Ninomiya-Victoir scheme : strong convergence, asymptotics for the normalized error and multilevel Monte Carlo methods

Anis Al Gerbi

UNIVERSITÉ PARIS-EST

École Doctorale: Mathématiques et Sciences et Technologies de l'Information et de la Communication

THÈSE DE DOCTORAT

Spécialité : Mathématiques appliquées

Présentée par

Anis Al Gerbi

Ninomiya-Victoir scheme: strong convergence, asymptotics for the normalized error and multilevel Monte Carlo methods

Thèse dirigée par Emmanuelle Clément et Benjamin Jourdain au CERMICS, École des Ponts ParisTech

Soutenue le 10 octobre 2016 devant un jury composé de :

Mme. Emmanuelle Clément	Université Paris-Est Marne-la-Vallée	Directeur de thèse
M. Pierre Henry-Labordère	Société Générale	Examinateur
M. Benjamin Jourdain	École des Ponts ParisTech	Directeur de thèse
M. Ahmed Kebaier	Université Paris 13	Examinateur
M. Antoine Lejay	INRIA	Rapporteur
M. Gilles Pagès	Université Pierre et Marie Curie	Examinateur
M. Fabien Panloup	Université d'Angers	Rapporteur

Cette thèse est dédiée à mes parents.

Acknowledgements

Bien évidemment, c'est à mes deux directeurs de thèse, Emmanuelle Clément et Benjamin Jourdain, que vont mes premiers remerciements. Je les remercie pour leur confiance, leur patience à mon égard, ainsi que pour leur disponibilité. Je tiens à souligner leur rigueur scientifique, leurs compétences ainsi que leur gentillesse. Ils ont su rendre ces années agréables et enrichissantes. Leurs larges connaissances scientifiques et leur soutien permanent ont été décisifs pour le bon déroulement de ma thèse.

Je suis très reconnaissant envers la chaire "Risques Financiers" de la fondation du risque pour avoir participé au financement de ma thèse. En particulier, je voudrais remercier Pierre Henry-Labordere pour avoir accepté de faire partie de mon jury.

Je tiens à adresser mes remerciements les plus sincères à Antoine Lejay et à Fabien Panloup pour l'intérêt qu'ils ont porté à mon travail en acceptant d'être les rapporteurs de ma thèse. Je suis également très honoré de la présence d'Ahmed Kebaier et de Gilles Pagès dans mon jury de thèse.

Je remercie et salue l'ensemble des personnes que j'ai côtoyées pendant ces trois années, les doctorants, post-doctorants et stagiaires. Mes remerciements vont également à tous les membres du CERMICS, en particuliers à Jean-François Delmas et à mes anciens professeurs Bernard Lapeyre, Tony Lelièvre, Aurélien Alfonsi, Gabriel Stoltz et Jean-Philippe Chancelier. Un grand merci à Isabelle Simunic pour sa sympathie, son soutien et son travail efficace et remarquable.

Je souhaite également remercier l'ensemble des membres de ma famille pour leur confiance et leur soutien. Ils sont la raison principale de cet accomplissement et je ne peux exprimer en quelques lignes la gratitude que j'ai pour eux.

Mes derniers remerciements vont enfin à l'ensemble de mes amis de Lyon, Paris, Londres ou d'ailleurs. Leur amitié m'a beaucoup apporté.

Abstract

This thesis is dedicated to the study of the strong convergence properties of the Ninomiya-Victoir scheme, which is based on the resolution of $d+1$ ordinary differential equations (ODEs) at each time step, to approximate the solution to a stochastic differential equation (SDE), where d is the dimension of the Brownian. This study is aimed at analysing the use of this scheme in a multilevel Monte Carlo estimator. Indeed, the optimal complexity of this method is driven by the order of convergence to zero of the variance between the two schemes used on the coarse and fine grids at each level, which is related to the strong convergence order between the two schemes.

In the second chapter, we prove strong convergence with order $1/2$ of the Ninomiya-Victoir scheme $X^{NV,\eta}$, with time step T/N, to the solution X of the limiting SDE. Recently, Giles and Szpruch proposed a modified Milstein scheme and its antithetic version, based on the swapping of each successive pair of Brownian increments in the scheme, permitting to construct a multilevel Monte Carlo estimator achieving the optimal complexity $O\left(\epsilon^{-2}\right)$ for the precision ϵ, as in a simple Monte Carlo method with independent and identically distributed unbiased random variables. In the same spirit, we propose a modified Ninomiya-Victoir scheme, which may be strongly coupled with order 1 to the Giles-Szpruch scheme at the finest level of a multilevel Monte Carlo estimator. This idea is inspired by Debrabant and Rössler who suggest to use a scheme with high order of weak convergence on the finest grid at the finest level of the multilevel Monte Carlo method. As the optimal number of discretization levels is related to the weak order of the scheme used in the finest grid at the finest level, Debrabant and Rössler manage to reduce the computational time, by decreasing the number of discretization levels. The coupling with the Giles-Szpruch scheme allows us to combine both ideas. By this way, we preserve the optimal complexity $O\left(\epsilon^{-2}\right)$ and we reduce the computational time, since the Ninomiya-Victoir scheme is known to exhibit weak convergence with order 2.

In the third chapter, we check that the normalized error defined by $\sqrt{N}\left(X - X^{NV,\eta}\right)$ converges to an affine SDE with source terms involving the Lie brackets between the Brownian vector fields. This result ensures that the strong convergence rate is actually $1/2$ when at least two Brownian vector fields do not commute. To link this result to the multilevel Monte Carlo estimator, it can be seen as a first step to adapt to the Ninomiya-Victoir scheme the central limit theorem of Lindeberg Feller type, derived recently by Ben Alaya and Kebaier for the multilevel Monte Carlo estimator based on the Euler scheme. When the Brownian vector fields commute, the limit vanishes. We then prove strong convergence with order 1 in this case.

The fourth chapter deals with the convergence of the normalized error process $N\left(X - X^{NV}\right)$, where X^{NV} is the Ninomiya-Victoir in the commutative case. We prove its stable convergence in law to an affine SDE with source terms involving the Lie brackets between the Brownian vector fields and the drift vector field. This result ensures that the strong convergence rate is actually 1 when the Brownian vector fields commute, but at least one of them does not commute with the Stratonovich drift vector field.

Finally, in the fifth chapter, we analyse the use of approximations for all ODEs in the Ninomiya-Victoir scheme. It is well known that if the numerical integration is accurate up to the order 3 for the Stratonovich drift vector filed and up to the order 6 for the Brownian vector fields, the weak order 2 is then preserved. In this last chapter, we prove that these conditions are suitable to obtain a strong convergence with order 1 between the Ninomiya-Victoir scheme and its approximation, induced by the use of approximations for all ODEs. This allows us to replace the Ninomiya-Victoir scheme by its numerical approximation, in our multilevel Monte Carlo estimators, when one of the ODEs has no closed-form solution and keep the computational complexity $O\left(\epsilon^{-2}\right)$.

Résumé

Cette thèse est consacrée à l'étude des propriétés de convergence forte du schéma de Ninomiya et Victoir. Les auteurs de ce schéma proposent d'approcher la solution d'une équation différentielle stochastique (EDS), notée X, en résolvant $d+1$ équations différentielles ordinaires (EDOs) sur chaque pas de temps, où d est la dimension du mouvement brownien. Le but de cette étude est d'analyser l'utilisation de ce schéma dans une méthode de Monte-Carlo multi-pas. En effet, la complexité optimale de cette méthode est dirigée par l'ordre de convergence vers 0 de la variance entre les schémas utilisés sur la grille grossière et sur la grille fine. Cet ordre de convergence est lui-même lié à l'ordre de convergence fort entre les deux schémas.

Nous montrons alors dans le chapitre 2, que l'ordre fort du schéma de Ninomiya-Victoir, noté $X^{NV,\eta}$ et de pas de temps T/N, est $1/2$. Récemment, Giles et Szpruch ont proposé un estimateur Monte-Carlo multi-pas réalisant une complexité $O\left(\epsilon^{-2}\right)$ à l'aide d'un schéma de Milstein modifié. Dans le même esprit, nous proposons un schéma de Ninomiya-Victoir modifié qui peut-être couplé à l'ordre fort 1 avec le schéma de Giles et Szpruch au dernier niveau d'une méthode de Monte-Carlo multi-pas. Cette idée est inspirée de Debrabant et Rossler. Ces auteurs suggèrent d'utiliser un schéma d'ordre faible élevé au niveau de discrétisation le plus fin. Puisque le nombre optimal de niveaux de discrétisation d'une méthode de Monte-Carlo multi-pas est dirigé par l'erreur faible du schéma utilisé sur la grille fine du dernier niveau de discrétisation, cette technique permet d'accélérer la convergence de la méthode Monte-Carlo multi-pas en obtenant une approximation d'ordre faible élevé. L'utilisation du couplage à l'ordre 1 avec le schéma de Giles-Szpruch nous permet ainsi de garder un estimateur Monte-Carlo multi-pas réalisant une complexité optimale $O\left(\epsilon^{-2}\right)$ tout en profitant de l'erreur faible d'ordre 2 du schéma de Ninomiya-Victoir.

Dans le troisième chapitre, nous nous sommes intéressés à l'erreur renormalisée définie par $\sqrt{N}\left(X - X^{NV,\eta}\right)$. Nous montrons la convergence en loi stable vers la solution d'une EDS affine, dont le terme source est formé des crochets de Lie entre les champs de vecteurs browniens. Ainsi, lorsqu'au moins deux champs de vecteurs browniens ne commutent pas, la limite n'est pas triviale. Ce qui assure que l'ordre fort $1/2$ est optimal. D'autre part, ce résultat peut être vu comme une première étape en vue de prouver un théorème de la limite centrale pour les estimateurs Monte-Carlo multi-pas. Pour cela, il faut analyser l'erreur en loi stable du schéma entre deux niveaux de discrétisation successifs. Ben Alaya et Kebaier ont prouvé un tel résultat pour le schéma d'Euler. Lorsque les champs de vecteurs browniens commutent, le processus limite est nul. Nous montrons que dans ce cas précis, que l'ordre fort est 1.

Dans le chapitre 4, nous étudions la convergence en loi stable de l'erreur renormalisée $N\left(X - X^{NV}\right)$ où X^{NV} est le schéma de Ninomiya-Victoir lorsque les champs de vecteurs browniens commutent. Nous démontrons la convergence du processus d'erreur renormalisé vers la solution d'une EDS affine. Lorsque le champ de vecteurs dritf ne commute pas avec au moins

un des champs de vecteurs browniens, la vitesse de convergence forte obtenue précédemment est optimale.

Le chapitre 5 de cette thèse est consacré à l'utilisation de schémas numériques pour les EDOs dans l'implémentation du schéma de Ninomiya et Victoir. Pour conserver l'ordre faible 2, il suffit d'utiliser un schéma numérique où l'erreur de troncature est d'ordre 3 pour l'EDO portant sur le champ de vecteurs drift et un schéma numérique où l'erreur de troncature est d'ordre 6 pour l'EDO portant sur les champs de vecteurs browniens. Nous montrons que cela suffit pour obtenir une erreur forte d'ordre 1 entre le schéma de Ninomiya et Victoir et son approximation numérique. On peut ainsi remplacer le schéma de Ninomiya et Victoir par son approximation numérique dans les estimateurs Monte-Carlo multi-pas tout en conservant la complexité $O\left(\epsilon^{-2}\right)$.

Table of contents

List of figures

Chapter 1

Introduction

En finance, les modèles d'évolution des prix des actifs financiers reposent en général sur des équations différentielles stochastiques. Pour décrire l'évolution des indices boursiers, les modèles les plus connus sont ceux de Black-Scholes-Merton [12] ou encore Heston [27]. Il existe aussi des modèles pour décrire l'évolution des taux d'intérêt, citons par exemple le modèle de Vasicek [54], le modèle Cox-Ingersoll-Ross [15] ou encore le modèle de Ho-Lee [28]. Les équations différentielles stochastiques sont aussi très utilisées dans d'autres disciplines, par exemple en physique et en biologie.

Précisons un peu plus les notations que nous allons utiliser dans cette introduction. Nous considérons une équation différentielle stochastique $n-$dimensionnelle, dirigée par un mouvement brownien standard $d-$dimensionnel noté W, avec un horizon de temps $T \in \mathbb{R}_+^*$, de la forme

$$\begin{cases} dX_t = b(X_t)dt + \sum_{j=1}^{d} \sigma^j(X_t)dW_t^j, t \in [0,T], \\ X_0 = x_0. \end{cases} \tag{1.0.1}$$

La condition initiale $x_0 \in \mathbb{R}^n$ est supposée déterministe, le champ de vecteurs $b : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ correspond au coefficient de dérive, le champ de vecteurs $\sigma^j : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, pour j$\in \{1,\ldots,d\}$, correspond à la j-ème colonne de la matrice de diffusion notée σ. D'après le théorème d'Itô [51], si les champs de vecteurs b et σ^j, pour $j \in \{1,\ldots,d\}$, sont globalement lipschitziens, alors il existe une unique solution forte à l'équation différentielle stochastique (1.0.1). En outre, à chaque instant $t \in [0,T]$, la solution X_t possède des moments d'ordre p, pour tout $p \in [1,+\infty)$. Ce résultat correspond à l'analogue du théorème de Cauchy-Lipschitz global pour les équations différentielles ordinaires [53].

Un exemple de problématique pratique est le calcul d'espérance, notée Y, d'une fonction, notée $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, de la solution de l'équation différentielle stochastique (1.0.1) à la date terminale T, c'est-à-dire

$$Y = \mathbb{E}\left[f\left(X_T\right)\right]. \tag{1.0.2}$$

En finance, lorsque X correspond au prix d'un actif, appelé sous-jacent, la quantité Y correspond au prix d'un produit dérivé associé à ce sous-jacent. Sauf cas très particulier, en général il n'y a pas de formule explicite pour Y. Nous avons donc besoin de recourir à une approximation numérique pour le calcul de la quantité Y. Pour obtenir une approximation numérique de Y, il y a essentiellement deux approches possibles. La première approche, déterministe, repose sur la formule de représentation de Feynman-Kac [51]. Cette formule établit un lien entre équations aux dérivées partielles paraboliques et calcul d'espérance. Plus précisément, considérons l'équation aux dérivées partielles rétrograde suivante

$$\begin{cases} \partial_t u\,(t,x) + \mathcal{L}u\,(t,x) = 0, \quad (t,x) \in [0,T] \times \mathbb{R}^n, \\ u\,(T,x) = f(x), \end{cases} \tag{1.0.3}$$

où

$$\mathcal{L}u\,(t,x) = \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{n} a^{i,k}\,(x)\,\partial^2_{x_i,x_k}u\,(t,x) + \sum_{i=1}^{n} b^i\,(x)\,\partial_{x_i}u\,(t,x)\,,$$

avec $a = \sigma\sigma^*$, c'est-à-dire,

$$\forall i,k \in \{1,\ldots,n\}\,, a^{i,j}\,(x) = \sum_{j=1}^{d} \sigma^{ij}\,(x)\,\sigma^{kj}\,(x)\,.$$

L'opérateur $\mathcal{L}$ est appelé générateur infinitésimal de la diffusion X. Sous certaines hypothèses de régularité (voir [47] ou [51] par exemple), la solution de l'équation aux dérivées partielles (1.0.3) admet la représentation probabiliste suivante

$$u\,(t,x) = \mathbb{E}\left[f\,(X_T)\,|X_t = x\right],$$

où X est solution de (1.0.1). En particulier, nous avons $Y = u\,(0,x_0)$. Ainsi, pour obtenir une estimation de la quantité Y, il suffit d'approcher numériquement la solution de l'équation aux dérivées partielles parabolique (1.0.3) par un schéma numérique de type éléments finis ou de type différences finies (voir [49] par exemple). Cette méthode est assez précise, mais assez difficile à mettre en œuvre en grande dimension. La seconde approche pour estimer la quantité Y est appelée méthode de Monte-Carlo. Contrairement à l'approche précédente, les méthodes de Monte-Carlo sont relativement simples à implémenter et restent efficaces même en grande dimension.

1.1 Discrétisation d'EDS et méthodes de Monte-Carlo

L'approximation de Y par la méthode de Monte-Carlo standard consiste à simuler $M \in \mathbb{N}^*$ copies indépendantes de la solution de l'équation différentielle stochastique (1.0.1), à l'instant terminal T, notées $X_T^1, \ldots, X_T^M$, puis calculer la moyenne empirique suivante

$$\bar{Y}_{MC} = \frac{1}{M}\sum_{k=1}^{M} f\left(X_T^k\right).$$

La convergence de cet estimateur vers la quantité Y, lorsque le nombre de tirages M tend vers l'infini, est assurée par la loi forte des grands nombres. En outre, si $\mathbb{E}\left[(f\,(X_T))^2\right] < \infty$, alors le théorème central limite fournit l'intervalle de confiance suivant

$$\left[\bar{Y}_{MC} - z_{1-\frac{\alpha}{2}}\frac{S_M}{\sqrt{M}}; \bar{Y}_{MC} + z_{1-\frac{\alpha}{2}}\frac{S_M}{\sqrt{M}}\right],$$

où, $\alpha \in (0,1)$ représente le risque de l'intervalle de confiance, $z_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1-\alpha/2$ d'une loi normale standard et

$$S_M^2 = \frac{1}{M-1} \sum_{k=1}^{M} \left(f\left(X_T^k\right) - \bar{Y}_{MC} \right)^2,$$

est un estimateur (consistant) de la variance $\mathbb{V}\left[f\left(X_T\right)\right]$. L'erreur de cette méthode est donc une erreur statistique. Ainsi, si nous voulons un intervalle de confiance de longueur inférieure à $\epsilon \in \mathbb{R}_+^*$, le nombre de tirages M doit être proportionnel à ϵ^{-2}. On dit alors que la complexité est en $O\left(\epsilon^{-2}\right)$.

Lorsque $n = d = 1$, Beskos, Papaspiliopoulos et Roberts [11] ont proposé une méthode de simulation exacte pour la solution de l'EDS (1.0.1). En revanche, sauf cas particulier, en dimension supérieure, les solutions d'équations différentielles stochastiques ne sont pas simulables. Ce qui nous amène donc à discuter de la discrétisation des équations différentielles stochastiques.

1.1.1 Discrétisation d'EDS

Pour discrétiser l'équation différentielle stochastique (1.0.1), nous considérons une grille temporelle uniforme, de pas de temps $h = T/N$ et nous notons $\left(t_k = kh\right)_{0 \leq k \leq N}$ les instants de discrétisation. La discrétisation la plus naturelle est donnée par le schéma d'Euler [34]

$$\begin{cases} X_{t_0}^E = x_0, \\ X_{t_{k+1}}^E = X_{t_k}^E + b\left(X_{t_k}^E\right) h + \displaystyle\sum_{j=1}^{d} \sigma^j \left(X_{t_k}^E\right) \Delta W_{t_{k+1}}^j, \end{cases}$$

où $\Delta W_{t_{k+1}}^j = W_{t_{k+1}}^j - W_{t_k}^j$. Ainsi, s'il n'existe pas de méthode de simulation exacte pour la solution de (1.0.1), on peut adapter la méthode de Monte-Carlo précédente en procédant de la façon suivante. On commence par choisir un pas de temps $h = T/N$. On simule ensuite M copies indépendantes du schéma d'Euler, à l'instant terminal T, notées $X_T^{E,1}, \ldots, X_T^{E,M}$, puis on calcule la moyenne empirique

$$\hat{Y}_{MC}^E = \frac{1}{M} \sum_{k=1}^{M} f\left(X_T^{E,k}\right).$$

Outre l'erreur statistique, l'introduction d'un schéma numérique d'approximation de la solution X induit un biais dû à l'erreur de discrétisation. Pour mesurer ce biais, il faut donc comparer le schéma de discrétisation à la solution X. Pour pouvoir comparer le schéma d'Euler à la solution X, il est assez commode d'utiliser l'interpolation suivante

$$\begin{cases} X_{t_0}^E = x_0, \\ dX_t^E = b\left(X_{\hat{\tau}_t}^E\right) dt + \displaystyle\sum_{j=1}^{d} \sigma^j \left(X_{\hat{\tau}_t}^E\right) dW_t^j, \end{cases}$$

où $\hat{\tau}_t$, désigne l'instant de discrétisation avant $t \in [0,T)$, c'est-à-dire, $\hat{\tau}_t = t_k$ si $t \in [t_k, t_{k+1})$. Un critère de convergence assez naturel est d'analyser l'erreur dite forte. Plus précisément, il s'agit d'étudier la dépendance en h de la norme L^{2p}, pour $p \in [1, +\infty)$, sur l'espace des trajectoires

$$\left(\mathbb{E}\left[\sup_{t\leq T}\left\|X_t - X_t^E\right\|^{2p}\right]\right)^{\frac{1}{2p}}.$$

Bien sûr, le schéma d'Euler X^E dépend du pas de temps h à travers le nombre N d'instants de discrétisations de la grille temporelle. D'après Kanagawa [32], lorsque les champs de vecteurs b et σ^j, pour j$\in \{1, \ldots, d\}$, sont globalement lipschitziens, nous avons l'estimation suivante

$$\mathbb{E}\left[\sup_{t\leq T}\left\|X_t - X_t^E\right\|^{2p}\right] \leq C_E h^p.$$

Ici, la constante $C_E \in \mathbb{R}_+^*$ ne dépend pas de N. On dit alors que le schéma d'Euler est d'ordre fort 1/2. Plus généralement, on dit que le schéma $\hat{X}$ est d'ordre γ si l'on a l'estimation

$$\mathbb{E}\left[\sup_{t\leq T}\left\|X_t - \hat{X}_t\right\|^{2p}\right] \leq Ch^{2\gamma p}, \tag{1.1.1}$$

pour une certaine constante $C \in \mathbb{R}_+^*$ indépendante de N. Cependant, pour étudier le biais d'un schéma numérique dans une méthode de Monte-Carlo, il faut plutôt regarder l'erreur faible. Cette erreur correspond à l'erreur en loi à l'instant terminal T, c'est-à-dire $\mathbb{E}\left[f\left(X_T\right) - f\left(\hat{X}_T\right)\right]$. En effet, l'erreur quadratique de l'estimateur Monte-Carlo

$$\hat{Y}_{MC} = \frac{1}{M}\sum_{k=1}^{M} f\left(\hat{X}_T^k\right),$$

où les $\left(\hat{X}_T^k\right)_{1\leq k\leq N}$ sont des copies indépendantes du schéma $\hat{X}$ à l'instant terminal T, se décompose de la façon suivante

$$\begin{aligned}\mathbb{E}\left[\left(\hat{Y}_{MC} - Y\right)^2\right] &= \left(\mathbb{E}\left[f\left(X_T\right) - f\left(\hat{X}_T\right)\right]\right)^2 + \mathbb{V}\left[\hat{Y}_{MC}\right]\\ &= \left(\mathbb{E}\left[f\left(X_T\right) - f\left(\hat{X}_T\right)\right]\right)^2 + \frac{1}{M}\mathbb{V}\left[f\left(\hat{X}_T\right)\right].\end{aligned}$$

Ainsi, en plus du terme d'erreur statistique, nous avons un terme de biais dirigé par l'erreur faible du schéma numérique. On dit que le schéma numérique $\hat{X}$, de pas de temps h, est d'ordre faible α si on a l'estimation

$$\left|\mathbb{E}\left[f\left(X_T\right) - f\left(\hat{X}_T\right)\right]\right| \leq \hat{C}_1 h^{\alpha},$$

pour une constante $\hat{C}_1$ indépendante de h et pour toute fonction f suffisamment régulière. Pour assurer une erreur en norme L^2 inférieure à $\epsilon \in \mathbb{R}_+^*$, il faut alors choisir le nombre d'instants de discrétisation N proportionnel à $\epsilon^{-\frac{1}{\alpha}}$. Le temps de calcul de cette méthode Monte-Carlo étant proportionnel à $N \times M$, où M est le nombre de tirages Monte-Carlo, on dit alors que la complexité est en $O\left(\epsilon^{-\left(2+\frac{1}{\alpha}\right)}\right)$ [19]. Cela explique pourquoi l'erreur faible a été largement étudiée au cours de ces dernières décennies.

En supposant les champs de vecteurs b et σ^j, pour j$\in \{1, \ldots, d\}$, de classe $\mathcal{C}^\infty$ avec des dérivées bornées, Talay et Tubaro [52] ont démontré l'existence d'une suite $(C_k)_{k\geq 1} \in \mathbb{R}^{\mathbb{N}}$ indépendante de h et un développement de l'erreur faible pour le schéma d'Euler de la forme

$$\forall L \in \mathbb{N}^*,\ \mathbb{E}\left[f\left(X_T\right)-f\left(X_T^E\right)\right]=\sum_{k=1}^{L} C_k h^k+O\left(h^{L+1}\right), \tag{1.1.2}$$

pour des fonctions f de classe $\mathcal{C}^\infty$ avec des dérivées à croissance polynomiale. Ainsi, en particulier $\alpha = 1$ pour le schéma d'Euler. En supposant une condition de non-dégénérescence de type Hörmander [29] sur le générateur infinitésimal de la diffusion X, Bally et Talay [7] [8] ont montré, à l'aide du calcul de Malliavin, que $\alpha = 1$ pour des fonctions f mesurables bornées. Guyon [24] a montré ce résultat pour des distributions tempérées f, en utilisant cette fois-ci une hypothèse, plus forte, d'uniforme ellipticité sur le générateur infinitésimal de la diffusion X. Ainsi, pour une classe assez large de fonctions f, la complexité de l'estimateur $\hat{Y}_{MC}^E$ est $O\left(\epsilon^{-3}\right)$.

Un autre schéma assez connu est celui de Milstein [42]. Celui-ci s'écrit

$$\begin{cases} X_{t_0}^M = x_0, \\ X_{t_{k+1}}^M = X_{t_k}^M + b\left(X_{t_k}^M\right) h + \displaystyle\sum_{j=1}^{d} \sigma^j\left(X_{t_k}^M\right) \Delta W_{t_{k+1}}^j \\ \qquad + \dfrac{1}{2} \displaystyle\sum_{j=1}^{d}\sum_{m=1}^{d} \partial\sigma^j\sigma^m\left(X_{t_k}^M\right)\left(\Delta W_{t_{k+1}}^j \Delta W_{t_{k+1}}^j - \mathbb{1}_{\{j=m\}} h - A_{t_{k+1}}^{j,m}\right), \end{cases}$$

où $\partial\sigma^j$ est la matrice jacobienne du champ de vecteurs σ^j définie par

$$\partial\sigma^j = \left(\partial_{x_k}\sigma^{ij}\right)_{1\le i,k\le n},$$

et $A_{t_{k+1}}^{j,m}$ représente l'aire de Lévy définie par

$$A_{t_{k+1}}^{j,m} = \int_{t_k}^{t_{k+1}} \left(W_s^j - W_{t_k}^j\right) dW_s^m - \int_{t_k}^{t_{k+1}} \left(W_s^m - W_{t_k}^m\right) dW_s^j.$$

Pour ce schéma, sous des hypothèses de régularité sur les coefficients de l'équation différentielle stochastique, l'ordre faible α vaut 1. L'avantage de ce schéma est qu'il est d'ordre fort $\gamma = 1$. En revanche, lorsque la dimension d du mouvement brownien est supérieure à 1, celui-ci fait intervenir des aires de Lévy que l'on ne sait pas simuler sauf en dimension $d = 2$. En utilisant la fonction caractéristique du triplet

$$\left(W_t^1, W_t^2, \int_0^t W_s^1 dW_s^2 - \int_0^t W_s^2 dW_s^1\right),$$

Gaines et Lyons [21] ont réussi à proposer une méthode de simulation du triplet et ainsi simuler le schéma de Milstein en dimension $d = 2$.

Pour accélérer la convergence de la méthode de Monte-Carlo, la construction de schémas d'ordre faible élevé a fait l'objet d'une recherche active ces dernières années, notamment motivée par les applications des méthodes de Monte-Carlo en finance. Récemment, Kusuoka [37] [36] a introduit une classe de schémas d'ordre faible arbitrairement élevé. Pour cela, il a utilisé des développements de Taylor stochastiques, introduits par Platen et Wagner [48], et remplacé les intégrales stochastiques itérées,

$$\int_{0\le s_1 \ldots \le s_m \le t} dW_{s_1}^{j_1} \ldots dW_{s_m}^{j_m},$$

qui apparaissent dans ces développements et que l'on ne sait pas simuler, par des variables aléatoires simulables préservant les moments de ces intégrales jusqu'à un certain ordre. Mentionnons aussi la notion de cubature sur l'espace de Wiener, introduite par Lyons et Victoir [39], qui a permis la construction de schémas d'ordre faible élevé [25] et notamment le schéma de Ninomiya et Victoir [44] qui sera largement étudié dans cette thèse.

1.1.2 Méthodes de Monte-Carlo multi-pas

L'objectif des méthodes de Monte-Carlo multi-pas est de réduire la complexité de la méthode de Monte-Carlo standard, la quantité cible à estimer étant toujours Y donnée par (1.0.2) En combinant astucieusement deux schémas d'Euler avec des pas de discrétisation différents, Kebaier [33] a réussi à réduire la complexité, et donc le temps de calcul, de l'estimation de Y. Sa méthode porte le nom de méthode de Romberg statistique et repose sur le principe suivant. On considère deux grilles de discrétisation, une grossière de pas de temps T/N_c, et une fine de pas de temps T/N_f telles que $N_c = N_f^R$ avec $R \in (0,1)$. L'estimateur de Romberg statistique s'écrit de la façon suivante

$$\hat{Y}_{SR}^{E} = \frac{1}{M_c}\sum_{k=1}^{M_c} f\left(X_T^{E,N_c,c,k}\right) + \frac{1}{M_f}\sum_{k=1}^{M_f}\left(f\left(X_T^{E,N_f,f,k}\right) - f\left(X_T^{E,N_c,f,k}\right)\right),$$

où les variables aléatoires $\left(X_T^{E,N_c,c,k}\right)_{0\le k\le M_c}$ sont des copies indépendantes du schéma d'Euler sur la grille grossière, les variables aléatoires $\left(X_T^{E,N_f,f,k}\right)_{0\le k\le M_f}$ et $\left(X_T^{E,N_c,f,k}\right)_{0\le k\le M_f}$ sont respectivement des copies indépendantes du schéma d'Euler sur la grille fine et grossière telles que

- pour tout $k \in \{1,\ldots,M_f\}$, les schémas $\left(X_T^{E,N_f,f,k}\right)$ et $\left(X_T^{E,N_c,f,k}\right)$ sont simulés avec les mêmes accroissements browniens,
- la suite $\left(X_T^{E,N_c,c,k}\right)_{0\le k\le M_c}$ est indépendante de la suite $\left(X_T^{E,N_f,f,k}, X_T^{E,N_c,f,k}\right)_{0\le k\le M_f}$.

La complexité, ou temps de calcul, de cette méthode est définie par

$$\mathcal{C}_{SR} = K_E\left(M_c N_c + M_f N_f\right),$$

où K_E est une constante qui dépend uniquement du schéma d'Euler. Intéressons-nous maintenant au biais de cette méthode. Puisque les schémas numériques sont simulés avec des accroissements browniens, nous avons $X_T^{E,N_c,f} \overset{\mathcal{L}}{=} X_T^{E,N_c,c}$. Ainsi, en prenant l'espérance de $\hat{Y}_{SR}^{E}$ on obtient

$$\mathbb{E}\left[\hat{Y}_{SR}^{E}\right] = \mathbb{E}\left[f\left(X_T^{E,N_f}\right)\right].$$

Le biais est donc dirigé par le schéma le plus fin. Ainsi, cette méthode peut être vue comme une méthode de variable de contrôle pour l'estimation de $\mathbb{E}\left[f\left(X_T^{E,N_f}\right)\right]$. En effet, le fait de simuler les schémas $\left(X_T^{E,N_f,f}\right)$ et $\left(X_T^{E,N_c,f}\right)$ avec les mêmes accroissements browniens réduit grandement la variance. En utilisant les propriétés d'erreur forte du schéma d'Euler, on montre facilement l'estimation suivante

$$\mathbb{V}\left[\hat{Y}_{SR}^{E}\right] \leq \frac{1}{M_c}\mathbb{V}\left[f\left(X_T^{E,N_c}\right)\right] + \frac{2L_{ip}(f)^2 C_E}{M_f}\left(\frac{T}{N_f}+\frac{T}{N_c}\right),$$

où $L_{ip}(f)$ est la constante de Lipschitz de la fonction f. Contrairement à une méthode de Monte-Carlo standard, la variance est une fonction décroissante des pas de discrétisations T/N_f et T/N_c. On peut donc optimiser les paramètres (N_c, N_f, M_c, M_f) pour minimiser la complexité, sous contrainte de précision ϵ. En choisissant astucieusement ces paramètres, Kebaier a obtenu une complexité optimale $\mathcal{C}_{SR}^{*} = O\left(\epsilon^{-5/2}\right)$.

Giles [22] a généralisé l'approche de Kebaier en construisant un estimateur dit Monte-Carlo multi-pas. Au lieu d'utiliser uniquement 2 grilles de discrétisation, Giles considère une suite de pas de temps géométrique $\left(h_l = T/2^l\right)_{0\leq l\leq L}$ où $L \in \mathbb{N}^*$ est le niveau de discrétisation le plus fin. La construction de cet estimateur repose sur l'égalité suivante

$$\mathbb{E}\left[f\left(X_T^{2^L}\right)\right] = \mathbb{E}\left[f\left(X_T^{1}\right)\right] + \sum_{l=1}^{L}\mathbb{E}\left[f\left(X_T^{2^l}\right) - f\left(X_T^{2^{l-1}}\right)\right],$$

où X^{2^l} est un schéma numérique d'approximation de l'équation différentielle stochastique (1.0.1) et de pas de temps h_l. En vue d'estimer Y, au lieu d'estimer le membre de gauche de l'égalité précédente, comme dans une méthode de Monte-Carlo standard, Giles propose d'estimer le membre de droite. Dans [1] nous avons écrit l'estimateur de Monte-Carlo multi-pas de façon générique

$$\hat{Y}_{MLMC} = \sum_{l=0}^{L}\frac{1}{M_l}\sum_{k=1}^{M_l} Z^{l,k},$$

où les variables aléatoires $\left(Z_k^l\right)_{0\leq l\leq L, 1\leq k\leq M_l}$ sont indépendantes et vérifient

$$\forall k \in \{1,\ldots,M_0\}, \mathbb{E}\left[Z^{0,k}\right] = \mathbb{E}\left[f\left(X_T^{1}\right)\right], \tag{1.1.3}$$

et

$$\forall l \in \{1,\ldots,L\}, \forall k \in \{1,\ldots,M_l\}, \mathbb{E}\left[Z^{l,k}\right] = \mathbb{E}\left[f\left(X_T^{2^l}\right) - f\left(X_T^{2^{l-1}}\right)\right], \tag{1.1.4}$$

et $M_l \in \mathbb{N}^*$ correspond au nombre de tirages pour le niveau de discrétisation $l \in \{0,\ldots,L\}$. Comme pour la méthode de Romberg statistique, le choix du niveau de discrétisation le plus fin L est dirigé par le biais du schéma numérique:

$$\left|\mathbb{E}\left[f\left(X_T^{2^l}\right)\right] - Y\right| \leq \frac{c_1}{2^{\alpha l}},$$

pour une constante c_1 indépendante de l. Dans [22], Giles a considéré un choix assez naturel donné par le schéma d'Euler. En notant $X^{E,2^l}$ le schéma d'Euler de pas de temps $h_l = T/2^l$, Giles propose

$$Z_E^0 = f\left(X_T^{E,1}\right),$$

et

$$\forall l \in \{1,\ldots,L\}, Z_E^l = f\left(X_T^{E,2^l}\right) - f\left(X_T^{E,2^{l-1}}\right),$$

où les deux schémas sont dirigés par le même mouvement brownien. Ainsi, la variance de Z_E^l, pour $l \in \{1, \ldots, L\}$, est reliée à l'erreur forte du schéma d'Euler. Plus précisément, nous avons l'estimation suivante

$$\forall l \in \{1, \ldots, L\}, \mathbb{V}\left[Z_E^l\right] \leq \frac{C'_E}{2^l},$$

avec C'_E une constante strictement positive et indépendante de l. Dans le cas général, il est donc naturel de supposer

$$\forall l \in \{1, \ldots, L\}, \mathbb{V}\left[Z^l\right] \leq \frac{c_2}{2^{\beta l}},$$

où c_2 une constante strictement positive et indépendante de l et β est l'ordre de convergence vers 0 de la variance $\mathbb{V}\left[Z^l\right]$. Pour une fonction f lipschitzienne, nous avons $\beta \geq 2\gamma$, où γ est l'ordre de convergence forte du schéma défini par (1.1.1). La complexité de cette méthode est donnée par

$$\mathcal{C}_{MLMC} = K \sum_{l=0}^{L} M_l 2^l,$$

où K est une constante qui dépend uniquement du schéma numérique. Pour assurer un biais inférieur à $\epsilon/\sqrt{2}$, il suffit de choisir

$$L^* = \left\lceil \frac{\log_2\left(\frac{\sqrt{2}c_1}{\epsilon}\right)}{\alpha} \right\rceil.$$

Un contrôle de la variance de l'estimateur $\hat{Y}_{MLMC}$ est donné par

$$\mathbb{V}\left[\hat{Y}_{MLMC}\right] \leq \frac{\mathbb{V}\left[Z^0\right]}{M_0} + c_2 \sum_{l=1}^{L^*} \frac{1}{M_l 2^{\beta l}}.$$

On peut donc minimiser la complexité sous la contrainte $\mathbb{V}\left[\hat{Y}_{MLMC}\right] \leq \epsilon^2/2$. La complexité optimale de cette méthode, pour assurer une erreur en norme L^2 inférieure à ϵ, dépend essentiellement de l'ordre β. Ainsi, Giles a montré dans [22]

$$\begin{cases} \mathcal{C}^*_{MLMC} = O\left(\epsilon^{-2}\right) \text{ si } \beta > 1, \\ \mathcal{C}^*_{MLMC} = O\left(\epsilon^{-2}\left(\log\left(\frac{1}{\epsilon}\right)\right)^2\right) \text{ si } \beta = 1, \\ \mathcal{C}^*_{MLMC} = O\left(\epsilon^{-2+\frac{\beta-1}{\alpha}}\right) \text{ si } \beta < 1. \end{cases}$$

Pour un schéma d'ordre fort 1/2, dans le cas d'un schéma d'Euler par exemple, nous avons une complexité en $O\left(\epsilon^{-2}\left(\log\left(\frac{1}{\epsilon}\right)\right)^2\right)$. En revanche, pour un schéma d'ordre fort strictement supérieur à 1/2, nous avons une complexité en $O\left(\epsilon^{-2}\right)$ comme pour une méthode de Monte-Carlo sans biais. Cependant, Clark et Cameron [14] ont démontré, dans un cas particulier, qu'il n'est pas possible de faire mieux que l'ordre 1/2 lorsque l'on utilise uniquement les accroissements browniens pour discrétiser l'équation différentielle stochastique.

Récemment, des améliorations ont été apportées à l'estimateur Monte-Carlo multi-pas utilisant le schéma d'Euler. Celles-ci sont présentées par ordre d'importance. Une première amélioration envisagée par Debrabant et Rössler est d'utiliser un schéma d'ordre faible élevé au niveau de discrétisation le plus fin. Plus précisément, ces auteurs considèrent l'estimateur suivant

$$\begin{aligned}\hat{Y}_{MLMC}^{DR} = \frac{1}{M_0}\sum_{k=1}^{M_0} f\left(X_T^{E,1,0,k}\right) + \sum_{l=1}^{L-1}\frac{1}{M_l}\sum_{k=1}^{M_l}\left(f\left(X_T^{E,2^l,l,k}\right) - f\left(X_T^{E,2^{l-1},l,k}\right)\right)\\ + \frac{1}{M_L}\sum_{k=1}^{M_L}\left(f\left(\hat{X}_T^{2^L,L,k}\right) - f\left(X_T^{E,2^{L-1},L,k}\right)\right),\end{aligned}$$

où

- les schémas $\left(X^{E,1,0,k}\right)_{1\le k\le M_0}$ sont des copies indépendantes du schéma d'Euler de pas de temps grossier $h_0 = 1$,
- pour tout $l \in \{1,\dots L-1\}$, les schémas $\left(X^{E,2^l,l,k}, X^{E,2^{l-1},l,k}\right)_{1\le k\le M_l}$ sont des copies indépendantes des schémas d'Euler $X^{E,2^l,l}$ et $X^{E,2^{l-1},l}$, de pas de temps h_l et h_{l-1} respectivement, dirigés par le même mouvement brownien,
- les schémas $\left(\hat{X}^{2^L,L,k}, X^{E,2^{L-1},L,k}\right)_{1\le k\le M_L}$ sont des copies indépendantes d'un schéma numérique d'ordre faible élevé $\hat{X}^{2^L,L}$, de pas de temps $h_L = T/2^L$, et du schéma d'Euler et $X^{E,2^{L-1},L}$, de pas de temps h_{L-1}, respectivement, dirigés par le même mouvement brownien.

En outre, les variables aléatoires $\left(X^{E,1,0,k}\right)_{1\le k\le M_0}$, $\left(X^{E,2^l,l,k}, X^{E,2^{l-1},l,k}\right)_{1\le k\le M_l}$ et $\left(\hat{X}^{2^L,L,k}, X^{E,2^{L-1},L,k}\right)_{1\le k\le M_L}$ sont supposées mutuellement indépendantes. Le biais de cet estimateur correspond au biais du schéma $\hat{X}$. On récupère ainsi une approximation d'ordre faible élevé, cela permet alors d'accélérer la convergence de la méthode Monte-Carlo multi-pas en diminuant le nombre de niveaux de discrétisation optimaux L^*. Il n'y a en revanche pas d'amélioration de la complexité qui reste en $O\left(\epsilon^{-2}\left(\log\left(\frac{1}{\epsilon}\right)\right)^2\right)$ puisque $\beta = 1$.

Dans le même esprit que la méthode de Monte-Carlo multi-pas classique, Lemaire et Pagès [38] ont combiné cette idée de Monte-Carlo multi-step avec l'extrapolation de Richardson Romberg généralisée par Pagès dans [46]. La méthode d'extrapolation de Richardson Romberg permet de récupérer une approximation d'ordre faible en exploitant efficacement le développement de l'erreur faible (1.1.2). En effet, par des combinaisons linéaires de schémas avec différents pas de temps, on peut annuler les termes de biais successifs dans (1.1.2). Ainsi, en combinant l'approche Monte-Carlo multi-pas classique avec l'extrapolation de Richardson Romberg, Lemaire et Pagès proposent d'estimer Y par

$$\hat{Y}_{ML2R} = \sum_{l=0}^{L}\frac{W_l}{M_l}\sum_{k=0}^{M_l} Z^{l,k},$$

où les $(W_l)_{0\le l\le L} \in \mathbb{R}^{L+1}$ sont des poids, les variables aléatoires $\left(Z^{l,k}\right)_{0\le l\le L, 1\le k\le M_l}$ sont indépendantes et vérifient (1.1.3), (1.1.4). On suppose de plus l'existence d'un développement de l'erreur faible pour le schéma numérique

$$\exists\alpha \in \mathbb{R}_+^*, \exists R \in \mathbb{N}^*, \exists c_1', \dots, c_R' \in \mathbb{R}, \forall l \in \mathbb{N}, \mathbb{E}\left[f\left(X_T^{2^l}\right)\right] - Y = \sum_{j=1}^{R} c_j' h_l^{\alpha j} + O\left(h_l^{\alpha(R+1)}\right). \quad (1.1.5)$$

Cette méthode peut être vue comme une version pondérée de la méthode de Monte-Carlo multi-pas classique. Ainsi, la complexité de la méthode de Lemaire et Pagès notée $\mathcal{C}_{ML2R}$, est définie de façon analogue à $\mathcal{C}_{MLMC}$. Lorsque l'ordre $\beta \leq 1$, la complexité de cette méthode est bien meilleure:

$$\begin{cases} \mathcal{C}^*_{ML2R} = O\left(\epsilon^{-2}\right) \text{ si } \beta > 1, \\ \mathcal{C}^*_{ML2R} = O\left(\epsilon^{-2}\log\left(\frac{1}{\epsilon}\right)\right) \text{ si } \beta = 1, \\ \mathcal{C}^*_{ML2R} = O\left(\epsilon^{-2}\exp\left(-\frac{\beta-1}{\sqrt{\alpha}}\sqrt{2\log(2)\log\left(\frac{1}{\epsilon}\right)}\right)\right) \text{ si } \beta < 1. \end{cases}$$

Lorsque $\beta > 1$, nous avons, comme dans le cas Monte-Carlo multi-pas classique, une complexité en $O\left(\epsilon^{-2}\right)$.

Pour atteindre cette complexité avec un schéma d'ordre fort 1/2, Giles et Szpruch [23] ont récemment proposé le schéma suivant

$$\begin{cases} X^{GS}_{t_0} = x_0, \\ X^{GS}_{t_{k+1}} = X^{GS}_{t_k} + b\left(X^{GS}_{t_k}\right)h + \sum_{j=1}^{d}\sigma^j\left(X^{GS}_{t_k}\right)\Delta W^j_{t_{k+1}} \\ \qquad + \frac{1}{2}\sum_{j=1}^{d}\sum_{m=1}^{d}\partial\sigma^j\sigma^m\left(X^{GS}_{t_k}\right)\left(\Delta W^j_{t_{k+1}}\Delta W^j_{t_{k+1}} - \mathbb{1}_{\{j=m\}}h\right). \end{cases}$$

Ce schéma correspond au schéma de Milstein sans les aires de Lévy. Il est donc parfaitement simulable, quelle que soit la dimension d du mouvement brownien, mais il est seulement d'ordre fort 1/2. Pour le choix de Z^l, pour $l \in \{1, \ldots, d\}$, Giles et Szpruch procèdent à un couplage dit antithétique:

$$Z^l_{GS} = \frac{1}{2}\left(f\left(X^{GS,2^l}_T\right) + f\left(\tilde{X}^{GS,2^l}_T\right)\right) - f\left(X^{GS,2^{l-1}}_T\right), \tag{1.1.6}$$

dans lequel $\tilde{X}^{GS,2^l}_T$ est obtenu à l'aide du même schéma numérique que $X^{GS,2^l}_T$, mais en transposant les paires d'accroissements browniens successifs. Lorsque les champs de vecteurs b et σ^j, pour $j \in \{1, \ldots, d\}$, sont de classe $\mathcal{C}^2$ avec dérivées bornées, cette opération permet d'obtenir le couplage à l'ordre fort 1 suivant

$$\exists C_{GS} \in \mathbb{R}^*_+, \forall l \in \mathbb{N}, \mathbb{E}\left[\left\|\frac{1}{2}\left(X^{GS,2^l}_T + \tilde{X}^{GS,2^l}_T\right) - X^{GS,2^{l-1}}_T\right\|^{2p}\right] \leq C_{GS}h_l^{2p}.$$

Pour une fonction f de classe $\mathcal{C}^2$ à dérivées bornées, on peut facilement obtenir l'estimation suivante à l'aide du développement de Taylor à l'ordre 2,

$$\mathbb{E}\left[\left(Z^l_{GS}\right)^{2p}\right] \leq c\left(\mathbb{E}\left[\left\|\frac{1}{2}\left(X^{GS,l}_T + \tilde{X}^{GS,l}_T\right) - X^{GS,l-1}_T\right\|^{2p}\right] + \mathbb{E}\left[\left\|X^{GS,l}_T - \tilde{X}^{GS,l}_T\right\|^{4p}\right]\right),$$

où c est une constante qui ne dépend que de p et de f. Cela permet donc à la variance de décroître à l'ordre $\beta = 2$ et ainsi d'atteindre la complexité $O\left(\epsilon^{-2}\right)$.

1.2 Schéma de Ninomiya et Victoir

Nous présentons le schéma sur une grille uniforme de pas de temps $h = T/N$. Nous commençons par réécrire l'équation différentielle stochastique (1.0.1). Au lieu de considérer l'équation écrite à l'aide de l'intégrale d'Itô, nous considérons l'intégrale de Stratonovich définie par

$$\int_0^t H_s \circ dW_s^j = \int_0^t H_s dW_s^j + \frac{1}{2}\left\langle H, W^j \right\rangle_t,$$

où $(H_s)_{0\leq s\leq T}$ est un processus d'Itô unidimensionnel adapté à filtration naturelle du mouvement brownien W^j. Ainsi, lorsque les champs de vecteurs browniens sont de classe $\mathcal{C}^1$, (1.0.1) se réécrit grâce à la formule d'Itô, de la façon suivante

$$dX_t = \sigma^0(X_t)dt + \sum_{j=1}^{d} \sigma^j(X_t) \circ dW_t^j,$$

où $\sigma^0 = b - \frac{1}{2}\sum_{j=1}^{d} \partial\sigma^j\sigma^j$ est le champ de vecteur que nous appellerons drift Stratonovich. On se donne ensuite une suite de variables aléatoires, notées $\eta = (\eta_k)_{k\geq 1}$, indépendantes et identiquement distribuées selon une loi de Rademacher définie par

$$\forall k \in \{1, \ldots, N\}, \mathbb{P}(\eta_k = -1) = \mathbb{P}(\eta_k = 1) = \frac{1}{2}.$$

On suppose de plus que la suite η est indépendante du mouvement brownien W. Le schéma de Ninomiya et Victoir est défini par

$$\begin{cases} X_{t_0}^{NV,\eta} = x_0, \\ X_{t_{k+1}}^{NV,\eta} = \mathbb{1}_{\{\eta_{k+1}=1\}} \exp\left(\frac{h}{2}\sigma^0\right) \exp\left(\Delta W_{t_{k+1}}^d \sigma^d\right) \ldots \exp\left(\Delta W_{t_{k+1}}^1 \sigma^1\right) \exp\left(\frac{h}{2}\sigma^0\right) X_{t_k}^{NV,\eta} \\ \qquad + \mathbb{1}_{\{\eta_{k+1}=-1\}} \exp\left(\frac{h}{2}\sigma^0\right) \exp\left(\Delta W_{t_{k+1}}^1 \sigma^1\right) \ldots \exp\left(\Delta W_{t_{k+1}}^d \sigma^d\right) \exp\left(\frac{h}{2}\sigma^0\right) X_{t_k}^{NV,\eta}, \end{cases}$$

où, pour V un champ de vecteur lipschitzien, $\exp(\theta V)\, y$ désigne le flot, à l'instant $\theta \in \mathbb{R}$, de l'équation différentielle ordinaire dans $\mathbb{R}^n$ suivante

$$\begin{cases} \dfrac{dx(t)}{dt} = V(x(t)) \\ x(0) = y. \end{cases}$$

Pour passer d'un instant de discrétisation t_k, avec $k \in \{0, \ldots, N-1\}$, à l'instant suivant t_{k+1}, on commence par tirer une variable de Rademacher η_{k+1}. La première étape consiste à résoudre, sur un demi-pas de temps, l'équation différentielle ordinaire portant sur le drift Stratonovich σ^0, avec comme condition initiale $X_{t_k}^{NV,\eta}$. Ensuite, selon le signe de la variable aléatoire η_{k+1}, on résout successivement les équations différentielles ordinaires portant sur les champs de vecteurs browniens σ^j, pour $j \in \{1, \ldots, d\}$, jusqu'à l'instant aléatoire $\Delta W_{t_{k+1}}^j$, avec comme condition initiale le résultat de l'étape précédente. Si η_{k+1} vaut 1, on intègre les champs de vecteurs browniens dans le sens croissant. Dans le cas contraire, on procède dans le sens décroissant. Enfin, la dernière étape consiste à résoudre, sur un demi-pas de temps, l'équation différentielle ordinaire portant sur le drift Stratonovich σ^0. Ainsi, le coût de simulation du schéma de Ninomiya et Victoir est de $(d+1)\,N$.

1.2.1 Résultats de convergence

Ninomiya et Victoir ont construit ce schéma pour obtenir un ordre faible $\alpha = 2$ [44]. Ils prouvent le résultat pour des champs de vecteur b et σ^j, $j \in \{1, \ldots, d\}$, réguliers et pour des fonctions f régulières. Ce résultat peut être facilement généralisé pour des fonctions f lipschitziennes sous une condition de non-dégénérescence de type Hörmander sur le générateur infinitésimal de la diffusion X (voir remarque 6 dans [44]). En fait, ce schéma peut être vu comme une formule de cubature sur l'espace de Wiener [39]. Un autre point de vue intéressant est développé par Alfonsi [4]. Il fait le lien entre la composition des $d+1$ équations différentielles ordinaires à résoudre et le découpage du générateur infinitésimal $\mathcal{L}$ de la diffusion X sous la forme

$$\mathcal{L} = \mathcal{L}_0 + \frac{1}{2}\sum_{j=1}^{d} \mathcal{L}_j^2,$$

avec

$$\mathcal{L}_0 u\,(t,x) = \left(b\,(x) - \frac{1}{2}\sum_{j=1}^{n} \partial\sigma^j\sigma^j\,(x) \right) . \ \nabla_x u\,(t,x) = \sigma^0\,(x) . \ \nabla_x u\,(t,x)$$

et pour $j \in \{1, \ldots, d\}$

$$\mathcal{L}_j u\,(t,x) = \sigma^j\,(x) . \ \nabla_x u\,(t,x) .$$

Il apparait alors que $\mathcal{L}_0$ est associé à la résolution, sur un pas de temps déterministe, de l'équation différentielle portant sur le champ de vecteur σ^0, alors que, pour $j \in \{1, \ldots, d\}$, l'opérateur $\frac{1}{2}\mathcal{L}_j^2$ est associé à la résolution, sur un pas de temps aléatoire correspondant à un accroissement brownien, de l'équation différentielle portant sur le champ de vecteur σ^j [4].

Notons que pour obtenir une estimation de l'erreur faible, il n'est pas nécessaire de simuler le schéma de Ninomiya et Victoir à l'aide d'accroissements browniens. Il suffit de remplacer la loi de l'accroissement brownien sur un pas de temps, donnée par $\sqrt{h}G$ où G est une variable aléatoire gaussienne standard, par $\sqrt{h}Z$ avec Z une variable aléatoire telle que,

$$\forall q \in \{1, \ldots 5\}, \mathbb{E}\left[G^q\right] = \mathbb{E}\left[Z^q\right].$$

Signalons aussi deux résultats récents de convergence faible du schéma de Ninomiya et Victoir. Le premier est fondé sur la théorie des rough paths, introduite par Lyons [40]. Bayer et Fritz [9] ont montré la convergence faible du schéma de Ninomiya et Victoir pour des champs de vecteur b et σ^j, $j \in \{1, \ldots, d\}$ réguliers et pour des fonctions f a-Höldérienne d'exposant $a \in (0, \frac{1}{2})$. En revanche, dans ce cas, ils n'ont pas exhibé d'ordre de convergence. Le second est basé sur une généralisation du calcul de Malliavin [41]. Bally et Rey [6] ont démontré la convergence en variation totale à l'ordre 2, sous une condition d'ellipticité et pour des fonctions f mesurables bornées.

En vue d'utiliser le schéma de Ninomiya et Victoir dans une méthode de Monte-Carlo multi-pas, nous avons étudié les propriétés de convergence forte du schéma. Dans ce cas, nous devons simuler le schéma de Ninomiya et Victoir à l'aide d'accroissements browniens. Pour effectuer l'analyse de l'erreur forte, la principale difficulté est d'interpoler de façon simple et efficace le schéma de Ninomiya et Victoir. En utilisant une interpolation astucieuse, et sous une hypothèse de type Lipschitz sur les champs de vecteur b, σ^j et $\partial\sigma^j$, pour $j \in \{1, \ldots, d\}$, nous montrons, dans le chapitre 2, une convergence forte d'ordre 1/2. Nous nous sommes ensuite intéressés à la convergence en loi stable de l'erreur renormalisée définie par $V^N = \sqrt{N}\left(X - X^{NV,\eta}\right)$. La notion de convergence en loi stable a été introduite par Rényi [50] et développée par Jacod [30] puis Jacod et Protter [31] et s'applique à l'erreur renormalisée du schéma d'Euler dans [35] et à celle du schéma de Milstein [55]. Sous des hypothèses de régularité qui seront précisées dans le chapitre 3, nous avons montré la convergence en loi stable suivante:

$$V^N \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} V,$$

où V est l'unique solution de l'équation différentielle stochastique affine

$$V_t = \sqrt{\frac{T}{2}} \sum_{j=1}^{d} \sum_{m=1}^{j-1} \int_0^t \left[\sigma^j, \sigma^m\right](X_s)\, dB_s^{j,m} + \int_0^t \partial b\,(X_s)\, V_s ds + \sum_{j=1}^{d} \int_0^t \partial\sigma^j\,(X_s)\, V_s dW_s^j,$$

avec $\left[\sigma^j, \sigma^m\right]$ le crochet de Lie entre les champs de vecteurs σ^j et σ^m défini par

$$\left[\sigma^j, \sigma^m\right] = \partial\sigma^m\sigma^j - \partial\sigma^j\sigma^m,$$

et $(B_t)_{0\le t\le T}$ est un mouvement brownien de dimension $d(d-1)/2$ indépendant de W. Ce résultat permet d'une part de s'assurer que l'ordre de convergence 1/2 est optimal. D'autre part, il peut être vu comme une première étape en vue de prouver un théorème de la limite centrale pour les estimateurs Monte-Carlo multi-pas. Pour cela, il faut analyser l'erreur en loi stable du schéma entre deux niveaux de discrétisation successifs. Ben Alaya et Kebaier [10] ont prouvé un tel résultat pour le schéma d'Euler. D'une part, on remarque que le processus limite V ne dépend pas de la suite η. D'autre part, lorsque les champs de vecteur brownien commutent, le processus V est nul. Cela suggère que dans ce cas, l'ordre de convergence est strictement supérieur à 1/2. Cela nous a conduits à analyser l'erreur forte dans le cas commutatif.

Lorsque les champs de vecteurs browniens commutent, l'intégration de ces champs ne dépend pas de l'ordre choisi. Le schéma se réécrit de la façon suivante

$$\begin{cases} X_{t_0}^{NV} = x_0, \\ X_{t_{k+1}}^{NV} = \exp\left(\frac{h}{2}\sigma^0\right)\exp\left(\Delta W_{t_{k+1}}^d \sigma^d\right)\dots\exp\left(\Delta W_{t_{k+1}}^1 \sigma^1\right)\exp\left(\frac{h}{2}\sigma^0\right) X_{t_k}^{NV}. \end{cases}$$

Nous montrons, dans le chapitre 3, que dans ce cas l'erreur forte est d'ordre 1. Dans le chapitre 4, nous nous intéressons ensuite à la convergence en loi stable de l'erreur renormalisée définie par $U^N = N\left(X - X^{NV}\right)$. Nous démontrons la convergence suivante

$$U^N \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} U,$$

où V est l'unique solution de l'équation différentielle stochastique affine

$$U_t = \frac{T}{2\sqrt{3}} \sum_{j=1}^{d} \int_0^t \left[\sigma^0, \sigma^j\right](X_s)\, d\tilde{B}_s^j + \int_0^t \partial b\,(X_s)\, U_s ds + \sum_{j=1}^{d} \int_0^t \partial\sigma^j\,(X_s)\, U_s dW_s^j,$$

avec $\left(\tilde{B}_t\right)_{0\le t\le T}$ un mouvement brownien de dimension d indépendant de W. Lorsque tous les champs de vecteur σ^j, pour $j \in \{0,\dots,d\}$, commutent, le processus limite U est nul. Notre résultat n'est pas surprenant puisque dans ce cas le schéma de Ninomiya et Victoir est exact [17] [18].

1.2.2 Application aux méthodes de Monte-Carlo multi-pas

Dans le chapitre 2, nous combinons l'idée de Debrabant et Rössler avec celle de Giles et Szpruch. Nous construisons un estimateur utilisant le couplage antithétique de Giles et Szpruch pour les niveaux de discrétisation $l \in \{0, \dots L-1\}$ et le schéma de Ninomiya et Victoir au niveau de discrétisation le plus fin L. Nous espérons ainsi profiter de l'erreur faible d'ordre 2 du schéma de Ninomiya et Victoir et ainsi diminuer le nombre de niveaux de discrétisation et donc le temps de calcul. Pour conserver la complexité $O\left(\epsilon^{-2}\right)$, nous couplons à l'ordre fort 1 les schémas de Ninomiya et Victoir et celui de Giles et Szpruch. En considérant $\bar{X}^{NV,\eta} = \frac{1}{2}\left(X^{NV,\eta} + X^{NV,-\eta}\right)$, nous montrons, dans le chapitre 2, le couplage à l'ordre fort 1 suivant

$$\exists C_{GS}^{NV} \in \mathbb{R}_+^*, \forall N \in \mathbb{N}^*,\ \mathbb{E}\left[\max_{0\leq k\leq N}\left\|\bar{X}_{t_k}^{NV,\eta} - X_{t_k}^{GS}\right\|^{2p}\middle|\eta\right] \leq C_{GS}^{NV} h^{2p}.$$

Ainsi, en considérant, pour $l \in \mathbb{N}^*$,

$$\begin{aligned}Z_{GS-NV}^{l} = \frac{1}{4}&\left(f\left(\tilde{X}_T^{NV,2^l,\eta}\right) + f\left(\tilde{X}_T^{NV,2^l,-\eta}\right) + f\left(X_T^{NV,2^l,\eta}\right) + f\left(X_T^{NV,2^l,-\eta}\right)\right)\\ &- f\left(X_T^{GS,2^{l-1}}\right),\end{aligned}$$

où $\tilde{X}^{NV,\eta}$ et $\tilde{X}^{NV,-\eta}$ sont les versions antithétiques obtenues à l'aide du même schéma que pour $X^{NV,\eta}$ et $X^{NV,-\eta}$ respectivement, mais en transposant les paires d'accroissements browniens successifs, nous conservons un ordre $\beta = 2$ de convergence de la variance de $Z^{GS-NV,l}$. De même pour

$$\begin{aligned}Z_{NV}^{l} = \frac{1}{4}&\left(f\left(\tilde{X}_T^{NV,2^l,\eta}\right) + f\left(\tilde{X}_T^{NV,2^l,-\eta}\right) + f\left(X_T^{NV,2^l,\eta}\right) + f\left(X_T^{NV,2^l,-\eta}\right)\right)\\ &- \frac{1}{2}\left(f\left(X_T^{NV,2^{l-1},\eta}\right) + f\left(X_T^{NV,2^{l-1},-\eta}\right)\right),\end{aligned}$$

pour $l \in \mathbb{N}^*$. Cela permet d'envisager la construction de nouveaux estimateurs Monte-Carlo multi-pas réalisant la complexité $O\left(\epsilon^{-2}\right)$.

- L'estimateur Monte-Carlo multi-pas utilisant le couplage antithétique de Giles et Szpruch pour les niveaux de discrétisation $l \in \{1, \dots L-1\}$ et un couplage entre le schéma de Giles et Szpruch et le schéma de Ninomiya et Victoir au niveau de discrétisation le plus fin L est défini par

  $$\hat{Y}_{MLMC}^{GS-NV} = \sum_{l=0}^{L-1}\frac{1}{M_l}\sum_{k=1}^{M_l} Z_{GS}^{l,k} + \frac{1}{M_L}\sum_{k=1}^{M_L} Z_{GS-NV}^{l,k},$$

  où la variable aléatoire $Z^{GS,l}$ est définie par (1.1.6), pour $l \in \{1, \dots, L\}$, et

  $$Z_{GS}^{0} = f\left(X_T^{GS,1}\right).$$

- L'estimateur Monte-Carlo multi-pas utilisant le couplage antithétique de Ninomiya et Victoir pour tous les niveaux de discrétisation $l \in \{1, \dots L\}$ est défini par

$$\hat{Y}^{NV}_{MLMC} = \sum_{l=0}^{L} \frac{1}{M} \sum_{k=1}^{M} Z^{l,k}_{NV},$$

  où Z^0_{NV} est définie par

$$Z^0_{NV} = f\left(X^{NV,1}_T\right).$$

- L'estimateur Monte-Carlo multi-pas combiné à l'extrapolation de Romberg (méthode de Lemaire et Pagès) utilisant le couplage antithétique de Ninomiya et Victoir pour tous les niveaux de discrétisation $l \in \{0, \dots L\}$ est défini par

$$\hat{Y}^{NV}_{ML2R} = \sum_{l=0}^{L} \frac{W_l}{M_l} \sum_{k=1}^{M_l} Z^{l,k}_{NV}.$$

Pour utiliser ce dernier, encore faut-il disposer d'un développement de l'erreur faible du type (1.1.5). Citons les récents travaux de Fujiwara [20] ainsi que Oshima, Teichmann, et Veluscek [45] qui vont dans ce sens.

En pratique, pour réaliser la complexité $O\left(\epsilon^{-2}\right)$, il faut que les paramètres des méthodes de Monte-Carlo multi-pas, L et $(M_l)_{0\leq l\leq L}$, soient optimaux. Nous présentons dans le chapitre 2, une procédure numérique pour le calcul de ces paramètres. Nous illustrons nos résultats sur des équations différentielles de dimension 2, dirigées par un mouvement brownien bidimensionnel. Nous analysons également le temps de calcul par niveau de discrétisation $l \in \{0, \dots, L\}$, noté τ_l. Plus précisément, τ_l correspond au temps de calcul de la quantité

$$\frac{1}{M^*_l} \sum_{k=1}^{M^*_l} Z^l_k,$$

où $M^*_l \in \mathbb{N}^*$ est le nombre de tirages optimal associé au niveau de discrétisation l. Nous montrons que pour $l \in \{1, \dots, L\}$,

$$\tau_l \propto 2^{-l(\beta-1)/2}$$

pour les estimateurs $\hat{Y}_{MLMC}$. Ainsi, nous montrons que l'introduction du couplage, entre le schéma de Giles et Szpruch et le schéma de Ninomiya et Victoir, au niveau de discrétisation le plus fin, est d'autant plus efficace que β est petit. En particulier, l'utilisation du schéma de Ninomiya et Victoir dans une méthode de Monte-Carlo multi-pas s'avère être très efficace pour une certaine classe de fonctions f non régulières, très largement utilisée en finance, pour laquelle les variables Z^{GS}, Z^{NV} et Z^{GS-NV} ont un ordre $\beta = 3/2$. Nous signalons également que les techniques de réduction de variance du premier niveau $l = 0$ rendent les estimateurs $\hat{Y}_{MLMC}$ plus performants. En effet, lorsque $\beta > 1$, le temps de calcul global de la méthode de Monte-Carlo multi-pas se concentre principalement sur le niveau $l = 0$. Réduire la variance sur ce niveau de discrétisation permet de diminuer le nombre de tirages $M^*_0 \in \mathbb{N}^*$ et ainsi diminuer le temps de calcul de la méthode.

La simulation du schéma de Ninomiya et Victoir repose sur la résolution de $(d+1)$ équations différentielles ordinaires sur chaque pas de temps de discrétisation. Il n'est pas rare que ces équations n'admettent pas de solution explicite. Le chapitre 5 de cette thèse est consacré à cette problématique. Pour conserver l'ordre faible $\alpha = 2$, Ninomiya et Ninomiya [43] ont montré qu'il suffit d'utiliser un schéma numérique où l'erreur de troncature est d'ordre 3 pour l'équation différentielle ordinaire portant sur le champ de vecteurs σ^0 et un schéma numérique où l'erreur de troncature est d'ordre 6 pour l'équation différentielle ordinaire portant sur les champs de vecteurs browniens σ^j, pour $j \in \{1, \ldots, d\}$. Plus précisément, l'erreur de troncature d'ordre $m \in \mathbb{R}_+^*$ d'une approximation, notée $\Psi(t, y)$, de l'équation différentielle ordinaire (1.2) est définie par

$$\|\exp(tV)\, y - \Psi(t, y)\| \leq c\,(1 + \|y\|)\, |t|^m \,,$$

où la constante $c \in \mathbb{R}_+^*$ ne dépend ni de l'instant $t \in \mathbb{R}$, ni de la condition initiale $y \in \mathbb{R}^n$. Nous montrons que cela suffit pour conserver un couplage à l'ordre fort 1, entre le schéma de Ninomiya et Victoir $X^{NV,\eta}$ et son approximation numérique $\hat{X}^{NV,\eta}$. On peut ainsi remplacer le schéma de Ninomiya et Victoir par son approximation numérique dans les estimateurs Monte-Carlo multi-pas tout en conservant la complexité $O\left(\epsilon^{-2}\right)$

La thèse que je présente s'organise de la façon suivante. Le deuxième chapitre traite des propriétés de convergence forte du schéma de Ninomiya et Victoir et de l'utilisation de ce schéma dans une méthode de Monte-Carlo multi-pas. Ce chapitre correspond à un article écrit avec mes directeurs de thèse Emmanuelle Clément et Benjamin Jourdain. Il a été publié dans la revue *Monte Carlo Methods and Applications.* Dans le troisième chapitre, nous nous intéressons à la convergence en loi du processus d'erreur renormalisé dans le cas général, ainsi qu'à la vitesse de convergence dans le cas commutatif, c'est-à-dire lorsque les champs de vecteurs browniens commutent. Ce chapitre correspond à un article écrit avec mes directeurs de thèse. Il a été soumis pour publication. Le quatrième chapitre est consarcé à l'étude du processus d'erreur renormalisé dans le cas commutatif. Ce chapitre a fait l'objet d'une prépublication. Enfin, le dernier chapitre traite de l'utilisation de schémas numériques pour les EDOs dans l'implémentation du schéma de Ninomiya et Victoir.

Chapter 2

Ninomiya-Victoir scheme: strong convergence, antithetic version and application to multilevel estimators

This chapter corresponds to an article written with Émmanuelle Clément and Benjamin Jourdain [1]. It has been published in the journal Monte Carlo Methods and Applications.

Abstract. In this paper, we are interested in the strong convergence properties of the Ninomiya-Victoir scheme which is known to exhibit weak convergence with order 2. We prove strong convergence with order 1/2. This study is aimed at analysing the use of this scheme either at each level or only at the finest level of a multilevel Monte Carlo estimator: indeed, the variance of a multilevel Monte Carlo estimator is related to the strong error between the two schemes used on the coarse and fine grids at each level. Recently, Giles and Szpruch proposed in [23] a scheme permitting to construct a multilevel Monte Carlo estimator achieving the optimal complexity $O\left(\epsilon^{-2}\right)$ for the precision ϵ. In the same spirit, we propose a modified Ninomiya-Victoir scheme, which may be strongly coupled with order 1 to the Giles-Szpruch scheme at the finest level of a multilevel Monte Carlo estimator. Numerical experiments show that this choice improves the efficiency, since the order 2 of weak convergence of the Ninomiya-Victoir scheme permits to reduce the number of discretization levels.

2.1 Introduction

This paper is dedicated to the computation of $Y = \mathbb{E}\left[f\left(X_T\right)\right]$, where $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a payoff function and X_T is the solution, at time $T \in \mathbb{R}_+^*$, to a multi-dimensional stochastic differential equation of the form

$$\begin{cases} dX_t = b(X_t)dt + \sum_{j=1}^{d} \sigma^j(X_t)dW_t^j, \ t \in [0,T] \\ X_0 = x. \end{cases} \tag{2.1.1}$$

Here, $x \in \mathbb{R}^n$ is the initial condition, $W = \left(W^1, \ldots, W^d\right)$ is a $d-$dimensional standard Brownian motion, $b : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is the drift coefficient and $\sigma^j : \mathbb{R}^n \longrightarrow \mathbb{R}^n, j \in \{1, \ldots, d\}$, are the diffusion coefficients.

The standard Monte Carlo method consists in estimating $\mathbb{E}\left[f\left(X_T\right)\right]$ by discretizing the stochastic differential equation with $N \in \mathbb{N}^*$ steps and approximating the expectation using $M \in \mathbb{N}^*$ independent path simulations. To be clear, the crude Monte Carlo estimator is given by

$$\hat{Y}_{CMC} = \frac{1}{M} \sum_{k=1}^{M} f\left(X_T^{N,k}\right),$$

where $X^{N,k}$ are independent copies of a numerical scheme X^N with time step T/N. Under some regularity assumptions on the coefficients of the SDE and for a smooth payoff, it is well known that to ensure a root mean-square-error ϵ, the computational cost of this method is $O\left(\epsilon^{-\left(2+\frac{1}{\alpha}\right)}\right)$, where α is the order of weak convergence of the numerical scheme (see theorem 1 in [19]). In [44], Ninomiya and Victoir proposed a numerical scheme, achieving $\alpha = 2$, which reduces the computational complexity compared to the Euler scheme, for which $\alpha = 1$. In the optimal complexity $O\left(\epsilon^{-\left(2+\frac{1}{\alpha}\right)}\right)$, the term $1/\alpha$ is due to the bias $\mathbb{E}\left[f\left(X_T\right)\right] - \mathbb{E}\left[f\left(X_T^N\right)\right]$.

To remove this term, Giles introduced in [22] a multilevel Monte Carlo estimator permitting telescopic cancellation of the bias. The multilevel Monte Carlo estimator is built as follows

$$\hat{Y} = \frac{1}{M_0} \sum_{k=1}^{M_0} f\left(X_T^{1,0,k}\right) + \sum_{l=1}^{L} \frac{1}{M_l} \sum_{k=1}^{M_l} \left(f\left(X_T^{2^l,l,k}\right) - f\left(X_T^{2^{l-1},l,k}\right)\right),$$

where $L \in \mathbb{N}^*$ is the last and finest level of discretization with time-step $T/2^L$, $(M_l)_{0 \le l \le L} \in (\mathbb{N}^*)^{L+1}$ is the vector of sample sizes at each level. Moreover, for all $l \in \{1, \ldots, L\}$, the two numerical schemes $X_T^{2^l,l}$ and $X_T^{2^{l-1},l}$ are simulated with the same Brownian motion. For each discretization level $l \in \{0, \ldots, L\}$, M_l independent and identically distributed path simulations independent from the other levels are used. The optimal complexity of this method is driven by the order β of convergence to zero of the variance $\mathbb{V}\left(f\left(X_T^{2^l,l}\right) - f\left(X_T^{2^{l-1},l}\right)\right)$, which is related to the strong convergence order γ of the scheme. For a Lipschitz payoff f, using the strong convergence properties of the scheme in the estimation of the variance, one gets $\beta \ge 2\gamma$. For $\beta > 1$, the optimal complexity is $O\left(\epsilon^{-2}\right)$. This complexity is the same as in a simple Monte Carlo method with independent and identically distributed unbiased random variables. The condition $\beta > 1$ is satisfied by the Milstein scheme for which $\gamma = 1$. Unfortunately, to simulate the Milstein scheme, one needs, in general, to simulate Lévy areas for which there is no known efficient method when the dimension of the Brownian motion d is larger than 2. Unless the diffusion coefficients $\sigma^j, j \in \{1, \ldots, d\}$, are constant, the strong order of the Euler scheme is $\gamma = 1/2$, which leads to $\beta = 1$ and to the optimal complexity $O\left(\epsilon^{-2}\left(\log\left(\frac{1}{\epsilon}\right)\right)^2\right)$.

Recently, two approaches have been developed to improve the case $\gamma = 1/2$. In [23], Giles and Szpruch introduced a modified Milstein scheme, with the Lévy areas set to zero, and its antithetic version based on the swapping of each successive pair of Brownian increments in the scheme. Regarding the multilevel Monte Carlo estimator, at each discretization level $l \in \{1, \ldots, L\}$, on the finest grid, instead of using a simple scheme, Giles and Szpruch employed the arithmetic average of the scheme and its antithetic version as follows

$$\hat{Y}_{MLMC} = \frac{1}{M_0}\sum_{k=1}^{M_0} f\left(X_T^{1,0,k}\right) + \sum_{l=1}^{L}\frac{1}{M_l}\sum_{k=1}^{M_l}\left(\frac{1}{2}\left(f\left(\tilde{X}_T^{2^l,l,k}\right)+f\left(X_T^{2^l,l,k}\right)\right) - f\left(X_T^{2^{l-1},l,k}\right)\right),$$

where $\tilde{X}^{2^l}$ denotes the antithetic version of the Giles-Szpruch scheme with time-step $T/2^l$. Under some regularity assumptions on the coefficients of the SDE and for a smooth payoff, Giles and Szpruch showed that despites γ is equal to $1/2$, β is equal to 2 which leads to an optimal complexity $O\left(\epsilon^{-2}\right)$. The second approach called multilevel Richardson-Romberg method and investigated by Lemaire and Pagès in [38], fully takes advantage of the existence of a weak error expansion while keeping the multilevel Monte Carlo estimator properties. The multilevel Richardson-Romberg estimator is a weighted version of the multilevel Monte Carlo method which integrates the multi-step Richardson-Romberg extrapolation developed by Pagès in [46]. Lemaire and Pagès obtained an optimal complexity $O\left(\epsilon^{-2}\log\left(\frac{1}{\epsilon}\right)\right)$ when $\beta = 1$ which improves the standard multilevel Monte Carlo method. When $\beta > 1$, the optimal complexity $O\left(\epsilon^{-2}\right)$ is preserved.

In this paper, we propose to use the Ninomiya-Victoir scheme, which is known to exhibit weak convergence with order 2, on the finest grid at the last level L of a multilevel Monte Carlo estimator. This idea is inspired by Debrabant and Rössler [16] who suggest to use a scheme with high order of weak convergence on the finest grid at the finest level L of the multilevel Monte Carlo method. By this way, Debrabant and Rössler reduce the constant in the computational complexity by decreasing the number of discretization levels. In section 2, to derive the strong convergence order of the Ninomiya-Victoir scheme, we provide a suitable interpolation between time grid points. Then we prove strong convergence with order $\gamma = 1/2$ under some regularity assumptions on the coefficients of the SDE. In section 3, we propose a modified Ninomiya-Victoir scheme, which may be strongly coupled with order 1 to the Giles-Szpruch scheme. This result allows us to derive an antithetic version of the Ninomiya-Victoir scheme and combine the ideas of Giles-Szpruch and Debrabant-Rössler by building the multilevel Monte Carlo estimator with the Giles-Szpruch scheme from level 0 to level $L-1$ and the coupling between the Ninomiya-Victoir scheme and the Giles-Szpruch scheme at the last level L. The efficiency of this estimator is confirmed in section 4, where we present and comment, in details, numerical experiments carried out on the Clark-Cameron SDE and Heston SDE as in [23].

2.2 Strong convergence of the Ninomiya-Victoir scheme

We begin this section by introducing some notations which will be used throughout this paper. To discretize (2.1.1) we consider a uniform grid with time step $h = T/N$ where $N \in \mathbb{N}^*$ and we denote:

- $(t_k = kh)_{0\leq k\leq N}$ is the subdivision of $[0,T]$ with equal time step h,
- $\hat{\tau}_s$ is the last time discretization before $s \in [0,T]$, ie $\hat{\tau}_s = t_k$ if $s \in (t_k, t_{k+1}]$, and for $s = t_0 = 0$, we set $\hat{\tau}_0 = t_0 = 0$,
- $\check{\tau}_s$ is the first time discretization after $s \in [0,T]$, ie $\check{\tau}_s = t_{k+1}$ if $s \in (t_k, t_{k+1}]$, and for $s = t_0 = 0$, we set $\check{\tau}_0 = t_0 = 0$,

- for all $j \in \{1, \dots, d\}$, for all $s \in (t_k, t_{k+1}]$, $\Delta W^j_s = W^j_s - W^j_{t_k}$,
- for all $s \in (t_k, t_{k+1}]$, $\Delta s = s - t_k$,
- $\eta = (\eta_1, \dots, \eta_N)$ is a sequence of independent, identically distributed Rademacher random variables independent of W,
- by a slight abuse of notation, we set $\eta_s = \eta_{k+1}$ if $s \in (t_k, t_{k+1}]$,
- for all $x \in \mathbb{R}_+$, $\lceil x \rceil$ is the unique $n \in \mathbb{N}$ satisfying $n - 1 < x \leq n$,
- for all $x \in \mathbb{R}_+$, $\lfloor x \rfloor$ is the unique $n \in \mathbb{N}$ satisfying $n \leq x < n + 1$.

Let $V : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be Lipschitz continuous and consider the ordinary differential equation in $\mathbb{R}^n$:

$$\begin{cases} \dfrac{dx(t)}{dt} = V\left(x(t)\right) \\ x(0) = x_0. \end{cases} \tag{2.2.1}$$

The solution of (2.2.1) at time t, $t \in \mathbb{R}$ is denoted by

$$x(t) = \exp(tV)x_0,$$

and the integral form of (2.2.1) is given by

$$x(t) = \exp(tV)x_0 = x_0 + \int_0^t V(x(s))ds = x + \int_0^t V\left(\exp(sV)x_0\right) ds.$$

We recall that in (2.1.1), each coordinate $i \in \{1, \dots, n\}$ evolves according to the following stochastic differential equation

$$dX^i_t = b^i(X_t)dt + \sum_{j=1}^d \sigma^{ij}(X_t)dW^j_t.$$

Then, assuming $\mathcal{C}^1$ regularity for the diffusion coefficients, one can write (2.1.1) in Stratonovich form

$$\begin{cases} dX_t = \sigma^0(X_t)dt + \displaystyle\sum_{j=1}^d \sigma^j(X_t) \circ dW^j_t \\ X_0 = x, \end{cases} \tag{2.2.2}$$

where $\sigma^0 = b - \frac{1}{2}\sum_{j=1}^d \partial\sigma^j\sigma^j$ and $\partial\sigma^j$ is the Jacobian matrix of σ^j defined as follows

$$\partial\sigma^j = \left(\partial_{x_k}\sigma^{ij}\right)_{1 \leq i,k \leq n}.$$

Now, we present the Ninomiya-Victoir scheme introduced in [44].

- Starting point: $X^{NV,\eta}_{t_0} = x$.
- For $k \in \{0 \dots, N-1\}$:
  if $\eta_{k+1} = 1$,

$$X^{NV,\eta}_{t_{k+1}} = \exp\left(\frac{h}{2}\sigma^0\right)\exp\left(\Delta W^d_{t_{k+1}}\sigma^d\right)\dots\exp\left(\Delta W^1_{t_{k+1}}\sigma^1\right)\exp\left(\frac{h}{2}\sigma^0\right)X^{NV,\eta}_{t_k}, \tag{2.2.3}$$

and if $\eta_{k+1} = -1$,

$$X_{t_{k+1}}^{NV,\eta} = \exp\left(\frac{h}{2}\sigma^0\right)\exp\left(\Delta W_{t_{k+1}}^1\sigma^1\right)\dots\exp\left(\Delta W_{t_{k+1}}^d\sigma^d\right)\exp\left(\frac{h}{2}\sigma^0\right)X_{t_k}^{NV,\eta}. \tag{2.2.4}$$

The Stratonovich form is preferred when we use the Ninomiya-Victoir scheme since the vector field σ^0, which corresponds to the Stratonovich drift appears in the definition of the scheme. Moreover, using Itô's formula, one has: for all $t, s \in \mathbb{R}_+, s \leq t$,

$$\exp\left(\left(W_t^j - W_s^j\right)V\right)y = y + \int_s^t V\left(\exp\left(\left(W_u^j - W_s^j\right)V\right)y\right)\circ dW_u^j.$$

Then, rewriting (2.2.3) and (2.2.4), one obtains

$$X_{t_{k+1}}^{NV,\eta} = X_{t_k}^{NV,\eta} + \sum_{j=1}^d \int_{t_k}^{t_{k+1}} \sigma^j\left(\bar{X}_s^{j,\eta}\right)\circ dW_s^j + \int_{t_k}^{t_{k+1}} \frac{1}{2}\left(\sigma^0\left(\bar{X}_s^{0,\eta}\right) + \sigma^0\left(\bar{X}_s^{d+1,\eta}\right)\right)ds, \tag{2.2.5}$$

where, for $s \in (t_k, t_{k+1}]$,

$$\bar{X}_s^{0,\eta} = \exp\left(\frac{\Delta s}{2}\sigma^0\right)\left(X_{t_k}^{NV,\eta}\mathbb{1}_{\{\eta_{k+1}=1\}} + \bar{X}_{t_{k+1}}^{1,\eta}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right),$$

for $s \in (t_k, t_{k+1}]$, $j \in \{1, \dots, d\}$,

$$\bar{X}_s^{j,\eta} = \exp\left(\Delta W_s^j\sigma^j\right)\left(\bar{X}_{t_{k+1}}^{j-1,\eta}\mathbb{1}_{\{\eta_{k+1}=1\}} + \bar{X}_{t_{k+1}}^{j+1,\eta}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right), \tag{2.2.6}$$

for $s \in (t_k, t_{k+1}]$,

$$\bar{X}_s^{d+1,\eta} = \exp\left(\frac{\Delta s}{2}\sigma^0\right)\left(\bar{X}_{t_{k+1}}^{d,\eta}\mathbb{1}_{\{\eta_{k+1}=1\}} + X_{t_k}^{NV,\eta}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right).$$

Denoting $\bar{X}_{t_{k+1}}^{-1,\eta} = \bar{X}_{t_{k+1}}^{d+2,\eta} = X_{t_k}^{NV,\eta}$, one gets an expression similar to (2.2.6) for $j \in \{0, d+1\}$ and $s \in (t_k, t_{k+1}]$

$$\bar{X}_s^{j,\eta} = \exp\left(\frac{\Delta s}{2}\sigma^0\right)\left(\bar{X}_{t_{k+1}}^{j-1,\eta}\mathbb{1}_{\{\eta_{k+1}=1\}} + \bar{X}_{t_{k+1}}^{j+1,\eta}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right).$$

Then, one can observe that the Ninomiya-Victoir scheme is obtained by replacing the exact solution X by one of the intermediate processes $\bar{X}^{j,\eta}$ in the Stratonovich formulation (2.2.2) of the SDE (2.1.1).

Remark 2.2.1 *The stochastic processes $\left(\bar{X}_t^{j,\eta}\right)_{0\leq t\leq T}$, for $j \in \{1, \dots d+1\}$, are not adapted to the natural filtration $\mathcal{F}_t = \sigma\left(W_s, s \leq t\right)$ of the Brownian motion. To get around this problem, we work with the following filtration*

$$\tilde{\mathcal{F}}_t^j = \sigma\left(\eta, W_s^j, s \leq t\right)\bigvee\left(\bigvee_{k\neq j}\sigma\left(W_s^k, s \leq T\right)\right).$$

Then, for $j \in \{1, \dots, d\}$, by independence, W^j is an $\tilde{\mathcal{F}}^j$ Brownian motion, and $\bar{X}^{j,\eta}$ is adapted to the filtration $\tilde{\mathcal{F}}^j$. This ensures that each stochastic integral is well defined.

In order to study the strong convergence, we have to build an interpolated scheme. Let $\left(X_t^{NV,\eta}\right)_{0\le t\le T}$ be the following Itô process

$$\begin{cases} dX_t^{NV,\eta} = \sum_{j=1}^{d} \sigma^j(\bar{X}_t^{j,\eta}) \circ dW_t^j + \frac{1}{2}\left(\sigma^0\left(\bar{X}_t^{0,\eta}\right) + \sigma^0\left(\bar{X}_t^{d+1,\eta}\right)\right) dt \\ X_0^{NV,\eta} = x. \end{cases}$$

Using (2.2.5) and forward induction, one can show that $\left(X_t^{NV,\eta}\right)_{0\le t\le T}$ is an interpolation of the Ninomiya-Victoir scheme $\left(X_{t_k}^{NV,\eta}\right)_{0\le k\le N}$. The Itô decomposition of $\left(X_t^{NV,\eta}\right)_{0\le t\le T}$ is given by

$$\begin{cases} dX_t^{NV,\eta} = \sum_{j=1}^{d} \sigma^j(\bar{X}_t^{j,\eta})dW_t^j + \frac{1}{2}\sum_{j=1}^{d} \partial\sigma^j\sigma^j\left(\bar{X}_t^{j,\eta}\right)dt + \frac{1}{2}\left(\sigma^0\left(\bar{X}_t^{0,\eta}\right) + \sigma^0\left(\bar{X}_t^{d+1,\eta}\right)\right) dt \\ X_0^{NV,\eta} = x. \end{cases} \tag{2.2.7}$$

Remark 2.2.2 *A natural and adapted interpolation for this scheme could be*

$$h_{\eta_t}\left(\frac{\Delta t}{2}, \Delta W_t, \frac{\Delta t}{2}; X_{\hat{\tau}_t}^{NV,\eta}\right),$$

where

$$h_{-1}\left(t_0,\ldots,t_{d+1};x\right) = \exp\left(t_0\sigma^0\right)\exp\left(t_1\sigma^1\right)\ldots\exp\left(t_d\sigma^d\right)\exp\left(t_{d+1}\sigma^0\right)x,$$

and

$$h_1\left(t_0,\ldots,t_{d+1};x\right) = \exp\left(t_0\sigma^0\right)\exp\left(t_d\sigma^d\right)\ldots\exp\left(t_1\sigma^1\right)\exp\left(t_{d+1}\sigma^0\right)x.$$

In both cases $\Delta W_t = \left(\Delta W_t^1,\ldots,\Delta W_t^d\right)$. In order to obtain the Itô decomposition of $X^{NV,\eta}$, we have to apply the Itô formula. To do so, we have to compute the derivatives of h_η. In the general case, the computation of derivatives of this function is quite complicated. That is why we will not focus on this interpolation.

2.2.1 Strong convergence

We recall that for all $j \in \{1,\ldots,d\}$, $\sigma^j \in \mathcal{C}^1\left(\mathbb{R}^n,\mathbb{R}^n\right)$, and we assume that the vector fields σ^0, σ^j and $\partial\sigma^j\sigma^j$, for all $j \in \{1,\ldots,d\}$, are Lipschitz continuous. Obviously, b is also Lipschitz continuous, since $b = \sigma^0 + \frac{1}{2}\sum_{j=1}^{d}\partial\sigma^j\sigma^j$. Let $L \in \mathbb{R}_+^*$ denote their common Lipschitz constant:

$$\left\|\sigma^j(x) - \sigma^j(y)\right\| \le L\left\|x-y\right\|, \text{ for all } j \in \{0,\ldots,d\} \text{ and all } x,y \in \mathbb{R}^n,$$
$$\left\|\partial\sigma^j\sigma^j(x) - \partial\sigma^j\sigma^j(y)\right\| \le L\left\|x-y\right\|, \text{ for all } j \in \{1,\ldots,d\} \text{ and all } x,y \in \mathbb{R}^n,$$

where the Euclidean norm is denoted by $\|.\|$.

Theorem 2.2.3 *Let $p \in [1,+\infty)$. Under the previous Lipschitz assumption, there exists a deterministic constant $C_{NV} \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$,*

$$\mathbb{E}\left[\sup_{t\le T}\left\|X_t - X_t^{NV,\eta}\right\|^{2p}\middle|\eta\right] \le C_{NV}\left(1+\|x\|^{2p}\right)h^p.$$

Of course, this result implies that for all $N \in \mathbb{N}^*$,

$$\mathbb{E}\left[\sup_{t\leq T}\left\|X_t - X_t^{NV,\eta}\right\|^{2p}\right] \leq C_{NV}\left(1+\|x\|^{2p}\right)h^p.$$

Obviously, $\left(X_t^{NV,\eta}\right)_{0\leq t\leq T}$ and h depend on N, but in order to keep the notations simple, the dependence on N is not made explicit. The following proposition will be used to prove the theorem.

Proposition 2.2.4 *Let $p \geq 1$ and let $Y = (Y_t)_{0\leq t\leq h}$ be the solution of the following n-dimensional SDE, driven by a d-dimensional Brownian motion, on the time interval $[0,h]$:*

$$\begin{cases} dY_s = \alpha(Y_s)ds + \beta(Y_s)dW_s \\ Y_0 \textit{ independent of } (W_t)_{t\in[0,h]} \textit{ such that } \mathbb{E}\left[\|Y_0\|^{2p}\right] < +\infty. \end{cases}$$

Assume that α and β are Lipschitz continuous functions, then there exists a constant $C_0 \in \mathbb{R}_+^*$ such that for all $t,s \in [0,h], s\leq t$,

$$\mathbb{E}\left[1+\|Y_t\|^{2p}\right] \leq \mathbb{E}\left[1+\|Y_0\|^{2p}\right]\exp\left(C_0 h\right), \tag{2.2.8}$$

$$\mathbb{E}\left[\|Y_t - Y_s\|^{2p}\right] \leq C_0\left(1+\mathbb{E}\left[\|Y_0\|^{2p}\right]\right)(t-s)^p, \tag{2.2.9}$$

and if $\beta = 0$, we have a better result:

$$\mathbb{E}\left[\|Y_t - Y_s\|^{2p}\right] \leq C_0\left(1+\mathbb{E}\left[\|Y_0\|^{2p}\right]\right)(t-s)^{2p}. \tag{2.2.10}$$

The constant C_0 only depends on $\|\alpha(0)\|, \|\beta(0)\|$, T, p, and the Lipschitz constants of the functions α and β.

All these results are well known (see [51] for example).

2.2.2 Intermediate results

By using the previous proposition, one can show that the scheme has uniformly bounded moments.

Lemma 2.2.5 *For all $p \geq 1$, there exists a constant $C_1 \in \mathbb{R}_+^*$ such that for all $t \in [0,T], N \in \mathbb{N}^*$ and all $j \in \{0,\ldots,d+1\}$,*

$$\mathbb{E}\left[1+\left\|\bar{X}_t^{j,\eta}\right\|^{2p}\middle|\eta\right] \leq \exp(C_1\check{\tau}_t)\left(1+\|x\|^{2p}\right).$$

Proof : Let $p\geq 1$ and $t \in [0,T]$. Then there exists an integer $k \in \{0,\ldots,N-1\}$ such that $t_k < t \leq t_{k+1}$. For $j=0$, $\left(\bar{X}_s^{0,\eta}\right)_{t_k<s\leq t_{k+1}}$ is the solution of the following ODE

$$\begin{cases} dZ_s = \dfrac{1}{2}\sigma^0(Z_s)ds \\ \\ Z_{t_k} = X_{t_k}^{NV,\eta}\mathbb{1}_{\{\eta_{k+1}=1\}} + \bar{X}_{t_{k+1}}^{1,\eta}\mathbb{1}_{\{\eta_{k+1}=-1\}}. \end{cases}$$

The independence between η and W combined with (2.2.8) ensures that:

$$\begin{aligned}\mathbb{E}\left[1+\left\|\bar{X}_t^{0,\eta}\right\|^{2p}\middle|\eta\right] &\leq \mathbb{E}\left[1+\left\|X_{t_k}^{NV,\eta}\mathbb{1}_{\{\eta_{k+1}=1\}}+\bar{X}_{t_{k+1}}^{1,\eta}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right\|^{2p}\middle|\eta\right]\exp\left(\frac{1}{2}C_0h\right)\\ &= \left(\mathbb{1}_{\{\eta_{k+1}=1\}}\mathbb{E}\left[1+\left\|X_{t_k}^{NV,\eta}\right\|^{2p}\middle|\eta\right]\right. \\ &\qquad \left.+\mathbb{1}_{\{\eta_{k+1}=-1\}}\mathbb{E}\left[1+\left\|\bar{X}_{t_{k+1}}^{1,\eta}\right\|^{2p}\middle|\eta\right]\right)\exp\left(\frac{1}{2}C_0h\right).\end{aligned} \tag{2.2.11}$$

Similarly, for $1\leq j\leq d$: $\left(\bar{X}_s^{j,\eta}\right)_{t_k<s\leq t_{k+1}}$ is the solution of the following SDE:

$$\begin{cases} dZ_s = \dfrac{1}{2}\partial\sigma^j\sigma^j\left(Z_s\right)ds+\sigma^j(Z_s)dW_s^j\\ Z_{t_k} = \bar{X}_{t_{k+1}}^{j-1,\eta}\mathbb{1}_{\{\eta_{k+1}=1\}}+\bar{X}_{t_{k+1}}^{j+1,\eta}\mathbb{1}_{\{\eta_{k+1}=-1\}}.\end{cases}$$

Using the same argument, one gets

$$\begin{aligned}\mathbb{E}\left[1+\left\|\bar{X}_t^{j,\eta}\right\|^{2p}\middle|\eta\right] \leq &\left(\mathbb{1}_{\{\eta_{k+1}=1\}}\mathbb{E}\left[1+\left\|\bar{X}_{t_{k+1}}^{j-1,\eta}\right\|^{2p}\middle|\eta\right]\right.\\ &\left.+\mathbb{1}_{\{\eta_{k+1}=-1\}}\mathbb{E}\left[1+\left\|\bar{X}_{t_{k+1}}^{j+1,\eta}\right\|^{2p}\middle|\eta\right]\right)\exp\left(C_0h\right).\end{aligned}$$

Obviously, for $j=d+1$, one has a similar result

$$\begin{aligned}\mathbb{E}\left[1+\left\|\bar{X}_t^{j,\eta}\right\|^{2p}\middle|\eta\right] \leq &\left(\mathbb{1}_{\{\eta_{k+1}=1\}}\mathbb{E}\left[1+\left\|\bar{X}_{t_{k+1}}^{d,\eta}\right\|^{2p}\middle|\eta\right]\right.\\ &\left.+\mathbb{1}_{\{\eta_{k+1}=-1\}}\mathbb{E}\left[1+\left\|X_{t_k}^{NV,\eta}\right\|^{2p}\middle|\eta\right]\right)\exp\left(\frac{1}{2}C_0h\right).\end{aligned} \tag{2.2.12}$$

The global Lipschitz constant L is the same for all vector fields, therefore, the same constant C_0 is involved in the three inequalities. In both ODEs, the vector field σ^0 is multiplied by $1/2$, it is equivalent to integrate the equation until $h/2$ and simply remove the multiplicative factor $1/2$. That is why one gets a factor $1/2$ in both inequalities (2.2.11) and (2.2.12). Since for all $k\in\{0,\dots,N\}$,

$$X_{t_k}^{NV,\eta}=\mathbb{1}_{\{\eta_k=1\}}\bar{X}_{t_k}^{d+1,\eta}+\mathbb{1}_{\{\eta_k=-1\}}\bar{X}_{t_k}^{0,\eta},$$

one can use forward induction on k combined with forward induction (respectively backward) on $j\in\{0,\dots,d+1\}$ if $\eta_{k+1}=1$ (respectively $\eta_{k+1}=-1$) to get

$$\mathbb{E}\left[1+\left\|\bar{X}_t^{j,\eta}\right\|^{2p}\middle|\eta\right]\leq\exp\left(C_1t_{k+1}\right)\left(1+\|x\|^{2p}\right),$$

where $C_1=(d+1)\,C_0$. ∎

The following lemma is a direct application of Proposition 2.2.4, together with Lemma 2.2.5.

Lemma 2.2.6 *For all $p \geq 1$, there exists a constant $C_2 \in \mathbb{R}_+^*$ such that for all $t \in [0,T], N \in \mathbb{N}^*$, and all $j \in \{1,\dots,d\}$,*

$$\mathbb{E}\left[\left\|\bar{X}_t^{j,\eta} - \bar{X}_{\check{\tau}_t}^{j-1,\eta}\mathbb{1}_{\{\eta_t=1\}} - \bar{X}_{\check{\tau}_t}^{j+1,\eta}\mathbb{1}_{\{\eta_t=-1\}}\right\|^{2p}\middle|\eta\right] \leq C_2\left(1+\|x\|^{2p}\right)h^p,$$

and for all $j \in \{0, d+1\}$,

$$\mathbb{E}\left[\left\|\bar{X}_t^{j,\eta} - \bar{X}_{\check{\tau}_t}^{j-1,\eta}\mathbb{1}_{\{\eta_t=1\}} - \bar{X}_{\check{\tau}_t}^{j+1,\eta}\mathbb{1}_{\{\eta_t=-1\}}\right\|^{2p}\middle|\eta\right] \leq C_2\left(1+\|x\|^{2p}\right)h^{2p},$$

where by convention $\bar{X}_{\check{\tau}_t}^{-1,\eta} = \bar{X}_{\check{\tau}_t}^{d+2,\eta} = X_{\hat{\tau}_t}^{NV,\eta}$.

Proof : Let $p \geq 1$, $t \in [0,T]$ and $j \in \{1,\dots,d\}$. Thanks to (2.2.9) in Proposition 2.2.4 we have

$$\begin{aligned}
E_t^j &:= \mathbb{E}\left[\left\|\bar{X}_t^{j,\eta} - \bar{X}_{\check{\tau}_t}^{j-1,\eta}\mathbb{1}_{\{\eta_t=1\}} - \bar{X}_{\check{\tau}_t}^{j+1,\eta}\mathbb{1}_{\{\eta_t=-1\}}\right\|^{2p}\middle|\eta\right] \\
&\leq C_0\left(1+\mathbb{E}\left[\left\|\bar{X}_{\check{\tau}_t}^{j-1,\eta}\mathbb{1}_{\{\eta_t=1\}} - \bar{X}_{\check{\tau}_t}^{j+1,\eta}\mathbb{1}_{\{\eta_t=-1\}}\right\|^{2p}\middle|\eta\right]\right)h^p.
\end{aligned}$$

Since

$$\begin{aligned}
1+\mathbb{E}\left[\left\|\bar{X}_{\check{\tau}_t}^{j-1,\eta}\mathbb{1}_{\{\eta_t=1\}} - \bar{X}_{\check{\tau}_t}^{j+1,\eta}\mathbb{1}_{\{\eta_t=-1\}}\right\|^{2p}\middle|\eta\right] &= \mathbb{1}_{\{\eta_t=1\}}\mathbb{E}\left[1+\left\|\bar{X}_{\check{\tau}_t}^{j-1,\eta}\right\|^{2p}\middle|\eta\right] \\
&\quad + \mathbb{1}_{\{\eta_t=-1\}}\mathbb{E}\left[1+\left\|\bar{X}_{\check{\tau}_t}^{j+1,\eta}\right\|^{2p}\middle|\eta\right],
\end{aligned}$$

combining this estimation with lemma 2.2.5 we get

$$E_t^j \leq C_0\exp(C_1\check{\tau}_t)\left(1+\|x\|^{2p}\right)h^p \leq C_0\exp(C_1 T)\left(1+\|x\|^{2p}\right)h^p.$$

Applying a similar argument, using (2.2.10) from Proposition 2.2.4, we get the same result for $\bar{X}^{0,\eta}$ and $\bar{X}^{d+1,\eta}$. We conclude by setting $C_2 = C_0\exp(C_1 T)$. ∎

The following lemma deals with the estimation of the difference between the scheme $X^{NV,\eta}$ and the intermediate process $\bar{X}^{j,\eta}$ for $j \in \{0,\dots,d+1\}$.

Lemma 2.2.7 *For all $p \geq 1$, there exists a constant $C_3 \in \mathbb{R}_+^*$ such that for all $t \in [0,T], N \in \mathbb{N}^*$ and all $j \in \{0,\dots,d+1\}$,*

$$\mathbb{E}\left[\left\|\bar{X}_t^{j,\eta} - X_{\hat{\tau}_t}^{NV,\eta}\right\|^{2p}\middle|\eta\right] \leq C_3\left(1+\|x\|^{2p}\right)h^p.$$

Proof : Let $p \geq 1, t \in [0,T]$ and $j \in \{1,\dots,d+1\}$. Using telescopic summation and convexity inequality, we get

$$\left\|\bar{X}_t^{j,\eta} - X_{\hat{\tau}_t}^{NV,\eta}\right\|^{2p} \leq (d+2)^{2p-1}\left(\left\|\bar{X}_t^{j,\eta} - \bar{X}_{\check{\tau}_t}^{j-1,\eta}\mathbb{1}_{\{\eta_t=1\}} - \bar{X}_{\check{\tau}_t}^{j+1,\eta}\mathbb{1}_{\{\eta_t=-1\}}\right\|^{2p}\right.$$
$$\left. + \sum_{\eta_t m < \eta_t j}\left\|\bar{X}_{\check{\tau}_t}^{m,\eta} - \bar{X}_{\check{\tau}_t}^{m-1,\eta}\mathbb{1}_{\{\eta_t=1\}} - \bar{X}_{\check{\tau}_t}^{m+1,\eta}\mathbb{1}_{\{\eta_t=-1\}}\right\|^{2p}\right).$$

Taking the conditional expectation, and using Lemma 2.2.6, we obtain

$$\mathbb{E}\left[\left\|\bar{X}_t^{j,\eta} - X_{\hat{\tau}_t}^{NV,\eta}\right\|^{2p}\middle|\eta\right] \leq (d+2)^{2p-1}\left(d+2T^p\right)C_2\left(1+\|x\|^{2p}\right)h^p = C_3\left(1+\|x\|^{2p}\right)h^p,$$

with $C_3 = (d+2)^{2p-1}\left(d+2T^p\right)C_2$. ∎

2.2.3 Proof of the strong convergence

Proof : Let $p \in [1,+\infty)$, $t \in [0,T]$ and $s \in [0,t]$. Subtracting (2.2.7) from (2.1.1), we can evaluate the difference between the exact solution and the scheme

$$\begin{aligned}X_s - X_s^{NV,\eta} = &\frac{1}{2}\left(\int_0^s\left(\sigma^0(X_u) - \sigma^0\left(\bar{X}_u^{0,\eta}\right)\right)du + \int_0^t\left(\sigma^0(X_u) - \sigma^0\left(\bar{X}_u^{d+1,\eta}\right)\right)du\right)\\ &+ \sum_{j=1}^d\int_0^s\left(\sigma^j(X_u) - \sigma^j(\bar{X}_u^{j,\eta})\right)dW_u^j\\ &+ \frac{1}{2}\sum_{j=1}^d\int_0^s\left(\partial\sigma^j\sigma^j\left(X_u\right) - \partial\sigma^j\sigma^j\left(\bar{X}_u^{j,\eta}\right)\right)du.\end{aligned}$$

Using a convexity inequality and taking the conditional expectation of the supremum, we get

$$\mathbb{E}\left[\sup_{s\leq t}\left\|X_s - X_s^{NV,\eta}\right\|^{2p}\middle|\eta\right] \leq \left(2\left(d+1\right)\right)^{2p-1}\left(\sum_{j=1}^d E_j + \frac{1}{2^{2p}}\sum_{j=0}^{d+1} I_j\right), \tag{2.2.13}$$

where

$$I_0 = \mathbb{E}\left[\sup_{s\leq t}\left\|\int_0^s\left(\sigma^0(X_u) - \sigma^0\left(\bar{X}_u^{0,\eta}\right)\right)du\right\|^{2p}\middle|\eta\right],$$

$$I_{d+1} = \mathbb{E}\left[\sup_{s\leq t}\left\|\int_0^s\left(\sigma^0(X_u) - \sigma^0\left(\bar{X}_u^{d+1,\eta}\right)\right)du\right\|^{2p}\middle|\eta\right],$$

and for $j \in \{1,\ldots,d\}$,

$$E_j = \mathbb{E}\left[\sup_{s\leq t}\left\|\int_0^s\left(\sigma^j(X_u) - \sigma^j(\bar{X}_u^{j,\eta})\right)dW_u^j\right\|^{2p}\middle|\eta\right],$$

$$I_j = \mathbb{E}\left[\sup_{s\leq t}\left\|\int_0^s\left(\partial\sigma^j\sigma^j\left(X_u\right) - \partial\sigma^j\sigma^j\left(\bar{X}_u^{j,\eta}\right)\right)du\right\|^{2p}\middle|\eta\right].$$

Let us focus on E_j and I_j, for $j \in \{1, \dots, d\}$. The independence between W and η permits to apply the Burkholder-Davis-Gundy inequality to obtain

$$E_j \leq K\ \mathbb{E}\left[\left(\int_0^t \left\|\sigma^j(X_u) - \sigma^j(\bar{X}_u^{j,\eta})\right\|^2 du\right)^p \middle| \eta\right] \leq KT^{p-1} \int_0^t \mathbb{E}\left[\left\|\sigma^j(X_u) - \sigma^j(\bar{X}_u^{j,\eta})\right\|^{2p} \middle| \eta\right] du,$$

where K is the constant that appears in the Burkholder-Davis-Gundy inequality. By the Lipschitz assumption

$$E_j \leq KT^{p-1}L^{2p} \int_0^t \mathbb{E}\left[\left\|X_u - \bar{X}_u^{j,\eta}\right\|^{2p} \middle| \eta\right] du. \tag{2.2.14}$$

Applying a convexity inequality, we obtain

$$I_j \leq T^{2p-1} \int_0^t \mathbb{E}\left[\left\|\partial\sigma^j\sigma^j\left(X_s\right) - \partial\sigma^j\sigma^j\left(\bar{X}_s^{j,\eta}\right)\right\|^{2p} \middle| \eta\right] ds.$$

Again, by the Lipschitz assumption, we also get

$$I_j \leq T^{2p-1}L^{2p} \int_0^t \mathbb{E}\left[\left\|X_u - \bar{X}_u^{j,\eta}\right\|^{2p} \middle| \eta\right] du. \tag{2.2.15}$$

Using the same approach, we get a similar result for I_0 and I_{d+1}. Combining (2.2.14) and (2.2.15), together with (2.2.13), we obtain

$$\mathbb{E}\left[\sup_{s\leq t}\left\|X_u - X_u^{NV,\eta}\right\|^{2p} \middle| \eta\right] \leq \alpha \sum_{j=0}^{d+1} \int_0^t \mathbb{E}\left[\left\|X_u - \bar{X}_u^{j,\eta}\right\|^{2p} \middle| \eta\right] du, \tag{2.2.16}$$

where $\alpha = \left(2\left(d+1\right)\right)^{2p-1} L^{2p}\left(KT^{p-1} + \frac{T^{2p-1}}{2^{2p}}\right)$. Now, we look at $\left\|X_u - \bar{X}_u^{j,\eta}\right\|$, for $u \in [0,t]$. Let $j \in \{0, \dots, d+1\}$ and $u \in [0,t]$. Introducing the solution X and the Ninomiya-Victoir scheme $X^{NV,\eta}$ at time $\hat{\tau}_u$, and using a convexity inequality we get

$$\mathbb{E}\left[\left\|X_u - \bar{X}_u^{j,\eta}\right\|^{2p} \middle| \eta\right] \leq 3^{2p-1}\ \mathbb{E}\left[\left\|X_u - X_{\hat{\tau}_u}\right\|^{2p} + \left\|X_{\hat{\tau}_u} - X_{\hat{\tau}_u}^{NV,\eta}\right\|^{2p} + \left\|X_{\hat{\tau}_u}^{NV,\eta} - \bar{X}_u^{j,\eta}\right\|^{2p} \middle| \eta\right].$$

Then, using the estimation (2.2.9) from Proposition 2.2.4

$$\mathbb{E}\left[\left\|X_u - X_{\hat{\tau}_u}\right\|^{2p} \middle| \eta\right] \leq C_0\left(1 + \|x\|^{2p}\right)(u - \hat{\tau}_u)^p \leq C_0\left(1 + \|x\|^{2p}\right)h^p,$$

and from Lemma 2.2.7

$$\mathbb{E}\left[\left\|\bar{X}_u^{j,\eta} - X_{\hat{\tau}_u}^{NV,\eta}\right\|^{2p} \middle| \eta\right] \leq C_3\left(1 + \|x\|^{2p}\right)h^p.$$

Moreover

$$\mathbb{E}\left[\left\|X_{\hat{\tau}_u} - X_{\hat{\tau}_u}^{NV,\eta}\right\|^{2p} \middle| \eta\right] \leq \mathbb{E}\left[\sup_{v\leq u}\left\|X_v - X_v^{NV,\eta}\right\|^{2p} \middle| \eta\right].$$

We finally get

$$\mathbb{E}\left[\sup_{s\leq t}\left\|X_s - X_s^{NV,\eta}\right\|^{2p}\middle|\eta\right] \leq \beta\int_0^t \mathbb{E}\left[\sup_{v\leq u}\left\|X_v - X_v^{NV,\eta}\right\|^{2p}\middle|\eta\right] du + \gamma\left(1+\|x\|^{2p}\right)h^p,$$

where

$$\beta = 3^{2p-1}\left(d+2\right)\alpha,$$

and

$$\gamma = \beta T\left(C_0 + C_3\right).$$

Before applying Gronwall's lemma, let us remark, by estimates (2.2.16) and (2.2.8) and Lemma 2.2.5, that $\mathbb{E}\left[\sup_{s\leq t}\left\|X_s - X_s^{NV,\eta}\right\|^{2p}\right]$ is finite. Thanks to Gronwall's lemma, we conclude that

$$\mathbb{E}\left[\sup_{t\leq T}\left\|X_t - X_t^{NV,\eta}\right\|^{2p}\middle|\eta\right] \leq \exp\left(\beta T\right)\gamma\left(1+\|x\|^{2p}\right)h^p.$$

∎

We conclude this section with a lemma which will be useful for the next section.

Lemma 2.2.8 *Let $F\in\mathcal{C}^2\left(\mathbb{R}^n,\mathbb{R}^n\right)$ and assume that its first and second order derivatives have a polynomial growth. Under the assumptions of Theorem 2.2.3 we have the following result. For all $p\in[1,+\infty)$ there exists a constant $C_4\in\mathbb{R}_+^*$ such that for all $N\in\mathbb{N}^*$ and all $j\in\{0,\ldots,d+1\}$,*

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\int_0^t F\left(\bar{X}_s^{j,\eta}\right) - F\left(X_{\hat{\tau}_s}^{NV,\eta}\right)ds\right\|^{2p}\middle|\eta\right] \leq C_4 h^{2p}.$$

Proof : Let $j\in\{0,\ldots,d+1\}, i\in\{1,\ldots,n\}$, and $t\in[0,T]$. Using the integration by parts formula

$$\begin{aligned}\int_0^t\left(F^i\left(\bar{X}_s^{j,\eta}\right) - F^i\left(X_{\hat{\tau}_s}^{NV,\eta}\right)\right)du &= \int_0^t\left(t\wedge\check{\tau}_s - s\right)d\left(F^i\left(\bar{X}_s^{j,\eta}\right)\right)\\ &+ \int_0^{\check{\tau}_t}\sum_{\eta_s m<\eta_s j}\left(t\wedge\check{\tau}_s - \hat{\tau}_s\right)d\left(F^i\left(\bar{X}_s^{m,\eta}\right)\right).\end{aligned}$$

Then, using the chain rule for $m\in\{0,d+1\}$, we get

$$d\left(F^i\left(\bar{X}_s^{m,\eta}\right)\right) = \frac{1}{2}\sigma^0\left(\bar{X}_s^{m,\eta}\right).\nabla F^i\left(\bar{X}_s^{m,\eta}\right)ds.$$

Applying Itô's formula for $m\in\{1,\ldots,d\}$, we obtain

$$\begin{aligned}d\left(F^i\left(\bar{X}_s^{m,\eta}\right)\right) = \frac{1}{2}\partial\sigma^m\sigma^m\left(\bar{X}_s^{m,\eta}\right).\nabla F^i\left(\bar{X}_s^{m,\eta}\right)ds + \sigma^m\left(\bar{X}_s^{m,\eta}\right).\nabla F^i\left(\bar{X}_s^{m,\eta}\right)dW_s^m\\ + \frac{1}{2}tr\left(\sigma^m\left(\sigma^m\right)^*\left(\bar{X}_s^{m,\eta}\right)\nabla^2 F^i\left(\bar{X}_s^{m,\eta}\right)\right)ds.\end{aligned}$$

In both cases, combining a convexity inequality, the Burkholder-Davis-Gundy inequality, the Holder inequality, the Lipschitz assumption on $\sigma^m, \partial\sigma^m\sigma^m$, for $m\in\{0,\ldots,d\}$, the polynomial growth assumption for the first and second order derivatives of F, and $t \wedge \check{\tau}_s - s \leq h$, for all $s \in [0, \check{\tau}_t]$, we get two constants $\gamma \in \mathbb{R}_+^*$ and $q \in \mathbb{N}^*$, independent of N, such that

$$\mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t F^i\left(\bar{X}_s^{j,\eta}\right) - F^i\left(X_{\hat{\tau}_s}^{NV,\eta}\right) ds\right|^{2p} \middle| \eta\right] \leq \gamma h^{2p} \sum_{m=0}^{d+1} \int_0^T \mathbb{E}\left[1 + \left\|\bar{X}_s^{m,\eta}\right\|^{2q} \middle| \eta\right] ds.$$

We conclude by using Lemma 2.2.5 and taking the Euclidean norm. ∎

2.3 Coupling with Giles-Szpruch scheme

In [23], Giles and Szpruch proposed a modified Milstein scheme defined as follows

$$\begin{cases} X_{t_{k+1}}^{GS} = X_{t_k}^{GS} + b\left(X_{t_k}^{GS}\right)(t_{k+1} - t_k) + \displaystyle\sum_{j=1}^{d} \sigma^j\left(X_{t_k}^{GS}\right)\Delta W_{t_{k+1}}^j \\ \qquad + \displaystyle\frac{1}{2}\sum_{j,m=1}^{d} \partial\sigma^j\sigma^m\left(X_{t_k}^{GS}\right)\left(\Delta W_{t_{k+1}}^j \Delta W_{t_{k+1}}^m - \mathbb{1}_{\{j=m\}}h\right) \\ X_{t_0}^{GS} = x. \end{cases} \tag{2.3.1}$$

In comparison with the Milstein scheme, the terms involving the Lévy areas

$$\int_{t_k}^{t_{k+1}} \Delta W_s^j dW_s^m - \int_{t_k}^{t_{k+1}} \Delta W_s^m dW_s^j$$

have been removed. According to Lemma 4.2 in [23], the strong order of convergence is $\gamma = 1/2$.

Lemma 2.3.1 *Assume that $b, \sigma^j \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$, for all $j \in \{1, \ldots, d\}$, with bounded first and second order derivatives, and that for all $j, m \in \{1, \ldots, d\}$, $\partial\sigma^j\sigma^m$ $\partial\sigma^j\sigma^m$ has bounded first order derivatives. Then there exists a constant $C_{GS} \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$*

$$\mathbb{E}\left[\max_{k\in\{0,\ldots,N\}}\left\|X_{t_k} - X_{t_k}^{GS}\right\|^{2p}\right] \leq C_{GS} h^p.$$

Giles and Szpruch also proposed an antithetic version of their scheme based on the swapping of each successive pair of the Brownian increments in the scheme. With regards to the multilevel Monte Carlo estimator, Giles and Szpruch use the arithmetic average of the scheme (2.3.1) and its antithetic version on the fine grids, at each level $l \in \{1, \ldots, L\}$ as follows

$$\hat{Y} = \frac{1}{M_0}\sum_{k=1}^{M_0} f\left(X_T^{1,0,k}\right) + \sum_{l=1}^{L}\frac{1}{M_l}\sum_{k=1}^{M_l}\left(\frac{1}{2}\left(f\left(\tilde{X}_T^{2^l,l,k}\right) + f\left(X_T^{2^l,l,k}\right)\right) - f\left(X_T^{2^{l-1},l,k}\right)\right).$$

The swapping of each successive pair of Brownian increments provides a strong convergence of order 1 between the schemes used on the coarse and fine grids, and so Giles and Szpruch obtained the convergence rate $\beta = 2$ of the variance

$$\mathbb{V}\left(\frac{1}{2}\left(f\left(\tilde{X}_T^{2^l,l,k}\right) + f\left(X_T^{2^l,l,k}\right)\right) - f\left(X_T^{2^{l-1},l,k}\right)\right),$$

when the payoff f is smooth. In this way, using this multilevel Monte Carlo estimator leads to the computational complexity $O\left(\epsilon^{-2}\right)$ for the mean-square-root error ϵ. To use the Ninomiya-Victoir scheme either at each level or only at the finest level of a multilevel Monte Carlo estimator, we study in this section the coupling between the Ninomiya-Victoir and Giles-Szpruch schemes. To keep $\beta = 2$, we suggest comparing the Giles-Szpruch scheme with the following modified Ninomiya-Victoir scheme

$$\bar{X}^{NV,\eta} = \frac{1}{2}\left(X^{NV,\eta} + X^{NV,-\eta}\right).$$

To be consistent with the interpolation of the Ninomiya-Victoir scheme, we define the interpolation of the scheme between the grid points as follows

$$\begin{aligned} X_s^{GS} = x &+ \int_0^s b\left(X_{\hat{\tau}_u}^{GS}\right) du + \sum_{j=1}^d \int_0^s \sigma^j\left(X_{\hat{\tau}_u}^{GS}\right) dW_u^j + \sum_{j=1}^d \int_0^s \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{GS}\right)\Delta W_u^j dW_u^j \\ &+ \frac{1}{2}\sum_{\substack{j,m=1\\ m\neq j}}^d \int_0^s \partial\sigma^j\sigma^m\left(X_{\hat{\tau}_u}^{GS}\right)\Delta W_{\check{\tau}_u}^m dW_u^j. \end{aligned} \tag{2.3.2}$$

Theorem 2.3.2 *Assume that $b \in \mathcal{C}^2\left(\mathbb{R}^n, \mathbb{R}^n\right)$ with bounded first and second order derivatives, for all $j \in \{1, \ldots, d\}, \sigma^j \in \mathcal{C}^3\left(\mathbb{R}^n, \mathbb{R}^n\right)$ with bounded first and second order derivatives and with polynomially growing third order derivatives, and that for all $j, m \in \{1, \ldots, d\}$, $\partial\sigma^j\sigma^m$ has bounded first order derivatives. Then there exists a constant $C \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$,*

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\bar{X}_t^{NV,\eta} - X_t^{GS}\right\|^{2p}\middle|\eta\right] \leq Ch^{2p}.$$

Proof : We denote by L the common Lipschitz constant of b, σ^j and $\partial\sigma^j\sigma^m$, for $j, m \in \{1, \ldots, d\}$. We also denote by M the global bound on first and second derivatives of b and σ^j, for $j \in \{1, \ldots, d\}$. Let $t \in [0, T]$ and $s \in [0, t]$. Writing $\bar{X}^{NV,\eta}$ in integral form, we get

$$\begin{aligned} \bar{X}_s^{NV,\eta} = x &+ \sum_{j=1}^d \int_0^s \frac{1}{2}\left(\sigma^j\left(\bar{X}_u^{j,\eta}\right) + \sigma^j\left(\bar{X}_s^{j,-\eta}\right)\right) dW_u^j \\ &+ \sum_{j=1}^d \int_0^s \frac{1}{4}\left(\partial\sigma^j\sigma^j\left(\bar{X}_u^{j,\eta}\right) + \partial\sigma^j\sigma^j\left(\bar{X}_u^{j,-\eta}\right)\right) du \\ &+ \int_0^s \frac{1}{4}\left(\sigma^0\left(\bar{X}_u^{0,\eta}\right) + \sigma^0\left(\bar{X}_u^{0,-\eta}\right)\right) du \\ &+ \int_0^s \frac{1}{4}\left(\sigma^0\left(\bar{X}_u^{d+1,\eta}\right) + \sigma^0\left(\bar{X}_u^{d+1,-\eta}\right)\right) du. \end{aligned}$$

Then using $\frac{1}{2}b - \frac{1}{4}\sum_{j=1}^d \partial\sigma^j\sigma^j - \frac{1}{2}\sigma^0 = 0$, we get

$$
\begin{aligned}
\bar{X}_s^{NV,\eta} = x &+ \sum_{j=1}^{d} \int_0^s \frac{1}{2}\left(\sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + \sigma^j\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\right) dW_u^j \\
&+ \int_0^s \frac{1}{2}\left(b\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + b\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\right) du \\
&+ \sum_{j=1}^{d} \int_0^s \frac{1}{2}\left(\sigma^j\left(\bar{X}_u^{j,\eta}\right) - \sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + \sigma^j\left(\bar{X}_u^{j,-\eta}\right) - \sigma^j\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\right) dW_u^j \\
&+ \sum_{j=1}^{d} \int_0^s \frac{1}{4}\left(\partial\sigma^j\sigma^j\left(\bar{X}_u^{j,\eta}\right) - \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + \partial\sigma^j\sigma^j\left(\bar{X}_u^{j,-\eta}\right) - \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\right) du \\
&+ \int_0^s \frac{1}{4}\left(\sigma^0\left(\bar{X}_u^{0,\eta}\right) - \sigma^0\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + \sigma^0\left(\bar{X}_u^{0,-\eta}\right) - \sigma^0\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\right) du \\
&+ \int_0^s \frac{1}{4}\left(\sigma^0\left(\bar{X}_u^{d+1,\eta}\right) - \sigma^0\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + \sigma^0\left(\bar{X}_u^{d+1,-\eta}\right) - \sigma^0\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\right) du.
\end{aligned}
$$

Subtracting (2.3.2) and using a convexity inequality, we obtain

$$
\mathbb{E}\left[\sup_{s\le t}\left\|\bar{X}_s^{NV,\eta} - X_s^{GS}\right\|^{2p}\middle|\eta\right] \le 3^{2p-1}(d+1)^{2p-1}\left(\sum_{j=1}^{d} I_j + \sum_{j=0}^{d} E_j + \sum_{j=0}^{d+1} R_j\right), \tag{2.3.3}
$$

where

$$
E_0 = \mathbb{E}\left[\sup_{s\le t}\left\|\int_0^s \left(\frac{1}{2}\left(b\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + b\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\right) - b\left(X_{\hat{\tau}_u}^{GS}\right)\right) du\right\|^{2p}\middle|\eta\right],
$$

$$
R_0 = \mathbb{E}\left[\sup_{s\le t}\left\|\int_0^s \frac{1}{4}\left(\sigma^0\left(\bar{X}_u^{0,\eta}\right) - \sigma^0\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + \sigma^0\left(\bar{X}_u^{0,-\eta}\right) - \sigma^0\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\right) du\right\|^{2p}\middle|\eta\right],
$$

$$
R_{d+1} = \mathbb{E}\left[\sup_{s\le t}\left\|\int_0^s \frac{1}{4}\left(\sigma^0\left(\bar{X}_u^{d+1,\eta}\right) - \sigma^0\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + \sigma^0\left(\bar{X}_u^{d+1,-\eta}\right) - \sigma^0\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\right) du\right\|^{2p}\middle|\eta\right],
$$

and for $j \in \{1, \ldots, d\}$,

$$
\begin{aligned}
I_j = \mathbb{E}\Bigg[\sup_{s\le t}\Bigg\|\int_0^s \frac{1}{2}\Bigg(&\sigma^j\left(\bar{X}_u^{j,\eta}\right) - \sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + \sigma^j\left(\bar{X}_u^{j,-\eta}\right) - \sigma^j\left(X_{\hat{\tau}_u}^{NV,-\eta}\right) \\
&- \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{GS}\right)\Delta W_u^j - \frac{1}{2}\sum_{\substack{m=1\\ m\ne j}}^{d} \partial\sigma^j\sigma^m\left(X_{\hat{\tau}_u}^{GS}\right)\Delta W_{\hat{\tau}_u}^m\Bigg) dW_u^j\Bigg\|^{2p}\Bigg|\eta\Bigg],
\end{aligned}
$$

$$
E_j = \mathbb{E}\left[\sup_{s\le t}\left\|\int_0^s \left(\frac{1}{2}\left(\sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + \sigma^j\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\right) - \sigma^j\left(X_{\hat{\tau}_u}^{GS}\right)\right) dW_u^j\right\|^{2p}\middle|\eta\right],
$$

$$
\begin{aligned}
R_j = \mathbb{E}\Bigg[\sup_{s\le t}\Bigg\|\int_0^s \frac{1}{4}\Bigg(&\partial\sigma^j\sigma^j\left(\bar{X}_u^{j,\eta}\right) - \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right) \\
&+ \partial\sigma^j\sigma^j\left(\bar{X}_u^{j,-\eta}\right) - \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\Bigg) du\Bigg\|^{2p}\Bigg|\eta\Bigg].
\end{aligned}
$$

Step 1: estimation of E_j, for $j \in \{0, \ldots, d\}$.

Let us start with the estimation of E_j, for $j \in \{0, \ldots, d\}$. We set for $F^0 = b$ and $F^j = \sigma^j$ for

$j \in \{1, \ldots, d\}$. Combining the Burkholder-Davis-Gundy inequality and a convexity inequality, we get

$$E_j \leq \max\left\{T^{2p-1}, KT^{p-1}\right\} \int_0^t \mathbb{E}\left[\left\|\frac{1}{2}\left(F^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + F^j\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\right) - F^j\left(X_{\hat{\tau}_u}^{GS}\right)\right\|^{2p} \middle| \eta\right] du,$$

where K is the constant that appears in the Burkholder-Davis-Gundy inequality. For $i \in \{1, \ldots, n\}$, denoting

$$Y_u = \frac{1}{2}\left(F^{ij}\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + F^{ij}\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\right) - F^{ij}\left(X_{\hat{\tau}_u}^{GS}\right),$$

and performing a second order Taylor series expansion, we obtain

$$\begin{aligned} Y_u &= F^{ij}\left(\bar{X}_{\hat{\tau}_u}^{NV,\eta}\right) - F^{ij}\left(X_{\hat{\tau}_u}^{GS}\right) \\ &\quad + \frac{1}{16}\left(X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}^{NV,-\eta}\right)^*\left(\nabla^2 F^{ij}\left(\xi_{\hat{\tau}_u}^1\right) + \nabla^2 F^{ij}\left(\xi_{\hat{\tau}_u}^2\right)\right)\left(X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}^{NV,-\eta}\right), \end{aligned}$$

where $\xi_{\hat{\tau}_u}^1$ and $\xi_{\hat{\tau}_u}^2$ are points between $X_{\hat{\tau}_u}^{NV,\eta}$ and $X_{\hat{\tau}_u}^{NV,-\eta}$. Then, we easily get

$$\|Y_u\|^{2p} \leq \alpha_1\left(\left\|\bar{X}_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}^{GS}\right\|^{2p} + \left\|X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}^{NV,-\eta}\right\|^{4p}\right),$$

where $\alpha_1 = 2^{2p-1}\left(L^{2p} + \left(\frac{M}{8}\right)^{2p}\right)$. Thus

$$\begin{aligned} E_j \leq \alpha_1 \max\left\{T^{2p-1}, KT^{p-1}\right\} \Bigg(&\int_0^t \mathbb{E}\left[\sup_{v \leq u}\left\|\bar{X}_v^{NV,\eta} - X_v^{GS}\right\|^{2p} \middle| \eta\right] du \\ &+ \int_0^t \mathbb{E}\left[\left\|X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}^{NV,-\eta}\right\|^{4p} \middle| \eta\right] du\Bigg). \end{aligned}$$

Introducing the solution X at time $\hat{\tau}_u$ and using a convexity inequality, we obtain

$$\mathbb{E}\left[\left\|X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}^{NV,-\eta}\right\|^{4p} \middle| \eta\right] \leq 2^{4p-1}\left(\mathbb{E}\left[\left\|X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}\right\|^{4p} \middle| \eta\right] + \mathbb{E}\left[\left\|X_{\hat{\tau}_u} - X_{\hat{\tau}_u}^{NV,-\eta}\right\|^{4p} \middle| \eta\right]\right).$$

Thanks to Theorem 2.2.3, we deduce that

$$\mathbb{E}\left[\left\|X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}^{NV,-\eta}\right\|^{4p} \middle| \eta\right] \leq 2^{4p} C_{NV}\left(1 + \|x\|^{4p}\right) h^{2p}.$$

It follows that

$$E_j \leq \beta_1\left(\int_0^t \mathbb{E}\left[\sup_{v \leq u}\left\|\bar{X}_v^{NV,\eta} - X_v^{GS}\right\|^{2p} \middle| \eta\right] du + h^{2p}\right), \tag{2.3.4}$$

where $\beta_1 = \alpha_1 \max\{T^{2p-1}, KT^{p-1}\} \max\left\{1, 2^{4p} C_{NV}\left(1 + \|x\|^{4p}\right)\right\}$.

Step 2: estimation of R_j, for $j \in \{0, \ldots, d\}$.

Turning to the estimation of R_j, for $j \in \{0, \ldots, d\}$, from Lemma 2.2.8 we get a constant

$\beta_2 \in \mathbb{R}_+^*$, such that

$$R_j \leq \beta_2 h^{2p}. \tag{2.3.5}$$

Step 3: estimation of I_j, for $j \in \{1, \ldots, d\}$.

It remains to estimate I_j, for $j \in \{1, \ldots, d\}$. Using the Burkholder-Davis-Gundy and convexity inequalities, we get

$$I_j \leq \frac{1}{2^{2p}} K T^{p-1} \int_0^t \mathbb{E}\Bigg[\Bigg\| \sigma^j\left(\bar{X}_u^{j,\eta}\right) - \sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right) + \sigma^j\left(\bar{X}_u^{j,-\eta}\right) - \sigma^j\left(X_{\hat{\tau}_u}^{NV,-\eta}\right) \\ - 2\partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{GS}\right)\Delta W_u^j - \sum_{\substack{m=1 \\ m\neq j}}^{d} \partial\sigma^j\sigma^m\left(X_{\hat{\tau}_u}^{GS}\right)\Delta W_{\check{\tau}_u}^m \Bigg\|^{2p} \Bigg| \eta\Bigg] ds.$$

Introducing

$$\Psi_u^j = \Psi_u^{j,\eta} + \Psi_u^{j,-\eta},$$

where

$$\Psi_u^{j,\eta} = \sigma^j\left(\bar{X}_u^{j,\eta}\right) - \sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right) - \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right)\Delta W_u^j - \sum_{\eta_u m < \eta_u j} \partial\sigma^j\sigma^m\left(X_{\hat{\tau}_u}^{NV,\eta}\right)\Delta W_{\check{\tau}_u}^m, \tag{2.3.6}$$

and

$$\Phi_u^j = \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right)\Delta W_u^j + \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\Delta W_u^j - 2\partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{GS}\right)\Delta W_u^j \\ + \sum_{\eta_u m < \eta_u j} \partial\sigma^j\sigma^m\left(X_{\hat{\tau}_u}^{NV,\eta}\right)\Delta W_{\check{\tau}_u}^m + \sum_{\eta_u m > \eta_u j} \partial\sigma^j\sigma^m\left(X_{\hat{\tau}_u}^{NV,-\eta}\right)\Delta W_{\check{\tau}_u}^m \\ - \sum_{\substack{m=1 \\ m\neq j}}^{d} \partial\sigma^j\sigma^m\left(X_{\hat{\tau}_u}^{GS}\right)\Delta W_{\check{\tau}_u}^m,$$

we obtain

$$I_j \leq \frac{1}{2} K T^{p-1}\left(\int_0^t \mathbb{E}\left[\left\|\Psi_u^j\right\|^{2p} + \left\|\Phi_u^j\right\|^{2p} \Big| \eta\right] du\right).$$

Step 3.1: estimation of $\mathbb{E}\left[\left\|\Psi_u^j\right\|^{2p} \Big| \eta\right]$, for $j \in \{1, \ldots, d\}$.

Applying Itô's formula in Equation (2.3.6) to compute $\sigma^j\left(\bar{X}_u^{j,\eta}\right) - \sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right)$, we get

$$\Psi_u^{j,\eta} = \int_{\hat{\tau}_u}^{u}\left(\partial\sigma^j\sigma^j\left(\bar{X}_v^{j,\eta}\right) - \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_v}^{NV,\eta}\right)\right) dW_v^j \\ + \sum_{\eta_u m < \eta_u j} \int_{\hat{\tau}_u}^{\check{\tau}_u}\left(\partial\sigma^j\sigma^m\left(\bar{X}_v^{m,\eta}\right) - \partial\sigma^j\sigma^m\left(X_{\hat{\tau}_v}^{NV,\eta}\right)\right) dW_v^m \\ + \frac{1}{2}\int_{\hat{\tau}_u}^{u} \partial F^{j,j}\sigma^j\left(\bar{X}_v^{j,\eta}\right) dv + \frac{1}{2}\sum_{\eta_u m < \eta_u j} \int_{\hat{\tau}_u}^{\check{\tau}_u} \partial F^{j,m}\sigma^m\left(\bar{X}_v^{m,\eta}\right) dv,$$

where $F^{j,m} = \partial\sigma^j\sigma^m$, for $j, m \in \{1, \ldots, d\}$. Note that the term

$$\frac{1}{2}\sum_{\eta_u m < \eta_u j} \int_{\hat{\tau}_u}^{\check{\tau}_u} \partial F^{j,m}\sigma^m\left(\bar{X}_v^{m,\eta}\right) dv$$

is equal to the sum of the drift contribution and the Itô correction due to the dynamics of $\bar{X}^{m,\eta}$. The assumptions on σ^j and Lemma 2.2.5, ensure that

- for all $j, m \in \{1, \ldots, d\}, \partial\sigma^j\sigma^m$ is Lipschitz continuous.
- for all $j, m \in \{1, \ldots, d\}, \partial F^{j,m}\sigma^m\left(\bar{X}_v^{m,\eta}\right)$ has uniformly bounded moments.

Using Lemma 2.2.7, the Burkholder-Davis-Gundy and a convexity inequalities, we obtain a constant $\gamma_3 \in \mathbb{R}_+^*$ such that

$$\mathbb{E}\left[\left\|\Psi_u^{j,\eta}\right\|^{2p}\middle|\eta\right] \leq \gamma_3 h^{2p}.$$

Obviously, we have the same inequality for $\Psi^{j,-\eta}$.

Step 3.2: estimation of $\mathbb{E}\left[\|\Phi_u^j\|^{2p}\middle|\eta\right]$, for $j \in \{1, \ldots, d\}$.

By the Lipschitz assumption

$$\begin{aligned}\left\|\Phi_u^j\right\|^{2p} \leq (d+1)^{2p-1} L^{2p}\Bigg(&\left\|X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}^{GS}\right\|^{2p}\left|\Delta W_u^j\right|^{2p} + \left\|X_{\hat{\tau}_u}^{NV,-\eta} - X_{\hat{\tau}_u}^{GS}\right\|^{2p}\left|\Delta W_u^j\right|^{2p} \\ &+ \sum_{\eta_u m<\eta_u j}\left\|X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}^{GS}\right\|^{2p}\mid \Delta W_{\check{\tau}_u}^m|^{2p} + \sum_{\eta_u m>\eta_u j}\left\|X_{\hat{\tau}_u}^{NV,-\eta} - X_{\hat{\tau}_u}^{GS}\right\|^{2p}\mid \Delta W_{\check{\tau}_u}^m|^{2p}\Bigg).\end{aligned} \tag{2.3.7}$$

By independence

$$\begin{aligned}\mathbb{E}\left[\left\|X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}^{GS}\right\|^{2p}\mid \Delta W_{\check{\tau}_u}^m|^{2p}\middle|\eta\right] &= \mathbb{E}\left[\left\|X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}^{GS}\right\|^{2p}\middle|\eta\right]\mathbb{E}\left[\mid \Delta W_{\check{\tau}_u}^m|^{2p}\right] \\ &\leq 2^{2p-1}\mathbb{E}\left[\mid \Delta W_{\check{\tau}_u}^m|^{2p}\right]\Bigg(\mathbb{E}\left[\left\|X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}\right\|^{2p}\middle|\eta\right] \\ &\quad + \mathbb{E}\left[\left\|X_{\hat{\tau}_u} - X_{\hat{\tau}_u}^{GS}\right\|^{2p}\middle|\eta\right]\Bigg).\end{aligned}$$

Then, using Theorem 2.2.3 and Lemma 2.3.1, we get

$$\mathbb{E}\left[\left\|X_{\hat{\tau}_u}^{NV,\eta} - X_{\hat{\tau}_u}^{GS}\right\|^{2p}\mid \Delta W_{\check{\tau}_u}^m|^{2p}\middle|\eta\right] \leq 2^{2p}\,\mathbb{E}\left[|G|^{2p}\right]\left(C_{NV}\left(1+\|x\|^{2p}\right) + C_{GS}\right)h^{2p},$$

where G is a normal random variable. Using the same approach, we get the same result for the other terms on the right-hand side of (2.3.7). Thus, we deduce that there exists a constant $\alpha_3 \in \mathbb{R}_+^*$ such that

$$\mathbb{E}\left[\left\|\Phi_u^j\right\|^{2p}\middle|\eta\right] \leq \alpha_3 h^{2p}.$$

Combining our different inequalities, we obtain

$$I_j \leq \beta_3 h^{2p}, \tag{2.3.8}$$

where $\beta_3 = \frac{1}{2} K T^p \left(\alpha_3 + 2^{2p}\gamma_3\right)$.

Step 4: conclusion.

Finally, by combining (2.3.4), (2.3.5), (2.3.8), together with (2.3.3), we complete the proof using Gronwall's lemma

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\bar{X}_t^{NV} - X_t^{GS}\right\|^{2p}\middle|\eta\right] \leq Ch^{2p},$$

where $C = 3^{2p-1}(d+1)^{2p-1}(d\beta_1 + (d+1)\beta_2 + (d+2)\beta_3)\exp\left(3^{2p-1}(d+1)^{2p-1} d\beta_1 T\right)$. ∎

2.4 Multilevel methods for SDEs

In this section, we are interested in the computation, by Monte Carlo methods, of the expectation $Y = \mathbb{E}[f(X_T)]$, where $X = (X_t)_{t\in[0,T]}$ is the solution of the stochastic differential equation (2.1.1) and $f : \mathbb{R}^n \mapsto \mathbb{R}$ a given function such that $\mathbb{E}\left[f(X_T)^2\right]$ is finite. We will focus on minimizing the computational complexity subject to a given target error ϵ. To measure the accuracy of an estimator $\hat{Y}$, we will consider the root mean square error

$$RMSE\left(\hat{Y}, Y\right) = \mathbb{E}^{\frac{1}{2}}\left[\left|Y - \hat{Y}\right|^2\right].$$

2.4.1 Multilevel Monte Carlo

The multilevel Monte Carlo method, introduced by Giles in [22], consists in combining multiple levels of discretization, using a geometric sequence of time steps $h_l = T/2^l$ for example. Denoting by X^N a numerical scheme, with time step T/N, the main idea of this technique is to use the following telescopic summation to control the bias

$$\mathbb{E}\left[f\left(X_T^{2^L}\right)\right] = \mathbb{E}\left[f\left(X_T^{1}\right)\right] + \sum_{l=1}^{L}\mathbb{E}\left[f\left(X_T^{2^l}\right) - f\left(X_T^{2^{l-1}}\right)\right].$$

Then, a generalized multilevel Monte Carlo estimator is built as follows

$$\hat{Y}_{MLMC} = \sum_{l=0}^{L}\frac{1}{M_l}\sum_{k=1}^{M_l} Z^{l,k},$$

where $\left(Z^{l,k}\right)_{0\leq l\leq L, 1\leq k\leq M_l}$ are independent random variables such that for, a given discretization level l with $l \in \{0, \ldots, L\}$, the sequence $\left(Z^{l,k}\right)_{1\leq k\leq M_l}$ is identically distributed and satisfies

$$\mathbb{E}\left[Z^0\right] = \mathbb{E}\left[f\left(X_T^1\right)\right], \tag{2.4.1}$$

and for all $l \in \{1, \ldots, L\}$,

$$\mathbb{E}\left[Z^l\right] = \mathbb{E}\left[f\left(X_T^{2^l}\right) - f\left(X_T^{2^{l-1}}\right)\right]. \tag{2.4.2}$$

Assume that, for a given discretization level $l \in \{0, \dots, L\}$, the computational cost of simulating one sample Z^l is $C\lambda_l 2^l$, where $C \in \mathbb{R}_+$ is a constant, depending only on the discretization scheme and $\lambda_l \in \mathbb{Q}_+^*$ is a weight, depending only on l. The computational complexity of $\hat{Y}_{MLMC}$, denoted by $\mathcal{C}_{MLMC}$, is given by

$$\mathcal{C}_{MLMC} = C \sum_{l=0}^{L} M_l \lambda_l 2^l.$$

The natural choice for $Z^l, l \in \{0, \dots, L\}$ considered in [22] is

$$Z^0 = f\left(X_T^1\right),$$

and for all $l \in \{1, \dots, L\}$,

$$Z^l = f\left(X_T^{2^l}\right) - f\left(X_T^{2^{l-1}}\right).$$

For this canonical choice, it is natural to take $\lambda_0 = 1$ and $\lambda_l = 3/2$, for all $l \in \{1, \dots, L\}$. According to Theorem 3.1 in [22] the optimal complexity $\mathcal{C}_{MLMC}^*$, depends on the order α of weak convergence of the scheme and the order β of convergence to 0 of the variance of Z^l. Here, we recall this complexity theorem.

Theorem 2.4.1 *Assume that*

$$\mathbb{E}\left[f\left(X_T^{2^l}\right)\right] - Y = \frac{c_1}{2^{\alpha l}} + o\left(\frac{1}{2^{\alpha l}}\right), \tag{2.4.3}$$

and

$$\mathbb{V}\left(Z^l\right) = \frac{c_2}{2^{\beta l}} + o\left(\frac{1}{2^{\beta l}}\right), \tag{2.4.4}$$

for some constants $c_1 \in \mathbb{R}^*$ and $c_2 \in \mathbb{R}_+^*$ independent of l. Then, by choosing

$$L^* = \left\lceil \frac{\log_2\left(\frac{\sqrt{2}|c_1|}{\epsilon}\right)}{\alpha} \right\rceil, \tag{2.4.5}$$

and for all $l \in \{0, \dots, L^*\}$

$$M_l^* = \left\lceil \frac{2}{\epsilon^2} \sqrt{\frac{\mathbb{V}\left(Z^l\right)}{\lambda_l 2^l}} \sum_{j=0}^{L^*} \sqrt{\lambda_j 2^j \mathbb{V}\left(Z^j\right)} \right\rceil, \tag{2.4.6}$$

we get an optimal computational complexity

$$\begin{cases} \mathcal{C}_{MLMC}^* = O\left(\epsilon^{-2}\right) \ \text{if } \beta > 1, \\ \mathcal{C}_{MLMC}^* = O\left(\epsilon^{-2}\left(\log\left(\frac{1}{\epsilon}\right)\right)^2\right) \ \text{if } \beta = 1, \\ \mathcal{C}_{MLMC}^* = O\left(\epsilon^{-2+\frac{\beta-1}{\alpha}}\right) \ \text{if } \beta < 1, \end{cases}$$

with $RMSE\left(\hat{Y}_{MLMC}, Y\right)$ bounded by ϵ.

To obtain the estimation (2.4.4), the key point is that the simulation of X^{2^l} and $X^{2^{l-1}}$ comes from the same Brownian path. We easily bound the variance convergence rate from above using the strong convergence rate γ of the numerical scheme, since in general, $\beta \geq 2\gamma$ for a smooth payoff. To attain $\gamma = 1$, one has, in general, to simulate iterated Brownian integrals involving Lévy areas, for which there is no known efficient method. To get around this difficulty, Giles and Szpruch introduced a Milstein scheme without Lévy areas and its antithetic version by swapping the Brownian increments. In a multilevel Monte Carlo method, using the arithmetic average of the modified Milstein scheme and its antithetic version in the finest grid, and the modified Milstein scheme in the coarsest grid leads to $\beta = 2$. By this way, Giles and Szpruch managed to improve the variance convergence rate without simulating the Lévy areas. To be precise, they choose Z^l as follows

$$Z^0_{GS} = f\left(X_T^{GS,1}\right), \tag{2.4.7}$$

and for all $l \in \{1, \ldots, L\}$

$$Z^l_{GS} = \frac{1}{2}\left(f\left(\tilde{X}_T^{GS,2^l}\right) + f\left(X_T^{GS,2^l}\right)\right) - f\left(X_T^{GS,2^{l-1}}\right). \tag{2.4.8}$$

Here, $X^{GS,2^l}$ is the Giles and Szpruch scheme defined by (2.3.1) using a grid with time step $h_l = T/2^l$ and $\tilde{X}^{GS,2^l}$ is an antithetic discretization defined by swapping each successive pair of Brownian increments in the scheme. To be more specific, we define two grids, a coarse grid with time step h_{l-1} and a fine grid with time step h_l. The discretization times $(t_k)_{0\leq k\leq 2^{l-1}}$ and $\left(t_{k+\frac{1}{2}}\right)_{0\leq k\leq 2^{l-1}-1}$ are defined by $t_k = kh_{l-1}$, for all $k \in \left\{0, \ldots, 2^{l-1}\right\}$, and $t_{k+\frac{1}{2}} = \left(k + \frac{1}{2}\right) h_{l-1}$, for all $k \in \left\{0, \ldots, 2^{l-1} - 1\right\}$. Then, on the coarsest grid, $\left(X_{t_{k+1}}^{GS,2^{l-1}}\right)_{0\leq k\leq 2^{l-1}}$ is defined inductively by

$$X_{t_0}^{GS,2^{l-1}} = x,$$

and

$$\begin{aligned} X_{t_{k+1}}^{GS,2^{l-1}} &= X_{t_k}^{GS,2^{l-1}} + b\left(X_{t_k}^{GS,2^{l-1}}\right) h_{l-1} + \sum_{j=1}^{d} \sigma^j\left(X_{t_k}^{GS,2^{l-1}}\right) \Delta W_{t_{k+1}}^{j,c} \\ &+ \frac{1}{2}\sum_{j,m=1}^{d} \partial\sigma^j\sigma^m\left(X_{t_k}^{GS,2^{l-1}}\right)\left(\Delta W_{t_{k+1}}^{j,c}\Delta W_{t_{k+1}}^{m,c} - \mathbb{1}_{\{m=j\}}h_{l-1}\right), \end{aligned}$$

where $\Delta W^c_{t_{k+1}} = W_{t_{k+1}} - W_{t_k}$. Similarly, on the finest grid, $\left(X_{t_{k+1}}^{GS,2^l}\right)_{0\leq k\leq 2^{l-1}}$ is defined inductively by

$$X_{t_0}^{GS,2^l} = x,$$

and

$$\begin{cases} X_{t_{k+\frac{1}{2}}}^{GS,2^l} &= X_{t_k}^{GS,2^l} + b\left(X_{t_k}^{GS,2^l}\right) h_l + \sum\limits_{j=1}^{d} \sigma^j\left(X_{t_k}^{GS,2^l}\right) \Delta W_{t_{k+\frac{1}{2}}}^{j,f} \\ &+\frac{1}{2}\sum\limits_{j,m=1}^{d} \partial\sigma^j\sigma^m\left(X_{t_k}^{GS,2^l}\right)\left(\Delta W_{t_{k+\frac{1}{2}}}^{j,f}\Delta W_{t_{k+\frac{1}{2}}}^{m,f} - \mathbb{1}_{\{m=j\}}h_l\right) \\ X_{t_{k+1}}^{GS,2^l} &= X_{t_{k+\frac{1}{2}}}^{GS,2^l} + b\left(X_{t_{k+\frac{1}{2}}}^{GS,2^l}\right) h_l + \sum\limits_{j=1}^{d} \sigma^j\left(X_{t_{k+\frac{1}{2}}}^{GS,2^l}\right) \Delta W_{t_{k+1}}^{j,f} \\ &+\frac{1}{2}\sum\limits_{j,m=1}^{d} \partial\sigma^j\sigma^m\left(X_{t_{k+\frac{1}{2}}}^{GS,2^l}\right)\left(\Delta W_{t_{k+1}}^{j,f}\Delta W_{t_{k+1}}^{m,f} - \mathbb{1}_{\{m=j\}}h_l\right), \end{cases}$$

where $\Delta W^f_{t_{k+\frac{1}{2}}} = W_{t_{k+\frac{1}{2}}} - W_{t_k}$, $\Delta W^f_{t_{k+1}} = W^f_{t_{k+1}} - W^f_{t_{k+\frac{1}{2}}}$. The antithetic scheme is defined by the same iterative equations, except that the Brownian increment $\Delta W^f_{t_{k+\frac{1}{2}}}$ and $\Delta W^f_{t_{k+1}}$ are swapped

$$\begin{cases} \tilde{X}^{GS,2^l}_{t_{k+\frac{1}{2}}} & = \tilde{X}^{GS,2^l}_{t_k} + b\left(\tilde{X}^{GS,2^l}_{t_k}\right)h_l + \sum_{j=1}^{d}\sigma^j\left(\tilde{X}^{GS,2^l}_{t_k}\right)\Delta W^{j,f}_{t_{k+1}} \\ & +\frac{1}{2}\sum_{j,m=1}^{d}\partial\sigma^j\sigma^m\left(\tilde{X}^{GS,2^l}_{t_k}\right)\left(\Delta W^{j,f}_{t_{k+1}}\Delta W^{m,f}_{t_{k+1}} - \mathbb{1}_{\{m=j\}}h_l\right) \\ \tilde{X}^{GS,2^l}_{t_{k+1}} & = \tilde{X}^{GS,2^l}_{t_{k+\frac{1}{2}}} + b\left(\tilde{X}^{GS,2^l}_{t_{k+\frac{1}{2}}}\right)h_l + \sum_{j=1}^{d}\sigma^j\left(\tilde{X}^{GS,2^l}_{t_{k+\frac{1}{2}}}\right)\Delta W^{j,f}_{t_{k+\frac{1}{2}}} \\ & +\frac{1}{2}\sum_{j,m=1}^{d}\partial\sigma^j\sigma^m\left(\tilde{X}^{GS,2^l}_{t_{k+\frac{1}{2}}}\right)\left(\Delta W^{j,f}_{t_{k+\frac{1}{2}}}\Delta W^{m,f}_{t_{k+\frac{1}{2}}} - \mathbb{1}_{\{m=j\}}h_l\right). \end{cases}$$

Theorem 4.10, Lemma 2.2 and Lemma 4.6 in [23] ensure that $\beta = 2$ under some regularity assumptions on f and the coefficients of the SDE.

Theorem 2.4.2 *Assume that $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ with bounded first and second order derivatives, $b, \sigma^j \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$, for all $j \in \{1, \dots, d\}$, with bounded first and second order derivatives, and that for all $j, m \in \{1, \dots, d\}$, $\partial\sigma^j\sigma^m$ has bounded first order derivatives. Then, for all $p \geq 1$, there exists a constant $c \in \mathbb{R}_+^*$ such that for all $l \in \mathbb{N}^*$,*

$$\mathbb{E}\left[\left|Z^l_{GS}\right|^{2p}\right] \leq \frac{c}{2^{2pl}},$$

where Z^l_{GS} is defined by (2.4.8).

To account for the use of three schemes in the levels $l \in \{1, \dots, L^*\}$ instead of one in level 0 we choose $\lambda_0 = 1$ and $\lambda_l = 5/2$ for all $l \in \{1, \dots, L^*\}$. Then, the multilevel Monte Carlo estimator

$$\hat{Y}^{GS}_{MLMC} = \sum_{l=0}^{L^*}\frac{1}{M^*_l}\sum_{k=0}^{M^*_l} Z^{l,k}_{GS},$$

where L^* and M^*_l are given by (2.4.5) and (2.4.6), respectively, achieves a complexity $O\left(\epsilon^{-2}\right)$.

In [16], Debrabant and Rössler improved the multilevel Monte Carlo method by using, in the last level L, a scheme with high order of weak convergence. Although this modified method attains the same complexity, it reduces the computation time by reducing the bias. We can follow this idea using the Ninomiya-Victoir scheme at the last level L, thereby taking advantage of its order 2 of weak convergence. More precisely, we propose to choose

$$Z^0_{GS} = f\left(X^{GS,1}_T\right),$$

for all $l \in \{1, \dots, L-1\}$,

$$Z^l_{GS} = \frac{1}{2}\left(f\left(\tilde{X}^{GS,2^l}_T\right) + f\left(X^{GS,2^l}_T\right)\right) - f\left(X^{GS,2^{l-1}}_T\right),$$

and

$$\begin{aligned} Z^L_{GS-NV} = & \frac{1}{4}\left(f\left(\tilde{X}^{NV,2^L,\eta}_T\right) + f\left(\tilde{X}^{NV,2^L,-\eta}_T\right) + f\left(X^{NV,2^L,\eta}_T\right) + f\left(X^{NV,2^L,-\eta}_T\right)\right) \\ & - f\left(X^{GS,2^{L-1}}_T\right). \end{aligned} \tag{2.4.9}$$

Here, $\tilde{X}^{NV,2^L,\eta}$ (respectively $\tilde{X}^{NV,2^L,-\eta}$) is the antithetic version of the Ninomiya-Victoir scheme $X^{NV,2^L,\eta}$ (respectively $X^{NV,2^L,-\eta}$), obtained by swapping each successive pair of Brownian increments. Theorem 2.3.2 ensures that in (2.4.4) the order of convergence of the variance at the last level L is 2.

Proposition 2.4.3 *Assume that $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ with bounded first and second order derivatives, $b \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first and second order derivatives, for all $j \in \{1, \ldots, d\}$, $\sigma^j \in \mathcal{C}^3(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first and second order derivatives and with polynomially growing third order derivatives, and that for all $j, m \in \{1, \ldots, d\}$, $\partial\sigma^j\sigma^m$ has bounded first order derivatives. Then, for all $p \geq 1$, there exists a constant $c \in \mathbb{R}_+^*$ such that for all $l \in \mathbb{N}^*$,*

$$\mathbb{E}\left[\left|Z^l_{GS-NV}\right|^{2p}\right] \leq \frac{c}{2^{2pl}},$$

where Z^l_{GS-NV} is defined by (2.4.9).

Proof : Let $p \geq 1$, adding and subtracting

$$\frac{1}{2}\left(f\left(\tilde{X}_T^{GS,2^l}\right) + f\left(X_T^{GS,2^l}\right)\right),$$

and using a convexity inequality, we get

$$\begin{aligned}\left|Z^l_{GS-NV}\right|^{2p} \leq \frac{3^{2p-1}}{2^{2p}}\Bigg(&\left|\frac{1}{2}\left(f\left(X_T^{NV,2^l,\eta}\right) + f\left(X_T^{NV,2^l,-\eta}\right)\right) - f\left(X_T^{GS,2^l}\right)\right|^{2p} \\ &+ \left|\frac{1}{2}\left(f\left(\tilde{X}_T^{NV,2^l,\eta}\right) + f\left(\tilde{X}_T^{NV,2^l,-\eta}\right)\right) - f\left(\tilde{X}_T^{GS,2^l}\right)\right|^{2p}\Bigg) + 3^{2p-1}\left|Z^l_{GS}\right|^{2p}.\end{aligned}$$

However, $\left(X_T^{NV,2^l,\eta}, X_T^{NV,2^l,-\eta}, X_T^{GS,2^l}\right)$ and $\left(\tilde{X}_T^{NV,2^l,\eta}, \tilde{X}_T^{NV,2^l,-\eta}, \tilde{X}_T^{GS,2^l}\right)$ have exactly the same distribution. Then, by taking the expectation we obtain

$$\begin{aligned}\mathbb{E}\left[\left|Z^l_{GS-NV}\right|^{2p}\right] \leq \frac{3^{2p-1}}{2^{2p-1}}&\mathbb{E}\left[\left|\frac{1}{2}\left(f\left(X_T^{NV,2^l,\eta}\right) + f\left(X_T^{NV,2^l,-\eta}\right)\right) - f\left(X_T^{GS,2^l}\right)\right|^{2p}\right] \\ &+ 3^{2p-1}\mathbb{E}\left[\left|Z^l_{GS}\right|^{2p}\right].\end{aligned}$$

Denoting

$$\bar{X}_T^{NV,2^l,\eta} = \frac{1}{2}\left(X_T^{NV,2^l,\eta} + X_T^{NV,2^l,-\eta}\right),$$

and performing a second order Taylor expansion as in Lemma 2.2 in [23], we get a constant $C \in \mathbb{R}_+^*$, which only depends on f and p, such that

$$\begin{aligned}\mathbb{E}\left[\left|Z^l_{GS-NV}\right|^{2p}\right] \leq C\Bigg(&\mathbb{E}\left[\left\|\bar{X}_T^{NV,2^l,\eta} - X_T^{GS,2^l}\right\|^{2p}\right] + \mathbb{E}\left[\left\|X_T^{NV,2^l,\eta} - X_T^{NV,2^l,-\eta}\right\|^{4p}\right] \\ &+ \mathbb{E}\left[\left|Z^l_{GS}\right|^{2p}\right]\Bigg).\end{aligned}$$

Introducing the exact solution X at time T in $\mathbb{E}\left[\left\|X_T^{NV,2^l,\eta} - X_T^{NV,2^l,-\eta}\right\|^{4p}\right]$, we get

$$\mathbb{E}\left[\left\|X_T^{NV,2^l,\eta} - X_T^{NV,2^l,-\eta}\right\|^{4p}\right] \leq 2^{4p-1}\left(\mathbb{E}\left[\left\|X_T^{NV,2^l,\eta} - X_T\right\|^{4p}\right] + \mathbb{E}\left[\left\|X_T - X_T^{NV,2^l,-\eta}\right\|^{4p}\right]\right).$$

Since $\left(X_T^{NV,2^l,\eta}, X_T\right)$ and $\left(X_T^{NV,2^l,-\eta}, X_T\right)$ have the same distribution, we deduce that

$$\mathbb{E}\left[\left\|X_T^{NV,2^l,\eta} - X_T^{NV,2^l,-\eta}\right\|^{4p}\right] \leq 2^{4p}\mathbb{E}\left[\left\|X_T^{NV,2^l,\eta} - X_T\right\|^{4p}\right].$$

Hence

$$\mathbb{E}\left[\left|Z_{GS-NV}^l\right|^{2p}\right] \leq 2^{4p}C\left(\mathbb{E}\left[\left\|\bar{X}_T^{NV,2^l,\eta} - X_T^{GS,2^l}\right\|^{2p}\right] + \mathbb{E}\left[\left\|X_T^{NV,2^l,\eta} - X_T\right\|^{4p}\right] + \mathbb{E}\left[\left|Z_{GS}^l\right|^{2p}\right]\right).$$

Then we conclude using Theorems 2.2.3, 2.3.2 and 2.4.2. ∎

Exploiting the telescoping summation, one can change the constraint (2.4.2) on the last level L and assume:

$$\mathbb{E}\left[Z^L\right] = \mathbb{E}\left[f\left(\hat{X}_T^{2^L}\right) - f\left(X_T^{2^{L-1}}\right)\right].$$

Here $\hat{X}$ is an other scheme, and to be consistent, (2.4.3) becomes

$$\mathbb{E}\left[f\left(\hat{X}_T^{2^l}\right)\right] - Y = \frac{c_1}{2^{\alpha l}} + o\left(\frac{1}{2^{\alpha l}}\right).$$

Then we propose to use the estimator

$$\hat{Y}_{MLMC}^{GS-NV} = \sum_{l=0}^{L-1}\frac{1}{M_l}\sum_{k=1}^{M_l} Z_{GS}^{l,k} + \frac{1}{M_L}\sum_{k=1}^{M_L} Z_{GS-NV}^{L,k}.$$

Of course, the bias of this estimator is given by the bias of the Ninomiya-Victoir scheme. Thanks to its weak order 2, we hope to decrease the value of L, and so to reduce the computation time. We can also use the Ninomiya-Victoir scheme at each level and choose $\left(Z_{NV}^l\right)_{0\leq l\leq L}$, as follows

$$Z_{NV}^0 = f\left(X_T^{NV,1,\eta}\right), \tag{2.4.10}$$

or

$$Z_{NV}^0 = \frac{1}{2}\left(f\left(X_T^{NV,1,\eta}\right) + f\left(X_T^{NV,1,-\eta}\right)\right), \tag{2.4.11}$$

and for all $l \in \{1, \ldots, L\}$,

$$\begin{aligned} Z_{NV}^l = &\frac{1}{4}\left(f\left(\tilde{X}_T^{NV,2^l,\eta}\right) + f\left(\tilde{X}_T^{NV,2^l,-\eta}\right) + f\left(X_T^{NV,2^l,\eta}\right) + f\left(X_T^{NV,2^l,-\eta}\right)\right) \\ &- \frac{1}{2}\left(f\left(X_T^{NV,2^{l-1},\eta}\right) + f\left(X_T^{NV,2^{l-1},-\eta}\right)\right). \end{aligned} \tag{2.4.12}$$

Actually, there is an abuse of notation in (2.4.12), we use the same notation η for the 2^l-dimensional vector $(\eta_1, \ldots, \eta_{2^l})$ of the independent and identically distributed Rademacher random variables needed to generate the Ninomiya-Victoir scheme on the fine grid with 2^l steps and for the 2^{l-1}-dimensional subvector $(\eta_1, \eta_3 \ldots, \eta_{2^l-1})$ used to generate the Ninomiya-Victoir scheme on the coarse grid with 2^{l-1} steps. The extraction of the 2^{l-1}-dimensional vector from the 2^l-dimensional one is aimed at reducing the variance. As previously, we obtain the same rates α and β, but the main drawback is the simulation of six schemes at each level $l \in \{1, \ldots, L-1\}$

instead of three. Reasoning like in the proof of Proposition 2.4.3, since

$$Z_{NV}^{l} = Z_{GS-NV}^{l} + f\left(\bar{X}_{T}^{NV,2^{l-1},\eta}\right) - \frac{1}{2}\left(f\left(X_{T}^{NV,2^{l-1},\eta}\right) + f\left(X_{T}^{NV,2^{l-1},-\eta}\right)\right) + f\left(X_{T}^{GS,2^{l-1}}\right) \\ - f\left(\bar{X}_{T}^{NV,2^{l-1},\eta}\right),$$

one obtains.

Proposition 2.4.4 *Assume that $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ with bounded first and second order derivatives, $b \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first and second order derivatives, for all $j \in \{1, \ldots, d\}$, $\sigma^j \in \mathcal{C}^3(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first and second order derivatives and with polynomially growing third order derivatives, and that for all $j, m \in \{1, \ldots, d\}$, $\partial\sigma^j\sigma^m$ has bounded first order derivatives. Then, for all $p \geq 1$, there exists a constant $c \in \mathbb{R}_+^*$ such that for all $l \in \mathbb{N}^*$,*

$$\mathbb{E}\left[\left|Z_{NV}^{l}\right|^{2p}\right] \leq \frac{c}{2^{2pl}},$$

where Z_{NV}^{l} is defined by (2.4.12).

2.4.2 Multilevel Richardson-Romberg extrapolation

Recently, in [38] Lemaire and Pagès developed a new method called multilevel Richardson-Romberg extrapolation (ML2R). This method combines the ideas behind the multilevel Monte Carlo approach and the multi-step Richardson-Romberg extrapolation introduced in [46]. Actually, the multilevel Richardson-Romberg extrapolation can be seen as a weighted version of the multilevel Monte Carlo estimator. Adapting the notation of Lemaire and Pagès [38], the multilevel Richardson-Romberg extrapolation estimator is built as follows

$$\hat{Y}_{ML2R} = \sum_{l=0}^{L} \frac{W_l}{M_l} \sum_{k=0}^{M_l} Z_k^l,$$

where $\left(Z_k^l\right)_{0 \leq l \leq L, 1 \leq k \leq M_l}$ are independent random variables satisfying (2.4.1), (2.4.2) and a bias error expansion: there exist $\alpha \in \mathbb{R}_+^*$, $R \in \mathbb{N}^*$ and $c'_1, \ldots, c'_R \in \mathbb{R}$ such that for all $l \in \mathbb{N}$,

$$\mathbb{E}\left[f\left(X_T^{2^l}\right)\right] - Y = \sum_{j=1}^{R} c'_j h_l^{\alpha j} + O\left(h_l^{\alpha(R+1)}\right), \tag{2.4.13}$$

where $h_l = T/2^l$ is the time step. As previously, α is the order of weak convergence of the discretization scheme. By introducing the weights $(W_l)_{0 \leq l \leq L}$, one can get a smaller bias[1] by canceling the successive bias terms in the expansion (2.4.13). Following [38], the computational complexity of $\hat{Y}_{ML2R}$, denoted by $\mathcal{C}_{ML2R}$ is defined as $\mathcal{C}_{MLMC}$, except that we do not take into account the weights $(\lambda_l)_{0 \leq l \leq L}$. Under some assumptions (see [38] for further information), the optimal complexity $\mathcal{C}_{ML2R}^*$ is given by Theorem 3.11 in [38], which states that $\mathcal{C}_{ML2R}^*$ depends on α, and the variance convergence rate [2] of Z^l, denoted as previously by β: there exists $c_2 \in \mathbb{R}_+$

[1] See [38] and [46] for more details.
[2] In [38], Lemaire and G. Pagès assume that for all $l \in \mathbb{N}^*$ there exists $V_1 \in \mathbb{R}_+$ such that

$$\mathbb{E}\left[\left\|f\left(X_T^{2^l}\right) - f\left(X_T\right)\right\|^2\right] \leq V_1 h_l^{\beta}.$$

One can easily adapt the proof with assumption (2.4.14).

such that for all $l \in \mathbb{N}^*$,

$$\mathbb{V}\left(Z^l\right) \leq \frac{c_2}{2^{\beta l}}. \tag{2.4.14}$$

Here, we recall optimal complexities for the multilevel Richardson-Romberg extrapolation estimator:

- $\mathcal{C}^*_{ML2R} = O\left(\epsilon^{-2}\right)$ if $\beta > 1$,
- $\mathcal{C}^*_{ML2R} = O\left(\epsilon^{-2}\log\left(\frac{1}{\epsilon}\right)\right)$ if $\beta = 1$,
- $\mathcal{C}^*_{ML2R} = O\left(\epsilon^{-2}\exp\left(-\frac{\beta-1}{\sqrt{\alpha}}\sqrt{2\log(2)\log\left(\frac{1}{\epsilon}\right)}\right)\right)$ if $\beta < 1$.

Similarly to the multilevel Monte Carlo method, the best complexity, obtained when $\beta > 1$, is the same as in a simple Monte Carlo method with independent and identically distributed unbiased random variables. With a view to achieving this complexity by applying Theorem 2.4.2 or Proposition 2.4.4 we will choose $\left(Z^l_{GS}\right)_{0\leq l\leq L}$ and $\left(Z^l_{NV}\right)_{0\leq l\leq L}$ with $Z^0_{NV} = f\left(X^{NV,1,\eta}_T\right)$. Here, we recall the asymptotic[3] optimal parameters for the multilevel Richardson-Romberg extrapolation estimator:

$$L^* = \left\lfloor \sqrt{\left(\frac{1}{2} + \log_2(T)\right)^2 + \frac{2}{\alpha}\log_2\left(\frac{\sqrt{1+4\alpha}}{\epsilon}\right)} + \log_2(T) - \frac{1}{2}\right\rfloor, \tag{2.4.15}$$

$$M^*_l = \lceil q^*_l N^* \rceil, \tag{2.4.16}$$

$$W_l = \sum_{j=l}^{L^*} w_j, \tag{2.4.17}$$

where

$$w_j = (-1)^{L^*-j}\,\frac{2^{-\frac{\alpha}{2}(L^*-j)(L^*-j+1)}}{\prod_{k=1}^{j}\left(1-2^{-k\alpha}\right)\prod_{k=1}^{L^*-j}\left(1-2^{-k\alpha}\right)}, \tag{2.4.18}$$

$$\begin{cases} q^*_0 \propto (1+\theta) \\ q^*_l \propto \theta\,|W_l|\,\frac{2^{-\frac{\beta}{2}l}+2^{-\frac{\beta}{2}(l-1)}}{\sqrt{2^l+2^{l-1}}}, \text{ for all } l \in \{1,\ldots,L^*\}, \\ \sum_{l=0}^{L^*} q^*_l = 1, \end{cases} \tag{2.4.19}$$

$$N^* = \left(1 + \frac{1}{2\alpha(L^*+1)}\right)\frac{\mathbb{V}(f(X_T))\left(1+\theta\left(1+\sum_{l=1}^{L^*}|W_l|\left(2^{-\frac{\beta}{2}l}+2^{-\frac{\beta}{2}(l-1)}\right)\sqrt{2^l+2^{l-1}}\right)\right)^2}{\epsilon^2\left(q^*_0+\sum_{l=1}^{L^*}q^*_l\left(2^l+2^{l-1}\right)\right)}, \tag{2.4.20}$$

and

$$\theta = T^{-\frac{\beta}{2}}\sqrt{\frac{c_2}{\mathbb{V}(f(X_T))}}. \tag{2.4.21}$$

2.4.3 Numerical experiments

In this section we present numerical tests in which we compare the multilevel Monte Carlo and the multilevel Richardson-Romberg estimators. Although we have not proved a theoretical

[3] When ϵ goes to 0.

expansion of the bias like (2.4.13) for the Ninomiya-Victoir and the Giles-Szpruch schemes, we will use these schemes in the multilevel Richardson-Romberg estimators (see [20] and [45] for extrapolation methods based on the Ninomiya-Victoir scheme). More precisely, we compare the following estimators:

- The multilevel Monte Carlo estimator with the Giles-Szpruch scheme

$$\hat{Y}^{GS}_{MLMC} = \sum_{l=0}^{L^*} \frac{1}{M_l^*} \sum_{k=1}^{M_l^*} Z^{l,k}_{GS},$$

  where Z^0_{GS} and Z^l_{GS} are respectively given by (2.4.7) and (2.4.8).

- The multilevel Monte Carlo estimator with the Ninomiya-Victoir scheme

$$\hat{Y}^{NV}_{MLMC} = \sum_{l=0}^{L^*} \frac{1}{M_l^*} \sum_{k=1}^{M_l^*} Z^{l,k}_{NV},$$

  where Z^0_{NV} and Z^l_{NV} are respectively given by (2.4.10) or[4] (2.4.11) and (2.4.12).

- The multilevel Monte Carlo estimator with the Giles-Szpruch scheme from level 0 to level $L^* - 1$, and the coupling between the Ninomiya-Victoir and the Giles-Szpruch scheme at the last level L^*

$$\hat{Y}^{GS-NV}_{MLMC} = \sum_{l=0}^{L^*-1} \frac{1}{M_l^*} \sum_{k=1}^{M_l^*} Z^{l,k}_{GS} + \frac{1}{M^*_{L^*}} \sum_{k=1}^{M^*_{L^*}} Z^{L^*,k}_{GS-NV},$$

  where $Z^{L^*}_{GS-NV}$ is given by (2.4.9).

- The multilevel Richardson-Romberg estimator with the Giles-Szpruch scheme

$$\hat{Y}^{GS}_{ML2R} = \sum_{l=0}^{L^*} \frac{W_l}{M_l^*} \sum_{k=1}^{M_l^*} Z^{l,k}_{GS}.$$

- The multilevel Richardson-Romberg estimator with the Ninomiya-Victoir scheme

$$\hat{Y}^{NV}_{ML2R} = \sum_{l=0}^{L^*} \frac{W_l}{M_l^*} \sum_{k=1}^{M_l^*} Z^{l,k}_{NV}.$$

  Here, Z^0_{NV} is given by (2.4.10).

Clark-Cameron SDE

For our first numerical test, we consider the Clark-Cameron SDE with drift which is defined as follows

$$\begin{cases} dU_t = S_t dW_t^1 \\ dS_t = \mu dt + dW_t^2, \end{cases} \tag{2.4.22}$$

[4]The choice of level 0 will be discussed later.

where $\mu \in \mathbb{R}$. In this $2-$dimensional stochastic differential equation, the diffusion coefficients are given by

$$\sigma^1 \begin{pmatrix} u \\ s \end{pmatrix} = \begin{pmatrix} s \\ 0 \end{pmatrix} \text{ and } \sigma^2 \begin{pmatrix} u \\ s \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and the drift coefficient is

$$b \begin{pmatrix} u \\ s \end{pmatrix} = \begin{pmatrix} 0 \\ \mu \end{pmatrix}.$$

The Stratonovich drift is given by

$$\sigma^0 \begin{pmatrix} u \\ s \end{pmatrix} = \left(b - \frac{1}{2}\left(\partial\sigma^1\sigma^1 + \partial\sigma^2\sigma^2\right)\right) \begin{pmatrix} u \\ s \end{pmatrix} = \begin{pmatrix} 0 \\ \mu \end{pmatrix} - \frac{1}{2}\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} s \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ \mu \end{pmatrix}.$$

These functions are smooth and satisfy the assumptions of Theorems 2.2.3, 2.3.2, Propositions 2.4.3 and 2.4.4. By a straightforward calculation, the Giles-Szpruch scheme is given by

$$\begin{cases} U^{GS}_{t_{k+1}} = U^{GS}_{t_k} + S^{GS}_{t_k}\left(W^1_{t_{k+1}} - W^1_{t_k}\right) + \frac{1}{2}\left(W^1_{t_{k+1}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_k}\right) \\ S^{GS}_{t_{k+1}} = S^{GS}_{t_k} + \mu\left(t_{k+1} - t_k\right) + \left(W^2_{t_{k+1}} - W^2_{t_k}\right), \end{cases}$$

and the Ninomiya-Victoir scheme is given by

$$\begin{cases} U^{NV,\eta}_{t_{k+1}} = U^{NV,\eta}_{t_k} + S^{NV,\eta}_{t_k}\left(W^1_{t_{k+1}} - W^1_{t_k}\right) + \frac{1}{2}\mu\left(t_{k+1} - t_k\right)\left(W^1_{t_{k+1}} - W^1_{t_k}\right) \\ \qquad\qquad + \mathbb{1}_{\{\eta_{k+1}=1\}}\left(W^1_{t_{k+1}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_k}\right) \\ S^{NV,\eta}_{t_{k+1}} = S^{NV,\eta}_{t_k} + \mu\left(t_{k+1} - t_k\right) + \left(W^2_{t_{k+1}} - W^2_{t_k}\right). \end{cases}$$

Before comparing these estimators, we will illustrate Theorems 2.2.3, 2.3.2, Propositions 2.4.3 and 2.4.4. In order to check the strong convergence rate of the Ninomiya-Victoir scheme, we will look at the expectation of the square L^2-norm of the difference, at time T, between the schemes with steps h_l, and h_{l-1}, simulated with the same Brownian path. Denoting by

$$X^{NV,2^l,\eta}_T = \left(U^{NV,2^l,\eta}_T, S^{NV,2^l,\eta}_T\right),$$

it follows, from Theorem 2.2.3 that

$$\mathbb{E}\left[\left\|X^{NV,2^l,\eta}_T - X^{NV,2^{l-1},\eta}_T\right\|^2\right] \leq \frac{c}{2^l}.$$

For the simulations, we choose the initial conditions $U_0 = V_0 = 0$, the final time $T = 1$ and the parameter $\mu = 1$. In figure 2.1, the blue line shows the behavior of

$$\log_2\left(\mathbb{E}\left[\left\|X^{NV,2^l,\eta}_T - X^{NV,2^{l-1},\eta}_T\right\|^2\right]\right)$$

and the red line shows the behavior of

$$\log_2\left(\mathbb{E}\left[\left\|\bar{X}^{NV,2^l,\eta}_T - X^{GS,2^l,\eta}_T\right\|^2\right]\right)$$

as a function of the discretization level l. These expectations are estimated with a standard Monte Carlo method with $M_l = 10^6$ samples for all l. This choice ensures that the confidence intervals are very tight, that is why they are not represented in our plot. The blue line illustrates the strong convergence order of the Ninomiya-Victoir scheme. As expected, we obtain a line with slope -1. The red line illustrates the strong convergence order of the coupling between the Ninomiya-Victoir and the Giles-Szpruch scheme. It follows, from Theorem 2.3.2 that

$$\mathbb{E}\left[\left\|\bar{X}_T^{NV,2^l,\eta} - X_T^{GS,2^l}\right\|^2\right] \leq \frac{c}{2^{2l}}.$$

Again, as expected, we obtain a line with slope -2. These numerical results are consistent with Theorems 2.2.3 and 2.3.2 stated and proved in this paper.

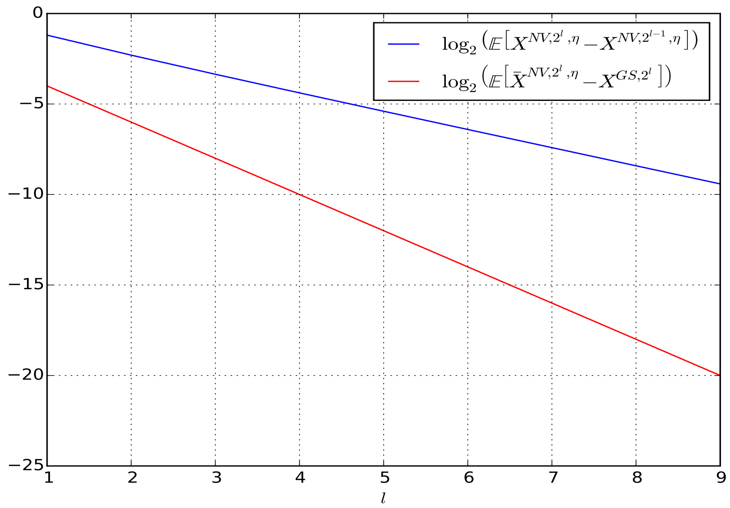

Fig. 2.1 Strong convergence order. Strong error (y-axis $\log_2$ scale) as a function of l (x-axis).

To illustrate Propositions 2.4.3 and 2.4.4, we choose a smooth payoff function, satisfying the assumptions of Propositions 2.4.3 and 2.4.4: $f(u,s) = \cos(u)$. In figure 2.2, the red line shows the behavior of

$$\log_2\left(\mathbb{E}\left[\left(Z^l_{GS-NV}\right)^2\right]\right)$$

defined by (2.4.9) whereas the blue line shows the behavior of

$$\log_2\left(\mathbb{E}\left[\left(Z^l_{NV}\right)^2\right]\right)$$

defined by (2.4.12). Both lines have slope -2.

By increasing the value of μ, we noticed that the theoretical rate of convergence is reached for larger and larger values of l. For small values of l, the variance decreases faster than the theoretical rate. Figure 2.3 shows this phenomenon for Z^l_{NV}, with the payoff $f(u,s) = u^2$. Actually, by choosing this payoff we can check that

$$\mathbb{E}\left[\left(Z^l_{NV}\right)^2\right] = 2^{-4l}\left(\frac{3}{16}\mu^4T^6 + \frac{9}{16}\mu^2T^5\right) + 2^{-3l}\left(\frac{11}{64}\mu^2T^5 + \frac{1545}{512}T^4\right) + 2^{-2l}\left(\frac{163}{1024}T^4\right). \tag{2.4.23}$$

The details of this tedious calculation are postponed to the Appendix. The previous formula (2.4.23) contains higher order terms which overshadow the theoretical behavior of the variance.

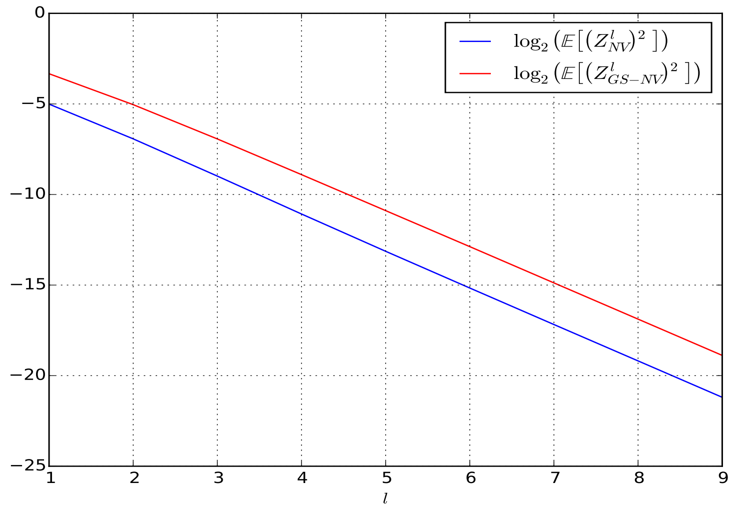

Fig. 2.2 Variance convergence order with $f(u,s) = cos(u)$. Second order moment (y-axis $\log_2$ scale) as a function of l (x-axis).

The following plot shows the behavior of $\log_2\left(\mathbb{E}\left[\left(Z^l_{NV}\right)^2\right]\right)$ as a function of l. For large values of μ and for small values of l, the ratio

$$\mathbb{E}\left[\left(Z^{l+1}_{NV}\right)^2\right] \Big/ \mathbb{E}\left[\left(Z^l_{NV}\right)^2\right]$$

is close to 16, which shows that the leading term is 2^{-4l}. Asymptotically, the slopes of the curves are 2. From a numerical point of view and given the structure of multilevel methods, this is an important point to emphasize. In particular, the choice (2.4.16) of parameters $(M^*_l)_{0\leq l\leq L^*}$ in the multilevel Richardson-Romberg estimator is based on asymptotic properties and will not be optimal when this asymptotic behavior fails for the first levels.

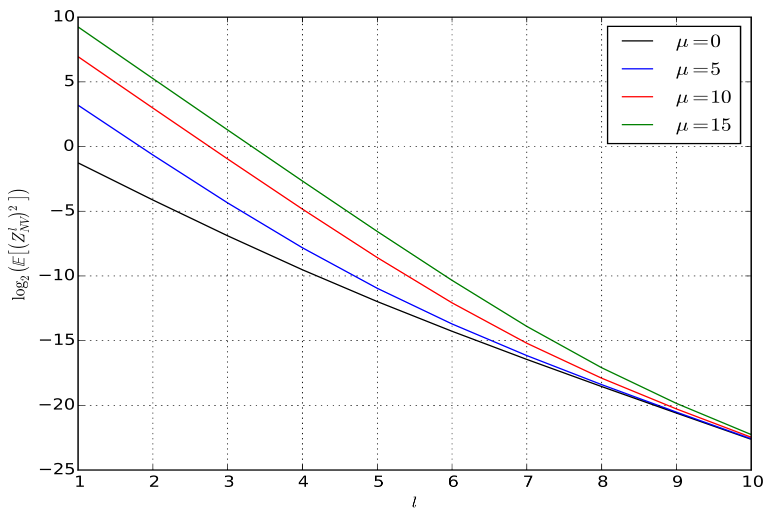

Fig. 2.3 Variance convergence order with $f(u,s) = u^2$. Second order moment (y-axis $\log_2$ scale) as a function of l (x-axis).

Now we present the practical procedure used to implement the multilevel estimators. Putting together the elements already discussed, the algorithm that we use for the multilevel Monte Carlo with the Ninomiya-Victoir scheme or the Giles-Szpruch scheme is as follows. We begin by estimating the weak error constant c_1 in (2.4.3), the constant c_2 which comes from the variance estimation (2.4.4) and checking the orders of weak and strong convergence. When the asymptotic behavior (2.4.3) of the bias of the scheme is satisfied, one has

$$\mathbb{E}\left[Z^l\right] \sim \frac{c_1\left(1-2^{\alpha}\right)}{2^{\alpha l}}.$$

Using a regression with few values of $\left(l, \left|\mathbb{E}\left[Z^l\right]\right|\right)$, we estimate c_1 and check the order α of weak convergence. In the same way, we estimate c_2 and check the strong order β of variance convergence to 0, using a regression in (2.4.4). Then we estimate $\mathbb{V}\left(Z^0\right)$ using a standard Monte Carlo estimator $\hat{V}_0$. After that, for a given ϵ we define L^* using (2.4.5) then we set

$$M_0^* = \left\lceil \frac{2}{\epsilon^2}\sqrt{\frac{\hat{V}^0}{\lambda_0}}\left(\sqrt{\lambda_0\hat{V}^0} + \sum_{j=1}^{L^*}\sqrt{c_2\lambda_j 2^{j(1-\beta)}}\right)\right\rceil, \tag{2.4.24}$$

and for all $l \in \{1,\ldots,L^*\}$,

$$M_l^* = \left\lceil \frac{2}{\epsilon^2}\sqrt{\frac{c_2}{\lambda_l 2^{l(\beta+1)}}}\left(\sqrt{\lambda_0\hat{V}^0} + \sum_{j=1}^{L^*}\sqrt{c_2\lambda_j 2^{j(1-\beta)}}\right)\right\rceil. \tag{2.4.25}$$

When we use the Ninomiya-Victoir scheme we have the choice between

$$Z^0_{NV} = f\left(X_T^{NV,1,\eta}\right),$$

and

$$Z^0_{NV} = \frac{1}{2}\left(f\left(X_T^{NV,1,\eta}\right) + f\left(X_T^{NV,1,-\eta}\right)\right).$$

The second choice reduces the variance of level 0 if $X_T^{NV,1,\eta}$ effectively depends on η. So, in general, using

$$Z^0_{NV} = \frac{1}{2}\left(f\left(X_T^{NV,1,\eta}\right) + f\left(X_T^{NV,1,-\eta}\right)\right)$$

reduces the sample size of the multilevel Monte Carlo estimator. Thus, although we use two schemes in the level 0, the method is slightly faster with this choice in practice. As already mentioned, for the Giles-Szpruch scheme we choose $\lambda_0 = 1$ and for all $l \in \{1,\ldots,L^*\}, \lambda_l = 5/2$ to balance the lower cost of level $l = 0$. Following this idea, for the Ninomiya-Victoir scheme we choose the same sequence if

$$Z^0_{NV} = \frac{1}{2}\left(f\left(X_T^{NV,1,\eta}\right) + f\left(X_T^{NV,1,-\eta}\right)\right),$$

and we propose to choose $\lambda_0 = 1$ and and for all $l \in \{1,\ldots,L^*\}, \lambda_l = 5$, if

$$Z^0_{NV} = f\left(X_T^{NV,1,\eta}\right).$$

Let us discuss the implementation of the multilevel Monte Carlo estimator with the Giles-Szpruch scheme from level 0 to level $L^* - 1$ and the coupling between the Ninomiya-Victoir and the Giles-Szpruch scheme at the last level L^*. The practical procedure is slightly different. As already discussed, in the case of $\hat{Y}^{GS-NV}_{MLMC}$ the bias is given by the bias of the Ninomiya-Victoir scheme, so we begin with the estimation of the weak error constant c_1 using the Ninomiya-Victoir scheme. The next step is to estimate the constant c_2 using the Giles-Szpruch scheme. Then, we estimate $\mathbb{V}\left(Z^0_{GS}\right)$ (respectively $\mathbb{V}\left(Z^{L^*}_{GS-NV}\right)$) using a standard Monte Carlo estimator $\hat{V}^0_{GS}$ (respectively $\hat{V}^{L^*}_{GS-NV}$). Finally, we define L^* using (2.4.5) and set

$$M_0^* = \left\lceil \frac{2}{\epsilon^2}\sqrt{\frac{\hat{V}^0_{GS}}{\lambda_0}}\left(\sqrt{\lambda_0 \hat{V}^0_{GS}} + \sum_{j=1}^{L^*-1}\sqrt{c_2\lambda_j 2^{j(1-\beta)}} + \sqrt{\lambda_{L^*}2^{L^*}\hat{V}^{L^*}_{GS-NV}}\right)\right\rceil, \tag{2.4.26}$$

for all $l \in \{1, \ldots, L^*\}$,

$$M_l^* = \left\lceil \frac{2}{\epsilon^2}\sqrt{\frac{c_2}{\lambda_l 2^{l(\beta+1)}}}\left(\sqrt{\lambda_0 \hat{V}^0_{GS}} + \sum_{j=1}^{L^*-1}\sqrt{c_2\lambda_j 2^{l(1-\beta)}} + \sqrt{\lambda_{L^*}2^{L^*}\hat{V}^{L^*}_{GS-NV}}\right)\right\rceil, \tag{2.4.27}$$

and

$$M_{L^*}^* = \left\lceil \frac{2}{\epsilon^2}\sqrt{\frac{\hat{V}^{L^*}_{GS-NV}}{\lambda_{L^*}2^{L^*}}}\left(\sqrt{\lambda_0 \hat{V}^0_{GS}} + \sum_{j=1}^{L^*-1}\sqrt{c_2\lambda_j 2^{j(1-\beta)}} + \sqrt{\lambda_{L^*}2^{L^*}\hat{V}^{L^*}_{GS-NV}}\right)\right\rceil. \tag{2.4.28}$$

We suggest choosing $\lambda_0 = 1$, $\lambda_l = 5/2$, for all $l \in \{1, \ldots, L^* - 1\}$, and $\lambda_{L^*} = 9/2$ to balance the higher cost of level L^*.

Since all parameters are explicit, implementing the Multilevel Richardson-Romberg estimator is quite simple. As noted in [38], we only need to estimate $\mathbb{V}\left(f\left(X_T\right)\right)$ and the constant c_2 in (2.4.4) which comes from the variance estimation. The variance $\mathbb{V}\left(f\left(X_T\right)\right)$ is estimated using a crude Monte Carlo method.

Now we present our numerical tests in which we compare the computing time of each estimator as a function of the upper bound, denoted by ϵ, on the root mean squared error. For our first test we choose a smooth payoff $f(u, s) = \cos(u)$. We estimate the two constants c_1 and c_2 using the above-mentioned procedure. To compute our regression, we estimate $\mathbb{E}\left[Z^l\right]$ and $\mathbb{V}\left[Z^l\right]$ for $l \in \{1, \ldots, 4\}$, using a standard Monte Carlo method. The sample size used must be adjusted to get a rather good estimate, but without spending too much time during this step. In our numerical experiment, we choose a sample size $M = 10^4$. Using this approach, we estimate the theoretical values of the orders of weak and variance convergences. More precisely we get $\alpha = 1$, $\beta = 2$ for the Giles-Szpruch scheme and $\alpha = 2$, $\beta = 2$ for the Ninomiya-Victoir scheme. In figure 2.4 is depicted the CPU-time in seconds (in $\log_2$ scale) of each multilevel method as a function of ϵ (in $\log_2$ scale). It provides a direct comparison of the performance of the different estimators. The red line is for $\hat{Y}^{GS-NV}_{MLMC}$. This line is below the other lines, which indicates clearly that, for this experiment, $\hat{Y}^{GS-NV}_{MLMC}$ is faster than the other estimators. Moreover, we observe a rather close behavior of the Multilevel Richardson-Romberg estimator and the Multilevel Monte Carlo estimator. Indeed the black line, representing $\hat{Y}^{NV}_{MLMC}$ is close to the black dashed line representing $\hat{Y}^{NV}_{ML2R}$. Similarly the blue line, representing $\hat{Y}^{GS}_{MLMC}$ is close to the blue dashed line representing $\hat{Y}^{GS}_{ML2R}$. Finally, one can notice that all slopes are equal to -2, which indicates that all these estimators achieve an $O\left(\epsilon^{-2}\right)$ complexity.

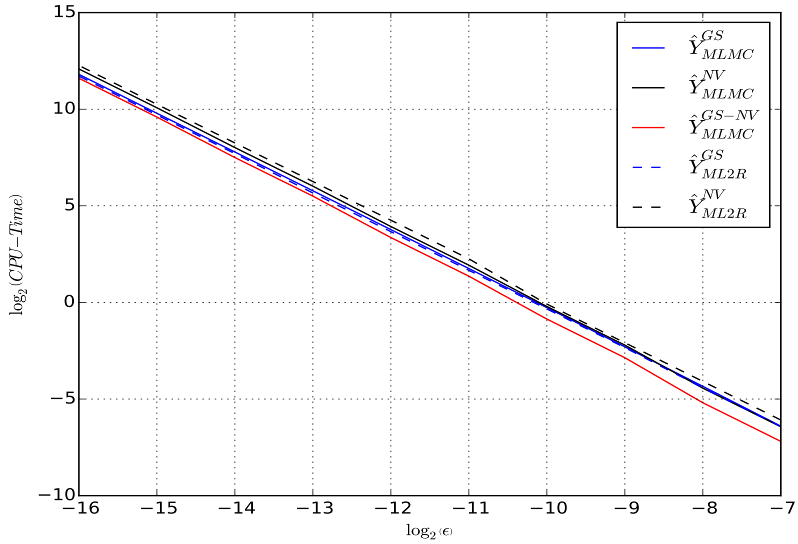

Fig. 2.4 Clark-Cameron SDE with $f(u,s) = cos(u)$, CPU-time in second (y-axis $\log_2$ scale) as a function of ϵ (x-axis $\log_2$ scale).

To measure the efficiency of $\hat{Y}^{GS-NV}_{MLMC}$ with respect to other estimators, we plot in figure 2.5 the following CPU-time ratios:

$$R = \frac{CPU - time\left(\hat{Y}\right)}{CPU - time\left(\hat{Y}^{GS-NV}_{MLMC}\right)}.$$

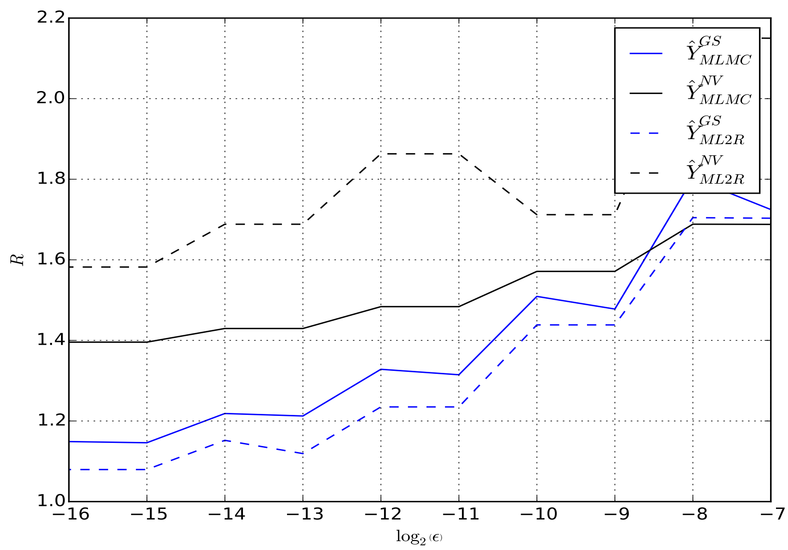

Fig. 2.5 Clark-Cameron SDE with $f(u,s) = cos(u)$, CPU-time ratios (y-axis) as a function of ϵ (x-axis $\log_2$ scale).

The estimator $\hat{Y}^{GS-NV}_{MLMC}$ is about 1.1 to 1.6 faster than $\hat{Y}^{GS}_{MLMC}$ or $\hat{Y}^{GS}_{ML2R}$ when ϵ goes from 2^{-16} to 2^{-7}. In comparison with the estimator $\hat{Y}^{GS-NV}_{MLMC}$, $\hat{Y}^{NV}_{MLMC}$ and $\hat{Y}^{NV}_{ML2R}$ perform poorly.

In order to understand what is going on, let us provide a theoretical calculation of the CPU-time for the multilevel Monte Carlo estimator. Denoting by τ_l the theoretical computing time of level $l \in \{0, \ldots, L^*\}$, one has

$$\tau^l \propto M_l^* 2^l.$$

Replacing[5] M_l^*, one can write

$$\tau^l = C_l(\epsilon) 2^{-l\left(\frac{\beta+1}{2}\right)} 2^l = C_l(\epsilon) 2^{-l\left(\frac{\beta-1}{2}\right)}.$$

The theoretical computing time, denoted by τ, is given by

$$\tau(\epsilon) = \sum_{l=0}^{L^*(\epsilon)} \tau_l = \sum_{l=0}^{L^*(\epsilon)} C_l(\epsilon) 2^{-l\left(\frac{\beta-1}{2}\right)}.$$

In the multilevel Monte Carlo estimator studied in this paper $C_l = C_1$, for all $l \in \{1, \ldots, L^* - 1\}$, then one has

$$\tau(\epsilon) = \begin{cases} C_0(\epsilon) + \frac{C_1(\epsilon)}{1-2^{-\frac{\beta-1}{2}}}\left(2^{-\frac{\beta-1}{2}} - 2^{-L^*(\epsilon)\frac{\beta-1}{2}}\right) + C_{L^*}(\epsilon) 2^{-L^*(\epsilon)\frac{\beta-1}{2}} & \text{if } \beta \neq 1, \\ C_0(\epsilon) + (L^*(\epsilon) - 2) C_1(\epsilon) + C_{L^*}(\epsilon) & \text{if } \beta = 1. \end{cases} \tag{2.4.29}$$

Now, it is easy to understand why $\hat{Y}_{MLMC}^{GS-NV}$ is faster than $\hat{Y}_{MLMC}^{GS}$. As a matter of fact, the two estimators are very close, and in our numerical experiments we observe that

$$C^{GS}(\epsilon) \approx C^{GS-NV}(\epsilon).$$

Since using a scheme with second order of weak convergence provides a lower optimal last level L^*, in view of (2.4.29), we understand why, in general, we can state that

$$\tau_{MLMC}^{GS-NV} \leq \tau_{MLMC}^{GS}.$$

The poor performance of $\hat{Y}_{MLMC}^{NV}$ or $\hat{Y}_{ML2R}^{NV}$ reflects the use of six schemes in Z_{NV}^l.

For our second experiment, we only change the payoff. We choose the non-smooth payoff $f(u, s) = u_+$. Theorem 5.2 in [23] gives the lower bound $\beta = 3/2$ for the Giles-Szpruch scheme. Their proof is, in some ways, generic and it can easily be adapted to the Ninomiya-Victoir scheme. This is enough to keep the $O(\epsilon^{-2})$ complexity. To determine the actual values of β and α, we rely on the numerical results. Using the same automatic process, we get for the Ninomiya-Victoir scheme $\alpha = 3/2$ and $\beta = 3/2$. The non-regularity of the payoff affects both the weak and the variance convergence rates. With regard to the Giles-Szpruch scheme, the regression procedure

[5] Obviously, the constant $C(\epsilon)$ depend on the estimator. For $\hat{Y}_{MLMC}^{GS}$ and $\hat{Y}_{MLMC}^{NV}$, the constants are given by formulas (2.4.24) and (2.4.25):

$$C_0(\epsilon) = \frac{2}{\epsilon^2}\sqrt{\frac{\hat{V}^0}{\lambda_0}}\left(\sqrt{\lambda_0 \hat{V}^0} + \sum_{j=1}^{L^*} \sqrt{c_2 \lambda_j 2^{j(1-\beta)}}\right),$$

and for all $l \in \{1, \ldots, L^*\}$,

$$C_l(\epsilon) = \frac{2}{\epsilon^2}\sqrt{\frac{c_2}{\lambda_l}}\left(\sqrt{\lambda_0 \hat{V}^0} + \sum_{j=1}^{L^*} \sqrt{c_2 \lambda_j 2^{j(1-\beta)}}\right).$$

For $\hat{Y}_{MLMC}^{GS-NV}$ the constans are given by formulas (2.4.26), (2.4.27) and (2.4.28).

leads to $\alpha = 1$ and $\beta = 2$, but the situation is quite confusing. Indeed, we noticed that the asymptotic rate $\beta = 3/2$ is reached for $l \geq \bar{l} = 5$. Figure 2.6 illustrates this inflection. The blue line is the estimation of $\left(\mathbb{V}\left(Z^l_{GS}\right)\right)_{1 \leq l \leq 7}$ whereas the red line is the regression on the first four values. The two lines diverge at level $\bar{l} = 5$, which show clearly the inflection.

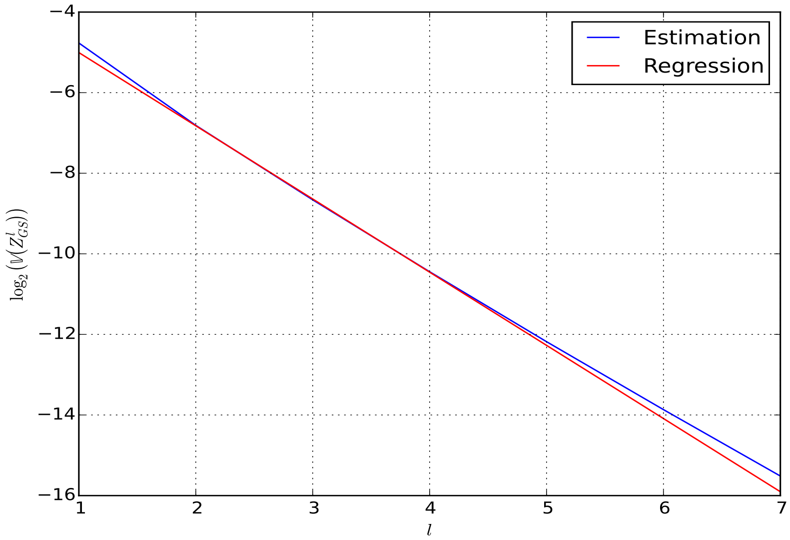

Fig. 2.6 Clark-Cameron SDE with $f(u,s) = u_+$, Variance of the Giles-Szpruch scheme (y-axis $\log_2$ scale) as a function of l (x-axis).

Here, assigning a value for (β, c_2) to implement $\hat{Y}^{GS}_{MLMC}$, $\hat{Y}^{GS}_{ML2R}$ and $\hat{Y}^{GS-NV}_{MLMC}$ by using respectively (2.4.24)-(2.4.25), (2.4.26) to (2.4.28), and (2.4.16) to (2.4.21) may not be convenient. We suggest applying the numerical procedure described in the following remark to implement the multilevel estimators.

Remark 2.4.5 *In the case of the Clark-Cameron SDE with $\mu = 1, U_0 = 0$, $S_0 = 0$ and for a smooth payoff, everything is going as expected, but in some cases (see the Heston model or the Clark-Cameron SDE with a large μ) estimating (β, c_2) may be difficult, especially when the theoretical rate of convergence is reached for a level $\bar{l} \geq 2$ and this may affect the efficiency of the multilevel methods. To get around this problem, a reasonable criterion is to compare $\bar{l}$ and the last level $L^*(\epsilon)$. If $L^*(\epsilon) < \bar{l}$, we decide to use the values obtained by the regression and use the usual formulas[6] to compute $(M^*_l)_{0 \leq l \leq L^*}$ for both methods. If $L^*(\epsilon) \geq \bar{l}$, then we estimate $\mathbb{V}\left[Z^l\right]$ for $l \in \left\{0, \ldots, \bar{l}\right\}$ using standard Monte Carlo Method. Then, we approximate, for $l \in \left\{\bar{l}+1, \ldots, L^*\right\}$, $\mathbb{V}\left[Z^l\right]$ by $2^{\beta(l-\bar{l})}\hat{\mathbb{V}}^{\bar{l}}$ where $\hat{\mathbb{V}}^{\bar{l}}$ is the estimation of $\mathbb{V}\left[Z^{\bar{l}}\right]$ and β is the theoretical order of convergence of the variance. Finally, we compute M^*_l for $l \in \{0, \ldots, L^*\}$ using (2.4.6). As regards the multilevel Richardson-Romberg estimator, we do not recommend its use in this case.*

In our second experiment, the Giles-Szpruch scheme only appears to be problematic. Indeed the values of $L^*(\epsilon)$ are given by

[6]Formulas (2.4.24)-(2.4.25) for $\hat{Y}^{GS}_{MLMC}$ and $\hat{Y}^{NV}_{MLMC}$, (2.4.26) to (2.4.28) for $\hat{Y}^{GS-NV}_{MLMC}$, and (2.4.16) to (2.4.21) for $\hat{Y}_{ML2R}$.

ϵ	2^{-7}	2^{-8}	2^{-9}	2^{-10}	2^{-11}	2^{-12}	2^{-13}	2^{-14}	2^{-15}	2^{-16}
$\hat{Y}^{GS}_{MLMC}$	6	7	8	9	10	11	12	13	14	15
$\hat{Y}^{GS-NV}_{MLMC}$	3	4	4	5	6	6	7	8	8	9
$\hat{Y}^{GS}_{ML2R}$	3	3	4	4	4	4	4	5	5	5

.

If $\epsilon \in \{2^{-14}, 2^{-15}, 2^{-16}\}$, for $\hat{Y}^{GS}_{MLMC}$, since $\bar{l}$ is exactly equal to $L(\epsilon)$, we are in a borderline situation. Nevertheless, we keep in the following figures the performance of this estimator for $\epsilon \in \{2^{-14}, 2^{-15}, 2^{-16}\}$. For the multilevel Monte Carlo estimators with the Giles-Szpruch scheme, we apply the modified procedure of Remark 2.4.5 if necessary. Figure 2.7 compares the computing time of the estimator, with the previous graphical conventions.

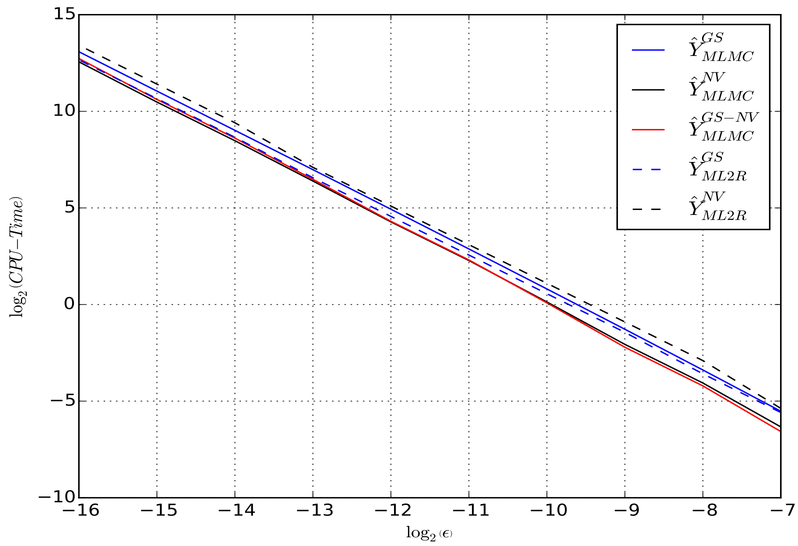

Fig. 2.7 Clark-Cameron SDE with $f(u, s) = u_+$, CPU-time in second (y-axis $\log_2$ scale) as a function of ϵ (x-axis $\log_2$ scale).

Unlike the previous experiment, the two fastest estimators are $\hat{Y}^{NV}_{MLMC}$ and $\hat{Y}^{GS-NV}_{MLMC}$. Although we lose the second order of weak convergence, the estimator $\hat{Y}^{GS-NV}_{MLMC}$ is about 1.3 to 2 faster than $\hat{Y}^{GS}_{MLMC}$. This is due to the degradation of the variance convergence order β from 2 to 3/2 in comparison with a smooth payoff. Indeed, thanks to formula (2.4.29), one can see that, in the multilevel Monte Carlo methods, all things being equal, the gain in computing time due to the introduction of a scheme with high order of weak convergence in the last level is all the more significant that β is small. This explains why $\hat{Y}^{NV}_{MLMC}$ performs very well. Despite using six schemes in Z^l_{NV}, $\hat{Y}^{NV}_{MLMC}$ goes up to 1.1 faster than $\hat{Y}^{GS-NV}_{MLMC}$ (see Figure 2.8). In contrast, the use of a scheme with high order of weak convergence like Z^l_{NV} in the multilevel Richardson-Romberg does not appear to counterbalance its complexity. This difference of behavior is related to the dependence of $L^*(\epsilon)$ on α . In the multilevel Monte Carlo estimators, the dependence is of the form $1/\alpha$ (see (2.4.5)) which provides better results as alpha increases than the multilevel Richardson-Romberg estimator where the dependence on α is given by (2.4.15).

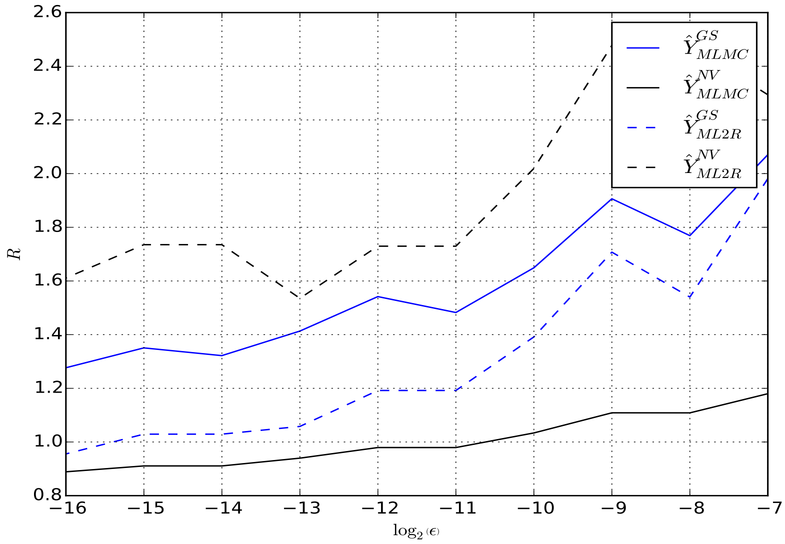

Fig. 2.8 Clark-Cameron SDE with $f(u, s) = u+$, CPU-time ratios (y-axis) as a function of ϵ (x-axis $\log_2$ scale).

Heston model

The Heston model is an asset price model which assumes that volatility, denoted by V, evolves according to an autonomous Cox-Ingersoll-Ross SDE:

$$\begin{cases} dU_t = (r - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t^1 \\ dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t^2. \end{cases} \tag{2.4.30}$$

The asset price S is given by $S_t = \exp(U_t)$. We assume, for simplicity, no correlation between the Brownian motion driving the asset price and the volatility process. We also assume that $2\kappa\theta \geq \sigma^2$ to ensure that the zero boundary is not attainable for the volatility process. The main difficulty is located in 0, where the square root is not Lipschitz. In this 2−dimensional model, the diffusion coefficients are given by

$$\sigma^1 \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \sqrt{v} \\ 0 \end{pmatrix} \text{ and } \sigma^2 \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ \sigma\sqrt{v} \end{pmatrix},$$

and the drift coefficient is

$$b \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} r - \frac{1}{2}v \\ \kappa\,(\theta - v) \end{pmatrix}.$$

The Stratonovich drift is given by

$$
\begin{aligned}
\sigma^0\begin{pmatrix}u\\v\end{pmatrix} &= \left(b-\frac{1}{2}\left(\partial\sigma^1\sigma^1+\partial\sigma^2\sigma^2\right)\right)\begin{pmatrix}u\\v\end{pmatrix}\\
&= \begin{pmatrix}r-\frac{1}{2}v\\ \kappa\left(\theta-v\right)\end{pmatrix}-\frac{1}{2}\left(\begin{pmatrix}0 & \frac{1}{2\sqrt{v}}\\ 0 & 0\end{pmatrix}\begin{pmatrix}\sqrt{v}\\ 0\end{pmatrix}+\begin{pmatrix}0 & 0\\ 0 & \frac{\sigma}{2\sqrt{v}}\end{pmatrix}\begin{pmatrix}0\\ \sigma\sqrt{v}\end{pmatrix}\right)\\
&= \begin{pmatrix}r-\frac{1}{2}v\\ \kappa\left(\theta-v\right)-\frac{\sigma^2}{4}\end{pmatrix}.
\end{aligned}
$$

Then, the Giles-Szpruch scheme is given by

$$
\begin{cases}
V^{GS}_{t_{k+1}} = V^{GS}_{t_k}+\kappa\left(\theta-V^{GS}_{t_k}\right)h+\sigma\sqrt{V^{GS}_{t_k}}\Delta W^2_{t_{k+1}}+\frac{1}{4}\sigma^2\left(\left(\Delta W^2_{t_{k+1}}\right)^2-h\right)\\
U^{GS}_{t_{k+1}} = U^{GS}_{t_k}+\left(r-\frac{1}{2}V^{GS}_{t_k}\right)h+\sqrt{V^{GS}_{t_k}}\Delta W^1_{t_{k+1}}+\frac{1}{4}\sigma\Delta W^1_{t_{k+1}}\Delta W^2_{t_{k+1}}.
\end{cases}
\tag{2.4.31}
$$

Setting $\xi=\theta-\frac{\sigma^2}{4\kappa}$, the Ninomiya-Victoir scheme is given by

$$
\begin{cases}
V^{NV,\eta}_{t_{k+\frac{1}{3}}} = \left(V^{NV,\eta}_{t_k}-\xi\right)\exp\left(-\frac{1}{2}\kappa h\right)+\xi\\
U^{NV,\eta}_{t_{k+\frac{1}{3}}} = U^{NV,\eta}_{t_k}+\frac{1}{2}\left(r-\frac{1}{2}\xi\right)h+\frac{1}{2\kappa}\left(V^{NV,\eta}_{t_k}-\xi\right)\left(\exp\left(-\frac{1}{2}\kappa h\right)-1\right)\\
V^{NV,\eta}_{t_{k+\frac{2}{3}}} = \left(\sqrt{V^{\eta,\eta}_{t_{k+\frac{1}{3}}}}+\frac{1}{2}\sigma\Delta W^2_{t_{k+1}}\right)^2\\
U^{NV,\eta}_{t_{k+\frac{2}{3}}} = U^{NV,\eta}_{t_{k+\frac{1}{3}}}+\sqrt{V^{NV,\eta}_{t_{k+\frac{1}{3}}}\mathbb{1}_{\{\eta_{k+1}=1\}}+V^{NV,\eta}_{t_{k+\frac{2}{3}}}\mathbb{1}_{\{\eta_{k+1}=-1\}}}\Delta W^1_{t_{k+1}}\\
V^{NV,\eta}_{t_{k+1}} = \left(V^{NV,\eta}_{t_{k+\frac{2}{3}}}-\xi\right)\exp\left(-\frac{1}{2}\kappa h\right)+\xi\\
U^{NV,\eta}_{t_{k+1}} = U^{NV,\eta}_{t_{k+\frac{2}{3}}}+\frac{1}{2}\left(r-\frac{1}{2}\xi\right)h+\frac{1}{2\kappa}\left(V^{NV,\eta}_{t_{k+\frac{2}{3}}}-\xi\right)\left(\exp\left(-\frac{1}{2}\kappa h\right)-1\right).
\end{cases}
\tag{2.4.32}
$$

In these formulas close to our implementation of the scheme, the evolution from

$$\left(U^{NV,\eta}_{t_k},V^{NV,\eta}_{t_k}\right),\text{ respectively }\left(U^{NV,\eta}_{t_{k+\frac{2}{3}}},V^{NV,\eta}_{t_{k+\frac{2}{3}}}\right),$$

to

$$\left(U^{NV,\eta}_{t_{k+\frac{1}{3}}},V^{NV,\eta}_{t_{k+\frac{1}{3}}}\right),\text{ respectively }\left(U^{NV,\eta}_{t_{k+1}},V^{NV,\eta}_{t_{k+1}}\right),$$

corresponds to the integration of the ODE directed by the vector field σ^0 on half a time step whereas the evolution from

$$\left(U^{NV,\eta}_{t_{k+\frac{1}{3}}},V^{NV,\eta}_{t_{k+\frac{1}{3}}}\right)$$

to

$$\left(U^{NV,\eta}_{t_{k+\frac{2}{3}}},V^{NV,\eta}_{t_{k+\frac{2}{3}}}\right)$$

corresponds to the integration of the Brownian vector fields. The Giles-Szpruch scheme and usual schemes such as the Euler scheme are not well defined since they can lead to negative values of the volatility process for which the square root is not defined at the next step. Assuming $\xi\geq 0$, the Ninomiya-Victoir scheme is well defined and the volatility process is always positive (see

[5]). For $\xi < 0$, in section 3.1 of [5], Alfonsi proposed a modification of the Ninomiya-Victoir scheme preserving the positivity of the volatility and the weak order two. For the simulation studies, we choose, as in [23], $S_0 = V_0 = 1, r = 0.05, T = 1,\ \kappa = 0.5, \theta = 0.9$ and $\sigma = 0.05$. Then $\xi = 0.89875$, so that the Ninomiya-Victoir scheme is well defined. Using this parameters, we do not observe negative values for the volatility with the Giles-Szpruch scheme. We choose to price the at the money call option. This corresponds to the payoff $f(u,v) = \exp(-rT)(\exp(u) - 1\ )_+$. Estimating the multilevel parameters, we obtain, $\alpha = 1$ and $\beta = 2$ and not 3/2 as predicted by the analysis. For the Nynomiya-Victoir scheme the estimation of (α, c_1) leads to $(2, -3.2\times 10^{-4})$. Since c_1 is very small, formula (2.4.5) can lead to negative values. If this occurs we set $L^*(\epsilon) = 1$. We also observe that $\mathbb{V}\left(Z^l_{NV}\right)$ decreases very quickly and faster than the theoretical rate for the first levels. Actually, an analogy can be drawn between the Ninomiya-Victoir scheme for the Clark-Cameron SDE and the Ninomiya-Victoir scheme for the Heston SDE since we have the same structure. As a matter of fact, when

$$\sqrt{V^{\eta,\eta}_{t_{k+\frac{1}{3}}}} + \frac{1}{2}\sigma\Delta W^2_{t_{k+1}} \geq 0,$$

$$V^{NV,\eta}_{t_{k+\frac{2}{3}}} = \left(\sqrt{V^{\eta,\eta}_{t_{k+\frac{1}{3}}}} + \frac{1}{2}\sigma\Delta W^2_{t_{k+1}}\right)^2$$

rewrites

$$\sqrt{V^{NV,\eta}_{t_{k+\frac{2}{3}}}} = \sqrt{V^{\eta,\eta}_{t_{k+\frac{1}{3}}}} + \frac{1}{2}\sigma\Delta W^2_{t_{k+1}}.$$

This equation, similar to

$$S^{NV,\eta}_{t_{k+1}} = S^{NV,\eta}_{t_k} + \mu\left(t_{k+1} - t_k\right) + \left(W^2_{t_{k+1}} - W^2_{t_k}\right)$$

in the Clark-Cameron SDE, is the only place where the Brownian increment ΔW^2 appears in the Ninomiya-Victoir scheme for the Heston model. In the dynamics of the U component the Brownian increment ΔW^1 is multiplied by $S^{NV,\eta}$ in the Clark-Cameron SDE and

$$\sqrt{V^{NV,\eta}_{t_{k+\frac{1}{3}}}\mathbb{1}_{\{\eta_{k+1}=1\}} + V^{NV,\eta}_{t_{k+\frac{2}{3}}}\mathbb{1}_{\{\eta_{k+1}=-1\}}}$$

in the Heston SDE. Then, the presence of a non-zero drift probably explains the existence of the higher order terms that disrupt the theoretical behavior like in formula (2.4.23). Figure 2.9 illustrate this phenomenon. The blue line is the estimation of $\left(\mathbb{V}\left(Z^l_{NV}\right)\right)_{1\leq l\leq 7}$ whereas the red line is the regression on the first four values. To be precise, we estimate $\mathbb{V}\left(Z^l_{NV}\right)$, for $l \in \{1,\ldots,7\}$ using $M = 10^7$ samples to get pretty good estimations, but in practice, $M = 10^6$ would be enough to implement the multilevel estimators. The regression leads to $\beta = 3$. As in the Clark-Cameron SDE with $f(u,s) = s_+$, the two lines diverge at level $\bar{l} = 5$.

So to implement the multilevel estimators, we compare $\bar{l}$ and $L^*(\epsilon)$ as already mentioned in Remark 2.4.5. The values of $L^*(\epsilon)$ are given by

ϵ	2^{-7}	2^{-8}	2^{-9}	2^{-10}	2^{-11}	2^{-12}	2^{-13}	2^{-14}	2^{-15}	2^{-16}
$\hat{Y}^{NV}_{MLMC}$	1	1	1	1	1	1	2	2	3	3
$\hat{Y}^{NV}_{ML2R}$	2	2	2	2	3	3	3	3	3	3

.

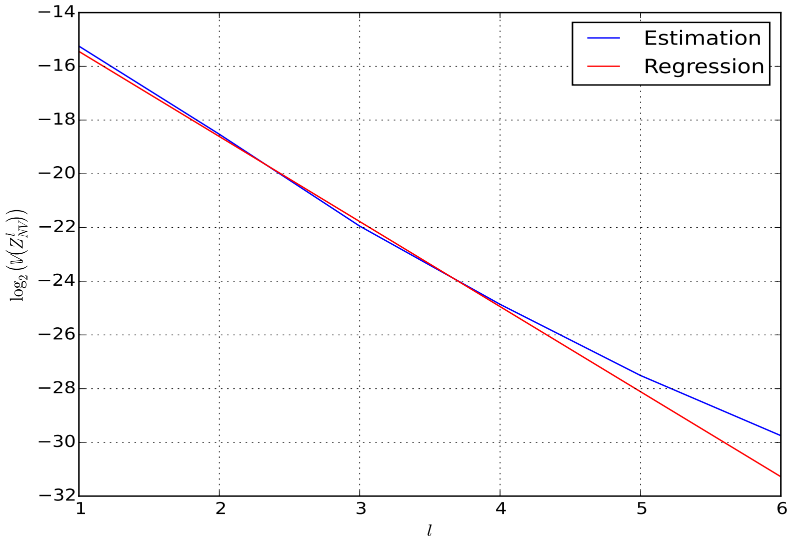

Fig. 2.9 Heston SDE with $f(u,v) = \exp(-rT)(\exp(u) - 1)_+$, Variance of the Ninomiya-Victoir scheme (y-axis $\log_2$ scale) as a function of l (x-axis).

We notice that even if ϵ is very small, $L(\epsilon) < \bar{l}$. So we can implement the multilevel estimators using the standard automatic procedure. With regard to the Ninomiya-Victoir scheme, we remark that:

$$\mathbb{V}\left(\frac{1}{2}\left(f\left(X_T^{NV,1,\eta}\right) + f\left(X_T^{NV,1,-\eta}\right)\right)\right) \approx \mathbb{V}\left(f\left(X_T^{NV,1,\eta}\right)\right),$$

so we naturally decide to implement the multilevel Monte Carlo with

$$Z^0_{NV} = f\left(X_T^{NV,1,\eta}\right).$$

In the following plots we compare the five estimators. This time, the fastest estimator is $\hat{Y}^{NV}_{ML2R}$. This is due to the outstanding value of β observed for the Ninomiya-Victoir scheme. The poor performance of $\hat{Y}^{NV}_{MLMC}$ is explained by the high variance at the level 0, while the variance at the higher levels are very small since the numerical value of β is 3.

Figure 2.11, which represents our CPU-time ratios defined as previously, emphasizes that $\hat{Y}^{NV}_{ML2R}$ is about 1.75 faster than $\hat{Y}^{NV}_{MLMC}$. $\hat{Y}^{NV}_{ML2R}$ is also 1.3 faster than $\hat{Y}^{NV-GS}_{MLMC}$ since the black dashed curve is always below 1.

2.5 Conclusion

In this paper, we have improved the multilevel Monte Carlo estimator of Giles and Szpruch [23] using a coupling between the Giles-Szpruch and Ninomiya-Victoir schemes at the last level of the MLMC estimator, which generalize their antithetic method. When the payoff is Lipschitz and piecewise smooth, which is very common in finance for example, the gain is amplified since $\beta = 3/2$. We have also highlighted a strange phenomenon: sometimes the numerical rate of convergence of the variance can be better than the theoretical one, at least for the levels used in the multilevel methods. This illustrates the presence of higher order terms which overshadows the theoretical behavior. Therefore, we emphasize that the estimation of the rate β and its associated constant c_2 should be done cautiously, since the optimal parameters of the multilevel estimators depend on them.

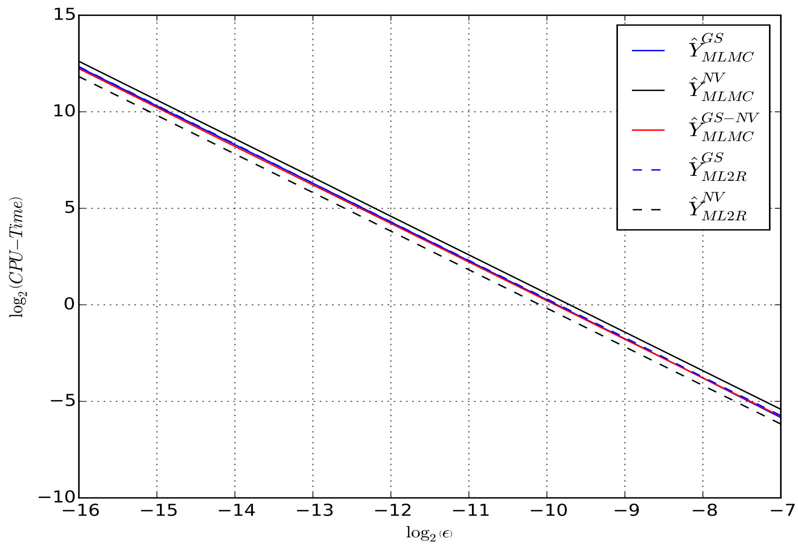

Fig. 2.10 Heston SDE with $f(u, v) = \exp(-rT)(\exp(u) - 1)_+$, CPU-time in second ($y$-axis $\log_2$ scale) as a function of ϵ (x-axis $\log_2$ scale).

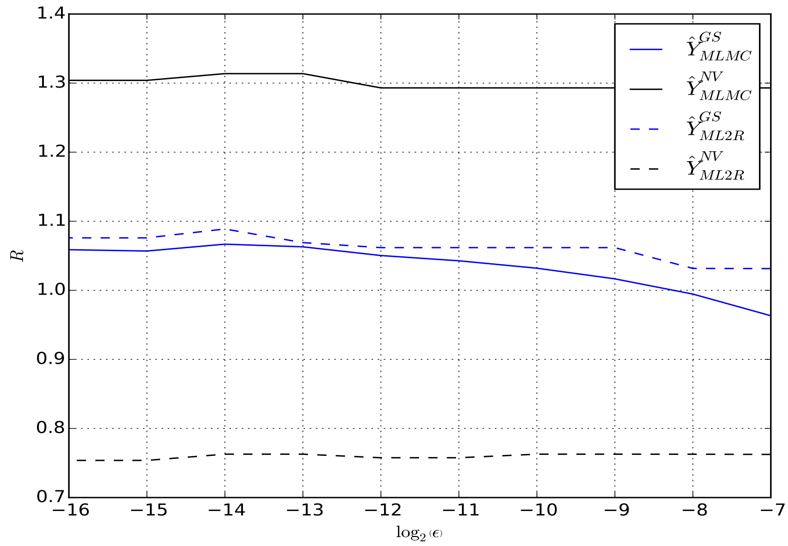

Fig. 2.11 Heston SDE with $f(u, v) = \exp(-rT)(\exp(u) - 1)_+$, CPU-time ratios ($y$-axis) as a function of ϵ (x-axis $\log_2$ scale).

2.6 Appendix

Let $l \in \mathbb{N}^*$ and $N = 2^{l-1}$. We recall that the Clark-Cameron SDE is defined as follows

$$\begin{cases} dU_t = S_t dW_t^1 \\ dS_t = \mu dt + dW_t^2, \end{cases}$$

and the Ninomiya-Victoir is given by
if $\eta_{t_{k+1}} = -1$

$$\begin{cases} U_{t_{k+1}}^{NV,\eta} = U_{t_k}^{NV,\eta} + S_{t_k}^{NV,\eta}\left(W_{t_{k+1}}^1 - W_{t_k}^1\right) + \frac{1}{2}\mu\left(t_{k+1}-t_k\right)\left(W_{t_{k+1}}^1 - W_{t_k}^1\right) \\ \qquad \left(W_{t_{k+1}}^1 - W_{t_k}^1\right)\left(W_{t_{k+1}}^2 - W_{t_k}^2\right) \\ S_{t_{k+1}}^{NV,\eta} = S_{t_k}^{NV,\eta} + \mu\left(t_{k+1}-t_k\right) + \left(W_{t_{k+1}}^2 - W_{t_k}^2\right), \end{cases}$$

and if $\eta_{t_{k+1}} = 1$

$$\begin{cases} U_{t_{k+1}}^{NV,\eta} = U_{t_k}^{NV,\eta} + S_{t_k}^{NV,\eta}\left(W_{t_{k+1}}^1 - W_{t_k}^1\right) + \frac{1}{2}\mu\left(t_{k+1}-t_k\right)\left(W_{t_{k+1}}^1 - W_{t_k}^1\right) \\ S_{t_{k+1}}^{NV,\eta} = S_{t_k}^{NV,\eta} + \mu\left(t_{k+1}-t_k\right) + \left(W_{t_{k+1}}^2 - W_{t_k}^2\right). \end{cases}$$

We define $\bar{U}^{NV,\eta}$ and $\bar{S}^{NV,\eta}$ as $\frac{1}{2}\left(U^{NV,\eta} + U^{NV,-\eta}\right)$, $\frac{1}{2}\left(S^{NV,\eta} + S^{NV,-\eta}\right)$ respectively.

Remark 2.6.1 *$S^{NV,\eta}$ does not depend on η, so $\bar{S}^{NV,\eta} = S^{NV,\eta}$.*

The evolution of $\left(\bar{U}^{NV,2^{l-1},\eta}, \bar{V}^{NV,2^{l-1},\eta}\right)$ on the coarse grid with time step $T/2^{l-1}$ is given by

$$\begin{cases} \bar{U}_{t_{k+1}}^{NV,2^{l-1},\eta} = \bar{U}_{t_k}^{NV,2^{l-1},\eta} + \bar{S}_{t_k}^{NV,2^{l-1},\eta}\left(W_{t_{k+1}}^1 - W_{t_k}^1\right) + \frac{1}{2}\mu\left(t_{k+1}-t_k\right)\left(W_{t_{k+1}}^1 - W_{t_k}^1\right) \\ \qquad + \frac{1}{2}\left(W_{t_{k+1}}^1 - W_{t_k}^1\right)\left(W_{t_{k+1}}^2 - W_{t_k}^2\right) \\ \bar{S}_{t_{k+1}}^{NV,2^{l-1},\eta} = \bar{S}_{t_k}^{NV,2^{l-1},\eta} + \mu\left(t_{k+1}-t_k\right) + \left(W_{t_{k+1}}^2 - W_{t_k}^2\right). \end{cases}$$

Similarly, the evolution of $\left(\bar{U}^{NV,2^l,\eta}, \bar{V}^{NV,2^l,\eta}\right)$ on the fine grid with time step $T/2^l$ is given by

$$\begin{cases} \bar{U}_{t_{k+\frac{1}{2}}}^{NV,2^l,\eta} = \bar{U}_{t_k}^{NV,2^l,\eta} + \bar{S}_{t_k}^{NV,2^l,\eta}\left(W_{t_{k+\frac{1}{2}}}^1 - W_{t_k}^1\right) + \frac{1}{2}\mu\left(t_{k+\frac{1}{2}}-t_k\right)\left(W_{t_{k+\frac{1}{2}}}^1 - W_{t_k}^1\right) \\ \qquad + \frac{1}{2}\left(W_{t_{k+\frac{1}{2}}}^1 - W_{t_k}^1\right)\left(W_{t_{k+\frac{1}{2}}}^2 - W_{t_k}^2\right) \\ \bar{S}_{t_{k+\frac{1}{2}}}^{NV,2^l,\eta} = \bar{S}_{t_k}^{NV,2^l,\eta} + \mu\left(t_{k+\frac{1}{2}}-t_k\right) + \left(W_{t_{k+\frac{1}{2}}}^2 - W_{t_k}^2\right), \end{cases}$$

and

$$\begin{cases} \bar{U}_{t_{k+1}}^{NV,2^l,\eta} = \bar{U}_{t_{k+\frac{1}{2}}}^{NV,2^l,\eta} + \bar{S}_{t_{k+\frac{1}{2}}}^{NV,2^l,\eta}\left(W_{t_{k+1}}^1 - W_{t_{k+\frac{1}{2}}}^1\right) + \frac{1}{2}\mu\left(t_{k+1}-t_{k+\frac{1}{2}}\right)\left(W_{t_{k+1}}^1 - W_{t_{k+\frac{1}{2}}}^1\right) \\ \qquad + \frac{1}{2}\left(W_{t_{k+1}}^1 - W_{t_{k+\frac{1}{2}}}^1\right)\left(W_{t_{k+1}}^2 - W_{t_{k+\frac{1}{2}}}^2\right) \\ \bar{S}_{t_{k+1}}^{NV,2^l,\eta} = \bar{S}_{t_{k+\frac{1}{2}}}^{NV,2^l,\eta} + \mu\left(t_{k+1}-t_{k+\frac{1}{2}}\right) + \left(W_{t_{k+1}}^2 - W_{t_{k+\frac{1}{2}}}^2\right). \end{cases}$$

By a straightforward calculation, we get

$$\begin{cases} \bar{U}^{NV,2^l,\eta}_{t_{k+1}} = \bar{U}^{NV,2^l,\eta}_{t_k} + \bar{S}^{NV,2^l,\eta}_{t_k}\left(W^1_{t_{k+1}} - W^1_{t_k}\right) \\ \qquad + \frac{1}{2}\mu\left(t_{k+1}-t_k\right)\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right) + \frac{1}{4}\mu\left(t_{k+1}-t_k\right)\left(W^1_{t_{k+1}} - W^1_{t_k}\right) \\ \qquad + \left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) \\ \qquad + \frac{1}{2}\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) + \frac{1}{2}\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right) \\ \bar{S}^{NV,2^l,\eta}_{t_{k+1}} = \bar{S}^{NV,2^l,\eta}_{t_k} + \mu\left(t_{k+1}-t_k\right) + \left(W^2_{t_{k+1}} - W^2_{t_k}\right). \end{cases}$$

The antithetic scheme $\left(\tilde{\bar{U}}^{NV,2^l,\eta}, \tilde{\bar{V}}^{NV,2^l,\eta}\right)$ is defined by swapping $\left(W_{t_{k+1}} - W_{t_{k+\frac{1}{2}}}\right)$ and $\left(W_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)$:

$$\begin{cases} \tilde{\bar{U}}^{NV,2^l,\eta}_{t_{k+1}} = \tilde{\bar{U}}^{NV,2^l,\eta}_{t_k} + \tilde{\bar{S}}^{NV,\eta,a}_{t_k}\left(W^1_{t_{k+1}} - W^1_{t_k}\right) \\ \qquad + \frac{1}{2}\mu\left(t_{k+1}-t_k\right)\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right) + \frac{1}{4}\mu\left(t_{k+1}-t_k\right)\left(W^1_{t_{k+1}} - W^1_{t_k}\right) \\ \qquad + \left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right) \\ \qquad + \frac{1}{2}\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) + \frac{1}{2}\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right) \\ \tilde{\bar{S}}^{NV,2^l,\eta}_{t_{k+1}} = \tilde{\bar{S}}^{NV,2^l,\eta}_{t_k} + \mu\left(t_{k+1}-t_k\right) + \left(W^2_{t_{k+1}} - W^2_{t_k}\right). \end{cases}$$

Now we define

$$\bar{\bar{U}}^{NV,2^l,\eta}_{t_k} := \frac{1}{2}\left(\bar{U}^{NV,2^l,\eta}_{t_{k+1}} + \tilde{\bar{U}}^{NV,2^l,\eta}_{t_{k+1}}\right),$$

and

$$\bar{\bar{S}}^{NV,2^l,\eta}_{t_k} := \frac{1}{2}\left(\bar{S}^{NV,2^l,\eta}_{t_{k+1}} + \tilde{\bar{S}}^{NV,2^l,\eta}_{t_{k+1}}\right).$$

Performing a straightforward calculation, we obtain

$$\begin{cases} \bar{\bar{U}}^{NV,2^l,\eta}_{t_{k+1}} = \bar{\bar{U}}^{NV,2^l,\eta}_{t_k} + \bar{\bar{S}}^{NV,2^l,\eta}_{t_k}\left(W^1_{t_{k+1}} - W^1_{t_k}\right) + \frac{1}{2}\mu\left(t_{k+1}-t_k\right)\left(W^1_{t_{k+1}} - W^1_{t_k}\right) \\ \qquad + \frac{1}{2}\left(W^1_{t_{k+1}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_k}\right) \\ \bar{\bar{S}}^{NV,2^l,\eta}_{t_{k+1}} = \bar{\bar{S}}^{NV,2^l,\eta}_{t_k} + \mu\left(t_{k+1}-t_k\right) + \left(W^2_{t_{k+1}} - W^2_{t_k}\right). \end{cases}$$

Then, using forward induction on k, we easily get that for all $k \in \left\{0, \ldots, 2^{l-1}-1\right\}$

$$\begin{cases} \bar{\bar{U}}^{NV,2^l,\eta}_{t_{k+1}} - \bar{U}^{NV,2^{l-1},\eta}_{t_{k+1}} = 0 \\ \bar{\bar{S}}^{NV,2^l,\eta}_{t_{k+1}} - \bar{S}^{NV,2^{l-1},\eta}_{t_{k+1}} = 0. \end{cases} \tag{2.6.1}$$

We want to calculate

$$Y = \mathbb{E}\left[\left(Z^l_{NV}\right)^2\right],$$

where

$$Z_{NV}^{l} = \frac{1}{4}\left(\left(U_T^{NV,2^l,\eta}\right)^2 + \left(\tilde{U}_T^{NV,2^l,\eta}\right)^2 + \left(U_T^{NV,2^l,-\eta}\right)^2 + \left(\tilde{U}_T^{NV,2^l,-\eta}\right)^2\right) - \frac{1}{2}\left(\left(U_T^{NV,2^{l-1},\eta}\right)^2 + \left(U_T^{NV,2^{l-1},-\eta}\right)^2\right).$$

Using

$$\begin{aligned}\frac{1}{4}\left(x^2+y^2+u^2+v^2\right) - \frac{1}{2}\left(z^2+w^2\right) = &\left(\frac{1}{4}\left(x+y+u+v\right)\right)^2 - \left(\frac{1}{2}\left(z+w\right)\right)^2 \\ &+ \frac{1}{64}\left(3x-y-u-v\right)^2 + \frac{1}{64}\left(3y-x-u-v\right)^2 \\ &+ \frac{1}{64}\left(3u-x-y-v\right)^2 + \frac{1}{64}\left(3v-x-y-u\right)^2 \\ &- \frac{1}{4}\left(z-w\right)^2,\end{aligned}$$

and (2.6.1), we get

$$Y = \frac{1}{16}\mathbb{E}\left[Z^2\right],$$

where

$$\begin{aligned}Z = &\frac{1}{16}\left(3U_T^{NV,2^l,\eta} - \tilde{U}_T^{NV,2^l,\eta} - U_T^{NV,2^l,-\eta} - \tilde{U}_T^{NV,2^l,-\eta}\right)^2 \\ &+ \frac{1}{16}\left(3\tilde{U}_T^{NV,2^l,\eta} - U_T^{NV,2^l,\eta} - U_T^{NV,2^l,-\eta} - \tilde{U}_T^{NV,2^l,-\eta}\right)^2 \\ &+ \frac{1}{16}\left(3U_T^{NV,2^l,-\eta} - U_T^{NV,2^l,\eta} - \tilde{U}_T^{NV,2^l,\eta} - \tilde{U}_T^{NV,2^l,-\eta}\right)^2 \\ &+ \frac{1}{16}\left(3\tilde{U}_T^{NV,2^l,-\eta} - U_T^{NV,2^l,\eta} - U_T^{NV,2^l,-\eta} - \tilde{U}_T^{NV,2^l,\eta}\right)^2 \\ &- \left(U_T^{NV,2^{l-1},\eta} - U_T^{NV,2^{l-1},-\eta}\right)^2.\end{aligned}$$

To lighten the previous expression, we introduce

$$Z_1 = \frac{1}{4}\left(3U_T^{NV,2^l,\eta} - \tilde{U}_T^{NV,2^l,\eta} - U_T^{NV,2^l,-\eta} - \tilde{U}_T^{NV,2^l,-\eta}\right),$$

$$Z_2 = \frac{1}{4}\left(3\tilde{U}_T^{NV,2^l,\eta} - U_T^{NV,2^l,\eta} - U_T^{NV,2^l,-\eta} - \tilde{U}_T^{NV,2^l,-\eta}\right),$$

$$Z_3 = \frac{1}{4}\left(3U_T^{NV,2^l,-\eta} - U_T^{NV,2^l,\eta} - \tilde{U}_T^{NV,2^l,\eta} - \tilde{U}_T^{NV,2^l,-\eta}\right),$$

$$Z_4 = \frac{1}{4}\left(3\tilde{U}_T^{NV,2^l,-\eta} - U_T^{NV,2^l,\eta} - U_T^{NV,2^l,-\eta} - \tilde{U}_T^{NV,2^l,\eta}\right),$$

and

$$Z_0 = U_T^{NV,2^{l-1},\eta} - U_T^{NV,2^{l-1},-\eta}.$$

In order to get an explicit expression for Z_i, for all $i \in \{0, \ldots, 4\}$, we compute the following differences

$$\begin{aligned}
U_{t_{k+1}}^{NV,2^l,\eta} - U_{t_{k+1}}^{NV,2^l,-\eta} &= U_{t_k}^{NV,2^l,\eta} - U_{t_k}^{NV,2^l,-\eta} \\
&- \eta_{k+\frac{1}{2}} \left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) \\
&- \eta_{k+1} \left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right),
\end{aligned}$$

$$\begin{aligned}
\tilde{U}_{t_{k+1}}^{NV,2^l,\eta} - \tilde{U}_{t_{k+1}}^{NV,2^l,-\eta} &= \tilde{U}_{t_k}^{NV,2^l,\eta} - \tilde{U}_{t_k}^{NV,2^l,-\eta} \\
&- \eta_{k+1} \left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) \\
&- \eta_{k+\frac{1}{2}} \left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right),
\end{aligned}$$

$$\begin{aligned}
U_{t_{k+1}}^{NV,2^l,\eta} - \tilde{U}_{t_{k+1}}^{NV,2^l,\eta} &= U_{t_k}^{NV,2^l,\eta} - \tilde{U}_{t_k}^{NV,2^l,\eta} + \frac{\mu}{2}\frac{T}{N}\left(W^1_{t_{k+1}} - 2W^1_{t_{k+\frac{1}{2}}} + W^1_{t_k}\right) \\
&+ \left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) \\
&- \left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right) \\
&+ \frac{1}{2}\left(\eta_{k+1} - \eta_{k+\frac{1}{2}}\right)\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) \\
&+ \frac{1}{2}\left(\eta_{k+\frac{1}{2}} - \eta_{k+1}\right)\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right),
\end{aligned}$$

$$\begin{aligned}
U_{t_{k+1}}^{NV,2^l,-\eta} - \tilde{U}_{t_{k+1}}^{NV,2^l,-\eta} &= U_{t_k}^{NV,2^l,-\eta} - \tilde{U}_{t_k}^{NV,2^l,-\eta} + \frac{\mu}{2}\frac{T}{N}\left(W^1_{t_{k+1}} - 2W^1_{t_{k+\frac{1}{2}}} + W^1_{t_k}\right) \\
&+ \left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) \\
&- \left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right) \\
&+ \frac{1}{2}\left(\eta_{k+\frac{1}{2}} - \eta_{k+1}\right)\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) \\
&+ \frac{1}{2}\left(\eta_{k+1} - \eta_{k+\frac{1}{2}}\right)\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right),
\end{aligned}$$

$$\begin{aligned}
U_{t_{k+1}}^{NV,2^l,\eta} - \tilde{U}_{t_{k+1}}^{NV,2^l,-\eta} &= U_{t_k}^{NV,2^l,\eta} - \tilde{U}_{t_k}^{NV,2^l,-\eta} + \frac{\mu}{2}\frac{T}{N}\left(W^1_{t_{k+1}} - 2W^1_{t_{k+\frac{1}{2}}} + W^1_{t_k}\right) \\
&+ \left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) \\
&- \left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right),
\end{aligned}$$

and

$$U^{NV,2^l,-\eta}_{t_{k+1}} - \tilde{U}^{NV,2^l,\eta}_{t_{k+1}} = U^{NV,2^l,-\eta}_{t_k} - \tilde{U}^{NV,2^l,\eta}_{t_k} + \frac{\mu}{2}\frac{T}{N}\left(W^1_{t_{k+1}} - 2W^1_{t_{k+\frac{1}{2}}} + W^1_{t_k}\right)$$
$$+ \left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right)$$
$$- \left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right).$$

Hence, we obtain by summation

$$Z_i = \sum_{k=0}^{N-1} z^i_k,$$

where

$$z^0_k = \left(W^1_{t_{k+1}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_k}\right),$$

$$\begin{aligned}
z^1_k = &\left(\frac{1}{2} - \frac{1}{8}\left(\eta_{k+\frac{1}{2}} - \eta_{k+1}\right)\right)\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right)\\
&- \left(\frac{1}{2} + \frac{1}{8}\left(\eta_{k+\frac{1}{2}} - \eta_{k+1}\right)\right)\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right)\\
&+ \mu\frac{T}{4N}\left(W^1_{t_{k+1}} - 2W^1_{t_{k+\frac{1}{2}}} + W^1_{t_k}\right)\\
&- \frac{1}{4}\eta_{k+\frac{1}{2}}\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right)\\
&- \frac{1}{4}\eta_{k+1}\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right),
\end{aligned}$$

$$\begin{aligned}
z^2_k = &-\left(\frac{1}{2} + \frac{1}{8}\left(\eta_{k+\frac{1}{2}} - \eta_{k+1}\right)\right)\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right)\\
&+ \left(\frac{1}{2} - \frac{1}{8}\left(\eta_{k+\frac{1}{2}} - \eta_{k+1}\right)\right)\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right)\\
&- \mu\frac{T}{4N}\left(W^1_{t_{k+1}} - 2W^1_{t_{k+\frac{1}{2}}} + W^1_{t_k}\right)\\
&- \frac{1}{4}\eta_{k+1}\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right)\\
&- \frac{1}{4}\eta_{k+\frac{1}{2}}\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right),
\end{aligned}$$

$$\begin{aligned}
z^3_k = &\left(\frac{1}{2} + \frac{1}{8}\left(\eta_{k+\frac{1}{2}} - \eta_{k+1}\right)\right)\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right)\\
&- \left(\frac{1}{2} - \frac{1}{8}\left(\eta_{k+\frac{1}{2}} - \eta_{k+1}\right)\right)\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right)\\
&+ \mu\frac{T}{4N}\left(W^1_{t_{k+1}} - 2W^1_{t_{k+\frac{1}{2}}} + W^1_{t_k}\right)\\
&+ \frac{1}{4}\eta_{k+\frac{1}{2}}\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right)\\
&+ \frac{1}{4}\eta_{k+1}\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right),
\end{aligned}$$

and

$$
\begin{aligned}
z_k^4 = &-\left(\frac{1}{2} - \frac{1}{8}\left(\eta_{k+\frac{1}{2}} - \eta_{k+1}\right)\right)\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) \\
&+ \left(\frac{1}{2} + \frac{1}{8}\left(\eta_{k+\frac{1}{2}} - \eta_{k+1}\right)\right)\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right) \\
&- \mu\frac{T}{4N}\left(W^1_{t_{k+1}} - 2W^1_{t_{k+\frac{1}{2}}} + W^1_{t_k}\right) \\
&+ \frac{1}{4}\eta_{k+1}\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) \\
&+ \frac{1}{4}\eta_{k+\frac{1}{2}}\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right).
\end{aligned}
$$

Then, one can express Z^2 as

$$
\begin{aligned}
Z^2 &= \left(\left(\sum_{k=0}^{N-1} z_k^1\right)^2 + \left(\sum_{k=0}^{N-1} z_k^2\right)^2 + \left(\sum_{k=0}^{N-1} z_k^3\right)^2 + \left(\sum_{k=0}^{N-1} z_k^4\right)^2 - \left(\sum_{k=0}^{N-1} z_k^0\right)^2\right)^2 \\
&= \left(\sum_{k=0}^{N-1}\left(z_k^1\right)^2 + \left(z_k^2\right)^2 + \left(z_k^3\right)^2 + \left(z_k^4\right)^2 - \left(z_k^0\right)^2\right. \\
&\qquad\qquad \left. + 2\sum_{(k,l)\in\Delta_N} z_1^k z_1^l + z_2^k z_2^l + z_3^k z_3^l + z_4^k z_4^l - z_0^k z_0^l\right)^2,
\end{aligned}
$$

where $\Delta_N = \left\{(k,l) \in \{0,\dots,N-1\}^2, k < l\right\}$.

Preliminary calculus:

We begin by writing z_k^i in generic form

$$
\begin{aligned}
z_k^i = \alpha_k^i &\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) + \beta_k^i\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right) \\
&+ \gamma_k^i\left(W^1_{t_{k+1}} - 2W^1_{t_{k+\frac{1}{2}}} + W^1_{t_k}\right) + \delta_k^i\left(W^1_{t_{k+\frac{1}{2}}} - W^1_{t_k}\right)\left(W^2_{t_{k+\frac{1}{2}}} - W^2_{t_k}\right) \\
&+ \omega_k^i\left(W^1_{t_{k+1}} - W^1_{t_{k+\frac{1}{2}}}\right)\left(W^2_{t_{k+1}} - W^2_{t_{k+\frac{1}{2}}}\right);
\end{aligned}
$$

where

$\alpha_k^1 = \left(\frac{1}{2} - \frac{1}{8}\left(\eta_{k+\frac{1}{2}} - \eta_{k+1}\right)\right)$	$\alpha_k^2 = \beta_k^1$	$\alpha_k^3 = -\beta_k^1$	$\alpha_k^4 = -\alpha_k^1$	$\alpha_k^0 = 1$
$\beta_k^1 = -\left(\frac{1}{2} + \frac{1}{8}\left(\eta_{k+\frac{1}{2}} - \eta_{k+1}\right)\right)$	$\beta_k^2 = \alpha_k^1$	$\beta_k^3 = -\alpha_k^1$	$\beta_k^4 = -\beta_k^1$	$\beta_k^0 = 1$
$\gamma_k^1 = \mu\frac{T}{4N}$	$\gamma_k^2 = -\gamma_k^1$	$\gamma_k^3 = \gamma_k^1$	$\gamma_k^4 = -\gamma_k^1$	$\gamma_k^0 = 0$
$\delta_k^1 = -\frac{1}{4}\eta_{k+\frac{1}{2}}$	$\delta_k^2 = \omega_k^1$	$\delta_k^3 = -\delta_k^1$	$\delta_k^4 = -\omega_k^1$	$\delta_k^0 = 1$
$\omega_k^1 = -\frac{1}{4}\eta_{k+1}$	$\omega_k^2 = \delta_k^1$	$\omega_k^3 = -\omega_k^1$	$\omega_k^4 = -\delta_k^1$	$\omega_k^0 = 1$

.

First, let us look at the expectation of $z_k^i z_k^j$

$$
\mathbb{E}\left[z_k^i z_k^j\right] = \frac{T^2}{4N^2}\mathbb{E}\left[\alpha_k^i\alpha_k^j + \beta_k^i\beta_k^j + \delta_k^i\delta_k^j + \omega_k^i\omega_k^j\right] + \frac{T}{N}\mathbb{E}\left[\gamma_k^i\gamma_k^j\right].
$$

Then, for $i = j = 0$

$$\mathbb{E}\left[\left(z_k^0\right)^2\right] = \frac{T^2}{N^2}.$$

For $i = j \in \{1, \dots, 4\}$

$$\mathbb{E}\left[\left(z_k^i\right)^2\right] = \frac{11T^2}{64N^2} + \frac{\mu^2 T^3}{16N^3}.$$

For $j = 0$ and $i \in \{1, \dots, 4\}$

$$\mathbb{E}\left[z_k^i z_k^0\right] = 0.$$

For $(i, j) \in \{(1, 2), (3, 4)\}$

$$\mathbb{E}\left[z_k^i z_k^j\right] = -\frac{7T^2}{64N^2} - \frac{\mu^2 T^3}{16N^3}.$$

For $(i, j) \in \{(1, 3), (2, 4)\}$

$$\mathbb{E}\left[z_k^i z_k^j\right] = \frac{5T^2}{64N^2} + \frac{\mu^2 T^3}{16N^3}.$$

For $(i, j) \in \{(1, 4), (2, 3)\}$

$$\mathbb{E}\left[z_k^i z_k^j\right] = -\frac{9T^2}{64N^2} - \frac{\mu^2 T^3}{16N^3}.$$

Now we look at the expectation of $\left(z_k^i\right)^2 \left(z_k^j\right)^2$

$$\mathbb{E}\left[\left(z_k^i\right)^2 \left(z_k^j\right)^2\right] = \frac{T^4}{16N^4}\left(9E_1^4 + E_2^4 + 3E_3^4\right) + \frac{3T^2}{N^2}E^2,$$

where

$$E_1^4 = \mathbb{E}\left[\left(\alpha_k^i\right)^2 \left(\alpha_k^j\right)^2 + \left(\beta_k^i\right)^2 \left(\beta_k^j\right)^2 + \left(\delta_k^i\right)^2 \left(\delta_k^j\right)^2 + \left(\omega_k^i\right)^2 \left(\omega_k^j\right)^2\right],$$

$$\begin{aligned} E_2^4 = &\ \mathbb{E}\left[\left(\alpha_k^i\right)^2 \left(\beta_k^j\right)^2 + \left(\alpha_k^j\right)^2 \left(\beta_k^i\right)^2 + \left(\delta_k^i\right)^2 \left(\omega_k^j\right)^2 + \left(\delta_k^j\right)^2 \left(\omega_k^i\right)^2\right] \\ &+ 4\left(\mathbb{E}\left[\left(\alpha_k^i\beta_k^i + \delta_k^i\omega_k^i\right)\left(\alpha_k^j\beta_k^j + \delta_k^j\omega_k^j\right)\right]\mathbb{E}\left[\left(\alpha_k^i\beta_k^j + \alpha_k^j\beta_k^i\right)\left(\delta_k^i\omega_k^j + \delta_k^j\omega_k^i\right)\right]\right), \end{aligned}$$

$$\begin{aligned} E_3^4 = &\ \mathbb{E}\left[\left(\left(\alpha_k^i\right)^2 + \left(\beta_k^i\right)^2\right)\left(\left(\delta_k^j\right)^2 + \left(\omega_k^j\right)^2\right)\right] + \mathbb{E}\left[\left(\left(\alpha_k^j\right)^2 + \left(\beta_k^j\right)^2\right)\left(\left(\delta_k^i\right)^2 + \left(\omega_k^i\right)^2\right)\right] \\ &+ 4\mathbb{E}\left[\left(\alpha_k^i\alpha_k^j + \beta_k^i\beta_k^j\right)\left(\delta_k^i\delta_k^j + \omega_k^i\omega_k^j\right)\right], \end{aligned}$$

$$
\begin{aligned}
E_1^3 = &\, \mathbb{E}\left[\left(\gamma_k^i\right)^2\left(\left(\alpha_k^j\right)^2+\left(\beta_k^j\right)^2+\left(\delta_k^j\right)^2+\left(\omega_k^j\right)^2-4\alpha_k^j\delta_k^j-4\beta_k^j\omega_k^j\right)\right] \\
&+\mathbb{E}\left[\left(\gamma_k^j\right)^2\left(\left(\alpha_k^i\right)^2+\left(\beta_k^i\right)^2+\left(\delta_k^i\right)^2+\left(\omega_k^i\right)^2-4\alpha_k^i\delta_k^i-4\beta_k^i\omega_k^i\right)\right] \\
&+4\mathbb{E}\left[\gamma_k^i\gamma_k^j\left(\alpha_k^i\alpha_k^j+\beta_k^i\beta_k^j+\delta_k^i\delta_k^j+\omega_k^i\omega_k^j-2\alpha_k^i\delta_k^j-2\alpha_k^j\delta_k^i-2\beta_k^i\omega_k^j-2\beta_k^j\omega_k^i\right)\right],
\end{aligned}
$$

$$
\begin{aligned}
E_2^3 = &\, \mathbb{E}\left[\left(\gamma_k^i\right)^2\left(\left(\alpha_k^j\right)^2+\left(\beta_k^j\right)^2+\left(\delta_k^j\right)^2+\left(\omega_k^j\right)^2\right)\right] \\
&+\mathbb{E}\left[\left(\gamma_k^j\right)^2\left(\left(\alpha_k^i\right)^2+\left(\beta_k^i\right)^2+\left(\delta_k^i\right)^2+\left(\omega_k^i\right)^2\right)\right] \\
&+4\mathbb{E}\left[\gamma_k^i\gamma_k^j\left(\alpha_k^i\alpha_k^j+\beta_k^i\beta_k^j+\delta_k^i\delta_k^j+\omega_k^i\omega_k^j\right)\right],
\end{aligned}
$$

and

$$
\mathbb{E}\left[\left(\gamma_k^i\right)^2\left(\gamma_k^j\right)^2\right].
$$

Then, by straightforward calculation, for $i = j = 0$

$$
\mathbb{E}\left[\left(z_k^0\right)^4\right] = \frac{9T^4}{N^4}.
$$

For $i = j \in \{1, \dots, 4\}$

$$
\mathbb{E}\left[\left(z_k^i\right)^4\right] = \frac{231T^4}{1024N^4} + \frac{33\mu^2T^5}{256N^5} + \frac{3\mu^4T^6}{256N^6}.
$$

For $j = 0$ and $i \in \{1, \dots, 4\}$

$$
\mathbb{E}\left[\left(z_k^i\right)^2\left(z_k^0\right)^2\right] = \frac{37T^4}{64N^4} + \frac{\mu^2T^5}{16N^5}.
$$

For $(i, j) \in \{(1, 2), (3, 4)\}$

$$
\mathbb{E}\left[\left(z_k^i\right)^2\left(z_k^j\right)^2\right] = \frac{127T^4}{1024N^4} + \frac{25\mu^2T^5}{256N^5} + \frac{3\mu^4T^6}{256N^6}.
$$

For $(i, j) \in \{(1, 3), (2, 4)\}$

$$
\mathbb{E}\left[\left(z_k^i\right)^2\left(z_k^j\right)^2\right] = \frac{71T^4}{1024N^4} + \frac{21\mu^2T^5}{256N^5} + \frac{3\mu^4T^6}{256N^6}.
$$

For $(i, j) \in \{(1, 4), (2, 3)\}$

$$
\mathbb{E}\left[z_k^i z_k^j\right] = \frac{159T^4}{1024N^4} + \frac{29\mu^2T^5}{256N^5} + \frac{3\mu^4T^6}{256N^6}.
$$

Now, we define

$$
a_k = \left(z_k^1\right)^2 + \left(z_k^2\right)^2 + \left(z_k^3\right)^2 + \left(z_k^4\right)^2 - \left(z_k^0\right)^2,
$$

and

$$
b_{kl} = z_k^1 z_l^1 + z_k^2 z_l^2 + z_k^3 z_l^3 + z_k^4 z_l^4 - z_k^0 z_l^0.
$$

Then Z^2 can be expressed as

$$Z^2 = \sum_{k=0}^{N-1} a_k^2 + 2 \sum_{(k,l)\in\Delta_N} a_k a_l + 4 \left(\sum_{(k,l),(i,j)\in\Delta_N} b_{kl} b_{ij} + 4 \sum_{(k,l)\in\Delta_N, 0\leq j\leq N-1} a_j b_{kl} \right).$$

Observing that

$$\mathbb{E}\left[a_j b_{kl}\right] = 0,$$

and

$$\mathbb{E}\left[\sum_{(k,l),(i,j)\in\Delta_N} b_{kl} b_{ij}\right] = \mathbb{E}\left[\sum_{(k,l)\in\Delta_N} b_{kl}^2\right],$$

we obtain

$$\mathbb{E}\left[Z^2\right] = \sum_{k=0}^{N-1} \mathbb{E}\left[a_k^2\right] + 2 \sum_{(k,l)\in\Delta_N} \mathbb{E}\left[a_k a_l\right] + 4 \sum_{(k,l)\in\Delta_N} \mathbb{E}\left[b_{kl}^2\right].$$

For $k \neq l$, by independence

$$\mathbb{E}\left[a_k a_l\right] = \mathbb{E}\left[a_k\right] \mathbb{E}\left[a_l\right] = \left(\mathbb{E}\left[a_k\right]\right)^2,$$

this leads to

$$\mathbb{E}\left[Z^2\right] = N\mathbb{E}\left[a_k^2\right] + N\left(N-1\right)\left(\mathbb{E}\left[a_k\right]\right)^2 + 2N\left(N-1\right)\mathbb{E}\left[b_{kl}^2\right].$$

To achieve our goal, we calculate the following expectations

$$\mathbb{E}\left[a_k\right] = 4\mathbb{E}\left[\left(z_k^1\right)^2\right] - \mathbb{E}\left[\left(z_k^0\right)^2\right] = \frac{\mu^2 T^3}{4N^3} - \frac{5T^2}{16N^2},$$

$$\begin{aligned}
\mathbb{E}\left[a_k^2\right] &= 4\mathbb{E}\left[\left(z_k^1\right)^4\right] + \mathbb{E}\left[\left(z_k^0\right)^4\right] - 8\mathbb{E}\left[\left(z_k^1\right)^2\left(z_k^0\right)^2\right] \\
&\quad + 4\left(\mathbb{E}\left[\left(z_k^1\right)^2\left(z_k^2\right)^2\right] + \mathbb{E}\left[\left(z_k^1\right)^2\left(z_k^3\right)^2\right] + \mathbb{E}\left[\left(z_k^1\right)^2\left(z_k^4\right)^2\right]\right) \\
&= \frac{3\mu^4 T^6}{16N^6} + \frac{19\mu^2 T^5}{16N^5} + \frac{427T^4}{64N^4},
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\left[b_{kl}^2\right] &= 4\left(\mathbb{E}\left[\left(z_k^1\right)^2\right]\right)^2 + \left(\mathbb{E}\left[\left(z_k^0\right)^2\right]\right)^2 - 8\left(\mathbb{E}\left[z_k^1 z_k^0\right]\right)^2 \\
&\quad + 4\left(\left(\mathbb{E}\left[z_k^1 z_k^2\right]\right)^2 + \left(\mathbb{E}\left[z_k^1 z_k^3\right]\right)^2 + \left(\mathbb{E}\left[z_k^1 z_k^4\right]\right)^2\right) \\
&= \frac{\mu^4 T^6}{16N^6} + \frac{\mu^2 T^5}{4N^5} + \frac{69T^4}{256N^4}.
\end{aligned}$$

Combining our results, we obtain

$$\mathbb{E}\left[Z^2\right] = \frac{1}{N^4}\left(\frac{3}{16}\mu^4 T^6 + \frac{9}{16}\mu^2 T^5\right) + \frac{1}{N^3}\left(\frac{11}{32}\mu^2 T^5 + \frac{1545}{256}T^4\right) + \frac{1}{N^2}\left(\frac{163}{256}T^4\right).$$

Then replacing N by 2^{l-1}

$$\mathbb{E}\left[Z^2\right] = 2^{-4l}\left(3\mu^4T^6 + 9\mu^2T^5\right) + 2^{-3l}\left(\frac{11}{4}\mu^2T^5 + \frac{1545}{32}T^4\right) + 2^{-2l}\left(\frac{163}{64}T^4\right).$$

Dividing by 16, we get the desired result

$$Y = 2^{-4l}\left(\frac{3}{16}\mu^4T^6 + \frac{9}{16}\mu^2T^5\right) + 2^{-3l}\left(\frac{11}{64}\mu^2T^5 + \frac{1545}{512}T^4\right) + 2^{-2l}\left(\frac{163}{1024}T^4\right).$$

Chapter 3

Asymptotics for the normalized error of the Ninomiya-Victoir scheme

This chapter corresponds to an article written with Émmanuelle Clément and Benjamin Jourdain [3]. It has been submitted for publication.

Abstract. In a previous work, we proved strong convergence with order $1/2$ of the Ninomiya-Victoir scheme $X^{NV,\eta}$ with time step T/N to the solution X of the limiting SDE. In this paper we check that the normalized error defined by $\sqrt{N}\left(X - X^{NV,\eta}\right)$ converges to an affine SDE with source terms involving the Lie brackets between the Brownian vector fields. The limit does not depend on the Rademacher random variables η. This result can be seen as a first step to adapt to the Ninomiya-Victoir scheme the central limit theorem of Lindeberg Feller type, derived recently by Ben Alaya and Kebaier for the multilevel Monte Carlo estimator based on the Euler scheme. When the Brownian vector fields commute, the limit vanishes. This suggests that the rate of convergence is greater than $1/2$ in this case and we actually prove strong convergence with order 1.

3.1 Introduction

We consider a general n-dimensional stochastic differential equation, driven by a d-dimensional standard Brownian motion $W = \left(W^1, \dots, W^d\right)$, of the form

$$\begin{cases} dX_t = b(X_t)dt + \sum_{j=1}^{d} \sigma^j(X_t)dW_t^j, \ t \in [0,T] \\ X_0 = x \end{cases} \tag{3.1.1}$$

where $x \in \mathbb{R}^n$ is the starting point, $b : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is the drift coefficient and $\sigma^j : \mathbb{R}^n \longrightarrow \mathbb{R}^n, j \in \{1, \dots, d\}$, are the Brownian vector fields. We are interested in the study of the normalized error process for the Ninomiya-Victoir scheme. To do so we will consider in the whole paper a regular time grid, with time step $h = T/N$, of the time interval $[0,T]$. We introduce some notations to define the Ninomiya-Victoir scheme. Let

- $(t_k = kh)_{0 \le k \le N}$ be the subdivision of $[0,T]$ with equal time step h,

- $\Delta W_s^j = W_s^j - W_{t_k}^j$, for $s \in (t_k, t_{k+1}]$ and $j \in \{1, \ldots, d\}$,
- $\Delta s = s - t_k$, for $s \in (t_k, t_{k+1}]$,
- $\eta = (\eta_k)_{k\geq 1}$ be a sequence of independent, identically distributed Rademacher random variables independent of W.

Let $V : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be Lipschitz continuous and let $\exp(tV)x_0$ denote the solution, at time $t \in \mathbb{R}$, of the following ordinary differential equation in $\mathbb{R}^n$

$$\begin{cases} \dfrac{dx(t)}{dt} = V\left(x(t)\right) \\ x(0) = x_0. \end{cases}$$

To deal with the Ninomiya-Victoir scheme, it is more convenient to rewrite the stochastic differential equation (3.1.1) in Stratonovich form. Assuming $\mathcal{C}^1$ regularity for the vector fields, the Stratonovich form of (3.1.1) is given by

$$\begin{cases} dX_t = \sigma^0(X_t)dt + \sum_{j=1}^{d} \sigma^j(X_t) \circ dW_t^j \\ X_0 = x \end{cases}$$

where $\sigma^0 = b - \frac{1}{2}\sum_{j=1}^{d} \partial\sigma^j\sigma^j$ and $\partial\sigma^j$ is the Jacobian matrix of σ^j defined as follows

$$\partial\sigma^j = \left(\partial_{x_k}\sigma^{ij}\right)_{1\leq i,k\leq n}. \tag{3.1.2}$$

Now, we present the Ninomiya-Victoir scheme introduced in [44].

- Starting point: $X_{t_0}^{NV,\eta} = x$.
- For $k \in \{0\ldots, N-1\}$,
  if $\eta_{k+1} = 1$,

  $$X_{t_{k+1}}^{NV,\eta} = \exp\left(\frac{h}{2}\sigma^0\right)\exp\left(\Delta W_{t_{k+1}}^d \sigma^d\right)\ldots\exp\left(\Delta W_{t_{k+1}}^1 \sigma^1\right)\exp\left(\frac{h}{2}\sigma^0\right)X_{t_k}^{NV,\eta},$$

  and if $\eta_{k+1} = -1$,

  $$X_{t_{k+1}}^{NV,\eta} = \exp\left(\frac{h}{2}\sigma^0\right)\exp\left(\Delta W_{t_{k+1}}^1 \sigma^1\right)\ldots\exp\left(\Delta W_{t_{k+1}}^d \sigma^d\right)\exp\left(\frac{h}{2}\sigma^0\right)X_{t_k}^{NV,\eta}.$$

The strong convergence properties of a numerical scheme, which approximates the diffusion (3.1.1), are useful to control the variance of the multilevel Monte Carlo estimator based on this scheme (see [22] and [38]). This motivated our study of the strong convergence of the Ninomiya-Victoir scheme in Chapter 2. More precisely, under some regularity assumptions on the coefficients of the SDE, we proved strong convergence with order $1/2$: for all $p \geq 1$, there exists a constant $C_{NV} \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$,

$$\mathbb{E}\left[\max_{0\leq k\leq N}\left\|X_{t_k} - X_{t_k}^{NV,\eta}\right\|^{2p}\right] \leq C_{NV}\left(1 + \|x\|^{2p}\right)h^p.$$

In this present paper, we focus on the convergence in law of the normalized error defined by $\sqrt{N}\left(X - X^{NV,\eta}\right)$. The asymptotic distribution of the normalized error for the continuous time Euler scheme was established by Kurtz and Protter in [35]. The asymptotic behavior of the normalized error process for the continuous time Milstein scheme [42], which is known to exhibit strong convergence with order 1, was studied by Yan in [55]. In both cases, the normalized error converges to the solution of an affine SDE with a source term involving additional randomness given by a Brownian motion independent of the one driving both the SDE and the scheme.

This paper is organized as follows. In Section 2, we recall basic facts about the theory of stable convergence in law, introduced by Rényi [50] and developed by Jacod [30] and Jacod-Protter [31]. In Section 3, we will discuss the interpolation between time grid points and then derive the asymptotic error distribution for the Ninomiya-Victoir scheme in the general case. More precisely, we prove the stable convergence in law of $\sqrt{N}\left(X - X^{NV,\eta}\right)$ to the solution of the following SDE

$$V_t = \sqrt{\frac{T}{2}} \sum_{j=1}^{d} \sum_{m=1}^{j-1} \int_0^t \left[\sigma^j, \sigma^m\right] (X_s)\, dB_s^{j,m} + \int_0^t \partial b\,(X_s)\, V_s ds + \sum_{j=1}^{d} \int_0^t \partial \sigma^j\,(X_s)\, V_s dW_s^j,$$

where $\left[\sigma^j, \sigma^m\right] = \partial\sigma^m \sigma^j - \partial\sigma^j \sigma^m$, for $j, m \in \{1, \ldots, d\}\,, m < j$, denotes the Lie bracket between the Brownian vector fields σ^j and σ^m, ∂b is the Jacobian matrix of b, defined analogously to (3.1.2), and $(B_t)_{0 \le t \le T}$ is a standard $d(d-1)/2$-dimensional Brownian motion independent of W. This result ensures that the strong convergence rate is actually $1/2$. Moreover, it can be seen as a first step to adapt to the Ninomiya-Victoir scheme the central limit theorem of Lindeberg Feller type, derived by Ben Alaya and Kebaier in [10] for the multilevel Monte Carlo estimator based on the Euler scheme. Their approach leads to an accurate description of the optimal choice of the parameters for the multilevel Monte Carlo estimator. When the Brownian vector fields commute, the limit vanishes, which suggests that the rate of convergence is greater than $1/2$. In Section 4, we focus on the commutative case and we provide a suitable interpolation between time grid points, to show strong convergence with order 1.

3.2 Stable convergence

We start with the definition of the stable convergence in law which is stronger than the convergence in law.

Definition 3.2.1 *Let $\left(Z^N\right)_{N \in \mathbb{N}}$ be a sequence of random variables all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with values in a metric space (E, d). Let $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ be an "extension" of $(\Omega, \mathcal{F}, \mathbb{P})$, and let Z be an E-valued variable on this extension. The sequence $\left(Z^N\right)_{N \in \mathbb{N}}$ stably converges in law to Z, and we write this convergence as follows*

$$Z^N \underset{N \to +\infty}{\overset{stably}{\Longrightarrow}} Z,$$

if, and only if, for all $f : E \longrightarrow \mathbb{R}$ bounded continuous and for all bounded random variable Ξ on $(\Omega, \mathcal{F}, \mathbb{P})$

$$\mathbb{E}\left[f\left(Z^N\right)\Xi\right] \underset{N \to +\infty}{\longrightarrow} \mathbb{E}^*\left[f\,(Z)\,\Xi\right].$$

We do not go into details of the definition of an "extension" (see [30] for more information). The purpose of this section is to recall basic facts about stable convergence to study a sequence of stochastic differential equations in $\mathbb{R}^n$ of the form

$$U_t^N = R_t^N + J_t^N + \left(\int_0^t H_s^{0,N} U_s^N ds + \sum_{j=1}^d \int_0^t H_s^{j,N} U_s^N dW_s^j \right), \tag{3.2.1}$$

where $H^{j,N}$, for $j \in \{0, \dots, d\}$, take values in $\mathbb{R}^n \times \mathbb{R}^n$, R^N is a remainder term and J^N a source term. This is motivated by the decomposition of the error process (3.3.7).

The following fundamental proposition will be used to study the stable convergence in law of a random sequence of couple of variables (see section 2-1 in [30]).

Proposition 3.2.2 *Let $\left(\Lambda^N\right)_{N\in\mathbb{N}}$ and $\left(\Gamma^N\right)_{N\in\mathbb{N}}$ be two sequences of random variables all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with values in a metric space (E, d), and Λ be a random variable on an extension, with values in (E, d). Let $\left(\Theta^N\right)_{N\in\mathbb{N}}$ be a sequence of random variables and Θ be a random variable all defined on $(\Omega, \mathcal{F}, \mathbb{P})$, with values in an other metric space (E', d'). Then*

$$\text{if } \Lambda^N \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} \Lambda \text{ and } d\left(\Lambda^N, \Gamma^N\right) \underset{N\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0 \text{ then } \Gamma^N \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} \Lambda, \tag{3.2.2}$$

$$\text{if } \Lambda^N \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} \Lambda \text{ and } d'\left(\Theta^N, \Theta\right) \underset{N\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0 \text{ then } \left(\Lambda^N, \Theta^N\right) \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} (\Lambda, \Theta), \tag{3.2.3}$$

for the product topology on $E \times E'$.

In the following, we work on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where

$$\mathbb{F} = \left(\mathcal{F}_t = \sigma\left(\eta, W_s, s \le t\right)\right)_{0\le t\le T}.$$

We consider the metric space $E = \mathcal{C}\left([0,T], \mathbb{R}^n\right)$ equipped with the supremum-norm. The following theorem, dedicated to the convergence of a sequence of semimartingales, is a simplified version of Theorem 2.1 in [30].

Theorem 3.2.3 *Let $\left(Y^N\right)_{N\in\mathbb{N}}$ be a sequence of continuous semimartingales with values in $\mathbb{R}^p$, such that $Y^N = M^N + A^N$, for all $N \in \mathbb{N}$, where M^N is a sequence of continuous $\mathbb{F}$-local martingales null at $t = 0$ and A^N is a sequence of $\mathbb{F}$-predictable continuous processes with finite variation. Assume that, there exist A and f such that*

$$\sup_{t\le T} \left\| A_t^N - A_t \right\| \underset{N\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0,$$

$$\text{for all } i,j \in \{1,\dots,p\}, \text{ and all } t \in [0,T], \left\langle M^{i,N}, M^{j,N} \right\rangle_t \underset{N\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} F_t^{ij} = \int_0^t f_s^{ij} ds,$$

$$\text{for all } \in \{1,\dots,p\}, \in \{1,\dots,d\}, \text{ and all } t \in [0,T], \left\langle M^{i,N}, W^k \right\rangle_t \underset{N\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0.$$

Then,

$$Y^N \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} Y,$$

where

$$Y_t = A_t + \int_0^t (f_s)^{\frac{1}{2}} dB_s,$$

and $(f_s)^{\frac{1}{2}}$ is the square root of the positive semi-definite matrix

$$f_s = \left(f_s^{ij}\right)_{1\le i,j\le p},$$

and B a p-dimensional standard Brownian motion defined on a Wiener space $\left(\Omega^B, \mathcal{F}^B, \mathbb{P}^B\right)$ and independent of W. The stable convergence takes place in the canonical Wiener extension of W, denoted by $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ defined as follows

$$\Omega^* = \Omega \times \Omega^B,\ \mathcal{F}^* = \mathcal{F} \otimes \mathcal{F}^B,\ \mathbb{P}^* = \mathbb{P} \otimes \mathbb{P}^B.$$

In comparison with Theorem 2.1 of [30], the assumption,

$$\left\langle M^i, N\right\rangle_t = 0, \text{ for all } i \in \{1, \dots, p\} \text{ and all } N \text{ a bounded martingale orthogonal to } W,$$

is obvious, since we can write M in terms of an Itô integral with respect to the Brownian motion W, by using the martingale representation theorem. We will use Theorem 3.2.3, together with the following proposition to study the source term J^N in the decomposition (3.2.1). This proposition is a consequence of Theorem 2.3 in [31] (see the proof of Theorem 2.5 (c) in [31]).

Proposition 3.2.4 *Let $\left(Y^N\right)_{N\in\mathbb{N}}$ be a sequence of continuous semimartingales with values in $\mathbb{R}^p$, such that $Y_t^N = Y_0^N + M_t^N + A_t^N$, for all $N \in \mathbb{N}, t \in [0,T]$, where M^N is a sequence of continuous $\mathbb{F}$-local martingales null at $t = 0$ and A^N is a sequence of $\mathbb{F}$-predictable continuous processes with finite variation null at $t = 0$. Assume that the sequence $\left(\left\langle M^N\right\rangle_T + \int_0^T \left|dA_s^N\right|\right)_{N\in\mathbb{N}}$ is tight. Then, for any sequence $\left(K^N\right)_{N\in\mathbb{N}}$ of $\mathbb{F}$-predictable, right-continuous and left-hand limited processes, with values in $\mathbb{R}^q \otimes \mathbb{R}^p$, such that the sequence $\left(K^N, Y^N\right)$ stably converges in law to a limit (K, Y) we have the following result*
Y is a semimartingale and with respect to the filtration generated by the limit process (K, Y) and

$$\left(K^N, Y^N, \int K^N dY^N\right) \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} \left(K, Y, \int K dY\right),$$

where $\int K^N dY^N = \left(\int_0^t K_s^N dY_s^N\right)_{0\le t\le T}$ *and* $\int K dY = \left(\int_0^t K_s dY_s\right)_{0\le t\le T}$.

The following theorem deals with a sequence of stochastic differential equations in $\mathbb{R}^n$ of the form

$$U_t^N = R_t^N + J_t^N + \sum_{j=0}^d \int_0^t H_s^{j,N} U_s^N dW_s^j,$$

where, by convention, $dW_s^0 = ds$, $\left(J^N\right)_{N\in\mathbb{N}}$ is a sequence of continuous adapted processes, and for $j\ \{0, \dots, d\}$, $\left(H^{j,N}\right)_{N\in\mathbb{N}}$ is a sequence of $\mathbb{F}$-predictable, right-continuous and left-hand limited processes, with values in $\mathbb{R}^n \times \mathbb{R}^n$.

Theorem 3.2.5 *Assume that there exist $(H^j)_{0\le j\le d}$ and J such that*

- *for all* $j \in \{0, \dots, d\}, \sup_{t\le T} \left\|H_t^{j,N} - H_t^j\right\| \underset{N\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0,$
- $J^N \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} J,$
- $\sup_{t\le T} \left\|R_t^N\right\| \underset{N\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0.$

Then, U^N stably converges in law towards U, where U is the unique solution of the following affine stochastic differential equation

$$U_t = J_t + \sum_{j=0}^{d} \int_0^t H_s^j U_s dW_s^j.$$

Proof : On the one hand, denoting by

$$V_t^N = \sum_{j=0}^{d} \int_0^t H_s^{j,N} dW_s^j,$$

the first assumption ensures that

$$\sup_{t \leq T} \left\| V_t^N - V_t \right\| \underset{N \to +\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0,$$

where

$$V_t = \sum_{j=0}^{d} \int_0^t H_s^j dW_s^j.$$

On the other hand, (3.2.2) from Proposition 3.2.2 gives us

$$R^N + J^N \underset{N \to +\infty}{\overset{stably}{\Longrightarrow}} J.$$

Then, applying (3.2.3) from Proposition 3.2.2, we have

$$\left(R^N + J^N, V^N\right) \underset{N \to +\infty}{\overset{stably}{\Longrightarrow}} (J, V).$$

Finally, since $\left(\sup_{t \leq T} \left\| H_t^N \right\|\right)_{N \in \mathbb{N}^*}$ is tight, we get the desired result using Theorem 2.5 (c) in [31]. ∎

3.3 Asymptotic error distribution for the Ninomiya-Victoir scheme in the general case

3.3.1 Main result

In order to study the stable convergence in law of the normalized error process, we have to build an interpolated scheme. Let us first introduce some more notation.

- Let $\hat{\tau}_s$ be the last time discretization before $s \in [0, T]$, ie $\hat{\tau}_s = t_k$ if $s \in (t_k, t_{k+1}]$, and for $s = t_0 = 0$, we set $\hat{\tau}_0 = t_0$.
- Let $\check{\tau}_s$ be the first time discretization after $s \in [0, T]$, ie $\check{\tau}_s = t_{k+1}$ if $s \in (t_k, t_{k+1}]$, and for $s = t_0 = 0$, we set $\check{\tau}_0 = 0$.
- By a slight abuse of notation, we set $\eta_s = \eta_{k+1}$ if $s \in (t_k, t_{k+1}]$.

- To lighten up the notation, $\|.\|$ will denote both the Euclidean norm in $\mathbb{R}^n$ and its associated operator norm in $\mathbb{R}^n \otimes \mathbb{R}^n$.
- For a finite-dimensional normed vector space $(\mathcal{S}, \|.\|_{\mathcal{S}})$, $LIP_{loc}^{pgc}(\mathcal{S})$ denotes the space of locally Lipschitz with polynomially growing Lipschitz constant functions from $\mathbb{R}^n$ to $\mathcal{S}$:

$$LIP_{loc}^{pgc}(\mathcal{S}) = \left\{F : \mathbb{R}^n \longrightarrow \mathcal{S}, \exists c \in \mathbb{R}_+^*, q \in \mathbb{N}, \forall x, y \in \mathbb{R}^n, \|F(x) - F(y)\|_{\mathcal{S}} \leq c\,(1 + \|x\|^q \vee \|y\|^q)\, \|x - y\|\right\}.$$

Remark 3.3.1 *If $F \in \mathcal{C}^1(\mathbb{R}^n, \mathcal{S})$ with polynomially growing first order derivatives, then $F \in LIP_{loc}^{pgc}(\mathcal{S})$.*

A natural and adapted interpolation, at time $t \in [0, T]$, for the Ninomiya-Victoir scheme could be defined as follows

$$h_{\eta_t}\left(\frac{\Delta t}{2}, \Delta W_t, \frac{\Delta t}{2}; X_{\hat{\tau}_t}^{NV,\eta}\right),$$

where $\Delta W_t = \left(\Delta W_t^1, \ldots, \Delta W_t^d\right)$,

$$h_{-1}(t_0, \ldots, t_{d+1}; x) = \exp\left(t_0\sigma^0\right)\exp\left(t_1\sigma^1\right)\ldots\exp\left(t_d\sigma^d\right)\exp\left(t_{d+1}\sigma^0\right)x,$$

and

$$h_1(t_0, \ldots, t_{d+1}; x) = \exp\left(t_0\sigma^0\right)\exp\left(t_d\sigma^d\right)\ldots\exp\left(t_1\sigma^1\right)\exp\left(t_{d+1}\sigma^0\right)x.$$

Here, to compute the Itô decomposition of $\left(h_{\eta_t}\left(\frac{\Delta t}{2}, \Delta W_t, \frac{\Delta t}{2}; X_{\hat{\tau}_t}^{NV,\eta}\right)\right)_{0\leq t\leq T}$ the main difficulty is to explicit the derivatives of h_η. In the general case, the computation of derivatives of this function is quite complicated. For this reason, in this paper, we use the interpolation of the Ninomiya-Victoir introduced in Chapter 2

$$\begin{cases} dX_t^{NV,\eta} = \sum_{j=1}^{d} \sigma^j(\bar{X}_t^{j,\eta})dW_t^j + \frac{1}{2}\sum_{j=1}^{d} \partial\sigma^j\sigma^j\left(\bar{X}_t^{j,\eta}\right)dt + \frac{1}{2}\left(\sigma^0\left(\bar{X}_t^{0,\eta}\right) + \sigma^0\left(\bar{X}_t^{d+1,\eta}\right)\right)dt \\ X_0^{NV,\eta} = x, \end{cases} \tag{3.3.1}$$

where for $s \in (t_k, t_{k+1}]$

$$\bar{X}_s^{0,\eta} = \exp\left(\frac{\Delta s}{2}\sigma^0\right)\left(X_{t_k}^{NV,\eta}\mathbb{1}_{\{\eta_{k+1}=1\}} + \bar{X}_{t_{k+1}}^{1,\eta}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right),$$

$$\text{for } j \in \{1, \ldots, d\}, \bar{X}_s^{j,\eta} = \exp\left(\Delta W_s^j\sigma^j\right)\left(\bar{X}_{t_{k+1}}^{j-1,\eta}\mathbb{1}_{\{\eta_{k+1}=1\}} + \bar{X}_{t_{k+1}}^{j+1,\eta}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right),$$

$$\bar{X}_s^{d+1,\eta} = \exp\left(\frac{\Delta s}{2}\sigma^0\right)\left(\bar{X}_{t_{k+1}}^{d,\eta}\mathbb{1}_{\{\eta_{k+1}=1\}} + X_{t_k}^{NV,\eta}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right).$$

Although the stochastic processes $\left(\bar{X}_t^{j,\eta}\right)_{0\leq t\leq T}$, for $j \in \{1, \ldots d\}$, are not adapted to the filtration $\mathbb{F}$, each stochastic integral is well defined in (3.3.1). Indeed, $\left(\bar{X}_t^{j,\eta}\right)_{0\leq t\leq T}$ is adapted with respect to the filtration

$$\left(\sigma\left(\eta, W_s^j, s \leq t\right)\bigvee\left(\bigvee_{k\neq j}\sigma\left(W_s^k, s \leq T\right)\right)\right)_{0\leq t\leq T},$$

for $j \in \{1, \dots d\}$. Then, by independence, W^j is a also a Brownian motion with respect to this filtration and the stochastic integral

$$\int_0^t \sigma^j(\bar{X}_s^{j,\eta})dW_s^j$$

is well defined for all $t \in [0,T]$. Using this interpolation, we proved in Chapter 2 the strong convergence with order 1/2.

Theorem 3.3.2 *Assume that*

- *for all* $j \in \{1,\dots,d\}, \sigma^j \in \mathcal{C}^1(\mathbb{R}^n,\mathbb{R}^n)$,
- σ^0, σ^j *and* $\partial\sigma^j\sigma^j$, *for all* $j \in \{1,\dots,d\}$, *are Lipschitz continuous functions.*

Then, for all $p \geq 1$, *there exists a constant* $C_{NV} \in \mathbb{R}_+^*$ *such that for all* $N \in \mathbb{N}^*$,

$$\mathbb{E}\left[\sup_{t\leq T}\left\|X_t - X_t^{NV,\eta}\right\|^{2p}\right] \leq C_{NV}\left(1+\|x\|^{2p}\right)h^p.$$

Then, the normalized error process is defined as follows

$$V^N = \sqrt{N}\left(X - X^{NV,\eta}\right).$$

In this section, we check that the normalized error V^N converges to an affine SDE with source terms. Here is the main result.

Theorem 3.3.3 *Assume that*

- $\sigma^0 \in \mathcal{C}^2(\mathbb{R}^n,\mathbb{R}^n)$ *and is a Lipschitz continuous function with polynomially growing second order derivatives,*
- *for all* $j \in \{1,\dots,d\}, \sigma^j \in \mathcal{C}^2(\mathbb{R}^n,\mathbb{R}^n)$ *and is Lipschitz continuous and its first order derivative* $\partial\sigma^j \in LIP_{loc}^{pgc}(\mathbb{R}^n \otimes \mathbb{R}^n)$,
- *for all* $j,m \in \{1,\dots,d\}, \partial\sigma^j\sigma^m$ *is Lipschitz continuous,*
- *for all* $j \in \{1,\dots,d\}, \partial\sigma^j\sigma^j \in \mathcal{C}^2(\mathbb{R}^n,\mathbb{R}^n)$ *with polynomially growing second order derivatives.*

Then

$$V^N \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} V,$$

where V *is the unique solution of the following affine equation*

$$V_t = \sqrt{\frac{T}{2}}\sum_{j=1}^{d}\sum_{m=1}^{j-1}\int_0^t\left[\sigma^j,\sigma^m\right](X_s)\,dB_s^{j,m} + \int_0^t \partial b(X_s)\,V_s ds + \sum_{j=1}^{d}\int_0^t \partial\sigma^j(X_s)\,V_s dW_s^j, \quad (3.3.2)$$

with $\left[\sigma^j,\sigma^m\right] = \partial\sigma^m\sigma^j - \partial\sigma^j\sigma^m$, *and* $(B_t)_{0\leq t\leq T}$ *is a standard* $d(d-1)/2$*-dimensional Brownian motion independent of* W.

3.3.2 Discrete scheme

To compute the asymptotic error distribution, the method consists in writing the normalized error in the form (3.2.1) . Since the interpolation (3.3.1) is not adapted to the natural filtration of the Brownian motion W, we were not able to derive a decomposition (3.2.1) with V^N replacing U^N. To get around this problem, we build an adapted approximation $\hat{X}^{D,\eta}$ of $X^{NV,\eta}$, with order $1-\epsilon$, for all $\epsilon > 0$, and introduce

$$U^N = \sqrt{N}\left(X - \hat{X}^{D,\eta}\right).$$

Then, we obtain the decomposition of the form (3.2.1) (see (3.3.7)) and study the satble convergence in law of U^N to deduce the convergence of V^N. The approximation is defined as follows

$$\begin{cases} \hat{X}_t^{D,\eta} = \hat{X}_{\hat{\tau}_t}^{D,\eta} + b\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)\Delta t + \sum_{j=1}^{d}\sigma^j\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)\Delta W_t^j \\ \qquad + \frac{1}{2}\sum_{j=1}^{d}\partial\sigma^j\sigma^j\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)\left(\left(\Delta W_t^j\right)^2 - \Delta t\right) + \sum_{\eta_t m < \eta_t j}\partial\sigma^j\sigma^m\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)\Delta W_t^m \Delta W_t^j \\ \hat{X}_0^{D,\eta} = x. \end{cases} \tag{3.3.3}$$

In the following proposition, we compare $X^{NV,\eta}$ and $\hat{X}^{D,\eta}$.

Proposition 3.3.4 *Let $p \geq 1$ and $\epsilon > 0$. Under the assumptions of Theorem 3.3.3, there exists a constant $C_D \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$,*

$$\mathbb{E}\left[\sup_{t\leq T}\left\|X_t^{NV,\eta} - \hat{X}_t^{D,\eta}\right\|^{2p}\right] \leq C_D\frac{1}{N^{2p-\epsilon}}.$$

The proof of this proposition is postponed to the Appendix.

The next lemma gives estimation of the moment of $\hat{X}^{D,\eta}$ and its increments. Its hypotheses are consequences of the ones of Theorem 3.3.3. We omit its standard proof.

Lemma 3.3.5 *Assume that*

- *$b \in \mathcal{C}^0(\mathbb{R}^n, \mathbb{R}^n)$ has an affine growth,*
- *for all $j \in \{1, \dots, d\}, \sigma^j$ has an affine growth,*
- *for all $j, m \in \{1, \dots, d\}, \partial\sigma^j\sigma^m$ has an affine growth.*

Then, for all $p \geq 1$, there exists a constant $\hat{C}_D \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\hat{X}_t^{D,\eta}\right\|^{2p}\right] \leq \hat{C}_D, \tag{3.3.4}$$

and

$$\text{for all } t \in [0,T], \mathbb{E}\left[\left\|\hat{X}_t^{D,\eta} - \hat{X}_{\hat{\tau}_t}^{D,\eta}\right\|^{2p}\right] \leq \hat{C}_D h^p. \tag{3.3.5}$$

3.3.3 Proof of the stable convergence

We recall that $U^N = \sqrt{N}\left(X - \hat{X}^{D,\eta}\right)$. By Proposition 3.3.4,

$$\sup_{t\leq T}\sqrt{N}\left\|\hat{X}_t^{D,\eta} - X_t^{NV,\eta}\right\|$$

converges in probability to 0 as N goes to $+\infty$. Since

$$V^N - U^N = \sqrt{N}\left(\hat{X}^{D,\eta} - X^{NV,\eta}\right),$$

Property (3.2.3) from Proposition 3.2.2 ensures that Theorem 3.3.3 is a consequence of the following proposition dedicated to the stable convergence in law of U^N.

Proposition 3.3.6 *Under the assumptions of Theorem 3.3.3*

$$U^N \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} V,$$

where V is the unique solution of (3.3.2).

Proof : We begin by describing the limiting process for $U^N = \sqrt{N}\left(X - \hat{X}^{D,\eta}\right)$. The differential of U^N can be written as

$$\begin{aligned}
dU_t^N &= \sqrt{N}\left(\left(b\left(X_t\right) - b\left(\hat{X}_t^{D,\eta}\right)\right)dt + \sum_{j=1}^d\left(\sigma^j\left(X_t\right) - \sigma^j\left(\hat{X}_t^{D,\eta}\right)\right)dW_t^j\right)\\
&+ \sqrt{N}\left(\left(b\left(\hat{X}_t^{D,\eta}\right) - b\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)\right)dt + \sum_{j=1}^d\left(\sigma^j\left(\hat{X}_t^{D,\eta}\right) - \sigma^j\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)\right)dW_t^j\right)\\
&- \sqrt{N}\sum_{j=1}^d \partial\sigma^j\sigma^j\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)\Delta W_t^j dW_t^j\\
&- \sqrt{N}\left(\sum_{\eta_t m<\eta_t j}\partial\sigma^j\sigma^m\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)\Delta W_t^m dW_t^j + \sum_{\eta_t m<\eta_t j}\partial\sigma^j\sigma^m\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)\Delta W_t^j dW_t^m\right).
\end{aligned} \tag{3.3.6}$$

Then, the proof will go through several steps.

Step 1: linearisation of the two terms in the first line of the right-hand side of (3.3.6).

$$\sqrt{N}\left(\left(b\left(X_t\right) - b\left(\hat{X}_t^{D,\eta}\right)\right)dt + \sum_{j=1}^d\left(\sigma^j\left(X_t\right) - \sigma^j\left(\hat{X}_t^{D,\eta}\right)\right)dW_t^j\right).$$

Let $j \in \{1,\ldots,d\}$ and $i \in \{1,\ldots,n\}$. By the mean value theorem, we get

$$\sigma^{ij}\left(X_t\right) - \sigma^{ij}\left(\hat{X}_t^{D,\eta}\right) = \nabla\sigma^{ij}\left(\xi_t^{ij}\right).\left(X_t - \hat{X}_t^{D,\eta}\right),$$

where $\xi_t^{ij} = \alpha_t^{ij}X_t + \left(1-\alpha_t^{ij}\right)\hat{X}_t^{D,\eta}$ for some $\alpha_t^{ij} \in [0,1]$. Using a compact matrix notation, we can write

$$\sigma^j\left(X_t\right) - \sigma^j\left(\hat{X}_t^{D,\eta}\right) = \partial\sigma_t^{j,N}\left(X_t - \hat{X}_t^{D,\eta}\right),$$

where

$$\left(\partial\sigma_t^{j,N}\right)_{i,m} = \partial_{x_m}\sigma^{ij}\left(\xi_t^{ij}\right).$$

Then, we obtain

$$\sqrt{N}\sum_{j=1}^{d}\left(\sigma^j\left(X_t\right) - \sigma^j\left(\hat{X}_t^{D,\eta}\right)\right)dW_t^j = \sum_{j=1}^{d}\partial\sigma_t^{j,N}U_t^N dW_t^j.$$

In the same way

$$\sqrt{N}\left(b\left(X_t\right) - b\left(\hat{X}_t^{D,\eta}\right)\right)dt = \partial b_t^N U_t^N dt,$$

where

$$\left(\partial b_t^N\right)_{i,m} = \partial_{x_m}b^i\left(\xi_t^{i0}\right)$$

and $\xi_t^{i0} = \alpha_t^{i0}X_t + \left(1-\alpha_t^{i0}\right)\hat{X}_t^{D,\eta}$ for some $\alpha_t^{i0}\in[0,1]$.

Step 2: decomposition of U^N.

Writing the fourth term in the right-hand side of (3.3.6), $\sigma^j\left(\hat{X}_t^{D,\eta}\right) - \sigma^j\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)$, as the sum of the dominant contribution

$$\sum_{m=1}^{d}\partial\sigma^j\sigma^m\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)\Delta W_t^m,$$

and the remainder

$$\sigma^j\left(\hat{X}_t^{D,\eta}\right) - \sigma^j\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right) - \sum_{m=1}^{d}\partial\sigma^j\sigma^m\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)\Delta W_t^m,$$

which is of order $1/N$, we deduce that

$$U_t^N = R_t^N + J_t^N + \left(\int_0^t \partial b_s^N U_s^N ds + \sum_{j=1}^{d}\int_0^t \partial\sigma_s^{j,N}U_s^N dW_s^j\right), \tag{3.3.7}$$

where

$$\begin{aligned}R_t^N = \sqrt{N}\Bigg(&\sum_{j=1}^{d}\int_0^t\left(\sigma^j\left(\hat{X}_s^{D,\eta}\right) - \sigma^j\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) - \sum_{m=1}^{d}\partial\sigma^j\sigma^m\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)\Delta W_s^m\right)dW_s^j \\ &+\int_0^t b\left(\hat{X}_s^{D,\eta}\right) - b\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)ds\Bigg),\end{aligned}$$

and

$$\begin{aligned}J_t^N = -\sqrt{N}\Bigg(&\sum_{j=1}^{d}\int_0^t\partial\sigma^j\sigma^j\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)\Delta W_s^j dW_s^j + \int_0^t\sum_{\eta_s m<\eta_s j}\partial\sigma^j\sigma^m\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)\Delta W_s^m dW_s^j \\ &+\int_0^t\sum_{\eta_s m<\eta_s j}\partial\sigma^j\sigma^m\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)\Delta W_s^j dW_s^m - \sum_{j=1}^{d}\sum_{m=1}^{d}\int_0^t\partial\sigma^j\sigma^m\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)\Delta W_s^m dW_s^j\Bigg).\end{aligned}$$

The expression of J^N can be arranged as follows

$$J_t^N = \sqrt{N}\sum_{j=1}^{d}\sum_{m=1}^{j-1}\int_0^t \left[\sigma^j,\sigma^m\right]\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)\frac{-1+\eta_s}{2}\Delta W_s^m dW_s^j$$
$$+\sqrt{N}\sum_{j=1}^{d}\sum_{m=1}^{j-1}\int_0^t \left[\sigma^j,\sigma^m\right]\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)\frac{1+\eta_s}{2}\Delta W_s^j dW_s^m.$$

Step 3: stable convergence in law of J^N.

To lighten up the notations, we introduce

- $K_t^{j,m,N} = \left[\sigma^j,\sigma^m\right]\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)$ for $j,m\,\{1,\ldots,d\}$, $m<j$,
- $\Psi_t^1 = \frac{-1+\eta_t}{2}$ and $\Psi_t^2 = \frac{1+\eta_t}{2}$,
- $Y_t^{j,m,N} = \sqrt{N}\left(\int_0^t \Psi_s^1\Delta W_s^m dW_s^j + \int_0^t \Psi_s^2\Delta W_s^j dW_s^m\right)$ for $j,m\,\{1,\ldots,d\}$, $m<j$.

Then, we have that

$$\Psi^1\Psi^2 = 0,$$
$$\left(\Psi^1\right)^2+\left(\Psi^2\right)^2 = 1,$$

and

$$J_t^N = \sum_{j=1}^{d}\sum_{m=1}^{j-1}\int_0^t K_s^{j,m,N}dY_s^{j,m,N}.$$

With a view to apply Proposition 3.2.4, in order to obtain the stable convergence in law of J^N, we first study the stable convergence in law of Y^N. By virtue of Theorem 3.2.3, it is enough to study the asymptotic behavior of $\left\langle Y^{j,m,N},W^k\right\rangle_t$, for $t\in[0,T]$, $j,m,k\in\{1,\ldots,d\}$ such that $m<j$, and $\left\langle Y^{j,m,N},Y^{l,k,N}\right\rangle_t$, for $t\in[0,T]$, $j,m,k,l\in\{1,\ldots,d\}$ such that $m<j$, $k<l$.

Step 3.1: asymptotic behavior of $\left\langle Y^{j,m,N},W^k\right\rangle_t$ for $j,m,k\in\{1,\ldots,d\}, m<j$ and $t\in[0,T]$.

Computing the quadratic covariation we get

$$\left\langle Y^{j,m,N},W^k\right\rangle_t = \sqrt{N}\left(\mathbb{1}_{\{j=k\}}\int_0^t \Psi_s^1\Delta W_s^m ds + \mathbb{1}_{\{m=k\}}\int_0^t \Psi_s^2\Delta W_s^j ds\right).$$

Then, computing the L^2-norm, we obtain

$$\left\|\left\langle Y^{j,m,N},W^k\right\rangle_t\right\|_2^2 = N\,\mathbb{E}\left[\left(\mathbb{1}_{\{j=k\}}\int_0^t \Psi_s^1\Delta W_s^m ds + \mathbb{1}_{\{m=k\}}\int_0^t \Psi_s^2\Delta W_s^j ds\right)^2\right]$$
$$= 2N\Bigg(\mathbb{1}_{\{j=k\}}\int_0^t\int_0^s \mathbb{E}\left[\Psi_s^1\Psi_u^1\right]\mathbb{E}\left[\Delta W_s^m\Delta W_u^m\right]du\ ds$$
$$+\mathbb{1}_{\{m=k\}}\int_0^t\int_0^s \mathbb{E}\left[\Psi_s^2\Psi_u^2\right]\mathbb{E}\left[\Delta W_s^j\Delta W_u^j\right]du\ ds\Bigg).$$

Since for all $u, s \in [0,t]$ such that $u \le s$,

$$\mathbb{E}\left[\Delta W_s^m \Delta W_u^m\right] = u - u \wedge \hat{\tau}_s \ge 0,$$

$$0 \le \mathbb{E}\left[\Psi_s^1 \Psi_u^1\right] = \mathbb{E}\left[\Psi_s^2 \Psi_u^2\right] \le \frac{1}{2},$$

and $m < j$, then it follows that

$$\begin{aligned}\left\|\left\langle Y^{j,m,N}, W^k\right\rangle_t\right\|_2^2 &\le N \int_0^t \int_0^s u - u \wedge \hat{\tau}_s \, du \, ds = N \int_0^t \int_{\hat{\tau}_s}^s u - \hat{\tau}_s \, du \, ds = \frac{1}{2} N \int_0^t (s - \hat{\tau}_s)^2 \, ds \\ &\le \frac{1}{2} N \int_0^T (s - \hat{\tau}_s)^2 \, ds = \frac{T^3}{6N} \underset{N \to +\infty}{\longrightarrow} 0.\end{aligned}$$

Step 3.2: asymptotic behavior of $\left\langle Y^{j,m,N}, Y^{l,k,N}\right\rangle_t$ for $j, m, k, l \in \{1, \dots, d\}, m < j, k < l$ and $t \in [0,T]$.

Computing the quadratic covariation we get

$$\begin{aligned}\left\langle Y^{j,m,N}, Y^{l,k,N}\right\rangle_t = N \Bigg(& \mathbb{1}_{\{j=l\}} \int_0^t \left(\Psi_s^1\right)^2 \Delta W_s^m \Delta W_s^k ds + \mathbb{1}_{\{m=k\}} \int_0^t \left(\Psi_s^2\right)^2 \Delta W_s^j \Delta W_s^l ds \\ & + \mathbb{1}_{\{j=k\}} \int_0^t \Psi_s^1 \Psi_s^2 \Delta W_s^l \Delta W_s^m ds + \mathbb{1}_{\{m=l\}} \int_0^t \Psi_s^1 \Psi_s^2 \Delta W_s^j \Delta W_s^k ds \Bigg).\end{aligned}$$

Since $\Psi^1 \Psi^2 = 0$, we obtain

$$\left\langle Y^{j,m,N}, Y^{l,k,N}\right\rangle_t = N \left(\mathbb{1}_{\{j=l\}} \int_0^t \left(\Psi_s^1\right)^2 \Delta W_s^m \Delta W_s^k ds + \mathbb{1}_{\{m=k\}} \int_0^t \left(\Psi_s^2\right)^2 \Delta W_s^j \Delta W_s^l ds \right).$$

Then, we distinguish three cases. In the case $j = l$ and $m \ne k$, the bracket is given by

$$\left\langle Y^{j,m,N}, Y^{l,k,N}\right\rangle_t = N \int_0^t \left(\Psi_s^1\right)^2 \Delta W_s^m \Delta W_s^k ds.$$

By independence, the L^2-norm of the bracket $\left\langle Y^{j,m}, Y^{j,k}\right\rangle_t$ is given by

$$\begin{aligned}\left\|\left\langle Y^{j,m,N}, Y^{j,k,N}\right\rangle_t\right\|_2^2 &= 2N^2 \int_0^t \int_0^s \mathbb{E}\left[\left(\Psi_s^1\right)^2 \left(\Psi_u^1\right)^2\right] \mathbb{E}\left[\Delta W_s^m \Delta W_u^m\right] \mathbb{E}\left[\Delta W_s^k \Delta W_u^k\right] du \, ds \\ &= 2N^2 \int_0^t \int_0^s \mathbb{E}\left[\left(\Psi_s^1\right)^2 \left(\Psi_u^1\right)^2\right] \left(\mathbb{E}\left[\Delta W_s^m \Delta W_u^m\right]\right)^2 du \, ds.\end{aligned}$$

using once again

$$\mathbb{E}\left[\Delta W_s^m \Delta W_u^m\right] = u - u \wedge \hat{\tau}_s \ge 0,$$

and

$$0 \le \mathbb{E}\left[\left(\Psi_s^1\right)^2 \left(\Psi_u^1\right)^2\right] \le \frac{1}{2},$$

for all $u, s \in [0,t], u \le s$, we get

$$\begin{aligned}\left\|\left\langle Y^{j,m,N}, Y^{j,k,N}\right\rangle_t\right\|_2^2 &\le N^2 \int_0^t \int_0^s (u - u \wedge \hat{\tau}_s)^2 du \, ds = N^2 \int_0^t \int_{\hat{\tau}_s}^s (u - \hat{\tau}_s)^2 du \, ds \\ &= \frac{1}{3} N^2 \int_0^t (s - \hat{\tau}_s)^3 ds \le \frac{1}{3} N^2 \int_0^T (s - \hat{\tau}_s)^3 ds = \frac{1}{12} \frac{T^4}{N} \underset{N \to +\infty}{\longrightarrow} 0.\end{aligned}$$

The second case $k = m$ and $j \neq l$, is similar to the first case since

$$\left\langle Y^{j,k,N}, Y^{l,k,N} \right\rangle_t = N \int_0^t \left(\Psi_s^2\right)^2 \Delta W_s^j \Delta W_s^l ds.$$

As previously, we have

$$\left\| \left\langle Y^{j,k,N}, Y^{l,k,N} \right\rangle_t \right\|_2 \underset{N \to +\infty}{\longrightarrow} 0.$$

The third and last case $k = m$ and $j = l$ provides a nonzero limit

$$\left\langle Y^{j,m,N}, Y^{j,m,N} \right\rangle_t = N \int_0^t \left(\left(\Psi_s^1\right)^2 (\Delta W_s^m)^2 + \left(\Psi_s^2\right)^2 \left(\Delta W_s^j\right)^2 \right) ds.$$

To identify the limit we proceed to a preliminary calculus of the expectation of this bracket

$$\mathbb{E}\left[\left\langle Y^{j,m,N}, Y^{j,m,N} \right\rangle_t \right] = N \int_0^t \left(\mathbb{E}\left[\left(\Psi_s^1\right)^2 \right] \mathbb{E}\left[(\Delta W_s^m)^2 \right] + \mathbb{E}\left[\left(\Psi_s^2\right)^2 \right] \mathbb{E}\left[\left(\Delta W_s^j\right)^2 \right] \right) ds.$$

Since $\left(\Psi^1\right)^2 + \left(\Psi^2\right)^2 = 1$, we get

$$\begin{aligned}
\mathbb{E}\left[\left\langle Y^{j,m,N}, Y^{j,m,N} \right\rangle_t \right] &= N \int_0^t (s - \hat{\tau}_s)\ ds = N \int_0^{\hat{\tau}_t} (s - \hat{\tau}_s)\ ds + O\left(\frac{1}{N}\right) \\
&= \frac{1}{2} Tt + O\left(\frac{1}{N}\right) \underset{N \to +\infty}{\longrightarrow} \frac{1}{2} Tt.
\end{aligned}$$

Now, we show the convergence in L^2 of $\left\langle Y^{j,m,N}, Y^{j,m,N} \right\rangle_t$ towards $Tt/2$. Computing the L^2-norm of the difference between the quadratic variation $\left\langle Y^{j,m}, Y^{j,m} \right\rangle_t$ and $Tt/2$, we obtain

$$\begin{aligned}
\left\| \left\langle Y^{j,m,N}, Y^{j,m,N} \right\rangle_t - \frac{1}{2} Tt \right\|_2^2 &= \left\| \left\langle Y^{j,m,N}, Y^{j,m,N} \right\rangle_t \right\|_2^2 - 2\mathbb{E}\left[\left\langle Y^{j,m,N}, Y^{j,m,N} \right\rangle_t \right] \frac{1}{2} Tt + \left(\frac{1}{2} Tt \right)^2 \\
&= \left\| \left\langle Y^{j,m,N}, Y^{j,m,N} \right\rangle_t \right\|_2^2 - \left(\frac{1}{2} Tt \right)^2 + O\left(\frac{1}{N}\right).
\end{aligned}$$

To prove the convergence in L^2, it suffices to show that

$$\left\| \left\langle Y^{j,m,N}, Y^{j,m,N} \right\rangle_t \right\|_2^2 \underset{N \to +\infty}{\longrightarrow} \left(\frac{1}{2} Tt \right)^2.$$

Computing the square of the L^2 norm of the bracket

$$\begin{aligned}
\left\| \left\langle Y^{j,m,N}, Y^{j,m,N} \right\rangle_t \right\|_2^2 = 2N^2 \Bigg(&\int_0^t \int_0^s \mathbb{E}\left[\left(\Psi_s^1\right)^2 \left(\Psi_u^1\right)^2 \right] \mathbb{E}\left[(\Delta W_s^m)^2 (\Delta W_u^m)^2 \right] du\ ds \\
&+ \int_0^t \int_0^s \mathbb{E}\left[\left(\Psi_s^1\right)^2 \left(\Psi_u^2\right)^2 \right] \mathbb{E}\left[(\Delta W_s^m)^2 \left(\Delta W_u^j\right)^2 \right] du\ ds \\
&+ \int_0^t \int_0^s \mathbb{E}\left[\left(\Psi_s^2\right)^2 \left(\Psi_u^1\right)^2 \right] \mathbb{E}\left[\left(\Delta W_s^j\right)^2 (\Delta W_u^m)^2 \right] du\ ds \\
&+ \int_0^t \int_0^s \mathbb{E}\left[\left(\Psi_s^2\right)^2 \left(\Psi_u^2\right)^2 \right] \mathbb{E}\left[\left(\Delta W_s^j\right)^2 \left(\Delta W_u^j\right)^2 \right] du\ ds \Bigg).
\end{aligned}$$

Since for all $u, s \in [0, t]$ such that $u \leq s$ and all $k, l \in \{1, \ldots, d\}$,

$$\mathbb{E}\left[\left(\Delta W_s^k\right)^2 \left(\Delta W_u^l\right)^2 \right] = O\left(\frac{1}{N^2}\right),$$

we get

$$\left\|\left\langle Y^{j,m,N},Y^{j,m,N}\right\rangle_t\right\|_2^2=2N^2\int_0^t\int_0^{\hat{\tau}_s}E\left(u\right)\left(u-\hat{\tau}_u\right)\left(s-\hat{\tau}_s\right)du\;ds+O\left(\frac{1}{N}\right),$$

where

$$E\left(u\right)=\mathbb{E}\left[\left(\left(\Psi_u^1\right)^2+\left(\Psi_u^2\right)^2\right)\left(\left(\Psi_s^1\right)^2+\left(\Psi_s^2\right)^2\right)\right].$$

Then, using $\left(\Psi^1\right)^2+\left(\Psi^2\right)^2=1$

$$\begin{aligned}\left\|\left\langle Y^{j,m,N},Y^{j,m,N}\right\rangle_t\right\|_2^2&=2N^2\int_0^t\int_0^{\hat{\tau}_s}\left(u-\hat{\tau}_u\right)\left(s-\hat{\tau}_s\right)du\;ds+O\left(\frac{1}{N}\right)\\&=2N^2\int_0^t\int_0^{s}\left(u-\hat{\tau}_u\right)\left(s-\hat{\tau}_s\right)du\;ds+O\left(\frac{1}{N}\right)\\&=N^2\left(\int_0^t\left(s-\hat{\tau}_s\right)ds\right)^2+O\left(\frac{1}{N}\right)\\&=\left(\frac{1}{2}Tt\right)^2+O\left(\frac{1}{N}\right)\underset{N\to+\infty}{\longrightarrow}\left(\frac{1}{2}Tt\right)^2.\end{aligned}$$

Step 3.3: conclusion of the step 3.

Applying Theorem 3.2.3 we conclude that $\sqrt{\frac{2}{T}}Y^N$ stably converges in law to a standard $d(d-1)/2$-dimensional Brownian motion B, independent of W. Now, it remains to prove the convergence in probability of K^N. We recall that for $j,m\,\{1,\ldots,d\}$, $m<j$,

$$K_t^{j,m,N}=\left[\sigma^j,\sigma^m\right]\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right).$$

From Proposition 3.3.4 and Theorem 3.3.2, together with the Lipschitz assumption on $[\sigma^j,\sigma^m]$, $j,m\,\{1,\ldots,d\}$, $m<j$ we get the following convergence in L^2

$$\sup_{t\leq T}\left\|K_t^{j,m,N}-\left[\sigma^j,\sigma^m\right]\left(X_{\hat{\tau}_t}\right)\right\|\underset{N\to+\infty}{\xrightarrow{L^2}}0. \tag{3.3.8}$$

Once again, the continuity of $[\sigma^j,\sigma^m]$, $j,m\,\{1,\ldots,d\}$, $m<j$, together with the continuity of the solution X ensure that

$$\sup_{t\leq T}\left\|\left[\sigma^j,\sigma^m\right]\left(X_{\hat{\tau}_t}\right)-\left[\sigma^j,\sigma^m\right]\left(X_t\right)\right\|\underset{N\to+\infty}{\xrightarrow{a.s}}0. \tag{3.3.9}$$

Then, combining (3.3.8) and (3.3.9), we obtain

$$\sup_{t\leq T}\left\|K_t^{j,m,N}-\left[\sigma^j,\sigma^m\right]\left(X_t\right)\right\|\underset{N\to+\infty}{\xrightarrow{\mathbb{P}}}0.$$

Thus, according to Proposition 3.2.2, we have the following convergence

$$\left(K^N,\sqrt{\frac{2}{T}}Y^N\right)\underset{N\to+\infty}{\overset{stably}{\Longrightarrow}}\left(\left(\left[\sigma^j,\sigma^m\right]\left(X\right)\right)_{j,m\in\{1,\ldots,d\},m<j},B\right).$$

The convergence of $\left\langle Y^N \right\rangle_T$ ensures its tightness. Then Proposition 3.2.4 leads us to

$$\left(K^N, \sqrt{\frac{2}{T}} Y^N, J^N\right) \underset{N \to +\infty}{\overset{stably}{\Longrightarrow}} \left(\left(\left[\sigma^j, \sigma^m\right](X)\right)_{j,m \in \{1,\dots,d\}, m<j}, B, J\right), \tag{3.3.10}$$

where, for $t \in [0, T]$,

$$J_t = \sqrt{\frac{T}{2}} \sum_{j=1}^{d} \sum_{m=1}^{j-1} \int_0^t \left[\sigma^j, \sigma^m\right](X_s)\, dB_s^{j,m}.$$

Step 4: convergence of R^N.

We show the convergence in L^2 of the remainder R^N towards 0. Applying a convexity inequality, then Doob's martingale inequality to each stochastic integral, we get

$$\mathbb{E}\left[\sup_{t \le T} \left\| R_t^N \right\|^2\right] \le N(d+1)\left(E_0 + 4 \sum_{j=1}^{d} E_j\right),$$

where

$$E_0 = \mathbb{E}\left[\sup_{t \le T} \left\| \int_0^t b\left(\hat{X}_s^{D,\eta}\right) - b\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) ds \right\|^2\right],$$

and, for $j \in \{1, \dots, d\}$,

$$E_j = \int_0^T \mathbb{E}\left[\left\| \sigma^j\left(\hat{X}_s^{D,\eta}\right) - \sigma^j\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) - \sum_{m=1}^{d} \partial\sigma^j \sigma^m \left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) \Delta W_s^m \right\|^2\right] ds.$$

Step 4.1: estimation of E_0.

Let $i \in \{1, \dots, n\}$. Using the integration by parts and Itô's formulae we get

$$\begin{aligned}
\int_0^t b^i\left(\hat{X}_s^{D,\eta}\right) - b^i\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) du = & \int_0^t (t \wedge \check{\tau}_s - s) \nabla b^i\left(\hat{X}_s^{D,\eta}\right) . \, b^i\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) ds \\
& + \sum_{j=1}^{d} \int_0^t (t \wedge \check{\tau}_s - s) \nabla b^i\left(\hat{X}_s^{D,\eta}\right) . \, \sigma^j\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) dW_s^j \\
& + \sum_{j=1}^{d} \int_0^t (t \wedge \check{\tau}_s - s) \nabla b^i\left(\hat{X}_s^{D,\eta}\right) . \, \partial\sigma^j \sigma^j\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) \Delta W_s^j dW_s^j \\
& + \int_0^t \sum_{\eta_s m < \eta_s j} (t \wedge \check{\tau}_s - s) \nabla b^i\left(\hat{X}_s^{D,\eta}\right) . \, \partial\sigma^j \sigma^m\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) \Delta W_s^m dW_s^j \\
& + \int_0^t \sum_{\eta_s m < \eta_s j} (t \wedge \check{\tau}_s - s) \nabla b^i\left(\hat{X}_s^{D,\eta}\right) . \, \partial\sigma^j \sigma^m\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) \Delta W_s^j dW_s^m \\
& + \frac{1}{2} \int_0^t (t \wedge \check{\tau}_s - s)\, Tr\left(H_s^* H_s \nabla^2 b^i\left(\hat{X}_s^{D,\eta}\right)\right) ds,
\end{aligned}$$

where $(H_s)_{0 \le s \le T}$ is a $n \times d$-dimensional process built with the following columns

$$\begin{aligned}
H_s^j = & \sigma^j\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) + \partial\sigma^j \sigma^j\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) \Delta W_s^j + \sum_{\eta_s m < \eta_s j} \partial\sigma^j \sigma^m\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) \Delta W_s^m \\
& + \sum_{\eta_s m > \eta_s j} \partial\sigma^m \sigma^j\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) \Delta W_s^m.
\end{aligned}$$

Since the first and second order derivatives of $b = \sigma^0 + \frac{1}{2}\sum_{j=1}^{d} \partial\sigma^j\sigma^j$ have a polynomial growth and $t \wedge \check{\tau}_s - s \leq \frac{T}{N}$, (3.3.4) from Lemma (3.3.5) ensures the existence of a constant $\beta \in \mathbb{R}_+^*$ independent of N such that

$$\mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t b^i\left(\hat{X}_s^{D,\eta}\right) - b^i\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) ds\right|^2\right] \leq \frac{\beta}{N^2}.$$

Step 4.2: estimation of $E_j, j \in \{1,\ldots,d\}$.

Let $i \in \{1,\ldots,n\}$. Denoting

$$\Phi_s^{ij} = \sigma^{ij}\left(\hat{X}_s^{D,\eta}\right) - \sigma^{ij}\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) - \sum_{m=1}^{d} \nabla\sigma^{ij}\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right).\,\sigma^m\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)\Delta W_s^m,$$

and applying the mean value theorem, we get a ζ_s^{ij} between $\hat{X}_s^{D,\eta}$ and $\hat{X}_{\hat{\tau}_s}^{D,\eta}$, such that:

$$\begin{aligned}\Phi_s^{ij} &= \sum_{m=1}^{d}\left(\nabla\sigma^{ij}\left(\zeta_s^{ij}\right) - \nabla\sigma^{ij}\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)\right).\,\sigma^m\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)\Delta W_s^m + \nabla\sigma^{ij}\left(\zeta_s^{ij}\right).\int_{\hat{\tau}_s}^{s} b\left(\hat{X}_{\hat{\tau}_u}^{D,\eta}\right) du \\ &+ \nabla\sigma^{ij}\left(\zeta_s^{ij}\right).\left(\sum_{\eta_s k<\eta_s m}\int_{\hat{\tau}_s}^{s}\partial\sigma^m\sigma^k\left(\hat{X}_{\hat{\tau}_u}^{D,\eta}\right)\Delta W_u^k dW_u^m\right. \\ &\left.+ \sum_{\eta_s k<\eta_s m}\int_{\hat{\tau}_s}^{s}\partial\sigma^m\sigma^k\left(\hat{X}_{\hat{\tau}_u}^{D,\eta}\right)\Delta W_u^m dW_u^k + \sum_{m=1}^{d}\int_{\hat{\tau}_s}^{s}\partial\sigma^m\sigma^m\left(\hat{X}_{\hat{\tau}_u}^{D,\eta}\right)\Delta W_u^m dW_u^m\right).\end{aligned}$$

Since $\partial\sigma^j$, for $j \in \{1,\ldots,d\}$, is locally Lipschitz with polynomially growing Lipschitz constant, there exist $c \in \mathbb{R}_+^*$ and $q \in \mathbb{N}$ such that for all $j \in \{1,\ldots,d\}$

$$\mathbb{E}\left[\left\|\nabla\sigma^{ij}\left(\zeta_s^{ij}\right)\right\|^2\right] \leq c\,\mathbb{E}\left[\left(1+\left\|\zeta_s^{ij}\right\|^q\right)\left\|\zeta_s^{ij}\right\|^2\right],$$

and

$$\mathbb{E}\left[\left\|\nabla\sigma^{ij}\left(\zeta_s^{ij}\right) - \nabla\sigma^{ij}\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)\right\|^2\right] \leq \mathbb{E}\left[\left(1+\left\|\zeta_s^{ij}\right\|^q \vee \left\|\hat{X}_{\hat{\tau}_s}^{D,\eta}\right\|^q\right)\left\|\zeta_s^{ij} - \hat{X}_{\hat{\tau}_s}^{D,\eta}\right\|^2\right].$$

Moreover b, σ^j and $\partial\sigma^j\sigma^m$, for $j, m \in \{1,\ldots,d\}$, are Lipschitz continuous. Then, combining a convexity inequality and the Cauchy-Schwarz inequality (3.3.4) and (3.3.5) from Lemma 3.3.5, we get a constant $\gamma \in \mathbb{R}_+$ independent of N such that

$$\mathbb{E}\left[\left|\Phi_s^{ij}\right|^2\right] \leq \frac{\gamma}{N^2}.$$

Then, it follows that

$$\int_0^T \mathbb{E}\left[\left|\sigma^{ij}\left(\hat{X}_s^{D,\eta}\right) - \sigma^{ij}\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right) - \sum_{m=1}^{d}\nabla\sigma^{ij}\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right).\,\sigma^m\left(\hat{X}_{\hat{\tau}_s}^{D,\eta}\right)\Delta W_s^m\right|^2\right] ds \leq \frac{\gamma T}{N^2}.$$

Step 4.3: conclusion of the step 4.

Taking the Euclidean norm, we conclude that

$$\mathbb{E}\left[\sup_{t\leq T}\left\|R_t^N\right\|^2\right] \underset{N\to+\infty}{\longrightarrow} 0. \tag{3.3.11}$$

Step 5: stable convergence in law of U^N.

We recall that $U_t^N = R_t^N + J_t^N + \left(\int_0^t \partial b_s^N U_s^N ds + \sum_{j=1}^d \int_0^t \partial\sigma_s^{j,N} U_s^N dW_s^j\right)$. Thanks to (3.3.10) and (3.3.11), we conclude using Theorem 3.2.5 since the continuity of ∂b and $\partial\sigma^j, j \in \{1,\ldots,d\}$, together with Proposition 3.3.4 ands Theorem 3.3.2, ensure that

$$\sup_{t\leq T}\left\|\partial b_t^N - \partial b\left(X_t\right)\right\| \underset{N\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0,$$

and for all $j \in \{1,\ldots,d\}$,

$$\sup_{s\leq T}\left\|\partial\sigma_t^{j,N} - \partial\sigma^j\left(X_t\right)\right\| \underset{N\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0.$$

∎

3.4 Particular case: the Brownian vector fields commute

In this section, we assume the following commutativity condition

$$\text{for all } j,m \in \{1,\ldots,d\}, \left[\sigma^j,\sigma^m\right] = \partial\sigma^m\sigma^j - \partial\sigma^j\sigma^m = 0. \tag{$\mathcal{C}$}$$

The commutativity of the Brownian vector fields implies the commutativity of the associated flows. Then, the order of integration of these fields no longer matters and η is useless. We also assume the following regularity assumptions

- for all $j \in \{1,\ldots,d\}, \sigma^j \in \mathcal{C}^1\left(\mathbb{R}^n,\mathbb{R}^n\right)$ with bounded first order derivatives,
- $\sigma^0 \in \mathcal{C}^2\left(\mathbb{R}^n,\mathbb{R}^n\right)$ with bounded first order derivatives and polynomially growing second order derivatives,
- $\sum\limits_{j=1}^d \partial\sigma^j\sigma^j$ is a Lipschitz continuous function,

Notice that $b = \sigma^0 - \frac{1}{2}\sum\limits_{j=1}^d \partial\sigma^j\sigma^j$ is also Lipschitz continuous. We denote by $L \in \mathbb{R}_+^*$ a common Lipschitz constant of σ^j, for $j \in \{0,\ldots,d\}$, b and $\sum\limits_{j=1}^d \partial\sigma^j\sigma^j$. When the vector fields corresponding to each Brownian coordinate in the SDE commute, the solution of (3.3.2) is zero. This suggests that the rate of convergence is greater than $1/2$ in this case. In fact, we prove strong convergence with order 1.

3.4.1 Interpolated scheme and strong convergence

In the commutative case we can define a smart interpolation. We define three intermediate processes. For $t \in (t_k, t_{k+1}]$:

$$\bar{X}_t^0 = \exp\left(\frac{\Delta t}{2}\sigma^0\right) X_{t_k}^{NV},$$

$$\bar{X}_t = \exp\left(\Delta W_t^d \sigma^d\right) \dots \exp\left(\Delta W_t^1 \sigma^1\right) \bar{X}_{t_{k+1}}^0,$$

$$\bar{X}_t^{d+1} = \exp\left(\frac{\Delta t}{2}\sigma^0\right) \bar{X}_{t_{k+1}},$$

$$X_{t_{k+1}}^{NV} = \bar{X}_{t_{k+1}}^{d+1}.$$

Proposition 3.4.1 *Let $t \in (t_k, t_{k+1}]$. The dynamics of $\left(\bar{X}_t\right)_{t_k < t \le t_{k+1}}$ is given by*

$$\bar{X}_t = \bar{X}_{t_{k+1}}^0 + \sum_{j=1}^d \int_{t_k}^t \sigma^j\left(\bar{X}_s\right) \circ dW_s^j.$$

Proof : Frobenius' theorem (see [18] or [17]) ensures that there exists a unique function $h \in \mathcal{C}^{1,2}\left(\mathbb{R}^n \times \mathbb{R}^d, \mathbb{R}^n\right)$ such that

$$\begin{cases} \partial_y h(x,y) = \sigma\left((h(x,y)\right) & \text{for all } (x,y) \in \mathbb{R}^n \times \mathbb{R}^d \\ \quad h(x,0) = x & \text{for all } x \in \mathbb{R}^n, \end{cases}$$

i.e. for all $j \in \{1, \dots, d\}$,

$$\partial_{y_j} h(x,y) = \sigma^j\left((h(x,y)\right) \quad \text{for all } (x,y) \in \mathbb{R}^n \times \mathbb{R}^d.$$

Then, it is clear that, by induction on $j \in \{1, \dots, d\}$

$$\exp\left(\Delta W_t^j \sigma^j\right) \dots \exp\left(\Delta W_t^1 \sigma^1\right) \bar{X}_{t_{k+1}}^0 = h\left(\bar{X}_{t_{k+1}}^0, \Delta W_t^1, \dots, \Delta W_t^j, 0, \dots, 0\right),$$

and

$$\bar{X}_t = h\left(\bar{X}_{t_{k+1}}^0, \Delta W_t\right).$$

Finally, we obtain the desired result by applying the chain rule for the Stratonovich integral. ∎

The interpolated scheme is defined as follows

$$X_t^{NV} = x + \frac{1}{2}\int_0^t \sigma^0\left(\bar{X}_s^0\right) ds + \sum_{j=1}^d \int_0^t \sigma^j\left(\bar{X}_s\right) \circ dW_s^j + \frac{1}{2}\int_0^t \sigma^0\left(\bar{X}_s^{d+1}\right) ds.$$

Theorem 3.4.2 *Let $p \in [1, +\infty)$. Under the commutativity assumption, and the regularity assumptions made in the beginning of the section, there exists a constant $C'_{NV} \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$,*

$$\mathbb{E}\left[\sup_{t \le T} \left\| X_t - X_t^{NV} \right\|^{2p}\right] \le C'_{NV} h^{2p}.$$

In Chapter 4, we prove that the normalized error process $N\left(X - X^{NV}\right)$ converges to an affine SDE with source terms involving the Lie brackets between the Brownian vector fields and the drift vector field. More precisely, under the commutativity condition ($\mathcal{C}$) and the following assumptions

- for all $j \in \{0, \ldots, d\}, \sigma^j \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first order derivatives, $\partial\sigma^j \in LIP^{pgc}_{loc}(\mathbb{R}^n \otimes \mathbb{R}^n)$ and $\partial^2\sigma^j \in LIP^{pgc}_{loc}(\mathbb{R}^n \otimes \mathbb{R}^n \otimes \mathbb{R}^n)$,
- for all $j \in \{1, \ldots, d\}, \partial\sigma^j\sigma^j$ is a Lipschitz continuous function,

we show that

$$N\left(X - X^{NV}\right) \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} U,$$

where U is the unique solution of the following affine equation

$$U_t = \frac{T}{2\sqrt{3}}\sum_{j=1}^{d}\int_0^t \left[\sigma^0, \sigma^j\right](X_s)\, d\tilde{B}^j_s + \int_0^t \partial b\,(X_s)\, U_s ds + \sum_{j=1}^{d}\int_0^t \partial\sigma^j(X_s)\, U_s dW^j_s,$$

and $\left(\tilde{B}_t\right)_{0\le t\le T}$ is a standard d-dimensional Brownian motion independent of W. This result ensures that the strong convergence rate is actually 1 when the Brownian vector fields commute, but at least one of them does not commute with the drift vector field. It is not surprising that the limit vanishes when all the vector fields σ^j, for $j \in \{0, \ldots, d\}$, commute, since the Ninomiya-Victoir scheme solves the SDE (3.1.1) in this case.

In order to prove Theorem 3.4.2, we first need to prove that the Ninomiya-Victoir scheme has uniformly bounded moments under the assumptions made in the beginning of this section. This is the aim of the following subsection.

3.4.2 Intermediate results

The following proposition will be used to prove the theorem.

Proposition 3.4.3 *Let $p \ge 1$, $Z = (Z_t)_{0\le t\le h}$ and $Y = (Y_t)_{0\le t\le h}$ be the solutions of the following $n-$dimensional SDEs, driven by a $d-$dimensional brownian motion, on the time interval $[0, h]$:*

$$\begin{cases} dZ_s = \alpha(Z_s)ds + \beta(Z_s)dW_s \\ Z_0 \text{ independent of } (W_t)_{t\in[0,h]} \text{ such that: } \mathbb{E}\left[\|Z_0\|^{2p}\right] < +\infty, \end{cases}$$

$$\begin{cases} dY_s = \gamma(Y_s)ds + \beta(Y_s)dW_s \\ Y_0 \text{ independent of } (W_t)_{t\in[0,h]} \text{ such that: } \mathbb{E}\left[\|Y_0\|^{2p}\right] < +\infty, \end{cases}$$

respectively. Assume that α, β and γ are Lipschitz continuous functions. Then, there exists a constant $C_0 \in \mathbb{R}^*_+$ such that for all $t, s \in [0, h], s \le t$,

$$\mathbb{E}\left[1 + \|Z_t\|^{2p}\right] \le \mathbb{E}\left[1 + \|Z_0\|^{2p}\right]\exp\left(C_0 h\right), \tag{3.4.1}$$

$$\begin{aligned}\mathbb{E}\left[\sup_{t\le h}\|Z_t - Y_t\|^{2p}\right] \le C_0\Bigg(&\mathbb{E}\left[\|Z_0 - Y_0\|^{2p}\right] + \mathbb{E}\left[1 + \|Y_0\|^{2p}\right]h^{2p} \\ &+ \mathbb{E}\left[1 + \|Z_0\|^{2p}\right]h^{2p}\Bigg)\exp\left(C_0 h\right),\end{aligned} \tag{3.4.2}$$

if $\alpha = \gamma$, we have

$$\mathbb{E}\left[\|Z_t - Y_t\|^{2p}\right] \leq \mathbb{E}\left[\|Z_0 - Y_0\|^{2p}\right] \exp\left(C_0 h\right), \tag{3.4.3}$$

$$\mathbb{E}\left[\|Z_t - Z_s\|^{2p}\right] \leq C_0\left(1 + \mathbb{E}\left[\|Z_0\|^{2p}\right]\right)(t-s)^p, \tag{3.4.4}$$

and if $\beta = 0$, we have a better estimation

$$\mathbb{E}\left[\|Z_t - Z_s\|^{2p}\right] \leq C_0\left(1 + \mathbb{E}\left[\|Z_0\|^{2p}\right]\right)(t-s)^{2p}. \tag{3.4.5}$$

The constant C_0 only depends on $\|\alpha(0)\|, \|\beta(0)\|, \|\gamma(0)\|$, T, p, and the Lipschitz constants of α, β and γ.

Proof : Only (3.4.2) requires a proof, the other results are well known (see [51]). Let $t \in [0, h]$, and $s \in [0, t]$. Applying a convexity inequality, we get

$$\begin{aligned}\|Z_s - Y_s\|^{2p} = 3^{2p-1}\Bigg(&\|Z_0 - Y_0\|^{2p} + s^{2p-1}\int_0^s \|\alpha(Z_u) - \gamma(Y_u)\|^{2p} du \\ &+ \left\|\int_0^s \beta(Z_u) - \beta(Y_u)\, dW_u\right\|^{2p}\Bigg).\end{aligned}$$

Taking the expectation of the supremum, and using the Burkholder-Davis-Gundy inequality, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{s\leq t}\|Z_s - Y_s\|^{2p}\right] \leq 3^{2p-1}\Bigg(&\mathbb{E}\left[\|Z_0 - Y_0\|^{2p}\right] + (2h)^{2p-1}\int_0^t \mathbb{E}\left[\|\alpha(Z_u)\|^{2p}\right] + \mathbb{E}\left[\|\gamma(Y_u)\|^{2p}\right] du \\ &+ Kh^{p-1}\int_0^t \mathbb{E}\left[\|\beta(Z_u) - \beta(Y_u)\|^{2p}\right] du\Bigg),\end{aligned}$$

where K is the constant that appears in the Burkholder-Davis-Gundy inequality. By the Lipschitz assumption, and by using (3.4.1), we obtain

$$\begin{aligned}\mathbb{E}\left[\sup_{s\leq t}\|Z_s - Y_s\|^{2p}\right] \leq 3^{2p-1}\Bigg(&\mathbb{E}\left[\|Z_0 - Y_0\|^{2p}\right] + R\left(\mathbb{E}\left[1 + \|Z_0\|^{2p}\right] + \mathbb{E}\left[1 + \|Y_0\|^{2p}\right]\right) h^{2p} \\ &+ KT^{p-1}L^{2p}\int_0^t \mathbb{E}\left[\sup_{v\leq u}\|Z_v - Y_v\|^{2p}\right] du\Bigg),\end{aligned}$$

where

$$R = 4^{2p-1}\left(\max\left\{L, \max\left\{\left|\alpha^i(0)\right|, i \in \{1,\ldots,n\}\right\}, \max\left\{\left|\gamma^i(0)\right|, i \in \{1,\ldots,n\}\right\}\right\}\right)^{2p} \exp\left(C_0 T\right)$$

and L is the common Lipschitz constant of α, β, and γ. We conclude using Gronwall's lemma and changing C_0. ∎

The next results are similar to Lemmas 2.5 and 2.6 in Chapter 2. However, they require slightly different regularity assumptions since the Brownian vector fields commute.

Lemma 3.4.4 *Assume that*

- *for all* $j \in \{0, \ldots, d\}, \sigma^j \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ *with bounded first derivatives,*
- $\sum_{j=1}^{d} \partial\sigma^j\sigma^j$ *is a Lipschitz continuous function.*

Then, for all $p \geq 1$, *there exists a constant* $C_1' \in \mathbb{R}_+^*$ *such that for all* $N \in \mathbb{N}^*, t \in [0, T]$,

$$\mathbb{E}\left[1 + \left\|\bar{X}_t^0\right\|^{2p}\right] \leq \exp(C_1'\check{\tau}_t)\left(1 + \|x\|^{2p}\right),$$

$$\mathbb{E}\left[1 + \left\|\bar{X}_t\right\|^{2p}\right] \leq \exp(C_1'\check{\tau}_t)\left(1 + \|x\|^{2p}\right),$$

$$\mathbb{E}\left[1 + \left\|\bar{X}_t^{d+1}\right\|^{2p}\right] \leq \exp(C_1'\check{\tau}_t)\left(1 + \|x\|^{2p}\right).$$

Proof : Let $p \geq 1$ and $t \in [0, T]$, then there exists an integer $k \in \{0, \ldots, N-1\}$ such that $t_k < t \leq t_{k+1}$. We recall that $\left(\bar{X}_s^0\right)_{t_k < s \leq t_{k+1}}$ is the solution of the following ODE:

$$\begin{cases} dZ_s = \dfrac{1}{2}\sigma^0(Z_s)ds \\ Z_{t_k} = X_{t_k}^{NV}, \end{cases}$$

$\left(\bar{X}_s\right)_{t_k < s \leq t_{k+1}}$ is the solution of the following SDE:

$$\begin{cases} dZ_s = \dfrac{1}{2}\sum_{j=1}^{d} \partial\sigma^j\sigma^j\left(Z_s\right)ds + \sum_{j=1}^{d}\sigma^j(Z_s)dW_s^j \\ Z_{t_k} = \bar{X}_{t_{k+1}}^0, \end{cases}$$

and $\left(\bar{X}_s^{d+1}\right)_{t_k < s \leq t_{k+1}}$ is the solution of the following ODE:

$$\begin{cases} dZ_s = \dfrac{1}{2}\sigma^0(Z_s)ds \\ Z_{t_k} = \bar{X}_{t_{k+1}}. \end{cases}$$

Applying (3.4.1) from the Proposition 3.4.3, we get, for all $t \in (t_k, t_{k+1}]$,

$$\mathbb{E}\left[1 + \left\|\bar{X}_t^0\right\|^{2p}\right] \leq \mathbb{E}\left[1 + \left\|X_{t_k}^{NV,\eta}\right\|^{2p}\right]\exp\left(C_0 h\right), \tag{3.4.6}$$

$$\mathbb{E}\left[1 + \left\|\bar{X}_t\right\|^{2p}\right] \leq \mathbb{E}\left[1 + \left\|\bar{X}_{t_{k+1}}^{0,\eta}\right\|^{2p}\right]\exp\left(C_0 h\right), \tag{3.4.7}$$

$$\mathbb{E}\left[1 + \left\|\bar{X}_t^{d+1}\right\|^{2p}\right] \leq \mathbb{E}\left[1 + \left\|\bar{X}_{t_{k+1}}\right\|^{2p}\right]\exp\left(C_0 h\right). \tag{3.4.8}$$

Using backward induction on (3.4.6), (3.4.7) and (3.4.8) we get

$$\mathbb{E}\left[1 + \left\|\bar{X}_t^j\right\|^{2p}\right] \leq \exp\left(3C_0 t_{k+1}\right)\left(1 + \|x\|^{2p}\right),$$

for $j \in \{0, d+1\}$, and

$$\mathbb{E}\left[1+\left\|\bar{X}_t\right\|^{2p}\right] \leq \exp\left(3C_0 t_{k+1}\right)\left(1+\|x\|^{2p}\right).$$

∎

The proof of the next Lemma is a straightforward consequence of Lemma 3.4.4, (3.4.4) and (3.4.5) from Proposition 3.4.3.

Lemma 3.4.5 *Under the assumptions of the previous Lemma we have the following result. For all $p \geq 1$, there exists a constant $C_2' \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*, t \in [0,T]$,*

$$\mathbb{E}\left[\left\|\bar{X}_t^0 - X_{\hat{\tau}_t}^{NV}\right\|^{2p}\right] \leq C_2'\left(1+\|x\|^{2p}\right)h^{2p}, \tag{3.4.9}$$

$$\mathbb{E}\left[\left\|\bar{X}_t - \bar{X}_{\check{\tau}_t}^0\right\|^{2p}\right] \leq C_2'\left(1+\|x\|^{2p}\right)h^{p}, \tag{3.4.10}$$

$$\mathbb{E}\left[\left\|\bar{X}_t^{d+1} - \bar{X}_{\check{\tau}_t}\right\|^{2p}\right] \leq C_2'\left(1+\|x\|^{2p}\right)h^{2p}. \tag{3.4.11}$$

3.4.3 Proof of the strong convergence in the commutative case

Proof : The error process is given by

$$X_s - X_s^{NV} = \frac{1}{2}\int_0^s \sigma^0\left(X_u\right) - \sigma^0\left(\bar{X}_u^0\right)du + \frac{1}{2}\int_0^s \sigma^0\left(X_u\right) - \sigma^0\left(\bar{X}_u^{d+1}\right)du + \sum_{j=1}^{d}\int_0^s\left(\sigma^j\left(X_u\right) - \sigma^j\left(\bar{X}_u\right)\right)\circ dW_u^j.$$

Then, using a convexity inequality, we get for $t \in [0,T]$ and $p \geq 1$

$$\mathbb{E}\left[\sup_{s\leq t}\left\|X_t - X_t^{NV}\right\|^{2p}\right] \leq 3^{2p-1}\left(E_0 + E_{d+1} + E\right), \tag{3.4.12}$$

where

$$E_0 = \mathbb{E}\left[\sup_{s\leq t}\left\|\frac{1}{2}\int_0^s \sigma^0\left(X_u\right) - \sigma^0\left(\bar{X}_u^0\right)du\right\|^{2p}\right],$$

$$E_{d+1} = \mathbb{E}\left[\sup_{s\leq t}\left\|\frac{1}{2}\int_0^s \sigma^0\left(X_u\right) - \sigma^0\left(\bar{X}_u^{d+1}\right)du\right\|^{2p}\right],$$

and

$$E = \mathbb{E}\left[\sup_{s\leq t}\left\|\sum_{j=1}^{d}\int_0^s\left(\sigma^j\left(X_u\right) - \sigma^j\left(\bar{X}_u\right)\right)\circ dW_u^j\right\|^{2p}\right].$$

For the reader's convenience, the proof of this theorem is split into intermediate steps.

Step 1: estimation of E_0.

Introducing $\sigma^0\left(X_{\hat{\tau}_u}\right)$ and $\sigma^0\left(X_{\hat{\tau}_u}^{NV}\right)$, and using convexity inequality, we obtain

$$\begin{aligned} E_0 \leq \frac{3^{2p-1}}{2^{2p}} \Bigg(& \mathbb{E}\left[\sup_{s\leq t}\left\|\int_0^s \sigma^0\left(X_u\right) - \sigma^0\left(X_{\hat{\tau}_u}\right) du\right\|^{2p}\right] \\ & + T^{2p-1}\int_0^t \mathbb{E}\left[\left\|\sigma^0\left(X_{\hat{\tau}_u}\right) - \sigma^0\left(X^{NV}_{\hat{\tau}_u}\right)\right\|^{2p}\right] du \\ & + T^{2p-1}\int_0^t \mathbb{E}\left[\left\|\sigma^0\left(X^{NV}_{\hat{\tau}_u}\right) - \sigma^0\left(\bar{X}^0_u\right)\right\|^{2p}\right] du\Bigg). \end{aligned}$$

The two last integrals are easy to estimate. On the one hand, since σ^0 is Lipschitz,

$$\int_0^t \mathbb{E}\left[\left\|\sigma^0\left(X_{\hat{\tau}_u}\right) - \sigma^0\left(X^{NV}_{\hat{\tau}_u}\right)\right\|^{2p}\right] du \leq L^{2p}\int_0^t \mathbb{E}\left[\sup_{v\leq u}\left\|X_v - X^{NV}_v\right\|^{2p}\right] dv.$$

On the other hand, using (3.4.9) from Lemma 3.4.5 together with the Lipschitz property of σ^0, we get

$$\begin{aligned} \int_0^t \mathbb{E}\left[\left\|\sigma^0\left(X^{NV}_{\hat{\tau}_u}\right) - \sigma^0\left(\bar{X}^0_u\right)\right\|^{2p}\right] du &\leq L^{2p}\int_0^t \mathbb{E}\left[\left\|\bar{X}^0_u - X^{NV}_{\hat{\tau}_u}\right\|^{2p}\right] du \\ &\leq L^{2p}C_2'T\left(1+\|x\|^{2p}\right)h^{2p}. \end{aligned}$$

Now we look at the first integral. For $i\in\{1,\ldots,n\}$, using the integration by parts formula, we have

$$\int_0^s \sigma^{i0}\left(X_u\right) - \sigma^{i0}\left(X_{\hat{\tau}_u}\right) du = \int_0^s \left(\check{\tau}_u\wedge s - u\right) d\left(\sigma^{i0}\left(X_u\right)\right).$$

Then, applying Itô's formula, we get

$$\begin{aligned} \int_0^s \sigma^{i0}\left(X_u\right) - \sigma^{i0}\left(X_{\hat{\tau}_u}\right) du = & \int_0^s \left(\check{\tau}_u\wedge s - u\right)\nabla\sigma^{i0}\left(X_u\right).\, b\left(X_u\right) du \\ & + \frac{1}{2}\int_0^s \left(\check{\tau}_u\wedge s - u\right) tr\left(\sigma\left(X_u\right)\sigma^*\left(X_u\right)\nabla^2\sigma^{i0}\left(X_u\right)\right) du \\ & + \int_0^s \left(\check{\tau}_u\wedge s - u\right)\sigma^*\left(X_u\right)\nabla\sigma^{i0}\left(X_u\right).\, dW_u, \end{aligned}$$

where $\sigma = \left(\sigma^{ij}\right)_{1\leq i\leq n, 1\leq j\leq d}$ is the diffusion matrix. Taking the expectation of the supremum and using a convexity inequality, the Burkholder-Davis-Gundy inequality and $\check{\tau}_u - u \leq h$, we get

$$\begin{aligned} \mathbb{E}\left[\sup_{s\leq t}\left|\int_0^s \sigma^{i0}\left(X_u\right) - \sigma^{i0}\left(X_{\hat{\tau}_u}\right) du\right|^{2p}\right] \leq 3^{2p-1}h^{2p}\Bigg(& T^{2p-1}\int_0^t \mathbb{E}\left[\left|\nabla\sigma^{i0}\left(X_u\right).\, b\left(X_u\right)\right|^{2p}\right] du \\ & + \frac{T^{2p-1}}{2^{2p}}\int_0^t \mathbb{E}\left[\left|tr\left(\sigma\left(X_u\right)\sigma^*\left(X_u\right)\nabla^2\sigma^{i0}\left(X_u\right)\right)\right|^{2p}\right] du \\ & + T^{p-1}K\int_0^t \mathbb{E}\left[\left\|\sigma^*\left(X_u\right)\nabla\sigma^{i0}\left(X_u\right)\right\|^{2p}\right] du\Bigg), \end{aligned}$$

where K is the constant that appears in the Burkholder-Davis-Gundy inequality. By the regularity assumption on σ^j, for $j\in\{0,\ldots d\}$, we easily get two constants $\alpha_0\in\mathbb{R}_+$ and $q_1\in\mathbb{N}^*$ which only depend on p, T, σ and σ^0, such that

$$\mathbb{E}\left[\sup_{s\leq t}\left|\int_0^s \sigma^{i0}\left(X_u\right) - \sigma^{i0}\left(X_{\hat{\tau}_u}\right) du\right|^{2p}\right] \leq \alpha_0 h^{2p}\int_0^t \mathbb{E}\left[1+\left\|X_u\right\|^{2q_1}\right] du.$$

Moreover, by the Lipschitz assumption on b, σ^j, for $j \in \{0, \dots d\}$, (3.4.1) from Proposition 3.4.3 ensures that

$$\mathbb{E}\left[1 + \|X_u\|^{2q_1}\right] < +\infty.$$

Finally, by combining our different inequalities, we obtain a constant $\beta_0 \in \mathbb{R}_+^*$ independent of N, such that

$$E_0 \leq \beta_0 \left(h^{2p} + \int_0^t \mathbb{E}\left[\sup_{v \leq u} \left\| X_v - X_v^{NV} \right\|^{2p}\right] du \right). \tag{3.4.13}$$

Step 2: estimation of E_{d+1}.

Introducing $\sigma^0\left(X_{\hat{\tau}_u}\right)$ and $\sigma^0\left(X_{\hat{\tau}_u}^{NV}\right)$, and using convexity inequality, we get

$$\begin{aligned}
E_{d+1} \leq \frac{3^{2p-1}}{2^{2p}} \Bigg(& \mathbb{E}\left[\sup_{s \leq t} \left\| \int_0^s \sigma^0\left(X_u\right) - \sigma^0\left(X_{\hat{\tau}_u}\right) du \right\|^{2p}\right] \\
& + T^{2p-1} \int_0^t \mathbb{E}\left[\left\| \sigma^0\left(X_{\hat{\tau}_u}\right) - \sigma^0\left(X_{\hat{\tau}_u}^{NV}\right)\right\|^{2p}\right] du \\
& + \mathbb{E}\left[\sup_{s \leq t} \left\| \int_0^t \sigma^0\left(\bar{X}_u^{d+1}\right) - \sigma^0\left(X_{\hat{\tau}_u}^{NV}\right) du \right\|^{2p}\right] \Bigg).
\end{aligned}$$

We have already dealt with the two first terms int the right-hand side in the estimation of E_{d+1}. Let $i \in \{1, \dots, n\}$, we use integration by parts and Itô's formula to get

$$\begin{aligned}
\int_0^s \sigma^{i0}\left(\bar{X}_u^{d+1}\right) - \sigma^{i0}\left(X_{\hat{\tau}_u}^{NV}\right) du = & \int_0^s \left(s \wedge \check{\tau}_u - u\right) d\left(\sigma^{i0}\left(\bar{X}_u^{d+1}\right)\right) + \int_0^{\check{\tau}_s} \left(\check{\tau}_u - u\right) d\left(\sigma^{i0}\left(\bar{X}_u\right)\right) \\
& + \int_0^{\check{\tau}_s} \left(\check{\tau}_u - u\right) d\left(\sigma^{i0}\left(\bar{X}_u^0\right)\right) \\
= & \frac{1}{2} \int_0^s \left(\check{\tau}_u - u\right) \nabla \sigma^{i0}\left(\bar{X}_u^{d+1}\right) . \, \sigma^0\left(\bar{X}_u^{d+1}\right) du \\
& + \sum_{j=1}^d \int_0^{\check{\tau}_s} \left(\check{\tau}_u - u\right) \left(\nabla \sigma^{i0}\left(\bar{X}_u\right) . \, \sigma^j\left(\bar{X}_u\right)\right) dW_u^j \\
& + \frac{1}{2} \sum_{j=1}^d \int_0^{\check{\tau}_s} \left(\check{\tau}_u - u\right) \left(\nabla \sigma^{i0}\left(\bar{X}_u\right) . \, \partial \sigma^j \sigma^j\left(\bar{X}_u\right)\right) du \\
& + \frac{1}{2} \int_0^{\check{\tau}_s} \left(\check{\tau}_u - u\right) tr\left(\sigma \sigma^* \nabla^2 \sigma^{i0}\left(\bar{X}_u\right)\right) du \\
& + \frac{1}{2} \int_0^{\check{\tau}_s} \left(\check{\tau}_u - u\right) \nabla \sigma^{i0}\left(\bar{X}_u^0\right) . \, \sigma^0\left(\bar{X}_u^0\right) du.
\end{aligned}$$

Once again, by the regularity assumption on σ^j, for $j \in \{1, \dots, d\}$, and by using Lemma 3.4.4 and $\check{\tau}_u - u \leq h$, we get a constant $\alpha_2 \in \mathbb{R}_+$ independent of N such that

$$\mathbb{E}\left[\sup_{s \leq t} \left| \int_0^s \sigma^{i0}\left(\bar{X}_u^{d+1}\right) - \sigma^{i0}\left(X_{\hat{\tau}_u}^{NV}\right) du \right|^{2p}\right] \leq \alpha_2 h^{2p}.$$

Finally, by summing up our different inequalities, we obtain $\beta_{d+1} \in \mathbb{R}_+^*$, independent of N, such that

$$E_{d+1} \leq \beta_{d+1} \left(h^{2p} + \int_0^t \mathbb{E}\left[\sup_{v \leq u} \left\| X_v - X_v^{NV} \right\|^{2p} \right] du \right). \tag{3.4.14}$$

Step 3: estimation of E.

Re-expressing the integrals in Itô's form, we get

$$\sum_{j=1}^{d} \int_0^s \left(\sigma^j \left(X_u \right) - \sigma^j \left(\bar{X}_u \right) \right) \circ dW_u^j = \sum_{j=1}^{d} \int_0^s \left(\sigma^j \left(X_u \right) - \sigma^j \left(\bar{X}_u \right) \right) dW_u^j + \frac{1}{2} \sum_{j=1}^{d} \int_0^s \left(\partial \sigma^j \sigma^j \left(X_u \right) - \partial \sigma^j \sigma^j \left(\bar{X}_u \right) \right) du.$$

Applying a convexity inequality, the Burkholder-Davis-Gundy inequality, together with the Lipschitz property of σ^j, for $j \in \{1, \ldots, d\}$ and $\sum_{j=1}^{d} \partial \sigma^j \sigma^j$, we get

$$\mathbb{E}\left[\sup_{s \leq t} \left\| \sum_{j=1}^{d} \int_0^s \left(\sigma^j \left(X_u \right) - \sigma^j \left(\bar{X}_u \right) \right) \circ dW_u^j \right\|^{2p} \right] \leq \alpha \int_0^t \mathbb{E}\left[\left\| X_u - \bar{X}_u \right\|^{2p} \right] du.$$

where $\alpha = d\left(2L\right)^{2p} \left(KT^{p-1} + \frac{T^{2p-1}}{2^{2p}} \right)$ and K is the constant that appears in the Burkholder-Davis-Gundy inequality. To estimate $\mathbb{E}\left[\left\| X_u - \bar{X}_u \right\|^{2p} \right]$, we introduce the following piecewise continuous process $\left(\check{X}_u \right)_{0 \leq u \leq T}$ such that for $u \in [\hat{\tau}_u, \check{\tau}_u)$

$$\check{X}_u = X_{\hat{\tau}_u} + \sum_{j=1}^{d} \int_{\hat{\tau}_u}^{u} \sigma^j \left(\check{X}_v \right) \circ dW_v^j.$$

Then, using a convexity inequality, we obtain

$$\mathbb{E}\left[\sup_{s \leq t} \left\| \sum_{j=1}^{d} \int_0^s \sigma^j \left(X_u \right) - \sigma^j \left(\bar{X}_u \right) \circ dW_u^j \right\|^{2p} \right] \leq \alpha \int_0^t \mathbb{E}\left[\left\| X_u - \check{X}_u \right\|^{2p} \right] + \mathbb{E}\left[\left\| \check{X}_u - \bar{X}_u \right\|^{2p} \right] du,$$

where $\alpha' = 2^{2p-1}\alpha$. An estimation of the first expectation is given by (3.4.2) in Proposition 3.4.3:

$$\mathbb{E}\left[\left\| X_u - \check{X}_u \right\|^{2p} \right] \leq 2C_0 \exp\left(C_0 h \right) \mathbb{E}\left[1 + \left\| X_{\hat{\tau}_u} \right\|^{2p} \right] h^{2p}.$$

We estimate the second expectation using (3.4.3) in Proposition 3.4.3:

$$\mathbb{E}\left[\left\| \check{X}_u - \bar{X}_u \right\|^{2p} \right] \leq \exp\left(C_0 h \right) \mathbb{E}\left[\left\| X_{\hat{\tau}_u} - \bar{X}_{\check{\tau}_u}^{0} \right\|^{2p} \right].$$

Introducing $X_{\hat{\tau}_u}^{NV}$ in the right-hand side of the inequality, we get

$$\mathbb{E}\left[\left\| \check{X}_u - \bar{X}_u \right\|^{2p} \right] \leq 2^{2p-1} \exp\left(C_0 T \right) \left(\mathbb{E}\left[\left\| X_{\hat{\tau}_u} - X_{\hat{\tau}_u}^{NV} \right\|^{2p} \right] + \mathbb{E}\left[\left\| X_{\hat{\tau}_u}^{NV} - \bar{X}_{\check{\tau}_u}^{0} \right\|^{2p} \right] \right).$$

Applying (3.4.9) from Lemma 3.4.5 to the last expectation, we obtain

$$\mathbb{E}\left[\left\|\check{X}_u - \bar{X}_u\right\|^{2p}\right] \leq 2^{2p-1}\exp\left(C_0 T\right)\left(\mathbb{E}\left[\sup_{v\leq u}\left\|X_v - X_v^{NV}\right\|^{2p}\right] + C_1'\left(1+\|x\|^{2p}\right)h^{2p}\right).$$

We conclude by combining our different results, we get a constant $\beta \in \mathbb{R}_+^*$ independent of N, such that

$$E \leq \beta\left(h^{2p} + \int_0^t \mathbb{E}\left[\sup_{v\leq u}\left\|X_v - X_v^{NV}\right\|^{2p}\right]du\right). \tag{3.4.15}$$

Step 4: conclusion.

Combining (3.4.12), (3.4.13), (3.4.14),(3.4.15), and Gronwall's lemma, we get the following estimation

$$\mathbb{E}\left[\sup_{t\leq T}\left\|X_t - X_t^{NV}\right\|^{2p}\right] \leq 3^{2p-1}\left(\beta_0+\beta_{d+1}+\beta\right)\exp\left(3^{2p-1}\left(\beta_0+\beta_{d+1}+\beta\right)T\right)h^{2p}.$$

■

3.5 Appendix

This section is devoted to the proof of Proposition 3.3.4. To compare $\hat{X}^{D,\eta}$ with $X^{NV,\eta}$, we introduce the following non-adapted interpolation of $\left(\hat{X}^{D,\eta}_{t_k}\right)_{0\leq k\leq N}$:

$$\begin{aligned} X_t^{D,\eta} = \hat{X}^{D,\eta}_{\hat{\tau}_t} + b\left(\hat{X}^{D,\eta}_{\hat{\tau}_t}\right)\Delta t + \sum_{j=1}^d \sigma^j\left(\hat{X}^{D,\eta}_{\hat{\tau}_t}\right)\Delta W_t^j + \frac{1}{2}\sum_{j=1}^d \partial\sigma^j\sigma^j\left(\hat{X}^{D,\eta}_{\hat{\tau}_t}\right)\left(\left(\Delta W_t^j\right)^2 - \Delta t\right) \\ + \sum_{\eta_t m<\eta_t j}\partial\sigma^j\sigma^m\left(\hat{X}^{D,\eta}_{\hat{\tau}_t}\right)\Delta W^m_{\check{\tau}_t}\Delta W_t^j. \end{aligned} \tag{3.5.1}$$

Proposition 3.3.4 is a consequence of the next lemma, which compares $\hat{X}^{D,\eta}$ and $X^{D,\eta}$, and the next proposition, which compares $X^{D,\eta}$ and $X^{NV,\eta}$.

Lemma 3.5.1 *Let $p \geq 1$ and $\epsilon > 0$. Under the assumptions of Lemma 3.3.5, there exists a constant $C_D' \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$,*

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\hat{X}_t^{D,\eta} - X_t^{D,\eta}\right\|^{2p}\right] \leq C_D'\frac{1}{N^{2p-\epsilon}}.$$

Proof : Let $p \geq 1, q, r > 1, \dfrac{1}{q}+\dfrac{1}{r} = 1$, and $t \in [0,T]$. Subtracting (3.5.1) from (3.3.3), we obtain

$$\begin{aligned} \hat{X}_t^{D,\eta} - X_t^{D,\eta} &= \sum_{\eta_t m<\eta_t j}\partial\sigma^j\sigma^m\left(\hat{X}^{D,\eta}_{\hat{\tau}_t}\right)\Delta W_t^j\left(\Delta W_t^m - \Delta W^m_{\check{\tau}_t}\right) \\ &= -\sum_{\eta_t m<\eta_t j}\partial\sigma^j\sigma^m\left(\hat{X}^{D,\eta}_{\hat{\tau}_t}\right)\Delta W_t^j\left(W^m_{\check{\tau}_t} - W_t^m\right). \end{aligned}$$

Then, combining a convexity inequality and the Hölder inequality, we get

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\hat{X}_t^{D,\eta}-X_t^{D,\eta}\right\|^{2p}\right]\leq\left(\frac{d^2-d}{2}\right)^{2p-1}\sum_{m<j}\Bigg(\mathbb{E}^{\frac{1}{q}}\left[\sup_{t\leq T}\left\|\partial\sigma^j\sigma^m\left(\hat{X}_{\hat{\tau}_t}^{D,\eta}\right)\right\|^{2pq}\right]$$
$$\mathbb{E}^{\frac{1}{r}}\left[\sup_{t\leq T}\left|\Delta W_t^j\left(W_{\check{\tau}_t}^m-W_t^m\right)\right|^{2pr}\right]\Bigg).$$

Since $\partial\sigma^j\sigma^m$ for $j,m\in\{1,\dots,d\}$, has an affine growth, using Lemma 3.3.5 we obtain a constant β, independent of N, such that

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\hat{X}_t^{D,\eta}-X_t^{D,\eta}\right\|^{2p}\right]\leq\beta\sum_{m<j}\mathbb{E}^{\frac{1}{r}}\left[\sup_{t\leq T}\left|\Delta W_t^j\left(W_{\check{\tau}_t}^m-W_t^m\right)\right|^{2pr}\right].$$

Then a straightforward calculation gives us

$$\begin{aligned}\mathbb{E}\left[\sup_{t\leq T}\left\|\hat{X}_t^{D,\eta}-X_t^{D,\eta}\right\|^{2p}\right]&\leq\beta\sum_{m<j}\mathbb{E}^{\frac{1}{r}}\left[\sup_{k\in\{1,\dots,N\}}\sup_{t_k<t\leq t_{k+1}}\left|\Delta W_t^j\left(W_{\check{\tau}_t}^m-W_t^m\right)\right|^{2pr}\right]\\&\leq\beta\sum_{m<j}\mathbb{E}^{\frac{1}{r}}\left[\sum_{k=1}^{N}\sup_{t_k<t\leq t_{k+1}}\left|\left(W_t^j-W_{t_k}^j\right)\left(W_{t_{k+1}}^m-W_t^m\right)\right|^{2pr}\right]\\&\leq\beta N^{\frac{1}{r}}\sum_{m<j}\mathbb{E}^{\frac{1}{r}}\left[\sup_{0<t\leq t_1}\left|W_t^j\right|^{2pr}\sup_{0<t\leq t_1}\left|W_{t_1}^m-W_t^m\right|^{2pr}\right]\\&\leq\beta N^{\frac{1}{r}}\sum_{m<j}\mathbb{E}^{\frac{2}{r}}\left[\sup_{0<t\leq t_1}\left|W_t^j\right|^{2pr}\right].\end{aligned}$$

Using Doob's submartingale inequality, we get

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\hat{X}_t^{D,\eta}-X_t^{D,\eta}\right\|^{2p}\right]\leq\gamma N^{\frac{1}{r}}\sum_{m<j}\mathbb{E}^{\frac{2}{r}}\left[\left|W_{t_1}^j\right|^{2pr}\right],$$

where $\gamma=\beta\left(\frac{r}{r-1}\right)^2$. Finally, we obtain

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\hat{X}_t^{D,\eta}-X_t^{D,\eta}\right\|^{2p}\right]\leq C'_D\frac{1}{N^{2p-\frac{1}{r}}},$$

where $C'_D=\dfrac{d^2-d}{2}\gamma T^{2p}\,\mathbb{E}^{\frac{2}{r}}\left[|G|^{2pr}\right]$ and G a normal random variable. ∎

Proposition 3.5.2 *Let $p\geq 1$. Under the assumptions of Theorem 3.3.3, there exists a constant $C''_D\in\mathbb{R}_+^*$ such that for all $N\in\mathbb{N}^*$,*

$$\mathbb{E}\left[\sup_{t\leq T}\left\|X_t^{NV,\eta}-X_t^{D,\eta}\right\|^{2p}\right]\leq C''_D h^{2p}.$$

Before proving this proposition, we recall some useful results stated and proved in Chapter 2.

Lemma 3.5.3 *Assume that*

- *for all* $j \in \{1, \ldots, d\}, \sigma^j \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$.
- σ^0, σ^j *and* $\partial\sigma^j\sigma^j$, *for all* $j \in \{1, \ldots, d\}$, *are Lipschitz continuous functions.*
- $F \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ *with polynomially growing first and second order derivatives.*

Then, for all $p \geq 1$ *there exists a constant* $C_1 \in \mathbb{R}_+^*$ *such that for all* $N \in \mathbb{N}^*$,

$$\text{for all } t \in [0, T], \mathbb{E}\left[1 + \left\|X_t^{NV,\eta}\right\|^{2p}\right] \leq C_1\left(1 + \|x\|^{2p}\right), \tag{3.5.2}$$

$$\text{for all } t \in [0, T], j \in \{0, \ldots, d+1\}, \mathbb{E}\left[1 + \left\|\bar{X}_t^{j,\eta}\right\|^{2p}\right] \leq C_1\left(1 + \|x\|^{2p}\right), \tag{3.5.3}$$

$$\text{for all } t \in [0, T], j \in \{0, \ldots, d+1\}, \mathbb{E}\left[\left\|\bar{X}_t^{j,\eta} - X_{\hat{\tau}_t}^{NV,\eta}\right\|^{2p}\right] \leq C_1\left(1 + \|x\|^{2p}\right)h^p, \tag{3.5.4}$$

$$\text{for all } j \in \{0, \ldots, d+1\}, \mathbb{E}\left[\sup_{t\leq T}\left\|\int_0^t F\left(\bar{X}_s^{j,\eta}\right) - F\left(X_{\hat{\tau}_s}^{NV,\eta}\right)ds\right\|^{2p}\right] \leq C_1 h^{2p}. \tag{3.5.5}$$

Now, we turn to the proof of Proposition 3.5.2.

Proof of Proposition 3.5.2 : We denote by $L \in \mathbb{R}_+^*$ a common Lipschitz constant of σ^j for $j \in \{0, \ldots, d\}$ and $\partial\sigma^j\sigma^m$ for $j, m \in \{1, \ldots, d\}$. Let $t \in [0, T]$ and $s \in [0, t]$. Rewriting (3.5.1) in integral form, we get

$$\begin{aligned} X_t^{D,\eta} = x &+ \int_0^t \sigma^0\left(X_{\hat{\tau}_s}^{D,\eta}\right)ds + \frac{1}{2}\sum_{j=1}^d \int_0^t \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_s}^{D,\eta}\right)ds + \sum_{j=1}^d \int_0^t \sigma^j\left(X_{\hat{\tau}_s}^{D,\eta}\right)dW_s^j \\ &+ \sum_{j=1}^d \int_0^t \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_s}^{D,\eta}\right)\Delta W_s^j dW_s^j + \int_0^t \sum_{\eta_s m < \eta_s j} \partial\sigma^j\sigma^m\left(X_{\hat{\tau}_s}^{D,\eta}\right)\Delta W_{\check{\tau}_s}^m dW_s^j. \end{aligned} \tag{3.5.6}$$

Subtracting (3.5.6) from (3.3.1), we get

$$\begin{aligned} X_s^{NV,\eta} - X_s^{D,\eta} = &\frac{1}{2}\int_0^s \sigma^0\left(\bar{X}_u^{0,\eta}\right) - \sigma^0\left(X_{\hat{\tau}_u}^{D,\eta}\right)du + \sum_{j=1}^d \int_0^s \partial\sigma^j\sigma^j\left(\bar{X}_u^{j,\eta}\right) - \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{D,\eta}\right)du \\ &+ \frac{1}{2}\int_0^s \sigma^0\left(\bar{X}_u^{d+1,\eta}\right) - \sigma^0\left(X_{\hat{\tau}_u}^{D,\eta}\right)du + \sum_{j=1}^d \int_0^s \sigma^j\left(\bar{X}_u^{j,\eta}\right) - \sigma^j\left(X_{\hat{\tau}_u}^{D,\eta}\right)dW_u^j \\ &- \sum_{j=1}^d \int_0^s \left(\partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{D,\eta}\right)\Delta W_u^j + \sum_{\eta_u m < \eta_u j} \partial\sigma^j\sigma^m\left(X_{\hat{\tau}_u}^{D,\eta}\right)\Delta W_{\check{\tau}_u}^m\right)dW_u^j. \end{aligned}$$

Using the Burkholder-Davis-Gundy inequality and a convexity inequality, we obtain

$$\mathbb{E}\left[\sup_{s\leq t}\left\|X_s^{NV,\eta} - X_s^{D,\eta}\right\|^{2p}\right] \leq (2d+2)^{2p-1}\left(1 + KT^{p-1}\right)\left(\sum_{j=0}^{d+1} I_j + \sum_{j=1}^d E_j\right), \tag{3.5.7}$$

where K is the constant that appears in the Burkholder-Davis-Gundy inequality,

$$E_j = \int_0^t \mathbb{E}\left[\left\|\sigma^j\left(\bar{X}_u^{j,\eta}\right) - \sigma^j\left(X_{\hat{\tau}_u}^{D,\eta}\right) - \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{D,\eta}\right)\Delta W_u^j - \sum_{\eta_u m < \eta_u j} \partial\sigma^j\sigma^m\left(X_{\hat{\tau}_u}^{D,\eta}\right)\Delta W_{\check{\tau}_u}^m\right\|^{2p}\right]du,$$

and

$$I_j = \mathbb{E}\left[\sup_{s\leq t}\left\|\int_0^s F^j\left(\bar{X}_u^{j,\eta}\right) - F^j\left(X_{\hat{\tau}_u}^{D,\eta}\right)du\right\|^{2p}\right],$$

with $F^0 = F^{d+1} = \sigma^0$ and $F^j = \partial\sigma^j\sigma^j$, for $j \in \{1,\dots,d\}$.

Step 1: estimation of E_j, for $j \in \{1,\dots,d\}$.

We first introduce the vector R_u^j, for $u \in [0,t]$, defined by

$$R_u^j = \sigma^j\left(\bar{X}_u^{j,\eta}\right) - \sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right) - \partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right)\Delta W_u^j - \sum_{\eta_u m<\eta_u j}\partial\sigma^j\sigma^m\left(X_{\hat{\tau}_u}^{NV,\eta}\right)\Delta W_{\check{\tau}_u}^m,$$

with coordinates $(R_u^{ij})_{1\leq i\leq n}$. Denoting

$$J_u^{j,\eta} = \sum_{\eta_u m<\eta_u j}\sigma^m\left(X_{\hat{\tau}_u}^{NV,\eta}\right)\Delta W_{\check{\tau}_u}^m + \sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right)\Delta W_u^j,$$

R_u^{ij} rewrites

$$R_u^{ij} = \sigma^{ij}\left(\bar{X}_u^{j,\eta}\right) - \sigma^{ij}\left(X_{\hat{\tau}_u}^{NV,\eta}\right) - \nabla\sigma^{ij}\left(X_{\hat{\tau}_u}^{NV,\eta}\right).\ J_u^{j,\eta}.$$

By the mean value theorem, there exists $\alpha_u^{ij} \in [0,1]$ such that

$$\sigma^{ij}\left(\bar{X}_u^{j,\eta}\right) - \sigma^{ij}\left(X_{\hat{\tau}_u}^{NV,\eta}\right) = \nabla\sigma^{ij}\left(\xi_u^{ij}\right).\left(\bar{X}_u^{j,\eta} - X_{\hat{\tau}_u}^{NV,\eta}\right),$$

with

$$\xi_u^{ij} = X_{\hat{\tau}_u}^{NV,\eta} + \alpha_u^{ij}\left(\bar{X}_u^{j,\eta} - X_{\hat{\tau}_u}^{NV,\eta}\right).$$

Hence, introducing

$$\bar{R}_u^j = \bar{X}_u^{j,\eta} - X_{\hat{\tau}_u}^{NV,\eta} - J_u^{j,\eta},$$

we get

$$R_u^{ij} = \nabla\sigma^{ij}\left(\xi_u^{ij}\right).\ \bar{R}_u^j + \left(\nabla\sigma^{ij}\left(\xi_u^{ij}\right) - \nabla\sigma^{ij}\left(X_{\hat{\tau}_u}^{NV,\eta}\right)\right).\ J_u^{j,\eta}.$$

Let us now estimate R^{ij}. Since $\partial\sigma^j$ is locally Lipschitz with polynomially growing Lipschitz constant, we get a constant $\alpha_0 \in \mathbb{R}_+$ independent of N such that

$$\left|R_u^{ij}\right|^{2p} \leq \alpha_0\left(\left\|\nabla\sigma^{ij}\left(\xi_u^{ij}\right)\right\|^{2p}\left\|\bar{R}_u^j\right\|^{2p} + \left\|\xi_u^{ij} - X_{\hat{\tau}_u}^{NV,\eta}\right\|^{2p}\left\|J_u^{j,\eta}\right\|^{2p}\right). \tag{3.5.8}$$

Then, using the Cauchy-Schwarz inequality, we get

$$\mathbb{E}\left[\left\|\xi_u^{ij} - X_{\hat{\tau}_u}^{NV,\eta}\right\|^{2p}\left\|J_u^{j,\eta}\right\|^{2p}\right] \leq \left(\mathbb{E}\left[\left\|\xi_u^{ij} - X_{\hat{\tau}_u}^{NV,\eta}\right\|^{4p}\right]\mathbb{E}\left[\left\|J_u^{j,\eta}\right\|^{4p}\right]\right)^{\frac{1}{2}}.$$

Applying (3.5.4) from Lemma 3.5.3, we obtain

$$\mathbb{E}\left[\left\|\xi_u^{ij} - X_{\hat{\tau}_u}^{NV,\eta}\right\|^{4p}\right] \leq C_1\left(1 + \|x\|^{4p}\right)h^{2p}.$$

Once again, combining the Cauchy-Schwarz inequality, the Lipschitz property for $\sigma^j, j \in \{1, \ldots, d\}$, and (3.5.2) from Lemma 3.5.3, we get $\beta_1 \in \mathbb{R}_+$ independent of N such that

$$\mathbb{E}\left[\left\|J_u^{j,\eta}\right\|^{4p}\right] \leq \beta_1 h^{2p}.$$

For the last term in the right-hand side of (3.5.8) we obtain

$$\mathbb{E}\left[\left\|\xi_u^{ij} - X_{\hat{\tau}_u}^{NV,\eta}\right\|^{2p}\left\|J_u^{j,\eta}\right\|^{2p}\right] \leq \alpha_1 h^{2p},$$

where $\alpha_1 = \left(\beta_1 C_1\left(1 + \|x\|^{4p}\right)\right)^{\frac{1}{2}}$. For the first term in the right-hand side of (3.5.8), by the Lipschitz property of σ^j, $\partial\sigma^j$ is bounded by a constant denoted by M, so it remains to evaluate $\bar{R}^j$. Writing $\bar{R}^j$ in integral form, we obtain

$$\begin{aligned}
\bar{R}_u^j &= \bar{X}_u^j - X_{\hat{\tau}_u}^{NV,\eta} - J_u^{j,\eta} \\
&= \frac{1}{2}\int_{\hat{\tau}_u}^{\check{\tau}_u} \mathbb{1}_{\{\eta_u=1\}}\sigma^0\left(\bar{X}_v^0\right) + \mathbb{1}_{\{\eta_u=-1\}}\sigma^0\left(\bar{X}_v^{d+1}\right)dv \\
&+ \sum_{\eta_u m < \eta_u j} \frac{1}{2}\int_{\hat{\tau}_u}^{\check{\tau}_u} \partial\sigma^m\sigma^m\left(\bar{X}_v^m\right)dv + \frac{1}{2}\int_{\hat{\tau}_u}^{u} \partial\sigma^j\sigma^j\left(\bar{X}_v^j\right)dv \\
&+ \sum_{\eta_u m < \eta_u j} \int_{\hat{\tau}_u}^{\check{\tau}_u} \sigma^m\left(\bar{X}_v^j\right) - \sigma^m\left(X_{\hat{\tau}_u}^{NV,\eta}\right)dW_v^m + \int_{\hat{\tau}_u}^{u} \sigma^j\left(\bar{X}_v^j\right) - \sigma^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right)dW_v^j.
\end{aligned}$$

Combining a convexity inequality, the Burkholder-Davis-Gundy inequality and the Lipschitz property of σ^m for $m \in \{0, \ldots, d\}$ and $\partial\sigma^m\sigma^m$ for $m \in \{1, \ldots, d\}$, together with (3.5.3) and (3.5.4) from Lemma 3.5.3, we get $\beta_2 \in \mathbb{R}_+$ independent of N such that

$$\mathbb{E}\left[\left\|\bar{R}_u^j\right\|^{2p}\right] \leq \beta_2 h^{2p}.$$

For the first term in the right-hand side of (3.5.8), we obtain

$$\mathbb{E}\left[\left\|\nabla\sigma^{i,j}\left(\xi_u^{ij}\right)\right\|^{2p}\left\|\bar{R}_u^j\right\|^{2p}\right] \leq \alpha_2 h^{2p},$$

where $\alpha_2 = M^{2p}\beta_2$. This leads us to the following estimation

$$\mathbb{E}\left[\left|R_u^{ij}\right|^{2p}\right] \leq \alpha_0 h^{2p},$$

where $\alpha_3 = \alpha_0\left(\alpha_1 + \alpha_2\right)$. Therefore

$$\mathbb{E}\left[\left\|R_u^j\right\|^{2p}\right] \leq n^p\alpha_3 h^{2p}.$$

Hence, by introducing R_u^j in the expression of E_j and by using the Lipschitz assumption

$$\mathbb{E}\left[\left\|\sigma^j\left(\bar{X}_u^{j,\eta}\right)-\sigma^j\left(X_{\hat{\tau}_u}^{D,\eta}\right)-\partial\sigma^j\sigma^j\left(X_{\hat{\tau}_u}^{D,\eta}\right)\Delta W_u^j-\sum_{\eta_u m<\eta_u j}\partial\sigma^j\sigma^m\left(X_{\hat{\tau}_u}^{D,\eta}\right)\Delta W_{\check{\tau}_u}^m\right\|^{2p}\right]$$
$$\leq (d+2)^{2p-1}\left(1+L^{2p}\right)\left(\mathbb{E}\left[\sup_{v\leq u}\left\|X_v^{NV,\eta}-X_v^{D,\eta}\right\|^{2p}\right]+\mathbb{E}\left[\left\|X_{\hat{\tau}_u}^{NV,\eta}-X_{\hat{\tau}_u}^{D,\eta}\right\|^{2p}\left|\Delta W_u^j\right|^{2p}\right]\right.$$
$$\left.+\sum_{m\neq j}\mathbb{E}\left[\left\|X_{\hat{\tau}_u}^{NV,\eta}-X_{\hat{\tau}_u}^{D,\eta}\right\|^{2p}|\Delta W_{\check{\tau}_u}^m|^{2p}\right]+\mathbb{E}\left[\left\|R_u^j\right\|^{2p}\right]\right).$$

Then, by independence, for all $m\in\{1,\ldots,m\}$

$$\mathbb{E}\left[\left\|X_{\hat{\tau}_u}^{NV,\eta}-X_{\hat{\tau}_u}^{D,\eta}\right\|^{2p}|\Delta W_{\check{\tau}_u}^m|^{2p}\right]=\mathbb{E}\left[\left\|X_{\hat{\tau}_u}^{NV,\eta}-X_{\hat{\tau}_u}^{D,\eta}\right\|^{2p}\right]\mathbb{E}\left[|\Delta W_{\check{\tau}_u}^m|^{2p}\right]$$
$$\leq \mathbb{E}\left[|G|^{2p}\right]T^p\,\mathbb{E}\left[\sup_{v\leq u}\left\|X_v^{NV,\eta}-X_v^{D,\eta}\right\|^{2p}\right],$$

where G is a normal random variable. Summing up these last inequalities, we get

$$E_j\leq\gamma_1\left(h^{2p}+\int_0^t\mathbb{E}\left[\sup_{v\leq u}\left\|X_{\hat{\tau}_v}^{NV,\eta}-X_{\hat{\tau}_v}^{D,\eta}\right\|^{2p}\right]du\right),\tag{3.5.9}$$

where $\gamma_1=(d+2)^{2p-1}\left(1+L^{2p}\right)\left(1+d\,\mathbb{E}\left[|G|^{2p}\right]T^p+n^p\alpha_3T\right)$.

Step 2: estimation of I_j, for $j\in\{0,\ldots,d+1\}$.

Let $j\in\{0,\ldots,d+1\}$.

$$\left\|\int_0^s F^j\left(\bar{X}_u^{j,\eta}\right)-F^j\left(X_{\hat{\tau}_u}^{D,\eta}\right)du\right\|^{2p}\leq 2^{2p-1}\left(\left\|\int_0^s F^j\left(\bar{X}_u^{j,\eta}\right)-F^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right)du\right\|^{2p}\right.$$
$$\left.+s^{2p-1}\int_0^s\left\|F^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right)-F^j\left(X_{\hat{\tau}_u}^{D,\eta}\right)\right\|^{2p}du\right).$$

Hence

$$I_j\leq\alpha_3\left(\mathbb{E}\left[\sup_{s\leq t}\left\|\int_0^s F^j\left(\bar{X}_u^{j,\eta}\right)-F^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right)du\right\|^{2p}\right]\right.$$
$$\left.+\int_0^t\mathbb{E}\left[\left\|F^j\left(X_{\hat{\tau}_u}^{NV,\eta}\right)-F^j\left(X_{\hat{\tau}_u}^{D,\eta}\right)\right\|^{2p}\right]du\right),$$

where $\alpha_3=2^{2p-1}\left(1+T^{2p-1}\right)$. Then, by using (3.5.5) from Lemma 3.5.3 for the first integral and the Lipschitz assumption for the second one, we get

$$I_j\leq\gamma_2\left(h^{2p}+\int_0^t\mathbb{E}\left[\sup_{v\leq u}\left\|X_{\hat{\tau}_v}^{NV,\eta}-X_{\hat{\tau}_v}^{D,\eta}\right\|^{2p}\right]du\right),\tag{3.5.10}$$

where $\gamma_2=\alpha_3\left(C_1+L^{2p}\right)$.

Step 3: conclusion

Finally, by combining (3.5.9), (3.5.10), together with (3.5.7), we obtain

$$\mathbb{E}\left[\sup_{s\le t}\left\|X_s^{NV,\eta}-X_s^{D,\eta}\right\|^{2p}\right]\le\gamma_3\left(h^{2p}+\int_0^t\mathbb{E}\left[\sup_{v\le u}\left\|X_{\hat{\tau}_v}^{NV,\eta}-X_{\hat{\tau}_v}^{D,\eta}\right\|^{2p}\right]du\right),$$

where $\gamma_3=(2d+2)^{2p-1}\left(1+KT^{p-1}\right)\left(d\gamma_1+(d+2)\,\gamma_2\right)$ and we complete the proof using Gronwall's lemma since Lemmas 3.3.5, 3.5.1 and 3.5.3 ensure that $\mathbb{E}\left[\sup_{s\le t}\left\|X_s^{NV,\eta}-X_s^{D,\eta}\right\|^{2p}\right]$ is finite.
■

Chapter 4

Asymptotic error distribution for the Ninomiya-Victoir scheme in the commutative case

This chapter corresponds to an article written with Émmanuelle Clément and Benjamin Jourdain [2]. It has been prepublished.

Abstract. In Chapter 3 we proved strong convergence with order 1 of the Ninomiya-Victoir scheme X^{NV} with time step T/N to the solution X of the limiting SDE when the Brownian vector fields commute. In this paper, we prove that the normalized error process $N\left(X - X^{NV}\right)$ converges to an affine SDE with source terms involving the Lie brackets between the Brownian vector fields and the drift vector field. This result ensures that the strong convergence rate is actually 1 when the Brownian vector fields commute, but at least one of them does not commute with the drift vector field. When all the vector fields commute the limit vanishes. Our result is consistent with the fact that the Ninomiya-Victoir scheme solves the SDE in this case.

4.1 Introduction

We consider a general n-dimensional stochastic differential equation, driven by a d-dimensional standard Brownian motion $W = \left(W^1, \ldots, W^d\right)$, of the form

$$\begin{cases} dX_t = b(X_t)dt + \sum_{j=1}^{d} \sigma^j(X_t)dW_t^j, \ t \in [0,T] \\ X_0 = x, \end{cases} \tag{4.1.1}$$

where $x \in \mathbb{R}^n$ is the starting point, $b : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is the drift coefficient and $\sigma^j : \mathbb{R}^n \longrightarrow \mathbb{R}^n, j \in \{1, \ldots, d\}$, are the Brownian vector fields. We are interested in the study of the normalized error process for the Ninomiya-Victoir scheme. To do so we will consider in the whole paper a regular time grid, with time step $h = T/N$, of the time interval $[0,T]$. We introduce some notations to define the Ninomiya-Victoir scheme. Let

- $(t_k = kh)_{0 \leq k \leq N}$ be the subdivision of $[0,T]$ with equal time step h,

- $\Delta W_s^j = W_s^j - W_{t_k}^j$, for $s \in (t_k, t_{k+1}]$ and $j \in \{1, \ldots, d\}$,
- $\Delta s = s - t_k$, for $s \in (t_k, t_{k+1}]$.

Let $V : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be Lipschitz continuous and let $\exp(tV)x_0$ denote the solution, at time $t \in \mathbb{R}$, of the following ordinary differential equation in $\mathbb{R}^n$

$$\begin{cases} \dfrac{dx(t)}{dt} = V\left(x(t)\right) \\ x(0) = x_0. \end{cases}$$

To deal with the Ninomiya-Victoir scheme, it is more convenient to rewrite the stochastic differential equation (4.1.1) in Stratonovich form. Assuming $\mathcal{C}^1$ regularity for the vector fields, the Stratonovich form of (4.1.1) is given by:

$$\begin{cases} dX_t = \sigma^0(X_t)dt + \displaystyle\sum_{j=1}^{d} \sigma^j(X_t) \circ dW_t^j \\ X_0 = x, \end{cases}$$

where $\sigma^0 = b - \dfrac{1}{2}\displaystyle\sum_{j=1}^{d} \partial\sigma^j\sigma^j$ and $\partial\sigma^j$ is the Jacobian matrix of σ^j defined as follows

$$\partial\sigma^j = \left(\partial_{x_k}\sigma^{ij}\right)_{1\le i,k\le n}. \tag{4.1.2}$$

We recall that the Ninomiya-Victoir scheme [44] is given by:

- starting point: $X_{t_0}^{NV,\eta} = x$,
- for $k \in \{0\ldots, N-1\}$,
  if $\eta_{k+1} = 1$:
  $$X_{t_{k+1}}^{NV,\eta} = \exp\left(\frac{h}{2}\sigma^0\right)\exp\left(\Delta W_{t_{k+1}}^d\sigma^d\right)\ldots\exp\left(\Delta W_{t_{k+1}}^1\sigma^1\right)\exp\left(\frac{h}{2}\sigma^0\right)X_{t_k}^{NV,\eta},$$
  and if $\eta_{k+1} = -1$:
  $$X_{t_{k+1}}^{NV,\eta} = \exp\left(\frac{h}{2}\sigma^0\right)\exp\left(\Delta W_{t_{k+1}}^1\sigma^1\right)\ldots\exp\left(\Delta W_{t_{k+1}}^d\sigma^d\right)\exp\left(\frac{h}{2}\sigma^0\right)X_{t_k}^{NV,\eta},$$

where $\eta = (\eta_k)_{k\ge1}$ is a sequence of independent, identically distributed Rademacher random variables independent of W. In Chapter 2, we proved strong convergence with order $1/2$: for all $p \ge 1$, there exists a constant $C_{NV} \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$,

$$\mathbb{E}\left[\max_{0\le k\le N}\left\|X_{t_k} - X_{t_k}^{NV,\eta}\right\|^{2p}\right] \le C_{NV}\left(1 + \|x\|^{2p}\right)h^p.$$

In Chapter 3, we studied the stable convergence in law of the normalized error defined by $V^N = \sqrt{N}\left(X - X^{NV,\eta}\right)$. The theory of stable convergence, introduced by Rényi [50] was developed by Kurtz-Protter[35], Jacod [30] and Jacod-Protter [31]. The asymptotic distribution of the normalized error for the Euler continuous scheme was established by Kurtz and Protter in [35]. The asymptotic behavior of the normalized error processes for the Milstein scheme

[42], which is known to exhibit strong convergence with order 1, was studied by Yan in [55]. In both cases, the normalized error converges to the solution of an affine SDE with a source term involving additional randomness given by a Brownian motion independent of the one driving both the SDE and the scheme. In Chapter 3, we showed the stable convergence in law of V^N to the solution V of the affine SDE with source terms involving the Lie brackets between the Brownian vector fields:

$$V_t = \sqrt{\frac{T}{2}} \sum_{j=1}^{d} \sum_{m=1}^{j-1} \int_0^t \left[\sigma^j, \sigma^m\right] (X_s)\, dB_s^{j,m} + \int_0^t \partial b\,(X_s)\, V_s ds + \sum_{j=1}^{d} \int_0^t \partial \sigma^j\,(X_s)\, V_s dW_s^j,$$

where $\left[\sigma^j, \sigma^m\right] = \partial \sigma^m \sigma^j - \partial \sigma^j \sigma^m$, for $j, m \in \{1, \ldots, d\}, m < j$, denotes the Lie bracket between the Brownian vector fields σ^j and σ^m, ∂b is the Jacobian matrix of b, defined analogously to (4.1.2), and $(B_t)_{0 \leq t \leq T}$ is a standard $d(d-1)/2$-dimensional Brownian motion independent of W. The limit vanishes when the Brownian vector fields commute: for all $j, m \in \{1, \ldots, d\}$,

$$\left[\sigma^j, \sigma^m\right] = \partial \sigma^m \sigma^j - \partial \sigma^j \sigma^m = 0. \tag{$\mathcal{C}$}$$

When the Brownian vector fields commute, the order of integration of these fields no longer matters, since Frobenius' theorem ensures (see [17] or [18]) the commutativity of the associated flows. The sequence η is then useless. Therefore, the Ninomiya-Victoir scheme may be written as follows

- starting point: $X_{t_0}^{NV} = x$,
- for $k \in \{0 \ldots, N-1\}$,

$$X_{t_{k+1}}^{NV} = \exp\left(\frac{h}{2}\sigma^0\right) \exp\left(\Delta W_{t_{k+1}}^d \sigma^d\right) \ldots \exp\left(\Delta W_{t_{k+1}}^1 \sigma^1\right) \exp\left(\frac{h}{2}\sigma^0\right) X_{t_k}^{NV}.$$

Under some regularity assumptions, we proved, in Chapter 3, strong convergence with order 1 of the Ninomiya-Victoir scheme when the commutativity condition ($\mathcal{C}$) holds. More precisely, we showed the following result.

Theorem 4.1.1 *Assume that*

- *for all* $j \in \{1, \ldots, d\}, \sigma^j \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ *with bounded first order derivatives,*
- $\sigma^0 \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ *with bounded first order derivatives and polynomially growing second order derivatives,*
- $\sum_{j=1}^{d} \partial \sigma^j \sigma^j$ *is a Lipschitz continuous function,*

and that the commutativity condition ($\mathcal{C}$) *holds. Then, for all* $p \geq 1$, *there exists a constant* $C'_{NV} \in \mathbb{R}_+^*$ *such that for all* $N \in \mathbb{N}^*$,

$$\mathbb{E}\left[\max_{0 \leq k \leq N} \left\| X_{t_k} - X_{t_k}^{NV} \right\|^{2p}\right] \leq C'_{NV} h^{2p}.$$

In the present paper, we assume that the commutativity condition ($\mathcal{C}$) holds and we focus on the convergence in law of the normalized error defined by $U^N = N\left(X - X^{NV}\right)$. This paper is organized as follows. In section 2, we define an adapted interpolation between time grid points and derive its Itô decomposition. Then, we provide a suitable decomposition of the normalized error $U^N = N\left(X - X^{NV}\right)$ of the form

$$U_t^N = Q_t^N + J_t^N + \left(\int_0^t H_s^{0,N} U_s^N ds + \sum_{j=1}^d \int_0^t H_s^{j,N} U_s^N dW_s^j\right),$$

where $H^{j,N}$, for $j \in \{0, \ldots, d\}$, take values in $\mathbb{R}^n \otimes \mathbb{R}^n$, Q^N is a remainder term and J^N a source term, to study its stable convergence in law. In section 3, we prove the stable convergence in law of U^N to the solution of the following SDE:

$$U_t = \frac{T}{2\sqrt{3}} \sum_{j=1}^d \int_0^t \left[\sigma^0, \sigma^j\right](X_s)\, d\tilde{B}_s^j + \int_0^t \partial b\,(X_s)\, U_s ds + \sum_{j=1}^d \int_0^t \partial\sigma^j\,(X_s)\, U_s dW_s^j,$$

where $\left(\tilde{B}_t\right)_{0 \le t \le T}$ is a standard d-dimensional Brownian motion independent of W. This result ensures that the strong convergence rate is actually 1 when the Brownian vector fields commute, but at least one of them does not commute with the drift vector field σ^0. It is not surprising that the limit vanishes when all the vector fields σ^j, for $j \in \{0, \ldots, d\}$, commute, since the Ninomiya-Victoir scheme solves the SDE (4.1.1) in this case.

Notation

In the following, we introduce some more notations which will be used throughout this paper.

- Let $\hat{\tau}_s$ be the last time discretization before $s \in [0, T]$, ie $\hat{\tau}_s = t_k$ if $s \in (t_k, t_{k+1}]$, and for $s = t_0 = 0$, we set $\hat{\tau}_0 = 0$.
- Let $\check{\tau}_s$ be the first time discretization after $s \in [0, T]$, ie $\check{\tau}_s = t_{k+1}$ if $s \in (t_k, t_{k+1}]$, and for $s = t_0 = 0$, we set $\check{\tau}_0 = 0$.
- For the vector field σ^j, $j \in \{0, \ldots, d\}$, $\partial^2\sigma^j$ denotes the $n \times n \times n$-tensor $\left(\partial^2\sigma^j\right)^{i,k,l} = \partial^2_{x_l x_k}\sigma^{ij}$.
- The tensor product, between a $m \times p \times q-$tensor A and a vector b in $\mathbb{R}^q$ is denoted by $A \odot b$: for all $i \in \{1, \ldots, m\}$, for all $k \in \{1, \ldots, p\}$

$$(A \odot b)_{i,k} = \sum_{l=1}^q A^{i,k,l} b^l.$$

- To lighten up the notation, $\|.\|$ will denote both the Euclidean norm in $\mathbb{R}^n$ and its associated operator norm in $\mathbb{R}^n \otimes \mathbb{R}^n$.
- For a finite-dimensional normed vector space $(\mathcal{S}, \|.\|_{\mathcal{S}})$, $LIP^{pgc}_{loc}(\mathcal{S})$ denotes the space of locally Lipschitz with polynomially growing Lipschitz constant functions from $\mathbb{R}^n$ to $\mathcal{S}$:

$$LIP^{pgc}_{loc}(\mathcal{S}) = \left\{F : \mathbb{R}^n \longrightarrow \mathcal{S}, \exists c \in \mathbb{R}^*_+, q \in \mathbb{N}, \forall x, y \in \mathbb{R}^n, \|F(x) - F(y)\|_{\mathcal{S}} \le c\,(1 + \|x\|^q \vee \|y\|^q)\, \|x - y\|\right\}.$$

Remark 4.1.2

- *If $F \in \mathcal{C}^1(\mathbb{R}^n, \mathcal{S})$ with polynomially growing first order derivatives, then $F \in LIP^{pgc}_{loc}(\mathcal{S})$.*

- *If $A \in LIP^{pgc}_{loc}(\mathbb{R}^m \otimes \mathbb{R}^p \otimes \mathbb{R}^q)$, $M \in LIP^{pgc}_{loc}(\mathbb{R}^p \otimes \mathbb{R}^q)$ and $b \in LIP^{pgc}_{loc}(\mathbb{R}^q)$, then $A \odot b \in LIP^{pgc}_{loc}(\mathbb{R}^m \otimes \mathbb{R}^p)$ and $Mb \in LIP^{pgc}_{loc}(\mathbb{R}^p)$.*

4.2 Adapted interpolation and main result

4.2.1 Main result

To study the normalized error process $U^N = N\left(X - X^{NV}\right)$, using the framework described in section 2 of Chapter 3, we provide the following adapted interpolation between time grid points:

$$\begin{cases} X_t^{NV} = h_{d+1}\left(\dfrac{\Delta t}{2}, \Delta W_t, \dfrac{\Delta t}{2}; X_{\hat{\tau}_t}^{NV}\right) \\ X_0^{NV} = x, \end{cases} \tag{4.2.1}$$

where $h_{d+1} : \mathbb{R}^{d+1} \longrightarrow \mathbb{R}^n$ is defined by

$$h_{d+1}(t_0, \dots, t_{d+1}; y) = \exp\left(t_{d+1}\sigma^0\right)\exp\left(t_d\sigma^d\right)\dots\exp\left(t_1\sigma^1\right)\exp\left(t_0\sigma^0\right)y,$$

for the initial condition $y \in \mathbb{R}^n$. The main result of this paper is the following theorem, which gives the stable convergence in law of the normalized error process U^N.

Theorem 4.2.1 *Assume that*

- *for all $j \in \{0, \dots, d\}, \sigma^j \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first order derivatives, $\partial\sigma^j \in LIP^{pgc}_{loc}(\mathbb{R}^n \otimes \mathbb{R}^n)$ and $\partial^2\sigma^j \in LIP^{pgc}_{loc}(\mathbb{R}^n \otimes \mathbb{R}^n \otimes \mathbb{R}^n)$,*
- *for all $j \in \{1, \dots, d\}, \partial\sigma^j\sigma^j$ is a Lipschitz continuous function,*

and that the commutativity condition ($\mathcal{C}$) holds. Then:

$$U^N = N\left(X - X^{NV}\right) \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} U,$$

where U is the unique solution of the following affine equation:

$$U_t = \frac{T}{2\sqrt{3}}\sum_{j=1}^{d}\int_0^t \left[\sigma^0, \sigma^j\right](X_s)\,d\tilde{B}_s^j + \int_0^t \partial b(X_s)\,U_s ds + \sum_{j=1}^{d}\int_0^t \partial\sigma^j(X_s)\,U_s dW_s^j,$$

and $\left(\tilde{B}_t\right)_{0\le t\le T}$ is a standard d-dimensional Brownian motion independent of W.

In order to prove this theorem we will proceed as follows. Firstly, we will derive the dynamics of the interpolated Ninomiya-Victoir scheme X^{NV} using Itô's formula. Then we will decompose the dynamics as follows

$$dX_t^{NV} = b(X_t^{NV})dt + \sum_{j=1}^{d}\sigma^j(X_t^{NV})dW_t^j + d\bar{J}_t^N + d\bar{Q}_t^N, \tag{4.2.2}$$

where $\bar{J}^N$ is a term with strong order 1 and $\bar{Q}^N$ is a remainder term such that $N\bar{Q}^N$ uniformly converges in probability to 0. Using (4.2.2) we can derive a convenient decomposition of the normalized error process U^N as

$$U_t^N = Q_t^N + J_t^N + \left(\int_0^t H_s^{0,N} U_s^N ds + \sum_{j=1}^d \int_0^t H_s^{j,N} U_s^N dW_s^j\right),$$

where $H^{j,N}$, for $j \in \{0,\ldots,d\}$ take values in $\mathbb{R}^n \otimes \mathbb{R}^n$, $J^N = -N\bar{J}^N$ is a source term and $Q^N = -N\bar{Q}^N$ is a remainder term. By analyzing the stable convergence in law of the source term J^N, we will be able to prove the above result using Theorem 2.5 in Chapter 3.

4.2.2 Itô decomposition of X^{NV}

To get the Itô decomposition of X^{NV}, the main difficulty is to explicit the derivatives of h_{d+1} given by (4.2.1). Even if the commutativity of the Brownian vector fields $\sigma^j, j \in \{1,\ldots,d\}$, simplifies the calculation a lot, it is still cumbersome. To compute the derivatives of h_{d+1}, we begin by writing h_{d+1} in integral form as follows. The Frobenius theorem (see [17] or [18]) ensures that for all permutation ξ on $\{1,\ldots,d\}$:

$$\begin{aligned} h_{d+1}(t_0,\ldots,t_{d+1};y) = y &+ \int_0^{t_0} \sigma^0\left(\exp\left(s\sigma^0\right)y\right)ds \\ &+ \sum_{k=1}^d \int_0^{t_{\xi(k)}} \sigma^{\xi(k)}\left(\exp\left(s\sigma^{\xi(k)}\right)\left(\prod_{m<k}\exp\left(t_{\xi(m)}\sigma^{\xi(m)}\right)\right)\exp\left(t_0\sigma^0\right)y\right)ds \\ &+ \int_0^{t_{d+1}} \sigma^0\left(h_{d+1}(t_0,\ldots,t_d,s;y)\right)ds. \end{aligned} \tag{4.2.3}$$

First order derivative with respect to t_0.

To compute the derivative with respect to t_0, we choose $\xi = Id$, and we introduce the functions

$$h_j(t_0,\ldots,t_j;y) = \exp(t_j\sigma^j)\ldots\exp(t_1\sigma^1)\exp(t_0\sigma^0)y,$$

for $j \in \{0,\ldots,d\}$. Then, with $\xi = Id$, (4.2.3) becomes

$$\begin{aligned} h_{d+1}(t_0,\ldots,t_{d+1};y) = y + \int_0^{t_0}\sigma^0\left(h_0(s;y)\right)ds + \sum_{k=1}^d \int_0^{t_k}\sigma^k\left(h_k(t_0,\ldots,t_{k-1},s;y)\right)ds \\ + \int_0^{t_{d+1}}\sigma^0\left(h_{d+1}(t_0,\ldots,t_d,s;y)\right)ds, \end{aligned}$$

and it follows that

$$\begin{aligned} \partial_{t_0}h_{d+1}(t_0,\ldots,t_{d+1};y) = \sigma^0\left(h_0(t_0;y)\right) & \\ &+ \sum_{j=1}^d \int_0^{t_j}\partial\sigma^j\left(h_j(t_0,\ldots,t_{j-1},s;y)\right)\partial_{t_0}h_j(t_0,\ldots,t_{j-1},s;y)\,ds \\ &+ \int_0^{t_{d+1}}\partial\sigma^0\left(h_{d+1}(t_0,\ldots,t_d,s;y)\right)\partial_{t_0}h_{d+1}(t_0,\ldots,t_d,s;y)\,ds. \end{aligned}$$

Moreover,

$$\begin{aligned}\partial_{t_0} h_j\left(t_0, \ldots, t_j; y\right) &= \sigma^0\left(h_0\left(t_0; y\right)\right) \\ &\quad + \sum_{k=1}^{j} \int_0^{t_k} \partial \sigma^k\left(h_k\left(t_0, \ldots, t_{k-1}, s; y\right)\right) \partial_{t_0} h_k\left(t_0, \ldots, t_{k-1}, s; y\right) ds.\end{aligned}$$

Then, solving these linear differential equations, we obtain by induction

$$\partial_{t_0} h_{d+1}\left(t_0, \ldots, t_{d+1}; y\right) = R^{d+1}\left(t_0, \ldots, t_{d+1}; y\right) R^d\left(t_0, \ldots, t_d; y\right) \ldots R^1\left(t_0, t_1; y\right) \sigma^0\left(h_0\left(t_0; y\right)\right),$$

where, for $j \in \{1, \ldots, d+1\}$, R^j is the solution of the following linear system

$$\begin{cases} \dfrac{d}{dt} R^j\left(t_0, \ldots, t_{j-1}, t; y\right) = \partial \sigma^j\left(h_j\left(t_0, \ldots, t_{j-1}, t; y\right)\right) R^j\left(t_0, \ldots, t_{j-1}, t; y\right) \\ \quad R^j\left(t_0, \ldots, t_{j-1}, 0; y\right) = Id_n, \end{cases}$$

with $\sigma^{d+1} = \sigma^0$ by convention.

First order derivative with respect to $t_j, j \in \{1, \ldots, d\}$.

Choosing ξ such that $\xi(d) = j$, we get

$$\begin{aligned}\partial_{t_j} h_{d+1}\left(t_0, \ldots, t_{d+1}; y\right) &= \sigma^j\left(h_d\left(t_0, \ldots, t_d; y\right)\right) \\ &\quad + \int_0^{t_{d+1}} \partial \sigma^0\left(h_{d+1}\left(t_0, \ldots, t_d, s; y\right)\right) \partial_{t_j} h_{d+1}\left(t_0, \ldots, t_d, s; y\right) ds.\end{aligned} \tag{4.2.4}$$

Solving this linear differential equation, we obtain

$$\partial_{t_j} h_{d+1}\left(t_0, \ldots, t_{d+1}; y\right) = R^{d+1}\left(t_0, \ldots, t_{d+1}; y\right) \sigma^j\left(h_d\left(t_0, t_1, \ldots, t_d; y\right)\right).$$

First order derivative with respect to t_{d+1}.

The derivative with respect to t_{d+1} is trivial:

$$\partial_{t_{d+1}} h_{d+1}\left(t_0, \ldots, t_{d+1}; y\right) = \sigma^0\left(h_{d+1}\left(t_0, \ldots, t_{d+1}; y\right)\right).$$

Second order derivative with respect to $t_j, j \in \{1, \ldots, d\}$.

Using (4.2.4) and

$$\partial_{t_j} h_d\left(t_0, \ldots, t_d; y\right) = \sigma^j\left(h_d\left(t_0, \ldots, t_d; y\right)\right),$$

we have

$$\begin{aligned}\partial^2_{t_j t_j} h_{d+1}\left(t_0, \ldots, t_{d+1}; y\right) &= \partial \sigma^j \sigma^j\left(h_d\left(t_0, \ldots, t_d; y\right)\right) \\ &\quad + \int_0^{t_{d+1}} \left(\left(\left(\partial^2 \sigma^0 \circ h_{d+1}\right) \odot \partial_{t_j} h_{d+1}\right) \partial_{t_j} h_{d+1}\right)\left(t_0, \ldots, t_d, s; y\right) ds \\ &\quad + \int_0^{t_{d+1}} \left(\left(\partial \sigma^0 \circ h_{d+1}\right) \partial^2_{t_j t_j} h_{d+1}\right)\left(t_0, \ldots, t_d, s; y\right) ds.\end{aligned}$$

Solving this linear differential equation, we get

$$\begin{aligned}\partial^2_{t_j t_j} h_{d+1}\left(t_0,\ldots,t_{d+1};y\right) &= R^{d+1}\left(t_0,\ldots,t_{d+1};y\right)\partial\sigma^j\sigma^j\left(h_d\left(t_0,\ldots,t_d;y\right)\right)\\ &\quad+\int_0^{t_{d+1}} G^{d+1}\left(t_0,\ldots,t_{d+1},s;y\right)g^{d+1}\left(t_0,\ldots,t_d,s;y\right)ds,\end{aligned}$$

where

$$G^{d+1}\left(t_0,\ldots,t_{d+1},s;y\right) = R^{d+1}\left(t_0,\ldots,t_{d+1};y\right)\left(R^{d+1}\right)^{-1}\left(t_0,\ldots,t_d,s;y\right),$$

and

$$g^{d+1}\left(t_0,\ldots,t_d,s;y\right) = \left(\left(\partial^2\sigma^0\circ h_{d+1}\odot\partial_{t_j}h_{d+1}\right)\partial_{t_j}h_{d+1}\right)\left(t_0,\ldots,t_d,s;y\right).$$

Itô's formula.

Using Itô's formula, we obtain:

$$\begin{aligned}dX^{NV}_t &= \frac{1}{2}\partial_{t_0}h_{d+1}\left(\frac{\Delta t}{2},\Delta W_t,\frac{\Delta t}{2};X^{NV}_{\hat\tau_t}\right)dt+\sum_{j=1}^{d}\partial_{t_j}h_{d+1}\left(\frac{\Delta t}{2},\Delta W_t,\frac{\Delta t}{2};X^{NV}_{\hat\tau_t}\right)dW^j_t\\ &\quad+\frac{1}{2}\partial_{t_{d+1}}h_{d+1}\left(\frac{\Delta t}{2},\Delta W_t,\frac{\Delta t}{2};X^{NV}_{\hat\tau_t}\right)dt+\frac{1}{2}\sum_{j=1}^{d}\partial^2_{t_j t_j}h_{d+1}\left(\frac{\Delta t}{2},\Delta W_t,\frac{\Delta t}{2};X^{NV}_{\hat\tau_t}\right)dt\\ &= \frac{1}{2}Z^0_t dt+\sum_{j=1}^{d}Y^{d+1}_{t,t}\sigma^j\left(\bar X_{t,t}\right)dW^j_t\\ &\quad+\frac{1}{2}\sigma^0\left(X^{NV}_t\right)dt+\frac{1}{2}\sum_{j=1}^{d}\left(Y^{d+1}_{t,t}\partial\sigma^j\sigma^j\left(\bar X_{t,t}\right)+Z^{d+1,j}_t\right)dt,\end{aligned}\tag{4.2.5}$$

where for $t\in[0,T]$

$$Z^0_t = Y^{d+1}_{t,t}\ldots Y^1_{t,t}\sigma^0\left(\bar X^0_t\right),$$

for $j\in\{1,\ldots,d\}, s\in[\hat\tau_t,t]$

$$Y^j_{t,s} = R^j\left(\frac{\Delta t}{2},\Delta W^1_t,\ldots,\Delta W^{j-1}_t,\Delta W^j_s;X^{NV}_{\hat\tau_t}\right),$$

$$Y^{d+1}_{t,s} = R^{d+1}\left(\frac{\Delta t}{2},\Delta W^1_t,\ldots,\Delta W^d_t,\frac{\Delta s}{2};X^{NV}_{\hat\tau_t}\right),$$

$$\bar X^0_t = h_0\left(\frac{\Delta t}{2};X^{NV}_{\hat\tau_t}\right),$$

$$\bar X_{t,s} = h_d\left(\frac{\Delta t}{2},\Delta W_s;X^{NV}_{\hat\tau_t}\right),$$

$$Z^{d+1,j}_t = \frac{1}{2}\int_{\hat\tau_t}^{t}Y^{d+1}_{t,t}\left(Y^{d+1}_{t,s}\right)^{-1}\left(\partial^2\sigma^0\left(\bar X^{d+1}_{t,s}\right)\odot\left(Y^{d+1}_{t,s}\sigma^j\left(\bar X_{t,t}\right)\right)\right)Y^{d+1}_{t,s}\sigma^j\left(\bar X_{t,t}\right)\,ds,$$

and

$$\bar X^{d+1}_{t,s} = h_{d+1}\left(\frac{\Delta t}{2},\Delta W_t,\frac{\Delta s}{2};X^{NV}_{\hat\tau_t}\right).$$

The next lemma gives estimations of the moment of the Ninomiya-Victoir scheme X^{NV} and its increments. It also compares X^{NV} to the intermediate processes $\bar{X}^0$, $\bar{X}$ and $\bar{X}^{d+1}$. This result is very similar to Lemmas 4.4 and 4.5 in Chapter 3, that is why we omit its proof.

Lemma 4.2.2 *Assume that*

- *for all* $j \in \{1, \dots, d\}, \sigma^j \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ *with bounded first order derivatives,*
- σ^0 *and* $\sum_{j=1}^{d} \partial\sigma^j\sigma^j$ *are Lipschitz continuous functions.*

Then, for all $p \geq 1$, *there exists a constant* $C_0 \in \mathbb{R}_+^*$ *such that for all* $N \in \mathbb{N}^*, t \in [0, T]$,

$$\mathbb{E}\left[\left\|X_t^{NV}\right\|^{2p}\right] \leq C_0, \tag{4.2.6}$$

$$\mathbb{E}\left[\left\|X_t^{NV} - X_{\hat{\tau}_t}^{NV}\right\|^{2p}\right] \leq C_0 \left(\Delta t\right)^p, \tag{4.2.7}$$

$$\mathbb{E}\left[\left\|X_t^{NV} - \bar{X}_s^0\right\|^{2p}\right] \leq C_0 \left(\Delta t\right)^p, \text{ for all } s \in [\hat{\tau}_t, t], \tag{4.2.8}$$

$$\mathbb{E}\left[\left\|X_t^{NV} - \bar{X}_{t,s}\right\|^{2p}\right] \leq C_0 \left(\Delta t\right)^p, \text{ for all } s \in [\hat{\tau}_t, t], \tag{4.2.9}$$

$$\mathbb{E}\left[\left\|X_t^{NV} - \bar{X}_{t,s}^{d+1}\right\|^{2p}\right] \leq C_0 \left(\Delta t\right)^{2p}, \text{ for all } s \in [\hat{\tau}_t, t]. \tag{4.2.10}$$

Moreover if the commutativity condition ($\mathcal{C}$) *holds:*

$$\mathbb{E}\left[\left\|X_t^{NV} - \bar{X}_{t,t}\right\|^{2p}\right] \leq C_0 \left(\Delta t\right)^{2p}. \tag{4.2.11}$$

4.2.3 Suitable decomposition of X^{NV}

Our derivation of a decomposition of the form (4.2.2) is carried out in two steps. The first step will consist in approximating the dynamics of X^{NV} with strong order $3/2$. In the second step we will identify the appropriate source term J^N.

Approximation with strong order $3/2$

The dynamics (4.2.5) of X^{NV} is not really tractable to study the normalized error process U^N. In the following, we provide an approximation with strong order $3/2$ of the theoretical dynamics of X^{NV}. The goal is to be able to write the dynamics of X^{NV} in the form (4.2.2). We begin by indicating a natural approximation with strong order 2, respectively $3/2$, of the intermediate processes Y^{d+1} and $\left(\sigma^j\left(\bar{X}_{t,t}\right)\right)_{0 \leq t \leq T}$, respectively $\left(\partial\sigma^j\sigma^j\left(\bar{X}_{t,t}\right)\right)_{0 \leq t \leq T}$ and $Z^{d+1,j}$.

Proposition 4.2.3 *Assume that*

- *for all* $j \in \{0, \dots, d\}, \sigma^j \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ *with bounded first order derivatives,* $\partial\sigma^j \in LIP_{loc}^{pgc}(\mathbb{R}^n \otimes \mathbb{R}^n)$ *and* $\partial^2\sigma^j \in LIP_{loc}^{pgc}(\mathbb{R}^n \otimes \mathbb{R}^n \otimes \mathbb{R}^n)$,
- $\sum_{j=1}^{d} \partial\sigma^j\sigma^j$ *is a Lipschitz continuous function.*

Then, for all $p \geq 1$, there exists a constant $C_1 \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*, t \in [0,T]$, and all $j \in \{1,\ldots,d\}$,

$$\mathbb{E}\left[\left\|Y_{t,s}^{d+1} - Id_n - \frac{1}{2}\Delta s\partial\sigma^0\left(X_t^{NV}\right)\right\|^{2p}\right] \leq C_1\left(\Delta t\right)^{4p}, \text{ for all } s \in [\hat{\tau}_t, t], \tag{4.2.12}$$

$$\mathbb{E}\left[\left\|Z_t^{d+1,j} - \frac{1}{2}\Delta t\left(\partial^2\sigma^0 \odot \sigma^j\right)\sigma^j\left(X_{\hat{\tau}_t}^{NV}\right)\right\|^{2p}\right] \leq C_1\left(\Delta t\right)^{3p}. \tag{4.2.13}$$

Moreover if the commutativity condition ($\mathcal{C}$) holds:

$$\mathbb{E}\left[\left\|\sigma^j\left(\bar{X}_{t,t}\right) - \sigma^j\left(X_t^{NV}\right) + \frac{1}{2}\Delta t\partial\sigma^j\sigma^0\left(X_t^{NV}\right)\right\|^{2p}\right] \leq C_1\left(\Delta t\right)^{4p}, \tag{4.2.14}$$

$$\mathbb{E}\left[\left\|\partial\sigma^j\sigma^j\left(\bar{X}_{t,t}\right) - \partial\sigma^j\sigma^j\left(X_t^{NV}\right) + \frac{1}{2}\Delta t\left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right)\right\|^{2p}\right] \leq C_1\left(\Delta t\right)^{3p}. \tag{4.2.15}$$

In (4.2.13) and (4.2.15), we have replaced by $X_{\hat{\tau}_t}^{NV}$ the argument of the functions which would naturally appears in Taylor expansion when Lemma 4.2.2 ensures that the strong order 3/2 is preserved. Although the above approximations are very intuitive, their proofs are both heavy and technical. That is why, the proof of this proposition is postponed to the Appendix. To obtain an approximation with strong order 3/2 of the form (4.2.2), it remains to estimate $\int_0^t Z_s^0 ds$, for $t \in [0,T]$.

Proposition 4.2.4 *Assume that*

- *$\sigma^0 \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first order derivatives, and $\partial\sigma^0 \in LIP_{loc}^{pgc}(\mathbb{R}^n \otimes \mathbb{R}^n)$,*
- *for all $j \in \{1,\ldots,d\}, \sigma^j \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first order derivatives, $\partial\sigma^j \in LIP_{loc}^{pgc}(\mathbb{R}^n \otimes \mathbb{R}^n)$, and $\partial^2\sigma^j \in LIP_{loc}^{pgc}(\mathbb{R}^n \otimes \mathbb{R}^n \otimes \mathbb{R}^n)$,*
- *for all $j \in \{1,\ldots,d\}, \partial\sigma^j\sigma^j$ is a Lipschitz continuous function.*

Then, denoting by,

$$\begin{aligned}\theta_t^0 = {} & \sigma^0\left(\bar{X}_t^0\right) + \sum_{j=1}^d \Delta W_t^j\partial\sigma^j\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right) + \frac{1}{2}\sum_{j=1}^d \Delta t\left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right) \\ & + \frac{1}{2}\Delta t\partial\sigma^0\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right),\end{aligned}$$

for $t \in [0,T]$,we have that for all $p \geq 1$, there exists a constant $C_2 \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$,

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\int_0^t\left(Z_s^0 - \theta_s^0\right)ds\right\|^{2p}\right] \leq C_2 h^{3p}.$$

This approximation is not really intuitive. Indeed, to get this result, we approximate the process Z^0, and then we use the integration by parts formula to identify the dominant contribution. The proof of this proposition is also postponed to the Appendix.

We are now able to approximate the dynamics of the Ninomiya-Victoir scheme. Using $b = \sigma^0 + \frac{1}{2}\sum_{j=1}^{d} \partial\sigma^j\sigma^j$, together with Propositions 4.2.3 and 4.2.4, it is easy to see that an adequate decomposition of the dynamics of X^{NV} is given by

$$\begin{aligned}
dX_t^{NV} &= b\left(X_t^{NV}\right)dt + \sum_{j=1}^{d}\sigma^j\left(X_t^{NV}\right)dW_t^j + \frac{1}{2}\left(\sigma^0\left(\bar{X}_t^0\right) - \sigma^0\left(X_t^{NV}\right)\right)dt \\
&+ \frac{1}{2}\sum_{j=1}^{d}\Delta t\left[\sigma^j,\sigma^0\right]\left(X_t^{NV}\right)dW_t^j + \frac{1}{2}\sum_{j=1}^{d}\Delta W_t^j\partial\sigma^j\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right)dt \\
&+ \frac{1}{4}\Delta t\left(\partial\sigma^0\sigma^0 + \sum_{j=1}^{d}\left(\partial^2\sigma^0\odot\sigma^j + \partial\sigma^0\partial\sigma^j\right)\sigma^j\right)\left(X_{\hat{\tau}_t}^{NV}\right)dt + d\bar{Q}_t^{1,N},
\end{aligned} \tag{4.2.16}$$

where, $\bar{Q}^{1,N}$ is defined by $\bar{Q}_0^{1,N} = 0$ and

$$\begin{aligned}
d\bar{Q}_t^{1,N} &= \frac{1}{2}\left(Z_t^0 - \theta_t^0\right) + \sum_{j=1}^{d}\left(Y_{t,t}^{d+1}\sigma^j\left(\bar{X}_{t,t}\right) - \sigma^j\left(X_t^{NV}\right) - \frac{1}{2}\Delta t\left[\sigma^j,\sigma^0\right]\left(X_t^{NV}\right)\right)dW_t^j \\
&+ \frac{1}{2}\sum_{j=1}^{d}\left(Y_{t,t}^{d+1}\partial\sigma^j\sigma^j\left(\bar{X}_{t,t}\right) - \partial\sigma^j\sigma^j\left(X_t^{NV}\right)\right. \\
&\qquad\left. - \frac{1}{2}\Delta t\left(\partial\sigma^0\partial\sigma^j\sigma^j - \left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\right)\left(X_{\hat{\tau}_t}^{NV}\right)\right)dt \\
&+ \frac{1}{2}\sum_{j=1}^{d}\left(Z_t^{d+1,j} - \frac{1}{2}\Delta t\left(\partial^2\sigma^0\odot\sigma^j\right)\sigma^j\left(X_{\hat{\tau}_t}^{NV}\right)\right)dt,
\end{aligned}$$

is a remainder term with strong order 3/2. The proof of the following proposition is also postponed to the Appendix.

Proposition 4.2.5 *Let $p \geq 1$,. Under the assumptions of Theorem 4.2.1, there exists a constant $C_3 \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$,*

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\bar{Q}_t^{1,N}\right\|^{2p}\right] \leq C_3 h^{3p}.$$

Identification of the source term J^N

The goal of this subsection is to identify the source term J^N. Deducing the strong convergence with order 1 of the Ninomiya-Victoir scheme from the equation (4.2.16) does not look straightforward. Indeed, at first glance, in (4.2.16), the terms

$$\frac{N}{2}\int_0^T\left(\sigma^0\left(X_t^{NV}\right) - \sigma^0\left(\bar{X}_t^0\right)\right)dt \tag{4.2.17}$$

and

$$\frac{N}{2}\sum_{j=1}^{d}\int_0^T\Delta W_t^j\partial\sigma^j\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right)dt \tag{4.2.18}$$

seem to diverge as N goes to infinity. Actually, using the integration by parts formula, we show that both terms, (4.2.17) and (4.2.18), are bounded by a constant independent of N in L^2, which is consistent with Theorem 4.1.1. More precisely, we have the following result.

Proposition 4.2.6 *Let*

$$I_t^0 = \sum_{j=1}^d \int_0^t (t \wedge \check{\tau}_s - s)\, \partial\sigma^0\sigma^j \left(X_s^{NV}\right) dW_s^j - \frac{1}{2}\int_0^t (t \wedge \check{\tau}_s - s)\, \partial\sigma^0\sigma^0 \left(X_{\hat{\tau}_s}^{NV}\right) ds$$
$$- \frac{1}{2}\sum_{j=1}^d \int_0^t (t \wedge \check{\tau}_s - s) \left(\partial^2\sigma^0 \odot \sigma^j + \partial\sigma^0\partial\sigma^j\right) \sigma^j \left(X_{\hat{\tau}_s}^{NV}\right) ds,\ t \in [0,T],$$

and for $j \in \{1, \dots, d\}$

$$I_t^j = \int_0^t (t \wedge \check{\tau}_s - s)\, \partial\sigma^j\sigma^0 \left(X_s^{NV}\right) dW_s^j,\ t \in [0,T].$$

Let $p \geq 1$,. *Under the assumptions of Theorem 4.2.1, there exists a constant* $C_4 \in \mathbb{R}_+^*$ *such that for all* $N \in \mathbb{N}^*, j \in \{1, \dots, d\}$:

$$\mathbb{E}\left[\sup_{t \leq T} \left\| \int_0^t \sigma^0 \left(X_s^{NV}\right) - \sigma^0 \left(\bar{X}_s^0\right) ds - I_t^0 \right\|^{2p}\right] \leq C_4 h^{3p}, \tag{4.2.19}$$

$$\mathbb{E}\left[\sup_{t \leq T} \left\| \int_0^t \Delta W_s^j \partial\sigma^j\sigma^0 \left(X_{\hat{\tau}_s}^{NV}\right) ds - I_t^j \right\|^{2p}\right] \leq C_4 h^{3p}. \tag{4.2.20}$$

Proof : We start by proving (4.2.20) and we denote by

$$\Phi_t^N = \int_0^t \Delta W_s^j \partial\sigma^j\sigma^0 \left(X_{\hat{\tau}_s}^{NV}\right) ds - I_t^j.$$

Using the integration by parts formula, we have

$$\int_0^t \Delta W_s^j \partial\sigma^j\sigma^0 \left(X_{\hat{\tau}_s}^{NV}\right) ds = \int_0^t (t \wedge \check{\tau}_s - s)\, \partial\sigma^j\sigma^0 \left(X_{\hat{\tau}_s}^{NV}\right) dW_s^j.$$

Therefore,

$$\Phi_t^N = -\int_0^t (t \wedge \check{\tau}_s - s) \left(\partial\sigma^j\sigma^0 \left(X_s^{NV}\right) - \partial\sigma^j\sigma^0 \left(X_{\hat{\tau}_s}^{NV}\right)\right) dW_s^j.$$

Combining the Burkholder-Davis-Gundy inequality and a convexity inequality, we get a constant α_1 independent of N such that

$$\mathbb{E}\left[\sup_{t \leq T} \left\|\Phi_t^N\right\|^{2p}\right] \leq \alpha_1 \int_0^T (t \wedge \check{\tau}_s - s)^{2p} E(s)\, ds,$$

where

$$E(s) = \mathbb{E}\left[\left\|\partial\sigma^j\sigma^0 \left(X_s^{NV}\right) - \partial\sigma^j\sigma^0 \left(X_{\hat{\tau}_s}^{NV}\right)\right\|^{2p}\right].$$

As the function $\partial\sigma^j\sigma^0 \in LIP_{loc}^{pgc}(\mathbb{R}^n)$ is locally Lipschitz with polynomially growing Lipschitz constant, there exist $c \in \mathbb{R}_+^*$ and $q \in \mathbb{N}$ such that

$$\left\|\partial\sigma^j\sigma^0 \left(X_s^{NV}\right) - \partial\sigma^j\sigma^0 \left(X_{\hat{\tau}_s}^{NV}\right)\right\| \leq c\left(1 + \left\|X_s^{NV}\right\|^q \vee \left\|X_{\hat{\tau}_s}^{NV}\right\|^q\right) \left\|X_s^{NV} - X_{\hat{\tau}_s}^{NV}\right\|.$$

Hence,

$$E(s) \le c^{2p}\mathbb{E}\left[\left(1+\left\|X_s^{NV}\right\|^q \vee \left\|X_{\hat{\tau}_s}^{NV}\right\|^q\right)^{2p}\left\|X_s^{NV}-X_{\hat{\tau}_s}^{NV}\right\|^{2p}\right].$$

Applying the Cauchy-Schwarz inequality together with (4.2.6) and (4.2.7) from Lemma 4.2.2, we easily get a constant $\alpha_2 \in \mathbb{R}_+^*$ independent of N such that:

$$E(s) \le \alpha_2 h^p,$$

and we conclude that

$$\mathbb{E}\left[\sup_{t\le T}\left\|\Phi_t^N\right\|^{2p}\right] \le \alpha_1\alpha_2 T h^{3p}.$$

Now, we focus on (4.2.19) and we denote

$$\Psi_t^N = \int_0^t \left(\sigma^0\left(X_s^{NV}\right)-\sigma^0\left(\bar{X}_s^0\right)\right)ds - I_t^0.$$

To get a clearer picture, the i-th coordinate, $i \in \{1,\dots,n\}$, $\Psi_t^{i,N}$ is given by

$$\begin{aligned}
\Psi_t^{i,N} &= \int_0^t \sigma^{i0}\left(X_s^{NV}\right)-\sigma^{i0}\left(\bar{X}_s^0\right)ds - \sum_{j=1}^d\int_0^t (t\wedge\check{\tau}_s - s)\nabla\sigma^{i0}\left(X_s^{NV}\right).\,\sigma^j\left(X_s^{NV}\right)dW_s^j \\
&-\frac{1}{2}\int_0^t (t\wedge\check{\tau}_s - s)\nabla\sigma^{i0}\left(X_{\hat{\tau}_s}^{NV}\right).\,\sigma^0\left(X_{\hat{\tau}_s}^{NV}\right)ds \\
&-\frac{1}{2}\sum_{j=1}^d\int_0^t (t\wedge\check{\tau}_s - s)\nabla\sigma^{i0}\left(X_{\hat{\tau}_s}^{NV}\right).\left(\partial\sigma^j\sigma^j\left(X_{\hat{\tau}_s}^{NV}\right)\right)\ ds \\
&-\frac{1}{2}\sum_{j=1}^d\sum_{k=1}^n\sum_{l=1}^n\int_0^t (t\wedge\check{\tau}_s - s)\,\partial^2_{x_kx_l}\sigma^{i0}\left(X_{\hat{\tau}_s}^{NV}\right)\sigma^{kj}\left(X_{\hat{\tau}_s}^{NV}\right)\sigma^{lj}\left(X_{\hat{\tau}_s}^{NV}\right)ds.
\end{aligned}$$

Using once again the integration by parts formula, since for all $s\in[0,T]$, $X_{\hat{\tau}_s}^{NV} = \bar{X}_{\hat{\tau}_s}^0$, we have

$$\int_0^t \left(\sigma^{i0}\left(X_s^{NV}\right)-\sigma^{i0}\left(\bar{X}_s^0\right)\right)ds = \int_0^t (t\wedge\check{\tau}_s - s)\,d\left(\sigma^{i0}\left(X_s^{NV}\right)-\sigma^{i0}\left(\bar{X}_s^0\right)\right).$$

Before applying Itô's formula, we recall that the dynamics of $\bar{X}^0$ is given by

$$d\bar{X}_t^0 = \frac{1}{2}\sigma^0\left(\bar{X}_t^0\right)dt,$$

and that the dynamics of X^{NV} is given by (4.2.16). Since for all $t\in[0,T]$ and all $s\le t$,

$$|t\wedge\check{\tau}_s - s| \ \le h,$$

we rewrite the dynamics of X^{NV} as the sum of its dominant contribution and a remainder term with strong order $1/2$. Using (4.2.16) together with $b = \sigma^0 + \frac{1}{2}\sum_{j=1}^d \partial\sigma^j\sigma^j$, we obtain

$$dX_t^{NV} = \frac{1}{2}\sigma^0\left(X_t^{NV}\right)dt + \sum_{j=1}^d\sigma^j\left(X_t^{NV}\right)dW_t^j + \frac{1}{2}\sum_{j=1}^d\partial\sigma^j\sigma^j\left(X_t^{NV}\right)dt + \frac{1}{2}\sigma^0\left(\bar{X}_t^0\right)dt + d\vartheta_t$$

where

$$d\vartheta_t = \frac{1}{2}\sum_{j=1}^{d} \Delta t \left[\sigma^j, \sigma^0\right]\left(X_t^{NV}\right) dW_t^j + \frac{1}{2}\sum_{j=1}^{d} \Delta W_t^j \partial\sigma^j\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right) dt$$
$$+ \frac{1}{4}\Delta t\left(\partial\sigma^0\sigma^0 + \sum_{j=1}^{d}\left(\partial^2\sigma^0 \odot \sigma^j + \partial\sigma^0\partial\sigma^j\right)\sigma^j\right)\left(X_{\hat{\tau}_t}^{NV}\right) dt + d\bar{Q}_t^{1,N}.$$

Notice that ϑ is a term with strong order 1 since we have already proved (4.2.20). Applying Itô's formula we get

$$\begin{aligned}
\int_0^t \sigma^{i0}\left(X_s^{NV}\right) - \sigma^{i0}\left(\bar{X}_s^0\right) ds = {} & \frac{1}{2}\int_0^t (t\wedge\check{\tau}_s - s)\nabla\sigma^{i0}\left(X_s^{NV}\right).\,\sigma^0\left(X_s^{NV}\right) ds \\
& + \sum_{j=1}^{d}\int_0^t (t\wedge\check{\tau}_s - s)\nabla\sigma^{i0}\left(X_s^{NV}\right).\,\sigma^j\left(X_s^{NV}\right) dW_s^j \\
& + \frac{1}{2}\sum_{j=1}^{d}\int_0^t (t\wedge\check{\tau}_s - s)\nabla\sigma^{i0}\left(X_s^{NV}\right).\,\partial\sigma^j\sigma^j\left(X_s^{NV}\right) ds \\
& + \frac{1}{2}\int_0^t (t\wedge\check{\tau}_s - s)\left(\nabla\sigma^{i0}\left(X_s^{NV}\right) - \nabla\sigma^{i0}\left(\bar{X}_s^0\right)\right).\,\sigma^0\left(\bar{X}_s^0\right) ds \\
& + \frac{1}{2}\sum_{j=1}^{d}\sum_{k=1}^{n}\sum_{l=1}^{n}\int_0^t (t\wedge\check{\tau}_s - s)\,\partial^2_{x_kx_l}\sigma^{i0}\sigma^{kj}\sigma^{lj}\left(X_s^{NV}\right) ds \\
& + \int_0^t (t\wedge\check{\tau}_s - s)\nabla\sigma^{i0}\left(X_s^{NV}\right).\,d\vartheta_s \\
& + \frac{1}{2}\sum_{k=1}^{n}\sum_{l=1}^{n}\int_0^t (t\wedge\check{\tau}_s - s)\,\partial^2_{x_kx_l}\sigma^{i0}\left(X_s^{NV}\right) d\langle\vartheta^k, \vartheta^l\rangle_s \\
& + \frac{1}{2}\sum_{j=1}^{d}\sum_{k=1}^{n}\sum_{l=1}^{n}\int_0^t (t\wedge\check{\tau}_s - s)\,\partial^2_{x_kx_l}\sigma^{i0}\sigma^{lj}\left(X_s^{NV}\right) d\langle\vartheta^k, W^j\rangle_s.
\end{aligned}$$

Then, it follows that

$$\begin{aligned}
\Psi_t^{i,N} = {} & \frac{1}{2}\int_0^t (t\wedge\check{\tau}_s - s)\left(\nabla\sigma^{i0}\left(X_s^{NV}\right).\,\sigma^0\left(X_s^{NV}\right) - \nabla\sigma^{i0}\left(X_{\hat{\tau}_s}^{NV}\right).\,\sigma^0\left(X_{\hat{\tau}_s}^{NV}\right)\right) ds \\
& + \frac{1}{2}\sum_{j=1}^{d}\int_0^t (t\wedge\check{\tau}_s - s)\left(\nabla\sigma^{i0}\left(X_s^{NV}\right).\,\partial\sigma^j\sigma^j\left(X_s^{NV}\right) - \nabla\sigma^{i0}\left(X_{\hat{\tau}_s}^{NV}\right).\,\partial\sigma^j\sigma^j\left(X_{\hat{\tau}_s}^{NV}\right)\right) ds \\
& + \frac{1}{2}\int_0^t (t\wedge\check{\tau}_s - s)\left(\nabla\sigma^{i0}\left(X_s^{NV}\right) - \nabla\sigma^{i0}\left(\bar{X}_s^0\right)\right).\,\sigma^0\left(\bar{X}_s^0\right) ds \\
& + \frac{1}{2}\sum_{j=1}^{d}\sum_{k=1}^{n}\sum_{l=1}^{n}\int_0^t (t\wedge\check{\tau}_s - s)\left(\partial^2_{x_kx_l}\sigma^{i0}\sigma^{kj}\sigma^{lj}\left(X_s^{NV}\right) - \partial^2_{x_kx_l}\sigma^{i0}\sigma^{kj}\sigma^{lj}\left(X_{\hat{\tau}_s}^{NV}\right)\right) ds \\
& + \int_0^t (t\wedge\check{\tau}_s - s)\nabla\sigma^{i0}\left(X_s^{NV}\right).\,d\vartheta_s \\
& + \frac{1}{2}\sum_{k=1}^{n}\sum_{l=1}^{n}\int_0^t (t\wedge\check{\tau}_s - s)\,\partial^2_{x_kx_l}\sigma^{i0}\left(X_s^{NV}\right) d\langle\vartheta^k, \vartheta^l\rangle_s \\
& + \frac{1}{2}\sum_{j=1}^{d}\sum_{k=1}^{n}\sum_{l=1}^{n}\int_0^t (t\wedge\check{\tau}_s - s)\,\partial^2_{x_kx_l}\sigma^{i0}\sigma^{lj}\left(X_s^{NV}\right) d\langle\vartheta^k, W^j\rangle_s.
\end{aligned}$$

Now, it is easy to see that (4.2.19) is a straightforward consequence of Lemma 4.2.2 and the regularity assumption on the vector fields σ^j for $j \in \{0, \ldots, d\}$. ∎

We easily obtain the following decomposition of X^{NV}

$$\begin{aligned} X_t^{NV} &= \int_0^t b\left(X_s^{NV}\right) ds \\ &+ \sum_{j=1}^{d} \int_0^t \sigma^j \left(X_s^{NV}\right) dW_s^j + \frac{1}{2}\sum_{j=1}^{d}\int_0^t \left(\Delta s - (t \wedge \check{\tau}_s - s)\right) \left[\sigma^j, \sigma^0\right] \left(X_s^{NV}\right) dW_s^j \\ &+ \frac{1}{4}\int_0^t \left(\Delta s - (t \wedge \check{\tau}_s - s)\right) \left(\partial\sigma^0\sigma^0 + \sum_{j=1}^{d}\left(\partial^2\sigma^0 \odot \sigma^j + \partial\sigma^0\partial\sigma^j\right)\sigma^j\right) \left(X_{\hat{\tau}_s}^{NV}\right) ds \\ &+ \bar{Q}_t^{2,N} + \bar{Q}_t^{1,N}, \end{aligned} \tag{4.2.21}$$

where

$$\bar{Q}_t^{2,N} = \frac{1}{2}\left(\int_0^t \left(\sigma^0\left(X_s^{NV}\right) - \sigma^0\left(\bar{X}_s^0\right)\right) ds - I_t^0\right) + \frac{1}{2}\sum_{j=1}^{d}\left(\int_0^t \Delta W_s^j \partial\sigma^j\sigma^0\left(X_{\hat{\tau}_s}^{NV}\right) ds - I_t^j\right)$$

is a remainder term with strong order 3/2 from Proposition 4.2.6. Actually, in (4.2.21), the term

$$\int_0^t \left(\Delta s - (t \wedge \check{\tau}_s - s)\right) \left(\partial\sigma^0\sigma^0 + \sum_{j=1}^{d}\left(\partial^2\sigma^0 \odot \sigma^j + \partial\sigma^0\partial\sigma^j\right)\sigma^j\right) \left(X_{\hat{\tau}_s}^{NV}\right) ds \tag{4.2.22}$$

is null. To lighten up this expression, we denote $F = \partial\sigma^0\sigma^0 + \sum_{j=1}^{d}\left(\partial^2\sigma^0 \odot \sigma^j + \partial\sigma^0\partial\sigma^j\right)\sigma^j$. Then (4.2.22) becomes:

$$\begin{aligned} \int_0^t \left(\Delta s - (t \wedge \check{\tau}_s - s)\right) F\left(X_{\hat{\tau}_s}^{NV}\right) ds &= \sum_{k=0}^{\lfloor \frac{Nt}{T} \rfloor - 1} F\left(X_{t_k}^{NV}\right) \int_{t_k}^{t_{k+1}} \left(2s - t_k - t_{k+1}\right) ds \\ &+ F\left(X_{\hat{\tau}_t}^{NV}\right) \int_{\hat{\tau}_t}^{t} \left(2s - \hat{\tau}_t - t\right) ds \\ &= 0. \end{aligned}$$

Therefore (4.2.21) can be simplified to our final decomposition:

$$X_t^{NV} = \int_0^t b\left(X_s^{NV}\right) ds + \sum_{j=1}^{d}\int_0^t \sigma^j\left(X_s^{NV}\right) dW_s^j + \bar{J}_t^N + \bar{Q}_t^N, \tag{4.2.23}$$

where

$$\bar{J}_t^N = \frac{1}{2}\sum_{j=1}^{d}\int_0^t \left(\Delta s - (t \wedge \check{\tau}_s - s)\right)\left[\sigma^j, \sigma^0\right]\left(X_s^{NV}\right) dW_s^j,$$

and

$$\bar{Q}_t^N = \bar{Q}_t^{2,N} + \bar{Q}_t^{1,N}.$$

According to Propositions 4.2.5 and 4.2.6, $\bar{Q}^N$ is a remainder term with strong order 3/2.

Proposition 4.2.7 *Let $p \geq 1$. Under the assumptions of Theorem 4.2.1, there exists a constant $C_5 \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$,*

$$\mathbb{E}\left[\sup_{t \leq T}\left\|\bar{Q}_t^N\right\|^{2p}\right] \leq C_5 h^{3p}.$$

In the following proposition we state the continuous version of Theorem 4.1.1.

Proposition 4.2.8 *Let $p \geq 1$. Under the assumptions of Theorem 4.2.1, there exists a constant $C''_{NV} \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$,*

$$\mathbb{E}\left[\sup_{t\leq T}\left\|X_t - X_t^{NV}\right\|^{2p}\right] \leq C''_{NV} h^{2p}.$$

Proof : Let $p \in [1,+\infty)$, $t \in [0,T]$ and $s \in [0,t]$. Subtracting (4.2.23) from (4.1.1), we can evaluate the difference between the exact solution and the scheme:

$$X_s - X_s^{NV} = \int_0^s \left(b\left(X_u\right) - b\left(X_u^{NV}\right)\right) du + \sum_{j=1}^d \int_0^s \left(\sigma^j\left(X_u\right) - \sigma^j\left(X_u^{NV}\right)\right) dW_u^j - \bar{J}_s^N - \bar{Q}_s^N.$$

Using a convexity inequality, taking the expectation of the supremum and applying the Burkholder-Davis-Gundy inequality, we get a constant $\alpha_0 \in \mathbb{R}_+^*$ independent of N, such that

$$\begin{aligned}\mathbb{E}\left[\sup_{s\leq t}\left\|X_s - X_s^{NV}\right\|^{2p}\right] \leq \alpha_0 \Bigg(& \int_0^t \mathbb{E}\left[\left\|b\left(X_u\right) - b\left(X_u^{NV}\right)\right\|^{2p}\right] du \\ & + \sum_{j=1}^d \int_0^t \mathbb{E}\left[\left\|\sigma^j\left(X_u\right) - \sigma^j\left(X_u^{NV}\right)\right\|^{2p}\right] du \\ & + \mathbb{E}\left[\sup_{s\leq t}\left\|\bar{J}_s^N\right\|^{2p}\right] + \mathbb{E}\left[\sup_{s\leq t}\left\|\bar{Q}_s^N\right\|^{2p}\right]\Bigg).\end{aligned}$$

By the Lipschitz assumption, we can rewrite the last inequality as follows

$$\begin{aligned}\mathbb{E}\left[\sup_{s\leq t}\left\|X_s - X_s^{NV}\right\|^{2p}\right] \leq \alpha_1 \Bigg(& \int_0^t \mathbb{E}\left[\sup_{v\leq u}\left\|X_v - X_v^{NV}\right\|^{2p}\right] du \\ & + \mathbb{E}\left[\sup_{s\leq t}\left\|\bar{J}_s^N\right\|^{2p}\right] + \mathbb{E}\left[\sup_{s\leq t}\left\|\bar{Q}_s^N\right\|^{2p}\right]\Bigg),\end{aligned} \tag{4.2.24}$$

where $\alpha_1 \in \mathbb{R}_+^*$ is a constant independent of N. On the one hand, applying the Burkholder-Davis-Gundy inequality, we obtain a constant $\beta_0 \in \mathbb{R}_+^*$ independent of N, such that

$$\mathbb{E}\left[\sup_{s\leq t}\left\|\bar{J}_s^N\right\|^{2p}\right] \leq \beta_0 \sum_{j=1}^d \int_0^t \left|\Delta s - (t \wedge \check{\tau}_s - s)\right|^{2p} \left\|\left[\sigma^j, \sigma^0\right]\left(X_s^{NV}\right)\right\|^{2p} du.$$

Since $\left[\sigma^j, \sigma^0\right] \in LIP_{loc}^{pgc}\left(\mathbb{R}^n\right)$, (4.2.6) from Lemma 4.2.2 ensures that

$$\mathbb{E}\left[\sup_{s\leq t}\left\|\bar{J}_s^N\right\|^{2p}\right] \leq \beta_1 h^{2p}, \tag{4.2.25}$$

for some constant $\beta_1 \in \mathbb{R}_+^*$ independent of N. On the other hand, from Proposition 4.2.7, we have

$$\mathbb{E}\left[\sup_{s\leq t}\left\|\bar{Q}_s^N\right\|^{2p}\right] \leq C_5 h^{3p}. \tag{4.2.26}$$

Then, combing (4.2.24), (4.2.25) and (4.2.26), we easily obtain

$$\mathbb{E}\left[\sup_{s\leq t}\left\|X_s - X_s^{NV}\right\|^{2p}\right] \leq \alpha_1 \int_0^t \mathbb{E}\left[\sup_{v\leq u}\left\|X_v - X_v^{NV}\right\|^{2p}\right] du + \alpha_1\left(\beta_1 + C_5 T^p\right) h^{2p}.$$

We conclude thanks to Grönwall's inequality. ∎

As a comparison with theorem 4.2 in Chapter 3, note that using the adapted interpolation leads to stronger assumptions, which justifies the use of the non-adapted interpolation in Chapter 3.

4.3 Proof of the stable convergence

Using (4.2.23), the normalized error process U^N can be written as:

$$U_t^N = N\left(\int_0^t \left(b\left(X_s\right) - b\left(X_s^{NV}\right)\right) ds + \sum_{j=1}^d \int_0^t \left(\sigma^j\left(X_s\right) - \sigma^j\left(X_s^{NV}\right)\right) dW_s^j\right) + J_t^N + Q_t^N, \tag{4.3.1}$$

where

$$J_t^N = -N\bar{J}_t^N = \frac{N}{2}\sum_{j=1}^d \int_0^t \left(\Delta s - \left(t \wedge \check{\tau}_s - s\right)\right)\left[\sigma^0, \sigma^j\right]\left(X_s^{NV}\right) dW_s^j,$$

and

$$Q^N = -N\bar{Q}^N.$$

As previously mentioned, by analyzing the stable convergence in law of the source term J^N, we prove the stable convergence in law of U^N. In the following, we provide a detailed proof of Theorem 4.2.1.

Proof of Theorem 4.2.1: For the reader's convenience, the proof will go through several steps.

Step 1: linearization of

$$N\int_0^t \left(b\left(X_s\right) - b\left(X_s^{NV}\right)\right) ds + N\sum_{j=1}^d \int_0^t \left(\sigma^j\left(X_s\right) - \sigma^j\left(X_s^{NV}\right)\right) dW_s^j.$$

Let $j \in \{1,\ldots,d\}$ and $i \in \{1,\ldots,n\}$, by the mean value theorem, we get

$$\sigma^{ij}\left(X_s\right) - \sigma^{ij}\left(X_s^{NV}\right) = \nabla\sigma^{ij}\left(\zeta_s^{ij}\right).\left(X_s - X_s^{NV}\right),$$

where $\zeta_s^{ij} = \alpha_s^{ij} X_s + \left(1-\alpha_s^{ij}\right) X_s^{NV}$ for some $\alpha_s^{ij} \in [0,1]$. Using a compact matrix notation, we can write

$$\sigma^j\left(X_s\right) - \sigma^j\left(X_s^{NV}\right) = \partial\sigma_s^{j,N}\left(X_s - X_s^{NV}\right),$$

where

$$\left(\partial\sigma_s^{j,N}\right)_{i,m} = \partial_{x_m}\sigma^{ij}\left(\zeta_s^{ij}\right).$$

In the same way,

$$b\left(X_s\right) - b\left(X_s^{NV}\right) = \partial b_s^N\left(X_s - X_s^{NV}\right),$$

where

$$\left(\partial b_s^N\right)_{i,m} = \partial_{x_m} b^i\left(\zeta_s^{i0}\right),$$

with $\zeta_s^{i0} = \alpha_s^{i0} X_s + \left(1 - \alpha_s^{i0}\right) X_s^{NV}$ for some $\alpha_s^{i0} \in [0,1]$. Then, it follows that

$$N \int_0^t \left(b\left(X_s\right) - b\left(X_s^{NV}\right)\right) ds = \int_0^t \partial b_s^N U_s^N ds,$$

and

$$N \sum_{j=1}^d \int_0^t \left(\sigma^j\left(X_s\right) - \sigma^j\left(X_s^{NV}\right)\right) dW_s^j = \sum_{j=1}^d \int_0^t \partial \sigma_s^{j,N} U_s^N dW_s^j.$$

Then, we can write the error process in three parts as follows:

$$U_t^N = Q_t^N + J_t^N + \left(\int_0^t \partial b_s^N U_s^N ds + \sum_{j=1}^d \int_0^t \partial \sigma_s^{j,N} U_s^N dW_s^j\right).$$

Step 2: stable convergence in law of the source term J^N.

To study the convergence of the source term J^N, we introduce the d-dimensional martingale M^N with coordinates

$$M_t^{j,N} = N \int_0^t \frac{1}{2}\left(\Delta s - \left(t \wedge \check{\tau}_s - s\right)\right) dW_s^j, \quad j \in \{1, \ldots, d\},$$

and $K^N = \left(K^{j,N} = \left[\sigma^0, \sigma^j\right]\left(X_s^{NV}\right)\right)_{1 \le j \le d}$ with values in $\mathbb{R}^n \otimes \mathbb{R}^d$, so that

$$J_t^N = \sum_{j=1}^d \int_0^t K_s^{j,N} dM_s^{j,N}.$$

Step 2.1: stable convergence in law of M^N.

By virtue of Theorem 2.3 in Chapter 3, to study the limit in law of M^N, we check for all $t \in [0,T], j, m, k \in \{1, \ldots, d\}$

- the convergence in probability of $\langle M^{j,N}, M^{m,N}\rangle_t$, as N goes to infinity,
- the convergence in probability to 0 of $\langle M^{j,N}, W^k\rangle_t$, as N goes to infinity.

If $j \neq m$ and $j \neq k$ then, obviously, $\langle M^{j,N}, M^{m,N}\rangle_t = \langle M^{j,N}, W^k\rangle_t = 0$. Now, if $j = m$, a straightforward calculation gives us:

$$\begin{aligned}
\langle M^{j,N}, M^{j,N}\rangle_t &= N^2 \int_0^t \frac{1}{4}\left(\Delta s - \left(t \wedge \check{\tau}_s - s\right)\right)^2 ds \\
&= \frac{1}{4} N^2 \left(\int_0^{\hat{\tau}_t} \left(\Delta s - \left(\check{\tau}_s - s\right)\right)^2 ds + \int_{\hat{\tau}_t}^t \left(s - \hat{\tau}_t - (t - s)\right)^2 ds\right) \\
&= \frac{1}{12} N^2 \left(\lfloor \frac{Nt}{T} \rfloor \frac{T^3}{N^3} + \left(t - \hat{\tau}_t\right)^3\right) \underset{N \to +\infty}{\longrightarrow} \frac{1}{12} t T^2.
\end{aligned}$$

If $j = k$

$$\langle M^{j,N}, W^k \rangle_t = N \int_0^t \frac{1}{2} \left(\Delta s - (t \wedge \check{\tau}_s - s) \right) ds = 0.$$

Applying Theorem 2.3 in Chapter 3 we conclude that $\frac{2\sqrt{3}}{T} M^N$ stably converges in law to a standard d-dimensional Brownian motion $\tilde{B}$, independent of W.

Step 2.2: convergence in probability of K^N.

Now, it remains to prove the convergence in probability of K^N. From Proposition 4.2.8, together with the continuity assumption on $[\sigma^0, \sigma^j]$, for $j \{1, \dots, d\}$, we get the following convergence in probability

$$\sup_{t \leq T} \left\| K_t^{j,N} - \left[\sigma^0, \sigma^j \right] (X_t) \right\| \xrightarrow[N \to +\infty]{\mathbb{P}} 0.$$

Step 2.3: conclusion of the step 2.

According to Proposition 2.2 in Chapter 3, we have the following convergence:

$$\left(K^N, \frac{2\sqrt{3}}{T} M^N \right) \underset{N \to +\infty}{\overset{stably}{\Longrightarrow}} \left(\left(\left[\sigma^0, \sigma^j \right] (X) \right)_{j \in \{1, \dots, d\}}, \tilde{B} \right).$$

The convergence of $\left\langle M^N \right\rangle_T$ ensures its tightness. Then Proposition 2.4 in Chapter 3 leads us to:

$$\left(K^N, \frac{2\sqrt{3}}{T} M^N, J^N \right) \underset{N \to +\infty}{\overset{stably}{\Longrightarrow}} \left(\left(\left[\sigma^0, \sigma^j \right] (X) \right)_{j \in \{1, \dots, d\}}, \tilde{B}, J \right), \tag{4.3.2}$$

where, for $t \in [0, T]$,

$$J_t = \frac{T}{2\sqrt{3}} \sum_{j=1}^{d} \int_0^t \left[\sigma^0, \sigma^j \right] (X_s) \, d\tilde{B}_s^j.$$

Step 3: convergence of Q^N.

We easily get the following convergence in L^2 from Proposition 4.2.7:

$$\sup_{t \leq T} \left\| Q_t^N \right\| \xrightarrow[N \to +\infty]{L^2} 0. \tag{4.3.3}$$

Step 4: stable convergence in law of U^N.

We recall that $U_t^N = Q_t^N + J_t^N + \left(\int_0^t \partial b_s^N U_s^N ds + \sum_{j=1}^{d} \int_0^t \partial \sigma_s^{j,N} U_s^N dW_s^j \right)$. Thanks to (4.3.2) and (4.3.3), we conclude using Theorem 2.5 in Chapter 3 since the continuity of ∂b and $\partial \sigma^j, j \in \{1, \dots, d\}$, together with Proposition 4.2.8, ensure that

$$\sup_{t \leq T} \left\| \partial b_t^N - \partial b \left(X_t \right) \right\| \xrightarrow[N \to +\infty]{\mathbb{P}} 0,$$

and for all $j \in \{1, \ldots, d\}$,

$$\sup_{s \leq T} \left\| \partial \sigma_t^{j,N} - \partial \sigma^j \left(X_t \right) \right\| \xrightarrow[N \to +\infty]{\mathbb{P}} 0.$$

■

4.4 Appendix

This section is devoted to the proof of Propositions 4.2.3, 4.2.4 and 4.2.5. Before proving Proposition 4.2.3, in the following lemma, we give intermediate estimations of the process Y^{d+1}.

Lemma 4.4.1 *Assume that for all $j \in \{0, \ldots, d\}, \sigma^j \in \mathcal{C}^1 \left(\mathbb{R}^n, \mathbb{R}^n \right)$ with bounded first order derivatives. Then, for all $p \geq 1$, there exists a constant $C_6 \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*, t \in [0, T]$ and all $s \in [\hat{\tau}_t, t]$,*

$$\mathbb{E}\left[\left\| Y_{t,s}^{d+1} - Id_n \right\|^{2p} \right] \leq C_6 \left(\Delta t \right)^{2p}, \tag{4.4.1}$$

$$\mathbb{E}\left[\left\| Y_{t,t}^{d+1} \left(Y_{t,s}^{d+1} \right)^{-1} - Id_n \right\|^{2p} \right] \leq C_6 \left(\Delta t \right)^{2p}. \tag{4.4.2}$$

Proof : Let $t \in [0, T]$ and $s \in [\hat{\tau}_t, t]$, we recall that $Y_{t,s}^{d+1}$ is the solution, at time s, of the following ODE:

$$\begin{cases} \dfrac{d}{du} \Xi_{t,\hat{\tau}_t,u} = \dfrac{1}{2} \partial \sigma^0 \left(\bar{X}_{t,u}^{d+1} \right) \Xi_{t,\hat{\tau}_t,u}, u \in [\hat{\tau}_t, t] \\ \Xi_{t,\hat{\tau}_t,\hat{\tau}_t} = Id_n, \end{cases} \tag{4.4.3}$$

and $Y_{t,t}^{d+1} \left(Y_{t,s}^{d+1} \right)^{-1}$ is the solution, at time t, of the same ODE, but with the initial condition:

$$\begin{cases} \dfrac{d}{du} \Xi_{t,s,u} = \dfrac{1}{2} \partial \sigma^0 \left(\bar{X}_{t,u}^{d+1} \right) \Xi_{t,\hat{\tau}_t,u}, u \in [s, t] \\ \Xi_{t,s,s} = Id_n. \end{cases}$$

Therefore, we deduce (4.4.2) similarly to (4.4.1). Since $\partial \sigma^0$ is bounded, there exists a constant $c \in \mathbb{R}_+^*$, independent of N, t and s, such that

$$\left\| Y_{t,s}^{d+1} \right\|^{2p} \leq \exp\left(c \left(s - \hat{\tau}_t \right) \right) \leq \exp\left(cT \right), \tag{4.4.4}$$

Writing $Y_{t,s}^{d+1}$ in integral form, we have

$$Y_{t,s}^{d+1} = Id_n + \frac{1}{2} \int_{\hat{\tau}_t}^{s} \partial \sigma^0 \left(\bar{X}_{t,u}^{d+1} \right) Y_{t,u}^{d+1} du. \tag{4.4.5}$$

Hence, using a convexity inequality,

$$\mathbb{E}\left[\left\| Y_{t,s}^{d+1} - Id_n \right\|^{2p} \right] \leq \frac{1}{2^{2p}} \exp\left(cT \right) \left(\Delta t \right)^{2p-1} \int_{\hat{\tau}_t}^{s} \left\| \partial \sigma^0 \left(\bar{X}_{t,u}^{d+1} \right) \right\|^{2p} du.$$

Since $\partial \sigma^0$ is bounded we easily get (4.4.1). ■

Proof of Proposition 4.2.3:

Proof of (4.2.12):

To prove (4.2.12), we only need to assume that

- for all $j \in \{0, \dots, d\}, \sigma^j \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first order derivatives,
- $\sum\limits_{j=1}^{d} \partial\sigma^j \sigma^j$ is a Lipschitz continuous function,
- $\partial\sigma^0 \in LIP_{loc}^{pgc}(\mathbb{R}^n \otimes \mathbb{R}^n)$.

Let $t \in [0, T]$, $s \in [\hat{\tau}_t, t]$, $p \geq 1$, and

$$\Theta_{t,s}^{d+1} = Y_{t,s}^{d+1} - Id_n - \frac{1}{2}\Delta s \partial\sigma^0\left(X_t^{NV}\right). \tag{4.4.6}$$

Adding and subtracting $\frac{1}{2}\int_{\hat{\tau}_t}^{s} \partial\sigma^0\left(X_t^{NV}\right) Y_{t,u}^{d+1} du$, we get

$$\Theta_{t,s}^{d+1} = \frac{1}{2}\int_{\hat{\tau}_s}^{s}\left(\partial\sigma^0\left(\bar{X}_{t,u}^{d+1}\right) - \partial\sigma^0\left(X_t^{NV}\right)\right) Y_{t,u}^{d+1} du + \frac{1}{2}\int_{\hat{\tau}_s}^{s} \partial\sigma^0\left(X_t^{NV}\right)\left(Y_{t,u}^{d+1} - Id_n\right) du.$$

Combining a convexity inequality, the Cauchy-Schwarz inequality and (4.4.4) from the above proof of Lemma 4.4.1 we obtain a constant $\alpha_0 \in \mathbb{R}_+^*$, independent of N, such that

$$\begin{aligned}\mathbb{E}\left[\left\|\Theta_{t,s}^{d+1}\right\|^{2p}\right] \leq \alpha_0 (\Delta t)^{2p-1} \Bigg(&\int_{\hat{\tau}_s}^{s} \mathbb{E}^{\frac{1}{2}}\left[\left\|\partial\sigma^0\left(\bar{X}_{t,u}^{d+1}\right) - \partial\sigma^0\left(X_t^{NV}\right)\right\|^{4p}\right] du \\ &+ \int_{\hat{\tau}_s}^{s}\left(\mathbb{E}\left[\left\|\partial\sigma^0\left(X_t^{NV}\right)\right\|^{4p}\right]\mathbb{E}\left[\left\|Y_{t,u}^{d+1} - Id_n\right\|^{4p}\right]\right)^{\frac{1}{2}} du\Bigg).\end{aligned}$$

Since $\partial\sigma^0$ is bounded and locally Lipschitz with polynomially growing Lipschitz constant, there exists a constant $c \in \mathbb{R}_+^*$ and $q \in \mathbb{N}$, independent of N, t and s, such that

$$\begin{aligned}\mathbb{E}\left[\left\|\Theta_{t,s}^{d+1}\right\|^{2p}\right] \leq c (\Delta t)^{2p-1} \Bigg(&\int_{\hat{\tau}_s}^{s} \mathbb{E}^{\frac{1}{2}}\left[\left(1 + \left\|\bar{X}_{t,u}^{d+1}\right\|^{q} \vee \left\|X_t^{NV}\right\|^{q}\right)\left\|\bar{X}_{t,u}^{d+1} - X_t^{NV}\right\|^{4p}\right] du \\ &+ \int_{\hat{\tau}_s}^{s}\left(\mathbb{E}\left[\left\|Y_{t,u}^{d+1} - Id_n\right\|^{4p}\right]\right)^{\frac{1}{2}} du\Bigg).\end{aligned}$$

Applying once again the Cauchy-Schwarz inequality, and using (4.2.6) and (4.2.10) from Lemma 4.2.2, we get a constant $\alpha_1 \in \mathbb{R}_+^*$, independent of N, t and s, such that

$$\mathbb{E}\left[\left\|\Theta_{t,s}^{d+1}\right\|^{2p}\right] \leq \alpha_1 (\Delta t)^{2p-1}\int_{\hat{\tau}_s}^{s}\left(\left(\mathbb{E}\left[\left\|\bar{X}_{t,u}^{d+1} - X_t^{NV}\right\|^{8p}\right]\right)^{\frac{1}{4}} + \left(\mathbb{E}\left[\left\|Y_{t,u}^{d+1} - Id_n\right\|^{8p}\right]\right)^{\frac{1}{4}}\right) du.$$

We conclude using (4.2.10) from Lemma 4.2.2, and (4.4.1) from Lemma 4.4.1.

Proof of (4.2.13):

To prove (4.2.13), we only need to assume that

- for all $j \in \{1, \ldots, d\}, \sigma^j \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first order derivatives,
- $\sum\limits_{j=1}^{d} \partial\sigma^j \sigma^j$ is a Lipschitz continuous function,
- $\sigma^0 \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first order derivatives and $\partial^2\sigma^0 \in LIP^{pgc}_{loc}(\mathbb{R}^n \otimes \mathbb{R}^n)$.

Let $t \in [0, T]$, $j \in \{1, \ldots, d\}$, $p \geq 1$, and

$$\theta_t^{d+1,j} = Z_t^{d+1,j} - \frac{1}{2}\Delta t \left(\partial^2\sigma^0 \odot \sigma^j\right)\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right),$$

and we recall that

$$Z_t^{d+1,j} = \frac{1}{2}\int_{\hat{\tau}_t}^t Y_{t,t}^{d+1}\left(Y_{t,s}^{d+1}\right)^{-1}\left(\partial^2\sigma^0\left(\bar{X}_{t,s}^{d+1}\right) \odot \left(Y_{t,s}^{d+1}\sigma^j\left(\bar{X}_{t,t}\right)\right)\right)Y_{t,s}^{d+1}\sigma^j\left(\bar{X}_{t,t}\right)\, ds.$$

Therefore,

$$\theta_t^{d+1,j} = \frac{1}{2}\int_{\hat{\tau}_t}^t \Bigg(Y_{t,t}^{d+1}\left(Y_{t,s}^{d+1}\right)^{-1}\left(\partial^2\sigma^0\left(\bar{X}_{t,s}\right) \odot Y_{t,s}^{d+1}\sigma^j\left(\bar{X}_{t,t}\right)\right)Y_{t,s}^{d+1}\sigma^j\left(\bar{X}_{t,t}\right)$$
$$- \left(\partial^2\sigma^0 \odot \sigma^j\right)\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\Bigg)ds.$$

Adding and subtracting some appropriate terms, we obtain:

$$\begin{aligned}
\theta_t^{d+1,j} &= \frac{1}{2}\int_{\hat{\tau}_t}^t Y_{t,t}^{d+1}\left(Y_{t,s}^{d+1}\right)^{-1}\left(\partial^2\sigma^0\left(\bar{X}_{t,s}\right) \odot Y_{t,s}^{d+1}\sigma^j\left(\bar{X}_{t,t}\right)\right)\left(Y_{t,s}^{d+1} - Id_n\right)\sigma^j\left(\bar{X}_{t,t}\right)ds \\
&+ \frac{1}{2}\int_{\hat{\tau}_t}^t \left(Y_{t,t}^{d+1}\left(Y_{t,s}^{d+1}\right)^{-1} - Id_n\right)\left(\partial^2\sigma^0\left(\bar{X}_{t,s}\right) \odot Y_{t,s}^{d+1}\sigma^j\left(\bar{X}_{t,t}\right)\right)\sigma^j\left(\bar{X}_{t,s}\right)ds \\
&+ \frac{1}{2}\int_{\hat{\tau}_t}^t \left(\partial^2\sigma^0\left(\bar{X}_{t,s}\right) \odot \left(Y_{t,s}^{d+1} - Id_n\right)\sigma^j\left(\bar{X}_{t,t}\right)\right)\sigma^j\left(\bar{X}_{t,t}\right)ds \\
&+ \frac{1}{2}\int_{\hat{\tau}_t}^t \left(\left(\partial^2\sigma^0\left(\bar{X}_{t,s}\right) - \partial^2\sigma^0\left(X_t^{NV}\right)\right) \odot \sigma^j\left(\bar{X}_{t,t}\right)\right)\sigma^j\left(\bar{X}_{t,t}\right)ds \\
&+ \frac{1}{2}\int_{\hat{\tau}_t}^t \left(\partial^2\sigma^0\left(X_t^{NV}\right) \odot \left(\sigma^j\left(\bar{X}_{t,t}\right) - \sigma^j\left(X_t^{NV}\right)\right)\right)\sigma^j\left(\bar{X}_{t,t}\right)ds \\
&+ \frac{1}{2}\int_{\hat{\tau}_t}^t \left(\partial^2\sigma^0\left(X_t^{NV}\right) \odot \sigma^j\left(X_t^{NV}\right)\right)\left(\sigma^j\left(\bar{X}_{t,t}\right) - \sigma^j\left(X_t^{NV}\right)\right)ds \\
&+ \frac{1}{2}\int_{\hat{\tau}_t}^t \left(\partial^2\sigma^0 \odot \sigma^j\right)\sigma^j\left(X_t^{NV}\right) - \left(\partial^2\sigma^0 \odot \sigma^j\right)\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)ds.
\end{aligned}$$

Note that since $\partial^2\sigma^0 \in LIP^{pgc}_{loc}(\mathbb{R}^n \otimes \mathbb{R}^n \otimes \mathbb{R}^n)$ and σ^j is Lipschitz continuous, then $\left(\partial^2\sigma^0 \odot \sigma^j\right)\sigma^j \in LIP^{pgc}_{loc}(\mathbb{R}^n)$. We easily get the desired result by combing a convexity inequality, the Cauchy-Schwarz inequality, (4.4.1), (4.4.2) and (4.4.4) from Lemma 4.4.1, (4.2.6) (4.2.7) and (4.2.9) from Lemma 4.2.2

Proof of (4.2.14):

To prove (4.2.14), we assume that

- for all $j \in \{0, \ldots, d\}, \sigma^j \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first order derivatives,
- $\sum\limits_{j=1}^{d} \partial\sigma^j \sigma^j$ is a Lipschitz continuous function,
- for all $j \in \{1, \ldots, d\}, \partial\sigma^j \in LIP_{loc}^{pgc}(\mathbb{R}^n \otimes \mathbb{R}^n)$,

and that the commutativity condition ($\mathcal{C}$) holds. Let $t \in [0, T]$, $j \in \{1, \ldots, d\}$, $i \in \{1, \ldots, n\}$, $p \geq 1$, and

$$\theta_t^j = \sigma^j\left(\bar{X}_{t,t}\right) - \sigma^j\left(X_t^{NV}\right) + \frac{1}{2}\Delta t \partial\sigma^j \sigma^0\left(X_t^{NV}\right).$$

The i-th component of θ^j is given by:

$$\theta_t^{ij} = \sigma^{ij}\left(\bar{X}_{t,t}\right) - \sigma^{ij}\left(X_t^{NV}\right) + \frac{1}{2}\Delta t \nabla\sigma^{ij}\left(X_t^{NV}\right) . \, \sigma^0\left(X_t^{NV}\right).$$

By the mean value theorem

$$\begin{aligned}\sigma^{ij}\left(\bar{X}_{t,t}\right) &= \sigma^{ij}\left(X_t^{NV}\right) + \nabla\sigma^{ij}\left(X_t^{NV}\right) . \left(\bar{X}_{t,t} - X_t^{NV}\right) \\ &\quad + \left(\nabla\sigma^{ij}\left(\xi_t^{ij}\right) - \nabla\sigma^{ij}\left(X_t^{NV}\right)\right) . \left(\bar{X}_{t,t} - X_t^{NV}\right),\end{aligned}$$

where $\xi_t^{ij} = X_t^{NV} + \alpha_t^{ij}\left(\bar{X}_{t,t} - X_t^{NV}\right)$ for some $\alpha_t^{ij} \in [0, 1]$. Then, it follows that

$$\begin{aligned}\theta_t^{ij} &= \nabla\sigma^{ij}\left(X_t^{NV}\right) . \left(\bar{X}_{t,t} - X_t^{NV}\right) \\ &\quad + \left(\nabla\sigma^{ij}\left(\xi_t^{ij}\right) - \nabla\sigma^{ij}\left(X_t^{NV}\right)\right) . \left(\bar{X}_{t,t} - X_t^{NV}\right) + \frac{1}{2}\Delta t \nabla\sigma^{ij}\left(X_t^{NV}\right) . \, \sigma^0\left(X_t^{NV}\right).\end{aligned}$$

Moreover, since

$$X_t^{NV} = \exp\left(\frac{1}{2}\Delta t \sigma^0\right) \bar{X}_{t,t},$$

we have that

$$\bar{X}_{t,t} - X_t^{NV} = -\frac{1}{2}\int_{\hat{\tau}_t}^{t} \sigma^0\left(\bar{X}_{t,s}^{d+1}\right) ds.$$

Therefore, we obtain

$$\begin{aligned}\theta_t^{ij} &= \frac{1}{2}\int_{\hat{\tau}_t}^{t} \nabla\sigma^{ij}\left(X_t^{NV}\right) . \left(\sigma^0\left(X_t^{NV}\right) - \sigma^0\left(\bar{X}_{t,s}^{d+1}\right)\right) ds \\ &\quad + \left(\nabla\sigma^{ij}\left(\xi_t^{ij}\right) - \nabla\sigma^{ij}\left(X_t^{NV}\right)\right) . \left(\bar{X}_{t,t} - X_t^{NV}\right).\end{aligned}$$

Using a convexity inequality

$$\begin{aligned}\mathbb{E}\left[\left|\theta_t^{ij}\right|^{2p}\right] &\leq 2^{2p-1}\Bigg(\frac{1}{2^{2p}}(\Delta t)^{2p-1}\int_{\hat{\tau}_t}^{t} \mathbb{E}\left[\left|\nabla\sigma^{ij}\left(X_t^{NV}\right) . \left(\sigma^0\left(X_t^{NV}\right) - \sigma^0\left(\bar{X}_{t,s}^{d+1}\right)\right)\right|^{2p}\right] ds \\ &\quad + \mathbb{E}\left[\left|\left(\nabla\sigma^{ij}\left(\xi_t^{ij}\right) - \nabla\sigma^{ij}\left(X_t^{NV}\right)\right) . \left(\bar{X}_{t,t} - X_t^{NV}\right)\right|^{2p}\right]\Bigg).\end{aligned}$$

Applying the Cauchy-Schwarz inequality, we get

$$\begin{aligned}\mathbb{E}\left[\left|\theta_t^{ij}\right|^{2p}\right] &\leq \frac{2^{2p-1}}{2^{2p}}(\Delta t)^{2p-1}\int_{\hat{\tau}_t}^{t} \left(\mathbb{E}\left[\left\|\nabla\sigma^{ij}\left(X_t^{NV}\right)\right\|^{4p}\right] \mathbb{E}\left[\left\|\sigma^0\left(X_t^{NV}\right) - \sigma^0\left(\bar{X}_{t,s}^{d+1}\right)\right\|^{4p}\right]\right)^{\frac{1}{2}} ds \\ &\quad + 2^{2p-1}\left(\mathbb{E}\left[\left\|\nabla\sigma^{ij}\left(\xi_t^{ij}\right) - \nabla\sigma^{ij}\left(X_t^{NV}\right)\right\|^{4p}\right] \mathbb{E}\left[\left\|\bar{X}_{t,t} - X_t^{NV}\right\|^{4p}\right]\right)^{\frac{1}{2}}\Bigg).\end{aligned}$$

Since σ^0 is Lipschitz continuous and $\partial\sigma^j$ is locally Lipschitz with polynomially growing Lipschitz constant, using (4.2.6) from Lemma 4.2.2 , we easily get a constant $\alpha_2 \in \mathbb{R}_+^*$ independent of N and t such that

$$\mathbb{E}\left[\left|\theta_t^{ij}\right|^{2p}\right] \leq \alpha_2\left((\Delta t)^{2p-1}\int_{\hat{\tau}_t}^t \left(\mathbb{E}\left[\left\|X_t^{NV} - \bar{X}_{t,s}^{d+1}\right\|^{4p}\right]\right)^{\frac{1}{2}} ds + \mathbb{E}\left[\left\|\bar{X}_{t,t} - X_t^{NV}\right\|^{4p}\right]\right).$$

Applying (4.2.10) and (4.2.11) from Lemma 4.2.2 , we obtain

$$\mathbb{E}\left[\left|\theta_t^{ij}\right|^{2p}\right] \leq \alpha_2\left(\sqrt{C_0} + C_0\right)(\Delta t)^{4p}. \tag{4.4.7}$$

Proof of (4.2.15):

The proof of (4.2.15) is very similar to (4.2.14). However, we need to assume further that for all $j \in \{1,\dots,d\}, \sigma^j \in \mathcal{C}^2(\mathbb{R}^n,\mathbb{R}^n)$ and $\partial^2\sigma^j \in LIP_{loc}^{pgc}(\mathbb{R}^n\otimes\mathbb{R}^n\otimes\mathbb{R}^n)$. Let $t \in [0,T]$, $j \in \{1,\dots,d\}$, $p \geq 1$, and

$$\bar{\theta}_t^j = \partial\sigma^j\sigma^j\left(\bar{X}_{t,t}\right) - \partial\sigma^j\sigma^j\left(X_t^{NV}\right) + \frac{1}{2}\Delta t F_j\left(X_{\hat{\tau}_t}^{NV}\right).$$

where

$$F_j = \left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0.$$

Since $\partial^2\sigma^j \in LIP_{loc}^{pgc}(\mathbb{R}^n\otimes\mathbb{R}^n\otimes\mathbb{R}^n)$, $\partial\sigma^j \in LIP_{loc}^{pgc}(\mathbb{R}^n\otimes\mathbb{R}^n)$ and σ^0 is Lipschitz continuous, then $F_j \in LIP_{loc}^{pgc}(\mathbb{R}^n)$. Hence, similarly to (4.4.7), there exists $\bar{\alpha}_2 \in \mathbb{R}_+^*$, independent of N and t, such that

$$\mathbb{E}\left[\left\|\partial\sigma^j\sigma^j\left(\bar{X}_{t,t}\right) - \partial\sigma^j\sigma^j\left(X_t^{NV}\right) + \frac{1}{2}\Delta t F_j\left(X_t^{NV}\right)\right\|^{2p}\right] \leq \bar{\alpha}_2(\Delta t)^{4p}. \tag{4.4.8}$$

Then, using a convexity inequality

$$\begin{aligned}\mathbb{E}\left[\left\|\bar{\theta}_t^j\right\|^{2p}\right] &\leq 2^{2p-1}\Bigg(\mathbb{E}\left[\left\|\partial\sigma^j\sigma^j\left(\bar{X}_{t,t}\right) - \partial\sigma^j\sigma^j\left(X_t^{NV}\right) + \frac{1}{2}\Delta t F_j\left(X_t^{NV}\right)\right\|^{2p}\right] \\ &\quad + \mathbb{E}\left[\left\|\frac{1}{2}\Delta t\left(F_j\left(X_{\hat{\tau}_t}^{NV}\right) - F_j\left(X_t^{NV}\right)\right)\right\|^{2p}\right]\Bigg) \\ &\leq 2^{2p-1}\left(\bar{\alpha}_2(\Delta t)^{4p} + \frac{1}{2^{2p}}(\Delta t)^{2p}\,\mathbb{E}\left[\left\|F_j\left(X_{\hat{\tau}_t}^{NV}\right) - F_j\left(X_t^{NV}\right)\right\|^{2p}\right]\right).\end{aligned}$$

Since F_j is locally Lipschitz with polynomially growing Lipschitz constant, we conclude using (4.2.7) from Lemma 4.2.2. ∎

Before proving Proposition 4.2.4, we introduce some intermediate processes. We define for $j \in \{1,\dots,d\}$, $t \in [0,T]$ and $s \in [\hat{\tau}_t, t]$

$$\bar{X}_{t,s}^j = h_j\left(\frac{\Delta t}{2}, \Delta W_t^1, \dots, \Delta W_t^{j-1}, \Delta W_s^j; X_{\hat{\tau}_t}^{NV}\right).$$

The next lemma, which is similar to Lemma 4.2.2, compares the Ninomiya-Victoir scheme to the intermediate process $\bar{X}^j, j \in \{1,\dots,d\}$. We omit its proof.

Lemma 4.4.2 *Assume that*

- σ^0 *is a Lipschitz continuous function,*
- *for all* $j \in \{1,\dots,d\}, \sigma^j \in \mathcal{C}^1(\mathbb{R}^n,\mathbb{R}^n)$ *with bounded first order derivatives,*
- *for all* $j \in \{1,\dots,d\}, \partial\sigma^j\sigma^j$ *is a Lipschitz continuous function.*

Then, for all $p \geq 1$, *there exists a constant* $C_7 \in \mathbb{R}_+^*$, *such that for all* $N \in \mathbb{N}^*, t \in [0,T], s, s_1, s_2 \in [\hat{\tau}_t, t]$, *and all* $j, m \in \{1,\dots,d\}$,

$$\mathbb{E}\left[\left\|X_{\hat{\tau}_t}^{NV} - \bar{X}_{t,\hat{\tau}_t}^j\right\|^{2p}\right] \leq C_7 (\Delta t)^p, \tag{4.4.9}$$

$$\mathbb{E}\left[\left\|X_t^{NV} - \bar{X}_{t,s}^j\right\|^{2p}\right] \leq C_7 (\Delta t)^p, \tag{4.4.10}$$

$$\mathbb{E}\left[\left\|\bar{X}_{t,s_1}^j - \bar{X}_{t,s_2}^m\right\|^{2p}\right] \leq C_7 (\Delta t)^p. \tag{4.4.11}$$

In order to derive the estimation (4.2.4) from Proposition 4.2.4, we also need the following lemma, which gives several approximations of the processes Y^j, for $j \in \{1,\dots,d\}$.

Lemma 4.4.3 *Assume that*

- σ^0 *is a Lipschitz continuous function,*
- *for all* $j \in \{1,\dots,d\}, \sigma^j \in \mathcal{C}^2(\mathbb{R}^n,\mathbb{R}^n)$ *with bounded first order derivatives,* $\partial\sigma^j \in LIP_{loc}^{pgc}(\mathbb{R}^n \otimes \mathbb{R}^n)$ *and* $\partial^2\sigma^j \in LIP_{loc}^{pgc}(\mathbb{R}^n \otimes \mathbb{R}^n \otimes \mathbb{R}^n)$,
- *for all* $j \in \{1,\dots,d\}, \partial\sigma^j\sigma^j$ *is a Lipschitz continuous function.*

Then, for all ≥ 1, *there exists a constant* $C_8 \in \mathbb{R}_+^*$ *such that for all* $t \in [0,T], s \in [\hat{\tau}_t, t]$ *and all* $N \in \mathbb{N}^*$

$$\mathbb{E}\left[\left\|Y_{t,s}^j - Id_n\right\|^{2p}\right] \leq C_8 (\Delta t)^p, \tag{4.4.12}$$

$$\mathbb{E}\left[\left\|Y_{t,s}^j - Id_n - \Delta W_s^j \partial\sigma^j\left(\bar{X}_{t,\hat{\tau}_t}^j\right)\right\|^{2p}\right] \leq C_8 (\Delta t)^{2p}, \tag{4.4.13}$$

$$\mathbb{E}\left[\left\|Y_{t,s}^j - Id_n - \Delta W_s^j \partial\sigma^j\left(\bar{X}_{t,\hat{\tau}_t}^j\right) - \frac{1}{2}\left(\Delta W_s^j\right)^2\left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\left(\bar{X}_{t,\hat{\tau}_t}^j\right)\right\|^{2p}\right] \leq C_8 (\Delta t)^{3p}. \tag{4.4.14}$$

Proof : Let $p \geq 1, t \in [0,T]$ and $s \in [\hat{\tau}_t, t]$, we recall that $Y_{t,s}^j$ is the solution, at time ΔW_s^j, of the following ODE:

$$\begin{cases} \dfrac{d}{du}\Xi_{t,\hat{\tau}_t,u} = \partial\sigma^j\left(H_{t,u}^j\right)\Xi_{t,\hat{\tau}_t,u}, u \in \mathbb{R} \\ \Xi_{t,\hat{\tau}_t,\hat{\tau}_t} = Id_n, \end{cases} \tag{4.4.15}$$

where $H_{t,u}^j = h_j\left(\frac{\Delta t}{2}, \Delta W_t^1, \dots, \Delta W_t^{j-1}, u; X_{\hat{\tau}_t}^{NV}\right)$. Since $\partial\sigma^j$ is bounded, there exists a constant $c \in \mathbb{R}_+^*$, independent of N, t and s, such that

$$\left\|Y_{t,s}^j\right\|^{2p} \leq \exp\left(c\left|W_s^j - W_{\hat{\tau}_t}^j\right|\right). \tag{4.4.16}$$

Therefore,

$$\mathbb{E}\left[\left\|Y_{t,s}^{j}\right\|^{2p}\right] \leq 2\exp\left(\frac{1}{2}c^{2}\left(s-\hat{\tau}_{t}\right)\right) \leq 2\exp\left(\frac{1}{2}c^{2}T\right). \tag{4.4.17}$$

Now, we are able to prove (4.4.12). Writing $Y_{t,s}^{j}$ in integral form, we have:

$$Y_{t,s}^{j} = Id_{n} + \int_{0}^{\Delta W_{s}^{j}} \partial\sigma^{j}\left(H_{t,u}^{j}\right)\bar{Y}_{t,u}^{j}du,$$

where $\bar{Y}_{t,u}^{j}$ is the solution at time $u \in \mathbb{R}$ to the ODE (4.4.15). Applying Itô's formula, we obtain

$$Y_{t,s}^{j} - Id_{n} = \int_{\hat{\tau}_{t}}^{s} \partial\sigma^{j}\left(\bar{X}_{t,u}^{j}\right)Y_{t,u}^{j}dW_{u}^{j} + \frac{1}{2}\int_{\hat{\tau}_{t}}^{s}\left(\partial^{2}\sigma^{j}\odot\sigma^{j} + \left(\partial\sigma^{j}\right)^{2}\right)\left(\bar{X}_{t,u}^{j}\right)Y_{t,u}^{j}du. \tag{4.4.18}$$

Note that as $\partial\sigma^{j}\sigma^{j}$ is Lipschitz continuous, its derivatives given by $\partial^{2}\sigma^{j}\odot\sigma^{j} + \left(\partial\sigma^{j}\right)^{2}$ is bounded. Thus, combining a convexity inequality, the Burkholder-Davis-Gundy inequality and (4.4.17) we deduce (4.4.12). We prove now (4.4.13). Using (4.4.18), we have that

$$\begin{aligned}
Y_{t,s}^{j} - Id_{n} - \Delta W_{s}^{j}\partial\sigma^{j}\left(\bar{X}_{t,\hat{\tau}_{t}}^{j}\right) &= \int_{\hat{\tau}_{t}}^{s}\left(\partial\sigma^{j}\left(\bar{X}_{t,u}^{j}\right)Y_{t,u}^{j} - \partial\sigma^{j}\left(\bar{X}_{t,\hat{\tau}_{t}}^{j}\right)\right)dW_{u}^{j} \\
&+ \frac{1}{2}\int_{\hat{\tau}_{t}}^{s}\left(\partial^{2}\sigma^{j}\odot\sigma^{j} + \left(\partial\sigma^{j}\right)^{2}\right)\left(\bar{X}_{t,u}^{j}\right)Y_{t,u}^{j}du \\
&= \int_{\hat{\tau}_{t}}^{s}\left(\partial\sigma^{j}\left(\bar{X}_{t,u}^{j}\right) - \partial\sigma^{j}\left(\bar{X}_{t,\hat{\tau}_{t}}^{j}\right)\right)Y_{t,u}^{j}dW_{u}^{j} \\
&+ \int_{\hat{\tau}_{t}}^{s}\partial\sigma^{j}\left(\bar{X}_{t,\hat{\tau}_{t}}^{j}\right)\left(Y_{t,u}^{j} - Id_{n}\right)dW_{u}^{j} \\
&+ \frac{1}{2}\int_{\hat{\tau}_{t}}^{s}\left(\partial^{2}\sigma^{j}\odot\sigma^{j} + \left(\partial\sigma^{j}\right)^{2}\right)\left(\bar{X}_{t,u}^{j}\right)Y_{t,u}^{j}du.
\end{aligned} \tag{4.4.19}$$

Since $\partial^{2}\sigma^{j}\odot\sigma^{j} + \left(\partial\sigma^{j}\right)^{2}$ is bounded and $\partial\sigma^{j}$ is bounded and locally Lipschitz with polynomially growing Lipschitz constant, combining a convexity inequality, the Burkholder-Davis-Gundy inequality, (4.4.11) from Lemma 4.4.2, (4.4.12) from Lemma 4.4.3 and (4.4.17), we obtain (4.4.13). We can now focus on the approximation (4.4.14) , with strong order 3/2, of Y^{j}. We denote by Θ^{j} the process such that

$$\Theta_{t,s}^{j} = Y_{t,s}^{j} - Id_{n} - \Delta W_{s}^{j}\partial\sigma^{j}\left(\bar{X}_{t,\hat{\tau}_{t}}^{j}\right) - \frac{1}{2}\left(\Delta W_{s}^{j}\right)^{2}\left(\partial^{2}\sigma^{j}\odot\sigma^{j} + \left(\partial\sigma^{j}\right)^{2}\right)\left(\bar{X}_{t,\hat{\tau}_{t}}^{j}\right).$$

Writing

$$\begin{aligned}
\frac{1}{2}\left(\Delta W_{s}^{j}\right)^{2}\left(\partial^{2}\sigma^{j}\odot\sigma^{j} + \left(\partial\sigma^{j}\right)^{2}\right)\left(\bar{X}_{t,\hat{\tau}_{t}}^{j}\right) &= \int_{\hat{\tau}_{t}}^{s}\Delta W_{u}^{j}\left(\partial^{2}\sigma^{j}\odot\sigma^{j} + \left(\partial\sigma^{j}\right)^{2}\right)\left(\bar{X}_{t,\hat{\tau}_{t}}^{j}\right)dW_{u}^{j} \\
&+ \frac{1}{2}\int_{\hat{\tau}_{t}}^{s}\left(\partial^{2}\sigma^{j}\odot\sigma^{j} + \left(\partial\sigma^{j}\right)^{2}\right)\left(\bar{X}_{t,\hat{\tau}_{t}}^{j}\right)du,
\end{aligned}$$

and using (4.4.19), we get

$$\begin{aligned}\Theta^j_{t,s} &= \int_{\hat{\tau}_t}^{s} \left(\left(\partial\sigma^j\left(\bar{X}^j_{t,u}\right) - \partial\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\right) Y^j_{t,u} - \Delta W^j_u \partial^2\sigma^j \odot \sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\right) dW^j_u \\ &+ \int_{\hat{\tau}_t}^{s} \left(\partial\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\left(Y^j_{t,u} - Id_n\right) - \Delta W^j_u \left(\partial\sigma^j\right)^2\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\right) dW^j_u \\ &+ \frac{1}{2}\int_{\hat{\tau}_t}^{s} \left(\left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\left(\bar{X}^j_{t,u}\right) Y^j_{t,u} - \left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\right) du.\end{aligned}$$

Adding and subtracting some appropriate terms, we obtain:

$$\begin{aligned}\Theta^j_{t,s} &= \int_{\hat{\tau}_t}^{s} \left(\partial\sigma^j\left(\bar{X}^j_{t,u}\right) - \partial\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right) - \Delta W^j_u \partial^2\sigma^j \odot \sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\right) Y^j_{t,u} dW^j_u \\ &+ \int_{\hat{\tau}_t}^{s} \Delta W^j_u \partial^2\sigma^j \odot \sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\left(Y^j_{t,u} - Id_n\right) dW^j_u \\ &+ \int_{\hat{\tau}_t}^{s} \partial\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\left(Y^j_{t,u} - Id_n - \Delta W^j_u \partial\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\right) dW^j_u \\ &+ \frac{1}{2}\int_{\hat{\tau}_t}^{s} \left(\left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\left(\bar{X}^j_{t,u}\right) - \left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\right) Y^j_{t,u} du \\ &+ \frac{1}{2}\int_{\hat{\tau}_t}^{s} \left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\left(Y^j_{t,u} - Id_n\right) du.\end{aligned} \tag{4.4.20}$$

On the one hand, by the mean value theorem

$$\begin{aligned}\partial\sigma^j\left(\bar{X}^j_{t,u}\right) - \partial\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right) &= \partial^2\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right) \odot \left(\bar{X}^j_{t,u} - \bar{X}^j_{t,\hat{\tau}_t}\right) \\ &+ \left(\partial^2\sigma^j\left(\xi^j_{t,u}\right) - \partial^2\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\right) \odot \left(\bar{X}^j_{t,u} - \bar{X}^j_{t,\hat{\tau}_t}\right),\end{aligned} \tag{4.4.21}$$

where $\xi^j_{t,u}$ is a matrix of intermediate points between $\bar{X}^j_{t,u}$ and $\bar{X}^j_{t,\hat{\tau}_t}$. On the other hand, since

$$\bar{X}^j_{t,u} = \exp\left(\Delta W^j_u \sigma^j\right) \bar{X}^j_{t,\hat{\tau}_t},$$

applying Itô's formula we have that

$$\bar{X}^j_{t,u} - \bar{X}^j_{t,\hat{\tau}_t} = \Delta W^j_u \sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right) + \gamma^j_{t,u}, \tag{4.4.22}$$

where

$$\gamma^j_{t,u} = \int_{\hat{\tau}_t}^{u} \left(\sigma^j\left(\bar{X}^j_{t,v}\right) - \sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\right) dW^j_v + \frac{1}{2}\int_{\hat{\tau}_t}^{u} \partial\sigma^j\sigma^j\left(\bar{X}^j_{t,v}\right) dv.$$

Using (4.4.11) from Lemma 4.4.2, together with the regularity assumptions on σ^j and $\partial\sigma^j\sigma^j$ it is easy to see that γ^j a remainder with strong order 1: there exists $\alpha \in \mathbb{R}^*_+$ independent of N, t and u, such that

$$\mathbb{E}\left[\left\|\gamma^j_{t,u}\right\|^{2p}\right] \leq \alpha\left(\Delta u\right)^{2p}. \tag{4.4.23}$$

Combining (4.4.20), (4.4.21) and (4.4.22), we get

$$\begin{aligned}\Theta^j_{t,s} &= \int_{\hat{\tau}_t}^{s} \left(\partial^2\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\odot\gamma^j_{t,u} + \left(\partial^2\sigma^j\left(\xi^j_{t,u}\right) - \partial^2\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\right)\odot\left(\bar{X}^j_{t,u} - \bar{X}^j_{t,\hat{\tau}_t}\right)\right) Y^j_{t,u} dW^j_u \\
&+ \int_{\hat{\tau}_t}^{s} \Delta W^j_u \partial^2\sigma^j\odot\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\left(Y^j_{t,u} - Id_n\right) dW^j_u \\
&+ \int_{\hat{\tau}_t}^{s} \partial\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\left(Y^j_{t,u} - Id_n - \Delta W^j_s \partial\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\right) dW^j_u \\
&+ \frac{1}{2}\int_{\hat{\tau}_t}^{s} \left(\left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\left(\bar{X}^j_{t,u}\right) - \left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\right) Y^j_{t,u} du \\
&+ \frac{1}{2}\int_{\hat{\tau}_t}^{s} \left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\left(Y^j_{t,u} - Id_n\right) du.\end{aligned}$$

Since σ^j is Lipschitz continuous, $\partial\sigma^j$ and $\partial^2\sigma^j$ are locally Lipschitz with polynomially growing Lipschitz constant, it is easy now to see that (4.4.14) is a straightforward consequence of (4.4.11) from Lemma 4.4.2, (4.4.12), (4.4.13) and (4.4.23). ∎

Proof of Proposition 4.2.4:

We recall the result of this proposition:

$$E = \mathbb{E}\left[\sup_{t\le T}\left\|\int_0^t \left(Z^0_s - \theta^0_s\right) ds\right\|^{2p}\right] \le C_2 h^{3p},$$

where, for $t \in [0, T]$,

$$Z^0_t = Y^{d+1}_{t,t}\dots Y^1_{t,t}\sigma^0\left(\bar{X}^0_t\right),$$

and

$$\begin{aligned}\theta^0_t &= \sigma^0\left(\bar{X}^0_t\right) + \sum_{j=1}^{d}\Delta W^j_t \partial\sigma^j\sigma^0\left(X^{NV}_{\hat{\tau}_t}\right) + \frac{1}{2}\sum_{j=1}^{d}\Delta t\left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X^{NV}_{\hat{\tau}_t}\right) \\
&+ \frac{1}{2}\Delta t\partial\sigma^0\sigma^0\left(X^{NV}_{\hat{\tau}_t}\right).\end{aligned}$$

To prove this proposition, we will proceed in two steps.

Step 1: approximation with strong order $3/2$ of Z^0.

The first step consists in naively computing and approximating the product:

$$Z^0_t = Y^{d+1}_{t,t}\dots Y^1_{t,t}\sigma^0\left(\bar{X}^0_t\right).$$

We replace Y^{d+1} and Y^j, for $j \in \{1, \dots, d\}$, by their approximation (4.2.12) from Lemma 4.2.3 and (4.4.14) from Lemma 4.4.3, respectively.

Let

$$\Gamma_t = \left(Id_n + \frac{1}{2}\Delta t \partial\sigma^0\left(X_t^{NV}\right)\right)$$
$$\left(Id_n + \Delta W_t^d \partial\sigma^d\left(\bar{X}_{t,\hat{\tau}_t}^d\right) + \frac{1}{2}\left(\Delta W_t^d\right)^2\left(\partial^2\sigma^d \odot \sigma^d + \left(\partial\sigma^d\right)^2\right)\left(\bar{X}_{t,\hat{\tau}_t}^d\right)\right)\dots$$
$$\left(Id_n + \Delta W_t^1 \partial\sigma^1\left(\bar{X}_{t,\hat{\tau}_t}^1\right) + \frac{1}{2}\left(\Delta W_t^1\right)^2\left(\partial^2\sigma^1 \odot \sigma^1 + \left(\partial\sigma^1\right)^2\right)\left(\bar{X}_{t,\hat{\tau}_t}^1\right)\right),$$

and

$$\tilde{\theta}_t^0 = \Gamma_t \sigma^0\left(\bar{X}_t^0\right).$$

It is clear that $\tilde{\theta}^0$ is an approximation of strong order $3/2$ of Z^0: there exists $\tilde{\alpha} \in \mathbb{R}_+^*$ independent of N and t such that:

$$\mathbb{E}\left[\left\|Z_t^0 - \tilde{\theta}_t^0\right\|^{2p}\right] \leq \tilde{\alpha} h^{3p}. \tag{4.4.24}$$

Computing $\tilde{\theta}^0$, we easily deduce the following approximation with strong order $3/2$ by dropping higher order terms:

$$\mathbb{E}\left[\left\|\hat{\theta}_t^0 - \tilde{\theta}_t^0\right\|^{2p}\right] \leq \hat{\alpha} h^{3p}, \tag{4.4.25}$$

where

$$\begin{aligned}\hat{\theta}_t^0 &= \sigma^0\left(\bar{X}_t^0\right) + \sum_{j=1}^{d} \Delta W_t^j \partial\sigma^j\left(\bar{X}_{t,\hat{\tau}_t}^j\right)\sigma^0\left(\bar{X}_t^0\right) \\ &+ \sum_{j=1}^{d}\sum_{m=1}^{j-1} \Delta W_t^j \Delta W_t^m \partial\sigma^j\left(\bar{X}_{t,\hat{\tau}_t}^j\right)\partial\sigma^m\left(\bar{X}_{t,\hat{\tau}_t}^m\right)\sigma^0\left(\bar{X}_t^0\right) \\ &+ \sum_{j=1}^{d}\frac{1}{2}\left(\Delta W_t^j\right)^2\left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\left(\bar{X}_{t,\hat{\tau}_t}^j\right)\Bigg)\sigma^0\left(\bar{X}_t^0\right) - \frac{1}{2}\Delta t \partial\sigma^0\left(X_t^{NV}\right)\sigma^0\left(\bar{X}_t^0\right),\end{aligned}$$

and $\hat{\alpha} \in \mathbb{R}_+^*$ is a constant independent of N and t. We will now approximate the term $\Delta W_t^j \partial\sigma^j\left(\bar{X}_{t,\hat{\tau}_t}^j\right)\sigma^0\left(\bar{X}_t^0\right)$, for $j \in \{1,\dots,d\}$, with strong order $3/2$. To do so, it suffices to approximate $\partial\sigma^j\left(\bar{X}_{t,\hat{\tau}_t}^j\right)\sigma^0\left(\bar{X}_t^0\right)$ with strong order 1. By the mean value theorem

$$\begin{aligned}\partial\sigma^j\left(\bar{X}_{t,\hat{\tau}_t}^j\right) &= \partial\sigma^j\left(X_{\hat{\tau}_t}^{NV}\right) + \partial^2\sigma^j\left(X_{\hat{\tau}_t}^{NV}\right) \odot \left(\bar{X}_{t,\hat{\tau}_t}^j - X_{\hat{\tau}_t}^{NV}\right) \\ &+ \left(\partial^2\sigma^j\left(\zeta_t^j\right) - \partial^2\sigma^j\left(X_{\hat{\tau}_t}^{NV}\right)\right) \odot \left(\bar{X}_{t,\hat{\tau}_t}^j - X_{\hat{\tau}_t}^{NV}\right),\end{aligned}$$

where ζ_t^j is a matrix of intermediate points between $\bar{X}_{t,\hat{\tau}_t}^j$ and $X_{\hat{\tau}_t}^{NV}$. On the one hand, since for all $t \in [0,T]$, $\bar{X}_t^0 = \bar{X}_{t,\hat{\tau}_t}^1$ and $\bar{X}_{t,\hat{\tau}_t}^{m+1} = \bar{X}_{t,t}^m$, using telescopic summation, we have

$$\begin{aligned}\partial\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right) &= \partial\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right) + \sum_{m=1}^{j-1}\left(\partial^2\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\odot\left(\bar{X}^{m+1}_{t,\hat{\tau}_t} - \bar{X}^m_{t,\hat{\tau}_t}\right)\right)\\
&+ \partial^2\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\odot\left(\bar{X}^0_t - X^{NV}_{\hat{\tau}_t}\right) + \left(\partial^2\sigma^j\left(\zeta^j_t\right) - \partial^2\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\right)\odot\left(\bar{X}^j_{t,\hat{\tau}_t} - X^{NV}_{\hat{\tau}_t}\right)\\
&= \partial\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right) + \sum_{m=1}^{j-1}\left(\partial^2\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\odot\left(\bar{X}^m_{t,t} - \bar{X}^m_{t,\hat{\tau}_t}\right)\right)\\
&+ \partial^2\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\odot\left(\bar{X}^0_t - X^{NV}_{\hat{\tau}_t}\right) + \left(\partial^2\sigma^j\left(\zeta^j_t\right) - \partial^2\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\right)\odot\left(\bar{X}^j_{t,\hat{\tau}_t} - X^{NV}_{\hat{\tau}_t}\right).\end{aligned} \tag{4.4.26}$$

On the other hand, we recall that from (4.4.22) we have

$$\bar{X}^m_{t,t} - \bar{X}^m_{t,\hat{\tau}_t} = \Delta W^m_t \sigma^m\left(\bar{X}^m_{t,\hat{\tau}_t}\right) + \gamma^m_{t,t},$$

where γ^m is a remainder with strong order 1:

$$\gamma^m_{t,t} = \int_{\hat{\tau}_t}^t \left(\sigma^m\left(\bar{X}^m_{t,v}\right) - \sigma^m\left(\bar{X}^m_{t,\hat{\tau}_t}\right)\right) dW^m_v + \frac{1}{2}\int_{\hat{\tau}_t}^t \partial\sigma^m\sigma^m\left(\bar{X}^m_{t,v}\right) dv.$$

This leads us to the following decomposition

$$\partial\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right) = \partial\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right) + \sum_{m=1}^{j-1}\Delta W^m_t\partial^2\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\odot\sigma^m\left(\bar{X}^m_{t,\hat{\tau}_t}\right) + \Gamma^j_t, \tag{4.4.27}$$

where

$$\begin{aligned}\Gamma^j_t &= \sum_{m=1}^{j-1}\Delta W^m_t\partial^2\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\odot\gamma^m_{t,t} + \partial^2\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\odot\left(\bar{X}^0_t - X^{NV}_{\hat{\tau}_t}\right)\\
&+ \left(\partial^2\sigma^j\left(\zeta^j_t\right) - \partial^2\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\right)\odot\left(\bar{X}^j_{t,\hat{\tau}_t} - X^{NV}_{\hat{\tau}_t}\right).\end{aligned}$$

Using (4.2.8) from Lemma 4.2.2 and (4.4.9) from Lemma 4.4.2, it is easy to see that Γ^j is a remainder with strong order 1. Then it follows that $\check{\theta}^0$, which is given by

$$\begin{aligned}\check{\theta}^0_t &= \sigma^0\left(\bar{X}^0_t\right) + \sum_{j=1}^d \Delta W^j_t\partial\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\sigma^0\left(\bar{X}^0_t\right)\\
&+ \sum_{j=1}^d\sum_{m=1}^{j-1}\Delta W^j_t\Delta W^m_t\partial\sigma^j\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\partial\sigma^m\left(\bar{X}^m_{t,\hat{\tau}_t}\right)\sigma^0\left(\bar{X}^0_t\right)\\
&+ \sum_{j=1}^d\sum_{m=1}^{j-1}\Delta W^j_t\Delta W^m_t\partial^2\sigma^j\left(X^{NV}_{\hat{\tau}_t}\right)\odot\sigma^m\left(\bar{X}^m_{t,\hat{\tau}_t}\right)\sigma^0\left(\bar{X}^0_t\right)\\
&+ \sum_{j=1}^d\frac{1}{2}\left(\Delta W^j_t\right)^2\left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\left(\bar{X}^j_{t,\hat{\tau}_t}\right)\Bigg)\sigma^0\left(\bar{X}^0_t\right) - \frac{1}{2}\Delta t\partial\sigma^0\left(X^{NV}_t\right)\sigma^0\left(\bar{X}^0_t\right),\end{aligned}$$

is an approximation of strong order 3/2 of $\hat{\theta}^0$: there exists $\check{\alpha}\in\mathbb{R}^*_+$ independent of N and t such that:

$$\mathbb{E}\left[\left\|\check{\theta}^0_t - \hat{\theta}^0_t\right\|^{2p}\right] \leq \check{\alpha}h^{3p}. \tag{4.4.28}$$

Then, we can replace by $X^{NV}_{\hat{\tau}_t}$ the argument of the functions in the expression of $\check{\theta}^0$ when Lemma 4.2.2 and Lemma 4.4.2 ensure that the strong order 3/2 is preserved. This leads us to the following approximation, given by $\bar{\theta}^0$,

$$\begin{aligned}\bar{\theta}^0_t &= \sigma^0\left(\bar{X}^0_t\right) + \sum_{j=1}^{d} \Delta W^j_t \partial\sigma^j\sigma^0\left(X^{NV}_{\hat{\tau}_t}\right) \\ &+ \sum_{j=1}^{d}\sum_{m=1}^{j-1} \Delta W^j_t \Delta W^m_t \left(\partial^2\sigma^j \odot \sigma^m + \partial\sigma^j\partial\sigma^m\right)\sigma^0\left(X^{NV}_{\hat{\tau}_t}\right) \\ &+ \sum_{j=1}^{d} \frac{1}{2}\left(\Delta W^j_t\right)^2 \left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X^{NV}_{\hat{\tau}_t}\right) + \frac{1}{2}\Delta t \partial\sigma^0\sigma^0\left(X^{NV}_{\hat{\tau}_t}\right),\end{aligned} \tag{4.4.29}$$

and there exists $\bar{\alpha} \in \mathbb{R}^*_+$ independent of N and t, such that

$$\mathbb{E}\left[\left\|\bar{\theta}^0_t - \check{\theta}^0_t\right\|^{2p}\right] \le \bar{\alpha} h^{3p}. \tag{4.4.30}$$

Adding and subtracting $\bar{\theta}^0$, and using a convexity inequality, we get:

$$\begin{aligned}E = \mathbb{E}\left[\sup_{t\le T}\left\|\int_0^t \left(Z^0_s - \theta^0_s\right) ds\right\|^{2p}\right] \le 1 \vee T^{2p-1}\Bigg(\int_0^T \mathbb{E}\left[\left\|Z^0_t - \bar{\theta}^0_t\right\|^{2p}\right] dt \\ + \mathbb{E}\left[\sup_{t\le T}\left\|\int_0^t \left(\bar{\theta}^0_s - \theta^0_s\right) ds\right\|^{2p}\right]\Bigg).\end{aligned}$$

Combining a convexity inequality, (4.4.24), (4.4.25), (4.4.28), and (4.4.30), we obtain an estimation of the first expectation in the right-hand side of the above inequality:

$$\mathbb{E}\left[\left\|Z^0_t - \bar{\theta}^0_t\right\|^{2p}\right] \le 4^{2p-1}\left(\tilde{\alpha} + \hat{\alpha} + \check{\alpha} + \bar{\alpha}?\right) h^{3p}.$$

To achieve our goal, it remains to estimate:

$$\mathbb{E}\left[\sup_{t\le T}\left\|\int_0^t \left(\bar{\theta}^0_s - \theta^0_s\right) ds\right\|^{2p}\right].$$

This is the aim of the following step.

Step 2: the integration by parts formula.

Let $t \in [0, T]$, subtracting (4.2.4) from (4.4.29), we have that

$$\begin{aligned}\int_0^t \left(\bar{\theta}^0_s - \theta^0_s\right) ds &= \sum_{j=1}^{d}\sum_{m=1}^{j-1}\int_0^t \Delta W^j_s \Delta W^m_s \left(\partial^2\sigma^j \odot \sigma^m + \partial\sigma^j\partial\sigma^m\right)\sigma^0\left(X^{NV}_{\hat{\tau}_s}\right) ds \\ &+ \frac{1}{2}\int_0^t \left(\left(\Delta W^j_s\right)^2 - \Delta s\right)\left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X^{NV}_{\hat{\tau}_s}\right) ds.\end{aligned}$$

To lighten up the notation, we denote

$$F_{j,m} = \left(\partial^2\sigma^j \odot \sigma^m + \partial\sigma^j\partial\sigma^m\right)\sigma^0,$$

for $j, m \in \{1, \ldots, d\}, m \leq j$. Using the integration by parts formula, we have:

$$\int_0^t \Delta W_s^j \Delta W_s^m F_{j,m}\left(X_{\hat{\tau}_s}^{NV}\right) ds = \int_0^t \left(t \wedge \check{\tau}_s - s\right) \Delta W_s^j F_{j,m}\left(X_{\hat{\tau}_s}^{NV}\right) dW_s^m + \int_0^t \left(t \wedge \check{\tau}_s - s\right) \Delta W_s^m F_{j,m}\left(X_{\hat{\tau}_s}^{NV}\right) dW_s^j.$$

Taking the expectation of the supremum and using a convexity inequality, the Burkholder-Davis-Gundy inequality, we easily get a positive constant α_1, independent of N such that

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\int_0^t \Delta W_s^j \Delta W_s^m F_{j,m}\left(X_{\hat{\tau}_s}^{NV}\right) ds\right\|^{2p}\right] \leq \alpha_1 \int_0^T \left(\check{\tau}_s - s\right)^{2p} \mathbb{E}\left[\left\|\Delta W_s^m F_{j,m}\left(X_{\hat{\tau}_s}^{NV}\right)\right\|^{2p}\right] ds.$$

Then, by independence

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\int_0^t \Delta W_s^j \Delta W_s^m F_{j,m}\left(X_{\hat{\tau}_s}^{NV}\right) ds\right\|^{2p}\right] \leq \alpha_1 h^{3p} \int_0^T \left(\check{\tau}_s - s\right)^{2p} \mathbb{E}\left[\left\|F_{j,m}\left(X_{\hat{\tau}_s}^{NV}\right)\right\|^{2p}\right] ds.$$

Since $F_{j,m} \in LIP_{loc}^{pgc}\left(\mathbb{R}^n\right)$, we get $\alpha_2 \in \mathbb{R}_+$ and $q \geq 1$ independent of N such that

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\int_0^t \Delta W_s^j \Delta W_s^m F_{j,m}\left(X_{\hat{\tau}_s}^{NV}\right) ds\right\|^{2p}\right] \leq \alpha_2 \mathbb{E}\left[|G|^{2p}\right]\left(\int_0^T \mathbb{E}\left[\left\|\left(X_{\hat{\tau}_s}^{NV}\right)\right\|^{2q}\right] ds\right) h^{3p}.$$

We conclude using (4.2.6) from Lemma 4.2.2. To estimate,

$$\int_0^t \left(\left(\Delta W_s^j\right)^2 - \Delta s\right) F_{j,j}\left(X_{\hat{\tau}_s}^{NV}\right) ds,$$

we use exactly the same arguments since the integration by parts formula gives us

$$\int_0^t \left(\left(\Delta W_s^j\right)^2 - \Delta s\right) F_{j,j}\left(X_{\hat{\tau}_s}^{NV}\right) ds = 2\int_0^t \left(t \wedge \check{\tau}_s - s\right) \Delta W_s^j F_{j,j}\left(X_{\hat{\tau}_s}^{NV}\right) dW_s^j.$$

This completes the proof. ∎

In the following, we give a detailed proof of Proposition 4.2.5, which is a consequence of Propositions 4.2.3, 4.2.4.

Proof of Proposition 4.2.5:

Combining a convexity inequality and the Burkholder-Davis-Gundy, we get a constant α_0 independent of N such that

$$\mathbb{E}\left[\sup_{t\leq T}\left\|\bar{Q}_t^{1,N}\right\|^{2p}\right] \leq \alpha_0\left(E + \sum_{j=1}^{d}\left(E_1^j + E_2^j + E_3^j\right)\right), \tag{4.4.31}$$

where

$$E = \mathbb{E}\left[\sup_{t\leq T}\left\|\int_0^t \left(Z_s^0 - \theta_s^0\right) ds\right\|^{2p}\right],$$

for $j \in \{1, \ldots, d\}$,

$$E_1^j = \int_0^T \mathbb{E}\left[\left\|Y_{t,t}^{d+1}\sigma^j\left(\bar{X}_{t,t}\right) - \sigma^j\left(X_t^{NV}\right) - \frac{1}{2}\Delta t\left[\sigma^j, \sigma^0\right]\left(X_t^{NV}\right)\right\|^{2p}\right] dt,$$

$$E_2^j = \int_0^T \mathbb{E}\left[\left\|Y_{t,t}^{d+1}\partial\sigma^j\sigma^j\left(\bar{X}_{t,t}\right) - \partial\sigma^j\sigma^j\left(X_t^{NV}\right)\right.\right.$$
$$\left.\left. - \frac{1}{2}\Delta t\left(\partial\sigma^0\partial\sigma^j\sigma^j - \left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\right)\left(X_{\hat{\tau}_t}^{NV}\right)\right\|^{2p}\right] dt,$$

and

$$E_3^j = \int_0^T \mathbb{E}\left[\left\|Z_t^{d+1,j} - \frac{1}{2}\Delta t\left(\partial^2\sigma^0 \odot \sigma^j\right)\sigma^j\left(X_{\hat{\tau}_t}^{NV}\right)\right\|^{2p}\right] dt.$$

To prove our claim, it suffices to estimate, with at least strong order $3/2$, each expectations in the right-hand side of (4.4.31).

An estimation with strong order $3/2$ **of** E is given by Proposition 4.2.4.

Estimation with strong order 2 **of** E_1^j**, for** $j \in \{1, \ldots, d\}$**.**

To estimate E_1^j, we estimate its integrand, and we denote

$$\epsilon_1^j(t) = \mathbb{E}\left[\left\|\lambda_t^j\right\|^{2p}\right],$$

where

$$\lambda_t^j = Y_{t,t}^{d+1}\sigma^j\left(\bar{X}_{t,t}\right) - \sigma^j\left(X_t^{NV}\right) - \frac{1}{2}\Delta t\left[\sigma^j, \sigma^0\right]\left(X_t^{NV}\right).$$

Since

$$\begin{aligned}\lambda_t^j = &\left(Y_{t,t}^{d+1} - Id_n - \frac{1}{2}\Delta t\partial\sigma^0\left(X_t^{NV}\right)\right)\left(\sigma^j\left(\bar{X}_{t,t}\right) - \sigma^j\left(X_t^{NV}\right) + \frac{1}{2}\Delta t\partial\sigma^j\sigma^0\left(X_t^{NV}\right)\right)\\
&+ \left(Y_{t,t}^{d+1} - Id_n - \frac{1}{2}\Delta t\partial\sigma^0\left(X_t^{NV}\right)\right)\left(\sigma^j\left(X_t^{NV}\right) - \frac{1}{2}\Delta t\partial\sigma^j\sigma^0\left(X_t^{NV}\right)\right)\\
&+ \left(Id_n + \frac{1}{2}\Delta t\partial\sigma^0\left(X_t^{NV}\right)\right)\left(\sigma^j\left(\bar{X}_{t,t}\right) - \sigma^j\left(X_t^{NV}\right) + \frac{1}{2}\Delta t\partial\sigma^j\sigma^0\left(X_t^{NV}\right)\right)\\
&- \frac{1}{4}(\Delta t)^2\partial\sigma^0\partial\sigma^j\sigma^0\left(X_t^{NV}\right),\end{aligned}$$

combining a convexity inequality and the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
\epsilon_1^j(t) \leq 4^{2p-1}\Bigg(\Bigg(&\mathbb{E}\left[\left\|Y_{t,t}^{d+1} - Id_n - \frac{1}{2}\Delta t \partial\sigma^0\left(X_t^{NV}\right)\right\|^{4p}\right] \\
&\qquad \mathbb{E}\left[\left\|\sigma^j\left(\bar{X}_{t,t}\right) - \sigma^j\left(X_t^{NV}\right) + \frac{1}{2}\Delta t \partial\sigma^j\sigma^0\left(X_t^{NV}\right)\right\|^{4p}\right]\Bigg)^{\frac{1}{2}} \\
&+ \Bigg(\mathbb{E}\left[\left\|Y_{t,t}^{d+1} - Id_n - \frac{1}{2}\Delta t \partial\sigma^0\left(X_t^{NV}\right)\right\|^{4p}\right] \\
&\qquad \mathbb{E}\left[\left\|\sigma^j\left(X_t^{NV}\right) - \frac{1}{2}\Delta t \partial\sigma^j\sigma^0\left(X_t^{NV}\right)\right\|^{4p}\right]\Bigg)^{\frac{1}{2}} \\
&+ \Bigg(\mathbb{E}\left[\left\|Id_n + \frac{1}{2}\Delta t \partial\sigma^0\left(X_t^{NV}\right)\right\|^{4p}\right] \\
&\qquad \mathbb{E}\left[\left\|\sigma^j\left(\bar{X}_{t,t}\right) - \sigma^j\left(X_t^{NV}\right) + \frac{1}{2}\Delta t \partial\sigma^j\sigma^0\left(X_t^{NV}\right)\right\|^{4p}\right]\Bigg)^{\frac{1}{2}} \\
&+ \mathbb{E}\left[\left\|\frac{1}{4}(\Delta t)^2 \partial\sigma^0\partial\sigma^j\sigma^0\left(X_t^{NV}\right)\right\|^{2p}\right]\Bigg).
\end{aligned}$$

The first two expectations in the right-hand side of the previous inequality are estimated using (4.2.12) and (4.2.14) from Proposition 4.2.4, respectively. Since σ^0 and σ^j have bounded first order derivatives, applying and (4.2.6) from Lemma 4.2.2, we obtain a constant $\beta_j \in \mathbb{R}_+^*$ independent of N such that

$$\begin{aligned}
\mathbb{E}\left[\left\|\sigma^j\left(X_t^{NV}\right) - \frac{1}{2}\Delta t \partial\sigma^j\sigma^0\left(X_t^{NV}\right)\right\|^{4p}\right] &\leq \beta_j, \\
\mathbb{E}\left[\left\|Id_n + \frac{1}{2}\Delta t \partial\sigma^0\left(X_t^{NV}\right)\right\|^{4p}\right] &\leq \beta_j, \qquad (4.4.32)
\end{aligned}$$

and

$$\mathbb{E}\left[\left\|\frac{1}{4}(\Delta t)^2 \partial\sigma^0\partial\sigma^j\sigma^0\left(X_t^{NV}\right)\right\|^{2p}\right] \leq \beta_j h^{4p}.$$

Then it follows that

$$\epsilon_1^j(t) \leq 4^{2p-1}\left(C_1 T^{4p} + 2\sqrt{C_1\beta_j} + \beta_j\right) h^{4p},$$

and then

$$E_1^j \leq 4^{2p-1} T\left(C_1 T^{4p} + 2\sqrt{C_1\beta_j} + \beta_j\right) h^{4p}.$$

Estimation with strong order $3/2$ of E_2^j, for $j \in \{1, \ldots, d\}$.

As previously, we denote

$$\epsilon_2^j(t) = \mathbb{E}\left[\left\|\mu_t^j\right\|^{2p}\right],$$

where

$$\begin{aligned}
\mu_t^j &= Y_{t,t}^{d+1}\partial\sigma^j\sigma^j\left(\bar{X}_{t,t}\right) - \partial\sigma^j\sigma^j\left(X_t^{NV}\right) \\
&\quad - \frac{1}{2}\Delta t\left(\partial\sigma^0\partial\sigma^j\sigma^j - \left(\partial^2\sigma^j \odot \sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\right)\left(X_{\hat{\tau}_t}^{NV}\right).
\end{aligned}$$

Since

$$\begin{aligned}
\mu_t^j = &\left(Y_{t,t}^{d+1} - Id_n - \frac{1}{2}\Delta t \partial\sigma^0\left(X_t^{NV}\right)\right)\\
&\qquad\left(\partial\sigma^j\sigma^j\left(\bar{X}_{t,t}\right) - \partial\sigma^j\sigma^j\left(X_t^{NV}\right) + \frac{1}{2}\Delta t\left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right)\right)\\
&+\left(Y_{t,t}^{d+1} - Id_n - \frac{1}{2}\Delta t \partial\sigma^0\left(X_t^{NV}\right)\right)\\
&\qquad\left(\partial\sigma^j\sigma^j\left(X_t^{NV}\right) - \frac{1}{2}\Delta t\left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right)\right)\\
&+\left(Id_n + \frac{1}{2}\Delta t \partial\sigma^0\left(X_t^{NV}\right)\right)\\
&\qquad\left(\partial\sigma^j\sigma^j\left(\bar{X}_{t,t}\right) - \partial\sigma^j\sigma^j\left(X_t^{NV}\right) + \frac{1}{2}\Delta t\left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right)\right)\\
&-\frac{1}{4}\left(\Delta t\right)^2\partial\sigma^0\left(X_t^{NV}\right)\\
&\qquad\left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right) + \frac{1}{2}\Delta t\left(\partial\sigma^0\partial\sigma^j\sigma^j\left(X_t^{NV}\right) - \partial\sigma^0\partial\sigma^j\sigma^j\left(X_{\hat{\tau}_t}^{NV}\right)\right),
\end{aligned}$$

combining a convexity inequality, the Cauchy-Schwarz inequality, (4.2.12) and (4.2.15) from Proposition 4.2.4, and (4.4.32), we obtain

$$\begin{aligned}
\epsilon_2^j(t) \leq 5^{2p-1}\Bigg(&C_1 h^{7p}\\
&+\sqrt{C_1}h^{4p}\left(\mathbb{E}\left[\left\|\partial\sigma^j\sigma^j\left(X_t^{NV}\right) - \frac{1}{2}\Delta t\left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right)\right\|^{4p}\right]\right)^{\frac{1}{2}}\\
&+\sqrt{C_1\beta_j}h^{3p} + \mathbb{E}\left[\left\|\frac{1}{4}\left(\Delta t\right)^2\partial\sigma^0\left(X_t^{NV}\right)\left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right)\right\|^{2p}\right]\\
&+\mathbb{E}\left[\left\|\frac{1}{2}\Delta t\left(\partial\sigma^0\partial\sigma^j\sigma^j\left(X_t^{NV}\right) - \partial\sigma^0\partial\sigma^j\sigma^j\left(X_{\hat{\tau}_t}^{NV}\right)\right)\right\|^{2p}\right]\Bigg).
\end{aligned} \tag{4.4.33}$$

Since σ^0 and σ^j have bounded first order derivatives and $\partial^2\sigma^j$ is locally Lipschitz with polynomially growing Lipschitz constant, applying and (4.2.6) from Lemma 4.2.2, we easily get a constant $\gamma_j \in \mathbb{R}_+^*$ independent of N such that:

$$\mathbb{E}\left[\left\|\partial\sigma^j\sigma^j\left(X_t^{NV}\right) - \frac{1}{2}\Delta t\left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right)\right\|^{4p}\right] \leq \gamma_j,$$

and

$$\mathbb{E}\left[\left\|\frac{1}{4}\left(\Delta t\right)^2\partial\sigma^0\left(X_t^{NV}\right)\left(\partial^2\sigma^j\odot\sigma^j + \left(\partial\sigma^j\right)^2\right)\sigma^0\left(X_{\hat{\tau}_t}^{NV}\right)\right\|^{2p}\right] \leq \gamma_j h^{4p}.$$

It remains to estimate the last expectation in the right-hand side of (4.4.33). As the function $F^j := \frac{1}{2}\partial\sigma^0\partial\sigma^j\sigma^j$ is locally Lipschitz with polynomially growing Lipschitz constant, we easily get a constant $\delta_j \in \mathbb{R}_+^*$ independent of N such that

$$\mathbb{E}\left[\left\|\Delta t\left(F^j\left(X_t^{NV}\right) - F^j\left(X_{\hat{\tau}_t}^{NV}\right)\right)\right\|^{2p}\right] \leq \delta_j h^{3p}.$$

Then it follows that

$$\epsilon_2^j(t) \leq 5^{2p-1}\left(C_1 T^{4p} + \sqrt{C_1} T \gamma_j + \sqrt{C_1 \beta_j} + \gamma_j T^p + \delta_j\right) h^{3p},$$

and we conclude that

$$E_2^j \leq 5^{2p-1} T \left(C_1 T^{4p} + \sqrt{C_1} T \gamma_j + \sqrt{C_1 \beta_j} + \gamma_j T^p + \delta_j\right) h^{3p}.$$

An estimation with strong order $3/2$ **of** E_3^j**, for** $j \in \{1, \ldots, d\}$ is given by (4.2.13) from Proposition 4.2.3 and this concludes the proof. ∎

Chapter 5

Discretization of the ODEs involved in the Ninomiya-Victoir scheme

5.1 Introduction

In Chapter 2, we proposed new multilevel Monte Carlo estimators based on the Ninomiya-Victoir scheme, which is known to exhibit weak convergence with order 2, to compute $Y = \mathbb{E}\left[f\left(X_T\right)\right]$, where $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a payoff function and X_T is the solution, at time $T \in \mathbb{R}_+^*$, to a multi-dimensional stochastic differential equation of the form

$$\begin{cases} dX_t = b(X_t)dt + \sum_{j=1}^{d} \sigma^j(X_t)dW_t^j, \ t \in [0,T] \\ X_0 = x. \end{cases} \tag{5.1.1}$$

Here, $x \in \mathbb{R}^n$ is the initial condition, $W = \left(W^1, \ldots, W^d\right)$ is a $d-$dimensional standard Brownian motion, $b : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is the drift coefficient and $\sigma^j : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, for $j \in \{1, \ldots, d\}$, are the diffusion coefficients.

The derivation of our multilevel Monte Carlo estimators is based on a coupling with strong order 1 between the Giles-Szpruch scheme and a modified Ninomiya-Victoir scheme. Using this coupling, we built three new multilevel Monte Carlo estimators based on the Ninomiya-Victoir scheme which achieve the optimal complexity $O\left(\epsilon^{-2}\right)$ for a precision ϵ (see Chapter 2 for more details). However, the Ninomiya-Victoir scheme, introduced in [44], is based on the resolution of $d+1$ ODEs. If one of the ODEs has no closed-form solution, to implement the Ninomiya-Victoir scheme in practice, one needs to use an approximation. That is why, in this chapter, we focus on the strong convergence property of an approximation of the Ninomiya-Victoir scheme induced by the use of approximations for all ODEs in the scheme. In section 2, we provide an abstract approximation of the Ninomiya-Victoir scheme, together with general assumptions on the approximations of the ODEs. With a view to use our multilevel Monte Carlo estimators, we prove strong convergence with order 1 between the Ninomiya-Victoir scheme and its approximation, when the numerical integration is accurate up to the order 2 for the drift vector field σ^0 and up to the order 4 for the Brownian vector fields σ^j for $j \in \{1, \ldots, d\}$. That means that a first and a third order methods have to be used, respectively. By this way, the coupling with strong order 1 between the Giles-Szpruch scheme and the mean of the approximation of the Ninomiya-Victoir scheme with Rademacher random variables η and the approximation of the Ninomiya-Victoir scheme with Rademacher random variables $-\eta$, defined in Chapter 2, is preserved. This allows us to replace the Ninomiya-Victoir scheme by its numerical approximation, in our multilevel Monte Carlo estimators, when one of the ODEs has no closed-form solution and keep the computational complexity $O\left(\epsilon^{-2}\right)$.

According to Section 3.1 in [44], to keep the order 2 of weak convergence of the Ninomiya-Victoir scheme, the numerical integration should be accurate up to the order 3 for the drift vector filed σ^0 and up to the order 6 for the Brownian vector fields σ^j for $j \in \{1, \ldots, d\}$ (see also [43] and [4]). With regards to our multilevel Monte Carlo estimators based on the Ninomiya-Victoir scheme, this choice is recommended to fully take advantage of the order 2 of weak convergence of the Ninomiya-Victoir scheme and then reduce the computational time by decreasing the number of discretization levels. That is why, in section 3, we illustrate our abstract approximation of the Ninomiya-Victoir scheme, by giving classical examples of second and fifth order methods. We also check that our general assumptions, made in section 2, are satisfied.

In the following, we recall some notations

- $(t_k = kh)_{0 \leq k \leq N}$ is the subdivision of $[0, T]$ with equal time step $h = T/N$,
- for all $j \in \{1, \ldots, d\}, k \in \{0, \ldots, N-1\}, \Delta W^j_{t_{k+1}} = W^j_{t_{k+1}} - W^j_{t_k}$,
- $\eta = (\eta_1, \ldots, \eta_N)$ is a sequence of independent, identically distributed Rademacher random variables independent of W,
- for $V : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ Lipschitz continuous, $\exp(tV)x_0$ is the solution, at time $t \in \mathbb{R}$, of the following ordinary differential equation in $\mathbb{R}^n$

$$\begin{cases} \frac{dx(t)}{dt} = V\left(x(t)\right) \\ x(0) = x_0. \end{cases} \tag{5.1.2}$$

The Ninomiya-Victoir scheme introduced in [44], is defined as follows

- starting point: $X^{NV,\eta}_{t_0} = x$,
- for $k \in \{0 \ldots, N-1\}$, if $\eta_{k+1} = 1$:

$$X^{NV,\eta}_{t_{k+1}} = \exp\left(\frac{h}{2}\sigma^0\right)\exp\left(\Delta W^d_{t_{k+1}}\sigma^d\right)\ldots\exp\left(\Delta W^1_{t_{k+1}}\sigma^1\right)\exp\left(\frac{h}{2}\sigma^0\right)X^{NV,\eta}_{t_k},$$

  and if $\eta_{k+1} = -1$:

$$X^{NV,\eta}_{t_{k+1}} = \exp\left(\frac{h}{2}\sigma^0\right)\exp\left(\Delta W^1_{t_{k+1}}\sigma^1\right)\ldots\exp\left(\Delta W^d_{t_{k+1}}\sigma^d\right)\exp\left(\frac{h}{2}\sigma^0\right)X^{NV,\eta}_{t_k}.$$

where $\sigma^0 = b - \frac{1}{2}\sum_{j=1}^{d}\partial\sigma^j\sigma^j$ and $\partial\sigma^j$ is the Jacobian matrix of σ^j defined as follows

$$\partial\sigma^j = \left(\partial_{x_k}\sigma^{ij}\right)_{1 \leq i,k \leq n}.$$

5.2 Numerical approximation of the Ninomiya-Victoir scheme

In this section, we analyse the use of approximations for all ODEs in the Ninomiya-Victoir scheme. In the following, we first define an abstract approximation of the Ninomiya-Victoir scheme. Then, we make general assumptions for the approximation of ODEs in the Ninomiya-Victoir scheme. We finally prove a strong convergence result between the Ninomiya-Victoir scheme and its numerical approximation.

5.2.1 Abstract approximation and main result

We start by giving some notations. The numerical scheme which approximates the solution of ODE (5.1.2), at time $\theta \in \mathbb{R}$, with $V = \sigma^j$, for $j \in \{0, \dots, d\}$, is denoted by $\Psi^j(\theta, x_0)$. We are now able to define a general numerical approximation, denoted by $\hat{X}^{NV,\eta}$, of the Ninomiya-Victoir scheme.

- Starting point: $\hat{X}^{NV,\eta}_{t_0} = x$.
- For $k \in \{0 \dots, N-1\}$,

$$\hat{X}^{0,\eta}_{t_{k+1}} = \Psi^0\left(\frac{h}{2}, \hat{X}^{NV,\eta}_{t_k}\mathbb{1}_{\{\eta_{k+1}=1\}} + \hat{X}^{1,\eta}_{t_{k+1}}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right),$$

for $j \in \{1, \dots, d\}$,

$$\hat{X}^{j,\eta}_{t_{k+1}} = \Psi^j\left(\Delta W^j_{t_{k+1}}, \hat{X}^{j-1,\eta}_{t_{k+1}}\mathbb{1}_{\{\eta_{k+1}=1\}} + \hat{X}^{j+1,\eta}_{t_{k+1}}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right),$$

and

$$\hat{X}^{NV,\eta}_{t_{k+1}} = \hat{X}^{d+1,\eta}_{t_{k+1}} = \Psi^0\left(\frac{h}{2}, \hat{X}^{d,\eta}_{t_{k+1}}\mathbb{1}_{\{\eta_{k+1}=1\}} + \hat{X}^{NV,\eta}_{t_k}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right).$$

Hypothesis

To estimate the strong error between X^{NV} and $\hat{X}^{NV}$, we make some assumptions. The first two assumptions concern the numerical integration of $\exp(\theta\sigma^0)x_0$, for some $\theta \in [0, T]$. To ensure that the numerical scheme $\hat{X}^{NV}$ has uniformly bounded moments, we assume that for all $p \in \mathbb{N}^*$, there exists $C_0 \in \mathbb{R}^*_+$ such that for all $x_0 \in \mathbb{R}^n, \theta \in [0, T]$,

$$1 + \left\|\Psi^0(\theta, x_0)\right\|^{2p} \leq \exp(C_0\theta)\left(1 + \|x_0\|^{2p}\right). \tag{$\mathcal{H}_1$}$$

We also assume that the approximation Ψ^0 has order $m_0 \in \mathbb{N}$, meaning that the local truncation error is of order of $m_0 + 1$. More precisely, we assume that for all $p \in \mathbb{N}^*$, there exist $c_0 \in \mathbb{R}^*_+$ and $q \in \mathbb{N}^*$ such that for all $x_0 \in \mathbb{R}^n, \theta \in [0, T]$,

$$\left\|\exp\left(\theta\sigma^0\right)x_0 - \Psi^0(\theta, x_0)\right\|^{2p} \leq c_0\left(1 + \|x_0\|^{2q}\right)\theta^{2p(m_0+1)}. \tag{$\mathcal{H}_2$}$$

The last two assumptions concern the numerical integration of $\exp\left(W^j_\theta\sigma^j\right)x_0$, for $j \in \{1, \dots, d\}$ and for some $\theta \in [0, T]$. Similarly to assumption ($\mathcal{H}_1$), to ensure that the numerical scheme $\hat{X}^{NV}$ has uniformly bounded moments, we assume that for all $p \in \mathbb{N}^*$, there exists $C_1 \in \mathbb{R}^*_+$ such that for all $x_0 \in \mathbb{R}^n, \theta \in [0, T]$ and all $j \in \{1, \dots, d\}$,

$$\mathbb{E}\left[1 + \left\|\Psi^j\left(W^j_\theta, x_0\right)\right\|^{2p}\right] \leq \exp(C_1\theta)\left(1 + \|x_0\|^{2p}\right). \tag{$\mathcal{H}_3$}$$

Similarly to assumption ($\mathcal{H}_2$), we also assume that the approximation Ψ^j satisfies the following property: for all $p \in \mathbb{N}^*$, there exist $c_1 \in \mathbb{R}_+^*$ and $q \in \mathbb{N}^*$ such that for all $x_0 \in \mathbb{R}^n, \theta \in [0,T]$ and all $j \in \{1,\dots,d\}$,

$$\mathbb{E}\left[\left\|\exp\left(W_\theta^j \sigma^j\right)x_0 - \Psi^j\left(W_\theta^j, x_0\right)\right\|^{2p}\right] \le c_1\left(1+\|x_0\|^{2q}\right)\theta^{p(m+1)}. \qquad (\mathcal{H}_4)$$

Lemma 5.2.1 *If the approximation Ψ^j satisfies ($\mathcal{H}_3$) and has order $m \in \mathbb{N}$, which means that for all $M \in \mathbb{R}_+^*$, $p \in \mathbb{N}^*$, there exist $c_1' \in \mathbb{R}_+^*$ and $q' \in \mathbb{N}^*$ such that for all $x_0 \in \mathbb{R}^n, \tau \in [-M, M]$,*

$$\left\|\exp\left(\tau\sigma^j\right)x_0 - \Psi^j\left(\tau, x_0\right)\right\|^{2p} \le c_1'\left(1+\|x_0\|^{2q'}\right)|\tau|^{2p(m+1)}, \qquad (5.2.1)$$

it then satisfies ($\mathcal{H}_4$).

Proof : In the following we assume that ($\mathcal{H}_3$) and (5.2.1) are satisfied and we give a detailed derivation of ($\mathcal{H}_4$), using a localization technique. Let $M \in \mathbb{R}_+^*$, and $p \in \mathbb{N}^*$, then

$$\begin{aligned}\mathbb{E}\left[\left\|\exp\left(W_\theta^j \sigma^j\right)x_0 - \Psi^j\left(W_\theta^j, x_0\right)\right\|^{2p}\right] &= \mathbb{E}\left[\mathbb{1}_{\{|W_\theta|\le M\}}\left\|\exp\left(W_\theta^j \sigma^j\right)x_0 - \Psi^j\left(W_\theta^j, x_0\right)\right\|^{2p}\right] \\ &\quad + \mathbb{E}\left[\mathbb{1}_{\{|W_\theta|> M\}}\left\|\exp\left(W_\theta^j \sigma^j\right)x_0 - \Psi^j\left(W_\theta^j, x_0\right)\right\|^{2p}\right].\end{aligned} \qquad (5.2.2)$$

Using (5.2.1) for the first expectation in the right-hand side of (5.2.2), we get

$$\mathbb{E}\left[\mathbb{1}_{\{|W_\theta|\le M\}}\left\|\exp\left(W_\theta^j \sigma^j\right)x_0 - \Psi^j\left(W_\theta^j, x_0\right)\right\|^{2p}\right] \le c_1'\mathbb{E}\left[|G|^{2p(m+1)}\right]\left(1+\|x_0\|^{2q'}\right)\theta^{p(m+1)},$$

where G is a normal random variable. Applying the Cauchy-Schwarz inequality for the last expectation in the right-hand side of (5.2.2), we have

$$\mathbb{E}\left[\mathbb{1}_{\{|W_\theta|> M\}}\left\|\exp\left(W_\theta^j \sigma^j\right)x_0 - \Psi^j\left(W_\theta^j, x_0\right)\right\|^{2p}\right] \le \sqrt{\mathbb{P}\left(|W_\theta| > M\right)E\left(\theta, x_0\right)},$$

where $E(\theta, x_0) = \mathbb{E}\left[\left\|\exp\left(W_\theta^j \sigma^j\right)x_0 - \Psi^j\left(W_\theta^j, x_0\right)\right\|^{4p}\right]$. On the one hand, using Markov's inequality, we obtain

$$\mathbb{P}\left(|W_\theta| > M\right) \le \frac{1}{M^{4p(m+1)}}\mathbb{E}\left[|G|^{4p(m+1)}\right]\theta^{2p(m+1)}.$$

On the other hand, using a convexity inequality, we have

$$E(\theta, x_0) \le 2^{4p-1}\left(\mathbb{E}\left[1+\left\|\exp\left(W_\theta^j \sigma^j\right)x_0\right\|^{4p}\right] + \mathbb{E}\left[1+\left\|\Psi^j\left(W_\theta^j, x_0\right)\right\|^{4p}\right]\right).$$

Making the link between $\exp\left(W_\theta^j \sigma^j\right)x_0$ and the solution of an SDE at time θ, starting from x_0, (see [18]), one can estimate the first expectation in the right-hand side of the above inequality as follows

$$\mathbb{E}\left[1+\left\|\exp\left(W_\theta^j \sigma^j\right)x_0\right\|^{4p}\right] \le \exp\left(C\theta\right)\left(1+\|x_0\|^{4p}\right),$$

where C is positive constant independent of θ and x_0. Moreover, using ($\mathcal{H}_3$) we get

$$\mathbb{E}\left[1+\left\|\Psi^j\left(W^j_\theta,x_0\right)\right\|^{4p}\right]\leq \exp\left(C_1\theta\right)\left(1+\|x_0\|^{4p}\right),$$

and then,

$$E\left(\theta,x_0\right)\leq 2^{4p}\exp\left(C'\theta\right)\left(1+\|x_0\|^{4p}\right),$$

where $C'=\max\left(C,C_1\right)$. Hence

$$\sqrt{E\left(\theta,x_0\right)}\leq 2^{2p}\exp\left(\frac{1}{2}C'T\right)\left(1+\|x_0\|^{2p}\right),$$

which leads to

$$\mathbb{E}\left[\mathbb{1}_{\{|W_\theta|>M\}}\left\|\exp\left(W^j_\theta\sigma^j\right)x_0-\Psi^j\left(W^j_\theta,x_0\right)\right\|^{2p}\right]\leq c''_1\left(1+\|x_0\|^{2p}\right)\theta^{p(m+1)},$$

where $c''_1=\dfrac{1}{M^{2p(m+1)}}2^{2p}\exp\left(\dfrac{1}{2}C'T\right)\sqrt{\mathbb{E}\left[|G|^{2p(m+1)}\right]}$. Therefore, we can conclude that ($\mathcal{H}_4$) is satisfied by setting $c_1=2\left(c'_1\mathbb{E}\left[|G|^{2p(m+1)}\right]+c''_1\right)$ and $q=\max\left(p,q'\right)$. ∎

The following theorem is the main result of this chapter.

Theorem 5.2.2 *Assume that ($\mathcal{H}_1$), ($\mathcal{H}_2$), ($\mathcal{H}_3$) and ($\mathcal{H}_4$) hold. Then , if $m_0\geq 1$ and $m\geq 3$ the strong error between the Ninomiya-Victoir scheme $X^{NV,\eta}$ and its numerical approximation $\hat{X}^{NV,\eta}$ converges to 0 with order 1. More precisely, for all $p\in\mathbb{N}^*$ and $x\in\mathbb{R}^n$, there exist $\hat{C}_{NV}\in\mathbb{R}^*_+$ and $q\in\mathbb{N}^*$ such that for all $N\in\mathbb{N}^*$,*

$$\mathbb{E}\left[\left\|X^{NV,\eta}_T-\hat{X}^{NV,\eta}_T\right\|^{2p}\right]\leq\hat{C}_{NV}\left(1+\|x\|^{2q}\right)h^{2p}.$$

Remark 5.2.3 *To get strong convergence with order 2, it suffices to choose $m_0\geq 2$ and $m\geq 5$. This ensures that the numerical approximation $\hat{X}^{NV,\eta}$ has weak convergence with order 2.*

5.2.2 Intermediate results

To analyse the strong error between the Ninomiya-Victoir scheme $X^{NV,\eta}$ and its numerical approximation $\hat{X}^{NV,\eta}$, we recall the following intermediate processes, introduced in Chapter 2

$$\bar{X}^{0,\eta}_{t_{k+1}}=\exp\left(\frac{h}{2}\sigma^0\right)\left(X^{NV,\eta}_{t_k}\mathbb{1}_{\{\eta_{k+1}=1\}}+\bar{X}^{1,\eta}_{t_{k+1}}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right),$$

for $j\in\{1,\dots,d\}$,

$$\bar{X}^{j,\eta}_{t_{k+1}}=\exp\left(\Delta W^j_{t_{k+1}}\sigma^j\right)\left(\bar{X}^{j-1,\eta}_{t_{k+1}}\mathbb{1}_{\{\eta_{k+1}=1\}}+\bar{X}^{j+1,\eta}_{t_{k+1}}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right),$$

and

$$X^{NV,\eta}_{t_{k+1}}=\bar{X}^{d+1,\eta}_{t_{k+1}}=\exp\left(\frac{h}{2}\sigma^0\right)\left(\bar{X}^{d,\eta}_{t_{k+1}}\mathbb{1}_{\{\eta_{k+1}=1\}}+X^{NV,\eta}_{t_k}\mathbb{1}_{\{\eta_{k+1}=-1\}}\right).$$

Notice that for $1 \leq j \leq d$: $\left(\bar{X}^{j,\eta}_s\right)_{t_k < s \leq t_{k+1}}$ is the solution of the following SDE:

$$\begin{cases} dZ_s = \frac{1}{2}\partial\sigma^j\sigma^j\left(Z_s\right)ds + \sigma^j(Z_s)dW^j_s \\ Z_{t_k} = \bar{X}^{j-1,\eta}_{t_{k+1}}\mathbb{1}_{\{\eta_{k+1}=1\}} + \bar{X}^{j+1,\eta}_{t_{k+1}}\mathbb{1}_{\{\eta_{k+1}=-1\}}. \end{cases}$$

The following Proposition gives the stability of the solution of SDEs with respect to the initial conditions.

Proposition 5.2.4 *Let $p \in [1, +\infty)$ and let $Z = (Z_t)_{0\leq t\leq h}$ and $Y = (Y_t)_{0\leq t\leq h}$ be the solutions of the following $n-$dimensional SDE, driven by a $d-$dimensional brownian motion, on the time interval $[0, h]$:*

$$dV_s = \alpha(V_s)ds + \beta(V_s)dW_s$$

that start from Z_0 independent of $(W_t)_{t\in[0,h]}$ such that $\mathbb{E}\left[\|Z_0\|^{2p}\right] < +\infty$ and Y_0 independent of $(W_t)_{t\in[0,h]}$ such that $\mathbb{E}\left[\|Y_0\|^{2p}\right] < +\infty$, respectively. Assume that α and β are Lipschitz continuous functions. Then, there exists $C_2 \in \mathbb{R}^*_+$ such that for all $t \in [0, h]$,

$$\mathbb{E}\left[\|Z_t - Y_t\|^{2p}\right] \leq \mathbb{E}\left[\|Z_0 - Y_0\|^{2p}\right]\exp\left(C_2 h\right). \tag{5.2.3}$$

The constant C_2 only depends on $\|\alpha(0)\|, \|\beta(0)\|$, T, p, and the Lipschitz constants of α and β.

This result is well known, see [51] or [34] for example. We use the above Proposition to show the stability of the Ninomiya-Victoir scheme with respect to its initial condition. In the following, when the dependence on η is omitted, $X^{NV,t_k,y}$ denotes the Ninomiya-Victoir scheme which starts from $y \in \mathbb{R}^n$ at time t_k for $k \in \{0, \ldots, N\}$.

Proposition 5.2.5 *Assume that*

- *σ^0 is Lipschitz continuous,*
- *for all $j \in \{1, \ldots, d\}, \sigma^j \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first order derivatives, and that $\partial\sigma^j\sigma^j$ is Lipschitz continuous.*

Then for all $p \in [1, +\infty)$, there exists a constant $C_3 \in \mathbb{R}^*_+$ such that for all $y, z \in \mathbb{R}^n$, $N \in \mathbb{N}^*$ and $k \in \{0, \ldots, N\}$

$$\mathbb{E}\left[\left\|X^{NV,t_k,y}_T - X^{NV,t_k,z}_T\right\|^{2p}\right] \leq C_3\|y - z\|^{2p}.$$

Proof : Without loss of generality, to prove this proposition one can assume that for all $k \in \{1 \ldots, N\}$,

$$\eta_k = 1.$$

Let $p \in \mathbb{N}^*, y, z \in \mathbb{R}^n$, $k \in \{0, \ldots, N\}$ and $l \in \{k, \ldots, N\}$. Denoting

$$\bar{X}^{0,t_k,y}_{t_{l+1}} = \exp\left(\frac{h}{2}\sigma^0\right)X^{NV,t_k,y}_{t_l},$$

$$\bar{X}^{0,t_k,z}_{t_{l+1}} = \exp\left(\frac{h}{2}\sigma^0\right)X^{NV,t_k,z}_{t_l},$$

and applying (5.2.3) from Proposition 5.2.4, one has

$$\mathbb{E}\left[\left\|\bar{X}^{0,t_k,y}_{t_{l+1}} - \bar{X}^{0,t_k,z}_{t_{l+1}}\right\|^{2p}\right] \leq \mathbb{E}\left[\left\|X^{NV,t_k,y}_{t_l} - X^{NV,t_k,z}_{t_l}\right\|^{2p}\right]\exp\left(\frac{1}{2}C_2 h\right).$$

Similarly, denoting

$$\bar{X}^{j,t_k,y}_{t_{l+1}} = \exp\left(\Delta W^j_{t_{l+1}} \sigma^j\right) \bar{X}^{j-1,t_k,y}_{t_{l+1}},$$

$$\bar{X}^{j,t_k,z}_{t_{l+1}} = \exp\left(\Delta W^j_{t_{l+1}} \sigma^j\right) \bar{X}^{j-1,t_k,z}_{t_{l+1}},$$

for $j \in \{1, \ldots, d\}$ and applying (5.2.3) from Proposition 5.2.4, one gets

$$\mathbb{E}\left[\left\|\bar{X}^{j,t_k,y}_{t_{l+1}} - \bar{X}^{j,t_k,z}_{t_{l+1}}\right\|^{2p}\right] \leq \mathbb{E}\left[\left\|\bar{X}^{j-1,t_k,y}_{t_{l+1}} - \bar{X}^{j-1,t_k,z}_{t_{l+1}}\right\|^{2p}\right] \exp\left(C_2 h\right).$$

By induction, one obtains

$$\mathbb{E}\left[\left\|\bar{X}^{d,t_k,y}_{t_{l+1}} - \bar{X}^{d,t_k,z}_{t_{l+1}}\right\|^{2p}\right] \leq \exp\left(\left(d + \frac{1}{2}\right) C_2 h\right) \mathbb{E}\left[\left\|X^{NV,t_k,y}_{t_l} - X^{NV,t_k,z}_{t_l}\right\|^{2p}\right].$$

Since

$$X^{NV,t_k,y}_{t_{l+1}} = \exp\left(\frac{h}{2}\sigma^0\right) \bar{X}^{d,t_k,y}_{t_{l+1}},$$

and

$$X^{NV,t_k,z}_{t_{l+1}} = \exp\left(\frac{h}{2}\sigma^0\right) \bar{X}^{d,t_k,z}_{t_{l+1}},$$

using once again (5.2.3) from Proposition 5.2.4, one has

$$\mathbb{E}\left[\left\|X^{NV,t_k,y}_{t_{l+1}} - X^{NV,t_k,z}_{t_{l+1}}\right\|^{2p}\right] \leq \exp\left(\left(d+1\right) C_2 h\right) \mathbb{E}\left[\left\|X^{NV,t_k,y}_{t_l} - X^{NV,t_k,z}_{t_l}\right\|^{2p}\right].$$

Moreover, by definition

$$X^{NV,t_k,y}_{t_k} = y,$$

and

$$X^{NV,t_k,z}_{t_k} = z,$$

then, one can conclude that

$$\mathbb{E}\left[\left\|X^{NV,t_k,y}_{T} - X^{NV,t_k,z}_{T}\right\|^{2p}\right] \leq \exp\left(\left(d+1\right) C_2 T\right) \left\|y - z\right\|^{2p}.$$

∎

Before proving Theorem 5.2.2, in the following lemma we give an estimation of the moments of $\hat{X}^{NV,\eta}$ and $\hat{X}^{j,\eta}, j \in \{0\ldots, d+1\}$.

Lemma 5.2.6 *Assume that* ($\mathcal{H}_1$) *and* ($\mathcal{H}_3$) *hold. Then, for all* $p \in \mathbb{N}^*$, *there exists* $C_4 \in \mathbb{R}^*_+$ *such that for all* $N \in \mathbb{N}^*, k \in \{0, \ldots, N\}, j \in \{0, \ldots, d+1\}$ *and* $x \in \mathbb{R}^n$,

$$\mathbb{E}\left[1 + \left\|\hat{X}^{j,\eta}_{t_k}\right\|^{2p}\right] \leq \exp\left(C_4 t_{k+1}\right)\left(1 + \|x\|^{2p}\right), \tag{5.2.4}$$

and

$$\mathbb{E}\left[1 + \left\|\hat{X}^{NV,\eta}_{t_k}\right\|^{2p}\right] \leq \exp\left(C_4 t_{k+1}\right)\left(1 + \|x\|^{2p}\right). \tag{5.2.5}$$

The above lemma is a straightforward consequence of assumptions ($\mathcal{H}_1$) and ($\mathcal{H}_3$).

5.2.3 Proof of the strong convergence

Proof of Theorem 5.2.2 : Let $p \in \mathbb{N}^*$. The proof will go through several steps. The first step consists in decomposing, using the stability of the Ninomiya-Victoir scheme with respect to its initial condition, the global error of $X_T^{NV,\eta} - \hat{X}_T^{NV,\eta}$ into the sum of the local error made on each step.

Step 1: decomposition of the global error.

We begin by writing $X_T^{NV,\eta} - \hat{X}_T^{NV,\eta}$ using a telescopic summation as follows

$$X_T^{NV,\eta} - \hat{X}_T^{NV,\eta} = \sum_{k=0}^{N-1} \left(X_T^{NV,t_k,\hat{X}_{t_k}^{NV,\eta}} - X_T^{NV,t_{k+1},\hat{X}_{t_{k+1}}^{NV,\eta}} \right). \tag{5.2.6}$$

Since $X_T^{NV,t_k,\hat{X}_{t_k}^{NV,\eta}} = X_T^{NV,t_{k+1},X_{t_{k+1}}^{NV,t_k,\hat{X}_{t_k}^{NV,\eta}}}$, (5.2.6) rewrites

$$X_T^{NV,\eta} - \hat{X}_T^{NV,\eta} = \sum_{k=0}^{N-1} \left(X_T^{NV,t_{k+1},X_{t_{k+1}}^{NV,t_k,\hat{X}_{t_k}^{NV,\eta}}} - X_T^{NV,t_{k+1},\hat{X}_{t_{k+1}}^{NV,\eta}} \right).$$

Using a convexity inequality, we have

$$\left\| X_T^{NV,\eta} - \hat{X}_T^{NV,\eta} \right\|^{2p} \leq N^{2p-1} \sum_{k=0}^{N-1} \left\| X_T^{NV,t_{k+1},X_{t_{k+1}}^{NV,t_k,\hat{X}_{t_k}^{NV,\eta}}} - X_T^{NV,t_{k+1},\hat{X}_{t_{k+1}}^{NV,\eta}} \right\|^{2p}.$$

Taking the expectation and using Proposition 5.2.5, we get

$$\mathbb{E}\left[\left\| X_T^{NV,\eta} - \hat{X}_T^{NV,\eta} \right\|^{2p} \right] \leq C_3 N^{2p-1} \sum_{k=0}^{N-1} \mathbb{E}\left[\left\| X_{t_{k+1}}^{NV,t_k,\hat{X}_{t_k}^{NV,\eta}} - \hat{X}_{t_{k+1}}^{NV,\eta} \right\|^{2p} \right]. \tag{5.2.7}$$

The second step consists in estimating the local error, defined by

$$\mathbb{E}\left[\left\| X_{t_{k+1}}^{NV,t_k,\hat{X}_{t_k}^{NV,\eta}} - \hat{X}_{t_{k+1}}^{NV,\eta} \right\|^{2p} \right],$$

made on each step $k \in \{0, \dots, N-1\}$. Without loss of generality, to estimate this quantity we can assume that for all $k \in \{1 \dots, N\}$,

$$\eta_k = 1.$$

Step 2: estimation of the local error.

Let $k \in \{1 \dots, N\}$. To analyse the local error, we introduce the following random variables

$$\bar{Z}_{t_{k+1}}^0 = \exp\left(\frac{h}{2}\sigma^0\right) \hat{X}_{t_k}^{NV,\eta},$$

for $j \in \{1, \dots, d\}$,

$$\bar{Z}_{t_{k+1}}^j = \exp\left(\Delta W_{t_{k+1}}^j \sigma^j\right) \hat{X}_{t_{k+1}}^{j-1,\eta},$$

and
$$\bar{Z}^{d+1}_{t_{k+1}} = \exp\left(\frac{h}{2}\sigma^0\right)\hat{X}^{d,\eta}_{t_{k+1}}.$$
To lighten up the notation, we also introduce
$$\bar{Y}^{0}_{t_{k+1}} = \exp\left(\frac{h}{2}\sigma^0\right)\hat{X}^{NV,\eta}_{t_k},$$
for $j \in \{1,\ldots,d\}$,
$$\bar{Y}^{j}_{t_{k+1}} = \exp\left(\Delta W^j_{t_{k+1}}\sigma^j\right)\bar{Y}^{j-1}_{t_{k+1}},$$
and
$$\bar{Y}^{d+1}_{t_{k+1}} = \exp\left(\frac{h}{2}\sigma^0\right)\bar{Y}^{d}_{t_{k+1}},$$
so that
$$\bar{Y}^{d+1}_{t_{k+1}} = X^{NV,t_k,\hat{X}^{NV,\eta}_{t_k}}_{t_{k+1}}.$$
Using the law of iterated expectations together with $\hat{X}^{NV,\eta}_{t_{k+1}} = \hat{X}^{d+1,\eta}_{t_{k+1}}$, we write the local error as follows
$$\mathbb{E}\left[\left\|\bar{Y}^{d+1}_{t_{k+1}} - \hat{X}^{d+1,\eta}_{t_{k+1}}\right\|^{2p}\right] = \mathbb{E}\left[\mathbb{E}\left[\left\|\bar{Y}^{d+1}_{t_{k+1}} - \hat{X}^{d+1,\eta}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right]\right]. \tag{5.2.8}$$
The task is now to estimate the conditional expectation
$$E^{d+1}\left(\hat{X}^{NV,\eta}_{t_k}\right) := \mathbb{E}\left[\left\|\bar{Y}^{d+1}_{t_{k+1}} - \hat{X}^{d+1,\eta}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right].$$

Step 2.1: estimation of $E^{d+1}\left(\hat{X}^{NV,\eta}_{t_k}\right)$.

Adding and subtracting $\bar{Z}^{d+1}_{t_{k+1}}$, and using a convexity inequality, we obtain
$$E^{d+1}\left(\hat{X}^{NV,\eta}_{t_k}\right) \leq 2^{2p-1}\left(\mathbb{E}\left[\left\|\bar{Y}^{d+1}_{t_{k+1}} - \bar{Z}^{d+1}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right] + \mathbb{E}\left[\left\|\bar{Z}^{d+1}_{t_{k+1}} - \hat{X}^{d+1,\eta}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right]\right).$$
On the one hand, using (5.2.3) from Proposition 5.2.4, we get
$$\mathbb{E}\left[\left\|\bar{Y}^{d+1}_{t_{k+1}} - \bar{Z}^{d+1}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right] \leq \exp\left(\frac{1}{2}C_2 h\right)E^d\left(\hat{X}^{NV,\eta}_{t_k}\right),$$
where
$$E^{d}\left(\hat{X}^{NV,\eta}_{t_k}\right) = \mathbb{E}\left[\left\|\bar{Y}^{d}_{t_{k+1}} - \hat{X}^{d,\eta}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right].$$
On the other hand, using $(\mathcal{H}_2)$
$$\begin{aligned}\mathbb{E}\left[\left\|\bar{Z}^{d+1}_{t_{k+1}} - \hat{X}^{d+1,\eta}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right] &= \mathbb{E}\left[\left\|\exp\left(\frac{h}{2}\sigma^0\right)\hat{X}^{d,\eta}_{t_{k+1}} - \Psi^0\left(\frac{h}{2},\hat{X}^{d,\eta}_{t_{k+1}}\right)\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right]\\ &\leq \frac{c_0}{2^{2p(m_0+1)}}\mathbb{E}\left[1+\left\|\hat{X}^{d,\eta}_{t_{k+1}}\right\|^{2q}\middle|\hat{X}^{NV,\eta}_{t_k}\right]h^{2p(m_0+1)}.\end{aligned}$$

Similarly to Lemma 5.2.6, we can show by induction that an estimation of the expectation in the right-hand side of the last inequality is given by

$$\mathbb{E}\left[1+\left\|\hat{X}^{d,\eta}_{t_{k+1}}\right\|^{2q}\middle|\hat{X}^{NV,\eta}_{t_k}\right] \leq \exp\left(C_4 T\right)\left(1+\left\|\hat{X}^{NV,\eta}_{t_k}\right\|^{2q}\right). \tag{5.2.9}$$

Hence

$$E^{d+1}\left(\hat{X}^{NV,\eta}_{t_k}\right) \leq 2^{2p-1}\exp\left(\frac{1}{2}C_2 h\right)E^{d}\left(\hat{X}^{NV,\eta}_{t_k}\right)+\gamma_0\left(1+\left\|\hat{X}^{NV,\eta}_{t_k}\right\|^{2q}\right)h^{2p(m_0+1)},$$

where $\gamma_0 = \dfrac{2^{2p-1}}{2^{2p(m_0+1)}}C_0\exp\left(C_4 T\right)$. In the same way, we denote for $j\in\{0,\ldots,d\}$,

$$E^{j}\left(\hat{X}^{NV,\eta}_{t_k}\right) = \mathbb{E}\left[\left\|\bar{Y}^{j}_{t_{k+1}}-\hat{X}^{j,\eta}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right].$$

Step 2.2: estimation of $E^{j}\left(\hat{X}^{NV,\eta}_{t_k}\right)$ for $j\in\{0,\ldots,d\}$.

We start with $j\in\{1,\ldots,d\}$. Adding and subtracting $\bar{Z}^{j}_{t_{k+1}}$, we have

$$E^{j}\left(\hat{X}^{NV,\eta}_{t_k}\right) \leq 2^{2p-1}\left(\mathbb{E}\left[\left\|\bar{Y}^{j}_{t_{k+1}}-\bar{Z}^{j}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right]+\mathbb{E}\left[\left\|\bar{Z}^{j}_{t_{k+1}}-\hat{X}^{j,\eta}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right]\right).$$

On the one hand, using (5.2.3) from Proposition 5.2.4, we have

$$\mathbb{E}\left[\left\|\bar{Y}^{j}_{t_{k+1}}-\bar{Z}^{j}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right] \leq \exp\left(C_2 h\right)E^{j-1}\left(\hat{X}^{NV,\eta}_{t_k}\right).$$

On the other hand, using ($\mathcal{H}_4$), together with the independence between $\Delta W^{j}_{t_{k+1}}$ and $\hat{X}^{j-1,\eta}_{t_{k+1}}$,

$$\mathbb{E}\left[\left\|\bar{Z}^{j}_{t_{k+1}}-\hat{X}^{j,\eta}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right] \leq c_1\mathbb{E}\left[1+\left\|\hat{X}^{j-1,\eta}_{t_{k+1}}\right\|^{2q}\middle|\hat{X}^{NV,\eta}_{t_k}\right]h^{p(m+1)}.$$

Similarly to Lemma 5.2.6 and (5.2.9), an estimation of the expectation in the right-hand side of the last inequality is given by

$$\mathbb{E}\left[1+\left\|\hat{X}^{j-1,\eta}_{t_{k+1}}\right\|^{2q}\middle|\hat{X}^{NV,\eta}_{t_k}\right] \leq \exp\left(C_4 T\right)\left(1+\left\|\hat{X}^{NV,\eta}_{t_k}\right\|^{2q}\right).$$

Hence

$$E^{j}\left(\hat{X}^{NV,\eta}_{t_k}\right) \leq 2^{2p-1}\exp\left(C_2 h\right)E^{j-1}\left(\hat{X}^{NV,\eta}_{t_k}\right)+\gamma\left(1+\left\|\hat{X}^{NV,\eta}_{t_k}\right\|^{2p}\right)h^{p(m+1)},$$

where $\gamma = 2^{2p-1}c_1\exp\left(C_4 T\right)$. The estimation

$$E^{0}\left(\hat{X}^{NV,\eta}_{t_k}\right) = \mathbb{E}\left[\left\|\bar{Y}^{0}_{t_{k+1}}-\hat{X}^{0}_{t_{k+1}}\right\|^{2p}\middle|\hat{X}^{NV,\eta}_{t_k}\right] \leq \gamma_0'\left(1+\left\|\hat{X}^{NV,\eta}_{t_k}\right\|^{2q}\right)h^{2p(m_0+1)},$$

where $\gamma_0' = \dfrac{c_0}{2^{2p((m_0+1)}}$, is given by ($\mathcal{H}_2$).

Step 2.3: conclusion of step 2

By induction, it follows that

$$\begin{aligned} E^{d+1}\left(\hat{X}_{t_k}^{NV,\eta}\right) &\leq \left(2^{(d+1)(2p-1)}\exp\left(\left(d+\frac{1}{2}\right)C_2T\right)\gamma_0'+\gamma_0\right)\left(1+\left\|\hat{X}_{t_k}^{NV,\eta}\right\|^{2q}\right)h^{2p(m_0+1)} \\ &\quad + (d+1)\left(2^{d(2p-1)}\exp\left(\left(d-\frac{1}{2}\right)C_2T\right)\gamma\right)\left(1+\left\|\hat{X}_{t_k}^{NV,\eta}\right\|^{2q}\right)h^{p(m+1)} \\ &\leq C_5\left(1+\left\|\hat{X}_{t_k}^{NV,\eta}\right\|^{2q}\right)\left(h^{2p(m_0+1)}+h^{p(m+1)}\right), \end{aligned} \tag{5.2.10}$$

where

$$C_5 = \left(2^{(d+1)(2p-1)}\exp\left(\left(d+\frac{1}{2}\right)C_2T\right)\gamma_0'+\gamma_0\right)+(d+1)\left(2^{d(2p-1)}\exp\left(\left(d-\frac{1}{2}\right)C_2T\right)\gamma\right).$$

Combining (5.2.8), (5.2.10) and (5.2.5) from Lemma 5.2.6, we get

$$\mathbb{E}\left[\left\|X_{t_{k+1}}^{NV,t_k,\hat{X}_{t_k}^{NV,\eta}}-\hat{X}_{t_{k+1}}^{NV,\eta}\right\|^{2p}\right] \leq C_5\exp\left(C_4T\right)\left(1+\|x\|^{2q}\right)\left(h^{2p(m_0+1)}+h^{p(m+1)}\right).$$

Step 3: conclusion

We conclude that the global error is given by

$$\mathbb{E}\left[\left\|X_T^{NV,\eta}-\hat{X}_T^{NV,\eta}\right\|^{2p}\right] \leq C_3N^{2p}C_5\exp\left(C_4T\right)\left(1+\|x\|^{2q}\right)\left(h^{2p(m_0+1)}+h^{p(m+1)}\right). \tag{5.2.11}$$

Therefore, to get

$$\mathbb{E}\left[\left\|X_T^{NV,\eta}-\hat{X}_T^{NV,\eta}\right\|^{2p}\right] \leq \hat{C}\left(1+\|x\|^{2q}\right)h^{2p},$$

for some constant $\hat{C}$ independent of N, it suffices to choose $m_0 \geq 1$ and $m \geq 3$, and this proves our claim. ∎

The choice $m_0 = 1$ and $m = 3$ is then suitable to keep the order 1 of strong convergence between the following modified Ninomiya-Victoir scheme defined as in Chapter 2 by

$$\hat{\hat{X}}^{NV,\eta} = \frac{1}{2}\left(\hat{X}^{NV,\eta}+\hat{X}^{NV,-\eta}\right)$$

and the Giles-Szpruch scheme [23]. With a view to use our multilevel Monte Carlo estimators based on the Ninomiya-Victoir scheme, this allows us to replace the Ninomiya-Victoir scheme by its numerical approximation when one of the ODEs has no closed-form solution and keep the computational complexity $O\left(\epsilon^{-2}\right)$ for a precision ϵ.

As already mentioned, to obtain strong convergence with order 2, estimate (5.2.11) ensures that the choice $m_0 = 2$ and $m = 5$ is suitable. The weak order 2 is then preserved. With regards to our multilevel Monte Carlo estimators based on the the Ninomiya-Victoir scheme, this choice is recommended to fully take advantage of the order 2 of weak of the Ninomiya-Victoir scheme and then reduce the computational time by decreasing the number of discretization levels. To be more specific, if $\exp\left(\theta\sigma^0\right)x_0$ or/and $\exp\left(\theta\sigma^j\right)x_0$, for some $j \in \{1,\ldots,d\}$, have no closed-form solution, a second order or/and a fifth order method have to be used, respectively. That is why, in the next section we mainly focus on second and fifth order methods for ordinary differential equations.

5.3 Numerical methods for ordinary differential equations

In this section, we illustrate our abstract approximation of the Ninomiya-Victoir scheme, by giving classical examples for the numerical approximation of ODEs. We check that our general assumptions $(\mathcal{H}_1)$, $(\mathcal{H}_2)$, $(\mathcal{H}_3)$, and $(\mathcal{H}_4)$ are satisfied. As already mentioned, from a practical point of view, it is interesting to keep the order 2 of weak convergence. That is why, in this section we focus on the Runge-Kutta methods of order $m_0 = 2$ and $m = 5$, to approximate Ninomiya-Victoir scheme. Before recalling the Runge-Kutta methods, we start by checking that our general assumptions are satisfied for the Euler method which is a first order Runge-Kutta method. The aim is to show that our assumptions are satisfied in a simple case. Then we can easily generalize our approach to the family of explicit Runge-Kutta methods.

5.3.1 The explicit Euler method

We consider the autonomous ODE (5.1.2) in $\mathbb{R}^n$, which is recalled here

$$\begin{cases} \frac{dx(t)}{dt} = V\left(x(t)\right) \\ x(0) = x_0. \end{cases}$$

where $x_0 \in \mathbb{R}^n$ is the initial condition and V is a Lipschitz continuous vector field. In the following, we denote by $x\left(\theta, x_0\right)$ the solution of the ODE (5.1.2) at $\theta \in [-T, T]$. The explicit Euler approximation of (5.1.2) at time θ is defined by

$$x^E\left(\theta, x_0\right) = x_0 + \theta V\left(x_0\right).$$

Proposition 5.3.1 *Assume that the vector field $V \in \mathcal{C}^1\left(\mathbb{R}^n, \mathbb{R}^n\right)$ with bounded first order derivatives. Then x^E satisfies $(\mathcal{H}_1)$, $(\mathcal{H}_2)$, $(\mathcal{H}_3)$ and $(\mathcal{H}_4)$.*

From Lemma 5.2.1, $(\mathcal{H}_4)$ is a consequence $(\mathcal{H}_2)$ and $(\mathcal{H}_3)$. Therefore, to prove Proposition 5.3.1, we only need to show that

- for all $p \in \mathbb{N}^*$, there exists $C_0 \in \mathbb{R}_+^*$ such that for all $x_0 \in \mathbb{R}^n, \theta \in [0, T]$,

$$1 + \left\|x^E\left(\theta, x_0\right)\right\|^{2p} \leq \exp\left(C_0\theta\right)\left(1 + \|x_0\|^{2p}\right), \tag{5.3.1}$$

- for all $p \in \mathbb{N}^*$, there exist $c_0 \in \mathbb{R}_+^*$ and $q \in \mathbb{N}^*$ such that for all $x_0 \in \mathbb{R}^n, \theta \in [-T, T]$,

$$\left\|x\left(\theta, x_0\right) - x^E\left(\theta, x_0\right)\right\|^{2p} \leq c_0\left(1 + \|x_0\|^{2p}\right)|\theta|^{4p}, \tag{5.3.2}$$

- for all $p \in \mathbb{N}^*$, there exists $C_1 \in \mathbb{R}_+^*$ such that for all $x_0 \in \mathbb{R}^n, \theta \in [0,T]$,

$$\mathbb{E}\left[1 + \left\|x^E\left(W_\theta^j, x_0\right)\right\|^{2p}\right] \leq \exp\left(C_1\theta\right)\left(1 + \|x_0\|^{2p}\right). \tag{5.3.3}$$

Proof : The solution $x \in \mathcal{C}^2\left([-T,T], \mathbb{R}^n\right)$ since

$$\frac{dx\left(t, x_0\right)}{dt} = V\left(x\left(t, x_0\right)\right),$$

and

$$\frac{d^2x\left(t, x_0\right)}{dt^2} = \partial V\left(x\left(t, x_0\right)\right) V\left(x\left(t, x_0\right)\right).$$

We start by checking (5.3.1)**.**

Let $p \in \mathbb{N}^*$, and $\theta \in [0,T]$. To obtain (5.3.1), we compute exactly the norm of $x^E\left(\theta, x_0\right)$

$$\left\|x^E\left(\theta, x_0\right)\right\|^2 = \|x_0\|^2 + \theta^2\left\|V\left(x_0\right)\right\|^2 + 2\theta\left\langle x_0, V\left(x_0\right)\right\rangle.$$

Using the binomial expansion, we get

$$\begin{aligned}\left\|x^E\left(\theta, x_0\right)\right\|^{2p} &= \sum_{k=0}^{p}\sum_{l=0}^{p-k}\binom{p}{k}\binom{p-k}{l}\|x_0\|^{2(p-k-l)}\theta^{2l}\left\|V\left(x_0\right)\right\|^{2l}2^k\theta^k\left(\left\langle x_0, V\left(x_0\right)\right\rangle\right)^k \\ &= \sum_{k=0}^{p}\sum_{l=0}^{p-k}\binom{p}{k}\binom{p-k}{l}2^k\theta^{2l+k}\|x_0\|^{2(p-k-l)}\left\|V\left(x_0\right)\right\|^{2l}\left(\left\langle x_0, V\left(x_0\right)\right\rangle\right)^k.\end{aligned} \tag{5.3.4}$$

Using the Cauchy-Schwarz inequality, we obtain

$$\left\|x^E\left(\theta, x_0\right)\right\|^{2p} \leq \sum_{k=0}^{p}\sum_{l=0}^{p-k}\binom{p}{k}\binom{p-k}{l}2^k\theta^{2l+k}\|x_0\|^{2(p-k-l)}\left\|V\left(x_0\right)\right\|^{2l+k}\|x_0\|^k.$$

By the Lipschitz assumption, there exists a constant $L \in \mathbb{R}_+^*$, which only depends on V, such that

$$\begin{aligned}\left\|x^E\left(\theta, x_0\right)\right\|^{2p} &\leq \sum_{k=0}^{p}\sum_{l=0}^{p-k}\binom{p}{k}\binom{p-k}{l}2^k\theta^{2l+k}\|x_0\|^{2(p-k-l)}L^{2l+k}\left(1 + \|x_0\|\right)^{2l+k}\|x_0\|^k \\ &= \sum_{k=0}^{p}\sum_{l=0}^{p-k}\sum_{m=0}^{2l+k}\binom{p}{k}\binom{p-k}{l}\binom{2l+k}{m}2^kL^{2l+k}\theta^{2l+k}\|x_0\|^{2p-m}.\end{aligned} \tag{5.3.5}$$

We first assume that $\|x_0\| \geq 1$. Then, it follows that

$$\begin{aligned}\left\|x^E\left(\theta, x_0\right)\right\|^{2p} &\leq \|x_0\|^{2p}\sum_{k=0}^{p}\sum_{l=0}^{p-k}\sum_{m=0}^{2l+k}\binom{p}{k}\binom{p-k}{l}\binom{2l+k}{m}2^kL^{2l+k}\theta^{2l+k} \\ &= \|x_0\|^{2p}\left(1 + P_1(\theta)\right),\end{aligned}$$

where

$$P_1(\theta) := \sum_{k=0}^{p}\sum_{l=0}^{p-k}\sum_{m=1}^{2l+k}\binom{p}{k}\binom{p-k}{l}\binom{2l+k}{m}2^kL^{2l+k}\theta^{2l+k},$$

with the convention that the sum over an empty set is 0. Hence

$$1+\left\|x^{E}\left(\theta, x_{0}\right)\right\|^{2p} \leq\left(1+\left\|x_{0}\right\|^{2p}\right)\left(1+P_{1}(\theta)\right).$$

Since P_1 is a polynomial such that $P_1(0)=0$, we can easily find a constant $C_0 \in \mathbb{R}_+^*$ independent of θ such that

$$1+\left\|x^{E}\left(\theta, x_{0}\right)\right\|^{2p} \leq\left(1+\left\|x_{0}\right\|^{2p}\right)\exp\left(C_0\theta\right).$$

We now assume that $\|x_0\|<1$. Using (5.3.5)

$$\begin{aligned}
\left\|x^{E}\left(\theta, x_{0}\right)\right\|^{2p} \leq & \left\|x_{0}\right\|^{2p}\sum_{k=0}^{p}\sum_{l=0}^{p-k}\binom{p}{k}\binom{p-k}{l}2^{k}L^{2l+k}\theta^{2l+k} \\
&+\sum_{k=0}^{p}\sum_{l=0}^{p-k}\sum_{m=1}^{2l+k}\binom{p}{k}\binom{p-k}{l}\binom{2l+k}{m}2^{k}L^{2l+k}\theta^{2l+k} \\
&\leq\left\|x_{0}\right\|^{2p}\left(1+P_{2}(\theta)\right)+P_{1}(\theta),
\end{aligned}$$

where

$$P_{2}(\theta):=\sum_{k=0}^{p}\sum_{l=0}^{p-k}\binom{p}{k}\binom{p-k}{l}2^{k}L^{2l+k}\theta^{2l+k}.$$

Hence

$$\begin{aligned}
1+\left\|x^{E}\left(\theta, x_{0}\right)\right\|^{2p} &\leq 1+\left\|x_{0}\right\|^{2p}\left(1+P_{2}(\theta)\right)+P_{1}(\theta) \\
&\leq\left(1+\left\|x_{0}\right\|^{2p}\right)\left(1+P_{2}(\theta)+P_{1}(\theta)\right).
\end{aligned}$$

Since P_1 and P_2 are polynomials such that $P_1(0)=P_2(0)=0$, we conclude as in the previous case.

We now check (5.3.2).

Let $p\in\mathbb{N}^*$, and $\theta\in[-T,T]$. The quantity $\left\|x\left(\theta,x_0\right)-x^E\left(\theta,x_0\right)\right\|$ is called the local truncation error. It is well known that the explicit Euler has a local error of order 2 [13] [26]. However, we exhibit the dependency with respect to the initial condition x_0. Performing a second order Taylor series expansion to the i-th coordinate of the solution $x(\theta,x_0)$, one has

$$x^{i}\left(\theta, x_{0}\right)=x_{0}^{i}+\theta V^{i}\left(x_{0}\right)+\frac{\theta^{2}}{2}\nabla V^{i}\left(x\left(\tau, x_{0}\right)\right).\,V\left(x\left(\tau, x_{0}\right)\right)$$

where $\tau=\alpha\theta$, for some $\alpha\in[0,1]$. Then, the i-th coordinate of the local truncation error of the explicit Euler method is given by

$$\epsilon^{i,E}\left(\theta, x_{0}\right)=\left|\frac{\theta^{2}}{2}\nabla V^{i}\left(x\left(\tau, x_{0}\right)\right).\,V\left(x\left(\tau, x_{0}\right)\right)\right|.$$

Since ∇V is bounded, there exists a constant $C\in\mathbb{R}_+^*$, independent of θ and x_0, such that

$$\epsilon^{i,E}\left(\theta, x_{0}\right)\leq C\left(1+\left\|x\left(\tau, x_{0}\right)\right\|\right)\frac{\theta^{2}}{2}.$$

Moreover, $x(\tau, x_0)$ is the solution to the ODE (5.1.2) at time $\tau \in [-T, T]$, then there exists a constant $C' \in \mathbb{R}_+^*$ independent of θ and x_0 such that

$$1 + \|x(\tau, x_0)\| \leq C' (1 + \|x_0\|).$$

Hence,

$$\epsilon^{i,E}(\theta, x_0) \leq CC' (1 + \|x_0\|) \frac{\theta^2}{2},$$

and we easily get (5.3.2) using a convexity inequality.

To finish, we check (5.3.3).

Let $p \in \mathbb{N}^*$, and $\theta \in [0, T]$. Using (5.3.4) taking the expectation and denoting

$$E_{2p}(\theta, x_0) = \mathbb{E}\left[\left\|x^E\left(W_\theta^j, x_0\right)\right\|^{2p}\right],$$

we have that

$$E_{2p}(\theta, x_0) = \sum_{k=0}^{p} \sum_{l=0}^{p-k} \binom{p}{k} \binom{p-k}{l} 2^k g_{l,k} \theta^{l+k/2} \|x_0\|^{2(p-k-l)} \|V(x_0)\|^{2l} (\langle x_0, V(x_0)\rangle)^k$$

where $g_{l,k} = \mathbb{E}\left[G^{2l+k}\right]$ and G is a normal random variable. The only term that seems to be problematic in this sum, which corresponds to $\sqrt{\theta}$, is obtained when $l = 0$ and $k = 1$. However, since $g_{0,1} = 0$, we can conclude as in (5.3.1). ∎

5.3.2 Explicit Runge-Kutta methods

A general Runge-Kutta [13] [26] method of order r can be written as

$$\Phi_k(\theta) = V\left(x_0 + \theta \sum_{l=1}^{r} a_{kl} \Phi_l(\theta)\right)$$

and

$$x^{RK}(\theta, x_0) = x_0 + \theta \sum_{k=1}^{r} b_k \Phi_k(\theta).$$

A Runge-Kutta method is often represented by a partitioned table, known as a Butcher tableau, of the form

$$\begin{array}{c|ccc} & a_{1,1} & \dots & a_{1,r} \\ & \vdots & & \vdots \\ & a_{r,1} & \dots & a_{r,r} \\ \hline & b_1 & \dots & b_r \end{array}$$

in which the matrix $(a_{kl})_{1\leq k,l,\leq r}$ is called the Runge-Kutta matrix, while the $(b_k)_{1\leq k\leq r}$ are known as the weights. When the Runge-Kutta matrix $(a_{kl})_{1\leq k,l,\leq r}$ is zero-diagonal lower triangular, the method is called explicit. Otherwise, the method is called implicit. Although the implicit methods are in general more stable than the explicit method, in the following, we only focus on the explicit method. Indeed, implementing a numerical approximation $\hat{X}^{NV,\eta}$, using an implicit Runge-Kutta method, induces that at each step, a system of algebraic equations has to be solved. This increases the computational cost considerably and affects the efficiency of the multilevel methods based on the Ninomiya-Victoir scheme.

Proposition 5.3.2 *Assume that the vector field V is Lipschitz continuous. Then, for all $r \in \mathbb{N}^*$, all zero-diagonal lower triangular matrix $(a_{kl})_{1\leq k,l,\leq r} \in \mathbb{R}^r \otimes \mathbb{R}^r$ and all vector $(b_k)_{1\leq k\leq r} \in \mathbb{R}^r$ such that $\sum_{k=1}^{r} b_k = 1$, x^{RK} satisfies $(\mathcal{H}_1)$ and $(\mathcal{H}_3)$.*

Proof : Let $r \in \mathbb{N}^*$, $(a_{kl})_{1\leq k,l,\leq r}$ be a zero-diagonal lower triangular matrix of $\mathbb{R}^r \otimes \mathbb{R}^r$, $(b_k)_{1\leq k\leq r}$ be a vector $\mathbb{R}^r$ such that $\sum_{k=1}^{r} b_k = 1$, and let $k \in \{2, \dots, r\}$. Since V is Lipschitz continuous, there exists a constant $C \in \mathbb{R}_+^*$, which only depends on V, such that for all $\tau \in \mathbb{R}$ and all $x_0 \in \mathbb{R}^n$

$$\|\Phi_1(\tau)\| = \|V(x_0)\| \leq C(1 + \|x_0\|),$$

and

$$\begin{aligned}\|\Phi_k(\tau) - V(x_0)\| &\leq C|\tau| \sum_{l=1}^{k-1} |a_{kl}| \|\Phi_l(\tau)\| \\ &\leq C|\tau| \sum_{l=1}^{k-1} |a_{kl}| (\|\Phi_l(\tau) - V(x_0)\| + \|V(x_0)\|).\end{aligned}$$

By induction, we get a constant $C' \in \mathbb{R}_+^*$, which only depends on V and a, such that

$$\|\Phi_k(\tau) - V(x_0)\| \leq C' \sum_{l=1}^{k-1} |\tau|^l (1 + \|x_0\|). \tag{5.3.6}$$

Writing

$$x^{RK}(\tau, x_0) = x_0 + \tau V(x_0) + \tau \sum_{k=2}^{r} b_k (\Phi_k(\tau) - V(x_0)),$$

and using (5.3.6), there exists a polynomial P with smallest exponent greater than 2 and positive coefficients, which only depends on V, a and b, such that

$$\left\|\tau \sum_{k=2}^{r} b_k (\Phi_k(\tau) - V(x_0))\right\| \leq P(|\tau|)(1 + \|x_0\|).$$

Now when $\tau = \theta \in [0, T]$ is deterministic, we easily get $(\mathcal{H}_1)$ using the binomial expansion of $\left\|x^{RK}\right\|^{2p}$ as in the proof of (5.3.1) from Proposition 5.3.1. When $\tau = W_\theta$, for some $\theta \in [0, T]$, we first perform the binomial expansion of $\left\|x^{RK}\right\|^{2p}$. Then, since the smallest exponent of the polynomial P is greater than 2 the only term that seems to be problematic in the binomial expansion of $\left\|x^{RK}\right\|^{2p}$ is

$$W_\theta \|x_0\|^{2(p-1)} \langle x_0, V(x_0)\rangle.$$

Taking its expectation, this term vanishes and $(\mathcal{H}_3)$ is then satisfied.

∎

To implement the multilevel methods based on the Ninomiya-Victoir scheme, we present a second and fifth order Runge-Kutta method. To approximate (5.1.2) when $V = \sigma^0$, we give the following Butcher tableau [13] of a second order Runge-Kutta method

$$\begin{array}{c|cc} & 0 & \\ & 1 & 0 \\ \hline & 1/2 & 1/2 \end{array}$$

in other words, an approximation of order 2 of (5.1.2), denoted by x^{RK2}, is defined by

$$x^{RK2}(\theta, x_0) = x_0 + \frac{\theta}{2}\sigma^0(x_0) + \frac{\theta}{2}\sigma^0\Big(x_0 + \theta\sigma^0(x_0)\Big).$$

Proposition 5.3.3 *Assume that the vector field $\sigma^0 \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first order derivatives and with polynomially growing second order derivatives. Then x^{RK2} satisfies ($\mathcal{H}_2$) with $m_0 = 2$.*

This method is also called Heun's method and its well known that it has a local error of order 3. However, we exhibit the dependency with respect to the initial condition x_0.

Proof : Let $\theta \in [0, T]$ and $x_0 \in \mathbb{R}^n$. On the one hand, performing a third order Taylor series expansion to the i-th coordinate of the solution $x(\theta, x_0)$, one has

$$\begin{aligned} x^i(\theta, x_0) = {} & x_0^i + \theta\sigma^{i0}(x_0) + \frac{\theta^2}{2}\sum_{k=1}^{n}\partial_{x_k}\sigma^{i0}(x_0)\,\sigma^{k0}(x_0) \\ & + \frac{\theta^3}{6}\sum_{k=1}^{n}\sum_{l=1}^{n}\partial^2_{x_k x_l}\sigma^{i0}(x(\tau, x_0))\,\sigma^{l0}(x(\tau, x_0))\,\sigma^{k0}(x(\tau, x_0)) \\ & + \frac{\theta^3}{6}\sum_{k=1}^{n}\sum_{l=1}^{n}\partial_{x_k}\sigma^{i0}(x(\tau, x_0))\,\partial_{x_l}\sigma^{k0}(x(\tau, x_0))\,\sigma^{l0}(x(\tau, x_0)) \end{aligned}$$

where $\tau = \alpha\theta$, for some $\alpha \in [0, 1]$. On the other hand, performing a second order Taylor series expansion to the i-th coordinate of $x^{RK2}(\theta, x_0)$, one gets

$$\begin{aligned} x^{RK2,i}(\theta, x_0) = {} & x_0^i + \theta\sigma^{i0}(x_0) + \frac{\theta^2}{2}\sum_{k=1}^{n}\partial_{x_k}\sigma^{i0}(x_0)\,\sigma^{k0}(x_0) \\ & + \frac{\theta^3}{4}\sum_{k=1}^{n}\sum_{l=1}^{n}\partial^2_{x_k x_l}\sigma^{i0}(\xi)\,\sigma^{l0}(x_0)\,\sigma^{k0}(x_0), \end{aligned}$$

where $\xi = \alpha'\theta\sigma^0(x_0) + (1 - \alpha')\,x_0$, for some $\alpha' \in [0, 1]$. Then, it follows that

$$\begin{aligned} \left|x^i(\theta, x_0) - x^{RK2,i}(\theta, x_0)\right| \leq {} & \frac{\theta^3}{6}\sum_{k=1}^{n}\sum_{l=1}^{n}\left|\partial^2_{x_k x_l}\sigma^{i0}(x(\tau, x_0))\,\sigma^{l0}(x(\tau, x_0))\,\sigma^{k0}(x(\tau, x_0))\right| \\ & + \frac{\theta^3}{6}\sum_{k=1}^{n}\sum_{l=1}^{n}\left|\partial_{x_k}\sigma^{i0}(x(\tau, x_0))\,\partial_{x_l}\sigma^{k0}(x(\tau, x_0))\,\sigma^{l0}(x(\tau, x_0))\right| \\ & + \frac{\theta^3}{4}\sum_{k=1}^{n}\sum_{l=1}^{n}\left|\partial^2_{x_k x_l}\sigma^{i0}(\xi)\,\sigma^{l0}(x_0)\,\sigma^{k0}(x_0)\right|. \end{aligned}$$

Since σ^0 has bounded first order derivatives and polynomially growing second order derivatives, one can easily get ($\mathcal{H}_2$). ∎

We now present a fifth order Runge-Kutta method, to approximate (5.1.2) when $V = \sigma^j$ for $j \in \{1, \ldots, d\}$. The Butcher tableau [13] of a fifth order Runge-Kutta method, denoted by x^{RK5} is given by

$$\begin{array}{c|cccccc}
 & 0 & & & & & \\
 & 1/4 & 0 & & & & \\
 & 1/8 & 1/8 & 0 & & & \\
 & 0 & 0 & 1/2 & 0 & & \\
 & 3/16 & -6/16 & 6/16 & 9/16 & 0 & \\
 & -3/7 & 8/7 & 6/7 & -12/7 & 8/7 & 0 \\
\hline
 & 7/90 & 0 & 32/90 & 12/90 & 32/90 & 7/90
\end{array}$$

Proposition 5.3.4 *Assume that the vector field $\sigma^j \in \mathcal{C}^5(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first and with polynomially growing derivatives up to the order 5. Then x^{RK5} satisfies ($\mathcal{H}_4$) with $m = 5$.*

It is well known that this Runge-Kutta method has a local error of order 6. However, to obtain the dependency with respect to the initial condition as in Proposition 5.3.3, one has to apply a sixth and fifth order Taylor series expansion to the solution x and the Runge-Kutta method x^{RK5} respectively, so that both expansion coincide up to the order 6. Then, the local truncation error is a sum involving σ^j and its derivatives up to the order 5 and its is easy to conclude thanks to the regularity assumptions on σ^j.

References

[1] Al Gerbi, A., Jourdain, B., and Clément, E. (2015). Ninomiya-victoir scheme: strong convergence, antithetic version and application to multilevel estimators. *Preprint, http://arxiv.org/abs/1508.06492*.

[2] Al Gerbi, A., Jourdain, B., and Clément, E. (2016a). Asymptotic error distribution for the ninomiya-victoir scheme in the commutative case. *Preprint, http://arxiv.org/abs/1605.08270*.

[3] Al Gerbi, A., Jourdain, B., and Clément, E. (2016b). Asymptotics for the normalized error of the ninomiya-victoir scheme. *Preprint, http://arxiv.org/abs/1601.05268*.

[4] Alfonsi, A. (2010). High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Math. Comp.*, 79(269):209–237.

[5] Alfonsi, A. (2015). *Affine diffusions and related processes: simulation, theory and applications*, volume 6 of *Bocconi & Springer Series*. Springer, Cham; Bocconi University Press, Milan.

[6] Bally, V. and Rey, C. (2016). Approximation of Markov semigroups in total variation distance. *Electron. J. Probab.*, 21:Paper No. 12, 44.

[7] Bally, V. and Talay, D. (1996a). The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function. *Probab. Theory Related Fields*, 104(1):43–60.

[8] Bally, V. and Talay, D. (1996b). The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density. *Monte Carlo Methods Appl.*, 2(2):93–128.

[9] Bayer, C. and Friz, P. K. (2013). Cubature on Wiener space: pathwise convergence. *Appl. Math. Optim.*, 67(2):261–278.

[10] Ben Alaya, M. and Kebaier, A. (2015). Central limit theorem for the multilevel Monte Carlo Euler method. *Ann. Appl. Probab.*, 25(1):211–234.

[11] Beskos, A., Papaspiliopoulos, O., and Roberts, G. O. (2006). Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli*, 12(6):1077–1098.

[12] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The journal of political economy*, pages 637–654.

[13] Butcher, J. C. (2008). *Numerical methods for ordinary differential equations*. John Wiley & Sons, Ltd., Chichester, second edition.

[14] Clark, J. M. C. and Cameron, R. J. (1980). The maximum rate of convergence of discrete approximations for stochastic differential equations. In *Stochastic differential systems (Proc. IFIP-WG 7/1 Working Conf., Vilnius, 1978)*, volume 25 of *Lecture Notes in Control and Information Sci.*, pages 162–171. Springer, Berlin-New York.

[15] Cox, J. C., Ingersoll Jr, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, pages 385–407.

[16] Debrabant, K. and Rößler, A. (2015). On the acceleration of the multi-level Monte Carlo method. *J. Appl. Probab.*, 52(2):307–322.

[17] Dieudonné, J. (1963). *Fondements de l'analyse moderne*. Traduit de l'anglais par D. Huet. Avant-propos de G. Julia. Cahiers Scientifiques, Fasc. XXVIII. Gauthier-Villars, Éditeur, Paris.

[18] Doss, H. (1977). Liens entre équations différentielles stochastiques et ordinaires. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 13(2):99–125.

[19] Duffie, D. and Glynn, P. (1995). Efficient Monte Carlo simulation of security prices. *Ann. Appl. Probab.*, 5(4):897–905.

[20] Fujiwara, T. (2006). Sixth order methods of kusuoka approximation. *Preprint UTMS, http://kyokan.ms.u-tokyo.ac.jp/users/ preprint/pdf/2006-7.pdf.*

[21] Gaines, J. G. and Lyons, T. J. (1994). Random generation of stochastic area integrals. *SIAM J. Appl. Math.*, 54(4):1132–1146.

[22] Giles, M. B. (2008). Multilevel Monte Carlo path simulation. *Oper. Res.*, 56(3):607–617.

[23] Giles, M. B. and Szpruch, L. (2014). Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation. *Ann. Appl. Probab.*, 24(4):1585–1620.

[24] Guyon, J. (2006). Euler scheme and tempered distributions. *Stochastic Process. Appl.*, 116(6):877–904.

[25] Gyurkó, L. G. and Lyons, T. J. (2011). Efficient and practical implementations of cubature on Wiener space. In *Stochastic analysis 2010*, pages 73–111. Springer, Heidelberg.

[26] Hairer, E., Nørsett, S. P., and Wanner, G. (1993). *Solving ordinary differential equations. I*, volume 8 of *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, second edition. Nonstiff problems.

[27] Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, 6(2):327–343.

[28] Ho, T. S. and LEE, S.-B. (1986). Term structure movements and pricing interest rate contingent claims. *the Journal of Finance*, 41(5):1011–1029.

[29] Hörmander, L. (1967). Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171.

[30] Jacod, J. (1997). On continuous conditional Gaussian martingales and stable convergence in law. In *Séminaire de Probabilités, XXXI*, volume 1655 of *Lecture Notes in Math.*, pages 232–246. Springer, Berlin.

[31] Jacod, J. and Protter, P. (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.*, 26(1):267–307.

[32] Kanagawa, S. (1989). The rate of convergence for approximate solutions of stochastic differential equations. *Tokyo J. Math.*, 12(1):33–48.

[33] Kebaier, A. (2005). Statistical Romberg extrapolation: a new variance reduction method and applications to option pricing. *Ann. Appl. Probab.*, 15(4):2681–2705.

[34] Kloeden, P. E. and Platen, E. (1992). *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York).* Springer-Verlag, Berlin.

[35] Kurtz, T. G. and Protter, P. (1991). Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs. In *Stochastic analysis*, pages 331–346. Academic Press, Boston, MA.

[36] Kusuoka, S. (2001). Approximation of expectation of diffusion process and mathematical finance. In *Taniguchi Conference on Mathematics Nara '98*, volume 31 of *Adv. Stud. Pure Math.*, pages 147–165. Math. Soc. Japan, Tokyo.

[37] Kusuoka, S. (2004). Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus. In *Advances in mathematical economics. Vol. 6*, volume 6 of *Adv. Math. Econ.*, pages 69–83. Springer, Tokyo.

[38] Lemaire, V., , and Pagès, G. (2014). Multilevel Richardson-Romberg extrapolation. *Preprint arXiv:1401.1177v3.*

[39] Lyons, T. and Victoir, N. (2004). Cubature on Wiener space. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041):169–198. Stochastic analysis with applications to mathematical finance.

[40] Lyons, T. J. (1998). Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310.

[41] Malliavin, P. (1997). *Stochastic analysis*, volume 313 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin.

[42] Milʹšteĭn, G. N. (1974). Approximate integration of stochastic differential equations. *Teor. Verojatnost. i Primenen.*, 19:583–588.

[43] Ninomiya, M. and Ninomiya, S. (2009). A new higher-order weak approximation scheme for stochastic differential equations and the Runge-Kutta method. *Finance Stoch.*, 13(3):415–443.

[44] Ninomiya, S. and Victoir, N. (2008). Weak approximation of stochastic differential equations and application to derivative pricing. *Appl. Math. Finance*, 15(1-2):107–121.

[45] Oshima, K., Teichmann, J., and Velušček, D. (2012). A new extrapolation method for weak approximation schemes with applications. *Ann. Appl. Probab.*, 22(3):1008–1045.

[46] Pagès, G. (2007). Multi-step Richardson-Romberg extrapolation: remarks on variance control and complexity. *Monte Carlo Methods Appl.*, 13(1):37–70.

[47] Pham, H. (2009). *Continuous-time stochastic control and optimization with financial applications*, volume 61 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin.

[48] Platen, E. and Wagner, W. (1982). On a Taylor formula for a class of Itô processes. *Probab. Math. Statist.*, 3(1):37–51 (1983).

[49] Quarteroni, A. and Valli, A. (1994). *Numerical approximation of partial differential equations*, volume 23 of *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin.

[50] Rényi, A. (1963). On stable sequences of events. *Sankhyā Ser. A*, 25:293 302.

[51] Revuz, D. and Yor, M. (1999). *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, third edition.

[52] Talay, D. and Tubaro, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991).

[53] Teschl, G. (2012). *Ordinary differential equations and dynamical systems*, volume 140 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI.

[54] Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of financial economics*, 5(2):177–188.

[55] Yan, L. (2005). Asymptotic error for the Milstein scheme for SDEs driven by continuous semimartingales. *Ann. Appl. Probab.*, 15(4):2706–2738.