https://doi.org/10.35662/unine-thesis-3222

Université de Neuchâtel

Faculty of Science
Mathematical Institute

On ergodic properties for systems of degenerate Stochastic Differential Equations

Thesis

Presented to the Faculty of Science
in fulfillment of the requirements for the degree of Doctor of Science in Mathematics

by

Jérémy Colombo

Accepted upon the recommendation of the jury:

Prof. Michel Benaïm	University of Neuchâtel, CH	Thesis Director
Prof. Christian Mazza	University of Fribourg, CH	Reviewer
Prof. Felix Schlenk	University of Neuchâtel, CH	Reviewer
Dr. Edouard Strickler	CNRS, University of Lorraine, FR	Reviewer

Successfully defended on 20 November 2025

unine
Université de Neuchâtel
Faculté des sciences

Rue Emile-Argand 11
CH-2000 Neuchâtel
doctorat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Monsieur Jérémy COLOMBO

Titre :

"On ergodic properties for systems of degenerate Stochastic Differential Equations"

sur le rapport des membres du jury composé comme suit :

- **Prof. Michel Benaïm,** directeur de thèse, Université de Neuchâtel, Suisse
- **Prof. Christian Mazza,** Université de Fribourg, Suisse
- **Prof. Felix Schlenk,** Université de Neuchâtel, Suisse
- **Dr Edouard Strickler,** CNRS, Université de Lorraine, France

Neuchâtel, le 25 novembre 2025

Le Doyen, Prof. P. Brunner

Abstract

The first part of this thesis focuses on the study of degenerate stochastic models typically represented by ecological SDEs. We first consider an ecological model of interacting species, seeking conditions for almost-sure persistence versus extinction of one or more species.

In particular, we focus on the classical two-dimensional Rosenzweig–MacArthur prey – predator model [RM63] under a degenerate stochastic perturbation, in which only the prey variable is subject to small environmental fluctuations.

Concretely, we analyze the Rosenzweig–MacArthur prey–predator model on $\mathbb{R}_+^2$ with degenerate noise affecting only the prey:

$$\begin{cases} \mathrm{d}x_1 = x_1 \left(F_1(x_1, x_2)\mathrm{d}t + \varepsilon \mathrm{d}B_t\right), \\ \mathrm{d}x_2 = x_2 F_2(x_1, x_2)\mathrm{d}t, \end{cases} \tag{1}$$

where $\varepsilon > 0$ is sufficiently small, $(B_t)_{t\geq 0}$ is a one-dimensional Brownian motion, x_1 represents the quantity of prey, and x_2 the quantity of predators.

We either adapt general results from [Ben18] or improve some of the ideas to exhibit the conditions on the model so that there is:

(i) *persistence* of the underlying solution to (1), and in particular the existence of a unique invariant probability measure Π called the *persistence measure*. It refers to a probability measure that let invariant the set where none of the species vanishes.

(ii) convergence almost-sure (respectively in total variation) of empirical occupation measure (respectively the law of the process) towards Π, at a polynomial rate, on the set of non-extinction.

(iii) *extinction* almost-sure of some or all species. In particular, new conditions based on unpublished work, which follows from [Ben18], are exhibit to ensure the convergence almost-sure of the process toward an extinction set.

All the refinements are based on our needs to illustrate the usefullness of the tools in the context of our model (1). This promising work is in collaboration with Prof. Michel Benaïm and Dr. Edouard Strickler, and has resulted in the article [BCS25] entitled *"Stochastic persistence and extinction for degenerate stochastic Rosenzweig-MacArthur model"*, published as a preprint in arXiv and currently under review in the *SIAM Journal on Mathematical Analysis*.

However, since a significant part of the work we accomplished is not in the published article, it takes a distinctive place in this thesis on Part A - *Persistence and extinction in ecological & degenerate SDE.*

The initial purpose of this PhD thesis was to address an open problem arising from the work of Prof. Michel Benaïm, Dr. Carl-Erik Gauthier and Dr. Ioana Ciotir on self-repelling diffusions. In their infinite-dimensional framework [BCG15], they established the existence of a unique strong solution to the following SDE defined on the Hilbert space $H := l^2 \times l^2 \times \mathbb{R}$,

$$\begin{cases} \mathrm{d}Y_t = F(Y_t)\mathrm{d}t + \sigma \mathrm{d}W_t, \\ Y_0 = y, \end{cases}$$

where $Y_t = \left((u_t^n)_{n\geq 0}, (v_t^n)_{n\geq 0}, X_t, \right)$, $\sigma = (1, 0, 0)$, $(W_t)_{t\geq 0}$ is a cylindrical Wiener process, and $F : H \to H$, is defined as

$$F \begin{pmatrix} (u^n)_{n\geq 0} \\ (v^n)_{n\geq 0} \\ x \end{pmatrix} = \begin{pmatrix} \left(a_n^{1/2} \cos(nx)\right)_{n\geq 0} \\ \left(a_n^{1/2} \sin(nx)\right)_{n\geq 0} \\ \left\langle \left(a_n^{1/2} n \sin(nx)\right)_{n\geq 0}, (u^n)_{n\geq 0} \right\rangle_{l^2} + \left\langle \left(a_n^{1/2} n \cos(nx)\right)_{n\geq 0}, (v^n)_{n\geq 0} \right\rangle_{l^2} \end{pmatrix}.$$

In particular, they showed that $(Y_t)_{t\geq 0}$ is a Markov process with the Feller property, and it admits an explicit invariant probability based on their finite-dimensional work in [BG17].

However, they left its uniqueness unproven: in fact, a direct counterexample showed that the strong Feller property does not hold in this infinite dimensionsal settings, which is a common tool used to prove unique ergodicity.

To this effect, the *asymptotic strong Feller property* (ASF) was introduced by Hairer and Mattingly in [HM06] and its subsequent refinements suggested a viable route to prove uniqueness. While strong Feller refers to the ability of the underlying semigroup to smooth functions, in the sense that it transforms measurable, bounded functions into continuous, bounded functions, the authors of [HM06] set up this new property to deal with specific models where they lack the strong Feller one, and in particular when the models are degenerate and infinite dimensionsal. One can see the asymptotic strong Feller property as a kind of strong Feller property for large time in the sense that the smoothing property is conserved on the long term. The core of this new property is that they achieved similar conclusions under ASF hypothesis, while being weaker than the original one.

However, despite significant progress, establishing the ASF property for our infinite-dimensional model remains elusive. Our progress is nonetheless interesting in the understanding and application of this recent theory to other degenerate models, I decided to integrate them as a second part of this thesis, on Part B - *Uniqueness of invariant measure in infinite-dimensional & degenerate SDE.*

Keywords: Markov process; stochastic differential equations; ecological model; Rosenzweig-MacArthur; degenerate noise; invariant probability measure; stochastic persistence; Hörmander condition; extinction; almost-sure convergence; rate of convergence; self-interacting diffusion; asymptotic strong Feller property; infinite-dimensional models; asymptotic coupling; unique ergodicity.

Résumé

La première partie de cette thèse se concentre sur l'étude de modèles stochastiques dégénérés représentant typidement des EDS écologiques. Nous considérons d'abord un modèle écologique d'espèces en interaction, en recherchant les conditions de persistance presque sûre, par opposition à l'extinction d'une ou plusieurs espèces.

En particulier, nous nous concentrons sur le modèle classique bidimensionnel Rosenzweig-MacArthur prédateur-proie [RM63] sous une perturbation stochastique dégénérée, dans lequel seule la variable proie est soumise à de légères fluctuations environnementales.

Concrètement, nous analysons le modèle prédateur-proie de Rosenzweig-MacArthur sur $\mathbb{R}_+^2$ avec un bruit dégénéré affectant uniquement la proie :

$$\begin{cases} \mathrm{d}x_1 = x_1 \left(F_1(x_1, x_2)\mathrm{d}t + \varepsilon \mathrm{d}B_t \right), \\ \mathrm{d}x_2 = x_2 F_2(x_1, x_2)\mathrm{d}t, \end{cases} \tag{2}$$

où $\varepsilon > 0$ est suffisamment petit, $(B_t)_{t\geq 0}$ est un mouvement brownien unidimensionnel, x_1 représente la quantité de proies et x_2 celle de prédateurs.

Nous adaptons les résultats généraux de [Ben18] et améliorons certaines des idées pour exposer les conditions du modèle afin qu'il y ait :

(i) *persistance* de la solution sous-jacente à (1), et en particulier l'existence d'une mesure de probabilité invariante unique Π appelée *mesure de persistance*. Elle désigne une mesure de probabilité qui laisse invariant l'ensemble où aucune des espèces ne disparaît.

(ii) convergence presque sûre (respectivement en variation totale) de la mesure d'occupation empirique (respectivement la loi du processus) vers Π, à un taux polynomial, sur l'ensemble des espèces non éteintes.

(iii) *extinction* presque sûre de certaines ou de toutes les espèces. En particulier, de nouvelles conditions basées sur les travaux non publiés qui font suite à [Ben18] sont présentées pour garantir la convergence presque sûre du processus vers un ensemble d'extinction.

Toutes les améliorations sont basées sur nos besoins d'illustrer l'utilité des outils dans le contexte de notre modèle (1). Ce travail prometteur a été mené en collaboration avec le Prof. Michel Benaïm et le Dr. Edouard Strinkler, et a donné lieu à l'article [BCS25] intitulé « Stochastic persistence and extinction for degenerate stochastic Rosenzweig-MacArthur model » (Persistance stochastique et extinction pour le modèle stochastique dégénéré de Rosenzweig-MacArthur), publié sous forme de prépublication dans arXiv et actuellement en cours d'examen dans le *SIAM Journal on Mathematical Analysis*.

Cependant, puisque qu'une partie significative du travail accompli ne se retrouve pas dans l'article publié, elle prend une place distinctive dans cette thèse, sous Partie A - *Persistance et extinction d'EDS écologiques & dégénérées*.

L'objectif initial de cette thèse de doctorat était de traiter un problème ouvert issu des travaux du professeur Michel Benaïm, Dr. Carl-Erik Gauthier et Dr. Ioana Ciotir sur les *self-repelling diffusions* ou diffusions auto-répulsives. Dans leur modèle en dimension infinie [BCG15], les auteurs ont établi l'existence d'une unique solution forte à l'équation différentielle stochastique suivante définie sur l'espace d'Hilbert $H := l^2 \times l^2 \times \mathbb{R}$,

$$\begin{cases} \mathrm{d}Y_t = F(Y_t)\mathrm{d}t + \sigma \mathrm{d}W_t, \\ Y_0 = y, \end{cases}$$

où $Y_t = ((u_t^n)_{n\geq 0}, (v_t^n)_{n\geq 0}, X_t)$, $\sigma = (1,0,0)$, $(W_t)_{t\geq 0}$ est un processus de Wiener cylindrique, et $F : H \to H$, est définie par

$$F\begin{pmatrix}(u^n)_{n\geq 0}\\ (v^n)_{n\geq 0}\\ x\end{pmatrix} = \begin{pmatrix}\left(a_n^{1/2}\cos(nx)\right)_{n\geq 0}\\ \left(a_n^{1/2}\sin(nx)\right)_{n\geq 0}\\ \langle\left(a_n^{1/2} n\sin(nx)\right)_{n\geq 0}, (u^n)_{n\geq 0}\rangle_{l^2} + \langle\left(a_n^{1/2} n\cos(nx)\right)_{n\geq 0}, (v^n)_{n\geq 0}\rangle_{l^2}\end{pmatrix}.$$

En particulier, ils ont montré que $(Y_t)_{t\geq 0}$ est un processus de Markov avec la propriété de Feller, et qu'il admet une probabilité invariante explicite basée sur leurs travaux en dimension finie dans [BG17].

Cependant, ils n'ont pas réussi à prouver son caractère unique: en effet, un contre-exemple direct a montré que la propriété Feller forte ne s'applique pas dans ce cadre à dimension infinie, qui est un outil couramment utilisé pour prouver l'ergodicité unique du semigroupe.

Pour y remédier, la propriété *asymptotiquement fortement Feller* (AFF) introduite par Hairer et Mattingly dans [HM06] et ses améliorations ultérieures ont suggéré une voie viable pour prouver l'unicité. Si la propriété Feller forte fait référence à la capacité du semigroupe sous-jacent à lisser les fonctions, dans le sens qu'il transforme les fonctions mesurables et bornées en fonctions continues et bornées, les auteurs de [HM06] ont présenté cette nouvelle propriété pour traiter des modèles spécifiques qui ne disposent pas de la propriété Feller forte, en particulier lorsque les modèles sont dégénérés et de dimension infinie. On peut considérer la propriété asymptotiquement fortement Feller comme une sorte de propriété Feller forte pour les temps longs, dans le sens où la propriété de lissage est conservée à long terme. Le coeur de cette nouvelle propriété réside dans sa capacité à parvenir à des conclusions similaires sous l'hypothèse AFF, tout en étant plus faible que la propriété originale.

Cependant, malgré des progrès significatifs, l'établissement de la propriété AFF pour notre modèle à dimension infinie demeure difficile à atteindre. Nos progrès sont néanmoins intéressants pour la compréhension et l'application de cette théorie récente à d'autres modèles dégénérés, j'ai donc décidé de les intégrer dans la deuxième partie de cette thèse, sous Partie B - *Unicité de la mesure invariante pour des EDS en dimension infinie & dégénérées*.

Mots-clés: Processus de Markov; équations différentielles stochastiques; modèle écologique; Rosenzweig-MacArthur; bruit dégénéré; mesure de probabilité invariante; persistance stochastique; condition de Hörmander; extinction; convergence presque-sûre; taux de convergence; diffusion auto-interactive; propriété asymptotique fortement Feller; modèles à dimension infinie; couplage asymptotique; ergodicité unique.

Remerciements

Après 5 années de doctorat et pas moins de 10 passées à l'Université de Neuchâtel, au départ comme étudiant paumé sur les bancs des petits salles de cours du bâtiment B, pour me retrouver de "l'autre côté" comme on dit, il est temps de clôturer cette partie de ma vie en apportant les derniers mots de cette thèse à travers quelques remerciements.

Mes premiers remerciements vont évidemment à mon directeur de thèse, Michel Benaïm. Cela a commencé par le plus grand des hasards, un concours de circonstance qui m'a amené à devenir votre "thésard" après avoir effectué un premier travail de Master à distance, coronavirus oblige, dont vous n'en avez peut-être même pas souvenir. J'avais déjà apprécié la liberté que vous m'aviez accordé afin d'étudier des modèles mathématiques appliquées à la finance durant quelques mois, pour finir par me dire qu'*"en fait, je n'y connais pas grand chose aux mathématiques financières"*. Un grand merci également pour votre spontanéité mathématique: une question qui vous passe par la tête, que vous trouvez intéressante sans que je puisse en comprendre le moindre mot, des notes prises à la va-vite et impossible à relire après l'un de nos rendez-vous, puis des jours (semaines?) de travail afin d'en tirer quelques apprentissages que vous arriverez à résumer en quelques mots et de manière si compréhensible lorsque que je viendrai vous présenter mes "avancées". Cela a toujours été un plaisir de pouvoir être l'assistant de vos différents cours de Bachelor et de Master, en m'impliquant dans la préparation de vos cours et de vos exercices. Merci pour cette opportunité que je suis aujourd'hui très heureux d'avoir pu et su saisir: je m'étais toujours dit que *"je ne ferai jamais de doctorat, ce n'est pas pour moi..."* et pourtant.

Je remercie également les membres de mon jury de thèse, en plus de mon directeur: Prof. Christian Mazza, Prof. Felix Schlenk et Dr. Edouard Strickler. Merci pour la chaleureuse discussion qui a suivi ma défense et pour l'intérêt que vous avez porté à mon travail.

Mais à Neuchâtel, on ne se contente pas d'être seulement l'assistant des cours de son directeur: c'est aussi l'occasion de travailler avec d'autres professeur(e)s et je tenais à remercier Alain Valette, avec qui nous avons assuré *feu* le cours de Probability Theory du Master en Statistiques. Je ne m'étais pas préparé à devoir donné un cours en anglais et après 4 assistanats successifs, je crois que j'y ai pris goût. Merci d'être toujours présent et investi, d'avoir fait et de continuer de faire vivre cet institut. Je te remercie également pour la précieuse invitation au *Portiragnes Mathematical Institute* et pour la semaine passée dans ta résidence: si suivre les exposés sur un domaine que je ne connaissais pas du tout et présenter un bout d'article dont je ne comprenais pas la moitié des mots ne furent pas de tout repos, j'en garderai toujours un excellent souvenir.

J'en profite pour remercier les autres professeur(e)s avec qui je n'ai pas eu la chance de collaborer autant mais avec qui j'ai toujours pris énormément de plaisir à échanger à la pause de midi ou autour d'un café au Saloon: Felix Schlenk, Elisa Gorla, Peter Feller et Jérémy Blanc. Mais aussi à ceux qui ont quitté l'Institut entre temps, Bruno Colbois et Paul Jolissaint.

Mais c'est bien avec les autres doctorant(e)s et post-doctorant(e)s que nous passons le plus de temps et créons le plus de souvenirs.

Mes premiers mots vont évidemment à Antoine, mon grand frère de thèse: merci pour ces discussions (y compris HORS maths quand on en pouvait plus, c'est-à-dire souvent) qui font du bien, pour les encouragements, pour l'aide et pour ton dévouement de tous les instants pour l'Institut. Je n'ai jamais vu un mec passer autant de temps en pause café alors qu'il n'en boit jamais, ni en pause clope alors qu'il ne fume pas !

En parlant de pause, c'est un bon moment pour remercier Alexandre, dit Bud, pour ces longues années passées à l'Institut et les nombreuses années à venir en dehors: partenaire rugbystique et de beuverie, grand amateur de pause et de café, un ami qui m'a fait l'honneur d'être le maître de cérémonie de son mariage. Merci pour tout, pour tes encouragements comme pour ta torture à quelques jours de

passer l'examen de systèmes dynamiques aléatoires. Un humour à toute épreuve, sauf de mon stress !

Heureusement que Tobias avait finalement été l'expert de cet examen ! Nous n'avons malheureusement pas eu le temps de collaborer ensemble, tu es parti continuer ton brillant parcours académique ailleurs. Cela a toujours été un plaisir d'échanger avec toi, de maths comme d'autres sujets. Et tu es un modèle pour tout ce qui touche à la présentation des mathématiques: toujours clair, on ressent que tu es passionné et que tu souhaites le partager avec ton auditoire.

J'en profite pour remercier également Edouard, (arrière?) grand frère de thèse qui m'a ouvert aux conférences en sortie de pandémie: tu as bien fait d'insister et d'obliger Antoine et moi-même à venir poursuivre ces mémorables Journées de probabilités au fin fond de l'Alsace. C'est un plaisir d'avoir pu collaborer avec Michel et toi-même sur cet article, en espérant que certaines bêtises que j'ai pu écrire ne t'aient pas brûler les yeux ! Tu as toujours trouvé le temps de m'aider, malgré une vie de chercheur et de papa bien rempli: je te souhaite plein de bonheur.

Même si nous n'avons pas eu le temps de beaucoup apprendre à se connaître mathématiquement après ton départ pour le terrible monde de la finance, un grand merci à William: un plaisir d'avoir pu partager ce bureau avec toi, quelques trajets entre La Tchaux et l'Uni en voiture le tout accompagné de bon rap français, et pour ces belles soirées au Coyote Bar.

Pour finir avec les probabilistes, merci aux récents arrivants Léo, Lucas et Pierre: en bon chercheur passionné, Léo, tu n'as pas hésité à venir avec tes idées et tes questions pour mieux comprendre ce que j'essayais de faire, chercher comment cela pourrait s'appliquer à tes propres problèmes et à mettre la main à la pâte ! Lucas, même si l'on a moins collaboré mathématiquement, cela fut un vrai plaisir de pouvoir parler de tout, de l'avenir (de notre futur startup, je ne l'ai pas oublié), de t'apprendre quelques règles du rugby et de boire quelques verres au Club-House. Je reviendrai bencher avec toi à la rentrée et avec Pierre évidemment.

Pierre, désolé de t'abandonner dans la jungle du doctorat alors que tu vas tout juste commencer cette aventure, mais je n'ai aucun doute que tu y arriveras avec panache. Je ne te suivrai pas sur un semi-marathon tout de suite, mais avec plaisir pour te voir au bord du terrain de rugby (peut-être sur le terrain un jour). Je suis fier d'être à mon tour grand frère de thèse et si tu as besoin de moi, tu sais où me trouver.

Il y a tellement d'autres (post-)doctorant(e)s à remercier ici: Laura, merci pour être si honnête et sans filtre. Il y a eu quelques discussions mémorables et animées ! Je me souviendrai aussi longtemps de cette soirée pour fêter ton doctorat, à 5h du matin au bord du Lac de Neuchâtel et une $n-$ième bouteille à finir. Merci également à Laurent, le plus musclé des mathématiciens que je connaisse (et le plus bel accent du gros de Vaud, aussi). Merci d'avoir quand même attendu que je me perde lors de mon célèbre exposé à Portiragnes pour me dire *"Ah mais j'aurai du te donner mon chapitre de thèse, ça explique tout !"*. Alors que t'aurais pu me faire *un petit audio, parce que ça sera plus simple*.

Merci Jannaud de m'avoir fait découvrir toujours plus de fromages et à Thomas pour le partage des expériences et bons-plans ferroviaires, en Suisse et ailleurs. Sans oublier Léonard: dire qu'on a commencé nos études en 2015 sans savoir où l'on avait mis les pieds pour finir en tant que doctorants à quelques bureaux l'un de l'autre. Merci d'avoir été là jusqu'à Portiragnes, pour que l'on partage notre désespoir mathématique du moment ainsi que pour ce long trajet en voiture que tu as assuré.

Je glisse également un mot à mes anciens camarades de Master: merci Emma d'avoir toujours été là pour m'encourager et m'accueillir dans ton bureau à Zürich. Dire que je me souviens encore de la galère de nos cours d'informatique au Lycée, tu as réussi avec brillo ton doctorat en science de R. Merci Antoine (l'autre) pour ta bonne humeur et tout ce temps que tu as passé à nous expliquer ce que l'on mettait 10x plus de temps à assimiler que toi. Bon courage pour la suite, je sais que tu y es presque et que tu deviendras facilement docteur. Je remercie également Alexis pour être toujours disponible à me trimballer en voiture lors de nos soirées de matheux, à Alexandre (le jurassien) pour ton accueil légendaire à Porrentruy et à Maël, que je suis sûr de recroiser en coup de vent dans tous les festivals du coin.

Il n'y a pas eu que les maths dans ma vie durant ces 5 dernières années et j'adresse un grand merci à

Xavier et Timothée, mes partenaires lors de mes aventures entrepreunariales à côté de ce doctorat. Merci d'avoir essayé de comprendre ce que je faisais, de m'encourager lorsque le temps me manquait pour tout faire et pour vous adapter à mes incessants changements d'horaire. Même si l'aventure que nous avions commencé il y a bien des années a pris un mauvais tournant ces derniers mois, je me réjouis déjà de la prochaine que nous lancerons. Merci pour l'accueil à Lausanne, à Berlin, pour les sessions workshops dans la grande maison des Eplatures-Grises ou pour un bain de minuit à Saint-Imier.

Une pensée à Jenny, merci de prendre toujours du temps pour faire un crochet par La Tchaux lorsque tu reviens de Paris. C'est bon, j'ai enfin fini mes études, t'as vu !

Je ne pourrai pas faire ces remerciements sans citer mes parents. Merci papa, merci maman de m'avoir toujours suivi dans mes idées et de m'avoir fait confiance toutes ces années. Merci pour votre compréhension lors de mes sessions d'examen, merci de m'avoir donné tout ce qu'il fallait pour que je puisse poursuivre ces (longues) études sans avoir à m'inquiéter de quoique ce soit, merci de toujours essayer de vous intéresser à ce que je fais malgré mes explications douteuses, merci de m'avoir toujours soutenu.

Merci également à mon petit frère, Quentin: le pull blanc que tu m'as offert fait toujours son effet à l'Institut.

Je remercie également mon coloc depuis 6 ans, Jimmy. Merci de m'écouter râler après une journée de m****, de toujours mettre de la musique de m**** pour assurer l'ambiance à l'appart', de me faire découvrir des jeux vidéos "insérer bruits de pistolet" pour se détendre et de passer toute la nuit à chasser les Pals ou vendre notre fameuse OG. Et merci de m'avoir initié au piment que je tank enfin !

Merci également à la merveilleuse bande de potes que j'ai la chance d'avoir à mes côtés: merci les bros de la *Bromance* d'être sûr de toujours trouver un compagnon de beuverie en fin de semaine (et tous les autres jours, en fait). Merci également aux potes de l'équipe de rugby, il n'y a pas meilleur environnement pour se changer les idées, sur ou en dehors du terrain.

Enfin, je tiens à adresser un grand merci à ma partenaire d'étude, de voyage et de vie depuis 10 ans: on a presque fait nos premiers pas à l'Université ensemble, tu avais une année d'avance mais je t'ai rattrapé ! En fin de compte, je garde presque un bon souvenir de ces sessions de révision passées avec toi. Merci mon amour pour toutes ces années passées et à venir, merci de m'écouter raconter mes problèmes de maths en essayant de me trouver des solutions, merci pour tes encouragements dans les moments de doute, pour tes félicitations dès qu'il y avait de quoi fêter quelque chose, pour tous ces moments où tu m'as rassuré, tu m'as changé les idées. Tu es un modèle de persévérance pour moi, tu n'as jamais rien lâché, tu as surmonté tes plus grandes peurs pour en arriver là où tu voulais. À toi de finaliser ta thèse: je n'ai aucun doute sur tes capacités à en faire quelque chose d'exceptionnelle et je serai à mon tour là pour toi.

Contents

Part A
Persistence and extinction in ecological & degenerate SDEs

Here, we investigate an important question arising in mathematical ecology: under which conditions a population of several groups of interacting species may coexist over a long period of time or, in contrast, one or multiple (or all) species may disappear. These biological situations are often mathematically modeled using differential equations: one can think of the Lotka–Volterra predator–prey model, originally studied in [Lot10], where first-order nonlinear differential equations described the dynamics of two species interacting with each other as a prey and a predator.

To take into account the effect of ecological fluctuations such as temperature variation or changing precipitation, it may be more relevant to consider *stochastic differential equations* as early proposed in [Tur77]. More recently, the survey paper [Sch17] or [BS19] studied the notion of persistence and extinction in stochastic ecological models by introducing a wide variety of examples.

Specifically, we consider a set of N species interacting with each other and denote by $x_1, \cdots, x_N$ their population density. Each of them is impacted by their own effect as well as the effects of other species, denoted $F_i(x_1, \cdots, x_N)$, and some environmental fluctuations, modeled with an $m-$dimensional standard Brownian motion$(B_t^1, \cdots, B_t^m)_{t\geq 0}$ whose impact differs for each species with a factor $\Sigma_i(x_1, \cdots, x_N)$. Those models are called *Kolmogorov stochastic differential equations* and are described by

$$dx_i(t) = x_i(t)\left(F_i(x(t))\mathrm{d}t + \sum_{j=1}^{m} \Sigma_i^j(x(t))\mathrm{d}B_t^j\right), \quad i = 1, \cdots, N,$$

where $x = (x_1, \cdots, x_N)$ is the density vector of the population, and $t \geq 0$ the time variable.

In particular, an interesting and challenging mathematical problems concerns the situation where the environmental fluctuations do not affect all species. Since this additional randomness tends to smooth the global behavior, its absence causes mathematical troubles to study: such models are said to be *degenerate* and they are at the core of this thesis.

The purpose of this work started with the goal to study the behavior of the classical two-dimensional Rosenzweig–MacArthur prey–predator model [RM63], where x_1 (respectively x_2) denotes prey density (respectively predator density), in which only the prey variable x_1 is subject to environmental fluctuations modeled by $\varepsilon\mathrm{d}B_t$. More precisely, the model is given by the following the degenerate 2-dimensional SDE on $\mathbb{R}_+^2$ in (1) with the corresponding drifts

$$F_1(x) = 1 - \frac{x_1}{\kappa} - \frac{x_2}{1+x_1}, \quad F_2(x) = -\alpha + \frac{x_1}{1+x_1}, \quad \alpha, \kappa > 0,$$

and the detailed model reads

$$\begin{cases} \mathrm{d}x_1 = x_1\left[\left(1 - \frac{x_1}{\kappa} - \frac{x_2}{1+x_1}\right)\mathrm{d}t + \varepsilon\mathrm{d}B_t\right], \\ \mathrm{d}x_2 = x_2\left(-\alpha + \frac{x_1}{1+x_1}\right)\mathrm{d}t. \end{cases} \tag{3}$$

where $\kappa, \varepsilon > 0$, $0 < \alpha < 1$, and $(B_t)_{t\geq 0}$ is a standard one-dimensional Brownian motion. Remark that if $\alpha > 1$, $x_2(t)$ goes to 0 almost surely as $t \to \infty$ since $\dot{x}_2(t) \leq x_2(t)(-\alpha + 1)$.

In case of absence of noise, i.e. $\varepsilon = 0$, the behavior of (3) is well-known (see e.g. [Smi08]):

- Every trajectory starting from $x_1 = 0$ converges to the origin.

- If $\alpha < \frac{\kappa}{\kappa+1}$, the ODE admits a unique equilibrium $p = \left(\frac{\alpha}{1-\alpha}, \frac{\kappa-(\kappa+1)\alpha}{\kappa(1-\alpha)^2}\right)$.

  In addition, if $\alpha < \frac{\kappa-1}{1+\kappa}$, p is a source and there exists a limit cycle $\gamma \subset \mathrm{int}(\mathbb{R}_+^2)$ surrounding p and every solution starting from $\mathrm{int}(\mathbb{R}_+^2) \setminus \{p\}$ has γ as its limit set.

- Otherwise, if $\alpha \geq \frac{\kappa-1}{1+\kappa}$, every solution starting from $\mathrm{int}(\mathbb{R}_+^2)$ converges to p.

For $0 < \varepsilon^2 < 2$, we define $k := \frac{2}{\varepsilon^2} - 1 > 0$, $\theta := \frac{\varepsilon^2\kappa}{2} < \kappa$, and

$$\gamma_{\varepsilon,\kappa}(x) = \frac{x^{k-1}e^{-x/\theta}}{\Gamma(k)\theta^k}, \quad x \geq 0,$$

which is the density of a Γ-distribution with parameters k, θ whose expectation is $k\theta = \kappa(1 - \frac{\varepsilon^2}{2})$ and variance $k\theta^2 = \frac{\kappa^2\varepsilon^2}{2}(1 - \frac{\varepsilon^2}{2})$. We also define

$$\Lambda(\varepsilon, \alpha, \kappa) = \int_0^{+\infty} \frac{x}{1+x}\gamma_{\varepsilon,\kappa}(x)dx - \alpha.$$

In particular, we can upper bound $\Lambda(\varepsilon, \alpha, \kappa)$ using Jensen's inequality applied to the concave function $\varphi(x) = \frac{x}{1+x}$ so that

$$\Lambda(\varepsilon, \alpha, \kappa) \leq \frac{\kappa(1 - \frac{\varepsilon^2}{2})}{1 + \kappa(1 - \frac{\varepsilon^2}{2})} - \alpha,$$

and lower bound it by Cauchy-Schwarz inequality which leads to

$$\Lambda(\varepsilon, \alpha, \kappa) \geq \frac{\kappa(1 - \frac{\varepsilon^2}{2})}{1 + \kappa(1 - \frac{\varepsilon^2}{2})} - k\frac{\varepsilon}{\sqrt{2}}\sqrt{1 - \frac{\varepsilon^2}{2}} - \alpha.$$

In particular, this concrete and complex underlying model (3) was chosen to present the usefulness of the tools developed in [Ben18] in a broad context. To this effect, some of the original general results have been either adapted to our situation, or improved with our more recent ideas.

Our interest in the Rosenzweig-MacArthur model (3) stems from the fact that it represents a degenerate SDE evolving on a non-compact state space. We will outline the conditions ensuring that degenerate models of the form of (3) reflect typical ergodic properties. Specifically for the Rosenzweig-MacArthur model (3), main properties can be summarized as follows:

(i) If $0 < \varepsilon^2 < 2$ and $\Lambda(\varepsilon, \alpha, \kappa) > 0$, $(x_1(t), x_2(t))_{t\geq 0}$ is *stochastically persistent* in the sense that the law of $(X_t^x)_{t\geq 0}$ converges almost surely to a unique invariant probability measure, supported on the interior of $\mathbb{R}_+^2$ and with a smooth density with respect to the Lebesgue measure, whenever $x = (x_1(0), x_2(0)) \in \mathrm{int}(\mathbb{R}_+^2)$ is the initial condition of the system.

(ii) If $0 < \varepsilon^2 < 2$ and $\Lambda(\varepsilon, \alpha, \kappa) < 0$, $(X_t^x)_{t\geq 0}$ is *stochastically nonpersistent* with respect to $\mathbb{R}_+^* \times \{0\}$, in the sense that $x_2(t)$ converges almost surely to 0.

(iii) If $\varepsilon^2 > 2$, $(X_t^x)_{t\geq 0}$ is *stochastically nonpersistent* with respect to $\{0\} \times \{0\}$ in the sense that $x_1(t)$ and $x_2(t)$ converge almost surely to 0.

The specific case $\Lambda(\varepsilon, \alpha, \kappa) = 0$ has been investigated in [NS20] as a critical case for (3) where the authors proved that the process goes on average to extinction.

To this effect, we state multiple hypotheses that will lead to those conclusions:

(i) Hypothesis 1-4 to show stochastic persistence, uniqueness of the invariant probability, and the convergence almost-sure and in Total variation;

(ii) Hypothesis 6 to exhibit a polynomial convergence in Total variation, in the specific case of non-compact extinction sets. We also detail how we lack an exponential rate of convergence with our framework while Hypothesis 5 is not verified for (3);

(iii) Hypotheses 7 and 8 to show the stochastic nonpersistence of the process and its relationship with almost-sure extinction.

Structure of this part

Chapter 1 focuses on defining the main tools used during the whole part, and how they will help us to reach our conclusion. In particular, we state Hypotheses 1-3 to ensure that the empirical measure is almost surely tight and each limit point is an invariant probability. To this effect, we will only focus on Kolmogorov stochastic differential equations and give a practical criterion in Proposition 2.1 ensuring that Hypotheses 1-3 are verified.

Chapter 2 introduces the notion of *stochastic persistence*, which will lead to the uniqueness of the invariant probability together with some convergence results under an additional Hypothesis 4. Under another stronger Hypothesis 5, the convergence rate will be proved to be exponential while under Hypothesis 6, the convergence rate is polynomial. Finally, we introduce the notion of *stochastic nonpersistence* as well as Hypotheses 7 and 8 that will be crucial to show the almost-sure extinction of some or all species. As in Chapter 1, we present a practical criterion to ensure that each Hypothesis holds true in the specific case of Kolmogorov SDE.

Chapter 3 presents two nondegenerate models that have to be considered as toy models: a one-dimensional logistic SDE and a two-dimensional nondegenerate SDE. They enable us to justify how to achieve the above criteria and also justify their relative accessibility. Some of the computations are useful too, since they will be re-used later on our core problem.

Chapter 4 presents the new results obtained on the Rosenzweig-MacArthur model in details. In particular, the choice of the parameters in (3) are crucial: indeed, assuming $\varepsilon^2 < 2$ and $\Lambda(\varepsilon, \alpha, \kappa) > 0$, Theorem 4.5 implies to stochastic persistence while when $\varepsilon^2 < 2$ and $\Lambda(\varepsilon, \alpha, \kappa) < 0$, Theorem 4.13 shows the extinction of predator species x_2. If we suppose $\varepsilon^2 > 2$, Theorem 4.17 leads to the extinction of both species.

Chapter 1

Introduction

We introduce here the general context and fix some notations that will be crucial in the first part of this thesis. It is inspired by the work of Prof. Michel Benaïm in his article *Stochastic persistence* [Ben18] together with the additional theory needed to achieve the improvements we discovered. Part of these notes is also inspired by the unpublished work of Prof. Michel Benaïm, Prof. Alexandru Hening, Prof. Dang Nguyen, Prof. Sebastian Schreiber and Dr. Edouard Strickler in [BHN+25]. More precisely, Chapter 1 goes as follows:

(i) Section 1.1 presents the mathematical background required for the following of the text by fixing some context and notations. You will also find the main standing assumption, Hypothesis 1.

(ii) Section 1.2 recalls a general theory about the *infinitesimal generator* associated to a Markov semigroup and its underlying *carré du champ* operator. In particular, we introduce extended versions of both tools to get rid of the limitations to bounded functions.

(iii) Section 1.3 recalls the notions of invariant and ergodic probability measures. In particular, we will focus on the *(mean) empirical measure* and how it is related to the existence, hence uniqueness of the invariant probability with the help of Lyapunov functions through Hypotheses 2 and 3.

1.1 General theory about Markov processes

Let (M, d) be a locally compact Polish space, which is a complete, separable and metrizable space, equipped with its Borel σ-algebra $\mathcal{B}(M)$. Since we focus on population dynamics, we will mainly consider $\mathbb{R}_+^n$ equipped with the usual Euclidean metric.

Let $(B(M), \|\cdot\|)$ be the Banach space of all real-valued, bounded, measurable functions on M endowed with the sup-norm $\|f\|_\infty = \sup_{x\in M} |f(x)|$.

We denote by $C_b(M)$ the Banach subspace of bounded continuous functions on M, and $C_0(M)$ for continuous functions vanishing at infinity.

Let $\mathcal{P}(M)$ be the space of probability measures on $(M, \mathcal{B}(M))$, equipped with the topology of weak convergence. For $\mu \in \mathcal{P}(M)$ and $f \in B(M)$, we set

$$\mu f := \int_M f(x)\mu(\mathrm{d}x).$$

Recall that a sequence $(\mu_n)_{n\geq 1} \subset \mathcal{P}(M)$ converges *weakly* to $\mu \in \mathcal{P}(M)$, written $\mu_n \Rightarrow \mu$, if

$$\mu_n f \to \mu f \quad \text{for all } f \in C_b(M).$$

For two probability measures $\alpha, \beta \in \mathcal{P}(M)$, we define the *Total variation distance* between them by

$$\|\alpha - \beta\|_{\mathrm{TV}} := \sup_{\|f\|_\infty \leq 1} |\alpha f - \beta f|.$$

From now on, we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, together with a complete, right-continuous filtration $(\mathcal{F}_t)_{t\geq 0}$, and a family of càdlàg (right-continuous with left limits) Markov processes $\{(X_t^x)_{t\geq 0} : x \in M\}$ adapted to this filtration. More precisely, recall that:

(i) For every $x \in M$, the random variable X_t^x is M-valued and $\mathcal{F}_t$-measurable, with $X_0^x = x$ $\mathbb{P}$-a.s., and the trajectory $t \mapsto X_t^x$ is càdlàg.

(ii) For each $f \in B(M)$, the mapping

$$(t, x) \in \mathbb{R}_+ \times M \longmapsto P_t f(x) := \mathbb{E}[f(X_t^x)]$$

is measurable, and the *Markov property* holds:

$$\mathbb{E}[f(X_{t+s}^x)|\mathcal{F}_t] = (P_s f)(X_t^x), \quad \mathbb{P}\text{-a.s.} \tag{4}$$

In addition, recall that the a process $(X_t)_{t\geq 0}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is *progressively measurable* if for every $t \geq 0$, the mapping

$$(s, \omega) \mapsto X_s(\omega), \quad \text{with } (s, \omega) \in [0, t] \times \Omega,$$

is measurable with respect to the product $\sigma-$algebra $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$.

We say that $(X_t)_{t\geq 0}$ is a *predictable process* if it is measurable with respect to the predictable $\sigma-$algebra generated by all left-continuous, adapted processes or equivalently by the sets of the form $]s, t[\times A$, where $A \in \mathcal{F}_s$, $\forall s \leq t$. The relationship between those properties can be summarized as follows:

$$\text{Predictable} \Rightarrow \text{Progressively measurable} \Rightarrow \text{Adapted},$$

but the converse does not always hold. However:

Proposition 1.1 ([KS98], Proposition 1.13)**.** *A right-continuous (or left-continuous) stochastic process adapted to $(\mathcal{F}_t)_{t\geq 0}$ is also progressively measurable with respect to $(\mathcal{F}_t)_{t\geq 0}$.*

Also, one can remark that any left-continuous adapted process is predictable, by definition of a predictable process.

The Markov property (4) of $(X_t)_{t\geq 0}$ implies the well-known *Chapman-Kolmogorov* equations:

$$\begin{aligned} &P_{t+s} = P_t \circ P_s = P_s \circ P_t, \\ &P_0 = Id. \end{aligned}$$

Then, the operators $(P_t)_{t\geq 0}$ defined above form a semigroup of contractions on $\mathcal{B}(M)$: $P_t(P_s f) = P_{t+s} f$ and $\|P_t f\|_\infty \leq \|f\|_\infty$ for all $f \in B(M)$.

We can also see $(P_t)_{t\geq 0}$ as a Markov kernel: for a set $A \subset M$, we write $\mathbf{1}_A$ for its indicator function. Then, for any $x \in M$ and $A \in \mathcal{B}(M)$, we define

$$P_t(x, A) = P_t \mathbf{1}_A(x) = \mathbb{P}_x(X_t \in A).$$

For $\mu \in \mathcal{P}(M)$, we let $\mathbb{P}_\mu$ be the law of the process $(X_t)_{t\geq 0}$ on the Skorokhod space $\mathbb{D}(\mathbb{R}_+, M)$ of càdlàg functions, under the initial condition μ, and $\mathbb{E}_\mu$ for its associated expected value. In particular, for the Dirac measure in one point δ_x, we write more commonly $\mathbb{P}_x$ and $\mathbb{E}_x$. Similarly, we denote by X^μ or X^x the process under the initial condition μ or x.

Definition 1.2 ($C_b(M)-$**Feller property**)**.** *A process $(X_t)_{t\geq 0}$ (or equivalently its semigroup $(P_t)_{t\geq 0}$) is said to $C_b(M)$-Feller if*

$$P_t\left(C_b(M)\right) \subset C_b(M), \quad \forall t \geq 0,$$

and

$$\lim_{t\downarrow 0} P_t f(x) = f(x), \quad \forall f \in C_b(M), \ \forall x \in M. \tag{5}$$

Remark 1.3. The term $C_b(M)$-Feller is used here to avoid confusion with the classical notion of a Feller semigroup, which requires that $(P_t)_{t\geq 0}$ is strongly continuous on $C_0(M)$ or, equivalently (see Proposition III.2.4, [RY04]), that the inclusion above and the limit property hold with $C_b(M)$ replaced by $C_0(M)$.When M is compact, the two definitions coincide; when M is non-compact, every Feller semigroup is $C_b(M)$-Feller (see e.g. Lemma 1.2 in [Str19]), but not conversely as it will be show through the one-dimensional logistic SDE (see Section 3.1 and Remark 3.1).

Also, we observe that by right-continuity of paths, condition (5) is automatically satisfied.

Our main and standing assumption is the following:

Hypothesis 1 (Standing assumption, Extinction set). There exists a closed set $M_0 \subset M$, called the *extinction set* of the semigroup $(P_t)_{t\geq 0}$, which is invariant under $(P_t)_{t\geq 0}$ in the sense that

$$P_t \mathbf{1}_{M_0} = \mathbf{1}_{M_0}, \quad \forall t \geq 0.$$

Define the *non-extinction set* by $M_+ = M \setminus M_0$. Note that M_+ is open and invariant, i.e.

$$P_t \mathbf{1}_{M_+} = \mathbf{1}_{M_+}, \quad t \geq 0.$$

Remark 1.4. The (non-)extinction set is indeed crucial to define the notion of persistence and hence extinction. As their names say, the persistence of the process $(X_t^x)_{t\geq 0}$ occurs when, for any initial $x \in M_+$, the process stays in M_+ with probability one while the extinction is defined when $(X_t^x)_{t\geq 0}$ goes almost surely to M_0 whenever it started in M_+.

Definition 1.5 (Accessibility). *A point $y \in M$ is* accessible *from $x \in M$ if, for every neighborhood U of y, there exists $t \geq 0$ with $P_t(x, U) > 0$. We denote by Γ_x the set of points accessible from x. For $A \subset M$ set $\Gamma_A = \cap_{x\in A}\Gamma_x$.*

Let G be the *1–resolvent kernel* (a discrete–time Markov kernel) associated to $(P_t)_{t\geq 0}$, defined as

$$Gf := \int_0^\infty e^{-t} P_t f \mathrm{d}t, \quad \forall f \in B(M). \tag{6}$$

It can be seen as the transition kernel of the discrete time chain $(Y_n)_{n\in\mathbb{N}}$ obtained by sampling $(X_t)_{t\geq 0}$ at random times exponentially distributed, in the sense that

$$Y_n^x = X_{T_n}^x, \tag{7}$$

where $T_0 = 0, \cdots, T_{n+1} = T_n + U_{n+1}$ and $U_1, U_2, \cdots$ is a sequence of independent identically distributed random variables having an exponential distribution with parameter 1 and independent of $(X_t^x)_{t\geq 0}$.

Then, accessibility is also recoverable through the 1–resolvent kernel G (see e.g Section 5.2.1 and Proposition 5.19 in [BH22]) as

$$\Gamma_x = \mathrm{supp}(G(x, \cdot)).$$

Now, we can state the core notions of (weak) Doeblin point:

Definition 1.6 (Weak Doeblin point). *A point $x^* \in M$ is a* weak Doeblin point *if there exist a neighborhood U of x^*, a non-zero measure ξ on M such that, for all $x \in U$, the smoothed kernel*

$$G(x, \cdot) = \int_0^\infty e^{-t} P_t(x, \cdot)\mathrm{d}t \geq \xi(\cdot).$$

Equivalently, U is a petite set in the sense of Meyn–Tweedie (see e.g. [MT09], Section 5.5).

Definition 1.7 (**Doeblin point**). *A point $x^* \in M$ is a* Doeblin point *if there exist a neighborhood U of x^*, a non-zero measure ξ on M and a time $t_* > 0$ such that*

$$P_{t_*}(x,\cdot) \geq \xi(\cdot), \quad x \in U. \tag{8}$$

In Meyn–Tweedie terminology (see e.g. [MT09], Section 5.2), U is a small set.

Remark 1.8. If x^* is an accessible Doeblin point, the minorization condition (8) extends to every compact set (see Lemma 4.9 in [Ben18]).

1.2 Generator and carré du champ

We now define the *infinitesimal generator* of $(P_t)_{t\geq 0}$ and its subsequent *domain*. It is a key tool to show the unique ergodicity of a Markov semigroup, as will highlighted in the corresponding Section 1.3 about invariant and ergodic probability measures.

Definition 1.9 (**Generator**). *The* generator *of $(P_t)_{t\geq 0}$ is the linear operator $\mathcal{L}\colon \mathcal{D} \to C_b(M)$, where the* domain *$\mathcal{D} \subset C_b(M)$ is the set of all $f \in C_b(M)$ satisfying*

(i) *the limit $\mathcal{L}f(x) := \lim_{t\downarrow 0} \frac{P_t f(x) - f(x)}{t}$ exists for every $x \in M$;*

(ii) *$\mathcal{L}f \in C_b(M)$;*

(iii) $\sup_{0<t\leq 1} \frac{1}{t}\|P_t f - f\|_\infty < \infty$.

Remark 1.10. Under the assumption that $(P_t)_{t\geq 0}$ is a $C_b(M)$−Feller semigroup, let $f \in C_b(M)$ and $\varepsilon > 0$. We set

$$\overline{f}_\varepsilon := \frac{1}{\varepsilon}\int_0^\varepsilon P_s f \mathrm{d}s.$$

Then $\overline{f}_\varepsilon \in \mathcal{D}$, $\mathcal{L}\overline{f}_\varepsilon = \frac{1}{\varepsilon}(P_\varepsilon f - f)$, and $\overline{f}_\varepsilon \to f$ pointwise as $\varepsilon \downarrow 0$. In particular, $\mathcal{D}$ is non-empty since it contains all functions $\overline{f}_\varepsilon$. Indeed, we can write

$$\begin{aligned}
\mathcal{L}\overline{f}_\varepsilon(x) &= \lim_{t\downarrow 0} \frac{P_t\overline{f}_\varepsilon(x) - \overline{f}_\varepsilon(x)}{t} \\
&= \lim_{t\downarrow 0} \frac{1}{t\varepsilon}\left(\int_0^\varepsilon P_{s+t}f(x)ds - \int_0^\varepsilon P_s f(x)ds\right) \\
&= \lim_{t\downarrow 0} \frac{1}{t\varepsilon}\left(\int_t^{\varepsilon+t} P_u f(x)du - \int_0^\varepsilon P_s f(x)ds\right) \\
&= \lim_{t\downarrow 0} \frac{1}{t\varepsilon}\left(\int_\varepsilon^{\varepsilon+t} P_u f(x)du - \int_0^t P_s f(x)ds\right),
\end{aligned}$$

where the last equality is obtained adding $0 = \int_\varepsilon^t P_u f(x)du + \int_t^\varepsilon P_s f(x)ds$. We can rewrite the first integral such as

$$\begin{aligned}
\int_\varepsilon^{\varepsilon+t} P_u f(x)du &= \int_\varepsilon^{\varepsilon+t} P_\varepsilon f(x)du + \int_\varepsilon^{\varepsilon+t} \left(P_u f(x) - P_\varepsilon f(x)\right)du \\
&= tP_\varepsilon f(x) + \int_\varepsilon^{\varepsilon+t} \left(P_u f(x) - P_\varepsilon f(x)\right)du,
\end{aligned}$$

and it follows that

$$\mathcal{L}\overline{f}_\varepsilon(x) = \lim_{t\downarrow 0} \frac{1}{t\varepsilon}\left(tP_\varepsilon f(x) + \int_\varepsilon^{\varepsilon+t} \left(P_u f(x) - P_\varepsilon f(x)\right)du - \int_0^t P_s f(x)ds\right).$$

By continuity of $u \mapsto P_u f(x)$, the second integral is of order $O(t)$ such that $\lim_{t\to 0} \frac{O(t)}{t} = 0$. By continuity, the last integral converges to

$$\begin{aligned}
\lim_{t\downarrow 0} \frac{1}{t}\int_0^t P_s f(x) ds &= \lim_{t\downarrow 0} \frac{1}{t}\left(\int_0^t f(x) ds - \int_0^t (P_s f(x) - f(x)) ds\right)\\
&= \lim_{t\downarrow 0} \frac{1}{t}\left(t f(x) - \int_0^t (P_s f(x) - f(x)) ds\right)\\
&= f(x), \quad \forall x \in M,
\end{aligned}$$

We conclude that

$$\mathcal{L}\overline{f}_\varepsilon(x) = \frac{1}{\varepsilon}\left(P_\varepsilon f(x) - f(x)\right), \quad \forall x \in M,$$

which is a bounded, continuous function over M, and $\overline{f}_\varepsilon \in \mathcal{D}$ since we also have

$$\begin{aligned}
\sup_{0<t\le 1} \frac{1}{t}\left\|P_t\overline{f}_\varepsilon - \overline{f}_\varepsilon\right\|_\infty &= \sup_{0<t\le 1} \frac{1}{t\varepsilon}\left\|\left(\int_\varepsilon^{\varepsilon+t} P_u f(x) du - \int_0^t P_s f(x) ds\right)\right\|_\infty\\
&\le \sup_{0<t\le 1} \frac{2t\|f\|_\infty}{t\varepsilon}\\
&\le \frac{2\|f\|_\infty}{\varepsilon}\\
&< \infty,
\end{aligned}$$

uniformly on t, since f is supposed to bounded. Finally, we can see that $\lim_{\varepsilon\downarrow 0} \overline{f}_\varepsilon(x) = f(x)$ for all $x \in M$ by using the same construction as above and

$$\lim_{\varepsilon\downarrow 0} \frac{1}{\varepsilon}\int_0^\varepsilon P_s f(x) ds = f(x), \quad \forall x \in M$$

From the definition above, we can see $\mathcal{L}$ as the derivative of $P_t f(x)$ for time $t = 0$ but it indeed describes the derivatives of $P_t f(x)$ at every time t, as summarized in the next proposition: (9) parallels Proposition 3.3 of [PR11] while (10) is a classical result from infinitesimal generator theory.

Proposition 1.11. *Let $f \in \mathcal{D}$. Then for every $t \ge 0$ one has $P_t f \in \mathcal{D}$, the mapping $t \mapsto P_t f(x)$ is C^1 for each $x \in M$ and*

$$\frac{\partial}{\partial t} P_t f(x) = \mathcal{L}(P_t f)(x) = P_t(\mathcal{L}f)(x). \tag{9}$$

Moreover, the process

$$M_t^f(x) := f(X_t^x) - f(x) - \int_0^t \mathcal{L}f(X_s^x) \mathrm{d}s, \quad t \ge 0, \tag{10}$$

is a càdlàg martingale with respect to $(\mathcal{F}_t)_{t\ge 0}$ and $\mathbb{P}_x$.

Remark 1.12. Formula (10) can also be viewed as a consequence of *Dynkin's formula* combined with the Markov property of $(X_t)_{t\ge 0}$ (see e.g. Proposition 14.10 and 14.13 in [Dav93].

Proof. Let $f \in \mathcal{D}$. We set

$$\Delta_\varepsilon f = \frac{P_\varepsilon f - f}{\varepsilon}, \quad 0 < \varepsilon \le 1.$$

Then $\sup_{0<\varepsilon\le 1} \|\Delta_\varepsilon P_t f\|_\infty \le \sup_{0<\varepsilon\le 1} \|\Delta_\varepsilon f\|_\infty < \infty$, and $\Delta_\varepsilon P_t f = P_t \Delta_\varepsilon f$, as we did in Remark 1.10. Hence

$$\lim_{\varepsilon\downarrow 0} \Delta_\varepsilon P_t f = \lim_{\varepsilon\downarrow 0} P_t \Delta_\varepsilon f = P_t(\mathcal{L}f),$$

where the last equality follows from dominated convergence. Consequently, $P_t f \in \mathcal{D}$ and $P_t(\mathcal{L}f) = \mathcal{L}(P_t f)$. Observe that the map $(t,x) \mapsto P_t f(x)$ is continuous because

$$\lim_{s\downarrow 0} \|P_{t+s}f - P_t f\|_\infty \leq \lim_{s\downarrow 0} \|P_s f - f\|_\infty = 0,$$

and

$$\lim_{s\downarrow 0} \|P_{t-s}f - P_t f\|_\infty = \lim_{s\downarrow 0} \|P_{t-s}(f - P_s f)\|_\infty \leq \lim_{s\downarrow 0} \|f - P_s f\|_\infty = 0.$$

Fix $x \in M$ and define

$$u(t) = P_t f(x) - f(x) - \int_0^t P_s(\mathcal{L}f)(x)\mathrm{d}s.$$

We claim that $u(t) = 0$ for all $t \geq 0$, from which it follows that $t \mapsto P_t f(x)$ is C^1 and satisfies (9).

To prove the claim, it suffices to show that for every $\delta > 0$ the set

$$O_\delta = \{t \geq 0 : |u(t)| > \delta t\}$$

is empty. Suppose, on the contrary, that $O_\delta \neq \emptyset$ for some $\delta > 0$ and let $t^* = \inf O_\delta$. By continuity of $t \mapsto u(t)$, it follows that $|u(t^*)| = \delta t^*$ and, using the form of u above,

$$\begin{aligned}
\lim_{\varepsilon\downarrow 0} \frac{u(t^*+\varepsilon) - u(t^*)}{\varepsilon} &= \lim_{\varepsilon\downarrow 0} \frac{P_{t^*+\varepsilon}f(x) - P_{t^*}f(x)}{\varepsilon} - \frac{1}{\varepsilon}\int_{t^*}^{t^*+\varepsilon} P_s(\mathcal{L}f)(x)ds \\
&= \lim_{\varepsilon\downarrow 0} P_{t^*}\frac{P_\varepsilon f(x) - f(x)}{\varepsilon} - P_{t^*}(\mathcal{L}f)(x) \\
&= P_{t^*}(\mathcal{L}f)(x) - P_{t^*}(\mathcal{L}f)(x) \\
&= 0, \quad \forall x \in M.
\end{aligned}$$

Thus, for every $\delta > 0$, there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon \leq \varepsilon_0$,

$$|u(t^*+\varepsilon)| \leq |u(t^*)| + \delta\varepsilon = \delta(t^*+\varepsilon),$$

which is a contradiction with the definition of O_δ and the continuity of $t \mapsto u(t)$.

It remains to prove that $\left(M_t^f(x)\right)_{t\geq 0}$ is a martingale. For all $t, s \geq 0$,

$$M_{t+s}^f(x) - M_t^f(x) = f(X_{t+s}^x) - f(X_t^x) - \int_0^s \mathcal{L}f(X_{t+u}^x)\mathrm{d}u.$$

By taking conditional expectation, using Fubini's theorem and Markov property, it yields

$$\mathbb{E}\left(M_{t+s}^f(x) - M_t^f(x) \mid \mathcal{F}_t\right) = P_s f(X_t^x) - f(X_t^x) - \int_0^s P_u\mathcal{L}f(X_t^x)\mathrm{d}u.$$

Given (9), it follows that

$$\int_0^s P_u\mathcal{L}f(X_t^x)\mathrm{d}u = \int_0^s \frac{\partial}{\partial u}P_u f(X_t^x)du = P_s f(X_t^x) - f(X_t^x),$$

so that $\mathbb{E}\left(M_{t+s}^f(x) - M_t^f(x) \mid \mathcal{F}_t\right) = 0$, and $\left(M_t^f(x)\right)_{t\geq 0}$ is a martingale. □

Let $f \in \mathcal{D}$ and $x \in M$. Suppose the martingale $(M_t^f(x))_{t\geq 0}$ defined by (10) is square-integrable, meaning that $\mathbb{E}_x\left(|M_t^f|^2\right) < \infty$ for all $t \geq 0$. Then there exists a unique, non-decreasing, right-continuous, adapted process $(\langle M^f(x)\rangle_t)_{t\geq 0}$ called the *predictable quadratic variation* such that

(i) $\langle M^f(x)\rangle_0 = 0$;

(ii) $\left(M_t^f(x)^2 - \langle M^f(x)\rangle_t\right)_{t\geq 0}$ is an $((\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_x)$–martingale.

The general existence of $\left(\langle M^f(x)\rangle_t\right)_{t\geq 0}$ follows from the Doob–Meyer decomposition (see e.g. [Mey62]). However, whenever both f and f^2 lie in $\mathcal{D}$, the quadratic variation can be expressed via the *carré-du-champ* operator.

Definition 1.13 (Carré du champ). *Let $\mathcal{D}^2 := \{f \in \mathcal{D} : f^2 \in \mathcal{D}\}$. For $f \in \mathcal{D}^2$ the* carré-du-champ *of f is defined by*

$$\Gamma(f) = \mathcal{L}f^2 - 2f\mathcal{L}f.$$

Note that

$$\Gamma(f) = \lim_{t\downarrow 0} \frac{1}{t}\left(P_t f^2 - (P_t f)^2\right) \geq 0.$$

Indeed, since both $f, f^2 \in \mathcal{D}$, by Proposition 1.11, the map $t \mapsto P_t f^2(x) - (P_t f(x))^2$ is C^1 for all fixed $x \in M$ and

$$\frac{d}{\mathrm{d}t}\left(P_t f^2(x) - (P_t f(x))^2\right) = P_t(\mathcal{L}f^2)(x) - 2P_t f(x)\cdot \frac{d}{\mathrm{d}t} P_t f(x) = P_t(\mathcal{L}f^2)(x) - 2P_t f(x) P_t(\mathcal{L}f(x)),$$

and in particular

$$\frac{d}{\mathrm{d}t}\left(P_t f^2(x) - (P_t f(x))^2\right)|_{t=0} = \mathcal{L}f^2(x) - 2f(x)\mathcal{L}f(x).$$

Since $P_0 f^2(x) - (P_0 f(x))^2 = 0$, it implies that

$$\begin{aligned}
\lim_{t\downarrow 0} \frac{P_t f^2(x) - (P_t f(x))^2}{t} &= \lim_{t\downarrow 0} \frac{P_t f^2(x) - (P_t f(x))^2 - \left(P_0 f^2(x) - (P_0 f(x))^2\right)}{t} \\
&= \frac{d}{\mathrm{d}t}\left(P_t f^2(x) - (P_t f(x))^2\right)|_{t=0} \\
&= \mathcal{L}f^2(x) - 2f(x)\mathcal{L}f(x) \\
&= \Gamma(f)(x).
\end{aligned}$$

Moreover, using Jensen's inequality, it is clear that

$$P_t f^2(x) - (P_t f(x))^2 = \mathbb{E}[f(X_t^x)^2] - (\mathbb{E}[f(X_t^x)])^2 \geq 0.$$

Proposition 1.14. *If $f \in \mathcal{D}^2$, then*

$$\langle M^f(x)\rangle_t = \int_0^t \Gamma(f)(X_s^x)\mathrm{d}s, \quad t \geq 0. \tag{11}$$

Proof. The map $t \mapsto \int_0^t \Gamma(f)(X_s^x)\mathrm{d}s$ is non-decreasing and continuous by Definition 1.13: let show that the process

$$(M_t^f(x)^2 - \textstyle\int_0^t \Gamma(f)(X_s^x)\mathrm{d}s)_{t\geq 0}$$

is a martingale. We define

$$Z_t = M_t^f(x)^2 - \int_0^t \Gamma(f)(X_s^x)\mathrm{d}s + 2f(x)M_t^f(x) - M_t^{f^2}(x).$$

Since $(M_t^f(x))_{t\geq 0}$, $(M_t^{f^2}(x))_{t\geq 0}$ are martingales, it is sufficient to show that $(Z_t)_{t\geq 0}$ is a martingale to conclude. Write $g_t = \mathcal{L}f(X_t^x)$ and $G_t = \int_0^t g_s \mathrm{d}s$, so that

$$
\begin{aligned}
Z_t &= 2f^2(x) + G_t^2 - 2f(X_t^x)\,(f(x)+G_t) + 2f(x)G_t + \int_0^t \mathcal{L}(f^2)(X_s^x)ds - \int_0^t \Gamma(f)(X_s^x)ds + 2f(x)M_t^f(x)\\
&= G_t^2 - 2f(X_t^x)G_t + \int_0^t \mathcal{L}(f^2)(X_s^x)ds - \int_0^t \Gamma(f)(X_s^x)ds\\
&= G_t^2 - 2f(X_t^x)G_t + \int_0^t \mathcal{L}(f^2)(X_s^x)ds - \int_0^t \mathcal{L}(f^2)(X_s^x)\mathrm{d}s + 2\int_0^t f(X_s^x)g_s\mathrm{d}s\\
&= G_t^2 - 2f(X_t^x)G_t + 2\int_0^t f(X_s^x)g_s ds\\
&= G_t^2 - 2(G_t + M_t^f)G_t + 2\int_0^t (G_s + M_s^f)g_s\mathrm{d}s\\
&= -G_t^2 - 2M_t^f G_t + 2\int_0^t (G_s + M_s^f)g_s\mathrm{d}s
\end{aligned}
$$

Using Fubini's theorem, $G_t^2 = 2\int_0^t G_s g_s \mathrm{d}s$, so that $Z_t = 2\left(\int_0^t M_s^f g_s \mathrm{d}s - M_t^f G_t\right)$. Therefore, for $u > 0$,

$$
\begin{aligned}
Z_{t+u} - Z_t &= 2\left(\int_t^{t+u} M_s g_s ds - G_{t+u}M_{t+u} + G_t M_t\right)\\
&= 2\left(\int_t^{t+u} (M_s - M_{t+u})g_s ds + M_{t+u}(G_{t+u} - G_t) - G_{t+u}M_{t+u} + G_t M_t\right)\\
&= 2\int_t^{t+u} (M_s - M_{t+u})g_s \mathrm{d}s + (M_t - M_{t+u})G_t.
\end{aligned}
$$

Taking conditional expectation with respect to $\mathcal{F}_t$ together with Fubini's theorem, and since $(M_t)_{t\geq 0}$ is a martingale, then $M_t - M_{t+u}$ has zero conditional expectation and similarly,

$$
\mathbb{E}[(M_s - M_{t+u})g_s \mid \mathcal{F}_t] = \mathbb{E}[\mathbb{E}[(M_s - M_{t+u})g_s \mid \mathcal{F}_s] \mid \mathcal{F}_t] = \mathbb{E}[g_s \mathbb{E}[M_s - M_{t+u} \mid \mathcal{F}_s] \mid \mathcal{F}_t] = 0,
$$

since $\mathcal{F}_s \subset \mathcal{F}_t$ and $\mathbb{E}[M_{t+u} \mid \mathcal{F}_s] = M_s$, for all $t \leq s \leq t+u$. It yields $\mathbb{E}[Z_{t+u} - Z_t \mid \mathcal{F}_t] = 0$. Hence $(Z_t)_{t\geq 0}$ is a martingale, proving the claim. □

Remark 1.15. As noticed in the proof of the preceding proposition, the mapping $(t,x) \mapsto P_t(f)(x)$ is continuous for all $f \in \mathcal{D}$, but there is no evidence that this continuity still holds for any $f \in C_b(M)$ even under the $C_b(M)$−Feller property, Definition 1.2. The next elementary result addresses this question and shows that this holds under a very mild condition.

Lemma 1.16. *Let $(t_0, x_0) \in \mathbb{R}_+ \times M$. The following conditions are equivalent:*

(i) *For all $f \in C_b(M)$, the map $(t,x) \mapsto P_t f(x)$ is continuous at (t_0, x_0).*

(ii) *For every sequence $(t_n, x_n) \in \mathbb{R}_+ \times M$ converging to (t_0, x_0), the family $\{P_{t_n}(x_n, \cdot)\}_{n\geq 0}$ is tight.*

In particular, condition ***(ii)*** will always be satisfied under the Lyapunov assumptions made in the next section, as explained in Remark 1.38. Under those hypotheses, it will implies the even stronger property $(t,x) \mapsto P_t(f)(x)$ is continuous, compared to Definition 1.2.

Proof. (i) implies (ii) because weak convergence implies tightness (see e.g. [Kal01], Lemma 5.8). Conversely, if $\{P_{t_n}(x_n, \cdot)\}_{n\geq 0}$ is tight, there exists, by Prokhorov's theorem (see e.g. [Kal01], Theorem 23.2), a subsequence $(t_{n_k}, x_{n_k})_{k\geq 0}$ and a probability measure μ on M such that

$$
P_{t_{n_k}}(x_{n_k}, \cdot) \Rightarrow \mu \quad \text{as } k \to \infty.
$$

For all $f \in \mathcal{D}$, we have $\mu(f) = P_{t_0}f(x_0)$ by continuity (see Remark 1.15). Thus, by Remark 1.10, if $f \in C_b(M)$, then $\mu(\bar{f}_\varepsilon) = P_{t_0}\bar{f}_\varepsilon(x_0)$ and the dominated convergence theorem, it implies that $\mu(f) = \mu(f) = P_{t_0}f(x_0)$ for all $f \in C_b(M)$. □

1.2.1 Extended generator and carré du champ

For many of our next results and our examples of interest, we will need to use functions f that are unbounded on M or M_+: one can think of $f : [0,\infty) \to [0,\infty) : x \mapsto x$ or $f : (0,\infty) \to (0,\infty) : x \mapsto \log x$. To this effect, we introduce an extension of the generator $\mathcal{L}$.

Recall that an adapted process $(M_t)_{t\geq 0}$ on $(\Omega, (\mathcal{F}_t)_{t\geq 0})$ with $M_0 = 0$ is called a $((\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_x)$*–local martingale* if there exists a non–decreasing sequence $(\tau_n)_{n\geq 0}$ of stopping times such that $\lim_{n\to\infty}\tau_n = \infty$ and, for every n, the stopped process $(M_t^{\tau_n})_{t\geq 0}$ defined by $M_t^{\tau_n} = M_{t\wedge\tau_n}$ is a $((\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_x)$–martingale. The sequence $(\tau_n)_{n\geq 0}$ is said to *reduce* $(M_t)_{t\geq 0}$.

Throughout we let $\mathcal{M}$ denote either of the sets M or M_+ and we write $C(\mathcal{M})$ for the vector space of continuous (possibly unbounded) maps $f\colon \mathcal{M} \to \mathbb{R}$. The following lemma justifies the definition of the extended generators.

Lemma 1.17. *There exists a vector space $\mathcal{D}_e^{\mathcal{M}} \subset C(\mathcal{M})$ and a linear map $\mathcal{L}_e^{\mathcal{M}}\colon \mathcal{D}_e^{\mathcal{M}} \to C(\mathcal{M})$ such that:*

(i) *For every $f \in \mathcal{D}_e^{\mathcal{M}}$ and $x \in \mathcal{M}$, the process*

$$M_t^f(x) = f(X_t^x) - f(x) - \int_0^t \mathcal{L}_e^{\mathcal{M}} f(X_s^x)\mathrm{d}s, \quad t \geq 0, \tag{12}$$

is a $((\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_x)$–local martingale.

(ii) *Let $(f,g) \in C(\mathcal{M})^2$ be such that, for every $x \in \mathcal{M}$, the process*

$$M_t^{f,g}(x) = f(X_t^x) - f(x) - \int_0^t g(X_s^x)\mathrm{d}s \tag{13}$$

is a $((\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_x)$–local martingale. Then $f \in \mathcal{D}_e^{\mathcal{M}}$ and $\mathcal{L}_e^{\mathcal{M}} f = g$.

Proof. Define $\mathcal{D}_e^{\mathcal{M}}$ as the set of $f \in C(\mathcal{M})$ for which there exists $g \in C(\mathcal{M})$ such that $M_t^{f,g}(x)$ is a $((\mathcal{F}_t)_{\geq 0}, \mathbb{P})$–local martingale for every x. This is a non–empty vector space since it contains $\mathcal{D}$. Suppose now that $\tilde{g} \in C(\mathcal{M})$ is another function such that $M_t^{f,\tilde{g}}(x)$ is also a $((\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_x)$–local martingale for every x. We claim that $g = \tilde{g}$, which then allows us to set $\mathcal{L}_e^{\mathcal{M}} f = g$.

Assume to the contrary that $\tilde{g}(x) - g(x) > 0$ for some $x \in \mathcal{M}$. Choose $\delta > 0$ and a compact neighborhood U of x such that $\tilde{g}(y) - g(y) \geq \delta$ for all $y \in U$. Let $\tau = \inf\{t \geq 0 : X_t^x \notin U\}$ and define the process

$$N_t := M_{t\wedge\tau}^{f,\tilde{g}} - M_{t\wedge\tau}^{f,g}(x) = \int_0^{t\wedge\tau} (\tilde{g}(X_s^x) - g(X_s^x))\,\mathrm{d}s, \quad t \geq 0.$$

Then $(N_t)_{t\geq 0}$ is a local martingale since both $M_t^{f,\tilde{g}}$, $M_t^{f,g}(x)$ are, and $0 \leq \delta(t\wedge\tau) \leq N_t \leq Ct$, where $C = \sup_{y\in U}|\tilde{g}(y) - g(y)|$ since $\tilde{g}, g$ are continuous functions taken over a compact.

Hence $(N_t)_{t\geq 0}$ is a true martingale, because if $(\tau_n)_{n\geq 1}$ is the sequence of stopping times that reduces $(N_t)_{t\geq 0}$, the family $N_t^{\tau_n}$ is uniformly integrable in n since $\mathbb{E}_x(N_t^{\tau_n}) \leq Ct$, and passing to the limit $n \to \infty$ in $\mathbb{E}_x[N_t^{\tau_n} \mid \mathcal{F}_s] = N_s^{\tau_n}$ yields

$$\mathbb{E}_x(N_t \mid \mathcal{F}_s) = \mathbb{E}_x(\lim_{n\to\infty} N_t^{\tau_n} \mid \mathcal{F}_s) = \lim_{n\to\infty}\mathbb{E}_x(N_t^{\tau_n} \mid \mathcal{F}_s) = \lim_{n\to\infty} N_s^{\tau_n} = N_s,$$

while we take out the limit out of expectation by dominated convergence theorem. Therefore, we have

$$0 \leq \mathbb{E}_x[N_t] \leq \mathbb{E}_x(N_0) = 0 \;\Rightarrow 0 = \mathbb{E}^x[N_t] \geq \delta \mathbb{E}_x[t \wedge \tau],$$

so that $\tau = 0$ $\mathbb{P}_x$–almost surely, which contradicts the right–continuity of paths of $(X_t^x)_{t\geq 0}$: around a small time interval $[0, \varepsilon[$, X_t^x should stay in U and τ would be strictly positive. Thus, we have $g = \tilde{g}$. □

Definition 1.18 (Extended generator). *The operator $\mathcal{L}_\mathrm{e}^{\mathcal{M}}$ is called the* extended generator *of $(P_t)_{t\geq 0}$ on $\mathcal{M}$, and $\mathcal{D}_\mathrm{e}^{\mathcal{M}}$ its domain. To shorten notation we write $(\mathcal{D}_\mathrm{e}, \mathcal{L}_\mathrm{e})$ (respectively $(\mathcal{D}_\mathrm{e}^+, \mathcal{L}_\mathrm{e}^+)$) instead of $\left(\mathcal{D}_\mathrm{e}^M, \mathcal{L}_\mathrm{e}^M\right)$ (respectively $\left(\mathcal{D}_\mathrm{e}^{M_+}, \mathcal{L}_\mathrm{e}^{M_+}\right)$).*

For shortness, we let

$$\mathcal{D}_e^M = \mathcal{D}_e, \; \mathcal{L}_e^M = \mathcal{L}_e, \; \text{ and } \; \mathcal{D}_e^{M_+^I} = \mathcal{D}_e^I, \; \mathcal{L}_e^{M_+^I} = \mathcal{L}_e^I.$$

When $I = \{1, \cdots, n\}$, we simply replace the I exponents by $+$.

Clearly, if $f \in \mathcal{D}_e$, we can see that $f|_{M_+} \in \mathcal{D}_e^+$ and in particular

$$\mathcal{L}_e^+(f|_{M_+}) = \mathcal{L}_e(f)|_{M_+}.$$

More precisely, we can assess the *extension* term through the following result.

Proposition 1.19. *$\mathcal{D} = \{f \in \mathcal{D}_e \cap C_b(M) : \mathcal{L}_e(f) \in C_b(M)\}$ and for all $f \in \mathcal{D}$, $\mathcal{L}_e(f) = \mathcal{L}(f)$.*

Proof. Since a martingale is a local martingale, the inclusion $\subset$ is direct.

Conversely, let $f \in \mathcal{D}_e \cap C_b(M)$ and $g = \mathcal{L}_e(f) \in C_b(M)$. By definition, $(M_t)_{t\geq 0} := (M_t^{f,g})_{t\geq 0}$ is a local martingale which satisfies

$$|M_t| \leq 2\|f\|_\infty + \|g\|_\infty t.$$

Since f and g are supposed to be bounded, the same construction as in Proposition 1.14 lets us show that $(M_t)_{t\geq 0}$ is in fact a true $((\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_x)$–martingale for all $x \in M$. Thus, by Fubini's theorem,

$$\mathbb{E}(M_t) = P_t f(x) - f(x) - \int_0^t P_s g(x) ds = 0,$$

so that $f \in \mathcal{D}$ and $g = \mathcal{L}f$ by Definition 1.9. □

Even in the case where $(M_t^f(x))_{t\geq 0}$ is a local martingale, we can still give sense to its predictable quadratic variation.

Proposition 1.20 ([Kal01], Prop. 23.1)**.** *Let $(M_t)_{t\geq 0}$ be a square-integrable local martingale (in the sense that any stopped process is a square-integrable martingale). Then, there exists a unique, increasing, predictable process $(\langle M\rangle_t)_{t\geq 0}$ with locally integrable variation such that*

(i) $\langle M\rangle_0 = 0$.

(ii) The process $(M_t^2 - \langle M\rangle_t)_{t\geq 0}$ is a local martingale.

By analogy, we can build an *extended carré du champ* for functions $f \in \mathcal{D}_e^{\mathcal{M}}$ such that $f^2 \in \mathcal{D}_e^{\mathcal{M}}$.

Definition 1.21 (Extended carré du champ). *Let*

$$\mathcal{D}_e^{2,\mathcal{M}} = \left\{f \in \mathcal{D}_e^{\mathcal{M}} : f^2 \in \mathcal{D}_e^{\mathcal{M}}\right\}.$$

For $f \in \mathcal{D}_e^{2,\mathcal{M}}$, the extended carré du champ is defined by

$$\Gamma_e^{\mathcal{M}}(f) = L_e^{\mathcal{M}}(f^2) - 2f L_e^{\mathcal{M}}(f).$$

Here again, we use the simplified notations

$$\mathcal{D}_e^{2,M} = \mathcal{D}_e^2,\ \Gamma_e^M = \Gamma_e, \text{ and } \mathcal{D}_e^{2,M_+^I} = \mathcal{D}_e^{2,I},\ \Gamma_e^{M_+^I} = \Gamma_e^I,$$

and when $I = \{1, \cdots, n\}$, we replace the I exponents by $+$.

Similarly to Proposition 1.14, one can also link the predictable quadratic variation of $(M_t^f(x))_{t\geq 0}$ to the extended carré du champ in the local martingale case:

Proposition 1.22. *Let $f \in \mathcal{D}_e^{2,\mathcal{M}}$. Then, $\Gamma_e^{\mathcal{M}}(f) \geq 0$ and for all $x \in \mathcal{M}$, the process*

$$\left(M_t^f(x)^2 - \int_0^t \Gamma_e^{\mathcal{M}}(f)(X_s^x)ds\right)_{t\geq 0},$$

is a $((\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ −local martingale. In particular,

$$\langle M^f(x)\rangle_t = \int_0^t \Gamma_e^{\mathcal{M}}(f)(X_s^x)ds.$$

Proof. We proceed like in the proof of Proposition 1.14. Let

$$Z_t = \left(M_t^f(x)^2 - \int_0^t \Gamma_e^{\mathcal{M}}(f)(X_s^x)ds\right) + 2f(x)M_t^f(x) - M_t^{f^2}(x).$$

Since $(M_t^f(x))_{t\geq 0}$ and $(M_t^{f^2}(x))_{t\geq 0}$ are local martingales, it suffices to show that $(Z_t)_{t\geq 0}$ is a local martingale. Let $g_t = \mathcal{L}_e^{\mathcal{M}}(f)(X_t^x)$ and $G_t = \int_0^t g_s ds$: as observed in the proof of Proposition 1.14,

$$Z_t = 2\left(\int_0^t M_s g_s ds - M_t G_t\right), \quad Z_{t+u} - Z_t = 2\left(\int_t^{t+u} (M_s^f - M_{t+u}^f) g_s ds + (M_t^f - M_{t+u}^f) G_t\right).$$

Let

$$\sigma_n = \inf\left\{t \geq 0 : \int_0^t |g_s|ds \geq n \text{ or } \int_0^t |M_s^f g_s|ds \geq n\right\}.$$

By the càdlàg continuity of g_s and M_s^f, the maps $t \mapsto \int_0^t |g_s|ds$ and $t \mapsto \int_0^t |M_s^f g_s|ds$ are continuous, so that $(\sigma_n)_{n\geq 0}$ is a sequence of stopping times with $\lim_{n\to\infty} \sigma_n = \infty$.

Let now $(\tau_n)_{n\geq 0}$ be a localizing sequence for $(M_t^f)_{t\geq 0}$ and $\tau_n' = \tau_n \wedge \sigma_n$. Before σ_n, we have $\int_0^t |g_s|ds < n$ and $\int_0^t |M_s^f g_s|ds < n$ so that

$$|Z_t^{\tau_n'}| \leq 2n(1 + |(M_t^f)^{\tau_n'}|),$$

so that $Z_t^{\tau_n'}$ is integrable in the sense that $\mathbb{E}(|Z_t^{\tau_n'}|) < \infty$, since $(M_t^f)^{\tau_n'}$ is, and

$$\begin{aligned}
\mathbb{E}\left(Z_{t+u}^{\tau_n'} - Z_t^{\tau_n'} \mid \mathcal{F}_t\right) &= 2\int_t^{t+u} \mathbb{E}\left[\left((M_s^f)^{\tau_n'} - (M_{t+u}^f)^{\tau_n'}\right) g_{s\wedge\tau_n'} \mid \mathcal{F}_t\right] ds + 2\mathbb{E}\left[\left((M_t^f)^{\tau_n'} - (M_{t+u}^f)^{\tau_n'}\right) G_{t\wedge\tau_n'} \mid \mathcal{F}_t\right] \\
&= 2\int_t^{t+u} \mathbb{E}\left[\mathbb{E}\left((M_s^f)^{\tau_n'} - (M_{t+u}^f)^{\tau_n'} \mid \mathcal{F}_{s\wedge\tau_n'}\right) g_{s\wedge\tau_n'} \mid \mathcal{F}_t\right] ds \\
&\quad + 2\mathbb{E}\left[\mathbb{E}\left(M_t^{\tau_n'} - M_{t+u}^{\tau_n'} \mid \mathcal{F}_{s\wedge\tau_n'}\right) G_{t\wedge\tau_n'} \mid \mathcal{F}_t\right] \\
&= 0.
\end{aligned}$$

This proves that $(Z_t)_{t\geq 0}$ is a $((\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_x)$-local martingale, for all $x \in M$.

Now, let's show that $\Gamma_e^{\mathcal{M}}(f) \geq 0$. Suppose to the contrary that $\Gamma_e^{\mathcal{M}}(f)(x) = -\delta < 0$ for some $x \in \mathcal{M}$. Let

$$T = \inf\left\{t \geq 0 : \Gamma_e^{\mathcal{M}}(f)(X_t^x) > -\delta/2\right\},$$

and let $(\tau_n)_{n\geq 0}$ be a localizing sequence for $\left(M_t^f(x)^2 - \int_0^t \Gamma_e^{\mathcal{M}}(f)(X_s^x)ds\right)_{t\geq 0}$. Then

$$-\frac{\delta}{2}\mathbb{E}_x(T \wedge \tau_n) \geq \mathbb{E}_x\left(\int_0^{T\wedge\tau_n} \Gamma_e^{\mathcal{M}}(f)(X_s^x)ds\right) = \mathbb{E}_x\left(M_{t\wedge\tau_n}^f(x)^2\right) \geq 0.$$

Thus $T = 0$, $\mathbb{P}_x$-almost surely, in contradiction with right-continuity of paths, similarly to the end of the proof of Proposition 1.14.

To conclude, recall that since $(X_t^x)_{t\geq 0}$ is adapted with càdlàg path, it is progressively measurable, which is preserved by composition with measurable functions. Since $\int_0^t H_s ds$ forms an adapted process for any progressively measurable process $(H_t)_{t\geq 0}$, it implies that $\int_0^t \Gamma_e^{\mathcal{M}}(f)(X_s^x)ds$ is also adapted since $\Gamma_e^{\mathcal{M}}(f)$ is continuous.

Since $\int_0^t \Gamma_e^{\mathcal{M}}(f)(X_s^x)ds$ is a continuous adapted process, it is left-continuous and, therefore, it is predictable. Moreover, since $\Gamma_e^{\mathcal{M}}(f) \geq 0$, it forms an increasing process starting at 0.

Then, with the sequence of stopping times defined as $\tau_n := \inf\{t \geq 0 : \int_0^t \Gamma_e^{\mathcal{M}}(f)(X_s^x)ds \geq n\}$, locally finite variation is verified and the conclusion follows directly by the uniqueness of the predictable quadratic variation. □

Definition 1.23 (Strong law). *Let $f \in \mathcal{D}_e^{\mathcal{M}}$. We say that f satisfies the* strong law *if, for every $x \in \mathcal{M}$,*

$$\lim_{t\to\infty} \frac{M_t^f(x)}{t} = 0 \quad \mathbb{P}_x\text{-a.s.}, \tag{14}$$

where $(M_t^f(x))_{t\geq 0}$ is the local martingale defined by (12).

By Proposition 1.14, we always have $\mathcal{D}^2 \subset \mathcal{D}_e^2$ but the converse is false. However, the following tool will be used to verify either $(M_t^f(x))_{t\geq 0}$ is a true martingale or if it satisfies the strong law. It has been proposed as part of the writing process of [BHN+25].

Proposition 1.24. *Let $f \in \mathcal{D}_e^{\mathcal{M}}$. Consider the following assertions:*

(i) *$f \in \mathcal{D}_e^{2,\mathcal{M}}$, and for all $x \in \mathcal{M}$,*

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t P_s(\Gamma_e^{\mathcal{M}}(f))(x)ds < \infty;$$

(ii) *For all $x \in \mathcal{M}$, $(M_t^f(x))_{t\geq 0}$ is a true L^2 $((\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_x)$−martingale, and*

$$\limsup_{t\to\infty} \frac{\mathbb{E}^x\left[(M_t^f(x))^2\right]}{t} < \infty; \tag{15}$$

(iii) *f satisfies the strong law.*

Then

***(i)* ⇒ *(ii)* ⇒ *(iii)*.**

Proof. Fix $x \in \mathcal{M}$, and, to shorten notation, set $M_t = M_t^f(x)$. Assume ***(i)***. Let $(\tau_n)_{n\geq 0}$ be a localizing sequence for the local martingale defined in Proposition 1.22. Let $M_t^n = M_{t\wedge\tau_n}$. Then, for t sufficiently large,

$$\mathbb{E}\left[(M_t^n)^2\right] = \mathbb{E}\left[\int_0^{t\wedge\tau_n} \Gamma_e^{\mathcal{M}}(f)(X_s^x)ds\right] \leq \int_0^t P_s\Gamma_e^{\mathcal{M}}(f)(x)ds \leq C_x t,$$

for some constant C_x using Fubini's theorem and the fact that $t \wedge \tau_n \leq t$ for $\Gamma_e^{\mathcal{M}}(f) \geq 0$. The sequence $(M_t^n)_{n\geq 1}$ is then bounded in L^2, hence uniformly integrable. It therefore converges in L^1, as $n \to \infty$, toward M_t. Since $(M_t^n)_{t\geq 0}$ is a martingale, by dominated convergence theorem, the L^1-convergence passes to the limit in the martingale property:

$$\mathbb{E}(M_t \mid \mathcal{F}_s) = \mathbb{E}(\lim_{n\to\infty} M_t^n \mid \mathcal{F}_s) = \lim_{n\to\infty} \mathbb{E}(M_t^n \mid \mathcal{F}_s) = \lim_{n\to\infty} M_s^n = M_s, \quad \text{for } 0 \leq s \leq t.$$

This proves that $(M_t)_{t\geq 0}$ is a martingale. By Fatou's lemma, the previous inequality also implies that

$$\mathbb{E}[M_t^2] \leq C_x t,$$

so that (15) is satisfied, which concludes the proof that ***(i)*** $\Rightarrow$ ***(ii)***.

We now show that ***(ii)*** $\Rightarrow$ ***(iii)***. For all integers n and $\varepsilon > 0$, Doob's inequality for right-continuous martingales ([Gal16], Proposition 3.15) implies that

$$\mathbb{P}_x\left(\sup_{2^n\leq t\leq 2^{n+1}} \frac{|M_t|}{t} \geq \varepsilon\right) \leq \mathbb{P}_x\left(\sup_{t\leq 2^{n+1}} |M_t| \geq \varepsilon 2^n\right) \leq \frac{\mathbb{E}_x\left[\sup_{t\leq 2^{n+1}} |M_t|^2\right]}{\varepsilon^2 2^{2n}} \leq \frac{4C_x 2^{n+1}}{\varepsilon^2 2^{2n}} \leq \frac{8C_x}{\varepsilon^2 2^n},$$

where the second inequality follows Markov's inequality, and the third one Doob's inequality in L^2. Thus, $\limsup_{t\to\infty} \frac{M_t}{t} = 0$, $\mathbb{P}_x$-almost surely, by Borel–Cantelli lemma. □

Remark 1.25. There are classical examples of local martingales that are bounded in L^2 but not true martingales, we can think about Exercise 5.33, (8) in [Gal16]. Thus condition (15) alone is not sufficient to ensure that $(M_t^f(x))_{t\geq 0}$ is a true martingale.

Indeed, given $(B_t)_{t\geq 0}$ a 3−dimensional Brownian motion started at $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, we can show by direct calculus that $(|B_t|^{-1})_{t\geq 0}$ is bounded in L^2. However, let's suppose that it is also a continuous true martingale. Then, combining $\lim_{t\to\infty} |B_t| = +\infty$ almost surely, and the fact that

$$\mathbb{E}(|B_\infty|^{-2}) = 0 = \mathbb{E}(|B_0|^{-2}) + \mathbb{E}(\langle |B|^{-1}, |B|^{-1}\rangle_\infty),$$

it is in contradiction with Theorem 4.13 of [Gal16] since both terms should be positive. Hence, $(|B_t|^{-1})_{t\geq 0}$ is only a continuous local martingale.

1.3 Invariant and ergodic probability measures

A probability measure $\mu \in \mathcal{P}(M)$ is called *stationary* or *invariant* if

$$\mu P_t = \mu, \quad t \geq 0,$$

that is, $\mu(P_t f) = \mu f$ for all $f \in B(M)$ (or $C_b(M)$) and all $t \geq 0$. Denote the set of invariant probability measures of $(P_t)_{t\geq 0}$ by $\mathcal{P}_{\text{inv}}(M)$. We also define

$$\mathcal{P}_{\text{inv}}(M_0) = \{\mu \in \mathcal{P}_{\text{inv}}(M) : \mu(M_0) = 1\},\ \mathcal{P}_{\text{inv}}(M_+) = \{\mu \in \mathcal{P}_{\text{inv}}(M) : \mu(M_+) = 1\}.$$

Previous properties are central to show the existence, hence uniqueness, of an invariant probability measure associated to the semigroup $(P_t)_{t\geq 0}$. For example, in the case of compact state space:

Proposition 1.26 ([Str19], Proposition 1.6)**.** *Let $(X_t)_{t\geq 0}$ be a $C_b(M)-$ Feller semigroup on M a compact space. Then, there exists at least one invariant probability measure.*

One the other hand, the uniqueness of the invariant probability measure can be deduced from the existence of an accessible Doeblin point.

Proposition 1.27 ([Bou23], Proposition 1.11)**.** *Let $(X_t)_{t\geq 0}$ be a $C_b(M)-$ Feller semigroup. We suppose that $\exists x^*$ an accessible Doeblin point. Then, there exists at most one invariant probability measure.*

Alternatively, invariant probability measures can be deduced from the infinitesimal generator introduced in Definition 1.9:

Proposition 1.28 ([Str19], Proposition 1.5)**.** *Let $(X_t)_{t\geq 0}$ be a $C_b(M)-$ Feller semigroup with generator $\mathcal{L}$, then $\mu \in \mathcal{P}_{\text{inv}}(M)$ if and only if $\mu \mathcal{L} f = 0$ for every $f \in \mathcal{D}$.*

Another point of view is given through the 1–resolvent kernel G defined in (6):

Proposition 1.29 ([BH22], Proposition 4.57)**.** *A probability measure μ is invariant for G if and only it is invariant for $(P_t)_{t\geq 0}$.*

Recall that a function $f \in B(M)$ is (G,μ)*–invariant* if $Gf = f$ μ–a.s. Likewise, a set $B \in \mathcal{B}(M)$ is (G,μ)–invariant if $\mathbf{1}_B$ is (G,μ)–invariant.

Lemma 1.30. *Let $\mu \in \mathcal{P}_{\text{inv}}(M)$ and $f \in B(M)$. Then f is (G,μ)–invariant if and only if the set*

$$\{(x,t) \in M \times \mathbb{R}_+ : P_t f(x) = f(x)\}$$

has full measure with respect to $\mu \otimes \mathrm{d}t$.

Proof. Assume f is (G,μ)–invariant. Then by Fubini's theorem,

$$\begin{aligned} 0 &\leq \int_0^\infty e^{-t}\mu((P_t f - f)^2)\mathrm{d}t \\ &\leq \int_0^\infty e^{-t}\mu(P_t f^2 - 2fP_t f + f^2)\mathrm{d}t \\ &= 2\mu(f^2 - fGf) \\ &= 0, \end{aligned}$$

for almost every $t \geq 0$. Hence, $\{(x,t) \in M \times \mathbb{R}_+ : P_t f(x) = f(x)\}$ has full measure with respect to $\mu \otimes \mathrm{d}t$ and the converse is immediate. □

Definition 1.31 (**Ergodic probability measure**)**.** *An invariant probability $\mu \in \mathcal{P}_{\text{inv}}(M)$ is called* ergodic *if every $(G,\mu)-$invariant map is $\mu-$almost surely constant.*

We denote the set of ergodic measures by $\mathcal{P}_{\text{erg}}(M)$ and we also defined $\mathcal{P}_{\text{erg}}(M_0) = \{\mu \in \mathcal{P}_{\text{erg}}(M) : \mu(M_0) = 1\}$.

Definition 1.32 (**Extremal probability measure**)**.** *$\mu \in \mathcal{P}_{\text{inv}}(M)$ is called* extremal *if it cannot be written as a non–trivial convex combination $\mu = \varepsilon\mu_1 + (1-\varepsilon)\mu_0$, $0 < \varepsilon < 1$, of two distinct invariant measures $\mu_0, \mu_1 \in \mathcal{P}_{\text{inv}}(M)$.*

The next proposition gathers results from [BH22], in particular Lemma 4.27, Proposition 4.29 and Theorem 4.51.

Proposition 1.33. *Let $\mu \in \mathcal{P}_{\text{inv}}(M)$.*

(i) *The following are equivalent:*

(a) μ *is ergodic;*

(b) *every* (G,μ)*–invariant set has* μ*–measure* 0 *or* 1*;*

(c) μ *is extremal.*

(ii) ***Ergodic decomposition.*** *There exists a Markov kernel* Q *on* M *such that*

$$\mu(\cdot) = \int_M Q(x,\cdot)\mu(\mathrm{d}x),$$

and $Q(x,\cdot)$ *is ergodic for* μ*–almost every* x*.*

1.3.1 Empirical occupation measures

We define the *empirical occupation measures* $(\Pi_t^x)_{t\geq 0}$ of the process $(X_t^x)_{t\geq 0}$ by

$$\Pi_t^x(B) = \frac{1}{t}\int_0^t \mathbf{1}_{\{X_s^x \in B\}}\mathrm{d}s, \quad B \in \mathcal{B}(M). \tag{16}$$

We can see $\Pi_t^x(B)$ as the fraction of time the process spends in B up to time t.

Remark 1.34. The existence of invariant probability measures highly rely on the tightness of $(\Pi_t^x)_{t\geq 0}$: it is naturally verified in Proposition 1.26 since we only consider compact state space M. Therefore, our next interest will be to weaken the conditions to achieve tightness of the empirical measure.

We also introduce the *mean* empirical measure $(\overline{\Pi}_t^x)_{t\geq 0}$ defined as

$$\overline{\Pi}_t^x(B) = \mathbb{E}[\Pi_t^x(B)] = \frac{1}{t}\int_0^t P_s(x,B)\mathrm{d}s, \quad B \in \mathcal{B}(M). \tag{17}$$

Recall that a continuous map $W\colon M \to \mathbb{R}$ is *proper* if the sublevel sets $\{x \in M : W(x) \leq R\}$ are compact for all $R > 0$. Proper maps are convenient for establishing tightness of $(\overline{\Pi}_t^x)_{t\geq 0}$ and that each of its limit point lies in $\mathcal{P}_{\mathrm{inv}}(M)$, as Proposition 4.15 in [BH22] or as below with Lemma 9.4 in [Ben18].

Lemma 1.35. *Let* W *be a non–negative proper function,* $C \geq 0$*, and* $(\mu_n) \subset \mathcal{P}(M)$ *with* $\limsup_{n\to\infty} \mu_n W \leq C$*. Then*

(i) (μ_n) *is tight and every limit point* μ *satisfies* $\mu W \leq C$*;*

(ii) *If* $H\colon M \to \mathbb{R}$ *is continuous with* $\frac{W}{1+|H|}$ *proper and* $\mu_n \Rightarrow \mu$*, then* $\mu_n H \underset{n\to\infty}{\longrightarrow} \mu H$*.*

Remark 1.36. Since $\frac{W}{1+|H|}$ is proper it follows that $H \in L^1(\mu)$: indeed, there exists a constant $K > 0$ and a compact set C_K such that outside C_K, we have

$$\frac{W}{1+|H|} > K \;\Rightarrow\; |H| < \frac{W}{K} - 1 \;\Rightarrow\; \mu|H| \leq \frac{1}{K}\mu(W) \leq \frac{C}{K},$$

by ***(i)***, while on C_K, H is continuous on a compact so $\mu|H| < \infty$.

This leads to our second hypothesis:

Hypothesis 2. There exist proper maps $W, \tilde{W}\colon M \to \mathbb{R}_+$, with $W \in \mathcal{D}_{\mathrm{e}}$, and a constant $C \geq 0$ such that

$$\mathcal{L}_{\mathrm{e}} W \leq -\tilde{W} + C.$$

Here, W and $\tilde{W}$ play the role of Lyapunov-type functions, and guarantee in particular the tightness of the mean empirical measure. The next proposition is adapted from Theorem 2.2 in [Ben18] with weaker conditions, also developed within the writing of [BHN+25].

Proposition 1.37. *Assume Hypothesis 2. Then*

(i) *For all $x \in M$,*

$$P_t W(x) + \int_0^t P_s \tilde{W}(x) \mathrm{d}s \leq W(x) + Ct.$$

In particular $\limsup_{t\to\infty} \overline{\Pi}_t^x(\tilde{W}) \leq C$.

(ii) *The family $(\overline{\Pi}_t^x)_{t\geq 0}$ is tight and every limit point lies in $\mathcal{P}_{\mathrm{inv}}(M)$.*

(iii) *$\mathcal{P}_{\mathrm{inv}}(M)$ is non–empty and compact, and $\mu\tilde{W} \leq C$ for all $\mu \in \mathcal{P}_{\mathrm{inv}}(M)$.*

Proof. Set

$$M_t(x) := W(X_t^x) - W(x) - \int_0^t \mathcal{L}_{\mathrm{e}} W(X_s^x) \mathrm{d}s, \quad t \geq 0. \tag{18}$$

Then, under Hypothesis 2 and in particular $W \in \mathcal{D}_e$, this is a càdlàg local martingale. Moreover, with $\mathcal{L}_{\mathrm{e}} W \leq -\tilde{W} + C$, and since $W, \tilde{W}$ are positive functions, we can rearranging the terms to obtain

$$0 \leq W(X_t^x) + \int_0^t \tilde{W}(X_s^x) \mathrm{d}s \leq W(x) + Ct + M_t(x). \tag{19}$$

For $n \geq 0$, let $(M_t^n(x))_{t\geq 0}$ defined as $M_t^n(x) = M_{t\wedge n}(x)$, which is a local martingale: therefore, there exists a non decreasing sequence of stopping times $(\tau_k)_{k\geq 0}$ with $\lim_{k\to\infty} \tau_k = \infty$ such that $(M_{t\wedge\tau_k}^n(x))_{t\geq 0}$ is a martingale and

$$\mathbb{E}(M_{t\wedge\tau_k(x)}^n) = \mathbb{E}(M_0^n(x)) = 0.$$

In addition, using (19), we can lower bound $(M_t^n(x))_{t\geq 0}$ by

$$M_t^n(x) \geq -W(x) - Cn.$$

By applying Fatou's lemma to the non-negative sequence $M_t^n(x) + W(x) + Cn$, we obtain

$$\mathbb{E}(M_t^n(x)) = \mathbb{E}\left(\lim_{k\to\infty} M_{t\wedge\tau_k}^n(x)\right) \leq \lim_{k\to\infty} \mathbb{E}\left(M_{t\wedge\tau_k}^n(x)\right) = \lim_{k\to\infty} \mathbb{E}(M_0^n(x)) = 0.$$

Thus, n being arbitrary, we have $\mathbb{E}(M_t(x)) \leq 0$ and by Fubini's theorem, (19) rewrites as

$$P_t W(x) + \int_0^t P_s \tilde{W}(x) ds \leq W(x) + Ct.$$

In addition, by dividing above inequality by t and rearranging the terms, we obtain

$$\overline{\Pi}_t^x(\tilde{W}) = \frac{1}{t}\int_0^t P_s(\tilde{W})(x) ds \leq -\frac{P_t W(x)}{t} + \frac{W(x)}{t} + C,$$

and taking the lim sup when $t \to \infty$ gives us

$$\limsup_{t\to\infty} \overline{\Pi}_t^x(\tilde{W}) \leq C,$$

so ***(i)*** is proved.

Then, by Lemma 1.35, it implies that the family $(\overline{\Pi}_t^x(\tilde{W}))_{t\geq 0}$ is tight. For $f \in C_b(M)$ and $s \geq 0$, we can notice that

$$\begin{aligned}
\lim_{t\to\infty} |\overline{\Pi}_t^x(f) - \overline{\Pi}_t^x(P_s f)| &\leq \lim_{t\to\infty} \frac{1}{t}\left(\left|\int_0^t P_u f(x)du\right| + \left|\int_0^t P_{s+u} f(x)du\right|\right) \\
&= \lim_{t\to\infty} \frac{1}{t}\left(\left|\int_0^s P_u f(x)du\right| + \left|\int_t^{t+s} P_u f(x)du\right|\right) \\
&\leq \lim_{t\to\infty} \frac{2s\|f\|_\infty}{t} \\
&= 0,
\end{aligned}$$

so that every limit points of $(\overline{\Pi}_t^x(\tilde{W}))_{t\geq 0}$ lies in $\mathcal{P}_{\mathrm{inv}}(M)$, and ***(ii)*** is proved.

In particular, $\mathcal{P}_{\mathrm{inv}}(M)$ is non-empty, and by Lemma 1.35, $\mu\tilde{W} \leq C$ for μ a limit point of $(\overline{\Pi}_t^x(\tilde{W}))_{t\geq 0}$.

Now, we use Jensen's inequality such that, for all $R > 0$,

$$0 \leq \frac{1}{t}\int_0^t P_s(\tilde{W} \wedge R)ds \leq \left(\frac{1}{t}\int_0^t P_s(\tilde{W})ds\right) \wedge R \leq \left(\frac{W}{t} + C\right) \wedge R.$$

Then, for every $\mu \in \mathcal{P}_{\mathrm{inv}}(M)$, we have

$$\mu(\tilde{W} \wedge R) \leq \mu\left(\left(\frac{W}{t} + C\right) \wedge R\right),$$

so letting $t, R \to \infty$, we get $\mu\tilde{W} \leq C$. In particular: for every $\mu \in \mathcal{P}_{\mathrm{inv}}(M)$, μ is tight.

To show compactness, by tightness of $\mathcal{P}_{\mathrm{inv}}(M)$, it is relatively compact by Prohorov's theorem. Then, by $C_b(M)-$Feller continuity, it implies that $P_{\mathrm{inv}}(M)$ is closed, and we can conclude about its compactness, which proves ***(iii)***. Indeed, for $(\mu_n)_{n\geq} \subset \mathcal{P}_{\mathrm{inv}}(M)$ that converges weakly to μ, we have

$$\int_M P_t f d\mu = \lim_{n\to\infty}\int_M P_t f d\mu_n = \lim_{n\to\infty}\int_M f d\mu_n = \int_M f d\mu, \quad \forall f \in C_b(M),$$

by weak convergence, invariance of μ_n for every $n \geq 0$ and the fact that $P_t f \in C_b(M)$, so that the limit μ of $(\mu_n)_{n\geq 0}$ also lies in $\mathcal{P}_{\mathrm{inv}}(M)$, which shows its closedness. □

Remark 1.38. Under Hypothesis 2, property ***(i)*** implies that the mapping $(t, x) \mapsto P_t f(x)$ is continuous for every $f \in C_b(M)$ by combining Lemmas 1.16 and Lemma 1.35.

Another useful consequence of Proposition 1.37 is the following.

Corollary 1.39. *Assume that Hypothesis 2 holds. Let $\mathcal{M}$ be one of the sets M or M_+. Let $f \in \mathcal{D}_e^{2,\mathcal{M}}$ be such that*

$$\Gamma_e^{\mathcal{M}}(f)(x) \leq \tilde{a}\tilde{W}(x) + \tilde{b},$$

for all $x \in \mathcal{M}$, where $\tilde{W}$ is as in Hypothesis 2, and $\tilde{a}, \tilde{b} \geq 0$. Then $(M_t^f(x))_{t\geq 0}$ is a true L^2 martingale and f satisfies the strong law.

Proof. By Proposition 1.37***(i)*** and since $W \geq 0$,

$$\frac{1}{t}\int_0^t P_s(\Gamma_e^{\mathcal{M}}(f)(x))ds \leq \frac{\tilde{a}W(x)}{t} + \tilde{a}C,$$

so condition ***(i)*** of Proposition 1.24 is satisfied. Hence the result. □

To ensure that limit points exist, we need the following assumption as a strengthening of Hypothesis 2.

Hypothesis 3. Hypothesis 2 holds and, in addition, either

(a) W satisfies the strong law (14); or

(b) $\tilde{W} = \alpha W$ for some $\alpha > 0$.

The next result states that, under Hypothesis 3, the empirical occupation measure $(\Pi_t^x)_{t\geq 0}$ is almost surely relatively compact and that its limit points are invariant probability measures. Earlier versions of this result are already proved in [SBA11], [EHS15] or [Ben18], and has also been developed during the writing of [BHN+25].

Theorem 1.40. *Assume Hypothesis 3.*

(i) *Under Hypothesis 3 (a), for all $x \in M$, almost surely*

$$\limsup_{t\to\infty} \Pi_t^x(\tilde{W}) \leq C.$$

(ii) *Under Hypothesis 3 (b),*

$$P_t\tilde{W} \leq e^{-\alpha t}(\tilde{W} - C) + C \leq e^{-\alpha t}\tilde{W} + C, \tag{20}$$

and, for all $x \in M$, almost surely

$$\limsup_{t\to\infty} \Pi_t^x\left(\sqrt{\tilde{W}}\right) \leq \frac{2-\sqrt{\rho}}{1-\sqrt{\rho}} C^{1/2},$$

with $\rho = e^{-\alpha}$.

(iii) *For all $x \in M$, $(\Pi_t^x)_{t\geq 0}$ is almost surely tight and every limit point lies in $\mathcal{P}_{\mathrm{inv}}(M)$.*

Remark 1.41. Under Hypothesis 3 both $\mathcal{P}_{\mathrm{inv}}(M)$ and $\mathcal{P}_{\mathrm{inv}}(M_0)$ are non-empty by Theorem 1.40. However, $\mathcal{P}_{\mathrm{inv}}(M_+)$ may still be empty because the set of probability measures supported in the open set M_+ is not closed for the weak topology. Along empirical measures Π_t^x for $x \in M_+$, we can still have $\Pi_{t_k}^x \Rightarrow \mu$ with $\mu(M_0) = 1$ compared to the case of M_0 being closed.

1.3.2 Proof of Theorem 1.40

Before stating the proof of Theorem 1.40, we introduce the following result which is the continuous version of a well-known one for discrete–time Feller chains. The proof is directly based on the discrete-time case.

Proposition 1.42. *For all $x \in M$, almost surely, every limit point of $(\Pi_t^x)_{t\geq 0}$ lies in $\mathcal{P}_{\mathrm{inv}}(M)$, $\mathbb{P}$-almost surely.*

Proof. Let $(Y_n^x)_{n\geq 0}$ be the discrete time chain (7) obtained by sampling the continuous one with G the 1-resolvent kernel defined in (6). We set

$$\widehat{\Pi}_n^x = \frac{1}{n}\sum_{k=0}^{n-1} \delta_{Y_k^x},$$

the empirical measure of the discrete chain $(Y_n^x)_{n\geq 0}$, where δ_z is the Dirac measure that assigns mass 1 to $\{z\}$.

For a discrete–time Feller chain, it is classical that the limit points of the empirical measures are invariant (see e.g. [BH22], Theorem 4.20). Since it is invariant with respect to its kernel G, limit points of $(\widehat{\Pi}_n^x)_{n\geq 0}$ lie almost surely in $\mathcal{P}_{\text{inv}}(M)$ (see e.g. [BH22], Proposition 4.57).

We now show that $(\Pi_t^x)_{t\geq 0}$ and $(\widehat{\Pi}_{\lfloor t\rfloor}^x)_{t\geq 0}$ have the same limit points almost surely, where $\lfloor\cdot\rfloor$ is the floor function. Because M is separable, there exists a countable family $\{f_n\}_{n\geq 0}\subset C_b(M)$ such that

$$D(\mu,\nu) := \sum_{k\geq 0}\frac{1}{2^k}\min(|\mu f_k - \nu f_k|, 1),$$

is a distance on $\mathcal{P}(M)$ that metrizes the topology of weak convergence on $\mathcal{P}(M)$ (see e.g. [BH22], Proposition 4.5). That is,

$$\mu_n \Rightarrow \mu \;\Leftrightarrow\; D(\mu_n,\mu) \underset{n\to\infty}{\longrightarrow} 0.$$

By Proposition 4.58 in [BH22], for every $f\in C_b(M)$,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t f(X_s)ds - \frac{1}{\lfloor t\rfloor}\sum_{k=0}^{\lfloor t\rfloor-1} Gf(Y_k) = 0,$$

almost surely, which is

$$\lim_{t\to\infty}\left|\Pi_t^x(f) - \widehat{\Pi}_{\lfloor t\rfloor}^x Gf\right| = 0 \quad \text{a.s.}$$

Moreover, let

$$M_n := \sum_{k=0}^{n-1} f(Y_{k+1}) - Gf(Y_k),$$

be a bounded martingale since $(M_n)_{n\geq 0}$ is adapted to the natural filtration $(\mathcal{F}_n)_{n\geq 0}$ defined by $\mathcal{F}_n = \sigma(Y_0,\cdots,Y_n)$, $\forall n\geq 0$, and the increments $D_{k+1} := f(Y_{k+1}) - Gf(Y_k)$ has zero expectation, so

$$\mathbb{E}(M_{n+1}|\mathcal{F}_n) = \mathbb{E}(M_n + D_{n+1}|\mathcal{F}_n) = M_n + \mathbb{E}\left[f(Y_{n+1}) - Gf(Y_n)|\mathcal{F}_n\right] = M_n + Gf(Y_n) - Gf(Y_n) = M_n.$$

By the strong law of large number for martingales with bounded increments (see e.g. [HH80], Theorem 2.18), it follows that

$$\lim_{n\to\infty}\left|\widehat{\Pi}_n^x f - \widehat{\Pi}_n^x Gf\right| = \lim_{n\to\infty}\frac{M_n}{n} = 0, \quad \mathbb{P}-\text{a.s.}$$

We now obtain

$$\lim_{t\to\infty}\left|\hat{\Pi}_t^x(f) - \hat{\Pi}_{\lfloor t\rfloor}^x f\right| \leq \lim_{t\to\infty}\left|\Pi_t^x(f) - \hat{\Pi}_{\lfloor t\rfloor}^x Gf\right| + \left|\hat{\Pi}_{\lfloor t\rfloor}^x Gf - \hat{\Pi}_{\lfloor t\rfloor}^x f\right| = 0,$$

therefore by dominated convergence, $\lim_{t\to\infty} D(\Pi_t^x, \widehat{\Pi}_{\lfloor t\rfloor}^x) = 0$ almost surely and the two families share the same limit points, which must be invariant. □

Proof of Theorem 1.40. ***(i)*** From inequality (19) we have

$$\Pi_t^x(\tilde{W}) \leq \frac{W(x)}{t} + C + \frac{M_t}{t},$$

where $(M_t)_{t\geq 0}$ is the local martingale defined in (18). Under Hypothesis 3(a), the strong law implies

$$\frac{M_t}{t} \underset{t\to\infty}{\rightarrow} 0, \quad a.s.$$

so $\limsup_{t\to\infty} \Pi_t^x(\tilde{W}) \le C$ for all $x \in M$, almost surely.

(ii) Set $V_t = e^{\alpha t}(W(X_t^x) - \frac{C}{\alpha})$. Because $(M_t)_{t\ge 0}$ is a local martingale, the process $(N_t)_{t\ge 0}$ defined by

$$N_t = V_t - V_0 - \int_0^t e^{\alpha s}(\alpha W(X_s^x) - C + \mathcal{L}_e W(X_s^x))\mathrm{d}s$$

is also a local martingale with $N_0 = 0$ (see e.g. [EK86], Proposition 3.2). Hypothesis 3(b) yields

$$V_t - V_0 \le N_t,$$

since $\mathcal{L}_e W + \alpha W - C \le 0$. For all $n \ge 0$, $(N_t^n)_{t\ge 0}$ defined as $N_t^n = N_{t\wedge n}$ is a local martingale bounded below by $\frac{C}{\alpha}(1 - e^{\alpha n}) - W(x)$.

By the same argument as in Proposition 1.37, applying Fatou's Lemma to $\mathbb{E}(\lim_{k\to\infty} N_{t\wedge\tau_k}^n) = \mathbb{E}(N_t^n)$ where $(\tau_k)_{k\ge 0}$ is the sequence of stopping times such that $(N_{t\wedge\tau_k}^n)_{t\ge 0}$ is a martingale, we have

$$\mathbb{E}(N_t^n) \le \mathbb{E}(N_0^n) = 0 \;\Rightarrow\; \mathbb{E}(V_{t\wedge n} - V_0) \le \mathbb{E}(N_t^n) \le 0,$$

which implies that

$$P_{t\wedge n}\tilde{W} \le e^{-\alpha(t\wedge n)}(\tilde{W} - C) + C \le e^{-\alpha(t\wedge n)}\tilde{W} + C,$$

and letting $n \to \infty$ gives (20).

Define $\nu_n = \frac{1}{n}\sum_{k=0}^{n-1} \delta_{X_k^x}$. By (20), it implies that

$$P_1\tilde{W} \le \rho\tilde{W} + C,$$

with $\rho = e^{-\alpha}$. Using [BH22], Corollary 4.23, we have

$$\limsup_{n\to\infty} \nu_n\left(\sqrt{\tilde{W}}\right) \le \frac{\sqrt{C}}{1 - \sqrt{\rho}}. \tag{21}$$

Writing

$$\Pi_n^x\left(\sqrt{\tilde{W}}\right) = \frac{1}{n}\int_0^n \sqrt{\tilde{W}(X_s^x)}ds = \frac{1}{n}\sum_{k=0}^{n-1}[\Delta_{k+1} + \int_0^1 P_s\sqrt{\tilde{W}(X_k^x)}\mathrm{d}s],$$

where

$$\Delta_{k+1} = \int_0^1 \left(\sqrt{\tilde{W}(X_{k+s}^x)} - P_s\sqrt{\tilde{W}(X_k^x)}\right)ds.$$

By Jensen's inequality, for all $s \ge 0$, we have

$$P_s\left(\sqrt{\tilde{W}}\right) \le \sqrt{P_s(\tilde{W})} \le e^{-\frac{\alpha s}{2}}\sqrt{\tilde{W}} + \sqrt{C} \le \sqrt{\tilde{W}} + \sqrt{C},$$

so that $\int_0^1 P_s\left(\sqrt{\tilde{W}}\right)ds \le \sqrt{\tilde{W}} + \sqrt{C}$. Thus,

$$\Pi_n^x\left(\sqrt{\tilde{W}}\right) \le \frac{1}{n}\sum_{k=0}^{n-1}\Delta_{k+1} + \nu_n\sqrt{\tilde{W}} + \sqrt{C}.$$

We now claim that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\Delta_{k+1} = 0,$$

$\mathbb{P}-$a.s. Combined with (21), it implies that

$$\limsup_{n\to\infty} \Pi_n^x\left(\sqrt{\tilde{W}}\right) \le \left(1 + \frac{1}{1-\sqrt{\rho}}\right)\sqrt{C},$$

which concludes the proof. To prove the claim, using Fubini's theorem and Markov property, we observe that

$$\mathbb{E}(\Delta_{k+1} \mid \mathcal{F}_k) = \mathbb{E}\left(\int_0^1 \left(\sqrt{\tilde{W}(X_{k+s}^x)} - P_s\sqrt{\tilde{W}(X_k^x)}\right) ds \mid \mathcal{F}_k\right) = 0,$$

and

$$\mathbb{E}(\Delta_{k+1}^2 \mid \mathcal{F}_k) \le \int_0^1 \mathbb{E}\left[\left(\sqrt{\tilde{W}(X_{k+s}^x)} - P_s\sqrt{\tilde{W}(X_k^x)}\right)^2 \mid \mathcal{F}_k\right] ds,$$

by applying Cauchy-Schwarz and Fubini's theorem. Finally, recalling that the conditional variance is defined by

$$\mathrm{Var}(Z \mid \mathcal{F}) = \mathbb{E}\left[(Z - \mathbb{E}(Z \mid \mathcal{F}))^2 \mid \mathcal{F}\right],$$

and with $\mathrm{Var}(Z \mid \mathcal{F}) \le \mathbb{E}(Z^2 \mid \mathcal{F})$, we have

$$\begin{aligned}
\mathbb{E}[\Delta_{k+1}^2 \mid \mathcal{F}_k] &\le \int_0^1 \mathbb{E}\left[\left(\sqrt{\tilde{W}(X_{k+s}^x)} - \mathbb{E}\left[\sqrt{\tilde{W}(X_{k+s}^x)} \mid \mathcal{F}_k\right]\right)^2 \mid \mathcal{F}_k\right] ds \\
&\le \int_0^1 \mathbb{E}[\tilde{W}(X_{k+s}^x) \mid \mathcal{F}_k] ds \\
&= \int_0^1 P_s\tilde{W}(X_k^x) ds \\
&\le \tilde{W}(X_k^x) + C,
\end{aligned}$$

where

$$\mathrm{Var}\left(\sqrt{\tilde{W}(X_{k+s}^x)} \mid \mathcal{F}_k\right) = \mathbb{E}\left[\left(\sqrt{\tilde{W}(X_{k+s}^x)} - \mathbb{E}\left[\sqrt{\tilde{W}(X_{k+s}^x)} \mid \mathcal{F}_k\right]\right)^2 \mid \mathcal{F}_k\right],$$

while the last inequality follows (20) and the fact that $e^{-\alpha} \ge 1 - \alpha$.

Using again previous bound together with (20), we have

$$\sum_{k=0}^n \mathbb{E}[\Delta_{k+1}^2] = \sum_{k=0}^n \mathbb{E}[\mathbb{E}[\Delta_{k+1}^2 | \mathcal{F}_k]] \le \sum_{k=0}^n P_k\tilde{W}(x) + nC \le \sum_{k=0}^n e^{-\alpha k}\tilde{W}(x) + 2nC.$$

By the strong law of large number for discrete time martingale, it is a sufficient condition to prove the claim (see e.g. [BH22], Theorem A.8 (iv)).

(iii) Either the lim sup bound from part ***(i)*** or ***(ii)*** combined with Lemma 1.35 imply that $(\Pi_t^x)_{t\ge 0}$ is almost surely tight. Proposition 1.42 tells us that any limit point is invariant, which concludes the proof. □

Chapter 2

Ecological Stochastic Differential Equation

This chapter is dedicated to the study of ecological SDEs, with a focus on Kolmogorov SDEs and many practical criteria to verify the hypotheses introduced in Chapter 1. Also, we will introduce the notions of *stochastic persistence and nonpersistence*, explain how to verify them, and the conclusions that can be reached under those new hypotheses. More precisely, Chapter 2 is organized as follows:

(i) Section 2.1 presents the setting of Kolmogorov Stochastic Differential Equations as well as a practical criterion to ensure Hypotheses 1-3 hold. It is a fundamental result since it also describes the sets $\mathcal{D}$, $\mathcal{D}^2$, $\mathcal{D}_e$, and $\mathcal{D}_e^2$ together with $\mathcal{L}$, Γ, $\mathcal{L}_e$, and Γ_e (and their restrictions to M_+).

(ii) Section 2.2 introduces the property of *stochastic persistence* and a practical criterion for Kolmogorov SDEs to achieve it through the related notions of *$H-$exponents* and *$H-$persistence* together with the additional Hypothesis 4.

(iii) Section 2.3 outlines how the $H-$persistence combined with the existence of an accessible point in the non-extinction set, which satisfies the Hörmander condition (respectively strong Hörmander condition), is related to the uniqueness of the invariant measure and the almost-sure convergence of the empirical measure (respectively the convergence in Total variation of $(P_t(x,\cdot))_{t\geq 0}$) to it.

(iv) Section 2.4 details the conditions to ensure either an exponential or a polynomial rate of convergence, under an additional Hypothesis 5 or under Hypothesis 6.

(v) Section 2.5 brings in a new extinction theory by defining the notion of *nonpersistence* together with additional Hypotheses 7 and 8. A practical criterion to ensure them is also proposed.

2.1 Kolmogorov Stochastic Differential Equations

From now on, we are interested by Kolmogorov Stochastic Differential Equations, which describes a system of SDEs on $M = \mathbb{R}_+^n$ of the form

$$dx_i(t) = x_i(t)\left(F_i(x(t))\mathrm{d}t + \sum_{j=1}^{m} \Sigma_i^j(x(t))\mathrm{d}B_t^j\right), \quad i = 1, \ldots, n, \tag{22}$$

where F_i, Σ_i^j are real valued locally Lipschitz maps on M, with F_i representing the per-capita growth rates of the species in absence of noise, and $(B_t^1, \cdots, B_t^m)_{t\geq 0}$ an $m-$dimensional standard Brownian motion that models the environmental noise affecting the growth rates.

For simplicity, we assume that Σ_i^j is bounded. This assumption can be relaxed under other conditions such as described in [HN18].

Let $x = (x_1, x_2, \ldots, x_n)$ denote the densities of n species or populations. For any $I \subset \{1, \ldots, n\}$, we set

$$M_0^I := \partial\mathbb{R}_+^{n,I} = \Big\{x \in M : \prod_{i\in I} x_i = 0\Big\}, \tag{23}$$

be the set corresponding to the extinction of at least one of the species $i \in I$, and

$$M_+^I := \{x \in M : x_i > 0,\ \forall i \in I\}. \tag{24}$$

For simplicity, when $I = \{1, \cdots, n\}$, we let $M_0^I = M_0$ and $M_+^I = M_+$.

We let $a(x)$ denote the positive semidefinite matrix defined by

$$a_{ij}(x) = \sum_{k=1}^{m} \Sigma_i^k(x)\Sigma_j^k(x). \tag{25}$$

Given (22) evolving on M, for all twice continuously differentiable functions $f\colon M \to \mathbb{R}$ we define

$$Lf(x) = \sum_{i=1}^{n} x_i F_i(x) \frac{\partial f}{\partial x_i}(x) + \frac{1}{2} \sum_{i,j=1}^{n} x_i x_j a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x), \tag{26}$$

and

$$\Gamma_L(f)(x) = \sum_{i,j=1}^{n} x_i x_j a_{ij}(x) \frac{\partial f}{\partial x_i}(x) \frac{\partial f}{\partial x_j}(x). \tag{27}$$

The next proposition ensures that Hypotheses 1 and 3 (hence 2) hold.

Proposition 2.1. *Assume there exists a C^2 proper map $U\colon M \to [1, \infty)$ and constants $a > 0$, $b \geq 0$ such that*

$$LU \leq -aU + b. \tag{28}$$

Then the following properties hold:

- ***(i)*** *For each $x \in M$, there exists a unique (strong) solution $(X_t^x)_{t\geq 0} \subset M$ to (22) with $X_0^x = x$, and X_t^x is continuous in (t, x). In particular, the $C_b(M)$–Feller continuity holds.*
- ***(ii)*** *Hypothesis 1 holds with $M_0^I = \{x \in M : \prod_{i\in I} x_i = 0\}$, for all $I \subset \{1, \ldots, n\}$.*
- ***(iii)*** *For all $f \in C^2(M)$ we have $f \in \mathcal{D}_e^2$, $\mathcal{L}_e f = Lf$ and $\Gamma_e f = \Gamma_L f$. In particular, Hypothesis 3(b) holds with $W = U$.*
- ***(iv)*** *For all $f \in C_c^2(M)$[1], $f \in \mathcal{D}^2$, $\mathcal{L}f = Lf$ and $\Gamma(f) = \Gamma_L(f)$.*
- ***(v)*** *If, in addition to (28), there exists $0 \leq \eta < 1$ such that*

$$\Gamma_L(U) \leq c(U^{2+\eta}) \tag{29}$$

for some constant $c \geq 0$, then Hypothesis 3(a) holds with $W = U^{\frac{1-\eta}{2}}$ and $\tilde{W} = (1 + cst)W$.

This result is a slightly improved version of Proposition 3.2 in [Ben18]. In particular, we use the notion of extended carré du champ defined in Definition 1.21 to detail the outcomes.

Proof. ***(i)*** Since the drift and the covariance are supposed to be locally Lipschitz continuous, we can use classical results on stochastic differential equations such that there exists for any $x \in \mathbb{R}^n$ a unique continuous process $(X_t)_{t\geq 0}$ defined on some interval $[0, \tau^x[$ solution to (22), with initial condition $X_0 = x$ and such that $t < \tau^x \iff \|X_t\| < \infty$. (see e.g. [RY04], Exercise 2.10 and [Hsu02], Theorem 1.1.8). We denote it $(X_t^x)_{0\leq t<\tau^x}$.

Furthermore, applying Itô's formula to $Y_{t,i}^x := \ln(X_{t,i}^x)$, rearranging the terms, and using the uniqueness of the solutions, then

$$X_{t,i}^x = x_i \exp\Big(\int_0^t [F_i(X_s^x) - \tfrac{1}{2} a_{ii}(X_s^x)] ds + \sum_{j=1}^{m} \int_0^t \Sigma_i^j(X_s) dB_s^j\Big).$$

[1] $C_c^2(M)$ denotes the set of C^2 functions in M with compact support in the sense that this is the restriction to M of a C^2 function $f : \mathbb{R}^n \to \mathbb{R}$ with compact support.

Thus

$$x_i > 0 \Longrightarrow X^x_{t,i} > 0,\ \forall t \in [0, \tau^x[, \tag{30}$$

and

$$x_i = 0 \Longrightarrow X^x_{t,i} = 0\ \forall t \in [0, \tau^x[. \tag{31}$$

Now, let's prove that $\tau^x = \infty$. For any C^2 function $\psi\colon M \to \mathbb{R}$, by Itô's formula,

$$\psi(X^x_t) - \psi(x) - \int_0^t L\psi(X^x_s)ds = \sum_{i=1}^n \Big(\int_0^t \frac{\partial \psi}{\partial x_i}(X^x_s)\big[X^x_{s,i}\sum_{j=1}^m \Sigma^j_i(X^x_s)dB^j_s\big]\Big). \tag{32}$$

Let $\tau^x_k = \inf\{t \geq 0 : U(X^x_t) \geq k\},\ k \in \mathbb{N}$. By assumption (28) on U, for all $x \in M$,

$$\begin{aligned}
k\mathbb{P}\left(\tau^x_k \leq t\right) &= \mathbb{E}\left[U\left(X^x_{\tau^x_k}\right)\mathbf{1}_{\{\tau^x_k \leq t\}}\right]\\
&\leq \mathbb{E}\left[U\left(X^x_{t\wedge\tau^x_k}\right)\right]\\
&= U(x) + \mathbb{E}\left[\int_0^{t\wedge\tau^x_k} LU\left(X^x_s\right)ds\right]\\
&\leq U(x) - a\mathbb{E}\left[\int_0^{t\wedge\tau^x_k} U\left(X^x_s\right)ds\right] + bt\\
&\leq U(x) + bt.
\end{aligned}$$

Hence

$$\mathbb{P}(\tau^x \leq t) = \mathbb{P}\Big(\cap_{k\geq 0}\{\tau^x_k \leq t\}\Big) = \lim_{k\to\infty} \mathbb{P}(\tau^x_k \leq t) = 0,$$

proving that $\tau^x = \infty$ almost surely.

Let $(P_t)_{t\geq 0}$ denote the Markov semigroup acting on bounded measurable functions $f\colon M \to \mathbb{R}$, defined by $P_t f(x) = \mathbb{E}(f(X^x_t))$. Then, $C_b(M)$–Feller continuity just follows from the dominated convergence theorem and the continuity in x of the solution $(X^x_t)_{t\geq 0}$, which implies that $x \mapsto P_t f(x)$ is continuous for every bounded continuous f.

(ii) It follows automatically from (30) and (31).

(iii) Let $f \in C^2(M)$, then the property $\mathcal{L}_{\mathrm{e}}(f) = Lf$ follows from Itô's formula (32) since the process

$$\sum_{i=1}^n \Big(\int_0^t \frac{\partial \psi}{\partial x_i}(X^x_s)\big[X^x_{s,i}\sum_{j=1}^m \Sigma^j_i(X^x_s)dB^j_s\big]\Big),$$

is a local martingale given $\psi \in C^2(M)$, by continuity of $x \mapsto X^x_t$ and the non-explosion of the solution.

For Γ_{e}, let $f \in C^2(M)$. Since

$$\begin{aligned}
\mathcal{L}_e(f^2) &= \sum_{i=1}^n x_i F_i(x)\frac{\partial f^2}{\partial x_i}(x) + \sum_{i,j=1}^n x_i x_j a_{ij}(x)\frac{\partial^2 f^2}{\partial x_i \partial x_j}(x)\\
&= \sum_{i=1}^n x_i F_i(x) 2f(x)\frac{\partial f}{\partial x_i}(x) + \sum_{i,j=1}^n x_i x_j a_{ij}(x)\left(2\frac{\partial f}{\partial x_i}(x)\frac{\partial f}{\partial x_j}(x) + 2f(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right),
\end{aligned}$$

and

$$2f\mathcal{L}_e(f) = \sum_{i=1}^n x_i F_i(x) 2f(x)\frac{\partial f}{\partial x_i}(x) + \sum_{i,j=1}^n x_i x_j a_{ij}(x) 2f(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x),$$

it follows that

$$\Gamma_e(f)(x) = \mathcal{L}_e(f^2) - 2f\mathcal{L}_e(f) = \sum_{i,j=1}^{n} x_i x_j a_{ij}(x)\frac{\partial f}{\partial x_i}(x)\frac{\partial f}{\partial x_j}(x) = \Gamma_L(f)(x).$$

In particular, since assumption (28) is satisfied, Hypothesis 3(b) is automatically satisfied since Hypothesis 2 holds with $W = U$ and $\tilde{W} = aU$.

(iv) Let $\psi \in C_c^2(M)$. By Itô's formula (32), $\psi(X_t^x) - \psi(x) - \int_0^t L\psi(X_s^x)ds$ is a martingale since the integrand on the right-side is bounded when ψ has compact support. Thus, by taking expectations and by Fubini's theorem, we obtain

$$P_t\psi(x) - \psi(x) = \int_0^t P_s(L\psi)(x)ds \;\Rightarrow\; |P_t\psi(x) - \psi(x)| \leq t\|L\psi\|_\infty,$$

thus

$$\mathcal{L}\psi(x) = \lim_{t\to 0}\frac{P_t\psi(x) - \psi(x)}{t} = \lim_{t\to 0}\frac{1}{t}\int_0^t P_s(L\psi(x))ds = L\psi(x),$$

by dominated convergence since $L\psi$ is bounded when $\psi \in C_c^2(M)$ as in Remark 1.10.

Replacing ψ by ψ^2 shows $\psi \in \mathcal{D}^2$ and $\Gamma(\psi) = \Gamma_L(\psi)$.

(v) For any smooth function $h\colon \mathbb{R}_+ \to \mathbb{R}$,

$$L(h(U)) = h'(U)LU + \tfrac{1}{2}h''(U)\Gamma_L(U).$$

If h is concave and nondecreasing, this gives

$$L(h(U)) \leq h'(U)LU \leq -ah'(U)U + bh'(U).$$

Set $h(t) = t^{\frac{1-\eta}{2}}$ and $W = h(U)$. Then $h'(t)t = \frac{1-\eta}{2}h(t)$ and $h'(t) = \frac{1-\eta}{2}t^{-\frac{1+\eta}{2}}$ so

$$L(W) \leq \tfrac{1-\eta}{2}(-aW + b),$$

since $U \geq 1$ by assumptions. Now we get

$$\Gamma_e(W) = h'(U)^2\Gamma_e(U) = \left(\tfrac{1-\eta}{2}\right)^2 U^{-\eta-1}\Gamma_e(U).$$

Thus, we use assumption (29) to get

$$\Gamma_e(W) \leq c(1 + U).$$

Therefore, Hypothesis 3(a) holds by Corollary 1.39. □

2.2 Stochastic persistence

From Remark 1.41, we now want to ensure that $\mathcal{P}_{\text{inv}}(M_+^I)$ is non-empty. To this effect, we introduce the notion of *stochastic persistence*.

The following definition is given in a general context: it is inspired by the work of Chesson in [Che78] and [Che82], follows the presentation of Schreiber in [Sch12].

Recall that for any $I \subset \{1, \cdots, n\}$, M_0^I denotes the extinction set of species $i \in I$, as in (23), while M_+^I is the non-extinction set of species $i \in I$, as in (24).

Definition 2.2 (**Stochastic persistence**)**.** *The family of processes* $\{(X_t^x)_{t\geq 0} : x \in M_+^I\}$ *is called* stochastically persistent *(with respect to* M_0^I*) if for every* $\varepsilon > 0$ *there exists a compact set* $K_\varepsilon \subset M_+^I$ *such that, for all* $x \in M_+^I$*,*

$$\mathbb{P}\Big(\liminf_{t\to\infty} \Pi_t^x(K_\varepsilon) \geq 1 - \varepsilon\Big) = 1.$$

where $\Pi_t^x(\cdot) = \frac{1}{t}\int_0^t \delta_{X_s^x}(\cdot)ds$ *is the empirical occupation measures defined in (16).*

Analyzing stochastic persistence requires control of the process near the extinction set. We introduce an additional Lyapunov–type assumption to this end.

Hypothesis 4. There exist continuous maps $V\colon \in M_+^I \to \mathbb{R}$ and $H\colon M \to \mathbb{R}$ such that

(i) $V \in \mathcal{D}_e^I$ and $\mathcal{L}_{\mathrm{e}}^I(V) = H|_{M_+^I}$;

(ii) V satisfies the strong law (Definition 1.23);

(iii) $\frac{\tilde{W}}{1+|H|}$ is proper, where $\tilde{W}$ is like in Hypothesis 3.

In some arguments, to control higher moments of H, we need the stronger condition

(iii′) $|H|^q \leq \mathrm{cst}(1 + \tilde{W})$ for some $q > 1$.

Under ***(iii)***, it implies that $\tilde{W}$ dominates $|H|$ outside a compact set, which implies $H \in L^1(\mu)$ for all $\mu \in \mathcal{P}_{\mathrm{inv}}(M)$ thanks to Proposition 1.37. Then, the following definition is meaningful.

Definition 2.3 (***H*-exponents**)**.** *Given V and H as in Hypothesis 4, we define the* lower *(respectively* upper*)* H–exponents *of* $(X_t)_{t\geq 0}$ *by*

$$\Lambda^-(H) := -\sup\{\mu H : \mu \in \mathcal{P}_{\mathrm{erg}}(M_0^I)\}, \quad \Big(\textit{respectively } \Lambda^+(H) := -\inf\{\mu H : \mu \in \mathcal{P}_{\mathrm{erg}}(M_0^I)\}\Big).$$

Remark 2.4. The key point in Hypothesis 4 is that H is defined on the whole space M, whereas V is only defined on M_+^I and typically $V(x) \to \infty$ as x approaches M_0^I. If, in fact, one could extend V to all of M while keeping condition ***(i)*** of Hypothesis 4 valid on M, then

$$\Lambda^-(H) = \Lambda^+(H) = 0,$$

since the integral of the infinitesimal generator against any invariant probability measure is 0, so $\mu H = 0$ on the whole M.

Definition 2.5 (***H*-persistence**)**.** *The family* $\{(X_t^x)_{t\geq 0} : x \in M_+^I\}$ *is called* H-persistent *if there exist* (V, H) *satisfying Hypothesis 4 with V positive and* $\Lambda^-(H) > 0$*.*

Our main result is that $H-$persistence implies the stochastic persistence, and in particular the almost-sure convergence of $(\Pi_t^x)_{t\geq 0}$ whose limit points lie in $\mathcal{P}_{\mathrm{inv}}(M_+^I)$.

Theorem 2.6 ([Ben18], Theorem 4.4)**.** *Assume Hypotheses 1–4 hold and that the family* $\{(X_t^x)_{t\geq 0} : x \in M_+^I\}$ *is H-persistent. Then*

(i) *For every* $x \in M_+^I$*, every weak limit point of* $(\Pi_t^x)_{t\geq 0}$ *lies in* $\mathcal{P}_{\mathrm{inv}}(M_+^I)$ *a.s.;*

(ii) *The process is stochastically persistent in the sense of Definition 2.2.*

This theorem has the following immediate consequence.

Corollary 2.7 ([Ben18], Corollary 4.5)**.** *Assume that the conditions of Theorem 2.6 hold, and in addition that* $\mathcal{P}_{\mathrm{inv}}(M_+^I)$ *contains at most one probability measure. Then* $\mathcal{P}_{\mathrm{inv}}(M_+^I)$ *consists of a single measure, denoted* $\{\Pi\}$*, and for every* $x \in M_+^I$*,*

$$\Pi_t^x \Longrightarrow \Pi \quad \textit{almost surely as } t \to \infty.$$

When $I = \{1, \cdots, n\}$, we call Π the *persistent measure*. In ecological models, Π characterises the long-term behavior of the coexisting species.

2.2.1 How to achieve stochastic persistence in practice?

Following the ideas from [BM24] and [Ben18], inspired by [Sch12], [SBA11] and [BHS08], we adapt the notion of *invasion* in the particular case of Kolmogorov SDE described in (22). For all $i \in \{1, \cdots, n\}$, let

$$\lambda_i(x) = F_i(x) - \frac{a_{ii}(x)}{2}, \tag{33}$$

be the invasion of species i with respect to x. For any $\mu \in \mathcal{P}_{\text{erg}}(M_0^I)$, we write

$$\mu\lambda_i := \int_M \lambda_i(x)\mu(\mathrm{d}x), \tag{34}$$

whenever $\lambda_i \in L^1(\mu)$. We call $\mu\lambda_i$ the *mean invasion rate* of species i with respect to μ. The next theorem extends the Hofbauer criterion (see e.g. the earlier work in [Hof81]) to possibly degenerate SDEs while earlier results such as [HN18] are limited to nondegenerate models.

Theorem 2.8 ([Ben18], Theorem 5.1)**.** *Let $(X_t^x)_{t\geq 0}$ be the process generated by (22) with $X_0 = x \in M_+^I$ and let U, η be as in Proposition 2.1. We assume*

$$\limsup_{\|x\|\to\infty} \frac{U^{\frac{1-\eta}{2}}(x)}{1 + \sum_{i\in I} |F_i(x)|} = \infty, \tag{35}$$

(i) *For every $\mu \in \mathcal{P}_{\text{erg}}(M_0^I)$ and $i \in \{1, \cdots, n\}$, one has $\lambda_i \in L^1(\mu)$ and*

$$\mu\lambda_i \neq 0 \quad \Longrightarrow \quad \text{supp}(\mu) \subset M_0^i := \{x \in M : x_i = 0\}.$$

(ii) *Suppose there exist positive numbers $\{p_i\}_{i\in I}$ such that*

$$\sum_{i\in I} p_i\mu\lambda_i > 0, \quad \textit{for all } \mu \in \mathcal{P}_{\text{erg}}(M_0). \tag{36}$$

Then $\{(X_t^x)_{t\geq 0} : x \in M_+^I\}$ is H–persistent.

Proof. ***(i)*** Since μ is supposed to be an ergodic probability measure, it follows from Birkhoff ergodic Theorem that $\Pi_t^x \Rightarrow \mu$ for μ almost every x and $\mathbb{P}_x$−almost surely. By Proposition 2.1***(v)***, $U^{\frac{1-\eta}{2}}$ satisfies Hypothesis 3***(a)*** and combining Theorem 1.40***(i)*** with Lemma 1.35 leads to

$$\mu\left(U^{\frac{1-\eta}{2}}\right) < \infty.$$

We now use assumption (35) to conclude that $\mu\lambda_i < \infty$ for all $i \in I$ and $\forall \mu \in \mathcal{P}_{\text{erg}}(M_0^I)$.

Let $i \in I$ and remark that $\lambda_i(x) = L(\log(x_i))$, so that the local martingale induced by $\log(x_i)$ verifies

$$\frac{1}{t}\int_0^t \lambda_i(x_i^x(s))ds = \frac{\log(x_i^x(t))}{t} - \frac{\log(x_i^x(0))}{t} - \frac{M_t^{\log(x_i)}(x)}{t}.$$

In particular, $\Gamma_e^i(\log(x_i)) \leq cst$ so that $\log(x_i)$ satisfies the strong law and for $\mu \in \mathcal{P}_{\text{erg}}(M_0^{(i)})$,

$$\mu(\lambda_i) = \limsup_{t\to\infty} \frac{1}{t}\int_0^t \lambda_i(x_i^x(s))ds = \limsup_{t\to\infty} \frac{\log(x_i^x(t))}{t}, \quad \forall x \in M_+^{(i)}$$

For $m \geq 1$, let $K_m = \left\{x \in M : x_i > \frac{1}{m}\right\} \cap [-m, m]$, so that $M_+^{(i)} = \cup_{m\geq 1} K_m$ and $M_0^{(i)} = \cap_{m\geq 1} K_m^c$. Since $\mu(M_0^{(i)}) = 0$, there exists $m^* \geq 1$ such that $\mu(K_m) \geq \frac{1}{2}$ and by Birkhoff ergodic Theorem, $(x_i^x(t))_{t\geq 0}$ visits K_{m^*} infinitely often, for μ−almost every x, $\mathbb{P}_x$−almost surely.

Since $\log(x_i)$ is bounded on K_{m^*}, then $\mu\lambda_i = 0$ on K_{m^*}, which extends to $M_+^{(i)} = \cup_{m\geq 1} K_m$. By taking the contrapositive, it follows that $\mu\lambda_i \neq 0$ implies $\mathcal{P}_{\text{erg}}(M_0^{(i)})$ and $\text{supp}(\mu) \subset M_0^{(i)}$.

(ii) We now suppose (36) and we show that conditions from Definition 2.5 are verified. Let $h(u) = \log\left(\frac{1}{u}\right)$, $v : \mathbb{R} \to \mathbb{R}_+$ a C^∞ function with bounded v', v'' such that $v(t) = t$, $\forall t \geq 1$, and

$$V(x) := v\left(\sum_{i\in I} p_i h(x_i)\right).$$

Remark that V coincides with $\sum_{i\in I} p_i h(x_i)$ on the subset $\{x \in M_+^I : \sum_{i\in I} p_i h(x_i) > 1\}$. In particular, $V(x) \to \infty$ as $x_i \to 0$ for $i \in I$ since $h(u) \underset{u\to 0}{\longrightarrow} \infty$ and $v(t) \underset{t\to\infty}{\longrightarrow} \infty$, so that V is not defined on M_0^I.

From Proposition 2.1, it yields

$$H|_{M_+^I}(x) = LV(x) = v'\left(\sum_{i\in I} p_i h(x_i)\right)\left(-\sum_{i\in I} p_i\lambda_i(x)\right) + \frac{1}{2}v''\left(\sum_{i\in I} p_i h(x_i)\right)\langle a(x)p, p\rangle_{\mathbb{R}_n},$$

for all $x \in M_+^I$.Then, $H|_{M_+^I}$ coincides $-\sum_{i\in I} p_i\lambda_i(x))$ on $\{x \in M_+^I : \sum_{i\in I} p_i h(x_i) > 1\}$.

Moreover, H extends continuously on M_0^I. Indeed, let $i \in I$ such that $x_i \to 0$. Since $h(x_i) \to \infty$, it yields $\sum_{i\in I} p_i h(x_i) > 1$ and in particular, H coincides with $-\sum_{i\in I} p_i\lambda_i(x)$ on M_0^I.

Then, $|H| \leq cst\left(\sum_{i\in I} p_i|\lambda_i| + \sum_{i\in I} p_i^2\right)$ which implies that $\frac{U^{\frac{1-\eta}{2}}}{1+|H|}$ is proper since

$$\frac{U^{\frac{1-\eta}{2}}}{1+|H|} \geq C \cdot \frac{U^{\frac{1-\eta}{2}}}{1+\sum_{i\in I}|F_i|}.$$

In particular, we also showed that $V \in \mathcal{D}_e^{2,I}$ since V is C^2 on M_+. Moreover, V is a positive continuous function by construction and

$$\Gamma_e V(x) = \sum_{i,j\in I} a_{ij}(x)p_i p_j \leq cst,$$

since Σ_i^j is bounded, so V satisfies the strong law by Corollary 1.39.

Then, for any $\mu \in \mathcal{P}_{erg}(M_0^I)$, condition (36) implies that $\mu H = -\sum_{i\in I} p_i\mu\lambda_i < 0$ since μ charges M_0^I and $\{(X_t^x)_{t\geq 0} : x \in M_+^I\}$ is H–persistent.

□

Remark 2.9. The proof in [Ben18] relies on the carré du champ operator rather the extended version, which adds some complexity that we prefer to set aside. In particular, the form of V in [Ben18] has to be with compact support, since Γ is well-defined for function on $C_c^2(M)$ as in Proposition 2.1, ***(iv)***: in our version, we only rely on $C^2(M)$ functions for which Γ_e makes sense, and the construction of V is simpler.

2.3 Almost-sure convergence and convergence in total variation

A sufficient (though not necessary) condition for $\mathcal{P}_{\text{inv}}(M_+^I)$ to contain at most one measure is ψ – irreducibility in the sense of Meyn–Tweedie (see e.g. [MT09], Section 4.2). A practical criterion implying ψ–irreducibility is the existence of an accessible weak Doeblin point. The following result is verified in a more general context than Kolmogorov SDEs:

Proposition 2.10 ([Ben18], Proposition 4.8)**.** *Assume there exists a weak Doeblin point $x^* \in \Gamma_{M_+^I}$ (hence $\Gamma_{M_+^I} \neq \emptyset$). Then:*

(i) *$\mathcal{P}_{\text{inv}}(M^I_+)$ contains at most one measure.*

(ii) *In addition, if the process is H–persistent, then $\mathcal{P}_{\text{inv}}(M^I_+) = \{\Pi\}$, $\Pi^x_t \Rightarrow \Pi$ almost surely for all $x \in M^I_+$, and for every $f \in L^1(\Pi)$ such that $\int_0^T f(X^x_s)ds < \infty$ for all $T > 0$ and for all $x \in M^I_+$,*

$$\lim_{t\to\infty} \Pi^x_t f = \Pi f, \quad \textit{almost surely.}$$

Using the Stratonovich formalism, (22) writes as

$$\mathrm{d}x_t = S^0(x_t)\mathrm{d}t + \sum_{j=1}^{m} S^j(x_t) \circ \mathrm{d}B^j_t \tag{37}$$

where, for all $j = 1, \dots, m$ and $i = 1, \dots, n$,

$$S^j_i(x) = x_i \Sigma^j_i(x), \quad \text{and} \quad S^0_i(x) = x_i F_i(x) - \frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{n} \frac{\partial S^j_i(x)}{\partial x_k} S^j_k(x).$$

We associate to this system the deterministic control system

$$\dot{y}(t) = S^0(y(t)) + \sum_{j=1}^{m} u_j(t) S^j(y(t)) \tag{38}$$

where the control function $u = (u_1, \dots, u_m) : \mathbb{R}_+ \to \mathbb{R}^m$ can be chosen to be piecewise continuous. Given such a control function, we let $y(u, x, \cdot)$ denote the maximal solution to (38) starting at x in the sense that $y(u, x, 0) = x$ (without any assumption of global integrability of the vector fields).

The following proposition easily follows from the well-known Stroock and Varadhan support theorem in [SV72] (see also Theorem 8.1, Chapter VI in [IW81]).

Proposition 2.11. *Let $x \in M$. A point $p \in M$ lies in Γ_x if and only if for every neighborhood O of p there exists a control u such that $y(u, x, \cdot)$ meets O (i.e. $y(u, x, t) \in O$ for some $t \geq 0$).*

Recall that the Lie bracket of two smooth vector fields $Y, Z : \mathbb{R}^n \to \mathbb{R}^n$ is defined by

$$[Y, Z](x) = DZ(x)Y(x) - DY(x)Z(x).$$

Given a family $\mathcal{X}$ of smooth vector fields on $\mathbb{R}^n$, let $[\mathcal{X}]_k$, $k \in \mathbb{N}$, and $[\mathcal{X}]$ be defined by

$$[\mathcal{X}]_0 = \mathcal{X}, \quad [\mathcal{X}]_{k+1} = [\mathcal{X}]_k \cup \{[Y, Z] : Y, Z \in [\mathcal{X}]_k\}, \quad [\mathcal{X}] = \cup_{k\in\mathbb{N}}[\mathcal{X}]_k,$$

and set $[\mathcal{X}](x) = \{Y(x) : Y \in [\mathcal{X}]\}$.

Definition 2.12 ((Strong) Hörmander condition). *In the context of (22) (equivalently (37)), we say that $x^* \in M$ satisfies the* Hörmander condition *(respectively the* strong Hörmander condition*) if*

$$[\{S^0, \dots, S^m\}](x^*),$$

(respectively $\{S^1(x^*), \dots, S^m(x^*)\} \cup \{[Y, Z](x^*) : Y, Z \in [\{S^0, \dots, S^m\}]\}$ spans $\mathbb{R}^n$.

The Hörmander condition (respectively strong Hörmander condition) will lead to the uniqueness of the invariant measure and the almost-sure convergence of $(\Pi^x_t)_{t\geq 0}$ (respectively the convergence in total variation of $(P_t(x, \cdot))_{t\geq 0}$ towards it.

Corollary 2.13 ([Ben18], Theorem 4.10). *We assume that $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ is $H-$persistent, where $(X_t^x)_{t\geq 0}$ is the solution of (22).*

(i) *If there exists $x^* \in \Gamma_{M_+} \cap M_+$ satisfying the Hörmander condition, then*

$$\mathcal{P}_{\mathrm{inv}}(M_+) = \{\Pi\}, \quad \Pi \ll \ell,$$

and $\Pi_t^x \Rightarrow \Pi$, $\mathbb{P}$–a.s. for all $x \in M_+$, while $\lim_{t\to\infty} \Pi_t^x f = \Pi f$, $\mathbb{P}-$a.s. for every $f \in L^1(\Pi)$ such that $\int_0^T f(X_s^x)ds < \infty$ for all $T > 0$ and for all $x \in M_+$.

(ii) *If x^* satisfies the strong Hörmander condition, then $(P_t(x,\cdot))_{t\geq 0}$ converges to Π in Total variation.*

(iii) *If $[\{S^1, \ldots, S^m\}](x)$ spans $\mathbb{R}^n$ for every $x \in M_+$, then $\Gamma_{M_+} \cap M_+ = M_+$.*

As usual, we call the diffusion (37) *elliptic* at x if $S_1(x), \cdots, S_m(x)$ span $\mathbb{R}^n$.

Remark 2.14. Assume that the diffusion (37) is elliptic at every $x \in M_+$, then the strong Hörmander condition holds on M_+. Moreover, by the classical Chow's theorem, every $x \in M_+$ is accessible from M_+ (see e.g. [BH22], Proposition 6.33).

2.4 Rate of convergence

Under additional assumptions, the rate of convergence in Corollary 2.13 can be shown to be exponential. We focus now on Kolmogorov SDEs as defined in (22).

2.4.1 Exponential convergence when M_0 is compact

Based on the work of [Ben18], the authors of [BM24] derived exponential convergence when M_0 is a compact. If it is not sufficient to conclude about our model of interest (3) since $M_0 = \partial R_+^2$, it will be useful to motivate the additional conditions needed later.

We use M_0 and M_+ to denote general extinction and non-extinction sets as in Hypothesis 1, where we assume M_0 is compact.

Remark 2.15. As noticed from Remark A.7 in [BM24], by compactness of M_0, if (V, H) is like in Hypothesis 4, let $\tilde{V}$ be a C^2 function on M_+ which coincides with V on a neighborhood of M_0. It implies that $(\tilde{V}, \tilde{H})$ satisfies Hypothesis 4 with $\tilde{H} = H\mathbf{1}_{M_0} + \mathcal{L}_e^+ \tilde{V}\mathbf{1}_{M_+}$. It follows that V (hence H) can be assumed to be zero outside a compact neighborhood of M_0, without loss of generality.

We typically assume that Hypothesis 1-4 hold.

Lemma 2.16 ([BM24], Lemma A.10*(i)*). *Let $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ be $H-$persistence with $H-$exponents $\Lambda^-(H) > 0$. Then, for any $0 < \Lambda^- < \Lambda^-(H)$, there exists $T > 0$ and U a neighborhood of M_0 such that for all $x \in U$,*

$$\frac{1}{T}\int_0^T P_s H(x)ds < -\Lambda^- < 0. \tag{39}$$

Proof. Since M_0 is an invariant compact set and by continuity of the mapping $x \mapsto P_t H(x)$, it is sufficient to prove the assertion for some $T > 0$ and for all $x \in M_0$ in order to extend it to U a neighborhood of M_0. We suppose the contrary, which is

$$\frac{1}{T}\int_0^T P_s H(x)ds > -\Lambda^-.$$

Then, let $(x_n)_{n\geq 1}$ be a sequence of M_0 and for all $n \geq 1$, let

$$\mu_n(\cdot) = \frac{1}{T_n}\int_0^{T_n} P_s(x_n, \cdot)ds,$$

where $(T_n)_{n\geq 1}$ is a sequence such that $T_n \uparrow \infty$ as $n \to \infty$. Then, we have $\mu_n(M_0) = 1$ for all $n \geq 1$ which is a tight sequence so that every limit point μ satisfies $\mu U \leq$ cst by compactness of M_0.

Let $t > 0$ be fixed, then for any $f \in C_b(M)$, we have

$$\int_0^t P_t f d\mu_n - \int_0^t f d\mu_n = \frac{1}{T_n}\left(\int_0^{T_n} P_{s+t}f(x_n)ds - \int_0^{T_n} P_s f(x_n)ds\right),$$

so that μ is invariant for $(P_t)_{t\geq 0}$ since

$$\left|\int P_t f d\mu_n - \int f d\mu_n\right| \leq \frac{2t}{T_n}\|f\|_\infty \underset{n\to\infty}{\longrightarrow} 0.$$

Since $\mu_n H > -\Lambda^-$ by assumption, it implies that

$$\mu H \geq -\Lambda^-.$$

Let $\mu = \int \nu\rho(d\nu)$ be the ergodic decomposition on $\mathcal{P}_{\text{erg}}(M_0)$: then,

$$\mu H = \int \nu H \rho(d\nu) \geq -\Lambda^-,$$

which implies $\exists \nu \in \mathcal{P}_{\text{erg}}(M_0)$ such that $-\nu H \leq \Lambda^- < \Lambda^-(H)$, which contradicts that $\Lambda^-(H) = -\sup\{\mu H : \mu \in \mathcal{P}_{\text{erg}}(M_0)\}$. □

Proposition 2.17 ([BM24], Prop. A.9*(i)*)**.** *Let $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ be $H-$persistence with $H-$exponents $\Lambda^-(H) > 0$. Then, for any $0 < \Lambda^- < \Lambda^-(H)$, there exists $T,\ \theta > 0$ and U a neighborhood of M_0 such that*

$$P_t\left(e^{\theta V}\right)(x) \leq e^{\theta V(x)}e^{-T\Lambda^-}, \quad \forall x \in U\backslash M_0, \tag{40}$$

and $\sup_{x\in M_+\backslash U} P_t\left(e^{-\theta V}\right)(x) < \infty$.

Furthermore, for all $x \in U\backslash M_0$, $(X_t^x)_{t\geq 0}$ eventually leaves U in the sense that $\mathbb{P}_x(\tau_U(x) < \infty) = 1$, where $\tau_U(x) := \inf\{t \geq 0 : X_t^x \notin U\}$ for all $x \in U\backslash M_0$.

Proof. As in the previous Lemma 2.16, we let

$$Y_T^x = \int_0^T H(X_s^x)ds - \int_0^T P_s H(x)ds,$$

and

$$\bar{H} = \sup_{x\in U}\int_0^T P_s H(x)ds < -T\Lambda^-,$$

after shrinking U if necessary so the bound holds on all of U.

Recall that given

$$V(X_T^x) = V(x) + \int_0^T \mathcal{L}_e^+ V(X_s^x)ds + M_T^V(x),$$

where $(M_t^V(x)_{t\geq 0}$ is the local martingale induced by the extended generator, and since

$$\int_0^T \mathcal{L}_e^+ V(X_s^x)ds = \int_0^T H(X_s^x)ds = Y_T^x + \int_0^T P_s H(x)ds,$$

then by Lemma 2.16, we deduce that for all $\theta > 0$ and $x \in U \backslash M_0$,

$$\exp(\theta V(X_T^x)) = \exp(\theta V(x)) \exp\left(\theta \int_0^T P_s H(x) ds\right) \exp(\theta Y_T^x) \exp(\theta M_T^V(x))$$
$$< \exp(\theta V(x)) \exp(\theta \bar{H}(x)) \exp(\theta Y_T^x) \exp(\theta M_T^V(x)).$$

Taking the inequality under expectation and applying Cauchy-Schwarz inequality yield

$$\mathbb{E}(e^{\theta V(X_T^x)}) \leq \exp(\theta V(x)) \exp(\theta \bar{H}(x)) \sqrt{\mathbb{E}(\exp(2\theta Y_T^x))} \sqrt{\mathbb{E}\left(\exp(2\theta M_T^V(x))\right)}. \tag{41}$$

Let

$$\|H\|_\infty = \sup_{x \in M} |H(x)|, \quad \|\Gamma_e^+(V)\|_\infty = \sup_{x \in M_+} \Gamma_e^+(V)(x),$$

which are well-defined by Remark 2.15, and let

$$C(T) = \max\left(4T^2 \|H\|_\infty^2,\ 2T \|\Gamma_e^+(V)\|_\infty\right).$$

Since $\mathbb{E}(Y_T^x) = 0$ by Fubini's and $|Y_T^x| \leq 2T\|H\|_\infty$, as a consequence of log-Laplace estimates (see e.g. [BLM13], Lemma 2.2 and the proof), we have

$$\log\left(\mathbb{E}[\exp(2\theta Y_T^x)]\right) \leq \frac{(2\theta)^2 (4T\|H\|_\infty)^2}{8} \leq 8\theta^2 T^2 \|H\|_\infty^2 \leq 2\theta^2 C(T) \ \Rightarrow\ \mathbb{E}[\exp(2\theta Y_T^x)] \leq \exp[2\theta^2 C(T)].$$

Now, let $Z_T(\theta, x) = \exp(2\theta M_T^V(x) - 2\theta^2 \langle M^V(x)\rangle_T)$, where

$$\langle M^V(x)\rangle_T = \int_0^T \Gamma_e^+(V)(X_s^x) ds \leq T\|\Gamma_e^+(V)\|_\infty \leq \frac{C(T)}{2},$$

by Proposition 1.22. It follows that

$$\mathbb{E}[\exp(2\theta M_T^V(x))] = \mathbb{E}[Z_T(\theta, x)] \exp(2\theta^2 \langle M^V(x)\rangle_T) \leq \exp(\theta^2 C(T)),$$

where $\mathbb{E}[Z_T(\theta, x)] = \mathbb{E}[Z_0(\theta, x)] = 1$ since the local martingale $2\theta M_T^V(x) - 2\theta^2 \langle M^V(x)\rangle_T$ is a true martingale by Novikov's condition. It yields

$$P_T(e^{\theta V})(x) \leq \exp(\theta V(x)) \exp(\theta \bar{H}(x)) \exp\left(\frac{3}{2}\theta^2 C(T)\right) \leq \exp(\theta V(x)) \exp(\theta(\bar{H}(x) + 2\theta C(T))).$$

Recalling that $\bar{H} < -T\Lambda^-$ and $C(T) < \infty$, then for θ small enough such that

$$\bar{H} + 2\theta C(T) \leq -T\Lambda^-,$$

which achieves to prove that (40) holds for all $x \in U$.

Now, if $x \notin U$, the contraction argument holds by replacing the estimate $\bar{H} < -T\Lambda^-$ by

$$\int_0^T P_s H(x) ds \leq T\|H\|_\infty.$$

Then, (41) writes

$$P_T(e^{\theta V})(x) \leq \exp(\theta V(x)) \exp(\theta T \|H\|_\infty) \exp(2\theta^2 C(T)), \forall x \notin U,$$

which is bounded on x since V is supposed to be bounded (see Remark 2.15) and the remaining terms on the right-hand side, depending only on θ and T, are also bounded by previous computations.

Finally, let $x \in U \backslash M_0$ and let $\tau = \min\{n \in \mathbb{N} : X^x_{tn} \notin U\}$. We have that the process $(S_n)_{n\geq 0}$ defined by

$$\exp\left(\theta\left(V(X^x_{(n\wedge\tau)T})\right) + (n \wedge \tau)T\Lambda^-)\right), \quad n \geq 1,$$

is a supermartingale. Indeed, let $(\mathcal{F}_t)_{t\geq 0}$ be the natural filtration induced by $(X^x_t)_{t\geq 0}$ and defined by $\mathcal{F}_t = \sigma(X^x_s : s \leq t)$. For $t \geq 0$ fixed, we let $\mathcal{G}_n = \mathcal{F}_{nt}$, then

$$\mathbb{E}(S_{n+1}|\mathcal{G}_n) \leq S_n.$$

It is direct whenever $n \geq \tau$ since we still have $n + 1 > \tau$ and $(n + 1) \wedge \tau = n \wedge \tau = \tau$. If $n < \tau$, then $n \wedge \tau = (n + 1) \wedge \tau \leq \tau$ and by Markov property and (40),

$$\mathbb{E}\left[\exp\left(\theta\left(V(X^x_{((n+1)\wedge\tau)T})\right)\right) \mid \mathcal{G}_n\right] = P_T(\exp(\theta V))(X^x_{nT}) \leq \exp\left(\theta V(X^x_{nT})\right)\exp(-T\Lambda^-),$$

and multiplying both sides by $\exp\left(((n + 1) \wedge \tau)T\Lambda^-\right)$ yields

$$\mathbb{E}[S_{n+1} \mid \mathcal{G}_n] \leq S_n.$$

Since $V \geq 0$ by $H-$persistence, it implies that

$$\mathbb{E}[\exp((n \wedge \tau)T\Lambda^-)] \leq \mathbb{E}[S_n] \leq \mathbb{E}[S_0] = \exp(\theta V(x)),$$

and by a monotone convergence argument, letting $n \to \infty$, we have

$$\mathbb{E}[\tau T\Lambda^-] = \lim_{n\to\infty} \mathbb{E}[\exp((n \wedge \tau)T\Lambda^-)] \leq \exp(\theta V(x)),$$

which is a.s. bounded by previous computations so that $\mathbb{P}(\tau < \infty) = 1$. □

Theorem 2.18 ([BM24], Prop. A.12)**.** *Let's suppose that $\Lambda^-(H) > 0$ and $\exists p \in M_+$ a Doeblin point, accessible from M_+. Then, there exists a unique invariant probability measure Π such that $\Pi(M_+) = 1$ and there exist positive constants C, a, and θ such that for every measurable function $f : M_+ \to \mathbb{R}$ and $\forall x \in M_+$,*

$$|P_t f(x) - \Pi(f)| \leq Ce^{-at}\left(1 + \max\{e^{\theta V(x)},\ W(x)\}\right)\|f\|_\theta, \tag{42}$$

where

$$\|f\|_\theta = \sup_{x\in M_+} \frac{|f(x)|}{1 + \max\{e^{\theta V(x)},\ W(x)\}},$$

and W comes from Hypothesis 2.

Proof. Since Hypothesis 3***(b)*** holds for $W = U$, $\tilde{W} = aW = aU$ from Proposition 2.1, we can use Theorem 1.40***(ii)*** with $\tilde{W} = aW$ so that

$$P_T W \leq e^{-aT} W + \frac{C}{a}.$$

Let $Z := \exp(\theta V) + W$ be defined on M_+, then by Proposition 2.17,

$$P_T Z \leq e^{-a'T} Z + \frac{C}{a},$$

where $a' = \min\{a, \Lambda^-\}$. The existence of a unique invariant probability measure and the exponential convergence (42) follows from Theorem 8.15 in [BH22]. □

Remark 2.19. The key assumption that M_0 is compact is used in 2 crucial arguments:

1. The sequence $(\mu_n)_{n\geq 1}$ of probability measures on M_0 defined by

$$\mu_n(\cdot) := \frac{1}{T_n}\int_0^{T_n} P_s(x_n,\cdot)ds, \quad n \geq 1,$$

is tight;

2. Inequality (39) is verified on all M_0, hence on a neighborhood U of M_0.

The first consequence of the compactness of M_0 may be achieved by another way and extended to non-compact extinction set. Indeed, it follows Proposition 1.37 applied to $W = U$, $\tilde{W} = aU$, with U from Proposition 2.1,

$$a\int_0^{T_n} P_sU(x_n)ds \leq U(x_n) + CT_n - P_{T_n}U(x_n) \;\Rightarrow\; \frac{1}{T_n}\int_0^{T_n} P_sU(x_n)ds \leq \frac{U(x_n)}{aT_n} + \frac{C}{a} \leq \text{cst},$$

taking $T_n > U(x_n)$ large enough if needed. Then, by Lemma 1.35, since the previous bound implies that $\limsup_{n\to\infty}\mu_n U \leq \text{cst}$, then $(\mu_n)_{n\geq 1}$ is tight.

By tightness, there exists a subsequence of $(\mu_n)_{n\geq 1}$ so that $(\mu_{n_k})_{k\geq 1}$ converges weakly to μ. Still by Lemma 1.35, since $\frac{U}{1+|H|}$ is proper by Hypothesis 4***(iii)***, then $\mu_{n_k}H \underset{k\to\infty}{\longrightarrow} \mu H$.

However, the second assertion may not be verified for non-compact M_0: we need a control near M_0 *and* at infinity, which will be detailed in the next subsection.

2.4.2 Exponential convergence when M_0 is non-compact - H-persistence at infinity

The major difference from the compact extinction set case is that the estimate (39) is verified on M_0 and can be extended to an open neighborhood by continuity of $x \mapsto P_tH(x)$ and invariance of M_0. When the extinction set is non-compact, this extension cannot be verified.

Throughout this section we will assume the following strengthening of Hypothesis 4,

Hypothesis 5. Given V from Hypothesis 4, there exists $\gamma < \infty$ such that

$$\langle M^V(x)\rangle_t \leq \gamma t \quad \text{for all } x \in M_+^I,\ t \geq 0,$$

where $(M_t^V(x))_{t\geq 0}$ is the martingale (respectively local martingale) defined in (10) (respectively in (12)).

Remark 2.20. Originally, in [Ben18], a second condition is presented in this stronger version of Hypothesis 4, where we assumed that the jumps of V are almost surely bounded. In our case, since the Kolmogorov SDEs we consider are typical diffusions, continuity of V guarantees that it remains true.

Definition 2.21 (**H-persistent, strong version**)**.** *The family* $\{(X_t^x)_{t\geq 0} : x \in M_+^I\}$ *is* H-persistent, strong version *if it is H-persistent and V satisfies Hypothesis 5. Additionally, if condition* ***(iii′)*** *in Hypothesis 4 is also verified, we will say that it is* H-persistent, strong version'.

Remark 2.22. Since $V \in D_e^I$ by Hypothesis 4 and in view of Proposition 1.22, a natural and sufficient condition ensuring Hypothesis 5 is that $V^2 \in D_e^I$ and

$$\gamma := \sup_{x\in M_+^I} \Gamma_e^I(V)(x) < \infty.$$

Although we do not see a direct link between Hypothesis 5 and any form of convergence rate, we can say more about: recall that

$$M_t^V(x) = V(X_t^x) - V(x) - \int_0^t \mathcal{L}_e^I V(X_s^x)ds, \quad \forall x \in M_+^I,$$

then the exponential process $(Z_\theta(t))_{t\geq 0}$ defined for any $\theta > 0$ as

$$Z_\theta(t) = \exp\left(\theta M_t^V(x) - \frac{\theta^2}{2}\langle M^V(x)\rangle_t\right),$$

is a supermartingale (see e.g. [Kal01], Lemma 23.19). It implies that

$$\mathbb{E}\left[\exp\left(\theta M_t^V(x) - \frac{\theta^2}{2}\gamma t\right)\right] \leq \mathbb{E}(Z_\theta(t)) \leq 1,$$

and there exists $\eta > 0$ depending on our choice of $\theta, T > 0$ large enough so that

$$P_T\left(e^{\theta V}\right)(x) \leq e^{-\eta T}e^{\theta V(x)},$$

for all $x \in C$ where C is a certain compact to be defined later.

In this situation, a Lyapunov function W controlling the behavior at infinity is not sufficient for exponential convergence; one must also control H *at infinity* (cf. [HN18], Remark 1.4 and Assumption 1.1, part (3)) so that the previous bound holds on $M_+^I \backslash C$ too. The whole idea is developed in Section 8.1 of [Ben18].

Definition 2.23. *We say that* $\{(X_t^x)_{t\geq 0} : x \in M_+^I\}$ *is* H-persistent at infinity *if it is H-persistent and the pair* (V, H) *of Definition 2.5 satisfies*

(a) V is proper;

(b) There exists a compact $C \subset M$ such that

$$\sup_{x\in M\backslash C} H(x) < 0. \tag{43}$$

Remark 2.24. One can read condition (43) as a natural extension from the case M_0 compact: it will play an important role to obtain a similar key estimate to (40). We will find a similar control of $P_t\left(e^{\theta V}\right)$ on $M \cap C$ while we will have a contraction outside of C by the additional assumption on H.

When ensuring to control the behavior near the extinction set and at infinity, the existence of a Doeblin point in the set of accessible points from M_+^I leads to the exponential convergence.

Theorem 2.25 ([Ben18], Theorem 4.13)**.** *Assume* $\{(X_t^x)_{t\geq 0} : x \in M_+^I\}$ *is H-persistent, strong version′ and at infinity and that there exists a Doeblin point* $x^* \in \Gamma_{M_+^I}$*. Then there exist* $\lambda > 0$, $\theta > 0$ *and* $cst > 0$ *such that, for all* $x \in M_+^I$ *and every measurable* $f : M_+^I \to \mathbb{R}$,

$$|P_t f(x) - \Pi f| \leq cst(1 + W_\theta(x))e^{-\lambda t}\|f\|_{W_\theta}$$

where $W_\theta = e^{\theta V}$ *and* $\|f\|_{W_\theta} = \sup_{x\in M_+^I} \frac{|f(x)|}{1+W_\theta(x)}$*. In particular,*

$$\left\|P_t(x, \cdot) - \Pi\right\|_{\mathrm{TV}} \leq cst(1 + W_\theta(x))e^{-\lambda t}.$$

Remark 2.26. Our function V already controls excursions to infinity via the Lyapunov drift. Recalling that $\mathcal{L}_e^+(V) = H|_{M_+^I}$, we require H to be negative off C to ensure exponential drift back towards C. However, the conditions ensuring stochastic persistence, hence our construction of (V, H), are focused on the behavior near the extinction set. There is no evidence that the additional conditions ensuring H-persistence at infinity are verified with our initial choice of V and H.

For example, recall that

$$V(x) = v\left(\sum_{i\in I} p_i h(x_i)\right),$$

where $h : \mathbb{R}_+ \to \mathbb{R} : u \mapsto \log\left(\frac{1}{u}\right)$. Since $h(u) \to -\infty$ when $u \to \infty$, having V proper would suggest that $v(-\infty) = \infty$, which clearly does not make part of our construction or assumption in the proof of Theorem 2.8.

However, the following practical tool will ensure that the conditions such that $H-$persistence, strong version$'$ with respect to M_0^I is also verified at infinity.

Proposition 2.27 ([Ben18], Proposition 4.14)**.** *Let $\{(X_t^x)_{t\geq 0} : x \in M_+^I\}$ be $H-$persistence, strong version$'$. We assume that there exists a pair of continuous functions $(\tilde{V}, \tilde{H})$ satisfying Hypothesis 4 with condition **(iii$'$)** and Hypothesis 5 such that:*

- ***(i)*** *$\tilde{V}$ is defined* on all *M and is proper;*
- ***(ii)*** *There exists $\varepsilon > 0$ such that*

$$\limsup_{\|x\|\to\infty} \varepsilon H(x) + \tilde{H}(x) < 0. \tag{44}$$

Then, $(X_t^x)_{t\geq 0}$ is $H-$persistence, strong version$'$ and at infinity.

Proof. The idea behind this trick is that we can verify that Hypotheses 4-5 and Definition 2.23 hold with $\varepsilon V + \tilde{V}$ and $\varepsilon H + \tilde{H}$. Indeed:

- Since the process is $H-$persistence, strong version$'$ there exists (V, H) verifying Hypotheses 4 with condition ***(iii$'$)*** and 5: since it is also true for $(\tilde{V}, \tilde{H})$, it is verified for $(\varepsilon V + \tilde{V}, \varepsilon H + \tilde{H})$.
- Since $\tilde{V}$ is defined on all M, by Remark 2.4, it implies that $\Lambda^-(\tilde{H}) = 0$ so that $\Lambda^-(\varepsilon H + \tilde{H}) = \varepsilon\Lambda^-(H) > 0$.
- Since $\tilde{V}$ is proper and $V \geq 0$ by Definition 2.5, $\varepsilon V + \tilde{V}$ is indeed proper.
- And since $\limsup_{\|x\|\to\infty} \varepsilon H(x) + \tilde{H}(x) < 0$, it follows that $\varepsilon H(x) + \tilde{H}(x) < 0$ on $M \setminus C$, where C is a (large enough) compact set.

□

For the Kolmogorov SDE we considered, a practical criteria to ensure H-persistent, strong version$'$ both at M_0^I and at infinity can also be deduced from the invasion rate (34). In particular, we suppose that $\eta = 0$, where η is the constant from (29):

Proposition 2.28. *We assume that conditions* (35) *and* (36) *from Theorem 2.8 hold. If* (35) *is strengthened to*

$$\left|\frac{LU}{U}\right|^q + \sum_{i\in I} |F_i|^q \leq cst\sqrt{U}, \quad \text{for some } q > 1, \tag{45}$$

and if, for $\tilde{V} := \log(U)$, $\tilde{H} := L\tilde{V}$, we assume in addition

$$\limsup_{\|x\|\to\infty} \varepsilon H(x) + \tilde{H}(x) < 0, \quad \text{for some } \varepsilon > 0, \tag{46}$$

then $(X_t^x)_{t\geq 0}$ is H–persistent, strong version$'$ and at infinity.

Proof. First, under (45), Hypothesis 4***(iii')*** holds for H constructed in the proof of Theorem 2.8. Indeed, recall that

$$|H| \leq \text{cst}\left(\sum_{i\in I} p_i|\lambda_i| + \sum_{i\in I} p_i^2\right) \Rightarrow |H|^q \leq \text{cst}\left(1 + \sum_{i\in I} |F_i|^q\right),$$

where the constant is changing from line to line and using $|\lambda_i| \leq |F_i| + \frac{|a_{ii}|}{2}$. By (45), it implies in particular that $|H|^q \leq \text{cst}(1+\sqrt{U})$, which proves that Hypothesis 4***(iii')*** holds since $W = \text{cst}(1+\sqrt{U})$, following Proposition 2.1***(v)***.

We also showed that $\Gamma_e^+(V) \leq \text{cst}$ so $\langle M^{\tilde{V}}\rangle_t \leq \text{cst}\cdot t$ by Proposition 1.22 which implies that Hypothesis 5 is satisfied so that $(X_t^x)_{t\geq 0}$ is $H-$persistent, strong version'.

Now, we use the practical criterion from Proposition 2.27. With $\tilde{V} = \log(U)$, which is indeed defined on all M since $U \geq 1$, and proper since U is. Moreover, $\tilde{V} \in C^2(M)$ so that $\tilde{V} \in \mathcal{D}_e^2$, and $\Gamma_e(\tilde{V}) = \frac{1}{U^2}\Gamma_e(U) \leq \text{cst}$ by Proposition 2.1***(v)*** again.

By Proposition 1.24, $\tilde{V}$ satisfies the strong law. Since $\tilde{H} = L\tilde{V}$,

$$\begin{aligned}
\tilde{H}(x) &= \sum_{i\in I} x_iF_i(x)\frac{\partial \log(U)}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j=1}^n x_ix_ja_{ij}(x)\frac{\partial^2 \log(U)}{\partial x_i\partial x_j}(x)\\
&= \sum_{i\in I} x_iF_i(x)\frac{1}{U(x)}\frac{\partial U}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j=1}^n x_ix_ja_{ij}(x)\frac{\partial\left(\frac{1}{U}\frac{\partial U}{\partial x_i}\right)}{\partial x_j}(x)\\
&= \frac{1}{U(x)}\sum_{i\in I} x_iF_i(x)\frac{\partial U}{\partial x_i}(x) - \frac{1}{2}\frac{1}{U^2(x)}\sum_{i,j=1}^n x_ix_ja_{ij}(x)\frac{\partial U}{\partial x_i}(x)\frac{\partial U}{\partial x_j}(x)\\
&\quad + \frac{1}{2}\frac{1}{U(x)}\sum_{i,j=1}^n x_ix_ja_{ij}(x)\frac{\partial^2 U}{\partial x_i\partial x_j}(x)\\
&= \frac{LU(x)}{U(x)} - \frac{1}{2U^2(x)}\Gamma_L(U)(x),
\end{aligned}$$

so that in particular,

$$|\tilde{H}|^q \leq \text{cst}\left(1 + \left|\tfrac{LU}{U}\right|^q\right). \tag{47}$$

Then, Hypothesis 4***(iii')*** (hence ***(iii)***) is verified by using (45). Since we assumed that condition (44) holds, we conclude that $(X_t^x)_{t\geq 0}$ is $H-$persistent, strong version' and at infinity. □

Remark 2.29. One of the original condition in [Ben18] to ensure $H-$persistent, strong version' and at infinity was

$$1 + \varepsilon F_i \geq 0, \quad \forall i \in I, \tag{48}$$

for some $\varepsilon > 0$. It is directly related to the condition (44) which can be deduced from it: from (48), we obtain

$$-\sum_{i\in I} p_iF_i(x) \leq \frac{1}{\varepsilon}\sum_{i\in I} p_i < \infty \Rightarrow H \leq C(1 - \sum_{i\in I} p_iF_i(x)) < \infty.$$

Since

$$\tilde{H} \leq \frac{LU}{U} - \frac{1}{2U^2}\Gamma_L(U),$$

using (28) and the fact that U is proper, we achieve a general bound

$$\limsup_{\|x\|\to\infty} \tilde{H}(x) = -a,$$

so that taking ε small enough so that

$$\limsup_{\|x\|\to\infty} 2a - \varepsilon H > 0 \;\Rightarrow\; a > \limsup_{\|x\|\to\infty} \frac{\varepsilon}{2} H,$$

we have

$$\limsup_{\|x\|\to\infty} \varepsilon H + \tilde{H} \le \limsup_{\|x\|\to\infty} \varepsilon H - a < \limsup_{\|x\|\to\infty} \varepsilon H - \frac{\varepsilon}{2} H < 0.$$

A typical condition ensuring (48) follows that F_i is bounded by below, in the sense that

$$\liminf_{\|x\|\to\infty} F_i(x) > -\infty.$$

However, the Rosenzweig-MacArthur model (3) does not verify bounded by below drifts: in particular, we remark that

$$F_1(x) = 1 - \frac{x_1}{\kappa} - \frac{x_2}{1 + x_1} \underset{\|x\|\to\infty}{\longrightarrow} -\infty.$$

2.4.3 Polynomial convergence rate

As an alternative to the exponential convergence rate, one can verify a polynomial convergence as presented in [BBN22]. We recall the general conditions from Proposition 2.1, that we suppose to be verified: $\exists U : M \to [1, \infty)$ a proper map such that (28) is verified and (29) holds for $\eta = 0$, which is

$$LU \le -aU + b, \quad \Gamma_L(U) \le cU^2,$$

for $a, c > 0$ and $b \ge 0$.

We fix $I = \{1, \cdots, n\}$ and recall conditions (35) and (36) from Theorem 2.8, ensuring $H-$persistence and that we suppose to be true, which is

$$\limsup_{\|x\|\to\infty} \frac{U^{\frac{1}{2}}(x)}{1 + \sum_{i=1}^n |F_i(x)|} = \infty, \quad \sum_{i=1}^n p_i \mu \lambda_i > 0 \quad \text{for all } \mu \in \mathcal{P}_{\text{erg}}(M_0),$$

for positive numbers $\{p_i\}_{i=1}^n$.

Hypothesis 6. In addition to conditions (28), (29), (35) and (36), we suppose that the function U satisfies

$$\liminf_{\|x\|\to\infty} \frac{U(x)}{\ln(\|x\|)} > 0, \tag{49}$$

$$\limsup_{\|x\|\to\infty} \left(LU(x) + p_0 \sum_{i=1}^n |F_i(x)| \right) < 0, \quad \text{for some } p_0 > 0, \tag{50}$$

and

$$\sum_{i=1}^n |F_i(x)| \le C U^{d_0}(x), \quad \text{for some } d_0 \ge 1 \text{ and } C > 0. \tag{51}$$

Starting from (36) from Theorem 2.8, note that $\sum_{i=1}^n p_i$ can be assumed to be sufficiently small without loss of generality. Then, in view of the definition of $H-$persistence, we modify the construction of V so that

$$V(x) = U(x) - \sum_{i=1}^n p_i \ln(x_i).$$

For $\sum_{i=1}^n p_i$ sufficiently small, it implies that V is positive on M_+ while

$$V(x) \to \infty \quad \text{as } \|x\| \to \infty,$$

and

$$H|_{M_+} = LV(x) = LU(x) + \sum_{i=1}^{n} p_i \lambda_i,$$

which can be extended continuously to M_0. In particular, since U is defined on all M, the persistence condition $\Lambda^-(H) > 0$ is still verified if $\sum_{i=1}^n p_i \mu\lambda_i > 0$ for all $\mu \in \mathcal{P}_{\text{erg}}(M_0)$ since $\Lambda^-(U) = 0$ (see Remark 2.4).

In this setting, let $q_0 > 1$ be a constant such that

$$-a + \frac{q_0 - 1}{2} c = 0,$$

where $a > 0$ is the constant from condition (28) and $c > 0$ the one from (29) in Proposition 2.1. Then, for $q \in]1, \min\{q_0, \frac{q_0+2}{2}\}[$, we define

$$W_q = \begin{cases} V^q + CU^q, & \text{if } 1 < q \leq 2, \\ V^q + CU^{2q-2} & \text{if } q > 2, \end{cases} \tag{52}$$

where C is the positive constant from (51).

Theorem 2.30 ([BBN22], Theorem 4.1)**.** *Assume that there exists a map $U : M \to [1, \infty)$ satisfying Hypothesis 6. Moreover, we suppose that there exists $x^* \in \Gamma_{M_+} \cap M_+$ which satisfies the strong Hörmander condition. Then, for all $q \in]1, \min\{q_0, \frac{q_0+2}{2}\}[$ and for all $1 \leq \beta \leq q$,*

$$\lim_{t\to\infty} t^{\beta-1} \|P_t(x, \cdot) - \Pi(\cdot)\|_{W_{\beta,q}} = 0, \quad \forall x \in M_+, \tag{53}$$

where $W_{\beta,q} = W_q^{1-\beta/q}$ with W_q defined as in (52), $\|f\|_{W_{\beta,q}} = \sup_{x\in M_+} \frac{|f(x)|}{1+W_{\beta,q}(x)}$, $\forall f \in \mathcal{B}(M_+)$, and

$$\|\mu\|_{W_{\beta,q}} = \sup_{|g|\leq W_{\beta,q}} |\mu(g)|,$$

for all signed measures μ.

An additional condition to Hypothesis 6 to ensure an exponential convergence rate exposed in [BBN22] is a strengthened version of (50),

$$\limsup_{\|x\|\to\infty} \left(L\ln(U) + p_0 \sum_{i=1}^{n} |F_i(x)| \right) < 0, \quad \text{for some } p_0 > 0. \tag{54}$$

This condition is similar to the $H-$persistence at infinity we introduced before: indeed, in the case where each F_i are bounded below, the existence of p_0 is equivalent to the existence of $\varepsilon > 0$ in Proposition 2.28 so that the pair of functions $(\ln(U) + p_0 V, L\ln(U) + p_0 H)$, where (V, H) satisfy the conditions of $H-$persistence M_0 (see Theorem 2.8), also fulfill the conditions of $H-$persistence at infinity (see Definition 2.23).

2.5 Extinction case

The study of extinction has already been quoted in [Ben18], underlying that the condition $\Lambda^+(H) < 0$ was necessary, as a complementary condition to persistence under $\Lambda^-(H) > 0$. However, we need additional conditions to ensure the extinction of the process, as detailed later. The following of the text develops the ideas introduced in [BHN+25] as well as in the appendix of [BM24] in the case M_0^I compact.

We recall the definition of the extinction sets

$$M_0^I = \{x \in M : \prod_{i\in I} x_i = 0\},$$

and of the non-extinction sets

$$M_+^I = \{x \in M : x_i > 0, \forall i \in I\}.$$

Definition 2.31 (H-nonpersistent). *The family* $\{(X_t^x)_{t\geq 0} : x \in M_+^I\}$ *is H–*nonpersistent *if there exists a pair* (V, H) *satisfying Hypothesis 4 such that*

(i) $\Lambda^+(H) < 0$*;*

(ii) $\lim_{v\to\infty} \sup\{\mathrm{dist}(x, M_0^I) : V(x) \geq v\} = 0$.

Remark 2.32. In view of Hypothesis 4, V need not to be non-negative, which is a requirement for the definition of H-persistence as stated in Definition 2.5. In the case of the H-nonpersistence defined above, V may be negative.

Let W be the Lyapunov function given by Hypothesis 2 and set $\tilde{U} = V - W$ defined on M_+^I. Let $I := H - \mathcal{L}_e W$ so that I is continuous and $\mathcal{L}_e^I \tilde{U} = I|_{M_+^I}$.

Let $0 < \Lambda^+ < -\Lambda^+(H)$. Since W is defined on the whole M and if W satisfies the strong law of large number, we have $\mu(\mathcal{L}_e W) = 0$ for all $\mu \in \mathcal{P}_{\mathrm{inv}}(M)$. Therefore, $\Lambda^+(I) = \Lambda^+(H)$.

Now, we state two additional hypotheses that will induce extinction as well as the main result of this section:

Hypothesis 7. For every compact set $C \subset M$, every $\varepsilon > 0$ and every $b > 0$ there exists $a > 0$ such that, for all $x \in C \cap M_+^I$,

$$\mathbb{P}_x\Big(|M_t^{\tilde{U}}| \leq a + bt, \forall t \geq 0\Big) \geq 1 - \varepsilon.$$

Hypothesis 8. I is uniformly integrable with respect to $\mathcal{P}_{\mathrm{inv}}(M_0^I)$, i.e.

$$\lim_{K\to\infty} \sup_{\mu\in\mathcal{P}_{\mathrm{inv}}(M_0^I)} \mu\left(|I|\mathbf{1}_{\{|I|>K\}}\right) = 0.$$

Theorem 2.33. *Assume the process is H–nonpersistent and that Hypotheses 3(a), 7 and 8 hold. Then for every* $0 < \Lambda^+ < -\Lambda^+(H)$*,* $\varepsilon > 0$ *and compact* $C \subset M$ *there exists* $u_{\varepsilon,C} > 0$ *such that, for all* $x \in O_{\varepsilon,C} := M_+^I \cap C \cap \{\tilde{U} > u_{\varepsilon,C}\}$*,*

$$\mathbb{P}_x\Big(\liminf_{t\to\infty} \tfrac{V(X_t)}{t} \geq \Lambda^+\Big) \geq 1 - \varepsilon. \tag{55}$$

Before starting the proof, we need the following Lemma:

Lemma 2.34. *For all* $r > 0$*, there exists a compact set* $C_r \subset M$ *such that* $I(x) \geq r$ *for all* $x \notin C_r$*.*

Proof. We have

$$I = H - \mathcal{L}_e W \geq H + \tilde{W} - C, \tag{56}$$

and the result follows from the fact that $\frac{\tilde{W}}{1+|H|}$ is proper. Indeed, it implies that for all $r > 0$, there exists a compact set C_r such that

$$\tilde{W}(x) > (1 + |H(x)|)(r + C + 1), \ \forall x \in M\backslash C_r,$$

which implies that $I(x) \geq H(x) + (1 + |H(x)|)(r + C + 1) - C > r$ on $M\backslash C_r$. □

Proof of Theorem 2.33. In fact, we will show that

$$\mathbb{P}_x\Big(\liminf_{t\to\infty} \tfrac{\tilde{U}(X_t)}{t} \geq \Lambda^+\Big) \geq 1-\varepsilon, \tag{57}$$

and (55) follows since $V \geq \tilde{U}$.

In fact, it is enough to assume that $\frac{|H|}{\tilde{W}}$ is bounded and then replace W by λW in the definition of $\tilde{U}$.
In particular, it implies that $\exists h_1 > 0$ such that $I \geq -h_1$ since $I \geq r > 0$ outside a compact set C_r and I reaches its minimum on C_r as a continuous function on a compact set, by Lemma 2.34.

Moreover, by Hypothesis 8, I is uniform integrable with respect to $\mathcal{P}_{\mathrm{inv}}(M_0^I)$ and there exists $h \geq h_1$ such that $\mu(I\wedge h) \geq (\Lambda^+ + 3\eta)$, where $\eta > 0$ satisfies $\Lambda^+ + 3\eta < -\Lambda^+(H)$. Indeed, let $\delta := -\Lambda^+(H) - \Lambda^+ > 0$ and $\eta < \delta/3$ so that $\delta - 3\eta > 0$. Then, since $I - I\wedge h = (I-h)_+$, by uniform integrability, for h large enough we have

$$\int_M I d\mu - \int_M (I\wedge h) d\mu = \int_M (I-h)_+ d\mu < \delta - 3\eta.$$

Then, by rearranging, we obtain

$$\int_M (I\wedge h) d\mu \geq \int_M I d\mu - (\delta - 3\eta) \geq (\Lambda^+ + \delta) - (\delta - 3\eta) = \Lambda^+ + 3\eta, \tag{58}$$

since $\mu(I) = \mu(H) - \mu(\mathcal{L}_e W) = \mu(H) \geq -\Lambda^+(H) = \Lambda^+ + \delta$ where we use that W is defined on all M and in particular $\Lambda^+ + 3\eta = -\Lambda^+(H) - \delta + 3\eta < -\Lambda^+(H)$.

Let $\tilde{I} = I \wedge h$, then $\tilde{I} \leq I$ and $-h_1 \leq \tilde{I} \leq h$. Without loss of generality, we can use Lemma 2.34 where we assume $r \geq h$ and h is large enough so that $h - h_1 \geq 2(\Lambda^+ + \eta)$ and $r \geq h \geq \Lambda^+ + \eta$. In particular, we have $\tilde{I}(x) = h$ for all $x \notin C_r$ since $I \geq r \geq h$ outside the compact C_r.

Recalling that $\mathcal{L}_e^I \tilde{U} = I|_{M_+^I}$, and by rearranging some terms,

$$\tilde{U}(X_t^x) = \tilde{U}(x) + \int_0^t I(X_s^x)\, ds + M_t^{\tilde{U}}(x) \geq \tilde{U}(x) + \int_0^t \tilde{I}(X_s^x)\, ds + M_t^{\tilde{U}}(x). \tag{59}$$

In addition, by Hypothesis 3 and Theorem 1.40, $\limsup_{t\to\infty} \Pi_t^x(\tilde{W}) \leq C$. Since $\tilde{I}$ is bounded and continuous, then $\lim_{t\to\infty} \Pi_t^x(\tilde{I}) = \mu\tilde{I}$ for $\mu \in \mathcal{P}_{\mathrm{inv}}(M_0^I)$ by Lemma 1.35. Using (58), then there exists T_0 large enough such that, for all $T \geq T_0$ and all $x \in C_r \cap M_0^I$,

$$\int_0^T \tilde{I}(X_s^x)\mathrm{d}s \geq (\mu\tilde{I} - \eta)T \geq (\Lambda^+ + 2\eta)T. \tag{60}$$

Moreover, by C_b–Feller continuity, for all $T_1 \geq T_0$, the mapping

$$(x,T) \mapsto \int_0^T \tilde{I}(X_s^x) ds,$$

is jointly continuous on the compact $C_r \times [T_0, T_1]$ so that we can enlarge (60) to an open set $\mathcal{N} \subset C_r$ of $C_r \cap M_0^I$. On the complement $C_r \backslash \mathcal{N}$, and by definition of $\tilde{U}$, it reaches a positive minimum denoted u_1: hence, for any $x \in C_r \cap \{\tilde{U} < u_1\}$, then $x \in \mathcal{N}$ so that $\forall T \in [T_0, T_1]$,

$$\int_0^T \tilde{I}(X_s^x)\mathrm{d}s \geq (\Lambda^+ + \eta)T. \tag{61}$$

Let $T_1 = 2T_0$. For each $n \geq 0$, we introduce

$$\xi_n = (n+1)T_1 \wedge \inf\{t \geq nT_1 : X_t^x \in C_r\},$$

and we set

$$\Delta_n = \int_{\xi_n}^{(n+1)T_1} \tilde{I}(X_s^x)\mathrm{d}s - \mathbb{E}\Big[\int_{\xi_n}^{(n+1)T_1} \tilde{I}(X_s^x)\mathrm{d}s \mid \mathcal{F}_{\xi_n}\Big],$$

so that by Fubini's theorem and Markov property, it writes

$$\Delta_n = \int_{\xi_n}^{(n+1)T_1} \tilde{I}(X_s^x)\mathrm{d}s - \int_0^{(n+1)T_1-\xi_n} P_s\left(\tilde{I}(X_{\xi_n}^x)\right)\mathrm{d}s.$$

Let

$$J_n = \int_{nT_1}^{(n+1)T_1} \tilde{I}(X_s^x)\mathrm{d}s - \Delta_n = \int_{nT_1}^{\xi_n} \tilde{I}(X_s^x)\mathrm{d}s + \int_0^{(n+1)T_1-\xi_n} P_s\tilde{I}(X_{\xi_n}^x)\mathrm{d}s.$$

Equation (59) yields that for all $t \in [nT_1, (n+1)T_1)$,

$$\tilde{U}(X_t) \geq \tilde{U}(x) + \int_{nT_1}^{t} \tilde{I}(X_s^x)\mathrm{d}s + \sum_{k=0}^{n-1} J_k + \sum_{k=0}^{n-1} \Delta_k + M_t^{\tilde{U}}(x).$$

The idea is that between nT_1 and ξ_n, $\tilde{I}$ is uniformly lower bounded by h. It will provide a relevant bound for J_n as follows:

Lemma 2.35. *For all $n \geq 0$ such that $X_{\xi_n}^x \notin C_r \cap \{\tilde{U} < u_1\}$, one has*

$$J_n \geq (\Lambda^+ + \eta)T_1.$$

Let C be a compact subset of M, $\varepsilon > 0$, and choose $b_1, b_2 > 0$ with $b_1 + \frac{b_2}{T_1} = \eta$. By Hypothesis 7, there exists $a > 0$ such that for all $x \in C \cap M_+^I$,

$$\mathbb{P}_x(A) \geq 1 - \frac{\varepsilon}{2},$$

where

$$A = \left\{|M_t^{\tilde{U}}| \leq a + b_1 t, \quad \forall t \geq 0\right\}.$$

Because $(\Delta_n)_{n\geq 0}$ is a bounded martingale difference with bound $2hT_1$, enlarging a if necessary, we also have that $\forall x \in C \cap M_+^I$

$$\mathbb{P}_x(B) \geq 1 - \frac{\varepsilon}{2},$$

where

$$B = \left\{\left|\sum_{k=0}^{n-1} \Delta_k\right| \leq a + b_1 n, \quad \forall n \geq 0\right\}$$

We claim there exists $u_{\varepsilon,C} > 0$ such that, for all $x \in M_+ \cap C \cap \{\tilde{U} > u_{\varepsilon,C}\}$, on the event $A \cap B$, we have $X_{\xi_n}^x \notin C_r \cap \{\tilde{U} < u_1\}$ for every $n \geq 0$. In particular, $J_n \geq (\Lambda^+ + \eta)T_1$ for all n by Lemma 2.35. Therefore,

$$\frac{\tilde{U}(X_t^x)}{t} \geq \frac{\tilde{U}(x)}{t} - \frac{h_1 T_1}{t} + \frac{n}{t}(\Lambda^+ + \eta)T_1 - \frac{2a}{t} - b_1 - \frac{n}{t}b_2,$$

for $t \in [nT_1, (n+1)T_1)$, which also implies $T_1 \geq \frac{n}{t} \geq \frac{n}{(n+1)T_1}$. Then, on $A \cap B$,

$$\frac{\tilde{U}(X_t^x)}{t} \geq \frac{\tilde{U}(x)}{t} - \frac{h_1 T_1}{t} + \frac{n}{n+1}(\Lambda^+ + \eta) - \frac{2a}{t} - b_1 - \frac{b_2}{T_1},$$

and taking the inferior limit when $t \to \infty$, hence $n \to \infty$, we have

$$\liminf_{t\to\infty} \frac{\tilde{U}(X_t^x)}{t} \geq \Lambda^+ + \eta - (b_1 + \frac{b_2}{T_1}) = \Lambda^+.$$

Then, $\mathbb{P}_x(A \cap B) \geq 1 - \varepsilon$, which proves (57) and hence (55).

To conclude, it remains to prove the claim. Let

$$u_{\varepsilon,C} = u_1 + h_1 T_1 + 2a + b_1 T_1. \tag{62}$$

Assume by contradiction that, for some $x \in M_+^I \cap C \cap \{\tilde{U} > u_{\varepsilon,C}\}$ and $\omega \in A \cap B$, there exists $n \geq 0$ with $X_{\xi_n}^x \in C_r \cap \{\tilde{U} < u_1\}$. Let

$$m = \inf\{n \in \mathbb{N} : X_{\xi_n}^x \in C_r \cap \{\tilde{U} < u_1\}\}.$$

Then $m < \infty$ and $J_k \geq (\Lambda^+ + \eta)T_1$ for all $k \leq m - 1$ by Lemma 2.35. For $t \in [mT_1, (m+1)T_1)$, we have

$$\begin{aligned}
\tilde{U}(X_t^x) &\geq \tilde{U}(x) - h_1 T_1 + m(\Lambda^+ + \eta)T_1 - 2a - b_1 t - mb_2 \\
&\geq u_{\varepsilon,C} - h_1 T_1 + m(\Lambda^+ + \eta)T_1 - 2a - b_1(m+1)T_1 - mb_2 \\
&\geq u_1 + 2a + mT_1(\Lambda^+ + \eta - b_1 - \frac{b_2}{T_1}) \\
&> u_1,
\end{aligned}$$

where the second inequality is obtained by substituting $\tilde{U}(x)$ by $u_{\varepsilon,C}$, the third by rearranging and using (62), and the last one using $\eta - b_1 - \frac{b_2}{T_1} = 0$, $a > 0$. It gives $\tilde{U}(X_t^x) > u_1$, and in particular with $t = \xi_m = mT_1$, while by the definition of m, we should have $\tilde{U}(X_{\xi_m}^x) < u_1$, which contradicts $m < \infty$. Hence the claim holds, completing the proof of Theorem 2.33. □

Proof of Lemma 2.35. By definition of ξ_n, if $X_{\xi_n}^x \notin C_r \cap \{\tilde{U} < u_1\}$, then either $\xi_n = (n+1)T_1$ or $X_{\xi_n}^x \in C_r \cap \{\tilde{U} \geq u_1\}$. We treat each case independently:

Case $\xi_n = (n+1)T_1$. Then $X_t^x \notin C_r$ for all $t \in [nT_1, (n+1)T_1)$, and since $\tilde{I} = h$ outside C_r, where we supposed $h \geq \Lambda^+ + \eta$, then

$$J_n = \int_{nT_1}^{(n+1)T_1} \tilde{I}(X_s^x)\mathrm{d}s \geq hT_1 \geq (\Lambda^+ + \eta)T_1.$$

Case $X_{\xi_n}^x \in C_r \cap \{\tilde{U} \geq u_1\}$. Since ξ_n is the first instant $t \geq nT_1$ where $X_t^x \in C_r$, then for $nT_1 < t < \xi_n$, $X_t^x \notin C_r$ so $\tilde{I} = h$ and we have

$$J_n \geq h(\xi_n - nT_1) + \int_0^{(n+1)T_1 - \xi_n} P_s \tilde{I}(X_{\xi_n}^x)\mathrm{d}s.$$

If $(n+1)T_1 - \xi_n \geq T_0$, then by Fubini's theorem and (61), we have

$$\int_0^{(n+1)T_1 - \xi_n} P_s \tilde{I}(X_{\xi_n})\mathrm{d}s \geq (\Lambda^+ + \eta)((n+1)T_1 - \xi_n),$$

so

$$J_n \geq h(\xi_n - nT_1) + (\Lambda^+ + \eta)((n+1)T_1 - \xi_n) \geq (\Lambda^+ + \eta)T_1,$$

recalling that $h \geq \Lambda^+ + \eta$.

Otherwise, if $(n+1)T_1 - \xi_n < T_0$, then we can use the earlier fact that there exists $h_1 > 0$ such that $I \geq -h_1$: since $\tilde{I} = I \wedge h \geq -h_1$ in our case, and by Fubini's theorem,

$$J_n \geq h(\xi_n - nT_1) - h_1((n+1)T_1 - \xi_n) \geq h(T_1 - T_0) - h_1 T_0 = (h - h_1)T_0 \geq 2(\Lambda^+ + \eta)T_0 \geq (\Lambda^+ + \eta)T_1,$$

where the second inequality is obtained using $(n+1)T_1 - \xi_n < T_0$, the third equality and last inequality with our definition $T_1 = 2T_0$, and the fourth one is due to our earlier choice of h which is $h - h_1 \geq 2(\Lambda^+ + \eta)$. □

Corollary 2.36. *Under the assumptions of Theorem 2.33 and* $\Gamma_{M_+^I} = M_+^I$,

$$X_t^x \underset{t\to\infty}{\overset{a.s}{\to}} M_0^I,$$

for any initial condition $x \in M_+^I$.

Proof. Let's focus on compact sets C of the form $\{\tilde{W}(x) \leq R\}$ for any $R > 0$ sufficiently large: Theorem 2.33 implies that, for any $\varepsilon > 0$, starting from a point $x \in O_{\varepsilon,C}$ sufficiently close to M_0^I, $(X_t^x)_{t\geq 0}$ will reach M_0^I for time t large enough with a probability of at least $1 - \varepsilon$.

Under the accessibility assumption, for any point $y \in M_+^I$ which may be far from M_0^I, we have a strictly positive probability to reach any open neighborhood O of $x \in O_{\varepsilon,C}$ starting from y, i.e. $\exists t \geq 0$ such that $P_t(y, O) > 0$.

We remark that the bound is uniform in y on any compact C. It implies that $\forall y \in C$, $\exists t \geq 0$ and $\delta > 0$ such that $P_t(y, O) \geq \delta > 0$. By Markov property, (57) becomes: $\forall \tilde{x} \in C := \{\tilde{W}(x) \leq R\}$,

$$\mathbb{P}_{\tilde{x}}\Big(\liminf_{t\to\infty} \tfrac{V(X_t)}{t} \geq \Lambda^+\Big) \geq \mathbb{P}_x\Big(\liminf_{t\to\infty} \tfrac{V(X_t)}{t} \geq \Lambda^+\Big) P_t(\tilde{x}, O_{\varepsilon,C}) \geq (1-\varepsilon)\delta > 0.$$

Now, let $x' \in M_+^I \backslash C = \{\tilde{W}(x) > R\}$: we will show that, starting from x', the process $(X_t^{x'})$ reaches C in a finite time with exponential decay, in the sense that if $\tau_C := \inf\{t \geq 0 : X_t^{x'} \in C\}$ is the hitting time of C, then τ_C has finite exponential moments, which is $\mathbb{E}_{x'}(e^{\lambda\tau_C}) < \infty$, for some $\lambda > 0$.

By Proposition 6.11 in [BH22], the condition

$$P\tilde{W} - \tilde{W} \leq -\lambda\tilde{W} \text{ on } M_+^I \backslash C \text{ for some } 0 < \lambda < 1, \tag{63}$$

implies that $\mathbb{E}_{x'}(e^{\lambda\tau_C}) < \frac{1}{1-\lambda} P\tilde{W}(x')$ for all $x' \in M_+^I \backslash C$. This result is considered on discrete Markov chains but combining a discretization argument with the continuity of X_t^x in (t, x) from Proposition 2.1***(i)*** to control the jumps size, the compactness of C and dominated convergence, the same holds true for the continuous process.

From Theorem 1.40, since Hypothesis 3(b) is verified as a consequence of Proposition 2.1***(iii)*** , then

$$P_t\tilde{W} \leq e^{-at}\tilde{W} + C \ \Rightarrow\ P_t\tilde{W} - \tilde{W} \leq -(1 - e^{-at})\tilde{W} + C,$$

holds on M where $a > 0$ is the constant from Hypothesis 3(b). In particular, on $M_+^I \backslash C$, since $\tilde{W}(x) > R$ with R large enough, the constant C may be absorbed so that $\exists 0 < \lambda < 1$ such that $P_t\tilde{W} - \tilde{W} \leq -\lambda e^{at}\tilde{W}$. For $t = 1$, condition (63) is verified.

Finally, since $X_{\tau_C}^{x'} \in C$, applying the Markov property (4) a second time leads to

$$\begin{aligned}
\mathbb{P}_{x'}\Big(\liminf_{t\to\infty} \tfrac{V(X_t)}{t} \geq \Lambda^+\Big) &\geq \mathbb{P}_{x'}\Big(\liminf_{t\to\infty} \tfrac{V(X_{t+\tau_C})}{t} \geq \Lambda^+\Big) \\
&= \mathbb{P}_{X_{\tau_C}^{x'}}\Big(\liminf_{t\to\infty} \tfrac{V(X_t)}{t} \geq \Lambda^+\Big) \\
&\geq (1-\varepsilon)\delta, \ \forall x' \in M_+^I \backslash C.
\end{aligned}$$

Let $\mathcal{A} := \{\liminf_{t\to\infty} \frac{V(X_t)}{t} \geq \Lambda^+\}$ so that $\forall x \in M_+^I$, $\mathbb{P}_x(\mathcal{A}) \geq (1-\varepsilon)\delta > 0$ by previous computations. Let $M_t := \mathbb{E}_x(\mathbf{1}_{\mathcal{A}} \mid \mathcal{F}_t)$ with $(\mathcal{F}_t)_{t\geq 0}$ is the natural filtration of $(X_t)_{t\geq 0}$ so that $(M_t)_{t\geq 0}$ is a martingale, bounded in $[0,1]$. In particular, $\mathbf{1}_{\mathcal{A}}$ is L^1 and $\mathcal{A} \in \mathcal{F}_\infty$ where $\mathcal{F}_\infty$ is the σ-field generated by $\cup_{t\geq 0}\mathcal{F}_t$. By Doob's martingale convergence theorem (see Theorem A.7 in [BH22]), M_t converges a.s. and in L^1 as $t \to \infty$ to M_∞ and since $\mathcal{A} \in \mathcal{F}_\infty$, it summarizes to

$$\lim_{t\to\infty} M_t = \lim_{t\to\infty} \mathbb{E}_x(\mathbf{1}_{\mathcal{A}} \mid \mathcal{F}_t) = \mathbb{E}_x(\mathbf{1}_{\mathcal{A}} \mid \mathcal{F}_\infty) = \mathbf{1}_{\mathcal{A}}.$$

Using Markov property (4),

$$\lim_{t\to\infty} \mathbb{E}_x(\mathbf{1}_{\mathcal{A}} \mid \mathcal{F}_t) = \lim_{t\to\infty} \mathbb{E}_{X_t^x}(\mathbf{1}_{\mathcal{A}}) = \lim_{t\to\infty} \mathbb{P}_{X_t^x}(\mathcal{A}) \geq (1-\varepsilon)\delta.$$

Since $\mathbf{1}_{\mathcal{A}} \in \{0,1\}$, above condition implies that $\mathbf{1}_{\mathcal{A}} = 1$ a.s., or alternatively

$$\mathbb{P}_x\Big(\liminf_{t\to\infty} \tfrac{V(X_t)}{t} \geq \Lambda^+\Big) = 1, \ \forall x \in M_+^I.$$

Finally, the process $(X_t^x)_{t\geq 0}$ being H-nonpersistent, we can conclude that $X_t^x \underset{t\to\infty}{\longrightarrow} M_0^I$ for all $x \in M_+^I$ a.s. by the Definition 2.31*(ii)*. □

2.5.1 A practical criterion for Hypotheses 7 and 8

Indeed, Hypotheses 7 and 8 may seen difficult to reach. However, in our settings, the following Lemma is a practical criterion to prove that Hypotheses 7 and 8 holds based on our earlier results.

Lemma 2.37. *Assume there exist proper maps $G, \tilde{G} : M \to \mathbb{R}_+$ such that $G \in \mathcal{D}_e$ and $\mathcal{L}_e G \leq -\tilde{G} + C$. If $\frac{\tilde{G}}{1+|\mathcal{L}_e W|}$ is proper, then Hypothesis 8 is satisfied. If furthermore $V \in D_e^{2,I}$, $W \in D_e^2$, and $\Gamma_e^I(V), \Gamma_e(W) \leq cst \cdot \tilde{G}$, then Hypothesis 7 is also satisfied.*

Before stating the proof of Lemma 2.37, we introduce this two intermediary results:

Lemma 2.38. *If there exists a function $c : M \to \mathbb{R}_+$, bounded on compact sets, and some $p > 1$ such that $(M_t^{\tilde{U}})_{t\geq 0}$ is an L^p-martingale and*

$$\sup_{t\geq 0} \frac{\mathbb{E}_x((M_t^{\tilde{U}})^p)}{t} \leq c(x), \tag{64}$$

then Hypothesis 7 holds true.

Proof. Let C be a compact subset of M and set $K := \sup_{x\in C} c(x) < \infty$. Observe that

$$\{\exists t \geq 0 : |M_t^{\tilde{U}}| \geq a + bt\} \subset \Big\{\sup_{0\leq t\leq 1} |M_t^{\tilde{U}}| \geq a\Big\} \cup \cup_{n\in\mathbb{N}}\Big\{\sup_{2^n\leq t\leq 2^{n+1}} |M_t^{\tilde{U}}| \geq a + b2^n\Big\}.$$

Then, for all $x \in C \cap M_+^I$, by Doob's L^p martingale inequality (see e.g. [RY04], Chapter II, Theorem 1.7) and (64),

$$\mathbb{P}_x\left(\sup_{2^n\leq t\leq 2^{n+1}} |M_t^{\tilde{U}}| \geq a + b2^n\right) \leq \frac{\mathbb{E}_x(|M_{2^{n+1}}^{\tilde{U}}|^p)}{(a+b2^n)^p} \leq \frac{c(x)2^{n+1}}{(a+b2^n)^p},$$

where the last inequality is due to the assumption on $(M_t^U)_{t\geq 0}$, and similarly

$$\mathbb{P}_x\left(\sup_{0\leq t\leq 1} |M_t^{\tilde{U}}| \geq a\right) \leq \frac{c(x)}{a^p}.$$

Hence,

$$\mathbb{P}_x(\exists t \geq 0 : |M_t^{\tilde{U}}| \geq a + bt) \leq Ka^{-p} + \sum_{n=0}^{\infty} \frac{K2^{n+1}}{(a + b2^n)^p}.$$

When $a \to \infty$, the first term goes to 0. In addition, since for all $n \geq 0$,

$$\lim_{a\to\infty} \frac{2^{n+1}}{(a + b2^n)^p} = 0,$$

and

$$\frac{2^{n+1}}{(a + b2^n)^p} \leq \frac{2}{b^p 2^{n(p-1)}}, \quad \forall a \geq 0,$$

dominated convergence implies that the second term also tends to 0 as $a \to \infty$, hence Hypothesis 7 holds since above bound is uniform in $x \in C \cap M_+^I$, so $\forall \varepsilon, b > 0$, $\exists 0 < a < \infty$ such that

$$\mathbb{P}_x(\exists t \geq 0 : |M_t^{\tilde{U}}| \geq a + bt) \leq \varepsilon \;\Rightarrow\; \mathbb{P}_x(\exists t \geq 0 : |M_t^{\tilde{U}}| \leq a + bt) \geq 1 - \varepsilon.$$

□

Lemma 2.39. *Let $f : M \to \mathbb{R}$ be continuous and suppose $\frac{\tilde{W}}{1+|f|}$ is proper. Then f is uniformly integrable with respect to $\mathcal{P}_{\text{inv}}(M)$, i.e.*

$$\lim_{K\to\infty} \sup_{\mu\in\mathcal{P}_{\text{inv}}(M)} \mu(|f|\mathbf{1}_{\{|f|>K\}}) = 0.$$

Proof. Let's assume that f is not uniformly integrable with respect to $\mathcal{P}_{\text{inv}}(M)$. Then there exists $\varepsilon > 0$ such that, for every $n \geq 0$, there is $\mu_n \in \mathcal{P}_{\text{inv}}(M)$ with

$$\mu_n(|f|\mathbf{1}_{\{|f|>n\}}) > \varepsilon.$$

Hence, for all $n_0 \geq 0$ and every $n \geq n_0$, $\mu_n(|f|\mathbf{1}_{\{|f|\geq n_0\}}) > \varepsilon$. Since $\mathcal{P}_{\text{inv}}(M)$ is compact by Proposition 1.37, there exists $\mu \in \mathcal{P}_{\text{inv}}(M)$ such that $\mu_n \Rightarrow \mu$, or at least one can extract a subsequence $(\mu_{n_k})_{k\geq 1}$ of $(\mu_n)_{n\geq 1}$ such that $\mu_{n_k} \Rightarrow \mu$. We claim that for all n_0,

$$\varepsilon \leq \limsup_{n\to\infty} \mu_n(|f|\mathbf{1}_{\{|f|\geq n_0\}}) \leq \mu(|f|\mathbf{1}_{\{|f|\geq n_0\}}),$$

contradicting $f \in L^1(\mu)$. Indeed, under the assumption that $\frac{\tilde{W}}{1+|f|}$ is a proper map, then $f \in L^1(\mu)$ by Lemma 1.35***(ii)*** (see Remark 1.36), so f is uniformly integrable with respect to $\mathcal{P}_{\text{inv}}(M)$.

We now prove the claim. Let $F \subset M$ be a closed set of M, and set $F_k := \left\{y : d(y, F) \leq \frac{1}{k}\right\}$ and

$$\varphi_k(x) := 1 - kd(x, F), \quad \forall x \in F_k,$$

while $\varphi_k(x) = 0$ otherwise. Then, φ_k is k Lipschitz such that $\mathbf{1}_F \leq \varphi_k \leq \mathbf{1}_{F_k}$. Moreover,

$$\frac{\tilde{W}}{1 + |f|\varphi_k} \geq \frac{\tilde{W}}{1 + |f|\mathbf{1}_{F_k}} \geq \frac{\tilde{W}}{1 + |f|},$$

then $\frac{\tilde{W}}{1+|f|\varphi_k}$ is proper since $\frac{\tilde{W}}{1+|f|}$ is, and $f\varphi_k$ is continuous. By Lemma 1.35***(ii)***, $\mu_n(|f|\varphi_k) \underset{n\to\infty}{\longrightarrow} \mu(|f|\varphi_k)$ and $\forall k \geq 0$,

$$\limsup_{n\to\infty} \mu_n(|f|\mathbf{1}_F) \leq \mu(|f|\varphi_k).$$

Since $\varphi_k \downarrow \mathbf{1}_F$ as $k \to \infty$, by monotone convergence it yields

$$\limsup_{n\to\infty} \mu_n(|f|\mathbf{1}_F) \leq \mu(|f|\mathbf{1}_F),$$

so applying this with $F = \{|f| \geq n_0\}$, it proves the claim and completes the proof. □

Remark 2.40. Because Hypothesis 4*(iii)* holds true for our choice of H, it follows that H is uniformly integrable with respect to $\mathcal{P}_{\text{inv}}(M)$.

Proof of Lemma 2.37. We assume there exist proper functions $G, \tilde{G} : M \to \mathbb{R}_+$ such that

$$G \in D_e, \quad \mathcal{L}_e G \leq -\tilde{G} + C, \text{ and } \frac{\tilde{G}}{1 + |\mathcal{L}_e W|} \text{ is proper,}$$

and in particular Hypothesis 2 holds for $G,\ \tilde{G}$. Then, Hypothesis 8 is satisfied by Lemma 2.39 and since H is uniformly integrable by Hypothesis 4 so that $I = H - \mathcal{L}_e W$ is uniformly integrable with respect to $\mathcal{P}_{\text{inv}}(M_0^I))$.

Now, we want to use the condition from Lemma 2.38, in particular we will show that $(M_t^{\tilde{U}})_{t\geq 0}$ is an L^2 martingale and the existence of a function $c : M \to \mathbb{R}_+$, bounded on compact set such that:

$$\sup_{t\geq 0} \frac{\mathbb{E}_x\left((M_t^{\tilde{U}})^2\right)}{t} \leq c(x),$$

hence Hypothesis 7 is satisfied.

Given that V satisfies Hypothesis 4, $V \in D_e^I$ and similarly, W satisfies Hypothesis 2 so that $W \in D_e$. Since in addition we have $V \in D_e^{2,+}$ and $W \in D_e^2$, then $\Gamma_e^I V$ and $\Gamma_e W$ are well-defined.

Given the conditions $\Gamma_e^I V, \Gamma_e W \leq \text{cst} \cdot \tilde{G}$, then by Corollary 1.39, $(M_t^V)_{t\geq 0}$ and $(M_t^W)_{t\geq 0}$ are L^2 martingales and satisfy the strong law since $\tilde{G}$ is as is Hypothesis 2.

By definition of $\mathcal{D}_e^I$, then for all $x \in M_+^I$, we have

$$\begin{aligned} M_t^{\tilde{U}}(x) &= \tilde{U}(X_t^x) - \tilde{U}(x) - \int_0^t \mathcal{L}_e^+ \tilde{U}(X_s^x) ds \\ &= V(X_t^x) - V(x) - (W(X_t) - W(x)) - \int_0^t \mathcal{L}_e^+ (V - W)(X_s^x) ds \\ &= M_t^V(x) - M_t^W(x). \end{aligned}$$

Then, $(M_t^{\tilde{U}})_{t\geq 0}$ is an L^2 martingale since:

$$\mathbb{E}_x(|M_t^{\tilde{U}}|^2) = \mathbb{E}_x(|M_t^V - M_t^W|^2) \leq \mathbb{E}_x(|M_t^V|^2) + \mathbb{E}_x(|M_t^W|^2) < +\infty$$

In fact, since the space of L^2 martingales is a vector space, $(M_t^U)_{t\geq 0}$ is naturally a L^2 martingale since $M_t^{\tilde{U}}(x) = M_t^V(x) - M_t^W(x)$.

Since we are working on a probability space, it has a finite measure and

$$M_t^{\tilde{U}} \in L^2(\Omega, \mathcal{F}, \mathbb{P}) \subset L^1(\Omega, \mathcal{F}, \mathbb{P}) \Rightarrow (M_t^{\tilde{U}})_{t\geq 0} \text{ is integrable.}$$

Since $\mathbb{E}_x(|M_t^{\tilde{U}}|^2) \leq \mathbb{E}_x(|M_t^V|^2) + \mathbb{E}_x(|M_t^W|^2)$, the last condition to show is the existence of a function $c : M_+ \to \mathbb{R}$, bounded on compact sets, such that

$$\sup_{t>0} \frac{\mathbb{E}_x(|M_t^V|^2)}{t} + \sup_{t>0} \frac{\mathbb{E}_x(|M_t^W|^2)}{t} \leq c(x).$$

In particular, by Proposition 1.22, since $V \in D_e^{2,+}$, we have that

$$(M_t^V)^2(x) - \int_0^t \Gamma_e^I(V)(X_s^x)ds, \quad t \geq 0,$$

is a local martingale for all $x \in M_+^I$. Let $(\tau_n)_{n\geq 1}$ be a sequence of stopping times such that for all $n \geq 1$, the process

$$(M_{t\wedge\tau_n}^V)^2(x) - \int_0^{t\wedge\tau_n} \Gamma_e^I(V)(X_s^x)ds, \quad t \geq 0$$

is a martingale for all $x \in M_+^I$ and $n \in \mathbb{N}$. We now have

$$\mathbb{E}_x\left((M_{t\wedge\tau_n}^V)^2 - \int_0^{t\wedge\tau_n} \Gamma_e^I(V)(X_s^x)ds\right) = 0,$$

which implies $\mathbb{E}_x\left((M_{t\wedge\tau_n}^V)^2\right) = \int_0^{t\wedge\tau_n} P_s(\Gamma_e^I(V)(x))ds$, for all $t \geq 0$, $x \in M_+^I$ and $n \geq 1$ by Fubini's theorem. For t sufficiently large,

$$\int_0^{t\wedge\tau_n} P_s(\Gamma_e^I(V)(x))ds \leq \int_0^t P_s(\Gamma_e^I(V)(x))ds.$$

In particular, by Fatou's Lemma, we finally have

$$\mathbb{E}_x\left((M_t^V)^2\right) \leq \int_0^t P_s(\Gamma_e^I(V)(x))ds,$$

and the same holds true for W so that $\mathbb{E}_x\left((M_t^W)^2\right) \leq \int_0^t P_s(\Gamma_e(W)(x))ds$ for all $x \in M$ in this case. By $\Gamma_e^I V,\ \Gamma_e W \leq \mathrm{cst}\cdot\tilde{G}$, then

$$\mathbb{E}_x\left((M_t^V)^2\right),\ \mathbb{E}_x\left((M_t^W)^2\right) \leq \int_0^t \mathrm{cst}\cdot P_s(\tilde{G}(x))ds.$$

Since $\tilde{G}$ satisfies Hypothesis 2, we can use Proposition 1.37*(i)* which implies that

$$P_tG(x) + \int_0^t P_s\tilde{G}(x)ds \leq G(x) + C_x t,$$

where $C_x > 0$ is a constant, which may depend on x, for all $x \in M$. If $t > 1$, recalling that G is supposed to be positive, then

$$\frac{1}{t}\int_0^t P_s\tilde{G}(x)ds \leq G(x) + C_x, \quad \forall x \in M.$$

If $0 \leq t \leq 1$, then

$$\frac{1}{t}\int_0^t P_s\tilde{G}(x)ds \leq \sup_{0\leq s\leq 1} P_s\tilde{G}(x), \quad \forall x \in M,$$

and since the map $t \mapsto P_t\tilde{G}(x)$ is continuous for each fixed $x \in M$, $\sup_{0\leq s\leq 1} P_s\tilde{G}(x) < \infty$. It implies that

$$\sup_{t\geq 0}\frac{\mathbb{E}_x\left((M_t^V)^2\right)}{t} \leq \mathrm{cst}\cdot\sup_{t\geq 0}\frac{1}{t}\int_0^t P_s\tilde{G}(x)ds \leq \mathrm{cst}\cdot\left(G(x) + C_x + \sup_{0\leq s\leq 1} P_s\tilde{G}(x)\right) =: c(x), \quad \forall x \in M.$$

In particular, $c(x)$ is bounded on every compact $C \subset M$ as $\sup_{x\in C} C_x \leq \tilde{C} < \infty$, $x \mapsto G(x)$ is continuous taken over C, and $(s, x) \mapsto P_s\tilde{G}(x)$ jointly continuous taken over the compact $[0, 1] \times C$.

The same holds true for $\sup_{t\geq 0}\frac{\mathbb{E}_x((M_t^W)^2)}{t}$ so that

$$\sup_{t\geq 0}\frac{\mathbb{E}_x\left((M_t^V)^2\right)}{t} + \sup_{t\geq 0}\frac{\mathbb{E}_x\left((M_t^W)^2\right)}{t} \leq 2c(x),$$

which implies that Hypothesis 7 holds true. □

Chapter 3

Motivating nondegenerate examples

This chapter is dedicated to the study of two nondegenerate models of SDEs that serve as toy models. Section 3.1 presents a one-dimensional logistic SDE and details how to achieve Hypotheses 1-5 in case of stochastic persistence as well as Hypotheses 7 and 8 for the extinction case.

Section 3.2 introduces a general two-dimensional nondegenerate SDE and focuses on how to show H−persistence using the practical criterion introduced in Theorem 2.8. In particular, we assume that Hypotheses 1-3 are verified by the model.

3.1 The one-dimensional logistic SDE

Consider the following SDE defined on $\mathbb{R}_+$,

$$dx_t = x_t\left((r - x_t)\mathrm{d}t + \sigma \mathrm{d}B_t\right), \quad x_0 = x, \tag{65}$$

with $r > 0$ a constant. (65) is called the one-dimensional logistic SDE.

First, we derive an explicit solution of (65): let $Z_t = \frac{1}{x_t}$, $M_t = \exp(Y_t)$ with $Y_t = (r - \frac{\sigma^2}{2})t + \sigma B_t$. Then, by Itô's formula,

$$\begin{aligned} dZ_t &= -\frac{dx_t}{x_t^2} + \frac{\sigma^2}{2}\frac{2x_t^2}{x_t^3}\mathrm{d}t \\ &= -\frac{1}{x_t^2}\Big(x_t((r - x_t)\mathrm{d}t + \sigma \mathrm{d}B_t)\Big) + \frac{\sigma^2}{x_t}\mathrm{d}t \\ &= \frac{1}{x_t}\left(-r + x_t + \sigma^2\right)\mathrm{d}t - \frac{\sigma}{x_t}\mathrm{d}B_t \\ &= Z_t\left(-r + x_t + \sigma^2\right)\mathrm{d}t - \sigma Z_t \mathrm{d}B_t, \end{aligned}$$

which implies that

$$\frac{dZ_t}{Z_t} = (-r + \sigma^2 + x_t)\mathrm{d}t - \sigma \mathrm{d}B_t = \left(-r + \sigma^2 + \frac{1}{Z_t}\right)\mathrm{d}t - \sigma \mathrm{d}B_t. \tag{66}$$

Similarly for M_t,

$$dM_t = \left(r - \frac{\sigma^2}{2}\right)M_t\mathrm{d}t + \frac{1}{2}\sigma^2 M_t \mathrm{d}t = M_t\left(r\mathrm{d}t + \sigma \mathrm{d}B_t\right), \tag{67}$$

so that $\frac{dM_t}{M_t} = r\mathrm{d}t + \sigma \mathrm{d}B_t$. By applying the product rule for Itô processes,

$$\begin{aligned} d(M_t Z_t) &= M_t dZ_t + Z_t dM_t + dM_t dZ_t \\ &= M_t dZ_t + Z_t dM_t + d[M, Z]_t \\ &= M_t Z_t\left(\frac{dZ_t}{Z_t} + \frac{dM_t}{M_t} + \frac{d[M, Z]_t}{M_t Z_t}\right), \end{aligned} \tag{68}$$

where $[M, Z]_t$ is the quadratic covariation between the processes $(M_t)_{t\geq 0}$ and $(Z_t)_{t\geq 0}$. Given the polarization identity (see e.g. [Kun12], Section 2.3), we can express the quadratic covariation in term of quadratic variation,

$$[X, Y]_t = \frac{1}{2}\left([X + Y]_t - [X]_t - [Y]_t\right).$$

Recall that for $(X_t)_{t\geq 0}$ an Itô process defined as $dX_t = \mu_X \mathrm{d}t + \sigma_X \mathrm{d}B_t$, its quadratic variation is defined as

$$[X]_t = \int_0^t \sigma_X^2 ds.$$

Given $(Y_t)_{t\geq 0}$ another Itô process with drift μ_Y and variance σ_Y, then $(X_t + Y_t)_{t\geq 0}$ is also an Itô process with drift $\mu_X + \mu_Y$ and variance $\sigma_X + \sigma_Y$, which implies

$$[X + Y]_t = \int_0^t (\sigma_X + \sigma_Y)^2 ds.$$

In our case, we have

$$[Z]_t = \int_0^t \sigma^2 Z_s^2 ds, \quad [M]_t = \int_0^t \sigma^2 M_s^2 ds, \quad [Z + M]_t = \int_0^t (\sigma M_s - \sigma Z_s)^2 ds,$$

which implies that

$$[M, Z]_t = \frac{1}{2}\left(\int_0^t (-\sigma Z_s + \sigma M_s)^2 ds - \int_0^t \sigma^2 Z_s^2 ds - \int_0^t \sigma^2 M_s^2 ds\right) = -\int_0^t \sigma^2 M_s Z_s ds,$$

so that $\frac{d[M,Z]_t}{M_t Z_t} = -\sigma^2 \mathrm{d}t$ and (68) becomes

$$d(M_t Z_t) = M_t Z_t \left[\left(-r + \sigma^2 - \frac{1}{Z_t}\right)\mathrm{d}t - \sigma \mathrm{d}B_t + r\mathrm{d}t + \sigma \mathrm{d}B_t - \sigma^2 \mathrm{d}t\right] = M_t Z_t \left(\frac{1}{Z_t}\mathrm{d}t\right) = M_t \mathrm{d}t.$$

So,

$$M_t Z_t = M_0 Z_0 + \int_0^t M_s ds \Rightarrow X_t = \frac{1}{Z_t} = \frac{M_t}{M_0 Z_0 + \int_0^t M_s ds},$$

which gives us with $M_0 = 1$, $Z_0 = \frac{1}{x_0} = \frac{1}{x}$,

$$x_t = \frac{x \exp\left[\left(r - \frac{\sigma^2}{2}\right)t + \sigma B_t\right]}{1 + x\int_0^t \exp\left[\left(r - \frac{\sigma^2}{2}\right)s + \sigma B_s\right]ds}. \tag{69}$$

Remark 3.1. Hence, for $x \to \infty$

$$x_t^\infty = \lim_{x\to\infty} \frac{x \exp\left[\left(r - \frac{\sigma^2}{2}\right)t + \sigma B_t\right]}{1 + x\int_0^t \exp\left[\left(r - \frac{\sigma^2}{2}\right)s + \sigma B_s\right]ds} = \frac{\exp\left[\left(r - \frac{\sigma^2}{2}\right)t + \sigma B_t\right]}{\int_0^t \exp\left[\left(r - \frac{\sigma^2}{2}\right)s + \sigma B_s\right]ds},$$

so that for $f \in C_0(\mathbb{R}_+)$ a function that vanishes at infinity and $t > 0$ fixed, $P_t f \notin C_0(\mathbb{R}_+)$ since

$$\lim_{x\to\infty} P_t f(x) = \mathbb{E}\left(f\left(\frac{\exp\left[\left(r - \frac{\sigma^2}{2}\right)t + \sigma B_t\right)}{\int_0^t \exp\left[\left(r - \frac{\sigma^2}{2}\right)s + \sigma B_s\right)ds}\right)\right) \neq 0,$$

with $f(x) = \frac{1}{1+x}$ for example. Thus, the semigroup induced by $(X_t^x)_{t\geq 0}$ is not a $C_0(\mathbb{R}_+)$−Feller semigroup but is $C_b(\mathbb{R}_+)$-Feller in the sense that $(P_t)_{t\geq 0}$ preserves the continuous, bounded functions. As mentioned in Remark 1.3, this is a good example of the usefulness of the $C_b(M)$-Feller property instead of the $C_0(M)$-Feller one.

3.1.1 Stochastic persistence of the one-dimensional logistic SDE

Let $M_0 = \{0\}$ be the extinction set and $M_+ = \mathbb{R}_+^*$: naturally, Hypothesis 1 is satisfied since

$$P_t \mathbf{1}_{M_0}(x) = 0, \quad \text{if } x \neq 0,$$

and

$$P_t \mathbf{1}_{M_0}(x) = \mathbb{E}(\mathbf{1}_{M_0}(X_t^0)) = \mathbf{1}_{M_0}\left(\frac{0}{1+0}\right) = 1, \quad \text{if } x = 0,$$

so $P_t \mathbf{1}_{\{0\}} = \mathbf{1}_{\{0\}}$.

Let $M_+ = \mathbb{R}_+ \backslash M_0 = \mathbb{R}_+^*$. We want to show that the condition (28) of Proposition 2.1 holds, which will imply that Hypothesis 3, (hence Hypothesis 2) is verified.

Let $U(x) := e^{\theta x} \geq 1$ with $0 < \theta < \frac{2}{\sigma^2}$ so that $a := 1 - \theta \frac{\sigma^2}{2} > 0$. By definition of L,

$$LU(x) = x(r-x)\theta e^{\theta x} + \frac{1}{2}\sigma^2 x^2 \theta^2 e^{\theta x} = \theta x U(x)\left[r - x\left(1 - \theta\frac{\sigma^2}{2}\right)\right] = \theta x U(x)\,(r - ax)\,.$$

Let

$$F(x) := \theta x U(x)(r - ax) + U(x) = U(x)G(x),$$

with $G(x) = \theta x(r - ax) + 1 = -a\theta x^2 + \theta r x + 1$. Then, F reaches its maximum when

$$\begin{aligned} F'(x) = 0 &\Leftrightarrow U'(x)G(x) + U(x)G'(x) = 0 \\ &\Leftrightarrow \theta e^{\theta x}\left(-a\theta x^2 + \theta r x + 1\right) + e^{\theta x}\left(-2a\theta x + \theta r\right) = 0 \\ &\Leftrightarrow -a\theta^2 x^2 + (\theta^2 r - 2a\theta)x + \theta(r+1) = 0 \\ &\Leftrightarrow x = \frac{-\theta^2 r + 2a\theta \pm \sqrt{(\theta^2 r - 2a\theta)^2 + 4a\theta^3(r+1)}}{-2a\theta^2}, \end{aligned}$$

which simplify to

$$x = \frac{\theta^2 r - 2a\theta \pm \theta\sqrt{\theta^2 r^2 + 4a\theta + 4a^2}}{2a\theta^2}.$$

With the hypothesis $a, r, \theta > 0$, then

$$\tilde{x} = \frac{\theta r - 2a + \sqrt{\theta^2 r^2 + 4a\theta + 4a^2}}{2a\theta} > 0,$$

so that $\tilde{x} \in \mathbb{R}_+^*$ and

$$F(x) \leq F(\tilde{x}) = U(\tilde{x})G(\tilde{x}) = e^{\theta \tilde{x}}\left(\theta \tilde{x}(r - a\tilde{x}) + 1\right) =: b,$$

which implies that

$$LU(x) = F(x) - U(x) \leq -U(x) + b.$$

By Proposition 2.1, Hypothesis 3(b) holds with $W(x) = U(x) = e^{\theta x}$, hence Hypothesis 2 too. In addition, we have

$$\Gamma_L(U)(x) = x^2\sigma^2\theta e^{\theta x}\theta e^{\theta x} = x^2\sigma^2\theta^2\left(e^{\theta x}\right)^2 = x^2\sigma^2\theta^2 U(x)^2.$$

Let $f(x) = x^2\theta^2 e^{-\eta\theta x}$ with $0 < \eta < 1$, which reaches its maximum on $\mathbb{R}_+$ when

$$\begin{aligned} f'(x) = 0 &\Leftrightarrow 2x\theta^2 e^{-\eta\theta x} - x^2\theta^3\eta e^{-\eta\theta x} = 0 \\ &\Leftrightarrow -\theta\eta x^2 + 2x = 0 \\ &\Leftrightarrow x^* = 0 \notin \mathbb{R}_+^* \text{ or } x^* = \frac{2}{\theta\eta}. \end{aligned}$$

In the second case,

$$f(x^*) = \frac{4\theta^2}{\theta^2\eta^2}e^{-2} = \frac{4e^{-2}}{\eta^2} =: c > 0.$$

Then, for $0 < \eta < 1$ fixed, we have

$$x^2\theta^2 \leq \frac{4e^{-2}}{\eta^2}\left(e^{\theta x}\right)^{\eta} \Leftrightarrow \Gamma_L(U)(x) \leq c\sigma^2\left(e^{\theta x}\right)^{2+\eta} = c\sigma^2 U(x)^{2+\eta}. \tag{70}$$

Then, (29) is verified so that, by Proposition 2.1, Hypothesis 3(a) holds with $W = U^{\frac{1-\eta}{2}}$ and $\tilde{W} = (1+\mathrm{cst})W$.

Now, to show the H-persistence, let V be a smooth function such that $V(x) = -\log(x)$ for $x \in \left(0, \frac{1}{2}\right)$ and $V(x) = 0$ for $x \geq 1$. By Proposition 2.1***(iii)***, then $V,\ V^2 \in \mathcal{D}_e^+$, and

$$\mathcal{L}_e^+(V)(x) = V'(x)x(r-x) + \frac{\sigma^2}{2}x^2V''(x) = -(r-x) + \frac{\sigma^2}{2}, \quad \text{on } \left(0, \frac{1}{2}\right).$$

It can be continuously extended to bounded map $H : \mathbb{R}_+ \to \mathbb{R}$ since $\lim_{x\to 0}\mathcal{L}_e^+(V)(x) = -r + \frac{\sigma^2}{2}$. We also get

$$\Gamma_e^+(V)(x) = (\sigma x V'(x))^2 = \sigma^2.$$

Then, V satisfies the strong law by Corollary 1.39, and since $W(x) = \tilde{W}(x) = e^{\theta x}$ by Hypothesis 3(b),

$$\frac{\tilde{W}(x)}{1+|H(x)|} > \frac{e^{\theta x}}{x + r + 1 + \frac{\sigma^2}{2}} \underset{x\to\infty}{\longrightarrow} \infty,$$

so that $\frac{\tilde{W}}{1+|H|}$ is a proper map. It implies that Hypothesis 4 holds.

Now, condition (35) from Theorem 2.8 follows

$$\limsup_{x\to\infty}\frac{U(x)^{\frac{1-\eta}{2}}}{1+|F(x)|} \geq \limsup_{x\to\infty}\frac{\exp\left(\frac{1-\eta}{2}\theta x\right)}{1+x+r} = \infty,$$

with $0 < \eta < 1$. It remains to show that (36) holds to show that the process is H-persistent, i.e. $\forall\mu \in \mathcal{P}_{\mathrm{erg}}(M_0)$, $\mu\lambda > 0$ where λ is the invasion rate

$$\lambda(x) = r - x - \frac{\sigma^2}{2}.$$

Since the only ergodic measure on M_0 is δ_0, then

$$\delta_0\lambda = \lambda(0) = r - \frac{\sigma^2}{2} > 0 \Leftrightarrow r > \frac{\sigma^2}{2}.$$

Under this assumption, it implies that the process is H-persistent. Since all the points of $M_+ = (0, +\infty)$ are nondegenerate, $\Gamma_{\mathbb{R}_+^*} = \mathbb{R}_+^*$ by Stroock and Varadhan support theorem in [SV72], and by Corollary 2.13, there exist a unique invariant probability on M_+, denoted Π, such that for all $x \in M_+$,

$$\lim_{t\to\infty}\|P_t(x,\cdot) - \Pi\|_{TV} = 0,$$

and whose convergence is exponential by Theorem 2.18.

Figure 1 illustrates the situation of persistence for the one-dimensional logistic SDE: given fixed parameters, with $1 =: r > \frac{\sigma^2}{2} := \frac{0.25^2}{2}$, its trajectory stays in the non-extinction set M_+.

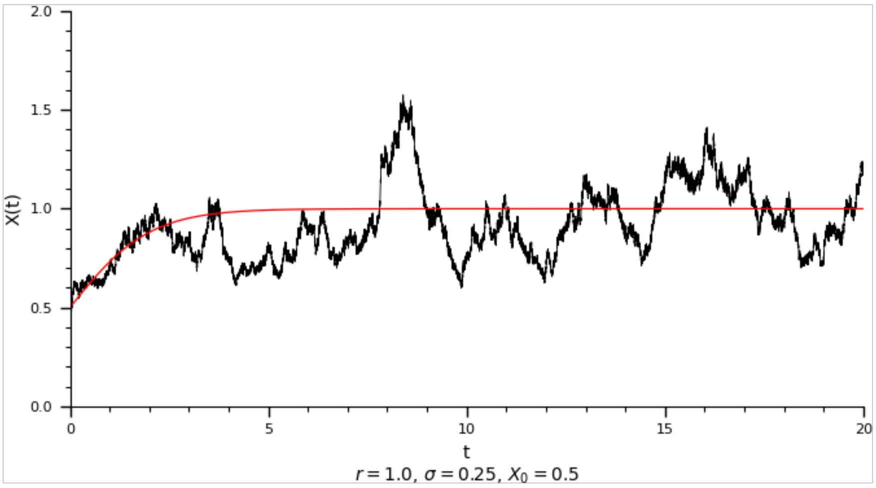

Figure 1: Python simulation of the trajectory of the deterministic logistic equation (respectively logistic SDE) in red (respectively black) with parameters $r = 1$, $\sigma = 0.25$ and $x_0 = 0.5$.

3.1.2 Stochastic nonpersistence

Now, if we suppose that $r < \frac{\sigma^2}{2}$ and recalling that stochastic nonpersistence does not require V to be nonnegative, let $V(x) = -\log(x)$ for all $x \in \mathbb{R}_+^*$. It implies that $L_e^+ V(x) = -(r-x) + \frac{\sigma^2}{2} =: H(x)$ is continuous on all $\mathbb{R}_+$ and we obviously have

$$\Lambda^+(H) = -\inf\{\mu H : \mu \mathcal{P}_{\text{erg}}(M_0)\} = -\delta_0(H) = (r-0) - \frac{\sigma^2}{2} < 0.$$

Moreover, $\log(x) \to +\infty$ as $x \to 0$ so $\{(X_t^x)_{t\geq 0} : x \in \mathbb{R}_+^*\}$ is H-nonpersistent in the sense of Definition 2.31.

Let's show that Hypotheses 7 and 8 holds true with the support of Lemma 2.37. In the first part, we show that (28) holds true with $U_\theta(x) = e^{\theta x}$ with $0 < \theta < \frac{2}{\sigma^2}$ hence Hypothesis 3 is satisfied with $W_\theta(x) = U_\theta(x)$. In addition,

$$\mathcal{L}_e W_\theta(x) = \theta x W_\theta(x)\left(r - x\left(1 - \theta\frac{\sigma^2}{2}\right)\right).$$

So let $G, \tilde{G}$ be defined as $G_{\theta'}(x) = W_{\theta'}(x)$ and $\tilde{G}_{\theta'}(x) = \tilde{W}_{\theta'}(x)$ with $0 < (2+\eta)\theta < \theta' < \frac{2}{\sigma^2}$, and where $0 < \eta < 1$ is fixed: since Hypothesis 2 holds for every $0 < \theta < \frac{2}{\sigma^2}$, then $\mathcal{L}_e G_{\theta'} \leq -\tilde{G}_{\theta'} + C$ is automatically satisfied. Furthermore,

$$\frac{\tilde{G}_{\theta'}(x)}{1 + |\mathcal{L}_e W_\theta(x)|} \geq \frac{\theta' x e^{\theta' x}\left(r - x\left(1 - \theta'\frac{\sigma^2}{2}\right)\right)}{1 + \theta x e^{\theta x}\left(r - x\left(1 - \theta\frac{\sigma^2}{2}\right)\right)} \underset{\theta' > \theta}{>} \frac{\theta' x e^{\theta' x}}{1 + \theta x e^{\theta x}} \underset{x\to\infty}{\longrightarrow} \infty,$$

also since $\theta' > \theta$, which implies that $\frac{\tilde{G}_{\theta'}}{1+|\mathcal{L}_e W_\theta|}$ is proper. By Lemma 2.37, Hypothesis 8 holds true.

Furthermore, we showed that $V \in \mathcal{D}_e^{2,+}$, and since $W_\theta = U_\theta \in C^2(\mathbb{R}_+)$, then $W_\theta \in \mathcal{D}_e^2$ by Proposition 2.1***(iii)***. Recalling (70), one has

$$\Gamma_e U_\theta(x) \leq c\sigma^2 U_\theta(x)^{2+\eta},$$

and since $(2+\eta)\theta < \theta'$, then $\Gamma_e(W_\theta) \leq \text{cst} \cdot U_{\theta'} \leq \text{cst} \cdot \tilde{G}_{\theta'}$. In addition, since

$$\Gamma_e^+(V) = (\sigma x V'(x))^2 = \sigma^2,$$

then $\Gamma_e(V) \leq \text{cst} \cdot \tilde{G}_{\theta'}$ so that by the second part of Lemma 2.37, Hypothesis 7 holds true.

By Theorem 2.33 and Corollary 2.36, it implies that $x_t^x \underset{t\to\infty}{\longrightarrow} 0,\ \forall x \in \mathbb{R}_+$.

Figure 2 illustrates the situation of extinction for the one-dimensional logistic SDE: given fixed parameters, with $1 =: r < \frac{\sigma^2}{2} := \frac{1.5^2}{2}$, its trajectory quickly goes to 0 which is the extinction set $M_0 = \{0\}$.

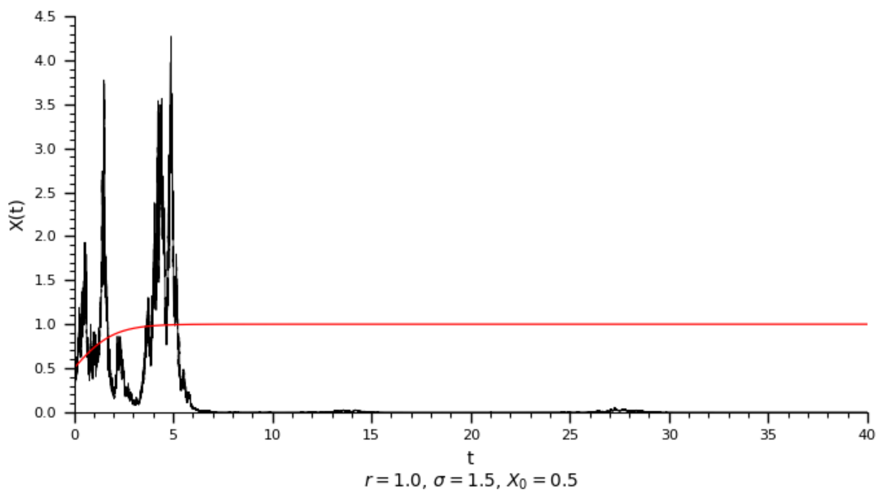

Figure 2: Python simulation of the trajectory of the deterministic logistic equation (respectively logistic SDE) in red (respectively black) with parameters $r = 1$, $\sigma = 1.5$ and $x_0 = 0.5$.

Remark 3.2. While $r - \frac{\sigma^2}{2} < 0$, note that

$$x e^{\left(r-\frac{\sigma^2}{2}\right)t+\sigma B_t} = x e^{t\left(r-\frac{\sigma^2}{2}+\frac{\sigma}{t}B_t\right)} \underset{t\to\infty}{\longrightarrow} 0, \quad \text{almost surely,}$$

since $\lim_{t\to\infty} \frac{B_t}{t} = 0$. In view of (69), $x_t^x \underset{t\to\infty}{\longrightarrow} 0$ and

$$\limsup_{t\to\infty} \frac{1}{t}\log(x_t^x) \leq \limsup_{t\to\infty} r - \frac{\sigma^2}{2} + \sigma\frac{B_t}{t} = r - \frac{\sigma^2}{2}. \tag{71}$$

3.2 A two-dimensional general SDE

In the spirit of the Rosenzweig-MacArthur model (3), we consider two interacting species whose dynamics follow the SDE on $\mathbb{R}_+^2$,

$$\begin{aligned} dx_1 &= x_1(F_1(x)dt + \sigma_1(x)dB_t^1), \\ dx_2 &= x_2(F_2(x)dt + \sigma_2(x)dB_t^2), \end{aligned} \tag{72}$$

where F_i, σ_i are C^∞, $\sigma_i > 0$ are bounded, and $(B_t^1)_{t\geq 0}, (B_t^2)_{t\geq 0}$ are independent standard Brownian motions. Assume the hypotheses of Proposition 2.1 and condition (35) (or its stronger version (45)) are satisfied.

To show that the process $(x_1(t), x_2(t))_{t\geq 0}$ is H−persistent, we need to define an extinction set M_0 and show that condition (36) holds true, i.e.

$$\exists p_1, p_2 > 0 \text{ such that } p_1\mu(\lambda_1) + p_2\mu(\lambda_2) > 0,\ \forall \mu \in \mathcal{P}_{\text{erg}}(M_0).$$

Let $M_+ = \{x \in \mathbb{R}^2_+ : x_1 > 0, x_2 > 0\}$ and $M_0 = \{x \in \mathbb{R}^2_+ : x_1 x_2 = 0\}$. The invasion rate is given by the formula

$$\lambda_i(x) = F_i(x) - \frac{\sigma_i^2(x)}{2}, \quad i = 1, 2.$$

Let's focus on the invariant face $\{x_2 = 0\}$: it is clear that, if $x_2 = 0$, then $\mathrm{d}x_2 = 0$ so that $(x_2(t))_{t\geq 0}$ is invariant along this face.

Then, the process admits the ergodic probability $\delta_{(0,0)}$ on M_0: indeed, if in addition we also have $x_1 = 0$, then $\mathrm{d}x_1 = 0$ so that $(x_1(t))_{t\geq 0}$ is also invariant, which implies that $P_t\delta_{(0,0)} = \delta_{(0,0)}$, i.e. $\delta_{(0,0)}$ is an invariant probability. Its ergodicity is direct since $(0, 0) \in M_0$ is absorbing.

But there exists also an invariant measure, not necessarily a probability, on $\{x_2 = 0, x_1 \neq 0\}$ given by the *speed measure* of a one-dimensional diffusion

$$\mu_1(\mathrm{d}x_1, \mathrm{d}x_2) = h_1(x_1)\mathrm{d}x_1\delta_0(\mathrm{d}x_2),$$

where

$$h_1(x_1) = \frac{2}{x_1^2\sigma_1^2(x_1, 0)} \exp\left\{\int_r^{x_1} \frac{2F_1(u, 0)}{u\sigma_1^2(u, 0)}\mathrm{d}u\right\}\mathbf{1}_{\{x_1>0\}}, \quad r > 0.$$

Indeed, consider a general one-dimensional diffusion

$$dX_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t,\ X_0 = x,$$

where b, σ are chosen so that $(X_t^x)_{t\geq 0}$ is a strong Markov process for all $x \in \mathbb{R}$ and the Markov semigroup $(P_t)_{t\geq 0}$ generated by the $(X_t^x)_{t\geq 0}$ is $C_b(\mathbb{R})$−Feller. It will be sufficient so that the next operations hold true and it will be also true for our example since the condition of Proposition 2.1 are satisfied, in particular Proposition 2.1***(i)*** holds true.

Let's suppose that $\sigma > 0$ to avoid doubt, which is also a condition for our based model (72). By the definition of its infinitesimal generator, we have

$$\mathcal{L}f(x) = b(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x), \quad f \in C_b^2(\mathbb{R}).$$

Its adjoint operator $\mathcal{L}^*$, acting on density functions ρ, is defined as

$$\mathcal{L}^*\rho(x) = -\frac{\partial}{\partial x}\left[b(x)\rho(x)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\sigma^2(x)\rho(x)\right].$$

It is well-known that its invariant measure $\rho(x)dx$ is the solution of $\mathcal{L}^*\rho = 0$, up to a constant, in the sense of the distribution: indeed, the measure $\rho(x)dx$ is invariant if

$$\int_{\mathbb{R}} (P_t f)(x)\rho(x)dx = \int_{\mathbb{R}} f(x)\rho(x)dx, \quad \forall t \geq 0,$$

and for all test functions $f \in \mathcal{B}_b(\mathbb{R})$. By taking the derivative with respect to t, which can be applied under the $\int_{\mathbb{R}}$-sign by standard arguments, the right-term is 0 and we have

$$\int_{\mathbb{R}} \frac{\partial}{\partial t}(P_t f)(x)\rho(x)dx = \int_{\mathbb{R}} \mathcal{L}(P_t f)(x)\rho(x)dx = \int_{\mathbb{R}} \mathcal{L}(f)(x)\rho(x)dx = 0, \quad \forall f \in \mathcal{B}_b(\mathbb{R}),$$

which is equivalent by taking the adjoint of $\mathcal{L}$, denoted $\mathcal{L}^*$, to

$$\int_{\mathbb{R}} f(x)\mathcal{L}^*(\rho)(x)dx = 0, \quad \forall f \in \mathcal{B}_b(\mathbb{R}).$$

Then, the condition $\mathcal{L}^*\rho = 0$ ensure that, for all test functions $f \in \mathcal{B}_b(\mathbb{R})$, $\int_{\mathbb{R}} f(x)\mathcal{L}^*(\rho)(x)dx = 0$ so that ρ is an invariant measure with respect to $(P_t)_{t\geq 0}$. In particular, let

$$\rho(x) = \frac{C}{\sigma^2(x)} \exp\left(2\int_0^x \frac{b(u)}{\sigma^2(u)}du\right),$$

where $C > 0$ is any constant, and let's verify that it is the solution of $\mathcal{L}^*\rho = 0$. We remark that

$$\sigma^2(x)\rho(x) = C\exp\left(2\int_0^x \frac{b(u)}{\sigma^2(u)}du\right) =: \alpha(x),$$

so we have $\frac{\partial}{\partial x}\alpha(x) = \frac{2\alpha(x)b(x)}{\sigma^2(x)}$, by using the fundamental theorem of calculus to compute $\frac{\partial}{\partial x}\int_0^x f(u)du = f(x)$. It follows that

$$\frac{\partial^2}{\partial x^2}\alpha(x) = \frac{4\alpha(x)b(x)^2}{\sigma^4(x)} + 2\alpha(x)\left(\frac{b'(x)\sigma^2(x) - 2b(x)\sigma'(x)\sigma(x)}{\sigma^4(x)}\right),$$

and

$$\begin{aligned}\frac{\partial}{\partial x}(b(x)\rho(x)) &= b'(x)\frac{\alpha(x)}{\sigma^2(x)} + b(x)\left(\frac{2\alpha(x)\frac{b(x)}{\sigma^2(x)}\sigma^2(x) - 2b(x)\sigma'(x)\sigma(x)}{\sigma^4(x)}\right)\\ &= b'(x)\frac{\alpha(x)}{\sigma^2(x)} + b(x)\left(\frac{2\alpha(x)b(x) - 2\alpha(x)\sigma'(x)\sigma(x)}{\sigma^4(x)}\right),\end{aligned}$$

so that

$$\begin{aligned}\mathcal{L}^*\rho(x) =& -\frac{b'(x)\alpha(x)}{\sigma^2(x)} - \frac{2\alpha(x)b(x)^2}{\sigma^4(x)} + \frac{2\alpha(x)b(x)\sigma'(x)\sigma(x)}{\sigma^4(x)} + \frac{2\alpha(x)b^2(x)}{\sigma^4(x)}\\ &+ \frac{\alpha(x)b'(x)}{\sigma^2(x)} - \frac{2\alpha(x)b(x)\sigma'(x)\sigma(x)}{\sigma^4(x)}\\ =& \ 0.\end{aligned}$$

In our case, since $b(x) = x_1F_1(x_1, 0)$ and $\sigma(x) = x_1\sigma_1(x_1, 0)$ for $x_1 > 0$, we obtain

$$\begin{aligned}h_1(x_1) &= \frac{C}{x_1^2\sigma_1^2(x_1,0)}\exp\left(2\int_0^{x_1}\frac{uF_1(u,0)}{u^2\sigma_1^2(u,0)}du\right)\mathbf{1}_{x_1>0}\\ &= \frac{1}{x_1^2\sigma_1^2(x_1,0)}\exp\left(2\int_r^{x_1}\frac{F_1(u,0)}{u\sigma_1^2(u,0)}du\right)\mathbf{1}_{x_1>0},\end{aligned}$$

where $r > 0$ in the lower integration bound corresponds to the weight of the constant C. It remains to show that this measure is finite so that we can normalize it to have our invariant probability: given our hypotheses, we have already a control for $x_1 \to \infty$ with a Lyapunov function. We control $x_1 \to 0$ to avoid explosion: by continuity of F_1, σ_1,

$$\frac{F_1(x_1,0)}{\sigma_1^2(x_1,0)} \underset{x_1\to 0}{\longrightarrow} \frac{F_1(0,0)}{\sigma_1^2(0,0)},$$

which leads to

$$\lim_{x_1\to 0} h_1(x_1) = \lim_{x_1\to 0} \frac{1}{x_1^2\sigma_1^2(0,0)} \exp\left(2\frac{F_1(0,0)}{\sigma_1^2(0,0)} \int_r^{x_1} \frac{1}{u} du\right) \leq \frac{1}{\sigma_1^2(0,0)} x_1^{2\frac{F_1(0,0)}{\sigma_1^2(0,0)}-2},$$

which is integrable for small x_1 if $x_1^{2\frac{F_1(0,0)}{\sigma_1^2(0,0)}-2}$ is integrable, i.e.

$$2\frac{F_1(0,0)}{\sigma_1^2(0,0)} - 2 > -1 \;\Leftrightarrow\; 2\frac{F_1(0,0)}{\sigma_1^2(0,0)} > 1,$$

which is equivalent to the condition on the invasion rate

$$\lambda_1(0,0) = F_1(0,0) - \frac{\sigma_1^2(0,0)}{2} > 0.$$

Otherwise, the integral will explode when $x_1 \to 0$: under this integrability hypothesis, we can normalize the invariant measure to get an invariant, ergodic probability on $\{x_1 > 0, x_2 = 0\}$ given by

$$\mu_1(\mathrm{d}x_1\mathrm{d}x_2) = \frac{h_1(x_1)}{\int_0^\infty h_1(u)du} \mathrm{d}x_1\delta_0(\mathrm{d}x_2).$$

The same can be applied to the invariant face $x_1 = 0$, which gives us the invariant, ergodic probability on $\{x_1 = 0, x_2 > 0\}$

$$\mu_2(\mathrm{d}x_1\mathrm{d}x_2) = \delta_0(\mathrm{d}x_1)\frac{h_2(x_2)}{\int_0^\infty h_2(u)du}\mathrm{d}x_2,$$

under the hypothesis $\lambda_2(0,0) > 0$ to ensure the integrability of h_2. When $x_1 = x_2 = 0$, then $\delta_{(0,0)}(\mathrm{d}x_1\mathrm{d}x_2)$ is the same ergodic probability than before.

To summarize the situation, we can describe $\mathcal{P}_{\mathrm{erg}}(M_0)$ as follows:

$$\mathcal{P}_{\mathrm{erg}}(M_0) = \begin{cases} \{\delta_{(0,0)}\}, \text{ if } \lambda_1(0,0) < 0, \lambda_2(0,0) < 0, \text{ when } x_1 = x_2 = 0, \\ \{\delta_{(0,0)}, \mu_1\}, \text{ if } \lambda_1(0,0) > 0, \lambda_2(0,0) < 0, \text{ when } x_1 > 0, x_2 = 0, \\ \{\delta_{(0,0)}, \mu_2\}, \text{ if } \lambda_1(0,0) < 0, \lambda_2(0,0) > 0, \text{ when } x_1 = 0, x_2 > 0, \\ \{\delta_{(0,0)}, \mu_1, \mu_2\}, \text{ if } \lambda_1(0,0) > 0, \lambda_2(0,0) > 0, \text{ when } x_1, x_2 > 0. \end{cases}$$

For each case, we can check condition (36):

(i) If $\lambda_1(0,0), \lambda_2(0,0) < 0$ the only ergodic probability is $\delta_{(0,0)}$ and

$$p_1\delta_{(0,0)}(\lambda_1) + p_2\delta_{(0,0)}(\lambda_2) = p_1\lambda_1(0,0) + p_2\lambda_2(0,0) < 0,$$

so (36) is not verified.

(ii) If $\lambda_1(0,0) > 0$, $\lambda_2(0,0) < 0$, we check with $\mu = \delta_{(0,0)}$ and

$$p_1\delta_{(0,0)}(\lambda_1) + p_2\delta_{(0,0)}(\lambda_2) = p_1\lambda_1(0,0) + p_2\lambda_2(0,0) > 0,$$

with p_1 sufficiently large compared to p_2, in the sense that $p_1 > p_2\frac{\lambda_2(0,0)}{\lambda_1(0,0)}$.

For $\mu = \mu_1$, by contrapositive of Theorem 2.8 *(i)*, $\mu_1(\lambda_1) = 0$ since $\mu_1(x_1 \neq 0) = 0$, i.e. $\mathrm{supp}(\mu_1) \not\subseteq \{x_1 = 0\}$. It remains

$$p_1\mu_1(\lambda_1) + p_2\mu_1(\lambda_2) = p_2\mu_1(\lambda_2) > 0, \text{ if } \mu_1(\lambda_2) > 0.$$

(iii) If $\lambda_1(0,0) < 0$, $\lambda_2(0,0) > 0$: the same reasoning applies to check the condition with $\delta_{(0,0)}$. And a similar argument shows that $\mu_2(\lambda_2) = 0$ so that

$$p_1\mu_2(\lambda_1) + p_2\mu_2(\lambda_2) = p_1\mu_2(\lambda_1) > 0, \text{ if } \mu_2(\lambda_1) > 0.$$

(iv) If $\lambda_1(0,0) > 0$, $\lambda_2(0,0) > 0$, the reasoning is still the same for $\delta_{(0,0)}$ and given the fact that $\mu_1(\lambda_1) = \mu_2(\lambda_2) = 0$, either

$$p_1\mu_1(\lambda_1) + p_2\mu_1(\lambda_2) = p_2\mu_1(\lambda_2) > 0 \text{ if } \mu_1(\lambda_2) > 0,$$

or

$$p_1\mu_2(\lambda_1) + p_2\mu_2(\lambda_2) = p_1\mu_2(\lambda_1) > 0, \text{ if } \mu_2(\lambda_1) > 0.$$

Finally, since we are in a nondegenerate case, we can apply the conclusions of Corollary 2.13 if:

(i) $\lambda_1(0,0), \lambda_2(0,0) < 0$, the condition (36) is not verified.

(ii) $\lambda_1(0,0) > 0, \lambda_2(0,0) < 0$ (respectively $\lambda_1(0,0) < 0, \lambda_2(0,0) > 0$, the condition (36) is verified if $\mu_1(\lambda_2) > 0$ (respectively $\mu_2(\lambda_1) > 0$).

(iii) $\lambda_1(0,0), \lambda_2(0,0) > 0$, the condition (36) is verified if $\mu_1(\lambda_2), \mu_2(\lambda_1) > 0$.

Whenever condition (36) is verified, there is a unique persistent measure Π on M_+ and, under the stronger Lyapunov conditions, convergence in Total variation and exponential ergodicity hold. Conclusions obtained here are similar to the conditions exhibited in [HN18], Example 2.4.

Chapter 4

The degenerate Rosenzweig-MacArthur model in details

This chapter is dedicated to the study of our main model, the two-dimensional degenerate Rosenzweig-MacArthur model (3) and covers the persistence and extinction cases.

More precisely, Chapter 4 goes as follows:

(i) Section 4.1 presents the well-posedness of (3) by detailing how Hypotheses 1-3 are verified, with respect to the parameters of the model.

(ii) Section 4.2 details the persistence case through the notion of $H-$persistence and shows the almost-sure convergence (respectively in Total variation) of the empirical measure (respectively $(P_t(x,\cdot))_{t\geq 0}$) to the persistence measure. We also exhibit a polynomial convergence rate and depict the situation through a Python simulation.

(iii) Section 4.3 outlines the nonpersistence case, in particular the extinction of the predator x_2. We detail how to show Hypotheses 7 and 8 and we depict the situation through a Python simulation.

(iv) Section 4.4 focuses on the extinction of both species by changing the assumptions on the parameters of the model and adapting our proof, compared to previous section.

Recall that our model of interest (3) is the degenerate SDE defined on $\mathbb{R}_+^2$ and following

$$\begin{cases} \mathrm{d}x_1 = x_1\left[\left(1 - \frac{x_1}{\kappa} - \frac{x_2}{1+x_1}\right)\mathrm{d}t + \varepsilon \mathrm{d}B_t\right], \\ \mathrm{d}x_2 = x_2\left(-\alpha + \frac{x_1}{1+x_1}\right)\mathrm{d}t, \end{cases}$$

where $\kappa, \varepsilon > 0$, $0 < \alpha < 1$, and $(B_t)_{t\geq 0}$ is a $1-$dimensional standard Brownian motion. Recall that for $0 < \varepsilon^2 < 2$, we define $k := \frac{2}{\varepsilon^2} - 1 > 0$, $\theta := \frac{\varepsilon^2\kappa}{2} < \kappa$, and

$$\gamma_{\varepsilon,\kappa}(x) = \frac{x^{k-1}e^{-x/\theta}}{\Gamma(k)\theta^k}, \quad x \geq 0,$$

which is the density of a Γ-distribution with parameters k, θ whose expectation is $k\theta = \kappa(1 - \frac{\varepsilon^2}{2})$ and variance $k\theta^2 = \frac{\kappa^2\varepsilon^2}{2}(1 - \frac{\varepsilon^2}{2})$. We also define

$$\Lambda(\varepsilon, \alpha, \kappa) = \int_0^{+\infty} \frac{x}{1+x}\gamma_{\varepsilon,\kappa}(x)dx - \alpha.$$

We also recall that the extinction set is defined as

$$M_0 = \{(x_1, x_2) \in \mathbb{R}_+^2 : x_1x_2 = 0\} = \partial\mathbb{R}_+^2.$$

4.1 Proof of Hypotheses 1-3

First of all, we can use Proposition 2.1 above to show that the main Hypotheses 1-3 are verified. Here, we use a form of U inspired by [Ben18] but the proof is more detailed and includes the tools newly integrated throughout this text.

Proposition 4.1. *In the context of (3), with $0 < \varepsilon^2 < 2$, conditions (28) and (29) are verified with $U(x) = 1 + (x_1 + x_2)^n$, for all $n > 2$. In particular Hypotheses 1-3 hold for (3).*

Proof. Using the formula (26),

$$
\begin{aligned}
LU(x_1,x_2) &= \left[x_1\left(1-\frac{x_1}{\kappa}-\frac{x_2}{1+x_1}\right)+x_2\left(-\alpha+\frac{x_1}{1+x_1}\right)\right]n(x_1+x_2)^{n-1}\\
&\quad+\frac{1}{2}\varepsilon^2 x_1^2 n(n-1)(x_1+x_2)^{n-2}\\
&= n(x_1+x_2)^{n-1}\left(x_1-\frac{x_1^2}{\kappa}-\alpha x_2\right)+n\left(\frac{n-1}{2}\right)x_1^2\varepsilon^2(x_1+x_2)^{n-2}\\
&\le n(x_1+x_2)^{n-1}\left(x_1\left(1-\frac{x_1}{\kappa}+\left(\frac{n-1}{2}\right)\varepsilon^2\right)-\alpha x_2\right)\\
&= -\alpha n(x_1+x_2)^n+n(x_1+x_2)^{n-1}\left(x_1\left(1-\frac{x_1}{\kappa}+\left(\frac{n-1}{2}\right)\varepsilon^2+\alpha\right)\right).
\end{aligned}
$$

Let

$$
M=\kappa\Big(1+\frac{n-1}{2}\varepsilon^2+\alpha\Big),\quad C_1=nM\left|1-\frac{M}{\kappa}+\frac{n-1}{2}\varepsilon^2+\alpha\right|.
$$

Then on $\{x_1 > M\}$ the second term is negative, so

$$
LU(x_1,x_2)\le -\alpha n(x_1+x_2)^n = -\alpha n U(x_1,x_2)+\alpha n.
$$

On $\{0 \le x_1 \le M\}$, we observe that

$$
LU(x_1,x_2)\le -\alpha n(x_1+x_2)^n + C_1(x_1+x_2)^{n-1}.
$$

Since the leading term $-\alpha n(x_1 + x_2)^n$ dominates $C_1(x_1 + x_2)^{n-1}$ as $x_1 + x_2 \to \infty$, there exists $R > 0$ so that for all $x_1 + x_2 > R$,

$$
-\alpha n(x_1+x_2)^n + C_1(x_1+x_2)^{n-1}\le -\frac{\alpha n}{2}(x_1+x_2)^n,
$$

hence on $\{0 \le x_1 \le M, x_1 + x_2 > R\}$,

$$
LU(x_1,x_2)\le -\frac{\alpha n}{2}U(x_1,x_2)+\frac{\alpha n}{2}.
$$

In fact, we can exhibit that $R = \frac{2C_1}{\alpha n} > 0$ is sufficient by direct computation. Finally, on the compact set $\{0 \le x_1 \le M, 0 \le x_1 + x_2 \le R\}$, LU is bounded above and thus

$$
LU(x_1,x_2)\le -\alpha n U(x_1,x_2)+\alpha n+\frac{2^{n-1}C_1^n}{\alpha^{n-1}n^{n-1}},\quad \forall x\in\mathbb{R}_+^2,
$$

so that (28) holds globally with $a = \frac{\alpha n}{2}$, $b = \alpha n + \frac{2^{n-1}C_1^n}{\alpha^{n-1}n^{n-1}}$.

In addition, using the formula (27), we have

$$
\Gamma_L(U)(x) = x_1^2\varepsilon^2 n(x_1+x_2)^{n-1}n(x_1+x_2)^{n-1}\le \varepsilon^2 n^2(x_1+x_2)^{2n}, \tag{73}
$$

so that (29) holds with $\eta = 0$ and $c := \varepsilon^2 n^2$. □

At this point, we have already that $(\Pi_t^x)_{t\ge 0}$ is almost surely tight, by Theorem 1.40, and both $\mathcal{P}_{\text{inv}}(M)$ and $\mathcal{P}_{\text{inv}}(M_0)$ are non-empty, by Remark 1.41.

Proposition 4.2. *Condition (28) is verified with* $U(x_1, x_2) = e^{\theta(x_1+x_2)}$, $\forall \theta < \theta^* := \frac{2}{\kappa\varepsilon^2}$

Remark 4.3. By a scaling the time t by a factor r in (65), we obtain

$$dx_s = x_s\left(\left(1 - \frac{x_s}{r}\right)ds + \frac{\sigma}{\sqrt{r}}dB_s\right).$$

Let $r = \kappa$ and $\sigma = \varepsilon\sqrt{r} = \varepsilon\sqrt{\kappa}$ so that it matches the x_1−coordinate of the Rosenzweig-MacArthur model (3) when x_2 is set to 0. In particular, the condition ensuring (28) holds with $U(x) = e^{\theta x}$ is $0 < \theta < \frac{2}{\sigma^2} = \frac{2}{\varepsilon^2\kappa}$.

Moreover, the H−persistent condition $r > \frac{\sigma^2}{2} > 0$ becomes $0 < \varepsilon^2 < 2$, while extinction is deducted from $r < \frac{\sigma^2}{2}$ which is $\varepsilon^2 > 2$.

Proof of Proposition 4.2. Remark that

$$\begin{aligned}LU(x_1, x_2) &= x_1\left(1 - \frac{x_1}{\kappa} - \frac{x_2}{1 + x_1}\right)\theta e^{\theta(x_1+x_2)} + x_2\left(-\alpha + \frac{x_1}{1 + x_1}\right)\theta e^{\theta(x_1+x_2)} + \frac{\varepsilon^2}{2}\theta^2 x_1^2 e^{\theta(x_1+x_2)} \\ &\le \theta x_1 e^{\theta(x_1+x_2)}\left(\left(1 - \frac{x_1}{\kappa}\right) + \frac{\varepsilon^2}{2}\theta x_1\right),\end{aligned}$$

so that $LU \le -aU + b$ holds for any $\theta < \theta^* = \frac{2}{\varepsilon^2\kappa}$ by analogy with the one-dimensional logistic SDE (65) (see Remark 4.3). □

Remark 4.4. Unfortunately, condition (29) fails for $U(x) = e^{\theta(x_1+x_2)}$ since

$$\Gamma(U(x_1, x_2)) = \varepsilon^2\theta^2 x_1^2 U^2(x_1, x_2).$$

However, the existence of such an exponential function such that (28) holds will be useful to derive some properties, for example about the type of functions that are integrable with respect to the persistent measure Π.

4.2 Persistence in the case $0 < \varepsilon^2 < 2$ and $\Lambda(\varepsilon, \alpha, \kappa) > 0$

We establish a direct condition on $\Lambda(\varepsilon, \alpha, \kappa)$ to ensure the persistence of the model. Figure 3, obtained in Python, illustrates the behavior of the process when $0 < \varepsilon^2 < 2$ $\Lambda(\varepsilon, \alpha, \kappa) > 0$. We start by announcing the final result of this subsection:

Theorem 4.5 (Persistence). *Suppose that* $0 < \varepsilon^2 < 2$ *and* $\Lambda(\varepsilon, \alpha, \kappa) > 0$. *Then, there exists a unique invariant probability measure* Π *on* M_+ *such that, for all initial condition* $x \in M_+$*:*

(i) $(\Pi_t^x)_{t\ge0} \Rightarrow \Pi$, *almost surely.*

(ii) *For all* $f \in L^1(\Pi)$ *such that* $\int_0^T f(X_s^x)ds < \infty$ *for all* $T > 0$,

$$\Pi_t^x f \underset{t\to\infty}{\longrightarrow} \Pi f, \quad \textit{almost surely.}$$

Moreover, for all $\theta < \frac{2}{\kappa\varepsilon^2}$, $(x_1, x_2) \mapsto e^{\theta(x_1+x_2)}$ *lies in* $L^1(\Pi)$, *i.e.* $\int_{M_+} e^{\theta(x_1+x_2)}\Pi(\mathrm{d}x_1\mathrm{d}x_2) < \infty$.

(iii) $(P_t(x, \cdot))_{t\ge0}$ *converges in Total variation towards* Π *at a polynomial rate, in the sense that there exists* $\lambda > 0$ *such that*

$$\lim_{t\to\infty} t^\lambda \|P_t(x, \cdot) - \Pi(\cdot)\|_{TV} = 0.$$

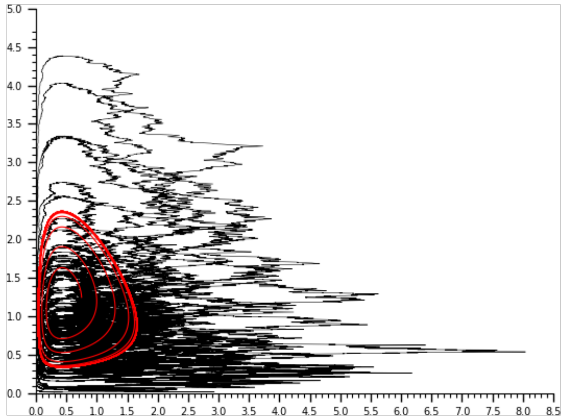

Figure 3: Simulation of (3) starting at $(x_1(0), x_2(0)) = (0.75, 1.25)$ in persistence case with $\Lambda(\varepsilon = 0.6,\ \alpha = 0.3,\ \kappa = 2.5) \approx 0.34 > 0$. The red trajectory is a trajectory of the unperturbed system (i.e. for $\varepsilon = 0$) while the black one describes the system (3) with an Euler–Maruyama scheme.

(iv) *Π has a smooth density (with respect to the Lebesgue measure), strictly positive, on M_+.*

Before proving Theorem 4.5, we need to prove some properties associated to (3). First, we show that under those conditions, (3) is $H-$persistence thanks to the Hofbauer criterion from Theorem 2.8.

Corollary 4.6. *In the context of (3), we suppose that $0 < \varepsilon^2 < 2$ and $\Lambda(\varepsilon, \alpha, \kappa) > 0$. Then, conditions (35) and (36) of Theorem 2.8 are verified and $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ is $H-$persistent.*

Proof. When we choose $U(x_1, x_2) = 1 + (x_1 + x_2)^n$, for $n > 2$, since $1 + \sum_{i=1}^{2} |F_i| \leq 2 + \frac{x_1}{\kappa} + \frac{x_2}{1+x_1} + \alpha + \frac{x_1}{1+x_1} \leq \left(1 + \frac{1}{\kappa}\right) \cdot (x_1 + x_2) + 2 + \alpha$, thus (35) holds with

$$\limsup_{\|x\|\to+\infty} \frac{U^{\frac{1}{2}}(x)}{1 + \sum_i |F_i(x)|} \geq \limsup_{\|x\|\to+\infty} \frac{(x_1 + x_2)^{\frac{n}{2}}}{\left(1 + \frac{1}{\kappa}\right) \cdot (x_1 + x_2) + 2 + \alpha} = +\infty.$$

When $U(x_1, x_2) = e^{\theta(x_1+x_2)}$, for $\theta < \frac{2}{\kappa\varepsilon^2}$, the same conclusion is direct.

It remains to verify the condition (36) of Theorem 2.8. We aim to evaluate $\mathcal{P}_{\text{erg}}(M_0^2)$. If we restrict to the invariant face $x_2 = 0 \Rightarrow \mathrm{d}x_2 = 0$, we are left with two possibilities:

(*i*) $x_1 = 0$ implies $\mathrm{d}x_1 = 0$ and the invariant probability measure is $\delta_{(0,0)}$, which is ergodic on $\{x \in \mathbb{R}_+^2 : x_1 x_2 = 0\}$ since it charges the entire set.

(*ii*) $x_1 \neq 0$ and we can analyze an invariant measure (not necessarily a probability) which is the speed measure associated to:

$$\mathrm{d}x_1 = x_1\left(1 - \frac{x_1}{k}\right)\mathrm{d}t + \varepsilon x_1 \mathrm{d}B_t.$$

As in the computation of model (72), the speed measure is given by

$$\begin{aligned}
h_1(x_1) &= \frac{2}{x_1^2\varepsilon^2} \exp\left(2\int_r^{x_1} \frac{u(1 - \frac{u}{\kappa})}{u^2\varepsilon^2} du\right) \\
&= \frac{2}{x_1^2\varepsilon^2} \exp\left(\frac{2}{\varepsilon^2\kappa}\left[\kappa\log(x_1) - x_1 - \kappa\log(r) + r\right]\right) \\
&= \frac{2 \cdot C}{\varepsilon^2} x_1^{\frac{2}{\varepsilon^2}-2} \exp\left(\frac{-2x_1}{\kappa\varepsilon^2}\right) \\
&= \frac{2 \cdot C}{\varepsilon^2} x_1^{\frac{2}{\varepsilon^2}-2} \exp\left(\frac{-x_1}{(\kappa\varepsilon^2)/2}\right),
\end{aligned}$$

where C is a constant coming from evaluating the integral at $u = r$. In particular, with $\frac{2}{\varepsilon^2} - 1 = k$, $\frac{\varepsilon^2 \kappa}{2} = \theta$, we obtain $h_1(x_1) = \frac{2 \cdot C}{\varepsilon^2} x_1^{k-1} \exp\left(-\frac{x_1}{\theta}\right)$.

We observe that this is a finite measure since $k - 1 > -1$ so that h_1 is finite for small x_1. And as $x_1 \to \infty$, h_1 is also finite since $\exp\left(-\frac{x_1}{\theta}\right) \to 0$ quickly than $x_1^{k-1} \to \infty$.

It remains to normalize the measure to obtain a probability. By a substitution $u = \frac{t}{\theta}$, the normalization constant is given by

$$\int_0^{+\infty} h_1(t)\mathrm{d}t = \int_0^{+\infty} t^{k-1} \exp\left(-\frac{t}{\theta}\right)\mathrm{d}t = \int_0^{+\infty} \theta^{k-1} u^{k-1} \exp(-u)\theta du = \Gamma(k) \cdot \theta^k.$$

Thus, we obtain the probability measure μ defined as

$$\begin{aligned} \mu_1(\mathrm{d}x_1 \mathrm{d}x_2) &= \frac{h_1(x_1)}{\int_0^{+\infty} h_1(t)\mathrm{d}t} \mathrm{d}x_1 \cdot \delta_0(\mathrm{d}x_2) \\ &= \frac{x_1^{k-1} \exp(-x_1/\theta)}{\Gamma(k)\theta^k} \mathrm{d}x_1 \cdot \delta_0(\mathrm{d}x_2) \\ &= \gamma_{\varepsilon,\kappa}(x_1)\mathrm{d}x_1 \cdot \delta_0(\mathrm{d}x_2) \end{aligned} \tag{74}$$

The same method can be applied to the face $x_1 = 0$ where we have $\mathrm{d}x_1 = 0$ and $\mathrm{d}x_2 = -\alpha x_2 \mathrm{d}t$, so $x_2(t) = x_2(0) \cdot e^{-\alpha t}$, which implies that the only invariant probability measure is $\delta_{(0,0)}$, which is also ergodic. Finally,

$$\mathcal{P}_{\mathrm{erg}}(M_0) = \begin{cases} \mu_1 & \text{if } x_2 = 0,\ x_1 > 0 \\ \delta_{(0,0)} & \text{if } x_1 = 0,\ x_2 \geq 0 \end{cases}$$

We treat each case independently:

(i) Let $\mu = \delta_{(0,0)} \in \mathcal{P}_{\mathrm{erg}}(M_0)$, then

$$\sum_i p_i \mu(\lambda_i) = p_1 \delta_{(0,0)}(\lambda_1) + p_2 \delta_{(0,0)}(\lambda_2) = p_1 \cdot \lambda_1(0,0) + p_2 \cdot \lambda_2(0,0) > 0.$$

By definition of λ_i in (33), it follows that

$$\lambda_1(0,0) = F_1(0,0) - \frac{\varepsilon^2}{2} = 1 - \frac{\varepsilon^2}{2} > 0, \quad \text{since } \varepsilon^2 < 2$$

And similarly,

$$\lambda_2(0,0) = F_2(0,0) = -\alpha < 0,$$

thus $p_1\mu(\lambda_1) + p_2\mu(\lambda_2) > 0$ for p_1 large enough, p_2 small enough, in the sense that $p_1 > \frac{p_2 \alpha}{1 - \frac{\varepsilon^2}{2}}$ for any $p_2 > 0$.

(ii) Let $\mu = \mu_1 \in \mathcal{P}_{\mathrm{erg}}(M_0)$. In particular, $\mathrm{supp}(\mu_1) \not\subseteq \{x \in \mathbb{R}^2 : x_1 = 0\}$, and by contrapositive of Theorem 2.8*(i)*, we have $\mu_1(\lambda_1) = 0$ so that condition (36) becomes $p_2\mu(\lambda_2) > 0 \iff \mu(\lambda_2) > 0$. Since $\lambda_2(x_1, x_2) = F_2(x_1, x_2) = -\alpha + \frac{x_1}{1+x_1}$, we have:

$$\mu(\lambda_2) = \int_0^{+\infty} \left(\frac{x}{1+x} - \alpha\right) \gamma_{\varepsilon,\kappa}(x) dx = \Lambda(\varepsilon, \alpha, \kappa),$$

which is supposed to be strictly positive.

Therefore, condition (36) is verified and we can conclude that $\{(X_t^x)_{t \geq 0} : x \in M_+\}$ is H-persistent by Theorem 2.8. □

Remark 4.7. When we evaluate the speed measure on the invariant face $x_2 = 0$, it also provides the unique ergodic probability measure on $x_1 > 0$ for the one-dimensional logistic SDE by Remark 4.3 with $\varepsilon = \frac{\sigma}{\sqrt{r}}$ and $\kappa = r$.

We now aim to prove the existence of points x^* satisfying Hörmander's condition (respectively the strong Hörmander condition) in Γ_{M_+}, in order to show almost-sure convergence (respectively in Total variation) of the occupation measure (respectively convergence of $(P_t(x, \cdot))_{t\geq 0}$) towards the unique invariant probability measure.

Proposition 4.8. *In the situation of (3), the strong Hörmander condition holds for every points $x^* \in M_+$.*

Proof. We rewrite (3) using the Stratonovich formalism introduced in (37) as

$$S_1^1(x) = x_1 \cdot \varepsilon,\ S_2^1(x) = S_1^2(x) = S_2^2(x) = 0,$$

and therefore

$$S^0(x) = \begin{pmatrix} x_1(F_1(x_1,x_2) - \varepsilon^2/2) \\ x_2 F_2(x_1,x_2) \end{pmatrix}, \quad S^1(x) = \begin{pmatrix} x_1\varepsilon \\ 0 \end{pmatrix}.$$

We obtain

$$[S^0, S^1](x) = DS^1(x)\cdot S^0(x) - DS^0(x)\cdot S^1(x) = \begin{pmatrix} -x_1^2\varepsilon \cdot \frac{\partial F_1}{\partial x_1}(x) \\ -x_1\varepsilon \cdot x_2 \cdot \frac{\partial F_2}{\partial x_1}(x) \end{pmatrix}.$$

In particular,

$$\begin{aligned} \det\Big([S^0, S^1](x), S^1(x)\Big) &= \det\begin{pmatrix} -x_1^2\varepsilon \cdot \frac{\partial F_1}{\partial x_1}(x) & x_1\varepsilon \\ -x_1\varepsilon \cdot x_2 \cdot \frac{\partial F_2}{\partial x_1}(x) & 0 \end{pmatrix} \\ &= x_1^2\varepsilon^2 x_2 \cdot \frac{1}{(1+x_1)^2} > 0, \quad \forall x \in M_+, \end{aligned}$$

which proves that, for all $x^* \in M_+$, the set $\{[S^0, S^1](x^*), S^1(x^*)\}$ spans $\mathbb{R}^2$, and the strong Hörmander condition (in particular, Hörmander condition) holds on M_+. □

The last step is to show that Γ_{M_+} contains M_+. Here, we use the notion of control system introduced in (38).

Proposition 4.9. *In the situation of (3), $\Gamma_{M_+} = M_+$ and every points of M_+ are accessible from anywhere in M_+.*

Proof. In our model, let's introduce a new control variable

$$v = \varepsilon u - \frac{\varepsilon^2}{2} \Leftrightarrow u = \frac{v}{\varepsilon} + \frac{\varepsilon}{2},$$

so that the associated control system $y(t)$ is the solution of the differential equation

$$\begin{aligned} \dot{y}(t) &= S^0(y(t)) + \sum_{j=1}^m u^j S^j(y(t)) \\ &= \begin{pmatrix} y_1(t)\Big(F_1(y_1(t), y_2(t)) - \varepsilon^2/2\Big) \\ y_2(t) F_2(y_1(t), y_2(t)) \end{pmatrix} + \begin{pmatrix} \varepsilon y_1(t) \\ 0 \end{pmatrix} \cdot \begin{pmatrix} u \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} y_1(t)\,(F_1(y_1(t), y_2(t)) + v) \\ y_2(t) F_2(y_1(t), y_2(t)) \end{pmatrix}, \end{aligned} \tag{75}$$

Let L be the vertical line defined by $x_1 = \frac{\alpha}{1-\alpha}$, and let $x_2 \geq 0$: it implies that, along this line, we have $\mathrm{d}x_2 = 0$ since

$$F_2(x_1, x_2) = F_2\left(\frac{\alpha}{1-\alpha}, x_2\right) = -\alpha + \alpha = 0.$$

Before this line, the drift associated to the second coordinate of the control system $y(t)$ is negative, i.e. $y_2(t)$ is decreasing, while it is increasing after the line.

Let $P_v(x_1, x_2)$ be the parabola defined by

$$(1 + v - \frac{x_1}{\kappa})(1 + x_1) = x_2 \iff -\frac{1}{\kappa}x_1^2 + \left(v - \frac{1}{\kappa} + 1\right)x_1 + 1 + v = x_2.$$
$$\iff F_1(x_1, x_2) + v = 0.$$

This last equality implies that below P_v, the drift associated to the first coordinate of the control system $y(t)$ is positive, and $y_1(t)$ is increasing, while it is decreasing above P_v.

It naturally implies that P_v reaches its maximum (x_1^*, x_2^*) at

$$x_1^* = \frac{1}{2}\left(\kappa v - 1 + \kappa\right),$$
$$x_2^* = \left(1 + v - \frac{(\kappa v - 1 + \kappa)}{2\kappa}\right)\left((1 + \frac{(\kappa v - 1 + \kappa)}{2}\right) = \frac{(\kappa(v + 1) + 1)^2}{4\kappa}.$$

Let $z = (z_1, z_2) \in M_+$ and O_z be an open neighborhood of z: we can choose v^* sufficiently large so that z is always below the graph of P_{v^*} and such that the x_1−coordinate of the maximum of P_{v^*} is after the line $\frac{\alpha}{1-\alpha}$.

For a small-enough open neighborhood of the origin O, $\forall x \in O \cap M_+$, the control system $t \mapsto y(v^*, x, t)$ remains below P_{v*} until it crosses P_{v*} near $(\kappa(1 + v^*), 0)$. Then, it remains above P_{v*} while being decreasing on x_1, increasing on x_2 until it crosses L, when it becomes decreasing on both x_1 and x_2. In particular, it crosses $x_2 = z_2$ while $x_1 > z_1$.

Let's construct a piecewise constant control $v(t)$ as follows:

- (i) $v(t) = -1$ until $t \mapsto y(v, x, t)$ enters O.
- (ii) Then, $v(t) = v^*$ until $t \mapsto y(v, x, t)$ crosses $x_2 = z_2$ (after crossing P_{v^*}).
- (iii) Finally, $v(t) = -R$ for R large enough so that $t \mapsto y(v, x, t)$ enters O_z, i.e. decreasing on x_1 as quickest as possible while remaining in a small-enough neighborhood of $x_2 = z_2$.

Figure 4 exhibit the structure of the associated control system through an example: starting at $x = (0.3, 0.3)$, $z = (1, 2)$ with $O_z = B_{0.15}(z)$, $O = B_{0.15}(0, 0)$, $\alpha = 0.3$, $\kappa = 0.5$ and $v^* = 3$.

Starting from x, the control system is sent to $O \cap \mathbb{R}_+^2$, then go to the right until it passes P_{v^*}, moves up to crosses the line $x_2 = z_2$ and is finally sent to O_z.

It implies by Stroock and Varadhan support theorem (Proposition 2.11) that $M_+ = \Gamma_{M_+}$ since any point $z \in M_+$ is accessible from any point $x \in M_+$, i.e. $\Gamma_{M_+} = \cap_{z \in M_+}\Gamma_z = M_+$. □

We are now interest by the convergence rate of $(P_t(x, \cdot))_{t\geq 0}$ towards the persistence measure $\Pi(\cdot)$, for all $x \in M_+$.

Proposition 4.10. *Hypothesis 6 holds with* $U(x_1, x_2) = 1 + (x_1 + x_2)^n$, *for* $n > 2$ *so that* $(P_t(x, \cdot))_{t\geq 0}$ *converges in Total variation towards* Π *at a polynomial rate.*

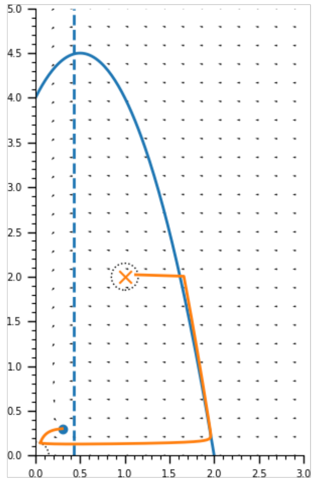

Figure 4: Example of (75) with $x = (0.3, 0.3)$, $z = (1, 2)$ with $O_z = B_{0.15}(z)$, $O = B_{0.15}(0, 0)$, $\alpha = 0.3$, $\kappa = 0.5$ and $v^* = 3$. The trajectory of $y(x, v, t)$ with the proper v as defined before is depicted in orange, with the blue parabola P_{v*} whose maximum is attained after the vertical blue dashed line L.

Proof. Condition (49) is direct since $\ln(\|x\|) = \frac{1}{2}\ln(x_1^2 + x_2^2) \ll x_1^n + x_2^n \leq U(x)$, since $n > 2$.

We also recall that $|F_1(x)| + |F_2(x)| \leq (2 + \frac{1}{\kappa} + \alpha)(1 + x_1 + x_2)$, so that (51) is verified for $d_0 = 1$ and $C > 2 + \frac{1}{\kappa} + \alpha$. In particular, since $LU \leq -aU + b$, by taking $p_0 < \frac{a-\delta}{C}$, for $0 < \delta < a$, it yields

$$LU(x) + p_0(|F_1(x)| + |F_2(x)|) < -\delta U(x) + b \underset{\|x\|\to\infty}{\longrightarrow} -\infty,$$

which achieves to prove that (50) is verified.

By Theorem 2.30, the convergence rate of $(P_t(x, \cdot))_{t\geq 0}$ towards Π is polynomial. □

Remark 4.11. We can even provide a precise estimate of the speed of convergence: regarding the proof of Proposition 4.1, we have $a = \frac{\alpha n}{2}$ and $c = \varepsilon^2 n^2$ in view of (73), for any $n > 2$. It yields

$$q_0 = 1 + \frac{\alpha}{\varepsilon^2 n},$$

and in particular

$$\min\left\{q_0, \frac{q_0 + 2}{2}\right\} = \begin{cases} q_0 = 1 + \frac{\alpha}{\varepsilon^2 n} & \text{if } \frac{\alpha}{\varepsilon^2} < \frac{n}{2}, \\ \frac{q_0+2}{2} = \frac{3}{2} + \frac{\alpha}{2\varepsilon^2 n} & \text{if } \frac{\alpha}{\varepsilon^2} > \frac{n}{2}, \end{cases}$$

and the speed parameter $\beta > 0$ from (53) follows

$$0 < \beta - 1 < \begin{cases} \frac{\alpha}{\varepsilon^2 n} & \text{if } \frac{\alpha}{\varepsilon^2} < \frac{n}{2}, \\ \frac{1}{2} + \frac{\alpha}{2\varepsilon^2 n} & \text{if } \frac{\alpha}{\varepsilon^2} > \frac{n}{2}. \end{cases}$$

In the first case, for any parameters α, ε in (3), we can take n as large as necessary so that $\frac{\alpha}{\varepsilon^2} < \frac{n}{2}$. In the second case, since $n > 2$, the condition is verified when $\frac{\alpha}{\varepsilon^2} > 1$ or $\alpha > \varepsilon^2$.

4.2.1 Proof of Theorem 4.5

Note that Hypotheses 1-3 are already verified by Proposition 4.1, while the H−persistence condition is also verified by Corollary 4.6.

(i) The strong Hörmander condition holds for (3) on all M_+ by Proposition 4.8, and every points of M_+ are accessible from M_+ by Proposition 4.9. Following Corollary 2.13, for every point $x^* \in M_+$,

$$\Pi_t^x \Rightarrow \Pi \ \mathbb{P}\text{–a.s.} \ .$$

(ii) The first convergence result is a direct consequence of Corollary 2.13***(ii)***, where the additional condition $\int_0^T f(X_s^x)ds$ for all $T > 0$ completes the proof of Proposition 4.8 in [Ben18].

In addition, by Theorem 1.40***(ii)***, even if we lack information about the persistence measure Π, we know that Π admits an exponential moment of order θ, hence its tails decay at least exponentially, since

$$\int_M e^{\theta(x_1+x_2)}\Pi(\mathrm{d}x_1, \mathrm{d}x_2) < \infty.$$

(iii) Also by Corollary 2.13, $(P_t(x, \cdot))_{t\geq 0}$ converges to Π in Total variation.

For the convergence rate, we rely on Proposition 4.10 to show that the conditions of Theorem 2.30 hold. It yields that the convergence rate is polynomial since there exists $x^* \in \Gamma_{\mathbb{R}^2_{++}} \cap \mathbb{R}^2_{++}$ satisfying the strong Hörmander condition.

(iii) Corollary 2.13 also implies that Π is absolutely continuous with respect to the Lebesgue measure on M_+, i.e. Π has a smooth density with respect to the Lebesgue measure on M_+, and is strictly positive. □

Remark 4.12. Unfortunately, we cannot exhibit an exponential convergence rate since the H-persistent *and* at infinity is not verified. We detail our trials to understand what is failing in (3).

Recalling that V, H are constructed as in the proof of Theorem 2.8, which is

$$V(x) = v\left(p_1 h(x_1) + p_2 h(x_2)\right)$$
$$LV(x) = v'\left(p_1 h(x_1) + p_2 h(x_2)\right)(-p_1\lambda_1(x) - p_2\lambda_2(x)) + \frac{1}{2}v''\left(p_1 h(x_1) + p_2 h(x_2)\right)\langle a(x)p, p\rangle_{\mathbb{R}^2},$$

where $h(u) = \log\left(\frac{1}{u}\right)$, $v : \mathbb{R} \to \mathbb{R}_+$ is a C^∞ function with bounded v', v'', and such that $v(t) = t$ for all $t \geq 1$. We let $H = LV$ on M_+, which extends continuously to

$$H(x) = -p_1\lambda_1(x) - p_2\lambda_2(x), \ \text{ on } M_0.$$

For our model (3), we remark that,

$$\lim_{x_1,x_2\to\infty} H(x) = \lim_{x_1,x_2\to\infty} -p_1 F(x_1,x_2) - p_2 F(x_1,x_2) + p_1\frac{\varepsilon^2}{2} \geq \lim_{x_1,x_2\to\infty} -p_1 + \frac{p_1 x_1}{\kappa} + \frac{p_1 x_2}{1+x_1} - p_2\frac{x_1}{1+x_1} > 0,$$

so that (V, H) is not $H-$persistent at infinity.

We consider the following modification: let $v(t) = -t$ for all $t \leq -1$, so that whenever $x_i \to \infty$, $v'\left(p_1 h(x_1) + p_2 h(x_2)\right) = -1$ and

$$LV(x) = p_1\lambda_2(x) + p_1\lambda_2(x), \quad \text{ whenever } x_1, x_2 \to \infty.$$

Based on the previous computations, it implies that

$$\lim_{x_1,x_2\to\infty} H(x) = \lim_{x_1,x_2\to\infty} p_1 F(x_1,x_2) + p_2 F(x_1,x_2) - p_1\frac{\varepsilon^2}{2} \leq \lim_{x_1,x_2\to\infty} p_1 - \frac{p_1 x_1}{\kappa} - \frac{p_1 x_2}{1+x_1} + p_1\frac{x_1}{1+x_1} < 0,$$

so that (V, H) may be $H-$persistent at infinity. However, since $H(x) = -p_1\lambda_1(x) - p_2\lambda_2(x)$, on M_0, if we fix $x_2 = 0$, we obtain

$$\lim_{x_1\to\infty} H(x) = \lim_{x_1\to\infty} -p_1 F(x_1, 0) - p_2 F(x_1, 0) + p_1\frac{\varepsilon^2}{2} \geq \lim_{x_1\to\infty} -p_1 + \frac{p_1 x_1}{\kappa} - p_2\frac{x_1}{1+x_1} > 0,$$

so that (V, H) is not $H-$persistent at infinity.

We modify the construction as follows: let $V^\star = \log(U) + V$ so that

$$H^\star = L(\log U) + LV = L(\log U) + H.$$

Since U is defined on all M, the $H-$persistence condition is still verified since

$$\Lambda^-(H^\star) = \Lambda^-(U) + \Lambda^-(H) = \Lambda^-(H).$$

Now, recalling that

$$L(\log(U)) = \frac{LU}{U} \le -a + \frac{b}{U},$$

we still have $\lim_{x_1, x_2 \to \infty} H(x) < 0$. For the case $x_2 = 0$, we also observe a better bound for $L(\log U)$ from the proof of Proposition 4.1 since

$$LU = nx_1^{n-1}\left(x_1 - \frac{x_1^2}{\kappa}\right) + n\left(\frac{n-1}{2}\right)x_1^2\varepsilon^2 x_1^{n-2} \;\Rightarrow\; \frac{LU}{U} \le n - \frac{nx_1^{n+1}}{\kappa(1+x_1^n)} + \frac{n(n-1)\varepsilon^2}{2},$$

and

$$\begin{aligned}\lim_{x_1\to\infty} H^\star(x) &\le \lim_{x_1\to\infty} n - \frac{nx_1^{n+1}}{\kappa(1+x_1^n)} + \frac{n(n-1)\varepsilon^2}{2} - p_1F(x_1,0) - p_2F(x_1,0) + p_1\frac{\varepsilon^2}{2}\\ &\le \lim_{x_1\to\infty} n - \frac{nx_1^{n+1}}{\kappa(1+x_1^n)} + \frac{n(n-1)\varepsilon^2}{2} + \frac{p_1x_1}{\kappa} + p_2\alpha + p_1\frac{\varepsilon^2}{2}\\ &< 0.\end{aligned}$$

However, if we fix $x_1 = 0$, observe that

$$LU = nx_2^{n-1}(-\alpha x_2) \;\Rightarrow\; \lim_{x_2\to\infty} \frac{LU(x_1,x_2)}{U(x_1,x_2)} = -\alpha n,$$

which implies that

$$\lim_{x_2\to\infty} H^\star(x) = \lim_{x_2\to\infty} -\alpha n + -p_1 + p_1x_2 - \alpha p_2 + p_1\frac{\varepsilon^2}{2} > 0,$$

so that (V, H) is not $H-$persistent at infinity.

As a last trial, if we consider the modification $V^\star = \log(U) + V + x_2$, the control whenever $x_1 = 0$ and $x_2 \to \infty$ or $x_2 = 0$ and $x_1 \to \infty$ holds true but we lack the property outside M_0, that is

$$\lim_{x_1,x_2\to\infty} H^\star(x_1,x_2) < 0.$$

4.3 Extinction of 1 species in the case $0 < \varepsilon^2 < 2$ and $\Lambda(\varepsilon, \alpha, \kappa) < 0$

Then, we are interested in the extinction condition. In particular, if the condition $\Lambda(\varepsilon, \alpha, \kappa) > 0$ is not respected, we can show the almost-sure convergence of $(X_t^x)_{t\ge0}$ towards $M_0^2 := \{x \in \mathbb{R}_+^2 : x_2 = 0\}$, $\forall x \in M_+$. Figure 5 illustrates this typical extinction behavior.

Here is the main result about this subsection.

Theorem 4.13. *If* $0 < \varepsilon^2 < 2$ *and* $\Lambda(\varepsilon, \alpha, \kappa) < 0$*, then* $\forall x \in \mathbb{R}_+^*$*,* $x_2^x(t) \underset{t\to\infty}{\longrightarrow} 0$*,* $\mathbb{P}-$*almost surely.*

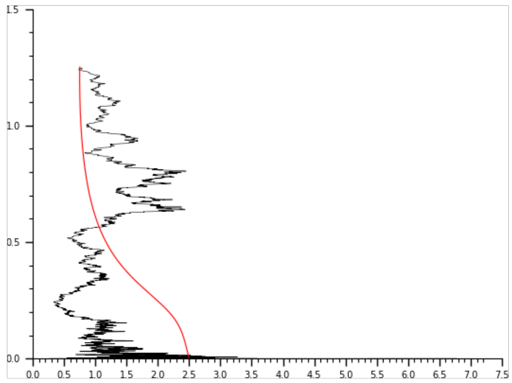

Figure 5: Simulation of (3) starting at $(x_1(0), x_2(0)) = (0.75, 1.25)$ in extinction case with $\Lambda(\varepsilon = 0.6,\ \alpha = 0.9,\ \kappa = 2.5) \approx -0.26 < 0$. The red trajectory is a trajectory of the unperturbed system (i.e. for $\varepsilon = 0$) while the black one describes the system (3) with an Euler–Maruyama scheme.

Proof. We fix $U(x_1, x_2) = 1 + (x_1 + x_2)^n$. Let $V(x_1, x_2) = -\log(x_2)$ be defined on $M_+^{(2)}$. Then, $V \in D_e^{(2)}$ follows Proposition 2.1, V satisfies the strong law by Corollary 1.39 since $\Gamma_e^{(2)}(V) = 0$.

Because the dynamic on x_2 is deterministic, we have

$$\mathcal{L}_e^{(2)} V(x_2) = \frac{V(x_2(t))}{\mathrm{d}t} = -\frac{\dot{x}_2}{x_2} = \alpha - \frac{x_1}{1 + x_1} =: H(x_1, x_2)|_{M_+^{(2)}},$$

is indeed continuous on M. We already showed in the proof of Corollary 4.6 that

$$\mu_1(H) = \mu_1(\lambda_2) = -\Lambda(\varepsilon, \alpha, \kappa) > 0,$$

where $\mu_1(\mathrm{d}x_1 \mathrm{d}x_2) = \gamma_{\varepsilon,\kappa}(x_1)\mathrm{d}x_1 \cdot \delta_0(x_2)\mathrm{d}x_2$ and $\delta_{(0,0)}(H) = \alpha > 0$ which implies

$$\Lambda^+(H) = -\inf\left\{\mu(H) : \mu \in \mathcal{P}_{\text{erg}}(M_0^2)\right\} < 0.$$

We also remark that $\frac{U^{\frac{1}{2}}}{1+|H|}$ is proper since H is bounded while U is proper. Moreover, the condition

$$\lim_{v \to +\infty} \sup_{x \in \mathbb{R}_{++}^2} \left\{\operatorname{dist}(x, M_0^{(2)}) : -\log(x_2) \geq v\right\} = 0,$$

is verified and $\{(X_t^x)_{t\geq 0} : x \in M_+^{(2)}\}$ is H-nonpersistent. It remains to verify Hypothesis 7 and Hypothesis 8; we can proceed via Lemma 2.37.

Since Proposition 2.1 and its consequences hold with $U(x_1, x_2) = 1 + (x_1 + x_2)^n$ for all $n > 2$, let's define $W_N(x_1, x_2) = U_N(x_1, x_2) = 1 + (x_1 + x_2)^N$, $\tilde{W}_N(x_1, x_2) = \tilde{\alpha} W_N(x_1, x_2)$ for $N > 2$ fixed, by Hypothesis 3 *(b)*, and let $G_{2N}(x_1, x_2) = 1 + (x_1 + x_2)^{2N}$. From what precedes, we already know that

$$LG_{2N}(x_1, x_2) = LU_{2N}(x_1, x_2) \underset{\text{by (28)}}{\leq} -\tilde{a} U_{2N}(x_1, x_2) + \tilde{b} =: -\tilde{G}_{2N}(x_1, x_2) + C,$$

with $\tilde{G}_{2N} = \tilde{a}\left(1 + (x_1 + x_2)^{2N}\right)$, $\tilde{b} = C > 0$ a constant.

We can check that $\frac{\tilde{G}}{1+|\mathcal{L}_e W|}$ is proper, because $\tilde{G}$ is of order $(x_1+x_2)^{2N}$ and $\mathcal{L}_e W$ is of order $(x_1+x_2)^{N+1}$, so that

$$\lim_{\|x\| \to +\infty} \frac{\tilde{G}(x_1, x_2)}{1 + |\mathcal{L}_e W(x_1, x_2)|} = +\infty$$

Indeed:

$$\begin{aligned}
|\mathcal{L}_e W(x_1, x_2)| &\leq |x_1 F_1(x) N(x_1 + x_2)^{N-1} + x_2 F_2(x) N(x_1 + x_2)^{N-1}| \\
&\quad + |\frac{N(N-1)}{2} \varepsilon^2 x_1^2 (x_1 + x_2)^{N-2}| \\
&\leq |N x_1 (x_1 + x_2)^{N-1}| + |\frac{x_1^2}{\kappa}(x_1 + x_2)^{N-1}| \\
&\quad + |\alpha x_2 N (x_1 + x_2)^{N-1}| + \left|\frac{N(N-1)}{2} \varepsilon^2 (x_1 + x_2)^N\right| \\
&\leq 2N(x_1 + x_2)^N + \frac{1}{\kappa}(x_1 + x_2)^{N+1} + \frac{N(N-1)}{2}\varepsilon^2 (x_1 + x_2)^N.
\end{aligned}$$

Therefore, for large enough x_1, x_2, we have $|\mathcal{L}eW(x_1, x_2)| \leq C(x_1 + x_2)^{N+1}$, for $C > 0$ a constant, so that $\frac{\tilde{G}}{1+|\mathcal{L}_e W|}$ is proper with $2N > N + 1$ and since

$$\lim_{\|x_1+x_2\|\to\infty} \frac{\tilde{G}_2 N(x_1, x_2)}{1 + |\mathcal{L}_e W(x_1, x_2)|} \geq \lim_{\|x_1+x_2\|\to\infty} C \frac{(x_1 + x_2)^{2N}}{(x_1 + x_2)^{N+1}} = +\infty,$$

Hypothesis 8 is therefore verified.

Then, $\Gamma_e^{(2)}(V) = 0 < \tilde{G}$. Similarly, $W \in C^2(M)$ so that

$$\begin{aligned}
\Gamma_e W(x_1, x_2) &= x_1^2 \varepsilon^2 \cdot N^2 \cdot (x_1 + x_2)^{2N-2} \\
&\leq C \cdot U_{2N}(x_1, x_2) \\
&= C \cdot \tilde{G}_2 N(x_1, x_2),
\end{aligned}$$

so Hypothesis 7 is also satisfied. We conclude by Corollary 2.36 that $X_t^x \underset{t\to\infty}{\longrightarrow} M_0^2$ almost surely. □

Figure 6 illustrates this situation: a particular attention near the origin demonstrates well that $x_2(t) \to 0$ as $t \to \infty$ before $x_1(t)$ goes extinct too.

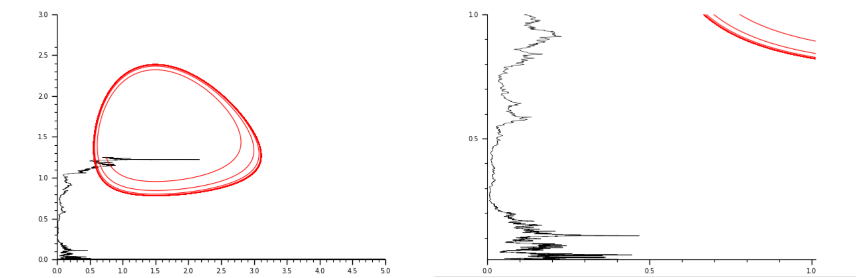

Figure 6: Left figure is a simulation of (3) starting at $(x_1(0), x_2(0)) = (0.75, 1.25)$ in global extinction case with $\Lambda(\varepsilon = 1.35,\ \alpha = 0.6,\ \kappa = 4.5) \approx -0.48 < 0$. The red trajectory is a trajectory of the unperturbed system (i.e. for $\varepsilon = 0$) while the black one describes the system (3) with an Euler–Maruyama scheme. Right figure is a close-up of the situation near the extinction set $x_2 = 0$ while x_1 does not reach 0.

Observe that the case $\alpha \geq 1$, ruled out in the introduction, naturally implies $\Lambda(\varepsilon, \alpha, \kappa) < 0$ and is thus covered by Theorem 4.13.

Remark 4.14. The extinction conclusion can also be recovered by direct computations: remark that

$$\mathrm{d}x_1 = x_1 \left[\left(1 - \frac{x_1(t)}{\kappa} - \frac{x_2(t)}{1 + x_1(t)}\right)\mathrm{d}t + \varepsilon \mathrm{d}B_t\right] \leq x_1 \left[\left(1 - \frac{x_1(t)}{\kappa}\right)\mathrm{d}t + \varepsilon \mathrm{d}B_t\right],$$

so that by a comparison theorem (see e.g. [IW77], Theorem 1.1), $x_1(t) \leq x_t$ where $(x_t)_{t\geq 0}$ is solution to the one-dimensional logistic SDE (65) starting at $x_0 = x_1(0)$ by a time change (see Remark 4.3). On the other hand,

$$x_2(t) = x_2(0)\exp\left(-\alpha t + \int_0^t \frac{x_1(s)}{1+x_1(s)}ds\right). \tag{76}$$

From the one-dimensional logistic SDE (65) and Remark 4.7 note that $\gamma_{\varepsilon,\kappa}$ is the unique invariant probability measure supported on $]0,+\infty[$ and $\Pi_t^x \Rightarrow \gamma_{\varepsilon,\kappa}$ for any $x > 0$ for (65). It follows that

$$\frac{1}{t}\int_0^t \frac{x(s)}{1+x(s)}ds \underset{t\to\infty}{\longrightarrow} \int_0^{+\infty} \frac{x}{1+x}\gamma_{\varepsilon,\kappa}(z)\mathrm{d}z, \quad \text{almost surely.}$$

Since $x_1(t) \leq x_t$ for all $t \geq 0$ almost surely, where x_t is the solution of (65) starting at $x_0 = x_1(0)$ and since $u \mapsto \frac{u}{1+u}$ is an increasing function on $\mathbb{R}_+$, then for all $t > 0$,

$$\frac{1}{t}\int_0^t \frac{x_1(s)}{1+x_1(s)}ds \leq \frac{1}{t}\int_0^t \frac{x(s)}{1+x(s)}ds.$$

Hence,

$$\limsup_{t\to\infty} -\alpha + \frac{1}{t}\int_0^t \frac{x_1(s)}{1+x_1(s)}ds \leq -\alpha + \int_0^{+\infty} \frac{x}{1+x}\gamma_{\varepsilon,\kappa}(x)\mathrm{d}x = \Lambda(\varepsilon,\alpha,\kappa), \quad \text{almost surely,}$$

and from (76), we obtain

$$\limsup_{t\to\infty} \frac{1}{t}\log(x_2(t)) \leq \Lambda(\varepsilon,\alpha,\kappa), \quad \text{almost surely.}$$

Proposition 4.15. *Under the assumption from Theorem 4.13, it yields*

$$\Pi_t^x \Rightarrow \gamma_{\varepsilon,\kappa}(x_1)\mathrm{d}x_1 \otimes \delta_0(x_2)\mathrm{d}x_2, \quad \textit{almost surely,}$$

for all $x \in M_+^{(1)} = \{x \in M : x_1 > 0\}$.

Proof. We will prove that $\{(X_t^x)_{t\geq 0} : x \in M_+^{(1)}\}$ is $H-$persistent. Indeed, Hypotheses 1-3 are verified by Proposition 4.1 and let V be a smooth function such that $V(x_1,x_2) = -\log(x_1)$ for $x_1 \in \left(0,\frac{1}{2}\right)$ and $V(x_1,x_2) = 0$ if $x_1 \geq 1$ so that V is positive, $\lim_{x_1\to 0} V(x) = +\infty$ and

$$LV(x_1,x_2)|_{M_+^{(1)}} = -\left(1 - \frac{x_1}{\kappa} - \frac{x_2}{1+x_1}\right) + \frac{\varepsilon^2}{2} =: H(x_1,x_2)|_{M_+^{(1)}}, \quad \forall 0 < x_1 < \frac{1}{2},$$

which extends continuously to $-1 + x_2 + \frac{\varepsilon^2}{2}$ at $x_1 = 0$. Since $\Gamma V(x_1,x_2)$ is bounded, V satisfies the strong law by Corollary 1.39. Moreover, $\frac{\sqrt{U}}{1+|H|}$ is proper since H is at most polynomial while U is exponential. Finally, since the only invariant probability measure on $M_0^{(1)}$ is $\delta_0(x_1)\otimes\delta_0(x_2)$, we have

$$\mu H = H(0,0) = -1 + \frac{\varepsilon^2}{2} < 0,$$

since $0 < \varepsilon^2 < 2$. By $H-$persistence and Corollary 2.7, it implies that $\Pi(\mathrm{d}x_1,\mathrm{d}x_2) = \gamma_{\varepsilon,\kappa}(x_1)\mathrm{d}x_1 \otimes \delta_0(x_2)\mathrm{d}x_2$ is the unique invariant probability measure supported on $M_+^{(1)}$ and $\forall x \in M_+^{(1)}$, $\Pi_t^x \Rightarrow \Pi$. □

Despite this result, there is no evidence that the law of $x_1(t)$ converges to $\gamma_{\varepsilon,\kappa}$.

Conjecture 4.16. *In the setting of Theorem 4.13,* $\forall x \in \mathbb{R}_+^*$, *the law of* $x_1^x(t)$ *converges weakly to* $\gamma_{\varepsilon,\kappa}(dx)$.

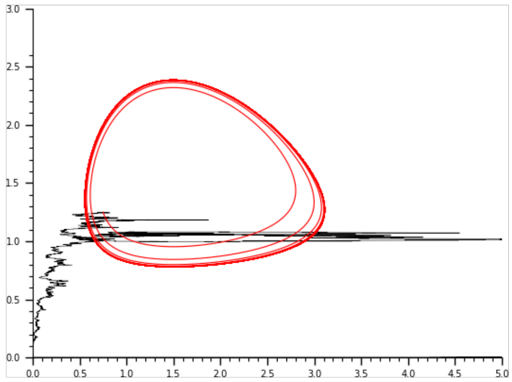

Figure 7: Simulation of (3) starting at $(x_1(0), x_2(0)) = (0.75, 1.25)$ in global extinction case with $\Lambda(\varepsilon = 1.5,\ \alpha = 0.6,\ \kappa = 4.5) \approx -0.79 < 0$. The red trajectory is a trajectory of the unperturbed system (i.e. for $\varepsilon = 0$) while the black one describes the system (3) with an Euler–Maruyama scheme.

4.4 Extinction of both species in the case $\varepsilon^2 > 2$

Finally, we can focus on the case where the condition $\varepsilon^2 < 2$ is not respected: in this situation where the environmental fluctuation is too large, it leads to the extinction of both species as depicted in Figure 7.

Theorem 4.17. *If $\varepsilon^2 > 2$, then $\forall x \in M_+$, $(X_t^x) \underset{t\to\infty}{\longrightarrow} (0,0)$ almost surely.*

Proof. From the one-dimensional logistic SDE (65), we know that $x_t \underset{t\to\infty}{\longrightarrow} 0$ since $\varepsilon^2 > 2$ (see Remark 4.3) so that $x_1(t) \underset{t\to\infty}{\longrightarrow} 0$ almost surely by Remark 4.14. Moreover, by (76), we obtain

$$\limsup_{t\to\infty} \frac{1}{t}\log(x_2(t)) \leq -\alpha + \limsup_{t\to\infty} \frac{1}{t}\int_0^t \frac{x_1(s)}{1+x_1(s)}ds = -\alpha.$$

It concludes to prove $(X_t^x) \underset{t\to\infty}{\longrightarrow} (0,0)$ $\mathbb{P}-$almost surely, $\forall x \in M_+$. The lim sup bounds are direct using above inequality and the logistic one (71) for x_1. □

As a summary of this situation, Figure 8 depicts the different situation we face by evaluating $\Lambda(\varepsilon, \alpha, \kappa)$ with $\alpha = 0.5$ fixed while varying ε and κ.

In the same spirit, Figure 9 depicts the different situation with $\varepsilon = 0.6$ fixed while varying α and κ.

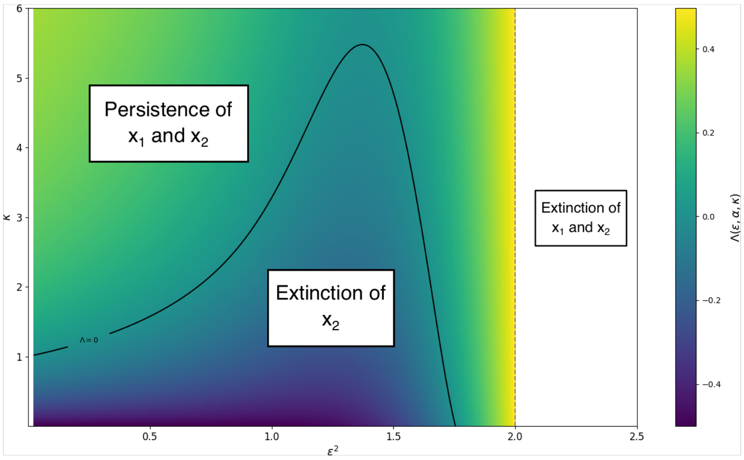

Figure 8: Evaluation of $\Lambda(\varepsilon,\ \alpha,\ \kappa)$ for $\alpha = 0.5$ fixed. The different zones of the graph detailed when we are in a persistence situation (above the $\Lambda = 0$ curve), general extinction (when $\varepsilon^2 > 2$) and extinction of only x_2 (below the $\Lambda = 0$ and when $\varepsilon^2 < 2$).

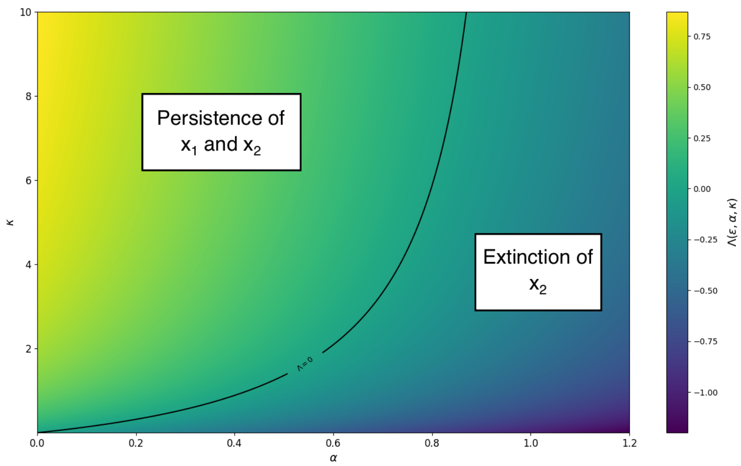

Figure 9: Evaluation of $\Lambda(\varepsilon,\ \alpha,\ \kappa)$ for $\varepsilon = 0.6$ fixed. The different zones of the graph detail when we are in a persistence situation (above the $\Lambda = 0$ curve) and extinction of only x_2 (below the $\Lambda = 0$).

4.5 Appendix - Python codes

4.5.1 Euler–Maruyama simulation

The purpose of the following Python code is to simulate a general two-dimensional system of SDE, with multiplicative noise on x, using a classical Euler–Maruyama scheme. Note that the simulation parameters are fixed with time $T = 700$ with a steps numbers $N = 20000$ to ensure the long-term behavior analysis is relevant while the simulation is robust near the extinction set when $x_1 = 0$ or $x_2 = 0$.

```
import numpy as np
import matplotlib.pyplot as plt
```

```python
import scipy.integrate as integrate
from math import gamma
from scipy.integrate import quad

# Simulation parameters
T = 700.0
N = 200000
\mathrm dt = T / N
t_vals = np.linspace(0, T, N+1)
alpha_test = 0.6
kappa_test = 4.5
epsilon_test = 1.5
epsilon_test_square = epsilon_test**2

# Initial conditions
x0, y0 = 0.75, 1.25
xd0, yd0 = 0.75, 1.25

# Compute the value of Lambda using numerical integration to infinity using scipy.integrate.quad.
def Lambda(epsilon, alpha, kappa):
    k = 2.0 / (epsilon**2) - 1.0
    theta = (epsilon**2 * kappa) / 2.0

    def gamma_pdf(x):
        if x <= 0:
            return 0.0
        return (x**(k - 1.0) * np.exp(-x / theta)) / ((gamma(k) * theta**k))

    def integrand(x):
        return (x / (1.0 + x)) * gamma_pdf(x)

    val, err = quad(integrand, 0, np.inf)
    return val - alpha

lambda_val = Lambda(epsilon_test, alpha_test, kappa_test)
print(f"$\Lambda$($\varepsilon^2$={epsilon_test_square}, $\alpha$={alpha_test}, $\kappa$={kappa_test}) = {lambda_val}")

# Define the determinisitc drift on each variable
def f(x, y):
    return x*(1-x/kappa_test - y/(1+x))

def g(x, y):
    return y*(-alpha_test + x/(1+x))

# Solutions storage - Stochastic case
x_stoch = np.zeros(N+1)
y_stoch = np.zeros(N+1)
x_stoch[0] = x0
y_stoch[0] = y0

# Solutions storage - Deterministic case
x_det = np.zeros(N+1)
y_det = np.zeros(N+1)
x_det[0] = xd0
```

```
y_det[0] = yd0

# Run the Euler-Maruyama scheme and the corresponding deterministic Euler scheme.
for i in range(N):

    # Brownian increments
    dW = np.sqrt(\mathrm dt) * np.random.randn(1)  # -> [N(0,1), N(0,1)]

    x_i, y_i = x_stoch[i], y_stoch[i]
    x_stoch[i+1] = x_i + f(x_i, y_i)*\mathrm dt + epsilon_test*x_i*dW[0]
    y_stoch[i+1] = y_i + g(x_i, y_i)*\mathrm dt

    # Deterministic case
    x_d_i, y_d_i = x_det[i], y_det[i]
    x_det[i+1] = x_d_i + f(x_d_i, y_d_i)*\mathrm dt
    y_det[i+1] = y_d_i + g(x_d_i, y_d_i)*\mathrm dt

# Plot generation
plt.plot(x_stoch, y_stoch, color='black', linewidth=0.4)
plt.plot(x_det, y_det, color='red', linewidth=0.8)

# Personalize axes
ax = plt.gca()
x_min = 0.0
x_max = ax.get_xlim()[1]
y_min = 0.0
y_max = ax.get_ylim()[1]
ax.set_xlim(x_min, x_max)
ax.set_ylim(y_min, y_max)
major_ticks_x = np.arange(x_min, x_max + 0.5, 0.5)
major_ticks_y = np.arange(y_min, y_max + 0.5, 0.5)
ax.set_xticks(major_ticks_x)
ax.set_yticks(major_ticks_y)
minor_ticks_x = np.arange(x_min, x_max + 0.1, 0.1)
minor_ticks_y = np.arange(y_min, y_max + 0.1, 0.1)
ax.set_xticks(minor_ticks_x, minor=True)
ax.set_yticks(minor_ticks_y, minor=True)
ax.tick_params(axis='both', which='major', length=6, width=1.0, labelsize=7)
ax.spines['top'].set_visible(False)
ax.spines['right'].set_visible(False)
ax.tick_params(axis='both', which='minor', length=3, width=0.8, labelbottom=False,
 ↪  labelleft=False)

# Print Lambda value on the figure
text_str = rf"$\Lambda(\epsilon^2={epsilon_test_square}, \alpha={alpha_test},
 ↪  \kappa={kappa_test}) = {lambda_val}$"
plt.figtext(0.5, 0.01, text_str, ha='center', va='bottom')

# Show the plot
print("det:", x_det[N], y_det[N], "et stoch:", x_stoch[N], y_stoch[N])
plt.show()
```

4.5.2 Evaluation of $\Lambda(\varepsilon, \alpha, \kappa)$

In the same spirit, the purpose of the following Python code was to evaluate the value of $\Lambda(\varepsilon, \alpha, \kappa)$ whenever one parameter is fixed and both other ones are evolving with respect to their own limitations. The result is the heatmap depicting the value Λ with respect to the variables that are not fixed, in particular the zones when it goes below or above the critical value $\Lambda = 0$. To speed up the computations of the integral, it uses a Gauss–Laguerre quadrature.

```python
import numpy as np
import mpmath as mp
import matplotlib.pyplot as plt
from numpy.polynomial.laguerre import laggauss

# Parameters
alpha = 0.5
u_min, u_max = 1e-6, 2.5
kap_min, kap_max = 1e-3, 6.0
n_u, n_kap   = 220, 200  # grid resolution
n_lag = 64               # Gauss-Laguerre order
u_label = 0.25

nodes, weights = laggauss(n_lag)

def Lambda_from_u(u, alpha, kappa):
    # Compute Lambda when 0 < u < 2; else NaN
    if not (u > 0 and u < 2 and kappa > 0):
        return np.nan
    k     = 2.0/u - 1.0
    theta = (u * kappa) / 2.0
    if k <= 0 or theta <= 0:
        return np.nan
    y = nodes
    w = weights
    f = np.power(y, k - 1.0) / (1.0 + theta * y)
    Einv = (w @ f) / float(mp.gamma(k))
    return 1.0 - Einv - alpha

u_vals = np.linspace(u_min, u_max, n_u)
kap_vals = np.linspace(kap_min, kap_max, n_kap)
Z = np.empty((n_kap, n_u), dtype=float)

for i, kap in enumerate(kap_vals):
    for j, u in enumerate(u_vals):
        Z[i, j] = Lambda_from_u(u, alpha, kap)

# Meshgrid for contours
U, KAP = np.meshgrid(u_vals, kap_vals)

# Plot design
plt.figure(figsize=(8, 6))
im = plt.imshow(Z, origin='lower', extent=[u_min, u_max, kap_min, kap_max],
 ↪  aspect='auto')
cbar = plt.colorbar(im)
cbar.set_label(r'$\Lambda(\varepsilon,\alpha,\kappa)$', fontsize=14, labelpad=10)

```

```python
# Lambda = 0 curve
CS1 = plt.contour(U, KAP, Z, levels=[0.0], colors='k', linewidths=1.5)

def kappa_on_contour_at_u(contour, u_target):
    for seg in contour.allsegs[0]:
        xs, ys = seg[:,0], seg[:,1]
        if xs.min() <= u_target <= xs.max():
            for i in range(len(xs)-1):
                x0, x1 = xs[i], xs[i+1]
                if (x0 <= u_target <= x1) or (x1 <= u_target <= x0):
                    t = (u_target - x0) / (x1 - x0)
                    return ys[i] + t*(ys[i+1] - ys[i])
    return None

kap_label = kappa_on_contour_at_u(CS1, u_label)
if kap_label is not None and kap_min <= kap_label <= kap_max:
    labels1 = plt.clabel(CS1, fmt={0.0: r'$\Lambda=0$'}, manual=[(u_label, kap_label)], fontsize=9)
    for txt in labels1:
        txt.set_rotation(0)

plt.axvline(2.0, linestyle='--', color='gray')

plt.title(rf'Graph of $\Lambda(\varepsilon,\alpha,\kappa)$ at $\alpha={alpha}$ fixed (x: $\varepsilon^2$)')
plt.xlabel(r'$\varepsilon^2$', fontsize=14)
plt.ylabel(r'$\kappa$', fontsize=14)

# Change the size of the graduation labels
plt.xticks(fontsize=12)
plt.yticks(fontsize=12)

plt.tight_layout()
plt.show()
```

Part B
Uniqueness of the invariant measure for infinite-dimensional and degenerate SDEs

When we are interested in Markov semigroups and their ergodic properties such as existence and uniqueness of the invariant probability measure, the well-known *strong Feller property* is a fundamental tool. However, there are many situations where we lack the strong Feller property: Hairer and Mattingly introduced a weaker version in [HM06], based on their earlier work on asymptotic couplings such as in [Hai02], called the *asymptotic strong Feller property*.

Their objective was to prove the ergodicity of the Navier-Stokes equations with degenerate noise evolving in an infinite-dimensional space. In particular, the asymptotic strong Feller property coupled with a weak form of irreducibility leads to the existence of at most one invariant probability measure.

In this work, our goal was to investigate a specific case where we lack the strong Feller property, especially when the stochastic differential equation involves a degenerate noise in an infinite-dimensional setting.

To this effect, let's start by focusing on the SDE given by

$$X_t = x + \int_0^t g(X_s)\mathrm{d}t - \int_0^t \int_0^s f'(X_s - X_r)drds + \beta_t, \tag{77}$$

where $x \in \mathbb{R}, \beta_t$ a one-dimensional Brownian motion, f and g are 2π-periodic functions, f with sufficient regularity, and g denotes the drift profile. This SDE is defined and studied in [BCG15].

As a concrete and motivating example of (77), we can turn to the field of physics. Indeed, this equation models the shape of growing polymers. One of the first models was introduced in [CD87] and in [NRW87], both in 1987. Then, [DR92] extended the model by studying

$$X_t = B_t + \int_0^t \int_0^s f(X_s - X_u)duds, \tag{78}$$

where $(B_t)_{t\geq 0}$ is a $d-$dimensional Brownian motion, and $f : \mathbb{R}^d \to \mathbb{R}^d$ a Lipschitz function.

Keeping in mind the physical motivation behind (78), the process $(X_t)_{t\geq 0}$ describes a *"growing polymer in which newly added units are repelled by existing ones"*, which is the reason why such a model is called *self-repelling diffusion*. Under relatively mild conditions, they showed in [BG17] the following results while the process evolves on a Riemannian manifold M:

(i) There exists a unique global strong solution;

(ii) strong Feller property holds;

(iii) The system admits a unique invariant measure which is given explicitly as the product of the uniform probability on M and a Gaussian probability on $\mathbb{R}^n$;

(iv) The law of the solution converges to the unique invariant probability exponentially fast.

A natural question then follows: what happens in an infinite-dimensional case?

In [BCG15], the authors studied (77) by choosing a particular form for the initial drift profile g. By taking the Fourier expansion of the function f, the SDE becomes equivalent to a system on a Hilbert space $H = \mathbb{R} \times l^2 \times l^2$. Thus, they aimed to show :

(i) There exists a unique strong solution to (77) having Markov property;

(ii) The Feller property holds;

(iii) There exists an explicit invariant probability measure μ, constructed as an extension of the finite-dimensional one, namely the product of the uniform distribution on the torus with two countably infinite products of normal distributions.

However, since the strong Feller property of the Markov semigroup does not hold in infinite dimensions, the question of the uniqueness of this invariant probability measure is still open.

Structure of this part

Chapter 5 fixes notation and introduces various mathematical tools needed to understand the strong Feller property. It is a requirement to understand how this property is related to unique ergodicity as well as its limitations in many different cases such as in finite- vs. infinite-dimensional state spaces and degenerate vs. nondegenerate noise.

In particular, we also introduce the model of interest in details and we show how the SDE we are focusing on lacks the strong Feller property.

Chapter 6 introduces a new mathematical background to define the *asymptotic strong Feller property*. Mirroring most of the results and ideas of Chapter 1, we explain how this property is weaker than the strong Feller one, their relations and differences as well as the consequences of the asymptotic strong Feller property, combined with a weak form of accessibility, in terms of unique ergodicity.

Chapter 7 details how we can achieve (asymptotic) strong Feller property through (asymptotic) coupling theory, by introducing a *log–Harnack inequality* in the strong Feller case, and a *modified* version for the asymptotic strong Feller property. Hypothesis 14 is introduced as a sufficient condition to prove that a modified log–Harnack inequality is verified, hence the asymptotic strong Feller property holds.

Since coupling is directly related to the notion of change of measure, we also fix some notations and recall some theoretical, fundamental results as Girsanov's theorem and Kolmogorov extension theorem.

Chapter 8 gives a broad range of examples, that may also have been introduced along previous chapters. A particular focus is given on a toy model introduced by Hairer in [Hai02] that is at the basis of the asymptotic coupling notion, even if the underlying Markov semigroup has the strong Feller property. Other examples show how the unique ergodicity may be achieved when we lack the strong Feller property, using the modified log–Harnack inequality.

Chapter 9 introduces new kinds of examples for which we successfully applied previous tools such as the modified log–Harnack inequality. We also detail how we try to apply all this work to our model of interest, without success. It also proposes some ideas to continue studying the unique ergodicity of this model.

Chapter 5

Introduction

We fix a similar setting as in Part A that we quickly recall for consistency.

Let (M, d) be a locally compact Polish (complete, separable, metrizable) space. We denote by $\mathcal{M}$ the σ-algebra on M, $C(M)$ (respectively $C_b(M)$) for the set of continuous (respectively bounded continuous) functions on M, and $\mathcal{B}(M)$ (respectively $\mathcal{B}_b(M)$) for the set of measurable (respectively bounded measurable) functions.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space together with a complete, right-continuous filtration $(\mathcal{F}_t)_{t\geq 0}$. Let $\{(X_t^x)_{t\geq 0} : x \in M\}$ be a family of càdlàg Markov processes, and let $(P_t)_{t\geq 0}$ be the associated *Markov semigroup* in the sense that, for each $f \in \mathcal{B}(M)$,

$$(t, x) \in \mathbb{R}_+ \times M \to P_t f(x) := \mathbb{E}\left[f(X_t^x)\right],$$

is a measurable mapping and the Markov property holds, i.e.

$$\mathbb{E}\left[f(X_{t+s}^x)|\mathcal{F}_t\right] = P_s f(X_t^x) \quad \mathbb{P} - a.s.$$

Let $\mathcal{P}(M)$ be the set of probability measures on $(M, \mathcal{M})$: $\mu \in \mathcal{P}(M)$ is called an *invariant probability measure* if

$$\mu P_t = \mu, \quad \forall t \geq 0,$$

and let $\mathcal{P}_{\text{inv}}(M)$ be the set of invariant probability measures. We say that $(P_t)_{t\geq 0}$ is *uniquely ergodic* if $\mathcal{P}_{\text{inv}}(M)$ has cardinality one.

For $\mu \in \mathcal{P}(M)$, we let $\mathbb{P}_\mu$ be the law of the process $(X_t)_{t\geq 0}$ on the Skorokhod space $\mathbb{D}(\mathbb{R}_+, M)$ of càdlàg functions, under the initial condition μ, and $\mathbb{E}_\mu$ for its associated expected value. In particular, for the Dirac measure at one point δ_x, we write more commonly $\mathbb{P}_x$ and $\mathbb{E}_x$. Similarly, we denote by X^μ or X^x the process under the initial condition μ or x.

A function $f \in \mathcal{B}(M)$ is *(G, μ)–invariant* if $Gf = f$ μ–a.s., where G denotes the *1–resolvent kernel* (a discrete–time Markov kernel) defined by

$$Gf := \int_0^\infty e^{-t} P_t f \mathrm{d}t, \quad \forall f \in \mathcal{B}(M). \tag{79}$$

Likewise, a set $B \in \mathcal{M}$ is (G, μ)–invariant if $\mathbf{1}_B$ is (G, μ)–invariant.

An invariant probability $\mu \in \mathcal{P}_{\text{inv}}(M)$ is called *ergodic* if every (G, μ)−invariant map is μ−almost surely constant. We denote the set of ergodic measures by $\mathcal{P}_{\text{erg}}(M)$.

Let $\mathcal{L}$ be the *infinitesimal generator* defined on $C_b(M)$ with $\mathcal{D}(\mathcal{L}) \subset C_b(M)$ its *domain*, for which

- **(i)** $\mathcal{L}f(x) := \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t}$ exists for all $x \in M$;
- **(ii)** $\mathcal{L}f \in C_b(M)$;
- **(iii)** $\sup\limits_{0<t\leq 1} \frac{1}{t}\|P_t f - f\|_\infty < \infty$.

5.1 General theory about (strong) Feller

The property of being (strong) Feller for a Markov operator leads to interesting results to show that there is a unique invariant probability associated to the Markov operator. In particular, the strong Feller property of a Markov semigroup $(P_t)_{t\geq 0}$, combined with some irreducibility of $(P_t)_{t\geq 0}$, implies the uniqueness of the invariant measure with respect to $(P_t)_{t\geq 0}$.

Definition 5.1 (**(Strong) Feller property**). *A Markov operator over M associated to the Markov semigroup $(P_t)_{t\geq 0}$ has the $C_b(M)-$Feller property (respectively strong Feller property) if $\forall \varphi \in C_b(M)$ (respectively $\mathcal{B}_b(M)$), $P_t\varphi \in C_b(M)$, $\forall t \geq 0$.*

Remark 5.2. **(i)** As stated by Hairer in [Hai21], a Markov operator P is $C_b(M)-$Feller if and only if the map $x \mapsto P(x, \cdot)$ is continuous in the topology of weak convergence, i.e.

$$\mu_n \underset{\text{weak}}{\Rightarrow} \mu \ \Leftrightarrow\ \mu_n f \underset{n\to\infty}{\to} \mu f,\ \forall f \in C_b(M).$$

(ii) Similarly, as stated in [Hai09], a Markov operator P is strong Feller iff the map $x \mapsto P(x, \cdot)$ is continuous in the topology of strong convergence, i.e.

$$\mu_n \underset{\text{strong}}{\Rightarrow} \mu \ \Leftrightarrow\ \mu_n f \underset{n\to\infty}{\to} \mu f,\ \forall f \in \mathcal{B}_b(M).$$

Example 5.3. We consider the following homogeneous Stochastic Differential Equation on $\mathbb{R}^d$,

$$dX_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t,\ X_0 = x, \tag{80}$$

where $(B_t)_{t\geq 0}$ is a $d-$dimensional Brownian motion. In addition, we suppose that:

(i) $a := \sigma \cdot \sigma^T$ is a continuous function;

(ii) $a(x)$ is positive definite for all $x \in \mathbb{R}^d$;

(iii) and $\exists k > 0$ such that $\forall 1 \leq i, j \leq d$ and $\forall x \in \mathbb{R}^d$,

$$|a^{ij}(x)| \leq k(1 + \|x\|^2), \quad |b^i(x)| \leq k(1 + \|x\|),$$

where a^{ij} is the $ij-$coordinate of the matrix ***(i)*** and b^i is the $i-$th coordinate of ***(ii)***.

Under these conditions, (80) is considered as a general nondegenerate $d-$dimensional diffusion process with standard boundary conditions. The following theorem is a well-known result stated by Stroock and Varadhan.

Theorem 5.4 ([SV71], Theorem 2.3)**.** *Under the assumption **(i)-(iii)**, there exists a unique solution $(X_t)_{t\geq 0}$ to (80) which is a Markov process with a semigroup $(P_t)_{t\geq 0}$ defined as*

$$P_t f(x) = \mathbb{E}_{\mathbb{Q}}\left[f(X_t) \mid X_0 = x\right], \quad \forall f \in \mathcal{B}_b(\mathbb{R}^d).$$

Moreover, $(P_t)_{t\geq 0}$ is a strong Feller Markov semigroup in the sense that $\forall f \in \mathcal{B}_b(\mathbb{R}^d)$, $P_t f \in C_b(\mathbb{R}^d)$ where $P_t f(x) = \int_{\mathbb{R}^d} P_t(x, dy) f(y)$.

Obviously, the strong Feller property is not as direct to prove unless specific examples as in (80). Many different techniques have been widely investigated depending on the type of models considered.

5.1.1 The finite-dimensional case - Hörmander condition

The Hörmander condition, formulated in [Hö85], is a great tool to achieve strong Feller property when $(P_t)_{t\geq 0}$ is generated by a diffusion with smooth coefficients on $\mathbb{R}^d$ or on a finite-dimensional manifold. The assumption is the following: by taking the successive Lie brackets of the vector fields, we suppose that they span the entire space, at each point. Then: the law of the Markov process possesses a density p which is C^∞ with respect to the Lebesgue measure. In particular, we have

$$P_t f(x) = \mathbb{E}\left(f(X_t^x)\right) = \int_{\mathbb{R}^d} p(t,x,dy)f(y), \quad \forall x \in \mathbb{R}^d,\ \forall t \geq 0,$$

and by dominated convergence,

$$x_n \underset{n\to\infty}{\longrightarrow} x \ \Rightarrow\ P_t f(x_n) \underset{n\to\infty}{\longrightarrow} P_t f(x), \quad \forall f \in \mathcal{B}_b(\mathbb{R}^d),\ \forall x \in \mathbb{R}^d, \text{ and } \forall t \geq 0,$$

which implies that $(P_t)_{t\geq 0}$ is strong Feller by Remark 5.2.

More precisely, following the presentation of Hörmander condition in [Gau16]: let $\mathcal{X}$ be a smooth, connected manifold of dimension d: one can think of $\mathcal{X} = \mathbb{R}^d$ or $\mathcal{X} = \mathbb{S}^d$. Let $S^0, \cdots, S^n$ be smooth vector fields taking values in $\mathcal{X}$, and let $(X_t^x)_{t\geq 0}$ be the solution to the SDE (under Stratonovich formalism)

$$dX_t = S^0(X_t)\mathrm{d}t + \sum_{k=1}^n S^k(X_t) \circ dB_t^{(k)}, \quad X_0 = x, \tag{81}$$

where $(B_t^{(1)}, \cdots, B_t^{(n)})_{t\geq 0}$ is a standard Brownian motion on $\mathbb{R}^n$ and $\circ$ denotes the Stratonovich integral. Under the same previous notations, given the SDE defined on $\mathbb{R}^d$ under the Itô's form

$$dX_t = f(X_t)\mathrm{d}t + \sum_{j=1}^n \sigma_j(X_t) dB_t^j,$$

we can recover its Stratonovich form using the relation

$$S_i^j(x) = \sigma_{ij}(x), \quad \text{and} \quad S_i^0 = f_i(x) - \frac{1}{2}\sum_{j=1}^n \sum_{k=1}^d \frac{\partial S_i^j(x)}{\partial x_k} S_k^j(x),$$

for all $j = 1, \cdots, n$ and $i = 1, \cdots, d$. Under the Stratonovich form, it is well-known that (81) has the infinitesimal generator $\mathcal{L}$ defined by

$$\mathcal{L} = \frac{1}{2}\sum_{k=1}^n S_k^2 + S_0.$$

Let $(P_t)_{t\geq 0}$ be the Markov semigroup induced by $\mathcal{L}$, which is defined in particular for bounded, continuous function by

$$P_t f(x) = \mathbb{E}\left[f(X_t^x)\right], \quad \forall f \in \mathcal{B}_b(\mathcal{X}),\ \forall x \in \mathcal{X},$$

and let $P_t(x, dy)$ be the law of $(X_t^x)_{t\geq 0}$ at time t.

We also recall that given Y, Z be two smooth vector fields on $\mathcal{X}$, the *Lie brackets* of Y and Z is the vector field on $\mathcal{X}$ defined as

$$[Y,Z](f) = Y(Z(f)) - Z(Y(f)), \quad \forall f \in C^\infty(\mathcal{X}).$$

In particular, when $\mathcal{X} = \mathbb{R}^d$, one has

$$[Y,Z](x) = DZ(x)Y(x) - DY(x)Z(x),$$

where D denotes the derivative at x. Now, let $\mathcal{S}_0 = \{S^0, \cdots S^n\}$ and define recursively

$$\mathcal{S}_{k+1} = \mathcal{S}_k \cup \{[Y, Z] : Y, Z \in \mathcal{S}_k\}, \quad \forall k \geq 0.$$

We set $\mathcal{S} = \cup_{k\geq 0}\mathcal{S}_K$ and $\forall x \in \mathcal{X}$,

$$\mathcal{S}(x) = \{Y(x) : Y \in \mathcal{S}\}.$$

Definition 5.5 (**Hörmander condition**). *We say that (81) satisfies the* Hörmander condition *if for all $x \in \mathcal{X}$, $\mathcal{S}$ spans $T_x\mathcal{X}$, which is the tangent plane at x of $\mathcal{X}$.*

We can interpret the Hörmander condition as a smoothness property in the sense that the noise term is sufficiently well spread in all the direction of the underlying process $(X_t^x)_{t\geq 0}$. It has the following well-known consequence:

Theorem 5.6 ([IK74], Theorem 3 and Lemma 5.1)**.** *Assume that the Hörmander condition holds for (81). Then, there exists a $C^\infty((0, \infty) \times \mathcal{X} \times \mathcal{X})$ function $p_t(x, y)$ such that*

(i) $P_t(x, dy) = p_t(x, y)dy$;

(ii) $\mathcal{L}p_t(\cdot, y) = \partial_t p_t(\cdot, y)$;

(iii) $\mathcal{L}^* p_t(x, \cdot) = \partial_t p_t(x, \cdot)$, *where $\mathcal{L}^*$ denotes the adjoint operator of $\mathcal{L}$ in $L^2(\mathcal{X}, dx)$;*

(iv) $(P_t)_{t\geq 0}$ *has the strong Feller property.*

We can emphasize that this result is true only when we consider finite-dimensional state space. We can still mention that, in infinite dimensionsal setting with only a finite vector fields modeled the noise term, similar results follow the work of Hairer and Mattingly in [HM11], following the ideas of [Hai11] which gives a proof of Hörmander's theorem using Malliavin calculus.

5.1.2 The infinite dimensionsal case - Bismut-Elworthy-Li formula

This method cannot be applied in infinite dimensions, where we lack a reference measure such as the Lebesgue measure. Moreover, in case of nondegenerate noise's covariance, one can rely on the Bismut-Elworthy-Li formula introduced in [EL94]. Using the settings of [PZ96] and [PZ92], we consider the SDE evolving on a separable Hilbert space $\mathbb{H}$ given by

$$dX_t = (AX_t + F(X_t))\,\mathrm{d}t + BdW_t, \quad X_0 = x, \tag{82}$$

where F, A are Lipschitz continuous, $(W_t)_{t\geq 0}$ is a cylindrical Wiener process on $\mathbb{H}$, and A is the infinitesimal generator of a strongly continuous semigroup $S(t)$, $t \geq 0$, on $\mathbb{H}$, in the sense that $S(0) = Id$, $S(t+s) = S(t)S(s)$, $\forall t, s \geq 0$, and $\lim_{t\downarrow 0} \|S(t)x - x\|_{\mathbb{H}} = 0$, $\forall x \in \mathbb{H}$.

Recall that if $(e_n)_{n\geq 1}$ is a Hilbert basis associated to $\mathbb{H}$, an operator $T : \mathbb{H} \to \mathbb{H}$ is said to be Hilbert–Schmidt if it has a finite square Hilbert–Schmidt norm, which is

$$\|T\|_{HS}^2 := \sum_{n\geq 1} \langle T^* T e_n, e_n\rangle_{\mathbb{H}} < \infty.$$

In addition, we assume the following:

Hypothesis 9. **(i)** $\forall z \in \mathbb{H}$, $B(z)$ is invertible and $\exists K > 0$ such that

$$\|B^{-1}(z)\|_{\mathbb{H}} \leq K.$$

(ii) For all $t \geq 0$, $S(t)$ is a Hilbert–Schmidt operator.

(iii) $\int_0^1 \|S(t)\|_{HS}^2 \mathrm{d}t < \infty$.

The following result follows from Theorem 9.32 and Lemmas 9.33-9.36 from [PZ92]:

Theorem 5.7. *Under Hypothesis 9, for any $\varphi \in \mathcal{B}_b(\mathbb{H})$, we obtain a Bismut-Elworthy-Li-like formula so that the directional derivatives of $\langle D_x P_t \varphi(x), h\rangle$ are given by*

$$\langle D_x P_t \varphi(x), h\rangle = \frac{1}{t}\mathbb{E}\left(\varphi(X_t^x)\int_0^t \langle B^{-1}(X_s^x)X_s^x, dW_s\rangle\right), \quad \mathbb{P}-a.s.$$

Moreover, for any $T > 0$ be fixed, there exists $C_T > 0$ such that $\forall \varphi \in \mathcal{B}_b(\mathbb{H})$, $\forall t \in [0, T]$, we have

$$|P_t\varphi(x) - P_t\varphi(y)| \leq \frac{C_T}{\sqrt{t}}\|\varphi\|_0\|x-y\|_{\mathbb{H}}, \quad \forall x, y \in \mathbb{H},$$

where $\|\cdot\|_0$ is the induced norm on $\mathbb{H}_0 := \mathrm{Range}\left(B^{\frac{1}{2}}\right)$ given by $\|h\|_0 = \left\|B^{-\frac{1}{2}}(h)\right\|$, for $h \in \mathbb{H}_0$.

In particular, $(P_t)_{t\geq 0}$ is strong Feller for all $t > 0$.

5.2 Links between strong Feller and unique ergodicity

The interest of the strong Feller property is directly related to the unique ergodicity of $(P_t)_{t\geq 0}$ when it is associated to some irreducibility conditions.

As stated by Da Prato and Zabczyk in [PZ96], we recall the notions of irreducibility and regularity for a Markov process, and the link with the strong Feller property.

Definition 5.8 (Irreducibility). *A Markov semigroup $(P_t)_{t\geq 0}$ is said to be irreducible at time $t_0 > 0$ if, for all open set $\Gamma \neq \emptyset$ of M, we have*

$$P_{t_0}(x, \Gamma) > 0, \quad \forall x \in M.$$

Definition 5.9 (Regularity). *A Markov semigroup $(P_t)_{t\geq 0}$, is said to be t_0-regular if all transition probabilities $P_{t_0}(x, \cdot)$, $\forall x \in M$, are mutually equivalent.*

Remark 5.10. **(i)** By the semigroup property, if a semigroup $(P_t)_{t\geq 0}$ is regular at time $t_0 > 0$, it is $s-$regular for arbitrary $s > t_0$, and all transition probability measures $P_s(x, \cdot)$ are equivalent, for all $s > t_0$, $x \in M$.

(ii) If μ is an invariant probability measure for a t_0-regular semigroup $(P_t)_{t\geq 0}$, then all transition probability measures $P_t(x, \cdot)$, for all $t > t_0$ and $x \in M$, are equivalent.

Proposition 5.11 ([PZ96], Proposition 4.1.1)**.** *If a Markov semigroup $(P_t)_{t\geq 0}$ is strong Feller at time $t_0 > 0$ and irreducible at time $s_0 > 0$, then it is regular at time $t_0 + s_0$.*

Finally, one of the main result about regular Markov semigroups is due to J.L. Doob in [Doo48], known as *Doob's Theorem* and is formulated as follows by Da Prato and Zabczyk.

Theorem 5.12 ([PZ96], Theorem 4.2.1)**.** *Let $(P_t)_{t\geq 0}$ be a stochastically continuous Markov semigroup, and μ be an invariant measure with respect to $(P_t)_{t\geq 0}$. If $(P_t)_{t\geq 0}$ is t_0-regular for some $t_0 \geq 0$, then*

(i) *μ is strongly mixing* [1]*, and for an arbitrary $x \in M$, and $\Gamma \in \mathcal{M}$,*

$$\lim_{t\to\infty} P_t(x,\Gamma) = \mu(\Gamma);$$

(ii) *μ is the unique invariant probability measure for the semigroup $(P_t)_{t\geq 0}$, $t \geq 0$;*

(iii) *μ is equivalent to all probability measures $P_t(x,\cdot)$, for all $t \geq t_0$ and $x \in M$.*

Remark 5.13. The notion of *stochastically continuous Markov semigroup* refers to Markov semigroups $(P_t)_{t\geq 0}$ such that

$$\lim_{t\to 0} P_t(x, B(x,\delta)) = 1 \; \forall x \in M, \; \forall \delta > 0,$$

where $B(x,\delta)$ is the open ball with radius δ and centred at x. Proposition 2.1.1 from [PZ96] gives equivalent conditions for a Markov semigroup to be stochastically continuous.

Another perspective is given by accessibility conditions such as in [BH22]. The motivation is that the condition of irreducibility is not easy to verify most of the time, so that we need to introduce this purely topological notion.

Recall that *topological support* of a measure μ, denoted $\mathrm{supp}(\mu)$, is the closed set defined as the intersection of all closed sets $F \subset M$ such that $\mu(M\backslash F) = 0$. It enjoys the following properties on any separable metric space, in particular Polish space as considered here:

(i) $\mu(M\backslash\mathrm{supp}(\mu)) = 0$,

(ii) $x \in \mathrm{supp}(\mu)$ if and only if $\mu(O) > 0$ for every open set O containing x.

Definition 5.14 (Accessibility). *For a Markov semigroup $(P_t)_{t\geq 0}$, a point $y \in M$ is* accessible *from $x \in M$ if, for every neighborhood U of y, there exists $t \geq 0$ with $P_t(x,U) > 0$. We denote by Γ_x the set of points accessible from x.*

For $C \subset M$, let $\Gamma_C = \cap_{x\in C}\Gamma_x$ the set of points that are accessible from C, and $\Gamma := \Gamma_M$ the set of accessible points.

Let G be the 1–resolvent kernel associated to $(P_t)_{t\geq 0}$ and defined in (79). Accessibility is also recoverable through G (see e.g Section 5.2.1 and Proposition 5.19 in [BH22]) as

$$\Gamma_x = \mathrm{supp}(G(x,\cdot)).$$

Remark 5.15. Γ_C is a closed but possibly empty set: we say that $(P_t)_{t\geq 0}$ is (topologically) *indecomposable* if $\Gamma \neq \emptyset$.

Definition 5.16 (ξ-irreducibility). *Let $(P_t)_{t\geq 0}$ be a Markov semigroup on M and ξ a nonzero measure on M. We say that $(P_t)_{t\geq 0}$ is ξ-*irreducible *if for every $A \subset \mathcal{M}$ and every $x \in M$,*

$$\xi(A) > 0 \;\Rightarrow\; \exists t > 0 \text{ such that } P_t(x,A) > 0.$$

Equivalently,

$$\xi(A) > 0 \;\Rightarrow\; G(x,A) > 0,$$

where G is the $1-$resolvent kernel.

[1] See [Kea72]: $\mu \in \mathcal{P}_{\mathrm{inv}}(M)$ is strongly mixing $\Leftrightarrow \forall f,g \in C(M)$ the set of continuous real-valued functions on M, $\mu(f \circ P^k g) \underset{k\to\infty}{\to} \mu(f)\mu(g)$.

In the continuous-time setting, we apply the following discrete-time results from [BH22] to any skeleton of the semigroup $(P_t)_{t\geq 0}$, that is, to the Markov operator P_{t_0} for a fixed $t_0 > 0$. All notions of irreducibility, indecomposability, and invariant measures are equivalent for the semigroup and for any of its skeletons (see e.g. [MT09, Theorem 6.2.3] or [PZ96, Chapter 4]) and the continuous-time setting can also treated via the resolvent kernel G, which plays an equivalent role.

Proposition 5.17 ([BH22], Proposition 5.8)**.** *Let $(P_t)_{t\geq 0}$ be a $C_b(M)$−Feller semigroup and fix $t_0 > 0$. Assume P_{t_0} is topologically indecomposable, and let Γ denote its accessible set. Then:*

(i) $P_{t_0}(x, \Gamma) = 1$ *for all $x \in \Gamma$;*

(ii) $\Gamma \subset \mathrm{supp}(\mu)$ *for every $\mu \in \mathcal{P}_{\mathrm{inv}}(M)$;*

(iii) *If Γ has nonempty interior, then* $\mathrm{supp}(\mu) = \Gamma$ *for all $\mu \in \mathcal{P}_{\mathrm{inv}}(M)$;*

(iv) *If Γ is compact, there exists $\mu \in \mathcal{P}_{\mathrm{inv}}(M)$ with* $\mathrm{supp}(\mu) = \Gamma$*;*

(v) *If Γ is compact and $g : \Gamma \to \mathbb{R}$ is continuous and P_{t_0}-harmonic ($P_{t_0} g = g$ on Γ), then g is constant.*

Proposition 5.18 ([BH22], Proposition 5.13)**.** *Assume M is compact. Fix $t_0 > 0$. If P_{t_0} is $C_b(M)$-Feller, Γ (for P_{t_0}) has nonempty interior, and for every bounded Lipschitz f the family $\{P^n_{t_0} f\}_{n\geq 1}$ is equicontinuous, then $(P_t)_{t\geq 0}$ is uniquely ergodic.*

Recall that a function $f : M \to \mathbb{R}$ is said to be *lower semicontinuous* (respectively *upper semicontinuous*) at a point $x_0 \in M$ if

$$f(x_0) \leq \liminf_{x\to x_0} f(x) \quad \left(\text{respectively } f(x_0) \geq \limsup_{x\to x_0} f(x)\right).$$

We can see that f is continuous at $x_0 \in M$ if and only if it is both lower and upper semicontinuous at x_0.

Proposition 5.19 ([BH22], Proposition 5.17)**.** *Suppose for some $t_0 > 0$ that P_{t_0} is topologically indecomposable. If there exists $x^* \in \Gamma$ such that for every $A \in \mathcal{M}$ the map $x \mapsto P_{t_0}(x, A)$ is lower semicontinuous at x^*, then $(P_t)_{t\geq 0}$ is ξ-irreducible with $\xi(\cdot) = P_{t_0}(x^*, \cdot)$. In particular, $(P_t)_{t\geq 0}$ admits at most one invariant probability measure.*

The assumption $x \mapsto P_t(x, A)$ is lower semi-continuous at x^*, $\forall t \geq 0$, is automatically satisfied if $(P_t)_{t\geq 0}$ is strong Feller since for all $A \in \mathcal{M}$, $x \mapsto \mathbf{1}_A(x)$ is a bounded, measurable function, which implies that $x \mapsto P_t \mathbf{1}_A(x) = P_t(x, A)$ is a bounded, continuous function, hence lower semicontinuous. It follows that:

Proposition 5.20 ([BH22], Proposition 5.18)**.** *If $(P_t)_{t\geq 0}$ is strong Feller, then:*

(i) *Two distinct ergodic invariant probabilities have disjoint supports;*

(ii) *The support of any non-ergodic invariant probability is disconnected;*

(iii) *If M is connected and $(P_t)_{t\geq 0}$ has an invariant probability with full support, then $(P_t)_{t\geq 0}$ is uniquely ergodic.*

5.3 A general criterion to achieve strong Feller property

A standard criterion explicited by Hairer and Mattingly in [HM06] is a direct condition to ensure strong Feller property of the Markov semigroup $(P_t)_{t\geq 0}$. Its interest lies in the fact that it will be directly related to the asymptotic strong Feller property: before stating the criterion, we need an introduction to the notion of pointwise Lipschitz constant:

5.3.1 Pointwise Lipschitz constant

Our purpose is to weaker condition of the Stepanov Differentiability Theorem, which is also an extension of the Rademacher Theorem. As in [BRZ04], the statement from Rademacher that a Lipschitz function on $\mathbb{R}^n$ is differentiable almost everywhere is extended by Stepanov where the assumption on Lipchitz continuity is replaced by the following condition: a function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at almost every $x \in S(f)$, where

$$S(f) = \left\{ x \in \mathbb{R}^n \ : \ \lim_{y \to x} \frac{|f(y) - f(x)|}{|y - x|} < \infty \right\}.$$

The natural next generalization is given by extending this methodology to any metric space. As stated in [DCJ10], [BH99], and [Lan13],

Definition 5.21 (**Pointwise Lipschitz constant**). *Let (E, d) be a metric space: for a function $f : E \to \mathbb{R}$, and $x \in E$, we define*

$$|\nabla f|(x) = \limsup_{y \to x} \frac{|f(x) - f(y)|}{d(x, y)}.$$

We called $|\nabla f|(x)$ the pointwise Lipschitz constant *of f at x.*

We also note $\|\nabla f\|_\infty = \sup_{x \in E} |\nabla f|(x)$, and we define

$$D(E) = \{ f : E \to \mathbb{R} \mid \|\nabla f\|_\infty < \infty \}.$$

We say that $f \in D(E)$ is a pointwise Lipschitz function.

If f is a (even locally) Lipschitz function, it is clear that $|\nabla f|(x)$ is well-defined through its (local) Lipschitz constant. This notion is thus a generalization of (local) Lipschitz property.

Remark 5.22. In the literature, we also find the notion $\mathrm{Lip} f(x) = |\nabla f|(x)$. In fact, the use of the symbol ∇ is an abuse of language: let $f \in C^1(\Omega)$ where Ω is an open subset from Euclidean space or a Riemannian manifold. Then, $\mathrm{Lip}\, f = \|\nabla f\|$ where ∇ is the classical gradient operator and $\|\cdot\|$ the norm associated with the inherent Euclidean space or Riemannian manifold.

In fact, this is clear that $D(E)$ contains the space of every Lipschitz functions $\mathrm{Lip}(E)$, but there is no evidence to the contrary.

Before diving into conditions for the equivalence and a counterexample, we need to define some tools. Recall that a *curve* γ on (E, d) is a continuous map $\gamma : [a, b] \to E$, its *image* is denoted $|\gamma| = \gamma([a, b])$ and its *length* is

$$l(\gamma) = \sup \left\{ \sum_{i=0}^{n-1} d\left(\gamma(t_i), \gamma(t_{i+1})\right) \right\},$$

where the sup is taken over all partitions $a = t_0 < t_1 < \cdots < t_n = b$ of the interval $[a, b]$.

Definition 5.23 (**Rectifiability**). *A curve γ is* rectifiable *if $l(\gamma) < \infty$.*

Definition 5.24 (**Length space**). *A metric space (E, d) is called a* length space *if for each pair of points $x, y \in E$, the distance $d(x, y)$ coincides with the infimum of all lengths of curves in E connecting x and y.*

As an interesting class of metric spaces that contains length spaces, we introduce the notion of *quasi-convex space*:

Definition 5.25 (**Quasi-convex space**)**.** *A metric space (E, d) is called a* quasi-convex space *if there exists a constant $C > 0$ such that for each pair of points $x, y \in E$, there exists a curve γ connecting x and y, which is rectifiable and such that*

$$l(\gamma) \leq C d(x, y).$$

Remark 5.26. Every normed vector space $(E, \|\cdot\|_{\mathbb{E}})$ is a length space, hence quasi-convex with constant $C = 1$. Indeed, for any $x, y \in E$, the straight-line path $\gamma : [0, 1] \to E : t \mapsto x + t(y - x)$ satisfies $l(\gamma) = \|y - x\|_{\mathbb{E}} = d(x, y)$.

In particular, this applies to all (pre-)Hilbert spaces and Banach spaces.

Lemma 5.27 ([DCJ10], Lemma 2.3)**.** *Let (E, d) be a metric space and $f \in D(E)$. Let $x, y \in E$ and suppose that there exists a rectifiable curve $\gamma : [a, b] \to E$ connecting x and y, which means that $\gamma(a) = x$ and $\gamma(b) = y$. Then,*

$$|f(x) - f(y)| \leq \|\nabla f\|_\infty \cdot l(\gamma). \tag{83}$$

As a consequence, the property of quasi-convex space as in Definition 5.25 implies that (83) rewrites as

$$|f(x) - f(y)| \leq C\|\nabla f\|_\infty \cdot d(x, y), \quad \forall x, y \in E, \tag{84}$$

for any function $f \in D(E)$ where (E, d) is a quasi-convex space.

Another interesting application is given by the notion of *geodesic space*, as defined in [Lan13]:

Definition 5.28 (**Geodesic (space)**)**.** *A curve $\gamma : [a, b] \to E$ on a metric space (E, d) is called a* geodesic *if γ has constant speed, i.e. there exists $\lambda \geq 0$ such that*

$$l(\gamma)|_{[t,t']} \leq \lambda|t - t'|, \quad \forall t, t' \in [a, b],\ t < t',$$

and if $l(\gamma)|_{[t,t']} = d(\gamma(t), \gamma(t'))$.

A metric space (E, d) is called a geodesic space *if for every pair of points $x, y \in E$, there exists a geodesic $\gamma : [0, 1] \to E$ connecting x and y.*

Remark 5.29. Every geodesic space is a length space but the converse is not necessarily true. The necessary conditions for equivalence are given by Hopf–Rinow Theorem (see e.g. [BH99], Proposition 3.7). In particular, a length space X which is complete and locally compact is a geodesic space. A direct counterexample is given by the length space $\mathbb{R}^2 \setminus \{(0, 0)\}$, which is neither a geodesic space, nor complete (but locally compact).

In the context of a geodesic space, (83) rewrites as

$$|f(x) - f(y)| \leq \|\nabla f\|_\infty \cdot d(x, y),\ \forall x, y \in E,$$

for any function $f \in D(E)$ where (E, d) is a geodesic space.

Example 5.30 ([BH99], Theorem I.1.6)**.** Every normed vector space $(E, \|\cdot\|_E)$ with metric $d(x, y) = \|x - y\|_E$ is a geodesic space: the straight line between two points x, y of E will be the geodesic to x and y. More precisely, the map

$$\gamma : [0, 1] \to E : t \to (1 - t)x + ty$$

defines a path which is a linearly reparameterized geodesic connecting x to y since

$$l(\gamma)|_{[t,t']} \leq \|x - y\|_E |t - t'| \leq \lambda|t - t'|, \quad \forall 0 \leq t \leq t' \leq 1.$$

The following corollary is a straightforward consequence of Proposition 5.27, and it gives sufficient condition to ensure the equivalence between $Lip(E)$ and $D(E)$.

Corollary 5.31 ([DCJ10], Corollary 2.4)**.** *If (E, d) is a quasi-convex space, then $Lip(E) = D(E)$.*

We can also gives a short example of the opposite case:

Example 5.32. [[DCJ10], Example 2.6] Let $E = [0, +\infty)$. For $n \geq 1$, we define $I_n = [n-1, n]$ such that $E = \cup_{n\geq 1} I_n$. Consider the sequence of functions $f_n : [0, 1] \to \mathbb{R}$ defined by

$$f_n(x) = \begin{cases} x & \text{if } x \in [0, \frac{1}{n}], \\ \frac{nx+n-1}{n^2} & \text{if } x \in [\frac{1}{n}, 1]. \end{cases}$$

Then, for each pair of points $x, y \in I_n$, we write $d_n(x, y) = f_n(|x - y|)$ and we define a metric of E as follows: given a pair of points $x, y \in E$, with $x < y$, $x \in I_n$ and $y \in I_m$, let

$$d(x, y) = \begin{cases} d_n(x, y), & \text{if } n = m, \\ d_n(x, n) + \sum_{i=n+1}^{m-1} d_i(i-1, i) + d_m(m-1, y), & \text{if } n < m. \end{cases}$$

We can show that d is in fact a metric which coincides locally with the Euclidean metric d_e. More precisely, if $x \in I_n$, let $J_x = \left(x - \frac{1}{n-1}, x + \frac{1}{n-1}\right)$, then $d|_{J_x} = d_e|_{J_x}$.

Now, let $g : E \to \mathbb{R}$ be a function defined as

$$g(x) = \begin{cases} 2k - x, & \text{if } x \in I_{2k}, \\ x - 2k, & \text{if } x \in I_{2k+1}. \end{cases}$$

For $x \in E$ fixed, we suppose that there exists $n \geq 1$ such that $x \in I_n$. If $y \in J_x$, we have that

$$\text{Lip}\,(g)\,(x) = \limsup_{y\to x} \frac{|g(x) - g(y)|}{d(x, y)} = \limsup_{y\to x} \frac{|x - y|}{|x - y|} = 1,$$

so $g \in D(E)$.

Otherwise, for each $n \geq 1$, we have $|g(n-1) - g(n)| = 1$ and $d(n-1, n) = f_n(1) = \frac{2n-1}{n^2}$. Then, since

$$\lim_{n\to\infty} \frac{|g(n-1) - g(n)|}{d(n-1, n)} = \lim_{n\to\infty} \frac{1}{\frac{2n-1}{n^2}} = \infty,$$

so that $g \notin Lip(E)$. In particular, since $Lip(X) \neq D(X)$, (X, d) is not a length space.

As highlighted in [DCJ10], in general, if E is not a compact space, then

$$Lip(E) \subsetneq Lip(E) \cap D(E) \begin{array}{c} \nearrow \\ \searrow \end{array} \begin{array}{c} Lip_{loc}(E) \\ D(E) \end{array} \begin{array}{c} \searrow \\ \nearrow \end{array} C(E),$$

where $Lip_{loc}(E)$ denotes the set of locally Lipschitz functions on E.

5.3.2 Hairer and Mattingly Criterion for strong Feller

The following standard criterion to achieve the strong Feller property is stated as Proposition 3.11 by Hairer and Mattingly in [HM06], and is directly based on Proposition 7.1.5 by Da Prato and Zabczyk in [PZ96].

Proposition 5.33 ([HM06], Proposition 3.11)**.** *A Markov semigroup $(P_t)_{t\geq 0}$ on a Hilbert space $\mathcal{H}$ is strong Feller if, $\forall\varphi : \mathcal{H} \to \mathbb{R}$ with $\|\varphi\|_\infty$ and $\|\nabla\varphi\|_\infty < \infty$, one has*

$$|\nabla P_t\varphi(x)| \leq C(\|x\|)\|\varphi\|_\infty, \tag{85}$$

where $C : \mathbb{R}_+ \to \mathbb{R}$ is a fixed nondecreasing function.

Before stating the proof, we recall the notion of ultra Feller:

Definition 5.34 (Ultra Feller). *A Markov semigroup $(P_t)_{t\geq 0}$ over a Polish space $\mathcal{X}$ is said to be ultra Feller if, for all $t \geq 0$, $x \mapsto P_t(x, \cdot)$ is continuous in the topology of Total variation convergence, i.e.*

$$\lim_{n\to\infty} d(x_n, x) = 0 \;\Rightarrow\; \lim_{n\to\infty} \|P_t(x_n, \cdot) - P_t(x, \cdot)\|_{TV} = 0.$$

Remark 5.35. **(i)** In fact, the ultra Feller property and the usefulness of this characterization has already been pointed out in [DM83]. Despite this, its use has become widespread among probabilists since Seidler popularised it only almost 20 years later in [Sei].

(ii) In particular, since the Total variation convergence is stronger than the strong convergence, it implies that any ultra Feller Markov semigroup is also strong Feller.

Proof. We show that the above inequality implies the continuity of the map $x \mapsto P_t(x, \cdot)$ in the topology of strong convergence.

Indeed, mixing properties from the inequalities (83) and (85), it leads to

$$\begin{aligned} |P_t\varphi(x) - P_t\varphi(y)| &\leq \|\nabla P_t\varphi\|_\infty d(x, y) \\ &\leq |\nabla P_t\varphi(x)| d(x, y) \\ &\leq C(\|x\|)\|\varphi\|_\infty d(x, y), \end{aligned}$$

which holds true for any measurable bounded functions φ. In particular, using the equivalence from Proposition 7.1.5 by Da Prato and Zabczyk in [PZ96], we have the continuity of $(P_t)_{t\geq 0}$ in the topology of the Total variation convergence, and $(P_t)_{t\geq 0}$ is in particular strong Feller as in Definition 5.34 and Remark 5.35. □

5.4 The setting for our infinite dimensionsal self-repelling diffusion

Going back to model (77), this infinite dimensionsal model is in fact a forward extension of the finite-dimensional approach investigated in [BG17], where the authors have considered a self-repelling diffusion evolving on the state space $M \times \mathbb{R}^n$, where M is a compact connected oriented Riemannian manifold.

To explore such stochastic processes, let (M, g) be a $n-$dimensional compact Riemannian manifold and let $p \in M$. To construct a Brownian motion $(B_t)_{t\geq 0}$ on M starting at p, one can view it as *rolling M* along a Brownian path in its tangent space at p, T_pM.

More precisely, let $(\beta_t)_{t\geq 0}$ be a Brownian motion in $\mathbb{R}^n$ starting at the origin. Using the isometry between (T_pM, g_p) and $(\mathbb{R}^n, \langle\cdot,\cdot\rangle_{\mathbb{R}^n})$, where $\langle\cdot,\cdot\rangle_{\mathbb{R}^n}$ denotes the Euclidean inner product on $\mathbb{R}^n$, we identify T_pM with $\mathbb{R}^n$ and apply the *stochastic development map* at p to $(\beta_t)_{t\geq 0}$: the resulting process $(B_t)_{t\geq 0}$ is the Brownian motion on M and it can be formally defined as the solution of the corresponding stochastic differential equation

$$dX_t = \sum_{i\in I} E_i(X_t) \circ d\beta_t^i,$$

where $E_1, \cdots, E_n$ is a local orthonormal frame field near p with respect to the Riemannian metric g, $\circ$ denotes the Stratonovich integral, and $\beta_t^1, \cdots, \beta_t^n$ are independent standard $\mathbb{R}-$valued Brownian motions. More details are given in the seminal work of Itô (see e.g. [Itô50] and [IM74]).

Analytically, Brownian motion is characterized by its infinitesimal generator: in $\mathbb{R}^n$, the generator of standard Brownian motion is given by $\frac{1}{2}\Delta$, where Δ denotes the Euclidean Laplacian. In the case of a Riemannian manifold (M, g), a Brownian motion is defined as the diffusion process whose generator is $\frac{1}{2}\Delta_{LB}$, where Δ_{LB} denotes the Laplace–Beltrami operator. For an intuitive exposition, see [Str96] and for a more technical and formal way, see [Ken87].

Indeed, using Itô's formalism, above SDE leads to

$$dX_t = \frac{1}{2}\sum_{i\in I} \nabla_{E_i} E_i(X_t)\mathrm{d}t + \sum_{i\in I} E_i(X_t) d\beta_t^i,$$

where ∇ is the Levi-Civita connection, which is exactly $\frac{1}{2}\Delta_{LB}$.

In [BG17], the authors proved that the induced Markov semigroup has the strong Feller property and has a unique invariant probability μ given as the product of the normalized Riemannian measure on M and a Gaussian measure on $\mathbb{R}^n$. More precisely, let M be a smooth Riemannian manifold, $V : M \times M \to \mathbb{R}$, and $w : [0, +\infty) \to [0, +\infty)$ be continuous functions. Then, we consider the solution $(X_t)_{t\geq 0}$ of the SDE

$$dX_t = \sigma \mathrm{d}B_t(X_t) - \nabla V_t(X_t)\mathrm{d}t,$$

where $\sigma > 0$, $(B_t)_{t\geq 0}$ is a Brownian vector field on M, and

$$V_t(x) = w_t \int_0^t V(X_s, x)ds.$$

In view of [BG17], we need further assumptions on the model:

Hypothesis 10. **(i)** $w_t = 1$;

(ii) M is finite-dimensional, compact, oriented, connected and without boundary;

(iii) V is a Mercer kernel, which means that $V(x, y) = V(y, x)$ and

$$\int_M \int_M V(x, y) f(x) f(y) dx dy \geq 0,$$

for all $f \in L^2(dx)$ where dx denotes the Riemannian measure;

(iv) By Mercer Theorem, V can be written as

$$V(x, y) = \sum_{i\geq 1} a_i e_i(x) e_i(y), \tag{86}$$

where $a_i \geq 0$ and $\{e_i\}_{i\geq 1}$ is an orthonormal family of eigenfunctions in $L^2(dx)$ of the operator $f \to Vf$, where $Vf(x) = \int_M V(x, y) f(y) dy$.

Then, we impose that the sum (86) is finite and that the $\{e_i\}_{i\geq 1}$ are eigenfunctions of the Laplace operator.

Theorem 5.36 ([BG17]). *Under Hypothesis 10, it follows that:*

- ***(i)*** (Proposition 1). *There exists a unique global strong solution;*
- ***(ii)*** (Theorem 5). *The strong Feller property holds;*
- ***(iii)*** (Theorem 5). *The system admits a unique invariant probability measure μ which is given explicitly as the product of the uniform probability on M and a Gaussian probability on $\mathbb{R}^n$;*
- ***(iv)*** (Theorem 6). *$(P_t)_{t\geq 0}$ converges to μ at an exponential rate.*

Let's extend the situation to the infinite dimensionsal setting: just like in [BCG15], we consider (77) and we assume that f is an even, $2\pi-$periodical function which is sufficiently regular such that the coefficients $(a_n)_{n\geq 1}$ of the corresponding Fourier series,

$$f(x) \sim \frac{a_0}{2} + \sum_{n\geq 1} a_n \cos(nx),$$

form a strictly positive rapidly decreasing sequence. Recall that the space of rapidly decreasing sequences of order k is defined by

$$O^k = \left\{ (a_n)_{n\geq 1} \ : \ \sum_{n=1}^{\infty} (1+n^2)^k \, a_n^2 < \infty \right\}.$$

Here, we will have to assume that the sequence $(a_n)_{n\geq 1}$ lies at least in O^5. We finally assume that

$$g(x) = \sum_{n\geq 1} a_n^{1/2} n \left(u_0^n \sin(nx) + v_0^n \cos(nx) \right),$$

where $(u_0^n)_{n\geq 1}$, $(v_0^n)_{n\geq 1} \in l^2$. We now are able to rewrite the SDE as a system on the Hilbert space $H = \mathbb{R} \times l^2 \times l^2$ as

$$\begin{cases} X_t = x + \displaystyle\int_0^t \sum_{n\geq 1} n \left(a_n^{1/2} \sin(nX_s) u_s^n + a_n^{1/2} \cos(nX_s) v_s^n \right) ds + \beta_t, \\ u_t^n = u_0^n + a_n^{1/2} \displaystyle\int_0^t \cos(nX_s) ds, \quad n \geq 1, \\ v_t^n = v_0^n - a_n^{1/2} \displaystyle\int_0^t \sin(nX_s) ds, \quad n \geq 1, \end{cases} \tag{87}$$

where β_t is a standard one-dimensional Brownian motion. This system can now be written under the form of a SDE as

$$\begin{cases} dY_t = F(Y_t)\mathrm{d}t + \sigma dW_t, \\ Y_0 = y, \end{cases} \tag{88}$$

where $Y_t = (X_t, (u_t^n)_n, (v_t^n)_n)$, $\sigma = (1,0,0)$, W_t is a cylindrical Wiener process, and $F : H \to H$, is defined as

$$F \begin{pmatrix} x \\ (u^n)_n \\ (v^n)_n \end{pmatrix} = \begin{pmatrix} \langle \left(a_n^{1/2} n \sin(nx)\right)_n, (u^n)_n \rangle_{l^2} + \langle \left(a_n^{1/2} n \cos(nx)\right)_n, (v^n)_n \rangle_{l^2} \\ \left(a_n^{1/2} \cos(nx)\right)_n \\ \left(a_n^{1/2} \sin(nx)\right)_n \end{pmatrix}.$$

As it is well demonstrated in [BCG15], we can summarize the most important results:

Theorem 5.37 ([BCG15], Propositions 1, 2, and 3). *Under the above assumptions for (88):*

(i) *There exists a unique strong solution in* $C\left((0,\infty]; H\right) \cap L^{\infty}_{loc}(0;\infty;H)$.[1]

(ii) *The induced Markov semigroup is defined by* $P_t\varphi(y) = \mathbb{E}\left(\varphi\left(Y_t\right) \,\middle|\, Y(0) = y\right)$ *for all* $\varphi \in \mathcal{B}_b(H)$, $t \geq 0$, *and it has the* $C_b(H)-$*Feller property.*

(iii) *The probability measure defined by*

$$\mu(dy) = \frac{dx}{2\pi} \otimes \prod_{n\geq 1} \mathcal{N}\left(0, \frac{1}{n^2}\right) du^n \otimes \prod_{n\geq 1} \mathcal{N}\left(0, \frac{1}{n^2}\right) dv^n$$

is invariant under $(P_t)_{t\geq 0}$.

Remark 5.38. In fact, the form of μ is inspired by the finite-dimensional case and using the strong convergence of a Galerkin type approximation of the solution of (87).

Sketch of the proof. Using Theorem 7.10 from [PZ92] since F is not globally but locally Lipschitz, we can show the existence and uniqueness of a strong solution under the following conditions:

(i) F is locally Lipschitz;

(ii) F is bounded on bounded sets;

(iii) There exists an increasing function $a : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\langle F(\tilde{y} + y), y^* \rangle_H \leq a(||\tilde{y}||_H)(1 + ||y||_{\mathbb{H}}), \quad \forall y, \tilde{y} \in H, \ \forall y^* \in \partial ||y||,$$

where $\partial || \cdot ||_H$ denotes the subdifferential of $|| \cdot ||_{\mathbb{H}}$.

Recall that, on a Banach space E with norm $|| \cdot ||_E$, the subdifferential $\partial ||x||_E$ of $|| \cdot ||_E$ at x is defined as the set

$$\partial ||x||_E = \{x^* \in E^* : ||x + y||_E - ||x||_E \geq \langle y, x^* \rangle_E, \ \forall y \in E\},$$

which is also the convex, closed, and nonempty set given by

$$\partial ||x||_E = \begin{cases} \{x \in E^* : \langle x, x^* \rangle_E = ||x||_E, \text{ with } ||x||_E = 1\} & \text{if } x \neq 0, \\ \{x \in E^* : ||x^*||_E \leq 1\} & \text{if } x = 0. \end{cases}$$

Then, letting $(y_k)_{k\geq 0}$ be a sequence of initial conditions in H such that $y_k \underset{k\to\infty}{\longrightarrow} y$, where $y \in H$ is the original initial condition, and

$$Y_k(t) = \left(X^k_t, (u^n_t)^k_{n\geq 1}, (v^n_t)^k_{n\geq 1}\right),$$

be the solution of (88) corresponding to every initial condition y_k, then for every $t > 0$,

$$||Y_k(t) - Y(t)||^2_H \underset{k\to\infty}{\longrightarrow} 0.$$

It implies the Feller property since, for every bounded and continuous function φ, the L^2 convergence of $Y_k(t)$ to $Y(t)$ implies the convergence in probability of $\mathbb{E}\left[\varphi\left(Y_k(t)\right)\right]$ to $\mathbb{E}\left[\varphi\left(Y(t)\right)\right]$, for any $t > 0$ fixed: hence, $P_t\varphi$ is also a bounded, continuous function.

[1] The *loc* as an index of L^∞ is used to indicate that this property holds on every compact subset of its domain.

Finally, by using the strong convergence of a Galerkin-type approximation, we can reuse the unique invariant probability measure in the finite-dimensional case to show that its infinite dimensionsal extension is invariant under $(P_t)_{t\geq 0}$. Namely, letting $H = H_N \times l^2 \times l^2$ where $H_N = \mathbb{R}\times\mathbb{R}^N\times\mathbb{R}^N$, and

$$\Pi_N : H \to H_N \times \{0\}^\infty \times \{0\}^\infty : (x,(u^n)_{n\geq 1},(v^n)_{n\geq 1}) \mapsto (x,(u_n)_{n=1}^N \times \{0\}^\infty,(v_n)_{n=1}^N \times \{0\}^\infty),$$

then for any $T > 0$,

$$\lim_{N\to\infty} \sup_{0\leq t\leq T} \|Y^{(N)}(t) - Y(t)\|_{\mathbb{H}} = 0,$$

where $(Y^{(N)}(t))_{t\geq 0}$ is the solution to

$$\begin{cases} dY_t^{(N)} = \Pi_N\left(F(Y_t^{(N)})\right)\mathrm{d}t + \sigma dW_t, \\ Y_0^{(N)} = \Pi_N(y), \end{cases}$$

which has a unique strong solution by classical results for SDE in finite-dimension. □

A natural question arises following this article:

Is μ the unique invariant probability measure with respect to $(P_t)_{t\geq 0}$?

Unfortunately, as said above, $(P_t)_{t\geq 0}$ does not have the strong Feller property. To see it, let $\mathcal{A} = \mathbb{R}\times O^1\times O^1 \subset H$ where O^1 is the space of rapidly decreasing sequences of order 1. Then,

Lemma 5.39. $\forall t\geq 0,\ y\in\mathcal{A} \Leftrightarrow Y(t)\in\mathcal{A}$. *In particular, $(P_t)_{t\geq 0}$ is not strong Feller.*

Proof. "⇒:" Let's suppose that $y = \left(x,(u_0^n)_n,(v_0^n)_n\right)\in\mathcal{A}$, which means

$$\sum_{n\geq 1}(1+n^2)\left(u_0^n\right)^2 < \infty \quad \left(\text{respectively } v_0^n\right).$$

By definition of $(u_t^n)_n$ (respectively $(v_t^n)_n$), we claim that $(u_t^n)_n \in\mathcal{A}$ (respectively $(v_t^n)_n\in\mathcal{A}$), $\forall t\geq 0$. In other words, it means that $Y(t)\in\mathcal{A}$.

Indeed, since

$$u_t^n = u_0^n + a_n^{1/2}\int_0^t \cos(nX_s)ds,$$

if $(u_0^n)_{n\geq 1}$ and $(a_n^{1/2})_n$ lie in O^1, then $(u_t^n)_n$ lies in O^1 $\forall t\geq 0$. We show that $(a_n^{1/2})_n \in O^1$. Since $(a_n)_{n\geq 1}\in O^5$ by assumption, then $\sum_{n\geq 1}(1+n^2)^5 a_n^2 < \infty$. By Cauchy-Schwartz inequality, it yields

$$\sum_{n\geq 1}(1+n^2)a_n \leq \sum_{n\geq 1}\left(1+n^2\right)^{-\frac{3}{2}}(1+n^2)^{\frac{5}{2}}a_n \leq \left(\sum_{n\geq 1}\left(1+n^2\right)^{-3}\right)^{\frac{1}{2}}\left(\sum_{n\geq 1}\left(1+n^2\right)^5 a_n^2\right)^{\frac{1}{2}} < \infty,$$

which concludes that $(a_n^{1/2})_{n\geq 1}\in O^1$. The same method can be applied to show $(v_t^n)_{n\geq 1}\in O^1$, $\forall t\geq 0$, and overall we have $Y(t)\in\mathcal{A}$ for all $t\geq 0$.

"⇐:" Now, let's suppose that $Y(t)\in\mathcal{A}$ for a fixed $t\geq 0$, and since

$$u_0^n = u_t^n - a_n^{1/2}\int_0^t \cos(nX_s)ds,$$

where $(u_t^n)_{n\geq 1}$ by assumption and $(a_n^{1/2})_{n\geq 1}\in O^1$ by Cauchy-Schwarz inequality as before. Since O^1 has a structure of vector space, then $(u_0^n)_{n\geq 1}\in O^1$. The same method can be applied to show $(v_0^n)_{n\geq 1}\in O^1$.

In particular, it yields $\mathbf{1}_{\mathcal{A}}$ is invariant under $(P_t)_{t\geq 0}$, $\forall t\geq 0$, a measurable but not continuous map, so that $(P_t)_{t\geq 0}$ is not strong Feller. □

Remark 5.40. The fact that $\mathbf{1}_{\mathcal{A}}$ is not continuous can be seen as follows. Note that l^2 is connected, by arc-connectedness. If $\mathbf{1}_{\mathcal{A}}$ is continuous, then both pre-images $\mathbf{1}_{\mathcal{A}}^{-1}(1) = \mathcal{A}$, $\mathbf{1}_{\mathcal{A}}^{-1}(0) = \mathcal{A}^c$ would be closed in H, and $\mathcal{A}^c$ would be also open in H as the complementary of a closed subset.

It is well-known that the only open-closed subset of a connected space is either $\emptyset$ or the whole set, in our case H. Since $\left(\frac{1}{n}\right)_{n\geq 1}$ lies in l^2 but not in O^1, $\mathcal{A}^c \neq \emptyset$. And since $\left(\frac{1}{n^2}\right)_{n\geq 1}$ lies both in l^2 and in O^1, then $\mathcal{A}^c \neq H$.

Chapter 6

The asymptotic strong Feller property

Based on the work of Hairer and Mattingly in [HM06], here is a presentation of the main tools needed with the aim to state an abstract ergodic result, the *asymptotic strong Feller property*. Since the results obtained are only stated completely in the original text [HM06] and for a seek of clarity, we also provide a complete proof for each of them.

We start with a Polish space $\mathcal{X}$ (complete, separable, metrizable).

Definition 6.1 (Pseudo-metric). *A pseudo-metric for $\mathcal{X}$ is a continuous function $d : \mathcal{X}^2 \to \mathbb{R}_+$, such that $d(x,x) = 0\ \forall x \in \mathcal{X}$, $d(x,y) = d(y,x)\ \forall x,y \in \mathcal{X}$, and such that the triangle inequality is satisfied.*

Remark 6.2. **(i)** In fact, a pseudo-metric d can be defined as a classical metric on $\mathcal{X}$ except for the condition $d(x,y) = 0 \Rightarrow x = y$.

(ii) We say that a pseudo-metric d_1 is larger than d_2 if $d_1(x,y) \geq d_2(x,y)$, $\forall (x,y) \in \mathcal{X}^2$. It gives meaning to the notion of increasing sequence of pseudo-metrics used in the next definition.

Definition 6.3 (Totally separating system of (pseudo-)metrics). *Let $\{d_n\}_{n\geq 0}$ be an increasing sequence of (pseudo-)metrics over $\mathcal{X}$. If*

$$\lim_{n\to\infty} d_n(x,y) = 1, \quad \forall x \neq y,$$

then $\{d_n\}_{n\geq 0}$ is said to be a totally separating system of (pseudo-)metrics *for $\mathcal{X}$.*

In other words, we impose that $\lim_{n\to\infty} d_n(x,y) = \mathbf{1}_{x\neq y}$.

Example 6.4. Let d be an arbitrary continuous (pseudo)-metric on $\mathcal{X}$, and let $(a_n)_{n\geq 0}$ be an increasing sequence in $\mathbb{R}$ such that $\lim_{n\to\infty} a_n = \infty$. Then the sequence $\{d_n\}_{n\geq 0}$, defined as

$$d_n(x,y) = \min\{1;\ a_n d(x,y)\},$$

is a totally separating system of (pseudo-)metrics.

Recall that, given a (pseudo)-metric d on $\mathcal{X}$, a function $\varphi : \mathcal{X} \to \mathbb{R}$ is said to be d−Lipschitz if, $\forall x,y \in \mathcal{X}$, there exists $C > 0$ a constant such that

$$|\varphi(x) - \varphi(y)| \leq C d(x,y).$$

Definition 6.5 (Seminorm). *Let d be a pseudo-metric. We define the* seminorm *on the set of d-Lipschitz functions $\varphi : \mathcal{X} \to \mathbb{R}$ by*

$$\|\varphi\|_d = \sup_{x\neq y} \frac{|\varphi(x) - \varphi(y)|}{d(x,y)}.$$

Remark 6.6. **(i)** Intuitively, this definition does not allow us to speak about a classical norm since $\|\varphi\|_d = 0$ if and only if φ is a constant function.

(ii) The seminorm must not be confused with the notion of pointwise Lipschitz function as in Definition 5.21.

It lets us define a dual seminorm on the space of finite signed Borel measures on $\mathcal{X}$ with vanishing integral by

$$|||\nu|||_d = \sup_{\|\varphi\|_d=1} \int_{\mathcal{X}} \varphi(x)\nu(dx).$$

Similarly for positive finite Borel measures,

Definition 6.7 (**Wasserstein distance**). *Let μ_1, μ_2 be two positive finite Borel measures on $\mathcal{X}$ with equal mass, and let $C(\mu_1, \mu_2)$ be the set of positive measures on $\mathcal{X}^2$ with marginals μ_1 and μ_2. We define*

$$\|\mu_1 - \mu_2\|_d = \inf_{\mu \in C(\mu_1,\mu_2)} \int_{\mathcal{X}^2} d(x, y)\mu(dx, dy). \tag{89}$$

It is known as the Wasserstein distance of order 1 between μ_1 and μ_2, *denoted $W_1^d(\mu_1, \mu_2)$ (see e.g. [Vil09], Section 6).*

We are now able to define the asymptotic strong Feller property as proposed by Hairer and Mattingly in [HM06].

Definition 6.8 (**Asymptotic strong Feller property**). *A Markov transition semigroup $(P_t)_{t\geq 0}$ on $\mathcal{X}$ is said to be* asymptotically strong Feller *at x if there exists $(d_n)_{n\geq 0}$ a totally separating system of pseudo-metrics for $\mathcal{X}$, and a sequence of positive times $(t_n)_{n\geq 0}$ such that*

$$\inf_{U\in\mathcal{U}_x} \limsup_{n\to\infty} \sup_{y\in U} \|P_{t_n}(x,\cdot) - P_{t_n}(y,\cdot)\|_{d_n} = 0, \tag{90}$$

where $\mathcal{U}_x$ is the set of all the open neighborhoods of x.

Remark 6.9. Let $B(x, \gamma)$ be the open ball of radius γ centered at x. Equivalently, $(P_t)_{t\geq 0}$ is asymptotically strong Feller at x if

$$\lim_{\gamma\to 0} \limsup_{n\to\infty} \sup_{y\in B(x,\gamma)} \|P_{t_n}(x,\cdot) - P_{t_n}(y,\cdot)\|_{d_n} = 0. \tag{91}$$

6.1 Asymptotic strong Feller property in practice

One of the main advantages of the asymptotic strong Feller property is that it implies similar ergodic results as the strong Feller one, see Theorem 6.15 for example, while being a weaker property.

Before diving in the ergodicity of asymptotically strong Feller Markov semigroup, we introduce a practical criterion that guarantees the asymptotic strong Feller property.

6.1.1 Practical criterion to verify asymptotic strong Feller property

As for the strong Feller property in Proposition 5.33, a useful criterion can be given through this proposition, stated by Hairer and Mattingly. Here, ∇ denotes the pointwise Lipschitz constant as defined in Definition 5.21.

Proposition 6.10 ([HM06], Proposition 3.12). *Let $(t_n)_{n\geq 1}$, $(\delta_n)_{n\geq 1}$ be two positive sequences, $(t_n)_{n\geq 1}$ a nondecreasing one and $(\delta_n)_{n\geq 1}$ converges to zero. A semigroup $(P_t)_{t\geq 0}$ on a Hilbert space $\mathcal{H}$ is asymptotically strong Feller at $x \in \mathcal{H}$ if, $\forall \varphi : \mathcal{H} \to \mathbb{R}$ with $\|\varphi\|_\infty$ and $\|\nabla\varphi\|_\infty < \infty$, one has*

$$|\nabla P_{t_n}\varphi(x)| \leq C(\|x\|)(\|\varphi\|_\infty + \delta_n\|\nabla\varphi\|_\infty), \quad \forall n \geq 1, \tag{92}$$

where $C : \mathbb{R}_+ \to \mathbb{R}$ is a fixed nondecreasing function.

Before stating the proof, we need the following as useful as crucial result.

Lemma 6.11 ([HM06], Lemma 3.3)**.** *Let d be a continuous pseudo-metric on $\mathcal{X}$, and let μ_1, μ_2 be two positive measures on $\mathcal{X}$ with equal mass. Then,*

$$|||\mu_1 - \mu_2|||_d = ||\mu_1 - \mu_2||_d.$$

Proof. This result is a direct consequence of the Monge-Kantorovich duality, especially in the setting of a separable metric space (see e.g. sections 5 and 6, [Vil09]).

In particular, we define the equivalence relation on $\mathcal{X}$ by $x \sim y \iff d(x, y) = 0$, such that by setting $\mathcal{X}_d = \mathcal{X}/\sim$, d is a well-defined metric on $\mathcal{X}_d$ and $(\mathcal{X}_d, d)$ is hence a separable metric space, possibly no longer complete.

Then, with $\pi : \mathcal{X} \to \mathcal{X}_d$ the quotient map defined as $\pi(x) = [x]$, both sides of (89) do not change if the measure μ_1 (resp. μ_2) is replaced by $\pi_\# \mu_1$ (resp. $\pi_\# \mu_2$). Then, as explained in Remark 6.5 in [Vil09], Monge-Kantorovich duality in Theorem 5.10(i) and the particular case 5.4 of separable metric space in [Vil09] together lead to the duality formula for the Kantorovich–Rubinstein distance, and

$$\begin{aligned} ||\mu_1 - \mu_2||_d &= \inf_{\mu \in C(\mu_1,\mu_2)} \int_{\mathcal{X}^2} d(x, y)\mu(dx, dy) \\ &= \sup_{||\phi||_{\mathrm{Lip}} \le 1} \left\{ \int_{\mathcal{X}} \phi(x)\mu_1(dx) - \int_{\mathcal{X}} \phi(x)\mu_2(dx) \right\}, \end{aligned}$$

where the sup is taken on the set of 1−Lipschitz functions. By a scaling argument, one can replace the 1−Lipschitz function by functions with dual seminorm $||\varphi||_d = 1$ so that letting $\varphi = \frac{\phi}{||\phi||_d}$, it yields

$$||\mu_1 - \mu_2||_d = |||\mu_1 - \mu_2|||_d.$$

□

Proof of Proposition 6.10. Let $\varepsilon > 0$, and let d_ε be the metric on $\mathcal{H}$ defined as

$$d_\varepsilon(x, y) = 1 \wedge \frac{||x - y||}{\varepsilon}, \quad \forall x, y \in \mathcal{H}.$$

Since $||\nabla\varphi||_\infty$ is finite, suppose that $||\nabla\varphi||_\infty < 1$ without loss of generality by scaling φ. Then, $||\varphi||_{d_\varepsilon} \le 1$, where $|| \cdot ||_{d_\varepsilon}$ is defined as in Definition 6.5, which leads to

$$\sup_{x,y\in\mathcal{H}, y\neq x} \frac{|\varphi(x) - \varphi(y)|}{d_\varepsilon(x, y)} \le 1.$$

By definition of d_ε, it follows that

$$|\varphi(x) - \varphi(y)| \le 1 \wedge \frac{||x - y||}{\varepsilon}, \quad \forall x, y \in \mathcal{H},\ y \neq x,$$

or equivalently

$$\frac{|\varphi(x) - \varphi(y)|}{||x - y||} \le \frac{1}{||x - y||} \wedge \frac{1}{\varepsilon}, \quad \forall x, y \in \mathcal{H},\ y \neq x.$$

By taking the limit when $y \to x$, we have the relation

$$|\nabla\varphi|(x) \le \frac{1}{\varepsilon}, \quad \forall x \in \mathcal{H}.$$

Combining with the assumption on $\nabla P_{t_n}\varphi(x)$ in Equation 92, we now have

$$\lim_{y\to x}\frac{|P_{t_n}\varphi(x) - P_{t_n}\varphi(y)|}{\|x-y\|} \leq C(\|x\|)(\|\varphi\|_\infty + \frac{\delta_n}{\varepsilon}).$$

Using the assumption that $\|\varphi\|_\infty < \infty$, we can bound $C(\|x\|)$ by another nondecreasing function C' such that

$$C(\|x\|)\|\varphi\|_\infty \leq C'(\|x\|).$$

Since C' is a nondecreasing function, then $C'(\|x\|) \leq C'(\|x\| \vee \|y\|)$, where $a \vee b = \max\{a, b\}$.

We now have

$$|P_{t_n}\varphi(x) - P_{t_n}\varphi(y)| \leq \|x-y\|C'(\|x\| \vee \|y\|)(1 + \frac{\delta_n}{\varepsilon}).$$

By definition of $\|\cdot\|_{d_\varepsilon}$ and by Lemma 6.11, we get

$$\begin{aligned}\|P_{t_n}(x,\cdot) - P_{t_n}(y,\cdot)\|_{d_\varepsilon} &= \sup_{\|\varphi\|_{d_\varepsilon}=1}\int_{\mathcal{H}}\varphi(h)\,(P_{t_n}(x,dh) - P_{t_n}(y,dh))\\ &= \sup_{\|\varphi\|_{d_\varepsilon}=1}|P_{t_n}\varphi(x) - P_{t_n}\varphi(y)|.\end{aligned}$$

Using the previous inequality for $|P_{t_n}\varphi(x) - P_{t_n}\varphi(y)|$ when $y \to x$, with $\varepsilon = a_n = \sqrt{\delta_n}$, $d_n := d_{a_n}$ is a totally separating system of pseudo-metrics since $a_n \downarrow 0$ and it remains that for any open neighborhood U of x,

$$\sup_{y\in U}\|P_{t_n}(x,\cdot) - P_{t_n}(y,\cdot)\|_{d_n} \leq \sup_{y\in U}\|x-y\|C'(\|x\| \vee \|y\|),$$

so $\lim_{\gamma\to 0}\limsup_{n\to\infty}\sup_{y\in B(x,\gamma)}\|P_{t_n}(x,\cdot) - P_{t_n}(y,\cdot)\|_{d_n} = 0$ and the asymptotic strong Feller property holds. □

Remark 6.12. **(i)** The assumption $\|\nabla\varphi\|_\infty < 1$ is sufficient since we can extend to all functions φ such that $\|\nabla\varphi\|_\infty$ is finite by a scaling argument.

(ii) It is also possible to directly show it for functions φ that are Fréchet Differentiable and to extend them to pointwise Lipschitz functions using an approximation with Fréchet Differentiable functions, which is possible due to [Cer99] or [PZ96]. A detailed paragraph is dedicated in [BH22], Remark 5.31.

6.1.2 A direct ergodic consequence

We start by recalling some consequences about ergodicity of a Markov semigroup under the strong Feller property.

Given a probability measure μ, we can see the elements of its support as follows,

Lemma 6.13 ([Par68], Theorem 2.1 or [HM06], Lemma 3.7)**.** *Given a separable metric space X and μ a measure in X, a point $x \in supp(\mu) \Leftrightarrow \mu(U) > 0$ for all open set U containing x.*

A useful consequence of strong Feller property about the support of a probability measure follows directly.

Proposition 6.14 ([BH22], Proposition 5.18 ***(i)***)**.** *Let $(P_t)_{t\geq 0}$ be a strong Feller Markov transition semigroup. If μ_1, μ_2 are two distinct ergodic invariant measures for $(P_t)_{t\geq 0}$, then $supp(\mu_1) \cap supp(\mu_2) = \emptyset$.*

Similarly, the asymptotic strong Feller property of a Markov transition semigroup leads to a characterization of the support of distinct ergodic invariant probability measures,

Theorem 6.15 ([HM06], Theorem 3.16)**.** *Let $(P_t)_{t\geq 0}$ be a Markov transition semigroup on $\mathcal{X}$, and let μ, ν be two distinct ergodic invariant probability measures for $(P_t)_{t\geq 0}$. If $(P_t)_{t\geq 0}$ is asymptotically strong Feller at x, then $x \notin supp(\mu) \cap supp(\nu)$.*

The proof is postponed to the next section since it relies on further results. As a direct consequence of Theorem 6.15, one can see the power of the asymptotic strong Feller property and the similarities about its consequences compared to the strong Feller one.

Corollary 6.16 ([HM06], Corollary 3.17)**.** *If $(P_t)_{t\geq 0}$ is an asymptotically strong Feller Markov transition semigroup on $\mathcal{X}$ and there exists a point $x \in \mathcal{X}$ such that $x \in supp(\mu)$ for every invariant probability measure μ of $(P_t)_{t\geq 0}$, then there exists at most one invariant probability measure.*

6.2 Links between the asymptotic strong and strong Feller Properties

Our purpose now is to show that the asymptotic strong Feller property leads to similar conclusions than the strong Feller property about the supports of distinct ergodic invariant probability measures.

We remark the following characterization of the limit of the seminorm associated to a family of continuous pseudo-metrics, thanks to the sense we gave to the Wasserstein distance in (89).

Lemma 6.17 ([HM06], Lemma 3.4)**.** *Let $\{d_n\}_{n\geq 0}$ be a bounded and increasing family of continuous pseudo-metrics on $\mathcal{X}$, and let*

$$d(x, y) = \lim_{n\to\infty} d_n(x, y),$$

for all $x, y \in \mathcal{X}$. Then,

$$\lim_{n\to\infty} \|\mu_1 - \mu_2\|_{d_n} = \|\mu_1 - \mu_2\|_d,$$

for any two positive measures μ_1, μ_2 with equal mass on $\mathcal{X}$.

Proof. Since the sequence is bounded and increasing, by a monotone convergence argument, the left hand-side limit exists: let's denote this limit by L.

By (89), and the increasing characteristic of $\{d_n\}_{n\geq 0}$, we have $\|\mu_1 - \mu_2\|_d \geq L$, so it is sufficient to show that the converse bound holds true.

Let μ_n be the measure in $C(\mu_1, \mu_2)$ that realize the infimum in (89) for d_n: the existence of such a minimizing coupling follows Theorem 4.1 in [Vil09].

Moreover, for any $n \geq 0$, the marginals of μ_n are constant, equal to μ_1 (respectively μ_2). Since μ_1 (respectively μ_2) is a positive, finite Borel measure on a complete, separable metric space $\mathcal{X}$, then μ_1 (respectively μ_2) is tight (see e.g. [Par68], Chapter II, Theorem 3.2).

By definition it implies that for any $\varepsilon > 0$, there exists of a compact K_1 (respectively K_2) such that $\mu_1(\mathcal{X}\backslash K_1) \leq \varepsilon$ (respectively $\mu_2(\mathcal{X}\backslash K_2) \leq \varepsilon$). Since $K_1 \times K_2$ is compact in $\mathcal{X}^2$ as a product of compact, then for any $\delta > 0$, we obtain $\mu_n(\mathcal{X}^2\backslash K_1 \times K_2) \leq \delta$ by setting $\delta = \frac{\varepsilon}{2}$ in the tightness equivalence for the marginals.

Then, by Prokhorov's Theorem, one can extract a weakly converging subsequence: let d_∞ be the limiting measure, so that for $m \geq n$,

$$\int_{\mathcal{X}^2} d_n(x, y)\mu_m(dx, dy) \leq \int_{\mathcal{X}^2} d_m(x, y)\mu_m(dx, dy) \leq L,$$

and by continuity of $\{d_n\}_{n\geq 0}$, the weak convergence of the subsequence when $m \to \infty$, we have

$$\int_{\mathcal{X}^2} d_n(x, y)\mu_\infty(dx, dy) \leq L.$$

Finally, by dominated convergence, we have $\int_{\mathcal{X}^2} d(x,y)\mu_\infty(dx,dy) \leq L$, so that $\|\mu_1 - \mu_2\|_d \leq L$. □

We also recall:

Definition 6.18 (Total variation norm). *The* Total variation norm *of a finite signed measure μ is defined as*

$$\|\mu\|_{TV} = \frac{1}{2}\left(\mu^+(\mathcal{X}) + \mu^-(\mathcal{X})\right),$$

where $\mu = \mu^+ - \mu^-$ is the Jordan decomposition of μ.

Then, as a consequence of Lemma 6.17:

Corollary 6.19 ([HM06], Corollary 3.5)**.** *Let $\{d_n\}_{n\geq 0}$ be a totally separating system of pseudo-metrics for $\mathcal{X}$. Then,*

$$\|\mu_1 - \mu_2\|_{TV} = \lim_{n\to\infty} \|\mu_1 - \mu_2\|_{d_n},$$

for any two positive measures μ_1, μ_2 with equal mass on $\mathcal{X}$.

Proof. Since the difference between two positive measures is a signed measure, with the help of Definition 6.18 and by definition of the Total variation norm, one has

$$\|\mu_1 - \mu_2\|_{TV} = \inf_{\mu\in C(\mu_1,\mu_2)} \mu\{(x,y) \mid x \neq y\},$$

where the right hand-side is exactly $\|\mu_1 - \mu_2\|_d$ with $d(x,y) = \mathbf{1}_{x\neq y}$. Indeed, using Kantorovich-Rubinstein Theorem (Theorem 1.14 in [Vil03]), with $d(x,y) = \mathbf{1}_{x\neq y}$, then

$$\inf_{\mu\in C(\mu_1,\mu_2)} \int_{\mathcal{X}^2} \mathbf{1}_{x\neq y} d\mu(x,y) = \sup_{0\leq f\leq 1} \int_{\mathcal{X}} f d\mu_1(x) - \int_{\mathcal{X}} f d\mu_2(x),$$

where the sup is taken on positive continuous functions $0 \leq f \leq 1$, using Remark 1.15 *(i)* from [Vil03]. Left-hand side writes

$$\inf_{\mu\in C(\mu_1,\mu_2)} \mu\{(x,y) : x \neq y\},$$

and right-hand side is exactly the Total variation norm of $\mu_1 - \mu_2$. Recalling that

$$\|\mu_1 - \mu_2\|_{TV} = \frac{1}{2}\sup_{\|f\|_\infty\leq 1} \int_{\mathcal{X}} f d\mu_1(x) - \int_{\mathcal{X}} f d\mu_2(x) = \frac{1}{2}\sup_{-1\leq f\leq 1} \int_{\mathcal{X}} f d\mu_1(x) - \int_{\mathcal{X}} f d\mu_2(x),$$

then for $-1 \leq g \leq 1$, we can use that $g = 2f - 1$ for $0 \leq f \leq 1$ so that we can rewrite

$$\begin{aligned}
\sup_{-1\leq g\leq 1} \int_{\mathcal{X}} g d\mu_1(x) - \int_{\mathcal{X}} g d\mu_2(x) &= \sup_{0\leq f\leq 1} \int_{\mathcal{X}} (2f-1) d\mu_1(x) - \int_{\mathcal{X}} (2f-1) d\mu_2(x) \\
&= 2\left(\sup_{0\leq f\leq 1} \int_{\mathcal{X}} f d\mu_1(x) - \int_{\mathcal{X}} f d\mu_2(x)\right) \\
&\quad + \mu_2(\mathcal{X}) - \mu_1(\mathcal{X}).
\end{aligned}$$

Since $\mu_1(\mathcal{X}) = \mu_2(\mathcal{X})$, it only remains that

$$\begin{aligned}
\sup_{0\leq f\leq 1} \int_{\mathcal{X}} f d\mu_1(x) - \int_{\mathcal{X}} f d\mu_2(x) &= \frac{1}{2}\left(\sup_{-1\leq g\leq 1} \int_{\mathcal{X}} g d\mu_1(x) - \int_{\mathcal{X}} g d\mu_2(x)\right) \\
&= \|\mu_1 - \mu_2\|_{TV}.
\end{aligned}$$

By Lemma 6.17, and since the limit of $\{d_n\}_{n\geq 0}$ is $d(x,y)$ by definition of a totally separating system of pseudo-metrics, the result is proven. □

Remark 6.20. We can exhibit a particular case where the asymptotically strong Feller property implies the strong Feller one: to this effect, let's suppose that (90) holds for $t_n = t > 0$ constant, $\forall n \geq 1$. In this case, $x \mapsto P_t(x, \cdot)$ is continuous in the topology of the Total variation, and thus P_s is strong Feller, $\forall s \geq t$. Indeed, by Corollary 6.19, for any $\varepsilon > 0$, there exists an open neighborhood U of x such that

$$\sup_{y \in U} \|P_t(x, \cdot) - P_t(y, \cdot)\|_{TV} \leq \varepsilon. \tag{93}$$

In the sense of the Total variation norm, we have that for any open neighborhood V of $P_t(x, \cdot)$, then $P_t(y, \cdot) \in V$, $\forall y \in U$ since

$$\|P_t(x, \cdot) - P_t(y, \cdot)\|_{TV} \leq \sup_{y \in U} \|P_t(x, \cdot) - P_t(y, \cdot)\|_{TV} \leq \varepsilon, \quad \forall y \in U$$

In other words, the function $x \mapsto P_t(x, \cdot)$ is continuous in the topology of Total variation. By definition,

$$\|\mu - \nu\|_{TV} = \sup\{|\mu f - \nu f| \mid f \in B(\mathcal{X}),\ \|f\|_\infty \leq 1\}, \tag{94}$$

we have that $P_t f$ is continuous in the topology of strong convergence since (94) holds for any $f \in B(\mathcal{X})$ with $\|f\|_\infty \leq 1$, which can be extend to all $f \in \mathcal{B}_b(\mathcal{X})$ (see Remark 5.35).

Thus, P_t is strong Feller and by the semigroup property, it is true for any $s \geq t$ because for all $\varepsilon > 0$,

$$P_{t+\varepsilon} f = P_t(P_\varepsilon f) \in C_b(\mathcal{X}),$$

since $P_\varepsilon f \in \mathcal{B}_b(\mathcal{X})$ for any $f \in \mathcal{B}_b(\mathcal{X})$.

Obviously, it will not be the case in every situation and a first example of why the asymptotic strong Feller property is weaker than the strong Feller one may be found later, in Example 6.28.

In fact, this property of being continuous in the topology of Total variation is very useful and powerful. As stated by Hairer in [Hai09], this characterization of continuity is known as the ultra Feller property, see Definition 5.34. And as we used it many times, ultra Feller semigroups are strong Feller. If the reverse is not always true, it is still possible to link strong Feller and ultra Feller properties.

Theorem 6.21 ([Hai09], Theorem 1.6.6)**.** *Let P and Q be two strong Feller Markov kernels over the same Polish space $\mathcal{X}$. Then, PQ is ultra Feller.*

We can now use this property in order to show that any strong Feller semigroup on a Polish space is also asymptotically strong Feller, so that the second property is weaker than the first one.

Proposition 6.22 ([BH22], Proposition 5.26)**.** *Let's suppose that $(P_t)_{t\geq 0}$ is a strong Feller Markov semigroup on a Polish space $\mathcal{X}$, with metric d^*. Then, the asymptotic strong Feller property holds for $(P_t)_{t\geq 0}$*

Proof. Since P_t is strong Feller, then $P_{2t} = P_t P_t$ is ultra Feller, i.e. continuous in the topology of the Total variation, by Theorem 6.21. The continuation of the proof follows the one in [BH22].

Let's construct a sequence $(d_n)_{n\geq 0}$ such that

$$d_n(x, y) = 1 \wedge (n d^*(x, y)), \quad \forall n \geq 0,\ \forall x, y \in \mathcal{X}.$$

Then, $(d_n)_{n\geq 0}$ is a nondecreasing sequence of continuous metrics such that

$$\lim_{n\to\infty} d_n(x, y) = d(x, y) = \begin{cases} 1, & \text{if } x \neq y, \\ 0, & \text{else.} \end{cases}$$

So $(d_n)_{n\geq 0}$ is a totally separating system of pseudo-metrics. Let $x \in \mathcal{X}$ be fixed, and let's define the sequence

$$f_n(y) = \|\delta_x P_t^2 - \delta_y P_t^2\|_{d_n},$$

where $\delta_x(\cdot)$ is the Dirac measure. The sequence $(f_n)_{n\geq 0}$ is nondecreasing, and dominated by

$$f(y) = \|\delta_x P_t^2 - \delta_y P_t^2\|_d.$$

Let $U \in \mathcal{U}_x$ be an open neighborhood of x. Then,

$$\limsup_{n\to\infty} \sup_{y\in U} f_n(y) \leq \sup_{y\in U} \limsup_{n\to\infty} f_n(y) \leq \sup_{y\in U} f(y).$$

The first inequality follows the fact that $(f_n)_{n\geq 0}$ is nondecreasing, the second one that $(f_n)_{n\geq 0}$ is dominated by f. Since P_t^2 is ultra Feller, i.e. $x \mapsto P_t^2(x,\cdot)$ is continuous in the sense of the Total variation norm, so for any $\varepsilon > 0$, there exists an open neighborhood U of x such that

$$\sup_{y\in U} f(y) \leq \varepsilon,$$

which implies

$$\inf_{U\in\mathcal{U}_x} \limsup_{n\to\infty} \sup_{y\in U} \|\delta_x P_{t_n} - \delta_y P_{t_n}\|_{d_n} \leq \varepsilon,$$

with $t_n = 2t$, $\forall n \geq 0$. Since $\delta_x P_t = P_t(x,\cdot)$ and every previous steps are verified for any $x \in \mathcal{X}$, then $(P_t)_{t\geq 0}$ is asymptotically strong Feller. □

6.2.1 Proof of Theorem 6.15

We want to reuse the links between the strong Feller property and its asymptotic version: thanks to Corollary 6.19, this proof is very similar to the corresponding one for the strong Feller property.

For every measurable set A, every $t > 0$, and every pseudo-metric d on $\mathcal{X}$ with $d \leq 1$, we claim that

$$\|\mu - \nu\|_d \leq 1 - \min\{\mu(A), \nu(A)\} \left(1 - \max_{y,z\in A} \|P_t(y,\cdot) - P_t(z,\cdot)\|_d\right). \tag{95}$$

By definition of the asymptotic strong Feller property, there exists $N > 0$, a sequence of totally separating pseudo-metrics $\{d_n\}_{n\geq 0}$, and an open set U containing x such that

$$\|P_{t_n}(y,\cdot) - P_{t_n}(z,\cdot)\|_{d_n} \leq \frac{1}{2}, \quad \forall n > N,\ \forall y,z \in U.$$

Let's assume by contradiction that there exists $x \in \mathrm{supp}(\mu)\cap\mathrm{supp}(\nu)$. By definition of totally separating pseudo-metrics, $d_n \leq 1$ for all $n \geq 0$ so that we can use (95) with $\alpha = \min(\mu(U), \nu(U)) > 0$ where $U \in \mathcal{U}_x$, $d = d_n$, and $t = t_n$, which implies

$$\|\mu - \nu\|_{d_n} \leq 1 - \frac{\alpha}{2}, \quad \forall n > N.$$

Since the inequality is independent of rank n, by Corollary 6.19, it implies

$$\|\mu - \nu\|_{TV} \leq 1 - \frac{\alpha}{2},$$

and in particular $\|\mu - \nu\|_{TV} < 1$. Hence, μ and ν are not mutually singular, which is in contradiction since they are distinct ergodic invariant probability measures (see e.g. [BH22], Proposition 4.29).

To prove the claim, let $\alpha := \min\{\mu(A), \nu(A)\}$ and we suppose $0 < \alpha < 1$: if $\alpha = 0$ or $\alpha = 1$, the assumption $d \leq 1$ automatically shows that (95) holds.

In particular $\mu(A), \nu(A) \geq \alpha > 0$: for any measurable set B, let

$$\mu_A(B) = \frac{\mu(A \cap B)}{\mu(A)}, \quad \nu_A(B) = \frac{\nu(A \cap B)}{\nu(A)},$$

so that μ_A, ν_A are probability measures such that $\mu_A(A) = \nu_A(A) = 1$. The following is true for both μ_A and ν_A: we are looking for $\tilde{\mu}$ (respectively $\tilde{\nu}$) such that $\mu = (1-\alpha)\tilde{\mu} + \alpha\mu_A$, so

$$\tilde{\mu} = \frac{\mu - \alpha\mu_A}{1-\alpha},$$

which is a probability measure with above assumptions, and similarly for $\tilde{\nu} = \frac{\nu - \alpha\nu_A}{1-\alpha}$. Since μ, ν are invariant, then

$$\begin{aligned}
\|\mu - \nu\|_d &= \|P_t\mu - P_t\nu\|_d \\
&\leq (1-\alpha)\|P_t\tilde{\mu} - P_t\tilde{\nu}\|_d + \alpha\|P_t\mu_A - P_t\nu_A\|_d \\
&\leq (1-\alpha) + \alpha \int_A \int_A \|P_t(y,\cdot) - P_t(z,\cdot)\|_d \, \mu_A(dy)\nu_A(dz) \\
&\leq 1 - \alpha\left(1 - \max_{y,z\in A} \|P_t(y,\cdot) - P_t(z,\cdot)\|_d\right),
\end{aligned}$$

for all $t \geq 0$, which proves (95). Here, the first inequality follows triangle inequality together with the forms of μ, ν using $\tilde{\mu}, \mu_A, \tilde{\nu}, \nu_A$, the second one from $d \leq 1$, and the last one since $\mu_A(A) = \nu_A(A) = 1$. □

6.3 On unique ergodicity for asymptotically strong Feller Markov semigroup

As for the strong Feller property, its asymptotic version has direct consequence for the unique ergodicity of the associated Markov semigroup. This paragraph is based on the summary made in [BH22] stated for discrete-time Markov semigroup. We start by a complementary result to Theorem 6.15:

Lemma 6.23 ([BH22], Theorem 5.32)**.** *Let (M, d) be a Polish space. If P is asymptotically strong Feller, then two distinct ergodic probability measures have disjoint support.*

For a Markov kernel P, let $\mathcal{P}_{\text{erg}}(M)$ be the set of ergodic probability measures for P. In addition to Theorem 6.15 and Lemma 6.23:

Corollary 6.24 ([BH22], Corollary 5.41)**.** *Let M be a Polish space and let P be asymptotically strong Feller at $x \in M$. Then there exists a neighborhood U of x and an ergodic measure ν such that $\pi(U) = 0$ for every $\pi \in \mathcal{P}_{\text{erg}}(M)\backslash\{\nu\}$.*

It is well-known that given a strong Feller Markov semigroup P on M supposed to be a connected space, if we suppose that P has an invariant probability measure having full support, then P is uniquely ergodic (see e.g. [BH22], Proposition 5.18 ***(iii)***). It has a counterpart for the asymptotic strong Feller property:

Proposition 6.25 ([BH22], Proposition 5.42)**.** *Let M be a Polish space and P be asymptotically strong Feller. Then:*

(i) *The set $\mathcal{P}_{\text{erg}}(M)$ is countable, and for every $P-$invariant probability measure μ, one has*

$$\mu(\cdot) = \sum_{\nu\in\mathcal{P}_{\text{erg}}(M)} \nu(\cdot)\mu\left(X(\nu)\right),$$

where $X(\nu) = \{x \in M : Q\left(x, \text{supp}(\nu)\right) = 1\}$ and Q is the Markov kernel from the ergodic decomposition, see Proposition 1.33.

(ii) *If P has an invariant probability measure having full support, then*

$$M = \cup_{\nu\in\mathcal{P}_{\text{erg}}(M)}\text{supp}(\nu),$$

where the supports supp(ν) *are pairwise disjoint for distinct ν.*

(iii) *If P has an invariant probability measure μ having full support and if M is connected, then $\mathcal{P}_{\text{erg}}(M)$ is either countably infinite, or $\mathcal{P}_{\text{erg}}(M) = \{\mu\}$.*

(iv) *If P has an invariant probability measure μ having full support and if, for every $\varepsilon > 0$, there exists a connected compact set $K \subset M$ such that $\mu(K) > 1 - \varepsilon$, then $\mathcal{P}_{\text{erg}}(M) = \{\mu\}$.*

Remark 6.26. However, those conclusions remain true for a continuous-time Markov semigroup $(P_t)_{t\geq 0}$ on a Polish space M. There are two equivalent way to extend the discrete-time statements for the continuous-time case:

(***i***) Since above results are a consequence of the original Theorem 6.15 in [HM06], we can adapt them by following the continuous-time proof of Hairer and Mattingly's theorem.

(***ii***) Alternatively, let's fix a step $\Delta > 0$ so that we consider the Δ–skeleton $Q := P_\Delta$: indeed, given an original continuous-time Markov process $(X_t)_{t\geq 0}$, we consider the discrete-time chain $(X_{n\Delta})_{n\geq 0}$ whose Markov semigroup is Q.

Standard result yield

$$\mathcal{P}_{\text{erg}}(M)_Q = \mathcal{P}_{\text{erg}}(M)_{(P_t)_{t\geq 0}},$$

where $\mathcal{P}_{\text{erg}}(M)_Q$ (respectively $\mathcal{P}_{\text{erg}}(M)_{(P_t)_{t\geq 0}}$) is the set of ergodic probability measures with respect to Q (respectively $(P_t)_{t\geq 0}$) (see e.g. [Kin63] or [TT79]). Moreover, if $(P_t)_{t\geq 0}$ is asymptotically strong Feller at some point, so then so does Q. Hence one may apply the above discrete-time results stated above to Q and conclude that the same hold for $(P_t)_{t\geq 0}$.

6.4 Examples of the (asymptotic) strong Feller property

Let's dive into some basic examples to better understand the difference between strong Feller and asymptotic strong Feller properties. In particular, we use criterion (85) (respectively (92)) to show that the strong Feller (respectively asymptotic strong Feller) property holds, by direct calculus.

Example 6.27. Consider the following SDE on $\mathbb{R}$, given by

$$dx_t \quad = -x_t\text{d}t + \text{d}B_t,$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion on $\mathbb{R}$, defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. The Markov semigroup $(P_t)_{t\geq 0}$ associated to the Markov process $(x_t)_{t\geq 0}$ is defined as

$$P_t f(x) = \mathbb{E}\left(f(x_t) \mid x_0 = x\right), \ \forall f \in B_b(\mathbb{R}), \ \forall x \in \mathbb{R}.$$

By standard theorems on continuity under the integral sign, we can see that $(P_t)_{t\geq 0}$ has the strong Feller property. First of all, we can explicit $(x_t)_{t\geq 0}$ by using a variation of the constant method on its deterministic part, which means

$$x_t = z_t e^{-t}, \quad z_0 = x_0,$$

where the process $(z_t)_{t\geq 0}$ is a function of t and $(B_t)_{t\geq 0}$ to be defined. By Itô's product rule,

$$dx_t = e^{-t} dz_t - x_t \mathrm{d}t = -x_t \mathrm{d}t + \mathrm{d}B_t.$$

Finally, we have

$$e^{-t} dz_t = \mathrm{d}B_t \Rightarrow z_t = z_0 + \int_0^t e^s dB_s \Rightarrow x_t = e^{-t}\left(x_0 + \int_0^t e^s dB_s\right).$$

Then, for $f \in \mathcal{B}_b(\mathbb{R})$, we have

$$P_t f(x) = \mathbb{E}\left(f(X_t) \mid X_0 = x\right) = \mathbb{E}\left[f\left(e^{-t}\left(x + \int_0^t e^s dB_s\right)\right)\right].$$

Since the probability density function Φ of a random variable $\mathrm{U} \sim \mathcal{N}(0,1)$ is

$$\Phi(\mathrm{U}) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{\mathrm{U}^2}{2}\right)},$$

we can link x_t and U. Indeed, using Itô isometry, the variance of the random variable x_t for $t \geq 0$ fixed is given by

$$\int_0^t \left(e^{-t} e^s\right)^2 ds = e^{-2t}\left(\frac{e^{2t}-1}{2}\right) = \frac{1-e^{-2t}}{2},$$

and since $\mathbb{E}_{\mathbb{P}}(x_t) = x_0 e^{-t}$, then we can write

$$\begin{aligned} P_t f(x) &= \int_{-\infty}^{+\infty} f\left(e^{-t}x + \sqrt{\frac{1-e^{-2t}}{2}}\mathrm{U}\right)\Phi(\mathrm{U}) d\mathrm{U} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f\left(e^{-t}x + \sqrt{\frac{1-e^{-2t}}{2}}\mathrm{U}\right) e^{\left(-\frac{\mathrm{U}^2}{2}\right)} d\mathrm{U}. \end{aligned}$$

Consider the following substitution,

$$V = e^{-t}x + \sqrt{\frac{1-e^{-2t}}{2}}\mathrm{U},\ dV = \sqrt{\frac{1-e^{-2t}}{2}} d\mathrm{U}.$$

Then, the previous integral rewrites as

$$\begin{aligned} P_t f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f\left(e^{-t}x + \sqrt{\frac{1-e^{-2t}}{2}}\mathrm{U}\right) e^{\left(-\frac{\mathrm{U}^2}{2}\right)} d\mathrm{U} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(V) e^{-\frac{(V-e^{-t}x)^2}{1-e^{-2t}}} \sqrt{\frac{2}{1-e^{-2t}}} d\mathrm{V} \end{aligned}$$

Under this form, and for $t > 0$ fixed, we remark that $P_t f$ is a function defined as

$$x \mapsto P_t f(x) = \frac{1}{\sqrt{(1-e^{-2t})\pi}} \int_{-\infty}^{+\infty} g(x, V) dV,$$

where $g(x, V) = f(V) e^{-\frac{(V-e^{-t}x)^2}{1-e^{-2t}}}$. In order to apply standard continuity under integral sign theorems, let's prove that (***i***) *(Claim 1) For almost all $x \in \mathbb{R}$, the function $V \mapsto g(x, \mathrm{V})$ is measurable.*

It follows directly from $V \mapsto f(V)$ being a bounded, measurable function and $V \mapsto e^{-\frac{(V-e^{-t}x)^2}{1-e^{-2t}}}$ being continuous.

(***ii***) *(Claim 2) For almost all V, the function $x \mapsto g(x, V)$ is continuous.*

For t, V fixed, the function $x \mapsto f(V)e^{-\frac{(V-e^{-t}x)^2}{1-e^{-2t}}}$ is clearly continuous since f is bounded and $x \mapsto e^{-x^2}$ is continuous.

(***iii***) *(Claim 3) There exists $h : \mathbb{R} \to \mathbb{R}$ an integrable function such that $\forall (x, V) \in \mathbb{R} \times \mathbb{R}$, we have $|g(x, V)| \leq h(V)$.*

We verify the claim for an open neighborhood U of $x_0 \in \mathbb{R}$, let say $U =]x_0 - M, x_0 + M[$ so that $|x| \leq 2(|x_0| + |M|) =: K_M$ on U. Then,

$$-(V - e^{-t}x)^2 = -V^2 + 2Ve^{-t}x - e^{-2t}x^2 \leq -V^2 + 2K_M|V|.$$

Remark that if $|V| \geq 2e^{-t}x$, then $-V^2 + 2K_M|V| \leq -\frac{V^2}{2}$ and $-(V - e^{-t}x)^2 \leq -\frac{V^2}{2}\left(1 + 4K_M^2\right)$, hence

$$|g(x, V)| \leq h(V) := \sup_{x\in\mathbb{R}} |f(x)| e^{\frac{-V^2(1+4K_M^2)}{2(1-e^{-2t})}},$$

which is integrable since $e^{-\frac{V^2}{2}}$ is.

Finally, by continuity under integral sign, $x \mapsto P_t f(x)$ is continuous at any $x_0 \in \mathbb{R}$, $\forall t > 0$, $\forall f \in \mathcal{B}_b(\mathbb{R})$, which prove that $(P_t)_{t\geq 0}$ is strong Feller.

We can extend this example as proposed by Hairer and Mattingly in [HM06]. It will show the power of the criterion 92 in comparison with the strong Feller one.

Example 6.28 ([HM06], Example 3.13)**.** Consider the following SDE

$$\begin{cases} dx_t = -x_t \mathrm{d}t + \mathrm{d}B_t, \\ dy_t = -y_t \mathrm{d}t, \end{cases}$$

defined on $\mathbb{R}^2$ and where $(B_t)_{t\geq 0}$ is a standard Brownian motion on $\mathbb{R}$, defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

One can see that the Markov semigroup associated $(P_t)_{t\geq 0}$ does not have strong Feller property. Let $\varphi(x, y) = \mathrm{sgn}(y)$. It is clear that $P_t\varphi(x, y) = \varphi(x, y) = \mathrm{sgn}(y)$ for any $(x, y) \in \mathbb{R}^2$ and $t \in [0, \infty)$, by definition of $(P_t)_{t\geq 0}$. In other words, $P_t\varphi = \varphi$, $\forall t \in [0, \infty)$.

It follows directly that since φ is bounded but not continuous, same conclusion holds for $P_t\varphi$, which implies that the corresponding Markov semigroup does not have the strong Feller property.

To show the asymptotic strong Feller property, let $\xi = (u_0, v_0) \in \mathbb{R}^2$ with $||\xi|| = 1$, and let $\varphi : \mathbb{R}^2 \to \mathbb{R}$ be a differentiable function. We define the linearized flow starting from ξ as $(u_t, v_t)_{t\geq 0}$, in other words

$$\begin{cases} du_t = -u_t \mathrm{d}t, \\ dv_t = -v_t \mathrm{d}t. \end{cases}$$

It implies that

$$\begin{aligned} |\nabla P_t\varphi(x, y) \cdot \xi| &= |\mathbb{E}\left[\nabla\varphi(x_t, y_t) \cdot (u_t, v_t)|(x_0, y_0) = (x, y)\right]| \\ &\leq ||\nabla\varphi||_\infty \mathbb{E}\left[|(u_t, v_t)|\right] \\ &\leq ||\nabla\varphi||_\infty e^{-t}. \end{aligned}$$

Finally, we can find a nondecreasing positive sequence $(t_n)_{n\geq 1}$ and $(\delta_n)_{n\geq 1}$ converging to zero and defined as $\delta_n = e^{-t_n}$ such that inequality (92) holds, in other words $(P_t)_{t\geq 0}$ is asymptotically strong Feller.

We can also consider this example by a complete computing approach: since $(y_t)_{t\geq 0}$ is deterministic, we can be expressed it as

$$y_t = y_0 e^{-t}, \quad \forall t \geq 0.$$

The form of $(x_t)_{t\geq 0}$ follows from Example 6.27, which is

$$x_t = e^{-t}\left(x_0 + \int_0^t e^s dB_s\right), \quad \forall t \geq 0.$$

For any bounded, measurable function f, since $P_t f(x,y) = \mathbb{E}\left[f(x_t, y_t) \mid (x_0, y_0) = (x,y)\right]$ and by definition of $(B_t)_{t\geq 0}$, we can explicit $P_t f(x,y)$ as

$$P_t f(x,y) = \int_{-\infty}^{+\infty} f\left(e^{-t}x + e^{-t}\int_0^t e^s \cdot \mathrm{U}\, ds,\ ye^{-t}\right)\Phi(\mathrm{U}) d\mathrm{U},$$

where Φ is the probability density function of a $\mathcal{N}(0,1)-$distribution, $\Phi(\mathrm{U}) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{\mathrm{U}^2}{2}\right)}$.
We can then write

$$P_t f(x,y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f\left(e^{-t}x + \sqrt{\frac{1-e^{-2t}}{2}}\mathrm{U},\ ye^{-t}\right) e^{\left(-\frac{\mathrm{U}^2}{2}\right)} d\mathrm{U}.$$

By applying the gradient in term of x and y, we see that a factor e^{-t} will appear in front of the integral whose absolute value can be also bounded by $||\nabla f||_\infty$. We finally find

$$\begin{aligned} |\nabla P_t f(x,y)| &\leq e^{-t} ||\nabla f||_\infty \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{\mathrm{U}^2}{2}\right)} d\mathrm{U} \\ &\leq e^{-t} ||\nabla f||_\infty. \end{aligned}$$

Let $(t_n)_{n\geq 0}$ be the positive nondecreasing sequence defined as $t_n = n$, $\forall n \geq 0$, and let $(\delta_n)_{n\geq 0}$ be defined as $\delta_n = e^{-t_n}$ which is converging to 0 as $n \to \infty$. We can conclude that

$$\begin{aligned} |\nabla P_{t_n} f(x,y)| &\leq e^{-t_n} ||\nabla f||_\infty \\ &\leq \delta_n ||\nabla f||_\infty \\ &\leq ||f||_\infty + \delta_n ||\nabla f||_\infty, \end{aligned}$$

which implies the asymptotic strong Feller property since (92) holds with $C := 1$ the constant function. To summarize, we have constructed a Markov semigroup $(P_t)_{t\geq 0}$ which lacks the strong Feller property but is asymptotically strong Feller.

We can finally see a new extension of this example by adding some randomness on the second coordinate.

Example 6.29. Consider the following SDE on $\mathbb{R}^2$,

$$\begin{cases} dx_t = -x_t \mathrm{d}t + \mathrm{d}B_t, \\ dy_t = -y_t \mathrm{d}t + d\tilde{B}_t, \end{cases}$$

where $(B_t)_{t\geq 0}, (\tilde{B}_t)_{t\geq 0}$ are two standard Brownian motions on $\mathbb{R}$, defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. As in Example 6.27, we can explicit $(x_t)_{t\geq 0}$ and $(y_t)_{t\geq 0}$ by

$$x_t = e^{-t}\left(x_0 + \int_0^t e^s dB_s\right), \quad y_t = e^{-t}\left(y_0 + \int_0^t e^s d\tilde{B}_s\right), \quad \forall t \geq 0,$$

and use a variable substitution the first and second coordinates to compute

$$\begin{aligned} P_t f(x,y) &= \mathbb{E}\left(f(x_t,y_t) \mid (x_0,y_0)=(x,y)\right) \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f\left(e^{-t}x + \sqrt{\frac{1-e^{-2t}}{2}}\mathrm{U}, e^{-t}y + \sqrt{\frac{1-e^{-2t}}{2}}\tilde{\mathrm{U}}\right)\Phi(U)\Phi(\tilde{U})dUd\tilde{U}. \end{aligned}$$

By using the same variables substitutions as in Examples 6.27 and 6.28, which means

$$V = e^{-t}x + \sqrt{\frac{1-e^{-2t}}{2}}\mathrm{U},\ dV = \sqrt{\frac{1-e^{-2t}}{2}}d\mathrm{U},$$

$$\tilde{V} = e^{-t}y + \sqrt{\frac{1-e^{-2t}}{2}}\tilde{\mathrm{U}},\ d\tilde{V} = \sqrt{\frac{1-e^{-2t}}{2}}d\tilde{\mathrm{U}},$$

we can see $P_t f(x,y)$ as a function

$$(x,y) \mapsto \frac{2}{1-e^{-2t}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(V,\tilde{V})\Phi\left(\frac{\sqrt{2}(V-e^{-t}x)}{\sqrt{1-e^{-2t}}}\right)\Phi\left(\frac{\sqrt{2}(\tilde{V}-e^{-t}y)}{\sqrt{1-e^{-2t}}}\right)dVd\tilde{V}.$$

For $t > 0$ fixed, we can use the same standard continuity under integral sign theorem, as in Example 6.27, to the function

$$(x,y) \mapsto \frac{1}{\pi(1-e^{-2t})}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y,V,\tilde{V})dVd\tilde{V},$$

where

$$g(x,y,V,\tilde{V}) = f(V,\tilde{V})e^{\frac{-\left(\sqrt{2}(V-e^{-t}x)\right)^2}{2(1-e^{-2t})}} e^{\frac{-\left(\sqrt{2}(\tilde{V}-e^{-t}y)\right)^2}{2(1-e^{-2t})}}.$$

It will show that $P_t f$ is continuous, $\forall f \in \mathcal{B}_b(\mathbb{R}\times\mathbb{R})$, and so that $(P_t)_{t\geq 0}$ is strong Feller.

Chapter 7

Log–Harnack inequalities and links with (asymptotic) strong Feller property

The notion of *Harnack inequalities* qualifies an inequality on the values of a harmonic function at two points. It was first introduced at the end of the 19th century for harmonic functions on a Euclidean space by Harnack in [Har87]. The use of this kind of inequality is well-known in many different fields of mathematics. Perelman's solution to the Poincaré conjecture, which will lead to a Millennium prize in 2006, is based on a Harnack inequality about the Ricci Flow, more precisely on the work of Hamilton in [Ham93].

Before stating the inequalities of interest, we need a short introduction to the change of measure, involving Girsanov's theorem as well as the Kolmogorov extension theorem.

7.1 Change of measure

The concept of a change of measure in probability theory provides a powerful framework for transforming stochastic processes, particularly in the context of Brownian motion. This subsection delves into the foundational theories and conditions that facilitate the change of measure in a stochastic calculus setting.

To start, let's recall the Radon–Nikodym theorem in a general setting:

Theorem 7.1 ([Obe17], Theorem 3.24)**.** *Let μ, ν be two $\sigma-$finite measures defined on* $(\Omega, \mathcal{F})$ *a measurable space. The following statements are equivalent:*

(i) *The measure ν is absolutely continuous with respect to μ, denoted as $\nu \ll \mu$, which means that for every $A \in \mathcal{F}$, if $\mu(A) = 0$, then it implies that $\nu(A) = 0$.*

(ii) *There exists a measurable function $\rho \in L^1(\Omega, \mathcal{F}, \mu)$ (the space of integrable functions with respect to μ) such that $\rho \geq 0$ and for every $A \in \mathcal{F}$, the measure $\nu(A)$ can be represented as the integral of ρ over A with respect to μ, which is $\nu(A) = \int_A \rho d\mu$.*

Definition 7.2 (**Radon-Nikodym derivative**)**.** *The function ρ in Theorem 7.1**(ii)** is the density of ν with respect to μ and is called the* Radon-Nikodym derivative *of ν with respect to μ. In particular, we write*

$$\rho = \frac{d\nu}{d\mu}.$$

If $\mu \gg \nu$ and $\nu \gg \mu$, we say μ and ν are *equivalent* and denote this as $\mu \sim \nu$.

In particular, we are interested in the case of probability measures on a path space: in other words, we have in addition a filtration which leads to a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. If we suppose that there exists a positive random variable ρ_∞ and $\mathbb{E}_{\mathbb{P}}(\rho_\infty) = 1$, we get a new probability measure

$$d\mathbb{Q} = \rho_\infty d\mathbb{P},$$

such that $\mathbb{P} \sim \mathbb{Q}$, and ρ_∞ is a density defined as

$$\rho_\infty = \frac{d\mathbb{Q}}{d\mathbb{P}}.$$

It follows that the process ρ defined as

$$\rho = (\rho_t)_{t\geq 0} = (\mathbb{E}_{\mathbb{P}}[\rho_\infty \mid \mathcal{F}_t])_{t\geq 0},$$

is a non-negative, uniformly integrable martingale under $\mathbb{P}$.

Although these tools are powerful, the construction of such an equivalent probability measure is nontrivial. This specific choice of measures change is handled by the famous Girsanov's theorem: before that, we recall the notion of Doleans-Dade exponential. It is directly suggested by Itô's formula applied to Brownian motion and to the function $x \mapsto \exp(x)$:

Theorem 7.3 ([Obe17], Theorem 3.5)**.** *Let $X = (X_t)_{t\geq 0}$ be an adapted, continuous semimartingale with $X_0 = 0$, then the process $\mathcal{E}(X) = (\mathcal{E}(X)_t)_{t\geq 0}$ called the* Doleans-Dade exponential of *X and defined as*

$$\mathcal{E}(X)_t := \exp\left(X_t - \frac{1}{2}\langle X\rangle_t\right), \quad \forall t \geq 0,$$

is a continuous semimartingale. Moreover, $\mathcal{E}(X)$ is the unique solution of the SDE

$$dZ_t = Z_t dX_t, \quad Z_0 = 1.$$

Then, Girsanov's theorem states the following:

Theorem 7.4 ([Obe17], Theorem 3.29)**.** *Let X be a continuous, local martingale under $\mathbb{P}$, and $K \in L^2(X)$*[1]*. We suppose that the process ρ defined as*

$$\rho = (\rho_t)_{t\geq 0} = \left(\mathcal{E}\left(\int_0^t K_s dX_s\right)\right)_{t\geq 0},$$

is a continuous, uniformly integrable martingale. Then:

(i) ρ_∞ exists, $\mathbb{E}_{\mathbb{P}}(\rho_\infty) = 1$, and $d\mathbb{Q} = \rho_\infty d\mathbb{P}$ satisfies $\mathbb{Q} \sim \mathbb{P}$.

(ii) There exists Y a local martingale under $\mathbb{Q}$ such that the process X can be written as

$$X_t = Y_t + \int_0^t K_s d\langle X\rangle_s, \quad \forall t \geq 0,\ \mathbb{P} - \textit{almost surely}.$$

Given X is a Brownian motion under $\mathbb{P}$, a modified version of Girsanov's theorem, called Cameron-Martin theorem, tells us that the transformed process is also a Brownian motion under the new measure:

Theorem 7.5 ([Obe17], Theorem 3.30)**.** *Let $B := (B_t)_{t\geq 0}$ be a Brownian motion under $\mathbb{P}$, and $K \in L^2(B)$. We suppose that the process ρ defined as*

$$\rho = (\rho_t)_{t\geq 0} = \left(\mathcal{E}\left(\int_0^t K_s dB_s\right)\right)_{t\geq 0},$$

is a continuous, uniformly integrable martingale. Then:

(i) ρ_∞ exists, $\mathbb{E}_{\mathbb{P}}(\rho_\infty) = 1$, and $d\mathbb{Q} = \rho_\infty d\mathbb{P}$ satisfies $\mathbb{Q} \sim \mathbb{P}$.

[1] $L^2(X)$ refers to the space of processes that are square integrable with respect to the quadratic variation of the local martingale X, i.e. $K \in L^2(X)$ if $\mathbb{E}_{\mathbb{Q}}\left[\int_0^T K_s^2 d\langle X\rangle_s\right) < \infty$, for every $T > 0$.

(ii) There exists $\tilde{B} := (\tilde{B}_t)_{t\geq 0}$ a Brownian motion under $\mathbb{Q}$ such that the Brownian motion B can be written as

$$B_t = \tilde{B}_t + \int_0^t K_s ds, \quad \forall t \geq 0,\ \mathbb{P} - \textit{almost surely}.$$

Obviously, the question about the property of being a martingale, and not just a local martingale, for the process ρ is tough. But since ρ is a stochastic exponential, some tools will be useful. Thus, we introduce the *Novikov's condition*:

Theorem 7.6 ([Obe17], Proposition 3.10). *Let $M \in \mathcal{M}_{c,loc}$ the space of continuous, local martingales. Then, if*

$$\mathbb{E}_{\mathbb{P}}\left(\exp\left(\frac{1}{2}\langle M\rangle_\infty\right)\right) < \infty, \tag{96}$$

then $\mathcal{E}(M)$ is a continuous, uniformly integrable martingale.

Another weaker condition is given by the so-called *Kazamaki's condition*: given M a continuous uniformly integrable martingale and

$$\mathbb{E}_{\mathbb{P}}\left[\exp\left(\frac{1}{2}M_\infty\right)\right] < \infty,$$

then $\mathcal{E}(M)$ is a continuous and uniformly integrable martingale. In particular, Novikov's condition implies Kazamaki's condition (see e.g. [Gal16], Theorem 5.23).

A similar question about the uniform integrability arises naturally, which can be solved with the *De la Vallée Poussin Theorem*: let Φ denote the class of nondecreasing, measurable functions $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ such that

$$\lim_{x\to+\infty} \frac{\varphi(x)}{x} = +\infty.$$

Theorem 7.7 ([HR11], Theorem 1.1). *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(X_i)_{i\in I}$ be a sequence of random variables. Then, $(X_i)_{i\in I}$ is uniformly integrable if and only if there exists $\varphi \in \Phi$ such that*

$$\sup_{i\in I} \mathbb{E}_{\mathbb{Q}}\left[\varphi(|X_i|)\right) < \infty.$$

Typical choices include $\varphi(x) = x\log(x)$ or $\varphi(x) = x^r$ for $r > 1$.

Remark 7.8. As stated in [HR11], original De la Vallée Poussin criterion (see e.g. [Mey66], Theorem T22) was restricted to convex monotone functions φ with $\varphi(0) = 0$ which can be weakened to above conditions.

Example 7.9. Let's study an application of Girsanov's theorem, especially in the setting of SDE. Let

$$dX_t = b(t, X_t)\mathrm{d}t + \mathrm{d}B_t,$$

where b is a bounded measurable function on $\mathbb{R}_+ \times \mathbb{R}$ and $B = (B_t)_{t\geq 0}$ is an $\mathcal{F}_t$−adapted Brownian motion under a probability measure $\mathbb{P}$. Also, we assume that there exists a function $g \in L^2(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), \mathrm{d}t)$ such that $|b(t, x)| \leq g(t)$ on $\mathbb{R}_+ \times \mathbb{R}$.

Let $(L_t)_{t\geq 0}$ be the continuous local martingale defined as

$$L_t = -\int_0^t b(s, X_s)dB_s.$$

Since $b(t,x)$ is bounded by $g(t)$ which is square integrable, Novikov's condition (96) holds and the associated stochastic exponential is a uniformly integrable martingale, i.e. $D = (D_t)_{t\geq 0}$, defined as

$$D_t = \mathcal{E}(L)_t = \exp\left(-\int_0^t b(s,X_s)dB_s - \frac{1}{2}\int_0^t b(s,X_s)^2 ds\right),$$

is a uniformly integrable martingale. We define $d\mathbb{Q}|_{\mathcal{F}_t} = D_t d\mathbb{P}|_{\mathcal{F}_t}$ for all $t \geq 0$: by Theorem 7.5, the process $\beta = (\beta_t)_{t\geq 0}$ defined as

$$\beta_t = B_t + \int_0^t b(s,X_s)ds,$$

is a $\mathcal{F}_t$−adapted Brownian motion under the new probability measure $\mathbb{Q}$ whose existence follows from Kolmogorov extension Theorem 7.12.

In other words, under the probability measure $\mathbb{Q}$, there exists a $\mathcal{F}_t$−adapted Brownian motion $(\beta_t)_{t\geq 0}$ such that the process $(X_t)_{t\geq 0} = (\beta_t)_{t\geq 0}$ solves the same SDE with drift under $\mathbb{Q}$.

7.1.1 Kolmogorov extension Theorem

Having explored the critical role of Girsanov's theorem in facilitating the change of measure and the transformation of Brownian motion within modified probability spaces, we now shift our focus to another foundational aspect of probability theory: the *Kolmogorov extension Theorem*.

While Girsanov's theorem provides a framework for altering the probabilistic characteristics of stochastic processes on a global scale, the Kolmogorov extension Theorem addresses a different but equally fundamental problem. It ensures the existence and uniqueness of a comprehensive probability measure based on consistent finite-dimensional marginals.

This theorem also plays a crucial role to solve the problem of the existence of a Markov process based on a given semigroup.

Let X be a set, and $(\Sigma_t)_{t\in T}$ be an increasing net[1] of σ-algebras on X. That is, if $t \leq t'$, then $\Sigma_t \subset \Sigma_{t'}$. The index set T may be infinite and generally represents a set of times period. For each finite subset F of T, let μ_F be a probability measure on (X, Σ_F).

Definition 7.10 (**Kolmogorov consistency**). *For each $t \in T$, let μ_t be a probability measure on Σ_t. The net $(\Sigma_t, \mu_t)_{t\in T}$ is said to be* Kolmogorov consistent *if*

$$\Sigma_t \subset \Sigma_{t'} \;\Rightarrow\; \mu_{t'}|_{\Sigma_t} = \mu_t,$$

where $\mu_{t'}|_{\Sigma_t}$ is the restriction of $\mu_{t'}$ to the σ-subalgebra Σ_t of $\Sigma_{t'}$.

Let $\mathcal{A} = \cup_{t\in T}\Sigma_t$.

Definition 7.11 (**Kolmogorov extension**). *A* Kolmogorov extension *of the net $(\mu_t)_{t\in T}$ is a probability measure μ on $\mathcal{A}$ such that $\mu|_{\Sigma_t} = \mu_t$, $\forall t \in T$.*

There are several versions of the Kolmogorov extension Theorem, here is a version focused on Polish spaces which which matters for our purposes:

Theorem 7.12 ([AB06], Corollary 15.27). *Let $(X_t)_{t\in T}$ be a family of Polish spaces equipped with their Borel σ-algebras. For each finite subset F of T, let μ_F be a Borel probability measure on $X_F = \Pi_{t\in F}X_t$ with its product σ-algebra Σ_F. Assume the distributions Σ_F are Kolmogorov consistent. Then, there is a unique Kolmogorov extension on the infinite product σ-algebra $\Sigma_T = \bigotimes_{t\in T}\Sigma_t$.*

A more general result follows from the *Generalized Kolmogorov extension Theorem* (see e.g. [AB06], Theorem 15.26).

[1] In the sense of a generalization of the notion of sequences, where the indexes are not necessary $\mathbb{N}$ but any directed set

7.2 A log–Harnack inequality to prove strong Feller property

The following part is based on [Wan13], which brings together main articles and results about Harnack inequalities, and their applications in the case of semigroups associated to Stochastic (or Partial, or Functional) Differential Equations, in possibly infinite-dimensional space.

The dimension-free Harnack inequality was first introduced in [Wan97], in more details in [ATW06]. Given a Markov semigroup $(P_t)_{t\geq 0}$ on a Polish space $\mathcal{X}$, the Harnack-type inequality considered in [ATW06] is of the form

$$\Phi(P_t f(x)) \leq [P_t\Phi(f)(y)]e^{C(t)\Psi(x,y)}, \quad x, y \in \mathcal{X},\ f \in \mathcal{B}_b^+(\mathcal{X}), \tag{97}$$

for Φ a nonnegative convex function on $[0, \infty)$, Ψ a nonnegative function on $\mathcal{X}^2$, and $C(t) > 0$ is explicitly determined. Here, $\mathcal{B}_b^+(\mathcal{X})$ denotes the set of nonnegative bounded measurable functions on $\mathcal{X}$. Thinking about Jensen's inequality, we may always take $\Psi(x, x) = 0$, $\forall x \in \mathcal{X}$, without loss of generality.

Here are two typical choices for Φ:

(i) *Harnack inequality with power*: Let $\Phi(r) = r^p$ for $p > 1$. Then, (97) rewrites as

$$[P_t f(x)]^p \leq [P_t f^p(y)]e^{C(t)\Psi(x,y)},\ x, y \in \mathcal{X},\ f \in \mathcal{B}_b^+(\mathcal{X}).$$

(ii) *Log–Harnack inequality*: Let $\Phi(r) = e^r$. In this case, we can replace f by $\log f$ so that (97) rewrites as

$$P_t[\log f](x) \leq \log[P_t f(y)] + C(t)\Psi(x, y),\ x, y \in \mathcal{X},\ f \in \mathcal{B}_b^+(\mathcal{X}).$$

This special type of Harnack inequality has been introduced in [Wan10] and [RW10].

We now follow the extensive review from [Wan13]. In particular, this type of Harnack inequalities are stronger than the strong Feller property.

Theorem 7.13 ([Wan13], Theorem 1.4.1)**.** *Let $(P_t)_{t\geq 0}$ be a Markov semigroup on a Polish space $\mathcal{X}$. We suppose $\Phi \in C^1([0, \infty))$ is a convex function satisfying $\Phi'(r) > 1$ for all $r \in [0, \infty)$ and $\lim_{r\to\infty} \Phi(r) = \infty$ such that (97) holds for $T \geq 0$ fixed and $\forall x, y \in \mathcal{X}$, $\forall f \in \mathcal{B}_b^+(\mathcal{X})$.*

(i) *If, in addition, $\lim_{y\to x}\Psi(x, y) = \lim_{y\to x}\Psi(y, x) = 0$, $\forall x \in \mathcal{X}$, then $(P_t)_{t\geq 0}$ is strong Feller at time T.*

(ii) *$(P_t)_{t\geq 0}$ has at most one invariant probability measure.*

(iii) *Let μ_0 be a quasi-invariant probability[1] measure for $(P_t)_{t\geq 0}$. Then, $(P_t)_{t\geq 0}$ possesses a density $\mathbf{p}$ with respect to μ_0.*

(iv) *For all $x, y \in E$, it holds that*

$$\int_E \mathbf{p}(x, \cdot)\Phi^{-1}\left(\frac{\mathbf{p}(x, \cdot)}{\mathbf{p}(y, \cdot)}\right) d\mu_0 \leq \Phi^{-1}\left(e^{\Psi(x,y)}\right),$$

where $\Phi^{-1}(\infty) = \infty$ by convention.

(v) *If $r\Phi^{-1}(r)$ is convex for $r \geq 0$, then the density $\mathbf{p}$ of $(P_t)_{t\geq 0}$ with respect to μ_0 satisfies*

$$\int_E \mathbf{p}(x, \cdot)\mathbf{p}(y, \cdot)d\mu_0 \geq e^{-\Psi(x,y)}, \quad \forall x, y \in E.$$

[1] Let μ be a probability measure on a measurable space $(E, \mathcal{B})$ and P be a linear bounded operator on $\mathcal{B}_b^+(E)$, then μ is called a *quasi-invariant probability* of P if μP is absolutely continuous with respect to μ, in the sense that $\mu P(A) = \mu(P\mathbf{1}_A) = 0$ for each $A \in \mathcal{B}$ such that $\mu(A) = 0$.

(vi) *If $(P_t)_{t\geq 0}$ possesses an invariant probability measure μ, it is absolutely continuous with respect to μ_0, the density* $\mathbf{p}$ *of $(P_t)_{t\geq 0}$ with respect to μ is strictly positive, and*

$$\sup_{f\in\mathcal{B}_b^+(E),\ \mu(\Phi(f))\leq 1} \Phi(P_t f(x)) \leq \frac{1}{\int_E e^{-\Psi(x,y)}\mu(dy)}, \quad \forall x\in E,\ \forall t\geq 0.$$

Here, we only prove the strong Feller argument but the other parts are fully proved in [Wan13].

Proof of (i). For $T>0$ fixed, let $\varepsilon>0$, and $f\in\mathcal{B}_b^+(\mathcal{X})$. If we apply (97) to $1+\varepsilon f$, then

$$\Phi(1+\varepsilon P_T f(x)) \leq P_T\Phi(1+\varepsilon f)(y)e^{C(T)\Psi(x,y)},\ \forall x,y\in\mathcal{X}.$$

By a Taylor expansion of order 1, it follows that

$$\Phi(1)+\varepsilon\Phi'(1)P_T f(x)+o(\varepsilon) \leq \left(\Phi(1)+\varepsilon\Phi'(1)P_T f(y)+o(\varepsilon)\right)e^{C(T)\Psi(x,y)},$$

for small $\varepsilon>0$. Then, letting $y\to x$, we obtain

$$\begin{aligned}\Phi(1)+\varepsilon\Phi'(1)P_T f(x)+o(\varepsilon) &\leq \liminf_{y\to x}\ \left(\Phi(1)+\varepsilon\Phi'(1)P_T f(y)+o(\varepsilon)\right)e^{C(T)\Psi(x,y)}\\ &\leq \Phi(1)+\varepsilon\Phi'(1)\liminf_{y\to x} P_T f(y)+o(\varepsilon),\end{aligned}$$

which holds true since $C(T)\Psi(x,y)\geq 0$, $\forall t\geq 0$, $\forall x,y\in\mathcal{X}$. Now, we obtain

$$\varepsilon\Phi'(1)P_T f(x) \leq \varepsilon\Phi'(1)\liminf_{y\to x} P_T f(y),\ \forall x\in\mathcal{X}.$$

Since $\varepsilon,\Phi'(1)>0$,

$$P_T f(x) \leq \liminf_{y\to x} P_T f(y),\ \forall x\in\mathcal{X}. \tag{98}$$

By the same way, replacing $y\to x$ by $x\to y$, it implies that

$$\begin{aligned}\limsup_{x\to y}\Phi(1)+\varepsilon\Phi'(1)P_T f(x)+o(\varepsilon) &\leq \limsup_{x\to y}\ \left(\Phi(1)+\varepsilon\Phi'(1)P_T f(y)+o(\varepsilon)\right)e^{C(T)\Psi(x,y)}\\ &\leq \Phi(1)+\varepsilon\Phi'(1)P_T f(y)+o(\varepsilon),\end{aligned}$$

which means that

$$\limsup_{y\to x} P_T f(y) \leq P_T f(x),\ \forall x\in\mathcal{X}, \tag{99}$$

and since $\liminf_{y\to x} P_T f(y) \leq \limsup_{y\to x} P_T f(y)$, $\forall x\in\mathcal{X}$, by combining (98) and (99), we conclude that

$$\limsup_{y\to x} P_T f(y) \leq P_T f(x) \leq \liminf_{y\to x} P_T f(y) \leq \limsup_{y\to x} P_T f(y), \quad \forall x\in\mathcal{X}.$$

Naturally, by combining above inequalities, it yields $P_T f$ is continuous since $\lim_{y\to x} P_T f(y)$ exists and equals $P_T f(x)$ so the strong Feller property holds true at any $T>0$. □

7.2.1 Coupling construction

To construct such inequalities, we will use the methodology from [Wan13] and the so-called strategy of coupling, also introduced in [ATW06]. Intuitively, given an initial Markov process $(X_t^x)_{t\geq 0}$ starting at $x\in\mathcal{X}$ on $(\Omega,\mathcal{F},\mathbb{P})$, our goal is to construct a modified version $(Y_t^y)_{t\geq 0}$ starting at a different initial point $y\in\mathcal{X}$ such that:

(i) there exists a time $T>0$ where both processes come together;

(ii) that there exists a probability measure $\mathbb{Q}$, absolutely continuous with respect to $\mathbb{P}$;

(iii) and the distribution of $(X_t^x)_{t\geq 0}$ under $\mathbb{P}$ and the one of $(Y_t^y)_{t\geq 0}$ under $\mathbb{Q}$ coincides.

This construction allows us to derive the strong Feller property for $(P_t)_{t\geq 0}$ at time $T > 0$. More precisely, we define a coupling of two probability measures as follows.

Definition 7.14 (**Coupling of measure**). *Let μ and ν be probability measures on a measurable space Ω. A probability measure ω on $\Omega \times \Omega$ is a* coupling of μ and ν *if its marginals are μ and ν, respectively.*

Now, we can extend the previous definition to the coupling of two random processes through a change of measure, in particular for Markov processes.

Definition 7.15 (**Coupling by change of measure**). *Let $(X_t^x)_{t\geq 0}$ and $(Y_t^y)_{t\geq 0}$ be two $\mathcal{X}$-valued Markov processes defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The joint process $(X_t, Y_t)_{t\geq 0}$ on $\mathcal{X} \times \mathcal{X}$ constitute a* coupling by change of measure for $(X_t^x)_{t\geq 0}$ and $(Y_t^y)_{t\geq 0}$ *if there exists a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ such that the distribution of $(X_t^x)_{t\geq 0}$ under $\mathbb{P}$ is equal in law to the distribution of $(Y_t^y)_{t\geq 0}$ under $\mathbb{Q}$.*

Remark 7.16. **(i)** The condition that the distributions coincide under different measures is analogous to stating that the Markov semigroups of $(X_t^x)_{t\geq 0}$ under $\mathbb{P}$ and $(Y_t^y)_{t\geq 0}$ under $\mathbb{Q}$ are the same. This implies that the transition probabilities of the processes match under their respective probabilities.

(ii) If $\mathbb{Q} = \mathbb{P}$, then the joint process $(X_t^x, Y_t^y)_{t\geq 0}$ on $\mathcal{X} \times \mathcal{X}$ represents a standard coupling of $(X_t^x)_{t\geq 0}$ and $(Y_t^y)_{t\geq 0}$, where no change in the underlying measure is required.

This specific notion of coupling (by change of measure) is particularly useful for deriving our log–Harnack inequalities of interest:

Theorem 7.17 ([Wan13], Theorem 1.1.1)**.** *Let $(X_t^x)_{t\geq 0}$ with $X_0 = x$, and $(Y_t^y)_{t\geq 0}$ with $Y_0 = y$ be two Markov process defined on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. We suppose that $(X_t^x, Y_t^y)_{t\geq 0}$ is a coupling by change of measure of the processes $(X_t^x)_{t\geq 0}$ and $(Y_t^y)_{t\geq 0}$ with changed probability $\mathbb{Q}$ that we suppose absolutely continuous with respect to $\mathbb{P}$, in the sense that $\exists R_t(\cdot)$ a nonnegative $\mathcal{F}_t$-measurable function defined on Ω such that*

$$d\mathbb{Q}|_{\mathcal{F}_t} = R_t d\mathbb{P}|_{\mathcal{F}_t}, \quad \forall t \geq 0.$$

If there exists a fixed $T > 0$ such that $X_T^x = Y_T^y$ almost surely under $\mathbb{Q}$, then

$$\left(P_T \log f\right)(y) \leq \left(\log P_T f\right)(x) + \mathbb{E}_{\mathbb{P}}\left[R_T \log R_T\right], \ \forall f \in \mathcal{B}_b^+(E), f > 0,$$

where we interpret $P_T f(\cdot)$ as

$$P_T f(x) = \mathbb{E}_{\mathbb{P}}\left[f\left(X_T\right) \mid X_0 = x\right], \quad P_T f(y) = \mathbb{E}_{\mathbb{Q}}\left[f\left(Y_T\right) \mid Y_0 = y\right].$$

Before proving this theorem, we need this useful application of the Jensen's inequality to get a good estimate of logarithm gradient:

Lemma 7.18 ([Str00], Lemma 6.45)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and φ be a nonnegative $\mathcal{F}$-measurable function on Ω with expected value* 1*. If $\psi : \Omega \to \mathbb{R}$ is a $\mathcal{F}$-measurable function such that $\psi\varphi$ is integrable, then*

$$\mathbb{E}_{\mathbb{P}}\left[\psi\varphi\right] \leq \mathbb{E}_{\mathbb{P}}\left[\varphi \log \varphi\right] + \log\left(\mathbb{E}_{\mathbb{P}}\left[e^{\psi}\right]\right).$$

Proof. It is sufficient to suppose that φ is positive, and φ, ψ are bounded.

With those assumptions, it follows from Theorem 7.1 that we can define another probability measure $\mathbb{Q}$ determined by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \varphi.$$

Since the function log is concave, we can apply the Jensen's inequality, which yields

$$\begin{aligned} \log\left(\mathbb{E}_{\mathbb{P}}\left[e^{\psi}\right]\right) &= \log\left(\mathbb{E}_{\mathbb{Q}}\left[\varphi^{-1}e^{\psi}\right]\right) \\ &\geq \mathbb{E}_{\mathbb{Q}}\left[\log\left(\varphi^{-1}e^{\psi}\right)\right] \\ &= -\mathbb{E}_{\mathbb{Q}}\left[\log(\varphi)\right] + \mathbb{E}_{\mathbb{Q}}\left[\psi\right] \\ &= -\mathbb{E}_{\mathbb{P}}\left[\varphi\log(\varphi)\right] + \mathbb{E}_{\mathbb{P}}\left[\varphi\psi\right]. \end{aligned}$$

□

Proof of Theorem 7.17. The fact that $d\mathbb{Q}|_{\mathcal{F}_t} = R_t d\mathbb{P}|_{\mathcal{F}_t}$ naturally implies that R_t is a $\mathcal{F}_t$-measurable function with expected value 1.

Let $f \in \mathcal{B}_b^+(E)$ be a positive function. We can then apply Lemma 7.18 with $\varphi = R_T$, $\psi = \log f$, such that

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}\left[R_T \log f\right] &\leq \mathbb{E}_{\mathbb{P}}\left[R_T \log R_T\right] + \log\left(\mathbb{E}_{\mathbb{P}}\left[e^{\log f}\right]\right) \\ \Rightarrow \mathbb{E}_{\mathbb{Q}}\left[\log f\right] &\leq \mathbb{E}_{\mathbb{P}}\left[R_T \log R_T\right] + \log\left(\mathbb{E}_{\mathbb{P}}\left[f\right]\right). \end{aligned}$$

In particular, the inequality holds for $(X_t^x)_{t\geq 0}$ with $X_0 = x$ and at time T, we obtain

$$\mathbb{E}_{\mathbb{Q}}\left[\log f\left(X_T\right)|X_0 = x\right] \leq \mathbb{E}_{\mathbb{P}}\left[R_T \log R_T\right] + \log\left(\mathbb{E}_{\mathbb{P}}\left[f\left(X_T\right)|X_0 = x\right]\right).$$

Since $X_T = Y_T$, $\mathbb{Q}$-almost surely, it follows that

$$\mathbb{E}_{\mathbb{Q}}\left[\log f\left(Y_T\right)|Y_0 = y\right] \leq \mathbb{E}_{\mathbb{P}}\left[R_T \log R_T\right] + \log\left(\mathbb{E}_{\mathbb{P}}\left[f\left(X_T\right)|X_0 = x\right]\right).$$

And by definition of $P_T f(x)$, $P_T f(y)$, we finally have

$$\left(P_T \log f\right)(y) \leq \left(\log P_T f\right)(x) + \mathbb{E}_{\mathbb{P}}\left[R_T \log R_T\right].$$

□

7.2.2 Examples of coupling construction and log–Harnack inequalities

Let's dive into some concrete examples. We start by showing how the strong Feller property can be recovered for a one-dimensional diffusion on $\mathbb{R}$ via the coupling construction and Theorem 7.13.

Example 7.19. We consider the following SDE on $\mathbb{R}$,

$$dX_t = f(X_t)\mathrm{d}t + \mathrm{d}B_t,\ X_0 = x, \tag{100}$$

where $f : \mathbb{R} \to \mathbb{R}$ is typically a Lipchitz function, and $(B_t)_{t\geq 0}$ is a standard one-dimensional Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

By classical results about finite-dimensional SDE with global Lipschitz drift, nondegenerate diffusion coefficient (see e.g. [Gal16], Theorems 8.3 and 8.6), the solution is well-defined, unique, and generates a Markov process $(X_t^x)_{t\geq 0}$ with Markov semigroup $(P_t)_{t\geq 0}$ defined as

$$P_t\varphi(x) = \mathbb{E}\left(\varphi(X_t^x)\right),\ \forall t \geq 0,\ \varphi \in \mathcal{B}_b(\mathbb{R}).$$

Let $T > 0$ be fixed, $\varepsilon \in \mathbb{R}$, and let's consider the following modified SDE on $\mathbb{R}$,

$$dY_t = f(X_t)\mathrm{d}t + \mathrm{d}B_t - \frac{1}{T}\varepsilon\mathrm{d}t, \quad Y_0 = y = x + \varepsilon, \tag{101}$$

where $(B_t)_{t\geq 0}$ is the same standard Brownian motion as considered above. In particular, $\varepsilon = x - y$.

We can compute

$$dY_t - dX_t = -\frac{1}{T}\varepsilon\mathrm{d}t \Rightarrow Y_t - Y_0 - (X_t - X_0) = \int_0^t -\frac{1}{T}\varepsilon ds = -\frac{t}{T}\varepsilon,$$

so that

$$Y_t - X_t = Y_0 - X_0 - \frac{t}{T}\varepsilon = x + \varepsilon - x - \frac{t}{T}\varepsilon = \frac{T-t}{T}\varepsilon. \tag{102}$$

We can then conclude that $Y_T = X_T$. We consider the following process $(\tilde{B}_t)_{0\leq t\leq T}$, generated by

$$d\tilde{B}_t = \mathrm{d}B_t + \left(f(X_t) - f(Y_t) - \frac{1}{T}\varepsilon\right)\mathrm{d}t, \quad \forall 0 \leq t \leq T.$$

Since

$$d\tilde{B}_t = \mathrm{d}B_t - K_t\mathrm{d}t,$$

where $K_t = f(Y_t) - f(X_t) + \frac{1}{T}\varepsilon$, a sufficient condition to apply Girsanov's theorem and to find a probability measure $\mathbb{Q}$ such that $(\tilde{B}_t)_{0\leq t\leq T}$ is a $\mathbb{Q}$–Brownian motion is the Novikov's condition, see Theorem 7.6. In our case, we have

$$\begin{aligned} |K_s|^2 &= |f(Y_s) - f(X_s) + \frac{1}{T}\varepsilon|^2 \\ &\leq |f(Y_s) - f(X_s)|^2 + \frac{1}{T^2}\varepsilon^2 \\ &\leq C|Y_s - X_s|^2 + \frac{1}{T^2}\varepsilon^2 \\ &\underset{102}{=} C\varepsilon^2\frac{(T-s)^2 + 1}{T^2}, \end{aligned}$$

where C is the Lipschitz constant of f. It implies that the process $\left(\mathcal{E}\left(\int_0^t K_s dB_s\right)\right)_{t\geq 0}$ is a continuous, uniformly integrable martingale. Thus, the assumptions of Girsanov's theorem are verified and it implies that $(\tilde{B}_t)_{0\leq t\leq T}$ is a $\mathbb{Q}$-Brownian motion for $d\mathbb{Q} = R_T d\mathbb{P}$, and where

$$\begin{aligned} R_T &= \mathcal{E}\left(\int_0^T K_s dB_s\right) \\ &= \exp\left(\int_0^T K_s dB_s - \frac{1}{2}\left\langle \int_0^T K_s dB_s \right\rangle_t\right) \\ &= \exp\left(\int_0^T K_s dB_s - \frac{1}{2}\int_0^T |K_s|^2 ds\right), \end{aligned}$$

where the quadratic variation of $\int_0^T K_s dB_s$ follows Lemma 15.11 in [Kal01]. We can now rewrite the modified SDE (101) as

$$\begin{aligned} f(Y_t)\mathrm{d}t + d\tilde{B}_t &= f(Y_t)\mathrm{d}t + \mathrm{d}B_t - K_t\mathrm{d}t \\ &= f(Y_t)\mathrm{d}t + \mathrm{d}B_t - \left(f(Y_t)\mathrm{d}t - f(X_t)\mathrm{d}t + \frac{1}{T}\varepsilon\right)\mathrm{d}t \\ &= f(X_t) + \mathrm{d}B_t - \frac{1}{T}\varepsilon\mathrm{d}t \\ &= dY_t, \quad \forall 0 \leq t \leq T, \end{aligned}$$

and $Y_t = X_t$, $\forall t > T$.

As before, the same conclusions hold for the solution of the modified SDE: it generates a well-defined, unique Markov process $(Y_t^y)_{t\geq 0}$, and

$$P_t\varphi(y) = \mathbb{E}_{\mathbb{Q}}\left(\varphi(Y_t^y)\right), \quad \forall \varphi \in \mathcal{B}_b(\mathbb{R}) \text{ and } y \in \mathbb{R}.$$

Since conditions of Theorem 7.17 are verified, we have the following inequality,

$$\left(P_T \log \varphi\right)(y) \leq \left(\log P_T\varphi\right)(x) + \mathbb{E}_{\mathbb{P}}\left[R_T \log R_T\right], \quad \forall \varphi \in \mathcal{B}_b^+(E) \text{ with } \varphi > 0 \text{ and } \forall x, y \in \mathbb{R}.$$

Now, we derive $\mathbb{E}_{\mathbb{P}}\left[R_T \log R_T\right]$ which yields

$$\mathbb{E}_{\mathbb{P}}\left[R_T \log R_T\right] = \mathbb{E}_{\mathbb{Q}}\left[\log R_T\right] = \mathbb{E}_{\mathbb{Q}}\left[\int_0^T K_s dB_s - \frac{1}{2}\int_0^T |K_s|^2 ds\right].$$

Using

$$d\tilde{B}_t = \mathrm{d}B_t - K_t \mathrm{d}t \Rightarrow \mathrm{d}B_t = d\tilde{B}_t + K_t \mathrm{d}t,$$

and recalling that $(\tilde{B}_t)_{0\leq t\leq T}$ is a Brownian motion under $\mathbb{Q}$, it remains that

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left[R_T \log R_T\right] &= \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{2}\int_0^T |K_s|^2 ds\right] \\
&\leq \frac{C\varepsilon^2}{2T^2}\mathbb{E}_{\mathbb{Q}}\left[\int_0^T (T-s)^2 + 1\right] \\
&= C\varepsilon^2\left(\frac{T}{6} + \frac{1}{2T}\right) \\
&= C|y-x|^2\left(\frac{T}{6} + \frac{1}{2T}\right),
\end{aligned}$$

recalling that $\varepsilon = x - y$. Since no particular assumption on T or ε was used, it yields

$$P_T \log \varphi(x) \leq \log P_T\varphi(y) + C|y-x|^2\left(\frac{T}{6} + \frac{1}{2T}\right), \quad \forall T > 0,\ \forall \varphi \in \mathcal{B}_b^+(\mathbb{R}) \text{ with } \varphi > 0, \text{ and } \forall x, y \in \mathbb{R}.$$

By Theorem 7.13, with $C(T)\Psi(x, y) = C|y-x|^2\left(\frac{T}{6} + \frac{1}{2T}\right)$ we conclude that $(X_t)_{t\geq 0}$ is strong Feller at time $T > 0$.

Remark 7.20. The assumption on f being $C-$Lipschitz may be replaced by a locally Lipschitz condition together with a control to avoid blow up such as a condition of linear growth,

$$|\sigma(t, x)| \leq K(1 + |x|), \quad |b(t, x)| \leq K(1 + |x|), \quad \forall t \geq 0,\ x \in \mathbb{R}.$$

This is a common construction where we localize the Lipschitz condition, well-developed when we are looking for existence and uniqueness of solutions to SDE (see e.g. [Gal16], Section 8.2). We keep the Lipschitz condition for simplicity in our computations.

Example 7.21. We can choose a different modified SDE to get another coupling. We consider the previous SDE (100) but on $\mathbb{R}^d$. As before, let $T > 0$ and $x \in \mathbb{R}^d$ be fixed, and

$$dY_t = f(Y_t)\mathrm{d}t + \eta(t)\frac{X_t - Y_t}{\|X_t - Y_t\|}\mathbf{1}_{[0,\tau)}(t)\mathrm{d}t + \mathrm{d}B_t, \quad Y_0 = y, \tag{103}$$

where $y \in \mathbb{R}^d$ is another initial condition, $(B_t)_{t\geq 0}$ is a standard $d-$dimensional Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, τ is the coupling time defined as

$$\tau = \inf\{s \geq 0 \ : \ X_s = Y_s\},$$

and η is a positive continuous function which will be defined later. In this model, the term $X_s - Y_s$ forces the modified process $(Y_t)_{t\geq 0}$ to move to $(X_t)_{t\geq 0}$. We can see that this equation has a unique solution up to the stopping time τ, and we let $Y_t = X_t$ for all $t \geq \tau$ (see e.g. [Hsu02], Theorem 1.1.8). Here, $\|\cdot\|$ stan for the standard Euclidean norm on $\mathbb{R}^d$.

Firstly, we want to show that such a τ is well-defined: let $(\tau_n)_{n\geq 1}$ be an increasing sequence of stopping times defined for each n as

$$\tau_n = \inf_{s\geq 0}\Big\{\|X_s - Y_s\| \leq 1/n\Big\},$$

such that $\tau_n \uparrow \tau$ as $n \to \infty$. By focusing on the stopped process $(Y_t^{\tau_n})_{t\geq 0} = (Y_{t\wedge\tau_n})_{t\geq 0}$, then

$$\|X_{t\wedge\tau_n} - Y_{t\wedge\tau_n}\| > 1/n,$$

and

$$dY_t^{\tau_n} = f(Y_t^{\tau_n})\mathrm{d}t + \eta(t\wedge\tau_n)\frac{X_t^{\tau_n} - Y_t^{\tau_n}}{\|X_t^{\tau_n} - Y_t^{\tau_n}\|}\mathrm{d}t + \mathrm{d}B_t,$$

has a unique solution up to τ_n, for each $n \geq 0$. This solution is a semimartingale according to standard results (see e.g. [Gal16], Definition 8.1). Indeed, for each $n > 0$ fixed, the additional drift is the function

$$(t, Y) \mapsto \frac{X_t - Y}{\|X_t - Y\|},$$

which is Lipschitz in Y on $[0, t\wedge\tau_n]$ since

$$\begin{aligned}
\Big\|\frac{X_t - Y}{\|X_t - Y\|} - \frac{X_t - Z}{\|X_t - Z\|}\Big\| &= \Big\|\frac{X_t - Y}{\|X_t - Y\|} + \frac{X_t - Z}{\|X_t - Y\|} - \frac{X_t - Z}{\|X_t - Y\|} - \frac{X_t - Z}{\|X_t - Z\|}\Big\| \\
&\leq \Big\|\frac{X_t - Y}{\|X_t - Y\|} - \frac{X_t - Z}{\|X_t - Y\|}\Big\| + \Big\|\frac{X_t - Z}{\|X_t - Y\|} - \frac{X_t - Z}{\|X_t - Z\|}\Big\|,
\end{aligned}$$

where the first term can be bounded by

$$\Big\|\frac{X_t - Y}{\|X_t - Y\|} - \frac{X_t - Z}{\|X_t - Y\|}\Big\| = \frac{\|Y - Z\|}{\|X_t - Y\|} \leq n\|Y - Z\|.$$

For the second term:

$$\begin{aligned}
\Big\|\frac{X_t - Z}{\|X_t - Y\|} - \frac{X_t - Z}{\|X_t - Z\|}\Big\| &= \Big\|X_t - Z\Big\|\Big|\frac{1}{\|X_t - Y\|} - \frac{1}{\|X_t - Z\|}\Big| \\
&= \Big\|X_t - Z\Big\|\Big|\frac{\|X_t - Z\| - \|X_t - Y\|}{\|X_t - Y\|\cdot\|X_t - Z\|}\Big| \\
&\leq \Big|\frac{\|X_t - Z\| - \|X_t - Y\|}{\|X_t - Y\|}\Big| \\
&\leq n\|Y - Z\|,
\end{aligned}$$

so that

$$\Big\|\frac{X_t - Y}{\|X_t - Y\|} - \frac{X_t - Z}{\|X_t - Z\|}\Big\| \leq 2n\|Y - Z\|_{\mathbb{R}^d}.$$

So for each fixed n, $(Y_t^{\tau_n})_{t\geq 0}$ is a semimartingale, and

$$Y_t^{\tau_n} = Y_0 + \int_0^{t\wedge\tau_n} f(Y_s) + \eta(s)\frac{X_s - Y_s}{\|X_s - Y_s\|}ds + \int_0^{t\wedge\tau_n} dB_s,$$

which is the unique solution of (103) up to the stopping time τ.

Now, the choice of $\eta(t)$ is major since it ensures that the force will be strong enough to make $(Y_t)_{t\geq 0}$ moving towards $(X_t)_{t\geq 0}$ before our fixed time $T > 0$ and in particular, $X_T = Y_T$. Let $X_s = Y_s$ for $s \geq \tau$. Under the assumption

$$\|f(x) - f(y)\| \leq K\|x - y\|, \quad \forall x, y \in \mathbb{R}^d,$$

and since the difference process $(X_t - Y_t)_{t\geq 0}$ has zero quadratic variation, then the classical chain rule applies so that

$$\begin{aligned}
\frac{d}{ds}\|X_s - Y_s\|^2 &= 2\left\langle \frac{d}{ds}(X_s - Y_s), X_s - Y_s \right\rangle \\
&\leq 2\left\langle f(X_s) - f(Y_s) - \eta(s)\frac{X_s - Y_s}{\|X_s - Y_s\|}, X_s - Y_s \right\rangle \\
&\leq 2K\|X_s - Y_s\|^2 - 2\eta(s)\|X_s - Y_s\|.
\end{aligned}$$

Since the inequality holds true for all $s < \tau$, we can restrict s so that $X_s \neq Y_s$. We can divide both side by $\|X_s - Y_s\|$, leading to

$$\frac{d}{ds}\|X_s - Y_s\| \leq 2K\|X_s - Y_s\| - 2\eta(s).$$

Let $g(s) = \|X_s - Y_s\|$ so that the inequality rewrites as

$$\frac{d}{ds}g(s) - 2Kg(s) \leq -2\eta(s).$$

The homogeneous case of this differential equation would be $g(s) = \lambda e^{2Ks}$, and we obtain the general case by using a variation of the constant,

$$\frac{d}{ds}\left(\lambda_s e^{2Ks}\right) - 2K\lambda_s e^{2Ks} = \frac{d}{ds}(\lambda_s)\, e^{2Ks} \leq -2\eta(s),$$

which means

$$\lambda_s \leq -2\int_0^s \eta(r)e^{-2Kr}dr.$$

Remarking that $g(0)e^{2Ks} \geq 0$, it leads to

$$\begin{aligned}
g(s) &\leq g(s) + g(0)e^{2Ks} \\
&= \lambda_s e^{2K_s} + \|Y_0 - X_0\|e^{2Ks} \\
&\leq e^{2Ks}\left(\|y - x\| - 2\int_0^s \eta(r)e^{-2Kr}dr\right).
\end{aligned}$$

We can still restrict s to $T \wedge \tau$. In this case, one have

$$e^{-K(T\wedge\tau)}\|X_{T\wedge\tau} - Y_{T\wedge\tau}\| \leq \|x - y\| - 2\int_0^{T\wedge\tau} e^{-2Ks}\eta(s)ds.$$

By taking a function $\eta(s)$ such that $\|x - y\| - 2\int_0^T e^{-2Ks}\eta(s)ds \leq 0$, for example

$$\eta(s) = \frac{\|x - y\|e^{-Ks}}{\int_0^T e^{-2Ks}ds}, \quad \forall s \geq 0,$$

so that $\tau < T$ unless we have a contradiction with $\|X_T - Y_T\| > 0$ when $T < \tau$ which implies $X_T \neq Y_T$. We conclude that τ is finite, $X_\tau = Y_\tau$, and $\tau < T$. We let $X_t = Y_t$ for all $t \geq \tau$. Thus,

$$\int_0^T e^{-2Ks}\eta(s)ds = \|x-y\|\frac{\int_0^T e^{-3Ks}ds}{\int_0^T e^{-2Ks}ds} \geq \frac{2}{3}\|x-y\|,$$

since the function $\phi(t) := \frac{1-e^{-3t}}{1-e^{-2t}}$, for $t \geq 0$, reaches its minimum in 1 when $t \to \infty$, so that

$$\|x-y\| - 2\int_0^T e^{-2Ks}\eta(s)ds \leq \left(1 - \frac{4}{3}\right)\|x-y\| \leq 0.$$

As in the previous Example 7.19, letting

$$R_t = \exp\left(\int_0^t K_s dB_s - \frac{1}{2}\int_0^t \|K_s\|^2 ds\right), \quad \forall 0 \leq t \leq T,$$

where

$$K_s = -\eta(s)\mathbf{1}_{[0,\tau)}(s)\frac{X_s - Y_s}{\|X_s - Y_s\|},$$

then $(\tilde{B}_t)_{0\leq t\leq T}$ defined as

$$\tilde{B}_t = B_t - \int_0^t K_s ds, \ \forall t \geq 0,$$

is a $d-$dimensional standard Brownian motion under the probability $\mathbb{Q}$ defined by

$$d\mathbb{Q}|_{\mathcal{F}_t} = R_t d\mathbb{P}|_{\mathcal{F}_t}, \quad \forall 0 \leq t \leq T,$$

by Theorem 7.5, since

$$\langle \int_0^t K_s dB_s \rangle_\infty = \int_0^\infty \|K_s\|^2 ds \leq \frac{\|x-y\| \int_0^\tau e^{-Ks}ds}{\int_0^T e^{-2Ks}ds} < \infty,$$

so that Novikov's condition (96) is verified. By Theorem 7.17, using $dB_s = d\tilde{B}_s + K_s ds$ and since $(\tilde{B}_t)_{0\leq t\leq T}$ is a $d-$dimensional standard $\mathbb{Q}-$Brownian motion, we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}(R_t \log R_t) &= \mathbb{E}_{\mathbb{Q}}(\log R_t) \\
&= \mathbb{E}_{\mathbb{Q}}\left[\int_0^t K_s dB_s - \frac{1}{2}\int_0^t \|K_s\|^2 ds\right) \\
&= \mathbb{E}_{\mathbb{Q}}\left[\int_0^t K_s d\tilde{B}_s + \frac{1}{2}\int_0^t \|K_s\|^2 ds\right) \\
&\leq \frac{1}{2}\int_0^T \|K_s\|^2 ds \\
&\leq \frac{2K\|x-y\|^2}{1-e^{-2KT}},
\end{aligned}$$

$\forall 0 \leq t \leq T$, and the following log–Harnack inequality holds true,

$$P_T(\log\varphi)(y) \leq (\log P_T\varphi)(x) + \frac{2K\|x-y\|^2}{1-e^{-2KT}}, \quad \forall T > 0, \ \forall \varphi \in \mathcal{B}_b^+(\mathbb{R}^d) \text{ with } \varphi > 0, \text{ and } \forall x, y \in \mathbb{R}^d.$$

Since it holds true for every $T > 0$, then $(P_t)_{t\geq 0}$ is strong Feller.

7.2.3 Semilinear Stochastic Partial Differential Equations in infinite dimensions

Let $(\mathbb{H}, \langle\cdot\rangle, \|\cdot\|)$ be a separable Hilbert Space, and let $\tilde{\mathbb{H}}$ be a larger Hilbert space into which $\mathbb{H}$ is densely and continuously embedded, similarly as when we construct a classical Wiener process on an infinite-dimensional Hilbert space (see e.g. [PZ92], Section 4.1).

Let $(A, \mathcal{D}(A))$ be a negative definite self-adjoint operator on $\mathbb{H}$ which generates a C_0 contraction semigroup with $S(t) = e^{At}$, $\forall t \geq 0$.

Let $\mathcal{L}_S(\mathbb{H})$ be the set of all densely defined closed linear operators $(L, \mathcal{D}(L))$ on $\mathbb{H}$ such that for every $s > 0$, $S(s)L$ extends to a unique Hilbert–Schmidt operator on $\mathbb{H}$: we keep the same notation $S(s)L$. We equip $\mathcal{L}_S(\mathbb{H})$ with the $\sigma-$algebra induced by $\{L \mapsto \langle (S(s)L)x, y\rangle \mid s > 0, x, y \in \mathbb{H}\}$.

Recall that the Hilbert-Schmidt norm of an operator $A : \mathbb{H} \to \mathbb{H}$ is defined as

$$\|A\|^2_{HS} := \sum_{i\geq 1} \|Ae_i\|^2,$$

where $\{e_i\}_{i\geq 1}$ is a orthonormal basis of $\mathbb{H}$, and an operator A is said to be an Hilbert-Schmidt operator if $\|A\|^2_{HS} < \infty$.

For fixed $T > 0$, let

$$b : [0, T] \times \mathbb{H} \to \tilde{\mathbb{H}}, \ \sigma : [0, T] \times \mathbb{H} \to \mathcal{L}_S(\mathbb{H})$$

be measurable maps. We admit $\|v\| = \infty$ for $v \notin \mathbb{H}$. We consider the following SPDE on $\mathbb{H}$,

$$dX_t = (AX_t + b(t, X_t))\, dt + \sigma(t, X_t) dW_t, \ t \in [0, T], \tag{104}$$

where $(W_t)_{t\geq 0}$ be a cylindrical Brownian motion on $\mathbb{H}$ with respect to a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

We follow [Wan13] and recall:

Definition 7.22 (Mild solution of SDE). *An $\mathbb{H}-$valued progressively measurable process $(X_t)_{t\in[0,T]}$ is called a* mild solution *of (104) if, for every $t \in [0, T]$,*

$$\int_0^t \mathbb{E}\left[\|S(t-s)b(s, X_s)\| + \|S(t-s)\sigma(s, X_s)\|^2_{HS}\right] ds < \infty,$$

and $\mathbb{P}-$almost surely,

$$X_t = S(t)X_0 + \int_0^t S(t-s)b(s, X_s)ds + \int_0^t S(t-s)\sigma(s, X_s)dW_s.$$

Then, to ensure the existence and uniqueness of the solution of (104), we assume the following:

Hypothesis 11. For every $s > 0$ and $t \in [0, T]$, $S(s)b(t, 0) \in \mathbb{H}$ with

$$\int_0^T \sup_{r\in[0,T]} \|S(s)b(r, 0)\|^2 ds < \infty,$$

and there exists a positive function $K_b \in C((0, T])$ with $\phi_b(t) := \int_0^t K_b(s)ds < \infty$, $\forall 0 \leq t \leq T$, such that

$$\|S(t)(b(s, x) - b(s, y))\|^2 \leq K_b(t)\|x - y\|^2, \ \forall s, t \in [0, T], \ \forall x, y \in \mathbb{H}.$$

Hypothesis 12. We suppose that

$$\int_0^T \sup_{r\in[0,T]} \|S(s)\sigma(r,0)\|_{HS}^2 ds < \infty,$$

and there exists a positive function $K_\sigma \in C((0,T])$ with $\phi_\sigma(t) := \int_0^t K_\sigma(s)ds < \infty$, $\forall 0 \le t \le T$, such that

$$\|S(t)(\sigma(s,x) - \sigma(s,y))\|_{HS}^2 \le K_\sigma(t)\|x-y\|^2, \ \forall s,t \in [0,T], \ \forall x,y \in \mathbb{H}.$$

Then,

Theorem 7.23 ([Wan13], Theorem 3.1.1)**.** *Let's assume Hypotheses 11 and 12.*

(i) For every $X_0 \in L^2(\Omega \to \mathbb{H}, \mathcal{F}_0, \mathbb{P})$, (104) has a unique mild solution $(x_t)_{t\ge0}$, and there exists a constant $t_0 \in (0,T]$ such that for every $n \ge 1$,

$$\begin{aligned}\sup_{t\in T\wedge(nt_0)} \mathbb{E}\left(\|X_t\|^2\right) &\le 6^n\mathbb{E}\left(\|X_0^2\|\right) + 12\left(\sum_{i\in I} 6^{n-i}\right)\int_0^{t_0} t_0 \sup_{r\in[0,T]} \|S(s)b(r,0)\|^2 ds \\ &\quad +12\left(\sum_{i\in I} 6^{n-i}\right)\int_0^{t_0} \sup_{r\in[0,T]} \|S(s)\sigma(r,0)\|_{HS}^2 ds.\end{aligned}$$

(ii) If there exists a constant $\varepsilon > 0$ such that $\mathbb{E}\left(\|X_0\|^{2(1+\varepsilon)}\right) < \infty$ and

$$\int_0^T \left(K_\sigma(s) + K_b(s) + \sup_{r\in[0,T]} \left(\|S(s)\sigma(r,0)\|_{HS}^2 + \|S(s)b(r,0)\|^2\right)\right)^{1+\varepsilon} ds < \infty,$$

then

$$\mathbb{E} \sup_{t\in[0,T]} \|X_t\|^{2(1+\varepsilon)} \le C\left(1 + \mathbb{E}(\|X_0\|^{2(1+\varepsilon)}\right).$$

Moreover, if the following inequality

$$\int_0^T s^{-\alpha}\left(K_\sigma(s) + \sup_{r\in[0,T]} \|S(s)\sigma(r,0)\|_{HS}^2\right)ds < \infty,$$

holds for some constant $\alpha \in (0,1)$ and all $r \in [0,T]$, $x \in \mathbb{H}$, the the solution has a continuous version.

Sketch of the proof: The existence and uniqueness parts are a consequence of a combination of Fixed-Point and Local Inversion Theorems (see e.g. [PZ92], Theorem 7.2).

The first inequality is a consequence of the definition of a mild solution for (104) combined with the estimates on $\|S(s)b(r,u)\|^2$ and $\|S(s)\sigma(r,u)\|_{HS}^2$ from Hypotheses 11 and 12.

Repeating the argument for $\|X_t\|^{2(1+\varepsilon)}$ instead of $\|X_t\|^2$, combined with the Hypothesis 12, we obtained the second inequality about the non-explosion.

Finally, the existence of a continuous version is given by Theorem 7.7 from [PZ92]. □

In addition, to study the properties of the associated semigroup $(P_t)_{t\in[0,T]}$, we need the following finite-dimensional approximation assumption:

Hypothesis 13. $(A, \mathcal{D}(A))$ has a discrete spectrum, so that there exists an orthonormal basis $\{e_n\}_{n\ge1} \subset \mathcal{D}(A)$ of $\mathbb{H}$ such that

$$-Ae_n = \lambda_n e_n, \ \forall n \ge 1,$$

where $\lambda_n \geq 0$ are all eigenvalues of $-A$ including multiplicities. For $n \geq 1$, let $\mathcal{P}_n$ be the projection operator from $\mathbb{H}$ into $\mathbb{H}_n := \text{span}\{e_1, \cdots, e_n\}$.

For simplicity, we let $\langle \cdot \rangle$ and $\| \cdot \|$ stand for the scalar product and norm on $\mathbb{H}$, while we use the subscript $\mathbb{H}_n$ to precise when we are on the projected space.

We remark that $\mathcal{P}_n$ commutes with the semigroup $S(t)$ for $t \geq 0$. Finally, let

$$A_n = A|_{\mathbb{H}_n},\ b_n = \mathcal{P}_n b,\ \sigma_n = \mathcal{P}_n \sigma.$$

Now, we can consider the system of stochastic differential equations on $\mathbb{H}_n$ given by

$$\begin{cases} dX_t^{(n)} = \left[A_n X_t^{(n)} + b_n\left(t, X_t^{(n)}\right)\right] \mathrm{d}t + \sigma_n\left(t, X_t^{(n)}\right) dW_t, \\ dX_0^{(n)} = \mathcal{P}_n X_0, \end{cases} \tag{105}$$

where

$$\sigma_n\left(t, X_t^{(n)}\right) dW_t = \sum_{i\geq 1}\left(\sum_{j\geq 1}\langle \sigma(t, X_t^{(n)}) e_j, e_i\rangle d\langle W_t, e_j\rangle\right) e_i.$$

Remark that Hypothesis 11 implies that $b_n(t, \cdot)$ is Lipschitz continuous uniformly in $t \in [0, T]$. Also, using Hypothesis 12, there exists a constant $C_n > 0$ such that

$$\begin{aligned} \sum_{i\in I}\sum_{j=1}^{\infty}\langle(\sigma(t,x) - \sigma(t,y))\, e_j, e_i\rangle^2 &\leq e^{2\lambda_n T}\|S(T)(\sigma(t,x) - \sigma(t,y))\|_{HS}^2 \\ &\leq C_n\left(1 + \|x - y\|_{\mathbb{H}_n}^2\right), \end{aligned}$$

and

$$\begin{aligned} \sum_{i\in I}\sum_{j=1}^{\infty}\langle\sigma(t,x) e_j, e_i\rangle^2 &\leq e^{2\lambda_n T}\|S(T)\sigma(t,x)\|_{HS}^2 \\ &\leq C_n\left(1 + \|x\|_{\mathbb{H}_n}^2\right), \end{aligned}$$

for all $x, y \in \mathbb{H}_n$ and $t \in [0, T]$. Therefore, it is well-known that for any initial data, (105) has a unique strong solution (see e.g. [Gal16], Theorem 8.3 and consecutive remark). In addition,

Theorem 7.24 ([Wan13], Theorem 3.1.2). *Let's assume Hypotheses 11-13. If* $\mathbb{E}\left(\|X_0\|^2\right) < \infty$, *then*

$$\lim_{n\to\infty} \mathbb{E}\left(\|X_t^{(n)} - X_t\|^2\right) = 0, \quad \forall t \in [0, T].$$

Let's focus on cases where the noise is additive, in the sense that $\sigma(t, x) = \sigma(t)$. Denote by σ^* the adjoint of σ in $\mathbb{H}$. Then, we get the following log–Harnack inequality for the solution of (104):

Theorem 7.25 ([Wan13], Theorem 3.2.1). *Let's assume Hypotheses 11-13, and let's suppose that* $\sigma\sigma^*$ *is invertible so that* $\hat{\sigma}(t) := \sigma^*(t)(\sigma\sigma^*)^{-1}(t)$ *is bounded by* $\lambda > 0$, *which is*

$$\|\hat{\sigma}(t)\|_{HS}^2 \leq \frac{1}{\lambda}, \quad \forall t \geq 0, \tag{106}$$

and there exists $K > 0$ *a constant such that*

$$\langle b(s,x) - b(s,y), x - y\rangle \leq K\|x - y\|^2, \quad \forall s \in [0, T],\ \forall x, y \in \mathbb{H}.$$

Then, for every strictly positive function $f \in \mathcal{B}_b(\mathbb{H})$ *and* $\forall x, y \in \mathbb{H}$, *we have*

$$P_T \log f(y) \leq \log P_T f(x) + \frac{K\|x - y\|^2}{\lambda(1 - e^{-2KT})}.$$

Remark 7.26. The condition on $\hat{\sigma}$ is weaker than assuming σ is invertible. If we assume so, then (106) is equivalent to

$$\|\sigma^{-1}(t)\|_{HS}^2 \leq \frac{1}{\lambda}.$$

Now, let $(e_i)_{i\geq 1}$ be an orthonormal basis of a Hilbert space $\mathbb{H}$ and let

$$\sigma e_1 = 0, \text{ and } \sigma e_i = e_{i-1}, \quad \forall i \geq 2.$$

It is clear that σ is not invertible: however, since $\sigma^* e_i = e_{i+1}$ for all $i \geq 1$, thus $\sigma\sigma^*$ is the identity operator and it is invertible.

Sketch of the proof: Since f is strictly positive and bounded, $\log f$ is well-defined and is also bounded. We will show the inequality only for continuous functions f with $\inf f > 0$, instead of all measurable ones: indeed, it is well-known that we can approximate a measurable function g by a decreasing sequence of continuous bounded functions $\{g_k\}_{k\geq 0}$ with $\inf g_k > 0$ so that $g_k \downarrow g$ pointwise, almost everywhere as $k \to \infty$.

Under these settings, it suffices to prove the inequality for $P_T^{(n)}$ instead of P_T: we'll extend the result from $\mathbb{H}_n$ to $\mathbb{H}$ by dominated convergence.

In the finite-dimensional setting, it is quite a similar construction as in Example 7.21 so we do not dive into too many details.

Let $x, y \in \mathbb{H}_n$, $T > 0$ be fixed. Let $X_t^{(n)}$ solves (105) with $X_0^{(n)} = x$, and let's consider the following modified SDE,

$$\begin{cases} dY_t^{(n)} = \left(A_n Y_t^{(n)} + b_n\left((t, Y_t^{(n)}\right)\right) \mathrm{d}t + \eta(t) \dfrac{X_t^{(n)} - Y^{(n)}}{\|X_t^{(n)} - Y^{(n)}\|_{\mathbb{H}_n}} \mathbf{1}_{[0,\tau)}(t) + \sigma_n(t) dW_t, \\ dY_0^{(n)} = y, \end{cases} \tag{107}$$

where $\eta \in C((0,\infty))$ is to be determined, and $\tau := \inf\{s > 0 \mid X_s^{(n)} = Y_s^{(n)}\}$ is the coupling time. This is a well-defined stopping time and the modified SDE possess a unique strong solution up to τ: we fix $X_t^{(n)} = Y_t^{(n)}$ for all $t \geq \tau$.

By choosing $\eta(t)$ such that $2\int_0^T \eta(t) e^{-2Kt} \mathrm{d}t \geq \|x - y\|_{\mathbb{H}_n}$, it follows that $\tau \leq T$ and $X_T^{(n)} = Y_T^{(n)}$. In particular, the construction of η uses the condition

$$\langle b_n(s,x) - b_n(s,y), x - y\rangle_{\mathbb{H}_n} \leq K\|x - y\|_{\mathbb{H}_n}, \quad \forall s \in [0,T], \; x, y \in \mathbb{H}_n,$$

inherited from the same condition for b on $\mathbb{H}$. Such an η can be defined as

$$\eta(t) = \frac{2Ke^{-Kt}\|x - y\|_{\mathbb{H}_n}}{1 - e^{-2KT}},$$

by analogy with Example 7.21.

Now, let $d\mathbb{Q}|_{\mathcal{F}_t} = R_t d\mathbb{P}|_{\mathcal{F}_t}$ for all $0 \leq t \leq T$, where

$$R_t := \exp\left(\int_0^t \langle \psi(s), dW_s\rangle - \frac{1}{2}\int_0^t \|\psi(s)\|_{\mathbb{H}_n}^2 ds\right),$$

and where

$$\psi(s) := -\eta(s)\mathbf{1}_{[0,\tau)}(s) \frac{\hat{\sigma}(s)\left(X_s^{(n)} - Y_s^{(n)}\right)}{\|X_s^{(n)} - Y_s^{(n)})\|_{\mathbb{H}_n}}.$$

Note that

$$\mathbb{E}_{\mathbb{Q}}\left[R_t \log R_t\right] \leq \frac{1}{2}\int_0^{\tau} \|\psi(s)\|^2_{\mathbb{H}_n} \mathrm{d}s \leq \frac{1}{2\lambda}\int_0^{T} \eta(s)^2 \mathrm{d}s = \frac{K\|x-y\|^2_{\mathbb{H}_n}}{\lambda(1-e^{-2KT})}, \tag{108}$$

$\forall 0 \leq t \leq T$. By Girsanov's theorem and thanks to Novikov's condition, $\mathbb{Q}$ is a probability measure under which

$$\tilde{W}_t = W_t - \int_0^t \psi(s) ds, \quad 0 \leq t \leq T,$$

is a cylindrical Brownian motion on $\mathbb{H}$. It follows from Girsanov's theorem provided the Novikov's condition holds, since for $\|K_t\|^2 = \eta(t)^2 \mathbf{1}_{[0,\tau)}(t)$, a sufficient condition can be expressed as

$$\mathbb{E}_{\mathbb{P}}\left(e^{\frac{1}{2}\int_0^T \eta(s)^2 ds}\right) < \infty,$$

because $\tau < T$, which is verified by (108) so that $(\tilde{W}_t)_{0\leq t\leq T}$ is a $\mathbb{Q}$-Brownian motion.

We can modify (107) so that

$$\begin{cases} dY_t^{(n)} = \left(A_n Y_t^{(n)} + b_n\left((t, Y_t^{(n)}\right)\right)\mathrm{d}t + \sigma_n(t) d\tilde{W}_t, \\ dY_0^{(n)} = y, \end{cases} \tag{109}$$

which implies that $\left(X_t^{(n)}, Y_t^{(n)}\right)_{t\geq 0}$ is a coupling by change of measure as defined in Theorem 7.17 and

$$\begin{aligned} P_T \log f(y) \leq & \quad \log P_T f(x) + \mathbb{E}_{\mathbb{Q}}\left[R_T \log R_T\right) \\ \leq & \quad \log P_T f(x) + \frac{K\|x-y\|^2_{\mathbb{H}_n}}{\lambda(1-e^{-2KT})}, \end{aligned} \tag{110}$$

$\forall f \in C_b(\mathbb{H}_n)$ with $f > 0$ and $\forall x, y \in \mathbb{H}_n$. By using Theorem 7.24, we can pass to the limit $n \to \infty$ and by dominated convergence, we obtain

$$P_T \log f(y) \leq \log P_T f(x) + \frac{K\|x-y\|^2}{\lambda(1-e^{-2KT})}, \quad \forall f \in C_b(\mathbb{H}) \text{ and } x, y \in \mathbb{H}.$$

Finally, the inequality holds true for $f \in \mathcal{B}_b(\mathbb{H})$ since there exists a decreasing sequence of continuous functions f_k, converging pointwise towards f, almost everywhere so that (110) holds true for all $f \in \mathcal{B}_b(\mathbb{H})$ by a monotone convergence argument.

Since $\Psi(x,y) := \frac{K\|x-y\|^2_{\mathbb{H}}}{\lambda(1-e^{-2KT})}$ is defined such that $\lim_{y\to x}\Psi(x,y) = \lim_{y\to x}\Psi(y,x) = 0$, then by Theorem 7.13, $(P_t)_{t\geq 0}$ is strong Feller at time T. Since previous computations hold true for any $T > 0$, $(P_t)_{t>0}$ is strong Feller. □

Remark 7.27. We can already state why such a construction will not work in our degenerate model (88): since the noise term does not affect every direction, it is not possible to construct a modified copy of our SDE such that

(i) modifications are only on the noisy coordinates;

(ii) and we ensure that $X_t \neq Y_t$ on all coordinates up to a time $\tau < \infty$, while $X_t = Y_t$ after τ.

Indeed, because of the degeneracy of the noise covariance, it is naturally possible to have $X_t^1 = Y_t^1$, where X_t^1, Y_t^1 denote the degenerate parts of $(X_t)_{t\geq 0}$, $(Y_t)_{t\geq 0}$, while $X_t^2 \neq Y_t^2$, where X_t^2, Y_t^2 denote the nondegenerate parts, so that the above construction has either non-sense or does not imply a coupling.

7.3 A modified log–Harnack inequality to prove asymptotic strong Feller property

Based on the previous results from [Wan13], one may ask whether such a finite meeting time exists. In the spirit of [Hai02], the author proposed a new coupling construction that overcomes this assumption: in the asymptotic setting, the goal is to show that the two coupled processes converge exponentially fast towards each other.

The shift from a finite coupling time to the exponential convergence may remind you of the transition from the strong Feller property to the asymptotic one: since we are aiming to demonstrate that a kind of strong Feller property holds asymptotically, the connection with asymptotic coupling makes sense.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(\mathcal{X}, d)$ be a Polish space. Following the idea of Xu in [Xu11], the asymptotic log–Harnack inequality is a weaker version of (97) and is defined as follows.

Definition 7.28 (**Asymptotic log–Harnack inequality**). *Let $(P_t)_{t\geq 0}$ be a Markov semigroup on a Polish metric space $(\mathcal{X}, d)$. We say that $(P_t)_{t\geq 0}$ satisfies an* asymptotic log–Harnack inequality *if $\exists \phi \in C(\mathcal{X}^2, \mathbb{R}), \Psi \in C([0,\infty) \times \mathcal{X}^2, \mathbb{R})$ two functions such that*

$$P_t \log f(x) \leq \log P_t f(y) + \phi(x,y) + \Psi(t,x,y) \|\nabla \log f\|_\infty, \tag{111}$$

for all bounded Lipschitz functions $f \geq 1$, $x, y \in \mathcal{X}$ and $t \geq 0$. Furthermore, ϕ and ψ satisfy the following conditions:

- ***(i)*** *$\phi(x,x) = 0$ for all $x \in \mathcal{X}$;*
- ***(ii)*** *$\Psi(t,x,x) = 0$ for all $x \in \mathcal{X}$ and $t \geq 0$;*
- ***(iii)*** *and there exists a constant $\delta > 0$ such that $\lim_{t\to\infty} \sup_{\{y\in\mathcal{X} | d(x,y)\leq\delta\}} \Psi(t,x,y) = 0$ for any $x \in \mathcal{X}$.*

Remark 7.29. The notation ∇ has to be considered as in the pointwise Lipschitz constant framework, see Definition 5.21.

The following result gives the link between the asymptotic log–Harnack inequality and the asymptotic strong Feller property.

Theorem 7.30 ([Xu11], Theorem 1.4). *If a Markov semigroup on a Polish metric space $\mathcal{X}$ satisfies an asymptotic log–Harnack inequality, then it is asymptotically strong Feller.*

Remark 7.31 ([Xu11], Remark 1.3). Since the asymptotically strong Feller property refers to a long-time behavior, to prove it by using the asymptotic log–Harnack inequality, we only need to verify (111) on $[T, \infty)$ for $T > 0$ fixed.

Proof. Let f be a bounded Lipschitz function. Without loss of generality, we can suppose that $f > 0$. Let $\varepsilon > 0$ be small enough such that $\varepsilon\|f\|_\infty < \frac{1}{2}$. Then, the function $2 + 2\varepsilon f$ is also bounded, Lipschitz, and $2 + 2\varepsilon f > 1$. If f takes negative values, we just have to take $-\varepsilon$ in the above construction.

Since the function $2 + 2\varepsilon f$ satisfies the condition of Definition 7.28, we can apply the asymptotic log–Harnack inequality such that

$$P_t \log(2 + 2\varepsilon f)(y) \leq \log P_t(2 + 2\varepsilon f)(x) + \varphi(x,y) + \psi(t,x,y)\|\nabla \log(2 + 2\varepsilon f)\|_\infty.$$

By linearity of $\mathbb{E}$ and the properties of the logarithm, we obtain

$$P_t \log(1 + \varepsilon f)(y) \leq \log(1 + P_t \varepsilon f(x)) + \varphi(x,y) + \psi(t,x,y)\|\nabla \log(2 + 2\varepsilon f)\|_\infty,$$

Since $\log(1+x) \le x$, we have $\|\nabla \log(2+2\varepsilon f)\|_\infty \le 2\varepsilon \|\nabla f\|_\infty$ and the previous inequality writes

$$P_t \log(1+\varepsilon f)(y) \le \log(1+P_t\varepsilon f(x)) + \varphi(x,y) + 2\varepsilon\psi(t,x,y)\|\nabla f\|_\infty. \tag{112}$$

The Taylor series expansion of second order of $\log(1+x)$ at $x=0$ is given by

$$\log(1+x) = x - \frac{x^2}{2} + o(x^3),$$

and converges for $|x|<1$. In our case, and recalling that we supposed $\varepsilon\|f\|_\infty < \frac{1}{2}$, we obtain:

(i) $P_t \log(1+\varepsilon f) = P_t\left(\varepsilon f - \frac{\varepsilon^2 f^2}{2} + o(\varepsilon^3)\right) = \varepsilon P_t f - \frac{\varepsilon^2}{2} P_t(f^2) + o(\varepsilon^3)$.

(ii) $\log(1+P_t\varepsilon f) = \varepsilon P_t f - \frac{\varepsilon^2}{2}(P_t f)^2 + o(\varepsilon^3)$.

Thus, it yields (112) writes

$$\begin{aligned}
\varepsilon\left(P_t f(y) - P_t f(x)\right) &\le \frac{\varepsilon^2}{2} P_t(f^2)(x) - \frac{\varepsilon^2}{2}\left(P_t f(y)\right)^2 + \varphi(x,y) + 2\varepsilon\psi(t,x,y)\|\nabla f\|_\infty + o(\varepsilon^3) \\
&= \frac{\varepsilon^2}{2}\left(P_t(f^2)(x) - (P_t f(y))^2\right) + \varphi(x,y) + 2\varepsilon\psi(t,x,y)\|\nabla f\|_\infty + o(\varepsilon^3) \\
&\le \frac{\varepsilon^2}{2} P_t(f^2)(x) + \varphi(x,y) + 2\varepsilon\psi(t,x,y)\|\nabla f\|_\infty + o(\varepsilon^3),
\end{aligned}$$

where the last inequality is given by $(P_t f(y))^2 \ge 0$. Since $P_t(f^2)(x)$ is bounded by $\|f\|_\infty$, we obtain

$$\begin{aligned}
\varepsilon\left(P_t f(y) - P_t f(x)\right) &\le \frac{\varepsilon^2}{2}\|f\|_\infty^2 + \varphi(x,y) + 2\varepsilon\psi(t,x,y)\|\nabla f\|_\infty + o(\varepsilon^3) \\
&\le \varepsilon^2\|f\|_\infty^2 + \varphi(x,y) + 2\varepsilon\psi(t,x,y)\|\nabla f\|_\infty + o(\varepsilon^3).
\end{aligned}$$

Dividing both side by $\varepsilon > 0$ and we obtain

$$|P_t f(y) - P_t f(x)| \le \varepsilon\|f\|_\infty^2 + \frac{\varphi(x,y)}{\varepsilon} + 2\psi(t,x,y)\|\nabla f\|_\infty + \frac{o(\varepsilon^3)}{\varepsilon}.$$

Let $\gamma > 0$ and let's define the following metric,

$$d_\gamma(x,y) = 1 \wedge \frac{1}{\gamma} d(x,y), \quad \forall x,y \in \mathcal{X},$$

where d is the metric of the Polish metric space $\mathcal{X}$. In particular, it implies that $\|d_\gamma\|_\infty := \sup_{x,y\in\mathcal{X}} d_\gamma(x,y) \le 1$.

From Definitions 6.5 and 6.7 as well as Lemma 6.11, it implies that for μ_1, μ_2 two positive measures on $\mathcal{X}$ with equal mass, then

$$\|\mu_1 - \mu_2\|_{d_\gamma} = |||\mu_1 - \mu_2|||_{d_\gamma} = \sup_{\|f\|_{d_\gamma}=1} \int_{\mathcal{X}} f(z)(\mu_1 - \mu_2)(dz).$$

By definition of d_γ, for f a bounded Lipschitz function such that $\|f\|_{d_\gamma} \le 1$, we have

$$\|f\|_\infty \le 1, \quad \|\nabla f\|_\infty \le \frac{1}{\gamma},$$

as in the proof of Proposition 6.10. Since $P_t(x,\cdot), P_t(y,\cdot)$ are positive measures on $\mathcal{X}$ with equal mass, then

$$\begin{aligned}
\|P_t(y,\cdot)-P_t(x,\cdot)\|_{d_\gamma} &= \sup_{\|f\|_{d_\gamma}=1}\int_{\mathcal{X}} f(z)\,(P_t(y,\cdot)-P_t(x,\cdot))\,(dz)\\
&= \sup_{\|f\|_{d_\gamma}=1} |P_tf(y)-P_tf(x)|\\
&\leq \sup_{\|f\|_{d_\gamma}=1} \varepsilon\|f\|_\infty^2+\frac{\varphi(x,y)}{\varepsilon}+2\psi(t,x,y)\|\nabla f\|_\infty+\frac{o(\varepsilon^3)}{\varepsilon}\\
&\leq \varepsilon+\frac{\varphi(x,y)}{\varepsilon}+\frac{2\psi(t,x,y)}{\gamma}+\frac{o(\varepsilon^3)}{\varepsilon}.
\end{aligned}$$

Let's define

$$\gamma_n=\sqrt{\sup_{\{y\in\mathcal{X}\ \mid\ d(y,x)\leq\delta\}}\psi(n,x,y)},\qquad \varepsilon=\sqrt{\varphi(x,y)}.$$

Then, for any $x,y\in\mathcal{X}$ such that $d(y,x)\leq\delta$, we have

$$\|P_n(y,\cdot)-P_n(x,\cdot)\|_{d_{\gamma_n}}\leq 2\sqrt{\varphi(x,y)}+2\sqrt{\sup_{\{y\in\mathcal{X}\ \mid\ d(y,x)\leq\delta\}}\psi(n,x,y)}+\frac{o\left(\varphi(x,y)^{\frac{3}{2}}\right)}{\sqrt{\varphi(x,y)}}.$$

The asymptotic Strong Feller property for $(P_t)_{t\geq0}$ follows (91) in Definition 6.8, since

$$\lim_{t\to\infty}\sup_{\{y\in\mathcal{X}|d(x,y)\leq\delta\}}\Psi(t,x,y)=0,$$

for any $x\in\mathcal{X}$, and

$$\varphi(x,y)\to0 \text{ and } \frac{o\left(\varphi(x,y)^{\frac{3}{2}}\right)}{\sqrt{\varphi(x,y)}}\to0 \ \text{ as } x\to y.$$

□

If this asymptotic log–Harnack inequality implies directly the asymptotic strong Feller property, we can go further in details through the following Theorem:

Theorem 7.32 ([BWY19], Theorem 2.1). *Let $(P_t)_{t\geq0}$ be a Markov semigroup satisfying the asymptotic log–Harnack inequality for some symmetric functions ϕ, $\Psi:\mathcal{X}\times\mathcal{X}\to\mathbb{R}^+$ with $\lim_{t\to\infty}\Psi(t,x,y)=0$.*

(i) *If, for any $x\in\mathcal{X}$,*

$$\Lambda(x)=\limsup_{y\to x}\frac{\phi(x,y)}{d(x,y)^2}<\infty,\ \textit{and}\ \Gamma_t(x)=\limsup_{y\to x}\frac{\Psi(t,x,y)}{d(x,y)}<\infty,$$

then for any $t>0$ and $f\in\mathcal{B}_b(\mathcal{X})$ a Lipschitz function, we have

$$\|\nabla P_tf(x)\|\leq\sqrt{2\Lambda(x)}\sqrt{P_t(f^2)(x)-(P_tf(x))^2}+\|\nabla f\|_\infty\Gamma_t(x)$$

for all $x\in\mathcal{X}$. In particular, since $\lim_{t\to\infty}\Gamma_t(x)=0$, $(P_t)_{t\geq0}$ is asymptotically strong Feller.

(ii) *If $(P_t)_{t\geq0}$ has an invariant probability measure μ, then for any $f\in\mathcal{B}_b^+(\mathcal{X})$ such that $\|\nabla f\|_\infty<\infty$,*

$$\limsup_{t\to\infty}P_tf(x)\leq\log\left(\frac{\mu e^f}{\int_{\mathcal{X}}e^{-\phi(x,y)}\mu(dy)}\right),\ x\in\mathcal{X}.$$

As a consequence, for any closed set $A\subset\mathcal{X}$ with $\mu(A)=0$, we have

$$\lim_{t\to\infty}P_t\mathbf{1}_A(x)=0,\ x\in\mathcal{X}.$$

(iii) $(P_t)_{t\geq 0}$ *has at most one invariant probability measure.*

(iv) *Let* $x \in \mathcal{X}$ *and* $A \subset \mathcal{X}$ *be a measurable set such that*

$$\delta(x,A) := \liminf_{t\to\infty} \mathcal{P}_t(x,A) > 0.$$

Then,

$$\liminf_{t\to\infty} \mathcal{P}_t(y,A_\varepsilon) > 0,$$

for all $y \in \mathcal{X}$*,* $\varepsilon > 0$*, and where*

$$A_\varepsilon = \{y \in \mathcal{X} \mid d(x,A) = \inf_{y\in A} d(x,y) < \varepsilon\}.$$

With this theoretical setting, a new range of problems has been widely investigate:

(i) If the uniqueness of the invariant measure associated to the two-dimensional Navier-Stokes equations driven by degenerate noise has been demonstrated in [HM06] using the notion of asymptotically strong Feller property, same results have been obtained by using the asymptotic log–Harnack inequality (see e.g. [Xu11]);

(ii) In [BWY19], other related problems have been treated through the use of asymptotic log–Harnack inequalities such that nondegenerate SDEs of infinite memory, semi-linear SPDEs of infinite memory, neutral SDEs of infinite memory, and stochastic Hamiltonian systems of infinite memory;

(iii) In [Moh20], the author considers the two and three-dimensional stochastic convective Brinkman-Forchheimer equations and their asymptotic behavior through the use of asymptotic log–Harnack inequality and via the asymptotic coupling methodology.

(iv) In [WHL21], the author investigated similarly the ergodicity for the three-dimensional Leray-α model with degenerate type noise;

(v) In [Ham23], the author investigates stochastic Volterra integral equations with both asymptotic log–Harnack inequality and asymptotic coupling methodology to show that the induced Markov semigroup possesses the asymptotic strong Feller property;

(vi) In [Med23], the same tools are used to study the ergodicity for the two-dimensional Stochastic Cahn-Hilliard-Navier-Stokes equations driven by an additive degenerate noise;

(vii) Some variations of these tools such as the notion of Generalized couplings have been investigated by various authors (see e.g. [KS18]). Some applications and similar results, for example on the two-dimensional Navier-Stokes Equation driven by degenerate noise, have been investigate in [BKS19].

7.3.1 Asymptotic coupling construction

As in the classical log–Harnack inequality, its asymptotic version can be recovered by use a modification of the coupling by change of measure introduced earlier: we present here the setting to construct such an *asymptotic coupling* by change of measure.

To this effect, we adapt the condition on the existence of a meeting time in Theorem 7.17 as follows.

Hypothesis 14. Let $(\mathcal{X},d)$ be a Polish metric space, and let's consider $(Z_t)_{t\geq 0}$, $(\tilde{Z}_t)_{t\geq 0}$ two stochastic processes that we supposed to be Markovian and non-explosive (typically some "nice" solutions to SDEs). We also suppose:

(i) $\exists\mathbb{Q}$ another probability measure on $(\Omega, \mathcal{F})$, defined by

$$d\mathbb{Q}_{|\mathcal{F}_t} = R_t d\mathbb{P}_{|\mathcal{F}_t},\ \forall t \geq 0,$$

or in other words,

$$R_t(\cdot) = \frac{d\mathbb{Q}_{|\mathcal{F}_t}(\cdot)}{d\mathbb{P}_{|\mathcal{F}_t}(\cdot)}, \quad \forall t \geq 0,\ \text{valued in } (\Omega, \mathcal{F}),$$

such that the distribution of $(Z_t)_{t\geq 0}$ under $\mathbb{P}$ and the one of $(\tilde{Z}_t)_{t\geq 0}$ under $\mathbb{Q}$ coincide. In other words, it means that

$$\mathbb{E}_{\mathbb{Q}}\left[\varphi(Z_t) \mid Z_0 = z\right) = \mathbb{E}_{\mathbb{Q}}\left[\varphi(\tilde{Z}_t) \mid \tilde{Z}_0 = z\right)$$

holds $\forall t \geq 0,\ \forall \varphi \in \mathcal{B}_b(\mathcal{X}),\ \forall z \in \mathcal{X}$.

(ii) $\forall z, \tilde{z} \in \mathcal{X}$, $\exists \eta > 0$ a constant such that $\forall t \geq 0$,

$$\mathbb{E}_{\mathbb{Q}}\left[d(\tilde{Z}_t^{\tilde{z}}, Z_t^z)\right) \leq e^{-\eta t} d(\tilde{z}, z).$$

In this line, we used the short writing $\tilde{Z}_t^{\tilde{z}}$, meaning $\tilde{Z}_t$ *knowing that* $\tilde{Z}_0 = \tilde{z}$.

Theorem 7.33. *Under Hypothesis 14, the following inequality*

$$P_t \log\left(f(\tilde{z})\right) \leq \log\left(P_t f(z)\right) + \mathbb{E}_{\mathbb{Q}}\left[R_t \log(R_t)\right) + \|\nabla \log(f)\|_\infty e^{-\eta t} d(\tilde{z}, z),$$

holds $\forall t \geq 0$, *for all bounded Lipschitz functions* $f \geq 1$, *and* $\forall \tilde{z}, z \in \mathcal{X}$.

The study of this particular inequality will be very useful to derive properties on $(P_t)_{t\geq 0}$ such that the asymptotic strong Feller property or the existence and uniqueness of the invariant probability measure.

Proof. Let f be a bounded, Lipschitz function with $f \geq 1$, then $\|\nabla \log(f)\|_\infty < \infty$ and $\log(f) \in \mathcal{B}_b(\mathcal{X})$. By Hypothesis 14***(i)***:

$$\begin{aligned}
P_t \log\left(f(\tilde{z})\right) &= \mathbb{E}_{\mathbb{Q}}\left[\log\left(f(Z_t^{\tilde{z}})\right)\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\log\left(f(\tilde{Z}_t^{\tilde{z}})\right)\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\log\left(f(\tilde{Z}_t^{\tilde{z}})\right)\right] + \mathbb{E}_{\mathbb{Q}}\left[\log\left(f(Z_t^z)\right)\right] - \mathbb{E}_{\mathbb{Q}}\left[\log\left(f(Z_t^z)\right)\right].
\end{aligned}$$

By reordering and using the relation between $\mathbb{Q}$ and $\mathbb{P}$, we obtain

$$\begin{aligned}
P_t \log\left(f(\tilde{z})\right) &= \mathbb{E}_{\mathbb{Q}}\left[\log\left(f(Z_t^z)\right)\right] + \mathbb{E}_{\mathbb{Q}}\left[\log\left(f(\tilde{Z}_t^{\tilde{z}})\right) - \log\left(f(Z_t^z)\right)\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[R_t \log\left(f(Z_t^z)\right)\right] + \mathbb{E}_{\mathbb{Q}}\left[\log\left(f(\tilde{Z}_t^{\tilde{z}})\right) - \log\left(f(Z_t^z)\right)\right].
\end{aligned}$$

As in the proof of Theorem 7.17, we use Lemma 7.18 to bound the first term.

In our case, let $f = R_t$, which is by Hypothesis 14***(i)*** an $\mathcal{F}_t$−measurable function with expected value equals to 1 since $d\mathbb{Q}_{|\mathcal{F}_t} = R_t d\mathbb{P}_{|\mathcal{F}_t},\ \forall t \geq 0$. In addition, let $\psi = \log(f)$: we remark that $\psi f = R_t \log(f)$ is indeed integrable with respect to $\mathbb{P}$ since

$$\mathbb{E}_{\mathbb{Q}}\left[R_t \log(f(Z_t^z))\right] = \mathbb{E}_{\mathbb{Q}}\left[\log\left(f(Z_t^z)\right)\right] = P_t \log\left(f(z)\right) < \infty,$$

since $\log f \in \mathcal{B}_b(E)$ and by definition of $(P_t)_{t\geq 0}$.

By applying this result to $\mathbb{E}_{\mathbb{Q}}\left[R_t \log\left(f(Z_t^z)\right)\right]$, it gives us

$$\mathbb{E}_{\mathbb{Q}}\left[R_t \log\left(f(Z_t^z)\right)\right] \leq \mathbb{E}_{\mathbb{Q}}\left[R_t \log(R_t)\right] + \log\left(\mathbb{E}_{\mathbb{P}}\left[f(Z_t^z)\right)\right].$$

Including it in the previous inequality and using the property (83) as well as Hypothesis 14***(ii)***, it yields

$$\begin{aligned}
P_t \log\left(f(\tilde{z})\right) &\leq \mathbb{E}_{\mathbb{Q}}\left[R_t \log(R_t)\right] + \log\left(\mathbb{E}_{\mathbb{P}}\left[f(Z_t^z)\right)\right] + \mathbb{E}_{\mathbb{Q}}\left[\log\left(f(\tilde{Z}_t^{\tilde{z}})\right) - \log\left(f(Z_t^z)\right)\right] \\
&= \log\left(P_t\left((f(z)\right)\right) + \mathbb{E}_{\mathbb{Q}}\left[R_t \log(R_t)\right] + \mathbb{E}_{\mathbb{Q}}\left[\log\left(f(\tilde{Z}_t^{\tilde{z}})\right) - \log\left(f(Z_t^z)\right)\right] \\
&\leq \log\left(P_t\left((f(z)\right)\right) + \mathbb{E}_{\mathbb{Q}}\left[R_t \log(R_t)\right] + \|\nabla \log(f)\|_\infty \mathbb{E}_{\mathbb{Q}}\left[d\left(\tilde{Z}_t^{\tilde{z}}, Z_t^z\right)\right] \\
&\leq \log\left(P_t\left((f(z)\right)\right) + \mathbb{E}_{\mathbb{Q}}\left[R_t \log(R_t)\right] + \|\nabla \log(f)\|_\infty e^{-\eta t} d(\tilde{z}, z),
\end{aligned}$$

which holds true $\forall \tilde{z}, z \in \mathcal{X}$ and $\forall t \geq 0$. □

That will be exactly our methodology: given $(Z_t)_{t\geq 0}$ the typical "nice" solution of an SDE, we can construct a modified SDE so that its solution $(\tilde{Z}_t)_{t\geq 0}$ and the original process $(Z_t)_{t\geq 0}$ satisfy the assumptions Hypothesis 14.

In this case, we can construct such an inequality thanks to Theorem 7.33. According to Theorem 7.32, we only require the existence of a symmetric function $\phi \in C(\mathcal{X}^2, \mathbb{R})$ such that $\phi(z,z) = 0 \ \forall z \in \mathcal{X}$ and $\mathbb{E}_{\mathbb{Q}}\left[R_t \log(R_t)\right) \leq \phi(z, \tilde{z})$ in order to get an asymptotic log–Harnack inequality, hence to prove that the semigroup $(P_t)_{t\geq 0}$ generated by $(Z_t)_{t\geq 0}$ is asymptotically strong Feller.

Example 7.34 (**Continuation of Example 6.28**)**.** We consider the following SDE defined on $\mathbb{R}^2$,

$$\begin{cases} dX_t = -X_t \mathrm{d}t + \mathrm{d}B_t, \\ dY_t = -Y_t \mathrm{d}t, \end{cases}$$

where $(B_t)_{t\geq 0}$ is a one-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. We have already seen that the solution $(Z_t)_{t\geq 0} = (X_t, Y_t)_{t\geq 0}$ lacks the strong Feller property. We also have seen that it's possible to derive the asymptotic strong Feller property by direct computations.

We can retrieve this property using the new methodology introduced. In this case, let's take a second copy of the SDE without any modification: its solution is $(\tilde{Z}_t)_{t\geq 0} = (\tilde{X}_t, \tilde{Y}_t)_{t\geq 0}$ with initial condition $\tilde{Z}_0 = \tilde{z} = (\tilde{x}, \tilde{y}) = (\tilde{X}_0, \tilde{Y}_0)$.

It is clear that starting with the same initial condition, meaning $z = \tilde{z}$, both processes have the same distribution under $\mathbb{P}$, and Hypothesis 14***(i)*** holds true with $\mathbb{Q} = \mathbb{P}$.

We consider the difference process $\rho(t) = \tilde{Z}_t - Z_t = (\rho_1(t), \rho_2(t)) = (\tilde{X}_t - X_t, \tilde{Y}_t - Y_t)$. In particular, the difference process is the solution of the following ODE,

$$\begin{cases} d\rho_1(t) = d\left(\tilde{X}_t - X_t\right) = -\left(\tilde{X}_t - X_t\right)\mathrm{d}t = -\rho_1(t)\mathrm{d}t, \\ d\rho_2(t) = d\left(\tilde{Y}_t - Y_t\right) = -\left(\tilde{Y}_t - Y_t\right)\mathrm{d}t = -\rho_2(t)\mathrm{d}t, \end{cases}$$

And its solution is given by

$$\rho_1(t) = \rho_1(0)e^{-t}, \ \rho_2(t) = \rho_2(0)e^{-t}.$$

In particular, with different initial conditions $z = (x, y)$ and $\tilde{z} = (\tilde{x}, \tilde{y})$, we get

$$\begin{aligned}
\|\tilde{Z}_t^{\tilde{z}} - Z_t^z\|^2 &= \rho_1(t)^2 + \rho_2(t)^2 \\
&= \left(|\rho_1(0)|^2 + |\rho_2(0)|^2\right) e^{-2t} \\
&= \|\tilde{z} - z\|^2 e^{-2t},
\end{aligned}$$

and finally $d(\tilde{Z}_t, Z_t) = \|\tilde{Z}_t - Z_t\| \leq d(\tilde{z}, z)e^{-t}$ hence Hypothesis 14***(ii)*** holds true.

In particular, we remark that $\mathbb{E}_{\mathbb{Q}}\left[R_t \log(R_t)\right) = 0$ by our choice of modified SDE and hence R_t, so our asymptotic log–Harnack inequality holds true for

$$\mathbb{E}_{\mathbb{Q}}\left[R_t \log(R_t)\right) = 0 \leq \phi(\tilde{z}, z) := \|\tilde{z} - z\|^2,$$

respecting the condition of (92). We can conclude that the following asymptotic log–Harnack inequality holds,

$$P_t \log(f(\tilde{z})) \leq \log(P_t f(z)) + \|\tilde{z} - z\|^2 + \|\nabla \log f\|_\infty \|\tilde{z} - z\| e^{-t},$$

for all bounded Lipschitz functions $f \geq 1$, and by Theorem 7.32, the semigroup $(P_t)_{t\geq 0}$ generated by $(Z_t)_{t\geq 0}$ is asymptotically strong Feller since

$$\Lambda(z) = \limsup_{\tilde{z}\to z} \frac{\phi(z, \tilde{z})}{d(z, \tilde{z})^2} = \limsup_{\tilde{z}\to x} \frac{\|z - \tilde{z}\|^2}{\|z - \tilde{z}\|^2} = 1 < \infty,$$

$$\Gamma_t(z) = \limsup_{\tilde{z}\to z} \frac{\Psi(t, z, \tilde{z})}{d(z, \tilde{z})} = \limsup_{\tilde{z}\to z} \frac{\|\tilde{z} - z\| e^{-t}}{\|\tilde{z} - z\|} = e^{-t} < \infty,$$

and $\lim_{t\to\infty} \Psi(t, z, \tilde{z}) = 0$. In particular, $(P_t)_{t\geq 0}$ is asymptotically strong Feller and possesses at most one invariant probability measure.

Remark 7.35. In fact, as stated in [HM06], Example 3.15, examples with pathwise contractive flows are asymptotically strong Feller. In our degenerate models of interest, given a deterministic degenerate part with a pathwise contractive flow, it automatically satisfies an exponential convergence of the difference process on the degenerate part.

Then, the existence of a function G that ensures that the difference process on the nondegenerate part converges exponentially is direct: the last step is to verify that there exists a probability measure $\mathbb{Q}$ such that the process $\tilde{B}_t = B_t + \int_0^t G(\tilde{X}_s, \tilde{Y}_s, X_s, Y_s)ds$ is a $\mathbb{Q}$–Brownian motion, which will follow from an integrability condition on G in the sense that $\mathbb{E}\left[\frac{1}{2}\exp\left(\int_0^t |G(\tilde{X}_s, \tilde{Y}_s, X_s, Y_s)|^2 ds\right)\right] < \infty$ so that Novikov's condition is verified. We will give more details on the construction of such G later in Chapter 4.

Example 7.36. We can see the limitations of our methodology by introducing a counterexample: indeed, let's consider a degenerate SDE with logistic drift on the degenerate part:

$$\begin{cases} dX_t &= X_t(1 - X_t)\mathrm{d}t, \\ dY_t &= -Y_t \mathrm{d}t + \mathrm{d}B_t, \end{cases}$$

evolving on $\mathbb{R}_+^2$ and where $(B_t)_{t\geq 0}$ is a one-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Since the drift on the first variable possesses two stationary points in $x = 0$ and $x = 1$, there exists two ergodic probability measures which are of the form $\delta_0 \times \mu$ and $\delta_1 \times \mu$, where μ is the speed measure associated to $dY_t = -Y_t \mathrm{d}t + \mathrm{d}B_t$, so

$$\mu(dy) = \frac{1}{\sqrt{\pi}} e^{-y^2} dy.$$

Indeed, given $d\tilde{X}_t = \tilde{X}_t(1 - \tilde{X}_t)\mathrm{d}t$, it implies that

$$\tilde{X}_t = \frac{\tilde{X}_0 e^t}{1 + \tilde{X}_0(e^t - 1)}, \quad \text{and} \quad X_t = \frac{X_0 e^t}{1 + X_0(e^t - 1)},$$

so that if $X_0 = 0,\ \tilde{X}_0 \neq 0$, we have

$$\rho_1(t) = \tilde{X}_t - X_t = \frac{\tilde{X}_0 e^t}{1 + \tilde{X}_0(e^t - 1)} \underset{t\to\infty}{\longrightarrow} 1,$$

and we lack the first asymptotic coupling assumption from Hypothesis 14.

Moreover, as stated in Remark 7.27, it is impossible to modify the nondegenerate part to avoid this problem: indeed, if $d\tilde{X}_t = \tilde{X}_t(1 - \tilde{X}_t) + G(X_t, Y_t, \tilde{X}_t, \tilde{Y}_t)$, we fail to have the same distribution for X_t and $\tilde{X}_t$: modifications can be made only on the degenerate part of the SDE of interest.

Chapter 8
Asymptotic coupling in practice

8.1 A toy model by Hairer

Given an original SDE and its solution $(Z_t)_{t\geq 0}$ generating a semigroup $(P_t)_{t\geq 0}$, the construction of the modified SDE such that its modified solution $(\tilde{Z}_t)_{t\geq 0}$ and $(Z_t)_{t\geq 0}$ satisfy Hypotheses 14 is the key part of our methodology. And also the pain point: such constructions are not easy to carry out.

To achieve this, we shift our focus to the asymptotic coupling methodology introduced by Hairer in [Hai02]. Our purpose will be to construct a modified version of our SDE by adding a function G to the original drift (to be specified).

Its goal will to "force" both original and modified SDEs to converge to each other. Keeping in mind both assumptions we want to reach, this additional drift has to take effect only on the nondegenerate part: if G acts on the degenerate part, then the possible Markov semigroup generated by the modified SDE will not have the same distribution as the one from the original SDE, as detailed at the end of the last Example 7.36.

Back to the idea of [Hai02], our goal is to split the process by considering the degenerate part of the SDE, called the *low-frequency part*, and the nondegenerate part, called the *high-frequency*, such that:

(i) the low-frequency part, which drives the instability, is finite-dimensional;

(ii) the high-frequency part, which represents the stability of the system, is finite- or infinite-dimensional;

(iii) and the long-time asymptotics of the dynamics driven by high-frequency part is completely dominated by the behavior of the low-frequency part.

The denomination low- and high-frequency parts are also used in subsequent articles such as [Xu11].

Now, let's study the following concrete example, which is a toy model introduced by Hairer in [Hai02].

Example 8.1. We consider the following SDE on $\mathbb{R}^2$,

$$\begin{cases} dX_t = \left(2X_t + Y_t - X_t^3\right) \mathrm{d}t + \mathrm{d}B_t, \\ dY_t = \left(2Y_t + X_t - Y_t^3\right) \mathrm{d}t, \end{cases} \tag{113}$$

where $(B_t)_{t\geq 0}$ is a standard one-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. As before, we write $Z_t = (X_t, Y_t)$ the unique strong solution of the SDE (113), which follows standard results on finite-dimensional degenerate SDE with locally Lipschitz drift where the explosion is controlled by a Lyapunov function, here $V(x,y) = x^2 + y^2$ (see e.g. [Kha12], Theorem 3.5).

If the context of this example allows us to use tools like Hörmander condition to get ergodic properties on the semigroup generated by $(Z_t)_{t\geq 0}$, it highly relies on the existence of a reference measure, here the Lebesgue measure. These tools will be useless in contexts such as infinite dimensions where the noise is sufficiently degenerate, so a reference measure like Lebesgue is unavailable. Furthermore, a classical coupling by change of measure is not possible in this model since both directions are linearly unstable so that a finite meeting time is not reachable.

To achieve the more general asymptotic strong Feller property through asymptotic log–Harnack inequalities, Hairer introduces the idea of asymptotic coupling. More precisely, we are looking for another probability measure $\mathbb{Q}$ and Brownian motion $(\tilde{B}_t)_{t\geq 0}$ defined on $(\Omega, \mathcal{F}, \mathbb{Q})$ so that the solution of the original SDE and this modified SDE,

$$\begin{cases} d\tilde{X}_t = \left(2\tilde{X}_t + \tilde{Y}_t - \tilde{X}_t^3\right) \mathrm{d}t + d\tilde{B}_t, \\ d\tilde{Y}_t = \left(2\tilde{Y}_t + \tilde{X}_t - \tilde{Y}_t^3\right) \mathrm{d}t, \end{cases} \tag{114}$$

denoted $\tilde{Z}_t = (\tilde{X}_t, \tilde{Y}_t)$, satisfy Hypothesis 14. In particular, the notion of asymptotic coupling is used to describe the long-term behavior of Z_t and $\tilde{Z}_t$ so that $\|\tilde{Z}_t^{\tilde{z}} - Z_t^z\| \to 0$ as $t \to \infty$, where $\tilde{Z}_0 = \tilde{z}$ and $Z_0 = z$.

In particular, the passage to the expected value will maintain the convergence toward zero, and Hypothesis 14***(ii)*** holds true.

8.1.1 Proof of Hypothesis 14*(ii)* for Example 8.1

In this part, we are going to focus on the construction of the modified SDE by introducing the asymptotic coupling setting. Indeed, the purpose is to construct a good candidate as well as possible for the modified Brownian motion: this will be realized in the next section.

First, we introduce a function G to be defined such that the modified SDE is now

$$\begin{cases} d\tilde{X}_t = \left(2\tilde{X}_t + \tilde{Y}_t - \tilde{X}_t^3\right) \mathrm{d}t + G(X_t, Y_t, \tilde{X}_t, \tilde{Y}_t)\mathrm{d}t + \mathrm{d}B_t, \\ d\tilde{Y}_t = \left(2\tilde{Y}_t + \tilde{X}_t - \tilde{Y}_t^3\right) \mathrm{d}t. \end{cases}$$

In this setting, our modified Brownian motion would be defined as

$$\tilde{B}_t = B_t + \int_0^t G(X_s, Y_s, \tilde{X}_s, \tilde{Y}_s) ds.$$

If G is small enough, in the sense that $\int_0^{+\infty} \|G(X_s, Y_s, \tilde{X}_s, \tilde{Y}_s)\|^2 ds$ is bounded with a high enough probability, then $\tilde{B}_t$ will be well-defined on a modified probability space $(\Omega, \mathcal{F}, \mathbb{Q})$.

Let's prove the asymptotic convergence of the difference process. We define $\rho(t) = \tilde{Z}_t - Z_t = (\tilde{X}_t - X_t, \tilde{Y}_t - Y_t) = (\rho_1(t), \rho_2(t))$, which is the solution of the following SDE,

$$\begin{cases} d\rho_1(t) = 2\rho_1(t)\mathrm{d}t + \rho_2(t)\mathrm{d}t - \rho_1(t)(X_t^2 + X_t\tilde{X}_t + \tilde{X}_t^2)\mathrm{d}t + G(X_t, Y_t, \tilde{X}_t, \tilde{Y}_t)\mathrm{d}t, \\ d\rho_2(t) = 2\rho_2(t)\mathrm{d}t + \rho_1(t)\mathrm{d}t - \rho_2(t)(Y_t^2 + Y_t\tilde{Y}_t + \tilde{Y}_t^2)\mathrm{d}t. \end{cases}$$

Indeed, we have

$$\begin{aligned} d\rho_1(t) &= d(\tilde{X}_t - X_t) \\ &= \left(2\tilde{X}_t + \tilde{Y}_t - \tilde{X}_t^3\right) \mathrm{d}t - (2X_t\mathrm{d}t + Y_t\mathrm{d}t - X_t^3\mathrm{d}t) + G(X_t, Y_t, \tilde{X}_t, \tilde{Y}_t)\mathrm{d}t \\ &= 2(\tilde{X}_t - X_t)\mathrm{d}t + (\tilde{Y}_t - Y_t)\mathrm{d}t - (\tilde{X}_t^3 - X_t^3)\mathrm{d}t + G(X_t, Y_t, \tilde{X}_t, \tilde{Y}_t)\mathrm{d}t \\ &= 2\rho_1(t)\mathrm{d}t + \rho_2(t)\mathrm{d}t - (\tilde{X}_t^3 - X_t^3)\mathrm{d}t + G(X_t, Y_t, \tilde{X}_t, \tilde{Y}_t)\mathrm{d}t. \end{aligned}$$

It only remains

$$\begin{aligned} -(\tilde{X}_t^3 - X_t^3) &= -(\tilde{X}_t - X_t)(X_t^2 + \tilde{X}_t^2) + \tilde{X}_t X_t^2 - X_t\tilde{X}_t^2 \\ &= -(\tilde{X}_t - X_t)(X_t^2 + X_t\tilde{X}_t + \tilde{X}_t^2) \\ &= -\rho_1(t)(X_t^2 + X_t\tilde{X}_t + \tilde{X}_t^2). \end{aligned}$$

And same arguments still hold for $\rho_2(t)$.

It appears easy to choose G such that $\rho_1(t) \to 0$, which does not imply that $\rho_2(t) \to 0$. So let's focus on ρ_2: in the expression of $\frac{d}{dt}\rho_2(t)$, it appears that if we force $\rho_1(t)$ to be very close from $-3\rho_2(t)$, then

$$d\rho_2(t) = -\rho_2(t)\mathrm{d}t + \varepsilon - \rho_2(t)(Y_t^2 + Y_t\tilde{Y}_t + \tilde{Y}_t^2)\mathrm{d}t,$$

with $\varepsilon > 0$ small enough, then the underlying differential equation is asymptotically stable, and such a G would be a good candidate. We introduce $\xi(t) = \rho_1(t) + 3\rho_2(t)$, which means that

$$d\xi(t) = (\cdots)\mathrm{d}t + G(X_t, Y_t, \tilde{X}_t, \tilde{Y}_t)\mathrm{d}t,$$

with $(\cdots)$ an expression of the order $\|\rho(t)\|(1 + \|Z_t\|^2 + \|\tilde{Z}_t\|^2)$ that we will investigate later.

With $G = -(\cdots) - 2\xi(t)$, it appears that

$$d\xi(t) = -2\xi(t)\mathrm{d}t \;\Rightarrow\; \xi(t) = \xi(0)e^{-2t}.$$

Recalling the definition of $\xi(t) = \rho_1(t) + 3\rho_2(t)$, the equation for $d\rho_2(t)$ rewrites as

$$d\rho_2(t) = -\rho_2(t)\mathrm{d}t + \xi(0)e^{-2t}\mathrm{d}t - \rho_2(t)(Y_t^2 + Y_t\tilde{Y}_t + \tilde{Y}_t^2)\mathrm{d}t,$$

which we can solve by using the variation of the constant method. The homogeneous equation is given by

$$d\rho_2(t) = -\rho_2(t)(1 + Y_t^2 + Y_t\tilde{Y}_t + \tilde{Y}_t^2)\mathrm{d}t \;\Rightarrow\; \rho_2^H(t) = C \cdot e^{-A(Y_t,\tilde{Y}_t)},$$

where $A(Y_t, \tilde{Y}_t) = \int_0^t 1 + Y_s^2 + Y_s\tilde{Y}_s + \tilde{Y}_s^2 ds = t + \int_0^t Y_s^2 + Y_s\tilde{Y}_s + \tilde{Y}_s^2 ds$.

Let's consider the particular solution $\rho_2^P(t) = C(t) \cdot e^{-A(Y_t,\tilde{Y}_t)}$, which has to be solution of

$$d\rho_2^P(t) = -\rho_2^P(t)\mathrm{d}t + \xi(0)e^{-2t}\mathrm{d}t - \rho_2^P(t)(Y_t^2 + Y_t\tilde{Y}_t + \tilde{Y}_t^2)\mathrm{d}t.$$

Replacing by the correct terms and simplifying both sides, the equation rewrites as

$$\frac{d}{\mathrm{d}t}C(t) \cdot e^{-A(Y_t,\tilde{Y}_t)} = \xi(0)e^{-2t} \;\Rightarrow\; \frac{d}{\mathrm{d}t}C(t) = \xi(0)e^{-2t}e^{A(Y_t,\tilde{Y}_t)}.$$

Then, the solution for $C(t)$ is given by $\xi(0)\int_0^t e^{-2s+A(Y_s,\tilde{Y}_s)}ds$, the particular solution writes

$$\rho_2^P(t) = \xi(0)e^{-A(Y_t,\tilde{Y}_t)}\int_0^t e^{-2s+A(Y_s,\tilde{Y}_s)}ds,$$

and the general solution is given by

$$\rho_2(t) = C \cdot e^{-A(Y_t,\tilde{Y}_t)} + \xi(0)e^{-A(Y_t,\tilde{Y}_t)}\int_0^t e^{-2s+A(Y_s,\tilde{Y}_s)}ds,$$

and with the initiation condition, $C = \rho_2(0)$.

We can remark that $A(Y_t, \tilde{Y}_t) = t + \int_0^t Y_s^2 + Y_s\tilde{Y}_s + \tilde{Y}_s^2 ds \geq t$, since the function

$$(x, y) \mapsto x^2 + xy + y^2$$

is positive for $x, y \in \mathbb{R}$. In addition, since for $s \leq t$, we have

$$A(Y_s, \tilde{Y}_s) - A(Y_t, \tilde{Y}_t) = s - t + \int_0^s Y_u^2 + Y_u\tilde{Y}_u + \tilde{Y}_u^2 ds - \int_0^t Y_v^2 + Y_v\tilde{Y}_v + \tilde{Y}_v^2 dv \leq s - t,$$

which implies

$$\begin{aligned}\Big|\int_0^t e^{-2s+A(Y_s,\tilde{Y}_s)-A(Y_t,\tilde{Y}_t)}ds\Big| &\leq \Big|\int_0^t e^{-2s+s-t}ds\Big| \\ &= \Big|e^{-t}\int_0^t e^{-s}ds\Big| \\ &= |e^{-t}(1-e^{-t})| \\ &\leq e^{-t}\end{aligned}$$

We finally can bound $\rho_2(t)$ by

$$\begin{aligned}|\rho_2(t)| &\leq |\rho_2(0)|e^{-A(Y_t,\tilde{Y}_t)} + \Big|\xi(0)\cdot\int_0^t e^{-2s+A(Y_s,\tilde{Y}_s)-A(Y_t,\tilde{Y}_t)}ds\Big| \\ &\leq |\rho_2(0)|e^{-t} + |\xi(0)|e^{-t} \\ &\leq 4(|\rho_1(0)| + \rho_2(0)|)e^{-t},\end{aligned}$$

since $\xi(t) = \rho_1(t) + 3\rho_2(t)$. Also, with the previous bounds for ξ, ρ_2, we get

$$|\rho_1(t)| \leq |\xi(t)| + 3|\rho_2(t)| \leq |\xi(0)|e^{-2t} + 3|\rho_2(0)|e^{-t} + 3|\xi(0)|e^{-t},$$

that we can rewrite as

$$|\rho_1(t)| \leq 4|\xi(0)|e^{-t} + 3|\rho_2(0)|e^{-t} \leq 15(|\rho_1(0)| + |\rho_2(0)|)e^{-t},$$

which leads to the following estimate,

$$\Big\|(\tilde{X}_t^{\tilde{x}}, \tilde{Y}_t^{\tilde{y}}) - (X_t^x, Y_t^y)\Big\|^2 = \rho_1(t)^2 + \rho_2(t)^2 \leq 482\|(\tilde{x},\tilde{y}) - (x,y)\|^2 e^{-2t},$$

and that Hypothesis 14***(ii)*** holds true.

Before proving that Hypothesis 14***(i)*** holds true, we can show the previous claim on the additional drift G, which tells that $G(t) = -(\cdots) - 2\xi(t)$ where $(\cdots)$ is an expression of the order $\|\rho(t)\|(1 + \|Z_t\|^2 + \|\tilde{Z}_t\|^2)$.

Since (...) is defined from the definition of ξ, ρ_1 and ρ_2, it gives us

$$\begin{aligned}(...) &= 2\rho_1(t) + \rho_2(t) - \rho_1(t)(X_t^2 + X_t\tilde{X}_t + \tilde{X}_t^2) + 6\rho_2(t) + 3\rho_1(t) - 3\rho_2(t)(Y_t^2 + Y_t\tilde{Y}_t + \tilde{Y}_t^2) \\ &= 5(\tilde{X}_t - X_t) + 7(\tilde{Y}_t - Y_t) - (\tilde{X}_t - X_t)(X_t^2 + X_t\tilde{X}_t + \tilde{X}_t^2) - 3(\tilde{Y}_t - Y_t)(Y_t^2 + Y_t\tilde{Y}_t + \tilde{Y}_t^2) \\ &= 5(\tilde{X}_t - X_t) + 7(\tilde{Y}_t - Y_t) - \tilde{X}_tX_t^2 - X_t\tilde{X}_t^2 - \tilde{X}_t^3 + X_t^3 + X_t^2\tilde{X}_t \\ &\quad + X_t\tilde{X}_t^2 - 3(\tilde{Y}_tY_t^2 + Y_t\tilde{Y}_t^2 + \tilde{Y}_t^3 - Y_t^3 - Y_t^2\tilde{Y}_t - Y_t\tilde{Y}_t^2) \\ &= 5(\tilde{X}_t - X_t) + 7(\tilde{Y}_t - Y_t) - (\tilde{X}_t^3 - X_t^3) - 3(\tilde{Y}_t^3 - Y_t^3).\end{aligned}$$

We can bound each part of (...): the first part is directly bounded using triangular inequality,

$$\begin{aligned}(|\tilde{X}_t - X_t| + |\tilde{Y}_t - Y_t|)^2 &= |\tilde{X}_t - X_t|^2 + 2|\tilde{X}_t - X_t|\cdot|\tilde{Y}_t - Y_t| + |\tilde{Y}_t - Y_t|^2 \\ &\leq 2|\tilde{X}_t - X_t|^2 + 2|\tilde{Y}_t - Y_t|^2,\end{aligned}$$

which implies

$$\begin{aligned}|\tilde{X}_t - X_t| + |\tilde{Y}_t - Y_t| &\leq \sqrt{2|\tilde{X}_t - X_t|^2 + 2|\tilde{Y}_t - Y_t|^2} \\ &= \sqrt{2}\sqrt{|\tilde{X}_t - X_t|^2 + |\tilde{Y}_t - Y_t|^2} \\ &= \sqrt{2}\|\rho(t)\|.\end{aligned}$$

Then, using

$$a^3 - b^3 = (a-b)(a^2 + ab + b^2) \leq \frac{3}{2}(a-b)(a^2 + b^2)$$

on the second part leads to

$$\begin{aligned} |\tilde{X}_t^3 - X_t^3| + |\tilde{Y}_t^3 - Y_t^3| &\leq \frac{3}{2}\left(|\tilde{X}_t - X_t| \cdot |\tilde{X}_t^2 + X_t^2|\right) + \frac{3}{2}\left(|\tilde{Y}_t - Y_t| \cdot |\tilde{Y}_t^2 + Y_t^2|\right) \\ &\leq \frac{3}{2}\left(|\tilde{X}_t - X_t| + |\tilde{Y}_t - Y_t|\right)\left(\tilde{X}_t^2 + X_t^2 + \tilde{Y}_t^2 + Y_t^2\right) \\ &\leq \frac{3}{\sqrt{2}}\|\rho(t)\|\left(\|\tilde{Z}_t\|^2 + \|Z_t\|^2\right), \end{aligned}$$

where the last step follows from

$$|a| + |b| \leq \sqrt{2} \cdot \sqrt{a^2 + b^2}.$$

Finally, we get the following bound for $(\cdots)$,

$$\begin{aligned} |(\cdots)| &\leq 7\left(|\tilde{X}_t - X_t| + |\tilde{Y}_t - Y_t| + |\tilde{X}_t^3 - X_t^3| + |\tilde{Y}_t^3 - Y_t^3|\right) \\ &\leq \left(7\sqrt{2} + \frac{21}{\sqrt{2}}\right)\|\rho(t)\|\left(1 + \|\tilde{Z}_t\|^2 + \|Z_t\|^2\right), \end{aligned}$$

and in particular G has the form of a locally Lipschitz function.

However, we need to show that the additional drift $G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)$ combined with the original Brownian motion $(B_t)_{t\geq 0}$ produced a new Brownian motion under a probability measure $\mathbb{Q}$ to be constructed, in other words the process defined as

$$d\tilde{B}_t = \mathrm{d}B_t + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\mathrm{d}t$$

is a Brownian motion under a probability measure $\mathbb{Q}$.

We will prove this assumption for the Toy Model of Hairer later in this section, and consider another simplest model to expose the required methodology. Indeed, if the locally Lipschitz form for the additional drift can be undercover, we can firstly focus on the Lipschitz case to avoid computational difficulties.

8.2 Methodology to prove Hypothesis 14*(i)*

Here is an overview of our methodology:

(i) We will show that the new process we constructed is locally a Brownian motion, using tools as Girsanov's theorem 7.4 and the underlying Novikov's condition as stated in 7.6. In particular, we want to use the Doleans-Dade Exponential of the process $(X_t)_{t\geq 0} = (\int_0^t G_s dB_s)_{t\geq 0}$ as stated in Theorem 7.3, that will be denoted $(R_t)_{t\geq 0}$.

(ii) We also recall the the so-called de la Vallée-Poussin Theorem 7.7, which gives us a powerful tool to study the uniform integrability for a family of random variables. Part of the intuition, the particular choice of the function $x \mapsto x\log(x)$ in the application of this result will lead to make the term $\mathbb{E}_{\mathbb{P}}(R_t \log(R_t))$ appear in the computation of (asymptotic) log–Harnack inequality.

(iii) To conclude about the overall existence of a unique probability measure $\mathbb{Q}$ starting from the local ones, we have to think to the Kolmogorov extension Theorem, here Theorem 7.12. The Kolmogorov consistency will be achieved thanks to the martingale property of the process $(R_t)_{t\geq 0}$.

Example 8.2. Let's consider the following SDE on $\mathbb{R}^2$,

$$\begin{cases} dX_t = Y_t, \\ dY_t = f(X_t, Y_t)\mathrm{d}t + \mathrm{d}B_t, \end{cases}$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ is a K-Lipschitz function, and $(B_t)_{t\geq 0}$ is a standard one-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P})$.

As before, we use the asymptotic coupling methodology to construct our modified SDE, so let

$$\begin{cases} d\tilde{X}_t = \tilde{Y}_t, \\ d\tilde{Y}_t = f(\tilde{X}_t, \tilde{Y}_t)\mathrm{d}t + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\mathrm{d}t + \mathrm{d}B_t, \end{cases}$$

where G is a function to be defined. For simplicity, we further denote $G(t)$ for $G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)$. Considering the difference process $\rho(t) = (\rho_1(t), \rho_2(t))$, it is a solution of

$$\begin{cases} d\rho_1(t) = \left(\tilde{Y}_t - Y_t\right)\mathrm{d}t, \\ d\rho_2(t) = f(\tilde{X}_t, \tilde{Y}_t)\mathrm{d}t - f(X_t, Y_t)\mathrm{d}t + G(t)\mathrm{d}t. \end{cases}$$

We are looking to construct G such that $\tilde{Y}_t - Y_t \sim -(\tilde{X}_t - X_t)$: let $\xi(t) = \tilde{Y}_t - Y_t + \tilde{X}_t - X_t = \rho_1(t) + \rho_2(t)$. By differentiate ξ, we get

$$d\xi(t) = (\tilde{Y}_t - Y_t)\mathrm{d}t + \left(f(\tilde{X}_t, \tilde{Y}_t) - f(X_t, Y_t) + G(t)\right)\mathrm{d}t,$$

so that with $G(t) = -\tilde{Y}_t - Y_t - \left(f(\tilde{X}_t, \tilde{Y}_t) - f(X_t, Y_t)\right) - 2\xi(t)$, we have

$$d\xi(t) = -2\xi(t)\mathrm{d}t \;\Rightarrow\; \xi(t) = \xi(0)e^{-2t}.$$

It appears that $d\rho_1(t)$ rewrites as

$$d\rho_1(t) = -\rho_1(t)\mathrm{d}t + \xi(0)e^{-2t}\mathrm{d}t,$$

which has the following solution,

$$\rho_1(t) = \rho_1(0)e^{-t} + \xi(0)e^{-t} - \xi(0)e^{-2t},$$

which can be bounded by

$$|\rho_1(t)| \leq 2\left(|\rho_1(0)| + |\xi(0)|\right)e^{-t} \leq 4\left(|\rho_1(0)| + |\rho_2(0)|\right)e^{-t}.$$

From the definition of $\xi(t) = \rho_1(t) + \rho_2(t) = \xi(0)e^{-2t}$, and the previous result for $\rho_1(t)$, we have

$$\begin{aligned} \rho_2(t) &= \xi(0)e^{-2t} - \rho_1(t) \\ &= \xi(0)e^{-2t} - \rho_1(0)e^{-t} - \xi(0)e^{-t} + \xi(0)e^{-2t} \\ &= 2\xi(0)e^{-2t} - \rho_1(0)e^{-t} - \xi(0)e^{-t}, \end{aligned}$$

which can be equivalently bounded by

$$|\rho_2(t)| \leq 3\left(|\rho_1(0)| + |\xi(0)|\right)e^{-t} \leq 6\left(|\rho_1(0)| + |\rho_2(0)|\right)e^{-t}.$$

Finally, we get the following bound

$$\begin{aligned} \|(\tilde{X}_t^{\tilde{x}}, \tilde{Y}_t^{\tilde{y}}) - (X_t^x, Y_t^y)\|_{\mathbb{R}^2}^2 &= \rho_1(t)^2 + \rho_2(t)^2 \\ &\leq 52\left(|\rho_1(0)| + |\rho_2(0)|\right)^2 e^{-2t} \\ &\leq 104\left(|\tilde{x} - x|^2 + |\tilde{y} - y|^2\right)e^{-2t}, \end{aligned}$$

hence Hypothesis 14***(ii)*** holds true.

Let $K(s) = -G(s)$ and let's consider the Doleans-Dade Exponential of the process $\int_0^t K(s)dB_s$, given by

$$R_t = \exp\left(\int_0^t K(s) \cdot dB_s - \frac{1}{2}\int_0^t |K(s)|^2 ds\right),$$

and let $T > 0$ be fixed.

We define the following stopping time: for $n > \|(x, y)\|$,

$$\tau_n := \inf\{t \geq 0 \mid \|(X_t, Y_t)\| \geq n\}.$$

Due to the non-explosion of (X_t, Y_t), we have that $\tau_n \uparrow \infty$ when $n \uparrow \infty$. And so the process $(X_t, Y_t)_{t\in[0,T\wedge\tau_n]}$ is bounded. The same can be applied to the modified SDE so that the process $(\tilde{X}_t, \tilde{Y}_t)_{t\in[0,T\wedge\tau_n]}$ is bounded.

It implies that on $[0, T \wedge \tau_n]$, Novikov's condition 7.6 holds true since

$$\mathbb{E}\left(\exp\left(\frac{1}{2}\int_0^{T\wedge\tau_n} |G(s)|^2 ds\right)\right) < \infty,$$

so that $(R(t \wedge \tau_n))_{t\in[0,T]}$ is a martingale. Now, we can rewrite both original and modified SDEs with respect to $(\tilde{B}_t)_{t\in[0,T\wedge\tau_n]}$ defined as

$$\tilde{B}_t = B_t - \int_0^t K(s)ds = B_t + \int_0^t G(s)ds, \quad \forall t \in [0, T \wedge \tau_n],$$

which gives us respectively

$$\begin{cases} dX_t = Y_t \mathrm{d}t, \\ dY_t = (f(X_t, Y_t) - G(t))\,\mathrm{d}t + d\tilde{B}_t, \end{cases} \quad \text{and} \quad \begin{cases} d\tilde{X}_t = \tilde{Y}_t, \\ dY_t = f(\tilde{X}_t, \tilde{Y}_t)\mathrm{d}t + d\tilde{B}_t, \end{cases}$$

$\forall t \in [0, T \wedge \tau_n]$ with initial values $X_0 = x$, $Y_0 = y$, $\tilde{X}_0 = \tilde{x}$, and $\tilde{Y}_0 = \tilde{y}$.

We now want to estimate $\mathbb{E}_P\left(R(t \wedge \tau_n)\log(R(t \wedge \tau_n))\right)$ for $t \in [0, T]$. By taking the supremum on $[0, T]$ and n, we have

$$\begin{aligned} \sup_{t\in[0,T],\ n} \mathbb{E}_{\mathbb{P}}\left(R(t \wedge \tau_n)\log(R(t \wedge \tau_n))\right) &= \sup_{t\in[0,T],\ n} \mathbb{E}_{\mathbb{Q}_{T,n}}\left(\log(R(t \wedge \tau_n))\right) \\ &\leq \frac{1}{2}\sup_{t\in[0,T],\ n}\int_0^{T\wedge n} \mathbb{E}_{\mathbb{Q}_{T,n}}\left(|G(s)|^2\right) ds. \end{aligned}$$

With our previous computations and the upper bounds for $\rho_1(t)$ and $\rho_2(t)$, we can bound $G(s)$ by

$$\begin{aligned} |G(s)| &\leq |\tilde{Y}_s - Y_s| + |\left(f(\tilde{X}_s, \tilde{Y}_s) - f(X_s, Y_s)\right)| + 2|\xi(t)| \\ &\leq |\rho_2(t)| + K\left(|\rho_1(t)| + |\rho_2(t)|\right) + 2|\xi(0)|e^{-2t} \\ &\leq 6\left(|\rho_1(0)| + |\rho_2(0)|\right)e^{-t} + 5K\left(|\rho_1(0)| + |\xi(0)|\right)e^{-t} + 2|\xi(0)|e^{-t} \\ &\leq (10K + 6)\left(|\tilde{x} - x| + |\tilde{y} - y|\right)e^{-t}. \end{aligned}$$

Finally, we can bound

$$|G(s)|^2 \leq 2(10K + 6)^2\left(|\tilde{x} - x|^2 + |\tilde{y} - y|^2\right)e^{-2t},$$

so that

$$\begin{aligned}
\sup_{[0,T],n} \mathbb{E}_{\mathbb{P}}\left(R(t\wedge\tau_n)\log(R(t\wedge\tau_n))\right) &\leq \frac{1}{2}\sup_{[0,T],n}\int_0^{T\wedge n} \mathbb{E}_{\mathbb{Q}_{T,n}}\left(|G(s)|^2\right)ds \\
&\leq (10K+6)^2\left(|\tilde{x}-x|^2+|\tilde{y}-y|^2\right)\sup_{[0,T],n}\int_0^{T\wedge n}\mathbb{E}_{\mathbb{Q}_{T,n}}\left(e^{-2s}\right)ds \\
&= (10K+6)^2\left(|\tilde{x}-x|^2+|\tilde{y}-y|^2\right)\sup_{[0,T],n}\left(\frac{1}{2}-\frac{1}{2}e^{-2(T\wedge\tau_n)}\right) \\
&\leq \frac{(10K+6)^2}{2}\left(|\tilde{x}-x|^2+|\tilde{y}-y|^2\right).
\end{aligned}$$

By De la Vallée Poussin's Theorem 7.7 about uniform integrability, the condition

$$\sup_{[0,T],n}\mathbb{E}_{\mathbb{P}}\left(R(t\wedge\tau_n)\log(R(t\wedge\tau_n))\right)<\infty$$

is equivalent to the uniform integrability of $(R(t\wedge\tau_n))_{t\in[0,T]}$.

So, $(R(t\wedge\tau_n))_{t\in[0,T]}$ is a uniformly integrable martingale, and by Girsanov's theorem 7.4, $(\tilde{B}_t)_{t\in[0,T\wedge\tau_n]}$ is a Brownian motion under the probability measure $\mathbb{Q}_{(T,n)}=R(T\wedge\tau_n)\mathbb{P}$.

In particular: for $0\leq s\leq t$, by dominated convergence argument and since $(R(t\wedge\tau_n))_{t\in[0,T]}$ is a martingale, we have

$$\begin{aligned}
\mathbb{E}(R_t\mid\mathcal{F}_s) &= \mathbb{E}(\lim_{n\to\infty}R(t\wedge\tau_n)\mid\mathcal{F}_s) \\
&= \lim_{n\to\infty}\mathbb{E}(R(t\wedge\tau_n\mid\mathcal{F}_s)) \\
&= R(s),
\end{aligned}$$

which means that $(R_t)_{t\in[0,T]}$ is a martingale. Recalling that

$$\begin{aligned}
\sup_{t\in[0,T]}\mathbb{E}(R_t\log(R_t)) &= \sup_{t\in[0,T]}\mathbb{E}_{\mathbb{Q}_T}(\log(R_t)) \\
&= \frac{1}{2}\mathbb{E}_{\mathbb{Q}_T}\left(\int_0^T|G(s)|^2ds\right),
\end{aligned}$$

combining with

$$\begin{aligned}
\liminf_{n\to\infty}\sup_{t\in[0,T]}\mathbb{E}_{\mathbb{Q}_{T,n}}(\log R(t\wedge\tau_n)) &= \liminf_{n\to\infty}\sup_{t\in[0,T]}\frac{1}{2}\mathbb{E}_{Q_{T,n}}\left(\int_0^{t\wedge\tau_n}|G(s)|^2ds\right) \\
&\geq \frac{1}{2}\sup_{t\in[0,T]}\mathbb{E}_{Q_T}\left(\int_0^t|G(s)|^2ds\right),
\end{aligned}$$

where the inequality is a consequence of Fatou's Lemma, we have

$$\sup_{t\in[0,T]}\mathbb{E}(R_t\log(R_t))\leq\frac{(10K+6)^2}{2}\left(|\tilde{x}-x|^2+|\tilde{y}-y|^2\right)<\infty.$$

By De la Vallée Poussin's Theorem 7.7, $(R_t)_{t\in[0,T]}$ is a uniformly integrable martingale, and by Girsanov's theorem 7.4, $(\tilde{B}_t)_{t\in[0,T]}$ is a Brownian motion under the probability measure $\mathbb{Q}_T$ defined as follows,

$$\mathbb{Q}_T=R_T\mathbb{P},$$

since we have $\forall A\in\mathcal{F}_{T\wedge\tau_n}$, $\mathbb{Q}_T(A)=\mathbb{Q}_{T,n}(A)$.

Also, it holds for any $T > 0$: let's extend our sequence of probability measures $(\mathbb{Q}_T)_{T>0}$ to a unique probability measure $\mathbb{Q}$ defined on $(\Omega, \mathcal{F}_\infty)$ such that the following restriction holds true,

$$d\mathbb{Q}|_{\mathcal{F}_T} = R_T d\mathbb{P}|_{\mathcal{F}_T}, \ \forall T > 0.$$

Let $\mathcal{G} \subset \mathcal{F}_T$, and $B \in \mathcal{G}$. Since $\mathbb{Q}_T(B) = \mathbb{E}_{\mathbb{Q}_T}(\mathbf{1}_B) = \mathbb{E}_{\mathbb{P}}(R_T \mathbf{1}_B)$, by definition of conditional expectation we have the relation

$$\mathbb{Q}_T(B) = \mathbb{E}_{\mathbb{P}}(R_T \mathbf{1}_B) = \mathbb{E}_{\mathbb{P}}(\mathbb{E}_{\mathbb{P}}(R_T \mid \mathcal{G})\mathbf{1}_B).$$

By Radon-Nikodym Theorem 7.1, there exists a unique $\mathcal{G}$-measurable function f such that for any $\mathcal{G}$-measurable set B,

$$\mathbb{Q}_T(B) = \int_B f d\mathbb{P},$$

and so that $\mathbb{Q}_T(B) = f\mathbb{P}(B)$. The choice of f is restricted to $\mathbb{E}_{\mathbb{P}}(R_T \mid \mathcal{G})$, since it is a $\mathcal{G}$-measurable function such that

$$\mathbb{Q}_T(B) = \mathbb{E}_{\mathbb{P}}(\mathbb{E}_{\mathbb{P}}(R_T \mid \mathcal{G})\mathbf{1}_B) = \int_B \mathbb{E}_{\mathbb{P}}(R_T \mid \mathcal{G}) d\mathbb{P}, \quad \forall B \in \mathcal{G}.$$

By the martingale property of $(R_t)_{t\in[0,T]}$, $\forall 0 < s \leq t \leq T$, we have

$$\mathbb{Q}_T|_{\mathcal{F}_s} \quad = \quad \mathbb{E}_{\mathbb{P}}(R_t \mid \mathcal{F}_s)\mathbb{P} = R(s)\mathbb{P} = \mathbb{Q}_s.$$

The sequence $(\mathbb{Q}_T)_{T>0}$ respecting Kolmogorov consistency from Definition 7.10, Theorem 7.12 allows us to conclude that there exists a unique probability measure $\mathbb{Q}$ so that

$$\frac{d\mathbb{Q}|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} = R_t, \ \forall t \geq 0,$$

and $(\tilde{B}_t)_{t\geq 0}$ is a Brownian motion by Girsanov's theorem. Finally, Hypothesis 14***(ii)*** is fulfilled by our choice of modified SDE construction. By Theorem 7.33, we have that

$$P_t \log(f(\tilde{z})) \leq \log(P_t f(z)) + \mathbb{E}_{\mathbb{Q}}[R_t \log(R_t)) + \sqrt{104}\|\nabla \log(f)\|_\infty e^{-t}\|\tilde{z} - z\|,$$

holds $\forall t \geq 0$, for all bounded Lipschitz functions $f \geq 1$, and $\forall \tilde{z}, z \in \mathbb{R}^2$.

From our previous computations, we get

$$\mathbb{E}_{\mathbb{Q}}[R_t \log(R_t)) \leq \frac{(10K+6)^2}{2}\|\tilde{z} - z\|^2,$$

so that the asymptotic log–Harnack inequality as defined in (111) holds true with

$$\phi(\tilde{z}, z) = \frac{(10K+6)^2}{2}\|\tilde{z} - z\|^2, \ \Psi(t, \tilde{z}, z) = \sqrt{104}\|\tilde{z} - z\|e^{-t},$$

and allows us to apply Theorem 7.32 since

$$\Lambda(\tilde{z}) = \limsup_{z\to\tilde{z}} \frac{\phi(\tilde{z}, z)}{d(\tilde{z}, z)^2} = \frac{(10K+6)^2}{2} < \infty,$$
$$\Gamma_t(\tilde{z}) = \limsup_{z\to\tilde{z}} \frac{\Psi(t, \tilde{z}, z)}{d(\tilde{z}, z)} = \sqrt{104}\, e^{-t} < \infty,$$

with in particular $\lim_{t\to\infty} \Gamma_t(\tilde{z}) = 0$. We can conclude that $(P_t)_{t\geq 0}$ is asymptotically strong Feller, and possesses at most one invariant probability measure.

8.2.1 Proof of Hypothesis 14*(i)* for Example 8.1

Back to Example 8.1 and similarly as in Example 8.2, let $K(s) = -G(s)$ and let

$$R_t := \exp\left(\int_0^t K_s dB_s - \frac{1}{2}\int_0^t |K(s)|^2 ds\right), \quad \forall t \geq 0,$$

be the Doleans-Dade exponential of the process $\int_0^t K(s)dB_s$.

Let $T > 0$: we define the following stopping time: for $n > \|(x, y)\|$, let

$$\tau_n := \inf\{t \geq 0 \mid \|(X_t, Y_t)\| \geq n\}.$$

Due to the non-explosion of $(X_t, Y_t)_{t\geq 0}$, we have that $\tau_n \uparrow \infty$ when $n \to \infty$ so that $(X_t, Y_t)_{t\in[0,T\wedge\tau_n]}$ is bounded. The same can be applied to the modified SDE so that the process $(\tilde{X}_t, \tilde{Y}_t)_{t\in[0,T\wedge\tau_n]}$ is bounded.

It implies that on $[0, T \wedge \tau_n]$, Novikov's condition holds true, which is

$$\mathbb{E}_{\mathbb{P}}\left(\exp\left(\int_0^{T\wedge\tau_n} \int_0^t |G(s)|^2 ds\right)\right) < \infty,$$

which implies that $(R(t \wedge \tau_n))_{t\in[0,T]}$ is a continuous, uniformly integrable martingale, and by Girsanov's theorem, $(\tilde{B}_t)_{t\in[0,T\wedge\tau_n]}$ is a Brownian motion under the probability measure $\mathbb{Q}_{(T,n)} = R(T \wedge \tau_n)\mathbb{P}$, where

$$\tilde{B}_t = B_t - \int_0^t K(s)ds = B_t + \int_0^t G(s)ds, \quad \forall t \in [0, T \wedge \tau_n].$$

Now, we can rewrite both original SDE (113) and the modified one (114) with respect to $(\tilde{B}_t)_{t\in[0,T\wedge\tau_n]}$ which gives us respectively

$$\begin{cases} dX_t = \left(2X_t + Y_t - X_t^3\right)\mathrm{d}t - G(t)\mathrm{d}t + d\tilde{B}_t, \\ dY_t = \left(2Y_t + X_t - Y_t^3\right)\mathrm{d}t, \end{cases} \quad \text{and} \quad \begin{cases} d\tilde{X}_t = \left(2\tilde{X}_t + \tilde{Y}_t - \tilde{X}_t^3\right)\mathrm{d}t + d\tilde{B}_t, \\ d\tilde{Y}_t = \left(2\tilde{Y}_t + \tilde{X}_t - \tilde{Y}_t^3\right)\mathrm{d}t, \end{cases}$$

$\forall t \in [0, T \wedge \tau_n]$ with initial values $X_0 = x$, $Y_0 = y$, $\tilde{X}_0 = \tilde{x}$, and $\tilde{Y}_0 = \tilde{y}$.

In particular: for $0 \leq s \leq t$, by Dominated Convergence and since $(R(t \wedge \tau_n))_{t\in[0,T]}$ is a martingale, we have

$$\mathbb{E}(R_t \mid \mathcal{F}_s) = \mathbb{E}(\lim_{n\to\infty} R(t \wedge \tau_n) \mid \mathcal{F}_s) = \lim_{n\to\infty} \mathbb{E}(R(t \wedge \tau_n \mid \mathcal{F}_s)) = R(s),$$

which means that $(R_t)_{t\in[0,T]}$ is a martingale.

We can estimate $\mathbb{E}_{\mathbb{P}}(R(t \wedge \tau_n)\log(R(t \wedge \tau_n)))$ for $t \in [0, T]$. By taking the supremum on $[0, T]$ and n, we have

$$\begin{aligned} \sup_{[0,T],n} \mathbb{E}_{\mathbb{P}}(R(t \wedge \tau_n)\log(R(t \wedge \tau_n))) &= \sup_{[0,T],n} \mathbb{E}_{\mathbb{Q}_{T,n}}(\log(R(t \wedge \tau_n))) \\ &\leq \frac{1}{2}\sup_{[0,T],n} \int_0^{T\wedge n} \mathbb{E}_{\mathbb{Q}_{T,n}}\left(|G(s)|^2\right) ds. \end{aligned}$$

Using previous bounds, we have

$$\begin{aligned} |G(t)| &\leq \left(7\sqrt{2} + \frac{21}{\sqrt{2}}\right)\|\rho(t)\|(1 + \|\tilde{Z}_t\|^2 + \|Z_t\|^2) + 2|\xi(0)|e^{-2t} \\ &\leq 35\sqrt{241}\|(\tilde{x}, \tilde{y}) - (x, y)\|e^{-t}(1 + \|\tilde{Z}_t\|^2 + \|Z_t\|^2) + 2|\xi(0)|e^{-2t}, \end{aligned}$$

and since both processes $(Z_t)_{t\geq 0}$, $(\tilde{Z}_t)_{t\geq 0}$ are non-explosive, i.e.

$$\mathbb{E}_{\mathbb{Q}}\left[\sup_{t\in[0,T]}(\|Z_t\|^2)\right) < \infty,\ E_{\mathbb{P}}\left(\sup_{t\in[0,T]}(\|\tilde{Z}_t\|^2)\right) < \infty,$$

then there exist a constant $C > 0$ such that

$$\begin{aligned}
\sup_{[0,T],n}\mathbb{E}_{\mathbb{P}}\left(R(t\wedge\tau_n)\log(R(t\wedge\tau_n))\right) &\leq \frac{1}{2}\sup_{[0,T],n}\int_0^{T\wedge n}\mathbb{E}_{\mathbb{Q}_{T,n}}\left(|G(s)|^2\right)ds \\
&\leq C\|(\tilde{x},\tilde{y})-(x,y)\|^2\sup_{[0,T],n}\int_0^{T\wedge n}\mathbb{E}_{\mathbb{Q}_{T,n}}(e^{-2s}+e^{-4s})ds \\
&= C\|(\tilde{x},\tilde{y})-(x,y)\|^2\sup_{[0,T],n}\left(\frac{3}{4}-\frac{1}{2}(e^{-2(T\wedge n)}+\frac{1}{2}e^{-4(T\wedge n)})\right) \\
&\leq C\|(\tilde{x},\tilde{y})-(x,y)\|^2,
\end{aligned}$$

where $C > 0$ is a constant that changes from line to line. Recalling that

$$\sup_{t\in[0,T]}\mathbb{E}(R_t\log(R_t)) = \sup_{t\in[0,T]}\mathbb{E}_{\mathbb{Q}_T}(\log(R_t)) = \frac{1}{2}\mathbb{E}_{\mathbb{Q}_T}\left(\int_0^T|G(s)|^2ds\right),$$

and combining with

$$\begin{aligned}
\liminf_{n\to\infty}\sup_{t\in[0,T]}\mathbb{E}_{\mathbb{Q}_{T,n}}(\log R(t\wedge\tau_n)) &= \liminf_{n\to\infty}\sup_{t\in[0,T]}\frac{1}{2}\mathbb{E}_{Q_{T,n}}\left(\int_0^{t\wedge\tau_n}|G(s)|^2ds\right) \\
&\geq \frac{1}{2}\sup_{t\in[0,T]}\mathbb{E}_{Q_T}\left(\int_0^t|G(s)|^2ds\right),
\end{aligned}$$

where the inequality is a consequence of Fatou's Lemma, we have

$$\sup_{t\in[0,T]}\mathbb{E}(R_t\log(R_t)) \leq C\|(\tilde{x},\tilde{y})-(x,y)\|^2 < \infty.$$

By De la Vallée Poussin's Theorem 7.7, $(R_t)_{t\in[0,T]}$ is a uniformly integrable martingale, and by Girsanov's theorem, $(\tilde{B}_t)_{t\in[0,T]}$ is a Brownian motion under the probability measure $\mathbb{Q}_T$ defined as follows,

$$\mathbb{Q}_T = R_T\mathbb{P},$$

since we have $\forall A\in\mathcal{F}_{T\wedge\tau_n}$, $\mathbb{Q}_T(A) = \mathbb{Q}_{T,n}(A)$.

Let's extend our sequence of probability measures $(\mathbb{Q}_T)_{T>0}$ to a unique probability measure $\mathbb{Q}$ defined on $(\Omega,\mathcal{F}_\infty)$ such that the following restriction holds true,

$$d\mathbb{Q}|_{\mathcal{F}_T} = R_T d\mathbb{P}|_{\mathcal{F}_T},\ \forall T > 0.$$

As before in Example 8.2, since

$$\mathbb{Q}_T(B) = \mathbb{E}_{\mathbb{P}}(R_T\mathbf{1}_B) = \mathbb{E}_{\mathbb{P}}(\mathbb{E}_{\mathbb{P}}(R_T\mid\mathcal{G})\mathbf{1}_B),$$

then by Radon-Nikodym Theorem, there exists a unique $\mathcal{G}$-measurable function $f := \mathbb{E}_{\mathbb{P}}(R_T\mid\mathcal{G})$ such that for any $\mathcal{G}$-measurable set B, $\mathbb{Q}_T(B) = f\mathbb{P}(B)$. By the martingale property of $(R_t)_{t\in[0,T]}$, the sequence $(\mathbb{Q}_T)_{T>0}$ respects Kolmogorov consistency, and by Theorem 7.12, there exists a unique probability measure $\mathbb{Q}$ so that

$$\frac{d\mathbb{Q}|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} = R_t,\ \forall t\geq 0,$$

and $(\tilde{B}_t)_{t\geq 0}$ is a Brownian motion by Girsanov's theorem. Finally, Hypothesis 14***(ii)*** is fulfilled by our choice of modified SDE construction.

By Theorem 7.33, we have that

$$P_t \log\left(f(\tilde{z})\right) \leq \log\left(P_t f(z)\right) + \mathbb{E}_{\mathbb{Q}}\left[R_t \log(R_t)\right) + \sqrt{482}\|\nabla \log(f)\|_\infty e^{-t}\|\tilde{z} - z\|,$$

holds $\forall t \geq 0$, for all bounded Lipschitz functions $f \geq 1$, and $\forall \tilde{z}, z \in \mathbb{R}^2$. By our previous computations, we have showed the following bound,

$$\mathbb{E}_{\mathbb{Q}}\left[R_t \log(R_t)\right) \leq C\|\tilde{z} - z\|^2,$$

where $C > 0$ is the previous constant, so that

$$\log\left(P_t f(z)\right) + C\|\tilde{z} - z\|^2 + \sqrt{482}\|\nabla \log(f)\|_\infty e^{-t}\|\tilde{z} - z\|.$$

We can then apply Theorem 7.32 so that $(P_t)_{t\geq 0}$ is asymptotically strong Feller, and has at most 1 invariant probability measure. Indeed, our construction leads to the following choice for ϕ and Ψ,

$$\phi(\tilde{z}, z) = C\|\tilde{z} - z\|^2, \ \Psi(t, z, \tilde{z}) = \sqrt{482}\|\tilde{z} - z\|e^{-t},$$

with in particular ϕ, Ψ are measurable symmetric functions such that

$$\Lambda(\tilde{z}) = \lim_{z\to\tilde{z}} \frac{\phi(\tilde{z}, z)}{d(\tilde{z}, z)^2} \leq C < \infty, \quad \Gamma_t(\tilde{z}) = \limsup_{z\to\tilde{z}} \frac{\Psi(t, \tilde{z}, z)}{d(\tilde{z}, z)} = \sqrt{482}e^{-t} < \infty,$$

with in particular $\lim_{t\to\infty} \Psi(t, \tilde{z}, z) = 0$. Thus, conditions of Theorem 7.32 are verified: $(P_t)_{t\geq 0}$ is asymptotically strong Feller and admits at most one invariant probability measure.

Chapter 9

Study of new examples with the asymptotic coupling strategy

Our main purpose is to extend the class of examples on which our asymptotic coupling strategy is effective. We start from Example 8.2, a well-known model which can be associated to Hamiltonian system (see e.g. [RQW24]) or Langevin equation (see e.g. [GRW24]).

9.1 Langevin model in $\mathbb{R}^{n+1}$

We fix integers $n \geq 1$, $m \geq n$ and let $A \in \mathbb{R}^{n\times m}$ a full row rank matrix, rank$(A) = n$, or equivalently AA^T is a symmetric, positive definite matrix. We consider the SDE on $\mathbb{R}^n \times \mathbb{R}^m$,

$$\begin{cases} dX_t &= AY_t \mathrm{d}t, \quad X_0 = x \\ dY_t &= f(X_t, Y_t)\mathrm{d}t + \mathrm{d}B_t, \quad Y_0 = y, \end{cases} \tag{115}$$

where $(B_t)_{t\geq 0}$ is a $m-$dimensional standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and $f : \mathbb{R}^{n+m} \to \mathbb{R}^m$ is supposed to be a Lipschitz function.

Theorem 9.1. *Under the above assumptions, there exists a unique strong solution $(Z_t^z)_{t\geq 0} = (X_t^x, Y_t^y)_{t\geq 0}$ to (115) starting at $Z_0 = z = (x_0, y_0) \in \mathbb{R}^{n+m}$, with Markov semigroup $(P_t)_{t\geq 0}$ defined by*

$$P_t\varphi(z) = \mathbb{E}_{\mathbb{P}}(\varphi(Z_t^z)), \quad \forall\varphi \in \mathcal{B}_b(\mathbb{R}^{n+m}),\ \forall z \in \mathbb{R}^{n+m},$$

such that $(P_t)_{t\geq 0}$ possesses the asymptotic strong Feller property. Moreover, $(P_t)_{t\geq 0}$ has at most one invariant probability measure.

Proof. The existence and uniqueness of a strong solution follows classical results on finite-dimensional SDE with global Lipschitz drift, as well as the existence of such a Markov semigroup $(P_t)_{t\geq 0}$ (see e.g. [Gal16], Theorems 8.3 and 8.6).

Now, let's define the following modified SDE on $\mathbb{R}^n \times \mathbb{R}^m$,

$$\begin{cases} d\tilde{X}_t &= A\tilde{Y}_t \mathrm{d}t, \quad \tilde{X}_0 = \tilde{x} \\ d\tilde{Y}_t &= f(\tilde{X}_t, \tilde{Y}_t)\mathrm{d}t + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\mathrm{d}t + \mathrm{d}B_t, \quad \tilde{Y}_0 = \tilde{y}, \end{cases} \tag{116}$$

where G is a function to be defined later. If the additional drift G is chosen to be Lipschitz, existence and uniqueness of a strong solution for the modified SDE (116) also follows from classical results.

Recall that (see e.g. [HJ12] Theorem 4.2.6 and Theorem 7.3.8. with its proof, in case of real-valued matrix) that since AA^T is a symmetric, strictly positive definite $n\times n$ matrix since A has full row rank and rank$(A) = n$, the eigenvalues of AA^T, denoted $\lambda_i(AA^T)$ for all $i = 1, \cdots, n$, are the squares of the singular values of A, denoted $\sigma_1(A) \geq \cdots \geq \sigma_n(A) > 0$. We let

$$\lambda_{min}(AA^T) = \sigma_n(A)^2, \quad \lambda_{max}(AA^T) = \sigma_1(A)^2, \quad \text{and} \quad \mu = \theta\lambda_{min}(AA^T) > 0.$$

Also (see e.g. [HJ12], Example 5.6.6 and the following remark), the spectral norm $\|\cdot\|_2$ satisfies

$$\|A\|_2 = \sqrt{\lambda_{max}(AA^T)}, \quad \|A^T\|_2 = \|A\|_2,$$

which is induced by the Euclidean vector norm $||\cdot||$ as

$$\sup_{||x||=1} ||Ax||^2 = \sigma_1^2(A) = \lambda_{max}(AA^T).$$

In particular, the spectral norm satisfies

$$||A||_2 := \sup_{||x||\neq 0} \frac{||Ax||}{||x||} = \sup_{||x||=1} ||Ax|| = \sqrt{\lambda_{max}(AA^T)}.$$

We consider the difference process $u(t) = \tilde{X}_t - X_t$ and $v(t) = \tilde{Y}_t - Y_t$, so that

$$du(t) = Av(t)\mathrm{d}t, \quad dv(t) = f(\tilde{X}_t, \tilde{Y}_t)\mathrm{d}t - f(X_t, Y_t)\mathrm{d}t + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\mathrm{d}t.$$

For $\theta > 0$, let $\xi(t) := v(t) + \theta A^T u(t)$, which implies that

$$d\xi(t) = f(\tilde{X}_t, \tilde{Y}_t)\mathrm{d}t - f(X_t, Y_t)\mathrm{d}t + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\mathrm{d}t + \theta A^T Av(t)\mathrm{d}t.$$

For $C > 0$ and I_m the $m \times m$ identity matrix, let

$$G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t) = -\left(f(\tilde{X}_t, \tilde{Y}_t) - f(X_t, Y_t)\right) - \left(\theta A^T A + CId_m\right) v(t) - C\theta A^T u(t),$$

which implies that

$$d\xi(t) = -Cv(t) - C\theta A^T u(t) = -C\xi(t) \;\Rightarrow\; \xi(t) = \xi(0)e^{-Ct}.$$

In particular, we have $v(t) = -\theta A^T u(t) + \xi(t)$ so that

$$du(t) = -\theta AA^T u(t) + A\xi(t) = -\theta AA^T u(t) + A\xi(0)e^{-Ct}.$$

In particular, we fix $C \neq \mu$ to avoid division by zero. By variation of the constant, it gives us

$$u(t) = e^{-\theta AA^T t}u(0) + \int_0^t e^{-\theta AA^T (t-s)}e^{-Cs}A\xi(0)ds.$$

Thus, previous equality is bounded by

$$\begin{aligned} ||u(t)|| &\leq e^{-\mu t}||u(0)|| + \sqrt{\lambda_{max}(AA^T)}\, ||\xi(0)|| \int_0^t e^{-\mu(t-s)}e^{-Cs}ds \\ &\leq e^{-\mu t}||u(0)|| + \sqrt{\lambda_{max}(AA^T)}\, ||\xi(0)|| \frac{e^{-Ct} - e^{-\mu t}}{\mu - C}, \end{aligned}$$

which implies

$$||u(t)|| \leq C_1 e^{-\lambda t}(||u(0)|| + ||\xi(0)||),$$

for $\lambda = \min\{\mu, C\}$ and $C_1 > 0$ a constant depending only on A, θ, and C (not on t). Now, since $v(t) = \xi(t) - \theta A^T u(t)$, it implies that

$$\begin{aligned} ||v(t)|| &\leq ||\xi(0)||e^{-Ct} + \theta||A^T|| \cdot ||u(t)|| \\ &\leq ||\xi(0)||e^{-Ct} + \theta\sqrt{\lambda_{max}(AA^T)}||u(t)|| \\ &\leq C_2 e^{-\lambda t}(||u(0)|| + ||\xi(0)||), \end{aligned}$$

since $||A^T|| = ||A||$, where $C_2 > 0$ is a constant depending on C_1, A, θ, and C. Finally, since

$$||\xi(0)|| = ||v(0) + \theta A^T u(0)|| \leq ||v(0)|| + \theta\sqrt{\lambda_{max}(AA^T)}||u(0)||,$$

we get that

$$\|u(t)\| + \|v(t)\| \leq Ke^{-\lambda t}(\|u(0)\| + \|v(0)\|),$$

where $K > 0$ is a constant depending on C_1, C_2, A, θ, and C. Equivalently, we have

$$\|u(t)\|^2 + \|v(t)\|^2 \leq 2K^2 e^{-2\lambda t}(\|u(0)\|^2 + \|v(0)\|^2),$$

hence Hypothesis 14***(ii)*** holds.

Moreover, with this choice of G, it follows that

$$\|G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\| \leq K_f(\|\tilde{X}_t - X_t\| + \|\tilde{Y}_t - Y_t\|) + (\theta\lambda_{max}(AA^T) + C)\|\tilde{Y}_t - Y_t\| + C\theta\sqrt{\lambda_{max}(AA^T)}\|\tilde{X}_t - X_t\|,$$

where K_f is the Lipschitz constant associated to the function f. In particular, using previous bounds for $\|u(t)\|$ and $\|v(t)\|$, we have

$$\|G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\|^2 \leq \tilde{C}\left((\tilde{x}_0 - x_0)^2 + (\tilde{y}_0 - y_0)^2\right)e^{-2\lambda t},$$

where $\tilde{C} > 0$ is a constant depending on K_f, K, C_1, C_2, A, θ, and C. With exactly the same methodology as in Example 8.2, using Novikov's condition and Girsanov's theorem, it yields the process $(\tilde{B}_t)_{t\geq 0}$ defined by

$$\tilde{B}_t = B_t + \int_0^t G(\tilde{X}_s, \tilde{Y}_s, X_s, Y_s)ds, \quad \forall t \geq 0,$$

is a $\mathbb{Q}-$Brownian motion, where $\mathbb{Q}$ is a new probability constructed as

$$d\mathbb{Q}|_{\mathcal{F}_t} = R_t d\mathbb{P}|_{\mathcal{F}_t}, \quad \forall t \geq 0,$$

and where

$$R_t = \exp\left(\int_0^t -G(\tilde{X}_s, \tilde{Y}_s, X_s, Y_s)dB_s - \frac{1}{2}\int_0^t \|G(\tilde{X}_s, \tilde{Y}_s, X_s, Y_s)\|^2 ds\right), \quad \forall t \geq 0.$$

In particular, since Hypothesis 14 holds true, by Theorem 7.33, there exists an asymptotic log–Harnack inequality of the form

$$\begin{aligned} P_t(\log f)(z) &\leq (\log P_t f)(\tilde{z}) + \mathbb{E}_{\mathbb{P}}(R_t \log R_t) + \|\nabla \log f\|_\infty \|\tilde{z} - z\| e^{-\eta t} \\ &\leq (\log P_t f)(\tilde{z}) + \frac{\tilde{C}}{4\lambda}\|\tilde{z} - z\|^2 + \sqrt{2}K\|\nabla \log f\|_\infty \|\tilde{z} - z\| e^{-\sqrt{\lambda} t}. \end{aligned}$$

for all bounded Lipschitz functions $f \geq 1$. In particular, conditions of Theorem 7.32 are also verified such that $(P_t)_{t\geq 0}$ is asymptotically strong Feller and has at most one invariant probability measure. □

Remark 9.2. The assumption that f is globally Lipschitz can be weakened to a locally Lipschitz condition together with linear growth rate or via a Lyapunov function: the computations are the same but it adds some complexity.

These cases with linear drift on the nondegenerate parts have been widely studied, the most general results may be found in [GW12] who focused on SDE models on $\mathbb{R}^m \times \mathbb{R}^d$ of the form

$$\begin{cases} dX_t = AY_t \mathrm{d}t, \\ dY_t = Z(X_t, Y_t)\mathrm{d}t + \sigma \mathrm{d}B_t, \end{cases} \tag{117}$$

where σ is supposed to be an invertible $d \times d-$matrix, A is a $m \times d-$matrix with rank m, $(B_t)_{t\geq 0}$ is a $d-$dimensional Brownian motion, and $Z_t \in C^1(\mathbb{R}^m \times \mathbb{R}^d, \mathbb{R}^d)$. A model of the form of (117) is referred to *kinetic Fokker-Plank equation* or *stochastic damping Hamiltonian system*.

In [GW12], the authors assume the following: there exists constants $C, \lambda > 0$, an increasing function U on $[0, \infty)$, and W a Lyapunov function such that $\mathcal{L}W \leq CW$ and

$$\|Z(x) - Z(y)\|^2 \leq \|x - y\|^2 \left(U(\|x - y\|) + \lambda W(y)\right), \quad \forall x, y \in \mathbb{R}^{m+d},$$

where

$$\mathcal{L} := \frac{1}{2} \sum_{i,j=1}^{d} (\sigma\sigma^*)_{ij} \frac{\partial^2}{\partial y_i \partial y_j} + \sum_{i=1}^{d} Z(x, y)_i \frac{\partial}{\partial y_i} + \sum_{l=1}^{m} (Ay)_l \frac{\partial}{\partial x_l}.$$

Then,

Theorem 9.3 ([GW12], Theorem 4.3)**.** *Under the above assumptions, there exists $c > 0$ such that*

$$P_t \log f(x) \leq \log P_t f(y) + c\|x - y\|^2 \left(\frac{1}{(1 \wedge t)^3} + \frac{U\left((1 \vee t^{-1})\|x - y\|\right) + W(y)}{t \wedge 1} \right),$$

holds $\forall t \geq 0$,, for all bounded Lipschitz functions $f \geq 1$, and for all $x, y \in \mathbb{R}^{m+d}$.

9.1.1 Extension to infinite-dimensional settings

We follow the methodology introduced in [Wan13] to treat the infinite-dimensional case as an extension of the finite-dimensional one: in particular, let U, H be two separable, real Hilbert spaces and let $A \in \mathcal{L}(U, H)$ be a bounded linear operator such that

$$AA^* \geq \mu Id_H,$$

with $\geq$ in the sense of Loewner order and where $\mu := \left(\inf_{\|u\|_U = 1} \|Au\|_H\right)^2 > 0$, A^* stands for the adjoint operator of A.

Let $\mathbb{H} := H \times U$ be the separable Hilbert space with induced norm $\|(x, y)\|_{\mathbb{H}}^2 = \|x\|_H^2 + \|y\|_U^2$. Now, let's consider the following SDE defined on $\mathbb{H}$,

$$\begin{cases} dX_t &= AY_t \mathrm{d}t, \\ dY_t &= f(X_t, Y_t)\mathrm{d}t + Q^{\frac{1}{2}} dW_t, \end{cases} \tag{118}$$

where $(W_t)_{t \geq 0}$ is a cylindrical Wiener process on U with a Hilbert–Schmidt diffusion operator $\Sigma h = (0, Q^{\frac{1}{2}} h)$ such that $\mathrm{Tr}(Q) < \infty$, and $f : \mathbb{H} \to U$ is supposed to be a globally Lipschitz function. Let $(X_0, Y_0) = (x, y) \in \mathbb{H}$ be an initial condition for (118).

Let $\{e_i\}_{i \geq 1}$ be an orthonormal basis of H and similarly $\{u_i\}_{i \geq 1}$ for U. For $n \geq 1$ fixed, let $\mathcal{P}_n$ be the projection operator from H to $H_n := \mathrm{span}(e_1, \cdots, e_n)$ and similarly for $m \geq 1$, $\mathcal{Q}_m$ from U to $U_m := \mathrm{span}(u_1, \cdots, u_m)$. Now, we can consider the finite-dimensional approximated SDE defined on $\mathbb{H}_{n,m} = H_n \times U_m$ by

$$\begin{cases} dX_t^{(n)} &= A_{n,m} Y_t^{(m)} \mathrm{d}t, \\ dY_t^{(m)} &= f_{n,m}(X_t^{(n)}, Y_t^{(m)})\mathrm{d}t + \mathcal{Q}_m Q^{\frac{1}{2}} dW_t, \end{cases} \tag{119}$$

where $A_{n,m} = \mathcal{P}_n A \mathcal{Q}_m$, $f_{n,m}(x, y) = \mathcal{Q}_m f(\mathcal{P}_n x, \mathcal{Q}_m y)$, with initial condition $(\mathcal{P}_n x, \mathcal{Q}_m y)$.

Lemma 9.4. *Under previous assumptions, there exists a unique mild (respectively strong) solution to (118) (respectively (119)) $(Z_t)_{t \geq 0} = (X_t, Y_t)_{t \geq 0}$ (respectively $(Z_t^{(n,m)})_{t \geq 0} = (X_t^{(n)}, Y_t^{(m)})_{t \geq 0}$) with a Markov semigroup $(P_t)_{t \geq 0}$ defined by*

$$P_t \varphi(z) = \mathbb{E}(\varphi(Z_t^z)), \quad \forall \varphi \in \mathcal{B}_b(\mathbb{H}), \ \forall z \in \mathbb{H},$$

(*respectively* $P_t^{(n,m)} \varphi(z) = \mathbb{E}(\varphi(Z_t^{(n,m)} | Z_0^{(n,m)} = z)), \ \forall \varphi \in \mathcal{B}_b(\mathbb{H}_{n,m}), \ \forall z \in \mathbb{H}_{n,m}$).

Proof. The existence and uniqueness of a strong solution in the finite-dimensional approximation follows from classical results (see e.g. [Gal16]) since the globally Lipschitz drift property is preserved through the approximation.

For (118), we can rewrite using the form of (104): We set the linear operator to be 0, hence $S(t) = I$ the identity operator is a C_0 contraction semigroup, and the drift as $b : \mathbb{H} \to \mathbb{H} : (x, y) \mapsto (Ay, f(x, y))$, which is globally Lipschitz since A is bounded and f is indeed globally Lipschitz. Thus,

$$dZ_t = b(Z_t)\mathrm{d}t + \Sigma dW_t,$$

where the diffusion operator term $\Sigma h = (0, Q^{\frac{1}{2}} h)$ is a Hilbert–Schmidt map since $\mathrm{Tr}(Q) < \infty$. Hence, one can verify that Hypotheses 11 and 12 hold true and by Theorem 7.23, there exists a unique mild solution to (118). □

Remark 9.5. It is also possible to prove it using the conditions from [PZ92], Chapter 7 in case of globally Lipschitz drift.

In the framework of [Wan13], Hypothesis 13 requires that the linear generator $(\mathcal{A}, D(\mathcal{A}))$ has a discrete spectrum with an orthonormal basis of eigenvectors commuting with the semigroup. In our model, the linear part may be taken as $\mathcal{A} = 0$ on $\mathbb{H}$ so that $S(t) \equiv I$ and the entire drift is put into the nonlinear term

$$b : \mathbb{H} \to \mathbb{H} : (x, y) \mapsto (Ay, f(x, y)),$$

so that $\mathrm{spec}(\mathcal{A}) = \{0\}$. Alternately one may take the following bounded block operator

$$\mathcal{A} = \left(\begin{smallmatrix} 0 & A \\ 0 & 0 \end{smallmatrix}\right),$$

defined on $\mathbb{H}$ and for which $S(t) = I + t\mathcal{A}$. Then, $\mathcal{A}^2 = 0$ and $\mathrm{spec}(\mathcal{A}) = \{0\}$, which conclude that no discrete spectrum assumption is verified in both cases.

However, since our purpose is to implement an approximation of the original infinite-dimensional process that converges to it, we can define the finite-rank orthogonal projections $\mathcal{P}_n : H \to H_n$ and $Q_m : U \to U_m$ with $\mathcal{P}_n \to Id_H, Q_m \to Id_U$ strongly, truncate the coefficients and the noise accordingly, and obtain strong convergence of the finite-dimensional solutions to the original solution.

Lemma 9.6. *Let $\{\mathcal{P}_n\}_{n\geq 1}$ and $\{Q_m\}_{m\geq 1}$ be orthogonal projections on H and U with finite rank, such that $\mathcal{P}_n \underset{n\to\infty}{\longrightarrow} Id_H$ and $Q_m \underset{m\to\infty}{\longrightarrow} Id_U$ strongly. Let $\mathcal{R}_{n,m} := \mathcal{P}_n \oplus Q_m : \mathbb{H} \to \mathbb{H}_{n,m}$, the truncated drift*

$$b_{n,m}(z) := \mathcal{R}_{n,m} b(\mathcal{R}_{n,m} z),$$

and the truncated noise operator $\Sigma_{n,m} h := \mathcal{R}_{n,m}\Sigma h = (0, Q_m Q^{1/2} h)$. Consider the finite-dimensional SDE on $\mathbb{H}_{n,m} := H_n \times U_m$,

$$dZ_t^{(n,m)} = b_{n,m}(Z_t^{(n,m)})\mathrm{d}t + \Sigma_{n,m} dW_t, \quad Z_0^{(n,m)} = \mathcal{R}_{n,m} Z_0. \tag{120}$$

Let $(Z_t)_{t\geq 0}$ be the (unique mild) solution to (118). Then for every $T > 0$, under the assumption $\mathbb{E}(\|Z_0\|^2_{\mathbb{H}}) < \infty$,

$$\lim_{n,m\to\infty} \mathbb{E}\Big(\sup_{t\in[0,T]} \big\|Z_t^{(n,m)} - Z_t\big\|^2_{\mathbb{H}} \Big) = 0.$$

Proof. We remark that the unique mild (respectively strong) solution to (118) (respectively (120)) has the uniform moment bound,

$$\sup_{n,m} \mathbb{E}\Big(\sup_{t\in[0,T]} \|Z_t^{(n,m)}\|^2_{\mathbb{H}} \Big) + \mathbb{E}\Big(\sup_{t\in[0,T]} \|Z_t\|^2_{\mathbb{H}} \Big) \leq C_T\Big(1 + \mathbb{E}(\|Z_0\|^2_{\mathbb{H}})\Big),$$

for $C_T > 0$ a constant depending on T, $\|A\|$, the Lipschitz constant of f, and $\|Q^{1/2}\|_{HS}$. It uses either Theorem 7.23 for the infinite-dimensional case, or classical results as in [Gal16] for the finite-dimensional approximation. Let $\Delta_t := Z_t - Z_t^{(n,m)}$. Then,

$$\begin{aligned}\Delta_t &= (Id_{\mathbb{H}} - \mathcal{R}_{n,m})Z_0 + \int_0^t (b(Z_s) - b_{n,m}(Z_s^{(n,m)}))ds + \int_0^t (\Sigma - \Sigma_{n,m})dW_s \\ &= (Id_{\mathbb{H}} - \mathcal{R}_{n,m})Z_0 + \int_0^t (b(Z_s) - b(Z_s^{(n,m)}) + b(Z_s^{(n,m)}) - b_{n,m}(Z_s^{(n,m)}))ds + \int_0^t (\Sigma - \Sigma_{n,m})dW_s \\ &= \int_0^t (b(Z_s) - b(Z_s^{(n,m)}))ds + \int_0^t (Id_{\mathbb{H}} - \mathcal{R}_{n,m})b(Z_s^{(n,m)})ds + \int_0^t (\Sigma - \Sigma_{n,m})dW_s + (Id_{\mathbb{H}} - \mathcal{R}_{n,m})Z_0.\end{aligned}$$

By the Lipschitz property of $z \mapsto b(z)$ combined with Burkholder-Davis-Gundy ([Gal16], Theorem 5.16) and Cauchy–Schwarz inequalities,

$$\begin{aligned}\mathbb{E}\Big(\sup_{t\le T}\|\Delta_t\|_{\mathbb{H}}^2\Big) \le C\int_0^T \mathbb{E}\Big(\sup_{r\le s}\|\Delta_r\|_{\mathbb{H}}^2\Big)ds + CT\int_0^T \mathbb{E}\Big(\Big\|(I - \mathcal{R}_{n,m})b(Z_s^{(n,m)})\Big\|_{\mathbb{H}}^2\Big)ds + C\int_0^T \|(\Sigma - \Sigma_{n,m})\|_{HS}^2 ds \\ + C\mathbb{E}\Big(\|(I - \mathcal{R}_{n,m})Z_0\|_{\mathbb{H}}^2\Big),\end{aligned}$$

for a constant C independent of n, m. Since $\sup_{n,m}\int_0^T \mathbb{E}\Big(\|b(Z_s^{(n,m)})\|_{\mathbb{H}}^2\Big)ds < \infty$ and $\mathcal{R}_{n,m} \to I$ strongly, then by a dominated convergence argument,

$$\int_0^T \mathbb{E}\Big(\Big\|(I - \mathcal{R}_{n,m})b(Z_s^{(n,m)})\Big\|_{\mathbb{H}}^2\Big)ds \underset{n,m\to\infty}{\longrightarrow} 0.$$

Since $\Sigma - \Sigma_{n,m} = (Id_{\mathbb{H}} - \mathcal{R}_{n,m})\Sigma$ with $\Sigma h = (0, Q^{\frac{1}{2}}h)$ and $\mathrm{Tr}(Q) < \infty$, we have

$$\|\Sigma - \Sigma_{n,m}\|_{HS}^2 = \|(Id_u - Q_m)Q^{\frac{1}{2}}\|_{HS}^2 \underset{m\to\infty}{\longrightarrow} 0,$$

by strong convergence of $\mathcal{R}_{n,m} \underset{n,m\to\infty}{\longrightarrow} Id_{\mathbb{H}}$ and definition of the Hilbert–Schmidt norm.

Finally,

$$\mathbb{E}\Big(\|(Id_{\mathbb{H}} - \mathcal{R}_{n,m})Z_0\|_{\mathbb{H}}^2\Big) \underset{n,m\to\infty}{\longrightarrow} 0,$$

since $\mathcal{R}_{n,m} \underset{n,m\to\infty}{\longrightarrow} Id_{\mathbb{H}}$ strongly and $Z_0 \in L^2$.

We can conclude by Gronwall's lemma: indeed, for each $t > 0$ fixed, let $\Phi_{n,m}(t) = \mathbb{E}\Big(\sup_{r\le t}\|\Delta_r\|_{\mathbb{H}}^2\Big)$, then previous inequality rewrites as

$$\Phi_{n,m}(t) \le C\int_0^t \Phi_{n,m}(s)ds + \varepsilon_{n,m}(t),$$

where $\forall t \le T$, we have

$$\varepsilon_{n,m}(t) \le CT\int_0^T \mathbb{E}\Big(\Big\|(Id_{\mathbb{H}} - \mathcal{R}_{n,m})b(Z_s^{(n,m)})\Big\|_{\mathbb{H}}^2\Big)ds + C\int_0^T \|(\Sigma - \Sigma_{n,m})\|_{HS}^2 ds \quad + C\mathbb{E}\Big(\|(Id_{\mathbb{H}} - \mathcal{R}_{n,m})Z_0\|_{\mathbb{H}}^2\Big).$$

By Gronwall's lemma, we have

$$\Phi_{n,m}(T) \le \varepsilon_{n,m}(T)e^{CT},$$

so that $\mathbb{E}\Big(\sup_{t\le T}\|\Delta_t\|_{\mathbb{H}}^2\Big) \to 0$ as $n, m \to \infty$ since $\varepsilon_{n,m}(T) \underset{n,m\to\infty}{\longrightarrow} 0$, for any $T > 0$ fixed, by dominated convergence and the strong convergence of $\mathcal{R}_{n,m} \to I$. □

We can then use both previous Lemmas to prove the following asymptotic log–Harnack inequality.

Theorem 9.7. *Under the above assumptions, the Markov semigroup $(P_t)_{t\geq 0}$ associated to (118) is asymptotically strong Feller, and has at most one invariant probability measure.*

Proof. Existence and uniqueness of a strong solution $(Z_t^{(n,m)})_{t\geq 0}$ to (119) and mild solution to (118) follows from Lemma 9.4.

Now, given the finite-dimension approximation (119), by Theorem 9.1, the following asymptotic log–Harnack inequality holds for all bounded Lipschitz functions $f \geq 1$ on $\mathbb{H}_{n,m}$,

$$P_t^{(n,m)}(\log f)(y) \leq (\log P_t^{(n,m)} f)(x) + C_{n,m}\|x-y\|^2_{\mathbb{H}_{n,m}} + C'_{n,m}\|\nabla \log f\|_\infty \|x-y\|_{\mathbb{H}_{n,m}} e^{-K_{n,m}t}, \tag{121}$$

where $C_{n,m}, C'_{n,m}$ and $K_{n,m}$ are positive fixed constants depending on the model's parameters. In view of the proof of Theorem 9.1, the constants above are dependent to n, m only through the operator norm of $A_{n,m}$. Here, $A_{n,m}$ being the restriction of a bounded operator A, their limits remain upper bounded: however, we need to exhibit a lower bound to avoid the decay rate to collapse.

In particular, for $n \geq 1$ fixed, we can choose m big enough so that $\mathrm{Ran}(\mathcal{P}_n A^*) \subset \mathrm{Ran}(Q_m)$, which implies that $\lambda_{min}(A_{n,m}A^*_{n,m}) \geq \mu$, and the decay rate $\mu_{n,m}$ of each approximation is bounded below by $\mu > 0$ (see e.g. [Con07], Theorem 2.7).

It implies that there exists $C, \tilde{C}$, and K strictly positive fixed constants such that

$$\lim_{n,m\to\infty} C_{n,m} = C, \quad \lim_{n,m\to\infty} \tilde{C}_{n,m} = \tilde{C}, \quad \text{and} \lim_{n,m\to\infty} K_{n,m} = K.$$

Let φ be a bounded L–Lipschitz function on $\mathbb{H}$ such that $\varphi \geq 1$, and let's consider the inclusion $\iota_{n,m} : \mathbb{H}_{n,m} \to \mathbb{H}$ and let $\varphi_{n,m} := \varphi \circ \iota_{n,m}$ a $\mathbb{H}_{n,m}-$ valued function. In particular, $\varphi_{n,m}$ inherits the properties from φ: indeed, boundness follows

$$\varphi_{n,m}(z_{n,m}) = \varphi(\iota_{n,m} z_{n,m}) = \varphi(z) < +\infty,$$

for all $z_{n,m} \in \mathbb{H}_{n,m}$ since $\iota_{n,m} z_{n,m}$ is an element z of $\mathbb{H}$. Similarly, $\varphi_{n,m}$ is bounded below by 1 since

$$\varphi_{n,m}(z_{n,m}) = \varphi(\iota_{n,m} z_{n,m}) = \varphi(z) \leq 1,$$

for all $z_{n,m} \in \mathbb{H}_{n,m}$ where $z \in \mathbb{H}$ is the same representative of $\iota_{n,m} z_{n,m}$ as before. Moreover, the same global Lipschitz constant L holds for $\varphi_{n,m}$ since

$$|\varphi_{n,m}(x_{n,m}) - \varphi_{n,m}(y_{n,m})| = |\varphi(\iota_{n,m} x_{n,m}) - \varphi(\iota_{n,m} y_{n,m})| \leq L\|\iota_{n,m}\| \cdot \|x_{n,m} - y_{n,m}\|_{\mathbb{H}_{n,m}} = L\|x_{n,m} - y_{n,m}\|_{\mathbb{H}_{n,m}},$$

since $\iota_{n,m}$ is the identity embedding so that $\|\iota_{n,m}\| = 1$.

Finally, $\lim_{n,m\to\infty} \varphi \circ \iota_{n,m} = \varphi$ pointwise by continuity of φ and $\lim_{n,m\to\infty} \iota_{n,m}(z) = z$ for all $z \in \mathbb{H}$. By dominated convergence, Lemma 9.6, and starting from (121), we obtain

$$\begin{aligned}
P_t \log \varphi(z) &= \lim_{n,m\to\infty} P_t^{(n,m)}(\log \varphi_{n,m})(\mathcal{R}_{n,m} z) \\
&\leq \lim_{n,m\to\infty} (\log P_t^{(n,m)} \varphi_{n,m})(\mathcal{R}_{n,m} \tilde{z}) + \lim_{n,m\to\infty} C_{n,m}\|\mathcal{R}_{n,m} z - \mathcal{R}_{n,m}\tilde{z}\|^2_{\mathbb{H}_{n,m}} \\
&\quad + \lim_{n,m\to\infty} C'_{n,m}\|\nabla \log \varphi_{n,m}\|_\infty \|\mathcal{R}_{n,m} z - \mathcal{R}_{n,m}\tilde{z}\|_{\mathbb{H}_{n,m}} e^{-K_{n,m}t} \\
&\leq \log P_t \varphi(\tilde{z}) + C\|z-\tilde{z}\|^2_{\mathbb{H}} + C'\|\nabla \log \varphi\|_\infty \|z-\tilde{z}\|_{\mathbb{H}} e^{-Kt},
\end{aligned}$$

for all $z, \tilde{z} \in \mathbb{H}$. By Theorem 7.32, it implies that $(P_t)_{t\geq 0}$ is asymptotically strong Feller, and has at most one invariant probability measure. □

9.1.2 The semi-linear drift case

The infinite-dimensional setting introduced above has also been investigated in view of adding non-linearties in our model of interest. However, linearity in the high-frequency component remains crucial so we consider a semi-linear drift. Roughly speaking, we suppose that the drift on high-frequency dynamics has a linear part and a non-linear one.

We will make the proper assumptions to ensure that the linear part is stronger than the nonlinearities and control them.

In [HM11], the authors investigated non-linear equations on Hilbert space $\mathcal{H}$ of the form

$$\partial_t u(x,t) + Lu(x,t) = N(u)(x,t) + \sum_{k=1}^{d} g_k(x)\dot{W}_k(t), \tag{122}$$

where L is a positive, self-adjoint operator, N is a polynomial nonlinearity in the sense that there exists $m \geq 1$ such that $N(u) = \sum_{k=1}^{m} N_k(u)$, where N_k is $k-$multilinear, $\{g_k\}_{1\leq k\leq d}$ is a finite collection of smooth, time independent functions dictating the directions in which the randomness is injected, and $\{W_k\}_{1\leq k\leq d}$ is a collection of mutually independent one-dimensional Wiener processes. Here, $\dot{W}_k$ has to be interpreted as the formal derivative of independent Wiener processes through classical Itô calculus. Furthermore, the authors assume the following:

Hypothesis 15. There exist $a \in [0,1[$ and $\gamma_\star, \beta_\star > -a$ with $\gamma_\star + \beta_\star > -1$ such that

(i) The operator $L : \mathcal{D}(L) \to \mathbb{H}$ is self-adjoint and satisfies $\langle u, Lu\rangle \geq \|u\|^2$. We denote $\mathcal{H}_a$, for $a \in \mathbb{R}$, as the domain of L^a with the graph norm, and $\mathcal{H}_{-a}$ its dual with respect to $\mathcal{H}$.

(ii) There exists $n \geq 1$ such that the nonlinearities from N belong to $\text{Poly}^n(\mathcal{H}_{\gamma+a}, \mathcal{H}_\gamma)$[1] for every $\gamma \in [-a, \gamma_\star[$.

(iii) For every $\beta \in [-a, \beta_\star[$, there exists $\gamma \in [0, \gamma_\star + 1[$ such that the adjoint $DN^*(u)$ of the derivative $DN(u)$ in $\mathcal{H}$ can be extended to a continuous map from $\mathcal{H}_\gamma$ to $\mathcal{L}(\mathcal{H}_{\beta+\alpha}, \mathcal{H}_\beta)$.

(iv) For every $1 \leq k \leq d$, $g_k \in \mathcal{H}_{\gamma_\star+1}$.

Hypothesis 16. The operator L has compact resolvent and there exists a measurable function $V : \mathcal{H} \to \mathbb{R}_+$ such that there exist constants $c, \alpha > 0$ such that

$$V(u) \geq c\|u\|^\alpha,$$

for all $u \in \mathcal{H}$. In addition, there exists a constant $C > 0$ and $\eta' \in [0,1[$ such that

$$\mathbb{E}\left[e^{V(u_1)}\right] \leq e^{\eta' V(u_0)}.$$

Hypothesis 17. For every $p > 0$ and every $\delta > 0$, there exists a constant C such that

$$\sup_{t\in[0,1]} \mathbb{E}(\|u_t\|^p) \leq Ce^{\delta V(u_0)}, \quad \mathbb{E}\left(\sup_{s,t\in[0,1]} \|J_{s,t}\|^p\right) \leq Ce^{\delta V(u_0)}, \text{ and } \sup_{s,t\in[0,1]} \mathbb{E}(\|J^{(2)}_{s,t}\|^p) \leq Ce^{\delta V(u_0)},$$

for all $u_0 \in \mathcal{H}$, where $J_{s,t}\varphi$ refers to the Jacobian induced by (122) in the direction $\varphi \in \mathcal{H}$, and $J^{(2)}_{s,t}(\varphi, \psi)$ is the second derivative of u_t in the directions φ and ψ.

We state an Hörmander-like bracket condition in the infinite-dimensional settings:

[1] $\text{Poly}^n(X)$ is the set of continuous maps $P : X \to X$ such that $P(U) = \sum_{k=1}^{n} P^{(k)}(u)$.

Hypothesis 18. The directions g_i are in $\mathcal{H}_\infty$[1], and the linear span $A_\infty = \cup_{n>0} A_n$ is dense in $\mathcal{H}$, where A_n is defined recursively by

$$A_0 = \{g_j : j = 1, \dots, d\}, \quad A_{k+1} = A_k \cup \{N_m(h_1, \cdots, h_m) : h_j \in A_k\}.$$

Then:

Theorem 9.8 ([HM11], Theorem 8.1)**.** *Consider (122) and assume Hypotheses 15-18. Then, the Markov semigroup $(P_t)_{t\geq 0}$ induced by (122) is asymptotically strong Feller.*

More recent results using the asymptotic log–Harnack inequalities setting have been also developped in [HLL20]. Let $\mathbb{H}, \langle\cdot,\cdot\rangle_{\mathbb{H}}, \|\cdot\|_{\mathbb{H}}$ be a real separable Hilbert space identified with its dual space $\mathbb{H}$. Let A be a non-negative, self-adjoint linear operator, and let $V = D(A^{\frac{1}{2}})$ which is a reflexive Banach space with norm $\|u\|_V = \|A^{\frac{1}{2}} u\|_{\mathbb{H}}$, continuously and densely embedded into $\mathbb{H}$.

We suppose that there exists an orthonormal basis $\{e_k\}_{k\geq 1}$ for $\mathbb{H}$ and an increasing eigenvalue sequence $0 \leq \lambda_1 \leq \cdots \leq \lambda_n \leq \cdots$ for eigenvectors of A and such that $\lim_{n\to\infty} \lambda_n = \infty$.

We obtain the Gelfand triple $V \subset \mathbb{H} \equiv \mathbb{H}^* \subset V^*$, where V^* is the dual space of V: let $\langle\cdot,\cdot\rangle$ be the dualization between V and V^* such that $\langle u, v\rangle = \langle u, v\rangle_{\mathbb{H}}$ for $u \in \mathbb{H},\ v \in V$.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a complete, filtered probability space and $(L_2(\mathbb{H}, \mathbb{H}), \|\cdot\|_{L^2})$ be the space of all Hilbert–Schmidt operators from $\mathbb{H}$ to $\mathbb{H}$.

For $u \in \mathbb{H}$, we denote $u_k = \langle u, e_k\rangle_{\mathbb{H}}$, $\forall k \geq 1$. For any $N \in \mathbb{N}$, we define the projection $P_N : \mathbb{H} \to \mathbb{H}$ by

$$P_N u = \sum_{|k|\leq N} u_k e_k, \quad u \in \mathbb{H}.$$

Now, we consider the following SDE on $\mathbb{H}$,

$$du(t) = -Au(t) + F(u(t))\mathrm{d}t + B(u(t))dW(t), \quad u(0) = x \in \mathbb{H}, \tag{123}$$

where $W(t)$ is a cylindrical Wiener process in $\mathbb{H}$ defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, and let $F : V \to V^*$, $B : V \to L_2(\mathbb{H}, \mathbb{H})$. We suppose the following:

Hypothesis 19. For all $u, v, w \in V$, assume that F and B satisfy:

(i) The mapping $s \mapsto \langle F(u + sv), w\rangle$ is continuous on $\mathbb{R}$. There exist constants $K_1, C, \gamma > 0$ such that the mapping $F : V \to V^*$ is one-sided Lipschitz in the sense that

$$2\langle F(u) - F(v), u - v\rangle \leq K_1 \|u - v\|^2_{\mathbb{H}},$$

and satisfies

$$\|F(u)\|_{V^*} \leq C(1 + \|u\|_V)(1 + \|u\|^\gamma_{\mathbb{H}}).$$

(ii) The mapping $B : V \to L_2(\mathbb{H}, \mathbb{H})$ is bounded and Lipschitz in the sense that there exists a constant $K_2 > 0$ such that

$$\sup_{u\in V} \|B(u)\|^2_{L_2} < \infty, \quad \text{and} \quad \|B(u) - B(v)\|^2_{L_2} \leq K_2 \|u - v\|^2_{\mathbb{H}}.$$

(iii) There exists a constant $N_0 \in \mathbb{N}$ such that, for all $u \in \mathbb{H}$, we have $P_{N_0}\mathbb{H} \subset \text{Range } B(u)$ and $B(u)x = 0$ if $x \in (Id - P_{N_0})\mathbb{H}$. Moreover, the corresponding pseudo-inverse operator $B(u)^{-1} : P_{N_0}\mathbb{H} \to P_{N_0}\mathbb{H}$ is uniformly bounded in the sense that there exists $\eta > 0$ such that

$$\sup_{u\in V} \|B(u)^{-1}\|_{L(P_{N_0}\mathbb{H}, P_{N_0}\mathbb{H})} \leq \eta,$$

where $L(P_{N_0}\mathbb{H}, P_{N_0}\mathbb{H})$ denotes the space of all linear bounded operators from $P_{N_0}\mathbb{H}$ to $P_{N_0}\mathbb{H}$.

[1] $\mathcal{H}_\infty = \cap_{a>0} \mathcal{H}_a$ where $\mathcal{H}_a$ is the domain of L^a endowed with the graph norm.

Remark 9.9. In fact, the last assumption can be seen as a nondegeneracy condition on the low-frequency dynamics in case of global degenerate SDE. In particular, we assumed that the degenerate part is well reinjected in the nondegenerate part and, combined with the control from the linear part on the degenerate part, asymptotic coupling construction will be efficient.

In addition, combining the second and third points assumed above, there exist constants β_1, β_2 such that

$$2\langle -Au + F(u), u\rangle + \|B(u)\|_{L_2}^2 \leq \beta_1 - 2\|u\|_V^2 + \beta_2\|u\|_{\mathbb{H}}^2. \tag{124}$$

Then,

Theorem 9.10 ([HLL20], Theorems 2.1 and 2.2)**.** *Under Hypothesis 19, and if in addition $\lambda_{N_0} > \frac{K_1+K_2}{2}$, then for any $x, y \in \mathbb{H}$ and for all bounded Lipschitz functions $f \geq 1$, we have*

$$P_t \log f(y) \leq \log P_t f(x) + \frac{\lambda_{N_0}^2 \eta^2}{4\left(\lambda_{N_0} - \frac{K_1+K_2}{2}\right)}\|x-y\|_{\mathbb{H}}^2 + e^{-\left(\lambda_{N_0} - \frac{K_1+K_2}{2}\right)t}\|x-y\|_{\mathbb{H}}^2 \|\nabla \log f\|_\infty, \quad \forall t > 0.$$

In particular, due to Theorem 7.32, the Markov semigroup $(P_t)_{t\geq 0}$ induced by (123) is asymptotically strong Feller.

Moreover, if the embedding $V \subset \mathbb{H}$ is compact and $\beta_2 < 2\lambda_1$ where β_2 follows from (124), then there exists a unique invariant probability measure.

Unfortunately, there is no example in [HLL20] of any models such that Hypothesis 19 is verified. However, based our asymptotic coupling settings, let's consider the following toy model

$$\begin{cases} dX_t &= (Y_t + f(X_t, Y_t))\mathrm{d}t, \quad X_0 = x, \\ dY_t &= g(X_t, Y_t)\mathrm{d}t + \mathrm{d}B_t, \quad Y_0 = y, \end{cases} \tag{125}$$

on $\mathbb{R}^2$, where $(B_t)_{t\geq 0}$ is a 1-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, and $f, g \in C_b^{1,2}(\mathbb{R}^2, \mathbb{R})$ which are globally Lipschitz continuous with Lipschitz constants $K_f < 1$ and K_g. In addition, we suppose that

$$\inf_{(x,y)\in\mathbb{R}^2} \partial_y f(x,y) \geq c > -1,$$

ensuring that $\frac{1}{1+\partial_y f(x,y)}$ is strictly positive. We suppose also that $\partial_y f$ is globally Lipschitz, hence $\partial_{yy} f$ is a bounded function.

Theorem 9.11. *In the setting of* (125) *with above assumptions, there exists a unique strong solution $Z_t = (X_t, Y_t)$ starting from $z = (x, y) \in \mathbb{R}^2$, with Markov semigroup*

$$P_t\varphi(z) = \mathbb{E}[\varphi(Z_t^z)], \quad \varphi \in \mathcal{B}_b(\mathbb{R}^2),\ z \in \mathbb{R}^2,\ t \geq 0.$$

Moreover, if $\mathbb{E}\left(\|Z_0\|^2\right) < \infty$, then $(P_t)_{t\geq 0}$ is asymptotically strong Feller and has at most one invariant probability measure.

Proof. Existence and uniqueness of a strong solution follow from classical results for SDEs with globally Lipschitz drift (see e.g. [Gal16]).

We consider the following modified version of (125),

$$\begin{cases} d\tilde{X}_t &= (\tilde{Y}_t + f(\tilde{X}_t, \tilde{Y}_t))\mathrm{d}t, \\ d\tilde{Y}_t &= (g(\tilde{X}_t, \tilde{Y}_t) + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t))\mathrm{d}t + \mathrm{d}B_t, \end{cases} \tag{126}$$

starting at $(\tilde{X}_0, \tilde{Y}_0) = (\tilde{x}, \tilde{y}) \in \mathbb{R}^2$ and where G will be chosen later: in particular, we claim that G is globally Lipschitz so that (126) has a unique strong solution as in Example 8.1. Now, let's set the difference process

$$\rho_1(t) := \tilde{X}_t - X_t, \quad \rho_2(t) := \tilde{Y}_t - Y_t,$$

and

$$\xi(t) := \rho_1(t) + \rho_2(t) + (f(\tilde{X}_t, \tilde{Y}_t) - f(X_t, Y_t)),$$

which yields

$$d\rho_1(t) = (\rho_2(t) + f(\tilde{X}_t, \tilde{Y}_t) - f(X_t, Y_t))\mathrm{d}t = -\rho_1(t)\mathrm{d}t + \xi(t)\mathrm{d}t. \tag{127}$$

By Itô's formula, we have

$$\begin{aligned}
d\xi(t) =& \left(g(\tilde{X}_t, \tilde{Y}_t) - g(X_t, Y_t) + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\right)\mathrm{d}t + df(\tilde{X}_t, \tilde{Y}_t) - df(X_t, Y_t) \\
&+ (\tilde{Y}_t - Y_t)\mathrm{d}t + \left(f(\tilde{X}_t, \tilde{Y}_t) - f(X_t, Y_t)\right)\mathrm{d}t \\
=& \left(g(\tilde{X}_t, \tilde{Y}_t) - g(X_t, Y_t) + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\right)\mathrm{d}t + \left(\tilde{Y}_t - f(\tilde{X}_t, \tilde{Y}_t)\right)\partial_x f(\tilde{X}_t, \tilde{Y}_t)\mathrm{d}t \\
&- (Y_t - f(X_t, Y_t))\,\partial_x f(X_t, Y_t)\mathrm{d}t + \left(g(\tilde{X}_t, \tilde{Y}_t) + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\right)\partial_y f(\tilde{X}_t, \tilde{Y}_t)\mathrm{d}t \\
&- g(X_t, Y_t)\partial_y f(X_t, Y_t)\mathrm{d}t + \frac{1}{2}\partial_{yy} f(\tilde{X}_t, \tilde{Y}_t)\mathrm{d}t - \frac{1}{2}\partial_{yy} f(X_t, Y_t)\mathrm{d}t \\
&+ \left(\partial_y f(\tilde{X}_t, \tilde{Y}_t) - \partial_y f(X_t, Y_t)\right)\mathrm{d}B_t + (\tilde{Y}_t - Y_t)\mathrm{d}t + \left(f(\tilde{X}_t, \tilde{Y}_t) - f(X_t, Y_t)\right)\mathrm{d}t.
\end{aligned}$$

Let $\sigma_t := \partial_y f(\tilde{X}_t, \tilde{Y}_t) - \partial_y f(X_t, Y_t)$: we fix a parameter $\kappa > 0$ so that letting

$$\begin{aligned}
G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t) :=& \frac{1}{1 + \partial_y f(\tilde{X}_t, \tilde{Y}_t)}\Big(g(\tilde{X}_t, \tilde{Y}_t) - g(X_t, Y_t) - (\tilde{Y}_t - f(\tilde{X}_t, \tilde{Y}_t))\partial_x f(\tilde{X}_t, \tilde{Y}_t) \\
&+ (Y_t - f(X_t, Y_t))\partial_x f(X_t, Y_t) - g(\tilde{X}_t, \tilde{Y}_t)\partial_y f(\tilde{X}_t, \tilde{Y}_t) + g(X_t, Y_t)\partial_y f(X_t, Y_t) \\
&- \tfrac{1}{2}\partial_{yy} f(\tilde{X}_t, \tilde{Y}_t) + \tfrac{1}{2}\partial_{yy} f(X_t, Y_t) - (\tilde{Y}_t - Y_t) - (f(\tilde{X}_t, \tilde{Y}_t) - f(X_t, Y_t)) - \kappa\xi(t)\Big),
\end{aligned}$$

all the drift terms in $d\xi(t)$ cancel and we get the scalar SDE

$$d\xi_t = -\kappa\xi_t\mathrm{d}t + \sigma_t\mathrm{d}B_t. \tag{128}$$

In particular, $\mathbb{E}(\xi_t) = \xi_0 e^{-\kappa t}$. We also remark that our construction of G leads to a globally Lipchitz function, which proves previous claim.

Moreover, since $\partial_y f$ is globally Lipschitz with constant L_y, then

$$|\sigma_t| \;\leq\; L_y(|\rho_1(t)| + |\rho_2(t)|). \tag{129}$$

By (127) and applying Cauchy–Schwarz inequality, we obtain

$$d\mathbb{E}\left(\rho_1^2(t)\right) = -2\mathbb{E}\left(\rho_1(t)^2\right)\mathrm{d}t + 2\mathbb{E}\left(\rho_1(t)\xi(t)\right)\mathrm{d}t \;\leq\; -\mathbb{E}\left(\rho_1(t)^2\right)\mathrm{d}t + \mathbb{E}\left(\xi^2(t)\right)\mathrm{d}t.$$

From (128) and Itô,

$$d\mathbb{E}\left(\xi^2(t)\right) = -2\kappa\mathbb{E}\left(\xi^2(t)\right)\mathrm{d}t + \mathbb{E}\left(|\sigma_t|^2\right)\mathrm{d}t.$$

One can also relate ρ_2 to ρ_1 and ξ since

$$|\rho_2(t)| \leq |\xi(t)| + |\rho_1(t)| + \left|f(\tilde{X}_t, \tilde{Y}_t) - f(X_t, Y_t)\right|$$

so

$$(1 - K_f)|\rho_2(t)| \;\leq\; |\xi(t)| + (1 + K_f)|\rho_1(t)| \;\Rightarrow\; \rho_2(t)^2 \;\leq\; a\xi(t)^2 + b\rho_1(t)^2, \tag{130}$$

with
$$a := \frac{2}{(1-K_f)^2} > 0, \text{ and } b := \frac{2(1+K_f)^2}{(1-K_f)^2} > 0.$$

Let $u(t) := \mathbb{E}\left(\rho_1(t)^2\right)$, $v(t) := \mathbb{E}\left(\xi(t)^2\right)$. By combining (129) and (130), we obtain the closed differential inequality

$$\begin{cases} u'(t) \;\leq\; -u(t) + v(t), \\ v'(t) \;\leq\; -2\kappa v(t) + 2L_y^2(1+b)u(t) + 2L_y^2 a v(t). \end{cases} \tag{131}$$

Define the linear comparison system

$$\dot{y}(t) = My(t), \;\; \text{with } \; y(0) = \begin{pmatrix} u(0) \\ v(0) \end{pmatrix} \text{ and } \; M := \begin{pmatrix} -1 & 1 \\ 2L_y^2(1+b) & -2\kappa + 2L_y^2 a \end{pmatrix}.$$

It implies that $y(t) = e^{tM}y(0)$. Since M has nonnegative off–diagonal entries, it is a Metzler matrix so that the vector field $F(y) = My$ respects

$$\frac{\partial F_i}{\partial y_j} = M_{ij} \geq 0, \quad \forall i \neq j,$$

or equivalently we say that F is cooperative. By the Kamke comparison principle for cooperative systems (see e.g. [Hir89] or originally [Kam32]), the flow induced by $\dot{y}(t) = My(t)$ is order-preserving and even strongly monotone since M is irreducible, which follows that the off-diagonal terms are strictly positive. Hence, in our case, the componentwise inequalities (131) imply

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} \;\leq\; \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} =: y(t), \tag{132}$$

for all $t \geq 0$. We now check stability of M. For a 2×2 matrix, M is said to be a Hurwitz matrix if and only if
$$\operatorname{tr}(M) < 0 \quad \text{and} \quad \det(M) > 0.$$

This ensures exponential stability of the linear comparison system. Here

$$\operatorname{tr}(M) = -1 - 2\kappa + 2L_y^2 a, \quad \det(M) = 2\kappa - 2L_y^2(a+1+b).$$

Hence a sufficient condition is
$$\kappa \;>\; L_y^2(a+1+b),$$

so that M is Hurwitz. It implies by Lyapunov stability (see e.g. [Hes18], Theorem 8.2) that there exist constants $C \geq 1$ and $\lambda > 0$ depending only on L_y, K_f, κ such that

$$\|e^{tM}\| \;\leq\; Ce^{-\lambda t} \quad \text{for all } t \geq 0,$$

which holds for any induced norm while the constants may change. Therefore, by considering the ℓ^1–norm which is $\|x\|_1 = |x_1| + |x_2|$ and recalling that $y_1(t) \geq u(t) \geq 0$ and similarly for $y_2(t)$, then $y(t)$ satisfies
$$y_1(t) + y_2(t) \;\leq\; Ce^{-\lambda t}(u(0) + v(0)).$$

By the comparison step (132), $u(t) + v(t) \leq y_1(t) + y_2(t)$, hence

$$u(t) + v(t) \;\leq\; Ce^{-\lambda t}(u(0) + v(0)).$$

Finally, by definition of u and v, we obtain

$$\mathbb{E}(|\rho_1(t)|^2 + |\rho_2(t)|^2) \;\leq\; C'e^{-\lambda t}(\rho_1(0)^2 + \rho_2(0)^2),$$

for some $C' > 0$ a constant, from which we deduce that

$$\mathbb{E}[|\tilde{X}_t - X_t| + |\tilde{Y}_t - Y_t|] \leq \tilde{C}e^{-\lambda t}(|\tilde{x} - x| + |\tilde{y} - y|). \tag{133}$$

From the explicit formula of G, there exists $K > 0$ such that

$$|G_t| \leq K(|\rho_1(t)| + |\rho_2(t)| + |\xi(t)|).$$

Integrating and using (133) and the bound on $v(t) = \mathbb{E}(\xi^2(t))$, for each $T > 0$,

$$\int_0^T |G_t|^2 \mathrm{d}t \leq K_T(\rho_1(0)^2 + \rho_2(0)^2 + \xi(0)^2) < \infty.$$

Hence Novikov's condition holds on $[0, T]$ by a localization argument as in Example 8.2 using the stopping time $(\tau_n)_{n\geq 1}$ defined by

$$\tau_n := \inf\{t \geq 0 : |\rho_1(t)| + |\rho_1(t)| + |\xi(t)| \geq n\},$$

and the Doleans-Dade exponential of the process $(\int_0^t -G(s)dB_s)_{t\in[0,T]}$ given by

$$R_T = \exp\left(-\int_0^T G_s dB_s - \frac{1}{2}\int_0^T |G_s|^2 ds\right)$$

is a uniformly integrable martingale on $[0, T]$. Under $\mathbb{Q}_T$ defined by $d\mathbb{Q}_T|_{\mathcal{F}_T} = R_T d\mathbb{P}|_{\mathcal{F}_T}$, the process $\tilde{B}_t := B_t + \int_0^t G_s ds$ for $0 \leq t \leq T$ is a $\mathbb{Q}_T$-Brownian motion so that $(\tilde{X}_t, \tilde{Y}_t)_{t\in[0,T]}$ solves the original system (125) driven by $(\tilde{B}_t)_{t\in[0,T]}$.

Since previous results hold for any $T > 0$, one can check that $(\mathbb{Q}_T)_{T>0}$ is Kolmogorov consistent, which implies there exists a unique probability measure $\mathbb{Q}$ such that $\mathbb{Q}|_{\mathcal{F}_T} = \mathbb{Q}_T$ for all $T > 0$ by Kolmogorov extension Theorem 7.12.

Therefore, Hypothesis 14 holds and by Theorem 7.32, $(P_t)_{t\geq 0}$ satisfies an asymptotic log–Harnack inequality, $(P_t)_{t\geq 0}$ is asymptotically strong Feller and admits at most one invariant probability measure. □

9.2 Degenerate SDE with non-linear drift on high-frequency dynamics

As you can see in the multiple examples we treated until now, having a linear drift term on the nondegenerate part, said to be the *high-frequency* one, makes the asymptotic coupling strategy effective: this is a particular subcase of linear drifts that are covered by our examples and most of the work of various authors. As noted in [WZ13], *when the degenerate part is non-linear, the study becomes much more complicated*.

Based on several attempts, we observed that a non-linear drift in the degenerate part renders the analysis of the difference process, and in particular its asymptotic decay, highly challenging.

9.2.1 The monotone drift case

Above results are non-applicable to (88): indeed, the drifts from the degenerate parts is formed neither by a linear function, nor a partially linear function.

When the degenerate component exhibits a nonlinear drift, the asymptotic coupling analysis becomes significantly more delicate. Here we consider the case of a monotone nonlinear drift, for which similar contraction estimates can still be established.

By using a new asymptotic coupling strategy, we can state similar results as in the (semi-)linear case when the drift is monotone.

Example 9.12. As usual, we start by considering the following SDE on $\mathbb{R}^2$:

$$\begin{cases} dX_t &= f(Y_t)\mathrm{d}t, \\ dY_t &= g(X_t, Y_t)\mathrm{d}t + \mathrm{d}B_t, \end{cases} \tag{134}$$

where $(B_t)_{t\geq 0}$ is a one-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, $g : \mathbb{R}^2 \to \mathbb{R}$ is supposed to be a globally Lipschitz function, and $f : \mathbb{R} \to \mathbb{R}$ is supposed to be a globally Lipschitz, $C_b^2(\mathbb{R})$ function. In addition, we suppose that f is a strictly monotone function in the sense that $\exists c > 0$ a constant such that $\inf_{x\in\mathbb{R}} f'(x) \geq c > 0$ (respectively $\sup_{x\in\mathbb{R}} f'(x) \leq c < 0$). Without loss of generality, let's suppose that $\inf_{x\in\mathbb{R}} f'(x) \geq c > 0$.

We also assume that $f',\ f''$ are also globally Lipschitz.

Theorem 9.13. *In the setting of (134), there exists a unique strong solution $(Z_t)_{t\geq 0} = (X_t, Y_t)_{t\geq 0}$ starting at $Z_0 = z = (x, y) \in \mathbb{R}^2$ with Markov semigroup $(P_t)_{t\geq 0}$ defined as*

$$P_t\varphi(z) = \mathbb{E}(\varphi(Z_t^z)), \quad \forall \varphi \in \mathcal{B}_b(\mathbb{R}^2),\ \forall z \in \mathbb{R}^2,\ \forall t \geq 0.$$

Moreover, $(P_t)_{t\geq 0}$ is asymptotically strong Feller and has at most one invariant probability measure.

Proof. The existence and uniqueness of a strong solution $(Z_t)_{t\geq 0}$ to (134) as well as the characterization of its Markov semigroup follow standard results about finite-dimensional SDE with global Lipschitz drift (see e.g. [Gal16]).

We next consider the modified SDE on $\mathbb{R}^2$

$$\begin{cases} d\tilde{X}_t &= f(\tilde{Y}_t)\mathrm{d}t, \\ d\tilde{Y}_t &= g(\tilde{X}_t, \tilde{Y}_t)\mathrm{d}t + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\mathrm{d}t + \mathrm{d}B_t, \end{cases} \tag{135}$$

where $G : \mathbb{R}^4 \to \mathbb{R}$ is a function to be defined. We suppose that it is a Lipschitz function so existence and uniqueness of a strong solution $(\tilde{Z}_t)_{t\geq 0} = (\tilde{X}_t, \tilde{Y}_t)_{t\geq 0}$ starting at $\tilde{Z}_0 = \tilde{z} = (\tilde{x}_0, \tilde{y}_0) \in \mathbb{R}^2$ with Markov semigroup $P_t\varphi(\tilde{z}) = \mathbb{E}(\varphi(\tilde{Z}_t^{\tilde{z}}))$, for all $\varphi \in \mathcal{B}_b(\mathbb{R}^2)$, $\tilde{z} \in \mathbb{R}^2$, and $t \geq 0$, follow from the same standard results than above.

We consider the difference process $(\rho(t))_{t\geq 0} = (\rho_1(t), \rho_2(t))_{t\geq 0}$ where $\rho_1(t) = \tilde{X}_t - X_t$ and $\rho_2(t) = \tilde{Y}_t - Y_t$, which satisfies

$$\begin{cases} d\rho_1(t) &= \left(f(\tilde{Y}_t) - f(Y_t)\right)\mathrm{d}t, \\ d\rho_2(t) &= \left(g(\tilde{X}_t, \tilde{Y}_t) - g(X_t, Y_t)\right)\mathrm{d}t + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\mathrm{d}t. \end{cases}$$

In particular, we want to study the behavior of $\xi(t) = f(\tilde{Y}_t) - f(Y_t) + \tilde{X}_t - X_t$. By applying Itô's formula to $f \in C^2(\mathbb{R})$, we have

$$\begin{aligned} d\xi(t) =& f'(\tilde{Y}_t)d\tilde{Y}_t + \frac{1}{2}f''(\tilde{Y}_t)\mathrm{d}t - f'(Y_t)dY_t + \frac{1}{2}f''(Y_t)\mathrm{d}t + d\tilde{X}_t - dX_t \\ =& f'(\tilde{Y}_t)\left(g(\tilde{X}_t, \tilde{Y}_t) + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\right)\mathrm{d}t - f'(Y_t)g(X_t, Y_t)\mathrm{d}t + \left(f'(\tilde{Y}_t) - f'(Y_t)\right)\mathrm{d}B_t \\ &+ \frac{1}{2}(f''(\tilde{Y}_t) - f''(Y_t))\mathrm{d}t + (f(\tilde{Y}_t) - f(Y_t))\mathrm{d}t. \end{aligned}$$

As in previous model (125), let $\kappa > 0$ and $\sigma_t := f'(\tilde{Y}_t) - f'(Y_t)$, then

$$|\sigma_t| \leq L_{f'}|\rho_2(t)|,$$

for $L_{f'}$ the Lipschitz constant of f'. Since we supposed $f' > 0$, we can define

$$\begin{aligned} G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t) =& -g(\tilde{X}_t, \tilde{Y}_t) - \frac{1}{f'(\tilde{Y}_t)}\Bigg[\frac{1}{2}(f''(\tilde{Y}_t) - f''(Y_t)) \\ &+ f(\tilde{Y}_t) - f(Y_t) - f'(Y_t)g(X_t, Y_t) + \kappa\xi(t)\Bigg], \end{aligned}$$

so that

$$d\xi(t) = -\kappa\xi(t)\mathrm{d}t + \sigma_t \mathrm{d}B_t. \tag{136}$$

Let $u(t) := \mathbb{E}\left(\rho_1(t)^2\right),\ v(t) := \mathbb{E}\left(\xi(t)^2\right)$. From $d\rho_1(t) = (f(\tilde{Y}_t) - f(Y_t))\mathrm{d}t = (\xi(t) - \rho_1(t))\mathrm{d}t$, we obtain

$$du(t) = 2\mathbb{E}[\rho_1(\xi - \rho_1)]\mathrm{d}t \leq -u(t)\mathrm{d}t + v(t)\mathrm{d}t.$$

From (136) and by Itô's formula,

$$dv(t) = -2\kappa v(t)\mathrm{d}t + \mathbb{E}[|\sigma_t|^2]\mathrm{d}t.$$

Using (129) and Young's inequality,

$$\mathbb{E}[|\sigma_t|^2] \ \leq \ \frac{2L_{f'}^2}{c^2}(u(t) + v(t)).$$

Hence we obtain the *linear* differential inequality

$$\begin{cases} \frac{du(t)}{\mathrm{d}t} \ \leq \ -u(t) + v(t), \\ \frac{dv(t)}{\mathrm{d}t} \ \leq \ -2\kappa v(t) + 2\alpha\big(u(t) + v(t)\big), \end{cases}$$

where $\alpha := \frac{L_{f'}^2}{c^2}$. We introduce the Metzler matrix

$$M := \begin{pmatrix} -1 & 1 \\ 2\alpha & -2\kappa + 2\alpha \end{pmatrix}.$$

If $y(t) = (y_1(t), y_2(t))$ denotes the solution of the linear

$$\dot{y}(t) = My(t), \quad y(0) = \begin{pmatrix} u(0) \\ v(0) \end{pmatrix},$$

then by Kamke comparison principle,

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} \ \leq \ \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}.$$

for all $t \geq 0$. Assuming the following sufficient condition on κ,

$$\kappa \ > \ 2\alpha \ = \ \frac{2L_{f'}^2}{c^2},$$

then $\mathrm{Tr}(M) = -1 - 2\kappa + 2\alpha < 0$ and $\det(M) = 2\kappa - 4\alpha > 0$, so M is a Hurwitz matrix which implies that there exist $C \geq 1$ and $\lambda > 0$ such that

$$\|e^{tM}\| \leq Ce^{-\lambda t}, \quad t \geq 0,$$

which holds true for any norm while the constants may change. Therefore, since $y(t) = e^{tM}y(0)$ and considering the ℓ^1–norm, since $u(t), v(t) \geq 0$, it follows that

$$u(t) + v(t) = \|(u(t), v(t))\|_1 \ \leq \ \|y(t)\|_1 \ \leq \ Ce^{-\lambda t}\|y(0)\|_1 = Ce^{-\lambda t}(u(0) + v(0)), \tag{137}$$

which implies that

$$|\rho_2(t)| \ \leq \ \frac{1}{c}(|\xi(t)| + |\rho_1(t)|) \quad \Longrightarrow \quad \mathbb{E}[\rho_1(t)^2 + \rho_2(t)^2] \ \leq \ C'(u(t) + v(t)),$$

and there exists $C'' > 0$ a constant such that (137) yields

$$\mathbb{E}(|\tilde{X}_t - X_t|^2 + |\tilde{Y}_t - Y_t|^2) \ \leq \ C''e^{-\lambda t}(|\tilde{x} - x|^2 + |\tilde{y} - y|^2). \tag{138}$$

By definition of G and the previous Lipschitz bounds, we have

$$|G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)| \ \leq \ K(|\rho_1(t)| + |\rho_2(t)| + |\xi(t)|)$$

for some $K > 0$ depending on c, $L_{f'}$, $\|f''\|_\infty$ and L_g the Lipschitz constant of g. In particular, G is a globally Lipschitz function which proves previous claim. By (138) and $v(t) = \mathbb{E}\left(\xi(t)^2\right) \leq u(t) + v(t)$, for each fixed $T > 0$,

$$\mathbb{E}\left(\int_0^T |G_t|^2 \mathrm{d}t\right) \ < \ \infty.$$

One can verify that Novikov's condition holds on $[0, T]$ through the same localization argument as model (125) with the sequence of stopping times $(\tau_n)_{n\geq 1}$ defined by

$$\tau_n := \inf\{t \geq 0 : |\rho_1(t)| + |\rho_2(t)| + |\xi(t)| \geq n\},$$

and the Doleans–Dade exponential of $\left(\int_0^t -G(s)dB_s\right)_{t\in[0,T]}$ given by

$$R_T = \exp\Big(-\int_0^T G(t)\mathrm{d}B_t - \tfrac{1}{2}\int_0^T |G(t)|^2 \mathrm{d}t\Big),$$

is a uniform integrable martingale. By Girsanov's theorem, there exists a probability $\mathbb{Q}_T$ on $\mathcal{F}_T$ under which $(\tilde{B}_t)_{t\in[0,T]}$, defined by $\tilde{B}_t := B_t + \int_0^t G_s ds$, is a $\mathbb{Q}_T$–Brownian motion and $(\tilde{X}_t, \tilde{Y}_t)_{t\in[0,T]}$ solves the original SDE (134).

Since previous computations hold true for any $T > 0$, we can use the Kolmogorov extension Theorem 7.12 to the family $(\mathbb{Q}_T)_{T\geq 0}$ which is Kolmogorov consistent so that there exists a unique probability measure $\mathbb{Q}$ such that $\mathbb{Q}|_{\mathcal{F}_T} = Q_T$, for all $T > 0$.

Thus, Hypothesis 14 holds, an asymptotic log–Harnack inequality is verified, and by Theorem 7.32 the semigroup $(P_t)_{t\geq 0}$ is asymptotically strong Feller and admits at most one invariant probability measure. □

Remark 9.14. As before, the condition that f, f'' and g are Lipschitz continuous may be replaced by a local Lipchitz condition together with linear drift grow, which just add complexity as above construction.

Also, the strong monoticity condition on f' is similar to the one in [WZ13], where the authors considered SDE on $\mathbb{R}^m \times \mathbb{R}^d$ following

$$\begin{cases} dX_t^{(1)} &= Z^{(1)}(X_t^{(1)}, X_t^{(2)})\mathrm{d}t, \\ dX_t^{(2)} &= Z^{(2)}(X_t^{(1)}, X_t^{(2)})\mathrm{d}t + \sigma \mathrm{d}B_t, \end{cases} \tag{139}$$

where $(X_t^{(1)})_{t\geq 0}$ takes value in $\mathbb{R}^m$, $(X_t^{(2)})_{t\geq 0}$ in $\mathbb{R}^d$, σ is an invertible $d \times d$ matrix, $(B_t)_{t\geq 0}$ is a $d-$dimensional Brownian motion, $Z^{(1)} \in C^1(\mathbb{R}^m \times \mathbb{R}^d, \mathbb{R}^m)$, and $Z^{(2)} \in C^1(\mathbb{R}^m \times \mathbb{R}^d, \mathbb{R}^d)$. The idea is then to split $Z^{(1)}$ into a linear and non-linear part to derive an explicit derivative formula by controlling the non-linear part using the linear one in an efficient way, in the sense that

$$\nabla^{(2)} Z^{(1)} = B_0 + B,$$

where $\nabla^{(2)}$ is the gradient operator with respect to the second component, B_0 is a constant matrix $m \times d$.

This monotonicity condition may be interesting to enlarge to study more general SDE. one can think of the framework used in [PZ92], where they study the existence and uniqueness of solutions to various types of SDE, from linear to nonlinear cases. In the particular case of nonlinear drift, they extend their results starting by studying dissipative nonlinearities before generalizing their findings. Here is the same, since we start by linear models we extend until monotone ones: we can hope that such a generalization exists, and that we are on the right track.

9.3 The path to show unique ergodicity to self-repelling diffusion (88)

Although the existence of a unique invariant measure for (88) remains open, we summarize below several directions and model reductions we explored, highlighting their limitations.

One can start by studying a finite-dimensional approximation as an illustrative example.

Example 9.15. We consider the following SDE on $\mathbb{R}^2$

$$\begin{cases} dY_t = \cos(X_t)\mathrm{d}t,\ Y_0 = x, \\ dX_t = Y_t \sin(X_t)\mathrm{d}t + \mathrm{d}B_t,\ X_0 = y, \end{cases}$$

where $(B_t)_{t\geq 0}$ is a standard one-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. If we try to construct a modified SDE as before, we obtain

$$\begin{cases} d\tilde{Y}_t = \cos(\tilde{X}_t)\mathrm{d}t,\ \tilde{Y}_0 = \tilde{y}, \\ d\tilde{X}_t = \tilde{Y}_t \sin(\tilde{X}_t)\mathrm{d}t + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\mathrm{d}t + \mathrm{d}B_t,\ \tilde{X}_0 = \tilde{x}, \end{cases}$$

where $G : \mathbb{R}^4 \to \mathbb{R}$ is to be defined. Then, the difference process $(\rho_1(t), \rho_2(t))_{t\geq 0}$ defined as $\rho_1(t) = \tilde{X}_t - X_t$, $\rho_2(t) = \tilde{Y}_t - Y_t$, for all $t \geq 0$, is the solution of

$$\begin{cases} d\rho_1(t) = (\cos(\tilde{Y}_t) - \cos(Y_t))\mathrm{d}t, \\ d\rho_2(t) = \left(\tilde{Y}_t \sin(\tilde{X}_t) - Y_t \sin(X_t)\right)\mathrm{d}t + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\mathrm{d}t. \end{cases}$$

As before, we start to study the behavior of ρ_1. A first attempt may be, by analogy to previous examples, to focus on

$$\xi(t) = (\tilde{Y}_t - Y_t) + (\tilde{X}_t - X_t),$$

which satisfies the equation

$$d\xi(t) = (\cos(\tilde{Y}_t) - \cos(Y_t))\mathrm{d}t + \left(\tilde{Y}_t \sin(\tilde{X}_t) - Y_t \sin(X_t)\right)\mathrm{d}t + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\mathrm{d}t,$$

so that with

$$G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t) = -(\cos(\tilde{Y}_t) - \cos(Y_t)) - \left(\tilde{Y}_t \sin(\tilde{X}_t) - Y_t \sin(X_t)\right) - 2\xi(t),$$

we have

$$d\xi(t) = -2\xi(t) \ \Rightarrow\ \xi(t) = \xi(0)e^{-2t},$$

and in particular $\tilde{Y}_t - Y_t = -(\tilde{X}_t - X_t) + \xi(t)$. One can try to bound the drift part of the equation form of $d\rho_1(t)$ so that

$$d\rho_1(t) = (\cos(\tilde{Y}_t) - \cos(Y_t))\mathrm{d}t \leq |\tilde{Y}_t - Y_t|\mathrm{d}t,$$

but we lose the negative additive term $-(\tilde{X}_t - X_t)$ to get an exponential convergence to 0. If we consider the difference process $|\rho_1(t)|$, then by applying Tanaka's formula, we get

$$d|\rho_1(t)| = \mathrm{sign}\,(\rho_1(t))\, d\rho_1(t) = \mathrm{sign}\,(\rho_1(t)) \left[\cos(\tilde{Y}_t) - \cos(Y_t)\right]\mathrm{d}t,$$

which leads to the same conclusion.

To take another approach, let

$$\xi(t) = \cos(\tilde{Y}_t) - \cos(Y_t) + \tilde{X}_t - X_t,$$

so that if $\xi(t) \leq Ke^{-Ct}$, then $\cos(\tilde{Y}_t) - \cos(Y_t) \leq -(\tilde{X}_t - X_t) + \xi(t) \underset{t\to\infty}{\longrightarrow} -(\tilde{X}_t - X_t)$, which would lead to the desired bound

$$d\rho_1(t) \leq -\rho_1(t)\mathrm{d}t \;\Rightarrow\; \rho_1(t) \leq \rho_1(0)e^{-t} = (\tilde{x}_0 - x_0)e^{-t}.$$

By Itô's formula, it yields

$$d\cos(\tilde{Y}_t) = -\sin(\tilde{Y}_t)\left(\tilde{Y}_t \sin(\tilde{X}_t)\mathrm{d}t + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\mathrm{d}t + \mathrm{d}B_t\right),$$

and similarly $d\cos(Y_t) = -\sin(Y_t)\,(Y_t \sin(X_t)\mathrm{d}t + \mathrm{d}B_t)$ so that

$$\begin{aligned} d\xi(t) = &-\sin(\tilde{Y}_t)\left(\tilde{Y}_t \sin(\tilde{X}_t)\mathrm{d}t + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\mathrm{d}t + \mathrm{d}B_t\right) \\ &+ \sin(Y_t)\,(Y_t \sin(X_t)\mathrm{d}t + \mathrm{d}B_t) + \cos(\tilde{Y}_t)\mathrm{d}t - \cos(Y_t)\mathrm{d}t. \end{aligned}$$

As in the monotone case, we can keep the noisy term that will be canceled below expectation. However, when $\tilde{Y}_t \mapsto \sin(\tilde{Y}_t)$ reaches 0, the effect of the modification G to ξ will be null and void in this case.

We can try to combine previous approaches: keeping the last version of $\xi(t)$, then

$$|\xi(t)| \leq |\cos(\tilde{Y}_t) - \cos(Y_t)| + |\tilde{X}_t - X_t| \leq |\tilde{Y}_t - Y_t| + |\tilde{X}_t - X_t|,$$

so that we can try to bound each term on the right to obtain a global bound to $\xi(t)$. By Tanaka's formula,

$$\begin{aligned} d|\tilde{Y}_t - Y_t| &= \mathrm{sign}(\tilde{Y}_t - Y_t)\left(\tilde{Y}_t \sin(\tilde{X}_t) - Y_t \sin(X_t) + G(\tilde{X}_t, \tilde{Y}_t, X_t, Y_t)\right)\mathrm{d}t \\ d|\tilde{X}_t - X_t| &= \mathrm{sign}(\tilde{X}_t - X_t)(\cos(\tilde{Y}_t) - \cos(Y_t))\mathrm{d}t. \end{aligned}$$

Since $\mathrm{sign}(x)\cdot x = |x|$, $\mathrm{sign}(x)\cdot|x| = x$, there is no interesting choice of G to control both dynamics independently.

Since our current issue is about $\tilde{Y}_t \mapsto \sin(\tilde{Y}_t)$ reaching 0, we can introduce a third variable mirroring the original model of interest (87): we now consider the following SDE on $\mathbb{R}^3$,

$$\begin{cases} dU_t = \cos(X_t)\mathrm{d}t,\; U_0 = u, \\ dV_t = \sin(X_t)\mathrm{d}t,\; V_0 = v, \\ dX_t = U_t \sin(X_t)\mathrm{d}t + V_t \cos(X_t) + \mathrm{d}B_t,\; X_0 = x, \end{cases}$$

together with the associated modified SDE on $\mathbb{R}^3$,

$$\begin{cases} d\tilde{U}_t = \cos(\tilde{X}_t)\mathrm{d}t,\; \tilde{U}_0 = \tilde{u}, \\ d\tilde{V}_t = \sin(\tilde{X}_t)\mathrm{d}t,\; \tilde{V}_0 = \tilde{v}, \\ d\tilde{X}_t = \tilde{U}_t \sin(X_t)\mathrm{d}t + \tilde{V}_t \cos(\tilde{X}_t) + G_t\mathrm{d}t + \mathrm{d}B_t,\; \tilde{X}_0 = \tilde{x}, \end{cases}$$

where G_t stands for $G(\tilde{U}_t, \tilde{V}_t, \tilde{X}_t, U_t, V_t, X_t)$. Let

$$\xi(t) = \cos(\tilde{X}_t) - \cos(X_t) + \sin(\tilde{X}_t) - \sin(X_t) + \tilde{X}_t - X_t.$$

By Itô's formula, we have

$$\begin{aligned} d\xi(t) = &-\sin(\tilde{X}_t)\left(\tilde{U}_t \sin(\tilde{X}_t)\mathrm{d}t + \tilde{V}_t \sin(\tilde{X}_t) + G_t\mathrm{d}t + \mathrm{d}B_t\right) \\ &+ \cos(\tilde{X}_t)\left(\tilde{U}_t \sin(\tilde{X}_t)\mathrm{d}t + \tilde{V}_t \sin(\tilde{X}_t) + G_t\mathrm{d}t + \mathrm{d}B_t\right) \\ &+ \sin(X_t)\,(U_t \sin(X_t)\mathrm{d}t + V_t \cos(X_t) + \mathrm{d}B_t) - \cos(X_t)\,(U_t \sin(X_t)\mathrm{d}t + V_t \cos(X_t) + \mathrm{d}B_t) \\ &+ \cos(\tilde{X}_t)\mathrm{d}t - \cos(X_t)\mathrm{d}t. \end{aligned}$$

Now, when $\tilde{X}_t \mapsto \sin(\tilde{X}_t)$ reaches 0, the function $\tilde{X}_t \mapsto \cos(\tilde{X}_t)$ is obviously different from 0. However, it is still possible to have $\sin(\tilde{X}_t) = \cos(\tilde{X}_t)$, a situation where the effect of G will be still null and void.

Bibliography

[AB06] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis. A Hitchhiker's Guide*. Springer Berlin, Heidelberg, 2006.

[ATW06] M. Arnaudon, A. Thalmaier, and F.-Y. Wang. Harnack inequality and heat kernel estimates on manifolds with curvature unbounded below. *Bulletin des Sciences Mathématiques*, 130:223–233, 2006.

[BBN22] M. Benaïm, A. Bourquin, and D. H. Nguyen. Stochastic persistence in degenerate stochastic lotka-volterra food chains. *Discrete and Continuous Dynamical Systems - B*, 27(11):6841–6863, 2022.

[BCG15] M. Benaïm, Ioana Ciotir, and C.-E. Gauthier. Self-repelling diffusions via an infinite dimensional approach. *Stochastic Partial Differential Equations: Analysis and Computations*, 3:506–530, 2015.

[BCS25] M. Benaïm, J. Colombo, and E. Strickler. Stochastic persistence and extinction for degenerate stochastic rosenzweig-macarthur model, 2025. Preprint, arXiv:2511.10113.

[Ben18] M. Benaïm. Stochastic persistence. `https://arxiv.org/abs/1806.08450`, 2018. Preprint, first version 2014.

[BG17] M. Benaïm and C.-E. Gauthier. Self-repelling diffusions on a Riemannian manifold. *Probability Theory and Related Fields*, 169:63–104, 2017.

[BH99] M. R. Bridson and A. Haefliger. *Metric Spaces of Non-Positive Curvature*. Grundlehren der mathematischen Wissenschaften. Springer Berlin, Heidelberg, 1999.

[BH22] M. Benaïm and T. Hurth. *Markov Chains on Metric Spaces: A Short Course*. Universitext. Springer, 2022.

[BHN+25] M. Benaïm, A. Hening, D. Nguyen, S. Schreiber, and E. Strickler. Mathematical Foundations of Stochastic Persistence, 2025. Unpublished notes.

[BHS08] M. Benaïm, J. Hofbauer, and W. H. Sandholm. Robust permanence and impermanence for the stochastic replicator dynamics. *Journal of Biological Dynamics*, 2(2):180–195, 2008.

[BKS19] O. Butkovsky, A. Kulik, and M. Scheutzow. Generalized couplings and ergodic rates for SPDEs and other Markov models. `https://arxiv.org/pdf/1806.00395.pdf`, 2019. Preprint.

[BLM13] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities : a non asymptotic theory of independence*. Oxford University Press, 2013.

[BM24] M. Benaïm and L. Miclo. The asymptotic behavior of fraudulent algorithms, 2024. Preprint, arXiv:2401.12605.

[Bou23] A. Bourquin. Chaînes alimentaires de type lotka-volterra bruitées et quasi-stationnarité avec frontières mobiles aléatoires. phd thesis, 2023.

[BRZ04] Z. M. Balogh, K. Rogovin, and T. Zürcher. The Stepanov Differentiability Theorem in Metric Measure Spaces. *The Journal of Geometric Analysis*, 14:405–422, 2004.

[BS19] M. Benaïm and S. J. Schreiber. Persistence and extinction for stochastic ecological models with internal and external variables. *Journal of Mathematical Biology*, pages 393—-431, 2019.

[BWY19] J. Bao, F.-Y. Wang, and C. Yuan. Asymptotic Log-Harnack inequality and applications for stochastic systems of infinite memory. *Stochastic Processes and their Applications*, 129:4576–4596, 2019.

[CD87] D. Coppersmith and P. Diaconis. Random walk with reinforcement, 1987. Unpublished manuscript.

[Cer99] S. Cerrai. Ergodicity for stochastic reaction-diffusion systems with polynomial coefficients. *Stochastics and Stochastic Reports.*, 67(1-2):17–51, 1999.

[Che78] P. L. Chesson. Predator–prey theory and variability. *Annual Review of Ecology and Systematics*, 9:323–347, 1978.

[Che82] P. L. Chesson. The stabilizing effect of a random environment. *Journal of Mathematical Biology*, 15:1–36, 1982.

[Con07] J. B. Conway. *A Course in Functional Analysis*. Springer New York, NY, 2 edition, 2007.

[Dav93] M. H. A. Davis. *Markov models and optimization*, volume 49. Chapman & Hall, 1993.

[DCJ10] E. Durand-Cartagena and J.A. Jaramillo. Pointwise Lipschitz functions on metric spaces. *Journal of Mathematical Analysis and Applications*, 363:525–548, 2010.

[DM83] C. Dellacharie and P.-A. Meyer. *Probabilités et potentiel. Chapitres IX à XI: Théorie discrète du potentiel.* Publications de l'Institut de Mathématiques de l'Université de Strasbourg. Actualités Scientifiques et Industrielles. Hermann, Paris, 1983.

[Doo48] J. L. Doob. Asymptotic properties of markoff transition probabilities. *Transactions of the American Mathematical Society*, 63:393–421, 1948.

[DR92] R. T. Durrett and Leonard C. G. Rogers. Asymptotic behavior of Brownian polymers. *Probability Theory and Related Fields*, 92:337–349, 1992.

[EHS15] S. N. Evans, A. Hening, and S. J. Schreiber. Protected polymorphisms and evolutionary stability of patch-selection strategies in stochastic environments. *Journal of Mathematical Biology*, 71(2):325–359, 2015.

[EK86] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. Wiley, 1986.

[EL94] K. D. Elworthy and X.-M. Li. Formulae for the Derivatives of Heat Semigroups. *Journal of Functional Analysis*, 125:252–286, 1994.

[Gal16] J.-F. Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*, volume 274 of *Graduate Texts in Mathematics*. Springer, 1 edition, 2016.

[Gau16] C.-E. Gauthier. Long term behaviour of self-interacting diffusions, 2016.

[GRW24] M. Grothaus, P. R., and F.-Y. Wang. Singular degenerate sdes: Well-posedness and exponential ergodicity. *Journal of Differential Equations*, 413:632–661, 2024.

[GW12] A. Guilin and F.-Y. Wang. Degenerate fokker–planck equations: Bismut formula, gradient estimate and harnack inequality. *Journal of Differential Equations*, 253:20–40, 2012.

[Hai02] M. Hairer. Exponential mixing properties of stochastic PDEs through asymptotic coupling. *Probability Theory and Related Fields*, 124:345–380, 2002.

[Hai09] Martin Hairer. Ergodic properties of a class of non-markovian processes. In *Trends in Stochastic Analysis*, volume 353 of *London Mathematical Society Lecture Note Series*, pages 65–98. Cambridge University Press, 2009.

[Hai11] M. Hairer. On Malliavin's proof of Hörmander's theorem. *Bull. Sci. Math.*, 135:650–666, 2011.

[Hai21] M. Hairer. Convergence of Markov Processes. `https://hairer.org/notes/Convergence.pdf`, 2021. Unpublished lecture notes.

[Ham93] R. S. Hamilton. The Harnack estimate for the Ricci flow. *Journal of Differential Geometry*, 37:225–243, 1993.

[Ham23] Y. Hamaguchi. Markovian lifting and asymptotic log-Harnack inequality for stochastic Volterra integral equations. `https://arxiv.org/pdf/2304.06683.pdf`, 2023. Preprint.

[Har87] A. Harnack. *Die Grundlagen der Theorie des logarithmischen Potentiales und der eindeutigen Potentialfunktion in der Ebene*. V. G. Teubner, 1887.

[Hes18] J. P. Hespanha. *Linear Systems Theory*. Princeton University Press, 2 edition, 2018.

[HH80] P. Hall and C.C. Heyde. *Martingales Limit Theory and Its Application*. Probability and mathematical statistics. Academic Press, New York, 1980.

[Hir89] M. W. Hirsch. Systems of differential equations that are competitive or cooperative. v. convergence in 3-dimensional systems. *Journal of Differential Equations*, 80:94–106, 1989.

[HJ12] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 2 edition, 2012.

[HLL20] W. Hong, S. Li, and W. Liu. Asymptotic log-harnack inequality and applications for spde with degenerate multiplicative noise. *Statistics and Probability Letters*, 164, 2020.

[HM06] M. Hairer and J. C. Mattingly. Ergodicity of the 2D Navier-Stokes Equations with Degenerate Stochastic Forcing. *Annals of Mathematics*, 164:993–1032, 2006.

[HM11] M. Hairer and J. C. Mattingly. A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs. *Electronic Journal of Probability*, 16:658–738, 2011.

[HN18] A. Hening and D. H. Nguyen. Coexistence and extinction for stochastic kolmogorov systems. *Annals of Applied Probability*, 28(3):1893–1942, 2018.

[Hof81] J. Hofbauer. A general cooperation theorem for hypercycles. *Monatshefte fur Mathematik*, pages 233–240, 1981.

[HR11] T.-C. Hua and A. Rosalsky. A note on the de La Vallée Poussin criterion for uniform integrability. *Statistics and Probability Letters*, 81:169–174, 2011.

[Hsu02] E. P. Hsu. *Stochastic Analysis on Manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002.

[Hö85] L. Hörmander. *Analysis of Linear Partial Differential Operators*. Springer-Verlag, New-York, 1985.

[IK74] K. Ichihara and H. Kunita. A classification of the second order degenerate elliptic operators and its probabilistic characterization. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 30(3):235–254, 1974.

[IM74] K. Itô and H. P. Jr. McKean. *Diffusion Processes and their Sample Paths.* Die Grundlehren der mathematischen Wissenschaften, Band 125. Springer-Verlag, Berlin-New York, 1974.

[Itô50] K. Itô. Stochastic differential equations in a differentiable manifold. *Nagoya Mathematical Journal*, 1:35–47, 1950.

[IW77] N. Ikeda and S. Watanabe. A comparison theorem for solutions of stochastic differential equations and its applications. *Osaka Journal of Mathematics*, 14(3):619 – 633, 1977.

[IW81] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland, 1981.

[Kal01] O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer, 2 edition, 2001.

[Kam32] E. Kamke. Zur theorie der systeme gewöhnlicher differentialgleichungen. ii. *Acta Math.*, 58:57–85, 1932.

[Kea72] M. Keane. Strongly Mixing g-measures. *Inventiones Mathematicae*, 16:309–324, 1972.

[Ken87] W. S. Kendall. Stochastic differential geometry: An introduction. *Acta Applicandae Mathematica*, 9:29–60, 1987.

[Kha12] R. Khasminskii. *Stochastic Stability of Differential Equations*. Springer-Verlag Berlin Heidelberg, 2 edition, 2012.

[Kin63] J. F. C. Kingman. Ergodic properties of continuous-time markov processes and their discrete skeletons. *Proceedings of the London Mathematical Society*, s3-13(1):593–604, 1963.

[KS98] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer New York, NY, 2 edition, 1998.

[KS18] A. Kulik and M. Scheutzow. Generalized couplings and convergence of transition probabilities. *Probab. Theory Relat. Fields*, 171:333–376, 2018.

[Kun12] M. Kunze. Stochastic Differential Equations. `https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.020/kunze/SDE/sde_skript.pdf`, 2012. Lecture Notes, SummerTerm 2012, University of Ulm (DE).

[Lan13] U. Lang. Length Spaces. `https://people.math.ethz.ch/~lang/LengthSpaces.pdf`, 2013. Unpublished lecture notes.

[Lot10] A. J. Lotka. Contribution to the theory of periodic reaction. *The Journal of Physical Chemistry*, 14, 1910.

[Med23] T. Tachim Medjo. Asymptotic Log-harnack Inequality for the 2d Stochastic Cahn-hilliard-navier-stokes System With Degenerate Noise. `https://doi.org/10.21203/rs.3.rs-2600198/v1`, 2023. Preprint.

[Mey62] P.-A. Meyer. A decomposition theorem for supermartingales. *Illinois J. Math.*, 6(1):193–205, 1962.

[Mey66] P.-A. Meyer. *Probabilités et potentiel*. Blaisdell Publishing Company, 1966.

[Moh20] M. T. Mohan. Asymptotic log-Harnack inequality for the stochastic convective Brinkman-Forchheimer equations with degenerate noise. `https://arxiv.org/pdf/2008.00955.pdf`, 2020. Preprint.

[MT09] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, 2 edition, 2009.

[NRW87] J. R. Norris, L. C. G. Rogers, and D. Williams. Self-avoiding random walk: A Brownian motion model with local time drift. *Probability Theory and Related Fields*, 74:271–287, 1987.

[NS20] D. H. Nguyen and E. Strickler. A method to deal with the critical case in stochastic population dynamics. *SIAM Journal on Applied Mathematics*, 80(3):1567–1589, 2020.

[Obe17] H. Oberhauser. Stochastic Differential Equations. `https://courses.maths.ox.ac.uk/node/view_material/37063`, 2017. Unpublished lecture notes.

[Par68] K.R. Parthasarathy. *Probability Measures on Metric Spaces*. Academic Press, 1968.

[PR11] G. Da Prato and M. Röckner. Cores for generators of some markov semigroups. *AIP Conference Proceedings*, 1329:87–97, 2011.

[PZ92] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1992. Online publication date: M.h 2010.

[PZ96] G. Da Prato and J. Zabczyk. *Ergodicity for Infinite Dimensional Systems*. London Mathematical Society. Cambridge University Press, 1996. Lecture Note Series 229.

[RM63] M. L. Rosenzweig and R. H. MacArthur. Graphical representation and stability conditions of predator-prey interactions. *The American Naturalist*, 97(895):209–223, 1963.

[RQW24] P. Renc, Z. Qian, and F.-Y. Wang. Entropy estimate for degenerate sdes with applications to nonlinear kinetic fokker-planck equations. *SIAM Journal on Mathematical Analysis*, 56:5330–5349, 2024.

[RW10] M. Röckner and F.-Y. Wang. Log-Harnack inequality for stochastic differential equations in Hilbert spaces and its consequences. *Infinite Dimensional Analysis Quantum Probability and Related Topics*, 13:27–37, 2010.

[RY04] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften*. Springer, 3 edition, 2004.

[SBA11] S. J. Schreiber, M. Benaïm, and K. A. S. Atchadé. Persistence in fluctuating environments. *Journal of Mathematical Biology*, 62:655–683, 2011.

[Sch12] S. J. Schreiber. Persistence for stochastic difference equations: A mini review. *Journal of Difference Equations and Applications*, 18:1381–1403, 2012.

[Sch17] S. J. Schreiber. Coexistence in the face of uncertainty. *Recent Progress and Modern Challenges in Applied Mathematics, Modeling and Computational Science*, pages 349–384, 2017.

[Sei] J. Seidler. A note on the strong feller property. 2001. Unpublished lecture notes.

[Smi08] Hal L Smith. The rosenzweig-macarthur predator-prey model. *School of Mathematical and Statistical Sciences, Arizona State University: Phoenix*, 2008.

[Str96] D. W. Stroock. Gaussian measures in traditional and not so traditional settings. *Bull. AMS*, 33:135–255, 1996.

[Str00] D. W. Stroock. *An Introduction to the Analysis of Paths on a Riemannian Manifold*. Mathematical Surveys and Monographs 74. American Mathematical Society, 2000.

[Str19] E. Strickler. Persistance de processus de markov déterministes par morceaux. phd thesis, 2019.

[SV71] D. W. Stroock and S. R. S. Varadhan. Diffusion processes with boundary conditions. *Communications on Pure and Applied Mathematics*, 24:147–225, 1971.

[SV72] D. W. Stroock and S. R. S. Varadhan. *On the Support of Diffusion Processes with Applications to the Strong Maximum Principle*. University of California Press, 1972.

[TT79] P. Tuominen and R. L. Tweedie. Exponential decay and ergodicity of general markov processes and their discrete skeletons. *Advances in Applied Probability*, 11(4):784–803, 1979.

[Tur77] M. Turelli. Random environments and stochastic calculus. *Theoretical Population Biology*, 12:140–178, 1977.

[Vil03] C. Villani. *Topics in Optimal Transportation.*, volume 58 of *Graduate Studies in Mathematics.* American Mathematical Society, 2003.

[Vil09] C. Villani. *Optimal Transport. Old and New.* Grundlehren der mathematischen Wissenschaften. Springer Berlin, Heidelberg, 2009.

[Wan97] F.-Y. Wang. Logarithmic Sobolev inequalities on noncompact Riemannian manifolds. *Probability Theory and Related Fields*, 109:417–424, 1997.

[Wan10] F.-Y. Wang. Harnack inequalities on manifolds with boundary and applications. *Journal de Mathématiques Pures et Appliquées*, 94:304–321, 2010.

[Wan13] F.-Y. Wang. *Harnack Inequalities for Stochastic Partial Differential Equations*. Springer-Verlag New York, 2013.

[WHL21] S. Li W. Hong and W. Liu. Asymptotic Log-Harnack Inequality and Ergodicity for 3D Leray-α Model with Degenerate Type Noise. *Potential Analysis*, 55:477–490, 2021.

[WZ13] F.-Y. Wang and X.-C. Zhang. Derivative formula and applications for degenerate diffusion semigroups. *Journal de Mathématiques Pures et Appliquées*, 99:726–740, 2013.

[Xu11] L. Xu. A modified log-Harnack inequality and asymptotically strong Feller property. *Journal of Evolution Equations*, 11:925–942, 2011.