Notice that now any random vector $(v,x) = (0,x)$ which x taking values in $(T\mathbb{Z})^{dN}$ is an equilibrium. We thus have

Proposition 2.4.27. In addition to (H2-2), assume that σ is T periodic and satisfies (H2-22). Then if the system (2.4.1) is uniformly $\mathbb{L}^{2,2}$ swarming (i.e. $\max_{i,j} \sup_{t \geqslant 0} |x_i(t) - x_j(t)| = M \in \mathbb{L}^2$), the system (2.4.1) is at equilibrium (all velocities are equal and the differences between the positions belong to $(T\mathbb{Z})^d$).

Hence in all situations we are able to handle, uniform swarming does not occur, unless the system is at equilibrium, telling us that for a random environment depending on the positions only, it seems difficult to swarm out of equilibrium. ♢

2.5 A general form of (2.1.7) for constant communication rates.

We have already seen that the particular form (2.1.8) of (2.1.7) with constant communication rate is a particular case of (2.1.6). Also notice that, still for constant communication rate, when $w_{i,j} = w_i$ for all j, the w_i being independent, and $\sigma_{i,j} = \sigma_i \, \psi_{i,j}$, we recognize (2.1.9). We shall now look at another case, namely

$$dv_i(t) = -\,\frac{\lambda\psi}{N}\,\sum_{j=1}^{N}(v_i(t)-v_j(t))\,dt \,+\, \frac{1}{N}\,\sum_{j=1}^{N}\sigma_{ij}(t)\,(v_i(t)-v_j(t))\,dw_{i,j}(t)\,, \tag{2.5.1}$$

where the w_{ij} are d-dimensional noises (here vw is the vector such that each coordinate $(vw)^k$ is given by $v^k w^k$). We shall assume that

$$\sigma_{ij}=\sigma_{ji} \quad , \quad w_{i,j}=w_{j,i} \quad , \quad \text{and } (w_{i,j})_{i<j} \text{ are independent.} \tag{2.5.2}$$

The meaning of these assumptions seems a little bit more natural that for the (2.1.9) model: each pair of individuals (i,j) are interacting symmetrically with a constant communication rate which is perturbed by some noise (we may include ψ in the σ_{ij}), all the interaction noises being independent. Since we are speaking of constant communication rate, we shall also assume that the σ_{ij} are constant (more general situations will be discussed in the next section).

As we did before we shall look at $v_{i,j}=v_i-v_j$ which solves

$$dv_{i,j}(t) = -\,\lambda\psi\,v_{i,j}(t)\,dt \,+\, \frac{1}{N}\,\sum_{l=1}^{N}\sigma_{il}\,v_{i,l}(t)\,dw_{i,l}(t) \,-\, \frac{1}{N}\,\sum_{m=1}^{N}\sigma_{jm}\,v_{j,m}(t)\,dw_{j,m}(t)\,.$$

As before we can look separately at each coordinate (v^k,x^k). For the sake of simplicity, we skip the superscript k in the sequel, or if one prefers we take $d=1$.
Hence if we define $z(t)=\sum_{1\leqslant i,j\leqslant N}(v_{i,j})^2(t)$ (we skip the $2N$ in (2.3.1)), we have (being careful with the indices for which the Brownian motions are independent on one hand or the same on the other hand)

$$\begin{aligned} dz(t) &= -\,2\,\lambda\psi\,z(t)\,dt + \frac{4}{N}\,\sum_{i,j,l=1}^{N}\sigma_{il}\,v_{i,l}(t)\,v_{i,j}(t)\,dw_{i,l}(t) \\ &\quad + \frac{2}{N^2}\,\left(\sum_{i,j,l=1}^{N}\sigma_{il}^2 v_{i,l}^2(t) + \sum_{i,j=1}^{N}\sigma_{ij}^2 v_{i,j}^2(t)\right)\,dt \\ &= -\,2\,\lambda\psi\,z(t)\,dt + 4\,\sum_{i,l=1}^{N}\sigma_{il}\,v_{i,l}(t)\,\hat{v}_i(t)\,dw_{i,l}(t) \,+\, \frac{2(N+1)}{N^2}\,\left(\sum_{i,j=1}^{N}\sigma_{ij}^2 v_{i,j}^2(t)\right)\,dt\,. \end{aligned} \tag{2.5.3}$$

It follows

$$u(t) := \mathbb{E}(z(t)) = u(0) \,-\, 2\,\lambda\psi\,\int_0^t u(s)\,ds + \frac{2(N+1)}{N^2}\,\int_0^t \left(\sum_{i,j=1}^{N}\sigma_{ij}^2\mathbb{E}(v_{i,j}^2(s))\right)\,ds\,,$$

from which we deduce

$$\left(\frac{(N+1)}{N^2}\,\min_{i,j}\sigma_{ij}^2 - \lambda\psi\right)\int_0^t u(s)\,ds \leqslant \frac{u(t)-u(0)}{2} \leqslant -\,\left(\lambda\psi - \frac{(N+1)}{N^2}\,\max_{i,j}\sigma_{ij}^2\right)\int_0^t u(s)\,ds\,.$$

The latter furnishes conditions for $\mathbb{L}^{2,2}$ flocking or non flocking.

For almost sure flocking we may consider as we did before $\ln(z(t))$ which solves

$$\begin{aligned} d(\ln(z(t))) &= -2\,\lambda\psi\,dt + 4\sum_{i,l=1}^{N} \frac{(\sigma_{il}\,v_{i,l}\,\hat{v}_i)(t)}{z(t)}\,dw_{i,l}(t) \qquad (2.5.4)\\ &\quad + \frac{2(N+1)}{N^2}\sum_{i,l=1}^{N} \frac{(\sigma_{il}^2\,v_{i,l}^2)(t)}{z(t)}\,dt - 4\sum_{i,l=1}^{N} \frac{(\sigma_{il}^2\,v_{i,l}^4)(t)}{z^2(t)}\,dt\,. \end{aligned}$$

The non constant part of the drift term can be rewritten

$$A(t) = \frac{2}{z^2(t)}\left(\frac{N+1}{N}\left[\sum_{i,l=1}^{N}(\sigma_{il}^2\,v_{i,l}^2)(t)\right]\left[\frac{1}{N}\sum_{i,l=1}^{N}(v_{i,l}^2)(t)\right] - 2\sum_{i,l=1}^{N}(\sigma_{il}^2\,v_{i,l}^4)(t)\right)$$

so that using again $\sum_{i=1}^{N}\beta_i^4 \geqslant \frac{1}{N}(\sum_{i=1}^{N}\beta_i^2)^2$ we get

$$A(t) \leqslant \frac{2}{N}\left(\frac{N+1}{N}\max_{i,j}\sigma_{ij}^2 - 2\min_{i,j}\sigma_{ij}^2\right).$$

Now we may argue as in the previous section. We have thus obtained

Theorem 2.5.5. Consider the system (2.5.1), under the assumption (2.5.2) and with constant σ_{ij}. Then

(1) If $\lambda\psi > \frac{N+1}{N^2}\max_{i,j}\sigma_{ij}^2$ the system is $\mathbb{L}^{2,2}$ flocking.

(2) If $\lambda\psi < \frac{N+1}{N^2}\min_{i,j}\sigma_{ij}^2$ the system is not $\mathbb{L}^2$ flocking.

(3) If $\lambda\psi - \frac{1}{N}\left(\frac{N+1}{N}\max_{i,j}\sigma_{ij}^2 - 2\min_{i,j}\sigma_{ij}^2\right) > 0$, the system is almost surely flocking. In particular if $\sigma_{ij} = \sigma$ for all pair (i,j), the system is always almost surely flocking.

Notice that the flocking properties are still the same if we consider bounded processes $\sigma_{ij}(.)$ instead of constants. Also note that we could improve the bounds for almost sure flocking by using a more accurate comparison between $\sum \sigma_{il}^2\,v_{i,l}^2$ and $\sum v_{i,l}^2$, but the present statement is easier.

Remark 2.5.6. If we compare with (2.1.7) in his (2.1.8) version, the correspondence is $\bar{\sigma} = \frac{\sigma}{N}$. The comparison for flocking is thus between $\lambda\psi$ and σ^2/N^2 and not with σ/N. Of course this is simply the observation that the variance of the noise is of order $1/N^2$ in (2.1.8) while it is of order $1/N$ here. ♢

2.6 General communication rate.

Since we are mainly interested in flocking or swarming properties, we shall only consider models for which such properties may hold for constant communication rate. [12] contains informations on (2.1.5) for which it is possible to show the existence of stationary solutions (using Ito-Nisio theory for stochastic delayed equations) as well as propagation of chaos when N grows to infinity (also see [2] for this latter point). If we consider models for random environment, we will only

look at the case where the environment depends on the velocity only. Hence we will focus on two type of systems.
First, noisy communication rates i.e.

$$dv_i(t) = -\frac{\lambda}{N}\sum_{j=1}^{N}\psi_{ij}(t)(v_i(t)-v_j(t))\,dt \,+\, \frac{1}{N}\sum_{j=1}^{N}\sigma_{ij}(t)\,(v_i(t)-v_j(t))\,dw_{i,j}(t)\,, \tag{2.6.1}$$

where the w_{ij} are d-dimensional noises (again vw is the vector such that each coordinate $(vw)^k$ is given by $v^k w^k$), $w_{i,j} = w_{j,i}$ and the $(w_{i,j})_{i<j}$ are independent Brownian motions.
Next, noisy environment

$$dv_i(t) = -\frac{\lambda}{N}\sum_{j=1}^{N}\psi_{ij}(t)(v_i(t)-v_j(t))\,dt \,+\, \sigma(v_i(t))\,dw(t)\,, \tag{2.6.2}$$

where w is a d-dimensional Brownian motion.

2.6.1 Study of (2.6.2).

Consider the model given by (2.6.2). We shall introduce assumptions ensuring first existence and uniqueness.

Proposition 2.6.3. Assume that

(1) The processes $\psi_{ij}(t)$ can be written $\psi_{ij}(t) = \psi_{ij}(v(t), x(t))$, where all the functions ψ_{ij} are local Lipschitz, non-negative and satisfy $\psi_{ij} = \psi_{ji}$,

(2) σ satisfies (H2-1) i.e. is globally K-Lipschitz or σ is local Lipschitz and bounded.

Then, for all initial state $(v(0), x(0)) \in \mathbb{L}^2$ the system (2.6.2) admits a unique non-explosive (global) strong solution.

Proof. Existence of a unique local strong solution is immediate thanks to our assumptions. The only thing to prove is that it is global. Actually it is enough to show that $v(.)$ does not explode and to this end, as usual, it is enough to show that for all $t \geqslant 0$,

$$\sup_{R>0} \mathbb{E}(|v(t\wedge T_R)|^2) < +\infty$$

where T_R denotes the first (stopping) time $|v(.)|$ hits the value R. Defining $V(.) = |v(.)|^2$ we have, using Ito's formula and (2.3.3), that for $t \leqslant T_R$,

$$\begin{aligned} dV(t) \quad &= \quad -\frac{\lambda}{N}\sum_{1\leqslant i,j\leqslant N}\psi_{ij}(t)|v_i(t)-v_j(t)|^2\,dt + \sum_{i=1}^{N} Trace\,(\sigma(v_i(t))\sigma^*(v_i(t)))\;dt \\ &\qquad +2\left(\sum_{i=1}^{N} v_i^*(t)\sigma(v_i(t))\right)\,dw(t)\,, \end{aligned}$$

where a^* denotes the transposed of the vector (or the matrix) a. When σ is K-Lipschitz,

$|\sigma^{k,l}(v_i)| \leqslant K|v_i| + c$ for all (k,l), so that

$$\mathbb{E}(|v(t\wedge T_R)|^2) \leqslant \mathbb{E}(|v(0)|^2) + C(N)\left(\int_0^t K\mathbb{E}(|v(s\wedge T_R)|^2)\,ds + ct\right),$$

and the result follows using Gronwall's lemma. When σ is bounded the result is immediate. □

Remark 2.6.4. It is worth noticing that if $v_i(0) = \bar{v}(0)$ for all i, the unique solution of (2.6.2) is given by $v_i(t) = \bar{v}(t)$, $x_i(t) = x_i(0) + \int_0^t \bar{v}(s)ds$, where $\bar{v}(.)$ solves

$$d\bar{v}(t) = \sigma(\bar{v}(t))\,dw(t)\,.$$

This is in full generality the only dynamic equilibrium of the system. ♢

We consider again

$$z(t) = \sum_{i=1}^N |v_i(t) - \bar{v}(t)|^2 = \frac{1}{2N}\sum_{1\leqslant i,j\leqslant N} |v_i(t) - v_j(t)|^2\,.$$

Using this time (2.3.1), Ito's formula and (2.3.3), we obtain

$$\begin{aligned} dz(t) &= -\frac{\lambda}{N}\sum_{1\leqslant i,j\leqslant N}\psi_{ij}(t)|v_i(t)-v_j(t)|^2\,dt + \sum_{i=1}^N Trace\,(\sigma(v_i(t))\sigma^*(v_i(t)))\;dt \\ &\quad - N\,Trace\left(\left(\frac{1}{N}\sum_{i=1}^N\sigma(v_i(t))\right)\left(\frac{1}{N}\sum_{i=1}^N\sigma^*(v_i(t))\right)\right)\;dt \\ &\quad + 2\left(\sum_{i=1}^N \hat{v}_i^*(t)\sigma(v_i(t))\right)\;dw(t)\,, \end{aligned}$$

(recall that a^* denotes the transposed of the vector (or the matrix) a). But since $\sum_i \hat{v}_i = 0$, we may replace $\sigma(v_i)$ by $\sigma(v_i) - \sigma(\bar{v})$ in the martingale term. After simple manipulations, it follows

$$\begin{aligned} dz(t) &= -\frac{\lambda}{N}\sum_{1\leqslant i,j\leqslant N}\psi_{ij}(t)|v_i(t)-v_j(t)|^2\,dt \qquad (2.6.5)\\ &\quad + 2\left(\sum_{i=1}^N \hat{v}_i^*(t)(\sigma(v_i(t)) - \sigma(\bar{v}(t)))\right)\;dw(t) \\ &\quad + \frac{1}{2N}\,Trace\left(\sum_{1\leqslant i,j\leqslant N}(\sigma(v_i(t))-\sigma(v_j(t)))(\sigma^*(v_i(t))-\sigma^*(v_j(t)))\right)\;dt\,, \end{aligned}$$

and

$$
\begin{aligned}
d(\ln z)(t) &= -\frac{\lambda}{N} \sum_{1\leqslant i,j\leqslant N} \psi_{ij}(t) \frac{|v_i(t)-v_j(t)|^2}{z(t)}\, dt \\
&\quad + 2 \left(\sum_{i=1}^N \frac{\hat{v}_i^*(t)(\sigma(v_i(t)) - \sigma(\bar{v}(t)))}{z(t)} \right) dw(t) \\
&\quad + \frac{1}{2N}\, Trace \left(\sum_{1\leqslant i,j\leqslant N} \frac{(\sigma(v_i(t)) - \sigma(v_j(t)))(\sigma^*(v_i(t)) - \sigma^*(v_j(t)))}{z(t)} \right) dt \\
&\quad - 2\, \frac{\left| \sum_{i=1}^N \hat{v}_i^*(t)(\sigma(v_i(t)) - \sigma(\bar{v}(t))) \right|^2}{z^2(t)}\, dt\,.
\end{aligned} \tag{2.6.6}
$$

Of course, except for the part of the drift involving the ψ_{ij}'s, these expressions are exactly the same as in subsection 2.4.1 (in a more compact form). Hence we know how to manage each term except this part of the drift. But of course if we define

$$
\psi_{min} = \inf_{i,j,v,x} \psi_{i,j}(v,x) \quad \text{and} \quad \psi_{max} = \sup_{i,j,v,x} \psi_{i,j}(v,x)\,, \tag{2.6.7}
$$

we may write, on one hand

$$
\begin{aligned}
\mathbb{E}(z(t)) &\leqslant \mathbb{E}(z(0)) - 2\,\lambda\psi_{min} \int_0^t \mathbb{E}(z(s))\, ds \\
&\quad + \frac{1}{2N} \int_0^t Trace \left(\sum_{1\leqslant i,j\leqslant N} (\sigma(v_i(s)) - \sigma(v_j(s)))(\sigma^*(v_i(s)) - \sigma^*(v_j(s))) \right) ds
\end{aligned}
$$

and on the other hand

$$
\begin{aligned}
\mathbb{E}(z(t)) &\geqslant \mathbb{E}(z(0)) - 2\,\lambda\psi_{max} \int_0^t \mathbb{E}(z(s))\, ds \\
&\quad + \frac{1}{2N} \int_0^t Trace \left(\sum_{1\leqslant i,j\leqslant N} (\sigma(v_i(s)) - \sigma(v_j(s)))(\sigma^*(v_i(s)) - \sigma^*(v_j(s))) \right) ds\,,
\end{aligned}
$$

so that we may argue exactly as in subsubsection 2.4.1 to study $\mathbb{L}^2$ flocking or non flocking. Similarly, we can get an upper bound for $\ln(z(t))$ replacing all $\psi_{ij}(t)$ by ψ_{min}, and a lower bound if σ is diagonal with linear diagonal terms as in (2.1.6), and argue exactly as in subsubsection 2.4.1 and theorem 2.2.2 (2) in order to study almost sure flocking. This yields the following two results

Theorem 2.6.8. Assume that $\psi_{ij}(t) = \psi_{ij}(v(t), x(t))$ where all the functions ψ_{ij} are local Lipschitz, non-negative and satisfy $\psi_{ij} = \psi_{ji}$, and that σ satisfies (H2-1) i.e. is globally K-Lipschitz. Define ψ_{min} and ψ_{max} as in (2.6.7). Then :

(1) if $2\lambda\psi_{min} > 4K^2d^2$ the system (2.6.2) is almost surely and $\mathbb{L}^{2,2}$ flocking. When σ is diagonal we may replace d^2 by d, if in addition the diagonal term $\sigma^{k,k}(v) = \sigma^{k,k}(v^k)$ we may replace d by 1.

(2) If σ is diagonal with linear entries, i.e. $\sigma^{k,k}(v) = D(v^k - v_e^k)$, the system is always almost surely flocking provided $D \neq 0$.

If $2\lambda\psi_{max} \leqslant D^2$, the system is not $\mathbb{L}^2$ flocking.

For (2) just remark that, when σ is diagonal with linear entries, it holds

$$Trace\left(\sum_{1\leqslant i,j\leqslant N} \frac{(\sigma(v_i(t)) - \sigma(v_j(t)))(\sigma^*(v_i(t)) - \sigma^*(v_j(t)))}{z(t)}\right) -$$
$$- \quad 2\,\frac{\left|\sum_{i=1}^N \hat{v}_i^*(t)(\sigma(v_i(t)) - \sigma(\bar{v}(t)))\right|^2}{z^2(t)} \leqslant -\,D^2\,,$$

so that we get almost sure flocking (looking at $\ln(z(t))$ as soon as $D \neq 0$.
For the $\mathbb{L}^2$ non-flocking property it is enough to look at the lower bound for $\mathbb{E}(z(t))$ since the second integral is explicit for this σ.

Remark 2.6.9. Since for positive constant communication rate the deterministic Cucker-Smale is always flocking, the introduction of noises in the previous section only introduced in some cases new ($\mathbb{L}^2$) non-flocking properties.
But here, for linear σ we obtain, whatever ψ and the initial condition are, almost sure flocking, so that this time the noise can help to (almost surely) flock, since for the classical communication rate (2.1.2), we only know that flocking holds true for some initial conditions in the deterministic case ($D = 0$) when $r > \frac{1}{2}$. ♢

Comparing swarming and flocking is also easy. Indeed, when (2.3.8) is satisfied, if the process is $\mathbb{L}^{2,2}$ swarming, the local martingale term of $\langle x_{ij}, v_{ij}\rangle$, given by

$$\int_0^t (x_i^*(s) - x_j^*(s))\ (\sigma(v_i(s)) - \sigma(v_j(s)))\ dw(s)$$

is a true $\mathbb{L}^2$ martingale once σ is globally Lipschitz (recall that swarming means boundedness for both the expectations of $|v_i - v_j|^2$ and $|x_i - x_j|^2$). In addition, it is easily seen that, if ψ is bounded, condition (2) in Lemma 2.3.11 is satisfied under the $\mathbb{L}^{2,2}$ swarming assumption (recall that this assumption includes $\sup_t \mathbb{E}(|v_i(t) - v_j(t)|^2) < +\infty$). Hence

Proposition 2.6.10. In the situation of theorem 2.6.8, assume that (2.3.8) is satisfied and that ψ is bounded. Then $\mathbb{L}^{2,2}$ swarming implies $\mathbb{L}^{2,2}$ flocking.

Of course (1) in theorem 2.6.8 is not fully satisfactory, since it is reasonable to consider models where the communication rate decays with the distance between particles as in (2.1.2). Let us consider such cases assuming that (2.3.8) is in force. Define

$$\psi_l(r) = \min_{0\leqslant u\leqslant r}\ \psi(u)\,, \tag{2.6.11}$$

and

$$T_r = \inf\{s \geqslant 0\,;\ \max_{i,j} |x_i(s) - x_j(s)| \geqslant r\}\,. \tag{2.6.12}$$

Remark 2.6.13. *Back to the deterministic model.*

Assume that $\sigma = 0$, hence consider the deterministic model. First of all $dz(t) \leqslant 0$, so that $z(t) \leqslant z(0)$ i.e. for all (i,j), $\sup_t |v_i(t) - v_j(t)| < +\infty$. Hence, according to Proposition 2.6.10 (where one can forget all the expectations and squares), the process is flocking as soon as $\sup_t |x_i(t) - x_j(t)| < +\infty$ for all (i,j). But for $t \leqslant T_r$,

$$dz(t) \leqslant -\,2\,\lambda\,\psi_l(r^2)\,z(t)$$

so that $z(t) \leqslant\, z(0)\,e^{-\,2\,\lambda\,\psi_l(r^2)\,t}$ and

$$\begin{aligned} |x_i(t) - x_j(t)| &\leqslant |x_i(0) - x_j(0)| + \int_0^t |v_i(s) - v_j(s)|\,ds \\ &\leqslant |x_i(0) - x_j(0)| + \int_0^t z^{\frac{1}{2}}(s)\,ds \\ &\leqslant |x_i(0) - x_j(0)| + \frac{z^{\frac{1}{2}}(0)}{\lambda\psi_l(r^2)}\,(1 - e^{-\,\lambda\psi_l(r^2)\,t})\,, \end{aligned}$$

i.e. for all (i,j),

$$\sup_{t\leqslant T_r} |x_i(t) - x_j(t)| \leqslant |x_i(0) - x_j(0)| + \frac{z^{\frac{1}{2}}(0)}{\lambda\psi_l(r^2)}\,. \tag{2.6.14}$$

In particular if for all (i,j), $|x_i(0) - x_j(0)| + \dfrac{z^{\frac{1}{2}}(0)}{\lambda\psi_l(r^2)} < r$ then $T_r = +\infty$ and the system is flocking.

Choosing $r_0 = \max_{i,j} |x_i(0) - x_j(0)|$ and some $C > 1$, it is thus enough that

$$z^{\frac{1}{2}}(0) \leqslant \lambda\,r_0\,(C-1)\,\psi_l(C^2\,r_0^2)\,. \tag{2.6.15}$$

We recover that if the decay to 0 of $\psi_l(r)$ is (strictly) slower than $r^{-\,\frac{1}{2}}$, the system is flocking for all initial conditions (we may let C go to infinity), while if it is faster, one has to choose the initial conditions in such a way that (2.6.15) (where one can optimize in C) is satisfied. Note that we are far from the optimal conditions, but the previous approach is completely elementary. ♢

In the stochastic case, for $t < T_r$ (which is now a random stopping time), we have (a.s.)

$$\ln(z(t)) - \ln(z(0)) \leqslant -\,(2\lambda\psi_l(r^2) - 4K^2 d^2)\,t + \ln(N(t))\,, \tag{2.6.16}$$

where

$$N_t = e^{M(t) - \frac{1}{2}\,\langle M\rangle(t)}$$

and $M_.$ is a martingale whose bracket satisfies $\langle M\rangle(t) = \int_0^t \alpha(s)\,ds$ with $|\alpha(t)| \leqslant 4K^2$. Remark that the remaining stochastic term is the logarithm of an exponential (true) martingale.

Of course, if

$$\theta(r,K) = 2\lambda\psi_l(r^2) - 4K^2 d^2 > 0\,, \tag{2.6.17}$$

(2.6.16) shows that $z(t) \to 0$ as $t \to +\infty$ almost surely on the set $\{T_r = +\infty\}$. To understand the behavior of T_r, write

$$|x_i(t\wedge T_r)-x_j(t\wedge T_r)|\leqslant |x_i(0)-x_j(0)|\,+\,z^{\frac{1}{2}}(0)\int_0^t e^{-\,(\lambda\psi_l(r^2)-2K^2d^2)\,s}\,N^{\frac{1}{2}}(s)\,\mathbf{1}_{s<T_r}\,ds\,.$$

What we have to do is to control the almost sure behavior of $N(t)$. To this end we first prove a lemma

Lemma 2.6.18. Let $M(t)$ be a martingale satisfying $\langle M\rangle(t)\leqslant Ct$. Define

$$S(a,b)=\inf\{t\geqslant 0\,,\,M(t)-b\langle M\rangle(t)\geqslant a\}\,.$$

Then

$$\mathbb{P}(S(a,b)<+\infty)\leqslant e^{-2ab}\,.$$

Proof. We know that under our assumptions, for all $\eta>0$, $e^{\eta M(t)-\frac{\eta^2}{2}\langle M\rangle(t)}$ is a martingale. Hence

$$\mathbb{E}\left(e^{\eta M(t\wedge S(a,b))-\frac{\eta^2}{2}\langle M\rangle(t\wedge S(a,b))}\right)=1\,.$$

Choose $\eta=2b$. This yields

$$\mathbb{E}\left(\mathbf{1}_{S(a,b)<+\infty}\,e^{2bM(t\wedge S(a,b))-2b^2\langle M\rangle(t\wedge S(a,b))}\right)\leqslant 1\,.$$

Using Lebesgue bounded convergence theorem we may let t go to infinity and obtain the desired result. □

Remark 2.6.19. If M is a standard Brownian motion, it is known that the inequality is an equality. ♢

We deduce from this lemma, that with probability larger than $1-e^{-2ab}$,

$$N(t)\leqslant e^{a+(b-\frac{1}{2})\langle M\rangle(t)}\leqslant e^{a+4(b-\frac{1}{2})K^2\,t}\,,$$

so that

$$|x_i(s\wedge T_r)-x_j(s\wedge T_r)|\leqslant |x_i(0)-x_j(0)|\,+\,\frac{z^{\frac{1}{2}}(0)\,e^{\frac{a}{2}}}{\lambda\psi_l(r^2)-2K^2d^2-2\,K^2(b-\frac{1}{2})}\,,$$

provided $\lambda\psi_l(r^2)>2K^2d^2+2K^2(b-\dfrac{1}{2})$.
Thus, on $\{S(a,b)=+\infty\}$ we may let s go to infinity and get that on $\{T_r<+\infty\}$,

$$r\leqslant |x_i(0)-x_j(0)|\,+\,\frac{z^{\frac{1}{2}}(0)e^{\frac{a}{2}}}{\lambda\psi_l(r^2)-2K^2d^2-2K^2(b-\frac{1}{2})}\,,\tag{2.6.20}$$

which is no more random. Hence, if (2.6.20) is not satisfied, we have

$$\mathbb{P}(T_r=+\infty\,,\,S(a,b)=+\infty)\geqslant 1-e^{-2ab}\,.$$

We have thus obtained:

Theorem 2.6.21. In the situation of theorem 2.6.8 assume in addition that (2.3.8) is in force. Let $r > 0$. Let $a, b > 0$. Assume that

- $\lambda\psi_l(r^2) > 2K^2(d^2+(b-\frac{1}{2}))$ (replace d^2 by 1 when σ is diagonal, and by 1 if the diagonal term only depends on the corresponding coordinate) where ψ_l is defined in (2.6.11),
- the initial condition satisfies, for all (i,j),

$$|x_i(0) - x_j(0)| + \frac{z^{\frac{1}{2}}(0)\, e^{\frac{a}{2}}}{\lambda\psi_l(r^2) - 2K^2(d^2 + (b - \frac{1}{2}))} < r$$

where $z(0) = \sum_{k=1}^{N} |v_k(0) - \bar{v}(0)|^2$.

Then the system (2.6.2) is flocking with a probability larger than $1 - e^{-2ab}$.

The previous result is apparently the first one dealing with "conditional flocking" (i.e. flocking for a subset of initial conditions) in a stochastic context (the results in [3] have some similarities but are actually different since they deal with approximate flocking before some stopping time).

Remark 2.6.22. Remark that when $K = 0$ corresponding to a constant σ, we may take any b going to infinity and a going to 0 so that ba goes to infinity. We thus obtain almost sure flocking under the same initial conditions than for the deterministic result (in particular for any initial condition if $r\,\psi_l(r^2) \to +\infty$ as $r \to +\infty$). This is not surprising since the microscopic variables satisfy the deterministic system of differential equations. Only the center of mass is driven by some Brownian motion.
Also notice that when ψ_l is bounded from below, we recover the almost sure statement in Theorem 2.6.8, taking $b = \frac{1}{2}$, $r = +\infty$ and finally letting a go to infinity.
Finally remark that on $\mathbb{T}_r = +\infty$, ψ_l is bounded from below by $\psi_l(r^2)$, so that according to (2.6.16) and the law of large numbers for the martingale N_t, $z(t)$ goes to 0 at an exponential (random) rate (depending on $\sup_t(N_t/t)$), or if one prefers, for any $\kappa < \lambda\psi_l(r^2) - 2K^2d^2$, there exists a random time τ_κ such that for $t > \tau_\kappa$ the decay of $z(t)$ to 0 is at least $Ce^{-\kappa(t-\tau_\kappa)}$. τ_κ is simply the last time N_t/t is bigger than $\lambda\psi_l(r^2) - 2K^2d^2 - \kappa$. $\diamondsuit$

2.6.2 Study of (2.6.1).

Let us turn to (2.6.1). Looking at the calculations (2.5.3) we see that we can mimic what we have just done with the following main modifications: replace $4K^2d^2$ by $\frac{2(N+1)}{N^2} \max_{i,j} \parallel \sigma_{i,j}^2 \parallel_\infty$ and for the variance of the martingale part $4K^2$ by $4\max_{i,j} \parallel \sigma_{i,j}^2 \parallel_\infty$. In the very particular case where for all (i,j), $\sigma_{ij} = \sigma$ for some constant σ, we can argue as in Theorem 2.5.5 (3).

Hence we only state a general result whose proof is left to the reader :

Theorem 2.6.23. Consider (2.6.1). Assume that the processes $\psi_{ij}(t) = \psi_{ij}(v(t), x(t))$ where all the functions ψ_{ij} are local Lipschitz, non-negative and satisfy $\psi_{ij} = \psi_{ji}$, that the processes $\sigma_{i,j}(t) = \sigma_{i,j}(x(t), v(t))$ where all the functions σ_{ij} are local Lipschitz, bounded and satisfy $\sigma_{ij} = \sigma_{ji}$. Define ψ_{min} and ψ_{max} as in (2.6.7). Then:

(1) for all initial state $(v(0), x(0)) \in \mathbb{L}^2$ the system admits a unique non-explosive (global) strong solution.

(2) If $\lambda\psi_{min} > \dfrac{N+1}{N^2} \max_{i,j} \parallel \sigma_{i,j}^2 \parallel_\infty$ the system is $\mathbb{L}^{2,2}$ flocking.

(3) If $\lambda\psi_{max} < \dfrac{N+1}{N^2} \min_{i,j} \inf_{t\geqslant 0} \sigma_{i,j}^2(t)$ then the system is not $\mathbb{L}^2$ flocking.

(4) If $\lambda\psi_{min} > \dfrac{N+1}{N^2} \max_{i,j} \parallel \sigma_{i,j}^2 \parallel_\infty - \dfrac{2}{N} \min_{i,j} \parallel \sigma_{i,j}^2 \parallel_\infty$ the system is almost surely flocking.

(5) If $\sigma_{ij} = \sigma$ for all pair (i,j) and some constant σ, the system is always almost surely flocking, whatever ψ is.

Assume in addition that ψ satisfies (2.3.8). Then

(6) if ψ is bounded, $\mathbb{L}^{2,2}$ swarming implies $\mathbb{L}^{2,2}$ flocking.

(7) Let $r > 0$, $a, b > 0$. Assume that

(a)
$$\lambda\psi_l(r^2) > \left(2b + \frac{N+1}{N^2}\right) \max_{i,j} \parallel \sigma_{i,j}^2 \parallel_\infty - \frac{2}{N} \min_{i,j} \inf_{t\geqslant 0} \sigma_{i,j}^2(t)$$

where ψ_l is defined in (2.6.11),

(b) the initial condition satisfies, for all (i,j),

$$|x_i(0) - x_j(0)| + \frac{z^{\frac{1}{2}}(0)\, e^{\frac{a}{2}}}{\lambda\psi_l(r^2) - \left(2b + \frac{N+1}{N^2}\right) \max_{i,j} \parallel \sigma_{i,j}^2 \parallel_\infty + \frac{2}{N} \min_{i,j} \inf_{t\geqslant 0} \sigma_{i,j}^2(t)} < r$$

where $z(0) = \displaystyle\sum_{k=1}^{N} |v_k(0) - \bar{v}(0)|^2$.

Then the system is flocking with a probability larger than $1 - e^{-2ab}$.

Once again when σ goes uniformly to 0 we recover the deterministic situation just by choosing a and b in an appropriate way.

2.6.3 A simple example with $N = 2$ for (2.6.2).

The reader certainly remarked that, when σ is constant in (2.6.2), changing $v(t)$ into $v(t) - \sigma w(t)$, the system obeys the deterministic dynamics (this is the favorite random situation for the non probabilists). Hence in this situation, conditional flocking or non flocking holds with probability 1, depending on the deterministic behavior.

It should be interesting to exhibit an example (even with two particles) where almost sure flocking holds with a strictly positive probability strictly less than 1. This seems to be a hard task. However we shall study in details simple examples to better understand what happens. For reasons we shall explain later, we shall consider the case $N = 2$ and $d = 1$.

An explicit deterministic example.

Take $N = 2$, $d = 1$ and look at the deterministic system

$$\begin{aligned} dv_1(t) &= -2\,\frac{v_1(t)-v_2(t)}{1+|x_1(t)-x_2(t)|^2}\,dt \\ dv_2(t) &= -2\,\frac{v_2(t)-v_1(t)}{1+|x_1(t)-x_2(t)|^2}\,dt\,, \end{aligned}$$

with an initial condition $v_1(0) = -v_2(0)$, $x_1(0) = -x_2(0)$. The unique solution satisfies $v_1(t) = -v_2(t)$, $x_1(t) = -x_2(t)$ and the difference $v(t) = v_1(t) - v_2(t) = 2v_1(t)$ satisfies

$$dv(t) = -\,\frac{v(t)}{1+|x(t)|^2}\,dt$$

so that

$$v(t) - v(0) = \arctan(x(0)) - \arctan(x(t))$$

and

$$x(t) - x(0) = \int_0^t (v(0) + \arctan(x(0)) - \arctan(x(s)))\,ds\,.$$

We confess that we do not know how to solve the O.D.E.

$$x'(t) = c - \arctan(x(t))\,.$$

Nevertheless we can study the qualitative behavior of the system. Indeed one can notice the following points

1. if $v(0) = 0$ the unique solution is $v(t) = 0$ and $x(t) = x(0)$.

2. It follows that if $v(0) \geqslant 0$, then the solution $v(t) \geqslant 0$ for all $t \geqslant 0$. Indeed if $v(.)$ reaches 0 then it is sticked at 0 according to the previous point.
   If one prefers, one can also write

   $$v(t) = v(0)\,e^{-\int_0^t \frac{ds}{1+|x(s)|^2}} \geqslant v(0)\,e^{-t}\,.$$

   Hence $x(.)$ is non decreasing, so that assuming that $x(0) \geqslant 0$, $\lim_{t\to+\infty} x(t) = x(\infty) \leqslant +\infty$.

Now consider a solution such that $x(0) = 0$ (for simplicity) and $v(0) \geqslant 0$. If $x(\infty) < +\infty$, since $x(t) \leqslant x(\infty)$, $(v)'(t) \leqslant -\dfrac{v(t)}{1+|x(\infty)|^2}$ so that $v(t) \to 0$ as $t \to +\infty$ at an exponential rate. Thus, $0 = v(0) - \arctan(x(\infty))$ by letting t go to infinity. Similarly if $x(\infty) = +\infty$, $\lim_{t\to+\infty} v(t) = v(0) - \dfrac{\pi}{2} \geqslant 0$ since $v(t) \geqslant 0$.
Hence

1. if $x(0) = 0$ and $0 \leqslant v(0) < \dfrac{\pi}{2}$, $x(\infty) < +\infty$ so that $v(t) \to 0$ and $x(t) \to \tan(v(0))$, the system is flocking,

2. if $x(0) = 0$ and $v(0) \geqslant \dfrac{\pi}{2}$, $x(t) \to +\infty$ and $v(t) \to v(0) - \dfrac{\pi}{2}$, so that the system is not flocking.

Back to the stochastic model.

Consider the general case with ψ satisfying (2.3.8). If we add a stochastic term such that $\sigma(-v) = -\sigma(v)$ (assuming as before that σ is K-Lipschitz) we still have $v_1(t) = -v_2(t)$, $x_1(t) = -x_2(t)$ and the difference $v(t)$ satisfies

$$dv(t) = -\psi(|x(t)|^2)\, v(t)\, dt + 2\sigma\left(\frac{v(t)}{2}\right) dw(t)\,.$$

Again the unique solution starting from $v(0) = 0$ and $x(0)$ is $v(t) = 0$, $x(t) = x(0)$, so that using the Markov property, if $v(0) \geqslant 0$, $v(t) \geqslant 0$ for all $t \geqslant 0$. For simplicity again we assume that $x(0) = 0$ and $v(0) > 0$.
Hence, up to the first time $v(.)$ reaches 0 (and then is sticked at 0) we may write

$$d(\ln(v(t))) = -\psi(|x(t)|^2)\, dt - 2\,\frac{\sigma^2(v(t)/2)}{v^2(t)}\, dt + 2\,\frac{\sigma(v(t)/2)}{v(t)}\, dw(t)\,. \tag{2.6.24}$$

Here again we have

$$v(t) = v(0)\, e^{-\int_0^t \psi(|x(s)|^2)\, ds}\, e^{N(t)-\frac{1}{2}\langle N\rangle(t)} \tag{2.6.25}$$

where $N(.)$ is a $\mathbb{L}^2$ martingale, so that $v(.)$ does not hit 0 in finite time a.s. But this representation allows us to obtain more information. Indeed lemma 2.6.18 tells us that for any $a > 0$,

$$\mathbb{P}\left(\sup_{t\geqslant 0}(N(t) - \frac{1}{2}\langle N\rangle(t)) \geqslant a\right) \leqslant e^{-a}\,.$$

Hence

$$\mathbb{P}\left(\limsup_{t\to+\infty} v(t) = +\infty\right) \leqslant \mathbb{P}\left(\limsup_{t\to+\infty}(N(t) - \frac{1}{2}\langle N\rangle(t)) = +\infty\right) = 0\,. \tag{2.6.26}$$

We know that the martingale term in (2.6.24) satisfies almost surely,

$$\lim_{t\to+\infty} \frac{1}{t}\int_0^t \frac{\sigma(v(s)/2)}{v(s)}\, dw(s) = 0\,.$$

Assume that

$$\sigma \text{ is of class } C^1 \text{ with a bounded derivative, } \sigma'(0) > 0, \text{ and } \sigma(v) > 0 \text{ for all } v > 0. \tag{2.6.27}$$

As a consequence

$$\inf_{0\leqslant v\leqslant a} \frac{\sigma(v)}{v} = \sigma_{min}(a) > 0\,.$$

Notice that (2.6.27) is satisfied in particular if $\sigma(v) \geqslant Cv$ for some $C > 0$ and all $v \geqslant 0$, which is nothing else than a simple extension of the linear case, since in this case, for $v \geqslant 0$, $Cv \leqslant \sigma(v) \leqslant Kv$.
Now for almost all given ω, $\limsup v(t)(\omega) = v_{max}(\omega) < +\infty$, so that

$$\frac{1}{t}\int_0^t \frac{\sigma^2(v(s)(\omega)/2)}{v^2(s)(\omega)}\, ds \geqslant \frac{1}{4}\,\sigma^2_{min}(v_{max}(\omega)/2)\,. \tag{2.6.28}$$

It follows that $\ln(v(t)) \to -\infty$ i.e. $v(t) \to 0$, and that the latter convergence is exponential (at least $e^{-\frac{1}{2}\sigma^2_{min}(v^2_{max}(\omega)/2)\,t}$), so that $x(t)$ is almost surely bounded, and the system is almost surely flocking. We have proved

Proposition 2.6.29. Consider (2.6.2) for $N = 2$, $d = 1$ with $\sigma(v) = -\sigma(-v)$, and assume that (2.3.8) and (2.6.27) are satisfied. Then the system is always almost surely flocking.

Remark 2.6.30. (1) In the previous proof, since we know that $v(t)$ goes to 0, using L'Hospital's and Cesaro's rules, we obtain

$$\lim_{t\to+\infty} \frac{1}{t}\int_0^t 2\,\frac{\sigma^2(v(s)/2)}{v^2(s)}\,ds \;=\; \frac{1}{2}\,(\sigma'(0))^2\,,$$

which is no more random. But one has to be careful because this limit is not uniform in ω.

(2) Of course what we have just done is to show (exponential) stability for some stochastic differential equation. Indeed, since we are in dimension 1 and the interaction term is non-positive, we know that $v(t) \leqslant u(t)$ where $u(.)$ solves

$$du(t) = \sigma(u(t))\,dw(t)\,.$$

Our proof shows that $u(t) \to 0$ at an exponential rate almost surely.

(3) Assume that σ is compactly supported, say by $[-M, M]$. Thus, (2.6.27) is not fulfilled. Take $\psi(u) = \dfrac{1}{1+u^2}$. If $v(0) > M + \dfrac{\pi}{2}$, then $v(.)$ behaves like the deterministic model (hence stays larger than M) and does not flock. Hence in (2.6.27) the behavior of $\sigma(v)/v$ near the origin is not sufficient to control flocking.

(4) However if we only skip the assumption $\sigma'(0) > 0$ in (2.6.27) and replace it by $\sigma'(0) = 0$, the previous proof shows that $\liminf_{t\to+\infty} v(t) = 0$. Indeed if not we get again a lower bound as in (2.6.28), by taking the minimum of $\sigma(v/2)/v$ on the interval $[v_{min} = \liminf v(t), v_{max} = \limsup v(t)]$.

Is it possible to get flocking while the process $u(.)$ in (2) does not flock, that is to get an example where the interaction ψ really does matter ? Here is almost one.
Choose

$$\sigma(v/2) = \frac{v^{\frac{3}{2}}}{1+v^2}\,.$$

Then if $0 \leqslant v \leqslant v_{max}$,

$$\frac{v}{(1+v^2_{max})^2} \leqslant \frac{\sigma^2(\frac{v}{2})}{v^2} \leqslant v\,.$$

Hence, since $v_{max} = \limsup v(t) < +\infty$ almost surely, if $\displaystyle\int_0^{+\infty} \frac{\sigma^2(v(s)/2)}{v^2(s)}\,ds < +\infty$ (resp. $= +\infty$), $\displaystyle\sup_t x(t) \leqslant \int_0^{+\infty} v(s)\,ds < +\infty$ (resp. $= +\infty$) almost surely, so that in all cases

$$\int_0^{+\infty} \left(\psi(|x(t)|^2) + 2\,\frac{\sigma^2(v(t)/2)}{v^2(t)}\right) dt = +\infty\,.$$

Come back to the expression (2.6.25). We know that $e^{N(t)-\frac{1}{2}\langle N\rangle(t)}$ is almost surely finite, so that if $\int_0^{+\infty} \psi(|x(t)|^2)\, dt = +\infty$, $v(t) \to 0$. In addition, lemma 2.6.18 tells us that for any $a > 0$,

$$\mathbb{P}\left(\sup_{t\geqslant 0}(N(t) - b\langle N\rangle(t)) \geqslant a\right) \leqslant e^{-2ab}\,,$$

so that for $b < \frac{1}{2}$, $e^{N(t)-b\langle N\rangle(t)}$ is almost surely finite. Thus if

$$\langle N\rangle(t) = 4\int_0^t \frac{\sigma^2(v(s)/2)}{v^2(s)}\, ds \to +\infty\,,$$

$e^{N(t)-\frac{1}{2}\langle N\rangle(t)}$ goes to 0 and so does $v(t)$ again.
But we do not know whether $\sup_t x(t)$ is always a.s. finite or not, so that we do not know whether the process is flocking or not. ♢

Finally, in the particular case $\psi(u) = \frac{1}{1+u^2}$, if $v(0) > \frac{\pi}{2}$, $x(0) = 0$, the system is not $\mathbb{L}^1$ flocking. Indeed, taking the expectation ($v(t) \geqslant 0$) we have

$$\mathbb{E}(v(t)) = v(0) - \mathbb{E}(\arctan(x(t))) \geqslant v(0) - \frac{\pi}{2} > 0\,.$$

So once again, $\mathbb{L}^p$ flocking is much more demanding.

2.7 Comments and simulations.

What kind of (temporary) conclusions can we draw after this study ?

1. All the models we have discussed in the introduction (except (2.1.9) for which we do not have a convincing interpretation) have their “reasonable” physical (or biological) interpretation and at the same time suffer potential criticism. They are only models and certainly not a description of reality.

2. Too independent noises destroy the collective behavior (without any politically correct reference).

3. Random environment depending in a certain way of the positions can also destroy the collective behavior.

4. Noises whose variances depend either linearly on the velocities or on the differences between velocities may help, at least at the almost sure level, to flock. But actually in many of these situations, the communication between individuals is simply a perturbation of a stochastic system which is already stable (though, except in a very few number of particular cases, one cannot reduce the study to the use of the theory of stability of S.D.E. as detailed in the book [10]).

5. Due to the previous item, $\mathbb{L}^2$ flocking is presumably more convincing.

We shall now illustrate our results (and the situations that are not covered by our results) with some simulations. First we shall consider the system (2.6.2)

$$dv_i(t) = -\frac{\lambda}{N}\sum_{j=1}^{N} \psi_{ij}(t)(v_i(t) - v_j(t))\,dt \;+\; \sigma(v_i(t))\,dw(t)\,.$$

In all the section we will choose

$$\psi_{ij}(t) = \psi(|x_i(t) - x_j(t)|) \quad \text{with} \quad \psi(u) = (1+u^2)^{-1}$$

in dimension $d = 2$ with $N = 9$ particles and communication intensity $\lambda = 10$.

We shall consider two basic sets of initial configurations $(x^1(0), v^1(0))$ and $(x^2(0), v^2(0))$ given by $x^1(0) = 0$

$$v^1(0) = \begin{pmatrix} -0.4 & 0.2 & -0.3 & -0.3 & -0.1 & -0.2 & 0.2 & 0.5 & 0.2 \\ \\ 0.4 & -0.1 & 0.2 & 0.5 & 0.3 & 0.1 & -0.3 & 0.2 & 0.3 \end{pmatrix}$$

$$x^2(0) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \\ -4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$v^2(0) = \begin{pmatrix} -0.3 & 2 & -0.5 & -1.5 & -0.1 & -0.2 & 1.2 & 0.5 & 1.5 \\ \\ 0.7 & -0.6 & 2.1 & 0.4 & 0.8 & 2.6 & -3.4 & -0.6 & 0.2 \end{pmatrix}$$

Define $z(0) = \sum_{k=1}^{N} |v_k(0) - \bar{v}(0)|^2$ and $M_x(0) = \max_{i,j} |x_i(0) - x_j(0)|$. Recall the discussion preceding (2.6.15) to ensure flocking starting from $(x(0), v(0))$, i.e. we want to find some $r > 0$ such that the function g defined by

$$g(r) = M_x(0) + \frac{\sqrt{z(0)}}{\lambda}\,(1 + r^2) - r$$

is negative at r. This is equivalent to the following

$$\sqrt{z(0)} < \frac{\lambda}{2}\left(\sqrt{M_x(0)^2 + \frac{1}{4}} - M_x(0)\right)$$

and it is easy to show that the first set of initial data satisfies this condition, while the second one does not (see Figure 2.1 below). In the sequel we shall use modified initial data of the form $(x^i(0), \theta v^i(0))$ for some given θ's and will plot the function g to see whether the corresponding initial data do satisfy the condition or not.

We shall now plot several simulations of the stochastic model or numerical approximations in the deterministic case. In both cases the numerical scheme is a simple explicit Euler scheme.

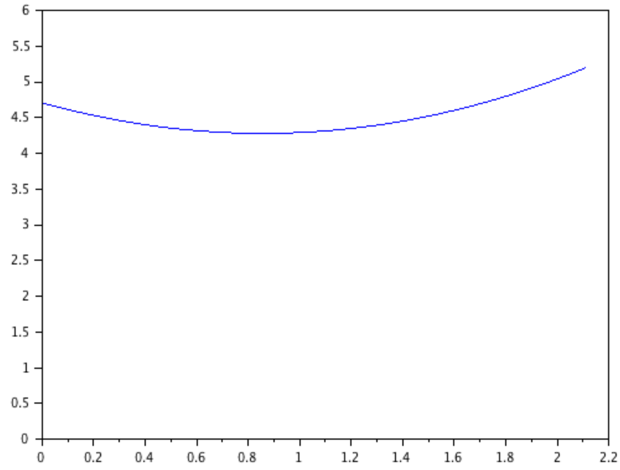

Figure 2.1: $r \mapsto g(r)$ in case $(x^2(0), v^2(0))$

On each Figure we draw the evolution in time of

$$t \mapsto \left(\sum_{i=1}^{N} |v_i(t) - \bar{v}(t)|^2 \right)^{\frac{1}{2}}$$

for both the stochastic and deterministic systems. Recall that we do not have theoretical results about the flocking property for the deterministic system once condition (2.6.15) is not satisfied.

In the next Figure 2.2 we choose $\sigma(v) = v$ and initial conditions $(x^2(0), v^2(0))$. According to Theorem 2.6.8 (2), we know that the stochastic system is almost surely flocking, but we do not know about $\mathbb{L}^2$ flocking.

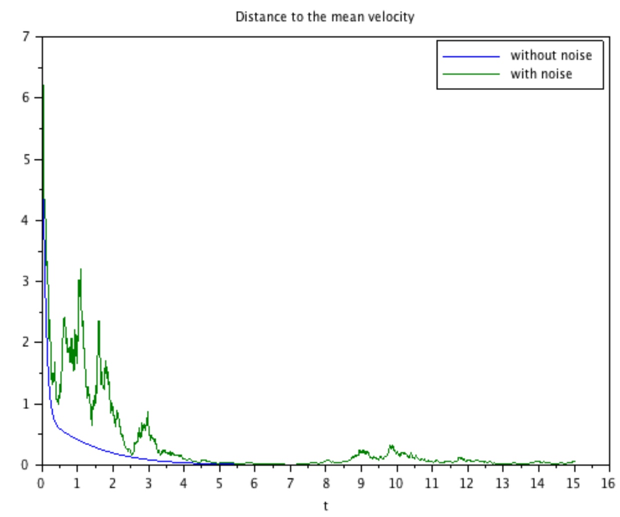

Figure 2.2: $t \mapsto \left(\sum_{i=1}^{N} |v_i(t) - \bar{v}(t)|^2 \right)^{\frac{1}{2}}$ for $\sigma(v) = v$ in case $(x^2(0), v^2(0))$

We observe that in this case the deterministic system flocks too and a reasonably quick convergence for the stochastic system.

Next, still with $\sigma(v) = v$, we change the initial configuration by choosing $(x^2(0), 5\, v^2(0))$. In this situation we see that the deterministic system does not flock anymore, while the stochastic system almost surely flocks. In Figure 2.3 we plot the evolution of the velocities on the right hand side, but also, on the left hand side, the evolution of $t \mapsto \max_{i,j} |x_i(t) - x_j(t)|$.

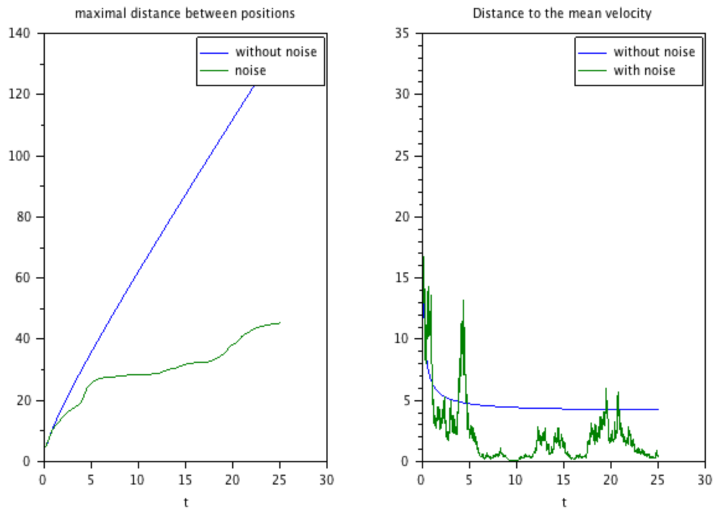

Figure 2.3: $(x^2(0), 5\, v^2(0))$, $\sigma = v$ first case.

The next two figures are obtained with the same data (be careful with the vertical scale which is not the same for each figure). The convergence to 0 in the stochastic case can be surprisingly quick (Figure 2.5), very slow (Figure 2.4 where the fluctuation size presumably indicates that there is no $\mathbb{L}^2$ flocking) or similar to the previous case (Figure 2.3 where we also observe a chaotic stabilization of the positions).

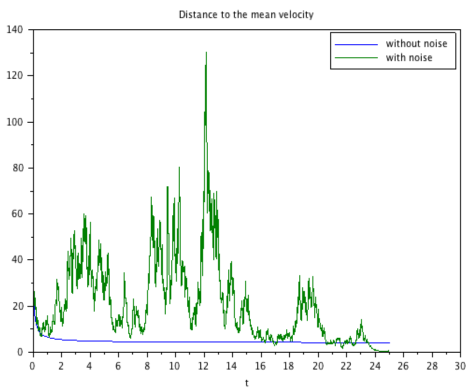

Figure 2.4: $(x^2(0), 5\, v^2(0))$, $\sigma = v$ second case.

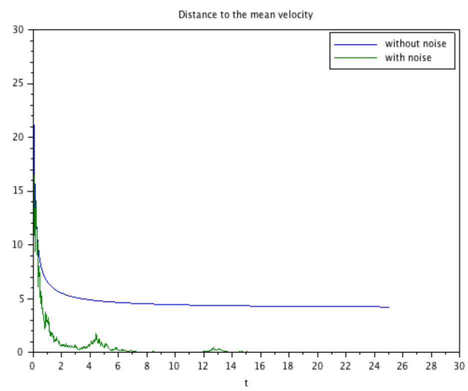

Figure 2.5: $(x^2(0), 5\, v^2(0))$, $\sigma = v$ third case.

The next situation we want to illustrate is the one of Theorem 2.6.21. To this end we choose $\sigma(v^i) = 1 + \sin v^i$ for each of the two coordinates v^i of v, hence a diagonal σ. Since σ is 1-

Lipschitz, we choose $b = 1/2$ and $a = \ln 2$ so that if the initial conditions satisfy the assumption in Theorem 2.6.21, the latter tells us that the stochastic system flocks with a probability larger than or equal to $\frac{1}{2}$.

To fulfill this assumption we choose this time $(x^1(0), 0.1\ v^1(0))$ as initial conditions. We thus know that the deterministic system is flocking. The next figure 2.6 plots the condition showing that some r can be found, while figure 2.7 presents an example of simulation. Actually in this case we were not able to obtain a non-flocking stochastic simulation, showing that, for sure, our result is far from optimal.

To observe something interesting we have to change the initial conditions and thus take $(x^2(0), 3\ v^2(0))$. If we still have the flocking property for the deterministic model, we have observed (as the two examples show) various cases in the stochastic setting, with or without flocking, indicating that flocking may occur with some probability strictly larger than 0 and strictly smaller than 1.

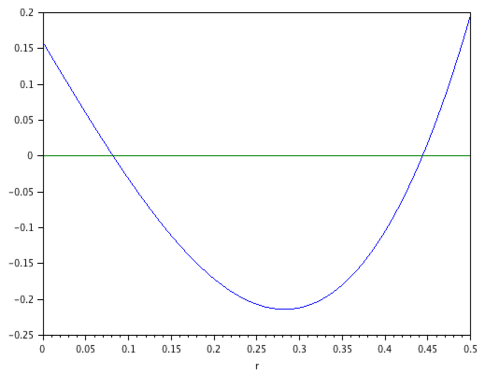

Figure 2.6: $(x^1(0), 0.1\ v^1(0))$ condition

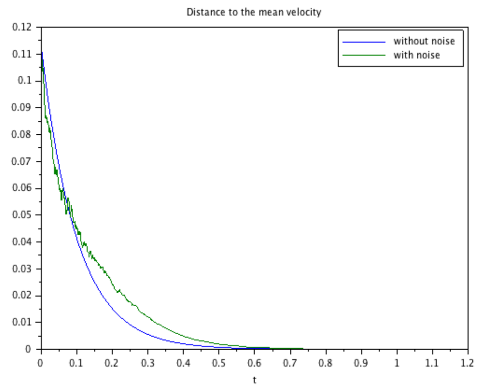

Figure 2.7: $(x^1(0), 0.1\ v^1(0))$ and $\sigma = 1 + \sin(v)$.

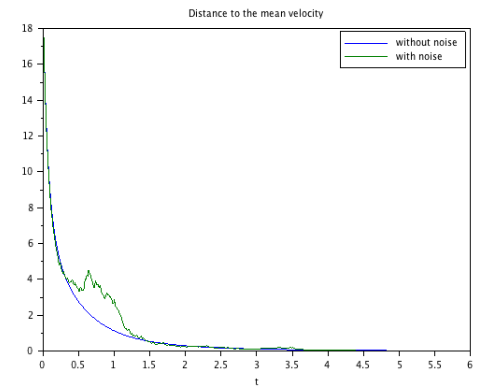

Figure 2.8: $(x^2(0), 3\ v^2(0))$, $\sigma(v) = 1 + \sin(v)$ with flock

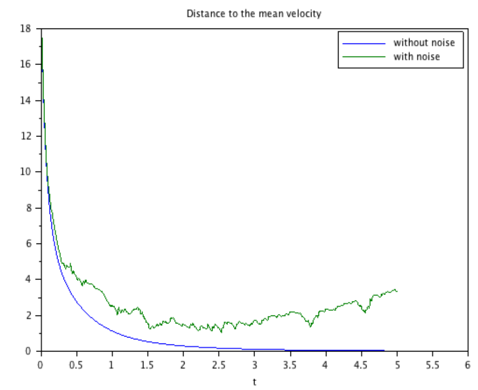

Figure 2.9: $(x^2(0), 3\ v^2(0))$, $\sigma(v) = 1 + \sin(v)$ with no flock

Finally, we show some simulations when σ is a function of x and no more of v. As we have seen, this situation is completely unclear, even for a constant communication rate. This chaotic behavior is illustrated by the final three pictures where, as before, we have drawn the behavior of the positions on the left hand side and of the velocities on the right hand side.

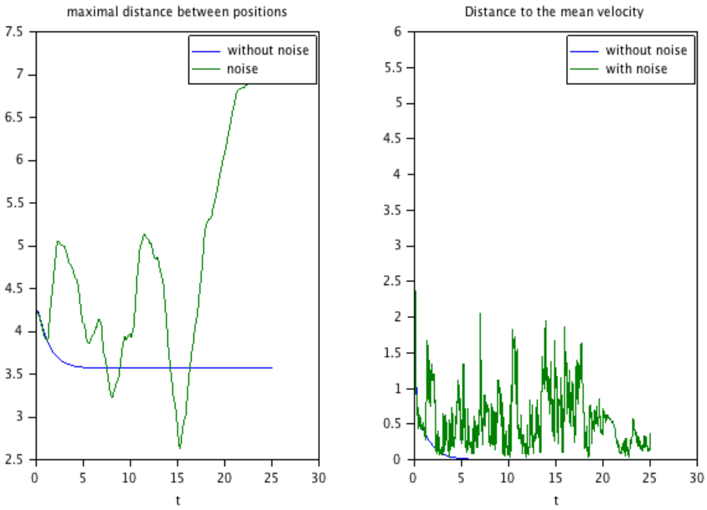

Figure 2.10: $(x^2(0), v^2(0))$, $\sigma = 1 + \sin(x)$ first case

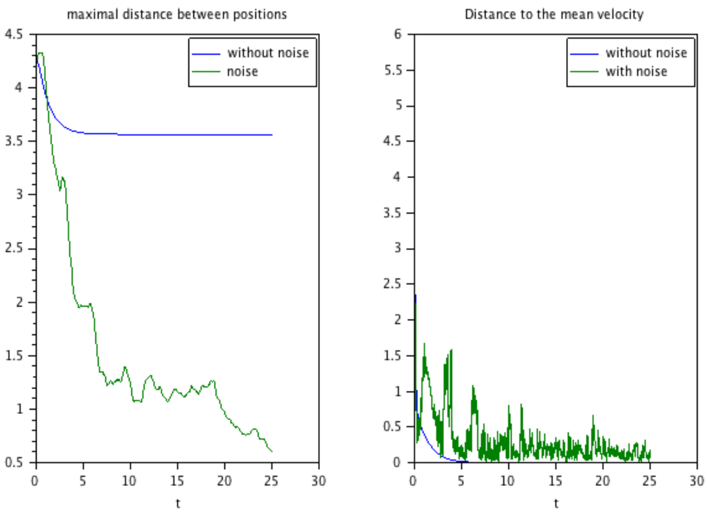

Figure 2.11: $(x^2(0), v^2(0))$, $\sigma = 1 + \sin(x)$ with flock

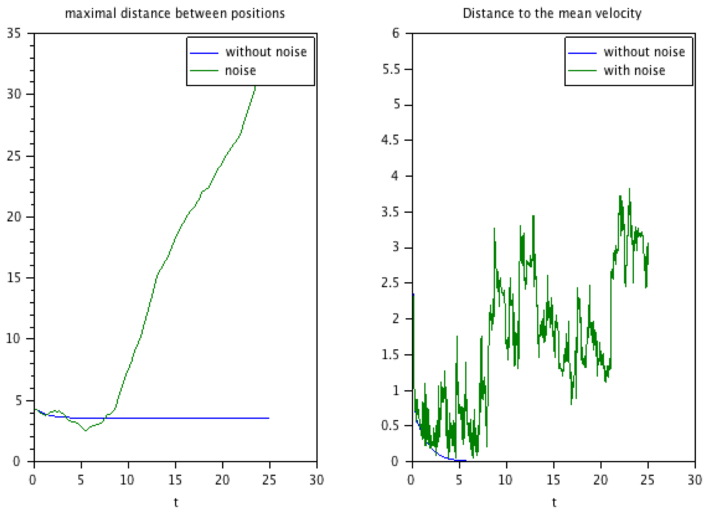

Figure 2.12: $(x^2(0), v^2(0))$, $\sigma = 1 + \sin(x)$ with no flock

Bibliography

[1] S. M. Ahn and S-Y Ha. Stochastic flocking dynamics of the Cucker-Smale model with multiplicative white noises. *J. Math. Physics*, 51:103301, 2010.

[2] F. Bolley, J. A. Cañizo, and J. A. Carrillo. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Math. Models Methods Appl. Sci.*, 11:2179–2210, 2011.

[3] F. Cucker and E. Mordecki. Flocking in noisy environments. *J. Math. Pures Appl.*, 89:278–296, 2008.

[4] F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Trans. Automat. Control*, 52:852–862, 2007.

[5] F. Cucker and S. Smale. On the mathematics of emergence. *Jpn. J. Math.*, 2:197–227, 2007.

[6] R. Erban, J. Haskovec, and Y. Sun. On Cucker-Smale model with noise and delay. Available on ArXiv 1507.04432 [DS]., 2015.

[7] S-Y. Ha, K. Lee, and D. Levy. Emergence of time-asymptotic flocking in a stochastic Cucker-Smale system. *Comm. Math. Sci.*, 7:453–469, 2009.

[8] S-Y. Ha and J-G. Liu. A simple proof of the Cucker-Smale flocking dynamics and mean-field limit. *Comm. Math. Sci.*, 7:297–325, 2009.

[9] S-Y. Ha and E. Tadmor. From particle to kinetic and hydrodynamic descriptions of flocking. *Kinet. Relat. Models*, 1:415–435, 2008.

[10] X. Mao. *Stochastic differential equations and applications.* Horwood Publishing Limited, 2007.

[11] S. Motsch and E. Tadmor. A new model for self-organized dynamics and its flocking behaviour. *J. Stat. Phys.*, 144:923–947, 2011.

[12] L. Pédèches. Asymptotic properties of some stochastic Cucker-Smale dynamics. Preprint., 2016.

[13] Y. Sun and W. Lin. A positive role of multiplicative noise on the emergence of flocking in a stochastic Cucker-Smale system. *Chaos*, 25:083118, 2015.

[14] T. V. Ton, N. T. H. Linh, and A. Yagi. Flocking and non flocking in a stochastic Cucker-Smale model. *Anal. Appl.*, 12(1):63–73, 2014.

[15] T. Vicsek and A. Zafeiris. Collective motion. *Physics Reports*, 517(3):71–140, 2012.

Chapter 3

Exponential ergodicity for a class of non-Markovian stochastic processes

This chapter was submitted, under the same title, in March 2017.

Abstract : We prove the exponential ergodicity of a class of solutions of stochastic differential equations with finite delay. This is done, in this non-Markovian setting, using the cluster expansion method, inspired from previous works. As a consequence, the results hold for small perturbations of ergodic diffusions.

Keywords : cluster expansion, SDE with delay, long-time behaviour.

Contents

3.1 Introduction

The aim of this paper is to prove that exponential ergodicity holds for a class of solutions of stochastic differential equations with finite delay and non regular drift.

We will consider $\mathbb{R}^d$-valued stochastic differential equations of the form :

$$dX_t = \Big(g(X_t) + \beta\, b((X)_{t-t_0}^t)\Big)\, dt + dW_t,$$

where $(X)_s^t := \{X_u | u \in [s,t]\}$ is the path of the process X between times s and t, and t_0 is a fixed positive number. We will make certain assumptions on the underlying semi-group of the reference process, weak solution of

$$dX_t = g(X_t)\, dt + \, dW_t$$

The additional drift term b, will only be required to be time-local, measurable and bounded by 1.

Our interest in those equations comes from possible applications for stochastic Cucker-Smale type models in $(\mathbb{R}^d)^N$ - such as the one presented by Ha, Lee and Levy in [10]. It is a N-particle mean-field system in $\mathbb{R}^d$, whose velocity $v(t) = (v_1(t), ..., v_N(t))$ satisfies, for all $t \geqslant 0$, the stochastic differential equation

$$dv(t) = -\frac{\lambda}{N}\, F((v)_0^t)\, dt + dW(t)$$

where for all $i \in \{1, ..., N\}$, $F_i((v)_0^t) = \sum_{j=1}^{N} \psi((v_j)_0^t, (v_i)_0^t)\ (v_i(t) - v_j(t))$ and W is a standard dN-dimensional Brownian motion. The function ψ, supposed to be non-negative and symmetric, is called communication rate and quantifies the interaction between each pair of particles.

Various results about the existence of invariant probability measures for stochastic differential equations with delay can be found in the literature, going back to the paper of [12], where is proven that, when the drift and diffusion coefficients are continuous, there exist stationary solutions for delayed processes. Since then, one can mention, among many others, the papers by [21],[3], or the book of [7], especially Chapter 10. General results on stochastic differential delay equations up to 2003 are gathered in a survey by Ivanov, Kazmerchuk and Swishchuk ([13]). However, they are mainly valid under strong regularity assumptions on the coefficients, despite the fact that non-regular coefficients appear in various fields, such as finance (see for instance [2], about pricing options) or physics (with bistable systems, in [25]) ; stability of non-regular processes is also a fixture in [18].

One notable exception is [24], where the author considers the equation

$$dx(t) = F((x)_{t-1}^t)\, dt + dW_t$$

for a function F measurable and locally bounded. Assuming the existence and uniqueness of a weak solution, and a restrictive recurrence condition (holding if some condition on certain Lyapunov functionals are met), the existence of – and the convergence in total variation distance towards – an invariant probability measure is proven, using the strong Markov property satisfied by $((x)_{t-1}^t)_{t\in\mathbb{R}_+}$. Nothing is said, however, about the rate of convergence for such processes. The true novelty of our work is the exponential rate of the ergodicity.

As we are dealing with non-Markovian processes, most standard methods of stochastic analysis are not available. Thus, our main tool here will be the so-called cluster expansion method, mainly used in statistical mechanics, in particular in Gibbs field theory. As a consequence, our results will hold for irregular but small (albeit not insignificant) perturbations of the reference process. Technical results for the adaptation of the cluster expansion methods to Gibbs random fields can be found in the book by [17]. Subsequent papers have implemented those methods for stochastic processes, for example, interacting diffusions systems or one-dimensional non-Markovian diffusions. It was done in [11], and, more recently, amongst others in [8], [9] or [19].

Our main result is the exponential ergodicity of the process. Moreover, with the same technique, one obtains that the decay of correlations is exponentially quick. It follows that a central limit theorem can be derived from the mixing properties implied by this inequality.

Contrary to what was done in [8], we do not require for the semi-group associated with the reference process to be ultracontractive, but we only need some strong form of hypercontractivity. One instance of a well-known process which is not ultracontractive but verifies our assumptions is the Ornstein-Uhlenbeck process. We will present some explicit results in this particular setting. Actually, the stochastic Cucker-Smale model can be seen as a mean-field perturbation of the Ornstein-Uhlenbeck process, and this led to this work. The lack of ultracontractive bounds for the underlying reference process introduces several new technical difficulties. We will therefore present a detailed proof for the cluster estimates. The use of these estimates to get to the final main theorem follows the lines of, for instance, what was done in [8] and [19] and we will go rather quickly over this part.

We start by introducing our framework, the objects we will encounter and the assumptions that will be needed, before giving our main result. Then, we obtain, in section 3, a cluster representation for the partition function defined in the first part. In section 4, we study the cluster estimates and show that they tend to 0 when β does. In section 5, we conclude the proof and present a few consequences of our convergence theorem. Finally, in section 6, we explicitly compute some of the bounds in the Ornstein-Uhlenbeck setting.

3.2 Framework and main theorem

We introduce here the process which will serve as reference in our work : a stochastic process sufficiently regular to be exponentially ergodic with respect to its invariant (and even reversible) probability measure. We present all the assumptions that will be necessary to extend this ergodicity to small perturbations of this process.

3.2.1 The reference process

First, we introduce the framework in which we are considering such a stochastic process :

- $\Omega = \mathcal{C}(\mathbb{R}, \mathbb{R}^d)$ shall be the canonical continuous $\mathbb{R}^d$-valued path space, for some $d \geqslant 1$, and $\mathcal{F}$ the canonical Borel σ-field on Ω. $(X_t)_{t\in\mathbb{R}}$ shall be, as usual, the canonical process.
- $\mathbb{W}$ shall be the Wiener measure on $(\Omega, \mathcal{F})$, the law of a standard d-dimensional Brownian motion $(W_t)_{t\in\mathbb{R}}$.

We consider the following stochastic differential equation, for any $t \in \mathbb{R}$,

$$dX_t = g(X_t)\, dt + dW_t \tag{3.2.1}$$

with $g : \mathbb{R}^d \to \mathbb{R}^d$ a smooth function (say $\mathcal{C}^k$, for a certain $k \geqslant 2$) and (W_t) a standard d-dimensional Brownian motion. We suppose that there exists a reversible probability measure, μ, for this process. We will soon give conditions to ensure that (3.2.1) has a unique stationary weak solution.
Let L be its associated infinitesimal generator, defined by

$$L = \frac{1}{2} \sum_{i,j} \partial^2_{ij} + \sum_i g_i\, \partial_i.$$

L is uniformly elliptic, and, as μ is reversible, symmetric in $L^2(\mu)$: for all f and g smooth enough, $\displaystyle\int f\, Lg\, d\mu = \int g\, Lf\, d\mu$. It is known that μ is then absolutely continuous with respect to the Lebesgue measure, with a positive density. Thus, μ is of the form $d\mu(x) = C\, e^{-V(x)}\, dx$; in addition, $V : \mathbb{R}^d \to \mathbb{R}$ is smooth (at least $\mathcal{C}^k$).
According to Kolmogorov's characterization of reversible diffusions, in [15], this even implies that g can be written as a gradient function of a potential function V : $g = -\frac{1}{2}\, \nabla V$. Thus, equation (3.2.1) becomes :

$$dX_t = -\frac{1}{2}\, \nabla V(X_t)\, dt + \, dW_t \tag{3.2.2}$$

The probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ shall denote the weak stationary solution of (3.2.2), with marginal law the invariant probability measure μ.

To ensure that the equation (3.2.2) indeed admits a stationary weak solution on $\mathbb{R}$, and in particular, that there is non-explosion in finite time, we further assume (see e.g. [23]) that one of the two following assertions on the potential V is true :

$$(1) \qquad V(x) \underset{|x|\to\infty}{\longrightarrow} +\infty \quad \text{and} \quad |\nabla V|^2 - \Delta V \quad \text{is bounded from below}$$

$$(2) \quad \text{There exist } a, b \in \mathbb{R} \text{ such that, for all } x, \quad x^* \, \nabla V(x) \geqslant -a\, |x|^2 - b$$

where x^* is the transpose of x.

In this case, there even exists a unique strong solution (Theorem 2.2.19 in [23]). The semi-group (P_t) admits a smooth transition density with respect to μ, denoted by $p(t, x, y)$. As the probability measure μ is reversible, $p(t, ., .)$ is symmetric :

$$\forall\, t, x, y, \quad p(t, x, y) = p(t, y, x).$$

We now introduce two assumptions which will be essential in the following :

- (H1) : μ satisfies a Poincaré inequality : there exists a constant C_P such that for all smooth functions f in $L^2(\mu)$,

$$\|\, f - \int f \; d\mu \,\|^2_{L^2(\mu)} \leqslant C_P \int |\nabla f|^2 \; d\mu$$

- (H2) : There exists $\delta \geqslant 0$ such that

$$\sup_{t\geqslant\delta} \|p(t, ., .)\|_{L^8(\mu\otimes\mu)} < \infty$$

Remark 12. It is well-known (see [1] for example) that hypothesis (H1) is equivalent to the exponential convergence of the semi-group towards μ, i.e. there exists a constant C_S such that for all $f \in L^2(\mu)$, for all $t \geqslant 0$,

$$\|\, P_t f - \int f \; d\mu \,\|_{L^2(\mu)} \leqslant e^{-C_S t} \|\, f - \int f \; d\mu \,\|_{L^2(\mu)}$$

and, moreover, $C_S = 1/C_P$ (see [6] for a more general statement).

In particular, if $\int f d\mu = 0$, then (H1) implies that for every $t \geqslant 0$,

$$\|P_t f\|_{L^2(\mu)} \leqslant e^{-t/C_P} \|f\|_{L^2(\mu)} \tag{3.2.3}$$

Remark 13. Using Cauchy-Schwarz's and Jensen's inequalities,

$$\|P_\delta f\|_{L^8(\mu)}^8 = \int \left(\int p(\delta,x,y) f(y)\, \mu(dy) \right)^8 \mu(dx)$$
$$\leqslant \int \left(\int |f(y)|^2\, \mu(dy) \right)^4 \left(\int p(\delta,x,y)^2\, \mu(dy) \right)^4 \mu(dx)$$
$$\leqslant \|f\|_{L^2(\mu)}^8 \int\int p(\delta,x,y)^8\, \mu(dy)\, \mu(dx)$$

This means that :

$$\|P_\delta f\|_{L^8(\mu)} \leqslant \|p(\delta,.,.)\|_{L^8(\mu\otimes\mu)}\, \|f\|_{L^2(\mu)}. \tag{3.2.4}$$

Thus, if (H2) is satisfied, $P_\delta : L^2(\mu) \to L^8(\mu)$ is a bounded operator.
If, in addition, (H1) is satisfied, for every $k \geqslant 2$, $P_t : L^2(\mu) \to L^k(\mu)$ is bounded by 1 when t is large enough.

Example 1. (H1) is satisfied, for instance, if V is uniformly convex outside of a compact set, that is if the Hessian matrix of V is a non-degenerate quadratic form outside of a compact set. It is, however, difficult to obtain a generic condition on the potential V for hypothesis (H2) to hold ; one can look at Section 3 of [5] to understand the underlying difficulties : in particular, condition (A4), introduced at the beginning of Section 2 in [5] is fairly close to our hypothesis (H2). In the special case an Ornstein-Uhlenbeck reference process, discussed in section 3.6, (H2) is proven thanks to the known explicit expression of the density function p.

We now prove a proposition, taking into account hypotheses (H1) and (H2) and yielding the assumption we will use in practice, rather than (H1) itself.

Proposition 22. *Under hypotheses (H1) and (H2), for $t \geqslant 2\delta$,*

$$\|p(t,.,.) - 1\|_{L^8(\mu\otimes\mu)} \leqslant \gamma_\delta(t) \left(\|p(\delta,.,.)\|_{L^8(\mu\otimes\mu)} \vee 1 \right)$$

where $\gamma_\delta(t) = 2M_\delta\, e^{-(t-2\delta)/C_P}$ with $M_\delta = \sup_{a\geqslant\delta} \|p(a,.,.)\|_{L^8(\mu\otimes\mu)} \vee 1$

In particular, $\gamma_\delta(t)$ goes to 0 exponentially fast when t goes to infinity.

Proof. The following lemma is essential for the proof of the proposition.

Lemma 6. Set $\delta > 0$. Suppose that (H1) holds true.
Then for all smooth f such that $\int f d\mu = 0$,

$$\forall t \geqslant \delta, \quad \|P_t f\|_{L^8(\mu)} \leqslant e^{-(t-\delta)/C_P}\, \|p(\delta,.,.)\|_{L^8(\mu\otimes\mu)}\, \|f\|_{L^8(\mu)}$$

Remark 14. It is possible for both sides of the above inequality to be infinite.

Proof. Let t be a positive number with $t \geqslant \delta$ and f a smooth function such that $\int f d\mu = 0$.

For any $t \geqslant \delta$, thanks to the inequality (3.2.4) proven in remark 13,

$$\|P_t f\|_{L^8(\mu)} = \|P_\delta(P_{t-\delta} f)\|_{L^8(\mu)} \leqslant \|p(\delta,.,.)\|_{L^8(\mu\otimes\mu)} \, \|P_{t-\delta} f\|_{L^2(\mu)}$$

As (H1) is supposed to be satisfied, so is (3.2.3) ; hence the conclusion of this proof :

$$\|P_t f\|_{L^8(\mu)} \leqslant e^{-(t-\delta)/C_P} \, \|p(\delta,.,.)\|_{L^8(\mu\otimes\mu)} \, \|f\|_{L^2(\mu)} \leqslant e^{-(t-\delta)/C_P} \, \|p(\delta,.,.)\|_{L^8(\mu\otimes\mu)} \, \|f\|_{L^8(\mu)}$$

□

We prove the proposition and start by expressing $p(t,x,y)-1$ using the semi-group :

$$\begin{aligned} p(t,x,y) - 1 = \int (p(t-\delta,x,z)\, p(\delta,z,y) - 1)\, \mu(dz) \\ = P_{t-\delta}(p(\delta,.,y))(x) - 1 = P_{t-\delta}(p(\delta,.,y)-1)(x) \end{aligned}$$

Thus, applying Lemma 6 for $f = p(\delta,.,y)-1$ at time $t-\delta$ ($\geqslant \delta$ as $t \geqslant 2\delta$),

$$\begin{aligned} \int (p(t,x,y)-1)^8 \mu(dx) = \int P^8_{t-\delta}(p(\delta,.,y)-1)(x)\, \mu(dx) \\ \leqslant e^{-8(t-2\delta)/C_P} \, \|p(\delta,.,.)\|^8_{L^8(\mu\otimes\mu)} \, \|p(\delta,.,y)-1\|^8_{L^8(\mu)} \\ \leqslant e^{-8(t-2\delta)/C_P} \, 2^8 \, \|p(\delta,.,.)\|^8_{L^8(\mu\otimes\mu)} \, \left(\|p(\delta,.,y)\|^8_{L^8(\mu)} \vee 1\right), \end{aligned}$$

which leads to :

$$\int (p(t,x,y)-1)^8 \mu(dx)\mu(dy) \leqslant e^{-8(t-2\delta)/C_P} \, 2^8 \, \|p(\delta,.,.)\|^8_{L^8(\mu\otimes\mu)} \, \left(\|p(\delta,.,.)\|^8_{L^8(\mu\otimes\mu)} \vee 1\right)$$

Hypothesis (H2) ensures that $\|p(u,.,.)\|_{L^8(\mu\otimes\mu)}$ is bounded uniformly in u for $u \geqslant \delta$, hence the result. □

Remark 15. As can be seen in the proof, M_δ is a priori not the optimal bound (although it corresponds with $\|p(\delta,.,.)\|_{L^8(\mu\otimes\mu)}$ as will be seen in the Ornstein-Uhlenbeck example in section 3.6) but will be good enough for our needs (and will simplify later computations).

3.2.2 The perturbed stochastic differential equation

We turn our attention to the stochastic differential equation with finite delay t_0, for all $t \in \mathbb{R}$,

$$dX_t = \left(-\frac{1}{2}\, \nabla V(X_t) + \beta\, b((X)^t_{t-t_0})\right) dt + \, dW_t \tag{3.2.5}$$

where the potential V and the Brownian motion (W_t) are as previously defined, β is a positive constant, which shall be small enough for the result to hold.

The perturbation drift, b shall satisfy the assumption (H3) detailed below :

(H3) : $b : \Omega \to \mathbb{R}^d$ is a measurable function, bounded by 1, and local, in the sense that there exists a delay $t_0 > 0$ such that, for any $u \in \Omega$, $b(u) = b((u)^t_{t-t_0})$.

Example 2. We give here a few examples for perturbation drifts b satisfying (H3) :

- we can consider b of the form $b((u)_{t-t_0}^t) = \int_{t-t_0}^t f(u_t, u_s)\, ds$ for any trajectory $u \in \Omega$ with f bounded by 1 and measurable ; for instance, $f(x,y) = \text{sign}(x-y)$ or $f(x,y) = \mathbb{1}_{y \in A}$ with A a subset of $\mathbb{R}^d$ (thus obtaining an occupation time);

- we may have a dependence in the past depending on a single time, of the form $b((u)_{t-t_0}^t) = g(u_{t-t_0})$ for a certain function g measurable and bounded by 1, but not necessarily continuous ;

- one can also consider a drift function with jumps, such as $b((u)_{t-t_0}^t) = \mathbb{1}_{(u)_{t-t_0}^t \in A}$ with A a subset of $\mathcal{C}([-t_0, 0], \mathbb{R}^d)$.

One of the main advantages of our method is that we only require from b that it satisfies (H3), without any stronger condition on its regularity.

Recall that the probability measure Q on Ω is said to be a weak solution of the stochastic differential system (3.2.5) if the process

$$\left(X_t - \int_0^t \left(-\frac{1}{2}\, \nabla V(X_s) + \beta\, b((X)_{s-t_0}^s) \right) ds \right)$$

is a Q-Brownian motion.

3.2.3 The main result

Our main theorem is the following convergence result for the stochastic differential equation with delay, for $t \in \mathbb{R}$,

$$dX_t = \left(-\frac{1}{2}\, \nabla V(X_t) + \beta\, b((X)_{t-t_0}^t) \right) dt + dW_t \tag{3.2.5}$$

considered as a perturbation of the reference process

$$dX_t = -\frac{1}{2}\, \nabla V(X_t)\, dt + \, dW_t. \tag{3.2.2}$$

Theorem 14. *Assume that the assumptions (H1) and (H2) are satisfied by the reference stochastic differential equation (3.2.2). Assume also that the perturbation drift b of equation (3.2.5) verifies (H3).*

Then, for β small enough,

(i) The stochastic differential equation (3.2.5) admits a unique weak stationary solution Q, and thus a unique invariant probability measure ν.

(ii) There is exponential ergodicity : there exist $\theta > 0$ and $C : \mathbb{R}^d \to \mathbb{R}_+$ such that for $|t|$ and $|t'|$ large enough, for every $x \in \mathbb{R}^d$, for every bounded measurable function f,

$$|\mathbb{E}_Q[f(X_t)|X_0 = x] - \mathbb{E}_Q[f(X_{t'})|X_0 = x]| \leqslant C(x)\, e^{-\theta\, |t-t'|}.$$

(iii) The decay of correlations is exponentially quick : there exist two positive constants θ_1 and θ_2,, such that for $|t|$ and $|t'|$ large enough, for all f and g measurable and bounded by 1, it holds :

$$|\mathbb{E}_Q[f(X_t)\, g(X_{t'})] - \mathbb{E}_Q[f(X_t)]\, \mathbb{E}_Q[g(X_{t'})]| \leqslant \theta_1\, e^{-\theta_2\, |t-t'|}.$$

The rest of the paper is devoted to the proof and the consequences of this theorem.

3.3 Approximation on finite-time windows and cluster representation

The main idea behind the proof is to build approximations on finite-time windows that will converge uniformly towards what will be the weak stationary solution of (3.2.5) ; the properties of these approximations will then be inherited by this limit.

3.3.1 Approximations

We set the following notations :

- a, a fixed positive number, destined to become quite large ;
- for every j in $\mathbb{Z}$, $I_j = [ja, (j+1)a]$;
- for every N in $\mathbb{N}^*$, $I(N) = [-Na, Na] = \bigcup_{j=-N}^{N-1} I_j$;
- for every u in Ω, $u^{(N)}(t) = u(Na)$ if $t \geqslant Na$, $u^{(N)}(t) = u(t)$ if $-Na \leqslant t \leqslant Na$, and $u^{(N)}(t) = u(-Na)$ if $t \leqslant -Na$. That is : $u^{(N)}$ is equal to u frozen outside of the interval $I(N)$.

Using Girsanov theorem (see e.g. [16]), we can show that the restriction to any finite time interval I of the law of the perturbed process is absolutely continuous with respect to the law of the reference process, $\mathbb{P}$, and that its density is of the form $\exp(-H_I(u))du$ where the associated Hamiltonian H_I is defined by

$$H_I(u) = -\int_I \beta\, b((u)^t_{t-t_0})^*\, dW_t + \frac{\beta^2}{2} \int_I |b((u)^t_{t-t_0})|^2\, dt \tag{3.3.1}$$

for every trajectory u in the path space Ω. We will denote $H_N = H_{I(N)}$.

To obtain the theorem, our main objective is to prove the convergence of the sequence of probability measures $(Q_N)_{N\in\mathbb{N}^*}$, defined on Ω by

$$Q_N(du) = \frac{1}{Z_N} \exp\left(-H_N(u^{(N)})\right) \mathbb{P}(du), \tag{3.3.2}$$

towards a weak solution of the equation (3.2.5) that will be time stationary. From this point, classical results of Gibbs theory shall lead to Theorem 14.

Remark 16. Under Q_N, the canonical process (X_t) is a weak solution of the stochastic differential system (3.2.5), for $t \in I(N)$, but not a stationary one.

3.3.2 The partition function and its cluster representation

The renormalization constant in (3.3.2), also called partition function, is given by

$$Z_N = \int_\Omega \exp\left(-H_N(u^{(N)})\right) \mathbb{P}(du)$$

The aim of our next section will be to expand Z_N with respect to β uniformly in N. Note that, contrary to Q_N, Z_N will not converge when N goes to infinity.

The bulk of the proof shall then be to control the different terms involved in this series expansion, to show that they are smaller than a certain function of β that vanishes when β goes to 0.

The cluster expansion method, very useful in statistical mechanics, shall then lead us first to the convergence of the sequence $(Q_N)_N$ towards a weak stationary solution Q of equation (3.2.5) and the existence of a invariant probability measure, second to the exponential ergodicity and Theorem 14.

First, however, we aim to expand the partition function into clusters, that is to obtain an expression of Z_N of the form :

$$Z_N = 1 + \sum_\tau \prod_i \Gamma_{\tau_i}$$

with the meaning and nature of each of τ, i and Γ_{τ_i} to be determined.

We start by conditioning the reference probability $\mathbb{P}$ on Ω with respect to the values of its marginals at times $-Na$, $-(N-1)a$, ..., 0, a, ..., Na :

$$
\begin{aligned}
Z_N = & \int_\Omega \exp\left(-H_N(u^{(N)})\right) \mathbb{P}(du) \\
& = \int_{\mathbb{R}^{(2N+1)d}} \int_\Omega \exp\left(-H_N(u^{(N)})\right) \mathbb{P}(du | X_{ja} = y_j, j = -N, ..., N) \\
& \qquad \otimes \mathbb{P}_{X_{-Na}}(dy_{-N}) \bigotimes_{j=-N}^{N-1} \mathbb{P}_{X_{(j+1)a}}(dy_{j+1} | X_{ja} = y_j)
\end{aligned}
$$

Let $\mathbb{P}_I^{a,b}$ denote the law of the stochastic bridge over I obtained by conditioning $\mathbb{P}$ so that $X_{\inf I} = a$ and $X_{\sup I} = b$. Then, on the interval $I(N)$,

$$
\mathbb{P}(\,.\,| X_{ja} = y_j, j = -N, ..., N) = \bigotimes_{j=-N}^{N-1} \mathbb{P}_{I_j}^{y_j, y_{j+1}}(\,.\,). \tag{3.3.3}
$$

Recall that by definition of the transition density p,

$$
\mathbb{P}_{X_{(j+1)a}}(dy_{j+1} | X_{ja} = y_j) = p(a, y_j, y_{j+1})\, \mu(dy_{j+1}) \tag{3.3.4}
$$

Combining (3.3.3) and (3.3.4), one obtains

$$
Z_N = \int_{\mathbb{R}^{(2N+1)d}} \int_\Omega \exp\left(-H_N(u^{(N)})\right) \bigotimes_{j=-N}^{N-1} \mathbb{P}_{I_j}^{y_j, y_{j+1}}(du) \bigotimes_{j=-N}^{N-1} p(a, y_j, y_{j+1}) \bigotimes_{j=-N}^{N} \mu(dy_j)
$$

Next, we re-order the terms in a convenient way :

$$
Z_N = \int_{\mathbb{R}^{(2N+1)d}} \int_\Omega \prod_{j=-N}^{N-1} \Big(\exp\left(-H_{I_j}(u^{(N)})\right) p(a, y_j, y_{j+1}) \Big) \bigotimes_{j=-N}^{N-1} \mathbb{P}_{I_j}^{y_j, y_{j+1}}(du) \bigotimes_{j=-N}^{N} \mu(dy_j)
$$

Contrary to what was done, by mistake, between equations (13) and (14) in [8], we cannot exchange the product and the integral over Ω. This can be corrected in a way by a different decomposition :

$$
Z_N = \int_{\mathbb{R}^{(2N+1)d}} \int_\Omega \prod_{j=-N}^{N-1} \alpha_j(a, y, u) \bigotimes_{j=-N}^{N-1} \mathbb{P}_{I_j}^{y_j, y_{j+1}}(du) \bigotimes_{j=-N}^{N} \mu(dy_j)
$$

where the coefficients α_j are defined, for $j \in \{-N+1, ..., N-2\}$, by

$$
\alpha_j(a, y, u) = \exp\left(-H_{I_j}(u^{(N)})\right) \sqrt{p(a, y_{j-1}, y_j)\, p(a, y_j, y_{j+1})}
$$

with the extremal cases $j = -N$ and $j = N-1$ as follows

$$
\alpha_{-N}(a, y, u) = \exp\left(-H_{I_{-N}}(u^{(N)})\right) \sqrt{p(a, y_{-N}, y_{-N+1})}
$$

$$
\alpha_{N-1}(a, y, u) = \exp\left(-H_{I_{N-1}}(u^{(N)})\right) \sqrt{p(a, y_{N-2}, y_{N-1})}\, p(a, y_{N-1}, y_N)
$$

In order to obtain a sum of a product of terms that are "temporally independent" from each other, we rewrite differently the product of the α_j :

$$\prod_{j=-N}^{N-1} \alpha_j(a,y,u) = \prod_{j=-N}^{N-1} (1+\alpha_j(a,y,u)-1) = 1 + \sum_S \prod_{j\in S} (\alpha_j(a,y,u)-1)$$

where the sum is taken on all non-empty subsets S of $\{-N,...,N-1\}$.
Thus,

$$\prod_{j=-N}^{N-1} \alpha_j(a,y,u) = 1 + \sum_{p\in\mathbb{N}^*} \sum_{\tau_1\sqcup...\sqcup\tau_p} \prod_{i=1}^{p} \prod_{j\in\tau_i} (\alpha_j(a,y,u)-1)$$

where $\sum_{\tau_1\sqcup...\sqcup\tau_p}$ is the sum over $\tau_1,...,\tau_p$ of the form $\tau=\{c,c+1,...,c+r\}$, with $r\geqslant 0$, $|c|\leqslant N$, $c+r\leqslant N$, and $d(\tau_i,\tau_j)\geqslant 2$ if $i\neq j$.

More precisely, these sets, called clusters, satisfy three conditions :

- $a\tau_i \subset I(N)$, in the sense that if $j\in\tau_i$, then $j\in\{-N,...,N\}$;
- if $j_1,j_2\in\tau_i$, with $j_1<j_2$, and $j_1\leqslant j_3\leqslant j_2$, then $j_3\in\tau_i$ (in some way, they are "connected sets", as subsets of $\mathbb{Z}$) ;
- if $j_1\in\tau_{i_1}$ and $j_2\in\tau_{i_2}$, with $i_1\neq i_2$, then $|j_1-j_2|\geqslant 2$ (they are "disjoint sets").

Notice that the sum over p is actually finite : according to the properties of the sets (τ_i), there are less than $2+Na$ of them, thus $p\leqslant 2+Na$.

Coming back to the expression of the partition function,

$$\begin{aligned} Z_N &= \int_{\mathbb{R}^{(2N+1)d}} \int_\Omega \left(1+\sum_{p\in\mathbb{N}^*} \sum_{\tau_1\sqcup...\sqcup\tau_p} \prod_{i=1}^{p} \prod_{j\in\tau_i} (\alpha_j(a,y,u)-1)\right) \\ &\qquad \times \bigotimes_{j=-N}^{N-1} \mathbb{P}_{I_j}^{y_j,y_{j+1}}(du) \bigotimes_{j=-N}^{N} \mu(dy_j) \\ &= 1+\sum_{p\in\mathbb{N}^*} \sum_{\tau_1\sqcup...\sqcup\tau_p} \int_{\mathbb{R}^{(2N+1)d}} \int_\Omega \prod_{i=1}^{p} \prod_{j\in\tau_i} (\alpha_j(a,y,u)-1) \\ &\qquad \times \bigotimes_{j=-N}^{N-1} \mathbb{P}_{I_j}^{y_j,y_{j+1}}(du) \bigotimes_{j=-N}^{N} \mu(dy_j) \end{aligned} \tag{3.3.5}$$

The decomposition of the product of the α_j was done to be able to invert the product for i from 1 to p and both integrals in the expression (3.3.5) just above. This is indeed now possible :

- Take a cluster $\tau_i=\{c_i,...,c_i+r_i\}$.

  As $\prod_{j\in\tau_i} (\alpha_j(a,y,u)-1)$ only depends on $u(t)$ for

  $$t\in\bigcup_{j\in\tau_i} [ja-t_0,(j+1)a] = [(c_ia-t_0))\wedge(-Na),(c_i+r_i+1)a] \subset I_{c_i-1}\cup...\cup I_{c_i+r_i}$$

and $(c_{i_1}+r_{i_1}+1)a < c_{i_2}a - t_0$ for $i_1, i_2 \in \{1, ..., p\}$, $i_1 \neq i_2$ and a large enough, we have

$$\left(I_{c_{i_1}-1} \cup ... \cup I_{c_{i_1}+r_{i_1}}\right) \cap \left(I_{c_{i_2}-1} \cup ... \cup I_{c_{i_2}+r_{i_2}}\right) = \emptyset.$$

This allows us to invert the product of the α_j with the integral over Ω : we thus have

$$Z_N = 1 + \sum_{p\in\mathbb{N}^*} \sum_{\tau_1 \sqcup ... \sqcup \tau_p} \int_{\mathbb{R}^{(2N+1)d}} \prod_{i=1}^{p} \left[\int_\Omega \prod_{j\in\tau_i} (\alpha_j(a,y,u)-1) \bigotimes_{k=c_i-1}^{c_i+r_i-1} \mathbb{P}_{I_k}^{y_k,y_{k+1}}(du) \right] \bigotimes_{j=-N}^{N} \mu(dy_j)$$

- Moreover, notice that the expression between the square brackets only depends on $y_{c_i-1}, y_{c_i}, ..., y_{c_i+r_i}$. As a consequence, we can interchange the integral over $\mathbb{R}^{2N+1}$ and the product in i.

Thus, we obtain the following cluster representation of the partition function Z_N :

$$Z_N = 1 + \sum_{p\in\mathbb{N}^*} \sum_{\tau_1 \sqcup ... \sqcup \tau_p \subset I(N)} \prod_{i=1}^{p} \Gamma_{\tau_i} \tag{3.3.6}$$

where

$$\Gamma_\tau = \int_{\mathbb{R}^{(|\tau|+1)d}} \int_\Omega \prod_{j\in\tau} (\alpha_j(a,y,u)-1) \bigotimes_{k=\min(\tau)-1}^{\max(\tau)-1} \mathbb{P}_{I_k}^{y_k,y_{k+1}} \bigotimes_{l=\min(\tau)-1}^{\max(\tau)} \mu(dy_l) \tag{3.3.7}$$

with $|\tau|$ the cardinal of τ.

3.4 Cluster estimates

Having obtained the quantities Γ_τ associated to a cluster τ, we now wish to control them. More specifically, we will show that, when the perturbation coefficient β is sufficiently small, there exists a positive function $\eta(\beta)$, which goes to 0 when β goes to 0, such that for a large enough,

$$|\Gamma_\tau| \leqslant \eta(\beta)^{|\tau|}. \tag{3.4.1}$$

3.4.1 First upper-bound for the clusters

In order to estimate this coefficient Γ_τ, we commute the integrals and the remaining product (over the elements of τ), to obtain the following inequality.

Proposition 23. *Setting*

$$A_j(a) = \int_{\mathbb{R}^{3d}} \int_\Omega (\alpha_j(a,y,u)-1)^4 \, \mathbb{P}_{I_{j-1}}^{y_{j-1},y_j}(du) \, \mathbb{P}_{I_j}^{y_j,y_{j+1}}(du) \, \mu(dy_{j-1}) \, \mu(dy_j) \, \mu(dy_{j+1}),$$

we have, for every cluster τ involved in the decomposition (3.3.6),

$$\Gamma_\tau \leqslant \prod_{j\in\tau} A_j(a)^{1/4}$$

Proof. The following lemma, taken from [20], is the main ingredient of the proof.

Lemma 7. Let $(\mu_x)_{x\in\mathcal{X}}$ be a family of probability measures, each one defined on a space E_x, where the elements x belong to some finite set $\mathcal{X}$. Let us also define a finite family $(f_i)_i$ of functions on $E_{\mathcal{X}} = \times_{x\in\mathcal{X}} E_x$ such that each f_i is $\mathcal{X}_i$-local for a certain $\mathcal{X}_i \subset \mathcal{X}$, in the sense that

$$f_i(e) = f_i(e_{|\mathcal{X}_i}), \text{ for } e = (e_x)_{x\in\mathcal{X}} \in E_{\mathcal{X}}.$$

Let $\rho_i > 0$ be numbers satisfying the following conditions :

$$\forall x \in \mathcal{X}, \sum_{\mathcal{X}_i \ni x} \frac{1}{\rho_i} \leqslant 1.$$

Then

$$\left| \int_{E_{\mathcal{X}}} \prod_i f_i \otimes_{x\in\mathcal{X}} d\mu_x \right| \leqslant \prod_i \left(\int_{E_{\mathcal{X}_i}} |f_i|^{\rho_i} \otimes_{x\in\mathcal{X}_i} d\mu_x \right)^{1/\rho_i}$$

We apply lemma 7 twice consecutively, first with respect to the integral over Ω, then with respect to the integral over $\mathbb{R}^{(|\tau|+1)d}$.

- For $\tau = \{c, ..., c+r\}$, set

$$I_\tau(y) = \int_\Omega \prod_{j=c}^{c+r} (\alpha_j(a,y,u) - 1) \bigotimes_{k=c-1}^{c+r-1} \mathbb{P}_{I_k}^{y_k, y_{k+1}}(du).$$

  Taking $\mathcal{X} = \{c-1, ..., c+r-1\}$, $\mathcal{X}_i = \{i-1, i\}$, $E_{\mathcal{X}} = \Omega$, $E_k = \mathcal{C}(I_k, \mathbb{R})$ and $d\mu_k = \mathbb{P}_{I_k}^{y_k, y_{k+1}}$, for $(\rho_j)_{j\in\tau}$ such that $\rho_j > 1$ and $\dfrac{1}{\rho_j} + \dfrac{1}{\rho_{j+1}} \leqslant 1$, by Lemma 7,

$$I_\tau(y) \leqslant \prod_{j=c}^{c+r} g_j(y_{j-1}, y_j, y_{j+1})^{1/\rho_j}$$

  where $\quad g_j(y_{j-1}, y_j, y_{j+1}) = \displaystyle\int_\Omega |\alpha_j(a,y,u) - 1|^{\rho_j} \, \mathbb{P}_{I_{j-1}}^{y_{j-1}, y_j}(du) \, \mathbb{P}_{I_j}^{y_j, y_{j+1}}(du).$

- Set now

$$\widetilde{\Gamma}_\tau = \int_{\mathbb{R}^{(r+2)d}} \prod_{j=c}^{c+r} g_j^{1/\rho_j} \bigotimes_{l=c-1}^{c+r} \mu(dy_l).$$

  Here we choose $\mathcal{X} = \{c-1, ..., c+r\}$, $\mathcal{X}_i = \{i-1, i, i+1\}$, $E_{\mathcal{X}} = \mathbb{R}^{(r+2)d}$, $E_x = \mathbb{R}^d$ and $d\mu_x = \mu(y_x)$, for $(\gamma_j)_{j\in\{-N,...,N\}}$ such that $\gamma_j > 1$ and $\dfrac{1}{\gamma_{j-1}} + \dfrac{1}{\gamma_j} + \dfrac{1}{\gamma_{j+1}} \leqslant 1$, lemma 7 ensures that

$$\widetilde{\Gamma}_\tau \leqslant \prod_{j=c}^{c+r} \left(\int_{\mathbb{R}^{3d}} |g_j|^{\gamma_j/\rho_j} \, \mu(dy_{j-1}) \, \mu(dy_j) \, \mu(dy_{j+1}) \right)^{1/\gamma_j}.$$

For every $i \in \tau$, every $j \in \{-N, ..., N\}$, we take $\rho_i = \gamma_j = 4$, and this concludes the proof. □

We now control this quantity and prove that it goes to 0, uniformly in j, and even independently of N, for a large enough time-scale a.

3.4.2 Decomposition of $A_j(a)$

Using that

$$xy - 1 = (x-1)\,y + (y-1) \quad \text{ and that } \quad (xy-1)^4 \leqslant 8\,((x-1)^4\,y^4 + (y-1)^4)$$

for non-negative x and y, and coming back to the expression of the α_j, we can decompose $A_j(a)$ in two parts that will be dealt with separately :

$$A_j(a) \leqslant 8\,B_j(a) + 8\,C_j(a)$$

- if $j \in \{-N+1, ..., N-2\}$,

$$\begin{aligned} B_j(a) := \int_{\mathbb{R}^{3d}} \int_{\Omega} &(e^{-H_{I_j}(u^{(N)})} - 1)^4\, p(a, y_{j-1}, y_j)^2\, p(a, y_j, y_{j+1})^2 \\ &\mathbb{P}_{I_{j-1}}^{y_{j-1}, y_j}(du)\, \mathbb{P}_{I_j}^{y_j, y_{j+1}}(du)\, \mu(dy_{j-1})\, \mu(dy_j)\, \mu(dy_{j+1}) \\ = \int_{\Omega} &(e^{-H_{I_j}(u^{(N)})} - 1)^4\, p(a, u((j-1)a), u(ja))\, p(a, u(ja), u((j+1)a))\, \mathbb{P}(du) \end{aligned}$$

$$\begin{aligned} C_j(a) := \int_{\mathbb{R}^{3d}} \int_{\Omega} &\left(\sqrt{p(a, y_{j-1}, y_j)\, p(a, y_j, y_{j+1})} - 1\right)^4 \\ &\times \mathbb{P}_{I_{j-1}}^{y_{j-1}, y_j}(du)\, \mathbb{P}_{I_j}^{y_j, y_{j+1}}(du)\, \mu(dy_{j-1})\, \mu(dy_j)\, \mu(dy_{j+1}) \\ &= \int_{\mathbb{R}^{3d}} \left(\sqrt{p(a,x,y)\, p(a,y,z)} - 1\right)^4 \mu(dx)\, \mu(dy)\, \mu(dz) \end{aligned}$$

- if $j = -N$,

$$B_{-N}(a) := \int_{\Omega} \left(e^{-H_{I_{-N}}(u^{(N)})} - 1\right)^4\, p(a, u(-Na), u((-N+1)a))\, \mathbb{P}(du)$$

$$C_{-N}(a) := \int_{\mathbb{R}^{2d}} \left(\sqrt{p(a,x,y)} - 1\right)^4 \mu(dx)\, \mu(dy)$$

- if $j = N-1$,

$$\begin{aligned} B_{N-1}(a) := \int_{\Omega} \left(e^{-H_{I_{N-1}}(u^{(N)})} - 1\right)^4\, &p(a, u((N-2)a), u((N-1)a)) \\ &\times\; p(a, u((N-1)a), u(Na))^2\, \mathbb{P}(du) \end{aligned}$$

$$C_{N-1}(a) := \int_{\mathbb{R}^{3d}} \left(\sqrt{p(a,x,y)}\, p(a,y,z) - 1\right)^4 \mu(dx)\, \mu(dy)\, \mu(dz)$$

We will now study separately the B_j and the C_j, without omitting the two boundary cases, especially the one when $j = N-1$, which will turn out to be the most troublesome.

3.4.3 Study of $B_j(a)$

Case $j \in \{-N+1, ..., N-2\}$

Using Cauchy-Schwarz's inequality, we again decompose the integral in two parts :

$$B_j(a) = \int_\Omega (e^{-H_{I_j}(u^{(N)})} - 1)^4 \, p(a, u((j-1)a), u(ja)) \, p(a, u(ja), u((j+1)a)) \, \mathbb{P}(du) \\ \leqslant \widetilde{B}_j(a) \, K_j(a)$$

with

$$\widetilde{B}_j(a) = \left(\int_\Omega p(a, u((j-1)a), u(ja))^2 \, p(a, u(ja), u((j+1)a))^2 \, \mathbb{P}(du) \right)^{1/2}$$

and

$$K_j(a) = \left(\int_\Omega (e^{-H_{I_j}(u^{(N)})} - 1)^8 \, \mathbb{P}(du) \right)^{1/2}.$$

Notice that $K_j(a)$ is bounded uniformly in j : indeed,

$$\begin{aligned} K_j(a)^4 &= \left(\int_\Omega p(a, u((j-1)a), u(ja))^2 \, p(a, u(ja), u((j+1)a))^2 \, \mathbb{P}(du) \right)^2 \\ &\leqslant \int_\Omega p(a, u((j-1)a), u(ja))^4 \, \mathbb{P}(du) \int_\Omega p(a, u(ja), u((j+1)a))^4 \, \mathbb{P}(du) \\ &= \mathbb{E}\left[p(a, y((j-1)a), y(ja))^4 \right] \mathbb{E}[p(a, y(ja), y((j+1)a))^4] \\ &= \left(\int_{\mathbb{R}^{2d}} p(a,x,y)^4 \, p(a,x,y) \, \mu(dx) \, \mu(dy) \right)^2 = \|p(a,.,.)\|^{10}_{L^5(\mu \otimes \mu)} \end{aligned}$$

The main goal of this subsection is to find an upper bound for

$$\widetilde{B}_j(a) = \left(\int_\Omega (e^{-H_{I_j}(u^{(N)})} - 1)^8 \, \mathbb{P}(du) \right)^{1/2},$$

depending on a and going to 0 as soon as a goes to infinity.

What follows is a direct adaptation of what was done in [22] and [8].

We start by noticing that for every $x \in \mathbb{R}$,

$$(e^{-x} - 1)^8 = x^8 \left(\int_0^1 e^{-tx} \, dt \right)^8 = x^8 \int_{[0,1]^8} e^{-(t_1 + ... + t_8)x} \, dt_1 \, ... \, dt_8$$

and thus

$$\widetilde{B}_j(a)^2 = \int_{[0,1]^8} \int_\Omega H_{I_j}(u^{(N)})^8 \, e^{-(t_1 + ... + t_8) \, H_{I_j}(u^{(N)})} \, \mathbb{P}(du) \, dt_1 \, ... \, dt_8$$

Set $L(z) = \int_\Omega e^{-zH_{I_j}(u^{(N)})} \mathbb{P}(du)$.
Then, L is an holomorphic function, and its eighth derivative is

$$\frac{\partial}{\partial z^8} L(z) = \int_\Omega H_{I_j}(u^{(N)})^8 \, e^{-z\, H_{I_j}(u^{(N)})} \, \mathbb{P}(du)$$

which means we can rewrite $\widetilde{B}_j$ as

$$\widetilde{B}_j(a)^2 = \int_{[0,1]^8} \frac{\partial}{\partial z^8} L(z) \,|_{z=t_1+...+t_8} \; dt_1 \,...\, dt_8 \tag{3.4.2}$$

Notice that

$$|L(z)| \leqslant \int_\Omega |e^{-z\, H_{I_j}(u^{(N)})}| \, \mathbb{P}(du) = \int_\Omega e^{-\,\mathfrak{Re}(z)\, H_{I_j}(u^{(N)})} \, \mathbb{P}(du) \;=\; L(\mathfrak{Re}(z))$$

Recall that the expression of the Hamiltonian H is given by equation (3.3.1).
For any real number x, we can obtain an alternative expression of L, using Cauchy-Schwarz inequality and the martingale property of $\exp(-H)$:

$$\begin{aligned}
L(x) &= \int_\Omega \exp\left(x \int_{I_j} \beta\, b((u)^t_{t-t_0})^* \, dW_t - \frac{x}{2} \int_{I_j} \beta^2 \, |b((u)^t_{t-t_0})|^2 \, dt \right) \mathbb{P}(du) \\
&= \int_\Omega \exp\left(x \int_{ja}^{(j+1)a} \beta\, b((u)^t_{t-t_0})^* \, dW_t - \, x^2 \int_{ja}^{(j+1)a} \beta^2 \, |b((u)^t_{t-t_0})|^2 \, dt \right) \\
&\qquad \times \, \exp\left(\frac{2x^2 - x}{2} \int_{ja}^{(j+1)a} \beta^2 \, |b((u)^t_{t-t_0})|^2 \, dt \right) \; \mathbb{P}(du) \\
&\leqslant \left(\int_\Omega \exp\left(2x \int_{ja}^{(j+1)a} \beta\, b((u)^t_{t-t_0})^* \, dW_t - 2x^2 \int_{ja}^{(j+1)a} \beta^2 \, |b((u)^t_{t-t_0})|^2 \, dt \right) \mathbb{P}(du) \right)^{1/2} \\
&\qquad \times \left(\int_\Omega \exp\left(x(2x-1) \int_{ja}^{(j+1)a} \beta^2 \, |b((u)^t_{t-t_0})|^2 \, dt \right) \mathbb{P}(du) \right)^{1/2} \\
&= \left(\int_\Omega \exp\left(x(2x-1) \int_{ja}^{(j+1)a} \beta^2 \, |b((u)^t_{t-t_0})|^2 \, dt \right) \mathbb{P}(du) \right)^{1/2}
\end{aligned}$$

We now apply Cauchy's inequality to L, for ρ such that L is well defined on $B(z,\rho) = \{v \in \mathbb{C}, |v - z| \leqslant \rho\}$:

$$\left| \frac{\partial}{\partial z^8} L(z) \right| \leqslant \frac{8!}{\rho^8} \sup_{v \in B(z,\rho)} |L(v)| \tag{3.4.3}$$

Thanks to the above expression of L,

$$|L(v)|^2 \leqslant \int_\Omega \exp\left(\mathfrak{Re}(v) \; (2\, \mathfrak{Re}(v) - 1) \int_{ja}^{(j+1)a} \beta^2 \, |b((u)^t_{t-t_0})|^2 \, dt \right) \mathbb{P}(du)$$

As $|v - z|^2 = \rho^2$, we have $(\mathfrak{Re}(v) - z) \leqslant \rho^2$ for $z = t_1 + ... + t_8$, it follows that $\mathfrak{Re}(v) \in \{x \in \mathbb{R} : |z - x| < \rho\}$, hence $\mathfrak{Re}(v) \leqslant z + \rho$, which implies $\mathfrak{Re}(v)\; (2\, \mathfrak{Re}(v) - 1) \leqslant 2\,(z+\rho)^2$.
Subsequently,

$$|L(v)|^2 \leqslant \int_\Omega \exp\left(2\,(\rho + z)^2) \, a\beta^2 \right) \, \mathbb{P}(du),$$

and thus,

$$\sup_{v\in\mathcal{C}(z,\rho)} |L(v)| \leqslant \exp\Big((\rho+z)^2\ a\beta^2)\Big) \tag{3.4.4}$$

where $\mathcal{C}(z,\rho)=\{v\in\mathbb{C}, |v-z|=\rho\}$.

Combining (3.4.2), (3.4.3) and (3.4.4),

$$\widetilde{B}_j(a)^2 \leqslant \int_{[0,1]^8} \frac{8!}{\rho^8}\ \exp\Big(\ (\rho+t_1+...+t_8)^2\ a\beta^2\Big)\ dt_1\ ...\ dt_8 \leqslant \frac{8!}{\rho^8}\ \exp\Big((\rho+8)^2\ a\beta^2)\Big)$$

It implies that, for every $\rho\geqslant 8$,

$$\widetilde{B}_j(a)^2 \leqslant \frac{8!}{\rho^8}\ e^{4\rho^2\ a\beta^2} \tag{3.4.5}$$

We want to determine which $\rho\geqslant 8$ will minimize the right hand side of this last inequality.

Let f be the function given by $f(\rho)=\dfrac{8!}{\rho^8}\ e^{4\rho^2\ a\beta^2}$. Then, $f^{'}(\rho)=\Big(-\dfrac{8}{\rho}+8\rho\ a\beta^2\Big)\ f(\rho)$.

Thus, $f^{'}(\rho)=0$ if and only if $\rho^2=\dfrac{1}{a\beta^2}$, which is larger than 8 if and only if

$$a\beta^2 \leqslant \frac{1}{8}, \tag{3.4.6}$$

and the optimal inequality for (3.4.5) is

$$\widetilde{B}_j(a)^2 \leqslant 8!\ e^4\ (a\beta^2)^4 \tag{3.4.7}$$

Finally coming back to the expression of B_j, we have obtained, under condition (3.4.6),

$$B_j(a) \leqslant \sqrt{8!}\ e^2\ \|p(a,.,.)\|^{5/2}_{L^5(\mu\otimes\mu)}\ (a\beta^2)^2 \tag{3.4.8}$$

Boundary cases, $j\in\{-N,N-1\}$

Remember that

$$B_{-N}(a)=\int_\Omega \Big(e^{-H_{I_{-N}}(u^{(N)})}-1\Big)^4\ p(a,u(-Na),u((-N+1)a))\ \mathbb{P}(du).$$

As in the previous case, we can write

$$B_{-N}(a)\leqslant K_{-N}(a)\Big(\int_\Omega (e^{-H_{I_{-N}}(u^{(N)})}-1)^8\ \mathbb{P}(du)\Big)^{1/2}$$

where

$$K_{-N}(a)^2=\int_\Omega p(a,u(-Na),u((-N+1)a))^2\ \mathbb{P}(du).$$

This square root can be dealt with in exactly the same fashion as is done above.

Furthermore,

$$K_{-N}(a)^2 = \mathbb{E}[p(a,y(-Na),y((-N+1)a))^2]$$
$$= \int_{\mathbb{R}^{2d}} p(a,x,y)^2\ p(a,x,y)\ \mu(dx)\ \mu(dy) = \|p(a,.,.)\|^3_{L^3(\mu\otimes\mu)}$$

Hence the following result :

$$B_{-N}(a) \leqslant \sqrt{8!}\ e^2\ \|p(a,.,.)\|^{3/2}_{L^3(\mu\otimes\mu)}\ (a\beta^2)^2 \tag{3.4.9}$$

We now turn our attention to

$$B_{N-1}(a) = \int_\Omega \left(e^{-H_{I_{N-1}}(u^{(N)})} - 1\right)^4\ p(a,u((N-2)a),u((N-1)a))\ p(a,u((N-1)a),u(Na))^2\ \mathbb{P}(du)$$

We proceed in a similar way to decompose $B_{N-1}(a)$ into the product of two terms and we have to study the quantity :

$$K_{N-1}(a) = \sqrt{\int_\Omega p(a,u((N-2)a),u((N-1)a))^2\ p(a,u((N-1)a),u(Na))^4\ \mathbb{P}(du)}$$

In order to obtain an upper bound for a moment of $p(a,.,.)$, with respect to $\mu\otimes\mu$, smaller than 8, Cauchy-Schwarz's inequality will not suffice : we have to apply Hölder's inequality. We choose the conjugated numbers 3 and 3/2 :

$$K_{N-1}(a)^2 \leqslant \left(\int_\Omega p(a,u((N-2)a),u((N-1)a))^6\ \mathbb{P}(du)\right)^{1/3}$$
$$\times \left(\int_\Omega p(a,u((N-1)a),u(Na))^6\ \mathbb{P}(du)\right)^{2/3}$$

which leads to

$$K_{N-1}(a) \leqslant \|p(a,.,.)\|^7_{L^7(\mu\otimes\mu)}$$

and subsequently to

$$B_{N-1}(a) \leqslant \sqrt{8!}\ e^2\ \|p(a,.,.)\|^{7/2}_{L^7(\mu\otimes\mu)}\ (a\beta^2)^2 \tag{3.4.10}$$

A bound for $B_j(a)$, uniform in N

From (3.4.8), (3.4.9) and (3.4.10), we deduce that, for every $j\in\{-N,...,N-1\}$,

$$B_j(a) \leqslant \sqrt{8!}\ e^2\ \|p(a,.,.)\|^{7/2}_{L^7(\mu\otimes\mu)}\ (a\beta^2)^2 \tag{3.4.11}$$

3.4.4 Study of $C_j(a)$

General case, $j \in \{-N+1, ..., N-2\}$

We remind that

$$C_j(a) = \int_{\mathbb{R}^{3d}} \left(\sqrt{p(a,x,y)\, p(a,y,z)} - 1\right)^4 \mu(dx)\, \mu(dy)\, \mu(dz).$$

Again, we seek an upper bound for $C_j(a)$ which vanishes when a goes to infinity.

It can be easily checked that for every positive real number U,

$$(\sqrt{1+U} - 1)^4 \leqslant \frac{1}{16}\, U^4, \tag{3.4.12}$$

that for positive x and y,

$$xy - 1 = (x-1)\,(y-1) + (x-1) + (y-1)$$

and that, thanks to the convexity of $u \mapsto u^4$, for any a, b and c,

$$(a+b+c)^4 \leqslant 27\,(a^4 + b^4 + c^4).$$

Subsequently,

$$\begin{aligned}
C_j(a) &\leqslant \frac{1}{16} \int_{\mathbb{R}^{3d}} (p(a,x,y)\, p(a,y,z) - 1)^4\ \mu(dx)\, \mu(dy)\, \mu(dz) \\
&\leqslant \frac{1}{16} \int_{\mathbb{R}^{3d}} [(p(a,x,y) - 1)\,(p(a,y,z) - 1) + (p(a,x,y) - 1) \\
&\qquad + (p(a,y,z) - 1)]^4\ \mu(dx)\, \mu(dy)\, \mu(dz) \\
&\leqslant \frac{27}{16} \int_{\mathbb{R}^{3d}} ((p(a,x,y) - 1)\,(p(a,y,z) - 1))^4\ \mu(dx)\, \mu(dy)\, \mu(dz) \\
&\qquad + \frac{27}{8} \int_{\mathbb{R}^{2d}} (p(a,x,y) - 1)^4\, \mu(dx)\, \mu(dy) \\
&\leqslant \frac{27}{16} \int_{\mathbb{R}^{2d}} (p(a,x,y) - 1)^8\, \mu(dx)\, \mu(dy)\ + \frac{27}{8} \int_{\mathbb{R}^{2d}} (p(a,x,y) - 1)^4\, \mu(dx)\, \mu(dy)
\end{aligned}$$

using, once more, Cauchy-Schwarz's inequality to obtain the final line.

Furthermore,

$$\int_{\mathbb{R}^{2d}} (p(a,x,y) - 1)^4\, \mu(dx)\, \mu(dy) \leqslant \sqrt{\int_{\mathbb{R}^{2d}} (p(a,x,y) - 1)^8\, \mu(dx)\, \mu(dy)}.$$

Thus, according to Proposition 22,

$$C_j(a) \leqslant \frac{27}{16}\, \gamma_\delta(a)^8\ \left(\|p(\delta,.,.)\|^8_{L^8(\mu\otimes\mu)} \vee 1\right) + \frac{27}{8}\, \gamma_\delta(a)^4\ \left(\|p(\delta,.,.)\|^4_{L^8(\mu\otimes\mu)} \vee 1\right).$$

Boundary cases, $j \in \{-N, N-1\}$

We can check that both boundary cases exhibit an analogous behaviour.

Indeed, on the one hand, recall that

$$C_{-N}(a) = \int_{\mathbb{R}^{2d}} \left(\sqrt{p(a,x,y)} - 1\right)^4 \mu(dx)\,\mu(dy).$$

Thus, thanks to (3.4.12) and Proposition 22,

$$\begin{aligned} C_{-N}(a) &\leqslant \frac{1}{16} \int_{\mathbb{R}^{2d}} (p(a,x,y)-1)^4 \;\mu(dx)\,\mu(dy) \\ &\leqslant \frac{1}{16} \sqrt{\int_{\mathbb{R}^{2d}} (p(a,x,y)-1)^8 \;\mu(dx)\,\mu(dy)} \\ &\leqslant \frac{1}{16}\,\gamma_\delta(a)^4 \left(\|p(\delta,.,.)\|^4_{L^8(\mu\otimes\mu)} \vee 1\right). \end{aligned}$$

On the other hand,

$$C_{N-1}(a) = \int_{\mathbb{R}^{3d}} \left(\sqrt{p(a,x,y)}\,p(a,y,z) - 1\right)^4 \mu(dx)\,\mu(dy)\,\mu(dz.)$$

Notice that

$$(\sqrt{p(a,x,y)}\,p(a,y,z)-1)^4 \leqslant 8\,(p(a,y,z)-1)^4\,p(a,x,y)^2 + 8\,(\sqrt{p(a,x,y)}-1)^4.$$

With Cauchy-Schwarz inequality and the computation of $C_{-N}(a)$ thrown in, it leads to

$$\begin{aligned} C_{N-1}(a) &\leqslant 8 \int_{\mathbb{R}^{3d}} (p(a,y,z)-1)^4\,p(a,x,y)^2\,\mu(dx)\,\mu(dy)\,\mu(dz) \\ &\quad + 8 \int_{\mathbb{R}^{2d}} (\sqrt{p(a,x,y)}-1)^4\,\mu(dx)\,\mu(dy) \\ &\leqslant 8 \sqrt{\int_{\mathbb{R}^{2d}} (p(a,x,y)-1)^8\,\mu(dx)\,\mu(dy)} \sqrt{\int_{\mathbb{R}^{2d}} p(a,x,y)^4\,\mu(dx)\,\mu(dy)} \\ &\quad + \frac{1}{2}\,\gamma_\delta(a)^4 \left(\|p(\delta,.,.)\|^4_{L^8(\mu\otimes\mu)} \vee 1\right) \\ &\leqslant 8\,\gamma_\delta(a)^4 \left(\|p(\delta,.,.)\|^4_{L^8(\mu\otimes\mu)} \vee 1\right) \|p(a,.,.)\|^2_{L^4(\mu\otimes\mu)} \\ &\quad + \frac{1}{2}\,\gamma_\delta(a)^4 \left(\|p(\delta,.,.)\|^4_{L^8(\mu\otimes\mu)} \vee 1\right) \\ &= \left(8\,\|p(a,.,.)\|^2_{L^4(\mu\otimes\mu)} + \frac{1}{2}\right) \gamma_\delta(a)^4 \left(\|p(\delta,.,.)\|^4_{L^8(\mu\otimes\mu)} \vee 1\right). \end{aligned}$$

Global upper bound for all C_j

Taking into account all three cases, when $a \geqslant 2\delta$, for every $j \in \{-N, ..., N-1\}$,

$$C_j(a) \leqslant \gamma_\delta(a)^4 \left(\|p(\delta,.,.)\|^4_{L^8(\mu\otimes\mu)} \vee 1\right) \left(\left(\frac{27}{16}\,\gamma_\delta(a)^4 + 8\right) \left(\|p(\delta,.,.)\|^4_{L^8(\mu\otimes\mu)} \vee 1\right) + 4\right) \tag{3.4.13}$$

To obtain the control of $|\Gamma_\tau|$ we are seeking, it remains to put all the pieces together and to determine its domain of validity with respect to a and b.

3.4.5 Back to the clusters

At last, the proposition below gives us the cluster estimates.

Proposition 24. *Assume that (H1) and (H2) are satisfied. Let ε be a positive number. There exist a minimal time-scale a_ε, defined in (3.4.22), and an upper-bound β_ε, given by (3.4.23), such that if $a \geqslant a_\varepsilon$ and $\beta \leqslant \beta_\varepsilon$, then, for every cluster τ, the quantity Γ_τ defined in (3.3.7) satisfies*

$$|\Gamma_\tau| \leqslant \varepsilon^{|\tau|}. \tag{3.4.14}$$

Proof. Suppose that $a \geqslant 2\delta$ and that (3.4.6) holds, i.e. $\beta \leqslant \dfrac{1}{\sqrt{8a}}$.

Recall that

$$M_\delta = \sup_{a \geqslant \delta} \|p(a,.,.)\|_{L^8(\mu\otimes\mu)} \vee 1 \tag{3.4.15}$$

Thus, we can obtain bounds for B_j and C_j easier to deal with : according to (3.4.11) and (3.4.13) respectively,

$$B_j(a) \leqslant \sqrt{8!}\; e^2\; M_\delta^{7/2}\; (a\beta^2)^2 \tag{3.4.16}$$

$$C_j(a) \leqslant M_\delta^4\; \left(4 + 2M_\delta^4\; \left(4 + \gamma_\delta(a)^4\right)\right) \gamma_\delta(a)^4$$

Remember that

$$|\Gamma_\tau(a)| \leqslant \prod_{j\in\tau} \left(8\; B_j(a) + 8\; C_j(a)\right)^{1/4},$$

so (3.4.14) will be satisfied if, for a sufficiently large, both $B_j(a)$ and $C_j(a)$ are smaller than $\varepsilon^4/16$.

- One can check, by solving a second order inequality in $\gamma_\delta(a)^4$, that for all a such that

  $$\gamma_\delta(a)^4 \leqslant \left(2 + \frac{1}{M_\delta^4}\right)\left(\sqrt{1 + \frac{\varepsilon^4}{32\;(1 + 2\; M_\delta^4)^2}} - 1\right) \tag{3.4.17}$$

  the condition $C_j(a) \leqslant \dfrac{\varepsilon^4}{16}$ is true.

  We remind that γ_δ was introduced in Proposition 22 and is defined by

  $$\gamma_\delta(a) = 2\; M_\delta\; e^{-(a-2\delta)/C_P},$$

  with C_P the constant associated with the Poincaré's inequality satisfied by μ, according to hypothesis (H1).

Using this expression, and setting

$$a_C(\varepsilon) = 2\,\delta - \frac{C_P}{4}\,\ln\left(\frac{1}{16\,M_\delta^4}\left(2+\frac{1}{M_\delta^4}\right)\left(\sqrt{1+\frac{\varepsilon^4}{32\,(1+2\,M_\delta^4)^2}}-1\right)\right) \tag{3.4.18}$$

it can be shown that (3.4.17) is equivalent to :

$$a \geqslant a_C(\varepsilon)$$

Thus, for every $a \geqslant a_C(\varepsilon)$, $C_j(a) \leqslant \frac{\varepsilon^4}{16}$.

It can be noticed that

$$a_C(\varepsilon) \geqslant 2\delta \text{ if and only if } \varepsilon \leqslant 2^{5/2}\,M_\delta^2\,(8\,M_\delta^8 + 2\,M_\delta^4 + 1)^{1/4}. \tag{3.4.19}$$

- From (3.4.16), it can be seen that $B_j(a) \leqslant \frac{\varepsilon^4}{16}$ if

$$\beta \leqslant \frac{1}{2\,\sqrt{e}\,(8!)^{1/8}\,M_\delta^{7/8}}\,\frac{\varepsilon}{\sqrt{a}} \tag{3.4.20}$$

  Notice that

$$\frac{\varepsilon}{2\,\sqrt{e}\,(8!)^{1/8}\,M_\delta^{7/8}} \leqslant \frac{1}{\sqrt{8}} \text{ if and only if } \varepsilon \leqslant \sqrt{\frac{e}{2}}\,(8!)^{1/8}\,M_\delta^{7/8}. \tag{3.4.21}$$

Thus, according to (3.4.6), (3.4.17) and (3.4.20), setting

$$a_\varepsilon = a_C(\varepsilon) \vee (2\delta)$$

that is

$$a_\varepsilon = 2\,\delta - \left[\frac{C_P}{4}\,\ln\left(\frac{1}{16\,M_\delta^4}\,\left(2+\frac{1}{M_\delta^4}\right)\,\left(\sqrt{1+\frac{\varepsilon^4}{32\,(1+2\,M_\delta^4)^2}}-1\right)\right)\right]_- \tag{3.4.22}$$

where $x_- = min(x,0)$, and

$$\beta_\varepsilon = \left(\frac{\varepsilon}{2\,\sqrt{e}\,(8!)^{1/8}\,M_\delta^{7/8}}\;\wedge\;\frac{1}{\sqrt{8}}\right)\,\frac{1}{\sqrt{a_\varepsilon}},$$

that is,

$$\beta_\varepsilon = \frac{\frac{\varepsilon}{2\,\sqrt{e}\,(8!)^{1/8}\,M_\delta^{7/8}}\;\wedge\;\frac{1}{\sqrt{8}}}{\sqrt{2\delta - \left[\frac{C_P}{4}\,\ln\left(\frac{1}{16\,M_\delta^4}\,\left(2+\frac{1}{M_\delta^4}\right)\left(\sqrt{1+\frac{\varepsilon^4}{32\,(1+2\,M_\delta^4)^2}}-1\right)\right)\right]_-}} \tag{3.4.23}$$

the proposition holds.

□

Remark 17. We recall hypothesis (H2) : there exists $\delta \geqslant 0$ such that

$$\sup_{t \geqslant \delta} \|p(t,.,.)\|_{L^8(\mu\otimes\mu)} < \infty.$$

One can notice that we only require $\|p(t,.,.)\|_{L^8(\mu\otimes\mu)}$ to be finite for certain values of t and it is not necessary for the supremum over t to be finite. However, the current form of (H2) simplifies the writing of the proofs, and it is satisfied by the important case of the Ornstein-Uhlenbeck process, as will be seen in section 3.6.

3.5 Completion of the proof of Theorem 14

In order to connect with the cluster expansion method, and to obtain an inequality of the form of (3.4.1), we have to show that ε can be expressed as a function of β that will go to 0 when β goes to 0.

Suppose that

$$\varepsilon \leqslant \left(2^{5/2}\, M_\delta^2\, (8M_\delta^8 + 2\, M_\delta^4 + 1)^{1/4}\right) \wedge \left(\sqrt{\frac{e}{2}}\, (8!)^{1/8}\, M_\delta^{7/8}\right) =: \varepsilon_0. \tag{3.5.1}$$

Then according to equivalences (3.4.19) and (3.4.21),

$$\beta_\varepsilon = \frac{\varepsilon}{2\,\sqrt{2e}\,(8!)^{1/8}\,M_\delta^{7/8}\sqrt{\delta - \frac{C_P}{8}\,\ln\left(\frac{1}{16\,M_\delta^4}\left(2+\frac{1}{M_\delta^4}\right)\left(\sqrt{1+\frac{\varepsilon^4}{32\,(1+2\,M_\delta^4)^2}}-1\right)\right)}} \tag{3.5.2}$$

One can see that $\varepsilon \mapsto \beta_\varepsilon$ is of the form

$$\frac{C_1\,\varepsilon}{\sqrt{1 - C_2\,\ln(C_3\,(\sqrt{1+C_4\,\varepsilon^4}-1))}},$$

for some C_1, C_2, C_3 and C_4 depending only on δ and C_P, and is thus an invertible function.

Compute the derivative of β_ε with respect to ε :

$$\beta_\varepsilon^{'} = \frac{\beta_\varepsilon}{\varepsilon}\left(1 + \frac{C_2\,C_4}{C_1^2}\frac{\varepsilon^2\,\beta_\varepsilon^2}{\sqrt{1+C_4\,\varepsilon^4}\left(\sqrt{1+C_4\,\varepsilon^4}-1\right)}\right)$$

$\beta_\varepsilon^{'}$ is positive for every ε in $(0,\varepsilon_0]$; thus, $\varepsilon \mapsto \beta_\varepsilon$ is (strictly) increasing from $(0,\varepsilon_0]$ to $(0,\beta_{\varepsilon_0}]$.

Therefore, $\varepsilon \mapsto \beta_\varepsilon$ admits an inverse function on $\eta : (0,\beta_{\varepsilon_0}]$ that we will call η : η is increasing and $\eta(x)$ goes to 0 when x goes to 0. For simplicity, denote by β_0 the bound β_{ε_0} given by (3.5.1) and (3.5.2).

We can now rewrite Proposition 24 in a more amenable way :

Proposition 25. *There exists β_0 such that if $\beta \leqslant \beta_0$, then, for every cluster Γ_τ,*

$$|\Gamma_\tau| \leqslant \eta(\beta)^{|\tau|} \tag{3.5.3}$$

where η, defined just above, is a function that goes to 0 in 0.

Remark 18. One can easily show that a first-order approximation of β_ε, when ε is small, is

$$\beta_\varepsilon \sim C\ \varepsilon\ (-\ln(\varepsilon))^{-1/2},$$

for a certain, explicit, constant C, depending only on the parameters δ and C_P, given by the hypotheses (H1) and (H2). Since, for any $\alpha > 0$, when ϵ is small enough, $-\ln(\varepsilon) \leqslant \dfrac{1}{\varepsilon^{2\alpha}}$, thus $\beta_\varepsilon \geqslant C\ \varepsilon^{1+\alpha}$, and Proposition 25 holds, with $\eta : y \mapsto C\ y^{1/(1+\alpha)}$.

Recall that we wish to prove the convergence of the sequence of measures $(Q_N)_N$, with

$$Q_N(du) = \frac{1}{Z_N}\ \exp(-H_N(u^{(N)}))\ \mathbb{P}(du),$$

towards a weak stationary solution Q of the perturbed equation (3.2.5).

Proposition 25, just above, is the key point to prove this convergence : the cluster representation (3.3.6) of the partition function Z_N and the cluster estimate (3.5.3) are the crucial elements in order to obtain in a canonical way an expansion for the measures Q_N (see [17]).
It has been explained in details in both [8] and [19] (see for instance paragraph 4.1.4 of [19], with lemma 10 and what follows) ; we give here an overview of the reasoning, adapted to our framework.

For a finite subset $\mathcal{S}$ of $\mathbb{Z}$, we associate $\mathcal{I}_\mathcal{S}$ such that $\mathcal{I}_\mathcal{S} = \cup_{k\in\mathcal{S}}\ I_k$.
We define the partition function on $\mathcal{I}$ by

$$Z_\mathcal{I} = 1 + \sum_{\tau_1\sqcup\ldots\sqcup\tau_p} \prod_{i=1}^{p} \Gamma_{\tau_i}$$

where the τ_i are defined as in subsection 3.3.2 : they are connected sets, disjoint from each other, such that $a\tau_i \subset \mathcal{I}$. Notice that $\mathcal{I}_{\{-N,\ldots,N-1\}} = I(N)$ and $Z_N = Z_{I(N)}$.

For $\mathcal{S}_1 \subset \mathcal{S}_2$, we define

$$f_{\mathcal{S}_1}^{\mathcal{S}_2} = \frac{Z_{\mathcal{I}_{\mathcal{S}_2}\setminus\bar{\mathcal{I}}_{\mathcal{S}_1}}}{Z_{\mathcal{I}_{\mathcal{S}_2}}}$$

where $\bar{\mathcal{I}}_S = \cup_{k\in\mathcal{S}}\ \bar{I}_k$ and $\bar{I}_k = I_{k-1} \cup I_k \cup I_{k+1}$.

The original version of the following lemma can be found in [17] ; here it is adapted to our needs.

Lemma 8. For β small enough,

(i) There exists a positive constant C_1, independent from $\mathcal{S}_1$ and $\mathcal{S}_2$, such that

$$|f_{\mathcal{S}_1}^{\mathcal{S}_2}| < C_1\ 2^{|\mathcal{S}_1|}. \tag{3.5.4}$$

(ii) The following assertion holds

$$f_{\mathcal{S}_1}^{\mathcal{S}_2} = 1 + \sum_{\tau_1,\ldots,\tau_p} C_{\mathcal{S}_1}(\tau_1,\ldots,\tau_p) \prod_{i=1}^{p} \Gamma_{\tau_i}, \tag{3.5.5}$$

where the sum is over every $\tau_1,\ldots,\tau_p$ for every possible integer p as defined in section 3.3.2, such that $\mathcal{I}_{\mathcal{S}_1} \cap a(\tau_1 \sqcup \ldots \sqcup \tau_p) \neq \emptyset$ and $a\tau_1 \sqcup \ldots \sqcup a\tau_p \subset \mathcal{I}_{\mathcal{S}_2}$. $C_{\mathcal{S}_1}(\tau_1,\ldots,\tau_p)$ is independent from $\mathcal{S}_2$. Furthermore, the series converges absolutely.

(iii) The expression (3.5.5) admits a limit $f_{\mathcal{S}_1}$ when $\mathcal{S}_2$ tends towards $\mathbb{Z}$ and it satisfies

$$f_{\mathcal{S}_1} = 1 + \sum_{\mathcal{I}_{\mathcal{S}_1} \cap (a\tau_1 \sqcup \ldots \sqcup a\tau_p) \neq \emptyset} C_{\mathcal{S}_1}(\tau_1,\ldots,\tau_p) \prod_{i=1}^{p} \Gamma_{\tau_i}. \tag{3.5.6}$$

(iv) There exists a positive constant C_2 such that

$$|f_{\mathcal{S}_1}^{\mathcal{S}_2} - f_{\mathcal{S}_1}| < C_2\ 2^{|\bar{\mathcal{S}}| - d(\mathcal{I}_{\mathcal{S}_1}, \mathcal{I}_{\mathcal{S}_2}^c)}$$

where $\bar{\mathcal{S}}$ is such that $\mathcal{I}_{\bar{\mathcal{S}}} = \bar{\mathcal{I}}_{\mathcal{S}_1}$.

(v) The exists a positive constant C_3 such that for any subsets $\mathcal{S}_1$ and $\hat{\mathcal{S}}_1$ of $\mathcal{S}_2$,

$$|f_{\mathcal{S}_1 \cup \hat{\mathcal{S}}_1}^{\mathcal{S}_2} - f_{\mathcal{S}_1}^{\mathcal{S}_2}\ f_{\hat{\mathcal{S}}_1}^{\mathcal{S}_2}| < C_3\ 3^{|\mathcal{S}_1| + |\hat{\mathcal{S}}_1|}\ \eta(\beta)^{d(\mathcal{I}_{\mathcal{S}_1}, \mathcal{I}_{\hat{\mathcal{S}}_1})},$$

$$|f_{\mathcal{S}_1 \cup \hat{\mathcal{S}}_1} - f_{\mathcal{S}_1}\ f_{\hat{\mathcal{S}}_1}| < C_3\ 3^{|\mathcal{S}_1| + |\hat{\mathcal{S}}_1|}\ \eta(\beta)^{d(\mathcal{I}_{\mathcal{S}_1}, \mathcal{I}_{\hat{\mathcal{S}}_1})}. \tag{3.5.7}$$

Remark 19. This is where, despite the explicit bounds obtained in (3.4.23) and (3.5.1), we must renounce to an explicit expression for the required smallness of β.

Let $\mathcal{I}$ be a finite interval and N large enough such that $\mathcal{I} \subset I(N)$. Let $F_{\mathcal{I}}$ be a $\mathcal{I}$-local bounded measurable function on Ω, i.e. for every u in Ω, $F_{\mathcal{I}}(u) = F(u_{\mathcal{I}})$. Our aim is to show that when β is small enough, the sequence $\left(\int F_{\mathcal{I}}\, dQ_N\right)_N$ converges.

Recall that

$$\int F_{\mathcal{I}}\, dQ_N = \frac{1}{Z_N} \int_\Omega F_{\mathcal{I}}(u)\ \exp(-H_N(u^{(N)}))\ \mathbb{P}(du)$$

From manipulations similar to those of section 3.3.2, one can establish that

$$\int F_{\mathcal{I}}\, dQ_N = \frac{1}{Z_N} \sum_{\tau_1,\dots,\tau_p} K_{\tau_1,\dots,\tau_p}(F_{\mathcal{I}})\ Z_{\mathcal{I}(N)\setminus(\mathcal{I}\cup\tau_1\sqcup\dots\sqcup\tau_p)}$$

with $\tau_1 \sqcup \dots \sqcup \tau_p \subset \{-N, \dots, N\}$ and where the coefficients $K_{\tau_1,\dots,\tau_p}(F_{\mathcal{I}})$ can be given explicitly, and do not depend on N.

The above expression can be written as

$$\int F_{\mathcal{I}}\, dQ_N = \sum_{\tau_1,\dots,\tau_p} K_{\tau_1,\dots,\tau_p}(F_{\mathcal{I}})\ f^{I(N)}_{\mathcal{I}\cup(\tau_1\sqcup\dots\sqcup\tau_p)}.$$

From (3.5.5), we have

$$\int F_{\mathcal{I}}\, dQ_N = \sum_{\tau_1,\dots\tau_p} K_{\tau_1,\dots,\tau_p}(F_{\mathcal{I}}) \left(1 + \sum_{\hat\tau_1,\dots,\hat\tau_q} C_{\mathcal{I}\cup(\tau_1\sqcup\dots\sqcup\tau_p)}(\hat\tau_1, \dots, \hat\tau_q) \prod_{j=1}^{q} \Gamma_{\hat\tau_j}\right)$$

with $\tau_1\sqcup\dots\sqcup\tau_p \subset \{-N, \dots, N\}$, $\hat\tau_1\sqcup\dots\sqcup\hat\tau_q \subset \{-N, \dots, N\}$ and $(\mathcal{I}\cup(\tau_1\sqcup\dots\sqcup\tau_p))\cap(\hat\tau_1\sqcup\dots\sqcup\hat\tau_q) \neq \emptyset$. From (3.5.4) and (3.5.6), we can conclude that there is absolute convergence of the series over $\hat\tau_1, \dots, \hat\tau_q$ when $\mathcal{S}_2$ converges towards $\mathbb{Z}$, so that

$$\lim_{N\to+\infty} \int F_{\mathcal{I}}\, dQ_N = \sum_{\tau_1,\dots\tau_p} K_{\tau_1,\dots,\tau_p}(F_{\mathcal{I}})\ f_{\mathcal{I}\cup(\tau_1\sqcup\dots\sqcup\tau_p)}$$

Setting $\int F_{\mathcal{I}}\, dQ := \lim_{N\to+\infty} \int F_{\mathcal{I}}\, dQ_N$, the following result holds.

Proposition 26. *Assume (H1) and (H2). For β small enough, there exists a unique stationary probability measure Q on Ω such that :*

$$Q = \lim_{N\to\infty} Q_N$$

The probability Q is the weak limit of the sequence (Q_N), in the sense that it is the limit for the topology of local convergence.

Due to the nature of the convergence in Proposition 26, properties that are satisfied by the approximations Q_N are inherited by the limit Q. Indeed, further classical results taken from Gibbs field theory (a combination of Proposition 2 and Lemma 4 in [8]) ensure that the probability measure Q is truly a weak stationary solution of the equation :

$$dX_t = \left(-\frac{1}{2}\ \nabla V(X_t) + \beta\ b((X)^t_{t-t_0})\right) dt + dW_t, \tag{3.2.5}$$

that is, under the probability measure Q, the canonical process (X_t) satisfy the stochastic system (3.2.5).

Hence our main result.

The property of exponential decorrelations, (iii) of Theorem 14,

$$\left|\int f(X_t)\, g(X_{t'})\, dQ - \int f(X_t)\, dQ \int g(X_{t'})\, dQ\right| \leqslant \theta_1\, e^{-\theta_2\, |t-t'|},$$

is a consequence of the inequality (3.5.7) and of a cluster representation for quantities of the form :

$$\int f(X_t)\, g(X_{t'})\, dQ_N - \int f(X_t)\, dQ_N \int g(X_{t'})\, dQ_N,$$

for $t, t' \in I(N)$.

As the correlations decay at an exponential rate, we have strong mixing properties, and, in particular, the central limit theorem below. Though this process is not Markovian, the proof of the following corollary is similar to that of the famous result obtained in [14] and expanded in [4].

Corollary 6. *If a smooth f is such that $\int f d\nu = 0$, then under Q,*

$$\frac{1}{\sqrt{t}} \int_0^t f(X_s)\, ds \underset{t\to+\infty}{\xrightarrow{(d)}} \mathcal{N}(0, \sigma_f^2)$$

with

$$\sigma_f^2 := 2 \int_{-\infty}^{+\infty} \mathbb{E}_Q[f(X_0)\, f(X_s)]\, ds = \int |\nabla f|^2\, d\nu$$

3.6 An example : the Ornstein-Uhlenbeck dynamics as reference process

Suppose the reference drift g is a linear one.
In order to simplify the writing of the computations, we restrict ourselves to the one-dimensional situation $d = 1$; the behaviour in higher dimensions is completely similar.

We are thus considering the one-dimensional Ornstein-Uhlenbeck process solution of :

$$dX_t = -\lambda\, X_t\, dt + \, dW_t$$

where λ is a positive parameter and (W_t) is a standard one-dimensional Brownian motion.

3.6.1 Verification of the assumptions

It is a process whose explicit expression and general behaviour are well-known ; in particular, it admits the Gaussian law $\mu = \mathcal{N}(0, 1/2\lambda)$, whose density is given by

$$\mu(dy) = \sqrt{\frac{\lambda}{\pi}}\, e^{-\lambda y^2}\, dy,$$

as its (unique) symmetric probability measure.

Furthermore, the transition density of (y_t) with respect to μ is given by

$$p(t,x,y) = \frac{1}{\sqrt{1-e^{-2\lambda t}}}\ \exp\left(-\frac{\lambda}{1-e^{-2\lambda t}}\ \left((x^2+y^2)\, e^{-2\lambda t}\ - 2\, x\, y\, e^{-\lambda t}\right)\right).$$

Thus, all the assumptions made at the beginning of section 3.2.1 are satisfied, as are hypotheses (H1) and (H2) :

Proposition 27. *In this setting, assumptions (H1) and (H2) are satisfied, with* $C_P = \dfrac{1}{2\lambda}$ *and any* $\delta > \dfrac{\ln(7)}{\lambda}$*. Furthermore, these bounds are optimal.*

Proof. Indeed, thanks to a well-known result (see, for instance, [1]) on Poincaré's inequalities verified by Gaussian measures, for a smooth function f,

$$Var_\mu(f) \leqslant \frac{1}{2\lambda} \int (f')^2\, d\mu$$

which implies that (H1) holds, with $C_P = 1/2\lambda$.

Moreover, (H2) follows from the lemma below :

Lemma 9. For every positive t and for $k \in \mathbb{N}^*$,

$$\int_{\mathbb{R}^2} p(t,x,y)^k\ \mu(dx)\ \mu(dy) = \frac{1}{(1-e^{-2\lambda t})^{k/2-1}\ \sqrt{(1+(k-1)\, e^{-2\lambda t})^2 - k^2\, e^{-2\lambda t}}} \tag{3.6.1}$$

We thus have immediately :

Corollary 7. *For every integer* k*,* $\|p(t,.,.)\|_{L^k(\mu\otimes\mu)}$ *goes to* 1 *when* t *goes to infinity, and for every* $K > 1$*, there exists* t_K *such that*

$$\sup_{t\geqslant t_K} \|p(t,.,.)\|_{L^k(\mu\otimes\mu)} \leqslant K$$

Proof. Set $I_k(t) = \displaystyle\int_{\mathbb{R}^2} p(t,x,y)^k\ \mu(dx)\ \mu(dy)$.

Then, letting $K_t = \lambda\,(1-e^{-2\lambda\,t})^{-1}$, $c_t = 1+(k-1)e^{-2\lambda\,t}$, and $d_t = k\,e^{-\lambda\,t}$,

$$\begin{aligned}
I_k(t) &= \frac{\lambda}{\pi}\,(1-e^{-2\lambda\,t})^{-k/2} \\
&\quad\times \int_{\mathbb{R}^2} \exp\Big(-\frac{\lambda}{1-e^{-2\lambda\,t}}\;\Big((1+(k-1)\,e^{-2\lambda\,t})(x^2+y^2)-2k\,x\,y\,e^{-\lambda t}\Big)\Big)\;dx\,dy \\
&= \frac{\lambda}{\pi}\,(1-e^{-2\lambda\,t})^{-k/2}\int_{\mathbb{R}^2}\exp\Big(-K_t\,c_t\,(x-\frac{d_t}{c_t}\;y)^2\Big)\;\exp\Big(-K_t\,c_t\,(1-\frac{d_t^2}{c_t^2})\;y^2\Big)\;dx\,dy \\
&= \frac{\lambda}{\pi}\,(1-e^{-2\lambda\,t})^{-k/2}\,\sqrt{\frac{\pi}{K_t\,c_t}}\,\sqrt{\frac{\pi}{K_t\,c_t\,(1-\frac{d_t^2}{c_t^2})}} \\
&= \lambda\,(1-e^{-2\lambda\,t})^{-k/2}\,\frac{1}{K_t\,\sqrt{c_t^2-d_t^2}} = (1-e^{-2\lambda\,t})^{1-k/2}\,\frac{1}{\sqrt{c_t^2-d_t^2}}
\end{aligned}$$

Hence the result we were looking for. □

In particular,

$$\|p(a,.,.)\|^8_{L^8(\mu\otimes\mu)} = (1-e^{-2\,\lambda a})^{-3}\,(1-50\,e^{-2\,\lambda a}+49\,e^{-4\,\lambda a})^{-1/2}$$

which is finite if and only $a > \dfrac{\ln(7)}{\lambda}$.

Besides, a study of the function $a \mapsto (1-e^{-2\lambda a})^{-3}\,(1-50\,e^{-2\,\lambda\,a}+49\,e^{-4\,\lambda\,a})^{-1/2}$ shows that it is decreasing towards 1 on the open interval $\Big(\dfrac{\ln(7)}{\lambda}, +\infty\Big)$.

Thus, for every $\delta > \dfrac{\ln(7)}{\lambda}$,

$$\sup_{a\geqslant\delta}\|p(a,.,.)\|_{L^8(\mu\otimes\mu)} < \infty$$

and, furthermore,

$$\sup_{a\geqslant\delta}\|p(a,.,.)\|_{L^8(\mu\otimes\mu)} = (1-e^{-2\lambda\delta})^{-3/8}\,(1-50\,e^{-2\lambda\delta}+49\,e^{-4\lambda\delta})^{-1/16} = M_\delta$$

where M_δ corresponds to the constant defined in (3.4.15). □

The perturbed equation is

$$dx_t = \Big(-\lambda x_t + \beta\,b((x)^t_{t-t_0})\Big)\,dt + \,dW_t$$

where $b : \mathbb{R}\times\mathcal{C}([-t_0,0],\mathbb{R}) \to \mathbb{R}$ is a measurable function, bounded by 1, satisfying the assumption (H3) introduced in subsection 3.2.2, and β is a positive number.

3.6.2 Numerical applications

The map $\varepsilon \mapsto \beta_\varepsilon$

We represent the function $\varepsilon \mapsto \beta_\varepsilon$, where β_ε is the bound defined in (3.4.23), taking $\delta = \dfrac{2}{\lambda}$ (as $\ln(7) \sim 1.95$).

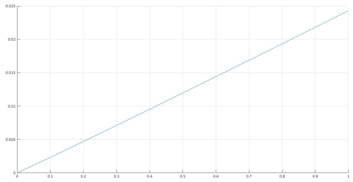

Figure 3.1: The map $\varepsilon \mapsto \bar{b}(\varepsilon)$ between 0 and 1.

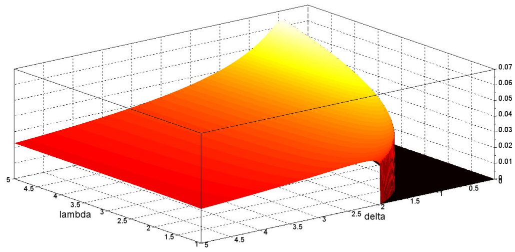

Figure 3.2: The surface $(\delta, \lambda) \mapsto B_{0.9}(\delta, \lambda)$ for $\delta \in [0, 10]$ and $\lambda \in [0, 5]$.

Then $M_\delta = (1 - e^{-4})^{-3/8}(1 - 50e^{-4} + 49e^{-8})^{-1/16} \simeq 1.16$.

This map can be seen in figure 3.1, for ε evolving between 0 and 1. This curve is not linear, and, as expected, non-decreasing : the smaller ε, the smaller β_ε, and it vanishes when ε vanishes.

Determination of β_ε

We seek to obtain the largest possible value β_ε, to have the largest possible window of choice for β satisfying Proposition 24. Indeed, one should note that its expression depends on the parameter δ, which appears in hypothesis (H2) and is not uniquely determined : in our case, any $\delta > \dfrac{\ln(7)}{\lambda}$ will do.

Consider $B_\varepsilon(.,.) = \beta_\varepsilon$ as a function of ε, δ and λ :

$$B_\varepsilon(\delta, \lambda) = \frac{\varepsilon}{2\sqrt{2e}(8!)^{1/8} M_\delta^{7/8}} \left(\delta - \frac{1}{16\lambda} \ln \left(\frac{2M_\delta^4 + 1}{16 M_\delta^8} \left(\sqrt{1 + \frac{\varepsilon^4}{32(1 + 2M_\delta^4)^2}} - 1 \right) \right) \right)^{-1/2}$$

with $M_\delta = (1 - e^{-2\delta})^{-3/8} \, (1 - 50 \, e^{-2\delta} + 49 \, e^{-4\delta})^{-1/16}$.

Set $\varepsilon = 0.9$ and $a = a_{0.9}$.
Differentiating B with respect to δ in order to find the points where the derivative vanishes, and thus the maxima of the function, looks a rather hopeless case.
We draw the map of $(\delta, \lambda) \mapsto B_{0.9}(\delta, \lambda)$ in Figure 3.2.

Looking closely at the relation between $\sup_\delta B_{0.9}(\delta, \lambda)$ and λ, one can conjecture that

$$\sup_{\delta > \frac{\ln(7)}{\lambda}} B_{0.9}(\delta, \lambda) \simeq 0.0291 \sqrt{\lambda}$$

Bibliography

[1] C. Ané, S Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques.* Société mathématique de France, 2000.

[2] M. Arriojas, Y. Hu, S. Mohammed, and G. Pap. A delayed Black and Scholes formula. *Stochastic Analysis and Applications*, 25(2):471–492, 2007.

[3] Yuri Bakhtin and Jonathan C. Mattingly. Stationary solutions of stochastic differential equations with memory and stochastic partial differential equations. *Commun. Contemp. Math.*, 7(5):553–582, 2005.

[4] P. Cattiaux, D. Chafaï, and A. Guillin. Central limit theorems for additive functionals of ergodic Markov diffusions processes. *ALEA Lat. Am. J. Probab. Math. Stat.*, 9(2):337–82, 2012.

[5] P. Cattiaux, P. Dai Pra, and S. Roelly. A constructive approach to a class of ergodic HJB equations with unbounded and nonsmooth cost. *SIAM J. Control Optim.*, 47(5):2598–2615, 2008.

[6] P. Cattiaux, A. Guillin, and P.A. Zitt. Poincaré inequality and hitting times. *Ann. Institut Henri Poincaré*, 49(1):95–118, 2013.

[7] G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems.* Cambridge University Press, Cambridge, 1996.

[8] P. Dai Pra and S. Roelly. An existence result for infinite-dimensional Brownian diffusions with non-regular and non-Markovian drift. *Markov Processes Relat. Fields*, 10:113–136, 2004.

[9] P. Dai Pra, S. Roelly, and H. Zessin. A Gibbs variatonal principle in space-time for infinite-dimensional diffusions. *Proba. Theory Relat. Fields*, 122(2):289–315, 2002.

[10] S-Y. Ha, K. Lee, and D. Levy. Emergence of time-asymptotic flocking in a stochastic Cucker-Smale system. *Commun. Math. Sci.*, 7(2):453–69, 2009.

[11] I. Ignatyuk, V.A. Malyshev, and V. Sidoravicius. Convergence of hte stochatic quantization method. I. *Theory Probab. Appl.*, 37(2):209–21, 1992.

[12] K. Itô and M. Nisio. On stationary solutions of a stochastic differential equation. *J. Math. Kyoto Univ.*, 4(1):1–75, 1964.

[13] A.F. Ivanov, Y.I. Kazmerchuk, and A.V. Swishchuk. Theory, Stochastic Stability and Applications of Stochastic Delay Differential Equations: a Survey of Recent Results. *Differential Equations and Dynamical Systems*, 11:55–115, 2003.

[14] C. Kipnis and S.R.S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19, 1986.

[15] A. Kolmogorov. Zur Umkehrbarkeit der statistischen Naturgesetze. *Mathematische Annalen*, 113(1):766–72, 1937.

[16] R.S. Lipster and A.N. Shiryayev. *Statistics of Random Processes I, General Theory.* Springer-Verlag, 1977.

[17] V.A. Malyshev and R.A. Minlos. *Gibbs Random Fields, Cluster Expansions.* Kluwer, 1991.

[18] X. Mao. *Stochastic differential equations and applications.* Woodhead Publishing, 2007.

[19] R.A. Minlos, S. Roelly, and H. Zessin. Gibbs states on space-time. *Potential Annal.*, 13(4):367–408, 2000.

[20] R.A. Minlos, A. Verbeure, and V. Zagrebnov. A quantum cristal model in the light mass limit : Gibbs states. *Rev. Math. Phys.*, 12(7):981–1032, 2000.

[21] S. E. A. Mohammed. Stability of linear delay equations under a small noise. *Proc. Edinburgh Math. Soc. (2)*, 29(2):233–254, 1986.

[22] S. Roelly and W.M. Ruszel. Propagation of Gibbsianness for Infinite-Dimensional Diffusions with Space-Time Interaction. *Markov Processes Relat. Fields*, 20:653–74, 2014.

[23] G. Royer. *Une initiation aux inégalités de Sobolev logarithmiques.* Société Mathématique de France, 1999.

[24] M. Scheutzow. Qualitative behaviour of stochastic delay equations with a bounded memory. *Stochastics*, 12:41–80, 1984.

[25] L.S. Tsimring and A. Pikovsky. Noise-induced dynamics in bistable systems with delay. *Phys. Rev. Lett.*, 87, 2001.

Chapter 4

The 2-D stochastic Keller-Segel particle model : existence and uniqueness.

This chapter, written in collaboration with Patrick Cattiaux, was published, under the same title, in ALEA, Lat. Am. J. Probab. Math. Stat. 13 (1), 447 – 463 (2016).

Abstract : We introduce a stochastic system of interacting particles which is expected to furnish, as the number of particles goes to infinity, a stochastic approach of the 2-D Keller-Segel model. In this note, we prove existence and some uniqueness for the stochastic model for the parabolic-elliptic Keller-Segel equation, for all regimes under the critical mass. Prior results for existence and weak uniqueness have been very recently obtained by Fournier and Jourdain.

Keywords : Keller-Segel model, diffusion processes, Bessel processes

Contents

4.1 Introduction and main results.

The (Patlak) Keller-Segel system introduced in [12], is a tentative model to describe chemo-taxis phenomenon, an attractive chemical phenomenon between organisms. In two dimensions, the classical 2-D parabolic-elliptic Keller-Segel model reduces to a single non linear P.D.E.,

$$\partial_t \rho_t(x) = \Delta_x \rho_t(x) + \chi \nabla_x.((K * \rho_t)\rho_t)(x) \tag{4.1.1}$$

with some initial ρ_0.
Here $\rho : \mathbb{R}_+ \times \mathbb{R}^2 \to \mathbb{R}$, $\chi > 0$ and $K : x \in \mathbb{R}^2 \mapsto \dfrac{x}{\| x \|^2} \in \mathbb{R}^2$ is the gradient of the harmonic kernel, i.e. $K(x) = \nabla \log(\| x \|)$.
It is not difficult to see that (4.1.1) preserves positivity and mass, so that we may assume that ρ_0 is a density of probability, i.e. $\rho_0 \geqslant 0$ and $\int \rho_t dx = \int \rho_0 dx = 1$.
With this choice, (4.1.1) is written in a non-dimensional form. In order to compare it to the usual formulation, the reader can think that the parameter χ is actually given by

$$\chi \, = \, \chi_0 \, \frac{\alpha \, m}{2\pi \, D}$$

where χ_0 is the chemotactic sensitivity, α is the rate of production of chemoattractant by the cells, m is the total mass and D is the product of the diffusivities.
As usual, ρ is modeling a density of cells, and $c_t = K * \rho_t$ is (up to some constant) the concentration of chemo-attractant.

A very interesting property of such an equation is a blow-up phenomenon

Theorem 4.1.2. Assume that $\rho_0 \log \rho_0 \in \mathbb{L}^1(\mathbb{R}^2)$ and that $(1+ \| x \|^2)\rho_0 \in \mathbb{L}^1(\mathbb{R}^2)$. Then if $\chi > 4$, the maximal time interval of existence of a classical solution of (4.1.1) is $[0, T^*)$ with

$$T^* \, \leqslant \, \frac{1}{2\pi \, \chi \, (\chi - 4)} \int \| x \|^2 \; \rho_0(x) \, dx \, .$$

If $\chi \leqslant 4$ then $T^* = +\infty$.

For this result, a wonderful presentation of what Keller-Segel models are and an almost up to date state of the art, we refer to the unpublished HDR document of Adrien Blanchet (available on Adrien's webpage [1]). We also apologize for not furnishing a more complete list of references on the topic, where beautiful results were obtained by brilliant mathematicians. But the present paper is intended to be a short note.

Actually (4.1.1) is nothing else but a Mc Kean-Vlasov type equation (non linear Fokker-Planck equation if one prefers), involving a potential which is singular at 0. Hence one can expect that the movement of a typical cell will be given by a non-linear diffusion process

$$\begin{aligned} dX_t &= \sqrt{2} \, dB_t \, - \, \chi \, (K * \rho_t)(X_t) \, dt \, , \\ \rho_t(x) \, dx &= \mathcal{L}(X_t) \, , \end{aligned} \tag{4.1.3}$$

where $\mathcal{L}(X_t)$ denotes the distribution of probability of X_t. A natural possible approach of (4.1.3) is through the limit, as N goes to infinity, of a linear system of stochastic differential equations in mean field interactions given for $i = 1, ..., N$ by

$$dX_t^{i,N} = \sqrt{2}\, dB_t^{i,N} - \frac{\chi}{N} \sum_{j \neq i}^{N} \frac{X_t^{i,N} - X_t^{j,N}}{\| X_t^{i,N} - X_t^{j,N} \|^2}\, dt\,, \tag{4.1.4}$$

for a well chosen initial distribution of the $X_0^{\cdot,N}$. Here the $B_\cdot^{i,N}$ are for each N independent standard 2-D Brownian motions. Under some exchangeability assumptions, it is expected that the distribution of any particle (say $X^{1,N}$) converges to a solution of (4.1.3) as $N \to \infty$, yielding a solution to (4.1.1). This strategy (including the celebrated propagation of chaos phenomenon) has been well known for a long time. One can see [16] for bounded and Lipschitz potentials, [15, 4] for unbounded potentials connected with the granular media equation.

The goal of the present note is the study of existence, uniqueness and non explosion for the system (4.1.4). That is, this is the very first step of the whole program we have described previously. Moreover we will see how the N-particle system is feeling the blow-up property of the Keller-Segel equation.

(4.1.4) can be viewed as a "modern" formulation of the microscopic description given by Keller and Segel themselves in [13]. The main difficulty is of course that the potentials explode when two particles are colliding. For such singular potentials very few is known.
Fournier, Hauray and Mischler [5] have tackled the case of the 2-D viscous vortex model, corresponding to $K(x) = \frac{x^\perp}{\| x \|^2}$ for which no blow-up phenomenon occurs. In the same spirit the sub-critical Keller-Segel model corresponding to $K(x) = \frac{x}{\| x \|^{2-\varepsilon}}$ for some $\varepsilon > 0$ is studied in [8]. The methods of both papers are close, and mainly based on some entropic controls. These methods seem to fail for the classical Keller-Segel model we are looking at. However, during the preparation of the manuscript, we received the paper by N. Fournier and B. Jourdain [6], who prove existence and some weak uniqueness by using approximations. Though some intermediate results are the same, we shall here give a very different and much direct approach, at least for existence and some uniqueness. However, we shall use one result in [6] to prove a more general uniqueness result. Also notice that a similar model (but with a different treatment after collisions) was studied from a numerical point of view in [9] and a theoretical one in [10].

It can also be noticed that when we replace the attractive potential K by a repulsive one (say $-K$), we find models connected with random matrix theory (like the Dyson Brownian motion).

Our main theorem in this paper is the following

Theorem 4.1.5. Let $M = \{$there exists at most one pair $i \neq j$ such that $X^i = X^j\}$. Then,

- for $N \geqslant 4$ and $\chi < 4\left(1 - \frac{1}{N-1}\right)$, there exists a unique (in distribution) non explosive solution of (4.1.4), starting from any $x \in M$. Moreover, the process is strong Markov, lives in M and admits a symmetric σ-finite, invariant measure given by

$$\mu(dX^1, ..., dX^N) = \Pi_{1 \leqslant i < j \leqslant N} \| X^i - X^j \|^{-\frac{\chi}{N}}\; dX^1 ... dX^N\,,$$

- for $N \geqslant 2$, if $\chi > 4$, the system (4.1.4) does not admit any global solution (i.e. defined on the whole time interval $\mathbb{R}^+$),

- for $N \geqslant 2$, if $\chi = 4$, either the system (4.1.4) explodes or the N particles are glued in finite time.

Let us explain a little bit more on the meaning of this statement. A (weak) solution of (4.1.4) up to a stopping time T and starting from x is a Probability measure $\mathbb{P}_x$ on the state of continuous paths from $\mathbb{R}^+$ into $(\mathbb{R}^2)^N$ such that for all $t > 0$,

$$M_t^{i,N} = X_{t\wedge T}^{i,N} - x^i - \int_0^{t\wedge T} \frac{\chi}{N} \sum_{j\neq i}^{N} \frac{X_s^{i,N} - X_s^{j,N}}{\| X_s^{i,N} - X_s^{j,N} \|^2} \, ds$$

is well defined and is a martingale with brackets

$$\langle M^{i_k,N}, M^{j_l,N} \rangle_t = 2\,(t \wedge T)\, \delta_{i_k = j_l} \,.$$

(Recall that $x^i = (x^{i_1}, x^{i_2})$). The supremum of all stopping times such that this property holds is the explosion time or the lifetime, we shall denote by ξ in the sequel.
The first part of the Theorem thus tells us that for any $x \in M$ such a solution $\mathbb{P}_x$ exists with $\xi = +\infty$, $\mathbb{P}_x$ almost surely. In addition the hitting time T_{M^c} of the complement of M is $\mathbb{P}_x$ almost surely infinite too. Finally, the family $(\mathbb{P}_x)_{x\in M}$ is a strong Markov family on $C^0(\mathbb{R}^+, M)$ that admits μ as a symmetric measure.
When $\chi \geqslant 4$ we shall not describe the explosion time ξ, but we shall prove that, denoting by T_N the first time of collision of all the N particles, $\xi \wedge T_N < +\infty$. If $\chi > 4$ we will see that a solution cannot exist after time T_N, hence $T_N \wedge \xi = \xi < +\infty$. For $\chi = 4$ the situation is more intricate but on $T_N < t < \xi$ (which can be empty) we will see that all the particles are glued (i.e. $X_t^{i,N} = X_t^{j,N}$ for all i and j). The situation for $4\left(1 - \dfrac{1}{N-1}\right) \leqslant \chi < 4$ is still more intricate, but some results are contained in [6].

The proof of this theorem is partly "pathwise", based on comparisons between one dimensional diffusion processes and the behavior of squared Bessel processes, partly based on Dirichlet forms theory and partly based on an uniqueness result for 2 dimensional skew Bessel processes obtained in [6]. The latter is only used to get rid of a non allowed polar set of starting positions which appears when using Dirichlet forms.
Recall that a set E is polar if for all $x \in M$ the hitting time T_E of E defined as

$$T_E = \inf\{t > 0\,,\, X_t \in E\}$$

is $\mathbb{P}_x$ almost surely infinite.

Since later we will be interested in the limit $N \to +\infty$, this theorem is in a sense optimal: for $\chi < 4$ we have no asymptotic explosion while for $\chi > 4$ the system explodes. Also notice that in the limiting case $\chi = 4$, we have (at least) explosion for the density of the stochastic system and not for the equation (4.1.1).
The remaining part of the whole program will be the aim of future works.

4.2 Study of the system (4.1.4).

Most of the proofs in this section will use comparison with squared Bessel processes. Let us recall some basic results on these processes.

Definition 4.2.1. Let $\delta \in \mathbb{R}$. The unique strong solution (up to some explosion time τ) of the following one dimensional stochastic integral equation

$$Z_t = z + 2 \int_0^t \sqrt{Z_s}\, dB_s + \delta\, t\,,$$

is called the (generalized) squared Bessel process of dimension δ starting from $z \geqslant 0$.

In general squared Bessel processes are only defined for $\delta \geqslant 0$, that is why we used the word generalized in the previous definition. For these processes the following properties are known

Proposition 4.2.2. Let Z be a generalized squared Bessel process of dimension δ. Let τ_0 the first hitting time of the origin.

- If $\delta < 0$, then τ_0 is almost surely finite and equal to the explosion time,
- if $\delta = 0$, then $\tau_0 < +\infty$ and $Z_t = 0$ for all $t \geqslant \tau_0$ almost surely,
- if $0 < \delta < 2$, then $\tau_0 < +\infty$ almost surely and the origin is instantaneously reflecting, i.e starting from 0 the hitting time of $]0, +\infty[$ is almost surely equal to 0,
- if $\delta \geqslant 2$, then the origin is polar (hence $\tau_0 = +\infty$ almost surely).

For all this see [17] chapter XI, proposition 1.5.

Now come back to (4.1.4). For simplicity we skip the index N in the definition of the process. Since all coefficients are locally Lipschitz outside the set

$$A = \Big\{\text{there exists (at least) a pair } i \neq j \text{ such that } X^i = X^j\Big\}$$

and bounded when the distance to A is bounded from below, the only problem is the one of collisions between particles. As usual we denote by ξ the lifetime of the process. For simplicity we also assume, for the moment, that the starting point does not belong to A, so that the lifetime is almost surely positive.

For $2 \leqslant k \leqslant N$ we define $K = \{1, ..., k\}$ and $\bar{K}^2 = \{(i,j) \in K^2 | i \neq j\}$. We shall say that a k-collision occurs at (a random) time T if $X^i_T = X^j_T$ for all $(i,j) \in \bar{K}^2$, $X^l_T \neq X^i_T$ for all $l > k$. Of course, there is no lack of generality when looking at the first k indices, and we can also assume that at this peculiar time T, any other collision involves at most k other particles.
In what follows we denote $D^{i,j} = X^i - X^j$, and

$$Z^k = \sum_{(i,j) \in \bar{K}^2} \parallel D^{i,j} \parallel^2 \ .$$

Of course a k-collision occurs at time T if and only if $T < \xi$ and $Z^k_T = 0$.

Let us study the process Z^k. Applying Ito's formula we get on $t < \xi$

$$\begin{aligned} Z_t^k &= Z_0^k + 2\sqrt{2} \int_0^t \sum_{(i,j)\in\bar{K}^2} D_s^{i,j}\,(dB_s^i - dB_s^j) + 4k(k-1)\left(2 - \frac{\chi}{N}\right) t \\ &\quad - \frac{2\chi}{N} \int_0^t \sum_{(i,j)\in\bar{K}^2} D_s^{i,j} \sum_{\substack{l=1\\ l\neq i,j}}^N \left(\frac{D_s^{i,l}}{\| D_s^{i,l} \|^2} + \frac{D_s^{l,j}}{\| D_s^{l,j} \|^2} \right) ds\,. \end{aligned} \tag{4.2.3}$$

We denote

$$\begin{aligned} dM_s^k &= \sum_{(i,j)\in\bar{K}^2} D_s^{i,j}\,(dB_s^i - dB_s^j) \\ E_s^k &= \sum_{(i,j)\in\bar{K}^2} D_s^{i,j} \sum_{\substack{l=1\\ l\neq i,j}}^N \left(\frac{D_s^{i,l}}{\| D_s^{i,l} \|^2} + \frac{D_s^{l,j}}{\| D_s^{l,j} \|^2} \right) \end{aligned}$$

the martingale part and the non-constant drift part.

4.2.1 Investigation of the martingale part M^k.

Let us compute the martingale bracket, using the immediate $D^{i,l} = -D^{l,i}$ and $D^{i,l} + D^{l,j} = D^{i,j}$.

$$\begin{aligned} d < M^k >_s &= \sum_{\substack{(i,j)\in\bar{K}^2\\ (l,m)\in\bar{K}^2}} < D_s^{i,j}\,(dB_s^i - dB_s^j), D_s^{l,m}\,(dB_s^l - dB_s^m) > \\ &= \sum_{\substack{(i,j)\in\bar{K}^2\\ (l,m)\in\bar{K}^2}} D_s^{i,j}\, D_s^{l,m}\,(\delta_{il} - \delta_{im} - \delta_{jl} + \delta_{jm})\, ds \\ &= \sum_{(i,j)\in\bar{K}^2} D_s^{i,j} \left(\sum_{\substack{m\in K\\ m\neq i}} D_s^{i,m} + \sum_{\substack{l\in K\\ l\neq i}} D_s^{i,l} + \sum_{\substack{l\in K\\ l\neq j}} D_s^{l,j} + \sum_{\substack{m\in K\\ m\neq j}} D_s^{m,j} \right) ds \\ &= 2 \sum_{(i,j)\in\bar{K}^2} D_s^{i,j} \left(\sum_{\substack{m\in K\\ m\neq i}} D_s^{i,m} + \sum_{\substack{m\in K\\ m\neq j}} D_s^{m,j} \right) ds \\ &= 2k \sum_{(i,j)\in\bar{K}^2} \| D_s^{i,j} \|^2 \end{aligned}$$

i.e. finally

$$d < M^k >_s = 2k\, Z_s^K\, ds\,. \tag{4.2.4}$$

According to Doob's representation theorem (applied to $\mathbb{1}_{s<\xi}\, d < M^k >_s$), there exists (on an extension of the initial probability space) a one dimensional Brownian motion W^k such that almost surely for $t < \xi$

$$2\sqrt{2} \int_0^t \sum_{(i,j)\in\bar{K}^2} D_s^{i,j}\,(dB_s^i - dB_s^j) \;=\; 4\sqrt{k} \int_0^t \sqrt{Z_s^k}\, dW_s^k\,. \tag{4.2.5}$$

4.2.2 Reduction of the drift term.

In order to study the drift term E_t^k we will divide it into two sums: the first one, C_t^k taking into consideration the i and j in K, i.e. the pair of particles which will be directly involved in the eventual k-collision; the other one R_t^k, involving the remaining indices.

More precisely $E_t^k = C_t^k + R_t^k$ with

$$\begin{aligned} C_t^k &= \sum_{(i,j)\in\bar K^2} \sum_{\substack{l\in K\\ l\neq i,j}} D_t^{i,j} \left(\frac{D_t^{i,l}}{\| D_t^{i,l} \|^2} + \frac{D_t^{l,j}}{\| D_t^{l,j} \|^2} \right), \\ R_t^k &= \sum_{(i,j)\in\bar K^2} \sum_{l=k+1}^{N} D_t^{i,j} \left(\frac{D_t^{i,l}}{\| D_t^{i,l} \|^2} + \frac{D_t^{l,j}}{\| D_t^{l,j} \|^2} \right). \end{aligned}$$

We will deal with R_t^k later. First we ought to simplify the expression of C_t^k. Indeed

$$\begin{aligned} C_t^k &= \sum_{(i,j)\in\bar K^2} \sum_{\substack{l\in K\\ l\neq i,j}} D_t^{i,j} \left(\frac{D_t^{i,l}}{\| D_t^{i,l} \|^2} + \frac{D_t^{l,j}}{\| D_t^{l,j} \|^2} \right) \\ &= \sum_{(i,l)\in\bar K^2} \left(\frac{D_t^{i,l}}{\| D_t^{i,l} \|^2} \sum_{\substack{j\in K\\ j\neq i,l}} D_t^{i,j} \right) + \sum_{(j,l)\in\bar K^2} \left(\frac{D_t^{l,j}}{\| D_t^{l,j} \|^2} \sum_{\substack{i\in K\\ i\neq j,l}} D_t^{i,j} \right) \\ &= \sum_{\substack{(i,l)\in\bar K^2\\ i<l}} \left(\frac{D_t^{i,l}}{\| D_t^{i,l} \|^2} \sum_{\substack{j\in K\\ j\neq i,l}} D_t^{i,j} + \frac{D_t^{l,i}}{\| D_t^{l,i} \|^2} \sum_{\substack{j\in K\\ j\neq i,l}} D_t^{l,j} \right) + \\ &\qquad + \sum_{\substack{(j,l)\in\bar K^2\\ j<l}} \left(\frac{D_t^{l,j}}{\| D_t^{l,j} \|^2} \sum_{\substack{i\in K\\ i\neq j,l}} D_t^{i,j} + \frac{D_t^{j,l}}{\| D_t^{j,l} \|^2} \sum_{\substack{j\in K\\ i\neq j,l}} D_t^{i,l} \right) \end{aligned}$$

so that using again $D^{l,i}D^{l,j} = D^{i,l}D^{j,l}$ the latter is still equal to

$$= \sum_{\substack{(i,l)\in\bar K^2\\ i<l}} \frac{D_t^{i,l}}{\| D_t^{i,l} \|^2} \left(\sum_{\substack{j\in K\\ j\neq i,l}} (D_t^{i,j} + D_t^{j,l}) \right) + \sum_{\substack{(j,l)\in\bar K^2\\ j<l}} \frac{D_t^{l,j}}{\| D_t^{i,l} \|^2} \left(\sum_{\substack{i\in K\\ i\neq j,l}} (D_t^{i,j} + D_t^{l,i}) \right)$$

and using again $D^{i,l} + D^{l,j} = D^{i,j}$, we finally arrive at

$$C_t^k \;=\; 2\,(k-2) \;\times\; \#\{(i,l) \in K^2 | i < l\} \;=\; k(k-1)(k-2)\,. \tag{4.2.6}$$

4.2.3 Back to the process Z^k.

With the results obtained in (4.2.5) and (4.2.6) we may simplify (4.2.3), writing (still on $t < \xi$)

$$Z_t^k = Z_t^0 + 4\sqrt{k} \int_0^t \sqrt{Z_s^k}\, dW_s^k + 2k(k-1)\left(4 - \frac{k\chi}{N}\right) t - \frac{2\chi}{N} \int_0^t R_s^k\, ds\,. \tag{4.2.7}$$

Hence defining

$$V_t^k = \frac{1}{4k}\, Z_t^k$$

the process V^k satisfies

$$dV_t^k = 2\sqrt{V_t^k}\, dW_t^k + (k-1)\left(2 - \frac{k\chi}{2N}\right) dt - \frac{\chi}{2kN} R_t^k\, dt\,, \tag{4.2.8}$$

i.e. can be viewed as a perturbation of a squared Bessel process of dimension

$$\delta = (k-1)\left(2 - \frac{k\chi}{2N}\right)\,,$$

we shall denote by U^k in the sequel.

4.2.4 The case of an N-collision.

If $k = N$, $R^N = 0$ so that V^N is exactly the squared Bessel process of dimension $\frac{N-1}{2}\,(4-\chi)$. Hence, according to proposition 4.2.2

- if $\chi > 4$ there is explosion in finite time for the process V^N (hence for X also),
- if $\chi = 4$, there is an almost sure N-collision in finite time, and then all the particles are glued, provided no explosion occurred before for the process X,
- if $4\left(1 - \frac{1}{N-1}\right) < \chi < 4$ there is an almost sure N-collision in finite time, provided no explosion occurred before for the process X,
- if $\chi \leqslant 4\left(1 - \frac{1}{N-1}\right)$ there is almost surely no N-collision (before explosion).

In particular we see that the particle system immediately feels the critical value $\chi = 4$, in particular explosion occurs in finite time as soon as $\chi > 4$.
For $4\left(1 - \frac{1}{N-1}\right) < \chi < 4$ we know that V^N is instantaneously reflected once it hits the origin, but it does not indicate whether all or only some particles will separate (we only know that they are not all glued). Notice that when $N = 2$ this condition reduces to $0 < \chi < 4$, and then both particles are separated. Hence in this very specific case, there is no explosion (for the distance between both particles) in finite time almost surely, but there are always 2-collisions.

4.2.5 Towards non explosion for $\chi \leqslant 4\left(1 - \frac{1}{N-1}\right)$.

As we said before, the lifetime ξ is greater than or equal to the first multiple collision time T. Since we shall consider V^k as a perturbation of U^k, what happens for the latter ?

- for $\chi > \frac{4N}{k}$, U^k reaches 0 in finite time a.s. and then explosion occurs,
- for $\chi = \frac{4N}{k}$, U^k reaches 0 and is sticked,
- for $\frac{4N}{k} > \chi > \frac{4N}{k}\left(1 - \frac{1}{k-1}\right)$, U^k reaches 0 and is instantaneously reflected,
- for $\chi \leqslant \frac{4N}{k}\left(1 - \frac{1}{k-1}\right)$, U^k does not hit 0 in finite time a.s.

Lemma 4.2.9. For all $3 \leqslant k \leqslant N$, it holds

$$\frac{4N}{k}\left(1-\frac{1}{k-1}\right) \geqslant 4\left(1-\frac{1}{N-1}\right).$$

Proof. Introduce the function

$$u \mapsto g(u) = \frac{4N}{u}\left(1-\frac{1}{u-1}\right)$$

defined for $u > 1$. Then

$$g'(u) = \frac{4N}{u(u-1)}\left(\frac{2-u}{u}+\frac{1}{u-1}\right)$$

is negative on $[2+\sqrt{2}, +\infty[$, so that the lemma is proved for $N \geqslant k \geqslant 4$. For $k = 3$, it amounts to $\dfrac{N}{6} \geqslant \dfrac{N-2}{N-1}$ which is true for all $N \geqslant 3$ (with equality for $N = 3$ and $N = 4$). $\square$

In particular, since $\chi \leqslant 4\left(1-\dfrac{1}{N-1}\right)$, U^k never hits 0 for $3 \leqslant k \leqslant N$, while it reaches 0 but is instantaneously reflected for $k = 2$. What we expect is that the same occurs for V^k.

In order to prove it, let T be the first multiple collision time. With our convention (changing indices if necessary) there exists some $2 \leqslant k \leqslant N$ such that T is the first k-collision time T^k. Note that this does not prevent other k'-collisions (with $k' \leqslant k$) possibly at the same time T for the particles with indices larger than $k+1$. But as we will see this will not change anything. The reasoning will be the same but the conclusion completely different for $k = 2$ and for $k \geqslant 3$.

No k-collisions for $k \geqslant 3$.

Introduce, for $\varepsilon > 0$, the random set

$$A_\varepsilon^k = \{T = T^k < +\infty \text{ and } \inf_{i\in K\,,\, l\geqslant k+1} \inf_{t\leqslant T} \parallel D_t^{i,l} \parallel \geqslant 2\varepsilon\}.$$

It holds

$$\{T = T^k < +\infty\} = \bigcup_{\varepsilon \in 1/\mathbb{N}} A_\varepsilon^k.$$

In particular if $\mathbb{P}(T = T^k < +\infty) > 0$ there exists some $\varepsilon > 0$ so that $\mathbb{P}(A_\varepsilon^k) > 0$. We shall see that this is impossible when $k \geqslant 3$.

Indeed recall that

$$R_t^k = \sum_{(i,j)\in \bar{K}^2} \sum_{l=k+1}^{N} D_t^{i,j}\left(\frac{D_t^{i,l}}{\parallel D_t^{i,l}\parallel^2}+\frac{D_t^{l,j}}{\parallel D_t^{l,j}\parallel^2}\right).$$

So on A^k_ε, we have, for $t \leqslant T$,

$$
\begin{aligned}
|R^k_t| &\leqslant \sum_{(i,j)\in \bar{K}^2} \sum_{l=k+1}^{N} \| D^{i,j}_t \| \left(\frac{1}{\| D^{i,l}_t \|^2} + \frac{1}{\| D^{l,j}_t \|^2} \right) \\
&\leqslant \sum_{(i,j)\in \bar{K}^2} \| D^{i,j}_t \| \; \frac{N-k}{\varepsilon} \\
&\leqslant \frac{N-k}{\varepsilon} \sqrt{k(k-1)} \sqrt{Z^k_t}
\end{aligned}
$$

the latter being a consequence of Cauchy-Schwarz inequality. Thus on A^k_ε, for $t \leqslant T$

$$
|R^k_t| \leqslant \frac{2}{\varepsilon} (N-k)k \sqrt{k-1} \sqrt{V^k_t} \,. \tag{4.2.10}
$$

Hence, on A^k_ε for $t \leqslant T$ the drift b^k (which is not Markovian) of V^k_t satisfies

$$
b^k \geqslant \hat{b}^k(v) = \frac{(k-1)}{2} \left(4 - \frac{N\chi}{k}\right) - \frac{2}{\varepsilon} (N-k)k \sqrt{k-1} \sqrt{v} \,. \tag{4.2.11}
$$

In particular for any $\theta > 0$,

$$
\hat{b}^k(v) \geqslant \frac{(k-1)}{2} \left(4 - \frac{N\chi}{k}\right) - \theta
$$

provided v is small enough. Thus the hitting time of the origin for the process with drift $\hat{b}^k$ is larger than the one for the corresponding squared Bessel process (thanks to well known comparison results between one dimensional Ito processes, see e.g. [11] Chap.6, Thm 1.1), and since this holds for all θ, finally is larger than the one of U^k. But as we already saw, U^k never hits the origin for $3 \leqslant k$. Using again the comparison theorem (this time with b^k and $\hat{b}^k(v)$), V^k does not hit the origin in finite time on A^k_ε which is in contradiction with $\mathbb{P}(A^k_\varepsilon) > 0$.

About 2-collisions.

Actually all we have done in the previous sub subsection is unchanged for $k = 2$, except the conclusion. Indeed U^2 reaches the origin but is instantaneously reflected. So V^2 (on A^2_ε) can reach the origin too, but is also instantaneously reflected. Actually using that

$$
b^k(v) \leqslant \bar{b}^k(v) = \frac{(k-1)}{2} \left(4 - \frac{N\chi}{k}\right) + \frac{2}{\varepsilon} (N-k)k \sqrt{k-1} \sqrt{v} \,,
$$

together with the Feller's explosion test, it is easily seen that V^2 will reach the origin with a (strictly) positive probability (presumably equal to one, but this is not important for us).

But this instantaneous reflection is not enough for the non explosion of the process X, because X^i is $\mathbb{R}^2$ valued. Before going further in the construction, let us notice another important fact: there are no multiple 2-collisions at the same time, i.e. starting from A^c the process lives in M at least up to the explosion time ξ. Of course this is meaningful provided $N \geqslant 4$.

To prove the previous sentence, first look at

$$
Y_t = \| D^{1,2}_t \|^2 + \| D^{3,4}_t \|^2 \,,
$$

assuming that $Y_T = 0$ and that no other 2-collision happens at time T. It is easily seen that (just take care that we had an extra factor 2 in our definition of Z_t^k)

$$\begin{aligned} Y_t &= Y_0 + 2\sqrt{2} \int_0^t \left(D_s^{1,2} \, (dB_s^1 - dB_s^2) + D_s^{3,4} \, (dB_s^3 - dB_s^4) \right) + 8 \left(2 - \frac{\chi}{N} \right) t \\ &\quad - \frac{\chi}{N} \int_0^t \left(D_s^{1,2} \sum_{\substack{l=1 \\ l \neq 1,2}}^{N} \left(\frac{D_s^{1,l}}{\parallel D_s^{1,l} \parallel^2} + \frac{D_s^{l,2}}{\parallel D_s^{l,2} \parallel^2} \right) + D_s^{3,4} \sum_{\substack{l=1 \\ l \neq 3,4}}^{N} \left(\frac{D_s^{3,l}}{\parallel D_s^{3,l} \parallel^2} + \frac{D_s^{l,4}}{\parallel D_s^{l,4} \parallel^2} \right) \right) ds \, , \end{aligned}$$

so that, defining $V_t = Y_t/4$ we get that

$$dV_t = 2 \sqrt{V_t} \, dW_t + 2 \left(2 - \frac{\chi}{N} \right) dt + \, R_t \, dt$$

where R_t is a remaining term we can manage just as we did for Z_t^k. Since for $N \geqslant 4$, $2 \left(2 - \frac{\chi}{N} \right) \geqslant 2$, V_t, hence Y_t does not hit the origin. Notice that if we consider $k (\geqslant 2)$ 2-collisions, the same reasoning is still true, just replacing 4 by $2k$, the final equation being unchanged except for R_t.

Non explosion.

According to all what precedes what we need to prove is the existence of the solution of (4.1.4) with an initial configuration satisfying $X_0 = x$ with $x^1 = x^2$, all other coordinates being different and different from $x^1 = x^2$. Indeed, on $\xi < +\infty$, $X_\xi \in \delta M$ the set of particles with exactly two glued particles, so that if we can prove that starting from any point of δM, we can build a strong solution on an interval $[0, S]$ for some strictly positive stopping time S, it will show that $\xi = +\infty$ almost surely. However we will not be able to prove the existence of such a strong solution. Actually we think that it does not exist. We will thus build some weak solution and show uniqueness in some specific sense.

This will be the goal of the next sections.

4.3 Building a solution.

4.3.1 Existence of a weak solution.

Writing

$$M = \cup_{i<j} \cap_{k \neq l \, ; \, l \neq i,j} \, \{ X^k \neq X^l \}$$

we see that M is an open subset of $\mathbb{R}^{2N}$.

Recall that $x \in \delta M$ means that exactly two coordinates coincide (say $x^1 = x^2$), all other coordinates being distinct and distinct from x^1. We may thus define

$$d_x = \min\{ i \geqslant 3 \, ; i \neq j \, ; \; j = 1, ..., N \, ; \parallel x^i - x^j \parallel \} \, > \, 0 \, ,$$

so that

$$\Omega_x = \Pi_{j=1}^N \, B(x^j, d_x/2) \, \subset \, M \, , \tag{4.3.1}$$

and points $y \in \Omega_x \cap \delta M$ will satisfy $y^1 = y^2$. If $x \notin \delta M$, we may similarly define $d_x = \min\{i \neq j\,;\, j = 1, ..., N\,;\, \| x^i - x^j \|\}$ and then Ω_x. In all cases the balls $B(.,.)$ are the open balls. Now if K is some compact subset of M we can cover K by a finite number of sets Ω_x, so that for any measure μ, a function g belongs to $\mathbb{L}^1_{loc}(M, \mu)$ if and only if $g \in \mathbb{L}^1(\Omega_x, \mu)$ for all x in M.

The natural measure to be considered is

$$\mu(dX^1, ..., dX^N) = \Pi_{1 \leqslant i < j \leqslant N} \ \| X^i - X^j \|^{-\frac{\chi}{N}} \ dX^1...dX^N \,, \tag{4.3.2}$$

since it is, at least formally, the symmetric measure for the system (4.1.4).

It is clear that for $x \notin \delta M$, μ is a bounded measure on Ω_x. When $x \in \delta M$, say that $x^1 = x^2$ and perform the change of variables

$$Y^1 = X^1 - X^2 \quad , \quad Y^2 = X^1 + X^2 \,.$$

In restriction to Ω_x, μ can be written

$$\mu(dX^1, ..., dX^N) = C(N, x) \ \| Y^1 \|^{-\frac{\chi}{N}} \ dY^1 \, dY^2 \, dX^3 ... dX^N \,,$$

hence is a bounded measure on Ω_x provided $\chi < 2N$ just looking at polar coordinates for Y^1. In this case it immediately follows that μ is a σ finite measure on M. Also remark that if f is compactly supported by K and belongs to $\mathbb{L}^2(d\mu)$ then it belongs to $\mathbb{L}^2(dX)$ and

$$\int_K f^2 \, dX \leqslant \sup_K (\| Y^1 \|^{\frac{\chi}{N}}) \int_K f^2 \, d\mu \,.$$

But we can say much more.

To this end consider the symmetric form

$$\mathcal{E}(f, g) = \int_M < \nabla f, \nabla g > \, d\mu \quad , \quad f, g \in C_0^\infty(M) \,. \tag{4.3.3}$$

First we check that this form is closable in the sense of [7]. To this end it is enough to show that it is a closable form when restricted to functions $f, g \in C_0^\infty(\Omega_x)$ for all $x \in M$. If $x \notin \delta M$ the form is equivalent to the usual scalar product on square Lebesgue integrable functions, so that it is enough to look at $x \in \delta M$.

Hence let f_n be a sequence of functions in $C_0^\infty(\Omega_x)$, converging to 0 in $\mathbb{L}^2(d\mu)$ and such that ∇f_n converges to some vector valued function g in $\mathbb{L}^2(d\mu)$. What we need to prove is that g is equal to 0. To this end consider a vector valued function h which is smooth and compactly supported in $\Omega_x \cup \{d(., \delta M) > \varepsilon\}$ for some $\varepsilon > 0$. Then a simple integration by parts shows that

$$\int < g, h > \, d\mu = \lim_n \int < \nabla f_n, h > \, d\mu = \lim_n \int f_n \, H \, d\mu$$

for some $H \in \mathbb{L}^2(d\mu)$ so is equal to 0. Hence g vanishes almost surely on $\Omega_x \cup \{d(., \delta M) > \varepsilon\}$, for all $\varepsilon > 0$, so that g is μ-almost everywhere equal to 0.

By construction, $\mathcal{E}$ is regular and local. Hence, its smallest closed extension $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a Dirichlet form, which is in addition regular and local. According to Theorem 6.2.2. in [7], there exists a μ-symmetric diffusion process $X_.$ whose form is given by $\mathcal{E}$. Notice that, integrating by parts, we see that the generator of this diffusion process coincides with the generator L given by

$$L = \sum_{i=1}^{N} \Delta_{x^i} - \frac{\chi}{N} \sum_{i=1}^{N} \left(\sum_{i \neq j} \frac{x^i - x^j}{\| x^i - x^j \|^2} \right) \nabla_{x^i} \tag{4.3.4}$$

for the functions f in $C_0^\infty(M)$ such that $Lf \in \mathbb{L}^2(\mu)$. This is a core for the domain $D(L)$. The Dirichlet form theory tells us that once $f \in D(L)$, $f(X_t) - f(X_0) - \int_0^t Lf(X_s)ds$ is a $\mathbb{P}_x$ martingale for quasi every starting point x, i.e. for all x out of some subset E of M which is of zero μ-capacity.

But remark that for any $x \in M - E$ and $t > 0$, the transition kernel $p_t(x,.)$ of the Markov semi-group is absolutely continuous with respect to μ. Indeed using the local Malliavin calculus as in [3] (or elliptic standard results), this transition kernel has a (smooth) density w.r.t. Lebesgue measure (hence w.r.t. μ) on each open subset of $M \cap \{d(.,\delta M) > \varepsilon\}$ for any $\varepsilon > 0$. Hence if $\mu(A) = 0$, $\mu(A \cap \{d(.,\delta M) > \varepsilon\}) = 0$ for all $\varepsilon > 0$ so that $p_t(x, A \cap \{d(.,\delta M) > \varepsilon\}) = 0$ and finally using monotone convergence, $p_t(x, A) = 0$.
Since $p_t(x,.)$ is absolutely continuous w.r.t. μ, we deduce from Theorem 4.3.4 in [7] that the sets of zero μ capacity are exactly the polar sets for the process.

Note that the function $x \mapsto x$ does not belong to $D(L)$, so that we cannot use the previous result. Nevertheless

Lemma 4.3.5. Assume that $\chi < N$. Then for all $x \in M - E$ and all $i = 1, ..., N$,

$$X_t^i - X_0^i - \int_0^t \frac{\chi}{N} \left(\sum_{i \neq j} \frac{X_s^i - X_s^j}{\| X_s^i - X_s^j \|^2} \right) ds$$

is a $\mathbb{P}_x$ martingale, and actually is $\mathbb{P}_x$ almost surely equal to $\sqrt{2}\, B_t^i$ for some Brownian motion.

Proof. To prove the lemma, for all $x \in M - E$ it is enough to show the martingale property starting from x up to the exit time $S(x)$ of Ω_x (since $X_{S(x)} \in M - E$ because E is polar, see the discussion above). In the sequel, for notational convenience, we do not write the exit time $S(x)$ (all times t have to be understood as $t \wedge S(x)$) and we simply write M instead of Ω_x.

To show this result it is actually enough to look locally in the neighborhood of a point $x \in \delta M$ such that $x^1 = x^2$, and with our previous notation to look at both coordinates of $y = x^1 - x^2$. Indeed $x \mapsto x^1 + x^2$ belongs (at least locally) to $D(L)$ as well as all other coordinates x^j for $j \geqslant 3$.

Let $g^j(x) = y_j$ for $j = 1, 2$ be the coordinate application of $y = x^1 - x^2$. Clearly $Lg^j \in \mathbb{L}^p(\mu)$ for $p < 2 - \dfrac{\chi}{N}$, hence belongs to $\mathbb{L}^1$ thanks to our assumption on χ/N. Introduce the function defined on $\mathbb{R}$ by,

$$h_\varepsilon(u) = \sin^2\left(\frac{\pi\, u}{2\varepsilon}\right) \mathbf{1}_{|u| \leqslant \varepsilon} + \mathbf{1}_{|u| > \varepsilon}\,.$$

h is of C^2 class except at $|u| = \varepsilon$. Now define $v_\varepsilon(x) = g^1(x)\, h_\varepsilon(g^1(x))$. We have

$$L(v_\varepsilon)(x) = \left[4h'_\varepsilon(g^1) + 2\, g^1\, h''_\varepsilon(g^1) - \frac{\chi}{N} \left(2h_\varepsilon(g^1)\, \frac{g^1}{|g^1 + g^2|^2} + 2h'_\varepsilon(g^1) +\, R_\varepsilon \right) \right](x)$$

the remaining term R_ε corresponding to the interactions with particles x^j for $j \geqslant 3$.

For all $\varepsilon > 0$, v_ε thus belongs to $D(L)$ and $v_\varepsilon(X_t) - v_\varepsilon(X_0) - \int_0^t Lv_\varepsilon(X_s)ds$ is a $\mathbb{P}_x$ martingale, with brackets $4\int_0^t |\nabla v_\varepsilon(X_s)|^2 \, ds$ for all $x \in M - E$.

But it is easily seen that Lv_ε converges to Lg^1 in $\mathbb{L}^1(\mu)$ as $\varepsilon \to 0$. Since v_ε converges to g^1 in $\mathbb{L}^1(\mu)$ too, we deduce that

$$\mathbb{E}_\mu \left(g^1(X_{t+h}) - g^1(X_t) - \int_t^{t+h} Lg^1(X_s)ds | \mathcal{F}_t \right) = 0 \quad \text{for all } t \geqslant 0 \text{ and } h \geqslant 0, \tag{4.3.6}$$

where $\mathcal{F}_t$ denotes the natural filtration on the probability space, since the same property is true for v_ε. Similarly the brackets converge to $4t$. Since the same holds for g^2, we get the desired result $\mathbb{P}_\mu$ almost surely. Actually this result holds true $\mathbb{P}_x$ almost surely for μ almost all $x \in M$.

But since $p_t(x,.)$ is absolutely continuous w.r.t. μ for $t > 0$, it immediately follows using the Markov property at time t, that (4.3.6) is true $\mathbb{P}_x$ a.s. for all $x \in M - E$, but only for $t > 0$. Hence for all $x \in M - E$ and all $t > 0$, $N_t^s = g^1(X_{t+s}) - g^1(X_t) - \int_t^{t+s} Lg^1(X_u)du$ is a martingale defined on $[t, +\infty[$, whose bracket is given by $4(s-t)$, i.e. is (2 times) a Brownian motion. In particular for a fixed t, $(N_t^s)_{0<s\leqslant t}$ is bounded in $\mathbb{L}^2(\mathbb{P}_x)$. Up to a sub-sequence it is thus weakly convergent in $\mathbb{L}^2(\mathbb{P}_x)$ as $s \to 0$ so that $N_t^0 = N_t$ is well defined $\mathbb{P}_x$ a.s., and satisfies (4.3.6) for all $t \geqslant 0$ this time. Thus it is a martingale with a linear bracket, i.e. 2 times a Brownian motion. □

The previous lemma shows that the diffusion $X_.$ is simply the μ symmetric solution of (4.1.4) killed when it hits the boundary ∂M.

Assume in addition that $\chi \leqslant 4\left(1 - \frac{1}{N-1}\right)$. Then the previous diffusion process never hits ∂M since the latter is exactly the set where either some k-collision occurs for some $k \geqslant 3$ or at least two 2-collisions occur at the same time. So it is actually the unique μ-symmetric Markov diffusion defined on $\bar{M}$ solving (4.1.4). Indeed we could associate to any markovian extension of $(\mathcal{E}, C_0^\infty(M))$ another diffusion process, which would coincide with the previous one up to the hitting time of ∂M which is almost surely infinite. We have thus obtained

Theorem 4.3.7. Assume that $\chi \leqslant 4\left(1 - \frac{1}{N-1}\right)$ and that $N \geqslant 4$.
Then there exists a unique μ-symmetric (see (4.3.2)) diffusion process $(X_t, \mathbb{P}_x)$ (i.e. a Hunt process with continuous paths), defined for $t \geqslant 0$ and $x \in M - E$ where $E \subset M$ is polar (or equivalently of μ capacity equal to 0) such that for all $f \in C_0^\infty(\mathbb{R}^{2N})$,

$$f(X_t) - f(x) - \int_0^t Lf(X_s)ds$$

is a $\mathbb{P}_x$ martingale (for the natural filtration) with L given by (4.3.4). Furthermore $X_.$ lives in M (never hits ∂M).

Proof. As for the previous lemma, it is enough to work locally in the neighborhood of the points in δM and to look at the new particles $(y = x^1 - x^2, z = x^1 + x^2, x^3, ..., x^N)$. Let $f \in C_0^\infty(M)$ be written in these new coordinates. Using a Taylor expansion in y (z and all the others x^j being fixed) and the fact that if the partial derivatives at $y = 0$ of a smooth function of y

are vanishing, then this function belongs to the domain of the generator, we see that proving the martingale property for f amounts to the corresponding martingale property for smooth functions g written as $g(y,z,x^j) = y\,h(z,x^j)$ i.e. amounts to the previous lemma (and of course the remaining particles for which there is no problem).

It remains to extend the martingale property we proved to hold for $f \in C_0^\infty(M)$ to $f \in C_0^\infty(\mathbb{R}^{2N})$. Take $f \in C_0^\infty(\mathbb{R}^{2N})$ and define S_ε as the first time the distance $d(X_., \partial M)$ is less than ε. Then replacing f by some $f_\varepsilon \in C_0^\infty(M)$ which coincides with f on $d(y,\partial M) \geqslant \varepsilon$, we see that $f(X_{t\wedge S_\varepsilon}) - f(x) - \int_0^{t\wedge S_\varepsilon} Lf(X_s)ds$ is a $\mathbb{P}_x$ martingale. Since S_ε growths to infinity the conclusion follows from Lebesgue theorem. □

Remark 4.3.8. The main disadvantage of the previous construction is that it is not explicit and that it does not furnish a solution starting from all $x \in M$ but only for all x except those in some unknown polar set. In particular, proving the regularity of the Markov transition kernels up to δM requires additional work. The advantage is that if we require μ-symmetry, we get uniqueness of the diffusion process.

This Theorem is to be compared with Theorem 7 in [6], where existence of a weak solution is shown by using approximation and tightness, in the same $\chi \leqslant 4\left(1-\frac{1}{N-1}\right)$ case (take care of the normalization of χ which is not the same here and therein). Note that the result in [6] is concerned with existence starting from some initial absolutely continuous density and does not furnish a diffusion process. ♢

4.3.2 Existence and uniqueness of a weak solution.

In this subsection we assume that $\chi \leqslant 4\left(1-\frac{1}{N-1}\right)$ and that $N \geqslant 4$. Our aim is to build a solution starting from any point in M, i.e. to get rid of the polar set E in the previous subsection. The construction will be very similar (still using Dirichlet forms) but we shall here use one result in [6], namely the uniqueness result for a 2 dimensional Bessel process.

We start with an important lemma

Lemma 4.3.9. Let $\mathbb{P}_x$ be the solution of (4.1.4) built in Theorem 4.3.7 and starting from some allowed point x. Then

$$\int_0^{+\infty} \mathbb{1}_{\delta M}(X_s)\,ds \;=\; 0\,, \quad \mathbb{P}_x \ a.s.$$

Proof. We can cover δM by an enumerable union of Ω_y $(y \in \delta M)$. It is thus immediate that the lemma will be proved once we prove that

$$\int_0^{+\infty} \mathbb{1}_{\delta M \cap \Omega_y}(X_s)\,ds \;=\; 0\,, \quad \mathbb{P}_x \ a.s.$$

But we have seen in the previous section that, when the process is in some Ω_y (where say $y^1 = y^2$), the process $\parallel D_t^{1,2} \parallel^2$ is larger than or equal to the square of a Bessel process U_t of index δ strictly between 0 and 2. But (see [17] proof of proposition 1.5 p.442), the time spent at the origin by the latter is equal to 0, i.e. $\int_0^{+\infty} \mathbb{1}_{U_s=0}\,ds = 0$ almost surely. The same necessarily holds for $D^{1,2}$, hence the result. □

We intend now to prove some uniqueness, when starting from a point in δM. Actually, using some standard tools of concatenation of paths, it is enough to look at the behavior of our process starting at some $y \in \delta M$ with $y^1 = y^2$, up to the exit time of Ω_y (or some open non empty subset of Ω_y). In this case the only difficulty is to control the pair $(X^1_., X^2_.)$ since all other coordinates are defined through smooth coefficients. Of course writing

$$D_t^{1,2} = X_t^1 - X_t^2 \quad , \quad S_t^{1,2} = X_t^1 + X_t^2$$

we have that

$$dD_t^{1,2} = 2\, dW_t^1 - \frac{2\chi}{N} \frac{D_t^{1,2}}{\parallel D_t^{1,2} \parallel^2} dt + b_1(X_t)\, dt \tag{4.3.10}$$

and

$$dS_t^{1,2} = 2\, dW_t^2 + b_2(X_t)\, dt \tag{4.3.11}$$

where b_1 and b_2 are smooth functions (in Ω_y), W^1 and W^2 being two independent 2 dimensional Brownian motions.

Define $\bar{\Omega}_y$ as we defined Ω_y (see (4.3.1), but replacing $d_y/2$ by $d_y/4$ and consider a smoothed version η of the indicator of $\bar{\Omega}_y$ i.e. a smooth non negative function such that

$$\mathbb{I}_{\bar{\Omega}_y} \leqslant \eta \leqslant \mathbb{I}_{\Omega_y} .$$

We may extend all coefficients (except $D/ \parallel D \parallel^2$) as smooth compactly supported functions outside Ω_y, and replace $D/ \parallel D \parallel^2$ by $\eta(X)\, D/ \parallel D \parallel^2$. If we can show uniqueness for this new system we will have shown uniqueness up to the exit time of $\bar{\Omega}_y$ for (4.3.10), (4.3.11) and the remaining part of the initial system.

Hence our problem amounts to the following one: prove uniqueness for $Y = (D, S, \bar{X}) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^{2(N-2)}$ solution of

$$\begin{aligned} dD_t &= 2\, dW_t - \frac{2\chi}{N} \frac{D_t}{\parallel D_t \parallel^2} dt + b(D_t, S_t, \bar{X}_t)\, dt\,, \\ dS_t &= 2\, dW'_t + b'(D_t, S_t, \bar{X}_t)\, dt\,, \\ d\bar{X}_t &= \sqrt{2}\, d\bar{B}_t + \bar{b}(D_t, S_t, \bar{X}_t)\, dt\,, \end{aligned} \tag{4.3.12}$$

where $b, b', \bar{b}$ are smooth and compactly supported in $\mathbb{R}^{2N}$.

Thus, after a standard Girsanov transform, we are reduced to prove uniqueness for

$$\begin{aligned} dD_t &= 2\, dW_t - \frac{2\chi}{N} \frac{D_t}{\parallel D_t \parallel^2} dt, \\ dS_t &= 2\, dW'_t\,, \\ d\bar{X}_t &= \sqrt{2}\, d\bar{B}_t\,, \end{aligned} \tag{4.3.13}$$

hence for $D_.$. $U_. = D_./2$ is some type of 2-dimensional skew Bessel process with dimension $\chi/2N$ (see [2] for the one dimensional version). Its squared norm $|U_.|^2$ is a squared Bessel process of dimension $\delta = 2 - \frac{\chi}{N}$, so that the origin is not polar for the process $D_.$.

As we did in the previous sub-section, we can directly prove the existence and uniqueness of a symmetric Hunt process (here the reference measure is $|D|^{-\chi/N}\, dD$) using the associated Dirichlet form, and since the origin is not polar, we know the existence of a solution starting from $D_0 = 0$. Here we only need $\chi < N$, but for the whole construction our initial assumption on χ is required. Finally we can check that the occupation time formula of Lemma 4.3.9 is still true.

But as before, if now we have existence starting from every initial point, we only have uniqueness in the sense of symmetric Markov processes. To get weak uniqueness we can use polar coordinates: the squared norm is a squared Bessel process, so that strong uniqueness holds (with the corresponding dimension we are looking at); the polar angle is much tricky to handle. This is the main goal of Lemmata 19 and 20 in [6], and the final weak uniqueness then follows from the proof of Theorem 17 in [6] and the occupation time formula.

Remark 4.3.14. It can be noticed that this result is out of reach of the method developed by Krylov and Röckner in [14] for a general Brownian motion plus drift b, since it requires that $b \in \mathbb{L}^p(dX)$ for some $p > 2$. Also notice that one cannot use standard Girsanov transform for solving (4.3.13), since for a 2-dimensional Brownian motion starting from the origin,

$$\int_0^t \frac{1}{|B_s|^2}\, ds = +\infty \quad a.s. \quad \text{for all } t > 0\,,$$

see [17]. ♢

Remark 4.3.15. Of course our construction of a solution is quite abstract and one should ask about the behavior of D_{0^+} in (4.3.12) when $D_0 = 0$. Since we do not have a strong solution, this question is out of reach at least rigorously. But since the solution starting from 0 is rotation invariant (this is easily seen), one can imagine that the particle starting from 0 will uniformly choose an angle in $[0, 2\pi]$ and start a (constant times) Bessel process of dimension $2 - (\chi/N)$ in this direction for an "infinitesimal" time $t = 0^+$. Of course for the initial process we also have to add the drift coming from the Girsanov transformation. This leads to a time discretization procedure which is different in nature from [9]. ♢

Bibliography

[1] A. Blanchet. Méthodes variationnelles appliquées en Biologie et en Economie. Document Habilitation à Diriger des Recherches. Available on Adrien's home page, 2012.

[2] S. Blei. On symmetric and skew Bessel processes. *Stoch. Proc. Appl.*, 122(9):3262–3287, 2012. MR2946442.

[3] P. Cattiaux. Hypoellipticité et hypoellipticité partielle pour les diffusions avec une condition frontière. *Ann. Inst. Henri Poincaré. Prob. Stat.*, 22(1):67–112, 1986. MR0838373.

[4] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non uniformly convex case. *Probab. Theory Relat. Fields*, 140(1):19–40, 2008. MR2357669.

[5] N. Fournier, M. Hauray, and S. Mischler. Propagation of chaos for the 2D viscous vortex model. *J. Eur. Math. Soc. (JEMS)*, 16(7):1423–1466, 2014. MR3254330.

[6] N. Fournier and B. Jourdain. Stochastic particle approximation of the Keller-Segel equation and two-dimensional generalization of Bessel processes. Available on ArXiv 1507.01087. [PR], 2015.

[7] M. Fukushima. *Dirichlet forms and Markov processes.* North Holland, 1980. MR0569058.

[8] D. Godinho and C. Quininao. Propagation of chaos for a sub-critical Keller-Segel model. *Ann. Inst. Henri Poincaré. Prob. Stat.*, 51(3):965–992, 2015. MR3365970.

[9] J. Haskovec and C. Schmeiser. Stochastic particle approximation for measure valued solutions of the 2D Keller-Segel system. *J. Stat. Phys.*, 135(1):133–151, 2009. MR2505729.

[10] J. Haskovec and C. Schmeiser. Convergence analysis of a stochastic particle approximation for measure valued solutions of the 2D Keller-Segel system. *Comm. PDE*, 36(6):940–960, 2011. MR2765424.

[11] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes.* North Holland, 1981. MR0637061.

[12] E. Keller and L. A. Segel. Initiation of slime mold aggregation viewed as an instability. *J. theor. Biol.*, 26:399–415, 1970.

[13] E. Keller and L. A. Segel. Model for chemotaxis. *J. theor. Biol.*, 30:225–234, 1971.

[14] N. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Relat. Fields*, 131(2):154–196, 2005. MR2117951.

[15] F. Malrieu. Logarithmic Sobolev inequalities for some nonlinear PDE's. *Stochastic Process. Appl.*, 95(1):109–132, 2001. MR1847094.

[16] S. Méléard. Asymptotic behaviour of some interacting particle systems: Mc Kean-Vlasov and Boltzmann models. *Probabilistic models for non linear partial differential equations. ed. Talay and Tubaro, Lecture Notes in Mathematics*, 1627:42–95, 1995. MR1431299.

[17] D. Revuz and M. Yor. *Continuous martingales and Brownian motion.* Springer, 1999. MR1725357.