

---

Document de synthèse pour l'obtention d'une

# Habilitation à diriger des recherches

---

Discipline : Mathématiques  
Spécialité : Probabilités et Statistiques

---

Contributions à l'étude en temps long

de processus stochastiques

---

présenté par

**Fabien Panloup**

---

au vu des rapports de

Pr Benjamin Jourdain    École Nationale des Ponts et Chaussées  
Pr Denis Talay            INRIA Sophia Antipolis  
Pr Frederi Viens          University of Purdue

---

Soutenue publiquement le 4 décembre 2014 devant le jury composé de

|                      |  |              |
|----------------------|--|--------------|
| Pr Fabrice Baudoin   | University of Purdue                   | Examineur    |
| Pr Patrick Cattiaux  | Université Paul Sabatier               | Président    |
| Pr Serge Cohen       | Université Paul Sabatier               | Coordinateur |
| Pr Benjamin Jourdain | École Nationale des Ponts et Chaussées | Rapporteur   |
| Pr Gilles Pagès      | Université Pierre et Marie Curie       | Examineur    |
| Pr Denis Talay       | INRIA Sophia Antipolis                 | Rapporteur   |

---

Université Paul Sabatier - Institut de Mathématiques de Toulouse

---



## Remerciements

Mes remerciements vont tout d'abord à Benjamin Jourdain, Denis Talay et Frederi Viens qui ont accepté d'être rapporteurs de ce mémoire. Je leur suis particulièrement reconnaissant pour leur lecture attentive, leurs suggestions et plus généralement, pour l'intérêt accordé à mon travail. Benjamin Jourdain et Denis Talay me font de surcroît l'honneur d'être aujourd'hui membres du jury et je les en remercie chaleureusement.

Fabrice Baudoin a également accepté de participer à ce jury. Je le connais principalement au travers de ses articles, notamment sur la convergence à l'équilibre d'équations de Fokker-Planck cinétiques et sur le théorème de Hörmander pour les équations fractionnaires, et suis donc très heureux qu'il soit présent à la soutenance.

Merci aussi aux membres locaux du jury : Patrick Cattiaux, dont les conseils ou réponses à des questions de recherche ont souvent été précieux et Serge Cohen, coordinateur de l'HDR, qui a joué un rôle important dans mon intégration à l'IMT, me proposant notamment de collaborer dès mon arrivée à Toulouse.

Enfin, je suis profondément reconnaissant envers Gilles Pagès pour tout ce qu'il m'a apporté. Je le remercie pour son implication de la première heure, pour l'énergie et l'envie qu'il a su me communiquer, et naturellement pour sa présence dans ce jury.

Depuis mon arrivée à Toulouse, je me fais souvent la réflexion que le travail y est particulièrement agréable. Ceci tient à la qualité des structures de recherche et d'enseignement mais aussi à certaines rencontres : Aldéric Joulin co-bureau recruté la même année à l'INSA qui, pour reprendre le phrasé d'un certain Florent M., est un peu mon dual topologique (dans un sens plutôt  $L^2$ , les initiés comprendront...), Sébastien Gadat, à l'enthousiasme communicatif avec qui, les collaborations sont toujours un pur plaisir et Clément Pellegrini, lui-aussi, collaborateur et ami.

Mes pensées vont ensuite à l'ensemble de mes collègues de l'IMT et du département GMM de l'INSA, qui sont aussi pour beaucoup dans le sentiment explicité ci-dessus. Il règne dans ces deux structures une ambiance à la fois conviviale et dynamique et je souhaite ici leur dire le plaisir que j'ai à travailler avec eux.

J'exprime également ma gratitude envers l'INSA pour le financement qui m'a été accordé pour partir 7 mois au CMM à Santiago de Chile. Cette expérience fut riche à tous les niveaux. J'en profite pour remercier Joaquin Fontbona pour son accueil et pour la collaboration que l'on a pu débiter là-bas.

Quelques autres personnes que je souhaitais particulièrement saluer dans ces remerciements : Laure Coutin, Béatrice Laurent, Laurent Miclo (animateur pas seulement scientifique de la place mathématique toulousaine), Vincent Lemaire, Jaime San Martin, Samy Tindel, et enfin Sofiane Saadane dont je co-dirige la thèse avec Sébastien Gadat et à qui je souhaite un futur universitaire épanouissant.

Enfin, quelques mots pour ma petite famille sans qui les choses n'auraient pas la même saveur : ma compagne Claire que je remercie pour son soutien et son implication dans mes/nos projets et mon fils Léon dont l'énergie débordante est une garantie contre toute tentative de dilatation excessive du temps de travail !







# Table des matières

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>   | <b>1</b>  |
| <b>2</b> | <b>Sur l'étude de certains processus hypoelliptiques</b>  | <b>9</b>  |
| 2.1      | Diffusions à gradient à mémoire . . . . .   | 9         |
| 2.1.1    | Un exemple . . . . .  | 11        |
| 2.1.2    | Préliminaires . . . . .   | 11        |
| 2.1.3    | Comportement en temps long . . . . .  | 14        |
| 2.1.3.1  | Le cas stable : $\liminf_{t \rightarrow +\infty} r(t) = r_\infty \in (0, +\infty]$ . . . . .    | 14        |
| 2.1.4    | Comportement en temps long lorsque $r_\infty = 0$ . . . . .                                     | 19        |
| 2.2      | Grandes Déviations pour le régime stationnaire de diffusions à gradient moyenné                 | 20        |
| 2.2.1    | Introduction . . . . .  | 20        |
| 2.2.2    | Tension exponentielle . . . . .   | 22        |
| 2.2.3    | Caractérisation des $(LD)$ -limites . . . . .   | 23        |
| 2.2.3.1  | $(LD)$ -limite et contrôle optimal . . . . .  | 23        |
| 2.2.3.2  | Unicité et caractérisation de $W(z_i^*)$ . . . . .  | 25        |
| 2.2.3.3  | Concentration de $(\nu_\varepsilon)_{\varepsilon \in (0,1]}$ sur les minima globaux ? . . . . . | 27        |
| 2.3      | Un modèle de bulle spéculative . . . . .  | 30        |
| 2.3.1    | Introduction . . . . .  | 30        |
| 2.3.2    | Reformulations du modèle . . . . .  | 33        |
| 2.3.3    | Majoration de $\mathbb{P}(\tau > t)$ . . . . .  | 35        |
| 2.3.4    | Minoration de $\mathbb{P}(\tau > t)$ . . . . .  | 36        |
| 2.3.4.1  | Approche spectrale . . . . .  | 36        |
| 2.3.4.2  | Approche probabiliste . . . . .   | 39        |
| 2.3.5    | Mesure quasistationnaire et persistance . . . . .   | 41        |
| 2.4      | Commentaires et perspectives . . . . .  | 41        |
| <b>3</b> | <b>Approximation du régime stationnaire de diffusions et problèmes associés</b>                 | <b>45</b> |
| 3.1      | État de l'art . . . . .   | 45        |
| 3.2      | Approximation fonctionnelle du régime stationnaire . . . . .                                    | 50        |
| 3.2.1    | Convergence de $(\bar{\nu}^{(n)}(\omega, d\alpha))_{n \geq 1}$ . . . . .                        | 52        |
| 3.2.2    | Vitesse de convergence de $(\bar{\nu}^{(n)}(\omega, d\alpha))_{n \geq 1}$ . . . . .             | 53        |
| 3.3      | Un algorithme d'approximation du régime stationnaire moins moyennisant . . . . .                | 59        |
| 3.3.1    | Vitesse de $(\bar{\mu}_n(\omega, f))_{n \geq 1}$ . . . . .                                      | 60        |
| 3.3.2    | Vitesse de convergence de $(\bar{\mu}^{(n)}(\omega, F))_{n \geq 1}$ . . . . .                   | 66        |
| 3.3.3    | Accélération de convergence par la méthode de Richardson-Romberg . . . . .                      | 67        |
| 3.4      | Unicité de la mesure invariante pour les diffusions dupliquées . . . . .                        | 70        |

|          |  |            |
|----------|--|------------|
| 3.4.1    | Propriétés générales . . . . .   | 71         |
| 3.4.2    | Unicité dans le cas unidimensionnel . . . . .  | 73         |
| 3.4.3    | Un contre-exemple. . . . .   | 75         |
| 3.4.4    | Critères d'unicité dans le cas multidimensionnel . . . . .                                 | 76         |
| 3.4.5    | Applications, extensions . . . . .   | 77         |
| 3.4.6    | Lien avec le contrôle optimal. . . . .   | 79         |
| 3.5      | Commentaires et perspectives . . . . .   | 80         |
| <b>4</b> | <b>Sur le comportement en temps long d'EDS de type fractionnaire</b>                       | <b>83</b>  |
| 4.1      | Introduction et rappels . . . . .  | 83         |
| 4.2      | Convergence à l'équilibre d'EDS fractionnaires multiplicatives d'ordre $H > 1/2$           | 86         |
| 4.2.1    | Résultat principal . . . . .   | 89         |
| 4.2.2    | Éléments de preuve . . . . .   | 91         |
| 4.3      | Approximation de solutions stationnaires d'EDS multiplicatives d'ordre $H > 1/2$ . . . . . | 92         |
| 4.4      | Existence et approximation de solutions stationnaires d'EDS gaussiennes . .                | 97         |
| 4.5      | Quelques perspectives . . . . .  | 100        |
|          | <b>Bibliographie</b>   | <b>103</b> |



# Chapitre 1

## Introduction

Ce manuscrit est consacré à une synthèse des principaux travaux de recherche effectués après ma thèse (soutenue en décembre 2006).

Comme l'indique le titre du document, ces travaux sont en général (à une exception près sur laquelle on reviendra en fin d'introduction) reliés d'assez près à l'étude du comportement en temps long d'un processus stochastique. Les problèmes considérés vont de l'étude (motivée par des questions d'optimisation ou de modélisation) de la stabilisation de certains types de systèmes hypoelliptiques à la convergence à l'équilibre de systèmes non markoviens en passant par le problème de l'approximation du régime stationnaire de tels systèmes.

L'objectif de ce rapport est donc de présenter et de mettre en perspective les différents thèmes abordés ainsi que d'évoquer les diverses pistes de recherche relatives. La rédaction choisie est plutôt de type "auto-contenu". Ainsi, bien que les commentaires et les quelques esquisses de preuves soient plutôt informels, le rapport est a priori rédigé en vue d'une lecture ne nécessitant pas de renvois trop fréquents aux articles associés.

Avant d'aller plus en détails, on propose ci-après une présentation générale relativement succincte des thèmes de recherche développés dans la suite.

Les thèmes abordés dans le chapitre 1 ont en commun d'être reliés à des processus de Markov d'ordre 2. Plus exactement, il s'agit de modèles de diffusions  $d$ -dimensionnelles à mémoire (dans le drift) qui, par transformations simples, peuvent être vues, comme des diffusions  $2d$ -dimensionnelles markoviennes, dégénérées de type hypoelliptique.

**Diffusions à gradient moyenné.** Dans les deux premières sections de ce chapitre, basées sur [A4] et [A5] (références données en page 6), l'objectif est de présenter et d'étudier certaines propriétés d'une version à temps continu d'un algorithme d'optimisation dit "de gradient moyenné". Initialement motivé par des expérimentations numériques (voir *e.g.* [61]), ce travail a pour objectif de tenter d'évaluer l'impact de l'injection de mémoire dans un algorithme de type "descente de gradient" pour la recherche des minima (si possible globaux) d'une fonction  $U$  donnée (coercive). L'idée retenue est de moyenniser le gradient sur le passé (de manière pondérée), ce qui permet de fournir de l'inertie déterministe à la particule lorsqu'elle s'approche de ses minima locaux. À temps continu, ceci nous ramène à l'étude d'un modèle  $d$ -dimensionnel de la forme suivante :

$$dX_t = - \left( \frac{1}{k(t)} \int_0^t \dot{k}(s) \nabla U(X_s) ds \right) dt + \sigma(X_t) dW_t, \quad (1.1)$$

où  $k$  est une fonction strictement croissante représentant la pondération choisie. En vue d'un algorithme d'optimisation, il serait ici naturel d'imposer un coefficient de diffusion dépendant de  $t$  asymptotiquement évanescent. Néanmoins, une étude fine semble alors techniquement difficile. Le choix est donc ici de considérer le problème d'optimisation de manière "triangulaire" en s'intéressant successivement au comportement en temps long à  $\sigma$  "fixé", puis en étudiant le comportement limite lorsque  $\sigma$  tend vers 0. Ces deux objectifs coïncident en substance avec les thèmes respectifs des deux premières sections.

Plus précisément, l'idée est dans un premier de temps de comprendre le comportement en temps long en fonction de la mémoire injectée dans le système pour des fonctions  $U$  et  $\sigma$  données. Pour ce type de système, on met en avant un phénomène de type transition de phase. Sous des hypothèses de type Lyapounov, le système a des propriétés de stabilisation en temps long si et seulement si la fonction  $k$  a une croissance au moins exponentielle. Dans ce dernier cas, on obtient des résultats de convergence de la mesure d'occupation ou en loi de la diffusion vers un régime stationnaire. Dans le cas contraire, on montre que le système explose à une vitesse explicite dans les cas simples.

Lorsque la pondération de la mémoire est exponentielle, *i.e.* lorsque  $k(t) = \exp(\lambda t)$  ( $\lambda > 0$ ), la diffusion associée à (1.1) est une diffusion homogène hypoelliptique admettant, sous des hypothèses raisonnables sur  $U$  et  $\sigma$ , une unique probabilité invariante. Dans l'optique de l'étude de l'algorithme de recuit simulé relatif à ce système, on cherche alors à montrer que cette probabilité invariante se concentre sur les minima globaux lorsque le coefficient de diffusion (supposé constant) tend vers 0. Alors que la concentration sur les minima locaux est quasi-évidente, celle sur les minima globaux requiert généralement un principe de grandes déviations (PGD) pour la probabilité invariante (en "petit bruit") dont la fonction de taux associée doit être bien identifiée. Dans la section 2, on établit donc ce type de résultat mais le caractère généralement non explicite de la probabilité invariante rend ce type d'étude difficile. Si l'obtention du PGD est possible sous des conditions raisonnables, l'identification de la fonction de taux ou plus exactement du quasi-potentiel associé pose plus de problème. Via une étude fine du coût nécessaire pour joindre deux minima de  $U$ , on réussit néanmoins à obtenir des bornes pour des modèles double-puits permettant d'assurer la concentration annoncée précédemment sous des hypothèses mettant en jeu la dérivée seconde de  $U$  et les hauteurs de puits.

**Modèle de bulle spéculative.** Dans la dernière section de ce chapitre (basée sur [A13]), l'objectif est de proposer et d'étudier un modèle simple de dynamique spéculative. Le principe de la modélisation est de traduire mathématiquement le comportement spéculatif standard qui consiste à tenter d'anticiper les prix futurs au vu des variations du prix passé. Comme dans le cas précédent, ceci nous conduit à modéliser l'évolution du prix par une dynamique à mémoire dans le drift, ce qui après transformation revient à étudier une diffusion gaussienne 2-dimensionnelle "tournante", *i.e.* dont la matrice  $2 \times 2$  relative au système déterministe a des valeurs propres complexes. Après une "justification particulière" de cette dynamique, on se focalise sur une étude fine du temps de retour à un prix donné. Plus exactement, on cherche à obtenir des bornes quantatives pour la *persistance* associée : en d'autres termes, il s'agit de montrer que la queue de distribution de ce temps de retour est bornée par des exponentielles dont on veut estimer quantitativement le paramètre. Une des contributions principales de ce travail est de mettre en avant une quasi-proportionnalité de la persistance vis-à-vis de la vitesse de rotation du système déterministe. La notion de persistance étant

fortement reliée à celle de mesure quasi-stationnaire, ce travail se termine par un résultat d’existence de mesure quasi-stationnaire dans ce contexte. La question de l’unicité pour ce type de modèle reste un problème ouvert et constitue un de mes travaux en cours (en collaboration avec Jaime San Martin). Parmi d’autres, elle est en particulier abordée à la fin du chapitre dans les perspectives de recherche.

Le chapitre suivant est consacré aux méthodes d’approximation du régime stationnaire de processus de Markov, essentiellement présentées dans le cadre des diffusions.

**Approximation fonctionnelle du régime stationnaire.** Dans la continuité de mes derniers travaux de thèse, on s’intéresse dans un premier temps à la construction et à l’étude d’un algorithme efficace d’approximation du régime stationnaire fonctionnel, *i.e.* de la loi du processus ayant pour condition initiale la probabilité invariante de la diffusion. Ce travail fait naturellement suite à l’étude de procédures similaires pour l’approximation du régime stationnaire marginal, *i.e.* de la probabilité invariante elle-même. À l’image de ces algorithmes (voir la section “État de l’art” pour plus de détails), la procédure d’approximation du régime stationnaire fonctionnel est basée sur une suite de mesures d’occupation relatives au schéma d’Euler de la diffusion, choisi de pas décroissant (lentement) vers 0 afin de gérer simultanément les erreurs de discrétisation et de “temps long” inhérentes au problème. Plus exactement, ces mesures d’occupation sont des moyennes pondérées de masses de Dirac en les trajectoires shiftées du schéma d’Euler. Dans la section 2 de ce chapitre (basée sur [A6] et [A10]), on établit la convergence de ce type de suites de mesures empiriques, puis pour une fonctionnelle  $F$  fixée, on obtient un résultat de vitesse de type TCL (ce résultat semblant également nouveau dans le cadre non discrétisé). La normalité asymptotique de cet algorithme est enfin testée numériquement sur le problème d’estimation de prix d’options barrière dans des modèles à volatilité stochastique stationnaire.

**Un algorithme moins moyennisant.** Construite en mimant la mesure d’occupation naturelle du processus (au sens moyenne infinitésimale), la procédure précédente requiert une “simulation de la fonctionnelle” à chaque pas de temps, ce qui peut s’avérer très coûteux. En particulier, excepté une classe de fonctionnelles “agréables”, la complexité de la procédure n’est pas linéaire. Dans la section 3 de ce chapitre (basée sur [A3]), on propose alors de considérer un algorithme moins moyennisant dit “mixed-step” basé sur le calcul de la fonctionnelle à des instants absolus notés  $nT$ ,  $n \geq 0$  où  $T$  est un réel strictement positif. En notant que ce type d’alternative peut aussi avoir un intérêt dans le cas marginal (lorsque que le calcul de la fonction considérée est coûteux), la vitesse de convergence de cet algorithme est étudiée dans les deux cadres : marginal et fonctionnel. Une des contributions principales est alors de mettre en avant que cette modification n’affecte pas l’ordre de la vitesse optimale de convergence (en  $N^{-1/3}$  dans le cas marginal où  $N$  désigne le nombre de points de discrétisation) et que la variance associée croît de manière raisonnable (linéairement en  $T$ ). Après une discussion sur l’existence de solutions régulières à l’équation de Poisson associée au problème (condition inhérente à tous les résultats de vitesse de ce chapitre), on termine cette section par la mise en oeuvre d’une méthode de type Romberg pour accélérer la convergence. On montre en particulier qu’il est possible d’accélérer la convergence (en  $N^{-2/5}$  dans le cas marginal) tout en préservant la variance de la procédure “mixed-step” originelle si l’on a unicité de la probabilité invariante pour la *diffusion dupliquée* construite avec des accroissements consistants (voir paragraphe suivant pour plus de clarté).

**Confluence faible.** La condition d’unicité ci-dessus constitue la motivation principale de la

dernière section de ce chapitre (basée sur [A2]). Pour une diffusion admettant une unique probabilité invariante  $\nu$ , on cherche à savoir si la diffusion dupliquée constituée de deux copies (a priori couplées) de cette même diffusion admet elle-aussi une unique probabilité invariante. Bien que l'on aborde ce problème de façon relativement générale, la plus grande partie de cette section est consacrée au cas où les browniens sont consistants, *i.e.* lorsque les browniens relatifs à chaque copie sont les mêmes. Cette focalisation, liée à la condition nécessaire à la préservation de la variance dans la méthode de Romberg décrite ci-dessus, s'explique également par son intérêt théorique. En effet, lorsque les deux copies sont identiques, le système est totalement dégénéré et les critères d'unicité classiques ne s'appliquent plus. En revanche, on remarque que, dans ce cas, il existe une probabilité invariante évidente :  $\nu_{\Delta} := \nu \circ (x \mapsto (x, x))^{-1}$ , *i.e.* la loi d'une condition initiale  $(X_0, X_0)$  où  $X_0 \sim \nu$ . Du point de vue "ergodique", montrer que cette mesure invariante est unique revient à montrer que la diffusion conflue faiblement, *i.e.* que ses trajectoires (issues de deux points distincts) se rapprochent à l'infini dans un sens statistique. Après avoir montré que la confluence faible (et souvent forte) est quasi-toujours vraie en dimension 1, on construit un contre-exemple en dimension 2, puis une série de critères de confluence faible basés sur un *exposant de Lyapounov non-infinitésimal*. On montre enfin que ces critères permettent de montrer la confluence faible dans des cas non triviaux (tels que des "diffusions-gradients" associés à des potentiels non convexes) et qu'ils peuvent être reliés à un problème de transport optimal.

Le dernier chapitre concentre des travaux liés au comportement en temps long d'EDS fractionnaires. Outre l'aspect intégration, le point captivant de ce type de problème est de comprendre comment gérer la mémoire du mouvement brownien fractionnaire (fBm) pour réussir à répondre dans ce cadre aux questions naturelles associées au temps long. La première pierre (le pluriel serait ici plus adapté!) est posée par Hairer (et ses collaborateurs, voir le chapitre pour les références) qui met en place une structure markovienne homogène adaptée à ce type d'EDS (fortement non markovienne) permettant une définition rigoureuse du concept de mesure invariante et l'établissement de critères d'unicité raisonnables pour cette dernière. Sont aussi développés des méthodes de couplage dit asymptotique permettant d'obtenir des bornes pour la vitesse en variation totale dans le cas additif, *i.e.* dans le cas où le coefficient de "diffusion" est constant. Après un état de l'art, on présente ici trois travaux sur le sujet dans les sections 2, 3 et 4 de ce chapitre (respectivement basés sur [A14], [A1] et [A7]).

**Vitesse de convergence à l'équilibre pour des EDS fractionnaires multiplicatives.** Dans la continuité de la fin du paragraphe précédent, on propose dans un premier temps d'aborder le problème de la convergence à l'équilibre (en variation totale) pour des EDS fractionnaires multiplicatives lorsque l'indice de Hurst  $H$  est strictement supérieur à  $1/2$ . Il s'agit donc du cadre où les intégrales peuvent être définies sans problème via l'intégrale de Young (à l'inverse, la mémoire est plus forte que dans le cas  $H < 1/2$ ). L'idée est donc ici de montrer que les méthodes de couplage développées dans le cadre additif peuvent s'étendre à ce cadre multiplicatif. Quand on cherche à mettre en place des méthodes de couplage (au sens variation totale du moins) dans un cadre à mémoire, on est confronté à deux difficultés majeures. D'une part, quand deux trajectoires se touchent, elles ne restent pas naturellement collées et d'autre part, chaque tentative de couplage a un coût : le "déphasage" nécessaire entre les deux fBms pour tenter de coller (ou de rester collé) pèsera sur le futur si la tentative rate. Hairer met alors en place un schéma de couplage en trois étapes lui

---

permettant d'obtenir une vitesse en  $t^{-\alpha}$  où  $\alpha \in (0, 1)$ . La contribution principale présentée dans cette section est de montrer, qu'il est possible de mettre en place des outils de type Lyapounov (non triviaux) pour ce type de modèle permettant de faire fonctionner ce schéma de couplage dans le cadre multiplicatif. En particulier, sous une hypothèse sur  $\sigma$  sur laquelle nous reviendrons plus tard, il nous permet de retrouver le même ordre de vitesse que dans le cas additif.

**Approximation de lois stationnaires d'EDS fractionnaires.** Les deux dernières sections sont consacrées à des méthodes d'approximation des *solutions stationnaires* relatives à ce type d'EDS, d'abord dans le cadre des EDS multiplicatives d'ordre  $H > 1/2$  puis dans celui des EDS additives mais dirigées par des processus gaussiens généraux. À l'image des procédures développées dans le chapitre 3, on construit des mesures empiriques basées sur des schémas d'Euler de ces EDS. Néanmoins, les schémas sont supposés ici de pas constant  $\gamma$  (pour des raisons numériques et techniques). Pour ce type de procédures, on obtient, sous hypothèses adéquates, la convergence à  $\gamma$  fixé vers une loi stationnaire discrétisée, puis lorsque  $\gamma$  tend vers 0, on prouve la convergence de ces lois stationnaires vers celles du processus. Ces résultats sont enfin illustrés par des simulations numériques. Là encore, les problèmes ouverts étant nombreux, on conclue ce chapitre comme les autres par des perspectives associées.

**Estimation de la volatilité instantanée.** Comme annoncé en début d'introduction, on termine ici par quelques mots sur un travail qui me semble un peu trop en marge des autres pour s'insérer de manière cohérente dans ce manuscrit. Il s'agit de la référence [A8]. Dans cet article, on s'intéresse à un problème classique de statistique des processus : celui de l'estimation de la volatilité. Alors que ce problème est souvent abordé dans un sens intégré, on propose ici un estimateur de la volatilité instantanée dans un modèle sous la forme  $dX_t = a_t dt + \sigma_t dB_t$ . L'idée de sa construction est de moyenniser les *power-variations* de  $(X_t)_{t \geq 0}$  (associées aux observations) sur des fenêtres de temps dont la longueur dépend du nombre d'observations. Dans un cadre général où  $(a_t)$  est un processus càdlàg (général) et  $(\sigma_t)$  est une semi-martingale càdlàg sans sauts prévisibles, on prouve la normalité asymptotique de l'estimateur et détermine le choix optimal de la fenêtre en fonctions des variations de  $\sigma$ .

## Publications

Cette section recense mes publications ou prépublications. Celles-ci sont toutes accessibles à la page suivante : <http://www.math.univ-toulouse.fr/~panloup/Recherche/>.

Publications dans des revues internationales avec comité de lecture :

- [A1] Approximation of stationary solutions to SDEs driven by multiplicative fractional noise. En collaboration avec Serge Cohen et Samy Tindel. *Stochastic Processes and Applications*. Volume 124 (13). 2014. 28 pages.
- [A2] Invariant distribution of duplicated diffusions and application to Richardson-Romberg extrapolation. En collaboration avec Vincent Lemaire and Gilles Pages. A paraître dans *Annales de l'IHP : Probabilités et Statistiques*. 39 pages. Accessible à la page suivante <http://hal.archives-ouvertes.fr/hal-00785766>.
- [A3] A mixed-step algorithm for the approximation of the stationary regime of a diffusion. En collaboration avec Gilles Pagès. *Stochastic Processes and their Applications*. Volume 124 (1). 2014. 34 pages.
- [A4] Large Deviation Principle for invariant distributions of Memory Gradient Diffusions. En collaboration avec Sébastien Gadat et Clément Pellegrini. 2012. *Electronic Journal of Probability* Volume 18 (81). 36 pages.
- [A5] Long Time Behavior and Stationary Regime of Memory Gradient Diffusions. En collaboration avec Sébastien Gadat. *Annales de l'IHP : Probabilités et Statistiques* Volume 50 (2). 2014. 39 pages.
- [A6] Ergodic Approximation of the distribution of a stationary diffusion : Rate of convergence. En collaboration avec Gilles Pagès. *Annals of Applied Probability*. Volume 22 (3). 2012. 41 pages.
- [A7] Approximation of Stationary Solutions of Gaussian Driven Stochastic Differential Equations. En collaboration avec Serge Cohen. *Stochastic Processes and Applications*. Volume 121 (12). 2011. 26 pages.
- [A8] Estimation of the instantaneous volatility. En collaboration avec Alexander Alvarez, Monique Pontier et Nicolas Savy. *Statistical Inference for Stochastic Processes*. Volume 15 (1) 2012. 33 pages.
- [A9] A connection between extreme value theory and long time approximation of SDEs. *Stochastic Processes and their Applications*. Volume 19 (10). 2009. 25 pages.
- [A10] Approximation of the distribution of a stationary Markov process with application to option pricing. En collaboration avec Gilles Pagès. *Bernoulli*. Volume 15 (1). 2009. 32 pages.

[A11] Computation of the invariant measure of a Lévy driven SDE : Rate of convergence. *Stochastic Processes and their Applications*. Volume 118 (8). 2008. 34 pages.

[A12] Recursive computation of the invariant measure of a SDE driven by a Lévy process. *Annals of Applied Probability*. Volume 18 (2). 2008. 48 pages.

**Prépublications :**

[A13] A stochastic model for speculative bubbles. En collaboration avec Sébastien Gadrat et Laurent Miclo. 49 pages. Accessible à la page suivante : <http://arxiv.org/pdf/1309.6287v1.pdf>.

[A14] Rate of convergence to equilibrium of fractional driven stochastic differential equations with multiplicative noise. En collaboration avec Joaquin Fontbona. Préprint. Accessible à la page suivante <http://hal.archives-ouvertes.fr/hal-00989414>.

**Thèse de doctorat**

[T] Approximation du regime stationnaire d'une EDS avec sauts. *Thèse de l'Université Paris VI*, 2006. Accessible à <http://tel.archives-ouvertes.fr/tel-00120508/fr>.





## Chapitre 2

### Sur l'étude de certains processus hypoelliptiques

Ce chapitre est divisé en trois parties. Les deux premières, basées sur les articles [35] et [36], concernent l'étude de certaines propriétés en temps long de diffusions à *gradient à mémoire* tandis que dans la dernière, issue de [34], on s'intéresse à la modélisation du phénomène de bulle spéculative via un processus hypoelliptique.

#### 2.1 Diffusions à gradient à mémoire

Pour une fonction  $U : \mathbb{R}^d \rightarrow \mathbb{R}$  donnée, coercive et continue, une question fondamentale d'optimisation est : comment construire une dynamique efficace permettant d'approcher son minimum global ? La réponse la plus simple à ce problème est l'algorithme de descente de gradient qui, dans un cadre à temps continu, s'écrit comme la solution de  $\dot{x}_t = -\nabla U(x_t)$ . Lorsque  $U$  est strictement convexe, cette dynamique donne une réponse satisfaisante à ce problème : la particule converge vers la cible souhaitée à une vitesse qui peut être extrêmement rapide si la fonction a de bonnes propriétés. Dès que la fonction  $U$  a plusieurs minima locaux, on est amené, pour espérer la convergence vers le minimum global, à faire des modifications de cette dynamique. La plus célèbre est celle dite du *recuit simulé*, qui consiste à introduire un aléa d'intensité bien calibrée afin de provoquer une excitation de la particule, qui assure l'évitement des minima locaux et la convergence (en probabilité) vers le minimum global. Malheureusement, l'infailibilité de ce nouvel algorithme a un coût : la vitesse de convergence qui s'en trouve fortement détériorée.

Plutôt que d'introduire un aléa, on peut aussi penser à modifier la dynamique déterministe, *i.e.* à considérer d'autres systèmes différentiels dont les points critiques coïncident avec les zéros de  $\nabla U$ . La descente de gradient à mémoire, introduite et étudiée dans [9] et [10] en est un exemple : il s'agit d'une modification où l'on remplace le gradient de la particule à l'instant  $t$  par une moyenne pondérée de sa valeur sur le passé :  $\dot{x}_t = -k(t)^{-1} \int_0^t k(s) \nabla U(x_s) ds$  où la fonction  $k$  est une fonction strictement croissante et régulière. Dans ce type de modèle, la mémoire induit une inertie qui permet à la particule de ne pas rester bloquée dès qu'elle rencontre un minimum local. Si un tel algorithme ne converge pas sans conditions vers le minimum global, on peut en revanche espérer, qu'à injection égale de bruit aléatoire, l'algorithme à gradient à mémoire (ou gradient moyenné) se trompe moins que l'algorithme classique. Cette remarque constitue l'un des points de départ des travaux décrits dans cette section et la suivante.

Outre l'aspect optimisation, la dynamique de descente de gradient à mémoire présente un

second intérêt en termes d'applications. En effet, ce type de système différentiel décrit, sous certaines conditions, le mouvement d'une boule roulant sur le graphe de  $U$ , soumise à l'apesanteur et à une certaine forme de frottement (en anglais, *Heavy Ball with Friction*). L'étude des propriétés asymptotiques d'une version bruitée de ce type de dynamique constitue donc une seconde motivation de notre travail.

En accord avec les remarques précédentes, nous considérons dans cette partie une version stochastique de la descente de gradient à mémoire. Plus précisément, étant donnée une fonction  $U : \mathbb{R}^d \mapsto \mathbb{R}_+^*$  régulière et satisfaisant les hypothèses suivantes,

$$\lim_{|x| \rightarrow +\infty} U(x) = +\infty \quad \text{et} \quad \liminf_{|x| \rightarrow +\infty} (x|\nabla U(x)) > 0, \quad (2.1)$$

nous considérons une EDS sous la forme suivante :

$$dX_t = - \left( \frac{1}{k(t)} \int_0^t \dot{k}(s) \nabla U(X_s) ds \right) dt + \sigma(X_t) dW_t. \quad (2.2)$$

La fonction  $\sigma : \mathbb{R}^d \rightarrow \mathbb{M}_{d,d}$  est continue et localement lipschitzienne,  $(W_t)_{t \geq 0}$  est un mouvement brownien standard  $d$ -dimensionnel et  $(k(t))_{t \geq 0}$  est une fonction déterministe positive,  $\mathcal{C}^2$  et strictement croissante, telle que  $k(0) = 1$ . Ce type de modèle est markovien d'ordre 2. Plus précisément, si l'on pose

$$Y_t = \frac{1}{k(t)} \int_0^t k'(s) \nabla U(X_s) ds,$$

alors, on remarque que  $(X_t, Y_t)_{t \geq 0}$  satisfait l'EDS suivante :

$$\begin{cases} dX_t = \sigma(X_t) dW_t - Y_t dt. \\ dY_t = r(t) (\nabla U(X_t) - Y_t) dt, \end{cases} \quad (2.3)$$

où  $r(t) = \frac{k'(t)}{k(t)}$ . Sous réserve d'existence et d'unicité des solutions (voir Proposition 2.4), le processus  $(Z_t)_{t \geq 0} := (X_t, Y_t)_{t \geq 0}$  ainsi construit est markovien. On peut remarquer que ce processus est de plus homogène si et seulement si  $t \mapsto r(t)$  est constante, *i.e.* si et seulement si

$$k(t) = e^{\lambda t} \quad \text{avec } \lambda > 0.$$

**Remarque 2.1** Lorsque  $r(t) = 1$ , ce système est en fait relativement proche d'une équation de type Fokker-Planck cinétique (cf *e.g.* [26], [5] ou [70]) dont la forme standard est donnée par

$$\begin{cases} dX_t = -Y_t dt. \\ dY_t = -(\nabla U(X_t) + Y_t) dt + \sigma(X_t) dW_t. \end{cases}$$

Les parties déterministes diffèrent "seulement" d'un signe tandis que la partie brownienne apparaît dans la position dans notre problème et dans la vitesse dans le second. Ces différences ne sont en réalité que d'apparence mineure.

La fonction  $t \mapsto r(t)$  caractérise la mémoire injectée dans le système et joue un rôle fondamental dans la suite. Pour simplifier, supposons pour l'instant que  $r(t) \rightarrow r_\infty \in \mathbb{R}_+ \cup \{+\infty\}$ . La quantité  $r_\infty$  est en quelque sorte, une fonction décroissante de la mémoire. On peut par exemple remarquer que

- $r_\infty = +\infty$  si  $k(t) = \exp(\lambda t^\alpha)$  avec  $\alpha > 1$  et  $\lambda > 0$ ,
- $r_\infty \in (0, +\infty)$  si  $k(t) = \exp(\lambda t)$  avec  $\lambda > 0$ ,
- $r_\infty = 0$  si  $k(t) = \exp(\lambda t^\alpha)$  avec  $\lambda > 0$  et  $\alpha \in (0, 1)$  ou  $k(t) = (1+t)^\alpha$  avec  $\alpha > 0$ .

Les cas considérés ci-dessus constituent dans la suite les classes principales de mémoire étudiées. Pour avoir une idée du comportement asymptotique relatif à chacune de ces classes considérons d'abord un cas simple.

### 2.1.1 Un exemple

On suppose ici que  $d = 1$ ,  $U(x) = ax^2/2$  avec  $a \in \mathbb{R}_+^*$ ,  $\sigma(x) = \sigma \in \mathbb{R}_+^*$  et  $r(t) = r_\infty \in (0, +\infty)$ . Dans ce cadre gaussien,  $(Z_t)_{t \geq 0}$  est donc solution de

$$dZ_t = AZ_t dt + BdW_t \quad \text{où,} \quad A = \begin{pmatrix} 0 & -1 \\ ar_\infty & -r_\infty \end{pmatrix}, \quad B = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}, \quad (2.4)$$

et  $(W_t)_{t \geq 0}$  est un mouvement Brownien en dimension 2. On peut alors montrer le résultat suivant :

**Proposition 2.2 ([35])** *Supposons que  $r(t) = r_\infty \in (0, +\infty)$ . Alors pour tout  $z \in \mathbb{R}^d \times \mathbb{R}^d$ ,*

$$\mathcal{L}(Z_t^z) \xrightarrow{t \rightarrow +\infty} \nu_\infty := \mathcal{N}(0, \Sigma_{r_\infty})$$

où

$$\Sigma_{r_\infty} = \frac{\sigma^2}{2} \begin{pmatrix} \frac{1}{a} + \frac{1}{r_\infty} & \frac{1}{a} \\ \frac{1}{a} & 1 \end{pmatrix}.$$

De plus,  $\nu_\infty$  est l'unique probabilité invariante de  $(Z_t)_{t \geq 0}$ .

**Remarque 2.3**  $\triangleright$  Dans ce cadre, on observe qu'asymptotiquement,  $(X_t)$  est bien une variable aléatoire concentrée autour de l'unique minimum de  $U$ . On constate également que  $\sigma_{r_\infty}^2 = \lim \mathbb{E}[X_t^2]$  dépend fortement de  $r_\infty$ . On peut dans un premier temps remarquer que la variance asymptotique de  $X_t$  est toujours supérieure à celle du système sans mémoire qui est égale à  $\sigma^2/(2a)$ . Lorsque  $r_\infty$  tend vers  $+\infty$ , on constate (naturellement) que  $\sigma_{r_\infty}^2$  converge vers la variance asymptotique du système sans mémoire tandis que  $\sigma_{r_\infty}^2$  tend vers  $+\infty$  lorsque  $r_\infty$  tend vers 0. Le système devient alors instable. Les propriétés ci-dessus décrivent relativement bien celles qui seront établies dans un cadre plus général dans la suite.

$\triangleright$  Il est aussi intéressant de noter que les valeurs propres de  $A$  sont réelles strictement négatives si  $r_\infty > 4a$  et qu'elles deviennent complexes de partie réelle strictement négative si  $r_\infty < 4a$ . Autrement dit, le système se met à osciller lorsque la mémoire augmente.

### 2.1.2 Préliminaires

- **Non-explosion en temps fini.** La fonction  $U$  n'étant en général pas sous-quadratique, il nous faut commencer par assurer l'existence de solutions et en particulier la non-explosion en temps fini de ces dernières.

**Proposition 2.4** ([35]) *Supposons que  $\text{Tr} [\sigma^* D^2 U \sigma] \leq CU$  ( $C > 0$ ). Alors, on a existence forte de solutions (et unicité trajectorielle) pour l'EDS (2.3). De plus, si  $(X_0, Y_0)$  satisfait  $\mathbb{E}[U(X_0) + |Y_0|^2] < +\infty$ , alors pour tout  $T > 0$ ,  $\sup_{t \in [0, T]} \mathbb{E}[U(X_t) + |Y_t|^2] < +\infty$ .*

Dans ces conditions, on notera  $\mathcal{A}$  le générateur infinitésimal du processus de Markov  $(X_t, Y_t, t)_{t \geq 0}$  défini par :

$$\mathcal{A}f(x, y, t) = -(y|\nabla_x f) + r(t)(\nabla U(x) - y|\nabla_y f) + \frac{1}{2}\text{Tr}(\sigma^*(x)D_x^2 f(x, y)\sigma(x)) + \partial_t f. \quad (2.5)$$

**Remarque 2.5** Pour ce résultat standard, l'idée de la preuve (lorsque  $U$  n'est pas à croissance linéaire) est d'introduire une fonction de Lyapounov qui permet de contrôler la fonction en temps fini. La fonction naturelle  $h(x, y, t) = U(x) + |y|^2/(2r(t))$  (fonction qui permet de "séparer les variables") satisfait pour tout  $T > 0$ ,

$$\mathcal{A}h(x, y, t) = \frac{1}{2}\text{Tr}(\sigma^*(x)D^2 U(x)\sigma(x)) + |y|^2 \left( -1 - \frac{r'(t)}{2r^2(t)} \right) \leq C_T h(x, y, t)$$

ce qui permet d'obtenir le résultat. En revanche, on peut constater que cette fonction ne peut pas jouer le rôle de fonction de Lyapounov en temps long car elle ne génère pas de rappel en la 1ère composante.

- **Hypoellipticité et unicité de la mesure invariante.** L'objectif de cette partie est d'obtenir des conditions raisonnables assurant l'unicité de la mesure invariante associée à (2.3) (dans le cadre homogène). On rappelle qu'un processus de Markov à valeurs dans un espace topologique mesuré admet une unique mesure invariante lorsqu'il est fortement fellerien et qu'il satisfait l'hypothèse d'irréductibilité topologique suivante : il existe un ouvert  $\mathcal{O}$  et  $T > 0$  tel que pour tout  $x$ ,  $P_T(x, \mathcal{O}) > 0$  (cf e.g. [22]). Les deux conditions ci-dessus sont en particulier satisfaites si le processus solution de (2.3) est hypoelliptique (dans un sens précisé ci-dessous) et (approximativement) contrôlable (approximativement signifie que pour tous  $z_1, z_2$  et pour tout  $\varepsilon > 0$ , on peut trouver  $z_3$  dans  $B(z_2, \varepsilon)$  tel que  $z_3$  est atteignable à l'aide du système dynamique issu de  $z_1$ ). L'objectif étant ici d'assurer l'unicité de la mesure invariante, on s'intéresse à l'hypoellipticité uniquement dans le cadre homogène  $r(t) = r_\infty, t \geq 0$  (voir [35] pour le cadre général). Sous sa forme Stratonovich, l'équation (2.3) peut se réécrire sous la forme :

$$dZ_t = A_0(Z_t)dt + \sum_{j=1}^d A_j(Z_t) \circ dW_t^j \quad (2.6)$$

où  $A_0, \dots, A_d$  sont des champs de vecteurs sur  $\mathbb{R}^d$  définis pour tout  $z = (x, y) \in \mathbb{R}^d \times \mathbb{R}^d$  par\* :

$$A_0(z) = \sum_{i=1}^d \left[ -y_i - \frac{1}{2} \sum_{l,j} \sigma_{l,j}(x) \partial_{x_j} \sigma_{i,l}(x) \right] \partial_{x_i} + r_\infty (\partial_{x_i} U(x) - y_i) \partial_{y_i}$$

---

\*. Par un abus de notation standard, on identifie champs de vecteurs et opérateurs différentiels associés.

et pour tout  $j \in \{1, \dots, d\}$  :

$$A_j(z) = \sum_{i=1}^d \sigma_{i,j}(x) \partial_{x_i}.$$

D'après [12] (Théorème 2.9), on sait que sous l'hypothèse de Hörmander suivante (avec l'hypothèse sous-jacente de régularité des coefficients) : il existe  $N \in \mathbb{N}^*$  tel que  $\forall z = (x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ ,

$$\dim(\text{Span}\{A_1(x), A_2(x), \dots, A_d(x), \text{Cr. de Lie de long.} \leq N \text{ des } A_j(z)\text{'s}, 0 \leq j \leq q\}) = 2d, \quad (2.7)$$

alors, pour tout  $t > 0$  et  $x \in \mathbb{R}^d$ ,  $P_t(x, \cdot)$  admet une densité  $p_t(x, \cdot)$  par rapport à la mesure de Lebesgue et que  $(x, y) \mapsto p_t(x, y)$  est  $C^\infty$  sur  $\mathbb{R}^d \times \mathbb{R}^d$ . En particulier,  $x \mapsto P_t(x, \cdot)$  est un semi-groupe fortement fellerien. Si la dynamique du processus empêche le processus de rester dans zones dégénérées, on peut légèrement affaiblir l'hypothèse (2.7) en autorisant que celle-ci ne soit satisfaite que sur un ensemble de complémentaire négligeable. Lorsque  $\sigma$  est non dégénérée, il est clair que  $\partial_{x_1}, \dots, \partial_{x_d}$  sont bien engendrés par  $A_1, \dots, A_d$ . En remarquant ensuite que pour tout  $i \in \{1, \dots, d\}$ ,

$$[\partial_{x_i}, A_0] = \sum_{j=1}^d D^2U(x)_{i,j} \partial_{y_j},$$

on en déduit que l'hypothèse (2.7) sera satisfaite sous une hypothèse de non-dégénérescence de la matrice Hessienne  $D^2U$ . En accord avec les remarques précédentes, on note

$$\mathcal{M}_U := \left\{ x \in \mathbb{R}^d, \det(D^2U(x)) = 0 \right\}, \quad (2.8)$$

et on introduit les hypothèses suivantes :

(**I**<sub>1</sub>) :  $\sigma$  et  $U$  sont de classe  $C^\infty$  et il existe  $\varepsilon_0 > 0$  tel que  $\sigma \sigma^* \geq \varepsilon_0 \text{Id}$ , ( $\sigma$  uniformément elliptique sur  $\mathbb{R}^d$ ).

(**I**<sub>2</sub>) :  $\dim(\mathcal{M}_U) \leq d - 1$ .

On obtient alors le résultat suivant :

**Proposition 2.6 ([35])** *Supposons (**I**<sub>1</sub>), (**I**<sub>2</sub>) satisfaites et  $r(t) = r_\infty$  pour tout  $t \geq 0$ . Alors, le semi-groupe  $(P_t)_{t \geq 0}$  associé à l'équation (2.3) admet une densité  $p_t(z, \cdot)$  (relativement à la mesure de Lebesgue sur  $\mathbb{R}^d \times \mathbb{R}^d$ ). De plus, pour tout  $z \in \mathbb{R}^d \times \mathbb{R}^d$ ,  $z' \mapsto p_t(z, z')$  est  $C^\infty$  sur  $\mathbb{R}^d \times \mathbb{R}^d$  et  $z \mapsto P_t(z, \cdot)$  est fortement fellerien. De plus, si  $\liminf_{|x| \rightarrow +\infty} \frac{U(x)}{|x|} > 0$ , alors pour tout  $z \in \mathbb{R}^d \times \mathbb{R}^d$ ,  $\text{Supp}P_t(z, \cdot) = \mathbb{R}^d \times \mathbb{R}^d$ . Par conséquent, il existe au plus une probabilité invariante  $\nu_{r_\infty}$  pour  $(P_t)_{t \geq 0}$ .*

*De plus, si  $\nu_{r_\infty}$  existe, alors  $\nu_{r_\infty}$  admet une densité notée  $p_{r_\infty}$  par rapport à la mesure de Lebesgue qui est l'unique solution positive de*

$$(y|\nabla_x p_{r_\infty}) + \frac{1}{2} \text{Tr}(\sigma^* D_x^2 p_{r_\infty} \sigma) + r_\infty [(y - \nabla U(x)|\nabla_y p_{r_\infty}) + p_{r_\infty}] = 0. \quad (2.9)$$

*satisfaisant  $\int_{\mathbb{R}^{2d}} p_{r_\infty}(x, y) dx dy = 1$ .*

**Remarque 2.7** Comme cela a été mentionné au début de cette section, l'objectif de ce résultat est de donner des conditions raisonnables (et pas nécessairement optimales) pour l'unicité de la mesure invariante. On peut noter que les hypothèses de régularité des coefficients et d'uniforme ellipticité du théorème peuvent clairement être affaiblies. Il en est de même pour l'hypothèse  $(\mathbf{I}_2)$ . On pourrait en effet s'appuyer sur les crochets de Lie d'ordre supérieur pour engendrer les  $\partial_{y_1}, \dots, \partial_{y_d}$  (voir [35] pour plus de détails). L'équation (2.9) est l'équation de Fokker-Planck stationnaire associée à l'EDS. Celle-ci s'obtient classiquement en utilisant que  $\int \mathcal{A}g(x, y)\nu_{r_\infty}(dx, dy) = 0$  pour toute fonction  $g \in \mathcal{C}^2$  à support compact (avec un petit abus de notation, on note à nouveau  $\mathcal{A}$  l'opérateur sans la dépendance en  $t$ ). On peut remarquer qu'en l'appliquant à  $g(x, y) = x$  et  $g(x, y) = y$ , on obtient

$$\int y\nu_{r_\infty}(dx, dy) = 0 \quad \text{et} \quad \int \nabla U(x)\nu_{r_\infty}(dx, dy) = 0,$$

ce qui montre que le processus est bien asymptotiquement centré sur les zéros de  $\nabla U$ . Contrairement au cadre cinétique introduit dans la remarque 2.1, il ne semble pas possible en général d'obtenir une forme explicite pour  $p_{r_\infty}$  même lorsque  $\sigma$  est constant.

Terminons enfin par quelques commentaires sur la propriété :  $\text{Supp}P_t(z, \cdot) = \mathbb{R}^d \times \mathbb{R}^d$ . Son obtention s'appuie sur une propriété (fine) de "contrôlabilité approximative" sur  $\mathbb{R}^d \times \mathbb{R}^d$  (voir lemme 4.2 de [35]). Lorsque  $\liminf_{|x| \rightarrow +\infty} \frac{U(x)}{|x|} = 0$  ou plus précisément lorsque  $\nabla U$  est bornée, on remarque que  $(Y_t)_{t \geq 0}$  est asymptotiquement borné (en tant que "moyenne du gradient", ses valeurs d'adhérence vivent dans l'enveloppe convexe de  $\{\nabla U(x), x \in \mathbb{R}^d\}$ ). La propriété de support devient alors fautive mais il semble que l'unicité de la mesure invariante puisse être conservée.

### 2.1.3 Comportement en temps long

#### 2.1.3.1 Le cas stable : $\liminf_{t \rightarrow +\infty} r(t) = r_\infty \in (0, +\infty]$

On s'intéresse ici au cas où le système a une faible mémoire (exponentielle ou sur-exponentielle). Dans ce cadre, on se propose de montrer que le système est asymptotiquement stable sous hypothèses de Lyapounov. On introduit donc l'hypothèse de Lyapounov suivante :

$(\mathbf{H}_1)$  : Il existe  $m \in (0, r_\infty)$  et  $\varepsilon \in (0, r_\infty - m)$  tels que

$$\limsup_{|x| \rightarrow +\infty} \left( -m(x|\nabla U(x)) + \frac{1}{2} \text{Tr}(\sigma^*(x)(D^2U(x) + (m + \varepsilon)I_d)\sigma(x)) \right) = -\infty.$$

Dans le cas où  $t \mapsto r(t)$  est variable, on a aussi besoin d'une seconde hypothèse sur le comportement asymptotique de  $(r(t))_{t \geq 0}$  :

$(\mathbf{R}_1)$  :  $\frac{r'(t)}{r^2(t)} \xrightarrow{t \rightarrow +\infty} 0$ .

**Remarque 2.8**  $\triangleright$  On peut noter que dans le cas  $r(t) = r_\infty$  pour tout  $t \geq 0$ , on peut prendre  $\varepsilon = 0$  dans  $(\mathbf{H}_1)$ . Par ailleurs, l'hypothèse de Lyapounov est en général très légèrement plus contraignante que celle intervenant dans le cas sans mémoire.

▷ L'hypothèse  $(\mathbf{R}_1)$  ne joue un rôle que dans le cas où  $(r(t))_{t \geq 0}$  n'est pas constante. Lorsque  $k(t) = \exp(t^\alpha)$  avec  $\alpha > 1$  (cas standard où  $r_\infty = +\infty$ ), on constate que celle-ci est bien satisfaite.

L'hypothèse  $(\mathbf{H}_1)$  est naturellement adaptée à une fonction de Lyapounov relative au système dynamique. Comme on l'a déjà mentionné, la fonction notée  $h$  définie dans la remarque 2.5 ne peut jouer ce rôle. Il nous faut en quelque sorte la compléter pour générer un rappel en la 1ère coordonnée. On introduit donc :

$$V(x, y, t) = U(x) + \frac{|y|^2}{2r(t)} + m + \varepsilon \left( \frac{|x|^2}{2} - \frac{(x|y)}{\rho(t)} \right), \quad (2.10)$$

où  $m > 0$ ,  $\varepsilon > 0$  et  $\rho : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  est définie par

$$\rho(t) = \left( \int_t^{+\infty} \frac{k(t)}{k(s)} ds \right)^{-1}. \quad (2.11)$$

On peut vérifier que la fonction  $\rho$  est bien définie si  $r_\infty \in (0, +\infty]$  et que dans le cas  $r(t) = r_\infty$ ,  $\rho$  n'est autre que la fonction constante  $\rho(t) = r_\infty$ . On reviendra sur le rôle et les propriétés de cette fonction dans la remarque 2.10.

**Proposition 2.9 ([35])** *Supposons que  $\liminf_{t \rightarrow +\infty} r(t) = r_\infty > 0$ . Alors, si  $(\mathbf{H}_1)$  et  $(\mathbf{R}_1)$  sont satisfaites, la fonction  $V : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}$  définie par (2.10) (avec les paramètres  $m$  et  $\varepsilon$  de l'hypothèse  $(\mathbf{H}_1)$ ) vérifie : il existe  $t_0 \geq 0$  tel que  $V$  est positive pour tout  $t \geq t_0$  et,*

$$\limsup_{|(x,y)| \rightarrow +\infty} \left( \sup_{t \geq t_0} \mathcal{A}V(x, y, t) \right) = -\infty. \quad (2.12)$$

Si de plus  $\lim_{t \rightarrow +\infty} r(t) < +\infty$ , alors il existe  $t_1 > 0$  tel que

$$\lim_{|(x,y)| \rightarrow +\infty} \left( \inf_{t \geq t_1} V(x, y, t) \right) = +\infty. \quad (2.13)$$

La fonction  $V$  peut ainsi être qualifiée de fonction de Lyapounov pour (2.3) lorsque  $\lim_{t \rightarrow +\infty} r(t) < +\infty$ . Dans le cas contraire, on verra que la fonction  $V$  conserve un rôle important dans l'obtention de la stabilité lorsque l'hypothèse de Lyapounov est légèrement renforcée.

**Remarque 2.10** La fonction  $\rho$  est calibrée pour avoir les deux propriétés suivantes : elle est solution de  $\dot{u}(t) = u^2(t) - r(t)u(t)$  et satisfait  $\rho(t) \sim r(t)$  lorsque  $t \rightarrow +\infty$ . La première joue un rôle important dans l'obtention de (2.12) (permettant d'annuler les termes produits) tandis que la seconde intervient plutôt dans (2.13).

Ce résultat a la conséquence suivante :

**Corollaire 2.11** *Supposons que  $r_\infty \in \mathbb{R}_+^* \cup \{+\infty\}$  et que  $(\mathbf{H}_0)$ ,  $(\mathbf{H}_1)$  et  $(\mathbf{R}_1)$  soient satisfaites. Alors, pour tout  $z = (x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ , la famille de probabilités  $(\mu_t^z)_{t \geq 1}$  définie par*

$$\mu_t^z(dx, dy) = \frac{1}{t} \int_0^t \mathbb{E}[\delta_{(X_s^z, Y_s^z)}(dx, dy)] ds, \quad t \geq 0$$

*est tendue sur  $\mathbb{R}^d \times \mathbb{R}^d$ . De plus, toute valeur d'adhérence  $\mu_\infty$  de  $(\mu_t^z)_{t \geq 1}$  (pour la convergence faible lorsque  $t \rightarrow +\infty$ ) satisfait*

(i) Si  $r_\infty = +\infty$ , la première marginale de  $\mu_\infty$  est une probabilité invariante pour l'EDS sans mémoire suivante :

$$d\xi_t = -\nabla U(\xi_t)dt + \sigma(\xi_t)dW_t. \quad (2.14)$$

(ii) Si  $r(t) \xrightarrow{t \rightarrow +\infty} r_\infty < +\infty$ , alors  $\mu_\infty$  est une probabilité invariante du processus de Markov homogène solution de (2.3) avec  $r(t) = r_\infty, \forall t \geq 0$ .

Ce résultat montre en particulier l'existence de mesure invariante. Comme dans celui qui va suivre, les assertions se transforment en propriétés de convergence lorsque l'on a unicité de la probabilité invariante.

$(\mu_t^z)_{t \geq 1}$ , construite à partir d'une moyennisation à la fois temporelle et spatiale peut être considérée comme la famille nécessitant les contraintes de Lyapounov les plus faibles pour approcher le régime stationnaire. On s'intéresse naturellement maintenant à la convergence des objets construits à partir exclusivement de moyennisation temporelle ou spatiale.

**Contrôle et convergence de la mesure d'occupation.** On s'intéresse au comportement de  $(\nu_t^z)_{t \geq 1}$  définie par

$$\nu_t^z(\omega, dx, dy) = \frac{1}{t} \int_0^t \delta_{(X_s^z, Y_s^z)}(dx, dy) ds, \quad t \geq 1.$$

Dans ce cadre, on introduit des hypothèses légèrement plus contraignantes :

**(H'\_a)** : Il existe  $a \in (0, 1]$ ,  $\beta \in \mathbb{R}$  et  $\alpha > 0$  tels que

- (i)  $-(x|\nabla U(x)) \leq \beta - \alpha (U(x) \vee |x|^2)^a, \forall x \in \mathbb{R}^d$
- (ii)  $(1 + \text{Tr}(\sigma\sigma^*)(x)) \left( 1 + \frac{|\nabla U(x)|^2}{U(x)} + \|D^2U(x)\| + \| \|D^3U(x)\| \| \right) |x|^{\rightarrow +\infty} o((U(x) \vee |x|^2)^a).$

**Remarque 2.12** L'hypothèse de rappel (i) dépend du paramètre  $a$  qui représente l'intensité de la force de rappel. La croissance de  $\sigma$  est ensuite contrôlée dans l'hypothèse (ii) (de manière adaptée à la première hypothèse). Dans le cas  $U(x) = |x|^q$  ( $q > 0$ ), on peut remarquer que **(H'\_a)** est satisfaite avec  $a = (q/2) \wedge 1$  si  $\|\sigma(x)\sigma^*(x)\| = o(|x|^{q \wedge 2})$  lorsque  $|x| \rightarrow +\infty$ . Contrairement à l'hypothèse **(H\_1)** on constate en revanche qu'elle ne permet pas de gérer le cas (critique)  $U(x) \sim \ln(1 + |x|)^\beta$ .

**Proposition 2.13 ([35])** Supposons que  $r_\infty \in \mathbb{R}_+^* \cup \{+\infty\}$  et que **(H'\_a)** et **(R\_1)** soient satisfaites. Alors, pour tout  $z = (x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ , pour tout  $p \geq 1$ ,

$$\sup_{t \geq 1} \frac{1}{t} \int_0^t \left( (U(X_s^z) \vee |X_s^z|^2)^{p+a-1} + |Y_s^z|^{2p} \right) ds < +\infty \quad p.s. \quad (2.15)$$

En particulier,  $(\nu_t^z(\omega))_{t \geq 1}$  est p.s. tendue sur  $\mathbb{R}^d \times \mathbb{R}^d$ . Si on note  $\nu_\infty^\omega$  une valeur d'adhérence de  $(\nu_t^z(\omega))_{t \geq 1}$  lorsque  $t \rightarrow +\infty$ , alors

- (i) Si  $r_\infty = +\infty$ , alors p.s.,  $\nu_\infty^\omega(dx, dy) = \delta_{\nabla U(x)}(dy)\pi(dx)$  où  $\pi$  une probabilité invariante de (2.14).
- (ii) Si  $r(t) \xrightarrow{t \rightarrow +\infty} r_\infty \in \mathbb{R}_+^*$ ,  $\nu_\infty^\omega$  est p.s une probabilité invariante de (2.3) avec  $r(t) = r_\infty \forall t \geq 0$ .



**Remarque 2.14**  $\triangleright$  Par un argument d'uniforme intégrabilité combiné à la propriété (2.15), on remarque que la convergence de  $\nu_t^z(\omega, f)$  vers  $\nu_\infty^\omega(f)$  est valable pour les fonctions continues à croissance polynomiale.

$\triangleright$  On peut noter que la condition sur  $D^3U$  dans  $(\mathbf{H}'_a)$  n'est en fait utile que lors de l'identification de la limite dans le cas  $r_\infty = +\infty$ .

**Convergence du semi-groupe lorsque  $r(t) = r_\infty$ .** Dans cette partie, on parlera uniquement du cas homogène  $r(t) = r_\infty$  sous les hypothèses d'hypoellipticité  $(\mathbf{I}_1)$  et  $(\mathbf{I}_2)$  introduites dans la proposition 2.6. Il s'agit ici d'établir un résultat de vitesse de convergence du semi-groupe  $(P_t)_{t \geq 0}$  vers l'unique mesure invariante  $\nu_{r_\infty}$ .

**Proposition 2.15 ([35])** *Supposons que  $\forall t \geq 0$ ,  $r(t) = r_\infty > 0$  et que les hypothèses  $(\mathbf{H}'_a)$  ( $a \in (0, 1]$ ),  $(\mathbf{I}_1)$   $(\mathbf{I}_2)$  sont satisfaites. Supposons également que  $\lim_{|x| \rightarrow +\infty} \frac{U(x)}{|x|} = +\infty$  et qu'il existe un minimum  $x^*$  de  $U$  pour lequel  $\det(D^2U(x^*)) \neq 0$ . Alors, pour tous  $z \in \mathbb{R}^d \times \mathbb{R}^d$ ,  $t \geq 0$  et  $p \geq 1$ ,*

$$\sup_{\{f, |f| \leq 1\}} |P_t^{r_\infty}(z_0, f) - \nu(f)| \leq C_{a,p,r_\infty} V_\infty^p(z_0) \begin{cases} \exp(-\gamma_{p,r_\infty} t) & \text{si } a = 1 \\ t^{-\frac{p+a-1}{1-a}} & \text{si } a \in (0, 1). \end{cases}$$

$V_\infty$  est la fonction définie par (2.10) qui s'écrit sous la forme  $V_\infty(z) = U(x) + \frac{r_\infty}{2} \left| x - \frac{y}{r_\infty} \right|^2$  lorsque  $r(t) = r_\infty$ , tandis que  $\gamma_{p,r_\infty}$  et  $C_{a,p,r_\infty}$  sont des constantes positives qui ne dépendent pas de  $z$  et  $t$ .

**Remarque 2.16** Les résultats ci-dessus sont obtenus par des méthodes de type Lyapounov (cf e.g. [30] et [29]). Elles ne permettent en particulier pas d'exhiber des bornes quantitatives. Afin d'obtenir des bornes plus explicites, il semble qu'on puisse s'inspirer des résultats récents de [70] pour les équations de type cinétique (voir Remarque 2.1) dans le cas  $\sigma$  constant. Néanmoins, le caractère non explicite de la mesure invariante dans notre cadre rend le problème plus difficile.

On peut remarquer que la vitesse obtenue n'est plus exponentielle dès que  $a < 1$ . Au vu de résultats obtenus dans l'article [36] (voir Lemme 3.3), il semble qu'on puisse en réalité conserver la vitesse exponentielle tant que  $a > 1/2$ , ce qui correspondrait alors à la "transition de phase" du cadre sans mémoire.

**Attraction du système sans mémoire lorsque  $r(t) \rightarrow +\infty$ .** Comme on l'a déjà remarqué lors de l'études des mesures d'occupation ou dans l'exemple gaussien étudié rapidement en début de section, le régime limite naturel des solutions de (2.3) lorsque  $r_\infty = +\infty$  est le régime stationnaire du régime sans mémoire (2.14). On souhaite donc dans cette partie obtenir un résultat de convergence de la loi de la marginale du processus vers la mesure invariante de (2.14). Les techniques précédentes étant moins naturelles dans un cadre non homogène, ce problème est abordé sous un angle différent du précédent. Nous faisons ici l'hypothèse de "convexité" suivante :

(AC) : Il existe  $\rho > 0$  tel que pour tous  $x_1, x_2 \in \mathbb{R}^d$ ,

$$(x_1 - x_2 | \nabla U(x_2) - \nabla U(x_1)) + \frac{1}{2} \text{Tr}((\tilde{\sigma} \tilde{\sigma}^*)(x_1, x_2)) \leq -\rho |x_2 - x_1|^2. \quad (2.16)$$

où  $\tilde{\sigma}(x, y) = \sigma(y) - \sigma(x)$ . Dans le cadre sans mémoire, cette hypothèse assure la confluence asymptotique (à vitesse exponentielle en distance de Wasserstein) des trajectoires construites avec le même Brownien (et ayant deux conditions initiales différentes) et implique (par conséquent) l'unicité de la mesure invariante. En notant  $W_2$  la distance de Wasserstein sur  $\mathcal{P}(\mathbb{R}^d)$  définie pour tout couple de probabilités  $(\nu_1, \nu_2)$ ,

$$W_2(\nu_1, \nu_2) = \inf_{\mu \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d), P_1\mu = \nu_1, P_2\mu = \nu_2} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \mu(dx, dy),$$

on obtient ici le résultat suivant :

**Proposition 2.17** ([35]) *Supposons que  $r_\infty = +\infty$ , que  $(\mathbf{H}'_1)$  et  $(\mathbf{AC})$  soient satisfaites et que  $\nabla U$  et  $\sigma$  sont localement lipschitziennes. Notons  $\pi$  l'unique mesure invariante de (2.14). Si  $t \rightarrow \int_0^t k(s)ds/k(t)$  est une fonction décroissante qui tend vers 0 lorsque  $t \rightarrow +\infty$ , alors, il existe des constantes (positives)  $C_1$  et  $C_2$  telles que pour tout  $t \geq 0$ ,*

$$\forall u \in [0, t] \quad W_2(\mathbb{P}_{X_t}, \pi) \leq \frac{C_1}{k(u)} \int_0^u k(v)dv + C_2 e^{-\rho(t-u)}. \quad (2.17)$$

En particulier,  $W_2(\mathbb{P}_{X_t}, \pi) \implies 0$  lorsque  $t \rightarrow +\infty$ .

En optimisant le choix de  $u$  dans l'inégalité précédente, on obtient le corollaire suivant :

**Corollaire 2.18** *Sous les conditions de la proposition 2.17, on a*

$$W_2(\mathbb{P}_{X_t}, \pi) \leq \frac{C_1}{r \circ H^{-1}(t)} + C_2 \left( 1 + \frac{\rho}{r \circ H^{-1}(t)} \right) e^{-\rho(t-H^{-1}(t))}$$

où  $H$  est définie par :

$$H(u) = u - \frac{1}{\rho} \log \left[ \frac{C_1}{C_2 \rho} \left( \frac{r(u)}{k(u)} \int_0^u k(v)dv + \frac{k(0)}{k(u)} - 1 \right) \right].$$

**Remarque 2.19**  $\triangleright$  Dans la proposition 2.17, l'instant  $u$  correspond au temps que l'on attend avant de comparer la dynamique de  $(X_t)_{t \geq 0}$  à celle de la diffusion sans mémoire. Plus précisément, on estime la distance  $L^2$  de  $X_t$  à  $(\xi_{t-u}^{X_u})$  où  $(\xi_s^{X_u})_{s \geq 0}$  désigne la solution de (2.14) de condition initiale  $X_u$ , construite avec les accroissements du Brownien du premier après l'instant  $u$ . En notant  $(\xi_s^\pi)_{s \in [0, t-u]}$ , une solution stationnaire de (2.14) construite encore une fois avec les mêmes accroissements, on a donc :

$$W_2(\mathbb{P}_{X_t}, \pi) \leq \mathbb{E}[|X_t - \xi_{t-u}^{X_u}|^2]^{\frac{1}{2}} + \mathbb{E}[|\xi_{t-u}^{X_u} - \xi_{t-u}^\pi|^2]^{\frac{1}{2}}.$$

Ces deux termes sont alors respectivement contrôlés à l'aide de  $(\mathbf{AC})$  par les deux termes du membre de droite de (2.17).

$\triangleright$  Les vitesses obtenues ci-dessus dépendent fortement de la vitesse de décroissance de  $t \rightarrow \int_0^t k(s)ds/k(t)$  vers 0 qui caractérise la vitesse de rapprochement à la dynamique sans mémoire. On propose ici de rendre plus explicites ces résultats dans des exemples caractéristiques :

- Si  $k(t) = \alpha t^{\alpha-1} e^{t^\alpha}$  avec  $\alpha > 1$ , alors  $r(t) \sim \alpha t^{2\alpha-2}$ ,  $H^{-1}(t) \sim t$  et  $t - H^{-1}(t) \sim \frac{\alpha}{\rho} \log t$ .  
On obtient alors une vitesse polynômiale :

$$W_2(\mathbb{P}_{X_t}, \pi) \leq C t^{1-\alpha}.$$

- Si  $k(t) = e^t e^{e^t}$ , alors  $H^{-1}(t) \sim \frac{\rho}{\rho+1} t$  ce qui mène à la borne suivante :

$$W_2(\mathbb{P}_{X_t}, \pi) \leq C e^{-\frac{\rho}{\rho+1} t}.$$

#### 2.1.4 Comportement en temps long lorsque $r_\infty = 0$

On s'intéresse enfin au comportement asymptotique dans le cas longue mémoire  $r(t) \rightarrow 0$ . Pour rappel, ce cas correspond aux exemples suivants :  $k(t) = (1+t)^\alpha$  avec  $\alpha > 0$  ou  $k(t) = e^{(1+t)^\alpha}$  avec  $\alpha \in (0, 1)$ . Comme cela a été déjà suggéré dans l'exemple gaussien, le système a tendance à exploser en temps long dans ce cadre. Cette propriété est précisée dans les deux résultats qui suivent, d'abord dans un cadre assez général puis dans le cadre gaussien unidimensionnel où l'on réussit à caractériser de manière précise le comportement asymptotique des moments d'ordre 2 de  $X_t$  et  $Y_t$  (lorsque  $k(t) = (1+t)^\alpha$ ).

**Proposition 2.20 ([35])** *Supposons que  $|\nabla U|^2 \leq C(1+U)$  et qu'il existe  $\lambda_0 > 0$  tel que  $\text{Tr}(\sigma^* D^2 U \sigma)(x) \geq \lambda_0 > 0$ . Alors, si  $r(t) \rightarrow 0$  et que l'on peut trouver  $t_0 \geq 0$  tel que  $r'(t) + 2r^2(t) \geq 0$  pour tout  $t \geq t_0$ , on a pour tout  $z = (x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ ,*

$$\limsup_{t \rightarrow +\infty} r(t) \mathbb{E}[|X_t^z|^2] > 0.$$

*En particulier, il existe une sous-suite  $(t_n)_{n \geq 1}$  telle que  $\mathbb{E}[|X_{t_n}^z|^2] \rightarrow +\infty$ .*

**Remarque 2.21**  $\triangleright$  On peut remarquer que la condition  $r'(t) + 2r^2(t) \geq 0$  pour  $t$  assez grand est satisfaite lorsque  $k(t) = e^{(1+t)^\alpha}$  avec  $\alpha \in (0, 1)$  et  $k(t) = (1+t)^\alpha$  avec  $\alpha > \frac{1}{2}$ . En particulier, cela contient le cas  $Y_t = \frac{1}{1+t} \int_0^t \nabla U(X_s) ds$ .

$\triangleright$  Le résultat ci-dessus est obtenu en supposant que  $U$  est sous-quadratique (propriété générée par l'hypothèse  $|\nabla U|^2 \leq C(1+U)$ ) et que la composante de bruit est non dégénérée. Sous ce type d'hypothèse, l'idée de la preuve consiste alors à exhiber une sorte d'"anti-fonction de Lyapounov" pour le système dynamique.

Considérons maintenant le cas gaussien unidimensionnel.

**Proposition 2.22 ([35])** *Supposons que  $d = 1$ ,  $U(x) = x^2/2$  et  $k(t) = (1+t)^\alpha$  avec  $\alpha > 1/2$ . Alors, (i)*

$$\lim_{t \rightarrow \infty} \mathbb{E} Y_t^2 = \frac{\alpha}{2\alpha + 1}, \quad \text{et} \quad \mathbb{E} X_t^2 \sim \frac{t}{2\alpha + 1} \quad \text{as } t \rightarrow +\infty.$$

(ii)

$$\left( \sqrt{\frac{2\alpha+1}{t}} X_t, \sqrt{\frac{2\alpha+1}{\alpha}} Y_t \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I_2) \quad \text{as } t \rightarrow +\infty.$$

**Remarque 2.23**  $\triangleright$  La propriété (ii) montre en particulier que  $\mathbb{E}[Y_t] \rightarrow 0$  lorsque  $t \rightarrow +\infty$ . On peut également prouver qu'il en est de même pour  $(\mathbb{E}[X_t])_{t \geq 0}$ .

$\triangleright$  La preuve de ce résultat est basée sur l'étude de l'évolution de la matrice de covariance. Plus exactement, on montre que le vecteur (renormalisé)  $z(t) = (r(t)\mathbb{E}[X_t^2], \mathbb{E}[Y_t^2], \mathbb{E}[X_t Y_t])'$  est solution d'un système différentiel (déterministe) asymptotiquement stationnaire. La limite s'obtient alors en identifiant le régime limite de ce système différentiel.

## 2.2 Grandes Déviations pour le régime stationnaire de diffusions à gradient moyenné

### 2.2.1 Introduction

Comme on l'a déjà mentionné au début de la section précédente, l'étude des diffusions à gradient moyenné est en partie motivée par des applications à des problèmes d'optimisation stochastique. Nous proposons ici de nous intéresser à une question d'intérêt pour le problème du recuit simulé : l'étude fine du comportement asymptotique de la mesure invariante du système dynamique lorsque le coefficient de diffusion  $\varepsilon$  tend vers 0. Plus précisément, l'objectif est d'obtenir un Principe de Grandes Déviations (PGD) pour cette famille de mesures invariantes et de montrer (sous conditions) que la mesure invariante se concentre (dans un sens défini dans la suite) sur l'ensemble des minima globaux du potentiel  $U$  considéré.

Afin de mieux comprendre les liens entre le problème de recuit simulé et les grandes déviations pour la mesure invariante, considérons dans un premier temps le modèle classique sans mémoire. Dans ce cadre, considérer le problème du recuit simulé dans un cadre continu revient à étudier le comportement en temps long de la diffusion inhomogène

$$d\xi_t = -\nabla U(\xi_t)dt + \varepsilon(t)dB_t$$

où  $t \mapsto \varepsilon(t)$  est une fonction (déterministe) décroissant vers 0. On peut alors montrer (cf [69]) que si  $\varepsilon(t) = k/(\ln t)$  avec  $k > c$  où  $c$  est une quantité semi-explicite, alors le processus se concentre (en probabilité) sur les minima globaux de  $U$ . Ce résultat peut être obtenu en considérant directement l'évolution de la loi du processus inhomogène  $(\xi_t)$ . Une seconde approche consiste à introduire la diffusion à "petit bruit" constant

$$d\xi_t^\varepsilon = \varepsilon dB_t - \nabla U(\xi_t^\varepsilon)dt, \quad (2.18)$$

et de faire une étude fine et séparée de

- la convergence de la loi à l'instant  $t$  de  $\xi_t^\varepsilon$  à la mesure invariante  $\nu_\varepsilon^\xi$  à  $\varepsilon$  fixé,
- la manière dont se concentre  $\nu_\varepsilon^\xi$  autour des zéros de  $\nabla U$  lorsque  $\varepsilon \rightarrow 0$ .

On propose ici de se focaliser sur le second point de l'étude ci-dessus lorsque le processus est une diffusion à gradient moyenné. En accord avec la section précédente, on introduit

$$dX_t^\varepsilon = \varepsilon dB_t - \left( \frac{1}{k(t)} \int_0^t k'(s) \nabla U(X_s^\varepsilon) ds \right) dt. \quad (2.19)$$

Contrairement à la section précédente, on considèrera uniquement le cadre homogène  $k(t) = \exp(\lambda t)$  ( $\lambda > 0$ ). Il s'agit en réalité du seul cadre nouveau adapté au problème. En effet, si

$r(t) = (\dot{k}/k)(t)$  tend vers  $+\infty$ , le régime limite est celui du système sans mémoire et a déjà été largement étudié tandis que dans le cas  $r(t) \rightarrow 0$ , il n'y a pas de régime stationnaire.

Cela nous conduit donc à l'étude du comportement de la famille  $(\nu_\varepsilon)_{\varepsilon>0}$  de probabilités invariantes de l'équation

$$\begin{cases} dX_t^\varepsilon = \varepsilon dB_t - Y_t^\varepsilon dt, \\ dY_t^\varepsilon = \lambda(\nabla U(X_t^\varepsilon) - Y_t^\varepsilon) dt. \end{cases} \quad (2.20)$$

où  $U : \mathbb{R}^d \rightarrow \mathbb{R}$  est une fonction positive régulière et coercive et  $\lambda$  est un réel strictement positif. Dans la suite, on notera  $Z^\varepsilon = (X^\varepsilon, Y^\varepsilon)$ ,  $(P_t^{(\varepsilon)}(z, \cdot))_{z,t}$  le semi-groupe associé à (2.20) et  $b$  le champ de vecteurs associé à ce système défini par :

$$b(x, y) = \begin{pmatrix} 0 & -y \\ \lambda \nabla U(x) & -\lambda y \end{pmatrix}. \quad (2.21)$$

On rappelle qu'à  $\varepsilon$  fixé strictement positif, on a existence et unicité de  $\nu_\varepsilon$  sous les hypothèses  $(\mathbf{H}_1)$ ,  $(\mathbf{I}_1)$ ,  $(\mathbf{I}_2)$  et  $\lim_{|x| \rightarrow +\infty} \frac{U(x)}{|x|} = +\infty$ . Afin que le problème soit bien posé, on travaillera a minima sous ces hypothèses.

Avant d'aller plus loin, rappelons quelques aspects de la théorie des grandes déviations (voir e.g. [25] pour une approche plus complète).

**Rappels sur les grandes déviations.** Soit  $(E, d)$  un espace métrique. Une famille de probabilités  $(\nu_\varepsilon)_{\varepsilon>0}$  sur cet espace satisfait un principe de grandes déviations (PGD) de vitesse  $r_\varepsilon$  et de fonction de taux  $I$  si pour tout ouvert  $O$  et pour tout fermé  $F$ ,

$$\liminf_{\varepsilon \rightarrow 0} r_\varepsilon \log(\nu_\varepsilon(O)) \geq - \inf_{x \in O} I(x) \quad \text{et} \quad \limsup_{\varepsilon \rightarrow 0} r_\varepsilon \log(\nu_\varepsilon(F)) \leq - \inf_{x \in F} I(x).$$

La fonction  $I$  est appelée *bonne fonction de taux* si pour tout  $c \in \mathbb{R}$ ,  $\{x \in E, I(x) \leq c\}$  est compact.

On rappelle également que  $(\nu_\varepsilon)_{\varepsilon>0}$  est dite *exponentiellement tendue* de vitesse  $r_\varepsilon$  si pour tout  $a > 0$ , il existe un compact  $K_a$  de  $E$  tel que

$$\limsup_{\varepsilon \rightarrow 0} r_\varepsilon \log(\nu_\varepsilon(K_a^c)) \leq -a.$$

On a alors un résultat similaire au théorème de Prokhorov pour la convergence en loi (voir [31]).

**Proposition 2.24** *Soit  $(S, d)$  un espace polonais. Supposons que  $(\nu_\varepsilon)_{\varepsilon \geq 0}$  soit une famille de probabilités exponentiellement tendue de vitesse  $r_\varepsilon$ . Alors, il existe une sous-suite  $(\varepsilon_k)_{k \geq 0}$  telle que  $\varepsilon_k \rightarrow 0$  et telle que  $(\nu_{\varepsilon_k})_{k \geq 0}$  satisfait un PGD de vitesse  $r_{\varepsilon_k}$ . De plus, la fonction de taux  $I$  associée est alors une bonne fonction de taux.*

Une telle sous-suite  $(\nu_{\varepsilon_k})_{k \geq 1}$  sera appelée une sous-suite *(LD)-convergente* et la fonction de taux associée, une *(LD)-limite*.

Dans le cadre sans mémoire, l'obtention du PGD pour les mesures invariantes est relativement aisée car la mesure invariante  $\mu_\varepsilon$  est explicite :

$$\mu_\varepsilon(dx) = \frac{\exp(-\frac{2U(x)}{\varepsilon^2})dx}{Z_\varepsilon} \quad \text{où} \quad Z_\varepsilon = \int \exp(-\frac{2U(x)}{\varepsilon^2})dx.$$

On peut alors montrer en s'appuyant sur la méthode de Laplace que  $(\mu_\varepsilon)_\varepsilon$  satisfait un PGD de vitesse  $\varepsilon^2$  et de fonction de taux  $I(x) = 2(U(x) - U_0)$  où  $U_0$  est le minimum global de  $U$ . De plus,  $\mu_\varepsilon$  se concentre (au sens de la convergence étroite) sur l'ensemble des minima globaux (voir [50]). Dans un cadre où la mesure invariante est non explicite, on doit comprendre le comportement de la mesure invariante au travers de la dynamique du processus. À cette fin, il nous est d'abord nécessaire d'introduire des hypothèses de stabilité adaptées au problème.

**Hypothèses de Lyapounov.** Les hypothèses de Lyapounov de la section précédente induisent simplement la tension de  $(\nu_\varepsilon)_{\varepsilon>0}$ . Afin d'obtenir une tension exponentielle, on travaillera ici sous les hypothèses plus fortes suivantes :

**(H<sub>Q+</sub>)** : Il existe  $\rho \in (0, 1)$ ,  $C > 0$ ,  $\beta \in \mathbb{R}$  et  $\alpha > 0$  tels que

$$(i) \quad -\langle x, \nabla U(x) \rangle \leq \beta - \alpha U(x), \forall x \in \mathbb{R}^d$$

$$(ii) \quad |\nabla U|^2 \leq C(1 + U^{2(1-\rho)}) \quad \text{et} \quad \lim_{|x| \rightarrow +\infty} \frac{\|D^2 U(x)\|}{U(x)} = 0.$$

**(H<sub>Q-</sub>)** : Il existe  $a \in (1/2, 1]$ ,  $C > 0$ ,  $\beta \in \mathbb{R}$  et  $\alpha > 0$  tels que

$$(i) \quad -\langle x, \nabla U(x) \rangle \leq \beta - \alpha |x|^{2a}, \forall x \in \mathbb{R}^d$$

$$(ii) \quad |\nabla U|^2 \leq C(1 + U) \quad \text{et} \quad \sup_{x \in \mathbb{R}^d} \|D^2 U(x)\| < +\infty.$$

**Remarque 2.25** Les hypothèses **(H<sub>Q+</sub>)** et **(H<sub>Q-</sub>)** correspondent respectivement aux potentiels  $U$  à croissance surquadratique et sous-quadratique. Par exemple, si l'on considère  $U(x) = (1 + |x|^2)^p$ , on constate que si  $p \geq 1$ , **(H<sub>Q+</sub>)** est satisfaite pour tout  $\rho \in (0, \frac{1}{2p})$  tandis que si  $p \in (1/2, 1]$ , **(H<sub>Q-</sub>)** est vraie avec  $a = p$ . Ces hypothèses ne sont en revanche pas adaptées aux potentiels à croissance sous-linéaire ou exponentielle.

### 2.2.2 Tension exponentielle

Afin d'obtenir la tension exponentielle de  $(\nu_\varepsilon)_{\varepsilon>0}$ , il nous est nécessaire d'obtenir un contrôle exponentiel de  $\nu_\varepsilon(|z| > R)$  (uniforme en  $\varepsilon$ ). Lorsque que la mesure invariante n'est pas explicite, il nous faut donc exploiter l'ergodicité du processus pour contrôler ce type de quantité. Si l'on fait l'hypothèse que pour tout  $\varepsilon > 0$ , le semi-groupe  $(P_t^{(\varepsilon)}(z, \cdot))_{t \geq 0}$  converge en loi vers  $\nu_\varepsilon$  et que  $\nu_\varepsilon$  est absolument continue par rapport à la mesure de Lebesgue, on peut par exemple s'appuyer sur la propriété suivante :

$$\forall R > 0, \quad \nu_\varepsilon(|z| > R) = \lim_{\rho \rightarrow 0} \mathbb{E} \left[ \int_0^{+\infty} e^{-\rho t} 1_{|Z_t^\varepsilon| > R} dt \right].$$

C'est le point de vue adopté par Puhalskii dans [87] (lemme 7) pour obtenir un critère de tension exponentielle basé sur le contrôle des temps de retour dans les compacts. Dans la proposition suivante, nous montrons que les hypothèses de ce critère sont bien satisfaites dans notre cadre.

**Proposition 2.26 ([36])** *Supposons que **(H<sub>Q+</sub>)** ou **(H<sub>Q-</sub>)** soit satisfaite. Alors, il existe un compact  $B$  de  $\mathbb{R}^d \times \mathbb{R}^d$  tel que le temps d'atteinte  $\tau_\varepsilon$  de  $B$  ( $\tau_\varepsilon := \inf\{t > 0, Z_t^\varepsilon \in B\}$ ) satisfait*

les trois propriétés suivantes :

(i) Pour tout compact  $K$  de  $\mathbb{R}^d \times \mathbb{R}^d$ ,

$$\limsup_{\varepsilon \rightarrow 0} \sup_{z \in K} \mathbb{E}_z[(\tau_\varepsilon)^2] < \infty. \quad (2.22)$$

(ii) Il existe  $\delta > 0$  tel que pour tout compact  $K$  de  $\mathbb{R}^d \times \mathbb{R}^d$ ,

$$\limsup_{\varepsilon \rightarrow 0} \sup_{z \in K} \sup_{t \geq 0} \mathbb{E}_z \left[ |Z_{t \wedge \tau_\varepsilon}^\varepsilon|^{\frac{\delta}{\varepsilon^2}} \right]^{\varepsilon^2} < +\infty. \quad (2.23)$$

(iii) Pour tout compact  $K$  de  $\mathbb{R}^d \times \mathbb{R}^d$  tel que  $K \cap B = \emptyset$ ,

$$\liminf_{\varepsilon \rightarrow 0} \inf_{z \in K} \mathbb{E}_z[\tau_\varepsilon] > 0. \quad (2.24)$$

**Remarque 2.27** La propriété (iii) est une conséquence assez directe de la convergence en loi du processus  $(Z_t^\varepsilon)_{t \in [0,1]}$  vers la solution de  $\dot{z} = b(z)$ . Les deux autres se déduisent d'arguments de type Lyapounov (uniformes en  $\varepsilon$ ) basés sur la fonction  $V$  définie par (2.10) (avec  $r(t) = \rho(t) = \lambda$ ). C'est naturellement à ce stade qu'interviennent les hypothèses  $(\mathbf{H}_{\mathbf{Q}_+})$  et  $(\mathbf{H}_{\mathbf{Q}_-})$  ainsi que la limitation à  $a > 1/2$ .

On en déduit le résultat suivant :

**Proposition 2.28** ([36]) *Supposons que  $(\mathbf{I}_1)$  et  $(\mathbf{I}_2)$  soient satisfaites, que  $\lim_{|x| \rightarrow +\infty} \frac{U(x)}{|x|} = +\infty$  et que  $(\mathbf{H}_{\mathbf{Q}_+})$  ou  $(\mathbf{H}_{\mathbf{Q}_-})$  le soit également. Alors,  $(\nu_\varepsilon)_{\varepsilon \in (0,1]}$  est exponentiellement tendue sur  $\mathbb{R}^{2d}$  de vitesse  $\varepsilon^2$ .*

### 2.2.3 Caractérisation des $(LD)$ -limites

En accord avec la proposition 2.24, il nous faut maintenant identifier les  $(LD)$ -limites, *i.e.* les fonctions de taux associées aux sous-suites  $(LD)$ -convergentes. Nous utilisons deux types de caractérisations. Dans la première, on montre que toute  $(LD)$ -limite est solution d'une équation de type Hamilton-Jacobi. Dans la seconde, on s'appuie sur les travaux de Freidlin et Wentzell. Ces deux approches sont évidemment liées mais la seconde donne une forme plus explicite et donc plus utilisable pour en déduire la concentration de  $(\nu_\varepsilon)_{\varepsilon > 0}$  sur les minima globaux de  $U$ .

#### 2.2.3.1 $(LD)$ -limite et contrôle optimal

Dans cette partie, nous fournissons une première caractérisation des  $(LD)$ -limites comme solutions d'un problème de contrôle optimal. On introduit auparavant des notations. On note  $\mathbb{H}(\mathbb{R}_+, \mathbb{R}^d)$  l'espace (de Cameron-Martin) des fonctions  $w : \mathbb{R}_+ \rightarrow \mathbb{R}^d$  absolument continues telles que  $\dot{w} \in L^{2,\text{loc}}(\mathbb{R}_+, \mathbb{R}^d)$ . Pour tout  $\varphi \in \mathbb{H}(\mathbb{R}_+, \mathbb{R}^d)$ ,  $\mathbf{z}_\varphi := (\mathbf{z}_\varphi(t))_{t \geq 0}$  et  $\tilde{\mathbf{z}}_\varphi := (\tilde{\mathbf{z}}_\varphi(t))_{t \geq 0}$  sont les solutions respectives de

$$\dot{\mathbf{z}}_\varphi = b(\mathbf{z}_\varphi) + \begin{pmatrix} \dot{\varphi} \\ 0 \end{pmatrix} \quad \text{et} \quad \dot{\tilde{\mathbf{z}}}_\varphi = -b(\tilde{\mathbf{z}}_\varphi) + \begin{pmatrix} \dot{\varphi} \\ 0 \end{pmatrix}. \quad (2.25)$$

Pour tout  $z \in \mathbb{R}^2$ , on note  $\mathbf{z}_\varphi(z, \cdot)$  et  $\tilde{\mathbf{z}}_\varphi(z, \cdot)$  les solutions issues de  $z$  (celles-ci sont bien définies de manière unique sur  $\mathbb{R}^+$  sous les hypothèses des résultats à venir). On remarque que  $\tilde{\mathbf{z}}_\varphi$  correspond au système contrôlé associé à la dynamique retournée en temps. Nous introduisons également une hypothèse sur les points critiques de  $U$  :

**(H<sub>D</sub>)** : L'ensemble des points critiques  $(x_i^*)_{i=1, \dots, \ell}$  de  $U$  est fini et  $\forall i \in \{1, \dots, \ell\}$ ,  $D^2U(x_i^*)$  est inversible.

On a alors la proposition suivante :

**Proposition 2.29** ([36]) *Supposons que les hypothèses de la proposition 2.28 soient satisfaites. Considérons une sous-suite  $(\nu_{\varepsilon_n})_{n \geq 1}$  (LD)-convergente de la famille (exponentiellement tendue)  $(\nu_\varepsilon)_{\varepsilon \in (0,1]}$  et notons  $W : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  la (bonne) fonction de taux associée. Alors,*

(i)  $W$  satisfait pour tout  $z \in \mathbb{R}^d \times \mathbb{R}^d$  :

$$\forall t \geq 0, \quad W(z) = \inf_{\varphi \in \mathbb{H}} \left[ \frac{1}{2} \int_0^t |\dot{\varphi}(s)|^2 ds + W(\tilde{\mathbf{z}}_\varphi(z, t)) \right]. \quad (2.26)$$

(ii) Si de plus, **(H<sub>D</sub>)** est satisfaite alors,

$$W(z) = \min_{1 \leq i \leq \ell} \left\{ \begin{array}{l} \inf_{\varphi \in \mathbb{H}} \left[ \frac{1}{2} \int_0^\infty |\dot{\varphi}(s)|^2 ds + W(z_i^*) \right] \\ \tilde{\mathbf{z}}_\varphi(z, +\infty) = z_i^* \end{array} \right\} \quad (2.27)$$

où  $\tilde{\mathbf{z}}_\varphi(z, +\infty) := \lim_{t \rightarrow +\infty} \tilde{\mathbf{z}}_\varphi(z, t)$  (si celle-ci existe) et  $\forall i \in \{1, \dots, \ell\}$ ,  $z_i^* := (x_i^*, 0)$ .

**Remarque 2.30**  $\triangleright$  Une solution  $W$  des équations ci-dessus peut aussi être vue comme une solution (de viscosité) stationnaire de l'équation de Hamilton-Jacobi-Bellman

$$\partial_t v(z, t) + \min_{u \in \mathbb{R}^d} (|\partial_z v(z, t)| - b(z) - (u, 0_{\mathbb{R}^d})' + \frac{1}{2}|u|^2) = 0$$

ou encore du problème de Cauchy non linéaire

$$\partial_t v(z, t) - (\partial_z v(z, t)|b(z)) - \frac{1}{2}|\partial_x v(z, t)|^2 = 0.$$

On renvoie par exemple à [3] pour plus de détails sur les liens entre ces différentes formulations du problème. Notons que les propriétés établies ci-dessus permettent donc en particulier d'assurer l'existence de solutions à ce type d'équations (via une approche probabiliste).

$\triangleright$  Pour passer de (i) à (ii), on montre qu'à  $z$  fixé, on peut construire  $\hat{\varphi} \in \mathbb{H}$  minimisant le problème (2.26) pour tout  $t \geq 0$  puis que la trajectoire contrôlée  $\tilde{\mathbf{z}}_{\hat{\varphi}}(z, \cdot)$  associée converge vers l'un des  $z_i^*$  lorsque  $t \rightarrow +\infty$ . Ce dernier point est relativement délicat car il s'agit d'une convergence pour la dynamique retournée qui n'a pas naturellement de propriétés de stabilisation. L'idée principale est de montrer que  $\tilde{\mathbf{z}}_{\hat{\varphi}}(z, \cdot)$  est une trajectoire pseudo-asymptotique associée à l'o.d.e.  $\dot{z} = \tilde{b}(z)$  et que  $\{z_i^*, i \in \{1, \dots, \ell\} = \{z, \tilde{b}(z) = 0\}$  est un attracteur pour l'ensemble des valeurs d'adhérence de  $\tilde{\mathbf{z}}_{\hat{\varphi}}(z, t)$  lorsque  $t \rightarrow +\infty$  (voir [36] pour plus de détails).



▷ La preuve de (i) se réalise en trois étapes. Dans un premier temps, on montre que si  $z_\varepsilon \rightarrow z$ , la famille  $(\mathbf{P}^{\varepsilon, z_\varepsilon})_{\varepsilon > 0}$  (où  $\mathbf{P}^{\varepsilon, z_\varepsilon}$  désigne la loi d'une solution de (2.20) issue  $z_\varepsilon$ ) satisfait un PGD sur  $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d \times \mathbb{R}^d)$  de fonction de taux

$$I_t^z(z') = \inf_{z(\cdot) \in \mathcal{Z}_t(z, z')} \mathcal{I}_z(z(\cdot)) \quad (2.28)$$

où  $\mathcal{Z}_t(z, z')$  est l'ensemble des fonctions absolument continues de  $\mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  telles que  $z_\varphi(0) = z$ , et  $z_\varphi(t) = z'$  et

$$\mathcal{I}_z((z(t))_{t \geq 0}) = \inf_{\varphi \in \mathbb{H}, z_\varphi(z, \cdot) = z(\cdot)} \frac{1}{2} \int_0^\infty |\dot{\varphi}(s)|^2 ds = \frac{1}{2} \int_0^\infty |\dot{x}(s) + y(s)|^2 ds.$$

Via une adaptation du Corollaire 1 de [87], on exploite dans un second temps la stationnarité au travers de l'égalité suivante :

$$\int h \frac{1}{\varepsilon^2} d\nu_\varepsilon = \int \mathbb{E}[h \frac{1}{\varepsilon^2} (Z_t^{\varepsilon, z})] d\nu_\varepsilon \quad (\text{où } h : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R} \text{ est une fonction mesurable bornée),}$$

pour en déduire que la fonction  $W$  définie ci-dessus satisfait

$$\forall t \geq 0, \quad \forall z_0 \in \mathbb{R}^d \times \mathbb{R}^d, \quad W(z_0) = \inf_{z \in \mathbb{R}^{2d}} (I_t(z, z_0) + W(z)). \quad (2.29)$$

La dernière étape consiste à retourner en temps l'expression ci-dessus.

### 2.2.3.2 Unicité et caractérisation de $W(z_i^*)$

La seconde formulation de  $W$  obtenue dans la proposition ci-dessus (voir (2.27)) permet d'obtenir l'unicité de  $W$  conditionnellement aux valeurs de  $W(z_i^*)$ ,  $i = 1, \dots, \ell$ . Afin d'obtenir un PGD, il nous faut maintenant établir l'unicité de  $(W(z_i^*))_{i \in \{1, \dots, \ell\}}$ . Notre objectif principal étant en fait de prouver que  $(\nu_\varepsilon)_{\varepsilon \in (0, 1]}$  se concentre sur les minima globaux, il nous est aussi nécessaire d'obtenir une caractérisation relativement explicite de  $W(z_i^*)$ ,  $i = 1, \dots, \ell$ . Notons qu'avec la formulation (2.27), on sait pour l'instant que  $W$  atteint son minimum (nécessairement égal à 0) sur  $\{W(z_i^*), i = 1, \dots, \ell\}$  et que les valeurs de  $W(z_i^*)$  sont interconnectées via le coût pour aller d'un point critique à un autre.

Pour aller plus loin, on s'appuie sur la théorie de Freidlin et Wentzell (voir [33]) que l'on adapte à notre cadre de travail. Le point de départ est de représenter la mesure invariante à partir des temps de retour dans le voisinages de l'ensemble des points critiques et d'une chaîne squelette associée. La construction est la suivante. Autour de chaque  $z_i^*$ , on considère la boule fermée  $g_i = \bar{B}(z_i^*, \rho)$  ( $\rho > 0$ ) dont la frontière est notée  $\partial g_i$ . Sous l'hypothèse  $(\mathbf{H}_D)$ , il existe  $\rho_0$  tel que pour tout  $\rho \in (0, \rho_0)$ ,  $g_i \cap g_j = \emptyset$  pour tout  $i \neq j$ . On note

$$\Gamma = \cup_{i=1}^{\ell} B(z_i^*, \rho_0)^c \quad \text{et} \quad g = \cup_{i=1}^{\ell} g_i.$$

On introduit alors les suites de temps d'arrêt  $(\tau'_n(\Gamma))_{n \geq 1}$  et  $(\tau_n(\partial g))_{n \geq 1}$  définies récursivement de la manière suivante :

$$\tau'_1(\Gamma) = \inf\{t \geq 0, Z_t^{\varepsilon, z} \in \partial \Gamma\}, \quad \tau_1(\partial g) = \inf\{t > \tau'_1(\Gamma), Z_t^{\varepsilon, z} \in \partial g\}. \quad (2.30)$$

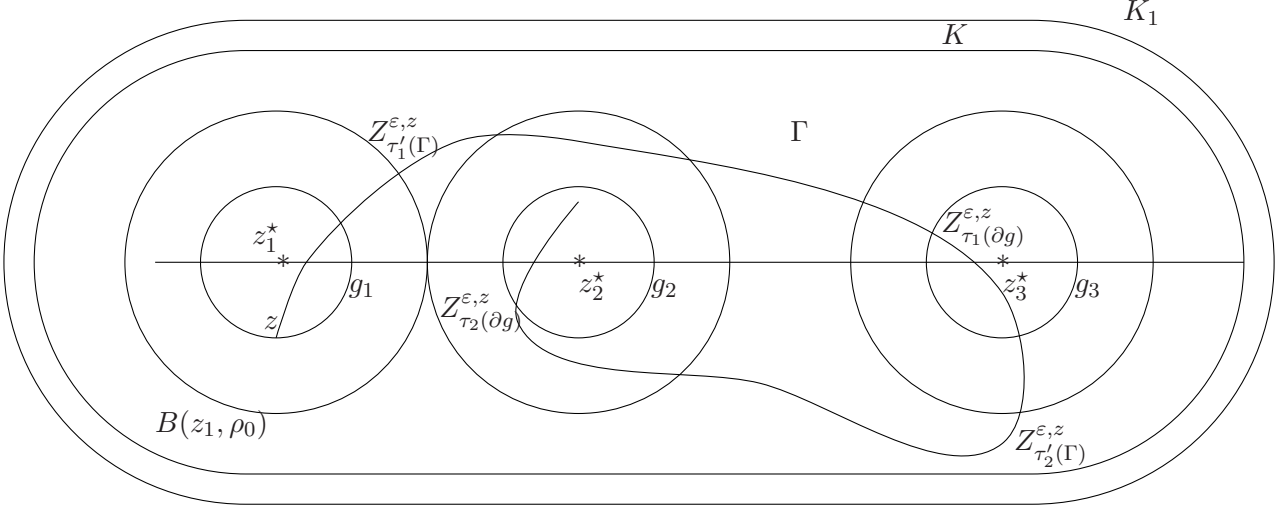


Figure 2.1 : Representation d'une trajectoire de  $(Z_t^{\epsilon, z})_{t \geq 0}$  et des temps d'arrêt  $\tau'_n(\Gamma)$  et  $\tau_n(\partial g)$ .

puis pour tout  $n \geq 2$ ,

$$\tau'_n(\Gamma) = \inf\{t > \tau_{n-1}(\partial g), Z_t^{\epsilon, z} \in \partial\Gamma\}, \quad \tau_n(\partial g) = \inf\{t > \tau'_n(\Gamma), Z_t^{\epsilon, z} \in \partial g\}.$$

Les notations ci-dessus sont illustrées par la figure 2.1 (les compacts  $K$  et  $K_1$  jouent un rôle technique sur lequel on n'insistera pas dans la suite).

On note  $(\tilde{Z}_n^\epsilon)_{n \geq 1}$  la chaîne squelette associée définie pour tout  $n \geq 1$  par  $\tilde{Z}_n^\epsilon = Z_{\tau_n(\partial g)}^{\epsilon, z}$ . Sans rentrer dans les détails, notons que cette chaîne est bien définie sous nos hypothèses (les temps d'arrêt sont finis *p.s.*) et qu'on a également la propriété :  $\sup_{z \in \partial g} \mathbb{E}_z^\epsilon[\tau_1(\partial g)] < \infty$ . La chaîne  $(\tilde{Z}_n^\epsilon)_{n \geq 1}$  admet une unique probabilité invariante notée  $\tilde{\mu}_\epsilon$ . À l'aide de la propriété de Markov forte, on peut alors obtenir la représentation suivante de la mesure invariante  $\nu_\epsilon$  (voir [53], Chapitre 4) :

**Lemme 2.31** *Supposons que les hypothèses de la proposition 2.28 soient satisfaites. Pour tous  $\epsilon > 0$  et  $\rho \in (0, \rho_0)$ ,  $\mu_\epsilon^{\partial g}$  définie pour tout  $A \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d)$  par :*

$$\mu_\epsilon^{\partial g}(A) = \int_{\partial g} \tilde{\mu}_\epsilon^{\partial g}(dz) \mathbb{E}_z \int_0^{\tau_1(\partial g)} \mathbf{1}_{Z_s^{\epsilon, z} \in A} ds \quad (2.31)$$

*est une mesure finie invariante pour  $(Z_t^\epsilon)_{t \geq 0}$ . En particulier  $\mu_\epsilon^{\partial g} = \lambda_\epsilon \nu_\epsilon$  où  $\lambda_\epsilon = \mu_\epsilon^{\partial g}(\mathbb{R}^d \times \mathbb{R}^d)$ .*

Si  $\rho_0$  est choisi suffisamment petit (indépendamment de  $\epsilon$  cependant), la probabilité partant de  $\partial g_i$  que  $\tau_1(\partial g)$  appartienne à  $\partial g_i$  est très grande devant les autres possibilités. Ceci implique en quelque sorte que  $\lambda_\epsilon$  joue un rôle "négligeable" dans le problème. Notons  $\rho_1$  un réel strictement positif tel que  $\rho < \rho_1 < \rho_0$ . En prenant en compte le caractère négligeable de  $\lambda_\epsilon$ , le lemme ci-dessus implique alors que les quantités  $\nu_\epsilon(B(z_i^*, \rho_1))$ ,  $i = 1, \dots, \ell$  sont asymptotiquement inter-connectées par les quantités  $\tilde{\mu}_\epsilon^{\partial g}(\partial g_i)$  et  $\mathbb{P}(\tau_1(\partial g) = \partial g_j | z \in \partial g_i)$ , *i.e.*

par la manière dont la chaîne squelette charge les voisinages de chaque point critique et par les probabilités d'aller d'un voisinage à un autre (voir [33] pour plus de détails). Il s'agit alors d'estimer ces probabilités de transition lorsque  $\varepsilon$  tend vers 0 :

**Proposition 2.32** *Pour tout  $\gamma > 0$ , il existe  $\rho$  et  $\rho_0$  tels que  $0 < \rho < \rho_0$  et tels que pour  $\varepsilon$  suffisamment petit,*

$$\forall (i, j) \in \{1 \dots \ell\}^2 \quad \forall z \in \partial g_i \quad e^{-\varepsilon^{-2}[I(z_i^*, z_j^*) + \gamma]} \leq \mathbb{P}(\tau_1(\partial g) = \partial g_j | z \in \partial g_i) \leq e^{-\varepsilon^{-2}[I(z_i^*, z_j^*) - \gamma]}.$$

où  $\forall z, z' \in \mathbb{R}^d \times \mathbb{R}^d$ ,  $I(z, z') = \inf_{T \geq 0} I_T(z, z')$ .

Comme on va le voir dans la suite, la fonction  $W$  peut alors être exprimée au travers des quantités,  $\{I(z_i^*, z_j^*), (i, j) \in \{1, \dots, \ell\}^2\}$  et des  $\{i\}$ -graphes associés à l'ensemble  $\{z_1^*, \dots, z_\ell^*\}$ . Rappelons-en la définition. Pour tout  $i \in \{1, \dots, \ell\}$ , on appelle  $\{i\}$ -graphe (associé à  $z_i^*$ ) un graphe orienté sur  $\{z_1^*, \dots, z_\ell^*\}$  ayant les propriétés suivantes :

- (i) Pour tout  $j \neq i$ ,  $z_j^*$  est l'origine d'exactlyement d'une arête (orientée) du graphe.
- (ii) Le graphe n'a pas de cycle.
- (iii) Pour tout  $j \neq i$ , il existe un (unique) chemin reliant  $z_j^*$  à  $z_i^*$ .

On note alors  $\mathcal{G}(i)$  l'ensemble des  $\{i\}$ -graphes associés à  $z_i^*$ . À l'aide des ingrédients introduits précédemment, on obtient alors le théorème suivant :

**Théorème 2.33 ([36])** *Supposons que  $(\mathbf{H}_D)$  et les hypothèses de la proposition 2.28 soient satisfaites. Pour toute sous-suite  $(\nu_{\varepsilon_n})_{n \geq 1}$  (LD)-convergente, la fonction de taux  $W$  associée satisfait :*

$$\forall i \in \{1 \dots \ell\} \quad W(z_i^*) = \mathcal{W}(z_i^*) - \min_{j \in \{1, \dots, \ell\}} \mathcal{W}(z_j^*) \quad (2.32)$$

où

$$\forall i \in \{1 \dots \ell\} \quad \mathcal{W}(z_i^*) := \min_{IG \in \mathcal{G}(i)} \sum_{(z_m^* \rightarrow z_n^*) \in IG} I(z_m^*, z_n^*).$$

Par conséquent,  $(\nu_\varepsilon)_{\varepsilon \in (0, 1]}$  satisfait un PGD de fonction de taux  $W$  caractérisée par (2.27) et (2.32) (et de vitesse  $\varepsilon^2$ ).

### 2.2.3.3 Concentration de $(\nu_\varepsilon)_{\varepsilon \in (0, 1]}$ sur les minima globaux ?

On aborde maintenant le problème initial posé dans cette section. La famille de mesures invariantes  $(\nu_\varepsilon)_{\varepsilon \in (0, 1]}$  se concentre-t-elle sur les zéros de  $b$  correspondant aux minima globaux lorsque  $\varepsilon \rightarrow 0$  ?

Notons d'abord que la concentration de  $(\nu_\varepsilon)_{\varepsilon \in (0, 1]}$  sur les zéros de  $b$  correspondant aux minima locaux ne pose pas de problème car ces derniers sont les seuls équilibres stables du système  $\dot{z} = b(z)$  (voir Proposition 4.1 de [36]). En revanche, pour répondre à la question posée ci-dessus, il faut (et il suffit de) prouver que  $W(z_i^*) = 0$  si et seulement si  $z_i^*$  est un minimum global pour  $U$ . Via (2.32), on est donc ramené à prouver que  $\text{Argmin} \mathcal{W}$  est exactement l'ensemble des minima globaux. La forme de  $\mathcal{W}$  (quasi-potentiel) nous suggère donc naturellement d'encadrer les quantités  $I(z_i^*, z_j^*)$ ,  $i \neq j$ .

On propose dans la suite une tentative de réponse dans le cas le plus simple :  $U$  potentiel (coercif) double-puits en dimension 1. On note  $z_1^*$  et  $z_2^*$  les deux minima locaux de  $U$  et  $z_3^*$

l'unique maximum local de  $U$ . On supposera également que  $U(z_1^*) < U(z_2^*)$ . Dans un premier temps, il est relativement simple de montrer que

$$I(z_3^*, z_1^*) = I(z_3^*, z_2^*) = 0 \quad (\text{cf Lemme 5.1 de [36]}).$$

Pour les autres quantités, l'étude se divise en deux parties consacrées à leur majoration et à leur minoration.

**Majoration du coût.** On veut majorer  $I(z_2^*, z_3^*)$ . Le but est de construire une trajectoire contrôlée  $z_\varphi(z_2^*, \cdot)$  permettant de relier  $z_2^*$  à  $z_3^*$ . (et générant si possible un coût proche de l'optimal). Notons que sous l'hypothèse  $(\mathbf{H}_D)$ , la dynamique du système est non dégénérée aux points d'équilibre (plus précisément, une condition d'hypoellipticité forte (dite de Kalman) est satisfaite) ce qui implique que le système dynamique est *localement exactement contrôlable* (à coût faible) au voisinage des  $z_j^*$ ,  $j = 1, \dots, 3$  (voir Lemme 4.3 de [36]). Ceci signifie que deux points d'un voisinage assez proche de  $z_j^*$  peuvent être (exactement) reliés via une trajectoire contrôlée (en un temps fini) et que le coût tend vers 0 lorsque le diamètre du voisinage tend également vers 0. Ceci nous permet de raccorder seulement les voisinages des points d'équilibres.

Dans le cadre sans mémoire, *i.e* lorsque  $(\xi_t)_{t \geq 0}$  est solution de  $d\xi_t = -\nabla U(\xi_t)dt + dB_t$ , un bon choix est de considérer la trajectoire contrôlée  $x_\varphi(\cdot)$  telle que  $\dot{\varphi}(t) = 2U'(x_\varphi(t))$ . En effet, supposons par exemple que  $x_2^* < x_3^*$  de sorte que  $U'(x) > 0$  sur  $(x_2^*, x_3^*)$ . Pour tous  $x_2^* < x_2^\varepsilon < x_3^\varepsilon < x_3^*$ , il existe  $T_\varepsilon$  tel que  $x_\varphi(x_2^\varepsilon, T_\varepsilon) = x_3^\varepsilon$ . De plus,

$$I(x_2^\varepsilon, x_3^\varepsilon) \leq \frac{1}{2} \int_0^{+\infty} (2U'(x_\varphi(s)))^2 ds = \frac{1}{2} \int_0^{+\infty} 4\dot{x}_\varphi(s)U'(x_\varphi(s)) ds \leq 2(U(x_3^*) - U(x_2^*)).$$

Au vu de la forme de la fonction de taux dans ce cadre (rappelée dans l'introduction), il s'agit d'une majoration optimale.

Dans notre cadre où le Brownien n'agit que sur la composante en  $x$ , on considère la trajectoire contrôlée  $z_\varphi$  telle que  $\dot{\varphi}(t) = 2y_\varphi(t)$ . On est ainsi ramené à l'étude de l'o.d.e.

$$\begin{cases} \dot{x}(t) = y(t) \\ \dot{y}(t) = \lambda(U'(x(t)) - y(t)) \end{cases}$$

pour laquelle on réussit à montrer la convergence vers  $z_3^*$  (partant de  $z_2^\varepsilon = (x_2^\varepsilon, 0)$ ). Ceci nous permet de retrouver la majoration du cas standard :

$$I(z_2^*, z_3^*) \leq 2(U(x_3^*) - U(x_2^*)).$$

La partie minoration est beaucoup plus délicate :

**Minoration du coût.** À nouveau, prenons en référence l'équation  $d\xi_t = -\nabla U(\xi_t)dt + dB_t$  (en dimension  $d$  cette fois). Dans ce cadre on peut faire usage de la minoration suivante :

$$|\dot{\varphi}(t)|^2 = |\dot{x}(t) + \nabla U(x(t))|^2 \leq 2(\dot{x}(t)|\nabla U(x(t)))$$

Par construction, celle-ci nous permet de minorer le coût de toute trajectoire contrôlée allant de  $x$  à  $x'$  par  $2(U(x') - U(x))$  et donc d'obtenir la "bonne" minoration. Si on veut adapter cette approche à notre cadre, on doit faire face à deux difficultés : on a un défaut d'ellipticité et surtout, le champ de vecteurs  $b$  n'est pas un gradient.

Une première idée consiste à forcer la propriété précédente à fonctionner : d'abord, on peut toujours se ramener au cadre elliptique car le coût dans ce cadre est forcément plus faible que dans le cas où le Brownien n'agit que sur la première composante (on a "plus" de trajectoires contrôlées dans le premier cas). Puis, on peut surmonter la seconde carence en projetant le champ de vecteurs  $b$  sur le gradient d'une fonction de Lyapounov bien choisie. Nous ne détaillerons pas plus cette idée (voir Proposition 5.1 de [36] pour plus de détails) et lui préférerons l'approche ci-dessous, qui n'est pas entièrement déconnectée à la première et qui génère de meilleurs résultats.

Dans notre cadre,

$$|\dot{\varphi}(t)|^2 = |\dot{x}_\varphi(t) + y_\varphi(t)|^2.$$

L'idée est alors de chercher une fonction  $\mathcal{L} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  telle que pour tous  $t$  et  $\varphi$ ,

$$|\dot{x}_\varphi(t) + y_\varphi(t)|^2 \geq 2(\nabla \mathcal{L}(x_\varphi(t), y_\varphi(t)) | (\dot{x}_\varphi(t), \dot{y}_\varphi(t))').$$

Comme  $\dot{y} = \lambda(\nabla U(x) - y)$ , on peut en posant  $u = \dot{x}$ , reformuler ce problème de la manière suivante : déterminer une fonction  $\mathcal{L} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  telle que

$$\forall (x, y, u) \in \mathbb{R}^3, \quad |u + y|^2 \geq 2(\nabla \mathcal{L}(x, y) | (u, y)'). \quad (2.33)$$

On obtiendrait ainsi pour tous  $z, z' \in \mathbb{R}^2$ ,

$$I(z, z') \geq 2(\mathcal{L}(z') - \mathcal{L}(z)).$$

Le but étant de maximiser le terme de droite, on cherche un choix optimal de  $\mathcal{L}$  dans une classe de fonctions dépendant de trois paramètres :

$$\mathcal{L}_{\alpha, \beta, \gamma}(z) = \mathcal{L}_{\alpha, \beta, \gamma}(x, y) := \alpha U(x) + \beta y^2 / 2 - \gamma y U'(x).$$

On obtient alors le résultat suivant :

**Proposition 2.34 ([36])** *Pour tout  $\alpha \in [0; 2[$ , il existe des constantes explicites  $m_\lambda(\alpha), \beta^*(\alpha), \gamma^*(\alpha)$  telles que (2.33) est vraie pour  $\beta = \beta^*(\alpha), \gamma = \gamma^*(\alpha)$  et pour tout potentiel double-puits  $U$  satisfaisant  $\|U''\|_\infty = M < m_\lambda(\alpha)$ . En particulier, on a pour tous  $j = 2, 3$ ,*

$$I(z_1^*, z_j^*) \geq \alpha \left[ U(x_3^*) - U(x_1^*) \right],$$

d'où l'on déduit le théorème (qui donne la condition voulue mais sous condition sur  $\|U''\|_\infty$ ).

**Théorème 2.35 ([36])** *Supposons  $U : \mathbb{R} \rightarrow \mathbb{R}$  soit un potentiel double puits tel que les hypothèses du théorème 2.33 soient satisfaites. Supposons également que  $U(z_1^*) < U(z_2^*)$ . Si  $\|U''\|_\infty < m_\lambda(\alpha)$  avec  $\alpha = 2 \frac{U(x_3^*) - U(x_2^*)}{U(x_3^*) - U(x_1^*)}$ , alors*

$$\mathcal{W}(z_1^*) < \mathcal{W}(z_2^*).$$

*En particulier,  $(\nu_\varepsilon)_{\varepsilon \in (0, 1]}$  converge en probabilité vers  $z_1^*$ .*

**Remarque 2.36**  $\triangleright$  Les arguments utilisés pourraient se généraliser au cadre multi-puits en dimension 1 (sous réserve de difficultés techniques). En revanche, le passage à la dimension supérieure semble lui, beaucoup plus délicat.

$\triangleright$  On peut regretter dans l'approche ci-dessus un côté "construction à la main" qui n'utilise pas pleinement les outils du contrôle optimal. On peut penser par exemple au principe du maximum de Pontryagin (principe d'optimisation sous contrainte). Malheureusement, ce type d'approche est pour l'instant resté infructueux.

$\triangleright$  Comme on peut le voir, le résultat dépend à la fois des hauteurs  $U(x_3^*) - U(x_i^*)$   $i = 1, 2$  et de  $\|U''\|_\infty$ . On propose dans la figure ci-dessous de représenter le rapport maximal entre les hauteurs sous lequel on a convergence vers le minimum global, en fonction de  $\|U''\|_\infty$ . L'approche dite "elliptique" évoquée plus haut (mais donnant de moins bons résultats) est aussi représentée. On constate que la contrainte sur  $\|U''\|_\infty$  décroît avec  $\lambda$ . Par exemple, si  $\alpha = 1$ , *i.e.* si  $U(x_3^*) - U(x_1^*) = 2(U(x_3^*) - U(x_2^*))$ , la valeur maximale de  $\|U''\|_\infty$  (pour assurer la concentration de  $(\nu_\varepsilon)$  sur  $x_1^*$ ) est de l'ordre de 8 lorsque  $\lambda = 10$  et de 0.7 lorsque  $\lambda = 1$ .

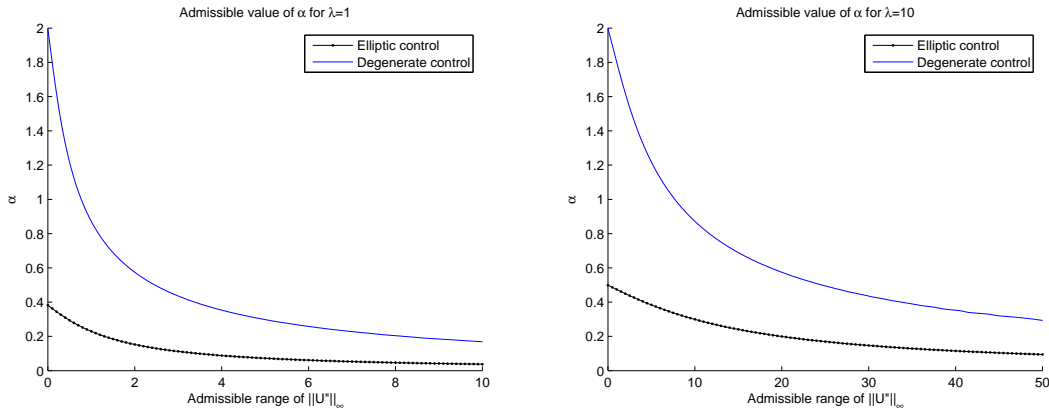


Figure 2.2 : Valeur de  $\alpha$  maximale en fonction de  $\|U''\|_\infty$  lorsque  $\lambda = 1$  et  $\lambda = 10$ .

## 2.3 Un modèle de bulle spéculative

### 2.3.1 Introduction

Il s'agit du dernier "volet" concernant des problèmes modélisés par des EDS hypoelliptiques. Néanmoins, cette section n'a aucun lien avec les précédentes.

Le but de ce travail <sup>†</sup> est de proposer un modèle simple de dynamique soumise au phénomène de bulle spéculative et d'en étudier certaines propriétés. Le modèle proposé est une diffusion à mémoire qui, après quelques transformations peut être vue comme un processus gaussien 2-dimensionnel tournant (par tournant, on sous-entend que la matrice  $2 \times 2$  relative au système déterministe associé au modèle a des valeurs propres complexes). Ce dernier est un processus markovien du second ordre non réversible et on montrera qu'il peut être construit comme limite lorsque  $N \rightarrow +\infty$  du comportement moyen de  $N$  investisseurs (de type spéculateur) en interaction.

<sup>†</sup>. "Motivé" initialement par une discussion sur les prix de l'immobilier Toulousain !

Nous nous consacrons ensuite à une étude fine du temps de retour à un prix donné et en particulier à la *persistance* associée : si l'on suppose pour simplifier que le temps de retour  $\tau$  a une loi exponentielle (ce qui n'est en général pas le cas dans notre cadre), alors la persistance correspond en substance au paramètre de cette loi (voir Section 2.3.5 pour une définition précise). Il s'agit donc d'un paramètre de vitesse de retour à un prix donné. Dans le résultat principal, on obtiendra à l'aide de méthodes spectrales ou probabilistes des bornes explicites pour cette quantité et on prouvera en particulier que la persistance est essentiellement proportionnelle à la fréquence de rotation  $\omega$  associée au modèle.

**Généralités.** On rappelle dans un premier temps que l'on parle communément de bulle financière ou spéculative, lorsqu'un prix s'éloigne (positivement) de sa valeur "réelle" ou *fondamentale* (même si celle-ci n'est pas clairement définie en général) à cause d'un phénomène de spéculation. Ce dernier peut être défini grossièrement de la manière suivante : les opérations de ventes ou achats sont réalisées principalement par tentative d'anticipation des prix futurs (et donc pas sur la seule base de la valeur "intrinsèque" de l'actif). On rappelle également que ce type de phénomène est assez récurrent, allant de la "tulipomanie" ((1634-1637) aux plus récentes bulle Internet (éclatée en 2000) ou immobilière aux Etats-Unis (2007). On renvoie à [37] pour un historique plus exhaustif.

**Modélisation.** Les tentatives de modélisation mathématique du phénomène de bulle spéculative sont assez nombreuses. On trouvera un certain nombre de références dans [91] ou dans les articles récents de Protter [86] et Kiselev [54].

La dynamique proposée ici s'inspire de l'article de Shiller [92] dans lequel le phénomène de bulle est décrit de la manière suivante : *"If asset prices start to rise strongly, the success of some investors attracts public attention that fuels the spread of the enthusiasm for the market : (often, less sophisticated) investors enter the market and bid up prices. This "irrational exuberance" heightens expectations of further price increases, as investors extrapolate recent price action far into the future. The markets meteoric rise is typically justified in the popular culture by some superficially plausible "new era" theory that validates the abandonment of traditional valuation metrics. But the bubble carries the seeds of its own destruction ; if prices begin to sag, pessimism can take hold, causing some investors to exit the market. Downward price motion begets expectations of further downward motion, and so on, until the bottom is eventually reached."*

De manière simplifiée, ce paragraphe peut être réinterprété comme suit : si les prix montent (resp. baissent), alors, les investisseurs anticipent une montée (resp. une baisse), ce qui amplifie chaque tendance. Cette citation suggère également que le phénomène de bulle spéculative a une dynamique du second ordre. Cela nous conduit à considérer le processus  $(X_t)_{t \geq 0}$  solution de l'EDS unidimensionnelle suivante :

$$\forall t \geq 0, \quad dX_t = -aX_t dt + \left( b \int_0^t \exp(b(s-t)) dX_s \right) dt + cdB_t \quad (2.34)$$

où  $a$ ,  $b$  et  $c$  sont des réels strictement positifs. Le premier terme est un terme de retour à l'équilibre (nécessaire pour éviter l'explosion des prix), le second correspond à la modélisation de la dynamique spéculative (moyenne pondérée des variations passées) tandis que le dernier contient l'aléa du problème. Afin d'assurer la prépondérance du phénomène de spéculation, on fera le plus souvent l'hypothèse que  $a \ll b$ . Comme on peut le voir, le terme associé à la spéculation est une moyenne pondérée sur le passé proche (la notion de

proximité dépendant de  $b$ ) des variations du prix. Notons également que  $(X_t)_{t \geq 0}$  pouvant prendre des valeurs négatives, il est plutôt à comprendre comme un processus de prix relatif (au prix intrinsèque).

Le modèle (2.34) peut être compris à partir d'une description "microscopique" du problème dont on rappelle ici rapidement l'heuristique. On considère  $N$  investisseurs (de type spéculateur) et on suppose que chacun a sa propre idée de l'évolution des prix. On note  $X(n) := (X_t(n))_{t \geq 0}$ , le prix fixé par chaque individu,  $n = 1, \dots, N$  et on note  $(\bar{X}_t^N)_{t \geq 0}$  le prix moyen

$$\forall t \geq 0, \quad \bar{X}_t^N := \frac{1}{N} \sum_{n \in \{1, \dots, N\}} X_t(n).$$

On fait l'hypothèse qu'à chaque instant  $t$ , chaque individu a accès à tout le passé de  $\bar{X}$  avant  $t$  et qu'il l'utilise en calculant le ratio  $(\bar{X}_t - \bar{X}_{t-\Upsilon(n)})/\Upsilon(n)$  pour "anticiper" le prix futur via la "règle" de spéculation rappelée ci-dessus.  $\Upsilon(n)$  est une variable aléatoire à valeurs dans  $(0, +\infty)$  qui représente la taille de la fenêtre du passé choisie par l'investisseur. Afin que la quantité précédente soit bien définie, on fera l'hypothèse que  $X_t(n) = 0$  sur  $\mathbb{R}_-$  pour tout  $n \in \{1, \dots, N\}$ . Avec ces notations, on suppose alors que pour tout  $n \in \{1, \dots, N\}$ ,

$$\forall t \geq 0, \quad dX_t(n) = -aX_t(n)dt + \frac{\bar{X}_t^N - \bar{X}_{t-\Upsilon(n)}^N}{\Upsilon(n)}dt + c\sqrt{N}dB_t(n)$$

où  $(B(n))_{n \in \{1, \dots, N\}}$  est une suite de browniens standards supposés indépendants. Le facteur  $\sqrt{N}$  (nécessaire pour justifier le passage à la limite) peut s'interpréter comme une amplification du bruit lorsque la taille de la population augmente. On en déduit donc que

$$\forall t \geq 0, \quad d\bar{X}_t^N = -a\bar{X}_t^N dt + \left( \frac{1}{N} \sum_{n \in \{1, \dots, N\}} \frac{\bar{X}_t^N - \bar{X}_{t-\Upsilon(n)}^N}{\Upsilon(n)} \right) dt + c \frac{1}{\sqrt{N}} \sum_{n \in \{1, \dots, N\}} dB_t(n). \quad (2.35)$$

Notons que  $\frac{1}{\sqrt{N}} \sum_{n \in \{1, \dots, N\}} B(n)$  est un Brownien standard. Si l'on suppose de plus que la suite  $(\Upsilon(n))_{n \in \mathbb{N}}$  est *i.i.d.* de même loi qu'une variable aléatoire notée  $\Upsilon$  et que  $\Upsilon$  a une densité  $p_\Upsilon$  (par rapport à la mesure de Lebesgue), on peut alors montrer, sous des hypothèses d'intégrabilité sur  $\Upsilon^{-1}$  que  $(\bar{X}^N)_N$  converge étroitement (pour la topologie de la convergence uniforme sur les compacts) vers  $\bar{X}$  solution de

$$\forall t \geq 0, \quad d\bar{X}_t = -a\bar{X}_t dt + \int_0^{+\infty} \frac{\bar{X}_t - \bar{X}_{t-u}}{u} p_\Upsilon(u) du + c dB_t \quad (2.36)$$

où  $(B_t)_{t \geq 0}$  est un mouvement Brownien standard. Bien qu'une justification plus précise ne soit pas envisagée ici, remarquons cependant que la convergence du terme de "spéculation" peut être formellement comprise via la loi des grands nombres appliquée à la suite  $(\Upsilon(n))_{n \in \mathbb{N}}$ . D'après la forme de l'équation ci-dessus, on en déduit donc que l'équation (2.34) peut être vue comme limite de ce type de système de particules si,

$$p.s., \quad \forall t \geq 0, \quad \int_0^{+\infty} \frac{\bar{X}_t - \bar{X}_{t-u}}{u} p_\Upsilon(u) du = b \int_0^t \exp(b(s-t)) dX_s. \quad (2.37)$$

On a alors la propriété suivante :



**Lemme 2.37** *Supposons que  $X = (X_t)_{t \in \mathbb{R}}$  soit une semi-martingale continue telle que  $X_t = 0$  pour  $t \leq 0$ . Alors, (2.37) est vraie si  $\Upsilon$  a une loi  $\Gamma_{2,b}$ , i.e. si,*

$$\forall t \geq 0, \quad \mathbb{P}_\Upsilon(dt) = b^2 t \exp(-bt) dt.$$

**Remarque 2.38** L'idée de la preuve est essentiellement une intégration par parties (Celle-ci explique aussi l'apparition d'une loi Gamma). La réciproque est vraie sous l'hypothèse  $\int t^{-1} \mathbb{P}_\Upsilon(dt) < +\infty$  (hypothèse assez naturelle puisqu'elle joue aussi un rôle pour assurer la convergence de  $(\bar{X}^N)_N$ ).

**Temps de retour à la valeur d'équilibre.** Une quantité (aléatoire) d'intérêt dans ce problème est le temps nécessaire au processus pour revenir à une valeur donnée. Pour simplifier, en supposant que  $X_0 = x_0 > 0$ , on s'intéresse simplement ici à

$$\tau := \inf\{t > 0, X_t = 0\}. \quad (2.38)$$

Dans notre problème, il s'agit de la valeur moyenne à l'équilibre du processus (voir (2.39) pour s'en convaincre).

L'objectif principal de la suite est d'obtenir des bornes quantitatives (asymptotiques et non asymptotiques) pour  $\mathbb{P}(\tau \geq t)$ . Ce travail est divisé en plusieurs parties. On commencera par effectuer une simplification du problème, puis on s'intéressera successivement à la majoration et à la minoration de  $\mathbb{P}(\tau > t)$ .

### 2.3.2 Reformulations du modèle

Le processus solution de (2.34) est markovien d'ordre 2. Plus précisément, à l'image des diffusions à gradient moyenné,  $(X_t)_{t \geq 0}$  peut être vu comme la première composante d'un processus markovien  $(X_t, Y_t)_{t \geq 0}$ . En posant

$$Y_t = b \int_0^t \exp(b(s-t)) dX_s - bX_t$$

on montre facilement que  $(Z_t)_{t \geq 0} := ((X_t, Y_t)^*)_{t \geq 0}$  est solution de

$$\forall t \geq 0, \quad dZ_t = AZ_t dt + C dB_t \quad (2.39)$$

où  $(B_t)_{t \geq 0}$  est encore un brownien unidimensionnel,

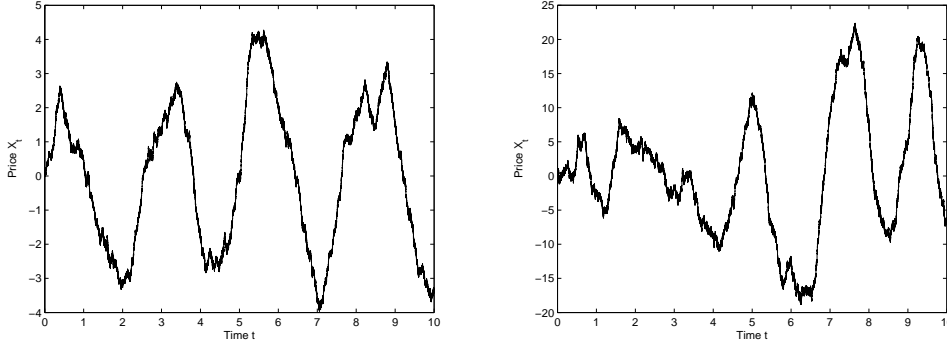
$$A := \begin{pmatrix} b-a & 1 \\ -b^2 & -b \end{pmatrix} \quad \text{et} \quad C := \begin{pmatrix} c \\ 0 \end{pmatrix}. \quad (2.40)$$

Le processus  $(Z_t)_{t \geq 0}$  est markovien et gaussien. Il est également hypoelliptique et converge lorsque  $t \rightarrow +\infty$  vers son unique mesure invariante  $\mu = \mathcal{N}(0_{\mathbb{R}^2}, \Sigma)$  où  $\Sigma$  est une matrice  $2 \times 2$  non dégénérée (voir [34] pour une forme explicite).

Si les résultats ci-dessus sont valables pour tous  $a, b, c$  strictement positifs, la dynamique du processus est néanmoins fortement dépendante des paramètres. En particulier, on remarque que :

- Si  $a > 4b$ ,  $A$  admet deux valeurs propres réelles,  $\lambda_{\pm} := (-a \pm \sqrt{a^2 - 4ab})/2$ .
- Si  $a = 4b$ ,  $A$  est semblable à la matrice de Jordan  $2 \times 2$  associée à la valeur propre  $-a/2$ .
- Si  $a < 4b$ ,  $A$  admet deux valeurs propres complexes conjuguées,  $\lambda_{\pm} := (-a \pm i\sqrt{4ab - a^2})/2$ .

Lorsque  $a \geq 4b$ , la dynamique du processus de  $(X_t)_{t \geq 0}$  est proche d'un Ornstein-Uhlenbeck. Le phénomène de bulle n'apparaît que lorsque  $a < 4b$ , *i.e.* lorsque la composante de spéculation devient "prépondérante" devant le terme de rappel à l'équilibre (voir figure 2.3 pour quelques exemples de trajectoires). On supposera donc toujours dans la suite que  $a < 4b$ .



**Figure 2.3** : Simulations de trajectoires de  $(X_t)_{t \geq 0}$  (gauche :  $a = 1, b = 5, c = 1$ , droite :  $a = 1, b = 10, c = 5$ ).

Notons

$$\omega = \frac{\sqrt{4ab - a^2}}{2} \quad \text{et} \quad \rho = \frac{a}{2}.$$

$\omega$  et  $\rho$  correspondent respectivement à la fréquence de rotation et au coefficient de contraction du système déterministe associé. Dans le lemme suivant, on propose via des changements de base et d'échelle appropriés de ramener l'étude à un modèle plus simple :

**Proposition 2.39** ([34]) *Soit  $(Z_t)_{t \geq 0}$  processus solution de (2.39). Supposons que  $a < 4b$ . Considérons  $v \in \mathbb{R}^2 \setminus \{0\}$  et notons  $P_v = (v, Bv)$  où  $B = \frac{1}{\omega}(A + \rho I_2)$ . Alors, pour tout  $\alpha \in \mathbb{R} \setminus \{0\}$ ,  $(\hat{Z}_t)_{t \geq 0}$  défini par  $\hat{Z}_t = \sqrt{\omega} \alpha P_v^{-1} Z_t$  est solution de*

$$d\hat{Z}_t = -\frac{a}{2\omega} \hat{Z}_t + J_2 \hat{Z}_t + \alpha c P_v^{-1} \Sigma dW_t \quad \text{où} \quad \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad (2.41)$$

et  $W$  est un mouvement Brownien standard 2-dimensionnel. En particulier, si  $v = (\frac{1}{b^2}(\frac{a}{2} - b), 1)^*$  et  $\alpha = \frac{\sqrt{2}\omega}{cb^2}$ , alors  $(\hat{Z}_t)_{t \geq 0} := ((U_t, V_t)^*)_{t \geq 0}$  est solution de

$$\begin{cases} dU_t = -\frac{a}{2\omega} U_t - V_t dt \\ dV_t = -\frac{a}{2\omega} V_t + U_t + \sqrt{2} dW_t. \end{cases} \quad (2.42)$$

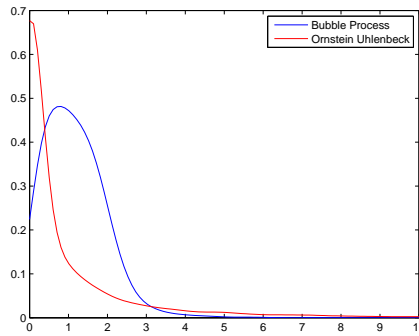
où  $W$  est maintenant un Brownien standard unidimensionnel. Par conséquent, en notant  $D = \{(x, y), x > 0\}$ , on a

$$\forall t \geq 0, \quad \mathbb{P}(\tau > t) = \mathbb{P}(\forall s \in [0, t], P_v \hat{Z}_{\omega s} \in D) = \mathbb{P}(\hat{\tau}_v > \omega t)$$

où  $\hat{\tau} := \inf\{t > 0, \hat{Z}_t \in P_v^{-1}D\}$ .

**Remarque 2.40**  $\triangleright$  Le dernier résultat montre que, sous réserve d'une rotation du domaine, on est ramené à étudier le problème pour (2.42). Les assertions suggèrent également une propriété de proportionalité vis-à-vis de la fréquence de rotation et une indépendance vis-à-vis du paramètre de variance  $c$  (la composante Brownienne dans (2.42) est normalisée). Ces remarques seront confirmées dans les résultats principaux. On peut pour l'instant illustrer ces propriétés via la figure 2.4 qui montre que la loi de  $\tau$  se démarque très nettement de celle relative au temps d'atteinte de 0 lorsque le processus est un Ornstein-Uhlenbeck.

$\triangleright$  La simplification obtenue ci-dessus jouera principalement un rôle dans la partie minoration dans laquelle les outils d'étude spectrale sont techniquement difficiles à porter dans un cadre général.



**Figure 2.4** : Approximation des densités de  $\tau$  relatives à  $(X_t)_{t \geq 0}$  et à un Ornstein-Uhlenbeck (avec même terme Brownien et paramètre de drift choisi pour que les mesures invariantes coïncident).

### 2.3.3 Majoration de $\mathbb{P}(\tau > t)$

Dans ce cadre tournant, l'idée de la majoration est assez simple. En effet, la période du système déterministe étant égale à  $2\pi/\omega$ , on en déduit facilement que

$$\forall x > 0, \quad \mathbb{E}[X_{\frac{\pi}{\omega}}] < 0 \quad \implies \quad \mathbb{P}(X_{\frac{\pi}{\omega}} \geq \frac{1}{2})$$

car  $X_{\frac{\pi}{\omega}}$  est une variable aléatoire gaussienne donc symétrique. Ainsi,

$$\sup_{(x,y) \in D} \mathbb{P}_{(x,y)}(\tau > \frac{\pi}{\omega}) \leq \frac{1}{2}.$$

Un raisonnement itératif s'appuyant sur la propriété de Markov forte permet alors d'en déduire :

**Théorème 2.41** ([34]) *Soit  $(Z_t)_{t \geq 0}$  solution de (2.39) avec  $a < 4b$ . Alors, pour tout  $(x_0, y_0) \in \mathbb{R}^2$  tel que  $x_0 > 0$ ,*

$$\mathbb{P}_{(x_0, y_0)}(\tau > t) \leq 2 \exp\left(-\frac{\log 2}{\pi} \omega t\right).$$

**Remarque 2.42** ▷ Dans l'inégalité ci-dessus, les points importants sont l'indépendance de la majoration vis-à-vis du point de départ et de la variance ainsi que la proportionnalité vis-à-vis de la fréquence de rotation (déjà mentionnée précédemment).

▷ Le résultat peut s'étendre au temps de sortie de secteur angulaires

$$\mathcal{S}_{\alpha_1, \alpha_2} = \{(x, y) \in D, \alpha_1 x < y < \alpha_2 x\} \quad (\text{avec } \alpha_1 < \alpha_2).$$

Au vu de l'esquisse de preuve ci-dessus, on pourrait s'attendre à remplacer  $\pi$  par  $\theta_2 - \theta_1$  dans l'inégalité. Cependant, la vitesse de rotation du système déterministe n'étant pas constante (à cause de la contraction),  $\pi$  est en fait remplacé par :  $\text{Arctan}(\frac{1}{\omega}(\frac{a}{2} - b - \alpha_2)) - \text{Arctan}(\frac{1}{\omega}(\frac{a}{2} - b - \alpha_1))$ .

### 2.3.4 Minoration de $\mathbb{P}(\tau > t)$

La minoration pose plus de problèmes. Pour cette partie, nous proposons deux approches : spectrale et probabiliste.

#### 2.3.4.1 Approche spectrale

Cette approche du problème est basée sur la proposition suivante :

**Proposition 2.43** Soit  $(Z_t)_{t \geq 0}$  un processus de Markov à valeurs dans  $\mathbb{R}^d$  de générateur infinitésimal  $\mathcal{L}$  et de distribution initiale  $m_0$ . Soit  $\mathcal{S}$  un domaine de  $\mathbb{R}^d$ . Supposons que  $m_0(\mathcal{S}) = 1$  et notons  $\tau := \inf\{t > 0, Z_t \in \mathcal{S}^c\}$ . Alors, s'il existe une fonction bornée  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  appartenant au domaine de  $\mathcal{L}$  et  $\lambda \in \mathbb{R}$  tels que

$$\begin{cases} f|_{\partial\mathcal{S}} = 0, & f|_{\mathcal{S}} > 0 \quad \text{et,} \\ \forall x \in \mathcal{S}, & \mathcal{L}f(x) \geq -\lambda f(x) \end{cases} \quad (2.43)$$

alors,  $\mathbb{E}_{m_0}[e^{\lambda\tau}] = +\infty$ . Par conséquent,

$$\limsup_{t \rightarrow +\infty} -\frac{1}{t} \log(\mathbb{P}_{m_0}(\tau \geq t)) \leq \lambda.$$

Vu sous cet angle, le problème revient donc à construire une solution de (2.43). D'après la proposition (2.39), on peut se ramener à l'étude de (2.42). Pour simplifier un peu plus l'étude, on considère en fait dans un premier temps la version elliptique (2-dimensionnelle) de (2.42)

$$d\xi_t = -\rho\xi_t + J_2\xi_t + dW_t, \quad (2.44)$$

où  $(W_t)_{t \geq 0}$  est un Brownien standard 2-dimensionnel. La dynamique déterministe nous suggère alors de considérer le générateur infinitésimal  $\mathcal{L}_e$  en coordonnées polaires. On a :

$$\mathcal{L}_e = -\rho r \partial_r + \partial_\theta + \partial_r^2 + \frac{1}{r} \partial_r + \frac{1}{r^2} \partial_\theta^2. \quad (2.45)$$

D'après la proposition ci-dessus, l'idée naturelle est de tenter de résoudre  $\mathcal{L}_e f = -\lambda f$  (avec les conditions sur  $f$  de (2.43)). Malheureusement, ce problème ne semblant pas soluble

explicitement, on raisonne dans un premier temps en “oubliant” les dérivées en  $r$ . Cela nous mène alors à l'équation du second ordre (à  $r$  fixé strictement positif)

$$r^{-2}G_r''(\theta) + G_r'(\theta) = -\lambda_r G_r(\theta), \quad -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}.$$

Cette équation admet une solution satisfaisant  $G_r(\theta) > 0 \forall \theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$   $G_r(\pi/2) = G_r(-\pi/2) = 0$  si et seulement si  $\lambda_r = \frac{1}{r^2} + \frac{r^2}{4}$ . Dans ce cas, l'unique solution est donnée par

$$\forall \theta \in [-\pi/2, \pi/2], \quad G_r(\theta) := e^{-\frac{r^2}{2}\theta} \cos \theta.$$

Lorsque que l'on “réinjecte” les dérivées en  $r$  dans le problème, on choisit de chercher  $g$  solution de  $\mathcal{L}_e g \geq -\lambda g$  sous la forme

$$g(r, \theta) = r e^{\beta(\theta)r^2} \cos(\theta), \quad \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \quad r \geq 0 \quad (2.46)$$

où  $\beta$  est négative sur  $[-\pi/2, \pi/2]$ . Notons que  $g$  est donc bien une fonction bornée nulle au bord et strictement positive à l'intérieur. Sous cette forme,  $\mathcal{L}_e g$  prend une forme relativement agréable :

$$\forall (r, \theta) \in \mathbb{R}_+^* \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \quad \mathcal{L}_e g(r, \theta) = [\psi_1(\theta)r^2 + \psi_2(\theta)] g(r, \theta)$$

où

$$\begin{aligned} \psi_1(\theta) &= -2\rho\beta(\theta) + \beta'(\theta) + (4\beta^2(\theta) + (\beta'(\theta))^2) \\ \psi_2(\theta) &= -\rho + 8\beta(\theta) - (1 + 2\beta'(\theta)) \tan \theta + \beta''(\theta). \end{aligned}$$

En particulier, cette formulation nous montre que si l'on veut minorer le ratio  $\frac{\mathcal{L}_e g}{g}$  (uniformément en  $(r, \theta)$ ), il nous faut construire  $\beta$  sur  $(-\pi/2, \pi/2)$  telle que  $\psi_1$  est positive sur  $(-\pi/2, \pi/2)$  et telle que

$$\limsup_{\theta \rightarrow \frac{\pi}{2}} 1 + 2\beta'(\theta) \leq 0 \quad \text{et} \quad \liminf_{\theta \rightarrow -\frac{\pi}{2}} 1 + 2\beta'(\theta) \geq 0, \quad (2.47)$$

afin que  $\psi_2$  soit minorée sur  $(-\pi/2, \pi/2)$ . Notre meilleure solution (*i.e* donnant le  $\lambda$  minimal) est alors donnée par la proposition suivante :

**Proposition 2.44 ([34])** (i) Soit  $\rho \geq 0$  et  $g$  donnée par (2.46) avec

$$\beta(\theta) = \begin{cases} \frac{1}{4}(1 - \sqrt{3}) & \text{if } \theta \in [-\frac{\pi}{2}, \frac{\pi}{4}] \\ \frac{1}{4}(\sin(2\theta) - \sqrt{3}) & \text{if } \theta \in [\frac{\pi}{4}, \frac{\pi}{2}]. \end{cases} \quad (2.48)$$

Alors, pour tout  $(r, \theta) \in \mathbb{R}_+^* \times [-\pi/2, \pi/2] \setminus \{\pi/4\}$ ,

$$\mathcal{L}_e g(r, \theta) \geq -\lambda_\rho g(r, \theta) \quad \text{où } \lambda_\rho = 2\sqrt{3} + \rho.$$

(ii) Soit  $\rho \geq 0$  et  $(\xi_t)_{t \geq 0}$  solution de (2.44). Alors, pour tout  $(x_0, y_0) \in \mathbb{R}^2$  tel que  $x_0 > 0$ ,

$$\limsup_{t \rightarrow +\infty} -\frac{1}{t} \log(\mathbb{P}_{(x_0, y_0)}(\tau \geq t)) \leq 2\sqrt{3} + \rho$$

où  $\tau := \inf\{t \geq 0, (\xi_t)_1 < 0\}$ .

Bien que l'approche elliptique ne fournisse pas d'outils concrets pour répondre à la question dans le cas hypoelliptique, elle permet néanmoins de donner une idée de la forme d'une "bonne" fonction et des difficultés du problème. Sur ce dernier point, on remarque notamment la condition (2.47) qui est particulièrement contraignante au voisinage de  $\pi/2$ . En effet, la dynamique déterministe "incite" le processus à sortir et il est alors beaucoup plus coûteux pour ce dernier de rester dans le domaine.

Si l'on considère maintenant le système hypoelliptique (2.42), on remarque que l'absence de Brownien sur la première composante amplifie la difficulté soulevée ci-dessus. Lorsque l'on est proche de  $\pi/2$ , la seule possibilité pour rester dans le domaine est d'avoir un Brownien très négatif pour repasser très vite dans le quart de plan inférieur  $\{(x, y), x > 0, y < 0\}$  (voir figure 2.5, partie gauche). Afin de moins subir ce problème (et aussi pour gérer la rotation mentionnée dans la remarque (2.40)), on propose de restreindre le domaine à

$$\mathcal{S} := \{(r \cos \theta, r \sin \theta), r > 0, \theta \in ]-\frac{\pi}{2}, \frac{\pi}{4}[\}. \quad (2.48)$$

**Remarque 2.45** Cette restriction a bien un sens pour la minoration puisque si  $A \subset B$  et  $\tau_A, \tau_B$  désignent respectivement les temps de sortie de  $A$  et  $B$ , alors pour tout  $x \in A$ ,  $\mathbb{P}_x(\tau_B > t) \geq \mathbb{P}_x(\tau_A > t)$ . Lorsque la loi initiale relative à  $\tau_B$  a un support non contenu dans  $A$  (comme c'est le cas dans la proposition ci-dessous), ce type d'argument peut encore être utilisé pour une minoration asymptotique sous réserve de possibilité de retourner dans  $A$  en un temps fini avec probabilité strictement positive (ce point pouvant se traiter via un argument de contrôlabilité).

Le générateur de  $\mathcal{L}_h$  de (2.42) (en coordonnées polaires) s'écrit (sous la forme moins agréable)

$$\mathcal{L}_h = -\rho r \partial_r + \partial_\theta + \frac{\sin^2 \theta}{2} \partial_{rr}^2 - \frac{\sin \theta \cos \theta}{r^2} \partial_\theta + \frac{\sin \theta \cos \theta}{r} \partial_{r\theta}^2 + \frac{\cos^2 \theta}{2r} \partial_r + \frac{\cos^2 \theta}{2r^2} \partial_\theta^2 \quad (2.49)$$

et on cherche cette fois  $g$  sous la forme

$$g(r, \theta) = r^n \gamma(\theta) e^{\beta(\theta)r^2}, \quad (2.50)$$

où  $\gamma$  est une fonction strictement positive sur  $[-\pi/2, \pi/4]$  et nulle au bord tandis que  $\beta$  est négative sur ce même intervalle. Via une décomposition de  $\mathcal{L}_h g/g$  similaire au cas elliptique (mais plus compliquée), on obtient le résultat suivant :

**Proposition 2.46** ([34]) Soit  $\rho \geq 0$ .

(i) Soit  $g$  définie par (2.50) avec  $n = 2$ ,

$$\gamma(\theta) = \begin{cases} -\sin(2\theta) & \text{si } \theta \in [-\frac{\pi}{2}, -\frac{\pi}{4}] \\ \cos^2(\pi/4 + \theta) & \text{si } \theta \in [-\frac{\pi}{4}, \frac{\pi}{4}] \end{cases} \quad (2.51)$$

et  $\beta(\theta) = -\frac{1}{2}$ . Alors, pour tous  $r > 0$  et  $\theta \in ]-\frac{\pi}{2}, \frac{\pi}{4}[\setminus\{-\frac{\pi}{4}\}$ ,

$$\mathcal{L}_h g(r, \theta) \geq -(3 + 2\rho)g(r, \theta).$$

(ii) Par conséquent, pour tout demi-plan ouvert  $H$  tel que  $S \subset H$ , pour toute mesure de probabilité  $m_0$  sur  $\mathbb{R}^2$  telle que  $m_0(H) = 1$ , on a

$$\limsup_{t \rightarrow +\infty} -\frac{1}{t} \log(\mathbb{P}_{m_0}(\tau_H \geq t)) \leq 3 + 2\rho.$$

Une représentation de  $\theta \mapsto g(r, \theta)$  est donnée dans la figure 2.5 (partie droite). La dissymétrie de cette fonction nous est techniquement nécessaire. Celle-ci peut elle-aussi s'interpréter comme une conséquence de la dynamique de rotation. À l'aide de la proposition 2.39, on

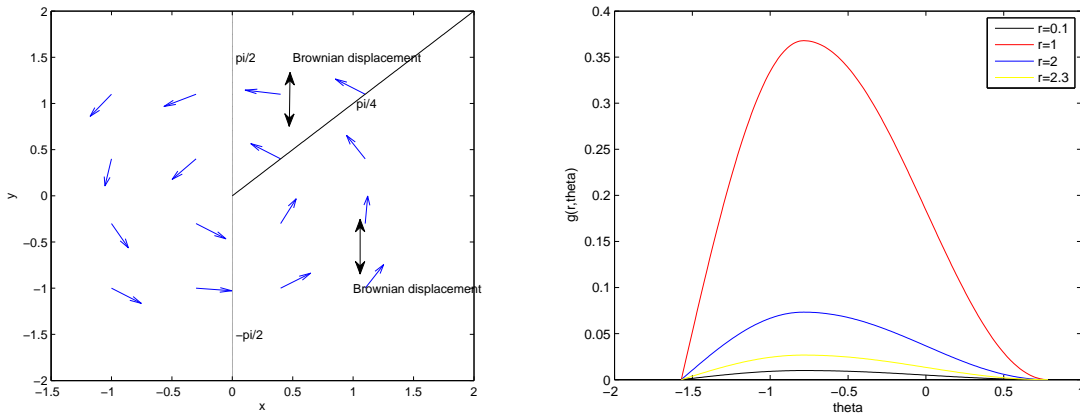


Figure 2.5 : Gauche : dynamique du système déterministe et direction possibles pour le Brownien. Droite :  $\theta \mapsto g(r, \theta)$  pour plusieurs valeurs de  $r$ .

peut maintenant énoncer le résultat principal de minoration pour  $\tau$  :

**Théorème 2.47 ([34])** Soit  $(Z_t)_{t \geq 0}$  processus solution de (2.39) avec  $(1 + \frac{1}{\sqrt{2}})a \leq b$ . Soit  $\tau = \inf\{t > 0, X_t = 0\}$ . Alors, pour toute loi initiale  $m_0$  sur  $\mathbb{R}^2$  telle que  $m_0(\{(x, y), x > 0\}) = 1$ ,

$$\limsup_{t \rightarrow +\infty} -\frac{1}{t} \log(\mathbb{P}_{m_0}(\tau \geq t)) \leq \left(3 + \frac{a}{\omega}\right) \omega.$$

En particulier, il existe une constante  $C$  positive (dépendant de  $m_0$  et des paramètres  $a$ ,  $b$  et  $c$ ) telle que

$$\forall t \geq 0, \quad \mathbb{P}_{m_0}[\tau > t] \leq C \exp(-4\omega t).$$

**Remarque 2.48** La seconde assertion se déduit de la première en remarquant que  $3 + \frac{a}{\omega} < 4$  lorsque  $(1 + \frac{1}{\sqrt{2}})a \leq b$ . On peut noter que la borne obtenue tend vers 3 lorsque  $a$  tend vers 0 (cadre où la composante spéculative est forte).

#### 2.3.4.2 Approche probabiliste

L'approche précédente ne permet pas réellement une compréhension dynamique du problème. En particulier, elle ne permet pas d'avoir une idée du comportement du processus conditionné à ne pas sortir. Sans répondre totalement à cette question, on propose dans

cette partie de retrouver une minoration en construisant explicitement un conditionnement qui assure une "survie" du processus.

Le principe de la construction suggérée ci-dessus est basé sur les ponts (de diffusion) associés à  $(\hat{Z}_t)_{t \geq 0} := (U_t, V_t)_{t \geq 0}$  solution de (2.42). Dans ce cadre gaussien, le point intéressant est que les ponts le sont également. On a :

$$\forall T > 0, \quad \forall u \in [0, 1], \quad \mathcal{L}(\hat{Z}_{uT}^{z_0} | \hat{Z}_T^{z_0} = z_T) = \mathcal{N}\left(\eta_u^{(T)}(z_0, z_T), \sigma_{uT}^{(T)}(z_0, z_T)\right)$$

où  $\eta_u^{(T)}(z_0, z_T) = (\eta_u^{(T),1}, \eta_u^{(T),2})(z_0, z_T) \in \mathbb{R}^2$  et  $\sigma_{uT}^{(T)}(z_0, z_T) \in \mathbb{M}_2$  sont explicitables (voir [34] pour plus de détails). En particulier, on peut obtenir le résultat suivant :

**Proposition 2.49 ([34])** *Pour tous  $z_0, z_T \in \mathbb{R}^2$  et  $u \in (0, 1)$ , on a*

$$\lim_{T \rightarrow 0_+} (\eta_{uT}^{(T),1}, \eta_{uT}^{(T),2})'(z_0, z_T) = \varphi_{z_0, z_T}(u) \quad \text{et} \quad \lim_{T \rightarrow 0_+} \sigma_{uT}^{(T)}(z_0, z_T) = 0$$

$$\text{où } \forall u \in [0, 1], \quad \varphi_{z_0, z_T}(u) := \begin{pmatrix} x_0 + (x_T - x_0)(3u^2 - 2u^3) \\ 6u(1-u)(x_0 - x_T) \end{pmatrix}.$$

**Remarque 2.50** Ce résultat peut s'étendre à une convergence fonctionnelle : si l'on note  $(\xi_{uT}^{(T)}(z_0, z_T))_{u \in [0,1]}$ , le processus solution de (2.42) issu de  $z_0$  et conditionné par  $\hat{Z}_T = z_T$ , alors

$$((\xi_{uT}^{(T),1}, T\xi_{uT}^{(T),2})'(z_0, z_T))_{u \in [0,1]} \xrightarrow{T \rightarrow 0_+} \varphi_{z_0, z_T} \quad \text{en probabilité,}$$

pour la convergence uniforme sur  $\mathcal{C}([0, 1], \mathbb{R}^2)$ .

La proposition précédente implique en substance que lorsque  $T$  est petit et que  $|x_T - x_0| \leq 1$ , la dynamique du processus est essentiellement verticale (avec grande probabilité puisque la variance tend vers 0). En s'appuyant sur des versions légèrement plus précises des assertions précédentes, on en déduit alors qu'on peut construire une boîte  $B$  incluse dans  $D = \{(x, y), x > 0\}$  telle que pour  $T$  suffisamment petit,

$$\inf_{z_0, z_T \in B} \mathbb{P}(\tau > T | \hat{Z}_T^{z_0} = z_T) \geq \varepsilon > 0.$$

La construction globale de trajectoires ne sortant pas de  $D$  est alors obtenue en itérant ce procédé via la propriété de Markov et en utilisant les propriétés complémentaires suivantes :

- $B$  est un ensemble accessible : pour tout  $z_0 \in D$ , il existe  $T_0 > 0$  tel que  $\mathbb{P}(\hat{Z}_{T_0}^{z_0} \in B) > 0$  (propriété basée sur un argument de contrôlabilité).
- la probabilité d'être dans  $B$  à l'instant  $T$  en partant d'un point de  $B$  est minorée :  $\forall T \in (0, 1], \inf_{z_0 \in B} \mathbb{P}_{z_0}(\hat{Z}_T \in B) > 0$ .

Ces ingrédients nous permettent alors de retrouver une version non quantitative de (2.47)(ii). Via le changement de variable proposé dans la proposition 2.39, on retrouve ensuite une versions elle-aussi non quantitative de (2.39) (voir Proposition 35 de [34] pour plus de détails).

**Remarque 2.51** L'obtention de bornes quantitatives compétitives pourrait être envisagée mais semble très difficile via cette approche. En particulier, le choix de  $T$  semble problématique. Si  $T$  est petit comme c'est le cas ici, les bornes obtenues sont automatiquement mauvaises (l'itération a un coût). Cependant, si  $T$  est plus grand, la quantité  $\inf_{z_0 \in B} \mathbb{P}_{z_0}(\hat{Z}_T \in B)$  se dégrade fortement (car le système tourne) et la probabilité que le pont ne sorte pas du domaine également.



### 2.3.5 Mesure quasistationnaire et persistance

Les problèmes abordés ci-dessus sont fortement reliés aux mesures quasistationnaires de  $(Z_t, \tau)_{t \geq 0}$ . On rappelle que une probabilité  $\nu_D$  sur  $D = \{(x, y), x > 0\}$  est dite quasistationnaire si

$$\forall t > 0, \quad \forall A \in \mathcal{B}(D), \quad \mathbb{P}_{\nu_D}(Z_t \in A | \tau > t) = \nu_D(A).$$

On montre alors via la propriété de Markov (voir *e.g.* [19]) que si  $\nu_D$  est une probabilité quasi-stationnaire, alors  $\tau$  a une loi exponentielle sous la loi  $\mathbb{P}_{\nu_D}$ . En d'autres termes, il existe  $\lambda_0(\nu_D) \in \mathbb{R}_+^*$  tel que

$$\forall t > 0, \quad \mathbb{P}_{\nu_D}(\tau > t) = \exp(-\lambda_0(\nu_D)t).$$

On appellera alors *persistance* ou *vitesse d'extinction* la quantité  $\lambda_0(\nu_D)$ . Notons que celle-ci est définie de manière intrinsèque lorsque  $\nu_D$  est unique. Ce type de propriété n'est pas démontré dans ce cadre (voir la section 2.4 pour plus de détails sur le sujet). Néanmoins, les bornes obtenues dans l'article nous permettent d'avoir des bornes uniformes pour cette quantité.

**Proposition 2.52** ([34]) *Supposons que  $(1 + \frac{1}{\sqrt{2}})a \leq b$ . Considérons  $(Z_t)_{t \geq 0}$  solution de (2.39) et  $\tau = \inf\{t > 0, X_t = 0\}$ . Le couple  $(Z, \tau)$  admet au moins une probabilité quasi-stationnaire  $\nu_D$ . Pour une telle probabilité, on a*

$$\frac{\ln 2}{\pi} \omega \leq \lambda_0(\nu_D) \leq 4\omega.$$

Comme cela a été annoncé au début de la section, on constate donc que dans ce type de modèle, la persistance est essentiellement proportionnelle à la vitesse de rotation du système déterministe.

## 2.4 Commentaires et perspectives

À propos du modèle de gradient moyenné. Les résultats obtenus sur ce modèle d'ordre 2 dégénéré sont de deux types. Dans le premier travail (cf section 2.1), on s'est attaché à comprendre de manière exhaustive le comportement asymptotique du processus en fonction de la mémoire tandis que dans le second (cf section 2.2), on s'est focalisé sur la convergence de la mesure invariante vers les minima globaux lorsque le processus est markovien ergodique, *i.e.* lorsque la mémoire est exponentielle. Bien que ces travaux développent certains outils nouveaux pour l'étude des modèles d'ordre 2, ils ne permettent pas pour l'instant de répondre à l'une des questions posées initialement : dans quels cas ce type de dynamique d'optimisation est-il efficace ? Outre le fait que la "concentration" de la mesure invariante ne soit établie dans un cas simple (et sous hypothèses), il manque aussi un outil important : des bornes quantitatives pour la convergence à l'équilibre du semi-groupe (la proposition 2.15 basée sur des méthodes de type Meyn-Tweedie fournit un résultat non quantitatif).

Afin d'obtenir une convergence (quantitative) dans  $L^2(\mu)$  (où  $\mu$  désigne la probabilité invariante), une méthode classique consiste à prouver une inégalité de Poincaré pour la mesure invariante. Une telle technique consiste à contrôler la norme  $L^2$  (ou plutôt sa variance) d'une fonction par son énergie, en l'occurrence la moyenne (sous la mesure invariante) du

carré de son gradient. En particulier, la meilleure constante dans l'inégalité de Poincaré donne le meilleur exposant dans la convergence exponentielle du semi-groupe vers sa mesure d'équilibre. Ce type de résultat n'est a priori pas accessible dans les modèles considérés ici. Néanmoins, depuis les travaux de Villani ([96]) dans le cadre *hypocoercif*, des outils efficaces ont été développés pour contourner l'inégalité de Poincaré (voir *e.g.* [70, 5]) afin de retrouver des convergences exponentielles dans  $L^2(\mu)$  (dans un sens affaibli). Il semble que certains de ces outils puissent être adaptés ici. Néanmoins, la difficulté principale semble résider dans le caractère non explicite de la mesure invariante dans notre modèle. Ceci pose en particulier problème pour obtenir une inégalité de Poincaré "classique" pour la mesure invariante, *i.e.* vis-à-vis de l'opérateur non dégénéré, propriété qui joue un rôle important dans ces travaux.

### À propos du modèle de bulle spéculative.

- *Un modèle un peu plus complexe ?* Le modèle de bulle spéculative présenté précédemment a pour intérêt de proposer une dynamique naturelle correspondant au phénomène de spéculation. Néanmoins, il est naturellement trop simple pour être réaliste. Parmi les imperfections de ce modèle, on peut penser en premier lieu au caractère non minoré du processus. Si pour la question du temps de retour en zéro étudiée ici (à comprendre dans un sens relatif), il ne s'agit pas d'un réel problème (puisque l'on s'intéresse à la dynamique avant le retour au prix intrinsèque), cela devient un peu plus gênant lorsque l'on cherche à modéliser cette dynamique de manière globale. D'un point de vue plus général, le modèle manque actuellement de dissymétrie, notamment aux sens suivants : d'une part, la dynamique de montée d'un phénomène de bulle est généralement différente de celle de descente dont la pente est souvent plus forte. D'autre part, on peut imaginer que la spéculation est un phénomène plus présent lorsque le prix dépasse le prix intrinsèque que dans le cas contraire. Notons que ce dernier point peut rejoindre le problème de positivité évoqué précédemment. Les pistes possibles pour gérer ces dissymétries sont nombreuses. Pour la question de la dynamique et de la valeur du processus lorsque celui est en-dessous du prix intrinsèque, on peut penser à considérer une fonction du prix relatif qui "écrase" les faibles valeurs et ses variations. De même, pour celle de la dissymétrie entre les montées et les descentes, une idée simple consiste à introduire de la non-linéarité dans le terme de spéculation, *i.e.* que la dynamique de spéculation dépende du signe des variations passées du processus. Du point de vue théorique, ces idées ont le défaut de faire disparaître le caractère gaussien du processus. Afin de préserver cette propriété, on pourrait aussi envisager de considérer des modèles d'ordre supérieur (voir [34] pour plus de détails sur ce sujet).
- *Mesures quasistationnaires.* La question de l'unicité de la probabilité quasi-stationnaire  $\nu_D$  introduite dans la section 2.3.5 n'est pas résolue. L'obtention d'une telle propriété est en général fortement reliée au *retour de l'infini* : si  $(X_t)_{t \geq 0}$  est une diffusion en dimension 1 et que  $\tau$  désigne le temps d'absorption en 0, l'unicité est quasi-équivalente (sous réserve de non-dégénérescence) à la condition (dite de retour de l'infini) :  $\sup_{x > 1} \mathbb{E}_x[\tau_1] < +\infty$  où  $\tau_1 = \inf\{t > 0, X_t = 1\}$ . Ceci implique en particulier qu'un processus de Ornstein-Uhlenbeck (O.U.) n'admet pas une unique mesure quasistationnaire. En effet, en dimension 1, revenir de l'infini "rapidement" n'est possible que si la force de rappel est plus que linéaire (on peut penser par exemple à la diffusion

dite *logistique*, voir [13]).

En dimension 2, le problème semble être fonction des valeurs propres de  $A$ . Lorsque les valeurs propres sont réelles, le comportement étant fortement relié à celui d'un O.U. réel, il est à peu près clair que l'on n'a pas unicité. En revanche, lorsque  $A$  est à valeurs propres complexes, le caractère tournant de la dynamique déterministe permet au processus de revenir très rapidement près de la frontière. Bien que cette propriété laisse penser que l'on puisse espérer l'unicité de la QSD dans ce cadre, elle n'est clairement pas suffisante. En effet, si l'on se réfère aux récents critères (probabilistes) établis dans [14], se rapprocher de la frontière rapidement est en quelque sorte une condition nécessaire mais pas suffisante. Il faut en fait montrer que l'on retourne rapidement dans un compact du domaine. Dans le cas où le domaine est l'ensemble  $D$  considéré dans la section précédente, la difficulté principale revient à prouver que lorsque l'on est proche du "point"  $(0, \pm\infty)$ , alors, conditionnellement au fait de rester dans le domaine, on revient rapidement dans un compact. Ce travail est actuellement en cours d'étude.



## Chapitre 3

### Approximation du régime stationnaire de diffusions et problèmes associés

Ce chapitre est une synthèse de travaux liés à l'approximation du régime stationnaire (au sens marginal ou fonctionnel) pour des processus de Markov et en particulier pour les diffusions. Après un état de l'art, on s'intéresse dans la section 3.2 à la convergence et à la vitesse de convergence d'un algorithme pour l'approximation fonctionnelle du régime stationnaire. Cette procédure étant a priori coûteuse (dans un cadre fonctionnel), on propose ensuite (Section 3.3) d'étudier un algorithme alternatif moins moyennisant dit *mixed-step* pour lequel on établit également des résultats de convergence et de vitesse. On développe en particulier une méthode de Richardson-Romberg dont l'optimalité du point de vue de la variance asymptotique dépend des propriétés de confluence de la diffusion. Plus exactement, cette dernière requiert une propriété de confluence asymptotique faible : l'unicité de la mesure invariante relative au système dupliqué, *i.e.* du système dynamique constitué de deux trajectoires issues de la diffusion originelle. L'étude de ce problème est l'objet de la section 3.4. Ces parties s'appuient principalement sur les articles [76], [78], [79] et [65].

#### 3.1 État de l'art

La connaissance du comportement à l'équilibre d'un système dynamique est une question importante pour la pratique mais souvent non soluble explicitement. Il est donc naturel de chercher à développer des algorithmes (efficaces) permettant son approximation. Dans le cadre des processus aléatoires markoviens, il s'agit d'approcher la (ou les) mesure(s) invariante(s) du système ou plus généralement, la loi du processus lorsqu'il évolue en régime stationnaire. L'idée de base d'approximation de ces probabilités est d'exploiter les propriétés de convergence en temps long vers ces dernières. Pour un processus de Markov ergodique  $(X_t)_{t \geq 0}$  de semi-groupe  $(P_t)_{t \geq 0}$  et de mesure invariante  $\nu$ , on peut envisager deux types de convergence (lorsque  $t \rightarrow +\infty$ ) :

- Convergence en loi :  $P_t(x, \cdot) \rightarrow \nu$  ou plus généralement,  $\mathcal{L}(X_{t+}^x) \rightarrow \mathbb{P}_\nu$  où  $X_{t+}^x = (X_{t+s}^x)_{s \geq 0}$  et  $\mathbb{P}_\nu$  désigne la loi de  $(X_t)_{t \geq 0}$  en régime stationnaire, *i.e.* de loi initiale  $\nu$ .
- Convergence *p.s.* de la mesure empirique (ou d'occupation) marginale ou fonctionnelle :

$$\frac{1}{t} \int_0^t \delta_{X_s^x} ds \Longrightarrow \nu \quad p.s. \quad \text{ou} \quad \frac{1}{t} \int_0^t \delta_{X_{s+}^x} ds \Longrightarrow \mathbb{P}_\nu \quad p.s.$$

où " $\Longrightarrow$ " désigne la convergence étroite.

L'ensemble des travaux qui vont suivre est basé sur la deuxième approche qui, du point de vue numérique, a un avantage certain : elle ne nécessite la simulation que d'une seule trajectoire. Dans ce cadre, le but est alors de construire une suite de mesures empiriques  $(\bar{\nu}_n)$  simulable telle que

$$\bar{\nu}_n(\omega, dy) \xrightarrow{n \rightarrow +\infty} \nu(dy) \quad p.s. \quad (3.1)$$

Dans le cadre des diffusions browniennes, les premiers travaux remontent au début des années 90. Dans [94], l'idée est de baser la construction de l'algorithme sur un schéma d'Euler (ou de Milstein) à pas constant  $\gamma$  associé à la diffusion, noté  $(\bar{X}_{n\gamma}^\gamma)_{n \geq 0}$ . Pour chaque  $\gamma > 0$ , on considère alors  $(\bar{\nu}_n^\gamma)_{n \geq 1}$  la suite de mesures d'occupations relative à ce schéma d'Euler. Sous des hypothèses de Lyapounov et d'uniforme ellipticité, Talay obtient la convergence lorsque  $n \rightarrow +\infty$  de  $(\bar{\nu}_n^\gamma)_{n \geq 1}$  vers l'unique mesure invariante  $\nu^\gamma$  de la chaîne de Markov  $(\bar{X}_{n\gamma}^\gamma)_{n \geq 0}$  et celle de  $\nu^\gamma$  vers  $\nu$  lorsque  $\gamma \rightarrow 0$ .

Afin d'éviter cette double convergence, on peut tenter de gérer conjointement l'erreur relative au temps long et celle liée à la discrétisation. Il s'agit alors d'introduire un schéma de discrétisation à pas décroissant vers 0. C'est le point de vue de Lamberton et Pagès ([58, 59]) qui obtiennent des résultats de convergence et de vitesse de convergence de la suite de mesures d'occupation relative à ce schéma via des arguments principalement basés sur des méthodes de martingales. Ces résultats ont ensuite été étendus par Lemaire ([62, 64]), à des diffusions à drift non lipschitzien par exemple, et par P. ([80, 82, 81]) à des EDS dirigées par des processus de Lévy. Avant de détailler les résultats les plus récents sur le sujet, on propose dans la suite de cette section de rappeler les principaux résultats existants (essentiellement dans le cadre des diffusions).

Considérons  $(X_t)_{t \geq 0}$  processus à valeurs dans  $\mathbb{R}^d$  solution de

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad (3.2)$$

où  $(W_t)_{t \geq 0}$  est un mouvement brownien standard  $q$ -dimensionnel et les coefficients  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  et  $\sigma : \mathbb{R}^d \rightarrow \mathbb{M}_{d,q}$  sont des fonctions continues ( $\mathbb{M}_{d,q}$  est l'ensemble des matrices à  $d$  lignes et  $q$  colonnes). Sous réserve d'existence forte et d'unicité des solutions,  $(X_t)_{t \geq 0}$  est un processus de Markov homogène de semi-groupe  $(P_t)_{t \geq 0}$  et de générateur infinitésimal  $\mathcal{A}$  défini pour toute fonction  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^2$  par

$$\forall x \in \mathbb{R}^d, \quad \mathcal{A}f(x) := (\nabla f(x)|b(x)) + \frac{1}{2} \text{Tr}(\sigma^*(x)D^2f(x)\sigma(x))$$

où  $\nabla f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  et  $D^2f : \mathbb{R}^d \rightarrow \mathbb{M}_{d,d}$  correspondent respectivement au gradient et à la hessienne de  $f$ . Lorsque  $X_0$  a pour loi  $\mu$ , on notera  $\mathbb{P}_\mu$ , la loi (globale) du processus  $(X_t)_{t \geq 0}$  (pour la topologie de la convergence uniforme sur tout compact sur  $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ , espaces des fonctions continues de  $\mathbb{R}_+$  de  $\mathbb{R}^d$ ), de loi initiale  $\mu$ .

**Schéma d'Euler à pas décroissant.** Soit  $(\gamma_n)_{n \geq 1}$  une suite décroissante de réels strictement positifs telle que

$$\lim_{n \rightarrow +\infty} \gamma_n = 0.$$

La suite  $(\gamma_n)_{n \geq 1}$  jouant le rôle des pas de discrétisation, on introduit également la suite  $(\Gamma_n)_{n \geq 0}$  des instants de discrétisation :  $\Gamma_0 = 0$  et

$$\forall n \geq 1, \quad \Gamma_n = \sum_{k=1}^n \gamma_k,$$

et on fait l'hypothèse que  $\Gamma_n \rightarrow +\infty$  lorsque  $n \rightarrow +\infty$ .

Le schéma d'Euler à pas décroissant associé  $(\bar{X}_{\Gamma_n})_{n \geq 0}$  est alors défini par

$$\bar{X}_0 = x \in \mathbb{R}^d, \quad \bar{X}_{\Gamma_{n+1}} = \bar{X}_{\Gamma_n} + \gamma_{n+1}b(\bar{X}_{\Gamma_n}) + \sigma(\bar{X}_{\Gamma_n})(W_{\Gamma_{n+1}} - W_{\Gamma_n}). \quad (3.3)$$

Comme on peut s'en douter, on peut, dans une partie importante des résultats ci-dessous, remplacer les accroissements du Brownien par une suite  $(\sqrt{\gamma_n}U_n)_{n \geq 1}$  où  $(U_n)_{n \geq 1}$  est un bruit blanc. On préférera ici se limiter au cadre le plus naturel.

On associe à ce schéma une suite de mesures empiriques  $(\bar{\nu}_n^\eta)_{n \geq 1}$  définie *p.s.* par

$$\bar{\nu}_n^\eta(\omega, dx) = \frac{1}{H_n} \sum_{k=1}^n \eta_k \delta_{\bar{X}_{\Gamma_{k-1}}} \quad (3.4)$$

où  $(\eta_k)_{k \geq 1}$  est une suite de réels strictement positifs telle que

$$H_n := \sum_{k=1}^n \eta_k \xrightarrow{n \rightarrow +\infty} +\infty.$$

**Remarque 3.1**  $\bar{\nu}_n^\eta$  est une discrétisation naturelle de la mesure d'occupation  $\Gamma_n^{-1} \int_0^{\Gamma_n} \delta_{X_s} ds$  lorsque  $\eta_n = \gamma_n$ . Bien que la question de l'optimalité du choix des poids soit relativement difficile, on peut au vu des travaux de Lemaire ([62]) remarquer que  $\eta_n = \gamma_n$  est aussi un "bon" choix. Enfin, supposer que  $\eta_n = \gamma_n$  simplifie souvent de manière importante les énoncés et les arguments de preuve. C'est pourquoi, les résultats seront souvent énoncés dans ce cadre. Pour terminer, notons cependant que, même dans ce cas, les mesures à poids généraux jouent un rôle non négligeable puisqu'elles apparaissent de manière naturelle dans l'étude des vitesses.

**Résultats de convergence.** On commence par rappeler des résultats de convergence *p.s.* pour la suite  $(\bar{\nu}_n^\eta(\omega, dx))_{n \geq 1}$  vers l'ensemble des probabilités invariantes de (3.2). Ceci suppose en particulier l'existence de ces dernières. Celle-ci est assurée par une hypothèse de Lyapounov construite via une classe de fonctions coercives dites *Essentiellement Quadratiques* : on note  $\mathcal{EQ}(\mathbb{R}^d)$ , l'ensemble des fonctions  $V : \mathbb{R}^d \rightarrow \mathbb{R}_+^*$  de classe  $\mathcal{C}^2$  telles que

$$\exists \rho > 0, \quad \liminf_{|x| \rightarrow +\infty} x^{-\rho} V(x) > 0, \quad |\nabla V|^2 \leq CV \quad \text{et} \quad D^2V \text{ est bornée.}$$

L'hypothèse  $\lim_{|x| \rightarrow +\infty} |x|^{-\rho} V(x) > 0$  implique la coercivité de  $V : \lim_{|x| \rightarrow +\infty} V(x) = +\infty$ . Cette dernière est en faite une hypothèse suffisante pour une bonne partie des résultats. Un des intérêts de l'hypothèse de coercivité renforcée introduite ci-dessus est d'avoir l'équivalence suivante :  $g$  est à croissance polynômiale si et seulement s'il existe  $r > 0$  tel que  $|g| \leq CV^r$ .

Pour une matrice symétrique  $S$  de  $\mathbb{M}_{d,d}$ , on note aussi  $\lambda_S^\pm := \max(0, \lambda_1, \dots, \lambda_d)$  où  $\lambda_1, \dots, \lambda_d$  représentent les valeurs propres de  $S$ . On peut maintenant introduire l'hypothèse de Lyapounov dépendant de  $a \in (0, 1]$  et  $p \in [1, +\infty)$  :

$(\mathbf{S}_{a,p})$  : Il existe  $V \in \mathcal{E}\mathcal{Q}(\mathbb{R}^d)$  telle que

$$(i) \quad \exists C_a > 0 \text{ tel que } |b|^2 + \text{Tr}(\sigma\sigma^*) \leq C_a V^a.$$

$$(ii) \quad \text{Il existe } \kappa_1 \in \mathbb{R} \text{ et } \kappa_2 > 0 \text{ tels que } \langle \nabla V, b \rangle + \lambda_p \text{Tr}(\sigma\sigma^*) \leq \kappa_1 - \kappa_2 V^a,$$

où  $\lambda_p := \frac{1}{2} \sup_{x \in \mathbb{R}^d} \lambda_{D^2V(x) + (p-1) \frac{\nabla V \otimes \nabla V}{V}}$ . Sous l'hypothèse  $(\mathbf{S}_{a,p})$ , la fonction  $V$  est appelée fonction de Lyapounov pour la diffusion  $(X_t)_{t \geq 0}$ .

Si  $(\mathbf{S}_{a,p})$  est vraie pour tout  $p \in [1, \infty)$ , l'hypothèse est alors notée  $(\mathbf{S}_{a,\infty})$ . On peut remarquer que  $(\mathbf{S}_{a,\infty})$  est satisfaite dès que  $|b|^2 \leq C_a V^a$ ,  $\langle \nabla V, b \rangle \leq \kappa_1 - \kappa_2 V^a$  ( $\kappa_2 > 0$ ) et  $\text{Tr}(\sigma\sigma^*) = o(V^a)$ .

**Remarque 3.2**  $\triangleright$  En s'appuyant sur le caractère symétrique de la gaussienne, on peut affiner l'hypothèse (ii) et la remplacer par : il existe  $a \in [0, 1]$ , il existe  $p > 0$ ,  $\kappa_p > 0$  et  $M_p > 0$  tels que

$$\forall |x| > M_p, \quad (\nabla V |b|)(x) + \frac{\lambda_1}{2} \text{Tr}(\sigma\sigma^*)(x) + \frac{(p-1)}{2} \frac{|\sigma^* \nabla V(x)|^2}{V(x)} \leq -\kappa_p V^a(x) \quad (3.5)$$

Ce raffinement permet en particulier de gérer des cas critiques tels que des diffusions à rappel (très) faible ou des systèmes dynamiques où l'intégrale brownienne génère elle-même sa stabilité (voir [80] pour plus de détails).

$\triangleright$  L'hypothèse  $(\mathbf{S}_{a,p})(i)$  implique en particulier que  $b$  est à croissance sous-linéaire. Parmi les solutions pour s'abstraire de cette hypothèse, on peut citer l'article [64] dans lequel l'introduction d'un pas adaptatif dans la procédure décrite ci-dessus permet de gérer l'erreur liée à la discrétisation de la partie drift lorsque celui a une croissance surlinéaire (qui est trop forte lorsque le pas est déterministe).

Rappelons maintenant un résultat de convergence associé à  $(\bar{v}_n^\eta(\omega, dx))_{n \geq 1}$  (issu de [59]).

**Proposition 3.3** *Supposons que  $(\mathbf{S}_{a,p})$  soit satisfaite avec  $p > 2$  et  $a \in (0, 1]$ . Supposons que  $(\eta_n/\gamma_n)_{n \geq 1}$  soit décroissante. Alors,*

(i) *La suite  $(\bar{v}_n^\eta(\omega, dx))_{n \geq 1}$  satisfait*

$$\sup_{n \geq 1} \bar{v}_n(\omega, V^{\frac{p}{2}+a-1}) < +\infty \quad p.s. \quad (3.6)$$

*En particulier,  $(\bar{v}_n^\eta(\omega, dx))_{n \geq 1}$  est p.s. tendue.*

(ii) *Toute valeur d'adhérence de  $(\bar{v}_n^\eta(\omega, dx))_{n \geq 1}$  est p.s. une probabilité invariante pour  $(X_t)_{t \geq 0}$ . En particulier, lorsque cette dernière (que l'on notera  $\nu$ ) est unique alors, pour toute fonction continue  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  telle que  $f(x) = o(V^{\frac{p}{2}+a-1}(x))$  lorsque  $|x| \rightarrow +\infty$ ,  $\bar{v}_n^\eta(\omega, f) \xrightarrow{n \rightarrow +\infty} \nu(f)$  p.s. Par conséquent, sous l'hypothèse  $(\mathbf{S}_{a,\infty})$ , la convergence précédente vaut pour toute fonction continue à croissance polynômiale.*

**Remarque 3.4** Dans le même sens que la remarque 3.2, on peut à nouveau noter que le cadre proposé ici peut être étendu. L'hypothèse de décroissance de  $(\eta_n/\gamma_n)_{n \geq 1}$  simplifie la preuve mais n'est pas indispensable. De même, on peut en renforçant légèrement l'hypothèse de Lyapounov étendre la convergence à des fonctions à croissance sous-exponentielle (voir [62] pour plus de détails).



**Vitesse de convergence.** En accord avec la remarque 3.1, on suppose dans cette partie que pour tout  $n \geq 1$ ,  $\eta_n = \gamma_n$  et on note  $(\bar{\nu}_n(\omega, dx))_{n \geq 1}$  la suite de mesures empiriques dans ce cas. On suppose également que la probabilité invariante notée  $\nu$  est unique. Le but du résultat ci-dessous est d'obtenir pour une fonction  $f$  donnée, un TCL pour la suite de variables aléatoires réelles  $(\bar{\nu}_n(\omega, f) - \nu(f))_{n \geq 1}$ . Ce type de résultat s'appuie sur une hypothèse d'existence de solution à l'équation de Poisson : on suppose qu'il existe  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  telle que  $\mathcal{A}g = f - \nu(f)$ . A temps continu, ce type d'hypothèse se comprend très simplement via la formule d'Itô (ou de Dynkin dans un cadre plus général). En effet, dans ce cas,

$$\frac{1}{t} \int_0^t f(X_s) ds - \nu(f) = \frac{1}{t} \int_0^t \mathcal{A}g(X_s) ds = \frac{g(X_t) - g(X_0)}{t} - \frac{1}{t} \int_0^t (\nabla g(X_s) | \sigma(X_s) dB_s) \quad (3.7)$$

et l'on déduit alors facilement (sous hypothèses sur  $\mathcal{A}$ ,  $f$  et  $g$  adéquates), via un TCL pour la martingale Brownienne, que

$$\frac{1}{\sqrt{t}} \left( \int_0^t f(X_s) ds - \nu(f) \right) \Longrightarrow \mathcal{N}(0, \sigma_f^2)$$

où

$$\int |\sigma^* \nabla g|^2 d\nu := \int |\sigma^*(x) \nabla g(x)|^2 \nu(dx) = -2D_\nu(g)$$

où  $D_\nu$  est la forme de Dirichlet sur  $L^2(\nu)$  définie par  $D_\nu(g) = \int g(x) \mathcal{A}g(x) \nu(dx)$ .

Dans un cadre discrétisé, l'idée rappelée ci-dessus (voir *e.g.* [8] pour plus de détails) joue toujours un rôle important mais le problème est plus délicat mettant en compétition l'erreur de discrétisation et l'erreur en temps long. Introduisons d'abord l'hypothèse d'existence de solution à l'équation de Poisson (relative à une fonction  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ) :

(**C(f, k)**) : Il existe une fonction de classe  $\mathcal{C}^k$   $g_f : \mathbb{R}^d \rightarrow \mathbb{R}$  solution de  $f - \nu(f) = \mathcal{A}g_f$  telle que  $f$ ,  $g_f$  et leurs dérivées partielles d'ordre inférieur à  $k$  sont dominées par  $V^r$  ( $r \geq 0$ ).

**Remarque 3.5** Lorsque  $f$  est à support compact, on peut vérifier (**C(f, k)**) en s'appuyant sur des résultats en domaine borné (de condition de bord  $g_f = 0$ ), cadre dans lequel les références sont assez nombreuses. En particulier, d'après [57] (Theorems III.1.1 and III.1.2), l'hypothèse (**C<sub>F</sub><sup>2</sup>**) est satisfaite lorsque la diffusion est uniformément elliptique et que  $f$  et les coefficients  $b$  et  $\sigma$  sont  $\mathcal{C}^{k-2, \rho}$  (avec  $\rho \in (0, 1)$ ), *i.e.* de classe  $\mathcal{C}^{k-2}$  dont les dérivées d'ordre  $k$  sont globalement  $\rho$ -höldériennes. Dans le cas général, le problème est abordé dans une série d'articles de Pardoux et Veretennikov ([83, 84, 85]) dans lesquels le problème est résolu sous des hypothèses d'ellipticité dans des espaces de Sobolev et où des contrôles de  $g$  et de ses premières dérivées sont donnés. On reviendra sur cette question à la fin de la section 3.3.1 où l'on donnera un critère dans un cas (éventuellement dégénéré) où la diffusion a des propriétés de *confluence asymptotique*.

Avant de rappeler le résultat de vitesse de convergence de  $(\bar{\nu}_n(\omega, dx))$ , introduisons quelques notations :

$$\forall r \in \mathbb{N}, \quad \Gamma_n^{(r)} = \sum_{k=1}^n \gamma_k^r.$$

Pour  $h : \mathbb{R}^d \rightarrow \mathbb{R}$  (suffisamment régulière) et  $r$  entier supérieur à 2, on note :

$$D^{(r)}h(x)y_1 \otimes \cdots \otimes y_r = \sum_{(i_1, \dots, i_r) \in \{1, \dots, d\}^r} \partial_{x_{i_1}, \dots, x_{i_r}}^r h(x)y_1^{i_1} \dots y_r^{i_r}.$$

On a alors le résultat suivant (adapté de [58] (théorème 10)).

**Proposition 3.6** *Supposons que  $(\mathbf{S}_{a, \infty})$  soit satisfaite avec  $a \in (0, 1]$  et que  $\nu$  soit l'unique probabilité invariante associée à  $(X_t)_{t \geq 0}$ . Soit  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  satisfaisant  $(\mathbf{C}(\mathbf{f}, \mathbf{5}))$ . Alors,*

- Si  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \xrightarrow{n \rightarrow +\infty} 0$ ,

$$\sqrt{\Gamma_n}(\bar{\nu}_n(\omega, f) - \nu(f)) \xrightarrow{(\mathbb{R})} \mathcal{N}\left(0; \int_{\mathbb{R}^d} |\sigma^* \nabla g_f|^2 d\nu\right) \quad \text{lorsque } n \rightarrow +\infty.$$

- Si  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \xrightarrow{n \rightarrow +\infty} \tilde{\beta} \in (0, +\infty]$ ,

$$\triangleright \sqrt{\Gamma_n}(\bar{\nu}_n(\omega, f) - \nu(f)) \xrightarrow{(\mathbb{R})} \mathcal{N}\left(\tilde{\beta} m_{g_f}^{(1)}; \int_{\mathbb{R}^d} |\sigma^* \nabla g_f|^2 d\nu\right) \quad \text{lorsque } n \rightarrow +\infty \quad \text{si } \tilde{\beta} \in (0, +\infty),$$

$$\triangleright \frac{\Gamma_n}{\Gamma_n^{(2)}}(\bar{\nu}_n(\omega, f) - \nu(f)) \xrightarrow{a.s.} m_{g_f}^{(1)} \quad \text{lorsque } n \rightarrow +\infty \quad \text{si } \tilde{\beta} = +\infty$$

où  $m_{g_f}^{(1)} = \int_{\mathbb{R}^d} \varphi_1 d\nu$  avec

$$\varphi_1(x) = \frac{1}{2} D^2 g_f(x) b(x)^{\otimes 2} + \frac{1}{2} \mathbb{E}[D^3 g_f(x) b(x) (\sigma(x) U_1)^{\otimes 2}] + \frac{1}{24} \mathbb{E}[D^4 g_f(x) (\sigma(x) U_1)^{\otimes 4}]. \quad (3.8)$$

**Remarque 3.7** Lorsque  $\gamma_n \propto n^{-\rho}$ ,  $0 < \rho < 1$ ,

$$\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \xrightarrow{n \rightarrow +\infty} 0 \iff \rho > \frac{1}{3}.$$

Comme pour tout  $\rho \in (0, 1)$ ,  $\Gamma_n \propto n^{1-\rho}$  et que pour tout  $\rho \in (0, 1/2)$ ,  $\Gamma_n^{(2)} \propto n^{1-2\rho}$ , on en déduit que la vitesse de convergence est de l'ordre de  $n^{-\varphi(\rho)}$  où

$$\varphi(\rho) = \begin{cases} \frac{1-\rho}{2} & \text{si } \rho \in (1/3, 1) \\ \rho & \text{si } \rho \in (0, 1/3). \end{cases}$$

En particulier, la vitesse optimale est de l'ordre de  $n^{-\frac{1}{3}}$ .

## 3.2 Approximation fonctionnelle du régime stationnaire

L'objectif de cette partie initiée dans [80] puis développée dans [76] et [78] est d'étudier une forme fonctionnelle de la suite  $(\bar{\nu}_n^\eta(\omega, dx))_{n \geq 1}$  définie dans (3.4), afin de tenter d'approcher le régime stationnaire global noté  $\mathbb{P}_\nu$ . De manière à bien isoler les outils nécessaires au passage du marginal au fonctionnel, on propose pour la partie convergence de se placer dans un cadre général : on considère  $(X_t)_{t \geq 0}$  un processus de Markov fellerien à valeurs dans  $\mathbb{R}^d$  de

semi-groupe  $(P_t)_{t \geq 0}$  et on suppose construit un schéma de discrétisation  $(\bar{X}_{\Gamma_n})_{n \geq 0}$ . On note  $(\bar{X}_t)_{t \geq 0}$  une version à temps continu de ce schéma de discrétisation qu'on suppose càdlàg (le schéma constant par morceaux par exemple) et pour tout  $k \geq 0$ , on note  $(\bar{X}_t^{(\Gamma_k)})_{t \geq 0}$  le schéma shifté de  $\Gamma_k$  défini par

$$\bar{X}_t^{(\Gamma_k)} := \bar{X}_{\Gamma_k+t}.$$

On peut alors introduire la suite  $(\bar{\nu}^{(n)}(\omega, d\alpha))_{n \geq 1}$  de mesures empiriques sur l'espace de Skorokhod  $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$  définie par :

$$\bar{\nu}^{(n)}(\omega, d\alpha) = \frac{1}{H_n} \sum_{k=1}^n \eta_k \mathbf{1}_{\{\bar{X}^{(\Gamma_{k-1})}(\omega) \in d\alpha\}}$$

où les poids  $(\eta_k)_{k \geq 1}$  sont définis de manière identique à la section précédente. On notera aussi  $\bar{\nu}_0^{(n)}(\omega, dx) = (\bar{\nu}^{(n)} \circ \pi_0^{-1})(\omega, dx)$  où  $\pi_0 : \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d) \rightarrow \mathbb{R}^d$  est définie par  $\pi_0(\alpha) = \alpha(0)$ . En d'autres termes,  $\bar{\nu}_0^{(n)}(\omega, dx)$  coïncide avec  $\bar{\nu}_n(\omega, dx)$  définie par (3.4).

**A propos de la simulation de  $(\bar{\nu}^{(n)}(\omega, F))_{n \geq 1}$ .** Pour une fonctionnelle  $F : \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d) \rightarrow \mathbb{R}$ , la suite  $(\bar{\nu}^{(n)}(\omega, F))_{n \geq 1}$  peut être formellement simulée via la relation de récurrence suivante :

$$\forall n \geq 1, \quad \bar{\nu}^{(n+1)}(\omega, F) = \bar{\nu}^{(n)}(\omega, F) + \frac{\eta_{n+1}}{H_{n+1}} (F(X^{(\Gamma_n)}(\omega)) - \bar{\nu}^{(n)}(\omega, F)). \quad (3.9)$$

Il nous faut donc naturellement simuler  $F(\bar{X}^{(\Gamma_n)})$  à chaque pas de temps. Supposons que  $F : \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d) \rightarrow \mathbb{R}$  soit une fonctionnelle dépendant seulement de la trajectoire entre 0 et  $T$ , et que (pour simplifier),  $(\bar{X}_t)_{t \geq 0}$  soit le schéma à temps continu constant par morceaux. En notant

$$N(n, T) := \inf\{k \geq n, \Gamma_{k+1} - \Gamma_n > T\} = \max\{k \geq 0, \Gamma_k - \Gamma_n \leq T\} \quad n \geq 0, T > 0, \quad (3.10)$$

on constate que la simulation de  $F(\bar{X}^{(\Gamma_n)})$  requiert la connaissance du vecteur  $(\bar{X}_{\Gamma_n}, \dots, \bar{X}_{\Gamma_{N(n, T)}})$ . Ainsi, pour  $n = 0$ , on simule le schéma jusqu'à l'indice  $N(0, T)$  puis après avoir calculé  $F(\bar{X}^{(0)})$ , on garde en mémoire le vecteur  $(\bar{X}_{\Gamma_1}, \dots, \bar{X}_{\Gamma_{N(0, T)}})$  et on simule le schéma entre les indices  $N(0, T) + 1$  et  $N(1, T)$ , et ainsi de suite (le caractère décroissant de  $(\gamma_n)_{n \geq 1}$  implique que  $n \mapsto N(n, T)$  est strictement croissante).

Le problème non soulevé dans ce qui précède est le coût de simulation de  $F(\bar{X}^{(\Gamma_n)})$ . En effet, le nombre de calculs à chaque pas de temps est a priori de l'ordre de  $N(n, T) - n$ , quantité qui explose avec  $n$ . L'algorithme n'est donc absolument pas linéaire. Néanmoins, comme les intervalles  $[\Gamma_n, \Gamma_{N(n, T)}]$  et  $[\Gamma_{n+1}, \Gamma_{N(n+1, T)}]$  ont un chevauchement important, il est possible, pour certaines classes de fonctionnelles d'introduire de la récursivité dans la simulation de la suite  $(F(X^{(\Gamma_n)}))_{n \geq 0}$ . C'est le cas par exemple, si

$$F(\alpha(t), t \in [0, T]) = \varphi \left( \int_0^T g(\alpha(s)) ds \right) \quad \text{ou si} \quad F(\alpha(t), t \in [0, T]) = \varphi \left( \sup_{t \in [0, T]} g(\alpha(s)) \right).$$

Dans la première situation (qui correspond au cadre des options asiatiques par exemple), la linéarité de l'intégrale implique que le nombre de calculs nécessaire à chaque pas de temps est de l'ordre de  $N(n+1, T) - N(n, T)$  de sorte que le nombre d'opérations effectuées jusqu'à l'instant  $n$  est de l'ordre de  $N(n, T)$ , i.e. du nombre de points de discrétisation nécessaires. (notons que lorsque  $\gamma_n = C_1 n^{-\rho}$  avec  $\rho \in (0, 1)$ ,  $N(n, T) = n + O(n^\rho)$ ).

Dans la seconde, la réponse est plus délicate dépendant de la stratégie de stockage de l'information mais on comprend relativement facilement que l'on peut exploiter la connaissance du supremum sur  $[\Gamma_n, \Gamma_{N(n,T)}]$  pour en déduire celui sur  $[\Gamma_{n+1}, \Gamma_{N(n+1,T)}]$ . On reviendra sur ces questions de simulation dans la section suivante où un nouvel algorithme sera proposé pour éviter cette difficulté à maîtriser la complexité.

### 3.2.1 Convergence de $(\bar{\nu}^{(n)}(\omega, d\alpha))_{n \geq 1}$

L'idée du résultat de convergence est de supposer, dans un cadre général, la convergence marginale acquise, puis d'exhiber les hypothèses supplémentaires nécessaires pour l'étendre au cadre fonctionnel. On introduit les hypothèses suivantes :

(C<sub>0</sub>) :  $(X_t)$  admet une unique probabilité invariante  $\nu$  et

$$\bar{\nu}_0^{(n)}(\omega, dx) \xrightarrow{n \rightarrow +\infty} \nu(dx) \quad a.s.$$

qui constitue l'hypothèse de base, à laquelle on ajoute (C<sub>1</sub>) et (C<sub>2</sub>) concernant respectivement la continuité en probabilité du flot  $x \mapsto (X_t^x)_{t \in [0, T]}$  et la convergence du schéma de discrétisation vers le "vrai" processus  $(X_t)$ .

(C<sub>1</sub>) : Pour tout  $x_0 \in \mathbb{R}^d$ ,  $\epsilon > 0$  et  $T > 0$ ,

$$\limsup_{x_0 \rightarrow x} \mathbb{P}(\sup_{0 \leq t \leq T} |X_t^x - X_t^{x_0}| \geq \epsilon) = 0. \quad (3.11)$$

(C<sub>2</sub>) :  $(\bar{X}_t)$  est un processus de Markov non homogène et pour tout  $n \geq 0$ , on peut construire une famille de processus stochastiques  $(Y_t^{(n,x)})_{x \in \mathbb{R}^d}$  telle que

i)  $\mathcal{L}(Y^{(n,x)}) \stackrel{\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)}{=} \mathcal{L}(\bar{X}^{(\Gamma_n)} | \bar{X}_0^{(\Gamma_n)} = x)$ .

ii) Pour tout compact  $K$  de  $\mathbb{R}^d$ , pour tout  $T \geq 0$ ,

$$\sup_{x \in K} \sup_{0 \leq t \leq T} |Y_t^{(n,x)} - X_t^x| \xrightarrow{n \rightarrow +\infty} 0 \quad \text{en probabilité.} \quad (3.12)$$

**Théorème 3.8 ([76])** *Supposons (C<sub>0</sub>), (C<sub>1</sub>), (C<sub>2</sub>) satisfaites et  $(\eta_n/\gamma_n)_{n \geq 1}$  décroissante. Alors, p.s., pour toute fonctionnelle  $F : \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d) \rightarrow \mathbb{R}$ , bornée continue (pour la topologie de Skorokhod),*

$$\nu^{(n)}(\omega, F) \xrightarrow{n \rightarrow +\infty} \int F(\alpha) \mathbb{P}_\nu(d\alpha). \quad (3.13)$$

Le résultat ci-dessus admet des extensions aux fonctionnelles non bornées et aux cas où  $\nu$  n'est pas unique. De même, lorsque le processus limite est à trajectoires continues, le résultat est valable pour la topologie de la convergence uniforme sur les compacts (*c.u.c.*). Cependant, comme les schémas considérés ne sont pas nécessairement continus, on choisit dans ce cas de travailler sur l'espace  $\mathbb{D}_{uc}(\mathbb{R}_+, \mathbb{R}^d)$ , i.e. l'espace des fonctions càdlàg muni de la topologie *c.u.c.* En utilisant le résultat et les remarques précédentes, on obtient le corollaire suivant pour les diffusions :

**Corollaire 3.9** ([76]) Soit  $(X_t)_{t \geq 0}$  solution de (3.2). Supposons que  $b$  et  $\sigma$  soient localement lipschitziennes, que  $(\mathbf{S}_{a,\infty})$  soit satisfaite avec  $a \in (0, 1]$  et que  $(\eta_n/\gamma_n)_{n \geq 1}$  soit décroissante. (a) La suite  $(\bar{\nu}^{(n)}(\omega, d\alpha))_{n \geq 1}$  est p.s. tendue sur  $\mathbb{D}_{uc}(\mathbb{R}_+, \mathbb{R}^d)$  et toute limite faible de  $(\bar{\nu}^{(n)}(\omega, d\alpha))_{n \geq 1}$  est p.s. égale à la loi d'un processus solution de (3.2) en régime stationnaire. En cas d'unicité de  $\nu$ , (3.13) est valable pour toute fonctionnelle continue bornée  $F : \mathbb{D}_{uc}(\mathbb{R}_+, \mathbb{R}^d) \rightarrow \mathbb{R}$ .

(b) (Fonctionnelles non bornées) De plus, s'il existe  $s \in (2, +\infty)$  et  $n_0 \in \mathbb{N}$  tels que

$$\left( \frac{N(k, T) - N(k-1, T)}{\gamma_k H_k^s} \right)_{n \geq n_0} \text{ est décroissante et } \sum_{k \geq 1} \frac{N(k, T) - N(k-1, T)}{H_k^s} < +\infty, \quad (3.14)$$

alors, pour tout  $T > 0$ , pour toute fonctionnelle  $F : \mathbb{D}_{uc}(\mathbb{R}_+, \mathbb{R}^d) \rightarrow \mathbb{R}$  continue, (3.13) est vraie pour  $F_T$  définie pour tout  $\alpha \in \mathbb{D}_{uc}(\mathbb{R}_+, \mathbb{R}^d)$  par  $F_T(\alpha) = F(\alpha(t \wedge T))$  si :

$$\exists r > 0 \text{ tel que } |F_T(\alpha)| \leq C \sup_{0 \leq t \leq T} V^r(\alpha_t) \quad \forall \alpha \in \mathbb{D}_{uc}(\mathbb{R}_+, \mathbb{R}^d).$$

**Remarque 3.10** On constate qu'une extension aux fonctionnelles non bornées est possible lorsque celles-ci ne dépendent de la trajectoire que sur un intervalle borné et qu'une condition supplémentaire sur les pas et les poids est satisfaite. Lorsque ces derniers sont polynômiaux, i.e. si  $\eta_n = C_1 n^{-\rho_1}$  et  $\gamma_n = C_2 n^{-\rho_2}$ , et que  $0 < \rho_2 \leq \rho_1 \leq 1$  (afin que  $(\eta_n/\gamma_n)_{n \geq 1}$  soit décroissante) alors (3.14) est satisfaite dès que  $\rho_1 < 1$  (ce qui constitue donc une très faible restriction).

### 3.2.2 Vitesse de convergence de $(\bar{\nu}^{(n)}(\omega, d\alpha))_{n \geq 1}$

Cette question est abordée uniquement dans le cadre des diffusions (bien qu'elle puisse s'étendre assez naturellement à des cadres plus généraux). On suppose donc que  $(X_t)_{t \geq 0}$  est solution de (3.2) et qu'il admet une unique probabilité invariante notée  $\nu$ . De plus, on supposera que  $\forall n \geq 1, \eta_n = \gamma_n$ .

Comme dans le cas marginal, l'idée est d'établir la vitesse de convergence de  $(\bar{\nu}^{(n)}(\omega, F))_{n \geq 1}$  vers  $\mathbb{P}_\nu(F)$  pour une certaine classe de fonctionnelles. On s'intéressera uniquement ici à des fonctionnelles  $F : \mathbb{D}_{uc}([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$ , i.e. dépendant seulement de la trajectoire entre 0 et  $T$  (où  $T > 0$ ). On fera l'hypothèse suivante sur  $F$

$(\mathbf{C}_F^1)$  :  $F : \mathbb{D}_{uc}([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$  est bornée et lipschitzienne.

Notons que  $F$  peut toujours être vue comme une application de  $\mathbb{D}_{uc}(\mathbb{R}_+, \mathbb{R}^d)$  vers  $\mathbb{R}$  en introduisant la fonctionnelle  $F_T : \mathbb{D}_{uc}(\mathbb{R}_+, \mathbb{R}^d) \rightarrow \mathbb{R}$  définie par  $F_T(\alpha) = F(\alpha^T)$  où  $\alpha^T(t) = \alpha(t \wedge T)$ . Avec un léger abus de notation, on choisit dans la suite d'identifier  $F$  et  $F_T$ . On note donc

$$f_F(x) = \mathbb{E}[F(X^x)] = \mathbb{E}[F(X_t^x, 0 \leq t \leq T)],$$

fonction pour laquelle on a la propriété suivante :

$$\nu(f_F) = \mathbb{P}_\nu(F) = \int \mathbb{E}[F(X^x)] \nu(dx).$$

On introduit le schéma d'Euler à temps continu noté  $(\xi_t)_{t \geq 0}$  associé à  $(\bar{X}_{\Gamma_n})_{n \geq 0}$  défini par : pour tout  $n \in \mathbb{N}$ ,  $\xi_{\Gamma_n} = \bar{X}_{\Gamma_n}$  et

$$\forall t \in [\Gamma_n, \Gamma_{n+1}), \quad \xi_t = \bar{X}_{\Gamma_n} + (t - \Gamma_n)b(\bar{X}_{\Gamma_n}) + \sigma(\bar{X}_{\Gamma_n})(W_t - W_{\Gamma_{n+1}}). \quad (3.15)$$

Si l'on note

$$\underline{t} = \Gamma_{N(\underline{t})} \quad \text{avec} \quad N(\underline{t}) = \min\{n \geq 0, \Gamma_{n+1} > \underline{t}\}. \quad (3.16)$$

alors ce schéma possède la propriété (agréable) suivante :

$$\forall t \geq 0, \quad \xi_t = \xi_0 + \int_0^t b(\xi_{\underline{s}})ds + \int_0^t \sigma(\xi_{\underline{s}})dW_s. \quad (3.17)$$

Comme on le sait, le schéma ci-dessus ne peut pas toujours être exploité dans la pratique car  $t \mapsto \xi_t$  n'est pas simulable de manière exacte (dans certains cas tels que les fonctions du supremum, on réussit néanmoins à simuler  $F(\xi_t, t \in [0, T])$  de manière exacte en exploitant des propriétés théoriques sur le Brownien). Les résultats principaux seront donc aussi énoncés pour le schéma constant par morceaux noté simplement  $(\bar{X}_t)_{t \geq 0}$ .

Afin d'unifier les notations de mesures empiriques relatives à chaque schéma, on introduit pour un processus càdlàg  $Y = (Y_t)_{t \geq 0}$  donné la mesure  $\bar{\nu}^{(n)}(Y(\omega), d\alpha)$  définie par

$$\bar{\nu}^{(n)}(Y(\omega), d\alpha) = \frac{1}{\Gamma_n} \sum_{k=1}^n \delta_{Y(\Gamma_{k-1})}(d\alpha) = \frac{1}{\Gamma_n} \int_0^{\Gamma_n} \delta_{Y(\underline{s})} ds$$

où pour tout  $s$  et  $t$  positifs,  $Y_t^{(s)} = Y_{s+t}$ .

**Décomposition de l'erreur.** Si l'on veut exploiter les résultats de vitesse du cadre marginal, on est naturellement tenté de s'appuyer sur la fonction  $f_F$  définie plus haut. Si l'on considère le problème pour le vrai processus, l'idée première est alors d'effectuer l'opération suivante :

$$\frac{1}{t} \int_0^t F(X^{(s)}) - \mathbb{P}_\nu(F) ds = \frac{1}{t} \int_0^t F(X^{(s)}) - \mathbb{E}[F(X^{(s)}) | \mathcal{F}_s] ds + \frac{1}{t} \int_0^t f_F(X_s) - \nu(f_F) ds$$

où  $\mathcal{F}_t$  désigne l'augmentation usuelle de  $\sigma(W_s, 0 \leq s \leq t)$  par les ensembles  $\mathbb{P}$ -négligeables. Bien que le second terme soit satisfaisant, cette décomposition n'est en réalité pas exploitable (a priori) car le premier terme ne possède pas de propriété de martingale (problème d'adaptation).

La question est donc la suivante : comment peut-on exhiber "la" martingale? L'idée est en fait de découper  $[0, \Gamma_n]$  en intervalles de longueur  $T$  puis d'effectuer la décomposition suivante pour tout  $k \in \{1, \dots, [\Gamma_n/T]\}$  :

$$\begin{aligned} \int_{k-1T}^{kT} F_T(\xi^{(\underline{s})}) - \mathbb{P}_\nu(F) ds &= \int_{k-1T}^{kT} F(\xi^{(\underline{s})}) - \mathbb{E}[F(\xi^{(\underline{s})}) | \mathcal{F}_{kT}] ds \\ &+ \int_{k-1T}^{kT} \mathbb{E}[F(\xi^{(\underline{s})}) | \mathcal{F}_{kT}] - \mathbb{E}[F(\xi^{(\underline{s})}) | \mathcal{F}_{\underline{s}}] ds \\ &+ \mathbb{E}[F(\xi^{(\underline{s})}) | \mathcal{F}_{\underline{s}}] - f_F(\xi_{\underline{s}}) ds \\ &+ \int_{k-1T}^{kT} f_F(\xi_{\underline{s}}) - \nu(f_F) ds. \end{aligned} \quad (3.18)$$

Les deux premiers termes sont des accroissements de martingale relativement à  $\mathcal{F}_{(k+1)T}$  et  $\mathcal{F}_{kT}$  respectivement. Le troisième est un terme d'erreur faible fonctionnelle tandis que le dernier est de type "marginal". Cette décomposition (ou plus exactement une variante de

cette dernière) permettent de lire le problème de manière similaire au cadre marginal en décomposant à nouveau l'erreur en une partie martingales et une partie erreur faible. Le second point à noter est qu'en sommant sur  $k$  la décomposition ci-dessus et en associant correctement les termes de type accroissement de martingale, on va construire une martingale relativement à la filtration  $(\mathcal{F}_{kT})_{k \geq 0}$ . Cette dernière est en fait la "bonne" filtration du problème.

Il nous faut maintenant introduire les hypothèses permettant de gérer le comportement asymptotique de chacun des termes.

**Hypothèses et résultats.** En accord avec le cadre marginal, il nous est d'abord nécessaire que  $f_F$  admette une solution à l'équation de Poisson :

$(C_F^2)$  : Il existe une fonction de classe  $\mathcal{C}^2$  notée  $g_F : \mathbb{R}^d \rightarrow \mathbb{R}$  à dérivées bornées et lipschitziennes telle que

$$\forall x \in \mathbb{R}^d, \quad f_F(x) - \nu(f_F) = \mathcal{A}g_F$$

Comme le suggère la décomposition (3.18) et les remarques qui suivent, on va être amené lors de l'étude du comportement asymptotique de la martingale du problème à considérer une nouvelle suite de mesures empiriques  $(\mu_n(\omega, dx))_{n \geq 1}$  définie par :

$$\forall n \geq 1, \quad \mu_n(\omega, dx) = \frac{1}{n} \sum_{k=1}^n \delta_{\xi_{(k-1)T}}.$$

$(\mu_n(\omega, dx))_{n \geq 1}$  est tendue sous l'hypothèse de Lyapounov  $(S_{a,p})$  ( $p > 2$ ) introduite dans la section précédente mais converge sous l'hypothèse d'unicité de mesure invariante suivante :

$(S_T^\nu)$  :  $\nu$  est l'unique probabilité invariante pour  $P_T$ .

Cette hypothèse implique l'unicité pour  $(P_t)_{t \geq 0}$  mais la réciproque n'est pas vraie. Cependant, l'unicité du premier cadre est souvent assurée via celle du deuxième. On peut penser aux critères classiques suivants (qui assurent l'unicité de  $P_T$ ) :

- Irréductibilité (basée sur l'(hypo)-ellipticité) : pour tout  $x \in \mathbb{R}^d$ ,  $P_T(x, dy)$  a une densité  $(p_T(x, y))_{y \in \mathbb{R}^d}$  par rapport à la mesure de Lebesgue  $\lambda_d$  et  $\lambda_d(dy) - p.s., p_T(x, y) > 0$  pour tout  $x \in \mathbb{R}^d$ .

- Confluence asymptotique\* : pour toute fonction lipschitzienne  $f$ , pour tout compact  $K$  de  $\mathbb{R}^d$ ,

$$\sup_{(x_1, x_2) \in K} |P_{kT}f(x_1) - P_{kT}f(x_2)| \xrightarrow{k \rightarrow +\infty} 0 \quad (\text{see e.g. [4]}).$$

Nous sommes maintenant en mesure d'énoncer le résultat principal :

**Théorème 3.11 ([78])** Soit  $T > 0$ . Supposons que  $(S_T^\nu)$  soit satisfaite, que  $b$  et  $\sigma$  soient des fonctions lipschitziennes satisfaisant  $(S_{a,p})$  pour  $V \in \mathcal{E}\mathcal{Q}(\mathbb{R}^d)$ ,  $a \in (0, 1]$  et  $p > 2$  et que de plus,  $V$  vérifie :

$$\liminf_{|x| \rightarrow +\infty} \frac{V^{p+a-1}(x)}{|x|} > 0. \quad (3.19)$$

Supposons enfin que  $(\gamma_n)_{n \geq 1}$  satisfasse

$$\sum_{k \geq 1} \frac{\gamma_k^{3/2}}{\sqrt{\Gamma_k}} < +\infty \quad (3.20)$$

---

\*. on reviendra sur la seconde notion dans la section 3.4

et considérons  $F : \mathbb{D}_{uc}([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$  vérifiant  $(\mathbf{C}_F^1)$  et  $(\mathbf{C}_F^2)$ . Alors,

(a)

$$\sqrt{\Gamma_n} \left( \bar{\nu}^{(n)}(\xi(\omega), F) - \mathbb{P}_\nu(F) \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma_F^2), \quad (3.21)$$

où

$$\sigma_F^2 = \frac{1}{T} \left( \int \mathbb{E} \left[ \left( \mathbb{E}(A_{2T}^x | \mathcal{F}_{2T}) - \mathbb{E}(A_T^x | \mathcal{F}_T) - \int_T^{2T} \sigma^* \nabla g_F(X_u^x) dW_u \right)^2 \right] \nu(dx) \right) \quad (3.22)$$

et  $A_t^x := \int_0^t (F(X_{u+}^x) - f_F(X_u^x)) du$ ,  $t \geq 0$ , (est  $\mathcal{F}_{t+T}$ -mesurable).

(b) schéma constant par morceaux : Si de plus, il existe  $\delta > 0$  tel que

$$\sum_{k \geq 1} \frac{\gamma_k^{\frac{3}{2}-\delta}}{\sqrt{\Gamma_k}} < +\infty, \quad (3.23)$$

alors,

$$\sqrt{\Gamma_n} \left( \bar{\nu}^{(n)}(\bar{X}(\omega), F) - \mathbb{P}_\nu(F) \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma_F^2). \quad (3.24)$$

**Remarque 3.12** Lorsque  $\gamma_n \propto n^{-\rho}$ ,  $0 < \rho < 1$ ,  $\sqrt{\Gamma_n} \propto n^{\frac{1-\rho}{2}}$  (voir Remarque 3.7). Comme dans le cas marginal, le choix de  $\rho$  est naturellement contraint par l'erreur de discrétisation qui apparaît au travers des conditions (3.20) ou (3.23). Ainsi,  $\rho$  doit être choisi strictement supérieur à  $1/2$ , ce qui mène à une vitesse "optimale" de l'ordre de  $n^{-\frac{1}{4}+\varepsilon}$  pour tout  $\varepsilon > 0$ . En particulier, on ne retrouve pas la vitesse optimale marginale en  $n^{-1/3}$ . Ceci est dû à l'erreur faible fonctionnelle qui est généralement plus forte que l'erreur faible marginale. Plus exactement, nous utilisons ici que pour un schéma d'Euler de pas  $\gamma$ , la première est contrôlée par  $C\sqrt{\gamma}$  (où la constante dépend en particulier du point initial) tandis que la seconde est de l'ordre de  $\gamma$  (voir [93]). Ces résultats peuvent en particulier être améliorés sous de meilleures estimations de l'erreur faible fonctionnelle. On renvoie au début de la section 3.3.2 pour une discussion sur ce sujet.

**Remarque 3.13** Sous cette forme, la variance a une forme relativement peu lisible. À l'aide d'une série de calculs exploitant la stationnarité du processus sous  $\mathbb{P}_\nu$ , on peut obtenir d'autres formulations pour  $\sigma_F^2$ . En particulier (voir [78], annexe A pour plus de détails),  $\sigma_F^2$  peut se réécrire sous la forme :

$$\sigma_F^2 = 2 \int_0^T \left(1 - \frac{v}{T}\right) C_F(v) dv - 2\mathbb{E}_\nu \left( F(X) \int_0^T \sigma^* \nabla g_F(X_u) dW_u \right) + \int_{\mathbb{R}^d} |\sigma^* \nabla g_F(x)|^2 \nu(dx), \quad (3.25)$$

où  $\mathbb{E}_\nu$  est l'espérance relative à  $\mathbb{P}_\nu$  et  $C_F$  est une fonction de type covariance définie par

$$C_F(u) = \mathbb{E}_\nu \left( (F(X_{u+}) - f_F(X_u))(F(X) - f_F(X_0)) \right). \quad (3.26)$$

Contrairement à (3.22), l'expression (3.25) n'est pas clairement positive mais a l'avantage de bien séparer la partie "marginale" de la variance de sa partie "fonctionnelle" (qui correspond aux deux premiers termes).

Par exemple, si  $F(\alpha) = \phi(\alpha(0))$  où  $\phi$  est une fonction bornée telle que  $\phi - \nu(\phi) = \mathcal{A}h$  (où  $h$  est une fonction bornée  $\mathcal{C}^2$  de dérivées bornées), alors  $f_F = \phi$  et on constate bien que



les deux premiers termes de (3.25) sont égaux à 0. Ainsi,  $\sigma_F^2 = \int_{\mathbb{R}^d} |\sigma^* \nabla g_F(x)|^2 \nu(dx)$  (avec  $g_F = h$ , ce qui correspond bien à la variance du *TCL* marginal).

Considérons enfin le cas où  $F_T$  est définie par  $F(\alpha) = \phi(\alpha(T))$  (avec  $\phi$  satisfaisant les mêmes hypothèses que précédemment). Dans ce cadre, on peut facilement d'un changement de variable (au moins dans le cas du vrai processus ci-après) que la variance limite est bien la même que dans le cas précédent, *i.e.*  $\int_{\mathbb{R}^d} |\sigma^* \nabla h(x)|^2 \nu(dx)$ . Retrouver ce résultat via l'une ou l'autre formulation des variances fonctionnelles est possible mais nécessite des calculs non triviaux (voir [78] annexe B). Ceci souligne le caractère complexe de la structure de la variance fonctionnelle.

On souhaite aussi souligner le fait que ce *TCL* fonctionnel est en particulier vrai pour le schéma "parfait", *i.e.* le processus lui-même. Contrairement au cas marginal, il semble que ce résultat soit nouveau :

**Théorème 3.14 ([78])** *Soit  $T > 0$ . Supposons que  $b$  et  $\sigma$  soient des fonctions lipschitziennes satisfaisant  $(S_{a,p})$  pour  $V \in \mathcal{EQ}(\mathbb{R}^d)$ ,  $a \in (0, 1]$  et  $p > 2$ . Supposons aussi que  $(S_T^\nu)$  soit satisfaite. Soit  $F : \mathbb{D}_{uc}([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$  une fonctionnelle satisfaisant  $(C_F^1)$  and  $(C_F^2)$ . Alors, pour tout  $x \in \mathbb{R}^d$ ,*

$$\sqrt{t} \left( \frac{1}{t} \int_0^t F(X_u^{(s),x}, 0 \leq u \leq T) ds - \mathbb{P}_\nu(F) \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma_F^2). \quad (3.27)$$

**Exemple d'application.** On propose de terminer cette section par une application à la valorisation de prix d'options "trajectoire-dépendantes" (*i.e.* dépendant de l'évolution du processus entre 0 et  $T$ ) dans un modèle à volatilité stochastique stationnaire. Supposer la volatilité stationnaire est une hypothèse assez peu usuelle. Pourtant, la volatilité est généralement supposée avoir des propriétés ergodiques et est aussi une quantité non observée a priori sur le marché. Il est donc naturel de supposer qu'elle évolue dans son régime stationnaire. Dans ce cadre, on peut alors, comme on va le voir dans la suite, utiliser la procédure étudiée ci-dessus pour valoriser des prix d'options.

Décrivons les choses plus précisément dans le cas du modèle de Heston à volatilité stationnaire. On rappelle que la dynamique de ce modèle est donnée par :

$$\begin{aligned} dS_t &= S_t(rdt + \sqrt{(1-\rho^2)}v_t dW_t^1 + \rho\sqrt{v_t}dW_t^2), \quad S_0 = s_0 > 0, \\ dv_t &= k(\theta - v_t)dt + \varsigma\sqrt{v_t}dW_t^2, \quad v_0 > 0, \end{aligned}$$

où  $r$  est le taux d'intérêt,  $(W^1, W^2)$  un brownien standard en dimension 2,  $\rho \in [-1, 1]$  est un coefficient de corrélation et  $k, \theta$  et  $\varsigma$  sont des entiers positifs. Lorsque  $2k\theta > \varsigma^2$  alors,  $(v_t)$  est un processus *p.s.* strictement positif admettant une unique probabilité invariante  $\nu$  de loi  $\Gamma(a, b)$  avec  $a = (2k)/\varsigma^2$  et  $b = (2k\theta)/\varsigma^2$ .

Lorsque  $(v_t)_{t \geq 0}$  évolue sous son régime stationnaire, alors pour une fonctionnelle  $F : \mathbb{D}_{uc}([0, T], \mathbb{R}) \rightarrow \mathbb{R}$ , on peut (artificiellement) exprimer  $F(S_t, 0 \leq t \leq T)$  comme une fonctionnelle d'un processus stationnaire. En effet, via la formule d'Itô,  $(S_t)$  satisfait

$$\forall t \geq 0, \quad S_t = s_0 \exp \left( rt - \frac{1}{2} \int_0^t v_s ds + \rho \int_0^t \sqrt{v_s} dW_s^2 + \sqrt{1-\rho^2} \int_0^t \sqrt{v_s} dW_s^1 \right). \quad (3.28)$$

Ainsi, en introduisant l'EDS,

$$\begin{cases} dy_t = -y_t dt + \sqrt{v_t} dW_t^1, \\ dv_t = k(\theta - v_t) dt + \varsigma \sqrt{v_t} dW_t^2, \end{cases} \quad (3.29)$$

et en utilisant que

$$\int_0^t \sqrt{v_s} dW_s^1 = y_t - y_0 + \int_0^t y_s ds \quad \text{et} \quad \int_0^t \sqrt{v_s} dW_s^2 = \frac{v_t - v_0 - k\theta t + k \int_0^t v_s ds}{\varsigma},$$

on en déduit qu'on peut construire  $\Phi$  de  $\mathbb{D}_{uc}(\mathbb{R}_+, \mathbb{R}^2)$  dans  $\mathbb{D}_{uc}(\mathbb{R}_+, \mathbb{R})$  telle que  $(S_t)_{t \geq 0} = \Phi((y_t, v_t)_{t \geq 0})$ . En notant  $\mu$  l'unique loi invariante du couple  $(y_t, v_t)_{t \geq 0}$ , on obtient

$$\mathbb{E}[F(S_t, 0 \leq t \leq T)] = \mathbb{E}_\mu[F \circ \Phi((y_t, v_t), 0 \leq t \leq T)],$$

ce qui permet alors de voir le problème de la valorisation d'options comme un calcul d'espérance sous régime stationnaire. Remarquons cependant qu'une application rigoureuse de la procédure étudiée précédemment nécessite quelques extensions théoriques que l'on ne fera pas ici mais que l'on souhaite néanmoins mentionner : les fonctionnelles considérées ici ne satisfont clairement pas l'hypothèse  $(C_F^1)$  tandis que la volatilité du modèle de Heston nécessite une discrétisation spécifique (on renvoie à [76, 78] pour plus détails).

On termine cette partie par un test numérique de la normalité asymptotique obtenue dans le théorème 3.11 lorsque la fonctionnelle considérée est relative au prix d'une option barrière (*Up and Out*) :

$$F(S_t, 0 \leq t \leq T) = e^{-rT} (S_T - K)_+ \mathbf{1}_{\{\sup_{0 \leq t \leq T} S_t \leq L\}} \quad (L > K > 0).$$

On souhaite donc identifier numériquement la loi de l'erreur normalisée

$$\mathcal{E}_N := \sqrt{\Gamma_N} \left( \bar{\nu}_N(\xi(\omega), F) - e^{-rT} \mathbb{E}[(S_T - K)_+ \mathbf{1}_{\{\sup_{0 \leq t \leq T} S_t \leq L\}}] \right). \quad (3.30)$$

Sans détailler la discrétisation, rappelons que pour ce type de fonctionnelles, on peut mettre en oeuvre le schéma noté  $(\xi_t)_{t \geq 0}$  en s'appuyant sur le fait que la loi de ce dernier conditionnée à  $(\xi_{\Gamma_n})_{n \geq 0}$  est explicite et simulable (suite de ponts Browniens indépendants, voir [78] pour plus de détails). Pour les paramètres suivants,

$$s_0 = 50, \quad r = 0.05, \quad T = 1, \quad \rho = 0.5, \quad \theta = 0.01, \quad \varsigma = 0.1, \quad k = 2, \quad K = 50, \quad L = 55, \quad (3.31)$$

une simulation longue donne l'approximation précise suivante :

$$e^{-rT} \mathbb{E}[(S_T - K)_+ \mathbf{1}_{\{\sup_{0 \leq t \leq T} S_t \leq L\}}] \approx 1,689.$$

En la réinjectant dans (3.30), on obtient pour  $N = 5 \cdot 10^5$  (points de discrétisation) et  $MC = 5 \cdot 10^3$  une approximation de la densité de  $\mathcal{E}_N$  (obtenue par convolution par un noyau gaussien) décrite dans la figure 3.1 (où  $\bar{\sigma}_F^2$  est la variance empirique obtenue) : Notons pour terminer que, outre l'illustration du *TCL* (à une vitesse raisonnable), ce test numérique suggère qu'une version locale (du *TCL*) pourrait sans doute être aussi démontrée (au sens "convergence de densité").

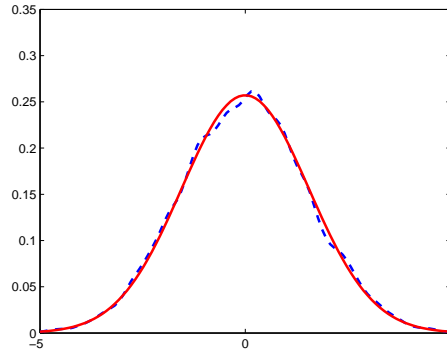


Figure 3.1 : Comparaison de la densité approchée de  $\mathcal{E}_N$  (en pointillé) avec celle de  $\mathcal{N}(0, \bar{\sigma}_F^2)$ ,  $N = 5 \cdot 10^5$ ,  $MC = 5 \cdot 10^3$ .

### 3.3 Un algorithme d'approximation du régime stationnaire moins moyennisant

Comme on l'a déjà mentionné, les suites de mesures empiriques étudiées précédemment sont (au moins dans le cas  $\eta_n = \gamma_n$ ) construites en mimant la mesure d'occupation du processus. En particulier, il s'agit d'une moyennisation de toute la trajectoire discrétisée, ce qui implique de simuler  $f(\bar{X}_{\Gamma_n})$  ou  $F(\bar{X}^{(\Gamma_n)})$  à chaque pas de temps. Or, si le coût global de calcul des quantités précédentes est élevé, la simulation globale de  $\bar{\nu}_n(\omega, f)$  ou de  $\bar{\nu}^{(n)}(\omega, F)$  s'en trouve lourdement impactée. Ce problème est très clair dans le second cas (fonctionnel) où, comme on l'a déjà vu, préserver la linéarité de la complexité de l'algorithme n'est faisable qu'à l'aide d'une simulation récursive de  $F(\bar{X}^{(\Gamma_n)})$  et sous conditions raisonnables sur la forme de la fonctionnelle. Ceci nous amène à une question naturelle : afin de limiter le nombre de calculs de  $f(\bar{X}_{\Gamma_n})$  ou  $F(\bar{X}^{(\Gamma_n)})$ , ne pourrait-on pas (au moins dans le cas fonctionnel) considérer de procédures moins moyennisantes, *i.e.* des suites de mesures empiriques construites comme mesures d'occupation relatives à des sous-suites de  $(\bar{X}_{\Gamma_n})_{n \geq 0}$  (ou de  $(\bar{X}^{(\Gamma_n)})_{n \geq 0}$ ) ?

En accord avec ce qui précède, on propose donc dans cette partie d'étudier des suites de mesures empiriques basées sur la moyennisation des valeurs du schéma prises à des instants (absolus)  $0, T, \dots, kT, \dots$  où  $T$  est un réel strictement positif. On considérera ici le problème uniquement pour la diffusion  $(X_t)_{t \geq 0}$  solution de (3.2) et pour le schéma  $(\xi_t)_{t \geq 0}$  défini par (3.15) (ce qui bien sûr ne change rien dans le cadre marginal). On construit alors un algorithme dit *mixed-step* et défini de la manière suivante : on fixe  $T > 0$  et on note respectivement  $(\bar{\mu}_n(\omega, dx))_{n \geq 1}$  et  $(\bar{\mu}^{(n)}(\omega, d\alpha))_{n \geq 1}$  les suites de mesures empiriques définies respectivement sur  $\mathbb{R}^d$  et  $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$  par :

$$\bar{\mu}_n(\omega, dx) = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{\xi_{kT}(\omega)}(dx), \quad n \geq 1,$$

$$\bar{\mu}^{(n)}(\omega, d\alpha) = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{\xi_{(kT)}(\omega)}(d\alpha), \quad n \geq 1,$$

où l'on rappelle que  $\xi^{(t)}$  est le schéma d'Euler à temps continu shifté de  $t$  :  $\forall s, t$  positifs,  $\xi_s^{(t)} = \xi_{t+s}$ .

Bien que plus délicates à obtenir, les propriétés de convergence relatives à ces suites sont respectivement similaires aux proposition 3.3 et théorème 3.8. La seule différence notable apparaît dans l'hypothèse d'unicité de la mesure invariante pour  $(P_t)_{t \geq 0}$  qui est à remplacer ici par l'hypothèse légèrement plus forte  $(S'_T)$  introduite plus haut (puisque dans ce cadre, les mesures limites sont naturellement des mesures invariantes pour  $P_T$  uniquement). Nous choisissons donc de ne pas rappeler ces résultats ici et de nous focaliser sur la vitesse de convergence de ces procédures, ou plus exactement de  $(\bar{\mu}_n(\omega, f))_{n \geq 1}$  et  $(\bar{\mu}^{(n)}(\omega, F))_{n \geq 1}$  vers  $\nu(f)$  et  $\mathbb{P}_\nu(F)$  respectivement.

**Remarque 3.15** Dans la partie convergence (non détaillée), le choix de  $T$  est complètement arbitraire. Dans la partie vitesse, le choix de  $T$  restera également arbitraire dans le cas marginal (au moins dans les résultats généraux). En revanche, dans le cas fonctionnel,  $T$  correspondra (principalement pour des raisons techniques) à la longueur de la trajectoire impliquée dans la fonctionnelle considérée. En d'autres termes, on s'intéressera au cas où la moyennisation est "minimale". Notons que dans ce cas, il n'y a aucun "chevauchement" entre les trajectoires impliquées dans le calcul de  $F(\xi^{(kT)})$  et  $F(\xi^{((k+1)T)})$  (pour tout  $k$ ), et que la linéarité de la complexité de l'algorithme est bien assurée.

### 3.3.1 Vitesse de $(\bar{\mu}_n(\omega, f))_{n \geq 1}$

On suppose dans la suite que  $T$  est un réel strictement positif et que l'hypothèse  $(S'_T)$  est satisfaite. On rappelle que l'on suppose la convergence *p.s.* de  $(\bar{\mu}_n(\omega, dx))_{n \geq 1}$  vers  $\nu$  établie et on s'intéresse maintenant pour une fonction  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  donnée, à la vitesse de convergence de  $\bar{\mu}_n(\omega, f)$  vers  $\nu(f)$ . À l'image de la procédure initiale, on a aussi besoin ici d'une hypothèse de type "existence de solution à l'équation de Poisson". Si l'on se réfère à l'heuristique donnée dans (3.7), on comprend alors assez bien qu'on aura ici besoin d'une version non infinitésimale de cette équation. Plus précisément, on travaillera dans la suite sous l'hypothèse :

$(\mathcal{P}_{m,T})$  : Il existe une fonction de classe  $C^m$  notée  $g_T : \mathbb{R}^d \rightarrow \mathbb{R}$  telle que

$$f(x) - \nu(f) = \frac{g_T(x) - P_T g_T(x)}{T}, \quad x \in \mathbb{R}^d. \quad (3.32)$$

Si de plus,  $g_T$  et ses dérivées (partielles d'ordre inférieur ou égal à  $m$ ) sont bornées, on dira que  $f$  satisfait  $(\mathcal{P}_{m,T}^b)$ . De même, si  $g_T$  et ses dérivées ont une croissance sous-polynômiale, on notera l'hypothèse précisée  $(\mathcal{P}_{m,T}^{\text{pol}})$ . On reviendra sur ces hypothèses à la fin de cette section.

Sous ce type d'hypothèse, on peut décomposer  $\bar{\mu}_n(\omega, f) - \nu(f)$  sous la forme suivante :

$$\bar{\mu}_n(\omega, f) - \nu(f) = \frac{1}{nT} \sum_{k=1}^n g_T(\xi_{(k-1)T}) - P_T g_T(\xi_{(k-1)T}) \quad (3.33)$$

$$= \frac{1}{nT} (g_T(\xi_0) - P_T g_T(\xi_{nT})) \quad (3.34)$$

$$+ \frac{1}{nT} \sum_{k=1}^{n-1} g_T(\xi_{kT}) - \mathbb{E}[g_T(\xi_{kT}) | \mathcal{F}_{(k-1)T}] \quad (3.35)$$

$$+ \frac{1}{nT} \sum_{k=1}^{n-1} \mathbb{E}[g_T(\xi_{kT}) | \mathcal{F}_{(k-1)T}] - P_T g_T(\xi_{(k-1)T}). \quad (3.36)$$

On constate que pour cette procédure les parties martingales et erreur faible sont très clairement séparées correspondant respectivement aux termes (3.35) et (3.36). En particulier, contrairement à l'algorithme originel, on aura ici besoin de "vrais" développements de l'erreur faible (mais dans un cadre où la suite de pas est variable). Avant de donner le résultat principal (théorème 3.18), on propose de rappeler quelques notions concernant cette thématique et d'expliciter quelques résultats dans ce cadre à pas non constant.

**À propos de l'erreur faible.** On s'intéresse ici au terme d'erreur faible (3.36). D'abord, on remarque que pour tout  $n \geq 1$ ,

$$\mathbb{E}[g_T(\xi_{nT}) | \mathcal{F}_{(n-1)T}] - P_T g_T(\xi_{(n-1)T}) = \mathcal{E}_T(g, x, \gamma^{(n-1)})$$

où  $\gamma^{(n)} := (\gamma_k^{(n)})_{k \geq 1}$  vérifie

$$\gamma_1^{(n)} = \Gamma_{N(nT)+1} - nT, \quad \gamma_k^{(n)} = \gamma_{N(nT)+k} \quad k \in \{2, \dots, N((n+1)T) - N(nT) - 1\}$$

et  $\gamma_{N((n+1)T) - N(nT) + 1}^{(n)} = (n+1)T - \Gamma_{N((n+1)T)}$ , et pour une suite de pas discrétisation  $\mathbf{h} = (h_k)_{k \geq 1}$  donnée,

$$\mathcal{E}_T(g, x, \mathbf{h}) = \mathbb{E}[g(\xi_T^{x, \mathbf{h}})] - \mathbb{E}[g(X_T^x)] \quad (3.37)$$

où  $(\xi_t^{x, \mathbf{h}})_{t \geq 0}$  le schéma d'Euler construit comme (3.15) avec valeur initiale  $x$  et suite de pas discrétisation  $\mathbf{h} = (h_k)_{k \geq 1}$ . Lorsque que l'on cherche à développer  $\mathcal{E}_T(g, x, \mathbf{h})$ , une idée de départ (voir [93]) est de décomposer cette quantité de la manière suivante :

$$\mathcal{E}_T(g, x, \mathbf{h}) = \sum_{k=1}^{k_T} \mathbb{E}[u_g(t_k, \xi_{t_k}^{x, \mathbf{h}}) - u_g(t_{k-1}, \xi_{t_{k-1}}^{x, \mathbf{h}})]. \quad (3.38)$$

où  $u_g(t, x) := \mathbb{E}[g(X_{T-t}^x)]$  est solution de

$$\begin{cases} \partial_t u_g(t, x) + \mathcal{A}u_g(t, x) = 0 & (t, x) \in [0, T] \times \mathbb{R}^d, \\ u_g(T, x) = g(x), & x \in \mathbb{R}^d. \end{cases} \quad (3.39)$$

Le développement à l'ordre 1, s'obtient ensuite à l'aide d'un développement d'Itô de chaque terme de la somme apparaissant dans (3.38).

Dans la proposition ci-dessous, on propose de rassembler quelques propriétés concernant  $u_g$  et le développement de  $\mathcal{E}_T(g, x, \gamma^{(n)})$ , qui joueront un rôle assez important dans la suite.

**Proposition 3.16** Soient  $m \in \mathbb{N}$  et  $\rho \in (0, 1]$  tels que  $b$  et  $\sigma$  sont de classe  $C^{m, \rho \dagger}$  sur  $\mathbb{R}^d$  à dérivées d'ordre  $\ell \leq m$  bornées. Soit  $g$  une fonction de classe  $C^m$  sur  $\mathbb{R}^d$  telle que  $g$  et ses dérivées d'ordre  $\ell \leq m$  sont à croissance polynômiale. Alors,

(i)  $x \mapsto u_g(t, x)$  est de classe  $C^m$  sur  $\mathbb{R}^d$  et  $t \mapsto u_g(t, x)$  est de classe  $C^{\lfloor m/2 \rfloor}$  sur  $\mathbb{R}_+$ . De plus, il existe  $r > 0$  tel que pour tout  $t \in [0, T]$ ,

$$|u_g(t, x)| \leq C_T(1 + |x|^r) \quad \text{et} \quad |\partial_{x_1^{\ell_1} \dots x_d^{\ell_d}}^\ell u_g(t, x)| \leq C_T(1 + |x|^r), \quad (3.40)$$

(ii) Si  $m = 4$ ,

$$|\mathcal{E}_T(g, x, \gamma^{(n)})| \leq C_T(1 + |x|^r) \sum_{k=1}^{k_T} (\gamma_k^{(n)})^2$$

(iii) Si  $m = 5$ , alors,

$$\mathcal{E}_T(g, x, \gamma^{(n)}) = \frac{1}{2} \gamma_{N(nT)} \int_0^T \mathbb{E}[\Phi_g(s, X_s^x)] ds + r_n(x)$$

où  $\Phi_g$  est définie par

$$\Phi_g(t, x) = \langle \mathcal{A}(\partial_x u_g)(t, x) - \partial_x(\mathcal{A}u_g)(t, x), b(x) \rangle + (\mathcal{A}(\partial_{x_2}^2 u_g)(t, x) - \partial_{x_2}^2(\mathcal{A}u_g)(t, x)) \sigma(x)^{\otimes 2} \quad (3.41)$$

et

$$|r_n(x)| \leq C_T(1 + |x|^r) \left( \gamma_{N(nT)}^{\frac{3}{2}} + (\gamma_{N(nT)} - \gamma_{N((n+1)T)}) \right).$$

(iv) Supposons que  $m = 9$  et que  $\Gamma_{N(nT)} = nT$  pour tout  $n \geq 1$ . Alors,

$$\mathcal{E}_T(g, x, \gamma^{(n)}) = \frac{1}{2} \gamma_{N(nT)} \int_0^T \mathbb{E}[\Phi_g(s, X_s^x)] ds + \gamma_{N(nT)}^2 \int_0^T \mathbb{E}[\chi_g(s, X_s^x)] ds + \tilde{\rho}_T(x, \gamma^{(n)}) \quad (3.42)$$

où  $\chi_g : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$  est une fonction continue (voir [79] pour une expression explicite) et

$$|\tilde{\rho}_T(x, \gamma^{(n)})| \leq C(1 + |x|^r) \left( \gamma_{N(nT)} - \gamma_{N((n+1)T)} + \gamma_{N(nT)}^{\frac{5}{2}} \right).$$

**Remarque 3.17** La décroissance de la suite  $(\gamma_n)_{n \geq 1}$  joue un rôle non négligeable dans l'obtention des développements d'ordre 1 et 2 (iii) et (iv). Le développement à l'ordre 2 obtenu dans (iv) ne jouera un rôle que dans la section 3.3.3 où une méthode de Romberg sera développée pour accélérer la vitesse de convergence. On remarque que pour ce développement, on ajoute l'hypothèse  $\Gamma_{N(nT)} = nT$  qui permet de tuer les termes de bord qui deviennent alors difficiles à gérer à l'ordre 2.

On peut maintenant énoncer un premier résultat principal :

**Théorème 3.18 ([79])** Supposons  $(\mathbf{S}_T^\nu)$  et  $(\mathbf{S}_{a,p})$ , avec  $a \in (0, 1]$  et  $p \in (2, +\infty]$ , satisfaites. Soit  $m \in \mathbb{N}$  et  $\rho \in (0, 1]$  tels que  $b$  et  $\sigma$  sont de classe  $C^{m, \rho}$  sur  $\mathbb{R}^d$  à dérivées partielles bornées (jusqu'à l'ordre  $m$ ). Soit  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  une fonction borélienne satisfaisant  $(\mathcal{P}_{m, \mathbf{T}}^b)$  si  $p < +\infty$  ou  $(\mathcal{P}_{m, \mathbf{T}}^{\text{pol}})$  si  $p = +\infty$ . Alors,

†. On rappelle que  $C^{m, \rho}$  désigne l'ensemble des fonctions de classe  $C^m$  dont les dérivées d'ordre  $m$  sont  $\rho$ -höldériennes.

(i) Si  $m = 4$  et  $\frac{1}{\sqrt{n}} \sum_{k=1}^n \gamma_{N(kT)} \xrightarrow{n \rightarrow +\infty} 0$ ,

$$\sqrt{nT} \left( \frac{1}{n} \sum_{k=1}^n f(\xi_{(k-1)T}) - \nu(f) \right) \xrightarrow{n \rightarrow +\infty} \mathcal{N}(0; \hat{\sigma}_T^2)$$

avec  $\hat{\sigma}_T^2 = \frac{1}{T} \int_{\mathbb{R}^d} (g_T^2(x) - (P_T g_T(x))^2) \nu(dx)$ .

(ii) Si  $m = 5$  et  $\frac{1}{\sqrt{n}} \sum_{k=1}^n \gamma_{N(kT)} \xrightarrow{n \rightarrow +\infty} +\infty$  et  $\gamma_{N(nT)} - \gamma_{N((n+1)T)} = o(\gamma_{N(nT)})$  lorsque  $n \rightarrow +\infty$ ,

$$\frac{n}{\sum_{k=1}^n \gamma_{N(kT)}} \left( \frac{1}{n} \sum_{k=1}^n f(\xi_{(k-1)T}) - \nu(f) \right) \xrightarrow{\mathbb{P}} m_T = \frac{1}{2T} \int_{\mathbb{R}^d} \int_0^T \mathbb{E}[\Phi_g(s, X_s^x)] ds \nu(dx)$$

lorsque  $n \rightarrow +\infty$ , où " $\xrightarrow{\mathbb{P}}$ " désigne la convergence en probabilité et  $\Phi_g$  est définie par (3.41).

(iii) Si  $m = 5$ ,  $\sqrt{\frac{T}{n}} \sum_{k=1}^n \gamma_{N(kT)} \xrightarrow{n \rightarrow +\infty} \beta_0 \in (0, +\infty)$  et  $\gamma_{N(nT)} - \gamma_{N((n+1)T)} = o(\gamma_{N(nT)})$  lorsque  $n \rightarrow +\infty$ , alors,

$$\sqrt{nT} \left( \frac{1}{n} \sum_{k=1}^n f(\xi_{(k-1)T}) - \nu(f) \right) \xrightarrow{n \rightarrow +\infty} \mathcal{N}(\beta_0 m_T; \hat{\sigma}_T^2).$$

**Comparaison avec  $(\bar{\nu}_n(\omega, dx))_{n \geq 1}$ .** Suite à ce résultat, on va maintenant tenter de comparer ce nouvel algorithme moins moyennisant à la suite  $(\bar{\nu}_n(\omega, dx))_{n \geq 1}$ . On considère successivement deux questions :

– l'ordre de la vitesse de convergence : supposons que  $\gamma_k = Ck^{-\rho}$ . On peut alors vérifier que  $N(t) = C^{-\frac{1}{1-\rho}} t^{\frac{1}{1-\rho}} + O(1)$ , puis que

$$\gamma_{N(nT)} = C^{\frac{1}{1-\rho}} T^{-\frac{\rho}{1-\rho}} n^{-\frac{\rho}{1-\rho}} + O(n^{-\frac{1+\rho}{1-\rho}}).$$

On en déduit que

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n \gamma_{N(kT)} \xrightarrow{n \rightarrow +\infty} 0 \iff \rho > \frac{1}{3}$$

et que  $\gamma_{N(nT)} - \gamma_{N((n+1)T)} = o(\gamma_{N(nT)})$  pour tout  $\rho \in (0, 1)$ .

À l'instant  $t_n = nT$ , l'erreur est donc de l'ordre de  $(t_n)^{-\frac{1}{2}}$  si  $\rho > 1/3$  et de l'ordre de  $(t_n)^{-\frac{\rho}{1-\rho}}$  si  $\rho \leq 1/3$ . D'un point de vue numérique, on doit cependant calculer l'erreur en fonction du nombre  $N$  d'instant de discrétisation. L'erreur est alors de l'ordre de

$$\begin{cases} \Gamma_N^{-\frac{1}{2}} \propto N^{-\frac{1-\rho}{2}} & \text{si } \rho \in (1/3, 1) \\ \Gamma_N^{-\frac{\rho}{1-\rho}} \propto N^{-\rho} & \text{si } \rho \in (0, 1/3]. \end{cases} \quad (3.43)$$

D'après la remarque 3.7, l'ordre de l'erreur est le même que pour la suite de mesures empiriques originelle  $(\bar{\nu}_N(\omega, dx))_{N \geq 1}$ . On retrouve en particulier la vitesse optimale en  $N^{-\frac{1}{3}}$ . Pour aller plus loin, il nous faut maintenant effectuer une

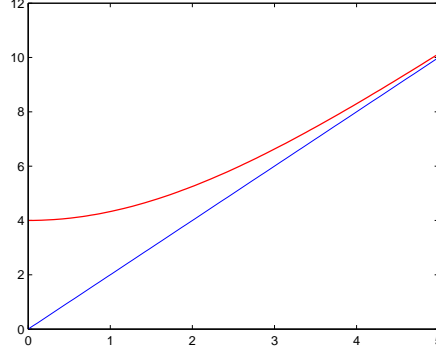


Figure 3.2 :  $T \mapsto \hat{\sigma}_T^2$  lorsque  $dX_t = -\frac{1}{2}X_t + dW_t$  et  $f(x) = x^2$

- *Comparaison des variances asymptotiques* : il s'agit donc de comparer  $\hat{\sigma}_T^2$  à  $\hat{\sigma}_0^2 = \int |\sigma^* \nabla g_0(x)|^2 \nu(dx)$  où  $g_0$  est solution de l'équation de Poisson :  $f - \nu(f) = -\mathcal{A}g_0$ . Commençons par un cas (très) particulier dans lequel on peut faire des calculs explicites : le processus d'Ornstein-Uhlenbeck  $dX_t = -\frac{1}{2}X_t + dW_t$  en dimension 1 et la fonction  $f(x) = x^2$ . Dans ce cas, on obtient

$$\hat{\sigma}_0^2 = 4 \quad \text{and } \forall T > 0, \quad \hat{\sigma}_T^2 = \frac{2T(1 + e^{-T})}{1 - e^{-T}} \geq 4.$$

$T \rightarrow \hat{\sigma}_T^2$  est continue croissante sur  $\mathbb{R}_+$  à croissance linéaire (voir Figure 3.2).

Dans un cadre général, certaines des propriétés précédentes peuvent être préservées. Si l'on suppose par exemple que le semi-groupe converge à vitesse exponentielle à l'équilibre, alors on peut montrer que pour toute fonction  $f$  de classe  $C^2$  telle que  $f$ ,  $\nabla f$  and  $D^2 f$  sont bornées, la propriété suivante est vraie (voir [79], Appendix B) :

$$\lim_{T \rightarrow 0} \hat{\sigma}_T^2 = \hat{\sigma}_0^2, \quad \lim_{T \rightarrow 0} \partial_T \hat{\sigma}_T^2 = 0 \quad \text{et} \quad \lim_{T \rightarrow +\infty} \partial_T \hat{\sigma}_T^2 = \int (f - \nu(f))^2 \nu(dx). \quad (3.44)$$

En conclusion, on observe bien un phénomène de croissance de la variance avec  $T$ , ce qui est statistiquement naturel (puisque l'on "jette des données"), mais que celle-ci est a priori raisonnable. La dérivée est nulle en 0 et la croissance asymptotique est linéaire et de vitesse donnée par la variance de  $f$  sous  $\nu$ . Comme cette dernière peut être approchée avec l'algorithme, on remarque que l'on pourrait envisager numériquement un choix de  $T$  adaptatif, fonction de l'estimation de la variance et du coût de calcul de  $f$ .

**À propos des solutions de l'équation de Poisson.** Avant de terminer cette section, on souhaite revenir sur l'hypothèse  $(\mathcal{P}_{m,T})$  et plus généralement sur la solvabilité de l'équation de Poisson sur  $\mathbb{R}^d$ , sur laquelle on s'appuie (classiquement) dans chaque résultat de vitesse de convergence. Il s'agit donc de compléter la remarque 3.5, notamment dans le cas non infinitésimal, qui est moins présent dans la littérature. On a d'abord le résultat général d'existence suivant :

**Lemme 3.19** *Supposons que  $(X_t)_{t \geq 0}$  soit un processus de Markov fellerien de semi-groupe  $(P_t)_{t \geq 0}$  tel que  $P_T$  admet une unique probabilité invariante  $\nu$ . Soit  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  une fonction continue satisfaisant*

$$\forall t \geq 0, \quad |P_t f(x) - \nu(f)| \leq C(x)\varphi(t)$$



avec  $x \mapsto C(x)$  localement bornée (positive) et  $\varphi(t) \rightarrow 0$  lorsque  $t \rightarrow +\infty$ . Alors,  
 (i) si  $\sum_{k \geq 1} \varphi(kT) < +\infty$ , la fonction  $g_T$  définie par

$$\forall x \in \mathbb{R}^d, \quad g_T(x) = \sum_{n \geq 0} T(P_{nT}f(x) - \nu(f))$$

est une solution continue de (3.32).

(ii) Supposons que pour tout  $x \in \mathbb{R}^d$ ,  $P_t f(x) \xrightarrow{t \rightarrow 0} f(x)$  (toujours vrai si  $f$  est continue bornée) et que  $\int_0^{+\infty} \varphi(s) ds < +\infty$ . La fonction  $g_0$  définie par

$$\forall x \in \mathbb{R}^d, \quad g_0(x) = \int_0^{+\infty} (P_s f(x) - \nu(f)) ds$$

est alors solution de  $\mathcal{A}g_0 = -(f - \nu(f))$ .

On constate donc que dès que le semi-groupe converge à vitesse intégrable vers l'équilibre, on a existence de solution à l'équation de Poisson (infinitésimale ou non). Au vu de la forme des solutions, on remarque également que d'un point de vue probabiliste, la régularité de  $x \mapsto g_T(x)$  ( $T \geq 0$ ) est liée à la vitesse de convergence des dérivées successives de l'application  $x \mapsto X_t^x$  vers 0. On propose ci-dessous un résultat dans un cas éventuellement dégénéré où le système dynamique a des propriétés de confluence asymptotique (forte) des trajectoires :

Pour une matrice réelle symétrique  $A$ , notons  $\bar{\lambda}_A := \max\{\lambda_1, \dots, \lambda_d\}$  où  $\lambda_1, \dots, \lambda_d$  sont les valeurs propres  $A$ . Pour tout  $x \in \mathbb{R}^d$ , on définit les matrices  $d \times d$   $A_\sigma(x)$  et  $B_\sigma(x)$  par

$$A_\sigma(x) := \sum_{k=1}^r \sum_{i=1}^d (\nabla \sigma_{i,k})^* \nabla \sigma_{i,k}(x) \quad \text{et} \quad B_\sigma(x) = \sum_{k=1}^r (\nabla \sigma^k)^* \nabla \sigma^k$$

avec  $\nabla \sigma^k = (\nabla \sigma_{1,k}, \dots, \nabla \sigma_{d,k})$ . Pour  $p \in [1, +\infty)$ , on introduit alors l'hypothèse

**(AC)<sub>p</sub>** : pour tout  $x \in \mathbb{R}^d$ ,  $\nabla b(x) + \frac{1}{2}A_\sigma(x) + (p-1)B_\sigma$  est une matrice définie négative et

$$\sup_{x \in \mathbb{R}^d} \bar{\lambda}_{\nabla b + \frac{1}{2}A_\sigma + (p-1)B_\sigma} := -c_p < 0.$$

**Remarque 3.20** Dans le cas  $d = 1$ , **(AC)<sub>p</sub>** se réécrit

$$\sup_{x \in \mathbb{R}} \left( b'(x) + \left( p - \frac{1}{2} \right) (\sigma'(x))^2 \right) < 0.$$

**Proposition 3.21** ([79]) Soit  $k \in \mathbb{N}^*$  et  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^k$  à dérivées partielles bornées (jusqu'à l'ordre  $k$ ). Supposons que  $b$  et  $\sigma$  soient de classe  $\mathcal{C}^{n,\rho}$  avec  $\rho \in (0, 1)$  et à dérivées bornées et que **(AC)<sub>2k</sub>** soit satisfaite. Alors, les fonctions  $g_T$  et  $g_0$  définies dans le lemme précédent sont de classe  $\mathcal{C}^k$ .

**Remarque 3.22** L'idée principale de la preuve de ces résultats est de montrer que si pour  $n, p \in \mathbb{N}^*$ , **(AC)<sub>np</sub>** est satisfaite, alors  $t \mapsto \nabla^{(n)} X_t^x$  converge exponentiellement vers 0 dans  $L^{2p}$ . Remarquons également que lorsque  $\sigma$  est constant, l'hypothèse **(AC)<sub>p</sub>** ne dépend pas de  $p$  et que les fonctions  $g_0$  et  $g_T$  sont alors  $\mathcal{C}^\infty$  dès que  $\sup_{x \in \mathbb{R}^d} \bar{\lambda}_{\nabla b(x)} < 0$ .

### 3.3.2 Vitesse de convergence de $(\bar{\mu}^{(n)}(\omega, F))_{n \geq 1}$

On considère maintenant le cadre fonctionnel qui, comme on l'a déjà signalé, constitue la motivation initiale de ce travail. Il s'agit donc d'évaluer la vitesse de convergence  $(\bar{\mu}^{(n)}(\omega, F))_{n \geq 1}$  vers  $\mathbb{P}_\nu(F) = \int \mathbb{E}[F(X^x)]\nu(dx)$ . On se limitera au cas où  $F$  dépend de la trajectoire entre 0 et  $T$  où  $T > 0$  et on identifiera avec un léger abus de notation, la fonctionnelle  $F : \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d) \mapsto \mathbb{R}$  à celle vue comme une application de  $\mathcal{C}([0, T], \mathbb{R}^d) \mapsto \mathbb{R}$ . Comme dans le cadre marginal, cette vitesse dépend fortement de l'erreur faible en horizon fini, *i.e.* de l'erreur entre  $\mathbb{E}[F(X^x)]$  et  $\mathbb{E}[F(\xi^{x, \mathbf{h}})]$ . Cependant, il n'existe pas à ce jour de théorie générale assurant un développement dans ce cadre fonctionnel. Nous introduisons donc l'hypothèse suivante :

$(\mathbf{C}_F(\alpha))$  ( $\alpha \in [\frac{1}{2}, 1]$ ) : Pour toute suite de réels strictement positifs  $\mathbf{h} := (h_k)_{1 \leq k \leq k_T}$  telle que  $\sum_{k=1}^{k_T} h_k = T$ , on a

$$\left| \mathbb{E}[F(X^x)] - \mathbb{E}[F(\xi^{x, \mathbf{h}})] \right| \leq C(1 + |x|^r) \|\mathbf{h}\|_\infty^\alpha$$

où  $\|\mathbf{h}\|_\infty = \max_{k=1}^{k_T} h_k$  et  $r > 0$ .

Lorsque  $F : \mathcal{C}([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$ ,  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  et  $\sigma : \mathbb{R}^d \mapsto \mathbb{M}_{d,q}$  sont lipschitziennes, l'hypothèse ci-dessus est satisfaite avec  $\alpha = 1/2$  et  $r = 1$  (car  $\mathbb{E}[\sup_{0 \leq t \leq T} |X_t^x - \xi_t^{x, \mathbf{h}}|] = O(\sqrt{h})$ ). Dans le cas unidimensionnel, il a été récemment obtenu dans [1] que, lorsque les coefficients sont suffisamment réguliers et que  $\sigma$  est uniformément elliptique, l'hypothèse  $(\mathbf{C}_F(\alpha))$  est en fait vraie avec  $\alpha = 3/4 - \varepsilon$  (pour tout  $\varepsilon > 0$ ). De plus, pour certaines fonctionnelles particulières, on peut trouver des résultats spécifiques. Par exemple, lorsque  $F(\alpha) = f(\int_0^T \alpha(s) ds)$ ,  $\alpha \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$  avec  $f$  lipschitzienne bornée, on peut montrer que  $(\mathbf{C}_F(\mathbf{1}))$  est satisfaite (voir [60]).

En accord avec ces remarques, on énonce le théorème, d'abord dans le cadre  $\alpha = 1/2$  puis avec un  $\alpha \in (\frac{1}{2}, 1]$  arbitraire.

**Théorème 3.23 ([79])** *Supposons que  $(\mathbf{S}_{a, \infty})$  soit satisfaite avec  $a \in (0, 1]$  et que  $(\mathbf{S}_T^\nu)$  soit également vérifiée. Soit  $\alpha \in (0, 1]$  tel que  $b$  et  $\sigma$  sont  $C^{4, \alpha}$  sur  $\mathbb{R}^d$  à dérivées bornées. Soit  $F : \mathcal{C}([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$  fonctionnelle lipschitzienne telle que  $f_F : \mathbb{R}^d \rightarrow \mathbb{R}$  définie par  $f_F(x) = \mathbb{E}[F(X^x)]$  satisfait  $(\mathcal{P}_{4, T}^{\text{pol}})$  avec une fonction associée notée  $g_F$ . Alors,*

(i) Si  $\frac{1}{\sqrt{n}} \sum_{k=1}^n \sqrt{\gamma_{N(kT)}} \xrightarrow{n \rightarrow +\infty} 0$ ,

$$\sqrt{nT} \left( \bar{\mu}^{(n)}(F) - \mathbb{E}_\nu(F) \right) = \sqrt{nT} \left( \frac{1}{n} \sum_{k=1}^n F(\xi^{((k-1)T)}) - \mathbb{E}_\nu(F) \right) \xrightarrow{n \rightarrow +\infty} \mathcal{N}(0; \tilde{\sigma}_F^2) \quad (3.45)$$

avec  $\tilde{\sigma}_F^2 = T \mathbb{E}_\nu \left[ \left( F(X^x) - \mathbb{E}[F(X^x)] + \frac{1}{T} (g_F(X_T^x) - \mathbb{E}[g_F(X_T^x)]) \right)^2 \right]$ .

(ii) Si  $(\mathbf{C}_F(\alpha))$  est satisfaite avec  $\alpha \in (\frac{1}{2}, 1]$  et  $\frac{1}{\sqrt{n}} \sum_{k=1}^n (\gamma_{N(kT)})^\alpha \xrightarrow{n \rightarrow +\infty} 0$ , alors (3.45) est encore vraie.

**Remarque 3.24** ▷ Comme dans le cas marginal, on cherche maintenant identifier la vitesse en fonction du nombre  $N = N(nT)$  de points de discrétisation. Soit  $\alpha \in [\frac{1}{2}, 1]$  et  $\gamma_n = Cn^{-\rho}$ . On peut alors vérifier que

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n (\gamma_{N(kT)})^\alpha \xrightarrow{n \rightarrow +\infty} 0 \iff \rho > \frac{1}{2\alpha + 1}.$$

Ainsi, la vitesse en fonction de  $N$  est de l'ordre de

$$\Gamma_N^{-\frac{1}{2}} \sim CN^{-\frac{1-\rho}{2}} \quad \text{si } \rho \in \left( \frac{1}{2\alpha + 1}, 1 \right).$$

Lorsque  $\alpha = 1/2$ , la vitesse optimale est donc de  $O(N^{\frac{1}{2}+\varepsilon})$  pour tout  $\varepsilon > 0$ . À nouveau, on retrouve la vitesse du cas marginal tout en améliorant (a priori) fortement le temps de calcul. Concernant la variance, la comparaison avec l'algorithme originel est plus délicate que dans le cas marginal. On peut simplement remarquer qu'en développant  $\tilde{\sigma}_F^2$ , on retrouve une formulation similaire à (3.25) avec un premier terme purement fonctionnel (via l'inégalité de Cauchy-Schwarz, on peut majorer ce premier terme de (3.25) par celui de  $\tilde{\sigma}_F^2$ ) et un dernier terme dans lequel on retrouve la variance asymptotique relative à  $(\bar{\mu}_n(\omega, f_F))_{n \geq 1}$  (en utilisant la stationnarité du processus sous  $\mathbb{P}_\nu$ ).

▷ Si l'on remplace le schéma  $(\xi_t)_{t \geq 0}$  par le schéma constant par morceaux  $(\bar{X}_t)_{t \geq 0}$  (toujours simulable), le TCL établi en (i) reste vrai sous l'hypothèse légèrement plus forte : il existe  $\delta > 0$ ,  $\frac{1}{\sqrt{n}} \sum_{k=1}^n (\gamma_{N(kT)})^{\frac{1}{2}-\delta} \rightarrow 0$  lorsque  $n \rightarrow +\infty$ . Cette restriction est exactement la même que celle apparaissant dans le théorème 3.11(b).

### 3.3.3 Accélération de convergence par la méthode de Richardson-Romberg

Comme dans tout problème basé sur une discrétisation de type Euler, on peut envisager de mettre en place une méthode de Richardson-Romberg (RR) pour réduire la contrainte liée à la discrétisation. On en rappelle le principe (dans le cadre le plus simple). Si l'on a une formule de développement d'erreur semblable à celle de la proposition 3.16(iii), alors en introduisant le schéma de "pas moitié" (à définir dans notre cadre), on peut, via une combinaison linéaire appropriée du schéma initial et de ce schéma de pas moitié "tuer" le premier terme du développement et réduire ainsi significativement le biais lié à l'erreur de discrétisation. Quand on met en oeuvre ce type de méthode dans un cadre aléatoire, on doit aussi être capable de gérer la variance induite par la modification de l'algorithme pour que cette dernière soit efficace. D'après [77], on sait que pour le problème de l'approximation de  $\mathbb{E}[f(X_T)]$ , il est possible de mettre en oeuvre une méthode RR sans augmenter la variance en utilisant des schémas consistants, *i.e.* construits avec le même Brownien<sup>‡</sup>.

Le but de cette section est donc de proposer une méthode RR pour l'approximation du régime stationnaire tout en cherchant à préserver la variance de l'algorithme originel. Nous nous contenterons ici de présenter cette méthode pour la suite  $(\bar{\mu}_n(\omega, dx))_{n \geq 1}$  et renvoyons à [65] pour les résultats concernant la suite  $(\bar{\nu}_n(\omega, dx))_{n \geq 1}$ . Le point de départ est de construire

‡. Ce résultat est a posteriori naturel puisqu'il correspond au cadre où l'on n'introduit pas d'aléa supplémentaire avec le second schéma. Notons néanmoins que cette interprétation ne s'étend pas de manière claire aux méthodes de Richardson-Romberg d'ordre supérieur où il semble que numériquement, le choix d'accroissements consistants n'est pas optimal.

un couple de schémas d'Euler associés à la diffusion (3.2). On introduit donc dans un premier temps un couple de diffusions dont les dynamiques "marginales" correspondent à celles de (3.2). Plus précisément, on fixe  $\rho \in \mathbb{M}_{q,q}$  et on suppose que  $I_q - \rho^* \rho$  est une matrice symétrique positive (*i.e.* tel que  $\rho^* \rho \leq I_q$  au sens des matrices symétriques positives). On note  $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$  un espace de probabilité filtré sur lequel on peut définir un  $(\mathcal{F}_t)$ -mouvement brownien  $2q$ -dimensionnel noté  $(W, \widetilde{W})$  (de sorte que  $W$  et  $\widetilde{W}$  sont deux  $(\mathcal{F}_t)_{t \geq 0}$ -mouvements browniens indépendants  $q$ -dimensionnels). On définit alors  $W^{(\rho)}$  par

$$W^{(\rho)} = \rho^* W + \sqrt{I_q - \rho^* \rho} \widetilde{W}, \quad (3.46)$$

de sorte que

$$\langle W^i, W^{(\rho);j} \rangle_t = \rho_{i,j} t, \quad t \geq 0.$$

Pour tout  $\rho \in \mathbb{M}_{q,q}$  tel que  $\rho^* \rho \leq I_q$ , on note alors  $\mathbb{X}^{(\rho)} := (X, X^{(\rho)})$  le processus  $2q$ -dimensionnel solution de

$$\begin{cases} dX_t = b(X_t)dt + \sigma(X_t) dW_t, & X_0 = x_1 \in \mathbb{R}^d, \\ dX_t^{(\rho)} = b(X_t^{(\rho)})dt + \sigma(X_t^{(\rho)}) dW_t^{(\rho)}, & X_0^{(\rho)} = x_2 \in \mathbb{R}^d. \end{cases} \quad (3.47)$$

On discrétise alors la première marginale via le schéma  $(\xi_t)_{t \geq 0}$  défini précédemment de pas  $(\gamma_n)_{n \geq 1}$  et de brownien associé  $W$  et la seconde via un schéma  $(\xi_t^{(\rho)})_{t \geq 0}$  de pas moitié construit à l'aide des accroissements de  $W^{(\rho)}$  et de la suite  $(\widetilde{\gamma}_n)_{n \geq 1}$  définie par

$$\forall n \geq 1, \quad \widetilde{\gamma}_{2n-1} = \widetilde{\gamma}_{2n} = \frac{\gamma_n}{2}$$

de sorte que  $\widetilde{\Gamma}_{2n} = \Gamma_n$  (où  $\widetilde{\Gamma}_n = \sum_{k=1}^n \widetilde{\gamma}_k$ ). Ainsi,  $(\xi_t^{(\rho)})_{t \geq 0}$  est défini récursivement par  $\xi_0^{(\rho)} = y$  et pour tout  $n \in \mathbb{N}$ , pour tout  $t \in [\widetilde{\Gamma}_n, \widetilde{\Gamma}_{n+1})$  par :

$$\xi_t^{(\rho)} = \xi_{\widetilde{\Gamma}_n}^{(\rho)} + (t - \widetilde{\Gamma}_n) b(\xi_{\widetilde{\Gamma}_n}^{(\rho)}) + \sigma(\xi_{\widetilde{\Gamma}_n}^{(\rho)}) (W_t^{(\rho)} - W_{\widetilde{\Gamma}_n}^{(\rho)}).$$

On note  $(\bar{\mu}_n^{(\rho)}(\omega, dy))_{n \geq 1}$ , la suite de mesures empiriques associée à  $(\xi_t^{(\rho)})_t$  définie par :

$$\forall n \geq 1, \quad \bar{\mu}_n^{(\rho)}(\omega, f) = \frac{1}{n} \sum_{k=1}^n f(\xi_{(k-1)\Gamma}^{(\rho)}(\omega)).$$

Afin de mettre en oeuvre une méthode RR, on introduit alors  $(\widetilde{\mu}_n^{(\rho)}(\omega, dx))_{n \geq 1}$  suite définie pour tout  $n \geq 1$  comme la combinaison linéaire suivante de  $\bar{\mu}_n$  et  $\bar{\mu}_n^{(\rho)}$  :

$$\widetilde{\mu}_n^{(\rho)}(\omega, f) = 2\bar{\mu}_n^{(\rho)}(\omega, f) - \bar{\mu}_n(\omega, f).$$

Sous les hypothèses  $(\mathbf{S}_{a,\infty})$  et  $(\mathbf{S}_T^\nu)$ ,  $(\bar{\mu}_n(\omega, dx))_{n \geq 1}$  et  $(\bar{\mu}_n^{(\rho)}(\omega, dx))_{n \geq 1}$  convergent *p.s.* vers  $\nu$  l'unique probabilité invariante associée à  $P_T$  (et donc à  $(P_t)_{t \geq 0}$ ), de sorte que  $(\widetilde{\mu}_n^{(\rho)})_{n \geq 1}$  converge également vers  $\nu$ . Afin d'obtenir un résultat de vitesse de convergence, on aura besoin d'une hypothèse supplémentaire : si l'on note  $(Q_t^{(\rho)})_{t \geq 0}$  le semi-groupe relatif à  $\mathbb{X}^{(\rho)}$ , il nous faudra supposer que  $Q_T^{(\rho)}$  admet une unique probabilité invariante notée  $\mu^{(\rho)}$ .

Cette question fera l'objet de la section suivante. Notons pour l'instant que l'existence de  $\nu$  implique l'existence de  $\mu^{(\rho)}$  mais que cela n'est pas vrai pour l'unicité (voir section 3.4

pour plus de détails sur ces propriétés). En revanche, l'unicité de  $\mu^{(\rho)}$  implique clairement l'unicité de  $\nu$  (de sorte que  $(\mathbf{S}_T^\nu)$  n'apparaît pas dans le résultat ci-dessous).

On peut maintenant énoncer le résultat de vitesse de convergence de  $(\tilde{\mu}_n^{(\rho)}(\omega, dx))_{n \geq 1}$  vers  $\nu$  :

**Théorème 3.25 ([79])** *Supposons que  $(\mathbf{S}_{a,\infty})$  soit satisfaite avec  $a \in (0, 1]$  et que  $\text{Tr}(\sigma\sigma^*(x)) = o(V^a(x))$  lorsque  $|x| \rightarrow +\infty$ . Soit  $\rho \in \mathbb{M}_{q,q}$  tel que  $\rho^*\rho \leq I_q$  et tel que  $Q_T^{(\rho)}$  admet une unique probabilité invariante  $\mu^{(\rho)}$ . Supposons que  $b$  et  $\sigma$  soient  $\mathcal{C}^{9,\alpha}$  ( $\alpha \in (0, 1]$ ) sur  $\mathbb{R}^d$  à dérivées bornées (jusqu'à l'ordre 9). Soit  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  satisfaisant  $(\mathcal{P}_{9,T}^{\text{pol}})$  (de fonction associée notée  $g_T$ ). Supposons que la fonction  $\varphi^{(1)} : \mathbb{R}^d \rightarrow \mathbb{R}$  définie par  $\varphi^{(1)}(x) = \frac{1}{2} \int_0^T \mathbb{E}[\Phi_{g_T}(X_s^x)] ds$  vérifie  $(\mathcal{P}_{5,T}^{\text{pol}})$  de fonction associée notée  $g_{\varphi^{(1)}}$  (solution de (3.32)).*

*Supposons enfin que  $(\gamma_n)_{n \geq 1}$  vérifie  $\sum_{k=1}^{+\infty} \gamma_{N(kT)}^2 = +\infty$  et que  $\gamma_{N(nT)} - \gamma_{N((n+1)T)} = o(\gamma_{N(nT)}^2)$ . Alors :*

(i) Si  $\frac{1}{\sqrt{nT}} \sum_{k=1}^n \gamma_{N(kT)}^2 \xrightarrow{n \rightarrow +\infty} 0$ ,

$$\sqrt{nT} \left( \tilde{\mu}_n^{(\rho)}(\omega, f) - \nu(f) \right) \xrightarrow{n \rightarrow +\infty} \mathcal{N}\left(0, (\hat{\sigma}_T^{(\rho)})^2\right)$$

$$\text{o } (\hat{\sigma}_T^{(\rho)})^2 = 5\hat{\sigma}_T^2 - \frac{4}{T} \int \left( g_T(x)g_T(y) - P_T g_T(x)P_T g_T(y) \right) \nu^{(\rho)}(dx, dy). \quad (3.48)$$

(ii) Si  $\sqrt{\frac{T}{n}} \sum_{k=1}^n \gamma_{N(kT)}^2 \xrightarrow{n \rightarrow +\infty} \beta_0 \in (0, +\infty]$  et si  $(\gamma_n)_{n \geq 1}$  est telle que  $N(nT) = nT$  pour tout  $n \geq 1$ , alors :

- Si  $\beta_0 = +\infty$ ,

$$\frac{n}{\sum_{k=1}^n \gamma_{N(kT)}^2} \left( \tilde{\mu}_n^{(\rho)}(\omega, f) - \nu(f) \right) \xrightarrow{\mathbb{P}} \tilde{m}_T = -\frac{1}{2T} \int_{\mathbb{R}^d} \int_0^T \mathbb{E} \left[ \frac{1}{T} \Phi_{g_{\varphi^{(1)}}}(X_s^x) + \chi_g(s, X_s^x) \right] ds \nu(dx)$$

lorsque  $n \rightarrow +\infty$ , où  $\chi_g$  est définie dans l'équation (3.42).

- Si  $\beta_0 \in (0, +\infty)$ ,

$$\sqrt{nT} \left( \tilde{\mu}_n^{(\rho)}(\omega, f) - \nu(f) \right) \xrightarrow{n \rightarrow +\infty} \mathcal{N}\left(\beta_0 \tilde{m}_T, (\hat{\sigma}_T^{(\rho)})^2\right).$$

**Remarque 3.26** • À l'aide de la stationnarité,  $(\hat{\sigma}_T^{(\rho)})^2$  peut être réécrite comme suit :

$$(\hat{\sigma}_T^{(\rho)})^2 = \frac{1}{T} \int_{\mathbb{R}^d \times \mathbb{R}^d} \text{Var}_{x,y} \left( 2g_T(X_T^{(\rho)}) - g_T(X_T) \right) \nu^{(\rho)}(dx, dy).$$

• Supposons que  $\gamma_n = \gamma_1 n^{-\beta}$  avec  $\beta \in (0, 1)$ . On peut montrer que

$$\sum_{k=1}^n \gamma_{N(kT)}^2 \xrightarrow{n \rightarrow +\infty} +\infty \iff \beta \in (0, 1/3)$$

puis que

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n \gamma_{N(kT)}^2 \xrightarrow{n \rightarrow +\infty} 0 \iff \rho > \frac{1}{5}.$$

L'ordre de l'erreur en fonction du nombre de points de discrétisation est alors de

$$\begin{cases} N^{-\frac{1-\rho}{2}} & \text{si } \rho \in (1/5, 1/3) \\ N^{-\rho} & \text{si } \rho \in (0, 1/5]. \end{cases}$$

Le minimum est atteint en  $\rho = 1/5$  et est alors de l'ordre de  $N^{-\frac{2}{5}}$ .

• L'identification de la limite dans (ii) (i.e. lorsque  $\rho \leq 1/5$ ), requiert une hypothèse sur les pas :  $N(nT) = nT$ . Cette hypothèse provient de la proposition 3.16(iv) dans laquelle l'obtention d'un développement d'ordre 2 nécessite de faire disparaître les termes de bords qui deviennent non négligeables.

**Minimisation de la variance.** Afin d'optimiser la méthode RR, il nous faut maintenant minimiser  $(\hat{\sigma}_T^{(\rho)})^2$ . Dans la proposition ci-dessous, on montre que la variance atteint son minimum en  $\rho = I_q$ , i.e. lorsque les accroissements sont consistants, sous la condition que dans ce cas, on ait unicité de la probabilité invariante pour  $Q_T^{(1)}$ . Avant d'énoncer le résultat, remarquons que lorsque  $W = W^{(\rho)}$ , on a une probabilité invariante évidente :

$$\nu_\Delta := \nu \circ \varphi_\Delta^{-1} \quad \text{où } \varphi_\Delta : \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^d \text{ est définie par } \varphi_\Delta(x) = (x, x),$$

i.e. la mesure image de  $\nu$  par l'application  $\varphi_\Delta$ . On a alors le résultat suivant :

**Proposition 3.27 ([79])** *Supposons  $(S_T^\nu)$  satisfaite.*

(i) *Pour tout  $\rho \in [-1, 1]$ ,  $\hat{\sigma}_T^2 \leq (\hat{\sigma}_T^{(\rho)})^2$ .*

(ii) *Si  $\nu_\Delta$  est l'unique probabilité invariante de  $Q_T^{(1)}$ , alors  $(\hat{\sigma}_T^{(1)})^2 = \hat{\sigma}_T^2$ .*

**Remarque 3.28** On remarque donc que la variance asymptotique de l'algorithme RR est toujours supérieure à celle de la procédure initiale mais que l'on retrouve la variance de l'algorithme originel lorsque les accroissements sont consistants et que  $\nu_\Delta$  est unique. La preuve de (i) est essentiellement basée sur l'inégalité de Cauchy-Schwarz appliquée au second terme de (3.48). La minimisation de la variance correspond en fait au cas d'égalité dans l'inégalité de Cauchy-Schwarz. Ce point permet en particulier de montrer que supposer les accroissements consistants est quasiment nécessaire pour retrouver la variance de l'algorithme d'origine. Enfin, notons que lorsque  $\rho = 0$ , i.e. lorsque  $W$  et  $W^{(\rho)}$  sont indépendants, l'unique mesure invariante est  $\nu \otimes \nu$  et que dans ce cas, la variance est maximale :  $(\hat{\sigma}_T^{(0)})^2 = 5\hat{\sigma}_T^2$ .

### 3.4 Unicité de la mesure invariante pour les diffusions dupliquées

Dans la continuité de la section précédente et en particulier des conditions requises pour l'applicabilité de la méthode de Richardson-Romberg et la minimisation de la variance asymptotique, on propose de s'intéresser dans cette partie au problème suivant : pour une diffusion  $(X_t)_{t \geq 0}$  à valeurs dans  $\mathbb{R}^d$  solution de (3.2) ayant une unique probabilité invariante  $\nu$ , en est-il de même pour le couple  $\mathbb{X}^{(\rho)} = (X, X^{(\rho)})$  constitué de deux copies (non nécessairement dépendantes) de la diffusion  $(X_t)_{t \geq 0}$  et défini comme la solution de (3.47) ?

Pour toute matrice  $\rho$  de  $\mathbb{M}_{q,q}$  telle que  $I_q - \rho^*\rho$  est une matrice positive, le processus  $\mathbb{X}^{(\rho)}$  est appelé dans la suite *diffusion dupliquée*. Dans cette classe de processus, on s'intéressera en priorité au cas  $\rho = I_q$ . En effet, il constitue le cas d'intérêt dans la proposition 3.27 (ii) (i.e. le seul le cas où l'on peut espérer appliquer la méthode RR sans augmenter la variance) et du point de vue théorique, il s'agit en quelque sorte du cas le plus difficile car "le plus dégénéré" (il s'agit du cas le moins diffusif). Lorsque  $\rho = I_q$ , le processus est aussi appelé dans la littérature *2-point motion* (voir [56] section 4.2 et [47]). Dans ce cadre, l'ergodicité de ce processus a déjà fait l'objet de nombreux articles, en particulier lorsque le flot stochastique vit dans une variété Riemannienne compacte  $M$ . On réfère notamment aux articles suivants [11, 6, 7, 28] dans lesquels le comportement en temps long du flot (sous son régime stationnaire) est lu au travers de l'exposant de Lyapounov (maximal) défini par

$$\lambda_1 := \limsup_{t \rightarrow +\infty} \frac{1}{t} \log \|D_x \Phi(x, t)\|$$

où  $\|D_x \Phi(x, t)\|$  correspond à la norme (d'opérateur) du processus tangent relatif au flot. Les auteurs s'intéressent le plus souvent au cas où l'exposant de Lyapounov est strictement positif ce qui d'une part, n'a un "sens ergodique" que dans un cadre compact, et qui d'autre part, constitue en quelque sorte le cas opposé à notre problème du point de vue dynamique. En effet, supposer que  $\lambda_1$  est strictement positif revient en quelque sorte à rendre la diagonale  $\Delta = \{(x, x), x \in M\}$  répulsive. Or, supposer que  $\nu_\Delta$  est l'unique mesure invariante relative à  $(X, X)$  implique naturellement que  $\Delta$  doit avoir un caractère attractif. En d'autres termes, dans notre cadre, nous avons besoin d'une propriété de *confluence asymptotique faible* des trajectoires.

La suite de cette section est structurée de la manière suivante. Après quelques énoncés généraux concernant l'existence de probabilité invariante et également, l'unicité dans le cas non dégénéré  $\rho^*\rho < I_q$  (i.e.  $I_q - \rho^*\rho$  définie positive), on se focalise ensuite sur le cas  $\rho = I_q$ . Sous cette hypothèse, on montre que l'unicité de la probabilité invariante relative à la diffusion dupliquée (confluence faible) est essentiellement toujours vraie en dimension 1. En dimension supérieure, après avoir exhibé un contre-exemple, on énonce une série de critères de confluence faible (de type intégral) mais aussi de *confluence trajectorielle presque sûre*, lisibles sur les coefficients de la diffusion au travers d'un *exposant de Lyapounov non-infinitésimal*. Ces critères permettent en particulier de traiter des cas non triviaux comme certaines classes de systèmes-gradients à potentiel (sur-quadratique) non convexe ou, à l'inverse, des systèmes pour lesquels la confluence est induite par le coefficient de diffusion. Enfin, on terminera en établissant un lien avec un problème de transport optimal.

### 3.4.1 Propriétés générales

Dans la suite, on suppose que  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  et  $\sigma : \mathbb{R}^d \rightarrow \mathcal{M}_{d,q}$  sont des fonctions continues telles que l'existence forte, l'unicité trajectorielle sont satisfaites pour (3.2) et telles que  $(X_t)_{t \geq 0}$  est un processus de Markov fellérien. Ces propriétés se transfèrent au processus  $\mathbb{X}^{(\rho)}$ . On note respectivement  $(P_t)_{t \geq 0}$  et  $(Q_t^{(\rho)})_{t \geq 0}$  les semi-groupes associés à  $X$  et  $\mathbb{X}^{(\rho)}$ , puis on fait l'hypothèse (par défaut) que  $(P_t)_{t \geq 0}$  admet une unique probabilité invariante  $\nu$ . Une probabilité invariante pour  $(Q_t^{(\rho)})_{t \geq 0}$  sera alors notée  $\mu^{(\rho)}$  puis  $\mu$  dans le cas particulier  $\rho = I_q$ . **Existence de  $\mu$ .** On a la propriété suivante : l'existence de  $\nu$  implique l'existence de  $\mu$ . En

effet, la famille de mesures d'occupations  $(\mu_t^{(\rho)})_{t>0}$  définie par

$$\forall t \geq 0, \quad \mu_t^{(\rho)} = \frac{1}{t} \int_0^t \nu^{\otimes 2}(dx_1, dx_2) Q_s^{(\rho)}((x_1, x_2), (dy_1, dy_2)) ds \quad (3.49)$$

est tendue car, via la propriété de stationnarité, ses marginales sont égales à  $\nu$  pour tout  $t \geq 0$ , donc sont trivialement tendues. Il est alors classique d'en déduire (en s'appuyant sur la propriété de Feller) que les valeurs d'adhérence de  $(\mu_t^{(\rho)})_{t>0}$  sont automatiquement des probabilités invariantes pour  $(Q_t^{(\rho)})_{t \geq 0}$ . Notons que l'ensemble des probabilités invariantes de  $(Q_t^{(\rho)})_{t \geq 0}$  est contenu dans le convexe  $\mathcal{C}$  de  $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$  défini par

$$\mathcal{C} := \{\mu \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d), \mu(dx \times \mathbb{R}^d) = \mu(\mathbb{R}^d \times dy) = \nu\}$$

où  $\mu(dx \times \mathbb{R}^d)$  et  $\mu(\mathbb{R}^d \times dy)$  sont respectivement les projections de  $\mu$  sur la première et la deuxième coordonnée. Comme  $\mathcal{C}$  est compact pour la topologie de la convergence étroite (car les marginales sont tendues), on peut en fait retrouver la propriété d'existence à l'aide du théorème de point fixe de Kakutani.

#### À propos de l'unicité.

- Comme on peut s'en douter, il est clair que l'on ne peut espérer l'unicité de  $\mu^{(\rho)}$  si l'on n'a pas l'unicité de  $\nu$ . On peut s'en convaincre rapidement via les arguments utilisés pour l'existence. En effet, si  $(P_t)_{t \geq 0}$  admet au moins deux probabilités invariantes distinctes  $\nu$  et  $\nu'$  (et donc une infinité en prenant les combinaisons convexes associées), alors, on peut construire à l'aide d'une variante de (3.49) deux probabilités invariantes distinctes de marginales respectives  $\nu$  et  $\nu'$  et,  $\nu'$  et  $\nu$ .
- Outre la probabilité invariante  $\nu_\Delta$  dans le cas  $\rho = I_q$ , on peut aussi constater que  $\nu \otimes \nu$  est aussi probabilité invariante évidente dans le cas opposé  $\rho = 0_{\mathbb{M}_{q,q}}$  (cas indépendant). De plus, si pour tout  $x \in \mathbb{R}^d$ ,  $P_t(x, \cdot) \rightarrow \nu$ , alors  $\nu \otimes \nu$  est la seule.
- Considérons maintenant le cas  $\rho^* \rho < I_q$ . Dans ce cadre, les propriétés classiques de non-dégénérescence telles que l'uniforme ellipticité ou l'hypoellipticité se transfèrent de la diffusion originelle à la diffusion dupliquée. Ainsi, l'unicité de la probabilité invariante se transfère également. Pour ce qui concerne l'uniforme ellipticité, on remarque que la matrice de diffusion  $\Sigma(X_t, X_t^{(\rho)})$  du couple  $(X_t, X_t^{(\rho)})$  satisfait

$$\Sigma(x_1, x_2) \Sigma(x_1, x_2)^* = \begin{bmatrix} \sigma \sigma^*(x_1) & \sigma(x_1) \rho \sigma^*(x_2) \\ \sigma(x_2) \rho^* \sigma^*(x_1) & \sigma \sigma^*(x_2) \end{bmatrix}.$$

On en déduit alors facilement en utilisant le caractère défini positif de la matrice  $I_q - \rho^* \rho$ , que s'il existe  $\lambda_0 > 0$  tel que pour tout  $x \in \mathbb{R}^d$ ,  $\sigma \sigma^*(x) \geq \lambda_0 I_d$ , alors, il existe  $\lambda_1 > 0$  tel que pour tout  $(x_1, x_2) \in \mathbb{R}^d \times \mathbb{R}^d$ ,  $\Sigma(x_1, x_2) \Sigma(x_1, x_2)^* \geq \lambda_1 I_{2d}$ .

Pour le cas hypoelliptique, on peut également montrer que la condition de Hörmander (faible) et la contrôlabilité du système dynamique se transfèrent bien à la diffusion dupliquée lorsque  $\rho^* \rho < I_q$ . Ces deux propriétés permettent également d'assurer la propriété de Feller forte et l'irréductibilité topologique des processus de Markov considérés. On renvoie à [65], Proposition A.1 pour un résultat plus précis sur le sujet.

En conclusion, lorsque  $\rho^* \rho < I_q$ , l'unicité de la probabilité invariante de la diffusion dupliquée ne pose pas vraiment de problème. Par conséquent, il en est de même pour l'applicabilité du théorème 3.25. Plutôt que de vouloir absolument traiter le cas  $\rho = I_q$ ,



on pourrait alors être tenté d'appliquer ce résultat avec par exemple  $\rho_\theta = \theta I_q$  où  $\theta = 1 - \varepsilon$  avec  $\varepsilon$  petit. En fait, on peut conjecturer que même lorsque l'on n'a pas unicité de  $\nu_\Delta$ , le choix  $\theta = 1$  reste optimal (sur une classe raisonnable de fonctions) au sens où le supremum des "variance asymptotiques" possibles lorsque  $\varepsilon = 0$  reste majoré par celle associée au cas  $\varepsilon > 0$ .

On se focalise maintenant sur le cas d'intérêt  $\rho = I_q$  §.

### 3.4.2 Unicité dans le cas unidimensionnel

Lorsque  $d = q = 1$ , l'unicité de  $\nu$  implique bien celle de  $\nu_\Delta$  sous des hypothèses très générales. L'énoncé précis du résultat est donné dans le théorème 3.29 ci-dessous. Celui-ci est divisé en deux parties. Dans la première, on établit l'unicité en s'appuyant sur des outils de type ergodique. Dans la seconde, on va un peu plus loin en énonçant, sous des hypothèses (très) légèrement plus contraignantes, une propriété de confluence asymptotique des trajectoires au travers de la fonction d'échelle  $p$  dont la définition est rappelée ci-dessous. Dans cette seconde partie du résultat, on revisite des travaux de Has'minskii ([48], théorème 2.2, p.308).

Avant d'énoncer le résultat, donnons quelques définitions. On note  $M$  la mesure de vitesse de la diffusion définie par  $M(d\xi) = (\sigma^2 p')^{-1}(\xi) d\xi$ , où  $p$  est la fonction d'échelle définie (à une constante près) par

$$p(x) = \int_{x_0}^x d\xi e^{-\int_{x_0}^{\xi} \frac{2b}{\sigma^2}(u) du}, \quad x \in \mathbb{R}.$$

Naturellement, on considère la fonction  $p$  lorsque celle-ci est une fonction à valeurs finies (ce qui est vrai si  $b/\sigma^2$  est localement intégrable sur  $\mathbb{R}$ ).

**Théorème 3.29 ([65])** *Supposons que  $b$  et  $\sigma$  soient des fonctions continues sur  $\mathbb{R}$  telles que existence forte et unicité trajectorielle sont satisfaites pour (3.2) et telles que  $(X_t)_{t \geq 0}$  est un processus de Markov féllérien. Supposons également qu'il existe  $\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  strictement croissante satisfaisant  $\lambda(0) = 0$  et  $\int_{0+} \lambda(u)^{-2} du = +\infty$  telle que pour tous  $x, y \in \mathbb{R}$ ,  $|\sigma(y) - \sigma(x)| \leq \lambda(|x - y|)$ . Alors,*

(a) *si  $(X_t)_{t \geq 0}$  admet une unique probabilité invariante  $\nu$ , la probabilité  $\nu_\Delta = \nu \circ (\xi \mapsto (\xi, \xi))^{-1}$  est l'unique probabilité invariante pour la diffusion dupliquée  $(X_t^{x_1}, X_t^{x_2})_{t \geq 0}$ .*

(b) *(Has'minskii) Supposons que  $p$  soit bien définie sur  $\mathbb{R}$  comme une fonction à valeurs réelles, et vérifie*

$$\lim_{x \rightarrow \pm\infty} p(x) = \pm\infty \quad \text{et} \quad M \text{ est finie .}$$

*Alors,  $\nu = M/M(\mathbb{R})$  est l'unique probabilité invariante de  $(X_t)_{t \geq 0}$  et  $(p(X_t))_{t \geq 0}$  est asymptotiquement trajectoriellement confluent :  $\mathbb{P}$ -p.s., pour tous  $x_1, x_2 \in \mathbb{R}$ ,  $p(X_t^{x_1}) - p(X_t^{x_2})$  tend vers 0 lorsque  $t \rightarrow +\infty$ .*

**Remarque 3.30**  $\triangleright$  Les hypothèses sur  $b$  et  $\sigma$  sont en particulier satisfaites si ces fonctions sont localement lipschitziennes et sous-linéaires.

---

§. Ceci signifie en particulier que l'on ne s'intéressera pas aux autres matrices du groupe orthogonal. On renvoie à [15] pour des résultats dans ce cadre.

▷ L'hypothèse impliquant  $\lambda$  et  $\sigma$  (issue de [51]) a pour rôle d'assurer la propriété de comparaison des solutions : si  $x \leq y$ , alors *p.s.*,  $\forall t \geq 0$ ,  $X_t^x \leq X_t^y$ . Cette propriété est en fait l'argument clé pour les preuves des deux résultats. Pour le point (a), cette propriété est l'outil principal permettant d'assurer que toute probabilité invariante extrémale de la diffusion dupliquée est nécessairement portée par l'un des trois sous-ensembles (disjoints) de  $\mathbb{R}^2$  suivants :  $\Delta^+ = \{(x, y), x > y\}$ ,  $\Delta = \{(x, x), x \in \mathbb{R}\}$  ou  $\Delta^- = \{(x, y), x < y\}$ . Le fait que les marginales soient toutes les deux égales à  $\nu$  implique alors un équilibre entre les deux coordonnées qui ne peut être assuré si la probabilité invariante est portée par  $\Delta^+$  ou  $\Delta^-$ . Il suit alors que la seule probabilité invariante extrémale est  $\nu_\Delta$ .

Pour le point (b), en s'appuyant sur le fait que la fonction d'échelle est croissante et que pour tout  $x$ ,  $(p(X_t^x))_{t \geq 0}$  est une martingale locale, on déduit de la propriété de comparaison que lorsque  $x_1 \geq x_2$ ,  $(p(X_t^{x_1}) - p(X_t^{x_2}))_{t \geq 0}$  est une surmartingale positive donc convergente. On utilise alors le point (a) pour obtenir que cette limite est nécessairement égale à 0. Comme cela est signalé dans l'énoncé, il s'agit à l'origine d'un résultat de Has'minskii. Néanmoins, la preuve proposée ici est complètement différente.

▷ L'unicité de la probabilité invariante de la diffusion dupliquée peut être comprise comme une propriété de confluence statistique des trajectoires. En effet, si l'on suppose acquise la convergence de la mesure d'occupation associée à  $(X^{x_1}, X^{x_2})$ , alors

$$\forall K \geq 0, \quad \frac{1}{t} \int_0^t |X_s^{x_2} - X_s^{x_1}| \wedge K ds \xrightarrow{t \rightarrow +\infty} \nu_\Delta(|x_2 - x_1| \wedge K) = 0.$$

On parle donc de confluence faible par opposition aux propriétés de confluence trajectorielle. On peut en particulier (facilement) montrer, en utilisant que  $p$  est strictement croissante, que la propriété de confluence trajectorielle de  $(p(X_t))_{t \geq 0}$  implique bien l'unicité de  $\nu_\Delta$ .

La question naturelle est maintenant de savoir si  $(X_t)_{t \geq 0}$  satisfait une propriété de confluence trajectorielle. On a le résultat suivant :

**Corollaire 3.31** (a) *Supposons que les hypothèses du théorème 3.29(b) sont satisfaites. Si de plus,*

$$\sigma \text{ ne s'annule pas} \quad \text{et} \quad \limsup_{|x| \rightarrow +\infty} \int_0^x \frac{b}{\sigma^2}(\xi) d\xi < +\infty$$

*alors,  $\mathbb{P}$ -p.s., pour tous  $x_1, x_2 \in \mathbb{R}$ ,*

$$X_t^{x_1} - X_t^{x_2} \longrightarrow 0 \quad \text{lorsque} \quad t \rightarrow +\infty.$$

(b) *La conclusion de (a) est en particulier vraie si  $\sigma$  ne s'annule pas et s'il existe  $K > 0$  tel que*

$$|x| > K \implies \text{sign}(x)b(x) \leq 0. \tag{3.50}$$

Ce résultat montre donc que la confluence trajectorielle est aussi une propriété très générale en dimension 1. En particulier, on notera que la condition (3.50) n'est nécessaire qu'en dehors d'un compact.

### 3.4.3 Un contre-exemple.

L'unicité de  $\nu_\Delta$  n'est pas vraie en général en dimension supérieure. On propose donc pour commencer un contre-exemple lorsque  $(X_t)_{t \geq 0}$  est une diffusion 2-dimensionnelle. À l'image du problème d'unicité de la quasi-stationnaire de diffusions gaussiennes évoqué dans la section 2.4, on exploite ici le fait qu'en dimension 2, on peut tourner. Si en particulier, la vitesse angulaire est indépendante de la position et que le rayon est ergodique et ne se stabilise pas sur l'origine, on comprend alors facilement que deux trajectoires issues de deux points initiaux ayant des angles différents auront des difficultés à confluer asymptotiquement. On considère donc le processus suivant  $X_t = (r_t \cos \varphi_t, r_t \sin \varphi_t)$ ,  $t \in \mathbb{R}_+$ , où  $(r_t, \varphi_t)_{t \geq 0}$  est solution de

$$dr_t = \min(r_t, 1)(1 - r_t)(dt + \vartheta dW_t^1), \quad r_0 \in \mathbb{R}_+ \quad (3.51)$$

$$d\varphi_t = c dW_t^2, \quad \varphi_0 \in [0, 2\pi), \quad (3.52)$$

où  $W = (W^1, W^2)$  est un brownien standard 2-dimensionnel et  $c$  et  $\vartheta$  sont des réels strictement positifs (voir [65] pour une formulation en coordonnées cartésiennes). À l'aide de la classification de Feller, on prouve alors que  $x_0 = r_0(\cos \varphi_0, \sin \varphi_0) \neq 0$  (i.e.  $r_0 > 0$ ) et si  $\vartheta \in (0, \sqrt{2})$  alors

$$r_t \longrightarrow 1 \text{ lorsque } t \rightarrow +\infty, \quad (3.53)$$

La composante angulaire étant un brownien standard (à constante près), on en déduit alors facilement que la mesure de Lebesgue sur la sphère unité  $\mathcal{S}_1$  de  $\mathbb{R}^2$  notée  $\lambda_{\mathcal{S}_1}$  est l'unique probabilité invariante de  $(X_t)_{t \geq 0}$ . Vu que  $\varphi_t - \varphi_0 = cW_t^2$ , il est également clair que si  $x = r_0 e^{i\varphi_0}$  et  $x' = r'_0 e^{i\varphi'_0}$ ,  $r_0, r'_0 \neq 0$ ,  $\varphi_0, \varphi'_0 \in [0, 2\pi)$ , alors

$$\lim_{t \rightarrow +\infty} |X_t^x - X_t^{x'}| = |e^{i(\varphi_0 - \varphi'_0)} - 1| \quad \mathbb{P}\text{-p.s.}$$

ce qui implique en particulier que pour tout  $K > 2$ ,

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \mathbb{E}[|X_s^x - X_s^{x'}| \wedge K] ds = |e^{i(\varphi_0 - \varphi'_0)} - 1|.$$

Cette quantité étant différente de 0 si  $\varphi_0 \neq \varphi'_0$ , on en déduit que  $\nu_\Delta$  ne peut être l'unique probabilité invariante (en utilisant que les valeurs d'adhérence de  $(\frac{1}{t} \int_0^t \mathbb{E}[\delta_{X_s^x, X_s^{x'}}] ds)_{t \geq 1}$  sont des probabilités invariantes pour la diffusion dupliquée). Dans cet exemple, on peut en fait caractériser l'ensemble des probabilités invariantes de la diffusion dupliquée :

**Proposition 3.32 ([65])** (a) Une probabilité  $\mu$  est invariante pour  $(Q_t)_{t \geq 0}$  si et seulement si

$$\mu = \mathcal{L}(e^{i\Theta}, e^{i(\Theta+V)}) \quad (3.54)$$

où  $\Theta$  suit la loi uniforme sur  $[0, 2\pi]$  et  $V$  est une variable aléatoire à valeurs dans  $[0, 2\pi)$  indépendante de  $\Theta$ . Enfin, une telle probabilité est extrémale (dans le convexe des probabilités invariantes) si et seulement si  $V = \delta_\varphi$  avec  $\varphi \in [0, 2\pi)$ .

**Remarque 3.33**  $\triangleright$  On retrouve  $\nu_\Delta$  en posant  $V = 0$  et  $\nu \otimes \nu$  en supposant que  $V$  est uniforme sur  $[0, 2\pi]$ .

▷ La propriété d'extrémalité (ou au moins l'implication directe) se comprend facilement en remarquant que pour tout  $\varphi \in [0, 2\pi]$ ,  $A_\varphi = \{((\theta_1, \theta_2) \in [0, 2\pi], \theta_2 - \theta_1 = \varphi)\} = 1$  est invariant par la dynamique.

▷ Le choix de  $(r_t)_{t \geq 0}$  est très particulier puisque que la probabilité invariante associée est  $\delta_1$ . On peut en fait généraliser à toute diffusion ergodique positive  $(r_t)_{t \geq 0}$  satisfaisant les propriétés suivantes :

- $(r_t)_{t \geq 0}$  admet une unique probabilité invariante  $\pi$  telle que  $\pi(\mathbb{R}_+^*) = 1$ .
- Pour tous  $x, y \in (0, +\infty)$ ,  $r_t^x - r_t^y \rightarrow 0$  p.s. lorsque  $t \rightarrow +\infty$  (voir Corollaire 3.31 pour un critère associé).

Dans ce cadre,  $\mu$  est invariante si et seulement si  $\mu = \mathcal{L}(Re^{i\Theta}, Re^{i(\Theta+V)})$  où  $R$  est une variable aléatoire de loi  $\pi$  indépendante de  $\Theta$  et  $V$  (définis comme précédemment).

### 3.4.4 Critères d'unicité dans le cas multidimensionnel

Comme le montre le contre-exemple ci-dessus, les trajectoires du processus ne confluent pas naturellement (même dans un sens statistique) en dimension supérieure. En d'autres termes, la diagonale  $\Delta$  n'est pas naturellement attractive. On propose donc dans cette partie d'exhiber des propriétés géométriques permettant d'assurer une telle propriété. Ce type de démarche est assez classique dans la littérature, permettant par exemple, sous de bonnes conditions (de convexité lorsque l'on a une diffusion-gradient) d'assurer une vitesse exponentielle de convergence à l'équilibre (au sens de la distance de Wasserstein). Les résultats que l'on présente ci-dessous peuvent s'apparenter à ce type de méthode mais les hypothèses pour assurer l'unicité seront beaucoup plus faibles.

Notons  $\nu$  l'unique probabilité invariante de  $(X_t)_{t \geq 0}$  sur  $\mathbb{R}^d$  et  $\mathcal{P}_{\nu, \nu}^*$  le sous-ensemble convexe de  $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$  défini par

$$\mathcal{P}_{\nu, \nu}^* = \left\{ m \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d), m(dx \times \mathbb{R}^d) = m(\mathbb{R}^d \times dy) = \nu, m(\Delta_{\mathbb{R}^d \times \mathbb{R}^d}) = 0 \right\}.$$

Ce sous-espace est conçu pour contenir toutes les mesures invariantes extrémales autres que  $\nu_\Delta$ .

Pour  $S \in \mathcal{S}^{++}(d, \mathbb{R})$  (l'ensemble des matrices symétriques définies positives), on note  $(\cdot | \cdot)_S$  et  $|\cdot|_S$ , les produits scalaires et normes induites définies par  $(x|y)_S = (x|Sy)$  et  $|x|_S^2 = (x|x)_S$  respectivement. De plus pour  $A \in \mathbb{M}_{d,d}$ , on note  $\|A\|_S^2 = \text{Tr}(A^*SA)$  (norme de Hilbert-Schmidt induite). On introduit également

$$[b]_{S,+} = \sup_{x \neq y} \frac{(b(x) - b(y)|x - y)_S}{|x - y|_S^2}.$$

Si  $[b]_{S,+} < +\infty$  et si  $\sigma$  est lipschitzienne, alors on a existence forte et unicité trajectorielle des solutions.

Les critères présentés ci-dessous se lisent au travers d'une sorte d'exposant de Lyapounov que nous définissons maintenant :

**Définition 3.34** Pour une matrice  $S \in \mathcal{S}^{++}(d, \mathbb{R})$ , l'exposant de Lyapounov non-infinitésimal (LIN) est une fonction de  $(\mathbb{R}^d \times \mathbb{R}^d) \setminus \Delta_{\mathbb{R}^d \times \mathbb{R}^d}$  dans  $\mathbb{R}$  définie par :  $\forall x, y \in \mathbb{R}^d, x \neq y$ ,

$$\Lambda_S(x, y) = \frac{(b(x) - b(y)|x - y)_S}{|x - y|_S^2} + \frac{1}{2} \frac{\|\sigma(x) - \sigma(y)\|_S^2}{|x - y|_S^2} - \left( \left| \frac{(\sigma^*(x) - \sigma^*(y))S(x - y)}{|x - y|_S^2} \right|^2 \right).$$

**Théorème 3.35 ([65])** Soit  $S \in \mathcal{S}^{++}(d, \mathbb{R})$  telle que  $[b]_{S,+} < +\infty$ . Supposons  $\sigma$  lipschitzienne et  $\mathcal{I}_{SDE} = \{\nu\}$  où  $\mathcal{I}_{SDE}$  désigne l'ensemble des probabilités invariantes associées à  $(X_t)_{t \geq 0}$ . Alors, (a) si

$$\forall m \in \mathcal{P}_{\nu, \nu}^*, \quad \int_{\mathbb{R}^d \times \mathbb{R}^d} \Lambda_S(x, y) m(dx, dy) < 0, \quad (3.55)$$

alors  $(X_t)_{t \geq 0}$  et sa diffusion dupliquée associée admettent respectivement  $\nu$  et  $\nu_\Delta$  comme unique probabilités invariantes.

(b) si de plus,  $(\frac{1}{t} \int_0^t P_s(x, dy) ds)_{t \geq 1}$  est tendue pour tout  $x \in \mathbb{R}^d$  et s'il existe  $c_0 > 0$  tel que

$$\forall x, y \in \mathbb{R}^d, x \neq y, |x - y|_S^2 \leq \varepsilon_0 \implies \Lambda_S(x, y) \leq -c_0 \quad (3.56)$$

alors le système est asymptotiquement confluent, *i.e.*

$$\forall x_1, x_2 \in \mathbb{R}^d, \quad X_t^{x_1} - X_t^{x_2} \longrightarrow 0 \quad p.s. \text{ lorsque } t \rightarrow +\infty.$$

**Remarque 3.36**  $\triangleright$  Les hypothèses de (a) et (b) ont une différence notable : la première est de type intégral tandis que la seconde est ponctuelle. En particulier, bien que l'hypothèse  $\Lambda_S(x, y) < 0$  pour tous  $x, y$  semble être le cas le plus naturel dans lequel (3.55) s'applique, on verra dans la section 3.4.5, que l'on peut traiter des cas où  $\Lambda_S(x, y)$  prend des valeurs strictement positives sur un sous-espace (non négligeable) de  $\mathbb{R}^d \times \mathbb{R}^d$ . Notons que lorsque  $\Lambda_S < 0$  sur  $\Delta_{\mathbb{R}^d \times \mathbb{R}^d}^c$ , alors l'unicité de  $\nu$  n'est pas requise. Plus précisément, si  $\mathcal{I}_{SDE}$  est compact non vide, alors ce critère prouve également l'unicité de  $\nu$ .

$\triangleright$  Le point de départ de la preuve de ce résultat (écrit sous une forme plus générale dans le théorème 3.1. de [65]) est d'introduire pour une fonction  $\theta : \mathbb{R}_+^* \mapsto \mathbb{R}$  continue, une pseudo-fonction d'échelle  $f_\theta$  définie par :

$$\forall u \in (0, +\infty), \quad f_\theta(u) = \int_1^u e^{\int_\xi^1 \frac{\theta(w)}{w} dw} d\xi \quad (3.57)$$

et de considérer  $(\varphi_\theta((X_t^{x_1}, X_t^{x_2})))_{t \geq 0}$  ( $x_1 \neq x_2$ ) où  $\varphi_\theta : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  est définie par  $\varphi(y_1, y_2) = f_\theta(|y_2 - y_1|_S^2)$  sur  $\Delta_{\mathbb{R}^d \times \mathbb{R}^d}^c$ . L'exposant NIL correspond en fait à  $\mathcal{A}^{(2)}\varphi_\theta$  (où  $\mathcal{A}^{(2)}$  désigne le générateur de la diffusion dupliquée) lorsque  $\theta = 1$ , *i.e.* lorsque  $f_\theta(u) = \log(u)$ . Les hypothèses (3.55) et (3.56) sont alors à comprendre comme des hypothèses de contraction (dans un sens faible). Pour le point (a) dans lequel l'ingrédient principal est le théorème de Birkhoff, le choix  $\theta = 1$  est alors "optimal", au sens où parmi les choix admissibles de  $\theta$ , il s'agit de celui qui génère l'hypothèse de contraction la plus faible. Pour l'assertion (b) basée sur la convergence d'une surmartingale, le choix de  $\theta$  est légèrement plus restrictif, ce qui génère la contrainte supplémentaire (3.56).

### 3.4.5 Applications, extensions

**Localisation autour de la diagonale.** Comme on l'a mentionné dans sa terminologie, l'exposant de Lyapounov est non-infinitésimal. Le rendre infinitésimal reviendrait en quelque sorte à montrer que seul le comportement dans un voisinage (infinitésimal) de la diagonale joue un rôle dans l'unicité de  $\nu_\Delta$ . Ce type de résultat n'est pour l'instant pas accessible. Néanmoins, dans le résultat suivant, on propose un résultat de type localisation en supposant une hypothèse dite d'ellipticité directionnelle, *i.e.* en supposant que le terme de bruit joue un rôle favorable dans l'attractivité de  $\Delta$ .

**Proposition 3.37** ([65]) *Supposons que  $[b]_{S,+} < +\infty$ , que  $\sigma$  soit lipschitzienne et que  $\mathcal{I}_{SDE} = \{\nu\}$ . S'il existe  $\varepsilon_0 > 0$  tel que*

$$\left\{ \begin{array}{l} (i) \text{ hyp. d'ellipticité directionnelle : } \eta_0 := \inf \left\{ |(\sigma^*(x) - \sigma^*(y))S(x-y)|, |x-y|_S \geq \varepsilon_0 \right\} > 0, \\ (ii) \text{ hyp. locale de contraction : } \forall m \in \mathcal{P}_{\nu,\nu}^*, \int_{|x-y|_S \leq \varepsilon_0} \Lambda_S(x,y)m(dx,dy) < 0, \end{array} \right.$$

alors les conclusions du théorème 3.35(a) restent vraies.

Dans les deux paragraphes suivants, on tente d'appliquer le théorème 3.35(a) dans le cas où  $\Lambda_S$  n'est pas strictement négatif partout.

**Le cas  $\Lambda_S \leq 0$ .** On commence par le cas où  $\Lambda_S$  s'annule sur un sous-ensemble de  $\mathbb{R}^d \times \mathbb{R}^d$ . On a alors la proposition assez générale suivante :

**Proposition 3.38** *Supposons que  $[b]_{S,+} < +\infty$ , que  $\sigma$  est lipschitzienne et que  $\mathcal{I}_{SDE} = \{\nu\}$  où  $\nu$  n'est pas à support compact. Alors,  $\nu_\Delta$  est l'unique probabilité invariante si*

$$\forall x, y \in \mathbb{R}^d, \Lambda_S(x, y) \leq 0 \text{ et } \exists R > 0 \text{ s.t. } \max(|x|_S, |y|_S) > R \implies \Lambda_S(x, y) < 0. \quad (3.58)$$

**Remarque 3.39** Ce résultat s'applique par exemple aux diffusions gradient  $dX_t = \nabla U(X_t)dt + \sigma dW_t$  où  $U$  est un potentiel convexe mais seulement strictement convexe en dehors d'un compact (on peut penser aux potentiels à fonds plat).

Pour prouver le résultat précédent, il suffit de remarquer que pour tout  $m \in \mathcal{P}_{\nu,\nu}^*$ ,  $m(|x|_S > R, |y|_S > R) \geq \nu(|x| > R) > 0$ . Bien qu'évidente, cette idée met l'accent sur le fait que pour exploiter le critère intégral, on est plus ou moins contraint de se ramener à  $\nu$  de manière uniforme sur l'ensemble  $\mathcal{P}_{\nu,\nu}^*$ . Dans le point suivant, ce type d'idée est mis en place de manière plus fine grâce à la connaissance explicite de  $\nu$ .

**$\Lambda_S > 0$  sur des sous-domaines de  $\mathbb{R}^d \times \mathbb{R}^d$ .** On essaie ici d'exhiber des exemples d'applications dans un cadre plus "hostile" où  $\Lambda_S$  peut être strictement positif. On se focalise sur les diffusions-gradient à potentiel non convexe.

**Proposition 3.40** ([65]) *Soit  $U : \mathbb{R}^d \rightarrow \mathbb{R}_+$  localement lipschitzienne, différentiable fonction satisfaisant*

$$\exists \gamma > 0 \text{ tel que } 0 < \liminf_{|x| \rightarrow +\infty} \frac{U(x)}{|x|^\gamma} < \limsup_{|x| \rightarrow +\infty} \frac{U(x)}{|x|^\gamma} < +\infty.$$

Alors, (on rappelle que) la diffusion brownienne solution de

$$dX_t^x = -\nabla U(X_t^x)dt + \sigma dW_t, \quad X_0^x = x, \quad (\sigma \in \mathbb{R}_+^*)$$

admet pour unique probabilité invariante  $\nu_\sigma(dx) = C_\sigma e^{-\frac{2U(x)}{\sigma^2}} dx$ .

De plus, si

$$\forall x, y \in \mathbb{R}^d, \quad \Lambda_{Id}(x, y) \leq \beta - \frac{\alpha}{2}(|x|^a + |y|^a) \quad \text{où } \beta \in \mathbb{R}, \alpha, a > 0 \quad (3.59)$$

alors, il existe  $\sigma_c > 0$  tel que, pour tout  $\sigma > \sigma_c$ ,  $\nu_\Delta$  est l'unique probabilité invariante de la diffusion dupliquée associée.

L'hypothèse (3.59) est une sorte de séparation de variables qui permet d'exploiter la connaissance des marginales de  $\nu_\Delta$ . On constate également que la conclusion attendue (unicité de  $\nu_\Delta$ ) est valable lorsque  $\sigma$  est suffisamment grand. Ce point peut s'interpréter comme suit : lorsque  $\sigma$  est grand, les zones (loin de zéro) où  $\Lambda_{Id} < 0$  sont plus souvent visitées ce qui rend la diagonale attractive. Néanmoins, comme on va le voir dans le corollaire ci-dessous, cela impose d'avoir des potentiels surquadratiques afin d'avoir une contractivité suffisante en dehors d'un compact.

**Corollaire 3.41 ([65])** (i) Supposons que  $U : \mathbb{R}^d \rightarrow \mathbb{R}_+$  soit définie par  $U(x) = C|x|^{2p} + \varepsilon(x)$  où  $p > 1$ ,  $C > 0$  et que  $\varepsilon$  soit une fonction de classe  $\mathcal{C}^1$  telle que  $\nabla\varepsilon$  est lipschitzienne. Alors, il existe  $\sigma_c > 0$  tel que, pour tout  $\sigma > \sigma_c$ ,  $\nu_\Delta$  est l'unique probabilité invariante de la diffusion dupliquée associée (à  $dX_t = -\nabla U(X_t)dt + \sigma dW_t$ ).

(ii) Supposons que  $U : \mathbb{R}^d \rightarrow \mathbb{R}_+$  soit définie par  $U(x) = \frac{(|x|^2-1)^2}{4}$ . Alors, la conclusion précédente est valable pour tout  $\sigma > 0$ .

**Remarque 3.42** Le cas (ii) (basé sur un calcul explicite) est assez frappant et pourrait laisser penser que l'on peut espérer mieux dans le cas général. Néanmoins, on peut aussi remarquer que ce cas reste assez favorable car l'ensemble des puits est connexe (lorsque  $d \geq 2$ ).

### 3.4.6 Lien avec le contrôle optimal.

En conclusion de cette partie, on souhaite remarquer que la question de la négativité (au sens intégral) de l'exposant LIN (autrement dit, l'hypothèse 3.35) peut être connectée à un problème de transport optimal (voir *e.g.* [95] pour des rappels sur cette théorie).

Plus précisément, notons  $\bar{\Lambda}_S : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  l'enveloppe semi-continue supérieurement (s.c.s) de  $\Lambda_S$ . Lorsque  $[b]_{S,+} < +\infty$  et  $\sigma$  est lipschitzienne,  $\bar{\Lambda}_S$  est à valeurs dans  $(-\infty, C_{b,\sigma}]$  où  $C_{b,\sigma}$  est une constante réelle. Notons également que lorsque  $b$  et  $\sigma$  sont de classe  $\mathcal{C}^1$ , le prolongement (s.c.s) sur la diagonale de  $\Lambda_S$  a une forme explicite donnée par

$$\underline{\Lambda}_S(x, x) = \frac{1}{2} \sup_{|u|_S=1} \left( u^* (S J_b(x) + J_b^*(x) S) u + \left\| (\nabla \sigma(x)|u) \right\|_S^2 - 2 \left| (\nabla \sigma^*(x) S u | u) \right|^2 \right). \quad (3.60)$$

Alors, si l'on renforce légèrement l'hypothèse (3.55) en supposant aussi que  $\int_{\mathbb{R}^d} \bar{\Lambda}_S(x, x) \nu(dx) < 0$ , le critère devient :

$$\forall m \in \mathcal{P}_{\nu,\nu}(\mathbb{R}^d \times \mathbb{R}^d), \quad \int_{\mathbb{R}^d \times \mathbb{R}^d} \bar{\Lambda}_S(x, y) m(dx, dy) < 0.$$

Bien que légèrement plus contraignante, cette nouvelle hypothèse est aussi mieux posée car  $\mathcal{P}_{\nu,\nu}(\mathbb{R}^d \times \mathbb{R}^d)$  est compact. L'application  $m \mapsto \int_{\mathbb{R}^d \times \mathbb{R}^d} \bar{\Lambda}_S(x, y) m(dx, dy)$  étant également s.c.s., le critère peut alors s'écrire

$$\max \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} \bar{\Lambda}_S(x, y) m(dx, dy), m \in \mathcal{P}_{\nu,\nu}(\mathbb{R}^d \times \mathbb{R}^d) \right\} < 0.$$

À l'aide du théorème de dualité de Kantorovich et de la symétrie de  $\bar{\Lambda}_S$ , on peut à nouveau reformuler cette hypothèse sous la forme suivante :

$$\inf \left\{ \int_{\mathbb{R}^d} \varphi d\nu, \varphi \in L^1(\nu), \varphi(x) + \varphi(y) \geq \bar{\Lambda}_S(x, y), (x, y) \in \mathbb{R}^d \times \mathbb{R}^d \right\} < 0. \quad (3.61)$$

Cette dernière formulation a pour avantage de ne dépendre que de la mesure  $\nu$ .

### 3.5 Commentaires et perspectives

**Un algorithme Romberg-multi-échelle pour le temps long.** Parmi les pistes pour accélérer la convergence, une idée naturelle consisterait à développer une extrapolation de Romberg à l'ordre supérieur. Néanmoins, ce type d'extension pose problème car si dans le cas développé ici, on peut espérer préserver la variance sous condition de confluence faible, cela devient semble-t-il faux aux ordres supérieurs. Néanmoins, on sait que grâce aux algorithmes multi-échelles (*multilevel*), il est possible de construire un moyen efficace de réduire la variance à l'aide d'une combinaison appropriée de schémas d'Euler de pas grossiers et plus fins (et en simulant un nombre de trajectoires inversement proportionnel à la finesse du schéma). Plus récemment (voir [66]), il a été démontré qu'il était possible d'associer l'approche Romberg et l'approche Multilevel en horizon fini.

La mise en oeuvre de telles méthodes en temps long constitue donc une piste exploitable pour accélérer la vitesse sans augmenter la variance. Néanmoins, dans le cadre ergodique, il n'existe pas à proprement parlé d'aspect Monte-Carlo puisqu'il n'y a qu'une seule trajectoire. La composante Monte-Carlo apparaît alors dans la longueur de la trajectoire simulée. Il semble donc que la construction d'une telle méthode pour l'approximation de la mesure invariante requiert la simulation de trajectoires sur des intervalles de longueur différente puis un ajustement de ces longueurs en fonctions des variances asymptotiques de chaque terme impliqué.

**Calcul d'estimateurs dans des modèles "sparses".** Les travaux effectués jusqu'à présent souffrent légèrement d'un manque de mise en application sur des problèmes concrets (voir cependant [80] pour des applications en finance ou sur des phénomènes de coalescence et fragmentation ou [63] pour des applications aux systèmes hamiltoniens). Parmi les nombreux cadres d'applications (comme la dynamique moléculaire par exemple), je m'intéresse actuellement (avec Sébastien Gadat, Sébastien Gerchinovitz et Arnaud Guillin) au calcul d'estimateurs EWA (*Exponentially Weighted Aggregation*) développés dans [23] dans des modèles épars. En effet, ce type d'estimateur est en fait défini comme l'espérance sous le régime stationnaire d'une diffusion-gradient mais en très grande dimension. Ainsi, outre la mise en oeuvre des algorithmes précédents, une estimation efficace nécessite une prise en compte de la particularité du modèle et notamment de l'aspect "sparse" du problème.

**Vers un nouveau critère de confluence faible?** Le point gênant de l'hypothèse de confluence faible (3.55) est qu'elle est écrite sur la base d'une famille de mesures que l'on ne connaît pas. L'idée développée pour les diffusions gradients ou de manière plus générale avec le théorème de dualité de Kantorovich est de tenter de ramener le problème à un critère ne dépendant que la mesure invariante  $\nu$ . Néanmoins, déterminer une fonction  $\varphi$  satisfaisant le critère (3.61) est généralement difficile en pratique. On pourrait donc envisager de développer des méthodes numériques associées à (3.61) permettant de tester la confluence faible.

Sur le plan "théorique", la piste la plus intéressante semble être celle de la localisation développée dans la proposition 3.37. En effet, certains éléments nous laissent penser que lorsque  $\sigma$  est constant par exemple, ce résultat reste vrai sans hypothèse d'ellipticité directionnelle. Cela signifie que le comportement au voisinage de la diagonale détermine à lui seul la propriété de confluence faible. Néanmoins, une telle localisation n'est vraiment exploitable que si l'on peut faire tendre  $\varepsilon_0$  vers 0 dans l'hypothèse (ii) de la proposition 3.37. Ainsi, on obtiendrait un critère "ultime" basé sur un "vrai" exposant de Lyapounov, *i.e.* sur la seule



---

propriété suivante :  $\int_{\mathbb{R}^d} \bar{\Lambda}_S(x, x) \nu(dx) < 0$ . La preuve de cette conjecture reste pour l'instant à l'état de projet.



## Chapitre 4

### Sur le comportement en temps long d'EDS de type fractionnaire

Ce dernier chapitre est consacré à certaines propriétés en temps long d'EDS fractionnaires. Après une introduction générale pointant en particulier les outils inhérents à ce type de système dynamique (pour l'étude du comportement en temps long), on s'intéresse à deux problèmes : dans la section 4.2, il s'agit de la vitesse de convergence à l'équilibre d'EDS multiplicatives tandis que les deux suivantes sont consacrées à la construction et à l'approximation de solutions stationnaires d'EDS dirigées par un brownien fractionnaire ou plus généralement par un processus gaussien à accroissements stationnaires.

#### 4.1 Introduction et rappels

Les équations différentielles dirigées par un mouvement brownien fractionnaire (fBm) ont été introduites pour modéliser l'évolution de phénomènes dont les accroissements du bruit sont fortement inter-dépendants. On trouve de nombreux exemples d'applications dans la littérature. On renvoie à [67] pour des applications en hydrologie et en télécommunications, à [75] pour des applications en biotechnologie ou encore à [40] et [52, 55], pour des modèles de finance ou de biophysique (respectivement).

Du point de vue mathématique, le fait que le fBm soit un processus à mémoire, ne possédant pas la propriété de martingale et ayant potentiellement des variations plus fortes (trajectoires rugueuses) que le brownien standard implique que la plupart des questions naturelles liées aux EDS associées nécessite une théorie propre.

Dans les travaux présentés ci-dessous, nous nous intéresserons à des EDS  $d$ -dimensionnelles de type

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t^H \quad (4.1)$$

où  $(B_t^H)_{t \geq 0}$  est un mouvement brownien fractionnaire d'ordre  $H > 1/2$  ou à des EDS plus générales de la forme

$$dX_t = b(X_t)dt + \sigma dZ_t \quad (4.2)$$

où  $(Z_t)_{t \geq 0}$  est un processus gaussien (général) à accroissements stationnaires mais dans le cas où  $\sigma$  est une matrice constante (les hypothèses précises sur les coefficients seront données dans la suite). Rappelons que l'on parle respectivement d'EDS multiplicatives ou additives selon que  $\sigma$  est variable ou constant.

En particulier, le fait que  $H$  soit strictement supérieur à  $1/2$  dans le premier cas et que  $\sigma$  soit constant dans le second signifient que les travaux abordés ici se focalisent plus sur les problèmes liés à la mémoire du processus plutôt qu'aux problèmes d'intégration du fBm. De telles extensions sont pour l'instant à l'état de projet. Rappelons que dans le premier cas, l'intégrale par rapport au fBm peut être définie via l'intégrale de Young puisque le fBm est  $(H - \varepsilon)$ -hölderien pour tout  $\varepsilon \in (0, H)$ .

Lorsque l'on aborde l'étude du comportement en temps long (sous des hypothèses de stabilité) de tels systèmes dynamiques, le premier problème à gérer est de définir correctement la notion de mesure invariante dans ce contexte non markovien. Historiquement, la première approche utilisable est celle des *Random Dynamical Systems* (RDS) qui s'appuie sur une vision quasi-déterministe du problème (voir [2, 21, 38]). L'idée générale consiste à injecter tout le bruit dans la représentation de l'évolution du système. Pour fixer les idées, considérons la suite  $(X_n)_{n \geq 0}$  à valeurs dans  $\mathbb{R}^d$  définie par

$$X_0 = x, \quad \text{et} \quad \forall n \geq 1, \quad X_{n+1} = F(X_n, \Delta_{n+1}), \quad (4.3)$$

où  $F$  est une application mesurable de  $\mathbb{R}^d \times \mathbb{R}^q \rightarrow \mathbb{R}^d$  et  $\Delta := (\Delta_n)_{n \in \mathbb{Z}}$  est une suite stationnaire de variables aléatoires à valeurs dans  $\mathbb{R}^q$ . On peut en particulier penser à un schéma d'Euler associé à (4.1) ou à (4.2) (voir sections 4.3 et 4.4).

Notons  $\Theta$  l'opérateur de shift sur  $\Omega = (\mathbb{R}^q)^{\mathbb{Z}}$  (muni de la topologie produit) défini pour tout  $w \in \Omega$  par :  $\forall j \in \mathbb{Z}, (\Theta(w))_j = w_{j+1}$ . Le système (4.3) peut alors être vu comme la projection sur la première coordonnée du système grossi  $(X_n, \Theta^n \Delta)_{n \geq 0}$  de condition initiale  $(x, \Delta)$  et d'évolution donnée par

$$\forall n \geq 0, \quad (X_{n+1}, \Theta^{n+1} \Delta) = \varphi(X_n, \Theta^n \Delta)$$

où  $\varphi : \mathbb{R}^d \times (\mathbb{R}^q)^{\mathbb{Z}} \rightarrow \mathbb{R}^d \times (\mathbb{R}^q)^{\mathbb{Z}}$  est définie par

$$\forall (x, w) \in \mathbb{R}^d \times (\mathbb{R}^q)^{\mathbb{Z}}, \quad \varphi(x, w) = (F(x, w_1), \Theta w).$$

Si l'on note  $(\mathcal{G}_n)_{n \geq 1}$  la filtration (constante) définie par  $\mathcal{G}_n = \sigma(\Delta_k, k \in \mathbb{Z})$ , le processus  $(X_n, \Theta^n \Delta)_{n \geq 0}$  peut alors être considéré comme un processus markovien homogène mais d'évolution ... déterministe (la transition est donnée par  $Q(x, w) = \delta_{\varphi(x, w)}$ ). Une mesure invariante pour  $(X_n, \Theta^n \Delta)_{n \geq 0}$  n'est autre qu'un point fixe  $\hat{\mu}$  pour l'application  $\mu \mapsto \mu \circ \varphi^{-1}$  et la loi de  $X_0$  associée s'obtient alors en projetant  $\hat{\mu}$  sur la première coordonnée (on peut noter que la seconde marginale est toujours la loi de  $(\Delta_n)_{n \in \mathbb{N}}$ ).

Bien que ce point de vue donne une définition rigoureuse de mesure invariante pour le système, il possède quelques défauts : du point de vue physique, cette modélisation du problème n'est pas très naturelle puisque l'on a connaissance de tout le futur du processus de bruit. En particulier, on peut ainsi facilement construire des exemples de modèles où la mesure invariante dépend elle-même de tout le futur du processus (voir fin de section 4.4 pour un exemple). Ce premier problème est néanmoins soluble en définissant les mesures invariantes de manière plus restrictive. En revanche, il apparaît un problème plus important lorsque l'on cherche à établir des critères d'unicité de cette même mesure invariante. L'évolution du système étant déterministe, on ne peut pas espérer étendre à ce type de système dynamique, les outils markoviens fondamentaux d'unicité de type "Feller fort+ irréductibilité". Via cette approche, il semble alors nécessaire d'utiliser des critères basés sur les propriétés

de confluence du processus, ce qui restreint fortement le champ d'application. La même difficulté apparaît également lorsque l'on cherche à étudier la convergence à l'équilibre du processus.

Afin de surmonter ces problèmes, Hairer [42] propose de représenter la dynamique de (4.2) (lorsque  $(Z_t)_{t \geq 0}$  est un fBm standard) au travers de *Stochastic Dynamical Systems* (SDS), le but étant d'obtenir une structure markovienne physiquement raisonnable, *i.e.* relativement à la filtration naturelle (éventuellement complétée) du bruit. Ceci est possible si le bruit peut lui-même être muni d'une structure markovienne homogène relativement à cette filtration. Rappelons le principe de la construction de cette structure dans le cadre (plus simple) à temps discret présenté ici. L'idée est de s'appuyer sur la représentation en *moyenne mobile* des suites gaussiennes stationnaires (ce qui correspond à temps continu à la représentation de Mandelbrot-Van Ness pour les accroissements du fBm, sur laquelle on reviendra dans la suite). On rappelle qu'une suite gaussienne stationnaire réelle *purement non déterministe* (voir [49] pour plus de détails) peut être réalisée de la manière suivante :

$$\forall n \in \mathbb{Z}, \quad \Delta_n = \sum_{k=0}^{+\infty} a_k \xi_{n-k} \quad a.s., \quad (4.4)$$

où  $(a_k)_{k \geq 0}$  est une suite de réels positifs tels que  $a_0 \neq 0$  et  $\sum a_k^2 < +\infty$ , et  $(\xi_k)_{k \in \mathbb{Z}}$  est une suite de variables *i.i.d.* centrées réduites. La suite  $(\xi_k)_{k \in \mathbb{Z}}$  est le *processus d'innovation* associé à  $(\Delta_n)_{n \geq 1}$ . Notons  $h : \mathbb{R}^{\mathbb{Z}^-} \rightarrow \mathbb{R}^{\mathbb{Z}^-}$  l'application définie par  $h(u) = (\sum_{\ell=1}^{+\infty} a_\ell u_{k-\ell})_{k \leq 0}$ . Cette application est linéaire bijective continue sur  $\ell^2(\mathbb{Z}^-, \mathbb{R})$  d'inverse (continue) notée  $h^{-1}$ . On a alors

$$\forall n \in \mathbb{Z}, \forall w \in \mathbb{R}^{\mathbb{Z}^-}, \quad \mathcal{L}(\Delta_{n+1} | (\Delta_{n+k})_{k \leq 0} = w) = P(w, d\delta)$$

où  $P(w, d\delta) = \mathcal{N}(\sum_{k=1}^{+\infty} a_k \xi_{n+k+1}^w, a_0^2)$  avec  $\xi^w = h^{-1}(w)$ . Notons que ce type de construction vaut aussi dans le cas  $q$ -dimensionnel si les coordonnées du processus de bruit sont indépendantes. Supposons donc construit une telle probabilité de transition sur  $\bar{\Omega} \times \mathcal{B}(\mathbb{R}^q)$  où  $\bar{\Omega} = (\mathbb{R}^q)^{\mathbb{Z}^-}$  et considérons l'opérateur de *concaténation* de  $\bar{\Omega} \times \mathbb{R}^q \rightarrow \bar{\Omega}$  défini par

$$\forall w \in \bar{\Omega}, \forall \delta \in \mathbb{R}^q, \quad (w \sqcup \delta)_k = \begin{cases} w_{k+1} & \text{if } k \leq -1 \\ \delta & \text{if } k = 0. \end{cases}$$

La suite  $(\Delta_{n+k})_{k \leq 0})_{n \in \mathbb{Z}}$ , est alors munie d'une structure markovienne donnée par : pour tout  $f : \bar{\Omega} \rightarrow \mathbb{R}$

$$\mathbb{E}[f((\Delta_{n+1+k})_{k \in \mathbb{Z}^-}) | (\Delta_{n+k})_{k \in \mathbb{Z}^-} = w] = \int f(w \sqcup \delta) P(w, d\delta).$$

Le SDS  $(X_n, (\Delta_{n+k})_{k \leq 0})_{n \geq 0}$  est donc un processus de Markov de transition  $Q$  défini par :

$$\int g(x', w') Q((x, w), (dx', dw')) = \int g(F(x, w \sqcup \delta), w \sqcup \delta) P(w, d\delta).$$

Une mesure  $\mu$  est dite invariante pour le SDS si elle l'est pour la transition  $Q$ .

**Remarque 4.1** Les approches RDS et SDS ont donc en commun de définir des mesures stationnaires pour  $(X_n)_{n \geq 0}$ , *i.e.* telles que la loi de  $(X_n)_{n \geq 0}$  de condition initiale  $\mu$  est invariante par translation. La seconde approche permet en revanche d'assurer que la mesure  $\mu$  soit une

condition initiale *admissible*, *i.e.* telle que le processus  $(X_n)_{n \geq 0}$  de condition initiale  $\mu$  soit adapté par rapport à la filtration  $(\mathcal{F}_n)_{n \geq 0}$ . Afin d'éviter de toujours travailler avec la structure markovienne (qui dans le cadre à temps continu fait appel à plus de technicité), ces deux propriétés joueront souvent (avec un léger abus) le rôle de définition dans la suite.

Comme cela a été mentionné précédemment, l'approche SDS permet de construire des critères d'unicité de la mesure invariante qui miment ceux du cadre markovien standard, l'idée étant de définir des notions de type "propriété de Feller forte" et "irréductibilité (topologique)" adaptés à ce type de processus. Cette approche est utilisée dans les articles [44], [41] et [45] pour démontrer successivement l'unicité (dans un sens légèrement différent du cadre classique) de la mesure invariante d'EDS fractionnaires multiplicatives elliptiques d'ordre  $H > 1/2$ , puis hypoelliptiques d'ordre  $H > 1/3$ . Enfin, dans [43], la même théorie est développée pour les systèmes à temps discret de type (4.3).

Enfin, l'approche SDS (ou plus exactement la représentation de Mandelbrot-Van Ness) permet également de développer des outils pour étudier la vitesse de convergence à l'équilibre via des méthodes de couplage. Dans [42], il est ainsi démontré (sous des hypothèses de stabilité) que lorsque  $\sigma$  est constant et  $H \in (0, 1)$ , alors la loi de  $(X_{t+}^x)_{t \geq 0}$  solution de  $dX_t = b(X_t)dt + \sigma dB_t^H$  converge vers la loi du processus en régime stationnaire à une vitesse bornée pour tout  $\varepsilon > 0$ , par  $C_\varepsilon t^{-(\alpha-\varepsilon)}$  où

$$\alpha = \begin{cases} \frac{1}{8} & \text{if } H \in (\frac{1}{4}, 1) \setminus \{\frac{1}{2}\} \\ H(1 - 2H) & \text{if } H \in (0, \frac{1}{4}]. \end{cases} \quad (4.5)$$

Ce résultat (qui permet aussi de démontrer l'unicité de la mesure invariante) constitue le point d'ancrage de la section suivante :

## 4.2 Convergence à l'équilibre d'EDS fractionnaires multiplicatives d'ordre $H > 1/2$

L'objectif de ce travail est de généraliser les résultats de [42] rappelés ci-dessus aux EDS fractionnaires multiplicatives d'ordre  $H > 1/2$ . Il s'agit a priori du premier travail sur ce thème dans le cadre des EDS fractionnaires.

On s'intéresse ici à  $(X_t)_{t \geq 0}$  processus à valeurs dans  $\mathbb{R}^d$  solution de

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t^H \quad (4.6)$$

où  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $\sigma : \mathbb{R}^d \rightarrow \mathbb{M}_{d,d}$  sont (au moins) continues (on rappelle que  $\mathbb{M}_{d,d}$  est l'espace des matrices  $d \times d$  à valeurs réelles) et  $H > 1/2$ .  $(B_t^H)_{t \geq 0}$  est un fBm  $d$ -dimensionnel d'indice de Hurst  $H \in (\frac{1}{2}, 1)$ . On supposera qu'existence forte et unicité trajectorielle des solutions issues de  $x_0 \in \mathbb{R}^d$  sont satisfaites (voir plus bas pour des conditions suffisantes sur les coefficients).

Le processus  $(B_t^H)_{t \geq 0}$  peut être représenté comme une intégrale par rapport à un mouvement brownien standard. Parmi ces représentations, on rappelle celle dite de Mandelbrot-Van Ness :

$$B_t = \alpha_H \int_{-\infty}^0 (-r)^{H-\frac{1}{2}} (dW_{r+t} - dW_r), \quad t \geq 0, \quad (4.7)$$

où  $(W_t)_{t \in \mathbb{R}}$  est un mouvement brownien  $d$ -dimensionnel (indexé par  $\mathbb{R}$ ) et  $\alpha_H$  est un coefficient de normalisation dépendant de  $H$ . Comme on l'a déjà mentionné, cette représentation de type "moyenne mobile" permet de munir le couple  $(X_t, (B_{s+t})_{s \leq 0})_{t \geq 0}$  d'une structure markovienne. Plus exactement, pour des réels  $\theta$  et  $\delta$  tels que  $\theta \in (1/2, H)$  et  $\theta + \delta \in (H, 1)$ , notons  $\mathcal{W}_{\theta, \delta}$  l'espace polonais construit par complétion de l'espace  $\mathcal{C}_0^\infty((-\infty, 0], \mathbb{R}^d)$  (espace des fonctions  $f : (-\infty, 0] \rightarrow \mathbb{R}^d$  de classe  $\mathcal{C}^\infty$  à support compact telles que  $f(0) = 0$ ) pour la norme

$$\|f\|_{\mathcal{W}_{\theta, \delta}} = \sup_{-\infty < s < t \leq 0} \frac{|f(t) - f(s)|}{|t - s|^\theta (1 + |t|^\delta + |s|^\delta)}. \quad (4.8)$$

À l'aide d'une construction similaire à celle introduite dans le cadre discret dans l'introduction, on peut montrer que  $(X_t, (B_{s+t})_{s \leq 0})_{t \geq 0}$  peut être réalisé sur l'espace  $\mathbb{R}^d \times \mathcal{W}_{\theta, \delta}$  au travers d'une transition fellerienne  $(Q_t)_{t \geq 0}$  (on renvoie à [44] pour des détails). Une condition initiale pour (4.6) est une probabilité  $\mu_0$  sur  $\mathbb{R}^d \times \mathcal{W}_{\theta, \delta}$ . En termes probabilistes, une condition initiale est la donnée d'un couple  $(X_0, (B_s)_{s \leq 0})$  où  $(B_s)_{s \leq 0}$  est un fBm à valeurs dans  $\mathbb{R}^d$  indexé par  $(-\infty, 0]$ . Celle-ci est dite invariante pour  $(X_t)_{t \geq 0}$  si elle l'est pour  $(Q_t)_{t \geq 0}$ . On dit qu'on a unicité à équivalence près de la mesure invariante  $\mu$  si la loi globale du processus en régime stationnaire est unique (voir [44] pour plus de détails et des critères).

Lorsque cette propriété d'unicité est satisfaite, une question d'intérêt est celle de la vitesse de convergence de la loi de  $X_t^{\mu_0}$  vers le régime stationnaire (au sens marginal ou fonctionnel). Du point de vue probabiliste, ce type de résultat peut s'obtenir via des méthodes de couplage dont on rappelle rapidement le principe : l'idée est de construire deux copies  $(X_t)_{t \geq 0}$  et  $(\tilde{X}_t)_{t \geq 0}$  issues respectivement de  $\mu_0$  et  $\mu$  (terminologie un peu abusive car dans ce cadre les conditions initiales sont couplées) et de tenter de les "coller". Supposons que ça soit possible et notons

$$\tau_\infty := \inf\{t > 0, \forall s \geq t, X_s = \tilde{X}_s\}.$$

$(\tilde{X}_t)_{t \geq 0}$  étant de loi initiale la mesure invariante, il est donc stationnaire et pour tout  $t \geq 0$ ,  $\mathcal{L}(\tilde{X}_{t+}) = \mathbb{P}_\mu$  où  $\mathbb{P}_\mu$  est l'unique loi du processus en régime stationnaire. Ainsi, via la définition de  $\tau_\infty$ ,

$$\|\mathcal{L}(X_{t+}) - \mathbb{P}_\mu\|_{TV} \leq \mathbb{P}(\tau_\infty > t).$$

Dans un cadre markovien, lorsque deux trajectoires se touchent, alors il suffit pour qu'elles restent collées que les accroissements futurs du bruit soient identiques. En d'autres termes, la majoration précédente reste valable avec  $\tau = \inf\{X_t = \tilde{X}_t\}$ . Dans un cadre non markovien, rester collé a un coût, *i.e.* nécessite un couplage non trivial. Dans [42], l'auteur montre (dans le cas  $\sigma$  constant) que ce coût peut être contrôlé, *i.e.* qu'il est possible de construire un couplage tel que la probabilité de rester collé jusqu'à l'infini est strictement positive (à chaque tentative). Chaque tentative est construite de la manière suivante : d'abord, on attend (classiquement) que les trajectoires ne soient pas trop loin l'une de l'autre (dans cette partie, les innovations de  $B^H$  et de  $\tilde{B}^H$  sont identiques). Une fois cette étape réalisée, on tente le couplage en deux étapes :

- Étape 1 : Sur un intervalle de longueur 1, on tente de coller  $X$  et  $\tilde{X}$ .
- Étape 2 : si l'étape 1 réussit, on tente de laisser les trajectoires collées jusqu'à l'infini (cette seconde étape se réalise en fait sur des intervalles successifs  $I_N$  dont la longueur croît exponentiellement :  $|I_N| = 2^N$ ).

Si l'une ou l'autre des étapes rate, alors on passe à l'étape 3, qui consiste à attendre suffisamment (en utilisant des innovations identiques) pour être dans de bonnes conditions

pour tenter un nouvel essai de couplage. Il y a en fait deux contraintes : il faut d'une part que les trajectoires se rapprochent à nouveau mais il faut aussi que les accroissements des fBms associés à chaque copie soient également assez proches. En effet, l'accroissement du fBm ayant une mémoire assez forte de son passé, le coût d'une tentative de couplage sera d'autant plus élevé si les passés de chaque fBms sont très éloignés. Il faut donc laisser le temps effacer (partiellement) la mémoire de la tentative précédente. En fait, cette seconde contrainte s'avère être la plus coûteuse. On peut noter en particulier que plus l'étape 2 échoue tardivement, plus le temps d'attente associé est élevé (de rapport exponentiellement proportionnel).

Dans cette construction, le fait que  $\sigma$  soit constant est fondamental pour assurer les propriétés suivantes :

- Si deux fBms  $B$  et  $\tilde{B}$  du fBm diffèrent d'un terme de drift, alors deux trajectoires  $X$  et  $\tilde{X}$  de (4.6) dirigées respectivement par  $B$  et  $\tilde{B}$  diffèrent également d'un terme de drift : cette propriété est fondamentale pour réaliser l'étape 1 à l'aide du théorème de Girsanov.
- Sous des hypothèses sur le drift de type "convexité en dehors d'un compact" deux trajectoires  $X$  et  $\tilde{X}$  dirigées par le même fBm (ou plus précisément par deux copies assez proches) se rapprochent naturellement en un temps contrôlable de manière déterministe (ou quasi-déterministe).

Ainsi, si l'on veut étendre le schéma de couplage à un cadre multiplicatif, il faut naturellement réussir à surmonter les difficultés que génère l'absence de telles propriétés. Concernant la première, l'idée est de restreindre à des coefficients  $\sigma$  pour lesquels, il existe une fonction injective  $h : \mathbb{R}^d \rightarrow \mathbb{R}^d$  tel que  $(h(X_t), h(\tilde{X}_t))$  possède la propriété décrite ci-dessus. Cela nous amènera à supposer que  $\sigma$  a la propriété suivante :  $x \mapsto \sigma^{-1}(x)$  est une matrice jacobienne ( $\sigma$  est donc en particulier inversible). Cette hypothèse sera détaillée et commentée plus bas.

En ce qui concerne la seconde propriété fortement basée sur le caractère constant de  $\sigma$ , l'idée naïve est de tenter de remplacer les arguments de "contraction déterministe en dehors d'un compact" par une approche de type Meyn-Tweedie dans le cadre fractionnaire : en d'autres termes, si les trajectoires ne se rapprochent pas de manière déterministe ou quasi-déterministe, on peut quand même espérer que sous des hypothèses de type Lyapounov, elles reviennent simultanément dans un compact en un temps contrôlé (en loi). Comme on peut s'y attendre, ce problème est plus compliqué dans ce cadre non markovien. Outre la difficulté d'estimer le temps de retour simultané des deux trajectoires, il faut aussi avoir un contrôle du bruit utilisé avant cet instant. Ces estimations semblent difficilement accessibles, on choisit de contourner le problème de la manière suivante : pendant l'étape 3, on laisse le temps nécessaire pour que la précédente tentative de couplage soit suffisamment oubliée et pour que la probabilité à la fin de cette étape, que les deux trajectoires soient toutes les deux dans un compact, soit minorée par un réel strictement positif. Cette minoration de la probabilité est réalisée à l'aide d'un contrôle de l'espérance d'une fonction de Lyapounov du système, mais conditionnée à l'échec des précédentes tentatives. Ce conditionnement implique en particulier de comprendre les "déformations" de la loi du fBm induites par les précédentes tentatives. Cette étape constitue l'une des difficultés principales du travail.



### 4.2.1 Résultat principal

On commence par introduire les hypothèses. Afin d'assurer la propriété d'existence forte et d'unicité trajectorielle pour les solutions de (4.6) (voir [74, 20] pour des rappels), on fera l'hypothèse suivante :

**(H<sub>0</sub>)** :  $b$  est localement lipschitzienne sous-linéaire et  $\sigma$  est une fonction bornée  $(1 + \gamma)$ -lipschitzienne avec  $(\gamma \in (\frac{1}{H} - 1, 1])$ .

**Remarque 4.2** On rappelle que  $\sigma : \mathbb{R}^d \rightarrow \mathbb{M}_{d,d}$  est dite  $(1 + \gamma)$ -lipschitzienne ( $\gamma \in (0, 1]$ ) si elle est de classe  $\mathcal{C}^1$  à dérivées partielles bornées et globalement  $\gamma$ -höldériennes, *i.e.* si pour tous  $i, j \in \{1, \dots, d\}$ , la norme suivante est finie :

$$\|\sigma_{i,j}\|_{1+\alpha} = \sup_{x \in \mathbb{R}^d} |\nabla \sigma_{i,j}(x)| + \sup_{x, y \in \mathbb{R}^d} \frac{|\nabla \sigma_{i,j}(x) - \nabla \sigma_{i,j}(y)|}{|x - y|^\gamma}. \quad (4.9)$$

On introduit ensuite une hypothèse de Lyapounov permettant d'assurer la stabilité en temps long du processus :

**(H<sub>1</sub>)** : Il existe une fonction  $V : \mathbb{R}^d \rightarrow \mathbb{R}_+^*$  de classe  $\mathcal{C}^1$  telle que

$$\liminf \frac{V(x)}{|x|^2} > 0, \quad \text{et} \quad |\nabla V| \leq C\sqrt{V} \quad (C > 0), \quad (4.10)$$

et telle qu'il existe  $\beta_0 \in \mathbb{R}_+$  et  $\kappa_0 > 0$  pour lesquels

$$\forall x \in \mathbb{R}^d, \quad (\nabla V(x)|b(x)) \leq \beta_0 - \kappa_0 V(x).$$

**Remarque 4.3** Outre la stabilité, cette hypothèse permet d'assurer que les trajectoires ont une probabilité (conditionnellement aux tentatives passées) minorée d'être dans un compact. On peut remarquer que les hypothèses sur  $b$  permettant d'assurer **(H<sub>0</sub>)** et **(H<sub>1</sub>)** sont assez contraignantes. Elle est adaptée aux champs de vecteurs  $b$  "attractifs" (en dehors d'un compact) dont la croissance est essentiellement linéaire. En particulier, dans ce cadre fractionnaire où l'on a absence de propriétés de martingales, il paraît techniquement délicat d'étendre aux forces de rappels à croissance strictement surlinéaire ou sous-linéaire comme on peut le faire dans le cas des diffusions par exemple.

On termine maintenant par l'hypothèse supplémentaire annoncée précédemment (pour pouvoir réaliser l'étape 1 avec un coût contrôlé) :

**(H<sub>2</sub>)**  $\forall x \in \mathbb{R}^d$ ,  $\sigma(x)$  est inversible et il existe  $h = (h_1, \dots, h_d) : \mathbb{R}^d \rightarrow \mathbb{R}^d$  de classe  $\mathcal{C}^1$  telle que sa jacobienne  $\nabla h = (\partial_{x_j} h_i)_{i,j \in \{1, \dots, d\}}$  est localement lipschitzienne sur  $\mathbb{R}^d$  et vérifie  $\nabla h = \sigma^{-1}$ .

**Remarque 4.4**  $\triangleright$  Sous les hypothèses **(H<sub>0</sub>)** et **(H<sub>2</sub>)**,  $h$  est un  $\mathcal{C}^1$ -difféomorphisme global de  $\mathbb{R}^d$  dans  $\mathbb{R}^d$ . En effet, sous ces hypothèses,  $\nabla h$  est partout inversible et  $x \mapsto [(\nabla h)(x)]^{-1} = \sigma(x)$  est bornée sur  $\mathbb{R}^d$ . La propriété de  $\mathcal{C}^1$ -difféomorphisme est alors une conséquence du théorème de Hadamard-Lévy (voir e.g. [89]). Cette hypothèse renvoie donc à un argument de type

“transformée de Lamperti  $d$ -dimensionnelle”. Il est cependant important de noter que cet argument ne joue un rôle “que” localement, *i.e.* que dans l'étape 1.

▷ Comme cela a déjà été mentionné, l'hypothèse  $(\mathbf{H}_2)$  est la restriction la plus gênante. Néanmoins, on peut remarquer que l'on a peu de restrictions sur la fonction  $h$ . Par exemple, la bornitude de  $\sigma^{-1}$  n'est pas requise, ce qui permet de considérer des cas où l'ellipticité est évanescence au voisinage de l'infini. On peut aussi facilement vérifier que les EDS satisfaisant : pour tout  $i \in \{1, \dots, d\}$ ,

$$dX_t^i = b_i(X_t^1, \dots, X_t^d)dt + \sigma_i(X_t^1, \dots, X_t^d)d(B_t^H)^i$$

où  $\sigma_i : \mathbb{R}^d \rightarrow \mathbb{R}$  est une fonction strictement positive de classe  $\mathcal{C}^1$ , vérifient l'hypothèse  $(\mathbf{H}_2)$ . Plus généralement,  $(\mathbf{H}_2)$  est adaptée (au moins) pour les matrices  $\sigma$  de la forme

$$\sigma(x) = P \text{Diag}(\sigma_1(x_1, \dots, x_d), \dots, \sigma_d(x_1, \dots, x_d))$$

où  $P$  est une matrice  $d \times d$  inversible (et  $\sigma_i$  a les mêmes propriétés que précédemment), *i.e.* les cas où la matrice  $\sigma$  peut s'écrire comme une transformée linéaire d'une matrice diagonale.

On peut maintenant énoncer le résultat principal. Notons  $\mathcal{L}((X_t^{\mu_0})_{t \geq 0})$  (resp.  $\mathbb{P}_\mu$ ) la loi du processus solution de (4.6) sur  $\mathcal{C}([0, +\infty), \mathbb{R}^d)$  de condition initiale  $\mu_0$  (resp. de condition initiale  $\mu$  où  $\mu$  est une probabilité invariante). Notons  $\bar{\mu}_0(dx)$  la première marginale de  $\mu_0(dx, dw)$ .

**Théorème 4.5** ([32]) *Soit  $H \in (1/2, 1)$ . Supposons  $(\mathbf{H}_0)$ ,  $(\mathbf{H}_1)$  et  $(\mathbf{H}_2)$  satisfaites. Alors, (4.6) admet une unique probabilité invariante  $\mu$  (à équivalence près). De plus, pour toute loi initiale  $\mu_0$  pour laquelle il existe  $r > 0$  telle que  $\int |x|^r \bar{\mu}_0(dx) < \infty$ , pour tout  $\varepsilon > 0$ , il existe  $C_\varepsilon > 0$  tel que*

$$\|\mathcal{L}((X_{t+s}^{\mu_0})_{s \geq 0}) - \mathbb{P}_\mu\|_{TV} \leq C_\varepsilon t^{-(\frac{1}{8}-\varepsilon)}.$$

**Remarque 4.6** ▷ On retrouve donc l'ordre des vitesses du cas additif. Il s'agit de la contribution principale de ce résultat. Les résultats d'existence et d'unicité ne sont pas réellement nouveaux. Cependant, si l'on compare aux hypothèses de [44], on peut noter que “mis à part l'hypothèse  $(\mathbf{H}_2)$ ” qui n'apparaît pas dans l'article cité, les autres hypothèses sont légèrement contraignantes. En particulier,  $b$  n'est ici que localement lipschitzienne sous-linéaire et  $x \mapsto \sigma^{-1}(x)$  ne nécessite pas de bornitude.

▷ Le fait que le taux ne dépende pas de  $H$  est assez surprenant. Ce taux est relié aux moments de la durée de chaque tentative. Alors que la durée de l'étape 2 a un moment d'ordre  $\alpha$  pour tout  $\alpha \in (0, H)$ , l'étape 3 de “démémorisation” nécessite un temps plus élevé dont le moment, après optimisation de la méthode est fini pour tout  $\alpha < 1/8$ . Si, au vu de la construction (et de la preuve) de l'étape 2, la limitation par  $H$  semble assez naturelle, on est assez tenté d'imaginer que  $1/8$  n'est pas une quantité significative. Terminons enfin en remarquant que lorsque l'on a un terme de drift (fortement) contractant partout (comme pour le Ornstein-Uhlenbeck par exemple), on peut facilement établir une vitesse de convergence exponentielle (plus naturellement en distance de Wasserstein mais aussi en variation totale). Il est bien entendu ce type de cas ne donne pas d'indication réelle sur la vitesse “naturelle” de convergence à l'équilibre de ce type d'EDS et que les cas d'intérêt sont ceux où le drift a des propriétés de contraction seulement en dehors d'un compact.

### 4.2.2 Éléments de preuve

Comme on l'a déjà expliqué, le point de départ consiste à considérer un couple de trajectoires  $(X, \tilde{X})$  dont les marginales sont solutions de (4.6) de fBms respectifs notés  $B$  et  $\tilde{B}$ . Les processus d'innovation associés (par la représentation de Mandelbrot-Van Ness) sont notés  $W$  et  $\tilde{W}$ . On suppose pour simplifier que  $X$  est issu de  $x \in \mathbb{R}^d$  déterministe et que par ailleurs  $\tilde{X}$  a pour loi initiale  $\mu$ . L'instant de première tentative de couplage est noté  $\tau_0$  (jusqu'à cet instant, les bruits sont supposés identiques, *i.e.* *p.s.*  $\forall t \in (-\infty, \tau_0), W_t = \tilde{W}_t$ ). Plus généralement, le début de la  $k$ -ième tentative (si les précédentes ont échoué) est noté  $\tau_{k-1}$  et sa durée  $\Delta\tau_k$ . On peut alors définir  $\tau_\infty$  par  $\tau_\infty := \tau_{k^*-1} + 1^*$  où  $k^* := \inf\{k \geq 1, \Delta\tau_k = +\infty\}$ . En remarquant que,

$$\mathbb{P}(\tau_\infty > t - 1) = \mathbb{P}\left(\tau_0 + \sum_{k=1}^{+\infty} \Delta\tau_k 1_{k^* > k} > t\right),$$

on déduit assez facilement (voir [32], section 5) que l'on peut ramener le problème au contrôle uniforme en  $k$  de :

$$\mathbb{P}(\Delta\tau_k < +\infty | \Delta\tau_{k-1} < +\infty, \mathcal{F}_{\tau_{k-1}}) \quad \text{et} \quad \mathbb{E}[|\Delta\tau_k|^p | \Delta\tau_k, \mathcal{F}_{\tau_{k-1}}]$$

où  $(\mathcal{F}_t)_{t \geq 0}$  est l'augmentation usuelle de  $((\sigma(W_s, \tilde{W}_s))_{s \leq t})_{t \geq 0}$  et  $p \in (0, 1)$  est à maximiser (comme le suggère le théorème 4.5 et la remarque qui suit,  $p < 1/8$ ).

**$(K, \alpha)$ -admissibilité.** Afin de contrôler les quantités ci-dessus, on doit introduire une condition d'admissibilité. Plus exactement, on ne peut tenter le couplage que si le système est dans de bonnes conditions. On parlera alors de  $(K, \alpha)$ -admissibilité. Pour définir ce concept, on fait maintenant l'hypothèse que  $W$  et  $\tilde{W}$  diffèrent d'un terme de drift noté  $g_w$  :  $dW_t = d\tilde{W}_t + g_w(t)dt$ ,  $g_w$  étant nulle avant la première tentative et durant chaque étape 3. Afin de quantifier l'impact de  $g_w$  sur les futures tentatives de couplage, on introduit un opérateur  $\mathcal{R}_T$  défini par : pour tout  $T \geq 0$  et pour toute fonction  $g : \mathbb{R} \rightarrow \mathbb{R}$  (lorsque l'expression ci-dessous a un sens),

$$(\mathcal{R}_T g)(t) = \int_{-\infty}^0 \frac{t^{\frac{1}{2}-H}(T-s)^{H-\frac{1}{2}}}{t+T-s} g(s) ds, \quad t \in (0, +\infty).$$

On peut alors définir la notion d'admissibilité de la manière suivante :

**Définition 4.7** Soient  $K$  et  $\alpha$  deux réels strictement positifs et  $\tau$  un  $(\mathcal{F}_t)_{t \geq 0}$ -temps d'arrêt. On dit alors que le système est  $(K, \alpha)$ -admissible à l'instant  $\tau$  si  $\tau(\omega) < +\infty$  et si  $(X_\tau^1(\omega), X_\tau^2(\omega), (W^1(\omega), W^2(\omega))_{t \leq \tau})$  vérifie :

$$\sup_{T \geq 0} \int_0^{+\infty} (1+t)^{2\alpha} |(\mathcal{R}_T g_w^\tau)(t)|^2 dt \leq 1 \quad (\text{où } g_w^\tau(\cdot) = g_w(\cdot + \tau)) \quad (4.11)$$

et si  $(X_\tau^1(\omega), X_\tau^2(\omega), (W^1(\omega), W^2(\omega))_{t \leq \tau}) \in \Omega_{K, \alpha, \tau}$  où

$$\Omega_{K, \alpha, \tau} := \{|X_\tau^1(\omega)| \leq K, |X_\tau^2(\omega)| \leq K, \varphi_{\tau, \varepsilon_\theta}(W^1(\omega)) \leq K \text{ et } \varphi_{\tau, \varepsilon_\theta}(W^2(\omega)) \leq K\}, \quad (4.12)$$

---

\*. Lorsque le couplage réussit, les trajectoires sont collées à partir de l'instant  $\tau_{k-1} + 1$  ce qui explique la définition.

avec  $\varepsilon_\theta = \frac{H-\theta}{2}$  avec  $\theta \in (1/2, H)$  et pour  $\varepsilon > 0$ ,

$$\varphi_{\tau,\varepsilon}(w) = \sup_{\tau \leq s \leq t \leq \tau+1} \left| \frac{1}{t-s} \int_{-\infty}^{\tau-1} (t-r)^{H-\frac{1}{2}} - (s-r)^{H-\frac{1}{2}} dw_r \right| + \|w\|_{\frac{1}{2}-\varepsilon}^{\tau-1,\tau}.$$

La condition nouvelle par rapport à [42] est la seconde, qui impose à la fois que les positions et le bruit antérieur soient contrôlés par une constante  $K$  à calibrer. Sans rentrer dans les détails, la condition de  $(K, \alpha)$ -admissibilité permet d'assurer que la probabilité que le couplage soit réussi (*i.e.* tel que les trajectoires restent collées jusqu'à l'infini) est minorée : il existe  $\delta_0 > 0$  tel que pour tout  $k \geq 1$ ,

$$\mathbb{P}(\Delta\tau_k = +\infty | \{\Delta\tau_{k-1} < +\infty\} \cap \Omega_{K,\alpha,\tau_{k-1}}) \geq \delta_0.$$

Une des difficultés principales consiste alors à montrer que la probabilité de  $(K, \alpha)$ -admissibilité est également minorée : il existe  $\delta_1 > 0$  tel que pour tout  $k \geq 1$ ,

$$\mathbb{P}(\Omega_{K,\alpha,\tau_k} | \Delta\tau_{k-1} < +\infty) \geq \delta_1.$$

La preuve de cette propriété s'effectue en deux étapes principales. Il faut d'abord contrôler les variations de  $W$  et de  $\tilde{W}$  avant l'instant  $\tau_k$ , conditionnellement aux échecs des tentatives précédentes. Ceci nécessite à la fois une connaissance assez précise de la construction des browniens à chaque étape du couplage pour pouvoir identifier les "déformations" générées par chaque scénario de couplage passé. Une réinjection de ces contrôles dans la représentation de Mandelbrot- Van Ness permet ensuite d'identifier l'impact de ces déformations sur les accroissements futurs du fBm (voir lemmes 4.5, 4.6 et 4.7 de [32] pour plus de détails). La seconde étape consiste à s'appuyer sur l'hypothèse de Lyapounov pour contrôler  $\mathbb{E}[V^r(X_{\tau_k}) + V^r(\tilde{X}_{\tau_k}) | \Delta\tau_k < +\infty]$  (où  $r$  est un réel strictement positif). Conditionnellement à la première étape, le point clé pour établir ce contrôle est le résultat de contraction suivant (proposition 4.4 de [32]) : si  $(\mathbf{H}_0)$  et  $(\mathbf{H}_1)$  sont satisfaites, alors, pour tout  $\theta \in (1/2, H)$ , il existe  $\bar{\rho} \in (0, 1)$ ,  $C > 0$  et  $r > 0$  tels que pour tous  $x \in \mathbb{R}^d$ ,

$$V^r(X_1) \leq \bar{\rho}V^r(x) + C(1 + \|B^H\|_\theta^{0,1}) \quad \text{où} \quad \|B\|_\theta^{0,1} = \sup_{0 < s < t < 1} \frac{B_t^H - B_s^H}{(t-s)^\theta}. \quad (4.13)$$

Par une itération de cette inégalité, on constate que le contrôle de  $(V^r(X_N))_{N \in \mathbb{N}}$  est alors possible via celui de la suite  $(\|\delta^N B^H\|_\theta^{0,1})_{N \in \mathbb{N}}$  où  $(\delta^N B^H)_t = B_t^H - B_N^H$ , qui lui est basé sur les résultats de l'étape précédente.

### 4.3 Approximation de solutions stationnaires d'EDS multiplicatives d'ordre $H > 1/2$

Dans cette section et celle qui suit, l'objectif est de construire des méthodes d'approximation du comportement en temps long d'EDS fractionnaires. En horizon fini, les propriétés de convergence de schémas de discrétisation de type Euler ont déjà fait l'objet de nombreux articles (voir par exemple les articles suivants ainsi que les références associées [73, 72, 24, 71, 27]). Pour l'approximation du comportement en temps long, les références sont plus rares. On peut néanmoins citer l'article [39], dans lequel la convergence d'une

procédure basée sur un schéma implicite est étudiée lorsque  $\sigma$  est constant et des propriétés de contraction forte sont supposées sur le drift (permettant de voir la mesure invariante comme un attracteur aléatoire au sens des RDS).

Dans les travaux présentés ici, l'approche est similaire à celle développée dans le chapitre 3, basée sur l'approximation de la mesure d'occupation du processus par celle relative à un schéma de discrétisation associé. En revanche, contrairement au chapitre 3, on choisit ici de baser l'étude sur un schéma d'Euler à pas constant (à la manière de [94]). Ce choix est dû à des difficultés techniques liées au pas décroissant dans ce cadre (mais qui sont a priori surmontables) mais aussi à un aspect simulation. En effet, contrairement à celle des accroissements (indépendants) du brownien standard, la simulation du fBm est coûteuse et non (raisonnablement) récursive (voir plus bas pour des détails sur l'aspect simulation). Ainsi, l'intérêt d'une procédure à pas décroissant s'en trouve réduit car il n'est alors pas possible d'exploiter une simulation précédente pour affiner l'approximation.

Dans toute cette section, on considère donc  $(X_t)_{t \geq 0}$  processus à valeurs dans  $\mathbb{R}^d$  solution de

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t^H \quad (4.14)$$

où  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $\sigma : \mathbb{R}^d \rightarrow \mathbb{M}_{d,q}$  sont continues (on rappelle que  $\mathbb{M}_{d,q}$  est l'espace des matrices  $d \times q$  à valeurs réelles) et  $H > 1/2$ .  $(B_t^H)_{t \geq 0}$  est un fBm  $q$ -dimensionnel d'indice de Hurst  $H \in (\frac{1}{2}, 1)$  (la seule différence avec (4.6) est qu'ici  $d$  n'est pas nécessairement égal à  $q$ ). On a pour objectif d'approcher (sur un espace adéquat) la probabilité  $\mathbf{P}$  du processus en régime stationnaire, qu'on appellera ici solution stationnaire. Avant de préciser la définition de solution stationnaire, on fait quelques rappels sur les espace hölderiens. On note  $\mathcal{C}^\theta(\mathbb{R}_+, \mathbb{R}^d)$ , l'ensemble des fonctions  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  globalement hölderiennes sur tout compact, *i.e.* telles que

$$\forall T > 0, \quad \|f\|_{\theta, T} := \sup_{0 \leq s < t \leq T} \frac{|f(t) - f(s)|}{(t - s)^\theta} < +\infty.$$

Muni de la distance  $\delta$  définie par

$$\delta_\theta(f, g) = \sum_{N \in \mathbb{N}} 2^{-N} \left( 1 \wedge \left( \sup_{0 \leq t \leq N} |f(t) - g(t)| + \|f - g\|_{\theta, N} \right) \right),$$

l'espace  $\mathcal{C}^\theta(\mathbb{R}_+, \mathbb{R}^d)$  est un espace métrique complet mais non séparable (voir par exemple [46]). Afin de le rendre séparable, on le restreint au sous-espace fermé des fonctions  $f$  telles que

$$\forall T > 0, \quad \omega_{\theta, T}(f, \delta) := \sup_{0 \leq s < t < T, 0 \leq |t-s| \leq \delta} \frac{|f(t) - f(s)|}{|t - s|^\theta} \xrightarrow{\delta \rightarrow 0} 0. \quad (4.15)$$

Ce nouvel espace est noté  $\bar{\mathcal{C}}^\theta(\mathbb{R}_+, \mathbb{R}^d)$ . Sous l'hypothèse  $(\mathbf{H}_0)$  introduite dans la section précédente, les solutions de (4.14) sont à valeurs dans  $\bar{\mathcal{C}}^\theta(\mathbb{R}_+, \mathbb{R}^d)$ . On peut alors définir les solutions stationnaires de (4.14) de la manière suivante (sans faire appel à la structure markovienne introduite plus haut) :

**Définition 4.8** Soit  $\theta \in (1/2, H)$  et  $(X_t)_{t \geq 0}$  solution de (4.14). Soit  $\mathbf{P}$  la loi de  $(X_t)_{t \geq 0}$  sur  $\bar{\mathcal{C}}^\theta(\mathbb{R}_+, \mathbb{R}^d)$ . Alors,  $\mathbf{P}$  est appelée solution stationnaire de (4.14) si  $\mathbf{P}$  est invariante par l'opérateur de shift (standard) sur  $\bar{\mathcal{C}}^\theta(\mathbb{R}_+, \mathbb{R}^d)$ . Une telle solution stationnaire est dite adaptée, si  $\forall t \geq 0$ ,  $(X_s)_{0 \leq s \leq t}$  et  $(B_s^H)_{s \geq t}$  sont indépendants conditionnellement à  $(B_s^H)_{s \leq t}$ .

**Remarque 4.9** Par un léger abus, on identifie le processus et sa loi. La définition donnée ici correspond à celle de [44]. La notion d'adaptation assure le caractère *physiquement admissible* des solutions stationnaires. Via la représentation de Mandelbrot-Van Ness, cette condition peut être reformulée de manière plus simple au travers du processus d'innovation. Supposer la solution adaptée revient exactement à supposer que pour tout  $t \geq 0$ ,  $(X_t)_{0 \leq s \leq t}$  et  $(W_s - W_t)_{s \geq t}$  sont indépendants (ce qui est en fait vrai dès que  $X_0$  est indépendant de  $(W_t)_{t \geq 0}$ ). En d'autres termes, cette condition exclue les solutions construites en faisant usage du futur des innovations. On verra dans la suite que la construction ergodique utilisée ici garantit ce type de propriété. Notons pour terminer que la notion d'unicité de mesure invariante à équivalence près évoquée dans l'introduction coïncide exactement avec la notion d'unicité de solution stationnaire adaptée.

Afin d'approcher la loi  $\mathbf{P}$ , on introduit la procédure suivante. On fixe  $\gamma$  un réel strictement positif puis on discrétise l'équation (4.14) par un schéma d'Euler (standard)  $(\bar{X}_t^\gamma)_{t \geq 0}$  défini par :  $\bar{X}_0^\gamma = x \in \mathbb{R}^d$  et pour tout  $n \geq 0$ ,

$$\forall t \in [n\gamma, (n+1)\gamma), \quad \bar{X}_t^\gamma = \bar{X}_{n\gamma}^\gamma + (t - n\gamma)b(\bar{X}_{n\gamma}^\gamma) + \sigma(\bar{X}_{n\gamma}^\gamma)(B_t^H - B_{n\gamma}^H). \quad (4.16)$$

On définit ensuite la suite  $(\mathcal{P}^{(n,\gamma)}(\omega, d\alpha))_{n \geq 1}$  sur  $\bar{\mathcal{C}}^\theta(\mathbb{R}_+, \mathbb{R}^d)$  avec  $\theta < H$  par

$$\forall n \geq 1, \quad \mathcal{P}^{(n,\gamma)}(\omega, d\alpha) = \frac{1}{n} \sum_{k=1}^n \delta_{\bar{X}_{\gamma(k-1)+}^\gamma}(\omega)(d\alpha)$$

où  $\bar{X}_{t+}^\gamma$  désigne le processus  $\bar{X}$  shifté de  $t$ .

**Remarque 4.10** On peut ici noter que le fait de considérer une mesure d'occupation fonctionnelle (et non marginale) n'est pas dû à un désir de généralité. En effet, dans ce cadre non markovien, la mesure empirique marginale  $\mathcal{P}^{(n,\gamma)}(\omega, dx) = \frac{1}{n} \sum_{k=1}^n \delta_{\bar{X}_{\gamma(k-1)}^\gamma}$  ne contient pas assez d'information pour que ses limites potentielles puissent être caractérisées. Le passage par le fonctionnel peut donc être perçu comme un outil quasi-obligatoire pour la convergence de la procédure même marginale.

La suite  $(\mathcal{P}^{(n,\gamma)}(\omega, d\alpha))_{n \geq 1}$  est en particulier construite pour approcher "la" solution stationnaire de l'équation discrétisée à  $\gamma$  fixé. Une telle solution est définie de manière quasi-identique au cas continu :

**Définition 4.11** Soit  $(Y_t^\gamma)_{t \geq 0}$  solution de :  $\forall n \geq 0$ ,

$$\forall t \in [n\gamma, (n+1)\gamma), \quad Y_t^\gamma = Y_{n\gamma}^\gamma + (t - n\gamma)b(Y_{n\gamma}^\gamma) + \sigma(Y_{n\gamma}^\gamma)(B_t^H - B_{n\gamma}^H). \quad (4.17)$$

et  $\mathbf{P}^\gamma$  sa loi sur  $\bar{\mathcal{C}}^\theta(\mathbb{R}_+, \mathbb{R}^d)$  où  $\theta \in (1/2, H)$ .  $\mathbf{P}^\gamma$  est appelée solution stationnaire de (4.17) si  $\mathbf{P}^\gamma$  est invariante par l'opérateur  $\Theta_\gamma$  défini sur  $\bar{\mathcal{C}}^\theta(\mathbb{R}_+, \mathbb{R}^d)$  par  $\Theta_\gamma(w) = w(\gamma + \cdot)$ . Une telle solution stationnaire est dite adaptée, si  $\forall t \geq 0$ ,  $(Y_t^\gamma)_{0 \leq s \leq t}$  et  $(B_s^H)_{s \geq t}$  sont indépendants conditionnellement à  $(B_s^H)_{s \leq t}$ .

La seule différence avec la définition (4.8) est qu'ici la stationnarité n'est exigée que dans un sens discret (en effet, comme tout schéma à temps continu, celui-ci ne peut être muni d'une structure homogène en temps continu).

**Unicité des solutions stationnaires adaptées.** On note  $(\mathbf{S}^\gamma)$  l'hypothèse :

$(\mathbf{S}^\gamma)$  ( $\gamma \geq 0$ ) : Il existe au plus une solution stationnaire adaptée de (4.14) (resp. de (4.17)) si  $\gamma = 0$  (resp. si  $\gamma > 0$ ). En accord avec la remarque 4.9, on rappelle que  $(\mathbf{S}^0)$  peut être obtenue via les critères de [44] ou par les résultats de la section précédente. Lorsque  $\gamma > 0$ , on a le résultat suivant (obtenu comme application de [43]) :

**Proposition 4.12 ([18])** Soit  $H \in (1/2, 1)$ . Supposons que  $d = q$  et que  $b$  et  $\sigma$  soient de classe  $\mathcal{C}^2$ . Supposons que  $\sigma$  soit inversible et que  $\sup_{x \in \mathbb{R}^d} \sigma^{-1}(x) < +\infty$ . Alors,  $(\mathbf{S}^\gamma)$  est vérifiée pour tout  $\gamma > 0$ .

**Convergence de  $(\mathcal{P}^{(n,\gamma)}(\omega, d\alpha))_{n \geq 1}$ .** On peut maintenant énoncer le théorème principal :

**Théorème 4.13 ([18])** Soit  $H \in (1/2, 1)$  et  $\theta \in (1/2, H)$ . Supposons que  $(\mathbf{H}_0)$  soit satisfaite et que  $(\mathbf{H}_1)$  le soit également avec  $V$  de classe  $\mathcal{C}^2$  de Hessienne bornée. Si  $(\mathbf{S}^\gamma)$  est vérifiée pour tout  $\gamma > 0$ , alors

(i) il existe  $\gamma_0 > 0$  tel que, pour tout  $\gamma \in (0, \gamma_0)$ ,

$$\lim_{n \rightarrow +\infty} \mathcal{P}^{(n,\gamma)}(\omega, d\alpha) = \mathbf{P}^\gamma(d\alpha) \quad p.s.,$$

pour la topologie de la convergence étroite induite par  $\bar{\mathcal{C}}^\theta(\mathbb{R}_+, \mathbb{R}^d)$  et  $\mathbf{P}^\gamma$  désigne l'unique solution stationnaire adaptée de (4.17).

(ii) Si de plus  $(\mathbf{S}^0)$  est vérifiée, alors,

$$\lim_{\gamma \rightarrow 0} \mathbf{P}^\gamma(d\alpha) = \mathbf{P}(d\alpha) \quad a.s.$$

où  $\mathbf{P}$  désigne l'unique solution stationnaire adaptée de (4.14).

**Remarque 4.14**  $\triangleright$  Mis à part l'adaptation sur laquelle on reviendra juste après, l'architecture de preuve est relativement standard. On établit dans un premier temps un contrôle uniforme de  $\mathbb{E}[V(\bar{X}_{n\gamma}^\gamma)]$  à l'aide de l'hypothèse de Lyapounov qui permet ensuite de déduire une tension p.s. marginale puis fonctionnelle de la suite  $(\mathcal{P}^{(n,\gamma)}(\omega, d\alpha))_{n \geq 1}$  ((voir par exemple [88], Théorème 1.4 pour des critères de tension sur  $\bar{\mathcal{C}}^\theta(\mathbb{R}_+, \mathbb{R}^d)$ ). Il suffit alors d'établir que les valeurs d'adhérence sont des solutions stationnaires de l'équation discrétisée pour obtenir le point (i). La stationnarité est quasi-automatique tandis que le fait que la limite soit solution de l'EDS discrétisée est une conséquence de l'ergodicité des accroissements du fBm. La preuve de (ii) est basée sur le même concept mais pose plus de problèmes. En effet, obtenir la tension de  $\mathbf{P}^\gamma$  implique indirectement de contrôler  $\mathbb{E}[V(\bar{X}_{n\gamma}^\gamma)]$  uniformément en  $n$  mais aussi en  $\gamma$ . L'obtention de la double uniformité (et notamment la seconde) constitue la difficulté principale et nécessite un contrôle très fin de l'évolution du schéma d'Euler sous l'hypothèse  $(\mathbf{H}_1)$ .

Concernant l'adaptation, la preuve se base sur des méthodes de martingales qui exploitent l'idée suivante : lorsque l'on initialise le schéma (ou le processus) en une condition initiale qui ne dépend pas du futur, alors, les translatés de cet objet conservent cette propriété. Ainsi, celle-ci se transfère aux limites de moyennes de (mesures de Dirac en) ces processus translatés.

$\triangleright$  La vitesse de convergence des mesures d'occupation vers le régime stationnaire (à temps discret ou continu) reste à ce jour un problème ouvert (voir section 4.5 pour plus de détails sur ce sujet).

**Simulations numériques.** On propose ici d'illustrer numériquement ce résultat sur un exemple en dimension 1. On considère  $(X_t)_{t \geq 0}$  solution de

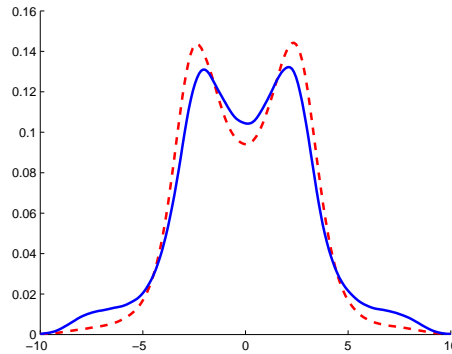
$$dX_t = -X_t dt + (4 + \cos(X_t)) dB_t^H$$

pour laquelle on cherche à estimer la densité de sa loi stationnaire marginale notée  $\nu^H$ . D'après le théorème 4.13, pour toute fonction  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ,

$$\lim_{\gamma \rightarrow 0} \lim_{n \rightarrow +\infty} \mathcal{P}_0^{(n, \gamma)}(\omega, f) = \nu^H(f)$$

où  $\mathcal{P}_0^{(n, \gamma)} = 1/n \sum_{k=1}^n \delta_{\bar{X}_{(k-1)\gamma}^\gamma}$ . L'approximation de densité s'effectue alors classiquement par convolution avec un noyau (choisi ici gaussien). Les accroissements du fBm sont simulés à l'aide de la méthode de Wood-Chan (voir [98]), basée sur le plongement des accroissements fractionnaires dans une matrice symétrique circulaire (dont les valeurs propres peuvent être calculées par "Fast Fourier Transform"). Notons que la complexité de cette méthode est en  $n \log n$  (et que, par conséquent, le temps de calcul reste raisonnable même pour des valeurs assez grandes de  $n$ ).

Dans la figure 4.1, on s'intéresse à la comparaison de  $\nu^H$  (lorsque  $H = 3/4$ ) avec  $\nu^{\frac{1}{2}}$  où  $\nu^{\frac{1}{2}}$  désigne la mesure invariante de la diffusion  $dX_t = -X_t dt + (4 + \cos(X_t)) dW_t$ . Dans ce cadre unidimensionnel,  $\nu^{\frac{1}{2}}$  est explicite (notons au passage que dans le cas des EDS fractionnaires, les seuls cas où la loi stationnaire est explicitable sont les processus de type Ornstein-Uhlenbeck, *i.e.* les cas où l'EDS est un processus gaussien, voir *e.g.* [16]). Dans la figure 4.2, on propose de tester numériquement la vitesse de convergence de la procédure.



**Figure 4.1 :** Densité approchée de  $\nu^{\frac{3}{4}}$  (en continu) comparée à celle de  $\nu_0^{\frac{1}{2}}$  (discontinu),  $n = 10^7$ ,  $\gamma = 0.05$

Dans la première figure, on constate que  $\nu^{\frac{3}{4}}$  a une queue de distribution plus épaisse. Ceci est naturel puisque la variance du fBm (au-delà de  $t = 1$ ) croît avec  $H$ . Du point de vue dynamique, on peut aussi y voir (de manière quasi-équivalente avec l'interprétation précédente) un lien avec la corrélation positive des accroissements du fBm lorsque  $H > 1/2$  qui accentue l'inertie de la particule.



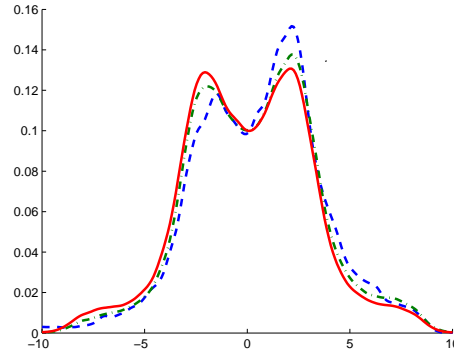


Figure 4.2 : Densité approchée de  $\nu^{\frac{3}{4}}$  pour  $n = 10^5$  (“dotted”),  $n = 10^6$  (“dash-dotted”),  $n = 10^7$  (continu)

#### 4.4 Existence et approximation de solutions stationnaires d'EDS gaussiennes

Il s'agit ici de s'intéresser au même problème pour une EDS  $d$ -dimensionnelle additive mais dirigée par un processus gaussien à accroissements stationnaires général noté  $Z$  :

$$\forall t \geq 0, \quad dX_t = b(X_t)dt + dZ_t, \quad (4.18)$$

où  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  est une fonction (au moins) continue. Notons que si  $\sigma$  est une matrice constante,  $\sigma Z$  est encore un processus gaussien. Ceci explique l'absence de coefficient devant le processus directeur. Le but de ce travail (antérieur à celui de la section précédente) est principalement de comprendre quels types d'hypothèses sont nécessaires sur la structure de covariance du processus directeur pour assurer l'existence de solution stationnaire. De plus, l'approche choisie étant la même que précédemment, elle procure à la fois une preuve constructive de l'existence et une méthode d'approximation de ces dernières.

On suppose donc ici que  $(Z_t)_{t \in \mathbb{R}} = (Z_t^1, \dots, Z_t^d)_{t \in \mathbb{R}}$  est un processus gaussien centré nul en 0 à accroissements stationnaires. Pour tout  $i \in \{1, \dots, d\}$ , on note  $c_i : \mathbb{R} \rightarrow \mathbb{R}_+$  la variance associée : pour tous  $s, t$ ,  $\mathbb{E}[(Z_t^i - Z_s^i)^2] = c_i(t - s)$ . On en déduit facilement que

$$\forall s < t, \quad \mathbb{E}[Z_s^i Z_t^i] = \frac{1}{2} (c_i(t) + c_i(s) - c_i(t - s))$$

puis que

$$\forall s < t, \quad \mathbb{E}[Z_s^i (Z_{t+s}^i - Z_t^i)] = \frac{1}{2} (c_i(t + s) - c_i(t) - (c_i(t) - c_i(t - s))).$$

Lorsque  $c_i$  est de classe  $\mathcal{C}^2$ , on comprend alors en appliquant la formule de Taylor que la dérivée seconde de  $c_i$  contient une information importante sur la structure de corrélation du processus. On fera ici l'hypothèse suivante :

(A<sub>1</sub>) Pour tout  $i \in \{1, \dots, d\}$ ,  $c_i$  est continue sur  $\mathbb{R}_+$  et  $\mathcal{C}^2$  sur  $(0, +\infty)$ . De plus, il existe  $a_i \in (0, 2)$  et  $b_i > 0$  tels que :

$$|c_i''(t)| \leq \begin{cases} Ct^{-a_i} & \forall t \in (0, 1) \\ Ct^{-b_i} & \forall t \geq 1. \end{cases} \quad (4.19)$$

Lorsque  $(Z_t^i)_{t \geq 0}$  est un fBm d'indice de Hurst  $H \in (0, 1)$ , ces hypothèses sont satisfaites avec  $a_i = b_i = 2 - 2H$ . La quantité  $b_i$  indique la vitesse de décorrélation des accroissements. En particulier, pour tout  $t_0 > 0$ ,  $\mathbb{E}[Z_{t_0}^i (Z_{t+t_0}^i - Z_{t_0}^i)] \rightarrow 0$  lorsque  $t \rightarrow +\infty$ . Via le théorème de Maruyama (voir [68, 97]), on en déduit alors que les accroissements du  $(Z_t^i)_{t \geq 0}$  sont ergodiques, *i.e.*

$$\frac{1}{t} \int_0^t \delta_{Z_{s+}^i, -Z_s^i} ds \xrightarrow{t \rightarrow \infty} \mathcal{L}((Z_t^i)_{t \geq 0}) \quad p.s.,$$

ce type de convergence étant aussi valable pour des moyennes discrètes.

D'autre part,  $a_i$  est une quantité reliée aux variations locales du processus. En particulier, (4.19) implique que dans un voisinage de 0,

$$c_i(t) \leq C \begin{cases} t & \text{si } a_i \in (0, 1), \\ t \ln t & \text{si } a_i = 1, \\ t^{2-a_i} & \text{si } a_i \in (1, 2). \end{cases} \quad (4.20)$$

Naturellement, plus  $a_i$  est grand et plus la variance "croît" fortement au voisinage de 0.

Pour tout  $\gamma > 0$ , on considère pour cette équation la suite empirique  $(\mathcal{P}^{(n, \gamma)}(\omega, d\alpha))_{n \geq 1}$  construite de manière identique à la section précédente mais cette fois basée sur le schéma d'Euler (à temps continu) de pas  $\gamma$  associé à (4.18). On travaillera de plus ici sur l'espace  $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$  des fonctions continues de  $\mathbb{R}_+$  dans  $\mathbb{R}^d$  muni de la topologie de la convergence uniforme sur les compacts.

Les notions de solutions stationnaires dans les cas continus et discrétisés (sur l'espace  $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ ) restent identiques à la section précédente. L'adaptation peut également être définie de manière identique (en remplaçant  $B^H$  par  $Z$ ).

On introduit également une hypothèse de Lyapounov similaire à la section précédente (dans le cas particulier  $V(x) = 1 + |x|^2$ ) :

**(A<sub>2</sub>)** :

- (i) Il existe  $C > 0$  tel que  $|b(x)| \leq C(1 + |x|) \quad \forall x \in \mathbb{R}^d$ .
- (ii) Il existe  $\beta \in \mathbb{R}$  et  $\alpha > 0$  tels que

$$\langle x, b(x) \rangle \leq \beta - \alpha |x|^2.$$

En revanche, dans ce cadre additif mais beaucoup plus général du point de vue du bruit directeur, on s'appuie ici sur une hypothèse supplémentaire qui joue un rôle important lorsque le pas de discrétisation tend vers 0 :

**(A<sub>3</sub>)** :  $b$  est lipschitzienne et il existe  $\alpha > 0$  et  $\beta \geq 0$  tels que  $\forall x, y \in \mathbb{R}^d$ ,

$$\langle b(x) - b(y), x - y \rangle \leq \beta - \alpha |x - y|^2.$$

**Remarque 4.15** On note que **(A<sub>3</sub>)** implique **(A<sub>2</sub>)**(ii). L'intérêt de ce type d'hypothèse (de contraction forte en dehors d'un compact) est de pouvoir contrôler la distance entre une solution de (4.18) et le processus d'Ornstein-Uhlenbeck  $(Y_t)_{t \geq 0}$  solution de  $dY_t = -Y_t dt + dZ_t$  (au sens continu ou discrétisé), pour lequel il est possible de faire des calculs fins. Plus

exactement, si l'on note  $(\bar{X}_{n\gamma})_{n \geq 1}$  et  $(\Sigma_{n\gamma})_{n \geq 1}$ , les schémas associés respectivement à (4.18) et à  $dY_t = -Y_t dt + dZ_t$ , alors, sous  $(\mathbf{A}_2)$ ,

$$\sup_{\gamma \in (0, \gamma_0)} \limsup_{n \rightarrow +\infty} \mathbb{E}[|\bar{X}_{n\gamma} - \Sigma_{n\gamma}|^2] < +\infty \quad (\gamma_0 > 0).$$

Cette propriété implique à son tour que pour toutes lois stationnaires marginales  $\nu^\gamma$  et  $\mu^\gamma$  associées respectivement aux schémas  $(\bar{X}_{n\gamma})_{n \geq 0}$  et  $(\Sigma_{n\gamma})_{n \geq 0}$ ,  $\sup_{\gamma \in (0, \gamma_0)} |\nu_\gamma(|x|^2) - \mu_\gamma(|x|^2)| < C < +\infty$  (où  $C$  ne dépend pas des lois stationnaires considérées). Or, à l'aide de la structure gaussienne du Ornstein-Uhlenbeck, on peut montrer, via un calcul quasi-explicite que sous  $(\mathbf{A}_1)$   $\sup_{\gamma \in (0, \gamma_0)} \mu_\gamma(|x|^2) < C$ , ce qui permet de conclure que  $\sup_{\gamma \in (0, \gamma_0)} (\nu_\gamma(|x|^2)) < 2C$  (où  $C$  ne dépend pas de la loi stationnaire considérée). Par conséquent, si l'on note  $\mathcal{U}_\gamma$  l'ensemble des lois stationnaires marginales associées à l'équation discrétisée de pas  $\gamma$ ,  $(\mathcal{U}_\gamma)_{\gamma \in (0, \gamma_0)}$  est tendu. Ce point est fondamental pour l'obtention du résultat principal de cette section énoncé ci-dessous :

**Théorème 4.16 ([17])** 1. *Supposons  $(\mathbf{A}_1)$  et  $(\mathbf{A}_2)$  satisfaites. Alors, il existe  $\gamma_0 > 0$  tel que pour  $\gamma \in (0, \gamma_0)$ ,  $(\mathcal{P}^{(n, \gamma)}(\omega, d\alpha))_{n \geq 1}$  est p.s. tendue sur  $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ . De plus, pour tout  $\gamma \in (0, \gamma_0)$ , toute valeur d'adhérence  $\mathbf{P}^\gamma$  est solution stationnaire de l'équation discrétisée (de pas  $\gamma$ ).*

2. *Supposons de plus  $(\mathbf{A}_3)$  satisfaite et pour tout  $\gamma \in (0, \gamma_0)$ , notons  $\mathcal{U}^{\infty, \gamma}(\omega)$  l'ensemble p.s. tendu des valeurs d'adhérence de  $(\mathcal{P}^{(n, \gamma)}(\omega, d\alpha))_{n \geq 1}$ . Il existe  $\gamma_1 \in (0, \gamma_0)$  tel que  $(\mathcal{U}^{\infty, \gamma}(\omega))_{\gamma \leq \gamma_1}$  est p.s. tendu sur  $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ . De plus, toute valeur d'adhérence lorsque  $\gamma \rightarrow 0$  de  $(\mathcal{U}^{\infty, \gamma}(\omega))_{\gamma \leq \gamma_1}$  est une solution stationnaire de (4.18).*

Contrairement au théorème 4.13, cet énoncé ne précise pas si les solutions stationnaires sont adaptées. Ce point sera abordé plus bas. D'autre part, on fait ici le choix de présenter un énoncé sans unicité de la loi stationnaire. Naturellement, les résultats ci-dessus se transforment en propriétés de convergence lorsque l'on a unicité. Via des techniques de confluence de trajectoires, on peut facilement montrer que l'unicité est satisfaite lorsque  $(\mathbf{A}_3)$  est vérifiée avec  $\beta = 0$ . Comme on l'a déjà mentionné, ce cadre est néanmoins restrictif. Lorsque  $(Z_t)_{t \in \mathbb{R}}$  est un processus gaussien non dégénéré, il est tout à fait raisonnable de penser que l'unicité de solution stationnaire adaptée puisse être obtenue via une adaptation des méthodes de type "Feller fort+irréductibilité topologique" présentés dans l'introduction du chapitre. Néanmoins, de tels résultats n'existent pas encore dans la littérature.

**Adaptation.** Comme on l'a déjà mentionné, il est a priori possible de construire des solutions stationnaires physiquement absurdes. Supposons par exemple que  $(Z_t)_{t \geq 0}$  soit un mouvement brownien standard. Alors,  $(X_t)_{t \geq 0}$  défini par  $X_t = -\int_t^{+\infty} e^{t-s} dZ_s$  est une solution stationnaire de  $dX_t = X_t dt + dZ_t$  dont la condition initiale  $X_0 = -\int_0^{+\infty} e^{-s} dZ_s$  n'est clairement pas adaptée (puisqu'elle dépend de tout le futur du processus de bruit). On remarque que cet exemple (qui n'est autre que le retourné en temps du processus d'Ornstein-Uhlenbeck stationnaire) ne satisfait pas l'hypothèse  $(\mathbf{A}_2)$ . Plus exactement, son caractère explosif lorsqu'il est issu d'une condition initiale déterministe (par exemple) implique qu'une telle solution ne peut être obtenue comme limite de mesures d'occupations issues d'un point initial déterministe.

Une preuve rigoureuse du caractère adapté des éléments de  $(\mathcal{U}^{\infty, \gamma}(\omega))$  et des valeurs d'adhérence  $((\mathcal{U}^{\infty, \gamma}(\omega)))_\gamma$  lorsque  $\gamma \rightarrow 0$  nécessite néanmoins un minimum d'information sur la

dynamique du processus du bruit. Dans la section 6.2 de [17], on démontre qu'une telle propriété est satisfaite lorsque la suite  $(\Delta_n^\gamma := Z_{n\gamma} - Z_{(n-1)\gamma})_{n \in \mathbb{Z}}$  (ou  $(Z_t)_{t \geq 0}$  dans le cas continu) admet une représentation en moyenne mobile (voir (4.4)). L'adaptation est alors obtenue en exploitant la construction moyennisée et des méthodes de martingale. Rappelons pour terminer que la représentation en moyenne mobile est valable si le processus est purement non déterministe (*i.e.* ne contient pas de partie dont la dynamique est prévisible). Cette restriction n'en est en fait pas vraiment une car dans le cas général, on sait d'après la décomposition de Wold qu'un tel processus gaussien peut être décomposé comme la somme de deux processus indépendants respectivement purement déterministe et non déterministe. Le "futur" du processus coïncide alors avec celui de la partie purement non déterministe et on est ramené au cas précédent.

**Dépendance entre  $X_0$  et  $(Z_t)_{t \geq 0}$ .** Comme on l'a déjà mentionné, une condition initiale d'une solution stationnaire adaptée est la donnée d'un couple  $(X_0, (Z_t)_{t \leq 0})$ . Sous l'influence des habitudes markoviennes, on peut être malgré tout tenté de se poser la question de savoir s'il s'agit réellement d'un couplage. Supposons donc que  $X_0$  soit indépendant de  $(Z_t)_{t \leq 0}$ . Dans ce cas,  $X_0$  est également indépendant de  $(Z_t)_{t \geq 0}$  et on a la proposition suivante :

**Proposition 4.17 ([17])** *Soit  $(X_t)_{t \geq 0}$  une solution stationnaire (pas nécessairement adaptée) de (4.18). Si  $X_0$  est indépendant de  $(Z_t)_{t \geq 0}$ , alors  $(Z_t)_{t \geq 0}$  est à accroissements indépendants.  $Z$  étant à accroissements stationnaires, on a par conséquent :  $Z = QW$  où  $W$  est un  $d$ -mouvement brownien standard et  $Q$  est une matrice déterministe.*

## 4.5 Quelques perspectives

A propos de la vitesse de convergence à l'équilibre. Le problème de la vitesse de convergence en loi (marginale ou fonctionnelle) vers la loi stationnaire est un problème où de nombreuses questions sont ouvertes. Concernant le problème particulier des EDS multiplicatives, il serait intéressant dans un premier temps de tenter, dans le cas  $H > 1/2$  de s'abstraire de l'hypothèse : " $\sigma^{-1}$  est une matrice jacobienne". Dans les techniques de preuve développées dans [32], le seul verrou est la question du couplage des positions à coût raisonnable (cf Étape 1 du couplage). Plus précisément, il faut être capable de coller les trajectoires à l'aide d'un terme de drift (dont l'intégrande a une régularité hôlderienne d'exposant adéquat). Ce problème est actuellement en cours de résolution, l'idée principale étant de baser la construction de ce terme de drift sur la famille de processus tangents associés à l'équation. Dans le cas  $H < 1/2$ , il semble alors que la difficulté principale soit de réussir à adapter les arguments de type Lyapounov permettant d'assurer la  $(K, \alpha)$ -admissibilité.

La question plus générale est celle de l'ordre "réel" de la vitesse. Au vu de la structure de dépendance du fBm, il semble raisonnable de penser que la méthode de couplage jusqu'à l'infini utilisée ici génère une vitesse réelle de type polynômial mais il serait néanmoins intéressant d'identifier le bon exposant. Si l'on considère maintenant le problème du point de vue marginal, *i.e.* celui de la distance de  $(\mathcal{L}(X_t^x)$  à  $\bar{\mu}$ ) (où  $\bar{\mu}$  est la loi stationnaire marginale), on peut aussi penser que l'approche de type "collage jusqu'à l'infini" est sans doute trop coûteuse pour identifier l'ordre réel de la vitesse associée. Par exemple, via cette approche, on ne peut a priori pas capturer la vitesse du brownien fractionnaire sur la sphère, qui est explicite. En effet, si  $X_t = e^{iB_t^H}$ , on peut facilement montrer, en utilisant un scaling et

la vitesse exponentielle de convergence du brownien standard vers la loi uniforme  $\lambda_S$  sur la sphère, que  $\|\mathcal{L}(X_t) - \lambda_S\|$  est de l'ordre de  $\exp(-t^{2H})$ . Notons que bien qu'il ne soit pas du tout clair que cet ordre soit représentatif, il a cependant l'intérêt de mettre en avant la dépendance de la vitesse à  $H$ .

Parmi les projets actuels pour tenter de mieux capter l'ordre de la vitesse, une idée serait de regarder le problème en distance de Wasserstein dans le cas simple unidimensionnel  $dX_t = -U'(X_t) + dB_t^H$  où  $U$  est un potentiel à fonds plat, strictement convexe en dehors d'un compact. Lorsque les fBms de chaque trajectoire sont égaux, les trajectoires sont confluentes. La vitesse de confluence est alors fonction du temps passé en dehors du compact. Dans ce cas particulier, l'intérêt est de pouvoir ramener le problème à celui des temps de retour de l'EDS fractionnaire dans les compacts.

**A propos des méthodes d'approximation.** Outre l'extension aux "vraies" EDS fractionnaires d'ordre  $H < 1/2$ , la question naturelle est celle de la vitesse de convergence de ces mesures d'occupation vers le régime stationnaire. Dans le cas où  $\sigma$  est constant, il existe une fonction pour laquelle la vitesse est explicitable : il s'agit du drift  $b$ . En effet,

$$t^{1-H} \left( \frac{1}{t} \int_0^t b(X_s) ds \right) = t^{1-H} \left( \frac{X_t - X_0}{t} - \sigma \frac{B_t^H}{t} \right) \implies \mathcal{N}(0, \sigma^2).$$

Dans le cas général, un résultat de normalité asymptotique de ce type semble inenvisageable. En revanche, d'après les récents résultats de [90], il est possible de prouver des résultats de concentration pour la mesure d'occupation (via l'obtention d'inégalités de Talagrand pour la loi des solutions d'EDS fractionnaires). Ainsi, au moins dans le cas continu, on en déduirait une estimation (non-asymptotique) de la distance entre  $1/t \int_0^t g(X_s) ds$  et  $\bar{\mu}(g)$  en combinant ce résultat avec un résultat de convergence à l'équilibre. Dans le cas où le drift est suffisamment contractant pour assurer une convergence exponentielle à l'équilibre, on retrouve a priori une erreur en  $t^{H-1}$ . Il semble donc que cette approche soit raisonnable pour envisager un contrôle de l'erreur entre la procédure d'approximation (discrétisée) et la loi stationnaire. Notons pour terminer que les méthodes d'approximation du régime stationnaire ont des liens assez forts avec le problème de l'estimation du drift, qui se base généralement sur une observation en temps long de la trajectoire. Ainsi, il semble que les outils développés ici puissent jouer un rôle important dans ce type d'étude (dans le cas fractionnaire). Ce dernier point fait aussi partie des projets actuels (en collaboration avec Bruno Saussereau et Samy Tindel).



## Bibliographie

- [1] A. Alfonsi, B. Jourdain, and A. Kohatsu-Higa. Pathwise optimal transport bounds between a one-dimensional diffusion and its Euler scheme. *Ann. Appl. Probab.*, 24(3) :1049–1080, 2014.
- [2] Ludwig Arnold. *Random dynamical systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.
- [3] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi*, volume 17 of *Mathématiques et applications*. Springer-Verlag, 1994.
- [4] G. K. Basak and R. N. Bhattacharya. Stability in distribution for a class of singular diffusions. *Ann. Probab.*, 20(1) :312–321, 1992.
- [5] Fabrice Baudoin. Bakry-emery meet villani. *Preprint available at <http://arxiv-web3.library.cornell.edu/pdf/1308.4938.pdf>*, pages 1–38, 2013.
- [6] P. H. Baxendale and D. W. Stroock. Large deviations and stochastic flows of diffeomorphisms. *Probab. Theory Related Fields*, 80(2) :169–215, 1988.
- [7] P.H. Baxendale. Statistical equilibrium and two-point motion for a stochastic flow of diffeomorphisms. In *Spatial stochastic processes*, volume 19 of *Progr. Probab.*, pages 189–218. Birkhäuser Boston, Boston, MA, 1991.
- [8] R. N. Bhattacharya. On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Z. Wahrsch. Verw. Gebiete*, 60(2) :185–201, 1982.
- [9] A. Cabot. Asymptotics for a gradient system with memory term. *Proc. of the American Mathematical Society*, 9 :3013–3024, 2009.
- [10] A. Cabot, H. Engler, and S. Gadat. On the long time behavior of second order differential equations with asymptotically small dissipation. *Trans. of the American Mathematical Society*, 361 :5983–6017, 2009.
- [11] A. Carverhill. Flows of stochastic dynamical systems : ergodic theory. *Stochastics*, 14(4) :273–317, 1985.
- [12] P. Cattiaux. Stochastic calculus and degenerate boundary value problems. *Ann. Inst. Fourier (Grenoble)*, 42(3) :541–624, 1992.
- [13] Patrick Cattiaux, Pierre Collet, Amaury Lambert, Servet Martínez, Sylvie Méléard, and Jaime San Martín. Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Probab.*, 37(5) :1926–1969, 2009.
- [14] Nicolas Champagnat and Denis Villemonais. Exponential convergence to quasi-stationary distribution and q-process. *Preprint available at <http://hal.inria.fr/hal-00973509>*, pages 1–46, 2014.

- [15] M. F. Chen and S. F. Li. Coupling methods for multidimensional diffusion processes. *Ann. Probab.*, 17(1) :151–177, 1989.
- [16] Patrick Cheridito, Hideyuki Kawaguchi, and Makoto Maejima. Fractional Ornstein-Uhlenbeck processes. *Electron. J. Probab.*, 8 :no. 3, 14 pp. (electronic), 2003.
- [17] Serge Cohen and Fabien Panloup. Approximation of stationary solutions of Gaussian driven stochastic differential equations. *Stochastic Process. Appl.*, 121(12) :2776–2801, 2011.
- [18] Serge Cohen, Fabien Panloup, and Samy Tindel. Approximation of stationary solutions to SDEs driven by multiplicative fractional noise. *Stochastic Process. Appl.*, 124(3) :1197–1225, 2014.
- [19] Pierre Collet, Servet Martínez, and Jaime San Martín. *Quasi-stationary distributions*. Probability and its Applications (New York). Springer, Heidelberg, 2013. Markov chains, diffusions and dynamical systems.
- [20] Laure Coutin. Rough paths via sewing lemma. To appear in ESAIM PS, 2012.
- [21] Hans Crauel. Non-Markovian invariant measures are hyperbolic. *Stochastic Process. Appl.*, 45(1) :13–28, 1993.
- [22] G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems*, volume 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.
- [23] A. S. Dalalyan and A. B. Tsybakov. Sparse regression learning by aggregation and Langevin Monte-Carlo. *J. Comput. System Sci.*, 78(5) :1423–1443, 2012.
- [24] A. M. Davie. Differential equations driven by rough paths : an approach via discrete approximation. *Appl. Math. Res. Express. AMRX*, (2) :Art. ID abm009, 40, 2007.
- [25] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.
- [26] L. Desvillettes and C. Villani. On the trend to global equilibrium for spatially inhomogeneous kinetic systems : the Boltzmann equation. *Invent. Math.*, 159(2) :245–316, 2005.
- [27] A. Deya, A. Neuenkirch, and S. Tindel. A Milstein-type scheme without Lévy area terms for SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.*, 48(2) :518–550, 2012.
- [28] D. Dolgopyat, V. Kaloshin, and L. Korolov. Sample path properties of the stochastic flows. *Ann. Probab.*, 32(1A) :1–27, 2004.
- [29] Randal Douc, Gersende Fort, and Arnaud Guillin. Subgeometric rates of convergence of  $f$ -ergodic strong Markov processes. *Stochastic Process. Appl.*, 119(3) :897–923, 2009.
- [30] D. Down, S. P. Meyn, and R. L. Tweedie. Exponential and uniform ergodicity of Markov processes. *Ann. Probab.*, 23(4) :1671–1691, 1995.
- [31] Jin Feng and Thomas G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2006.



- [32] Joaquin Fontbona and Fabien Panloup. Rate of convergence to equilibrium of fractional driven stochastic differential equations with some multiplicative noise. *Preprint available at <http://hal.archives-ouvertes.fr/hal-00989414>*, pages 1–34, 2014.
- [33] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, second edition, 1979. Translated from the 1979 Russian original by Joseph Szücs.
- [34] Sébastien Gadat, Laurent Miclo, and Fabien Panloup. A stochastic model for speculative bubbles. *Preprint available at <http://hal.archives-ouvertes.fr/hal-00937447>*, pages 1–50, 2013.
- [35] Sébastien Gadat and Fabien Panloup. Long time behaviour and stationary regime of memory gradient diffusions. *Annales de l'Institut Henri Poincaré*, to appear, 50 :564–601, 2014.
- [36] Sébastien Gadat, Fabien Panloup, and Clément Pellegrini. Large deviation principle for invariant distributions of memory gradient diffusions. *Electron. J. Probab.*, 18 :no. 81, 34, 2013.
- [37] P. M. Garber. Famous first bubbles. *Journal of Economic Perspectives*, 4 :35–54, 1990.
- [38] María J. Garrido-Atienza, Peter E. Kloeden, and Andreas Neuenkirch. Discretization of stationary solutions of stochastic systems driven by fractional Brownian motion. *Appl. Math. Optim.*, 60(2) :151–172, 2009.
- [39] María J. Garrido-Atienza, Peter E. Kloeden, and Andreas Neuenkirch. Discretization of stationary solutions of stochastic systems driven by fractional Brownian motion. *Appl. Math. Optim.*, 60(2) :151–172, 2009.
- [40] Paolo Guasoni. No arbitrage under transaction costs, with fractional Brownian motion and beyond. *Math. Finance*, 16(3) :569–582, 2006.
- [41] M. Hairer and N. S. Pillai. Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(2) :601–628, 2011.
- [42] Martin Hairer. Ergodicity of stochastic differential equations driven by fractional Brownian motion. *Ann. Probab.*, 33(2) :703–758, 2005.
- [43] Martin Hairer. Ergodic properties of a class of non-Markovian processes. In *Trends in stochastic analysis*, volume 353 of *London Math. Soc. Lecture Note Ser.*, pages 65–98. Cambridge Univ. Press, Cambridge, 2009.
- [44] Martin Hairer and Alberto Ohashi. Ergodic theory for SDEs with extrinsic memory. *Ann. Probab.*, 35(5) :1950–1977, 2007.
- [45] Martin Hairer and Natesh S. Pillai. Regularity of laws and ergodicity of hypoelliptic SDEs driven by rough paths. *Ann. Probab.*, 41(4) :2544–2598, 2013.
- [46] D. Hamadouche and C. Suquet. Weak Hölder convergence of processes with application to the perturbed empirical process. *Appl. Math. (Warsaw)*, 26(1) :63–83, 1999.
- [47] T. E. Harris. Brownian motions on the homeomorphisms of the plane. *Ann. Probab.*, 9(2) :232–254, 1981.
- [48] R. Z. Has'minskiĭ. *Stochastic stability of differential equations*, volume 7 of *Monographs and Textbooks on Mechanics of Solids and Fluids : Mechanics and Analysis*. Sijthoff & Noordhoff, Alphen aan den Rijn, 1980. Translated from the Russian by D. Louvish.

- [49] Takeyuki Hida and Masuyuki Hitsuda. *Gaussian processes*, volume 120 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1993. Translated from the 1976 Japanese original by the authors.
- [50] Chii-Ruey Hwang. Laplace's method revisited : weak convergence of probability measures. *Ann. Probab.*, 8(6) :1177–1182, 1980.
- [51] N. Ikeda and S. Watanabe. A comparison theorem for solutions of stochastic differential equations and its applications. *Osaka J. Math.*, 14(3) :619–633, 1977.
- [52] Jae-Hyung Jeon, Vincent Tejedor, Stas Burov, Eli Barkai, Christine Selhuber-Unkel, Kirstine Berg-Sørensen, Lene Oddershede, and Ralf Metzler. *In Vivo* anomalous diffusion and weak ergodicity breaking of lipid granules. *Phys. Rev. Lett.*, 106 :048103, Jan 2011.
- [53] R.Z. Khas'minskiĭ. *Stochastic stability of differential equations*. Sijthoff & Noordhoff, Alphen aan den Rijn (The Netherlands), 1980.
- [54] A. Kiselev and L. Ryzhik. A simple model for asset price bubble formation and collapse. *Technical Report*, pages 1–30, 2010.
- [55] S. C. Kou. Stochastic modeling in nanoscale biophysics : subdiffusion within proteins. *Ann. Appl. Stat.*, 2(2) :501–535, 2008.
- [56] H. Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.
- [57] Olga A. Ladyzhenskaya and Nina N. Ural'tseva. *Linear and quasilinear elliptic equations*. Translated from the Russian by Scripta Technica, Inc. Translation editor : Leon Ehrenpreis. Academic Press, New York-London, 1968.
- [58] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion. *Bernoulli*, 8(3) :367–405, 2002.
- [59] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion : the case of a weakly mean reverting drift. *Stoch. Dyn.*, 3(4) :435–451, 2003.
- [60] B Lapeyre and E Temam. Competitive monte carlo methods for the pricing of asian options. *Comp. Finance*, 5(1) :39–59, 2001.
- [61] Kim-Anh Lê Cao, Olivier Gonçalves, Philippe Besse, and Sébastien Gadat. Selection of biologically relevant genes with a wrapper stochastic algorithm. *Stat. Appl. Genet. Mol. Biol.*, 6 :Art. 29, 23, 2007.
- [62] V. Lemaire. *Estimation récursive de la mesure invariante d'un processus de diffusion*. Thèse de doctorat, Université de Marne-la-Vallée (France), 2005.
- [63] V. Lemaire. *Estimation récursive de la mesure invariante d'un processus de diffusion*. Thèse de doctorat, Université de Marne-la-Vallée (France), 2005.
- [64] Vincent Lemaire. An adaptive scheme for the approximation of dissipative systems. *Stochastic Process. Appl.*, 117(10) :1491–1518, 2007.
- [65] Vincent Lemaire, Gilles Pagès, and Fabien Panloup. Invariant distribution of duplicated diffusions and application to richardson-romberg extrapolation. *Ann. Inst. Henri Poincaré Probab. Stat.*, to appear.

- [66] Vincent Lemaire and Gilles Pagès. Multilevel richardson-romberg extrapolation. *Preprint available at <http://arxiv.org/abs/1401.1177>*, pages 1–20, 2014.
- [67] Benoit B. Mandelbrot and John W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 10 :422–437, 1968.
- [68] Gisirō Maruyama. The harmonic analysis of stationary stochastic processes. *Mem. Fac. Sci. Kyūsyū Univ. A.*, 4 :45–106, 1949.
- [69] Laurent Miclo. Recuit simulé sur  $\mathbf{R}^n$ . Étude de l'évolution de l'énergie libre. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques*, 28(2) :235–266, 1992.
- [70] Pierre Monmarché. Hypocoercive relaxation to equilibrium for some kinetic models. *Kinet. Relat. Models*, 7(2) :341–360, 2014.
- [71] A. Neuenkirch, S. Tindel, and J. Unterberger. Discretizing the fractional Lévy area. *Stochastic Process. Appl.*, 120(2) :223–254, 2010.
- [72] Andreas Neuenkirch and Ivan Nourdin. Exact rate of convergence of some approximation schemes associated to SDEs driven by a fractional Brownian motion. *J. Theoret. Probab.*, 20(4) :871–899, 2007.
- [73] Ivan Nourdin. Schémas d'approximation associés à une équation différentielle dirigée par une fonction höldérienne; cas du mouvement brownien fractionnaire. *C. R. Math. Acad. Sci. Paris*, 340(8) :611–614, 2005.
- [74] David Nualart and Aurel Răşcanu. Differential equations driven by fractional Brownian motion. *Collect. Math.*, 53(1) :55–81, 2002.
- [75] David J. Odde, Elly M. Tanaka, Stacy S. Hawkins, and Helen M. Buettner. Stochastic dynamics of the nerve growth cone and its microtubules during neurite outgrowth. *Biotechnology and Bioengineering*, 50(4) :452–461, 1996.
- [76] G. Pagès and F. Panloup. Approximation of the distribution of a stationary Markov process with application to option pricing. *Bernoulli*, 15(1) :146–177, 2009.
- [77] Gilles Pagès. Multi-step Richardson-Romberg extrapolation : remarks on variance control and complexity. *Monte Carlo Methods Appl.*, 13(1) :37–70, 2007.
- [78] Gilles Pagès and Fabien Panloup. Ergodic approximation of the distribution of a stationary diffusion : rate of convergence. *Ann. Appl. Probab.*, 22(3) :1059–1100, 2012.
- [79] Gilles Pagès and Fabien Panloup. A mixed-step algorithm for the approximation of the stationary regime of a diffusion. *Stochastic Process. Appl.*, 124(1) :522–565, 2014.
- [80] F. Panloup. *Approximation récursive du régime stationnaire d'une Équation Différentielle Stochastique avec sauts*. Thèse de doctorat, Université Pierre et Marie Curie (Paris 6, France), 2006.
- [81] F. Panloup. Computation of the invariant measure for a Lévy driven SDE : Rate of convergence. *Stochastic Processes and their Applications*, 118(8) :1351–1384, 2008.
- [82] F. Panloup. Recursive computation of the invariant measure of a stochastic differential equation driven by a Lévy process. *Annals of Applied Probability*, 18(2) :379–426, 2008.
- [83] E. Pardoux and A. Yu. Veretennikov. On the Poisson equation and diffusion approximation. I. *Ann. Probab.*, 29(3) :1061–1085, 2001.

- [84] È. Pardoux and A. Yu. Veretennikov. On Poisson equation and diffusion approximation. II. *Ann. Probab.*, 31(3) :1166–1192, 2003.
- [85] E. Pardoux and A. Yu. Veretennikov. On the Poisson equation and diffusion approximation. III. *Ann. Probab.*, 33(3) :1111–1133, 2005.
- [86] Philip Protter. A mathematical theory of financial bubbles. In *Paris-Princeton Lectures on Mathematical Finance 2013*, volume 2081 of *Lecture Notes in Mathematics*, pages 1–108. Springer, 2013.
- [87] Anatolii Puhalskii. On large deviation convergence of invariant measures. *J. Theoret. Probab.*, 16(3) :689–724, 2003.
- [88] Alfredas Račkauskas and Charles Suquet. Central limit theorem in Hölder spaces. *Probab. Math. Statist.*, 19(1, Acta Univ. Wratislav. No. 2138) :133–152, 1999.
- [89] Sorin Rădulescu and Marius Rădulescu. An application of Hadamard-Lévy’s theorem to a scalar initial value problem. *Proc. Amer. Math. Soc.*, 106(1) :139–143, 1989.
- [90] Bruno Sausseureau. Transportation inequalities for stochastic differential equations driven by a fractional Brownian motion. *Bernoulli*, 18(1) :1–23, 2012.
- [91] J. Scheinkman and W. Xiong. Overconfidence and speculative bubbles. *Journal of Political Economy*, 6 :1183–1219, 2003.
- [92] Robert J. Shiller. *Irrational Exuberance*. Princeton University Press, 3 Market place, Woodstock, Oxfordshire, 2000.
- [93] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4) :483–509 (1991), 1990.
- [94] Denis Talay. Second order discretization schemes of stochastic differential systems for the computation of the invariant law. *Stoch. Stoch. Rep.*, 29(1) :13–36, 1990.
- [95] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
- [96] Cédric Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950) :iv+141, 2009.
- [97] Michel Weber. Sur un théorème de Maruyama. In *Seminar on Probability, XIV (Paris, 1978/1979) (French)*, volume 784 of *Lecture Notes in Math.*, pages 475–488. Springer, Berlin, 1980.
- [98] Andrew T. A. Wood and Grace Chan. Simulation of stationary Gaussian processes in  $[0, 1]^d$ . *J. Comput. Graph. Statist.*, 3(4) :409–432, 1994.