



HAL
open science

Interacting run-and-tumble particles as piecewise deterministic Markov processes : invariant distribution and convergence

Léo Hahn

► **To cite this version:**

Léo Hahn. Interacting run-and-tumble particles as piecewise deterministic Markov processes: invariant distribution and convergence. Probability [math.PR]. Université Clermont Auvergne, 2024. English. NNT: 2024UCFA0084. tel-04792665

HAL Id: tel-04792665

<https://theses.hal.science/tel-04792665v1>

Submitted on 20 Nov 2024

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

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



THÈSE DE DOCTORAT DE
L'UNIVERSITÉ CLERMONT AUVERGNE

Opérée au sein du :

Laboratoire de mathématiques Blaise Pascal

École Doctorale 178
Sciences Fondamentales

Spécialité de doctorat : Mathématiques appliquées

Soutenue publiquement le 17/09/2024, par :

Léo HAHN

Interacting Run-and-Tumble Particles as
Piecewise Deterministic Markov Processes
Invariant Distribution and Convergence

Devant le jury composé de :

M. Florent MALRIEU Professeur des universités, Université de Tours	Rapporteur
M. Grigorios PAVLIOTIS Professor, Imperial College London	Rapporteur
M. Arnaud GUILLIN Professeur des universités, Université Clermont Auvergne	Directeur de thèse
Mme Manon MICHEL Chargée de recherche, Université Clermont Auvergne	Directrice de thèse
M. Richard BLYTHE Professor, University of Edinburgh	Examineur
M. Thierry BODINEAU Directeur de recherche, Institut des Hautes Études Scientifiques	Examineur
M. Jonathan MATTINGLY Professor, Duke University	Examineur
Mme Marielle SIMON Professeure des universités, Université Lyon 1	Présidente du jury
M. Michel BENAÏM Professor, Université de Neuchâtel	Invité

Abstract

This thesis investigates the long-time behavior of run-and-tumble particles (RTPs), a model for bacteria's moves and interactions in out-of-equilibrium statistical mechanics, using piecewise deterministic Markov processes (PDMPs). The motivation is to improve the particle-level understanding of active phenomena, in particular motility induced phase separation (MIPS). The invariant measure for two jamming RTPs on a 1D torus is determined for general tumbling and jamming, revealing two out-of-equilibrium universality classes. Furthermore, the dependence of the mixing time on model parameters is established using coupling techniques and the continuous PDMP model is rigorously linked to a known on-lattice model. In the case of two jamming RTPs on the real line interacting through an attractive potential, the invariant measure displays qualitative differences based on model parameters, reminiscent of shape transitions and universality classes. Sharp quantitative convergence bounds are again obtained through coupling techniques. Additionally, the explicit invariant measure of three jamming RTPs on the 1D torus is computed. Finally, hypocoercive convergence results are extended to RTPs, achieving sharp L^2 convergence rates in a general setting that also covers kinetic Langevin and sampling PDMPs.

Résumé

Cette thèse étudie le comportement en temps long des particules run-and-tumble (RTPs), un modèle pour les bactéries en physique statistique hors équilibre, en utilisant des processus de Markov déterministes par morceaux (PDMPs). La motivation est d'améliorer la compréhension au niveau particulaire des phénomènes actifs, en particulier la séparation de phase induite par la motilité (MIPS). La mesure invariante pour deux RTPs avec jamming sur un tore 1D est déterminée pour mécanismes de tumble et jamming généraux, révélant deux classes d'universalité hors équilibre. De plus, la dépendance du temps de mélange en fonction des paramètres du modèle est déterminée en utilisant des techniques de couplage et le modèle continu PDMP est rigoureusement relié à un modèle sur réseau connu. Dans le cas de deux RTPs avec jamming sur la droite réelle et interagissant à travers un potentiel attractif, la mesure invariante présente des différences qualitatives en fonction des paramètres du modèle, rappelant des transitions de forme et des classes d'universalité. Des taux de convergence fins sont à nouveau obtenus par des techniques de couplage. Par ailleurs, la mesure invariante explicite de trois RTPs se bloquant sur le tore 1D est calculée. Enfin, les résultats de convergence hypocoercive sont étendus aux RTPs, obtenant ainsi des taux de convergence L^2 fins dans un cadre général qui couvre également les PDMPs utilisés pour l'échantillonnage et Langevin cinétique.

Bicycle, bicycle, bicycle
“Bicycle Race” *by* Queen

Remerciements

On a trop rarement l'occasion de remercier les personnes qui nous entourent et nous soutiennent. Je n'ai pas fait cette thèse seul et je suis heureux de pouvoir exprimer ma gratitude envers toutes celles et ceux qui m'ont accompagné dans cette aventure.

Tout d'abord merci à mes directeurs de thèse, Arnaud et Manon, pour le sujet qui aurait pu être fait sur mesure, pour tout ce que vous m'avez appris, pour les opportunités que vous m'avez données, pour vos encouragements et votre soutien dans les moments difficiles. Je vous dois beaucoup.

Ich möchte mich ganz herzlich bei Andreas Eberle und Francis Lörler für die bereichernde Zusammenarbeit bedanken.

Merci à mes professeurs de m'avoir transmis leur passion. Merci à Mme Schepers de m'avoir fait découvrir les maths et de m'avoir tant aidé. Merci à Olivier Hénard d'avoir cru en moi. Merci à Dominique Hulin pour sa gentillesse et ses cours d'une clarté lumineuse. Merci à Reika Fukuizumi de m'avoir fait découvrir les probabilités. Merci à Quentin Berger pour ses conseils bienveillants.

I would also like to thank Michel Benaïm, Richard Blythe, Thierry Bodineau, Florent Malrieu, Jonathan Mattingly, Grigorios Pavliotis and Marielle Simon for being part of my thesis committee and taking the time to review my work.

Eh bonjour! Et surtout merci à mes co-bureaux pour l'expérience partagée pendant ces trois ans. Je suis heureux d'avoir fait ce bout de chemin avec vous. Merci Athina de m'avoir tout de suite intégré quand je suis arrivé à Clermont et d'avoir tant fait pour la vie du labo et la bonne ambiance chez les doctorants. Merci Baptiste de m'avoir appris la différence entre le lion léopardé et le léopard lionné. Merci Sue pour les ankylosaures et l'analyse fonctionnelle, on reparlera des semigroupes! Merci Tristan pour les jeux de mots et les discussions sur les éditeurs de texte. Merci Hoa pour les repas partagés. Merci Rémi de ne pas m'en vouloir de parler pendant les films. Merci Aho d'avoir lu mon futur dans les cartes, j'espère qu'il ne se réalisera jamais.

Un énorme merci à tous les doctorants du LMBP pour votre soutien et votre amitié! Merci Vincent pour ton goût vestimentaire sûr et ton anecdote sur le village de Mirefleurs. Merci Arhur pour toutes les cigarettes que je t'ai taxées! Merci Martin TN pour ta bonne humeur, promis, je t'appelle la prochaine fois que je suis à Barcelone! Merci Martin PAS pour les discussions sur les stats fréquentistes et bayésiennes. Merci Florian d'avoir surmonté ton appréhension de la piscine. Merci Julian pour episode, framed et les jeux du New York Times. Merci Émilien pour ton humour engagé.

Merci Richard, Maxime, Loulou, César et Abel pour les jeux et les conseils de vie.

Merci aux copains de l'impro pour les spectacles et les fous rires. Merci Max, Claire, Jean-Baptiste, Charlotte, Joris, Benoît, Julie, Benjamin, Marceau, Jean-Bernard et Agnès!

Merci Raphaël pour l'horoscope de la machine à café, les repas à l'INMG et ton soutien quand je choisis du jambon. Merci Orane pour la peluche pigeon, le BBQ végétarien et les concerts. Vous êtes des énormes Shaka Ponk! Merci Laurent, Christian, Emiliano, Madeleine et Garance d'être encore là huit ans après.

Merci à mes parents et à mon frère pour leur soutien inconditionnel et leurs encouragements constants, merci à mon grand-père Jean-Michel pour avoir toujours cru en moi.

Merci Alice d'être dans ma vie. Te rencontrer est la plus belle chose qui me soit arrivée ces trois dernières années.

Contents

1	Introduction	7
1.1	Background	7
1.2	PDMPs	10
1.2.1	Examples	10
1.2.2	Probability semigroup and generator	13
1.3	Invariant measure	14
1.3.1	Fokker-Planck equation	16
1.3.2	Laplace transform	17
1.3.3	Stopping time	18
1.4	Convergence	19
1.4.1	Diffusions: an important class of examples	19
1.4.2	Couplings	19
1.4.3	Uniform minorization condition	20
1.4.4	Foster-Lyapunov criterion	21
1.4.5	Poincaré inequality	23
1.5	Hypoconvexity	24
1.5.1	An example where Poincaré fails	24
1.5.2	A hypoconvex approach	24
1.6	Contributions	26
1.6.1	Two run-and-tumble particles on the 1D torus	26
1.6.2	Two run-and-tumble particles on the real line	32
1.6.3	Three run-and-tumble particles on the 1D torus	35
1.6.4	Lifts with sticky boundaries and general velocity dynamics	37
1.7	Perspectives	39
2	Introduction en français	40
2.1	Revue bibliographique	40
2.2	PDMPs	43
2.2.1	Exemples	43
2.2.2	Semigroupe et générateur	46
2.3	Mesure invariante	48
2.3.1	Équation de Fokker-Planck	49
2.3.2	Transformée de Laplace	50
2.3.3	Temps d'arrêt	51
2.4	Convergence	52
2.4.1	Les processus de diffusion: une classe importante d'exemples	52
2.4.2	Couplage	53
2.4.3	Condition de minoration uniforme	54
2.4.4	Foster-Lyapunov	55
2.4.5	Inégalité de Poincaré	56
2.5	Hypoconvexité	57
2.5.1	Un exemple sans Poincaré	57
2.5.2	Une approche hypoconvexe	58

2.6	Contributions	59
2.6.1	Deux particules run-and-tumble sur le tore 1D	60
2.6.2	Deux particules run-and-tumble sur la droite réelle	66
2.6.3	Trois particules run-and-tumble sur le tore 1D	69
2.6.4	Lifts et hypocoercivité	70
2.7	Perspectives	73
3	RTP universality classes	74
3.1	Introduction	74
3.2	RTPs as PDMPs	75
3.3	Constraints	77
3.4	Universality classes	78
3.5	Discussion	79
3.6	Conservation conditions	80
3.7	Detailed-jamming	81
3.8	Global-jamming	83
3.8.1	Bulk evolution equation	83
3.8.2	Spectral analysis of bulk matrix B	84
3.8.3	General global-jamming solution	86
3.8.4	Case of a single non-null eigenvalue	89
4	Two RTPs on the 1D torus	91
4.1	Introduction	91
4.1.1	Introduction	91
4.1.2	Models and main results	92
4.2	PDMP formulation	96
4.3	Continuous space limit	99
4.3.1	Convergence in law	99
4.3.2	Convergence of the discrete invariant measures	105
4.4	Mixing times	109
4.4.1	Lower bound	109
4.4.2	Upper bound for the instantaneous tumble process	112
4.4.3	Upper bound for the finite tumble process	113
4.5	Invariant probability	117
4.5.1	Symmetries of the invariant measure	117
4.5.2	Invariant measure of the continuous process	118
5	Two RTPs on the line	121
5.1	Introduction	121
5.1.1	Model	123
5.1.2	Main results	125
5.2	Invariant measures	127
5.2.1	Instantaneous linear process	127
5.2.2	Finite linear process	129
5.2.3	Instantaneous harmonic process	136
5.3	Convergence	138
5.3.1	Convergence of the instantaneous and finite linear process	138
5.3.2	Convergence of the instantaneous harmonic process	144
5.3.3	Instantaneous linear process: lower bound on the total variation distance to the invariant measure	146
5.3.4	Finite linear process: lower bound on the total variation distance to the invariant measure	148

6	Three RTPs	152
6.1	Introduction	152
6.2	Model description	152
6.3	State space reduction	155
6.4	Invariant measure	158
7	Hypocoercivity via lifts	160
7.1	Introduction	160
7.2	Lift and collapse	162
7.2.1	Definition	162
7.2.2	Examples	163
7.2.3	Rate of convergence to equilibrium of lift/collapse	165
7.3	Main results	166
7.3.1	Divergence lemma and space-time property of the lift	169
7.3.2	Proof of the flow Poincaré inequality	171
7.3.3	Proof of Lemma 7.9	173
7.4	Applications	180
7.4.1	Underdamped Langevin equation	180
7.4.2	PDMP Sampling algorithm	182
7.4.3	Run-and-Tumble particles with sticky walls or jamming	185

Chapter 1

Introduction

1.1 Background

Run-and-tumble particles (RTPs) perform a series of straight ‘runs’ separated by random reorientations known as ‘tumbles’. The result is a stochastic movement that is very different from diffusions and closely models the motion of bacteria such as *E. Coli* [Ber04]. Because they consume energy at the particle level to achieve this persistent movement, bacteria are driven out of thermodynamical equilibrium, offering a simple model to study phenomena arising in out-of-equilibrium statistical mechanics such as collective motion [VZ12], pattern formation [BB95], accumulation at boundaries [EG15] and motility induced phase separation (MIPS) [CT15]. Understanding this last phenomenon, namely the separation into a dense and a dilute phase, is the initial motivation of this manuscript. The basic mechanism behind it is the following: persistent particles with repulsive interactions slow down in regions with high particle density, which in turn increases the density in these parts of the space. This creates a feedback loop that leads to the formation of clusters, even though the interactions are purely repulsive [CT15]. We refine this broad picture using the mathematical framework of piecewise deterministic Markov processes (PDMPs) [Dav93], a class of stochastic processes combining deterministic motion and discrete random jumps. The related question of convergence speed is addressed utilizing the many mathematical tools designed to study the long-time behavior of Markov processes.

An important preliminary step to understand MIPS is the study of a single run-and-tumble particle and how it differs from its passive, diffusing counterparts, especially in terms of invariant measure. In [Ang17] the invariant measure of a single one-dimensional RTP confined in an interval is found to have Dirac mass at the boundary. This is a strong form of accumulation near boundaries, a feature shared by most single RTP systems. In [MJK⁺18, Fry21] the presence of additional thermal noise weakens the accumulation and the Diracs broaden and become exponential terms peaked near the boundary. Interestingly, there are qualitative changes in the invariant measure of a RTP in the confining potential $V(x) = \alpha|x|^p$ ($p > 1$) depending on the model parameters [DKM⁺19, BMR⁺20]. For certain sets of parameters, the invariant measure diverges at the edges of the state space, a strong form of out-of-equilibrium accumulation. For other parameter choices, a shape transition occurs and the invariant measure vanishes at the edges, making it closer to that of passive particles.

Potential signs of MIPS are seen on the invariant measure of two discrete interacting RTPs [SEB16], which has increased mass near jamming configuration in the form of decaying exponentials and Diracs. In the continuous-space limit, where the lattice spacing goes to zero, the exponential terms become additional Diracs indicating strong effective attraction. When an internal state is added to this discrete-space model to take into account the non-zero duration of tumbling [SEB17], a second decaying exponential appears and survives in the continuous space-limit. Finally, when the discrete-space particles instantly jump apart on collisions [MEB23, MEB22], the effective interaction can be both attractive, i.e. in line with clustering, or repulsive depending on model parameters. In each case, the model is

tractable because it reduces to the single particle setting. Indeed, it suffices to consider the relative particle, which is again an active particle albeit displaying a more intricate velocity dynamics. This reduction transforms clustering into the already discussed accumulation at boundaries. Working directly in continuous space leads to decaying exponentials in the presence of thermal noise [DDK20a, LDMS21] and to Dirac terms in its absence [LDMS21]. Two continuous RTPs subjected to the harmonic interaction potential reduce to the single RTP model in [BMR⁺20] and thus display the same vanishing or diverging behavior at the edges of the space.

A common feature in these RTP models is that the invariant measure frequently consists of a combination of uniform, exponential, and Dirac terms [SEB16, SEB17, Ang17, EM18, MJK⁺18, LDMS21, Fry21]. This is explained by the fact that, except at the boundaries, the continuity equations take the form of a linear ODE so the density is given by a matrix exponential. Establishing the correct boundary conditions is a recurring challenge, especially in the zero temperature and jamming interaction case, prompting the development of various approaches. First, one can discretize space and consider the limit where the lattice spacing goes to zero. Second, one can add thermal noise and let the strength of the noise go to zero. Finally, one can consider increasingly peaked short range repulsive potentials. The common theme here is the approximation by another system where the boundary conditions are easier to understand. However, except in terms of boundary conditions, these approximation turn out to be more complicated than the original process and induce an additional step where an appropriate limit has to be computed. The mathematical formalism of PDMPs solves this issue and allows the direct study of continuous-space models with zero thermal noise and jamming by using the generator of the process. This new point of view unlocks many invariant measures and allows us to identify two out-of-equilibrium universality classes for jamming RTPs with general tumbling and jamming mechanisms [HGM23]. These classes correspond to either detailed or global probability flow preservation, indicating the proximity to equilibrium and determining the behavior of the invariant measure through a simple criterion. The PDMP framework also allows us to rigorously prove that continuous-space models are the limit of on-lattice models. Furthermore, we find the invariant distribution of three jamming RTPs on the 1D torus, an encouraging step towards solving the many-particle case, which remains an open problem.

Once the invariant measure is known, the question at which time scale it becomes relevant arises naturally. Indeed, at very short time scales, the process is strongly influenced by its initial distribution. The question ‘when does the asymptotic behavior take over?’ is thus of crucial importance. For this, one can turn to the spectral decomposition of the generator: the asymptotic exponential convergence rate is given by the second eigenvalue. In the discrete setting, the generator is simply a matrix and its complete spectral decomposition is provided in the case of a single and of two interacting RTPs [MBE19]. In continuous space, the generator is an unbounded operator and the eigenvalue problem takes the form of a differential equation [MJK⁺18, DDK20a]. In cases where the Laplace transform of fixed-time distributions can be computed [AG19, DKM⁺19], the eigenvalues can instead be obtained as the poles of the Laplace transform. In other cases, the exponential convergence speed can be evaluated by approximating the fixed-time distribution [SSK20].

These spectral decompositions often contain a numerical component, which prevents them from being fully explicit. Another drawback of the spectral approach is that the diagonalization of unbounded operators is technical and the eigenvalue problem only yields a subset of the spectrum. Furthermore, in the non-reversible case, one can only hope to obtain the convergence speed $Ce^{-\lambda t}$ where $-\lambda$ is the second eigenvalue of the generator and C is a constant that cannot be controlled by λ , which is a problem when $C \gg 1$. Here again, the hurdles are overcome by adopting a different point of view. Sharp rates are achieved for a one-dimensional RTP with asymmetric tumbling rates biased towards zero [FGM12] by reducing the problem to the computation of the Laplace transform of a hitting time. General biased tumbling rates are treated in [FGM16] at the price of less precise bounds. Also using coupling arguments, we give sharp non-asymptotic bounds for the convergence

speed of two RTPs on the 1D torus [GHM24] and on the line [Hah24]. We obtain matching lower bounds through concentration inequalities, showing the optimality of our results and fully determining scaling of the speed of convergence as a function of model parameters. On the 1D torus, this reveals a persistent and a diffusive regime where the mixing time scale differently in terms of model parameters.

These convergence results, along with the understanding of invariant measures and related universality classes, find their roots in the dialog between the RTP models from out-of-equilibrium statistical mechanics and the mathematical formalism of piecewise deterministic Markov processes (PDMPs). The study of these processes, defined by deterministic motion punctuated by stochastic jumps, is strongly motivated by their use in sampling [MKK14, BCVD18, FBPR18, MDS20, LW22, MGM23] within the fields of statistical mechanics and Bayesian statistics. In this context, the speed of convergence is closely linked to the efficiency of the sampling algorithms, a recurring bottleneck in applications. A unified treatment of PDMPs was first proposed by Davis [Dav84, Dav93] and the subsequent developments of the theory can be found in the reviews [Mal16, ABG⁺14] and references therein. The general theory was then extended in [DC99, CD08], where stability properties such as Harris recurrence were characterized in terms of an embedded chain, and the Feller property along with a tractable core of the Feller generator were established in [DGM21]. Interestingly, PDMPs do not share the regularizing properties of diffusions and are essentially never strong Feller.

Markovian switching between ODEs in \mathbb{R}^d constitutes an important subclass of PDMPs, only differing from RTP processes by the absence of sticky boundaries. It was shown in [BH12, BLBMZ15] that the invariant measure of these switched dynamical systems has a density with respect to the Lebesgue measure provided the vector fields driving the ODEs satisfy Hörmander bracket conditions. While the regularity of general switched random dynamical systems is an open question, C^k -regularity is obtained [BT22, BT23] when the switching rates are sufficiently high. The smoothness of the invariant density for random switching between two smooth vector fields on the torus is shown in [BHLM18].

The question of the long-time behavior, which arises naturally in the context of MIPS, is also relevant in the broader framework of PDMPs. Even simple PDMPs can display unexpected blow-ups [LMR13, BLBMZ14]. One of the more surprising results is that Markovian switching between two linear ODEs $\partial_t x = A_0 x$ and $\partial_t x = A_1 x$ can lead to $x(t) \rightarrow +\infty$ almost surely, even for A_i having only negative eigenvalues. In more favorable settings, e.g. compact state space, classical coupling arguments yield convergence in total variation distance [BLBMZ15] under Hörmander bracket conditions or a more tractable condition on the barycentric combination of the vector fields [BHS18]. Contracting deterministic dynamics [BLBMZ12, CH15] lead to quantitative bounds in a Wasserstein-like metric. Specific models such as the transmission control protocol [DGR02, MZ09, OK08, LVL08], used for internet data transmission, have also been thoroughly investigated [CMP10, BCG⁺13].

The convergence of PDMPs can also be established using hypocoercivity techniques [Vil06, DMS09]. Recently, sharp quantitative bounds for sampling PDMPs were obtained [LW22] by utilizing a hypocoercivity approach [AAMN19, CLW23] originally developed for kinetic Langevin. This is enabled by the fact that many PDMPs relevant to applications, such as RTP systems [CRS14] and sampling processes [ADNR21], are kinetic processes in which the randomness exclusively affects the speed. The approach was then simplified in [EL24a] by introducing a continuous-space analog of lifted Markov chains [CLP99, DHN00]. We adapt these techniques to two RTPs on the 1D torus, extending them to a setting with sticky boundaries and very different velocity dynamics. In fact, we establish a general result that provides sharp rates for kinetic Langevin, sampling PDMPs, and two RTPs on the 1D torus.

Turning now to the outline of the thesis, the remainder Chapter 1 introduces the PDMP framework and standard tools for studying Markov processes' asymptotic behavior. It ends with a description of our contributions and is translated into French in Chapter 2. In Chapter 3, the invariant measure of two RTPs on a 1D torus interacting via jamming is

computed for general tumbling and jamming by casting it as a PDMP. This reveals the existence of two out-of-equilibrium universality classes for this model conjectured to be the continuous-space limit of a known on-lattice model. Chapter 4 rigorously establishes this continuous-space limit. The scaling of the mixing time is fully determined as a function of model parameters using coupling techniques and concentration inequalities. Chapter 5 examines the long-time behavior of two RTPs on a line with an attractive potential and jamming. Changes in model parameters induce qualitative changes in the invariant measure and sharp exponential decay rates are obtained through coupling. Chapter 6 provides the explicit invariant measure for three RTPs on a 1D torus with jamming by plugging a physical ansatz for the invariant measure into the generator characterization of invariance. Chapter 7 presents a sharp hypocoercive convergence result with a unified proof for kinetic Langevin, sampling PDMPs and RTPs based on a general divergence lemma and a flow Poincaré inequality.

1.2 Piecewise deterministic Markov processes

We start with a quick informal description of general piecewise deterministic Markov processes [Dav84, Dav93], which we will heavily use to formalize out-of-equilibrium statistical mechanics models throughout the manuscript. In a few words, a PDMP $X(t)$ follows a differential flow (ϕ_t) on its state space E and jumps with rate $\lambda : E \rightarrow \mathbb{R}_+$ and post-jump distribution given by a Markov kernel $(Q(x, \cdot))_{x \in E}$. More precisely, when the process starts from $x_0 \in E$ we have

$$X(t) = \phi_t(x_0) \text{ for } t \in [0, T)$$

where T is a random time with cumulative distribution function

$$\mathbb{P}(T \leq t) = 1 - \exp\left(-\int_0^t \lambda(\phi_s(x_0)) ds\right).$$

At time T the process jumps to a new state x_1 distributed according to $Q(\phi_T(x_0), \cdot)$. The process then restarts from x_1 and everything repeats. The differential flow can exit the state space through a point $x_\partial \in \partial E$, in which case the process instantly jumps back into E with distribution $Q_\partial(x_\partial, \cdot)$ where $(Q_\partial(x, \cdot))_{x \in \partial E}$ is a second transition kernel.

1.2.1 Examples

The next three sections complete this picture by providing examples of PDMPs.

Instantaneous process

The first example, modeling two persistent random walkers on a 1D torus interacting through jamming collisions, will later illustrate techniques to compute the invariant measure and prove convergence to it. It is a simplified version of the PDMP model introduced in [HGM23] to work directly with jamming RTPs in continuous space, which allows efficient computation of the invariant measure, highlights two out-of-equilibrium universality classes, and enables the determination of the scaling of the mixing time as a function of model parameters. A discrete-space version was first introduced in out-of-equilibrium statistical physics [SEB16] as a way to study clustering phenomena.

Consider two particles on a 1D torus of length L described by their positions $x_1, x_2 \in \mathbb{R}/L\mathbb{Z}$ and velocities $\sigma_1, \sigma_2 \in \{-1, +1\}$. The σ_i are two independent Markov jump processes switching between $+1$ and -1 with rate ω . The positions evolve as follows

- if $x_1 \neq x_2$ (see Figure 1.1a) then $\partial_t x_i = \sigma_i$,
- if $x_1 = x_2$ and if the velocities of the particles are
 - pushing them together (see Figure 1.1b) then $\partial_t x_i = 0$,

– identical (see Figure 1.1c) or driving them apart (see Figure 1.1d) then $\partial_t x_i = \sigma_i$.

The two particles are so-called run-and-tumble particles (RTPs): they alternate between long phases of constant-velocity linear motion (runs) and short erratic movement resulting in reorientation (tumbles). In this simplified model, the reorientation (switch between +1 and -1) is instantaneous.

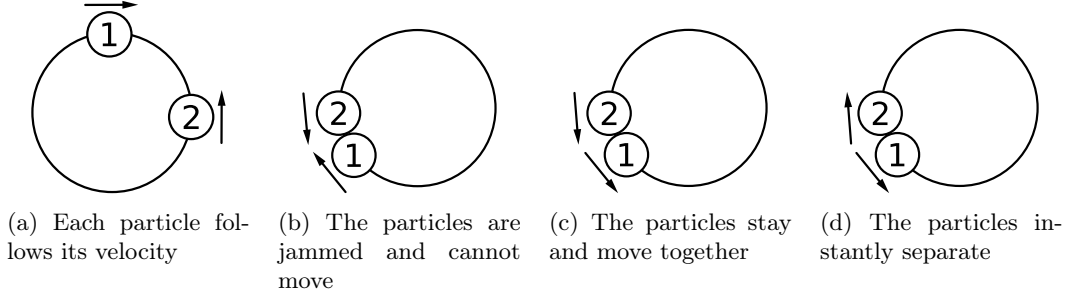


Figure 1.1: Two RTPs with jamming interactions

Thanks to model symmetries, it is enough to keep track of x , the distance from x_1 to x_2 , and the relative velocity $\sigma = \sigma_2 - \sigma_1$. This leads to the following stochastic process, a realization of which is provided in Figure 1.3.

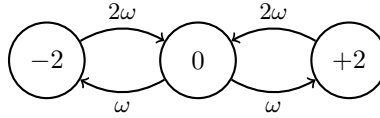


Figure 1.2: Relative velocity jump rates

Definition 1.1 (Instantaneous process). *Consider the state space $E = [0, L] \times \Sigma$ where $L > 0$ and $\Sigma = \{-2, 0, 2\}$. For any initial state $(x_0, \sigma_0) \in E$ let*

- $\sigma(t)$ be a Markov jump process with the rates of Figure 1.2, initial state σ_0 and jump times $0 = T_0 < T_1 < \dots$,
- $x(t)$ be recursively defined by $x(0) = x_0$ and

$$x(t) = \min[L, \max[0, x(T_i) + \sigma(T_i)(t - T_i)]] \text{ for } T_i \leq t \leq T_{i+1}.$$

We call $X(t) = (x(t), \sigma(t))$ the instantaneous process.

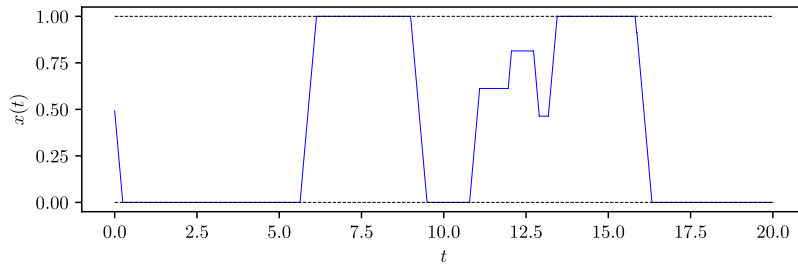


Figure 1.3: Realization of the instantaneous process ($\omega = 0.5, L = 1$)

Random switching between ODEs

The next example is a class of PDMPs relevant to molecular biology [GMK10], where the process switches between different ODEs.

Definition 1.2 (Random switching process). *Let Σ be a finite set and*

- $F_\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be a Lipschitz vector field for all $\sigma \in \Sigma$,
- $q_{\sigma, \bar{\sigma}} : \mathbb{R}^d \rightarrow \mathbb{R}_+$ be a C^0 jump rate for all $\sigma, \bar{\sigma} \in \Sigma$.

Let $(x_0, \sigma_0) \in \mathbb{R}^d \times \Sigma$ be given. For $i \geq 0$ we recursively define

- ξ_i to be the solution of $\partial_t \xi_i = F_{\sigma_i}(\xi_i)$ with $\xi_i(0) = \xi_{i-1}(E_{i-1})$ (resp. $\xi_0(0) = x_0$ for $i = 0$),
- E_i to be a positive random variable depending only on σ_i with survivor function

$$\mathbb{P}(E_i > t) = \exp\left(-\int_0^t \sum_{\sigma} q_{\sigma_i, \sigma}(\xi_i(s)) ds\right),$$

- σ_{i+1} to be a random variable depending only on E_i and σ_i such that

$$\mathbb{P}(\sigma_{i+1} = \sigma) = \frac{q_{\sigma_i, \sigma}(\xi_i(E_i))}{\sum_{\sigma} q_{\sigma_i, \sigma}(\xi_i(E_i))}.$$

Finally set $T_i = \sum_{k=0}^{i-1} E_k$ and define

$$x(t) = \xi_i(t - T_i) \text{ and } \sigma(t) = \sigma_i \text{ for } t \in [T_i, T_{i+1}).$$

We call $X(t) = (x(t), \sigma(t))$ the random switching process.

Key aspects of this large class of PDMPs, such as the regularity of the invariant measure [BH12, BLBMZ15], stability properties [LMR13, BLBMZ14], and convergence behavior [BLBMZ12, CH15], have been thoroughly investigated. Note that the instantaneous process of the previous section is almost a random switching process, but its sticky boundaries set it apart.

Randomized Hamiltonian Monte Carlo

The third example is motivated by the crucial role played by the Boltzmann distribution

$$\pi = \frac{1}{Z} e^{-U(x)} dx \tag{1.1}$$

in statistical mechanics and offers a way to approximately sample from it. In this section, we assume that $U : \mathbb{R}^d \rightarrow \mathbb{R}$ is a C^1 potential and that the normalization $Z := \int_{\mathbb{R}^d} e^{-U(x)} dx$ is finite.

Definition 1.3 (Randomized Hamiltonian Monte Carlo [BRSS17]). *Let $(x_0, v_0) \in \mathbb{R}^d \times \mathbb{R}^d$ be given. Recursively define ξ_i to be the solution of*

$$\begin{aligned} \dot{x} &= v, \\ \dot{v} &= -\nabla_x U(x), \end{aligned}$$

with initial condition $\xi_i(0) = (p_x(\xi_{i-1}(E_{i-1})), V_i)$ (resp. $\xi_0(0) = (x_0, v_0)$), where $(E_i)_{i \geq 0}$ is a sequence of independent exponential random variables with parameter $\lambda > 0$, $(V_i)_{i \geq 1}$ is a sequence of independent d -dimensional Gaussians and p_x is the projection on the x -component. Set $T_i = \sum_{k=0}^{i-1} E_k$.

We call

$$X(t) = \xi_i(t - T_i) \text{ for } [T_i, T_{i+1})$$

the randomized Hamiltonian Monte Carlo process.

In a few words, the process follows the Hamiltonian flow and refreshes its velocity at exponentially distributed time intervals. The randomized Hamiltonian Monte Carlo process admits $\pi \otimes \mathcal{N}(0, I_d)$ as its unique invariant measure and is ergodic [LW22]. Hence, sampling from its fixed-time distribution at sufficiently large times is a way to approximately sample from π .

In practice, it is often challenging to precisely simulate Hamiltonian dynamics, which is a practical concern in many applications. Other PDMPs used in sampling include forward event-chain Monte Carlo [MKK14], the Zig-Zag process [BFR19] and the bouncy particle sampler [PdW12].

1.2.2 Probability semigroup and generator

The study of PDMPs relies on the general theory of Markov processes, whose fundamental tools are succinctly outlined in this section. We adopt the framework of Feller processes [Lig10, Definition 3.1], a broad class of Markov processes with good regularity properties. Although this framework does not encompass all PDMPs as described by Davis [Dav93], it includes all the processes in this manuscript.

A key tool for studying PDMPs, motivated by the complexity inherent to the transition probabilities of continuous-time Markov processes, is the probability semigroup and its generator.

Theorem 1.4 (Feller probability semigroup [Lig10, Theorem 3.15]). *Given a Feller process $X(t)$ define*

$$P_t f(x) = \mathbb{E}_x[f(X(t))]$$

for $f \in C(E)$. Then P_t is a probability semigroup in the sense of [Lig10, Definition 3.4].

If we define μP_t by

$$(\mu P_t)(f) = \mu(P_t f)$$

we see that P_t acts on the space of measures by duality. In fact, μP_t is the law at time t of the process started with initial measure μ , which shows that the process is entirely determined by its semigroup. Studying the semigroup can be thought of a way to study the process through its action on the space of continuous functions vanishing at infinity. This will allow us to use tools from functional analysis and spectral theory.

Definition 1.5 (Feller probability generator [Lig10, Theorem 3.16]). *Given a probability semigroup P_t define \mathcal{L} by*

$$\mathcal{L}f = \lim_{t \rightarrow 0} \frac{P_t f - f}{t}$$

on

$$D(\mathcal{L}) = \{f \in C(E) : \text{the (strong) limit above exists}\}.$$

Then \mathcal{L} is a probability generator in the sense of [Lig10, Definition 3.12].

The generator often constitutes the main computational tool in the study of Markov processes. It can, for example, be used to compute invariant measures (see Section 1.3).

Proposition 1.6. *The Feller probability generator of the instantaneous process is given by*

$$\mathcal{L}f(x, \sigma) = \sigma \partial_x f(x, \sigma) + \sum_{\tilde{\sigma}} q_{\sigma, \tilde{\sigma}} (f(x, \tilde{\sigma}) - f(x, \sigma))$$

where

$$\begin{pmatrix} q_{-2, -2} & q_{-2, 0} & q_{-2, +2} \\ q_{0, -2} & q_{0, 0} & q_{0, +2} \\ q_{+2, -2} & q_{+2, 0} & q_{+2, +2} \end{pmatrix} = \begin{pmatrix} -2\omega & 2\omega & 0 \\ \omega & -2\omega & \omega \\ 0 & 2\omega & -2\omega \end{pmatrix}$$

and has domain

$$D(\mathcal{L}) = \{f(\cdot, \pm 2) \in C^1\} \cap \{\partial_x f(L, +2) = \partial_x f(0, -2) = 0\}.$$

Proposition 1.7. *The set $C^1(\mathbb{R}^d)$ constitutes a core of the Feller probability generator \mathcal{L} of the random switching process and*

$$\mathcal{L}f(x, \sigma) = \underbrace{\langle F_\sigma(x), \nabla_x f(x, \sigma) \rangle}_{\text{deterministic transport}} + \overbrace{\sum_{\tilde{\sigma}} q_{\sigma, \tilde{\sigma}}(x) (f(x, \tilde{\sigma}) - f(x, \sigma))}^{\text{random jumps}} \text{ for all } f \in C^1(E).$$

Proposition 1.8. *The set $C^1(\mathbb{R}^d)$ constitutes a core of the Feller probability generator \mathcal{L} of the randomized Hamiltonian Monte Carlo process and*

$$\mathcal{L}f(x, v) = \underbrace{\langle v, \nabla_x f(x, v) \rangle - \langle \nabla_x U(x), \nabla_v f(x, v) \rangle}_{\text{deterministic Hamiltonian dynamics}} + \lambda \underbrace{\int f(x, \tilde{v}) - f(x, v) d\kappa(\tilde{v})}_{\text{velocity refresh}}$$

where κ is the standard Gaussian in \mathbb{R}^d .

Remark 1.9. *Other definitions of the generator exist outside the Feller setting.*

- *Instead of considering the action of the process on the space of continuous functions, one can consider its action on the space $L^2(\pi)$, leading to a different semigroup and generator. Because $L^2(\pi)$ is a Hilbert space, this is often the right setting for spectral analysis and is particularly useful when the process is reversible i.e. when the L^2 generator is self-adjoint.*
- *The Feller probability generator \mathcal{L} possesses important property that for all $f \in D(\mathcal{L})$*

$$f(X(t)) - f(X(0)) - \int_0^t \mathcal{L}f(X(s)) ds$$

is a martingale under all initial distributions $\mu \in \mathcal{P}(E)$. Taking this property as the definition of the generator yields the extended generator. This extends the domain of the generator to a much larger class of functions and enables the use of probabilistic instead of functional analytic tools.

1.3 Invariant probability measure

Continuing our tour of fundamental concepts in the study of Markov processes, we turn to the objects which characterize their long-time behavior: invariant measures. A measure π is said to be invariant if

$$\pi P_t = \pi \text{ for all } t \geq 0. \quad (1.2)$$

This means that, provided the initial distribution is equal to π , the law of $X(t)$ remains equal to π for all $t \geq 0$. The following theorem collects results in [Par06, Section 7.6] to illustrate how these measures describe the asymptotic behavior. The

Theorem 1.10 (Asymptotic behavior). *Let $X(t)$ be an irreducible Markov jump process with finite state space E then there exists a unique invariant measure which we denote by π .*

(i) Markovian law of large numbers: *For all $f : E \rightarrow \mathbb{R}$ we have*

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t f(X(s)) ds = \pi(f)$$

\mathbb{P}_μ -almost surely for all $\mu \in \mathcal{P}(E)$.

(ii) Markovian central limit theorem: For all $f : E \rightarrow \mathbb{R}$ we have

$$\sqrt{t} \left(\frac{1}{t} \int_0^t f(X(s)) ds - \pi(f) \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2(f))$$

starting from any initial distribution with

$$\sigma^2(f) = \int \Gamma(g) d\pi$$

where g any solution of $\mathcal{L}g = f - \pi(f)$ and $\Gamma(g) = \mathcal{L}(g^2) - 2g\mathcal{L}g$.

(iii) Convergence of fixed-time distributions: There exist $C, \gamma > 0$ such that

$$\|\mu P_t - \pi\|_{\text{TV}} \leq C e^{-\gamma t} \|\mu - \pi\|_{\text{TV}}$$

for all initial distributions $\mu \in \mathcal{P}(E)$ where $\|\cdot\|_{\text{TV}}$ is the total variation distance (see Definition 1.16).

Remark 1.11. We only consider finite state spaces for simplicity, but these types of results can be extended to a large class of continuous-time Markov processes.

Formally, the evolution of the law of $X(t)$ is dictated by the measure-valued autonomous linear differential equation

$$\frac{d}{dt} \mu_t = \mathcal{L}^* \mu_t \tag{1.3}$$

where \mathcal{L}^* is the adjoint with respect to the dual pairing of measures and functions. This suggests that Equation (1.2) can be rewritten as

$$\mathcal{L}^* \pi = 0.$$

The following theorem, which is the main computational tool in the search for explicit invariant measures, makes this rigorous.

Theorem 1.12 (Generator characterization of invariance [Lig10, Theorem 3.37]). *If D is a core of the generator \mathcal{L} then*

$$\pi \text{ is invariant} \iff \int \mathcal{L}f d\pi = 0 \text{ for all } f \in D. \tag{1.4}$$

In light of Equation (1.3), the links between the invariance property of Equation (1.2) and the long-time behavior of $X(t)$ established in Theorem 1.10 can be thought of as an illustration of the following general fact: contracting flows converge towards their unique fixed point. The close relationship between long-time behavior and invariant measures is further illustrated by the following partial converse to Theorem 1.10.

Theorem 1.13. *If $X(t)$ is a Feller Markov process and $X(t) \xrightarrow{\mathcal{L}} \pi$ then π is invariant.*

Note that if we can leverage our physical understanding of the process to obtain an ansatz for the invariant measure, we can prove its correctness using Theorem 1.12. Let us illustrate the strength of this seemingly naive approach using our running example. We know that the instantaneous process spends a positive amount of time in the finite set $\{0, L\} \times \Sigma$ and thus its invariant measure must assign positive mass to this set. Furthermore, equilibrium statistical physics tells us that two RTPs with reflecting instead of jamming collisions should have invariant measure $(\frac{1}{L} dx) \otimes (\frac{1}{4} \delta_{-2} + \frac{1}{2} \delta_0 + \frac{1}{4} \delta_{+2})$. Hence

$$\pi = \sum_{\sigma} (d_{\sigma}^0 \delta_0 + a_{\sigma} dx + d_{\sigma}^L \delta_L) \otimes \delta_{\sigma}$$

is a natural ansatz for the invariant measure of the instantaneous process. Plugging this ansatz into Equation (1.4) proves its correctness and yields

$$\begin{aligned} d_{-2}^0 &= \frac{1/2}{2 + \omega L}, & a_{-2} &= \frac{\omega/4}{2 + \omega L}, & d_{-2}^L &= 0, \\ d_0^0 &= \frac{1/2}{2 + \omega L}, & a_0 &= \frac{\omega/2}{2 + \omega L}, & d_0^L &= \frac{1/2}{2 + \omega L}, \\ d_{+2}^0 &= 0, & a_{+2} &= \frac{\omega/4}{2 + \omega L}, & d_{+2}^L &= \frac{1/2}{2 + \omega L}. \end{aligned}$$

Crucially, when σ takes on more values, this ansatz fails and the invariant measure exhibits rich out-of-equilibrium behavior [HGM23]. Hence, strategies to go from the simple verification of invariant measures to their computation are needed. The next three sections are devoted to finding the invariant measure without having to come up with an ansatz. Even though the three approaches are closely related, they are different enough that one might succeed when the others fail or lead to long computations. Each method will be illustrated using our running example.

1.3.1 Fokker-Planck equation

Solving the Fokker-Planck equation in the interior of the state space is the most versatile approach. Let π be an invariant measure of the instantaneous process, which we write

$$\pi = \sum_{\sigma} (d_{\sigma}^0 \delta_0 + \pi_{\sigma} + d_{\sigma}^L \delta_L) \otimes \delta_{\sigma} \quad (1.5)$$

where π_{σ} is the restriction of π to $(0, L) \times \{\sigma\}$. Note that, unlike in the previous approach, we do not assume anything on the form of π_{σ} . This is a general decomposition valid for any measure on E and not an ansatz.

Consider the test function

$$f(x, \sigma) = \begin{cases} f_{\sigma}(x) & \text{if } 0 < x < L, \\ 0 & \text{otherwise,} \end{cases}$$

where the $f_{\sigma} \in C_c^{\infty}(0, L)$ are arbitrary but fixed. Plugging this particular test function into Equation (1.4) yields

$$0 = \sum_{\sigma} \pi_{\sigma} \left(v_{\sigma} f'_{\sigma} + \sum_{\bar{\sigma}} q_{\sigma, \bar{\sigma}} f_{\bar{\sigma}} \right) = \sum_{\sigma} \left(-v_{\sigma} \pi'_{\sigma} + \sum_{\bar{\sigma}} q_{\bar{\sigma}, \sigma} \pi_{\bar{\sigma}} \right) (f_{\sigma})$$

with differentiation in the sense of distributions. Because the f_{σ} were arbitrary, this leads to $-V\Pi' + \mathcal{Q}^t\Pi = 0$ where

$$\Pi = \begin{pmatrix} \pi_{-2} \\ \pi_0 \\ \pi_{+2} \end{pmatrix}, \quad \mathcal{Q} = \begin{pmatrix} -2\omega & 2\omega & 0 \\ \omega & -2\omega & \omega \\ 0 & 2\omega & -2\omega \end{pmatrix}, \quad V = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & +2 \end{pmatrix},$$

and thus

$$\Pi = a(1, 2, 1)^t dx + b(\omega x + 1, 2\omega x, \omega x - 1)^t dx$$

for some constants a and b . Different angles of attack can then be used to find a, b and $d_{\sigma}^0, d_{\sigma}^L$

1. use the symmetries of the process,
2. use physical arguments to find boundary conditions,

3. use the fact that the σ -marginal of the invariant measure is the invariant measure of the Σ -valued Markov jump process $\sigma(t)$,

and finally

4. plug the information gained in the previous steps back into Equation (1.4).

This approach is very well suited to compute the invariant measure of one-dimensional processes [DKM⁺19, BMR⁺20, LDMS21, Fry21] because one can make use of the established strategies to solve systems of differential equations. However, this approach fails when $-VII' + Q^t\Pi = 0$ is not explicitly solvable.

1.3.2 Laplace transform

The next approach is based on the Laplace transform and is the continuous-space version of the strategy in [SEB16, SEB17]. We illustrate it using again the instantaneous process. Let

$$\pi = \sum_{\sigma} (d_{\sigma}^0 \delta_0 + \pi_{\sigma} + d_{\sigma}^L \delta_L) \otimes \delta_{\sigma} \quad (1.6)$$

be an invariant measure and assume this time that $\pi_{\sigma} = \pi_{\sigma}(x)dx$ has a smooth density w.r.t. the Lebesgue measure.

Step 1: The first step is to obtain a system of linear equations for the Laplace transform of Equation (1.6). Plugging the test function

$$f^{\lambda, \sigma_0}(x, \sigma) = \begin{cases} e^{\lambda x} & \text{if } \sigma = \sigma_0, \\ 0 & \text{otherwise,} \end{cases}$$

and Equation (1.6) into Equation (1.4) yields

$$\int_0^L v_{\sigma} \lambda e^{\lambda x} \pi_{\sigma_0}(x) dx + \sum_{\sigma} q_{\sigma, \sigma_0} \left(d_{\sigma}^0 + \int_0^L e^{\lambda x} \pi_{\sigma}(x) dx + e^{\lambda L} d_{\sigma}^L \right) = 0.$$

Putting the equations obtained for $\sigma_0 = -2, 0, 2$ into vector form yields

$$Q^t D^0 + \lambda V \hat{\Pi} + Q^t \hat{\Pi} + e^{\lambda L} Q^t D^L = 0, \quad (1.7)$$

where $\hat{\Pi} = (\hat{\pi}_{-2}, \hat{\pi}_0, \hat{\pi}_{+2})^t$, $\hat{\pi}_{\sigma}(\lambda) = \int_0^L e^{\lambda x} \pi_{\sigma}(x) dx$ and $D^x = (d_{-2}^x, d_0^x, d_{+2}^x)^t$ for $x = 0, L$.

Step 2: Define $\Sigma^* = \{\sigma \in \Sigma : V_{\sigma, \sigma} \neq 0\}$. The next step is to rewrite Equation (1.7) to express $(\hat{\pi}_{\sigma})_{\sigma \in \Sigma^*}$ in terms of $(d_{\sigma}^x)_{\sigma \in \Sigma^*}$

$$\hat{\pi}_{-2} = \hat{\pi}_{+2} = \frac{\omega (d_{+2}^L e^{\lambda L} - d_{-2}^0)}{2\lambda}.$$

Step 3: The fact that $\hat{\pi}_{+2}(\lambda) = \int_0^L e^{\lambda x} \pi_{+2}(x) dx$ is an entire function implies that the right hand side in the previous step has no poles. This can be used to obtain linear constraints on the $(d_{\sigma}^x)_{\sigma \in \Sigma^*}$

$$\text{Res}_{\lambda=0}(\hat{\pi}_{+2}) = 0 \iff \frac{\omega}{2} (d_{+2}^L - d_{-2}^0) = 0 \iff d_{+2}^L = d_{-2}^0.$$

Step 4: Using Equation (1.7) we can express the Laplace transforms $d_{\sigma}^0 + \hat{\pi}_{\sigma} + e^{\lambda L} d_{\sigma}^L$ for $\sigma \notin \Sigma^*$ in terms of the Laplace transforms for $\sigma \in \Sigma^*$

$$d_0^0 + \hat{\pi}_0(\lambda) + e^{\lambda L} d_0^L = \frac{1}{2} (d_{-2}^0 + \hat{\pi}_{-2}(\lambda) + e^{\lambda L} d_{-2}^L) + \frac{1}{2} (d_{+2}^0 + \hat{\pi}_{+2}(\lambda) + e^{\lambda L} d_{+2}^L)$$

so that all that is left to do is invert the Laplace transforms and impose the normalization $\pi(E) = 1$.

1.3.3 Stopping time

The last approach is based on a probabilistic characterization of the invariant measure that does not involve the generator. It is therefore very different from the previous approaches.

Theorem 1.14 (Stopping time characterization [BH22, Theorem 6.26]). *We suppose that τ is a stopping time, which we assume finite in the sense that $\mathbb{P}_{X_0}(\tau < +\infty)$ for all $X_0 \in E$. We may therefore define the kernels*

$$S(X_0, \cdot) = \mathbb{P}_{X_0}(X(\tau) \in \cdot),$$

and

$$G(X_0, \cdot) = \mathbb{E}_{X_0} \left[\int_0^\tau \delta_{X(s)}(\cdot) ds \right],$$

the former being a Markov kernel. We assume in addition that π_0 is a stationary distribution of S . Then the measure $\pi_0 G$ is stationary for $X(t)$. In particular, if we also have that $\mathbb{E}_{\pi_0}[\tau] < +\infty$, then the probability measure

$$\pi := \frac{\pi_0 G}{\mathbb{E}_{\pi_0}[\tau]}$$

is stationary for $X(t)$.

Remark 1.15. *This characterization is mostly used to prove qualitative properties of the invariant measure such as its existence [Kha60, Kli87] or its regularity [BT22] and not to obtain an explicit formula.*

Let us illustrate this last approach using the instantaneous process. If $E_0 = \{0, L\} \times \Sigma$ and we define

$$\tau = \inf\{t \geq 0 : X(t) \neq X(0) \text{ and } X(t) \in E_0\}$$

then E_0 is invariant under S in the sense that $S((x, \sigma), E_0) = 1$ for all $(x, \sigma) \in E_0$. Hence S can be seen as a Markov kernel on E_0 . Its one-step transition matrix

$$P = \begin{matrix} & \begin{matrix} (0, -2) & (0, 0) & (0, +2) & (L, -2) & (L, 0) & (L, +2) \end{matrix} \\ \begin{matrix} (0, -2) \\ (0, 0) \\ (0, +2) \\ (L, -2) \\ (L, 0) \\ (L, +2) \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{L\omega}{L\omega+2} & 0 & 0 & 0 & 0 & \frac{2}{L\omega+2} \\ \frac{2}{L\omega+2} & 0 & 0 & 0 & 0 & \frac{L\omega}{L\omega+2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix} \quad (1.8)$$

can be computed by solving systems of ODEs [Dav93, Chapter 3]. Its invariant measure π_0 is given by

$$\begin{aligned} \pi_0(0, -2) &= 1/5, & \pi_0(0, 0) &= 1/5, & \pi_0(0, +2) &= 1/10, \\ \pi_0(L, -2) &= 1/10, & \pi_0(L, 0) &= 1/5, & \pi_0(L, +2) &= 1/5. \end{aligned}$$

If we start from $(x, \sigma) \in E_0 - \{(L, -2); (0, +2)\}$ then $\tau = \inf\{t \geq 0 : X(t) \neq X(0)\}$ so that τ has an exponential distribution and $\delta_{(x, \sigma)} G = \frac{1}{2\omega} \delta_{(x, \sigma)}$. Solving systems of ODEs yields

$$\begin{aligned} \delta_{(L, -2)} G &= \frac{\omega x + 2}{2(L\omega + 2)} dx \otimes \delta_{-2} + \frac{\omega x + 1}{L\omega + 2} \otimes \delta_0 + \frac{\omega x}{2(L\omega + 2)} \otimes \delta_{+2}, \\ \delta_{(0, +2)} G &= \frac{\omega(L - x)}{2(L\omega + 2)} dx \otimes \delta_{-2} + \frac{\omega(L - x) + 1}{L\omega + 2} \otimes \delta_0 + \frac{\omega(L - x) + 2}{2(L\omega + 2)} \otimes \delta_{+2}, \end{aligned}$$

and makes the invariant probability measure $\pi = \frac{\pi_0 G}{\pi_0 G(E)}$ fully explicit.

1.4 Convergence towards the invariant measure

Theorem 1.10 shows that invariant measures are a crucial tool in the study of Markov processes because they dictate their long-time behavior. However, the theorem holds no clue as to when the asymptotic behavior takes over. This question turns out to be of paramount importance in applications. For example, a quantitative understanding of the speed of convergence towards the invariant measure can be used to compare the performance of different Markov Chain Monte Carlo sampling algorithms, especially in terms dimension dependence.

Many approaches exist and we do not try to be exhaustive. Each technique yields a different form of convergence and is applicable to different classes of Markov processes. For this reason, we start by highlighting an important type of Markov processes, on which we will illustrate convergence techniques that are not directly applicable to PDMPs. Some techniques are applicable to both, in which case we use the simpler example.

1.4.1 Diffusions: an important class of examples

Diffusions are strong Markov processes with continuous sample paths possessing the Feller property and will be used as an alternative to PDMPs to demonstrate convergence techniques in the remainder of this section. They constitute a large class of continuous-time Markov processes and can be defined as the solution of the stochastic differential equation [Oks13, LG16]

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t$$

where $b, \sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d$. Their generator is given by

$$\mathcal{L}f(x) = \sum_i b_i(x)\partial_i f(x) + \sum_{ij} \left(\sigma(x)\sigma(x)^t\right)_{ij} \partial_{ij} f(x).$$

We will use the Ornstein–Uhlenbeck process (OUP)

$$dZ_t = -Z_t dt + \sqrt{2}dB_t$$

associated to the generator

$$\mathcal{L}f(x) = -xf'(x) + f''(x)$$

to illustrate classical convergence techniques. Its invariant measure is the standard Gaussian.

1.4.2 Couplings

With our new example in hand, we turn to the first convergence technique which directly leverages the characterization of the total variation and Wasserstein distance in terms of couplings.

Definition 1.16 (Total variation distance [Jou24, Definition 1.3.1]). *The total variation distance between two measures μ and ν is defined as*

$$\|\mu - \nu\|_{\text{TV}} = \inf_{\Pi \in C(\mu, \nu)} \int 1_{\{x \neq y\}} d\Pi(x, y)$$

where $C(\mu, \nu)$ is the set of all couplings of μ and ν .

This distance makes $\mathcal{P}(E)$ a complete metric space and convergence w.r.t. this distance implies convergence in law. If we have a coupling (X, \tilde{X}) where X is Markov process with initial distribution δ_x and \tilde{X} is the same Markov process started from the invariant distribution π then it follows from the definition of the total variation distance that

$$\|\delta_x P_t - \pi\|_{\text{TV}} \leq \mathbb{P}(\tau > t)$$

where $\tau = \inf\{t \geq 0 : X(t) = \tilde{X}(t)\}$.

The following distance offers an alternative way to quantify the distance between two measures.

Definition 1.17 (Wasserstein distance [Vil09, Definition 6.1]). (i) Define the set of measures possessing a first moment

$$\mathcal{P}_1(E) := \{\mu \in \mathcal{P}(E) : \mu(d(x_0, \cdot)) < +\infty\}$$

where $x_0 \in E$ is arbitrary but fixed. (The definition is independent of this choice.)

(ii) The Wasserstein distance between $\mu, \nu \in \mathcal{P}_1(E)$ is defined as

$$\|\mu - \nu\|_{\mathcal{W}^1} = \inf_{\Pi \in C(\mu, \nu)} \int d(x, y) d\Pi(x, y)$$

where $C(\mu, \nu)$ is again the set of all couplings of μ and ν .

The Wasserstein distance makes $\mathcal{P}_1(E)$ a complete metric space and convergence w.r.t. this distance implies convergence in law. It is again apparent from the definition that if we have a coupling (X, \tilde{X}) where $X(t)$ (resp. $\tilde{X}(t)$) is a Markov process started with initial distribution δ_x (resp. π) then

$$\|\delta_x P_t - \pi\|_{\mathcal{W}^1} \leq \mathbb{E}[d(X(t), \tilde{X}(t))].$$

The strength of the coupling approach lies in its simplicity and its generality. However, to get sharp bounds, one has to work on a case by case basis.

Illustration using the OUP Let (X, \tilde{X}) be the solution of the SDE

$$\begin{aligned} dX_t &= -X_t dt + \sqrt{2} dB_t, \\ d\tilde{X}_t &= -\tilde{X}_t dt + \sqrt{2} dB_t, \end{aligned}$$

i.e. two OUPs driven by the same Brownian motion. This is known as the synchronous coupling. Then $D_t := X_t - \tilde{X}_t$ follows the deterministic differential equation $dD_t = -D_t dt$ so that $D_t = e^{-t} D_0$. Hence taking $X_0 \sim \delta_x$ and $\tilde{X}_0 \sim \pi$ we get

$$\|\delta_x P_t - \pi\|_{\mathcal{W}^1} \leq \mathbb{E}|X_t - \tilde{X}_t| = e^{-t} \mathbb{E}|D_0| \leq e^{-t} \left(x + \sqrt{\frac{2}{\pi}} \right).$$

1.4.3 Uniform minorization condition

The next method also yields convergence in total variation distance and boils down to the Banach fixed-point theorem [Jou24, Section 1.3].

Definition 1.18. We say that the uniform minorization condition is satisfied if there exist $\nu \in \mathcal{P}(E)$ and $\alpha \in (0, 1)$ such that

$$\inf_{x \in E} \delta_x P_T(\cdot) \geq \alpha \nu(\cdot)$$

for some $T > 0$.

This condition is also called the Doeblin condition.

Theorem 1.19. If the uniform minorization condition is satisfied then there exists a unique invariant measure π and

$$\|\mu P_t - \pi\|_{\text{TV}} \leq (1 - \alpha)^{\lfloor t/T \rfloor}$$

for all $\mu \in \mathcal{P}(E)$.

In fact, under the uniform minorization condition, we have

$$\|\mu P_T - \nu P_T\|_{\text{TV}} \leq (1 - \alpha) \|\mu - \nu\|_{\text{TV}} \text{ for all } \mu, \nu \in \mathcal{P}(E).$$

Hence P_T is a contraction mapping. So, because the total variation distance makes $\mathcal{P}(E)$ a complete metric space, Theorem 1.19 follows from the Banach fixed-point theorem. This convergence result is interesting because of its simplicity and the fact that it is quantitative, although it rarely yields sharp bounds. Also note that it often fails when the state space is not compact. For example, in the case of the OUP

$$\inf_{x \in E} \delta_x P_T(A) = 0 \tag{1.9}$$

for all measurable sets A and all $T > 0$ implying that this convergence technique cannot be applied. For this reason, we illustrate it using the instantaneous process instead.

Illustration using the instantaneous process The process $\sigma(t)$ on its own is a Markov jump process so, by solving the Kolmogorov backward equations, one can compute

$$\inf_{(x_0, \sigma_0) \in E} \mathbb{P}_{(x_0, \sigma_0)}(\sigma(1/\omega) = -2) = \frac{(e^2 - 1)^2}{4e^4}.$$

Also observe that if the instantaneous process starts with velocity $\sigma_0 = -2$ and the velocity does not change during a time $L/2$ then the process arrives at the state $(0, -2)$. This implies

$$\inf_{\sigma_0 = -2} \mathbb{P}_{(x_0, \sigma_0)}(X(L/2) = (0, -2)) \geq e^{-2\omega(L/2)}.$$

The Markov property then implies

$$\begin{aligned} & \mathbb{P}_{(x_0, \sigma_0)}(X(1/\omega + L/2) = (0, -2)) \\ & \geq \left(\inf_{(x_0, \sigma_0) \in E} \mathbb{P}_{\sigma_0}(\sigma(1/\omega) = -2) \right) \left(\inf_{\sigma_0 = -2} \mathbb{P}_{(x_0, \sigma_0)}(X(L/2) = (0, -2)) \right) \\ & \geq \frac{(e^2 - 1)^2}{4e^4} e^{-\omega L} \end{aligned}$$

which is of course Definition 1.18 with $T = 1/\omega + L/2$, $\alpha = \frac{(e^2 - 1)^2}{4e^4} e^{-\omega L}$ and $\nu = \delta_{(0, -2)}$.

1.4.4 Foster-Lyapunov criterion

This section closely follows [HM11] and [Jou24]. The general spirit of the approach originates in [MT93]. The following two conditions generalize the uniform minorization condition and can often be applied even in the non-compact state space setting where the uniform minorization condition fails as in Equation (1.9).

Definition 1.20. *We say that the geometric drift condition is satisfied when there exists a function $V : E \rightarrow \mathbb{R}_+$ and constants $K \geq 0$ and $\gamma \in (0, 1)$ such that*

$$P_T V \leq \gamma V + K$$

for some $T > 0$.

Definition 1.21. *We say that the local uniform minorization condition is satisfied if there exist $\nu \in \mathcal{P}(E)$ and $\alpha \in (0, 1)$ such that*

$$\inf_{x \in C} \delta_x P_T(\cdot) \geq \alpha \nu(\cdot)$$

with $C = \{x \in E : V(x) \leq R\}$ for some $R > 2K/(1 - \gamma)$ where K, γ and T are the constants from the geometric drift condition.

The geometric drift condition ensures that the process often visits the level set C where the local uniform minorization condition is satisfied. In the spirit of [MT93], it can be replaced by the stronger condition

$$\mathcal{L}V \leq -cV + K.$$

Definition 1.22 (Modified V -norm). *Define the set of measures*

$$\mathcal{X} := \{\mu \in \mathcal{P}(E) : \mu(V) < +\infty\}.$$

For all $\mu, \nu \in \mathcal{X}$ define the modified V -norm

$$\rho_\beta(\mu, \nu) := \int_E (1 + \beta V(x)) d|\mu - \nu|(x)$$

where $|\mu - \nu|$ is the total variation measure.

This norm makes \mathcal{X} a complete metric space.

Theorem 1.23 ([HM11, Theorem 1.3]). *If both the geometric drift condition and the local uniform minorization condition are verified, then there exists $\bar{\alpha} \in (0, 1)$ and $\beta > 0$ so that*

$$\rho_\beta(P_T\mu, P_T\nu) \leq \bar{\alpha}\rho_\beta(\mu, \nu) \text{ for all } \mu, \nu \in \mathcal{X}.$$

In particular, for any $\alpha_0 \in (0, \alpha)$ and $\gamma_0 \in (\gamma + 2K/R, 1)$ one can choose $\beta = \alpha_0/K$ and

$$\bar{\alpha} = \max\left(1 - (\alpha - \alpha_0), \frac{2 + R\beta\gamma_0}{2 + R\beta}\right).$$

The following theorem again follows from the Banach fixed-point theorem.

Theorem 1.24. *If both the geometric drift condition and the local uniform minorization condition are satisfied, then there exists a unique invariant probability π . Furthermore $\pi(V) \leq \frac{K}{1-\gamma}$ and*

$$\rho_\beta(\mu P_t, \pi) \leq \bar{\alpha}^{\lfloor t/T \rfloor} \rho_\beta(\mu, \pi) \text{ for all } \mu \in \mathcal{X}.$$

The strength of this convergence result lies in the fact that it yields a quantitative bound and that, contrary to the uniform minorization condition, it can be applied in a wide range of settings, for example for the OUP. However, the bounds obtained through this method are rarely tight.

Illustration using the OUP Taking $V(x) = x^2$ leads to

$$\mathcal{L}V(x) = -xV'(x) + V''(x) = -2x^2 + 2 = -2V(x) + 2$$

so that the geometric drift condition is satisfied. Furthermore the transition probabilities are known to be

$$\delta_x P_t = \sqrt{\frac{1}{2\pi(1 - e^{-2t})}} \exp\left(-\frac{(y - xe^{-t})^2}{1 - e^{-2t}}\right) dy =: p(t, x, y) dy$$

and thus, for any fixed $t > 0$, we have

$$\inf_{x^2 \leq R} \inf_{0 \leq y \leq 1} p(t, x, y) := \alpha(t, R) > 0$$

implying the local uniform minorization condition

$$\delta_x P_t(\cdot) \geq \alpha(t, R)(1_{\{0 \leq y \leq 1\}} dy)(\cdot).$$

Theorem 1.24 then yields convergence in modified V -norm.

1.4.5 Poincaré inequality

The next approach is more functional analytic than probabilistic in flavor and yields exponential decay of the probability semigroup in $L^2(\pi)$ [GGSC06, Section 2.1].

Assumption (A) Assume that \mathcal{L} admits a Hilbert basis $(f_n)_{n \in \mathbb{N}}$ of eigenfunctions in $L^2(\pi)$ corresponding to the eigenvalues $(-\lambda_n)_{n \in \mathbb{N}}$. Further assume $0 = \lambda_0 < \lambda_1 < \dots$ and $f_0 = 1$.

Remark 1.25. *If \mathcal{L} is self-adjoint in $L^2(\pi)$ and has compact resolvent then it admits a Hilbert basis of eigenfunctions.*

It is well known that in this setting the variance decreases exponentially with a decay rate given by the spectral gap λ_1 .

Lemma 1.26. *If (A) is satisfied then $\text{Var}(P_t f) \leq e^{-2\lambda_1 t} \text{Var}(f)$.*

The optimality of the decay rate can be seen by taking $f = f_1$.

Proof. We have

$$\text{Var}_\pi(P_t f) = \left\| \sum_{n=1}^{+\infty} a_n e^{-\lambda_n t} f_n \right\|_{L^2(\pi)}^2 = \sum_{n=1}^{+\infty} a_n^2 e^{-2\lambda_n t} \leq e^{-2\lambda_1 t} \sum_{n=1}^{+\infty} a_n^2 = e^{-2\lambda_1 t} \text{Var}_\pi(f).$$

□

The following alternative proof is also instructive.

Proof. For $f = \sum_{n=0}^{+\infty} a_n f_n \in D(\mathcal{L})$ we have

$$\langle f, \mathcal{L}f \rangle_{L^2(\pi)} = - \sum_{n=0}^{+\infty} \lambda_n a_n^2 \leq -\lambda_1 \sum_{n=1}^{+\infty} a_n^2 = -\lambda_1 \text{Var}_\pi(f)$$

and therefore

$$\frac{d}{dt} \text{Var}_\pi(P_t f) = 2 \langle P_t f, \mathcal{L}P_t f \rangle_{L^2(\pi)} \leq -2\lambda_1 \text{Var}_\pi(P_t f)$$

so the desired exponential decrease is a consequence of Grönwall's inequality. □

Notice that in the alternative proof, the spectral decomposition is not really needed, we actually just need $\langle f, \mathcal{L}f \rangle_{L^2(\pi)} \leq -\lambda_1 \text{Var}_\pi(f)$. This motivates the following definition.

Definition 1.27. *We say that the Poincaré inequality with constant $C > 0$ is satisfied if*

$$\text{Var}_\pi(f) \leq -C \langle f, \mathcal{L}f \rangle_{L^2(\pi)} \text{ for all } f \in D(\mathcal{L}).$$

Note that in the reversible case, the optimal Poincaré constant is the inverse of the spectral gap. The following theorem completes the picture of the relationship between Poincaré constant and asymptotic decay rate.

Theorem 1.28. (i) *The Poincaré inequality with constant C is equivalent to*

$$\text{Var}_\pi(P_t f) \leq e^{-\frac{2}{C}t} \text{Var}_\pi(f)$$

for all $t \geq 0$ and $f \in L^2(\pi)$ (see [Ebe09, Theorem 7.22]).

(ii) *If \mathcal{L} is reversible, then the Poincaré inequality with constant C is further equivalent to the existence of $K > 1$ such that*

$$\text{Var}_\pi(P_t f) \leq K e^{-\frac{2}{C}t} \text{Var}_\pi(f)$$

for all $t \geq 0$ and $f \in L^2(\pi)$ (see [CGZ13, Lemma 2.14]).

In the non-reversible case, if $\mathcal{L} = \mathcal{S} + \mathcal{A}$ where \mathcal{S} is symmetric and \mathcal{A} is anti-symmetric we have $\langle f, \mathcal{L}f \rangle_{L^2(\pi)} = \langle f, \mathcal{S}f \rangle_{L^2(\pi)}$. Therefore we can only hope to harness the decay coming from \mathcal{S} and will necessarily lose out on the decay coming from \mathcal{A} and its interaction with \mathcal{S} . In fact, some processes do not satisfy the Poincaré inequality, as will be illustrated in the next section.

Illustration using the OUP Hermite polynomials defined by

$$\text{He}_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}$$

are known to satisfy the differential equation

$$\text{He}_n'' - X\text{He}_n' = -n\text{He}_n$$

and form an orthogonal basis of $L^2(\pi)$. Therefore, Lemma 1.26 implies

$$\text{Var}_\pi(P_t f) \leq e^{-t} \text{Var}_\pi(f).$$

1.5 Kinetic processes and hypocoercivity

This section follows [Vil06, Section 3] and [Mon14, Section 1.4]. Instead of trying to present a general theory, we will first give an example where the Poincaré approach of Section 1.4.5 fails in an instructive way and then show how the approach can be adapted.

1.5.1 An example where Poincaré fails

The kinetic Langevin process associated to the potential $U(x) = \frac{1}{2}x^2$ is the solution of the stochastic differential equation

$$\begin{aligned} dx_t &= v_t dt, \\ dv_t &= -x_t dt - v_t dt + \sqrt{2} dB_t. \end{aligned}$$

It has generator

$$\mathcal{L}f(x, v) = \underbrace{v\partial_x f(x, v) - x\partial_v f(x, v)}_{\text{Hamiltonian flow}} + \underbrace{-v\partial_v f + \partial_v^2 f(x, v)}_{\text{OUP in } v}$$

and invariant measure $\pi = \frac{1}{2\pi} e^{-\frac{x^2+v^2}{2}} dx dv$. The term ‘kinetic’ refers to the fact that there are two variables x and v and that the random noise is restricted to the second variable. The instantaneous process is also a kinetic process because all its randomness comes from its velocity σ . This reduced level of noise is a source of added difficulty. For instance, in this specific example, the Poincaré inequality takes the form

$$\text{Var}_\pi(f) \leq C \|\partial_v f\|_{L^2(\pi)}^2,$$

which cannot be satisfied since the right hand side vanishes when f depends only on x . Hence, we cannot use the Poincaré approach to obtain convergence towards the invariant measure.

1.5.2 A hypocoercive approach

In an attempt to extend the Poincaré inequality to this setting, one can consider the H^1 norm

$$\|f\|_{H^1(\pi)}^2 = \|f\|_{L^2(\pi)}^2 + \|\partial_x f\|_{L^2(\pi)}^2 + \|\partial_v f\|_{L^2(\pi)}^2$$

and try to prove an inequality of the form

$$\frac{d}{dt} \|P_t f\|_{H^1(\pi)}^2 \leq -C \|P_t f\|_{H^1(\pi)}^2.$$

However, we have

$$\begin{aligned} \frac{d}{dt} \|\partial_x P_t f\|_{L^2(\pi)}^2 &= -2 \langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)} - 2 \|\partial_{xv} P_t f\|_{L^2(\pi)}^2, \\ \frac{d}{dt} \|\partial_v P_t f\|_{L^2(\pi)}^2 &= -2 \|\partial_v P_t f\|_{L^2(\pi)}^2 + 2 \langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)} - 2 \|\partial_v^2 P_t f\|_{L^2(\pi)}^2, \end{aligned}$$

so that

$$\frac{d}{dt} \|P_t f\|_{H^1(\pi)} = -4 \|\partial_v P_t f\|_{L^2(\pi)} - 2 \|\partial_v^2 P_t f\|_{L^2(\pi)} - 2 \|\partial_{xv} P_t f\|_{L^2(\pi)}$$

and thus $\frac{d}{dt} \|P_t f\|_{H^1(\pi)}^2$ vanishes again if f is a non-constant function of x .

Including the cross term Taking inspiration from Hörmander's hypoelliptic regularity theorem [Hör67] where the vector fields interact through their Lie brackets to create the missing directions, we consider the cross term $\langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)}$ which has derivative

$$\begin{aligned} \frac{d}{dt} \langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)} &= \\ \underbrace{\|\partial_x P_t f\|_{L^2(\pi)}^2 - \|\partial_v P_t f\|_{L^2(\pi)}^2}_{\text{desired term}} &- \langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)} - 2 \langle \partial_{xv} P_t f, \partial_v^2 P_t f \rangle_{L^2(\pi)}. \end{aligned}$$

Now notice that the norm

$$\|f\|_\epsilon^2 = \|f\|_{L^2(\pi)}^2 + \epsilon \left(\|\partial_x f\|_{L^2(\pi)}^2 + \|\partial_v f\|_{L^2(\pi)}^2 + \langle \partial_x f, \partial_v f \rangle_{L^2(\pi)} \right)$$

is equivalent to $\|\cdot\|_{H^1(\pi)}$ for all $\epsilon > 0$ and compute

$$\begin{aligned} \frac{d}{dt} \|P_t f\|_\epsilon^2 &= -\epsilon \|\partial_x P_t f\|_{L^2(\pi)}^2 + (-2 - \epsilon) \|\partial_v P_t f\|_{L^2(\pi)}^2 + \epsilon \langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)} \\ &\quad - 2\epsilon \|\partial_v^2 P_t f\|_{L^2(\pi)}^2 - 2\epsilon \|\partial_{xv} P_t f\|_{L^2(\pi)}^2 + 2\epsilon \langle \partial_{xv} P_t f, \partial_v^2 P_t f \rangle_{L^2(\pi)}. \end{aligned}$$

Because

$$\begin{aligned} \langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)} &\leq \frac{1}{2} \|\partial_x P_t f\|^2 + \frac{1}{2} \|\partial_v P_t f\|^2, \\ \langle \partial_{xv} P_t f, \partial_v^2 P_t f \rangle_{L^2(\pi)} &\leq \frac{1}{2} \|\partial_{xv} P_t f\|^2 + \frac{1}{2} \|\partial_v^2 P_t f\|^2, \end{aligned}$$

we get

$$\frac{d}{dt} \|P_t f\|_\epsilon^2 \leq -\frac{\epsilon}{2} \|\partial_x P_t f\|_{L^2(\pi)}^2 - \left(2 + \frac{\epsilon}{2}\right) \|\partial_v P_t f\|_{L^2(\pi)}^2$$

so using the Poincaré inequality for π (w.r.t. the full gradient) and the equivalence of $\|\cdot\|_\epsilon$ and $\|\cdot\|_{H^1(\pi)}$ we get that there exists $C > 0$ such that

$$\frac{d}{dt} \|P_t f\|_\epsilon^2 \leq -C \|P_t f\|_\epsilon^2.$$

Grönwall's inequality then implies the exponential decay in H^1 . One can then recover exponential decay in L^2 using regularity estimates (e.g. [Vil09, Theorem A.8]).

Selected related results The overall strategy can be summed up as follows: find another norm to which Grönwall’s inequality can be applied. The many alternative approaches designed to deal with similarly degenerate processes form the theory of hypocoercivity. Changing the norm is also at the heart of the Dolbeault-Mouhot-Schmeiser approach to hypocoercivity [DMS09, DMS15]. The main difference with the derivation above is that the norm to which Grönwall’s inequality is applied is equivalent to the L^2 norm and not the H^1 norm. This L^2 approach is used in [ADNR21] to obtain exponential convergence for PDMPs used in sampling with explicit dimension dependence of the decay rate.

Using Poincaré inequalities in both space and time introduced in [AAMN19], sharp quantitative bounds for the convergence of kinetic Langevin are given in [CLW23]. In particular, after optimizing over the friction parameter, the optimal \sqrt{m} dependence of the spectral gap on the spectral gap m of the corresponding overdamped Langevin is achieved in the convex potential case. The method is then revisited to obtain exponential convergence of PDMPs coming from sampling applications with a sharp explicit decay rates [LW22]. These results are simplified in [EL24a] by reinterpreting them in terms of a continuous analog of lifts of Markov chains [DHN00, CLP99].

1.6 Contributions

This section gives an overview of the original contributions in this thesis, all of which are concerned with the long-time behavior of RTPs.

- The invariant measure of two RTPs on the 1D torus as well as on the real line are computed using the generator characterization of invariance. This perspective on boundary conditions reveals two out-of-equilibrium universality classes on the 1D torus and qualitative differences in the invariant measure on the line. Both translate different degrees of deviation from the equilibrium setting. An ansatz for the three-particle measure on the 1D torus is derived by leveraging the understanding of the two-particle case and model symmetries and then proven to be correct using the generator. This represents a step toward addressing the general many-particle measure, which remains an open problem.
- Sharp quantitative bounds on the speed of convergence towards the invariant measure are obtained using coupling techniques. Reverse bounds are derived through concentration inequalities, proving the optimality of these results and fully determining the speed of convergence as a function of model parameters. On the 1D torus, a persistent and a diffusive regime, with different scalings of the mixing time in terms of model parameters, are observed.
- Hypocoercive convergence results based on second-order lifts and flow Poincaré inequalities are extended and unified to obtain sharp quantitative convergence bounds in a framework covering both previous results, such as underdamped Langevin and PDMP sampling processes, and RTPs.

1.6.1 Two run-and-tumble particles on the 1D torus



Figure 1.4: Instantaneous tumble

Preexisting discrete model The following process was introduced in [SEB16] to study the clustering properties of two run-and-tumble particles interacting through jamming, a model for bacteria such as *E. Coli*. Consider two particles on a one-dimensional periodic lattice represented by their positions $y_1^N, y_2^N \in \mathbb{Z}/N\mathbb{Z}$ and their velocities $\sigma_1, \sigma_2 \in \{\pm 1\}$. These velocities are two independent copies of the Markov jump process of Figure 1.4a and there is an independent Poisson clock with rate $\gamma > 0$ attached to each particle. When their clock rings, the particles jump from their current position y_i^N to the new position $y_i^N + \sigma_i$, except if the other particle is already there. Hence the particles cannot cross and interact through jamming. The interplay of random position jumps and velocity flips is illustrated in Figure 1.5.

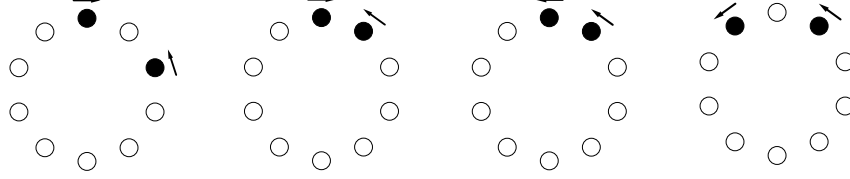


Figure 1.5: Sequence of random position jumps and velocity flips

It suffices to keep track of the distance $y^N \in \{1, \dots, N\}$ from the first to the second particle as well as the relative velocity $\sigma = \sigma_2 - \sigma_1$. The process $\sigma(t)$ is a Markov jump process with the transition rates of Figure 1.4b and the couple $Y^N(t) = (y^N(t), \sigma(t))$ is the Markov jump process with rates of Figure 1.6 and state space $E_N = \{1, \dots, N\} \times \{-2, 0, +2\}$.

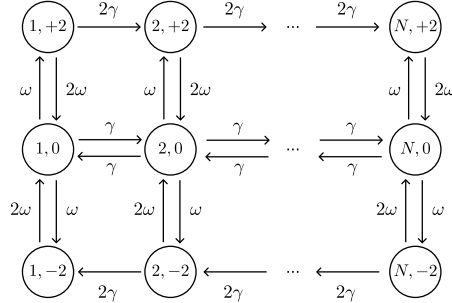


Figure 1.6: Transitions of the discrete RTP process

In [SEB16], the explicit expression for the invariant measure π_N of $Y^N(t)$, which has increased mass at the boundary in the form of Diracs and decaying exponentials, is found¹. Furthermore, the continuous-space limit of π_N is identified as the invariant measure

$$\pi = \sum_{\sigma \in \Sigma} (d_\sigma^0 \delta_0 + a_\sigma dx + d_\sigma^L \delta_L) \otimes \delta_\sigma$$

of the instantaneous process of Section 1.2.1 computed in Section 1.3. However, the exact sense in which this convergence takes place (convergence in law, total variation, Wasserstein...) is not entirely clear and the proof of convergence is not very detailed.

Continuous limit of discrete processes The fact that $\pi_N \rightarrow \pi$ in the continuous-space limit leads to the following question: is it possible to show that the entire process $Y^N(t)$ converges to the instantaneous process $X(t)$?

¹The process in [SEB16] is actually $(y^N(t), \sigma_1(t), \sigma_2(t))$ and not $Y^N(t) = (y^N(t), \sigma(t))$ but all results easily translate between the models.

To answer this question, we start by defining the injection $i_N : E_N \rightarrow E$ by

$$i_N(y, \sigma) = \left(L \frac{y-1}{N-1}, \sigma \right)$$

which allows us to see the state space E_N as a discretization of the state space of the instantaneous process $E = [0, L] \times \{-2, 0, +2\}$.

The continuous-space limit is the limit where the lattice spacing $\frac{L}{N-1}$ goes to 0 while the average speed of the particles $\gamma \frac{L}{N-1}$ and the velocity flip rate ω stay constant i.e.

$$\left. \begin{array}{l} L \rightarrow +\infty \\ \gamma = \frac{N-1}{L} \\ \omega \text{ constant} \end{array} \right\}. \quad (*)$$

The following proposition then shows that the instantaneous process $X(t)$ is the distributional limit of $Y^N(t)$ rigorously linking the on-lattice model in [SEB16] and its continuous-space version.

Proposition 1.29 (Discrete-continuous limit). *(i) There exists a coupling of the instantaneous process X and a sequence of discrete processes $(Y^N)_{N \geq 2}$ with parameters scaling as in (*) such that for all $T > 0$*

$$\mathbb{P} \left(\sup_{t \leq T} |i_N(y^N(t)) - x(t)| \geq \epsilon \right) \leq \frac{1}{\epsilon} \left(\frac{L}{N-1} + 8 \sqrt{\frac{T((\omega T)^2 + 3\omega T + 1)L}{N-1}} \right).$$

(ii) We have

$$\left((i_N(y^N(t)), \sigma(t)) \right)_{t \geq 0} \xrightarrow{\mathcal{L}} \left(X(t) \right)_{t \geq 0}$$

in the Skorokhod space $\mathcal{D}[0, +\infty)$.

The main idea is to construct a joint probability space for $X(t) = (x(t), \sigma(t))$ and $Y^N(t) = (y^L(t), \sigma^N(t))$ where

$$\sigma(t) = \sigma^N(t) \text{ a.s. for all } t \geq 0.$$

In other words, the velocities are identical and we have to examine the convergence of $i_N(y^N(t))$, which makes more and more frequent jumps of smaller and smaller size, towards the deterministic motion of $x(t)$. On any time interval where $\sigma = +2$ is constant, the situation essentially reduces to the toy example $\tilde{y}^N(t)$ of Figure 1.7, where the state space is $\{0, \dots, N\}$, the process jumps from k to $k+1$ with rate N if $k \leq N-1$ and N is absorbing.

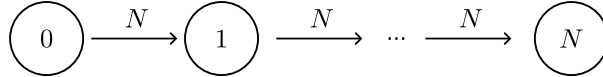


Figure 1.7: Transition rates of $\tilde{y}^N(t)$ in toy example

The following proposition shows that this toy example converges towards the right translation on $[0, 1]$ with absorbing boundary at $x = 1$.

Proposition 1.30 (Toy example). *Let $\tilde{y}^N(t)$ be the Markov jump process with state space $\{0, \dots, N\}$ and rates of Figure 1.7 starting from $\tilde{y}^N(0) = 0$. If $\tilde{x}(t) = \min(1, t)$ then we have*

$$\mathbb{P} \left(\sup_{t \leq T} |\tilde{y}^N(t)/N - \tilde{x}(t)| \geq \epsilon \right) \leq \frac{T}{\epsilon^2 N}.$$

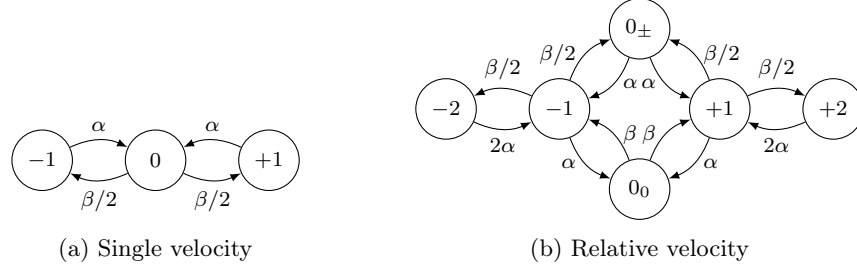


Figure 1.8: Instantaneous tumble

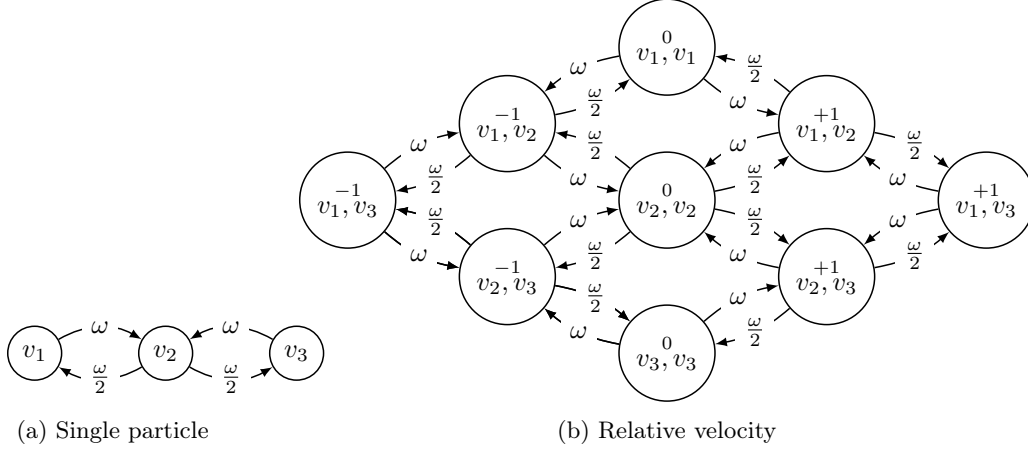


Figure 1.9: Anisotropic velocities

Proof. Notice that $\tilde{y}^N(t)$ has the same law as $\min(U^N(t), N)$ where $U^N(t)$ is a Poisson counting process with rate N . Hence, using the fact that $x \mapsto \min(x, 1)$ is 1-Lipschitz

$$\begin{aligned} \mathbb{P}\left(\sup_{t \leq T} |\tilde{y}^N(t)/N - \tilde{x}(t)| \geq \epsilon\right) &= \mathbb{P}\left(\sup_{t \leq T} |\min(U^N(t)/N, 1) - \min(t, 1)| \geq \epsilon\right), \\ &\leq \mathbb{P}\left(\sup_{t \leq T} |U^N(t)/N - t| \geq \epsilon\right), \\ &= \mathbb{P}\left(\sup_{t \leq T} |U^N(t) - Nt| \geq N\epsilon\right). \end{aligned}$$

Doob's martingale inequality then yields

$$\mathbb{P}\left(\sup_{t \leq T} |\tilde{y}^N(t)/N - \tilde{x}(t)| \geq \epsilon\right) \leq \frac{\mathbb{E}\left[(U^N(T) - NT)^2\right]}{(\epsilon N)^2} = \frac{T}{\epsilon^2 N}.$$

□

Establishing that the convergence of Y^N towards π_N is uniform in N then allows the exchange the continuous-space and long-time limit and rigorously show the convergence

$$\|i_N \# \pi_N - \pi\|_{\mathcal{V}^1} \rightarrow 0$$

giving further theoretical underpinning to the discretization approach in [SEB16, SEB17, MEB22, MEB23].

Invariant measure and universality classes In [SEB17] the computation of the invariant measure and its continuous-space limit is extended to the setting where each particle follows the tumbling rates of Figure 1.8a. The extra state $\sigma_i = 0$ corresponds to tumbling which now has a finite duration, more closely modeling the biological reality of bacteria such as E. Coli because the tumble duration is not neglected [SSB12]. Note that, in this case, the process $\sigma = \sigma_2 - \sigma_1$ is not Markovian unless the state $\sigma_2 - \sigma_1 = 0$ is divided into two states 0_{\pm} (corresponding to $\sigma_1 = \sigma_2 = \pm 1$) and 0_0 (corresponding to $\sigma_1 = \sigma_2 = 0$) as done in Figure 1.8b.

We call the continuous-space version of this process the finite process (because of the finite tumble duration). Using the Fokker-Planck approach of Section 1.3.1 to compute the invariant measure yields

$$\pi = \sum_{\sigma \in \Sigma} (d_{\sigma}^0 \delta_0 + \pi_{\sigma}(x) dx + d_{\sigma}^L \delta_L) \otimes \delta_{\sigma} \quad (1.10)$$

where

$$\pi_{\sigma}(x) = a_{\sigma} + b_{\sigma}^c \cosh(\kappa(x - L/2)) + b_{\sigma}^s \sinh(\kappa(x - L/2))$$

and $a_{\sigma}, b_{\sigma}^c, b_{\sigma}^s$ are explicit in terms model parameters. This is a more compact version of the continuous-space limit in [SEB17] and highlights model symmetries. Up to a scalar factor, the vector of uniform contributions $(a_{\sigma})_{\sigma \in \Sigma}$ coincides with the invariant measure of the velocity process $\sigma(t)$.

Note that, unlike in [SEB16], there is exponential accumulation at the boundaries even at the continuous-space level. In the case of anisotropic velocities of Figure 1.9 we have

$$\pi_{\sigma}(x) = a_{\sigma} + \sum_{i=1,2} [b_{\sigma}^{c,i} \cosh(\kappa_i(x - L/2)) + b_{\sigma}^{s,i} \sinh(\kappa_i(x - L/2))]$$

where the coefficients are again explicit in terms of model parameters. Here the exponential accumulation at the boundaries takes place at two different length scales.

In fact, the PDMP formalism together with the use of model symmetries allows the computation of the invariant measure of a much larger class of processes where the velocity σ can follow any Markov jump process with rates that can differ in the bulk $(0, L) \times \Sigma$ and at the boundary $\{0, L\} \times \Sigma$, i.e. rates depending on whether the particles are at contact or not. The invariant measure always takes the form of Equation (1.10) and, just as in Section 1.3.1, the vector of densities w.r.t. the Lebesgue measure $(\pi_{\sigma}(x))_{\sigma \in \Sigma}$ is obtained by solving a linear ODE so it is given by a matrix exponential. Fixing the coefficients in front of the Diracs then yields the explicit invariant measure, generalizing [SEB17].

It remains true in general that the vector of uniform terms $(a_{\sigma})_{\sigma \in \Sigma}$ is set by the invariant measure of the velocity process. The ODE is determined by the stochastic dynamics in the bulk while the boundary conditions are set by the tumbling behavior at the boundary, yielding a neat separation of the contributions to the invariant measure of the jamming and tumbling aspects of the dynamics.

All models share the following active counterpart to global balance

$$\sum_{\sigma \in \Sigma} \sigma \pi_{\sigma}(x) = 0 \text{ for all } x \in (0, L), \quad (1.11)$$

which we use to identify two out-of-equilibrium universality classes. The detailed jamming class is characterized by the stricter condition

$$\pi_{-\sigma}(x) = \pi_{\sigma}(x) \text{ for all } (x, \sigma) \in (0, L) \times \Sigma, \quad (1.12)$$

whereas in the global jamming class only Equation (1.11) is satisfied. This serves as an analog to detailed and global balance respectively. We have thus arrived at a simple criterion, which determines the behavior of the invariant measure, and remains valid when the state space of the velocity process $\sigma(t)$ is continuous. For processes in the detailed jamming

class such as the instantaneous process, we have that the $\pi_\sigma(x) = a_\sigma$ are uniform and that the restriction $\sum_\sigma \pi_\sigma(x) dx \otimes \delta_\sigma$ of invariant measure to the bulk is a product measure. On the contrary, processes in the global jamming class, such as the finite process and the process with anisotropic velocities, have invariant measures whose restriction to the bulk are neither uniform nor product measures. Comparing this with the equilibrium process with the same stochastic dynamics in the bulk but reflecting instead of jamming collision, which has invariant measure $\pi = \sum_\sigma a_\sigma dx \otimes \delta_\sigma$, we see that the detailed jamming class is closer to equilibrium while the global jamming class is further away. Interestingly Equation (1.12) is equivalent to

$$\pi_{-\sigma}(0) = \pi_\sigma(0) \text{ for all } \sigma \in \Sigma$$

so it can be inferred from the jamming behavior.

Remark 1.31. *The rigorous discrete-to-continuous limits for the invariant measure and the entire process can also be established for the process with the relative velocity of Figure 1.8b.*

Mixing time To obtain upper bounds on the mixing time, we apply a coupling argument to the synchronous coupling.

Definition 1.32 (Synchronous coupling). *Let $(x_0, \sigma_0), (\tilde{x}_0, \tilde{\sigma}_0) \in [0, L] \times \Sigma$ be given. Consider the coupling of two Markov jump processes $\sigma(t)$ and $\tilde{\sigma}(t)$ with rates of Figure 1.4b (resp. Figure 1.8b) and initial states σ_0 and $\tilde{\sigma}_0$ such that $\sigma(t)$ and $\tilde{\sigma}(t)$ are independent until the stopping time $\tau_\sigma = \inf\{t \geq 0 : \sigma(t) = \tilde{\sigma}(t)\}$ and then equal forever. Further recursively define $x(t)$ (resp. $\tilde{x}(t)$) as in Definition 1.1 with initial state x_0 (resp. \tilde{x}_0) and driving noise $\sigma(t)$ (resp. $\tilde{\sigma}(t)$). We call the couple (X, \tilde{X}) with $X(t) = (x(t), \sigma(t))$ and $\tilde{X}(t) = (\tilde{x}(t), \tilde{\sigma}(t))$ the synchronous coupling.*

Proposition 1.33 (Upper bound on mixing time). *(i) In the case of the instantaneous process, for all $\epsilon \in (0, 1)$ there exists $C(\epsilon) > 0$ such that*

$$t_{\text{mix}}(\epsilon) \leq C(\epsilon) \frac{1}{\omega} ((\omega L)^2 + 1)$$

where

$$t_{\text{mix}}(\epsilon) = \inf \left\{ t \geq 0 : \sup_{\mu \in \mathcal{P}(E)} \|\mu P_t - \pi\|_{\text{TV}} \leq \epsilon \right\}.$$

(ii) In the case of the finite process, for all $\epsilon \in (0, 1)$ there exists $C(\epsilon) > 0$ such that

$$t_{\text{mix}}(\epsilon) \leq C(\epsilon) \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) ((\alpha L)^2 + 1).$$

Sketch of proof. It follows from the recursive definition of $x(t)$ and $\tilde{x}(t)$ that we have the order preservation property

$$\begin{aligned} x(\tau_\sigma) \leq \tilde{x}(\tau_\sigma) &\implies x(t) \leq \tilde{x}(t) \text{ for all } t \geq \tau_\sigma, \\ x(\tau_\sigma) \geq \tilde{x}(\tau_\sigma) &\implies x(t) \geq \tilde{x}(t) \text{ for all } t \geq \tau_\sigma. \end{aligned}$$

Therefore, assuming without loss of generality that $x(\tau_\sigma) \leq \tilde{x}(\tau_\sigma)$, if $\tilde{x}(t^*) = 0$ for some $t^* \geq \tau_\sigma$ then

$$0 \leq x(t^*) \leq \tilde{x}(t^*) = 0 \implies x(t^*) = \tilde{x}(t^*).$$

Hence, for the synchronous coupling $\mathbb{P}(X(t) \neq \tilde{X}(t))$ can be bounded by controlling the tails of τ_σ and $\tau_0 = \inf\{t \geq 0 : x(t) = 0\}$, which are explicitly computable for the instantaneous process and can be estimated for the finite process. \square

Proposition 1.34 (Lower bound on mixing time). *(i) In the case of the instantaneous process, for all $\epsilon \in (0, 1)$ there exists $\tilde{C}(\epsilon) > 0$ such that*

$$t_{\text{mix}}(\epsilon) \geq \tilde{C}(\epsilon) \frac{1}{\omega} ((\omega L)^2 + 1).$$

(ii) In the case of the finite process, for all $\epsilon \in (0, 1)$ there exists $\tilde{C}(\epsilon) > 0$ such that

$$t_{\text{mix}}(\epsilon) \leq \tilde{C}(\epsilon) \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) ((\alpha L)^2 + 1).$$

Sketch of proof. We identify two obstacles to mixing, i.e. two obstacles to the fixed-time distribution being close to the invariant measure. First, enough time must have elapsed for at least one velocity jump to occur. This leads to

$$t_{\text{mix}}(\epsilon) \geq \frac{\log(1/\epsilon)}{2\omega} \quad \text{and} \quad t_{\text{mix}}(\epsilon) \geq \frac{\log(1/\epsilon) \max(1/\alpha, 1/\beta)}{2}$$

for the instantaneous and finite process respectively. Second, enough time must have elapsed for a change of order L in the position x to occur. Establishing concentration results on $\int_0^t \sigma(s) ds$ yields

$$t_{\text{mix}}(\epsilon) \geq C'(\epsilon) \omega L^2 \quad \text{and} \quad t_{\text{mix}}(\epsilon) \geq C'(\epsilon) (1/\alpha + 1/\beta) (\alpha L)^2$$

for the instantaneous process in the regime $\omega L \gg 1$ and for the finite process in the regime $\alpha L \gg 1$ respectively. \square

Note that the upper and lower bounds on the mixing time match up to a multiplicative constant meaning that we have fully determined its dependence on model parameters. The scaling of the mixing time can be understood as a crossover between persistent regime where $\omega L \ll 1$ (resp. $\alpha L \ll 1$) and a diffusive regime where $\omega L \gg 1$ (resp. $\alpha L \gg 1$). In the persistent regime, L order of the time needed to traverse the state space if there is no velocity flip is lower than $1/\omega$ (resp. $1/\alpha + 1/\beta$) the order of the mean time between two velocity flips. Hence the mixing time is essentially the time it takes for at least one velocity flip to occur. On the contrary, in the diffusive regime, traversing the state space takes much longer than velocity flips and hence the mixing time is essentially the time it takes to see a meaningful change in the position.

1.6.2 Two run-and-tumble particles on the real line

Model We extend our analysis beyond the compact state space setting to consider two RTPs on the real line, subjected to an attractive interaction potential V and jamming interactions. The particles' speeds, σ_1 and σ_2 , are independent copies of the same Markov jump process, while their positions, x_1 and x_2 , evolve according to the ODEs

$$\partial_t x_1 = -V'(x_1 - x_2) + v\sigma_1(t) \quad \text{and} \quad \partial_t x_2 = -V'(x_2 - x_1) + v\sigma_2(t).$$

Thus, the relative position $x = x_2 - x_1$ is governed by the ODE

$$\partial_t x = -2V'(x) + v\sigma(t)$$

for $x > 0$, where $\sigma = \sigma_2 - \sigma_1$ is the relative velocity, with state space denoted by Σ . Along with jamming interactions, this forms the PDMP $X(t) = (x(t), \sigma(t))$ with state space $[0, +\infty) \times \Sigma$.

We focus on three specific instances of this model

- the instantaneous linear process where $V(x) = c|x|$ and σ follows the rates in Figure 1.4b,

- the finite linear process where $V(x) = c|x|$ and σ follows the rates in Figure 1.8b,
- the instantaneous harmonic process where $V(x) = \mu x^2/2$ and σ follows the rates in Figure 1.4b.

These models are similar to those in [LDMS21], but with key differences: the jamming interactions prevent particle crossings, and the finite linear process incorporates a biologically relevant tumbling mechanism where the tumble duration is explicitly considered [SEB17, SSB12].

Invariant measure In the instantaneous linear case, the Fokker-Planck approach laid out in Section 1.3.1 involves solving a linear vector ODE $\dot{X} = AX$. The invariant measure is then given in terms of the matrix exponential of A which is computed using its spectral decomposition. Integrability on $[0, +\infty)$ implies that only negative eigenvalues and the associated eigenvectors appear in the invariant measure. The Dirac masses on the boundary can be fixed using the fact that the σ -marginal of the invariant measure is the invariant measure of the process $\sigma(t)$.

Proposition 1.35. *The unique invariant measure of the instantaneous linear process is*

$$\pi = \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + a_\sigma e^{\zeta x} dx) \otimes \delta_\sigma$$

where $\zeta = -\frac{2c\omega}{v^2 - c^2}$ and

$$\begin{aligned} d_2 &= 0, & d_0 &= \frac{c}{c+v}, & d_{-2} &= \frac{cv}{(c+v)^2}, \\ a_2 &= \frac{c\omega}{2(v^2 - c^2)}, & a_0 &= \frac{c\omega}{(c+v)^2}, & a_{-2} &= \frac{c(v-c)\omega}{2(c+v)^3}. \end{aligned}$$

The same approach can be used for the finite linear process. However, the matrix A is larger and depends on more parameters so its explicit spectral decomposition demands greater care.

Proposition 1.36. *The unique invariant measure of the finite linear process is*

$$\pi = \begin{cases} \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + c_2 a_\sigma(\zeta_2) e^{\zeta_2 x} dx) \otimes \delta_\sigma & \text{when } c < v \leq 2c, \\ \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + c_2 a_\sigma(\zeta_2) e^{\zeta_2 x} dx + c_3 a_\sigma(\zeta_3) e^{\zeta_3 x} dx) \otimes \delta_\sigma & \text{when } v > 2c, \end{cases}$$

where c_i, ζ_i, d_σ and a_σ are explicit functions of α, β, c and v . Furthermore, when $v < c \leq 2v$, we have

$$d_{+2} = 0 \text{ and } d_\sigma > 0 \text{ otherwise,}$$

whereas when $c > 2v$, we have

$$d_{+2} = d_{+1} = 0 \text{ and } d_\sigma > 0 \text{ otherwise.}$$

Notice that when $c \leq 2v$ the invariant measure is given by a single exponential term in the interior of the space $\{x > 0\}$, whereas when $c > 2v$ an additional exponential term appears at the same time as a Dirac disappears on the boundary. This qualitative change indicates a jump from a stronger to a weaker form of clustering, as indicated by the amount of mass at and around $x = 0$ under the invariant measure. This is reminiscent of the shape transition of the invariant measure in [DKM⁺19, BMR⁺20] where the invariant measure can either vanish or diverge at the edge of the state space depending on the parameters, indicating varying extents of departure from equilibrium.

In the instantaneous harmonic case, the fact that the system of differential equations for the invariant measure is no longer linear makes the Fokker-Planck approach more challenging. Fortunately, once restricted to the interior of the state space $\{x > 0\}$, the process

coincides with the restriction of the process in [BMR⁺20]. Because both processes enter this part of their state space through the same point $(0, +2)$, their invariant measures must coincide on $\{x > 0\}$ up to a multiplicative constant. The Dirac masses on the boundary can again be fixed using the σ -marginal leading to the following result.

Proposition 1.37. *The unique invariant measure of the instantaneous harmonic process takes the form*

$$\pi = \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + \pi_\sigma(x) dx) \otimes \delta_\sigma$$

where π_σ vanishes outside the compact interval $[0, v/\mu]$ and at

- $x = 0$ the x -marginal $\sum_\sigma \pi_\sigma$ diverges for $\mu \geq 2\omega$ and reaches a finite limit for $\mu < 2\omega$,
- $x = v/\mu$ the x -marginal $\sum_\sigma \pi_\sigma$ diverges for $\mu > 2\omega$, reaches a positive limit for $\mu = 2\omega$ and vanishes for $\mu < 2\omega$.

Furthermore $d_2 = 0$ and $d_0, d_{-2} > 0$ for all parameter choices.

The densities π_σ are combinations of hypergeometric functions. Proposition 1.37 shows that the clustering again takes different forms depending on model parameters.

Convergence Applying a coupling argument to the synchronous coupling yields explicit non-asymptotic convergence bounds involving the Laplace transform of the hitting time $\tau_0 = \inf\{t \geq 0 : x(t) = 0\}$. These bounds imply the following asymptotic bounds which are easier to state and interpret.

Proposition 1.38. (i) *In the instantaneous linear case*

$$\frac{\omega}{2} \frac{c^2}{v^2} \leq \liminf_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}}.$$

for all $(x, \sigma) \in [0, +\infty) \times \Sigma$.

(ii) *In the finite linear case*

$$\frac{3 - \sqrt{5}}{8} \min \left(\alpha \left(1 + \frac{\alpha}{\beta} \right) \frac{c^2}{v^2}, \alpha \frac{c}{v} \right) \leq \liminf_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}}$$

for all $(x, \sigma) \in [0, +\infty) \times \Sigma$.

Sketch of proof. We consider again the synchronous coupling where the two copies σ and $\tilde{\sigma}$ of the velocity coalesce exponentially fast and then remain equal forever. The same order preservation argument as in the sketch of proof of Proposition 1.33 transforms the quantitative understanding of the tail of τ_0 into convergence bounds. Noticing that

$$\tau_0 > t \implies x(0) + v \int_0^t \sigma(s) ds - 2ct > 0 \implies \frac{1}{t} \int_0^t \sigma(s) ds > \frac{2c}{v} - \frac{x(0)}{vt}$$

allows the use of non-asymptotic large deviation results [Wu00, CG08, Lez01] to obtain exponential tail bounds for τ_0 . \square

We investigate the optimality of the bounds above by establishing corresponding upper bounds.

Proposition 1.39. (i) *In the instantaneous linear case*

$$\overline{\lim}_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \leq 4\omega \frac{c^2}{v^2}$$

for all $(x, \sigma) \in [0, +\infty) \times \Sigma$.

(ii) In the finite linear case

$$\overline{\lim}_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \leq 4\alpha \left(1 + \frac{\alpha}{\beta}\right) \frac{c^2}{v^2}$$

for all $(x, \sigma) \in [0, +\infty) \times \Sigma$.

Sketch of proof. The key observation is that the process takes a long time to explore the part of the state space where x is large. By definition of the total variation distance we have

$$\|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \geq \pi(\{x > at\}) - \mathbb{P}_{(x,\sigma)}(x(t) > at)$$

so if we can find $a > 0$ such that

$$\underbrace{\lim_{t \rightarrow +\infty} -\frac{1}{t} \log \mathbb{P}(x(t) > at)}_{=:\lambda(a)} > \underbrace{\lim_{t \rightarrow +\infty} -\frac{1}{t} \log \pi(\{x > at\})}_{=:\lambda_\pi(a)}$$

then

$$\lim_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \leq \lambda_\pi(a).$$

The value of $\lambda_\pi(a)$ can be read off the explicit expression for the invariant measure. Setting $\tau_L = \inf\{t \geq 0 : x(t) \geq L\}$ we get that for all $u < 0$

$$\mathbb{P}(x(t) > at) \leq \mathbb{P}(\tau_{at} < t) \leq e^{-ut} \mathbb{E}[e^{u\tau_{at}}],$$

and $\mathbb{E}[e^{u\tau_{at}}]$ can be computed by solving a system of ODEs. \square

Notice that bounds on the exponential convergence rate of the instantaneous linear process match up to a multiplicative constant. This is not exactly the case for the finite linear process but

$$\alpha \frac{c}{v} \ll \alpha \left(1 + \frac{\alpha}{\beta}\right) \frac{c^2}{v^2}$$

can only happen when $\alpha \gg \beta$ which would mean that the RTPs spend the overwhelming majority of their time tumbling and not running, which is of little physical interest.

In the instantaneous harmonic case, the deterministic dynamics is contracting in the sense that

$$|x(t_2) - \tilde{x}(t_2)| \leq e^{-\mu(t_2-t_1)} |x(t_1) - \tilde{x}(t_1)|,$$

if $\sigma(t) = \tilde{\sigma}(t)$ for all $t \in [t_1, t_2]$. This can be exploited to obtain a quantitative convergence result in Wasserstein-type distance following [BLBMZ12].

1.6.3 Three run-and-tumble particles on the 1D torus

Model We now turn to three run-and-tumble particles on the 1D torus of length L interacting through jamming. This corresponds to the instantaneous process but with three particles instead of two.

Because of cyclic order preservation, it is enough to keep track of the distances between pairs of consecutive particles as well as their velocities. We denote x_i the distance from particle i to particle $i+1$ (adopting the convention that particle 4 is particle 1) and $\sigma_i \in \{\pm 1\}$ the velocity of particle i . The state space of the process is $E = \Delta \times \Sigma$ where

$$\Delta = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1, x_2, x_3 \geq 0 \text{ and } x_1 + x_2 + x_3 = L\},$$

and $\Sigma = \{\pm 1\}^3$. We denote $x = (x_1, x_2, x_3) \in \Delta$ and $\sigma = (\sigma_1, \sigma_2, \sigma_3) \in \Sigma$ the distance and velocity vectors respectively.

Each σ_i independently switches between $+1$ and -1 with rate ω . When all particles are at positive distance from each other, each particle follows its velocity so that $\partial_t x_i = \sigma_{i+1} - \sigma_i$.

When exactly two particles are at contact, the particles behave as in the two particle case (see Figure 1.1). Finally, simultaneous collisions between three particles, is summarized in Figure 1.10.

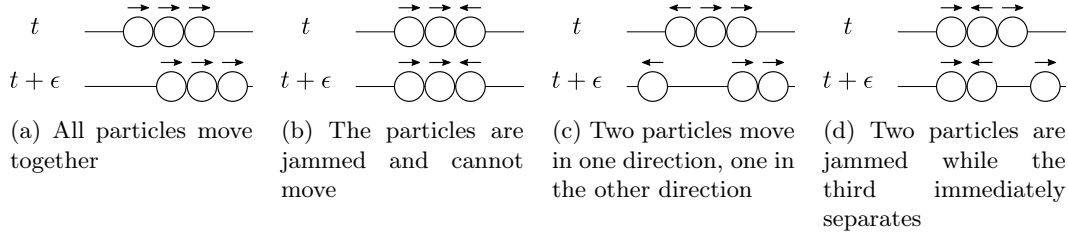


Figure 1.10: Complete list of three particle collisions

Denoting $A := (L, 0, 0)$, $B := (0, L, 0)$, $C := (0, 0, L)$ the vertices of the simplex, and after observing that x remains constant while $\sigma = \pm(1, 1, 1)$, the deterministic dynamics on the rest of the state space can be summarized as in Figure 1.11.

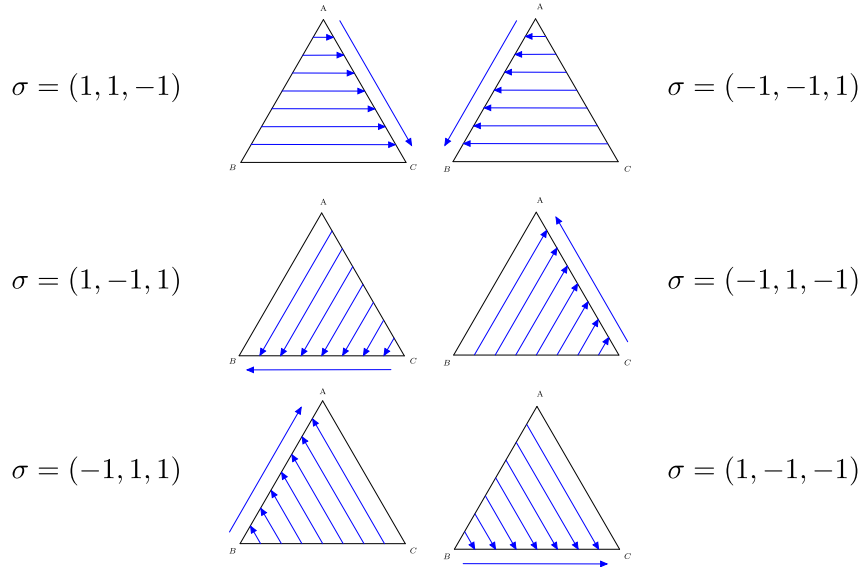


Figure 1.11: Deterministic dynamics on $\Delta \times \{\sigma\}$ for $\sigma \neq \pm(1, 1, 1)$

Invariant measure Noticing that two particles at contact behave like a single effective particle with changed tumbling rates, together with the understanding of the two-particle case, leads to the ansatz

$$\pi = \sum_{\sigma \in \Sigma} \left(\sum_{F \in \mathcal{F}} a_{\sigma, F} \lambda_F + \sum_{X \neq Y} b_{\sigma, XY} \epsilon_{XY} \right) \otimes \delta_{\sigma} \quad (1.13)$$

for the invariant measure. Here $\mathcal{F} = \{A, B, C, AB, BC, CA, ABC\}$ is the set of non-empty faces of the polytope Δ , λ_F is the $\dim(F)$ -dimensional Lebesgue measure on the face $F \in \mathcal{F}$ and for $X, Y \in \{A, B, C\}$

$$\epsilon_{XY}(f) = \int_0^{\|Y-X\|} f \left(X + u \frac{Y-X}{\|Y-X\|} \right) e^{\kappa u} du \neq \epsilon_{YX}(f)$$

is the one-dimensional measure with exponential density.

Plugging this ansatz into the generator characterization of invariance confirms its validity and determines the parameters. The two key ingredients here are the explicit characterization of the extended generator of PDMPs [Dav93] and the use of model symmetries. Note that the difficult question of the boundary conditions of the invariant measure is implicitly answered by the boundary conditions that $f \in D(\mathcal{L})$ must satisfy.

1.6.4 Lifts with sticky boundaries and general velocity dynamics

Second-order lifts The following continuous-time analog of the lift of discrete Markov chains [CLP99, DHN00] was introduced by Andreas Eberle and Francis Lörler [EL24a] to simplify the hypocoercive convergence results in [AAMN19, LW22, CLW23].

Let $p : \hat{E} \rightarrow E$ be a surjection between two Polish spaces \hat{E} and E . Given process \hat{X} with strong L^2 generator $\hat{\mathcal{L}}$ and invariant measure $\hat{\pi}$ taking its values on the larger state space \hat{E} , the idea is to define a new process X with strong L^2 generator \mathcal{L} on the smaller state space E . The invariant measure of X should be the pushforward of $\hat{\pi}$ by p , which we denote π .

Definition 1.40 (Second-order lift [EL24a]). *If there exists a core C of \mathcal{L} such that*

$$\{f \circ p : f \in C\} \subset D(\hat{\mathcal{L}})$$

and the first order condition

$$\left\langle \hat{\mathcal{L}}(f \circ p), g \circ p \right\rangle_{L^2(\hat{\pi})} = 0$$

as well as the second order condition

$$\langle \mathcal{L}f, g \rangle_{L^2(\pi)} = -\frac{1}{2} \left\langle \hat{\mathcal{L}}(f \circ p), \hat{\mathcal{L}}(g \circ p) \right\rangle_{L^2(\hat{\pi})}$$

are satisfied for all $f, g \in C$ then \hat{X} is a lift of X . We say that X is the collapse of \hat{X} .

For example, if $U : \mathbb{R}^d \rightarrow \mathbb{R}$ is a C^1 potential with $Z := \int e^{-U} dx < +\infty$, the underdamped Langevin process

$$\hat{\mathcal{L}}f(x, v) = v \cdot \nabla_x f(x, v) - \nabla_x U(x) \cdot \nabla_v f(x, v) - \gamma v \cdot \nabla_v f(x, v) + \gamma \Delta_v f(x, v)$$

is a lift of the overdamped Langevin process

$$\mathcal{L}f(x) = -\frac{1}{2} \nabla_x U(x) \cdot \nabla_x f(x) + \frac{1}{2} \Delta_x f(x).$$

Here we have $\hat{E} = \mathbb{R}^d \times \mathbb{R}^d$, $E = \mathbb{R}^d$ and p is the projection of onto the x -component. The invariant measure $\pi = \frac{1}{Z} e^{-U} dx$ of X is the pushforward of the invariant measure $\hat{\pi} = \pi \otimes \kappa$ of \hat{X} where κ is the d -dimensional standard Gaussian. Because $\hat{\mathcal{L}}(f \circ p)(x, v) = v \cdot \nabla_x f(x)$ we can check that

$$\int \hat{\mathcal{L}}(f \circ \pi)(g \circ \pi) d\hat{\pi} = \int g(x) \nabla_x f(x) \cdot \left[\int v d\kappa(v) \right] d\pi(x) = 0$$

and

$$\begin{aligned} -\frac{1}{2} \int \hat{\mathcal{L}}(f \circ \pi) \hat{\mathcal{L}}(g \circ \pi) d\hat{\pi} &= -\frac{1}{2} \int \int (v \cdot \nabla_x f(x)) (v \cdot \nabla_x g(x)) d\kappa(v) d\pi(x) \\ &= -\frac{1}{2} \int \nabla_x f(x) \cdot \nabla_x g(x) d\pi(x) \\ &= \int \left(-\frac{1}{2} \nabla_x U(x) \cdot \nabla_x f(x) + \frac{1}{2} \Delta_x f(x) \right) g(x) d\pi(x). \end{aligned}$$

Many sampling processes such as the bouncy particle sampler [BCVD18], zigzag [BFR19] and forward event-chain Monte Carlo [MKK14] are second-order lifts of diffusions.

Abstract hypocoercive setting Starting from a collapse \mathcal{L} with a spectral gap m i.e.

$$\|f - \pi(f)\|_{L^2(\pi)}^2 \leq \frac{1}{m} \langle f, (-\mathcal{L})f \rangle_{L^2(\pi)}$$

and a lift $\hat{\mathcal{L}}$, the first step to establish convergence results in [EL24a] is to find a splitting $\hat{\mathcal{L}} = \hat{\mathcal{L}}_d + \gamma \hat{\mathcal{L}}_v$ where $\gamma > 0$, $\hat{\mathcal{L}}_d$ is itself a lift of \mathcal{L} which is antisymmetric and $\hat{\mathcal{L}}_v$ is a generator acting only on the velocity. For example, the overdamped Langevin dynamics

$$\hat{\mathcal{L}}f(x, v) = \underbrace{v \cdot \nabla_x f(x, v) - \nabla_x U(x) \cdot \nabla_v f(x, v)}_{=: \hat{\mathcal{L}}_d} - \underbrace{\gamma v \cdot \nabla_v f(x, v) + \gamma \Delta_v f(x, v)}_{=: \gamma \hat{\mathcal{L}}_v}$$

naturally splits into the Hamiltonian dynamics $\hat{\mathcal{L}}_d$ and the velocity OUP $\hat{\mathcal{L}}_v$. For the instantaneous process, this splitting and even the computation of the collapsed process is less obvious. In fact, there are two main challenges to adapting the results in [AAMN19, LW22, CLW23, EL24a] (1) the state space has boundaries (2) the velocity refreshment $\hat{\mathcal{L}}_v$ is neither the full refresh of [LW22, EL24a] nor the OUP of [CLW23].

We nonetheless compute the collapse of the instantaneous process, which turns out to be sticky Brownian motion, and find a suitable splitting $\hat{\mathcal{L}} = \hat{\mathcal{L}}_d + \gamma \hat{\mathcal{L}}_v$. We further identify general assumptions on $\hat{\mathcal{L}}_d$ and $\hat{\mathcal{L}}_v$ which are reminiscent of the assumptions in the DMS framework [DMS09, DMS15] and allow us to establish a unified hypocoercive convergence result covering the instantaneous process as well as overdamped Langevin and many sampling processes. The key intermediary result is the following flow Poincaré inequality [AAMN19, LW22, CLW23, EL24a] along the flow of the Fokker-Planck equation.

Proposition 1.41 (Flow Poincaré inequality). *For $f_0 \in L_0^2(\hat{\pi})$ and $f(t, x, v) = P_t f_0(x, v)$ we have*

$$\int_0^T \|f\|_{L^2(\hat{\pi})}^2 dt \leq (\gamma^2 C_0 + C_1) \int_0^T \langle f, (-\hat{\mathcal{L}}_v f) \rangle_{L^2(\hat{\pi})} dt$$

where C_0 and C_1 depend only on T and the spectral gap of the collapse m .

Following the broad proof structure in [LW22, CLW23, EL24a], this inequality is established using a quantitative divergence lemma. Note that in the general setting, the lemma is stated in terms of the collapse \mathcal{L} instead of the standard divergence.

Lemma 1.42 (Divergence lemma). *For $f \in L_0^2(\lambda \otimes \mu)$ there exist $h, g \in L^2(\lambda \otimes \mu)$ with*

$$\begin{aligned} h(t, \cdot), g(t, \cdot) &\in \text{Dom}(\mathcal{L}) \text{ for } \lambda\text{-a.e. } t, \\ h(\cdot, x) &\in H^1(0, T) \text{ for } \mu\text{-a.e. } x, \\ g(\cdot, x) &\in H^2(0, T) \text{ for } \mu\text{-a.e. } x, \end{aligned}$$

such that

$$\partial_t h - 2\mathcal{L}g = f,$$

with boundary conditions

$$\begin{aligned} h(t=0, \cdot) &= h(t=T, \cdot) = 0, \\ g(t=0, \cdot) &= g(t=T, \cdot) = 0, \end{aligned}$$

and

$$\begin{aligned} \|2\mathcal{L}g\|_{L^2(\lambda \otimes \pi)} &\leq c_1 \|f\|_{L^2(\lambda \otimes \pi)}, \\ \sqrt{2\mathcal{E}_T(h)} &\leq c_2 \|f\|_{L^2(\lambda \otimes \pi)}, \\ \sqrt{2\mathcal{E}_T(g)} &\leq c_3 \|f\|_{L^2(\lambda \otimes \pi)}, \\ \sqrt{2\mathcal{E}_T(\partial_t g)} &\leq c_4 \|f\|_{L^2(\lambda \otimes \pi)}, \end{aligned}$$

where $\lambda = 1_{[0, T]} dx$ and $\mathcal{E}_T(f) = \langle f, (-\mathcal{L})f \rangle_{L^2(\lambda \otimes \pi)}$.

For the instantaneous process, this hypocoercive strategy yields sharp exponential decay rate in agreement with the scaling of the total variation mixing time.

Proposition 1.43. *There exists $C(\omega, L) > 0$ such that the instantaneous process satisfies*

$$\|P_t f\|_{L^2(\pi)} \leq C(\omega, L) \exp\left(-\frac{K\omega}{(\omega L)^2 + 1} t\right) \|f\|_{L^2(\pi)}$$

for all $f \in L^2(\pi)$ with $\pi(f) = 0$, where $K > 0$ is an absolute constant.

For the overdamped Langevin with convex potential we recover the sharp quantitative rates of [CLW23] i.e.

$$\frac{\gamma m}{\gamma^2 + m}.$$

1.7 Perspectives

The following open questions are possible directions for future work.

Many-particle systems Computing the invariant measure of two general RTPs helped further our understanding of the invariant measure of the instantaneous process [HGM23]. Thus, exploring the invariant measure of three particle models with different tumbling rates could yield valuable insights. A deeper understanding of these three-particle systems might inspire an ansatz for the four-particle invariant measure, and potentially the n -particle measure, contributing to a precise quantitative understanding of MIPS.

Higher-dimensional systems The PDMP framework allows the rigorous modeling of higher-dimensional active particles with sticky collisions, a critical feature given that numerous physical systems exist in dimensions two or three. Such collisions result in particles skidding along each other instead of equilibrium, billiard-like reflections. This alternative to considering increasingly peaked soft repulsive potentials [AdPLVW19, AdPvW23] might enable the computation or approximation of the invariant measure. Difficult questions about boundary conditions should again find their answer in the domain of the generator.

Qualitative properties of invariant measures When the models get too complex, the strategies to explicitly compute the invariant measure laid out in the introduction break down. However, each characterization of the invariant measure possibly holds clues to its qualitative properties. This is especially true for the regularity of Markovian switching between ODEs [BH12, BLBMZ15], which remains an area of active research.

Relaxed assumptions for hypocoercive convergence The hypocoercive convergence results in the last chapter is proven under the assumption that the velocity possesses strong mixing properties. The question of whether the results can be extended to processes with noise of varying intensity or even vanishing noise under geometric control condition [DHHM22] is worth investigating.

Chapter 2

Introduction en français

2.1 Revue bibliographique et contexte de la recherche

Les particules run-and-tumble (RTPs) effectuent une série de ‘runs’ en ligne droite séparés par des réorientations aléatoires appelées ‘tumbles’. Le résultat est un mouvement aléatoire très différent des diffusions qui est un bon modèle pour le mouvement des bactéries comme *E. Coli* [Ber04]. Parce qu’elles consomment de l’énergie au niveau particulaire pour réaliser ce mouvement persistant, les bactéries sont hors de l’équilibre thermodynamique, offrant un modèle simple pour étudier les phénomènes en mécanique statistique hors équilibre tels que le mouvement collectif [VZ12], la formation de motifs [BB95], l’accumulation aux frontières [EG15] et la séparation de phase induite par la motilité (MIPS) [CT15]. Comprendre ce dernier phénomène, à savoir la séparation en une phase dense et une phase diluée, est la motivation théorique de ce manuscrit. Le mécanisme de base derrière ce phénomène est le suivant : les particules persistantes avec des interactions répulsives ralentissent dans les régions à forte densité de particules, ce qui augmente encore la densité dans ces parties de l’espace. Cela crée une boucle de rétroaction qui conduit à la formation de clusters, même si les interactions sont purement répulsives [CT15]. Nous précisons cette image générale en utilisant le cadre mathématique des processus de Markov déterministes par morceaux (PDMPs) [Dav93], une classe de processus stochastiques combinant un mouvement déterministe et des sauts aléatoires discrets. La question naturellement liée de la vitesse de convergence est abordée en utilisant les nombreux outils mathématiques conçus pour étudier le comportement en temps long des processus de Markov.

Une étape préliminaire importante pour comprendre MIPS est l’étude d’une seule particule run-and-tumble et de comment elle diffère de ses équivalents passifs et diffusifs, en particulier en termes de mesure invariante. Dans [Ang17], il est démontré que la mesure invariante d’un RTP unidimensionnel confiné dans un intervalle a un Dirac sur le bord. Il s’agit d’une forme forte d’accumulation près des frontières, une caractéristique partagée par la plupart des systèmes avec une seule RTP. Dans [MJK⁺18, Fry21], la présence de bruit thermique supplémentaire affaiblit l’accumulation aux frontières et les Diracs se diluent et deviennent des termes exponentiels concentrés près de la frontière. Notons qu’il y a des changements qualitatifs dans la mesure invariante d’un RTP dans le potentiel de confinement $V(x) = \alpha|x|^p$ ($p > 1$) en fonction des paramètres du modèle [DKM⁺19, BMR⁺20]. Pour certains ensembles de paramètres, la mesure invariante diverge aux bords de l’espace d’état, une forme forte d’accumulation hors équilibre. Pour d’autres choix de paramètres, une transition de forme se produit et la mesure invariante s’annule aux bords, la rapprochant de celle des particules passives.

Des signes potentiels de MIPS sont observés dans la mesure invariante de deux RTPs discrets en interaction [SEB16], qui a une masse accrue près des configurations de blocage sous forme d’exponentielles décroissantes et de Diracs. Dans la limite espace continu, où le pas de discrétisation tend vers zéro, les termes exponentiels deviennent des Diracs supplémentaires indiquant une attraction effective forte. Lorsqu’un état interne est ajouté à ce modèle

d'espace discret pour tenir compte de la durée non nulle du tumbling [SEB17], une seconde exponentielle décroissante apparaît et survit dans la limite espace continu. Enfin, lorsque les particules discrètes sautent pour se séparer instantanément lors des collisions [MEB23, MEB22], l'interaction effective peut être soit attractive, c'est-à-dire en cohérence avec une agrégation, ou répulsive en fonction des paramètres du modèle. Dans chaque cas, le modèle est traitable car il se réduit au cadre d'une seule particule. En effet, il suffit de considérer la particule relative, qui est à nouveau une particule active affichant cependant une dynamique de vitesse plus complexe. Cette réduction transforme l'agrégation en l'accumulation aux frontières déjà mentionnée. Travailler directement en espace continu conduit à des exponentielles décroissantes en présence de bruit thermique [DDK20a, LDMS21] et à des Dirac en son absence [LDMS21]. Deux RTPs continus soumis au potentiel d'interaction harmonique se réduisent au modèle de RTP unique dans [BMR⁺20] et affichent donc le même comportement de divergence ou d'annulation aux bords de l'espace.

Une caractéristique commune de ces modèles de RTP est que la mesure invariante consiste fréquemment en une combinaison de termes uniformes, exponentiels et de Dirac [SEB16, SEB17, Ang17, EM18, MJK⁺18, LDMS21, Fry21]. Cela s'explique par le fait que, sauf aux bords, les équations de continuité prennent la forme d'une EDO linéaire de sorte que la densité est donnée par une exponentielle matricielle. Établir les bonnes conditions aux bords est un problème récurrent, en particulier en température nulle avec blocage, ce qui a conduit au développement de plusieurs approches. D'abord, on peut discrétiser l'espace et considérer la limite où le pas de discrétisation tend vers zéro. Ensuite, on peut ajouter du bruit thermique et faire tendre la force du bruit vers zéro. Enfin, on peut considérer des potentiels répulsifs de courte portée de plus en plus piqués. L'idée générale ici est l'approximation par un autre système où les conditions aux bords sont plus faciles à comprendre. Cependant, à part en termes de conditions aux bords, ces approximations s'avèrent plus compliquées que le processus original et induisent une étape supplémentaire où une limite doit être calculée. Le formalisme mathématique des PDMPs résout ce problème et permet l'étude directe des modèles blocants en espace continu sans bruit thermique en utilisant le générateur du processus. Ce nouveau point de vue débloque de nombreuses mesures invariantes et nous permet d'identifier deux classes d'universalité hors équilibre pour les RTPs avec blocage et des mécanismes généraux de tumbling et de blocage [HGM23]. Ces classes correspondent soit à la conservation détaillée soit à la conservation globale du flux de probabilité, indiquant la proximité de l'équilibre et déterminant le comportement de la mesure invariante à travers un critère simple. Le cadre des PDMPs nous permet également de prouver rigoureusement que les modèles d'espace continu sont la limite des modèles sur réseau. De plus, nous trouvons la distribution invariante de trois RTPs se bloquant sur le tore 1D, une étape encourageante vers la résolution du cas n particules, qui reste un problème ouvert.

Une fois la mesure invariante connue, la question de l'échelle de temps où elle devient pertinente se pose naturellement. En effet, en temps court, le processus est fortement influencé par sa distribution initiale. La question 'quand le comportement asymptotique prend-il le dessus ?' est donc d'une importance cruciale. Pour cela, on peut se tourner vers la décomposition spectrale du générateur : le taux de convergence exponentielle asymptotique est donné par la deuxième valeur propre. Dans le cadre discret, le générateur est simplement une matrice et sa décomposition spectrale complète est fournie dans le cas d'un RTP unique et de deux RTPs interagissant [MBE19]. En espace continu, le générateur est un opérateur non borné et le problème aux valeurs propres prend la forme d'une équation différentielle [MJK⁺18, DDK20a]. Dans les cas où la transformée de Laplace des distributions à temps fixé peut être calculée [AG19, DKM⁺19], les valeurs propres peuvent être obtenues comme les pôles de la transformée de Laplace. Dans d'autres cas, la vitesse de convergence exponentielle peut être évaluée en approximant la distribution à temps fixe [SSK20].

Ces décompositions spectrales ont souvent une partie numérique, ce qui les empêche d'être totalement explicites. Un autre inconvénient de l'approche spectrale est que la diagonalisation des opérateurs non bornés est technique et que le problème aux valeurs propres ne donne qu'un sous-ensemble du spectre. De plus, dans le cas non réversible, on ne peut

espérer obtenir que la vitesse de convergence $Ce^{-\lambda t}$ où $-\lambda$ est la deuxième valeur propre du générateur et C est une constante qui ne peut être contrôlée, ce qui pose un problème lorsque $C \gg 1$. Ici encore, les obstacles sont surmontés en adoptant un point de vue différent. Des taux précis sont obtenus pour un RTP unidimensionnel avec des taux de tumbling asymétriques biaisés vers zéro [FGM12] en réduisant le problème au calcul de la transformée de Laplace d'un temps de premier passage. Des taux de tumbling biaisés plus généraux sont traités dans [FGM16] au prix de bornes moins fines. En utilisant également des arguments de couplage, nous obtenons des bornes non asymptotiques précises pour la vitesse de convergence de deux RTPs sur le tore 1D [GHM24] et sur la droite [Hah24]. Nous obtenons des bornes inférieures correspondantes grâce à des inégalités de concentration, montrant l'optimalité de nos résultats et déterminant complètement la dépendance de la vitesse de convergence en fonction des paramètres du modèle. Sur le tore 1D, cela révèle un régime persistant et un régime diffusif où le temps de mélange varie différemment en fonction des paramètres du modèle.

Ces résultats de convergence, ainsi que la compréhension des mesures invariantes et des classes d'universalité associées, trouvent leurs racines dans le dialogue entre les modèles de RTP de la mécanique statistique hors équilibre et le formalisme mathématique des processus de Markov déterministes par morceaux (PDMPs). L'étude de ces processus, définis par un mouvement déterministe ponctué de sauts stochastiques, est fortement motivée par leur utilisation en échantillonnage [MKK14, BCVD18, FBPR18, MDS20, LW22, MGM23] en physique statistique et en statistique bayésienne. Dans ce contexte, la vitesse de convergence est étroitement liée à l'efficacité des algorithmes d'échantillonnage, un goulot d'étranglement récurrent dans les applications. Un traitement unifié des PDMPs a été proposé pour la première fois par Davis [Dav84, Dav93] et les développements ultérieurs de la théorie peuvent être trouvés dans les reviews [Mal16, ABG⁺14] et les références qui y sont citées. La théorie générale a ensuite été étendue dans [DC99, CD08], où des propriétés de stabilité telles que la récurrence au sens de Harris ont été caractérisées en termes d'une chaîne squelette, et la propriété de Feller ainsi qu'un noyau du générateur de Feller ont été établies dans [DGM21]. Il est intéressant de noter que les PDMPs ne partagent pas les propriétés de régularisation des diffusions et ne sont en général pas fortement Feller.

Les sauts markoviens entre des EDOs dans \mathbb{R}^d constitue une sous-classe importante des PDMPs, ne différant des processus de RTP que par l'absence de frontières. Il a été montré dans [BH12, BLBMZ15] que la mesure invariante de ces processus a une densité par rapport à la mesure de Lebesgue, à condition que les champs de vecteurs correspondants aux EDOs satisfassent des conditions de crochet à la Hörmander. Bien que la régularité de ces processus soit une question ouverte en générale, une régularité C^k est obtenue [BT22, BT23] lorsque les taux de sauts sont suffisamment élevés. La régularité de la densité invariante pour des sauts entre deux champs de vecteurs lisses sur le tore est démontrée dans [BHLM18].

La question du comportement en temps long, qui se pose naturellement dans le cadre MIPS, est également pertinente dans le cadre plus large des PDMPs. Même des PDMPs simples peuvent présenter des explosions inattendues [LMR13, BLBMZ14]. L'un des résultats les plus surprenants est que des sauts markoviens entre deux EDOs linéaires $\partial_t x = A_0 x$ et $\partial_t x = A_1 x$ peut conduire à $x(t) \rightarrow +\infty$ presque sûrement, même pour A_i ayant uniquement des valeurs propres négatives. Dans des contextes plus favorables, par exemple un espace d'état compact, des arguments de couplage classiques conduisent à la convergence en distance de variation totale [BLBMZ15] sous des conditions de crochet de Hörmander ou une condition plus facile à vérifier sur la combinaison barycentrique des champs de vecteurs [BHS18]. Des dynamiques déterministes contractantes [BLBMZ12, CH15] conduisent à des bornes quantitatives dans une métrique de type Wasserstein. Des modèles spécifiques tels que le TCP [DGR02, MZ09, OK08, LVL08], utilisé pour la transmission de données sur Internet, ont également été étudiés [CMP10, BCG⁺13].

La convergence des PDMPs peut également être établie en utilisant des techniques d'hypocoercivité [Vil06, DMS09]. Récemment, des bornes quantitatives précises pour les PDMPs d'échantillonnage ont été obtenues [LW22] en utilisant une approche hypocoercive

[AAMN19, CLW23] développée à l'origine pour Langevin cinétique. Ceci est rendu possible par le fait que de nombreux PDMPs dans les applications, tels que les systèmes RTP [CRS14] et les processus d'échantillonnage [ADNR21], sont des processus cinétiques dans lesquels l'aléa affecte exclusivement la vitesse. L'approche a ensuite été simplifiée dans [EL24a] en introduisant un analogue en espace continu des chaînes de Markov liftées [CLP99, DHN00]. Nous adaptons ces techniques à deux RTPs sur le tore 1D, en les étendant à un cadre avec des bords et des dynamiques de vitesse très différentes. Nous établissons un résultat général qui fournit des taux précis pour le Langevin cinétique, les PDMPs d'échantillonnage, et deux RTPs sur le tore 1D.

En ce qui concerne contenu de la thèse, le chapitre 1 introduit le cadre PDMP et les outils standards pour étudier le comportement asymptotique des processus de Markov. Il se termine par une description de nos contributions et est traduit en français dans ce chapitre 2. Dans le chapitre 3, la mesure invariante de deux RTPs blocantes sur un tore 1D est calculée pour des mécanismes généraux de tumbling et de blocage en les formulant comme un PDMP. Cela révèle l'existence de deux classes d'universalité hors équilibre pour ce modèle conjecturé être la limite en espace continu d'un modèle sur réseau connu. Le chapitre 4 établit rigoureusement cette limite en espace continu. L'ordre de grandeur du temps de mélange est entièrement déterminée en fonction des paramètres du modèle en utilisant des techniques de couplage et des inégalités de concentration. Le chapitre 5 examine le comportement en temps long de deux RTPs sur une ligne avec potentiel d'interaction attractif et blocage. Des changements dans les paramètres du modèle induisent des changements qualitatifs dans la mesure invariante et des taux de décroissance exponentielle précis sont obtenus par couplage. Le chapitre 6 fournit la mesure invariante explicite pour trois RTPs sur un tore 1D avec blocage en utilisant un ansatz physique pour la mesure invariante dans la caractérisation par le générateur de l'invariance. Le chapitre 7 présente un résultat de convergence hypocoercif fin avec une preuve unifiée pour Langevin cinétique, les PDMPs d'échantillonnage et les RTPs basé sur un lemme de divergence général et une inégalité de Poincaré le long du flot.

2.2 Processus de Markov déterministes par morceaux

Nous commençons par une description rapide et informelle des processus de Markov déterministes par morceaux généraux [Dav84, Dav93], que nous utiliserons systématiquement pour formaliser des modèles de physique statistique hors équilibre tout au long du manuscrit. En quelques mots, un PDMP $X(t)$ suit un flux différentiel (ϕ_t) sur son espace d'états E et saute avec un taux $\lambda : E \rightarrow \mathbb{R}_+$ et une distribution post-saut donnée par un noyau de Markov $(Q(x, \cdot))_{x \in E}$. Plus précisément, lorsque le processus commence en $x_0 \in E$, nous avons

$$X(t) = \phi_t(x_0) \text{ pour } t \in [0, T)$$

où T est un temps aléatoire avec une fonction de répartition

$$\mathbb{P}(T \leq t) = 1 - \exp\left(-\int_0^t \lambda(\phi_s(x_0)) ds\right).$$

Au temps T , le processus saute vers un nouvel état x_1 distribué selon $Q(\phi_T(x_0), \cdot)$. Le processus redémarre alors de x_1 et tout se répète. Le flux différentiel peut sortir de l'espace d'états par un point $x_\partial \in \partial E$, auquel cas le processus saute instantanément dans E avec une distribution $Q_\partial(x_\partial, \cdot)$ où $(Q_\partial(x, \cdot))_{x \in \partial E}$ est un deuxième noyau de transition.

2.2.1 Exemples

Les trois sections suivantes complètent cette description générale en fournissant des exemples de PDMPs.

Processus instantané

Le premier exemple, modélisant deux marcheurs aléatoires persistants sur un tore 1D interagissant par des collisions bloquantes, illustrera plus tard des techniques pour calculer la mesure invariante et prouver la convergence vers celle-ci. Il s'agit d'une version simplifiée du modèle PDMP introduit dans [HGM23] pour travailler directement avec des RTPs bloquants en espace continu, ce qui permet un calcul efficace de la mesure invariante, met en évidence deux classes d'universalité hors équilibre, et permet de déterminer le temps de mélange en fonction des paramètres du modèle. Une version en espace discret avait précédemment été introduite en physique statistique hors équilibre [SEB16] comme moyen d'étudier des phénomènes d'agglomération.

Considérons deux particules sur un tore 1D de longueur L décrites par leurs positions $x_1, x_2 \in \mathbb{R}/L\mathbb{Z}$ et leurs vitesses $\sigma_1, \sigma_2 \in \{-1, +1\}$. Les σ_i sont deux processus markoviens de sauts indépendants alternant entre $+1$ et -1 avec un taux ω . Les positions évoluent comme suit

- si $x_1 \neq x_2$ (voir figure 2.1a) alors $\partial_t x_i = \sigma_i$,
- si $x_1 = x_2$ et si les vitesses des particules
 - les poussent l'une vers l'autre (voir figure 2.1b) alors $\partial_t x_i = 0$,
 - sont identiques (voir figure 2.1c) ou les éloignent l'une de l'autre (voir figure 2.1d) alors $\partial_t x_i = \sigma_i$.

Les deux particules sont ce que l'on appelle des particules run-and-tumble (RTPs) : elles alternent entre de longues phases de mouvement linéaire à vitesse constante (runs) et de courts mouvements erratiques entraînant une réorientation (tumbles). Dans ce modèle simplifié, la réorientation (basculement entre $+1$ et -1) est instantanée.

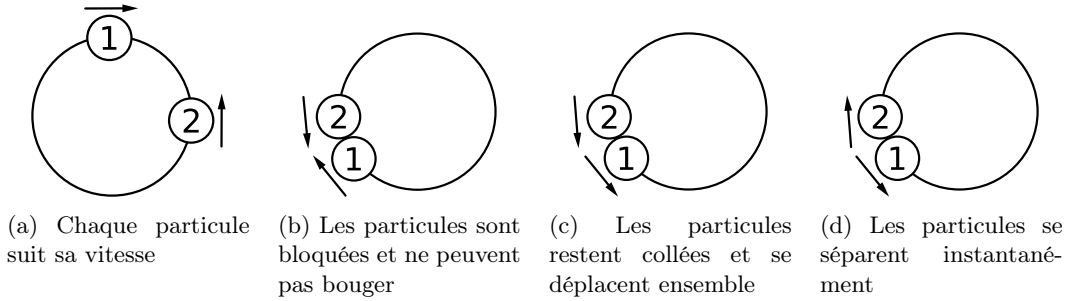


Figure 2.1: Deux RTPs bloquantes

Grâce aux symétries du modèle, il suffit de considérer x , la distance de x_1 à x_2 , et la vitesse relative $\sigma = \sigma_2 - \sigma_1$. Ce qui donne le processus stochastique suivant, dont une réalisation est donnée par la figure 2.3.

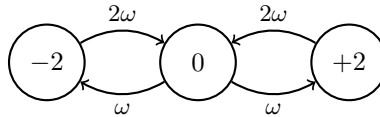


Figure 2.2: Taux de saut de vitesse relative

Definition 2.1 (Processus instantané). *Considérons l'espace d'états $E = [0, L] \times \Sigma$ où $L > 0$ et $\Sigma = \{-2, 0, 2\}$. Pour tout état initial $(x_0, \sigma_0) \in E$, soit*

- $\sigma(t)$ un processus markovien de sauts avec les taux de la figure 2.2, état initial σ_0 et temps de saut $0 = T_0 < T_1 < \dots$,

- $x(t)$ défini de manière récursive par $x(0) = x_0$ et

$$x(t) = \min[L, \max[0, x(T_i) + \sigma(T_i)(t - T_i)]] \text{ pour } T_i \leq t \leq T_{i+1}.$$

Nous appelons processus instantané $X(t) = (x(t), \sigma(t))$.

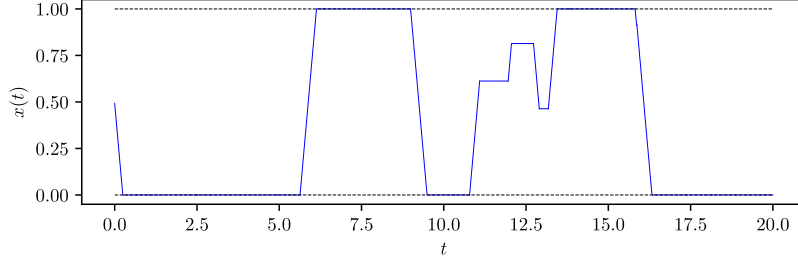


Figure 2.3: Réalisation du processus instantané ($\omega = 0.5, L = 1$)

Sauts markoviens entre EDOs

Le prochain exemple est une classe de PDMPs avec des applications en biologie moléculaire [GMK10], où le processus saute entre différentes EDOs de façon markovienne.

Definition 2.2 (Sauts markoviens entre EDOs). *Soit Σ un ensemble fini et*

- $F_\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d$ un champ vectoriel Lipschitz pour tout $\sigma \in \Sigma$,
- $q_{\sigma, \tilde{\sigma}} : \mathbb{R}^d \rightarrow \mathbb{R}_+$ un taux de saut C^0 pour tout $\sigma, \tilde{\sigma} \in \Sigma$.

Soit $(x_0, \sigma_0) \in \mathbb{R}^d \times \Sigma$ donné. Pour $i \geq 0$ nous définissons récursivement

- ξ_i comme étant la solution de $\partial_t \xi_i = F_{\sigma_i}(\xi_i)$ avec $\xi_i(0) = \xi_{i-1}(E_{i-1})$ (resp. $\xi_0(0) = x_0$ pour $i = 0$),
- E_i comme étant une variable aléatoire positive dépendant uniquement de σ_i avec une fonction de répartition

$$\mathbb{P}(E_i \leq t) = 1 - \exp\left(-\int_0^t \sum_{\sigma} q_{\sigma_i, \sigma}(\xi_i(s)) ds\right),$$

- σ_{i+1} comme étant une variable aléatoire dépendant uniquement de E_i et σ_i telle que

$$\mathbb{P}(\sigma_{i+1} = \sigma) = \frac{q_{\sigma_i, \sigma}(\xi_i(E_i))}{\sum_{\sigma} q_{\sigma_i, \sigma}(\xi_i(E_i))}.$$

Enfin, posons $T_i = \sum_{k=0}^{i-1} E_k$ et définissons

$$x(t) = \xi_i(t - T_i) \text{ et } \sigma(t) = \sigma_i \text{ pour } t \in [T_i, T_{i+1}).$$

Nous appelons processus de sauts markoviens entre EDOs $X(t) = (x(t), \sigma(t))$.

De nombreux aspects de cette classe de PDMPs, tels que la régularité de la mesure invariante [BH12, BLBMZ15], les propriétés de stabilité [LMR13, BLBMZ14], et le comportement de convergence [BLBMZ12, CH15], ont été largement étudiés. Notez que le processus instantané de la section précédente est presque un processus de sauts markoviens entre EDOs, mais son comportement au bord de l'espace d'états l'en sépare.

Randomized Hamiltonian Monte Carlo

Le troisième exemple est motivé par le rôle crucial joué par la distribution de Boltzmann

$$\pi = \frac{1}{Z} e^{-U(x)} dx \quad (2.1)$$

en physique statistique et permet un échantillonnage approché selon celle-ci. Dans cette section, nous supposons que $U : \mathbb{R}^d \rightarrow \mathbb{R}$ est un potentiel C^1 et que la normalisation $Z := \int_{\mathbb{R}^d} e^{-U(x)} dx$ est finie.

Definition 2.3 (Randomized Hamiltonian Monte Carlo [BRSS17]). *Soit $(x_0, v_0) \in \mathbb{R}^d \times \mathbb{R}^d$ donné. Définissons récursivement ξ_i comme étant la solution de*

$$\begin{aligned} \dot{x} &= v, \\ \dot{v} &= -\nabla_x U(x), \end{aligned}$$

avec la condition initiale $\xi_i(0) = (p_x(\xi_{i-1}(E_{i-1})), V_i)$ (resp. $\xi_0(0) = (x_0, v_0)$), où $(E_i)_{i \geq 0}$ est une suite de variables aléatoires exponentielles indépendantes de paramètre $\lambda > 0$, $(V_i)_{i \geq 1}$ est une suite de gaussiennes d -dimensionnelles indépendantes et p_x est la projection sur x .

Posons $T_i = \sum_{k=0}^{i-1} E_k$ et appelons

$$X(t) = \xi_i(t - T_i) \text{ pour } [T_i, T_{i+1})$$

randomized Hamiltonian Monte Carlo.

En quelques mots, le processus suit le flot hamiltonien et retire aléatoirement sa vitesse à des intervalles de temps distribués exponentiellement. Le randomized Hamiltonian Monte Carlo admet $\pi \otimes \mathcal{N}(0, I_d)$ comme sa mesure invariante unique et est ergodique [LW22]. Ainsi, échantillonner à partir de sa distribution à un temps fixé suffisamment grand permet d'échantillonner approximativement selon π .

En pratique, il est souvent difficile de simuler précisément les dynamiques hamiltoniennes, ce qui peut poser problème dans les applications. D'autres PDMPs utilisés pour l'échantillonnage incluent le forward event-chain Monte Carlo [MKK14], le processus Zig-Zag [BFR19] et le bouncy particle sampler [PdW12].

2.2.2 Semigroupe et générateur

L'étude des PDMPs repose sur la théorie générale des processus de Markov, dont les outils fondamentaux sont brièvement décrits dans cette section. Nous adoptons le cadre des processus de Feller [Lig10, Definition 3.1], une large classe de processus de Markov avec de bonnes propriétés de régularité. Bien que ce cadre n'englobe pas tous les PDMPs décrits par Davis [Dav93], il inclut tous les processus de ce manuscrit.

Un outil clé pour l'étude des PDMPs, motivé par la complexité inhérente aux probabilités de transition des processus de Markov en temps continu, est le semigroupe et son générateur.

Theorem 2.4 (Semigroupe Feller [Lig10, Theorem 3.15]). *Étant donné un processus Feller $X(t)$, définissons*

$$P_t f(x) = \mathbb{E}_x[f(X(t))]$$

pour $f \in C(E)$. Alors (P_t) est un semigroupe au sens de [Lig10, Definition 3.4].

Si nous définissons μP_t par

$$(\mu P_t)(f) = \mu(P_t f)$$

nous voyons que P_t agit sur l'espace des mesures par dualité. En fait, μP_t est la loi au temps t du processus avec mesure initiale μ , ce qui montre que le processus est entièrement déterminé par son semigroupe. Étudier le semigroupe peut être vu comme une façon d'étudier le

processus à travers son action sur l'espace des fonctions continues s'annulant à l'infini. Cela permet d'utiliser des outils d'analyse fonctionnelle et de théorie spectrale.

Definition 2.5 (Générateur Feller [Lig10, Theorem 3.16]). *Étant donné un semigroupe (P_t) , définissons \mathcal{L} par*

$$\mathcal{L}f = \lim_{t \rightarrow 0} \frac{P_t f - f}{t}$$

sur

$$D(\mathcal{L}) = \{f \in C(E) : \text{la limite (forte) ci-dessus existe}\}.$$

Alors \mathcal{L} est un générateur de probabilité au sens de [Lig10, Definition 3.12].

Le générateur constitue souvent l'outil de calcul principal dans l'étude des processus de Markov. Il peut, par exemple, être utilisé pour calculer des mesures invariantes (voir Section 2.3).

Proposition 2.6. *Le générateur Feller du processus instantané est donné par*

$$\mathcal{L}f(x, \sigma) = \sigma \partial_x f(x, \sigma) + \sum_{\tilde{\sigma}} q_{\sigma, \tilde{\sigma}} (f(x, \tilde{\sigma}) - f(x, \sigma))$$

où

$$\begin{pmatrix} q_{-2, -2} & q_{-2, 0} & q_{-2, +2} \\ q_{0, -2} & q_{0, 0} & q_{0, +2} \\ q_{+2, -2} & q_{+2, 0} & q_{+2, +2} \end{pmatrix} = \begin{pmatrix} -2\omega & 2\omega & 0 \\ \omega & -2\omega & \omega \\ 0 & 2\omega & -2\omega \end{pmatrix}$$

et a pour domaine

$$D(\mathcal{L}) = \{f(\cdot, \pm 2) \in C^1\} \cap \{\partial_x f(L, +2) = \partial_x f(0, -2) = 0\}.$$

Proposition 2.7. *L'ensemble $C^1(\mathbb{R}^d)$ constitue un cœur du générateur Feller \mathcal{L} du processus de sauts markoviens entre EDOs et*

$$\mathcal{L}f(x, \sigma) = \underbrace{\langle F_\sigma(x), \nabla_x f(x, \sigma) \rangle}_{\text{transport déterministe}} + \overbrace{\sum_{\tilde{\sigma}} q_{\sigma, \tilde{\sigma}}(x) (f(x, \tilde{\sigma}) - f(x, \sigma))}^{\text{sauts aléatoires}} \text{ pour tout } f \in C^1(E).$$

Proposition 2.8. *L'ensemble $C^1(\mathbb{R}^d)$ constitue un cœur du générateur Feller \mathcal{L} du randomized Hamiltonian Monte Carlo et*

$$\mathcal{L}f(x, v) = \underbrace{\langle v, \nabla_x f(x, v) \rangle - \langle \nabla_x U(x), \nabla_v f(x, v) \rangle}_{\text{dynamique Hamiltonienne déterministe}} + \lambda \underbrace{\int f(x, \tilde{v}) - f(x, v) d\kappa(\tilde{v})}_{\text{rafraîchissement de vitesse}}$$

où κ est la Gaussienne standard dans \mathbb{R}^d .

Remark 2.9. *D'autres définitions du générateur existent en dehors du cadre de Feller.*

- *Au lieu de considérer l'action du processus sur l'espace des fonctions continues, on peut considérer son action sur l'espace $L^2(\pi)$, ce qui conduit à un semigroupe et un générateur différents. Parce que $L^2(\pi)$ est un espace de Hilbert, cela est souvent le cadre approprié pour l'analyse spectrale et est particulièrement utile lorsque le processus est réversible, c'est-à-dire lorsque le générateur L^2 est autoadjoint.*
- *Le générateur Feller \mathcal{L} possède la propriété importante que pour tout $f \in D(\mathcal{L})$*

$$f(X(t)) - f(X(0)) - \int_0^t \mathcal{L}f(X(s)) ds$$

est une martingale pour toute distribution initiale $\mu \in \mathcal{P}(E)$. Prendre cette propriété comme la définition du générateur donne le générateur étendu. Cela étend le domaine du générateur à une classe de fonctions beaucoup plus large et permet l'utilisation d'outils probabilistes au lieu d'outils d'analyse fonctionnelle.

2.3 Mesure de probabilité invariante

Pour poursuivre notre présentation des concepts fondamentaux dans l'étude des processus de Markov, nous nous tournons vers les objets qui caractérisent leur comportement en temps long : les mesures invariantes. Une mesure π est dite invariante si

$$\pi P_t = \pi \text{ pour tout } t \geq 0. \quad (2.2)$$

Cela signifie que, si la distribution initiale est égale à π , la loi de $X(t)$ reste égale à π pour tout $t \geq 0$. Le théorème suivant regroupe des résultats dans [Par06, Section 7.6] pour illustrer comment ces mesures décrivent le comportement asymptotique.

Theorem 2.10 (Comportement asymptotique). *Soit $X(t)$ un processus markovien de sauts irréductible avec un espace d'états fini E , alors il existe une unique mesure invariante que nous notons π .*

(i) Loi des grands nombres markovienne: Pour tout $f : E \rightarrow \mathbb{R}$, nous avons

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t f(X(s)) ds = \pi(f)$$

\mathbb{P}_μ -presque sûrement pour tout $\mu \in \mathcal{P}(E)$.

(ii) Théorème central limite markovien: Pour tout $f : E \rightarrow \mathbb{R}$, nous avons

$$\sqrt{t} \left(\frac{1}{t} \int_0^t f(X(s)) ds - \pi(f) \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2(f))$$

en partant de n'importe quelle distribution initiale avec

$$\sigma^2(f) = \int \Gamma(g) d\pi$$

où g est une solution de $\mathcal{L}g = f - \pi(f)$ et $\Gamma(g) = \mathcal{L}(g^2) - 2g\mathcal{L}g$.

(iii) Convergence de la loi à temps fixé: Il existe $C, \gamma > 0$ tels que

$$\|\mu P_t - \pi\|_{\text{TV}} \leq C e^{-\gamma t} \|\mu - \pi\|_{\text{TV}}$$

pour toutes les distributions initiales $\mu \in \mathcal{P}(E)$ où $\|\cdot\|_{\text{TV}}$ est la distance en variation totale (voir Définition 2.16).

Remark 2.11. *Nous considérons uniquement des espaces d'états finis pour des raisons de simplicité, mais ces résultats peuvent être étendus à une large classe de processus de Markov en temps continu.*

Formellement, l'évolution de la loi de $X(t)$ est dictée par l'équation différentielle linéaire dans l'espace des mesures

$$\frac{d}{dt} \mu_t = \mathcal{L}^* \mu_t \quad (2.3)$$

où \mathcal{L}^* est l'adjoint pour la dualité mesures/fonctions. Cela suggère que l'équation (2.2) peut être réécrite comme

$$\mathcal{L}^* \pi = 0.$$

Le théorème suivant, qui est l'outil de calcul principal dans la recherche de mesures invariantes explicites, rend cela rigoureux.

Theorem 2.12 (Caractérisation par le générateur de l'invariance [Lig10, Theorem 3.37]). *Si D est un cœur du générateur \mathcal{L} alors*

$$\pi \text{ est invariante} \iff \int \mathcal{L}f d\pi = 0 \text{ pour tout } f \in D. \quad (2.4)$$

En voyant l'équation (2.3), les liens entre la propriété d'invariance de l'équation (2.2) et le comportement en temps long de $X(t)$ établis dans le théorème 2.10 peuvent être considérés comme une illustration du fait général suivant: les flots contractants convergent vers leur unique point fixe. La relation étroite entre le comportement en temps long et les mesures invariantes est aussi illustrée par la réciproque partielle suivante du théorème 2.10.

Theorem 2.13. *Si $X(t)$ est un processus Feller et $X(t) \xrightarrow{\mathcal{L}} \pi$, alors π est invariante.*

Notons que si nous pouvons tirer parti de notre compréhension physique du processus pour obtenir un ansatz pour la mesure invariante, nous pouvons prouver sa validité en utilisant le théorème 2.12. Illustrons la force de cette approche en apparence naïve en utilisant notre exemple récurrent. Nous savons que le processus instantané passe un temps strictement positif dans l'ensemble fini $\{0, L\} \times \Sigma$ et donc sa mesure invariante doit assigner une masse strictement positive à cet ensemble. De plus, la physique statistique à l'équilibre nous dit que deux RTPs avec des collisions réfléchissantes au lieu de bloquantes devraient avoir pour mesure invariante $(\frac{1}{L}dx) \otimes (\frac{1}{4}\delta_{-2} + \frac{1}{2}\delta_0 + \frac{1}{4}\delta_{+2})$. Ainsi

$$\pi = \sum_{\sigma} (d_{\sigma}^0 \delta_0 + a_{\sigma} dx + d_{\sigma}^L \delta_L) \otimes \delta_{\sigma}$$

est un ansatz naturel pour la mesure invariante du processus instantané. En insérant cet ansatz dans l'équation (2.4), on prouve sa validité et on obtient

$$\begin{aligned} d_{-2}^0 &= \frac{1/2}{2 + \omega L}, & a_{-2} &= \frac{\omega/4}{2 + \omega L}, & d_{-2}^L &= 0, \\ d_0^0 &= \frac{1/2}{2 + \omega L}, & a_0 &= \frac{\omega/2}{2 + \omega L}, & d_0^L &= \frac{1/2}{2 + \omega L}, \\ d_{+2}^0 &= 0, & a_{+2} &= \frac{\omega/4}{2 + \omega L}, & d_{+2}^L &= \frac{1/2}{2 + \omega L}. \end{aligned}$$

Il est important de noter que lorsque σ prend plus de valeurs, cet ansatz échoue et la mesure invariante présente un comportement hors équilibre riche [HGM23]. Par conséquent, des stratégies pour passer de la simple vérification des mesures invariantes à leur calcul sont nécessaires. Les trois sections suivantes sont consacrées à la recherche de la mesure invariante sans avoir à proposer un ansatz. Bien que les trois approches soient étroitement liées, elles sont suffisamment différentes pour que l'une puisse aboutir lorsque les autres échouent ou mènent à des calculs longs. Chaque méthode sera illustrée en utilisant notre exemple récurrent, le processus instantané.

2.3.1 Équation de Fokker-Planck

Résoudre l'équation de Fokker-Planck à l'intérieur de l'espace d'états est l'approche la plus polyvalente. Soit π une mesure invariante du processus instantané, que nous écrivons

$$\pi = \sum_{\sigma} (d_{\sigma}^0 \delta_0 + \pi_{\sigma} + d_{\sigma}^L \delta_L) \otimes \delta_{\sigma} \quad (2.5)$$

où π_{σ} est la restriction de π à $(0, L) \times \{\sigma\}$. Notons que, contrairement à l'approche précédente, nous ne faisons aucune hypothèse sur la forme de π_{σ} . Il s'agit d'une décomposition générale valable pour toute mesure sur E et non d'une hypothèse.

Considérons la fonction test

$$f(x, \sigma) = \begin{cases} f_{\sigma}(x) & \text{si } 0 < x < L, \\ 0 & \text{sinon,} \end{cases}$$

où les $f_{\sigma} \in C_c^{\infty}(0, L)$ sont arbitraires mais fixes. En utilisant cette fonction test particulière dans l'équation (2.4), on obtient

$$0 = \sum_{\sigma} \pi_{\sigma} \left(v_{\sigma} f'_{\sigma} + \sum_{\bar{\sigma}} q_{\sigma, \bar{\sigma}} f_{\bar{\sigma}} \right) = \sum_{\sigma} \left(-v_{\sigma} \pi'_{\sigma} + \sum_{\bar{\sigma}} q_{\bar{\sigma}, \sigma} \pi_{\bar{\sigma}} \right) (f_{\sigma})$$

où la différenciation est faite au sens des distributions. Parce que les f_σ étaient arbitraires, cela conduit à $-VII' + \mathcal{Q}^t \Pi = 0$ où

$$\Pi = \begin{pmatrix} \pi_{-2} \\ \pi_0 \\ \pi_{+2} \end{pmatrix}, \quad \mathcal{Q} = \begin{pmatrix} -2\omega & 2\omega & 0 \\ \omega & -2\omega & \omega \\ 0 & 2\omega & -2\omega \end{pmatrix}, \quad V = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & +2 \end{pmatrix},$$

et donc

$$\Pi = a(1, 2, 1)^t dx + b(\omega x + 1, 2\omega x, \omega x - 1)^t dx$$

pour certaines constantes a et b . Différents angles d'attaque peuvent ensuite être utilisés pour trouver a, b et d_σ^0, d_σ^L

1. utiliser les symétries du processus,
2. utiliser des arguments physiques pour trouver les conditions aux bords,
3. utiliser le fait que la marginale en σ de la mesure invariante est la mesure invariante du processus markovien de sauts $\sigma(t)$ à valeurs dans Σ ,

et enfin

4. réinjecter les informations obtenues dans les étapes précédentes dans l'équation (2.4).

Cette approche est bien adaptée pour calculer la mesure invariante des processus unidimensionnels [DKM⁺19, BMR⁺20, LDMS21, Fry21] car on peut utiliser les stratégies établies pour résoudre les systèmes d'équations différentielles. Cependant, cette approche échoue lorsque $-VII' + \mathcal{Q}^t \Pi = 0$ n'est pas explicitement résoluble.

2.3.2 Transformée de Laplace

L'approche suivante est basée sur la transformée de Laplace et est la version en espace continu de la stratégie dans [SEB16, SEB17]. Nous l'illustrons à nouveau en utilisant le processus instantané. Soit

$$\pi = \sum_{\sigma} (d_\sigma^0 \delta_0 + \pi_\sigma + d_\sigma^L \delta_L) \otimes \delta_\sigma \quad (2.6)$$

une mesure invariante et supposons cette fois que $\pi_\sigma = \pi_\sigma(x) dx$ a une densité lisse par rapport à la mesure de Lebesgue.

Étape 1: La première étape consiste à obtenir un système d'équations linéaires pour la transformée de Laplace de l'équation (2.6). En utilisant la fonction test

$$f^{\lambda, \sigma_0}(x, \sigma) = \begin{cases} e^{\lambda x} & \text{si } \sigma = \sigma_0, \\ 0 & \text{sinon,} \end{cases}$$

et l'équation (2.6) dans l'équation (2.4), on obtient

$$\int_0^L v_\sigma \lambda e^{\lambda x} \pi_{\sigma_0}(x) dx + \sum_{\sigma} q_{\sigma, \sigma_0} \left(d_\sigma^0 + \int_0^L e^{\lambda x} \pi_\sigma(x) dx + e^{\lambda L} d_\sigma^L \right) = 0.$$

Mettre les équations obtenues pour $\sigma_0 = -2, 0, 2$ sous forme vectorielle donne

$$\mathcal{Q}^t D^0 + \lambda V \hat{\Pi} + \mathcal{Q}^t \hat{\Pi} + e^{\lambda L} \mathcal{Q}^t D^L = 0, \quad (2.7)$$

où $\hat{\Pi} = (\hat{\pi}_{-2}, \hat{\pi}_0, \hat{\pi}_{+2})^t$, $\hat{\pi}_\sigma(\lambda) = \int_0^L e^{\lambda x} \pi_\sigma(x) dx$ et $D^x = (d_{-2}^x, d_0^x, d_{+2}^x)^t$ pour $x = 0, L$.

Étape 2: Soit $\Sigma^* = \{\sigma \in \Sigma : V_{\sigma, \sigma} \neq 0\}$. On réécrit ensuite l'équation (2.7) pour exprimer $(\hat{\pi}_\sigma)_{\sigma \in \Sigma^*}$ en termes de $(d_\sigma^x)_{\sigma \in \Sigma^*}$

$$\hat{\pi}_{-2} = \hat{\pi}_{+2} = \frac{\omega (d_{+2}^L e^{\lambda L} - d_{-2}^0)}{2\lambda}.$$

Étape 3: Le fait que $\hat{\pi}_{+2}(\lambda) = \int_0^L e^{\lambda x} \pi_{+2}(x) dx$ soit une fonction entière implique que le membre droit dans l'étape précédente n'a pas de pôle. On en déduit des contraintes linéaires sur $(d_\sigma^x)_{\sigma \in \Sigma^*}$

$$\operatorname{Res}_{\lambda=0}(\hat{\pi}_{+2}) = 0 \iff \frac{\omega}{2} (d_{+2}^L - d_{-2}^0) = 0 \iff d_{+2}^L = d_{-2}^0.$$

Étape 4: En utilisant l'équation (2.7), nous pouvons exprimer les transformées de Laplace $d_\sigma^0 + \hat{\pi}_\sigma + e^{\lambda L} d_\sigma^L$ pour $\sigma \notin \Sigma^*$ en termes des transformées de Laplace pour $\sigma \in \Sigma^*$

$$d_0^0 + \hat{\pi}_0(\lambda) + e^{\lambda L} d_0^L = \frac{1}{2} (d_{-2}^0 + \hat{\pi}_{-2}(\lambda) + e^{\lambda L} d_{-2}^L) + \frac{1}{2} (d_{+2}^0 + \hat{\pi}_{+2}(\lambda) + e^{\lambda L} d_{+2}^L)$$

de sorte qu'il ne reste plus qu'à inverser les transformées de Laplace et à imposer la normalisation $\pi(E) = 1$.

2.3.3 Temps d'arrêt

La dernière approche repose sur une caractérisation probabiliste de la mesure invariante qui n'implique pas le générateur. Elle est donc très différente des approches précédentes.

Theorem 2.14 (Caractérisation par temps d'arrêt [BH22, Théorème 6.26]). *Soit τ un temps d'arrêt, supposé fini au sens où $\mathbb{P}_{X_0}(\tau < +\infty)$ pour tout $X_0 \in E$. On peut donc définir les noyaux*

$$S(X_0, \cdot) = \mathbb{P}_{X_0}(X(\tau) \in \cdot),$$

et

$$G(X_0, \cdot) = \mathbb{E}_{X_0} \left[\int_0^\tau \delta_{X(s)}(\cdot) ds \right],$$

le premier étant un noyau de Markov. On suppose de plus que π_0 est une distribution stationnaire de S . Alors la mesure $\pi_0 G$ est stationnaire pour $X(t)$. En particulier, si $\mathbb{E}_{\pi_0}[\tau] < +\infty$, alors la mesure de probabilité

$$\pi := \frac{\pi_0 G}{\mathbb{E}_{\pi_0}[\tau]}$$

est stationnaire pour $X(t)$.

Remark 2.15. *Cette caractérisation est principalement utilisée pour prouver des propriétés qualitatives de la mesure invariante telles que son existence [Kha60, Kli87] ou sa régularité [BT22] et non pour obtenir une formule explicite.*

Illustrons cette dernière approche en utilisant le processus instantané. Si $E_0 = \{0, L\} \times \Sigma$ et que nous définissons

$$\tau = \inf\{t \geq 0 : X(t) \neq X(0) \text{ et } X(t) \in E_0\}$$

alors E_0 est invariant sous S au sens où $S((x, \sigma), E_0) = 1$ pour tout $(x, \sigma) \in E_0$. Ainsi, S peut être vu comme un noyau de Markov sur E_0 . Sa matrice de transition en une étape

$$P = \begin{matrix} & \begin{matrix} (0, -2) & (0, 0) & (0, +2) & (L, -2) & (L, 0) & (L, +2) \end{matrix} \\ \begin{matrix} (0, -2) \\ (0, 0) \\ (0, +2) \\ (L, -2) \\ (L, 0) \\ (L, +2) \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{L\omega}{L\omega+2} & 0 & 0 & 0 & 0 & \frac{2}{L\omega+2} \\ \frac{2}{L\omega+2} & 0 & 0 & 0 & 0 & \frac{L\omega}{L\omega+2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix} \quad (2.8)$$

peut être calculée en résolvant des systèmes d'EDO [Dav93, Chapitre 3]. Sa mesure invariante π_0 est donnée par

$$\begin{aligned}\pi_0(0, -2) &= 1/5, & \pi_0(0, 0) &= 1/5, & \pi_0(0, +2) &= 1/10, \\ \pi_0(L, -2) &= 1/10, & \pi_0(L, 0) &= 1/5, & \pi_0(L, +2) &= 1/5.\end{aligned}$$

Si l'état initial est $(x, \sigma) \in E_0 - \{(L, -2); (0, +2)\}$ alors $\tau = \inf\{t \geq 0 : X(t) \neq X(0)\}$ de sorte que τ a une distribution exponentielle et $\delta_{(x,\sigma)}G = \frac{1}{2\omega}\delta_{(x,\sigma)}$. Résoudre des systèmes d'EDO donne

$$\begin{aligned}\delta_{(L,-2)}G &= \frac{\omega x + 2}{2(L\omega + 2)}dx \otimes \delta_{-2} + \frac{\omega x + 1}{L\omega + 2}dx \otimes \delta_0 + \frac{\omega x}{2(L\omega + 2)}dx \otimes \delta_{+2}, \\ \delta_{(0,+2)}G &= \frac{\omega(L-x)}{2(L\omega + 2)}dx \otimes \delta_{-2} + \frac{\omega(L-x) + 1}{L\omega + 2}dx \otimes \delta_0 + \frac{\omega(L-x) + 2}{2(L\omega + 2)}dx \otimes \delta_{+2},\end{aligned}$$

et rend la mesure de probabilité invariante $\pi = \frac{\pi_0 G}{\pi_0 G(E)}$ entièrement explicite.

2.4 Convergence vers la mesure invariante

Le théorème 2.10 montre que les mesures invariantes sont un outil crucial dans l'étude des processus de Markov car elles dictent leur comportement en temps long. Cependant, le théorème ne donne aucune indication quant au moment où ce comportement asymptotique prend le dessus. Cette question s'avère être importante pour les applications. Par exemple, une compréhension quantitative de la vitesse de convergence vers la mesure invariante peut être utilisée pour comparer les performances de différents algorithmes d'échantillonnage par chaînes de Markov, en particulier la dépendance en la dimension.

De nombreuses approches existent et nous ne chercherons pas à être exhaustifs. Chaque technique donne une forme différente de convergence et est adaptée à une classe de processus de Markov différente. Pour cette raison, nous commençons par mettre en évidence un type important de processus de Markov, sur lequel nous illustrerons des techniques de convergence qui ne sont pas directement applicables aux PDMPs. Certaines techniques sont applicables aux deux classes, auquel cas nous utilisons l'exemple le plus simple.

2.4.1 Les processus de diffusion: une classe importante d'exemples

Les diffusions sont des processus de Markov forts avec des trajectoires continues possédant la propriété de Feller et seront utilisées comme alternative aux PDMPs pour démontrer des techniques de convergence dans la suite de cette section. Elles constituent une large classe de processus de Markov en temps continu et peuvent être définies comme la solution de l'équation différentielle stochastique [Oks13, LG16]

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t$$

où $b, \sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d$. Leur générateur est donné par

$$\mathcal{L}f(x) = \sum_i b_i(x)\partial_i f(x) + \sum_{ij} \left(\sigma(x)\sigma(x)^t\right)_{ij} \partial_{ij} f(x).$$

Nous utiliserons le processus d'Ornstein–Uhlenbeck (OUP)

$$dZ_t = -Z_t dt + \sqrt{2}dB_t$$

associé au générateur

$$\mathcal{L}f(x) = -xf'(x) + f''(x)$$

pour illustrer les techniques classiques de convergence. Sa mesure invariante est la loi Gaussienne standard.

2.4.2 Couplage

Une fois notre nouvel exemple en main, nous nous tournons vers la première technique de convergence qui exploite directement la caractérisation de la distance en variation totale et de la distance de Wasserstein en termes de couplages.

Definition 2.16 (Distance en variation totale [Jou24, Definition 1.3.1]). *La distance en variation totale entre deux mesures μ et ν est définie comme*

$$\|\mu - \nu\|_{\text{TV}} = \inf_{\Pi \in C(\mu, \nu)} \int 1_{\{x \neq y\}} d\Pi(x, y)$$

où $C(\mu, \nu)$ est l'ensemble de tous les couplages de μ et ν .

Cette distance fait de $\mathcal{P}(E)$ un espace métrique complet et la convergence par rapport à cette distance implique la convergence en loi. Si nous avons un couplage (X, \tilde{X}) où X est un processus de Markov de distribution initiale δ_x et \tilde{X} est le même processus de Markov démarré avec la distribution invariante π , il découle de la définition de la distance en variation totale que

$$\|\delta_x P_t - \pi\|_{\text{TV}} \leq \mathbb{P}(\tau > t)$$

où $\tau = \inf\{t \geq 0 : X(t) = \tilde{X}(t)\}$.

La distance suivante quantifie également la distance entre deux mesures.

Definition 2.17 (Distance de Wasserstein [Vil09, Definition 6.1]). (i) *Soit l'ensemble des mesures possédant un premier moment*

$$\mathcal{P}_1(E) := \{\mu \in \mathcal{P}(E) : \mu(d(x_0, \cdot)) < +\infty\}$$

où $x_0 \in E$ est arbitraire mais fixé. (La définition est indépendante de ce choix.)

(ii) *La distance de Wasserstein entre $\mu, \nu \in \mathcal{P}_1(E)$ est définie comme*

$$\|\mu - \nu\|_{\mathcal{W}^1} = \inf_{\Pi \in C(\mu, \nu)} \int d(x, y) d\Pi(x, y)$$

où $C(\mu, \nu)$ est encore l'ensemble de tous les couplages de μ et ν .

La distance de Wasserstein fait de $\mathcal{P}_1(E)$ un espace métrique complet et la convergence par rapport à cette distance implique la convergence en loi. Il découle à nouveau de la définition que si nous avons un couplage (X, \tilde{X}) où $X(t)$ (resp. $\tilde{X}(t)$) est un processus de Markov de mesure initiale δ_x (resp. π) alors

$$\|\delta_x P_t - \pi\|_{\mathcal{W}^1} \leq \mathbb{E}[d(X(t), \tilde{X}(t))].$$

La force de l'approche par couplage réside dans sa simplicité et sa généralité. Cependant, pour obtenir des bornes précises, il faut travailler au cas par cas.

Illustration à l'aide du OUP Soit (X, \tilde{X}) la solution de l'EDS

$$\begin{aligned} dX_t &= -X_t dt + \sqrt{2} dB_t, \\ d\tilde{X}_t &= -\tilde{X}_t dt + \sqrt{2} dB_t, \end{aligned}$$

c'est-à-dire deux OUPs issus du même mouvement brownien, aussi appelé couplage synchrone. Alors $D_t := X_t - \tilde{X}_t$ suit l'équation différentielle déterministe $dD_t = -D_t dt$ de sorte que $D_t = e^{-t} D_0$. Ainsi, en prenant $X_0 \sim \delta_x$ et $\tilde{X}_0 \sim \pi$, nous obtenons

$$\|\delta_x P_t - \pi\|_{\mathcal{W}^1} \leq \mathbb{E}|X_t - \tilde{X}_t| = e^{-t} \mathbb{E}|D_0| \leq e^{-t} \left(x + \sqrt{\frac{2}{\pi}} \right).$$

2.4.3 Condition de minoration uniforme

La méthode suivante donne également une convergence en distance en variation totale et repose sur le théorème du point fixe de Picard [Jou24, Section 1.3].

Definition 2.18. *La condition de minoration uniforme est satisfaite s'il existe $\nu \in \mathcal{P}(E)$ et $\alpha \in (0, 1)$ tels que*

$$\inf_{x \in E} \delta_x P_T(\cdot) \geq \alpha \nu(\cdot)$$

à $T > 0$ fixé.

Cette condition est également appelée condition de Doeblin.

Theorem 2.19. *Si la condition de minoration uniforme est satisfaite, alors il existe une unique mesure invariante π et*

$$\|\mu P_t - \pi\|_{\text{TV}} \leq (1 - \alpha)^{\lfloor t/T \rfloor}$$

pour tout $\mu \in \mathcal{P}(E)$.

Sous la condition de minoration uniforme, nous avons

$$\|\mu P_T - \nu P_T\|_{\text{TV}} \leq (1 - \alpha) \|\mu - \nu\|_{\text{TV}} \text{ pour tout } \mu, \nu \in \mathcal{P}(E).$$

Ainsi, P_T est une application contractante. Donc, puisque la distance en variation totale fait de $\mathcal{P}(E)$ un espace métrique complet, le théorème 2.19 découle du théorème du point fixe de Picard. Ce résultat de convergence est intéressant en raison de sa simplicité et de son aspect quantitatif, bien qu'il donne rarement des bornes fines. Notons également qu'il échoue souvent lorsque l'espace d'états n'est pas compact. Par exemple, dans le cas de l'OUP

$$\inf_{x \in E} \delta_x P_T(A) = 0 \tag{2.9}$$

pour tout ensemble mesurable A et tout $T > 0$ donc cette technique de convergence n'est pas applicable. C'est pourquoi nous l'illustrons en utilisant le processus instantané à la place.

Illustration à l'aide du processus instantané Le processus $\sigma(t)$ est un processus markovien de sauts donc les probabilités de transition sont données par une exponentielle matricielle et on peut calculer

$$\inf_{(x_0, \sigma_0) \in E} \mathbb{P}_{(x_0, \sigma_0)}(\sigma(1/\omega) = -2) = \frac{(e^2 - 1)^2}{4e^4}.$$

Notons également que si le processus instantané démarre avec une vitesse $\sigma_0 = -2$ et que la vitesse ne change pas, le processus atterrit forcément dans l'état $(0, -2)$ après un temps $L/2$. Ceci implique

$$\inf_{\sigma_0 = -2} \mathbb{P}_{(x_0, \sigma_0)}(X(L/2) = (0, -2)) \geq e^{-2\omega(L/2)}.$$

La propriété de Markov implique alors

$$\begin{aligned} & \mathbb{P}_{(x_0, \sigma_0)}(X(1/\omega + L/2) = (0, -2)) \\ & \geq \left(\inf_{(x_0, \sigma_0) \in E} \mathbb{P}_{\sigma_0}(\sigma(1/\omega) = -2) \right) \left(\inf_{\sigma_0 = -2} \mathbb{P}_{(x_0, \sigma_0)}(X(L/2) = (0, -2)) \right) \\ & \geq \frac{(e^2 - 1)^2}{4e^4} e^{-2\omega(L/2)} \end{aligned}$$

ce qui correspond bien à la définition 2.18 avec $T = 1/\omega + L/2$, $\alpha = \frac{(e^2 - 1)^2}{4e^4} e^{-\omega L}$ et $\nu = \delta_{(0, -2)}$.

2.4.4 Foster-Lyapunov

Cette section suit [HM11] et [Jou24], l'esprit général de cette approche vient de [MT93]. Les deux conditions suivantes généralisent la condition de minoration uniforme et peuvent souvent être appliquées même dans le cadre des espaces d'états non compacts où la condition de minoration uniforme échoue comme dans l'équation (2.9).

Definition 2.20. *La condition de dérive géométrique est satisfaite lorsqu'il existe une fonction $V : E \rightarrow \mathbb{R}_+$ et des constantes $K \geq 0$ et $\gamma \in (0, 1)$ telles que*

$$P_T V \leq \gamma V + K$$

pour un certain $T > 0$.

Definition 2.21. *La condition de minoration uniforme locale est satisfaite s'il existe $\nu \in \mathcal{P}(E)$ et $\alpha \in (0, 1)$ tels que*

$$\inf_{x \in C} \delta_x P_T(\cdot) \geq \alpha \nu(\cdot)$$

avec $C = \{x \in E : V(x) \leq R\}$ pour un $R > 2K/(1 - \gamma)$ où K, γ et T sont les constantes de la condition de dérive géométrique.

La condition de dérive géométrique garantit que le processus visite souvent l'ensemble de niveau C où la condition de minoration uniforme locale est satisfaite. Dans l'esprit de [MT93], elle peut être remplacée par la condition plus forte

$$\mathcal{L}V \leq -cV + K.$$

Definition 2.22 (*V-norme modifiée*). *Soit l'ensemble de mesures*

$$\mathcal{X} := \{\mu \in \mathcal{P}(E) : \mu(V) < +\infty\}.$$

Pour tous $\mu, \nu \in \mathcal{X}$, on définit la *V-norme modifiée*

$$\rho_\beta(\mu, \nu) := \int_E (1 + \beta V(x)) d|\mu - \nu|(x)$$

où $|\mu - \nu|$ est la variation.

Cette norme fait de \mathcal{X} un espace métrique complet.

Theorem 2.23 ([HM11, Theorem 1.3]). *Si la condition de dérive géométrique et la condition de minoration uniforme locale sont vérifiées, alors il existe $\bar{\alpha} \in (0, 1)$ et $\beta > 0$ tels que*

$$\rho_\beta(P_T \mu, P_T \nu) \leq \bar{\alpha} \rho_\beta(\mu, \nu) \text{ pour tous } \mu, \nu \in \mathcal{X}.$$

En particulier, pour tout $\alpha_0 \in (0, \alpha)$ et $\gamma_0 \in (\gamma + 2K/R, 1)$, on peut choisir $\beta = \alpha_0/K$ et

$$\bar{\alpha} = \max \left(1 - (\alpha - \alpha_0), \frac{2 + R\beta\gamma_0}{2 + R\beta} \right).$$

Le théorème suivant découle à nouveau du théorème du point fixe de Picard.

Theorem 2.24. *Si la condition de dérive géométrique et la condition de minoration uniforme locale sont satisfaites, alors il existe une unique probabilité invariante π . De plus, $\pi(V) \leq \frac{K}{1-\gamma}$ et*

$$\rho_\beta(\mu P_t, \pi) \leq \bar{\alpha}^{\lfloor t/T \rfloor} \rho_\beta(\mu, \pi) \text{ pour tout } \mu \in \mathcal{X}.$$

La force de ce résultat de convergence réside dans le fait qu'il fournit une borne quantitative et que, contrairement à la condition de minoration uniforme, il peut souvent être appliqué même quand la condition de minoration uniforme n'est pas vérifiée, par exemple pour l'OUP. Cependant, les bornes obtenues par cette méthode sont rarement précises.

Illustration avec l'OUP En prenant $V(x) = x^2$, on obtient

$$\mathcal{L}V(x) = -xV'(x) + V''(x) = -2x^2 + 2 = -2V(x) + 2$$

de sorte que la condition de dérive géométrique est satisfaite. De plus, les probabilités de transition sont connues et données par

$$\delta_x P_t = \sqrt{\frac{1}{2\pi(1-e^{-2t})}} \exp\left(-\frac{(y-xe^{-t})^2}{1-e^{-2t}}\right) dy =: p(t, x, y) dy$$

et donc, pour tout $t > 0$ fixé, nous avons

$$\inf_{x^2 \leq R} \inf_{0 \leq y \leq 1} p(t, x, y) := \alpha(t, R) > 0$$

impliquant la condition de minoration uniforme locale

$$\delta_x P_t(\cdot) \geq \alpha(t, R)(1_{\{0 \leq y \leq 1\}} dy)(\cdot).$$

Le théorème 2.24 entraîne alors la convergence en V -norme modifiée.

2.4.5 Inégalité de Poincaré

L'approche suivante est plus proche de l'analyse fonctionnelle que des probabilités et donne la décroissance exponentielle du semigroupe dans $L^2(\pi)$ [GGSC06, Section 2.1].

Hypothèse (A) Supposons que \mathcal{L} admette une base hilbertienne $(f_n)_{n \in \mathbb{N}}$ de fonctions propres dans $L^2(\pi)$ correspondant aux valeurs propres $(-\lambda_n)_{n \in \mathbb{N}}$. Supposons de plus $0 = \lambda_0 < \lambda_1 < \dots$ et $f_0 = 1$.

Remark 2.25. Si \mathcal{L} est auto-adjoint dans $L^2(\pi)$ et que sa résolvante est compacte, il admet une base hilbertienne de fonctions propres.

Dans ce contexte, la variance diminue de manière exponentielle avec un taux de décroissance donné par le trou spectral λ_1 .

Lemma 2.26. Si (A) est satisfaite alors $\text{Var}(P_t f) \leq e^{-2\lambda_1 t} \text{Var}(f)$.

L'optimalité du taux de décroissance se déduit en prenant $f = f_1$.

Proof. On a

$$\text{Var}_\pi(P_t f) = \left\| \sum_{n=1}^{+\infty} a_n e^{-\lambda_n t} f_n \right\|_{L^2(\pi)}^2 = \sum_{n=1}^{+\infty} a_n^2 e^{-2\lambda_n t} \leq e^{-2\lambda_1 t} \sum_{n=1}^{+\infty} a_n^2 = e^{-2\lambda_1 t} \text{Var}_\pi(f).$$

□

La preuve alternative suivante est instructive.

Preuve. Pour $f = \sum_{n=0}^{+\infty} a_n f_n \in D(\mathcal{L})$, on a

$$\langle f, \mathcal{L}f \rangle_{L^2(\pi)} = - \sum_{n=0}^{+\infty} \lambda_n a_n^2 \leq -\lambda_1 \sum_{n=1}^{+\infty} a_n^2 = -\lambda_1 \text{Var}_\pi(f)$$

et donc

$$\frac{d}{dt} \text{Var}_\pi(P_t f) = 2 \langle P_t f, \mathcal{L}P_t f \rangle_{L^2(\pi)} \leq -2\lambda_1 \text{Var}_\pi(P_t f)$$

ainsi la décroissance exponentielle souhaitée découle de Grönwall. □

Notons que dans la preuve alternative, la décomposition spectrale n'est pas réellement nécessaire, nous avons en fait simplement besoin de $\langle f, \mathcal{L}f \rangle_{L^2(\pi)} \leq -\lambda_1 \text{Var}_\pi(f)$. Ceci motive la définition suivante.

Definition 2.27. *L'inégalité de Poincaré de constante $C > 0$ est satisfaite si*

$$\text{Var}_\pi(f) \leq -C \langle f, \mathcal{L}f \rangle_{L^2(\pi)} \text{ pour tout } f \in D(\mathcal{L}).$$

Notons que dans le cas réversible, la constante de Poincaré optimale est l'inverse du trou spectral. Le théorème suivant complète le tableau de la relation entre la constante de Poincaré et le taux de décroissance asymptotique.

Theorem 2.28. (i) *L'inégalité de Poincaré de constante C est équivalente à*

$$\text{Var}_\pi(P_t f) \leq e^{-\frac{2}{C}t} \text{Var}_\pi(f)$$

pour tout $t \geq 0$ et $f \in L^2(\pi)$ (voir [Ebe09, Theorem 7.22]).

(ii) *Si \mathcal{L} est réversible, alors l'inégalité de Poincaré de constante C est de plus équivalente à l'existence de $K > 1$ telle que*

$$\text{Var}_\pi(P_t f) \leq K e^{-\frac{2}{C}t} \text{Var}_\pi(f)$$

pour tout $t \geq 0$ et $f \in L^2(\pi)$ (voir [CGZ13, Lemma 2.14]).

Dans le cas non réversible, si $\mathcal{L} = \mathcal{S} + \mathcal{A}$ où \mathcal{S} est symétrique et \mathcal{A} est anti-symétrique, on a $\langle f, \mathcal{L}f \rangle_{L^2(\pi)} = \langle f, \mathcal{S}f \rangle_{L^2(\pi)}$. Et donc, on ne peut espérer exploiter que la décroissance venant de \mathcal{S} et on perd forcément la décroissance venant de \mathcal{A} et de son interaction avec \mathcal{S} . En réalité, certains processus ne satisfont pas l'inégalité de Poincaré, ce qui sera illustré dans la section suivante.

Illustration avec l'OUP Les polynômes d'Hermite définis par

$$\text{He}_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}$$

satisfont l'équation différentielle

$$\text{He}_n'' - X \text{He}_n' = -n \text{He}_n$$

et forment une base orthogonale de $L^2(\pi)$. Par conséquent, le lemme 2.26 implique

$$\text{Var}_\pi(P_t f) \leq e^{-2t} \text{Var}_\pi(f).$$

2.5 Processus cinétiques et hypocoercivité

Cette section suit [Vil06, Section 3] et [Mon14, Section 1.4]. Plutôt que d'essayer de présenter une théorie générale, nous donnerons d'abord un exemple instructif où l'approche basée sur l'inégalité de Poincaré de la Section 2.4.5 échoue, puis nous montrerons comment l'approche peut être adaptée.

2.5.1 Un exemple sans Poincaré

Le processus de Langevin cinétique associé au potentiel $U(x) = \frac{1}{2}x^2$ est la solution de l'équation différentielle stochastique

$$\begin{aligned} dx_t &= v_t dt, \\ dv_t &= -x_t dt - v_t dt + \sqrt{2} dB_t. \end{aligned}$$

Il a pour générateur

$$\mathcal{L}f(x, v) = \underbrace{v\partial_x f(x, v) - x\partial_v f(x, v)}_{\text{flot hamiltonien}} - \underbrace{v\partial_v f + \partial_v^2 f(x, v)}_{\text{OUP en } v}$$

et pour mesure invariante $\pi = \frac{1}{2\pi} e^{-\frac{x^2+v^2}{2}} dx dv$. Le terme ‘cinétique’ se réfère au fait qu’il y a deux variables x et v et que le bruit aléatoire est restreint à la deuxième variable. Le processus instantané est également un processus cinétique car tout l’aléatoire provient de sa vitesse σ . Ce faible niveau de bruit est une source de difficulté supplémentaire. Par exemple, pour Langevin cinétique, l’inégalité de Poincaré prend la forme

$$\text{Var}_\pi(f) \leq C \|\partial_v f\|_{L^2(\pi)}^2,$$

qui ne peut pas être satisfaite puisque le membre droit s’annule lorsque f dépend uniquement de x . Ainsi, nous ne pouvons pas utiliser Poincaré pour obtenir la convergence vers la mesure invariante.

2.5.2 Une approche hypocoercive

Dans une tentative d’étendre l’inégalité de Poincaré à ce contexte, on peut considérer la norme H^1

$$\|f\|_{H^1(\pi)}^2 = \|f\|_{L^2(\pi)}^2 + \|\partial_x f\|_{L^2(\pi)}^2 + \|\partial_v f\|_{L^2(\pi)}^2$$

et essayer de prouver une inégalité de la forme

$$\frac{d}{dt} \|P_t f\|_{H^1(\pi)}^2 \leq -C \|P_t f\|_{H^1(\pi)}^2.$$

Cependant, on a

$$\begin{aligned} \frac{d}{dt} \|\partial_x P_t f\|_{L^2(\pi)}^2 &= -2 \langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)} - 2 \|\partial_{xv} P_t f\|_{L^2(\pi)}^2, \\ \frac{d}{dt} \|\partial_v P_t f\|_{L^2(\pi)}^2 &= -2 \|\partial_v P_t f\|_{L^2(\pi)}^2 + 2 \langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)} - 2 \|\partial_v^2 P_t f\|_{L^2(\pi)}^2, \end{aligned}$$

de sorte que

$$\frac{d}{dt} \|P_t f\|_{H^1(\pi)} = -4 \|\partial_v P_t f\|_{L^2(\pi)} - 2 \|\partial_v^2 P_t f\|_{L^2(\pi)} - 2 \|\partial_{xv} P_t f\|_{L^2(\pi)}$$

et donc $\frac{d}{dt} \|P_t f\|_{H^1(\pi)}^2$ s’annule à nouveau si f dépend uniquement de x .

Terme croisé En s’inspirant du théorème de régularité hypoelliptique de Hörmander [Hör67] où les champs de vecteurs interagissent à travers leurs crochets de Lie pour créer les directions manquantes, nous considérons le terme croisé $\langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)}$ dont la dérivée est

$$\begin{aligned} \frac{d}{dt} \langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)} &= \\ \underbrace{\|\partial_x P_t f\|_{L^2(\pi)}^2}_{\text{terme souhaité}} - \|\partial_v P_t f\|_{L^2(\pi)}^2 - \langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)} - 2 \langle \partial_{xv} P_t f, \partial_v^2 P_t f \rangle_{L^2(\pi)}. \end{aligned}$$

Remarquons maintenant que la norme

$$\|f\|_\epsilon^2 = \|f\|_{L^2(\pi)}^2 + \epsilon \left(\|\partial_x f\|_{L^2(\pi)}^2 + \|\partial_v f\|_{L^2(\pi)}^2 + \langle \partial_x f, \partial_v f \rangle_{L^2(\pi)} \right)$$

est équivalente à $\|\cdot\|_{H^1(\pi)}$ pour tout $\epsilon > 0$ et calculons

$$\begin{aligned} \frac{d}{dt} \|P_t f\|_\epsilon^2 &= -\epsilon \|\partial_x P_t f\|_{L^2(\pi)}^2 + (-2 - \epsilon) \|\partial_v P_t f\|_{L^2(\pi)}^2 + \epsilon \langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)} \\ &\quad - 2\epsilon \|\partial_v^2 P_t f\|_{L^2(\pi)}^2 - 2\epsilon \|\partial_{xv} P_t f\|_{L^2(\pi)}^2 + 2\epsilon \langle \partial_{xv} P_t f, \partial_v^2 P_t f \rangle_{L^2(\pi)}. \end{aligned}$$

Puisque

$$\begin{aligned} \langle \partial_x P_t f, \partial_v P_t f \rangle_{L^2(\pi)} &\leq \frac{1}{2} \|\partial_x P_t f\|^2 + \frac{1}{2} \|\partial_v P_t f\|^2, \\ \langle \partial_{xv} P_t f, \partial_v^2 P_t f \rangle_{L^2(\pi)} &\leq \frac{1}{2} \|\partial_{xv} P_t f\|^2 + \frac{1}{2} \|\partial_v^2 P_t f\|^2, \end{aligned}$$

nous obtenons

$$\frac{d}{dt} \|P_t f\|_\epsilon^2 \leq -\frac{\epsilon}{2} \|\partial_x P_t f\|_{L^2(\pi)}^2 - \left(2 + \frac{\epsilon}{2}\right) \|\partial_v P_t f\|_{L^2(\pi)}^2$$

donc en utilisant l'inégalité de Poincaré pour π (par rapport au gradient complet) et l'équivalence de $\|\cdot\|_\epsilon$ et $\|\cdot\|_{H^1(\pi)}$, nous obtenons qu'il existe $C > 0$ tel que

$$\frac{d}{dt} \|P_t f\|_\epsilon^2 \leq -C \|P_t f\|_\epsilon^2.$$

L'inégalité de Grönwall implique alors la décroissance exponentielle dans H^1 . On peut ensuite récupérer la décroissance exponentielle dans L^2 en utilisant des estimations de régularité (par exemple [Vil09, Theorem A.8]).

Sélection de résultats connexes La stratégie globale peut être résumée comme suit: trouver une norme à laquelle l'inégalité de Grönwall peut être appliquée. Les nombreuses approches alternatives conçues pour traiter des processus similairement dégénérés forment la théorie de l'hypo-coercivité. Changer la norme est également au cœur de l'approche hypo-coercive Dolbeault-Mouhot-Schmeiser [DMS09, DMS15]. La principale différence avec la dérivation ci-dessus est que la norme à laquelle l'inégalité de Grönwall est appliquée est équivalente à la norme L^2 et non à la norme H^1 . Cette approche L^2 est utilisée dans [ADNR21] pour obtenir la convergence exponentielle pour des PDMPs d'échantillonnage avec une dépendance explicite du taux de décroissance en la dimension.

En utilisant l'inégalité de Poincaré en espace et en temps introduites dans [AAMN19], des bornes fines pour la convergence de Langevin cinétique sont données dans [CLW23]. En particulier, après optimisation du coefficient de friction, la dépendance optimale \sqrt{m} du trou spectral en fonction du trou spectral m du Langevin suramorti correspondant est atteinte pour les potentiels convexes. La méthode est ensuite revisitée pour obtenir la convergence exponentielle de PDMPs d'échantillonnage avec des taux de décroissance fins [LW22]. Ces résultats sont simplifiés dans [EL24a] en utilisant un analogue continu des chaînes de Markov liftées [DHN00, CLP99].

2.6 Contributions

Cette section donne un aperçu des contributions originales de cette thèse, qui concernent toutes le comportement en temps long de RTPs.

- La mesure invariante de deux RTPs sur le tore 1D ainsi que la droite réelle est calculée en utilisant la caractérisation générateur de l'invariance. Ce point de vue sur les conditions aux bords révèle deux classes d'universalité hors équilibre sur le tore 1D et des différences qualitatives dans la mesure invariante sur la ligne. Les deux traduisent différents degrés de déviation par rapport à l'équilibre. Un ansatz pour la mesure à trois particules sur le tore 1D est obtenu en exploitant la compréhension du cas deux particules et les symétries du modèle, puis démontré correct en utilisant le générateur. Ceci est un pas vers la mesure à n particules, qui reste un problème ouvert.
- Des bornes fines sur la vitesse de convergence vers la mesure invariante sont obtenues en utilisant des techniques de couplage. Des bornes inverses sont déduites d'inégalités de concentration, prouvant l'optimalité de ces résultats et déterminant entièrement

la dépendance de la vitesse de convergence en fonction des paramètres du modèle. Sur le tore 1D, un régime persistant et un régime diffusif, avec des temps de mélange différents, sont observés.

- Les résultats de convergence hypocoercive basés sur des processus de Markov liftés et des inégalités de Poincaré le long du flot sont étendus et unifiés pour obtenir des bornes de convergence fines dans un cadre couvrant à la fois les résultats précédents, tels que Langevin cinétique et les PDMPs d'échantillonnage, et les RTPs.

2.6.1 Deux particules run-and-tumble sur le tore 1D



Figure 2.4: Virage instantané

Modèle discret préexistant Le processus suivant a été introduit dans [SEB16] pour étudier l'agglomération de deux particules run-and-tumble bloquantes, un modèle pour les bactéries comme E. Coli. Considérons deux particules sur un réseau périodique unidimensionnel représentées par leurs positions $y_1^N, y_2^N \in \mathbb{Z}/N\mathbb{Z}$ et leurs vitesses $\sigma_1, \sigma_2 \in \{\pm 1\}$. Ces vitesses sont deux copies indépendantes du processus markovien de sauts de la Figure 2.4a et une horloge de Poisson indépendante de taux $\gamma > 0$ est attachée à chaque particule. Lorsque leur horloge sonne, les particules sautent de leur position actuelle y_i^N à la nouvelle position $y_i^N + \sigma_i$, sauf si l'autre particule s'y trouve déjà. Ainsi, les particules ne peuvent pas se croiser et interagissent par blocage. L'interaction entre les sauts de position aléatoires et les retournements de vitesse est illustrée dans la Figure 2.5.

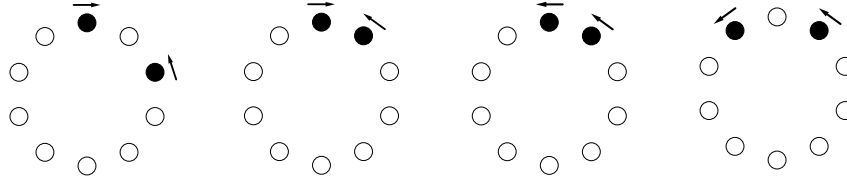


Figure 2.5: Séquence de sauts de position aléatoires et de retournements de vitesse

Il suffit de considérer la distance $y^N \in \{1, \dots, N\}$ de la première à la deuxième particule ainsi que la vitesse relative $\sigma = \sigma_2 - \sigma_1$ pour modéliser le système. Le processus $\sigma(t)$ est un processus markovien de sauts avec les taux de transition de la Figure 2.4b et le couple $Y^N(t) = (y^N(t), \sigma(t))$ est le processus markovien de sauts avec les taux de la Figure 2.6 et l'espace d'états $E_N = \{1, \dots, N\} \times \{-2, 0, +2\}$.

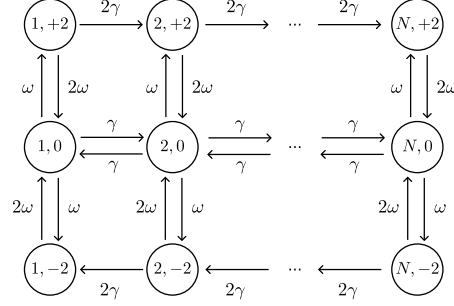


Figure 2.6: Transitions du processus RTP discret

Dans [SEB16], l'expression explicite de la mesure invariante π_N de $Y^N(t)$, qui présente une accumulation de masse aux bords sous la forme de Diracs et d'exponentielles décroissantes, est déterminée¹. De plus, la limite espace continu de π_N est identifiée comme la mesure invariante

$$\pi = \sum_{\sigma \in \Sigma} (d_\sigma^0 \delta_0 + a_\sigma dx + d_\sigma^L \delta_L) \otimes \delta_\sigma$$

du processus instantané de la Section 2.2.1 calculée dans la Section 2.3. Cependant, le sens exact de cette convergence (convergence en loi, variation totale, Wasserstein...) n'est pas entièrement clair et la preuve de la convergence n'est pas très détaillée.

Limite espace continu Le fait que $\pi_N \rightarrow \pi$ dans la limite espace continu conduit à la question suivante: est-il possible de montrer que l'ensemble du processus $Y^N(t)$ converge vers le processus instantané $X(t)$?

Pour répondre à cette question, nous commençons par définir l'injection $i_N : E_N \rightarrow E$ par

$$i_N(y, \sigma) = \left(L \frac{y-1}{N-1}, \sigma \right)$$

ce qui nous permet de voir l'espace d'états E_N comme une discrétisation de l'espace d'états du processus instantané $E = [0, L] \times \{-2, 0, +2\}$.

La limite en espace continu est la limite où l'espacement des points du réseau $\frac{L}{N-1}$ tend vers 0 tandis que la vitesse moyenne des particules $\gamma \frac{L}{N-1}$ et le taux de retournement de vitesse ω restent constants, c'est-à-dire

$$\left. \begin{array}{l} L \rightarrow +\infty \\ \gamma = \frac{N-1}{L} \\ \omega \text{ constant} \end{array} \right\}. \quad (*)$$

La proposition suivante montre alors que le processus instantané $X(t)$ est la limite de distribution de $Y^N(t)$, reliant rigoureusement le modèle sur réseau dans [SEB16] et sa version en espace continu.

Proposition 2.29 (Limite discrète-continue). (i) *Il existe un couplage du processus instantané X et une séquence de processus discrets $(Y^N)_{N \geq 2}$ avec des paramètres qui évoluent comme dans (*) tels que pour tout $T > 0$*

$$\mathbb{P} \left(\sup_{t \leq T} |i_N(y^N(t)) - x(t)| \geq \epsilon \right) \leq \frac{1}{\epsilon} \left(\frac{L}{N-1} + 8 \sqrt{\frac{T \left((\omega T)^2 + 3\omega T + 1 \right) L}{N-1}} \right).$$

¹Le processus dans [SEB16] est en fait $(y^N(t), \sigma_1(t), \sigma_2(t))$ et non $Y^N(t) = (y^N(t), \sigma(t))$ mais tous les résultats se traduisent facilement entre les modèles.

(ii) Nous avons

$$\left((i_N(y^N(t)), \sigma(t)) \right)_{t \geq 0} \xrightarrow{\mathcal{L}} \left(X(t) \right)_{t \geq 0}$$

dans l'espace de Skorokhod $\mathcal{D}[0, +\infty)$.

L'idée principale est de construire un espace de probabilité commun pour $X(t) = (x(t), \sigma(t))$ et $Y^N(t) = (y^L(t), \sigma^N(t))$ où

$$\sigma(t) = \sigma^N(t) \text{ p.s. pour tout } t \geq 0.$$

En d'autres termes, les vitesses sont identiques et nous devons examiner la convergence de $i_N(y^N(t))$, qui effectue des sauts de plus en plus fréquents et de plus en plus petite taille, vers le mouvement déterministe de $x(t)$. Sur tout intervalle de temps où $\sigma = +2$ est constant, la situation se réduit essentiellement à l'exemple jouet $\tilde{y}^N(t)$ de la Figure 2.7, où l'espace d'états est $\{0, \dots, N\}$, le processus saute de k à $k+1$ avec un taux N si $k \leq N-1$, et N est absorbant.

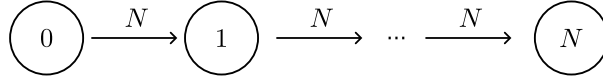


Figure 2.7: Taux de transition de l'exemple jouet $\tilde{y}^N(t)$

La proposition suivante montre que cet exemple jouet converge vers la translation vers la droite sur $[0, 1]$ avec une frontière absorbante en $x = 1$.

Proposition 2.30 (Exemple jouet). *Soit $\tilde{y}^N(t)$ le processus de sauts de Markov avec l'espace d'états $\{0, \dots, N\}$, les taux de transition de la Figure 2.7 et l'état initial $\tilde{y}^N(0) = 0$. Si $\tilde{x}(t) = \min(1, t)$, alors on a*

$$\mathbb{P} \left(\sup_{t \leq T} |\tilde{y}^N(t)/N - \tilde{x}(t)| \geq \epsilon \right) \leq \frac{T}{\epsilon^2 N}.$$

Proof. Remarquons que $\tilde{y}^N(t)$ a la même loi que $\min(U^N(t), N)$, où $U^N(t)$ est un processus de Poisson de taux N . Ainsi, en utilisant le fait que $x \mapsto \min(x, 1)$ est 1-Lipschitz, on obtient

$$\begin{aligned} \mathbb{P} \left(\sup_{t \leq T} |\tilde{y}^N(t)/N - \tilde{x}(t)| \geq \epsilon \right) &= \mathbb{P} \left(\sup_{t \leq T} |\min(U^N(t)/N, 1) - \min(t, 1)| \geq \epsilon \right), \\ &\leq \mathbb{P} \left(\sup_{t \leq T} |U^N(t)/N - t| \geq \epsilon \right), \\ &= \mathbb{P} \left(\sup_{t \leq T} |U^N(t) - Nt| \geq N\epsilon \right). \end{aligned}$$

L'inégalité de Doob donne alors

$$\mathbb{P} \left(\sup_{t \leq T} |\tilde{y}^N(t)/N - \tilde{x}(t)| \geq \epsilon \right) \leq \frac{\mathbb{E} \left[(U^N(T) - NT)^2 \right]}{(\epsilon N)^2} = \frac{T}{\epsilon^2 N}.$$

□

Établir que la convergence de Y^N vers π_N est uniforme en N permet ensuite d'échanger la limite espace continu et la limite temps long et de montrer rigoureusement la convergence

$$\|i_N \# \pi_N - \pi\|_{\mathcal{W}^1} \rightarrow 0$$

renforçant le cadre théorique de l'approche de discrétisation dans [SEB16, SEB17, MEB22, MEB23].

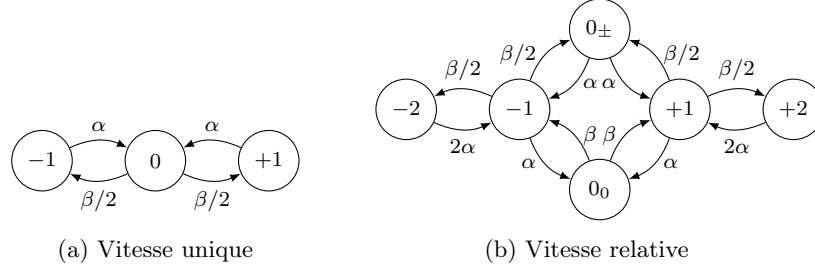


Figure 2.8: Tumble instantané

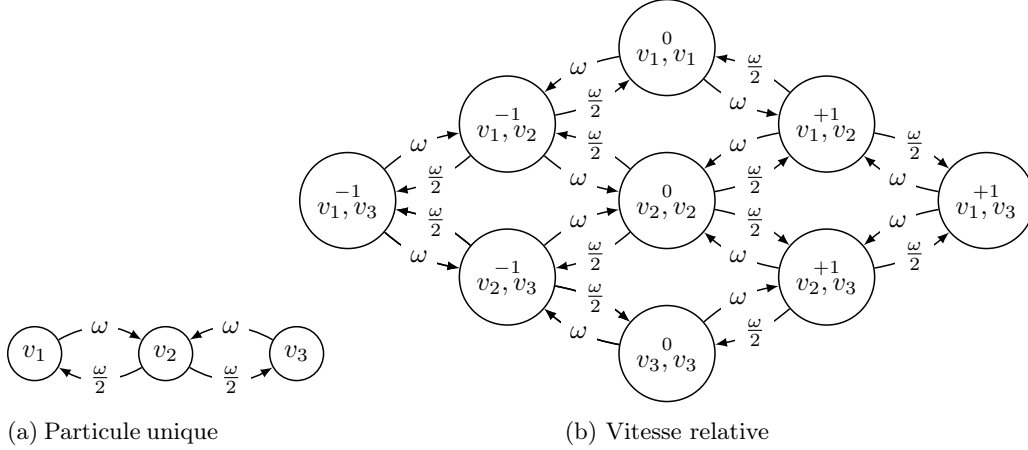


Figure 2.9: Vitesses anisotropes

Mesure invariante et classes d'universalité Dans [SEB17], le calcul de la mesure invariante et de sa limite espace continu est étendu au cadre où chaque particule suit les taux de la Figure 2.8a. L'état supplémentaire $\sigma_i = 0$ correspond au tumble qui a désormais une durée strictement positive, reflétant plus fidèlement la réalité biologique des bactéries comme *E. Coli* car la durée du tumble n'est pas négligée [SSB12]. Notons que, dans ce cas, le processus $\sigma = \sigma_2 - \sigma_1$ n'est pas markovien à moins que l'état $\sigma_2 - \sigma_1 = 0$ ne soit divisé en deux états 0_{\pm} (correspondant à $\sigma_1 = \sigma_2 = \pm 1$) et 0_0 (correspondant à $\sigma_1 = \sigma_2 = 0$), comme c'est fait dans la Figure 2.8b.

Nous appelons 'processus fini' la version espace continu (en raison de la durée finie du tumble). En utilisant l'approche Fokker-Planck de la Section 2.3.1 pour calculer la mesure invariante, on obtient

$$\pi = \sum_{\sigma \in \Sigma} (d_{\sigma}^0 \delta_0 + \pi_{\sigma}(x) dx + d_{\sigma}^L \delta_L) \otimes \delta_{\sigma}, \quad (2.10)$$

où

$$\pi_{\sigma}(x) = a_{\sigma} + b_{\sigma}^c \cosh(\kappa(x - L/2)) + b_{\sigma}^s \sinh(\kappa(x - L/2)),$$

et $a_{\sigma}, b_{\sigma}^c, b_{\sigma}^s$ sont des fonctions explicites des paramètres du modèle. Il s'agit d'une version plus compacte de la limite espace continu dans [SEB17] et qui met en évidence les symétries du modèle. À un facteur scalaire près, le vecteur des termes uniformes $(a_{\sigma})_{\sigma \in \Sigma}$ coïncide avec la mesure invariante du processus de vitesse $\sigma(t)$.

Notons que, contrairement à [SEB16], l'accumulation exponentielle aux bords est présente même en espace continu. Dans le cas des vitesses anisotropes de la Figure 2.9, on a

$$\pi_{\sigma}(x) = a_{\sigma} + \sum_{i=1,2} [b_{\sigma}^{c,i} \cosh(\kappa_i(x - L/2)) + b_{\sigma}^{s,i} \sinh(\kappa_i(x - L/2))]$$

où les coefficients sont à nouveau explicites en fonction des paramètres du modèle. On voit ici une accumulation exponentielle à deux échelles différentes.

Le formalisme PDMP et les symétries du modèle permettent le calcul de la mesure invariante d'une classe beaucoup plus large de processus où la vitesse σ peut suivre n'importe quel processus markovien de sauts avec des taux qui peuvent différer à l'intérieur de l'espace d'états $(0, L) \times \Sigma$ et aux bords $\{0, L\} \times \Sigma$, c'est-à-dire des taux dépendant du contact ou non des particules. La mesure invariante prend toujours la forme de l'Équation (2.10) et, tout comme dans la Section 2.3.1, le vecteur des densités par rapport à la mesure de Lebesgue $(\pi_\sigma(x))_{\sigma \in \Sigma}$ est obtenu en résolvant une EDO linéaire, il est donc donné par une exponentielle matricielle. Fixer les poids des Diracs sur le bord donne alors une formule explicite pour la mesure invariante, ce qui généralise [SEB17].

Il est vrai en général que le vecteur des termes uniformes $(a_\sigma)_{\sigma \in \Sigma}$ est déterminé par la mesure invariante de $\sigma(t)$. L'EDO est déterminée par le comportement de $\sigma(t)$ à l'intérieur de l'espace d'états tandis que les conditions aux bords sont fixées par le mécanisme de blocage, ce qui donne une séparation nette des contributions du blocage et du tumble à la mesure invariante.

Tous les modèles partagent le pendant actif suivant du bilan détaillé

$$\sum_{\sigma \in \Sigma} \sigma \pi_\sigma(x) = 0 \text{ pour tout } x \in (0, L), \quad (2.11)$$

que nous utilisons pour identifier deux classes d'universalité hors équilibre. La classe detailed jamming est caractérisée par la condition plus stricte

$$\pi_{-\sigma}(x) = \pi_\sigma(x) \text{ pour tout } (x, \sigma) \in (0, L) \times \Sigma, \quad (2.12)$$

alors que pour la classe global jamming, uniquement l'Équation (2.11) est satisfaite.

On a donc un analogue du bilan détaillé et du bilan global respectivement qui donne un critère simple pour déterminer le comportement de la mesure invariante. Ceci reste valide lorsque l'espace d'état du processus de vitesse $\sigma(t)$ est continu. Pour les processus dans la classe detailed jamming, comme le processus instantané, on a que les $\pi_\sigma(x) = a_\sigma$ sont uniformes et que la restriction $\sum_\sigma \pi_\sigma(x) dx \otimes \delta_\sigma$ de la mesure invariante à l'intérieur de l'espace d'états est une mesure produit. En revanche, les processus dans la classe global jamming, tels que le processus fini et le processus avec des vitesses anisotropes, ont des mesures invariantes dont la restriction à l'intérieur n'est ni uniforme ni une mesure produit. En comparant avec le processus à l'équilibre avec la même dynamique à l'intérieur mais avec des collisions réfléchissantes au lieu de blocantes, qui a une mesure invariante $\pi = \sum_\sigma a_\sigma dx \otimes \delta_\sigma$, on voit que la classe detailed jamming est plus proche de l'équilibre tandis que la classe global jamming en est plus éloignée. Notons que l'Équation (2.12) est équivalente à

$$\pi_{-\sigma}(0) = \pi_\sigma(0) \text{ pour tout } \sigma \in \Sigma,$$

la classe du processus se déduit donc du mécanisme de blocage.

Remark 2.31. *Les limites discrètes-continues rigoureuses pour la mesure invariante et l'ensemble du processus peuvent également être établies lorsque $\sigma(t)$ suit la Figure 2.8b.*

Here is the translation of the provided LaTeX passage into French:

Temps de mélange Pour obtenir des bornes supérieures sur le temps de mélange, nous appliquons un argument de couplage avec le couplage synchrone.

Definition 2.32 (Couplage synchrone). *Soient $(x_0, \sigma_0), (\tilde{x}_0, \tilde{\sigma}_0) \in [0, L] \times \Sigma$. Considérons le couplage de deux processus de saut de Markov $\sigma(t)$ et $\tilde{\sigma}(t)$ avec des taux de la Figure 2.4b (resp. Figure 2.8b) et des états initiaux σ_0 et $\tilde{\sigma}_0$ tels que $\sigma(t)$ et $\tilde{\sigma}(t)$ sont indépendants jusqu'au temps d'arrêt $\tau_\sigma = \inf\{t \geq 0 : \sigma(t) = \tilde{\sigma}(t)\}$ et ensuite égaux pour toujours. Définissons ensuite récursivement $x(t)$ (resp. $\tilde{x}(t)$) comme dans la Définition 2.1 avec l'état initial x_0 (resp. \tilde{x}_0) et la vitesse $\sigma(t)$ (resp. $\tilde{\sigma}(t)$). On appelle couplage synchrone le couple (X, \tilde{X}) avec $X(t) = (x(t), \sigma(t))$ et $\tilde{X}(t) = (\tilde{x}(t), \tilde{\sigma}(t))$.*

Proposition 2.33 (Borne supérieure sur le temps de mélange). (i) Dans le cas du processus instantané, pour tout $\epsilon \in (0, 1)$ il existe $C(\epsilon) > 0$ tel que

$$t_{\text{mix}}(\epsilon) \leq C(\epsilon) \frac{1}{\omega} ((\omega L)^2 + 1)$$

où

$$t_{\text{mix}}(\epsilon) = \inf \left\{ t \geq 0 : \sup_{\mu \in \mathcal{P}(E)} \|\mu P_t - \pi\|_{\text{TV}} \leq \epsilon \right\}.$$

(ii) Dans le cas du processus fini, pour tout $\epsilon \in (0, 1)$ il existe $C(\epsilon) > 0$ tel que

$$t_{\text{mix}}(\epsilon) \leq C(\epsilon) \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) ((\alpha L)^2 + 1).$$

Idée de preuve. La préservation de l'ordre de x et \tilde{x} après τ_σ découle de définition récursive

$$\begin{aligned} x(\tau_\sigma) \leq \tilde{x}(\tau_\sigma) &\implies x(t) \leq \tilde{x}(t) \text{ pour tout } t \geq \tau_\sigma, \\ x(\tau_\sigma) \geq \tilde{x}(\tau_\sigma) &\implies x(t) \geq \tilde{x}(t) \text{ pour tout } t \geq \tau_\sigma. \end{aligned}$$

Par conséquent, en supposant sans perte de généralité que $x(\tau_\sigma) \leq \tilde{x}(\tau_\sigma)$, si $\tilde{x}(t^*) = 0$ pour un certain $t^* \geq \tau_\sigma$ alors

$$0 \leq x(t^*) \leq \tilde{x}(t^*) = 0 \implies x(t^*) = \tilde{x}(t^*).$$

Ainsi, pour le couplage synchrone, $\mathbb{P}(X(t) \neq \tilde{X}(t))$ peut être bornée en contrôlant les queues de distribution de τ_σ et $\tau_0 = \inf\{t \geq 0 : x(t) = 0\}$, qui sont explicitement calculables pour le processus instantané et peuvent être estimées pour le processus fini. \square

Proposition 2.34 (Borne inférieure sur le temps de mélange). (i) Dans le cas du processus instantané, pour tout $\epsilon \in (0, 1)$ il existe $\tilde{C}(\epsilon) > 0$ tel que

$$t_{\text{mix}}(\epsilon) \geq \tilde{C}(\epsilon) \frac{1}{\omega} ((\omega L)^2 + 1).$$

(ii) Dans le cas du processus fini, pour tout $\epsilon \in (0, 1)$ il existe $\tilde{C}(\epsilon) > 0$ tel que

$$t_{\text{mix}}(\epsilon) \leq \tilde{C}(\epsilon) \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) ((\alpha L)^2 + 1).$$

Idée de preuve. On identifie deux obstacles au mélange, c'est-à-dire deux obstacles à ce que la distribution à temps fixe soit proche de la mesure invariante. Premièrement, suffisamment de temps doit s'être écoulé pour qu'au moins un saut de vitesse se produise. Cela conduit à

$$t_{\text{mix}}(\epsilon) \geq \frac{\log(1/\epsilon)}{2\omega} \text{ et } t_{\text{mix}}(\epsilon) \geq \frac{\log(1/\epsilon) \max(1/\alpha, 1/\beta)}{2}$$

pour les processus instantané et fini respectivement. Deuxièmement, suffisamment de temps doit s'être écoulé pour qu'un changement d'ordre L dans la position x se produise. Établir des résultats de concentration sur $\int_0^t \sigma(s) ds$ donne

$$t_{\text{mix}}(\epsilon) \geq C'(\epsilon) \omega L^2 \text{ et } t_{\text{mix}}(\epsilon) \geq C'(\epsilon) (1/\alpha + 1/\beta) (\alpha L)^2$$

pour le processus instantané dans le régime $\omega L \gg 1$ et pour le processus fini dans le régime $\alpha L \gg 1$ respectivement. \square

Notons que les bornes supérieures et inférieures sur le temps de mélange sont les mêmes, à une constante multiplicative près, ce qui signifie que nous avons entièrement déterminé sa dépendance en fonction des paramètres du modèle. Le temps de mélange peut être compris comme la combinaison de deux temps de mélanges correspondants respectivement à un régime persistant où $\omega L \ll 1$ (resp. $\alpha L \ll 1$) et un régime diffusif où $\omega L \gg 1$ (resp. $\alpha L \gg 1$). Dans le régime persistant, l'ordre de grandeur du temps nécessaire pour traverser l'espace d'état si aucun renversement de vitesse ne se produit L est inférieur à $1/\omega$ (resp. $1/\alpha + 1/\beta$), l'ordre de grandeur du temps moyen entre deux changements de vitesse. Par conséquent, le temps de mélange est essentiellement le temps nécessaire pour qu'au moins un changement de vitesse se produise. Au contraire, dans le régime diffusif, traverser l'espace d'état prend beaucoup plus de temps que changer de vitesse et donc le temps de mélange est essentiellement le temps nécessaire pour voir un changement significatif de position.

2.6.2 Deux particules run-and-tumble sur la droite réelle

Modèle On sort maintenant du cadre espace d'états compact pour considérer deux particules run-and-tumble bloquantes sur la droite réelle, soumises à un potentiel d'interaction attractif V . Les vitesses des particules, σ_1 et σ_2 , sont des copies indépendantes du même processus markovien de sauts, tandis que leurs positions, x_1 et x_2 , évoluent selon les EDOs

$$\partial_t x_1 = -V'(x_1 - x_2) + v\sigma_1(t) \text{ et } \partial_t x_2 = -V'(x_2 - x_1) + v\sigma_2(t).$$

Ainsi, la position relative $x = x_2 - x_1$ est régie par l'EDO

$$\partial_t x = -2V'(x) + v\sigma(t)$$

pour $x > 0$, où $\sigma = \sigma_2 - \sigma_1$ est la vitesse relative, dont l'espace d'états est noté Σ . En ajoutant le blocage des particules, on obtient un PDMP $X(t) = (x(t), \sigma(t))$ avec l'espace d'états $[0, +\infty) \times \Sigma$.

Nous nous concentrons sur trois instances spécifiques de ce modèle

- le processus linéaire instantané où $V(x) = c|x|$ et σ suit les taux de la Figure 2.4b,
- le processus linéaire fini où $V(x) = c|x|$ et σ suit les taux de la Figure 2.8b,
- le processus harmonique instantané où $V(x) = \mu x^2/2$ et σ suit les taux de la Figure 2.4b.

Ces modèles sont similaires à ceux de [LDMS21], mais avec deux différences importantes: les interactions de blocage empêchent les croisements de particules, et la vitesse du processus linéaire fini est plus proche de la réalité biologique car la durée du tumble n'est pas négligée [SEB17, SSB12].

Mesure invariante Dans le cas linéaire instantané, l'approche de Fokker-Planck décrite à la Section 2.3.1 implique la résolution d'une EDO vectorielle linéaire $\dot{X} = AX$. La mesure invariante est alors donnée en fonction de l'exponentielle matricielle de A , qui est calculée en utilisant sa décomposition spectrale. L'intégrabilité sur $[0, +\infty)$ implique que seules les valeurs propres négatives et les vecteurs propres associés apparaissent dans la mesure invariante. Les masses de Dirac sur le bord peuvent être fixées en utilisant le fait que la marginale en σ de la mesure invariante est la mesure invariante du processus $\sigma(t)$.

Proposition 2.35. *L'unique mesure invariante du processus linéaire instantané est*

$$\pi = \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + a_\sigma e^{\zeta x} dx) \otimes \delta_\sigma$$

où $\zeta = -\frac{2c\omega}{v^2-c^2}$ et

$$\begin{aligned} d_2 &= 0, & d_0 &= \frac{c}{c+v}, & d_{-2} &= \frac{cv}{(c+v)^2}, \\ a_2 &= \frac{c\omega}{2(v^2-c^2)}, & a_0 &= \frac{c\omega}{(c+v)^2}, & a_{-2} &= \frac{c(v-c)\omega}{2(c+v)^3}. \end{aligned}$$

La même approche peut être utilisée pour le processus linéaire fini. Cependant, la matrice A est de dimension plus grande et dépend de plus de paramètres, donc sa décomposition spectrale explicite est plus délicate.

Proposition 2.36. *L'unique mesure invariante du processus linéaire fini est*

$$\pi = \begin{cases} \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + c_2 a_\sigma(\zeta_2) e^{\zeta_2 x} dx) \otimes \delta_\sigma & \text{quand } c < v \leq 2c, \\ \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + c_2 a_\sigma(\zeta_2) e^{\zeta_2 x} dx + c_3 a_\sigma(\zeta_3) e^{\zeta_3 x} dx) \otimes \delta_\sigma & \text{quand } v > 2c, \end{cases}$$

où c_i, ζ_i, d_σ et a_σ sont des fonctions explicites de α, β, c et v . De plus, quand $v < c \leq 2v$, on a

$$d_{+2} = 0 \text{ et } d_\sigma > 0 \text{ sinon,}$$

tandis que quand $c > 2v$, on a

$$d_{+2} = d_{+1} = 0 \text{ et } d_\sigma > 0 \text{ sinon.}$$

Remarquons que lorsque $c \leq 2v$, la mesure invariante est donnée une unique exponentielle à l'intérieur de l'espace $\{x > 0\}$, tandis que lorsque $c > 2v$, une exponentielle supplémentaire apparaît en même temps qu'une Dirac disparaît du bord. Ce changement qualitatif indique un passage d'une forme d'agglomération plus forte à une forme plus faible, comme l'indique la masse autour de $x = 0$ sous la mesure invariante. Ceci ressemble à la transition de forme de la mesure invariante dans [DKM⁺19, BMR⁺20] où la mesure invariante peut soit s'annuler, soit diverger au bord de l'espace d'états en fonction des paramètres, indiquant différents degrés de déviation de l'équilibre.

Dans le cas harmonique instantané, le fait que le système d'équations différentielles pour la mesure invariante ne soit plus linéaire rend l'approche de Fokker-Planck plus difficile. Heureusement, une fois restreint à l'intérieur de l'espace d'états $\{x > 0\}$, le processus coïncide avec la restriction du processus dans [BMR⁺20]. Étant donné que les deux processus entrent toujours dans cette partie de leur espace d'états par le même point $(0, +2)$, leurs mesures invariantes doivent coïncider sur $\{x > 0\}$, à une constante multiplicative près. Les Diracs sur le bord peuvent à nouveau être fixés en utilisant la marginale σ , ce qui conduit au résultat suivant.

Proposition 2.37. *L'unique mesure invariante du processus harmonique instantané est*

$$\pi = \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + \pi_\sigma(x) dx) \otimes \delta_\sigma$$

où π_σ s'annule en dehors de l'intervalle $[0, v/\mu]$. En

- $x = 0$ la marginale $x \sum_{\sigma} \pi_\sigma$ diverge pour $\mu \geq 2\omega$ et admet une limite finie pour $\mu < 2\omega$,
- $x = v/\mu$ la marginale $x \sum_{\sigma} \pi_\sigma$ diverge pour $\mu > 2\omega$, admet une limite positive pour $\mu = 2\omega$ et s'annule pour $\mu < 2\omega$.

De plus $d_2 = 0$ et $d_0, d_{-2} > 0$ pour tous les choix de paramètres.

Les densités π_σ sont des combinaisons de fonctions hypergéométriques. La Proposition 2.37 montre que le phénomène d'agglomération peut prendre différentes formes en fonction des paramètres du modèle.

Convergence L'argument de couplage avec le couplage synchrone donne des bornes de convergence explicites et non asymptotiques qui dépendent de la transformée de Laplace du temps d'atteinte $\tau_0 = \inf\{t \geq 0 : x(t) = 0\}$. Ces bornes impliquent les bornes asymptotiques suivantes, plus faciles à énoncer et à interpréter.

Proposition 2.38. (i) *Dans le cas linéaire instantané*

$$\frac{\omega}{2} \frac{c^2}{v^2} \leq \liminf_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}}.$$

pour tout $(x, \sigma) \in [0, +\infty) \times \Sigma$.

(ii) *Dans le cas linéaire fini*

$$\frac{3 - \sqrt{5}}{8} \min \left(\alpha \left(1 + \frac{\alpha}{\beta} \right) \frac{c^2}{v^2}, \alpha \frac{c}{v} \right) \leq \liminf_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}}$$

pour tout $(x, \sigma) \in [0, +\infty) \times \Sigma$.

Idée de preuve. On considère à nouveau le couplage synchrone où les deux copies σ et $\tilde{\sigma}$ de la vitesse coïncident après un temps exponentiel et restent ensuite égales pour toujours. Le même argument de préservation de l'ordre que dans l'idée de preuve de la Proposition 2.33 transforme la compréhension quantitative de la queue de distribution de τ_0 en bornes de convergence. Remarquer

$$\tau_0 > t \implies x(0) + v \int_0^t \sigma(s) ds - 2ct > 0 \implies \frac{1}{t} \int_0^t \sigma(s) ds > \frac{2c}{v} - \frac{x(0)}{vt}$$

permet l'utilisation de résultats de grandes déviations non asymptotiques [Wu00, CG08, Lez01] pour obtenir des bornes exponentielles pour de la queue de distribution τ_0 . \square

On étudie l'optimalité des bornes ci-dessus en établissant des bornes supérieures correspondantes.

Proposition 2.39. (i) *Dans le cas linéaire instantané*

$$\overline{\lim}_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \leq 4\omega \frac{c^2}{v^2}$$

pour tout $(x, \sigma) \in [0, +\infty) \times \Sigma$.

(ii) *Dans le cas linéaire fini*

$$\overline{\lim}_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \leq 4\alpha \left(1 + \frac{\alpha}{\beta} \right) \frac{c^2}{v^2}$$

pour tout $(x, \sigma) \in [0, +\infty) \times \Sigma$.

Idée de preuve. Le point clé est que le processus met longtemps pour explorer la partie de l'espace d'états où x est grand. Par définition de la distance en variation totale, on a

$$\|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \geq \pi(\{x > at\}) - \mathbb{P}_{(x,\sigma)}(x(t) > at)$$

donc si on peut trouver $a > 0$ tel que

$$\underbrace{\lim_{t \rightarrow +\infty} -\frac{1}{t} \log \mathbb{P}(x(t) > at)}_{=:\lambda(a)} > \underbrace{\lim_{t \rightarrow +\infty} -\frac{1}{t} \log \pi(\{x > at\})}_{=:\lambda_\pi(a)}$$

alors

$$\lim_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \leq \lambda_\pi(a).$$

La valeur de $\lambda_\pi(a)$ se déduit de l'expression explicite de la mesure invariante. En définissant $\tau_L = \inf\{t \geq 0 : x(t) \geq L\}$, on obtient que pour tout $u < 0$

$$\mathbb{P}(x(t) > at) \leq \mathbb{P}(\tau_{at} < t) \leq e^{-ut} \mathbb{E}[e^{u\tau_{at}}],$$

et $\mathbb{E}[e^{u\tau_{at}}]$ peut être calculé en résolvant un système d'EDO. \square

Remarquons que les bornes sur le taux de convergence exponentiel du processus linéaire instantané sont les mêmes, à une constante multiplicative près. Ce n'est pas exactement le cas pour le processus linéaire fini mais

$$\alpha \frac{c}{v} \ll \alpha \left(1 + \frac{\alpha}{\beta}\right) \frac{c^2}{v^2}$$

ne peut se produire que lorsque $\alpha \gg \beta$, ce qui signifierait que les RTPs passent la majorité de leur temps en tumble et non en run, ce qui ne correspond pas à la réalité physique.

Dans le cas harmonique instantané, la dynamique déterministe est contractante au sens où

$$|x(t_2) - \tilde{x}(t_2)| \leq e^{-\mu(t_2-t_1)} |x(t_1) - \tilde{x}(t_1)|,$$

si $\sigma(t) = \tilde{\sigma}(t)$ pour tout $t \in [t_1, t_2]$. Ceci peut être exploité pour obtenir un résultat de convergence quantitatif en distance de type Wasserstein [BLBMZ12].

2.6.3 Trois particules run-and-tumble sur le tore 1D

Modèle On s'intéresse maintenant à trois particules run-and-tumble bloquantes sur le tore 1D de longueur L . Ceci correspond au processus instantané mais avec trois particules au lieu de deux.

Il suffit de considérer les distances entre les paires de particules consécutives ainsi que leurs vitesses. Nous notons x_i la distance entre la particule i et la particule $i+1$ (en adoptant la convention que la particule 4 est la particule 1) et $\sigma_i \in \{\pm 1\}$ la vitesse de la particule i . L'espace d'états du processus est $E = \Delta \times \Sigma$ où

$$\Delta = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1, x_2, x_3 \geq 0 \text{ et } x_1 + x_2 + x_3 = L\},$$

et $\Sigma = \{\pm 1\}^3$. Nous notons $x = (x_1, x_2, x_3) \in \Delta$ et $\sigma = (\sigma_1, \sigma_2, \sigma_3) \in \Sigma$ les vecteurs de distances et de vitesses respectivement.

Chaque σ_i alterne indépendamment entre $+1$ et -1 avec un taux ω . Lorsque toutes les particules sont à distance positive les unes des autres, chaque particule suit sa vitesse de sorte que $\partial_t x_i = \sigma_{i+1} - \sigma_i$. Lorsque deux particules sont en contact, elles se comportent comme dans le cas deux particules (voir la Figure 2.1). Enfin, les collisions simultanées entre trois particules sont résumées dans la Figure 2.10.

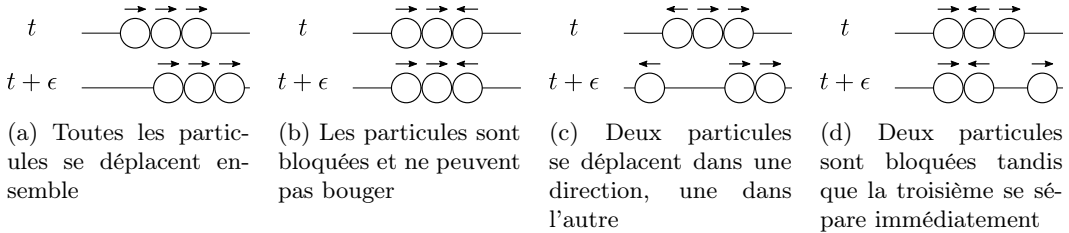


Figure 2.10: Liste complète des collisions entre trois particules

En notant $A := (L, 0, 0)$, $B := (0, L, 0)$, $C := (0, 0, L)$ les sommets du simplexe, et après avoir observé que x reste constant lorsque $\sigma = \pm(1, 1, 1)$, la dynamique déterministe sur le reste de l'espace d'états est résumé dans la Figure 2.11.

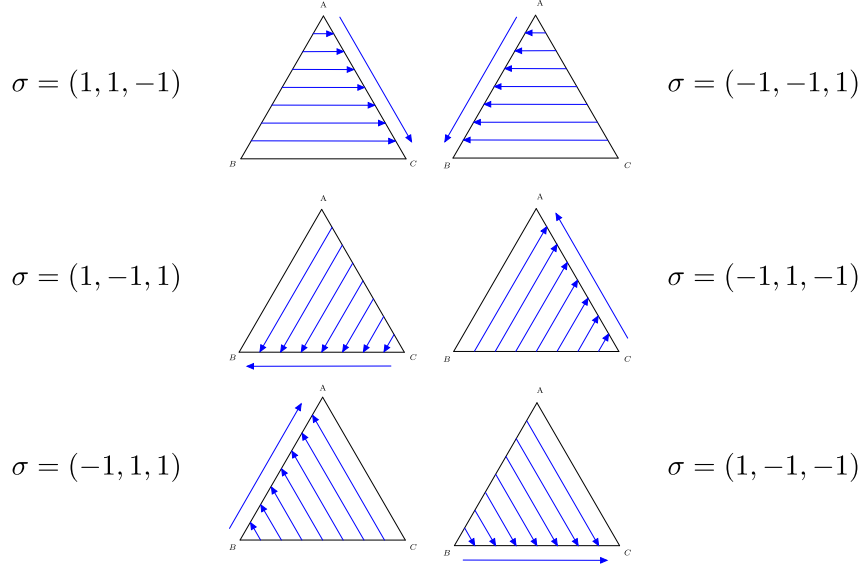


Figure 2.11: Dynamiques déterministes sur $\Delta \times \{\sigma\}$ pour $\sigma \neq \pm(1, 1, 1)$

Mesure invariante En remarquant que deux particules en contact se comportent comme une seule particule effective avec des taux de sauts modifiés, et en utilisant les résultats dans le cas deux particules, on obtient l'ansatz

$$\pi = \sum_{\sigma \in \Sigma} \left(\sum_{F \in \mathcal{F}} a_{\sigma, F} \lambda_F + \sum_{X \neq Y} b_{\sigma, XY} \epsilon_{XY} \right) \otimes \delta_{\sigma} \quad (2.13)$$

pour la mesure invariante. Ici $\mathcal{F} = \{A, B, C, AB, BC, CA, ABC\}$ est l'ensemble des faces non vides du polytope Δ , λ_F est la mesure de Lebesgue de dimension $\dim(F)$ sur la face $F \in \mathcal{F}$ et pour $X, Y \in \{A, B, C\}$

$$\epsilon_{XY}(f) = \int_0^{\|Y-X\|} f \left(X + u \frac{Y-X}{\|Y-X\|} \right) e^{\kappa u} du \neq \epsilon_{YX}(f)$$

est la mesure unidimensionnelle de densité exponentielle.

En insérant cet ansatz dans la caractérisation générateur de l'invariance, on confirme sa validité et on détermine les paramètres. Les deux ingrédients clés ici sont la caractérisation explicite du générateur étendu des PDMPs [Dav93] et l'utilisation des symétries du modèle. Notons que la question difficile des conditions aux bords de la mesure invariante est implicitement traitée par les conditions aux bords que $f \in D(\mathcal{L})$ doit satisfaire.

2.6.4 Lifts et hypocoercivité

Lifts L'analogie en temps continu du lift des chaînes de Markov discrètes [CLP99, DHN00] a été introduit par Andreas Eberle et Francis Lörler [EL24a] pour simplifier les résultats de convergence hypocoercive dans [AAMN19, LW22, CLW23].

Soit $p : \hat{E} \rightarrow E$ une surjection entre deux espaces polonais \hat{E} et E . Étant donné un processus \hat{X} avec $\hat{\mathcal{L}}$ comme générateur L^2 et mesure invariante $\hat{\pi}$ prenant ses valeurs dans le grand espace d'états \hat{E} , l'idée est de définir un nouveau processus X avec \mathcal{L} comme

générateur L^2 dans le petit espace d'état E . On voudrait que π la mesure invariante de X soit l'image par p de $\hat{\pi}$.

Definition 2.40 (Lift [EL24a]). *S'il existe un noyau C de \mathcal{L} tel que*

$$\{f \circ p : f \in C\} \subset D(\hat{\mathcal{L}})$$

et que la condition de premier ordre

$$\left\langle \hat{\mathcal{L}}(f \circ p), g \circ p \right\rangle_{L^2(\hat{\pi})} = 0$$

ainsi que la condition de second ordre

$$\langle \mathcal{L}f, g \rangle_{L^2(\pi)} = -\frac{1}{2} \left\langle \hat{\mathcal{L}}(f \circ p), \hat{\mathcal{L}}(g \circ p) \right\rangle_{L^2(\hat{\pi})}$$

sont satisfaites pour tout $f, g \in C$, alors \hat{X} est un lift de X . On dit aussi que X est le collapse de \hat{X} .

Par exemple, si $U : \mathbb{R}^d \rightarrow \mathbb{R}$ est un potentiel C^1 avec $Z := \int e^{-U} dx < +\infty$, Langevin cinétique

$$\hat{\mathcal{L}}f(x, v) = v \cdot \nabla_x f(x, v) - \nabla_x U(x) \cdot \nabla_v f(x, v) - \gamma v \cdot \nabla_v f(x, v) + \gamma \Delta_v f(x, v)$$

est un lift du processus de Langevin suramorti

$$\mathcal{L}f(x) = -\frac{1}{2} \nabla_x U(x) \cdot \nabla_x f(x) + \frac{1}{2} \Delta_x f(x).$$

Ici, nous avons $\hat{E} = \mathbb{R}^d \times \mathbb{R}^d$, $E = \mathbb{R}^d$ et p est la projection sur x . La mesure invariante $\pi = \frac{1}{Z} e^{-U} dx$ de X est l'image de la mesure invariante $\hat{\pi} = \pi \otimes \kappa$ de \hat{X} où κ est la gaussienne d -dimensionnelle standard. Puisque $\hat{\mathcal{L}}(f \circ p)(x, v) = v \cdot \nabla_x f(x)$, on vérifie que

$$\int \hat{\mathcal{L}}(f \circ p)(g \circ p) d\hat{\pi} = \int g(x) \nabla_x f(x) \cdot \left[\int v d\kappa(v) \right] d\pi(x) = 0$$

et

$$\begin{aligned} -\frac{1}{2} \int \hat{\mathcal{L}}(f \circ p) \hat{\mathcal{L}}(g \circ p) d\hat{\pi} &= -\frac{1}{2} \int \int (v \cdot \nabla_x f(x)) (v \cdot \nabla_x g(x)) d\kappa(v) d\pi(x) \\ &= -\frac{1}{2} \int \nabla_x f(x) \cdot \nabla_x g(x) d\pi(x) \\ &= \int \left(-\frac{1}{2} \nabla_x U(x) \cdot \nabla_x f(x) + \frac{1}{2} \Delta_x f(x) \right) g(x) d\pi(x). \end{aligned}$$

De nombreux processus d'échantillonnage, tels que le bouncy particle sampler [BCVD18], le zigzag [BFR19] et le forward event-chain Monte Carlo [MKK14], sont des lifts de diffusions.

Cadre hypocoercif abstrait En partant d'un collapse \mathcal{L} de trou spectral m , c'est-à-dire

$$\|f - \pi(f)\|_{L^2(\pi)}^2 \leq \frac{1}{m} \langle f, (-\mathcal{L})f \rangle_{L^2(\pi)}$$

et de son lift $\hat{\mathcal{L}}$, la première étape pour établir les résultats de convergence dans [EL24a] est de trouver une décomposition $\hat{\mathcal{L}} = \hat{\mathcal{L}}_d + \gamma \hat{\mathcal{L}}_v$ où $\gamma > 0$, $\hat{\mathcal{L}}_d$ est aussi un lift de \mathcal{L} qui est antisymétrique et $\hat{\mathcal{L}}_v$ est un générateur agissant uniquement sur la vitesse. Par exemple, pour Langevin cinétique

$$\hat{\mathcal{L}}f(x, v) = \underbrace{v \cdot \nabla_x f(x, v) - \nabla_x U(x) \cdot \nabla_v f(x, v)}_{=: \hat{\mathcal{L}}_d} \underbrace{- \gamma v \cdot \nabla_v f(x, v) + \gamma \Delta_v f(x, v)}_{=: \gamma \hat{\mathcal{L}}_v}$$

se décompose naturellement en la dynamique hamiltonienne $\hat{\mathcal{L}}_d$ et $\hat{\mathcal{L}}_v$, un OUP en vitesse. Pour le processus instantané, cette décomposition, et même le calcul du collapse, sont moins évidents. Il y a en effet deux obstacles pour adapter les résultats dans [AAMN19, LW22, CLW23, EL24a], à savoir (1) les bords de l'espace d'états et (2) la dynamique de la vitesse $\hat{\mathcal{L}}_v$ qui n'est ni le rafraîchissement complet de [LW22, EL24a] ni l'OUP de [CLW23].

On arrive quand même à calculer le collapse du processus instantané, qui s'avère être le brownien sticky, et on trouve quand même une décomposition appropriée $\hat{\mathcal{L}} = \hat{\mathcal{L}}_d + \gamma\hat{\mathcal{L}}_v$. On identifie également des hypothèses générales sur $\hat{\mathcal{L}}_d$ et $\hat{\mathcal{L}}_v$, qui rappellent les hypothèses DMS [DMS09, DMS15], et qui permettent d'établir un résultat de convergence hypocoercif unifié couvrant le processus instantané ainsi que Langevin cinétique et de nombreux processus d'échantillonnage. Le résultat intermédiaire clé est l'inégalité de Poincaré le long du flot de l'équation de Fokker-Planck [AAMN19, LW22, CLW23, EL24a].

Proposition 2.41 (Inégalité de Poincaré le long du flot). *Pour $f_0 \in L^2_0(\hat{\pi})$ et $f(t, x, v) = P_t f(x, v)$, nous avons*

$$\int_0^T \|f\|_{L^2(\hat{\pi})}^2 dt \leq (\gamma^2 C_0 + C_1) \int_0^T \left\langle f, (-\hat{\mathcal{L}}_v f) \right\rangle_{L^2(\hat{\pi})} dt$$

où C_0 et C_1 dépendent uniquement de T et du trou spectral du collapse m .

Suivant la structure de preuve générale dans [LW22, CLW23, EL24a], cette inégalité est établie à l'aide d'un lemme de divergence quantitatif. Notons que dans le cadre général, le lemme est énoncé en termes du collapse \mathcal{L} au lieu de la divergence standard.

Lemma 2.42 (Lemme de divergence). *Pour $f \in L^2_0(\lambda \otimes \mu)$, il existe $h, g \in L^2(\lambda \otimes \mu)$ tels que*

$$\begin{aligned} h(t, \cdot), g(t, \cdot) &\in \text{Dom}(\mathcal{L}) \text{ pour } \lambda\text{-presque partout } t, \\ h(\cdot, x) &\in H^1(0, T) \text{ pour } \mu\text{-presque partout } x, \\ g(\cdot, x) &\in H^2(0, T) \text{ pour } \mu\text{-presque partout } x, \end{aligned}$$

tels que

$$\partial_t h - 2\mathcal{L}g = f,$$

avec les conditions aux bords

$$\begin{aligned} h(t = 0, \cdot) &= h(t = T, \cdot) = 0, \\ g(t = 0, \cdot) &= g(t = T, \cdot) = 0, \end{aligned}$$

et

$$\begin{aligned} \|2\mathcal{L}g\|_{L^2(\lambda \otimes \pi)} &\leq c_1 \|f\|_{L^2(\lambda \otimes \pi)}, \\ \sqrt{2\mathcal{E}_T(h)} &\leq c_2 \|f\|_{L^2(\lambda \otimes \pi)}, \\ \sqrt{2\mathcal{E}_T(g)} &\leq c_3 \|f\|_{L^2(\lambda \otimes \pi)}, \\ \sqrt{2\mathcal{E}_T(\partial_t g)} &\leq c_4 \|f\|_{L^2(\lambda \otimes \pi)}, \end{aligned}$$

où $\lambda = 1_{[0, T]} dx$ et $\mathcal{E}_T(f) = \langle f, (-\mathcal{L})f \rangle_{L^2(\lambda \otimes \pi)}$.

Pour le processus instantané, cette approche hypocoercive donne un taux de décroissance exponentielle précis en accord avec le temps de mélange.

Proposition 2.43. *Il existe $C(\omega, L) > 0$ tel que le processus instantané satisfait*

$$\|P_t f\|_{L^2(\pi)} \leq C(\omega, L) \exp\left(-\frac{K\omega}{(\omega L)^2 + 1} t\right) \|f\|_{L^2(\pi)}$$

pour tout $f \in L^2(\pi)$ avec $\pi(f) = 0$, où $K > 0$ est une constante absolue.

Pour Langevin cinétique avec un potentiel convexe, on retrouve les bons taux de [CLW23], c'est-à-dire

$$\frac{\gamma m}{\gamma^2 + m}.$$

2.7 Perspectives

Les questions ouvertes suivantes pourront faire l'objet de futurs travaux.

Systèmes à n particules Le calcul de la mesure invariante de deux RTPs générales a permis de mieux comprendre la mesure invariante du processus instantané [HGM23]. Il est donc naturel d'explorer la mesure invariante de modèles à trois particules avec différents taux de sauts. Une compréhension approfondie de ces systèmes à trois particules pourrait donner un ansatz pour la mesure invariante à quatre particules, et potentiellement pour la mesure à n particules qui contribuerait à la compréhension quantitative de MIPS.

Systèmes en dimensions supérieures Le cadre PDMP permet la modélisation rigoureuse de particules actives de dimensions supérieures avec des collisions sticky, une application importante étant donné que de nombreux systèmes physiques existent en deux ou trois dimensions. Ces collisions correspondent à un glissement des particules les unes sur les autres au lieu de réflexions de type billard de l'équilibre. Cette alternative aux potentiels répulsifs de plus en plus piqués [AdPLVW19, AdPvW23] pourrait permettre le calcul ou l'approximation de la mesure invariante. Les questions difficiles concernant les conditions aux bords devraient à nouveau trouver leur réponse dans le domaine du générateur.

Propriétés qualitatives des mesures invariantes Lorsque les modèles deviennent trop complexes, les stratégies pour calculer explicitement la mesure invariante présentées dans l'introduction échouent. Cependant, chaque caractérisation de la mesure invariante contient potentiellement des indices sur ses propriétés qualitatives. Ceci est particulièrement vrai pour la régularité de la mesure invariante pour les sauts markoviens entre EDOs [BH12, BLBMZ15], qui reste un domaine de recherche actif.

Assouplissement des hypothèses pour la convergence hypocoercive Les résultats de convergence hypocoercive du dernier chapitre sont prouvés sous l'hypothèse que la vitesse possède des propriétés de mélange fortes. La question de savoir si les résultats peuvent être étendus à des processus avec un bruit d'intensité variable ou même un bruit qui s'annule sous une condition de contrôle géométrique [DHHM22] se pose.

Chapter 3

Jamming pair of general run-and-tumble particles: Exact results, symmetries and universality classes

This chapter is based on the preprint [HGM23] with Manon Michel and Arnaud Guillin.

3.1 Introduction

Active matter systems such as bacterial colonies [Sch93, WL00, Cat12, EWG15] or robot swarms [MNS⁺19, WPL⁺21] are characterized by the breaking at the microscopic scale of energy conservation and time reversibility [Ram10, MJR⁺13, OKTvW22]. Compared to their passive counterparts, these systems exhibit rich behaviors including motility-induced phase separation [CT15, CMP20], emergence of patterns [SNLK01] or collective motion [VZ12, BCD⁺13]. Their study remains challenging because many tools from equilibrium statistical mechanics do not apply. In particular there is no guarantee the steady state follows some Boltzmann distribution, even up to the definition of an effective potential. This has fueled interesting developments in both physics and mathematics, but which typically rely on a coarse-grained approach [TT98, MJR⁺13, TC08, FKB15, SFK17, LEB18]. While they successfully recover some out-of-equilibrium macroscopic behaviors, they leave open the question of the microscopic origin of a given phenomenon, a question strongly linked to the fundamental one of universality.

It motivates the derivation of an exact theory of active-particle systems directly from the microscopic scale of the individual particle, without any coarse-graining apart the stochastic description of individual particle motion. Recent works [KHEBT18, Eri21] deal with the derivation of large-scale limit of active lattice-gas models [KL98], typically aiming at deriving the steady state. While being exact, such hydrodynamic limit requires strong conditions (symmetric jumps, particular exclusion rule), in order to ensure a rigorous limit derivation. Furthermore, any results at a large but finite size requires a subtle computation of fluctuating terms, as done in Macroscopic Fluctuation Theory [Der07, BDSG⁺02]. Whereas such terms impacts the existence of metastable solutions, their derivation remains challenging [ARKL21].

Another approach focuses on the derivation of exact results for microscopic models [SEB16, SEB17], with a strong motivation to decipher the impact of the reversibility breaking. They consider discrete models of two Run-and-Tumble Particles (RTP) with a hardcore jamming interaction on a 1D-ring. The RTP is a paradigmatic active matter model mimicking the moving pattern of bacteria such as E. Coli [Sch93]. RTPs perform a series of straight

runs separated by stochastic refreshments of the run direction called tumbles. Even a single RTP displays interesting out-of-equilibrium features [BMR⁺20, LDMS19, MLDMS21]. In [SEB16, SEB17], the steady state distribution is exactly determined from a Master equation on a lattice by a generating function approach, respectively for the case of instantaneous and finite tumbles. The continuous limit is also discussed. By doing so, they identify the Dirac jamming contributions to the stationary distribution and how a new lengthscale, linked to an exponential decay or growth, appears in the case of the finite tumble. While these results are a first, questions remain open regarding the rigorous continuum limit of such discrete lattice models, the portability of such involved methods to more general systems and dimensions and the actual capacity of lattice models to capture the particular role that domain boundaries seem to play in such piecewise-ballistic systems. More importantly, the definition of some general universality classes ruling the steady state remains yet open.

In this work, we answer this precise question and show that a general periodic two-RTP system with jamming interactions is entirely described by two explicit universality classes that we name the *detailed-jamming* and *global-jamming* classes. They identify with the preservation or not of a detailed symmetry at the level of the dynamical internal states between probability flows entering and exiting jamming configurations. The key idea is to build a formalism of the RTP dynamics incorporating all the present symmetries (particle indistinguishability, space homogeneity and periodicity) and directly at the continuous-space and -time level thanks to Piecewise deterministic Markov processes (PDMP) [Dav84, Dav93]. It then gives a direct access to the impact of any dynamical skewness introduced by the jamming as a boundary constraint on the bulk dynamical relaxation ruled by the tumbles. This approach is inspired by developments in the sampling of equilibrium systems by non-reversible and continuous-time Markov processes, known as *Event-chain Monte Carlo* [BKW09, MKK14] and also formalized as PDMP [MGM23]. Such stochastic process comes down in particle systems to generating an artificial RTP dynamics. Time reversibility is broken but the equilibrium Boltzmann distribution is still left invariant by the preservation at the particle collisions, i.e. on the boundary of the set of valid configurations, of some key invariances, be it pairwise [MKK14], translational [HMKK17] or rotational [MDS20]. This motivated a similar investigation of which symmetry is preserved on the boundaries, but in an active system, and whether it is sufficient to fully determine the steady state.

Thus, we provide in Section 3.2 a full description of the two-RTP process as PDMP, including their generator (an infinitesimal description of the evolution). This is the basis to obtain in Section 3.3 the conservation constraints on the exact steady state distribution. A careful analysis leads in Section 3.4 to the explicit definition of the two universality classes and their respective steady-state expressions for general continuous models, which are consistent with the continuous-space limit of the particular discrete cases studied in [SEB16, SEB17]. Finally we discuss in Section 3.5 the light now shed on the interplay between reversibility breaking and jamming and the perspectives it brings.

3.2 Two-RTP system as PDMP

We consider a pair of interacting RTPs on a one-dimensional torus of length L . A single-particle possible internal state is described by the variable v , which typically codes for the velocity of the particle and commonly takes values in $\{-1, +1\}$ (two-state RTP or instantaneous tumble) or in $\{-1, 0, +1\}$ (three-state RTP or finite tumble). Aiming at universality class definition, we adopt a general setting and, therefore, we do not restrict the value of v , which can be continuous. Both particles change their internal state independently according to a Poisson jump process set by the transition rate $\omega(v) > 0$ and kernel q , which are general but homogeneous in space and yielding ergodicity in the internal-state space. The particles interact with each other through hardcore interactions leading to jammed states where the two RTPs are colliding against each other until a tumble allows the particles to escape, as illustrated in Fig. 3.1.

We now show how to formalize the system and all its symmetries in a continuous-

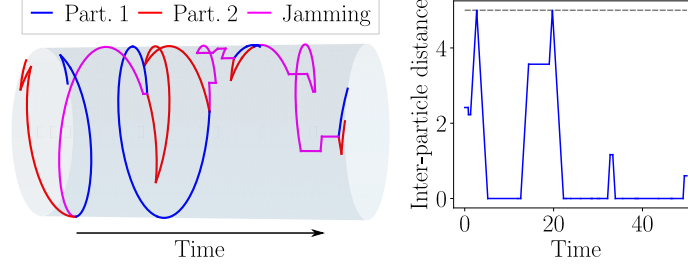


Figure 3.1: Trajectory of two RTPs undergoing instantaneous tumble (**left**), and of their corresponding periodic interdistance (**right**) ($\omega = 0.2$, $L = 10$).

time and continuous-space setting by using PDMP. First, we note $x_{t,1}, x_{t,2}$ the positions of the particles, and $v_{t,1}$ and $v_{t,2}$ the particle internal states. Thanks to the homogeneity, periodicity and particle indistinguishability, the system can entirely be described at time t by the periodic interdistance $0 \leq r_t \leq L/2$ between particles, defined as $\min(|x_{t,1} - x_{t,2}|, L - |x_{t,1} - x_{t,2}|)$. As shown in Fig. 3.1, the interdistance r_t then undergoes the following evolution, depending if it is located in the *bulk* (a), at the *periodic* (b) or *jamming* (c) boundaries:

(a - Run) While in the *bulk* ($0 < r_t < L/2$), the interdistance r_t is updated through a deterministic evolution in three possible regimes: increasing $((v_{t,1} - v_{t,2})(x_{t,1} - x_{t,2}) \geq 0$ if $|x_{t,1} - x_{t,2}| \leq L/2$), decreasing $((v_{t,1} - v_{t,2})(x_{t,1} - x_{t,2}) \leq 0$ if $|x_{t,1} - x_{t,2}| \leq L/2$) and stalling ($v_{t,1} = v_{t,2}$).

(a - Tumble) The transitions inside and between regimes stem from a particle tumble and are ruled by the superposed Poisson process of rate $\omega(v_1) + \omega(v_2)$. No direct transition from two states of the stalling regime can occur, as it would require a simultaneous tumble from both particles, which is of null probability.

(b) When reaching the *periodic* boundary ($r_t = L/2$) in an increasing regime, the interdistance automatically switches, by periodicity, from an increasing to a decreasing regime.

(c) Finally, and most interestingly, when the *jamming* boundary ($r_t = 0$) is reached in a decreasing regime, the interdistance is stalling in 0 until a tumble occurs allowing the particles to separate from each other and making r_t re-enter the *bulk* in an increasing regime.

We now characterize the stochastic dynamics of r_t by a PDMP [Dav84, Dav93], which is entirely encoded:

(a) by its infinitesimal generator in the bulk, $0 < r_t < L/2$,

$$\mathcal{A}_{\mathcal{B}} = \underbrace{\nu \phi(s) \partial_r}_{\text{Run phase}} + \underbrace{\tilde{\omega}(s)(Q((\nu, s), \cdot) - \text{Id})}_{\text{Tumble phase}}, \quad (3.1)$$

where we introduce the two auxiliary dynamical variables $(\nu_t, s_t) \in \mathcal{V}$ (so that the complete bulk set is $\mathcal{B} =]0, L/2[\times \mathcal{V}$): The variable $\nu_t \in \{-1, 0, 1\}$ sets the nature of the regime as defined in (a) and \mathcal{V} can be decomposed as $\mathcal{V}_{-1} \cup \mathcal{V}_0 \cup \mathcal{V}_1$. The variable $s_t = \{v_{t,1}, v_{t,2}\}$, encoding the particle indistinguishability, sets the amplitude of the evolution by $\phi(s_t) = |v_{t,1} - v_{t,2}|$ and the tumble by its rate $\tilde{\omega}(s_t) = \omega(v_{t,1}) + \omega(v_{t,2})$ and its transition Markov kernel $Q((\nu, s), \cdot)$,

$$Q((\nu, \{v, \tilde{v}\}), (\nu', \{v', \tilde{v}'\})) = q(v, v') \left(\frac{1}{2} 1_{\{0\}}(\nu) 1_{\{-1, 1\}}(\nu') \right. \\ \left. + \frac{\omega(v)}{\tilde{\omega}(\{v, \tilde{v}\})} 1_{\{-1, 1\}}(\nu) 1_{\{\text{sign}(\nu\nu')\}} \left(\text{sign}\left(\frac{v' - \tilde{v}}{v - \tilde{v}}\right) \right) \right),$$

all other transitions being of probability 0, as they would involve tumble of both particles. Thus, the particle indistinguishability leads to an isotropic symmetry for Q ,

$$Q((\nu, s), (\nu', s')) = Q((- \nu, s), (- \nu', s')), \quad (3.2)$$

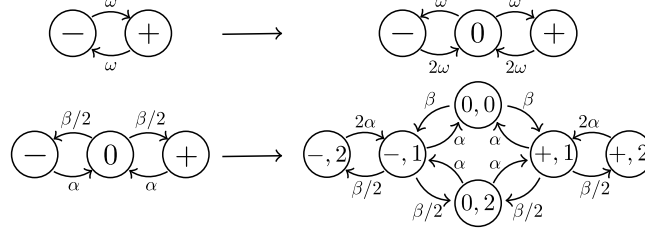


Figure 3.2: Transitions between the internal states at the single particle (**left**) and interdistance level (**right**) in the case of instantaneous (**top**, and finite tumble (**left**), in a reduced representation for s thanks to the single-particle tumble isotropy (instantaneous: s can be omitted, finite: s codes for the number of moving particles).

as illustrated on Fig. 3.2 for the instantaneous and finite tumbles. The symmetry (3.2) will have an important impact on the possible invariant distributions and justifies the introduction of the representation (ν, s) .

(b) by its behavior at the periodic boundary: at $(r, \nu, s) \in \{L/2\} \times \mathcal{V}_1 = \Gamma_P^*$, the periodic boundary kernel $Q_P((r, \nu, s), (r', \nu', s')) = 1_{\{(r, -\nu, s)\}}(r', \nu', s')$ codes for a switch from an increasing ($\nu = 1$) to a decreasing ($\nu' = -1$) regime.

(c) and finally by its behavior at the jamming boundary: for $(r, \nu, s) \in \{0\} \times \mathcal{V}_{-1} = \Gamma_J^*$, the jamming boundary kernel, $Q_J((r, \nu, s), (r^*, \nu^*, s^*)) = 1_{\{(r, \nu, s)\}}(r^*, \nu^*, s^*)$ makes the system jumps into a jammed configuration $(r^*, \nu^*, s^*) \in \mathcal{J} = \{0\} \times \mathcal{V}_{-1} \cup \mathcal{V}_0$, which is ruled by the following infinitesimal generator,

$$\mathcal{A}_{\mathcal{J}} = \tilde{\omega}(s^*)(Q((\nu^*, s^*), \cdot) - \text{Id}), \quad (3.3)$$

which translates the persistence of the jamming state until the tumbles lead to an unjamming state ($\nu = +1$), which, rigorously combined with an unjamming boundary kernel analogous to $Q_{\mathcal{J}}$, generates an increasing regime in the bulk.

3.3 Conservation constraints in the steady state and jamming boundary role

The steady-state distribution of the periodic interdistance r of two passive hardcore particles is the uniform density over $[0, L/2]$, with in particular a null probability to find r in the exact state 0. The addition of non-equilibrium features (activity, jamming) may introduce deviations from this equilibrium behavior, which we now study. The measure π left invariant by the considered PDMP should satisfy the following condition,

$$\int_{\mathcal{B}} \mathcal{A}_{\mathcal{B}} f(r, \nu, s) d\pi(r, \nu, s) + \int_{\mathcal{J}} \mathcal{A}_{\mathcal{J}} f(0, \nu, s) d\pi(0, \nu, s) = 0, \quad (3.4)$$

with f a suitably smooth test function. As the considered process is trivially irreducible thanks to the ergodicity of the tumbles, the distribution π is unique. Now, as f can be any test function, we obtain by integration by part and boundary conditions, as detailed in Appendix 3.6, the equivalent condition system,

$$\begin{cases} (\text{C}_{\mathcal{B}}) \int_{\mathcal{V}} \pi(r, d\nu' ds') \tilde{\omega}(s') Q((\nu', s'), (\nu, s)) \\ \quad = \tilde{\omega}(s) \pi(r, \nu, s) + \nu \phi(s) \partial_r \pi(r, \nu, s) \\ (\text{C}_{\mathcal{P}}) \pi(L/2, 1, s) = \pi(L/2, -1, s) \\ (\text{C}_{\mathcal{J}}) \int_{\mathcal{J}} d\pi(0, \nu', s') \tilde{\omega}(s') Q((\nu', s'), (\nu, s)) \\ \quad = 1_{\{-1, 0\}}(\nu) \tilde{\omega}(s) \pi(0, \nu, s) + \nu \phi(s) \pi(0^+, \nu, s) \end{cases} \quad (3.5)$$

This condition system completely constrains the invariant measure π . It can be understood as a set of conservation equations of probability flows in the bulk \mathcal{B} , and on the jamming and periodic boundaries. It leads to the following conclusions:

(i) (C_P) is a direct translation of the periodicity of the system and is not impacting the physical meaning of the form of π , in particular potential deviation from the passive equilibrium form.

(ii) In the bulk, (C_B) is setting a balance between the Run and Tumble contributions. Any product measure of the form $\frac{2}{L}1_{]0,L/2[}(r)\mu_{\mathcal{B}}(\cdot)$, with $\tilde{\omega}(\cdot)\mu_{\mathcal{B}}(\cdot)$ the unique invariant measure of the tumble kernel Q , is satisfying (C_B) with null Run term $\phi(\cdot)\partial_r\pi(r, \pm 1, \cdot)$. Furthermore, the isotropy of Q (3.2) leads to $\mu_{\mathcal{B}}(\nu, s) = \mu_{\mathcal{B}}(-\nu, s)$, satisfying (C_P) . Reciprocally, any distribution such that $\pi(r, \nu, s) = \pi(r, -\nu, s)$ is obeying (C_B) only if $\partial_r\pi = 0$ for all $r > 0$, i.e. only if $\tilde{\omega}(\cdot)\pi(\cdot)$ is invariant by Q and then identifies with $\tilde{\omega}(\cdot)\mu_{\mathcal{B}}(\cdot)$.

(iii) The jamming condition (C_J) creates Dirac components in $r = 0$ in π for finite ω , which are not present in the passive counterpart.

(iv) More interestingly, it appears clearly now that the jamming interaction can impact the bulk behavior by the term $\phi(s)\pi(0^+, \pm 1, s)$ in (C_J) . It can be understood as a boundary source term for the bulk equation set by (C_B) , which can force the system away from the uniform distribution in r in the bulk. This term comes from the integration by part, which stems from the Run term in \mathcal{A}_B . It is indeed the activity of the particles that pushes the exploration up to the boundary of the state space.

(v) In spite of the non-equilibrium features, we can derive some *active global* balance condition. Indeed, (C_J) yields the global balance of the probability flows in and out of jamming, $\int ds\phi(s)\pi(0^+, -1, s) = \int ds\phi(s)\pi(0^+, 1, s)$ and (C_B) imposes the global balance at any $r > 0$ $\int ds\phi(s)\partial_r\pi(r, -1, s) = \int ds\phi(s)\partial_r\pi(r, 1, s)$. Both combined, we obtain the following *active global* balance of the probability flows at any $r > 0$,

$$\int ds\phi(s)\pi(r, -1, s) = \int ds\phi(s)\pi(r, 1, s). \quad (3.6)$$

From (ii), any distribution π obeying the stricter *detailed* counterpart of (3.6), i.e., for any r, s ,

$$\pi(r, -1, s) = \pi(r, 1, s), \quad (3.7)$$

must identify with $\frac{2}{L}1_{]0,L/2[}(r)\mu_{\mathcal{B}}(\cdot)$ in the bulk. Actually, by symmetry argument, any distribution π obeying (3.7) in at least a single point r satisfies (3.7) in all r .

Thus, analyzing the conservation conditions in (3.5) unravels the potential impact of the non-equilibrium features on the invariant measure: First, the presence of Dirac terms in $r = 0$ for finite ω ; Second, the deviation from a product measure uniform in r in the bulk for unjamming scenarios which sets boundary conditions $\pi(0^+, \cdot)$ different from the tumble stationary distribution $\mu_{\mathcal{B}}(\cdot)$.

3.4 Universality classes and expression of the steady state

Therefore, we identify two universality classes, based on the deviation or not of $\pi(0^+, \pm 1, s)$ from $\mu_{\mathcal{B}}(\pm 1, s)$, which can always be explicitly tested in (C_J) . It is equivalent to the satisfaction or breaking of the detailed symmetry upon entering and exiting jamming for any s ,

$$\pi(0^+, 1, s) = \pi(0^+, -1, s). \quad (3.8)$$

Detailed-jamming class. The detailed symmetry (3.8) is satisfied. As a consequence, the global-flow balance (3.6) condition is obeyed through the detailed-flow one (3.7) and the non-equilibrium nature is only apparent through the presence of Dirac term at the jamming point in the invariant distribution $\pi_{\text{det}}(r, \nu, s)$, which is of the form,

$$w_{\mathcal{J}}1_{\{0\}}(r)\mu_{\mathcal{J}}(\nu, s) + w_{\mathcal{B}}\frac{2}{L}1_{]0,L/2[}(r)\mu_{\mathcal{B}}(\nu, s), \quad (3.9)$$

with $\mu_{\mathcal{J}}$ the invariant jamming distribution and $w_{\mathcal{J}}, w_{\mathcal{B}}$ the average time ratio respectively spent in jamming and the bulk, details on their derivation can be found in Appendix 3.7. Very close to an equilibrium counterpart, it follows in the bulk a product form, which, once marginalized over the dynamical variables, identifies with the passive uniform steady-state. In particular, this is the case of the instantaneous tumble [SEB16] but also of any systems with only two internal states $\pm v$.

Global-jamming class. The detailed symmetry (3.8) is not satisfied. This is for instance the case when the particles unjam at the smallest possible velocity when they jam together at different ones, implying some energy dissipation. This class then gathers systems further away from equilibrium, as we now explain. First, while the global-flow balance (3.6) still holds, there is no detailed-flow balance (3.7) at any point $r > 0$ but at the periodic boundary $r = L/2$. Sign of a stronger non-equilibrium nature, the stationary distribution $\pi_{\text{glob}}(r, \nu, s)$ cannot longer be put under the form of Dirac contributions in $r = 0$ and a product form in the bulk $r > 0$, and writes itself,

$$w_{\mathcal{J}} \mathbf{1}_{\{0\}}(r) \mu_{\mathcal{J}}(\nu, s) + w_{\mathcal{B}} \frac{2}{L} \mathbf{1}_{]0, L/2[}(r) \mu(r, \nu, s), \quad (3.10)$$

where $\mu(r, \nu, s) = w_{\text{eq}} \mu_{\mathcal{B}}(\nu, s) + w_{\text{rel}} \gamma(r, \nu, s)$ and $\gamma(\cdot)$ stands for the relaxation in the bulk from the non-detailed constraint on its jamming boundary to the detailed constraint on its periodic one. Supposing that the internal states are of discrete values and as detailed in Appendix 3.8.1, the bulk condition ($C_{\mathcal{B}}$) gives that $(\gamma(r, 0, s))_s$ is determined by the sum of $\gamma_{\pm}(r) = (\gamma(r, \pm 1, s))_s$, which obeys some relaxation equation,

$$\partial_r(\gamma_+(r), \gamma_-(r)) = B(\gamma_+(r), \gamma_-(r)), \quad (3.11)$$

where, by the symmetry (3.2), the bulk matrix B is of the form $B = \begin{pmatrix} B^+ & B^- \\ -B^- & -B^+ \end{pmatrix}$, leading to the uncoupled second-order system,

$$\begin{cases} \partial_r^2(\gamma_+ + \gamma_-) = (B^+ - B^-)(B^+ + B^-)(\gamma_+ + \gamma_-) \\ \partial_r^2(\gamma_+ - \gamma_-) = (B^+ + B^-)(B^+ - B^-)(\gamma_+ - \gamma_-) \end{cases} \quad (3.12)$$

The symmetry (3.2) also constrains the Jordan form of B , as it admits as eigenvalues symmetrical finite $(\pm \lambda_k)_k$ of respective eigenvectors $(\boldsymbol{\mu}_{k,1}^{\pm}, \boldsymbol{\mu}_{k,1}^{\mp})$ and necessarily 0 of eigenvector $(\boldsymbol{\mu}_{\mathcal{B}}, \boldsymbol{\mu}_{\mathcal{B}})$, see Appendix 3.8.2.

The general solution of (3.11), whose precise expression is in Appendix 3.8.3, derives from the spectral properties of B and decomposes over a sum of exponentials of rate $(\pm \lambda_k)_k$, modulated by polynomial terms stemming from possible degeneracies of the λ_k and the boundary conditions. As B admits 0 as an eigenvalue, the general solution also admits a pure polynomial term.

An interesting case is when B does not admit Jordan blocks of size bigger than 3, as, from (C_P), any polynomial decay disappears and we recover an exact symmetry between decaying and increasing exponentials and such for any jamming scenario, leading to,

$$\gamma^{\pm}(r) = \sum_k a_k \left[\cosh\left(\lambda_k \left(\frac{L}{2} - r\right)\right) (\boldsymbol{\mu}_{k,1}^+ + \boldsymbol{\mu}_{k,1}^-) \mp \sinh\left(\lambda_k \left(\frac{L}{2} - r\right)\right) (\boldsymbol{\mu}_{k,1}^+ - \boldsymbol{\mu}_{k,1}^-) \right], \quad (3.13)$$

with $(a_k)_k$ set by ($C_{\mathcal{J}}$). It yields a catenary-like relaxation once marginalized over (ν, s) , as hinted by (3.12). A particular case is any system with particle internal states symmetrical enough so that B only admits a single non-null $\pm \lambda_k$, so that $\lambda_k^2 \propto \text{Tr}((B^+ + B^-)(B^+ - B^-))$, as for the finite tumble (see Appendix 3.8.4), which simplifies the formula in [SEB17].

3.5 Discussion

The approach presented here gives a direct continuous-time and continuous-space setting, which encodes the fundamental system symmetries (periodicity, particle indistinguishability). By adequately capturing the jamming behavior as some boundary effect on the bulk,

we are able to define the two universality classes of the steady-state behavior for systems of two RTPs with homogeneous tumbles. Interestingly, they are entirely defined from the satisfaction or not of a detailed symmetry (3.8) at the jamming point, which translates in the bulk into the satisfaction of the necessary active global balance (3.6) in a stricter –and closer to equilibrium– detailed manner (3.7) or not. Then, the further away from the detailed symmetry (3.8) the jamming scenario constrains the system, the more involved structure and relaxation behavior set by (3.11) are eventually exhibited in the bulk, see Appendix 3.8.3. Indeed, the system is analogous to a catenary problem, where the active global balance acts as the fixed-length constraint on a cable or equivalently the RTP activity acts as some tension, while the tumble process acts as a weight force.

Thus, the impact of the time-reversibility breaking by the particle activity simply translates itself into the capacity of the system to explore its boundaries. There, the now *seen* jamming interactions may further impact the steady state, depending on some symmetry preservation. It is indeed reminiscent of the situation in non-reversible sampling by ECMC, where particle collisions can also be understood as infinite-rate jamming preserving some key symmetries.

The presented conclusions straightforwardly generalize into more general jamming scenario \mathcal{A}_J , set by some tumbling rates and kernel different from the bulk, as it only impacts the exact form of $w_{\mathcal{J}}$ and $\mu_{\mathcal{J}}(\cdot)$. Some future work lies in the study of the impact of non-homogeneity, with the tumbling depending on the RTP positions, as in presence of an external potential or soft interactions. The question of the relaxation of the system to its steady state should also find in the PDMP formalism a powerful approach that should be explored. More importantly, a compelling research perspective lies in the use of PDMPs and their faculty to isolate the impacts of boundary effects on the bulk to model more than two RTPs with hardcore jamming interactions. An explicit formula or a good quantitative understanding of the invariant measure would provide a new perspective on clustering phenomena such as MIPS, with a direct understanding of the impact of the microscopic details.

3.6 Derivation of the system of conservation conditions

Piecewise deterministic Markov processes (PDMP) have been introduced in [Dav84, Dav93]. A PDMP refers to a Markov process which evolves ballistically according to a deterministic differential flow, until a jump following some Poisson process occurs or the process reaches the domain boundary. We consider the PDMP defined by the infinitesimal generators \mathcal{A}_B and \mathcal{A}_J and by the periodic and jamming boundaries, as defined in (a), (b), (c).

The generator is an efficient tool to study the invariance of a measure π (e.g. [Dav93, Prop. 34.7]), as done in [MGM23] for PDMP generated in particle systems by ECMC methods. The measure π is indeed left invariant by the considered PDMP if the following condition is satisfied,

$$\int_{\mathcal{B}} \mathcal{A}_B f(r, \nu, s) d\pi(r, \nu, s) + \int_{\mathcal{J}} \mathcal{A}_J f(0, \nu, s) d\pi(0, \nu, s) = 0, \quad (3.14)$$

with f a suitably smooth test function (here we do not detail the description of the extended domain of the generator and a suitable core, as required by [Dav93, Th. 5.5,]). Furthermore, stemming from the periodic boundary kernel Q_P , f shall verify the condition on the periodic boundary,

$$f(L/2, \nu, s) = f(L/2, -\nu, s). \quad (3.15)$$

Similar conditions at the jamming boundary are required, i.e.

$$f(0^+, \nu, s) = f(0, \nu, s), \quad (3.16)$$

where we omit out of simplicity the superscript $*$ characterizing the jamming state. Now, expliciting the bulk term in 3.14,

$$\begin{aligned} & \int_{\mathcal{V}_{-1} \cup \mathcal{V}_0} \mathcal{A}_J f(0, \nu, s) \pi(0, d\nu, ds) + \int_{]0, L/2[\times \mathcal{V}} d\pi(r, \nu, s) \nu \phi(s) \partial_r f(r, \nu, s) \\ & + \int_{]0, L/2[\times \mathcal{V}} d\pi(r, \nu, s) \tilde{\omega}(s) \int_{\mathcal{V}} Q((\nu, s), (d\nu' ds')) (f(r, \nu', s') - f(r, \nu, s)) = 0, \end{aligned} \quad (3.17)$$

we obtain by integration by parts in the second term,

$$\begin{aligned} & \int_{\mathcal{V}_{-1} \cup \mathcal{V}_0} \mathcal{A}_J f(0, \nu, s) \pi(0, d\nu, ds) - \int_{\mathcal{V}_{-1} \cup \mathcal{V}_1} \nu \phi(s) \pi(0^+, d\nu, ds) f(0, \nu, s) \\ & + \int_{\mathcal{V}_{-1} \cup \mathcal{V}_1} \nu \phi(s) \pi(L/2^-, d\nu, ds) f(L/2, \nu, s) \\ & + \int_{]0, L/2[\times \mathcal{V}} dr d\nu ds f(r, \nu, s) \left[\int_{\mathcal{V}} \pi(r, d\nu', ds') \tilde{\omega}(s') Q((\nu', s'), (\nu, s)) \right. \\ & \quad \left. - \pi(r, \nu, s) \tilde{\omega}(s) - \nu \phi(s) \partial_r \pi(r, \nu, s) \right] = 0. \end{aligned}$$

As, $f(L/2, \nu, s) = f(L/2, -\nu, s)$ by the condition on the periodic boundary, $\pi(L/2, -\nu, s) = \pi(L/2, \nu, s)$ by definition of the periodic interdistance, we have the cancellation of the term in $L/2$ by periodicity,

$$\int_{\mathcal{V}_1} \nu \phi(s) \pi(L/2, d\nu, ds) f(L/2, \nu, s) + \int_{\mathcal{V}_{-1}} -(-\nu) \phi(s) \pi(L/2, -d\nu, ds) f(L/2, -\nu, s) = 0. \quad (3.18)$$

Thus we have, expliciting the jamming generator,

$$\begin{aligned} & \int_{\mathcal{V}} d\nu ds f(0, \nu, s) \left[\int_{\mathcal{V}_{-1} \cup \mathcal{V}_0} \pi(0, d\nu', ds') \tilde{\omega}(s') Q((\nu', s'), (\nu, s)) \right. \\ & \quad \left. - 1_{\{-1, 0\}}(\nu) \pi(0, \nu, s) \tilde{\omega}(s) - \nu \phi(s) \pi(0^+, \nu, s) \right] \\ & + \int_{]0, L/2[\times \mathcal{V}} dr d\nu ds f(r, \nu, s) \left[\int_{\mathcal{V}} \pi(r, d\nu', ds') \tilde{\omega}(s') Q((\nu', s'), (\nu, s)) \right. \\ & \quad \left. - \pi(r, \nu, s) \tilde{\omega}(s) - \nu \phi(s) \partial_r \pi(r, \nu, s) \right] = 0. \end{aligned}$$

As f can be any test function and recalling the periodic condition, we obtain the following equivalent condition system,

$$\left\{ \begin{array}{l} (C_B) \quad \int_{\mathcal{V}} \pi(r, d\nu, ds') \tilde{\omega}(s') Q((\nu', s'), (\nu, s)) \\ \quad = \pi(r, \nu, s) \tilde{\omega}(s) + \nu \phi(s) \partial_r \pi(r, \nu, s) \text{ for any } (r, \nu, s) \in \mathcal{B}, \\ (C_P) \quad \pi(L/2, 1, s) = \pi(L/2, -1, s) \text{ for any } (\pm 1, s) \in \mathcal{V}_{\pm 1}, \\ (C_J) \quad \int_{\mathcal{V}_{-1} \cup \mathcal{V}_0} \pi(0, d\nu' ds') \tilde{\omega}(s') Q((\nu', s'), (\nu, s)) \\ \quad = 1_{\{-1, 0\}}(\nu) \pi(0, \nu, s) \tilde{\omega}(s) + \nu \phi(s) \pi(0^+, \nu, s) \text{ for any } (\nu, s) \in \mathcal{V}, \end{array} \right. \quad (3.19)$$

which identifies with the system (3.5), whose consequences are further discussed in Section 3.3.

3.7 Derivation of the detailed-jamming steady state

We now determine in more details the steady-state distribution in the detailed-jamming class as introduced in (3.9) in Section 3.4 and as recalled here,

$$\pi_{\text{det}}(r, \nu, s) = w_{\mathcal{J}} 1_{\{0\}}(r) \mu_{\mathcal{J}}(\nu, s) + w_{\mathcal{B}} \frac{2}{L} 1_{]0, L/2[}(r) \mu_{\mathcal{B}}(\nu, s).$$

Plugging the detailed-jamming symmetry (3.8) into (C_J) (3.5), we obtain,

$$\int_{\mathcal{V}_{-1} \cup \mathcal{V}_0} d\mu_{\mathcal{J}}(\nu', s') \tilde{\omega}(s') Q((\nu', s'), (1, s)) = \phi(s) \pi(0^+, 1, s) = -(-1) \phi(s) \pi(0^+, -1, s),$$

and derive an explicit condition on the jamming equivalent to (3.8),

$$(C_{\text{det}}) \begin{cases} \int_{\mathcal{V}_{-1} \cup \mathcal{V}_0} d\mu_{\mathcal{J}}(\nu', s') \tilde{\omega}(s') Q((\nu', s'), (0, s)) = \tilde{\omega}(s) \mu_{\mathcal{J}}(0, s) \\ \int_{\mathcal{V}_{-1} \cup \mathcal{V}_0} d\mu_{\mathcal{J}}(\nu', s') \tilde{\omega}(s') (Q((\nu', s'), (-1, s)) + Q((\nu', s'), (1, s))) = \tilde{\omega}(s) \mu_{\mathcal{J}}(-1, s) \\ \int_{\mathcal{V}_{-1} \cup \mathcal{V}_0} w_{\mathcal{J}} d\mu_{\mathcal{J}}(\nu', s') \tilde{\omega}(s') Q((\nu', s'), (1, s)) = \phi(s) w_{\mathcal{B}} \frac{2}{L} \mu_{\mathcal{B}}(1, s) \text{ for any } (1, s) \in \mathcal{V}_1 \end{cases} \quad (3.20)$$

so that the jamming invariant measure $\mu_{\mathcal{J}}$ is completely characterized by the first two lines and only depends on the jamming scenario. Then, the ratio $w_{\mathcal{J}}/w_{\mathcal{B}}$ of time spent respectively while jamming and in the bulk is set by the last line. The ratio $w_{\mathcal{J}}/w_{\mathcal{B}}$ does not depend on s as the jamming exit probability in state $(1, s)$ (left term of the last line) is indeed proportional to the invariant bulk probability $\mu_{\mathcal{B}}(\pm 1, s)$ of this state, which is known from (C_B), as we recall,

$$\int_{\mathcal{V}} \mu_{\mathcal{B}}(d\nu' ds') \tilde{\omega}(s') Q((\nu', s'), d\nu ds) = \mu_{\mathcal{B}}(d\nu, ds) \tilde{\omega}(s).$$

An interesting particular case is systems with two internal states v_1, v_2 , which includes the case of the instantaneous tumble ($v_1, v_2 = -1, +1$, $\omega(v_1) = \omega(v_2)$), as illustrated in Figure 3.2. These systems are always detailed-jamming ones, as there is only one entry state into jamming and one exit state out of jamming, realizing the detailed-jamming symmetry (3.8). For these systems, we obtain,

$$(C_{\text{det}}) \begin{cases} \mu_{\mathcal{J}}(-1, \{v_1, v_2\}) \omega(v_1) = 2\omega(v_2) \mu_{\mathcal{J}}(0, \{v_2, v_2\}) \\ \mu_{\mathcal{J}}(-1, \{v_1, v_2\}) \omega(v_2) = 2\omega(v_1) \mu_{\mathcal{J}}(0, \{v_1, v_1\}) \\ w_{\mathcal{J}} \omega(v_1) \mu_{\mathcal{J}}(0, \{v_1, v_1\}) + w_{\mathcal{J}} \omega(v_2) \mu_{\mathcal{J}}(0, \{v_2, v_2\}) = w_{\mathcal{B}} \frac{2|v_1 - v_2|}{L} \mu_{\mathcal{B}}(1, \{v_1, v_2\}) \\ \iff \frac{1}{2} w_{\mathcal{J}} (\omega(v_1) + \omega(v_2)) \mu_{\mathcal{J}}(-1, \{v_1, v_2\}) = w_{\mathcal{B}} \frac{2|v_1 - v_2|}{L} \mu_{\mathcal{B}}(1, \{v_1, v_2\}) \end{cases} \quad (3.21)$$

and,

$$\begin{cases} \omega(v_2) \mu_{\mathcal{B}}(\pm 1, \{v_1, v_2\}) = \omega(v_1) \mu_{\mathcal{B}}(0, \{v_1, v_1\}) \\ \omega(v_1) \mu_{\mathcal{B}}(\pm 1, \{v_1, v_2\}) = \omega(v_2) \mu_{\mathcal{B}}(0, \{v_2, v_2\}) \end{cases}.$$

We then get,

$$\begin{cases} \mu_{\mathcal{J}}(0, \{v_1, v_1\}) = \frac{\omega(v_2)^2}{\mathcal{N}_{\mathcal{J}}} \\ \mu_{\mathcal{J}}(0, \{v_2, v_2\}) = \frac{\omega(v_1)^2}{\mathcal{N}_{\mathcal{J}}} \\ \mu_{\mathcal{J}}(-1, \{v_1, v_2\}) = 2 \frac{\omega(v_1) \omega(v_2)}{\mathcal{N}_{\mathcal{J}}} \\ \mathcal{N}_{\mathcal{J}} = (\omega(v_1) + \omega(v_2))^2 \end{cases}, \quad \begin{cases} \mu_{\mathcal{B}}(0, \{v_1, v_1\}) = \frac{\omega(v_2)^2}{\mathcal{N}_{\mathcal{B}}} \\ \mu_{\mathcal{B}}(0, \{v_2, v_2\}) = \frac{\omega(v_1)^2}{\mathcal{N}_{\mathcal{B}}} \\ \mu_{\mathcal{B}}(\pm 1, \{v_1, v_2\}) = \frac{\omega(v_1) \omega(v_2)}{\mathcal{N}_{\mathcal{B}}} \\ \mathcal{N}_{\mathcal{B}} = (\omega(v_1) + \omega(v_2))^2 \end{cases}, \quad (3.22)$$

$$\text{and } \begin{cases} w_{\mathcal{B}} = \frac{1}{1 + \frac{2|v_1 - v_2|}{L(\omega(v_1) + \omega(v_2))}} \\ w_{\mathcal{J}}/w_{\mathcal{B}} = \frac{2|v_1 - v_2|}{L(\omega(v_1) + \omega(v_2))} \end{cases}.$$

where the ratio $w_{\mathcal{J}}/w_{\mathcal{B}}$ of time spent in jamming versus in the bulk increases with the strength of the ballistic nature of the considered RTP dynamics (decrease of the time needed to cover the length L of the torus, decrease of the total tumble rate). Computations for systems with more than two internal states are similar. We finally also stress that the conditions determined here also applied for a continuous distribution of internal states.

3.8 Derivation of the global-jamming steady state

We now determine in more details for discrete distributions of internal states the steady-state distribution in the global-jamming class as introduced in (3.10) in Section 3.4 and as recalled here,

$$\pi_{\text{glob}}(r, \nu, s) = w_{\mathcal{J}} 1_{\{0\}}(r) \mu_{\mathcal{J}}(\nu, s) + w_{\mathcal{B}} \frac{2}{L} 1_{]0, L/2[}(r) (w_{\text{eq}} \mu_{\mathcal{B}}(\nu, s) + w_{\text{rel}} \gamma(r, \nu, s)).$$

We do so by first deriving in Section 3.8.1 the bulk relaxation equation, which is ruled by the bulk relaxation matrix B . We then study in Section 3.8.2 the spectral properties of B , which are strongly constrained by the symmetry (3.2) of Q . We then obtain a general explicit expression for $\gamma(\cdot)$ in Section 3.8.3, which we completely determine in particular cases including the finite tumble in Section 3.8.4.

3.8.1 Bulk evolution equation

We obtain, by plugging π_{glob} into the bulk condition ($C_{\mathcal{B}}$) and exploiting the symmetry (3.2) of Q ,

$$\left\{ \begin{array}{l} \gamma(r, 0, s) \tilde{\omega}(s) = \int ds' 1_{\mathcal{V}_1}((1, s')) (\gamma(r, -1, s') + \gamma(r, 1, s')) \tilde{\omega}(s') Q((1, s'), (0, s)) \\ \pm \partial_r \gamma(r, \pm 1, s) = -\phi(s)^{-1} \gamma(r, \pm 1, s) \tilde{\omega}(s) \\ \quad + \phi(s)^{-1} \int ds' 1_{\mathcal{V}_1}((1, s')) \left[\gamma(r, 1, s') Q((1, s'), (\pm 1, s)) + \gamma(r, -1, s') Q((-1, s'), (\pm 1, s)) \right. \\ \quad \left. + (\gamma(r, 1, s') + \gamma(r, -1, s')) \int ds'' 1_{\mathcal{V}_0}((0, s'')) Q((1, s'), (0, s'')) Q((0, s''), (1, s)) \right] \tilde{\omega}(s') \end{array} \right. \quad (3.23)$$

We note,

$$\gamma(r) = \begin{pmatrix} \gamma^+(r) = (\gamma(1, s))_{(1, s) \in \mathcal{V}_1} \\ \gamma^-(r) = (\gamma(-1, s))_{(-1, s) \in \mathcal{V}_{-1}} \end{pmatrix}.$$

and summarize (3.23) into,

$$\partial_r \gamma(r) = B \gamma \quad (3.24)$$

with B the bulk relaxation matrix,

$$B = \begin{pmatrix} B^+ & B^- \\ -B^- & -B^+ \end{pmatrix} \in M_{2|\mathcal{V}_1| \times 2|\mathcal{V}_1|}(\mathbb{R}) \quad (3.25)$$

and,

$$B^+ = D_{\phi}^{-1} (Q^+ - I) D_{\omega}, \quad B^- = D_{\phi}^{-1} Q^- D_{\omega} \in M_{|\mathcal{V}_1| \times |\mathcal{V}_1|}(\mathbb{R}) \quad (3.26)$$

$$D_{\phi, ss'} = \phi(s) \delta_{ss'} \quad \text{and} \quad D_{\omega, ss'} = \omega(s) \delta_{ss'}, \quad (3.27)$$

and,

$$Q_{ss'}^{\pm} = Q((\pm 1, s'), (1, ds)) + \sum_{s''} Q((\pm 1, s'), (0, ds'')) Q((0, s''), (1, ds)). \quad (3.28)$$

As Q is a Markov kernel and as there is no tumble between two states with $\nu = 0$ since it will require a tumble of each particle at the same time, we have for any $(1, s') \in \mathcal{V}_1$ the conservation identity,

$$\sum_s (Q_{ss'}^+ + Q_{ss'}^- - \delta_{ss'}) = 0 \quad (3.29)$$

i.e.,

$$\sum_s \phi(s) (B^+ + B^-)_{ss'} = 0, \quad (3.30)$$

which is consistent with the active global balance condition (3.6), as clearly appearing when rewritten as,

$$\sum_s \phi(s) (\partial_r \gamma_s^+(r) - \partial_r \gamma_s^-(r)) = 0 = \sum_s \phi(s) (B^+ + B^-)_{ss'} (\gamma_{s'}^+(r) + \gamma_{s'}^-(r)). \quad (3.31)$$

3.8.2 Spectral analysis of bulk matrix B

Jordan reduction of B and symmetry constraints

We now consider the Jordan reduction of $B = P^{-1}JP$, with J a Jordan matrix. For the Jordan block k of size p_k , we note the corresponding eigenvalue λ_k and the Jordan chain of linearly-independent generalized eigenvectors $\boldsymbol{\mu}_{k,n} = (\boldsymbol{\mu}_{k,n}^+, \boldsymbol{\mu}_{k,n}^-)$, with $1 \leq n \leq p_k$, $\boldsymbol{\mu}_{k,1}$ being an eigenvector and $B\boldsymbol{\mu}_{k,n} = \lambda_k\boldsymbol{\mu}_{k,n} + \boldsymbol{\mu}_{k,n-1}$ for $n > 1$.

By the symmetry of B , any block k of eigenvalue $\lambda_k > 0$ is mirrored by a block k^* of eigenvalue $\lambda_{k^*} = -\lambda_k$ and generalized eigenvectors $\boldsymbol{\mu}_{k^*,n} = (-1)^{n+1}(\boldsymbol{\mu}_{k,n}^-, \boldsymbol{\mu}_{k,n}^+)$, as $B(\boldsymbol{\mu}_{k,1}^-, \boldsymbol{\mu}_{k,1}^+) = -\lambda_k(\boldsymbol{\mu}_{k,1}^-, \boldsymbol{\mu}_{k,1}^+)$ and, for $n > 1$,

$$(-1)^{n+1}B(\boldsymbol{\mu}_{k,n}^-, \boldsymbol{\mu}_{k,n}^+) = -\lambda_k(-1)^{n+1}(\boldsymbol{\mu}_{k,n}^-, \boldsymbol{\mu}_{k,n}^+) + (-1)^n(\boldsymbol{\mu}_{k,n-1}^-, \boldsymbol{\mu}_{k,n-1}^+).$$

Furthermore, as $\hat{D}_\phi^\pm B \hat{D}_\omega^{-1}$, with $\hat{D}_\phi^\pm = \begin{pmatrix} D_\phi & 0 \\ 0 & -D_\phi \end{pmatrix}$ and $\hat{D}_\omega^{-1} = \begin{pmatrix} D_\omega^{-1} & 0 \\ 0 & D_\omega^{-1} \end{pmatrix}$, identifies with the tumble Markov kernel collapsed in $\nu = 0$ of the considered process, it is diagonalizable with real eigenvalues ranging from 0 to -1 and a unique eigenvector $\boldsymbol{\mu}$ so that $\hat{D}_\phi^\pm B \hat{D}_\omega^{-1} \boldsymbol{\mu} = 0$. The vector $\boldsymbol{\mu}$ identifies with the unique stationary bulk solution as $\boldsymbol{\mu} = \hat{D}_\omega(\boldsymbol{\mu}_B, \boldsymbol{\mu}_B)$. As \hat{D}_ϕ is invertible, $(\boldsymbol{\mu}_B, \boldsymbol{\mu}_B)$ is the unique eigenvector of eigenvalue 0 of B . Then, for the unique block k with $\lambda_k = 0$, the generalized eigenvectors are of the following form $\boldsymbol{\mu}_{k,n} = (\boldsymbol{\mu}_{k,n}^+, (-1)^{n+1}\boldsymbol{\mu}_{k,n}^+)$, as,

$$\begin{cases} \boldsymbol{\mu}_{k,1} = (\boldsymbol{\mu}_B, \boldsymbol{\mu}_B) \\ B\boldsymbol{\mu}_{k,n} = \boldsymbol{\mu}_{k,n-1}, \quad n > 1 \end{cases} \quad \text{gives} \quad \begin{cases} (B^+ + (-1)^{n+1}B^-)\boldsymbol{\mu}_{k,n}^+ = \boldsymbol{\mu}_{k,n-1}^+ \\ -((-1)^{n+1}B^+ + B^-)\boldsymbol{\mu}_{k,n}^+ = (-1)^n\boldsymbol{\mu}_{k,n-1}^+ \end{cases}$$

It is finally noteworthy that $(\boldsymbol{\mu}_B, \boldsymbol{\mu}_B)$ is the only symmetrical eigenvector of B , as any symmetrical eigenvector is of eigenvalue 0, as can be seen directly from B or from an equivalent argument of linear independence for $\lambda_k \neq 0$ of the eigenvectors $(\boldsymbol{\mu}_{k,n}^+, \boldsymbol{\mu}_{k,n}^-)$ (e.v. λ_k) and $(\boldsymbol{\mu}_{k,n}^-, \boldsymbol{\mu}_{k,n}^+)$ (e.v. $-\lambda_k$). This is consistent with the fact that a symmetrical eigenvector is by definition a stationary solution of the detailed-jamming class which is unique and of eigenvalue 0. Hence the complete characterization only by the detailed-symmetry condition ($\pi(r, \nu, s) = -\pi(r, -\nu, s)$) of the detailed-jamming class, which is equivalent to $\pi(r, \nu, s) = \mu_Q(\nu, s)$.

Thus, the number of Jordan blocks of J is odd and is noted $2d + 1$. We rank the blocks from $-d$ to d in a symmetrical manner: The k -th block of eigenvalue λ_k and generalized eigenvectors $(\boldsymbol{\mu}_{k,n}^+, \boldsymbol{\mu}_{k,n}^-)$ is mirrored by the $(-k)$ -th block of eigenvalue $-\lambda_k$, generalized eigenvectors $((-1)^{n+1}\boldsymbol{\mu}_{k,n}^-, (-1)^{n+1}\boldsymbol{\mu}_{k,n}^+)$. We then set the position of the block linked to the eigenvector $(\boldsymbol{\mu}_{0,1}, \boldsymbol{\mu}_{0,1}) = (\boldsymbol{\mu}_B, \boldsymbol{\mu}_B)$ to the 0-th rank. It is composed of generalized eigenvectors of the form $(\boldsymbol{\mu}_{0,n}, (-1)^{n+1}\boldsymbol{\mu}_{0,n})$.

It is noteworthy that, as B is of even size, the Jordan block associated to $\lambda_0 = 0$ must be of an even size $p_0 > 0$, so that B is never diagonalizable. This is a direct consequence of the symmetry imposed on Q by the particle indistinguishability and space homogeneity.

Spectral relationship with submatrices B_+ and B_- and catenary equation

The family $((\boldsymbol{\mu}_{k,n}^+, \boldsymbol{\mu}_{k,n}^-)_{k,n}, (\boldsymbol{\mu}_{0,n}, (-1)^{n+1}\boldsymbol{\mu}_{0,n})_n, (\boldsymbol{\mu}_{k,n}^-, \boldsymbol{\mu}_{k,n}^+)_{k,n})$ naturally forms a basis of $\mathbb{R}^{2|\mathcal{V}_1|}$. Furthermore, the vectors $(\boldsymbol{\mu}_{k,n}^+ + \boldsymbol{\mu}_{k,n}^-)_{k,n}$ (resp. $(\boldsymbol{\mu}_{k,n}^+ - \boldsymbol{\mu}_{k,n}^-)_{k,n}$) are linearly independent. Indeed, if it exists $\{b_{k,n}\}$ so that $\sum_{k,n} b_{k,n}\boldsymbol{\mu}_{k,n}^+ = -\sum_{k,n} b_{k,n}\boldsymbol{\mu}_{k,n}^-$ (resp. $= \sum_{k,n} b_{k,n}\boldsymbol{\mu}_{k,n}^-$), then we have $\sum_{k,n} b_{k,n}\boldsymbol{\mu}_{k,n} = -\sum_{k,n} b_{k,n}(-1)^{n+1}\boldsymbol{\mu}_{-k,n}$ (resp. $= \sum_{k,n} b_{k,n}\boldsymbol{\mu}_{k,n}$), which is not possible. Eventually, by considering the size of the sets $(\boldsymbol{\mu}_{k,n}^+ \pm \boldsymbol{\mu}_{k,n}^-)$, they form two basis of $\mathbb{R}^{|\mathcal{V}_1|}$.

In addition, we have the relationships, noting $M^+ = (B^+ + B^-)$ and $M^- = (B^+ - B^-)$,
 For $k > 0$,
$$\begin{cases} M^+(\boldsymbol{\mu}_{k,1}^+ + \boldsymbol{\mu}_{k,1}^-) = \lambda_k(\boldsymbol{\mu}_{k,1}^+ - \boldsymbol{\mu}_{k,1}^-) \\ M^-(\boldsymbol{\mu}_{k,1}^+ - \boldsymbol{\mu}_{k,1}^-) = \lambda_k(\boldsymbol{\mu}_{k,1}^+ + \boldsymbol{\mu}_{k,1}^-) \end{cases} \quad \text{and} \quad \begin{cases} M^+\boldsymbol{\mu}_{0,1} = 0 \\ M^-\boldsymbol{\mu}_{0,2} = \boldsymbol{\mu}_{0,1} \end{cases}, \quad (3.32)$$

so that, the vectors $(\boldsymbol{\mu}_{k,1}^+ \pm \boldsymbol{\mu}_{k,1}^-)_{k>1}$ are respectively the eigenvectors of $M^\mp M^\pm$ of eigenvalues λ_k^2 and $\boldsymbol{\mu}_{0,1}$ (resp. $\boldsymbol{\mu}_{0,2}$) is the eigenvector of M^-M^+ (resp. M^+M^-) of eigenvalue 0.

And furthermore, for $n > 1$, we have, for $k > 0$

$$\begin{cases} M^+(\boldsymbol{\mu}_{k,n}^+ + \boldsymbol{\mu}_{k,n}^-) = \lambda_k(\boldsymbol{\mu}_{k,n}^+ - \boldsymbol{\mu}_{k,n}^-) + (\boldsymbol{\mu}_{k,n-1}^+ - \boldsymbol{\mu}_{k,n-1}^-) \\ M^-(\boldsymbol{\mu}_{k,n}^+ - \boldsymbol{\mu}_{k,n}^-) = \lambda_k(\boldsymbol{\mu}_{k,n}^+ + \boldsymbol{\mu}_{k,n}^-) + (\boldsymbol{\mu}_{k,n-1}^+ + \boldsymbol{\mu}_{k,n-1}^-) \end{cases} \quad \text{and} \quad \begin{cases} M^+\boldsymbol{\mu}_{0,2n+1} = \boldsymbol{\mu}_{0,2n} \\ M^-\boldsymbol{\mu}_{0,2n} = \boldsymbol{\mu}_{0,2n-1} \end{cases}. \quad (3.33)$$

Giving, for $p < n$, for $k > 0$

$$\begin{cases} (M^-M^+ - \lambda_k^2)^p(\boldsymbol{\mu}_{k,n}^+ + \boldsymbol{\mu}_{k,n}^-) \neq \mathbf{0} \\ (M^+M^- - \lambda_k^2)^p(\boldsymbol{\mu}_{k,n}^+ - \boldsymbol{\mu}_{k,n}^-) \neq \mathbf{0} \end{cases} \quad \text{and} \quad \begin{cases} (M^-M^+)^p\boldsymbol{\mu}_{0,2n+1} \neq \mathbf{0} \\ (M^+M^-)^p\boldsymbol{\mu}_{0,2n} \neq \mathbf{0} \end{cases} \\ \begin{cases} (M^-M^+ - \lambda_k^2)^{n-1}(\boldsymbol{\mu}_{k,n}^+ + \boldsymbol{\mu}_{k,n}^-) = \boldsymbol{\mu}_{k,1}^+ + \boldsymbol{\mu}_{k,1}^- \\ (M^+M^- - \lambda_k^2)^{n-1}(\boldsymbol{\mu}_{k,n}^+ - \boldsymbol{\mu}_{k,n}^-) = \boldsymbol{\mu}_{k,1}^+ - \boldsymbol{\mu}_{k,1}^- \end{cases} \quad \text{and} \quad \begin{cases} (M^-M^+)^{n-1}\boldsymbol{\mu}_{0,2n-1} = \boldsymbol{\mu}_{0,1} \\ (M^+M^-)^{n-1}\boldsymbol{\mu}_{0,2n} = \boldsymbol{\mu}_{0,2}, \end{cases} \quad (3.34)$$

so that,

$$((\boldsymbol{\mu}_{k,n}^+ + \boldsymbol{\mu}_{k,n}^-)_{k,1 < n \leq p_k}, (\boldsymbol{\mu}_{0,2n-1})_{1 < n \leq p_0/2}),$$

respectively

$$((\boldsymbol{\mu}_{k,n}^+ - \boldsymbol{\mu}_{k,n}^-)_{k,1 < n \leq p_k}, (\boldsymbol{\mu}_{0,2n})_{1 < n \leq p_0/2}),$$

are generalized eigenvectors of M^-M^+ (resp. of M^+M^-).

More generally, the spectral properties of B and M^+M^- , M^-M^+ are strongly related. Indeed, the characteristic polynomial of B writes itself in terms of B^2 , as,

$$\begin{aligned} B^{p_0} \prod_{k=1}^d (B - \lambda_k I)^{p_k} (B + \lambda_k I)^{p_k} &= (B^2)^{p_0/2} \prod_{k=1}^d (B^2 - \lambda_k^2 I)^{p_k} = 0 \\ B^{2(p_0/2)} \sum_{m=0}^d B^{2m} \sum_{\{i_k\}_{k=1}^{d-m}} \prod_{k=1}^{d-m} \lambda_{i_k}^2 &= 0. \end{aligned} \quad (3.35)$$

And, as,

$$B^{2m} = \frac{1}{2} \begin{pmatrix} (M^-M^+)^m + (M^+M^-)^m & (M^-M^+)^m - (M^+M^-)^m \\ (M^-M^+)^m - (M^+M^-)^m & (M^-M^+)^m + (M^+M^-)^m \end{pmatrix}$$

it is equivalent to,

$$\begin{cases} (M^+M^-)^{p_0/2} \sum_{m=0}^d (M^+M^-)^m \sum_{\{i_k\}_{k=1}^{d-m}} \prod_{k=1}^{d-m} \lambda_{i_k}^2 = 0 \\ (M^-M^+)^{p_0/2} \sum_{m=0}^d (M^-M^+)^m \sum_{\{i_k\}_{k=1}^{d-m}} \prod_{k=1}^{d-m} \lambda_{i_k}^2 = 0. \end{cases} \quad (3.36)$$

So that, the characteristic polynomials of M^+M^- and M^-M^+ are,

$$(M^\pm M^\mp)^{p_0/2} \prod_{k=1}^d (M^\pm M^\mp - \lambda_k^2 I)^{p_k} = 0. \quad (3.37)$$

A similar derivation can be carried out regarding the minimal polynomial and, thus, the fact that B admits no generalized eigenvectors apart the necessary $\boldsymbol{\mu}_{0,2}$ (i.e. $p_0 = 2$, $p_{k \neq 0} = 1$) is equivalent to M^+M^- and M^-M^+ both diagonalizable.

Therefore, studying the squared global bulk matrix B^2 or equivalently the submatrices M^+M^- and M^-M^+ , seems the natural way to deal with the bulk evolution. Indeed, it underlines the fact that, instead of considering the global bulk evolution $\partial_r \gamma = B\gamma$, one can consider the system composed of the symmetrical bulk evolution,

$$\partial_r(\gamma_+ + \gamma_-) = M^-(\gamma_+ - \gamma_-), \quad (3.38)$$

and of the antisymmetrical bulk evolution,

$$\partial_r(\gamma_+ - \gamma_-) = M^+(\gamma_+ + \gamma_-) \quad (3.39)$$

which then gives the uncoupled second-order system,

$$\begin{cases} \partial_r^2(\gamma_+ + \gamma_-) = M^-M^+(\gamma_+ + \gamma_-) \\ \partial_r^2(\gamma_+ - \gamma_-) = M^+M^-(\gamma_+ - \gamma_-) \end{cases}, \quad (3.40)$$

which leads to catenary-like solutions in common cases, as the finite tumble.

3.8.3 General global-jamming solution

Bulk general form

Exploiting the Jordan form of B , the solution to Eq. 3.24 is,

$$\begin{aligned} \begin{pmatrix} \gamma^+ \\ \gamma^- \end{pmatrix} (r) &= \sum_{n=1}^{p_0} \sum_{l=n}^{p_0} a_{0,l} \frac{r^{l-n}}{(l-n)!} \begin{pmatrix} \boldsymbol{\mu}_{0,n} \\ (-1)^{n+1} \boldsymbol{\mu}_{0,n} \end{pmatrix} \\ &+ \sum_{k=1}^d \left[\exp(-\lambda_k (\frac{L}{2} - r)) \sum_{n=1}^{p_k} \sum_{l=n}^{p_k} a_{k,l} \frac{r^{l-n}}{(l-n)!} \begin{pmatrix} \boldsymbol{\mu}_{k,n}^+ \\ \boldsymbol{\mu}_{k,n}^- \end{pmatrix} \right. \\ &\left. + \exp(\lambda_k (\frac{L}{2} - r)) \sum_{n=1}^{p_k} \sum_{l=n}^{p_k} a_{-k,l} \frac{r^{l-n}}{(l-n)!} (-1)^{n+1} \begin{pmatrix} \boldsymbol{\mu}_{k,n}^- \\ \boldsymbol{\mu}_{k,n}^+ \end{pmatrix} \right], \end{aligned} \quad (3.41)$$

with $(a_{\pm k,l})_{k,l}$ real constants and $a_{0,1} = 0$ (no uniform component along $(\boldsymbol{\mu}_B, \boldsymbol{\mu}_B)$). It is straightforward to check that if $\gamma^+(r) = \gamma^-(r)$ for all r , then $\gamma^+(r) = \gamma^-(r) = 0$ and we recover the detailed-jamming case. We can also check for the active global balance,

$$\begin{aligned} \sum_s \phi(s)(\gamma^+(r) - \gamma^-(r))_s &= 0, \\ \sum_{k=1}^d \sum_{n=1}^{p_k} \sum_{l=n}^{p_k} \frac{r^{l-n}}{(l-n)!} (a_{k,l} e^{-\lambda_k (\frac{L}{2} - r)} - (-1)^{n+1} a_{-k,l} e^{\lambda_k (\frac{L}{2} - r)}) \sum_s \phi(s)(\boldsymbol{\mu}_{k,n}^+(s) - \boldsymbol{\mu}_{k,n}^-(s)) \\ &+ \sum_{n=1}^{p_0/2} \sum_{l=2n}^{p_0} \frac{r^{l-2n}}{(l-2n)!} 2a_{0,l} \sum_s \phi(s) \boldsymbol{\mu}_{0,2n}(s) = 0. \end{aligned} \quad (3.42)$$

This is verified as, we have for all $k \geq 1$ ($\lambda_k \neq 0$) by summing along $\phi(s)$ in Eq. 3.33 and reminiscing Eq. 3.30,

$$\begin{aligned} \sum_s \phi(s)(\boldsymbol{\mu}_{k,n}^+(s) - \boldsymbol{\mu}_{k,n}^-(s)) &= -\frac{1}{\lambda_k} \sum_s \phi(s)(\boldsymbol{\mu}_{k,n-1}^+(s) - \boldsymbol{\mu}_{k,n-1}^-(s)) \\ &= \frac{(-1)^{n-1}}{\lambda_k^{n-1}} \sum_s \phi(s)(\boldsymbol{\mu}_{k,1}^+(s) - \boldsymbol{\mu}_{k,1}^-(s)) \\ &= \frac{(-1)^{n-1}}{\lambda_k^n} \sum_{s'} \underbrace{\left[\sum_s \phi(s)(B^+ + B^-)_{ss'} \right]}_{=0 \text{ see Eq. 3.30}} (\boldsymbol{\mu}_{k,1}^+(s') + \boldsymbol{\mu}_{k,1}^-(s')) \\ &= 0, \end{aligned}$$

And for $k = 0$,

$$\sum_s \phi(s) \boldsymbol{\mu}_{0,2n}(s) = \sum_{s'} \underbrace{\left[\sum_s \phi(s) (B^+ + B^-)_{ss'} \right]}_{=0 \text{ see Eq. 3.30}} \boldsymbol{\mu}_{0,2n+1}(s').$$

Periodic boundary constraint

We now constrain the constants $(a_{\pm k,n})_{k,n}$ by the boundary conditions. First, rewriting the general solution (3.41) into,

$$\begin{aligned} \begin{pmatrix} \gamma^+ \\ \gamma^- \end{pmatrix} (r) &= \sum_{n=1}^{p_0} \sum_{l=n}^{p_0} \left[\sum_{m=l}^{p_0} a_{0,m} \frac{\left(\frac{L}{2}\right)^{m-l}}{(m-l)!} \right] \frac{\left(r - \frac{L}{2}\right)^{l-n}}{(l-n)!} \begin{pmatrix} \boldsymbol{\mu}_{0,n} \\ (-1)^{n+1} \boldsymbol{\mu}_{0,n} \end{pmatrix} \\ &+ \sum_{k=1}^d \left[\exp(-\lambda_k \left(\frac{L}{2} - r\right)) \sum_{n=1}^{p_k} \sum_{l=n}^{p_k} \left[\sum_{m=l}^{p_k} a_{k,m} \frac{\left(\frac{L}{2}\right)^{m-l}}{(m-l)!} \right] \frac{\left(r - \frac{L}{2}\right)^{l-n}}{(l-n)!} \begin{pmatrix} \boldsymbol{\mu}_{k,n}^+ \\ \boldsymbol{\mu}_{k,n}^- \end{pmatrix} \right. \\ &\left. + \exp(\lambda_k \left(\frac{L}{2} - r\right)) \sum_{n=1}^{p_k} \sum_{l=n}^{p_k} \left[\sum_{m=l}^{p_k} (-1)^{n+1} a_{-k,m} \frac{\left(\frac{L}{2}\right)^{m-l}}{(m-l)!} \right] \frac{\left(r - \frac{L}{2}\right)^{l-n}}{(l-n)!} \begin{pmatrix} \boldsymbol{\mu}_{k,n}^- \\ \boldsymbol{\mu}_{k,n}^+ \end{pmatrix} \right], \end{aligned} \quad (3.43)$$

the periodic boundary condition $\gamma^+(L/2^-) = \gamma^-(L/2^-)$ yields,

$$\begin{aligned} \sum_{n=1}^{p_0/2} 2 \left[\sum_{m=2n}^{p_0} a_{0,m} \frac{\left(\frac{L}{2}\right)^{m-2n}}{(m-2n)!} \right] \boldsymbol{\mu}_{0,2n} \\ + \sum_{k=1}^d \sum_{n=1}^{p_k} \left[\sum_{m=n}^{p_k} (a_{k,m} - (-1)^{n+1} a_{-k,m}) \frac{\left(\frac{L}{2}\right)^{m-n}}{(m-n)!} \right] (\boldsymbol{\mu}_{k,n}^+ - \boldsymbol{\mu}_{k,n}^-) = 0. \end{aligned}$$

As the vectors $(\boldsymbol{\mu}_{k,n}^+ - \boldsymbol{\mu}_{k,n}^-)_{k,n}, (\boldsymbol{\mu}_{0,2n})_n$ are linearly independent, we obtain the condition,

$$\begin{cases} \sum_{m=2n}^{p_0} a_{0,m} \frac{\left(\frac{L}{2}\right)^{m-2n}}{(m-2n)!} = 0, \text{ for all } n \leq p_0/2 \\ \sum_{m=n}^{p_k} (a_{k,m} - a_{-k,m} (-1)^{n+1}) \frac{\left(\frac{L}{2}\right)^{m-n}}{(m-n)!} = 0, \text{ for all } n \leq p_k \end{cases}, \quad (3.44)$$

constraining γ into the following symmetrical decomposition in respect to $\gamma^+(r)$ and $\gamma^-(L-r)$,

$$\begin{aligned} \begin{pmatrix} \gamma^+(r) \\ \gamma^-(r) \end{pmatrix} &= \sum_{l=0}^{p_0/2-1} b_{0,2l+1} \frac{\left(\frac{L-r}{2}\right)^{2l}}{(2l)!} \begin{pmatrix} \boldsymbol{\mu}_{\mathcal{B}} \\ \boldsymbol{\mu}_{\mathcal{B}} \end{pmatrix} \\ &+ \sum_{n=1}^{p_0/2-1} \sum_{l=n}^{p_0/2-1} b_{0,2l+1} \frac{\left(\frac{L-r}{2}\right)^{2(l-n)}}{(2(l-n))!} \left[\begin{pmatrix} \boldsymbol{\mu}_{0,2n+1} \\ \boldsymbol{\mu}_{0,2n+1} \end{pmatrix} - \frac{\left(\frac{L-r}{2}\right)}{2(l-n)+1} \begin{pmatrix} \boldsymbol{\mu}_{0,2n} \\ -\boldsymbol{\mu}_{0,2n} \end{pmatrix} \right] \\ &+ \sum_{k=1}^d \sum_{n=1}^{p_k} \sum_{l=n}^{p_k} b_{k,l} \frac{\left(\frac{L-r}{2}\right)^{l-n}}{(l-n)!} \left[(-1)^{l-n} e^{-\lambda_k \left(\frac{L}{2} - r\right)} \begin{pmatrix} \boldsymbol{\mu}_{k,n}^+ \\ \boldsymbol{\mu}_{k,n}^- \end{pmatrix} + e^{\lambda_k \left(\frac{L}{2} - r\right)} \begin{pmatrix} \boldsymbol{\mu}_{k,n}^- \\ \boldsymbol{\mu}_{k,n}^+ \end{pmatrix} \right], \end{aligned} \quad (3.45)$$

with,

$$\begin{cases} b_{0,2l+1} = \sum_{m=2l+1}^{p_0-1} a_{0,m} \frac{\left(\frac{L}{2}\right)^{m-2l-1}}{(m-2l-1)!}, \text{ for } 1 \leq 2l+1 \leq p_0-1 \\ a_{0,1} = 0 \text{ and } a_{0,2l} = -\sum_{m=2l+1}^{p_0} a_{0,m} \frac{\left(\frac{L}{2}\right)^{m-2l}}{(m-2l)!}, \text{ for } 2 \leq 2l \leq p_0-2 \\ b_{k,l} = \sum_{m=l}^{p_k} a_{k,m} \frac{\left(\frac{L}{2}\right)^{m-l}}{(m-l)!} = (-1)^{l+1} \sum_{m=l}^{p_k} a_{-k,m} \frac{\left(\frac{L}{2}\right)^{m-l}}{(m-l)!}, \text{ for } 1 \leq l \leq p_k \end{cases}. \quad (3.46)$$

Jamming boundary constraint

We now consider the constraint imposed by the jamming interaction. The condition in 0^+ set by $(C_{\mathcal{J}})$ leads to,

$$\begin{cases} w_{\mathcal{J}} B^+ \boldsymbol{\mu}_{\mathcal{J}}^- = -\boldsymbol{\pi}_{\text{glob}}(0^+, -1) \\ w_{\mathcal{J}} B^- \boldsymbol{\mu}_{\mathcal{J}}^- = \boldsymbol{\pi}_{\text{glob}}(0^+, 1) \end{cases}, \quad (3.47)$$

with $\boldsymbol{\mu}_{\mathcal{J}}^- = (\boldsymbol{\mu}_{\mathcal{J}}(-1, s))_s$, $\boldsymbol{\pi}_{\text{glob}}(0^+, -1) = (\boldsymbol{\pi}_{\text{glob}}(0^+, -1, s))_s$, $\boldsymbol{\pi}_{\text{glob}}(0^+, 1) = (\boldsymbol{\pi}_{\text{glob}}(0^+, 1, s))_s$. It decomposes into,

$$\begin{cases} \frac{w_{\mathcal{J}}}{w_{\mathcal{B}}}(B^+ + B^-)\boldsymbol{\mu}_{\mathcal{J}}^- = w_{\text{rel}}(\boldsymbol{\gamma}^+(0^+) - \boldsymbol{\gamma}^-(0^+)) \\ -\frac{w_{\mathcal{J}}}{w_{\mathcal{B}}}(B^+ - B^-)\boldsymbol{\mu}_{\mathcal{J}}^- = (2w_{\text{eq}}\boldsymbol{\mu}_{\mathcal{B}} + w_{\text{rel}}(\boldsymbol{\gamma}^+(0^+) + \boldsymbol{\gamma}^-(0^+))) \end{cases} \quad (3.48)$$

It clearly appears that recovering a detailed-jamming solution (i.e. $w_{\text{rel}} = 0$) imposes $(B^+ + B^-)\boldsymbol{\mu}_{\mathcal{J}}^- = 0$, i.e. $\boldsymbol{\mu}_{\mathcal{J}}^- \propto \boldsymbol{\mu}_{\mathcal{B}}$, which, combined with the second condition, then makes the detailed-jamming condition equivalent to $\boldsymbol{\mu}_{\mathcal{B}}$ being the unique eigenvector shared by B^+ and B^- with opposite eigenvalues. It also imposes that $(B^+ - B^-)\boldsymbol{\mu}_{\mathcal{J}}^- \propto \boldsymbol{\mu}_{\mathcal{B}}$, i.e. $\boldsymbol{\mu}_{\mathcal{J}}^- \propto \boldsymbol{\mu}_{0,2}$, which leads to the condition $\boldsymbol{\mu}_{0,2} \propto \boldsymbol{\mu}_{\mathcal{B}}$.

Now assuming $w_{\text{rel}} \neq 0$, i.e. $\boldsymbol{\mu}_{\mathcal{J}}^- \not\propto \boldsymbol{\mu}_{\mathcal{B}}$ nor $\boldsymbol{\mu}_{\mathcal{J}}^- \not\propto \boldsymbol{\mu}_{0,2}$, we decompose $\boldsymbol{\mu}_{\mathcal{J}}^-$ along the two basis of general eigenvectors,

$$\begin{aligned} \boldsymbol{\mu}_{\mathcal{J}}^- &= \sum_{n=1}^{p_0/2} c_{0,2n} \boldsymbol{\mu}_{0,2n} + \sum_{k=1}^d \sum_{n=1}^{p_k} c_{k,n}^- (\boldsymbol{\mu}_{k,n}^+ - \boldsymbol{\mu}_{k,n}^-) \\ &= \sum_{n=0}^{p_0/2-1} c_{0,2n+1} \boldsymbol{\mu}_{0,2n+1} + \sum_{k=1}^d \sum_{n=1}^{p_k} c_{k,n}^+ (\boldsymbol{\mu}_{k,n}^+ + \boldsymbol{\mu}_{k,n}^-), \end{aligned} \quad (3.49)$$

so that,

$$\begin{aligned} -(B^+ - B^-)\boldsymbol{\mu}_{\mathcal{J}}^- &= -\sum_{n=1}^{p_0/2} c_{0,2n} \boldsymbol{\mu}_{0,2n-1} - \sum_{k=1}^d \left[\lambda_k c_{k,p_k}^- (\boldsymbol{\mu}_{k,p_k}^+ + \boldsymbol{\mu}_{k,p_k}^-) \right. \\ &\quad \left. + \sum_{n=1}^{p_k-1} (\lambda_k c_{k,n}^- + c_{k,n+1}^-) (\boldsymbol{\mu}_{k,n}^+ + \boldsymbol{\mu}_{k,n}^-) \right], \\ (B^+ + B^-)\boldsymbol{\mu}_{\mathcal{J}}^- &= \sum_{n=1}^{p_0/2-1} c_{0,2n+1} \boldsymbol{\mu}_{0,2n} + \sum_{k=1}^d \left[\lambda_k c_{k,p_k}^+ (\boldsymbol{\mu}_{k,p_k}^+ - \boldsymbol{\mu}_{k,p_k}^-) \right. \\ &\quad \left. + \sum_{n=1}^{p_k-1} (\lambda_k c_{k,n}^+ + c_{k,n+1}^+) (\boldsymbol{\mu}_{k,n}^+ - \boldsymbol{\mu}_{k,n}^-) \right]. \end{aligned} \quad (3.50)$$

From the expression of $\boldsymbol{\gamma}^{\pm}$ (3.45), we get that,

$$\begin{aligned} \boldsymbol{\gamma}^+(0) + \boldsymbol{\gamma}^-(0) &= \sum_{n=0}^{p_0/2-1} \sum_{l=n}^{p_0/2-1} 2b_{0,2l+1} \frac{(\frac{L}{2})^{2(l-n)}}{(2(l-n))!} \boldsymbol{\mu}_{0,2n+1} \\ &\quad + \sum_{k=1}^d \sum_{n=1}^{p_k} \sum_{l=n}^{p_k} b_{k,l} \frac{(\frac{L}{2})^{l-n}}{(l-n)!} \left[\exp(\lambda_k \frac{L}{2}) + (-1)^{l-n} \exp(-\lambda_k \frac{L}{2}) \right] (\boldsymbol{\mu}_{k,n}^+ + \boldsymbol{\mu}_{k,n}^-), \end{aligned} \quad (3.51)$$

and,

$$\begin{aligned} \boldsymbol{\gamma}^+(0) - \boldsymbol{\gamma}^-(0) &= -\sum_{n=1}^{p_0/2-1} \sum_{l=n}^{p_0/2-1} 2b_{0,2l+1} \frac{(\frac{L}{2})^{2(l-n)+1}}{(2(l-n)+1)!} \boldsymbol{\mu}_{0,2n} \\ &\quad + \sum_{k=1}^d \sum_{n=1}^{p_k} \sum_{l=n}^{p_k} b_{k,l} \frac{(\frac{L}{2})^{l-n}}{(l-n)!} \left[-\exp(\lambda_k \frac{L}{2}) + (-1)^{l-n} \exp(-\lambda_k \frac{L}{2}) \right] (\boldsymbol{\mu}_{k,n}^+ - \boldsymbol{\mu}_{k,n}^-). \end{aligned} \quad (3.52)$$

Plugging those expression in (3.48), we obtain,

$$\begin{cases} w_{\mathcal{J}} c_{0,2} = -2w_{\mathcal{B}}(w_{\text{eq}} + w_{\text{rel}} \sum_{l=0}^{p_0/2-1} b_{0,2l+1} \frac{(L/2)^{2l}}{(2l)!}) \\ w_{\mathcal{J}} c_{0,2(n+1)} = -2w_{\mathcal{B}} w_{\text{rel}} \sum_{l=n}^{p_0/2-1} b_{0,2l+1} \frac{(L/2)^{2(l-n)}}{(2(l-n))!} \text{ for } n \geq 1 \\ w_{\mathcal{J}} c_{0,2n+1} = -2w_{\mathcal{B}} w_{\text{rel}} \sum_{l=n}^{p_0/2-1} b_{0,2l+1} \frac{(L/2)^{2(l-n)+1}}{(2(l-n)+1)!} \text{ for } n \geq 1 \\ w_{\mathcal{J}} (\lambda_k c_{k,n}^- + c_{k,n+1}^-) = -w_{\mathcal{B}} w_{\text{rel}} \sum_{l=n}^{p_k} b_{k,l} \frac{(L/2)^{l-n}}{(l-n)!} \left[e^{\lambda_k \frac{L}{2}} + (-1)^{l-n} e^{-\lambda_k \frac{L}{2}} \right] \text{ for } n < p_k \\ w_{\mathcal{J}} (\lambda_k c_{k,n}^+ + c_{k,n+1}^+) = -w_{\mathcal{B}} w_{\text{rel}} \sum_{l=n}^{p_k} b_{k,l} \frac{(L/2)^{l-n}}{(l-n)!} \left[e^{\lambda_k \frac{L}{2}} - (-1)^{l-n} e^{-\lambda_k \frac{L}{2}} \right] \text{ for } n < p_k \\ w_{\mathcal{J}} \lambda_k c_{k,p_k}^- = -w_{\mathcal{B}} w_{\text{rel}} b_{k,p_k} \left[e^{\lambda_k \frac{L}{2}} + e^{-\lambda_k \frac{L}{2}} \right] = w_{\mathcal{J}} \lambda_k c_{k,p_k}^+ \coth(\lambda_k \frac{L}{2}) \end{cases} \quad (3.53)$$

Recovering a catenary-like solution (i.e. no polynomial modulation) then requires

$$\begin{cases} w_{\mathcal{J}}c_{0,2} = -2w_{\mathcal{B}}w_{\text{eq}} \\ c_{0,2(n+1)} = c_{0,2n+1} = 0 \text{ for } n \geq 1 \\ w_{\mathcal{J}}\lambda_k c_{k,1}^- = -w_{\mathcal{B}}w_{\text{rel}}b_{k,1} \left[\exp\left(\lambda_k \frac{L}{2}\right) + \exp\left(-\lambda_k \frac{L}{2}\right) \right] = w_{\mathcal{J}}\lambda_k c_{k,1}^+ \text{cotanh}\left(\lambda_k \frac{L}{2}\right) \\ c_{k,n}^- = c_{k,n}^+ = b_{k,n} = 0 \text{ for } n > 1 \end{cases}, \quad (3.54)$$

which, following (3.49), is possible only if the generalized eigenvector $\mu_{0,2}$ can be decomposed over $(\mu_{\mathcal{B}}, (\mu_{k,1}^+ + \mu_{k,1}^-)_{k \in N})$, N some subset of $\llbracket 1, d \rrbracket$, as well as the corresponding vectors $(\mu_{k,1}^+ - \mu_{k,1}^-)_{k \in N}$ so that the third relation in (3.54) is obeyed.

More generally, the more vectors of $(\mu_{\mathcal{B}}, (\mu_{k,n}^+ + \mu_{k,n}^-)_{k,n})$ are needed to decompose $\mu_{0,2}$ over, the richer and less close to its equilibrium counterpart the bulk relaxation behavior will be. A special case is the detailed-jamming situation where $\mu_{0,2}$ is proportional to $\mu_{\mathcal{B}}$ and $N = \emptyset$ or the catenary-like relaxation where only the eigenvectors are involved in the decomposition.

That analysis of the impact of the jamming boundary constraint on the bulk relaxation can be generalized to a tumble Markov kernel Q' different from Q , leading to different matrices B'^+ , B'^- in (3.48) and the consideration of their eigenvectors and generalized ones.

3.8.4 Case of a single non-null eigenvalue

As in the case of the finite tumble represented in Figure 3.2, an interesting situation is when the particles have an internal state space and tumbling mechanism presenting enough symmetries so that it can be summarized into a four-state space. In this collapsed representation, the corresponding matrix B_{coll} is of size 4, and, as the Jordan block associated to 0 is at least of size 2, can only admit two Jordan blocks of size 1 linked to the eigenvalues $\pm\lambda \neq 0$. From the analysis carried on in Appendix 3.8.2, λ^2 is the only non-null eigenvalue of $M_{\text{coll}}^+ M_{\text{coll}}^-$ (or $M_{\text{coll}}^- M_{\text{coll}}^+$), so that $\lambda^2 = \text{Tr}(M_{\text{coll}}^+ M_{\text{coll}}^-)$.

Recovering the finite-tumble case, we now detail the case of such a symmetrical three-state RTP system. We note the three states $\{v_1, v_2, v_3\}$. The symmetry leading to a possible collapse to a four-state space is obeyed if two pair states, e.g. $\{v_1, v_2\}$ and $\{v_2, v_3\}$ are equivalent, i.e. present the same run and tumble mechanisms, leading to $\pi(r, \pm 1, \{v_2, v_3\}) = \pi(r, \pm 1, \{v_1, v_2\})$, and more precisely framed by the conditions,

$$\begin{cases} q(v_3, v_1) = q(v_1, v_3) \\ q(v_2, v_1) = q(v_2, v_3) = \frac{1}{2} \\ q(v_1, v_2) = q(v_3, v_2) \\ |v_3 - v_2| = |v_2 - v_1| \\ \omega(v_3) = \omega(v_1) \end{cases}. \quad (3.55)$$

In the finite-tumble case, the central state v_2 is the state 0. This leads to the following matrices,

$$\begin{aligned} B^+ + B^- &= \frac{(1 + q(v_1, v_3))}{|v_1 - v_2|} \begin{pmatrix} -\omega(v_1)(1 - q(v_1, v_3)) & \frac{1}{4}\omega(v_2) & \frac{1}{4}\omega(v_2) \\ \omega(v_1)(1 - q(v_1, v_3)) & -\frac{1}{2}(\omega(v_1) + \omega(v_2)) & \frac{1}{2}\omega(v_1) \\ \omega(v_1)(1 - q(v_1, v_3)) & \frac{1}{2}\omega(v_1) & -\frac{1}{2}(\omega(v_1) + \omega(v_2)) \end{pmatrix}, \\ B^+ - B^- &= \frac{1}{|v_1 - v_2|} \begin{pmatrix} -\omega(v_1) & \frac{1}{4}\omega(v_2) & \frac{1}{4}\omega(v_2) \\ \omega(v_1)(1 - q(v_1, v_3)) & -(\omega(v_1) + \omega(v_2)) & -\omega(v_1)q(v_1, v_3) \\ \omega(v_1)(1 - q(v_1, v_3)) & -\omega(v_1)q(v_1, v_3) & -(\omega(v_1) + \omega(v_2)) \end{pmatrix}, \end{aligned} \quad (3.56)$$

so that the eigenvalues of $M^+ M^-$ and $M^- M^+$ are 0 and

$$\lambda^2 = (\omega(v_1)(1 - q(v_1, v_3)) + \omega(v_2))(\omega(v_1) + \frac{1}{2}\omega(v_2)) \frac{1 + q(v_1, v_3)}{(v_1 - v_2)^2}. \quad (3.57)$$

In particular, in the case of the finite tumble, i.e. $v_1 = -1, v_2 = 0, v_3 = 1$ and $\omega(v_1) = \omega(v_3) = \alpha$, $\omega(v_2) = \beta$ and $q(v_1, v_3) = q(v_3, v_1) = 0$, the typical inverse lengthscale of the catenary relaxation is,

$$\lambda = \sqrt{(\alpha + \beta)(\alpha + \frac{1}{2}\beta)}. \quad (3.58)$$

More generally, if the persistent run dominates over the tumbles ($|v_1 - v_2| \gg \omega(v_1) + \omega(v_2)$), then $\lambda \rightarrow 0$ and the system does not decorrelate from the jamming constraint. In the opposite situation where the tumbles dominates over the run ($|v_1 - v_2| \ll \omega(v_1) + \omega(v_2)$), then $\lambda \rightarrow \infty$ and the density immediately relaxes to the bulk uniform solution set by μ_B .

Chapter 4

Long time analysis of a pair of on-lattice and continuous run-and-tumble particles with jamming

This chapter is based on an upcoming article with Manon Michel and Arnaud Guillin.

4.1 Introduction, models and main results

4.1.1 Introduction

Run-and-tumble particles (RTPs), which randomly switch between phases of linear motion (runs) and localized chaotic motion (tumbles) resulting in reorientation, are a prime example of active matter [Ram10]. Their self-propulsion breaks time reversibility and energy conservation at the microscopic scale, two fundamental properties of equilibrium systems. This leads to rich out-of-equilibrium behavior such as collective motion [VZ12], pattern formation [BB95] and motility induced phase separation (MIPS) [CT15].

Accumulation at obstacles and MIPS both result in spatially inhomogeneous particle distributions, making the derivation of the steady state in RTP systems a major research focus. In continuous space, jamming interactions between particles result in boundaries in the state space. This poses a challenge when computing the invariant measure, because it necessitates a thorough understanding of boundary conditions. This motivates the introduction of two RTPs on the discrete periodic lattice interacting through jamming [SEB16, SEB17], where there are no boundary conditions. The continuous-space invariant measure, obtained by making the lattice spacing go to zero, displays increased mass at jamming configurations. This increased mass takes the form of Diracs and, in the case of non-zero tumble duration [SEB17], exponentials, indicating effective attraction in line with MIPS. The same discretization scheme is revisited for models with collision induced recoil [MEB22, MEB23], leading to either effective attraction or repulsion are possible depending on model parameters. In more direct approach, the piecewise deterministic Markov process (PDMP) [Dav93] framework with sticky boundaries [BGMS23] makes boundary conditions transparent and allows to work directly with a continuous-space process [HGM23]. This greatly simplifies computations and allows the computation of the invariant measure of RTPs with general tumbling and jamming behavior, revealing the existence of two universality classes corresponding detailed and global probability flow preservation.

The computation of the invariant measures naturally leads to the question of the speed of convergence towards it, which, in the context of RTPs, this is linked to the time it takes

for MIPS to emerge from an arbitrary configuration. Indeed, while the invariant measure of Markov processes determines their behavior at time scales of the order of their mixing time, it becomes irrelevant at smaller time scales. The most commonly used approach is spectral decomposition: the distance to the invariant measure decays exponentially, with an asymptotic rate determined by the second eigenvalue of the generator. The full spectral decomposition of two discrete RTPs on the 1D torus with jamming is computed in [MBE19], revealing eigenvalue crossings and a non-analytic spectral gap. In continuous space with added thermal noise [DDK20a], the generator is a differential operator so the eigenvalue problem takes the form of an ODE. The full spectrum is again computed and the spectral gap is found to freeze to the spectral gap of the velocity dynamics as the length of the 1D torus goes to zero. This occurs because the time it takes to explore position space goes to zero while the time to explore velocity space remains constant. In [AG19] the eigenvalues are obtained as the poles of the Laplace transform in time of the position distribution and the same freezing phenomenon is observed. Unfortunately, when the generator is an unbounded operator, the set of eigenvalues can be a proper subset of the spectrum, complicating the computation of the spectral gap. Furthermore, because RTP systems are non-reversible, there is an uncontrolled prefactor to the exponential decay and the convergence results are only true in the asymptotic regime. For this reason, this article adopts the coupling approach [FGM12] to obtain sharp non-asymptotic bounds.

The first result of this paper is that the continuous-space limit, where the lattice spacing tends to zero, of the Markov jump processes in [SEB16, SEB17] is a variant of the PDMP in [HGM23], rigorously linking the two models. The convergence is made quantitative by relying on a coupling strategy rather than the usual generator approach [EK86]. A uniformity result on the mixing times of the discrete-space processes is then used to prove that the on-lattice invariant measures converge to the continuous invariant measure for the Wasserstein distance, thus giving further theoretical underpinning to [SEB16, SEB17].

An order preservation argument, in the general spirit of [LMT96], reduces the coupling approach to hitting time computations and provides quantitative upper bounds on the mixing time. Matching lower bounds are established by identifying obstacles to mixing, which are then made quantitative by concentration inequalities. This shows that the bounds are optimal and fully capture the dependence of the mixing time on model parameters, incidentally revealing the existence of a persistent and diffusive regime similar to the freezing in [DDK20a, AG19].

In the remainder of this section, the discrete models from [SEB16, SEB17] as well as their continuous-space limit are introduced and the main results are presented. The section ends with a plan of the paper.

4.1.2 Models and main results

Discrete process

The first step is to define the Markov jump process introduced in [SEB16, SEB17] which is the starting point of the analysis in this article. The goal is to show that this discrete model converges in law to a continuous-space stochastic process.

All processes in this paper aim to describe the collective ballistic motion of so-called *run-and-tumble particles* which randomly switch between phases of linear motion (runs) and localized chaotic motion resulting in reorientation (tumbles). In this paper, particles always jam when they collide. It is this particular feature of this particles system which mimics the MIPS phenomenon. A particular emphasis is then at the level of the invariant measure of the system which should exhibit a Dirac mass at the collisions to express this self organisation.

The discrete process of interest in corresponds to the setting of two run-and-tumble particles on a periodic one-dimensional lattice. Each particle is endowed with a velocity and an independent Poisson clock with rate γ . When their clock rings, the particles jump to the neighboring position in the direction of their velocity except if that position is occupied

by the other particle. The particles thus display the persistent motion characteristic of run-and-tumble particles as well as jamming interactions.

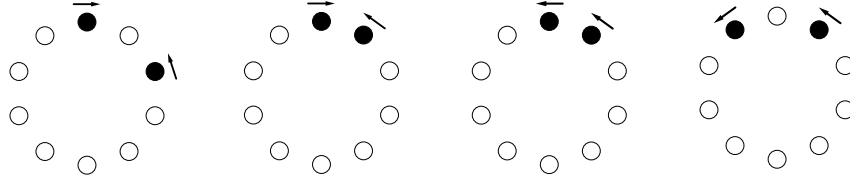


Figure 4.1: Sequence of random position jumps and velocity flips

In [SEB16] the tumbles are approximated by instantaneous reorientations and the velocities are independent Markov jump processes taking the values 1 (moving counter-clockwise) and -1 (moving clockwise) and following the transition rates of figure 4.2a. In [SEB17] the tumbles have a finite nonzero duration so that the velocities can take the additional value 0 (tumbling and not moving) and follow the transition rates of figure 4.2b.

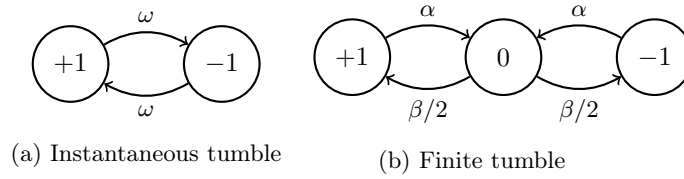


Figure 4.2: Transition rates of the velocity of a single particle

Because the process is translation invariant, it is enough to consider the inter-particle distance y^L and the velocities σ^1 and σ^2 of the two particles. The process $Y^L(t) = (y^L(t), \sigma^1(t), \sigma^2(t))$ is a Markov jump process and its construction is laid out in the following definition.

Definition 4.1 (Discrete process). *Let L be a positive integer and let $y_0^L \in \{1, \dots, L\}$. Let σ^1 and σ^2 be two independent Markov jump processes with the transition rates of figure 4.2a (resp. 4.2b) and let N_1 and N_2 be two independent Poisson clocks with rate γ independent of the σ^i .*

Recursively define $y^L(t)$ as follows

- $y^L(0) = y_0^L$
- when N_1 rings
 - if $y^L(t-) - \sigma^1(t-) \in \{1, \dots, L\}$ then y^L jumps to $y^L(t-) - \sigma^1(t-)$,
 - if $y^L(t-) - \sigma^1(t-) \notin \{1, \dots, L\}$ then nothing happens,
- when N_2 rings
 - if $y^L(t-) + \sigma^2(t-) \in \{1, \dots, L\}$ then y^L jumps to $y^L(t-) + \sigma^2(t-)$,
 - if $y^L(t-) + \sigma^2(t-) \notin \{1, \dots, L\}$ then nothing happens.

To emphasize which σ^i , N_i , y_0^L and L were used, it is sometimes useful to denote the process constructed in definition 4.1 by $\mathcal{Y}(\sigma^1, \sigma^2, N_1, N_2, y_0^L, L)$. The process Y^L is referred to as discrete instantaneous tumble process (DITP) when the velocities follow figure 4.2a and as discrete finite tumble process (DFTP) when the velocities follow figure 4.2b.

Continuous process

Next the process that turns out to be the continuous-space limit of the discrete process under the appropriate rescaling is presented.

Consider two point particles, each endowed with a velocity following a Markov jump process with the transition rates of figure 4.2a (resp. 4.2b), that move around a 1D torus of length $\ell > 0$ as follows

- when their velocity is 1, the particles move counter-clockwise at constant unit velocity,
- when their velocity is -1 , the particles move clockwise at constant unit velocity,
- when their velocity is 0, the particles do not move.

Furthermore, the particles jam when they collide so they can never pass through each other.

Let $x \in [0, \ell]$ be the distance from the first to the second particle and let σ^1 and σ^2 be the velocities of the particles. Further let $[t_1, t_2]$ be a time interval on which the velocities are constant.

- If $\sigma^2(t_1) - \sigma^1(t_1) > 0$ then x increases at constant velocity $\sigma^2(t_1) - \sigma^1(t_1)$ until it reaches ℓ (which corresponds to a configuration where the two particles are jammed) hence

$$x(t) = \min [\ell, x(t_1) + (\sigma^2(t_1) - \sigma^1(t_1))(t - t_1)] \text{ for } t \in [t_1, t_2].$$

- If $\sigma^2(t_1) - \sigma^1(t_1) < 0$ then x decreases at constant velocity $\sigma^2(t_1) - \sigma^1(t_1)$ until it reaches 0 (which corresponds to a configuration where the two particles are jammed) hence

$$x(t) = \max [0, x(t_1) + (\sigma^2(t_1) - \sigma^1(t_1))(t - t_1)] \text{ for } t \in [t_1, t_2].$$

- If $\sigma^2(t_1) - \sigma^1(t_1) = 0$ then x is constant on $[t_1, t_2]$.

This leads to the following recursive construction for the continuous process.

Definition 4.2. (*Continuous process*) Let $x_0 \in [0, \ell]$ and $\sigma_0^1, \sigma_0^2 \in \{-1, 1\}$ (resp. $\sigma_0^1, \sigma_0^2 \in \{-1, 0, 1\}$) be given. Consider two independent Markov jump processes σ^1 and σ^2 with the transition rates of figure 4.2a (resp. 4.2b) and respective initial states σ_0^1 and σ_0^2 .

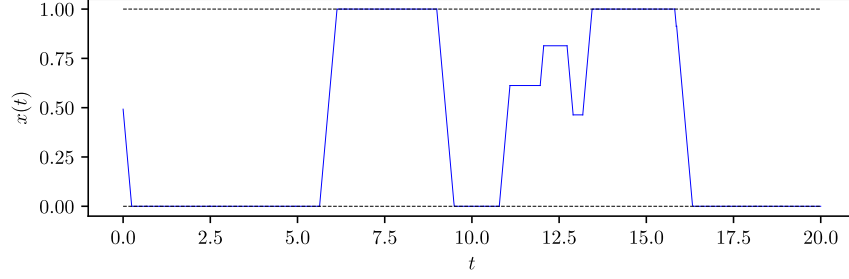
Using the jump times $(T_i)_{i \geq 0}$ of the couple (σ^1, σ^2) recursively define $x(t)$ by $x(0) = x_0$ and

$$wx(t) = \max [0, \min [\ell, x(T_i) + (\sigma^2(T_i) - \sigma^1(T_i))(t - T_i)]] \text{ for } t \in [T_i, T_{i+1}].$$

The process $X(t) = (x(t), \sigma^1(t), \sigma^2(t))$ is referred to as continuous process.

In order to emphasize which σ^i , x_0 and ℓ were used, the process constructed in the previous definition is sometimes denoted by $\mathcal{X}(\sigma^1, \sigma^2, x_0, \ell)$. When the σ^i follow the rates of figure 4.2a, X is called the continuous instantaneous tumble process (CITP) and when the σ^i follow the rates of figure 4.2b X is called the continuous finite tumble process (CFTP).

If definition 4.2 was chosen as a starting point, the proof of the strong Markov property and the characterization of the generator would require careful work. Instead, it is better to construct the process using the formalism of piecewise deterministic Markov processes (PDMPs) developed in [Dav93] so that the strong Markov property and the characterization of the generator are immediate. Borrowing an idea from [BGvdMS], a bijection ι mapping the state space of the PDMP to a more convenient state space is used to come back to definition 4.2. It is the subject of Section 4.2.

Figure 4.3: Realization of x for the CITP with $\omega = 0.5$ and $\ell = 1$

Main results

The following theorem is a summary of the findings obtained in this paper.

Theorem 4.3. *Let $\gamma_L = (L - 1)/\ell$.*

- (1) *The discrete in space processes DITP (resp. DFTP) converge, with quantitative control, when L tends to infinity, in Skorohod topology towards the continuous-space piecewise deterministic Markov process CITP (resp. CFTP).*
- (2) *The processes DITP and DFTP are exponentially ergodic with invariant measures converging as L tends to infinity to the unique invariant measures of the CITP and CFTP.*
- (3) *The invariant measures of CITP and CFTP have a positive mass at $\{0\} \times \{-1, 0, 1\}$ and $\{\ell\} \times \{-1, 0, 1\}$.*
- (4) *Let us define for a Markov process with semigroup P_t and invariant measure its mixing time as $t_{mix}(\epsilon) = \inf \{t \geq 0; \sup_{\mu} \|\mu P_t - \pi\|_{TV} \leq \epsilon\}$.*
 - (a) *The mixing time $t_{mix}(\epsilon)$ of CITP is of the order $C'(\epsilon)(1 + \omega^2 l^2)/\omega$ for some $C'(\epsilon)$.*
 - (b) *The mixing time $t_{mix}(\epsilon)$ of CFTP is of the order $C'(\epsilon)(\frac{1}{\alpha} + \frac{1}{\beta})(1 + \alpha^2 l^2)$ for some $C'(\epsilon)$.*

The precise statements of points (1) and (2) and (3) are given in Theorem 4.8, Corollary 4.14 and in the appendix. It enables to justify the analysis of [SEB16, SEB17] who have calculated the invariant measure of the DITP and DFTP and have passed to the limit in order to assert that this processes exhibit a MIPS phenomenon, i.e. only their intrinsic movement allows for agglomerate of particles asserted by a positive mass at 0 and ℓ . This result not only ensures the validity of their analysis but also provides a direct way to build processes (PDMP) at the continuous level. Note that the explicit invariant measures obtained in the Appendix is a particular case of the general study of [HGM23] where universality classes (for the invariant measures) were derived for general pairs of RTP depending on the behavior at the jamming boundary. The coupling strategy used here to derive the discrete to continuous limit is quite general and is not limited to the DITP/DFTP processes of [SEB16, SEB17].

The mixing times were not considered before, despite the importance of the understanding of the speed of convergence to stationarity. The precise statements of the results are given in Theorem 4.20. Most interestingly, it enables to identify two regimes

- the *persistent* regime $\omega \ell \ll 1$ (resp. $\alpha \ell \ll 1$) where the mean time between two stochastic velocity changes is much greater than the time needed to reach a jamming configuration,
- the *diffusive* regime $\omega \ell \gg 1$ (resp. $\alpha \ell \gg 1$) where the mean time between two stochastic velocity changes is much lower than the time needed to reach a jamming configuration.

This allows for a heuristic interpretation of the mixing time in terms of slow observables. In the persistent regime, the slowest observables are the velocities σ^i so the mixing time is of order $1/\omega$ (resp. $1/\alpha + 1/\beta$). In the diffusive regime, the position x is the slowest observable and its dynamic is approximately Brownian leading to the mixing time $\omega\ell^2$ (resp. $(1/\alpha + 1/\beta)\alpha^2\ell^2$) with the quadratic dependence in ℓ characteristic of diffusions. Notice that, in the case of the CFTP, the diffusivity is measured by $\alpha\ell$ alone. This can also be seen in the fact that x -marginal of the invariant measure of the CFTP converges to a sum of Dirac measures as $\alpha\ell \rightarrow 0$, and to the uniform measure as $\alpha\ell \rightarrow +\infty$, both limits being uniform in β .

Plan of the paper

In the following, in section 4.2, the stochastic processes CITP and CFTP are constructed considering the formulation of Davis [Dav93], with a particular emphasis on their generator, which are helpful to derive the explicit form of the invariant measure but also for mixing time evaluation. In section 4.3, it is shown that, under the right rescaling, the discrete process converges in law to the continuous process with respect to the Skorokhod metric. It is also proven that the on-lattice invariant measures converge to the continuous invariant measure with respect to the Wasserstein distance. Proofs are based on coupling between the discrete state process and the continuous one. In section 4.4, the scaling of the mixing time of the continuous process as a function of model parameters is determined. The arguments leading to such precise controls on the mixing time relies on sharp controls of expectation of particular hitting times of the continuous-space processes, as well as concentration of additive functionals of the velocity processes. Finally, for the sake of completeness, the explicit formula for the invariant measure of the continuous process is derived in the appendix, directly for CITP and CFTP, following [HGM23].

4.2 PDMP formulation of the continuous-space processes

Piecewise deterministic Markov processes follow a deterministic dynamic interspersed with random jumps. Their strength lies in the wide range of processes that fall into this category and their straightforward construction. The complete characterization of their generator and its domain is known (even if sometimes in an intrinsic way), which greatly facilitates the search for their invariant measures. Furthermore, the formalism conveniently includes a way of explicitly computing the expected value and the Laplace transform of hitting times by solving systems of differential equations.

The construction of the PDMP as in [Dav93] is followed here. Let first precise the state space and deterministic dynamics. Set $\Sigma = \{-1, 1\}^2$ (resp. $\Sigma = \{-1, 0, 1\}^2$) as well as

$$\Sigma_+ = \{(\sigma^1, \sigma^2) \in \Sigma : \sigma^2 - \sigma^1 \geq 0\} \text{ and } \Sigma_- = \{(\sigma^1, \sigma^2) \in \Sigma : \sigma^2 - \sigma^1 \leq 0\}.$$

Further define $\partial K = \{*_\sigma^\ell : \sigma \in \Sigma_+\} \cup \{*_\sigma^0 : \sigma \in \Sigma_-\}$ and $K = \Sigma \cup \partial K$.

The state space is built from the family of measurable sets $(E_\nu^0)_{\nu \in K}$ defined by

$$E_\sigma^0 = (0, \ell) \times \{\sigma\} \text{ for } \sigma \in \Sigma, E_{*_\sigma^\ell}^0 = \{\ell\} \times \{*_\sigma^\ell\} \text{ for } \sigma \in \Sigma_+, E_{*_\sigma^0}^0 = \{0\} \times \{*_\sigma^0\} \text{ for } \sigma \in \Sigma_-.$$

The elements of E_ν^0 are always denoted $(x, \nu) \in [0, \ell] \times K$ (and also the elements of E_ν which are defined).

Also consider the family of vector fields $(\mathfrak{X}_\nu)_{\nu \in K}$ where

- for $\sigma \in \Sigma$ the vector field \mathfrak{X}_σ defined on E_σ^0 corresponds to the ODE $\partial_t x = \sigma^2 - \sigma^1$,
- for $\nu \in \partial K$ the vector field \mathfrak{X}_ν defined on E_ν^0 is null.

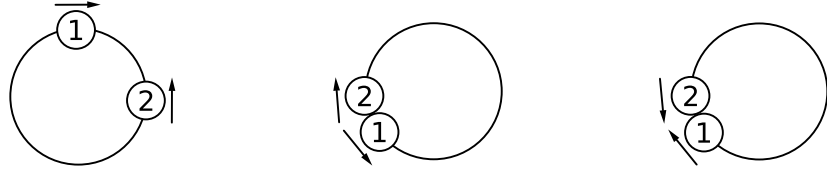
One can now add the entry-non-exit points to get the family of measurable sets $(E_\nu)_{\nu \in K}$ defined by

- $E_{(\sigma^1, \sigma^2)} = [0, \ell] \times \{(\sigma^1, \sigma^2)\}$ for $(\sigma^1, \sigma^2) \in \Sigma$ such that $\sigma^2 - \sigma^1 > 0$,

- $E_{(\sigma^1, \sigma^2)} = (0, \ell) \times \{(\sigma^1, \sigma^2)\}$ for $(\sigma^1, \sigma^2) \in \Sigma$ such that $\sigma^2 - \sigma^1 = 0$,
- $E_{(\sigma^1, \sigma^2)} = (0, \ell] \times \{(\sigma^1, \sigma^2)\}$ for $(\sigma^1, \sigma^2) \in \Sigma$ such that $\sigma^2 - \sigma^1 < 0$,
- $E_\nu = E_\nu^0$ for $\nu \in \partial K$

The process takes its values in the state space $E_D = \bigcup_{\nu \in K} E_\nu$ which can be divided into a bulk and a boundary as follows

- the bulk $\bigcup_{\sigma \in \Sigma} E_\sigma$ corresponds to states where
 - $0 < x < \ell$ meaning that the particles are at positive distance (ex: figure 4.4a),
 - $x \in \{0, \ell\}$ meaning that the particles are at contact and their velocities push them apart (ex: figure 4.4b),
- the boundary $\bigcup_{\nu \in \partial K} E_\nu$ corresponds to states where $x \in \{0, \ell\}$ meaning that the particles are at contact and their velocities push them together so they jam (ex: figure 4.4c).



(a) $0 < x < \ell$ and $\nu = (-1, 1)$ (b) $x = \ell$ and $\nu = (1, -1)$ (c) $x = 0$ and $\nu = *_{(1, -1)}^0$

Figure 4.4: Bulk and boundary states of the continuous process

The x component is always the distance from the first to the second particle whereas ν tells us both the value of the velocities σ^i and whether or not the particles are jammed. In the bulk, the deterministic dynamics correspond to the ODE $\partial_t x = \sigma^2 - \sigma^1$. In the jammed states of the boundary, the distance x is constant.

Let us now define the stochastic jumps. Let $\mathcal{Q} = (q_{\sigma, \sigma'})_{\sigma, \sigma' \in \Sigma}$ be the discrete generator of the couple (σ^1, σ^2) where σ^1 and σ^2 are independent Markov jump processes with the transition rates of figure 4.2a (resp. 4.2b).

For example, if the σ^i follow figure 4.2a then

$$\mathcal{Q} = \begin{pmatrix} -2\omega & \omega & \omega & 0 \\ \omega & -2\omega & 0 & \omega \\ \omega & 0 & -2\omega & \omega \\ 0 & \omega & \omega & -2\omega \end{pmatrix}.$$

Define the jump rate λ by $\lambda(x, \sigma) = \lambda(x, *_{\sigma}^0) = \lambda(x, *_{\sigma}^{\ell}) = -q_{\sigma, \sigma}$ and the jump kernel on E_D by

$$Q((x, \sigma); \cdot) = Q((x, *_{\sigma}^0); \cdot) = Q((x, *_{\sigma}^{\ell}); \cdot) = \sum_{\sigma' \neq \sigma} -\frac{q_{\sigma, \sigma'}}{q_{\sigma, \sigma}} \sum_{\nu = \sigma', *_{\sigma'}^0, *_{\sigma'}^{\ell}} 1_{\{(x, \nu) \in E\}} \delta_{(x, \nu)}.$$

It is important to notice here that exactly one of the elements of $\{(x, \sigma'); (x, *_{\sigma'}^0); (x, *_{\sigma'}^{\ell})\}$ in the state space E_D is such that the jump kernel on E_D is indeed a collection of probability measures.

Recalling that the states (x, σ^1, σ^2) , $(0, *_{(\sigma^1, \sigma^2)}^0)$ and $(\ell, *_{(\sigma^1, \sigma^2)}^{\ell})$ all correspond to the first particle having velocity σ^1 and the second particle having velocity σ^2 , one notices that the jump rate and jump kernels just defined capture the fact that the velocities of the

particles are independent Markov jump processes with the transition rates of figure 4.2a (resp. 4.2b).

It remains to define the boundary jumps. On the boundary $\Gamma^* = \{(0, \sigma) : \sigma \in \Sigma_- - \Sigma_+\} \cup \{(\ell, \sigma) : \sigma \in \Sigma_+ - \Sigma_-\}$ the jump kernel is given by

$$Q((0, \sigma); \cdot) = \delta_{(0, *^0_\sigma)} \text{ for } \sigma \in \Sigma_- - \Sigma_+ \text{ and } Q((\ell, \sigma); \cdot) = \delta_{(\ell, *^\ell_\sigma)} \text{ for } \sigma \in \Sigma_+ - \Sigma_-.$$

This captures the fact that when the deterministic dynamics leads us to a state where the particles are jammed, the process jumps to the corresponding boundary state.

To emphasize the difference with definition 4.2, denote this process $X_D(t) = (x_D(t), \nu(t))$.

Many important properties of PDMPs, such as the characterization of their generator, are true under the set of ‘standard conditions’ (24.8) of [Dav93]. The continuous process clearly satisfies conditions (1)–(3) and condition (4) follows from proposition (24.6) of [Dav93]. Theorem (26.14) of [Dav93] hence yields the following characterization of the generator.

Proposition 4.4 (Generator of the continuous process). *Let \mathcal{L} be the generator of the continuous process. Its domain $D(\mathcal{L})$ is the set of bounded measurable functions $f : E_D \rightarrow \mathbb{R}$ such that*

- for $\sigma \in \Sigma_+ - \Sigma_-$ the function $f(\cdot, \sigma)$ is absolutely continuous on $[0, \ell]$ and

$$\lim_{x \rightarrow \ell} f(x, \sigma) = f(\ell, *^\ell_\sigma),$$

- for $\sigma \in \Sigma_- - \Sigma_+$ the function $f(\cdot, \sigma)$ is absolutely continuous on $(0, \ell]$ and

$$\lim_{x \rightarrow 0} f(x, \sigma) = f(0, *^0_\sigma).$$

For $f \in D(\mathcal{L})$ one has that

- in the bulk $\cup_{\sigma \in \Sigma} E_\sigma$ the generator comprises a transport term and a jump term

$$\mathcal{L}f(x, \sigma) = \underbrace{(\sigma^2 - \sigma^1) \partial_x f(x, \sigma)}_{\text{transport term}} + \lambda(x, \nu) \underbrace{\int_E (f(\tilde{x}, \tilde{\nu}) - f(x, \sigma)) dQ((x, \sigma); (\tilde{x}, \tilde{\nu}))}_{\text{jump term}},$$

- whereas for boundary states $(x, \nu) \in \cup_{\nu \in \partial K} E_\nu$ the generator is given only by the jump term

$$\mathcal{L}f(x, \nu) = \lambda(x, \nu) \int_E (f(\tilde{x}, \tilde{\nu}) - f(x, \nu)) dQ((x, \nu); (\tilde{x}, \tilde{\nu}))$$

Next, the link between Definition 4.2 and the PDMP construction is established to be able to use them both.

Definition 4.5. Define $\iota : E_D \rightarrow [0, \ell] \times \Sigma$ by

$$\iota(x, \sigma) = \iota(x, *^0_\sigma) = \iota(x, *^\ell_\sigma) = (x, \sigma)$$

and define p_{σ^1} and p_{σ^2} by

$$\begin{aligned} p_{\sigma^1}(x, \sigma^1, \sigma^2) &= p_{\sigma^1}(x, *^0_{(\sigma^1, \sigma^2)}) = p_{\sigma^1}(x, *^\ell_{(\sigma^1, \sigma^2)}) = \sigma^1, \\ p_{\sigma^2}(x, \sigma^1, \sigma^2) &= p_{\sigma^2}(x, *^0_{(\sigma^1, \sigma^2)}) = p_{\sigma^2}(x, *^\ell_{(\sigma^1, \sigma^2)}) = \sigma^2. \end{aligned}$$

Proposition 4.6. *The mapping ι is a bijection.*

Proof. This follows from the fact that for all $x \in [0, \ell]$ and all $\sigma \in \Sigma$ exactly one element of the set $\{(x, \sigma); (x, *^0_\sigma); (x, *^\ell_\sigma)\}$ is an element of E_D .

Indeed

- if $x \in (0, \ell)$ then $(x, \sigma) \in E_D$ and $\{(x, *^0_\sigma); (x, *^\ell_\sigma)\} \cap E_D = \emptyset$,

- if $x = 0$ then
 - if $\sigma^2 - \sigma^1 \geq 0$ then $(0, \sigma) \in E_D$ and $\{(0, *_{\sigma}^0); (0, *_{\sigma}^{\ell})\} \cap E_D = \emptyset$,
 - if $\sigma^2 - \sigma^1 < 0$ then $(0, *_{\sigma}^0) \in E_D$ and $\{(0, \sigma); (0, *_{\sigma}^{\ell})\} \cap E_D = \emptyset$,
- if $x = \ell$
 - if $\sigma^2 - \sigma^1 \leq 0$ then $(0, \sigma) \in E_D$ and $\{(0, *_{\sigma}^0); (0, *_{\sigma}^{\ell})\} \cap E_D = \emptyset$,
 - if $\sigma^2 - \sigma^1 > 0$ then $(0, *_{\sigma}^{\ell}) \in E_D$ and $\{(0, \sigma); (0, *_{\sigma}^0)\} \cap E_D = \emptyset$.

□

Proposition 4.7. *Let $X_D(t) = (x_D(t), \nu(t))$ be as in Section 4.2.*

(i) *One has that $\sigma^1(t) = p_{\sigma^1}(X_D(t))$ and $\sigma^2(t) = p_{\sigma^2}(X_D(t))$ are two independent Markov jump processes with the transition rates of figure 4.2a (resp. 4.2b).*

(ii) *If one recursively defines*

$$x(t) = \max [0, \min [\ell, x(T_n) + (\sigma^2(t) - \sigma^1(t))(t - T_n)]] \text{ for } T_n \leq t \leq T_{n+1}$$

where $0 = T_0 < T_1 < \dots$ are the jump times of the couple (σ^1, σ^2) then $\iota(X_D(t)) = (x(t), \sigma^1(t), \sigma^2(t))$.

Proof. (i) For all fixed $(s^1, s^2) \in \Sigma$ the function $f(x, \nu) = \mathbb{P}_{(x, \nu)}(X_D(t) \in p_{\sigma^1}^{-1}(s^1) \cap p_{\sigma^2}^{-1}(s^2))$ can be computed explicitly using its characterization as the solution of a system of differential equations (see section 3 of [Dav93]). One gets

$$\mathbb{P}_{(x, \nu)}(X_D(t) \in p_{\sigma^1}^{-1}(s_1) \cap p_{\sigma^2}^{-1}(s_2)) = \mathbb{P}_{p_{\sigma^1}(x, \nu)}(\varsigma_1(t) = s_1) \mathbb{P}_{p_{\sigma^2}(x, \nu)}(\varsigma_2(t) = s_2)$$

for all $(s_1, s_2) \in \Sigma$ where the ς_i are Markov jump processes with the transition rates of figure 4.2a (resp. 4.2b).

(ii) Follows from the recursive construction in section 23 of [Dav93]. □

4.3 From discrete to continuous space

4.3.1 Convergence in law

In this section, it is shown that under the right rescaling the discrete process of [SEB16, SEB17], i.e. DITP (resp. DFTP), converges to the continuous process CITP (resp. CFTP).

First, the discrete and continuous process have different state spaces. The continuous process X takes its values in $[0, \ell] \times \Sigma$ and one can think of $\{1, \dots, L\}$ as a discrete subset of the interval $[0, \ell]$ using the injection

$$i_L(k) := \ell \frac{k-1}{L-1}.$$

This allows to see the discrete process as taking values in $[0, \ell] \times \Sigma$ as well by considering the process $(i_L(y^L(t)), \sigma^1(t), \sigma^2(t))$ instead of $(y^L(t), \sigma^1(t), \sigma^2(t))$.

It remains to understand how the parameters ω (resp. α and β), γ and L should scale. Thinking of the discrete process as an approximation of the continuous process where the interval $[0, \ell]$ is replaced by its discrete counterpart $\{1, \dots, L\}$, the natural scalings

$$(*) \left\{ \begin{array}{l} \omega \text{ (resp. } \alpha \text{ and } \beta) \text{ constant} \\ L \rightarrow +\infty \\ \gamma_L = (L-1)/\ell \end{array} \right.$$

correspond to taking increasingly fine discretisations of the position $y^L(t)$ while preserving the dynamics of the σ^i and the physical velocity of the particles $\ell\gamma/(L-1)$.

The convergence of Markov processes is often shown using approaches based on generators or on martingale problems (see [EK86]). This article takes a different approach and instead uses a coupling argument that requires only basic probabilistic tools. It also provides quantitative estimates for the approximation procedure. The key insight is that, under the scalings described above, the stochastic jumps of $y^L(t)$ become both smaller in size and more frequent in a way that makes them converge to the constant velocity translations of $x(t)$.

Theorem 4.8 (Scaling limit). *(i) There exists a coupling of the continuous process X and a sequence of discrete processes $(Y^L)_{L \geq 2}$ with parameters scaling as in (*) such that for all $T > 0$*

$$\mathbb{P} \left(\sup_{t \leq T} |i_L(y^L(t)) - x(t)| \geq \epsilon \right) \leq \frac{1}{\epsilon} \left(\frac{\ell}{L-1} + 8\sqrt{\frac{T((\eta T)^2 + 3\eta T + 1)\ell}{L-1}} \right)$$

where $\eta = 2\omega$ or $\eta = 2 \max(\alpha, \beta)$ depending on the transition rates of the σ^i .

(ii) One has

$$\left((i_L(y^L(t)), \sigma^1(t), \sigma^2(t)) \right)_{t \geq 0} \xrightarrow{\text{Law}} \left(X(t) \right)_{t \geq 0}$$

in the Skorokhod space $\mathcal{D}[0, +\infty)$.

The supremum distance is finer than the Skorokhod distance so assertion (i) implies that for all $T > 0$

$$\left((i_L(y^L(t)), \sigma^1(t), \sigma^2(t)) \right)_{t \in [0, T]} \xrightarrow{\text{Law}} \left(X(t) \right)_{t \in [0, T]}$$

in the sense of random variables taking their values in the Skorokhod space $\mathcal{D}[0, T]$. Hence assertion (ii) follows by theorem 16.7 of [Bil99].

The rest of this subsection is dedicated to the proof of assertion (i). First, consider a toy problem that proves useful in its own right.

Lemma 4.9. *Let σ^1 and σ^2 be two independent Markov jump processes with the transition rates of figure 4.2a (resp. 4.2b) and let N_1^L and N_2^L be two independent Poisson processes with rate γ_L independent of the σ^i .*

If one sets

$$S^L(t) = \sum_{s \leq t} \sum_{i=1,2} 1_{\{N_i^L(s+) \neq N_i^L(s-)\}} (-1)^i \sigma^i(s).$$

then one has

$$\mathbb{P} \left(\sup_{t \leq T} \left| \frac{1}{\gamma_L} S^L(t) - \int_0^t (\sigma^2(s) - \sigma^1(s)) ds \right| \geq \epsilon \right) \leq \frac{2}{\epsilon} \sqrt{\frac{T}{\gamma_L}}.$$

The proof is included for the sake of completeness.

Proof. First notice that $(S^L(t), \sigma^1(t), \sigma^2(t))$ is a $(\mathbb{Z} \times \Sigma)$ -valued Markov jump process and its generator \mathcal{L}^L is given by

$$\begin{aligned} \mathcal{L}^L f(z, \sigma^1, \sigma^2) &= \gamma_L f(z - \sigma^1, \sigma^1, \sigma^2) + \gamma_L f(z + \sigma^2, \sigma^1, \sigma^2) - 2\gamma_L f(z, \sigma^1, \sigma^2) \\ &+ \sum_{(s^1, s^2) \in \Sigma} q_{(\sigma^1, \sigma^2), (s^1, s^2)} f(z, s^1, s^2) \end{aligned}$$

where $(q_{\sigma, \sigma'})_{\sigma, \sigma' \in \Sigma}$ is the generator of the couple (σ^1, σ^2) . Hence Dynkin's formula implies that $S^L(t) - \gamma_L \int_0^t (\sigma^2(s) - \sigma^1(s)) ds$ is a martingale and by Doob's martingale inequality

$$\begin{aligned} & \mathbb{P} \left(\sup_{t \leq T} \left| \frac{1}{\gamma_L} S^L(t) - \int_0^t (\sigma^2(s) - \sigma^1(s)) ds \right| \geq \epsilon \right) \\ & \leq \frac{\mathbb{E} \left| S^L(T) - \gamma_L \int_0^T (\sigma^2(s) - \sigma^1(s)) ds \right|}{\epsilon \gamma_L}, \\ & \leq \frac{\sum_{i=1,2} \mathbb{E} \left| \sum_{s \leq t} 1_{\{N_i^L(s+) \neq N_i^L(s-)\}} \sigma^i(s) - \gamma_L \int_0^T \sigma^i(s) ds \right|}{\epsilon \gamma_L}, \\ & \leq \frac{2}{\epsilon \gamma_L} \sqrt{\mathbb{E} \left(\sum_{s \leq t} 1_{\{N_1^L(s+) \neq N_1^L(s-)\}} \sigma^1(s) - \gamma_L \int_0^T \sigma^1(s) ds \right)^2}. \end{aligned}$$

It remains to show that $\mathbb{E} \left(\sum_{s \leq t} 1_{\{N_1^L(s+) \neq N_1^L(s-)\}} \sigma^1(s) - \gamma_L \int_0^T \sigma^1(s) ds \right)^2 \leq T \gamma_L$. Let $(T_n)_{n \geq 0}$ be the jump times of the couple (σ^1, σ^2) . The process

$$M(t) = \sum_{s \leq t} 1_{\{N_1^L(s+) \neq N_1^L(s-)\}} \sigma^1(s) - \gamma_L \int_0^t \sigma^1(s) ds$$

is also a martingale by Dynkin's formula so that repeated use of the optional stopping theorem yields

$$\mathbb{E} [M(T)^2] = \sum_{k=1}^N \mathbb{E} [(M(T_k \wedge T) - M(T_{k-1} \wedge T))^2] + \mathbb{E} [(M(T) - M(T_N \wedge T))^2].$$

Taking $N \rightarrow +\infty$ and using the dominated convergence theorem one gets

$$\mathbb{E} [(M(T) - M(T \wedge T_N))^2] \rightarrow 0$$

and $\mathbb{E} [M(T)^2] = \sum_{k=1}^{+\infty} \mathbb{E} [(M(T_k \wedge T) - M(T_{k-1} \wedge T))^2]$ so that

$$\begin{aligned} \mathbb{E} [M(T)^2] &= \sum_{k=1}^{+\infty} \mathbb{E} [\sigma^1(T_{k-1} \wedge T)^2 (N_1^L(T_k \wedge T) - N_1^L(T_{k-1} \wedge T) - \gamma_L(T_k \wedge T - T_{k-1} \wedge T))^2], \\ &\leq \sum_{k=1}^{+\infty} \mathbb{E} \left[(N_1^L(T_k \wedge T) - N_1^L(T_{k-1} \wedge T) - \gamma_L(T_k \wedge T - T_{k-1} \wedge T))^2 \right] \\ &= \mathbb{E} \left[(N_1^L(T) - \gamma_L T)^2 \right] = \gamma_L T. \end{aligned}$$

□

It is easy to see that one can also obtain directly an upper bound in the previous lemma such as $8T/(\epsilon^2 \gamma_L)$ (in fact even exponential inequality can be derived), deteriorating in T with however a better behavior in γ_L .

Let us assume $\gamma_L \rightarrow +\infty$. The jumps of S^L/γ_L then have decreasing size $1/\gamma_L$ and occur with increasing rate γ_L so that the physical velocity $\gamma_L/\gamma_L = 1$ remains constant. Lemma 4.9 implies that the stochastic jumps of S^L/γ_L converge to the constant velocity translations of $\int_0^t (\sigma^2(s) - \sigma^1(s)) ds$ and, crucially, the convergence does not take place for a single t but for the supremum all $t \leq T$ thanks to Doob's martingale inequality. Notice that the only difference between the S^L of the toy problem and the y^L of the discrete process is that S^L is neither jammed at 0 nor L so it takes its values in \mathbb{Z} instead of $\{1, \dots, L\}$. Assertion (i)

of theorem 4.8 can thus be seen as a more involved analog of lemma 4.9. Armed with the understanding of the toy problem, the convergence of the discrete process to the continuous process. may now be easier to handle. The key is the following coupling.

Definition 4.10 (Discrete-continuous coupling). *Let $X(t) = (x(t), \sigma^1(t), \sigma^2(t))$ be a continuous process with fixed parameters ω (resp. α and β) and ℓ . Use $\sigma^1(t) = p_{\sigma^1}(X(t))$ and $\sigma^2(t) = p_{\sigma^2}(X(t))$ to construct a sequence of stochastic processes $(Y^L)_{L \geq 2}$ which later turns out to be DITPs (resp. DFTPs). Recall that proposition 4.7 implies that the σ^i are independent Markov jump processes with the transition rates of figure 4.2a (resp. 4.2b). Denote by $0 = T_0 < T_1 < \dots$ the jump times of the couple (σ^1, σ^2) . For $L \geq 2$, let N_1^L and N_2^L be two independent Poisson processes with parameter $\gamma_L = (L-1)/\ell$ independent of the σ^i . Recursively define y^L by setting $y^L(0) = \lfloor (L-1)x(0)/\ell \rfloor + 1$ and*

- if $(\sigma^1(T_k), \sigma^2(T_k)) \neq \pm(1, 1)$

$$y^L(t) = \max \left[0, \min \left[L, y^L(T_k) + \sum_{i=1,2} (-1)^i \sigma^i(T_k) [N_i^L(t) - N_i^L(T_k)] \right] \right]$$

for $t \in [T_k, T_{k+1}]$,

- if $(\sigma^1(T_k), \sigma^2(T_k)) = \pm(1, 1)$

$$y^L(t) = p_L \left(y^L(T_k) + \sum_{i=1,2} (-1)^i \sigma^i(T_k) [N_i^L(t) - N_i^L(T_k)] \right) \text{ for } t \in [T_k, T_{k+1}]$$

where $p_L : \mathbb{Z} \rightarrow \{1, \dots, L\}$ is given by

$$p_L(k + 2mL) = \begin{cases} k & \text{for } m \in \mathbb{Z} \text{ and } 1 \leq k \leq L, \\ 2L + 1 - k & \text{for } m \in \mathbb{Z} \text{ and } L + 1 \leq k \leq 2L. \end{cases}$$

Finally, set $Y^L(t) = (y^L(t), \sigma^1(t), \sigma^2(t))$.

Let first prove that this definition indeed provides a coupling, i.e. Y^L alone is indeed a DITP (or FTP).

Lemma 4.11. *Let X and $(Y^L)_{L \geq 2}$ be as in definition 4.10.*

(i) *Each Y^L is a DITP (resp. DFTP).*

(ii) *Let $0 = T_0 < T_1 < \dots$ be the jump times of the couple (σ^1, σ^2) . For all $t \geq 0$ one has*

$$|i_L(y^L(t)) - x(t)| \leq |i_L(y^L(0)) - x(0)| + \frac{1}{\gamma_L} \sum_{k=1}^{n(t)} |D_k(t)|$$

where $n(t) = \sum_{k=0}^{+\infty} 1_{\{T_k \leq t\}}$ and

$$D_k(t) = \sum_{i=1,2} (-1)^i \sigma^i(T_{k-1}) \left[N_i^L(T_k \wedge t) - N_i^L(T_{k-1} \wedge t) - \gamma_L (T_k \wedge t - T_{k-1} \wedge t) \right].$$

It would seem more natural in definition 4.10 to construct Y^L like in definition 4.1 instead of using p_L . Although this would allow us to avoid the proof of assertion (i) of lemma 4.11, this alternative construction is necessary for the bound of assertion (ii) to hold. The main obstacle to showing the same kind of uniform bound as in proposition 4.9 is the fact that $(i_L(y^L(t)) - x(t))$ is not a martingale so one cannot apply Doob's martingale inequality. Fortunately, assertion (ii) allows us to bound $\sup_{t \leq T} |i_L(y^L(t)) - x(t)|$ in terms

of $\sup_{t \leq T} |N_i^L(t) - \gamma_L t|$. Hence, because $(N_i^L(t) - \gamma_L t)$ is a martingale, a uniform bound similar to proposition 4.9 is attainable.

Proof of lemma 4.11. (i) Let $L \geq 2$ be arbitrary but fixed. Let Y^L be as in definition 4.10 and let σ^i and N_i^L be the corresponding velocities and Poisson processes. Set $\tilde{Y}^L = \mathcal{Y}(\sigma^1, \sigma^2, N_1^L, N_2^L, L)$.

To show that Y^L and \tilde{Y}^L have the same law, it suffices to show that y^L and \tilde{y}^L are identically distributed conditional on the velocities (σ^1, σ^2) . Let $s : [0, +\infty) \rightarrow \Sigma$ be an arbitrary but fixed càdlàg function. In the remainder of the proof of (i), take the conditional probability $\mathbb{Q} = \mathbb{P}(\cdot | (\sigma^1, \sigma^2) = s)$ to be the reference probability so that functionals of the σ^i such as jump times are deterministic. Under the probability \mathbb{Q} both y^L and \tilde{y}^L are time-inhomogeneous Markov processes with the same initial distribution so that it suffices to prove that their transition probabilities are the same.

Recall that $0 = T_0 < T_1 < \dots$ are the jump times of the couple (σ^1, σ^2) . For $t \in [T_k, T_{k+1}]$

- if $(\sigma^1(T_k), \sigma^2(T_k)) \neq \pm(1, 1)$ then

$$y_L(t) = \max \left[0, \min \left[L, y^L(T_k) + \sum_{i=1,2} (-1)^i \sigma^i(T_k) (N_i^L(t) - N_i^L(T_k)) \right] \right],$$

$$\tilde{y}_L(t) = \max \left[0, \min \left[L, \tilde{y}^L(T_k) + \sum_{i=1,2} (-1)^i \sigma^i(T_k) (N_i^L(t) - N_i^L(T_k)) \right] \right],$$

so that the transition rates are the same on the time interval $[T_k, T_{k+1}]$,

- if $(\sigma^1(T_k), \sigma^2(T_k)) = \pm(1, 1)$ then p_L transform the symmetric random walk with rate γ_L on \mathbb{Z} into a Markov jump process with the same transition rates as \tilde{y}^L so that the transition probabilities are again the same on the time interval $[T_k, T_{k+1}]$.

(ii) The velocities σ^i are constant on the time interval $[T_{n(t)}, t]$.

- If $\sigma^1(T_{n(t)}) \neq \sigma^2(T_{n(t)})$ one has

$$\begin{aligned} & |i_L(y^L(t)) - x(t)| \\ &= \left| \max \left[0, \min \left[\ell, i_L(y^L(T_{n(t)})) + \frac{1}{\gamma_L} \sum_{i=1,2} (-1)^i \sigma^i(T_{n(t)}) (N_i^L(t) - N_i^L(T_{n(t)})) \right] \right] \right. \\ &\quad \left. - \max \left[0, \min \left[\ell, x(T_{n(t)}) + (\sigma^2(T_{n(t)}) - \sigma^1(T_{n(t)}))(t - T_{n(t)}) \right] \right] \right|, \\ &\leq |i_L(y^L(t)) - x(T_{n(t)})| \\ &\quad + \frac{1}{\gamma_L} \left| \sum_{i=1,2} (-1)^i \sigma^i(T_{n(t)}) (N_i^L(t) - N_i^L(T_{n(t)}) - \gamma_L(t - T_{n(t)})) \right|. \end{aligned}$$

- If $\sigma^1(T_{n(t)}) = \sigma^2(T_{n(t)}) = 0$ one has $|i_L(y^L(t)) - x(t)| = |i_L(y^L(T_{n(t)})) - x(T_{n(t)})|$.
- If $\sigma^1(T_{n(t)}) = \sigma^2(T_{n(t)}) = \pm 1$, because p_L is 1-Lipschitz and i_L is γ_L^{-1} -Lipschitz, one

has

$$\begin{aligned}
 & |i_L(y^L(t)) - x(t)| \\
 & \leq |i_L(y^L(T_{n(t)})) - x(T_{n(t)})| + |i_L(y^L(t)) - i_L(y^L(T_{n(t)}))|, \\
 & = |i_L(y^L(T_{n(t)})) - x(T_{n(t)})| \\
 & \quad + \frac{1}{\gamma_L} \left| p_L \left(y^L(T_{n(t)}) + \sum_{i=1,2} (-1)^i \sigma^i(T_{n(t)}) (N_i^L(t) - N_i^L(T_{n(t)})) \right) - p_L(y^L(T_{n(t)})) \right|, \\
 & \leq |i_L(y^L(T_{n(t)})) - x(T_{n(t)})| \\
 & \quad + \frac{1}{\gamma_L} \left| \sum_{i=1,2} (-1)^i \sigma^i(T_{n(t)}) (N_i^L(t) - N_i^L(T_{n(t)}) - \gamma_L(t - T_{n(t)})) \right|.
 \end{aligned}$$

So in any case

$$|i_L(y^L(t)) - x(t)| \leq |i_L(y^L(T_{n(t)})) - x(T_{n(t)})| + \frac{1}{\gamma_L} |D_{n(t)}(t)|$$

and iterating the same argument yields the desired bound. \square

The proof of theorem 4.8 (i) is now possible. The fact that the upper bound (ii) of lemma 4.11 contains the term $\sum |D_k|$ instead of $|\sum D_k|$ leads to a slightly worse bound compared to the toy problem. In particular, unlike for the toy problem, the bound of theorem 4.8 (i) depends on ω (resp. α and β).

Proof of theorem 4.8 (i). Let X and $(Y^L)_{L \geq 2}$ be as in definition 4.10. Keeping the same notation as in lemma 4.11 one has for $t \leq T$

$$\begin{aligned}
 |D_k(t)| & = \left| \sum_{i=1,2} (-1)^i \sigma^i(T_k) \left[N_i^L(T_k \wedge t) - N_i^L(T_{k-1} \wedge t) - \gamma_L(T_k \wedge t - T_{k-1} \wedge t) \right] \right|, \\
 & \leq \sum_{i=1,2} (|N_i^L(T_k \wedge t) - \gamma_L(T_k \wedge t)| + |N_i^L(T_{k-1} \wedge t) - \gamma_L(T_{k-1} \wedge t)|), \\
 & \leq 2 \sum_{i=1,2} \sup_{\theta \leq T} |N_i^L(\theta) - \gamma_L \theta|.
 \end{aligned}$$

So claim (ii) of lemma 4.11 implies

$$\sup_{t \leq T} |i_L(y^L(t)) - x(t)| \leq |i_L(y^L(0)) - x(0)| + \frac{2n(T)}{\gamma_L} \sum_{i=1,2} \sup_{t \leq T} |N_i^L(t) - \gamma_L t|.$$

Using Markov's inequality first and then the Cauchy-Schwarz inequality one gets

$$\begin{aligned}
 \mathbb{P} \left(\sup_{t \leq T} |i_L(y^L(t)) - x(t)| \geq \epsilon \right) & \leq \frac{1}{\epsilon} \left(\mathbb{E} |i_L(y^L(0)) - x(0)| + \frac{4}{\gamma_L} \mathbb{E} \left[n(T) \sup_{t \leq T} |N_1^L(t) - \gamma_L t| \right] \right), \\
 & \leq \frac{1}{\epsilon} \left(\frac{\ell}{L-1} + \frac{4}{\gamma_L} \sqrt{\mathbb{E} [n(T)^2]} \sqrt{\mathbb{E} \left[\sup_{t \leq T} (N_1^L(t) - \gamma_L t)^2 \right]} \right).
 \end{aligned}$$

It follows from Doob's maximal inequality that

$$\mathbb{E} \left[\sup_{t \leq T} (N_1^L(t) - \gamma_L t)^2 \right] \leq 4 \mathbb{E} [(N_1^L(T) - \gamma_L T)^2] = 4\gamma_L T.$$

Finally, a coupling argument shows that one has the stochastic domination

$$\mathbb{P}(n(T) \geq x) \leq \mathbb{P}(1 + \tilde{n} \geq x) \text{ for all } x \geq 0,$$

where \tilde{n} is a Poisson random variable with parameter ηT . Hence

$$\mathbb{E} [n(T)^2] \leq \mathbb{E} [(1 + \tilde{n})^2] = \mathbb{E} [\tilde{n}^2] + 2\mathbb{E} [\tilde{n}] + 1 = (\eta T)^2 + 3\eta T + 1$$

and the desired result follows. \square

4.3.2 Convergence of the discrete invariant measures

Let X and $(Y^L)_{L \geq 2}$ be as in definition 4.10. Because Y^L is irreducible and has finite state space, it has a unique invariant measure π_L and there exist $C_L, \rho_L > 0$ such that

$$\sup_{\mu_L} \|\mu_L P_t^L - \pi_L\|_{\text{TV}} \leq C_L e^{-\rho_L t} \text{ for all } t \geq 0$$

where (P_t^L) is the semigroup associated to Y^L . Similarly, the continuous process satisfies the Doeblin property ensuring the existence of a unique invariant measure as well as exponential convergence towards that measure. However, the bounds on the speed of convergence towards the invariant measure are quantitatively poor.

Proposition 4.12 (Doeblin). *Let X_D , with semigroup (P_t) , be as in Section 4.2 with fixed parameters ω (resp. α and β) and ℓ . One has the existence of a positive a such that*

$$\inf_{(x, \nu) \in E_D} \mathbb{P}_{(x, \nu)} \left(X_{1+\ell/2} = \left(0, *_{(1, -1)}^0 \right) \right) > a.$$

In particular, X_D admits a unique invariant probability π and there exist $C, \rho > 0$ such that

$$\sup_{\mu} \|\mu P_t - \pi\|_{\text{TV}} \leq C e^{-\rho t} \text{ for all } t \geq 0.$$

Proof. Using the Markov property, one gets

$$\begin{aligned} & \mathbb{P}_{(x, \nu)} \left(X_D(1 + \ell/2) = \left(0, *_{(1, -1)}^0 \right) \right) \\ & \geq \mathbb{P}_{(x, \nu)} \left(X_D(1 + \ell/2) = \left(0, *_{(1, -1)}^0 \right) \text{ and } X_D(1) \in E_{(1, -1)} \cup E_{*_{(1, -1)}^0} \right), \\ & \geq \left(\inf_{(x, \nu) \in E_{(1, -1)} \cup E_{*_{(1, -1)}^0}} \mathbb{P}_{(x, \nu)} \left(X_D(\ell/2) = \left(0, *_{(1, -1)}^0 \right) \right) \right) \\ & \quad \times \left(\inf_{(x, \nu) \in E_D} \mathbb{P}_{(x, \nu)} \left(X_D(1) \in E_{(1, -1)} \cup E_{*_{(1, -1)}^0} \right) \right). \end{aligned}$$

Because the velocity process (σ^1, σ^2) is irreducible, one has

$$\inf_{(x, \nu) \in E_D} \mathbb{P}_{(x, \nu)} \left(X_D(1) \in E_{(1, -1)} \cup E_{*_{(1, -1)}^0} \right) = \inf_{\sigma \in \Sigma} \mathbb{P}_{\sigma} (\sigma(1) = (1, -1)) > 0.$$

Starting from any initial condition $(x, \nu) \in E_{(1, -1)} \cup E_{*_{(1, -1)}^0}$ one has

$$\{\text{there is no jump before time } t = \ell/2\} \subset \{X_D(\ell/2) = *_{(1, -1)}^0\}$$

so that

$$\inf_{(x, \nu) \in E_{(1, -1)} \cup E_{*_{(1, -1)}^0}} \mathbb{P}_{(x, \nu)} \left(X_D(\ell/2) = \left(0, *_{(1, -1)}^0 \right) \right) \geq e^{\frac{\ell}{2} q_{(1, -1), (1, -1)}} > 0.$$

\square

The convergence, under an appropriate rescaling, of Y^L towards X thus invites the question: does π_L converge to π ? This amounts to interchanging the $L \rightarrow +\infty$ and the $t \rightarrow +\infty$ limits and was a crucial implicit assumptions in [SEB16, SEB17]. This kind of limit interchange requires some sort of uniformity result, which in this case takes the form of proposition 4.13.

Proposition 4.13. *There exists $C, \rho > 0$ such that*

$$\sup_L \sup_{\mu_L} \|\mu_L P_t^L - \pi_L\|_{\text{TV}} \leq C e^{-\rho t} \text{ for all } t \geq 0.$$

The following proposition is a direct consequence of proposition 4.13.

Corollary 4.14. *One has $W_1(i_L \# \pi_L, \pi) \rightarrow 0$.*

Proof. Theorem 4.8 together with the boundedness of $E = [0, \ell] \times \Sigma$ implies that for all $t \geq 0$

$$\lim_{L \rightarrow +\infty} W_1(i_L \# ((s_L \# \pi) P_t^L), \pi P_t) = 0$$

where $s_L : E \rightarrow \{1, \dots, L\} \times \Sigma$ is given by

$$s_L(x, \sigma^1, \sigma^2) = (\lfloor (L-1)x/\ell \rfloor + 1, \sigma^1, \sigma^2).$$

Because $\|X - X'\| \leq \sqrt{4^2 + 4^2 + \ell^2} =: A$ for all $X, X' \in [0, \ell] \times \Sigma$ one has $W_1(\mu, \nu) \leq A \|\mu - \nu\|_{\text{TV}}$ for all measure μ, ν on $[0, \ell] \times \Sigma$. One gets for all $t \geq 0$

$$\begin{aligned} W_1(i_L \# \pi_L, \pi) &\leq W_1(i_L \# \pi_L, i_L \# ((s_L \# \pi) P_t^L)) + W_1(i_L \# ((s_L \# \pi) P_t^L), \pi P_t), \\ &\leq A \|i_L \# \pi_L - i_L \# ((s_L \# \pi) P_t^L)\|_{\text{TV}} + W_1(i_L \# ((s_L \# \pi) P_t^L), \pi P_t), \\ &\leq A \|\pi_L - (s_L \# \pi) P_t^L\|_{\text{TV}} + W_1(i_L \# ((s_L \# \pi) P_t^L), \pi P_t), \\ &\leq A C e^{-\rho t} + W_1(i_L \# ((s_L \# \pi) P_t^L), \pi P_t). \end{aligned}$$

Hence $\limsup_{L \rightarrow +\infty} W_1(i_L \# \pi_L, \pi) \leq A C e^{-\rho t}$. The result follows since $t \geq 0$ is arbitrary. \square

The remainder of this section is dedicated to the proof of proposition 4.13 using a coupling argument. Coupling two discrete processes $Y^L = (y^L, \sigma^1, \sigma^2)$ and $\tilde{Y}^L = (\tilde{y}^L, \tilde{\sigma}^1, \tilde{\sigma}^2)$ boils down to coupling their velocities and their Poisson clocks, which are taken to be identical and their velocities to be independent until they meet for the first time and then identical. Let us formalize this in the following definition.

Definition 4.15 (Discrete-discrete coupling). *Let $Y_0^L = (y_0^L, \sigma_0^1, \sigma_0^2)$ and $\tilde{Y}_0^L = (\tilde{y}_0^L, \tilde{\sigma}_0^1, \tilde{\sigma}_0^2)$ in $\{1, \dots, L\} \times \Sigma$ be given. Now turn to constructing a coupling between two discrete processes Y^L and \tilde{Y}^L with respective initial states Y_0^L and \tilde{Y}_0^L . Let $s^1, s^2, \tilde{s}^1, \tilde{s}^2$ be four independent Markov jump processes with the transition rates of figure 4.2a (resp. 4.2b) and respective initial states $\sigma_0^1, \sigma_0^2, \tilde{\sigma}_0^1, \tilde{\sigma}_0^2$. Define $\tau_i = \inf\{t \geq 0 : \sigma^i(t) = \tilde{\sigma}^i(t)\}$ and set*

$$\sigma^i(t) = s^i(t) \text{ for all } t \geq 0, \quad \tilde{\sigma}^i(t) = \tilde{s}^i(t) \text{ for } t \leq \tau_i \text{ and } \tilde{\sigma}^i(t) = s^i(t) \text{ for } t \geq \tau_i.$$

Let N_1^L and N_2^L be two independent Poisson clocks with parameter $\gamma_L = (L-1)/\ell$ independent of the σ^i .

Finally, set $Y^L = \mathcal{Y}(\sigma^1, \sigma^2, N_1^L, N_2^L, y_0^L, L)$ and $\tilde{Y}^L = \mathcal{Y}(\tilde{\sigma}^1, \tilde{\sigma}^2, N_1^L, N_2^L, \tilde{y}_0^L, L)$.

Remark 4.16. *With this coupling*

$$\begin{aligned} \sigma^i(t_0) = \tilde{\sigma}^i(t_0) &\implies \sigma^i(t) = \tilde{\sigma}^i(t) \text{ for all } t_0 \geq t, \\ Y^L(t_0) = \tilde{Y}^L(t_0) &\implies Y^L(t) = \tilde{Y}^L(t) \text{ for all } t_0 \geq t. \end{aligned}$$

In the rest of this section, the notations

$$S^L(t) := \sum_{s \leq t} \sum_{i=1,2} (-1)^i \sigma^i(s) 1_{\{N_i^L(s-) \neq N_i^L(s+)\}} \quad \text{and} \quad I(t) := \int_0^t (\sigma^2(s) - \sigma^1(s)) ds$$

are used.

Recall the coupling characterization of the total variation distance

$$\sup_{\mu_L} \|\mu_L P_t^L - \pi_L\|_{\text{TV}} \leq \sup_{Y_0^L, \tilde{Y}_0^L} \mathbb{P}_{Y_0^L, \tilde{Y}_0^L} \left(Y^L(t) \neq \tilde{Y}^L(t) \right).$$

Hence, it is left to understand the stopping time $\inf\{t \geq 0 : Y^L(t) = \tilde{Y}^L(t)\}$. Understanding the first time where both velocities are identical $\tau_\sigma = \max_i \tau_i$ is not a problem because it does not depend on L and so the main obstacle is understanding the time that elapses between τ_σ and $\tau_y = \inf\{t \geq \tau_\sigma : y^L(t) = \tilde{y}^L(t)\}$. The following deterministic lemma allows to do just that by linking $\tau_y - \tau_\sigma$ to the variations of $S^L(t)$.

Lemma 4.17. *Let Y^L and \tilde{Y}^L be as in definition 4.15 and assume that $\sigma_0^1 = \tilde{\sigma}_0^1$ and $\sigma_0^2 = \tilde{\sigma}_0^2$. For any fixed realization, one has*

$$|S^L(T)| \geq L - 1 \implies \inf\{t \geq 0 : y^L(t) = \tilde{y}^L(t)\} \leq T.$$

Proof. Assume without loss of generality that $y^L(0) \leq \tilde{y}^L(0)$ and $S^L(T) \geq L - 1$. Because $\sigma^1 = \tilde{\sigma}^1$ and $\sigma^2 = \tilde{\sigma}^2$ one has that $y^L(0) \leq \tilde{y}^L(0)$ implies that $y^L(t) \leq \tilde{y}^L(t)$ for all $t \geq 0$.

Assume by contradiction that $y^L(t) \neq \tilde{y}^L(t)$ for all $t \leq T$. If $y^L(t) = L$ for some $t \leq T$ then $L = y^L(t) \leq \tilde{y}^L(t) \leq L$ which would contradict $y^L(t) \neq \tilde{y}^L(t)$ for all $t \leq T$. Hence $y^L(t) < L$ for all $t \leq T$.

It can be shown by induction that $y^L(t) < L$ for all $t \leq T$ implies $y^L(t) \geq y^L(0) + S^L(t)$ for all $t \leq T$. Hence $y^L(T) \geq y^L(0) + S^L(T) \geq L$ which is a contradiction. \square

The following lemma establishes a crude quantitative result on the variations of $I(t)$ and then uses the convergence of $\frac{1}{\gamma_L} S^L(t)$ to $I(t)$ to deduce a result on the variations of $S^L(t)$ for large L .

Lemma 4.18. (i) *For all $\delta > 0$ one has $\liminf_{t \rightarrow +\infty} \inf_\sigma \mathbb{P}_\sigma (|I(t)| \geq \delta) \geq 1/3$.*

(ii) *There exist $T > 0$ and $L_0 \in \mathbb{N}$ such that for all $L \geq L_0$ one has*

$$\inf_{\sigma \in \Sigma} \mathbb{P}_\sigma (|S^L(T)| \geq L - 1) \geq \frac{1}{9}.$$

Proof. (i) For the sake of conciseness, only consider the case of the DITP. The same arguments apply mutatis mutandis to the DFTP.

Set $I_k(t) = \int_0^t \sigma^k(s) ds$ for $k = 1, 2$. By Feynman-Kac formulae

$$\begin{pmatrix} \mathbb{E}_{-1} [e^{\lambda I_k(t)}] \\ \mathbb{E}_1 [e^{\lambda I_k(t)}] \end{pmatrix} = e^{t(\mathcal{Q} + \lambda V)} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{with } \mathcal{Q} = \begin{pmatrix} -\omega & \omega \\ \omega & -\omega \end{pmatrix} \text{ and } V = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence one gets

$$E_{\pm 1} [e^{\lambda I_k(t)}] = e^{-t\omega} \left(\frac{(\omega \pm \lambda) \sinh(t\sqrt{\lambda^2 + \omega^2})}{\sqrt{\lambda^2 + \omega^2}} + \cosh(t\sqrt{\lambda^2 + \omega^2}) \right).$$

For $k = 1, 2$ and $\sigma_0^k \in \{-1, 1\}$, differentiating and taking $t \rightarrow +\infty$ yields

$$\begin{aligned} \lim_{t \rightarrow +\infty} \frac{\mathbb{E}_{\sigma_0^k} [I_k(t)]}{t^{1/2}} &= 0, & \lim_{t \rightarrow +\infty} \frac{\mathbb{E}_{\sigma_0^k} [I_k(t)^2]}{t} &= \frac{1}{\omega}, \\ \lim_{t \rightarrow +\infty} \frac{\mathbb{E}_{\sigma_0^k} [I_k(t)^3]}{t^{3/2}} &= 0, & \lim_{t \rightarrow +\infty} \frac{\mathbb{E}_{\sigma_0^k} [I_k(t)^4]}{t^2} &= \frac{3}{\omega^2}. \end{aligned}$$

Using $I(t) = I_2(t) - I_1(t)$ and the independence of $I_1(t)$ and $I_2(t)$ one sees that for all $\sigma \in \Sigma$

$$\lim_{t \rightarrow +\infty} \mathbb{E}_\sigma [I(t)^2] = +\infty \text{ and } \lim_{t \rightarrow +\infty} \frac{\mathbb{E}_\sigma [I(t)^2]^2}{\mathbb{E}_\sigma [I(t)^4]} = \frac{1}{3}.$$

Finally, when t is large enough $\frac{\delta^2}{\mathbb{E}[I(t)^2]} \in (0, 1)$ so by the Paley-Zygmund inequality

$$\mathbb{P}_\sigma (|I(t)| \geq \delta) = \mathbb{P}_\sigma \left(I(t)^2 \geq \frac{\delta^2}{\mathbb{E}_\sigma [I(t)^2]} \mathbb{E}_\sigma [I(t)^2] \right) \geq \left(1 - \frac{\delta^2}{\mathbb{E}_\sigma [I(t)^2]} \right)^2 \frac{\mathbb{E}_\sigma [I(t)^2]^2}{\mathbb{E}_\sigma [I(t)^4]}$$

and the result follows.

(ii) Assertion (i) implies that there exists $T > 0$ such that $\inf_{\sigma \in \Sigma} \mathbb{P}_\sigma (|I(T)| \geq \ell + 1) \geq \frac{2}{9}$ and by proposition 4.9 there exists $L_0 \in \mathbb{N}$ such that for all $L \geq L_0$

$$\sup_{\sigma \in \Sigma} \mathbb{P}_\sigma \left(\left| \frac{\ell}{L-1} S^L(T) - I(T) \right| \geq 1 \right) \leq \frac{1}{9}.$$

Hence for $L \geq L_0$ and $\sigma \in \Sigma$

$$\mathbb{P}_\sigma (|S^L(t)| \geq L-1) \geq \mathbb{P}_\sigma (|I(T)| \geq \ell + 1) - \mathbb{P}_\sigma \left(\left| \frac{\ell}{L-1} S^L(T) - I(T) \right| \geq 1 \right) \geq \frac{2}{9} - \frac{1}{9}.$$

□

The key uniformity result can now be proven.

Proof of proposition 4.13. For all $L \geq 2$ there exist $C_L, \rho_L > 0$ such that

$$\sup_{\mu_L} \left\| \mu_L P_t^L - \pi_L \right\|_{\text{TV}} \leq C_L e^{-\rho_L t}$$

so it is enough to show that there exists $L_0 \in \mathbb{N}$ such that

$$\sup_{L \geq L_0} \sup_{\mu_L} \left\| \mu_L P_t^L - \pi_L \right\|_{\text{TV}} \leq C e^{-\rho t}$$

for some $C, \rho > 0$.

Lemma 4.18 guarantees that there exist $T > 0$ and $L_0 \in \mathbb{N}$ such that for all $L \geq L_0$

$$\inf_{\sigma \in \Sigma} \mathbb{P}_\sigma (|S^L(T)| \geq L-1) \geq \frac{1}{9}.$$

Recall $\tau_i = \inf\{t \geq 0 : \sigma^i(t) = \tilde{\sigma}^i(t)\}$ and set $\tau_\sigma = \max_{i=1,2} \tau_i$, $\tau_y = \inf\{t \geq \tau_\sigma : y^L(t) = \tilde{y}^L(t)\}$ and

$$p^* = \inf_{\sigma_0^1, \tilde{\sigma}_0^1} \mathbb{P}_{(\sigma_0^1, \tilde{\sigma}_0^1)} (\sigma^1(1) = \tilde{\sigma}^1(1)).$$

For all $k \in \mathbb{N}$ and $L \geq L_0$

$$\begin{aligned} \sup_{\mu_L} \left\| \mu_L P_{(T+1)k}^L - \pi_L \right\|_{\text{TV}} &\leq \sup_{Y_0^L, \tilde{Y}_0^L} \mathbb{P}_{Y_0^L, \tilde{Y}_0^L} \left(Y((T+1)k) \neq \tilde{Y}^L((T+1)k) \right), \\ &\leq \sup_{Y_0^L, \tilde{Y}_0^L} \mathbb{P}_{Y_0^L, \tilde{Y}_0^L} (\tau_y > (T+1)k), \\ &\leq \sup_{Y_0^L, \tilde{Y}_0^L} \left[\mathbb{P}_{Y_0^L, \tilde{Y}_0^L} (\tau_y - \tau_\sigma > kT) \right. \\ &\quad \left. + \mathbb{P}_{Y_0^L, \tilde{Y}_0^L} (\tau_1 > k) + \mathbb{P}_{Y_0^L, \tilde{Y}_0^L} (\tau_2 > k) \right], \\ &\leq (8/9)^k + (1-p^*)^k + (1-p^*)^k, \end{aligned}$$

where lemma 4.17 and 4.18 are iterated to bound $\mathbb{P}(\tau_y - \tau_\sigma > kT)$ by $(8/9)^k$.

This concludes the proof since $t \mapsto \sup_{L \geq L_0} \sup_{\mu_L} \left\| \mu_L P_t^L - \pi_L \right\|_{\text{TV}}$ is decreasing. □

Remark 4.19. *The set $E = [0, \ell] \times \Sigma$ is compact w.r.t. convergence in law and convergence in the Wasserstein metric is equivalent to convergence in law and convergence of the first moment hence the sequence (π^L) lives in a compact set. Thus, it suffices to prove that the invariant measure π is its unique accumulation point to show*

$$W_1(i_L \# \pi_L, \pi) \rightarrow 0.$$

Let $i_{L_k} \# \pi^{L_k} \xrightarrow{k} \pi^\infty$ then the proof of Theorem 4.8 shows $i_{L_k} \# (\pi^{L_k} P_t^{L_k}) \rightarrow \pi^\infty P_t$ so that

$$\pi^\infty P_t = \lim i_{L_k} \# (\pi^{L_k} P_t^{L_k}) = \lim i_{L_k} \# \pi^{L_k} = \pi^\infty.$$

Hence $\pi^\infty = \pi$ is the unique invariant measure of the continuous process.

4.4 Mixing times

In this section, the dependence of the mixing time

$$t_{\text{mix}}(\epsilon) = \inf \left\{ t \geq 0 : \sup_{\mu} \|\mu P_t - \pi\|_{\text{TV}} \leq \epsilon \right\}$$

of the continuous process on the model parameters ω (resp. α and β) and ℓ is determined up to a multiplicative constant. Denote by $t_{\text{mix}}^i(\epsilon)$ (resp. $t_{\text{mix}}^f(\epsilon)$) the mixing time of the CITP (resp. CFTP).

Theorem 4.20 (Mixing time). *(i) For all $\epsilon \in (0, 1)$ there exists $C(\epsilon), C'(\epsilon) > 0$ such that*

$$C(\epsilon) \frac{1}{\omega} (1 + \omega^2 \ell^2) \leq t_{\text{mix}}^i(\epsilon) \leq C'(\epsilon) \frac{1}{\omega} (1 + \omega^2 \ell^2).$$

(ii) For all $\epsilon \in (0, 1)$ there exists $C(\epsilon), C'(\epsilon) > 0$ such that

$$C(\epsilon) \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) (1 + \alpha^2 \ell^2) \leq t_{\text{mix}}^f(\epsilon) \leq C'(\epsilon) \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) (1 + \alpha^2 \ell^2).$$

By standard arguments, e.g. [LPW09], one has that $C'(\epsilon) \sim \log(1/\epsilon)$. However the expression for $C(\epsilon)$ is more involved. Notice that matching lower and upper bounds are obtained meaning that the dependence on the model parameters is optimal. The remainder of this section is dedicated to the proof of theorem 4.20. The lower and upper bound of assertions (i) and (ii) are proven separately.

The exact forms of the invariant measure of the CITP and the CFTP are used during the proofs of the theorems, noticeably the fact that they exhibit Dirac masses at 0 and ℓ . While derived in the discrete case in [SEB16, SEB17], universality classes have been obtained in [HGM23] characterizing the exact form of the invariant measure in the continuous case in a general setting. For the sake of completeness, the invariant measure of the CITP is explicated in Proposition 4.31 and of the CFTP in Proposition 4.32.

4.4.1 Lower bound

The lower bounds in assertions (i) and (ii) of theorem 4.20 can be shown using identical arguments. For the sake of conciseness, only the lower bound of assertion (ii) is proven.

Following a general approach to prove lower bounds for mixing times, slow observables which constitute obstacles to mixing are identified. In the following lemma, a first lower bound is deduced by looking at the velocities.

Lemma 4.21. *For all $\epsilon \in (0, 1)$ one has*

$$\frac{\log(1/\epsilon) \max(1/\alpha, 1/\beta)}{2} \leq t_{\text{mix}}^f(\epsilon).$$

Proof. Set $X_0 = (\ell/2, 1, 1)$. If $t \geq 0$ is such that $\sup_{\mu} \|\mu P_t - \pi\|_{\text{TV}} \leq \epsilon$ then

$$\mathbb{P}_{X_0}(X(t) = X_0) = 1 - (1 - \mathbb{P}_{X_0}(X(t) = X_0)) = \pi(\{X_0\}^c) - \mathbb{P}_{X_0}(X(t) \neq X_0) \leq \epsilon.$$

using the explicit formula for π (see Proposition 4.32) which implies $\pi(\{X_0\}^c) = 0$. Furthermore

$$\mathbb{P}_{X_0}(X(t) = X_0) \geq \mathbb{P}_{X_0}(\sigma^1(s) = \sigma^2(s) = 1 \text{ for all } s \leq t) = e^{-2\alpha t} \implies t_{\text{mix}}^f(\epsilon) \geq \frac{\log(1/\epsilon)}{2\alpha}.$$

The same argument with $X_0 = (\ell/2, 0, 0)$ yields $t_{\text{mix}}^f(\epsilon) \geq \frac{\log(1/\epsilon)}{2\beta}$ and the result follows. \square

Exploring a positive fraction of the state space requires an order ℓ change in position. The two following lemmas show that, in the diffusive regime $\alpha\ell \gg 1$, this takes order ℓ^2 time leading to a lower bound with quadratic dependence on ℓ .

Lemma 4.22. *Let σ^1 and σ^2 be two independent Markov processes with the transition rates of figure 4.2b.*

For all $\epsilon > 0$ there exists $u(\epsilon), A(\epsilon) > 0$ such that for all $\delta > 0$, choosing $t^ = u(\epsilon)(1/\alpha + 1/\beta)\alpha^2\delta^2$,*

$$\alpha\delta \geq A(\epsilon) \implies \max_{\sigma \in \Sigma} \mathbb{P}_{\sigma} \left(\sup_{0 \leq t \leq t^*} \left| \int_0^t (\sigma^2(s) - \sigma^1(s)) ds \right| > \delta \right) \leq \epsilon.$$

Proof. One has

$$\mathbb{P}_{\sigma} \left(\sup_{0 \leq t \leq t^*} \left| \int_0^t (\sigma^2(s) - \sigma^1(s)) ds \right| > \delta \right) \leq 2\mathbb{P}_{\sigma_0^1} \left(\sup_{0 \leq t \leq t^*} \left| \int_0^t \sigma^1(s) ds \right| > \delta/2 \right)$$

so that it suffices to prove that for all $\epsilon \in (0, 1)$ there exist $u(\epsilon), A(\epsilon) > 0$ such that for $t^* = u(\epsilon)(1/\alpha + 1/\beta)\alpha^2\delta^2$

$$\alpha\delta \geq A(\epsilon) \implies \max_{\sigma_0^1 = -1, 0, 1} \mathbb{P}_{\sigma_0^1} \left(\sup_{0 \leq t \leq t^*} \left| \int_0^t \sigma^1(s) ds \right| > \delta \right) \leq \epsilon.$$

Define $\tilde{\tau}_{-1} = 0$ as well as $\tau_n = \inf\{t \geq \tilde{\tau}_{n-1} : \sigma^1(t) = 0\}$ and $\tilde{\tau}_n = \inf\{t \geq \tau_n : \sigma^1(t) \neq 0\}$ for $n \geq 0$. Denote $D_i = \int_{\tau_{i-1}}^{\tau_i} \sigma^1(s) ds$ and observe that

$$\sup_{t \in [0, \tau_n]} \left| \int_0^t \sigma^1(s) ds \right| = \max_{0 \leq k \leq n} \left| \int_0^{\tau_k} \sigma^1(s) ds \right| \leq \left| \int_0^{\tau_0} \sigma^1(s) ds \right| + \max_{1 \leq k \leq n} \left| \sum_{i=1}^k D_i \right|.$$

Set $n = \lceil 2u\alpha^2\delta^2 \rceil$ and $t^* = u(1/\alpha + a/\beta)\alpha^2\delta^2$. One has

$$\begin{aligned} & \mathbb{P}_{\sigma_0^1} \left(\sup_{0 \leq t \leq t^*} \left| \int_0^t \sigma^1(s) ds \right| > \delta \right) \\ & \leq \mathbb{P}_{\sigma_0^1} \left(\left| \int_0^{\tau_0} \sigma^1(s) ds \right| > \delta/2 \text{ or } \max_{1 \leq k \leq n} \left| \sum_{i=1}^k D_i \right| > \delta/2 \text{ or } t^* > \tau_n \right) \\ & \leq \underbrace{\mathbb{P}_{\sigma_0^1} \left(\left| \int_0^{\tau_0} \sigma^1(s) ds \right| > \delta/2 \right)}_{(I)} + \underbrace{\mathbb{P}_{\sigma_0^1} \left(\max_{1 \leq k \leq n} \left| \sum_{i=1}^k D_i \right| > \delta/2 \right)}_{(II)} \\ & \quad + \underbrace{\mathbb{P}_{\sigma_0^1} (t^* > \tau_n)}_{(III)}. \end{aligned}$$

It remains to show that there exist $u = u(\epsilon) > 0$ and $A = A(\epsilon) > 0$ such that each of the terms (I)–(III) is bounded by $\epsilon/3$ if $\alpha\delta \geq A$.

Bounding (I). If $\sigma_0^1 = 0$ then (I) is zero. If $\sigma_0^1 = \pm 1$ then Markov's inequality yields

$$\mathbb{P}_{\sigma_0^1} \left(\left| \int_0^{\tau_0} \sigma^1(s) ds \right| > \delta/2 \right) \leq \frac{\mathbb{E}_{\sigma_0^1} \left| \int_0^{\tau_0} \sigma^1(s) ds \right|}{\delta/2} = \frac{2}{\alpha\delta}$$

so choosing $A \geq 6/\epsilon$ leads to the desired bound.

Bounding (II). Notice that (D_i) are i.i.d. random variables so that by Kolmogorov's inequality

$$\mathbb{P}_{\sigma_0^1} \left(\max_{1 \leq k \leq n} \left| \sum_{i=1}^k D_i \right| > \delta/2 \right) \leq \frac{\mathbb{E}_{\sigma_0^1} (\sum_{i=1}^n D_i)^2}{\delta^2/4} = \frac{4n\mathbb{E}_{\sigma_0^1} [D_1^2]}{\delta^2} = 8 \frac{[2u\alpha^2\delta^2]}{\alpha^2\delta^2}.$$

If $2u\alpha^2\delta^2 \geq 1$ one has $[2u\alpha^2\delta^2] \leq 4u\alpha^2\delta^2$ so that

$$\mathbb{P}_{\sigma_0^1} \left(\max_{1 \leq k \leq n} \left| \sum_{i=1}^k D_i \right| > \delta/2 \right) \leq 32u$$

and the desired bound holds if $u \leq \epsilon/96$ and $A \geq 1/\sqrt{2u}$.

Bounding (III). Notice that the random variables $(\tau_i - \tau_{i-1})_{i \geq 1}$ are i.i.d. and have mean $1/\alpha + 2/\beta$ and variance $1/\alpha^2 + 4/\beta^2$. Chebyshev's inequality gives us

$$\begin{aligned} \mathbb{P}_{\sigma_0^1} (t^* > \tau_n) &\leq \mathbb{P}_{\sigma_0^1} \left(t^* > \sum_{i=1}^n (\tau_i - \tau_{i-1}) \right) \\ &= \mathbb{P}_{\sigma_0^1} \left(t^* - n(1/\alpha + 2/\beta) > \sum_{i=1}^n (\tau_i - \tau_{i-1}) - n(1/\alpha + 2/\beta) \right), \\ &\leq \frac{n\text{Var}(\tau_1 - \tau_0)}{(t^* - n(1/\alpha + 2/\beta))^2} \leq \frac{[2u\alpha^2\delta^2] (1/\alpha^2 + 4/\beta^2)}{u^2(1/\alpha + 1/\beta)^2\alpha^4\delta^4}. \end{aligned}$$

If $2u\alpha^2\delta^2 \geq 1$ one has $[2u\alpha^2\delta^2] \leq 4u\alpha^2\delta^2$ so that

$$\mathbb{P}_{\sigma_0^1} (t^*(u) > \tau_n) \leq \left(\sup_{\alpha, \beta > 0} \frac{1/\alpha^2 + 4/\beta^2}{(1/\alpha + 1/\beta)^2} \right) \frac{4}{u\alpha^2\delta^2} = \frac{16}{u\alpha^2\delta^2}$$

hence the desired bound for (III) holds if $A \geq \max \left(1/\sqrt{2u}, \sqrt{\frac{48}{\epsilon u}} \right)$. \square

Lemma 4.23. *For all $\epsilon \in (0, 1)$ there exist $\tilde{A}(\epsilon), \tilde{u}(\epsilon) > 0$ such that*

$$\alpha\ell \geq \tilde{A}(\epsilon) \implies t_{\text{mix}}^f(\epsilon) \geq \tilde{u}(\epsilon)(1/\alpha + 1/\beta)\alpha^2\ell^2.$$

Proof. Set $M = ((1-a)\ell/2, (1+a)\ell/2) \times \Sigma$ where $a \in (0, 1)$ is fixed later. The explicit formula for the invariant measure π implies $\pi(M^c) \geq 1-a$. It is clear from definition 4.2 that

$$\mathbb{P}_{(\ell/2, 1, 1)} (X(t) \in M^c) \leq \mathbb{P}_{(1, 1)} \left(\sup_{0 \leq \theta \leq t} \left| \int_0^\theta (\sigma^2(s) - \sigma^1(s)) ds \right| \geq \frac{a\ell}{2} \right).$$

By lemma 4.22, with $t^* = u \left(\frac{1-\epsilon}{4} \right) (1/\alpha + 1/\beta)\alpha^2\delta^2$,

$$\alpha\delta \geq A \left(\frac{1-\epsilon}{4} \right) \implies \mathbb{P}_{(1, 1)} \left(\sup_{0 \leq t \leq t^*} \left| \int_0^t (\sigma^2(s) - \sigma^1(s)) ds \right| \geq \delta \right) \leq \frac{1-\epsilon}{4}.$$

Taking $\delta = a\ell/2$ one gets with $t^* = \frac{a^2}{4}u\left(\frac{1-\epsilon}{4}\right)(1/\alpha + 1/\beta)\alpha^2\ell^2$,

$$\alpha\ell \geq \frac{2}{a}A\left(\frac{1-\epsilon}{4}\right) \implies \mathbb{P}_{(1,1)}\left(\sup_{0 \leq t \leq t^*} \left| \int_0^t (\sigma^2(s) - \sigma^1(s)) ds \right| \geq \frac{a\ell}{2}\right) \leq \frac{1-\epsilon}{4}.$$

Set $a = \frac{1-\epsilon}{2}$, $\tilde{A}(\epsilon) = \frac{2}{a}A\left(\frac{1-\epsilon}{4}\right)$, $\tilde{u}(\epsilon) = \frac{a^2}{4}u\left(\frac{1-\epsilon}{4}\right)$ and $t^* = \tilde{u}(\epsilon)(1/\alpha + 1/\beta)\alpha^2\ell^2$. If $\alpha\ell \geq \tilde{A}(\epsilon)$ and $t^* \geq t_{\text{mix}}^i(\epsilon)$ then

$$\begin{aligned} 1 - \frac{1-\epsilon}{2} - \frac{1-\epsilon}{4} &\leq 1 - a - \mathbb{P}_{(1,1)}\left(\sup_{0 \leq \theta \leq t^*} |I(\theta)| \geq \frac{a\ell}{2}\right) \\ &\leq \pi(M^c) - \mathbb{P}_{(\ell/2, 1, 1)}(X(t^*) \in M^c) \leq \sup_{\mu} \|\mu P_{t^*} - \pi\|_{\text{TV}} \leq \epsilon \end{aligned}$$

which is a contradiction because $\epsilon \in (0, 1)$. Hence $\alpha\ell \geq \tilde{A}(\epsilon)$ implies $t_{\text{mix}}^i(\epsilon) \geq \tilde{u}(\epsilon)(1/\alpha + 1/\beta)\alpha^2\ell^2$. \square

Assertion (ii) of theorem 4.20 now follows from lemma 4.21 and lemma 4.23.

4.4.2 Upper bound for the instantaneous tumble process

The upper bound of theorem 4.20 is shown using coupling techniques. The upper bound of assertion (i) is the focus of this section and the upper bound of assertion (ii) is treated in the next section. The continuous analog of the coupling in definition 4.15 is used for the proof of both assertions.

Definition 4.24 (Continuous-continuous coupling). *Let $(x_0, \sigma_0^1, \sigma_0^2), (\tilde{x}_0, \tilde{\sigma}_0^1, \tilde{\sigma}_0^2) \in [0, \ell] \times \Sigma$ be given. Let $s^1, s^2, \tilde{s}^1, \tilde{s}^2$ be four independent Markov jump processes with the transition rates of figure 4.2a (resp. figure 4.2b) and respective initial states $\sigma_0^1, \sigma_0^2, \tilde{\sigma}_0^1, \tilde{\sigma}_0^2$. Set $\tau_i = \inf\{t \geq 0 : s_i(t) = \tilde{s}_i(t)\}$ and define*

$$\sigma^i(t) = s^i(t) \text{ for } t \geq 0, \quad \tilde{\sigma}^i(t) = \tilde{s}^i(t) \text{ for } t \leq \tau_i \text{ and } \tilde{\sigma}^i(t) = s^i(t) \text{ for } t \geq \tau_i.$$

Finally set $X = \mathcal{X}(\sigma^1, \sigma^2, x_0, \ell)$ and $\tilde{X} = \mathcal{X}(\tilde{\sigma}^1, \tilde{\sigma}^2, \tilde{x}_0, \ell)$.

The goal for the construction of this coupling is that

$$\begin{aligned} \sigma^i(t_0) = \tilde{\sigma}^i(t_0) &\implies \sigma^i(t) = \tilde{\sigma}^i(t) \text{ for all } t \geq t_0, \\ X(t_0) = \tilde{X}(t_0) &\implies X(t) = \tilde{X}(t) \text{ for all } t \geq t_0. \end{aligned}$$

Let us also do the following observation: as was already the case for the coupling of the discrete process, the challenge is understanding how much time passes between $\tau_\sigma = \max_i \tau_i$ and the moment the positions become identical. The key ingredient here is that once the velocities are identical, the order between the positions is preserved in the sense that

$$\begin{aligned} x(\tau_\sigma) \leq \tilde{x}(\tau_\sigma) &\implies x(t) \leq \tilde{x}(t) \text{ for all } t \geq \tau_\sigma, \\ x(\tau_\sigma) \geq \tilde{x}(\tau_\sigma) &\implies x(t) \geq \tilde{x}(t) \text{ for all } t \geq \tau_\sigma. \end{aligned}$$

In particular, if $x(\tau_\sigma) \leq \tilde{x}(\tau_\sigma)$ and $t \geq \tau_\sigma$ is such that $\tilde{x}(t) = 0$ then $0 \leq x(t) \leq \tilde{x}(t) = 0$ and so the two processes must have coupled before time t . Deduce

$$\inf\{t \geq 0 : X(t) = \tilde{X}(t)\} \leq \inf\{t \geq \tau_\sigma : \tilde{x}(t) = 0\} \vee \{t \geq \tau_\sigma : \tilde{x}(t) = 0\} =: \tau_{\text{coupling}}.$$

Hence upper bounding the mixing times reduces to upper bounding a hitting time.

Section 4.2 and proposition 4.7 allows us to explicitly compute hitting times by solving systems of differential equations using the theory laid out in section 3 of [Dav93].

Lemma 4.25. *Let $\tau = \inf\left\{t \geq 0 : X_D(t) = (0, *_{(1, -1)}^0)\right\}$. There exists $C > 0$ such that*

$$\sup_{X_0 \in E_D} E_{X_0}[\tau] \leq C \frac{1}{\omega} (1 + \omega^2 \ell^2).$$

Proof. The mean hitting time $f(x, \sigma) = E_{(x, \sigma)}[\tau]$ satisfies the differential-algebraic system of equations

$$\left. \begin{aligned} -2\omega f(x, 1, 1) + \omega f(x, 1, -1) + \omega f(x, -1, 1) + 1 &= 0 \\ -2\partial_x f(x, 1, -1) - 2\omega f(x, 1, -1) + \omega f(x, 1, 1) + \omega f(x, -1, -1) + 1 &= 0 \\ 2\partial_x f(x, -1, 1) - 2\omega f(x, -1, 1) + \omega f(x, 1, 1) + \omega f(x, -1, -1) + 1 &= 0 \\ -2\omega f(x, -1, -1) + \omega f(x, 1, -1) + \omega f(x, -1, 1) + 1 &= 0 \end{aligned} \right\} \text{ for all } x \in (0, \ell)$$

with boundary conditions $f(0, 1, -1) = 0$ and $f(\ell, -1, 1) = f(\ell, 1, -1) + 2/\omega$ yielding

$$\begin{aligned} \mathbb{E}_{(x, 1, 1)}[\tau] &= \frac{2(\ell + x)\omega + (2\ell x - x^2)\omega^2 + 3}{2\omega}, & \mathbb{E}_{(x, 1, -1)}[\tau] &= \frac{4x + (2\ell x - x^2)\omega}{2}, \\ \mathbb{E}_{(x, -1, 1)}[\tau] &= \frac{4\ell\omega + (2\ell x - x^2)\omega^2 + 4}{2\omega}, & \mathbb{E}_{(x, -1, -1)}[\tau] &= \frac{2(\ell + x)\omega + (2\ell x - x^2)\omega^2 + 3}{2\omega}. \end{aligned}$$

The result follows by optimizing over x . \square

An upper bound for the mixing time of the continuous instantaneous tumble process can now be given.

Proof of the upper bound of theorem 4.20 (i). Let $(x_0, \sigma_0^1, \sigma_0^2), (\tilde{x}_0, \tilde{\sigma}_0^1, \tilde{\sigma}_0^2) \in [0, \ell] \times \Sigma$ be arbitrary but fixed. Let X, \tilde{X} and τ_1, τ_2 be as in definition 4.24.

Set $\tau_0 = \inf\{t \geq \tau_\sigma : x(t) = 0\}$ and $\tilde{\tau}_0 = \inf\{t \geq \tau_\sigma : \tilde{x}(t) = 0\}$. Using lemma 4.25 one gets

$$\begin{aligned} \mathbb{E}_{X_0, \tilde{X}_0}[\tau_{\text{coupling}}] &= \mathbb{E}_{X_0, \tilde{X}_0}[\max(\tau_0, \tilde{\tau}_0) - \max(\tau_1, \tau_2)] + \mathbb{E}_{X_0, \tilde{X}_0}[\max(\tau_1, \tau_2)] \\ &\leq \mathbb{E}_{X_0, \tilde{X}_0}[\tau_0 - \max(\tau_1, \tau_2)] + \mathbb{E}_{X_0, \tilde{X}_0}[\tilde{\tau}_0 - \max(\tau_1, \tau_2)] + \mathbb{E}_{X_0, \tilde{X}_0}[\tau_1] \\ &\quad + \mathbb{E}_{X_0, \tilde{X}_0}[\tau_2] \\ &\leq C \frac{1}{\omega} (1 + \omega^2 \ell^2) + C \frac{1}{\omega} (1 + \omega^2 \ell^2) + \frac{1}{2\omega} + \frac{1}{2\omega} \\ &= (2C + 1) \frac{1}{\omega} (1 + \omega^2 \ell^2). \end{aligned}$$

Let $\epsilon > 0$ be arbitrary but fixed and set $t^* = (2C + 1)(1 + \omega^2 \ell^2)/(\epsilon\omega)$. One has

$$\sup_{\mu} \|\mu P_{t^*} - \pi\|_{\text{TV}} \leq \sup_{X_0, \tilde{X}_0} \mathbb{P}_{X_0, \tilde{X}_0}(\tau_{\text{coupling}} > t^*) \leq \sup_{X_0, \tilde{X}_0} \frac{\mathbb{E}_{X_0, \tilde{X}_0}[\tau_{\text{coupling}}]}{t^*} \leq \epsilon. \quad \square$$

4.4.3 Upper bound for the finite tumble process

In order to upper bound the mixing time of the continuous finite tumble process, the same coupling as in the previous section is used. First, a control for $\tau_\sigma = \max_i \tau_i$ is needed.

Lemma 4.26. *Let $\sigma = (\sigma^1, \sigma^2), \tilde{\sigma} = (\tilde{\sigma}^1, \tilde{\sigma}^2)$ and τ_1, τ_2 be as in definition 4.24. One has*

$$\sup_{\sigma_0, \tilde{\sigma}_0 \in \Sigma} \mathbb{E}_{\sigma_0, \tilde{\sigma}_0}[\max(\tau_1, \tau_2)] \leq \frac{3}{\alpha}.$$

Proof. One has $\sup_{\sigma_0, \tilde{\sigma}_0} \mathbb{E}_{\sigma_0, \tilde{\sigma}_0}[\max(\tau_1, \tau_2)] \leq \sup_{\sigma_0, \tilde{\sigma}_0} \mathbb{E}[\tau_1 + \tau_2] \leq 2 \sup_{\sigma_0^1, \tilde{\sigma}_0^1} \mathbb{E}_{\sigma_0^1, \tilde{\sigma}_0^1}[\tau_1]$. Hence, it is enough to show $\sup_{\sigma_0^1, \tilde{\sigma}_0^1} \mathbb{E}_{(\sigma_0^1, \tilde{\sigma}_0^1)}[\tau_1] \leq \frac{3}{2\alpha}$.

Set $k_{\sigma_0^1, \tilde{\sigma}_0^1} = \mathbb{E}_{(\sigma_0^1, \tilde{\sigma}_0^1)}[\tau_1]$. By first step analysis

$$\begin{aligned} \sum_{(s^1, \tilde{s}^1)} q_{(\sigma_0^1, \tilde{\sigma}_0^1), (s^1, \tilde{s}^1)} k_{(s^1, \tilde{s}^1)} + 1 &= 0 \text{ if } \sigma_0^1 \neq \tilde{\sigma}_0^1, \\ k_{\sigma_0^1, \tilde{\sigma}_0^1} &= 0 \text{ if } \sigma_0^1 = \tilde{\sigma}_0^1, \end{aligned}$$

where q is the discrete generator of the couple $(\sigma^1, \tilde{\sigma}^1)$. Solving this system yields

$$\begin{aligned}\mathbb{E}_{(-1,-1)}[\tau_1] &= \mathbb{E}_{(0,0)}[\tau_1] = \mathbb{E}_{(1,1)}[\tau_1] = 0, \\ \mathbb{E}_{(0,-1)}[\tau_1] &= \mathbb{E}_{(0,1)}[\tau_1] = \mathbb{E}_{(-1,0)}[\tau_1] = \mathbb{E}_{(1,0)}[\tau_1] = \frac{1}{\alpha} \frac{4r+1}{4r+2}, \\ \mathbb{E}_{(-1,1)}[\tau_1] &= \mathbb{E}_{(1,-1)}[\tau_1] = \frac{1}{\alpha} \frac{3r+1}{2r+1},\end{aligned}$$

so that the result follows from optimizing over $r > 0$. \square

The next step is again understanding the amount of time that passes between τ_σ and the first moment where the positions coincide. Because a direct computation seems out of reach, it is best to return to the strategy that was used for the discrete process and focus on variations of the process $\int_0^t (\sigma^2(s) - \sigma^1(s)) ds$. First, an improved continuous version of the deterministic lemma 4.17 that relates the coupling time to the variations of $I(t)$ is needed.

Lemma 4.27. *Let $x(t)$ and $\tilde{x}(t)$ be as in definition 4.24 and assume $(\sigma_0^1, \sigma_0^2) = (\tilde{\sigma}_0^1, \tilde{\sigma}_0^2)$. Then for all $T > 0$ and all fixed realization of (σ^1, σ^2) one has*

$$\sup_{t_1 \leq t_2 \leq T} \left| \int_{t_1}^{t_2} (\sigma^2(s) - \sigma^1(s)) ds \right| \geq \ell \implies \inf\{t \geq 0 : x(t) = \tilde{x}(t)\} \leq T.$$

Proof. Assume without loss of generality that $x(0) \leq \tilde{x}(0)$ and that there exists $t_1 \leq t_2 \leq T$ such that $\int_{t_1}^{t_2} (\sigma^2(s) - \sigma^1(s)) ds \geq \ell$. A recursion shows that $x(0) \leq \tilde{x}(0)$ implies $x(t) \leq \tilde{x}(t)$ for all $t \geq 0$. Hence, if $x(t) = \ell$ then $\ell = x(t) \leq \tilde{x}(t) \leq \ell$ so that $x(t) = \tilde{x}(t)$.

Assume by contradiction that $x(t) \neq \tilde{x}(t)$ for all $t \leq T$. The previous remark implies $x(t) < \ell$ for all $t \leq T$. A recursion shows that this leads to $x(t) \geq x(t_1) + \int_{t_1}^t (\sigma^2(s) - \sigma^1(s)) ds$ for all $t \in [t_1, T]$. Thus $x(t_2) \geq x(t_1) + \int_{t_1}^{t_2} (\sigma^2(s) - \sigma^1(s)) ds \geq \ell$ which is impossible. \square

The previous deterministic result are completed by showing that variation of order ℓ after order $(1/\alpha + 1/\beta)(1 + \alpha^2 \ell^2)$ time. Instead of considering $\int_{t_1}^{t_2} (\sigma^2(s) - \sigma^1(s)) ds$, let turn to $\int_{\tau_0}^{\tau_n} (\sigma^2(s) - \sigma^1(s)) ds$ where the τ_i are return times of (σ^1, σ^2) to a set to be determined later. It permits to write

$$\int_{\tau_0}^{\tau_n} (\sigma^2(s) - \sigma^1(s)) ds = \sum_{i=1}^n \int_{\tau_{i-1}}^{\tau_i} (\sigma^2(s) - \sigma^1(s)) ds =: \sum_{i=1}^n D_i$$

and exploit the fact that the D_i are i.i.d. and more easily lend themselves to computations.

Lemma 4.28. *If σ^1 and σ^2 are independent Markov jump processes with the transition rates 4.2b then*

$$\inf_{\sigma} \mathbb{P}_{\sigma} \left(\sup_{t_1 \leq t_2 \leq t^*} \left| \int_{t_1}^{t_2} (\sigma^2(s) - \sigma^1(s)) ds \right| \geq \ell \right) \geq \frac{1}{144}$$

where $t^* = 288 \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) (1 + \alpha^2 \ell^2)$.

Proof. Define $\tilde{\tau}_n = \inf\{t \geq \tau_{n-1} : \sigma^1(t) \neq \sigma^2(t)\}$ and $\tau_n = \inf\{t \geq \tilde{\tau}_n : \sigma^1(t) = \sigma^2(t)\}$ for $n \geq 0$ with the convention that $\tau_{-1} = 0$.

Step 1. Solving the adequate system of linear equations, as in Lemma 4.25, yields

$$\begin{aligned}\mathbb{E}_{(0,0)}[\tau_0] &= \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) \frac{2r^2 + 5r + 1}{4r^2 + 6r + 2}, \\ \mathbb{E}_{(1,0)}[\tau_0] &= \mathbb{E}_{(-1,0)}[\tau_0] = \mathbb{E}_{(0,1)}[\tau_0] = \mathbb{E}_{(0,-1)}[\tau_0] = \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) \frac{4r + 1}{4r^2 + 6r + 2}, \\ \mathbb{E}_{(1,1)}[\tau_0] &= \mathbb{E}_{(-1,-1)}[\tau_0] = \mathbb{E}_{(-1,1)}[\tau_0] = \mathbb{E}_{(1,-1)}[\tau_0] = \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) \frac{3r + 1}{2r^2 + 3r + 1},\end{aligned}$$

so that optimizing over $r > 0$ leads to $\sup_{\sigma} \mathbb{E}_{\sigma} [\tau_0] \leq 1/\alpha + 1/\beta$.

Step 2. First step analysis yields that for $n = 0$ and $\sigma \in \Sigma$

$$\mathbb{P}_{\sigma}(\sigma(\tau_n) = (0, 0)) = \frac{2\alpha}{2\alpha + \beta} \text{ and } \mathbb{P}_{\sigma}(\sigma(\tau_n) = (1, 1)) = \mathbb{P}_{\sigma}(\sigma(\tau_n) = (-1, -1)) = \frac{\beta}{2\alpha + \beta}$$

and a simple iteration yields the same result for $n \geq 1$ and $\sigma \in \Sigma$. Hence the sequence $(\tau_i - \tau_{i-1})_{i \geq 1}$ is identically distributed and for all $\sigma \in \Sigma$

$$\mathbb{E}_{\sigma} [\tau_n - \tau_0] = \sum_{i=1}^n \mathbb{E}_{\sigma} [\tau_i - \tau_{i-1}] = n \sum_{s=-1,0,1} \mathbb{P}_{\sigma}(\sigma(\tau_0) = (s, s)) \mathbb{E}_{(s,s)} [\tau_0] = n \frac{\alpha^2 + 2\alpha\beta + \beta^2}{2\alpha^2\beta + \alpha\beta^2}.$$

Step 3. Define $D_i = \int_{\tau_{i-1}}^{\tau_i} (\sigma^2(s) - \sigma^1(s)) ds$ for $i \geq 1$ as well as $S_n = \sum_{i=1}^n D_i = \int_{\tau_0}^{\tau_n} (\sigma^2(s) - \sigma^1(s)) ds$ for $n \geq 1$. Start by computing $\mathbb{E}_{\sigma} [e^{\lambda \int_0^{\tau_0} (\sigma^2(s) - \sigma^1(s)) ds}]$ for all $\sigma \in \Delta = \{(\sigma^1, \sigma^2) \in \Sigma : \sigma^1 = \sigma^2\}$.

Let $\tilde{\tau}_0 = \inf\{t \geq 0 : \sigma^1(t) = \sigma^2(t)\}$ be the first hitting time of Δ and notice that τ_0 is the first *return* time of Δ so that $\tau_0 = \tilde{\tau}_0$ whenever the initial state of (σ^1, σ^2) is outside Δ . Furthermore, if one sets $k_{\sigma} = \mathbb{E}_{\sigma} [e^{\lambda \int_0^{\tau_0} (\sigma^2(s) - \sigma^1(s)) ds}]$ then, by first step analysis, the vector $(k_{\sigma})_{\sigma \in \Sigma}$ satisfies the system of linear equations

$$\begin{aligned} \sum_{(s^1, s^2) \in \Sigma} q_{(\sigma^1, \sigma^2), (s^1, s^2)} k_{(s^1, s^2)} + \lambda(\sigma^2 - \sigma^1) k_{(\sigma^1, \sigma^2)} &= 0 \text{ for } (\sigma^1, \sigma^2) \notin \Delta, \\ k_{(\sigma^1, \sigma^2)} &= 1 \text{ for } (\sigma^1, \sigma^2) \in \Delta \end{aligned}$$

where $(q_{\sigma, \sigma'})_{\sigma, \sigma' \in \Sigma}$ is the generator of the couple (σ^1, σ^2) .

Using this system and using $\mathbb{E}_{\sigma} [e^{\lambda \int_0^{\tau_0} (\sigma^2(s) - \sigma^1(s)) ds}] = \sum_{\sigma' \neq \sigma} -\frac{q_{\sigma, \sigma'}}{q_{\sigma, \sigma}} k_{\sigma'}$ for all $\sigma \in \Delta$ yields

$$\mathbb{E}_{\sigma} [e^{\lambda \int_0^{\tau_0} (\sigma^2(s) - \sigma^1(s)) ds}] = \frac{4\alpha^4 + 4\alpha^3\beta + \alpha^2\beta^2 - 2(2\alpha^2 + 3\alpha\beta + \beta^2)\lambda^2}{4\alpha^4 + 4\alpha^3\beta + \alpha^2\beta^2 + 4\lambda^4 - 4(2\alpha^2 + 3\alpha\beta + \beta^2)\lambda^2} := \phi(\lambda)$$

for all $\sigma \in \Delta$ so that for all $\sigma = (\sigma_0^1, \sigma_0^2) \in \Sigma$,

$$\begin{aligned} \mathbb{E}_{\sigma} [e^{\sum_{i=1}^n \lambda_i D_i}] &= \mathbb{E}_{\sigma} [e^{\sum_{i=1}^{n-1} \lambda_i D_i} \mathbb{E}_{(\sigma^1(\tau_n), \sigma^2(\tau_n))} [e^{\lambda \int_0^{\tau_0} (\sigma^2(s) - \sigma^1(s)) ds}]] \\ &= \mathbb{E}_{\sigma} [e^{\sum_{i=1}^{n-1} \lambda_i D_i}] \phi(\lambda_n), \end{aligned}$$

which can be iterated to obtain $\mathbb{E}_{\sigma} [e^{\sum_{i=1}^n \lambda_i D_i}] = \prod_{i=1}^n \phi(\lambda_i)$.

The D_i are thus independent and identically distributed and differentiating ϕ gives us

$$\mathbb{E}_{\sigma} [D_i] = \mathbb{E}_{\sigma} [D_i^3] = 0, \quad \mathbb{E}_{\sigma} [D_i^2] = \frac{4(\alpha + \beta)}{2\alpha^3 + \alpha^2\beta}, \quad \mathbb{E}_{\sigma} [D_i^4] = \frac{96(\alpha^2 + 4\alpha\beta + 2\beta^2)}{4\alpha^6 + 4\alpha^5\beta + \alpha^4\beta^2}.$$

Step 4. Set $n = \lceil \alpha^2 \ell^2 \rceil$. For all $\sigma = (\sigma_0^1, \sigma_0^2) \in K$ one has

$$\mathbb{E}_{\sigma} [S_n^2] = n \mathbb{E}_{\sigma} [D_i^2] = \lceil \alpha^2 \ell^2 \rceil \frac{4(\alpha + \beta)}{2\alpha^3 + \alpha^2\beta} \geq \alpha^2 \ell^2 \frac{4(\alpha + \beta)}{2\alpha^3 + \alpha^2\beta} = \ell^2 \frac{4r + 4}{2r + 1} \geq 2\ell^2.$$

Hence, by the Paley-Zygmund inequality

$$\mathbb{P}_{\sigma} (S_n^2 \geq \ell^2) \geq \mathbb{P}_{\sigma} \left(S_n^2 \geq \frac{1}{2} \mathbb{E}_{\sigma} [S_n^2] \right) \geq \left(1 - \frac{1}{2} \right)^2 \frac{\mathbb{E}_{\sigma} [S_n^2]^2}{\mathbb{E}_{\sigma} [S_n^4]}.$$

Making use of step 3 one gets

$$\begin{aligned} \frac{\mathbb{E}_\sigma [S_n^2]^2}{\mathbb{E}_\sigma [S_n^4]} &= \frac{n^2 \mathbb{E}_\sigma [D_i^2]^2}{n \mathbb{E}_\sigma [D_i^4] + 3n(n-1) \mathbb{E}_\sigma [D_i^2]^2} = \frac{1}{3} \left(1 + \frac{r^2 + 6r + 3}{n(r+1)^2} \right)^{-1} \\ &\geq \frac{1}{3} \left(1 + \frac{r^2 + 6r + 3}{(r+1)^2} \right)^{-1} \geq \frac{1}{12}. \end{aligned}$$

Putting it all together, one gets

$$\mathbb{P}_\sigma \left(\left| \int_{\tau_0}^{\tau_n} (\sigma^2(s) - \sigma^1(s)) ds \right| \geq \ell \right) \geq \frac{1}{48} \text{ for all } \sigma \in \Sigma.$$

Step 5. From step 2, one gets

$$\mathbb{E}_\sigma [\tau_n - \tau_0] = \lceil \alpha^2 \ell^2 \rceil \frac{\alpha^2 + 2\alpha\beta + \beta^2}{2\alpha^2\beta + \alpha\beta^2} \leq \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) (1 + \alpha^2 \ell^2) \frac{r+1}{2r+1} \leq \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) (1 + \alpha^2 \ell^2).$$

So, if one sets $t^* = 288(1/\alpha + 1/\beta)(1 + \alpha^2 \ell^2)$ then

$$\begin{aligned} &\mathbb{P}_\sigma \left(\sup_{t_1 \leq t_2 \leq t^*} \left| \int_{t_1}^{t_2} (\sigma^2(s) - \sigma^1(s)) ds \right| \geq \ell \right) \\ &\geq \mathbb{P}_\sigma \left(\left| \int_{\tau_0}^{\tau_n} (\sigma^2(s) - \sigma^1(s)) ds \right| \geq \ell \right) - \mathbb{P}_\sigma \left(\tau_0 \geq 144 \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) \right) \\ &\quad - \mathbb{P}_\sigma \left(\tau_n - \tau_0 \geq 144 \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) (1 + \alpha^2 \ell^2) \right) \\ &\geq \frac{1}{48} - \frac{1}{144} - \frac{1}{144} = \frac{1}{144}. \end{aligned}$$

using Markov's inequality for the last step. \square

The upper bound for the mixing time of the continuous finite tumble process can now be proven.

Proof of the upper bound of theorem 4.20 (ii). Let $\epsilon \in (0, 1)$ be given and let X, \tilde{X} and τ_1, τ_2 be as in definition 4.24. Set $t^* = a \frac{3}{\alpha} + 288k(1/\alpha + 1/\beta)(1 + \alpha^2 \ell^2)$ for $a > 0$ and $k \in \mathbb{N}$ to be fixed later and define $\tau_x = \inf\{t \geq \max(\tau_1, \tau_2) : x(t) = \tilde{x}(t)\}$.

One has

$$\begin{aligned} &\sup_{\mu} \|\mu P_t - \pi\|_{\text{TV}} \\ &\leq \sup_{X_0, \tilde{X}_0} \mathbb{P} \left(X(t) \neq \tilde{X}(t) \right) \leq \sup_{X_0, \tilde{X}_0} \mathbb{P} (\tau_x > t^*) \\ &\leq \sup_{X_0, \tilde{X}_0} \left\{ \mathbb{P} \left(\tau_x - \max(\tau_1, \tau_2) > 288k \left(\frac{1}{\alpha} + \frac{1}{\beta} \right) (1 + \alpha^2 \ell^2) \right) + \mathbb{P} \left(\max(\tau_1, \tau_2) > a \frac{3}{\alpha} \right) \right\} \\ &\leq \frac{1}{a} + \left(\frac{143}{144} \right)^k, \end{aligned}$$

where, in the last step, lemma 4.26 and Markov's inequality were used for the first term and lemma 4.27 and lemma 4.28 were combined and iterated for the second term. \square

4.5 Invariant probability measure of the continuous process

In [HGM23], the general form of the invariant measures, as well as universality classes, were derived for general velocity processes. For the sake of completeness, the explicit invariant measure π of the continuous process is derived here. Note that it was already computed in [SEB16, SEB17] as the limit of the discrete invariant measures, an approach made rigorous by proposition 4.14.

The key in the search for its invariant measure is the construction of the continuous process as a PDMP in Section 4.2. This allows us to use the characterization of the invariant measure in terms of its generator

$$\pi \text{ is invariant} \iff \int \mathcal{L}f d\pi = 0 \text{ for all } f \in D(E).$$

following from (34.7) and (34.11) of [Dav93].

4.5.1 Symmetries of the invariant measure

The mappings

$$\begin{aligned} \rho_1 &= \iota^{-1} \circ [(x, \sigma^1, \sigma^2) \mapsto (\ell - x, \sigma^2, \sigma^1)] \circ \iota, \\ \rho_2 &= \iota^{-1} \circ [(x, \sigma^1, \sigma^2) \mapsto (\ell - x, -\sigma^1, -\sigma^2)] \circ \iota, \\ \rho_3 &= \iota^{-1} \circ [(x, \sigma^1, \sigma^2) \mapsto (x, -\sigma^2, -\sigma^1)] \circ \iota \end{aligned}$$

are such that $f \mapsto f \circ \rho_i$ is a one-to-one mapping of $D(\mathcal{L})$ onto itself and satisfy the relations

$$\rho_1^2 = \rho_2^2 = \text{id}_E \text{ and } \rho_1 \circ \rho_2 = \rho_2 \circ \rho_1 = \rho_3.$$

The goal of this section is to show the following proposition, which proves useful in the search for the explicit form of the invariant probability.

Proposition 4.29. *For $i = 1, 2, 3$ one has $\pi = \rho_i \# \pi$.*

The proof of the previous proposition rests on the following lemma, the proof of which is omitted for the sake of brevity.

Lemma 4.30. *For $i = 1, 2, 3$ one has $(\mathcal{L}f) \circ \rho_i = \mathcal{L}(f \circ \rho_i)$ for all $f \in D(\mathcal{L})$.*

Proof of proposition 4.29. Because $f \mapsto f \circ \rho_i$ is a one-to-one mapping of $D(\mathcal{L})$ onto itself, one has that $\int \mathcal{L}f d\pi = 0$ for all $f \in D(\mathcal{L})$ implies

$$0 = \int \mathcal{L}(f \circ \rho_i) d\pi = \int (\mathcal{L}f) \circ \rho_i d\pi = \int \mathcal{L}f d(\rho_i \# \pi) \text{ for all } f \in D(\mathcal{L})$$

so that $\rho_i \# \pi$ is also an invariant probability measure. The uniqueness of π implies $\pi = \rho_i \# \pi$. \square

One has shown that the action of the group $G = \{\text{id}_E, \rho_1, \rho_2, \rho_3\} \cong (\mathbb{Z}/2\mathbb{Z})^2$ commutes with the generator. Denoting $p_G : E \rightarrow E/G$ the projection operator, this is a strong indication that $p_G(X_D(t))$ is also a Markov process which can be identified as the process in [HGM23].

4.5.2 Invariant measure of the continuous process

Introduce the following notation for convenience

$$\kappa = \sqrt{(\alpha + \beta)(2\alpha + \beta)/2}, \quad r = \alpha/\beta, \quad \tilde{r} = \kappa/\beta = \sqrt{(r+1)(2r+1)/2}$$

as well as

$$\begin{aligned} \lambda_d &= \frac{1}{r} \left(1 - \frac{\tilde{r}}{2\tilde{r} + (2r+1) \tanh(\frac{\kappa\ell}{2})} \right), \\ \lambda_a &= \frac{\kappa}{4\tilde{r}} \left(1 - \frac{1}{2r+2 + 2\tilde{r} \tanh(\frac{\kappa\ell}{2})} \right), \\ \lambda_b &= \frac{\kappa}{4(r+1) \cosh(\frac{\kappa\ell}{2}) (2\tilde{r} + (2r+1) \tanh(\frac{\kappa\ell}{2}))}. \end{aligned}$$

σ	d_σ^0	a_σ	d_σ^ℓ
(1, 1)	$\frac{1}{4(2+\omega\ell)}$	$\frac{\omega}{4(2+\omega\ell)}$	$\frac{1}{4(2+\omega\ell)}$
(1, -1)	$\frac{1}{2(2+\omega\ell)}$	$\frac{\omega}{4(2+\omega\ell)}$	0
(-1, 1)	0	$\frac{\omega}{4(2+\omega\ell)}$	$\frac{1}{2(2+\omega\ell)}$
(-1, -1)	$\frac{1}{4(2+\omega\ell)}$	$\frac{\omega}{4(2+\omega\ell)}$	$\frac{1}{4(2+\omega\ell)}$

(a) Instantaneous tumble

σ	d_σ^0	d_σ^ℓ	a_σ	b_σ^c	b_σ^s
(1, 1)	$\frac{1}{4r}$	$\frac{1}{4r}$	λ_a	$\frac{2r+1}{r} \lambda_b$	0
(1, 0)	1	0	$2r \lambda_a$	$(4r+2) \lambda_b$	$2\tilde{r} \lambda_b$
(1, -1)	λ_d	0	λ_a	$-(2r+1) \lambda_b$	$-2\tilde{r} \lambda_b$
(0, 1)	0	1	$2r \lambda_a$	$(4r+2) \lambda_b$	$-2\tilde{r} \lambda_b$
(0, 0)	r	r	$4r^2 \lambda_a$	$4r(2r+1) \lambda_b$	0
(0, -1)	1	0	$2r \lambda_a$	$(4r+2) \lambda_b$	$2\tilde{r} \lambda_b$
(-1, 1)	0	λ_d	λ_a	$-(2r+1) \lambda_b$	$2\tilde{r} \lambda_b$
(-1, 0)	0	1	$2r \lambda_a$	$(4r+2) \lambda_b$	$-2\tilde{r} \lambda_b$
(-1, -1)	$\frac{1}{4r}$	$\frac{1}{4r}$	λ_a	$\frac{2r+1}{r} \lambda_b$	0

(b) Finite tumble

Figure 4.5: Coefficients of the continuous invariant measure

Proposition 4.31. *The invariant probability π of the continuous instantaneous tumble process is given by*

$$\pi = \iota^{-1} \# \left(\sum_{\sigma \in \Sigma} (d_\sigma^0 \delta_0 + a_\sigma dx + d_\sigma^\ell \delta_\ell) \otimes \delta_\sigma \right)$$

with the coefficients summarized in figure 4.5a.

Proposition 4.32. *The unnormalized invariant measure π of the continuous finite tumble process is given by*

$$\begin{aligned} \pi = \iota^{-1} \# \left(\sum_{\sigma \in \Sigma} \left(d_\sigma^0 \delta_0 + d_\sigma^\ell \delta_\ell + \left(a_\sigma + b_\sigma^s \sinh \left(\kappa \left(x - \frac{\ell}{2} \right) \right) \right. \right. \right. \\ \left. \left. \left. + b_\sigma^c \cosh \left(\kappa \left(x - \frac{\ell}{2} \right) \right) \right) dx \right) \otimes \delta_\sigma \right) \end{aligned}$$

with the coefficients summarized in figure 4.5b.

Let first establish and solve a system ODEs for the density of the invariant measure in the bulk. First no boundary conditions are taken into account when solving the system so the solution depends on constants that remain to be fixed. One then looks at the boundary conditions to determine these constants. The proof of proposition 4.31 is very similar to the proof of proposition 4.32 so is not included here.

Proof of proposition 4.32. System of ODEs in the bulk. For $\sigma \in \Sigma$ let π_σ be the restriction of π to $(0, \ell) \times \{\sigma\}$ seen as measure on $(0, \ell)$. For arbitrary but fixed $f_\sigma \in C_c^\infty((0, \ell))$ one has that $f : E_D \rightarrow \mathbb{R}$ defined by

$$f(x, \sigma) = f_\sigma(x) \text{ for } (x, \sigma) \in (0, \ell) \times \Sigma \text{ and } f(x, \nu) = 0 \text{ otherwise}$$

is in the domain $D(\mathcal{L})$ of the generator. Rewriting $\int \mathcal{L}f d\pi = 0$ in terms of f_σ and π_σ and using the fact that the f_σ are arbitrary, one gets $-VF' + Q^t F = 0$ where

$$F = \left(\pi_{(1,1)} \quad \pi_{(1,0)} \quad \pi_{(1,-1)} \quad \pi_{(0,1)} \quad \pi_{(0,0)} \quad \pi_{(0,-1)} \quad \pi_{(-1,1)} \quad \pi_{(-1,0)} \quad \pi_{(-1,-1)} \right)^t,$$

$$Q = \begin{pmatrix} -2\alpha & \alpha & 0 & \alpha & 0 & 0 & 0 & 0 & 0 \\ \frac{\beta}{2} & -\alpha - \beta & \frac{\beta}{2} & 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & \alpha & -2\alpha & 0 & 0 & \alpha & 0 & 0 & 0 \\ \frac{\beta}{2} & 0 & 0 & -\alpha - \beta & \alpha & 0 & \frac{\beta}{2} & 0 & 0 \\ 0 & \frac{\beta}{2} & 0 & \frac{\beta}{2} & -2\beta & \frac{\beta}{2} & 0 & \frac{\beta}{2} & 0 \\ 0 & 0 & \frac{\beta}{2} & 0 & \alpha & -\alpha - \beta & 0 & 0 & \frac{\beta}{2} \\ 0 & 0 & 0 & \alpha & 0 & 0 & -2\alpha & \alpha & 0 \\ 0 & 0 & 0 & 0 & \alpha & 0 & \frac{\beta}{2} & -\alpha - \beta & \frac{\beta}{2} \\ 0 & 0 & 0 & 0 & 0 & \alpha & 0 & \alpha & -2\alpha \end{pmatrix},$$

and $V = \text{Diagonal}(0, -1, -2, 1, 0, -1, 2, 1, 0)$. Note that differentiation is used in distributional sense here.

This implies

$$\pi_{(1,1)} + \pi_{(-1,-1)} = \frac{1}{4r} (\pi_{(1,0)} + \pi_{(0,1)} + \pi_{(-1,0)} + \pi_{(0,-1)}),$$

$$\pi_{(0,0)} = \frac{r}{2} (\pi_{(1,0)} + \pi_{(0,1)} + \pi_{(-1,0)} + \pi_{(0,-1)}),$$

and $G' = AG$ in the distributional sense where

$$G = \begin{pmatrix} \pi_{(-1,1)} \\ \pi_{(-1,0)} + \pi_{(0,1)} \\ \pi_{(1,0)} + \pi_{(0,-1)} \\ \pi_{(1,-1)} \end{pmatrix} \text{ and } A = \begin{pmatrix} -\alpha & \frac{\beta}{4} & 0 & 0 \\ 2\alpha & -\frac{\alpha}{2} - \frac{3\beta}{4} & \frac{\alpha}{2} + \frac{\beta}{4} & 0 \\ 0 & -\frac{\alpha}{2} - \frac{\beta}{4} & \frac{\alpha}{2} + \frac{3\beta}{4} & -2\alpha \\ 0 & 0 & -\frac{\beta}{4} & \alpha \end{pmatrix}.$$

Hence

$$G = c_1 \begin{pmatrix} -2r + 2\tilde{r} - 1 \\ 8r - 4\tilde{r} + 4 \\ 8r + 4\tilde{r} + 4 \\ -2r - 2\tilde{r} - 1 \end{pmatrix} e^{\kappa x} dx + c_2 \begin{pmatrix} -2r - 2\tilde{r} - 1 \\ 8r + 4\tilde{r} + 4 \\ 8r - 4\tilde{r} + 4 \\ -2r + 2\tilde{r} - 1 \end{pmatrix} e^{-\kappa x} dx$$

$$+ c_3 \begin{pmatrix} 1 \\ 4r \\ 4r \\ 1 \end{pmatrix} dx + c_4 \left[\begin{pmatrix} -2/\alpha \\ -4/\beta \\ 4/\beta \\ 2/\alpha \end{pmatrix} + x \begin{pmatrix} 1 \\ 4r \\ 4r \\ 1 \end{pmatrix} \right] dx.$$

The symmetry $\pi = \rho_1 \# \pi$ implies $\pi_{(-1,0)} + \pi_{(0,1)} = (x \mapsto \ell - x) \# (\pi_{(1,0)} + \pi_{(0,-1)})$ which forces $c_2 = e^{\kappa\ell} c_1$ and $c_4 = 0$. Furthermore, the symmetry $\pi = \rho_3 \# \pi$ implies

$$\begin{aligned}\pi_{(-1,0)} &= \pi_{(0,1)} = \frac{1}{2} (\pi_{(-1,0)} + \pi_{(0,1)}), \\ \pi_{(1,0)} &= \pi_{(0,-1)} = \frac{1}{2} (\pi_{(1,0)} + \pi_{(0,-1)}), \\ \pi_{(1,1)} &= \pi_{(-1,-1)} = \frac{1}{8r} (\pi_{(-1,0)} + \pi_{(0,1)} + \pi_{(1,0)} + \pi_{(0,-1)}),\end{aligned}$$

so that $F = c_1 E_b^+ e^{\kappa x} dx + c_1 E_b^- e^{\kappa(\ell-x)} dx + c_3 E_a dx$ where

$$E_b^+ = \begin{pmatrix} \frac{2r+1}{r} \\ 4r + 2\tilde{r} + 2 \\ -2r - 2\tilde{r} - 1 \\ 4r - 2\tilde{r} + 2 \\ 4r(2r+1) \\ 4r + 2\tilde{r} + 2 \\ -2r + 2\tilde{r} - 1 \\ 4r - 2\tilde{r} + 2 \\ \frac{2r+1}{r} \end{pmatrix}, \quad E_b^- = \begin{pmatrix} \frac{2r+1}{r} \\ 4r - 2\tilde{r} + 2 \\ -2r + 2\tilde{r} - 1 \\ 4r + 2\tilde{r} + 2 \\ 4r(2r+1) \\ 4r - 2\tilde{r} + 2 \\ -2r - 2\tilde{r} - 1 \\ 4r + 2\tilde{r} + 2 \\ \frac{2r+1}{r} \end{pmatrix}, \quad E_a = \begin{pmatrix} 1 \\ 2r \\ 1 \\ 2r \\ 4r^2 \\ 2r \\ 1 \\ 2r \\ 1 \end{pmatrix}.$$

Boundary conditions. Because the measure is unique, it is ergodic and hence

$$d_{(0,1)}^0 = \lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t 1_{\{X(s)=(0,0,1)\}} ds = 0$$

since the set $\{s \geq 0 : X(s) = (0, 0, 1)\}$ is almost surely discrete. Similarly $d_{(-1,1)}^0 = d_{(-1,0)}^0 = 0$.

Let $x_\sigma \in \mathbb{R}$ be arbitrary but fixed for $\sigma \in \Sigma$. It is easy to check that

$$f = \iota^{-1} \circ \left[(x, \sigma) \mapsto \left(x_\sigma \frac{\ell - x}{\ell}, \sigma \right) \right] \circ \iota$$

is in $D(\mathcal{L})$. Injecting the formula for π into $\int \mathcal{L} f d\pi$ and using the fact that the x_σ are arbitrary, one gets

$$\mathcal{Q}^t D^0 - c_1 V E_b^+ - c_1 e^{\kappa\ell} V E_b^- - c_3 V E_a = 0 \quad (*)$$

where

$$D^0 = (d_\sigma^0)_{\sigma \in \Sigma} = \left(d_{(1,1)}^0, d_{(1,0)}^0, d_{(1,-1)}^0, 0, d_{(0,0)}^0, d_{(0,-1)}^0, 0, 0, d_{(-1,-1)}^0 \right)^t.$$

The system of linear equations (*) uniquely determines c_1, c_3 and D^0 if one imposes $d_{(1,0)}^0 = 1$. Finally, the symmetry $\pi = \rho_1 \# \pi$ implies $d_{(\sigma^1, \sigma^2)}^\ell = d_{(\sigma^2, \sigma^1)}^0$ for all $(\sigma^1, \sigma^2) \in \Sigma$. \square

Chapter 5

Invariant measures and mixing of two interacting run-and-tumble particles with jamming

This chapter is based on an upcoming article.

5.1 Introduction

Run-and-tumble particles (RTPs), used to model bacteria such as E. Coli [Ber04, BB72, Sch93], are characterized by random piecewise linear motion. In order to self-propel, RTPs consume energy at the particle level making them a prime example of active particles. This leads to out-of-equilibrium behavior such as clustering [CT15] and accumulation at boundaries [EWG15], two subjects of ongoing research. In fact, the two features are strongly linked since cluster configurations are boundary configurations in the joint position space.

Accumulation at boundaries is already visible on the non-equilibrium steady state (NESS) of a single one-dimensional run-and-tumble particle. Indeed, the NESS of a RTP subjected to thermal noise and confined in a bounded interval differs from the Boltzmann measure and displays exponential accumulation at the boundaries [MJK⁺18]. In the limit where the strength of the thermal noise vanishes, Dirac masses appear on the boundary, showing an even stronger form of accumulation. The same phenomenon can also be observed on a one-dimensional RTP subjected to a confining potential [DKM⁺19, BMR⁺20]. Strongly departing from the Boltzmann case, the support of the NESS is a bounded interval $[x_-, x_+]$. Furthermore, depending on model parameters two very different behaviors can be observed at the boundaries x_{\pm} . There are the close-to-passive configurations where the density of the NESS vanishes at the boundary and strongly out-of-equilibrium configurations where the NESS diverges at the boundaries, a strong mark of accumulation.

Clustering, on the other hand, is more challenging as it requires the study of multiple interacting active particles. In recent years, some answers have been provided by simulations [DDK⁺20b, TW21, DBG⁺18] but exact results remain limited. The two particle case constitutes a notable exception because it can be reduced to the one particle case by considering the relative particle. This comes at the price of dealing with a more involved dynamics for the speed. Clustering is visible on the NESS of two RTPs interacting through attractive potentials [LDMS21]. The potential $V(x) = \mu x^2/2$ maps to [BMR⁺20] and all hallmarks of accumulation at boundaries can be reinterpreted as clustering. The case $V(x) = c|x|$ leads again to a departure from the Boltzmann measure and displays both exponential accumulation and Dirac masses at the boundary, weak and strong telltale signs of clustering respectively. Clustering is also present in the NESS of two RTPs on a 1D torus interacting through jamming both when particles have two [SEB16] and three internal states [SEB17].

In fact, considering arbitrary tumbling mechanisms [HGM23] reveals the existence of two distinct universality classes: in the close-to-equilibrium detailed jamming class, clustering takes the form of Dirac masses, whereas in the strongly out-of-equilibrium global jamming class brings exponential concentration of mass in addition.

This paper is dedicated to three processes modeling two RTPs on the real line subjected to attractive interaction potentials. We consider the relative particle speeds $-2, 0, +2$ together with the potentials $V(x) = c|x|$ (instantaneous linear process) and the potential $V(x) = \mu x^2/2$ (instantaneous harmonic process). In the $V(x) = c|x|$ case, we also consider the relative particle speeds $-2, -1, 0_{\pm}, 0_0, +1, +2$ where 0_{\pm} and 0_0 are two distinct 0 velocity internal states with different tumbling rates (finite linear process). This model shares some features with [LDMS21] but differs in two crucial ways: the particles interact through jamming that impedes particles crossings and, in the linear potential case, we consider richer tumbling behavior, which is closer to the movement of bacteria such as *E. Coli* [SEB17, SSB12]. Jamming constitutes a tractable alternative to potentials with short range repulsion such as Lennard-Jones or Weeks-Chandler-Andersen for which the NESS is determined by a differential equation [LDMS21, Section IV] which seems challenging to solve explicitly. We compute the NESS of all processes using the mathematical framework of piecewise deterministic Markov processes (PDMPs) and in particular the explicit characterization of their generator [Dav93, Dav84]. This allows us to work directly at zero temperature without discretizing space [SEB16, SEB17] or looking at the limit of vanishing thermal noise strength [MJK⁺18]. All three processes show strong marks of clustering in the form of Dirac masses on the boundary. The instantaneous and finite linear process also display exponential concentration of mass and the density of the instantaneous harmonic process can diverge at the boundary for certain parameter choices. Furthermore, the behavior of the NESS of the finite linear process is particularly rich as its form changes depending on the parameters. When the ratio of the velocity of the particles and the strength of the attractive interaction c is under a certain threshold, the NESS is a product measure outside of the boundary. This independence is lost and an additional Dirac mass appears when the threshold is crossed, which is reminiscent of the shape transition in [DKM⁺19, BMR⁺20] and the separation into close-to-equilibrium and strongly out-of-equilibrium universality classes in [HGM23].

Finally, we tackle the question of the speed of convergence of these processes toward their NESS. Indeed, while the NESS determines their long time behavior, it is irrelevant at small time scales. Understanding precisely when the asymptotic behavior takes over thus is of paramount importance. We depart from the spectral gap approach in [MBE19, MJK⁺18] which provides asymptotic results in the form of exponential decay with an uncontrolled prefactor. Instead, we obtain non-asymptotic bounds in total variation by constructing an explicit coupling (see [Mal16] and references therein for a review of convergence results for PDMPs). Even though the processes in this article are not stochastically ordered [LMT96], they possess a weak form of order preservation that essentially reduces the study of convergence to the study of a hitting time. The Laplace transform of this hitting time, which gives the decay rate in total variation distance [LMT96], is bounded using non-asymptotic large deviation results [Wu00, Lez01, CG08]. We also show the sharpness of our convergence results. The method is general and can be extended to arbitrary tumbling mechanisms and linear potentials. The proof strategy shares similarities with [FGM12] where biased constant tumbling rates drive a single RTP toward 0, thus ensuring the existence of an invariant measure. Position dependent tumbling rates are tackled in [FGM16] using coupling arguments closer to [MT93] and continuous velocity distributions are addressed in [CRS14] using hypocoercivity techniques. The increased complexity of the last two models comes at the price of explicitness and sharpness. In the case of the instantaneous harmonic process, we stick to coupling methods but work in Wasserstein instead of total variation distance to take advantage of the contractivity of the underlying deterministic dynamics. The proof is adapted from the coupling argument in [BLBMZ12], which was extended in [CH15].

The rest of the article is organized as follows. The remainder of this section gives a detailed description of the mathematical model and summarizes the main results. In

Section 5.2 explicit formulae for the NESS are established. Finally, Section 5.3 quantifies the speed of convergence towards the invariant measures.

5.1.1 Model

We consider two point particles on the real line interacting through an attractive potential. Their motion is also subjected to telegraphic noise and jamming interactions. Using the positions x_1, x_2 and the velocities σ_1, σ_2 of the particles this can be modeled as follows

- the σ_i are independent Markov processes with the transition rates of figure 5.1a (resp. 5.1b),
- the positions x_i evolve according to the ODEs

$$\partial_t x_1 = -V'(x_1 - x_2) + v\sigma_1(t), \quad \partial_t x_2 = -V'(x_2 - x_1) + v\sigma_2(t),$$

where V is an interaction potential satisfying $V(-x) = V(x)$.



Figure 5.1: Single-particle velocity transition rates

The positions perform a null recurrent random walk so they do not reach a steady state. Thus we focus on the relative position $x = x_2 - x_1$ and the relative velocity $\sigma = \sigma_2 - \sigma_1$. The relative position is governed by the ODE

$$\partial_t x = -2V'(x) + v\sigma(t)$$

when $x > 0$ and by jamming interactions which will be detailed later when $x = 0$. The relative velocity follows the transition rates 5.2a (resp. 5.2b). We will always denote $\Sigma = \{-2, 0, 2\}$ (resp. $\Sigma = \{-2, -1, 0_{\pm}, 0_0, 1, 2\}$) the state space of σ and adopt the convention $v \cdot 0_{\pm} = v \cdot 0_0 = 0$.

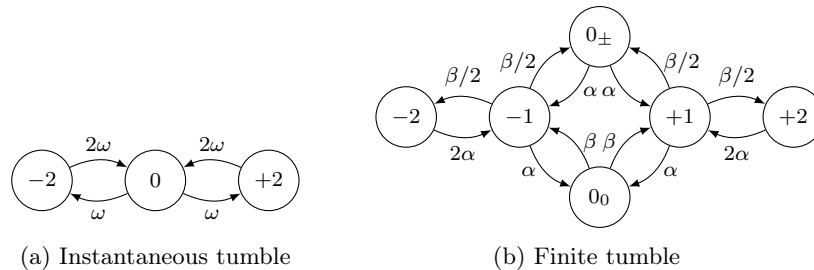


Figure 5.2: Relative velocity transition rates

Notice that if the σ_i follow figure 5.1b then $\sigma_2 - \sigma_1$ is not Markovian. Therefore, we split the state $\sigma_2 - \sigma_1 = 0$ into the states 0_{\pm} (corresponding to $\sigma_1 = \sigma_2 = \pm 1$) and 0_0 (corresponding to $\sigma_1 = \sigma_2 = 0$) to recover a Markov jump process (see figure 5.2b).

Assume without loss of generality that $x(0) = x_2(0) - x_1(0) \geq 0$ and define the behavior of the process when the particles collide, i.e. when $x(t) = 0$, as follows

- if $-2V'(0) + v\sigma(t) > 0$ the noise overcomes the attractive forces and pushes the particles away so x follows the ODE $\partial_t x = -2V'(x) + v\sigma(t)$ and immediately becomes positive,

- if $-2V'(0) + v\sigma(t) \leq 0$ the attractive forces overcome the noise so the particles are ‘glued’ together and x remains 0.

This yields jamming interactions leading to $x(t) \geq 0$ for all $t \geq 0$, which is also equivalent to the absence of particle crossings. The process can thus be recursively defined as follows.

Definition 5.1 (Interacting run-and-tumble process). *Let σ be a Markov jump process with the transition rates 5.2a (resp. 5.2b) and recursively construct*

$$x(t) = \max \left\{ 0, \phi_{t-T_n}^{\sigma(T_n)}(x(T_n)) \right\} \text{ for } t \in [T_n, T_{n+1}]$$

where $0 = T_0 < T_1 < \dots$ are the jump times of σ and $(\phi_t^\sigma)_{t \in \mathbb{R}}$ is the flow associated to the ODE $\partial_t x = -2V'(x) + v\sigma$. We call $(\mathbb{R}_+ \times \Sigma)$ -valued process $X(t) = (x(t), \sigma(t))$ interacting run-and-tumble process.

This paper is dedicated to the long-time behavior of three special cases of the interacting run-and-tumble process

- the *instantaneous linear* process where the particles are assumed to have instantaneous tumbles, i.e. σ follows figure 5.2a, and to interact through the linear potential $V(x) = c|x|$ (see figure 5.3a),
- the *finite linear* process where the particles are assumed to have finite tumble duration, i.e. σ follows figure 5.2b, and to interact through the linear potential $V(x) = c|x|$ (see figure 5.3b),
- the *instantaneous harmonic* process where the particles are assumed to have instantaneous tumbles, i.e. σ follows figure 5.2a, and to interact through the harmonic potential $V(x) = \frac{\mu}{2}x^2$ (see figure 5.3c).

Note that the linear potential (which resembles the one-dimensional Coulomb interaction [LDMS21]) and the harmonic potential are part of the larger class $V(x) = a|x|^p$ of physical interactions also studied in [DKM⁺19, GMS23]. It would be interesting to consider the *finite harmonic* process but computing its invariant measure requires solving a system of differential equations that seems intractable (see Section 5.2.1).

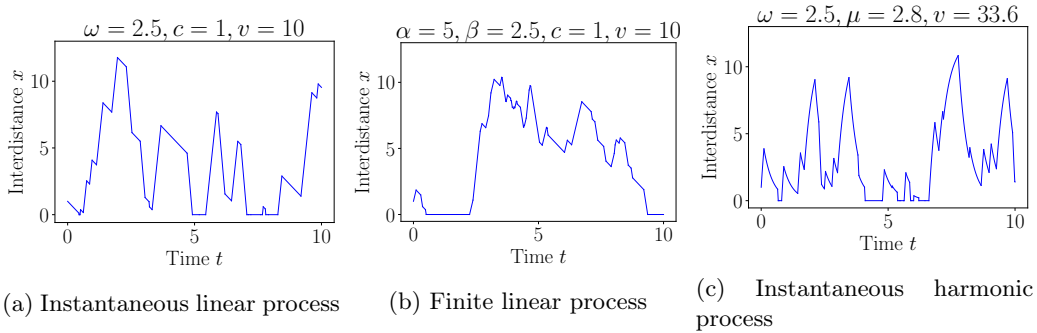


Figure 5.3: Realization of the interdistance x

Let us define the transition-rate matrices corresponding to the Markov jump processes 5.2a and 5.2b.

Definition 5.2. *Let $\mathcal{Q} = (q_{\sigma, \tilde{\sigma}})_{\sigma, \tilde{\sigma} \in \Sigma}$ be the discrete generator of the Σ -valued jump process σ . In the instantaneous case*

$$\mathcal{Q} = \begin{matrix} & \tilde{\sigma} = 2 & \tilde{\sigma} = 0 & \tilde{\sigma} = -2 \\ \begin{matrix} \sigma = 2 \\ \sigma = 0 \\ \sigma = -2 \end{matrix} & \begin{pmatrix} -2\omega & 2\omega & 0 \\ \omega & -2\omega & \omega \\ 0 & 2\omega & -2\omega \end{pmatrix} \end{matrix},$$

while in the finite case

$$\mathcal{Q} = \begin{matrix} & \tilde{\sigma} = 2 & \tilde{\sigma} = 1 & \tilde{\sigma} = 0_{\pm} & \tilde{\sigma} = 0_0 & \tilde{\sigma} = -1 & \tilde{\sigma} = -2 \\ \begin{matrix} \sigma = 2 \\ \sigma = 1 \\ \sigma = 0_{\pm} \\ \sigma = 0_0 \\ \sigma = -1 \\ \sigma = -2 \end{matrix} & \begin{pmatrix} -2\alpha & 2\alpha & 0 & 0 & 0 & 0 \\ \frac{1}{2}\beta & -\alpha - \beta & \frac{1}{2}\beta & \alpha & 0 & 0 \\ 0 & \alpha & -2\alpha & 0 & \alpha & 0 \\ 0 & \beta & 0 & -2\beta & \beta & 0 \\ 0 & 0 & \frac{1}{2}\beta & \alpha & -\alpha - \beta & \frac{1}{2}\beta \\ 0 & 0 & 0 & 0 & 2\alpha & -2\alpha \end{pmatrix} \end{matrix}.$$

The rigorous construction of the interacting run-and-tumble process is identical to the construction of the sticky PDMP samplers in [BGMS23] and leads to the following proposition.

Proposition 5.3. (i) *The interacting run-and-tumble process is a strong Markov Feller process with state space $E = \mathbb{R}_+ \times \Sigma$.*

(ii) *The domain $D(\mathcal{L})$ of its extended generator is the set of bounded measurable functions $f : E \rightarrow \mathbb{R}$ such that for all $(x, \sigma) \in E$*

$$t \mapsto f(\max(0, \phi_t^\sigma(x)), \sigma) \text{ is absolutely continuous}$$

and for $f \in D(\mathcal{L})$

$$\mathcal{L}f(x, \sigma) = 1_{\{x>0 \text{ or } -2V'(0)+v\sigma \geq 0\}} (-2V'(x) + v\sigma) \partial_x f(x, \sigma) + \sum_{\tilde{\sigma}} q_{\sigma, \tilde{\sigma}} f(x, \tilde{\sigma}).$$

5.1.2 Main results

This paper is devoted to the study of the long-time behavior of the instantaneous linear, finite linear and instantaneous harmonic process. Our main results are explicit formulae for their invariant measure and quantitative bounds for the speed of convergence towards these measures.

Using the PDMP formalism gives us access to the generator \mathcal{L} and enables us to use the characterization

$$\pi \text{ invariant} \iff \int \mathcal{L}f d\pi = 0 \text{ for all } f \in D(\mathcal{L}).$$

This dual formulation of the Fokker-Planck equations allows us to completely determine the invariant measure of the three processes considered in this article.

Theorem 5.4. (i) *The unique invariant measure of the instantaneous linear process is given by*

$$\pi = \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + a_\sigma e^{\zeta x} dx) \otimes \delta_\sigma$$

where ζ, d_σ and a_σ are explicit functions of ω, c and v (see Proposition 5.6).

(ii) *The unique invariant measure of the finite linear process is given by*

$$\pi = \begin{cases} \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + c_2 a_\sigma(\zeta_2) e^{\zeta_2 x} dx) \otimes \delta_\sigma & \text{when } c < v \leq 2c, \\ \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + c_2 a_\sigma(\zeta_2) e^{\zeta_2 x} dx + c_3 a_\sigma(\zeta_3) e^{\zeta_3 x} dx) \otimes \delta_\sigma & \text{when } v > 2c, \end{cases}$$

where c_i, ζ_i, d_σ and a_σ are explicit functions of α, β, c and v (see Proposition 5.10). Furthermore, when $v < c \leq 2v$, we have

$$d_{+2} = 0 \text{ and } d_\sigma > 0 \text{ otherwise,}$$

whereas when $c > 2v$, we have

$$d_{+2} = d_{+1} = 0 \text{ and } d_\sigma > 0 \text{ otherwise.}$$

- (iii) The unique invariant measure of the instantaneous harmonic process is an explicit linear combination Dirac masses and hypergeometric functions (see Proposition 5.12). Its support is the compact set $[0, v/\mu] \times \Sigma$.

Observe that in the bulk $\{x > 0\}$ the instantaneous linear invariant measure is a re-weighting of the ‘linear interaction without thermal noise’ case in [LDMS21]. This is because the processes have the same behavior in the bulk and they enter the bulk the same way (always through the point $(x, \sigma) = (0, 2)$). However, the coefficients of the Dirac masses on the boundary $\{x = 0\}$ differ. Furthermore, the state space of the processes in this article are of the form $\mathbb{R}_+ \times \Sigma$ and not $\mathbb{R} \times \Sigma$ since the particles cannot cross and therefore their interdistance remains non-negative. Also note that, unlike in [LDMS21, Section IV], the invariant measure is unique. The central observation that allows us to determine the invariant measure of the instantaneous linear process is that, in the bulk, the Fokker-Planck equations are a system of ordinary linear equations, which implies that the invariant measure is given by a matrix exponential.

It is noteworthy that the richer, biologically relevant [SSB12], relative velocity transition rates 5.2b lead to different formulae for the invariant measure depending on whether $c \leq v \leq 2c$ or $v > 2c$. This behavior comes from the fact that when $(x, \sigma) = (0, 1)$

- if $c < v \leq 2c$ then the attractive interaction between the particles is stronger than their self-propulsion so the process is ‘glued’ to $(0, 1)$ until a velocity change occurs,
- whereas if $v > 2c$ then the self-propulsion outweighs the attractive interaction and the particles immediately separate.

Hence the nature of the boundary point $(x, \sigma) = (0, 1)$ depends on c and v leading to the presence of a Dirac mass only when $v \leq 2c$. Note that, once restricted to the bulk, the invariant measure is a product measure for $v \leq 2c$ but not for $v > 2c$. This is reminiscent of the detailed and global jamming classes in [HGM23]. The strategy to determine the invariant measure remains the same as for the instantaneous linear process. As was already the case for the instantaneous linear process, the invariant measure of the instantaneous harmonic process coincides with [LDMS21] in the bulk after rescaling.

We also quantify the speed of convergence towards the invariant measure. In the linear potential case, we delay the statement of precise non-asymptotic convergence bounds until Theorem 5.16 and instead give asymptotic bounds which are easier to state and interpret.

Theorem 5.5. (i) *In the instantaneous linear case*

$$\frac{1}{2}\omega \frac{c^2}{v^2} \leq \liminf_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \leq \overline{\lim}_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \leq 4\omega \frac{c^2}{v^2}$$

for all $(x, \sigma) \in E$ where

$$\|\eta - \tilde{\eta}\|_{\text{TV}} = \inf \{ \mathbb{P}((x, \sigma) \neq (\tilde{x}, \tilde{\sigma})) : (x, \sigma) \sim \eta \text{ and } (\tilde{x}, \tilde{\sigma}) \sim \tilde{\eta} \}$$

is the total variation distance (see Theorem 5.16 and Proposition 5.19).

(ii) *In the finite linear case*

$$\begin{aligned} \frac{3 - \sqrt{5}}{8} \min \left(\alpha \left(1 + \frac{\alpha}{\beta} \right) \frac{c^2}{v^2}, \alpha \frac{c}{v} \right) &\leq \liminf_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \\ &\leq \overline{\lim}_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \leq 4\alpha \left(1 + \frac{\alpha}{\beta} \right) \frac{c^2}{v^2} \end{aligned}$$

for all $(x, \sigma) \in E$ (see Theorem 5.16 and Proposition 5.19).

(iii) In the instantaneous harmonic case, for $1 \leq p < q$

$$\mathcal{W}_p(\eta P_t, \pi) \leq 2 \left(\left(\int \max(x, v/\mu)^q d\eta(x) \right)^{1/q} + v/\mu \right) e^{-\rho(p)t} + 2e^{-2\omega t}.$$

where

$$\mathcal{W}_p(\eta, \tilde{\eta}) := \inf \left\{ (\mathbb{E} |x - \tilde{x}|^p)^{\frac{1}{p}} + \mathbb{P}(\sigma \neq \tilde{\sigma}) : (x, \sigma) \sim \eta \text{ and } (\tilde{x}, \tilde{\sigma}) \sim \tilde{\eta} \right\}$$

is a Wasserstein-type distance (see Definition 5.24) and $\rho(p) = \frac{2\mu\omega(1-p/q)}{\mu p + 2\omega(1-p/q)}$ (see Theorem 5.25).

Notice that the upper and lower bounds for the exponential decay rate only differ by a constant factor in the instantaneous linear case. Hence we have fully determined its dependence on all model parameters. The lower bound is shown using synchronous coupling. The upper bound is deduced from concentration inequalities reflecting the fact that the process is slow to explore parts of the state space where the interdistance x is large. The prefactor of the decay grows exponentially with x .

In the finite linear case there is a mismatch between the upper and lower bound when

$$\alpha \left(1 + \frac{\alpha}{\beta} \right) \frac{c^2}{v^2} \gg \alpha \frac{c}{v} \iff \left(1 + \frac{\alpha}{\beta} \right) \frac{c}{v} \gg 1.$$

Because $0 < c/v < 1$ this can only happen when $\alpha/\beta \gg 1$ meaning that the particles spend the overwhelming majority of their time tumbling and not running. Hence the mismatch only occurs in the least relevant regime from a modeling perspective. Outside this part of the parameter space, i.e. when $(1 + \alpha/\beta)(c/v)$ is upper bounded, the upper and lower bound differ at most by a constant factor. The core ideas of the proofs remain the same as for the instantaneous linear process. The prefactor of the decay is again exponential in x .

We also show the exponential ergodicity of the instantaneous harmonic process following the coupling approach in [BLBMZ12]. The idea is to use the exponential contractivity of the deterministic dynamics to obtain exponential decay in a Wasserstein-type distance. Had we chosen Foster-Lyapunov techniques [HM11], we would obtain exponential decay in total variation distance but the bounds would likely be loose.

5.2 Invariant measures

We delay the proof of the existence and uniqueness of invariant probabilities for our models (see Theorems 5.16 and 5.25) and turn to the derivation of exact formulae for these probabilities under the assumption that they exist and are unique.

5.2.1 Instantaneous linear process

We start by outlining the general approach we will use to compute the invariant measure of all our models. This method is similar to [Fry21, LDMS21] and always follow the same basic steps

1. write down differential equations for the density of the invariant measure in the bulk $\mathbb{R}_{>0} \times \Sigma$,
2. find the general solution of these equations (which contains constants that remain to be fixed),
3. use integrability and symmetry constraints to reduce the number of constants,

4. use the fact that the σ -marginal of π is the invariant measure of the Σ -valued process with generator \mathcal{Q} to determine the remaining constants and the weights of the Dirac masses on the boundary.

While these steps can be applied to a wide range of one-dimensional systems, they do not always enable us to fix all the constants. However, the following additional step always enables us to fully determine the invariant measure

5. reinject the result into the generator characterization of invariance $\int \mathcal{L}f d\pi = 0$.

Interestingly this approach can also be applied to systems with Gaussian noise [Fry21, MJK⁺18]. The main challenge is always finding an analytical solution for the system of differential equations in the bulk. This is best exemplified by considering a harmonic potential together with the relative velocity transition rates 5.2b. Indeed, the process is well defined and Theorem 5.25 can be adapted to show that it admits a unique invariant probability but writing it down explicitly seems challenging. With this in mind, we turn to the computation of the invariant measure of the instantaneous linear process.

Proposition 5.6 (Invariant measure of the instantaneous linear process). *The unique invariant measure of the instantaneous linear process is given by*

$$\pi = \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + a_\sigma e^{\zeta x} dx) \otimes \delta_\sigma$$

where $\zeta = -\frac{2c\omega}{v^2 - c^2}$ and

$$\begin{aligned} d_2 &= 0, & d_0 &= \frac{c}{c+v}, & d_{-2} &= \frac{cv}{(c+v)^2}, \\ a_2 &= \frac{c\omega}{2(v^2 - c^2)}, & a_0 &= \frac{c\omega}{(c+v)^2}, & a_{-2} &= \frac{c(v-c)\omega}{2(c+v)^3}. \end{aligned}$$

Proof. Let π be the unique invariant measure of the instantaneous linear process and let π_2, π_0, π_{-2} be the Borel measures on $\mathbb{R}_{>0}$ such that

$$\pi(1_{\{x>0\}}f) = \sum_{\sigma \in \Sigma} \int_0^{+\infty} f(x, \sigma) d\pi_\sigma(x)$$

for all bounded measurable $f : E \rightarrow \mathbb{R}$.

Let $f_{-2}, f_0, f_2 \in C_c^\infty(\mathbb{R}_{>0})$ be arbitrary but fixed and let the function $f : E \rightarrow \mathbb{R}$ be defined by $f_\sigma(x, \sigma) = f_\sigma(x)$ if $x > 0$ and $f(x, \sigma) = 0$ when $x = 0$. We have $f \in D(\mathcal{L})$ and using the explicit expression of \mathcal{L} in Proposition 5.3 yields

$$\begin{aligned} \int \mathcal{L}f d\pi &= \sum_{\sigma \in \Sigma} \left(\int_0^{+\infty} \left((v\sigma - 2c) \partial_x f_\sigma(x) + \sum_{\bar{\sigma} \in \Sigma} q_{\sigma, \bar{\sigma}} f_{\bar{\sigma}}(x) \right) d\pi_\sigma(x) \right) \\ &= \sum_{\sigma \in \Sigma} \left[-(v\sigma - 2c) \pi'_\sigma + \sum_{\bar{\sigma} \in \Sigma} q_{\bar{\sigma}, \sigma} \pi_{\bar{\sigma}} \right] (f_\sigma) \end{aligned}$$

where π'_σ is the distributional derivative of π_σ . By [Lig10, Theorem 3.37] we have $\int \mathcal{L}f d\pi = 0$ so, because the f_σ are arbitrary, we get

$$-\mathcal{V}\Pi' + \mathcal{Q}^t\Pi = 0$$

where $\mathcal{V} = v\text{Diag}(2, 0, -2) - 2cI_3$, $\mathcal{Q} = (q_{\sigma, \sigma'})_{\sigma, \sigma' \in \Sigma}$ is the discrete generator of the Markov jump process followed by the σ -marginal and $\Pi = (\pi_2, \pi_0, \pi_{-2})^t$. Because \mathcal{V} is invertible,

this is equivalent to $\Pi' = \mathcal{A}\Pi$ where $\mathcal{A} = \mathcal{V}^{-1}\mathcal{Q}^t$ the Jordan form of which reads

$$\mathcal{A} = \begin{pmatrix} 1 & 1 & 1 \\ -\frac{2v}{c} & \frac{2(c+v)}{v-c} & 2 \\ -1 & \frac{(c+v)^2}{(v-c)^2} & 1 \end{pmatrix} \begin{pmatrix} \frac{c}{v} & 0 & 0 \\ 0 & -\frac{2c\omega}{v^2-c^2} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ -\frac{2v}{c} & \frac{2(c+v)}{v-c} & 2 \\ -1 & \frac{(c+v)^2}{(v-c)^2} & 1 \end{pmatrix}^{-1}.$$

Hence the general (distributional) solution of $\Pi' = \mathcal{A}\Pi$ is

$$\begin{pmatrix} \pi_{-2} \\ \pi_0 \\ \pi_2 \end{pmatrix} = a \begin{pmatrix} 1 \\ \frac{2(c+v)}{v-c} \\ \frac{(c+v)^2}{(v-c)^2} \end{pmatrix} e^{-\frac{2c\omega}{v^2-c^2}x} dx + b \begin{pmatrix} 1 \\ -\frac{2v}{c} \\ -1 \end{pmatrix} e^{\frac{c}{v}x} dx + c \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} dx$$

where $a, b, c \in \mathbb{R}$ are constants that remain to be fixed. Because $\mathbb{R}_{>0}$ has finite mass under the π_σ , we know that $b = c = 0$ so that only a remains to be fixed.

Because the σ -marginal of the process is a Markov jump process with generator \mathcal{Q} we have

$$\pi([0, +\infty) \times \{2\}) = \frac{1}{4}, \quad \pi([0, +\infty) \times \{0\}) = \frac{1}{2}, \quad \pi([0, +\infty) \times \{-2\}) = \frac{1}{4}.$$

The measure π is unique so it is ergodic and almost surely

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t 1_{\{X(s)=(0,2)\}} ds = \pi(\{(0,2)\}).$$

Because the process instantaneously leaves the point $(0,2)$ and cannot come back before a stochastic velocity flip occurs, we have $\int_0^t 1_{\{X(s)=(0,2)\}} ds = 0$ for all $t \geq 0$ almost surely so that $\pi(\{(0,2)\}) = 0$.

Setting $\zeta = -\frac{2c\omega}{v^2-c^2}$ we get

$$\frac{1}{4} = \pi([0, +\infty) \times \{2\}) = a \int_0^{+\infty} e^{\zeta x} dx = \frac{a}{-\zeta} \implies a = \frac{c\omega}{2(v^2 - c^2)},$$

and thus

$$\pi_2 = \frac{c\omega}{2(v^2 - c^2)} e^{\zeta x} dx, \quad \pi_0 = \frac{c\omega}{(c+v)^2} e^{\zeta x} dx, \quad \pi_{-2} = \frac{c(v-c)\omega}{2(c+v)^3} e^{\zeta x} dx.$$

Finally

$$\begin{aligned} \pi(\{(0,0)\}) &= \frac{1}{2} - \pi_0(\mathbb{R}_{>0}) = \frac{c}{c+v}, \\ \pi(\{(0,-2)\}) &= \frac{1}{4} - \pi_{-2}(\mathbb{R}_{>0}) = \frac{cv}{(c+v)^2} \end{aligned}$$

which concludes our explicit characterization of π . \square

5.2.2 Finite linear process

Recall that when σ follows 5.2b its generator is given by

$$\mathcal{Q} = \begin{matrix} & \begin{matrix} \tilde{\sigma} = 2 & \tilde{\sigma} = 1 & \tilde{\sigma} = 0_{\pm} & \tilde{\sigma} = 0_0 & \tilde{\sigma} = -1 & \tilde{\sigma} = -2 \end{matrix} \\ \begin{matrix} \sigma = 2 \\ \sigma = 1 \\ \sigma = 0_{\pm} \\ \sigma = 0_0 \\ \sigma = -1 \\ \sigma = -2 \end{matrix} & \begin{pmatrix} -2\alpha & 2\alpha & 0 & 0 & 0 & 0 \\ \frac{1}{2}\beta & -\alpha - \beta & \frac{1}{2}\beta & \alpha & 0 & 0 \\ 0 & \alpha & -2\alpha & 0 & \alpha & 0 \\ 0 & \beta & 0 & -2\beta & \beta & 0 \\ 0 & 0 & \frac{1}{2}\beta & \alpha & -\alpha - \beta & \frac{1}{2}\beta \\ 0 & 0 & 0 & 0 & 2\alpha & -2\alpha \end{pmatrix} \end{matrix},$$

and define

$$\mathcal{V} = \begin{matrix} & \tilde{\sigma} = 2 & \tilde{\sigma} = 1 & \tilde{\sigma} = 0_{\pm} & \tilde{\sigma} = 0_0 & \tilde{\sigma} = -1 & \tilde{\sigma} = -2 \\ \begin{matrix} \sigma = 2 \\ \sigma = 1 \\ \sigma = 0_{\pm} \\ \sigma = 0_0 \\ \sigma = -1 \\ \sigma = -2 \end{matrix} & \begin{pmatrix} 2v - 2c & 0 & 0 & 0 & 0 & 0 \\ 0 & v - 2c & 0 & 0 & 0 & 0 \\ 0 & 0 & -2c & 0 & 0 & 0 \\ 0 & 0 & 0 & -2c & 0 & 0 \\ 0 & 0 & 0 & 0 & -v - 2c & 0 \\ 0 & 0 & 0 & 0 & 0 & -2v - 2c \end{pmatrix} \end{matrix}.$$

To find the invariant measure of the finite linear process, we will apply the same strategy as in the instantaneous case. As before, the key will be solving the linear ODE $-\mathcal{V}\Pi' + \mathcal{Q}^t\Pi = 0$ with the added difficulty of increased matrix size and non-invertibility of \mathcal{V} when $v = 2c$.

We start with a preliminary lemma related to the eigenvalues of $\mathcal{A} = \mathcal{V}^{-1}\mathcal{Q}^t$, the proof of which is delayed until the end of the section. We have

$$\det(\mathcal{V}) \det(-XI + \mathcal{A}) = \det(-X\mathcal{V} + \mathcal{Q}^t) = 8XP_2P_3$$

where the polynomials P_2 and P_3 are given by

$$\begin{aligned} P_2 &= c(v^2 - c^2)X^2 + ((2\alpha + \beta)c^2 - \beta v^2)X - \alpha(\alpha + \beta)c, \\ P_3 &= c(8c^2 - 2v^2)X^3 + 2(\alpha v^2 - (8\alpha + 4\beta)c^2)X^2 + 2c(5\alpha^2 + 5\alpha\beta + \beta^2)X \\ &\quad - (2\alpha^3 + 3\alpha^2\beta + \alpha\beta^2). \end{aligned}$$

Their roots will play a crucial role in the computation of the invariant measure.

Lemma 5.7. *For $v > c > 0$ and $\alpha, \beta > 0$ it holds that*

- (i) P_2 is a harmonic polynomial with one positive and one negative root,
- (ii) if $c < v < 2c$ then P_3 is a cubic polynomial with no negative roots,
- (iii) if $v = 2c$ then P_3 is a harmonic polynomial with no negative roots,
- (iv) if $v > 2c$ then P_3 is a cubic polynomial with a single negative root,
- (v) P_3 has only real roots,
- (vi) P_2 and P_3 have distinct negative roots when $v > 2c$.

Notation 5.8. *The following notations will be useful*

- denote ζ_2 the unique negative root of P_2 ,
- denote ζ_3 the unique negative root of P_3 when $v > 2c$,
- denote $\pi_{\mathcal{Q}}$ the invariant measure of the Markov jump process with generator \mathcal{Q} given by

$$\begin{pmatrix} \pi_{\mathcal{Q}}(\{2\}) \\ \pi_{\mathcal{Q}}(\{1\}) \\ \pi_{\mathcal{Q}}(\{0_{\pm}\}) \\ \pi_{\mathcal{Q}}(\{0_0\}) \\ \pi_{\mathcal{Q}}(\{-1\}) \\ \pi_{\mathcal{Q}}(\{-2\}) \end{pmatrix} = \frac{1}{(\alpha + \beta)^2} \begin{pmatrix} \frac{1}{4}\beta^2 \\ \alpha\beta \\ \frac{1}{2}\beta^2 \\ \alpha^2 \\ \alpha\beta \\ \frac{1}{4}\beta^2 \end{pmatrix}.$$

The next lemma gives explicit formulae for the eigenvectors corresponding to ζ_2 and ζ_3 . We also delay its proof to the end of the section.

Lemma 5.9. *If $\zeta = \zeta_2$ (resp. $v > 2c$ and $\zeta = \zeta_3$) the kernel of $-\zeta\mathcal{V} + \mathcal{Q}^t$ is spanned by*

$$a(\zeta) = \begin{pmatrix} -(3c\zeta - 2\alpha - \beta)\alpha^2\beta^2 \\ 4(3c\zeta - 2\alpha - \beta)(c\zeta - v\zeta - \alpha)\alpha^2\beta \\ -2(4c^2\zeta^2 - 6cv\zeta^2 + 2v^2\zeta^2 - 6\alpha c\zeta - 2\beta c\zeta + 4\alpha v\zeta + 2\beta v\zeta + 2\alpha^2 + \alpha\beta)(c\zeta - \beta)\alpha\beta \\ -4(4c^2\zeta^2 - 6cv\zeta^2 + 2v^2\zeta^2 - 6\alpha c\zeta - 2\beta c\zeta + 4\alpha v\zeta + 2\beta v\zeta + 2\alpha^2 + \alpha\beta)(c\zeta - \alpha)\alpha^2 \\ 4\rho(\zeta)(3c\zeta - 2\alpha - \beta)(c\zeta + v\zeta - \alpha)\alpha^2\beta \\ -\rho(\zeta)(3c\zeta - 2\alpha - \beta)\alpha^2\beta^2 \end{pmatrix}$$

where

$$\rho(\zeta) = \frac{4c^2\zeta^2 - 6cv\zeta^2 + 2v^2\zeta^2 - 6\alpha c\zeta - 2\beta c\zeta + 4\alpha v\zeta + 2\beta v\zeta + 2\alpha^2 + \alpha\beta}{4c^2\zeta^2 + 6cv\zeta^2 + 2v^2\zeta^2 - 6\alpha c\zeta - 2\beta c\zeta - 4\alpha v\zeta - 2\beta v\zeta + 2\alpha^2 + \alpha\beta}.$$

We can now give explicit formulae for the invariant measure of the finite linear process.

Proposition 5.10 (Invariant measure of the finite linear process). (i) When $c < v \leq 2c$ the unique invariant measure of the finite linear process is given by

$$\pi = \sum_{\sigma \in \Sigma} (d_\sigma \delta_0 + c_2 a_\sigma(\zeta_2) e^{\zeta_2 x} dx) \otimes \delta_\sigma$$

where

$$c_2 = \frac{-\zeta_2}{4(-3c\zeta_2 + 2\alpha + \beta)(\alpha + \beta)^2 \alpha^2}$$

and

$$d_\sigma = \begin{cases} 0 & \text{if } \sigma = 2, \\ \pi_{\mathcal{Q}}(\{\sigma\}) + \frac{c_2 a_\sigma(\zeta_2)}{\zeta_2} & \text{if } \sigma \neq 2. \end{cases}$$

(ii) When $v > 2c$ the unique invariant measure of the finite linear process is given by

$$\pi = \sum_{\sigma \in \mathbb{V}} (d_\sigma \delta_0 + c_2 a_\sigma(\zeta_2) e^{\zeta_2 x} dx + c_3 a_\sigma(\zeta_3) e^{\zeta_3 x} dx) \otimes \delta_\sigma$$

where

$$c_2 = -\frac{\zeta_2 \zeta_3}{4\alpha^2(\alpha + \beta)^2(3c\zeta_2 - 2\alpha - \beta)(\zeta_2 - \zeta_3)},$$

$$c_3 = \frac{\zeta_2 \zeta_3}{4\alpha^2(\alpha + \beta)^2(3c\zeta_3 - 2\alpha - \beta)(\zeta_2 - \zeta_3)},$$

and

$$d_\sigma = \begin{cases} 0 & \text{if } \sigma = 1, 2, \\ \pi_{\mathcal{Q}}(\{\sigma\}) + \frac{c_2 a_\sigma(\zeta_2)}{\zeta_2} + \frac{c_3 a_\sigma(\zeta_3)}{\zeta_3} & \text{if } \sigma \neq 1, 2. \end{cases}$$

Notice that the invariant measure has a rich behavior and differs when $c < v \leq 2c$ and when $v > 2c$. Because $\pi(\mathbb{R}_+ \times \{\sigma\}) < +\infty$ only the negative eigenvalues of $\mathcal{A} = \mathcal{V}^{-1}\mathcal{Q}^t$ can enter the invariant measure. Hence, the different formulae can be explained by the fact that the characteristic polynomial of \mathcal{A} equals $\frac{8}{\det(\mathcal{V})}XP_2P_3$ and

- when $c < v \leq 2c$ the polynomial P_2 has a unique negative root while P_3 has no negative root,
- when $v > 2c$ the polynomials P_2 and P_3 each have a single negative root,

as stated in Lemma 5.7. The process also differs qualitatively. As outlined in the introduction

- when $c < v \leq 2c$ the process stays ‘glued’ to $(0, 1)$ whenever it visits the state,

- when $v > 2c$ the process immediately leaves the state $(0, 1)$ whenever it passes through it.

As a result $\pi(\{(0, 1)\}) > 0$ when $c < v \leq 2c$ and $\pi(\{(0, 1)\}) = 0$ when $v > 2c$.

Proof of Proposition 5.10. Let π be the unique invariant measure of the finite linear process and for $\sigma \in \Sigma$ let $(\pi_\sigma)_{\sigma \in \Sigma}$ be the Borel measures on $\mathbb{R}_{>0}$ such that

$$\pi(1_{\{x>0\}}f) = \sum_{\sigma \in \Sigma} \int_0^{+\infty} f(x, \sigma) d\pi_\sigma(x).$$

for all bounded measurable $f : E \rightarrow \mathbb{R}$.

The same argument as in the proof of Proposition 5.6 yields

$$-\mathcal{V}\Pi' + \mathcal{Q}^t\Pi = 0 \quad (*)$$

in the distributional sense where $\Pi = (\pi_\sigma)_{\sigma \in \Sigma}$.

Case $v \neq 2c$. In this case, \mathcal{V} is invertible so $(*)$ is equivalent to $\Pi' = \mathcal{A}\Pi$ where $\mathcal{A} = \mathcal{V}^{-1}\mathcal{Q}^t$. Hence solving $(*)$ boils down to the spectral analysis of \mathcal{A} .

If we denote

- K the number of Jordan blocks of \mathcal{A} ,
- N_k the size of the k -th Jordan block,
- λ_k the eigenvalue associated to the k -th Jordan block,
- $(\mathbf{a}_{k,n})_{n=0}^{N_k-1}$ the k -th Jordan chain of generalized eigenvectors, i.e.

$$\mathcal{A}\mathbf{a}_{k,0} = \lambda_k\mathbf{a}_{k,0} \text{ and } \mathcal{A}\mathbf{a}_{k,n} = \lambda_k\mathbf{a}_{k,n} + \mathbf{a}_{k,n-1} \text{ for } n \geq 1,$$

then

$$e^{t\mathcal{A}} \left(\sum_{k=1}^K \sum_{n=0}^{N_k-1} b_{k,n} \mathbf{a}_{k,n} \right) = \sum_{k=1}^K e^{\lambda_k t} \sum_{n=0}^{N_k-1} P_k^{(n)}(t) \mathbf{a}_{k,n}$$

where $P_k = \sum_{n=0}^{N_k-1} b_{k,n} \frac{X^n}{n!}$.

We have

$$\det(\mathcal{A} - XI) = \frac{1}{\det(\mathcal{V})} \det(-X\mathcal{V} + \mathcal{Q}^t) = \frac{8}{\det(\mathcal{V})} XP_2P_3 \quad (**)$$

so Lemma 5.7 implies that all eigenvalues of \mathcal{A} are real. Because the π_σ are integrable we have $b_{k,n} = 0$ for all k such that $\lambda_k \geq 0$ and all $n \in \{0, \dots, N_k - 1\}$. Thus we need only consider the negative eigenvalues of \mathcal{A} and the associated generalized eigenvectors.

Subcase $c < v < 2c$. Lemma 5.7 and $(**)$ imply that ζ_2 is the only negative eigenvalue of \mathcal{A} . It has multiplicity one and

$$\mathcal{V} \text{ invertible} \implies \ker(\mathcal{A} - \zeta_2 I) = \ker(-\zeta_2 \mathcal{V} + \mathcal{Q}^t)$$

so Lemma 5.9 gives us the associated eigenvector $\mathbf{a}(\zeta_2)$. Hence $\Pi = c_2 \mathbf{a}(\zeta_2) e^{\zeta_2 x} dx$ where c_2 is a constant that remains to be determined.

Because the σ -marginal of the process is a Markov jump process with generator \mathcal{Q} we have $\pi(\mathbb{R}_+ \times \{\sigma\}) = \pi_{\mathcal{Q}}(\{\sigma\})$ for $\sigma \in \Sigma$. The ergodicity argument in the proof of Proposition 5.6 shows $d_2 = \pi(\{(0, 2)\}) = 0$ so

$$\begin{aligned} \pi(\mathbb{R}_{>0} \times \{2\}) = \pi_{\mathcal{Q}}(\{2\}) &\iff c_2 \frac{a_2(\zeta_2)}{-\zeta_2} = \frac{\beta^2}{4(\alpha + \beta)^2} \\ &\iff c_2 = \frac{-\zeta_2}{4(-3c\zeta_2 + 2\alpha + \beta)(\alpha + \beta)^2 \alpha^2}. \end{aligned}$$

For $\sigma \neq 2$ we have

$$\pi(\{(0, \sigma)\}) + \pi(\mathbb{R}_{>0} \times \{\sigma\}) = \pi_{\mathcal{Q}}(\{\sigma\}) \iff d_\sigma = \pi(\{(0, \sigma)\}) = \pi_{\mathcal{Q}}(\{\sigma\}) + \frac{c_2 a_\sigma(\zeta_2)}{\zeta_2}.$$

Subcase $v > 2c$. Lemma 5.7 and (**) imply that ζ_2 and ζ_3 are the only negative eigenvalues of \mathcal{A} . Both have multiplicity one and Lemma 5.9 yields the associated eigenvectors. Therefore $\Pi = c_2 a(\zeta_2) e^{\zeta_2 x} dx + c_3 a(\zeta_3) e^{\zeta_3 x} dx$ where c_2 and c_3 are constants that remain to be fixed.

The usual ergodicity argument implies $d_2 = \pi(\{(0, 2)\}) = 0$ and $d_1 = \pi(\{(0, 1)\}) = 0$. Furthermore $\pi(\mathbb{R}_+ \times \{\sigma\}) = \pi_{\mathcal{Q}}(\{\sigma\})$ so

$$\pi(\mathbb{R}_{>0} \times \{2\}) = \pi_{\mathcal{Q}}(\{2\}) \text{ and } \pi(\mathbb{R}_{>0} \times \{1\}) = \pi_{\mathcal{Q}}(\{1\})$$

which can be rewritten as $M \begin{pmatrix} c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} \pi_{\mathcal{Q}}(\{2\}) \\ \pi_{\mathcal{Q}}(\{1\}) \end{pmatrix}$ with $M = \begin{pmatrix} \frac{a_2(\zeta_2)}{-\zeta_2} & \frac{a_2(\zeta_3)}{-\zeta_3} \\ \frac{a_1(\zeta_2)}{-\zeta_2} & \frac{a_1(\zeta_3)}{-\zeta_3} \end{pmatrix}$.

We have

$$\det(M) = \frac{4(3c\zeta_2 - 2\alpha - \beta)(3c\zeta_3 - 2\alpha - \beta)\alpha^4\beta^3(c - v)(\zeta_2 - \zeta_3)}{\zeta_2\zeta_3} \neq 0$$

so we can solve the system and obtain

$$\begin{aligned} c_2 &= -\frac{\zeta_2\zeta_3}{4\alpha^2(\alpha + \beta)^2(3c\zeta_2 - 2\alpha - \beta)(\zeta_2 - \zeta_3)}, \\ c_3 &= \frac{\zeta_2\zeta_3}{4\alpha^2(\alpha + \beta)^2(3c\zeta_3 - 2\alpha - \beta)(\zeta_2 - \zeta_3)}. \end{aligned}$$

For $\sigma \neq 1, 2$ we have

$$\begin{aligned} \pi(\{(0, \sigma)\}) + \pi(\mathbb{R}_{>0} \times \{\sigma\}) &= \pi_{\mathcal{Q}}(\{\sigma\}) \iff \\ d_\sigma = \pi(\{(0, \sigma)\}) &= \pi_{\mathcal{Q}}(\{\sigma\}) + \frac{c_2 a_\sigma(\zeta_2)}{\zeta_2} + \frac{c_3 a_\sigma(\zeta_3)}{\zeta_3}. \end{aligned}$$

Case $v = 2c$. In this case the matrix \mathcal{V} is no longer invertible. It follows from lemma 5.7 and lemma 5.9 that $\Pi = c_2 a(\zeta_2) e^{\zeta_2 x} dx$ is a solution of (*). We now show that the solutions of (*) that are integrable on $\mathbb{R}_{>0}$ form a one-dimensional vector space.

When $v = 2c$ equation (*) reads

$$\begin{aligned} \frac{1}{2}\beta f_1(x) - 2\alpha f_2(x) - f_2'(x)v &= 0, \\ \beta f_{0_0}(x) + \alpha f_{0_\pm}(x) - (\alpha + \beta)f_1(x) + 2\alpha f_2(x) &= 0, \\ -2\alpha f_{0_\pm}(x) + \frac{1}{2}\beta f_1(x) + \frac{1}{2}\beta f_{-1}(x) + f_{0_\pm}'(x)v &= 0, \\ -2\beta f_{0_0}(x) + \alpha f_1(x) + \alpha f_{-1}(x) + f_{0_0}'(x)v &= 0, \\ \beta f_{0_0}(x) + \alpha f_{0_\pm}(x) - (\alpha + \beta)f_{-1}(x) + 2\alpha f_{-2}(x) + 2f_{-1}'(x)v &= 0, \\ \frac{1}{2}\beta f_{-1}(x) - 2\alpha f_{-2}(x) + 3f_{-2}'(x)v &= 0. \end{aligned}$$

Hence the second line can be used to remove f_1 from the system of equations and obtain a system of linear differential equations which can be solved. This is equivalent to the observation that (*) implies $\Pi = M_{5 \rightarrow 6} M_{6 \rightarrow 5} \Pi$ and $(M_{6 \rightarrow 5} \Pi)' = \tilde{\mathcal{A}}(M_{6 \rightarrow 5} \Pi)$ where $\tilde{\mathcal{A}} = (M_{6 \rightarrow 5} \mathcal{V} M_{5 \rightarrow 6})^{-1} M_{6 \rightarrow 5} \mathcal{Q}^t M_{5 \rightarrow 6}$ and

$$M_{5 \rightarrow 6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{2\alpha}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} & \frac{\beta}{\alpha+\beta} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad M_{6 \rightarrow 5} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Assume by contradiction that Π_1 and Π_2 are two linearly independent solutions of (*) that are integrable on $\mathbb{R}_{>0}$. We have

$$\sum_{i=1,2} a_i M_{6 \rightarrow 5} \Pi_i = 0 \implies \sum_{i=1,2} a_i M_{5 \rightarrow 6} M_{6 \rightarrow 5} \Pi_i = 0 \implies \sum_{i=1,2} a_i \Pi_i = 0 \implies a_1 = a_2 = 0$$

so $\tilde{\Pi}_1 = M_{6 \rightarrow 5} \Pi_1$ and $\tilde{\Pi}_2 = M_{6 \rightarrow 5} \Pi_2$ are two linearly independent solutions of $\tilde{\Pi}' = \tilde{\mathcal{A}} \tilde{\Pi}$ that are integrable on $\mathbb{R}_{>0}$.

On the other hand the characteristic polynomial of $\tilde{\mathcal{A}}$ is given by $\frac{4}{3(\alpha+\beta)v^5} X P_2 P_3$ so Lemma 5.7 implies that $\tilde{\mathcal{A}}$ has a single negative eigenvalue. This is a contradiction.

Hence $\Pi = c_2 a(\zeta_2) e^{\zeta_2 x} dx$ and we conclude using the same arguments as in the case $c < v < 2c$. \square

We now turn to the postponed proof of Lemma 5.7 and 5.9.

Proof of Lemma 5.7. (i) The leading coefficient and the constant term of P_2 have opposite signs.

(ii) When $v \in (c, 2c)$ we have

$$\begin{aligned} 2(2c+v)(2c-v)c &> 0 \\ -16\alpha c^2 - 8\beta c^2 + 2\alpha v^2 &= -2(2\sqrt{2}c+v)(2\sqrt{2}c-v)\alpha - 8c^2\beta < 0 \\ 2(5\alpha^2 + 5\alpha\beta + \beta^2)c &> 0 \\ -(2\alpha + \beta)(\alpha + \beta)\alpha &< 0 \end{aligned}$$

so $x < 0$ implies $P_3(x) < 0$.

(iii) Follows from the same argument as (ii).

(iv) When $v \in (2c, +\infty)$ we have

$$2(2c+v)(2c-v)c < 0, \quad 2(5\alpha^2 + 5\alpha\beta + \beta^2)c > 0, \quad -(2\alpha + \beta)(\alpha + \beta)\alpha < 0,$$

so by Descartes' rule of signs P_3 has a unique negative root.

(v) When $v \neq 2c$ the polynomial P_3 is a cubic with discriminant

$$\begin{aligned} \Delta &= 128\alpha^3\beta^3c^6 + 64\alpha^2\beta^4c^6 + 64\alpha^6(c+v)^2(c-v)^2v^2 + 128\alpha^5\beta c^4v^2 + 288\alpha^2\beta^4c^4v^2 \\ &\quad + 256\alpha\beta^5c^4v^2 + 64\beta^6c^4v^2 + 32\alpha^5\beta c^2v^4 + 404\alpha^4\beta^2c^2v^4 + 280\alpha^3\beta^3c^2v^4 \\ &\quad + 52\alpha^2\beta^4c^2v^4 + 96\alpha^5\beta v^6 + 32(2c^2 + v^2)\alpha^4\beta^2(c+v)^2(c-v)^2 > 0. \end{aligned}$$

When $v = 2c$ the polynomial P_3 has degree 2 and discriminant

$$\Delta = (9\alpha^4 + 10\alpha^3\beta + 3\alpha^2\beta^2 + 2\alpha\beta^3 + \beta^4)v^2 > 0.$$

(vi) Assume by contradiction that the negative roots of P_2 and P_3 are identical. Then the following resultant must vanish

$$\begin{aligned} \text{Res}_X(P_2, P_3) &= \left(\beta^2 (c^2 - v^2)^2 (c^2 + v^2) + 4\alpha c^4 v^2 + 4\alpha\beta c^2 v^4 \right) \\ &\quad \times (\alpha (4v^2 - c^2) + 2\beta (v^2 - c^2)) (\alpha + \beta)(\alpha - \beta)\alpha c. \end{aligned}$$

All the factors in the previous expression are positive except for $\alpha - \beta$. Hence $\alpha = \beta$ and

$$\begin{aligned} P_2 &= ((c^2 - v^2)X - 2\alpha c)(cX - \alpha), \\ P_3 &= 2((4c^2 - v^2)X^2 - 8X\alpha c + 3\alpha^2)(cX - \alpha). \end{aligned}$$

Hence P_2 and P_3 share the positive root $\frac{\alpha}{c}$ and since

$$\text{Res}_X\left(\frac{P_2}{cX - \alpha}, \frac{P_3}{cX - \alpha}\right) = 6\alpha^2 (c^2 + v^2)^2 > 0$$

there can be no shared negative root. This is a contradiction. \square

Proof of Lemma 5.9. Set $\mathcal{B} = -\zeta\mathcal{V} + \mathcal{Q}^t$ and denote

- r_k the k -th row of \mathcal{B} for $k = 1, \dots, 6$,
- \mathcal{B}_k the matrix obtained by replacing the k -th row of \mathcal{B} by the k -th basis vector e_k for $k = 1, \dots, 6$.

We now prove

$$\mathcal{B}_k \text{ invertible} \implies u_k := (\text{Adj } \mathcal{B}_k) e_k \text{ spans } \ker(\mathcal{B}) = (\text{Vect}(r_l)_l)^\perp.$$

If \mathcal{B}_k is invertible then

$$u_k = (\text{Adj } \mathcal{B}_k) e_k \implies \mathcal{B}_k u_k = \det(\mathcal{B}_k) e_k \implies u_k \neq 0 \text{ and } \langle r_l, u_k \rangle = 0 \text{ for } l \neq k.$$

Because \mathcal{B}_k is invertible we have that the vectors $(r_l)_{l \neq k}$ span a vector space of dimension 5. Thus $\ker(\mathcal{B})$ is of dimension at most one. It remains to show $u_k \in (\text{Vect}(r_l)_l)^\perp$ i.e. $\langle r_k, u_k \rangle = 0$. Because $\det(-X\mathcal{V} + \mathcal{Q}^t) = 8XP_2P_3$ Lemma 5.7 ensures that the family $(r_l)_l$ spans a vector space of dimension at most 5. We deduce $r_k \in \text{Vect}(r_l)_{l \neq k}$ and $\langle r_k, u_k \rangle = 0$.

Working with the resultants

$$\text{Res}(\det(\mathcal{B}_1), P_2), \quad \text{Res}(\det(\mathcal{B}_1), P_3), \quad \text{Res}(\det(\mathcal{B}_6), P_2), \quad \text{Res}(\det(\mathcal{B}_6), P_3),$$

as in the proof of Lemma 5.7 (v) and using $\zeta = \zeta_2$ (resp. $v \in (2c, +\infty)$ and $\zeta = \zeta_3$) yields $\det(\mathcal{B}_1) \neq 0$ and $\det(\mathcal{B}_6) \neq 0$. Hence

$$\begin{aligned} u_1 &= \begin{pmatrix} * \\ * \\ -2(4c^2\zeta^2 + 6cv\zeta^2 + 2v^2\zeta^2 - 6\alpha c\zeta - 2\beta c\zeta - 4\alpha v\zeta - 2\beta v\zeta + 2\alpha^2 + \alpha\beta)(c\zeta - \beta)\alpha\beta \\ -4(4c^2\zeta^2 + 6cv\zeta^2 + 2v^2\zeta^2 - 6\alpha c\zeta - 2\beta c\zeta - 4\alpha v\zeta - 2\beta v\zeta + 2\alpha^2 + \alpha\beta)(c\zeta - \alpha)\alpha^2 \\ 4(3c\zeta - 2\alpha - \beta)(c\zeta + v\zeta - \alpha)\alpha^2\beta \\ -(3c\zeta - 2\alpha - \beta)\alpha^2\beta^2 \end{pmatrix}, \\ u_6 &= \begin{pmatrix} -(3c\zeta - 2\alpha - \beta)\alpha^2\beta^2 \\ 4(3c\zeta - 2\alpha - \beta)(c\zeta - v\zeta - \alpha)\alpha^2\beta \\ -2(4c^2\zeta^2 - 6cv\zeta^2 + 2v^2\zeta^2 - 6\alpha c\zeta - 2\beta c\zeta + 4\alpha v\zeta + 2\beta v\zeta + 2\alpha^2 + \alpha\beta)(c\zeta - \beta)\alpha\beta \\ -4(4c^2\zeta^2 - 6cv\zeta^2 + 2v^2\zeta^2 - 6\alpha c\zeta - 2\beta c\zeta + 4\alpha v\zeta + 2\beta v\zeta + 2\alpha^2 + \alpha\beta)(c\zeta - \alpha)\alpha^2 \\ * \\ * \end{pmatrix}, \end{aligned}$$

both span $\ker(\mathcal{B})$ where we omit the entries marked by $*$ for conciseness. Hence u_1 and u_6 are collinear eigenvectors corresponding to the eigenvalue ζ and the result follows from

$$4c^2\zeta^2 + 6cv\zeta^2 + 2v^2\zeta^2 - 6\alpha c\zeta - 2\beta c\zeta - 4\alpha v\zeta - 2\beta v\zeta + 2\alpha^2 + \alpha\beta \neq 0$$

which can again be shown by computing a resultant. \square

5.2.3 Instantaneous harmonic process

We now turn to the invariant probability of the instantaneous harmonic process. The key step of writing down and solving Fokker-Planck in the bulk was already done in [BMR⁺20]. Hence, we only have to determine the boundary Dirac masses arising from the ‘stickiness’ of our process but we include a sketch of proof for the sake of completeness. An important feature of the instantaneous harmonic process is that once it enters the set $[0, v/\mu] \times \Sigma$ it stays there indefinitely. In particular the invariant measure has compact support.

Notation 5.11. *The following special functions are needed to write down the invariant probability of the instantaneous harmonic process*

- ${}_pF_q$ the hypergeometric function (see equation 16.2.1 in [Olw10]),
- ${}_p\tilde{F}_q$ the regularized hypergeometric function given by

$${}_p\tilde{F}_q(a_1, \dots, a_p; b_1, \dots, b_q; z) := \frac{{}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; z)}{\Gamma(b_1) \cdots \Gamma(b_q)},$$

- K the Legendre’s complete elliptic integral of the first kind (see equation 19.2.8 in [Olw10]),
- $G_{m,n}^{p,q}$ the Meijer G -function (see equation 16.17.1 in [Olw10]).

The parameter $b := 2\frac{\omega}{\mu}$ will determine the shape of the invariant measure.

Proposition 5.12 (Invariant measure of the instantaneous harmonic process). *Let π be the invariant measure of the instantaneous harmonic process. If $b \neq 1$ then the x -marginal of π is*

$$d_0\delta_0 + Cp(x)dx$$

where

$$d_0 = \left(\frac{b\Gamma\left(\frac{1}{2} - \frac{b}{2}\right)\Gamma\left(\frac{b}{2} + 1\right)\Gamma\left(b + \frac{1}{2}\right) {}_3\tilde{F}_2\left(\frac{3}{2}, 1 - \frac{b}{2}, \frac{b+2}{2}, \frac{b+3}{2}, \frac{b+4}{2}; 1\right)}{2\sqrt{\pi}} + 2b {}_3F_2\left(\frac{1}{2}, \frac{1}{2} - b, 1 - \frac{b}{2}; \frac{3}{2}, \frac{1}{2} - \frac{b}{2}; 1\right) + 2 \right)^{-1}$$

and

$$p(x) = {}_2\tilde{F}_1\left(\frac{3}{2} - b, 1 - \frac{b}{2}; \frac{3-b}{2}; \frac{x^2\mu^2}{v_0^2}\right) - \frac{\Gamma\left(b - \frac{1}{2}\right)\left(\frac{\mu^2 x^2}{v_0^2}\right)^{\frac{b-1}{2}} {}_2\tilde{F}_1\left(\frac{1}{2}, 1 - \frac{b}{2}; \frac{b+1}{2}; \frac{x^2\mu^2}{v_0^2}\right)}{\sqrt{\pi}}$$

as well as

$$C = \frac{\sqrt{\pi}b(2b-1)\mu\Gamma\left(\frac{3}{2} - \frac{b}{2}\right)}{v_0} \left(2\sqrt{\pi}(b-1) \left({}_3F_2\left(\frac{1}{2}, \frac{1}{2} - b, 1 - \frac{b}{2}; \frac{3}{2}, \frac{1}{2} - \frac{b}{2}; 1\right) + 1 \right) - b\Gamma\left(\frac{3}{2} - \frac{b}{2}\right)\Gamma\left(\frac{b}{2} + 1\right)\Gamma\left(b + \frac{1}{2}\right) {}_3\tilde{F}_2\left(\frac{3}{2}, 1 - \frac{b}{2}, \frac{b+2}{2}; \frac{b+3}{2}, \frac{b+4}{2}; 1\right) \right)^{-1}.$$

If $b = 1$ then the x -marginal of the invariant measure is

$$\frac{8}{8 + \pi^2} \delta_0 + \frac{4\mu}{v_0(8 + \pi^2)} K \left(1 - \frac{x^2 \mu^2}{v_0^2} \right).$$

Sketch of proof. Following [BMR⁺20] we set $b = 2\frac{\gamma}{\mu}$ as well as

$$P = \pi_2 + \pi_0 + \pi_{-2}, \quad Q = \pi_2 + \pi_{-2}, \quad R = \pi_2 - \pi_{-2},$$

where π_2, π_0, π_{-2} are the Borel measures such that

$$\pi(1_{\{x>0\}}f) = \sum_{\sigma \in \Sigma} \int_0^{+\infty} f(x, \sigma) d\pi_\sigma(x).$$

It is shown in [BMR⁺20] that when $b \neq 1$

$$\begin{aligned} P &= C_1 \left({}_2F_1 \left(\frac{3}{2} - b, 1 - \frac{b}{2}; \frac{3-b}{2}; \frac{x^2 \mu^2}{v_0^2} \right) + \right. \\ &\quad \left. \frac{2\Gamma\left(\frac{3-b}{2}\right) \Gamma\left(b + \frac{1}{2}\right) \left(\frac{\mu^2 x^2}{v_0^2}\right)^{\frac{b-3}{2}+1} {}_2F_1\left(\frac{b-3}{2} - b + \frac{5}{2}, \frac{b-3}{2} - \frac{b}{2} + 2; \frac{b-3}{2} + 2; \frac{x^2 \mu^2}{v_0^2}\right)}{\sqrt{\pi}(1-2b)\Gamma\left(\frac{b+1}{2}\right)} \right) dx, \\ Q &= C_1 \left(\frac{(1-b) {}_2F_1\left(\frac{1}{2} - b, 1 - \frac{b}{2}, \frac{1-b}{2}; \frac{x^2 \mu^2}{v_0^2}\right)}{1-2b} + \right. \\ &\quad \left. \frac{2\Gamma\left(\frac{3-b}{2}\right) \Gamma\left(b + \frac{1}{2}\right) \left(\frac{\mu^2 x^2}{v_0^2}\right)^{\frac{b-1}{2}+1} {}_2F_1\left(\frac{b-1}{2} - b + \frac{3}{2}, \frac{b-1}{2} - \frac{b}{2} + 2; \frac{b-1}{2} + 2; \frac{x^2 \mu^2}{v_0^2}\right)}{\sqrt{\pi}(1-2b)(b+1)\Gamma\left(\frac{b+1}{2}\right)} \right) dx, \end{aligned}$$

where C_1 is a constant that remains to be fixed. It is also shown that $R = \frac{\mu x}{v_0} P$ so that

$$\pi_2 = \frac{1}{2} \left(Q + \frac{\mu x}{v_0} P \right), \quad \pi_0 = P - Q, \quad \pi_{-2} = \frac{1}{2} \left(Q - \frac{\mu x}{v_0} P \right).$$

Considering the σ -marginal yields

$$\pi(\mathbb{R}_+ \times \{2\}) = \frac{1}{4}, \quad \pi(\mathbb{R}_+ \times \{0\}) = \frac{1}{2}, \quad \pi(\mathbb{R}_+ \times \{-2\}) = \frac{1}{4}.$$

and $\pi(\{(0, 2)\}) = 0$ follows from the usual ergodicity argument. Denoting $C_2 = \pi(\{(0, 0)\})$ and $C_3 = \pi(\{(0, -2)\})$ this translates to

$$\left. \begin{aligned} \frac{1}{2} \left(Q + \frac{\mu x}{v_0} P \right) (\mathbb{R}_{>0}) &= \frac{1}{4} \\ C_2 + (P - Q) (\mathbb{R}_{>0}) &= \frac{1}{2} \\ C_3 + \left(Q - \frac{\mu x}{v_0} P \right) (\mathbb{R}_{>0}) &= \frac{1}{4} \end{aligned} \right\} \quad (*)$$

which allows us to determine C_1, C_2 and C_3 .

It is shown in [BMR⁺20] that when $b = 1$

$$P = \frac{2C_1}{\pi} K \left(1 - \frac{x^2 \mu^2}{v_0^2} \right) dx, \quad Q = C_1 G_{2,2}^{2,0} \left(\frac{\mu^2 x^2}{v_0^2} \middle| \frac{1}{2}, \frac{3}{2} \right) dx$$

where C_1 is a constant that remains to be fixed. We can again use (*) to compute the C_i and deduce the desired result. \square

Remark 5.13. *The identities*

$$\pi_2 = \frac{1}{2} \left(Q + \frac{\mu x}{v_0} P \right), \quad \pi_0 = P - Q, \quad \pi_{-2} = \frac{1}{2} \left(Q - \frac{\mu x}{v_0} P \right)$$

fully determine π (not just its x -marginal). However, the formulae for the complete invariant probability are somewhat unwieldy and are thus omitted.

Having found explicit representations for the invariant probabilities, we turn to the study of relaxation towards them. This will answer the crucial question: when does the asymptotic behavior take over?

5.3 Convergence towards the invariant measure

This section is dedicated to the proof of Theorem 5.5.

5.3.1 Convergence of the instantaneous and finite linear process

In this section we give quantitative upper bounds for $\|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}}$ using coupling and large deviation techniques that allow for a unified treatment of the instantaneous and finite linear process. In fact, our computations could be extended to relative velocity transition rates other than 5.2a and 5.2b.

The key is to couple the processes in a ‘synchronous’ manner, meaning that after some time the velocities of the two copies $X(t) = (x(t), \sigma(t))$ and $\tilde{X}(t) = (\tilde{x}(t), \tilde{\sigma}(t))$ are always identical. Once the velocities are the same, the order between x and \tilde{x} is preserved. If, for example, $\sigma(t) = \tilde{\sigma}(t)$ and $x(t) \leq \tilde{x}(t)$ for all $t \geq T$ then $\tilde{x}(t_0) = 0$ for some $t_0 \geq T$ implies

$$0 \leq x(t_0) \leq \tilde{x}(t_0) = 0 \implies X(t_0) = \tilde{X}(t_0).$$

Hence it suffices to prove the exponential decay of $\mathbb{P}(\tau_0 > t)$ where $\tau_0 = \inf\{t \geq 0 : x(t) = 0\}$. The last crucial observation is

$$\tau_0 > t \implies x(0) - 2ct + v \int_0^t \sigma(s) ds > 0 \implies \frac{1}{2t} \int_0^t \sigma(s) ds > \frac{c}{v} - \frac{x(0)}{2vt}$$

so we can use large deviations results for additive functionals of Markov processes to obtain exponential decay.

Notation 5.14. *We denote*

- \mathcal{Q} the discrete generator of 5.2a (resp. 5.2b) and $\pi_{\mathcal{Q}}$ the associated invariant probability,
- $\tilde{\mathcal{V}} = \frac{1}{2} (\sigma 1_{\{\sigma=\tilde{\sigma}\}})_{\sigma, \tilde{\sigma} \in \Sigma}$ the matrix with the normalized velocities on the diagonal,
- $\mathcal{S} = (\pi_{\mathcal{Q}}(\sigma) 1_{\{\sigma=\tilde{\sigma}\}})_{\sigma, \tilde{\sigma} \in \Sigma}$ the matrix with the invariant measure on the diagonal,
- $\lambda_{\mathcal{Q}}$ the spectral gap of \mathcal{Q} ,
- $\lambda_{\pi} := \lim_{L \rightarrow +\infty} -\frac{1}{L} \log \pi(\{x > L\})$ the exponential decay rate of the invariant measure.

The quantitative aspect of our results comes from non-asymptotic large deviation bounds for additive functionals of Markov processes. Our main tool will be the following lemma which is an immediate consequence of [Wu00, Theorem 1].

Lemma 5.15. *If σ follows the transition rates 5.2a (resp. 5.2b) and $R > 0$ then*

$$\mathbb{P}_{\mu} \left(\frac{1}{2t} \int_0^t \sigma(s) ds \geq R \right) \leq \left\| \frac{\mu}{\pi_{\mathcal{Q}}} \right\|_{L^2(\pi_{\mathcal{Q}})} e^{-tI(R)}$$

where I is Legendre transform of

$$\Lambda(u) = \sup_{\langle \mathcal{S}f, f \rangle = 1} \langle \mathcal{S}(\mathcal{Q} + u\tilde{\mathcal{V}})f, f \rangle.$$

We can now state the main result of this section.

Theorem 5.16 (Geometric ergodicity). *(i) The instantaneous (resp. finite) linear process has a unique invariant measure π .*

(ii) For all $\lambda \in (0, I(\frac{c}{v}))$ there exist $C_1, C_2 > 0$ independent of t and (x, σ) such that

$$\begin{aligned} & \|\delta_{(x,\sigma)}P_t - \pi\|_{\text{TV}} \\ & \leq \left(e^{\frac{\lambda}{2c}x} + \frac{\lambda}{I(\frac{c}{v}) - \lambda} \frac{e^{\left(\frac{I'(\frac{c}{v})}{2v} - \frac{I(\frac{c}{v}) - \lambda}{2c}\right)x}}{\pi_{\mathcal{Q}}(\{\sigma\})} + 2 + \pi(f) + 2t \left(\max_{x \in \Sigma} \mathcal{L}f(0, \sigma) \right) \right) e^{-\frac{\lambda_{\mathcal{Q}}\lambda}{\lambda_{\mathcal{Q}} + \lambda}t} \end{aligned}$$

where $I'(\frac{c}{v})$ is any subderivative, \mathcal{L} is the generator of the process and

$$f(x, \sigma) = \mathbb{E}_{(x,\sigma)} \left[e^{\lambda \inf\{t \geq 0 : x(t) = 0\}} \right].$$

Remark 5.17. *In the instantaneous linear case $I(R) = 2\omega(1 - \sqrt{1 - R^2})$ so $I'(\frac{c}{v})$ exists in the usual sense. We state Theorem 5.16 in terms of subderivatives of the convex function I to emphasize the fact that it is not necessary to assume that it is differentiable, which might not be the case if we considered transition rates other than 5.2a and 5.2b.*

Remark 5.18. *The prefactor in Theorem 5.16 (ii) has an exponential x -dependence. We would expect the same x -dependence if we turned to Foster-Lyapunov techniques [MT93].*

The following lemma, the proof of which is delayed until the end of the section, makes the exponential decay rates explicit in terms of model parameters.

Proposition 5.19 (Decay rates as a function of model parameters). *(i) For the instantaneous linear process $\lambda_{\mathcal{Q}} = 2\omega$ and $I(R) = 2\omega(1 - \sqrt{1 - R^2})$.*

(ii) For the finite linear process $\lambda_{\mathcal{Q}} = \alpha$ and

$$I(R) \geq \frac{3 - \sqrt{5}}{4} \min\left(\alpha R, \alpha \left(1 + \frac{\alpha}{\beta}\right) R\right).$$

Combining Theorem 5.16 and Proposition 5.19 and using the inequality $\frac{ab}{a+b} \geq \frac{1}{2} \min(a, b)$ for $a, b > 0$ yields

$$\underline{\lim} -\frac{1}{t} \log \|\delta_{(x,\sigma)}P_t - \pi\|_{\text{TV}} \geq \omega \left(1 - \sqrt{1 - \frac{c^2}{v^2}}\right) \geq \frac{\omega}{2} \frac{c^2}{v^2}$$

in the instantaneous case and

$$\underline{\lim} -\frac{1}{t} \log \|\delta_{(x,\sigma)}P_t - \pi\|_{\text{TV}} \geq \frac{3 - \sqrt{5}}{8} \min\left(\alpha \frac{c}{v}, \alpha \left(1 + \frac{\alpha}{\beta}\right) \frac{c^2}{v^2}\right)$$

in the finite case, which are the lower bounds from Theorem 5.5. The rest of the section is dedicated to proving Theorem 5.16 and Proposition 5.19. We start by deducing quantitative bounds for $\mathbb{E}_{(x,\sigma)} [e^{\lambda\tau_0}]$ from Lemma 5.15.

Lemma 5.20 (Lyapunov function). *For all $\lambda \in (0, I(\frac{c}{v}))$ we have*

$$\mathbb{E}_{(x,\sigma)} [e^{\lambda\tau_0}] \leq e^{\frac{\lambda}{2c}x} + \frac{\lambda}{I(\frac{c}{v}) - \lambda} \frac{e^{\left(\frac{I'(\frac{c}{v})}{2v} - \frac{I(\frac{c}{v}) - \lambda}{2c}\right)x}}{\pi_{\mathcal{Q}}(\{\sigma\})}$$

where $\tau_0 = \inf\{t > 0 : x(t) = 0\}$ and $I'(\frac{c}{v})$ is any subderivative.

Proof. We have

$$\tau_0 > t \implies x(0) - 2ct + v \int_0^t \sigma(s) ds > 0 \implies \frac{1}{2t} \int_0^t \sigma(s) ds > \frac{c}{v} - \frac{x(0)}{2vt},$$

so that

$$\begin{aligned} \mathbb{E}_{(x,\sigma)} [e^{\lambda\tau_0}] &= 1 + \int_0^{+\infty} \lambda e^{\lambda t} \mathbb{P}_{(x,\sigma)}(\tau_0 > t) dt \\ &\leq 1 + \int_0^{\frac{x}{2c}} \lambda e^{\lambda t} dt + \int_{\frac{x}{2c}}^{+\infty} \lambda e^{\lambda t} \mathbb{P}_{(x,\sigma)} \left(\frac{1}{2t} \int_0^t \sigma(s) ds \geq \frac{c}{v} - \frac{x}{2vt} \right) ds \\ &\leq e^{\frac{\lambda}{2c}x} + \int_{\frac{x}{2c}}^{+\infty} \lambda e^{\lambda t} \left\| \frac{\delta_\sigma}{\pi_\mathcal{Q}} \right\|_{L^2(\pi_\mathcal{Q})} e^{-I\left(\frac{c}{v} - \frac{x}{2vt}\right)t} dt < +\infty \end{aligned}$$

using Lemma 5.15 for the last inequality. Because $I' \left(\frac{c}{v} \right)$ is a subderivative

$$I \left(\frac{c}{v} - \frac{x}{2vt} \right) \geq I \left(\frac{c}{v} \right) - I' \left(\frac{c}{v} \right) \frac{x}{2vt} \iff -tI \left(\frac{c}{v} - \frac{x}{2vt} \right) \leq -tI \left(\frac{c}{v} \right) + I' \left(\frac{c}{v} \right) \frac{x}{2v}$$

so the desired result follows from $\left\| \frac{\delta_\sigma}{\pi_\mathcal{Q}} \right\|_{L^2(\pi_\mathcal{Q})} = \frac{1}{\pi_\mathcal{Q}(\{\sigma\})}$. \square

The following lemma bounds the exponential rates appearing in Lemma 5.20. In particular, it ensures that $(x, \sigma) \mapsto \mathbb{E}_{(x,\sigma)} [e^{\lambda\tau_0}]$ is π -integrable (this is also a consequence of [MT93, Theorem 4.3]).

Lemma 5.21. *We have*

$$\frac{I \left(\frac{c}{v} \right)}{2c} \leq \frac{I' \left(\frac{c}{v} \right)}{2v} \leq \lambda_\pi$$

where $I' \left(\frac{c}{v} \right)$ is any subderivative.

Proof. By definition of $I' \left(\frac{c}{v} \right)$ we have $I(R) \geq I \left(\frac{c}{v} \right) + (R - \frac{c}{v}) I' \left(\frac{c}{v} \right)$ for all $R \in \mathbb{R}$. Because $I(0) = 0$ this is equivalent to the first inequality when $R = 0$.

We now turn in to the second inequality, which we prove in the instantaneous case to simplify notation. If we set $\tau_{(0,2)} := \inf\{t > 0 : (x(t), \sigma(t)) = (0, 2)\}$ then we have the representation

$$\pi(A) = \frac{\mathbb{E}_{(0,2)} \left[\int_0^{\tau_{(0,2)}} \mathbf{1}_{\{X(s) \in A\}} dt \right]}{\mathbb{E}_{(0,2)} [\tau_{(0,2)}]}.$$

for all measurable sets A (see [Kha60, Theorem 2.1] and its proof). In particular

$$\pi(\{x > a\}) = \frac{\mathbb{E}_{(0,2)} \left[\int_0^{\tau_{(0,2)}} \mathbf{1}_{\{x(t) > a\}} dt \right]}{\mathbb{E}_{(0,2)} [\tau_{(0,2)}]}.$$

So using

$$x(s) > a \text{ and } t < \tau_{(0,2)} \implies -2ct + v \int_0^t \sigma(s) ds > a \implies \frac{1}{2t} \int_0^t \sigma(s) ds > \frac{c}{v} + \frac{a}{2tv}$$

and Lemma 5.15 we deduce

$$\begin{aligned} \mathbb{E}_{(0,2)} \left[\int_0^{\tau_{(0,2)}} \mathbf{1}_{\{x(t) > a\}} dt \right] &= \int_0^{+\infty} \mathbb{P}_{(0,2)}(x(t) > a \text{ and } t < \tau_{(0,2)}) dt \\ &\leq \left\| \frac{\delta_2}{\pi_\mathcal{Q}} \right\|_{L^2(\pi_\mathcal{Q})} \int_0^{+\infty} e^{-tI\left(\frac{c}{v} + \frac{a}{2tv}\right)} dt. \end{aligned}$$

Because $I'(\frac{c}{v})$ is a subderivative

$$I\left(\frac{c}{v} + \frac{a}{2tv}\right) \geq I\left(\frac{c}{v}\right) + I'\left(\frac{c}{v}\right) \frac{a}{2tv} \implies -tI\left(\frac{c}{v}\right) - I'\left(\frac{c}{v}\right) \frac{a}{2v} \geq -tI\left(\frac{c}{v} + \frac{a}{2tv}\right)$$

and thus

$$\int_0^{+\infty} e^{-tI(\frac{c}{v} + \frac{a}{2tv})} dt \leq \frac{e^{-\frac{I'(\frac{c}{v})}{2v}a}}{I(\frac{c}{v})}.$$

Hence the decay rate of $\pi(\{x > a\})$ is at least $\frac{I'(\frac{c}{v})}{2v}$ which is precisely the second inequality. \square

This last preliminary lemma constructs the synchronous coupling and shows that it has two important properties: after a short period of time (this is made precise in the lemma) the velocities of the two copies $X(t) = (x(t), \sigma(t))$ and $\tilde{X}(t) = (\tilde{x}(t), \tilde{\sigma}(t))$ coincide and the order between x and \tilde{x} is preserved.

Lemma 5.22 (Linear synchronous coupling). *For all $(\sigma_0, x_0), (\tilde{x}_0, \tilde{\sigma}_0) \in E$ there exists a coupling $\mathbb{Q}_{(\sigma_0, x_0), (\tilde{x}_0, \tilde{\sigma}_0)}$ of $X(t) = (x(t), \sigma(t))$ and $\tilde{X}(t) = (\tilde{x}(t), \tilde{\sigma}(t))$ such that*

(i) X and \tilde{X} are instantaneous (resp. finite) linear processes with initial states (x_0, σ_0) and $(\tilde{x}_0, \tilde{\sigma}_0)$,

(ii) $\mathbb{P}(\sigma(s) = \tilde{\sigma}(s) \text{ for all } s \geq t) \geq 1 - 2e^{-\lambda_{\mathcal{Q}}t}$ where $\lambda_{\mathcal{Q}}$ is the spectral gap of \mathcal{Q} ,

(iii) if $\sigma(s) = \tilde{\sigma}(s)$ for all $s \geq t$ then

$$\begin{aligned} x(t) \leq \tilde{x}(t) &\implies x(s) \leq \tilde{x}(s) \text{ for all } s \geq t, \\ \tilde{x}(t) \leq x(t) &\implies \tilde{x}(s) \leq x(s) \text{ for all } s \geq t, \end{aligned}$$

meaning that the ordering of x and \tilde{x} is preserved.

Proof. (i) Let $f : \{\pm 1\}^2 \rightarrow \Sigma$ (resp. $f : \{0, \pm 1\}^2 \rightarrow \Sigma$) be defined by $f(\sigma_1, \sigma_2) = \sigma_2 - \sigma_1$ (resp. $f(\sigma_1, \sigma_2) = \sigma_2 - \sigma_1$ when $\sigma_1 \neq \sigma_2$ as well as $f(\pm 1, \pm 1) = 0_{\pm}$ and $f(0, 0) = 0_0$) and choose $\sigma_{1,0}, \sigma_{2,0}, \tilde{\sigma}_{1,0}, \tilde{\sigma}_{2,0}$ such that $\sigma_0 = f(\sigma_{1,0}, \sigma_{2,0})$ and $\tilde{\sigma}_0 = f(\tilde{\sigma}_{1,0}, \tilde{\sigma}_{2,0})$.

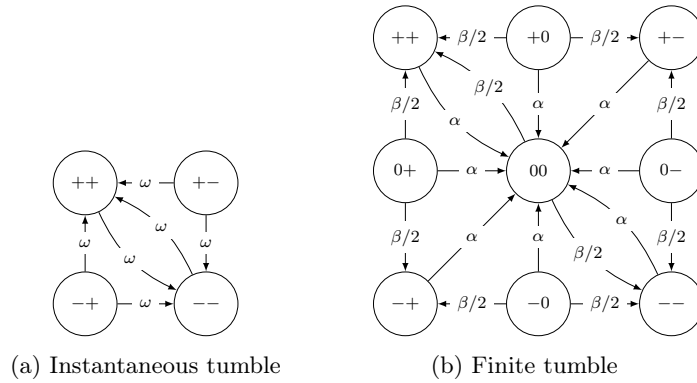


Figure 5.4: Single velocity coupling

Let $(\sigma_1, \tilde{\sigma}_1)$ and $(\sigma_2, \tilde{\sigma}_2)$ be independent Markov jump processes with the transition rates of figure 5.4a (resp. 5.4b) and initial states $(\sigma_{1,0}, \tilde{\sigma}_{1,0})$ and $(\sigma_{2,0}, \tilde{\sigma}_{2,0})$. Then it can be checked using [BY93, Theorems 2.3 and 2.4] that $\sigma = f(\sigma_1, \sigma_2)$ and $\tilde{\sigma} = f(\tilde{\sigma}_1, \tilde{\sigma}_2)$ are Markov jump processes with the transition rates of figure 5.2a (resp. 5.2b).

Finally, recursively construct

$$\begin{aligned} x(t) &= \max(0, x_0 + (v\sigma_0 - 2c)t) \text{ for } t \leq T_1, \\ x(t) &= \max(0, x(T_1) + (v\sigma(T_1) - 2c)(t - T_1)) \text{ for } t \in [T_1, T_2], \text{ etc,} \end{aligned}$$

and similarly for $\tilde{x}(t)$ and set

$$X(t) = (x(t), \sigma(t)), \quad \tilde{X}(t) = (\tilde{x}(t), \tilde{\sigma}(t)).$$

(ii) Recall that $\lambda_{\mathcal{Q}} = 2\omega$ in the instantaneous case and $\lambda_{\mathcal{Q}} = \alpha$ in the finite case by Proposition 5.19. Set $\tau_{\sigma}^i = \inf\{t \geq 0 : \sigma_i(t) = \tilde{\sigma}_i(t)\}$ and observe that $\sigma(s) = \tilde{\sigma}(s)$ for all $s \geq \max\{\tau_{\sigma}^1, \tau_{\sigma}^2\}$ hence it is enough to show $\mathbb{P}(\tau_{\sigma}^i > t) \leq e^{-\lambda_{\mathcal{Q}}t}$.

In the instantaneous case, if $\sigma_{i,0} \neq \tilde{\sigma}_{i,0}$ then we have $\sigma_i = \tilde{\sigma}_i$ after the first jump of $(\sigma_i, \tilde{\sigma}_i)$ which is exponentially distributed with parameter 2ω (see figure 5.4a). Hence $\mathbb{P}(\tau_{\sigma}^i > t) \leq e^{-2\omega t}$.

In the finite case, notice that all states other than $(0, 0)$ have a transition with rate α to the state $(0, 0)$ (see figure 5.4b). Hence, if $(\sigma_{i,0}, \tilde{\sigma}_{i,0}) \neq (0, 0)$ then $\tau_{00}^i := \inf\{t \geq 0 : \sigma_i(t) = \tilde{\sigma}_i = 0\}$ is exponentially distributed with parameter α . Hence $\mathbb{P}(\tau_{\sigma}^i > t) \leq \mathbb{P}(\tau_{00}^i > t) \leq e^{-\alpha t}$.

(iii) Immediately follows from the recursive construction of x and \tilde{x} and the fact that

$$x \mapsto \max(0, x + (v\sigma(T_k) - 2c)(t - T_k))$$

is always a nondecreasing function. \square

Remark 5.23. *The coupling constructed in Lemma 5.22 is not jointly Markovian. Indeed, if $(\sigma_i, \tilde{\sigma}_i)$ has the transition rates 5.4b then, even though $\sigma = f(\sigma_1, \sigma_2)$ and $\tilde{\sigma} = f(\tilde{\sigma}_1, \tilde{\sigma}_2)$ are Markovian, the joint process $(\sigma, \tilde{\sigma})$ is not Markovian. We could fix this by directly coupling the velocities σ and $\tilde{\sigma}$ but then we would have to describe a Markov jump process of size 36 in the finite linear case which would not be particularly illuminating.*

Proof of Theorem 5.16. (i) To simplify notations, we prove the proposition only for the instantaneous process. Lemma 5.20 and [MT93, Theorem 4.1] applied to the closed petite set $C = \{(0, 2)\}$, the test function $f(x, \sigma) = 1$ and any shift $\delta > 0$ imply the existence of a unique invariant probability measure.

(ii) Let $(x, \sigma) \in E$ be arbitrary but fixed and set

$$\mathbb{P} = \int \mathbb{Q}_{(x, \sigma), (\tilde{x}, \tilde{\sigma})} d\pi(\tilde{x}, \tilde{\sigma})$$

where $\mathbb{Q}_{(x, \sigma), (\tilde{x}, \tilde{\sigma})}$ is the coupling constructed in Lemma 5.22. We denote $X = (x(t), \sigma(t))$ (resp. $\tilde{X} = (\tilde{x}(t), \tilde{\sigma}(t))$) the first (resp. second) marginal of \mathbb{P} and observe that this is a coupling of two instantaneous (resp. finite) linear processes with initial distributions $\delta_{(x, \sigma)}$ and π .

Fix $\lambda \in (0, I(\frac{c}{v}))$ and set $\theta = \frac{\lambda}{\lambda_{\mathcal{Q}} + \lambda}$. For any time $T \in \mathbb{R}_+$ define $\tau_0^T := \{t \geq T : x(t) = 0\}$ and $\tilde{\tau}_0^T := \{t \geq T : \tilde{x}(t) = 0\}$. If

$$\sigma(s) = \tilde{\sigma}(s) \text{ for all } s \geq T, \quad x(T) \leq \tilde{x}(T), \quad \tilde{x}(s_0) = 0 \text{ for some } s_0 \geq T,$$

then, because the ordering is preserved,

$$0 \leq x(s_0) \leq \tilde{x}(s_0) = 0 \implies x(s_0) = \tilde{x}(s_0)$$

and similarly if $\tilde{x}(T) \leq x(T)$ and $x(s_0) = 0$ for some $s_0 \geq T$. This implies

$$\begin{aligned} \|\delta_{(x, \sigma)} P_t - \pi\|_{\text{TV}} &\leq \mathbb{P}(\sigma(s) = \tilde{\sigma}(s) \text{ for all } s \geq \theta t) + \mathbb{P}(\tau_0^{\theta t} > t) + \mathbb{P}(\tilde{\tau}_0^{\theta t} > t) \\ &\leq 2e^{-\frac{\lambda_{\mathcal{Q}}\lambda}{\lambda_{\mathcal{Q}} + \lambda}t} + \mathbb{P}(\tau_0^{\theta t} > t) + \mathbb{P}(\tilde{\tau}_0^{\theta t} > t) \end{aligned}$$

using Lemma 5.22 for the second inequality. The Markov property yields

$$\mathbb{P}(\tau_0^{\theta t} > t) = \mathbb{E}_{(x,\sigma)} [\mathbb{P}_{X(\theta t)}(\tau_0^0 > (1-\theta)t)] \leq \mathbb{E}_{(x,\sigma)} \left[\frac{\mathbb{E}_{X(\theta t)}[e^{\lambda\tau_0^0}]}{e^{\lambda(1-\theta)t}} \right] = e^{-\frac{\lambda\mathcal{Q}\lambda}{\lambda\mathcal{Q}+\lambda}t} \mathbb{E}_{(x,\sigma)}[P_{\theta t}f]$$

where $f(X_0) = \mathbb{E}_{X_0}[e^{\lambda\tau_0^0}]$. We have

$$\mathcal{L}f(x, \sigma) + \lambda f(x, \sigma) = 0 \implies \mathcal{L}f(x, \sigma) = -\lambda f(x, \sigma) \leq 0 \text{ for all } (x, \sigma) \in \mathbb{R}_{>0} \times \Sigma.$$

Hence $\mathcal{L}f$ is upper bounded and we have $P_t f = f + \int_0^t P_s \mathcal{L}f ds \leq f + Ct$ for

$$C = \max_{\sigma \in \Sigma} \mathcal{L}f(0, \sigma).$$

Combining this with the previous inequality we get

$$\mathbb{P}(\tau_0^{\theta t} > t) \leq (f(x, \sigma) + Ct) e^{-\frac{\lambda\mathcal{Q}\lambda}{\lambda\mathcal{Q}+\lambda}t}$$

and the same arguments show

$$\mathbb{P}(\tilde{\tau}_0^{\theta t} > t) \leq (\pi(f) + Ct) e^{-\frac{\lambda\mathcal{Q}\lambda}{\lambda\mathcal{Q}+\lambda}t}.$$

The result now follows from replacing $f(x, \sigma)$ by the upper bound in Lemma 5.20 and combining Lemma 5.20 and Lemma 5.21 to see $\pi(f) < +\infty$. \square

Proof of Proposition 5.19. (i) Recall that in the instantaneous case, we have

$$\mathcal{Q} = \begin{pmatrix} -2\omega & 2\omega & 0 \\ \omega & -2\omega & \omega \\ 0 & 2\omega & -2\omega \end{pmatrix}, \quad \tilde{\mathcal{V}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \mathcal{S} = \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix}.$$

The eigenvalues of \mathcal{Q} are 0, -2ω and -4ω so $\lambda_{\mathcal{Q}} = 2\omega$. Notice that $\mathcal{A} := \mathcal{Q} + u\tilde{\mathcal{V}}$ is symmetric w.r.t. the scalar product $\langle \mathcal{S}\cdot, \cdot \rangle$ (i.e. $\mathcal{A} = \mathcal{S}^{-1}\mathcal{A}^t\mathcal{S}$) so that

$$\Lambda(u) = \sup_{\langle \mathcal{S}f, f \rangle = 1} \langle \mathcal{S}\mathcal{A}f, f \rangle = -2\omega + \sqrt{u^2 + 4\omega^2}$$

is the largest eigenvalue of \mathcal{A} and the result follows by taking the Legendre transform.

(ii) An explicit computation shows that the second eigenvalue of \mathcal{Q} is $-\alpha$ and thus $\lambda_{\mathcal{Q}} = \alpha$. By [Lez01, Theorem 1.1 and Remark 1.2]

$$I(R) \geq \alpha \frac{2(1+r)R^2}{\left(1 + \sqrt{1 + 4(1+r)R}\right)^2} =: \tilde{I}(R)$$

where $r = \frac{\alpha}{\beta}$ and, with the help of a computer algebra system, it can be seen that

$$\inf_{(1+r)R \geq 1} \frac{\tilde{I}(R)}{\alpha R} = \inf_{(1+r)R \leq 1} \frac{\tilde{I}(R)}{\alpha(1+r)R^2} = \frac{3 - \sqrt{5}}{4}.$$

\square

5.3.2 Convergence of the instantaneous harmonic process

In the harmonic case, the underlying deterministic dynamics are contracting in the sense that for any fixed $\sigma_0 \in \Sigma$ we have

$$\partial_t x = -\mu x + v\sigma_0 \text{ and } \partial_t \tilde{x} = -\mu \tilde{x} + v\sigma_0 \implies x(t) - \tilde{x}(t) = e^{-\mu t}(x(0) - \tilde{x}(0)).$$

The random switching of σ and the jamming interactions preserve this contractivity and hence the instantaneous harmonic process displays exponential convergence towards its invariant measure in a Wasserstein-type distance.

The same kind of contractivity is exploited in [BLBMZ12] to obtain quantitative bounds for the convergence of processes randomly switching between contracting deterministic dynamics. These results can also be extended to random switching between diffusions with contracting drifts (see [CH15, Section 4.1]). The following distance, which combines the Wasserstein distance of the x -marginal and the total variation distance for the σ -marginal, is naturally contracting for the instantaneous harmonic process.

Definition 5.24 (Mixed distance [BLBMZ12, Definition 1.4]). *For all probability laws $\eta, \tilde{\eta}$ on $[0, +\infty) \times \Sigma$ such that the x -marginal of η and $\tilde{\eta}$ have a finite moment of order p , we define the mixed distance*

$$\mathcal{W}_p(\eta, \tilde{\eta}) := \inf \left\{ \left(\mathbb{E} |x - \tilde{x}|^p \right)^{\frac{1}{p}} + \mathbb{P}(\sigma \neq \tilde{\sigma}) : (x, \sigma) \sim \eta \text{ and } (\tilde{x}, \tilde{\sigma}) \sim \tilde{\eta} \right\}.$$

Adapting [BLBMZ12, Theorem 1.10 and Corollary 1.11] to our setting yields the following theorem.

Theorem 5.25 (Invariant measure and convergence of the instantaneous harmonic process). *(i) For all $1 \leq p < q$ we have*

$$\mathcal{W}_p(\eta P_t, \tilde{\eta} P_t) \leq 2 \left(\left(\int \max(x, v/\mu)^q d\eta(x) \right)^{1/q} + \left(\int \max(x, v/\mu)^q d\tilde{\eta}(x) \right)^{1/q} \right) e^{-\rho(p)t} + 2e^{-2\omega t}$$

where $\rho(p) = \frac{2\mu\omega(1-p/q)}{\mu p + 2\omega(1-p/q)}$.

(ii) The instantaneous harmonic process has a unique invariant probability π and

$$\mathcal{W}_p(\eta P_t, \pi) \leq 2 \left(\left(\int \max(x, v/\mu)^q d\eta(x) \right)^{1/q} + v/\mu \right) e^{-\rho(p)t} + 2e^{-2\omega t}.$$

The theorem is stated for the instantaneous harmonic process but can be extended to other tumbling mechanisms (for example 5.2b) and arbitrary contracting potentials, including higher-dimensional processes such as [SLDMS22]. Our setting allows for a somewhat simplified proof, which we include for the sake of completeness. We again turn to a coupling argument and start by adapting Lemma 5.22 to the harmonic setting.

Lemma 5.26 (Quadratic synchronous coupling). *For all $(x_0, \sigma_0), (\tilde{x}_0, \tilde{\sigma}_0) \in E$ there exists a coupling $\mathbb{Q}_{(x_0, \sigma_0), (\tilde{x}_0, \tilde{\sigma}_0)}$ of $X(t) = (x(t), \sigma(t))$ and $\tilde{X}(t) = (\tilde{x}(t), \tilde{\sigma}(t))$ such that*

(i) X and \tilde{X} are instantaneous harmonic processes with initial states (x_0, σ_0) and $(\tilde{x}_0, \tilde{\sigma}_0)$,

(ii) $\mathcal{P}(\sigma(s) = \tilde{\sigma}(s) \text{ for all } s \geq t) \geq 1 - 2e^{-2\omega t}$,

(iii) if $\sigma(s) = \tilde{\sigma}(s)$ for all $s \geq t$ then

$$|x(s) - \tilde{x}(s)| \leq e^{-\mu(s-t)} |x(t) - \tilde{x}(t)| \text{ for all } s \geq t.$$

Proof. We couple the velocities $\sigma, \tilde{\sigma}$ like in the proof of Lemma 5.22. For $\sigma = 2, 0, -2$ define $x_\sigma^* = \frac{\mu\sigma}{2v}$ and recursively construct $x(t)$ by setting

$$\begin{aligned} x(t) &= \max\left(0, x_{\sigma_0}^* + e^{-\mu t} (x_0 - x_{\sigma_0}^*)\right) \text{ for } t \leq T_1, \\ x(t) &= \max\left(0, x_{\sigma(T_1)}^* + e^{-\mu(t-T_1)} (x(T_1) - x_{\sigma(T_1)}^*)\right) \text{ for } t \in [T_1, T_2], \text{ etc.} \end{aligned}$$

where $0 = T_0 < T_1 < \dots$ are the combined jump times of σ and $\tilde{\sigma}$. We define $\tilde{x}(t)$ similarly and finally set $X(t) = (x(t), \sigma(t))$ and $\tilde{X}(t) = (\tilde{x}(t), \tilde{\sigma}(t))$. Assertions (i) and (ii) follow like in the proof of Lemma 5.22.

(iii) Let $t \geq 0$ be such that $\sigma(s) = \tilde{\sigma}(s)$ for all $s \geq t$ and let $k \in \mathbb{N}$ be such that $T_k \leq t \leq T_{k+1}$. Then we have $\sigma(T_k) = \tilde{\sigma}(T_k)$ and hence

$$\begin{aligned} x(T_{k+1}) &= \max\left(0, x_{\sigma(T_k)}^* + e^{-\mu(T_{k+1}-t)} (x(t) - x_{\sigma(T_k)}^*)\right), \\ \tilde{x}(T_{k+1}) &= \max\left(0, x_{\sigma(T_k)}^* + e^{-\mu(T_{k+1}-t)} (\tilde{x}(t) - x_{\sigma(T_k)}^*)\right). \end{aligned}$$

So because $x \mapsto \max(0, x)$ is 1-Lipschitz we have

$$\begin{aligned} |x(T_{k+1}) - \tilde{x}(T_{k+1})| &\leq \left| \left(x_{\sigma(T_k)}^* + e^{-\mu(T_{k+1}-t)} (x(t) - x_{\sigma(T_k)}^*) \right) \right. \\ &\quad \left. - \left(x_{\sigma(T_k)}^* + e^{-\mu(T_{k+1}-t)} (\tilde{x}(t) - x_{\sigma(T_k)}^*) \right) \right|, \\ &= e^{-\mu(T_{k+1}-t)} |x(t) - \tilde{x}(t)|. \end{aligned}$$

Iterating the same computation shows $|x(T_l) - \tilde{x}(T_l)| \leq e^{-\mu(T_l-t)} |x(t) - \tilde{x}(t)|$ for all $l \geq k+1$. Applying it once more yields $|x(s) - \tilde{x}(s)| \leq e^{-\mu(s-t)} |x(t) - \tilde{x}(t)|$ for all $s \geq t$. \square

We can now return to the proof of the main theorem of the section.

Proof of Theorem 5.25. (i) Using the coupling of Lemma 5.26 we construct the coupling

$$\int \mathbb{Q}_{(x,\sigma),(\tilde{x},\tilde{\sigma})} d(\eta \otimes \tilde{\eta})(x, \sigma, \tilde{x}, \tilde{\sigma})$$

with initial distribution $\eta \otimes \tilde{\eta}$. Fix $\beta \in (0, 1)$ to be optimized later and define the event

$$A_t = \{\sigma(s) = \tilde{\sigma}(s) \text{ for all } s \geq t\}.$$

Notice that $t \mapsto x(t)$ is non-increasing up to time $T = \inf\{t \geq 0 : x(t) \leq v/\mu\}$ and $x(t) \leq v/\mu$ for all $t \geq T$. Hence $0 \leq x(t) \leq \max(x(0), v/\mu)$ for all $t \geq 0$ and we have

$$\begin{aligned} \mathbb{E} [|x(t) - \tilde{x}(t)|^p 1_{A_{\beta t}}] &\leq e^{-(1-\beta)p\mu t} \mathbb{E} [|x(\beta t) - \tilde{x}(\beta t)|^p] \\ &\leq e^{-(1-\beta)p\mu t} \left(\|\max(x(0), v/\mu)\|_p + \|\max(\tilde{x}(0), v/\mu)\|_p \right)^p \\ &\leq e^{-(1-\beta)p\mu t} \left(\|\max(x(0), v/\mu)\|_q + \|\max(\tilde{x}(0), v/\mu)\|_q \right)^p \end{aligned}$$

where we used Lemma 5.26 (iii) for the first inequality. Furthermore, if $s > 1$ is such that $\frac{q}{p} + \frac{1}{s} = 1$ then by Hölder's inequality

$$\begin{aligned} \mathbb{E} [|x(t) - \tilde{x}(t)|^p 1_{A_{\beta t}^c}] &\leq \mathbb{E} [|x(t) - \tilde{x}(t)|^q]^{\frac{p}{q}} \mathbb{P}(A_{\beta t}^c)^{\frac{1}{s}} \\ &\leq \left(\|x(t)\|_q + \|\tilde{x}(t)\|_q \right)^p e^{-2\omega \frac{\beta t}{s}} \\ &\leq \left(\|\max(x(0), v/\mu)\|_q + \|\max(\tilde{x}(0), v/\mu)\|_q \right)^p e^{-2\omega \frac{\beta t}{s}} \end{aligned}$$

Taking $\beta = \frac{p\mu}{2\omega/s+p\mu}$ leads to $(1-\beta)p\mu = 2\omega\beta/s = \frac{2\omega p\mu/s}{2\omega/s+p\mu}$ and thus

$$\begin{aligned} \mathcal{W}_p(\eta P_t, \tilde{\eta} P_t) &\leq \|x(t) - \tilde{x}(t)\|_p + \mathbb{P}(\sigma(t) \neq \tilde{\sigma}(t)) \\ &\leq 2 \left(\|\max(x(0), v/\mu)\|_q + \|\max(\tilde{x}(0), v/\mu)\|_q \right) \exp\left(-\frac{2\omega\mu/s}{2\omega/s+p\mu}t\right) + 2e^{-2\omega t}. \end{aligned}$$

(ii) Because an instantaneous harmonic process started inside the compact set $[0, v/\mu] \times \Sigma$ stays inside that set almost surely, the Krylov–Bogolyubov argument implies that an invariant measure exists. Furthermore, the contractivity shown in assertion (i) implies its uniqueness and the bound for $\mathcal{W}_p(\eta P_t, \pi)$. \square

Remark 5.27. *Exponential convergence towards the invariant measure in total variation distance could be obtained using Foster–Lyapunov techniques [MT93]. However, we do not expect the bounds to be quantitatively useful.*

5.3.3 Instantaneous linear process: lower bound on the total variation distance to the invariant measure

In this section, we prove that for all $(x, \sigma) \in E$

$$\overline{\lim}_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \leq 4\omega \frac{c^2}{v^2}$$

thus showing that the lower bound for the exponential decay rate obtained in section 5.3.1 can only be improved by a constant factor.

In order to do this, we want to exploit the fact that the process takes a ‘long time’ to explore the parts of the state space where the interdistance x is large. The central observation is that, for any fixed $a > 0$, because

$$\|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \geq \pi(\{x > at\}) - \mathbb{P}_{(x,\sigma)}(x(t) > at)$$

we have that

$$\lim_{t \rightarrow +\infty} -\frac{1}{t} \log \mathbb{P}_{(x,\sigma)}(x(t) > at) > a\lambda_\pi \quad (*)$$

implies

$$\overline{\lim}_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \leq a\lambda_\pi.$$

The lower the constant a the sharper the upper bound $a\lambda_\pi$. Hence we want to take $a > 0$ as small as possible while also ensuring that $\mathbb{P}_{(x,\sigma)}(x(t) > at)$ is negligible in the sense of (*).

Lemma 5.28. *Set $\tau_L := \inf\{t \geq 0 : x(t) > L\}$. For all $u < 0$ and $(x, \sigma) \in E$ we have*

$$\lim_{L \rightarrow +\infty} \frac{1}{L} \log \mathbb{E}_{(x,\sigma)} [e^{u\tau_L}] = \frac{cu - 2c\omega - \sqrt{u^2v^2 - 4uv^2\omega + 4c^2\omega^2}}{2(c^2 - v^2)} =: \Lambda(u).$$

We postpone the proof of this lemma until the end of the section.

Lemma 5.29. *For all $a > \frac{1}{\Lambda(0)}$ we have*

$$\lim_{t \rightarrow +\infty} -\frac{1}{t} \log \mathbb{P}_{(x,\sigma)}(x(t) > at) > a\lambda_\pi$$

for all $(x, \sigma) \in E$. Furthermore

$$\overline{\lim}_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(x,\sigma)} P_t - \pi\|_{\text{TV}} \leq 4\omega \frac{c^2}{v^2}.$$

Proof. Let $(x, \sigma) \in E$ be given. For all $u < 0$

$$\mathbb{P}_{(x, \sigma)}(x(t) \geq at) \leq \mathbb{P}_{(x, \sigma)}(\tau_{at} \leq t) \leq e^{-ut} \mathbb{E}_{(x, \sigma)}[e^{u\tau_{at}}]$$

so that by Lemma 5.28 we have

$$\overline{\lim}_{t \rightarrow +\infty} \frac{1}{t} \log \mathbb{P}_{(x, \sigma)}(x(t) \geq at) \leq a\Lambda(u) - u =: G(u).$$

Because $G(0) = -a\lambda_\pi$ if $G'(0) > 0 \iff a > \frac{1}{\Lambda'(0)}$ then

$$a\Lambda(u) - u < -a\lambda_\pi$$

when $u \rightarrow 0-$.

A simple computation yields

$$\frac{\lambda_\pi}{\Lambda'(0)} = 4\omega \frac{c^2}{c^2 + v^2} \leq 4\omega \frac{c^2}{v^2}.$$

□

Proof of Lemma 5.28. Let $\omega, c, v > 0$, $u < 0$ and $(x_0, \sigma_0) \in E$ be given. Fix $L > x_0$ and set $f_\sigma(x) = \mathbb{E}_{(x, \sigma)}[e^{u\tau_L}]$ for $\sigma \in \Sigma$ and $x \in [0, L]$. Further define $F(x) = (f_2(x), f_0(x), f_{-2}(x))^t$. We know that

$$\mathcal{V}F' + \mathcal{Q}F + uF = 0 \iff F' = \mathcal{A}F$$

on $(0, L)$ where $\mathcal{A} = -\mathcal{V}^{-1}(\mathcal{Q} + uI)$. Together with the boundary conditions $f_2(L) = 1$ and

$$\left. \begin{array}{l} \omega f_2(0) + \omega f_{-2}(0) - 2\omega f_0(0) + u f_0(0) = 0 \\ 2\omega f_0(0) - 2\omega f_{-2}(0) + u f_{-2}(0) = 0 \end{array} \right\} \iff \begin{pmatrix} f_2(0) \\ f_0(0) \\ f_{-2}(0) \end{pmatrix} = f_{-2}(0) \underbrace{\begin{pmatrix} 1 + \frac{u^2}{2\omega^2} - \frac{2u}{\omega} \\ 1 - \frac{u}{2\omega} \\ 1 \end{pmatrix}}_{:=F_0}$$

this fully determines F .

We deduce

$$F(x_0) = f_{-2}(0)e^{x_0\mathcal{A}}F_0$$

where $f_{-2}(0)$ depends on L but $e^{x_0\mathcal{A}}F_0$ does not. Hence

$$\lim_L \frac{1}{L} \log \mathbb{E}_{(x_0, \sigma_0)}[e^{u\tau_L}] = \lim_L \frac{1}{L} \log f_{-2}(0).$$

Furthermore the boundary condition $f_2(L) = 1$ can be rewritten as

$$f_{-2}(0) \underbrace{[(1, 0, 0)e^{L\mathcal{A}}F_0]}_{:=\Gamma(L)} = 1.$$

The characteristic polynomial $P_{\mathcal{A}}$ of \mathcal{A} satisfies $-\det(\mathcal{V})P_{\mathcal{A}} = P_1P_2$ where

$$P_1 = 2Xc - u + 2\omega \text{ and } P_2 = 4(c^2 - v^2)X^2 - 4(cu - 2c\omega)X + u^2 - 4u\omega.$$

P_1P_2 is a cubic polynomial with discriminant

$$\Delta = 256(u^2v^2 - 4uv^2\omega + 4c^2\omega^2 + 4v^2\omega^2)^2(u^2v^2 - 4uv^2\omega + 4c^2\omega^2)$$

so whenever $u < 0$ we have $\Delta > 0$ and $P_{\mathcal{A}}$ has three distinct real roots. Furthermore, when $u < 0$ the root of P_1 is negative and P_2 has one negative and one positive root (because its

leading coefficient and constant term have opposite signs).

Let λ_1 be the negative root of P_1 and λ_2^+, λ_2^- be the positive and negative root of P_2 respectively. Because all eigenvalues of \mathcal{A} are distinct, there exist $a_1, a_2^+, a_2^- \in \mathbb{R}$ that do not depend on L such that

$$\Gamma(L) = a_1 e^{\lambda_1 L} + a_2^+ e^{\lambda_2^+ L} + a_2^- e^{\lambda_2^- L}.$$

Assume by contradiction that $a_2^+ < 0$. This would imply $\Gamma(L) \rightarrow -\infty$ and hence $f_{-2}(0) < 0$ for large L , which is absurd. Assume by contradiction that $a_2^+ = 0$ then $|\Gamma(L)| \rightarrow 0$ and hence $|f_{-2}(0)| \rightarrow +\infty$ which is also absurd. Hence $a_2^+ > 0$ and $\lim_t \frac{1}{L} \log f_{-2}(0) = -\lambda_2^+$. A simple computation yields $-\lambda_2^+ = \Lambda(u)$. \square

5.3.4 Finite linear process: lower bound on the total variation distance to the invariant measure

In this section, we show

$$\overline{\lim}_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(\sigma, x)} P_t - \pi\|_{\text{TV}} \leq 4\alpha \left(1 + \frac{\alpha}{\beta}\right) \frac{c^2}{v^2} \text{ for all } (x, \sigma) \in E.$$

As in the previous section, we want to exploit the fact that the process is slow to explore regions of the state space where the interdistance x is large. Recall that for all $a > 0$ such that

$$\underline{\lim}_{t \rightarrow +\infty} -\frac{1}{t} \log \mathbb{P}_{(x, \sigma)}(x(t) \geq at) > a\lambda_\pi \quad (*)$$

we have $\overline{\lim}_t -\frac{1}{t} \log \|\delta_{(x, \sigma)} P_t - \pi\|_{\text{TV}} \leq a\lambda_\pi$. The following lemma is the first step to make this upper bound quantitative.

Lemma 5.30. *We have $\lambda_\pi = -\zeta_2$ (see Notation 5.8).*

Proof. If $v \leq 2c$ then Proposition 5.10 implies $\pi(\{x > t\}) = a_2 e^{\zeta_2 t}$ for some $a_2 \in \mathbb{R}$ and the result follows immediately. If $v > 2c$ then Proposition 5.10 implies $\pi(\{x > t\}) = a_2 e^{\zeta_2 t} + a_3 e^{\zeta_3 t}$ for some $a_2, a_3 \in \mathbb{R}$ so it suffices to show that $\zeta_3 < \zeta_2$.

It follows from the implicit function theorem that ζ_2 and ζ_3 are continuously differentiable functions of (α, β, c, v) (the invertibility condition in the implicit function theorem is guaranteed by the fact that P_2 and P_3 have simple roots as shown in Lemma 5.7). Assume by contradiction that there exist two sets of parameters $\Theta_0 = (\alpha_0, \beta_0, c_0, v_0)$ and $\Theta_1 = (\alpha_1, \beta_1, c_1, v_1)$ such that $\zeta_2(\Theta_0) < \zeta_3(\Theta_0)$ and $\zeta_2(\Theta_1) > \zeta_3(\Theta_1)$. Applying the intermediate value theorem to the function $t \mapsto \zeta_2(\Theta_0 + t(\Theta_1 - \Theta_0)) - \zeta_3(\Theta_0 + t(\Theta_1 - \Theta_0))$ yields that there exists $t^* \in (0, 1)$ such that $\zeta_2(\Theta_0 + t^*(\Theta_1 - \Theta_0)) = \zeta_3(\Theta_0 + t^*(\Theta_1 - \Theta_0))$ which contradicts Lemma 5.7. Hence the ordering of ζ_2 and ζ_3 is always the same and considering any particular choice of parameters shows that we always have $\zeta_3 < \zeta_2$. \square

We now estimate $\lim_t \frac{1}{t} \log \mathbb{P}_{(x, \sigma)}(x(t) > at)$ to find a small $a > 0$ such that $(*)$ holds. We could try to mimic the bounds of the previous section by computing

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \log \mathbb{E}_{(x, \sigma)} [e^{u\tau_t}]$$

but the size of the systems of differential equations involved makes this challenging and it is not clear how to adapt Lemma 5.28 to the finite linear case. To overcome this difficulty, we will exploit the fact that the velocity σ is the difference of two simpler Markov jump processes to reduce the size of the systems of differential equations in our computations.

Definition 5.31 (Single speed coupling). *Let $(x_0, \sigma_0) \in [0, +\infty) \times \Sigma$ be given. And let $f : \{-1, 0, 1\}^2 \rightarrow \Sigma$ be defined by*

$$f(\sigma_1, \sigma_2) = \begin{cases} \sigma_2 - \sigma_1 & \text{if } \sigma_1 \neq \sigma_2, \\ 0_{\pm} & \text{if } \sigma_1 = \sigma_2 = \pm 1, \\ 0_0 & \text{if } \sigma_1 = \sigma_2 = 0. \end{cases}$$

Let σ_1 and σ_2 be independent Markov jump processes following figure 5.1b with initial states $\sigma_{1,0}$ and $\sigma_{2,0}$ such that $f(\sigma_{1,0}, \sigma_{2,0}) = \sigma_0$. This makes $\sigma(t) = f(\sigma_1(t), \sigma_2(t))$ a Markov jump process following figure 5.2b with initial state σ_0 .

Set $x(0) = x_0$ and $y_1(0) = y_2(0) = \frac{x_0}{2}$ and recursively define

$$\begin{aligned} x(t) &= \max[0, x(T_n) + (-2c + v\sigma(T_n))(t - T_n)], \\ y_1(t) &= \max[0, y_1(T_n) + (-c + v\sigma_1(T_n))(t - T_n)], \\ y_2(t) &= \max[0, y_2(T_n) + (-c - v\sigma_2(T_n))(t - T_n)], \end{aligned}$$

for $t \in [T_n, T_{n+1}]$ where $0 = T_0 < T_1 < \dots$ are the combined jump times of σ_1, σ_2 and σ .

The process $X(t) = (x(t), \sigma(t))$ constructed in Definition 5.31 is clearly a finite linear process and the two auxiliary processes $y_1(t)$ and $y_2(t)$ provide a useful upper bound as shown by the following lemma.

Lemma 5.32 (Order preservation). *If $x(t), y_1(t)$ and $y_2(t)$ are as in Definition 5.31 then $x(t) \leq y_1(t) + y_2(t)$ for all $t \geq 0$ almost surely.*

Proof. We can decompose $[0, +\infty)$ into a sequence of intervals $[t_0, t_1], [t_1, t_2], \dots$ such that both x and y_i have constant derivative on $[t_k, t_{k+1}]$. By induction, it suffices to prove that $x(t_k) \leq y_1(t_k) + y_2(t_k)$ implies $x(t) \leq y_1(t) + y_2(t)$ for all $t \in [t_k, t_{k+1}]$.

Recall that, for $t \in [t_k, t_{k+1}]$

- if $x(t_k) > 0$ then $\partial_t x(t) = -2c + v(\sigma_2(t_k) - \sigma_1(t_k))$,
- if $x(t_k) = 0$ and $-2c + v(\sigma_2(t_k) - \sigma_1(t_k)) \geq 0$ then $\partial_t x(t) = -2c + v(\sigma_2(t_k) - \sigma_1(t_k))$,
- if $x(t_k) = 0$ and $-2c + v(\sigma_2(t_k) - \sigma_1(t_k)) < 0$ then $\partial_t x(t) = 0$,

and similarly

- if $y_i(t_k) > 0$ then $\partial_t y_i(t) = -c + v(-1)^i \sigma_i(t_k)$,
- if $y_i(t_k) = 0$ and $-c + v(-1)^i \sigma_i(t_k) \geq 0$ then $\partial_t y_i(t) = -c + v(-1)^i \sigma_i(t_k)$,
- if $y_i(t_k) = 0$ and $-c + v(-1)^i \sigma_i(t_k) < 0$ then $\partial_t y_i(t) = 0$.

Case 1. If $\partial_t x = -2c + \sigma_2(t_k) - \sigma_1(t_k)$ then, by the preceding remark, we necessarily have $\partial_t x \leq \partial_t y_1 + \partial_t y_2$ so $x(t_k) \leq y_1(t_k) + y_2(t_k)$ implies $x(t) \leq y_1(t) + y_2(t)$ for all $t \in [t_k, t_{k+1}]$.

Case 2. If $\partial_t x > -2c + \sigma_2 - \sigma_1$ then $x(t) = 0$ for all $t \in [t_k, t_{k+1}]$. We always have $y_i \geq 0$ so this implies $x(t) \leq y_1(t) + y_2(t)$ for all $t \in [t_k, t_{k+1}]$. \square

The following lemma will play the same role as Lemma 5.28.

Lemma 5.33. *Set $\tau_L^{(i)} := \inf\{s \geq 0 : y_i(s) \geq L\}$. There exists $U = U(\alpha, \beta, c, v) < 0$ such that for all $u \in (U, 0)$ and all $(y_i, \sigma_i) \in \mathbb{R}_+ \times \{1, 0, -1\}$ we have*

$$\lim_{L \rightarrow \infty} \frac{1}{L} \log \mathbb{E}_{(y_i, \sigma_i)} \left[e^{u\tau_L^{(i)}} \right] = -\max(\text{spec } \mathcal{A}_1(u))$$

where $\mathcal{A}_1(u) = -\mathcal{V}_1^{-1}(\mathcal{Q}_1 + uI)$ and

$$\mathcal{V}_1 = \begin{pmatrix} -c+v & 0 & 0 \\ 0 & -c & 0 \\ 0 & 0 & -c-v \end{pmatrix}, \quad \mathcal{Q}_1 = \begin{pmatrix} -\alpha & \alpha & 0 \\ \beta/2 & -\beta & \beta/2 \\ 0 & \alpha & -\alpha \end{pmatrix}.$$

Proof. If we can show that for $u \approx 0$ and $u < 0$ we have that \mathcal{A}_1 has three distinct eigenvalues, one of which is positive and two of which are negative then the result follows from the same arguments as in Lemma 5.28. The eigenvalues of $\mathcal{A}_1(0)$ are the roots of

$$\det(\mathcal{V}) \det(XI - \mathcal{A}_1(0)) = X(c(v^2 - c^2)X^2 - (2\alpha c^2 + \beta c^2 - \beta v^2)X - \alpha^2 c - \alpha \beta c)$$

so $\mathcal{A}_1(0)$ has three distinct real eigenvalues $X_+ > 0, X_0 = 0, X_- < 0$. For $u \approx 0$, these can be extended in a continuously differentiable manner to three distinct eigenvalues $X_+(u), X_0(u), X_-(u)$ of $\mathcal{A}_1(u)$ using the implicit function theorem (the invertibility condition is guaranteed by the fact that the eigenvalues are distinct). Clearly $X_+(u) > 0$ and $X_-(u) < 0$ for u small enough and the implicit function theorem also yields that the derivative of $X_0(u)$ at $u = 0$ is $\frac{1}{c}$. So we also have $X_0(u) < 0$ for $u \approx 0$ and $u < 0$. \square

We will also use the following lemma, the proof of which we include for the sake of completeness.

Lemma 5.34. *Let Y_1 and Y_2 be independent a.s. positive random variables. For all $A \geq 0$ and all $n \geq 1$ we have*

$$\mathbb{P}(Y_1 + Y_2 \geq A) \leq \sum_{k=0}^n \mathbb{P}\left(Y_1 \geq \frac{k}{n+1}A\right) \mathbb{P}\left(Y_2 \geq \frac{n-k}{n+1}A\right).$$

Proof. If we can show

$$\{y_1 + y_2 \geq A\} \cap \{y_1, y_2 \geq 0\} \subset \bigcup_{k=0}^n \left\{y_1 \geq \frac{k}{n+1}A\right\} \cap \left\{y_2 \geq \frac{n-k}{n+1}A\right\}$$

then we can conclude using the union bound and the independence of Y_1 and Y_2 .

We have

$$y_1 \geq \frac{k}{n+1}A \iff \frac{(n+1)y_1}{A} \geq k$$

so setting $k^* := \left\lfloor \frac{(n+1)y_1}{A} \right\rfloor$ guarantees $y_1 \geq \frac{k^*}{n+1}A$. Furthermore, $y_1 + y_2 \geq A$ implies

$$\frac{n - k^*}{n+1}A = \frac{n - \left\lfloor \frac{(n+1)y_1}{A} \right\rfloor}{n+1}A \leq \frac{n - \frac{(n+1)y_1}{A} + 1}{n+1}A = A - y_1 \leq y_2.$$

Finally $k^* \geq 0$ and $k^* \leq n$ except if $y_1 = A$ and $y_2 = 0$ but in this case we can take $k^* = n$. \square

We can now adapt the computations of the previous section and obtain an upper bound for the exponential decay rate.

Lemma 5.35. (i) *For all $n \geq 1$, $u \in (U, 0)$ and $(x, \sigma) \in E$ we have*

$$\overline{\lim}_{t \rightarrow +\infty} \frac{1}{t} \log \mathbb{P}_{(x, \sigma)}(x(t) > at) \leq \frac{n}{n+1} a \Lambda_1(u) - 2u$$

where $\Lambda_1(u)$ equals minus the largest eigenvalue of $\mathcal{A}_1(u)$.

(ii) For all $a > \frac{2}{\Lambda_1'(0)}$ we have

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \log \mathbb{P}_{(x,\sigma)}(x(t) > at) > a\lambda_\pi.$$

(iii) We have $\overline{\lim}_{t \rightarrow +\infty} -\frac{1}{t} \log \|\delta_{(\sigma,x)} P_t - \pi\|_{\text{TV}} \leq 4\alpha \left(1 + \frac{\alpha}{\beta}\right) \frac{c^2}{v^2}$.

Proof. (i) Lemma 5.34 implies

$$\begin{aligned} \mathbb{P}(x(t) \geq at) &\leq \mathbb{P}(y_1(t) + y_2(t) \geq at) \\ &= \sum_{k=0}^n \mathbb{P}\left(y_1(t) \geq \frac{k}{n+1}at\right) \mathbb{P}\left(y_2(t) \geq \frac{n-k}{n+1}at\right) \\ &\leq \sum_{k=0}^n \mathbb{P}\left(\tau_{\frac{k}{n+1}at}^{(1)} < t\right) \mathbb{P}\left(\tau_{\frac{n-k}{n+1}at}^{(2)} < t\right) \\ &\leq \sum_{k=0}^n e^{-2ut} \mathbb{E}\left[e^{u\tau_{\frac{k}{n+1}at}^{(1)}}\right] \mathbb{E}\left[e^{u\tau_{\frac{n-k}{n+1}at}^{(2)}}\right]. \end{aligned}$$

(ii) Follows from the same arguments as Lemma 5.29.

(iii) Assertion (ii) and Lemma 5.30 imply that $\frac{2}{\Lambda_1'(0)}(-\zeta_2)$ is an upper bound for the exponential decay rate. Using the implicit value theorem to compute $\Lambda_1'(0)$ yields

$$\begin{aligned} \frac{2}{\Lambda_1'(0)}(-\zeta_2) &= 2\alpha \left((4r^2 + 2r - 1)R^2 + R^4 \left(\sqrt{(4r^2 - 2)R^2 + R^4 + 1} + 4r(2r^2 + r - 1) - 1 \right) \right. \\ &\quad \left. - \sqrt{(4r^2 - 2)R^2 + R^4 + 1} + (2r + 1)R^6 + 1 \right) \\ &\quad \times (2r((4r^2 - 1)R^4 + (4r - 1)R^2 + R^6 + 1))^{-1} \end{aligned}$$

where $r = \frac{\alpha}{\beta}$ and $R = \frac{c}{v}$. Optimizing over $r > 0$ and $0 < R < 1$ with the help of a computer algebra system we get

$$\sup \frac{\frac{2}{\Lambda_1'(0)}(-\zeta_2)}{\alpha(1+r)R^2} = 4$$

which implies the desired result. \square

Chapter 6

Three interacting run-and-tumble particles

6.1 Introduction

In out-of-equilibrium statistical mechanics, effective attraction can arise between particles that exhibit persistent movement and repulsive interactions, resulting in cluster formation [CT15]. Essentially, particles in denser regions decelerate, which further increases the density and sets up a feedback loop leading to aggregation. Simulations have confirmed this phenomenon [DDK⁺20b], but exact results in the form of invariant measures are only known for systems with up to two interacting particles [SEB16, DDK20a, LDMS21]. Note that for two particles, the problem reduces to that of a single particle by considering the relative particle, albeit with more complex velocity dynamics. However, this method fails for systems with three or more particles, which would be useful for understanding clustering.

A major challenge lies in properly understanding the boundary conditions for the invariant measure, even in the two-particle case. One approach to address this is discretizing the state space and considering the limit as the discretization step approaches zero [SEB16, MEB22]. Alternatively, one can examine the limit of vanishing thermal noise [MJK⁺18] or study increasingly peaked soft repulsive interaction potentials [AdPvW23]. However, these methods complicate the original process and introduce an additional step where a limit must be computed. A Laplace transform approach [Ang17] can mitigate some of these issues but is difficult to generalize to systems with more particles.

In this chapter, we propose a model for three one-dimensional run-and-tumble particles (RTPs) with jamming interactions using the piecewise deterministic Markov process (PDMP) formalism. We then determine its invariant measure by utilizing the explicit knowledge of the extended generator [Dav93] and the generator characterization of invariance. This approach is enabled by leveraging insights from two-particle systems with arbitrary tumbling mechanisms [HGM23] to formulate an ansatz for the invariant measure. This is a step toward the general many-particle case, which remains an open problem.

6.2 Model description

When generalizing two RTPs on the 1D torus with jamming to three particles, the choice of the state space is a natural first question. Translational symmetry and cyclic order preservation allow modeling the system at the level of the distance between two consecutive particles and particle velocities. We denote x_i the distance between particle i and particle $i + 1$ adopting the convention that particle 4 is particle 1. Hence the position vector $x = (x_1, x_2, x_3)$ is an element of the set

$$\Delta = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1, x_2, x_3 \geq 0 \text{ and } x_1 + x_2 + x_3 = L\}.$$

Each particle is endowed with a velocity $\sigma_i \in \{\pm 1\}$ so we obtain the state space $E = \Delta \times \Sigma$, the elements of which we denote $(x, \sigma) = (x_1, x_2, x_3, \sigma_1, \sigma_2, \sigma_3)$, and where $\Sigma = \{\pm 1\}^3$. We denote the vertices of the simplex

$$A = (L, 0, 0), \quad B = (0, L, 0), \quad C = (0, 0, L).$$

The three particle process $X(t) = (x(t), \sigma(t))$ modeling the random evolution of the system is a PDMP meaning that it combines random jumps and deterministic motion. Each σ_i is an independent Markov jump process switching between $+1$ and -1 with rate $\omega > 0$. When $\sigma_1 = \sigma_2 = \sigma_3 = \pm 1$ all particles move around the 1D torus in exactly the same way so the distances x_i remain the same.

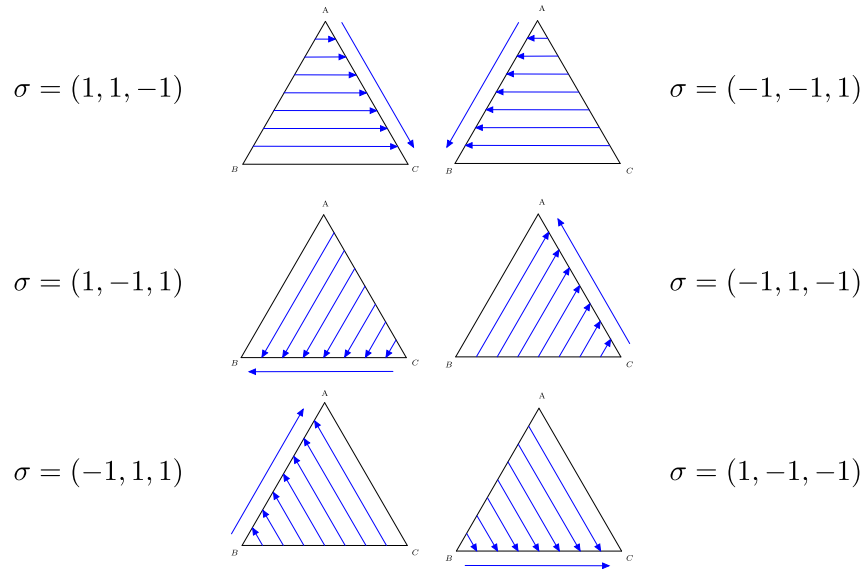
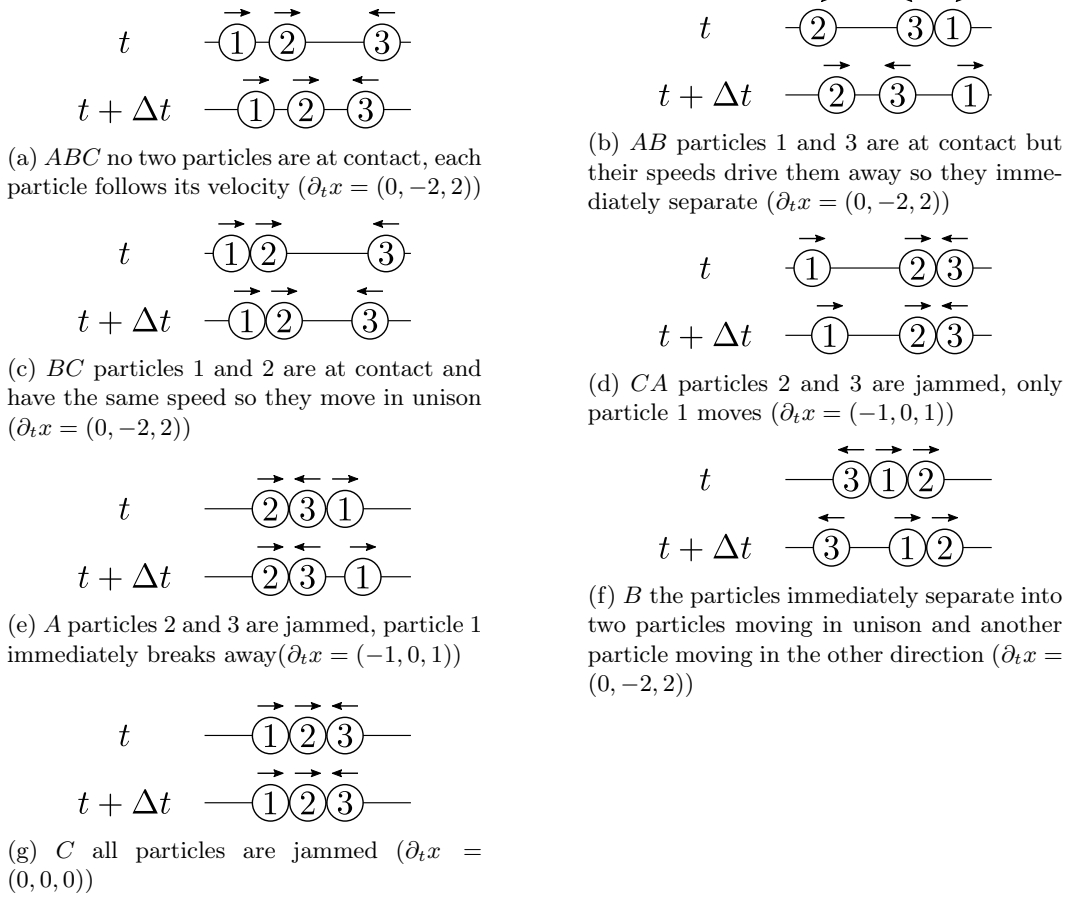


Figure 6.1: Deterministic dynamics on $\Delta \times \{\sigma\}$ for $\sigma \neq \pm(1, 1, 1)$

The deterministic dynamics when $\sigma \neq \pm(1, 1, 1)$ are summarized in Figure 6.1. To deduce them for $\sigma = (1, 1, -1)$, decompose Δ into its vertices A, B, C , open edges $(AB), (BC), (CA)$ and interior $ABC = \Delta$. The dynamics of x can then be obtained separately on each part of Δ separately, as in Figure 6.2.

Figure 6.2: Deterministic dynamics for $\sigma = (1, 1, -1)$

In summary, we get the following deterministic flow $(\phi_t^0)_{t \geq 0}$ when $\sigma = (1, 1, -1)$

- starting at *C* i.e. $x = (0, 0, L)$ then $\phi_t^0(x) = C$ for all $t \geq 0$,
- starting on $[A, C)$ i.e. $x = (x_1, 0, L - x_1)$ for some $x_1 > 0$ then

$$\phi_t^0(x) = \begin{cases} (x_1 - t, 0, L - x_1 + t) & \text{for } t \leq x_1, \\ (0, 0, L) & \text{for } t \geq x_1, \end{cases}$$

- starting in $\Delta - [A, C]$ i.e. $x = (x_1, x_2, x_3)$ for some $x_2 > 0$

$$\phi_t^0(x) = \begin{cases} (x_1, x_2 - 2t, x_3 + 2t) & \text{for } t \leq x_2/2, \\ (x_1 - (t - x_2/2), 0, x_3 + x_2 + (t - x_2/2)) & \text{for } x_2/2 \leq t \leq x_2/2 + x_1, \\ (0, 0, L) & \text{for } t \geq x_2/2 + x_1. \end{cases}$$

Carrying out the same analysis on each of the $\Delta \times \{\sigma\}$ for $\sigma \neq \pm(1, 1, 1)$ leads to the same deterministic dynamics, up to permutation of the coordinates (see Figure 6.1).

More precisely, notice that for all $(\sigma_1, \sigma_2, \sigma_3) \neq \pm(1, 1, 1)$ the (unjammed) relative velocity vector $(\sigma_2 - \sigma_1, \sigma_3 - \sigma_2, \sigma_1 - \sigma_3)$ is equal to $(0, -2, 2)$ up to coordinate permutation and if $\tau_\sigma : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is the unique coordinate permutation such that

$$\tau_\sigma(\sigma_2 - \sigma_1, \sigma_3 - \sigma_2, \sigma_1 - \sigma_3) = (0, -2, 2)$$

then the deterministic dynamics on $\Delta \times \{\sigma\}$ is given by

$$\phi_t(x, \sigma) = ([\tau_\sigma^{-1} \circ \phi_t^0 \circ \tau_\sigma](x), \sigma).$$

The flow $(\phi_t)_{t \geq 0}$ naturally extends to $\Delta \times \Sigma$ by

$$\phi_t(x, \pm(1, 1, 1)) = (x, \pm(1, 1, 1)) \text{ for all } t \geq 0.$$

Davis [Dav93] provides the extended generator of the process together with its domain.

Proposition 6.1. *The three particle process is a Feller, strong Markov process with extended generator*

$$\mathcal{L}f(x, \sigma) = \omega \sum_{i=1}^3 [f(x, (-1)^{\delta_{i1}} \sigma_1, (-1)^{\delta_{i2}} \sigma_2, (-1)^{\delta_{i3}} \sigma_3) - f(x, \sigma)]$$

where δ_{ij} is the Kronecker delta if $\sigma = \pm(1, 1, 1)$ and

$$\mathcal{L}f(x, \sigma) = \tau_\sigma^{-1}(v(\tau_\sigma(x))) \cdot \nabla_x f(x, \sigma) + \omega \sum_{i=1}^3 [f(x, (-1)^{\delta_{i1}} \sigma_1, (-1)^{\delta_{i2}} \sigma_2, (-1)^{\delta_{i3}} \sigma_3) - f(x, \sigma)]$$

where

$$v(x) = \begin{cases} (0, -2, 2) & \text{for } x \in \Delta - [A, C], \\ (-1, 0, 1) & \text{for } x \in [A, C], \\ (0, 0, 0) & \text{for } x = C, \end{cases}$$

if $\sigma \neq \pm(1, 1, 1)$. Its domain is

$$D(\mathcal{L}) = \{t \mapsto f(\phi_t(x, \sigma)) \text{ is absolutely continuous for all } (x, \sigma) \in E\}.$$

6.3 State space reduction

In this section, the $(\Delta \times \Sigma)$ -valued three particle process $X(t) = (x(t), \sigma(t))$ is transformed into a process with simpler state space. This is done in three steps such that the original invariant measure can be deduced from the more easily computable invariant measure of the new process.

Trajectorial cutting and pasting The first step is to remove the velocity configurations $\sigma = \pm(1, 1, 1)$. This means erasing the parts of the trajectory of the original process where $\sigma(t) = \pm(1, 1, 1)$ and glue the remaining parts together. To this end, define

$$T(t) = \inf \left\{ s \geq 0 : \int_0^s 1_{\{\sigma_\theta \neq \pm(1, 1, 1)\}} d\theta > t \right\}$$

and consider the new process $Y_1(t) := X(T(t))$. Starting from $(x_0, \sigma_0) \in \Delta \times \Sigma$ such that $\sigma_0 \neq \pm(1, 1, 1)$ this is equivalent to the following trajectorial construction

1. follow $X(t)$ until the first time τ such that $\sigma(\tau) = \pm(1, 1, 1)$,
2. jump to the position $X(\tilde{\tau})$ where $\tilde{\tau}$ is the first time after τ such that $\sigma(\tilde{\tau}) \neq \pm(1, 1, 1)$,
3. repeat.

The process Y_1 is a PDMP with state space $E_1 = \Delta \times (\Sigma - \{\pm(1, 1, 1)\})$ and the same deterministic dynamics as X , only the random jumps change. Up to renormalization, the invariant measure of Y_1 is equal to the restriction to E_1 of the invariant measure of X , as can be seen from the mean occupancy time characterization of invariant measures.

It remains to understand how the jumps are affected by the removal of $\sigma = \pm(1, 1, 1)$. This transformation is illustrated in Figure 6.3 using the example of a jump starting from $\sigma = (1, 1, -1)$.

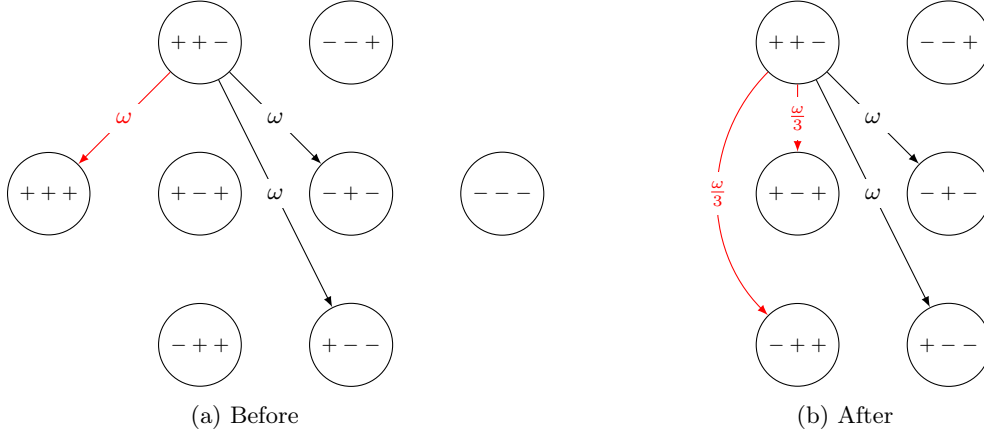


Figure 6.3: Transitions of $\sigma(t)$ from $(1, 1, -1)$ before and after removing $\sigma = \pm(1, 1, 1)$

When X jumps from a state in E_1 to a state in E_1 , the process Y_1 does the same. This translates into the fact that the black transitions in Figure 6.3 are identical before and after the removal. When X jumps from a state in E_1 to a state outside E_1 then Y_1 jumps immediately to the position that X will occupy once it returns to E_1 . This means that the single red transition in Figure 6.3a is replaced by the two red transitions in figure 6.3b. Notice that the sum of the rates of the two red arrows in Figure 6.3b is strictly less than the rate of the original red arrow. This comes from the fact that when X jumps outside E_1 , it has a probability of $\frac{1}{3}$ of coming back to the exact same position when it reenters E_1 , leading to Y_1 not jumping at all.

Transforming velocity flips into coordinate permutations Further reduction of the state space hinges on the observation that, up to coordinate permutation, the six copies of Δ that make up E_1 are identical in terms of deterministic dynamics (see Figure 6.1) and stochastic jumps. It is therefore possible to forget about σ and transform all velocity flips into permutations of the x_i .

To do this, we define $p : E_1 \rightarrow \Delta$ by

$$p(x_1, x_2, x_3, \sigma_1, \sigma_2, \sigma_3) = \tau(x_1, x_2, x_3)$$

where τ is the unique coordinate permutation such that

$$\tau(\sigma_2 - \sigma_1, \sigma_3 - \sigma_2, \sigma_1 - \sigma_3) = (0, -2, 2)$$

and consider the Δ -valued process $Y_2(t) = p(Y_1(t))$. Note that

$$p(x_1, x_2, x_3, 1, 1, -1) = (x_1, x_2, x_3),$$

enables the identification of $\Delta = E_2$ with $\Delta \times \{(1, 1, -1)\} \subset E_1$. Seen through this identification, Y_2 is again a PDMP with the same deterministic dynamics as Y_1 , only the stochastic jumps change. Up to rescaling, the invariant measure of Y_2 is the restriction to $\Delta \times \{(1, 1, -1)\} \approx E_2$ of the invariant measure of Y_1 .

The construction of p implies that the velocity jumps of Y_1 are transformed into coordinate permutation for Y_2 . For example, the jump with rate $\lambda_1 = \omega$ from $\sigma = (1, 1, -1)$ to $\tilde{\sigma} = (-1, 1, -1)$ becomes the permutation

$$\tau_1(x_1, x_2, x_3) = (x_3, x_2, x_1)$$

because τ_1 is the unique permutation such that

$$\tau_1(\sigma_2 - \sigma_1, \sigma_3 - \sigma_2, \sigma_1 - \sigma_3) = (0, -2, 2).$$

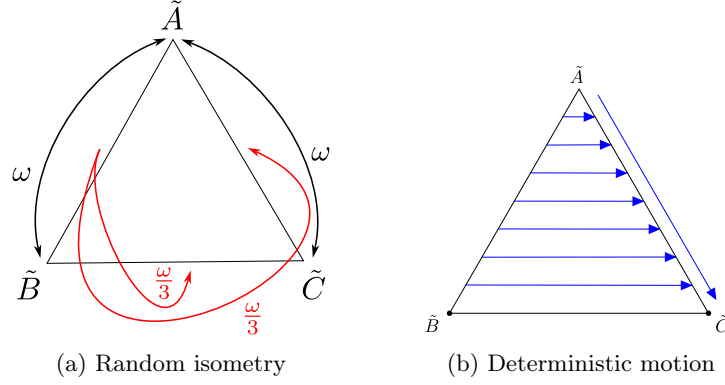


Figure 6.4: Illustration of the reduced process

Similarly the jump with rate

- $\lambda_2 = \omega$ from $(1, 1, -1)$ to $(1, -1, -1)$ becomes the permutation

$$\tau_2(x_1, x_2, x_3) = (x_2, x_1, x_3),$$

- $\lambda_3 = \frac{\omega}{3}$ from $(1, 1, -1)$ to $(1, -1, 1)$ becomes the permutation

$$\tau_3(x_1, x_2, x_3) = (x_2, x_3, x_1),$$

- $\lambda_4 = \frac{\omega}{3}$ from $(1, 1, -1)$ to $(-1, 1, 1)$ becomes the permutation

$$\tau_4(x_1, x_2, x_3) = (x_3, x_1, x_2).$$

Planar projection In the last step, the simplex $\Delta \subset \mathbb{R}^3$ is isometrically mapped to $\tilde{\Delta} = \tilde{A}\tilde{B}\tilde{C} \subset \mathbb{R}^2$ where

$$\tilde{A} = (\sqrt{1/2}L, \sqrt{3/2}L), \quad \tilde{B} = (0, 0), \quad \tilde{C} = (0, \sqrt{2}L).$$

This transforms $v(x)$ into

$$\tilde{v}(x) = \begin{cases} (2\sqrt{2}, 0) & \text{for } x \in \tilde{A}\tilde{B}\tilde{C} - [\tilde{A}, \tilde{C}], \\ (\sqrt{2}/2, -\sqrt{6}/2) & \text{for } x \in [\tilde{A}, \tilde{C}], \\ (0, 0) & \text{for } x = \tilde{C}, \end{cases}$$

and

- τ_1 into the linear isometry $\tilde{\tau}_1 : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ exchanging A and B and leaving C fixed,
- τ_2 into the linear isometry $\tilde{\tau}_2 : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ exchanging A and B and leaving C fixed,
- τ_3 into the linear isometry $\tilde{\tau}_3$ satisfying $\tilde{\tau}_3(A) = B$, $\tilde{\tau}_3(B) = C$ and $\tilde{\tau}_3(C) = A$,
- τ_4 into the linear isometry $\tilde{\tau}_4$ satisfying $\tilde{\tau}_4(A) = C$, $\tilde{\tau}_4(B) = A$ and $\tilde{\tau}_4(C) = B$.

The final process, illustrated by Figure 6.4, has state space $\tilde{E} = \tilde{A}\tilde{B}\tilde{C}$ and extended generator

$$\tilde{\mathcal{L}}f(x) = \tilde{v}(x) \cdot \nabla_x f(x) + \sum_{i=1}^4 \lambda_i (f(\tilde{\tau}_i(x)) - f(x))$$

with domain

$$D(\tilde{\mathcal{L}}) = \{t \mapsto f(\tilde{\phi}_t(x)) \text{ is absolutely continuous for all } x \in \tilde{E}\}$$

where $(\tilde{\phi}_t)_{t \geq 0}$ is the flow induced on \tilde{E} by $(\phi_t)_{t \geq 0}$.

6.4 Invariant measure

We now turn to the computation of the invariant measure. The approach is to first formulate an ansatz for the invariant measure based on our physical understanding of the process and then verify that it is correct using the generator characterization

$$\pi \text{ is invariant} \iff \int \mathcal{L}f d\pi = 0 \text{ for all } f \in D(\mathcal{L}).$$

Ansatz The invariant measure must take the form of a weighted Dirac on the vertices A, B, C . Because the deterministic dynamics immediately leaves A and B , only the Dirac at C has nonzero weight. The invariant measure assigns zero mass to the edge (AB) for the same reason. Simulations show that the invariant measure is uniform in the interior of the state space Δ so that it suffices to obtain an ansatz for the invariant measure on (AC) and (BC) . Notice that states on AC and BC correspond to configurations where two particles stay at zero distance because they are jammed or because they move together respectively.

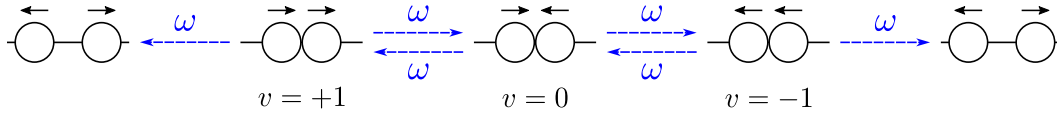


Figure 6.5: Velocity transition rates of effective particle (in the leftmost and rightmost configurations the particles immediately separate and there is no effective velocity)

Hence, zooming in on the velocity jumps and deterministic dynamics when the process is in (AC) and (BC) (see Figure 6.5) reveals that the two particles at contact act like a single effective particle. Therefore, the model behaves like two RTPs with different velocity dynamics. Our study of two RTPs with arbitrary velocity dynamics in Section 1.6.1 shows that the invariant measure of such a system is given by a matrix exponential, i.e. uniform and exponential terms (possibly with polynomial prefactors). The simplest version of this is the ansatz

$$\pi = \sum_{F \in \mathcal{F}} a_{\sigma, F} \lambda_F + \sum_{X \neq Y} b_{\sigma, XY} \epsilon_{XY} \quad (6.1)$$

where $\mathcal{F} = \{A, B, C, AB, BC, CA, ABC\}$ is the set of non-empty faces of the polytope Δ , λ_F is the $\dim(F)$ -dimensional Lebesgue measure on the face $F \in \mathcal{F}$ and for $X, Y \in \{A, B, C\}$

$$\epsilon_{XY}(f) = \int_0^{\|Y-X\|} f\left(X + u \frac{Y-X}{\|Y-X\|}\right) e^{\kappa u} du$$

is the exponential measure. Note that $\lambda_{XY} = \lambda_{YX}$ but $\epsilon_{XY} \neq \epsilon_{YX}$.

Generator characterization of invariance In this section, we work with the reduced planar process of the previous section and drop all tilde in our notation.

Injecting the ansatz and the explicit extended generator into

$$\int \mathcal{L}f d\pi = 0 \text{ for all } f \in D(\mathcal{L})$$

is simple but involves a lot of computations, which we do not include here for conciseness. We just expand the term $\int \mathcal{L}f d\epsilon_{AC}$ as an example. Defining $g : [0, \sqrt{2}L] \rightarrow \mathbb{R}$ by

$$g(u) = f\left(A + u \frac{C-A}{\|C-A\|}\right)$$

we get

$$\begin{aligned}
& \int \mathcal{L} f d\epsilon_{AC} = \\
& \int_0^{\sqrt{2}L} \left\{ \sqrt{2}g'(u) + \sum_{i=0}^4 \lambda_i \left[f \left(\tau_i \left(A + u \frac{C-A}{\|C-A\|} \right) \right) - f \left(A + u \frac{C-A}{\|C-A\|} \right) \right] \right\} e^{\kappa u} du = \\
& [\sqrt{2}g(u)e^{\kappa u}]_0^{\sqrt{2}L} - \int_0^{\sqrt{2}L} \sqrt{2}g(u)(\kappa e^{\kappa u}) du - \frac{8\omega}{3} \int_0^{\sqrt{2}L} f \left(A + u \frac{C-A}{\|C-A\|} \right) e^{\kappa u} du \\
& + \sum_{i=0}^4 \lambda_i \int_0^{\sqrt{2}L} f \left(\tau_i(A) + u \frac{\tau_i(C) - \tau_i(A)}{\|\tau_i(C) - \tau_i(A)\|} \right) e^{\kappa u} du \\
& = \sqrt{2}f(C)e^{\sqrt{2}\kappa L} - \sqrt{2}f(A) - \sqrt{2}\kappa \int f d\epsilon_{AC} - \frac{8\omega}{3} \int f d\epsilon_{AC} \\
& + \omega \int f d\epsilon_{BC} + \omega \int f d\epsilon_{CA} + \frac{\omega}{3} \int f d\epsilon_{BA} + \frac{\omega}{3} \int f d\epsilon_{CB}.
\end{aligned}$$

Expanding everything and grouping the terms $\int f d\nu$ for $\nu = \lambda_F$ and $\nu = \epsilon_{XY}$ leads to a system of linear equations, the non-trivial solution of which is unique up to normalization and given by

$$\begin{aligned}
\kappa &= 2\omega, \quad d_C = (5\sqrt{2} + 7)e^{2\sqrt{2}\omega L} + \sqrt{2} - 1, \\
a_{AC} &= \frac{\omega}{3} \left[(8\sqrt{2} + 11)e^{2\sqrt{2}\omega L} - 2\sqrt{2} + 1 \right], \quad a_{BC} = \frac{1}{2}a_{AC}, \quad a_{ABC} = \frac{\omega}{\sqrt{6}}a_{AC}, \\
b_{AC} &= \omega, \quad b_{CA} = (2\sqrt{2} + 3)\omega, \quad b_{BC} = \frac{\sqrt{2}}{4}\omega, \quad b_{CB} = -\left(\frac{3\sqrt{2}}{4} + 1 \right)\omega.
\end{aligned}$$

Invariant measure on $\Delta \times \Sigma$ Going back to the original process, we get the invariant measure

$$\pi = \left(\frac{1}{3} \sum_{\sigma \in \Sigma_1} \tau_\sigma \# \pi_0 \right) \otimes \delta_{(1,1,1)} + \left(\frac{1}{3} \sum_{\sigma \in \Sigma_{-1}} \tau_\sigma \# \pi_0 \right) \otimes \delta_{(-1,-1,-1)} + \sum_{\sigma \neq \pm(1,1,1)} (\tau_\sigma \# \pi_0) \otimes \delta_\sigma$$

where π_0 is the isometric pushforward from $\tilde{\Delta} \subset \mathbb{R}^2$ to $\Delta \subset \mathbb{R}^3$ of the invariant measure above and

$$\Sigma_{\pm 1} = \{(\sigma_1, \sigma_2, \sigma_3) \in \Sigma : \sigma_1 + \sigma_2 + \sigma_3 = \pm 1\}.$$

Note that the invariant measure combines Lebesgue measures of different dimensions and assigns positive mass some edges and vertices. Furthermore, it cannot be written a product over pairs of particles.

Chapter 7

Convergence to equilibrium of some irreversible Markov processes via lifting, Flow Poincaré inequality and a general divergence lemma

This chapter is the result of joint work with Andreas Eberle, Arnaud Guillin, Francis Lörler and Manon Michel.

7.1 Introduction

Convergence to equilibrium of Markov processes, being in discrete time or in continuous time, is a longstanding problematic in probability, being for processes modeling physical or biological systems for example or for sampling (MCMC methods). Various techniques have emerged, and it would be an impossible task to be exhaustive on the literature, such as Lyapunov conditions [MT93], coupling techniques [Ebe16], functional inequalities [BGL14]. If the Lyapunov conditions technique is more qualitative than quantitative, functional inequalities have provided sharp rates of convergence in the reversible case via Poincaré inequalities (or spectral gap) or logarithmic Sobolev inequalities, however requiring to find the optimal constant in those inequalities. In the non reversible case, much less is known, despite an accelerated rate of convergence to stationarity precisely due to this irreversibility, as demonstrated for example for Ornstein-Uhlenbeck type processes [HHMS93, LNP13, GM16], or more generally in [HHMS05, CH13] or in [EGZ19] for underdamped Langevin processes. For Markov chains, the irreversibility was cleverly considered in [DHN00, CLP99] to lift, i.e. extending the state space, a process, such as random walk, to get better rates of convergence. In physics [MKK14, NMKH15] or statistics, for sampling procedure, piecewise deterministic Markov processes such as Event Chain Monte Carlo, Bouncy particle Sampler or Zig-Zag process [FBPR18, BRZ19, DGM20, BGMS23] have been developed showing remarkable numerical dimension dependence. However quantitative rates of convergence have been considered only very recently benefiting of a serie of results for the underdamped Langevin process. Indeed, a common feature of the underdamped Langevin process and those PDMP samplers is to exhibit a state space with position and velocity, and thus showing diffusive-to-ballistic speed up.

For the underdamped Langevin process, hypocoercivity has been introduced to get quantitative rate of convergence [H07, Vil09] via dissipation of H^1 type norms. In parallel,

[DMS15] rather considered a mollified L^2 norm to capture rate of convergence which has shown to be even adapted to PDMP sampler [ADNR21]. Even more recently, a variational approach was build in [AAMN19] successfully applied in [CLW23] to get sharp rates for the underdamped Langevin, retrieving the result obtained by coupling of [EGZ19], when the spectral gap is small or to PDMP sampler [LW22] exhibiting sharp dimensional dependence.

As a prelude, two of the authors [EL24b] formalized the idea behind this variational approach by adopting the notion of lift however reversing the concept. Indeed, in [DHN00, CLP99], the idea is to consider a reversible process whose convergence to equilibrium may prove slow and to extend its value space, i.e. to lift it, to get, through irreversibility, an acceleration of the convergence. Here, in fact we have some irreversible process, that we consider as a lifted process and we want to find the collapse process and exploit some of its properties to show that the lifted one has nicer convergence properties. It is done via the second order lift property introduced in the next section. It was implicitly considered in [AAMN19, LW22, CLW23] but here we aim at a general point of view. We pursue this goal by implementing the general scheme of these papers, considering a space-time approach, at an abstract level: we prove a flow Poincaré inequality, specially tailored to consider the dissipation scheme simpler to set than the space-time Poincaré inequality used in [CLW23], via a divergence lemma requiring only Dirichlet forms at the level of the collapse process under the assumption that its spectrum is purely discrete. Previously such a divergence lemma and thus a space-time Poincaré inequality were only considered for the collapse process being the reversible overdamped Langevin process. Our framework allows to retrieve the results of [LW22, CLW23] but also to prove new results for PDMP sampler such as the Forward process or the coordinate Zig-Zag process. Moreover we consider new applications to Run-and-Tumble particles with jamming: consider two RTPs on the 1D torus with velocities $\{-1, 1\}$ switching at exponential rates which jam when they collide. Such models have been considered at the discrete level by [SEB16, SEB17, MBE19] and by [HGM23] at the continuous level as a proxy for the Motility Induced Phase Separation phenomenon, as the invariant measure of the interdistance process exhibits Dirac mass at the contact of particles (i.e. at 0 and L). We will use our approach to obtain sharp rate of convergence showing in the process that the collapse is a sticky Brownian Motion [Fel52, Fel54, EP14].

The paper is divided in three more sections. Section 2 presents the concept of second order lift which is illustrated by two examples: the underdamped Langevin equation and a run-and-tumble particle confined by hard walls. Then in Section 3, we present our main results on convergence to equilibrium of lifts via Flow Poincaré inequality. It also contains the proofs and in particular a new version of divergence lemma which relies only on Dirichlet form. Finally, the last section presents applications: the underdamped Langevin process serves as a warm up to verify our set of assumptions and then we consider new processes: PDMP for sampling via the Forward Process or the one-coordinate Zig-Zag process, the Run-and-Tumble particle confined by hard walls and two Run-and-Tumble particles with jamming.

Notations

Let us end this section by commonly used notations.

- the state space of the lift is $\hat{\mathcal{S}} = \mathcal{S} \times \mathcal{V}$ where \mathcal{S} is the state space of the collapse seen as ‘position’ and \mathcal{V} is the state space of the ‘velocities’
- Position measure of equilibrium $\mu = \frac{1}{Z} e^{-U(x)} dx$ where $Z = \int e^{-U(x)} dx$
- Conditional velocity invariant measure of equilibrium $\kappa_x(dv)$
- Lifted process : generator $\hat{\mathcal{L}} = \hat{\mathcal{L}}_d + \gamma \hat{\mathcal{L}}_v$ where $\hat{\mathcal{L}}_v$ stands for the "refreshment" part of the velocity:

- For the underdamped Langevin process

$$\hat{\mathcal{L}}_d = v \cdot \nabla_x - \nabla_x U \cdot \nabla_v, \quad \hat{\mathcal{L}}_v = \Delta_v - v \cdot \nabla_v$$

- For PDMP sampling type process

$$\hat{\mathcal{L}}_d f(x, v) = v \cdot \nabla_x f(x, v) + (v \cdot \nabla_x U)_+ (Rf(x, v) - f(x, v))$$

$$\hat{\mathcal{L}}_v f(x, v) = \lambda_r (Kf(x, v) - f(x, v))$$

where for example $Kf(x, v) = \Pi_v f = \int f(x, v') d\kappa(v')$ for κ a probability measure (gaussian, ...) and R an "event" kernel (reflection, flip of velocity, ...)

- Collapsed process of generator \mathcal{L}
- time extended generator $A = -\partial_t + \hat{\mathcal{L}}_d$
- Different Dirichlet forms
 - for $f \in L^2(\mu)$ set $\mathcal{E}(f) = \langle f, -\mathcal{L}f \rangle_{L^2(\mu)}$
 - for $f \in L^2(\hat{\mu})$ set $\mathcal{E}_v(f) = \langle f, -\hat{\mathcal{L}}_v f \rangle_{L^2(\hat{\mu})}$
 - for $f \in L^2(\lambda \otimes \mu)$ set $\mathcal{E}_T(f) = \langle f, -\mathcal{L}f \rangle_{L^2(\lambda \otimes \mu)}$
 - for $f \in L^2(\lambda \otimes \hat{\mu})$ set $\mathcal{E}_{v,T}(f) = \langle f, -\hat{\mathcal{L}}_v f \rangle_{L^2(\lambda \otimes \hat{\mu})}$
- Π_v the conditional expectation i.e.

$$\Pi_v f(x, v) = \int f(x, \tilde{v}) d\kappa_x(\tilde{v}) \text{ for } f \in L^2(\hat{\mu}),$$

and similarly

$$\Pi_v f(t, x, v) = \int f(t, x, \tilde{v}) d\kappa_x(\tilde{v}) \text{ for } f \in L^2(\lambda \otimes \hat{\mu}).$$

7.2 Lift and collapse

7.2.1 Definition

Inspired by the notion of lifts of discrete Markov chains [DHN00, CLP99], the following concept of second-order lifts was introduced in [EL24a] and used to simplify the variational approach for hypocoercivity in [CLW23, LW22, AAMN19]. Let $\pi : \hat{\mathcal{S}} \rightarrow \mathcal{S}$ be a surjection between two Polish spaces $\hat{\mathcal{S}}$ and \mathcal{S} . Given a Markov process (\hat{Z}_t) on the enlarged space $\hat{\mathcal{S}}$, the idea is to define a new Markov process (Z_t) on the reduced state space \mathcal{S} , which may be seen as an averaged version of the original process. Denote

- $\hat{\mu}$ the invariant measure of (Z_t) and μ its pushforward by π ,
- $\hat{\mathcal{L}}$ (resp. \mathcal{L}) the strong generator of (\hat{Z}_t) (resp. (Z_t)) in $L^2(\hat{\mu})$ (resp. $L^2(\mu)$),
- (\hat{P}_t) (resp. (P_t)) the semigroup associated to $\hat{\mathcal{L}}$ (resp. \mathcal{L}),
- $\mathcal{E}(f, g) := \langle f, (-\mathcal{L})g \rangle_{L^2(\mu)}$ the Dirichlet form of (Z_t) .

The following definition is taken from [EL24a].

Definition 7.1 (Second-order lift). *The process (\hat{Z}_t) is a second-order lift of (Z_t) if there exists a core C of \mathcal{L} such that*

$$f \circ \pi \in \text{Dom}(\hat{\mathcal{L}}) \text{ for all } f \in C$$

and

(i) first order condition: for all $f \in C$ and $g \in L^2(\mu)$

$$\left\langle \hat{\mathcal{L}}(f \circ \pi), g \circ \pi \right\rangle_{L^2(\hat{\mu})} = 0,$$

(ii) second order condition: for all $f, g \in C$

$$\left\langle \hat{\mathcal{L}}(f \circ \pi), \hat{\mathcal{L}}(g \circ \pi) \right\rangle_{L^2(\hat{\mu})} = \langle f, (-2\mathcal{L})g \rangle_{L^2(\mu)} = 2\mathcal{E}(f, g).$$

If, in addition,

$$\{f \in L^2(\mu) : f \circ \pi \in \text{Dom}(\hat{\mathcal{L}})\} = \text{Dom}(\mathcal{E})$$

then (\hat{Z}_t) is a strong second order lift.

To get some understanding of this second order lift property, note that it is equivalent to the following limit

$$\int_{\mathcal{V}} \hat{P}_t(g \circ \pi(x, v)) \kappa_x(dv) = P_{t^2}g(x) + o(t^2)$$

in $L^2(\mu)$ as t goes to 0.

In the rest of this paper, second-order lift are simply called ‘lifts’ and (Z_t) is called the collapse of (\hat{Z}_t) . Note that (ii) implies that μ is the invariant measure of (Z_t) . Only the setting where $\hat{\mathcal{S}}$ is the product space $\mathcal{S} \times \mathcal{V}$ and π is the projection on \mathcal{S} will be considered. One can think of $x \in \mathcal{S}$ as the position and of $v \in \mathcal{V}$ as the velocity. The following examples illustrate the concept.

7.2.2 Examples

Underdamped and overdamped Langevin process

Let $\hat{\mathcal{S}} = \mathcal{S} \times \mathcal{V} = \mathbb{R}^d \times \mathbb{R}^d$ and $U \in C^1(\mathbb{R}^d)$ be a potential such that $U(x) \rightarrow +\infty$ when $|x| \rightarrow +\infty$ and $\int e^{-U(x)} dx < +\infty$.

The underdamped Langevin process $\hat{Z}_t = (X_t, V_t)$ is the $\hat{\mathcal{S}}$ -valued solution of the SDE

$$\begin{aligned} dX_t &= V_t dt, \\ dV_t &= -\nabla_x U(X_t) dt - \gamma V_t dt + \sqrt{2\gamma} dB_t, \end{aligned}$$

which has invariant measure $\hat{\mu} = \mu \otimes \kappa$ where $\mu \propto e^{-U(x)} dx$ and κ is the standard Gaussian measure on \mathbb{R}^d and its generator is

$$\hat{\mathcal{L}}f(x, v) = v \cdot \nabla_x f - \nabla_x U(x) \cdot \nabla_v f - \gamma v \cdot \nabla_v f + \gamma \Delta_v f$$

with domain $\text{Dom}(\hat{\mathcal{L}})$.

The overdamped Langevin process Z_t is the \mathcal{S} -valued solution of the SDE

$$dZ_t = -\frac{1}{2} \nabla_x U(Z_t) dt + dB_t$$

which has invariant measure μ and generator

$$\mathcal{L}f(x) = -\frac{1}{2} \nabla_x U(x) \cdot \nabla_x f + \frac{1}{2} \Delta_x f = -\frac{1}{2} \nabla_x^* \nabla_x f$$

where $*$ denotes the adjoint with respect to μ and domain $\text{Dom}(\mathcal{L})$ for which a core is constituted by all smooth compact functions.

Because $\hat{\mathcal{L}}(f \circ \pi) = v \cdot \nabla_x f$, we have

$$\int_{\hat{\mathcal{S}}} \hat{\mathcal{L}}(f \circ \pi)(g \circ \pi) d\hat{\mu} = \int_{\mathcal{S}} g(x) \left(\int_{\mathcal{V}} v \cdot \nabla_x f(x) d\kappa(v) \right) d\mu(x) = 0$$

and

$$\begin{aligned} \int_{\hat{\mathcal{S}}} \hat{\mathcal{L}}(f \circ \pi) \hat{\mathcal{L}}(g \circ \pi) d\hat{\mu} &= \int_{\mathcal{S}} \int_{\mathcal{V}} (v \cdot \nabla_x f(x)) (v \cdot \nabla_x g(x)) d\kappa(v) d\mu(x) \\ &= \int_{\mathcal{S}} \nabla_x f(x) \cdot \nabla_x g(x) d\mu(x) \\ &= \int_{\mathcal{S}} f(-2\mathcal{L}g) d\mu \end{aligned}$$

so that (\hat{Z}_t) is a lift of (Z_t) .

Run-and-tumble particle confined by hard walls

Let $L > 0$ and set $\mathcal{S} = [0, L]$, $\mathcal{V} = \{-2, 2\}$ and $\hat{\mathcal{S}} = \mathcal{S} \times \mathcal{V}$. Consider the process $\hat{Z}(t) = (x(t), v(t))$ where the velocity $v(t)$ is a Markov jump process switching between $+2$ and -2 with rate ω . If the velocity jump times are denoted by $0 = T_0 < T_1 < \dots$ then the position $x(t)$ is recursively defined by

$$x(t) = \max[0, \min[L, x(T_n) + v(T_n)(t - T_n)]] \text{ for } t \in [T_n, T_{n+1}].$$

In a few words, between two velocity jumps, the position $x(t)$ moves at constant speed $v(t) \in \{+2, -2\}$ until it hits either $x = 0$ or $x = L$, at which point it stops and stays ‘glued’ to that value. The process $\hat{Z}(t) = (x(t), v(t))$ will be called the single RTP process.

Its Feller generator $\hat{\mathcal{L}}_F$ is given by

$$\hat{\mathcal{L}}_F f(x, v) = v \partial_x f(x, v) + \omega(f(x, -v) - f(x, v))$$

with domain

$$\text{Dom}(\hat{\mathcal{L}}_F) = \{f(\cdot, \pm 2) \in C^1([0, L]) \text{ and } \partial_x f(L, 2) = \partial_x f(0, -2) = 0\}$$

and it has invariant measure

$$\hat{\mu} = \left(\frac{1}{2 + \omega L} \delta_L + \frac{\omega/2}{2 + \omega L} dx \right) \otimes \delta_2 + \left(\frac{\omega/2}{2 + \omega L} dx + \frac{1}{2 + \omega L} \delta_0 \right) \otimes \delta_{-2}.$$

The Feller domain $\text{Dom}(\hat{\mathcal{L}}_F)$ is a subset of the $L^2(\hat{\mu})$ domain $\text{Dom}(\hat{\mathcal{L}})$ which is dense in $L^2(\hat{\mu})$ and stable by the $L^2(\hat{\mu})$ semigroup, so by [EN01, Chapter II Proposition 1.7] it is a core of the strong $L^2(\hat{\mu})$ generator $(\hat{\mathcal{L}}, \text{Dom}(\hat{\mathcal{L}}))$.

The next step is to identify a candidate for the collapse of the single RTP process. For $f \in C_c^\infty(0, L)$ we have

$$\hat{\mathcal{L}}(f \circ \pi) = v \partial_x f \text{ and } \hat{\mathcal{L}}^* \hat{\mathcal{L}}(f \circ \pi) = v^2 \partial_x^2 f + \omega(-v \partial_x f(x, -v) - v \partial_x f(x, v))$$

so that the second order condition yields

$$\mathcal{L}f(x) = -\frac{1}{2} \int \hat{\mathcal{L}}^* \hat{\mathcal{L}}f(x, v) d\kappa_x(v) = 2\partial_x f(x).$$

This suggests that the collapse is a diffusion. Because

$$\mu = \frac{1}{2 + \omega L} \delta_0 + \frac{\omega}{2 + \omega L} dx + \frac{1}{2 + \omega L} \delta_L,$$

is invariant for the collapse, one has for all $f \in \text{Dom}(\mathcal{L})$ that

$$\begin{aligned} 0 &= \int 2\partial_x^2 f d\mu, \\ &= \frac{1}{2 + \omega L} (2\partial_x^2 f(0)) + \frac{\omega}{2 + \omega L} \int_0^L 2\partial_x^2 f(x) dx + \frac{1}{2 + \omega L} (2\partial_x^2 f(L)), \\ &= \frac{2}{2 + \omega L} (\partial_x^2 f(0) - \omega \partial_x f(0) + \partial_x^2 f(L) + \omega \partial_x f(L)), \end{aligned}$$

pointing to the sticky Brownian motion with boundary conditions $\partial_x^2 f(0) = \omega \partial_x f(0)$ and $\partial_x^2 f(L) = -\omega \partial_x f(L)$.

The sticky Brownian motion is thus a natural candidate for the collapse. It is sufficient to check that it satisfies Definition 7.1. The Feller generator of the sticky Brownian motion

$$\mathcal{L}_F f(x) = 2\partial_x^2 f(x)$$

is shown to have domain

$$\text{Dom}(\mathcal{L}_F) = \{f \in C^2([0, L]) : \partial_x^2 f(0) = \omega \partial_x f(0) \text{ and } \partial_x^2 f(L) = -\omega \partial_x f(L)\}$$

in [Lig10] so, again by [EN01, Chapter II Proposition 1.7], $\text{Dom}(\mathcal{L}_F)$ is a core of the strong $L^2(\mu)$ generator \mathcal{L} of the sticky Brownian motion.

For $f \in \text{Dom}(\mathcal{L}_F)$ one has $f \circ \pi \in \text{Dom}(\hat{\mathcal{L}})$ and $\hat{\mathcal{L}}(f \circ \pi) = 1_{\{0 < x < L\}} v f$ and thus

$$\left\langle \hat{\mathcal{L}}(f \circ \pi), g \circ \pi \right\rangle_{L^2(\hat{\mu})} = \sum_{v=\pm 2} \frac{\omega}{2 + \omega L} \int_0^L (v \partial_x f) g dx = 0$$

and

$$\begin{aligned} & -\frac{1}{2} \left\langle \hat{\mathcal{L}}(f \circ \pi), \hat{\mathcal{L}}(g \circ \pi) \right\rangle_{L^2(\hat{\mu})} \\ &= -\frac{1}{2} \sum_{v=\pm 2} \frac{\omega}{2 + \omega L} \int_0^L (v \partial_x f)(v \partial_x g) dx \\ &= \frac{-2\omega}{2 + \omega L} \int_0^L \partial_x f \partial_x g dx \\ &= \frac{-2\omega}{2 + \omega L} \left[\partial_x f(L)g(L) - \partial_x f(0)g(0) - \int_0^L \partial_x^2 f(x)g(x) dx \right] \\ &= \frac{-2\omega}{2 + \omega L} \left[-\frac{1}{\omega} \partial_x^2 f(L)g(L) - \frac{1}{\omega} \partial_x^2 f(0)g(0) - \int_0^L \partial_x^2 f(x)g(x) dx \right] \\ &= \int (\mathcal{L}f)g d\mu. \end{aligned}$$

This confirms that the sticky Brownian motion is the collapse of the single RTP process.

7.2.3 Rate of convergence to equilibrium of lift/collapse

By introducing formally the notion of lift/collapse process in L^2 , Eberle and Lörler [EL24a] have also presented a nice property linking the rate of convergence to equilibrium of the lift process with the spectral gap of the collapse process.

Theorem 7.2 (Eberle-Lörler [EL24a, Theorem 11]). *Suppose that $(P_t)_{t \geq 0}$ with generator $(\mathcal{L}, \text{Dom}(\mathcal{L}))$ is the transition semigroup on $L^2(S, \mu)$ of a reversible Markov process with spectral gap $\text{Gap}(\mathcal{L})$, and that $(\hat{P}_t)_{t \geq 0}$ with generator $(\hat{\mathcal{L}}, \text{Dom}(\hat{\mathcal{L}}))$ is the transition semigroup*

on $L^2(\mathcal{S} \times \mathcal{V}, \hat{\mu})$ of an arbitrary second order lift of \mathcal{L} . Assume that there exists positive C and ν such that for all $\hat{\mu}$ -centered f and $t > 0$

$$\|\hat{P}_t f\|_{L^2(\hat{\mu})} \leq C e^{-\nu t} \|f\|_{L^2(\hat{\mu})}.$$

Then one has

$$\nu \leq C \sqrt{2 \text{Gap}(\mathcal{L})}. \quad (7.1)$$

From (7.1), a clear limitation is given to the rate of convergence to equilibrium, depending on the spectral gap of the collapse process. It also characterizes lifts which are optimal in the sense that ν is of order of $\sqrt{\text{Gap}(\mathcal{L})}$ when $\text{Gap}(\mathcal{L})$ is small. We will see in applications that it is precisely the case for the underdamped Langevin process with a well chosen friction parameter.

7.3 Main results

In this section, the variational approach for convergence results in [AAMN19, CLW23, LW22, EL24a] is extended to treat a broad class of examples in a unified way. One may start by listing the assumptions on a lift $\hat{\mathcal{L}}$ of \mathcal{L} in order to prove exponential convergence.

The first set of assumptions ensures that the lift can be split into an antisymmetric part acting on both $x \in \mathcal{S}$ and $v \in \mathcal{V}$ and a refresh term which acts only on v .

Assumption (A) Assume $\hat{\mathcal{L}} = \hat{\mathcal{L}}_d + \gamma \hat{\mathcal{L}}_v$ where $\gamma \geq 0$ and

- the generator $\hat{\mathcal{L}}_d$ is a lift of \mathcal{L} that is weakly antisymmetric i.e. there exists a core C of \mathcal{L} such that

$$g \in \text{Dom}(\mathcal{L}) \implies g \circ \pi \in \text{Dom}(\hat{\mathcal{L}}_d^*)$$

and

$$\hat{\mathcal{L}}_d^*(g \circ \pi) = -\hat{\mathcal{L}}_d(g \circ \pi).$$

- the generator $\hat{\mathcal{L}}_v$ acts only on v and leaves the conditional distribution of v invariant in the sense that

$$\int \hat{\mathcal{L}}_v f d\kappa_x(v) = 0 \text{ for all } x \in \mathcal{S}$$

and there is a spectral gap in v meaning that

$$\|f - \Pi_v f\|_{L^2(\hat{\mu})}^2 \leq c_2^* \mathcal{E}_v(f).$$

This splitting into two distinct dynamics is natural for the underdamped Langevin dynamics

$$\hat{\mathcal{L}}f(x, v) = \underbrace{v \cdot \nabla_x f - \nabla_x U(x) \cdot \nabla_v f}_{=\hat{\mathcal{L}}_d} \underbrace{-\gamma v \cdot \nabla_v f + \gamma \Delta_v f}_{=\gamma \hat{\mathcal{L}}_v}$$

where $\hat{\mathcal{L}}_d$ is the generator of the deterministic Hamiltonian dynamics and $\hat{\mathcal{L}}_v$ is the generator of the Ornstein-Uhlenbeck dynamics in v .

Remark 7.3. If $\hat{\mathcal{L}} = \hat{\mathcal{L}}_d + \gamma \hat{\mathcal{L}}_v$ and $\hat{\mathcal{L}}_v$ only acts on v in the sense that $\hat{\mathcal{L}}_v(f \circ \pi) = 0$ then $\hat{\mathcal{L}}$ is a lift of \mathcal{L} if and only if $\hat{\mathcal{L}}_d$ is a lift of \mathcal{L} .

In line with the general principles of hypocoercivity, the idea is that the dynamics in v has strong relaxation properties that are then propagated to the variable x by the rest of the dynamics. It is thus natural to ask for a Poincaré inequality for the refreshment part of the full dynamics.

It is also required that the collapse satisfies a Poincaré inequality and also the additional assumption that it has a purely discrete spectral decomposition.

Assumption (B) Assume that the spectrum of \mathcal{L} is discrete and that

$$\|f - \mu(f)\|_{L^2(\mu)}^2 \leq \frac{1}{m} \mathcal{E}(f)$$

for all $f \in \text{Dom}(\mathcal{E})$, i.e. $m = \text{Gap}(\mathcal{L})$.

This assumption of a purely discrete spectrum, which was also considered in [CLW23], is surely too strong but allows for simple calculations and most importantly permits to prove a divergence lemma not limited to the overdamped Langevin process but for general reversible Markov process, as will be seen in the proof. It is precisely the goal of [BS23] which aims at removing this technical assumption. The Poincaré inequality is however natural for exponential convergence to equilibrium and allows us to differentiate the regimes of convergence of the overdamped and underdamped Fokker-Planck equation. In order to get subexponential convergence to equilibrium, it may be relaxed to a weak Poincaré inequality assumption which is one of the purpose of [BSWW24].

The last set of assumptions, which may seem technical, ensures good interactions between $\hat{\mathcal{L}}_d$ and the refreshment $\hat{\mathcal{L}}_v$. It is reminiscent of the assumptions (H_3) and (H_4) in the seminal work of [DMS15].

Assumption (C) Assume there exist $c_0^*, c_1^* > 0$ such that

$$\left| \left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\hat{\mu})} \right| \leq c_0^* \|2\mathcal{L}g\|_{L^2(\mu)} \|f - \Pi_v f\|_{L^2(\hat{\mu})}, \quad (\text{C1})$$

and

$$\left| \left\langle \hat{\mathcal{L}}_v f, \hat{\mathcal{L}}_d(g \circ \pi) \right\rangle_{L^2(\hat{\mu})} \right| \leq c_1^* \sqrt{\mathcal{E}_v(f)} \sqrt{2\mathcal{E}(g)}. \quad (\text{C2})$$

Let us give a few remarks about this last assumption (C).

Remark 7.4. *This first remark gives a setting where (C1) may be easily verified. Note that if $\hat{\mathcal{L}}_d(g \circ \pi) \in \text{Dom}(\hat{\mathcal{L}}_d^*)$ for all $g \in \text{Dom}(\mathcal{L})$ then the second order lift condition becomes*

$$\int [\hat{\mathcal{L}}_d^* \hat{\mathcal{L}}_d(g \circ \pi)](x, v) d\kappa_x(v) = -2\mathcal{L}g(x) \iff \Pi_v \hat{\mathcal{L}}_d^* \hat{\mathcal{L}}_d(g \circ \pi) = -(2\mathcal{L}g) \circ \pi$$

and thus

$$\begin{aligned} & \left| \left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\hat{\mu})} \right| \\ &= \left| \left\langle \hat{\mathcal{L}}_d^* \hat{\mathcal{L}}_d(g \circ \pi), f - \Pi_v f \right\rangle_{L^2(\hat{\mu})} \right| \\ &= \left| \left\langle \hat{\mathcal{L}}_d^* \hat{\mathcal{L}}_d(g \circ \pi) - \Pi_v \hat{\mathcal{L}}_d^* \hat{\mathcal{L}}_d(g \circ \pi), f - \Pi_v f \right\rangle_{L^2(\hat{\mu})} \right| \\ &= \left| \left\langle \hat{\mathcal{L}}_d^* \hat{\mathcal{L}}_d(g \circ \pi) + (2\mathcal{L}g) \circ \pi, f - \Pi_v f \right\rangle_{L^2(\hat{\mu})} \right| \\ &\leq \| \hat{\mathcal{L}}_d^* \hat{\mathcal{L}}_d(g \circ \pi) + (2\mathcal{L}g) \circ \pi \|_{L^2(\hat{\mu})} \|f - \Pi_v f\|_{L^2(\hat{\mu})}. \end{aligned}$$

So if in addition

$$\| \hat{\mathcal{L}}_d^* \hat{\mathcal{L}}_d(g \circ \pi) + (2\mathcal{L}g) \circ \pi \|_{L^2(\hat{\mu})} \leq c_0^* \|2\mathcal{L}g\|_{L^2(\mu)},$$

then

$$\left| \left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\hat{\mu})} \right| \leq c_0^* \|2\mathcal{L}g\|_{L^2(\mu)} \|f - \Pi_v f\|_{L^2(\hat{\mu})}.$$

Remark 7.5. *This second remark provides sufficient condition for (C2). Note that*

$$\left| \left\langle \hat{\mathcal{L}}_v f, \hat{\mathcal{L}}_d(g \circ \pi) \right\rangle_{L^2(\hat{\mu})} \right| \leq \|\hat{\mathcal{L}}_v f\|_{L^2(\mu)} \|\hat{\mathcal{L}}_d(g \circ \pi)\|_{L^2(\hat{\mu})} = \|\hat{\mathcal{L}}_v f\|_{L^2(\mu)} \sqrt{2\mathcal{E}(g)}$$

so when $\hat{\mathcal{L}}_v f = \Pi_v f - f$ is the full refresh w.r.t. the conditional expectation of v , we have

$$\|\hat{\mathcal{L}}_v f\|_{L^2(\hat{\mu})}^2 = \left\langle f, -\hat{\mathcal{L}}_v f \right\rangle_{L^2(\hat{\mu})} = \mathcal{E}_v(f)$$

and the second inequality in assumption (C) holds with $c_1^* = 1$.

Remark 7.6. *As will be seen from the proof, one may have alternative versions of (C1). One appears most naturally when the underlying potential U is not convex, while using Bochner's equality:*

$$\left| \left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\hat{\mu})} \right| \leq \tilde{c}_0^* \left(\|2\mathcal{L}g\|_{L^2(\mu)} + \sqrt{\mathcal{E}(g)} \right) \|f - \Pi_v f\|_{L^2(\hat{\mu})}. \quad (C^*1)$$

It leads to slightly different order of constants. We refer to the applications for further details.

To get an exponential decay in L^2 norm, one often uses the principle of energy dissipation, which amounts to differentiating along the L^2 norm along the flow and to control this dissipation by the original L^2 norm. In the elliptic (often reversible) setting, it is precisely where a Poincaré inequality is needed. However such a Poincaré inequality is no longer valid in the hypoelliptic setting such as the underdamped Fokker-Planck equation. In [DMS15, Vil09], the choice is made to add a "mixing" term to the L^2 norm to ensure the control of the dissipation. Here, following [AAMN19, CLW23, EL24a], we consider the integral wrt time of the L^2 norm and thus need and integrated Poincaré inequality which is called "Flow Poincaré inequality".

Theorem 7.7 (Flow Poincaré inequality). *If (A), (B) and (C) hold, then for $f_0 \in L_0^2(\hat{\mu})$ and $f(t, x, v) = (P_t f_0)(x, v)$ we have*

$$\|f\|_{L^2(\lambda \otimes \hat{\mu})}^2 \leq (\gamma^2 C_0(T) + C_1(T)) \mathcal{E}_{v,T}(f),$$

where

$$\begin{aligned} C_0(T) &= C(c_1^*(T + m^{-1/2}))^2, \\ C_1(T) &= Cc_2^*(1 + c_0^*)^2. \end{aligned}$$

Notice that this Flow Poincaré inequality differs from the one encountered in [CLW23, LW22, EL24a], called space-time Poincaré inequality, because it is only valid along the flow. This is a small price to pay, as it is meant to be applied to a solution of the Fokker-Planck equation, but allows us to extend it to a broad class of refresh mechanisms. Our scheme of proof allows us also to prove such a space-time Poincaré inequality.

One of the main tool to get such a Flow Poincaré inequality is called divergence lemma. There are now many different works [CLW23, BS23, EL24a] which aims at providing a neat divergence lemma under general hypotheses. The goal here is to take a Dirichlet form approach enabling to get quite general assumption on the collapsed process, however imposing (not so) strong assumption, namely spectral gap and discrete spectrum, by adapting [EL24a] to our setting.

With all these ingredients one may now present the main result, adopting the terminology of [EL24a].

Theorem 7.8 (Exponential decay). *The semigroup \hat{P}_t generated by $\hat{\mathcal{L}} = \hat{\mathcal{L}}_d + \gamma \hat{\mathcal{L}}_v$ is exponentially contractive in T -average with rate*

$$\nu = \frac{\gamma}{\gamma^2 C_0(T) + C_1(T)}.$$

i.e. for all centered f wrt $\hat{\mu}$,

$$\frac{1}{T} \int_t^{t+T} \|P_s f\|_{L^2(\hat{\mu})} ds \leq e^{-\nu t} \|f\|_{L^2(\hat{\mu})}. \quad (7.2)$$

Note that this implies that for all t

$$\|P_t f\|_{L^2(\hat{\mu})} \leq e^{-\nu(t-T)} \|f\|_{L^2(\hat{\mu})}.$$

Let us prove quickly how the Flow Poincaré inequality may be used to get Theorem 7.8.

Proof. Let $f \in L_0^2(\hat{\mu}) = \{f \in L^2(\hat{\mu}) \text{ and } \int f d\hat{\mu} = 0\}$. Then

$$\begin{aligned} \frac{d}{dt} \int_t^{t+T} \|\hat{P}_s f\|_{L^2(\hat{\mu})}^2 ds &= \|\hat{P}_{t+T} f\|_{L^2(\hat{\mu})}^2 - \|\hat{P}_t f\|_{L^2(\hat{\mu})}^2 \\ &= \int_t^{t+T} \frac{d}{ds} \|\hat{P}_s f\|_{L^2(\hat{\mu})}^2 ds \\ &= 2 \int_t^{t+T} \left\langle \hat{P}_s f, (\hat{\mathcal{L}}_d + \gamma \hat{\mathcal{L}}_v) \hat{P}_s f \right\rangle_{L^2(\hat{\mu})} ds \\ &\leq 2\gamma \int_t^{t+T} \left\langle \hat{P}_s f, \hat{\mathcal{L}}_v \hat{P}_s f \right\rangle_{L^2(\hat{\mu})} ds \\ &= -2\gamma \mathcal{E}_{v,T}(\hat{P}_{t+} f). \end{aligned}$$

Theorem 7.7 states

$$\|\hat{P}_{t+} f\|_{L^2(\lambda \otimes \hat{\mu})}^2 \leq (\gamma^2 C_0 + C_1) \mathcal{E}_{v,T}(\hat{P}_{t+} f)$$

so that

$$\frac{d}{dt} \int_t^{t+T} \|\hat{P}_s f\|_{L^2(\hat{\mu})}^2 ds \leq - \underbrace{\frac{2\gamma}{\gamma^2 C_0 + C_1}}_{=: \nu} \int_t^{t+T} \|\hat{P}_s f\|_{L^2(\hat{\mu})}^2 ds.$$

Grönwall's inequality yields

$$\frac{1}{T} \int_t^{t+T} \|\hat{P}_s f\|_{L^2(\hat{\mu})}^2 ds \leq e^{-2\nu t} \|f\|_{L^2(\hat{\mu})}^2.$$

Finally, the Cauchy-Schwarz inequality yields

$$\frac{1}{T} \int_t^{t+T} \|\hat{P}_s f\|_{L^2(\hat{\mu})} ds \leq \left(\frac{1}{T} \int_t^{t+T} \|\hat{P}_s f\|_{L^2(\hat{\mu})}^2 ds \right)^{1/2} \leq e^{-\nu t} \|f\|_{L^2(\hat{\mu})}$$

and thus the ν -exponential contractivity in T -average. \square

The proof of the Flow Poincaré inequality requires first to work on the collapse process via the divergence lemma.

7.3.1 Divergence lemma and space-time property of the lift

The following technical lemma is at the heart of the proof of the flow Poincaré inequality. As alluded to before, this divergence lemma, which has fundamental importance in fluid dynamics [Gal11], was investigated recently by [CLW23, BS23, EL24a]. The goal is to provide a Dirichlet form version, enabling thus to be used with minimal structure.

Lemma 7.9 (Divergence lemma). *For $f \in L_0^2(\lambda \otimes \mu)$ there exist $h, g \in L^2(\lambda \otimes \mu)$ with*

$$\begin{aligned} h(t, \cdot), g(t, \cdot) &\in \text{Dom}(\mathcal{L}) \text{ for } \lambda\text{-a.e. } t, \\ h(\cdot, x) &\in H^1(0, T) \text{ for } \mu\text{-a.e. } x, \\ g(\cdot, x) &\in H^2(0, T) \text{ for } \mu\text{-a.e. } x, \end{aligned}$$

such that

$$\partial_t h - 2\mathcal{L}g = f$$

with boundary conditions

$$\begin{aligned} h(t=0, \cdot) &= h(t=T, \cdot) = 0, \\ g(t=0, \cdot) &= g(t=T, \cdot) = 0, \end{aligned}$$

and

$$\begin{aligned} \|2\mathcal{L}g\|_{L^2(\lambda \otimes \mu)} &\leq c_1 \|f\|_{L^2(\lambda \otimes \mu)}, \\ \sqrt{2\mathcal{E}_T(h)} &\leq c_2 \|f\|_{L^2(\lambda \otimes \mu)}, \\ \sqrt{2\mathcal{E}_T(g)} &\leq c_3 \|f\|_{L^2(\lambda \otimes \mu)}, \\ \sqrt{2\mathcal{E}_T(\partial_t g)} &\leq c_4 \|f\|_{L^2(\lambda \otimes \mu)}. \end{aligned}$$

The c_i can be taken to be

$$\begin{aligned} c_1 &= C, & c_2 &= C, \\ c_3 &= C(T + m^{-1/2}), & c_4 &= C(1 + T^{-1}m^{-1/2}), \end{aligned}$$

where C denotes different absolute constants.

Its proof is delayed until the next section. One may now present what is called space-time property of lifts and constitutes the last ingredient for the proof of the flow Poincaré inequality.

Lemma 7.10 (Space-time property of lifts). *We have*

$$\left\langle A(f \circ \pi), h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} = -\langle h, \partial_t f \rangle_{L^2(\lambda \otimes \mu)} + 2\mathcal{E}(f, g)$$

where $Af = -\partial_t f + \hat{\mathcal{L}}_d f$.

Proof. We have

$$\begin{aligned} \left\langle h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi), A(f \circ \pi) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} &= \langle h \circ \pi, -\partial_t(f \circ \pi) \rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &\quad + \left\langle h \circ \pi, \hat{\mathcal{L}}_d(f \circ \pi) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &\quad + \left\langle \hat{\mathcal{L}}_d(g \circ \pi), -\partial_t(f \circ \pi) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &\quad + \left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f \circ \pi) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \end{aligned}$$

and it follows from the first order property of lifts that

$$\left\langle h \circ \pi, \hat{\mathcal{L}}_d(f \circ \pi) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} = \left\langle \hat{\mathcal{L}}_d(g \circ \pi), -\partial_t(f \circ \pi) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} = 0.$$

□

7.3.2 Proof of the flow Poincaré inequality

The constants C_0 and C_1 can be taken to be

$$\begin{aligned} C_0 &= 2(c_1^* c_3)^2, \\ C_1 &= c_2^*(1 + 2(c_4 + c_2 + c_0^* c_1)^2), \end{aligned}$$

as we now show.

Proof of Theorem 7.7. Let $\tilde{f} \in L^2(\lambda \otimes \mu)$ be such that $\Pi_v f = \tilde{f} \circ \pi$. Let h and g be as in Lemma 7.9 for $f = \tilde{f}$. Because Π_v is an orthogonal projection, we have

$$\|f\|_{L^2(\lambda \otimes \hat{\mu})}^2 = \|\tilde{f}\|_{L^2(\lambda \otimes \mu)}^2 + \|f - \Pi_v f\|_{L^2(\lambda \otimes \hat{\mu})}^2.$$

Furthermore

$$\begin{aligned} \|\tilde{f}\|_{L^2(\lambda \otimes \mu)}^2 &= \left\langle \tilde{f}, \tilde{f} \right\rangle_{L^2(\lambda \otimes \mu)} \\ &= \left\langle \partial_t h - 2\mathcal{L}g, \tilde{f} \right\rangle_{L^2(\lambda \otimes \mu)} \\ &= - \left\langle h, \partial_t \tilde{f} \right\rangle_{L^2(\lambda \otimes \mu)} + 2\mathcal{E}(f, g) \\ &= \left\langle h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi), A(\tilde{f} \circ \pi) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &= \underbrace{\left\langle h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi), Af \right\rangle_{L^2(\lambda \otimes \hat{\mu})}}_{\text{(I)}} \\ &\quad + \underbrace{\left\langle h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi), A(\Pi_v f - f) \right\rangle_{L^2(\lambda \otimes \hat{\mu})}}_{\text{(II)}} \end{aligned}$$

using first integration by parts with $h(t=0, \cdot) = h(t=T, \cdot) = 0$ and then the space-time property of lifts 7.10.

Because $f(t, x, v) = (P_t f_0)(x, v)$ we have $Af = -\partial_t f + \hat{\mathcal{L}}_d f = -\gamma \hat{\mathcal{L}}_v f$ so

$$\langle h \circ \pi, Af \rangle_{L^2(\lambda \otimes \hat{\mu})} = -\gamma \int_0^T \int_S h(x) \underbrace{\int_{\mathcal{V}} \hat{\mathcal{L}}_v f d\kappa_x(v)}_{=0} d\mu(x) dt = 0$$

by assumption (A) and

$$\begin{aligned} \left\langle h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi), Af \right\rangle_{L^2(\lambda \otimes \hat{\mu})} &= -\gamma \left\langle \hat{\mathcal{L}}_v f, \hat{\mathcal{L}}_d(g \circ \pi) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &\leq c_1^* \gamma \sqrt{\mathcal{E}_{v,T}(f)} \sqrt{2\mathcal{E}_T(g)} \\ &\leq c_1^* c_3 \gamma \|\tilde{f}\|_{L^2(\lambda \otimes \mu)} \sqrt{\mathcal{E}_{v,T}(f)} \end{aligned}$$

using (C2) and Cauchy-Schwarz.

Turning to the estimation of (II)

$$\begin{aligned} &\left\langle h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi), A(\Pi_v f - f) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &= \left\langle h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi), -\partial_t(\Pi_v f - f) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &\quad + \left\langle h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(\Pi_v f - f) \right\rangle_{L^2(\lambda \otimes \hat{\mu})}. \end{aligned}$$

Because

$$\begin{aligned} h(t=0, \cdot) &= h(t=T, \cdot) = 0, \\ g(t=0, \cdot) &= g(t=T, \cdot) = 0, \end{aligned}$$

and because $\partial_t h \circ \pi$ does not depend on v we have

$$\begin{aligned} &\left\langle h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi), -\partial_t(\Pi_v f - f) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &= \left\langle \partial_t h \circ \pi + \partial_t \hat{\mathcal{L}}_d(g \circ \pi), \Pi_v f - f \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &= \left\langle \partial_t \hat{\mathcal{L}}_d(g \circ \pi), \Pi_v f - f \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &\leq \|\partial_t \hat{\mathcal{L}}_d(g \circ \pi)\|_{L^2(\lambda \otimes \hat{\mu})} \|f - \Pi_v f\|_{L^2(\lambda \otimes \hat{\mu})}. \end{aligned}$$

Furthermore, using the fact that ∂_t commutes with $\hat{\mathcal{L}}_d(\cdot \circ \pi)$ we get

$$\left\| \partial_t \hat{\mathcal{L}}_d(g \circ \pi) \right\|_{L^2(\lambda \otimes \hat{\mu})} = \left\| \hat{\mathcal{L}}_d((\partial_t g) \circ \pi) \right\|_{L^2(\lambda \otimes \hat{\mu})} = \sqrt{\mathcal{E}_T(\partial_t g)}.$$

We deduce

$$\begin{aligned} &\left\langle h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi), -\partial_t(\Pi_v f - f) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &\leq \sqrt{\mathcal{E}_T(\partial_t g)} \|f - \Pi_v f\|_{L^2(\lambda \otimes \hat{\mu})} \\ &\leq c_4 \left\| \tilde{f} \right\|_{L^2(\lambda \otimes \mu)} \|f - \Pi_v f\|_{L^2(\lambda \otimes \mu)} \\ &\leq \sqrt{c_2^*} c_4 \left\| \tilde{f} \right\|_{L^2(\lambda \otimes \mu)} \sqrt{\mathcal{E}_{v,T}(f)}. \end{aligned}$$

Finally

$$\begin{aligned} &\left\langle h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &= \left\langle h \circ \pi, \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} + \left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &= \left\langle \hat{\mathcal{L}}_d^*(h \circ \pi), f - \Pi_v f \right\rangle_{L^2(\lambda \otimes \hat{\mu})} + \left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &= \left\langle -\hat{\mathcal{L}}_d(h \circ \pi), f - \Pi_v f \right\rangle_{L^2(\lambda \otimes \hat{\mu})} + \left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &\leq \left\| \hat{\mathcal{L}}_d(h \circ \pi) \right\|_{L^2(\lambda \otimes \hat{\mu})} \|f - \Pi_v f\|_{L^2(\lambda \otimes \hat{\mu})} \\ &\quad + c_0^* \|2\mathcal{L}g\|_{L^2(\lambda \otimes \mu)} \|f - \Pi_v f\|_{L^2(\lambda \otimes \mu)} \end{aligned}$$

using the weak antisymmetry of $\hat{\mathcal{L}}_d$ and assumption (C). Hence, using the lemma 7.9, we get

$$\begin{aligned} &\left\langle h \circ \pi + \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\lambda \otimes \hat{\mu})} \\ &\leq \sqrt{2\mathcal{E}_T(h)} \|f - \Pi_v f\|_{L^2(\lambda \otimes \hat{\mu})} + c_0^* c_1 \left\| \tilde{f} \right\|_{L^2(\lambda \otimes \mu)} \|f - \Pi_v f\|_{L^2(\lambda \otimes \mu)} \\ &\leq (c_2 + c_0^* c_1) \left\| \tilde{f} \right\|_{L^2(\lambda \otimes \mu)} \|f - \Pi_v f\|_{L^2(\lambda \otimes \hat{\mu})} \\ &\leq \sqrt{c_2^*} (c_2 + c_0^* c_1) \left\| \tilde{f} \right\|_{L^2(\lambda \otimes \mu)} \sqrt{\mathcal{E}_{v,T}(f)}. \end{aligned}$$

Combining everything and using the inequality $(a + b)^2 \leq 2a^2 + 2b^2$ leads to

$$\begin{aligned} \|f\|_{L^2(\lambda \otimes \hat{\mu})}^2 &\leq \|\Pi_v f - f\|_{L^2(\lambda \otimes \hat{\mu})}^2 + \left(c_1^* c_3 \gamma + \sqrt{c_2^*} c_4 + \sqrt{c_2^*} (c_2 + c_0^* c_1) \right)^2 \mathcal{E}_{v,T}(f) \\ &\leq \left(c_2^* + \left(c_1^* c_3 \gamma + \sqrt{c_2^*} c_4 + \sqrt{c_2^*} (c_2 + c_0^* c_1) \right)^2 \right) \mathcal{E}_{v,T}(f) \\ &\leq \left[2(c_1^* c_3 \gamma)^2 + c_2(1 + 2[c_4 + c_2 + c_0^* c_1]^2) \right] \mathcal{E}_{v,T}(f) \end{aligned}$$

as announced. \square

7.3.3 Proof of Lemma 7.9

The proof follows the strategy of [CLW23, EL24b] adapted to our setting, see also [EL24a] for a declinaison on the case of the overdamped Langevin with reflection on a domain.

Proof of Lemma 7.9. Consider an orthonormal basis $(e_k : k \geq 0)$ of eigenfunctions of \mathcal{L} with eigenvalues $-\alpha_k^2$ where e_0 is constant and $\alpha_0 = 0$. In particular, $\inf\{\alpha_k^2 : k > 0\} = m$ is the spectral gap of \mathcal{L} . Then $(h_k^a : k \geq 0) \cup (h_k^s : k > 0)$ given by

$$\begin{aligned} h_0^a(t, x) &= (t - (T - t))e_0(x), \\ h_k^a(t, x) &= \left(e^{-\alpha_k t} - e^{-\alpha_k(T-t)} \right) e_k(x) \text{ for } k > 0, \\ h_k^s(t, x) &= \left(e^{-\alpha_k t} + e^{-\alpha_k(T-t)} \right) e_k(x) \text{ for } k > 0, \end{aligned}$$

defines an orthogonal basis of

$$\begin{aligned} \mathbb{H} &= \{u(\cdot, x) \in H^2(0, T) \text{ for } \mu\text{-a.e. } x \text{ and } u(t, \cdot) \in \text{Dom}(\mathcal{L}) \text{ for } \lambda\text{-a.e. } t\} \\ &\cap \{\partial_{tt}u + \mathcal{L}u = 0 \text{ and } (\lambda \otimes \mu)(u) = 0\}. \end{aligned}$$

One may decompose the space of 0-mean L^2 -functions into

$$L_0^2(\lambda \otimes \mu) = \mathbb{H} \oplus \mathbb{H}^\perp = \mathbb{H}_{l,a} \oplus \mathbb{H}_{l,s} \oplus \mathbb{H}_{h,a} \oplus \mathbb{H}_{h,s} \oplus \mathbb{H}^\perp$$

where

$$\begin{aligned} \mathbb{H}_{l,a} &= \text{Span} \left\{ h_k^a : \alpha_k \leq \frac{2}{T} \right\}, \\ \mathbb{H}_{l,s} &= \text{Span} \left\{ h_k^s : \alpha_k \leq \frac{2}{T} \right\}, \\ \mathbb{H}_{h,a} &= \text{Span} \left\{ h_k^a : \alpha_k > \frac{2}{T} \right\}, \\ \mathbb{H}_{h,s} &= \text{Span} \left\{ h_k^s : \alpha_k > \frac{2}{T} \right\}, \end{aligned}$$

and prove the lemma separately on each subspace.

Note that, in the definition of \mathbb{H} , there can be boundary conditions in x coming from the domain of \mathcal{L} but, crucially, there are no conditions on the boundary conditions at $t = 0$ or $t = T$.

Case 1: $f \in \mathbb{H}^\perp$ On the set of 0-mean functions, one may construct $\bar{G} = (-\partial_{tt} - 2\mathcal{L})^{-1}$ using Lax-Milgram (with reflecting i.e. Neumann boundary conditions in t and boundary conditions in x coming from the domain of \mathcal{L}) and then set

$$u = \bar{G}f \text{ i.e. } u \text{ is the unique solution of } -\partial_{tt}u - 2\mathcal{L}u = f$$

as well as $g = u$ and $h = -\partial_t u$.

Because of the reflecting boundary conditions in t one has by construction that

$$h(t = 0, \cdot) = h(t = T, \cdot) = 0.$$

For $v \in \mathbb{H}$ one has

$$\begin{aligned} 0 &= \langle f, v \rangle_{L^2(\lambda \otimes \mu)} = \langle -\partial_{tt}u - 2\mathcal{L}u, v \rangle_{L^2(\lambda \otimes \mu)} \\ &= \int_S u(T, x) \partial_t v(T, x) - u(0, x) \partial_t v(0, x) d\mu(x) + \langle u, -\partial_{tt}v - 2\mathcal{L}v \rangle_{L^2(\lambda \otimes \mu)} \\ &= \int_S u(T, x) \partial_t v(T, x) - u(0, x) \partial_t v(0, x) d\mu(x) \end{aligned}$$

using integration by parts and $-\partial_{tt}v - 2\mathcal{L}v = 0$ for the last two equalities. It follows

$$u|_{t=0} = u|_{t=T} = 0$$

and thus the boundary conditions for g are satisfied as well.

On the set of 0-mean functions, the operators $-\partial_{tt}$ and $-2\mathcal{L}$ are two commuting operators with positive discrete spectrum. Hence the operator norm of

$$2\mathcal{L}\bar{G} = 2\mathcal{L}(-\partial_{tt} - 2\mathcal{L})^{-1} \text{ and } \partial_{tt}\bar{G} = \partial_{tt}(-\partial_{tt} - 2\mathcal{L})^{-1}$$

is at most one, as can be seen by considering a common spectral decomposition. In particular, one can take $c_1 = 1$. Furthermore $\|\bar{G}\| \leq \max(1/m, T^2/\pi^2)$, so one also has

$$\begin{aligned} \sqrt{2\mathcal{E}_T(g)} &= \sqrt{\langle g, -2\mathcal{L}g \rangle_{L^2(\lambda \otimes \mu)}} \\ &= \sqrt{\langle \bar{G}f, -2\mathcal{L}\bar{G}f \rangle_{L^2(\lambda \otimes \mu)}} \\ &\leq \sqrt{\max\left(\frac{1}{m}, \frac{T^2}{\pi^2}\right)} \|f\|_{L^2(\lambda \otimes \mu)} \end{aligned}$$

and, using the commutativity of ∂_t and \mathcal{L} and integrating by parts,

$$\begin{aligned} \sqrt{2\mathcal{E}_T(\partial_t g)} &= \sqrt{\langle \partial_{tt}g, -2\mathcal{L}g \rangle_{L^2(\lambda \otimes \mu)}} \\ &= \sqrt{\langle -\partial_{tt}\bar{G}f, -2\mathcal{L}\bar{G}f \rangle_{L^2(\lambda \otimes \mu)}} \\ &\leq \|f\|_{L^2(\lambda \otimes \mu)}. \end{aligned}$$

Finally

$$\sqrt{2\mathcal{E}_T(h)} \leq \|-2\mathcal{L}\bar{G}f\|_{L^2(\lambda \otimes \mu)}^{\frac{1}{2}} \|\partial_{tt}\bar{G}f\|_{L^2(\lambda \otimes \mu)}^{\frac{1}{2}} \leq \|f\|_{L^2(\lambda \otimes \mu)}.$$

Hence the c_i are given by

$$c_1 = 1, \quad c_2 = 1, \quad c_3 = \sqrt{\max\left(\frac{1}{m}, \frac{T^2}{\pi^2}\right)}, \quad c_4 = 1.$$

Case 2: $f \in \mathbb{H}_{l,a}$ Set $h(t, x) = \int_0^t f(s, x) ds$ and $g(t, x) = 0$. Due to the antisymmetry of f one has $h(0, \cdot) = 0 = h(T, \cdot)$. We obtain

$$\begin{aligned} \|h\|_{L^2(\lambda \otimes \mu)}^2 &= \int_0^T \left\| \int_0^t f(s, \cdot) ds \right\|_{L^2(\mu)}^2 dt \\ &\leq \int_0^T t \int_0^t \|f(s, \cdot)\|_{L^2(\mu)}^2 ds dt \leq \frac{T^2}{2} \|f\|_{L^2(\lambda \otimes \mu)}^2 \end{aligned}$$

and

$$\begin{aligned}\|\mathcal{L}h\|_{L^2(\lambda\otimes\mu)}^2 &\leq \int_0^T t \int_0^t \|\mathcal{L}f(s, \cdot)\|_{L^2(\lambda\otimes\mu)}^2 ds dt \\ &\leq \int_0^T t \int_0^t \frac{16}{T^4} \|f(s, \cdot)\|_{L^2(\lambda\otimes\mu)}^2 ds dt \leq \frac{8}{T^2} \|f\|_{L^2(\lambda\otimes\mu)}^2\end{aligned}$$

using the fact that $f \in \mathbb{H}_{l,a}$ for the second inequality. Thus

$$\sqrt{2\mathcal{E}_T(h)} \leq \sqrt{2\sqrt{\frac{T^2}{2}} \|f\|_{L^2(\lambda\otimes\mu)} \times \sqrt{\frac{8}{T^2}} \|f\|_{L^2(\lambda\otimes\mu)}} = 2\|f\|_{L^2(\lambda\otimes\mu)}$$

so overall

$$c_1 = 0, \quad c_2 = 2, \quad c_3 = 0, \quad c_4 = 0.$$

Case 3: $f \in \mathbb{H}_{l,s}$ Since f is symmetric, one may decompose it into a part f_0 that integrates to 0 in time and a part f_1 with Dirichlet boundary values

$$\begin{aligned}f_0(t, x) &= f(0, x) \cos\left(\frac{2\pi t}{T}\right) \\ f_1(t, x) &= f(t, x) - f_0(t, x).\end{aligned}$$

Note that $f(t, \cdot), f_0(t, \cdot), f_1(t, \cdot) \in \text{Span}(e_k : \alpha_k \leq \frac{2}{T})$. Let $G = -\frac{1}{2}\mathcal{L}^{-1}$ on $\text{Span}(e_k : \alpha_k \leq \frac{2}{T})$ i.e. $Ge_k = \frac{1}{2\alpha_k^2}e_k$. If one sets $h(t, x) = \int_0^t f_0(s, x)ds$ and $g(x) = Gf_1(t, \cdot)(x)$ then $f_0 = \partial_t h$ and $f_1 = -2\mathcal{L}g$.

To obtain estimates one needs to control $\|f_0\|_{L^2(\lambda\otimes\mu)}$ by $\|f\|_{L^2(\lambda\otimes\mu)}$. We have

$$\|f_0\|_{L^2(\lambda\otimes\mu)}^2 = \|f(0, \cdot)\|_{L^2(\mu)}^2 \int_0^T \cos\left(\frac{2\pi t}{T}\right)^2 dt = \frac{T}{2} \|f(0, \cdot)\|_{L^2(\mu)}^2. \quad (7.3)$$

Furthermore

$$f(t, x) = \sum_{\alpha_k \leq \frac{2}{T}} b_k h_k^s(t, x) = \sum_{\alpha_k \leq \frac{2}{T}} b_k \left(e^{-\alpha_k t} + e^{-\alpha_k(T-t)} \right) e_k(x), \quad (7.4)$$

where the decomposition is orthogonal in $L^2(\mu)$ for every t and

$$\|f(0, \cdot)\|_{L^2(\mu)}^2 = \sum_{\alpha_k \leq \frac{2}{T}} b_k^2 (1 + e^{-\alpha_k T})^2 \leq 4 \sum_{\alpha_k \leq \frac{2}{T}} b_k^2. \quad (7.5)$$

If $\alpha_k \leq \frac{2}{T}$ then

$$\begin{aligned}\int_0^T \left(e^{-\alpha_k t} + e^{-\alpha_k(T-t)} \right)^2 dt &= T \frac{2e^{-\alpha_k T} (\alpha_k T + \sinh(\alpha_k T))}{\alpha_k T} \\ &\geq 2Te^{-\alpha_k T} \geq \frac{2T}{e^2}.\end{aligned}$$

So that

$$\|f\|_{L^2(\lambda\otimes\mu)}^2 = \sum_{\alpha_k \leq \frac{2}{T}} b_k^2 \int_0^T \left(e^{-\alpha_k t} + e^{-\alpha_k(T-t)} \right)^2 dt \geq \frac{2T}{e^2} \sum_{\alpha_k \leq \frac{2}{T}} b_k^2. \quad (7.6)$$

Combining (7.5) and (7.6) yields $\|f(0, \cdot)\|_{L^2(\mu)}^2 \leq \frac{2}{T} e^2 \|f\|_{L^2(\lambda\otimes\mu)}^2$ so that

$$\|f_0\|_{L^2(\lambda\otimes\mu)}^2 \leq \frac{T}{2} \|f(0, \cdot)\|_{L^2(\mu)}^2 \leq e^2 \|f\|_{L^2(\lambda\otimes\mu)}^2. \quad (7.7)$$

This yields the bounds

$$\begin{aligned}\|h\|_{L^2(\lambda \otimes \mu)}^2 &\leq \frac{T^2}{2} \|f_0\|_{L^2(\lambda \otimes \mu)}^2 \leq \frac{T^2 e^2}{2} \|f\|_{L^2(\lambda \otimes \mu)}^2, \\ \|\mathcal{L}h\|_{L^2(\lambda \otimes \mu)}^2 &\leq \frac{8}{T^2} \|f_0\|_{L^2(\lambda \otimes \mu)}^2 \leq \frac{8e^2}{T^2} \|f\|_{L^2(\lambda \otimes \mu)}^2,\end{aligned}$$

using the same estimates as in step 3. Therefore one also has

$$\sqrt{2\mathcal{E}_T(h)} \leq \sqrt{2\sqrt{\frac{T^2 e^2}{2}} \|f\|_{L^2(\lambda \otimes \mu)} \sqrt{\frac{8e^2}{T^2}} \|f\|_{L^2(\lambda \otimes \mu)}} = 2e \|f\|_{L^2(\lambda \otimes \mu)}$$

and one can take $c_2 = 2e$.

One must now also bound the terms involving $g = Gf_1$. First notice

$$\|f_1\|_{L^2(\lambda \otimes \mu)} \leq \|f\|_{L^2(\lambda \otimes \mu)} + \|f_0\|_{L^2(\lambda \otimes \mu)} \leq (1+e) \|f\|_{L^2(\lambda \otimes \mu)}.$$

Clearly

$$\|2\mathcal{L}g\|_{L^2(\lambda \otimes \mu)} = \|2\mathcal{L}Gf_1\|_{L^2(\lambda \otimes \mu)} = \|f_1\|_{L^2(\lambda \otimes \mu)} \leq (1+e) \|f\|_{L^2(\lambda \otimes \mu)},$$

so one can take $c_1 = 1+e$. Furthermore

$$\|g\|_{L^2(\lambda \otimes \mu)} = \|Gf_1\|_{L^2(\lambda \otimes \mu)} \leq \frac{1}{2m} \|f_1\|_{L^2(\lambda \otimes \mu)} \leq \frac{1}{2m} (1+e) \|f\|_{L^2(\lambda \otimes \mu)},$$

so

$$\sqrt{2\mathcal{E}_T(g)} = \sqrt{\langle g, -2\mathcal{L}g \rangle_{L^2(\lambda \otimes \mu)}} \leq \frac{1+e}{\sqrt{2m}} \|f\|_{L^2(\lambda \otimes \mu)}$$

and one can take $c_3 = \frac{1+e}{\sqrt{2m}}$.

Because ∂_t commutes with \mathcal{L} one has

$$\sqrt{2\mathcal{E}_T(\partial_t g)} = \sqrt{\langle \partial_t g, 2\mathcal{L}g \rangle_{L^2(\lambda \otimes \mu)}} \leq \sqrt{\|\partial_t g\|_{L^2(\lambda \otimes \mu)} \|2\mathcal{L}g\|_{L^2(\lambda \otimes \mu)}}$$

so it remains to bound $\|\partial_{tt}g\|_{L^2(\lambda \otimes \mu)}$. Using the fact that ∂_t commutes with G yields

$$\|\partial_{tt}g\|_{L^2(\lambda \otimes \mu)} = \|G\partial_{tt}f_1\|_{L^2(\lambda \otimes \mu)} \leq \frac{1}{2m} \|\partial_{tt}f_1\|_{L^2(\lambda \otimes \mu)}.$$

so that it suffices to estimate $\|\partial_{tt}f_1\|_{L^2(\lambda \otimes \mu)}$.

Note that $\|\partial_{tt}f_1\|_{L^2(\lambda \otimes \mu)} \leq \|\partial_{tt}f\|_{L^2(\lambda \otimes \mu)} + \|\partial_{tt}f_0\|_{L^2(\lambda \otimes \mu)}$. Use now $\partial_{tt}f_0 = -\left(\frac{2\pi}{T}\right)^2 f_0$ to get

$$\|\partial_{tt}f_0\|_{L^2(\lambda \otimes \mu)} = \left(\frac{2\pi}{T}\right)^2 \|f_0\|_{L^2(\lambda \otimes \mu)} \leq e \left(\frac{2\pi}{T}\right)^2 \|f\|_{L^2(\lambda \otimes \mu)}.$$

Furthermore, (7.4) yields

$$\partial_{tt}f = \sum b_k \alpha_k^2 \left(e^{-\alpha_k t} + e^{-\alpha_k(T-t)} \right) e_k(x)$$

so that

$$\|\partial_{tt}f\|_{L^2(\lambda \otimes \mu)}^2 = \sum_{\alpha_k \leq \frac{2}{T}} b_k^2 \alpha_k^4 \int_0^T \left(e^{-\alpha_k t} + e^{-\alpha_k(T-t)} \right)^2 dt \leq \left(\frac{2}{T}\right)^4 \|f\|_{L^2(\lambda \otimes \mu)}^2.$$

Overall

$$\begin{aligned}\|\partial_{tt}f_1\|_{L^2(\lambda \otimes \mu)} &\leq \frac{4(\pi^2 e + 1)}{T^2} \|f\|_{L^2(\lambda \otimes \mu)} \\ \|\partial_{tt}g\|_{L^2(\lambda \otimes \mu)} &\leq \frac{1}{m} \frac{2(\pi^2 e + 1)}{T^2} \|f\|_{L^2(\lambda \otimes \mu)}.\end{aligned}$$

Finally, by construction

$$h|_{t=0} = h|_{t=T} = 0 \text{ and } g|_{t=0} = g|_{t=T} = 0$$

so that the boundary conditions are satisfied.

And the c_i are

$$c_1 = 1 + e, \quad c_2 = 2e, \quad c_3 = \frac{1 + e}{\sqrt{2m}}, \quad c_4 = \sqrt{\frac{2(\pi^2 e + 1)(1 + e)}{mT^2}}.$$

Case 4: $f \in \mathbb{H}_{h,a}$ One may use the expansion $f(t, x) = \sum_{\alpha_k > 2/T} b_k h_k^a(t, x)$ and write $u_k(t) = \frac{e^{-\alpha_k t} - e^{-\alpha_k(T-t)}}{e^{-\alpha_k t} - e^{-\alpha_k(T-t)}}$ so that $h_k^a(t, x) = u_k(t)e_k(x)$. We employ the ansatz

$$u_k = \dot{v}_k + w_k \text{ with } v_k(0) = v_k(T) = w_k(0) = w_k(T) = 0.$$

Then setting $h_k = v_k e_k$ and $g_k = \frac{1}{2\alpha_k^2} w_k e_k$ yields $h_k^a = \partial_t h_k - 2\mathcal{L}g_k$. Note that

$$\begin{aligned} \mathcal{E}_T \left(\sum_{\alpha_k > 2/T} b_k h_k \right) &= \left\langle \sum_{\alpha_k > 2/T} b_k h_k, \mathcal{L} \sum_{\alpha_k > 2/T} b_k h_k \right\rangle_{L^2(\lambda \otimes \mu)} \\ &= \sum_{k,l} b_k b_l \alpha_l \langle h_k, h_l \rangle_{L^2(\lambda \otimes \mu)} \\ &= \sum_k \alpha_k b_k^2 \|h_k\|_{L^2(\lambda \otimes \mu)}^2 \\ &= \sum_{\alpha_k > 2/T} b_k^2 \mathcal{E}_T(h_k) \end{aligned}$$

because $\langle h_k, h_l \rangle_{L^2(\lambda \otimes \mu)} = 0$ if $k \neq l$ and hence it is enough to show the bound

$$\sqrt{2\mathcal{E}_T(h_k)} \leq c_2 \|h_k^a\|_{L^2(\lambda \otimes \mu)} \text{ for all } k \text{ with } \alpha_k > 2/T$$

to show $\sqrt{2\mathcal{E}(h)} \leq c_2 \|f\|_{L^2(\lambda \otimes \mu)}$ for all $f \in \mathbb{H}_{h,a}$. The same is true for the other desired inequalities. Also note that the boundary conditions

$$h(t=0, \cdot) = h(t=T, \cdot) = g(t=0, \cdot) = g(t=T, \cdot) = 0$$

are satisfied by construction.

One may now construct v_k and w_k . For $t \in [0, T/2]$ set

$$\begin{aligned} v_k(t) &= \varphi_k(t) \int_0^t u_k(s) ds, \\ w_k(t) &= u_k(t) - \dot{v}_k(t) = (1 - \varphi_k(t))u_k(t) - \dot{\varphi} \int_0^t u_k(s) ds, \end{aligned}$$

where

$$\varphi_k(t) = (\alpha_k t - 1)^2 \mathbf{1}_{\{0 \leq t \leq 1/\alpha_k\}}$$

with $|\varphi_k| \leq 1$, $|\dot{\varphi}_k| \leq 2\alpha_k$ and $|\ddot{\varphi}_k| \leq 2\alpha_k^2$. Here the second derivative should be understood in the H^2 sense. Analogously, for $t \in [T/2, T]$ one may set

$$\begin{aligned} v_k(t) &= -\varphi_k(T-t) \int_t^T u_k(s) ds, \\ w_k(t) &= u_k(t) - \dot{v}_k(t) = (1 - \varphi_k(T-t))u_k(t) - \dot{\varphi}_k(T-t) \int_t^T u_k(s) ds. \end{aligned}$$

Then $v_k, w_k \in H^1[0, T]$ with $v_k(0) = v_k(T) = 0$, $v_k(T/2) = \dot{v}_k(T/2) = 0$, and

$$\begin{aligned} \dot{v}_k(0) &= \varphi_k(0)u_k(0) = u_k(0), & \dot{v}_k(T) &= \varphi_k(0)u_k(T) = u_k(T), \\ w_k(0) &= u_k(0) - \dot{v}_k(0) = 0, & w_k(T) &= u_k(T) - \dot{v}_k(T) = 0, \end{aligned}$$

so that the boundary conditions of v_k and w_k are satisfied.

The first step is to estimate

$$\int_0^T v_k(t)^2 dt, \quad \int_0^T w_k(t)^2 dt \quad \text{and} \quad \int_0^T [\dot{w}_k(t)]^2 dt.$$

We have

$$\begin{aligned} \int_0^T v_k(t)^2 dt &= \int_0^{1/\alpha_k} \varphi_k(t)^2 \left(\int_0^t u_k(s) ds \right)^2 dt \\ &\leq \int_0^{1/\alpha_k} (\alpha_k t - 1)^4 t \int_0^{T/2} u_k(s)^2 ds dt \\ &= \frac{1}{30\alpha_k^2} \int_0^{T/2} u_k(t)^2 dt. \end{aligned}$$

For $w_k = (1 - \varphi_k)u_k - \dot{\varphi}_k \int_0^t u_k(s) ds$ one gets

$$\begin{aligned} \int_0^{T/2} w_k(t)^2 dt &\leq \left[\left(\int_0^{1/\alpha_k} \dot{\varphi}_k(t)^2 \left(\int_0^t u_k(s) ds \right)^2 dt \right)^{1/2} \right. \\ &\quad \left. + \left(\int_0^{T/2} (1 - \varphi_k(s))u_k(t)^2 dt \right)^{1/2} \right]^2 \\ &\leq \left[1 + \left(\int_0^{1/\alpha_k} (2\alpha_k(\alpha_k t - 1))^2 t \right)^{1/2} \right]^2 \int_0^{T/2} u_k(t)^2 dt \\ &= \left(1 + \frac{1}{\sqrt{3}} \right)^2 \int_0^{T/2} u_k(t)^2 dt. \end{aligned}$$

Since $\dot{w}_k = (1 - \varphi_k)\dot{u}_k - 2\dot{\varphi}_k u_k - \ddot{\varphi}_k \int_0^t u_k(s) ds$ one needs to estimate $\int_0^T \dot{u}_k^2 dt$ by $\int_0^T u_k^2 dt$. Both can explicitly be computed

$$\int_0^{T/2} \dot{u}_k^2(t) dt = \alpha_k^2 \int_0^{T/2} \left(e^{-\alpha_k t} + e^{-\alpha_k(T-t)} \right)^2 dt = \alpha_k e^{-\alpha_k T} (\sinh(\alpha_k T) + \alpha_k T)$$

and

$$\int_0^{T/2} u_k(t)^2 dt = \alpha_k^{-1} e^{-\alpha_k T} (\sinh(\alpha_k T) - \alpha_k T)$$

so that

$$\int_0^T \dot{u}_k^2 dt \leq \alpha_k^2 \frac{\sinh(2) + 2}{\sinh(2) - 2}$$

using the fact that $x \mapsto (\sinh(x) + x)/(\sinh(x) - x)$ is a decreasing function. Similar estimations as for w_k then yield

$$\int_0^{T/2} \dot{w}_k^2 dt \leq \alpha_k^2 \left(\frac{\sinh(2) + 2}{\sinh(2) - 2} + 4 + \sqrt{2} \right)^2 \int_0^{T/2} u_k^2 dt.$$

Combining these results with similar computations on $[T/2, T]$ yields

$$\begin{aligned} \int_0^T v_k(t)^2 dt &\leq \frac{1}{30\alpha_k^2} \int_0^T u_k(t)^2 dt, \\ \int_0^T w_k(t)^2 dt &\leq \left(1 + \frac{1}{\sqrt{3}}\right)^2 \int_0^T u_k^2(t) dt, \\ \int_0^T [\dot{w}_k(t)]^2 dt &\leq \alpha_k^2 \left(4 + \sqrt{2} + \sqrt{\frac{\sinh(2) + 2}{\sinh(2) - 2}}\right)^2 \int_0^T u_k(t)^2 dt. \end{aligned}$$

Hence

$$\|2\mathcal{L}g_k\|_{L^2(\lambda \otimes \mu)} = \|w_k\|_{L^2(\lambda \otimes \mu)} \leq \left(1 + \frac{1}{\sqrt{3}}\right) \|h_k^a\|_{L^2(\lambda \otimes \mu)},$$

and

$$\begin{aligned} \mathcal{E}_T(h_k) &= \langle h_k, -\mathcal{L}h_k \rangle_{L^2(\lambda \otimes \mu)} \\ &= \alpha_k^2 \int_0^T v_k^2(t) dt \\ &\leq \frac{1}{30} \int_0^T u_k^2(t) dt \\ &= \frac{1}{30} \|h_k^a\|_{L^2(\lambda \otimes \mu)}^2. \end{aligned}$$

We also have

$$\mathcal{E}_T(g_k) = 1/(2\alpha_k^2) \int_0^T w_k^2(t) dt \leq 1/(2\alpha_k^2) \left(1 + \frac{1}{\sqrt{3}}\right)^2 \|h_k^a\|_{L^2(\lambda \otimes \mu)}^2,$$

and finally

$$\begin{aligned} \mathcal{E}_T(\partial_t g_k) &= 1/(2\alpha_k^2) \int_0^T \dot{w}_k^2(t) dt \\ &\leq \frac{1}{2} \left(4 + \sqrt{2} + \sqrt{\frac{\sinh(2) + 2}{\sinh(2) - 2}}\right)^2 \|h_k^a\|_{L^2(\lambda \otimes \mu)}^2. \end{aligned}$$

One can thus choose

$$\begin{aligned} c_1 &= 1 + \frac{1}{\sqrt{3}}, & c_2 &= \frac{1}{\sqrt{15}}, \\ c_3 &= \frac{1}{2} \left(1 + \frac{1}{\sqrt{3}}\right) T, & c_3 &= 4 + \sqrt{2} + \sqrt{\frac{\sinh(2) + 2}{\sinh(2) - 2}}. \end{aligned}$$

on $\mathbb{H}_{h,a}$.

Case 5: $f \in \mathbb{H}_{h,s}$ Let $f \in \mathbb{H}_{h,s}$. As in the previous case, one may use the expansion $f(t, x) = \sum_{\alpha_k > 2/T} b_k h_k^s(t, x)$ and now write $u_k(t) = e^{-\alpha_k t} + e^{-\alpha_k(T-t)}$, so that $h_k^s(t, x) = u_k(t)e_k(x)$. We again employ the ansatz

$$u_k = \dot{v}_k + w_k \text{ with } v_k(0) = v_k(T) = w_k(0) = w_k(T) = 0,$$

so that setting $h_k = v_k e_k$ and $g_k = \frac{1}{\alpha_k} w_k e_k$ yields $h_k^s = \partial_t h_k - \mathcal{L}g_k$. Finally, setting $h = \sum_{\alpha_k > 2/T} b_k h_k$ and $g = \sum_{\alpha_k > 2/T} b_k g_k$ yields $f = \partial_t h - \mathcal{L}g$ with the same bounds by orthogonality, as well as the boundary conditions

$$h(t=0, \cdot) = h(t=T, \cdot) = g(t=0, \cdot) = g(t=T, \cdot) = 0.$$

The only thing that changes compared to the previous case is the bound on $\int_0^T \dot{u}_k^2(t) dt$. In this case, one has

$$\begin{aligned} \int_0^T \dot{u}_k^2(t) dt &= \alpha_k \int_0^T \left(-e^{-\alpha_k t} + e^{-\alpha_k(T-t)} \right)^2 dt \\ &= \alpha_k^2 \frac{e^{-\alpha_k T} (\sinh(\alpha_k T) - \alpha_k T)}{\alpha_k} \end{aligned}$$

whereas

$$\int_0^{T/2} u_k^2(t) dt = \frac{e^{-\alpha_k T} (\sinh(\alpha_k T) + \alpha_k T)}{\alpha_k}.$$

Hence

$$\int_0^T \dot{u}_k^2(t) dt = \alpha_k^2 \frac{\sinh(\alpha_k T) - \alpha_k T}{\sinh(\alpha_k T) + \alpha_k T} \int_0^T u_k^2(t) dt \leq \alpha_k^2 \int_0^T u_k^2(t)^2 dt.$$

□

7.4 Applications

7.4.1 Underdamped Langevin equation

As a warm up, the potential U is supposed smooth for simplicity, and it is needed to check our assumptions (A), (B) and (C). Recall from the example of Section 2:

$$\begin{aligned} \hat{\mathcal{L}}f(x, v) &= v \cdot \nabla_x f - \nabla_x U(x) \cdot \nabla_v f - \gamma v \cdot \nabla_v f + \gamma \Delta_v f \\ &= \hat{\mathcal{L}}_d f + \gamma \hat{\mathcal{L}}_v f \end{aligned}$$

Checking Assumption (A) The fact that $\hat{\mathcal{L}}$ is a lift of the overdamped Langevin process has already been checked. So, by Remark 7.3, the antisymmetric generator $\hat{\mathcal{L}}_d$ is as well. Here, $\kappa_x = \kappa$ is the Gaussian measure, which is invariant for the Ornstein–Uhlenbeck process and which satisfies the Poincaré inequality with constant 1, see for example [BGL14].

Checking Assumption (B) For μ related to the generator \mathcal{L} , one can use Hardy’s type condition inequality in dimension one, curvature condition or Lyapunov conditions for example, for which a well known sufficient condition is derived: there exists $a \in (0, 1)$ such that $a|\nabla U|^2 - \Delta U > 0$ outside of a large ball see [BBCG08, BGL14]. However to ensure that the spectrum is discrete (or that the operator is compact), one can impose the sufficient condition (see [BGW15, Theorem 5]): there exists $\beta > 1$

$$\lim_{|x| \rightarrow \infty} \frac{\nabla U(x) \cdot x}{|x|^\beta} = c > 0.$$

Suppose moreover for simplicity that U is convex.

Checking Assumption (C1) By Remark 7.3 it is enough to prove

$$\|\hat{\mathcal{L}}_d^* \hat{\mathcal{L}}_d(g \circ \pi) + (2\mathcal{L}g) \circ \pi\|_{L^2(\hat{\mu})} \leq c_0^* \|2\mathcal{L}g\|_{L^2(\mu)}.$$

We have

$$\begin{aligned} \hat{\mathcal{L}}_d^* \hat{\mathcal{L}}_d(g \circ \pi) &= -v^t (\text{Hess}_x g) v - \nabla_x U \cdot \nabla_x g \\ 2\mathcal{L}g &= \Delta_x g + \nabla_x U \cdot \nabla_x g \end{aligned}$$

so that

$$\begin{aligned} & \int \left([\hat{\mathcal{L}}_d^* \hat{\mathcal{L}}_d(g \circ \pi)](x, v) + 2\mathcal{L}g(x) \right)^2 d\kappa_x(v) \\ &= \int (v^t(\text{Hess}_x g)v - \Delta_x g)^2 d\kappa_x(v), \\ &= \int (v^t(\text{Hess}_x g)v)^2 d\kappa_x(v) - (\Delta_x g)^2, \end{aligned}$$

using the fact $\int v^t(\text{Hess}_x g)v d\kappa_x(v) = \Delta_x g$.

Furthermore

$$\begin{aligned} \int (v^t(\text{Hess}_x g)v)^2 d\kappa_x(v) &= \sum_{i,j,k,l=1}^d (\partial_{ij}g)(\partial_{kl}g) \int v_i v_j v_k v_l d\kappa_x(v) \\ &= \sum_{i,k=1}^d (\partial_i^2 g)(\partial_k^2 g) \int v_i^2 v_k^2 d\kappa_x(v) \\ &\quad + 2 \sum_{\substack{i,j=1 \\ i \neq j}}^d (\partial_{ij}g)^2 \int v_i^2 v_j^2 d\kappa_x(v) \\ &= (\Delta_x g)^2 + 2 \sum_{i,j=1}^d (\partial_{ij}g)^2 \end{aligned}$$

using the fact that v_i, v_j, v_k, v_l are independent standard Gaussians.

Recall that, by Bochner's formula, we have

$$\sum_{ij} \|\partial_{ij}g\|_{L^2(\mu)}^2 = \|2\mathcal{L}\|_{L^2(\mu)}^2 - \int (\nabla_x g)^t (\text{Hess}_x U) \nabla_x g d\mu$$

and hence $\sum_{ij} \|\partial_{ij}g\|_{L^2(\mu)}^2 \leq \|2\mathcal{L}\|_{L^2(\mu)}^2$ under the assumption that the potential U is convex.

Overall we get

$$\|\hat{\mathcal{L}}_d^* \hat{\mathcal{L}}_d(g \circ \pi) + (2\mathcal{L}g) \circ \pi\|_{L^2(\hat{\mu})}^2 = 2 \sum_{ij} \|\partial_{ij}g\|_{L^2(\mu)}^2 \leq 2\|2\mathcal{L}g\|_{L^2(\mu)}^2,$$

and (C1) holds with $c_0^* = \sqrt{2}$.

Remark 7.11. *The assumption of convexity may easily be relaxed to $\text{Hess}_x U \geq -K \text{Id}$ for some positive K to get*

$$\sum_{ij} \|\partial_{ij}g\|_{L^2(\mu)}^2 \leq \|2\mathcal{L}\|_{L^2(\mu)}^2 + K\mathcal{E}(g).$$

As this last term may also be controlled successfully via the divergence lemma, straightforward modifications of the proof lead to a new version of the Flow Poincaré inequalities.

Checking Assumption (C2) Here we have, $\hat{\mu}$ being a product measure

$$\left\langle \hat{\mathcal{L}}_v f, \hat{\mathcal{L}}_d(g \circ \pi) \right\rangle_{L^2(\hat{\mu})} = \langle \nabla_v^* \nabla_v f, v \cdot \nabla_x g \rangle_{L^2(\hat{\mu})} = \langle \nabla_v f, \nabla_x g \rangle_{L^2(\hat{\mu})}$$

and then apply Cauchy-Schwarz to get the desired inequality with $c_1^* = 1$.

Conclusion: Convergence rate for the underdamped Langevin equation We recover here the result of [CLW23, Theorem 1] in the convex case.

Corollary 7.12. *Let us suppose that U is convex, that m is the Poincaré constant associated to $d\mu = e^{-U}$, and that there exists $\beta > 1$ such that*

$$\lim_{|x| \rightarrow \infty} \frac{\nabla U(x) \cdot x}{|x|^\beta} = c > 0.$$

Then choosing $T = m^{-1/2}$, the semigroup \hat{P}_t associated to \hat{L} is exponentially contractive in T -average with constant

$$O\left(\frac{\gamma m}{\gamma^2 + m}\right).$$

This is the rate obtained by [CLW23]. By straightforward adaptations of (C1) and (C2) one may get the other statements of [CLW23, Theorem 1].

7.4.2 PDMP Sampling algorithm

Piecewise Deterministic Processes (PDMP) have been recently studied as irreversible, and particularly efficient, sampler in physics and statistics fields, for which proving exponential ergodicity is of crucial importance. Let us cite for example, without possible exhaustivity, [MKK14] for the Event Chain Monte Carlo, [BRZ19] for the Zig Zag sampler, [MDS20] for the Forward process, [DGM20] for the Bouncy particle sampler. However those results were hardly quantitative. Very recently, using the Dolbeault-Mouhot-Schmeiser approach, dimensionally interesting results were established in [ADNR21], completed using Armstrong-Mourrat's method in [LW22] for the Zig-Zag process and the Bouncy Particle Sampler. We aim here at completing these results by our method.

Forward process

Let us consider a differentiable potential U associated to the measure $d\mu = e^{-U} dx$. For simplicity, let us suppose that U is convex, that \mathcal{L} has a purely discrete spectrum (condition as in the underdamped Langevin process), and that $\forall i, j, |\partial_{ij} U| \leq L$ and let $\kappa(dv)$ be the Gaussian measure. The forward process, following [MDS20], may be defined by the generator

$$\hat{\mathcal{L}}f(x, v) = \underbrace{v \cdot \nabla_x f + (v \cdot \nabla U)_+ (B_- f - f)}_{=: \hat{\mathcal{L}}_d} + \underbrace{\gamma(\Pi_v f - f)}_{=: \gamma \hat{\mathcal{L}}_v}$$

where

$$B_\pm f = \frac{1}{Z} \int (\omega \cdot \nabla U)_\pm f(x, \omega) \kappa(d\omega).$$

As for the underdamped Langevin process, the invariant measure for the forward process is $\hat{\mu} = \mu \otimes \kappa$. Due to the structure of the process, quite similar to the underdamped Langevin process, assumptions (A) and (B) may be verified quite similarly leading to the collapse process being the overdamped Langevin process.

One may thus focus on Assumption (C).

Checking Assumption (C1) Remark first that

$$\hat{\mathcal{L}}_d(g \circ \pi) = v \cdot \nabla_x g, \quad \hat{\mathcal{L}}_d^* = -v \cdot \nabla_x + (v \cdot \nabla U)_- (B_+ - I).$$

Thus

$$\hat{\mathcal{L}}_d^*(v \cdot \nabla_x g) = -v^t \text{Hess } g \cdot v + \frac{1}{Z} (v \cdot \nabla U)_- \left(\frac{1}{2} \nabla U \cdot \nabla g - Zv \cdot \nabla g \right).$$

We may then proceed as follow to get (C1)

$$\begin{aligned} \left| \left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle \right| &= \left| \left\langle \hat{\mathcal{L}}_d^*(v \cdot \nabla_x g), f - \Pi_v f \right\rangle \right| \\ &\leq \| \hat{\mathcal{L}}_d^*(v \cdot \nabla_x g) \|_{L^2(\hat{\mu})} \| f - \Pi_v f \|_{L^2(\hat{\mu})}. \end{aligned}$$

One thus has to control $\| \hat{\mathcal{L}}_d^*(v \cdot \nabla_x g) \|_{L^2(\hat{\mu})}$ which can be explicitly controlled (c being an absolute positive constant that may change from line to line)

Starting with the Cauchy-Schwarz inequality,

$$\begin{aligned} \int (\hat{\mathcal{L}}_d^*(v \cdot \nabla_x g))^2 \kappa(dv) &\leq 2 \int (v \cdot \nabla^2 g v)^2 \kappa(dv) \\ &\quad + \frac{2}{Z^2} \int (v \cdot \nabla U)_-^2 \left(\frac{1}{2} \nabla U \cdot \nabla g - Z v \cdot \nabla g \right)^2 \kappa(dv) \\ &= 4 \| \nabla^2 g \|^2 + 2 (\Delta g)^2 \\ &\quad + 2 \left(3 + \frac{\pi}{4} \right) (\nabla U \cdot \nabla g)^2 + | \nabla U |^2 | \nabla g |^2. \end{aligned}$$

First remark, $\nabla U \cdot \nabla g = -2\mathcal{L}g + \Delta g$, so that

$$(\nabla U \cdot \nabla g)^2 \leq 2(2\mathcal{L}g)^2 + 2(\Delta g)^2.$$

It remains to control the term $| \nabla U |^2 | \nabla g |^2$. By integration by parts,

$$\begin{aligned} \int | \nabla U |^2 | \nabla g |^2 d\mu &= \sum_{i,j=1}^d \int \partial_i U \partial_i U \partial_j g \partial_j g d\mu \\ &= \sum_{i,j=1}^d \int \partial_{ii} U (\partial_j g)^2 d\mu + 2 \sum_{i,j=1}^d \int \partial_i U \partial_{ij} g \partial_j g d\mu \\ &= \int \Delta U | \nabla g |^2 d\mu + 2 \int \nabla(\nabla U \cdot \nabla g) \cdot \nabla g d\mu \\ &\quad - 2 \sum_{i,j=1}^d \int \partial_{ij} U \partial_i g \partial_j g d\mu \\ &\leq cLd \int | \nabla g |^2 d\mu + 8 \int (\nabla U \cdot \nabla g)^2 d\mu + 8 \int (Lg)^2 d\mu. \end{aligned}$$

One may then use Bochner's inequality (as in the underdamped Langevin case), with U convex to get

$$\int \hat{\mathcal{L}}_d^*(v \cdot \nabla_x g) d\hat{\mu} \leq cd \left(\int (\mathcal{L}g)^2 d\mu + \int | \nabla g |^2 d\mu \right).$$

Finally, there exists $c > 0$ such that

$$\left| \left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle \right| \leq c\sqrt{d} \left(\| \mathcal{L}g \|_{L^2(\hat{\mu})} + \sqrt{\mathcal{E}(g)} \right) \| f - \Pi_v f \|_{L^2(\hat{\mu})},$$

and it is a slightly different version of (C1) which is established, i.e. (C*1). However, using divergence lemma 7.9, we may control this extra term to get

$$\begin{aligned} C_0 &= 2(c_1^* c_3)^2, \\ C_1 &= c_2^* (1 + 2(c_4 + c_2 + \tilde{c}_0^*(c_1 + c_3))^2), \end{aligned}$$

where $c_0^* = c\sqrt{d}$. Note that one may also easily consider the case where $\text{Hess } U \geq -KId$.

Checking Assumption (C2) Assumption (C2) is readily verified through Remark 7.5.

Conclusion: Convergence rate for the Forward process

Corollary 7.13. *Let us suppose that U is convex and that $|\partial_{i,j}U| \leq L$, that m is the Poincaré constant associated to $d\mu = e^{-U}$, and that there exists $\beta > 1$ such that*

$$\lim_{|x| \rightarrow \infty} \frac{\nabla U(x) \cdot x}{|x|^\beta} = c > 0.$$

Then choosing $T = m^{-1/2}$, the semigroup \hat{P}_t associated to \hat{L} is exponentially contractive in T -average with rate

$$\nu = O\left(\frac{\gamma m}{\gamma^2 + m + d}\right).$$

Here, the same dimension dependance as for the Bouncy Particle Sampler in [LW22] is recovered.

Zig-Zag process in coordinate direction

In [LW22], the convergence to equilibrium for the Zig-Zag process was considered with velocity refreshed as a Gaussian, and then at events flip of velocity in the considered direction, which is usually not the sampler considered in practice. One may thus focus on a coordinate change with velocity $\pm e_i$ (denoting (e_1, \dots, e_d) the base of \mathbb{R}^d). Here also, we suppose for simplicity

$$L \geq \text{Hess } U \geq 0.$$

Let us so consider the generator of the Zig-Zag process: $x \in \mathbb{R}^d$, $v \in \{\pm e_i\}_{i=1}^d$

$$\begin{aligned} \hat{L}f(x, v) &= v \cdot \nabla_x f + \sum_{i=1}^d (v_i \partial_i U)_+ (f(x, -v) - f(x, v)) \\ &\quad + \gamma \left(\int f(x, w) \kappa(dw) - f(x, v) \right) \\ &= \hat{L}_d f(x, v) + \gamma \hat{L}_v f(x, v) \end{aligned}$$

where κ is the uniform measure on $\{\pm e_i\}_{i=1}^d$. Since only one component of v is nonzero, \hat{L}_d can be written more concisely as

$$\hat{L}_d f(x, v) = v \cdot \nabla_x f + (v \cdot \nabla U)_+ (f(x, -v) - f(x, v))$$

Assumption (A) and (B) follows from the same arguments than for the underdamped Langevin equation and for the forward process. We thus impose the same conditions on U .

Checking Assumption (C1) We have

$$\hat{L}_d^* f(x, v) = -v \cdot \nabla_x f + (v \cdot \nabla U)_- (f(x, -v) - f(x, v))$$

and thus

$$\begin{aligned} \int (\hat{L}_d^* (v \cdot \nabla g))^2 d\kappa &= \frac{1}{d} \sum_{k=1}^d \left((\partial_{kk} g - \partial_k U \partial_k g)^2 + (\partial_k U)^2 (\partial_k g)^2 \right) \\ &\leq \frac{1}{d} \sum_{k=1}^d (2(\partial_{kk} g)^2 + 3(\partial_k U)^2 (\partial_k g)^2). \end{aligned}$$

An integration by parts and $2ab \leq \frac{1}{2}a^2 + 2b^2$ gives

$$\begin{aligned} \sum_{k=1}^d \int (\partial_k U)^2 (\partial_k g)^2 d\mu &= \sum_{k=1}^d \int (\partial_k g)^2 \partial_{kk} U d\mu + 2 \sum_{k=1}^d \int \partial_k U \partial_{kk} g \partial_k g d\mu \\ &\leq \sum_{k=1}^d \int (\partial_k g)^2 \partial_{kk} U d\mu \\ &\quad + \sum_{k=1}^d \int \left(\frac{1}{2} (\partial_k U)^2 (\partial_k g)^2 + 2 (\partial_{kk} g)^2 \right) d\mu, \end{aligned}$$

so that rearranging yields

$$\sum_{k=1}^d \int (\partial_k U)^2 (\partial_k g)^2 d\mu \leq 2 \sum_{k=1}^d \int \partial_{kk} U (\partial_k g)^2 d\mu + 4 \sum_{k=1}^d \int (\partial_{kk} g)^2 d\mu.$$

Then by $\nabla^2 U \leq L$,

$$\int (\hat{\mathcal{L}}_d^*(v \cdot \nabla g))^2 d\kappa \leq \frac{1}{d} \left(6L \mathcal{E}(g) + 14 \sum_{k=1}^d \int (\partial_{kk} g)^2 d\mu \right).$$

We finally get

$$\left| \langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \rangle \right| \leq cd^{-1/2} (\|\mathcal{L}g\|_{L^2(\hat{\mu})} + \sqrt{\mathcal{E}(g)}) \|f - \Pi_v f\|_{L^2(\hat{\mu})}.$$

Assumption (C2) is readily verified through Remark 7.5.

Conclusion: Convergence rate for the coordinate Zig-Zag process

Corollary 7.14. *Let us suppose that U is convex and that $\nabla^2 U \leq L$, that m is the Poincaré constant associated to $d\mu = e^{-U}$, and that there exists $\beta > 1$ such that*

$$\lim_{|x| \rightarrow \infty} \frac{\nabla U(x) \cdot x}{|x|^\beta} = c > 0.$$

Then choosing $T = m^{-1/2}$, the semigroup \hat{P}_t associated to \hat{L} is exponentially contractive in T -average with a rate independent of the dimension.

Remark that the same method, leading to the same result (up to constants), can be applied to the Zig-Zag process on $\mathbb{R}^d \times \{-1, 1\}^d$, considered in [BFR19].

7.4.3 Run-and-Tumble particles with sticky walls or jamming

First, consider a simpler process than the Run-and-Tumble particles with jamming. As seen in Section 2, the collapsed processes is not the usual overdamped Langevin process but rather sticky Brownian motion [Fel52, Fel54, EP14].

RTP with two velocities and sticky walls

One may continue here the application to a single one-dimensional run-and-tumble particle confined by hard walls. The first step in order to obtain exponential convergence is to split the generator into an antisymmetric lift and a velocity refreshment. By analogy with the Langevin case, it would be tempting to set $\hat{\mathcal{L}}_d$ the deterministic motion and $\hat{\mathcal{L}}_v$ the velocity jumps. However, for this splitting $\mathcal{L} = \hat{\mathcal{L}}_d + \gamma \hat{\mathcal{L}}_v$, the measure

$$\mu = \frac{1}{2 + \omega L} \delta_0 + \frac{\omega}{2 + \omega L} dx + \frac{1}{2 + \omega L} \delta_L$$

is neither invariant for $\hat{\mathcal{L}}_d$ nor for $\hat{\mathcal{L}}_v$. Instead, we take

$$\hat{\mathcal{L}}_v f(x, v) = \int (f(x, \tilde{v}) - f(x, v)) d\kappa_x(\tilde{v})$$

to be the full refresh with respect to the conditional measure of v . We then take $\hat{\mathcal{L}}_d = \hat{\mathcal{L}} - \gamma \hat{\mathcal{L}}_v$, where $\gamma > 0$ is the largest possible value such that $\hat{\mathcal{L}}_d$ is the generator of a stochastic process i.e. $\gamma = \omega$ and

$$\hat{\mathcal{L}}_d f(x, v) = \begin{cases} v \partial_x f(x, v) & \text{when } (x, v) \notin \{(2, L); (0, -2)\}, \\ \omega (f(L-, -2) - f(L, 2)) & \text{when } (x, v) = (2, L), \\ \omega (f(0+, 2) - f(0, -2)) & \text{when } (x, v) = (0, -2). \end{cases}$$

Checking assumption (A) The generator $\hat{\mathcal{L}}_d$ is a lift by Remark 7.3 and

$$f \in \text{Dom}(\mathcal{L}) \text{ implies } f \circ \pi \in \text{Dom}(\hat{\mathcal{L}}_d^*) \text{ and } \hat{\mathcal{L}}_d^*(f \circ \pi) = -\hat{\mathcal{L}}_d(f \circ \pi).$$

We have $\hat{\mathcal{L}}_v f = \Pi_v f - f$ so $\int \hat{\mathcal{L}}_v(f \circ \pi) d\kappa_x(v) = 0$ for all $x \in S$ and the dynamics in v has spectral gap $c_2^* = 1$.

Checking assumption (B) Let us first prove the Poincaré inequality of the collapsed process.

Lemma 7.15 (Poincaré inequality for sticky Brownian motion). *We have*

$$\|f - \mu(f)\|_{L^2(\mu)}^2 \leq \frac{1}{m_{\text{sti}}} \left[\frac{\omega}{2 + \omega L} \int_0^L 2(\partial_x f)^2 dx \right]$$

for all $f \in \text{Dom}(\mathcal{L}_{\text{sti}})$ where

$$\frac{1}{m_{\text{sti}}} = \frac{C}{\omega^2} ((\omega L)^2 + \omega L)$$

and C is an absolute constant.

Proof. Let us begin by a well known result on variance of mixtures. Denote $\mu_p = \int \Pi_\theta dp(\theta)$ where Π_θ are probability laws and p a probability (discrete for us but continuous would do). We have

$$\text{Var}_{\mu_p}(f) = \int \text{Var}_{\Pi_\theta} p(d\theta) + \text{Var}_p \left(\int f d\Pi_\theta \right).$$

Recall also that

$$\text{Var}_{\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1}(g) = \frac{1}{4}(g(0) - g(1))^2 = \frac{1}{4} \left(\int_0^L g'(x) dx \right)^2.$$

Rather than applying this to the triple mixture we have, let us do twice on mixture of two components by rewriting

$$\Pi = \frac{\omega L}{2 + \omega L} dU + \frac{2}{2 + \omega L} \left(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_L \right).$$

So that

$$\begin{aligned}\mathrm{Var}_{\Pi}(f) &= \frac{\omega L}{2 + \omega L} \mathrm{Var}_U(f) + \frac{2}{2 + \omega L} \mathrm{Var}_{\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1}(f) \\ &\quad + \frac{2\omega L}{(2 + \omega L)^2} \left(\frac{1}{L} \int_0^L f(x) dx - \frac{1}{2}f(0) - \frac{1}{2}f(L) \right)^2 \\ &= \frac{\omega L}{2 + \omega L} \mathrm{Var}_U(f) + \frac{1/2}{2 + \omega L} \left(\int_0^L f'(x) dx \right)^2 \\ &\quad + \frac{2\omega L}{(2 + \omega L)^2} \left(\int_0^L f'(x) \frac{x - L/2}{L} dx \right)^2.\end{aligned}$$

We then apply Poincaré inequality for the uniform measure and Cauchy-Schwartz for the other two to get

$$\mathrm{Var}_{\Pi}(f) \leq \left[\frac{L}{\pi^2} \frac{\omega L}{2 + \omega L} + \frac{L}{2(2 + \omega L)} + \frac{2\omega L}{(2 + \omega L)^2} \cdot \frac{L}{12} \right] \int_0^L f'^2(x) dx$$

using $\int_0^L \left(\frac{x-L/2}{L}\right)^2 dx = \frac{L}{12}$. Now $\frac{1}{m_{\mathrm{sti}}} = \frac{C}{\omega^2} ((\omega L)^2 + \omega L)$ follows from optimizing

$$\left[\frac{L}{\pi^2} \frac{\omega L}{2 + \omega L} + \frac{L}{2(2 + \omega L)} + \frac{2\omega L}{(2 + \omega L)^2} \cdot \frac{L}{12} \right] \frac{2 + \omega L}{\omega}$$

over ω and L . □

Discrete spectrum. Let us now prove that the collapsed process has a discrete spectrum. It is surely known but we did not find a reference. To do this we will prove a Nash inequality for the Dirichlet form associated to the sticky Brownian motion and invariant measure μ . Starting from [Bre11, Chapter 8, eq. (42)], a Nash inequality holds for any interval in dimension 1 for the Lebesgue measure, i.e. for some positive constant c

$$\left(\int_0^L f^2 dx \right)^{1/2} \leq c \left(\int_0^L |f| dx \right)^{2/3} \left(\int_0^L f^2 dx \right)^{1/2} + \left(\int_0^L (f')^2 dx \right)^{1/2} \right)^{1/3}.$$

By [Wan00] or [Bre11, Chapter 8, eq. (43)], this Nash inequality is equivalent to a super Poincaré inequality: for all $s > 0$

$$\int_0^L f^2 dx \leq s \int_0^L (f')^2 dx + \beta(s) \left(\int_0^L |f| dx \right)^2,$$

with $\beta(s) \sim s^{-3/8}$ going to infinity when s goes to 0 which captures Nash inequality (eventually retrieving worse constant), so that we may assume $\beta(s) > 1$. Now, let us prove that it implies the same type of super Poincaré inequality for the collapse Dirichlet form. Denoting $a_0 = \frac{1}{2 + \omega L}$ and $b_0 = \omega a_0$, for all $s > 0$ we have

$$\begin{aligned}\int f^2 d\mu &= a_0(f(0)^2 + f(L)^2) + b_0 \int_0^L f^2 dx, \\ &\leq a_0(f(0)^2 + f(L)^2) + b_0 s \int_0^L (f')^2 dx + b_0 \beta(s) \left(\int_0^L |f| dx \right)^2, \\ &\leq b_0 s \int_0^L (f')^2 dx + b_0 \max(b_0^{-2}, a_0^{-1}) \beta(s) \left(\int |f| d\mu \right)^2.\end{aligned}$$

Therefore the sticky Brownian satisfies a Nash inequality, thus is compact and has discrete spectrum, see for example [Cou96].

Checking assumption (C) We have that $u = f - \Pi_v f$ satisfies $\Pi_v u = 0$ so in particular $u(0, -2) = u(L, +2) = 0$ and

$$\frac{1}{2}u(x, +2) + \frac{1}{2}u(x, -2) = 0 \iff u(x, -2) = -u(x, +2).$$

It follows that

$$\begin{aligned} \langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \rangle &= \frac{\omega/2}{2 + \omega L} \int_0^L (2\partial_x g(x))(2\partial_x u(x, +2)) dx \\ &\quad + \frac{\omega/2}{2 + \omega L} \int_0^L (-2\partial_x g(x)) \underbrace{(-2\partial_x u(x, -2))}_{=-u(x, +2)} dx \\ &= 0 \end{aligned}$$

so the first inequality of assumption (C) holds with $c_0^* = 0$. The second inequality holds with $c_1^* = 1$ by Remark 7.5.

Conclusion: rate of convergence to equilibrium for the RPT with sticky walls

Taking $T = m^{-1/2}$ in Theorem 7.7 yields

$$\begin{aligned} C_0 &= 2(c_1^* c_3)^2 = C \frac{1}{m_{\text{sti}}} = \frac{C}{\omega^2} ((\omega L)^2 + \omega L), \\ C_1 &= c_2^* (1 + 2(c_4 + c_2 + c_0^* c_1)^2) = C, \end{aligned}$$

where C denotes different absolute constants. Theorem 7.8 then yields

Corollary 7.16. *Considering $T = m^{-1/2}$, we have that the RTP with hard walls is exponentially contractive in T -average with rate*

$$\nu = \frac{C\omega}{(\omega L)^2 + 1}.$$

Instantaneous RTP with jamming

In the run-and-tumble model of the previous section, the hard walls were a challenge that necessitated rewriting the exponential decay proof in a more intrinsic way, but the refresh mechanism remained the full refresh with respect to the conditional probability in v . The following process, coming from out-of-equilibrium statistical mechanics, combines hard walls and a refresh mechanism which is not the full refresh. It comes from a study in the discrete setting of [SEB16, SEB17] to consider Motility Induced Phase Separation (MIPS), i.e. to simplify concentration of particles only due to their mechanism of propulsion, or by [HGM23] in the continuous setting. It corresponds to two Run-and-Tumble particles on the 1D torus with jamming: once in contact they jam. Considering the interdistance process, its invariant measure exhibit an absolutely continuous part but also Dirac mass at the boundaries, seen as a caricature of MIPS. We will complete here its study of convergence to equilibrium by bounding the relaxation time.

Definition 7.17 (Instantaneous RTP). *Let $L > 0$ and set $\mathcal{S} = [0, L]$, $\mathcal{V} = \{-2, 0, 2\}$ and $\hat{\mathcal{S}} = \mathcal{S} \times \mathcal{V}$. Consider the process with generator*

$$\begin{aligned} \hat{\mathcal{L}}f(x, 2) &= 1_{\{x < L\}} 2\partial_x f(x, 2) + 2\omega(f(x, 0) - f(x, 2)), \\ \hat{\mathcal{L}}f(x, 0) &= \omega f(x, 2) + \omega f(x, -2) - 2\omega f(x, 0), \\ \hat{\mathcal{L}}f(x, -2) &= 1_{\{x > 0\}} (-2)\partial_x f(x, -2) + 2\omega(f(x, 0) - f(x, -2)), \end{aligned}$$

and invariant measure

$$\hat{\mu} = \sum_{\sigma=2,0,-2} (d_\sigma^0 \delta_0 + a_\sigma dx + d_\sigma^L \delta_L) \otimes \delta_\sigma$$

where

$$\begin{aligned} d_2^0 &= 0, & a_2 &= \frac{\omega/4}{2 + \omega L}, & d_2^L &= \frac{1/2}{2 + \omega L}, \\ d_0^0 &= \frac{1/2}{2 + \omega L}, & a_0 &= \frac{\omega/2}{2 + \omega L}, & d_0^L &= \frac{1/2}{2 + \omega L}, \\ d_{-2}^0 &= \frac{1/2}{2 + \omega L}, & a_{-2} &= \frac{\omega/4}{2 + \omega L}, & d_{-2}^L &= 0. \end{aligned}$$

The Dirichlet form of the collapse

$$\mathcal{E}(f) = \frac{\omega}{2 + \omega L} \int_0^L (\partial_x f)^2 dx$$

is the same as in the previous section, up to a factor 2, and the collapsed invariant measure

$$\mu = \frac{1}{2 + \omega L} \delta_0 + \frac{\omega}{2 + \omega L} dx + \frac{1}{2 + \omega L} \delta_L$$

is identical.

This time, we take

$$\begin{aligned} \hat{\mathcal{L}}_d f(x, +2) &= 1_{\{x < L\}} (+2) \partial_x f(x, +2) + 1_{\{x=L\}} \omega (f(x, 0) - f(x, +2)), \\ \hat{\mathcal{L}}_d f(x, 0) &= 1_{\{x=0\}} \omega (f(x, +2) - f(x, 0)) + 1_{\{x=L\}} \omega (f(x, -2) - f(x, 0)), \\ \hat{\mathcal{L}}_d f(x, -2) &= 1_{\{x > 0\}} (-2) \partial_x f(x, -2) + 1_{\{x=0\}} \omega (f(x, 0) - f(x, -2)), \end{aligned}$$

as well as $\gamma = \omega$ and

$$\begin{aligned} \hat{\mathcal{L}}_v f(x, +2) &= 1_{\{x < L\}} 2(f(x, 0) - f(x, +2)) + 1_{\{x=L\}} (f(x, 0) - f(x, +2)), \\ \hat{\mathcal{L}}_v f(x, 0) &= 1_{\{x > 0\}} (f(x, +2) - f(x, 0)) + 1_{\{x < L\}} (f(x, -2) - f(x, 0)), \\ \hat{\mathcal{L}}_v f(x, -2) &= 1_{\{x > 0\}} 2(f(x, 0) - f(x, -2)) + 1_{\{x=0\}} (f(x, 0) - f(x, -2)). \end{aligned}$$

Checking assumption (A) The only difference with the confined RTP is the Poincaré inequality in v of $\hat{\mathcal{L}}_v$ which we now show with constant $c_2^* = \frac{1}{2}$. We have

$$\mathcal{E}_v(f) = \int \int f(-\hat{\mathcal{L}}_v f) d\kappa_x d\mu$$

so it suffices to show that

$$(\kappa_x(f(x, \cdot)^2) - \kappa_x(f(x, \cdot)))^2 \leq \frac{1}{2} \int f(x, v) (-\hat{\mathcal{L}}_v f)(x, v) d\kappa_x(v) \quad (*)$$

for all $x \in [0, L]$. For $x \in (0, L)$ rewrite (*) as

$$\|f - \Pi f\|_S^2 \leq \frac{1}{2} \langle f, (-\mathcal{Q})f \rangle_S$$

where $\langle x, y \rangle_S := x^t S y$,

$$f = \begin{pmatrix} f(x, 2) \\ f(x, 0) \\ f(x, -2) \end{pmatrix}, \quad S = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{pmatrix}, \quad \mathcal{Q} = \begin{pmatrix} -2 & 2 & 0 \\ 1 & -2 & 1 \\ 0 & 2 & -2 \end{pmatrix},$$

and Π is the orthogonal projection on the kernel of \mathcal{Q} with respect to $\langle \cdot, \cdot \rangle_S$.

The matrix \mathcal{Q} is symmetric with respect to the scalar product $\langle \cdot, \cdot \rangle_S$ (i.e. $\mathcal{Q} = S^{-1} \mathcal{Q}^t S$) so it admits an orthonormal basis of eigenvectors. Its eigenvalues are 0, -2, -4 thus (*) follows.

The same arguments can be used for the cases $x = 0$ and $x = L$.

Checking assumption (B) The same arguments as for the confined RTP apply.

Checking assumption (C1) We have that $u = f - \Pi_v f$ satisfies $\Pi_v u = 0$ so in particular

$$\frac{1}{4}u(x, 2) + \frac{1}{2}u(x, 0) + \frac{1}{4}u(x, -2) = 0 \iff u(x, 2) + u(x, -2) = -2u(x, 0).$$

Therefore

$$\begin{aligned} & \left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\hat{\mu})} \\ &= \frac{\omega/4}{2 + \omega L} \int_0^L (+2)\partial_x g(x)(+2)\partial_x u(x, +2)dx \\ & \quad + \frac{\omega/4}{2 + \omega L} \int_0^L (-2)\partial_x g(x)(-2)\partial_x u(x, -2)dx \\ &= \frac{\omega}{2 + \omega L} \int_0^L \partial_x g(x)(\partial_x u(x, +2) + \partial_x u(x, -2))dx \\ &= \frac{-2\omega}{2 + \omega L} \int_0^L \partial_x g(x)\partial_x u(x, 0)dt \end{aligned}$$

so we have

$$\begin{aligned} \left| \left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\hat{\mu})} \right| &= \left| \langle 2\mathcal{L}g, u(\cdot, 0) \rangle_{L^2(\mu)} \right| \\ &\leq \|2\mathcal{L}g\|_{L^2(\mu)} \|u(\cdot, 0)\|_{L^2(\mu)}. \end{aligned}$$

Finally

$$\begin{aligned} \|u(\cdot, 0)\|_{L^2(\mu)}^2 &= \frac{1}{2 + \omega L} u(0, 0)^2 + \frac{\omega}{2 + \omega L} \int_0^L u(x, 0)^2 dx + \frac{1}{2 + \omega L} u(L, 0)^2 \\ &\leq 2 \left[\frac{1}{2 + \omega L} \left(\frac{1}{2} u(0, -2)^2 + \frac{1}{2} u(0, 0)^2 \right) \right. \\ & \quad + \frac{\omega}{2 + \omega L} \int_0^L \left(\frac{1}{4} u(x, -2)^2 + \frac{1}{2} u(x, 0)^2 + \frac{1}{4} u(x, +2)^2 \right) dx \\ & \quad \left. + \frac{1}{2 + \omega L} \left(\frac{1}{2} u(L, +2)^2 + \frac{1}{2} u(L, 0)^2 \right) \right] \\ &\leq 2 \|u\|_{L^2(\hat{\mu})} \end{aligned}$$

and hence

$$\left\langle \hat{\mathcal{L}}_d(g \circ \pi), \hat{\mathcal{L}}_d(f - \Pi_v f) \right\rangle_{L^2(\hat{\mu})} \leq 2 \|2\mathcal{L}g\|_{L^2(\mu)} \|u\|_{L^2(\hat{\mu})}$$

and we can take $c_0^* = 2$.

Checking assumption (C2) Let us show that inequality (C2) is satisfied with $c_1^* = 2$.
By Cauchy-Schwarz

$$\left| \left\langle \hat{\mathcal{L}}_v f, \hat{\mathcal{L}}_d(g \circ \pi) \right\rangle_{L^2(\hat{\mu})} \right| \leq \|\hat{\mathcal{L}}_v f\|_{L^2(\hat{\mu})} \|\hat{\mathcal{L}}_d(g \circ \pi)\|_{L^2(\hat{\mu})} = \|\hat{\mathcal{L}}_v f\|_{L^2(\hat{\mu})} \sqrt{2\mathcal{E}(g)}$$

so it suffices to show that $\|\hat{\mathcal{L}}_v f\|_{L^2(\hat{\mu})} \leq 2\sqrt{\mathcal{E}_v(f)}$.

We have

$$\|\hat{\mathcal{L}}_v f\|_{L^2(\hat{\mu})}^2 = \int \int (\hat{\mathcal{L}}_v f)^2 d\kappa_x d\mu \quad \text{and} \quad \mathcal{E}_v(f) = \int \int f(-\hat{\mathcal{L}}_v f) d\kappa_x d\mu$$

so it is enough to show

$$\int [\hat{\mathcal{L}}_v f(x, v)]^2 d\kappa_x(v) \leq 4 \int f(x, v) (-\hat{\mathcal{L}}_v f)(x, v) d\kappa_x(v)$$

for all $x \in [0, L]$. For $x \in (0, L)$ this is equivalent to

$$\|\mathcal{Q}f\|_S^2 \leq 4 \langle f, -\mathcal{Q}f \rangle_S$$

with the same notation as in the verification of assumption (A). The inequality again follows from the spectral decomposition of \mathcal{Q} in an orthonormal basis. The cases $x = 0$ and $x = L$ follow from the same arguments.

Conclusion: rate of convergence to equilibrium for the interdistance of two RTPs on the 1D torus with jamming. Taking $T = m^{-1/2}$ in Theorem 7.7 yields

$$\begin{aligned} C_0 &= 2(c_1^* c_3)^2 = C \frac{1}{m_{\text{sti}}} = \frac{C}{\omega^2} ((\omega L)^2 + \omega L), \\ C_1 &= c_2^* (1 + 2(c_4 + c_2 + c_0^* c_1)^2) = C, \end{aligned}$$

where C denotes different absolute constants. Theorem 7.8 then yields

Corollary 7.18. *Considering $T = m^{-1/2}$, we have that the interdistance process of two RTPs on the 1D torus with jamming is exponentially contractive in T -average with rate*

$$\nu = \frac{C\omega}{(\omega L)^2 + 1}.$$

Let us remark that the rate obtained implies that the relaxation time of the interdistance process is of the same order than the mixing time obtained in [Hah24].

Bibliography

- [AAMN19] Dallas Albritton, Scott Armstrong, Jean-Christophe Mourrat, and Matthew Novack. Variational methods for the kinetic fokker-planck equation. *arXiv preprint arXiv:1902.04037*, 2019.
- [ABG⁺14] Romain Azaïs, Jean-Baptiste Bardet, Alexandre Génadot, Nathalie Krell, and Pierre-André Zitt. Piecewise deterministic markov process—recent results. In *Esaim: Proceedings*, volume 44, pages 276–290. EDP Sciences, 2014.
- [ADNR21] Christophe Andrieu, Alain Durmus, Nikolas Nüsken, and Julien Roussel. Hypocoercivity of piecewise deterministic markov process-monte carlo. *The Annals of Applied Probability*, 31(5):2478–2517, 2021.
- [AdPLVW19] Thibaut Arnoux de Pirey, Gustavo Lozano, and Frédéric Van Wijland. Active hard spheres in infinitely many dimensions. *Physical review letters*, 123(26):260602, 2019.
- [AdPvW23] Thibaut Arnoux de Pirey and Frédéric van Wijland. A run-and-tumble particle around a spherical obstacle: the steady-state distribution far-from-equilibrium. *Journal of Statistical Mechanics: Theory and Experiment*, 2023(9):093202, 2023.
- [AG19] L Angelani and R Garra. Run-and-tumble motion in one dimension with space-dependent speed. *Physical Review E*, 100(5):052147, 2019.
- [Ang17] Luca Angelani. Confined run-and-tumble swimmers in one dimension. *Journal of Physics A: Mathematical and Theoretical*, 50(32):325601, 2017.
- [ARKL21] Tal Agranov, Sunghan Ro, Yariv Kafri, and Vivien Lecomte. Exact fluctuating hydrodynamics of active lattice gases—typical fluctuations. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(8):083208, 2021.
- [BB72] Howard C Berg and Douglas A Brown. Chemotaxis in escherichia coli analysed by three-dimensional tracking. *nature*, 239(5374):500–504, 1972.
- [BB95] Elena O Budrene and Howard C Berg. Dynamics of formation of symmetrical patterns by chemotactic bacteria. *Nature*, 376(6535):49–53, 1995.
- [BBCG08] Dominique Bakry, Franck Barthe, Patrick Cattiaux, and Arnaud Guillin. A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Electron. Commun. Probab.*, 13:60–66, 2008.
- [BCD⁺13] A. Bricard, J. B. Caussin, N. Desreumaux, O. Dauchot, and D. Bartolo. Emergence of macroscopic directed motion in populations of motile colloids. *Nature*, 2013.
- [BCG⁺13] Jean-Baptiste Bardet, Alejandra Christen, Arnaud Guillin, Florent Malrieu, and Pierre-André Zitt. Total variation estimates for the tcp process. 2013.

- [BCVD18] Alexandre Bouchard-Côté, Sebastian J Vollmer, and Arnaud Doucet. The bouncy particle sampler: A nonreversible rejection-free markov chain monte carlo method. *Journal of the American Statistical Association*, 113(522):855–867, 2018.
- [BDSG⁺02] Lorenzo Bertini, Alberto De Sole, Davide Gabrielli, Giovanni Jona-Lasinio, and Claudio Landim. Macroscopic fluctuation theory for stationary non-equilibrium states. *Journal of Statistical Physics*, 107:635–675, 2002.
- [Ber04] Howard C Berg. *E. Coli in Motion*. Springer, 2004.
- [BFR19] Joris Bierkens, Paul Fearnhead, and Gareth Roberts. The Zig-Zag process and super-efficient sampling for Bayesian analysis of big data. *The Annals of Statistics*, 47(3):1288 – 1320, 2019.
- [BGL14] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.
- [BGMS23] Joris Bierkens, Sebastiano Grazi, Frank van der Meulen, and Moritz Schauer. Sticky PDMP samplers for sparse and local inference problems. *Statistics and Computing*, 33(1):8, 2023.
- [BGvdMS] J. Bierkens, S. Grazi, F. van der Meulen, and M. Schauer. Sticky PDMP samplers for sparse and local inference problems. Arxiv:2103.08478.
- [BGW15] François Bolley, Arnaud Guillin, and Xinyu Wang. Non ultracontractive heat kernel bounds by Lyapunov conditions. *Discrete Contin. Dyn. Syst.*, 35(3):857–870, 2015.
- [BH12] Yuri Bakhtin and Tobias Hurth. Invariant densities for dynamical systems with random switching. *Nonlinearity*, 25(10):2937, 2012.
- [BH22] Michel Benaïm and Tobias Hurth. *Markov Chains on Metric Spaces: A Short Course*. Springer Nature, 2022.
- [BHLM18] Yuri Bakhtin, Tobias Hurth, Sean D Lawley, and Jonathan C Mattingly. Smooth invariant densities for random switching on the torus. *Nonlinearity*, 31(4):1331, 2018.
- [BHS18] Michel Benaïm, Tobias Hurth, and Edouard Strickler. A user-friendly condition for exponential ergodicity in randomly switched environments. 2018.
- [Bil99] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [BKW09] Etienne P Bernard, Werner Krauth, and David B Wilson. Event-chain monte carlo algorithms for hard-sphere systems. *Physical Review E*, 80(5):056704, 2009.
- [BLBMZ12] Michel Benaïm, Stéphane Le Borgne, Florent Malrieu, and Pierre-André Zitt. Quantitative ergodicity for some switched dynamical systems. *Electronic Communications in Probability*, 2012.
- [BLBMZ14] Michel Benaïm, Stéphane Le Borgne, Florent Malrieu, and Pierre-André Zitt. On the stability of planar randomly switched systems. 2014.

- [BLBMZ15] Michel Benaïm, Stéphane Le Borgne, Florent Malrieu, and Pierre-André Zitt. Qualitative properties of certain piecewise deterministic markov processes. In *Annales de l'IHP Probabilités et statistiques*, volume 51, pages 1040–1075, 2015.
- [BMR⁺20] Urna Basu, Satya N Majumdar, Alberto Rosso, Sanjib Sabhapandit, and Grégory Schehr. Exact stationary state of a run-and-tumble particle with three internal states in a harmonic trap. *Journal of Physics A: Mathematical and Theoretical*, 53(9):09LT01, 2020.
- [Bre11] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.
- [BRSS17] Nawaf Bou-Rabee and Jesús María Sanz-Serna. Randomized Hamiltonian Monte Carlo. *The Annals of Applied Probability*, 27(4):2159 – 2194, 2017.
- [BRZ19] Joris Bierkens, Gareth O. Roberts, and Pierre-André Zitt. Ergodicity of the zigzag process. *Ann. Appl. Probab.*, 29(4):2266–2301, 2019.
- [BS23] Giovanni Brigati and Gabriel Stoltz. How to construct decay rates for kinetic fokker–planck equations? *arXiv preprint arXiv:2302.14506*, 2023.
- [BSWW24] Giovanni Brigati, Gabriel Stoltz, Andi Wang, and Lihan Wang. Explicit convergence rates of underdamped langevin dynamics under weighted and weak poincaré–lions inequalities. *preprint, <https://arxiv.org/pdf/2407.16033>*, 2024.
- [BT22] Michel Benaïm and Oliver Tough. Regularity of the stationary density for systems with fast random switching. *arXiv preprint arXiv:2212.03632*, 2022.
- [BT23] Michel Benaïm and Oliver Tough. On invariant distributions of feller markov chains with applications to dynamical systems with random switching. *arXiv e-prints*, pages arXiv–2310, 2023.
- [BY93] Frank Ball and Geoffrey F Yeo. Lumpability and marginalisability for continuous-time Markov chains. *Journal of Applied Probability*, 30(3):518–528, 1993.
- [Cat12] Michael E Cates. Diffusive transport without detailed balance in motile bacteria: does microbiology need statistical physics? *Reports on Progress in Physics*, 75(4):042601, 2012.
- [CD08] Oswaldo LV Costa and François Dufour. Stability and ergodicity of piecewise deterministic markov processes. *SIAM Journal on Control and Optimization*, 47(2):1053–1077, 2008.
- [CG08] Patrick Cattiaux and Arnaud Guillin. Deviation bounds for additive functionals of markov processes. *ESAIM: Probability and Statistics*, 12:12–29, 2008.
- [CGZ13] Patrick Cattiaux, Arnaud Guillin, and Pierre André Zitt. Poincaré inequalities and hitting times. In *Annales de l'IHP Probabilités et statistiques*, volume 49, pages 95–118, 2013.
- [CH13] Ting-Li Chen and Chii-Ruey Hwang. Accelerating reversible Markov chains. *Statist. Probab. Lett.*, 83(9):1956–1962, 2013.
- [CH15] Bertrand Cloez and Martin Hairer. Exponential ergodicity for Markov processes with random switching. *Bernoulli*, 21(1), 2015.

- [CLP99] Fang Chen, László Lovász, and Igor Pak. Lifting markov chains to speed up mixing. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 275–281, 1999.
- [CLW23] Yu Cao, Jianfeng Lu, and Lihan Wang. On explicit l_2 -convergence rate estimate for underdamped langevin dynamics. *Archive for Rational Mechanics and Analysis*, 247(5):90, 2023.
- [CMP10] Djali Chafaï, Florent Malrieu, and Katy Paroux. On the long time behavior of the tcp window size process. *Stochastic Processes and their Applications*, 120(8):1518–1534, 2010.
- [CMP20] Lorenzo Caprini, U Marini Bettolo Marconi, and Andrea Puglisi. Spontaneous velocity alignment in motility-induced phase separation. *Physical review letters*, 124(7):078001, 2020.
- [Cou96] Thierry Coulhon. Ultracontractivity and Nash type inequalities. *J. Funct. Anal.*, 141(2):510–539, 1996.
- [CRS14] Vincent Calvez, Gaël Raoul, and Christian Schmeiser. Confinement by biased velocity jumps: aggregation of Escherichia Coli. *arXiv preprint arXiv:1404.0643*, 2014.
- [CT15] Michael E Cates and Julien Tailleur. Motility-induced phase separation. *Annual Review of Condensed Matter Physics*, 6(1):219–244, 2015.
- [Dav84] Mark HA Davis. Piecewise-deterministic markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 46(3):353–376, 1984.
- [Dav93] Mark HA Davis. *Markov models & optimization*, volume 49. CRC Press, 1993.
- [DBG⁺18] Antoine Deblais, Thomas Barois, T Guerin, Pierre-Henri Delville, Rémi Vaudaine, Juho S Lintuvuori, Jean-François Boudet, Jean-Christophe Baret, and H Kellay. Boundaries control collective dynamics of inertial self-propelled robots. *Physical review letters*, 120(18):188002, 2018.
- [DC99] François Dufour and Oswaldo LV Costa. Stability of piecewise-deterministic markov processes. *SIAM Journal on Control and Optimization*, 37(5):1483–1502, 1999.
- [DDK20a] Arghya Das, Abhishek Dhar, and Anupam Kundu. Gap statistics of two interacting run and tumble particles in one dimension. *Journal of Physics A: Mathematical and Theoretical*, 53(34):345003, 2020.
- [DDK⁺20b] Pritha Dolai, Arghya Das, Anupam Kundu, Chandan Dasgupta, Abhishek Dhar, and K Vijay Kumar. Universal scaling in active single-file dynamics. *Soft Matter*, 16(30):7077–7087, 2020.
- [Der07] Bernard Derrida. Non-equilibrium steady states: fluctuations and large deviations of the density and of the current. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(07):P07023, 2007.
- [DGM20] Alain Durmus, Arnaud Guillin, and Pierre Monmarché. Geometric ergodicity of the Bouncy Particle Sampler. *The Annals of Applied Probability*, 30(5), October 2020.

- [DGM21] Alain Durmus, Arnaud Guillin, and Pierre Monmarché. Piecewise deterministic markov processes and their invariant measures. In *Annales de l'Institut Henri Poincaré (B) Probabilités et statistiques*, volume 57, pages 1442–1475. Institut Henri Poincaré, 2021.
- [DGR02] Vincent Dumas, Fabrice Guillemin, and Philippe Robert. A markovian analysis of additive-increase multiplicative-decrease algorithms. *Advances in Applied Probability*, 34(1):85–111, 2002.
- [DHHM22] Helge Dietert, Frédéric Hérau, Harsha Hutridurga, and Clément Mouhot. Quantitative geometric control in linear kinetic theory. *arXiv preprint arXiv:2209.09340*, 2022.
- [DHN00] Persi Diaconis, Susan Holmes, and Radford M Neal. Analysis of a nonreversible markov chain sampler. *Annals of Applied Probability*, pages 726–752, 2000.
- [DKM⁺19] Abhishek Dhar, Anupam Kundu, Satya N Majumdar, Sanjib Sabhapandit, and Grégory Schehr. Run-and-tumble particle in one-dimensional confining potentials: Steady-state, relaxation, and first-passage properties. *Physical Review E*, 99(3):032132, 2019.
- [DMS09] Jean Dolbeault, Clément Mouhot, and Christian Schmeiser. Hypocoercivity for kinetic equations with linear relaxation terms. *Comptes Rendus. Mathématique*, 347(9-10):511–516, 2009.
- [DMS15] Jean Dolbeault, Clément Mouhot, and Christian Schmeiser. Hypocoercivity for linear kinetic equations conserving mass. *Transactions of the American Mathematical Society*, 367(6):3807–3828, 2015.
- [Ebe09] Andreas Eberle. Markov processes. *Lecture Notes at University of Bonn*, 2009.
- [Ebe16] Andreas Eberle. Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields*, 166(3-4):851–886, 2016.
- [EG15] Jens Elgeti and Gerhard Gompper. Run-and-tumble dynamics of self-propelled particles in confinement. *Europhysics Letters*, 109(5):58003, 2015.
- [EGZ19] Andreas Eberle, Arnaud Guillin, and Raphael Zimmer. Couplings and quantitative contraction rates for Langevin dynamics. *Ann. Probab.*, 47(4):1982–2010, 2019.
- [EK86] S.N. Ethier and T. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.
- [EL24a] Andreas Eberle and Francis Lörler. Non-reversible lifts of reversible diffusion processes and relaxation times. *arXiv preprint arXiv:2402.05041*, 2024.
- [EL24b] Andreas Eberle and Francis Lörler. Quantitative divergence lemmas in space-time domains, 2024. Private communications.
- [EM18] Martin R Evans and Satya N Majumdar. Run and tumble particle under resetting: a renewal approach. *Journal of Physics A: Mathematical and Theoretical*, 51(47):475003, 2018.
- [EN01] Klaus-Jochen Engel and Rainer Nagel. One-parameter semigroups for linear evolution equations. *Semigroup Forum*, 63:278–280, 06 2001.

- [EP14] Hans-Jürgen Engelbert and Goran Peskir. Stochastic differential equations for sticky Brownian motion. *Stochastics*, 86(6):993–1021, 2014.
- [Eri21] Clément Erignoux. Hydrodynamic limit for an active exclusion process. *Mémoires de la Société Mathématique de France*, 169:1–206, 2021.
- [EWG15] Jens Elgeti, Roland G Winkler, and Gerhard Gompper. Physics of microswimmers—single particle motion and collective behavior: a review. *Reports on progress in physics*, 78(5):056601, 2015.
- [FBPR18] Paul Fearnhead, Joris Bierkens, Murray Pollock, and Gareth O Roberts. Piecewise deterministic markov processes for continuous-time monte carlo. *Statistical Science*, 33(3):386–412, 2018.
- [Fel52] William Feller. The parabolic differential equations and the associated semi-groups of transformations. *Ann. of Math. (2)*, 55:468–519, 1952.
- [Fel54] William Feller. Diffusion processes in one dimension. *Trans. Amer. Math. Soc.*, 77:1–31, 1954.
- [FGM12] Joaquin Fontbona, Hélène Guérin, and Florent Malrieu. Quantitative estimates for the long-time behavior of an ergodic variant of the telegraph process. *Advances in Applied Probability*, 44(4):977–994, 2012.
- [FGM16] Joaquin Fontbona, Hélène Guérin, and Florent Malrieu. Long time behavior of telegraph processes under convex potentials. *Stochastic Processes and their Applications*, 126(10):3077–3101, 2016.
- [FKB15] Thomas FF Farage, Philip Krinninger, and Joseph M Brader. Effective interactions in active brownian suspensions. *Physical Review E*, 91(4):042310, 2015.
- [Fry21] Derek Frydel. Generalized run-and-tumble model in 1d geometry for an arbitrary distribution of drift velocities. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(8):083220, 2021.
- [Gal11] G. P. Galdi. *An introduction to the mathematical theory of the Navier-Stokes equations*. Springer Monographs in Mathematics. Springer, New York, second edition, 2011. Steady-state problems.
- [GGSC06] Evarist Giné, Geoffrey R Grimmett, and Laurent Saloff-Coste. *Lectures on probability theory and statistics: Ecole d’été de probabilités de Saint-Flour XXVI-1996*. Springer, 2006.
- [GHM24] Arnaud Guillin, Leo Hahn, and Manon Michel. Long time analysis of a pair of on-lattice and continuous run-and-tumble particles with jamming, 2024. In preparation, see Chapter 4 for a preprint.
- [GM16] Arnaud Guillin and Pierre Monmarché. Optimal linear drift for the speed of convergence of an hypoelliptic diffusion. *Electron. Commun. Probab.*, 21:Paper No. 74, 14, 2016.
- [GMK10] Wainrib Gilles, Thieullen Michele, and Pakdaman Khashayar. Intrinsic variability of latency to first-spike. *Biological cybernetics*, 103:43–56, 2010.
- [GMS23] Mathis Guéneau, Satya N Majumdar, and Grégory Schehr. Active particle in a harmonic trap driven by a resetting noise: an approach via kesten variables. *Journal of Physics A: Mathematical and Theoretical*, 56(47):475002, 2023.

- [H07] Frédéric Hérau. Short and long time behavior of the Fokker-Planck equation in a confining potential and applications. *J. Funct. Anal.*, 244(1):95–118, 2007.
- [Hah24] Leo Hahn. Invariant measures and mixing of two interacting run-and-tumble particles with jamming, 2024. In preparation, see Chapter 5 for a preprint.
- [HGM23] Leo Hahn, Arnaud Guillin, and Manon Michel. Jamming pair of general run-and-tumble particles: Exact results and universality classes. *arXiv preprint arXiv:2306.00831*, 2023.
- [HHMS93] Chii-Ruey Hwang, Shu-Yin Hwang-Ma, and Shuenn Jyi Sheu. Accelerating Gaussian diffusions. *Ann. Appl. Probab.*, 3(3):897–913, 1993.
- [HHMS05] Chii-Ruey Hwang, Shu-Yin Hwang-Ma, and Shuenn-Jyi Sheu. Accelerating diffusions. *Ann. Appl. Probab.*, 15(2):1433–1444, 2005.
- [HM11] Martin Hairer and Jonathan C Mattingly. Yet another look at Harris’ ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI: Centro Stefano Franscini, Ascona, May 2008*, pages 109–117. Springer, 2011.
- [HMKK17] Julian Harland, Manon Michel, Tobias A Kampmann, and Jan Kierfeld. Event-chain monte carlo algorithms for three-and many-particle interactions. *Europhysics Letters*, 117(3):30001, 2017.
- [Hör67] Lars Hörmander. Hypoelliptic second order differential equations. 1967.
- [Jou24] Benjamin Jourdain. Algorithmes de Monte Carlo par chaînes de Markov et méthodes particulières, 2024. <http://cermics.enpc.fr/~jourdain/M2SU/polymcmc.pdf>.
- [Kha60] Rafail Z Khas’minskii. Ergodic properties of recurrent diffusion processes and stabilization of the solution to the cauchy problem for parabolic equations. *Theory of Probability & Its Applications*, 5(2):179–196, 1960.
- [KHEBT18] Mourtaza Kourbane-Houssene, Clément Erignoux, Thierry Bodineau, and Julien Tailleur. Exact hydrodynamic description of active lattice gases. *Physical review letters*, 120(26):268003, 2018.
- [KL98] Claude Kipnis and Claudio Landim. *Scaling limits of interacting particle systems*, volume 320. Springer Science & Business Media, 1998.
- [Kli87] Wolfgang Kliemann. Recurrence and invariant measures for degenerate diffusions. *The annals of probability*, pages 690–707, 1987.
- [LDMS19] Pierre Le Doussal, Satya N Majumdar, and Grégory Schehr. Noncrossing run-and-tumble particles on a line. *Physical Review E*, 100(1):012113, 2019.
- [LDMS21] Pierre Le Doussal, Satya N Majumdar, and Grégory Schehr. Stationary nonequilibrium bound state of a pair of run and tumble particles. *Physical Review E*, 104(4):044103, 2021.
- [LEB18] E Ó Laighléis, Martin R Evans, and Richard A Blythe. Minimal stochastic field equations for one-dimensional flocking. *Physical Review E*, 98(6):062127, 2018.
- [Lez01] Pascal Lezaud. Chernoff and Berry–Esséen inequalities for Markov processes. *ESAIM: Probability and Statistics*, 5:183–201, 2001.

- [LG16] Jean-François Le Gall. *Brownian motion, martingales, and stochastic calculus*. Springer, 2016.
- [Lig10] Thomas Milton Liggett. *Continuous time Markov processes: an introduction*, volume 113. American Mathematical Soc., 2010.
- [LMR13] Sean D Lawley, Jonathan C Mattingly, and Michael C Reed. Sensitivity to switching rates in stochastically switched odes. *arXiv preprint arXiv:1310.2525*, 2013.
- [LMT96] Robert B Lund, Sean P Meyn, and Richard L Tweedie. Computable exponential convergence rates for stochastically ordered markov processes. *The Annals of Applied Probability*, 6(1):218–237, 1996.
- [LNP13] T. Lelièvre, F. Nier, and G. A. Pavliotis. Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion. *J. Stat. Phys.*, 152(2):237–274, 2013.
- [LPW09] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.
- [LVL08] Andreas H Löpker and Johan SH Van Leeuwaarden. Transient moments of the tcp window size process. *Journal of applied probability*, 45(1):163–175, 2008.
- [LW22] Jianfeng Lu and Lihan Wang. On explicit l_2 -convergence rate estimate for piecewise deterministic markov processes in mcmc algorithms. *The Annals of Applied Probability*, 32(2):1333–1361, 2022.
- [Mal16] Florent Malrieu. Some simple but challenging Markov processes. *Annales de la Faculté des sciences de Toulouse : Mathématiques*, 24(4):857–883, January 2016.
- [MBE19] Emil Mallmin, Richard A Blythe, and Martin R Evans. Exact spectral solution of two interacting run-and-tumble particles on a ring lattice. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(1):013204, 2019.
- [MDS20] Manon Michel, Alain Durmus, and Stéphane Sénécal. Forward event-chain monte carlo: Fast sampling by randomness control in irreversible markov chains. *Journal of Computational and Graphical Statistics*, 29(4):689–702, 2020.
- [MEB22] Matthew J Metson, Martin R Evans, and Richard A Blythe. From a microscopic solution to a continuum description of interacting active particles. *Physical Review E*, 2022.
- [MEB23] Matthew J Metson, Martin R Evans, and Richard A Blythe. Tuning attraction and repulsion between active particles through persistence. *Europhysics Letters*, 141(4):41001, 2023.
- [MGM23] Athina Monemvassitis, Arnaud Guillin, and Manon Michel. PDMP characterisation of event-chain monte carlo algorithms for particle systems. *Journal of Statistical Physics*, 190(3):66, 2023.
- [MJK⁺18] Kanaya Malakar, V Jemseena, Anupam Kundu, K Vijay Kumar, Sanjib Sabhapandit, Satya N Majumdar, S Redner, and Abhishek Dhar. Steady state, relaxation and first-passage properties of a run-and-tumble particle in one-dimension. *Journal of Statistical Mechanics: Theory and Experiment*, 2018(4):043215, 2018.

- [MJR⁺13] M Cristina Marchetti, Jean-François Joanny, Sriram Ramaswamy, Tanniemola B Liverpool, Jacques Prost, Madan Rao, and R Aditi Simha. Hydrodynamics of soft active matter. *Reviews of modern physics*, 85(3):1143, 2013.
- [MKK14] Manon Michel, Sebastian C Kapfer, and Werner Krauth. Generalized event-chain monte carlo: Constructing rejection-free global-balance algorithms from infinitesimal steps. *The Journal of chemical physics*, 140(5):054116, 2014.
- [MLDMS21] Francesco Mori, Pierre Le Doussal, Satya N Majumdar, and Grégory Schehr. Condensation transition in the late-time position of a run-and-tumble particle. *Physical Review E*, 103(6):062134, 2021.
- [MNS⁺19] Siddharth Mayya, Gennaro Notomista, Dylan Shell, Seth Hutchinson, and Magnus Egerstedt. Non-uniform robot densities in vibration driven swarms using phase separation theory. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4106–4112. IEEE, 2019.
- [Mon14] Pierre Monmarché. *Hypocoercivité: approches alternatives et applications aux algorithmes stochastiques*. PhD thesis, Université de Toulouse, Université Toulouse III-Paul Sabatier, 2014.
- [MT93] Sean P Meyn and Richard L Tweedie. Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 25(3):518–548, 1993.
- [MZ09] Krishanu Maulik and Bert Zwart. An extension of the square root law of tcp. *Annals of Operations Research*, 170(1):217–232, 2009.
- [NMKH15] Yoshihiko Nishikawa, Manon Michel, Werner Krauth, and Koji Hukushima. Event-chain algorithm for the Heisenberg model: evidence for $z \simeq 1$ dynamic scaling. *Phys. Rev. E (3)*, 92(6):063306, 5, 2015.
- [OK08] Teunis J Ott and JHB Kemperman. Transient behavior of processes in the tcp paradigm. *Probability in the Engineering and Informational Sciences*, 22(3):431–471, 2008.
- [Oks13] Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.
- [OKTvW22] Jérémy O’Byrne, Yariv Kafri, Julien Tailleur, and Frédéric van Wijland. Time irreversibility in active matter, from micro to macro. *Nature Reviews Physics*, 4(3):167–183, 2022.
- [Olv10] Frank WJ Olver. *NIST handbook of mathematical functions hardback and CD-ROM*. Cambridge university press, 2010.
- [Par06] Etienne Pardoux. Processus de markov et applications, 2006. <https://www.association-tremplin.org/Documents/Eleves/Apres/CoursChainedeMarkovM1Pardoux.pdf>.
- [PdW12] Elias AJF Peters and G de With. Rejection-free monte carlo sampling for general potentials. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 85(2):026703, 2012.
- [Ram10] Sriram Ramaswamy. The mechanics and statistics of active matter. *Annual Review of Condensed Matter Physics*, 1(1):323–345, 2010.
- [Sch93] Mark J Schnitzer. Theory of continuum random walks and application to chemotaxis. *Physical Review E*, 48(4):2553, 1993.

- [SEB16] AB Slowman, MR Evans, and RA Blythe. Jamming and attraction of interacting run-and-tumble random walkers. *Physical review letters*, 116(21):218101, 2016.
- [SEB17] AB Slowman, MR Evans, and RA Blythe. Exact solution of two interacting run-and-tumble random walkers with finite tumble duration. *Journal of Physics A: Mathematical and Theoretical*, 50(37):375601, 2017.
- [SFK17] Stefano Steffenoni, Gianmaria Falasco, and Klaus Kroy. Microscopic derivation of the hydrodynamics of active-brownian-particle suspensions. *Physical Review E*, 95(5):052142, 2017.
- [SLDMS22] Naftali R Smith, Pierre Le Doussal, Satya N Majumdar, and Grégory Schehr. Exact position distribution of a harmonically confined run-and-tumble particle in two dimensions. *Physical Review E*, 106(5):054133, 2022.
- [SNLK01] T. Surrey, F. Nédélec, S. Leibler, and E. Karsenti. Physical properties determining self-organization of motors and microtubules. *Science*, 2001.
- [SSB12] Jonathan Saragosti, Pascal Silberzan, and Axel Buguin. Modeling E. coli tumbles by rotational diffusion. implications for chemotaxis. *PloS one*, 7(4):e35412, 2012.
- [SSK20] Prashant Singh, Sanjib Sabhapandit, and Anupam Kundu. Run-and-tumble particle in inhomogeneous media in one dimension. *Journal of Statistical Mechanics: Theory and Experiment*, 2020(8):083207, 2020.
- [TC08] J Tailleur and ME Cates. Statistical mechanics of interacting run-and-tumble bacteria. *Physical review letters*, 100(21):218103, 2008.
- [TT98] John Toner and Yuhai Tu. Flocks, herds, and schools: A quantitative theory of flocking. *Physical review E*, 58(4):4828, 1998.
- [TW21] Francesco Turci and Nigel B Wilding. Phase separation and multibody effects in three-dimensional active brownian particles. *Physical Review Letters*, 126(3):038002, 2021.
- [Vil06] Cédric Villani. Hypocoercive diffusion operators. In *International Congress of Mathematicians*, volume 3, pages 473–498, 2006.
- [Vil09] Cédric Villani. *Hypocoercivity*, volume 202. American Mathematical Society, 2009.
- [VZ12] Tamás Vicsek and Anna Zafeiris. Collective motion. *Physics reports*, 517(3-4):71–140, 2012.
- [Wan00] Feng-Yu Wang. Functional inequalities for empty essential spectrum. *J. Funct. Anal.*, 170(1):219–245, 2000.
- [WL00] Xiao-Lun Wu and Albert Libchaber. Particle diffusion in a quasi-two-dimensional bacterial bath. *Physical review letters*, 84(13):3017, 2000.
- [WPL⁺21] Gao Wang, Trung V Phan, Shengkai Li, Michael Wombacher, Junle Qu, Yan Peng, Guo Chen, Daniel I Goldman, Simon A Levin, Robert H Austin, et al. Emergent field-driven robot swarm states. *Physical review letters*, 126(10):108002, 2021.
- [Wu00] Liming Wu. A deviation inequality for non-reversible markov processes. In *Annales de l’Institut Henri Poincaré (B) Probability and Statistics*, volume 36, pages 435–445. Elsevier, 2000.

