



**HAL**  
open science

## Estimation récursive de la mesure invariante d'un processus de diffusion.

Vincent Lemaire

► **To cite this version:**

Vincent Lemaire. Estimation récursive de la mesure invariante d'un processus de diffusion.. Mathématiques [math]. Université de Marne la Vallée, 2005. Français. NNT: . tel-00011281

**HAL Id: tel-00011281**

**<https://theses.hal.science/tel-00011281>**

Submitted on 3 Jan 2006

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

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

THÈSE DE DOCTORAT  
DE L'UNIVERSITÉ DE MARNE-LA-VALLÉE

présentée par  
Vincent LEMAIRE

pour obtenir le grade de  
DOCTEUR DE L'UNIVERSITÉ DE MARNE-LA-VALLÉE  
*Spécialité : Mathématiques Appliquées*

---

Estimation récursive de la mesure invariante  
d'un processus de diffusion

---

Soutenu le 08 décembre 2005  
devant le jury composé des professeurs :

Damien LAMBERTON	Directeur de thèse
Dominique LÉPINGLE	Rapporteur
Gilles PAGÈS	Directeur de thèse
Bernard ROYNETTE	Examineur
Christian SOIZE	Examineur
Denis TALAY	Rapporteur



*À Jacques,  
À Marcel.*



# Remerciements

En premier lieu, je tiens à remercier mes directeurs de thèse Damien Lambertson et Gilles Pagès. Merci à Gilles Pagès qui a su me guider et me faire découvrir les probabilités numériques dès ma maîtrise lors de mon travail d'étude et de recherche sur les réseaux de neurones artificiels et les algorithmes stochastiques. Il m'a réitéré sa confiance en acceptant d'encadrer ma thèse et je lui en suis très reconnaissant. Merci aussi à Damien Lambertson qui a été d'un véritable soutien pendant ces quatre années et sans qui cette expérience n'aurait pu avoir lieu. Ce travail leur doit beaucoup.

Je souhaite également remercier Dominique Lépingle et Denis Talay pour avoir examiné ce mémoire. Leurs travaux sont pour moi une grande source d'inspiration et je suis très honoré de leur présence dans le jury.

J'exprime toute ma gratitude à Bernard Roynette et Christian Soize qui ont également accepté d'être membres du jury. Je remercie particulièrement Christian Soize de m'avoir encouragé à poursuivre l'étude de mon algorithme consacré aux systèmes hamiltoniens perturbés.

Merci aussi à toute l'équipe du laboratoire d'analyse et de mathématiques appliquées de l'université de Marne-la-Vallée. Je pense notamment à tous les thésards avec qui j'ai passé de très bons moments.

---

# Avant-propos

L'objet de cette thèse est l'étude d'un algorithme, simple d'implémentation et récursif, permettant de calculer l'intégrale d'une fonction par rapport à la probabilité invariante d'un processus solution d'une équation différentielle stochastique de dimension finie.

Dans la première partie, nous introduisons tout d'abord le sujet en rappelant les résultats connus et en exposant brièvement les résultats obtenus durant la thèse. Puis dans le second chapitre, nous présentons la méthode de démonstration, utilisée pour l'étude du schéma d'approximation, dans le cas continu (pour le processus de diffusion lui-même).

Dans la seconde partie, nous nous concentrons sur le schéma d'Euler de pas décroissant étudié par Lamberton et Pagès. Nous étendons dans un premier temps les résultats de convergence aux fonctions exponentielles et pour une classe de diffusion plus importante (tout le travail est fait sous une condition de stabilité faible). Nous proposons aussi une alternative à la méthode de martingales pour prouver la convergence dans le cas des diffusions  $(V, \alpha)$ -confluente. Nous montrons dans le même temps la convergence presque sûre et dans  $\mathbf{L}^p$  du schéma vers la diffusion. L'exposé se poursuit avec une étude précise des vitesses de convergence de l'algorithme, notamment en donnant un développement à tout ordre de l'erreur pour une classe de fonctions tests. Nous finissons l'étude de ce schéma lorsqu'il y a présence de multiples mesures invariantes. Nous donnons le comportement de l'algorithme en dimension 1, et faisons un lien entre classification de Feller et fonctions de Lyapounov.

Dans la troisième partie, nous introduisons un nouvel algorithme adaptatif permettant de considérer des problèmes plus généraux tels que les systèmes Hamiltoniens ou les systèmes monotones. L'algorithme est explicite, récursif et de complexité faible. Il s'agit de considérer les mesures empiriques d'un schéma d'Euler construit à partir d'une suite de pas aléatoires adaptés dominée par une suite décroissant vers 0.



# Table des matières

<b>Partie 1</b>	<b>Introduction</b>	<b>1</b>
<b>I</b>	<b>Introduction</b>	<b>3</b>
<b>II</b>	<b>Stabilité et méthodes de martingales</b>	<b>9</b>
1	Stabilité . . . . .	9
2	Tension des mesures empiriques . . . . .	14
3	Identification de la limite . . . . .	17
4	Conclusion . . . . .	19
5	Annexe . . . . .	20
<b>Partie 2</b>	<b>Étude du schéma d'Euler de pas décroissant</b>	<b>25</b>
<b>III</b>	<b>Intégration de fonctions exponentielles</b>	<b>27</b>
1	Schéma d'Euler de pas décroissant . . . . .	27
2	Stabilité . . . . .	29
3	Mesures empiriques pondérées . . . . .	36
3.1	Tension . . . . .	36
3.2	Identification de la limite . . . . .	40
4	Conclusion . . . . .	43
5	Annexe . . . . .	43
<b>IV</b>	<b>Approche ergodique pour une classe de diffusion</b>	<b>47</b>
1	Diffusion $(V, \alpha)$ -confluente . . . . .	47
2	Convergence en moyenne du schéma d'Euler de pas décroissant . . . . .	50
3	Convergence <i>p.s.</i> et vitesses . . . . .	55
4	Convergence des mesures empiriques . . . . .	58
<b>V</b>	<b>Vitesses de convergence</b>	<b>61</b>
1	Introduction . . . . .	61
2	Résultats préliminaires . . . . .	65
3	Développement de l'erreur . . . . .	68
4	Loi du logarithme itéré . . . . .	75

TABLE DES MATIÈRES

---

5	Simulations numériques . . . . .	80
5.1	Choix des paramètres et commentaires . . . . .	83
5.2	Tests numériques . . . . .	86
6	Extrapolation de Richardson . . . . .	89
6.1	Rappel sur l'extrapolation de Richardson . . . . .	90
6.2	Application de cette méthode à notre algorithme . . . . .	92
7	Annexe . . . . .	96
<b>VI Étude en présence de mesures invariantes multiples</b>		<b>99</b>
1	Diffusion unidimensionnelle . . . . .	100
1.1	Fonction d'échelle et mesure de vitesse . . . . .	100
1.2	Classification de Feller . . . . .	102
1.3	Comportement des mesures empiriques : le point de vue ergodique . . . . .	104
1.4	Attractivité et fonctions de Lyapounov . . . . .	108
2	Comportement de l'algorithme en dimension 1 . . . . .	113
2.1	Schéma d'Euler . . . . .	114
2.2	Mesures empiriques . . . . .	116
3	Exemple numérique en dimension supérieure . . . . .	118
<b>Partie 3 Algorithme adaptatif</b>		<b>121</b>
<b>VII Adaptive scheme</b>		<b>123</b>
1	Introduction . . . . .	123
2	Framework and algorithm . . . . .	124
3	Preliminary results . . . . .	126
4	Convergence of empirical measures . . . . .	128
4.1	A.s. tightness of empirical measures . . . . .	129
4.2	Identification of the limit . . . . .	130
5	Monotone and dissipative problems . . . . .	133
6	Dissipative Hamiltonian systems . . . . .	140
7	Numerical experiments . . . . .	148
7.1	Lorenz equations . . . . .	148
7.2	Perturbed Hamiltonian system . . . . .	150
<b>VIII Étude numérique, choix des paramètres</b>		<b>153</b>
1	Influence du paramètre $\gamma_0$ , extrapolation de Richardson . . . . .	154
2	À propos du bruit blanc $U_n$ . . . . .	160
2.1	Gaussienne ou Bernoulli? . . . . .	160
2.2	Extrapolation de Richardson : même suite $(U_n)_{n \geq 1}$ ? . . . . .	161
3	Estimation des quantités $m_{f,q^*}$ et $s_f^2$ . . . . .	161
4	Réduction de variance . . . . .	163

**Première partie**

**Introduction**



L'approximation du régime stationnaire d'une équation différentielle stochastique (EDS) est nécessaire dans de nombreux problèmes issus de mécanique aléatoire : étude des effets aléatoires du vent, de la houle ou des séismes sur les structures, vibrations d'éléments de transport de fluides turbulents. . . Ces problèmes se ramènent souvent à la résolution de systèmes dynamiques vectoriels de dimension finie, non linéaires, du second ordre et à excitation aléatoire externe, de la forme

$$\begin{cases} \frac{d^2Q}{dt^2}(t) + f\left(Q(t), \frac{dQ}{dt}(t)\right) = g\left(Q(t), \frac{dQ}{dt}(t)\right) \frac{dB(t)}{dt}, & t > 0, \\ Q(0) = p_0 \in \mathbf{R}^d, \quad \text{et} \quad \frac{dQ}{dt}(0) = q_0 \in \mathbf{R}^d, \end{cases} \quad (\text{sys. dyn.})$$

où  $f$  et  $g$  sont des applications de  $\mathbf{R}^d \times \mathbf{R}^d$  dans  $\mathbf{R}^d$ , et  $B$  est le processus de Wiener normalisé à valeurs dans  $\mathbf{R}^d$ . En passant dans l'espace des phases, c'est à dire en introduisant le processus  $x$  tel que  $\forall t \geq 0, x_t = \left(Q(t), \frac{dQ}{dt}(t)\right)$ , l'équation (sys. dyn.) se ramène à une EDS de la forme :

$$\begin{cases} dx_t = b(x_t)dt + \sigma(x_t)dB_t, & t > 0, \\ x(0) = x_0 \in \mathbf{R}^{2d}, \end{cases} \quad (\text{eds})$$

où  $b(x_t) = \left(\frac{dQ}{dt}(t), -f(x_t)\right)$  et  $\sigma(x_t) = (0, g(x_t))$ . Ainsi, le système Hamiltonien (sys. dyn.) se ramène à une EDS non-linéaire dont le coefficient de diffusion est dégénéré. Le but de la thèse est de fournir et d'étudier un algorithme efficace pour déterminer le régime stationnaire de tels systèmes, et d'EDS possédant les mêmes caractéristiques : condition de stabilité (existence d'une fonction de Lyapounov), non-linéarité, possiblement dégénérée.

Il est d'intérêt pour le mécanicien de connaître le comportement asymptotique des solutions stationnaires de (eds). Dans son livre [Soi94], Christian Soize donne des expressions analytiques exactes des densités stationnaires associées à certaines EDS, mais dans de nombreux cas il semble impossible de résoudre explicitement ces problèmes. Un article de synthèse sur les méthodes numériques utilisées dans les problèmes issus de la mécanique aléatoire est [BF94]. Les méthodes de simulation type Monte-Carlo se révèlent très coûteuses numériquement et les méthodes analytiques utilisant la résolution de l'équation de Fokker-Planck associée sont très difficilement implémentables en grande dimension (difficultés encore accrues lorsque le coefficient de diffusion est dégénéré). D'autres méthodes numériques plus spécifiques à un problème donné sont satisfaisantes mais trop restreintes.

Denis Talay a étudié un algorithme (cf. [Tal90]) permettant un calcul approché de la réponse stationnaire le long d'une fonction  $f$  pour une classe assez générale de diffusions. L'approximation

mation de  $\int f \nu$ , où  $\nu$  désigne l'unique mesure invariante, se fait en utilisant l'ergodicité d'un schéma d'approximation  $(X_k^h)_{k=1, \dots, N}$  de  $(x_t)_{t \geq 0}$ . Le cadre de cet article est le suivant :  $b$  et  $\sigma$  sont  $\mathcal{C}^\infty$  de dérivées bornées,  $\sigma$  est bornée, le générateur associé à  $(x_t)_{t \geq 0}$  est uniformément elliptique et la condition de stabilité suivante est vérifiée

$$\exists \alpha > 0, \exists \beta \geq 0, \quad \forall x \in \mathbf{R}^d, \quad \langle x, b(x) \rangle \leq -\alpha |x|^2 + \beta.$$

Sous ces conditions, Talay prouve l'ergodicité du schéma d'Euler ainsi que du schéma de Milshtein et de tout schéma d'ordre 2. Plus précisément, pour un pas de discrétisation  $h > 0$  suffisamment petit, la chaîne de Markov  $(X_k^h)_{k \geq 1}$  est ergodique et possède une unique mesure invariante  $\nu^h$  donc

$$\forall f \in \mathbf{L}^1(\nu^h), \quad \frac{1}{N} \sum_{k=1}^N f(X_k^h) \xrightarrow{p.s.} \int f d\nu^h \quad \text{quand } N \rightarrow +\infty.$$

Il prouve ensuite que pour toute fonction  $f \in \mathcal{C}^\infty(\mathbf{R}^d, \mathbf{R})$  à croissance polynômiale à l'infini ainsi que toutes ses dérivées, l'erreur  $\nu^h(f) - \nu(f)$  est de l'ordre de  $h$  lorsque le schéma est d'Euler ou de Milshtein, et de l'ordre de  $h^2$  lorsque le schéma est d'ordre 2. Plus récemment, cette méthode a été prolongée au cas d'un système Hamiltonien similaire à (syst. dyn.) où  $f$  (donc  $b$ ) n'est pas à dérivée bornée et  $g$  est constante (cf. [Tal02]). Le schéma d'approximation considéré dans ce cas est un schéma d'Euler implicite afin de pallier l'instabilité du schéma d'Euler explicite. D'autre part, dans [MSH02] Mattingly, Stuart et Higham ont montré l'ergodicité de schémas d'approximation implicites dans certains exemples où le drift n'est pas localement Lipschitz, sans étudier la convergence vers la mesure stationnaire limite.

Une autre alternative pour le calcul de  $\nu(f)$  est donnée par Damien Lambertson et Gilles Pagès dans [LP02]. Il s'agit cette fois d'approcher la diffusion par un schéma dont le pas de discrétisation tend vers 0 lorsque le nombre d'itérations (donc le temps) augmente. Ainsi, plus on s'approche du régime stationnaire et plus le schéma est précis. Le schéma est une chaîne de Markov inhomogène, et ne possède donc pas de mesure stationnaire. L'approche précédente ne peut donc pas s'appliquer avec un tel schéma. L'étude est faite lorsque le schéma  $(X_n)_{n \geq 0}$  est un schéma d'Euler construit à partir d'une suite positive  $(\gamma_n)_{n \geq 1}$  telle que  $\sum_{n \geq 1} \gamma_n = +\infty$  et  $\lim_n \gamma_n = 0$  :

$$\forall n \geq 0, \quad X_{n+1} = X_n + \gamma_{n+1} b(X_n) + \sqrt{\gamma_{n+1}} \sigma(X_n) U_{n+1} \quad \text{et} \quad X_0 = x_0. \quad (*)$$

Le bruit blanc  $(U_n)_{n \geq 1}$  a des moments polynômiaux et est normé. Notons qu'une telle approximation de la diffusion a déjà été utilisée par Georg Pflug dans [LPW92]. Il montre, en dimension 1, que  $(X_n)_{n \geq 0}$  converge étroitement vers la mesure invariante  $\nu$  en supposant le coefficient de diffusion borné et  $b$  à dérivée bornée. De même, dans [BHW97] les auteurs prouvent cette convergence étroite dans un cadre multidimensionnel lorsque  $\mathcal{A}$  est uniformément elliptique, mais aussi dans le cas des diffusion  $(V, \alpha)$ -confluentes<sup>1</sup>. Notons aussi que dans un cadre compact, Piccioni et Scarlatti étudient les mesures empiriques d'un tel schéma et obtiennent des résultats de convergence dans  $\mathbf{L}^2$  (cf. [PS94]).

Dans [LP02], le point de vue est de considérer les mesures empiriques pondérées  $(\nu_n^j)_{n \geq 1} =$

---

<sup>1</sup>notion définie dans le chapitre IV appelée "asymptotic flatness" par les auteurs

---

$\left(\frac{1}{\sum_{k=1}^n \eta_k} \sum_{k=1}^n \eta_k \delta_{X_{k-1}}\right)_{n \geq 1}$  du schéma  $(X_n)_{n \geq 0}$  défini par (\*); la suite  $(\eta_n)_{n \geq 1}$  est une suite de poids positive et terme général d'une série divergente. Ils supposent qu'il existe une fonction de Lyapounov  $V$  sous-quadratique <sup>2</sup> telle que

$$\exists \alpha > 0, \exists \beta \geq 0, \quad \forall x \in \mathbf{R}^d, \quad \mathcal{A}V(x) \leq -\alpha V(x) + \beta.$$

que  $b$  et  $\sigma$  sont continues et sous-linéaires<sup>3</sup>, et montrent que

$$\frac{1}{\sum_{k=1}^n \eta_k} \sum_{k=1}^n \eta_k f(X_{k-1}) \xrightarrow{p.s.} \int f d\nu,$$

pour toute fonction  $f$  à croissance polynômiale (si  $\nu$  est l'unique mesure invariante). Ils obtiennent ce résultat en utilisant des techniques de martingales et en identifiant la limite à l'aide du théorème d'Echeverria-Weiss<sup>4</sup>. Il est important de noter que le générateur infinitésimal  $\mathcal{A}$  n'est pas nécessairement uniformément elliptique et qu'il peut exister plusieurs mesures invariantes. Dans ce cas, toute valeur d'adhérence des mesures empiriques de  $(X_n)_{n \geq 0}$  est mesure invariante de  $(x_t)_{t \geq 0}$ . Dans un second papier [LP03], ils étendent leurs résultats à une classe de diffusion plus importante en affaiblissant la condition de rappel sur la fonction de Lyapounov : ils supposent qu'il existe  $V$  sous-quadratique et  $a \in ]0, 1]$  tels que

$$\exists \alpha > 0, \exists \beta \geq 0, \quad \forall x \in \mathbf{R}^d, \quad \mathcal{A}V(x) \leq -\alpha V^a(x) + \beta.$$

C'est cette seconde approche que nous allons étudier et étendre tout au long de la thèse.

## Plan et résultats

**Dans la première partie** (chapitre II), nous introduisons la notion de condition de stabilité (forte et faible) et nous exposons la méthode de martingales pour l'étude des mesures empiriques du processus lui-même, par opposition à l'approche ergodique (cf. Pagès [Pag01]). Toute l'étude se fait donc sur  $(x_t)_{t \geq 0}$  et non sur le schéma de discrétisation. La condition de stabilité faible, que nous proposons, se rapproche de la condition introduite par Douc, Fort, Moulines et Soulier dans le cadre des chaînes de Markov (cf. [DFMS04]). Dans ce cadre très général et sans hypothèse sur la croissance de  $b$  et de  $V$  (fonction de Lyapounov apparaissant dans la condition de stabilité) nous montrons, par la méthode de martingales, d'une part la tension des mesures empiriques de  $(x_t)_{t \geq 0}$  et d'autre part que toute valeur d'adhérence est mesure invariante. Un autre résultat, moins classique, concerne le comportement presque sûr de la diffusion.

**Dans la seconde partie** (qui contient les chapitres III à VI), nous concentrons sur le schéma d'Euler de pas décroissant étudié par Lamberton et Pagès.

Dans le chapitre III, nous étendons leurs résultats à un cadre plus général en supposant que la diffusion vérifie une condition de stabilité faible. De plus, les résultats de convergence sont obtenus pour des fonctions à croissance exponentielle, plus précisément pour des fonctions

---

<sup>2</sup>i.e.  $|\nabla V|^2 = \mathcal{O}(V)$  et  $\|\mathbb{D}^2 V\|_\infty < +\infty$

<sup>3</sup>plus exactement  $|b|^2 + \text{Tr}(\sigma\sigma^*) = \mathcal{O}(V)$

<sup>4</sup>cf. théorème II.2 page 18

négligeables devant  $\exp(\lambda V^\varsigma)$  où  $\lambda$  et  $\varsigma$  sont des paramètres strictement positifs dépendant de la croissance du drift et du coefficient de diffusion.

Dans le chapitre IV, nous proposons une alternative à la méthode de martingales pour prouver la convergence de l'algorithme. Le cadre est plus restreint car nous nous intéressons uniquement aux diffusions  $(V, \alpha)$ -confluentes (toutes les trajectoires convergent vers la trajectoire stationnaire). Il nous est alors possible de montrer la convergence presque sûre et dans  $\mathbf{L}^p$  du schéma vers la diffusion. Cette convergence se faisant suffisamment vite, les mesures empiriques du schéma et celles de la diffusion convergent vers la même cible, c'est à dire  $\nu(f)$  (par le théorème ergodique pour la diffusion). L'avantage de cette méthode de démonstration, qui relève donc d'une approche de type ergodique, est d'obtenir la convergence pour une classe de fonctions encore plus large que celle du chapitre précédent.

Dans le chapitre V, nous étudions précisément les vitesses de convergence de l'algorithme. Nous donnons un développement à tout ordre de l'erreur  $\nu_n^\eta(f) - \nu(f)$  pour une classe de fonctions tests  $f$  (fonctions admettant une solution régulière à l'équation de Poisson). Nous prouvons que pour des pas rapidement décroissants, l'erreur vérifie un théorème de la limite centrale (TLC) et que pour des pas lentement décroissants il y a convergence avec biais. Une loi de logarithme itéré accompagne notre TLC. Cette étude permet notamment de déterminer certains paramètres nécessaires à la mise en œuvre de l'algorithme. Pour une implémentation, il est naturel de considérer une suite de pas  $(\gamma_n)_{n \geq 1}$  de la forme  $\gamma_n = \gamma_0 n^{-r}$  avec  $\gamma_0 > 0$  et  $r \in ]0, 1[$ , et une suite de poids  $(\eta_n)_{n \geq 1}$  de la forme  $\eta_n = n^{-s}$  avec  $s \in ]-\infty, 1[$ . À l'issue de ce chapitre, nous trouvons les paramètres  $r$  et  $s$  optimaux et donnons la vitesse de convergence associée (il s'agit d'un théorème de la limite centrale avec biais). Un exemple numérique illustre le choix des paramètres. Nous finissons le chapitre, en montrant qu'une extrapolation de type Richardson se prête très bien à l'algorithme afin de gagner un ordre de convergence. La détermination du paramètre  $\gamma_0$  est un problème délicat car il dépend du biais et de la variance du TLC optimal, quantités pour lesquelles nous n'avons pas de formule explicite et dont l'approximation numérique est relativement complexe.

Nous finissons l'étude de ce schéma en nous concentrant sur une situation particulière : que se passe-t-il pour les mesures empiriques pondérées du schéma lorsqu'il y a plusieurs mesures invariantes ? Cette question, à laquelle nous ne répondons pour l'instant qu'en dimension un, nous a amenés à mettre en évidence une correspondance entre classification de Feller et fonctions de Lyapounov.

**Dans la troisième partie**, nous exposons un nouvel algorithme à pas adapté qui permet de nous affranchir des hypothèses trop restrictive sur  $b$  et  $V$ . Comme nous l'avons déjà souligné, les schémas d'Euler explicites à pas constant et à pas décroissant sont instables lorsque le drift  $b$  n'est pas à dérivée bornée. Il s'agit pourtant d'une situation importante car très fréquente dans les problèmes de mécanique aléatoire. Donnons plus de précisions sur les travaux, déjà mentionnés, où des schémas implicites (à pas constant) sont considérés. Dans [Tal02], l'auteur prouve la convergence de son algorithme (moyennisation le long d'une trajectoire à horizon fini d'un schéma implicite) dans le cas d'un système Hamiltonien à coefficient de diffusion constant. La généralisation à un coefficient non constant apparaît cependant très complexe. Dans [MSH02], les auteurs prouvent uniquement l'ergodicité de leurs schémas implicites sans étudier la convergence vers la mesure invariante. Leur travail se fait dans un cadre général et ils considèrent plusieurs exemples tels que les systèmes Hamiltoniens et les systèmes mono-

tones. Dans d'autres travaux ([LMA03] et [MS04]), Mattingly, Stuart et Lamba introduisent un schéma explicite dont le pas de discrétisation est adapté en fonction de l'erreur faite entre le schéma d'Euler et le schéma de Heun de l'équation différentielle sous-jacente  $\dot{x} = b(x)$ . Ils prouvent la stabilité de ce schéma explicite et la convergence vers la diffusion sans aborder la question de l'approximation du régime stationnaire.

L'algorithme que nous proposons est un algorithme explicite, récursif et très simple d'implémentation. Il repose sur les moyennes empiriques pondérées d'un schéma d'Euler explicite de pas  $(\tilde{\gamma}_n)_{n \geq 1}$  où  $\tilde{\gamma}_n$  est une variable aléatoire adaptée bornée par  $\gamma_n$  (la suite  $(\gamma_n)_{n \geq 1}$  tend vers 0). Plus précisément, le pas de discrétisation  $\tilde{\gamma}_n$  est de la forme  $\tilde{\gamma}_n = \gamma_n \wedge \chi_{n-1}$  où  $\chi_{n-1}$  est une variable aléatoire de la tribu engendrée par  $(X_0, \dots, X_{n-1})$ . La variable aléatoire  $\chi_{n-1}$  sera convenablement choisie, à chaque itération, afin de jouer le rôle de force de rappel, ce qui fera de ce schéma adaptatif un schéma stable. Intuitivement, il faut comprendre ce schéma de la façon suivante : lorsque  $(X_n)_{n \geq 0}$  tend à exploser, le pas de discrétisation devient très petit afin de mieux approcher l'EDS et ainsi de garder ses propriétés qualitatives, notamment la force de rappel (condition de stabilité). Le critère permettant de trouver une bonne variable aléatoire  $\chi_n$  repose sur le résultat concernant le comportement presque sûr de la diffusion obtenu dans le chapitre II. Un choix naturel pour  $\chi_n$  est par exemple  $\chi_n = \frac{1}{|b(X_n)|^2 \vee 1}$ . De plus, nous montrons qu'à partir d'un temps aléatoire fini presque sûrement, l'événement  $\{\chi_{n-1} < \gamma_n\}$  ne se produit plus, c'est-à-dire  $\tilde{\gamma}_n = \gamma_n$ . Ainsi, nous obtenons des résultats de convergence similaires à ceux du chapitre III dans le cas de systèmes Hamiltoniens très généraux (coefficient de diffusion non borné) ou encore dans le cas de systèmes monotones (équation de Lorenz perturbée par exemple).

Nous terminons cette partie par un chapitre numérique dans lequel nous discutons des différents paramètres liés à l'implémentation de l'algorithme. Nous testons aussi l'algorithme couplé à l'extrapolation de Richardson pour cinq exemples variés.

## Rappels et notations

Rappelons tout d'abord que l'équation différentielle stochastique

$$\forall t \geq 0, \quad dx_t = b(x_t)dt + \sigma(x_t)dB_t, \quad \text{et } x_0 \in \mathbf{R}^d,$$

admet une solution unique <sup>5</sup>  $(x_t)_{t \geq 0}$ , c'est à dire un processus continu  $(\mathcal{F}_t)_{t \geq 0}$ -adapté à valeurs dans  $\mathbf{R}^d$ , si les coefficients  $b$  et  $\sigma$  sont localement lipschitziens *i.e.*

$$\forall |x| \leq n, |y| \leq n, \quad |\sigma(x) - \sigma(y)| \vee |b(x) - b(y)| \leq L_n |x - y| \quad (0.1)$$

et qu'il existe une fonction de Lyapounov  $V$  positive,  $\mathcal{C}^2$  qui tend vers l'infini à l'infini, telle que

$$\exists \lambda > 0, \forall x \in \mathbf{R}^d, \quad \langle b(x), \nabla V(x) \rangle + \frac{1}{2} \text{Tr}(\sigma^* D^2 V \sigma)(x) \leq \lambda V(x). \quad (0.2)$$

Pour une preuve de ce résultat on pourra consulter [Has80] (théorème 4.1 page 84). De plus, la solution  $(x_t)_{t \geq 0}$  est un processus de Markov homogène de fonction de transition  $P_t(y, A)$  qui est un semi-groupe de Feller défini par  $P_t(y, A) = \mathbf{P}_y[x(t) \in A]$  pour tout  $y \in \mathbf{R}^d$  et  $A$

---

<sup>5</sup>à l'indistinguabilité près

ouvert de  $\mathbf{R}^d$ . Le générateur infinitésimal associé au processus de diffusion  $(x_t)_{t \geq 0}$  et à l'EDS est défini pour toute  $f \in \mathcal{C}^2(\mathbf{R}^d, \mathbf{R})$  par

$$\begin{aligned} \forall x \in \mathbf{R}^d, \quad \mathcal{A}f(x) &= \sum_{i=1}^d b_i(x) \frac{\partial f}{\partial x_i}(x) + \frac{1}{2} \sum_{i,j=1}^d (\sigma \sigma^*)_{i,j}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x), \\ &= \langle b, \nabla f \rangle(x) + \frac{1}{2} \text{Tr}(\sigma^* \mathbf{D}^2 f \sigma)(x). \end{aligned} \quad (0.3)$$

Tout processus solution d'une telle EDS sera appelé par la suite processus de diffusion (associé au générateur  $\mathcal{A}$ ).

On appelle mesure invariante pour le processus  $(x_t)_{t \geq 0}$  toute probabilité  $\nu$  sur  $\mathbf{R}^d$  vérifiant

$$\forall t \geq 0, \quad \nu P_t = \nu.$$

Si  $x_0$  est de loi  $\nu$  alors pour tout  $t \geq 0$ ,  $x_t$  est de loi  $\nu$  est on dit que le processus  $(x_t)_{t \geq 0}$  est stationnaire.

Dans ce chapitre, on expose en détail la notion de stabilité d'une EDS et la méthode de martingales permettant d'étudier le comportement asymptotique du processus solution  $(x_t)_{t \geq 0}$ . On définit tout d'abord la notion de "condition de stabilité" puis on montre la tension des mesures empiriques sous cette condition. On conclut en prouvant que toute valeur d'adhérence des mesures empiriques est probabilité invariante de  $(x_t)_{t \geq 0}$  en utilisant le théorème d'Echeverria-Weiss.

## 1. Stabilité

On considère l'équation différentielle stochastique suivante

$$dx_t = b(x_t)dt + \sigma(x_t)dB_t, \quad x_0 \in \mathbf{R}^d, \quad (1.1)$$

où  $b : \mathbf{R}^d \rightarrow \mathbf{R}^d$  et  $\sigma : \mathbf{R}^d \rightarrow \mathcal{M}(d \times q)$  ( $q \leq d$ ) sont continues et localement lipschitziennes, et  $(B_t)_{t \geq 0}$  est un mouvement brownien standard sur  $\mathbf{R}^q$ . On note  $\mathcal{A}$  le générateur associé à cette équation (1.1), défini pour toute  $f \in \mathcal{C}^2(\mathbf{R}^d, \mathbf{R})$  par

$$\forall x \in \mathbf{R}^d, \quad \mathcal{A}f(x) = \sum_{i=1}^d b_i(x) \frac{\partial f}{\partial x_i}(x) + \frac{1}{2} \sum_{i,j=1}^d (\sigma \sigma^*)_{i,j}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x),$$

ou encore

$$\mathcal{A}f = \langle \nabla f, b \rangle + \frac{1}{2} \text{Tr}(\sigma^*(D^2V)\sigma). \quad (1.2)$$

On suppose que l'équation (1.1) a une solution  $(x_t)_{t \geq 0}$  qui est un processus continu, fortement markovien. Dans l'introduction, on a rappelé qu'une condition suffisante pour l'existence d'une telle solution est l'existence d'une fonction  $V \in \mathcal{C}^2(\mathbf{R}^d, \mathbf{R}_+^*)$  telle que  $\lim_{|x| \rightarrow +\infty} V(x) = +\infty$  et  $\mathcal{A}V \leq cV$  avec  $c > 0$ .

**Exemple.** En dimension 1, si  $b(x) = -x^3$  et  $\sigma(x) = x$  alors l'EDS a une unique solution. En effet, si on considère la fonction  $V(x) = x^2 + 1$  alors on a

$$\mathcal{A}V(x) = -2x^4 + x^2 \leq V(x).$$

La solution n'explose pas et est définie sur  $\mathbf{R}_+$  si  $x_0 \in \mathbf{R}_+$ , ou bien  $\mathbf{R}_-$  si  $x_0 \in \mathbf{R}_-$ . \*

La notion de stabilité introduite dans cette section est proche de la condition de Hajek pour les chaînes de Markov. La condition de Hajek est une condition assez forte qui implique l'ergodicité géométrique. Notre hypothèse est plus faible et est similaire à la condition sur le drift (dans le cadre chaîne de Markov) introduite dans [DFMS04] qui implique l'ergodicité sous-géométrique.

**Définition II.1 (Condition de stabilité).** Une diffusion de générateur  $\mathcal{A}$  satisfait une *condition de stabilité* s'il existe deux fonctions  $V \in \mathcal{C}^2(\mathbf{R}^d, [1, +\infty[)$  et  $\phi \in \mathcal{C}([1, +\infty[, \mathbf{R}_+)$  telles que

1.  $\lim_{|x| \rightarrow +\infty} V(x) = +\infty$ ,
2.  $\phi$  vérifie :  $\forall x \in [1, +\infty[, \phi(x) \leq x$  et  $\lim_{x \rightarrow +\infty} \phi(x) = +\infty$ ,
3.  $\exists \alpha > 0$  et  $\exists \beta > 0$  tels que  $\mathcal{A}V \leq -\alpha(\phi \circ V) + \beta$ .

La condition de stabilité est dite *forte* si  $\phi(x) = x$  et *faible* sinon. ♣

**Note:** Par abus, on dira par la suite que le générateur  $\mathcal{A}$  vérifie une condition de stabilité de type  $(V, \phi, \alpha)$  si les hypothèses de la définition II.1 sont satisfaites avec les fonctions  $V$  et  $\phi$ , et la constante  $\alpha$ . Lorsque la connaissance du  $\alpha$  n'est pas importante, on parlera de condition de stabilité de type  $(V, \phi)$ . Le fait que  $\alpha > 0$  est par contre primordial. \*

Il est important de noter que si  $\mathcal{A}$  vérifie une condition de stabilité, alors il existe une mesure invariante associée à (1.1) (cf. [Has80] théorème 5.1 page 90).

Contrairement à la condition introduite dans [DFMS04], la fonction  $\phi$  n'est pas nécessairement concave. Cependant, il est classique de considérer la fonction  $\phi$  de la forme  $\phi(x) = x^a$  avec  $a \in ]0, 1]$  ou bien  $\phi(x) = \ln(x)$ . Pour un exposé des méthodes de stabilité lorsque  $\phi(x) = x^a$  avec  $a \in ]0, 1]$ , on pourra se reporter à [Pag01]. Ces méthodes sont généralisées dans ce chapitre et légèrement améliorées notamment par le théorème suivant.

**Théorème II.1:**

On suppose qu'une diffusion  $(x_t)_{t \geq 0}$  de générateur infinitésimal  $\mathcal{A}$  vérifie une condition de stabilité de type  $(V, \phi)$ . Alors,

– si  $\theta \in \mathcal{C}^1(\mathbf{R}_+; \mathbf{R}_+^*)$  est décroissante telle que  $\int_0^{+\infty} \theta(s) ds < +\infty$ , on a

$$\mathbf{E} \left[ \int_0^{+\infty} \theta(s) (\phi \circ V)(x_s) ds \right] < +\infty \quad \text{et} \quad \lim_{t \rightarrow +\infty} \theta(t) V(x_t) = 0 \quad p.s. \quad (1.3)$$

– si la condition de stabilité est forte,  $\sup_{t \geq 0} \mathbf{E}[V(x_t)] < +\infty$ . ★

Ce théorème permet d'avoir une bonne connaissance du comportement du processus  $(x_t)_{t \geq 0}$  sous une unique hypothèse : la condition de stabilité. La preuve repose sur le lemme suivant.

**Lemme II.1:**

Soit  $V \in \mathcal{C}^2$  et  $\mathcal{V}$  deux fonctions positives de  $\mathbf{R}^d$  vérifiant  $\mathcal{A}V \leq -\alpha\mathcal{V} + \beta$  pour  $\alpha > 0$  et  $\beta > 0$ . Si  $\theta \in \mathcal{C}^1(\mathbf{R}_+; \mathbf{R}_+^*)$  est décroissante telle que  $\int_0^{+\infty} \theta(s) ds < +\infty$ , alors

$$\mathbf{E} \left[ \int_0^{+\infty} \theta(s) \mathcal{V}(x_s) ds \right] < +\infty, \quad (1.4)$$

et  $\lim_{t \rightarrow +\infty} \theta(t)V(x_t) = 0$  *p.s.* ◆

Preuve. Soit  $(S_t)_{t \geq 0}$  la surmartingale positive définie par

$$\forall t \geq 0, \quad S_t = \int_0^t \theta(s) dV(x_s) + \theta(0)V(x_0) + \int_0^t \alpha \theta(s) \mathcal{V}(x_s) ds + \beta \int_t^{+\infty} \theta(s) ds, \quad (1.5)$$

où  $x_0$  est défini en (1.1).

- Montrons tout d'abord que le processus  $(S_t)_{t \geq 0}$  est positif. Une intégration par parties donne

$$\forall t \geq 0, \quad \int_0^t \theta(s) dV(x_s) + \theta(0)V(x_0) = \theta(t)V(x_t) - \int_0^t \theta'(s)V(x_s) ds \geq 0, \quad (1.6)$$

car  $\theta$  est décroissante positive et  $V$  est positive. Les fonctions  $\mathcal{V}$  et  $\theta$  étant positives, on déduit aisément de (1.5) et (1.6) que  $S_t \geq 0$ , pour tout  $t \geq 0$ .

- D'autre part  $dV(x_t) = \mathcal{A}V(x_t)dt + ((\nabla V)^* \sigma)(x_t)dB_t$ , et en posant  $\Theta = \int_0^{+\infty} \theta(s) ds$  on a

$$S_t = \int_0^t \theta(s) (\mathcal{A}V(x_s) + \alpha \mathcal{V}(x_s) - \beta) ds + \int_0^t \theta(s) ((\nabla V)^* \sigma)(x_s) dB_s + \beta \Theta + \theta(0)V(x).$$

Comme  $\mathcal{A}V \leq -\alpha \mathcal{V} + \beta$ , on en déduit que  $(S_t)_{t \geq 0}$  est une surmartingale locale positive. De plus  $S_0 = \beta \Theta + \theta(0)V(x_0) > 0$  est déterministe donc par le lemme de Fatou conditionnel,  $(S_t)_{t \geq 0}$  est une surmartingale positive, et de ce fait converge presque sûrement vers une variable aléatoire  $S_\infty$  vérifiant  $\mathbf{E}[S_\infty] \leq \beta \Theta + \theta(0)V(x_0) < +\infty$ . En combinant (1.5) et (1.6), on obtient alors

$$\alpha \mathbf{E} \left[ \int_0^{+\infty} \theta(s) \mathcal{V}(x_s) ds \right] \leq \mathbf{E}[S_\infty] < +\infty. \quad (1.7)$$

- De plus, on en déduit que  $\int_0^{+\infty} \theta(s) dV(x_s) \leq S_\infty < +\infty$  *p.s.*, et comme la fonction  $\theta$  décroît vers 0, le lemme de Kronecker s'applique et donne

$$\lim_{t \rightarrow +\infty} \theta(t) \int_0^t dV(x_s) = \lim_{t \rightarrow +\infty} \theta(t) (V(x_t) - V(x)) = 0 \quad \textit{p.s.},$$

d'où le résultat. #

Preuve (théorème II.1).

- Pour obtenir (1.3) et  $\lim_{t \rightarrow +\infty} \theta(t)V(x_t) = 0$  *p.s.*, il suffit d'appliquer le lemme II.1 avec les fonctions  $V$  et  $\mathcal{V} = (\phi \circ V)$ .
- Dans le cas où la condition de stabilité est forte ( $\phi = \text{Id}$ ), on applique la formule d'Itô à  $(x_t)_{t \geq 0}$  avec la fonction  $g(t, x) = \exp(\alpha t)V(x)$ , qui donne

$$\begin{aligned} \exp(\alpha t)V(x_t) &= V(x_0) + \int_0^t (\exp(\alpha s) \mathcal{A}V(x_s) + \alpha \exp(\alpha s)V(x_s)) ds \\ &\quad + \int_0^t \exp(\alpha s) ((\nabla V)^* \sigma)(x_s) dB_s, \quad (1.8) \end{aligned}$$

et on localise la martingale locale en considérant une suite croissante  $(\tau_n)_{n \geq 0}$  de temps d'arrêt qui tend vers l'infini. En utilisant la condition de stabilité forte, on obtient pour tout  $t \geq 0$ ,

$$\mathbf{E}[\exp(\alpha(t \wedge \tau_n))V(x_{t \wedge \tau_n})] \leq V(x_0) + \beta \mathbf{E} \left[ \int_0^{t \wedge \tau_n} \exp(\alpha s) ds \right].$$

Comme  $V$  est positive, le lemme de Fatou entraîne

$$\exp(\alpha t) \mathbf{E}[V(x_t)] \leq V(x_0) + \frac{\beta}{\alpha} (\exp(\alpha t) - 1),$$

d'où  $\sup_{t \geq 0} \mathbf{E}[V(x_t)] \leq \frac{\beta}{\alpha} \vee V(x_0)$ . #

La condition de stabilité seule ne suffit pas à obtenir les résultats de tension souhaités. Il faut lui ajouter une hypothèse de domination du coefficient de diffusion et du gradient de la fonction de Lyapounov dans la direction de  $\sigma$  (cf. condition (1.9) ci-dessous)

**Proposition II.1:**

Supposons que  $\mathcal{A}$  vérifie une condition de stabilité de type  $(V, \phi, \alpha)$ . S'il existe  $\varsigma \in [0, 1]$  et  $C_{V,\sigma} > 0$  tels que

$$\text{Tr}(\sigma^*(\nabla V)^{\otimes 2} \sigma) \leq C_{V,\sigma} (\phi \circ V) V^{1-\varsigma}, \quad (1.9)$$

alors  $\mathcal{A}$  vérifie une condition de stabilité de type  $(\Psi \circ V, \phi_\Psi)$  où

$$\forall v \in ]1, +\infty[, \quad \Psi(v) = \begin{cases} v^p & \text{avec } p \in \left] 1, 1 + \frac{2\alpha}{C_{V,\sigma}} \right], \quad \text{si } \varsigma = 0, \\ \exp(\lambda v^\varsigma) & \text{avec } \lambda < \frac{2\alpha}{\varsigma C_{V,\sigma}}, \quad \text{si } \varsigma \in ]0, 1], \end{cases} \quad (1.10)$$

et  $\phi_\Psi(x) = x \left( \frac{\phi \circ \Psi^{-1}}{\Psi^{-1}} \right) (x)$  (où  $\Psi^{-1}$  est la réciproque de  $\Psi$ ). ◆

C'est un résultat très important pour la suite. Le fait que la condition de stabilité peut se "transmettre" à des fonctions plus grandes (polynômiales voir exponentielles) sous une condition de domination sur le coefficient de diffusion est fondamentale. Par exemple, on sait par le théorème II.1 que si  $\mathcal{A}$  vérifie une condition de stabilité forte ( $\phi(x) = x$ ) avec  $V$  on a

$$\lim_{t \rightarrow +\infty} \frac{V(x_t)}{t \ln^{1+\varepsilon}(t)} = 0 \quad p.s.$$

pour tout  $\varepsilon > 0$ , car  $\int_1^{+\infty} \frac{1}{t \ln^{1+\varepsilon}(t)} dt < +\infty$ . Or si le coefficient de diffusion est constant (cas où  $\varsigma = 1$ ), la proposition précédente permet d'affirmer que

$$\lim_{t \rightarrow +\infty} \frac{\exp(\lambda V(x_t))}{t \ln^{1+\varepsilon}(t)} = 0 \quad p.s.$$

Cela permet d'avoir une meilleure connaissance du comportement  $(x_t)_{t \geq 0}$  mais aussi des mesures empiriques  $(\nu_t)_{t \geq 0}$  comme on le verra dans la section suivante.

**Remarque:** Le paramètre  $\varsigma$  sert de jauge au coefficient de diffusion : plus  $\varsigma$  diminue, plus  $\sigma$  peut avoir une croissance importante. Si l'on a une condition de stabilité forte avec  $V(x) = |x|^2 + 1$  et  $\varsigma = 0$ , l'hypothèse (1.9) est impliquée par  $\text{Tr}(\sigma \sigma^*)(x) \leq C|x|^2$ . \*

**Exemple.** En dimension 1, si  $b(x) = \frac{x}{x^2+1} \ln\left(\frac{1}{x^2+1}\right)$  et  $\sigma(x) = \sqrt{\ln(x^2+1)}$ , alors  $\mathcal{A}$  vérifie une condition de stabilité faible de type  $(V, \phi, \alpha)$  avec  $V(x) = x^2 + 1$ ,  $\phi = \ln(x)$  et  $\alpha < 1$ . De plus, la condition (1.9) est satisfaite avec  $\varsigma = 0$  et  $C_{V,\sigma} = 4$ .

Dans cet exemple, le drift est très petit à l'infini et la vitesse de convergence du semi-groupe vers la mesure invariante est plus lente que dans le cas où la condition de stabilité est forte (cf. [GRS99] dans le cas  $\sigma$  constant). \*

**Preuve** (de la proposition II.1). On montre tout d'abord, un résultat un peu plus général. Montrons que si  $\Psi \in \mathcal{C}^2([1, +\infty[)$  et s'il existe  $y_0 \geq 1$  tel que  $\Psi$  est croissante, bijective sur  $[v_0, +\infty[$  et  $\forall x \in V^{-1}([v_0, +\infty[)$ ,  $\text{Tr}(\sigma^*(\nabla V)^{\otimes 2}\sigma)(x) \leq \varepsilon(\phi \circ V)(x) \frac{\Psi' \circ V}{\Psi'' \circ V}(x)$  pour un  $\varepsilon < 2\alpha$ , alors il existe  $\tilde{\alpha} > 0$  et  $\tilde{\beta} > 0$  tels que

$$\forall x \in \mathbf{R}^d, \quad \mathcal{A}(\Psi \circ V)(x) \leq -\tilde{\alpha}(\tilde{\phi}_\Psi \circ (\Psi \circ V))(x) + \tilde{\beta}, \quad (1.11)$$

avec  $\tilde{\phi}_\Psi = (\Psi' \phi) \circ \Psi^{-1}$ , où  $\Psi^{-1}$  est la réciproque de  $\Psi$ .

- Il est clair que sur  $\mathbf{R}^d$

$$\mathcal{A}(\Psi \circ V) = (\Psi' \circ V)\mathcal{A}V + \frac{1}{2}(\Psi'' \circ V) \text{Tr}(\sigma^*(\nabla V)^{\otimes 2}\sigma),$$

et comme  $\mathcal{A}$  vérifie une condition de stabilité avec  $V$  et  $\phi$ , on a

$$\mathcal{A}(\Psi \circ V) \leq -\alpha(\Psi' \circ V)(\phi \circ V) + \beta(\Psi' \circ V) + \frac{1}{2}(\Psi'' \circ V) \text{Tr}(\sigma^*(\nabla V)^{\otimes 2}\sigma).$$

Comme  $V$  et  $\Psi$  sont  $\mathcal{C}^2$ , la fonction  $\mathcal{A}(\Psi \circ V)$  est majorée par  $C > 0$  sur le compact  $V^{-1}([1, v_0])$ .

Sur l'ensemble  $V^{-1}([v_0, +\infty[)$ , on a d'après la condition de croissance sur  $\text{Tr}(\sigma^*(\nabla V)^{\otimes 2}\sigma)$

$$\mathcal{A}(\Psi \circ V) \leq -\alpha((\Psi' \phi) \circ V) + \beta(\Psi' \circ V) + \varepsilon \frac{1}{2}((\Psi' \phi) \circ V),$$

et puisque  $\varepsilon < 2\alpha$ ,  $\mathcal{A}(\Psi \circ V) \leq -\bar{\alpha}((\Psi' \phi) \circ V) + \beta(\Psi' \circ V)$  avec  $\bar{\alpha} = \frac{\varepsilon}{2} - \alpha > 0$ . Pour  $\tilde{\alpha} \in ]0, \bar{\alpha}[$ , on note  $R(V, \phi) = ((\tilde{\alpha} - \bar{\alpha})(\phi \circ V) + \beta)(\Psi' \circ V)$  et on a donc

$$\forall x \in V^{-1}([v_0, +\infty[), \quad \mathcal{A}(\Psi \circ V)(x) \leq -\tilde{\alpha}((\Psi' \phi) \circ V)(x) + R(V, \phi)(x). \quad (1.12)$$

La fonction  $R(V, \phi)$  est continue et négative hors d'un compact de  $V^{-1}([v_0, +\infty[)$  car  $\Psi' \circ V$  est positive et  $\lim_{|x| \rightarrow +\infty} (\phi \circ V)(x) = +\infty$ , donc  $R(V, \phi)$  est majorée par  $\tilde{\beta}$  sur  $V^{-1}([v_0, +\infty[)$ .

On a ainsi montré (1.11). (Si  $\phi$  n'est pas continue mais monotone, le résultat reste vrai en considérant la réciproque généralisée.)

- Dans le cas  $\varsigma = 0$ , on considère la fonction  $\Psi(v) = v^p$  et  $v_0 = 1$ . Alors, pour tout  $v \in [1, +\infty[$ , on a  $\Psi'(v) = pv^{p-1}$  et  $\Psi''(v) = p(p-1)v^{p-2}$ , et d'après (1.9),

$$\begin{aligned} \text{Tr}(\sigma^*(\nabla V)\sigma) &\leq C_{V,\sigma}(\phi \circ V)V, \\ &= C_{V,\sigma}(p-1)(\phi \circ V) \frac{\Psi' \circ V}{\Psi'' \circ V}. \end{aligned}$$

Comme  $p \in ]1, 1 + \frac{2\alpha}{C_{V,\sigma}}[$ , on a  $\varepsilon = C_{V,\sigma}(p-1) \in ]0, 2\alpha[$  donc d'après (1.11)

$$\exists \tilde{\alpha} > 0, \exists \tilde{\beta} > 0, \quad \mathcal{A}V^p \leq -\tilde{\alpha} \left( \tilde{\phi}_\Psi \circ (\Psi \circ V) \right) + \tilde{\beta}, \quad (1.13)$$

où  $\tilde{\phi}_\Psi = px^{\frac{p-1}{p}} \phi(x^{\frac{1}{p}})$ . Or  $\phi_\Psi = x \frac{\phi(x^{1/p})}{x^{1/p}} = \frac{1}{p} \tilde{\phi}_\Psi$ , donc (1.13) reste vrai en remplaçant  $\tilde{\phi}_\Psi$  par  $\phi_\Psi$ . De plus, on a pour tout  $x \in [1, +\infty[$ ,  $\phi_\Psi(x) \leq x$  et  $\lim_{x \rightarrow +\infty} \phi_\Psi(x) = +\infty$  donc  $\mathcal{A}$  vérifie une condition de stabilité de type  $(V^p, \phi_\Psi)$ .

- Dans le cas  $\varsigma \in ]0, 1]$ , on considère la fonction  $\Psi(v) = \exp(\lambda v^\varsigma)$  et  $v_0 = \left(\frac{1-\varsigma}{\lambda}\right)^{1/\varsigma} \vee 1$ . Alors pour tout  $x \in [V^{-1}(v_0), +\infty[$ ,  $\Psi'(x) = \lambda \varsigma x^{\varsigma-1} \Psi(x)$  et

$$0 \leq \Psi''(x) = \lambda \varsigma (\varsigma - 1) x^{\varsigma-2} \Psi(x) + \lambda^2 \varsigma^2 x^{2(\varsigma-1)} \Psi(x) \leq \lambda^2 \varsigma^2 x^{2(\varsigma-1)} \Psi(x),$$

car  $\varsigma \leq 1$ . Ainsi, d'après (1.9), on a sur  $[V^{-1}(v_0), +\infty[$

$$\begin{aligned} \text{Tr}(\sigma^*(\nabla V)\sigma) &\leq C_{V,\sigma}(\phi \circ V)V^{1-\varsigma}, \\ &\leq C_{V,\sigma} \lambda \varsigma (\phi \circ V) \frac{\Psi' \circ V}{\Psi'' \circ V}. \end{aligned}$$

Comme  $\lambda < \frac{2\alpha}{\varsigma C_{V,\sigma}}$ , on déduit de (1.11) que

$$\exists \tilde{\alpha} > 0, \exists \tilde{\beta} > 0, \quad \mathcal{A} \exp(\lambda V^\varsigma) \leq -\tilde{\alpha} \tilde{\phi}_\Psi(\exp(\lambda V^\varsigma)) + \tilde{\beta}, \quad (1.14)$$

où  $\tilde{\phi}_\Psi(x) = \lambda \varsigma (\Psi^{-1}(x))^{\varsigma-1} x \phi(\Psi^{-1}(x))$  et  $\Psi^{-1}(x) = \left(\frac{\log(x)}{\lambda}\right)^{1/\varsigma}$ . Or

$$\forall x \in [e^\lambda, +\infty[, \quad \tilde{\phi}_\Psi(x) = \lambda \varsigma (\Psi^{-1}(x))^\varsigma x \left( \frac{\phi(\Psi^{-1}(x))}{\Psi^{-1}(x)} \right) \geq \lambda \varsigma \phi_\Psi(x), \quad (1.15)$$

donc la relation (1.14) est encore vérifiée en remplaçant la fonction  $\tilde{\phi}_\Psi$  par  $\phi_\Psi$ . De plus, on a pour tout  $x \in [1, +\infty[$ ,  $\phi_\Psi(x) \leq x$  et  $\lim_{x \rightarrow +\infty} \phi_\Psi(x) = +\infty$  donc  $\mathcal{A}$  vérifie une condition de stabilité de type  $(\exp(\lambda V^\varsigma), \phi_\Psi)$ . #

Remarque: Si  $\mathcal{A}$  vérifie une condition de stabilité du type  $(V, \phi)$  et que  $\tilde{\phi}$  est une fonction de  $\mathcal{C}([1, +\infty[, \mathbf{R}_+)$  vérifiant  $\tilde{\phi}(x) \leq x$ ,  $\lim_{x \rightarrow +\infty} \tilde{\phi}(x) = +\infty$  et  $\tilde{\phi} = \mathcal{O}(\phi)$ , alors  $\mathcal{A}$  vérifie une condition de stabilité de type  $(V, \tilde{\phi})$ . \*

## 2. Tension des mesures empiriques

La combinaison des conditions de stabilité et de domination sur  $\text{Tr}(\sigma^*(\nabla V)^{\otimes 2}\sigma)$  permet d'établir la tension des mesures empiriques  $(\nu_t)_{t \geq 0}$  définies par

$$\forall t \geq 0, \quad \nu_t(dx) = \frac{1}{t} \int_0^t \delta_{x_s}(dx) ds. \quad (2.1)$$

En fait, la condition de stabilité seule suffit à prouver la tension de  $(\nu_t)_{t \geq 0}$  (cf. [EK86]) et donc l'existence d'une mesure stationnaire (cf. Introduction). Mais sous la condition de domination supplémentaire, on montre que toute valeur d'adhérence de  $(\nu_t)_{t \geq 0}$  est une probabilité sous laquelle on peut intégrer des fonctions à croissance polynômiale, voir exponentielle. Plus précisément, on montre le résultat suivant.

**Proposition II.2:**

Si  $\mathcal{A}$  vérifie une condition de stabilité de type  $(V, \phi, \alpha)$ , et s'il existe  $\varsigma \in [0, 1]$  et  $C_{V,\sigma} > 0$  tels que

$$\text{Tr}(\sigma^*(\nabla V)^{\otimes 2}\sigma) \leq C_{V,\sigma}(\phi \circ V)V^{1-\varsigma}, \quad (2.2)$$

alors

$$\begin{cases} \forall p < \frac{1}{2} + \frac{\alpha}{C_{V,\sigma}}, & \overline{\lim}_{t \rightarrow +\infty} \frac{1}{t} \int_0^t ((\phi \circ V)V^{p-1})(x_s)ds \leq C \quad p.s. \quad \text{si } \varsigma = 0, \\ \forall \lambda < \frac{\alpha}{\varsigma C_{V,\sigma}}, & \overline{\lim}_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \left( \frac{\phi \circ V}{V} \exp(\lambda V^\varsigma) \right)(x_s)ds \leq C \quad p.s. \quad \text{si } \varsigma \in ]0, 1]. \end{cases} \quad \blacklozenge$$

Remarque: Notons que les conditions sur  $p$  et  $\lambda$  sont plus restrictives que celles de (1.10). Il est difficile de faire mieux en utilisant les méthodes de martingales, ce facteur 2 intervenant dans l'utilisation du théorème de convergence pour les martingales de crochet intégrable. \*

Preuve. On va prouver uniquement le cas  $\varsigma \in ]0, 1]$ , le cas  $\varsigma = 0$  étant semblable.

Pour tout  $\lambda < \frac{2\alpha}{\varsigma C_{V,\sigma}}$ , on note  $W_{\lambda,\varsigma} = \exp(\lambda V^\varsigma)$  et  $\phi_{\lambda,\varsigma}(x) = x \frac{\phi((\log(x)/\lambda)^{\frac{1}{\varsigma}})}{(\log(x)/\lambda)^{\frac{1}{\varsigma}}}$  de sorte que  $\phi_{\lambda,\varsigma}(W_{\lambda,\varsigma}) = \left( \frac{\phi \circ V}{V} \right) W_{\lambda,\varsigma}$ .

Soit  $\lambda_0 < \frac{\alpha}{\varsigma C_{V,\sigma}}$ . D'après la proposition II.1, on sait que  $\mathcal{A}$  vérifie une condition de stabilité de type  $(W_{\lambda_0,\varsigma}, \phi_{\lambda_0,\varsigma})$ , donc il existe  $\tilde{\alpha} > 0$  et  $\tilde{\beta} > 0$  tels que  $\mathcal{A}W_{\lambda_0,\varsigma} \leq -\tilde{\alpha}(\phi_{\lambda_0,\varsigma} \circ W_{\lambda_0,\varsigma}) + \tilde{\beta}$ . Par le lemme d'Itô, on a donc

$$\begin{aligned} W_{\lambda_0,\varsigma}(x_t) &= W_{\lambda_0,\varsigma}(x) + \int_0^t \mathcal{A}W_{\lambda_0,\varsigma}(x_s)ds + \int_0^t ((\nabla W_{\lambda_0,\varsigma})^*\sigma)(x_s)dB_s, \\ &\leq W_{\lambda_0,\varsigma}(x) - \tilde{\alpha} \int_0^t (\phi_{\lambda_0,\varsigma} \circ W_{\lambda_0,\varsigma})(x_s)ds + \tilde{\beta}t + \int_0^t ((\nabla W_{\lambda_0,\varsigma})^*\sigma)(x_s)dB_s, \end{aligned}$$

et on en déduit que

$$\frac{W_{\lambda_0,\varsigma}(x_t)}{\tilde{\alpha}t} + \frac{1}{t} \int_0^t (\phi_{\lambda_0,\varsigma} \circ W_{\lambda_0,\varsigma})(x_s)ds \leq \frac{\tilde{\beta}}{\tilde{\alpha}} + \frac{W_{\lambda_0,\varsigma}(x)}{\tilde{\alpha}t} + \frac{1}{\tilde{\alpha}t} \int_0^t ((\nabla W_{\lambda_0,\varsigma})^*\sigma)(x_s)dB_s. \quad (2.3)$$

- Prouvons que

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_1^t ((\nabla W_{\lambda_0,\varsigma})^*\sigma)(x_s)dB_s = 0 \quad p.s. \quad (2.4)$$

On introduit la martingale locale  $(M_t)_{t \geq 1} = \left( \int_1^t \frac{((\nabla W_{\lambda_0, \varsigma})^* \sigma)(x_s)}{s} dB_s \right)_{t \geq 1}$  de crochet

$$\forall t \geq 1, \quad \langle M \rangle_t = \int_1^t \frac{\text{Tr}(\sigma^*(\nabla W_{\lambda_0, \varsigma})^{\otimes 2} \sigma)(x_s)}{s^2} ds.$$

Comme  $W_{\lambda_0, \varsigma} = \exp(\lambda_0 V^\varsigma)$ , on déduit de (2.2) que

$$\begin{aligned} \text{Tr}(\sigma^*(\nabla W_{\lambda_0, \varsigma})^{\otimes 2} \sigma) &= \text{Tr}(\sigma^*(\nabla V)^{\otimes 2} \sigma) (\lambda_0 \varsigma V^{\varsigma-1} \exp(\lambda_0 V^\varsigma))^2, \\ &\leq \lambda_0^2 \varsigma^2 C_{V, \sigma} (\phi \circ V) V^{\varsigma-1} \exp(2\lambda_0 V^\varsigma), \\ &\leq \lambda_0^2 \varsigma^2 \tilde{C}_{V, \sigma} \left( \frac{\phi \circ V}{V} \right) \exp(\tilde{\lambda} V^\varsigma), \end{aligned}$$

avec  $\tilde{C}_{V, \sigma} > C_{V, \sigma}$  et  $\tilde{\lambda} \in ]2\lambda_0, \frac{2\alpha}{\varsigma C_{V, \sigma}}[$  (intervalle non vide car  $\lambda_0 < \frac{\alpha}{\varsigma C_{V, \sigma}}$ ). On a donc

$$\langle M \rangle_\infty \leq C \int_1^\infty \frac{1}{s^2} \left( \frac{\phi \circ V}{V} \right) (x_s) \exp(\tilde{\lambda} V^\varsigma(x_s)) ds. \quad (2.5)$$

De plus,  $\tilde{\lambda} < \frac{2\alpha}{\varsigma C_{V, \sigma}}$  donc d'après la proposition II.1 on sait que  $\mathcal{A}$  vérifie une condition de stabilité de type  $(\exp(\tilde{\lambda} V^\varsigma), \phi_{\tilde{\lambda}, \varsigma})$ . Le théorème II.1 appliqué avec la fonction  $\theta(t) = \frac{1}{t^2}$ , et combiné avec (2.5) donne alors  $\mathbf{E}[\langle M \rangle_\infty] < +\infty$ . Par suite, la martingale locale  $(M_t)_{t \geq 1}$  converge presque sûrement. En appliquant le lemme de Kronecker, on obtient

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t ((\nabla W_{\lambda_0, \varsigma})^* \sigma)(x_s) dB_s = 0 \quad p.s. \quad (2.6)$$

- Ainsi, on déduit de (2.6) et (2.3) que

$$\overline{\lim}_{t \rightarrow +\infty} \left( \frac{W_{\lambda_0, \varsigma}(x_t)}{\tilde{\alpha} t} + \frac{1}{t} \int_0^t (\phi_{\lambda_0, \varsigma} \circ W_{\lambda_0, \varsigma})(x_s) ds \right) \leq \frac{\tilde{\beta}}{\tilde{\alpha}}, \quad (2.7)$$

d'où le résultat car  $W_{\lambda_0, \varsigma} \geq 0$ . #

*Remarque:* On pourrait sans difficulté remplacer la condition (1.9) par une condition du type  $\text{Tr}(\sigma^*(\nabla V)^{\otimes 2} \sigma) \leq C_{V, \sigma} (\phi \circ V) V \varepsilon$  avec  $\varepsilon : \mathbf{R}^d \rightarrow ]0, 1]$  tendant vers 0 à l'infini. Dans ce cas, les mesures empiriques admettent des moments exponentiels. \*

Pour l'utilisation de la martingale  $(M_t)_{t \geq 0}$  dans la preuve de la tension des mesures empiriques, on pourra se reporter à [Duf96] et à ses références. Classiquement, la convergence de la martingale est obtenue comme conséquence de  $\sup_{t \geq 0} \mathbf{E}[W_{\lambda, \varsigma}(x_t)] < +\infty$  (cf. (2.5)) ce que l'on sait en se plaçant sous une condition de stabilité forte. Dans le cas plus général d'une condition de stabilité faible, on a pu montrer la convergence de la martingale en utilisant le lemme II.1 (c'est à dire le résultat (1.3) du théorème II.1).

Une conséquence de la proposition II.2 est le corollaire suivant qui permet d'affiner la connaissance que l'on a sur le comportement asymptotique de la diffusion.

**Corollaire II.1:**

Sous les hypothèses de la proposition II.2, on a

$$\overline{\lim}_{t \rightarrow +\infty} \frac{V(x_t)}{t^q} \leq C \text{ p.s. si } \varsigma = 0, \quad \text{et} \quad \overline{\lim}_{t \rightarrow +\infty} \frac{V(x_t)}{\ln(t)^{\frac{1}{\varsigma}}} \leq C \text{ p.s. si } \varsigma \in ]0, 1],$$

avec  $q > \frac{2C_{V,\sigma}}{C_{V,\sigma} + 2\alpha}$ . •

Preuve. D'après la proposition II.1,  $\mathcal{A}$  vérifie une condition de stabilité de type  $(\Psi \circ V, \phi_\Psi)$  où  $\Psi$  est définie en (1.10) et  $\phi_\Psi(x) = x \left( \frac{\phi \circ \Psi^{-1}}{\Psi^{-1}} \right) (x)$  et d'après (2.7) on a

$$\overline{\lim}_{t \rightarrow +\infty} \frac{(\Psi \circ V)(x_t)}{t} \leq C \text{ p.s.} \quad (2.8)$$

- Si  $\varsigma = 0$ , alors  $\Psi(v) = v^p$  avec  $p \in ]1, \frac{1}{2} + \frac{\alpha}{C_{V,\sigma}}[$ , donc par continuité de  $(x \mapsto x^{1/p})$

$$\forall p < \frac{1}{2} + \frac{\alpha}{C_{V,\sigma}}, \quad \overline{\lim}_{t \rightarrow +\infty} \frac{V(x_t)}{t^{1/p}} \leq C^{\frac{1}{p}} \text{ p.s.}$$

- Si  $\varsigma \in ]0, 1]$ , alors  $\Psi(v) = \exp(\lambda v^\varsigma)$  avec  $\lambda < \frac{\alpha}{\varsigma C_{V,\sigma}}$ . Ainsi, pour tout  $\varepsilon > 0$  il existe  $N \geq 2$  tel que pour tout  $n \geq N$  et tout  $t \geq n$ ,  $\frac{\exp(\lambda V^\varsigma(x_t))}{t} \leq C + \varepsilon$  p.s., et par convexité de  $(x \rightarrow x^{1/\varsigma})$  on en déduit

$$\forall n \geq N, \forall t \geq n, \quad V(x_t) \leq \lambda^{-1} 2^{\frac{1}{\varsigma}-1} \left( \log(C + \varepsilon)^{\frac{1}{\varsigma}} + \log(t)^{\frac{1}{\varsigma}} \right) \text{ p.s.}, \quad (2.9)$$

d'où le résultat. #

### 3. Identification de la limite

Dans la section précédente, on a prouvé que les mesures empiriques de la diffusion solution de (1.1) sont tendues, et admettent des moments polynômiaux voir exponentiels dès que le coefficient de diffusion  $\sigma$  est suffisamment petit. Il faut maintenant identifier la limite d'une sous-suite de  $(\nu_t)_{t \geq 0}$ .

Il est connu que toute valeur d'adhérence de  $(\nu_t)_{t \geq 0}$  est une probabilité stationnaire  $\nu$  de  $(x_t)_{t \geq 0}$  i.e.

$$\forall f \in \mathcal{C}_K^2, \forall t \geq 0, \quad \int P_t f d\nu = \int f d\nu. \quad (3.1)$$

(cf. proposition II.4 en annexe). Cependant cette caractérisation de la mesure invariante par le semi-groupe est difficilement utilisable lors de l'étude des mesures empiriques du schéma d'Euler à pas décroissant. On va donc utiliser un critère plus simple à vérifier que la définition (3.1) lorsque l'on étudie l'ergodicité du schéma. Ce critère est donné par le théorème d'Echeverria-Weiss (voir [EK86] Theorem 9.17).

Soit  $E$  un espace métrique séparable et  $\mathcal{C}_0(E)$  l'ensemble des fonctions continues sur  $E$  s'annulant à l'infini. On rappelle qu'un opérateur  $\mathfrak{A} : \mathcal{D}(\mathfrak{A}) \subset \mathcal{C}_0(E) \rightarrow \mathcal{C}_0(E)$  vérifie le

principe du maximum positif si

$$\forall f \in \mathcal{D}(\mathfrak{A}), \quad f(x_0) = \max_{x \in E} f(x) \Rightarrow \mathfrak{A}f(x) \leq 0. \quad (3.2)$$

**Théorème II.2 (Echeverria-Weiss):**

Soit  $E$  un espace métrique séparable compact et  $\mathfrak{A} : \mathcal{D}(\mathfrak{A}) \subset \mathcal{C}(E) \rightarrow \mathcal{C}(E)$  un opérateur linéaire vérifiant le principe du maximum positif, où  $\mathcal{D}(\mathfrak{A})$  est une sous-algèbre dense de  $\mathcal{C}(E)$  vérifiant  $1 \in \mathcal{D}(\mathfrak{A})$  et  $\mathfrak{A}1 = 0$ . Soit  $\mu$  une probabilité sur  $E$  telle que  $\forall f \in \mathcal{D}(\mathfrak{A}), \int_E \mathfrak{A}f d\mu = 0$ .

Alors il existe une solution stationnaire au problème de martingale pour  $(\mathfrak{A}, \mu)$ , c'est-à-dire un processus  $(x_t)_{t \geq 0}$  tel que pour toute  $f \in \mathcal{D}(\mathfrak{A}), \left( f(x_t) - \int_0^t \mathfrak{A}f(x_s) ds \right)_{t \geq 0}$  est une martingale par rapport à la filtration naturelle  $\left( \mathcal{F}_{t+}^x \right)_{t \geq 0}$ , et de loi initiale  $\mu$ . ★

Le passage du cas compact au cas localement compact (qui est notre cadre car on considère des équations de  $\mathbf{R}^d$ ) ne pose pas de problème mis à part qu'il faut vérifier que la solution  $(x_t)_{t \geq 0}$  n'explose pas, c'est à dire que

$$\mathbf{P} \left[ \forall t \geq 0, x_t \in \mathbf{R}^d \right] = 1. \quad (3.3)$$

On donne d'abord une proposition qui étend le théorème d'Echeverria-Weiss au cas localement compact, puis on donnera une condition sur  $\mathcal{A}$  sous laquelle on peut appliquer cette proposition (cf. corollaire II.2).

**Proposition II.3 (Extension au cas localement compact):**

Soit  $E$  est un espace métrique séparable localement compact, et  $\mathfrak{A} : \mathcal{D}(\mathfrak{A}) \subset \mathcal{C}_0(E) \rightarrow \mathcal{C}_0(E)$  un opérateur linéaire vérifiant le principe du maximum positif, où  $\mathcal{D}(\mathfrak{A})$  est une sous-algèbre dense de  $\mathcal{C}_0(E)$ .

S'il existe une suite de fonctions strictement positives  $(V_n)_{n \geq 0}$  telle que

- $(V_n)_{n \geq 0}$  croît vers une fonction  $V > 0$  telle que  $\lim_{p \rightarrow +\infty} \inf_{y \notin L_p} V(y) = +\infty$  pour toute suite de compacts  $(L_p)_{p \geq 1}$  tendant vers  $E$ .
- $\forall n \geq 0, \exists K_n$  compact de  $E$  tel que  $\forall x \in E \setminus K_n, V_n(x) = C_n$ , et  $V_n - C_n \in \mathcal{D}(\mathfrak{A})$ ,
- $\exists c > 0, \forall n \geq 0, \mathfrak{A}(V_n - C_n) \leq cV_n$ ,

et s'il existe une probabilité  $\mu$  sur  $E$  telle que  $\int_E V d\mu < +\infty$  et  $\forall f \in \mathcal{D}(\mathfrak{A}), \int_E \mathfrak{A}f d\mu = 0$ , alors il existe une solution stationnaire au problème de martingale pour  $(\mathfrak{A}, \mu)$ . ◆

La preuve de cette extension est donnée en annexe, section 5.

**Corollaire II.2:**

Soit  $(x_t)_{t \geq 0}$  une diffusion de générateur  $\mathcal{A}$ . S'il existe une fonction  $V \in \mathcal{C}^2(\mathbf{R}^d, [1, +\infty[)$  telle que  $\lim_{|x| \rightarrow +\infty} V(x) = +\infty$ ,

$$\exists c > 0, \mathcal{A}V \leq cV, \quad \text{et} \quad \lim_{|x| \rightarrow +\infty} V(x) = +\infty, \quad (3.4)$$

et s'il existe une probabilité  $\mu$  telle que  $\int_{\mathbf{R}^d} V d\mu < +\infty$  et pour toute  $f \in \mathcal{C}_K^2(\mathbf{R}^d)$ ,  $\int_{\mathbf{R}^d} \mathcal{A}f d\mu = 0$ , alors  $\mu$  est une mesure invariante pour  $(x_t)_{t \geq 0}$ . •

**Preuve.** Il s'agit de vérifier les hypothèses de la proposition II.3 en considérant l'opérateur  $\mathcal{A}$  et le domaine  $\mathcal{C}_K^2(\mathbf{R}^d)$  des fonctions de classe  $\mathcal{C}^2$  à support compact dans  $\mathbf{R}^d$ . L'opérateur de diffusion  $\mathcal{A}$  est linéaire et vérifie le principe du maximum positif (cf. [RY99]). De plus, la sous-algèbre  $\mathcal{C}_K^2(\mathbf{R}^d)$  est dense dans  $\mathcal{C}_0(\mathbf{R}^d)$ .

Soit  $(\Psi_n)_{n \geq 1}$  la suite de fonctions définie pour tout  $n \geq 1$  par  $\Psi_n \in \mathcal{C}^2$  telle que

$$\Psi_n(x) = \begin{cases} x & \text{si } x \in [1, n], \\ x - (x - n)^3 + \frac{1}{2}(x - n)^4 & \text{si } x \in [n, n + 1], \\ n + \frac{1}{2} & \text{si } x \geq n + 1. \end{cases} \quad (3.5)$$

On vérifie que  $\Psi_n$  est concave (car  $\Psi_n'' \leq 0$ ) et que  $\Psi_n'(x) \in [0, 1]$  pour tout  $x \geq 1$ . On pose  $V_n = \Psi_n \circ V$  de sorte que la suite  $(V_n)_{n \geq 1}$  est croissante et converge vers  $V$ . Pour tout  $n \geq 1$ , la formule d'Itô donne

$$\mathcal{A}V_n = \Psi_n'(V)\mathcal{A}V + \Psi_n''(V)(\nabla V)^{\otimes 2} \leq c\Psi_n'(V)V, \quad (3.6)$$

et comme  $x\Psi_n'(x) \leq \Psi_n(x)$  pour tout  $x \geq 1$ , on obtient  $\mathcal{A}V_n \leq cV_n$ .

De plus, la fonction  $V_n - (n + 1)$  est bien  $\mathcal{C}^2$  à support compact et  $\mathcal{A}(V_n - (n + 1)) = \mathcal{A}V_n \leq cV_n$ . La proposition II.3 peut alors s'appliquer. #

## 4. Conclusion

L'étude par les méthodes de martingales du comportement ergodique de  $(x_t)_{t \geq 0}$  donne de nombreuses informations sur le comportement asymptotique de  $(x_t)_{t \geq 0}$  ainsi que sur les mesures empiriques  $(\nu_t)_{t \geq 0}$ . Ces techniques permettent de retrouver un théorème ergodique sans utiliser le théorème ergodique ponctuel (rappelé en introduction). De plus, on obtient un résultat de convergence même en présence de multiples mesures invariantes (situation qui peut se produire si le coefficient de diffusion est dégénéré).

### **Théorème II.3:**

Supposons que  $\mathcal{A}$  vérifie une condition de stabilité de type  $(V, \phi, \alpha)$  et qu'il existe  $\varsigma \in [0, 1]$  et  $C_{V, \sigma} > 0$  tels que  $\text{Tr}(\sigma^*(\nabla V)^{\otimes 2} \sigma) \leq C_{V, \sigma}(\phi \circ V)V^{1-\varsigma}$ . Alors l'ensemble  $\mathcal{I}$  des mesures invariantes de (1.1) est non vide, et

$$\lim_{t \rightarrow +\infty} d_e(\nu_t, \mathcal{I}) = 0 \quad p.s. \quad (4.1)$$

où  $d_e$  est la distance de la convergence étroite. De plus, si (1.1) a une unique mesure invariante  $\nu$ , alors pour toute fonction mesurable  $f$  telle que  $|f| = o\left(\frac{\phi \circ V}{V}(\sqrt{\Psi} \circ V)\right)$  (où  $\Psi$  est définie en (1.10)) on a

$$\frac{1}{t} \int_0^t f(x_s) ds \xrightarrow{p.s.} \int_{\mathbf{R}^d} f(x) \nu(dx). \quad \star$$

**Preuve.** Par la proposition II.2, on a  $\limsup_t \nu_t \left( \frac{\phi \circ V}{V}(\sqrt{\Psi} \circ V) \right) < +\infty$  *p.s.*, d'où la tension de la suite  $(\nu_t)_{t \geq 0}$ . Le théorème de Prohorov (cf. [Bil99]) donne alors la relative compacité de  $(\nu_t)_{t \geq 0}$ . Soit  $\nu_\infty$  une valeur d'adhérence de  $(\nu_t)_{t \geq 0}$  et une sous-suite  $(\nu_{a(t)})_{t \geq 0}$  qui converge étroitement vers  $\nu_\infty$ . Alors, pour toute fonction  $f \in \mathcal{C}_K^2(\mathbf{R}^d)$  on a

$$\frac{1}{a(t)} \int_0^{a(t)} \mathcal{A} f(x_s) ds = \frac{f(x_{a(t)}) - f(x)}{a(t)} - \frac{1}{a(t)} \int_0^{a(t)} ((\nabla f)^* \sigma)(x_s) dB_s. \quad (4.2)$$

Comme  $f$  est bornée, on a  $\lim_{t \rightarrow +\infty} \frac{f(x_{a(t)}) - f(x)}{a(t)} = 0$ , et en considérant la martingale  $\left( \int_1^{a(t)} \frac{1}{a(s)} (\nabla f^* \sigma)(x_s) dB_s \right)_{t \geq 1}$  on montre que

$$\int_{\mathbf{R}^d} \mathcal{A} f d\nu_\infty = \lim_{t \rightarrow +\infty} \frac{1}{a(t)} \int_0^{a(t)} ((\nabla f)^* \sigma)(x_s) dB_s = 0. \quad (4.3)$$

D'après le corollaire II.2, la valeur d'adhérence  $\nu_\infty$  est donc une mesure invariante de (1.1). L'unicité de la mesure invariante implique alors la convergence étroite de  $(\nu_t)_{t \geq 0}$  vers  $\nu$ .  $\#$

Les méthodes introduites dans ce chapitre sont efficaces et ont l'avantage de s'appliquer à un schéma de discrétisation tel que le schéma d'Euler de pas décroissant. C'est l'objet du chapitre III. Mais il n'est *a priori* pas possible d'obtenir un théorème ergodique classique, c'est à dire valable pour toutes les fonctions intégrables par rapport à la mesure invariante  $\nu$ . C'est là une des restrictions de ces méthodes.

**Exemple.** Si on considère le processus d'Ornstein-Uhlenbeck  $(x_t)_{t \geq 0}$  solution de l'EDS de drift  $b(x) = -\frac{1}{2}x$  et  $\sigma(x) = 1$ , alors le générateur  $\mathcal{A}$  associé vérifie une condition de stabilité forte de type  $(V, \text{Id}, 1)$  avec  $V(x) = x^2 + 1$ . De plus,

$$(V')^2(x) \sigma^2(x) = 4x^2 \leq 4V(x) \quad (4.4)$$

donc la condition de domination est satisfaite avec  $C_{V,\sigma} = 4$  et  $\varsigma = 1$ . Ainsi, pour toute fonction  $f$  négligeable devant  $\left( x \rightarrow \exp\left(\frac{x^2}{4}\right) \right)$ , on a  $\lim_t \frac{1}{t} \int_0^t f(x_s) ds = \int_{\mathbf{R}} f d\nu$  *p.s.* où  $\nu$  est la mesure gaussienne centrée normalisée.  $*$

## 5. Annexe

### Proposition II.4:

Si les mesures empiriques  $(\nu_t)_{t \geq 0}$  de  $(x_t)_{t \geq 0}$  sont tendues, alors tout candidat limite est une mesure invariante pour  $(x_t)_{t \geq 0}$ .  $\blacklozenge$

**Preuve.** Soit  $t_0 > 0$  fixé et  $\mathcal{G}_k$  la filtration définie pour tout  $k \geq 0$  par  $\mathcal{G}_k = \sigma(x_s, s \leq kt_0)$ . Soit  $f \in \mathcal{C}_K(\mathbf{R}^d, \mathbf{R})$ . Montrons dans un premier temps que

$$\frac{1}{t} \int_0^t f(x_s) ds - \frac{1}{t} \int_0^t P_{t_0} f(x_s) ds \xrightarrow{p.s.} 0. \quad (5.1)$$

Pour cela, on considère la variable aléatoire  $Z_k = \int_{(k-1)t_0}^{kt_0} f(x_s) ds$ ,  $\mathcal{G}_k$ -mesurable, de sorte

que

$$\frac{1}{nt_0} \int_{t_0}^{nt_0} f(x_s) ds = \frac{1}{nt_0} \sum_{k=2}^{n+1} Z_k. \quad (5.2)$$

Par la convergence de la martingale  $\left( \sum_{k=2}^n \frac{1}{kt_0} (Z_k - \mathbf{E}[Z_k | \mathcal{G}_{k-1}]) \right)_{n \geq 0}$  et le lemme de Krockner, on montre que  $\lim_n \frac{1}{nt_0} \sum_{k=2}^n (Z_k - \mathbf{E}[Z_k | \mathcal{G}_{k-1}]) = 0$  p.s.. De même, en réitérant cet argument, on montre que

$$\frac{1}{nt_0} \sum_{k=2}^{n+1} (Z_k - \mathbf{E}_{Z_k}[\mathcal{G}_{k-2}]) \xrightarrow{p.s.} 0. \quad (5.3)$$

Pour tout  $k \geq 2$ ,  $\mathbf{E}[Z_k | \mathcal{G}_{k-2}] = \int_{(k-1)t_0}^{kt_0} \mathbf{E}[f(x_s) | \mathcal{G}_{k-2}] ds$ , et pour tout  $s \in [(k-1)t_0, kt_0]$ ,  $\mathcal{G}_{k-2} \subset \mathcal{F}_{s-t_0}$  donc

$$\mathbf{E}[f(x_s) | \mathcal{G}_{k-2}] = \mathbf{E}[\mathbf{E}[f(x_s) | \mathcal{F}_{s-t_0}] | \mathcal{G}_{k-2}] = \mathbf{E}[P_{t_0} f(x_{s-t_0}) | \mathcal{G}_{k-2}].$$

Ainsi on a

$$\mathbf{E}[Z_k | \mathcal{G}_{k-2}] = \mathbf{E} \left[ \int_{(k-1)t_0}^{kt_0} P_{t_0} f(x_{s-t_0}) ds \middle| \mathcal{G}_{k-2} \right] = \mathbf{E} \left[ \int_{(k-2)t_0}^{(k-1)t_0} P_{t_0} f(x_s) ds \middle| \mathcal{G}_{k-1} \right],$$

et on montre comme précédemment que

$$\frac{1}{nt_0} \sum_{k=2}^{n+1} \int_{(k-1)t_0}^{kt_0} (P_{t_0} f(x_s) ds - \mathbf{E}[Z_k | \mathcal{G}_{k-2}]) \xrightarrow{p.s.} 0. \quad (5.4)$$

En combinant (5.2), (5.3) et (5.4), on obtient  $\lim_n \frac{1}{nt_0} \int_{t_0}^{nt_0} f(x_s) ds - \frac{1}{nt_0} \int_{t_0}^{nt_0} P_{t_0} f(x_s) ds = 0$  p.s., d'où l'on déduit aisément (5.1) ( $f$  étant continue à support compact donc bornée).

On conclut en prenant une suite extraite  $(\nu_{a(t)})_{t \geq 0}$  de  $(\nu_t)_{t \geq 0}$  qui converge étroitement vers une limite  $\nu_\infty$  vérifiant alors

$$\int_{\mathbf{R}^d} f(x) \nu_\infty(dx) = \int_{\mathbf{R}^d} P_{t_0} f(x) \nu_\infty(dx),$$

pour tout  $t_0 \geq 0$ . La mesure  $\nu_\infty$  est donc invariante pour  $(x_t)_{t \geq 0}$ . #

**Preuve** (de la proposition II.3). On note  $E_\Delta = E \cup \{\Delta\}$  le compactifié d'Alexandroff construit en ajoutant le point  $\Delta$  à  $E$ . La topologie de  $E$  est étendue à  $E_\Delta$  en définissant les ouverts contenant  $\Delta$  comme les complémentaires des compacts de  $E$ . Cette topologie est métrisable et fait de  $E_\Delta$  un compact.

- On définit  $\mathfrak{A}_\Delta$  sur  $\mathcal{D}(\mathfrak{A}_\Delta) = \{f \in \mathcal{C}(E_\Delta), (f - f(\Delta))|_E \in \mathcal{D}(\mathfrak{A})\}$ , sous-algèbre dense de

$\mathcal{C}(E_\Delta)$ , par

$$(\mathfrak{A}_\Delta f)|_E = \mathfrak{A}((f - f(\Delta))|_E), \quad \mathfrak{A}_\Delta f(\Delta) = 0,$$

pour tout  $f \in \mathcal{C}(E_\Delta)$  tel que  $(f - f(\Delta))|_E \in \mathcal{D}(\mathfrak{A})$ , et on étend la mesure  $\mu$  à  $E_\Delta$  en posant  $\mu(\{\Delta\}) = 0$  de telle sorte que

$$\forall f \in \mathcal{D}(\mathfrak{A}_\Delta), \quad \int_{E_\Delta} \mathfrak{A}_\Delta f d\mu = 0.$$

On vérifie aisément que  $\mathcal{D}(\mathfrak{A}_\Delta)$  est dense dans  $\mathcal{C}(E_\Delta)$ , et que  $1 \in \mathcal{D}(\mathfrak{A}_\Delta)$  et  $\mathfrak{A}_\Delta 1 = 0$ . Par le théorème II.2, il existe un processus stationnaire  $(x_t)_{t \geq 0}$  solution du problème de martingale pour  $(\mathfrak{A}_\Delta, \mu)$  à valeurs dans  $E_\Delta$ .

- Pour conclure, il suffit de montrer que le processus  $(x_t)_{t \geq 0}$  est en fait à valeurs dans  $E$ , c'est à dire que  $\mathbf{P}[\exists t \geq 0, x_t = \Delta] = 0$ . Soit  $(L_p)_{p \geq 1}$  une suite de compacts tels que  $L_p \subset \overset{\circ}{L}_{p+1}$  et  $\bigcup_{p \geq 1} L_p = E$ , et  $\tau_{L_p} = \inf\{t \geq 0; x_t \notin L_p\}$ . Montrons que

$$\mathbf{P} \left[ \lim_{p \rightarrow +\infty} \tau_{L_p} = \infty \right] = 1. \quad (5.5)$$

Soit  $n \geq 0$  et  $\varphi_n = V_n - C_n \in \mathcal{D}(\mathfrak{A})$ . On prolonge par continuité  $V_n$  sur  $E_\Delta$  en posant  $\tilde{V}_n(x) = V_n(x)$  si  $x \in E$  et  $\tilde{V}_n(\Delta) = C_n$ . Alors  $\tilde{V}_n$  est une fonction de  $\mathcal{C}(E_\Delta)$  telle que  $(\tilde{V}_n - \tilde{V}_n(\Delta))|_E = \varphi_n$  donc  $\tilde{V}_n \in \mathcal{D}(\mathfrak{A}_\Delta)$ , et le processus  $\left(\tilde{V}_n(x_t) - \int_0^t \mathfrak{A}_\Delta \tilde{V}_n(x_s) ds\right)_{t \geq 0}$  est une  $(\mathcal{F}_{t+}^x)_{t \geq 0}$  martingale continue. Par définition de  $\mathfrak{A}_\Delta$ , on a  $(\mathfrak{A}_\Delta \tilde{V}_n)|_E = \mathfrak{A}\varphi_n$  donc

$$(M_t)_{t \geq 0} = \left(\tilde{V}_n(x_t) - \int_0^t (\mathfrak{A}\varphi_n)(x_s) ds\right)_{t \geq 0} \quad (5.6)$$

est une  $(\mathcal{F}_{t+}^x)_{t \geq 0}$  martingale continue à droite (càd). Ainsi, par la formule d'Itô, on a pour tout  $t \geq 0$ ,

$$\begin{aligned} \exp(-ct)\tilde{V}_n(x_t) &= \tilde{V}_n(x_0) + \int_0^t \exp(-cs) \left(-c\tilde{V}_n(x_s) + (\mathfrak{A}\varphi_n)(x_s)\right) ds \\ &\quad + \int_0^t \exp(-cs) dM_s, \end{aligned}$$

où  $\left(\int_0^t \exp(-cs) dM_s\right)_{t \geq 0}$  est une  $(\mathcal{F}_{t+}^x)_{t \geq 0}$  martingale càd. Comme  $\mathfrak{A}\varphi_n \leq cV_n$  et  $V_n \leq \tilde{V}_n$ , on en déduit que  $\left(\exp(-ct)\tilde{V}_n(x_t)\right)_{t \geq 0}$  est une surmartingale locale positive de valeur initiale intégrable car

$$\mathbf{E}_{\mu_\Delta} \left[ \tilde{V}_n(x_0) \right] = \int_{E_\Delta} \tilde{V}_n d\mu_\Delta = \int_E V_n d\mu \leq \int_E V d\mu < +\infty.$$

Par le théorème d'arrêt appliqué au temps d'arrêt borné  $t \wedge \tau_{L_p}$ , on a pour tout  $n \geq 1$  et tout  $p \geq 1$ ,

$$\mathbf{E}_{\mu_\Delta} \left[ \exp(-c(t \wedge \tau_{L_p})) \tilde{V}_n(x_{t \wedge \tau_{L_p}}) \right] \leq \int_E V d\mu, \quad (5.7)$$

et puisque  $t \wedge \tau_{L_p} \leq t$  et  $V_n \leq \tilde{V}_n$ , on a  $\mathbf{E}_{\mu_\Delta} [V_n(x_{t \wedge \tau_{L_p}})] \leq \exp(ct) \int_E V d\mu$ . Le théorème de Beppo Levi donne alors  $\mathbf{E}_{\mu_\Delta} [V(x_{t \wedge \tau_{L_p}})] \leq \exp(ct) \int_E V d\mu$ . De plus, on a  $V(x_{t \wedge \tau_{L_p}}) \geq \mathbf{1}_{\{\tau_{L_p} \leq t\}} V(x_{\tau_{L_p}})$  et  $x_{\tau_{L_p}} \notin \overset{\circ}{L}_p$ , donc

$$\mathbf{P} [\tau_{L_p} \leq t] \leq \frac{\exp(ct) \int_E V d\mu}{\inf_{y \notin \overset{\circ}{L}_p} V(y)} \leq \frac{\exp(ct) \int_E V d\mu}{\inf_{y \notin L_{p-1}} V(y)}, \quad (5.8)$$

car  $L_{p-1} \subset \overset{\circ}{L}_p$ . En faisant tendre  $p$  vers l'infini on obtient (5.5). #



Deuxième partie

Étude du schéma d'Euler de pas  
décroissant



Dans ce chapitre, on montre qu'il est possible d'intégrer numériquement des fonctions exponentielles par rapport à la probabilité invariante d'un processus  $(x_t)_{t \geq 0}$  solution d'une équation différentielle stochastique. La probabilité invariante est limite des mesures empiriques pondérées d'un schéma d'Euler de pas décroissant. Pour prouver la convergence, on étudie donc les propriétés d'ergodicité du schéma de discrétisation, et cela se fait en utilisant principalement les mêmes techniques que pour montrer l'ergodicité de  $(x_t)_{t \geq 0}$  par des méthodes de martingales (méthodes introduites dans le chapitre II).

Dans tout ce chapitre, on suppose que le processus  $(x_t)_{t \geq 0}$  est solution de l'EDS (de drift  $b$  et de coefficient de diffusion  $\sigma$ ), et que le générateur  $\mathcal{A}$  satisfait une condition de stabilité de type  $(V, \phi, \alpha)$  (cf. définition II.1). Les fonctions  $b$  et  $\sigma$  sont continues et localement Lipschitziennes. De plus, on suppose que la condition de domination  $\text{Tr}(\sigma^*(\nabla V)^{\otimes 2}\sigma) \leq C_{V,\sigma}(\phi \circ V)V^{1-\varsigma}$  est vérifiée avec  $\varsigma \in ]0, 1]$ , de sorte que toute mesure invariante  $\nu$  de l'EDS ait des moments exponentiels (plus précisément,  $\nu(\exp(\lambda V^\varsigma)) < +\infty$  p.s. pour  $\lambda < \lambda_0$ ). Le cas  $\varsigma = 0$  a été traité dans [LP02] avec  $\phi(x) = x$  (condition de stabilité forte) et dans [LP03] avec  $\phi(x) = x^a$  ( $a \in ]0, 1]$ ).

## 1. Schéma d'Euler de pas décroissant

On considère une suite  $(\gamma_n)_{n \geq 1}$ , appelée *suite de pas*, qui tend vers 0 et telle que  $\sum_{n \geq 1} \gamma_n = +\infty$ . On note  $\Gamma_n$  la somme partielle en  $n$ , i.e.  $\Gamma_n = \sum_{k=1}^n \gamma_k$ , et  $\bar{\gamma} = \max_{n \geq 1} \gamma_n$ . La suite  $(\Gamma_n)_{n \geq 1}$  permet de discrétiser le temps, et cette discrétisation est de plus en plus fine car la suite  $(\gamma_n)_{n \geq 1}$  tend vers 0. Il n'est pas nécessaire de supposer que la suite est décroissante pour l'étude du schéma  $(X_n)_{n \geq 0}$  construit de la façon suivante :

$$X_{n+1} = X_n + \gamma_{n+1}b(X_n) + \sqrt{\gamma_{n+1}}\sigma(X_n)U_{n+1}, \quad X_0 = x \in \mathbf{R}^d, \quad (1.1)$$

où  $(U_n)_{n \geq 1}$  est une suite de variables aléatoires indépendantes à valeurs dans  $\mathbf{R}^q$ , centrées et de matrice de corrélation  $\text{Id}_q$ . La suite  $(\sqrt{\gamma_n}U_n)_{n \geq 1}$  sert à discrétiser le brownien  $(B_t)_{t \geq 0}$  entre deux instants  $\Gamma_n$  et  $\Gamma_{n+1}$ , donc on pourrait supposer la variable aléatoire  $U_1$  de loi gaussienne. Cependant, on va travailler sous une hypothèse plus large qui nous permet de considérer (par exemple)  $U_1$  bornée. Dans toute la suite, on suppose que  $U_1$  est une variable

aléatoire sous-gaussienne (ou gaussienne généralisée, cf. [Sto74]), *i.e.*

$$\exists \kappa > 0, \forall \theta \in \mathbf{R}^d, \quad \mathbf{E}[\exp(\langle \theta, U_1 \rangle)] \leq \exp\left(\frac{\kappa|\theta|^2}{2}\right). \quad (1.2)$$

La suite  $(U_n)_{n \geq 1}$  sera appelée bruit blanc sous-gaussien.

On rappelle que  $U_1$ , en tant que variable aléatoire sous-gaussienne, admet un moment exponentiel d'ordre 2 au voisinage de 0, *i.e.*

$$\exists \tau > 0, \quad \mathbf{E}\left[\exp\left(\tau|U_1|^2\right)\right] < +\infty. \quad (1.3)$$

La preuve est donnée en annexe (section 5), ainsi que la preuve du lemme suivant utilisé par la suite.

**Lemme III.1:**

Il existe  $K > 0$  tel que pour tout  $\theta \in [0, \tau]$  et tout  $h \in ]0, 1[$ ,

$$\forall v \in \mathbf{R}^q, \quad \mathbf{E}\left[\exp\left(\sqrt{h}\langle v, U_1 \rangle + h\theta|U_1|^2\right)\right] \leq \exp\left(\frac{h}{1-h}\frac{\kappa}{2}|v|^2 + Kh\right). \quad \blacklozenge$$

Avant d'étudier les propriétés de stabilité du schéma, commençons par donner un résultat préliminaire sur le contrôle des accroissements de  $(X_n)_{n \geq 0}$ .

**Lemme III.2:**

Si  $V$  est une fonction positive vérifiant  $|\nabla V|^2 \leq C_V V$  et telle que  $|b|^2 \leq C_b(\phi \circ V)$  et  $\text{Tr}(\sigma\sigma^*) \leq C_\sigma(\phi \circ V)$  où  $\phi$  est une fonction positive, alors

$$\forall n \geq 1, \quad |V(X_n) - V(X_{n-1})| \leq C\sqrt{\gamma_n}\sqrt{(\phi \circ V)V}(X_{n-1})\left(1 + |U_n|^2\right) \quad \blacklozenge$$

Preuve. Par le théorème des valeurs intermédiaires, il existe  $t \in ]0, 1[$  tel que pour tout  $n \geq 1$ ,

$$V(X_n) - V(X_{n-1}) = \langle \nabla V(\xi_n), \Delta X_n \rangle, \quad (1.4)$$

avec  $\xi_n = tX_n + (1-t)X_{n-1}$ . Donc  $|\Delta V(X_n)| \leq C_V\sqrt{V}(\xi_n)|\Delta X_n|$ . Comme  $|\nabla V|^2 \leq C_V V$ ,  $\sqrt{V}$  est Lipschitz de coefficient  $L_{\sqrt{V}}$  et on a

$$|\Delta V(X_n)| \leq C_V\sqrt{V}(X_{n-1})|\Delta X_n| + C_V L_{\sqrt{V}}|\Delta X_n|^2.$$

D'autre part, puisque  $|b|^2 \leq C_b(\phi \circ V)$  et  $\text{Tr}(\sigma\sigma^*) \leq C_\sigma(\phi \circ V)$  on a

$$\begin{aligned} |\Delta X_n| &\leq \gamma_n|b(X_{n-1})| + \sqrt{\gamma_n}|\sigma(X_{n-1})U_n|, \\ &\leq C\sqrt{\gamma_n}\sqrt{\phi \circ V}(X_{n-1})(1 + |U_n|). \end{aligned}$$

La convergence vers 0 de  $(\gamma_n)_{n \geq 1}$  permet de conclure. #

Remarque: Dès ce lemme apparaît une condition de domination sur le drift  $b$  qui n'était jamais présente dans le chapitre II. Le contrôle des accroissements, et la gestion du reste d'ordre 2

dans la formule de Taylor sont spécifiques au cas discret et nous obligent à travailler sous des conditions plus restrictives sur  $b$  et  $\sigma$ . Typiquement, la fonction de Lyapounov  $V$  est quadratique i.e.  $V(x) = |x|^2 + 1$ , et  $b$  et  $\sigma$  sont sous-linéaires i.e.  $|b(x)| \leq C_b(|x| + 1)$  et  $|\sigma(x)| \leq C_\sigma(|x| + 1)$ . \*

## 2. Stabilité

On étend tout d'abord la notion de stabilité définie pour la diffusion  $(x_t)_{t \geq 0}$  au schéma d'approximation  $(X_n)_{n \geq 0}$ .

**Définition III.1.** Un schéma d'approximation  $(X_n)_{n \geq 0}$  construit à partir d'une suite de pas  $(\gamma_n)_{n \geq 1}$  vérifie une *condition de stabilité* de type  $(V, \phi, n_0)$  s'il existe une fonction  $V \in \mathcal{C}^2(\mathbf{R}^d, [1, +\infty[)$  et une fonction  $\phi \in \mathcal{C}([1, +\infty[, \mathbf{R}_+)$  telles que

1.  $\lim_{|x| \rightarrow +\infty} V(x) = +\infty$ ,
2.  $\phi$  vérifie :  $\forall x \in [1, +\infty[$ ,  $\phi(x) \leq x$  et  $\lim_{x \rightarrow +\infty} \phi(x) = +\infty$ ,
3.  $\exists \alpha > 0$  et  $\exists \beta \geq 0$  tels que  $\forall n \geq n_0$ , 
$$\frac{\mathbf{E}[V(X_{n+1}) | \mathcal{F}_n] - V(X_n)}{\gamma_{n+1}} \leq -\alpha(\phi \circ V)(X_n) + \beta.$$

Comme pour la diffusion, on dira que la condition de stabilité est forte lorsque  $\phi = \text{Id}$ .

Remarque: On rappelle que le générateur  $\mathcal{A}$  de la diffusion  $(x_t)_{t \geq 0}$  de semi-groupe de transition  $(P_t)_{t \geq 0}$  vérifie

$$\forall f \in \mathcal{C}_K^2(\mathbf{R}^d, \mathbf{R}), \quad \mathcal{A}f(x) = \lim_{t \rightarrow 0} \frac{P_t f(x) - f(x)}{t} \quad (2.1)$$

Cela souligne l'analogie entre la condition de stabilité de la diffusion et celle du schéma, et l'importance par la suite de la quantité 
$$\frac{\mathbf{E}[V(X_{n+1}) | \mathcal{F}_n] - V(X_n)}{\gamma_{n+1}}.$$
 \*

Un schéma vérifiant une condition de stabilité peut s'étudier avec les mêmes méthodes que la diffusion. Commençons par montrer un équivalent au théorème II.1.

### **Théorème III.1:**

On suppose que le schéma d'Euler construit à partir de la suite  $(\gamma_n)_{n \geq 1}$  vérifie une condition de stabilité de type  $(V, \phi, n_0)$ . Alors,

- si  $(\theta_n)_{n \geq 0}$  est une suite positive décroissante telle que  $\sum_{n \geq 1} \theta_n \gamma_n < +\infty$ , on a

$$\mathbf{E} \left[ \sum_{n \geq n_0+1} \theta_n \gamma_n (\phi \circ V)(X_{n-1}) \middle| \mathcal{F}_{n_0} \right] < +\infty \quad p.s.$$

et  $\lim_{n \rightarrow +\infty} \theta_n V(X_n) = 0$  p.s.

- si  $\mathbf{E}[V(X_{n_0})] < +\infty$  et que la condition de stabilité est forte  $\sup_{n \geq 0} \mathbf{E}[V(X_n)] < +\infty$ .

★

On prouve ce théorème à partir du lemme suivant.

**Lemme III.3:**

Soit  $V$  et  $\mathcal{V}$  deux fonctions positives de  $\mathbf{R}^d$  vérifiant : il existe  $n_0 \geq 0$ ,  $\alpha > 0$  et  $\beta \geq 0$  tels que

$$\forall n \geq n_0, \quad \mathbf{E}[V(X_{n+1}) | \mathcal{F}_n] - V(X_n) \leq -\alpha\gamma_n\mathcal{V}(X_n) + \beta\gamma_n. \quad (2.2)$$

Si  $(\theta_n)_{n \geq 0}$  est une suite décroissante positive telle que  $\sum_{n \geq 1} \theta_n \gamma_n < +\infty$ , alors

$$\mathbf{E} \left[ \sum_{n \geq n_0+1} \theta_n \gamma_n \mathcal{V}(X_{n-1}) \middle| \mathcal{F}_{n_0} \right] < +\infty \quad p.s.$$

et  $\lim_{n \rightarrow +\infty} \theta_n V(X_n) = 0$  p.s. ◆

Preuve. On considère la suite  $(S_n)_{n \geq n_0}$  définie pour tout  $n \geq n_0$  par

$$S_n = Z_n + \alpha \sum_{k=n_0+1}^n \theta_k \gamma_k \mathcal{V}(X_{k-1}) + \beta \sum_{k \geq n+1} \theta_k \gamma_k,$$

$$\text{où } Z_n = \sum_{k=n_0+1}^n \theta_k (V(X_k) - V(X_{k-1})) + \theta_{n_0} V(X_{n_0}).$$

- Montrons tout d'abord que pour tout  $n \geq n_0$ ,  $S_n \geq 0$ . Comme  $\mathcal{V}$  est positive et que les suites  $(\theta_n)_{n \geq 0}$  et  $(\gamma_n)_{n \geq 1}$  sont positives, il suffit de montrer que pour tout  $n \geq n_0$ ,  $Z_n \geq 0$ . Par la transformation d'Abel, on a

$$Z_n = \sum_{k=n_0+1}^n (\theta_{k-1} - \theta_k) V(X_{k-1}) + \theta_n V(X_n).$$

La décroissance de  $(\theta_n)_{n \geq 0}$  et la positivité de  $V$  permettent de conclure.

- D'autre part, par la condition (2.2) on a pour tout  $n \geq n_0$ ,

$$\mathbf{E}[Z_{n+1} | \mathcal{F}_n] \leq Z_n - \alpha\theta_{n+1}\gamma_{n+1}\mathcal{V}(X_n) + \beta\theta_{n+1}\gamma_{n+1}, \quad (2.3)$$

ce qui implique  $\mathbf{E}[S_{n+1} | \mathcal{F}_n] \leq S_n$ . Ainsi,  $(S_n)_{n \geq n_0}$  est une surmartingale locale positive qui converge alors presque sûrement vers une variable aléatoire positive  $S_\infty$  finie presque sûrement. De plus, pour tout  $n \geq n_0$ ,  $\mathbf{E}[S_n | \mathcal{F}_{n_0}] < +\infty$  p.s. La première partie du lemme est donc prouvée.

- On déduit aussi de la convergence de  $(S_n)_{n \geq n_0}$  la convergence presque sûre de  $(Z_n)_{n \geq n_0}$ . Comme la suite  $(\theta_n)_{n \geq 0}$  est décroissante et convergente vers 0, le lemme de Kronecker s'applique et donne la convergence presque sûre de  $(\theta_n V(X_n))_{n \geq 0}$  vers 0. #

Preuve (du théorème III.1).

- Le premier résultat s'obtient en appliquant le lemme précédent avec les fonctions  $V$  et  $\mathcal{V} = \phi \circ V$ .
- Comme la condition de stabilité est forte ( $\phi = \text{Id}$ ), il existe  $\alpha > 0$  et  $\beta > 0$  tels que

$$\forall n \geq n_0, \quad \mathbf{E}[V(X_{n+1}) | \mathcal{F}_n] \leq (1 - \alpha\gamma_{n+1})V(X_n) + \beta\gamma_{n+1},$$

En prenant l'espérance, on obtient par récurrence

$$\sup_{n \geq n_0} \mathbf{E} [V(X_n)] \leq \frac{\beta}{\alpha} \vee \mathbf{E} [V(X_{n_0})]. \quad \#$$

La proposition suivante est primordiale car elle fait le lien entre la condition de stabilité de la diffusion et celle du schéma. De plus, on voit en comparant ce résultat à la proposition II.1 qu'on peut obtenir une condition de stabilité pour le schéma avec la même classe de fonction que celle obtenue pour la diffusion.

**Proposition III.1:**

Supposons que  $\mathcal{A}$  vérifie une condition de stabilité de type  $(V, \phi, \alpha)$ . Si la fonction  $V$  est essentiellement quadratique, i.e.

$$\exists C_V > 0, |\nabla V|^2 \leq C_V V \quad \text{et} \quad \rho = \sup_{x \in \mathbf{R}^d} \|D^2 V\| < +\infty, \quad (2.4)$$

et si les coefficients  $b$  et  $\sigma$  vérifient les conditions de domination suivantes :

$$\exists C_b > 0, |b|^2 \leq C_b(\phi \circ V) \quad \text{et} \quad \exists \varsigma \in ]0, 1], \exists C_\sigma > 0, \text{Tr}(\sigma\sigma^*) \leq C_\sigma \frac{\phi \circ V}{V^\varsigma}, \quad (2.5)$$

alors le schéma d'Euler de pas décroissant  $(\gamma_n)_{n \geq 1}$  vérifie une condition de stabilité de type

$$(\exp(\lambda V^\varsigma), \phi_{\lambda, \varsigma}) \quad \text{avec} \quad \lambda < \frac{2\alpha}{\varsigma C_V C_\sigma \kappa} \wedge \frac{2\tau}{\varsigma \rho C_\sigma} \quad \text{et} \quad \phi_{\lambda, \varsigma}(x) = x \frac{\phi\left((\log(x)/\lambda)^{\frac{1}{\varsigma}}\right)}{(\log(x)/\lambda)^{\frac{1}{\varsigma}}}. \quad \blacklozenge$$

Notons que les hypothèses (2.4) et (2.5) sont bien plus fortes que l'hypothèse de domination  $\text{Tr}(\sigma^*(\nabla V)^{\otimes 2}\sigma) = \mathcal{O}((\phi \circ V)V^{1-\varsigma})$  apparaissant dans la proposition II.1. La condition (2.4) oblige le gradient de  $V$  à être Lipschitz, et la condition (2.5) implique que le drift doit être sous-linéaire. Cependant, de nombreuses diffusions vérifient ces hypothèses.

Remarque: Malgré la similarité de ce résultat avec la proposition II.1, la preuve est relativement différente. En effet, on ne peut pas remplacer la formule d'Itô appliqué à  $\exp(\lambda V^\varsigma)$  par la formule de Taylor et obtenir le résultat en contrôlant le reste d'ordre 2 : les hypothèses sur  $U_1$  seraient trop restrictives et n'inclueraient pas le cas gaussien. On procède donc différemment en appliquant la formule de Taylor à  $V$  et en utilisant la concavité de  $(x \mapsto x^\varsigma)$  et la convexité de la fonction exponentielle. \*

La condition  $\lambda < \frac{2\tau}{\varsigma \rho C_\sigma}$  est une condition technique qui n'est pas restrictive lorsque  $U_1$  est borné car dans ce cas  $\tau = +\infty$ .

Preuve. On rappelle que  $\varsigma \in ]0, 1]$ . Par concavité de la fonction  $(x \mapsto x^\varsigma)$  on a

$$V^\varsigma(X_{n+1}) - V^\varsigma(X_n) \leq \varsigma V^{\varsigma-1}(X_n) (V(X_{n+1}) - V(X_n)), \quad (2.6)$$

et la formule de Taylor appliquée à  $V$  entre  $X_n$  et  $X_{n+1}$  donne

$$V(X_{n+1}) = V(X_n) + \langle \nabla V(X_n), \Delta X_{n+1} \rangle + \frac{1}{2} D^2 V(\xi_{n+1}) \cdot (\Delta X_{n+1})^{\otimes 2},$$

avec  $\xi_{n+1} = tX_{n+1} + (1-t)X_n$ ,  $t \in ]0, 1[$ . Comme  $\rho = \sup_{x \in \mathbf{R}^d} \|D^2V(x)\| < +\infty$  on a

$$V(X_{n+1}) \leq V(X_n) + \langle \nabla V(X_n), \Delta X_{n+1} \rangle + \frac{\rho}{2} |\Delta X_{n+1}|^2. \quad (2.7)$$

Par définition du schéma  $(X_n)_{n \geq 0}$  on a

$$\begin{aligned} \Delta X_{n+1} &= \gamma_{n+1} b(X_n) + \sqrt{\gamma_{n+1}} \sigma(X_n) U_{n+1}, \\ |\Delta X_{n+1}|^2 &= \gamma_{n+1}^2 |b(X_n)|^2 + 2\gamma_{n+1}^{\frac{3}{2}} \langle b(X_n), \sigma(X_n) U_{n+1} \rangle + \gamma_{n+1} |\sigma(X_n) U_{n+1}|^2, \end{aligned}$$

donc d'après (2.7) et (2.6) on a

$$\begin{aligned} V^\varsigma(X_{n+1}) &\leq V^\varsigma(X_n) + \varsigma V^{\varsigma-1}(X_n) \left( \gamma_{n+1} \langle \nabla V, b \rangle(X_n) + \sqrt{\gamma_{n+1}} \langle \nabla V(X_n), \sigma(X_n) U_{n+1} \rangle \right. \\ &\quad \left. + \frac{\rho}{2} \gamma_{n+1}^2 |b(X_n)|^2 + \rho \gamma_{n+1}^{\frac{3}{2}} \langle b(X_n), \sigma(X_n) U_{n+1} \rangle \right. \\ &\quad \left. + \frac{\rho}{2} \gamma_{n+1} \operatorname{Tr}(\sigma \sigma^*)(X_n) |U_{n+1}|^2 \right). \end{aligned}$$

- Comme  $\mathcal{A}$  vérifie une condition de stabilité de type  $(V, \phi, \alpha)$  on a

$$\begin{aligned} \gamma_{n+1} \langle \nabla V, b \rangle(X_n) &\leq -\alpha \gamma_{n+1} (\phi \circ V)(X_n) + \beta \gamma_{n+1} - \frac{1}{2} \gamma_{n+1} \operatorname{Tr}(\sigma^*(D^2V)\sigma)(X_n), \\ &\leq -\alpha \gamma_{n+1} (\phi \circ V)(X_n) + \beta \gamma_{n+1} + C \gamma_{n+1} \left( \frac{\phi \circ V}{V^\varsigma} \right)(X_n), \end{aligned}$$

d'après (2.4) et (2.5). De plus,  $|b(x)|^2 \leq C_b (\phi \circ V)(x)$  donc en notant  $\psi_V = \frac{\phi \circ V}{V}$ , on a  $(\phi \circ V) V^{\varsigma-1} = \psi_V V^\varsigma$  et

$$\begin{aligned} V^\varsigma(X_{n+1}) &\leq V^\varsigma(X_n) - \varsigma \alpha \gamma_{n+1} (\psi_V V^\varsigma)(X_n) + \beta \varsigma \gamma_{n+1} V^{\varsigma-1}(X_n) + C \gamma_{n+1} \psi_V(X_n) \\ &\quad + \frac{\varsigma \rho C_b}{2} \gamma_{n+1}^2 (\psi_V V^\varsigma)(X_n) + \Lambda_{n+1}(X_n, U_{n+1}), \quad (2.8) \end{aligned}$$

avec

$$\Lambda_n(x, y) = \varsigma \sqrt{\gamma_n} V^{\varsigma-1}(x) \langle \nabla V(x) + \rho \gamma_n b(x), \sigma(x) y \rangle + \frac{\varsigma \rho}{2} \gamma_n V^{\varsigma-1}(x) \operatorname{Tr}(\sigma \sigma^*)(x) |y|^2.$$

Comme pour tout  $x \in \mathbf{R}^d$ ,  $\operatorname{Tr}(\sigma \sigma^*)(x) \leq C_\sigma (\phi \circ V)(x) V^{-\varsigma}(x)$ , on a

$$\begin{aligned} \Lambda_n(x, y) &\leq \varsigma \sqrt{\gamma_n} V^{\varsigma-1}(x) \langle \nabla V(x) + \rho \gamma_n b(x), \sigma(x) y \rangle + C \gamma_n \left( \frac{\phi \circ V}{V} \right)(x) |y|^2, \\ &\leq \varsigma \sqrt{\gamma_n} V^{\varsigma-1}(x) \langle \nabla V(x) + \rho \gamma_n b(x), \sigma(x) y \rangle + C \gamma_n \psi_V(x) |y|^2. \quad (2.9) \end{aligned}$$

Soit  $\lambda \in ]0, \frac{2\tau}{\varsigma\rho C_\sigma} \wedge \frac{2\alpha}{\varsigma C_V C_\sigma \kappa}]$ . En combinant (2.9) et (2.8) on obtient pour tout  $n \geq 0$ ,

$$\mathbf{E} \left[ \exp(\lambda V^\varsigma(X_{n+1})) \middle| \mathcal{F}_n \right] \leq \exp \left( \lambda V^\varsigma(X_n) - \lambda \varsigma \alpha \gamma_{n+1} (\psi_V V^\varsigma)(X_n) + C \gamma_{n+1} \psi_V(X_n) \right. \\ \left. + \lambda \varsigma \beta \gamma_{n+1} V^{\varsigma-1}(X_n) + \frac{\lambda \varsigma \rho C_b}{2} \gamma_{n+1}^2 (\psi_V V^\varsigma)(X_n) \right) L_{n+1}(X_n) \quad (2.10)$$

avec  $L_n(x) = \mathbf{E} \left[ \exp \left( \lambda \varsigma \sqrt{\gamma_n} V^{\varsigma-1}(x) \langle \nabla V(x) + \rho \gamma_n b(x), \sigma(x) U_1 \rangle + \frac{\lambda \varsigma \rho C_\sigma}{2} \gamma_n \psi_V(x) |U_1|^2 \right) \right]$ .

- Majorons maintenant  $L_n(x)$ . La suite  $(\gamma_n)_{n \geq 1}$  converge vers 0 donc il existe  $\tilde{n}_0 \geq 0$  tel que  $\forall n \geq \tilde{n}_0$ ,  $\gamma_n < 1$ . De plus,  $\lambda < \frac{2\tau}{\varsigma\rho C_\sigma}$  donc on peut appliquer le lemme III.1 avec  $v = \frac{\lambda \varsigma}{\psi_V(x)} V^{\varsigma-1}(x) \sigma^*(x) (\nabla V(x) + \rho \gamma_n b(x))$ ,  $h = \gamma_n \psi_V(x)$  (car  $\psi_V \leq 1$ ) et  $\theta = \frac{\lambda \varsigma \rho C_\sigma}{2}$  ce qui donne l'existence de  $K > 0$  tel que pour tout  $n \geq \tilde{n}_0$  et tout  $x \in \mathbf{R}^d$

$$L_n(x) \leq \exp \left( \frac{\gamma_n \psi_V(x)}{1 - \gamma_n \psi_V(x)} \frac{\kappa}{2} |v|^2 + K \gamma_n \psi_V(x) \right). \quad (2.11)$$

D'une part, on a

$$\psi_V(x) |v|^2 \leq \lambda^2 \varsigma^2 V^{2(\varsigma-1)}(x) \text{Tr}(\sigma \sigma^*)(x) \left( |\nabla V(x)|^2 + 2\rho \gamma_n \langle \nabla V, b \rangle(x) + \rho^2 \gamma_n^2 |b(x)|^2 \right), \\ \left( \frac{\phi \circ V}{V} \right)(x) |v|^2 \leq \lambda^2 \varsigma^2 C_\sigma V^{\varsigma-2}(x) (\phi \circ V)(x) \left( C_V V(x) + 2\rho \beta \gamma_n + \rho^2 \gamma_n^2 C_b (\phi \circ V)(x) \right),$$

et en posant  $C_1 = \lambda^2 \varsigma^2 C_\sigma \rho^2 C_b$  et  $C_2 = 2\lambda^2 \varsigma^2 C_\sigma \rho \beta$ , on a

$$|v|^2 \leq \lambda^2 \varsigma^2 C_\sigma C_V V^\varsigma(x) + C_1 \gamma_n^2 (\psi_V V^\varsigma)(x) + C_2 \gamma_n V^{\varsigma-1}(x), \\ \leq \lambda^2 \varsigma^2 C_\sigma C_V V^\varsigma(x) + C_1 \gamma_n^2 (\psi_V V^\varsigma)(x) + C_2 \gamma_n, \quad (2.12)$$

car  $\psi_V \leq 1$ ,  $V \geq 1$  et  $\varsigma \leq 1$ .

D'autre part, on a

$$\frac{\gamma_n \psi_V}{1 - \gamma_n \psi_V} = \gamma_n \psi_V + \frac{\gamma_n^2 \psi_V^2}{1 - \gamma_n \psi_V}.$$

Par la décroissance de la suite  $(\gamma_n)_{n \geq 1}$  et par (2.12), on déduit de (2.11) qu'il existe  $C_1 > 0$  et  $C_2 > 0$  tels que pour tout  $n \geq \tilde{n}_0$  et tout  $x \in \mathbf{R}^d$

$$L_n(x) \leq \exp \left( \frac{\lambda^2 \varsigma^2 C_\sigma C_V \kappa}{2} \gamma_n (\psi_V V^\varsigma)(x) + C_1 \gamma_n^2 (\psi_V V^\varsigma)(x) + C_2 \gamma_n \psi_V(x) \right). \quad (2.13)$$

- En combinant (2.13) et (2.10), on montre aisément qu'il existe  $C_1 > 0$  et  $C_2 > 0$  (nouvelles constantes) telles que pour tout  $n \geq n_0$  et tout  $x \in \mathbf{R}^d$ ,

$$\mathbf{E} \left[ \exp(\lambda V^\varsigma(X_{n+1})) \middle| \mathcal{F}_n \right] \leq \exp \left( \lambda V^\varsigma(X_n) - \lambda \varsigma \alpha \gamma_{n+1} (\psi_V V^\varsigma)(X_n) + \lambda \varsigma \beta \gamma_{n+1} V^{\varsigma-1}(X_n) \right. \\ \left. + \frac{\lambda^2 \varsigma^2 C_\sigma C_V \kappa}{2} \gamma_{n+1} (\psi_V V^\varsigma)(X_n) + C_1 \gamma_{n+1}^2 (\psi_V V^\varsigma)(x) + C_2 \gamma_{n+1} \psi_V(X_n) \right).$$

Comme  $\lambda < \frac{2\alpha}{aC_V C_\sigma \kappa}$ , on pose  $\bar{\alpha} = \zeta\alpha - \frac{\lambda\zeta^2 C_\sigma C_V \kappa}{2} > 0$  et on a pour tout  $n \geq n_0$ ,

$$\mathbf{E} \left[ \exp(\lambda V^\zeta(X_{n+1})) \middle| \mathcal{F}_n \right] \leq \exp \left( \lambda V^\zeta(X_n) - \lambda \bar{\alpha} \gamma_{n+1} (\psi_V V^\zeta)(X_n) + \lambda \zeta \beta \gamma_{n+1} V^{\zeta-1}(X_n) \right. \\ \left. + C_1 \gamma_{n+1}^2 (\psi_V V^\zeta)(X_n) + C_2 \gamma_{n+1} \psi_V(X_n) \right).$$

Soit  $\tilde{\alpha} \in ]0, \bar{\alpha}[$ . On réécrit alors cette inégalité de la façon suivante

$$\mathbf{E} \left[ \exp(\lambda V^\zeta(X_{n+1})) \middle| \mathcal{F}_n \right] \leq \exp \left( (1 - \tilde{\alpha} \gamma_{n+1} \psi_V(X_n)) \lambda V^\zeta(X_n) \right. \\ \left. + \tilde{\alpha} \gamma_{n+1} \psi_V(X_n) R_{n+1}(X_n) \right),$$

en posant  $R_n(x) = \frac{1}{\tilde{\alpha}} \left( \lambda (\tilde{\alpha} - \bar{\alpha}) V^\zeta(x) + \lambda \zeta \beta \frac{V^{\zeta-1}(x)}{\psi_V(x)} + C_1 \gamma_n V^\zeta(x) + C_2 \right)$ . Soit  $n_1 \geq n_0$  tel que  $\forall n \geq n_1, \gamma_n \leq \tilde{\alpha}^{-1}$ . Alors, par convexité de la fonction exponentielle on pour tout  $n \geq n_1$ ,

$$\mathbf{E} \left[ \exp(\lambda V^\zeta(X_{n+1})) \middle| \mathcal{F}_n \right] \leq (1 - \tilde{\alpha} \gamma_{n+1} \psi_V(X_n)) \exp(\lambda V^\zeta(X_n)) \\ + \tilde{\alpha} \gamma_{n+1} \psi_V(X_n) \exp(R_{n+1}(X_n)). \quad (2.14)$$

- Il reste à montrer que pour  $n$  suffisamment grand, la fonction  $R_n$  est majorée. Comme  $(\gamma_n)_{n \geq 1}$  tend vers 0, et que  $\tilde{\alpha} < \bar{\alpha}$ , il existe  $n_2 \geq 0$  et  $\varepsilon > 0$  tels que

$$\forall n \geq n_2, \quad R_n(x) \leq R_\varepsilon(x) = -\varepsilon V^\zeta(x) + \frac{\lambda \zeta \beta}{\tilde{\alpha}} \frac{V^\zeta}{(\phi \circ V)(x)} + \frac{C_2}{\tilde{\alpha}}.$$

La fonction  $R_\varepsilon$  est continue et négative hors d'un compact car  $\phi \circ V$  et  $V$  tendent vers l'infini à l'infini. Ainsi,  $R_\varepsilon$  est majorée par  $\bar{\beta}$  sur  $\mathbf{R}^d$ . À partir de (2.14), on obtient le résultat annoncé en prenant  $\tilde{\beta} = \tilde{\alpha} \exp(\bar{\beta})$  et  $n \geq n_1 \wedge n_2$ . #

**Remarque:** Le cas  $\zeta = 0$  se démontre exactement de la même façon. Dans ce cas, le schéma d'Euler  $(X_n)_{n \geq 0}$  construit à partir de la suite  $(\gamma_n)_{n \geq 1}$  vérifie une condition de stabilité de type  $(V^p, \phi_p)$  avec  $p \in ]1, 1 + \frac{2\alpha}{C_V C_\sigma}[$ . Il est intéressant de remarquer que lorsque  $C_\sigma$  tend vers 0 (c'est à dire que le coefficient de diffusion est négligeable devant  $\sqrt{\phi \circ V}$ ), alors  $p$  tend vers l'infini. \*

Contrairement au cas polynômial sous une condition de stabilité forte traité dans [LP02], il n'est pas immédiat d'obtenir  $\mathbf{E}[W(X_n)] < +\infty$  où  $W$  désigne  $\exp(\lambda V^\zeta)$  dans notre cadre et  $V^p$  dans le cadre polynômial. En effet, l'inégalité (2.14) n'est pas exploitable car elle est vraie pour  $n \geq n_1$  et on n'a *a priori* aucune information sur  $\mathbf{E}[W(X_{n_1})]$ . Il est toutefois possible d'obtenir la  $\mathbf{L}^1$ -bornitude des  $W(X_n)$  pour des  $W = \exp(\lambda V^\zeta)$  où  $\lambda$  est suffisamment petit. C'est l'objet du corollaire suivant.

**Corollaire III.1:**

On se place sous les hypothèses de la proposition III.1. Si la condition de stabilité est forte, alors pour tout  $\lambda < \frac{2\alpha}{\varsigma C_V C_{\sigma\kappa}} \wedge \frac{2\tau}{\varsigma \rho C_{\sigma}}$ ,  $\sup_{n \geq 0} \mathbf{E} [\exp(\lambda V^{\varsigma}(X_n))] < +\infty$ . •

Preuve. On note  $\bar{\lambda} = \frac{2\alpha}{\varsigma C_V C_{\sigma\kappa}} \wedge \frac{2\tau}{\varsigma \rho C_{\sigma}}$ .

- Montrons tout d'abord que pour tout  $n \geq 0$

$$\forall \lambda \in ]0, \bar{\lambda}[, \exists \lambda' \in ]\lambda, \bar{\lambda}[, \quad \mathbf{E} \left[ \exp(\lambda V^{\varsigma}(X_{n+1})) \middle| \mathcal{F}_n \right] \leq C \exp(\lambda' V^{\varsigma}(X_n)). \quad (2.15)$$

Soit  $n \geq 0$ . Par le théorème des valeurs intermédiaires appliqué à la fonction  $(x \mapsto \exp(\lambda x))$ , il existe  $Y_{n+1} \in (V^{\varsigma}(X_n), V^{\varsigma}(X_{n+1}))$  tel que

$$\exp(\lambda V^{\varsigma}(X_{n+1})) - \exp(\lambda V^{\varsigma}(X_n)) = \lambda \exp(\lambda Y_{n+1}) (V^{\varsigma}(X_{n+1}) - V^{\varsigma}(X_n))$$

Or, par le lemme III.2, on sait que  $|\Delta V(X_n)| \leq C \sqrt{\gamma_n} V(X_{n-1}) (1 + |U_n|^2)$  donc par concavité de  $(x \mapsto x^{\varsigma})$  on obtient

$$\exp(\lambda V^{\varsigma}(X_{n+1})) - \exp(\lambda V^{\varsigma}(X_n)) \leq \lambda \varsigma C \sqrt{\gamma_{n+1}} V^{\varsigma}(X_n) \exp(\lambda Y_{n+1}) (1 + |U_{n+1}|^2).$$

Par croissance et positivité de l'exponentielle,  $\exp(\lambda Y_{n+1}) \leq \exp(\lambda V^{\varsigma}(X_n)) + \exp(\lambda V^{\varsigma}(X_{n+1}))$ , donc en injectant cette majoration dans l'inégalité précédente et en prenant l'espérance sachant  $\mathcal{F}_n$ , il existe  $C_1 > 0$  et  $C_2 > 0$  tels que

$$\begin{aligned} \mathbf{E} \left[ \exp(\lambda V^{\varsigma}(X_{n+1})) \middle| \mathcal{F}_n \right] - \exp(\lambda V^{\varsigma}(X_n)) &\leq C_1 V^{\varsigma}(X_n) \exp(\lambda V^{\varsigma}(X_n)) \\ &\quad + C_2 V^{\varsigma}(X_n) \mathbf{E} \left[ \exp(\lambda V^{\varsigma}(X_{n+1})) (1 + |U_{n+1}|^2) \middle| \mathcal{F}_n \right]. \end{aligned} \quad (2.16)$$

Par l'inégalité de Young, on a pour tout  $\varepsilon > 0$ ,

$$\mathbf{E} \left[ \exp(\lambda V^{\varsigma}(X_{n+1})) (1 + |U_{n+1}|^2) \middle| \mathcal{F}_n \right] \leq \frac{1}{1 + \varepsilon} \mathbf{E} \left[ \exp(\lambda(1 + \varepsilon) V^{\varsigma}(X_{n+1})) \middle| \mathcal{F}_n \right] + \frac{\varepsilon}{1 + \varepsilon} K_{\varepsilon},$$

avec  $K_{\varepsilon} = \mathbf{E} \left[ (1 + |U_1|^2)^{(1+\varepsilon)/\varepsilon} \right]$ . Pour  $\varepsilon$  tel que  $\lambda(1 + \varepsilon) < \bar{\lambda}$ , on a par la proposition III.1

$$\mathbf{E} [\exp(\lambda(1 + \varepsilon) V^{\varsigma}(X_{n+1})) \mid \mathcal{F}_n] \leq \exp(\lambda(1 + \varepsilon) V^{\varsigma}(X_n)) + \beta \gamma_{n+1}.$$

Par ces deux inégalités combinées avec (2.16), on en déduit (2.15) (avec  $\lambda' \in ]\lambda(1 + \varepsilon), \bar{\lambda}[$ ).

- En prenant l'espérance, et par une simple récurrence, on déduit de (2.15) que pour tout  $\lambda \in ]0, \bar{\lambda}[$  et pour tout  $n \geq 0$ ,  $\mathbf{E} [\lambda V^{\varsigma}(X_n)] < +\infty$ . Le corollaire est obtenu en appliquant le théorème III.1. #

Remarque: Ce résultat n'est pas forcément nécessaire pour l'étude des mesures empiriques pondérées mais il est tout de même important de noter que  $(V(X_n))_{n \geq 0}$  a des moments exponentiels. \*

### 3. Mesures empiriques pondérées

On introduit maintenant les mesures empiriques pondérées du schéma  $(X_n)_{n \geq 0}$ . On se donne une suite  $(\eta_n)_{n \geq 1}$  de réels positifs tels que  $\lim_{n \rightarrow +\infty} H_n = +\infty$  où  $H_n = \sum_{k=1}^n \eta_k$ , et on définit la suite  $(\nu_n^\eta)_{n \geq 1}$  des mesures empiriques de  $(X_n)_{n \geq 0}$  pondérées par  $(\eta_n)_{n \geq 1}$  de la façon suivante :

$$\forall n \geq 1, \quad \nu_n^\eta(dx) = \frac{1}{H_n} \sum_{k=1}^n \eta_k \delta_{X_{k-1}}(dx). \quad (3.1)$$

On appellera souvent  $(\eta_n)_{n \geq 1}$  la *suite de poids* associée au schéma  $(X_n)_{n \geq 0}$ .

**Remarque:** La suite  $(\eta_n)_{n \geq 1}$  n'est pas supposée convergente vers 0. Un choix naturel est de prendre  $\eta_n = 1$  pour tout  $n \geq 1$ , ou bien  $(\eta_n)_{n \geq 1} = (\gamma_n)_{n \geq 1}$ . Le fait d'étudier la convergence de  $(\nu_n^\eta)_{n \geq 1}$  pour toute une classe de suite de poids est motivé par l'étude des vitesses de convergence qui permet d'obtenir une suite de poids optimale. \*

Pour  $s \in ]1, 2]$ , on introduit les conditions suivantes sur la suite de pas  $(\gamma_n)_{n \geq 1}$  et la suite de poids  $(\eta_n)_{n \geq 1}$  :

$$\sum_{n \geq 1} \frac{1}{H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ < +\infty \quad \text{et} \quad \sum_{n \geq 1} \left( \frac{\eta_n}{H_n \sqrt{\gamma_n}} \right)^s < +\infty, \quad (C_s^{\eta, \gamma})$$

et

$$\left( \frac{1}{\gamma_n H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ \right)_{n \geq 1} \quad \text{et} \quad \left( \frac{1}{\gamma_n} \left( \frac{\eta_n}{H_n \sqrt{\gamma_n}} \right)^s \right)_{n \geq 1} \quad \text{sont décroissantes.} \quad (D_s^{\eta, \gamma})$$

La condition  $(C_s^{\eta, \gamma})$  est suffisante pour montrer la tension lorsque la condition de stabilité est forte ( $\phi = \text{Id}$ ). Dans le cas contraire, il faut lui ajouter la condition  $(D_s^{\eta, \gamma})$  afin de pouvoir appliquer le théorème III.1.

Ces conditions ne sont pas restrictives, et de nombreux choix sont possibles. Par exemple si on considère des suites de la forme "puissance" i.e.  $\gamma_n = n^{-p}$  avec  $p \in ]0, 1]$  et  $\eta_n = n^{-q}$  avec  $q \leq 1$ , les conditions  $(C_s^{\eta, \gamma})$  et  $(D_s^{\eta, \gamma})$  sont vérifiées si et seulement si

$$(q, p) \in \left( 0, \frac{2(s-1)}{s} \right) \times (-\infty, 1] \cup \left\{ \left( \frac{2(s-1)}{s}, 1 \right) \right\}.$$

#### 3.1. Tension

La tension s'obtient en suivant le même schéma que la diffusion. La proposition ci-dessous est l'analogie de la proposition II.2.

##### **Proposition III.2:**

On suppose que  $\mathcal{A}$  vérifie une condition de stabilité de type  $(V, \phi, \alpha)$ , et que les hypothèses de croissance (2.4) et (2.5) de la proposition III.1 sont satisfaites. Soit  $\bar{\lambda} = \frac{2\alpha}{\varsigma C_V C_\sigma \kappa} \wedge \frac{2\tau}{\varsigma \rho C_\sigma}$

et  $\lambda < \bar{\lambda}/s$ . Si les suites  $(\gamma_n)_{n \geq 1}$  et  $(\eta_n)_{n \geq 1}$  vérifient pour un  $s \in ]1, 2]$ ,

$$\begin{cases} (C_s^{\eta, \gamma}) & \text{si } \sup_{n \geq 0} \mathbf{E}[\exp(\lambda V^s(X_n))] < +\infty, \\ (C_s^{\eta, \gamma}) \text{ et } (D_s^{\eta, \gamma}) & \text{sinon,} \end{cases} \quad (3.2)$$

alors

$$\sup_{n \geq 1} \frac{1}{H_n} \sum_{k=1}^n \eta_k \left( \frac{\phi \circ V}{V} \exp(\lambda V^s) \right) (X_{k-1}) < +\infty \quad p.s. \quad \blacklozenge$$

L'analogie de la martingale continue  $(M_t)_{t \geq 0}$  utilisée dans la preuve de la proposition II.2 est la martingale discrète  $(M_n)_{n \geq 0}$  nulle en 0 et définie pour tout  $n \geq 1$  par

$$M_n = \sum_{k=1}^n \frac{\eta_k}{H_k \gamma_k} \left( W_{\lambda_0, s}(X_k) - \mathbf{E}[W_{\lambda_0, s}(X_k) | \mathcal{F}_{k-1}] \right).$$

Pour les martingales discrètes, le théorème de Chow (rappelé en annexe) permet d'obtenir la convergence de la martingale sous une condition plus faible que la finitude du crochet (critère que l'on a utilisé dans le cas continu). Pour cette raison, il est possible de prouver la tension pour des fonctions plus grandes que dans le cas continu. Cela introduit par contre une restriction sur les pas et les poids : c'est le rôle du paramètre  $s$  introduit dans les hypothèses  $(C_s^{\eta, \gamma})$  et  $(D_s^{\eta, \gamma})$ . Une des difficultés de la preuve est la vérification de la condition du théorème de Chow.

**Preuve.** Pour tout  $\lambda < \bar{\lambda}$ , on note  $W_{\lambda, s} = \exp(\lambda V^s)$  et  $\phi_{\lambda, s}$  de sorte que  $\phi_{\lambda, s}(W_{\lambda, s}) = \left( \frac{\phi \circ V}{V} \right) W_{\lambda, s}$ , c'est-à-dire  $\phi_{\lambda, s}(x) = x \frac{\phi \left( (\log(x)/\lambda)^{\frac{1}{s}} \right)}{(\log(x)/\lambda)^{\frac{1}{s}}}$ .

Par la proposition III.1, on sait que le schéma vérifie une condition de stabilité de type  $(W_{\lambda, s}, \phi_{\lambda, s})$  pour tout  $\lambda < \bar{\lambda}$ .

Soit  $\lambda_0 < \bar{\lambda}/s$ , alors il existe  $n_0 \geq 0$ ,  $\tilde{\alpha} > 0$  et  $\tilde{\beta} > 0$  tels que

$$\forall n \geq n_0, \quad \frac{\mathbf{E} \left[ W_{\lambda_0, s}(X_{n+1}) \middle| \mathcal{F}_n \right] - W_{\lambda_0, s}(X_n)}{\gamma_n} \leq -\tilde{\alpha} (\phi_{\lambda_0, s} \circ W_{\lambda_0, s})(X_n) + \tilde{\beta}, \quad (3.3)$$

donc

$$\frac{1}{H_n} \sum_{k=1}^n \eta_k (\phi_{\lambda_0, s} \circ W_{\lambda_0, s})(X_{k-1}) \leq \frac{1}{\tilde{\alpha} H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} \left( W_{\lambda_0, s}(X_{k-1}) - \mathbf{E} \left[ W_{\lambda_0, s}(X_k) \middle| \mathcal{F}_{k-1} \right] \right) + \frac{\tilde{\beta}}{\tilde{\alpha}}.$$

Ainsi, il suffit de montrer que

$$\sup_{n \geq 1} \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} \left( W_{\lambda_0, s}(X_{k-1}) - \mathbf{E} \left[ W_{\lambda_0, s}(X_k) \middle| \mathcal{F}_{k-1} \right] \right) < +\infty \quad p.s. \quad (3.4)$$

Pour cela, on introduit  $W_{\lambda_0, s}(X_k)$  (le retranchant et l'ajoutant) dans la différence et on étudie les deux sommes moyennisées obtenues.

- D'une part, par la transformation d'Abel on a, en posant  $\eta_0 = 0$ ,

$$\begin{aligned} \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (W_{\lambda_0, \varsigma}(X_{k-1}) - W_{\lambda_0, \varsigma}(X_k)) &= \frac{1}{H_n} \sum_{k=1}^n \left( \Delta \frac{\eta_k}{\gamma_k} \right) W_{\lambda_0, \varsigma}(X_{k-1}) - \frac{\eta_n}{\gamma_n} W_{\lambda_0, \varsigma}(X_n), \\ &\leq \frac{1}{H_n} \sum_{k=1}^n \left( \Delta \frac{\eta_k}{\gamma_k} \right)_+ W_{\lambda_0, \varsigma}(X_{k-1}). \end{aligned} \quad (3.5)$$

Prouvons que  $\sum_{n \geq 1} \frac{1}{H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ W_{\lambda_0, \varsigma}(X_{n-1}) < +\infty$  *p.s.*

- Si  $\sup_n \mathbf{E}[W_{\lambda_0, \varsigma}(X_n)] < +\infty$ , alors

$$\sum_{n \geq 1} \frac{1}{H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ \mathbf{E}[W_{\lambda_0, \varsigma}(X_{n-1})] \leq \left( \sum_{n \geq 1} \frac{1}{H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ \right) \sup_{n \geq 0} \mathbf{E}[W_{\lambda_0, \varsigma}(X_n)],$$

et par  $(C_s^{\eta, \gamma})$  on a  $\sum_{n \geq 1} \frac{1}{H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ \mathbf{E}[W_{\lambda_0, \varsigma}(X_{n-1})] < +\infty$ . On conclut par le théorème de convergence monotone.

- Dans le cas général, on considère  $\lambda \in ]\lambda_0, \bar{\lambda}[$ . La condition de stabilité de type  $(W_{\lambda, \varsigma}, \phi_{\lambda, \varsigma})$  et les hypothèses  $(C_s^{\eta, \gamma})$  et  $(D_s^{\eta, \gamma})$  impliquent, par le théorème III.1 appliqué avec

$$(\theta_n)_{n \geq 1} = \left( \frac{1}{\gamma_n H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ \right)_{n \geq 1},$$

$$\sum_{n \geq 1} \frac{1}{H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ (\phi_{\lambda, \sigma} \circ W_{\lambda, \varsigma})(X_{n-1}) < +\infty \quad p.s.$$

On conclut aisément en remarquant que  $W_{\lambda_0, \varsigma} = o(\phi_{\lambda, \sigma} \circ W_{\lambda, \varsigma})$ .

Le lemme de Kronecker appliqué avec la suite  $(H_n)_{n \geq 1}$  donne alors

$$\lim_{n \rightarrow +\infty} \frac{1}{H_n} \sum_{k=1}^n \left( \Delta \frac{\eta_k}{\gamma_k} \right)_+ W_{\lambda_0, \varsigma}(X_{k-1}) = 0 \quad p.s.$$

et d'après (3.5),

$$\overline{\lim}_{n \rightarrow +\infty} \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (W_{\lambda_0, \varsigma}(X_{k-1}) - W_{\lambda_0, \varsigma}(X_k)) \leq 0 \quad p.s.$$

- Montrons maintenant que  $\lim_{n \rightarrow +\infty} \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (W_{\lambda_0, \varsigma}(X_k) - \mathbf{E}[W_{\lambda_0, \varsigma}(X_k) | \mathcal{F}_{k-1}]) = 0$  *p.s.*. Pour cela, on considère la martingale  $(M_n)_{n \geq 0}$  nulle en 0 et définie pour tout  $n \geq 1$  par

$$M_n = \sum_{k=1}^n \frac{\eta_k}{H_k \gamma_k} (W_{\lambda_0, \varsigma}(X_k) - \mathbf{E}[W_{\lambda_0, \varsigma}(X_k) | \mathcal{F}_{k-1}]). \quad (3.6)$$

Par le théorème de Chow (cf. théorème III.3), la martingale  $(M_n)_{n \geq 0}$  converge presque sûre-

ment si

$$\sum_{n \geq 1} \left( \frac{\eta_n}{H_n \gamma_n} \right)^s \mathbf{E} \left[ \left( W_{\lambda_0, \varsigma}(X_n) - \mathbf{E}[W_{\lambda_0, \varsigma}(X_n) | \mathcal{F}_{n-1}] \right)^s \middle| \mathcal{F}_{n-1} \right] < +\infty \quad p.s. \quad (3.7)$$

Il suffit donc de prouver (3.7) et d'appliquer le lemme de Kronecker à (3.6) avec la suite  $(H_n)_{n \geq 0}$ , pour obtenir le résultat souhaité.

- Montrons (3.7). En notant  $\Delta W_{\lambda_0, \varsigma}(X_n) = W_{\lambda_0, \varsigma}(X_n) - W_{\lambda_0, \varsigma}(X_{n-1})$ , on a par convexité de  $(x \mapsto x^s)$  ( $s > 1$ )

$$\left| W_{\lambda_0, \varsigma}(X_n) - \mathbf{E}[W_{\lambda_0, \varsigma}(X_n) | \mathcal{F}_{n-1}] \right|^s \leq 2^{s-1} \left( \left| \Delta W_{\lambda_0, \varsigma}(X_n) \right|^s + \left| \mathbf{E}[\Delta W_{\lambda_0, \varsigma}(X_n) | \mathcal{F}_{n-1}] \right|^s \right),$$

et par l'inégalité de Jensen

$$\left| \mathbf{E}[\Delta W_{\lambda_0, \varsigma}(X_n) | \mathcal{F}_{n-1}] \right|^s \leq \mathbf{E} \left[ \left| \Delta W_{\lambda_0, \varsigma}(X_n) \right|^s \middle| \mathcal{F}_{n-1} \right],$$

donc

$$\mathbf{E} \left[ \left| W_{\lambda_0, \varsigma}(X_n) - \mathbf{E}[W_{\lambda_0, \varsigma}(X_n) | \mathcal{F}_{n-1}] \right|^s \middle| \mathcal{F}_{n-1} \right] \leq 2^s \mathbf{E} \left[ \left| \Delta W_{\lambda_0, \varsigma}(X_n) \right|^s \middle| \mathcal{F}_{n-1} \right]. \quad (3.8)$$

La formule de Taylor appliquée à la fonction  $(x \rightarrow \exp(\lambda_0 x))$  entre  $V^\varsigma(X_n)$  et  $V^\varsigma(X_{n-1})$  donne

$$\left| \Delta W_{\lambda_0, \varsigma}(X_n) \right| \leq \lambda_0 (\exp(\lambda_0 V^\varsigma(X_n)) + \exp(\lambda_0 V^\varsigma(X_{n-1}))) \left| \Delta V^\varsigma(X_n) \right|. \quad (3.9)$$

Par le lemme III.2, on sait que  $|\Delta V(X_n)| \leq C \sqrt{\gamma_n} \sqrt{(\phi \circ V)V}(X_{n-1}) (1 + |U_n|^2)$  donc par le théorème des accroissements finis appliqué avec  $(x \mapsto x^\varsigma)$ , par  $V \geq 1$  et par  $\phi(x) \leq x$  on a

$$\left| \Delta V^\varsigma(X_n) \right| \leq \varsigma C \sqrt{\gamma_n} V(X_{n-1}) (1 + |U_n|^2).$$

Ainsi, de (3.9) on obtient

$$\begin{aligned} \mathbf{E} \left[ \left| \Delta W_{\lambda_0, \varsigma}(X_n) \right|^s \middle| \mathcal{F}_{n-1} \right] &\leq C \gamma_n^{\frac{\varsigma}{2}} \left( V^\varsigma(X_{n-1}) \mathbf{E} \left[ W_{\lambda_0, \varsigma}^s(X_n) (1 + |U_n|^2)^s \middle| \mathcal{F}_{n-1} \right] \right. \\ &\quad \left. + V^\varsigma(X_{n-1}) W_{\lambda_0, \varsigma}^s(X_{n-1}) \mathbf{E} \left[ (1 + |U_n|^2)^s \middle| \mathcal{F}_{n-1} \right] \right). \end{aligned} \quad (3.10)$$

Soit  $\lambda \in ]s\lambda_0, \bar{\lambda}[$ . Il est clair  $V^s W_{\lambda_0, \varsigma}^s = \mathcal{O}(\exp(\lambda V^\varsigma))$  donc

$$V^\varsigma(X_{n-1}) W_{\lambda_0, \varsigma}^s(X_{n-1}) \mathbf{E} \left[ (1 + |U_n|^2)^s \middle| \mathcal{F}_{n-1} \right] \leq C \exp(\lambda V^\varsigma(X_{n-1})). \quad (3.11)$$

D'autre part, par l'inégalité de Young, on a pour tout  $\varepsilon > 0$

$$\mathbf{E} \left[ W_{\lambda_0, \varsigma}^s(X_n) (1 + |U_n|^2)^s \middle| \mathcal{F}_{n-1} \right] \leq \frac{1}{1 + \varepsilon} \mathbf{E} \left[ W_{\lambda_0, \varsigma}^{s(1+\varepsilon)}(X_n) \middle| \mathcal{F}_{n-1} \right] + \frac{\varepsilon}{1 + \varepsilon} K_\varepsilon, \quad (3.12)$$

avec  $K_\varepsilon = \mathbf{E} \left[ (1 + |U_1|^2)^{s(1+\varepsilon)/\varepsilon} \right]$ . On choisit  $\varepsilon$  tel que  $\lambda_0 s(1 + \varepsilon) < \lambda$  de sorte que l'al-

gorithme vérifie une condition de stabilité de type  $(W_{\lambda_0 s(1+\varepsilon), \varsigma}, \phi_{\lambda_0 s(1+\varepsilon), \varsigma})$ . Or, il est clair que

$$V^s \left( \frac{\phi \circ V}{V} \right) \exp(\lambda_0 s(1+\varepsilon)V^\varsigma) \leq C \exp(\lambda V^\varsigma), \quad (3.13)$$

donc en insérant (3.13), (3.12) et (3.11) dans (3.10), on obtient

$$\mathbf{E} [|\Delta W_{\lambda_0, \varsigma}(X_n)|^s \mid \mathcal{F}_{n-1}] \leq C_1 \gamma_n^{s/2} \exp(\lambda V^\varsigma(X_{n-1})) + C_2 \gamma_n^{s/2}.$$

Soit par le théorème de convergence monotone combiné avec  $(C_s^{\eta, \gamma})$ , soit par le théorème III.1 combiné avec  $(C_s^{\eta, \gamma})$  et  $(D_s^{\eta, \gamma})$ , on en déduit que

$$\sum_{n \geq 1} \left( \frac{\eta_n}{H_n \gamma_n} \right)^s \mathbf{E} [|\Delta W_{\lambda_0, \varsigma}(X_n)|^s \mid \mathcal{F}_{n-1}] < +\infty,$$

c'est à dire (3.7). #

### 3.2. Identification de la limite

Il reste à prouver que toute valeur d'adhérence de  $(\nu_n^\eta)_{n \geq 1}$  est probabilité invariante du processus  $(x_t)_{t \geq 0}$ . Soit  $\nu_\infty$  une valeur d'adhérence de  $(\nu_n^\eta)_{n \geq 1}$ . D'après le théorème d'Echeverria-Weiss, exprimé en terme de fonction de Lyapounov dans le corollaire II.2, il suffit de montrer que pour toute fonction  $f \in C_K^2(\mathbf{R}^d)$ ,  $\int_{\mathbf{R}^d} \mathcal{A}f d\nu_\infty = 0$ . C'est l'objet de la proposition suivante.

#### Proposition III.3:

Sous les hypothèses de la proposition III.2, on a

$$\forall f \in \mathcal{C}_K^2(\mathbf{R}^d), \quad \lim_{n \rightarrow +\infty} \frac{1}{H_n} \sum_{k=1}^n \eta_k \mathcal{A}f(X_{k-1}) = 0 \quad p.s. \quad \blacklozenge$$

On commence par prouver le lemme suivant.

#### Lemme III.4:

Sous les hypothèses de la proposition III.3, on a

$$\forall f \in \mathcal{C}_K^2(\mathbf{R}^d), \quad \lim_{n \rightarrow +\infty} \frac{1}{H_n} \sum_{k=1}^n \eta_k \left( \frac{\mathbf{E}[f(X_k) \mid \mathcal{F}_{k-1}] - f(X_{k-1})}{\gamma_k} - \mathcal{A}f(X_k) \right) = 0 \quad p.s. \quad \blacklozenge$$

Preuve. Par la formule de Taylor appliquée à  $f$  entre  $X_{n-1}$  et  $X_n$  on a

$$\begin{aligned} \mathbf{E}[f(X_n) \mid \mathcal{F}_{n-1}] - f(X_{n-1}) &= \gamma_n \mathcal{A}f(X_{n-1}) + \frac{1}{2} \gamma_n^2 D^2 f(X_{n-1}) \cdot b(X_{n-1})^{\otimes 2} \\ &\quad + \mathbf{E}[R_2(X_{n-1}, X_n) \mid \mathcal{F}_{n-1}], \end{aligned}$$

avec  $R_2(x, y) = f(y) - f(x) - \langle \nabla f(x), y - x \rangle - \frac{1}{2} D^2 f(x) \cdot (y - x)^{\otimes 2}$ . Il faut donc montrer que

$$\lim_{n \rightarrow +\infty} \frac{1}{H_n} \sum_{k=1}^n \eta_k \left( \gamma_k D^2 f(X_{k-1}) \cdot b(X_{k-1})^{\otimes 2} + \frac{\mathbf{E}[R_2(X_{k-1}, X_k) | \mathcal{F}_{k-1}]}{\gamma_k} \right) = 0 \quad p.s.$$

- Comme  $D^2 f$  est une fonction à support compact, on a

$$\left| \frac{1}{H_n} \sum_{k=1}^n \eta_k \gamma_k D^2 f(X_{k-1}) \cdot b^{\otimes 2}(X_{k-1}) \right| \leq \|D^2 f \cdot b^{\otimes 2}\|_{\infty} \frac{1}{H_n} \sum_{k=1}^n \eta_k \gamma_k \xrightarrow{p.s.} 0, \quad (3.14)$$

car  $(\gamma_n)_{n \geq 1}$  tend vers 0.

- Il reste à montrer que  $\lim_n \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} \mathbf{E}[R_2(X_{k-1}, X_k) | \mathcal{F}_{k-1}] = 0$  *p.s.* On introduit la fonction continue  $r_2$  définie sur  $\mathbf{R}^d \times \mathbf{R}_+$  par  $r_2(x, \delta) = \sup_{z \in \mathbf{R}^d, |z-x| < \delta} |D^2 f(z) - D^2 f(x)|$ . La fonction  $r_2$  est croissante en  $\delta$  et bornée, et vérifie

$$\forall (x, y) \in \mathbf{R}^{2d}, \quad |R_2(x, y)| \leq \frac{1}{2} r_2(x, |y - x|) |y - x|^2$$

Comme  $|X_k - X_{k-1}|^2 \leq C \gamma_k V(X_{k-1}) (1 + |U_1|^2)$ , il suffit de prouver que

$$\frac{1}{H_n} \sum_{k=1}^n \eta_k V(X_{k-1}) \int_{\mathbf{R}^q} r_2(X_{k-1}, \delta_k(u)) (1 + |u|^2) \mu(du) \xrightarrow{p.s.} 0, \quad (3.15)$$

où  $\delta_k(u) = \gamma_k b(X_{k-1}) + \sqrt{\gamma_k} \sigma(X_{k-1}) u$  et où  $\mu(du)$  est la loi de  $U_1$ . D'une part, pour  $A > 0$  fixé, la suite  $(\delta_k(u) \mathbf{1}_{\{|X_{k-1}| \leq A\}})_{k \geq 1}$  converge presque sûrement vers 0, donc

$$\lim_{k \rightarrow +\infty} r_2(X_{k-1}, \delta_k(u)) \mathbf{1}_{\{|X_{k-1}| \leq A\}} = 0 \quad p.s.$$

De plus,  $r_2$  est bornée et  $U_1$  admet un moment d'ordre 2, donc par le théorème de convergence dominée on en déduit que

$$\lim_{k \rightarrow +\infty} \int_{\mathbf{R}^q} r_2(X_{k-1}, \delta_k(u)) \mathbf{1}_{\{|X_{k-1}| \leq A\}} (1 + |u|^2) \mu(du) = 0 \quad p.s.$$

et par suite

$$\frac{1}{H_n} \sum_{k=1}^n \eta_k V(X_{k-1}) \int_{\mathbf{R}^q} r_2(X_{k-1}, \delta_k(u)) \mathbf{1}_{\{|X_{k-1}| \leq A\}} (1 + |u|^2) \mu(du) \xrightarrow{p.s.} 0.$$

D'autre part, il existe  $\lambda > 0$  tel que  $\sup_{n \geq 0} \nu_n^\eta(\exp(\lambda V)) < +\infty$  *p.s.* donc

$$\begin{aligned} \overline{\lim}_n \frac{1}{H_n} \sum_{k=1}^n \eta_k V(X_{k-1}) \int_{\mathbf{R}^q} r_2(X_{k-1}, \delta_k(u)) \mathbf{1}_{\{|X_{k-1}| > A\}} (1 + |u|^2) \mu(du) \\ \leq C \|r_2\|_{\infty} \sup_{|x| > A} \left( \frac{V(x)}{\exp(\lambda V(x))} \right) \sup_n \nu_n^\eta(\exp(\lambda V)). \end{aligned}$$

On conclut en faisant tendre  $A$  tend vers l'infini. #

Preuve (proposition III.3). D'après le lemme III.4, il suffit de prouver que

$$\lim_{n \rightarrow +\infty} \frac{1}{H_n} \sum_{k=1}^n \eta_k \frac{\mathbf{E}[f(X_k) | \mathcal{F}_{k-1}] - f(X_{k-1})}{\gamma_k} = 0 \quad p.s.$$

pour toute fonction  $f \in \mathcal{C}_K^2(\mathbf{R}^d)$ .

- Tout d'abord montrons que  $\lim_n \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (f(X_k) - f(X_{k-1})) = 0$  p.s. En posant  $\eta_0 = 0$ , on obtient par la transformation d'Abel

$$\begin{aligned} \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (f(X_{k-1}) - f(X_k)) &= \frac{1}{H_n} \left( \sum_{k=1}^n \left( \Delta \frac{\eta_k}{\gamma_k} \right) f(X_{k-1}) - \frac{\eta_n}{\gamma_n} f(X_n) \right), \\ &= \frac{1}{H_n} \left( \sum_{k=1}^n \left( \Delta \frac{\eta_k}{\gamma_k} \right) \tilde{f}(X_{k-1}) - \frac{\eta_n}{\gamma_n} \tilde{f}(X_n) \right), \end{aligned}$$

avec  $\tilde{f} = f + \|f\|_\infty$ . Alors  $\tilde{f} \geq 0$  et on a

$$\frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (f(X_{k-1}) - f(X_k)) \leq \frac{1}{H_n} \sum_{k=1}^n \left( \Delta \frac{\eta_k}{\gamma_k} \right)_+ \tilde{f}(X_{k-1})$$

On conclut en utilisant la bornitude de  $\tilde{f}$ ,  $(C_s^{\eta, \gamma})$  et le lemme de Kronecker (appliqué avec la suite  $(H_n)_{n \geq 0}$ ).

- On introduit la martingale  $(M_n)_{n \geq 0}$  nulle en 0 et définie pour tout  $n \geq 1$  par

$$M_n = \sum_{k=1}^n \frac{\eta_k}{H_k \gamma_k} (\mathbf{E}[f(X_k) | \mathcal{F}_{k-1}] - f(X_k)).$$

Cette martingale se gère de la même façon que dans la preuve de la proposition III.2. On montre par convexité de  $(x \rightarrow x^s)$  et l'inégalité de Jensen que

$$\sum_{n \geq 1} \mathbf{E}[|\Delta M_n|^s | \mathcal{F}_{n-1}] \leq 4 \sum_{n \geq 1} \left( \frac{\eta_n}{H_n \gamma_n} \right)^s \mathbf{E}[(f(X_n) - f(X_{n-1}))^s | \mathcal{F}_n].$$

Or, la fonction  $f$  est Lipschitz, et  $|\Delta X_n| \leq C \sqrt{\gamma_n} (\phi \circ V)(X_{n-1}) (1 + |U_n|^2)$ . Donc

$$\sum_{n \geq 1} \mathbf{E}[|\Delta M_n|^s | \mathcal{F}_{n-1}] \leq C \sum_{n \geq 1} \left( \frac{\eta_n}{H_n \sqrt{\gamma_n}} \right)^s (\phi \circ V)(X_{n-1}) \quad p.s.$$

On conclut comme dans III.2 via le théorème III.1. #

Remarque: Faire intervenir la fonction  $\tilde{f}$  dans la première partie de la preuve nous permet d'obtenir l'identification de la limite sous les mêmes conditions sur les suites  $(\gamma_n)_{n \geq 1}$  et  $(\eta_n)_{n \geq 1}$  que celles nécessaires à la tension de  $(\nu_n^\eta)_{n \geq 0}$ . Notamment, la condition (peu restrictive)

$\lim_n \frac{1}{H_n} \sum_{k=1}^n \left| \Delta \frac{\eta_k}{\gamma_k} \right| = 0$  requise dans [LP02] et [LP03] n'est plus nécessaire. \*

## 4. Conclusion

On vient de voir que la notion de stabilité et les méthodes de martingales introduites pour l'étude de la diffusion se transposent très bien à l'étude d'un schéma de discrétisation tel que le schéma d'Euler de pas décroissant. La décroissance du pas de discrétisation vers 0 permet d'approcher de façon directe le régime stationnaire de la diffusion. On obtient notamment le théorème ergodique suivant qui résume l'étude faite dans ce chapitre.

### Théorème III.2:

Supposons que  $\mathcal{A}$  vérifie une condition de stabilité de type  $(V, \phi, \alpha)$  avec  $V$  essentiellement quadratique (i.e. vérifiant (2.4)) et que les conditions suivantes sont vérifiées :

$$\exists C_b > 0, |b|^2 \leq C_b(\phi \circ V) \quad \text{et} \quad \exists \varsigma \in ]0, 1], \exists C_{V,\sigma} > 0, \text{Tr}(\sigma\sigma^*) \leq C_\sigma \frac{\phi \circ V}{V^\varsigma}.$$

Si les suites  $(\gamma_n)_{n \geq 1}$  et  $(\eta_n)_{n \geq 1}$  vérifient les conditions  $(C_s^{\eta,\gamma})$  et  $(D_s^{\eta,\gamma})$  pour un  $s \in ]1, 2]$ , alors toute valeur d'adhérence des mesures empiriques pondérées  $(\nu_n^\eta)_{n \geq 1}$  est mesure invariante de l'EDS.

Si la mesure invariante  $\nu$  est unique, alors pour toute fonction mesurable  $f$  telle que  $f = o\left(\frac{\phi \circ V}{V} \exp(\lambda V^\varsigma)\right)$  avec  $\lambda < \frac{1}{s} \left( \frac{2\alpha}{\varsigma C_V C_\sigma \kappa} \wedge \frac{2\tau}{\varsigma \rho C_\sigma} \right)$ , on a

$$\frac{1}{H_n} \sum_{k=1}^n \eta_k f(X_{k-1}) \xrightarrow{p.s.} \int_{\mathbf{R}^d} f(x) \nu(dx). \quad \star$$

On rappelle que la condition  $(D_s^{\eta,\gamma})$  n'est pas nécessaire si la condition de stabilité est forte (cf. corollaire III.1).

Ces techniques de martingales sont donc très efficaces pour obtenir des approximations de la mesure invariante. De plus, elles peuvent aussi être utilisées lorsque l'EDS est dirigée par un processus de Lévy à la place d'un brownien (voir les travaux de F. Panloup dans cette direction).

*A priori*, il est impossible par ces méthodes de montrer un "vrai" théorème ergodique, c'est à dire de moyenniser toute fonction de  $\mathbf{L}^1(\nu)$ . Cependant, on peut s'approcher d'un tel résultat pour une classe de diffusion restreinte en prouvant que le schéma d'Euler converge presque sûrement et suffisamment vite vers la diffusion. C'est l'objet du chapitre suivant.

## 5. Annexe

### Lemme III.5 (de Kronecker):

Soient  $(a_n)_{n \geq 0}$  et  $(b_n)_{n \geq 0}$  deux suites réelles. Si  $(b_n)_{n \geq 0}$  est une suite positive croissante vers  $+\infty$  telle que la série  $\sum_{n \geq 0} \frac{a_n}{b_n}$  converge, alors  $\lim_n \frac{1}{a_n} \sum_{k=0}^n a_k = 0$ .  $\blacklozenge$

**Théorème III.3 (de Chow):**

Soit  $(M_n)_{n \geq 0}$  une  $(\mathcal{F}_n)_{n \geq 0}$ -martingale. Alors

$$\forall s \in ]1, 2], \quad M_n \xrightarrow{p.s} M_\infty \text{ sur l'événement } \left\{ \sum_{n \geq 1} \mathbf{E}[|\Delta M_n|^s | \mathcal{F}_{n-1}] < +\infty \right\}. \quad \star$$

**Rappels sur une variable aléatoire sous-gaussienne**

Montrons qu'une variable aléatoire sous-gaussienne  $U_1$  de  $\mathbf{R}^q$  admet un moment exponentiel d'ordre 2 au voisinage de 0, *i.e.*

$$\exists \tau > 0, \quad \mathbf{E} \left[ \exp \left( \tau |U_1|^2 \right) \right] < +\infty. \quad (5.1)$$

D'après la définition (1.2), on a pour tout  $\theta \in \mathbf{R}^q$  et pour tout  $\beta > 0$

$$\begin{aligned} \forall t \geq 0, \quad \mathbf{P} \left[ |\langle \theta, U_1 \rangle| \geq t \right] &= \mathbf{P} \left[ \exp(\beta |\langle \theta, U_1 \rangle|) \geq \exp(\beta t) \right], \\ &\leq \exp \left( -\beta t + \kappa \beta^2 \frac{|\theta|^2}{2} \right). \end{aligned}$$

En optimisant le terme de droite de l'inégalité précédent en  $\beta$ , on obtient la majoration

$$\forall \theta \in \mathbf{R}^q, \quad \forall t \geq 0, \quad \mathbf{P} \left[ |\langle \theta, U_1 \rangle| \geq t \right] \leq \exp \left( -\frac{t^2}{2\kappa|\theta|^2} \right). \quad (5.2)$$

Le nom de variable aléatoire sous-gaussienne provient de cette inégalité qui stipule que la queue de distribution de  $U_1$  se trouve sous celle de la distribution gaussienne.

D'autre part, l'inégalité Hölder donne

$$\mathbf{E} \left[ \exp \left( \tau |U_1|^2 \right) \right] \leq \prod_{i=1}^d \left( \mathbf{E} \left[ \exp \left( d\tau (U_1^i)^2 \right) \right] \right)^{1/d} \quad (5.3)$$

avec  $U_1^i = \langle U_1, e_i \rangle$  où  $e_i$  est le  $i$ -ème vecteur de la base canonique orthonormée de  $\mathbf{R}^d$ . Or, pour tout  $i \in \{1, \dots, d\}$

$$\mathbf{E} \left[ \exp \left( d\tau (U_1^i)^2 \right) \right] = \int_1^{+\infty} \mathbf{P} \left[ |U_1^i| \geq \sqrt{(d\tau)^{-1} \ln(t)} \right] dt,$$

et d'après (5.2) appliqué avec  $\theta = e_i$ , on obtient

$$\forall i \in \{1, \dots, d\}, \quad \mathbf{E} \left[ \exp \left( d\tau (U_1^i)^2 \right) \right] \leq \int_1^{+\infty} \exp \left( -\frac{\ln(t)}{2\kappa d\tau} \right) dt.$$

Pour tout  $\tau < \frac{1}{2\kappa d}$ , l'intégrale majorante dans l'inégalité précédente est finie, et d'après (5.3) on a (5.1).

On prouve maintenant le lemme III.1 qui est une simple application de l'inégalité de Hölder et de l'existence d'un moment d'ordre 2 au voisinage de 0 pour  $U_1$ .

Preuve (lemme III.1). Par l'inégalité de Hölder, on a

$$\mathbf{E} \left[ \exp \left( \sqrt{h} \langle v, U_1 \rangle + h\theta |U_1|^2 \right) \right] \leq \left( \mathbf{E} \left[ \exp \left( \frac{\sqrt{h}}{1-h} \langle v, U_1 \rangle \right) \right] \right)^{1-h} \left( \mathbf{E} \left[ \exp \left( \theta |U_1|^2 \right) \right] \right)^h.$$

Comme  $U_1$  est une sous-gaussienne, alors

$$\mathbf{E} \left[ \exp \left( \frac{\sqrt{h}}{1-h} \langle v, U_1 \rangle \right) \right] \leq \exp \left( \frac{h}{(1-h)^2} \frac{\kappa}{2} |v|^2 \right),$$

et puisque  $U_1$  satisfait (2.4) et  $\theta \leq \tau$ ,  $\mathbf{E} \left[ \exp \left( \theta |U_1|^2 \right) \right] < +\infty$ . Le lemme est obtenu en posant  $K = \log \left( \mathbf{E} \left[ \exp \left( \tau |U_1|^2 \right) \right] \right)$ . #

**Exemples.**

1. Une variable aléatoire gaussienne est sous-gaussienne avec  $\kappa = 1$ .
2. Une variable aléatoire bornée est sous-gaussienne avec  $\kappa = +\infty$  (lemme de Hoeffding).\*



---

# Approche ergodique pour une classe de diffusion

---

IV

La classe de diffusion que nous considérons dans ce chapitre a été introduite et étudiée par Basak et Bhattacharya dans [BB92]. Ces diffusions (appelées par la suite  $(V, \alpha)$ -confluentes) ont la propriété d'oublier leur point de départ. Toutes les trajectoires convergent presque sûrement vers une unique trajectoire. Cette propriété implique l'unicité de la mesure invariante (même dans des situations où le coefficient de diffusion est dégénéré). On rappelle que par le théorème ergodique (cf. [Has80])

$$\forall f \in \mathbf{L}^1(\nu), \quad \frac{1}{t} \int_0^t f(x_s) ds \xrightarrow{p.s.} \int f d\nu,$$

où  $(x_t)_{t \geq 0}$  est une diffusion  $(V, \alpha)$ -confluente et  $\nu$  son unique mesure invariante. L'idée est de montrer que le schéma d'Euler de pas décroissant  $(X_n)_{n \geq 0}$  vérifie

$$\frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k f(X_{k-1}) - \frac{1}{\Gamma_n} \int_0^{\Gamma_n} f(x_s) ds \xrightarrow{p.s.} 0 \quad (*)$$

pour une classe de fonctions  $f$ , plus grande que celle considérée dans le théorème III.2 (la restriction sur le  $\lambda$  optimal due à l'utilisation du théorème de Chow est ici levée).

Dans [BHW97], les auteurs montrent que le schéma d'Euler de pas décroissant converge en loi vers la mesure invariante. Nous allons montrer que si le bruit blanc est construit à partir des accroissements du brownien dirigeant l'EDS, alors le schéma d'Euler converge dans  $\mathbf{L}^p$  et presque sûrement vers la diffusion. De plus, la convergence est suffisamment rapide pour obtenir (\*) et ainsi un théorème ergodique pour les mesures empiriques du schéma d'Euler.

## 1. Diffusion $(V, \alpha)$ -confluente

On considère l'équation différentielle stochastique suivante

$$dx_t = b(x_t)dt + \sigma(x_t)dB_t, \quad (1.1)$$

où  $(B_t)_{t \geq 0}$  est un brownien standard de  $\mathbf{R}^q$  et  $b$  et  $\sigma$  sont des applications lipschitziennes respectivement de  $\mathbf{R}^d$  dans  $\mathbf{R}^d$  et de  $\mathbf{R}^d$  dans  $\mathcal{M}(d \times q)$  l'ensemble des matrices de  $d$  lignes et  $q$  colonnes.

**Définition IV.1 (( $V, \alpha$ )-confluente).** La diffusion solution de l'EDS (1.1) de drift  $b$  et de coefficient de diffusion  $\sigma$  est dite ( $V, \alpha$ )-confluente si

- $V \in \mathcal{C}^2(\mathbf{R}^d, \mathbf{R}_+)$  est une fonction nulle uniquement en 0 vérifiant  $\lim_{|x| \rightarrow +\infty} V(x) = +\infty$ ,
- pour tout  $x \in \mathbf{R}^d$  et tout  $y \in \mathbf{R}^d$ ,

$$\langle b_{x,y}, \nabla V(x-y) \rangle + \frac{1}{2} \text{Tr}(\sigma_{x,y}^* \text{D}^2 V(x-y) \sigma_{x,y}) \leq -\alpha V(x-y),$$

où  $b_{x,y} = b(x) - b(y)$  et  $\sigma_{x,y} = \sigma(x) - \sigma(y)$ . ♣

On peut voir cette condition comme une condition de stabilité pour le processus  $(x_t^x - x_t^y)_{t \geq 0}$  où  $(x_t^x)_{t \geq 0}$  est une solution de (1.1) partant de  $x$  et  $(x_t^y)_{t \geq 0}$  une solution partant de  $y$ . Il est très important de noter que  $V$  est nulle uniquement en 0. Dans de nombreux cas,  $V$  est quadratique de la forme  $V(x) = |x|^2$  ou bien  $V(x) = \langle x, Sx \rangle$  où  $S$  est une matrice définie positive. On prouve les résultats de convergence dans le cas plus général où  $V$  est sous-quadratique.

On introduit la notation

$$\forall f \in \mathcal{C}^2, \quad \tilde{\mathcal{A}}f(x, y) = \langle b_{x,y}, \nabla f(x-y) \rangle + \frac{1}{2} \text{Tr}(\sigma_{x,y}^* \text{D}^2 f(x-y) \sigma_{x,y}), \quad (1.2)$$

qui correspond au générateur infinitésimal de  $(x_t^x - x_t^y)_{t \geq 0}$ .

Tout d'abord, montrons qu'une diffusion ( $V, \alpha$ )-confluente vérifie une condition de stabilité lorsque  $V$  est sous-quadratique.

**Proposition IV.1:**

Si une diffusion est ( $V, \alpha$ )-confluente avec  $V$  sous-quadratique i.e.

$$|\nabla V|^2 \leq C_V V \quad \text{et} \quad \|\text{D}^2 V\|_\infty < +\infty, \quad (1.3)$$

et que  $\text{Tr}(\sigma\sigma^*) \leq CV$ , alors elle vérifie une condition de stabilité forte<sup>1</sup> de type  $(V + K, \text{Id}, \bar{\alpha})$  avec  $\bar{\alpha} < \alpha$  et  $K \geq 1$ . ♦

Preuve. Par définition de la ( $V, \alpha$ )-confluence, on a pour tout  $x \in \mathbf{R}^d$ ,

$$\langle b(x) - b(0), \nabla V(x) \rangle + \frac{1}{2} \text{Tr}(\sigma_{x,0}^* \text{D}^2 V(x) \sigma_{x,0}) \leq -\alpha V(x),$$

avec  $\sigma_{x,0} = \sigma(x) - \sigma(0)$ . De plus, par linéarité de l'application  $\text{Tr}$ , on a

$$\mathcal{A}V(x) = \tilde{\mathcal{A}}V(x, 0) + \langle b(0), \nabla V(x) \rangle + \text{Tr}(\sigma^*(0) \text{D}^2 V(x) \sigma(x)) - \frac{1}{2} \text{Tr}(\sigma^*(0) \text{D}^2 V(x) \sigma(0))$$

donc d'après (1.3)

$$\mathcal{A}V(x) \leq -\alpha V(x) + |b(0)| \sqrt{C_V} \sqrt{V}(x) + C \sqrt{V}(x) + \frac{1}{2} d \|\text{D}^2 V\|_\infty |\sigma(0)|^2.$$

---

<sup>1</sup>cf. définition II.1

Comme  $V$  tend vers l'infini à l'infini, on en déduit aisément que  $\mathcal{A}$  vérifie une condition de stabilité forte de type  $(V + K, \text{Id}, \bar{\alpha})$  avec  $\bar{\alpha} < \alpha$  et  $K \geq 1$ . #

Dans ce cadre, on sait qu'il existe au moins une solution stationnaire à l'EDS (1.1) et par le théorème II.1, on sait de plus que  $\sup_{t \geq 0} \mathbf{E}[V(x_t)] < +\infty$  et que  $\lim_{t \rightarrow +\infty} \theta(t)V(x_t) = 0$  p.s. pour

toute fonction  $\theta \in \mathcal{C}^1(\mathbf{R}_+; \mathbf{R}_+^*)$  décroissante telle que  $\int_0^{+\infty} \theta(s)ds < +\infty$ .

Le fait que  $(x_t)_{t \geq 0}$  soit une diffusion  $(V, \alpha)$ -confluente est une condition bien plus forte, et implique l'unicité de la probabilité invariante. En effet, on prouve que toute solution de (1.1) partant d'un point quelconque de  $\mathbf{R}^d$  converge en moyenne vers la solution stationnaire (d'où l'appellation "asymptotic flatness of the stochastic flow" introduite dans [BB92]). Plus précisément on montre que pour tout compact  $K$  de  $\mathbf{R}^d$ ,

$$\lim_{t \rightarrow +\infty} \sup_{x, y \in K} \mathbf{E}[V(x_t^x - x_t^y)] = 0. \quad (1.4)$$

Pour obtenir ce résultat, il suffit d'appliquer la formule d'Itô au processus  $(z_t^{x,y})_{t \geq 0} = (x_t^x - x_t^y)_{t \geq 0}$  avec la fonction  $g(t, x) = \exp(\alpha t)V(x)$ . Alors

$$\mathbf{E}[\exp(\alpha(t \wedge \tau_n)V(z_{t \wedge \tau_n}^{x,y}))] \leq V(x - y), \quad (1.5)$$

où  $\tau_n = \inf\{t \geq 0, |z_t^{x,y}| \geq n\}$ , et on conclut par le lemme de Fatou.

Cela traduit la principale propriété d'une diffusion  $(V, \alpha)$ -confluente : 0 est un point attractif asymptotiquement stable en moyenne pour tout processus de la forme  $(x_t^x - x_t^y)_{t \geq 0}$  avec  $x, y \in \mathbf{R}^d$ . C'est cette propriété qui permet d'obtenir des résultats de convergence trajectoirelle du schéma d'Euler de pas décroissant vers la diffusion.

Remarque: La condition de  $(V, \alpha)$ -attractivité implique l'unicité de la mesure invariante même lorsque le coefficient de diffusion peut s'annuler, c'est à dire dans le cas dégénéré. \*

**Exemple.** Si  $b(x) = Ax$  avec  $A$  une matrice attractive (i.e. toute les valeurs propres ont une partie réelle strictement négative) et si  $\sigma$  est constante (possiblement dégénérée), alors la diffusion est  $(V, \alpha)$ -confluente avec  $V(x) = \langle x, Sx \rangle$  où

$$S = \int_0^{+\infty} \exp(sA^*) \exp(sA) ds, \quad (1.6)$$

et  $\alpha = \frac{1}{\lambda_S}$  où  $\lambda_S$  est la plus grande valeur propre de  $S$ .

En effet,  $S$  est l'unique solution de l'équation de Lyapounov  $SA + A^*S = -\text{Id}_d$  (cf. [Duf96]). De plus,  $S$  est symétrique donc  $2\langle Sx, Ax \rangle = \langle x, (SA + A^*S)x \rangle = -|x|^2 \leq -\frac{1}{\lambda_S}V(x)$ . Ainsi pour tous  $x$  et  $y$ , on a

$$\langle b(x) - b(y), \nabla V(x - y) \rangle \leq -\frac{1}{\lambda_S}V(x - y). \quad (1.7)$$

Si  $\sigma$  est lipschitzienne de constante de Lipschitz  $[\sigma]_1$ , alors pour  $[\sigma]_1$  suffisamment petit la diffusion est  $(V, \tilde{\alpha})$ -confluente avec  $\tilde{\alpha} < \alpha$ . \*

## 2. Convergence en moyenne du schéma d'Euler de pas décroissant

Soit  $(\gamma_n)_{n \geq 1}$  une suite tendant vers 0 terme principal d'une série divergente. On note  $\Gamma_n = \sum_{k=1}^n \gamma_k$  (et  $\Gamma_0 = 0$ ), et on introduit le schéma construit à partir de la suite de pas  $(\gamma_n)_{n \geq 1}$  :

$$X_{\Gamma_{n+1}} = X_{\Gamma_n} + \gamma_{n+1}b(X_{\Gamma_n}) + \sigma(X_{\Gamma_n})U_{n+1}, \quad X_{\Gamma_0} = x,$$

où  $(U_n)_{n \geq 1}$  est la suite définie pour tout  $n \geq 1$  par  $U_n = B_{\Gamma_n} - B_{\Gamma_{n-1}}$ . Dans ce chapitre, afin d'obtenir des résultats de convergence trajectorielle on considère que le bruit blanc est construit à partir des accroissements du brownien  $(B_t)_{t \geq 0}$  entre deux instants  $\Gamma_n$  et  $\Gamma_{n+1}$ . Il devrait être possible de généraliser les résultats qui suivent à un bruit blanc gaussien en utilisant les techniques de plongements browniens de type Skorokhod.

La différence de notation " $X_{\Gamma_n}$ " avec le schéma " $X_n$ " introduit dans les chapitres précédents est motivée par la construction d'un schéma d'Euler à temps continu. On construit le schéma d'Euler continu en extrapolant le schéma discret  $(X_{\Gamma_n})_{n \geq 0}$ . Ainsi pour tout  $t \in [\Gamma_n, \Gamma_{n+1}[$  on définit

$$X_t = X_{\Gamma_n} + b(X_{\Gamma_n})(t - \Gamma_n) + \sigma(X_{\Gamma_n})(B_t - B_{\Gamma_n}).$$

En introduisant la notation  $\underline{t} = \Gamma_n$  si  $t \in [\Gamma_n, \Gamma_{n+1}[$ , le schéma s'écrit

$$\forall t \geq 0, \quad dX_t = b(X_{\underline{t}})dt + \sigma(X_{\underline{t}})dB_t.$$

En outre, on note :  $E_t = V(X_t - X_{\underline{t}})$  et  $\mathcal{E}_t = V(x_t - X_t)$ . On rappelle que  $V$  s'annule en 0, donc  $(E_t)_{t \geq 0}$  désigne l'erreur (mesurée par  $V$ ) entre le schéma d'Euler discret et le schéma continu, et  $(\mathcal{E}_t)_{t \geq 0}$  désigne l'erreur analogue entre le schéma continu et la diffusion  $(x_t)_{t \geq 0}$ . On obtient tout d'abord un contrôle de  $(E_t)_{t \geq 0}$  en utilisant le fait que le schéma d'Euler vérifie une condition de stabilité forte sous une hypothèse de croissance sur  $b$  et  $\sigma$  (résultats obtenus dans le chapitre précédent).

**Proposition IV.2:**

Si la diffusion est  $(V, \alpha)$ -confluente avec  $V$  sous-quadratique et si les coefficients  $b$  et  $\sigma$  vérifient

$$|b(x) - b(y)|^2 \leq C_b V(x - y) \quad \text{et} \quad \text{Tr}(\sigma_{x,y}^* \sigma_{x,y}) \leq C_\sigma V(x - y), \quad (2.1)$$

alors pour tout  $p \in [1, 1 + \frac{2\alpha}{C_b C_\sigma}[$ , on a  $\sup_{t \geq 0} \mathbf{E}[V^p(X_{\underline{t}})] < +\infty$  et

$$\forall n \geq 0, \quad \mathbf{E} \left[ \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} E_t^p \right] = \mathcal{O}(\gamma_{n+1}^p) \quad \blacklozenge \quad (2.2)$$

Preuve.

- Tout d'abord, on montre que pour tout  $\tilde{C}_b > C_b$  et tout  $\tilde{C}_\sigma > C_\sigma$ , il existe  $K \geq 1$  tels que pour tout  $x \in \mathbf{R}^d$ .

$$|b(x)|^2 \leq \tilde{C}_b (V(x) + K) \quad \text{et} \quad \text{Tr}(\sigma \sigma^*)(x) \leq \tilde{C}_\sigma (V(x) + K). \quad (2.3)$$

En appliquant (2.1) avec  $x$  et  $y = 0$  on obtient

$$|b(x) - b(0)|^2 = |b(x)|^2 + |b(0)|^2 + 2\langle b(x), b(0) \rangle \leq C_b V(x)$$

donc pour tout  $\varepsilon > 0$ ,  $|b(x)|^2 \leq C_b V(x) - |b(0)|^2 + 2\varepsilon |b(x)|^2 + \frac{2}{\varepsilon} |b(0)|^2$ . Ainsi, pour tout  $\varepsilon > 0$  on a

$$(1 - 2\varepsilon)|b(x)|^2 \leq C_b \left( V(x) + \frac{2/\varepsilon - 1}{C_b} |b(0)|^2 \right),$$

c'est à dire  $|b(x)|^2 \leq \tilde{C}_b (V(x) + K)$  avec  $\tilde{C}_b = \frac{C_b}{1-2\varepsilon}$  et  $K = 1 + \frac{2/\varepsilon - 1}{(1-2\varepsilon)C_b} |b(0)|^2$ . La majoration de  $\text{Tr}(\sigma\sigma^*)$  s'obtient exactement de la même façon.

- D'autre part, comme  $V$  est sous-quadratique, on sait que la diffusion vérifie une condition de stabilité forte de type  $(V + K, \text{Id}, \bar{\alpha})$  pour tout  $\bar{\alpha} < \alpha$  (proposition IV.1). D'après la proposition III.1 et la remarque qui suit, on sait que le schéma d'Euler  $(X_{\Gamma_n})_{n \geq 0}$  vérifie alors une condition de stabilité forte de type  $(V^p, \text{Id})$  pour tout  $p \in ]1, 1 + \frac{2\bar{\alpha}}{C_V C_\sigma}[$  avec  $\bar{\alpha} < \alpha$  et  $\tilde{C}_\sigma > C_\sigma$ . De plus, on prouve que  $(X_{\Gamma_n})_{n \geq 0}$  a des moments polynômiaux de tout ordre en utilisant les techniques de [Fau92] développées pour le schéma d'Euler à pas constant (lemme du contrôle de  $|X_t|^{2p}$  en utilisant Itô, suivi du lemme de Gronwall). Ainsi, comme  $V$  est sous-quadratique, on a  $\mathbf{E}[V^p(X_{\Gamma_{n_0}})] < +\infty$ , et d'après le théorème III.1 on obtient  $\sup_{n \geq 0} \mathbf{E}[V^p(X_{\Gamma_n})] < +\infty$ .
- Il reste à contrôler l'erreur  $(E_t)_{t \geq 0}$ . On a  $V(x) \leq C|x|^2$  donc

$$\sup_{t \in [\Gamma_n, \Gamma_{n+1}[} E_t^p \leq C \left( \gamma_{n+1} |b(X_{\Gamma_n})| + \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} \left| \int_{\Gamma_n}^t \sigma(X_{\underline{s}}) dB_s \right| \right)^{2p},$$

donc d'après l'inégalité de Burkholder-Davis-Gundy (cf. [RY99]), il existe  $C_1 > 0$  et  $C_2 > 0$  telles que pour tout  $n \geq 0$

$$\begin{aligned} \mathbf{E} \left[ \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} E_t^p \right] &\leq C_1 \gamma_{n+1}^{2p} \mathbf{E} \left[ |b(X_{\Gamma_n})|^{2p} \right] + C_2 \mathbf{E} \left[ \left( \int_{\Gamma_n}^t |\sigma(X_{\underline{s}})|^2 ds \right)^p \right], \\ &\leq C_1 \gamma_{n+1}^{2p} \mathbf{E} \left[ |b(X_{\Gamma_n})|^{2p} \right] + C_2 \gamma_{n+1}^p \mathbf{E} \left[ |\sigma(X_{\Gamma_n})|^{2p} \right]. \end{aligned}$$

On conclut en utilisant (2.3) et  $\sup_{n \geq 0} \mathbf{E}[V^p(X_{\Gamma_n})] < +\infty$ . #

Le contrôle de l'erreur entre le schéma continu et le schéma d'Euler discret utilise uniquement la condition de stabilité forte. En revanche, l'étude de l'erreur  $(\mathcal{E}_t)_{t \geq 0}$  entre la diffusion et le schéma continu nécessite l'hypothèse de  $(V, \alpha)$ -confluence.

**Théorème IV.1:**

Si  $(x_t)_{t \geq 0}$  est  $(V, \alpha)$ -confluente avec  $V$  sous-quadratique et  $b$  et  $\sigma$  tels que

$$|b(x) - b(y)|^2 \leq C_b V(x - y) \quad \text{et} \quad \text{Tr}(\sigma_{x,y}^* \sigma_{x,y}) \leq C_\sigma V(x - y), \quad (2.4)$$

alors pour tout  $p \in [1, 1 + \frac{2\alpha}{C_V C_\sigma}]$ ,  $(\mathcal{E}_t)_{t \geq 0}$  converge vers 0 dans  $\mathbf{L}^p$ . Plus précisément,

$$\forall n \geq 0, \quad \mathbf{E} \left[ \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} \mathcal{E}_t^p \right] = \mathcal{O}(\gamma_{n+1}^p). \quad (2.5) \quad \star$$

On commence par prouver le lemme suivant.

**Lemme IV.1:**

Sous les hypothèses du théorème IV.1, pour tout  $p \in [1, 1 + \frac{2\alpha}{C_V C_\sigma}]$  et tout  $\tilde{\alpha} \in ]0, \alpha - \frac{p-1}{2} C_V C_\sigma[$  il existe  $C > 0$  tel que

$$\forall s \leq t, \quad \mathcal{E}_t^p \leq \mathcal{E}_s^p - \tilde{\alpha} \int_s^t \mathcal{E}_u^p du + C \int_s^t E_u^p du + \int_s^t H_u dB_u, \quad (2.6)$$

avec  $H_u = p \mathcal{E}_u^{p-1} (\nabla V(x_u - X_u))^* (\sigma(x_u) - \sigma(X_u))$ .  $\blacklozenge$

Preuve. On rappelle que  $\mathcal{E}_t = V(x_t - X_t)$ . Par la formule d'Itô, on a

$$\begin{aligned} d\mathcal{E}_t^p &= p \mathcal{E}_t^{p-1} \langle \nabla V(x_t - X_t), b(x_t) - b(X_t) \rangle dt + \frac{p}{2} \mathcal{E}_t^{p-1} \text{Tr} \left( \sigma_{x_t, X_t}^* D^2 V(x_t - X_t) \sigma_{x_t, X_t} \right) dt \\ &\quad + \frac{p(p-1)}{2} \mathcal{E}_t^{p-2} \text{Tr} \left( \sigma_{x_t, X_t}^* (\nabla V^{\otimes 2}(x_t - X_t)) \sigma_{x_t, X_t} \right) dt \\ &\quad + p \mathcal{E}_t^{p-1} (\nabla V(x_t - X_t))^* (\sigma(x_t) - \sigma(X_t)) dB_t. \end{aligned} \quad (2.7)$$

- Soit  $S$  une matrice symétrique. Alors pour tout  $(x, y, z) \in \mathbf{R}^{3d}$  on a  $\sigma_{x,z} = \sigma_{x,y} + \sigma_{y,z}$  donc

$$\text{Tr}(\sigma_{x,z}^* S \sigma_{x,z}) = \text{Tr}(\sigma_{x,y}^* S \sigma_{x,y}) + \text{Tr}(\sigma_{y,z}^* S \sigma_{y,z}) + 2 \text{Tr}(\sigma_{y,z}^* S \sigma_{x,y}).$$

D'une part, comme  $(A \mapsto \sqrt{\text{Tr}(A^* A)})$  est une norme sur l'espace des matrices, il existe  $C > 0$  tel que

$$\begin{aligned} \text{Tr}(\sigma_{y,z}^* S \sigma_{x,y}) &\leq C \|\sigma_{y,z}^* S \sigma_{x,y}\|, \\ &\leq C \|\sigma_{y,z}^*\| \|S\| \|\sigma_{x,y}\|, \end{aligned}$$

où  $\|\cdot\|$  est la norme d'application linéaire. La lettre  $C$  désigne une constante changeant d'une ligne à l'autre. Par équivalence des normes,

$$\text{Tr}(\sigma_{y,z}^* S \sigma_{x,y}) \leq C \sqrt{\text{Tr}(\sigma_{y,z}^* \sigma_{y,z})} \sqrt{\text{Tr}(\sigma_{x,y}^* \sigma_{x,y})} \sqrt{\text{Tr}(S^* S)}.$$

En utilisant l'inégalité  $2ab \leq \frac{a^2}{\varepsilon} + b^2 \varepsilon$  vraie pour tout  $(a, b) \in \mathbf{R}^2$  et  $\varepsilon > 0$ , on obtient

$$2 \text{Tr}(\sigma_{y,z}^* S \sigma_{x,y}) \leq C \left( \varepsilon \text{Tr}(\sigma_{x,y}^* \sigma_{x,y}) + \frac{1}{\varepsilon} \text{Tr}(\sigma_{y,z}^* \sigma_{y,z}) \right) \sqrt{\text{Tr}(S^* S)}.$$

D'autre part, on a  $\text{Tr}(\sigma_{y,z}^* S \sigma_{y,z}) \leq C \text{Tr}(\sigma_{y,z}^* \sigma_{y,z}) \sqrt{\text{Tr}(S^* S)}$ , et en combinant ces différentes

majorations, il existe  $C > 0$  et  $C_\varepsilon > 0$  (changeant d'une ligne à l'autre) telles que

$$\mathrm{Tr}(\sigma_{x,z}^* S \sigma_{x,z}) \leq \mathrm{Tr}(\sigma_{x,y}^* S \sigma_{x,y}) + \varepsilon C \mathrm{Tr}(\sigma_{x,y}^* \sigma_{x,y}) \sqrt{\mathrm{Tr}(S^* S)} + C_\varepsilon \mathrm{Tr}(\sigma_{y,z}^* \sigma_{y,z}) \sqrt{\mathrm{Tr}(S^* S)}.$$

Comme pour tout  $(x, y) \in \mathbf{R}^2$ , on a  $\mathrm{Tr}(\sigma_{x,y}^* \sigma_{x,y}) \leq C_\sigma V(x - y)$  alors

$$\mathrm{Tr}(\sigma_{x,z}^* S \sigma_{x,z}) \leq \mathrm{Tr}(\sigma_{x,y}^* S \sigma_{x,y}) + \varepsilon C V(x - y) \sqrt{\mathrm{Tr}(S^* S)} + C_\varepsilon V(y - z) \sqrt{\mathrm{Tr}(S^* S)}. \quad (2.8)$$

– En prenant  $S = D^2V(x_t - X_t)$  dans (2.8), on a  $\sqrt{\mathrm{Tr}(S^* S)} \leq C\rho$  (où  $\rho = \|D^2V\|_\infty$ ) et

$$\begin{aligned} \mathrm{Tr}(\sigma_{x_t, X_t}^* D^2V(x_t - X_t) \sigma_{x_t, X_t}) &\leq \mathrm{Tr}(\sigma_{x_t, X_t}^* D^2V(x_t - X_t) \sigma_{x_t, X_t}) \\ &\quad + \varepsilon C V(x_t - X_t) + C_\varepsilon V(X_t - X_t). \end{aligned}$$

– Si  $S = (\nabla V)^{\otimes 2}(x_t - X_t)$ , alors  $\sqrt{\mathrm{Tr}(S^* S)} \leq |\nabla V|^2(x_t - X_t)$  donc  $\sqrt{\mathrm{Tr}(S^* S)} \leq C_V V(x_t - X_t)$  et d'après (2.8) on obtient

$$\begin{aligned} \mathrm{Tr}(\sigma_{x_t, X_t}^* (\nabla V)^{\otimes 2}(x_t - X_t) \sigma_{x_t, X_t}) &\leq (1 + C\varepsilon) C_V C_\sigma V^2(x_t - X_t) \\ &\quad + C_\varepsilon V(x_t - X_t) V(X_t - X_t). \end{aligned}$$

On combine ces inégalités avec (2.7) et on rappelle que  $\mathcal{E}_t = V(x_t - X_t)$  et  $E_t = V(X_t - X_t)$ . Alors il existe  $C > 0$  et  $C_\varepsilon > 0$  telles que pour tout  $s \leq t$ ,

$$\begin{aligned} \mathcal{E}_t^p &\leq \mathcal{E}_s^p + \int_s^t p \mathcal{E}_u^{p-1} \left( \langle \nabla V(x_u - X_u), b_{x_u, X_u} \rangle + \frac{1}{2} \mathrm{Tr}(\sigma_{x_u, X_u}^* D^2V(x_u - X_u) \sigma_{x_u, X_u}) \right) du \\ &\quad + \int_s^t \frac{p}{2} ((p-1) C_V C_\sigma + C_\varepsilon) \mathcal{E}_u^p du + \int_s^t C_\varepsilon \mathcal{E}_u^{p-1} E_u du + \int_s^t H_u dB_u, \quad (2.9) \end{aligned}$$

avec  $H_u = p \mathcal{E}_u^{p-1} (\nabla V(x_u - X_u))^* (\sigma(x_u) - \sigma(X_u))$ .

- D'autre part,  $b$  vérifie  $|b(x) - b(y)| \leq C_b V(x - y)$  donc pour tout  $\varepsilon > 0$  on a

$$\langle \nabla V(x_t - X_t), b(X_t) - b(X_t) \rangle \leq \frac{\varepsilon}{2} |\nabla V(x_t - X_t)|^2 + \frac{C_b}{2\varepsilon} V(X_t - X_t),$$

et  $|\nabla V|^2 \leq C_V V$  donc  $|\nabla V(x_t - X_t)|^2 \leq C_V V(x_t - X_t)$  et

$$\langle \nabla V(x_t - X_t), b(x_t) - b(X_t) \rangle \leq \langle \nabla V(x_t - X_t), b(x_t) - b(X_t) \rangle + \frac{\varepsilon}{2} C_V \mathcal{E}_t + \frac{C_b}{2\varepsilon} E_t.$$

D'après (2.9), on en déduit que pour tout  $\varepsilon > 0$  il existe  $C > 0$  et  $C_\varepsilon > 0$  telles que pour tout  $s \leq t$

$$\begin{aligned} \mathcal{E}_t^p &\leq \mathcal{E}_s^p + \int_s^t p \mathcal{E}_u^{p-1} \left( \langle \nabla V(x_u - X_u), b_{x_u, X_u} \rangle + \frac{1}{2} \mathrm{Tr}(\sigma_{x_u, X_u}^* D^2V(x_u - X_u) \sigma_{x_u, X_u}) \right) du \\ &\quad + \int_s^t \frac{p}{2} ((p-1) C_V C_\sigma + C_\varepsilon) \mathcal{E}_u^p du + \int_s^t C_\varepsilon \mathcal{E}_u^{p-1} E_u du + \int_s^t H_u dB_u, \quad (2.10) \end{aligned}$$

et comme la diffusion est  $(V, \alpha)$ -confluente, on a pour tout  $s \leq t$ ,

$$\mathcal{E}_t^p \leq \mathcal{E}_s^p + \int_s^t \left( -\alpha p + \frac{p}{2}((p-1)C_V C_\sigma + C\varepsilon) \right) \mathcal{E}_u^p du + \int_s^t C_\varepsilon \mathcal{E}_u^{p-1} E_u du + \int_s^t H_u dB_u,$$

En posant  $\bar{\alpha} = \alpha - \frac{p-1}{2}C_V C_\sigma > 0$  (car  $p < 1 + \frac{2\alpha}{C_V C_\sigma}$ ), on obtient

$$\forall s < t, \quad \mathcal{E}_t^p \leq \mathcal{E}_s^p + \int_s^t (-\bar{\alpha}p + C\varepsilon) \mathcal{E}_u^p du + \int_s^t C_\varepsilon \mathcal{E}_u^{p-1} E_u du + \int_s^t H_u dB_u.$$

De plus, si  $p > 1$  l'inégalité de Young donne

$$\forall \tilde{\varepsilon} > 0, \quad \mathcal{E}_u^{p-1} E_u \leq \frac{p-1}{p} \tilde{\varepsilon}^{\frac{p}{p-1}} \mathcal{E}_u^p + \frac{1}{p\tilde{\varepsilon}^p} E_u^p. \quad (2.11)$$

Ainsi pour tout  $p \geq 1$ , en prenant  $\varepsilon$  et  $\tilde{\varepsilon}$  suffisamment petits, il existe  $\tilde{\alpha} > 0$  et  $C > 0$  tels que

$$\forall s \leq t, \quad \mathcal{E}_t^p \leq \mathcal{E}_s^p - \tilde{\alpha} \int_s^t \mathcal{E}_u^p du + \int_s^t C E_u^p du + \int_s^t H_u dB_u,$$

d'où le lemme. #

Preuve (du théorème IV.1).

- Soit  $M > 0$ . Pour le temps d'arrêt  $\tau_M = \inf \{t \geq 0; |x_t| > M\}$ , le processus  $\left( \int_0^{t \wedge \tau_M} H_s dB_s \right)_{t \geq 0}$  est une martingale et d'après le lemme IV.1 on a pour tout  $\tilde{\alpha} \in ]0, \alpha[$  et  $t \geq 0$

$$\begin{aligned} \forall s \leq t, \quad \mathbf{1}_{\{s < \tau_M\}} \left( \mathcal{E}_{t \wedge \tau_M}^p + \int_0^{t \wedge \tau_M} \tilde{\alpha} \mathcal{E}_u^p du \right) \\ \leq \mathbf{1}_{\{s < \tau_M\}} \left( \mathcal{E}_s^p + \int_0^s \tilde{\alpha} \mathcal{E}_u^p du + \int_s^{t \wedge \tau_M} C E_u^p du + \int_s^{t \wedge \tau_M} H_u dB_u \right). \end{aligned}$$

Le lemme de Fatou conditionnel implique alors que pour tout  $s \leq t$ ,

$$\mathbf{E} \left[ \mathcal{E}_t^p + \int_0^t \tilde{\alpha} \mathcal{E}_u^p du \mid \mathcal{F}_s \right] \leq \mathcal{E}_s^p + \int_0^s \tilde{\alpha} \mathcal{E}_u^p du + \lim_{M \rightarrow +\infty} \mathbf{E} \left[ \mathbf{1}_{\{s < \tau_M\}} \int_s^{t \wedge \tau_M} C E_u^p du \mid \mathcal{F}_s \right].$$

En utilisant  $\mathbf{1}_{\{s < \tau_M\}} = 1 - \mathbf{1}_{\{s \geq \tau_M\}}$  et  $\lim_{M \rightarrow +\infty} \mathbf{1}_{\{s \geq \tau_M\}} = 0$  p.s., et  $\mathbf{1}_{\{t < \tau_M\}} \leq 1$  on en déduit que

$$\mathbf{E} [\mathcal{E}_t^p \mid \mathcal{F}_s] \leq \mathcal{E}_s^p + \mathbf{E} \left[ \int_s^t (-\tilde{\alpha} \mathcal{E}_u^p + C E_u^p) du \mid \mathcal{F}_s \right].$$

Soit  $g(t) = \mathbf{E} [\mathcal{E}_t^p]$  et  $\psi(t) = \mathbf{E} [E_t^p]$ , on a alors

$$\forall s \leq t, \quad g(t) \leq g(s) - \tilde{\alpha} \int_s^t g(u) du + C \int_s^t \psi(u) du.$$

On en déduit que pour tout  $s \leq t$ ,  $\exp(\tilde{\alpha}t)g(t) - \exp(\tilde{\alpha}s)g(s) \leq C \int_s^t \exp(\tilde{\alpha}u)\psi(u)du$ , et donc

pour tout  $n \geq 0$

$$\begin{aligned} g(\Gamma_{n+1}) &\leq \exp(-\tilde{\alpha}\gamma_{n+1})g(\Gamma_n) + C \exp(-\tilde{\alpha}\Gamma_{n+1}) \int_{\Gamma_n}^{\Gamma_{n+1}} \exp(\alpha u) \psi(u) du, \\ &\leq \exp(-\tilde{\alpha}\gamma_{n+1})g(\Gamma_n) + (1 - \exp(-\tilde{\alpha}\gamma_{n+1})) (C/\tilde{\alpha}) \sup_{u \in [\Gamma_n, \Gamma_{n+1}[} \psi(u). \end{aligned}$$

Comme  $(1 - \exp(-\tilde{\alpha}\gamma_n))_{n \geq 1}$  tend vers 0, on en déduit qu'il existe  $C > 0$  telle que

$$\forall n \geq 0, \quad \mathbf{E} \left[ \mathcal{E}_{\Gamma_{n+1}}^p \right] \leq C \sup_{u \in [\Gamma_n, \Gamma_{n+1}[} \mathbf{E} [E_u^p]. \quad (2.12)$$

• D'autre part, on a

$$\sup_{t \in [\Gamma_n, \Gamma_{n+1}[} \mathcal{E}_t^p \leq \mathcal{E}_{\Gamma_n}^p + C\gamma_{n+1} \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} E_t^p + \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} \left| \int_{\Gamma_n}^t H_u dW_u \right|, \quad (2.13)$$

avec  $H_u = 2p\mathcal{E}_u^{p-1} (x_u - X_u)^* (\sigma(x_u) - \sigma(X_u))$ . Par l'inégalité de Burkholder-Davis-Gundy, on a

$$\mathbf{E} \left[ \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} \left| \int_{\Gamma_n}^t H_u dW_u \right| \right] \leq C \mathbf{E} \left[ \sqrt{\int_{\Gamma_n}^{\Gamma_{n+1}} |H_u|^2 du} \right]. \quad (2.14)$$

Or  $|H_t|^2 \leq 8p^2 \mathcal{E}_t^{2p-1} (|\sigma(x_t) - \sigma(X_t)|^2 + |\sigma(X_t) - \sigma(X_{\underline{t}})|^2) \leq 8p^2 C_\sigma (\mathcal{E}_t^{2p} + \mathcal{E}_t^{2p-1} E_t)$ , et par l'inégalité de Young il existe  $C_1 > 0$  et  $C_2 > 0$  tels que

$$|H_t|^2 \leq C_1 \mathcal{E}_t^{2p} + C_2 E_t^{2p}.$$

En injectant cette majoration dans (2.14) on obtient

$$\mathbf{E} \left[ \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} \left| \int_{\Gamma_n}^t H_u dW_u \right| \right] \leq C \sqrt{\gamma_{n+1}} \mathbf{E} \left[ \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} \sqrt{C_1 \mathcal{E}_t^{2p} + C_2 E_t^{2p}} \right].$$

Donc d'après (2.13), il existe  $C_1 > 0$  et  $C_2 > 0$  tels que

$$\begin{aligned} \mathbf{E} \left[ \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} \mathcal{E}_t^p \right] &\leq \mathbf{E} [\mathcal{E}_{\Gamma_n}^p] + C\gamma_{n+1} \mathbf{E} \left[ \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} E_t^p \right] \\ &\quad + C_1 \sqrt{\gamma_{n+1}} \mathbf{E} \left[ \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} E_t^p \right] + C_2 \sqrt{\gamma_{n+1}} \mathbf{E} \left[ \sup_{t \in [\Gamma_n, \Gamma_{n+1}[} \mathcal{E}_t^p \right], \end{aligned}$$

On conclut via la proposition IV.2. #

### 3. Convergence *p.s.* et vitesses

Le théorème IV.1 combiné au lemme de Borel-Cantelli permet d'obtenir la convergence *p.s.* pour des schémas construits à partir de certaines suites de pas. En effet, si la suite de pas

$(\gamma_n)_{n \geq 1}$  à vérifie la condition

$$\sum_{n \geq 1} \gamma_n^p < +\infty \quad \text{avec } p \in \left] 1, 1 + \frac{2\alpha}{C_V C_\sigma} \right[ , \quad (3.1)$$

alors  $(\mathcal{E}_t)_{t \geq 0}$  converge *p.s.* vers 0.

La condition (3.1) imposée sur les pas est très restrictive, et il est difficile d'obtenir de cette façon des vitesses pour la convergence presque sûre. Le théorème suivant, qui repose sur le lemme IV.1, affaiblit la condition sur les pas et donne dans le même temps des vitesses de convergence.

**Théorème IV.2:**

Supposons que  $(x_t)_{t \geq 0}$  est  $(V, \alpha)$ -confluente avec  $V$  sous-quadratique et que  $b$  et  $\sigma$  vérifient

$$|b(x) - b(y)|^2 \leq C_b V(x - y) \quad \text{et} \quad \text{Tr}(\sigma_{x,y}^* \sigma_{x,y}) \leq C_\sigma V(x - y). \quad (3.2)$$

Si la suite de pas  $(\gamma_n)_{n \geq 1}$  sur laquelle est construit le schéma  $(X_t)_{t \geq 0}$  satisfait

$$\sum_{n \geq 1} \gamma_n^{p+1} v(\Gamma_n) < +\infty, \quad \text{avec } p \in \left[ 1, 1 + \frac{2\alpha}{C_V C_\sigma} \right[ \quad (3.3)$$

où  $v$  est une fonction  $\mathcal{C}^1(\mathbf{R}_+, \mathbf{R}_+)$  strictement croissante vérifiant

$$\exists \tilde{\alpha} \in ]0, \alpha[, \exists t_0 \geq 0, \forall t \geq t_0, \quad v'(t) \leq \tilde{\alpha} v(t), \quad (3.4)$$

alors  $\lim_{t \rightarrow +\infty} v(t) \mathcal{E}_t^p = 0$  *p.s.* ★

Preuve. On rappelle que d'après le lemme IV.1, il existe  $K > 0$  tel que

$$\forall s \leq t, \quad \mathcal{E}_t^p \leq \mathcal{E}_s^p - \tilde{\alpha} \int_s^t \mathcal{E}_u^p du + K \int_s^t E_u^p du + \int_s^t H_u dB_u, \quad (3.5)$$

avec  $H_u = 2p \mathcal{E}_u^{p-1} (x_u - X_u)^* (\sigma(x_u) - \sigma(X_u))$ .

- Montrons tout d'abord que  $\int_0^{+\infty} v(u) \mathbf{E}[E_u^p] du < +\infty$ . Pour tout  $t \in [\Gamma_n, \Gamma_{n+1}[$ , on a

$$\int_0^t v(u) \mathbf{E}[E_u^p] du \leq \sum_{k=0}^n \int_{\Gamma_k}^{\Gamma_{k+1}} v(u) \mathbf{E}[E_u^p] du,$$

donc, par la proposition IV.2, il existe  $C > 0$  tel que pour tout  $n \geq 0$  et tout  $t \in [\Gamma_n, \Gamma_{n+1}[$

$$\int_0^t v(u) \mathbf{E}[E_u^p] du \leq C \sum_{k=0}^n \gamma_{k+1}^p \int_{\Gamma_k}^{\Gamma_{k+1}} v(u) du.$$

On conclut en utilisant la croissance de la fonction  $v$  et la condition (3.3).

- Soit  $(S_t)_{t \geq 0}$  le processus défini pour tout  $t \geq 0$  par  $S_t = v(t) \mathcal{E}_t^p + K \mathbf{E} \left[ \int_t^{+\infty} v(u) E_u^p du \mid \mathcal{F}_t \right]$ . C'est un processus positif (car  $v$  est positive), et on va montrer que  $(S_t)_{t \geq 0}$  est une surmar-

tingale locale positive. Pour tout  $s \leq t$ , on a

$$\mathbf{E}[S_t | \mathcal{F}_s] = S_s + \mathbf{E}[v(t)\mathcal{E}_t^p - v(s)\mathcal{E}_s^p | \mathcal{F}_s] + K\mathbf{E}\left[\int_s^t v(u)E_u^p du \middle| \mathcal{F}_s\right],$$

et  $v(t)\mathcal{E}_t^p - v(s)\mathcal{E}_s^p = \int_s^t v'(u)\mathcal{E}_u^p du + \int_s^t v(u)d\mathcal{E}_u^p$  donc d'après (3.5), on a

$$\mathbf{E}[S_t | \mathcal{F}_s] \leq S_s + \mathbf{E}\left[\int_s^t (v'(u) - \tilde{\alpha}v(u)) \mathcal{E}_u^p du \middle| \mathcal{F}_s\right] + \mathbf{E}\left[\int_s^t v(u)H_u dB_u \middle| \mathcal{F}_s\right].$$

La fonction  $v$  vérifie (3.4) pour  $t_0 \geq 0$ , donc on en déduit que le processus  $(S_t)_{t \geq t_0}$  est une surmartingale locale positive. On a donc convergence presque sûre de  $(S_t)_{t \geq t_0}$  et donc convergence presque sûre de  $(v(t)\mathcal{E}_t^p)_{t \geq 0}$ . #

### Application à des pas "puissance"

On suppose maintenant que la suite de pas  $(\gamma_n)_{n \geq 1}$  est définie pour tout  $n \geq 1$  par

$$\gamma_n = \gamma_0 n^{-r}, \quad \text{avec } r \in ]0, 1] \quad \text{et } \gamma_0 > 0. \quad (3.6)$$

La condition (3.3) se traduit alors par une condition sur  $r$ , et la fonction  $v$  dépend de  $r$ . Pour une suite de pas du type  $\gamma_n = \gamma_0/n$ , on obtient une convergence exponentielle de  $\mathcal{E}_t$  vers 0 presque sûrement.

#### Corollaire IV.1:

Sous les hypothèses du théorème IV.2, le schéma  $(X_t)_{t \geq 0}$  converge presque sûrement vers la diffusion  $(x_t)_{t \geq 0}$  dès que  $r > \frac{1}{2 + \frac{2\alpha}{C_V C_\sigma}}$ . En outre,

$$v_r(t)V(x_t - X_t) \xrightarrow{p.s.} 0, \quad (3.7)$$

où

$$v_r(t) = \begin{cases} \exp(\tilde{\alpha}t) & \text{avec } \tilde{\alpha} < \alpha \wedge \frac{1}{\gamma_0} & \text{si } r = 1 \\ t^q & \text{avec } 0 < q < \frac{r}{1-r} - \frac{1}{1 + \frac{2\alpha}{C_V C_\sigma}} & \text{si } r \in ]0, 1[ \end{cases} \frac{1}{2 + \frac{2\alpha}{C_V C_\sigma}}, 1[. \quad (3.8) \quad \bullet$$

Preuve. Il s'agit de vérifier la condition (3.3) du théorème précédent pour une bonne fonction vitesse  $v$ .

- Si  $r = 1$ , alors  $\gamma_n = \gamma_0/n$  pour tout  $n \geq 1$ , et  $\Gamma_n \sim \gamma_0 \ln(n)$ . En posant  $v(t) = \exp(\tilde{\alpha}t)$  avec  $\tilde{\alpha} < \alpha \wedge 1$ , on a alors (pour  $p = 1$ )

$$\gamma_n^2 v_r(\Gamma_n) \sim \frac{\gamma_0^{2+q}}{n^2} \exp(\tilde{\alpha}\gamma_0 \ln(n)) = \frac{\gamma_0^2}{n^{2-\tilde{\alpha}\gamma_0}},$$

donc  $\sum_{n \geq 1} \gamma_n^2 v(\Gamma_n)$  converge car  $2 - \tilde{\alpha}\gamma_0 > 1$ .

- Si  $r \in \left] \frac{1}{2 + \frac{2\alpha}{C_V C_\sigma}}, 1 \right[$ , alors pour tout  $\tilde{q} < \left( \frac{r}{1-r} \right) \left( 1 + \frac{2\alpha}{C_V C_\sigma} \right) - 1$  et pour tout  $p \in \left] (1+\tilde{q})\frac{1-r}{r}, 1 + \frac{2\alpha}{C_V C_\sigma} \right[$  (intervalle non vide d'après les bornes sur  $r$  et  $\tilde{q}$ ), on pose  $\tilde{v}_r(t) = t^{\tilde{q}}$  et on obtient

$$\gamma_n^{p+1} \tilde{v}_r(\Gamma_n) \sim \left( \frac{\gamma_0^{p+1+\tilde{q}}}{(1-r)^{\tilde{q}}} \right) \frac{1}{n^{r(p+1)+(r-1)\tilde{q}}}.$$

Comme  $r(p+1) + (r-1)\tilde{q} > 1$ , la série  $\sum_{n \geq 1} \gamma_n^{p+1} \tilde{v}_r(\Gamma_n)$  converge. De plus, on vérifie que  $p \in \left[ 1, 1 + \frac{2\alpha}{C_V C_\sigma} \right[$  et que  $\tilde{q} > 0$ , donc d'après le théorème IV.2 on a

$$t^{\tilde{q}} \mathcal{E}_t^p \xrightarrow{p.s.} 0.$$

En posant  $q = \tilde{q}/p$ , on prouve le corollaire. #

Remarque: La vitesse de convergence est d'autant plus grande que  $r$  est grand, c'est à dire que la décroissance des pas  $(\gamma_n)_{n \geq 1}$  est rapide. De plus, on remarque que si  $C_\sigma$  est aussi petit que l'on veut (i.e.  $\text{Tr}(\sigma_{x,y} \sigma_{x,y}^*)$  négligeable devant  $V(x-y)$ ) alors on a convergence pour tout  $r \in ]0, 1]$ . \*

#### 4. Convergence des mesures empiriques

On va maintenant utiliser les résultats de convergence trajectorielle obtenus précédemment pour montrer la convergence des mesures empiriques du schéma d'Euler. On les définit de la façon suivante :

$$\nu_n(dx) = \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \delta_{X_{\Gamma_{k-1}}}(dx).$$

Le but est d'étendre les résultats du chapitre précédent lorsque la diffusion est  $(V, \alpha)$ -confluente, plus précisément d'obtenir la convergence de  $\nu_n^\eta(f)$  vers  $\nu(f)$  pour des fonctions  $f$  aussi grande que possible ( $\nu$  étant l'unique mesure invariante).

##### **Théorème IV.3:**

Supposons que  $(x_t)_{t \geq 0}$  est  $(V, \alpha)$ -confluente où  $V(x) = \langle x, Sx \rangle$  et  $S$  est une matrice définie positive, et que les coefficients  $b$  et  $\sigma$  vérifient :  $\exists C_b > 0$ ,  $\exists C_\sigma > 0$  et  $\exists \varsigma \in ]0, 1]$  tels que

$$|b(x) - b(y)|^2 \leq C_b V(x-y) \quad \text{et} \quad \text{Tr}(\sigma_{x,y}^* \sigma_{x,y}) \leq C_\sigma V^{1-\varsigma}(x-y).$$

Si la suite de pas  $(\gamma_n)_{n \geq 1}$  sur laquelle est construit le schéma  $(X_{\Gamma_n})_{n \geq 0}$  est définie par  $\gamma_n = \gamma_0 n^{-r}$  avec  $\gamma_0 > 0$  et  $r \in ]0, 1]$ , alors pour tout  $\lambda < \frac{\alpha \wedge 1}{2\lambda_{S^c} C_\sigma}$  et toute fonction  $f$  mesurable dominée par  $\exp(\lambda V^\varsigma)$  on a

$$\frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k f(X_{\Gamma_{k-1}}) \xrightarrow{p.s.} \int_{\mathbf{R}^d} f(x) \nu(dx). \quad \star$$

Preuve. Tout d'abord, il est clair que  $V$  vérifie (1.3) avec  $C_V = 4\lambda_S$  et  $\rho = 2\lambda_S$  avec  $\lambda_S = \max_i \{\lambda_i\}$  où  $\lambda_i$  est la  $i$ -ème valeur propre de  $S$ . D'autre part, on prouve de la même façon que dans la proposition IV.2 qu'il existe  $\tilde{C}_b > C_b$ ,  $\tilde{C}_\sigma > C_\sigma$  et  $K \geq 1$  tels que

$$|b(x)|^2 \leq \tilde{C}_b (V(x) + K) \quad \text{et} \quad \text{Tr}(\sigma\sigma^*)(x) \leq \tilde{C}_\sigma (V(x) + K)^{1-\varsigma}, \quad (4.1)$$

donc d'après la proposition II.1, la diffusion  $(x_t)_{t \geq 0}$  vérifie une condition de stabilité forte de type  $(\exp(\lambda(V + K)^\varsigma), \text{Id})$  avec  $K \geq 1$  et  $\lambda < \frac{\alpha}{\varsigma \tilde{C}_\sigma 2\lambda_S}$ . Ainsi on a

$$\sup_{t \geq 0} \frac{V(x_t)}{\ln(t)^{1/\varsigma}} < +\infty \quad p.s. \quad (4.2)$$

- Montrons maintenant que  $\lim_{t \rightarrow +\infty} V(x_t) - V(X_t) = 0$  *p.s.* On a

$$V(x_t) - V(X_t) = \langle \nabla V(x_t), x_t - X_t \rangle + \frac{1}{2} S(x_t - X_t)^{\otimes 2},$$

donc il existe  $C_1 > 0$  et  $C_2 > 0$  telles que

$$\begin{aligned} |V(x_t) - V(X_t)| &\leq C_1 V(x_t) \sqrt{V}(x_t - X_t) + C_2 V(x_t - X_t), \\ &\leq C_1 \frac{V(x_t)}{\ln(t)^{1/\varsigma}} \ln(t)^{1/\varsigma} \sqrt{V}(x_t - X_t) + C_2 V(x_t - X_t). \end{aligned}$$

D'après (4.2) et d'après les vitesses de convergence obtenues par le corollaire IV.1 lorsque  $\gamma_n = \gamma_0 n^{-r}$ , on en déduit que  $|V(x_t) - V(X_t)| \xrightarrow{p.s.} 0$ .

- Comme  $\varsigma \in ]0, 1]$  on en déduit que  $|V^\varsigma(x_t) - V^\varsigma(X_t)| \xrightarrow{p.s.} 0$ . Donc pour tout  $\varepsilon > 0$  il existe  $t_0 > 0$  tel que  $\forall t \geq t_0$ ,  $|V^\varsigma(x_t) - V^\varsigma(X_t)| \leq \varepsilon$  et pour tout  $\lambda > 0$ ,

$$\begin{aligned} \left| \frac{1}{t} \int_0^t \exp(\lambda V^\varsigma(x_s)) - \exp(\lambda V^\varsigma(X_s)) ds \right| &\leq \frac{1}{t} \left| \int_0^{t_0} \exp(\lambda V^\varsigma(x_s)) - \exp(\lambda V^\varsigma(X_s)) ds \right| \\ &\quad + \frac{\lambda \varepsilon (\exp(\lambda \varepsilon) + 1)}{t} \int_{t_0}^t \exp(\lambda V^\varsigma(x_s)) ds, \end{aligned}$$

car pour tout  $x \in \mathbf{R}$ ,  $|1 - \exp(x)| \leq |x|(\exp(x) + 1)$ . On en conclut que

$$\forall \lambda > 0, \quad \lim_{t \rightarrow +\infty} \left| \frac{1}{t} \int_0^t \exp(\lambda V^\varsigma(x_s)) ds - \frac{1}{t} \int_0^t \exp(\lambda V^\varsigma(X_s)) ds \right| \xrightarrow{p.s.} 0.$$

Par le théorème ergodique (cf. [Has80]), on sait que pour tout  $\lambda > 0$  tel que  $\exp(\lambda V^\varsigma) \in \mathbf{L}^1(\nu)$  on a  $\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \exp(\lambda V^\varsigma)(x_s) ds = \int_{\mathbf{R}^d} \exp(\lambda V^\varsigma(x)) \nu(dx)$ , donc on en déduit que pour toute fonction  $f$  mesurable,  $f = \mathcal{O}(\exp(\lambda V^\varsigma))$ ,

$$\frac{1}{t} \int_0^t f(X_s) ds \xrightarrow{p.s.} \int_{\mathbf{R}^d} f(x) \nu(dx).$$

- On conclut aisément en remarquant que pour tout  $n \geq 0$ ,

$$\frac{1}{\Gamma_n} \int_0^{\Gamma_n} f(X_s) ds = \frac{1}{\Gamma_n} \sum_{k=1}^n \int_{\Gamma_{k-1}}^{\Gamma_k} f(X_s) ds = \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k f(X_{\Gamma_{k-1}}). \quad \#$$

Dans ce chapitre, nous étudions la vitesse de convergence des lois empiriques pondérées  $(\nu_n^n)_{n \geq 0}$  du schéma d'Euler  $(X_n)_{n \geq 0}$  (construit à partir d'une suite de pas  $(\gamma_n)_{n \geq 1}$  convergeant vers 0) vers l'unique probabilité invariante  $\nu$  de  $(x_t)_{t \geq 0}$ . Nous donnons un développement à tout ordre de l'erreur entre  $\nu_n^n(f)$  et  $\nu(f)$  pour une fonction  $f$  admettant une solution à l'équation de Poisson suffisamment régulière. Ce développement de l'erreur permet d'obtenir soit un théorème de la limite centrale (TLC), soit une convergence presque sûre vers un biais. Ces deux modes de convergence apparaissent en fonction du choix de la suite de pas  $(\gamma_n)_{n \geq 1}$ . Pour les pas rapidement décroissants, le TLC obtenu est accompagné par une loi du logarithme itéré (LLI), conséquence de la LLI pour les martingales.

Pour les pas lentement décroissants, la convergence est presque sûre et une extrapolation de Richardson (ou Romberg-Richardson) peut être mise en œuvre. Nous concluons le chapitre en obtenant des vitesses de convergence lorsque l'algorithme est couplé à cette méthode d'extrapolation.

## 1. Introduction

Rappelons tout d'abord le résultat dû à Bhattacharya [Bha82] sur le comportement asymptotique des mesures empiriques d'un processus de Markov stationnaire et ergodique. Il montre le théorème de la limite centrale fonctionnelle suivant.

### **Théorème V.1:**

Soit  $(x_t)_{t \geq 0}$  un processus de Markov, stationnaire et ergodique de mesure invariante  $\nu$  et de générateur infinitésimal  $\mathcal{A}$ . Si  $f = \mathcal{A}\varphi$  avec  $\varphi$  fonction  $\mathcal{C}^2$  à support compact, alors

$$\frac{1}{\sqrt{n}} \int_0^{nt} f(x_s) ds \xrightarrow{(U_K)} s_f W_t,$$

où  $s_f = -2 \int_{\mathbf{R}^d} f \varphi d\nu$  ( $\xrightarrow{(U_K)}$  désigne la convergence en loi fonctionnelle sur  $\mathcal{C}(\mathbf{R}_+, \mathbf{R}^d)$ ). ★

La preuve de ce théorème repose en partie sur le théorème de la limite centrale fonctionnelle pour les martingales de carré intégrable à accroissements stationnaires et ergodiques. L'idée d'utiliser l'équation de Poisson pour ramener le problème à un problème de martingale est due à Gordin et Lifšic dans [GL78]. Il est clair que ce théorème fonctionnel implique le TLC suivant

$$\frac{1}{\sqrt{t}} \int_0^t f(x_s) ds \xrightarrow{\mathcal{L}} \mathcal{N}(0, s_f^2). \quad (1.1)$$

Nous allons établir un TLC similaire pour les mesures empiriques pondérées  $(\nu_n^\eta)_{n \geq 1}$  du schéma d'Euler discrétisant la solution d'une EDS vérifiant une condition de stabilité forte. On suppose que le processus  $(x_t)_{t \geq 0}$  est solution de l'équation différentielle stochastique suivante

$$dx_t = b(x_t)dt + \sigma(x_t)dB_t, \quad x_0 \in \mathbf{R}^d,$$

de générateur infinitésimal  $\mathcal{A}$ . Le processus  $(B_t)_{t \geq 0}$  est un mouvement Brownien sur  $\mathbf{R}^m$  et les fonctions  $b$  et  $\sigma$  sont supposées continues. De plus, on suppose que la diffusion  $(x_t)_{t \geq 0}$  vérifie une condition de stabilité forte de type  $(V, \text{Id}, \alpha)$  avec  $V$  sous-quadratique, et qu'il existe une unique mesure invariante  $\nu$  pour  $(x_t)_{t \geq 0}$ . Les coefficients  $b$  et  $\sigma$  vérifient de plus  $|b|^2 + \text{Tr}(\sigma\sigma^*) = \mathcal{O}(V)$ .

On discrétise la diffusion par un schéma d'Euler de pas décroissant construit à partir d'une suite  $(\gamma_n)_{n \geq 1}$  (convergente vers 0 et terme général d'une série divergente) et d'un bruit blanc  $(U_n)_{n \geq 1}$  normalisé. Ce schéma est noté  $(X_n)_{n \geq 0}$ , et d'après l'étude faite au chapitre III, on sait qu'il vérifie une condition de stabilité forte de type  $(W, \text{Id}, n_0)$  où  $W$  est une fonction dépendant de  $V$  et des moments de  $U_1$ . Plus précisément, il existe  $W$  telle que

$$\exists \alpha > 0, \exists \beta > 0, \exists n_0 \in \mathbf{N}, \forall n \geq n_0, \quad \frac{\mathbf{E}[W(X_{n+1}) | \mathcal{F}_n] - W(X_n)}{\gamma_{n+1}} \leq -\alpha W(X_n) + \beta, \quad (1.2)$$

$$\forall n \geq 1, \mathbf{E}[W(X_n)] < +\infty \quad \text{et} \quad \mathbf{E}\left[\left|\sqrt{W}(X_n) - \sqrt{W}(X_{n-1})\right|^2\right] = \mathcal{O}(\gamma_n). \quad (1.3)$$

De plus, si  $U_1$  est sous-gaussien et que  $\text{Tr}(\sigma\sigma^*) \leq C_\sigma V^{1-\varsigma}$ , on sait que l'on peut choisir la fonction  $W$  de la forme  $W = \exp(\lambda V^\varsigma)$  pour un  $\lambda > 0$ . Si  $U_1$  a un moment d'ordre  $2(p+1)$  ( $p \geq 2$ ) alors  $W = V^{p+1}$  est le meilleur choix possible. La fonction  $W$  que l'on considère dans la suite est la plus grande fonction possible qui satisfait les conditions (1.2) et (1.3).

On rappelle que, dans un tel cadre, le comportement de  $(X_n)_{n \geq 0}$  est régi par la proposition suivante (conséquence directe du théorème III.1).

**Proposition V.1:**

Soit  $(a_n)_{n \geq 1}$  une suite réelle positive. Si  $(\gamma_n)_{n \geq 1}$  tend vers 0 alors

- a.  $\sup_n \mathbf{E}[W(X_n)] < +\infty$ .
- b. Si  $(a_n)_{n \geq 1}$  décroît et que la série  $\sum_{n \geq 1} a_n \gamma_n$  converge alors la suite  $(a_n W(X_n))_{n \geq 1}$  converge presque sûrement vers 0.  $\blacklozenge$

Remarque: On utilisera souvent le résultat de la proposition V.1a. couplé avec le théorème de convergence monotone. Ainsi, on aura convergence presque sûre et dans  $\mathbf{L}^2$  de la série  $\sum_{n \geq 1} a_n \sqrt{W}(X_{n-1})$  si  $\sum_{n \geq 1} a_n < +\infty$ . \*

Les mesures empiriques pondérées de ce schéma sont construites à partir d'une suite de réels positifs  $(\eta_n)_{n \geq 1}$  et sont notées  $(\nu_n^\eta)_{n \geq 1}$ . Ainsi, pour toute fonction mesurable  $f$  et pour tout  $n \geq 1$ ,

$$\nu_n^\eta(f) = \frac{1}{\sum_{k=1}^n \eta_k} \sum_{k=1}^n \eta_k f(X_{k-1}).$$

De plus, on introduit la notion de couple pas-poids moyennisant qui servira tout au long de ce chapitre.

**Définition V.1.** Un couple pas-poids  $(\gamma, \eta)$  est dit *moyennisant* si les suites  $(\gamma_n)_{n \geq 0}$  et  $(\eta_n)_{n \geq 1}$  sont positives, termes général d'une série divergente et telles que

$$\lim_n \gamma_n = 0, \quad \sum_{n \geq 1} \frac{1}{H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ < +\infty \quad \text{et} \quad \sum_{n \geq 1} \left( \frac{\eta_n}{H_n \sqrt{\gamma_n}} \right)^2 < +\infty, \quad (1.4)$$

avec  $H_n = \sum_{k=1}^n \eta_k$ . ♣

Remarque: Un couple pas-poids  $(\gamma, \eta)$  est moyennisant si les suites  $(\gamma_n)_{n \geq 1}$  et  $(\eta_n)_{n \geq 1}$  vérifient la condition  $(C_s^{\eta, \gamma})$  avec  $s = 2$  (condition introduite dans le chapitre III). L'étude dans le cas général  $s \in ]1, 2]$  est similaire. \*

On rappelle le théorème de convergence obtenu dans le chapitre III :

**Théorème V.2:**

Si le couple  $(\gamma, \eta)$  est moyennisant et que les conditions (1.2) et (1.3) sont vérifiées, alors  $\sup_n \frac{1}{H_n} \sum_{k=1}^n \eta_k \sqrt{W}(X_{k-1}) < +\infty$  presque sûrement. De plus, pour toute fonction mesurable  $f$  négligeable devant  $\sqrt{W}$  (à l'infini) on a

$$\frac{1}{H_n} \sum_{k=1}^n \eta_k f(X_{k-1}) \xrightarrow{p.s.} \int_{\mathbf{R}^d} f d\nu. \quad (1.5) \quad \star$$

Nous allons étudier la vitesse de la convergence (1.5) afin de déterminer le couple pas-poids  $(\gamma, \eta)$  optimisant cette vitesse. L'étude de la vitesse de convergence se fait, comme pour le théorème V.1, pour une fonction  $f$  définie sur  $\mathbf{R}^d$ , admettant une solution  $\varphi$  à l'équation de Poisson

$$\mathcal{A}\varphi = -(f - \nu(f)).$$

Le développement de l'erreur  $\nu_n^\eta(f) - \nu(f)$  à l'ordre  $p \geq 2$  est obtenu lorsque la fonction  $\varphi$  est suffisamment régulière. Plus précisément, on définit l'ensemble  $E_p^W$  par

$$E_p^W = \left\{ \varphi \in \mathcal{C}^p(\mathbf{R}^d; \mathbf{R}), \forall j \in \{0, \dots, p\}, \forall x \in \mathbf{R}^d, \quad |D^j \varphi(x)|^2 = o\left(\frac{W(x)}{V^j(x)}\right) \right\}.$$

et on suppose que  $\varphi \in E_p^W$  et que  $D^p \varphi$  est Lipschitz. En pratique, les résultats obtenus sur la régularité de la solution de l'équation de Poisson sont de type Sobolev (cf. [PV01]).

La régularité imposée ici sur la fonction  $\varphi$  permet de définir les fonctions  $D_q$  ( $3 \leq q \leq p$ ) par

$$\forall x \in \mathbf{R}^d, \quad D_q(x) = \sum_{j \geq q/2}^q \frac{C_j^{q-j}}{j!} D^j \varphi(x) \cdot \left( (b(x))^{\otimes (q-j)}, \mathbf{E} \left[ (\sigma(x) U_1)^{\otimes (2j-q)} \right] \right). \quad (1.6)$$

Le développement de l'erreur d'ordre  $p$  pour  $p \geq 3$  s'écrit à l'aide de ces fonctions  $D_q$ . On a maintenant tous les éléments en main pour énoncer le théorème principal obtenu dans ce chapitre.

**Théorème V.3:**

Soit un entier  $p \geq 2$  tel que  $U_1 \in \mathbf{L}^{2(p+1)}$ , et  $\varphi \in E_p^W$  solution de l'équation de Poisson telle que  $D^p\varphi$  est Lipschitz. On définit  $q^*$  par

$$q^* = \min_{q \in \{3, \dots, p\}} \{D_q \neq 0\} \wedge (p+1).$$

Soit  $\theta_n = \eta_n^2/\gamma_n$  et  $\Theta_n = \sum_{k=1}^n \theta_k$ . On suppose que les couples  $(\gamma, \eta)$  et  $(\gamma, \theta)$  sont moyennisants et que  $(\gamma_n)_{n \geq 1}$  est décroissante. De plus, on pose  $\Upsilon_n^* = \sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1}$ .

Si la suite  $\left(\frac{\Upsilon_n^*}{\sqrt{\Theta_n}}\right)_{n \geq 1}$  admet une limite  $\xi$  dans  $\mathbf{R}_+ \cup \{+\infty\}$ , alors

- si  $\sum_{n \geq 1} \frac{1}{\sqrt{\Theta_n}} \left| \Delta \frac{\eta_n}{\gamma_n} \right| < +\infty$ , et  $\sum_{n \geq 1} \frac{\eta_n^2}{\Theta_n} < +\infty$ , on a

$$\frac{H_n}{\sqrt{\Theta_n}} (\nu_n^\eta(f) - \nu(f)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, s_f^2) \quad \text{si } \xi = 0, \quad (1.7)$$

$$\frac{H_n}{\sqrt{\Theta_n}} (\nu_n^\eta(f) - \nu(f)) \xrightarrow{\mathcal{L}} \mathcal{N}(\xi m_{f, q^*}, s_f^2) \quad \text{si } 0 < \xi < +\infty \text{ et } q^* \leq p, \quad (1.8)$$

- si  $\sum_{n \geq 1} \frac{1}{\Upsilon_n^*} \left| \Delta \frac{\eta_n}{\gamma_n} \right| < +\infty$ , et  $\sum_{n \geq 1} \frac{\eta_n^2}{(\Upsilon_n^*)^2} < +\infty$ , on a

$$\frac{H_n}{\Upsilon_n^*} (\nu_n^\eta(f) - \nu(f)) \xrightarrow{\mathbf{P}} m_{f, q^*} \quad \text{si } \xi = +\infty \text{ et } q^* \leq p, \quad (1.9)$$

avec

$$m_{f, q^*} = \int_{\mathbf{R}^d} D_{q^*} d\nu \quad \text{et} \quad s_f^2 = \int_{\mathbf{R}^d} |\sigma^* \nabla \varphi|^2 d\nu.$$

De plus, si  $\left(\frac{\eta_n}{\Upsilon_n^* \gamma_n}\right)_{n \geq 1}$  décroît et si  $\sum_{n \geq 1} \frac{\theta_n}{(\Upsilon_n^*)^2} < +\infty$ , alors la convergence dans (1.9) est presque sûre. ★

Remarque: Ce théorème généralise les théorèmes 9 et 10 de Lamberton et Pagès dans [LP02]. Le cas (a) du théorème 9 correspond à (1.7) avec  $p = 2$ , et le cas (b) correspond à (1.8) et (1.9) avec  $q^* = 3$  et  $p = 4$ . Le théorème 10 correspond au cas particulier  $D_3 = 0$  i.e.  $q^* = 4$ . Nos hypothèses sur le bruit  $U_1$  sont plus fortes que dans [LP02] mais ce théorème reste vrai dans le cas où  $U_1 \in \mathbf{L}^{2p}$ . Cependant, dans un souci de clarté nous prouvons ce résultat à la suite d'une étude dans  $\mathbf{L}^2$  du développement de l'erreur où nous montrons que le reste tend vers 0 (dans  $\mathbf{L}^2$ ). En fait, la convergence du reste dans  $\mathbf{L}^1$  (et p.s.) suffit pour obtenir le théorème V.3 auquel cas les hypothèses sur le bruit deviendraient identique. \*

Dans toute la suite de ce chapitre, on suppose que la suite  $(\gamma_n)_{n \geq 1}$  est décroissante et que  $(\gamma, \eta)$  et  $(\gamma, \theta)$  sont des couples pas-poids moyennisants, avec  $(\theta_n)_{n \geq 1}$  la suite définie par

$$\forall n \geq 1, \quad \theta_n = \frac{\eta_n^2}{\gamma_n}.$$

Remarque: Dans ce cadre, on a

$$\frac{1}{H_n} \sup_{k \leq n} \eta_k \longrightarrow 0 \quad \text{et} \quad \frac{1}{\Theta_n} \sup_{k \leq n} \theta_k \longrightarrow 0. \quad (1.10)$$

En effet,  $\sup_{k \leq n} \eta_k \leq \sup_{k \leq n} \left( \frac{\eta_k}{\gamma_k} \right) \gamma_1$  et en posant  $\eta_0 = 0$ , on en déduit  $\sup_{k \leq n} \eta_k \leq \gamma_1 \sum_{k=1}^n \left( \Delta \frac{\eta_k}{\gamma_k} \right)_+$ . Le lemme de Kronecker appliqué avec la suite  $(H_n)_{n \geq 0}$  et combiné avec (1.4) donne le résultat. \*

Il existe une large classe de couples  $(\eta, \gamma)$  entrant dans notre cadre de travail. Les hypothèses faites jusqu'à présent ne sont pas très restrictives. Des conditions nécessaires et suffisantes pour les vérifier lorsque les pas et les poids sont des suites "puissances" sont données par la proposition suivante (dont la preuve est donnée en annexe, section 7).

**Proposition V.2:**

Si  $\gamma_n = \gamma_0 n^{-r}$  et  $\eta_n = \eta_0 n^{-s}$  alors les couples  $(\gamma, \eta)$  et  $(\gamma, \theta)$  sont moyennisants si et seulement si

$$\left( 0 < r < 1 \quad \text{et} \quad s \leq \frac{1+r}{2} \right) \quad \text{ou} \quad (r = s = 1). \quad \blacklozenge$$

Remarque: Lorsque les suites  $(\gamma_n)_{n \geq 1}$  et  $(\eta_n)_{n \geq 1}$  sont de la forme  $\gamma_n = \gamma_0 n^{-r}$  et  $\eta_n = \eta_0 n^{-s}$ , le TLC (1.7) a lieu lorsque  $\frac{1}{q^*-1} < r < 1$  et  $s \leq \frac{1+r}{2}$ , et la vitesse est en  $n^{\frac{1-r}{2}}$  i.e.

$$c_{\mathcal{L}}(\gamma_0, r, s) n^{\frac{1-r}{2}} (\nu_n^\eta(f) - \nu(f)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, s_f^2),$$

où  $c_{\mathcal{L}}(\gamma_0, r, s) = \sqrt{\gamma_0} \frac{\sqrt{1-2s+r}}{1-s}$ .

D'un autre côté, si  $0 < r < \frac{1}{q^*-1}$  et  $s \leq \frac{1+r}{2}$ , on a alors convergence presque sûre

$$c_{p.s.}(\gamma_0, r, s, q^*) n^{r \left( \frac{q^*}{2} - 1 \right)} (\nu_n^\eta(f) - \nu(f)) \xrightarrow{p.s.} m_{f, q^*},$$

où  $c_{p.s.}(\gamma_0, r, s, q^*) = \gamma_0^{1-\frac{q^*}{2}} \left( 1 - \frac{r \left( \frac{q^*}{2} - 1 \right)}{1-2} \right)$ .

Plus de détails sont donnés dans la section 5. \*

## 2. Résultats préliminaires

Nous donnons tout d'abord un critère simple pour savoir si un couple est moyennisant. Ce lemme est prouvé en annexe.

**Lemme V.1:**

Si  $(\gamma, \eta)$  est moyennisant et que  $(\gamma_n)_{n \geq 1}$  est décroissante, alors pour tout  $q \geq 0$

$$\sum_{n \geq 1} \eta_n \gamma_n^q = +\infty \quad \Rightarrow \quad (\gamma, \eta \gamma^q) \text{ est moyennisant.} \quad \blacklozenge$$

Les deux lemmes suivant nous seront aussi très utiles tout au long de l'étude de l'erreur. Le premier est la clef pour établir l'analogie du théorème V.1 pour les mesures empiriques pondérées.

**Lemme V.2:**

Soit  $g \in \mathcal{C}(\mathbf{R}^d; \mathbf{R}^m)$  telle que  $|g|^2 = o(W)$ . Alors

$$\frac{1}{\sqrt{\Theta_n}} \sum_{k=1}^n \sqrt{\theta_k} \langle g(X_{k-1}), U_k \rangle \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \int |g|^2 d\nu\right). \quad (2.1) \quad \blacklozenge$$

Preuve. Le lemme découle du théorème de la limite centrale pour les martingales (voir [HH80] ou [Duf97]). On considère la martingale  $(M_n)_{n \geq 0}$ , nulle en 0, définie par

$$\forall n \geq 1, \quad M_n = \sum_{k=1}^n \sqrt{\theta_k} \langle g(X_{k-1}), U_k \rangle.$$

Elle est de carré intégrable et son crochet est défini pour  $n \geq 1$  par  $\langle M \rangle_n = \sum_{k=1}^n \theta_k |g(X_{k-1})|^2$ . Le couple  $(\gamma, \theta)$  est moyennisant, et la fonction  $|g|^2$  est continue et négligeable devant  $W$  à l'infini, donc par le théorème V.2 on a

$$\frac{\langle M \rangle_n}{\Theta_n} \xrightarrow{p.s.} \int_{\mathbf{R}^d} |g|^2 d\nu.$$

Pour appliquer le TLC pour les martingales, il suffit de vérifier la condition de Lindeberg. Soit  $\varepsilon > 0$ . Il faut montrer que

$$L_n^\varepsilon := \frac{1}{\Theta_n} \sum_{k=1}^n \theta_k |g(X_{k-1})|^2 K\left(X_{k-1}, \varepsilon \sqrt{\frac{\Theta_n}{\theta_k}}\right) \xrightarrow{\mathbf{P}} 0,$$

avec

$$K(x, y) = \mathbf{E} \left[ |U_1|^2 \mathbf{1}_{\{|\langle g(x), U_1 \rangle| \geq y\}} \right].$$

Il est clair que pour tout  $A > 0$ ,  $\lim_{y \rightarrow +\infty} \sup_{|x| \leq A} K(x, y) = 0$ . Soit  $A > 0$ . Alors

$$\begin{aligned} L_n^\varepsilon &= \frac{1}{\Theta_n} \sum_{k=1}^n \theta_k |g(X_{k-1})|^2 K\left(X_{k-1}, \varepsilon \sqrt{\frac{\Theta_n}{\theta_k}}\right) \mathbf{1}_{\{|X_{k-1}| \leq A\}} \\ &\quad + \frac{1}{\Theta_n} \sum_{k=1}^n \theta_k |g(X_{k-1})|^2 K\left(X_{k-1}, \varepsilon \sqrt{\frac{\Theta_n}{\theta_k}}\right) \mathbf{1}_{\{|X_{k-1}| > A\}}, \\ &\leq \nu_n^\theta(|g|^2) \sup_{|x| \leq A} K\left(x, \varepsilon \sqrt{\frac{\Theta_n}{\sup_{k \leq n} \theta_k}}\right) + \sup_{|x| > A} \left(\frac{|g(x)|^2}{W(x)}\right) \nu_n^\theta(W). \end{aligned}$$

Par (1.10), on a  $\lim_n \frac{\Theta_n}{\sup_{k \leq n} \theta_k} = +\infty$ , et comme  $(\gamma, \theta)$  est moyennisant on a  $\sup_n \nu_n^\theta(|g|^2)$  fini presque sûrement donc

$$0 \leq \lim_n L_n^\varepsilon \leq \sup_{|x| > A} \left( \frac{|g(x)|^2}{W(x)} \right) \sup_n \nu_n^\theta(W) \quad p.s.$$

Comme  $|g|^2 = o(W)$ , on conclut en faisant tendre  $A$  vers l'infini. On a montré plus que la condition de Lindeberg car on a convergence presque sûre.  $\#$

Ce second lemme, qui est une conséquence directe du théorème de convergence pour les martingales de carré intégrable et du lemme de Kronecker, sera souvent utilisé.

**Lemme V.3:**

Soit  $\Psi : \mathbf{R}^d \times \mathbf{R}^m \rightarrow \mathbf{R}$  vérifiant : il existe une fonction  $h \in \mathbf{L}^2(\mathbf{P}_{U_1})$  et une constante  $C > 0$  telles que

$$\forall (x, u) \in \mathbf{R}^d \times \mathbf{R}^m, \quad |\Psi(x, u)|^2 \leq CW(x)h^2(u). \quad (2.2)$$

Soit  $(a_n)_{n \geq 1}$  une suite réelle positive et  $(b_n)_{n \geq 1}$  une suite réelle positive décroissant vers 0, telles que la série  $\sum_{n \geq 1} (a_n b_n)^2$  converge. Alors

$$b_n \sum_{k=1}^n a_k \Psi(X_{k-1}, U_k) - b_n \sum_{k=1}^n a_k \mathbf{E}[\Psi(x, U_1)]_{|x=X_{k-1}} \xrightarrow[p.s.]{\mathbf{L}^2} 0. \quad \blacklozenge$$

Preuve. On considère la martingale  $(M_n)_{n \geq 0}$ , nulle en 0, définie par

$$\forall n \geq 1, \quad M_n = \sum_{k=1}^n b_k a_k \left( \Psi(X_{k-1}, U_k) - \mathbf{E}[\Psi(X_{k-1}, U_k) | \mathcal{F}_{k-1}] \right).$$

Son crochet prévisible vérifie

$$\langle M \rangle_n \leq \sum_{k=1}^n (b_k a_k)^2 \mathbf{E}[\Psi^2(X_{k-1}, U_k) | \mathcal{F}_{k-1}],$$

et d'après la condition (2.2) et l'hypothèse sur la série de terme général  $a_n b_n$ , il vient

$$\mathbf{E}[\langle M \rangle_\infty] \leq C \left( \sum_{n \geq 1} (b_n a_n)^2 \right) \left( \sup_n \mathbf{E}[W(X_n)] \right) \mathbf{E}[h^2(U_1)] < +\infty.$$

La martingale  $(M_n)_{n \geq 0}$  converge donc presque sûrement et dans  $\mathbf{L}^2$ . On conclut via le lemme de Kronecker appliqué avec la suite décroissante vers 0  $(b_n)_{n \geq 1}$  que

$$b_n \sum_{k=1}^n a_k \left( \Psi(X_{k-1}, U_k) - \mathbf{E}[\Psi(X_{k-1}, U_k) | \mathcal{F}_{k-1}] \right) \xrightarrow[p.s.]{\mathbf{L}^2} 0. \quad (2.3)$$

D'autre part,

$$\left\| b_n \sum_{k=1}^n a_k \left( \Psi(X_{k-1}, U_k) - \mathbf{E}[\Psi(X_{k-1}, U_k) | \mathcal{F}_{k-1}] \right) \right\|_2^2 \leq b_n^2 \sum_{k=1}^n a_k^2 \mathbf{E}[\Psi^2(X_{k-1}, U_k)]$$

et comme  $\lim_n b_n^2 \sum_{k=1}^n a_k^2 = 0$  (par Kronecker) alors la convergence (2.3) est vraie dans  $\mathbf{L}^2$ . #

### 3. Développement de l'erreur

Le but de cette section est de prouver le théorème V.3 qui est l'analogue du théorème V.2 pour les mesures empiriques pondérées par  $(\eta_n)_{n \geq 1}$ . On rappelle que tous nos résultats sur les vitesses de convergence de l'algorithme sont obtenus pour des fonctions tests  $f$  continues qui admettent une solution  $\varphi$  suffisamment régulière à l'équation de Poisson

$$\mathcal{A}\varphi = -(f - \nu(f)). \quad (3.1)$$

On établit tout d'abord un développement de l'erreur  $\nu_n^\eta(f) - \nu(f)$  fondée sur la formule de Taylor (cf. proposition V.3). Puis on étudiera l'erreur dans  $\mathbf{L}^2$  (cf. proposition V.4) afin de montrer le théorème annoncé et de préparer à l'étude de l'erreur lorsqu'une extrapolation de Richardson est mise en œuvre (cf. section 6).

On rappelle que la diffusion vérifie une condition de stabilité forte de type  $(V, \text{Id}, \alpha)$  et que  $|b|^2 + \text{Tr}(\sigma\sigma^*) = \mathcal{O}(V)$ .

**Proposition V.3:**

Soit  $\varphi : \mathbf{R}^d \rightarrow \mathbf{R}$  une application de classe  $\mathcal{C}^p$  ( $p \geq 2$ ) solution de l'équation de Poisson (3.1) et telle que  $D^p\varphi$  est Lipschitz. Alors

$$\begin{aligned} \nu_n^\eta(f) - \nu(f) &= -\frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (\Delta\varphi(X_k)) + \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\sqrt{\gamma_k}} \langle \nabla\varphi(X_{k-1}), \sigma(X_{k-1})U_k \rangle \\ &+ \frac{1}{2H_n} \sum_{k=1}^n \eta_k \left( D^2\varphi(X_{k-1}) \cdot (\sigma(X_{k-1})U_k)^{\otimes 2} - \text{Tr}(D^2\varphi(X_{k-1})(\sigma^*\sigma)(X_{k-1})) \right) \\ &+ \sum_{q=3}^{2p} \left( \frac{1}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} d_{q,p}(X_{k-1}, U_k) \right) + R_n^{(p)} \end{aligned}$$

avec

$$d_{q,p}(x, u) = \sum_{j \geq q/2} \frac{C^{q-j}}{j!} D^j\varphi(x) \cdot \left( (b(x))^{\otimes(q-j)}, (\sigma(x)u)^{\otimes(2j-q)} \right), \quad (3.2)$$

et

$$\left| R_n^{(p)} \right| \leq \frac{C}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{p-1}{2}} V^{\frac{p+1}{2}}(X_{k-1}) (1 + |U_k|^{p+1}). \quad (3.3) \quad \blacklozenge$$

Preuve. Soit  $k \geq 1$ . La formule de Taylor appliquée à  $\varphi$  entre  $X_{k-1}$  et  $X_k$  donne

$$\varphi(X_k) = \varphi(X_{k-1}) + \sum_{j=1}^p \frac{1}{j!} D^j\varphi(X_{k-1}) \cdot (\Delta X_k)^{\otimes j} + r_p(X_{k-1}, X_k) \quad (3.4)$$

avec

$$\left| r_p(X_{k-1}, X_k) \right| \leq \sup_{t \in ]0,1[} |D^p\varphi(X_{k-1} + t\Delta X_k) - D^p\varphi(X_{k-1})| \frac{|\Delta X_k|^p}{p!}. \quad (3.5)$$

Pour tout  $x \in \mathbf{R}^d$ ,  $D^j\varphi(x)$  est une application multilinéaire symétrique donc

$$D^j\varphi(x) \cdot (\Delta X_k)^{\otimes j} = \sum_{l=0}^j C_j^l D^j\varphi(x) \cdot \left( (\gamma_k b(X_{k-1}))^{\otimes l}, (\sqrt{\gamma_k} \sigma(X_{k-1}) U_k)^{\otimes (j-l)} \right)$$

et par homogénéité

$$\sum_{j=1}^p \frac{1}{j!} D^j\varphi(x) \cdot (\Delta X_k)^{\otimes j} = \sum_{j=1}^p \sum_{l=0}^j \frac{C_j^l}{j!} \gamma_k^{\frac{j+l}{2}} D^j\varphi(x) \cdot \left( (b(X_{k-1}))^{\otimes l}, (\sigma(X_{k-1}) U_k)^{\otimes (j-l)} \right).$$

Par interversion des deux signes sommes puis en utilisant le changement de variable  $q = j + l$  on obtient

$$\begin{aligned} \sum_{j=1}^p \frac{1}{j!} D^j\varphi(x) \cdot (\Delta X_k)^{\otimes j} &= \sum_{l=0}^p \sum_{j=1}^p \mathbf{1}_{\{l \leq j\}} \frac{C_j^l}{j!} \gamma_k^{\frac{j+l}{2}} D^j\varphi(x) \cdot \left( (b(X_{k-1}))^{\otimes l}, (\sigma(X_{k-1}) U_k)^{\otimes (j-l)} \right), \\ &= \sum_{q=1}^{2p} \sum_{j \geq q/2} \frac{C_j^{q-j}}{j!} \gamma_k^{\frac{q}{2}} D^j\varphi(x) \cdot \left( (b(X_{k-1}))^{\otimes (q-j)}, (\sigma(X_{k-1}) U_k)^{\otimes (2j-q)} \right). \end{aligned}$$

Avec la fonction  $d_{p,q}$  définie en (3.2) et d'après (3.4) on a

$$\Delta\varphi(X_k) = \sum_{q=1}^{2p} \gamma_k^{\frac{q}{2}} d_{q,p}(X_{k-1}, U_k) + r_p(X_{k-1}, X_k). \quad (3.6)$$

On remarque que pour tout  $p \geq 2$  on a

$$\begin{aligned} d_{1,p}(X_{k-1}, U_k) &= \langle \nabla\varphi(X_{k-1}), \sigma(X_{k-1}) U_k \rangle, \\ d_{2,p}(X_{k-1}, U_k) &= \langle \nabla\varphi(X_{k-1}), b(X_{k-1}) \rangle + \frac{1}{2} D^2\varphi(X_{k-1}) \cdot (\sigma(X_{k-1}) U_k)^{\otimes 2}, \end{aligned}$$

et comme la fonction  $\varphi$  est solution de l'équation  $-(f - \nu(f)) = \mathcal{A}\varphi$  alors

$$\begin{aligned} \Delta\varphi(X_k) &= \sqrt{\gamma_k} \langle \nabla\varphi(X_{k-1}), \sigma(X_{k-1}) U_k \rangle - \gamma_k (-f(X_{k-1}) - \nu(f)) \\ &\quad + \frac{1}{2} D^2\varphi(X_{k-1}) \cdot (\sigma(X_{k-1}) U_k)^{\otimes 2} - \frac{1}{2} \text{Tr}(D^2\varphi(X_{k-1})(\sigma^* \sigma)(X_{k-1})) \\ &\quad + \sum_{q=3}^{2p} \gamma_k^{\frac{q}{2}} d_{q,p}(X_{k-1}, U_k) + r_p(X_{k-1}, X_k). \end{aligned}$$

Ainsi en réarrangeant les termes et en sommant on obtient

$$\begin{aligned} \nu_n^\eta(f) - \nu(f) &= -\frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (\Delta\varphi(X_k)) + \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\sqrt{\gamma_k}} \langle \nabla\varphi(X_{k-1}), \sigma(X_{k-1})U_k \rangle \\ &\quad + \frac{1}{2H_n} \sum_{k=1}^n \eta_k \left( D^2\varphi(X_{k-1}) \cdot (\sigma(X_{k-1})U_k)^{\otimes 2} - \text{Tr}(D^2\varphi(X_{k-1})(\sigma^*\sigma)(X_{k-1})) \right) \\ &\quad + \frac{1}{H_n} \sum_{k=1}^n \sum_{q=3}^{2p} \eta_k \gamma_k^{\frac{q}{2}-1} d_{q,p}(X_{k-1}, X_k) + R_n^{(p)}, \end{aligned}$$

avec  $R_n^{(p)} = \frac{1}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{-1} r_p(X_{k-1}, X_k)$  et  $r_p$  vérifiant (3.5). Comme  $|b|^2 + \text{Tr}(\sigma\sigma^*) = \mathcal{O}(V)$ , il existe  $C > 0$  tel que  $|\Delta X_k| \leq C\sqrt{\gamma_k}\sqrt{V}(X_{k-1})(1 + |U_k|)$  et  $D^p\varphi$  est Lipschitz donc

$$|r_p(X_{k-1}, X_k)| \leq C\gamma_k^{\frac{p+1}{2}} V^{\frac{p+1}{2}}(X_{k-1})(1 + |U_k|^{p+1}).$$

On en déduit aisément la majoration (3.3) du reste  $R_n^{(p)}$  et donc la proposition. #

La proposition qui suit est importante pour connaître le comportement de l'erreur dans  $\mathbf{L}^2$  et presque sûrement. Sous des conditions peu restrictives sur les suites  $(\gamma_n)_{n \geq 1}$  et  $(\eta_n)_{n \geq 1}$ , cette proposition permet de ramener l'étude de  $\nu_n^\eta(f) - \nu(f)$  à la somme d'un terme martingale et d'un reste.

On rappelle que  $V$  est la fonction de Lyapounov associée à l'EDS et que  $W$  est la plus grande fonction possible qui permet de garder une structure de rappel pour le schéma  $(X_n)_{n \geq 0}$ . De plus, s'il existe  $p \geq 2$  tel que  $U_1 \in \mathbf{L}^{2(p+1)}$  alors  $V^{p+1} = \mathcal{O}(W)$ , et on introduit l'ensemble de fonctions suivant

$$E_p^W = \left\{ \varphi \in \mathcal{C}^p(\mathbf{R}^d; \mathbf{R}), \forall j \in \{0, \dots, p\}, \forall x \in \mathbf{R}^d, \quad |D^j\varphi(x)|^2 = o\left(\frac{W(x)}{V^j(x)}\right) \right\}.$$

**Proposition V.4:**

Soit  $p \geq 2$  tel que  $U_1 \in \mathbf{L}^{2(p+1)}$ , et  $\varphi$  une fonction de  $E_p^W$  solution de l'équation de Poisson (3.1) telle que  $D^p\varphi$  est Lipschitz. Si  $(v_n)_{n \geq 1}$  est une suite réelle positive telle que  $(v_n/H_n)_{n \geq 1}$  décroisse vers 0 et que

$$\sum_{n \geq 1} \frac{v_n}{H_n} \left| \Delta \frac{\eta_n}{\gamma_n} \right| < +\infty, \quad \sum_{n \geq 1} \left( \frac{v_n \eta_n}{H_n} \right)^2 < +\infty \quad \text{et} \quad \lim_n \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{p-1}{2}} = 0, \quad (3.7)$$

alors la suite  $(v_n(\nu_n^\eta(f) - \nu(f)) - \varepsilon_n)_{n \geq 1}$  converge dans  $\mathbf{L}^2$  vers 0, où  $(\varepsilon_n)_{n \geq 1}$  est définie pour tout  $n \geq 1$  par

$$\varepsilon_n = \frac{v_n}{H_n} \sum_{k=1}^n \frac{\eta_k}{\sqrt{\gamma_k}} \langle \nabla\varphi(X_{k-1}), \sigma(X_{k-1})U_k \rangle + \sum_{q=3}^p \left( \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_q(X_{k-1}) \right),$$

avec

$$D_q(x) := \mathbf{E} [d_{q,q}(x, U_1)] = \sum_{j \geq q/2}^q \frac{C_j^{q-j}}{j!} D^j \varphi(x) \cdot \left( (b(x))^{\otimes(q-j)}, \mathbf{E} \left[ (\sigma(x) U_1)^{\otimes(2j-q)} \right] \right).$$

De plus, si la suite  $\left( \frac{v_n \eta_n}{H_n \gamma_n} \right)_{n \geq 1}$  est décroissante vers 0 et la série  $\sum_{n \geq 1} \left( \frac{v_n \eta_n}{H_n} \right)^2 \frac{1}{\gamma_n}$  est convergente alors la convergence est presque sûre.  $\blacklozenge$

Preuve. On utilise la proposition V.3 pour obtenir un développement à l'ordre  $p$  de l'erreur  $\nu_n^\eta(f) - \nu(f)$  en  $n$ .

- Montrons tout d'abord que

$$\frac{v_n}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (\Delta \varphi(X_k)) \xrightarrow{\mathbf{L}^2} 0, \quad (3.8)$$

et que cette convergence est presque sûre lorsque l'on suppose la suite  $\left( \frac{v_n \eta_n}{H_n \gamma_n} \right)_{n \geq 1}$  décroissante vers 0 et la série  $\sum_{n \geq 1} \left( \frac{v_n \eta_n}{H_n} \right)^2 \frac{1}{\gamma_n}$  finie. D'une part, la transformation d'Abel donne

$$\frac{v_n}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (\Delta \varphi(X_k)) = \frac{v_n \eta_n}{H_n \gamma_n} \varphi(X_n) - \frac{v_n}{H_n} \sum_{k=1}^n \left( \Delta \frac{\eta_k}{\gamma_k} \right) \varphi(X_{k-1}).$$

D'autre part, la série  $\sum_{n \geq 1} \frac{v_n}{H_n} \left| \Delta \frac{\eta_n}{\gamma_n} \right|$  est supposée convergente et  $|\varphi|^2 = o(W)$  (car  $\varphi \in \mathbf{E}_p^W$ ) donc d'après la remarque suivant la proposition V.1, la série  $\sum_{n \geq 1} \frac{v_n}{H_n} \left| \Delta \frac{\eta_n}{\gamma_n} \right| |\varphi(X_{n-1})|$  converge presque sûrement et dans  $\mathbf{L}^2$ . En appliquant le lemme de Kronecker avec la suite  $(v_n/H_n)_{n \geq 1}$  décroissante vers 0, on obtient alors

$$\frac{v_n}{H_n} \sum_{k=1}^n \left( \Delta \frac{\eta_k}{\gamma_k} \right) \varphi(X_{k-1}) \xrightarrow[p.s.]{\mathbf{L}^2} 0.$$

Ainsi il reste uniquement le terme  $\frac{v_n \eta_n}{H_n \gamma_n} \varphi(X_n)$  à gérer. La convergence dans  $\mathbf{L}^2$  vers 0 de ce terme est triviale du fait de la convergence de la suite  $\left( \frac{v_n \eta_n}{H_n \gamma_n} \right)_{n \geq 1}$  vers 0 et la  $\mathbf{L}^2$ -bornitude de  $\varphi(X_n)$ . Pour la convergence  $p.s.$ , il suffit d'appliquer la proposition V.1b. avec la suite  $\left( \left( \frac{v_n \eta_n}{H_n \gamma_n} \right)^2 \right)_{n \geq 1}$ .

- Montrons maintenant que

$$\frac{v_n}{2H_n} \sum_{k=1}^n \eta_k \left( D^2 \varphi(X_{k-1}) \cdot (\sigma(X_{k-1}) U_k)^{\otimes 2} - \text{Tr}(D^2 \varphi(X_{k-1}) (\sigma^* \sigma)(X_{k-1})) \right) \xrightarrow[p.s.]{\mathbf{L}^2} 0. \quad (3.9)$$

D'une part  $\varphi \in E_p^W$  et  $\text{Tr} \sigma \sigma^* = \mathcal{O}(V)$ , donc  $\left( D^2 \varphi(x) \cdot (\sigma(x))^{\otimes 2} \right)^2 = o(W(x))$ . D'autre part, on sait que  $\mathbf{E} [U_1^{\otimes 2}] = \text{Id}$  donc la fonction  $\Psi(x, u) = D^2 \varphi(x) \cdot (\sigma(x) u)^{\otimes 2} - \text{Tr}(D^2 \varphi(x) \sigma^* \sigma(x))$  satisfait (2.2) avec  $h(u) = u$ . De plus, la série  $\sum_{n \geq 1} \left( \frac{v_n \eta_n}{H_n} \right)^2$  converge donc on peut appliquer le lemme V.3 avec les suites  $(a_n)_{n \geq 1} = (\eta_n)_{n \geq 1}$  et  $(b_n)_{n \geq 1} = (v_n/H_n)_{n \geq 1}$  et on obtient (3.9).

- Toujours d'après le développement de Taylor obtenu par la proposition V.3, on va montrer

que

$$\sum_{q=3}^{2p} \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} d_{q,p}(X_{k-1}, U_k) - \sum_{q=3}^p \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_q(X_{k-1}) \xrightarrow[p.s.]{\mathbf{L}^2} 0,$$

où  $D_q$  est définie en (1.6).

Soit  $q \in \{3, \dots, 2p\}$ . Comme  $\varphi \in E_p^W$  et  $|b|^2 + \text{Tr}(\sigma\sigma^*) = \mathcal{O}(V)$  alors

$$|d_{q,p}(x, u)|^2 = o\left(W(x)(1 + |u|^{2p})\right). \quad (3.10)$$

On applique le lemme V.3 avec les suites  $(a_n)_{n \geq 1} = (\eta_n \gamma_n^{\frac{q}{2}-1})_{n \geq 1}$  et  $(b_n)_{n \geq 1} = (v_n/H_n)_{n \geq 1}$  et la fonction  $\Psi(x, u) = d_{q,p}(x, u)$  et on obtient

$$\frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} d_{q,p}(X_{k-1}, U_k) - \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_{q,p}(X_{k-1}) \xrightarrow[p.s.]{\mathbf{L}^2} 0.$$

avec  $D_{q,p}(x) = \mathbf{E}[d_{q,p}(x, U_1)]$ . En notant  $D_q = D_{q,q}$  et en remarquant que  $D_{q,p} \mathbf{1}_{\{q \leq p\}} = D_q \mathbf{1}_{\{q \leq p\}}$  on a

$$\sum_{q=3}^{2p} \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_{q,p}(X_{k-1}) = \sum_{q=3}^p \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_q(X_{k-1}) + \tilde{R}_n^{(p)},$$

avec

$$\tilde{R}_n^{(p)} = \sum_{q=p+1}^{2p} \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_{q,p}(X_{k-1}).$$

Pour conclure il faut donc montrer que  $(\tilde{R}_n^{(p)})_{n \geq 1}$  converge presque sûrement et dans  $\mathbf{L}^2$  vers 0. D'après (3.10), pour tout  $q \in \{3, \dots, 2p\}$  on a  $|D_{q,p}|^2 = o(W)$  et comme la suite  $(\gamma_n)_{n \geq 1}$  tend vers 0 alors pour tout  $q \geq p+1$ ,  $\gamma_k^{\frac{q}{2}-1} D_{q,p}(X_{k-1}) = o\left(\gamma_k^{\frac{p-1}{2}} \sqrt{W}(X_{k-1})\right)$ .

Ainsi,

$$\left| \tilde{R}_n^{(p)} \right| \leq C \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{p-1}{2}} \sqrt{W}(X_{k-1}). \quad (3.11)$$

La convergence dans  $\mathbf{L}^2$  vers 0 est alors triviale car on a supposé  $\lim_n \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{p-1}{2}} = 0$  et que l'on a  $\sup_n \mathbf{E}[W(X_n)] < +\infty$  (voir proposition V.1a.). Pour la convergence presque sûre, on utilise l'alternative suivante :

- soit la série  $\sum_{n \geq 1} \eta_n \gamma_n^{\frac{p-1}{2}}$  converge, alors la proposition V.1a. et le théorème de convergence monotone donnent

$$\sum_{n \geq 1} \eta_n \gamma_n^{\frac{p-1}{2}} \sqrt{W}(X_{n-1}) < +\infty \quad p.s.$$

- soit la série diverge, et par le lemme V.1 le couple  $(\gamma, \eta \gamma^{\frac{p-1}{2}})$  est moyennisant ce qui implique

$$\sum_{k=1}^n \eta_k \gamma_k^{\frac{p-1}{2}} \sqrt{W}(X_{k-1}) = \mathcal{O}\left(\sum_{k=1}^n \eta_k \gamma_k^{\frac{p-1}{2}}\right) \quad p.s.$$

On conclut en utilisant  $\lim_n \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{p-1}{2}} = 0$ .

- Il reste à montrer la convergence *p.s.* et  $\mathbf{L}^2$  de  $v_n R_n^{(p)}$  vers 0. On rappelle (3.3) *i.e.*

$$\left| R_n^{(p)} \right| \leq \frac{C}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{p-1}{2}} V^{\frac{p+1}{2}}(X_{k-1}) \left( 1 + |U_k|^{p+1} \right),$$

et d'après le lemme V.3 appliqué avec la fonction  $\Psi(x, u) = V^{\frac{p+1}{2}}(x) \left( 1 + |u|^{p+1} \right)$  il suffit de montrer que

$$\frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{p-1}{2}} V^{\frac{p+1}{2}}(X_{k-1}) \xrightarrow[p.s.]{\mathbf{L}^2} 0. \quad (3.12)$$

Or  $V^{p+1} = \mathcal{O}(W)$  donc on est ramené à (3.11). #

D'après la proposition précédent, seuls deux termes sont importants pour une étude dans  $\mathbf{L}^2$  et presque sûre de l'erreur  $\nu_n^\eta(f) - \nu(f)$ . Il s'agit d'un terme martingale de la forme

$$M_n = \frac{v_n}{H_n} \sum_{k=1}^n \frac{\eta_k}{\sqrt{\gamma_k}} \langle \nabla \varphi(X_{k-1}), \sigma(X_{k-1}) U_k \rangle,$$

et d'un reste

$$S_n = \sum_{q=3}^p \left( \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_q(X_{k-1}) \right).$$

En fonction des suites  $(\gamma_n)_{n \geq 1}$  et  $(\eta_n)_{n \geq 1}$  utilisées pour discrétiser la diffusion et pondérer les mesures empiriques, on a soit une domination de  $M_n$ , soit une domination de  $S_n$ . Dans le premier cas, le lemme V.2 implique que l'erreur vérifie un théorème de la limite centrale. Dans le second cas, on obtient une vitesse de convergence en probabilité ou presque sûre. Les conditions sur les pas et les poids sont données par le théorème V.3, que nous montrons après ces quelques remarques.

Remarque: Lorsque  $(U_n)_{n \geq 0}$  est un bruit blanc symétrique, alors pour tout  $q \geq 1$ ,  $D_{2q+1} = 0$ . En effet, d'après (1.6) on a

$$D_{2q+1}(x) = \sum_{j=q+1}^{2q+1} \frac{C_j^{2q+1-j}}{j!} D^j \varphi(x) \cdot \left( (b(x))^{\otimes (2q+1-j)}, \mathbf{E} \left[ (\sigma(x) U_1)^{\otimes (2j-2q-1)} \right] \right).$$

Notamment,  $D_3 = 0$  si  $\mathbf{E} [U_1^{\otimes 3}] = 0$ . L'indice générique de départ de la somme  $S_n$  est alors 4 (l'exception correspondant au cas où  $\varphi$  est une fonction particulière annulant  $D_4$ ). \*

S'il existe un plus petit indice  $q$  de  $\{3, \dots, p\}$  tel que  $D_q$  soit non nulle alors on le note  $q^*$ . Sinon on note  $q^* = p + 1$ .

Preuve (théorème V.3). On rappelle que  $\xi$  est la limite de  $\frac{\Upsilon_n^*}{\sqrt{\Theta_n}}$  où  $\Upsilon_n^* = \sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1}$  et  $\Theta_n = \sum_{k=1}^n \frac{\eta_k^2}{\gamma_k}$ .

- Supposons  $\xi = 0$ . Pour tout  $q \in \{3, \dots, p\}$ ,  $D_q = o(\sqrt{W})$  donc il est clair que

$$\sum_{q=q^*}^p \frac{1}{\sqrt{\Theta_n}} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_q(X_{k-1}) \xrightarrow{\mathbf{L}^2} 0.$$

Comme  $\Upsilon_n^* = o(\sqrt{\Theta_n})$  et que  $q^* \leq p+1$ , alors  $\sum_{k=1}^n \eta_k \gamma_k^{\frac{p-1}{2}} = o(\sqrt{\Theta_n})$ . Ainsi, la suite  $(v_n)_{n \geq 1} = (H_n / \sqrt{\Theta_n})_{n \geq 1}$  vérifie (3.7) et la proposition V.4 peut s'appliquer. On obtient alors

$$\frac{H_n}{\sqrt{\Theta_n}} (\nu_n^\eta(f) - \nu(f)) - \frac{1}{\sqrt{\Theta_n}} \sum_{k=1}^n \frac{\eta_k}{\sqrt{\gamma_k}} \langle \nabla \varphi(X_{k-1}), \sigma(X_{k-1}) U_k \rangle \xrightarrow{\mathbf{L}^2} 0. \quad (3.13)$$

On conclut par le lemme V.2 appliqué avec la fonction continue  $\sigma^* \nabla \varphi$  négligeable devant  $W$  (car  $\varphi \in \mathbf{E}_p^W$ ).

- Si  $\xi < +\infty$ , alors il est facile de vérifier que

$$\sum_{q=q^*+1}^p \frac{1}{\sqrt{\Theta_n}} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_q(X_{k-1}) \xrightarrow{\mathbf{L}^2} 0.$$

Comme  $q^* \leq p$  et que  $\Upsilon_n^* \sim \xi \sqrt{\Theta_n}$ , il est facile de vérifier que les conditions (3.7) sont satisfaites avec la suite  $(v_n)_{n \geq 1} = (H_n / \sqrt{\Theta_n})_{n \geq 1}$ . La proposition V.4 donne alors

$$\begin{aligned} & \frac{H_n}{\sqrt{\Theta_n}} (\nu_n^\eta(f) - \nu(f)) \\ & - \frac{1}{\sqrt{\Theta_n}} \sum_{k=1}^n \frac{\eta_k}{\sqrt{\gamma_k}} \langle \nabla \varphi(X_{k-1}), \sigma(X_{k-1}) U_k \rangle - \frac{1}{\sqrt{\Theta_n}} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1} D_{q^*}(X_{k-1}) \xrightarrow{\mathbf{L}^2} 0. \end{aligned} \quad (3.14)$$

On a supposé  $\lim_n \Theta_n = +\infty$  (car  $(\gamma, \theta)$  est moyennisant) et  $\xi = \lim_n \frac{\Upsilon_n^*}{\sqrt{\Theta_n}} \in ]0, +\infty[$ , donc  $\lim_n \Upsilon_n^* = +\infty$  et d'après le lemme V.1 le couple  $(\gamma, \eta \gamma^{\frac{q^*}{2}-1})$  est moyennisant. Ainsi,

$$\frac{1}{\Upsilon_n^*} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1} D_{q^*}(X_{k-1}) \xrightarrow{p.s} \int_{\mathbf{R}^d} D_{q^*} d\nu.$$

Le lemme V.2 permet de nouveau de conclure.

- On suppose maintenant  $\xi = +\infty$ . Cela implique  $\lim_n \Upsilon_n^* = +\infty$ , donc

$$\frac{1}{\Upsilon_n^*} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1} D_{q^*}(X_{k-1}) \xrightarrow{p.s} \int_{\mathbf{R}^d} D_{q^*} d\nu.$$

D'autre part, on a

$$\left\| \frac{1}{\Upsilon_n^*} \sum_{k=1}^n \sqrt{\theta_k} \langle \nabla \varphi(X_{k-1}), \sigma(X_{k-1}) U_k \rangle \right\|_2 \leq \frac{\sqrt{\Theta_n}}{\Upsilon_n^*} \sup_{k=0, \dots, n-1} \|(\sigma^* \nabla \varphi)(X_k)\|_2 \longrightarrow 0$$

Ainsi, en appliquant la proposition V.4 avec la suite  $(v_n)_{n \geq 1} = (H_n / \Upsilon_n^*)_{n \geq 1}$  ( $q^* \leq p$ ) on obtient

$$\frac{H_n}{\Upsilon_n^*} (\nu_n^\eta(f) - \nu(f)) - \frac{1}{\Upsilon_n^*} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1} D_{q^*}(X_{k-1}) \xrightarrow{\mathbf{L}^2} 0, \quad (3.15)$$

et la convergence en probabilité annoncée en (1.9) (en effet on a deux convergences de natures différentes : convergence presque sûre et convergence en moyenne quadratique, mais elles impliquent toutes les deux la convergence en probabilité).

Il reste à établir la convergence presque sûre dans (1.9). Pour cela on suppose la série  $\sum_{n \geq 1} \frac{\theta_n}{(\Upsilon_n^*)^2}$  convergente. On peut alors appliquer le lemme V.3 avec les suites  $(a_n)_{n \geq 1} = (\theta_n)_{n \geq 1}$  et  $(b_n)_{n \geq 1} = (\Upsilon_n^*)_{n \geq 1}$  pour obtenir

$$\frac{1}{\Upsilon_n^*} \sum_{k=1}^n \sqrt{\theta_k} \langle \nabla \varphi(X_{k-1}), \sigma(X_{k-1}) U_k \rangle \xrightarrow{p.s.} 0.$$

En combinant cette convergence presque sûre et la proposition V.4 appliqué avec la suite  $(v_n)_{n \geq 1} = (H_n / \Upsilon_n^*)_{n \geq 1}$  on obtient

$$\frac{H_n}{\Upsilon_n^*} (\nu_n^\eta(f) - \nu(f)) - \sum_{q=q^*}^p \frac{1}{\Upsilon_n^*} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_q(X_{k-1}) \xrightarrow{p.s.} 0. \quad (3.16)$$

Pour conclure, il suffit alors de montrer que pour tout  $q \geq q^* + 1$ ,

$$\frac{1}{\Upsilon_n^*} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_q(X_{k-1}) \xrightarrow{p.s.} 0.$$

Cela se fait de la même façon que pour prouver  $\tilde{R}_n^{(p)} \xrightarrow{p.s.} 0$  dans la démonstration de la proposition précédente. #

## 4. Loi du logarithme itéré

Dans cette section, on établit une loi du logarithme itéré (LLI) pour l'erreur  $(\nu_n^\eta(f) - \nu(f))_{n \geq 1}$  lorsque les pas et les poids choisis impliquent le TLC (1.7). On suppose que  $f$  est une fonction continue et que  $\varphi$  est une solution de l'équation de Poisson  $\mathcal{A}\varphi = -(f - \nu(f))$  telle que  $\varphi \in E_p^W$  et  $D^p\varphi$  Lipschitz. Les pas  $(\gamma_n)_{n \geq 1}$  et poids  $(\eta_n)_{n \geq 1}$  vérifient

$$\sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1} = \mathcal{O}(\sqrt{\Theta_n}), \quad \sum_{n \geq 1} \frac{1}{\sqrt{\Theta_n}} \left| \Delta \frac{\eta_n}{\gamma_n} \right| < +\infty \quad \text{et} \quad \sum_{n \geq 1} \frac{\eta_n^2}{\Theta_n} < +\infty \quad (4.1)$$

où  $q^* = \min_{3 \leq q \leq p} \{D_q \neq 0\} \wedge (p+1)$ . Dans la décomposition de l'erreur (voir proposition V.3), les termes importants pour obtenir une LLI sont les deux premiers. Notamment le second qui est un terme martingale.

Rappelons la loi du logarithme itéré pour les martingales de carré intégrable (voir théorème 4.7 dans [HH80]). On note  $\phi$  la fonction définie pour tout  $x > e$  par

$$\phi(x) = \sqrt{2x \log \log x},$$

et égale à 1 pour  $x \in [0, e]$ .

**Théorème V.4:**

Soit  $(M_n)_{n \geq 0}$  une martingale nulle en 0 de carré intégrable,  $(W_n)_{n \geq 1}$  une suite croissante de variables aléatoires strictement positives et  $(Z_n)_{n \geq 1}$  une suite de variables aléatoires positives. Pour tout  $n \geq 1$ , on note  $Y_n = \Delta M_n$  et on suppose que  $W_n$  et  $Z_n$  sont  $\mathcal{F}_{n-1}$ -mesurable. Si

$$\frac{1}{\phi(W_n^2)} \sum_{k=1}^n \left( Y_k \mathbf{1}_{\{|Y_k| > Z_k\}} - \mathbf{E} \left[ Y_k \mathbf{1}_{\{|Y_k| > Z_k\}} \middle| \mathcal{F}_{k-1} \right] \right) \xrightarrow{p.s.} 0, \quad (4.2)$$

$$\frac{1}{W_n^2} \sum_{k=1}^n \left( \mathbf{E} \left[ Y_k^2 \mathbf{1}_{\{|Y_k| \leq Z_k\}} \middle| \mathcal{F}_{k-1} \right] - \mathbf{E} \left[ Y_k \mathbf{1}_{\{|Y_k| \leq Z_k\}} \middle| \mathcal{F}_{k-1} \right]^2 \right) \xrightarrow{p.s.} 1, \quad (4.3)$$

$$\sum_{n \geq 1} \frac{1}{W_n^4} \mathbf{E} \left[ Y_n^4 \mathbf{1}_{\{|Y_n| \leq Z_n\}} \middle| \mathcal{F}_{n-1} \right] < +\infty \quad p.s. \quad (4.4)$$

et

$$\frac{W_n}{W_{n+1}} \xrightarrow{p.s.} 1 \quad \text{et} \quad W_n \xrightarrow{p.s.} +\infty, \quad (4.5)$$

alors

$$\overline{\lim}_n \frac{M_n}{\phi(W_n^2)} = 1 \quad p.s. \quad \text{et} \quad \underline{\lim}_n \frac{M_n}{\phi(W_n^2)} = -1 \quad p.s. \quad \star$$

La loi du logarithme itéré obtenue pour l'erreur  $(\nu_n^\eta(f) - \nu(f))$  est donnée par le théorème suivant. Comme attendu, la vitesse est en  $(H_n/\phi(\Theta_n))_{n \geq 1}$ , mais des conditions supplémentaires sont nécessaires sur les suites  $(\gamma_n)_{n \geq 1}$  et  $(\eta_n)_{n \geq 1}$  (cf. (4.6)).

**Théorème V.5:**

Soit un entier  $p \geq 2$  et  $\varphi$  une fonction de  $E_p^W$  solution de l'équation de Poisson telle que  $D^p \varphi$  est Lipschitz. On suppose qu'il existe  $\delta \in ]0, 2[$  tel que

$$|\sigma^* \nabla \varphi|^{2+\delta} = \mathcal{O}(W) \quad \text{et} \quad \mathbf{E} \left[ |U_1|^{2+\delta} \right] < +\infty,$$

et que les suites  $(\gamma_n)_{n \geq 1}$  et  $(\eta_n)_{n \geq 0}$  vérifient (4.1),

$$\sum_{n \geq 1} \left( \frac{\theta_n}{\Theta_n \ln \ln \Theta_n} \right)^{1+\frac{\delta}{2}} \gamma_n^{-\delta/2} < +\infty, \quad \text{et} \quad \left( \left( \frac{\theta_n}{\Theta_n \ln \ln \Theta_n} \right) \frac{1}{\gamma_n} \right)_{n \geq 1} \text{ décroît.} \quad (4.6)$$

De plus, on suppose que  $\gamma_n = \mathcal{O}((\ln \ln(\Theta_n))^{-2})$ . Alors, presque sûrement on a

$$\overline{\lim}_n \frac{H_n}{\phi(\Theta_n)} (\nu_n^\eta(f) - \nu(f)) = s_f, \quad \text{et} \quad \underline{\lim}_n \frac{H_n}{\phi(\Theta_n)} (\nu_n^\eta(f) - \nu(f)) = -s_f. \quad \star$$

La preuve repose sur le lemme suivant qui est simplement l'application du théorème V.4 au terme martingale apparaissant dans le développement de l'erreur.

**Lemme V.4:**

Soit  $g \in \mathcal{C}(\mathbf{R}^d; \mathbf{R}^m)$ . On suppose qu'il existe  $\delta \in ]0, 2[$  tel que

$$|g|^{2+\delta} = \mathcal{O}(W), \quad \text{et} \quad \mathbf{E} \left[ |U_1|^{2+\delta} \right] < +\infty,$$

et que la suite  $(\theta_n)_{n \geq 1}$  vérifie

$$\sum_{n \geq 1} \left( \frac{\theta_n}{\Theta_n \ln \ln(\Theta_n)} \right)^{1 + \frac{\delta}{2}} \ln \ln^\delta(\Theta_n) < +\infty, \quad \text{et} \quad \lim_n \frac{\theta_n \ln \ln(\Theta_n)}{\Theta_n} = 0. \quad (4.7)$$

Alors, presque sûrement on a

$$\overline{\lim}_n \frac{1}{\phi(\Theta_n)} \sum_{k=1}^n \sqrt{\theta_k} \langle g(X_{k-1}), U_k \rangle = \sigma_g \quad \text{et} \quad \underline{\lim}_n \frac{1}{\phi(\Theta_n)} \sum_{k=1}^n \sqrt{\theta_k} \langle g(X_{k-1}), U_k \rangle = -\sigma_g,$$

où  $\sigma_g = \sqrt{\int_{\mathbf{R}^d} |g|^2 d\nu}$ . ♦

Preuve. Ce lemme est une conséquence du théorème V.4 appliqué à la martingale  $(M_n)_{n \geq 0}$  nulle en 0 définie pour tout  $n \geq 1$  par

$$M_n = \sum_{k=1}^n \sqrt{\theta_k} \langle g(X_{k-1}), U_k \rangle,$$

avec les suites déterministes  $(W_n)_{n \geq 1}$  et  $(Z_n)_{n \geq 1}$  définies par

$$W_n^2 = \sigma_g^2 \Theta_n \quad \text{et} \quad Z_n = \sqrt{\frac{\sigma_g^2 \Theta_n}{\log \log \Theta_n}},$$

où  $\sigma_g^2 = \nu(|g|^2)$ .

- Montrons tout d'abord (4.5). On rappelle que le couple  $(\gamma, \theta)$  est moyennisant, donc  $\lim_n \Theta_n = +\infty$ . On en déduit avec la seconde condition de (4.7) que  $\lim_n \frac{\theta_n}{\Theta_n} = 0$ . Or,

$$\left( \frac{W_{n+1}}{W_n} \right)^2 = \frac{\Theta_{n+1}}{\Theta_n} = \frac{\Theta_{n+1}}{\Theta_{n+1} - \theta_{n+1}} = \frac{1}{1 - \frac{\theta_{n+1}}{\Theta_{n+1}}},$$

donc (4.5) est vérifiée.

- On montre (4.4) *i.e.* la convergence presque sûre de la série

$$\sum_{n \geq 1} \left( \frac{\theta_n^2}{\sigma_g^4 \Theta_n^2} \right) \mathbf{E} \left[ |\langle g(X_{n-1}), U_n \rangle|^4 \mathbf{1}_{\{|\sqrt{\theta_n} \langle g(X_{n-1}), U_n \rangle| \leq Z_n\}} \middle| \mathcal{F}_{n-1} \right].$$

Cette série est majorée par

$$\sum_{n \geq 1} \left( \frac{\theta_n^2}{\sigma_g^4 \Theta_n^2} \right) \left( \frac{Z_n}{\sqrt{\theta_n}} \right)^{2-\delta} |g(X_{n-1})|^{2+\delta} \mathbf{E} \left[ |U_1|^{2+\delta} \right],$$

et par définition de  $(Z_n)_{n \geq 0}$  la série converge presque sûrement si

$$\sum_{n \geq 1} \left( \frac{\theta_n^{1+\delta/2}}{\Theta_n^{1+\delta/2} \log \log^{1-\delta/2} \Theta_n} \right) |g(X_{n-1})|^{2+\delta} < +\infty \quad p.s.$$

car  $\mathbf{E} \left[ |U_1|^{2+\delta} \right] < +\infty$ . On a supposé  $|g|^{2+\delta} = \mathcal{O}(W)$  donc  $\sup_n \mathbf{E} \left[ |g(X_n)|^{2+\delta} \right] < +\infty$  et on

conclut aisément car  $\sum_{n \geq 1} \left( \frac{\theta_n^{1+\delta/2}}{\Theta_n^{1+\delta/2} \log \log^{1-\delta/2} \Theta_n} \right) < +\infty$  (hypothèse (4.7)).

- Passons à (4.3). En posant

$$h(x, y) = \mathbf{E} \left[ \langle g(x), U_1 \rangle^2 \mathbf{1}_{\{|\langle g(x), U_1 \rangle| > y\}} \right] + \mathbf{E} \left[ \langle g(x), U_1 \rangle \mathbf{1}_{\{|\langle g(x), U_1 \rangle| \leq y\}} \right]^2,$$

il faut montrer que

$$\frac{1}{\sigma_g^2 \Theta_n} \sum_{k=1}^n \theta_k \mathbf{E} \left[ \langle g(X_{k-1}), U_k \rangle^2 \mid \mathcal{F}_{k-1} \right] - \frac{1}{\sigma_g^2 \Theta_n} \sum_{k=1}^n \theta_k h \left( X_{k-1}, \frac{Z_k}{\sqrt{\theta_k}} \right) \xrightarrow{p.s.} 1.$$

Comme  $|g|^{2+\delta} = \mathcal{O}(W)$ , la fonction  $(x \mapsto \mathbf{E} [\langle g(x), U_1 \rangle^2] = |g(x)|^2)$  est négligeable devant  $W$  et, le couple  $(\gamma, \theta)$  étant moyennisant, on a

$$\frac{1}{\sigma_g^2 \Theta_n} \sum_{k=1}^n \theta_k \mathbf{E} \left[ \langle g(X_{k-1}), U_k \rangle^2 \mid \mathcal{F}_{k-1} \right] \xrightarrow{p.s.} \frac{1}{\sigma_g^2} \int_{\mathbf{R}^d} \mathbf{E} \left[ \langle g(x), U_1 \rangle^2 \right] \nu(dx). \quad (4.8)$$

Or  $\sigma_g^2 = \nu(|g|^2)$  et  $U_1$  est centrée et normée (i.e. de matrice de covariance  $\text{Id}_m$ ) donc la limite dans (4.8) est égale à 1. Il faut donc montrer que

$$\frac{1}{\sigma_f^2 \Theta_n} \sum_{k=1}^n \theta_k h \left( X_{k-1}, \frac{Z_k}{\sqrt{\theta_k}} \right) \xrightarrow{p.s.} 0. \quad (4.9)$$

Comme  $\mathbf{E}[U_1] = 0$ , on a  $\mathbf{E} [\langle g(x), U_1 \rangle \mathbf{1}_{\{|\langle g(x), U_1 \rangle| \leq y\}}] = -\mathbf{E} [\langle g(x), U_1 \rangle \mathbf{1}_{\{|\langle g(x), U_1 \rangle| > y\}}]$ , et par Cauchy-Schwartz on obtient

$$\forall (x, y) \in \mathbf{R}^d \times \mathbf{R}^m, \quad h(x, y) \leq 2 \mathbf{E} \left[ \langle g(x), U_1 \rangle^2 \mathbf{1}_{\{|\langle g(x), U_1 \rangle| > y\}} \right].$$

Soit  $A > 0$  fixé. On a alors

$$\begin{aligned} \frac{1}{\Theta_n} \sum_{k=1}^n \theta_k h \left( X_{k-1}, \frac{Z_k}{\sqrt{\theta_k}} \right) &\leq \frac{1}{\Theta_n} \sum_{k=1}^n \theta_k h \left( X_{k-1}, \frac{Z_k}{\sqrt{\theta_k}} \right) \mathbf{1}_{\{|X_{k-1}| \leq A\}} \\ &\quad + \frac{1}{\Theta_n} \sum_{k=1}^n \theta_k h \left( X_{k-1}, \frac{Z_k}{\sqrt{\theta_k}} \right) \mathbf{1}_{\{|X_{k-1}| > A\}}, \\ &\leq 2 \left( \sup_{|x| \leq A} |g(x)|^2 \right) \left( \frac{1}{\Theta_n} \sum_{k=1}^n \theta_k \sup_{|x| \leq A} \mathbf{E} \left[ |U_1|^2 \mathbf{1}_{\{|\langle g(x), U_1 \rangle| > \frac{Z_k}{\sqrt{\theta_k}}\}} \right] \right) \\ &\quad + 2 \sup_{|x| > A} \left( \frac{|g(x)|^2}{W(x)} \right) \nu_n^\theta(W). \end{aligned}$$

Comme  $\lim_n \frac{Z_n}{\sqrt{\theta_n}} = +\infty$ , il est clair que

$$\lim_n \sup_{|x| \leq A} \mathbf{E} \left[ |U_1|^2 \mathbf{1}_{\{|\langle g(x), U_1 \rangle| > \frac{Z_n}{\sqrt{\theta_n}}\}} \right] = 0,$$

et par Césaro on obtient alors

$$0 \leq \overline{\lim}_n \frac{1}{\Theta_n} \sum_{k=1}^n \theta_k h \left( X_{k-1}, \frac{Z_k}{\sqrt{\theta_k}} \right) \leq 2 \sup_{|x|>A} \left( \frac{|g(x)|^2}{W(x)} \right) \sup_n \nu_n^\theta(W).$$

En faisant tendre  $A$  vers l'infini on obtient (4.9) car  $|g|^2 = o(W)$ .

- Il reste à montrer (4.2). Pour cela on introduit la martingale  $(N_n)_{n \geq 1}$  nulle en 0 définie par

$$N_n = \sum_{k=1}^n \frac{\sqrt{\theta_k}}{\phi(\sigma_g^2 \Theta_k)} \langle g(X_{k-1}), \varepsilon_k \rangle$$

avec

$$\varepsilon_k = U_k \mathbf{1}_{\{|\sqrt{\theta_k} \langle g(X_{k-1}), U_k \rangle| > Z_k\}} - \mathbf{E} \left[ U_k \mathbf{1}_{\{|\sqrt{\theta_k} \langle g(X_{k-1}), U_k \rangle| > Z_k\}} \middle| \mathcal{F}_{k-1} \right].$$

On a  $\mathbf{E} \left[ |\varepsilon_k|^2 \middle| \mathcal{F}_{k-1} \right] \leq \mathbf{E} \left[ |U_k|^2 \mathbf{1}_{\{|\sqrt{\theta_k} \langle g(X_{k-1}), U_k \rangle| > Z_k\}} \middle| \mathcal{F}_{k-1} \right]$  donc le crochet de  $(N_n)_{n \geq 1}$  vérifie

$$\begin{aligned} \langle N \rangle_n &\leq \sum_{k=1}^n \frac{\theta_k}{\phi^2(\sigma_g^2 \Theta_k)} |g(X_{k-1})|^2 \mathbf{E} \left[ |U_k|^{2+\delta} |U_k|^{-\delta} \mathbf{1}_{\{|\sqrt{\theta_k} \langle g(X_{k-1}), U_k \rangle| > Z_k\}} \middle| \mathcal{F}_{k-1} \right], \\ &\leq \mathbf{E} \left[ |U_1|^{2+\delta} \right] \sum_{k=1}^n \frac{\theta_k}{\phi^2(\sigma_g^2 \Theta_k)} |g(X_{k-1})|^{2+\delta} \theta_k^{\delta/2} Z_k^{-\delta}. \end{aligned}$$

Comme  $\phi^2(x) = 2x \ln \ln x$  et que  $Z_k^{-\delta} = (\sigma_g^2 \Theta_k)^{-\delta/2} \ln \ln^{\frac{\delta}{2}}(\Theta_k)$ , on a

$$\langle N \rangle_\infty \leq \mathbf{E} \left[ |U_1|^{2+\delta} \right] \sum_{n \geq 1} \frac{\theta_n^{1+\frac{\delta}{2}}}{2(\sigma_g^2 \Theta_n)^{1+\frac{\delta}{2}} \ln \ln^{1-\frac{\delta}{2}}(\Theta_n)} |g(X_{n-1})|^{2+\delta}.$$

On a supposé  $\sum_{n \geq 1} \left( \frac{\theta_n}{\Theta_n \ln \ln^{\frac{2-\delta}{2+\delta}}(\Theta_n)} \right)^{1+\delta/2} < +\infty$  et  $|g|^{2+\delta} = \mathcal{O}(W)$  donc  $\langle N \rangle_\infty < +\infty$  p.s..

Par le lemme de Kronecker, on obtient (4.2). #

**Preuve** (théorème V.5). La preuve est très proche de celle de la proposition V.4. Toutefois, on ne peut appliquer celle-ci directement car on veut obtenir la convergence presque sûre vers 0 du premier terme  $\frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (\Delta \varphi(X_{k-1}))$  apparaissant dans la décomposition de la proposition V.3. Les hypothèses sur les poids dans la proposition V.4 sont trop fortes et peuvent ici être relaxées car les hypothèses sur la domination de  $\varphi$  et les moments de  $U_1$  sont plus fortes.

- Montrons que

$$\frac{1}{\phi(H_n)} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (\Delta \varphi(X_k)) \xrightarrow{p.s.} 0. \quad (4.10)$$

Par la transformation d'Abel, il suffit de montrer que  $\frac{\eta_n}{\phi(H_n) \gamma_n} \varphi(X_n)$  converge vers 0 presque sûrement. En effet, on a supposé  $\sum_{n \geq 1} \frac{1}{\sqrt{\Theta_n}} \left| \Delta \frac{\eta_n}{\gamma_n} \right| < +\infty$  et comme  $\phi(\Theta_n) \geq \sqrt{\Theta_n}$  alors par Kronecker  $\lim_n \frac{1}{\phi(H_n)} \sum_{k=1}^n \left( \Delta \frac{\eta_k}{\gamma_k} \right) \varphi(X_{k-1}) = 0$  p.s.

On remarque que

$$\frac{\eta_n}{\phi(\Theta_n)\gamma_n} = \sqrt{\frac{\theta_n}{\phi^2(\Theta_n)\gamma_n}} = \sqrt{\frac{\theta_n}{2\Theta_n \ln \ln \Theta_n} \frac{1}{\gamma_n}} \quad (4.11)$$

et comme  $|\varphi|^{2+\delta} = \mathcal{O}(W)$ , la proposition V.1b. appliquée avec la suite  $\left( \left( \frac{\theta_n}{\phi^2(\Theta_n)\gamma_n} \right)^{1+\frac{\delta}{2}} \right)_{n \geq 1}$  donne

$$\left( \frac{\theta_n}{\phi^2(\Theta_n)\gamma_n} \right)^{1+\frac{\delta}{2}} |\varphi(X_n)|^{2+\delta} \xrightarrow{p.s.} 0. \quad (4.12)$$

- Sous l'hypothèse  $\gamma_n = \mathcal{O}(\ln \ln^{-2}(\Theta_n))$  il est aisé de voir que les conditions (4.6) impliquent les conditions (4.7). Ainsi on peut appliquer le lemme V.4 qui donne

$$\overline{\lim} \frac{1}{\phi(\Theta_n)} \sum_{k=1}^n \frac{\eta_k}{\sqrt{\gamma_k}} \langle (\sigma^* \nabla \varphi)(X_{k-1}), U_k \rangle = s_f \quad p.s.$$

avec  $s_f = \sqrt{\int_{\mathbf{R}^d} |\sigma^* \nabla \varphi|^2 d\nu}$ . De même pour la limite inférieure qui est égale à  $s_f$  presque sûrement.

- Tous les autres termes se gèrent de la même façon que dans la proposition V.4. Comme  $\phi(\Theta_n) \geq \sqrt{\Theta_n}$ , les conditions (4.1) impliquent les conditions (3.7) prises avec  $v_n = \frac{H_n}{\phi(\Theta_n)}$ . #

## 5. Simulations numériques

Dans cette section on se propose de valider numériquement nos résultats sur les vitesses de convergence des moyennes empiriques pondérées. On suppose dans cette section que le bruit blanc  $(U_n)_{n \geq 1}$  est symétrique et admet des moments à tout ordre. Ainsi, pour tout  $p \geq 0$  on a  $\mathbf{E} \left[ U_1^{\otimes 2p+1} \right] = 0$  si bien que pour tout  $q \geq 1$ ,  $D_{2q+1} = 0$ . On discrétise la diffusion en utilisant une suite  $(\gamma_n)_{n \geq 1}$  de la forme  $\gamma_n = \gamma_0 n^{-r}$  avec  $r \in ]0, 1]$  et  $\gamma_0 > 0$ , et on pondère les mesures empiriques à l'aide d'une suite  $(\eta_n)_{n \geq 1}$  de la forme  $\eta_n = \eta_0 n^{-s}$  avec  $s \in ]-\infty, 1]$  et  $\eta_0 > 0$ . On suppose que les couples  $(\gamma, \eta)$  et  $(\gamma, \theta)$  sont moyennisants, c'est à dire que

$$\left( 0 < r < 1 \text{ et } s \leq \frac{1+r}{2} \right) \quad \text{ou} \quad (r = s = 1), \quad (5.1)$$

(cf. proposition V.2).

Sous ces hypothèses, le théorème V.3 s'énonce ainsi :

### Proposition V.5:

Soit un entier  $p \geq 2$  tel que  $U_1 \in \mathbf{L}^{2(p+1)}$ , et  $\varphi \in E_p^W$  solution de l'équation de Poisson telle que  $D^p \varphi$  est Lipschitz. On définit  $q^*$  par

$$q^* = \min_{q \in \{4, \dots, p\}} \{D_q \neq 0\} \wedge (p+1). \quad (5.2)$$

Alors,

– si  $r = s = 1$

$$\sqrt{\gamma_0} \sqrt{\ln(n)} (\nu_n^\eta(f) - \nu(f)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, s_f^2),$$

– si  $\frac{1}{q^*-1} < r < 1$

$$c_{\mathcal{L}} n^{\frac{1-r}{2}} (\nu_n^\eta(f) - \nu(f)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, s_f^2),$$

– si  $r = \frac{1}{q^*-1}$  et  $q^* \leq p$

$$c_{\mathcal{L}} n^{\frac{1-r}{2}} (\nu_n^\eta(f) - \nu(f)) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\frac{c_{\mathcal{L}}}{c_{p.s.}} m_{f,q^*}, s_f^2\right),$$

– si  $0 < r < \frac{1}{q^*-1}$  et  $q^* \leq p$

$$\begin{aligned} c_{p.s.} n^{r\left(\frac{q^*}{2}-1\right)} (\nu_n^\eta(f) - \nu(f)) &\xrightarrow{\mathbf{P}} m_{f,q^*} && \text{si } s < -r \left(\frac{q^*}{2} - 1\right), \\ c_{p.s.} n^{r\left(\frac{q^*}{2}-1\right)} (\nu_n^\eta(f) - \nu(f)) &\xrightarrow{p.s.} m_{f,q^*} && \text{si } s \geq -r \left(\frac{q^*}{2} - 1\right), \end{aligned}$$

avec

$$c_{\mathcal{L}} = \sqrt{\gamma_0} \frac{\sqrt{1-2s+r}}{1-s} \quad \text{et} \quad c_{p.s.} = \gamma_0^{1-\frac{q^*}{2}} \left(1 - \frac{r\left(\frac{q^*}{2}-1\right)}{1-s}\right). \quad (5.3) \quad \blacklozenge$$

Remarque: La définition de  $q^*$  en (5.2) est différente de celle utilisée dans le théorème V.3 car le bruit blanc  $(U_n)_{n \geq 1}$  est maintenant supposé symétrique. \*

Le cas  $r = s = 1$  peut apparaître comme anecdotique à la vue de cette proposition, mais c'est en fait le cadre naturel du TLC presque sûre pour des variables aléatoires i.i.d. (cf section 5. de [LP02]). Ainsi, la proposition ci-dessus fournit un TLC pour le TLC presque sûre à la vitesse  $\sqrt{\ln(n)}$ .

Preuve. On montre cette proposition pour  $r < 1$  (et donc  $s < 1$ ), le cas  $r = s = 1$  étant similaire. En utilisant les équivalences suivantes

$$H_n \sim \eta_0 \frac{n^{1-s}}{1-s}, \quad \Theta_n \sim \frac{\eta_0^2}{\gamma_0} \frac{n^{1-2s+r}}{1-2s+r}, \quad \text{et} \quad \Upsilon_n^* \sim \eta_0 \gamma_0^{\frac{q^*}{2}-1} \frac{n^{1-s-r\left(\frac{q^*}{2}-1\right)}}{1-s-r\left(\frac{q^*}{2}-1\right)},$$

on en déduit que

$$\frac{\Upsilon_n^*}{\sqrt{\Theta_n}} \sim \gamma_0^{\frac{q^*-1}{2}} \frac{\sqrt{1-2s+r}}{1-s-r\left(\frac{q^*}{2}-1\right)} n^{\frac{1}{2}-r\left(\frac{q^*-1}{2}\right)}.$$

Ainsi  $\xi = \lim_n \frac{\Upsilon_n^*}{\sqrt{\Theta_n}} < +\infty$  si  $r \geq \frac{1}{q^*-1}$  et  $\xi = +\infty$  si  $r < \frac{1}{q^*-1}$ . Vérifions les hypothèses du théorème V.3.

- Dans le cas  $r \geq \frac{1}{q^*-1}$ , on a

$$\begin{aligned} \sum_{n \geq 1} \frac{1}{\sqrt{\Theta_n}} \left| \Delta \frac{\eta_n}{\gamma_n} \right| < +\infty &\Leftrightarrow \sum_{n \geq 1} \frac{1}{n^{\frac{1}{2}(1-2s+r)}} |r-s| n^{-s+r-1} < +\infty \\ &\Leftrightarrow (r = s \text{ ou } r < 1), \end{aligned}$$

et

$$\sum_{n \geq 1} \frac{\eta_n^2}{\Theta_n} < +\infty \Leftrightarrow \sum_{n \geq 1} n^{-2s-(1-2s+r)} < +\infty \Leftrightarrow r > 0,$$

donc les conditions sont clairement vérifiées et on a le TLC annoncé à la vitesse  $\frac{H_n}{\sqrt{\Theta_n}} \sim c_{\mathcal{L}} n^{\frac{1-r}{2}}$ .

- Supposons  $r < \frac{1}{q^*-1}$ . On a

$$\sum_{n \geq 1} \frac{1}{\Upsilon_n^*} \left| \Delta \frac{\eta_n}{\gamma_n} \right| < +\infty \Leftrightarrow \sum_{n \geq 1} \frac{n^{-s+r-1}}{n^{1-s-r(\frac{q^*}{2}-1)}} < +\infty \Leftrightarrow r < \frac{2}{q^*},$$

et

$$\sum_{n \geq 1} \frac{\theta_n}{(\Upsilon_n^*)^2} < +\infty \Leftrightarrow \sum_{n \geq 1} \frac{n^{-2s+r}}{n^{2-2s-r(q^*-2)}} < +\infty \Leftrightarrow r < \frac{1}{q^*-1},$$

donc les conditions du théorème V.3 sont satisfaites et on a la convergence en probabilité pour tout  $s \leq \frac{1+r}{2}$ .

Pour la convergence presque sûre, il faut montrer la décroissance de la suite  $\left( \frac{\eta_n}{\Upsilon_n^* \gamma_n} \right)_{n \geq 1}$ . Comme  $s \geq -r \left( \frac{q^*}{2} - 1 \right)$  alors  $\Upsilon_n^* = \sum_{k=1}^n k^{-\beta}$  avec  $\beta = s + r \left( \frac{q^*}{2} - 1 \right) \geq 0$ . Ainsi en utilisant une comparaison série-intégrale on a

$$\forall n \geq 1, \quad \frac{(n+1)^{1-\beta} - 1}{1-\beta} \leq \Upsilon_n^* \leq \frac{n^{1-\beta}}{1-\beta},$$

et par le théorème des valeurs intermédiaires et la décroissance de  $(x \mapsto x^{-\beta})$  on obtient

$$\forall n \geq 1, \quad (n+1)^{-\beta} n \leq \Upsilon_n^* \leq \frac{n^{1-\beta}}{1-\beta},$$

Donc, pour tout  $n \geq 1$

$$\frac{\Upsilon_{n+1}^* \gamma_{n+1} \eta_n}{\Upsilon_n^* \gamma_n \eta_{n+1}} \geq \left( 1 + \frac{1}{n} \right)^{-r \frac{q^*}{2}} (1-\beta).$$

On en conclut que la suite  $\left( \frac{\eta_n}{\Upsilon_n^* \gamma_n} \right)_{n \geq 1}$  est décroissante à partir du rang  $n_0 = \frac{1}{\left( \frac{1}{1-\beta} \right)^{-r \frac{q^*}{2}} - 1}$ ,  
et donc la convergence presque sûre. #

## 5.1. Choix des paramètres et commentaires

D'après la proposition V.5 on sait que l'erreur  $(\nu_n^\eta(f) - \nu(f))_{n \geq 1}$  vérifie un TLC centré de vitesse  $n^{\frac{1-r}{2}}$  lorsque  $r > \frac{1}{q^*-1}$ , et que cette erreur converge *p.s.* à la vitesse  $n^{r(\frac{q^*}{2}-1)}$  vers  $m_{f,q^*}$  lorsque  $r < \frac{1}{q^*-1}$ . De plus, il est intéressant de remarquer que l'ordre de la vitesse ne dépend que de  $r$  et qu'il est optimal pour  $r = \frac{1}{q^*-1}$ . Ce résultat est représenté par la figure V.1 pour deux valeurs de  $q^*$  : 4 et 6.

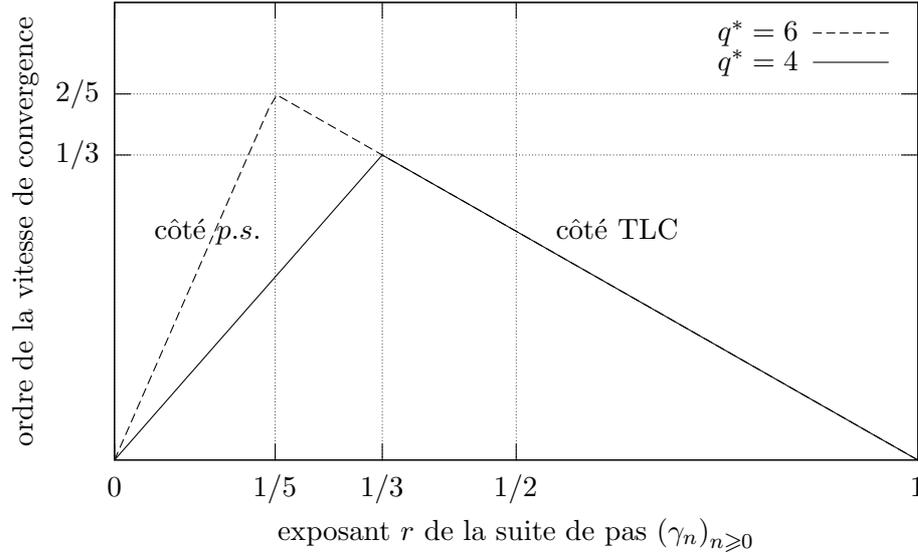


Figure V.1 – Ordre de la vitesse de convergence en fonction du paramètre  $r$ .

Les paramètres  $\gamma_0$  et  $s$  interviennent uniquement dans les constantes  $c_{\mathcal{L}}$  et  $c_{p.s.}$ . Il n'est pas étonnant de voir que la constante  $\eta_0$  n'intervient pas dans les vitesses car elle n'intervient même pas dans l'algorithme. Il disparaît par homogénéité des mesures empiriques pondérées.

Remarque: Lorsque  $r$  est fixé, les constantes  $c_{\mathcal{L}}$  et  $c_{p.s.}$  sont des fonctions de  $\gamma_0$  et  $s$  définies sur  $]0, +\infty[ \times ]-\infty, \frac{1+r}{2}[$  par

$$c_{\mathcal{L}}(\gamma_0, s) = \sqrt{\gamma_0} \frac{\sqrt{1-2s+r}}{1-s} \quad \text{et} \quad c_{p.s.}(\gamma_0, s) = \gamma_0^{1-\frac{q^*}{2}} \left( 1 - \frac{r(\frac{q^*}{2}-1)}{1-s} \right). \quad *$$

À la vue de ces résultats, une question naturelle est : Comment choisir  $\gamma_0$  et  $s$  lorsque l'on se fixe  $r$  (c'est à dire l'ordre de la vitesse) ?

Cas  $r = \frac{1}{q^*-1}$

Dans le cas où la vitesse est optimale, l'erreur vérifie un TLC décentré de vitesse  $n^{\frac{q^*-2}{2(q^*-1)}}$ . Ainsi, pour  $n$  suffisamment grand on peut supposer que

$$n^{\frac{q^*-2}{2(q^*-1)}} \left( \nu_n^\eta(f) - \nu(f) \right) \stackrel{\mathcal{L}}{\approx} \mathcal{N} \left( \frac{m_{f,q^*}}{c_{p.s.}}, \frac{s_f^2}{c_{\mathcal{L}}^2} \right), \quad (5.4)$$

dans le sens où, pour  $a$  fixé et  $n$  grand,

$$\mathbf{P} \left[ n^{\frac{q^*-2}{2(q^*-1)}} |\nu_n^\eta(f) - \nu(f)| \leq a \right] \simeq \mathbf{P} \left[ \left| \frac{sf}{c_{\mathcal{L}}} G + \frac{m_{f,q^*}}{c_{p.s.}} \right| \leq a \right]$$

avec  $G$  de loi  $\mathcal{N}(0, 1)$ . On pose  $p \in [0, 1]$  et on cherche à minimiser en  $\gamma_0$  et  $s$  la taille de l'intervalle de confiance de niveau  $p$  dans lequel se trouve l'erreur  $\nu_n^\eta(f) - \nu(f)$ . C'est-à-dire, on cherche  $\gamma_0$  et  $s$  minimisant  $a(\gamma, s)$  lorsque  $a(\gamma, s)$  vérifie

$$\mathbf{P} \left[ \left| \frac{sf}{c_{\mathcal{L}}(\gamma, s)} G + \frac{m_{f,q^*}}{c_{p.s.}(\gamma, s)} \right| \leq a(\gamma, s) \right] = p.$$

Lorsque  $m_{f,q^*} = 0$  c'est la démarche classique pour optimiser les paramètres en fonction de la variance du TLC, mais lorsqu'on a un biais il est difficile de trouver par cette méthode les paramètres optimaux  $\gamma_0$  et  $s$ .

*Remarque: L'hypothèse (5.4) est justifiée par le fait que le TLC que l'on obtient découle du TLC pour les martingales et que les hypothèses de Berry-Essen pour les martingales sont vérifiées. De plus, les autres termes apparaissant dans le développement de l'erreur sont de l'ordre de  $n$  et deviennent assez vite négligeable devant le biais.* \*

Raisonnons en norme  $\mathbf{L}^2$  pour déterminer les paramètres  $\gamma_0$  et  $s$ . D'après la proposition V.4 appliquée avec la suite  $v_n = n^{\frac{q^*-2}{2(q^*-1)}}$ , on sait que

$$\left\| n^{\frac{q^*-2}{2(q^*-1)}} (\nu_n^\eta(f) - \nu(f)) - M_n - S_n \right\|_2 \longrightarrow 0$$

avec  $M_n = \frac{v_n}{H_n} \sum_{k=1}^n \sqrt{\theta_k} \langle \nabla \varphi(X_{k-1}), \sigma(X_{k-1}) U_k \rangle$  et  $S_n = \sum_{q=q^*}^p \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_q(X_{k-1})$ . Ainsi, on a

$$n^{\frac{q^*-2}{2(q^*-1)}} \|\nu_n^\eta(f) - \nu(f)\|_2 = \|M_n + S_n\|_2 + o(1).$$

On a déjà prouvé que

$$M_n \xrightarrow{\mathcal{L}} \mathcal{N} \left( 0, \frac{s_f^2}{c_{\mathcal{L}}^2} \right) \quad \text{et} \quad S_n \xrightarrow{p.s.} \frac{m_{f,q^*}}{c_{p.s.}}.$$

On rappelle que dans ce cadre optimal, on a  $\sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1} \sim \frac{H_n}{v_n}$  et  $\sum_{k=1}^n \theta_k \sim \frac{H_n^2}{v_n^2}$ . Si on suppose  $\varphi$  à support compact, alors les fonctions  $\nabla \varphi$  et  $D_q$  pour  $q = q^*, \dots, p$  sont à support compact et par convergence dominée on en déduit que

$$M_n \xrightarrow{\mathbf{L}^2} \frac{sf}{c_{\mathcal{L}}} \quad \text{et} \quad S_n \xrightarrow{\mathbf{L}^2} \frac{m_{f,q^*}}{c_{p.s.}}.$$

Or  $\|M_n + S_n\|_2^2 = \|M_n\|_2^2 + \|S_n\|_2^2 + 2\mathbf{E}[M_n S_n]$  et  $2\mathbf{E}[M_n S_n] \leq \|M_n\|_2 \left\| S_n - \frac{m_{f,q^*}}{c_{p.s.}} \right\|_2$  car  $\mathbf{E}[M_n] = 0$ . Ainsi, on a  $\|\nu_n^\eta(f) - \nu(f)\|_2 = n^{-\frac{q^*-2}{2(q^*-1)}} \sqrt{\|M_n\|_2 + \|S_n\|_2} + o(1)$  et

$$\|\nu_n^\eta(f) - \nu(f)\|_2 \sim n^{-\frac{q^*-2}{2(q^*-1)}} \sqrt{\left(\frac{s_f}{c_{\mathcal{L}}}\right)^2 + \left(\frac{m_{f,q^*}}{c_{p.s.}}\right)^2}.$$

Trouver les paramètres  $\gamma_0$  et  $s$  qui minimisent l'erreur  $\nu_n^\eta(f) - \nu(f)$  en norme  $\mathbf{L}^2$  revient donc à trouver les minimums de la fonction  $g(\gamma, s)$  définie comme suit

$$g : ]0; +\infty[ \times ]-\infty, \frac{1+r}{2}[ \rightarrow \mathbf{R}^+ \\ (\gamma, s) \mapsto \frac{s_f^2}{c_{\mathcal{L}}^2(\gamma, s)} + \frac{m_{f,q^*}^2}{c_{p.s.}^2(\gamma, s)}.$$

Le gradient de  $g$ , qui s'écrit

$$\nabla g(\gamma, s) = \begin{pmatrix} -\frac{1}{\gamma^2} \frac{(1-s)^2}{1-2s+r} s_f^2 + (q^*-2)\gamma^{q^*-3} \left( \frac{1-s}{1-s-r(\frac{q^*}{2}-1)} \right)^2 m_{f,q^*}^2 \\ -\frac{1}{\gamma} \frac{2(1-s)(r-s)}{(1-2s+r)^2} s_f^2 + \gamma^{q^*-2} \frac{r(q^*-2)(1-s)}{(1-s-r(\frac{q^*}{2}-1))^3} m_{f,q^*}^2 \end{pmatrix},$$

s'annule en un unique point  $(\gamma_0, s^*)$  de  $]0; +\infty[ \times ]-\infty, \frac{1+r}{2}[$  (cf. preuve en annexe),

$$\begin{cases} s^* = 0, \\ \gamma_0 = \left( 2\sqrt{1 - \frac{1}{q^*} \frac{|m_{f,q^*}|}{s_f}} \right)^{-\frac{2}{q^*-1}}. \end{cases} \quad (5.5)$$

Et on vérifie que ce point  $(\gamma_0, s^*)$  est un minimum global de  $g$  sur  $]0, +\infty[ \times ]-\infty, \frac{1+r}{2}[$ . Ainsi lorsque  $r = \frac{1}{q^*-1}$  il faut choisir  $s^*$  et  $\gamma_0$  définis en (5.5) pour avoir une erreur minimale dans  $\mathbf{L}^2$ . Il est intéressant de remarquer que  $s^* = 0$ , c'est à dire que la suite de poids optimale à considérer est la suite constante égale à 1,  $\eta_n = 1$ . D'autre part, notons que  $\gamma_0$  dépend de  $q^*$  mais aussi de  $m_{f,q^*}$  et  $s_f$ . Ces deux quantités  $m_{f,q^*}$  et  $s_f$  sont *a priori* inconnues et dépendent de la diffusion et de la fonction  $f$  à intégrer. De plus, il est difficile d'approcher numériquement ces constantes. On reparlera de ce problème dans le chapitre VIII.

**Remarque:** Si  $q^* = 4$  alors  $r = 1/3$ ,  $s^* = 0$  et  $\gamma_0 = \left( \sqrt{3} \frac{|m_{f,q^*}|}{s_f} \right)^{-2/3}$ . Si  $q^* = 6$  alors  $r = 1/5$ ,

$$s^* = 0 \text{ et } \gamma_0 = \left( 2\sqrt{\frac{5}{6}} \frac{|m_{f,q^*}|}{s_f} \right)^{-1/2}. \quad *$$

Dans le cas où la vitesse est optimale en ordre ( $r = \frac{1}{q^*-1}$ ),  $\gamma_0$  est le seul paramètre à déterminer pour minimiser l'erreur quadratique. De plus, d'après le rôle de ce paramètre dans le TLC, on sait que choisir  $\gamma_0$  grand réduit la variance mais décentre significativement le TLC, et choisir  $\gamma_0$  petit augmente la variance.

Cas  $r > \frac{1}{q^*-1}$

Dans ce cas, l'erreur vérifie un TLC centré. On pourrait donc raisonner de la même façon que ce qui suit (5.4). Cela nous conduirait à chercher à réduire la variance  $\frac{s_f^2}{c_{\mathcal{L}}^2}$  c'est à dire à maximiser  $c_{\mathcal{L}}(\gamma, s)$ . Cependant si  $r$  est proche de  $\frac{1}{q^*-1}$ , le biais s'annule très lentement et on ne peut pas considérer l'erreur comme une gaussienne centrée, même pour  $n$  assez grand. Dans cette situation il est difficile de déterminer des paramètres  $\gamma$  et  $s$  optimaux et ne dépendant pas de  $n$ .

Si  $r$  est suffisamment grand devant  $\frac{1}{q^*-1}$ , le biais s'annule plus vite et on peut alors chercher à minimiser la variance. On rappelle que

$$c_{\mathcal{L}}(\gamma, s) = \sqrt{\gamma} \frac{\sqrt{1 - 2s + r}}{1 - s}.$$

Sur  $]0, +\infty[ \times ]0, \frac{1+r}{2}[$ , le supremum de  $c_{\mathcal{L}}(\gamma, s)$  est atteint pour  $s = r$  et lorsque  $\gamma$  tend vers  $+\infty$ . Ainsi, on prendra  $s^* = r$  et  $\gamma_0 > \left(2\sqrt{1 - \frac{1}{q^*} \left|\frac{m_{f,q^*}}{s_f}\right|}\right)^{-\frac{2}{q^*-1}}$ . Attention tout de même de ne pas prendre  $\gamma_0$  trop grand, car le biais mettra alors d'autant plus de temps à s'annuler.

Cas  $r < \frac{1}{q^*-1}$

Dans ce cas, on a une convergence presque sûre de l'erreur à la vitesse  $n^r$  vers  $\frac{m_{f,q^*}}{c_{p.s.}(\gamma,s)}$ . Minimiser l'erreur revient alors à maximiser  $c_{p.s.}(\gamma, s)$  défini sur  $]0, +\infty[ \times ]-\infty, \frac{1+r}{2}[$  par

$$c_{p.s.}(\gamma, s) = \gamma^{1-\frac{q^*}{2}} \left(1 - \frac{r\left(\frac{q^*}{2} - 1\right)}{1 - s}\right).$$

Le suprémum est atteint lorsque  $\gamma$  tend vers 0 et  $s$  tend vers  $-\infty$ . Cependant, de la même façon que dans le cas  $r > \frac{1}{q^*-1}$ , si  $r$  est proche de  $\frac{1}{q^*-1}$  la variance due au TLC est encore présente. On ne peut pas considérer qu'elle est négligeable donc il faut choisir  $\gamma_0$  et  $s$  aussi en fonction de  $c_{\mathcal{L}}(\gamma, s)$ .

Pour conclure, on donne une façon de choisir  $s$  en fonction de  $r$  lorsque  $q^* = 4$  (cf. figure V.2). Ce choix repose sur les commentaires précédents. On définit  $s$  de la façon suivante

$$s : ]0, 1] \rightarrow ]-\infty, 1],$$

$$r \mapsto s(r) = \begin{cases} r & \text{si } r \geq \frac{1}{2}, \\ 3r - 1 & \text{si } r \in \left[\frac{1}{3}, \frac{1}{2}\right[, \\ 1 - \frac{1}{3r} & \text{si } r < \frac{1}{3}. \end{cases}$$

## 5.2. Tests numériques

Le but de cette sous-section est de tester numériquement notre algorithme et de valider nos résultats théoriques obtenus sur les vitesses de convergence. Pour cela, on considère un oscillateur linéaire de dimension 2 soumis à une perturbation aléatoire externe. L'équation

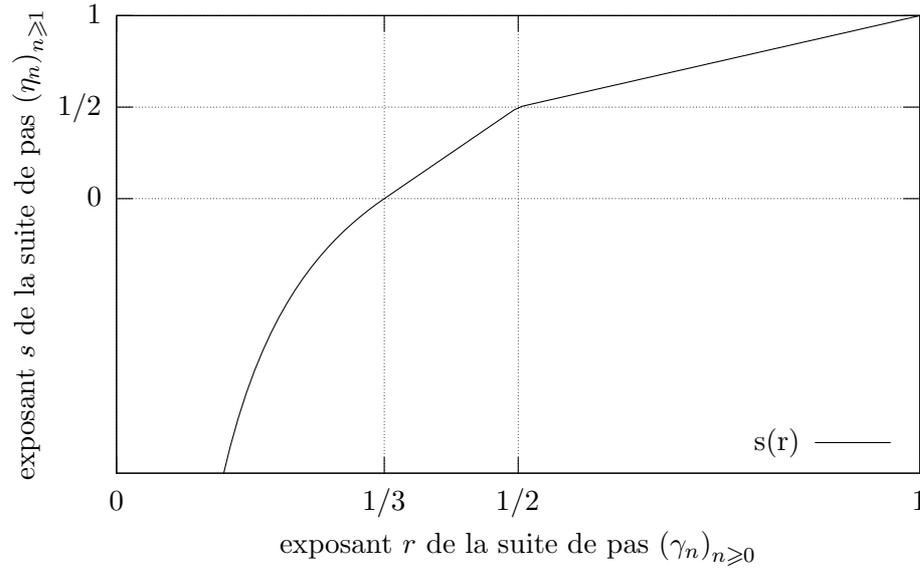


Figure V.2 – Choix de la suite de poids en fonction de la suite de pas dans le cas  $q^* = 4$ .

d'un tel système s'écrit

$$M_0 \frac{d^2 q}{dt^2}(t) + D_0 \frac{dq}{dt}(t) + K_0 q(t) = S_0 \frac{dB(t)}{dt}, \quad q(0) \in \mathbf{R}^2.$$

où  $(B_t)_{t \geq 0}$  est un  $\mathbf{R}^2$ -mouvement brownien,  $M_0$  est la matrice de la masse du système et  $K_0$  est la matrice de rigidité qui sont toutes deux des matrices réelles  $2 \times 2$  constantes et symétriques définies positive. Les matrices  $D_0$  et  $S_0$  sont des matrices réelles  $2 \times 2$  constantes et symétriques.

On considère le système d'ordre 1 équivalent à cette équation, c'est-à-dire le système stochastique suivant

$$\begin{cases} dq(t) = M_0^{-1} p(t) dt, \\ dp(t) = -D_0 M_0^{-1} p(t) dt - K_0 q(t) dt + S_0 dB_t, \end{cases} \quad (5.6)$$

$q(0) \in \mathbf{R}^2$  et  $p(0) \in \mathbf{R}^2$ . C'est un système Hamiltonien et l'Hamiltonien  $H$  s'écrit

$$H(q, p) = \frac{1}{2} \langle M_0^{-1} p, p \rangle + \frac{1}{2} \langle K_0 q, q \rangle.$$

À l'aide de l'équation de Fokker-Planck, on vérifie aisément que la mesure à densité  $\rho$  définie par

$$\rho(q, p) = c_2 \exp(-2H(q, p)) \quad \text{avec } c_2 > 0 \text{ constante de normalisation,} \quad (5.7)$$

est une mesure invariante du système.

Les valeurs numériques utilisées pour les simulations sont les suivantes

$$D_0 = S_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad M_0 = \begin{pmatrix} \frac{5}{4} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} \quad \text{et} \quad K_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

On sait alors que le système a une unique mesure stationnaire (voir [Soi94] pages 208-209). Cette mesure stationnaire est donc la mesure à densité  $\rho$  donnée en (5.7).

On considère une suite de pas  $(\gamma_n)_{n \geq 1} = (\gamma_0 n^{-r})_{n \geq 1}$  et le schéma de discrétisation  $(q_n, p_n)_{n \geq 1}$  défini par

$$\begin{cases} q_{n+1} = \gamma_{n+1} M_0^{-1} p_n, & q_0 = (0.5 \ 0.5), \\ p_{n+1} = -\gamma_{n+1} (D_0 M_0^{-1} p_n + K_0 q_n) + S_0 \sqrt{\gamma_{n+1}} U_{n+1}, & p_0 = (0.5 \ 0.5). \end{cases} \quad (5.8)$$

Le bruit blanc  $(U_n)_{n \geq 1}$  est gaussien.

La suite  $(\eta_n)_{n \geq 1}$  utilisée pour la pondération des mesures empiriques est la suite  $(n^{-s})_{n \geq 1}$  avec  $s \leq \frac{1+r}{2}$ . La fonction  $f$  que l'on intègre est la fonction carrée,  $f : (q, p) \rightarrow |q|^2 + |p|^2$ . L'intégrale  $\nu(f)$  vaut 2.125 et  $q^* = 4$ .

Sauf mention contraire, on considère dans les simulations qui suivent un échantillon de 500 trajectoires  $\nu_n^\eta(f)$  indépendantes avec  $n$  variant jusqu'à  $10^6$ . On calcule la moyenne empirique  $m_n$  et la variance empirique  $s_n^2$ , et on compare pour différents paramètres l'erreur en norme  $\mathbf{L}^2$  définie par

$$e_n(\gamma, r, s) = \sqrt{(m_n - 2.125)^2 + s_n^2}.$$

Dans un premier temps, on représente l'erreur en norme  $\mathbf{L}^2$  en fonction de  $r$  (voir figure V.3). Le minimum est atteint lorsque  $r = 1/3$ , ce qui confirme le fait que la vitesse optimale

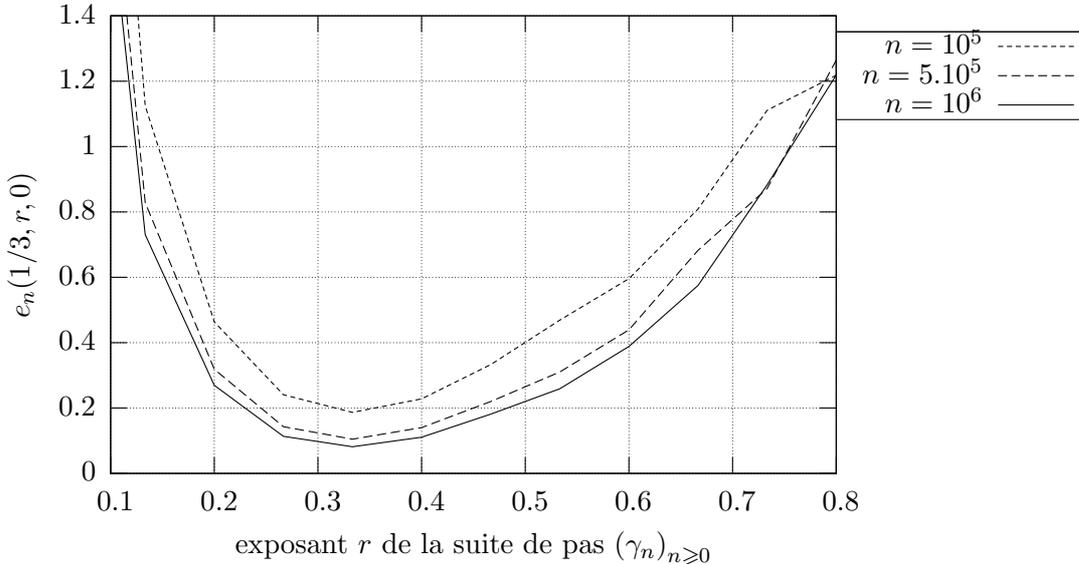


Figure V.3 – Erreur  $\mathbf{L}^2$  en fonction du paramètre  $r$  pour différentes valeurs de  $n$  ( $r = 1/3$  et  $\gamma_0 = 1/3$ ).

est obtenue en prenant  $r = \frac{1}{q^*-1}$ . De plus, on voit que l'erreur augmente considérablement

lorsque l'on s'éloigne de cet optimum.

On va maintenant se placer dans le cas  $r = \frac{1}{3}$  et représenter l'erreur en fonction des deux paramètres restants :  $s$  et  $\gamma_0$ . La figure V.4 montre l'influence de la suite de poids considérée sur l'erreur obtenue.

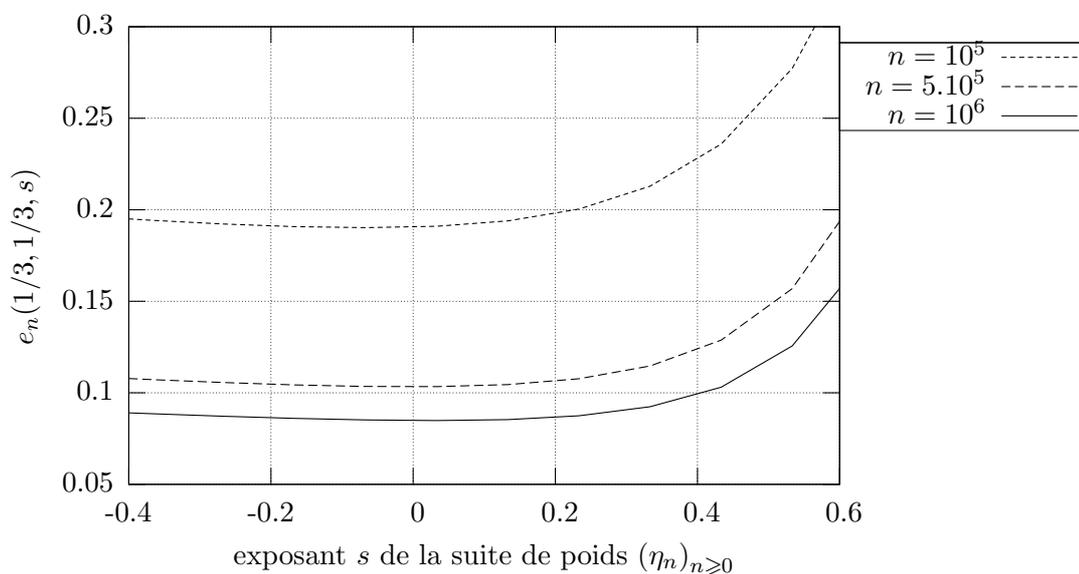


Figure V.4 – Erreur  $L^2$  en fonction du paramètre  $s$  pour différentes valeurs de  $n$  ( $r = 1/3$  et  $\gamma_0 = 1/3$ ).

Dans le cas  $q^* = 4$ , on sait qu'il faut prendre  $\gamma_0 = \left(\sqrt{3} \frac{|m_{f,q^*}|}{s_f}\right)^{-2/3}$ . Cependant il est difficile de connaître  $\frac{|m_{f,q^*}|}{s_f}$  et très coûteux numériquement d'avoir une bonne estimation de cette quantité. On va donc représenter l'évolution de l'erreur en fonction d'un paramètre  $\theta$  remplaçant ce rapport biais-variance inconnu. Le résultat est donné par la figure V.5. On voit que le minimum de l'erreur est atteint pour  $\theta \simeq 1.5$  et que l'erreur varie très peu pour  $\theta \in [1, 2]$ . De plus, cette variation diminue d'autant plus que  $n$  augmente. Ainsi, il n'est pas nécessaire de chercher à approcher  $\frac{|m_{f,q^*}|}{s_f}$  avec une bonne précision.

On finit cette partie numérique en représentant sur la figure V.6 l'évolution de l'intervalle de confiance empirique à 95% en fonction de  $n$ . Il s'agit donc de l'intervalle  $[m_n - 1.96s_n; m_n + 1.96s_n]$ . Ces résultats valident les vitesses données par la proposition V.5.

## 6. Extrapolation de Richardson

On rappelle que le développement de l'erreur fait apparaître un terme martingale et un reste composé d'une somme de la forme

$$\sum_{q=3}^p \frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} D_q(X_{k-1}). \quad (6.1)$$

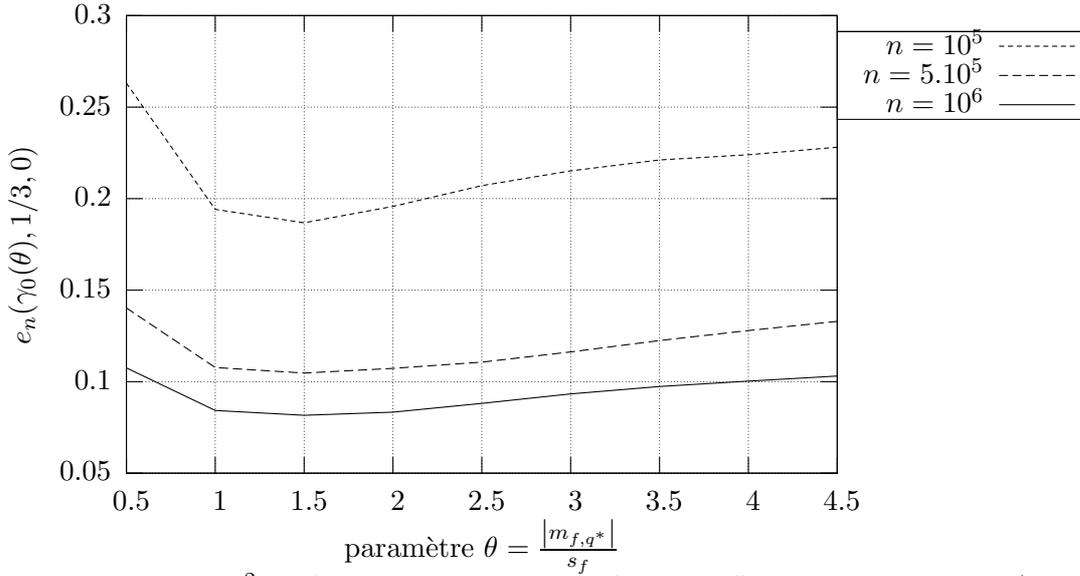


Figure V.5 – Erreur  $\mathbf{L}^2$  en fonction du paramètre  $\theta$  pour différentes valeurs de  $n$  ( $r = 1/3$ ,  $s = 0$  et  $\gamma_0 = (\sqrt{3}\theta)^{-2/3}$ ).

De plus, on a vu que la vitesse de convergence dépend du premier indice  $q^* \in \{3, \dots, p\}$  tel que  $\frac{v_n}{H_n} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1} D_{q^*}(X_{k-1})$  soit non nul. Plus  $q^*$  est grand, plus l'ordre de la vitesse de convergence de l'algorithme est élevé. L'idée est donc de mettre en oeuvre une extrapolation de Richardson afin d'annuler le premier terme (non nul) de cette somme.

Nous commençons par un petit rappel sur cette technique d'extrapolation.

### 6.1. Rappel sur l'extrapolation de Richardson

L'extrapolation de Richardson est une technique souvent utilisée pour améliorer les méthodes numériques. Supposons que l'on ait une approximation  $Q(h)$  (pour  $h > 0$ ) d'une certaine quantité  $Q$  de la forme suivante

$$Q(h) = Q + Ah^p + \mathcal{O}(h^{p+1}) \quad \text{avec } A \text{ ne dépendant pas de } h.$$

L'idée de l'extrapolation de Richardson est d'éliminer le terme  $Ah^p$  en prenant deux valeurs différentes de  $h$ . Prenons  $h$  et  $\frac{h}{\lambda}$  avec  $\lambda > 0$ . On a

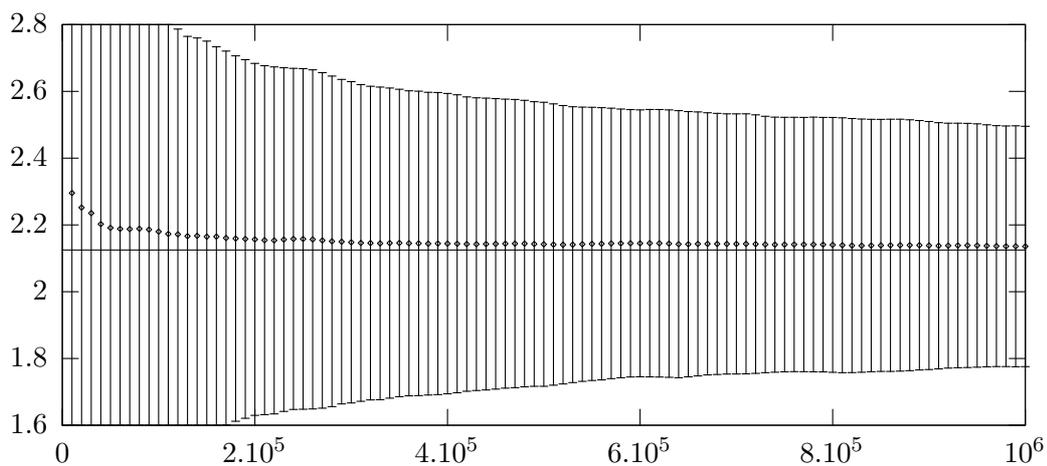
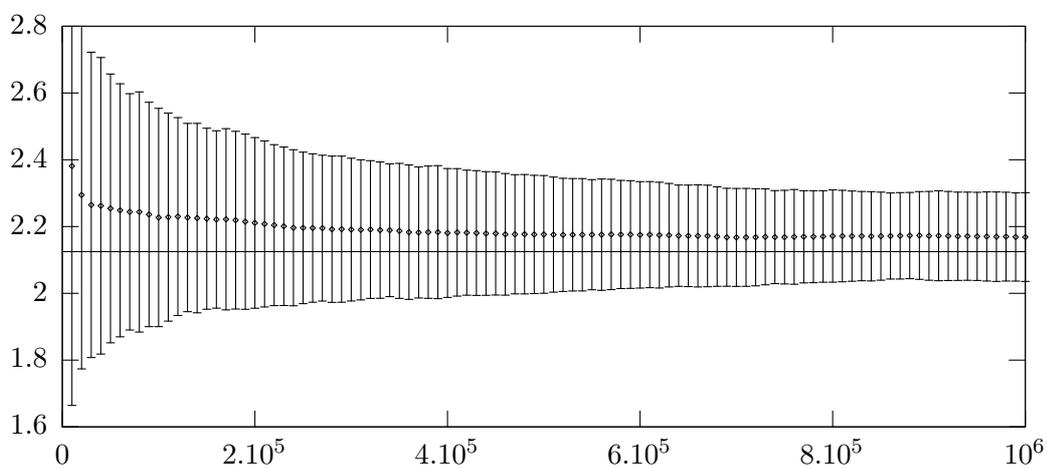
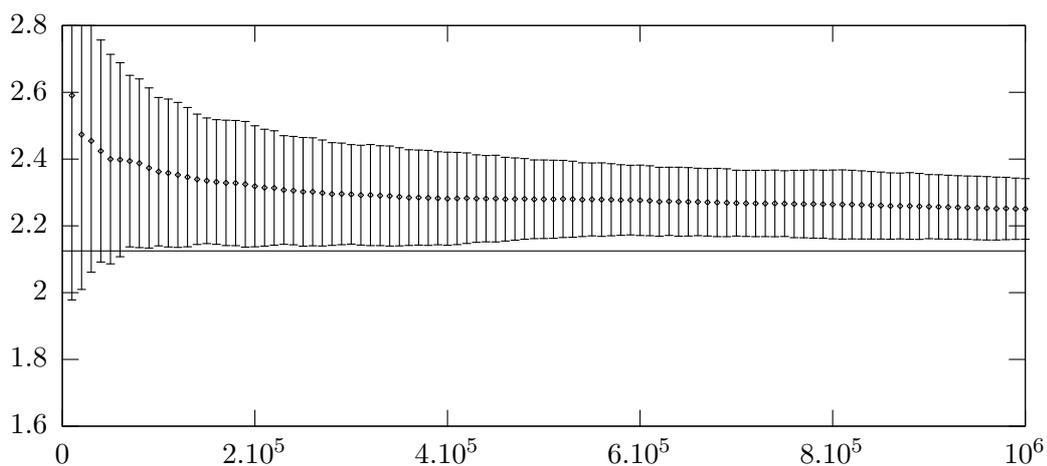
$$Q\left(\frac{h}{\lambda}\right) = Q + A\left(\frac{h}{\lambda}\right)^p + \mathcal{O}(h^{p+1}),$$

et pour éliminer le terme  $Ah^p$ , on pose

$$\tilde{Q}_1^\lambda(h) = \frac{\lambda^p}{\lambda^p - 1} Q\left(\frac{h}{\lambda}\right) - \frac{1}{\lambda^p - 1} Q(h).$$

Ainsi  $\tilde{Q}_1^\lambda(h) = Q + \mathcal{O}(h^{p+1})$ . On a donc gagné un ordre de convergence.

Si  $\tilde{Q}_1^\lambda = \tilde{Q}^\lambda$  vérifie  $Q = \tilde{Q}_1^\lambda(h) + Bh^{p+1} + \mathcal{O}(h^{p+2})$  alors par le même mécanisme, on peut

(a)  $r = 1/2$ ,  $s = 1/2$  et  $\gamma_0 = 1/3$  (TLC centré à la vitesse  $n^{1/4}$ )(b)  $r = 1/3$ ,  $s = 0$  et  $\gamma_0 = 1/3$  (TLC décentré à la vitesse  $n^{1/3}$ )(c)  $r = 1/4$ ,  $s = -1/3$  et  $\gamma_0 = 1/3$  (convergence *p.s.* à la vitesse  $n^{1/4}$ )Figure V.6 – Évolution de l'intervalle de confiance empirique à 95% en fonction du nombre d'itérations  $n$ , pour différentes valeurs de  $r$ .

gagner un ordre de convergence en considérant l'approximation  $\tilde{Q}_2^\lambda(h)$  définie par

$$\begin{aligned}\tilde{Q}_2^\lambda(h) &= \frac{\lambda^{p+1}}{\lambda^{p+1}-1} \tilde{Q}_1^\lambda\left(\frac{h}{\lambda}\right) - \frac{1}{\lambda^{p+1}-1} \tilde{Q}_1^\lambda(h), \\ &= \frac{1}{(\lambda^{p+1}-1)(\lambda^p-1)} \left( \lambda^{2p+1} Q\left(\frac{h}{\lambda^2}\right) - (\lambda^{p+1} + \lambda^p) Q\left(\frac{h}{\lambda}\right) + Q(h) \right).\end{aligned}$$

Dans ce cas, on a  $\tilde{Q}_2^\lambda(h) = Q + \mathcal{O}(h^{p+2})$ .

Cette extrapolation peut être réitérée afin d'obtenir un ordre de convergence  $q$  souhaité. Cependant la constante apparaissant devant le terme d'ordre  $q$  et la complexité augmentent à chaque itération.

## 6.2. Application de cette méthode à notre algorithme

Dans cette sous-section, on suppose que le bruit blanc  $(U_n)_{n \geq 1}$  est symétrique et admet des moments à tout ordre. Ainsi pour tout  $p \geq 2$ , on a  $V^p = \mathcal{O}(W)$ . Dans la suite, on note la dépendance en  $\varphi$  de la fonction  $D_q$  dans (1.6) *i.e.* on pose

$$D_q(\varphi)(x) = \sum_{q/2 \leq j \leq q} \frac{C_j^{q-j}}{j!} D^j \varphi(x) \cdot \left( (b(x))^{\otimes (q-j)}, \mathbf{E} \left[ (\sigma(x) U_1)^{\otimes (2j-q)} \right] \right).$$

*Remarque:* Lorsque le bruit blanc est symétrique, pour toute fonction  $f$  ( $q$  fois dérivable) on a  $D_q(f) = 0$  si  $q$  est impair. \*

On va utiliser l'extrapolation de Richardson en considérant deux suites de pas différentes. On se place sous les hypothèses sur  $(\gamma_n)_{n \geq 1}$  et  $(\eta_n)_{n \geq 1}$  qui induisent une vitesse de convergence en probabilité (ou presque sûre). Les pas et les poids vérifient donc  $\sqrt{\Theta_n} = o\left(\sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1}\right)$  où  $q^*$  est le plus petit  $q$  indice de  $\{4, \dots, p\}$  n'annulant pas  $D_q(\varphi)$ . L'extrapolation de Richardson permet de pousser plus loin la vitesse de convergence en annulant le terme en  $D_{q^*}$ . Le prochain terme non nul apparaissant dans le développement de l'erreur est alors en  $D_{q^*+2}$ . On se donne un paramètre  $\lambda > 0$  et on définit le schéma  $(X_n^\lambda)_{n \geq 0}$  par

$$X_{n+1}^\lambda = X_n^\lambda + \frac{\gamma_{n+1}}{\lambda} b(X_n^\lambda) + \sqrt{\frac{\gamma_{n+1}}{\lambda}} \sigma(X_n^\lambda) U_{n+1}, \quad X_0 = x_0 \in \mathbf{R}^d.$$

où  $(U_n)_{n \geq 1}$  est le même bruit blanc que celui utilisé pour construire  $(X_n)_{n \geq 0}$ .

La suite des mesures empiriques pondérées par  $\eta = (\eta_n)_{n \geq 1}$  de ce schéma est notée  $(\nu_n^{\eta, \lambda})_{n \geq 1}$ , et on introduit la suite  $(\tilde{\nu}_n^\lambda)_{n \geq 1}$  définie par

$$\forall n \geq 1, \quad \tilde{\nu}_n^\lambda = \frac{1}{\lambda^{\frac{q^*}{2}-1} - 1} \left( \lambda^{\frac{q^*}{2}-1} \nu_n^{\eta, \lambda} - \nu_n^\eta \right). \quad (6.2)$$

Pour obtenir des résultats de convergence lorsqu'une telle extrapolation est mise en œuvre, il faut d'importantes hypothèses de régularité sur la fonction  $f$  et la fonction  $\varphi$  solution de l'équation de Poisson  $\mathcal{A}\varphi = -(f - \nu(f))$ . Comme dans le théorème V.3, on suppose que

$\varphi \in E_p^W$ , que  $D^p\varphi$  est Lipschitz et on note  $q^* = \min_{q \in \{4, \dots, p\}} \{D_q(\varphi) \neq 0\} \wedge (p+1)$ . On suppose de plus que  $D_{q^*}(\varphi)$  admet une solution  $\psi$  à l'équation de Poisson

$$\mathcal{A}\psi = -(D_{q^*}(\varphi) - \nu(D_{q^*}(\varphi))), \quad (6.3)$$

et que  $\psi$  vérifie :  $\exists \bar{p} \geq 4$ ,  $\psi \in E_{\bar{p}}^W$  et  $D^{\bar{p}}\psi$  Lipschitz.

D'autre part, pour obtenir le TLC dans le théorème qui suit, on doit supposer que la diffusion  $(y_t)_{t \geq 0} = (y_t^1, y_t^2)_{t \geq 0}$  définie sur  $\mathbf{R}^{2d}$  par

$$\begin{cases} dy_t^1 = \frac{b(y_t^1)}{\lambda} dt + \frac{\sigma(y_t^2)}{\sqrt{\lambda}} dB_t, \\ dy_t^2 = b(y_t^1) dt + \sigma(y_t^2) dB_t, \end{cases} \quad (6.4)$$

vérifie une condition de stabilité forte de type  $(\tilde{V}_\lambda, \text{Id})$  et qu'elle possède une unique mesure invariante  $\nu_\lambda$ . De plus, on suppose que  $|\sigma^* \nabla \varphi|^2 = o(\sqrt{\tilde{V}})$ .

Sous ces hypothèses techniques, on obtient le résultat suivant.

**Théorème V.6:**

On suppose que  $\sqrt{\Theta_n} = o\left(\sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1}\right)$  et on pose  $\tilde{\Upsilon}_n^* = \sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}}$ . Si la suite  $\left(\frac{\tilde{\Upsilon}_n^*}{\sqrt{\Theta_n}}\right)_{n \geq 1}$  admet une limite  $\xi$  dans  $\mathbf{R}_+ \cup \{+\infty\}$ , alors

– si  $\sum_{n \geq 1} \frac{1}{\sqrt{\Theta_n}} \left| \Delta \frac{\eta_n}{\gamma_n} \right| < +\infty$ , et  $\sum_{n \geq 1} \frac{\eta_n^2}{\Theta_n} < +\infty$ , on a

$$\frac{H_n}{\sqrt{\Theta_n}} (\tilde{\nu}_n^\lambda(f) - \nu(f)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tilde{s}_{f,\lambda}^2) \quad \text{si } \xi = 0, \quad (6.5)$$

$$\frac{H_n}{\sqrt{\Theta_n}} (\tilde{\nu}_n^\lambda(f) - \nu(f)) \xrightarrow{\mathcal{L}} \mathcal{N}(\xi \tilde{m}_{f,q^*,\lambda}, \tilde{s}_{f,\lambda}^2) \quad \text{si } 0 < \xi < \infty, \quad (6.6)$$

– si  $\sum_{n \geq 1} \frac{1}{\tilde{\Upsilon}_n^*} \left| \Delta \frac{\eta_n}{\gamma_n} \right| < +\infty$ , et  $\sum_{n \geq 1} \frac{\eta_n^2}{(\tilde{\Upsilon}_n^*)^2} < +\infty$ , on a

$$\frac{H_n}{\tilde{\Upsilon}_n^*} (\tilde{\nu}_n^\lambda(f) - \nu(f)) \xrightarrow{\mathbf{P}} \tilde{m}_{f,q^*,\lambda} \quad \text{si } \xi = +\infty, \quad (6.7)$$

avec

$$\tilde{m}_{f,q^*,\lambda} = \left( \frac{\lambda^{-1} - 1}{\lambda^{\frac{q^*}{2}-1} - 1} \right) \int_{\mathbf{R}^d} D_{q^*+2} \nu \quad \text{et} \quad \tilde{s}_{f,\lambda}^2 = \int_{\mathbf{R}^{2d}} |g(y^1, y^2)|^2 \nu_\lambda(dy),$$

$$\text{où } g_\lambda(y^1, y^2) = \frac{\lambda^{\frac{q^*}{2}-1} \sqrt{\lambda}}{\lambda^{\frac{q^*}{2}-1} - 1} (\sigma^* \nabla \varphi)(y^1) - \frac{1}{\lambda^{\frac{q^*}{2}-1} - 1} (\sigma^* \nabla \varphi)(y^2).$$

De plus, si la suite  $\left(\frac{\eta_n}{\tilde{\Upsilon}_n^* \gamma_n}\right)_{n \geq 1}$  décroît vers 0 et que  $\sum_{n \geq 1} \frac{\eta_n^2}{(\tilde{\Upsilon}_n^*)^2 \gamma_n} < +\infty$ , alors la convergence dans (6.7) est presque sûre. ★

Preuve. La preuve se fait de la même façon que celle du théorème V.3. Dans le théorème V.3, soit le terme martingale domine le premier terme non nul (terme en  $q^*$ ) du reste du

développement et on a un TLC, soit le terme en  $q^*$  domine le terme martingale et on a une convergence en probabilité. Ici, c'est exactement la même situation sauf que la procédure de Richardson "tue" le terme en  $q^*$  et qu'il y a donc compétition entre le terme martingale et le terme en  $q^* + 2$ .

- Montrons tout d'abord que la procédure de Richardson fonctionne bien sous nos hypothèses, c'est à dire que le terme en  $q^*$  s'annule. La fonction  $D_{q^*}(\varphi)$  a une solution  $\psi$  à l'équation de Poisson (6.3) telle qu'il existe  $\tilde{p} \geq 4$  vérifiant  $\psi \in E_{\tilde{p}}^W$  et  $D^{\tilde{p}}\psi$  bornée. Dans ce cas,  $D_4(\psi)$  existe et comme  $\sqrt{\Theta_n} = o\left(\sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1}\right)$  alors le théorème V.3 appliqué avec la suite de poids  $\tilde{\eta} = \left(\eta_n \gamma_n^{\frac{q^*}{2}-1}\right)_{n \geq 1}$  donne

$$\frac{1}{\sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}}} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1} (D_{q^*}(\varphi) - \nu(D_{q^*}(\varphi))) \xrightarrow{\mathbf{P}} m_{D_{q^*}(\varphi), 4}. \quad (6.8)$$

Cette convergence est presque sûre lorsque  $\left(\frac{\eta_n \gamma_n^{\frac{q^*}{2}-1}}{\tilde{\Upsilon}_n^* \gamma_n}\right)_{n \geq 1}$  décroît et  $\sum_{n \geq 1} \frac{\eta_n^2 \gamma_n^{q^*-2}}{(\tilde{\Upsilon}_n^*)^2 \gamma_n} < +\infty$ .

En écrivant

$$D_{q^*}(\varphi)(X_{k-1}^\lambda) - D_{q^*}(\varphi)(X_{k-1}) = D_{q^*}(\varphi)(X_{k-1}^\lambda) - \nu(D_{q^*}(\varphi)) + \nu(D_{q^*}(\varphi)) - D_{q^*}(\varphi)(X_{k-1}),$$

on obtient

$$\frac{1}{\tilde{\Upsilon}_n^*} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q^*}{2}-1} \frac{1}{\lambda^{\frac{q^*}{2}-1} - 1} \left( D_{q^*}(X_{k-1}^\lambda) - D_{q^*}(X_{k-1}) \right) \xrightarrow{\mathbf{P}} 0. \quad (6.9)$$

Comme la suite  $(\gamma_n)_{n \geq 0}$  est décroissante, on remarque que les conditions donnant la convergence *p.s.* de (6.7) dans l'énoncé du théorème impliquent la convergence *p.s.* de (6.9).

L'extrapolation de Richardson fonctionne. Il faut maintenant montrer les différentes convergences en fonction de  $\xi$ . La preuve reprend essentiellement celle du théorème V.3.

- Supposons  $\xi = 0$ . Alors la proposition V.4 appliquée avec la suite  $(v_n)_{n \geq 1} = \left(\frac{H_n}{\sqrt{\Theta_n}}\right)_{n \geq 1}$  donne

$$\frac{H_n}{\sqrt{\Theta_n}} \left( \tilde{v}_n^\lambda(f) - \nu(f) \right) - \left( M_n^\lambda + S_n^\lambda \right) \xrightarrow{\mathbf{L}^2} 0, \quad (6.10)$$

avec

$$M_n^\lambda = \frac{1}{\sqrt{\Theta_n}} \sum_{k=1}^n \frac{\eta_k}{\sqrt{\gamma_k}} \left( \frac{\lambda^{\frac{q^*}{2}-1} \sqrt{\lambda}}{\lambda^{\frac{q^*}{2}-1} - 1} \langle (\sigma^* \nabla \varphi)(X_{k-1}^\lambda), U_k \rangle - \frac{1}{\lambda^{\frac{q^*}{2}-1} - 1} \langle (\sigma^* \nabla \varphi)(X_{k-1}), U_k \rangle \right),$$

$$S_n^\lambda = \sum_{q=q^*}^p \left( \frac{1}{\sqrt{\Theta_n}} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} \left( \frac{\lambda^{\frac{q^*}{2}-q}}{\lambda^{\frac{q^*}{2}-1} - 1} D_q(\varphi)(X_{k-1}^\lambda) - \frac{1}{\lambda^{\frac{q^*}{2}-1} - 1} D_q(\varphi)(X_{k-1}) \right) \right).$$

Comme  $D_q(\varphi) = o(\sqrt{W})$  pour tout  $q \in \{q^* + 2, \dots, p\}$  ( $D_{q^*+1}(\varphi) = 0$ ), et que  $\Upsilon_n^* =$

$o(\sqrt{\Theta_n})$  alors on a

$$\sum_{q=q^*+2}^p \left( \frac{1}{\sqrt{\Theta_n}} \sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} \left( \frac{\lambda^{\frac{q^*-q}{2}}}{\lambda^{\frac{q^*}{2}-1} - 1} D_q(\varphi)(X_{k-1}^\lambda) - \frac{1}{\lambda^{\frac{q^*}{2}-1} - 1} D_q(\varphi)(X_{k-1}) \right) \right) \xrightarrow{\mathbf{L}^2} 0.$$

D'après (6.9), on en déduit que  $(S_n^\lambda)_{n \geq 0}$  converge en probabilité vers 0.

D'autre part, le couple  $(\gamma, \theta)$  est moyennisant et  $|\sigma^* \varphi|^2 = o(\sqrt{\tilde{V}_\lambda})$  donc le lemme V.2 s'applique et donne

$$M_n^\lambda \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tilde{s}_{f,\lambda}^2).$$

La convergence (6.5) est prouvée.

- Supposons maintenant  $\xi \in ]0, \infty[$ . La proposition V.4 appliquée avec la suite  $(v_n)_{n \geq 1} = \left( \frac{H_n}{\sqrt{\Theta_n}} \right)_{n \geq 1}$  donne (6.10).

Pour tout  $q \geq q^* + 4$ , on a  $\sum_{k=1}^n \eta_k \gamma_k^{\frac{q}{2}-1} = o(\sqrt{\Theta_n})$  et  $D_q(\varphi) = o(\sqrt{W})$  donc la somme allant de  $q^* + 4$  à  $p$  présente dans  $S_n^\lambda$  converge vers 0 dans  $\mathbf{L}^2$ . Puisque l'on a (6.9) et que le couple  $(\gamma, \eta \gamma^{\frac{q^*}{2}})$  est moyennisant, on obtient

$$S_n^\lambda \xrightarrow{\mathbf{P}} \xi \left( \frac{\lambda^{-1} - 1}{\lambda^{\frac{q^*}{2}-1} - 1} \right) m_{f,q^*+2} = \xi \tilde{m}_{f,q^*,\lambda}.$$

Le terme  $M_n^\lambda$  se gère comme dans le cas  $\xi = 0$ , et on obtient le TLC décentré (6.6).

- Si  $\xi = +\infty$ , *i.e.*  $\sqrt{\Theta_n} = o(\tilde{\Upsilon}_n^*)$  alors la proposition V.4 appliquée avec la suite  $(v_n)_{n \geq 1} = \left( \frac{H_n}{\tilde{\Upsilon}_n^*} \right)_{n \geq 1}$  donne

$$\frac{H_n}{\tilde{\Upsilon}_n^*} \left( \tilde{\nu}_n^\lambda(f) - \nu(f) \right) - \frac{\sqrt{\Theta_n}}{\tilde{\Upsilon}_n^*} \left( M_n^\lambda + S_n^\lambda \right) \xrightarrow{\mathbf{L}^2} 0.$$

La convergence en probabilité vers 0 du terme martingale découle de la majoration dans  $\mathbf{L}^2$  suivante

$$\left\| M_n^\lambda \right\|_2 \leq \sup_{k=0, \dots, n-1} \left\{ \left\| (\sigma^* \nabla \varphi)(X_k^\lambda) \right\|_2 + \left\| (\sigma^* \nabla \varphi)(X_k) \right\|_2 \right\}. \quad (6.11)$$

Pour la convergence presque sûre, on applique le lemme V.3 avec les suites  $(a_n)_{n \geq 1} = (\theta_n)_{n \geq 1}$  et  $(b_n)_{n \geq 1} = (\tilde{\Upsilon}_n^*)_{n \geq 1}$ . D'autre part, on prouve comme précédemment que

$$\frac{\sqrt{\Theta_n}}{\tilde{\Upsilon}_n^*} S_n^\lambda \xrightarrow{\mathbf{P}} \tilde{m}_{f,q^*,\lambda},$$

et que l'on a convergence presque sûre sous les conditions annoncées. #

L'extrapolation de Richardson permet donc d'augmenter la vitesse de convergence de  $H_n/\Upsilon_n^*$  à  $H_n/\tilde{\Upsilon}_n^*$ . Par exemple, si  $q^* = 4$  et si les suites de pas et de poids sont de la forme suivante

$$\forall n \geq 1, \quad \gamma_n = \gamma_0 n^{-r} \quad \text{et} \quad \eta_n = \eta_0 n^{-s}, \quad (6.12)$$

alors le TLC optimal (biaisé) est réalisé pour  $r = \frac{1}{5}$  et la vitesse est en  $c_{\mathcal{L}} n^{\frac{2}{5}}$  où  $c_{\mathcal{L}} = \sqrt{\gamma} \frac{\sqrt{1-2s+r}}{1-s}$ . Plus généralement, quand une extrapolation de Richardson est mise en œuvre, la proposition V.5 est vraie en remplaçant  $q^*$  par  $q^*+2$ , le biais  $m_{f,q^*}$  par  $\tilde{m}_{f,q^*,\lambda}$  et la variance  $s_f^2$  par  $\tilde{s}_{f,\lambda}^2$ . Ces nouveaux biais et variance sont en général assez complexe à calculer.

Les paramètres  $\gamma_0$  et  $\lambda$  sont liés et difficile à déterminer. On s'intéresse à ce problème d'un point de vue numérique dans le chapitre VIII.

On pourrait envisager une réitération de l'extrapolation pour pousser plus loin la vitesse de convergence (passer de  $q^*$  à  $q^* + 4$  par exemple), mais l'augmentation de la complexité de l'algorithme ainsi que de la variance ne laisse rien présager de bon.

## 7. Annexe

Preuve (lemme V.1). La suite  $(\gamma_n)_{n \geq 1}$  est décroissante et  $q \geq 0$  donc pour tout  $k \leq n$ ,  $\gamma_n^q \leq \gamma_k^q$  et l'on a

$$\frac{\gamma_n^q}{\sum_{k=1}^n \eta_k \gamma_k^q} \leq \frac{1}{\sum_{k=1}^n \eta_k}.$$

Ainsi

$$\sum_{n \geq 1} \left( \frac{\eta_n \gamma_n^q}{(\sum_{k=1}^n \eta_k \gamma_k^q) \sqrt{\gamma_n}} \right)^2 \leq \sum_{n \geq 1} \left( \frac{\eta_n}{H_n \sqrt{\gamma_n}} \right)^2 < +\infty.$$

D'autre part, on a

$$\begin{aligned} \Delta \frac{\eta_n \gamma_n^q}{\gamma_n} &= \frac{\eta_n \gamma_n^q}{\gamma_n} - \frac{\eta_{n-1} \gamma_{n-1}^q}{\gamma_{n-1}}, \\ &= \left( \Delta \frac{\eta_n}{\gamma_n} \right) \gamma_n^q + \frac{\eta_{n-1}}{\gamma_{n-1}} (\Delta \gamma_n^q), \end{aligned}$$

et comme  $\Delta \gamma_n^q \leq 0$ ,

$$\sum_{n \geq 1} \frac{1}{\sum_{k=1}^n \eta_k \gamma_k^q} \left( \Delta \frac{\eta_n \gamma_n^q}{\gamma_n} \right)_+ \leq \sum_{n \geq 1} \frac{1}{\sum_{k=1}^n \eta_k} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ < +\infty. \quad \#$$

Preuve (proposition V.2). Montrons que le couple  $(\gamma, \eta)$  est moyennisant. Tout d'abord, il est évident que

$$\begin{aligned} \lim_n \gamma_n = 0 &\Leftrightarrow r > 0, \\ \lim_n H_n = +\infty &\Leftrightarrow s \leq 1. \end{aligned}$$

- Si  $s = 1$ , alors on a  $H_n \sim \ln(n)$  et

$$\Delta \frac{\eta_n}{\gamma_n} = \frac{\eta_0}{\gamma_0} (n^{r-1} - (n-1)^{r-1}) \sim \frac{\eta_0}{\gamma_0} (r-1)n^{r-2},$$

donc

$$\sum_{n \geq 1} \frac{1}{H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ < +\infty \Leftrightarrow \sum_{n \geq 1} \frac{1}{\ln n} (r-1)_+ n^{r-2} < +\infty \Leftrightarrow r \leq 1.$$

De plus, on a

$$\sum_{n \geq 1} \left( \frac{\eta_n}{H_n \sqrt{\gamma_n}} \right)^2 < +\infty \Leftrightarrow \sum_{n \geq 1} \frac{1}{n^{2-r} \ln n} < +\infty \Leftrightarrow r \leq 1.$$

- Si  $s < 1$ , alors en utilisant les équivalences suivantes

$$H_n \sim \eta_0 n^{1-s} \quad \text{et} \quad \Delta \frac{\eta_n}{\gamma_n} = \frac{\eta_0}{\gamma_0} ((n+1)^{-s+r} - n^{-s+r}) \sim \frac{\eta_0}{\gamma_0} (r-s) n^{r-s-1},$$

on a

$$\begin{aligned} \sum_{n \geq 1} \frac{1}{H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ < +\infty &\Leftrightarrow \left( \left( r \geq s \text{ et } \sum_{n \geq 1} \frac{n^{r-s-1}}{n^{1-s}} < +\infty \right) \text{ ou } (r < s) \right) \\ &\Leftrightarrow r < 1. \end{aligned}$$

De même, on a

$$\sum_{n \geq 1} \left( \frac{\eta_n}{H_n \sqrt{\gamma_n}} \right)^2 < +\infty \Leftrightarrow \sum_{n \geq 1} \frac{1}{n^{2-r}} < +\infty \Leftrightarrow r < 1.$$

On déduit de toutes ces équivalences que le couple  $(\gamma, \eta)$  est moyennisant si et seulement si

$$(r, s) \in ]0, 1[ \times ]-\infty, 1] \cup \{(1, 1)\}.$$

On rappelle que  $\theta_n = \frac{\eta_n^2}{\gamma_n} = \frac{\eta_0^2}{\gamma_0} n^{-(2s-r)}$ . Par similarité avec le couple  $(\gamma, \eta)$ , on sait que  $(\gamma, \theta)$  est moyennisant si et seulement si

$$(r, 2s-r) \in ]0, 1[ \times ]-\infty, 1] \cup \{(1, 1)\}$$

c'est à dire  $(0 < r < 1 \text{ et } s \leq \frac{1+r}{2})$  ou  $r = s = 1$ .

#

Montrons que le gradient de  $g$  (fonction définie dans la section 5.1) s'annule uniquement en  $s = 0$  et  $\gamma_0 = \left( 2\sqrt{1 - \frac{1}{q^*} \frac{|m_{f,q^*}|}{s_f}} \right)^{-\frac{2}{q^*-1}}$  sur  $]0, +\infty[ \times ]-\infty, \frac{1+r}{2}[$ . On résoud donc le système suivant :

$$\begin{cases} -\frac{1}{\gamma^2} \frac{(1-s)^2}{1-2s+r} s_f^2 + (q^*-2) \gamma^{q^*-3} \left( \frac{1-s}{1-s-r(\frac{q^*}{2}-1)} \right)^2 m_{f,q^*}^2 = 0, \\ -\frac{1}{\gamma} \frac{2(1-s)(r-s)}{(1-2s+r)^2} s_f^2 + \gamma^{q^*-2} \frac{r(q^*-2)(1-s)}{(1-s-r(\frac{q^*}{2}-1))^3} m_{f,q^*}^2 = 0, \end{cases}$$

En multipliant la première égalité par  $\frac{-2\gamma(r-s)}{(1-s)(1-2s+r)}$  et sommant avec la seconde on obtient le système équivalent

$$\begin{cases} -\frac{1}{\gamma^2} \frac{(1-s)^2}{1-2s+r} s_f^2 + (q^* - 2)\gamma^{q^*-3} \left( \frac{1-s}{1-s-r(\frac{q^*}{2}-1)} \right)^2 m_{f,q^*}^2 = 0, \\ -2 \left( \frac{(1-s)(r-s)}{(1-2s+r)(1-s-r(\frac{q^*}{2}-1))^2} + r \frac{(1-s)}{(1-s-r(\frac{q^*}{2}-1))^3} \right) (q^* - 2)\gamma^{q^*-2} m_{f,q^*}^2 = 0. \end{cases}$$

Cette seconde équation est équivalente à

$$s^2 - s \left( 1 - r \left( \frac{q^*}{2} - 1 \right) \right) - \frac{1}{2} r^2 (q^* - 1) + \frac{r}{2} = 0$$

Or, on est dans le cas  $r = \frac{1}{q^*-1}$  donc cette équation admet deux solutions  $s = 0$  et  $s = \frac{1+r}{2}$ . La seule solution dans le domaine d'étude est  $s = 0$ . La première équation donne alors

$$\gamma_0 = \left( \frac{\left( 1 - r \left( \frac{q^*}{2} - 1 \right) \right)^2 s_f^2}{(q^* - 2)(1+r) m_{f,q^*}^2} \right)^{\frac{1}{q^*-1}}.$$

---

# Étude en présence de mesures invariantes multiples

VI

---

L'objet de ce chapitre est d'étudier le comportement des mesures empiriques pondérées du schéma d'Euler de pas décroissant lorsque la diffusion admet plusieurs mesures invariantes. C'est une situation qui peut notamment se présenter lorsque le drift et le coefficient de diffusion s'annulent simultanément.

Dans un premier temps, on se place en dimension 1 et on s'intéresse aux diffusions dont le drift et le coefficient de diffusion s'annulent en un même point  $\Delta$  : ce point est alors un point frontière pour la diffusion qui peut admettre une mesure invariante à gauche  $\nu_-$  et à droite  $\nu_+$  de ce point. De plus, il est clair que  $\delta_\Delta$  est aussi une probabilité invariante pour la diffusion, puisque la fonction constante égale à  $\Delta$  est une solution de l'EDS. La question qui se pose est : "qu'en est-il pour le schéma d'Euler et ses mesures empiriques ?" Pour le schéma d'approximation,  $\Delta$  n'est pas un point frontière. En effet, le schéma, n'étant pas continu, peut "sauter" au-dessus de ce point. Les mesures empiriques pondérées convergent-elles vers l'une des trois mesures invariantes extrémales de la diffusion ( $\nu_-$ ,  $\delta_\Delta$  et  $\nu_+$ ), ou bien vers une combinaison convexe de celles-ci ?

Motivés par ces questions, nous rappelons partiellement l'outil essentiel pour l'étude d'une diffusion unidimensionnelle qu'est la classification de Feller. Nous introduisons en outre les notions, naturelles dans notre cadre, de point attractif, répulsif et fortement répulsif. On établit ainsi une classification en fonction de la nature des points frontières pour le comportement ergodique de la diffusion. Puis on donne des conditions nécessaires et suffisantes sur la nature des points frontières en terme de fonction de Lyapounov. Cela permet d'obtenir une caractérisation plus simple à utiliser pour l'étude du schéma d'Euler. On prouve alors que pour une large gamme de pas, le schéma se comporte comme la diffusion et reste à partir d'un temps fini (presque sûrement) d'un côté du point  $\Delta$ . De plus, dans ce cas, les mesures empiriques pondérées du schéma ont le même comportement que les mesures empiriques de  $(x_t)_{t \geq 0}$ .

Un exemple numérique en dimension supérieure clôt le chapitre.

## 1. Diffusion unidimensionnelle

Soit  $I$  un intervalle ouvert (non vide) de  $\mathbf{R}$  noté  $I = ]l, r[$ , et l'équation différentielle stochastique suivante

$$dx_t = b(x_t)dt + \sigma(x_t)dB_t, \quad (1.1)$$

où  $x_0$  est une variable aléatoire sur  $I$  et  $(B_t)_{t \geq 0}$  est un mouvement brownien standard sur  $\mathbf{R}$ . On suppose que  $b$  et  $\sigma$  sont continues sur  $\bar{I}$  à valeurs dans  $\mathbf{R}$ , et que  $\sigma$  est non dégénérée sur  $I$  i.e.  $\forall x \in I, \sigma^2(x) > 0$ . Il existe alors une unique solution  $(x_t)_{t \geq 0}$  adaptée à la filtration brownienne complétée, continue sur  $[0, \zeta[$  où  $\zeta$  est le temps d'explosion de la diffusion  $\zeta = \inf \{t \geq 0, x_t = l \text{ ou } x_t = r\}$ .

**Définition VI.1.** Un point *frontière* pour l'équation (1.1) est un point  $\Delta$  infini ( $\Delta = -\infty$  ou  $\Delta = +\infty$ ) ou un point fini vérifiant  $b(\Delta) = \sigma(\Delta) = 0$ . ♣

Dans toute ce chapitre, on suppose que  $l$  et  $r$  (les bornes de  $I$ ) sont des points frontières pour l'EDS. Dans ce cadre, nous allons introduire la fonction d'échelle et la mesure de vitesse du processus de diffusion  $(x_t)_{t \geq 0}$  solution de (1.1). Les deux premières sections sont inspirées par différents exposés de la classification de Feller trouvés dans [KT81], [KS91], [Bre92] et [RY99].

## 1.1. Fonction d'échelle et mesure de vitesse

**Définition VI.2 (Fonction d'échelle).** Une *fonction d'échelle* de l'EDS (1.1) est définie, pour un  $c, c \in I$  par

$$\forall x \in I, \quad p(x) = \int_c^x \exp\left(-\int_c^y \frac{2b(z)}{\sigma^2(z)} dz\right) dy. \quad \clubsuit$$

C'est une fonction strictement croissante définie à une constante près sur  $I$ . Deux fonctions d'échelle se déduisent l'une de l'autre par une affinité strictement croissante. Par abus de langage on parle de la fonction d'échelle de l'EDS.

Par continuité de  $b$  et  $\sigma$ , et non dégénérescence de  $\sigma$  sur  $I$ , la fonction d'échelle est de classe  $\mathcal{C}^2$  sur  $I$  à valeurs dans  $\mathbf{R}$ . C'est une solution de l'équation différentielle ordinaire

$$\forall x \in I, \quad b(x)p'(x) + \frac{1}{2}\sigma^2(x)p''(x) = 0 \quad \text{i.e.} \quad \mathcal{A}p(x) = 0,$$

et cette propriété la caractérise. Ainsi la probabilité d'atteindre un point  $a$  de  $I$  avant un point  $b$  de  $I$  partant de  $x$  dans l'intervalle de bornes  $a$  et  $b$  s'exprime grâce à la fonction d'échelle  $p$ . Pour tout  $a \in I$  on note  $T_a$  le temps d'atteinte de  $(x_t)_{t \geq 0}$  en  $a$  i.e.  $T_a = \inf \{t \geq 0, x_t = a\}$  et on considère un intervalle non vide  $]a, b[ \subset I$  (inclusion stricte). La fonction  $u(x)$  définie sur  $]a, b[$  par  $\forall x \in ]a, b[, u(x) = \mathbf{P}_x [x_{T_a \wedge T_b} = b]$  est solution du système

$$\begin{cases} \mathcal{A}u = 0 & \text{sur } ]a, b[ \\ u(a) = 0 & \text{et } u(b) = 1, \end{cases}$$

donc pour tout  $x \in ]a, b[$ ,

$$u(x) = \mathbf{P}_x [T_b < T_a] = \frac{p(x) - p(a)}{p(b) - p(a)} \quad (1.2)$$

Cette caractérisation est souvent utilisée pour définir la fonction d'échelle dans un cadre plus général *i.e.* pour des processus continus fortement markoviens et réguliers au sens de Dynkin ( $\forall x \in I, \forall y \in I, \mathbf{P}_x [T_y < +\infty] > 0$ ) (cf. [RW94] ou [RY99]).

Une autre caractérisation de la fonction d'échelle est donnée par la proposition suivante. On rappelle que  $\zeta$  est le temps d'explosion de la diffusion  $(x_t)_{t \geq 0}$ .

**Proposition VI.1:**

Le processus  $\left(p(x_t^\zeta)\right)_{t \geq 0}$  est une martingale locale si et seulement si  $p$  est la fonction d'échelle. ♦

Preuve. Voir proposition (VII.3.5) de [RY99] dans un cadre plus général.

- Si  $\left(p(x_t^\zeta)\right)_{t \geq 0}$  est une martingale locale, alors pour tout  $a < x < b$  le processus  $\left(p(x_t^{T_a \wedge T_b})\right)_{t \geq 0}$  est une martingale bornée et par le théorème d'arrêt on a

$$p(x) = p(a)\mathbf{P}_x [T_a < T_b] + p(b)\mathbf{P}_x [T_b < T_a],$$

d'où (1.2).

- Si  $p$  est la fonction d'échelle alors  $\mathcal{A}p = 0$ , et par la formule d'Itô appliquée à  $(x_t^\zeta)_{t \geq 0}$  avec  $p$  on en déduit que  $\left(p(x_t^\zeta)\right)_{t \geq 0}$  est une martingale locale. #

La proposition précédente est très utile car elle permet de ramener l'étude de toute diffusion unidimensionnelle à l'étude d'une martingale locale.

D'autre part, le processus  $(y_t)_{t \geq 0}$  défini pour tout  $t \geq 0$  par  $y_t = p(x_t)$  est une solution de l'équation sans drift :

$$y_t = y_0 + \int_0^t g(y_s)dB_s, \quad (1.3)$$

où  $y_0 = p(x_0) \in p(I)$  *p.s.* et  $g$  définie par

$$g(y) = \begin{cases} ((\sigma p') \circ p^{-1})(y) & \text{si } y \in p(I), \\ 0 & \text{sinon.} \end{cases}$$

Le processus  $(y_t)_{t \geq 0}$  apparaît donc comme un brownien changé de temps. La mesure de vitesse du processus  $(y_t)_{t \geq 0}$  évalue la façon dont le changement de temps affecte le temps moyen de sortie d'un intervalle ouvert borné. Soit  $\tilde{\mathcal{A}}$  le générateur associé à  $(y_t)_{t \geq 0}$  et  $v$  la fonction définie sur  $J = ]\tilde{a}, \tilde{b}[ \subset p(I)$  par  $v(y) = \mathbf{E}_y [\tilde{T}_{\tilde{a}} \wedge \tilde{T}_{\tilde{b}}]$  où  $\tilde{T}_z = \inf \{t \geq 0, y_t = z\}$ . La fonction  $v$  est solution du système

$$\begin{cases} \tilde{\mathcal{A}}v(y) = -1, & \tilde{a} < y < \tilde{b}, \\ v(\tilde{a}) = v(\tilde{b}) = 0, \end{cases} \quad (1.4)$$

et d'après (1.3), on a  $\tilde{\mathcal{A}}v(y) = \frac{1}{2}g^2(y)v''(y)$  pour tout  $y \in ]\tilde{a}, \tilde{b}[$  donc

$$\forall y \in ]\tilde{a}, \tilde{b}[, \quad v(y) = \int_{\tilde{a}}^{\tilde{b}} G_{\tilde{a}, \tilde{b}}(y, z) \frac{2}{g^2(z)} dz, \quad (1.5)$$

où  $G_{\tilde{a},\tilde{b}}(y, z)$  est la fonction de Green définie par  $G_{\tilde{a},\tilde{b}}(y, z) = \frac{(y \wedge z - \tilde{a})(\tilde{b} - y \vee z)}{\tilde{b} - \tilde{a}}$ , pour tout  $(y, z) \in ]\tilde{a}, \tilde{b}[^2$ .

**Définition VI.3 (Mesure de vitesse).** La mesure de vitesse du brownien changé de temps  $(p(x_t))_{t \geq 0}$  défini en (1.3) est la mesure  $\tilde{M}(dx)$  de densité  $\tilde{m} = 2g^{-2}$  par rapport à la mesure de Lebesgue.

La mesure de vitesse du processus  $(x_t)_{t \geq 0}$  est la mesure image de  $\tilde{M}$  par  $p^{-1}$ . C'est la mesure de densité  $m = \frac{2}{\sigma^2 p'}$  par rapport à Lebesgue. ♣

Comme la mesure de vitesse de  $(x_t)_{t \geq 0}$  est la mesure image de la mesure de vitesse de  $(y_t)_{t \geq 0}$  par  $p^{-1}$ , on a d'après (1.5)

$$\forall y \in ]\tilde{a}, \tilde{b}[, \quad v(y) = \mathbf{E}_y \left[ \tilde{T}_{\tilde{a}} \wedge \tilde{T}_{\tilde{b}} \right] = \int_{p^{-1}(\tilde{a})}^{p^{-1}(\tilde{b})} G_{\tilde{a},\tilde{b}}(y, p(z)) m(z) dz.$$

En notant  $a = p^{-1}(\tilde{a})$  et  $b = p^{-1}(\tilde{b})$ , alors pour tout  $x \in ]a, b[$

$$v(p(x)) = \int_a^b G_{p(a),p(b)}(p(x), p(z)) m(z) dz.$$

D'autre part,  $p$  est bijective donc il est clair que  $\mathbf{E}_{p(x)} \left[ \tilde{T}_{p(a)} \wedge \tilde{T}_{p(b)} \right] = \mathbf{E}_x [T_a \wedge T_b]$ . Ainsi, en utilisant (1.2) on montre aisément que pour tout  $]a, b[ \subset I$  et tout  $x \in ]a, b[$

$$\begin{aligned} \mathbf{E}_x [T_a \wedge T_b] &= (1 - \mathbf{P}_x [T_b < T_a]) \int_a^x (p(y) - p(a)) m(y) dy \\ &\quad + \mathbf{P}_x [T_b < T_a] \int_x^b (p(b) - p(y)) m(y) dy. \end{aligned} \quad (1.6)$$

De plus, la fonction d'échelle et la mesure de vitesse permettent d'obtenir une expression très utile du générateur infinitésimal  $\mathcal{A}$  associé à l'EDS (1.1). En effet, on vérifie que

$$\forall f \in \mathcal{C}^2(I, \mathbf{R}), \quad \mathcal{A}f(x) = b(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x) = \frac{1}{m(x)} \left( \frac{f'(x)}{p'(x)} \right)'. \quad (1.7)$$

## 1.2. Classification de Feller

L'étude des processus de diffusion unidimensionnels est due en partie à Feller, notamment par ses travaux [Fel52] et [Fel54]. Parallèlement, l'école russe a établi des résultats similaires, mais les deux terminologies ne coïncident pas toujours. Pour une comparaison et une synthèse de ces travaux, on pourra consulter [KT81].

Une première notion qui permet de rendre compte du comportement du processus  $(x_t)_{t \geq 0}$  solution de (1.1) au voisinage d'un point frontière de  $I = ]l, r[$  est l'attractivité.

**Définition VI.4 (Attractivité).** Un point frontière  $\Delta$  ( $\Delta = l$  ou  $\Delta = r$ ) est dit *attractif* si  $\lim_{\substack{b \rightarrow \Delta \\ b \in I}} |p(b)| < +\infty$ . ♣

La fonction  $p$  est définie à une transformation affine strictement croissante près, mais le fait que la limite de  $p$  soit finie (ou infinie) en  $\Delta$  n'en dépend pas. De même, pour tout  $(x, y)$ ,

$|p(x) - p(y)| < +\infty$  et  $p$  est strictement croissante donc

$$\begin{aligned} l \text{ est attractif} &\Leftrightarrow \forall x \in I, \quad \lim_{b \rightarrow l} (p(x) - p(b)) < +\infty, \\ r \text{ est attractif} &\Leftrightarrow \forall x \in I, \quad \lim_{b \rightarrow r} (p(b) - p(x)) < +\infty. \end{aligned}$$

Ces équivalences sont parfois utilisées pour définir l'attractivité. Nous montrons maintenant la proposition suivante qui justifie l'appellation de "point attractif".

**Proposition VI.2:**

Si  $\Delta$  est un point frontière attractif alors pour tout  $a \in I$  et tout point  $x$  de l'intervalle ouvert d'extrémités  $a$  et  $\Delta$  on a

$$\mathbf{P}_x [T_\Delta \leq T_a] > 0,$$

où  $T_\Delta$  est défini par  $T_\Delta = \lim_{\substack{b \rightarrow \Delta \\ b \in I}} T_b$ . ◆

Preuve. On montre la proposition dans le cas où  $a < x < \Delta$ . Tout d'abord, comme  $(x_t)_{t \geq 0}$  est continue, on a par le théorème des valeurs intermédiaires la stricte croissance de la fonction  $(b \mapsto T_b)$  (pour  $b > x$ ). Donc  $T_\Delta$  est la limite strictement croissante de  $T_b$  lorsque  $b$  croît vers  $\Delta$ . Ainsi, on a

$$\bigcap_{\substack{b \in I \cap \mathbf{Q} \\ x < b < \Delta}} \downarrow \{T_b < T_a\} = \{T_\Delta \leq T_a\},$$

et  $\mathbf{P}_x [T_\Delta \leq T_a] = \lim_{b \rightarrow \Delta} \mathbf{P}_x [T_b < T_a]$ . Par (1.2) on en déduit que

$$\mathbf{P}_x [T_\Delta \leq T_a] = \lim_{b \rightarrow \Delta} \frac{p(x) - p(a)}{p(b) - p(a)}.$$

Or  $\Delta$  est attractif, donc  $\lim_{b \rightarrow \Delta} (p(b) - p(a)) < +\infty$ , d'où le résultat. #

**Définition VI.5 (Répulsivité).** Un point frontière  $\Delta$  est dit *répulsif* s'il n'est pas attractif, i.e.  $\lim_{\substack{b \rightarrow \Delta \\ b \in I}} |p(b)| = +\infty$ . ♣

Comme  $p$  est strictement croissante et que  $p$  est finie en tout point de  $I$ , il est clair que

$$\begin{aligned} l \text{ est répulsif} &\Leftrightarrow \forall x \in I, \quad \lim_{b \rightarrow l} (p(x) - p(b)) = +\infty, \\ r \text{ est répulsif} &\Leftrightarrow \forall x \in I, \quad \lim_{b \rightarrow r} (p(b) - p(x)) = +\infty. \end{aligned}$$

Ainsi, on vérifie d'après la proposition VI.2 que si  $\Delta$  est répulsif, alors pour tout  $a \in I$  et tout point  $x$  de l'intervalle d'extrémités  $a$  et  $\Delta$ ,  $\mathbf{P}_x [T_\Delta > T_a] = 1$ .

**Atteignabilité**

Si  $\Delta$  est un point frontière attractif alors, avec une probabilité non nulle, une trajectoire issue de  $x \in I$  passe par  $\Delta$  avant un autre point  $a \in I$ . Mais cet événement se produit-il en un temps fini? Oui, si le point  $\Delta$  est atteignable.

**Définition VI.6 (Atteignabilité).** Un point frontière  $\Delta$  est dit *atteignable* si pour tout  $a \in I$ , et pour tout  $x$  de l'intervalle ouvert d'extrémités  $a$  et  $\Delta$  on a

$$\lim_{b \rightarrow \Delta} \mathbf{E}_x [T_b \wedge T_a] < +\infty.$$

On peut montrer qu'un point frontière atteignable est attractif, et que si  $\Delta$  est attractif, alors  $\Delta$  est atteignable si et seulement si  $\mathbf{P}_x [T_\Delta < +\infty] > 0$  (cf. lemme 6.2 dans [KT81]). Un point non atteignable est dit *inatteignable*.

Ces deux notions, "attractivité" et "atteignabilité", permettent de déterminer le comportement de la diffusion au voisinage des points frontières.

D'autres notions interviennent dans la classification de Feller que nous n'évoquerons pas ici : point régulier (réflectif, absorbant, adhésif), point de sortie, point naturel, point d'entrée.

### 1.3. Comportement des mesures empiriques : le point de vue ergodique

Par la classification de Feller, on connaît entièrement le comportement trajectorien de la diffusion, mais si l'on s'intéresse aux mesures empiriques et à la récurrence du processus la notion de point frontière répulsif n'est pas assez précise. En effet, plusieurs situations peuvent se produire : les points infinis sont répulsifs pour le mouvement brownien et le processus d'Ornstein-Uhlenbeck, or le premier est récurrent nul (en dimension 1) et le second est récurrent positif. On introduit alors une nouvelle notion : la forte répulsivité.

**Définition VI.7.** Un point frontière répulsif  $\Delta$  est dit *fortement répulsif* si pour tout  $c \in I$  on a  $|\int_\Delta^c m(y)dy| < +\infty$  ♣

Un point fortement répulsif est donc un point répulsif qui a une mesure de vitesse finie au voisinage de  $\Delta$ . Il est important de définir cette notion à partir de la répulsivité car le fait que la mesure de vitesse soit finie dans tout voisinage de  $\Delta$  n'implique pas que la fonction d'échelle soit infinie en  $\Delta$  (cf. exemple suivant).

#### Exemples.

1. Soit  $I = ]0, +\infty[$  et  $b$  et  $\sigma$  continues sur  $\bar{I}$  définies sur  $[0, 1]$  par  $b(x) = \frac{1}{2}\sqrt{x}$  et  $\sigma(x) = cx^{3/4}$  avec  $c \in ]1, \sqrt{2}[$ . Alors

$$\forall x \in ]0, 1[, \quad p'(x) = \exp\left(-\int_1^x \frac{\sqrt{y}}{c^2 y^{3/2}} dy\right) = x^{-\frac{1}{c^2}},$$

et comme  $c > 1$ , on a pour tout  $x \in ]0, 1[, p(x) = \frac{1}{1-1/c^2} x^{1-\frac{1}{c^2}} - \frac{1}{1-1/c^2}$  donc  $\lim_{x \rightarrow 0} p(x) = -\frac{1}{1-1/c^2}$ . Le point 0 est donc attractif. De plus, la mesure de vitesse est finie au voisinage de 0 car

$$\int_0^1 m(y)dy = \int_0^1 \frac{2}{c^2 y^{3/2} y^{-1/c^2}} dy = \frac{2}{c^2} \frac{1}{1/c^2 - 1/2}, \quad (1.8)$$

et  $c \in ]1, \sqrt{2}[$ .

2. Dans le cas du processus d'Ornstein-Uhlenbeck défini sur  $\mathbf{R}$  par

$$dx_t = -\frac{1}{2}x_t dt + dB_t, \quad x_0 = x \in \mathbf{R}, \quad (1.9)$$

la mesure de vitesse  $m(x)dx$  est la probabilité gaussienne donc les points  $-\infty$  et  $+\infty$  sont fortement répulsifs. \*

On rappelle tout d'abord le résultat principal obtenu sur le comportement ergodique d'une solution  $(x_t)_{t \geq 0}$  de (1.1) en dimension 1. Un tel processus  $(x_t)_{t \geq 0}$  est dit récurrent si pour tous point  $a$  et  $b$  dans  $I$  on a  $\mathbf{P}_a [T_b < +\infty] = 1$ . De plus, on dit qu'il est récurrent positif si  $\mathbf{E}_a [T_b] < +\infty$  et  $\mathbf{E}_b [T_a] < +\infty$ , et récurrent nul sinon.

**Théorème VI.1:**

On suppose que  $(x_t)_{t \geq 0}$  (solution de (1.1) et de mesure de vitesse  $m$ ) est récurrent sur  $I$ . Soient  $f$  et  $g$  deux fonctions mesurables réelles positives telles que

$$\int_I (f(x) + g(x)) m(dx) < +\infty, \quad \int_I g(x)m(dx) \neq 0.$$

Alors

$$\frac{\int_0^t f(x_s) ds}{\int_0^t g(x_s) ds} \xrightarrow{p.s.} \frac{\int_I f(x)m(dx)}{\int_I g(x)m(dx)}. \quad \star$$

Pour une preuve détaillée, on peut consulter [RW94] ou [IM74]. Dans la suite, on note  $(\nu_t)_{t \geq 0}$  les mesures empiriques de la diffusion, *i.e.*

$$\forall t > 0, \quad \nu_t(dx) = \frac{1}{t} \int_0^t \delta_{x_s}(dx) ds$$

Muni de ce théorème et de la notion de point fortement répulsif, on établit la classification suivante du comportement ergodique de la diffusion.

**Théorème VI.2:**

On suppose que  $(x_t)_{t \geq 0}$  vit dans un intervalle  $I$  tel que  $\bar{I} = [l, r]$ .

- si  $l$  est attractif et  $r$  est répulsif, alors  $x_t^\zeta \xrightarrow{p.s.} l$ ,
- si  $l$  et  $r$  sont attractifs, alors

$$\mathbf{P} \left[ \lim_{t \rightarrow +\infty} x_t^\zeta = l \right] = 1 - \mathbf{P} \left[ \lim_{t \rightarrow +\infty} x_t^\zeta = r \right] = \frac{p(r^-) - p(x_0)}{p(r^-) - p(l^+)}.$$

- si  $l$  et  $r$  sont répulsifs, alors la diffusion est récurrente et n'explose pas ( $\zeta = +\infty$  *p.s.*).  
Plus précisément,

\* si  $l$  et  $r$  sont fortement répulsifs (*i.e.* la mesure de vitesse est finie), alors la diffusion est récurrente positive et  $\nu_t \Rightarrow \nu$  où  $\nu$  est la mesure de vitesse renormalisée. De plus,

$$\forall f \in \mathbf{L}^1(\nu), \quad \frac{1}{t} \int_0^t f(x_s) ds \Rightarrow \int_{\mathbf{R}} f(x) \nu(dx).$$

\* si  $l$  est fortement répulsif et  $r$  simplement répulsif, alors la diffusion est récurrente nulle et si les mesures empiriques sont tendues on a

$$\frac{1}{t} \int_0^t \delta_{x_s} ds \Rightarrow \delta_r.$$

\* si  $l$  et  $r$  ne sont pas fortement répulsifs, alors la diffusion est récurrente nulle et si les mesures empiriques sont tendues alors toute valeur d'adhérence de  $\left(\frac{1}{t} \int_0^t \delta_{x_s} ds\right)_{t \geq 0}$  est une mesure de support  $\{l, r\}$ . ★

**Lemme VI.1:**

Si les deux points frontières  $l$  et  $r$  sont répulsifs, alors la diffusion est récurrente positive si et seulement si sa mesure de vitesse est finie. ◆

Preuve. Par définition, la diffusion est récurrente positive si et seulement si pour tous points  $a$  et  $b$  de  $I$ ,  $\mathbf{E}_a [T_b] < +\infty$  et  $\mathbf{E}_b [T_a] < +\infty$ . Soient  $l < a < b < r$ . Par symétrie, il suffit de montrer que

$$\mathbf{E}_a [T_b] < +\infty \Leftrightarrow \int_l^a m(y) dy < +\infty. \quad (1.10)$$

Tout d'abord, comme  $l$  est un point répulsif et que  $T_l = \lim_{x \rightarrow l} T_x$  on a  $\mathbf{E}_a [T_b] = \mathbf{E}_a [T_b \wedge T_l] = \lim_{x \rightarrow l} \mathbf{E}_a [T_b \wedge T_x]$ . D'autre part, d'après (1.6) on a

$$\begin{aligned} \forall x \in ]l, a[, \quad \mathbf{E}_a [T_b \wedge T_x] &= \mathbf{P}_a [T_b < T_x] \int_a^b (p(b) - p(y)) m(y) dy \\ &\quad + \mathbf{P}_a [T_x \leq T_b] \int_x^a (p(y) - p(x)) m(y) dy, \end{aligned}$$

et comme  $\int_a^b (p(b) - p(y)) m(y) dy$  est une quantité finie ne dépendant pas de  $x$ ,  $\mathbf{E}_a [T_b \wedge T_x]$  a une limite finie lorsque  $x$  tend vers  $l$  si et seulement si

$$\lim_{x \rightarrow l} \left( \mathbf{P}_a [T_x \leq T_b] \int_x^a (p(y) - p(x)) m(y) dy \right) < +\infty.$$

Or,  $\mathbf{P}_a [T_x \leq T_b] = \frac{p(b) - p(a)}{p(b) - p(x)}$  donc

$$\begin{aligned} \mathbf{P}_a [T_x \leq T_b] \int_x^a (p(y) - p(x)) m(y) dy &= (p(b) - p(a)) \int_x^a \frac{p(y) - p(x)}{p(b) - p(x)} m(y) dy, \\ &= (p(b) - p(a)) \int_x^a \mathbf{P}_y [T_b < T_x] m(y) dy. \end{aligned}$$

Ainsi, on en déduit que  $\mathbf{E}_a [T_b] < +\infty$  si et seulement si  $\lim_{x \rightarrow l} \int_x^a \mathbf{P}_y [T_b < T_x] m(y) dy < +\infty$ . Comme  $T_x$  croît strictement vers  $T_l$ ,  $\mathbf{P}_y [T_b < T_x]$  croît vers  $\mathbf{P}_y [T_b < T_l]$  c'est à dire 1 car  $l$  est répulsif. On conclut alors par convergence monotone que  $\mathbf{E}_a [T_b]$  est finie si et seulement si  $\int_l^a m(y) dy$  est finie. #

Preuve (Théorème VI.2). Les deux premiers items sont prouvés dans [KS91] (proposition 5.22). On montre le premier item mais pas le second.

- On suppose  $l$  attractif et  $r$  répulsif. Par définition de la fonction d'échelle on a pour tout  $l < a < x < b < r$

$$\mathbf{P}_x \left[ \inf_{0 \leq t < \zeta} x_t \leq a \right] \geq \mathbf{P}_x [x_{T_a \wedge T_b} = a] = \frac{p(b) - p(x)}{p(b) - p(a)}.$$

En faisant croître  $b$  vers  $r$ , on obtient (du fait que  $r$  est répulsif)

$$\forall a \in ]l, r[, \quad \mathbf{P}_x \left[ \inf_{0 \leq t < \zeta} x_t \leq a \right] = 1.$$

La limite lorsque  $a$  décroît vers  $l$  donne alors

$$\mathbf{P}_x \left[ \inf_{0 \leq t < \zeta} x_t = l \right] = 1. \quad (1.11)$$

D'autre part, on a

$$\mathbf{P}_x \left[ \sup_{0 \leq t < \zeta} x_t = b \right] \leq \mathbf{P}_x [x_{T_l \wedge T_b} = b] = \lim_{a \rightarrow l} \frac{p(x) - p(a)}{p(b) - p(a)},$$

et en prenant la limite lorsque  $b$  croît vers  $r$  on obtient que  $\sup_{0 \leq t < \zeta} x_t < r$  p.s. Il reste à montrer que le processus  $(x_t^\zeta)_{t \geq 0}$  converge presque sûrement. On sait que  $(p(x_t^\zeta))_{t \geq 0}$  est une martingale locale et que  $l$  est attractif donc  $(p(x_t^\zeta) - \lim_{a \rightarrow l} p(a))_{t \geq 0}$  est une martingale locale continue positive et par Fatou une surmartingale continue positive. Elle converge alors presque sûrement. La fonction  $p$  a une réciproque continue donc la diffusion arrêtée en  $\zeta$  converge presque sûrement vers  $l$ .

- Si  $l$  et  $r$  sont deux points répulsifs, alors de la même façon que l'on obtient (1.11) on a

$$\inf_{0 \leq t < \zeta} x_t = l \text{ p.s.} \quad \text{et} \quad \sup_{0 \leq t < \zeta} x_t = r \text{ p.s.}$$

La diffusion est donc récurrente sur  $I$  et n'explose pas i.e.  $\zeta = +\infty$  p.s. De plus, on sait par le lemme VI.1 que la diffusion est récurrente positive si et seulement si la mesure de vitesse est finie.

– Si les deux points frontières sont fortement répulsifs alors la mesure de vitesse est finie, et par le théorème VI.1 on a

$$\forall f \in \mathbf{L}^1(\nu), \quad \frac{1}{t} \int_0^t f(x_s) ds \Rightarrow \int_{\mathbf{R}} f d\nu,$$

où  $\nu$  est la mesure de vitesse renormalisée.

– Si  $l$  est fortement répulsif et  $r$  simplement répulsif, alors la diffusion est récurrente nulle. En considérant une suite croissante de fonctions continues à support compact  $(g_n(x))_{n \geq 1}$  telle que  $g_n(x) \rightarrow 1$  et  $\forall n \geq 1, \int_{\mathbf{R}} g_n(x) m(dx) \neq 0$ , on obtient par le théorème VI.1

$$\forall f \in \mathbf{L}^1(m), \quad \frac{1}{t} \int_0^t f(x_s) ds \xrightarrow{p.s.} 0. \quad (1.12)$$

D'autre part, on considère une sous-suite  $(\nu_{a(t)})_{t \geq 0}$  de  $(\nu_t)_{t \geq 0}$  qui converge étroitement vers une mesure  $\nu$  (les mesures empiriques sont tendues donc une telle suite existe). Soit  $f$  une fonction continue à support compact telle que  $\text{supp}(f) \subset [l, r[$ . Comme  $\nu_{a(t)} \Rightarrow \nu$ , on a

$$\frac{1}{a(t)} \int_0^{a(t)} f(x_s) ds \xrightarrow{p.s.} \int f d\nu.$$

Le point frontière  $l$  est fortement répulsif et  $\text{supp}(f) \subset [l, r[$ , donc  $f$  est intégrable par rapport à  $m$ , et d'après (1.12) on a alors  $\int f d\nu = 0$ . L'intervalle  $[l, r[$  est le plus grand ouvert  $O$  de  $\bar{I}$  tel que

$$\forall f \in \mathcal{C}_c(\bar{I}), \text{supp}(f) \subset O \Rightarrow \nu(f) = 0,$$

donc le support de  $\nu$  est  $\{r\}$ . Comme  $\nu$  est normalisée, on en déduit que  $\nu = \delta_r$ . On obtient ainsi la convergence étroite des mesures empiriques vers  $\delta_r$ .

– De la même façon, si deux points frontières ne sont pas fortement répulsifs alors la mesure limite d'une sous-suite des mesures empiriques a pour support  $\{l, r\}$ . #

#### 1.4. Attractivité et fonctions de Lyapounov

Dans la littérature, les liens entre la notion d'attractivité (ou de répulsivité) et la notion de stabilité (cf. définition II.1) ont été peu étudiés. Cela semble naturel, car la classification de Feller est un outil uniquement unidimensionnel, et les fonctions de Lyapounov ont surtout un intérêt en dimension grande. Cependant, afin d'étendre le concept d'attractivité aux dimensions supérieures et d'étudier le comportement de notre schéma aux point frontières, il est utile de faire ce lien.

On rappelle que  $I$  est l'intervalle de référence dans lequel vit la diffusion  $(x_t)_{t \geq 0}$ . Dans la suite, on note  $J_\Delta \subset I$  un voisinage époinché de  $\Delta$  i.e. un intervalle ouvert non vide inclus dans  $I$  et ayant le point  $\Delta$  comme extrémité.

##### **Proposition VI.3:**

Soit  $\Delta$  un point frontière (fini ou infini, borne gauche ou droite) de l'intervalle  $I$  dans lequel vit  $(x_t)_{t \geq 0}$ . On a les équivalences suivantes :

1.  $\Delta$  est un point répulsif de  $I$  si et seulement s'il existe un voisinage  $J_\Delta \subset I$  de  $\Delta$  et une fonction  $V \in \mathcal{C}^2(J_\Delta, \mathbf{R}_+)$  strictement monotone tels que

$$\lim_{x \rightarrow \Delta} V(x) = +\infty \quad \text{et} \quad \forall x \in J_\Delta, \quad \mathcal{A}V(x) \leq 0,$$

2.  $\Delta$  est un point fortement répulsif de  $I$  si et seulement s'il existe un voisinage  $J_\Delta \subset I$  de  $\Delta$  et une fonction  $V \in \mathcal{C}^2(J_\Delta, \mathbf{R}_+)$  strictement monotone tels que

$$\exists \varepsilon > 0, \forall x \in J_\Delta, \quad \mathcal{A}V(x) \leq -\varepsilon$$

3.  $\Delta$  est un point attractif de  $I$  si et seulement s'il existe un voisinage  $J_\Delta \subset I$  de  $\Delta$  et

une fonction  $V \in \mathcal{C}^2(J_\Delta, \mathbf{R}_+)$  strictement monotone tels que

$$\sup_{x \in J_\Delta} V(x) = V(\Delta) < +\infty \quad \text{et} \quad \forall x \in J_\Delta, \quad \mathcal{A}V(x) \geq 0, \quad \blacklozenge$$

Preuve. On prouve cette proposition lorsque  $\Delta$  est un point frontière à droite, l'autre cas étant similaire. Ainsi le voisinage  $J_\Delta$  est un intervalle de la forme  $]c, \Delta[$ ,  $c \in I$ .

- – Supposons qu'il existe un voisinage  $J_\Delta$  de  $\Delta$  et une fonction  $V \in \mathcal{C}^2(J_\Delta, \mathbf{R}_+)$  telle que  $\lim_{x \rightarrow \Delta} V(x) = +\infty$  et  $\mathcal{A}V \leq 0$  sur  $J_\Delta$ . Pour tout  $x \in J_\Delta$ , on a

$$\mathcal{A}V(x) = \frac{1}{m(x)} \left( \frac{V'(x)}{p'(x)} \right)' \leq 0$$

donc  $V'/p'$  décroît sur  $J_\Delta$ . Il existe alors  $C > 0$  tel que  $V'(x) \leq Cp'(x)$  pour tout  $x \in ]c, \Delta[$ . On en déduit aisément que  $\lim_{x \rightarrow \Delta} p(x) - p(c) = +\infty$  car  $V$  tend vers l'infini en  $\Delta$ .

– Pour la réciproque, il faut trouver la bonne fonction de Lyapounov  $V$ . Soit  $c > 0$  tel que  $p(c) > 0$  ( $c$  existe car  $\Delta$  est répulsif donc  $\lim_{x \rightarrow \Delta} p(x) = +\infty$ ). Par stricte croissance de  $p$ , on a pour tout  $x \in ]c, \Delta[$ ,  $p(x) > p(c) > 0$ . On définit alors la fonction  $V$  sur  $]c, \Delta[$  par

$$\forall x \in ]c, \Delta[, \quad V(x) = p(x) - p(c).$$

Le point  $\Delta$  est répulsif donc  $V$  croît vers l'infini lorsque  $x$  tend vers  $\Delta$ . De plus,  $V \in \mathcal{C}^2(]c, \Delta[, \mathbf{R}_+)$  et  $\mathcal{A}V = 0$ .

- – Soit  $J_\Delta$ ,  $V \in \mathcal{C}^2(J_\Delta, \mathbf{R}_+)$  infinie en  $\Delta$ , et  $\varepsilon > 0$  tels que  $\mathcal{A}V \leq -\varepsilon$ . Alors on a

$$\int_{J_\Delta} \mathcal{A}V(x)m(x)dx \leq -\varepsilon \int_{J_\Delta} m(x)dx, \quad (1.13)$$

et comme  $\mathcal{A}V(x) = \frac{1}{m(x)} \left( \frac{V'(x)}{p'(x)} \right)'$ , on a pour tout point  $c \in J_\Delta$ ,

$$\int_c^\Delta \mathcal{A}V(x)m(x)dx = \lim_{x \rightarrow \Delta} \frac{V'(x)}{p'(x)} - \frac{V'(c)}{p'(c)}. \quad (1.14)$$

Donc d'après (1.13), on a

$$\int_c^\Delta m(x)dx \leq \frac{1}{\varepsilon} \left( \frac{V'(c)}{p'(c)} - \lim_{x \rightarrow \Delta} \frac{V'(x)}{p'(x)} \right),$$

et comme  $\lim_{x \rightarrow \Delta} \frac{V'(x)}{p'(x)} \geq 0$  (car  $V$  et  $p$  sont croissantes sur  $J_\Delta$ ), on obtient  $\int_c^\Delta m(x)dx \leq C$  c'est à dire la forte répulsivité de  $\Delta$ .

– On montre maintenant la réciproque et on suppose que le point  $\Delta$  est fortement répulsif *i.e.* pour tout  $c \in I$ ,  $\int_c^\Delta m(x)dx < +\infty$ . Soit  $c \in I$  et  $V$  la fonction définie sur  $]c, \Delta[$  par

$$\forall x \in ]c, \Delta[, \quad V(x) = \int_c^x \left( p'(y) \int_y^\Delta m(z)dz \right) dy.$$

Il est clair que  $V \in \mathcal{C}^2(]c, \Delta[, \mathbf{R}_+)$  et que pour tout  $x \in ]c, \Delta[, V'(x) = p'(x) \int_x^\Delta m(z) dz$ . De plus,

$$\forall x \in ]x, \Delta[, \quad \mathcal{A}V(x) = -1$$

- On prouve 3. de la même façon que 1. Le sens direct utilise l'expression du générateur  $\mathcal{A}f = \frac{1}{m} \left( \frac{f'}{p'} \right)'$ , et pour la réciproque on considère la fonction  $V(x) = p(x) - p(c)$  sur  $]c, \Delta[$  avec  $c$  tel que  $p(c) > 0$ . #

La proposition précédente permet d'obtenir un critère simple à vérifier pour déterminer la nature d'un point frontière. Ce critère est notamment plus facile à vérifier lorsque  $\Delta$  est un point fini. De plus, les conditions obtenues sont proches des conditions classiques utilisées pour définir la nature d'un point critique d'une équation différentielle ordinaire.

**Corollaire VI.1:**

Soit  $\Delta$  un point frontière de  $I$ .

1.  $\Delta$  est un point répulsif de  $I$  si et seulement s'il existe un voisinage  $J_\Delta \subset I$  de  $\Delta$  et une fonction  $v \in \mathcal{C}^2(\bar{J}_\Delta, \mathbf{R}_+)$  strictement monotone, nulle en  $\Delta$ , tels que

$$\forall x \in J_\Delta, \quad \mathcal{A}v(x) \geq \frac{1}{2} \sigma^2(x) \frac{(v'(x))^2}{v(x)}, \quad (1.15)$$

2.  $\Delta$  est un point fortement répulsif de  $I$  si et seulement s'il existe un voisinage  $J_\Delta \subset I$  de  $\Delta$  et une fonction  $v \in \mathcal{C}^2(\bar{J}_\Delta, \mathbf{R}_+)$  strictement monotone, minimale en  $\Delta$ , tels que

$$\exists \varepsilon > 0, \quad \forall x \in J_\Delta, \quad \mathcal{A}v(x) \geq \frac{1}{2} \sigma^2(x) \frac{(v'(x))^2}{v(x)} + \varepsilon v(x), \quad (1.16)$$

3.  $\Delta$  est un point attractif de  $I$  si et seulement s'il existe un voisinage  $J_\Delta \subset I$  de  $\Delta$  et une fonction  $v \in \mathcal{C}^2(\bar{J}_\Delta, \mathbf{R}_+)$  strictement monotone, minimale en  $\Delta$ , tels que

$$\forall x \in J_\Delta, \quad \mathcal{A}v(x) < \frac{1}{2} \sigma^2(x) \frac{(v'(x))^2}{v(x)}. \quad (1.17)$$

•

Preuve. Soit  $J_\Delta$  un voisinage de  $\Delta$  inclus (strictement) dans  $I$ . On prouve le corollaire dans le cas où  $\Delta$  est le point frontière droite de  $I$ , donc  $J_\Delta = ]c, \Delta[$  avec  $c \in I$ .

Soit  $L > 0$  et  $\phi_L$  la fonction définie par

$$\begin{aligned} \phi_L : [0, \exp(L)[ &\rightarrow \mathbf{R}_+, \\ x &\mapsto -\ln(x) + L. \end{aligned}$$

Alors  $\phi_L$  est une bijection strictement décroissante et de classe  $\mathcal{C}^\infty$ . De plus, pour toute fonction  $v$  de classe  $\mathcal{C}^2$  à valeurs dans  $[0, \exp(L)[$  on a

$$\mathcal{A}(\phi_L \circ v) = -\frac{\mathcal{A}v}{v} + \frac{1}{2} \sigma^2 \frac{(v')^2}{v^2}. \quad (1.18)$$

- – Montrons le premier item. On suppose qu'il existe une fonction  $v \in \mathcal{C}^2([c, \Delta], \mathbf{R}_+)$  strictement monotone nulle en  $\Delta$  et vérifiant (1.15). Alors, sur  $]c, \Delta[$  on définit la fonction  $V$  par  $\forall x \in ]c, \Delta[, V(x) = (\phi_L \circ v)(x)$  avec  $L = \ln(v(c))$ . C'est une fonction strictement monotone qui tend vers l'infini lorsque  $x$  tend vers  $\Delta$ , et d'après (1.15) et (1.18) on a  $\mathcal{A}V \leq 0$  sur  $]c, \Delta[$ . Donc par la proposition (VI.3), le point  $\Delta$  est répulsif.
  - Réciproquement, si le point est répulsif alors il existe  $V \in \mathcal{C}^2(]c, \Delta[, \mathbf{R}_+)$  strictement monotone vers  $+\infty$  lorsque  $x$  tend vers  $\Delta$ . On pose alors  $v = \phi_L^{-1} \circ V$  sur  $]c, \Delta[$  avec  $L = \inf_{x \in ]c, \Delta[} V(x)$ , et on la prolonge par continuité sur  $[c, \Delta]$  en posant  $v(c) = 1$  et  $v(\Delta) = 0$ . Par  $\mathcal{A}V \leq 0$  et (1.18), on en déduit que  $\mathcal{A}v \geq \frac{1}{2}\sigma^2 \frac{(v')^2}{v}$  sur  $]c, \Delta[$ .
- – Soit  $v$  strictement monotone sur  $[c, \Delta]$  et nulle en  $\Delta$ . Pour  $\theta \geq 0$  et  $L = \ln(v(c) + \theta)$  on pose  $V(x) = (\phi_{L, \theta} \circ v)(x)$  pour tout  $x \in ]c, \Delta[$ . Alors  $V$  est strictement monotone sur  $]c, \Delta[$  et admet une limite (finie ou infinie) en  $\Delta$ . Des relations (1.16) et (1.18), on en déduit que

$$\forall x \in ]c, \Delta[, \quad \mathcal{A}V(x) \leq -\varepsilon.$$

- Pour la réciproque, on considère la fonction  $v = \phi_L^{-1} \circ V$  sur  $]c, \Delta[$  avec  $L = \inf_{x \in ]c, \Delta[} V(x)$ , prolongée par continuité par  $v(c) = 1$  et  $v(\Delta) = \lim_{x \rightarrow \Delta} \exp(-V(x))$ .
- Le dernier point se démontre de façon analogue aux deux autres.

### Lien avec l'équation différentielle sous-jacente

Les rappels qui suivent sont contenus dans [HW99]. On se place en dimension 1 par homogénéité avec le début de la section, mais tout ce qui suit a un intérêt principalement en dimension supérieure.

Soit  $u' = f(u)$  une équation différentielle réelle admettant un point critique  $\Delta$  (*i.e.*  $f(\Delta) = 0$ ). On rappelle qu'une fonction  $F$  à valeurs réelles, de classe  $\mathcal{C}^1$ , définie sur un voisinage  $\mathcal{U}_\Delta$  de  $\Delta$  est une fonction de Lyapounov sur  $\mathcal{U}_\Delta$  pour l'équation  $u' = f(u)$  si

1.  $F$  possède un unique minimum en  $\Delta$ ,
2. pour tout  $u \in \mathcal{U}_\Delta \setminus \{\Delta\}$ ,  $F'f(u) < 0$ .

La fonction  $F$  est donc décroissante le long des trajectoires de  $u' = f(u)$ . On a le théorème suivant :

#### **Théorème VI.3:**

*Si une équation différentielle admet une fonction de Lyapounov au voisinage d'un de ses points critiques, alors ce point critique est un puits, c'est à dire qu'il existe un voisinage  $\mathcal{U}_\Delta$  de  $\Delta$  tel que si  $u(t)$  est une solution vérifiant  $u(t_0) \in \mathcal{U}_\Delta$ , alors  $u(t) \in \mathcal{U}_\Delta$  pour tout  $t \geq t_0$  et  $\lim_{t \rightarrow +\infty} u(t) = \Delta$ . ★*

Un puits est aussi appelé point critique asymptotiquement stable. De même, on peut définir une *source* comme un point critique  $\Delta$  pour lequel il existe un voisinage  $\mathcal{U}_\Delta$  et  $\Delta$  et une fonction  $F$  définie sur  $\mathcal{U}_\Delta$ , admettant un unique minimum en  $\Delta$  et vérifiant :  $\forall u \in \mathcal{U}_\Delta$ ,  $F'f(u) \geq 0$ .

À la vue du corollaire VI.1, il existe donc un lien étroit entre la nature du point critique  $\Delta$  pour l'EDO  $u' = b(u)$  et la nature du point frontière  $\Delta$  pour l'EDS  $dx_t = b(x_t)dt + \sigma(x_t)dB_t$ .

En effet, si  $\Delta$  est un puits et que sa fonction de Lyapounov  $F$  associée est  $\mathcal{C}^2$  au voisinage de  $\Delta$  et telle que  $F'/F$  soit strictement décroissante, et si  $\sigma(\Delta) = 0$ , alors  $\Delta$  est un point frontière attractif pour la diffusion (on applique le corollaire VI.1 avec  $v = F$ ). Si  $\Delta$  est une source (ou point d'équilibre instable) plusieurs situations peuvent se produire : si  $\sigma(\Delta) = 0$ ,  $\Delta$  est un point frontière qui peut être répulsif, fortement répulsif et même attractif (cf. exemple ci-dessous).

**Exemple.** On considère la fonction  $V : \mathbf{R} \rightarrow \mathbf{R}_+$  définie par

$$V(x) = \begin{cases} (x - 3 \operatorname{sgn}(x))^2 & \text{si } |x| \geq 3, \\ \frac{1}{72}(x^2 - 9)^2 & \text{si } |x| \leq 3 \end{cases} \quad \text{et} \quad b(x) = \begin{cases} -2(x - 3 \operatorname{sgn}(x)) & \text{si } |x| \geq 3, \\ -\frac{1}{18}x^3 + \frac{1}{2}x & \text{si } |x| \leq 3 \end{cases},$$

de sorte que  $b = -V'$ . L'équation différentielle  $u' = b(u)$  a 3 points critiques :  $-3, 0$  et  $3$ . Les points  $-3$  et  $3$  sont des puits et le point  $0$  est une source.

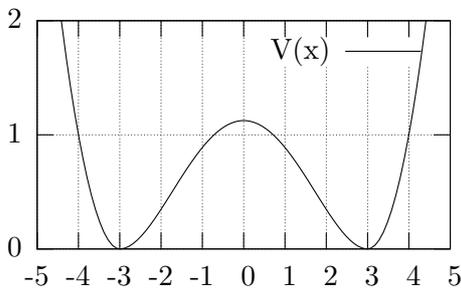


Figure VI.1 – Fonction  $V$

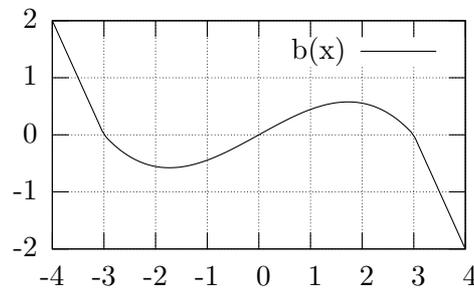


Figure VI.2 – Drift  $b = -V'$

On pose  $\sigma(x) = cx$  où  $c \in ]0, 2[$  est un paramètre permettant de régler la taille de  $\sigma$  au voisinage de  $0$ , et on considère le processus  $(x_t)_{t \geq 0}$  solution de l'EDS  $dx_t = b(x_t)dt + \sigma(x_t)dB_t$ . Il est clair que le point  $0$  est un point frontière pour  $(x_t)_{t \geq 0}$ . De plus, il est facile de vérifier que

$$\mathcal{A}V(x) = -(4 - c^2)x^2 + 12 \operatorname{sgn}(x) \quad \text{si } |x| \geq 3,$$

ce qui implique, d'après la proposition VI.3, que les points  $-\infty$  et  $+\infty$  sont fortement répulsifs. D'autre part, on utilise le corollaire VI.1 pour déterminer la nature du point  $0$  en fonction de  $c$ . La situation étant symétrique pour  $] - \infty, 0[$  et  $]0, +\infty[$  on suppose que  $I = ]0, +\infty[$ . Soit  $v$  la fonction nulle en  $0$  définie sur  $]0, +\infty[$  par  $v(x) = x^2$ . Alors on a

$$\forall x \in ]0, 3[, \quad \mathcal{A}v(x) = (1 + c^2)x^2 - \frac{1}{9}x^4 \quad \text{et} \quad \frac{1}{2}\sigma^2(x) \frac{(v'(x))^2}{v(x)} = 2c^2x^2. \quad (1.19)$$

On en déduit que

- si  $c < 1$  le point  $0$  est fortement répulsif (pour  $]0, +\infty[$  mais aussi pour  $] - \infty, 0[$  car la situation est symétrique). En effet, la condition (1.16) est vérifiée pour  $\varepsilon = \frac{1-c^2}{2}$  et  $J_\Delta = ]0, 3\sqrt{\frac{1-c^2}{2}}[$ .
- si  $c > 1$  le point  $0$  est attractif.
- si  $c = 1$ , on considère la fonction  $v(x) = x \exp(x)$ , et on vérifie que  $0$  est un point répulsif.

Ainsi, la nature du point frontière 0 (qui est un point source pour l'EDO sous-jacente) change en fonction de  $c$  (la force du bruit mis sur les deux autres points critiques de l'EDO). Par le théorème VI.2, il est alors aisé de connaître le comportement ergodique de  $(x_t)_{t \geq 0}$  :

- si  $c \in ]1, 2[$ , alors  $x_t \xrightarrow{p.s.} 0$  (pour tout point de départ  $x_0$ ),
- si  $c = 1$ , alors  $\frac{1}{t} \int_0^t \delta_{x_s}(dx) \Rightarrow \delta_0$ ,
- si  $c < 1$ , alors pour toute fonction  $f \in \mathbf{L}^1(m)$

$$\frac{1}{t} \int_0^t f(x_s) ds \xrightarrow{p.s.} \begin{cases} \int f d\nu_- & \text{si } x_0 \in ]0, +\infty[ \\ f(0) & \text{si } x_0 = 0 \\ \int f d\nu_+ & \text{si } x_0 \in ]-\infty, 0[ \end{cases}$$

où  $\nu_-$  est la mesure invariante de  $(x_t)_{t \geq 0}$  portée sur  $] -\infty, 0[$  et  $\nu_+$  la mesure invariante sur  $]0, +\infty[$  (les mesures  $\nu_+$  et  $\nu_-$  sont les mesures de vitesse renormalisées donc des probabilités). \*

## 2. Comportement de l'algorithme en dimension 1

On considère le schéma d'Euler  $(X_n)_{n \geq 0}$  construit à partir de la suite de pas  $(\gamma_n)_{n \geq 1}$ , suite convergente vers 0 et terme général d'une série divergente. Dans toute la suite, le bruit blanc  $(U_n)_{n \geq 1}$  utilisé pour construire  $(X_n)_{n \geq 0}$  est supposé sous-gaussien (cf. chapitre III).

On suppose que la diffusion  $(x_t)_{t \geq 0}$  vérifie une condition de stabilité forte de type  $(V, \text{Id}, \alpha)$  où  $V$  est sous-quadratique *i.e.*

$$\exists C_V > 0, \quad |V'|^2 \leq C_V V, \quad \text{et} \quad \sup_{x \in \mathbf{R}} |V''(x)| < +\infty. \quad (2.1)$$

D'après la proposition VI.3, la condition de stabilité implique que les points  $-\infty$  et  $+\infty$  sont fortement répulsifs pour la diffusion. Toute l'étude pourrait se faire sous une hypothèse de stabilité faible mais on se contentera d'exposer le cas de stabilité forte.

On suppose que  $b$  et  $\sigma$  sont continues sur  $\mathbf{R}$  et vérifient les conditions de domination suivante

$$\exists C_b > 0, \quad |b|^2 \leq C_b V, \quad \text{et} \quad \exists C_\sigma > 0, \quad |\sigma|^2 \leq C_\sigma V, \quad (2.2)$$

alors on sait que les mesures empiriques pondérées  $(\nu_n^n)_{n \geq 1}$  du schéma d'Euler sont tendues, et que toute valeur d'adhérence est une probabilité invariante pour la diffusion  $(x_t)_{t \geq 0}$ . Le coefficient de diffusion  $\sigma$  est supposé dégénéré en un point  $\Delta$  qui annule aussi  $b$ . Ainsi comme on l'a vu précédemment, le point  $\Delta$  est un point frontière pour la diffusion. Par contre le schéma d'Euler peut *a priori* traverser ce point indéfiniment. Le théorème d'identification de la limite entraîne alors la convergence des mesures empiriques vers une combinaison convexe des mesures invariantes de  $(x_t)_{t \geq 0}$ . On va montrer que, sous certaines conditions sur la nature du point frontière et sur  $(\gamma_n)_{n \geq 1}$ , cette situation ne se produit pas : le schéma reste –à partir d'un rang  $n_1$ – d'un côté  $I$  du point  $\Delta$  et les mesures empiriques convergent étroitement vers la mesure invariante portée par cet intervalle  $I$ .

## 2.1. Schéma d'Euler

Avant d'étudier  $(\nu_n^n)_{n \geq 1}$ , on donne le résultat principal concernant le comportement du schéma d'Euler au voisinage d'un point frontière.

**Théorème VI.4:**

Soit  $\Delta$  l'unique point de  $\mathbf{R}$  vérifiant  $b(\Delta) = \sigma(\Delta) = 0$ . Le coefficient  $\sigma$  est non dégénéré sur  $] - \infty, \Delta[ \cup ] \Delta, +\infty[$ . On suppose qu'il existe  $\mathcal{U}_\Delta$  de la forme  $\mathcal{U}_\Delta = ]\Delta - \varepsilon, \Delta + \varepsilon[$  avec  $\varepsilon > 0$ , et une fonction  $v \in \mathcal{C}^2(\mathcal{U}_\Delta, \mathbf{R})$  convexe, admettant un unique minimum  $v_*$  en  $\Delta$ , tels que

$$\forall x \in \mathcal{U}_\Delta, \quad (v'b)(x) \geq 0 \quad \text{et} \quad |(v'\sigma)(x)| \leq c_\sigma(v(x) - v_*). \quad (2.3)$$

Si la suite de pas  $(\gamma_n)_{n \geq 1}$  vérifie :  $\forall C > 0, \sum_{n \geq 1} \exp\left(-\frac{C}{\gamma_n}\right) < +\infty$ , alors le schéma d'Euler traverse un nombre fini de fois le point frontière  $\Delta$  i.e.

$$\mathbf{P}[\exists n_0 \geq 0, \forall n \geq n_0, X_n \in ] - \infty, \Delta[ ] + \mathbf{P}[\exists n_0 \geq 0, \forall n \geq n_0, X_n \in ] \Delta, +\infty[ ] = 1. \quad \star$$

Preuve. Soit  $A_{n+1}$  l'événement  $\{\Delta \in [X_n, X_{n+1}]\}$ . On va montrer que

$$\sum_{n \geq 1} \mathbf{P}[A_{n+1} | \mathcal{F}_n] < +\infty,$$

afin d'appliquer le lemme de Borel-Cantelli conditionnel.

- On se place tout d'abord sur  $E_n = \{X_n \in \mathcal{U}_\Delta\}$ . Tout d'abord, par continuité de  $v$

$$\begin{aligned} A_{n+1} &= \{\exists t \in [0, 1], X_n + t(X_{n+1} - X_n) = \Delta\}, \\ &\subset \{\exists t \in [0, 1], v(X_n + t(X_{n+1} - X_n)) = v(\Delta)\}, \end{aligned} \quad (2.4)$$

Soit  $t \in [0, 1]$  tel que  $\Delta = X_n + t(X_{n+1} - X_n)$ . Comme  $v$  est de classe  $\mathcal{C}^2$  sur  $\mathcal{U}_\Delta$ , la formule de Taylor donne

$$v(X_n + t(X_{n+1} - X_n)) = v(X_n) + v'(X_n)t(X_{n+1} - X_n) + \frac{v''(\xi_{n+1})}{2}t^2(X_{n+1} - X_n)^2,$$

avec  $\xi_{n+1} \in (X_n, X_{n+1})$ , et la convexité de  $v$  entraîne

$$v(X_n + t(X_{n+1} - X_n)) \geq v(X_n) + t\gamma_{n+1}(v'b)(X_n) + t\sqrt{\gamma_{n+1}}(v'\sigma)(X_n)U_{n+1}. \quad (2.5)$$

De plus, on a  $v'b \geq 0$  sur  $\mathcal{U}_\Delta$  donc d'après (2.4) et (2.5)

$$\begin{aligned} A_{n+1} \cap E_n &\subset \{\exists t \in [0, 1], v(X_n) + t\sqrt{\gamma_{n+1}}(v'\sigma)(X_n)U_{n+1} \leq v(\Delta)\} \cap E_n, \\ &\subset \{\exists t \in [0, 1], t\sqrt{\gamma_{n+1}}|(v'\sigma)(X_n)U_{n+1}| \geq v(X_n) - v(\Delta)\} \cap E_n \end{aligned}$$

Ainsi on a pour tout  $n \geq 0$ ,

$$\begin{aligned} \mathbf{P}[A_{n+1} \cap E_n \mid \mathcal{F}_n] &\leq \mathbf{P} \left[ \left\{ |U_{n+1}| \geq \frac{v(X_n) - v(\Delta)}{\sqrt{\gamma_{n+1}} |v'\sigma(X_n)|} \right\} \cap E_n \mid \mathcal{F}_n \right], \\ &\leq \mathbf{P} \left[ \left\{ |U_{n+1}| \geq \frac{1}{c_\sigma \sqrt{\gamma_{n+1}}} \right\} \cap E_n \mid \mathcal{F}_n \right], \end{aligned} \quad (2.6)$$

d'après l'hypothèse de domination (2.3) sur  $v'\sigma$ .

- On se place maintenant sur l'événement complémentaire  $E_n^c = \{X_n \notin \mathcal{U}_\Delta\}$ . Alors on a

$$\begin{aligned} A_{n+1} &= \left\{ \exists t \in [0, 1], U_{n+1} = \frac{\Delta - X_n}{t \sqrt{\gamma_{n+1}} \sigma(X_n)} - \sqrt{\gamma_{n+1}} \frac{b(X_n)}{\sigma(X_n)} \right\}, \\ &\subset \left\{ |U_{n+1}| \geq \frac{|\Delta - X_n|}{\sqrt{\gamma_{n+1}} |\sigma(X_n)|} - \sqrt{\gamma_{n+1}} \frac{|b(X_n)|}{|\sigma(X_n)|} \right\}. \end{aligned}$$

Le drift  $b$  est dominé par  $\sqrt{V}$  donc

$$A_{n+1} \subset \left\{ |U_{n+1}| \geq \left( \frac{|\Delta - X_n|}{\sqrt{\gamma_{n+1}} \sqrt{V}(X_n)} - \sqrt{\gamma_{n+1}} \sqrt{C_b} \right) \frac{\sqrt{V}(X_n)}{|\sigma(X_n)|} \right\},$$

et comme  $\sqrt{V}$  est Lipschitz, on a sur  $E_n^c$

$$\frac{|\Delta - X_n|}{\sqrt{V}(X_n)} \geq \frac{|\Delta - X_n|}{\sqrt{V}(\Delta) + C|\Delta - X_n|} \geq \frac{1}{C + \frac{\sqrt{V}(\Delta)}{\varepsilon}}.$$

Ainsi, il existe  $n_1 \geq 0$  et  $C > 0$  tel que pour tout  $n \geq n_1$ ,

$$\begin{aligned} \mathbf{P}[A_{n+1} \cap E_n^c \mid \mathcal{F}_n] &\leq \mathbf{P} \left[ \left\{ |U_{n+1}| \geq \frac{C}{\sqrt{\gamma_{n+1}}} \frac{\sqrt{V}(X_n)}{|\sigma(X_n)|} \right\} \cap E_n^c \mid \mathcal{F}_n \right], \\ &\leq \mathbf{P} \left[ \left\{ |U_{n+1}| \geq \frac{C}{\sqrt{C_\sigma} \sqrt{\gamma_{n+1}}} \right\} \cap E_n^c \mid \mathcal{F}_n \right], \end{aligned} \quad (2.7)$$

car  $|\sigma| \leq \sqrt{C_\sigma} \sqrt{V}$ .

- En combinant (2.6) et (2.7), on obtient

$$\exists n_1 \geq 0, \exists C > 0, \forall n \geq n_1, \quad \mathbf{P}[A_{n+1} \mid \mathcal{F}_n] \leq \mathbf{P} \left[ |U_1| \geq \frac{C}{\sqrt{\gamma_{n+1}}} \right],$$

et comme  $U_1$  est sous-gaussien, on a

$$\sum_{n \geq n_1} \mathbf{P}[A_{n+1} \mid \mathcal{F}_n] \leq 2 \sum_{n \geq n_1} \exp \left( -\frac{C^2}{2\gamma_{n+1}} \right).$$

Par la condition sur les pas et par le lemme de Borel-Cantelli conditionnel on en déduit que l'événement  $A_n$  ne se produit qu'un nombre fini de fois. #

Remarque: La condition sur la suite de pas  $(\gamma_n)_{n \geq 1}$  n'est pas restrictive car elle permet de considérer toute la gamme des pas puissances :  $\gamma_n = \gamma_0 n^{-r}$  avec  $\gamma_0 > 0$  et  $r \in ]0, 1]$ . On peut

aussi considérer des pas de la forme  $\gamma_n = \log(n)^{-r}$  avec  $r > 1$ . \*

L'hypothèse “ $v$  convexe” est une hypothèses technique mais peu restrictive en pratique. Il est intéressant de remarquer qu'une des conditions nécessaires au théorème précédent est  $v'b \geq 0$  ce qui implique que le point  $\Delta$  est un point source pour l'EDO  $u' = b(u)$ . Mais comme on l'a déjà vu dans l'exemple clôturant la section précédente, le point  $\Delta$  peut être de différentes natures pour l'EDS. Si le point  $\Delta$  est attractif (mais point source pour  $u' = b(u)$ ) alors le schéma  $(X_n)_{n \geq 0}$  converge presque sûrement vers  $\Delta$  mais reste d'un seul côté de  $\Delta$  à partir d'un temps fini presque sûrement.

**Remarque:** Il est immédiat d'étendre le théorème précédent à une situation où l'on a un nombre fini  $l$  de points frontières  $\Delta_i$  distincts. Si pour tout  $\Delta_i$ , il existe un voisinage  $U_{\Delta_i}$  et une fonction  $v_i \in \mathcal{C}^2(U_{\Delta_i}, \mathbf{R})$  convexe, minimale en  $\Delta$ , et vérifiant les conditions (2.3), alors

$$\mathbf{P} [\exists i \in \{0, \dots, l\}, \exists n_0 \geq 0, \forall n \geq n_0, X_n \in ]\Delta_i, \Delta_{i+1}[ ] = 1,$$

où  $\Delta_0 = -\infty$  et  $\Delta_{l+1} = +\infty$ . \*

## 2.2. Mesures empiriques

Les mesures empiriques pondérées, par une suite de poids  $(\nu_n^\eta)_{n \geq 1}$  positive terme principal d'une série divergente, du schéma d'Euler  $(X_n)_{n \geq 0}$  sont définies par

$$\forall \eta \geq 1, \quad \nu_n^\eta(dx) = \frac{1}{\sum_{k=1}^n \eta_k} \sum_{k=1}^n \eta_k \delta_{X_{k-1}}. \quad (2.8)$$

Sous les hypothèses de cette section –la diffusion vérifie une condition de stabilité forte de type  $(V, \text{Id}, \alpha)$  avec  $V$  sous-quadratique et les coefficients  $b$  et  $\sigma$  vérifient (2.2)–, on sait que  $(\nu_n^\eta)_{n \geq 1}$  est tendue et que tout candidat limite d'une sous-suite est probabilité invariante de  $(x_t)_{t \geq 0}$ . On obtient alors, comme conséquence du théorème VI.4, la proposition suivante qui précise la convergence des mesures empiriques pondérées en fonction du comportement de  $b$  et  $\sigma$  au voisinage du point frontière  $\Delta$ .

### Proposition VI.4:

Soit  $\Delta$  l'unique point de  $\mathbf{R}$  tel que  $b(\Delta) = \sigma(\Delta) = 0$ . On suppose qu'au voisinage de  $\Delta$  on a  $b(x) = \text{sgn}(x - \Delta)\rho_b(x)$  et  $\sigma(x) = \rho_\sigma(x)$  où  $\rho_b \geq 0$ ,

$$\rho_b(x) \sim c_b |x - \Delta|^\beta \quad \text{et} \quad \rho_\sigma(x) \sim c_\sigma |x - \Delta|^\zeta, \quad (2.9)$$

avec  $\beta, \zeta, c_b$  et  $c_\sigma$  réels et  $\zeta \geq 1$ . On suppose que le schéma d'Euler  $(X_n)_{n \geq 0}$  est construit à partir de la suite positive  $(\gamma_n)_{n \geq 1}$  vérifiant  $\forall C > 0, \sum_{n \geq 1} \exp(-C/\gamma_n) < +\infty$ . Alors

- si  $1 + \beta - 2\zeta > 0$ ,  $\Delta$  est attractif et  $\nu_n^\eta \Rightarrow \delta_\Delta$ ,
- si  $1 + \beta - 2\zeta = 0$  et  $c_\sigma > \sqrt{2c_b}$ , alors  $\Delta$  est attractif et  $\nu_n^\eta \Rightarrow \delta_\Delta$ ,
- si  $1 + \beta - 2\zeta = 0$ ,  $c_\sigma < \sqrt{2c_b}$  et  $\beta = 1$  (ce qui implique  $\zeta = 1$ ) alors  $\Delta$  est fortement répulsif et  $\nu_n^\eta \Rightarrow \nu_+$  ou  $\nu_n^\eta \Rightarrow \nu_-$ ,

- si  $1 + \beta - 2\zeta < 0$ ,  $c_\sigma \leq \sqrt{2c_b}$  et  $\beta \in ]0, 1]$  alors  $\Delta$  est fortement répulsif et  $\nu_n^\eta \Rightarrow \nu_+$  ou  $\nu_n^\eta \Rightarrow \nu_-$ ,

où  $\nu_+$  est la mesure invariante portée par  $]\Delta, +\infty[$  et  $\nu_-$  la mesure invariante portée par  $]-\infty, \Delta[$ .  $\blacklozenge$

Preuve. Par commodité on se place dans le cas où  $\Delta$  est le point 0.

- On considère un voisinage  $\mathcal{U}_\Delta$  de  $\mathbf{R}$  et la fonction  $v(x) = x^2$ . Alors pour tout  $x \in \mathcal{U}_\Delta$  on a

$$v'b = 2x \operatorname{sgn}(x) \rho_b(x) \geq 0.$$

De plus,  $v'\sigma \sim c_\sigma |x|^{\zeta+1}$  avec  $\zeta \geq 1$  donc il existe  $C > 0$  tel que  $|v'\sigma| \leq Cv$ . Le théorème VI.4 s'applique, et on sait alors que le schéma reste après un temps fini dans  $]-\infty, \Delta[$  ou dans  $]\Delta, +\infty[$ .

- Toujours en considérant la fonction  $v(x) = x^2$ , on a

$$\forall x \in \mathcal{U}_\Delta, \quad \mathcal{A}v(x) = 2|x|\rho_b(x) + \sigma^2(x) = 2c_b|x|^{1+\beta} + c_\sigma^2|x|^{2\zeta} + o(|x|^{(1+\beta)\vee(2\zeta)}), \quad (2.10)$$

et  $\frac{1}{2}\sigma^2(x) \frac{(v'(x))^2}{v(x)} = 2\sigma^2(x) \sim 2c_\sigma^2|x|^{2\zeta}$ .

– Si  $1 + \beta > 2\zeta$  alors il existe un voisinage de 0 dans lequel  $\mathcal{A}v < \frac{1}{2}\sigma \frac{(v')^2}{v}$ , donc par le corollaire VI.1 le point 0 est attractif. Les mesures empiriques pondérées du schéma d'Euler sont tendues sur  $[0, +\infty[$  et sur  $]-\infty, 0]$  donc par le théorème d'identification de la limite, toute sous-suite est mesure invariante. Or l'unique mesure invariante portée par  $[0, +\infty[$  ou  $]-\infty, 0]$  est  $\delta_0$ . Ainsi, on a convergence étroite de  $(\nu_n^\eta)_{n \geq 0}$  vers  $\delta_0$ .

– Si  $1 + \beta = 2\zeta$  alors  $\mathcal{A}v(x) = (2c_b + c_\sigma^2)|x|^{2\zeta} + o(|x|^{2\zeta})$ , donc si  $c_\sigma > \sqrt{2c_b}$  il existe un voisinage de 0 dans lequel  $\mathcal{A}v < 2\sigma^2(x)$  c'est à dire 0 est un point attractif. On conclut comme précédemment.

D'autre part, d'après le corollaire VI.1, le point 0 est fortement répulsif si  $\mathcal{A}v(x) \geq 2\sigma^2(x) + \varepsilon x^2$  pour un  $\varepsilon > 0$  et pour tout  $x$  dans un voisinage de 0. Si  $\beta = 1$  alors  $\zeta = 1$  et pour tout  $x$  dans un voisinage de 0, il existe  $\varepsilon > 0$  tel que  $\mathcal{A}v(x) = (2c_b + c_\sigma^2)x^2 + o(x^2) \leq 2\sigma^2(x) + \varepsilon x^2$  dès que  $c_b < \sqrt{2c_b}$ . Ainsi 0 est fortement répulsif et comme toute valeur d'adhérence de  $(\nu_n^\eta)_{n \geq 0}$  est portée par  $]-\infty, 0[$  ou bien  $]0, +\infty[$ , on en déduit que  $\nu_n^\eta \Rightarrow \nu_-$  ou bien  $\nu_n^\eta \Rightarrow \nu_+$  (on rappelle que les points infinis sont fortement répulsifs).

– Le cas  $1 + \beta < 2\zeta$ ,  $c_\sigma \leq \sqrt{2c_b}$  et  $\beta \in ]0, 1]$  se traite de la même façon.  $\#$

Pour illustrer ce résultat, reprenons l'exemple de la section précédente.

**Exemple.** Soit  $b$  et  $\sigma$  les fonctions définies par

$$b(x) = \begin{cases} -2(x - 3 \operatorname{sgn}(x)) & \text{si } |x| \geq 3, \\ -\frac{1}{18}x^3 + \frac{1}{2}x & \text{si } |x| \leq 3 \end{cases}, \quad \text{et } \sigma(x) = cx \text{ avec } c \in ]0, 2[.$$

On rappelle que la nature du point frontière 0 pour la diffusion  $(x_t)_{t \geq 0}$  solution de l'EDS de drift  $b$  et de coefficient de diffusion  $\sigma$  change en fonction du paramètre  $c$ . Si  $c \in ]1, 2[$  le point 0 est attractif, si  $c = 1$  il est répulsif, et si  $c \in ]0, 1[$  il est fortement répulsif.



Ce système non-linéaire de dimension 2 a le point  $(0, 0)$  comme source (ou “foyer répulsif”) et admet un cycle limite attractif. Ainsi, toute trajectoire tend vers une oscillation périodique. En bruitant suffisamment tout le plan sans briser l’origine, on s’aperçoit que le point  $(0, 0)$  est un point attractif pour le système stochastique ainsi créé, et que la mesure invariante de ce système est  $\delta_{(0,0)}$ .

On note  $b(x, y) = \begin{pmatrix} y \\ (1 - x^2)y - x \end{pmatrix}$  et  $\sigma(x, y) = \begin{pmatrix} cx & 0 \\ 0 & cy \end{pmatrix}$ , et on considère l’équation de Van der Pol perturbée

$$du_t = b(u_t)dt + \sigma(u_t)dB_t$$

où  $u_t = (x_t, y_t) \in \mathbf{R}^2$ . On discrétise la solution  $(x_t)_{t \geq 0}$  par un schéma d’Euler  $(X_n)_{n \geq 0}$  de pas décroissant  $(\gamma_n)_{n \geq 1}$  défini par  $\gamma_n = 0.5n^{-1/3}$ . Pour que ce schéma de discrétisation soit stable, on casse la non-linéarité lorsque  $x$  est grand en remplaçant la fonction  $b$  par la fonction  $\tilde{b}(x, y) = \begin{pmatrix} y \\ (1 - x^2 \wedge 4)y - x \end{pmatrix}$ . Le schéma considéré est donc défini par  $X_0 = (1, 1)$  et pour tout  $n \geq 0$

$$X_{n+1} = X_n + \gamma_n \tilde{b}(X_n) + \sqrt{\gamma_n} \sigma(X_n) U_{n+1},$$

où les  $U_n$  sont des gaussiennes centrées réduites bidimensionnelles.

On représente une approximation de la densité de  $\nu$  en utilisant un histogramme de  $\nu_n^\eta$  pour  $n = 10^6$  et  $\eta_n = 1$ . L’histogramme est réalisé en découpant d’un pas  $h = 0.2$  le plan d’extrémités  $(-3, -3)$  et  $(3, 3)$ . Les résultats sont donnés dans la figure VI.4.

On remarque que lorsque le bruit est faible ( $c = 0.5$ ), la mesure invariante ne semble pas charger le point  $(0, 0)$ . Plus le bruit augmente, et plus la mesure invariante charge le point  $(0, 0)$ . Le point source du système déterministe devient attractif. Pour des valeurs plus grandes de  $c$ , la mesure invariante semble être la Dirac en  $(0, 0)$ .

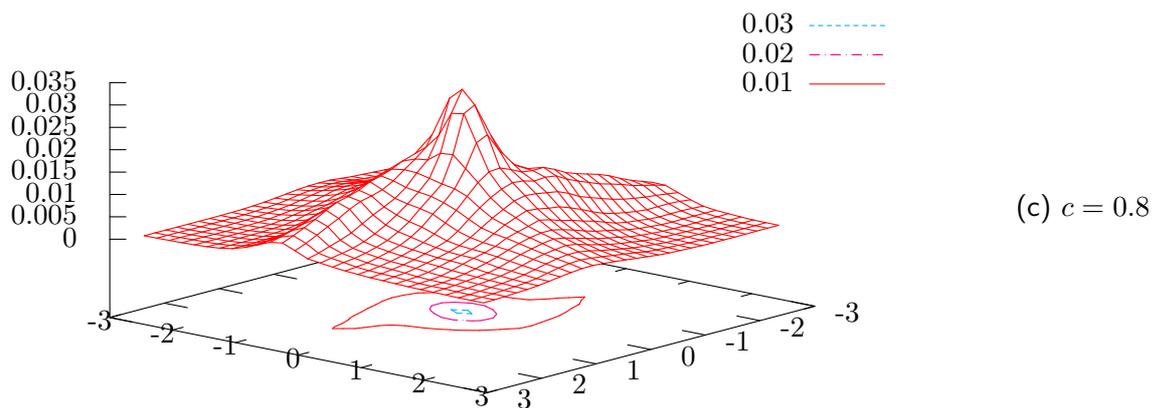
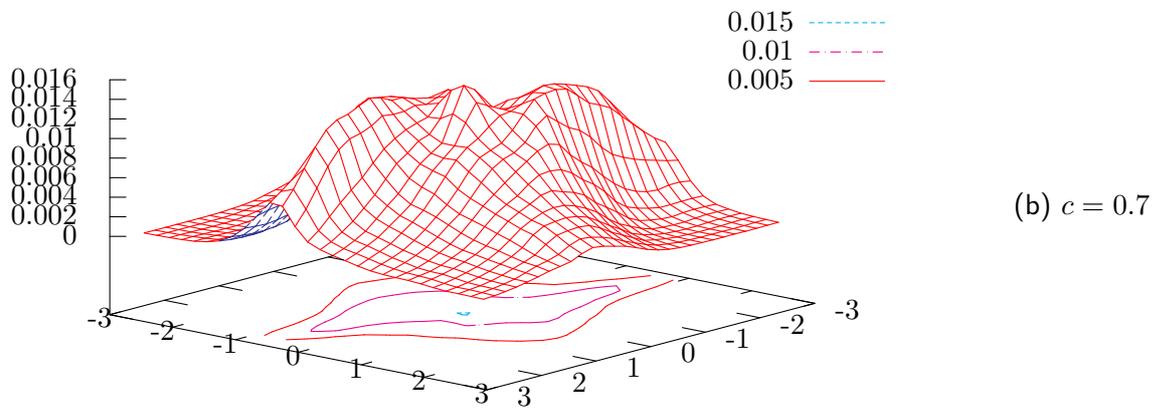
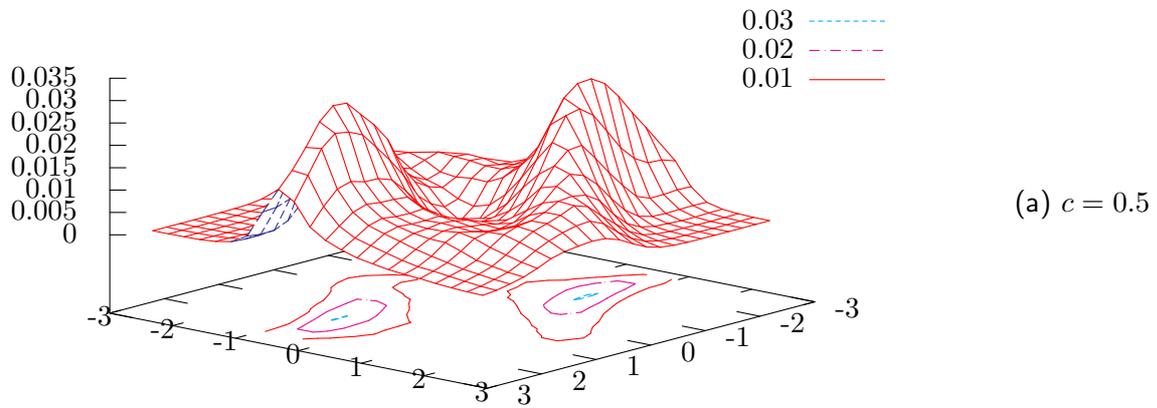


Figure VI.4 – Approximation de la densité de l'équation de Van der Pol perturbée pour différentes valeurs de  $c$ .

Troisième partie

Algorithme adaptatif



## 1. Introduction

We consider the following stochastic differential equation

$$dx_t = b(x_t)dt + \sigma(x_t)dB_t, \quad x(0) = x_0 \in \mathbf{R}^d, \quad (1.1)$$

where  $b : \mathbf{R}^d \rightarrow \mathbf{R}^d$  is a locally Lipschitz continuous vector field and  $\sigma$  is locally Lipschitz continuous on  $\mathbf{R}^d$ , with values in the set of  $d \times m$  matrices and  $B$  is an  $m$ -dimensional Brownian motion. Assume that  $(x_t)_{t \geq 0}$  has a Lyapounov function  $V$  *i.e.* a positive regular function decreasing along trajectories (precise conditions are given by Assumption 1 in Section 2), so that there exists at least one invariant measure.

Until recently, the approximation of the stationary mode of the diffusion has been studied under the assumption that  $V$  is *essentially quadratic* *i.e.*

$$|\nabla V|^2 = \mathcal{O}(V) \quad \text{and} \quad \sup_{x \in \mathbf{R}^d} \|D^2V\| < +\infty, \quad (1.2)$$

and  $|b|^2 = \mathcal{O}(V)$  (which implies sublinear growth for  $b$ ). When  $\sigma$  is bounded and the diffusion is uniformly strictly elliptic, the invariant measure  $\nu$  is unique and Talay proposed in [Tal90] a method for the computation of  $\nu$  based on the *constant step* Euler scheme. He proved the convergence of the invariant measure of the scheme to  $\nu$ . On the other hand, Lamberton and Pagès studied in [LP02] the ergodic properties of the weighted empirical measures  $(\nu_n^\eta)_{n \geq 1}$  of a *decreasing step* Euler scheme. They proved the almost sure tightness of  $(\nu_n^\eta)_{n \geq 1}$  and that any weak limit is a stationary distribution for the diffusion.

However, the conditions (1.2) and  $|b|^2 = \mathcal{O}(V)$  are too restrictive for studying systems used in random mechanics (see Soize [Soi94]). Indeed, the drift vector field  $b$  is generally locally Lipschitz and in many cases  $V$  is not essentially quadratic. This framework has been recently investigated by Talay in [Tal02] and by Mattingly et al. in [MSH02]. In these papers, *implicit* Euler schemes with constant steps are used for the approximation of the diffusion. In recent work, Lamba, Mattingly and Stuart have introduced on finite time interval  $[0; T]$  an adaptive explicit Euler scheme (see [LMA03] and [MS04]). The step is adapted according to the error between the Euler and Heun approximations of the ODE  $\dot{x} = b(x)$ . They prove strong mean-quadratic convergence of the scheme on over finite time intervals and ergodicity when the noise is non-degenerate. We propose a completely different *explicit scheme based on a stochastic step sequence* and we obtain the almost sure convergence of its weighted empirical

measures to the invariant measure of (1.1).

The key to prove the almost sure tightness of the weighted empirical measures of the decreasing Euler scheme  $(X_n)_{n \geq 0}$  introduced in [LP02] is that the scheme satisfies a *stability condition* i.e. there exist  $\tilde{\alpha} > 0$  and  $\tilde{\beta} > 0$  such that

$$\frac{\mathbf{E}[V(X_{n+1}) | \mathcal{F}_n] - V(X_n)}{\gamma_{n+1}} \leq -\tilde{\alpha}V(X_n) + \tilde{\beta}, \quad (1.3)$$

where  $(\gamma_n)_{n \geq 1}$  is the deterministic decreasing step sequence. Without assumptions (1.2) we can no longer prove the stability condition (1.3) for this scheme. Our scheme is built in order to satisfy (1.3). We proceed as follows. Firstly, we start from a deterministic  $X_0 = x_0 \in \mathbf{R}^d$  and set

$$X_{n+1} = X_n + \tilde{\gamma}_{n+1}b(X_n) + \sqrt{\tilde{\gamma}_{n+1}}\sigma(X_n)U_{n+1}, \quad n \geq 0, \quad (1.4)$$

where  $(U_n)_{n \geq 1}$  is a  $\mathbf{R}^m$ -white noise more precisely defined in Section 2 and  $\tilde{\gamma}_{n+1} = \gamma_{n+1} \wedge \chi_n$  with  $(\gamma_n)_{n \geq 1}$  a positive nonincreasing sequence and  $\chi_n$  a  $\sigma(U_1, \dots, U_n)$ -measurable random variable.

The basic main idea is to choose  $\chi_n$  small when the scheme starts to explode. In this case the discretization is finer and the stability condition of the diffusion prevents the explosion. Furthermore we prove that the scheme satisfies a similar condition with (1.3). A non-optimal –although natural– choice for  $\chi_n$  may be  $\chi_n = \frac{1}{|b(X_n)|^2 \vee 1}$ . For the two studied examples, optimal choices depend on the Lyapounov function (see (5.4) and (6.7)).

A crucial feature of our algorithm is the existence of an almost surely finite time  $n_1$  such that for every  $n \geq n_1$ ,  $\tilde{\gamma}_n = \gamma_n$  i.e. the event  $\{\chi_{n-1} < \gamma_n\}$  does not occur any more.

Numerically the algorithm is very simple to implement and the complexity is the same as that of a regular Euler scheme. Another interest is that the scheme is explicit, which is a big advantage on implicit schemes for high dimensional problems. Indeed a fixed point algorithm is not needed in our case. Moreover, we will see that wrong convergence problem due to fixed point algorithm may be avoided using our algorithm.

The paper is organized as follows. We introduce the framework and the algorithm in Section 2. In Section 3 are presented some preliminary results about the approximation scheme of  $(X_n)_{n \geq 0}$  defined in (1.4). In Section 4 we extend some results of [LP02] and give conditions for the almost sure tightness of the empirical measure and for its weak convergence to an invariant measure of (1.1). Section 5 is devoted to the study of monotone systems and Section 6 of stochastic Hamiltonian dissipative systems. The numerical experiments are in Section 7 including some comparison with recently introduced implicit scheme. We confirm the non-explosion and the convergence of the scheme.

## 2. Framework and algorithm

We will denote by  $\mathcal{A}$  the infinitesimal generator of (1.1). The following assumption will be needed throughout the paper.

**Assumption 1.** There is a  $\mathcal{C}^2$  function  $V$  on  $\mathbf{R}^d$  with values in  $[1, +\infty[$  such that  $\lim_{|x| \rightarrow +\infty} V(x) =$

$+\infty$  and satisfying

$$\exists \alpha > 0, \exists \beta > 0, \quad \langle \nabla V, b \rangle \leq -\alpha V + \beta, \quad (2.1)$$

$$\begin{aligned} \exists C_{V,\sigma} > 0, \exists a \in (0, 1], \quad \text{Tr}(\sigma^*(\nabla V)^{\otimes 2}\sigma) &\leq C_{V,\sigma} V^{2-a}, \\ \text{and} \quad \exists C > 0, \quad \sup_{x \in \mathbf{R}^d} \text{Tr}(\sigma^* D^2 V \sigma)(x) &\leq C \sup_{x \in \mathbf{R}^d} \text{Tr}(\sigma^* \sigma)(x). \end{aligned} \quad (2.2) \quad *$$

**Remark:** If  $V$  is essentially quadratic (i.e. satisfies (1.2)) then (2.2) is satisfied as soon as there exists  $C_\sigma > 0$  and  $a \in (0, 1]$  such that  $\text{Tr}(\sigma^* \sigma) \leq C_\sigma V^{1-a}$ . \*

Under this assumption, there exists a global solution to equation (1.1) and (at least) one invariant measure. An important point to note here is that all invariant measures have exponential moments. Indeed, an easy computation shows that for all  $\lambda < \frac{\alpha}{aC_{V,\sigma}}$  we have

$$\exists \tilde{\alpha} > 0, \exists \tilde{\beta} > 0, \quad \mathcal{A} \exp(\lambda V^a) \leq -\tilde{\alpha} \exp(\lambda V^a) + \tilde{\beta},$$

and this implies  $\nu(\exp(\lambda V^a))$  is finite for all invariant measures  $\nu$ .

For the approximation of the diffusion, we assume that  $(U_n)_{n \geq 1}$  is a sequence of i.i.d. random variables defined on a probability space  $(\Omega, \mathfrak{A}, \mathbf{P})$ , with values in  $\mathbf{R}^m$ , and such that  $U_1$  is a generalized Gaussian (see Stout [Sto74]) i.e.

$$\exists \kappa > 0, \forall \theta \in \mathbf{R}^d, \quad \mathbf{E}[\exp(\langle \theta, U_1 \rangle)] \leq \exp\left(\frac{\kappa |\theta|^2}{2}\right), \quad (2.3)$$

and that  $\text{var}(U_1) = \text{Id}_m$ . We will call  $(U_n)_{n \geq 1}$  a  $\mathbf{R}^m$ -valued generalized Gaussian white noise. The condition (2.3) implies that  $U_1$  is centered and satisfies

$$\exists \tau > 0, \quad \mathbf{E}[\exp(\tau |U_1|^2)] < +\infty. \quad (2.4)$$

Moreover, the condition  $\text{var}(U_1) = \text{Id}_m$  implies that  $\kappa \geq 1$ . In the sequel,  $\mathcal{F}_n$  denotes, for  $n \geq 1$ , the  $\sigma$ -field generated on  $\Omega$  by the random variables  $U_1, \dots, U_n$ , and  $\mathcal{F}_0$  the trivial  $\sigma$ -field.

**Remark:** The assumptions made on the white noise  $(U_n)_{n \geq 1}$  are not restrictive for numerical implementation. Indeed, centered Gaussian and centered bounded random variables satisfy (2.3). \*

The *stochastic step* sequence  $\tilde{\gamma} = (\tilde{\gamma}_n)_{n \geq 1}$  is defined by

$$\forall n \geq 1, \quad \tilde{\gamma}_n = \gamma_n \wedge \chi_{n-1}, \quad (2.5)$$

where  $(\gamma_n)_{n \geq 1}$  is a *deterministic nonincreasing* sequence of positive numbers satisfying

$$\lim_n \gamma_n = 0 \quad \text{and} \quad \sum_{n \geq 1} \gamma_n = +\infty,$$

and  $(\chi_n)_{n \geq 1}$  is an  $(\mathcal{F}_n)_{n \geq 0}$ -adapted sequence of positive *random* variables. It is important to note that the step sequence  $\tilde{\gamma}$  is  $(\mathcal{F}_n)_{n \geq 0}$ -predictable.

Now we introduce the *weighted empirical measures* like Lambertson and Pagès in [LP02]. Given a sequence  $\eta = (\eta_n)_{n \geq 1}$  of positive numbers satisfying  $\sum_{n \geq 1} \eta_n = +\infty$ , we denote by  $\nu_n^\eta$  the random probability measure on  $\mathbf{R}^d$  defined by

$$\nu_n^\eta = \frac{1}{H_n} \sum_{k=1}^n \eta_k \delta_{X_{k-1}}, \quad \text{with} \quad H_n = \sum_{k=1}^n \eta_k.$$

Throughout the paper,  $|\cdot|$  denotes the Euclidean norm and  $\|\cdot\|$  denotes the natural matrix norm induced by  $|\cdot|$  i.e. for every square matrix  $A$ ,  $\|A\| = \sup_{|x|=1} |Ax|$ . The letter  $C$  is used to denote a positive constant, which may vary from line to line.

### 3. Preliminary results

In this section, we prove results which are the keys to study the Euler scheme with predictable random step defined in the introduction. Proposition VII.1 contains two results : the first one (3.2) provides a substitute for the  $L^p$ -boundedness of  $(V(X_n))_{n \geq 0}$  used in [LP02]. The second one (3.3) is a new consequence of the stability condition (3.1) and is used to prove the fundamental proposition VII.2 which ensures the existence of an almost surely finite time  $n_1$  such that for every  $n \geq n_1$ ,  $\tilde{\gamma}_n = \gamma_n$ .

**Proposition VII.1:**

Let  $W$  be a nonnegative function and  $(\tilde{\gamma}_n)_{n \geq 1}$  be a  $(\mathcal{F}_n)_{n \geq 0}$ -predictable sequence of positive and finite random variables satisfying : there exist  $\alpha > 0$ ,  $\beta > 0$ ,  $n_0 \in \mathbf{N}$ , such that

$$\forall n \geq n_0, \quad \frac{\mathbf{E}[W(X_{n+1}) | \mathcal{F}_n] - W(X_n)}{\tilde{\gamma}_{n+1}} \leq -\alpha W(X_n) + \beta. \quad (3.1)$$

Suppose  $(\theta_n)_{n \geq 1}$  is a positive nonincreasing sequence such that  $\mathbf{E}[\sum_{n \geq 0} \theta_n \tilde{\gamma}_n]$  is finite, then

$$\mathbf{E}\left[\sum_{n \geq n_0+1} \theta_n \tilde{\gamma}_n W(X_{n-1}) | \mathcal{F}_{n_0}\right] < +\infty. \quad (3.2)$$

If, in addition,  $\lim_n \theta_n = 0$  then

$$\lim_n \theta_n W(X_n) = 0 \quad a.s. \quad (3.3)$$

The above proposition is related to Robbins-Siegmund's theorem (see Theorem 1.3.12 in [Duf97] and the references therein).

**Proof.** Let  $R_n = \sum_{k=0}^n \theta_k \tilde{\gamma}_k$  and  $R_\infty = \lim_n R_n \in \mathbf{L}^1(\mathbf{P})$ .

We consider the sequence  $(Z_n)_{n \geq n_0}$  defined by

$$\forall n \geq n_0, \quad Z_{n+1} = Z_n + \theta_{n+1}(W(X_{n+1}) - W(X_n)), \quad Z_{n_0} = \theta_{n_0} W(X_{n_0}).$$

We first prove that for every  $n \geq n_0$ ,  $Z_n \geq 0$ . Indeed, an Abel transform yields for every

$n \geq n_0$ ,

$$\begin{aligned} Z_n &= \sum_{k=n_0}^{n-1} \theta_{k+1} (W(X_{k+1}) - W(X_k)) + \theta_{n_0} W(X_{n_0}), \\ &= \sum_{k=n_0}^{n-1} (\theta_k - \theta_{k+1}) W(X_k) + \theta_n W(X_n). \end{aligned}$$

The sequence  $(\theta_n)_{n \geq 0}$  is nonincreasing and the function  $W$  is nonnegative, then  $(Z_n)_{n \geq n_0}$  is positive.

Let  $(S_n)_{n \geq n_0}$  denote the process defined for every  $n \geq n_0$  by

$$S_n = Z_n + \alpha \sum_{k=n_0+1}^n \theta_k \tilde{\gamma}_k W(X_{k-1}) + \beta (\mathbf{E}[R_\infty | \mathcal{F}_n] - R_n).$$

Since  $\mathbf{E}[R_\infty | \mathcal{F}_n] - R_n \geq 0$ , the sequence  $(S_n)_{n \geq n_0}$  is nonnegative. Moreover, as  $W$  satisfies (3.1) we have

$$\forall n \geq n_0, \quad \mathbf{E}[Z_{n+1} | \mathcal{F}_n] \leq Z_n - \alpha \theta_{n+1} \tilde{\gamma}_{n+1} W(X_n) + \beta \theta_{n+1} \tilde{\gamma}_{n+1}.$$

Then it follows from this and from the  $\mathcal{F}_n$ -measurability of  $\tilde{\gamma}_{n+1}$  that

$$\forall n \geq n_0, \quad \mathbf{E}[S_{n+1} | \mathcal{F}_n] \leq S_n.$$

Thus  $(S_n)_{n \geq 0}$  converges *a.s.* to a nonnegative finite random variable  $S_\infty$ , and we have

$$\mathbf{E} \left[ \sum_{n \geq n_0+1} \theta_n \tilde{\gamma}_n W(X_{n-1}) | \mathcal{F}_{n_0} \right] < +\infty.$$

From the almost sure convergence of  $(S_n)_{n \geq n_0}$ , we also deduce the almost sure convergence of the series

$$\sum_{n \geq 1} \theta_n (W(X_n) - W(X_{n-1})).$$

Since  $(\theta_n)_{n \geq 0}$  is nonincreasing and converges to 0, Kronecker's lemma implies the almost sure convergence of  $(\theta_n W(X_n))_{n \geq 0}$  to 0. #

**Remark:** A substitute for the  $L^p$ -boundedness of  $(V(X_n))_{n \geq 0}$  has been already found in Lemma 4 of [LP03] but does not apply in our case. Indeed, we have no information on the expectation of the random variable  $\tilde{\gamma}_n$ . \*

The following Proposition is a fundamental consequence of (3.3) and says that for a “good choice” of the process  $(\chi_n)_{n \geq 0}$ , we can choose a sequence  $(\gamma_n)_{n \geq 1}$  such that the random step  $\tilde{\gamma}$  becomes deterministic after an almost surely finite time. The existence of this finite time  $n_1$  is a significant property of our scheme.

**Proposition VII.2:**

Let  $f : [1, +\infty) \rightarrow (0, +\infty)$  be a decreasing one-to-one continuous function with  $\lim_{x \rightarrow +\infty} f(x) =$

0 and  $\lim_{x \rightarrow 1} f(x) = +\infty$ , and  $W$  a positive function with values in  $[1, +\infty)$  satisfying (3.1). If

$$\chi_n \geq (f \circ W)(X_n),$$

and if  $(\gamma_n)_{n \geq 1}$  is subject to the condition

$$\sum_{n \geq 1} \frac{\gamma_n}{f^{-1}(\gamma_n)} < +\infty,$$

where  $f^{-1}$  is the inverse of  $f$ , then there is an almost surely finite random variable  $n_1$  such that  $\tilde{\gamma}_n = \gamma_n$  for every  $n \geq n_1$ .  $\blacklozenge$

**Proof.** Let  $(\theta_n)_{n \geq 1}$  the sequence defined by

$$\forall n \geq 1, \quad \theta_n = \frac{1}{f^{-1}(\gamma_n)}. \quad (3.4)$$

Since  $(\gamma_n)_{n \geq 1}$  is nonincreasing and  $1/f^{-1}$  is increasing on  $\mathbf{R}_+$ , then  $(\theta_n)_{n \geq 1}$  is nondecreasing and we have  $\sum_n \theta_n \tilde{\gamma}_n \leq \sum_n \frac{\gamma_n}{f^{-1}(\gamma_n)} < +\infty$ . Moreover  $\lim_n \theta_n = 1/f^{-1}(0) = 0$  so that, by Proposition VII.1, we have

$$\lim_n \theta_{n-1} W(X_{n-1}) = 0 \quad a.s.$$

Hence there is an *a.s.* finite random variable  $n_1$  such that

$$\forall n \geq n_1, \quad W(X_{n-1}) < f^{-1}(\gamma_{n-1}) \leq f^{-1}(\gamma_n) \quad a.s.$$

By the lower bound on  $\chi_n$  and the monotony of  $f$ , we have

$$\forall n \geq n_1, \quad \chi_{n-1} > \gamma_n \quad a.s.$$

which completes the proof.  $\#$

## 4. Convergence of empirical measures

In this section, we give conditions for the almost sure tightness of  $(\nu_n^n)_{n \geq 1}$  and the weak convergence to an invariant distribution of (1.1). To this end, we assume the existence of an almost surely finite random variable  $n_1$  such that for every  $n \geq n_1$ ,  $\tilde{\gamma}_n = \gamma_n$ . In practice, this means that  $\tilde{\gamma}_n$  is defined by (2.5) with  $\gamma_n$  and  $\chi_n$  satisfying conditions of Proposition VII.2. Under this assumption the proofs are very close to those in [LP02] and [LP03].

From now on we make the assumption :

**Assumption 2.** The deterministic sequences  $(\gamma_n)_{n \geq 1}$  and  $(\eta_n)_{n \geq 1}$  satisfy

$$\begin{aligned} \lim_n \frac{1}{H_n} \sum_{k=1}^n \left| \Delta \frac{\eta_k}{\gamma_k} \right| &= 0, \\ \left( \frac{1}{\gamma_n H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ \right) &\text{ is nonincreasing and } \sum_{n \geq 1} \frac{1}{H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ < +\infty, \end{aligned} \quad (4.1)$$

and there exists  $s \in (1, 2]$  such that

$$\left( \frac{1}{\gamma_n} \left( \frac{\eta_n}{H_n \sqrt{\gamma_n}} \right)^s \right) \text{ is nonincreasing and } \sum_{n \geq 1} \left( \frac{\eta_n}{H_n \sqrt{\gamma_n}} \right)^s < +\infty. \quad (4.2)$$

**Remark:** In practice the above conditions on  $(\gamma_n)_{n \geq 1}$  and  $(\eta_n)_{n \geq 1}$  are not restrictive. Setting  $\gamma_n = n^{-p}$ ,  $0 < p \leq 1$  and  $\eta_n = n^{-q}$ ,  $q \leq 1$ , Assumption 2 is satisfied if and only if

$$(p, q) \in \left( 0, \frac{2(s-1)}{s} \right) \times (-\infty, 1] \cup \left\{ \left( \frac{2(s-1)}{s}, 1 \right) \right\}.$$

#### 4.1. A.s. tightness of empirical measures

We begin with proving the almost surely tightness of  $(\nu_n^\eta)_{n \geq 1}$  when we have a control on the scheme  $(X_n)_{n \geq 0}$ .

##### Theorem VII.1:

We assume that  $(\gamma_n)_{n \geq 1}$  and  $(\eta_n)_{n \geq 1}$  satisfy Assumption 2 and that there exists an almost surely finite random variable  $n_1$  such that  $\tilde{\gamma}_n = \gamma_n$  for every  $n \geq n_1$ . Suppose that  $W$  is a positive function satisfying (3.1) and

$$\sum_{n \geq 1} \left( \frac{\eta_n}{H_n \gamma_n} \right)^s \mathbf{E} \left[ |W(X_n) - \mathbf{E}[W(X_n) | \mathcal{F}_{n-1}]|^s | \mathcal{F}_{n-1} \right] < +\infty \quad a.s.,$$

for some  $s \in (1, 2]$ . Then,

$$\sup_{n \geq 1} \nu_n^\eta(W) < +\infty \quad a.s.$$

**Proof.** Since  $W$  satisfies (3.1), there exist  $\alpha > 0$ ,  $\beta > 0$  and  $n_0 \geq 0$  such that

$$\forall n \geq n_0, \quad W(X_n) \leq \frac{W(X_n) - \mathbf{E}[W(X_{n+1}) | \mathcal{F}_n]}{\tilde{\gamma}_{n+1} \alpha} + \frac{\beta}{\alpha}.$$

Hence, for every  $n \geq n_0 \vee n_1 + 1$ ,

$$\frac{1}{H_n} \sum_{k=n_0 \vee n_1+1}^n \eta_k W(X_{k-1}) \leq \frac{1}{H_n} \sum_{k=n_0 \vee n_1+1}^n \frac{\eta_k}{\gamma_k \alpha} (W(X_{k-1}) - \mathbf{E}[W(X_k) | \mathcal{F}_{k-1}]) + \frac{\beta}{\alpha}.$$

It suffices to prove that

$$\sup_{n \geq 1} \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (W(X_{k-1}) - \mathbf{E}[W(X_k) | \mathcal{F}_{k-1}]) < +\infty \quad a.s.$$

An Abel transform, setting  $\eta_0 = 0$ , yields

$$\begin{aligned} \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (W(X_{k-1}) - W(X_k)) &= \frac{1}{H_n} \sum_{k=1}^n \left( \Delta \frac{\eta_k}{\gamma_k} \right) W(X_{k-1}) - \frac{\eta_n}{\gamma_n} W(X_n) \\ &\leq \frac{1}{H_n} \sum_{k=1}^n \left( \Delta \frac{\eta_k}{\gamma_k} \right)_+ W(X_{k-1}), \end{aligned}$$

and it follows from condition (4.1) and Proposition VII.1 applied with  $\theta_n = \frac{1}{\gamma_n H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+$  that

$$\sum_{n \geq 1} \frac{1}{H_n} \left( \Delta \frac{\eta_n}{\gamma_n} \right)_+ W(X_{n-1}) < +\infty \quad a.s.$$

Applying Kronecker's lemma, we get

$$\limsup_n \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (W(X_{k-1}) - W(X_k)) \leq 0 \quad a.s.$$

It remains to prove that

$$\sup_{n \geq 1} \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (W(X_k) - \mathbf{E}[W(X_k) | \mathcal{F}_{k-1}]) < +\infty \quad a.s.$$

We introduce the martingale  $(M_n)_{n \geq 1}$  defined by

$$\forall n \geq 1, \quad M_n := \sum_{k=1}^n \frac{\eta_k}{H_k \gamma_k} \left( W(X_k) - \mathbf{E}[W(X_k) | \mathcal{F}_{k-1}] \right), \quad M_0 := 0.$$

By the Chow theorem (see [HH80]), the *a.s.* convergence of  $(M_n)_{n \geq 1}$  will follow from the *a.s.* convergence of

$$\sum_{n \geq 1} \left( \frac{\eta_n}{H_n \gamma_n} \right)^s \mathbf{E} \left[ |W(X_n) - \mathbf{E}[W(X_n) | \mathcal{F}_{n-1}]|^s | \mathcal{F}_{n-1} \right] < +\infty \quad a.s.$$

and the Kronecker lemma completes the proof. #

#### 4.2. Identification of the limit

We now prove that any weak limit of  $(\nu_n^\eta)_{n \geq 1}$  is an invariant distribution for the diffusion (1.1). We use the same method as Lamberton and Pagès. By the Echeverria-Weiss theorem (see [EK86]) it suffices to prove that  $\lim_n \nu_n^\eta(\mathcal{A}f) = 0$  for any twice continuously differentiable function  $f$  with compact support. The existence of the Lyapounov function  $V$  implies the regularity of the process  $(x_t)_{t \geq 0}$  so that the Echeverria-Weiss applies –although  $b$  has not sublinear growth–.

##### **Proposition VII.3:**

Suppose that there exists  $K > 0$  such that

$$\forall n \geq 1, \quad |X_n - X_{n-1}| \leq K \sqrt{\gamma_n} \sqrt{V}(X_{n-1}) (1 + |U_n|), \quad (4.3)$$

and a function  $W$  satisfying (3.1),  $V = o(W)$  and  $\sup_{n \geq 1} \nu_n^\eta(W) < +\infty$ . Then for every

twice continuously differentiable function  $f$  with compact support,

$$\lim \nu_n^\eta(\mathcal{A}f) = 0 \quad a.s.$$

The following Lemma is useful to prove this Proposition.

**Lemma VII.1:**

Under the assumptions of Proposition VII.3, for every bounded Lipschitz continuous function  $f : \mathbf{R}^d \rightarrow \mathbf{R}$  we have

$$\lim_n \frac{1}{H_n} \sum_{k=1}^n \eta_k \frac{\mathbf{E}[f(X_k) | \mathcal{F}_{k-1}] - f(X_{k-1})}{\gamma_k} = 0 \quad a.s.$$

Proof. First, we prove that

$$\lim_n \frac{1}{H_n} \sum_{k=1}^n \eta_k \frac{\mathbf{E}[f(X_k) | \mathcal{F}_{k-1}] - f(X_k)}{\gamma_k} = 0 \quad a.s. \quad (4.4)$$

To this end, we introduce the martingale  $(M_n)_{n \geq 0}$  defined by

$$\forall n \geq 1, \quad M_n = \sum_{k=1}^n \frac{\eta_k}{H_k \gamma_k} (\mathbf{E}[f(X_k) | \mathcal{F}_{k-1}] - f(X_k)) \quad \text{and} \quad M_0 = 0.$$

We have

$$(\mathbf{E}[f(X_k) | \mathcal{F}_{k-1}] - f(X_k))^2 = (\mathbf{E}[f(X_k) - f(X_{k-1}) | \mathcal{F}_{k-1}] + f(X_{k-1}) - f(X_k))^2,$$

and using  $(a + b)^2 \leq 2(a^2 + b^2)$ , Jensen's inequality and  $f$  Lipschitz we get

$$\begin{aligned} \langle M \rangle_\infty &\leq 4 \sum_{n \geq 1} \left( \frac{\eta_n}{H_n \gamma_n} \right)^2 \mathbf{E} \left[ (f(X_n) - f(X_{n-1}))^2 | \mathcal{F}_{n-1} \right], \\ &\leq C \sum_{n \geq 1} \left( \frac{\eta_n}{H_n \sqrt{\gamma_n}} \right)^2 V(X_{n-1}). \end{aligned} \quad (4.5)$$

Since  $s \leq 2$  and  $(\gamma_n)_n$  is nonincreasing, the sequence  $(\theta_n)_n = \left( \frac{1}{\gamma_n} \left( \frac{\eta_n}{H_n \sqrt{\gamma_n}} \right)^2 \right)_{n \geq 1}$  is nonincreasing and by (4.2) we have  $\sum_{n \geq 1} \theta_n \gamma_n < +\infty$ . Moreover  $W$  satisfies (3.1) and  $\tilde{\gamma}_n = \gamma_n$  for every  $n \geq n_1$ , then the Proposition VII.1 applied with  $\gamma_n$  and  $\theta_n$  yields

$$\sum_{n \geq n_0 \vee n_1 + 1} \theta_n \gamma_n W(X_{n-1}) < +\infty \quad a.s. \quad (4.6)$$

By (4.5),  $V = o(W)$  and (4.6) we obtain the almost sure convergence of the increasing process of  $(M_n)_{n \geq 0}$ . Thus the martingale converges almost surely. The Kronecker lemma gives (4.4).

Finally we prove that

$$\lim_n \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} (f(X_k) - f(X_{k-1})) = 0 \quad a.s.,$$

using Abel's transform, the boundedness of  $f$  and  $\lim_n \frac{1}{H_n} \sum_{k=1}^n \left| \Delta \frac{\eta_k}{\gamma_k} \right| = 0$ . #

**Proof** (Proposition VII.3). By Taylor's formula applied to  $f$  between  $X_{n-1}$  and  $X_n$  we have

$$\begin{aligned} \mathbf{E}[f(X_n) | \mathcal{F}_{n-1}] - f(X_{n-1}) &= \tilde{\gamma}_n \mathcal{A} f(X_{n-1}) + \frac{1}{2} \tilde{\gamma}_n^2 D^2 f(X_{k-1}) \cdot b(X_{n-1})^{\otimes 2} \\ &\quad + \mathbf{E}[R_2(X_{n-1}, X_n) | \mathcal{F}_{n-1}], \end{aligned} \quad (4.7)$$

with  $R_2(x, y) = f(y) - f(x) - \langle \nabla f(x), y - x \rangle - \frac{1}{2} D^2 f(x) \cdot (y - x)^{\otimes 2}$ . On the one hand, by the above lemma

$$\lim_n \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\tilde{\gamma}_k} \left( \mathbf{E}[f(X_k) | \mathcal{F}_{k-1}] - f(X_{k-1}) \right) = 0 \quad a.s. \quad (4.8)$$

and on the other hand we have, using that  $\tilde{\gamma}_n \leq \gamma_n$  and that  $D^2 f$  has compact support,

$$\left| \frac{1}{H_n} \sum_{k=1}^n \eta_k \tilde{\gamma}_k D^2 f(X_{k-1}) \cdot b^{\otimes 2}(X_{k-1}) \right| \leq \|D^2 f \cdot b^{\otimes 2}\|_{\infty} \frac{1}{H_n} \sum_{k=1}^n \eta_k \gamma_k$$

and since  $(\gamma_n)_{n \geq 1}$  is decreasing to 0, we obtain

$$\lim_n \frac{1}{H_n} \sum_{k=1}^n \eta_k \tilde{\gamma}_k D^2 f(X_{k-1}) \cdot b^{\otimes 2}(X_{k-1}) = 0 \quad a.s. \quad (4.9)$$

From (4.7), (4.8) and (4.9), it follows that

$$\lim_n \left( \nu_n^{\eta}(\mathcal{A} f) + \frac{1}{H_n} \sum_{k=1}^n \frac{\eta_k}{\gamma_k} \mathbf{E}[R_2(X_{k-1}, X_k) | \mathcal{F}_{k-1}] \right) = 0 \quad a.s. \quad (4.10)$$

We introduce the continuous bounded function

$$r_2(x, \delta) = \frac{1}{2} \sup_{z \in \mathbf{R}^d, |z-x| < \delta} |D^2 f(z) - D^2 f(x)|,$$

which is nondecreasing in  $\delta$  and satisfies

$$\forall (x, y) \in \mathbf{R}^{2d}, \quad |R_2(x, y)| \leq r_2(x, |y-x|) |y-x|^2.$$

Thus, from (4.10) and (4.3) it suffices to prove that

$$\lim_n \frac{1}{H_n} \sum_{k=1}^n \eta_k V(X_{k-1}) \mathbf{E} \left[ r_2(X_{k-1}, \delta_k(U_k)) (1 + |U_k|^2) | \mathcal{F}_{k-1} \right] = 0 \quad a.s. \quad (4.11)$$

where  $\delta_k(u) = |\tilde{\gamma}_k b(X_{k-1}) + \sqrt{\tilde{\gamma}_k} \sigma(X_{k-1}) u|$ .

Let  $u \in \mathbf{R}^m$  and  $A > 0$ . It is clear that the sequence  $(\delta_k(u) 1_{\{|X_{k-1}| \leq A\}})_{k \geq 1}$  converges almost surely to 0 and thus

$$\lim_k r_2(X_{k-1}, \delta_k(u)) 1_{\{|X_{k-1}| \leq A\}} = 0 \quad a.s.$$

Since

$$\frac{1}{H_n} \sum_{k=1}^n \eta_k V(X_{k-1}) r_2(X_{k-1}, \delta_k(u)) (1 + |u|^2) 1_{\{|X_{k-1}| \leq A\}} \leq K \|r_2\|_\infty (1 + |u|^2),$$

we have by the dominated convergence theorem

$$\lim_n \frac{1}{H_n} \sum_{k=1}^n \eta_k V(X_{k-1}) \int_{\mathbf{R}^m} r_2(X_{k-1}, \delta_k(u)) (1 + |u|^2) \mu(du) 1_{\{|X_{k-1}| \leq A\}} = 0 \quad a.s. \quad (4.12)$$

where  $\mu$  is the law of  $U_1$ . On the other hand, we have

$$\begin{aligned} \overline{\lim}_n \frac{1}{H_n} \sum_{k=1}^n \eta_k V(X_{k-1}) \mathbf{E} [r_2(X_{k-1}, \delta_k(U_k)) (1 + |U_k|^2) | \mathcal{F}_{k-1}] 1_{\{|X_{k-1}| > A\}} \\ \leq (1 + m) \|r_2\|_\infty \sup_{|x| > A} |V(x)/W(x)| \sup_n \nu_n^\eta(W), \end{aligned}$$

and letting  $A \rightarrow +\infty$  and combining with (4.12) we obtain (4.11). #

## 5. Monotone and dissipative problems

We now apply our results to monotone problems. In this section we assume that  $V$  is essentially quadratic but the drift  $b$  need not be globally Lipschitz.

**Assumption 3.** The function  $V$  satisfies

$$\exists C_V > 0, \quad |\nabla V|^2 \leq C_V V, \quad \text{and} \quad \|D^2 V\|_\infty := \sup_{x \in \mathbf{R}^d} \|D^2 V\| < +\infty. \quad (5.1)$$

Under this condition and Assumption 1, Mattingly, Stuart and Higham proved the geometric ergodicity of (1.1) when  $\sigma$  is constant (see [MSH02] for details). We assume that  $\sigma$  satisfies

$$\exists C_\sigma > 0, \exists a \in (0, 1], \quad \text{Tr}(\sigma \sigma^*) \leq C_\sigma V^{1-a}, \quad (5.2)$$

so that condition (2.2) is checked.

We consider the Markov chain  $(X_n)_{n \geq 0}$  built by the recursive procedure (1.4) with the random step sequence  $(\tilde{\gamma}_n)_{n \geq 1}$  defined by

$$\forall n \geq 1, \quad \tilde{\gamma}_n = g_n(X_{n-1}) \quad (5.3)$$

with  $g_n : \mathbf{R}^d \rightarrow \mathbf{R}_+^*$  defined by  $g_n(x) = \gamma_n \wedge \chi(x)$  where  $\chi$  is a function on  $\mathbf{R}^d$  with values in  $\mathbf{R}_+^*$ .

The main result of this section is the following Theorem.

**Théorème VII.1:**

Let  $(\gamma_n)_{n \geq 1}$  and  $(\eta_n)_{n \geq 1}$  satisfy Assumption 2. Suppose that there exist  $l > 1$  and  $C_b \geq 1$

such that  $|b(x)|^2 \leq C_b V^l(x)$  and that  $\chi$  satisfies

$$\forall x \in \mathbf{R}^d, \quad \zeta V^{-p}(x) \leq \chi(x) \leq \left( \frac{2\delta}{\|D^2V\|_\infty} \right) \frac{V(x)}{|b(x)|^2 \vee 1}, \quad (5.4)$$

with  $\zeta \leq \frac{2\delta}{C_b \|D^2V\|_\infty}$ ,  $p \geq l - 1$  and  $\delta \in (0, \alpha)$ . If  $(\gamma_n)_{n \geq 1}$  satisfies

$$\sum_{n \geq 1} \gamma_n \exp(-\lambda \zeta^{\frac{a}{p}} \gamma_n^{-\frac{a}{p}}) < +\infty, \quad (5.5)$$

and  $\lambda_0 = \frac{2\tau}{a C_\sigma \|D^2V\|_\infty} \wedge \frac{2(\alpha - \delta)}{\kappa a C_V C_\sigma}$  then

- there exists an almost surely finite random variable  $n_1$  such that  $\forall n \geq n_1$ ,  $\tilde{\gamma}_n = \gamma_n$ ,
- for every  $\lambda < \frac{\lambda_0}{s}$  (whith  $s$  as in (4.2)),

$$\sup_n \nu_n^\eta(\exp(\lambda V^a)) < +\infty \quad a.s.$$

and any weak limit of  $(\nu_n^\eta)_{n \geq 1}$  is an invariant distribution for (1.1). ★

**Remark:** The condition (5.5) is not restrictive. For example, it is satisfied if  $\gamma_n \leq \lambda^{\frac{p}{a}} \zeta \ln^{-\frac{p}{a}}(n)$  and  $p > a$ , or  $\gamma_n \leq \lambda^{\frac{p}{a}} \zeta \ln^{-(1+\varepsilon)}(n)$  for some  $\varepsilon > 0$  and  $p \leq a$ . \*

To prove this Theorem, it suffices essentially to check that the function  $W = \exp(\lambda V^a)$  satisfies condition (3.1) and the assumptions of Theorem VII.1. This is the aim of Lemmas VII.2 and VII.3 respectively, which will be proved later.

**Lemma VII.2:**

Assume that

$$\forall x \in \mathbf{R}^d, \quad \chi(x) \leq \left( \frac{2\delta}{\|D^2V\|_\infty} \right) \frac{V(x)}{|b(x)|^2 \vee 1},$$

where  $\delta \in (0, \alpha)$ . Let  $\lambda_0$  be as in Theorem VII.1. Then for every  $\lambda < \lambda_0$  there exists  $\tilde{\alpha} > 0$ ,  $\tilde{\beta} > 0$ , and  $n_0 \geq 0$  such that for every  $n \geq n_0$ ,

$$\frac{\mathbf{E}[\exp(\lambda V^a(X_{n+1})) | \mathcal{F}_n] - \exp(\lambda V^a(X_n))}{\tilde{\gamma}_{n+1}} \leq -\tilde{\alpha} \exp(\lambda V^a(X_n)) + \tilde{\beta}.$$

**Lemma VII.3:**

Assume that there exists an almost sure finite random variable  $n_1$  such that  $\tilde{\gamma}_n = \gamma_n$  for every  $n \geq n_1$ . Let  $\lambda_0$  be as in Theorem VII.1. Then for every  $\lambda < \frac{\lambda_0}{s}$  the series

$$\sum_{n \geq 1} \left( \frac{\eta_n}{H_n \gamma_n} \right)^s \mathbf{E} \left[ \left| \exp(\lambda V^a(X_n)) - \mathbf{E}[\exp(\lambda V^a(X_n)) | \mathcal{F}_{n-1}] \right|^s | \mathcal{F}_{n-1} \right]$$

is almost surely finite. ◆

**Proof** (Theorem VII.1). By Lemma VII.2 the function  $W = \exp(\lambda V^a)$  satisfies condition (3.1). We consider the function  $f$  defined on  $[1, +\infty)$  by

$$\forall x \geq 1, \quad f(x) = \zeta \lambda^{\frac{p}{a}} \ln^{-\frac{p}{a}}(x).$$

This is a decreasing one-to-one continuous function with  $\lim_{x \rightarrow 1} f(x) = +\infty$  and  $\lim_{x \rightarrow +\infty} f(x) = 0$ . We check that  $f \circ W = \zeta V^{-p}$  and by (5.4) we have  $\chi(X_n) \geq (f \circ W)(X_n)$ . The inverse of  $f$  is the function  $f^{-1}$  defined by

$$\forall x \in ]0, +\infty[, \quad f^{-1}(x) = \exp\left(\lambda \zeta^{\frac{a}{p}} x^{-\frac{a}{p}}\right), \quad (5.6)$$

so that the condition (5.5) on  $\gamma_n$  is  $\sum_{n \geq 0} \frac{\gamma_n}{f^{-1}(\gamma_n)} < +\infty$ . The Proposition VII.2 applied with  $f$  and  $W$  gives the existence of an almost surely finite random variable  $n_1$  such that  $\tilde{\gamma}_n = \gamma_n$  for every  $n \geq n_1$ .

By Lemma VII.2 the conditions of Theorem VII.1 are fulfilled and we have for every  $\lambda < \lambda_0/s$

$$\sup_{n \geq 1} \nu_n^\eta(\exp(\lambda V^a)) < +\infty \quad a.s.$$

It remains to prove that any weak limit of  $(\nu_n^\eta)_{n \geq 1}$  is an invariant distribution for the diffusion. By subsection 4.2 and Proposition VII.3, it suffices to check (4.3). On the one hand, the definition of the algorithm yields

$$|X_n - X_{n-1}| \leq \tilde{\gamma}_n |b(X_{n-1})| + \sqrt{\tilde{\gamma}_n} |\sigma(X_{n-1}) U_n|.$$

On the other hand, we have  $\tilde{\gamma}_n = g_n(X_{n-1})$  and

$$g_n(x) |b(x)| \leq \sqrt{g_n(x)} \sqrt{\chi(x)} |b(x)| \leq \sqrt{g_n(x)} \sqrt{\frac{2\delta}{\|D^2V\|_\infty}} \sqrt{V}(x), \quad (5.7)$$

thus  $\tilde{\gamma}_n |b(X_{n-1})| \leq \sqrt{\tilde{\gamma}_n} \sqrt{2\delta/\|D^2V\|_\infty} \sqrt{V}(X_{n-1})$ . Since  $\text{Tr}(\sigma\sigma^*) \leq C_\sigma V^{1-a}$ ,  $a > 0$  and  $V \geq 1$  then there exists  $C > 0$  such that

$$|X_n - X_{n-1}| \leq C \sqrt{\tilde{\gamma}_n} \sqrt{V}(X_{n-1}) (1 + |U_n|),$$

and this, combined with  $\tilde{\gamma}_n \leq \gamma_n$ , implies (4.3). #

**Remark:** The control of  $\tilde{\gamma}_n |b(X_{n-1})|$  given by (5.7) is an important property of our scheme and will be often used throughout the section. \*

For the proof of Lemmas VII.2 and VII.3 we will need the following consequence of conditions (2.3) and (2.4) on  $U_1$ .

**Lemma VII.4:**

There exists  $K > 0$  such that  $\forall \theta \in [0, \tau], \forall h \in (0, 1), \forall v \in \mathbf{R}^d$ ,

$$\mathbf{E} \left[ \exp(\sqrt{h} \langle v, U_1 \rangle + h\theta |U_1|^2) \right] \leq \exp\left(\frac{h}{1-h} \frac{\kappa}{2} |v|^2 + Kh\right)$$

**Proof** (Lemma VII.4). By Hölder's inequality, we have

$$\mathbf{E}\left[\exp(\sqrt{h}\langle v, U_1 \rangle + h\theta|U_1|^2)\right] \leq \left(\mathbf{E}\left[\exp\left(\frac{\sqrt{h}}{1-h}\langle v, U_1 \rangle\right)\right]\right)^{1-h} \left(\mathbf{E}\left[\exp(\theta|U_1|^2)\right]\right)^h.$$

Since the random variable  $U_1$  is a generalized Gaussian, we have

$$\mathbf{E}\left[\exp\left(\frac{\sqrt{h}}{1-h}\langle v, U_1 \rangle\right)\right] \leq \exp\left(\frac{h}{(1-h)^2} \frac{\kappa}{2} |v|^2\right).$$

Since  $U_1$  satisfies (2.4) and  $\theta \leq \tau$ ,  $\mathbf{E}\left[\exp(\theta|U_1|^2)\right] < +\infty$ . By setting

$$K = \log\left(\mathbf{E}\left[\exp(\tau|U_1|^2)\right]\right),$$

the formula is established. #

**Proof** (Lemma VII.2). We recall that  $a \in ]0, 1]$ . By concavity of the function  $(x \mapsto x^a)$  we have

$$V^a(X_{n+1}) - V^a(X_n) \leq aV^{a-1}(X_n)(V(X_{n+1}) - V(X_n)). \quad (5.8)$$

The Taylor formula applied to  $V$  between  $X_n$  and  $X_{n+1}$  yields

$$V(X_{n+1}) \leq V(X_n) + \langle \nabla V(X_n), \Delta X_{n+1} \rangle + \frac{1}{2}\rho|\Delta X_{n+1}|^2, \quad (5.9)$$

where  $\rho = \|D^2V\|_\infty$ .

From the stability condition (2.1) and  $\Delta X_{n+1} = \tilde{\gamma}_{n+1}b(X_n) + \sqrt{\tilde{\gamma}_{n+1}}\sigma(X_n)U_{n+1}$ , there exist  $\alpha > 0$  and  $\beta > 0$  such that

$$\langle \nabla V(X_n), \Delta X_{n+1} \rangle \leq -\alpha\tilde{\gamma}_{n+1}V(X_n) + \beta\tilde{\gamma}_{n+1} + \Lambda_{n+1}^{(1)}(X_n, U_{n+1}), \quad (5.10)$$

with the notation

$$\Lambda_{n+1}^{(1)}(x, u) = \sqrt{g_{n+1}(x)}\langle \nabla V(x), \sigma(x)u \rangle.$$

On the other hand we write

$$\Lambda_{n+1}^{(2)}(x, u) = 2(g_{n+1}(x))^{3/2}\langle b(x), \sigma(x)u \rangle + g_{n+1}(x) \operatorname{Tr}(\sigma\sigma^*)(x)|u|^2,$$

so that

$$|\Delta X_{n+1}|^2 \leq \tilde{\gamma}_{n+1}^2|b(X_n)|^2 + \Lambda_{n+1}^{(2)}(X_n, U_{n+1}).$$

From (5.7) it follows that

$$|\Delta X_{n+1}|^2 \leq \frac{2\delta}{\rho}\tilde{\gamma}_{n+1}V(X_n) + \Lambda_{n+1}^{(2)}(X_n, U_{n+1}). \quad (5.11)$$

Combining (5.8), (5.9), (5.10) and (5.11) gives

$$\begin{aligned} V^a(X_{n+1}) \leq & V^a(X_n) - a(\alpha - \delta)\tilde{\gamma}_{n+1}V^a(X_n) + a\beta\tilde{\gamma}_{n+1} \\ & + aV^{a-1}(X_n)(\Lambda_{n+1}^{(1)}(X_n, U_{n+1}) + \frac{\rho}{2}\Lambda_{n+1}^{(2)}(X_n, U_{n+1})). \end{aligned} \quad (5.12)$$

Let  $\lambda \in (0, \lambda_0)$ . From (5.12) we deduce that for every  $n \geq 0$

$$\begin{aligned} \mathbf{E} \left[ \exp(\lambda V^a(X_{n+1})) \mid \mathcal{F}_n \right] \\ \leq \exp \left( \lambda V^a(X_n) - \lambda a(\alpha - \delta)\tilde{\gamma}_{n+1}V^a(X_n) + \lambda a\beta\tilde{\gamma}_{n+1} \right) Y_{n+1} \end{aligned} \quad (5.13)$$

where

$$Y_{n+1} = \mathbf{E} \left[ \exp \left( \lambda a V^{a-1}(X_n)(\Lambda_{n+1}^{(1)}(X_n, U_{n+1}) + \frac{\rho}{2}\Lambda_{n+1}^{(2)}(X_n, U_{n+1})) \right) \mid \mathcal{F}_n \right].$$

We fix  $n$  such that  $\gamma_n < 1$  which will ensure  $g_n(x) < 1$  for every  $x \in \mathbf{R}^d$ , and define  $\phi_n(x)$  by

$$\phi_n(x) = \mathbf{E} \left[ \exp \left( \lambda a V^{a-1}(x)(\Lambda_n^{(1)}(x, U_1) + \frac{\rho}{2}\Lambda_n^{(2)}(x, U_1)) \right) \right],$$

so that  $Y_{n+1} = \phi_{n+1}(X_n)$ . Setting  $v = \lambda a V^{a-1}(x)\sigma^*(x)(\nabla V(x) + \rho g_n(x)b(x))$  we obtain by the definition of  $\Lambda_n^{(1)}$  and  $\Lambda_n^{(2)}$ ,

$$\phi_n(x) = \mathbf{E} \left[ \exp \left( \sqrt{g_n(x)} \langle v, U_1 \rangle + \frac{\lambda a \rho}{2} V^{a-1}(x) g_n(x) \text{Tr}(\sigma \sigma^*)(x) |U_1|^2 \right) \right].$$

Since  $\text{Tr}(\sigma \sigma^*) \leq C_\sigma V^{1-a}$  and  $\lambda < \lambda_0 \leq \frac{2\tau}{a\rho C_\sigma}$ , we are able to apply Lemma VII.4 with  $\theta = \frac{\lambda a \rho C_\sigma}{2}$  and  $h = g_n(x)$  which yields the existence of  $K > 0$  such that

$$\phi_n(x) \leq \exp \left( \frac{g_n(x)}{1 - g_n(x)} \frac{\kappa}{2} |v|^2 + K g_n(x) \right). \quad (5.14)$$

It remains to handle  $|v|^2$ . From  $|\nabla V|^2 \leq C_V V$  and  $\langle \nabla V, b \rangle \leq \beta$  we obtain

$$|v|^2 \leq \lambda^2 a^2 V^{2a-2}(x) \text{Tr}(\sigma \sigma^*)(x) (C_V V(x) + 2\rho\beta g_n(x) + \rho^2 g_n^2(x) |b(x)|^2)$$

From (5.2), (5.7) and  $V \geq 1$  we have

$$|v|^2 \leq \lambda^2 a^2 C_\sigma V^{a-1}(x) (C_V V(x) + C g_n(x) V(x)),$$

with  $C = 2\rho(\delta + \beta)$ . In the following inequalities, the letter  $C$  is used to denote a positive constant. Since

$$\frac{g_n(x)}{1 - g_n(x)} = g_n(x) + \frac{g_n^2(x)}{1 - g_n(x)} \leq g_n(x) + \frac{g_n^2(x)}{1 - \gamma_{n_0}},$$

where  $n_0 = \min\{n; \gamma_n < 1\}$ , it follows that

$$\frac{g_n(x)}{1 - g_n(x)} |v|^2 \leq \lambda^2 a^2 C_V C_\sigma g_n(x) V^a(x) + C g_n^2(x) V^a(x). \quad (5.15)$$

Combining (5.15) with (5.14) yields

$$\phi_n(x) \leq \exp\left(\lambda^2 a^2 C_V C_\sigma \frac{\kappa}{2} g_n(x) V^a(x) + C g_n^2(x) V^a(x) + K g_n(x)\right), \quad (5.16)$$

and this inequality is true for every  $n \geq n_0$ .

Let  $\bar{\alpha} = a(\alpha - \delta) - \lambda \frac{\kappa a^2 C_V C_\sigma}{2}$ . Since  $\lambda < \lambda_0 \leq \frac{2(\alpha - \delta)}{\kappa a C_V C_\sigma}$ , we have  $\bar{\alpha} > 0$ , and combining (5.16) with (5.13) we get for every  $n \geq n_0$ ,

$$\begin{aligned} \mathbf{E}\left[\exp(\lambda V^a(X_{n+1})) \mid \mathcal{F}_n\right] &\leq \exp\left(\lambda V^a(X_n) - \lambda \bar{\alpha} \tilde{\gamma}_{n+1} V^a(X_n) \right. \\ &\quad \left. + C \tilde{\gamma}_{n+1}^2 V^a(X_n) + \bar{K} \tilde{\gamma}_{n+1}\right), \end{aligned}$$

where  $\bar{K} = \lambda a \beta + K$ . Setting  $n'_0 = \min\{n; \gamma_n < \lambda \bar{\alpha} / (2C)\}$  and  $\tilde{\alpha} = \bar{\alpha} / 2$  we have for every  $n \geq n_0 \vee n'_0$ ,

$$\mathbf{E}\left[\exp(\lambda V^a(X_{n+1})) \mid \mathcal{F}_n\right] \leq \exp\left(\lambda V^a(X_n) - \lambda \tilde{\alpha} \tilde{\gamma}_{n+1} V^a(X_n) + K \tilde{\gamma}_{n+1}\right).$$

With the notation  $\tilde{\beta} = \tilde{\alpha} \exp(K/\tilde{\alpha})$  and  $n''_0 = \min\{n; \gamma_n < 1/\tilde{\alpha}\}$  we have by convexity of the exponential function : for every  $n > n_0 \vee n'_0 \vee n''_0$ ,

$$\mathbf{E}\left[\exp(\lambda V^a(X_{n+1})) \mid \mathcal{F}_n\right] \leq (1 - \tilde{\alpha} \tilde{\gamma}_{n+1}) \exp(\lambda V^a(X_n)) + \tilde{\beta} \tilde{\gamma}_{n+1}, \quad (5.17)$$

which is the desired conclusion. #

**Proof** (Lemma VII.3). Let  $W = \exp(\lambda V^a)$  with  $\lambda < \lambda_0/s$ . As  $s > 1$ , we have by convexity

$$\begin{aligned} |W(X_n) - \mathbf{E}[W(X_n) \mid \mathcal{F}_{n-1}]|^s &\leq 2^{s-1} \left( |W(X_n) - W(X_{n-1})|^s + |\mathbf{E}[W(X_n) - W(X_{n-1}) \mid \mathcal{F}_{n-1}]|^s \right), \end{aligned}$$

and by Jensen's inequality

$$|\mathbf{E}[W(X_n) - W(X_{n-1}) \mid \mathcal{F}_{n-1}]|^s \leq \mathbf{E}[|W(X_n) - W(X_{n-1})|^s \mid \mathcal{F}_{n-1}].$$

Thus it suffices to prove

$$\sum_{n \geq 1} \left( \frac{\eta_n}{H_n \gamma_n} \right)^s \mathbf{E}\left[|W(X_n) - W(X_{n-1})|^s \mid \mathcal{F}_{n-1}\right] < +\infty \quad a.s.$$

Taylor's formula applied to the convex function ( $x \mapsto \exp(\lambda x)$ ) between  $V^a(X_n)$  and  $V^a(X_{n-1})$  yields

$$|W(X_n) - W(X_{n-1})| \leq \lambda (\exp(\lambda V^a(X_n)) + \exp(\lambda V^a(X_{n-1}))) |\Delta V^a(X_n)|, \quad (5.18)$$

where  $\Delta V^a(X_n) = V^a(X_n) - V^a(X_{n-1})$ . As in the proof of Theorem VII.1, from (5.7) and (5.2) we deduce that

$$|X_n - X_{n-1}| \leq C\sqrt{\gamma_n}\sqrt{V}(X_{n-1})(1 + |U_n|),$$

and by Lemma 2. (a) of [LP02] we get

$$|V^a(X_n) - V^a(X_{n-1})| \leq C\sqrt{\gamma_n}(\sqrt{V}(X_{n-1}))^{2a\vee 1} (1 + |U_n|^{2a\vee 1}). \quad (5.19)$$

To simplify notation, set  $A_{n-1} = (\sqrt{V}(X_{n-1}))^{2a\vee 1}$  and  $B_n = 1 + |U_n|^{2a\vee 1}$ . Plugging (5.19) into (5.18) we obtain

$$|W(X_n) - W(X_{n-1})| \leq \lambda C\sqrt{\gamma_n}(W(X_n) + W(X_{n-1}))A_{n-1}B_n,$$

and

$$\begin{aligned} \mathbf{E}[|\Delta W(X_n)|^s | \mathcal{F}_{n-1}] &\leq \lambda^s C^s \gamma_n^{s/2} \left( W^s(X_{n-1}) A_{n-1}^s \mathbf{E}[B_n^s | \mathcal{F}_{n-1}] \right. \\ &\quad \left. + A_{n-1}^s \mathbf{E}[W^s(X_n) B_n^s | \mathcal{F}_{n-1}] \right). \end{aligned} \quad (5.20)$$

By Young's inequality for any  $\varepsilon > 0$

$$\mathbf{E}[W^s(X_n) B_n^s | \mathcal{F}_{n-1}] \leq \frac{1}{1+\varepsilon} \mathbf{E}[W^{s(1+\varepsilon)}(X_n) | \mathcal{F}_{n-1}] + \frac{\varepsilon}{1+\varepsilon} K_\varepsilon$$

with  $K_\varepsilon = \mathbf{E}[B_n^{s(1+\varepsilon)/\varepsilon} | \mathcal{F}_{n-1}]$ . We choose  $\varepsilon$  such that  $\lambda s(1+\varepsilon) < \lambda_0$  and we get from Lemma VII.2 that for every  $n \geq n_0$ ,

$$\begin{aligned} \mathbf{E}[W^s(X_n) B_n^s | \mathcal{F}_{n-1}] &\leq \frac{1}{1+\varepsilon} \mathbf{E}[\exp(\lambda s(1+\varepsilon)V^a(X_n)) | \mathcal{F}_{n-1}] + \frac{\varepsilon}{1+\varepsilon} K_\varepsilon, \\ &\leq \frac{1}{1+\varepsilon} \left( \exp(\lambda s(1+\varepsilon)V^a(X_{n-1})) + \tilde{\beta}\tilde{\gamma}_n + K_\varepsilon \varepsilon \right), \end{aligned}$$

Combining this with (5.20) we obtain

$$\begin{aligned} \mathbf{E}[|\Delta W(X_n)|^s | \mathcal{F}_n] &\leq \lambda^s C^s \gamma_n^{s/2} \left( W^s(X_{n-1}) A_{n-1}^s \mathbf{E}[B_1^s] \right. \\ &\quad \left. + \frac{1}{1+\varepsilon} W^{s(1+\varepsilon)}(X_{n-1}) A_{n-1}^s + C_\varepsilon A_{n-1}^s \right). \end{aligned}$$

Keeping in mind that  $A_{n-1} = (\sqrt{V}(X_{n-1}))^{2a\vee 1}$  one checks that there exists  $\bar{\lambda} \in (\lambda s(1+\varepsilon), \lambda_0)$  and  $\tilde{K} > 0$  such that

$$\mathbf{E}[|W(X_n) - W(X_{n-1})|^s | \mathcal{F}_{n-1}] \leq \tilde{K} \gamma_n^{s/2} \exp(\bar{\lambda} V^a(X_{n-1})).$$

Consequently, it remains to prove that

$$\sum_{n \geq 1} \left( \frac{\eta_n}{H_n \sqrt{\gamma_n}} \right)^s \exp(\bar{\lambda} V^a(X_{n-1})) < +\infty \quad a.s. \quad (5.21)$$

We consider the nonincreasing sequence  $(\theta_n)_{n \geq 1} = \left(\frac{1}{\gamma_n} \left(\frac{n_n}{H_n \sqrt{\gamma_n}}\right)^s\right)_{n \geq 1}$ . From (4.1) and  $\tilde{\gamma}_n = \gamma_n$  for every  $n \geq n_1$ , we can apply the Proposition VII.1 with  $\gamma_n$ ,  $\theta_n$  and  $W = \exp(\bar{\lambda}V^a)$  which gives (5.21). #

## 6. Dissipative Hamiltonian systems

In this section, we consider a stochastic differential system of the type

$$\begin{cases} dq_t = \partial_p H(q_t, p_t) dt, \\ dp_t = -\partial_q H(q_t, p_t) dt - F(q_t, p_t) \partial_p H(q_t, p_t) dt + c(q_t, p_t) dW_t, \end{cases} \quad (6.1)$$

with  $(q(0), p(0)) = (q_0, p_0) \in \mathbf{R}^{2d}$ . The process  $(W_t)_{t \geq 0}$  is a  $d$ -dimensional brownian motion and  $c$  is continuous on  $\mathbf{R}^{2d}$  with values in the set of  $d \times d$  matrices. The *Hamiltonian*  $H$  is of class  $\mathcal{C}^2$  on  $\mathbf{R}^{2d}$  with real values and the function  $F$  is of class  $\mathcal{C}^2$  on  $\mathbf{R}^{2d}$  with values in the set of  $d \times d$  matrices.

We write this system in the abstract form (1.1) where

$$x = \begin{pmatrix} q \\ p \end{pmatrix} \in \mathbf{R}^{2d}, \quad b(x) = \begin{pmatrix} \partial_p H(x) \\ -\partial_q H(x) - F(x) \partial_p H(x) \end{pmatrix} = \begin{pmatrix} b_1(x) \\ b_2(x) \end{pmatrix},$$

$$\text{and } \sigma(x) = \begin{pmatrix} 0 \\ c(x) \end{pmatrix}.$$

We recall that we work always under Assumption 1. The existence of a Lyapounov function  $V$  is a natural hypothesis for dissipative Hamiltonian systems, and in many cases we can determine  $V$  using the Hamiltonian  $H$ . For example, for the Langevin equation (equation (6.1) with  $F(x) = \gamma > 0$ ,  $c(x) = c$  invertible and  $H(q, p) = \frac{|p|^2}{2} + g(q)$  with  $g \in \mathcal{C}^\infty(\mathbf{R}^d; \mathbf{R})$  a polynomial function growing at infinity like  $|q|^{2l}$ ,  $l \geq 1$ ), the Lyapounov function is defined for every  $x = (q, p) \in \mathbf{R}^{2d}$  by  $V(x) = H(x) + \frac{\gamma}{2} \langle p, q \rangle + \frac{\gamma^2}{4} |q|^2 + 1$  (see [MSH02]).

In this section we assume that :

**Assumption 4.** The function  $V$  satisfies

$$\exists C_V > 0, \quad |\partial_p V|^2 \leq C_V V \quad \text{and} \quad \sup_{(q,p) \in \mathbf{R}^{2d}} \|\partial_{pp}^2 V(q, p)\| < +\infty. \quad (6.2)$$

\*

Typical Lyapounov function of Hamiltonian systems may have arbitrary polynomial growth with respect to  $q$  but are essentially quadratic with respect to  $p$ . So in such a framework the following assumption is natural

$$\text{Tr}(\sigma^*(\nabla V)^{\otimes 2} \sigma) = \text{Tr}(c^*(\partial_p V)^{\otimes 2} c) \leq C_V V \text{Tr}(c^* c),$$

and

$$\sup_{x \in \mathbf{R}^{2d}} \text{Tr}(\sigma^* D^2 V \sigma)(x) = \sup_{x \in \mathbf{R}^{2d}} \text{Tr}(c^* \partial_{pp}^2 V c)(x) \leq \rho \sup_{x \in \mathbf{R}^{2d}} \text{Tr}(c^* c)(x),$$

where  $\rho = \sup_{(q,p) \in \mathbf{R}^{2d}} |\partial_{pp}^2 V(q,p)|$ . Thus condition (2.2) about  $\sigma$  is satisfied as soon as

$$\exists C_\sigma > 0, a \in ]0, 1], \quad \text{Tr}(cc^*) = \text{Tr}(\sigma\sigma^*) \leq C_\sigma V^{1-a}. \quad (6.3)$$

These assumptions are very weak and are satisfied by a large class of examples derived from perturbed Hamiltonian systems. For a general model for the Hamiltonian and many examples (essentially multidimensional oscillators) we refer to [Soi94] (page 10 for hypothesis on the Hamiltonian).

Our scheme  $(X_n)_{n \geq 0}$  is built applying the recursive procedure (1.4) with the random step sequence  $(\tilde{\gamma}_n)_{n \geq 1}$  defined by

$$\forall n \geq 1, \tilde{\gamma}_n = g_n(X_{n-1}) \quad (6.4)$$

with  $g_n(x) = \gamma_n \wedge \chi_n(x)$  where  $\chi_n : \mathbf{R}^{2d} \rightarrow \mathbf{R}_+^*$ . Throughout the section, we consider the following function  $\psi_n(x) : \mathbf{R}^{2d} \rightarrow \mathbf{R}$  defined by

$$\forall x \in \mathbf{R}^d, \quad \psi_n(x) = \sup_{\bar{q} \in (q, q + \gamma_n b_1(x))} \|D^2 V(\bar{q}, p)\| \vee 1. \quad (6.5)$$

We introduce the notation  $X_n = (Q_n, P_n)$  so that

$$\begin{cases} Q_{n+1} = Q_n + \tilde{\gamma}_{n+1} b_1(X_n), \\ P_{n+1} = P_n + \tilde{\gamma}_{n+1} b_2(X_n) + \sqrt{\tilde{\gamma}_{n+1}} c(X_n) U_{n+1}, \end{cases} \quad (6.6)$$

and  $(Q_0, P_0) = (q_0, p_0)$ . Remark that  $(Q_n)_{n \geq 0}$  is  $(\mathcal{F}_n)_{n \geq 0}$ -predictable.

The principal result is the following Theorem.

**Theorem VII.2:**

Let  $(\gamma_n)_{n \geq 1}$  and  $(\eta_n)_{n \geq 1}$  satisfy Assumption 2. Suppose that there exist  $l > 1$  and  $C_b \geq 1$  such that  $\psi_n(x) |b(x)|^2 \leq C_b V^l(x)$  and that  $\chi$  satisfies

$$\forall x \in \mathbf{R}^d, \quad \zeta V^{-p}(x) \leq \chi_n(x) \leq \frac{2\delta V(x)}{\psi_n(x) (|b(x)|^2 \vee 1)}, \quad (6.7)$$

with  $\zeta \leq \frac{2\delta}{C_b}$ ,  $p \geq l - 1$  and  $\delta \in (0, \alpha/4)$ . If  $(\gamma_n)_{n \geq 1}$  satisfies

$$\sum_{n \geq 1} \gamma_n \exp(-\lambda \zeta^{\frac{a}{p}} \gamma_n^{-\frac{a}{p}}) < +\infty, \quad (6.8)$$

and  $\lambda_0 = \frac{\tau}{a C_\sigma \|\partial_{pp}^2 V\|_\infty} \wedge \frac{2(\alpha - 4\delta)}{\kappa a C_V C_\sigma}$  then

- there exists an almost surely finite random variable  $n_1$  such that  $\forall n \geq n_1, \tilde{\gamma}_n = \gamma_n$ ,
- for every  $\lambda < \frac{\lambda_0}{s}$  (whith  $s$  as in (4.2)),

$$\sup_n \nu_n^\eta(\exp(\lambda V^a)) < +\infty \quad a.s.$$

and any weak limit of  $(\nu_n^\eta)_{n \geq 1}$  is an invariant distribution for (1.1).

The proof of this Theorem is essentially the same as the proof of Theorem VII.1. Lemma VII.5 gives condition on  $\chi$  so that  $W = \exp(\lambda V^a)$  satisfy (3.1) and Lemma VII.6 allows to apply Theorem VII.1.

**Lemma VII.5:**

Assume that

$$\forall x \in \mathbf{R}^{2d}, \quad \chi_n(x) \leq \frac{2\delta V(x)}{\psi_n(x)(|b(x)|^2 \vee 1)},$$

where  $\delta \in (0, \alpha/4)$ . Let  $\lambda_0$  be as in Theorem VII.2. Then for every  $\lambda < \lambda_0$  there exist  $\tilde{\alpha} > 0$ ,  $\tilde{\beta} > 0$ , and  $n_0 \geq 0$  such that for every  $n \geq n_0$ ,

$$\frac{\mathbf{E}[\exp(\lambda V^a(X_{n+1})) | \mathcal{F}_n] - \exp(\lambda V^a(X_n))}{\tilde{\gamma}_{n+1}} \leq -\tilde{\alpha} \exp(\lambda V^a(X_n)) + \tilde{\beta}.$$

**Lemma VII.6:**

Assume that there exists an almost sure finite random variable  $n_1$  such that  $\tilde{\gamma}_n = \gamma_n$  for every  $n \geq n_1$ . Let  $\lambda_0$  be as in Theorem VII.2. Then for every  $\lambda < \frac{\lambda_0}{s}$  the series

$$\sum_{n \geq 1} \left( \frac{\eta_n}{H_n \gamma_n} \right)^s \mathbf{E} \left[ \left| \exp(\lambda V^a(X_n)) - \mathbf{E}[\exp(\lambda V^a(X_n)) | \mathcal{F}_{n-1}] \right|^s | \mathcal{F}_{n-1} \right]$$

is almost surely finite. ♦

A first property of our scheme is given by the following Lemma.

**Lemma VII.7:**

If  $\delta \leq \alpha$  and

$$\forall x \in \mathbf{R}^{2d}, \quad \chi_n(x) \leq \frac{2\delta V(x)}{\psi_n(x)(|b(x)|^2 \vee 1)}, \quad (6.9)$$

then for every  $x = (q, p) \in \mathbf{R}^{2d}$ ,

$$V(q + g_n(x)b_1(x), p) \leq (1 + \sqrt{2\delta C_V} \sqrt{g_n(x)})V(x) + g_n(x)\beta.$$

Proof. By the Taylor's formula we obtain

$$\begin{aligned} V(q + g_n(x)b_1(x), p) - V(x) &\leq g_n(x) \langle \partial_q V(x), b_1(x) \rangle \\ &\quad + \frac{1}{2} \|\partial_{qq}^2 V(\bar{q}_n, p)\| |g_n(x)b_1(x)|^2, \end{aligned} \quad (6.10)$$

where  $\bar{q}_n \in (q, q + g_n(x)b_1(x))$ . Since  $g_n(x) = \gamma_n \wedge \chi_n(x)$  we have

$$g_n^2(x) |b_1(x)|^2 \leq g_n(x) \chi_n(x) |b_1(x)|^2 \leq g_n(x) \frac{2\delta V(x)}{\psi_n(x)}. \quad (6.11)$$

Moreover,  $\|\partial_{qq}^2 V(\bar{q}_n, p)\| \leq \|D^2 V(\bar{q}_n, p)\| \leq \psi_n(x)$  and by (6.10)

$$V(q + g_n(x)b_1(x), p) - V(x) \leq g_n(x) \langle \partial_q V(x), b_1(x) \rangle + g_n(x) \delta V(x). \quad (6.12)$$

Writing  $\langle \partial_q V(x), b_1(x) \rangle = \langle \nabla V, b \rangle(x) - \langle \partial_p V, b_2 \rangle(x)$  and using the stability condition (2.1) we get

$$V(q + g_n(x)b_1(x), p) - V(x) \leq g_n(x)(\beta - \alpha V(x) - \langle \partial_p V(x), b_2(x) \rangle) + g_n(x)\delta V(x).$$

Since  $\delta \leq \alpha$  and  $|\partial_p V|^2 \leq C_V V$  we have

$$V(q + g_n(x)b_1(x), p) - V(x) \leq g_n(x)\sqrt{C_V}\sqrt{V(x)}|b_2(x)| + g_n(x)\beta,$$

and if  $|b_2(x)| \leq \frac{\sqrt{2\delta}}{\sqrt{g_n(x)}}\sqrt{V(x)}$  then

$$V(q + g_n(x)b_1(x), p) - V(x) \leq \sqrt{g_n(x)}\sqrt{2\delta C_V}V(x) + g_n(x)\beta,$$

else  $\sqrt{V(x)} < \frac{\sqrt{g_n(x)}}{\sqrt{2\delta}}|b_2(x)|$  and

$$\begin{aligned} V(q + g_n(x)b_1(x), p) - V(x) &\leq g_n(x)\sqrt{g_n(x)}\sqrt{\frac{C_V}{2\delta}}|b_2(x)|^2 + g_n(x)\beta, \\ &\leq \sqrt{g_n(x)}\sqrt{\frac{C_V}{2\delta}}\chi_n(x)|b_2(x)|^2 + g_n(x)\beta. \end{aligned}$$

From (6.9) we deduce that

$$V(q + g_n(x)b_1(x), p) - V(x) \leq \sqrt{g_n(x)}\sqrt{2\delta C_V}V(x) + g_n(x)\beta,$$

and the Lemma is proved. #

Proof (Lemma VII.5). First, remark that by the concavity of the function  $(x \mapsto x^a)$  we have

$$V^a(X_{n+1}) - V^a(X_n) \leq aV^{a-1}(X_n)(V(X_{n+1}) - V(X_n)), \quad (6.13)$$

and that

$$V(X_{n+1}) - V(X_n) = V(X_{n+1}) - V(Q_{n+1}, P_n) + V(Q_{n+1}, P_n) - V(X_n).$$

In the proof of Lemma VII.7 we proved (6.12) i.e.

$$V(Q_{n+1}, P_n) - V(X_n) \leq \tilde{\gamma}_{n+1}\langle \partial_q V(X_n), b_1(X_n) \rangle + \tilde{\gamma}_{n+1}\delta V(X_n). \quad (6.14)$$

We now study  $V(X_{n+1}) - V(Q_{n+1}, P_n)$ . We apply Taylor's formula to  $V$  which gives

$$\begin{aligned} V(X_{n+1}) - V(Q_{n+1}, P_n) &\leq \langle \partial_p V(Q_{n+1}, P_n), \Delta P_{n+1} \rangle + \\ &\quad \frac{1}{2} \sup_{p \in (P_n, P_{n+1})} \|\partial_{pp}^2 V(Q_{n+1}, p)\| |\Delta P_{n+1}|^2. \end{aligned} \quad (6.15)$$

On the one hand, we have

$$\begin{aligned} \langle \partial_p V(Q_{n+1}, P_n), \Delta P_{n+1} \rangle &= \tilde{\gamma}_{n+1} \langle \partial_p V(X_n), b_2(X_n) \rangle \\ &\quad + \tilde{\gamma}_{n+1} \langle \partial_p V(Q_{n+1}, P_n) - \partial_p V(X_n), b_2(X_n) \rangle \\ &\quad + \sqrt{\tilde{\gamma}_{n+1}} \langle \partial_p V(Q_{n+1}, P_n), c(X_n) U_{n+1} \rangle, \end{aligned}$$

and using

$$\begin{aligned} |\partial_p V(Q_{n+1}, P_n) - \partial_p V(X_n)| &\leq \|\partial_{qp}^2 V(q, P_n)\| |\Delta Q_{n+1}|, \\ &\leq \tilde{\gamma}_{n+1} \|D^2 V(q, P_n)\| |b_1(X_n)|, \end{aligned}$$

where  $q \in (Q_n, Q_{n+1})$  we obtain

$$\begin{aligned} \langle \partial_p V(Q_{n+1}, P_n), \Delta P_{n+1} \rangle &\leq \tilde{\gamma}_{n+1} \langle \partial_p V(X_n), b_2(X_n) \rangle \\ &\quad + \tilde{\gamma}_{n+1}^2 \|D^2 V(q, P_n)\| |b_1(X_n)| |b_2(X_n)| \\ &\quad + \sqrt{\tilde{\gamma}_{n+1}} \langle \partial_p V(Q_{n+1}, P_n), c(X_n) U_{n+1} \rangle. \end{aligned}$$

Since  $\tilde{\gamma}_{n+1} = \gamma_{n+1} \wedge \chi_{n+1}(X_n)$  we have

$$\tilde{\gamma}_{n+1} \leq \chi_{n+1}(X_n) \leq \frac{2\delta V(X_n)}{\sup_{q \in (Q_n, Q_{n+1})} \|D^2 V(q, P_n)\| (|b(X_n)|^2 \vee 1)} \quad (6.16)$$

and from  $|b_1(X_n)| |b_2(X_n)| \leq \frac{1}{2} |b(X_n)|^2$  we obtain

$$\begin{aligned} \langle \partial_p V(Q_{n+1}, P_n), \Delta P_{n+1} \rangle &\leq \tilde{\gamma}_{n+1} \langle \partial_p V(X_n), b_2(X_n) \rangle + \tilde{\gamma}_{n+1} \delta V(X_n) \\ &\quad + \sqrt{\tilde{\gamma}_{n+1}} \langle \partial_p V(Q_{n+1}, P_n), c(X_n) U_{n+1} \rangle. \end{aligned} \quad (6.17)$$

On the other hand, setting  $\rho = \sup_{x \in \mathbf{R}^{2d}} \|\partial_{pp}^2 V(x)\|$  and using (6.6) we have

$$\begin{aligned} \sup_{p \in (P_n, P_{n+1})} \|\partial_{pp}^2 V(Q_{n+1}, p)\| |\Delta P_{n+1}|^2 \\ \leq 2\tilde{\gamma}_{n+1}^2 \sup_{p \in (P_n, P_{n+1})} \|D^2 V(Q_{n+1}, p)\| |b_2(X_n)|^2 \\ + 2\tilde{\gamma}_{n+1} \rho \operatorname{Tr}(cc^*)(X_n) |U_{n+1}|^2. \end{aligned}$$

By the definition of  $\tilde{\gamma}_{n+1}$  it follows that

$$\begin{aligned} \sup_{p \in (P_n, P_{n+1})} \|\partial_{pp}^2 V(Q_n, p)\| |\Delta P_{n+1}|^2 &\leq 4\tilde{\gamma}_{n+1} \delta V(X_n) \\ &\quad + 2\tilde{\gamma}_{n+1} \rho \operatorname{Tr}(cc^*)(X_n) |U_{n+1}|^2. \end{aligned} \quad (6.18)$$

Finally, combining (6.15) with (6.17) and (6.18) we obtain

$$V(X_{n+1}) - V(Q_{n+1}, P_n) \leq \tilde{\gamma}_{n+1} \langle \partial_p V(X_n), b_2(X_n) \rangle + 3\delta \tilde{\gamma}_{n+1} V(X_n) + Y_{n+1}, \quad (6.19)$$

where

$$Y_{n+1} = \sqrt{\tilde{\gamma}_{n+1}} \langle \partial_p V(Q_{n+1}, P_n), c(X_n) U_{n+1} \rangle + \rho \tilde{\gamma}_{n+1} \operatorname{Tr}(cc^*)(X_n) |U_{n+1}|^2.$$

From (6.14) and (6.19) we have

$$V(X_{n+1}) - V(X_n) \leq \tilde{\gamma}_{n+1} \langle \nabla V, b \rangle(X_n) + 4\delta \tilde{\gamma}_{n+1} V(X_n) + Y_{n+1},$$

and by the stability condition (2.1) there exists  $\alpha > 0$  such that

$$\begin{aligned} V(X_{n+1}) &\leq V(X_n) - \alpha \tilde{\gamma}_{n+1} V(X_n) + 4\delta \tilde{\gamma}_{n+1} V(X_n) + \beta \tilde{\gamma}_{n+1} + Y_{n+1}, \\ &= V(X_n) - (\alpha - 4\delta) \tilde{\gamma}_{n+1} V(X_n) + \beta \tilde{\gamma}_{n+1} + Y_{n+1}. \end{aligned}$$

By (6.13) and  $V \geq 1$  we get (using that  $V^{a-1} \leq 1$ )

$$V^a(X_{n+1}) \leq V^a(X_n) - a(\alpha - 4\delta) \tilde{\gamma}_{n+1} V^a(X_n) + a\beta \tilde{\gamma}_{n+1} + aV^{a-1}(X_n) Y_{n+1}.$$

Let  $\lambda \in (0, \lambda_0)$ . Denoting

$$Z_{n+1} = \mathbf{E} \left[ \exp(a\lambda V^{a-1}(X_n) Y_{n+1}) \mid \mathcal{F}_n \right],$$

we have

$$\begin{aligned} &\mathbf{E} \left[ \exp(\lambda V^a(X_{n+1})) \mid \mathcal{F}_n \right] \\ &\leq \exp(\lambda V^a(X_n) - a\lambda(\alpha - 4\delta) \tilde{\gamma}_{n+1} V^a(X_n) + a\lambda\beta \tilde{\gamma}_{n+1}) Z_{n+1}. \end{aligned} \quad (6.20)$$

It remains to study  $Z_n$ . We define  $\phi_n(x)$  by

$$\phi_n(x) = \mathbf{E} \left[ \exp(\sqrt{g_n(x)} \langle v, U_1 \rangle + \lambda a \rho g_n(x) V^{a-1}(x) \operatorname{Tr}(cc^*)(x) |U_1|^2) \right] \quad (6.21)$$

with  $v = \lambda a V^{a-1}(x) c^*(x) \partial_p V(q + g_n(x) b_1(x), p)$ , so that  $Z_{n+1} = \phi_{n+1}(X_n)$ . Using  $\operatorname{Tr}(cc^*) \leq C_\sigma V^{1-a}$  we have

$$\phi_n(x) \leq \mathbf{E} \left[ \exp(\sqrt{g_n(x)} \langle v, U_1 \rangle + \lambda a \rho g_n(x) C_\sigma |U_1|^2) \right].$$

Let  $n_0 = \min\{n \geq 0; \gamma_n < 1\}$ . Since  $\lambda < \lambda_0 \leq \frac{\tau}{a\rho C_\sigma}$ , we are able to apply lemma VII.4 with  $\theta = \lambda a \rho C_\sigma$  and  $h = g_n(x)$  which gives the existence of  $K > 0$  such that for every  $n \geq n_0$

$$\phi_n(x) \leq \exp\left(\frac{g_n(x)}{1 - g_n(x)} \frac{\kappa}{2} |v|^2 + K g_n(x)\right). \quad (6.22)$$

Moreover, it follows from  $|\partial_p V| \leq C_V V$  that

$$\begin{aligned} |v|^2 &\leq \lambda^2 a^2 C_\sigma V^{a-1}(x) |\partial_p V(q + g_n(x) b_1(x), p)|^2, \\ &\leq \lambda^2 a^2 C_\sigma C_V V^{a-1}(x) V(q + g_n(x) b_1(x), p), \end{aligned}$$

The Lemma VII.7 gives

$$|v|^2 \leq \lambda^2 a^2 C_\sigma C_V V^a(x) + C \sqrt{g_n(x)} V^a(x) + g_n(x) \beta.$$

In the same manner as in the proof of Lemma VII.2 there exists  $C > 0$  and  $K > 0$  such that for every  $n \geq n_0$

$$\phi_n(x) \leq \exp\left(\frac{\kappa}{2} \lambda^2 a^2 C_\sigma C_V g_n(x) V^a(x) + C g_n^{3/2}(x) V^a(x) + K g_n(x)\right)$$

then from (6.20) we have

$$\begin{aligned} \mathbf{E}\left[\exp(\lambda V^a(X_{n+1})) \mid \mathcal{F}_n\right] &\leq \exp(\lambda V^a(X_n)) \\ &\quad - a\lambda(\alpha - 4\delta - \frac{\kappa}{2} \lambda a C_\sigma C_V) \tilde{\gamma}_{n+1} V^a(X_n) \\ &\quad + C \tilde{\gamma}_{n+1}^{3/2} V^a(X_n) + K \tilde{\gamma}_{n+1}. \end{aligned}$$

Let  $\bar{\alpha} = a(\alpha - 4\delta - \frac{\kappa}{2} \lambda a C_\sigma C_V)$ . Since  $\lambda < \lambda_0 \leq \frac{2(\alpha - 4\delta)}{\kappa a C_\sigma C_V}$  we have  $\bar{\alpha} > 0$ . Setting  $n'_0 = \min\{n; \sqrt{\gamma_n} < a\lambda\bar{\alpha}/(2C)\}$  and  $\tilde{\alpha} = (\lambda a \bar{\alpha})/2$  we have for every  $n \geq n_0 \vee n'_0$

$$\mathbf{E}\left[\exp(\lambda V^a(X_{n+1})) \mid \mathcal{F}_n\right] \leq \exp(\lambda V^a(X_n) - \tilde{\alpha} \tilde{\gamma}_{n+1} V^a(X_n) + C \tilde{\gamma}_{n+1}),$$

which proves the lemma (by the convexity of the exponential). #

**Proof (Lemma VII.6).** Let  $W = \exp(\lambda V^a)$  with  $\lambda < \lambda_0/s$ . We recall that  $Q_n$  is  $\mathcal{F}_{n-1}$  measurable. Since  $s > 1$ , the convexity of  $(x \mapsto x^s)$  implies that

$$\mathbf{E}\left[|W(X_n) - \mathbf{E}[W(X_n) \mid \mathcal{F}_{n-1}]|^s \mid \mathcal{F}_{n-1}\right] \leq 2^s \mathbf{E}\left[|W(X_n) - W(Q_n, P_{n-1})|^s \mid \mathcal{F}_{n-1}\right].$$

First we prove that

$$|V^a(X_n) - V^a(Q_n, P_{n-1})| \leq C \sqrt{\gamma_n} A_{n-1} B_n \tag{6.23}$$

with

$$\begin{aligned} A_{n-1} &= (\sqrt{V}(Q_n, P_{n-1}))^{(2a-1)_+} (\sqrt{V}(X_{n-1}))^{(2a\vee 1)}, \\ B_n &= 1 + |U_n|^{2a\vee 1}. \end{aligned}$$

Since  $\Delta P_n = \tilde{\gamma}_n b_2(X_{n-1}) + \sqrt{\gamma_n} c(X_{n-1}) U_n$  and  $\text{Tr}(cc^*) \leq C_\sigma V^{1-a}$ , we have

$$|\Delta P_n| \leq \tilde{\gamma}_n |b_2(X_{n-1})| + \sqrt{\tilde{\gamma}_n} C_\sigma (\sqrt{V}(X_{n-1}))^{1-a} |U_n|.$$

Using  $\tilde{\gamma}_n \leq \chi_n(X_{n-1}) \leq \frac{2\delta V(X_{n-1})}{|b(X_{n-1})|^2 \vee 1}$  we obtain

$$\begin{aligned} |\Delta P_n| &\leq \sqrt{\tilde{\gamma}_n} \sqrt{2\delta} \sqrt{V}(X_{n-1}) + \sqrt{\tilde{\gamma}_n} C_\sigma (\sqrt{V}(X_{n-1}))^{1-a} |U_n|, \\ &\leq C \sqrt{\tilde{\gamma}_n} \sqrt{V}(X_{n-1}) (1 + |U_n|). \end{aligned} \tag{6.24}$$

From  $|\partial_p V|^2 \leq C_V V$  we deduce that  $|\partial_p V^a| \leq a\sqrt{C_V}V^{a-\frac{1}{2}}$  so that the application  $(p \mapsto V^a(q, p))$  is Lipschitz for every  $a \leq \frac{1}{2}$ . Hence, if  $a \leq \frac{1}{2}$  we have (6.23) with  $A_{n-1} = \sqrt{V}(X_{n-1})$  and  $B_n = 1 + |U_n|$ . If  $a > \frac{1}{2}$ , by Taylor's formula we have

$$\begin{aligned} V^a(X_n) - V^a(Q_n, P_{n-1}) &= \langle \partial_p V^a(Q_n, \bar{p}_n), \Delta P_n \rangle, \\ &\leq a\sqrt{C_V}(\sqrt{V}(Q_n, \bar{p}_n))^{2a-1} |\Delta P_n|, \end{aligned} \quad (6.25)$$

with  $\bar{p}_n \in (P_{n-1}, P_n)$  and since  $2a - 1 \leq 1$

$$\begin{aligned} (\sqrt{V}(Q_n, \bar{p}_n))^{2a-1} &\leq (\sqrt{V}(Q_n, P_{n-1}) + C|\Delta P_n|)^{2a-1}, \\ &\leq (\sqrt{V}(Q_n, P_{n-1}))^{2a-1} + C|\Delta P_n|^{2a-1}. \end{aligned} \quad (6.26)$$

Plugging (6.26) in (6.25) we get

$$\begin{aligned} |V^a(X_n) - V^a(Q_n, P_{n-1})| &\leq C\sqrt{\gamma_n}(\sqrt{V}(Q_n, P_{n-1}))^{2a-1}\sqrt{V}(X_{n-1})(1 + |U_n|) \\ &\quad + C\gamma_n^a(\sqrt{V}(X_{n-1}))^{2a}(1 + |U_n|)^{2a} \end{aligned}$$

Since  $V \geq 1$  and  $2a \geq 1$ , it is easy to check that (6.23) is satisfied with  $B_n = 1 + |U_n|^{2a}$  and  $A_{n-1} = (\sqrt{V}(Q_n, P_{n-1}))^{2a-1}(\sqrt{V}(X_{n-1}))^{2a}$ .

By (6.23) and Taylor's formula applied to the convex function  $(x \mapsto \exp(\lambda x))$  between  $V^a(X_n)$  and  $V^a(Q_n, P_{n-1})$  we have

$$|W(X_n) - W(Q_n, P_{n-1})| \leq \lambda C\sqrt{\gamma_n}(W(X_n) + W(Q_n, P_{n-1}))A_{n-1}B_n,$$

and then

$$\begin{aligned} \mathbf{E}[|W(X_n) - W(Q_n, P_{n-1})|^s | \mathcal{F}_{n-1}] &\leq \lambda^s C^s \gamma_n^{s/2} A_{n-1}^s \mathbf{E}[W^s(X_n) B_n^s | \mathcal{F}_{n-1}] \\ &\quad + \lambda^s C^s \gamma_n^{s/2} A_{n-1}^s W^s(Q_n, P_{n-1}) \mathbf{E}[B_n^s]. \end{aligned} \quad (6.27)$$

As in the proof of Lemma VII.3 we prove that for every  $\varepsilon > 0$  there exists  $C_\varepsilon > 0$  such that

$$\mathbf{E}[W^s(X_n) B_n^s | \mathcal{F}_{n-1}] \leq \frac{1}{1 + \varepsilon} \exp(\lambda s(1 + \varepsilon)V^a(X_{n-1})) + C_\varepsilon. \quad (6.28)$$

To complete the proof we apply Lemma VII.7 which gives

$$V(Q_n, P_{n-1}) \leq (1 + \sqrt{2\delta C_V} \sqrt{\tilde{\gamma}_n})V(X_{n-1}) + \tilde{\gamma}_n \beta,$$

and the concavity of  $(x \mapsto x^a)$  implies

$$V^a(Q_n, P_{n-1}) \leq (1 + a\sqrt{2\delta C_V} \sqrt{\tilde{\gamma}_n})V^a(X_{n-1}) + \tilde{\gamma}_n a \beta.$$

There exists  $n_0 > 0$  such that for every  $n \geq n_0$ ,  $\lambda s(1 + a\sqrt{2\delta C_V} \sqrt{\tilde{\gamma}_n}) \leq \bar{\lambda} < \lambda_0$ , and thus

there exists  $K > 0$  such that

$$W^s(Q_n, P_{n-1}) \leq K \exp(\bar{\lambda} V^a(X_{n-1})).$$

Plugging this and (6.28) in (6.27) and using the condition (4.2) and Lemma VII.1 we obtain the result. #

## 7. Numerical experiments

The aim of this section is to validate our scheme numerically. We consider two problems : the Lorenz equations under external random excitation and a perturbed Hamiltonian system, which illustrate results in Sections 4 and 5 respectively. In both cases, we compute  $\nu_n^\eta(f)$  for a given function  $f$  using our scheme and we compare it with  $\mathbf{E}[f(\bar{X}_h^T)]$  where  $(\bar{X}_h^T)$  is a discretization scheme with constant step  $h$  (an Euler scheme or an implicit Euler scheme). The approximation of  $\mathbf{E}[f(\bar{X}_h^T)]$  is given by a Monte-Carlo procedure with 10000 paths. We give a representation of the stochastic sequence  $(\tilde{\gamma}_n)_{n \geq 1}$  in Fig. VII.2 and VII.5.

The programs are in C using BLAS/LAPACK (see <http://www.netlib.org>) for linear algebra routine and the GSL library (see <http://www.gnu.org>). In particular, the approximation of the fixed point needed in the implicit Euler scheme is done by the function `gsl_multiroot_solver`. The random generator is a Mersenne twister generator of period  $2^{19937} - 1$  taken from <http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/emt.html>. All simulations are achieved using a Gaussian white noise for  $(U_n)_{n \geq 1}$ . The deterministic part  $\gamma_n$  of the implemented step sequence is  $\gamma_n = \gamma_0 n^{-1/3}$  (with  $\gamma_0 > 0$ ) and the weight sequence used is  $(\eta_n)_{n \geq 1} = (\gamma_n)_{n \geq 1}$ .

### 7.1. Lorenz equations

It is a dissipative problem which is related to Section 4. The equations are

$$\begin{cases} dx_t = 10(y_t - x_t)dt + dW_t^1, \\ dy_t = (28x_t - y_t - x_t z_t)dt + dW_t^2, \\ dz_t = (x_t y_t - (8/3)z_t)dt \end{cases}$$

and a Lyapounov function for this system is  $V(u) = |u|^2 + 1$ . The function  $\chi$  used for simulations is defined for every  $u \in \mathbf{R}^3$  by  $\chi(u) = \frac{2V(u)}{|b(u)|^2 \vee 1}$ .

The stochastic step sequence used for our scheme is thus

$$\forall n \geq 1, \quad \tilde{\gamma}_n = (\gamma_0 n^{-\frac{1}{3}}) \wedge \left( \frac{2V(X_{n-1})}{|b(X_{n-1})|^2 \vee 1} \right). \quad (7.1)$$

The weight sequence  $(\eta_n)_{n \geq 1}$  is equal to  $(\gamma_n)_{n \geq 1}$  and we compute for  $f(x) = |x|^2$  and  $n \leq 10^7$

$$\nu_n^\eta(f) = \frac{1}{\sum_{k=1}^n k^{-\frac{1}{3}}} \sum_{k=1}^n k^{-\frac{1}{3}} f(X_{k-1}). \quad (7.2)$$

The result is given by the figure VII.1. The scheme seems to have a better behaviour when

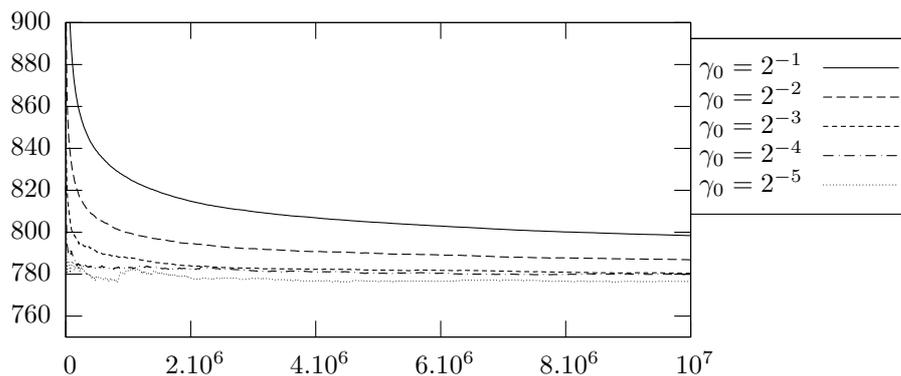


Figure VII.1 – One path of  $(\nu_n^\eta(f))_{1 \leq n \leq 10^7}$  for different value of  $\gamma_0$ .

$\gamma_0$  is equals to  $2^{-4}$  or  $2^{-5}$ . For bigger values of  $\gamma_0$ , the rate of convergence is poor but the Lorenz problem is a difficult numerical problem and the parameter  $\gamma_0$  is hard to fix. Other numerical methods have the same problem. The important point to note here is that the scheme does not explode (for any  $\gamma_0$ ) and appears convergent to the same limit.

Figure VII.2 gives a representation of the stochastic step sequence  $(\tilde{\gamma}_n)_{n \geq 1}$  when  $\gamma_0 = 0.5$ . We show that the bigger  $n$  is and the less the stochastic part  $\chi(X_{n-1})$  is used. For this path, after 20 000 iterations we have  $\tilde{\gamma}_n = \gamma_n$  (at least until  $n = 10^7$ ). Moreover before the 20 000th iteration the event  $\{\chi(X_{n-1}) < \gamma_n\}$  occurs only 924 times (4.62% of time).

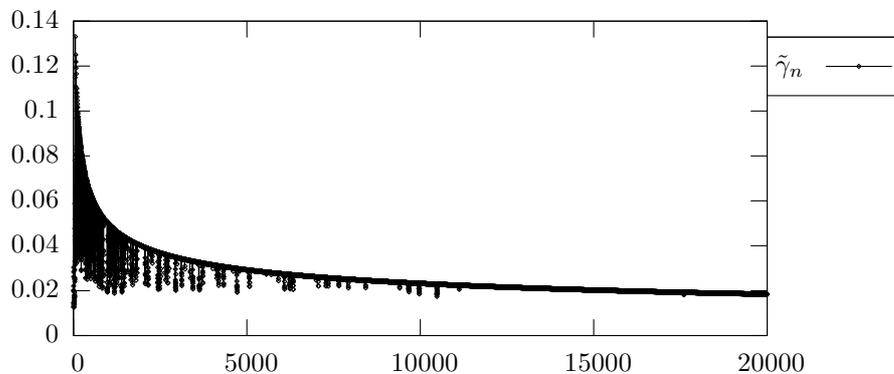
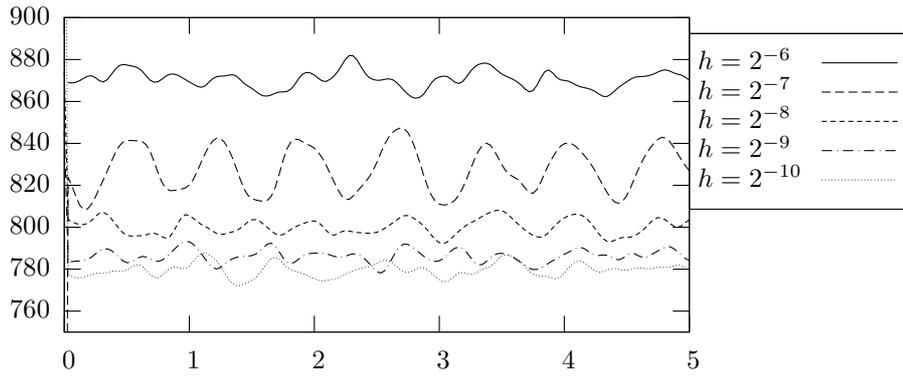


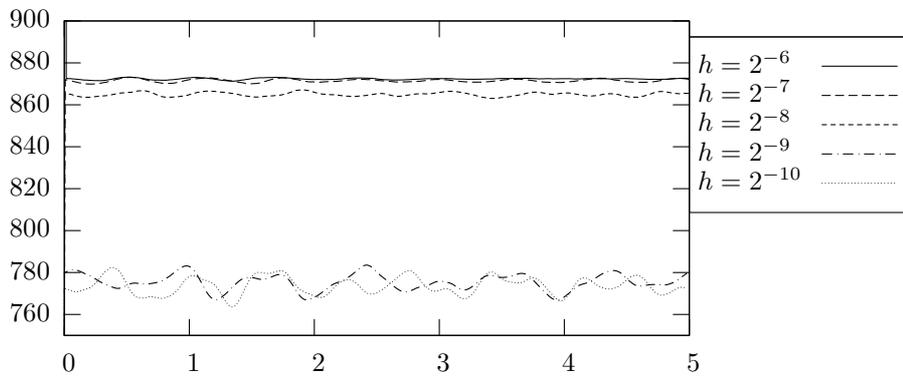
Figure VII.2 – Stochastic step sequence  $(\tilde{\gamma}_n)_{n \geq 1}$ .

We compare our results with the approximation of  $\mathbf{E}[f(\bar{X}_h^T)]$  where  $(\bar{X}_h^T)$  is a regular Euler scheme (figure VII.3(a)) or a implicit Euler scheme (figure VII.3(b)). We represent the results only for  $h \leq 2^{-6}$  because for bigger values the empirical expectation (based on 10 000 paths) explodes with the regular Euler scheme. For the implicit scheme the expectation remains bounded but the behaviour is very poor when  $h \leq 2^{-8}$ . The parameters  $h$ ,  $T$  and the number of paths used for the Monte-Carlo procedure are hard to fix.

For the implicit Euler scheme (figure VII.3(b)), the jump between  $h = 2^{-8}$  and  $h = 2^{-9}$  is due to the fact that the scheme remains trapped in the neighborhood of only one attractor (Lorenz equation has two attractors). This behavior does not occur with the regular Euler scheme and with our scheme.



(a) Euler scheme :  $\bar{X}_{k+1,h}^T = \bar{X}_{k,h}^T + hb(\bar{X}_{k,h}^T) + \sqrt{h}\sigma U_{k+1}$



(b) Implicit Euler scheme :  $\bar{X}_{k+1,h}^T = \bar{X}_{k,h}^T + hb(\bar{X}_{k+1,h}^T) + \sqrt{h}\sigma U_{k+1}$

Figure VII.3 – Computation of  $\mathbf{E}[f(\bar{X}_h^T)]$  by a Monte-Carlo procedure on 10000 paths and for  $T = 5$

## 7.2. Perturbed Hamiltonian system

The second example is a perturbed Hamiltonian system derived from a multidimensional linear oscillator under external random excitation. It is a 3-DOF (degree of freedom) system studied by Ibrahim and Li (see [IL89]) and Soize (see [Soi94] chap. XIII.6). We have thus the following equation in  $\mathbf{R}^6$

$$\begin{cases} dq_t = \partial_p H(q_t, p_t) dt, \\ dy_t = -\partial_q H(q_t, p_t) dt - f_0 D_0 \partial_p H(q_t, p_t) dt + g_0 S_0 dW_t \end{cases}$$

where  $H(q_t, p_t) = \frac{1}{2} \langle M(q)^{-1} p, p \rangle + \frac{1}{2} \langle K_0 q, q \rangle$  and

$$D_0 = S_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad K_0 = \begin{pmatrix} 1.3 & 0 & 0 \\ 0 & 0.154 & 0 \\ 0 & 0 & 0.196 \end{pmatrix}, \quad g_0 = 0.5, \quad f_0 = 0.9965,$$

and

$$M(q) = \begin{pmatrix} 1.3 & v_2 + 0.3 & v_3 \\ v_2 + 0.3 & v_2^2 + 2v_2 + 0.314 & v_2v_3 + v_3 + 0.0375 \\ v_3 & v_2v_3 + v_3 + 0.0375 & v_3^2 + 0.1 \end{pmatrix},$$

with  $v_2(q) = -1.61(0.375q_2 + q_3)$  and  $v_3(q) = -1.61q_2$ . This numerical example is taken from [Soi94] (page 257–264). This is the first damping model case with the external excitation applied to DOF 1 and the system parameter equal to 0.7. In this case we have an analytic expression of the density of the invariant measure and we can calculate the mean-square response for the DOF 1. We consider the function  $f_1 : (q, p) \mapsto q_1^2$  and we have

$$\int_{\mathbf{R}^6} f_1(x) \nu(dx) \simeq 0.0965.$$

The stochastic sequence used in the following simulations is defined  $\forall n \geq 1$  by

$$\tilde{\gamma}_n = (\gamma_0 n^{-1/3}) \wedge \left( \frac{1}{|b(X_{n-1})|^2 \vee 1} \right).$$

The results for different value of  $\gamma_0$  are given in figure VII.4. The convergence seems better when  $\gamma_0$  is big. Our scheme behaves very well and a representation of the stochastic step sequence  $(\tilde{\gamma}_n)_{n \geq 1}$  is given in figure VII.5.

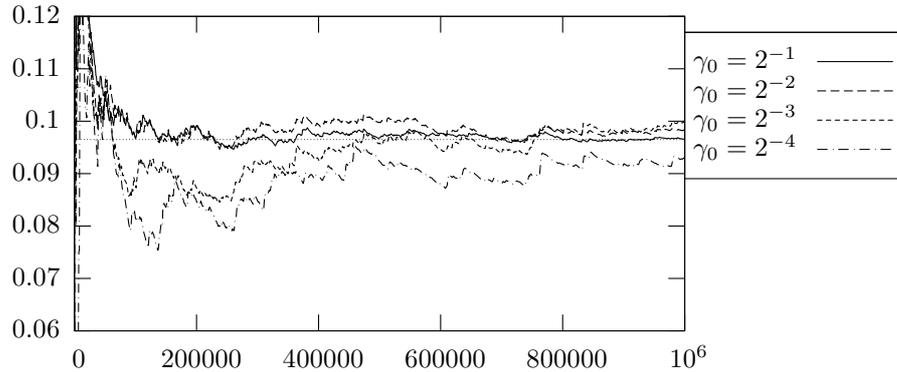


Figure VII.4 – One path of  $(\nu_n^\eta(f))_{1 \leq n \leq 10^6}$  for different value of  $\gamma_0$ .

We do not represent the approximation of  $\mathbf{E}[f(\bar{X}_h^T)]$  where  $(\bar{X}_h^T)$  is the explicit Euler scheme because the empirical expectation (based on 10 000 paths) explodes for  $h < 2^{-3}$ .

Figure VII.6 gives results using the approximation of  $\mathbf{E}[f(\bar{X}_h^T)]$  where  $(\bar{X}_h^T)$  is the implicit Euler scheme.

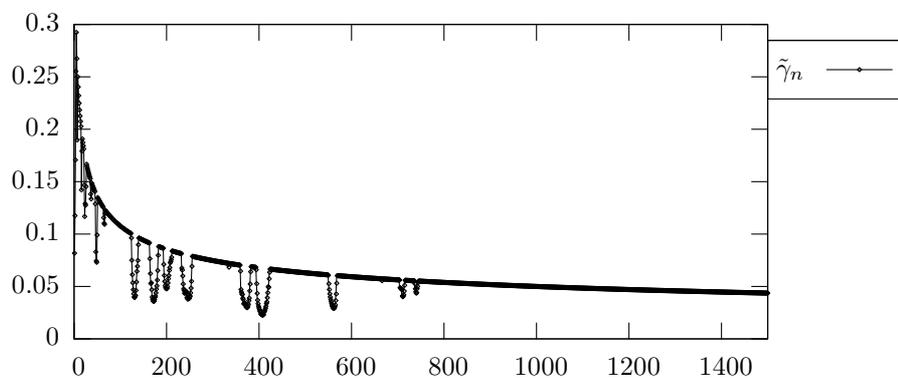


Figure VII.5 – One representation of the stochastic step sequence  $(\tilde{\gamma}_n)_{n \geq 1}$ .

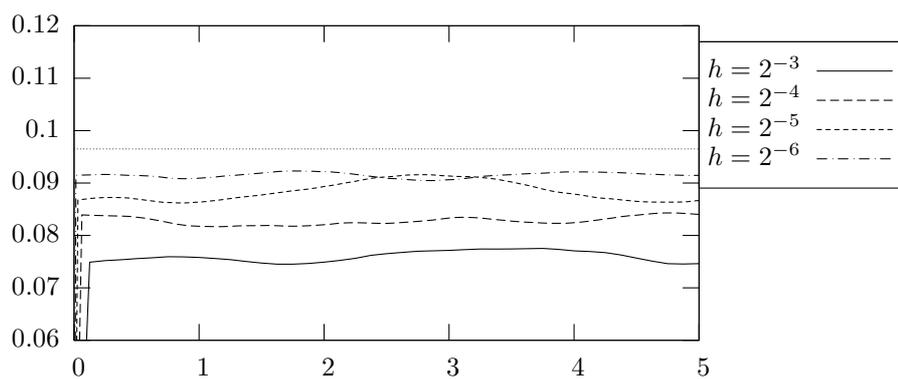


Figure VII.6 – Computation of  $\mathbf{E}[f(\bar{X}_h^T)]$  by a Monte-Carlo procedure on 10 000 paths and for  $T = 5$  where  $\bar{X}_h^T$  is a implicit Euler scheme.

---

# Étude numérique, choix des paramètres VIII

---

Ce chapitre vient en complément du chapitre V sur les vitesses de convergence de l'algorithme. On adopte un point de vue plus numérique et plus pratique afin d'aider à l'implémentation (au choix des paramètres) de l'algorithme. On rappelle que si la suite de pas  $(\gamma_n)_{n \geq 1}$  est de la forme  $\gamma_n = \gamma_0 n^{-r}$  pour  $n \geq 1$  avec  $\gamma_0 > 0$  et  $r \in ]0, 1]$ , et que la suite de poids  $(\eta_n)_{n \geq 1}$  est de la forme  $\eta_n = \eta_0 n^{-s}$  pour  $n \geq 1$  avec  $\eta_0 > 0$  et  $s \leq \frac{1+r}{2}$  alors la meilleure vitesse de convergence (en ordre) est obtenue en prenant  $r = \frac{1}{3}$  et  $s = 0$  (dans le cas où  $U_1$  est symétrique et  $q^* = 4$ <sup>1</sup>). Pour ces paramètres, on rappelle que le théorème de la limite centrale suivant est vérifié

$$\frac{2\sqrt{\gamma_0}}{\sqrt{3}} n^{\frac{1}{3}} (\nu_n^n(f) - \nu(f)) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\sqrt{3}\gamma_0^{\frac{3}{2}} m_{f,4}, s_f^2\right),$$

où

$$m_{f,4} = \int_{\mathbf{R}^d} \left( \frac{1}{2} D^2 \varphi(x) \cdot b(x)^{\otimes 2} + \frac{1}{2} D^3 \varphi(x) \cdot (b(x), \mathbf{E}[(\sigma(x)U_1)^{\otimes 2}]) \right. \\ \left. + \frac{1}{24} D^4 \varphi(x) \cdot \mathbf{E}[(\sigma(x)U_1)^{\otimes 4}] \right) \nu(dx)$$

et  $s_f^2 = \int_{\mathbf{R}^d} |\sigma^* \nabla \varphi|^2 d\nu$  avec  $\varphi$  solution de l'équation de Poisson  $\mathcal{A}\varphi = -(f - \nu(f))$ .

Lorsque  $m_{f,4} = 0$  ou qu'une extrapolation de Richardson est mise en œuvre, la vitesse optimale est en  $n^{\frac{2}{5}}$  et est obtenue pour  $r = \frac{1}{5}$  et  $s = 0$ . Tout au long du chapitre, la suite de poids considérée sera donc la suite constante égale à 1. De plus, pour davantage de clarté on note la dépendance en  $\gamma_0$  et en  $r$  pour le schéma de discrétisation et les mesures empiriques :

$$\forall n \geq 1, \quad \nu_n^{\gamma_0, r} = \frac{1}{n} \sum_{k=1}^n \delta_{X_{k-1}^{\gamma_0, r}}, \quad (0.1)$$

où  $(X_n^{\gamma_0, r})_{n \geq 0}$  est soit le schéma d'Euler de pas décroissant  $\gamma_0 n^{-r}$ , soit le schéma adaptatif introduit dans le chapitre VII construit à partir de la suite  $\tilde{\gamma}_n = \chi_{n-1} \wedge \gamma_0 n^{-r}$ .

Quand  $r$  et  $s$  sont fixés, on est réduit au choix d'un unique paramètre :  $\gamma_0$  qui fixe la taille des pas de discrétisation (la vitesse de décroissance est fixée par  $r$ ). D'après l'étude faite dans

---

<sup>1</sup>cf. théorème V.3 pour la définition de  $q^*$ . Le cas  $q^* = 4$  correspond au cas standard *i.e.* lorsque la solution  $\varphi$  de l'équation de Poisson n'annule pas  $D_4$ .

le chapitre V, on sait qu'il faut le choisir la façon suivante (si  $q^* = 4$ )

$$\gamma_0 = \left( \sqrt{3} \frac{|m_{f,4}|}{s_f} \right)^{-2/3}, \quad (0.2)$$

afin de minimiser l'erreur  $\nu_n^{\gamma_0}(f) - \nu(f)$  dans  $\mathbf{L}^2$  lorsque  $f$  et  $\varphi$  sont à support compact. Cependant, le biais  $m_{f,4}$  et la variance  $s_f^2$  sont inconnus et difficilement calculables (cf. section 3

Dans un premier temps, on fait une étude qualitative et quantitative de l'influence du paramètre  $\gamma_0$ . On en profite pour tester numériquement l'efficacité de notre algorithme lorsqu'il est couplé à l'extrapolation de Richardson (cf. section 6 du chapitre V). Ensuite, nous répondons brièvement à deux questions naturelles concernant le bruit blanc  $(U_n)_{n \geq 1}$  utilisé dans les simulations : Doit-on utiliser des gaussiennes ou des Bernoulli ? Dans le cas d'une extrapolation de Richardson, doit-on utiliser le même bruit pour la trajectoire changée de temps ou un bruit indépendant ? Puis nous utilisons des estimateurs consistants pour le biais et la variance, basés sur l'extrapolation de Richardson et sur le TLC presque sûr pour les martingales, afin d'établir une procédure permettant de choisir efficacement  $\gamma_0$ . On poursuit l'étude numérique en proposant une technique de réduction de variance utilisant les fonctions  $c\mathcal{A}g$  comme variables de contrôles (elles sont centrées sous  $\nu$ ).

## 1. Influence du paramètre $\gamma_0$ , extrapolation de Richardson

Le paramètre  $\gamma_0$  fixe la taille des pas de discrétisation et intervient dans le TLC biaisé optimal de deux façons : lorsque  $\gamma_0$  est proche de 0, le biais est petit mais la variance grande et lorsque  $\gamma_0$  croît la variance décroît mais le biais augmente. Ce comportement se révèle aussi très bien si on ne regarde qu'une trajectoire  $\nu_n^{\gamma_0}(f)$ . La figure VIII.1 illustre ce phénomène

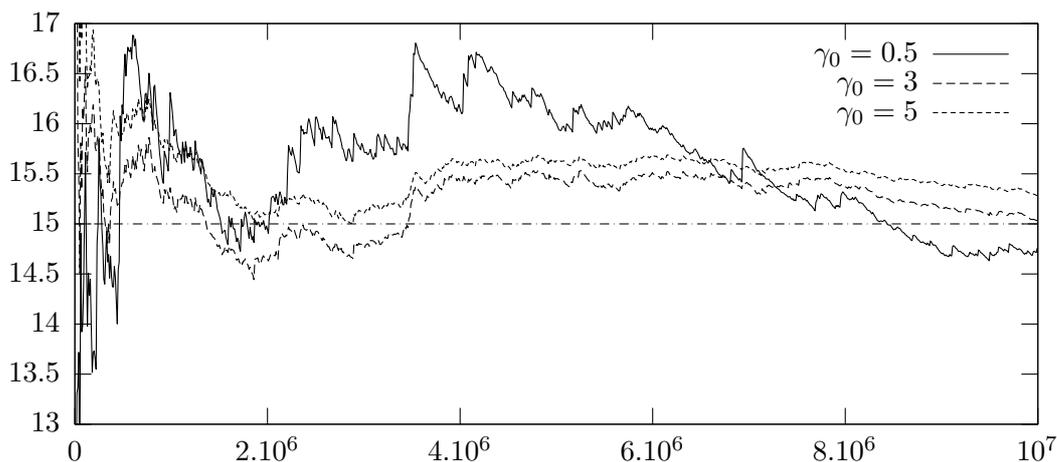


Figure VIII.1 – Comparaison de trois trajectoires de  $\nu_n^{\gamma_0}(x^6)$  construites à partir du même bruit blanc  $(U_n)_{n \geq 1}$  pour différentes valeurs de  $\gamma_0$  ( $n$  variant de 1 à  $10^7$ ,  $r = \frac{1}{3}$ )

(la diffusion considérée est un Ornstein-Uhlenbeck standard et la fonction  $f(x) = x^6$ , donc  $\nu(f) = 15$ ). Lorsque  $\gamma_0$  est petit (cas  $\gamma_0 = 0.5$ ) la trajectoire fluctue beaucoup : la variation

quadratique est grande; et lorsque  $\gamma_0$  est grand ( $\gamma_0 = 5$ ) la trajectoire est plus stable mais semble converger avec un biais plus grand que dans le cas  $\gamma_0 = 3$ . Le même bruit blanc  $(U_n)_{n \geq 1}$  a bien sûr été utilisé pour les trois simulations. Il est important de noter que le comportement d'une trajectoire  $\nu_n^{\gamma_0}$  donne des informations qualitatives sur le type d'erreur (biais élevé ou variance élevée) que l'on fait.

Comparons de la même façon, notre algorithme sans et avec extrapolation de Richardson de paramètre  $\lambda > 1$ . Il s'agit donc de comparer une trajectoire de  $\nu_n^{\gamma_0, r}(f)$  à une trajectoire de  $\tilde{\nu}_n^{\gamma_0, r, \lambda}(f)$  définie pour tout  $n \geq 1$  par

$$\tilde{\nu}_n^{\gamma_0, r, \lambda}(f) = \frac{1}{\lambda - 1} \left( \lambda \nu_n^{\frac{\gamma_0}{\lambda}, r}(f) - \nu_n^{\gamma_0, r}(f) \right).$$

On compare ces deux algorithmes pris dans leur situation optimale donc la vitesse de décroissance du pas est fixée à  $r = \frac{1}{3}$  sans extrapolation de Richardson, et à  $r = \frac{1}{5}$  avec extrapolation de Richardson. De plus le même bruit blanc a été utilisé pour les trois schémas  $X_n^{\gamma_0, \frac{1}{3}}$ ,  $X_n^{\gamma_0, \frac{1}{5}}$  et  $X_n^{\frac{\gamma_0}{\lambda}, \frac{1}{5}}$ .

Le résultat lorsque la diffusion est un Ornstein-Uhlenbeck standard (unidimensionnel) et que la fonction  $f$  est  $f(x) = x^6$ , est donné par la figure VIII.2. La trajectoire construite avec l'extrapolation de Richardson semble plus stable et converger plus rapidement.

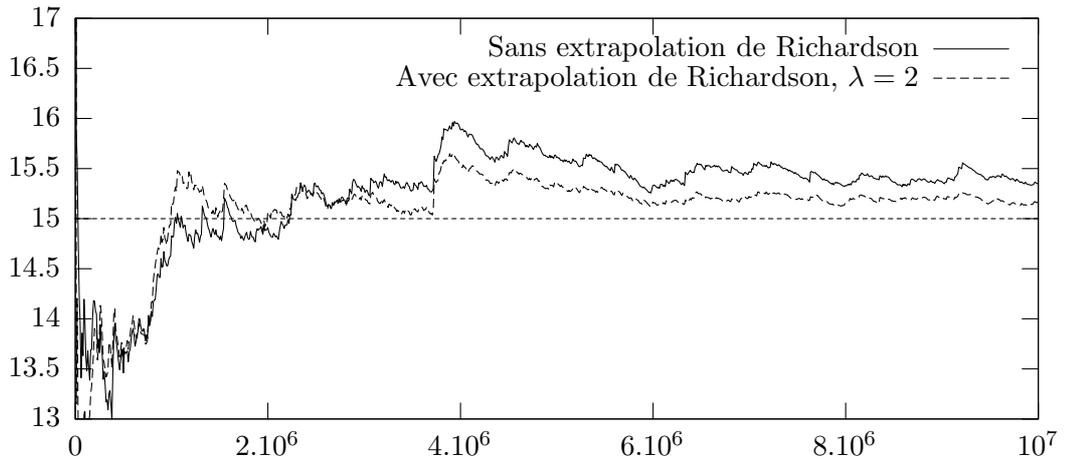


Figure VIII.2 – Comparaison d'une trajectoire de  $\nu_n^{\gamma_0, \frac{1}{3}}(x^6)$  avec une trajectoire de  $\tilde{\nu}_n^{\gamma_0, \frac{1}{5}, 2}(x^6)$  ( $n$  variant de 1 à  $10^7$  et  $\gamma_0 = 2$ )

**Remarque:** Le paramètre  $\lambda$  peut être choisit arbitrairement car il joue un rôle similaire à  $\gamma_0$ . Chercher  $\lambda$  optimal revient à chercher le paramètre  $\gamma_0$  optimal lorsque  $q^* = 6$ , que le biais est  $\tilde{m}_{f, q^*, \lambda}$  et que la variance est  $\tilde{s}_{f, \lambda}^2$ . Voir le théorème V.6 et les remarques qui suivent pour plus de détails. \*

Procédons maintenant à une étude quantitative qui permettra de statuer sur l'efficacité de l'extrapolation de Richardson et sur le choix d'une valeur de  $\gamma_0$ .

Dans tous les exemples qui suivent, on cherche le moment d'ordre 2 de la mesure invariante, *i.e.* on calcule  $\int_{\mathbf{R}^d} |x|^2 \nu(dx)$ . On compare l'algorithme sans extrapolation de Richard-

son (Standard), à l'algorithme couplé à une extrapolation (R1) de paramètre  $\lambda = 2$  et à l'algorithme couplé à deux extrapolations (R2) de paramètre  $\lambda = 2$ . Tous les algorithmes sont comparés avec leur paramètre  $r$  optimal, donc il s'agit d'étudier  $\nu_n^{\gamma_0, \frac{1}{3}}(f)$ ,  $\tilde{\nu}_n^{\gamma_0, \frac{1}{5}, 2}(f)$  et  $\bar{\nu}_n^{\gamma_0, \frac{1}{7}, 2}(f)$  où

$$\forall n \geq 1, \quad \nu_n^{\gamma_0, \frac{1}{3}}(f) = \frac{1}{n} \sum_{k=1}^n f(X_{k-1}^{\gamma_0, \frac{1}{3}}), \quad (\text{Standard})$$

$$\forall n \geq 1, \quad \tilde{\nu}_n^{\gamma_0, \frac{1}{5}, 2}(f) = \frac{1}{n} \sum_{k=1}^n \left( 2f(X_{k-1}^{\frac{\gamma_0}{2}, \frac{1}{5}}) - f(X_{k-1}^{\gamma_0, \frac{1}{5}}) \right), \quad (\text{R1})$$

$$\forall n \geq 1, \quad \bar{\nu}_n^{\gamma_0, \frac{1}{7}, 2}(f) = \frac{1}{n} \sum_{k=1}^n \frac{1}{3} \left( 4f(X_{k-1}^{\frac{\gamma_0}{4}, \frac{1}{7}}) - 2f(X_{k-1}^{\frac{\gamma_0}{2}, \frac{1}{7}}) + f(X_{k-1}^{\gamma_0, \frac{1}{7}}) \right). \quad (\text{R2})$$

Dans tous les tests qui suivent, on considère un bouquet de 500 trajectoires de  $(U_k)_{1 \leq k \leq n}$  indépendantes, et on calcule pour chaque trajectoire une approximation de  $\nu(|\cdot|^2)$  en utilisant (Standard), (R1) et (R2) jusqu'à l'itération  $n = 5.10^5$ . On calcule alors la moyenne empirique  $m_n$  et la variance empirique  $s_n^2$  sur ces 500 trajectoires. Puis on donne une estimation de l'erreur dans  $\mathbf{L}^2$

$$e_n = \sqrt{\left(m_n - \nu(|\cdot|^2)\right)^2 + s_n^2}$$

où  $\nu(|\cdot|^2)$  est soit la vraie valeur lorsqu'une expression analytique de la densité est connue, soit une approximation donnée par l'algorithme (Standard) poussé jusqu'à l'itération  $n = 10^9$ . Ces trois algorithmes sont toujours comparés pour six valeurs de  $\gamma_0$ , et tous les schémas  $X_n^{\gamma_0, r}$  sont initialisés avec un vecteur dont toutes les composantes sont égales à 2.

### Exemple 1 : Ornstein Uhlenbeck en dimension 1

On discrétise la solution de l'EDS de drift  $b$  et de coefficient de diffusion  $\sigma$  définis par

$$b(x) = -\frac{1}{2}x, \quad \sigma(x) = 1, \quad (\text{Ex. 1})$$

en utilisant le schéma d'Euler de pas décroissant. Le résultat est donné dans le tableau VIII.1. Dans le cas (Standard), l'erreur dans  $\mathbf{L}^2$  est faible et le  $\gamma_0$  optimal semble proche de 2 ou 3. Notons que pour  $\gamma_0 = 1$  ou  $\gamma_0 = 4$ , l'erreur est plus importante d'environ 25%. Comme attendu, on remarque aussi que le biais augmente avec  $\gamma_0$  et que la variance diminue. Dans le cas (R1), le  $\gamma_0$  optimal semble situé entre 3 et 4 et le gain obtenu sur l'erreur est d'environ 36% par rapport à (Standard). Dans le cas (R2), le  $\gamma_0$  optimal est proche de 1 et l'erreur est plus grande que (Standard). La vitesse est de l'ordre de  $n^{\frac{3}{7}}$  mais la variance est plus importante, l'effet sur l'erreur  $\mathbf{L}^2$  apparaît donc plus tard. Pour un nombre d'itérations important, il peut être envisageable d'utiliser deux extrapolations.

D'après cet exemple, il est clair que le choix de  $\gamma_0$  est important pour l'efficacité de l'algorithme et qu'une extrapolation de Richardson nécessite de choisir un nouveau paramètre  $\gamma_0$  en fonction du nouveau biais et de la nouvelle variance.

1. INFLUENCE DU PARAMÈTRE  $\gamma_0$ , EXTRAPOLATION DE RICHARDSON

$\gamma_0$		0.5	1	2	3	4	5
(Standard)	Erreur $\mathbf{L}^2$	0.0313	0.0217	<b>0.0172</b>	0.0184	0.0216	0.0259
	Moyenne	1.0004	1.0030	1.0084	1.0137	1.0188	1.0240
	Variance	0.0010	0.0005	0.0002	0.0002	0.0001	0.0001
(R1)	Erreur $\mathbf{L}^2$	0.0257	0.0178	0.0128	0.0110	<b>0.0108</b>	0.0123
	Moyenne	0.9978	0.9985	0.9982	0.9968	0.9945	0.9910
	Variance	0.0007	0.0003	0.0002	0.0001	0.0001	0.0001
(R2)	Erreur $\mathbf{L}^2$	0.0224	<b>0.0180</b>	0.0237	0.0346	0.0476	0.0627
	Moyenne	1.0033	1.0093	1.0210	1.0333	1.0469	1.0623
	Variance	0.0005	0.0002	0.0001	0.0001	0.0001	0.0001

Tableau VIII.1 – Comparaison des trois algorithmes  $\nu_n^{\gamma_0, \frac{1}{3}}(|\cdot|^2)$ ,  $\tilde{\nu}_n^{\gamma_0, \frac{1}{5}, 2}(|\cdot|^2)$  et  $\bar{\nu}_b^{\gamma_0, \frac{1}{7}, 2}(|\cdot|^2)$  lorsque  $b$  et  $\sigma$  vérifient (Ex. 1) pour différentes valeurs de  $\gamma_0$ .

Au niveau du temps de calcul, en 1 seconde on effectue environ  $10^6$  itérations avec l’algorithme (Standard), 660 000 itérations avec l’algorithme (R1) et 525 000 itérations avec l’algorithme (R2).

Exemple 2 : Modèle école en dimension 2

On considère maintenant une EDS fortement bruitée de drift  $b$  et de coefficient de diffusion  $\sigma$  définis par :

$$b(x, y) = \begin{pmatrix} -x + \arctan(x) \\ -2y \end{pmatrix}, \quad \sigma(x, y) = \frac{1}{2} \begin{pmatrix} \sqrt{1+y^2} & 1 \\ 1 & \sqrt{1+x^2} \end{pmatrix}. \quad (\text{Ex. 2})$$

Les résultats obtenus sont présentés dans le tableau (VIII.2).

$\gamma_0$		0.4	0.6	0.8	1	1.2	1.4
(Standard)	Erreur $\mathbf{L}^2$	0.0347	0.0285	0.0247	0.0223	0.0207	<b>0.0197</b>
	Moyenne	1.0709	1.0719	1.0730	1.0741	1.0753	1.0766
	Variance	0.0012	0.0008	0.0006	0.0005	0.0004	0.0003
(R1)	Erreur $\mathbf{L}^2$	0.0290	0.0234	0.0203	0.0182	0.0168	<b>0.0162</b>
	Moyenne	1.0681	1.0673	1.0667	1.0660	1.0652	1.0631
	Variance	0.0008	0.0005	0.0004	0.0003	0.0003	0.0002
(R2)	Erreur $\mathbf{L}^2$	0.0256	0.0216	<b>0.0203</b>	0.0206	0.0222	0.0308
	Moyenne	1.0737	1.0761	1.0790	1.0823	1.0859	1.0951
	Variance	0.0006	0.0004	0.0003	0.0002	0.0002	0.0003

Tableau VIII.2 – Comparaison des trois algorithmes  $\nu_n^{\gamma_0, \frac{1}{3}}(|\cdot|^2)$ ,  $\tilde{\nu}_n^{\gamma_0, \frac{1}{5}, 2}(|\cdot|^2)$  et  $\bar{\nu}_b^{\gamma_0, \frac{1}{7}, 2}(|\cdot|^2)$  lorsque  $b$  et  $\sigma$  vérifient (Ex. 2) pour différentes valeurs de  $\gamma_0$ .

Comme dans l’exemple 1, l’algorithme (R1) a la plus petite erreur pour le paramètre  $\gamma_0 = 1.4$ . Le gain est d’environ 17%. Le paramètre optimal pour (R2) est situé entre 0.6 et 0.8 mais l’erreur est légèrement supérieure à (Standard).

De plus, en 1 seconde on calcul 550 000 itérations avec (Standard), 360 000 itérations avec (R1) et 260 000 itérations avec (R2). L'algorithme (R2) demande donc deux fois plus de temps de calcul que l'algorithme (Standard).

Exemple 3 : Oscillateur linéaire perturbé en dimension 4

On considère maintenant un oscillateur vérifiant une EDS de coefficients :

$$b(q, p) = \begin{pmatrix} p_1 - \frac{1}{2}p_2 \\ -\frac{1}{2}p_1 + \frac{5}{4}p_2 \\ -p_1 + \frac{1}{2}p_2 - q_1 \\ -q_2 \end{pmatrix}, \quad \sigma(q, p) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0.2|q_2| \end{pmatrix}. \quad (\text{Ex. 3})$$

Les résultats sont donnés dans le tableau VIII.3. L'erreur  $\mathbf{L}^2$  est nettement plus grande que dans les exemples précédents. En particulier la variance est plus grande. L'algorithme (R1) est logiquement moins performant que dans les exemples précédents car la nouvelle variance est encore plus grande, et l'effet de la vitesse en  $n^{\frac{2}{5}}$  n'est pas perceptible à l'itération 500 000. Le paramètre  $\gamma_0$  semble aussi important à déterminer. En effet, en passant de  $\gamma_0 = 0.4$  à

		$\gamma_0$	0.05	0.1	0.2	0.3	0.4	0.6
(Standard)	Erreur $\mathbf{L}^2$		0.2535	0.2003	0.1440	0.1279	<b>0.1267</b>	0.1533
	Moyenne		2.3376	2.3187	2.3334	2.3535	2.3698	2.4102
	Variance		0.0610	0.0386	0.0179	0.0109	0.0080	0.0066
(R1)	Erreur $\mathbf{L}^2$		0.3103	0.2289	0.1656	0.1417	<b>0.1262</b>	0.1543
	Moyenne		2.3043	2.3110	2.2903	2.2687	2.2404	2.1681
	Variance		0.0957	0.0514	0.0273	0.0200	0.0144	0.0113
(R2)	Erreur $\mathbf{L}^2$		0.2833	0.2150	<b>0.1876</b>	0.2030	0.2469	0.4319
	Moyenne		2.3387	2.3569	2.3941	2.4415	2.5013	2.7002
	Variance		0.0768	0.0403	0.0222	0.0151	0.0120	0.0099

Tableau VIII.3 – Comparaison des trois algorithmes  $\nu_n^{\gamma_0, \frac{1}{3}}(|\cdot|^2)$ ,  $\tilde{\nu}_n^{\gamma_0, \frac{1}{5}, 2}(|\cdot|^2)$  et  $\bar{\nu}_b^{\gamma_0, \frac{1}{7}, 2}(|\cdot|^2)$  lorsque  $b$  et  $\sigma$  vérifient (Ex. 3) pour différentes valeurs de  $\gamma_0$ .

$\gamma_0 = 0.6$  l'erreur augmente de 21%, et elle augmente de 58% lorsque  $\gamma_0$  varie de 0.4 à 0.1 (pour l'algorithme (Standard)). Pour obtenir une meilleure précision, il est nécessaire d'augmenter le nombre d'itérations. L'algorithme (R1) semble être le meilleur à considérer bien qu'il nécessite un temps de calcul 50% plus long (environ).

En 1 seconde, on calcule 625 000 itérations avec l'algorithme (Standard), 415 000 avec l'algorithme (R1) et 312 000 avec l'algorithme (R2).

Exemple 4 : Équation de Langevin

On considère l'équation de Langevin en dimension 2 dont  $b$  et  $\sigma$  vérifient

$$b(q, p) = \begin{pmatrix} p \\ -(\delta(q)p + F'(q)) \end{pmatrix}, \quad \sigma(q, p) = \begin{pmatrix} 0 \\ \sigma(q) \end{pmatrix}. \quad (\text{Ex. 4})$$

## 1. INFLUENCE DU PARAMÈTRE $\gamma_0$ , EXTRAPOLATION DE RICHARDSON

avec  $F(q) = \frac{1}{4}(1 - q^2)^2$ ,  $\sigma(q) = \frac{4(5q^2 + 1)}{5(q^2 + 1)}$  et  $\delta = \frac{\sigma^2(q)}{2}$ . Cette équation est discrétisée en utilisant notre schéma adaptatif défini dans le chapitre VII. La suite de pas stochastique considérée est définie pour tout  $n \geq 1$  par  $\tilde{\gamma}_n = \gamma_n \wedge \chi(X_{n-1}^{\gamma_0, r})$ , où  $\chi(x) = \frac{|x|^2 + 1}{|b(x)|^2 \sqrt{1}}$ . Les résultats sont donnés dans le tableau VIII.4.

	$\gamma_0$	0.2	0.3	0.4	0.5	0.6	0.7
(Standard)	Erreur $\mathbf{L}^2$	0.0534	0.0455	0.0426	<b>0.0406</b>	0.0415	0.0447
	Moyenne	2.0511	2.0559	2.0610	2.0648	2.0698	2.0759
	Variance	0.0027	0.0018	0.0014	0.0010	0.0008	0.0007
(R1)	Erreur $\mathbf{L}^2$	0.0532	0.0403	0.0365	0.0322	0.0302	<b>0.0300</b>
	Moyenne	2.0373	2.0378	2.0363	2.0352	2.0338	2.0314
	Variance	0.0028	0.0016	0.0013	0.0010	0.0009	0.0008
(R2)	Erreur $\mathbf{L}^2$	0.0515	<b>0.0496</b>	0.0565	0.0679	0.0832	0.1001
	Moyenne	2.0611	2.0741	2.0864	2.1018	2.1190	2.1369
	Variance	0.0022	0.0013	0.0010	0.0008	0.0007	0.0006

Tableau VIII.4 – Comparaison des trois algorithmes  $\nu_n^{\gamma_0, \frac{1}{3}}(|\cdot|^2)$ ,  $\tilde{\nu}_n^{\gamma_0, \frac{1}{5}, 2}(|\cdot|^2)$  et  $\tilde{\nu}_b^{\gamma_0, \frac{1}{7}, 2}(|\cdot|^2)$  lorsque  $b$  et  $\sigma$  vérifient (Ex. 4) pour différentes valeurs de  $\gamma_0$ .

L'extrapolation de Richardson semble aussi très bien fonctionner lorsque la diffusion est discrétisée avec le schéma adaptatif. De plus, notons qu'avec un pas de grande taille ( $\gamma_0 \geq 0.5$ ) presque toute les trajectoires discrétisées par le schéma déterministe classique explose. Or, ce sont pour ces grands pas que l'erreur  $\mathbf{L}^2$  est la plus petite.

Les temps de calculs sont les suivants : en 1 seconde, on effectue environ 525 000 itérations avec l'algorithme (Standard), 320 000 avec l'algorithme (R1) et 230 000 avec l'algorithme (R2).

### Exemple 5 : Équation de Lorenz (dimension 3)

On considère maintenant l'équation de Lorenz perturbée. Les fonctions  $b$  et  $\sigma$  sont définies par

$$b(x, y, z) = \begin{pmatrix} 10(y - x) \\ 28x - y - xz \\ xy - \frac{8}{3}z \end{pmatrix}, \quad \sigma(x, y, z) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad (\text{Ex. 5})$$

et le schéma de discrétisation utilisé est le schéma adaptatif. La suite  $(\tilde{\gamma}_n)_{n \geq 0}$  est définie par  $\tilde{\gamma}_n = \gamma_n \wedge \chi(X_{n-1}^{\gamma_0, r})$ , où  $\chi(u) = \frac{|u|^2 + 1}{|b(u)|^2 \sqrt{1}}$ . Les résultats sont donnés dans le tableau VIII.5. C'est un exemple numérique très compliqué, et le choix du paramètre  $\gamma_0$  semble primordial pour espérer obtenir de bons résultats. Sans extrapolation de Richardson, l'erreur minimale est de 7.22 pour  $\gamma_0 = 0.05$  et avec extrapolation de Richardson, l'erreur minimale est de 1.88 pour  $\gamma_0 = 0.1$ . Ce gain de 384% est dû en particulier au biais qui semble bien plus petit. Il faut relativiser ce résultat, car le nombre de trajectoires (500) n'est peut-être pas suffisant, mais surtout la vraie valeur n'est pas connue et a été calculée par notre algorithme l'algorithme (Standard) jusqu'à l'itération  $10^{10}$  en prenant  $\gamma_0 = 0.05$ .

En ce qui concerne les temps de calculs, en 1 seconde on effectue environ 500 000 itérations avec l'algorithme (Standard), 312 500 itérations avec (R1), et 220 000 itérations avec (R2).

		$\gamma_0$	0.001	0.005	0.01	0.05	0.1	0.2
(Standard)	Erreur $\mathbf{L}^2$		68.56	13.91	7.73	<b>7.22</b>	12.88	25.43
	Moyenne		836.41	784.80	780.38	781.94	787.84	800.42
	Variance		929.81	97.50	30.92	3.96	1.23	0.45
(R1)	Erreur $\mathbf{L}^2$		58.60	12.09	7.72	2.03	<b>1.88</b>	3.19
	Moyenne		817.86	781.79	778.97	774.80	776.48	778.10
	Variance		1597.10	100.02	43.90	4.08	1.34	0.53
(R2)	Erreur $\mathbf{L}^2$		56.36	13.37	<b>9.13</b>	12.87	25.83	81.89
	Moyenne		816.38	783.71	781.87	787.75	800.81	856.88
	Variance		1464.53	103.01	36.24	3.03	1.05	0.44

Tableau VIII.5 – Comparaison des trois algorithmes  $\nu_n^{\gamma_0, \frac{1}{3}}(|\cdot|^2)$ ,  $\tilde{\nu}_n^{\gamma_0, \frac{1}{5}, 2}(|\cdot|^2)$  et  $\bar{\nu}_b^{\gamma_0, \frac{1}{7}, 2}(|\cdot|^2)$  lorsque  $b$  et  $\sigma$  vérifient (Ex. 5) pour différentes valeurs de  $\gamma_0$ .

## 2. À propos du bruit blanc $U_n$

On se propose de répondre numériquement à deux questions naturelles pour l'implémentation de l'algorithme. Quel bruit blanc considérer ? Dans le cas d'une extrapolation de Richardson, doit-on utiliser le même bruit blanc pour les deux trajectoires  $X_n^{\gamma_0, r}$  et  $X_n^{\frac{20}{\lambda}, r}$  ?

### 2.1. Gaussienne ou Bernoulli ?

Dans le développement de l'erreur obtenu dans la proposition V.3, on remarque que le terme

$$\frac{1}{2H_n} \sum_{k=1}^n \eta_k \left( D^2\varphi(X_{k-1}) \cdot (\sigma(X_{k-1})U_k)^{\otimes 2} - \text{Tr}(D^2\varphi(X_{k-1})(\sigma^*\sigma)(X_{k-1})) \right),$$

est nul lorsque  $U_n$  est un vecteur aléatoire dont chaque composante est de loi de Bernoulli prenant la valeur 1 avec probabilité 1/2 et la valeur  $-1$  sinon. Plus généralement (par exemple si  $U_n \sim \mathcal{N}(0, \text{Id}_m)$ ), on a montré que ce terme converge dans  $\mathbf{L}^2$  et *p.s.* vers 0.

On compare l'algorithme (Standard) lorsque  $U_1$  est une gaussienne centrée réduite et lorsque  $U_1$  est un vecteur dont chaque composante est de loi de Bernoulli de paramètre  $\frac{1}{2}$  à valeurs dans  $\{-1, 1\}$ . Le résultat, lorsque la diffusion est un Ornstein-Uhlenbeck et  $f(x) = x^2$ , est donné par le tableau VIII.6. On note aucune différence significative entre les deux schémas.

		$\gamma_0$	0.5	1	2	3	4	5
gaussienne	Erreur $\mathbf{L}^2$		0.0313	0.0217	<b>0.0172</b>	0.0184	0.0216	0.0259
	Moyenne		1.0004	1.0030	1.0084	1.0137	1.0188	1.0240
	Variance		0.0010	0.0005	0.0002	0.0002	0.0001	0.0001
Bernoulli	Erreur $\mathbf{L}^2$		0.0310	0.0222	<b>0.0178</b>	0.0187	0.0217	0.0258
	Moyenne		0.9998	1.0029	1.0084	1.0136	1.0187	1.0239
	Variance		0.0010	0.0005	0.0002	0.0002	0.0001	0.0001

Tableau VIII.6 – Comparaison entre l'algorithme (Standard) construit à partir d'un bruit blanc Gaussien et à partir d'un bruit blanc de type Bernoulli.

Une Bernoulli étant toujours plus rapide à simuler qu'une gaussienne, on recommande donc l'usage de Bernoulli comme bruit blanc.

## 2.2. Extrapolation de Richardson : même suite $(U_n)_{n \geq 1}$ ?

On compare maintenant l'extrapolation de Richardson lorsque les deux trajectoires utilisées  $X_n^{\gamma_0, r}$  et  $X_n^{\frac{\gamma_0}{\lambda}, r}$  sont construites à partir d'une même trajectoire  $(U_n)_{n \geq 1}$  et lorsqu'elles sont construites à partir de deux trajectoires  $(U_n)_{n \geq 1}$  et  $(\tilde{U}_n)_{n \geq 1}$  indépendantes. Le résultat est donné dans le tableau VIII.7. Il semble que le choix d'un bruit indépendant pour la trajectoire

		$\gamma_0$	0.5	1	2	3	4	5
$\tilde{U}_n = U_n$	Erreur $\mathbf{L}^2$		0.0257	0.0178	0.0128	0.0110	<b>0.0108</b>	0.0123
	Moyenne		0.9978	0.9985	0.9982	0.9968	0.9945	0.9910
	Variance		0.0007	0.0003	0.0002	0.0001	0.0001	0.0001
$\tilde{U}_n \perp U_n$	Erreur $\mathbf{L}^2$		0.0396	0.0279	0.0199	0.0167	<b>0.0156</b>	0.0162
	Moyenne		0.9936	0.9959	0.9967	0.9957	0.9935	0.9900
	Variance		0.0015	0.0008	0.0004	0.0003	0.0002	0.0002

Tableau VIII.7 – Comparaison entre deux algorithmes (R1) : l'un dont les deux trajectoires sont corrélées, et l'autre dont les deux trajectoires sont indépendantes.

changé de temps  $X_n^{\frac{\gamma_0}{\lambda}, r}$  ne soit pas un bon choix.

Plus généralement, on a vu que la variance obtenue dans le TLC lorsqu'il y a extrapolation de paramètre  $\lambda = 2$  est de la forme (en dimension 1 pour simplifier les notations)

$$\tilde{s}_{f,2}^2 = 9s_f^2 - 4\sqrt{2} \int_{\mathbf{R}^2} (\sigma\varphi')(y_1)(\sigma\varphi')(y_2)\nu_\lambda(dy),$$

si les deux bruits sont identiques, où  $\nu_\lambda$  est l'unique mesure invariante du système

$$\begin{cases} dy_t^1 = \frac{b(y_t^1)}{\lambda} dt + \frac{\sigma(y_t^1)}{\sqrt{\lambda}} dB_t, \\ dy_t^2 = b(y_t^2) dt + \sigma(y_t^2) dB_t. \end{cases}$$

De la même façon il est aisé de montrer que si les deux bruits sont indépendants, la variance est de la forme

$$\bar{s}_{f,2}^2 = 9s_f^2 - 4\sqrt{2} \left( \int_{\mathbf{R}} (\sigma\varphi')(y)\nu(dy) \right)^2.$$

Pour statuer sur cette question, il faut donc comparer les deux quantités

$$\int_{\mathbf{R}^2} (\sigma\varphi')(y_1)(\sigma\varphi')(y_2)\nu_\lambda(dy) \quad \text{et} \quad \left( \int_{\mathbf{R}^2} (\sigma\varphi')(y)\nu(dy) \right)^2.$$

## 3. Estimation des quantités $m_{f,q^*}$ et $s_f^2$

L'estimation du biais  $m_{f,q^*}$  et de la variance  $s_f^2$  est primordiale pour choisir le paramètre  $\gamma_0$  optimal. Un estimateur fortement consistant de la variance a été introduit dans [Bar05].

Il repose sur un principe d'invariance logarithmique (cf. [Cha02] et [Maa01]). De la même façon, en utilisant les résultats similaires pour les martingales à temps discret (cf. [CM00]), et le développement de l'erreur que l'on a obtenu dans le chapitre V on montre que, si  $q^* = 4$  et  $r = \frac{1}{3}$

$$\hat{s}_n^2 = \gamma_0(1+r) \frac{1}{\log(n)} \sum_{k=1}^n \frac{1}{k} k^{1-r} (\nu_k^{\gamma_0, r}(f) - \nu(f))^2 \xrightarrow{p.s} s_f^2, \quad (\text{avec } r = \frac{1}{3}),$$

et

$$\hat{m}_n = \frac{1-r}{\gamma_0} \frac{1}{\log(n)} \sum_{k=1}^n \frac{1}{k} k^{\frac{1-r}{2}} (\nu_k^{\gamma_0, r}(f) - \nu(f)) \xrightarrow{p.s} m_{f,4} \quad (\text{avec } r = \frac{1}{3}).$$

Numériquement ces estimateurs ne sont *a priori* pas très convaincants car leur vitesse est logarithmique.

Pour calculer le biais, il est possible d'utiliser un autre estimateur obtenu par une technique similaire à l'extrapolation de Richardson. On se place dans le cas  $q^* = 4$ . On sait que pour tout  $r < \frac{1}{3}$ ,

$$\frac{1-r}{\gamma_0} n^r (\nu_n^{\gamma_0, r}(f) - \nu(f)) \xrightarrow{p.s} m_{f,4},$$

donc un estimateur fortement consistant du biais est donné, pour tout  $r < \frac{1}{3}$ , par

$$\bar{m}_n^r = \frac{2(1-r)}{\gamma_0} n^r \left( \nu_n^{\gamma_0, r}(f) - \nu_n^{\frac{\gamma_0}{2}, r}(f) \right) \xrightarrow{p.s} m_{f,4}.$$

Pour  $r = \frac{1}{5}$ , on vérifie que cet estimateur converge à la vitesse  $n^{\frac{1}{15}}$  et vérifie un TLC biaisé. Ainsi, on obtient deux estimateurs pour le paramètre  $\gamma_0$  optimal, noté  $\gamma^*$  dans la suite pour éviter la confusion avec le paramètre  $\gamma_0$  utilisé pour discrétiser la diffusion. Le premier construit en utilisant  $\hat{m}_n$  et le second en utilisant  $\bar{m}_n^r$  :

$$\hat{\gamma}_n^* = \left( \sqrt{3} \frac{\hat{m}_n}{\hat{s}_n} \right)^{-\frac{2}{3}} \quad \text{et} \quad \bar{\gamma}_n^{*,r} = \left( \sqrt{3} \frac{\bar{m}_n^r}{\hat{s}_n} \right)^{-\frac{2}{3}}$$

On teste ces estimateurs numériquement en considérant une diffusion d'Ornstein-Uhlenbeck unidimensionnelle et la fonction  $f(x) = x^2$ . Les résultats sont donnés dans le tableau VIII.8. La moyenne et la variance sont calculées sur 1000 trajectoires, et  $n = 10^6$ . Comme attendu,

Estimateurs	$\hat{s}_n^2$	$\hat{m}_n$	$\bar{m}_n^{\frac{1}{5}}$	$\hat{\gamma}_n^*$	$\bar{\gamma}_n^{*,\frac{1}{5}}$
Erreur $\mathbf{L}^2$	7.67	0.467	0.164	2.664	8.227
Moyenne	10.32	0.425	0.258	3.053	5.254
Variance	52.58	0.043	0.019	7.017	61.524

Tableau VIII.8 – Estimateurs de la variance, du biais et du  $\gamma_0$  optimal ( $\gamma^*$ ) (1000 trajectoires et  $n = 10^6$ )

les estimateurs basés sur le principe d'invariance logarithmique ne sont pas très efficaces. Au contraire l'estimateur  $\bar{m}_n^{\frac{1}{5}}$  semble plus performant, bien que sa complexité soit supérieure à celle de  $\hat{m}_n$ .

Il est très intéressant de noter que le meilleur estimateur de  $\gamma^*$  est  $\hat{\gamma}_n^*$ . Sa moyenne est relativement proche de  $\gamma^* = 2.774$ , mais la variance est assez grande. Le second estimateur est à proscrire du fait d'une variance beaucoup trop élevée. Notons que  $\hat{s}_n^2$  et  $\hat{m}_n$  ont été construits à partir de la même trajectoire  $\nu_n^{\gamma_0, r}(f)$ .

#### 4. Réduction de variance

Une idée classique pour réduire la variance lorsque la convergence de l'erreur vérifie un TLC, est d'utiliser une variable de contrôle. Une bonne variable de contrôle doit réduire la variance sans ajouter trop de biais. Dans le cadre des mesures invariantes, il existe une classe de bonnes variables de contrôle : toutes les fonctions de la forme  $c\mathcal{A}g$  où  $g$  est  $\mathcal{C}^2(\mathbf{R}^d, \mathbf{R})$ . En effet,  $\nu(c\mathcal{A}g) = 0$  et le calcul de  $\nu(f)$  peut se ramener au calcul de

$$\nu(f + c\mathcal{A}g).$$

On note  $f_{c,g} = f + c\mathcal{A}g$ . Il s'agit maintenant de trouver le bon paramètre  $c$  et la bonne fonction  $g$  pour réduire la variance. On rappelle que  $s_f^2 = \int_{\mathbf{R}^d} |\sigma^* \nabla \varphi|^2 d\nu$  où  $\varphi$  est la solution à l'équation de Poisson  $\mathcal{A}\varphi = -(f - \nu(f))$ . Pour simplifier les notations, on se place en dimension 1. Il est clair que

$$\mathcal{A}(\varphi - cg) = -(f_{c,g} - \nu(f_{c,g})),$$

donc la variance apparaissant dans le TLC lorsque l'on approche  $\nu(f_{c,g})$  est

$$\begin{aligned} s_{f_{c,g}}^2 &= \int_{\mathbf{R}} \sigma^2 (\varphi' - cg')^2 d\nu, \\ &= s_f^2 + c \left( c \int_{\mathbf{R}} \sigma^2 (g')^2 d\nu - 2 \int_{\mathbf{R}} \sigma^2 g' \varphi' d\nu \right). \end{aligned}$$

Ainsi, dès que  $c \leq \frac{2 \int_{\mathbf{R}} \sigma^2 g' \varphi' d\nu}{\int_{\mathbf{R}} \sigma^2 (g')^2 d\nu}$  on réduit la variance, et le paramètre optimal à choisir est

$$c^* = \frac{\int_{\mathbf{R}} \sigma^2 g' \varphi' d\nu}{\int_{\mathbf{R}} \sigma^2 (g')^2 d\nu}. \text{ Dans ce cas optimal, on réduit la variance de } Gain_g^* = \frac{(\int_{\mathbf{R}} \sigma^2 g' \varphi' d\nu)^2}{\int_{\mathbf{R}} \sigma^2 (g')^2 d\nu}.$$

Il reste à choisir la fonction  $g$ . Par Hölder, il est clair que

$$0 \leq Gain_g^* \leq \int_{\mathbf{R}} \sigma^2 (\varphi')^2 d\nu$$

et le cas d'égalité (dans l'inégalité de droite) est obtenu pour  $g' = \varphi'$ . Avec une telle fonction, la variance est nulle ce qui est tout à fait naturel car si on connaît la solution  $\varphi$  à l'équation de Poisson, on connaît  $\nu(f)$  :  $\nu(f) = f + \mathcal{A}\varphi$ .

Pour trouver une bonne variable de contrôle, il faut donc trouver une fonction  $g$  explicite, calculable et telle que  $g'$  soit proche de  $\varphi'$ .

Remarque: Notons qu'une variable de contrôle possible est le drift  $b$ . En effet  $b = \mathcal{A} \text{Id}$ . L'intérêt d'utiliser cette variable est qu'elle est connue et calculable. Son efficacité n'est

toutefois pas garantie lorsque  $\varphi'$  n'est pas constante. \*

Nous allons effectuer quelques tests en dimension 1, afin de valider ou non l'utilisation de ces variables de contrôles. La fonction  $g$  considérée dans toutes les simulations qui suivent vérifie

$$g'(x) = \frac{-f(x)}{b^2(x) + 1} b(x).$$

Ce choix est motivé par le fait de vouloir avoir  $\mathcal{A}g$  proche de  $-(f - \nu(f))$ . Le paramètre  $c$  est pris égal à 1, et si  $b$  est dérivable alors

$$f + \mathcal{A}g = \frac{f(x)}{b^2(x) + 1} - \frac{1}{2}\sigma^2(x) \left( \frac{f'(x)b(x) + f(x)b'(x)}{b^2(x) + 1} - \frac{2f(x)b'(x)b^2(x)}{(b^2(x) + 1)^2} \right)$$

Dans un premier temps, on considère un processus d'Ornstein-Uhlenbeck et la fonction  $f(x) = x^2$ . Les résultats sont donnés dans le tableau VIII.9. La technique semble bien fonc-

$\gamma_0$		0.5	1	2	3	4	5
$f$	Erreur $\mathbf{L}^2$	0.0313	0.0217	<b>0.0172</b>	0.0184	0.0216	0.0259
	Moyenne	1.0004	1.0030	1.0084	1.0137	1.0188	1.0240
	Variance	0.0010	0.0005	0.0002	0.0002	0.0001	0.0001
$f + \mathcal{A}g$	Erreur $\mathbf{L}^2$	0.0203	0.0142	<b>0.0112</b>	0.0118	0.0138	0.0163
	Moyenne	1.0003	1.0020	1.0055	1.0087	1.0119	1.0150
	Variance	0.0004	0.0002	0.0001	0.0001	0.0000	0.0000

Tableau VIII.9 – Comparaison de l'algorithme (Standard) sans et avec variable de contrôle ( $b(x) = -0.5x$  et  $f(x) = x^2$ ).

tionner car le gain obtenu est de 35% sur l'erreur  $\mathbf{L}^2$ . On remarque que la variance est réduite mais aussi que le biais est plus petit.

On considère maintenant la fonction  $f(x) = x^6$  et toujours le processus d'Ornstein-Uhlenbeck. Les résultats sont donnés dans le tableau VIII.10. Le gain obtenu en utilisant la variable de

$\gamma_0$		0.5	1	2	3	4	5
$f$	Erreur $\mathbf{L}^2$	1.8439	1.2715	0.9664	<b>0.9402</b>	1.0542	1.3678
	Moyenne	14.9832	15.1022	15.3349	15.5760	15.8323	16.1634
	Variance	3.3997	1.6064	0.8217	0.5522	0.4187	0.5174
$f + \mathcal{A}g$	Erreur $\mathbf{L}^2$	1.2807	0.8895	<b>0.6876</b>	0.6995	0.8065	0.9692
	Moyenne	14.9962	15.0908	15.2821	15.4744	15.6710	15.8799
	Variance	1.6403	0.7830	0.3932	0.2642	0.2002	0.1651

Tableau VIII.10 – Comparaison de l'algorithme (Standard) sans et avec variable de contrôle ( $b(x) = -0.5x$  et  $f(x) = x^6$ ).

contrôle  $\mathcal{A}g$  est de 26% et la variance comme le biais sont réduits.

On considère maintenant la diffusion suivante

$$dx_t = -x_t^3 dt + dB_t,$$

que l'on discrétise en utilisant le schéma adaptatif avec la suite de pas définie pour tout  $n \geq 1$  par  $\tilde{\gamma}_n = \gamma_n \wedge \chi(X_{n-1}^{\gamma_0, r})$  où  $\chi(x) = \frac{|x|^2+1}{|b(x)|^2\sqrt{1}}$ . La fonction  $f$  est  $f(x) = x^6$ , et les résultats sont donnés dans le tableau VIII.11. Le gain sur l'erreur  $\mathbf{L}^2$  est alors de plus de 65% ce qui semble

		$\gamma_0$	0.2	0.5	0.7	1	1.2	1.5
$f$	Erreur $\mathbf{L}^2$		0.0587	<b>0.0421</b>	0.0428	0.0509	0.0590	0.0743
	Moyenne		0.7206	0.7350	0.7444	0.7592	0.7698	0.7871
	Variance		0.0034	0.0014	0.0011	0.0008	0.0007	0.0006
$f + \mathcal{A}g$	Erreur $\mathbf{L}^2$		0.0380	0.0236	0.0200	0.0170	0.0158	<b>0.0147</b>
	Moyenne		0.7146	0.7173	0.7184	0.7199	0.7209	0.7224
	Variance		0.0014	0.0006	0.0004	0.0003	0.0002	0.0002

Tableau VIII.11 – Comparaison de l'algorithme (Standard) sans et avec variable de contrôle ( $b(x) = -x^3$  et  $f(x) = x^6$ ).

très satisfaisant, d'autant plus que le coût numérique est négligeable. Notons aussi que le biais est très nettement diminué dans cet exemple (la vraie valeur étant 0.7169).



# Bibliographie

- [Bar05] O. BARDOU : *Contrôle dynamique des erreurs de simulation et d'estimation de processus de diffusion*. Thèse de doctorat, Université de Nice-Sophia Antipolis, 2005.
- [BB92] G. K. BASAK et R. N. BHATTACHARYA : Stability in distribution for a class of singular diffusions. *Ann. Probab.*, 20(1):312–321, 1992.
- [BF94] Ph. BRESSOLETTE et M. FOGLI : Quelques méthodes de résolution de problèmes de dynamique stochastique non linéaire. *Ann. Math. Blaise Pascal*, 1(1):43–60, 1994.
- [Bha82] 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.
- [BHW97] G. K. BASAK, I. HU et C.-Z. WEI : Weak convergence of recursions. *Stochastic Process. Appl.*, 68(1):65–82, 1997.
- [Bil99] P. BILLINGSLEY : *Convergence of probability measures*. Wiley Series in Probability and Statistics : Probability and Statistics. John Wiley & Sons Inc., New York, second édition, 1999. A Wiley-Interscience Publication.
- [Bre92] L. BREIMAN : *Probability*, volume 7 de *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992. Corrected reprint of the 1968 original.
- [Cha02] F. CHAABANE : Invariance principles with logarithmic averaging for continuous local martingales. *Statist. Probab. Lett.*, 59(2):209–217, 2002.
- [CM00] F. CHAABANE et F. MAAOUIA : Théorèmes limites avec poids pour les martingales vectorielles. *ESAIM Probab. Statist.*, 4:137–189 (electronic), 2000.
- [DFMS04] R. DOUC, G. FORT, E. MOULINES et P. SOULIER : Practical drift conditions for subgeometric rates of convergence. *Ann. Appl. Probab.*, 14(3):1353–1377, 2004.
- [Duf96] M. DUFLO : *Algorithmes stochastiques*, volume 23 de *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 1996.

- [Duf97] M. DUFLO : *Random iterative models*, volume 34 de *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1997. Translated from the 1990 French original by Stephen S. Wilson and revised by the author.
- [EK86] S. N. ETHIER et T. G. KURTZ : *Markov processes*. Wiley Series in Probability and Mathematical Statistics : Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.
- [Fau92] O. FAURE : *Simulation du mouvement brownien et des diffusions*. Thèse de doctorat, École nationale des ponts et chaussées, février 1992.
- [Fel52] W. FELLER : The parabolic differential equations and the associated semi-groups of transformations. *Ann. of Math. (2)*, 55:468–519, 1952.
- [Fel54] W. FELLER : Diffusion processes in one dimension. *Trans. Amer. Math. Soc.*, 77:1–31, 1954.
- [GL78] M. I. GORDIN et B. A. LIFŠIC : Central limit theorem for stationary Markov processes. *Dokl. Akad. Nauk SSSR*, 239(4):766–767, 1978.
- [GRS99] H. GANIDIS, B. ROYNETTE et F. SIMONOT : Convergence rate of some semi-groups to their invariant probability. *Stochastic Process. Appl.*, 79(2):243–263, 1999.
- [Has80] R. Z. HAS'MINSKIĬ : *Stochastic stability of differential equations*, volume 7 de *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.
- [HH80] P. HALL et C. C. HEYDE : *Martingale limit theory and its application*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980. Probability and Mathematical Statistics.
- [HW99] J. HUBBARD et B. WEST : *Équations différentielles et systèmes dynamiques*. Cassini, Paris, 1999.
- [IL89] R. A. IBRAHIM et W. LI : Principal internal resonances in 3-dof systems subjected to wide-band random excitation. *J. Sound Vibration*, 131(2):305–321, 1989.
- [IM74] K. ITÔ et H. P. MCKEAN, Jr. : *Diffusion processes and their sample paths*. Springer-Verlag, Berlin, 1974. Second printing, corrected, Die Grundlehren der mathematischen Wissenschaften, Band 125.
- [KS91] I. KARATZAS et S. E. SHREVE : *Brownian motion and stochastic calculus*, volume 113 de *Graduate Texts in Mathematics*. Springer-Verlag, New York, second édition, 1991.
- [KT81] S. KARLIN et H. M. TAYLOR : *A second course in stochastic processes*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1981.

- [LMA03] H. LAMBA, J.C MATTINGLY et Stuart A.M. : Strong convergence of an adaptive euler-maruyama scheme for stochastic differential equations i. 2003.
- [LP02] D. LAMBERTON et G. PAGÈS : Recursive computation of the invariant distribution of a diffusion. *Bernoulli*, 8(3):367–405, 2002.
- [LP03] D. LAMBERTON et 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.
- [LPW92] L. LJUNG, G. PFLUG et H. WALK : *Stochastic approximation and optimization of random systems*, volume 17 de *DMV Seminar*. Birkhäuser Verlag, Basel, 1992.
- [Maa01] F. MAAOUI : Principes d’invariance par moyennisation logarithmique pour les processus de Markov. *Ann. Probab.*, 29(4):1859–1902, 2001.
- [MS04] J.C. MATTINGLY et A.M. STUART : Strong convergence of an adaptive euler-maruyama scheme for stochastic differential equations ii. 2004.
- [MSH02] J. C. MATTINGLY, A. M. STUART et D. J. HIGHAM : Ergodicity for SDEs and approximations : locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2):185–232, 2002.
- [Pag01] G. PAGÈS : Sur quelques algorithmes récurrents pour les probabilités numériques. *ESAIM Probab. Statist.*, 5:141–170 (electronic), 2001.
- [PS94] M. PICCIONI et S. SCARLATTI : An iterative Monte Carlo scheme for generating Lie group-valued random variables. *Adv. in Appl. Probab.*, 26(3):616–628, 1994.
- [PV01] E. PARDOUX et A. Yu. VERETENNIKOV : On the Poisson equation and diffusion approximation. I. *Ann. Probab.*, 29(3):1061–1085, 2001.
- [RW94] L. C. G. ROGERS et D. WILLIAMS : *Diffusions, Markov processes, and martingales. Vol. 1*. Wiley Series in Probability and Mathematical Statistics : Probability and Mathematical Statistics. John Wiley & Sons Ltd., Chichester, second édition, 1994. Foundations.
- [RY99] D. REVUZ et M. YOR : *Continuous martingales and Brownian motion*, volume 293 de *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third édition, 1999.
- [Soi94] C. SOIZE : *The Fokker-Planck equation for stochastic dynamical systems and its explicit steady state solutions*, volume 17 de *Series on Advances in Mathematics for Applied Sciences*. World Scientific Publishing Co. Inc., River Edge, NJ, 1994.
- [Sto74] W. F. STOUT : *Almost sure convergence*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1974. Probability and Mathematical Statistics, Vol. 24.

## BIBLIOGRAPHIE

---

- [Tal90] D. TALAY : Second order discretization schemes for stochastic differential systems for the computation of the invariant law. *Stochastics and Stochastics Reports*, 29:13–36, 1990.
- [Tal02] D. TALAY : Stochastic Hamiltonian systems : exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields*, 8(2):163–198, 2002. Inhomogeneous random systems (Cergy-Pontoise, 2001).