UNIVERSITÉ DE GENÈVE
Section de mathématiques

FACULTÉ DES SCIENCES
Dr. Gilles VILMART

Algebraic Tools and Multiscale Methods for the Numerical Integration of Stochastic Evolutionary Problems

THÈSE

Présentée à la Faculté des Sciences de l'Université de Genève
pour obtenir le grade de Docteur ès Sciences, mention Mathématiques

par

Adrien LAURENT

de

Flers (France)

Thèse N° 5570

GENÈVE
Atelier d'impression ReproMail
2021

UNIVERSITÉ DE GENÈVE

FACULTÉ DES SCIENCES

DOCTORAT ÈS SCIENCES, MENTION MATHÉMATIQUES

Thèse de Monsieur Adrien LAURENT

intitulée :

«Algebraic Tools and Multiscale Methods for the Numerical Integration of Stochastic Evolutionary Problems»

La Faculté des sciences, sur le préavis de Monsieur G. VILMART, docteur et directeur de thèse (Section de mathématiques), Monsieur M. GANDER, professeur ordinaire (Section de mathématiques), Monsieur T. LELIEVRE, professeur (CERMICS, Ecole des Ponts Paris Tech, Champs sur Marne, France), Monsieur H. Z. MUNTHE-KAAS, professeur (Department of Mathematics, University of Bergen, Bergen, Norway), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 2 juillet 2021

Thèse - 5570 -

Le Doyen

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

À mes parents,

À mes grand-parents,

Avec toute mon affection.

Abstract

The aim of the work presented in this thesis is the construction and the study of numerical integrators in time to solve stochastic differential equations (SDEs) and stochastic partial differential equations (SPDEs). More precisely, we are interested in the convergence of the methods, in the weak sense and for sampling the invariant measure in the case of ergodic dynamics, in the geometric properties of the integrators, paying particular attention to the preservation of invariants and constraints, as well as to their robustness in the case of multiscale problems.

First, we study ergodic SDEs in large dimension, with an emphasis on Langevin dynamics in $\mathbb{R}^d$ or constrained on manifolds, the study of which is central in molecular dynamics. Inspired by several recent works that proposed examples of high-order numerical schemes for the invariant measure with low weak order, we propose a new formalism called exotic aromatic B-series, based on the popular tool that are the Butcher series and their recent generalization with aromatic trees, and prove that it is suitable for the computation of order conditions for the invariant measure for a class of Runge-Kutta methods. In particular, we introduce a new kind of edge, named liana, that allows us to represent isometric equivariant differential operators such as the Laplacian, and reveals a convenient tool for the computation of order conditions. We emphasize that the proposed tree formalism does not depend on the dimension of the problem. We translate into simple operations on graphs the tedious calculations necessary for the Talay-Tubaro expansion, for the integration by parts of differential operators against the invariant measure, as well as for the composition of operators.

Secondly, we apply the exotic aromatic B-series formalism to the effective computation of order conditions for the sampling of the invariant measure of overdamped Langevin dynamics in $\mathbb{R}^d$ or constrained on a manifold. The methodology is valid for any order and does not grow in complexity when the dimension of the problem increases. We extend this methodology in the context of $\mathbb{R}^d$ to the computation of order conditions for partitioned methods and to the use of postprocessors or non-reversible perturbations. We also present the conditions of order up to order three in the weak sense, as well as for the invariant measure. For the dynamics constrained on a manifold, we obtain the conditions of order two for a new class of methods of Runge-Kutta type. To illustrate the analysis, we present several examples of order two methods for solving constrained problems, as well as the first method of order two for the sampling of constrained Langevin dynamics that does not reduce to a splitting method. We confirm the theoretical results with numerical experiments on the sphere, the torus and the special linear group.

Thirdly, we study highly-oscillatory S(P)DEs whose oscillation is driven by a white noise. These equations appear naturally in the context of optical fibers via the spatial discretization of the non-linear Schrödinger equation with dispersive white noise. In this context, applying a standard method has a cost and a precision depending directly on the high frequency of the

oscillations. We introduce the concept of revolution times, a generalization of the period of the oscillations in the stochastic context, and, following the idea of deterministic multirevolution schemes, we apply it to construct a new second-order integrator whose cost and precision are independent of the high frequency of the oscillations. We also present a geometric variant that preserves quadratic invariants, and we test our schemes on stiff Kubo-type oscillators, as well as on the nonlinear Schrödinger equation with dispersive white noise.

Finally, we give an overview of the new results introduced in this thesis, as well as an outlook on future work. In line with the new ideas and tools introduced in this work, we present in particular a recent work on a uniformly accurate (UA) method for solving penalized Langevin dynamics.

Résumé de la thèse

Le but des travaux présentés dans cette thèse est la construction et l'étude d'intégrateurs numériques en temps pour résoudre des équations différentielles stochastiques (EDS) et des équations aux dérivées partielles stochastiques (EDPS). Plus précisément, nous nous intéressons à la convergence des méthodes, au sens faible et pour l'échantillonnage de la mesure invariante dans le cas de dynamiques ergodiques, aux propriétés géométriques de nos intégrateurs, en portant une attention particulière à la préservation des invariants et des contraintes, ainsi qu'à leur robustesse dans le cas de problèmes multiéchelles.

Dans un premier temps, nous étudions des EDS ergodiques en grande dimension, en mettant l'accent sur des dynamiques de Langevin dans $\mathbb{R}^d$ ou contraintes sur des variétés, dont l'étude est centrale en dynamique moléculaire. Plusieurs travaux récents ayant montré qu'il est possible de créer des schémas numériques d'ordre élevé pour la mesure invariante et d'ordre faible bas, nous proposons un nouveau formalisme nommé exotic aromatic B-series, basé sur l'outil populaire que sont les Butcher series et leur récente généralisation avec les aromatic trees, et nous prouvons qu'il est adapté au calcul de conditions d'ordre pour la mesure invariante pour une classe de méthodes de Runge-Kutta. En particulier, nous introduisons un nouveau type d'arête, appelé liane, qui permet de représenter des opérateurs différentiels isometric-équivariants tel le laplacien, et se révèle être un outil efficace pour le calcul de conditions d'ordre. Nous soulignons que le formalisme d'arbres proposé ne dépend pas de la dimension du problème. Nous traduisons en opérations simples sur des graphes les calculs fastidieux nécessaires au développement de Talay-Tubaro, à l'intégration par parties d'opérateurs différentiels contre la mesure invariante, ainsi qu'à la composition d'opérateurs.

Dans un second temps, nous appliquons le formalisme des exotic aromatic B-series au calcul effectif de conditions d'ordre pour l'échantillonnage de la mesure invariante de dynamiques de Langevin overdamped dans $\mathbb{R}^d$ ou contraintes sur une variété. La méthodologie est valable à tout ordre et ne gagne pas en complexité lorsque la dimension du problème augmente. Nous étendons cette méthodologie dans le contexte de $\mathbb{R}^d$ au calcul des conditions d'ordre pour des méthodes partitionnées et à l'utilisation de postprocesseurs ou de perturbations non-réversibles. Nous présentons également les conditions d'ordre jusqu'à l'ordre trois au sens faible, ainsi que pour la mesure invariante. Pour les dynamiques contraintes sur une variété, nous obtenons les conditions d'ordre deux pour une nouvelle classe de méthodes de type Runge-Kutta. Afin d'illustrer l'analyse, nous présentons plusieurs exemples de méthodes d'ordre deux pour résoudre des problèmes avec contraintes, ainsi que la première méthode d'ordre deux pour l'échantillonnage de dynamiques de Langevin contraintes qui ne se réduit pas à une méthode de splitting. Nous confirmons les résultats théoriques par des expériences numériques sur la sphère, le tore et le groupe spécial linéaire.

Dans un troisième temps, nous étudions des ED(P)S hautement oscillantes dont l'oscillation

est guidée par un bruit blanc. Ces équations apparaissent naturellement dans le contexte de fibres optiques via la discrétisation spatiale de l'équation de Schrödinger non linéaire avec bruit blanc dispersif. Dans ce contexte, l'application d'une méthode standard a un coût et une précision dépendant directement de la haute fréquence des oscillations. Nous introduisons le concept de temps de révolution, une généralisation de la période des oscillations dans le contexte stochastique, et, en suivant l'idée des schémas multirévolution déterministes, nous l'appliquons pour construire un nouvel intégrateur d'ordre deux dont le coût et la précision sont indépendants de la haute fréquence des oscillations. Nous présentons également une variante géométrique préservant les invariants quadratiques, et nous testons nos schémas sur des oscillateurs raides de type Kubo, ainsi que sur l'équation de Schrödinger non linéaire avec bruit blanc dispersif.

Finalement, nous donnons un résumé des nouveaux résultats introduits dans cette thèse, ainsi que quelques perspectives sur des futurs travaux. Dans la lignée des nouvelles idées et outils introduits dans ce travail, nous présentons en particulier un travail récent sur une méthode uniformément précise pour résoudre des dynamiques de Langevin pénalisées.

Remerciements

Une particule lancée seule à travers l'espace, sans subir aucune force, suit une trajectoire rectiligne, optimale et parfaitement ennuyeuse. Cette même particule, soumise à des contraintes et lancée au milieu d'autres particules, a une trajectoire chaotique. Elle accélère, ralentit, se perd, orbite un instant autour d'une particule, repart et se promène ainsi dans tout l'espace qui lui est alloué, allant de rencontre en découverte. Fort heureusement, ma thèse a suivi la seconde trajectoire, et j'aimerais profiter des quelques lignes suivantes pour remercier toutes les fantastiques particules qui ont rendu ces dernières années inoubliables.

Mes premiers remerciements vont bien entendu vers mon maître de thèse Gilles Vilmart. Merci pour ta constante disponibilité, ton écoute attentive, tes innombrables conseils, tes encouragements et ta confiance. Merci d'avoir partagé avec moi ta passion pour l'analyse numérique stochastique et de m'en avoir enseigné les outils, tout en me laissant libre de m'intéresser à ce que je voulais. Merci pour les voyages et les conférences, ainsi que pour ton soutien dans les choix et démarches venant après la thèse. Merci surtout pour ton amitié. Je suis très fier d'avoir fait ma thèse sous ta direction.

Ma gratitude s'adresse ensuite à Martin Gander, Hans Z. Munthe-Kaas et Tony Lelièvre qui m'ont fait l'honneur d'être membres de mon jury de soutenance. Je les remercie en particulier pour l'intérêt qu'ils ont porté à ce travail, pour leur bienveillance et pour leurs précieux retours.

Merci aux chercheurs qui ont croisé mon chemin pendant la thèse et qui m'ont beaucoup appris. Je remercie Hans Munthe-Kaas et Olivier Verdier pour leurs conseils et références sur une version précédente des travaux présentés dans les chapitres 3 et 4 de cette thèse, ainsi que Georg Gottwald pour nos stimulantes discussions sur le contenu du chapitre 5. Un grand merci à Jialin Hong, Julien Vovelle, Cui Jianbo, Hans Z. Munthe-Kaas et Charles-Edouard Bréhier pour leurs invitations à Pékin, Lyon, Atlanta, au séminaire CODYSMA et au One-World Seminar. Merci à Martin Gander et Bart Vandeyrecken pour leur bienveillance et leur implication pour rendre l'équipe d'analyse numérique de Genève si accueillante. Merci à Assyr Abdulle, David Cohen, Arnaud Debussche, Kurusch Ebrahimi-Fard, Ludovic Goudenège, Ernst Hairer, Hélène Hivert, Greg Pavliotis, Gabriel Stoltz, Gerhard Wanner et Peter Wittwer pour les discussions que nous avons partagées. Merci aussi à tous mes professeurs. Chacun de vous m'a transmis un peu de curiosité, de passion et surtout d'amusement. Comment aurais-je pu imaginer en 2013 que l'introduction aux Butcher-series vue lors de ce dernier cours d'équations différentielles avec Philippe Chartier m'entraînerait dans une si exaltante aventure ?

Dans l'entourage du mathématicien, il y a de nombreux collègues indispensables œuvrant chaque jour de leur façon propre à la vie du chercheur. Merci aux bibliothécaires de la section de mathématiques pour m'avoir prêté tant de vos trésors. Merci aux secrétaires Annick, Joselle et Nathalie pour leur soutien autant administratif que moral, ainsi que pour avoir trouvé des fonds pour le séminaire des doctorants. Merci à toute l'équipe du cluster BAOBAB de Genève

pour leur aide à lancer mes calculs sur des milliers de processeurs, et à Peter Pasche pour s'être occupé de nos machines linux. Merci enfin au Fond National Suisse, pour avoir financé ma thèse et mes voyages.

Au début de mon doctorat, j'ai entendu cette rumeur disant qu'on ne peut pas être bon chercheur et bon enseignant en même temps. J'ai eu à cœur de contredire cette affirmation en donnant autant d'énergie à la recherche, à l'enseignement et à la vulgarisation, trop souvent oubliée. En conséquence, l'enseignement fut bien plus qu'une charge obligatoire fixée par mon contrat, ce furent des heures de joie et d'aventure ! Je garderai surtout un excellent souvenir du cours de master d'intégration numérique géométrique. Parmi mes étudiants, je remercie particulièrement Adrien (Junior), Agnès, Léo M., Roberto pour la grimpe, Alessandro, Massimiliano, Michaël, Patricia pour les pauses scopa/durak/jas, et Laurent, Louis, Lucie, Vera pour toutes les autres pauses, les projets et jeux partagés ensemble. À tous mes étudiant(e)s, merci pour votre présence, votre bonne humeur, vos retours bienveillants, et vos questions innombrables. Vous m'avez aussi appris beaucoup.

Je remercie toutes les personnes qui m'ont laissé m'amuser avec leurs joujoux mathématiques et m'ont initié aux joies de la vulgarisation mathématique. Un immense merci à Fathi pour tes 50 conseils pour donner un exposé, pour m'avoir fait jouer dans ta première pièce de théâtre mathématique et m'avoir aidé à écrire la mienne. Merci aussi pour m'avoir fait entrer dans l'aventure MATh.en.JEANS avec Sabrina et Murielle. Merci à l'équipe du Mathscope et en particulier à Élise, Joana, Pierre-Alain et Shaula pour m'avoir fait découvrir les plaisirs de la vulgarisation mathématique, et pour avoir encouragé ma créativité. Merci pour la joie de construire le MegaMenger, pour votre enthousiasme, et votre soutien pour mon projet d'escape game mathématique. Pour l'escape, ce fut vraiment une expérience extraordinaire de gérer notre équipe de bricoleuses de grand talent : Audrey, Kamila, Laura, Maëllie et Syakira. Merci spécial à Léo (Tanuki) pour toute l'énergie que tu as donné à cet escape.

Ma gratitude va aussi à tous les formidables amis (non déjà cités au dessus) qui m'ont accompagné pendant ces années ! Merci aux rôlistes Conor, Léa, Michal et Thibaut pour toutes les aventures fantastiques, surtout pendant les confinements ! Merci à l'éternelle team Ker Lann et aux amis de Rennes : Alexandre, Anne-Élisabeth, Baptiste, Benoît, Clément, Corentin, Emily, Kévin, Léa, Manon, Marine, Paul, Thibaut, Yann, pour votre aide, vos invitations et votre amitié. Merci Grégoire, pour tes invitations à Paris et nos passionnantes discussions statistiques. Merci à Florian et Joackim de m'avoir invité à parler sur Rennes, merci à Nicolas de m'avoir invité à Dijon et Lise de m'avoir invité à parler au séminaire du LJLL. Merci à mes frères de thèse Eugen, Guillaume, Ibrahim et Nicolas M. pour le petit groupe sympathique que nous avons formé. Merci aux amis du YMCA de Genève, et en particulier à Margherita, pour leur immense soutien dans mes premiers temps sur Genève. Merci à tous les amis que je me suis fait un peu partout dans le monde et sans qui les conférences auraient été beaucoup moins drôles. Merci enfin aux amis doctorants et postdocs de Genève et en particulier à Aitor, Arina, Benjamin, Ding, Eiichi, Fayçal, Giancarlo, Grégoire, Jhih-Huang, Julian, Justine, Louis-Hadrien, Nicolas H., Pablo, Parisa, Pascaline, Pierre, Pratik, Raphaël B., Raphaël D., Renaud, Rik, Solenn, Tommaso, Virginie, Yaroslav et tous ceux que j'aurais pu oublier, pour votre soutien, mais aussi pour les jeux, les soirées, les balades,... la vie quoi !

Je remercie tout particulièrement les fantastiques personnes avec qui j'ai partagé le bureau 610c, avant le déménagement, et le bureau 3-04 dans la nouvelle section de Mathématiques de Genève. Merci Sandie pour ta bonne humeur et ton énergie. Ta présence irradiait le bureau de motivation ! Merci Marco pour les balades, les cafés, les livres, films, comics et aussi pour m'avoir suivi aveuglément dans des expériences théâtrales douteuses. Ton application dans

le travail et ta persévérance resteront des exemples pour moi. Merci Ausra pour ta présence délirante, pour les jeux, et pour ton support qui m'a tant aidé dans les derniers mois de cette thèse.

Merci enfin à ma chère amie Christine. Tu entres dans nombre des catégories citées au dessus, mais un remerciement spécial s'impose ! Merci pour ta motivation, ta créativité sans fin, ta bonne humeur, ta gentillesse, ton swing, tes jeux, ta rixme, tes déguisements, tes escapes... Les maths à Genève auraient été bien moins amusantes sans toi !

Merci à ma famille, à Papa, Maman, Simon et Camille qui m'ont toujours tant encouragé, et ce dans toutes mes entreprises. Merci à mes grands-pères bricoleurs et mes grands-mères aimantes. Cette thèse est en particulier dédiée à pépé André et mamie Marthe, décédés trop tôt pour lire ces lignes, et qui sont toujours restés humblement curieux de tout, même sans avoir eu la possibilité de faire de longues études. Merci à Marie-Agnès, Ninette et Pierre qui m'ont beaucoup accueilli sur Lyon pendant les premiers temps. De tout cœur, MERCI !

Enfin, merci à toi, lecteur, lectrice, de donner de ton temps pour lire cette thèse. Si, par mégarde, tu te perds dans la jungle luxuriante dans laquelle tu t'aventures en cet instant, sache que je dispense mon aide et ma passion sans limites, et que je me ferai un plaisir de répondre à n'importe laquelle de tes interrogations. Bonne lecture à toi !

Pour conclure ces remerciements, j'aimerais rappeler la place importante qu'a ce court chapitre dans cette thèse. En effet, il peut sembler y avoir bien peu de liens entre la vie du mathématicien et la rigueur des théorèmes qu'il produit. Mais c'est une grossière erreur. Les résultats présentés dans ce document ne sont pas apparus magiquement dans une étincelle de compréhension. Ils ont évolué, ont été modifié plusieurs fois pour aboutir à un résultat satisfaisant. Et ce lent processus de mûrissement n'a été possible que grâce à toutes les personnes qui m'ont encouragé, m'ont aidé, ont pointé les faiblesses ou les qualités d'un résultat, ont mentionné une citation ou m'ont juste proposé une pause café. Et en cela, cette thèse contient un peu de vous tous. Merci !

Contents

CHAPTER 1

Introduction and main results

The aim of the work presented in this thesis is the construction and study of numerical integrators for solving stochastic differential equations (SDEs) and stochastic partial differential equations (SPDEs). In particular, we focus on the accuracy, both in the weak sense and for sampling the invariant measure in the case of ergodic dynamics, and on the geometric properties of our integrators, with an emphasis on the preservation of invariants and constraints.

The thesis is organized as follows. In Chapter 2, we present a concise introduction to the standard equations, assumptions, notations and tools we use. The new results of the thesis, published in [83, 84, 85], are presented in Chapter 3, Chapter 4 and Chapter 5. In Chapter 6, we summarize the new results, and give an outlook on future work. In particular, we present a recent work [82] in preparation on a uniformly accurate (UA) integrator for penalized Langevin dynamics, that uses a handful of the new ideas introduced in this thesis. We give a detailed overview of the three main contributions of this work below, and more details in the sections that follow.

Exotic aromatic B-series for computing order conditions (Chap. 3): We introduce a new algebraic framework based on a modification (called exotic) of aromatic Butcher-series for the systematic study of the accuracy of numerical integrators for the invariant measure of overdamped Langevin dynamics in $\mathbb{R}^d$ and on manifolds. In particular, this formalism allows us to conveniently compute and compose the operators in the Talay-Tubaro expansion and to integrate them by parts with respect to the invariant measure. We also show that the introduced exotic aromatic B-series satisfy an isometric equivariance property.

High order integrators for the invariant measure (Chap. 4): We derive a new methodology for the construction of high order integrators for sampling the invariant measure of ergodic stochastic differential equations with dynamics constrained on a manifold. We obtain the order conditions for sampling the invariant measure for a class of Runge-Kutta methods applied to the overdamped Langevin equation in $\mathbb{R}^d$ and constrained on a manifold. The proposed analysis is valid for arbitrarily high order, includes the cases of partitioned methods and postprocessed methods in $\mathbb{R}^d$, and relies on the exotic aromatic Butcher-series formalism. To illustrate the methodology, a new method of order two for constrained dynamics is introduced, and numer-

ical experiments on the sphere, the torus and the special linear group confirm the theoretical findings.

Multirevolution integrators for highly-oscillatory SDEs (Chap. 5): We introduce a new methodology based on the multirevolution idea for constructing integrators for stochastic differential equations in the situation where the fast oscillations themselves are driven by a Stratonovich noise. Applications include in particular highly-oscillatory Kubo oscillators and spatial discretizations of the nonlinear Schrödinger equation with fast white noise dispersion. We construct a method of weak order two with computational cost and accuracy both independent of the stiffness of the oscillations. A geometric modification that conserves exactly quadratic invariants is also presented.

1.1 Exotic aromatic B-series for the computation of order conditions

The aim of this work is to introduce a new unified algebraic framework based on aromatic trees and B-series for the systematic study of the order conditions for the invariant measure of a class of numerical integrators that includes Runge-Kutta type schemes for overdamped Langevin dynamics in $\mathbb{R}^d$ of the form

$$dX(t) = f(X(t))dt + \sigma dW(t), \quad X(0) = X_0, \tag{1.1.1}$$

where the solution $X(t)$ is in $\mathbb{R}^d$, $\sigma > 0$, $f = -\nabla V$ with $V\colon \mathbb{R}^d \to \mathbb{R}$ a smooth potential and $W(t)$ is a standard d-dimensional Wiener process. Equations of the form (1.1.1) typically appear when modeling the motion of particles in a fluid in a high friction regime. The dimension d of the problem represents the degrees of freedom of the physical system and is ideally of the order of the Avogadro number $N_A \sim 10^{24}$, but at most of the order of 10^7 in current practical calculations. If the system of particles is subject to constraints, such as fixed angles in molecules or strong covalent bonds between atoms, we get constrained Langevin dynamics of the form

$$dX(t) = \Pi_{\mathcal{M}}(X(t))f(X(t))dt + \sigma\Pi_{\mathcal{M}}(X(t)) \circ dW(t), \quad X(0) = X_0 \in \mathcal{M}, \tag{1.1.2}$$

where $\mathcal{M} = \{x \in \mathbb{R}^d, \zeta(x) = 0\}$ is a smooth manifold, ζ is a smooth constraint, and $\Pi_{\mathcal{M}}$ is the orthogonal projection on the tangent bundle of $\mathcal{M}$. In this context, the noise term is of Stratonovich type, of dimension d, and projected on the manifold. To simplify the analysis, we assume that $\mathcal{M}$ is a compact manifold of codimension one. Hence $\zeta\colon \mathbb{R}^d \to \mathbb{R}$ is a scalar constraint and $\Pi_{\mathcal{M}} = I_d - G^{-1}gg^T$ with $g = \nabla\zeta$ and $G = g^Tg$. The new formalism we present is also suitable for computing order conditions for the constrained dynamics (1.1.2).

Sampling from overdamped Langevin dynamics in $\mathbb{R}^d$ or on manifolds allows us to compute integrals in high dimension of the form $\int_{\mathbb{R}^d} \phi\rho_\infty dx$ or $\int_{\mathcal{M}} \phi\rho_\infty d\sigma_{\mathcal{M}}$, where ϕ is a smooth test function, ρ_∞ is the density of the invariant measure and $d\sigma_{\mathcal{M}}$ is the canonical measure on the manifold $\mathcal{M}$ induced by the Euclidean metric of $\mathbb{R}^d$. We refer to the textbook [90], and references therein, for further details on the physical applications of sampling Langevin dynamics. In this work, we introduce a new formalism of B-series for studying the accuracy of Runge-Kutta integrators for sampling the invariant measure of (1.1.1) or (1.1.2). The numerical applications of this work are presented in Chapter 4, and include the order conditions up to order three for a class of Runge-Kutta methods, with extensions to partitioned methods,

the use of postprocessors and of non-reversible perturbations in the context of $\mathbb{R}^d$. We also introduce a new class of Runge-Kutta methods for sampling (1.1.2), with the order conditions up to order two, and several examples of order two methods for solving constrained problems.

In [5, 6], a methodology for the analysis and design of high order integrators for the invariant measure is introduced in the context of $\mathbb{R}^d$. First, we assume that the numerical integrator (X_n) has a weak Taylor expansion, called the Talay-Tubaro expansion [121], of the form

$$\mathbb{E}[\phi(X_1)|X_0 = x] = \phi(x) + \sum_{j=1}^{N} h^j \mathcal{A}_{j-1}\phi(x) + \dots,$$

where h is the timestep and the $\mathcal{A}_j$'s, $j = 0, 1, 2, \dots$, are linear differential operators with coefficients depending smoothly on f, g and their (high order) derivatives (and depending on the choice of the integrator). If we denote by $\mathcal{L}$ the generator of the SDE, then, under technical assumptions, the integrator has at least weak order r if

$$\mathcal{A}_{j-1} = \mathcal{L}^j/j!, \qquad j = 1, \dots, r.$$

We use the following result for building high-order integrators for the invariant measure. It was first introduced in [5] in the context of $\mathbb{R}^d$, and we generalize it in the context of manifolds in Section 4.2.

Theorem 1.1.1 (See Sect. 2.3 and Theorem 2.3.2)**.** *Under technical assumptions, if the numerical scheme is consistent (that is, if $\mathcal{A}_0 = \mathcal{L}$), and if it satisfies,*

$$\mathcal{A}_j^* \rho_\infty = 0, \quad j = 1, \dots, p-1, \tag{1.1.3}$$

where $\mathcal{A}_j^*$ is the adjoint of the operator $\mathcal{A}_j$ in $L^2(d\sigma_{\mathcal{M}})$, then it has order p for the invariant measure.

The formalism of exotic aromatic B-series is convenient for deriving the operators $\mathcal{A}_j$ with forests, and to integrate them by parts with respect to the invariant measure, in order to find conditions to satisfy (1.1.3).

1.1.1 The formalism of exotic aromatic B-series

Originally introduced by Hairer and Wanner in [66], and based on the work of Butcher [24], B-series have proved to be a powerful tool for the numerical analysis of deterministic differential equations. In the stochastic context, several works extended B-series for the computation of order conditions for strong and weak approximations on a finite time interval, but no such algebraic framework was known for the accuracy of sampling the invariant measure of ergodic dynamics. In particular, the first stochastic trees and B-series were introduced in [20, 21, 78], and were later followed by the works of Rößler [110, 111, 112, 113, 114] and Debrabant and Kværnø [46, 45, 47]. It is known for large classes of SDEs that a scheme of weak order p is automatically at least of order p for the invariant measure. It is however possible to create integrators with a high order p for the invariant measure and a low weak order of accuracy (see, for instance, [14, 86, 87, 6] in the context of Langevin dynamics). Inspired by this possibility, we define a new B-series formalism for the computation of order conditions for the invariant measure, based on the aromatic B-series from [36], the standard P-series and with the graph approach of [13], that is suitable for computing specifically the order conditions for the invariant

measure. The root of the forests (always drawn at the bottom of the rightmost tree) represents the test function ϕ, while the other black nodes represent f, as in the following examples

$$F(\bullet)(\phi) = \phi' f, \qquad F(\circ\!\!<\!\!\bullet)(\phi) = \mathrm{div}(f)\phi''(f,f), \qquad F\Big(\circ\!\!<\!\!\bullet\;\bullet\Big)(\phi) = \sum_{i,j=1}^{d} \partial_j f_i \partial_i f_j \phi' f' f.$$

In the context of manifolds, we add white nodes that represent the gradient of the constraint g, as in the following examples,

$$F(\circ\!\!<\!\!\bullet\;\bullet)(\phi) = \sigma^2 G^{-1}\,\mathrm{div}(g)\phi' g' f, \qquad F\Big(\bullet\!\!<\!\!\bullet\;\bullet\Big)(\phi) = \sigma^2 G^{-1}\phi''(g, g' f),$$

where we apply a renormalization by the constant $\sigma > 0$ and the map G that simplifies the expressions, to be explained in Section 3.2.2. Then, we introduce a new kind of edge, called liana, that, in particular, allows us to represent the Laplacian, and double edges for representing the scalar product. These new edges shall prove strong tools for the computation of order conditions for the invariant measure, and allow us to represent more complicated differentials as the following,

$$F(\circ)(\phi) = \sum_{i=1}^{d} \phi''(e_i, e_i) = \sigma^2 \Delta\phi,$$

$$F(\bullet\!\!-\!\!\circ\;\bullet)(\phi) = \sigma^4 G^{-2}(g, g' g) \sum_{i=1}^{d} \phi''(g'(e_i), e_i),$$

$$F\Big(\bullet\!\!-\!\!\bullet\;\circ\Big)(\phi) = \sigma^2 G^{-1}(g, f)(\Delta\phi)' g,$$

where (e_i) is the canonical basis of $\mathbb{R}^d$. We call these graphs the exotic aromatic forests, gather them in the set $\mathcal{EAT}$, and denote $|\gamma| \in \mathbb{N}$ their order. We refer the reader to Section 3.2 for the detailed definition of the exotic aromatic forests, and mention that the forests do not depend on the dimension of the problem d. Note that the generator $\mathcal{L}$ in the manifold case can be written with exotic aromatic forests as

$$\mathcal{L}\phi = F\Big(\bullet - \bullet\!\!-\!\!\bullet - \frac{1}{2}\,\circ\!\!<\!\!\bullet\;\bullet + \frac{1}{2}\,\bullet\!\!-\!\!\circ\;\bullet + \frac{1}{2}\,\circ - \frac{1}{2}\,\bullet\!\!<\!\!\bullet\Big)(\phi).$$

In $\mathbb{R}^d$, we have $g = 0$, so that the trees with white nodes vanish and $\mathcal{L}$ is given by

$$\mathcal{L}\phi = F\Big(\bullet + \frac{1}{2}\,\circ\Big)(\phi) = \phi' f + \frac{\sigma^2}{2}\Delta\phi.$$

If we restrict ourselves to black nodes, then we show in Section 3.2.4 that the exotic aromatic B-series satisfy an isometric equivariance property in the spirit of [105, 100]. This property of equivariance means that the method is transparent with respect to applying an affine coordinate transformation (see the precise definition in Section 3.2.4).

1.1.2 Computation of order conditions with exotic aromatic forests

We apply the new formalism for computing the order conditions on Runge-Kutta methods for solving (1.1.1) or (1.1.2) (see Section 3.1 for the details on the class of Runge-Kutta methods). In Section 3.3.1, we use the Isserlis theorem [73] in the field of combinatorics to prove that computing the expectation of a B-series amounts to link crosses together by pairs with lianas. In particular, we deduce that any such stochastic Runge-Kutta method satisfies an expansion in exotic aromatic B-series.

Theorem 1.1.2 (See Sect. 3.3.1 and Theorem 3.3.3)**.** *For a Runge-Kutta method with initial condition $X_0 = x$ assumed deterministic for simplicity, the weak approximations can be formally written with exotic aromatic B-series, that is, there exist two maps e, $a\colon \mathcal{EAT} \to \mathbb{R}$ such that*

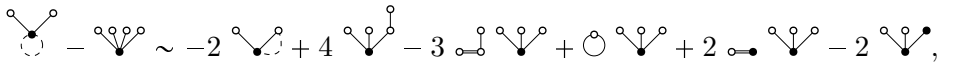
$$\mathbb{E}[\phi(X(h))] = \sum_{\gamma\in\mathcal{EAT}} h^{|\gamma|} e(\gamma) F(\gamma)(\phi)(x),$$
$$\mathbb{E}[\phi(X_1)] = \sum_{\gamma\in\mathcal{EAT}} h^{|\gamma|} a(\gamma) F(\gamma)(\phi)(x),$$

and where the operators in the Talay-Tubaro expansion are given by

$$\frac{\mathcal{L}^j}{j!} = F\Big(\sum_{\gamma\in\mathcal{EAT},|\gamma|=j} e(\gamma)\gamma\Big), \quad \mathcal{A}_{j-1} = F\Big(\sum_{\gamma\in\mathcal{EAT},|\gamma|=j} a(\gamma)\gamma\Big).$$

If $e(\gamma) = a(\gamma)$ for all $\gamma \in \mathcal{EAT}$ with $1 \leqslant |\gamma| \leqslant p$, then the integrator has at least weak order p.

According to Theorem 1.1.1, we integrate by parts the integrators $\mathcal{A}_j$ to obtain the order conditions such that (1.1.3) is satisfied. In Section 3.3.2, we show that the integration by parts process can be written as a direct operation of plugging/unplugging edges on exotic aromatic forests. We represent this operation on exotic aromatic forests with the equivalence symbol $\sim$. For instance, we have the equivalence of the linear combinations of forests



which represents the following integration by parts calculation,

$$\begin{aligned}\int_{\mathcal{M}} \Big[\sigma^4 G^{-1}\Delta\phi''(g,g) - \sigma^4 G^{-2}\phi^{(4)}(g,g,g,g)\Big]\rho_\infty d\sigma_{\mathcal{M}} \\ = \int_{\mathcal{M}} \Big[-2\sigma^4 G^{-1}\sum_i \phi^{(3)}(g,\partial_i g, ei) + 4\sigma^4 G^{-2}\phi^{(3)}(g,g,g'g) \\ - 3\sigma^4 G^{-3}(g,g'g)\phi^{(3)}(g,g,g) + \sigma^4 G^{-2}\operatorname{div}(g)\phi^{(3)}(g,g,g) \\ + 2\sigma^2 G^{-2}(g,f)\phi^{(3)}(g,g,g) - 2\sigma^2 G^{-1}\phi^{(3)}(g,g,f)\Big]\rho_\infty d\sigma_{\mathcal{M}}.\end{aligned}$$

We then deduce the following result, that allows us to compute the order conditions for the invariant measure in $\mathbb{R}^d$ or on manifolds, that we present in Chapter 4.

Theorem 1.1.3 (See Sect. 3.3.2 and Theorem 3.3.8)**.** *Consider a consistent ergodic Runge-Kutta method. We denote $\mathcal{A}_i = F(\gamma_i)$ with $\gamma_i \in \mathcal{EAT}$. If $\gamma_i \sim \gamma_i^0$ and $F(\gamma_i^0) = 0$ for $1 \leqslant i < p$, then the method has at least order p for the invariant measure.*

In addition, we present in Section 3.3.3 a new unpublished methodology for composing differential operators via exotic aromatic forests. It allows us to conveniently compute $\mathcal{L}^j/j!$ for getting weak order conditions, and the Lie bracket $[\mathcal{L},\mathcal{A}] = \mathcal{LA} - \mathcal{AL}$ for the construction of high order methods with postprocessors.

The exotic aromatic B-series formalism proves to be a crucial tool for the calculations of order conditions for the invariant measure as the number of terms grows exponentially with the desired order. For instance, for the order two in the codimension one manifold case of a class of Runge-Kutta methods, we present in Appendix C the decomposition into exotic aromatic forests of the operator $\mathcal{A}_1$ and the operator $\mathcal{A}_1^0$ obtained after integration by parts of $\mathcal{A}_1$. It

consists of a table that is eight pages long for $\mathcal{A}_1$ and five pages long for $\mathcal{A}_1^0$. The calculations of the order conditions are close to the limit of what can be done by hand, so that the use of a strong formalism for representing and integrating by parts differential operators becomes a necessity.

1.2 High order methods for sampling the invariant measure of ergodic SDEs

We give a methodology, based on the algebraic framework defined in Chapter 3 and on the methodology introduced in [5], that we extend in the context of manifolds, to study integrators of arbitrary high order for sampling the invariant measure of ergodic SDEs in $\mathbb{R}^d$ and on manifolds. We study in particular the following overdamped Langevin dynamics in $\mathbb{R}^d$,

$$dX(t) = f(X(t))dt + \sigma dW(t), \quad X(0) = X_0, \tag{1.2.1}$$

or Langevin dynamics constrained on a manifold $\mathcal{M}$,

$$dX(t) = \Pi_{\mathcal{M}}(X(t))f(X(t))dt + \sigma\Pi_{\mathcal{M}}(X(t)) \circ dW(t), \quad X(0) = X_0 \in \mathcal{M}, \tag{1.2.2}$$

where $\sigma > 0$, $f = -\nabla V$ with $V\colon \mathbb{R}^d \to \mathbb{R}$ a smooth potential with appropriate growth assumptions, $W(t)$ is a standard d-dimensional Wiener process, $\mathcal{M} = \{x \in \mathbb{R}^d, \zeta(x) = 0\}$ is a smooth compact manifold, $\zeta\colon \mathbb{R}^d \to \mathbb{R}$ is a smooth scalar constraint and $\Pi_{\mathcal{M}} = I_d - G^{-1}gg^T$ is the orthogonal projection on the tangent bundle of $\mathcal{M}$ with $g = \nabla\zeta$ and $G = g^Tg$. We study ergodic SDEs, that is, equations where there exists a unique invariant measure $d\mu_\infty$ that has a density ρ_∞ with respect to the Lebesgue measure dx, respectively to $d\sigma_{\mathcal{M}}$, the canonical measure on $\mathcal{M}$ induced by the Euclidean metric of $\mathbb{R}^d$ in the context of manifolds, such that for all test functions ϕ,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \phi(X(t))dt = \int \phi(x)d\mu_\infty(x) \quad \text{almost surely.}$$

We emphasize that in the context of manifolds, $d\mu_\infty$ is singular with respect to the Lebesgue measure on $\mathbb{R}^d$, and the methodology in $\mathbb{R}^d$ does not generalize straightforwardly.

A natural method for building integrators with high order for the invariant measure is to use integrators of high weak order since a method of weak order r has at least order $p \geqslant r$ for the invariant measure. However, it is possible to create computationally cheaper schemes of high order for the invariant measure with low weak order, such as in [14, 86, 87, 6] for the Langevin equation. Inspired by this possibility, we propose a methodology for building integrators of any order for the invariant measure and with low weak order (typically weak order one). We consider Runge-Kutta methods for solving (1.2.1) in the context of $\mathbb{R}^d$ of the form

$$\begin{aligned} Y_i &= X_n + h\sum_{j=1}^{s} a_{ij}f(Y_j) + \sigma\sqrt{h}\sum_{k=1}^{l} d_i^{(k)}\xi_n^{(k)}, \qquad i = 1,\dots,s, \\ X_{n+1} &= X_n + h\sum_{i=1}^{s} b_i f(Y_i) + \sigma\sqrt{h}\xi_n^{(1)}, \end{aligned} \tag{1.2.3}$$

and the following new class of projection methods for solving (1.2.2) on a manifold,

$$Y_i = X_n + h\sum_{j=1}^{s} a_{ij}f(Y_j) + \sigma\sqrt{h}\sum_{k=1}^{l} d_i^{(k)}\xi_n^{(k)} + \lambda_i\sum_{j=1}^{s}\widehat{a}_{ij}g(Y_j), \quad i = 1,\dots,s,$$

$$\begin{aligned} &\zeta(Y_i) = 0 \quad \text{if} \quad \delta_i = 1, \quad i = 1, \dots, s, \\ &X_{n+1} = Y_s, \end{aligned} \tag{1.2.4}$$

where $A = (a_{ij}), \widehat{A} = (\widehat{a}_{ij}) \in \mathbb{R}^{s\times s}$, $b = (b_i) \in \mathbb{R}^s$, $d^{(k)} = (d_i^{(k)}) \in \mathbb{R}^s$, $\delta_i = \sum_{j=1}^s \widehat{a}_{ij} \in \{0, 1\}$ are the given Runge-Kutta coefficients, and $\xi_n^{(k)} \sim \mathcal{N}(0, I_d)$ are independent random vectors, typically Gaussian or discrete with suitable first moments. In the manifold case (1.2.4), we fix $\delta_s = 1$ so that $X_{n+1} \in \mathcal{M}$ and we ask that if $\delta_i = 0$, then $\widehat{a}_{ij} = 0$ for $j = 1, \dots, s$ (internal stages without projection, $Y_i \notin \mathcal{M}$ almost surely). In particular, in the manifold context, the widely used Euler integrators (see, for instance, [37, 88, 90, 91]) can be written in the form (1.2.4). In this context, the Euler integrator with implicit projection direction is given by

$$X_{n+1} = X_n + hf(X_n) + \sigma\sqrt{h}\xi_n + \lambda g(X_{n+1}), \quad \zeta(X_{n+1}) = 0. \tag{1.2.5}$$

In addition to deriving the methodology for the construction of integrators of any order, we calculate the order conditions up to order three of Runge-Kutta methods of the form (1.2.3), and we present extensions to partitioned methods, postprocessed methods and methods using non-reversible perturbations. In the context of manifolds, we give the order two Runge-Kutta conditions for the class of schemes (1.2.4) and we present several examples of methods of order two for constrained problems for illustrating the analysis.

1.2.1 High order sampling integrators in $\mathbb{R}^d$

Using our methodology, we compute the order conditions for Runge-Kutta methods of the form (1.2.3). The conditions of order three are detailed in Theorem 4.3.1.

Theorem 1.2.1 (See Sect. 4.3.2 and Theorem 4.3.1)**.** *Consider an ergodic Runge-Kutta method of the form* (1.2.3) *for solving* (1.2.1)*, then the conditions for order one and two for the invariant measure are given by Table 1.1, with the corresponding exotic aromatic forest.*

Order	Tree τ	$F(\tau)(\phi)$	Order condition
1		$\phi' f$	$\sum b_i = 1$
2		$\phi' f' f$	$\sum b_i c_i - 2\sum b_i d_i = -\frac{1}{2}$
		$\sigma^2 \phi' \Delta f$	$\sum b_i d_i^2 - 2\sum b_i d_i = -\frac{1}{2}$

Table 1.1: Runge-Kutta order conditions for the invariant measure (See Theorem 1.2.1). The sums are over all involved indices.

In the spirit of [123], we also combine our approach with the idea of processing from Butcher [23], to design efficient postprocessed integrators with high order for the invariant measure at a negligible overcost compared to standard low order schemes. At the last step, we apply a costless correction to the integrator, called a postprocessor, and we study this corrected result instead. The following result gives the order conditions of the postprocessed integrator.

Theorem 1.2.2 (See Sect. 4.3.3 and Theorem 4.3.4)**.** *Consider an ergodic Runge-Kutta method of the form* (1.2.3) *and of order $p \geqslant 1$ for sampling the invariant measure of* (1.2.1)*, and the following associated postprocessor*

$$\overline{Y_i} = X_n + h \sum_{j=1}^{s} \overline{a_{ij}} f(\overline{Y_j}) + \overline{d_i}\sigma\sqrt{h}\,\overline{\xi_n}, \qquad i = 1, \ldots, s,$$
$$\overline{X_n} = X_n + h \sum_{i=1}^{s} \overline{b_i} f(\overline{Y_i}) + \overline{d_0}\sigma\sqrt{h}\,\overline{\xi_n}.$$

Under technical assumptions, if the conditions of order two in Table 4.3 are verified, then the postprocessor $\overline{X_n}$ has order two for sampling the invariant measure.

Order	Tree τ	Order condition
2		$\sum b_i c_i - 2\sum b_i d_i - 2\sum \overline{b_i} + 2\overline{d_0}^2 = -\frac{1}{2}$
		$\sum b_i d_i^2 - 2\sum b_i d_i - \sum \overline{b_i} + \overline{d_0}^2 = -\frac{1}{2}$

Table 1.2: Order conditions for Runge-Kutta method with Runge-Kutta postprocessor (See Theorem 1.2.2). The sums are over all involved indices.

Example. *Assuming ergodicity, the following Runge-Kutta method, introduced in [123], is of order two for sampling the invariant measure of* (1.1.1)*,*

$$\begin{aligned} X_{n+1} &= X_n + hf(X_{n+1} + \tfrac{-1+\sqrt{2}}{2}\sigma\sqrt{h}\xi_n) + \sigma\sqrt{h}\xi_n, \\ \overline{X_n} &= X_n + h\tfrac{\sqrt{2}}{2} f(\overline{X_n}) + \tfrac{\sqrt{4\sqrt{2}-1}}{2}\sigma\sqrt{h}\,\overline{\xi_n}. \end{aligned}$$

We refer to Section 4.3 for more applications of the formalism to modified equations (Section 4.3.1), partitioned methods (Section 4.3.4), and the use of non-reversible perturbations (Section 4.3.5).

1.2.2 High order sampling integrators on manifolds

In the context of manifolds, we derive the order two conditions for a method of the form (1.2.4). There are eleven order conditions for the invariant measure if each substage of the method is projected on the manifold, and the conditions are detailed in Theorem 4.4.8. We propose a new order two method of the form (1.2.4) with four stages and the coefficients given in Appendix A. For the toy problem where $\mathcal{M}$ is the sphere in $\mathbb{R}^d$, the order two conditions for the invariant measure with projected substages ($\delta_i = 1$ for $i = 1, \ldots, s$) reduce to

$$\begin{aligned} &\textstyle\sum \widehat{b}_i d_i = \sum b_i d_i = 1 \pm \frac{\sqrt{2}}{2}, \\ &\textstyle\sum b_i c_i = \sum b_i d_i^2 = \sum \widehat{b}_i c_i = 2\sum \widehat{b}_i d_i - \frac{1}{2}, \\ &\textstyle\sum \widehat{b}_i d_i \widehat{a}_{ij} c_j = 2\sum \widehat{b}_i d_i \widehat{a}_{ij} d_j, \end{aligned}$$

and an order two method is given by

$$Y_1 = X_n + h\left(\frac{3}{2} - \sqrt{2}\right) f(X_{n+1}) + \sigma\sqrt{h}\left(1 - \frac{\sqrt{2}}{2}\right)\xi_n + \lambda_1(2Y_1 - X_{n+1}), \quad \zeta(Y_1) = 0,$$

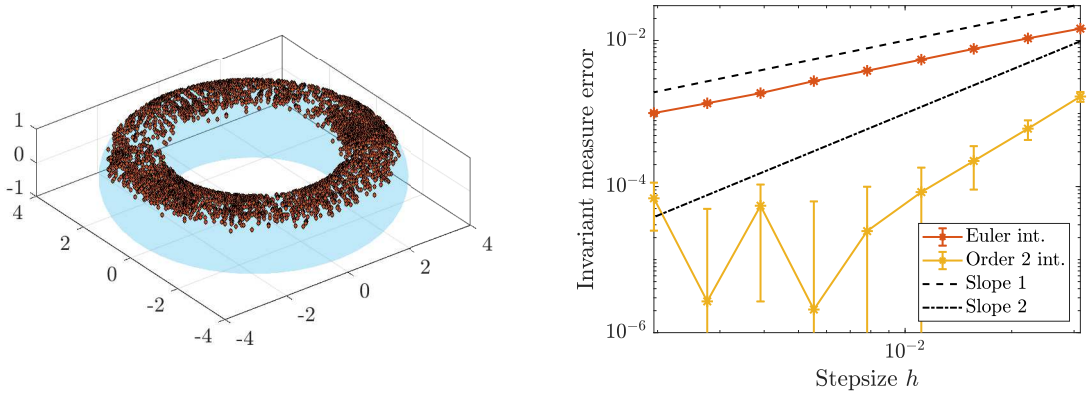


Figure 1.1: (From [85]) A trajectory of the order two method (left) and the convergence curve for the torus for the invariant measure (right) with the potential $V(x) = 25(x_3 - r)^2$, $\phi(x) = x_3^2$, a final time $T = 20$ and $M = 10^7$ trajectories.

$$X_{n+1} = X_n + hf(Y_1) + \sigma\sqrt{h}\xi_n + \lambda_2 Y_1, \quad \zeta(X_{n+1}) = 0.$$

To numerically check the order two of the new Runge-Kutta integrator, we compare it with the Euler scheme (1.2.5) on a torus defined by the constraint $\zeta(x) = (x_1^2 + x_2^2 + x_3^2 + R^2 - r^2)^2 - 4R^2(x_1^2 + x_2^2)$ with $R = 3$ and $r = 1$. In Figure 1.1, we plot the error for the invariant measure versus the timestep h. We observe order two for the proposed integrator, which confirms the theoretical findings. We present further experiments on a sphere and on the special linear group in Section 4.5.

1.3 Multirevolution integrators for SDEs with fast stochastic oscillations

The highly-oscillatory nonlinear Schrödinger equation (NLS) with white noise dispersion is used to describe the propagation of a signal in optical fibers (see [7, 8, 58]),

$$du(t) = \frac{i}{\sqrt{\varepsilon}}\Delta u(t) \circ dW(t) + F(u(t))dt, \quad u(t = 0) = u_0. \tag{1.3.1}$$

The highly-oscillatory behavior ($\varepsilon \ll 1$) appears naturally when observing the propagation in long time with a small nonlinearity or the propagation of a small initial data in an optical fiber with a polynomial nonlinearity. The goal of this work is to develop efficient and cheap numerical methods that can model the propagation of pulses in this context, in order to observe some specific behaviors and, ultimately, to build enhanced fibers.

More precisely, we apply a spatial discretization of the SPDE (1.3.1) to obtain the following class of highly-oscillatory SDEs in $\mathbb{R}^d$ driven by a one-dimensional Stratonovich noise

$$dX(t) = \frac{1}{\sqrt{\varepsilon}}AX(t) \circ dW(t) + F(X(t))dt, \quad t > 0, \quad X(0) = X_0, \tag{1.3.2}$$

where $W(t)$ is a standard one-dimensional Wiener process, the function $F : \mathbb{R}^d \to \mathbb{R}^d$ is a smooth nonlinear map, the stiff parameter $\varepsilon > 0$ is fixed and assumed small, and $A \in \mathbb{R}^{d\times d}$ is a given matrix satisfying $e^A = \mathrm{Id}$ (equivalently A is diagonalizable and has all its eigenvalues in $2i\pi\mathbb{Z}$). In the deterministic setting, this last property yields that the solution $x(t) =$

$\exp(\varepsilon At)x_0$ of $\frac{dx}{dt} = \varepsilon^{-1}Ax$ is ε-periodic. For stochastic oscillations, it means that the solution $X(t) = \exp(\varepsilon^{-1/2}AW(t))X_0$ of $dX = \varepsilon^{-1/2}AX \circ dW$ satisfies $X(T) = X(0)$ for a random time $T = \inf\{t > 0, \varepsilon^{-1/2}\,|W(t)| = 1\}$ of mean ε. The class of SDEs (1.3.2) includes in particular highly-oscillatory Kubo oscillators (see [39])

$$dX(t) = \frac{2\pi}{\sqrt{\varepsilon}}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}X(t)\circ dW(t) + \begin{pmatrix}0 & -a\\ a & 0\end{pmatrix}X(t)dt, \quad a \in \mathbb{R}. \tag{1.3.3}$$

Applying standard SDE integrators to solve equation (1.3.2) requires in general a time stepsize $h \leqslant \varepsilon$ to be accurate, which makes these methods dramatically expensive when ε is small. The goal of this work is to create robust numerical methods, i.e., numerical integrators whose cost and accuracy do not deteriorate when ε becomes small. Several classes of methods have already been developed for highly-oscillatory SDEs with a deterministic fast oscillation (see, for instance, [122]), but not in the case where the stiff oscillatory part is applied to the noise itself. To numerically face this challenge, we introduce in this work a new methodology to develop methods of any high weak order to approximate the solution of equation (1.3.2) that are robust with respect to the stiff parameter ε. In particular, we propose a method of weak order two, and a geometric modification of this algorithm that preserves quadratic invariants.

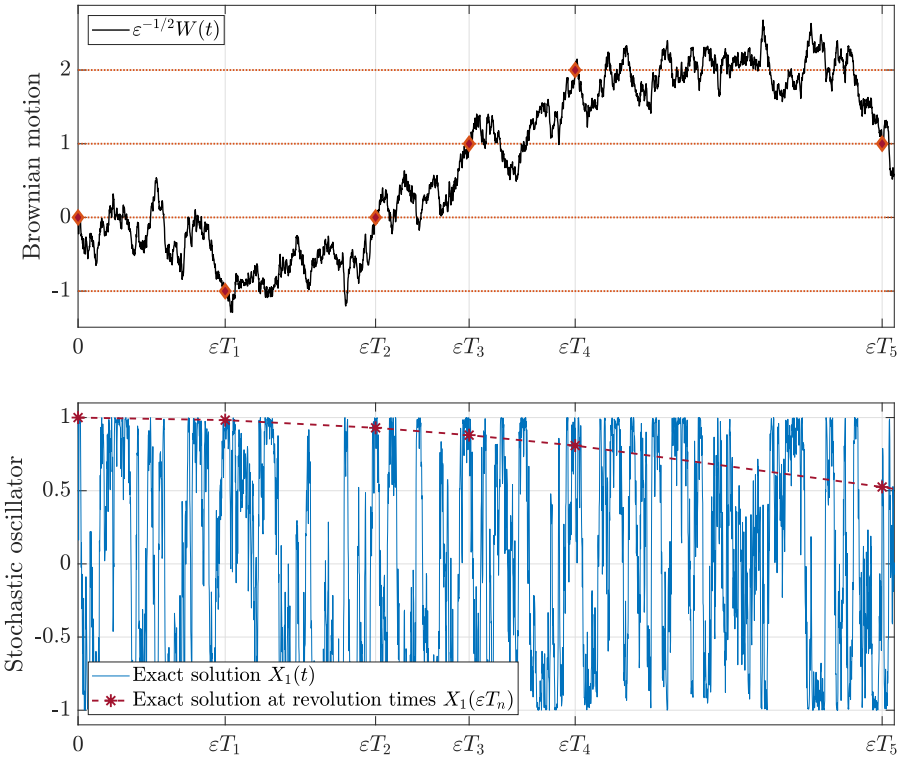


Figure 1.2: (From [84]) Revolution times (1.3.4) of a Brownian path (top) and exact solution evaluated at revolution times for the Kubo oscillator (1.3.3) with $a = 1$ and $\varepsilon = 10^{-1}$ (bottom).

1.3.1 New integrators for highly-oscillatory SDEs

Initially created in the context of celestial mechanics and later extended using geometric integration, multirevolution methods represent a class of numerical methods used for solving highly-oscillatory differential equations while reducing the cost of computation. For stochastic oscillations, the solution $X(t) = e^{\varepsilon^{-1/2}AW(t)}X_0$ of $dX(t) = \varepsilon^{-1/2}AX(t) \circ dW(t)$ is not periodic, but satisfies $X(\varepsilon T_N) = X_0$ where the T_N are random variables called revolution times and defined by $T_0 = 0$ and

$$T_{N+1} = \inf\left\{t > T_N, \varepsilon^{-1/2}\left|W(\varepsilon t) - W(\varepsilon T_N)\right| \geqslant 1\right\}, \quad N = 0, 1, 2, \dots \tag{1.3.4}$$

If X is the solution of (1.3.2), $X(\varepsilon T_N)$ is a perturbation of identity, that is, X satisfies the estimate $X(\varepsilon T_N) = X_0 + \mathcal{O}(\varepsilon N)$, thus the solution loses its highly-oscillatory feature when evaluated at revolution times, as shown in Figure 1.2. The idea of multirevolution is to approximate $X(\varepsilon T_N)$ with $N = \mathcal{O}(\varepsilon^{-1})$, with a cost independent of ε.

We propose the two following new multirevolution methods of weak second order for integrating equation (1.3.2) at the revolution times εT_{Nm} for $m = 0, 1, 2, \dots$ with cost in $H = N\varepsilon = \mathcal{O}(1)$ independent of ε. Method B is a geometric modification of Method A to preserve quadratic invariants. Methods A and B involve a Fourier decomposition of the following functions that are 1-periodic with respect to θ,

$$\begin{aligned} g^0_\theta(y) &= e^{-A\theta}F(e^{A\theta}y) = \sum_{k\in\mathbb{Z}} c^0_k(y)e^{2i\pi k\theta}, \\ g^1_\theta(y)(z) &= e^{-A\theta}F'(e^{A\theta}y)(e^{A\theta}z) = \sum_{p\in\mathbb{Z}} c^1_p(y)(z)e^{2i\pi p\theta}, \end{aligned} \tag{1.3.5}$$

with Fourier coefficients $(c^0_k(y))_{k\in\mathbb{Z}}$ and $(c^1_p(y))_{p\in\mathbb{Z}}$, respectively. In general, the series appearing in (1.3.5) have an infinite number of terms. For a practical implementation of the new methods, we truncate these series up to an even number of modes K_t, while inducing an exponentially small error. For each timestep, we also introduce bounded discrete random variables $(\widehat{\alpha}^N_k)_k$, and deterministic sequences $(\widehat{\beta}^N_{p,k})_{p,k}$ and $(\widehat{\widehat{\beta}}^N_{p,k})_{p,k}$, that are cheap to compute and whose definition is omitted here for brevity.

Method A (Explicit integrator of weak order two in $H = N\varepsilon$ to approximate the solution of equation (1.3.2) at times εT_{Nm} for $m = 0, 1, 2, \dots$)

$Y_0 = X_0$
for $m \geqslant 0$ **do**

$$Y_{m+1} = Y_m + H\sum_{k=-K_t/2}^{K_t/2-1} c^0_k(Y_m)\widehat{\alpha}^N_k + H^2\sum_{p,k=-K_t/2}^{K_t/2-1} c^1_p(Y_m)(c^0_k(Y_m))\widehat{\beta}^N_{p,k} \tag{1.3.6}$$

end for

Theorem 1.3.1 (See Sect. 5.4 and Theorem 5.4.1)**.** *Under technical assumptions, Method A has weak order two for solving* (1.3.2)*, that is, for all $T > 0$, for all test functions ϕ, there exists $H_0 > 0$ such that for all $H \leqslant H_0$, for all $m \geqslant 0$ such that $mN\varepsilon = mH \leqslant T$, there exist two positive constants K and C, both independent of ε and N such that*

$$\left|\mathbb{E}[\phi(X(\varepsilon T_{Nm}))] - \mathbb{E}[\phi(Y_m)]\right| \leqslant CH^2(1 + \mathbb{E}[|X_0|^K]). \tag{1.3.8}$$

Method B (Geometric integrator of weak order two in $H = N\varepsilon$ to approximate the solution of equation (1.3.2) at times εT_{Nm} for $m = 0, 1, 2, \ldots$ while preserving quadratic invariants)

$Y_0 = X_0$
for $m \geqslant 0$ **do**

$$Y_{m+1} = Y_m + H \sum_{k=-K_t/2}^{K_t/2-1} c_k^0 \left(\frac{Y_m + Y_{m+1}}{2} \right) \widehat{\alpha}_k^N + H^2 \sum_{p,k=-K_t/2}^{K_t/2-1} c_p^1 \left(\frac{Y_m + Y_{m+1}}{2} \right) \left(c_k^0 \left(\frac{Y_m + Y_{m+1}}{2} \right) \right) \widehat{\widehat{\beta}}_{p,k}^N \tag{1.3.7}$$

end for

Theorem 1.3.2 (See Sect. 5.4 and Theorem 5.4.2)**.** *Under technical assumptions, Method B has weak order two for solving* (1.3.2) *(i.e., it satisfies an estimate of the form* (1.3.8)*). In addition, it preserves quadratic invariants, that is, $Q(Y_{m+1}) = Q(Y_m)$ almost surely if Q is a quadratic invariant of* (1.3.2)*.*

1.3.2 Numerical experiments

We solve numerically the nonlinear Schrödinger equation (1.3.1) on the torus $\mathbb{T} = [-\pi, \pi]$ with the polynomial nonlinearity $F(u) = i\,|u|^{2\sigma} u$ and the stiffness parameter $\varepsilon = 10^{-2}$. We apply Methods A and B to a spectral discretization in space of this SPDE with $\sigma = 4$, and we compare their performance to the performance of the following one-step explicit deterministic integrator called the Euler scheme

$$y_0 = X_0, \quad y_{m+1} = y_m + H c_0^0(y_m), \tag{1.3.9}$$

where c_0^0 is defined in (1.3.5). Its cost is independent of ε and N, and it has weak order one with respect to H. On Figure 1.3 (left picture), we observe the evolution in time of one trajectory given by Method B. On Figure 1.3 (right picture), we plot the evolution of the discrete L^2 norm

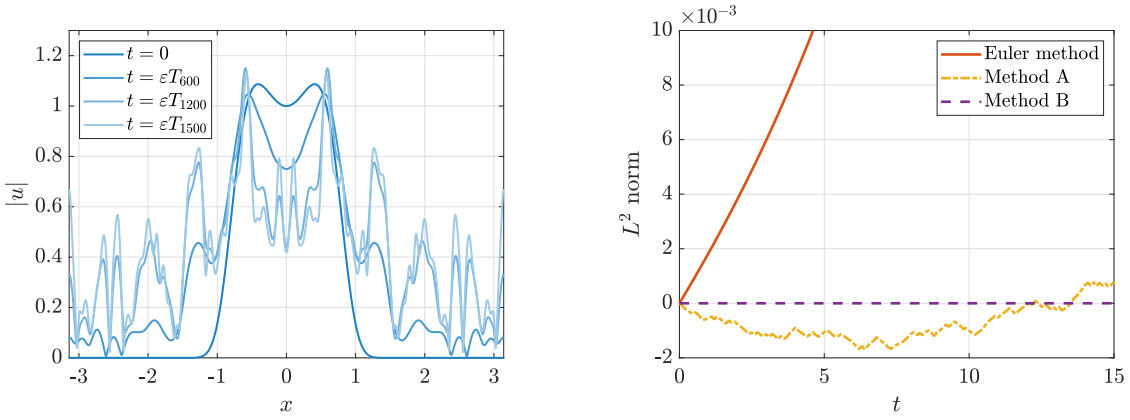


Figure 1.3: (From [84]) Approximation by Method B of $|u(t)|$ (left) and evolution in long time of the quantity $\|U_t\|_{L^2} - \|u_0\|_{L^2}$ (right) with u solution of a spatial discretization with $K_x = 2^7$ modes of the nonlinear Schrödinger equation with white noise dispersion (1.3.1) on the torus $\mathbb{T} = [-\pi, \pi]$ with the parameters $\varepsilon = 10^{-2}$ and $\sigma = 4$, and U_t the computed approximation.

of one trajectory given by our two algorithms and the Euler method (1.3.9). The Euler method quickly blows up, and the L^2 norm of Method A is not preserved. In contrast, Method B preserves the L^2 norm according to Theorem 1.3.2. Further numerical simulations also hint that a blow-up in the H^1 norm always happens for all considered methods. This behavior agrees with the blow-up conjecture for $\varepsilon = 1$ and $\sigma \geqslant 4$ presented in [11], and suggests that the conjecture persists in the highly-oscillatory regime $\varepsilon \ll 1$. We present in Section 5.5 further numerical experiments that confirm the results of Theorem 1.3.1 and Theorem 1.3.2.

CHAPTER 2

Preliminaries

We introduce in this chapter the notations, assumptions and standard tools we use in the rest of this work for the study of numerical integrators for solving SDEs. In particular, Section 2.1 is devoted to the definition of the SDEs we study here, with the physical context and the associated numerical challenges. In Section 2.2, we introduce the standard material for computing weak averages, including the Kolmogorov equation and the Talay-Tubaro expansion. In Section 2.3, we define and analyze the ergodicity property in $\mathbb{R}^d$ and on manifolds, with an emphasis on Langevin dynamics, and we present the conditions in terms of the Talay-Tubaro expansion for finding the accuracy for the invariant measure of a numerical integrator. Section 2.4 presents a concise introduction to the tools of geometric numerical integration that we use in this work.

2.1 Stochastic differential equations: examples and numerical challenges

We integrate numerically systems of Itô SDEs in $\mathbb{R}^d$ of the form

$$dX(t) = f(X(t))dt + \Sigma(X(t))dW(t), \quad X(0) = X_0, \tag{2.1.1}$$

where the solution $X(t)$ is in $\mathbb{R}^d$, $f\colon \mathbb{R}^d \to \mathbb{R}^d$ and $\Sigma\colon \mathbb{R}^d \to \mathbb{R}^{d\times k}$ are smooth globally Lipschitz maps, $W(t)$ is a standard k-dimensional Brownian motion in $\mathbb{R}^d$ on a probability space equipped with a filtration and fulfilling the usual assumptions. We also study SDEs in the sense of Stratonovich that are constrained on a compact smooth manifold $\mathcal{M}$,

$$dX(t) = \Pi_{\mathcal{M}}(X(t))f(X(t))dt + \Pi_{\mathcal{M}}(X(t))\Sigma(X(t)) \circ dW(t), \quad X(0) = X_0 \in \mathcal{M}, \tag{2.1.2}$$

where f is a Lipschitz map in a neighborhood of $\mathcal{M}$, the solution $X(t)$ lies on $\mathcal{M}$ for all $t > 0$ due to the projection operator $\Pi_{\mathcal{M}}\colon \mathbb{R}^d \to \mathbb{R}^{d\times d}$ on the tangent bundle of the manifold $\mathcal{M}$, and where $\circ$ denotes the Stratonovich product. In both contexts, the Lipschitz condition guaranties the existence and uniqueness of a solution to (2.1.1)-(2.1.2) for all times $t > 0$. The study of numerical integrators for solving SDEs with a non-Lipschitz f is interesting, but is out of the scope of the present work. We refer for instance to [108, 53] for further details on the standard theory of SDEs.

Stochastic differential equations have a wide range of applications in a variety of fields that include physics, mathematical finance or biology. In particular, we mention the overdamped Langevin equation in molecular dynamics (obtained in the particular case where $\Sigma(x) = \sigma I_d$ is a constant homothety),

$$dX(t) = f(X(t))dt + \sigma dW(t), \quad X(0) = X_0, \tag{2.1.3}$$

where $\sigma > 0$, $f = -\nabla V$ with $V\colon \mathbb{R}^d \to \mathbb{R}$ a smooth potential and $W(t)$ is a standard d-dimensional Wiener process. The overdamped Langevin equation is widely used to model the motion of a set of particles subject to a potential V in a high friction regime.

If the particles are submitted to constraints such as strong covalent bonds between atoms, or fixed angles in molecules, the solution of the SDE lies on a manifold of the form $\mathcal{M} = \{x \in \mathbb{R}^d, \zeta(x) = 0\}$, where $\zeta\colon \mathbb{R}^d \to \mathbb{R}^q$ is the smooth constraint. In this setting, the particles satisfy the constrained overdamped Langevin equation

$$dX(t) = \Pi_{\mathcal{M}}(X(t))f(X(t))dt + \sigma\Pi_{\mathcal{M}}(X(t)) \circ dW(t), \quad X(0) = X_0 \in \mathcal{M}. \tag{2.1.4}$$

In particular, equations (2.1.3) and (2.1.4) are identical in the case of $\mathbb{R}^d$ as the projection $\Pi_{\mathcal{M}}$ is the identity matrix. For $\zeta\colon \mathbb{R}^d \to \mathbb{R}$ a scalar constraint, we denote $g = \nabla\zeta$ its gradient, and $G(x) = g^T(x)g(x) = |g(x)|^2$ the Gram function related to the manifold $\mathcal{M} = \{x \in \mathbb{R}^d, \zeta(x) = 0\}$, where we denote by $|x| = (x^T x)^{1/2}$ the Euclidean norm in $\mathbb{R}^d$. For the sake of simplicity, we assume in the rest of the thesis that $\mathcal{M}$ is either $\mathbb{R}^d$ or a compact and smooth manifold of codimension one embedded in $\mathbb{R}^d$. In the manifold case, we suppose in addition that the Gram function G is strictly positive on $\mathcal{M}$, $G(x) \geqslant \alpha > 0$ for all $x \in \mathcal{M}$. With these notations, the projection $\Pi_{\mathcal{M}}$ on the tangent bundle is given by $\Pi_{\mathcal{M}}(x) = I - G(x)^{-1}g(x)g(x)^T$. Sampling from the constrained overdamped Langevin equation allows us to compute the so-called free energy, which is a key quantity in thermodynamics (see, for instance, [38, 90, 91] and references therein). Equations of the form (2.1.4) appear naturally when studying conservative SDEs, that is, SDEs possessing an invariant H conserved almost surely by all realizations of (2.1.4). The solution of conservative SDEs are subject to the constraint $\zeta(X) = 0$ with $\zeta(x) = H(x) - H(X_0)$. Drawing samples on a manifold also has many applications in statistics (see [18, 50] and references therein).

Finally, we mention the following SDE with a highly-oscillatory linear term driven by a one-dimensional Stratonovich white noise that arises in fiber optics models (see [7, 8, 58]),

$$dX(t) = \frac{1}{\sqrt{\varepsilon}}AX(t) \circ dW(t) + F(X(t))dt, \ t > 0, \ X(0) = X_0, \tag{2.1.5}$$

where the function $F : \mathbb{R}^d \to \mathbb{R}^d$ is a smooth possibly non-linear map, and $A \in \mathbb{R}^{d\times d}$ is a given matrix satisfying $e^A = \mathrm{Id}$ (equivalently A is diagonalizable and has all its eigenvalues in $2i\pi\mathbb{Z}$). In the deterministic setting, this last property yields that the solution $x(t) = \exp(\varepsilon At)x_0$ of $\frac{dx}{dt} = \varepsilon^{-1}Ax$ is ε-periodic. For stochastic oscillations, it means that the solution $X(t) = \exp(\varepsilon^{-1/2}AW(t))X_0$ of $dX = \varepsilon^{-1/2}AX \circ dW$ satisfies $X(T) = X(0)$ for a random time $T = \inf\{t > 0, \left|\varepsilon^{-1/2}W(t)\right| = 1\}$ of mean ε. The parameter $\varepsilon > 0$ is fixed and represents the high frequency of the oscillations. If ε is small, the solution has an highly-oscillatory behavior, and the problem (2.1.5) is stiff.

In this thesis, we introduce new efficient one-step integrators of the form

$$X_{n+1} = \Phi(X_n, h, \xi_n), \tag{2.1.6}$$

where the ξ_n are independent random variables and h is the numerical timestep. The simplest example of a consistent integrator of the form (2.1.6) for solving (2.1.1) is the Euler-Maruyama method,

$$X_{n+1} = X_n + hf(X_n) + \sqrt{h}\Sigma(X_n)\xi_n, \tag{2.1.7}$$

where the $\xi_n \sim \mathcal{N}(0, I_d)$ are independent standard Gaussian random vectors in $\mathbb{R}^d$.

Depending on the properties of the SDE we study and the quantities we wish to approximate, there corresponds a variety of numerical properties that we focus on for our numerical schemes to be more efficient. For problem (2.1.5), a standard method, such as the Euler-Maruyama method (2.1.7), would require a time stepsize $h \leqslant \varepsilon$ to be accurate, which makes it dramatically expensive when ε is small. For equation (2.1.4), as the solution lies on the manifold $\mathcal{M}$, the numerical scheme should also lie on $\mathcal{M}$, and this calls for specific geometric methods. Moreover, the Langevin dynamics (2.1.3)-(2.1.4) satisfy an ergodicity property, that is, the solution behaves in long time according to an invariant measure, and one is often more interested in this context in computing averages with respect to the invariant measure than weak averages. We recall that a natural way to achieve high order p for the invariant measure is to consider a numerical scheme with high standard weak order r, as it is known for large classes of SDEs that $p \geqslant r$. However it is possible to reach a high order p for the invariant measure, while keeping a low standard weak order of accuracy, typically $r = 1$, as presented for instance in [14, 86, 87, 6]. The analysis of weak order conditions and conditions for the invariant measure differ in many ways, and one would like to build integrators with high order for the invariant measure and low weak order.

2.2 Numerical integration of stochastic dynamics in finite time

In this section, we present the standard material for studying errors in the weak context, with an emphasis on Langevin dynamics. For more details, we refer the reader to [104].

For each problem we presented in Section 2.1, one wants to approximate different quantities that are related to the physical context. It can be the exact trajectory of the solution for a given realisation of W, the law of the process, or the stationary law of the process if it exists. Indeed, there are different ways to approximate the solution of the SDE problem (2.1.1). A strong approximation focuses on approaching the realisation of a single trajectory of (2.1.1) for a given realisation of W. A weak approximation approaches the average of functionals of the solution, that is, quantities of the form $\mathbb{E}[\phi(X(t))|X_0]$ for ϕ a given test function.

A numerical approximation (2.1.6) is said to have local weak order r for approximating the solution of (2.1.1) if for every test function ϕ, for all h small enough and for $X_0 = x$ assumed deterministic for simplicity, there exists a positive constant $C(\phi, x)$ such that

$$\left|\mathbb{E}[\phi(X_1)] - \mathbb{E}[\phi(X(h))]\right| \leqslant C(\phi, x)h^{r+1}.$$

The integrator has global weak order r if, for all $T > 0$, for all test function ϕ, for all h small enough and all $N \in \mathbb{N}$ such that $T = Nh$, there exists a positive constant $C(\phi, x)$ such that

$$\left|\mathbb{E}[\phi(X_N)|X_0 = x] - \mathbb{E}[\phi(X(T))|X_0 = x]\right| \leqslant C(\phi, x)h^r.$$

Under stability assumptions, a numerical scheme of local weak order r automatically has global weak order r. Sufficient stability properties are proposed in Assumption 2.2.3.

The choice of the test functions space depends on the context. It typically is a set of the form $\mathcal{C}^k(\mathbb{R}^d, \mathbb{R})$ or $\mathcal{C}^k_P(\mathbb{R}^d, \mathbb{R})$ for $k = 0, 1, \ldots, \infty$, where we recall that $\phi \in \mathcal{C}^k_P(\mathbb{R}^d, \mathbb{R})$ if and only if for every $j \leqslant k$, there exists constants C_j, $K_j > 0$ such that for all $x \in \mathbb{R}^d$,

$$\left|\phi^{(j)}(x)\right| \leqslant C_j(1 + |x|^{K_j}).$$

Unless specified otherwise, for dynamics in $\mathbb{R}^d$, we take $\mathcal{C}^\infty_P(\mathbb{R}^d, \mathbb{R})$, and for dynamics on a compact manifold $\mathcal{M}$, we use the space $\mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$.

A classical tool for the study of (2.1.1) is the generator. It is a differential operator of order two, that is given, for ϕ a test function, by

$$\mathcal{L}\phi = \phi' f + \frac{1}{2}\sum_{i=1}^{k} \phi''(\Sigma e_i, \Sigma e_i), \tag{2.2.1}$$

where (e_i) is the canonical basis of $\mathbb{R}^k$, and where we use the following notation for differentials in $\mathbb{R}^d$,

$$\phi^{(m)}(a^1, \ldots, a^m) = \sum_{i_1,\ldots,i_m=1}^{d} \partial_{i_1,\ldots,i_m}\phi\, a^1_{i_1} \ldots a^m_{i_m} = \sum_{i_1,\ldots,i_m=1}^{d} \frac{\partial^m \phi}{\partial x_{i_1} \ldots \partial x_{i_m}}\, a^1_{i_1} \ldots a^m_{i_m}.$$

In particular, the generator of equation (2.1.3) is given in $\mathbb{R}^d$, for ϕ a test function, by

$$\mathcal{L}\phi = \phi' f + \frac{\sigma^2}{2}\Delta\phi, \tag{2.2.2}$$

and for the Langevin dynamics (2.1.4) constrained on the manifold $\mathcal{M}$, it is given by

$$\begin{aligned}\mathcal{L}\phi = \phi' f - G^{-1}(g, f)\phi' g - \frac{\sigma^2}{2} G^{-1}\operatorname{div}(g)\phi' g + \frac{\sigma^2}{2} G^{-2}(g, g' g)\phi' g + \frac{\sigma^2}{2}\Delta\phi \\ - \frac{\sigma^2}{2} G^{-1}\phi''(g, g) = \frac{\sigma^2}{2}\exp\Big(\frac{2}{\sigma^2}V\Big)\operatorname{div}_{\mathcal{M}}\Big(\exp\Big(-\frac{2}{\sigma^2}V\Big)\nabla_{\mathcal{M}}\phi\Big),\end{aligned} \tag{2.2.3}$$

where $\nabla_{\mathcal{M}}\psi := \Pi_{\mathcal{M}}\nabla\psi$, $\operatorname{div}_{\mathcal{M}}(H) := \operatorname{div}(H) - G^{-1}(g, H'(g))$.

From now on and until the end of Section 2.3, we assume that W is a d-dimensional Brownian motion, and we make an additional regularity assumption on the generator $\mathcal{L}$. On a compact manifold $\mathcal{M}$, we follow the framework of [55]. In particular, we rely on the construction of the local orthogonal coordinates. In a neighbourhood $N_{\mathcal{M}}$ of the manifold $\mathcal{M}$, there exists an atlas of local orthogonal coordinate systems $(y, z) \in (V \subset \mathbb{R}^{d-1}) \times (-\varepsilon, \varepsilon)$ for $\varepsilon > 0$, with respect to local charts $\psi \colon U \subset N_{\mathcal{M}} \to (V \subset \mathbb{R}^{d-1}) \times (-\varepsilon, \varepsilon)$, such that if $\psi(x) = (y, z)$, then $z = \zeta(x)$. We take over the following sufficient assumptions, in the spirit of [68, 5] in $\mathbb{R}^d$ and [55] on manifolds, to carry out our analysis, and we emphasize that finding the most general assumptions is not the topic of this work.

Assumption 2.2.1. *In the context of $\mathbb{R}^d$, the functions f and Σ have bounded derivatives of any order, Σ is bounded and there exist $C_1, C_2 > 0$ such that for all $x \in \mathbb{R}^d$,*

$$x^T f(x) \leqslant -C_1 x^T x + C_2.$$

Moreover, the generator $\mathcal{L}$ is uniformly elliptic, that is, there exists a constant $C > 0$ such that, for all x, $v \in \mathbb{R}^d$,

$$v^T \Sigma(x)^T \Sigma(x) v \geqslant C v^T v.$$

In the context of manifolds, on an open neighbourhood $N_{\mathcal{M}}$ of $\mathcal{M}$ in $\mathbb{R}^d$, there exists a constant $C > 0$ such that for all $x \in N_{\mathcal{M}}$ and $(y, z) = \psi(x)$, for all one-form field $v \colon T\mathcal{M} \to \mathbb{R}$ on $\mathcal{M}$ of norm one, we have

$$\sum_{i,j=1}^{d-1} \sum_{k=1}^{d} \widetilde{\Sigma}_{ik}(y,z)\widetilde{\Sigma}_{jk}(y,z)v_i(\widetilde{x})v_j(\widetilde{x}) \geqslant C,$$

where $\widetilde{x} \in \mathcal{M}$ is such that $\psi(\widetilde{x}) = (y, 0)$ and, for $k = 1, \ldots, d$, $(\widetilde{\Sigma}_{ik}(y,z))_i \in \mathbb{R}^{d-1}$ is defined as the restriction of the vector $(\Pi_{\mathcal{M}}(\widetilde{x})\Sigma_{ik}(x))_i \in \mathbb{R}^d$ to the tangent space $T_{\widetilde{x}}\mathcal{M}$ of $\mathcal{M}$, written in the local orthogonal coordinate system.

In the context of Langevin dynamics (2.1.3)-(2.1.4), it is sufficient to assume the following.

Assumption 2.2.2. *The vector field f is a globally Lipschitz smooth gradient, i.e. there exists a C^∞ potential $V : \mathbb{R}^d \to \mathbb{R}$ such that $f(x) = -\nabla V(x)$ is globally Lipschitz. Moreover, in the context of $\mathbb{R}^d$, there exist two constants $C_1 > 0$ and C_2 such that for all $x \in \mathbb{R}^d$, the estimate $V(x) \geqslant C_1 x^T x - C_2$ holds.*

The regularity assumption 2.2.1 yields that the function $u(x,t) = \mathbb{E}[\phi(X(t))|X(0) = x]$ satisfies the backward Kolmogorov equation (see [79, 80] in $\mathbb{R}^d$ and [55] on manifolds):

$$\frac{\partial u}{\partial t}(x,t) = \mathcal{L}u(x,t), \quad u(x,0) = \phi(x), \quad x \in N_{\mathcal{M}}, \quad t > 0. \tag{2.2.4}$$

The backward Kolmogorov equation (2.2.4) allows us to write the following expansion of the weak average $u(x,h) = \mathbb{E}[\phi(X(h))|X_0 = x]$ for h small enough,

$$u(x,h) = \phi(x) + \sum_{j=1}^{N} \frac{h^j}{j!}\mathcal{L}^j\phi(x) + h^{N+1}R_N^h(\phi,x), \quad x \in N_{\mathcal{M}}, \tag{2.2.5}$$

where $N_{\mathcal{M}}$ is an open neighbourhood of $\mathcal{M}$ in $\mathbb{R}^d$ (that is, $N_{\mathcal{M}} = \mathbb{R}^d$ if $\mathcal{M} = \mathbb{R}^d$) and the remainder satisfies the estimate $\left|R_N^h(\phi,x)\right| \leqslant C_N(\phi,x)$ where the constant $C_N(\phi,x)$ is independent of h.

For the numerical integrator (2.1.6), we make the following stability assumption.

Assumption 2.2.3. *The integrator* (2.1.6) *has bounded moments of any order along time, i.e. for all integer $k \geqslant 0$ and all final time $T > 0$, for all h and N such that $Nh = T$,*

$$\sup_{n \leqslant N} \mathbb{E}\big[|X_n|^{2k}\big] < \infty.$$

Remark 2.2.4. *A convenient sufficient condition to satisfy Assumption 2.2.3 is given in [104, Lemma 2.2.2]: if X_0 is deterministic or has bounded moments of all order and the Markov chain $(X_n)_n$ satisfies*

$$|\mathbb{E}[X_{n+1} - X_n|X_n]| \leqslant C(1 + |X_n|)h, \qquad |X_{n+1} - X_n| \leqslant M_n(1 + |X_n|)\sqrt{h},$$

for C a constant independent of h and M_n a random variable whose moments are all bounded uniformly with respect to h small enough, then the numerical scheme satisfies Assumption 2.2.3. Hence Runge-Kutta type schemes introduced in Chapter 4 satisfy Assumption 2.2.3 naturally.

We assume in addition that $\mathbb{E}[\phi(X_1)|X_0 = x]$, the numerical analog of $u(x,h)$, can be expanded in power series with respect to h, as we did for the exact solution in (2.2.5). This weak expansion of the integrator was first introduced in [121], and is therefore called the Talay-Tubaro expansion.

Assumption 2.2.5. *For all test functions ϕ, the numerical integrator* (2.1.6) *has a weak Taylor expansion of the form*

$$\mathbb{E}[\phi(X_1)|X_0 = x] = \phi(x) + \sum_{j=1}^{N} h^j \mathcal{A}_{j-1}\phi(x) + h^{N+1} R_N^h(\phi, x), \quad x \in N_{\mathcal{M}}, \tag{2.2.6}$$

for all h assumed small enough, and where $N_{\mathcal{M}}$ is an open neighbourhood of $\mathcal{M}$ in $\mathbb{R}^d$ and the remainder satisfies $\left|R_N^h(\phi, x)\right| \leqslant C_N(\phi, x)$ where the constant $C_N(\phi, x)$ is independent of h. The $\mathcal{A}_j$'s, $j = 0, 1, 2, \ldots$ are linear differential operators with coefficients depending smoothly on f, g and their (high order) derivatives (and depending on the choice of the integrator).

With all the previously introduced material, we get the following result on global weak order.

Theorem 2.2.6 ([121])**.** *Under Assumptions 2.2.1, 2.2.3 and 2.2.5, if*

$$\mathcal{A}_{j-1} = \mathcal{L}^j/j!, \qquad j = 1, \ldots, r,$$

then the integrator (2.1.6) *has at least weak order r for solving* (2.1.1)*.*

Example. *The Euler-Maruyama method* (2.1.7) *has weak order 1 for solving* (2.1.1)*. The following θ-method is of weak order two for solving* (2.1.3) *when $\theta = 1/2$,*

$$X_{n+1} = X_n + h(1-\theta)f(X_n) + h\theta f(X_{n+1}) + \sigma\sqrt{h}\xi_n.$$

Indeed, this method naturally satisfies Assumption 2.2.3, and the Talay-Tubaro operators are given by $\mathcal{A}_0 = \mathcal{L}$ and

$$\mathcal{A}_1\phi = \frac{1}{2}\phi' f' f + \frac{1}{2}\phi''(f,f) + \frac{1}{4}\phi'\Delta f + \frac{1}{2}\sum_{i,j=1}^{d} \partial_{ij}\phi\partial_i f_j + \frac{1}{2}(\Delta\phi)' f + \frac{1}{8}\Delta^2\phi = \frac{1}{2}\mathcal{L}^2\phi.$$

2.3 Ergodicity and high order integration for the invariant measure

We saw two different kinds of approximations of the solution of a SDE: strong and weak. In this section, we present the ergodicity property of stochastic processes, and this leads us to the approximation for the invariant measure. We introduce a theorem for building high order integrators for the invariant measure in $\mathbb{R}^d$ and on manifolds, and a method for increasing the order of a scheme with a postprocessor.

We first assume the existence and uniqueness of an invariant measure, as well as an additional regularity property on $\mathcal{L}$, in the spirit of [48, Hypotheses H1-H2] in the context of $\mathbb{R}^d$.

Assumption 2.3.1. *There exists an open neighbourhood $N_{\mathcal{M}}$ of $\mathcal{M}$ in $\mathbb{R}^d$ and a unique positive function $\rho_\infty \in \mathcal{C}^\infty(N_{\mathcal{M}}, \mathbb{R})$ satisfying $\int_{\mathcal{M}} \rho_\infty d\sigma_{\mathcal{M}} = 1$ and $\mathcal{L}^*\rho_\infty = 0$ on $N_{\mathcal{M}}$. Moreover, for all $\phi \in \mathcal{C}^\infty(N_{\mathcal{M}}, \mathbb{R})$ such that $\int_{\mathcal{M}} \phi d\sigma_{\mathcal{M}} = 0$, there exists a unique solution $\rho \in \mathcal{C}^\infty(N_{\mathcal{M}}, \mathbb{R})$ to the Poisson problem $\mathcal{L}^*\rho = \phi$ that satisfies $\int_{\mathcal{M}} \rho d\sigma_{\mathcal{M}} = 0$.*

The existence and uniqueness of the invariant measure are in particular satisfied for the overdamped Langevin dynamics (2.1.3)-(2.1.4) (see [55, Sect. 2.3] for further details). Assumption 2.3.1 yields the ergodicity of the process $X(t)$ solution of (2.1.1) or (2.1.2), that is, there exists a unique invariant measure $d\mu_\infty$ on $\mathcal{M}$ that has a density ρ_∞ with respect to $d\sigma_\mathcal{M}$, the canonical measure on $\mathcal{M}$ induced by the Euclidean metric of $\mathbb{R}^d$ (that is, the Lebesgue measure in the context of $\mathbb{R}^d$), such that for all test functions ϕ,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \phi(X(t))dt = \int_\mathcal{M} \phi(x)d\mu_\infty(x) \quad \text{almost surely.} \tag{2.3.1}$$

In the context of Langevin dynamics (2.1.3)-(2.1.4), the density of the invariant measure $d\mu_\infty = \rho_\infty d\sigma_\mathcal{M}$ is given by $\rho_\infty = \frac{1}{Z}\exp\left(-\frac{2}{\sigma^2}V\right)$ with Z such that $d\mu_\infty$ is a probability measure. We emphasize that in the context of manifolds, $d\mu_\infty$ is singular with respect to the Lebesgue measure on $\mathbb{R}^d$, thus the theory on $\mathbb{R}^d$ does not directly apply.

To proceed further, we shall assume that the integrator (2.1.6) is ergodic, that is, there exists a measure $d\mu_h$ that has a density with respect to $d\sigma_\mathcal{M}$ such that

$$\lim_{N\to\infty} \frac{1}{N+1}\sum_{n=0}^{N} \phi(X_n) = \int_\mathcal{M} \phi d\mu^h \quad \text{almost surely.} \tag{2.3.2}$$

We refer to [118, 119, 98, 120] in the Euclidean case, and to [55] in the manifold case, and references therein, for further details on the ergodicity of numerical integrators.

An ergodic integrator (2.1.6) is said to have order p for the invariant measure if for all test functions ϕ, there exists a positive constant $C(\phi, x)$ such that

$$e(\phi,h) \leqslant C(\phi)h^p \quad \text{where} \quad e(\phi,h) = \left|\lim_{N\to\infty} \frac{1}{N+1}\sum_{n=0}^{N} \phi(X_n) - \int_\mathcal{M} \phi d\mu_\infty\right|. \tag{2.3.3}$$

A natural way to achieve high order p for the invariant measure is to consider a numerical scheme with high standard weak order r, and it is known for large classes of SDEs that $p \geqslant r$, see in particular [98] in the context of locally Lipschitz vector fields with multiplicative noise. Note analogously that the strong order of convergence is in general lower than or equal to the weak order r of convergence. There are interestingly many schemes in the literature for which $p > r$ and a high order p for the invariant measure is obtained, while the standard weak order of accuracy remains low, typically of order $r = 1$, i.e. the scheme is consistent in the weak convergence sense. This is the case in particular for the Langevin equation [14, 86, 87, 6]. In [5, 6], a methodology for the analysis and design of high order integrators for the invariant measure is introduced in the context of $\mathbb{R}^d$. In particular, the papers introduce a criterion for high order for the invariant measure that we generalize in the context of manifolds in Chapter 4 in the following result.

Theorem 2.3.2. *Under Assumption 2.2.5, if the numerical scheme is consistent (that is, if $\mathcal{A}_0 = \mathcal{L}$) and ergodic, and if it satisfies in $L^2(d\sigma_\mathcal{M})$*

$$\mathcal{A}_j^* \rho_\infty = 0, \quad j = 1, \dots, p-1,$$

then it has order p for sampling the invariant measure of (2.1.1) *and the numerical error* (2.3.3) *satisfies, for $h \to 0$,*

$$\begin{aligned} e(\phi,h) &= h^r \int_{\mathcal{M}} \phi(x)\rho_r(x)d\sigma_{\mathcal{M}}(x) + \mathcal{O}(h^{r+1}) \\ &= h^r \int_0^\infty \int_{\mathcal{M}} u(x,t)\mathcal{A}_r^*\rho_\infty(x)d\sigma_{\mathcal{M}}(x)dt + \mathcal{O}(h^{r+1}), \end{aligned}$$

where $\rho_r \in \mathcal{C}^\infty(N_{\mathcal{M}},\mathbb{R})$ *is the unique solution of the Poisson problem* $\mathcal{L}^*\rho_r = -\mathcal{A}_r^*\rho_\infty$ *in* $N_{\mathcal{M}}$ *that satisfies* $\int_{\mathcal{M}} \rho_r d\sigma_{\mathcal{M}} = 0$*, with* $N_{\mathcal{M}}$ *an open neighbourhood of* $\mathcal{M}$ *in* $\mathbb{R}^d$*.*

Theorem 2.3.2 states the result for times $t \to \infty$. In practice, we use the Monte-Carlo estimator $\bar{J} = \frac{1}{M}\sum_{m=1}^M \phi(X_n^{(m)}) \simeq \mathbb{E}[\phi(X_N)]$ with M the number of trajectories and $X_n^{(m)}$ the m-th realisation of the integrator at time $t_n = nh$. Then, a bound of the error at finite time $t_n = nh$ is typically given by the following estimate (see [55, 99, 48])

$$\mathbb{E}\left[\left|\frac{1}{M}\sum_{m=1}^M \phi(X_n^{(m)}) - \int_{\mathcal{M}} \phi(x)d\mu_\infty(x)\right|\right] \leqslant C_1 h^p + \frac{C_2}{\sqrt{M}} + C_3 e^{-\mu t_n},$$

where the constant $\mu > 0$ is the spectral gap of a differential operator that depends on the numerical integrator. In Chapter 3 and Chapter 4, we are interested in increasing the order of the discretization p in the error term C_1h^p. The error term $\frac{C_2}{\sqrt{M}}$ is called the Monte-Carlo error and appears with the variance of the estimator. A handful of techniques exists for improving this error term, such as multilevel Monte-Carlo methods (see [59, 60] and references therein), the use of perturbations [51, 4, 117], or couplings [107]. The recent works [89, 51, 4] proposed numerical methods in $\mathbb{R}^d$ that improve the rate of convergence at infinity $C_3e^{-\mu t_n}$. The B-series formalism that we present in Chapter 3 is compatible with these techniques. We present in Section 4.3.5 a methodology that combines high-order schemes and the use of a non-reversible perturbation [89, 51, 117]. The reversible case [4] could also be tackled with the exotic aromatic B-series formalism.

In [123] for finite dimensions and in [15] in the context of parabolic stochastic partial differential equations, this approach is combined with the idea of processing from Butcher [23], to design efficient postprocessed integrators with high order for the invariant measure at a negligible overcost compared to standard low order schemes. The postprocessor methodology is extended in [1] for a class of explicit stabilized schemes of order two for the invariant measure and with optimally large stability domains. The following extension of Theorem 2.3.2 permits to combine an integrator (2.1.6) with a postprocessor to achieve high order for the invariant measure at a negligible overcost compared to a standard scheme. Note that extending this theorem in the context of manifolds is matter for future work.

Theorem 2.3.3 ([123])**.** *Assume the hypotheses of Theorem 2.3.2 in* $\mathbb{R}^d$ *and consider a postprocessor*

$$\overline{X}_n = G_n(X_n)$$

that admits the following weak Taylor expansion for all $\phi \in C_P^\infty(\mathbb{R}^d,\mathbb{R})$*,*

$$\mathbb{E}[\phi(G_n(x))] = \phi(x) + \sum_{i=1}^{p-1} \alpha_i h^i \mathcal{L}^i\phi(x) + h^p\overline{\mathcal{A}_p}\phi(x) + \dots, \tag{2.3.4}$$

for some constants α_i *and a linear differential operator* $\overline{\mathcal{A}_p}$*. Assume further that*

$$(\mathcal{A}_p + [\mathcal{L},\overline{\mathcal{A}}_p])^*\rho_\infty = 0 \tag{2.3.5}$$

where $[\mathcal{L},\overline{\mathcal{A}}_p] = \mathcal{L}\overline{\mathcal{A}}_p - \overline{\mathcal{A}}_p\mathcal{L}$ *is the Lie bracket. Then* $\overline{X}_n$ *yields an approximation of order* $p+1$ *for sampling the invariant measure of* (2.1.1)*.*

We study the hypotheses of Theorem 2.3.2 and Theorem 2.3.3 with a new algebraic formalism in Chapter 3 and 4.

2.4 Insights on deterministic geometric numerical integration

In this section, we recall several standard tools from geometric numerical integration, and we redefine the aromatic Butcher-series using a graph approach (rather than a tree approach), to prepare the introduction of lianas, a new type of edge on aromatic trees.

2.4.1 Invariant-preserving Runge-Kutta methods

We consider ordinary differential equations (ODEs) of the form

$$\frac{dy}{dt}(t) = f(y(t)), \quad y(0) = y_0, \tag{2.4.1}$$

where $f\colon \mathbb{R}^d \to \mathbb{R}^d$ is a smooth Lipschitz vector field. For solving the ODE (2.4.1), we consider Runge-Kutta methods of the following form, equivalently written with the associated Butcher tableau,

$$\begin{aligned} Y_i &= X_n + h \sum_{j=1}^{s} a_{ij} f(Y_j), \quad i = 1, \dots, s, \\ X_{n+1} &= X_n + h \sum_{i=1}^{s} b_i f(Y_i), \end{aligned} \qquad \begin{array}{c|c} c & A \\ \hline & b \end{array} \tag{2.4.2}$$

where $A = (a_{ij}) \in \mathbb{R}^{s\times s}$, $b = (b_i) \in \mathbb{R}^s$ and $c = A\mathbb{1}$ with $\mathbb{1} = (1, \dots, 1)^T$ are the given Runge-Kutta coefficients. For instance, the implicit midpoint method can be written as a Runge Kutta method with the following Butcher tableau.

$$X_{n+1} = X_n + hf\left(\frac{X_n + X_{n+1}}{2}\right), \qquad \begin{array}{c|c} \frac{1}{2} & \frac{1}{2} \\ \hline & 1 \end{array} \tag{2.4.3}$$

In various contexts, the solution of the ODE (2.4.1) leaves a quantity unchanged along time. This conserved quantity, often called first integral or invariant, can be the energy in the context of Hamiltonian dynamics, the angular momentum for the N-body problem or a constraint that implicitly defines a manifold that holds the trajectory. This additional information on the geometry of the ODE problem (2.4.1) calls for integrators that preserve these invariants, or that lie on the manifold. In this section, we introduce in the deterministic context the necessary standard tools that we use in this thesis for building integrators that preserve invariants, with an emphasis on quadratic invariants and projection methods. We refer to [65, Chap. IV] for more details on invariant-preserving integrators.

Definition 2.4.1. *A smooth non-constant function $Q(y)$ is a first integral of the ODE* (2.4.1) *if for all $y \in \mathbb{R}^d$,*

$$Q'(y)f(y) = 0.$$

This definition implies that if Q is a first integral of (2.4.1), then for any initial condition y_0, the solution $y(t)$ satisfies $Q(y(t)) = Q(y_0)$, hence the name invariant. If the invariant is a quadratic function, that is, if it can be written as $Q(y) = y^T S y$ with $S \in \mathbb{R}^{d\times d}$ a given symmetric matrix, and if $y^T S f(y) = 0$ for all $y \in \mathbb{R}^d$, we say that Q is a quadratic invariant.

To build integrators that preserve the invariant Q, that is, that satisfy $Q(X_{n+1}) = Q(X_n)$, we derive conditions on the coefficients of Runge-Kutta methods of the form (2.4.2) that ensure the preservation of the invariant. If Q is linear, then all Runge-Kutta methods preserve Q. The following condition specifies which Runge-Kutta methods preserve quadratic invariants.

Theorem 2.4.2 ([43])**.** *If the coefficients of the Runge-Kutta method satisfy*

$$b_i a_{ij} + b_j a_{ji} = b_i b_j, \qquad i, j = 1, \dots, s,$$

then it conserves all quadratic invariants.

For instance, the implicit midpoint method (2.4.3) preserves quadratic invariants. When applied to Hamiltonian systems, integrators that preserve quadratic invariants play a key role in geometric numerical integration, as they satisfy the property of symplecticity (see [65, Chap. VI]). Symplectic methods preserve a modified Hamiltonian close to the original Hamiltonian[1], and are widely used for the long-term integration of Hamiltonian systems in a variety of fields such as celestial mechanics or molecular dynamics.

For matrix systems of ODEs, one is also interested in the polynomial invariant $Q = \det$, that is related to the preservation of volume. The aromatic Butcher series were originally introduced in [36, 72] to compute conveniently the divergence of B-series, in order to build volume-preserving integrators (see [65, Sec. VI.9]).

When the invariant is a more general map, there is no particular method to find a Runge-Kutta method that preserves Q. In particular, we recall that there exists no Runge-Kutta method that preserves every polynomial invariant (see [65, Chap. VI.3]). An alternative solution is to project each step of the method on the manifold $\mathcal{M} = \{y \in \mathbb{R}^d, Q(y) = 0\}$. Let $\widetilde{X}_{n+1} = \Phi(X_n, h)$ be a consistent one-step approximation for solving (2.4.1). Let $Q_1, \dots, Q_q$ be the q invariants of the problem and $\lambda_1, \dots, \lambda_q$ be the Lagrange multipliers, then a consistent approximation that preserves the invariants Q_i is given by

$$\begin{aligned} X_{n+1} &= \widetilde{X}_{n+1} + \sum_{i=1}^{q} Q_i'(X_{n+1})\lambda_i, \\ Q_i(X_{n+1}) &= Q_i(X_n), \qquad i = 1, \dots, q. \end{aligned} \tag{2.4.4}$$

Integrators of this form are called projection methods (see [65, Sect. IV.4]). Projection methods exactly preserve the invariants Q_i of the system (2.4.1). Note that these methods are not strictly better than the standard Runge-Kutta methods, as explained in [64]. To solve the implicit problem (2.4.4), one can apply a few iterations of a Newton method under regularity assumptions on the functions Q_i. The term $Q_i'(X_{n+1})$ in (2.4.4) could be replaced by $Q_i'(X_n)$ or $Q_i'(x)$ with any one-step approximation x of X_n.

2.4.2 Aromatic Butcher-series for the computation of order conditions

Originally introduced in [24, 66], B-series have proved to be a powerful standard tool for the numerical analysis of deterministic differential equations, as presented, for instance, in the textbooks [65, 25, 26]. In this section, we give the definition of the aromatic Butcher-series, an extension of the standard B-series introduced in [36, 72] in the context of volume-preserving integrators. In the spirit of [13], we see the Butcher trees as graphs instead of considering the

[1] Note that this non-trivial result uses B-series for the construction of the modified equation (see [29, 63] and [65, Chap. IX]).

standard recursive definition of [65]. The two definitions are equivalent, but the graph approach is closer to the definition we shall use in Chapter 3 for the definition of the exotic aromatic B-series. We then apply the formalism for computing the order conditions of Runge-Kutta methods of the form (2.4.2).

B-series were originally introduced in [24, 66] for computing order conditions for Runge-Kutta methods. Consider a method of the form (2.4.2), then a Taylor expansion gives

$$X_1 = X_0 + h\sum_{i=1}^{s} b_i f(X_0) + h^2 \sum_{i=1}^{s} b_i c_i f'f(X_0) \\ + h^3\left(\frac{1}{2}\sum_{i=1}^{s} b_i c_i^2 f''(f,f) + \sum_{i,j=1}^{s} b_i a_{ij} c_j f'f'f\right)(X_0) + \mathcal{O}(h^4).$$

On the other hand, an expansion of the exact solution of (2.4.1) is given by

$$X(h) = X_0 + hf(X_0) + \frac{h^2}{2}f'f(X_0) + \frac{h^3}{6}(f''(f,f) + f'f'f)(X_0) + \mathcal{O}(h^4).$$

Thus, if the coefficients of the Runge-Kutta method satisfy

$$\sum_{i=1}^{s} b_i = 1, \quad \sum_{i=1}^{s} b_i c_i = \frac{1}{2}, \quad \sum_{i=1}^{s} b_i c_i^2 = \frac{1}{3}, \quad \sum_{i,j=1}^{s} b_i a_{ij} c_j = \frac{1}{6},$$

then, under stability assumptions, the method has at least order three for solving (2.4.1). For methods of higher orders, these Taylor expansions become more complex and require a strong algebraic formalism to be dealt with. The idea of B-series is to represent the terms that appear in the Taylor expansions with rooted trees. For a tree γ, we denote by $F(\gamma)$ the function that maps a tree to its associated differential, $|\gamma|$ the order of the tree, and a (resp. e) the coefficient of the Taylor expansion of the numerical (resp. exact) solution associated to a tree. For instance, for a Runge-Kutta method of the form (2.4.2), the first terms are collected in Table 2.1.

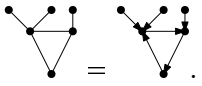
Tree γ	Order $\|\gamma\|$	Differential $F(\gamma)(f)$	Numerical coef. $a(\gamma)$	Exact coef. $e(\gamma)$
	1	f	$\sum_i b_i$	1
	2	$f'f$	$\sum_i b_i c_i$	$\frac{1}{2}$
	3	$f''(f,f)$	$\frac{1}{2}\sum_i b_i c_i^2$	$\frac{1}{6}$
	3	$f'f'f$	$\sum_{ij} b_i a_{ij} c_j$	$\frac{1}{6}$

Table 2.1: First terms of the expansion in B-series of a Runge-Kutta method of the form (2.4.2) and of the exact solution of (2.4.1).

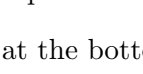
With these notations, the Taylor expansions of the exact and numerical solutions can be written as series indexed over trees, that we call B-series, and it is proved in [65, Chap. III] that any Runge-Kutta method of the form (2.4.2) can be written as a B-series. The maps e and a can be computed conveniently with straightforward calculations on graphs (see [65, Chap. III]) in order to avoid tedious Taylor expansions. The following result allows us to compute the order of a Runge-Kutta method conveniently with the help of B-series.

Theorem 2.4.3 ([65]). *If $a(\gamma) = e(\gamma)$ for $|\gamma| \leqslant p$, then the Runge-Kutta method has (at least) order p for solving* (2.4.1).

Now let us give a proper definition of an extension, called aromatic, of the standard Butcher trees and B-series. The aromatic extension was introduced in [36] to study volume-preserving integrators. In particular, this extension allows us to represent the divergence of standard B-series, which is a key tool in the study of volume-preserving methods (see [65, Sec. VI.9]). We follow the graph definition of aromatic B-series of [13]. We consider directed graphs $\gamma = (V, E)$ with V a finite set of nodes and $E \subset V \times V$ the set of directed edges. If $(v, w) \in E$, we say that the edge is going from v to w, and v is called a predecessor of w. Two directed graphs (V_1, E_1) and (V_2, E_2) are equivalent if there exists a bijection $\varphi : V_1 \to V_2$ with $(\varphi \times \varphi)(E_1) = E_2$. For brevity of notation, to avoid drawing arrows on the forests, an edge linking two nodes goes from the top node to the bottom one. If there is a possible cycle, by convention, the arrows on it are going in the clockwise direction. For example,



We call aromatic forests the equivalence classes of directed graphs where each node has at most one outgoing edge. The connected components making an aromatic forest are called aromatic trees. According to the above definition, there are two types of trees:

- *aromas* are aromatic trees[2] with exactly one cycle:  , ...
- *rooted trees* do not have a cycle ; they have a unique node that has no outgoing edge and that is called the root, graphically represented at the bottom:  , ...

Thus, an aromatic forest is a collection of aromas and rooted trees. We call $\mathcal{AT} = \{$  , ...$\}$ the set of aromatic forests containing exactly one rooted tree, and we name its elements the aromatic rooted forests. The order of a directed graph $\gamma = (V, E)$ is the number of nodes of γ, and is denoted as $|\gamma|$.

Definition 2.4.4 (Elementary differentials, [13]). *Let $\gamma = (V, E) \in \mathcal{AT}$, and let $f : \mathbb{R}^d \to \mathbb{R}^d$ be a smooth function. We denote $\pi(v) = \{w \in V, (w, v) \in E\}$ the set of all predecessors of the node $v \in V$ and r the root of γ. We also call $V^0 = V \setminus \{r\} = \{v_1, \dots, v_m\}$ the other nodes of γ. Finally we introduce the notation $I_{\pi(v)} = (i_{q_1}, \dots, i_{q_s})$ where the q_k are the predecessors of v, and we use the notation*

$$\partial_{I_{\pi(v)}} f = \frac{\partial^s f}{\partial x_{i_{q_1}} \dots \partial x_{i_{q_s}}}.$$

Then $F(\gamma)$ is defined as

$$F(\gamma)(f) = \sum_{i_{v_1}, \dots, i_{v_m} = 1}^{d} \left(\prod_{v \in V^0} \partial_{I_{\pi(v)}} f_{i_v} \right) \partial_{I_{\pi(r)}} f.$$

[2] Such graphs with one cycle are not strictly speaking "trees", they are however called aromatic trees in the literature as an analogy with carbon chemistry.

Example. *Let* $\gamma =$  *and* $\widetilde{\gamma} =$  *in* $\mathcal{AT}$ *where we added indices to apply the formula of Definition 2.4.4. Note that there is no index for the root. Then, the associated differentials are*

$$F(\gamma)(f) = \sum_{i,j,k,l,m=1}^{d} \partial_m f_m f_i \partial_i f_j f_k \partial_{j,k} f_l \partial_l f = \mathrm{div}(f) \cdot f' f''(f' f, f),$$

$$F(\widetilde{\gamma})(f) = \sum_{i,j,k,l,m=1}^{d} \partial_l f_m \partial_{m,k} f_l f_k f_i \partial_i f_j \partial_j f = \sum_{m=1}^{d} f'_m((\partial_m f)'(f)) \cdot f' f' f.$$

We add the empty tree $\varnothing$ that satisfies $F(\varnothing)(f) = I_d$. An aromatic Butcher series is then defined as the following.

Definition 2.4.5. *Let* $a : \mathcal{AT} \cup \{\varnothing\} \to \mathbb{R}$ *a map that satisfies* $a(\varnothing) = 1$*, and let* $f : \mathbb{R}^d \to \mathbb{R}^d$ *be a smooth functions, then the aromatic B-series* $B(a)(f)$ *is a formal series indexed over* $\mathcal{AT} \cup \{\varnothing\}$ *defined by*

$$B(a)(f) = \sum_{\gamma \in \mathcal{AT} \cup \{\varnothing\}} h^{|\gamma|} a(\gamma) F(\gamma)(f).$$

We extend the definition of F *on* $\mathrm{Span}(\mathcal{AT} \cup \{\varnothing\})$ *by writing*

$$F\Big(\sum_{\gamma \in \mathcal{AT} \cup \{\varnothing\}} h^{|\gamma|} a(\gamma) \gamma \Big)(f) = B(a)(f).$$

Remark 2.4.6. *B-series and aromatic B-series satisfy geometric properties [105, 100, 56] and have a structure of Hopf algebras [31, 13] (see also [17]). The new extension of aromatic B-series, named exotic aromatic B-series, that we present in Chapter 3 satisfies an isometric equivariance property (see Section 3.2.4). The study of the algebraic and geometric properties of the exotic aromatic B-series is matter for future work.*

CHAPTER 3

Exotic aromatic B-series for the computation of order conditions

Note: This chapter is based on the articles [83] ($\mathbb{R}^d$ case) and [85] (manifold case), both in collaboration with G. Vilmart. Section 3.3.3 is an unpublished work that presents a methodology for computing the composition of differential operators with exotic aromatic B-series.

3.1 Introduction

The aim of this chapter is to provide a new unified algebraic framework based on aromatic trees and B-series, with a set of trees independent of the dimension d of the problem, for the systematic study of the order conditions for the invariant measure of a class of numerical integrators that includes Runge-Kutta type schemes for Langevin dynamics. This new unified framework permits to recover some schemes, to simplify the calculations in [5] and in [123, 15, 1] for postprocessed integrators, to obtain the order conditions of a class of Runge-Kutta methods and to study partitioned systems or non-reversible perturbations. The reader only interested in the numerical methods can jump directly to Chapter 4, where we present the high order integrators for sampling the invariant measure, the order conditions and the numerical experiments.

The usage of trees and B-series[1] is known to be as a powerful standard tool for the numerical analysis of differential equations. Originally introduced by Hairer and Wanner in [66], and based on the works of Butcher [22, 24], B-series have proved to be a powerful standard tool for the numerical analysis of deterministic differential equations, as presented, for instance, in the textbooks [65, 25, 26]. In the last decades, several works extended B-series to the stochastic context. We mention in particular Burrage and Burrage [20, 21] and Komori, Mitsui and Sugiura [78] who first introduced stochastic trees and B-series for studying the order conditions of strong convergence of SDEs, Rößler [110, 111, 112, 113, 114] and Debrabant and Kværnø [46, 45, 47] for the design and analysis of high order weak and strong integrators on a finite time interval, [10] for creating schemes preserving quadratic invariants, and [81], where

[1] originally named Butcher-series

tree series were applied to a class of stochastic differential algebraic equations (SDAEs) for the computation of strong order conditions. We mention that no such algebraic framework was known for the calculation of order conditions for sampling the invariant measure of ergodic dynamics. In [36, 72], an extension of the original B-series, called aromatic B-series, was used to study deterministic volume-preserving integrators. It allowed in particular to represent the divergence of a B-series. We also recall that Butcher series have important links with other fields of mathematics, in particular the Hopf algebra of trees [42] from A. Connes and D. Kreimer in the theory of renormalization in quantum physics in the 1980s, and the theory of regularity structure [19] of M. Hairer in the context of SPDEs in the 2010s (see the review article [101]). B-series and aromatic B-series were also studied in [105, 100, 56] for their geometric properties, and in [31, 13] for their algebraic structure of Hopf algebras. The formalism that we present in this chapter is a new extension of the aromatic B-series, and is the first algebraic framework specifically designed for the computation of order conditions for the invariant measure.

Analogously to [111] (we study here the additive noise case), we consider Runge-Kutta methods for solving (2.1.3) in the context of $\mathbb{R}^d$ of the form[2]

$$\begin{aligned} Y_i &= X_n + h\sum_{j=1}^{s} a_{ij} f(Y_j) + \sigma\sqrt{h}\sum_{k=1}^{l} d_i^{(k)}\xi_n^{(k)}, \qquad i = 1,\dots,s, \\ X_{n+1} &= X_n + h\sum_{i=1}^{s} b_i f(Y_i) + \sigma\sqrt{h}\xi_n^{(1)}, \end{aligned} \tag{3.1.1}$$

and the following class of projection methods for solving (2.1.4) on a manifold,

$$\begin{aligned} &Y_i = X_n + h\sum_{j=1}^{s} a_{ij} f(Y_j) + \sigma\sqrt{h}\sum_{k=1}^{l} d_i^{(k)}\xi_n^{(k)} + \lambda_i \sum_{j=1}^{s} \widehat{a}_{ij} g(Y_j), \quad i = 1,\dots,s, \\ &\zeta(Y_i) = 0 \quad \text{if} \quad \delta_i = 1, \quad i = 1,\dots,s, \\ &X_{n+1} = Y_s, \end{aligned} \tag{3.1.2}$$

where $A = (a_{ij}), \widehat{A} = (\widehat{a}_{ij}) \in \mathbb{R}^{s\times s}$, $b = (b_i) \in \mathbb{R}^s$, $d^{(k)} = (d_i^{(k)}) \in \mathbb{R}^s$, $\delta_i = \sum_{j=1}^{s} \widehat{a}_{ij} \in \{0,1\}$ are the given Runge-Kutta coefficients, and $\xi_n^{(k)} \sim \mathcal{N}(0, I_d)$ are independent Gaussian random vectors. In the manifold case (3.1.2), we fix $\delta_s = 1$ so that $X_{n+1} \in \mathcal{M}$ and we ask that if $\delta_i = 0$, then $\widehat{a}_{ij} = 0$ for $j = 1,\dots,s$ (internal stages without projection, $Y_i \notin \mathcal{M}$ a.s.). Further details on the choice of these discretizations and their implementation is given in Chapter 4. We highlight once again that we focus in this work on the high order p of accuracy for the invariant measure, while the order r of accuracy in the weak sense can remain low (typically $r = 1$). The analysis we present applies to the classes of methods (3.1.1)-(3.1.2) for any number l of random vectors in the internal stages. However we shall often consider $l = 1$ random vector per internal stage[3], which is sufficient to achieve order two for the invariant measure. In particular, we shall consider the θ-method as an illustrative example in this chapter and recover known results on its accuracy. It is defined for θ fixed as

$$X_{n+1} = X_n + h(1-\theta) f(X_n) + h\theta f(X_{n+1}) + \sigma\sqrt{h}\xi_n, \tag{3.1.3}$$

For $\theta = 0$, we get the explicit Euler-Maruyama method (2.1.7) while the scheme is implicit

[2] Note that the internal stages Y_i depend on n, but this dependence is omitted for brevity of the notation.

[3] In this case, we denote $d = d^{(1)}$ and $\xi_n = \xi_n^{(1)}$.

for $\theta \neq 0$. It can be put in Runge-Kutta form (3.1.1) for $l = 1$ with the following coefficients.

$$\begin{array}{c|c|c} c & A & d \\ \hline & b & \end{array} = \begin{array}{c|cc|c} 0 & 0 & 0 & 0 \\ 1 & 1-\theta & \theta & 1 \\ \hline & 1-\theta & \theta & \end{array}$$

This chapter is organized as follows. In Section 3.2, we introduce a new generalization of B-series, called "exotic aromatic B-series" by considering an additional new type of edge called "liana" compared to standard aromatic B-series. We also show that these new exotic aromatic B-series satisfy an isometric equivariance property. In Section 3.3, we explain how this new algebraic framework applies for the long time accuracy analysis of stochastic integrators for ergodic problems by writing with exotic aromatic forests the operations of taking the expectation, integrating by parts and composing differential operators.

3.2 Exotic aromatic B-series: definition and algebraic properties

In this section, we introduce a new modification, called exotic aromatic B-series, of the standard aromatic B-series defined in Section 2.4.2 and originally introduced in [36, 72], that is well suited for computing order conditions for sampling the invariant measure of ergodic SDEs. We rely on the graph presentation of the formalism in the spirit of [13].

3.2.1 Exotic aromatic forests

In this subsection, we introduce a new kind of edge, called liana, and non-oriented edges. These additions allow us to represent new terms with our graphs, such as the Laplacian or the scalar product.

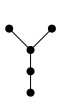
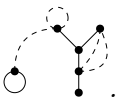
We split the set of edges into $E = E_0 \cup E_S$ where E_0 are the standard oriented edges as defined in Section 2.4.2, and where E_S is a new set of non-oriented edges represented as double horizontal straight lines. If $(v, w) = (w, v) \in E_S$, we consider this edge as an outgoing edge for both v and w, but v and w are not predecessors of each other. If $(v, w) \in E_S$, we denote $S(v) = w$ and $S(v) = v$ otherwise. We again consider graphs where each node has exactly one outgoing edge, except one node, called the root r, that has none. Such a graph (V, E) can still be decomposed into two kinds of connected components: one that contains the root, that we name the rooted tree, and the other components that we name aromas.

We also introduce a new kind of edge, called a liana, for the aromatic forests. The corresponding generalization is called exotic aromatic forests. Let (V, E) be a directed graph as previously defined and L be a finite list of pairs of elements of V (possibly with duplicates). We say that two such directed graphs (V^1, E^1, L^1) and (V^2, E^2, L^2) are equivalent if there exists a bijection $\varphi : V^1 \to V^2$ such that

$$\varphi(V^1) = V^2, \quad (\varphi \times \varphi)(E^1) = E^2, \quad (\varphi \times \varphi)(E^1_S) = E^2_S \quad (\varphi \times \varphi)(L^1) = L^2.$$

We call exotic aromatic forests the equivalence classes of these directed graphs $\gamma = (V, E, L)$. The elements of L are called lianas and correspond to non-oriented edges between any two nodes of the forest. We graphically represent them with a dashed edge linking the two given nodes. As we authorize duplicates, there can be several lianas between two given nodes. Also lianas can link a node to itself. For a node v, $\Gamma(v)$ denotes the list of the lianas (also with

possible duplicates) linked to v. The predecessors of v only take in account the edges of E. An exotic aromatic tree of an exotic aromatic forest $\gamma = (V, E, L)$ is a connected component of the associated aromatic forest (V, E). We call $\mathcal{EAT}$ the set of exotic aromatic forests with exactly one rooted tree, and name its elements exotic aromatic rooted forests.

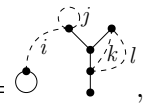
Example. *The lianas can link different trees of an aromatic forest and thus yield an exotic aromatic forest. For instance, linking the aroma* *and the rooted tree*  *gives*  .

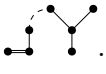
The elementary differential associated with an exotic aromatic forest is given by the following definition.

Definition 3.2.1. *Let $\gamma = (V, E, L) \in \mathcal{EAT}$, and let f be a smooth function. We denote $l_1, \dots, l_s$ the elements of L, $v_1, \dots, v_m$ the elements of $V \setminus \{r\}$ and $\delta_{i,j}$ the Kronecker delta. For $v \in V$, we denote $I_{\pi(v)} = (i_{q_1}, \dots, i_{q_s})$ where $\pi(v) = \{q_1, \dots, q_s\}$ are the predecessors of v, and $J_{\Gamma(v)} = (j_{l_{x_1}}, \dots, j_{l_{x_t}})$ where $\Gamma(v) = \{l_{x_1}, \dots, l_{x_t}\}$ are the lianas linked to v. Then $F(\gamma)$ is defined as*

$$F(\gamma)(f) = \sum_{i_{v_1},\dots,i_{v_m}=1}^{d} \sum_{j_{l_1},\dots,j_{l_s}=1}^{d} \left(\prod_{v \in V} \delta_{i_v, i_{S(v)}} \partial_{I_{\pi(v)}} \partial_{J_{\Gamma(v)}} f_{i_v} \right) \partial_{I_{\pi(r)}} \partial_{J_{\Gamma(r)}} f.$$

Examples. *The differential that corresponds to the rooted tree*  *with a single node and a single liana is* $F($$)(f) = \Delta f$. *We can also represent as exotic aromatic forest more complicated derivatives. For instance, we have*

$$F(\gamma)(f) = \sum_{i,j,k=1}^{d} \operatorname{div}(\partial_i f) \cdot f'((\partial_{kl} f)'(f''(\partial_{ijj} f, \partial_{kl} f))) \quad \textit{for} \quad \gamma = \text{  },$$

$$F(\gamma)(f) = \sum_{i=1}^{d} (f, f'(\partial_i f)) f' f''(\partial_i f, f) \quad \textit{for} \quad \gamma = \text{  }.$$

3.2.2 Exotic aromatic forests for the computation of order conditions

For studying the order conditions for the invariant measure of numerical integrators, we introduce a few modifications to exotic aromatic forests. We decompose the set of nodes into $V = V_f \cup V_g \cup V_\times \cup \{r\}$ where V_f are the nodes representing a function $f : \mathbb{R}^d \to \mathbb{R}^d$ and are represented with black disks, V_g are the nodes representing a function $g : \mathbb{R}^d \to \mathbb{R}^d$ and are represented with white disks. The set $V_\times$ gathers the grafted nodes, graphically represented by a cross, that represent a vector $\xi \in \mathbb{R}^d$. These nodes have exactly one outgoing edge and no ingoing edge (including lianas). We call these graphs the grafted exotic aromatic rooted forests, that we denote $\mathcal{EAT}_\times$. For simplicity, we call exotic aromatic forests, and denote $\mathcal{EAT}$, the set of grafted exotic aromatic rooted forests satisfying $V_\times = \varnothing$. Note that these new trees can be seen as multicoloured trees in the context of P-series (see [65, Chap. 3]), where the nodes represented with crosses cannot have predecessors.

We write $N_f(\gamma)$ the number of elements of V_f (respectively $N_g(\gamma)$ the number of elements of V_g and $N_\times(\gamma)$ the number of elements of $V_\times$), $N_l(\gamma)$ the number of lianas and we assume

that $N_g(\gamma)$ is even. The order of a directed graph $\gamma = (V, E, L)$ is defined as

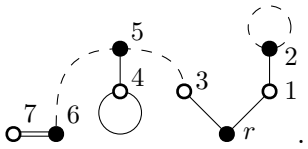
$$|\gamma| = N_f(\gamma) + N_l(\gamma) + \frac{N_g(\gamma) + N_\times(\gamma)}{2} - |E_S|\,.$$

For instance, the graph $\gamma = (V, E, L)$ with

$$V_f = \{v_2, v_5, v_6\}, \quad V_g = \{v_1, v_3, v_4, v_7\}, \quad E_S = \{(v_6, v_7)\}, \tag{3.2.1}$$

$$E_0 = \{(v_1, r), (v_2, v_1), (v_3, r), (v_4, v_4), (v_5, v_4)\}, \quad L = \{(v_2, v_2), (v_3, v_5), (v_5, v_6)\},$$

satisfies $|\gamma| = 7$ and is represented as



In addition, we need a different set of rooted forests where the root is in V_f or V_g. We call them exotic aromatic vector fields and gather them together in the set $\mathcal{EAV}$. The elementary differential associated with an exotic aromatic forest is given by the following definition.

Definition 3.2.2. *Let $\gamma = (V, E, L) \in \mathcal{EAT}$, $\sigma \in \mathbb{R}$, $\xi \in \mathbb{R}^d$, and let f, $g\colon \mathbb{R}^d \to \mathbb{R}^d$ and $\phi : \mathbb{R}^d \to \mathbb{R}$ be smooth functions, and the Gram function $G = g^T g \in \mathbb{R}^*$. The associated elementary differential of γ is, with the same notation as in Definition 3.2.1,*

$$F(\gamma)(f, g, \xi, \phi) = \sigma^{2(|\gamma| - N_f(\gamma))} G^{-N_g(\gamma)/2} \sum_{i_{v_1}, \dots, i_{v_m} = 1}^{d} \sum_{j_{l_1}, \dots, j_{l_s} = 1}^{d} \left(\prod_{v \in V_f} \delta_{i_v, i_{S(v)}} \partial_{I_{\pi(v)}} \partial_{J_{\Gamma(v)}} f_{i_v} \right)$$
$$\cdot \left(\prod_{v \in V_g} \delta_{i_v, i_{S(v)}} \partial_{I_{\pi(v)}} \partial_{J_{\Gamma(v)}} g_{i_v} \right) \cdot \left(\prod_{v \in V_\times} \delta_{i_v, i_{S(v)}} \xi_{i_v} \right) \partial_{I_{\pi(r)}} \partial_{J_{\Gamma(r)}} \phi.$$

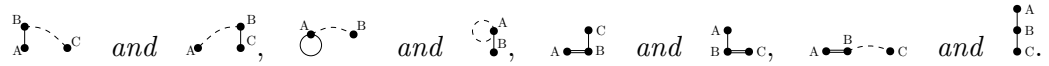
Example. *The differential associated with the exotic aromatic forests*  *and γ given by the example* (3.2.1) *are* $F($  $)(f, g, \xi, \phi) = \phi'(f''(\xi, \xi))$ *and*

$$F(\gamma)(f, g, \xi, \phi) = \sigma^8 G^{-2} \sum_{i_{v_1}, \dots, i_{v_7} = 1}^{d} \sum_{j_{l_1}, \dots, j_{l_3} = 1}^{d} \partial_{j_{l_1} j_{l_1}} f_{i_{v_2}} \partial_{j_{l_2} j_{l_3}} f_{i_{v_5}} \delta_{i_{v_6}, i_{v_7}} \partial_{j_{l_3}} f_{i_{v_6}}$$
$$\cdot \partial_{i_{v_2}} g_{i_{v_1}} \partial_{j_{l_2}} g_{i_{v_3}} \partial_{i_{v_4} i_{v_5}} g_{i_{v_4}} \delta_{i_{v_7}, i_{v_6}} g_{i_{v_7}} \partial_{i_{v_1} i_{v_3}} \phi.$$

If V_g and E_S are empty, we recover the exotic aromatic forests from [83]. For the rest of the thesis (with the exception of Section 3.2.4), we will use only Definition 3.2.2 for the differential associated with an exotic aromatic forest. For brevity of notation, we also write $F(\gamma)(\phi)$ instead of $F(\gamma)(f, g, \xi, \phi)$. Note that $\phi \to F(\gamma)(\phi)$ is a linear differential operator (that depends on f, g and ξ).

The functions f and g are gradient vector fields, according to Assumption 2.2.2. Different forests thus represent the same differential. The following simplification rules show how to identify such forests.

Proposition 3.2.3 (Simplification rules)**.** *The two forest patterns gathered in each of the following pairs represent the same differential:*



One can replace the nodes A, B, C with white nodes and the result remains.

Proof. For the first pair of patterns, the associated differentials have the following respective forms: $\sum_{i=1}^{d} \partial_j f_k \partial_i f_j \partial_i f_l$ and $\sum_{j=1}^{d} \partial_j f_k \partial_j f_i \partial_i f_l$. As f is a gradient, f' is a symmetric matrix and $\partial_i f_j = \partial_j f_i$. The two differentials are then equal. The other identifications are proved in the same way. □

Example. *Applying Proposition 3.2.3 yields the following equality of differentials,*

3.2.3 Exotic aromatic B-series

In this section, we adapt the formalism of aromatic B-series of [105] to grafted exotic aromatic forests, in order to use it as a numerical tool for weak Taylor expansions.

Definition 3.2.4. *Let $a : \mathcal{EAT}_\times \to \mathbb{R}$ a map, then the grafted exotic aromatic B-series $B(a)(\phi)$ is a formal series indexed over $\mathcal{EAT}_\times$ defined by*

$$B(a)(\phi) = \sum_{\gamma \in \mathcal{EAT}_\times} h^{|\gamma|} a(\gamma) F(\gamma)(\phi).$$

We extend the definition of F on $\mathrm{Span}(\mathcal{EAT}_\times)$ by writing

$$F\left(\sum_{\gamma \in \mathcal{EAT}_\times} h^{|\gamma|} a(\gamma) \gamma \right)(\phi) = B(a)(\phi).$$

The variable h is formal and thus can be chosen to be equal to 1. One can make a similar definition for the non-grafted (exotic) aromatic rooted forests of Section 3.2.1 by replacing $F(\gamma)(\phi)$ with the $F(\gamma)(f)$ introduced in Definition 3.2.1. We focus on these non-grafted exotic aromatic B-series in Section 3.2.4.

Remark 3.2.5. *The coefficients $a(\gamma)$ of standard B-series are sometimes renormalized as $\frac{a(\gamma)}{\rho(\gamma)}$ where ρ is a function determined by the symmetries of the associated forest. If ρ is appropriately chosen, it greatly simplifies the composition laws of (aromatic) B-series (see [65, 31, 13]). Finding the best definition of ρ for this exotic extension of B-series is matter for future work.*

3.2.4 Isometric equivariance of exotic aromatic forests

In this subsection, we show that the exotic aromatic B-series satisfy an isometric equivariance property in the spirit of [105, 100]. We consider exotic aromatic rooted forests γ with $N_g(\gamma) = 0$, and we add a new tree: the empty tree $\varnothing$. The function F is then extended on $\mathcal{EAT} \cup \{\varnothing\}$ by $F(\varnothing)(f) = \mathrm{Id}_{\mathbb{R}^d}$. Then, for a function $a : \mathcal{EAT} \cup \{\varnothing\} \to \mathbb{R}$, the associated exotic aromatic B-series is

$$B(a)(f) = \sum_{\gamma \in \mathcal{EAT} \cup \{\varnothing\}} a(\gamma) F(\gamma)(f).$$

We study (exotic) aromatic B-series $B(a)$ with $a(\varnothing) = 1$. We call these (exotic) aromatic B-series methods. Let G be a subgroup of $\mathrm{GL}_d(\mathbb{R}) \ltimes \mathbb{R}^d$, let the action of an element $(A, b) \in G$ on $\mathbb{R}^d$ be $x \mapsto Ax + b$, and let the action on a vector field $f : \mathbb{R}^d \to \mathbb{R}^d$ be

$$((A, b) * f)(x) := Af(A^{-1}(x - b)).$$

We simplify the notation by writing $A * f := (A, 0) * f$. We recall the definition of equivariance from [105]. The property of equivariance means the method stays unchanged when applying an affine coordinate transformation. Let Φ be a differential operator and let G be a subgroup of $\mathrm{GL}_d(\mathbb{R}) \ltimes \mathbb{R}^d$, then Φ is called G-equivariant if

$$\forall (A, b) \in G, \quad \forall f \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^d), \quad \Phi((A, b) * f) = (A, b) \circ \Phi(f) \circ (A, b)^{-1}.$$

In particular, Φ is said to be affine equivariant if $G = \mathrm{GL}_d(\mathbb{R}) \ltimes \mathbb{R}^d$ and isometric equivariant if $G = \mathrm{O}_d(\mathbb{R}) \ltimes \mathbb{R}^d$.

Theorem 3.2.6. *Consider an exotic aromatic B-series method $B(a)$, then $B(a)$ is isometric equivariant.*

Proof. By an argument of linearity, it suffices to prove the $\mathrm{O}_d(\mathbb{R})$-equivariance of every exotic aromatic forest. Let $\gamma = (V, E, L) \in \mathcal{EAT}$, f a smooth function. We name r the root and $V^0 = V \setminus \{r\} = \{v_1, \dots, v_m\}$ the other nodes. We denote $l_1, \dots, l_s$ the elements of L. Then

$$F(\gamma)(f) = \sum_{i_{v_1}, \dots, i_{v_m} = 1}^{d} \sum_{j_{l_1}, \dots, j_{l_s} = 1}^{d} \left(\prod_{v \in V^0} \delta_{i_v, i_{S(v)}} \partial_{I_{\pi(v)}} \partial_{J_{\Gamma(v)}} f_{i_v} \right) \partial_{I_{\pi(r)}} \partial_{J_{\Gamma(r)}} f.$$

Let $A \in \mathrm{O}_d(\mathbb{R})$ and $x \in \mathbb{R}^d$, we would like to prove that

$$F(\gamma)(A * f)(x) = (F(\gamma)(Af(A^{-1}.)))(x) = A(F(\gamma)(f))(A^{-1}x). \tag{3.2.2}$$

On the first hand, we have

$$\begin{aligned}(F(\gamma)(A * f))_{j_0}(x) = &\sum_{i_{v_1}, \dots, i_{v_m} = 1}^{d} \sum_{j_{l_1}, \dots, j_{l_s} = 1}^{d} \\ &\left(\prod_{v \in V^0} \sum_{k=1}^{d} \sum_{P_{\pi(v)}, Q_{\Gamma(v)}} a_{i_v, k} \delta_{i_v, i_{S(v)}} a_{I_{\pi(v)}, P_{\pi(v)}} a_{J_{\Gamma(v)}, Q_{\Gamma(v)}} \partial_{P_{\pi(v)}} \partial_{Q_{\Gamma(v)}} f_k \right)(A^{-1}x) \\ &\cdot \left(\sum_{k=1}^{d} \sum_{P_{\pi(r)}, Q_{\Gamma(r)}} a_{j_0, k} a_{I_{\pi(r)}, P_{\pi(r)}} a_{J_{\Gamma(r)}, Q_{\Gamma(r)}} \partial_{P_{\pi(r)}} \partial_{Q_{\Gamma(r)}} f_k \right)(A^{-1}x),\end{aligned}$$

where we denote $a_{I,J} := \prod_{i \in I, j \in J} a_{i,j}$ and $\sum_{P_{\pi(v)}, Q_{\Gamma(v)}} := \sum_{v' \in \pi(v), l' \in \Gamma(v)} \sum_{p_{v'}=1}^{d} \sum_{q_{l'}=1}^{d}$.

First step: simplification of edges. For every edge of the form $(v, v') \in E_0$, the following sum appears

$$\sum_{i_v, k, p_v} a_{i_v, k} a_{i_v, p_v} (\partial_{p_v} g_l) h_k = \sum_{p_v} (\partial_{p_v} g_l) h_{p_v},$$

where g and h are some derivatives of f, and where we used the isometric property of A. Moreover, for every edge of the form $(v, v') \in E_S$, the following sum appears

$$\sum_{i_v, i_{v'}, k_1, k_2} \delta_{i_v, i_{v'}} a_{i_v, k_1} a_{i_{v'}, k_2} g_{k_1} h_{k_2} = \sum_{k_1, k_2} \delta_{k_1, k_2} g_{k_1} h_{k_2},$$

where g and h are again certain derivatives of f. Thus, we can simplify our expression into

$$\begin{aligned}(F(\gamma)(A * f))_{j_0}(x) = \sum_{p_{v_1}, \dots, p_{v_m}=1}^{d} \sum_{j_{l_1}, \dots, j_{l_s}=1}^{d} \\ \left(\prod_{v \in V^0} \sum_{Q_{\Gamma(v)}} \delta_{p_v, p_{S(v)}} a_{J_{\Gamma(v)}, Q_{\Gamma(v)}} \partial_{P_{\pi(v)}} \partial_{Q_{\Gamma(v)}} f_{p_v} \right)(A^{-1}x) \\ \cdot \left(\sum_{k=1}^{d} \sum_{Q_{\Gamma(r)}} a_{j_0, k} a_{J_{\Gamma(r)}, Q_{\Gamma(r)}} \partial_{P_{\pi(r)}} \partial_{Q_{\Gamma(r)}} f_k \right)(A^{-1}x).\end{aligned}$$

Second step: simplification of lianas. If l is a liana linking v and v', then the following sum appears

$$\sum_{j_l, q_l, q_l'} a_{j_l, q_l} (\partial_{q_l} g_{p_v}) a_{j_l, q_l'} (\partial_{q_l'} h_{p_{v'}}) = \sum_{q_l} (\partial_{q_l} g_{p_v})(\partial_{q_l} h_{p_{v'}})$$

where g and h are some derivatives of f. Note that the number of sums over indexes of the form q_l is two times the number of lianas, and we identify these pairs of sums according to lianas. We find

$$\begin{aligned}(F(\gamma)(A * f))_{j_0}(x) = \sum_{p_{v_1}, \dots, p_{v_m}=1}^{d} \sum_{q_{l_1}, \dots, q_{l_s}=1}^{d} \\ \left(\prod_{v \in V^0} \delta_{p_v, p_{S(v)}} \partial_{P_{\pi(v)}} \partial_{Q_{\Gamma(v)}} f_{p_v} \right)(A^{-1}x) \\ \cdot \left(\sum_{k=1}^{d} a_{j_0, k} \partial_{P_{\pi(r)}} \partial_{Q_{\Gamma(r)}} f_k \right)(A^{-1}x).\end{aligned}$$

This shows (3.2.2), and thus the isometric equivariance property. □

Note that, if $E_S = \varnothing$, we do not need A to be isometric to write the first step of the proof, assuming A invertible is enough. If we do not assume A isometric, the term $a_{I_{\pi(v)}, P_{\pi(v)}}$ becomes $b_{P_{\pi(v)}, I_{\pi(v)}}$, where $B = A^{-1}$, and the sums still simplify. Thus the first step proves the affine equivariance of aromatic B-series methods.

Remark 3.2.7. *It is proved in [105] that standard B-series methods are exactly the affine equivariant methods preserving the stationary states of ODEs. Analogously, it would be interesting to characterize the isometric equivariant maps.*

3.3 Analysis of invariant measure order conditions using exotic aromatic forests

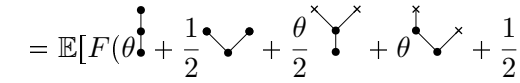
In this section, we show how the exotic aromatic B-series formalism presented in Section 3.2 applies for the study of order conditions for the invariant measure.

3.3.1 Weak Taylor expansion using exotic aromatic forests

Let us begin this subsection with the example of the θ-method (3.1.3) for solving (2.1.3) in the context of $\mathbb{R}^d$. We apply the standard methodology to expand in Taylor series $\mathbb{E}[\phi(X_1)|X_0 = x]$ as $h \to 0$. We refer to [126, 2] for other examples of analogous calculations performed without exotic aromatic forests. Under $X_0 = x$, we have

$$X_1 = x + \sqrt{h}\sigma\xi + hf + h\sqrt{h}\theta\sigma f'\xi + h^2\theta f'f + h^2\frac{\theta\sigma^2}{2}f''(\xi,\xi) + \dots$$

Then we deduce $\mathbb{E}[\phi(X_1)|X_0 = x] = \phi(x) + h\mathcal{L}\phi(x) + h^2\mathcal{A}_1\phi(x) + \dots$, where

$$\begin{aligned}\mathcal{A}_1\phi &= \mathbb{E}[\theta\phi'f'f + \frac{1}{2}\phi''(f,f) + \frac{\theta\sigma^2}{2}\phi'f''(\xi,\xi) + \theta\sigma^2\phi''(f'\xi,\xi)\\ &\quad + \frac{\sigma^2}{2}\phi^{(3)}(f,\xi,\xi) + \frac{\sigma^4}{24}\phi^{(4)}(\xi,\xi,\xi,\xi)]\\ &= \mathbb{E}[F(\theta$$  $$)(\phi)].\end{aligned} \tag{3.3.1}$$

All the forests with an odd number of grafted nodes vanished because odd moments of a centred Gaussian random variable are zero. The expectation of the differential of a forest with exactly two grafted nodes comes straightforwardly:

$$\mathbb{E}[F($$  $$)(\phi)] = \mathbb{E}[\sigma^2\phi'(f''(\xi,\xi))] = \sigma^2\sum_{i,j,k}\partial_i\phi\partial_{jk}f_i\mathbb{E}[\xi_j\xi_k] = \sigma^2\sum_{i,j}\partial_i\phi\partial_{jj}f_i = F($$  $$)(\phi),$$

where $\mathbb{E}[\xi_j\xi_k] = 0$ for $j \neq k$ by using the independence of the ξ_i's. We see that taking the expectation of the differential associated with a grafted tree amounts to linking the grafted nodes with lianas in all possible manners. For instance, for the following example with four grafted nodes:

$$\begin{aligned}\mathbb{E}[F($$  $$)(\phi)] &= \sigma^4\mathbb{E}[\phi^{(4)}(\xi,\xi,\xi,\xi)] = \sigma^4\sum_{i,j,k,l}\partial_{i,j,k,l}\phi\mathbb{E}[\xi_i\xi_j\xi_k\xi_l]\\ &= \sigma^4\sum_i\partial_{i,i,i,i}\phi\mathbb{E}[\xi_i^4] + 3\sigma^4\sum_{\substack{i,j\\ i\neq j}}\partial_{i,i,j,j}\phi\mathbb{E}[\xi_i^2]\mathbb{E}[\xi_j^2] = 3\sigma^4\sum_{i,j}\partial_{i,i,j,j}\phi\\ &= 3F($$ $$)(\phi).\end{aligned}$$

Let us now comment this computation. The interesting fact is that $\mathbb{E}[\xi_i^4] = 3$ corresponds exactly to the number of ways to gather the indices i, j, k and l in pairs. This observation makes an exotic aromatic tree naturally appear. However, we took here only four grafted nodes and the differential form was symmetric in the arguments ξ. We need to study the expectation of general exotic aromatic forest elementary differentials. This is the aim of the following result.

Theorem 3.3.1. *Let $\gamma \in \mathcal{EAT}_\times$ be a grafted exotic aromatic forest with an even number of grafted nodes $2n$, let $\phi : \mathbb{R}^d \to \mathbb{R}$ be a smooth function, and let $V^\times = \{c_1, \dots, c_{2n}\}$ be the set of grafted nodes of γ. We call $\mathcal{P}_2(2n)$ the set of partitions by pair of $\{1, \dots, 2n\}$, i.e. the set of surjections $p : \{1, \dots, 2n\} \to \{1, \dots, n\}$ such that the preimage of each singleton has exactly two elements and the minima of those preimages follow an ascending order ($\min(p^{-1}(\{i\})) < \min(p^{-1}(\{j\}))$ for $i < j$). Finally we define $\varphi_\gamma : \mathcal{P}_2(2n) \to \mathcal{EAT}$ the application that maps the*

partition p of γ to the aromatic forest where the grafted nodes are linked by lianas according to p. Then, the expectation of $F(\gamma)(\phi)$ is given by

$$\mathbb{E}[F(\gamma)(\phi)] = \sum_{p \in \mathcal{P}_2(2n)} F(\varphi_\gamma(p))(\phi).$$

This theorem states that taking the expectation of the differential associated with a forest amounts to sum the forests obtained by linking the grafted nodes together pairwise using lianas in all possible manners and to take the associated differential.

Example. *Let us take $\gamma \in \{\bullet, \, \vee, \, \vee\!\!\vee, \dots\}$ the tree with only a root, $2n$ grafted nodes and no liana, then*

$$\mathbb{E}[F(\gamma)(\phi)] = \frac{\sigma^{2n}(2n)!}{2^n n!} \Delta^n \phi$$

The integer $\frac{(2n)!}{2^n n!}$ is exactly the number of ways to gather the grafted nodes by pairs. An other example is

$$\mathbb{E}[F(\;\vee\!\!\vee\;)(\phi)] = 3F(\;\overset{\bullet}{\bullet}\;)(\phi),$$

where the coefficient 3 accounts for the number of choices for linking the grafted nodes pairwise. There is a subtlety for forests with double edges, where Proposition 3.2.3 allows some simplifications. In particular, we have

$$\mathbb{E}[F(\;\bullet\!\!-\!\!\bullet\;)(\phi)] = \mathbb{E}[F(\;\bullet\!\!=\!\!\bullet\;)(\phi)] = F(\;\bullet\!\!=\!\!\bullet\;)(\phi) = F(\;\overset{\bullet}{\bullet}\;)(\phi) \quad \textit{and} \quad \mathbb{E}[F(\;\bullet\!\!-\!\!\bullet\;)(\phi)] = F(\;\bigcirc\!\!\bullet\;)(\phi).$$

Theorem 3.3.1 follows from the following lemma, which is an extension of the Isserlis theorem [73] to the case of multilinear mappings. The Isserlis theorem states that if χ is a $2n$-dimensional Gaussian random vector with mean zero and arbitrary covariance, then

$$\mathbb{E}\left[\prod_{i=1}^{2n} \chi_i\right] = \sum_{p \in \mathcal{P}_2(2n)} \prod_{\substack{i<j \\ p(i)=p(j)}} \mathbb{E}[\chi_i \chi_j].$$

For $n = 2$, it gives $\mathbb{E}[\chi_1\chi_2\chi_3\chi_4] = \mathbb{E}[\chi_1\chi_2]\mathbb{E}[\chi_3\chi_4] + \mathbb{E}[\chi_1\chi_3]\mathbb{E}[\chi_2\chi_4] + \mathbb{E}[\chi_1\chi_4]\mathbb{E}[\chi_2\chi_3]$.

Lemma 3.3.2. *Let $B : \mathbb{R}^d \times \cdots \times \mathbb{R}^d = \mathbb{R}^{2nd} \to \mathbb{R}$ be a $2n$-multilinear form, and let ξ be a Gaussian vector $\mathcal{N}(0, I_d)$, then*

$$\mathbb{E}[B(\xi, \dots, \xi)] = \sum_{p \in \mathcal{P}_2(2n)} \sum_{i_1, \dots, i_n = 1}^{d} B(e_{i_p}), \tag{3.3.2}$$

with $e_{i_p} = (e_{i_{p(1)}}, \dots, e_{i_{p(2n)}})$, and we recall that $e_1, \dots, e_d$ denotes the canonical basis of $\mathbb{R}^d$.

Proof. For the case of an elementary multilinear form $B_\sigma : (x_1, \dots, x_{2n}) \mapsto \prod_{j=1}^{2n} (x_j)_{\sigma(j)}$ where $\sigma : \{1, \dots, 2n\} \to \{1, \dots, d\}$ is a given mapping and $(x_j)_{\sigma(j)}$ denotes the $\sigma(j)$'s component of $x_j \in \mathbb{R}^d$, the identity (3.3.2) reduces to the Isserlis theorem. As any multilinear form can be decomposed as a linear combination of such elementary multilinear forms, the result (3.3.2) is proved by linearity with respect to B. □

Proof of Theorem 3.3.1. We consider $F(\gamma)(\phi)$ as a $2n$-multilinear form $B_{\gamma,\phi}$ evaluated in the vectors $(\xi,\dots,\xi)$ (see Definition 3.2.2). Lemma 3.3.2 gives

$$\mathbb{E}[B_{\gamma,\phi}(\xi,\dots,\xi)] = \sum_{p\in\mathcal{P}_2(2n)} \sum_{i_1,\dots,i_n=1}^{d} B_{\gamma,\phi}(e_{i_p}) = \sum_{p\in\mathcal{P}_2(2n)} F(\varphi_\gamma(p))(\phi),$$

because $B_{\gamma,\phi}(e_{i_p})$ is the differential $F(\gamma)$ where we differentiate the non-grafted nodes linked to grafted nodes in the directions given by e_{i_p}, and $F(\varphi_\gamma(p))(\phi)$ is obtained by summing $B_{\gamma,\phi}(e_{i_p})$ over all the indices. □

A straightforward corollary of Theorem 3.3.1 is that the operators $\mathcal{L}^j/j!$ and $\mathcal{A}_j$ can be written with exotic aromatic forests.

Theorem 3.3.3. *Consider a Runge-Kutta method of the form* (3.1.1) *(resp.* (3.1.2)*), then the expansions* (2.2.5) *and* (2.2.6) *can be formally written with exotic aromatic B-series, that is, there exist two maps e and a over $\mathcal{EAT}$ such that*

$$\mathbb{E}[\phi(X(h))|X(0)=x] = B(e)(\phi)(x), \quad \mathbb{E}[\phi(X_1)|X_0=x] = B(a)(\phi)(x),$$

and where the operators are given by

$$\frac{\mathcal{L}^j}{j!} = F\Big(\sum_{\gamma\in\mathcal{EAT},|\gamma|=j} e(\gamma)\gamma\Big), \quad \mathcal{A}_{j-1} = F\Big(\sum_{\gamma\in\mathcal{EAT},|\gamma|=j} a(\gamma)\gamma\Big).$$

If $e(\gamma)=a(\gamma)$ for all $\gamma\in\mathcal{EAT}$ with $1\leqslant|\gamma|\leqslant p$, then the integrator has at least weak order p for solving (2.1.3) *(resp.* (2.1.4)*).*

For instance, the generator $\mathcal{L}$ of the SDE (2.1.4) in (2.2.3) can be written with exotic aromatic forests as

$$\begin{aligned}\mathcal{L}\phi &= \phi' f - G^{-1}(g,f)\phi' g - \frac{\sigma^2}{2}G^{-1}\operatorname{div}(g)\phi' g + \frac{\sigma^2}{2}G^{-2}(g,g'g)\phi' g + \frac{\sigma^2}{2}\Delta\phi - \frac{\sigma^2}{2}G^{-1}\phi''(g,g)\\ &= F\Big(\bullet - \bullet\!\!-\!\!\bullet - \frac{1}{2}\,\circ\!\!-\!\!\bullet + \frac{1}{2}\,\bullet\!\!-\!\!\bullet\;\bullet + \frac{1}{2}\,\bullet\;\circ - \frac{1}{2}\,\bullet\!\!<\!\!\circ\Big)(\phi).\end{aligned}$$

Note that removing the forests with white nodes yields the expression of the generator in $\mathbb{R}^d$, that is,

$$\mathcal{L}\phi = \phi' f + \frac{\sigma^2}{2}\Delta\phi = F\Big(\bullet + \frac{1}{2}\,\circ\Big)(\phi).$$

Application. *The operator $\mathcal{A}_1$ of the θ-method* (3.1.3) *is now convenient to write with exotic aromatic forests. Applying Theorem 3.3.1 to* (3.3.1)*, we deduce $\mathcal{A}_1 = F(\gamma)$ with*

$$\gamma = \theta\,\bullet + \frac{1}{2}\,\bullet\!\!<\!\!\bullet + \frac{\theta}{2}\,\circ\!\!-\!\!\bullet + \theta\,\bullet\;\bullet + \frac{1}{2}\,\bullet\;\circ + \frac{1}{8}\,\circ\;\circ. \tag{3.3.3}$$

3.3.2 Integration by parts of the exotic aromatic forests

In this subsection, we integrate by parts integrals of the form $\int_{\mathcal{M}} F(\gamma)(\phi)d\mu_\infty$ for $\gamma\in\mathcal{EAT}$. The idea is to transform a high order differential operator $\mathcal{A}:\phi\to F(\gamma)(\phi)$ into a differential operator $\mathcal{A}^0$ of lower order in ϕ such that $\int_{\mathcal{M}}\mathcal{A}\phi d\mu_\infty = \int_{\mathcal{M}}\mathcal{A}^0\phi d\mu_\infty$. The tree formalism previously defined makes this task systematic and very convenient.

Let us first begin with an example in $\mathbb{R}^d$.

$$\int_{\mathbb{R}^d} F(\bullet)(\phi)\rho_\infty dx = \sigma^2 \sum_{i,j} \int_{\mathbb{R}^d} \frac{\partial^3 \phi}{\partial x_i \partial x_j \partial x_j} f_i \rho_\infty dx \tag{3.3.4}$$
$$= -\sum_{i,j} \left[\sigma^2 \int_{\mathbb{R}^d} \frac{\partial \phi}{\partial x_i \partial x_j} \frac{\partial f_i}{\partial x_j} \rho_\infty dx + \sigma^2 \int_{\mathbb{R}^d} \frac{\partial \phi}{\partial x_i \partial x_j} f_i \frac{\partial \rho_\infty}{\partial x_j} dx \right],$$

where we integrated by parts. Note that the boundary term vanishes using the growth assumptions on ϕ. As we assumed f to be a gradient, the invariant measure satisfies $\nabla \rho_\infty = \frac{2}{\sigma^2} f \rho_\infty$ and the equality (3.3.4) can be written using exotic aromatic forests as

$$\int_{\mathbb{R}^d} F(\bullet)(\phi)\rho_\infty dx = -\sum_{i,j} \left[\sigma^2 \int_{\mathbb{R}^d} \frac{\partial \phi}{\partial x_i \partial x_j} \frac{\partial f_i}{\partial x_j} \rho_\infty dx + 2 \int_{\mathbb{R}^d} \frac{\partial \phi}{\partial x_i \partial x_j} f_i f_j \rho_\infty dx \right]$$
$$= -\int_{\mathbb{R}^d} F(\bullet)(\phi)\rho_\infty dx - 2 \int_{\mathbb{R}^d} F(\bullet)(\phi)\rho_\infty dx.$$

We notice that integrating by parts $F(\bullet)(\phi)$ with respect to the invariant measure $d\mu_\infty$ amounts to unplug a liana from the root and to replug it either to all the other nodes of the forests or to add a new edge linking to a new black node.

On a manifold, the integration by parts is a corollary of the Green theorem (see, for instance, [116, Chap. II]). It is a crucial tool for deriving order conditions for the invariant measure.

Lemma 3.3.4 (Integration by parts on $\mathcal{M}$). *If $\psi : \mathbb{R}^d \to \mathbb{R}$ and $H : \mathbb{R}^d \to \mathbb{R}^d$ are smooth functions, then*

$$\int_{\mathcal{M}} (\nabla_{\mathcal{M}} \psi, H) d\sigma_{\mathcal{M}} = -\int_{\mathcal{M}} \psi \operatorname{div}_{\mathcal{M}}(\Pi_{\mathcal{M}} H) d\sigma_{\mathcal{M}},$$

where $\nabla_{\mathcal{M}} \psi := \Pi_{\mathcal{M}} \nabla \psi$ and $\operatorname{div}_{\mathcal{M}}(H) := \operatorname{div}(H) - G^{-1}(g, H'g)$. In addition, with the invariant measure $d\mu_\infty = \rho_\infty d\sigma_{\mathcal{M}}$ and $k \geqslant 0$, we obtain

$$\int_{\mathcal{M}} \Big[G^{-k} \psi' H - G^{-(k+1)}(g, H)\psi' g \Big] d\mu_\infty = \int_{\mathcal{M}} \Big[G^{-(k+1)}(g, H'g)\psi \tag{3.3.5}$$
$$- (2k+1) G^{-(k+2)}(g, g'g)(g, H)\psi - G^{-k} \operatorname{div}(H)\psi + 2k G^{-(k+1)}(g, g'H)\psi$$
$$+ G^{-(k+1)} \operatorname{div}(g)(g, H)\psi + \frac{2}{\sigma^2} G^{-(k+1)}(g, f)(g, H)\psi - \frac{2}{\sigma^2} G^{-k}(f, H)\psi \Big] d\mu_\infty.$$

For instance, applying identity (3.3.5) with $\psi = \sigma^4 \Delta \phi'(e_i)$, $H = e_i$ and $k = 0$, and then summing on $i = 1, \ldots, d$ yields

$$\int_{\mathcal{M}} \Big[\sigma^4 \Delta^2 \phi - \sigma^4 G^{-1} \Delta \phi''(g, g) \Big] d\mu_\infty = \int_{\mathcal{M}} \Big[-\sigma^4 G^{-2}(g, g'g) \Delta \phi' g \tag{3.3.6}$$
$$+ \sigma^4 G^{-1} \operatorname{div}(g) \Delta \phi' g + 2\sigma^2 G^{-1}(g, f) \Delta \phi' g - 2\sigma^2 \Delta \phi' f \Big] d\mu_\infty.$$

Note that this calculation can be written with exotic aromatic forests as

$$\int_{\mathcal{M}} F(\bullet\!\bullet - \bullet\!\!<\!\!\bullet\!\!>\!\!\bullet)(\phi) d\mu_\infty = \int_{\mathcal{M}} F(-\bullet\!\!-\!\!\bullet\!\!<\!\!\bullet + \circ\!\!<\!\!\bullet + 2\,\bullet\!\!-\!\!\bullet\!\!<\!\!\bullet - 2\,\bullet\!\!<\!\!\bullet)(\phi) d\mu_\infty.$$

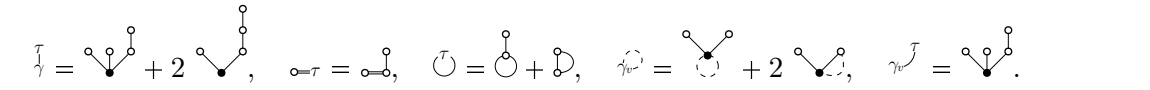
From these examples, we get the intuition that there exists a convenient method to integrate by parts exotic aromatic forests with graph operations such as unplugging/plugging edges and lianas. We shall make this intuition rigorous after defining a few notations in the spirit of the Butcher product on trees [65, Chap. III].

Notation 3.3.5. *Let γ be an exotic aromatic forest/vector field, τ be an exotic aromatic vector field and v a node of γ, then we define the following operators on forests.*

1.  : *sum of all exotic aromatic forests/vector fields obtained by linking the root of τ to a node of γ with a new edge in E_0*
2.  *(resp.*  *): aroma obtained by linking the root of τ to a white node (resp. a black node) with a new edge in E_S*
3. : *sum of all aromas obtained by linking the root of τ to a node of τ with a new edge in E_0*
4. : *sum of all exotic aromatic forests/vector fields obtained by linking the node v to a node of γ with a new liana*
5. : *forest obtained by linking the root of τ to the node v of γ with a new edge in E_0*

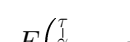
For simplicity, we combine multiple operations on a same forest as in  *and*  *, where operation 1 is always applied first.*

For example, let $\gamma =$  , $\tau =$ and $v = r$ the root of γ, then we get



The integration by parts (3.3.5) can be written conveniently with exotic aromatic forests.

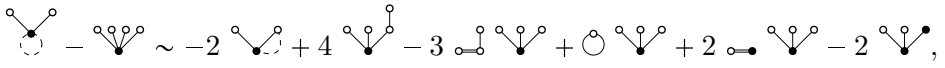
Lemma 3.3.6. *Let $\gamma \in \mathcal{EAT}$ and $\tau \in \mathcal{EAV}$, then the process of integration by parts can be written as*

$$\int_{\mathcal{M}} F\Big(\text{  }\Big)(\phi)d\mu_\infty = \int_{\mathcal{M}} F\Big(\text{} - (N_g(\gamma)+N_g(\tau)+1)\ \text{  } + (N_g(\gamma)+N_g(\tau))\ \text{} + \text{} + 2\ \text{} - 2\ \text{}\Big)(\phi)d\mu_\infty, \tag{3.3.7}$$

$$\int_{\mathcal{M}} F\Big(\text{}\Big)(\phi)d\mu_\infty = \int_{\mathcal{M}} F\Big(- (N_g(\gamma)+1)\ \text{} + N_g(\gamma)\ \text{} + \text{} + 2\ \text{} - 2\ \text{}\Big)(\phi)d\mu_\infty. \tag{3.3.8}$$

We write $\gamma \sim \widetilde{\gamma}$ if it is possible to go from $\gamma \in \mathcal{EAT}$ to $\widetilde{\gamma} \in \mathrm{Span}(\mathcal{EAT})$ with the processes of integration by parts (3.3.7) or (3.3.8), and by taking account of the simplification rules of Proposition 3.2.3. We extend this relation by linearity on $\mathrm{Span}(\mathcal{EAT})$ and make it symmetric so that $\sim$ becomes an equivalence relation on $\mathrm{Span}(\mathcal{EAT})$. For example, the integration by parts (3.3.6) can be written with exotic aromatic forests by using (3.3.8) with $\gamma =$. It yields

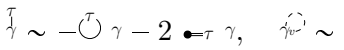
Similarly, applying (3.3.8) with $\gamma =$  gives



which represents the integration by parts

$$\int_{\mathcal{M}} \Big[\sigma^4 G^{-1}\Delta\phi''(g,g) - \sigma^4 G^{-2}\phi^{(4)}(g,g,g,g)\Big]d\mu_\infty \tag{3.3.9}$$
$$= \int_{\mathcal{M}} \Big[-2\sigma^4 G^{-1}\sum_i \phi^{(3)}(g,\partial_i g, ei) + 4\sigma^4 G^{-2}\phi^{(3)}(g,g,g'g)$$
$$- 3\sigma^4 G^{-3}(g,g'g)\phi^{(3)}(g,g,g) + \sigma^4 G^{-2}\operatorname{div}(g)\phi^{(3)}(g,g,g)$$
$$+ 2\sigma^2 G^{-2}(g,f)\phi^{(3)}(g,g,g) - 2\sigma^2 G^{-1}\phi^{(3)}(g,g,f)\Big]d\mu_\infty.$$

Remark 3.3.7. *In the Euclidean case $\mathbb{R}^d$, that is, for a forest $\gamma \in \mathcal{EAT}$ and a vector field $\tau \in \mathcal{EAV}$ with $N_g(\gamma) = N_g(\tau) = 0$ and $g = 0$, Lemma 3.3.6 reduces to the two following equations:*



We refer to [83] for more details on the integration by parts of exotic aromatic forests in $\mathbb{R}^d$, in particular in the case where f is not a gradient.

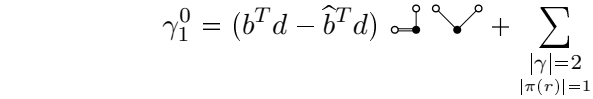
We can now revisit the statement of Theorem 2.3.2 in terms of B-series.

Theorem 3.3.8. *Consider a consistent ergodic Runge-Kutta method of the form* (3.1.1) *(respectively of the form* (3.1.2)*). We denote $\mathcal{A}_i = F(\gamma_i)$ with $\gamma_i \in \mathcal{EAT}$. If $\gamma_i \sim \gamma_i^0$ and $F(\gamma_i^0) = 0$ for $1 \leqslant i < p$, then the method has at least order p for sampling the invariant measure of* (2.1.3) *(respectively* (2.1.4)*).*

Application. *Applying the process of integration by parts to the operator $\mathcal{A}_1$ of the θ-method given in* (3.1.3) *yields $\int_{\mathbb{R}^d}\mathcal{A}_1\phi d\mu_\infty = \int_{\mathbb{R}^d}\mathcal{A}_1^0\phi d\mu_\infty$ with $\mathcal{A}_1^0 = F(\gamma^0)$ and*

In particular, the θ-method (3.1.3) *is of order two for the invariant measure if $\theta = \frac{1}{2}$.*

Remark 3.3.9. *We call $\mathcal{EAT}^0$ the subset of exotic aromatic forests whose root has only one predecessor (that is, the forests associated with an order one operator) or that have a rooted tree of the form*  ,  ,  , *... Then, if $\gamma \in \mathcal{EAT}$, there exists $\gamma^0 \in \mathcal{EAT}^0$ such that $\gamma \sim \gamma^0$. For instance, for a consistent method of the form* (3.1.2)*, the operator $\mathcal{A}_1^0 = F(\gamma_1^0)$ has the form*



so that $\gamma_1^0 \in \mathcal{EAT}^0$, and $\mathcal{A}_1^0$ is a differential operator of order one if the condition $b^T d = \widehat{b}^T d$ holds. Moreover, in the context of $\mathbb{R}^d$, one can always integrate by parts a differential operator $\mathcal{A}$ that can be represented with exotic aromatic forests into an operator $\mathcal{A}_0$ of order one.

3.3.3 Composition of differential operators with exotic aromatic forests

In this subsection, we compose differential operators with the help of exotic aromatic forests. The applications include the computation of $\mathcal{L}^p$ and of the Lie bracket $[\mathcal{L}, \overline{\mathcal{A}_p}]$ that we need for getting weak order conditions or for using Theorem 2.3.3. The composition of exotic aromatic forests is done according to the following theorem.

Theorem 3.3.10. *Let $\gamma_1 = (V^1, E^1, L^1)$ and $\gamma_2 = (V^2, E^2, L^2)$ be two exotic aromatic forests with roots r_1 and r_2, and let $\phi : \mathbb{R}^d \to \mathbb{R}$ be a smooth function. For $\varphi : \pi(r_2) \to V^1$ and $\psi : \Gamma(r_2) \to V^1$, we build $\gamma_{\varphi,\psi}$ by plugging all the edges connected to r_2 to the nodes of γ_1 according to φ, and all the lianas (counting multiplicity) according to ψ. Then the composition of forests is given by*

$$F(\gamma_2)(F(\gamma_1)(\phi)) = \sum_{\substack{\varphi:\pi(r_2)\to V^1 \\ \psi:\Gamma(r_2)\to V^1}} F(\gamma_{\varphi,\psi})(\phi).$$

Various composition rules for B-series and aromatic B-series have been studied in the literature (see [65, 31, 13] and the references therein). The main difference with these previous works is that we compose only the roots of exotic aromatic rooted forests, as this corresponds to composing linear differential operators in our context. The study of new operations with the lianas could lead to a structure of Hopf algebra of the exotic aromatic B-series (in the spirit of [31, 13]), and is matter for future work. For the sake of simplicity, we give a proof of Theorem 3.3.10 in the context of $\mathbb{R}^d$, that is, for forests such that V_g and E_S are empty. The extension to any exotic aromatic forests and hence to exotic aromatic B-series is straightforward.

Proof. Using Definition 3.2.2, we have

$$F(\gamma_1)(\phi) = \sum_{i_{v_1^{(1)}},\dots,i_{v_{m_1}^{(1)}}} \sum_{j_{l_1^{(1)}},\dots,j_{l_{s_1}^{(1)}}} \left(\prod_{v\in V_1^0} \partial_{I_{\pi(v)}} \partial_{J_{\Gamma(v)}} f_{i_v} \right) \partial_{I_{\pi(r_1)}} \partial_{J_{\Gamma(r_1)}} \phi.$$

Then we replace ϕ by $F(\gamma_1)(\phi)$ and use the Leibniz rule to distribute the partial derivatives

$$\begin{aligned}
F(\gamma_2)(F(\gamma_1)(\phi)) = {} & \sum_{\substack{i_{v_1^{(1)}},\dots,i_{v_{m_1}^{(1)}} \\ i_{v_1^{(2)}},\dots,i_{v_{m_2}^{(2)}}}} \sum_{\substack{j_{l_1^{(1)}},\dots,j_{l_{s_1}^{(1)}} \\ j_{l_1^{(2)}},\dots,j_{l_{s_2}^{(2)}}}} \left(\prod_{v\in V_2^0} \partial_{I_{\pi(v)}} \partial_{J_{\Gamma(v)}} f_{i_v} \right) \\
& \cdot \partial_{I_{\pi(r_2)}} \partial_{J_{\Gamma(r_2)}} \left[\left(\prod_{v\in V_1^0} \partial_{I_{\pi(v)}} \partial_{J_{\Gamma(v)}} f_{i_v} \right) \partial_{I_{\pi(r_1)}} \partial_{J_{\Gamma(r_1)}} \phi \right] \\
= {} & \sum_{\substack{\varphi:\pi(r_2)\to V_1 \\ \psi:\Gamma(r_2)\to V_1}} \sum_{\substack{i_{v_1^{(1)}},\dots,i_{v_{m_1}^{(1)}} \\ i_{v_1^{(2)}},\dots,i_{v_{m_2}^{(2)}}}} \sum_{\substack{j_{l_1^{(1)}},\dots,j_{l_{s_1}^{(1)}} \\ j_{l_1^{(2)}},\dots,j_{l_{s_2}^{(2)}}}} \left(\prod_{v\in V_2^0} \partial_{I_{\pi(v)}} \partial_{J_{\Gamma(v)}} f_{i_v} \right) \\
& \cdot \left(\prod_{v\in V_1^0} \partial_{I_{\pi(v)\cup\varphi^{-1}(\{v\})}} \partial_{J_{\Gamma(v)\cup\psi^{-1}(\{v\})}} f_{i_v} \right) \partial_{I_{\pi(r_1)\cup\varphi^{-1}(\{r_1\})}} \partial_{J_{\Gamma(r_1)\cup\psi^{-1}(\{r_1\})}} \phi.
\end{aligned}$$

Using Definition 3.2.2, we deduce that

$$F(\gamma_2)(F(\gamma_1)(\phi)) = \sum_{\substack{\varphi:\pi(r_2)\to V_1 \\ \psi:\Gamma(r_2)\to V_1}} F(\gamma_{\varphi,\psi})(\phi).$$

Hence the result. □

Using Theorem 3.3.10, we can calculate conveniently standard operations on differentials such as the Laplacian or the divergence. For instance, using Notation 3.3.5, for $\tau \in \mathcal{EAV}$ an exotic aromatic vector field such that V_g and E_S are empty, the divergence of τ is given by .

Application. *We recall that, in the context of* $\mathbb{R}^d$*, the generator* (2.2.2) *of equation* (2.1.3) *is given by* $\mathcal{L} = F($  $)$*. We can therefore compute* $\mathcal{L}^2\phi$*. Using Theorem 3.3.10, we obtain*

$$F(\text{•}\!\downarrow\!\text{•})(\mathcal{L}\phi) = F(\text{•}\!\downarrow\!\text{•} + \text{•}\!\diagup\!\diagdown\!\text{•} + \frac{1}{2}\,\text{•}\!\downarrow\!\circ)(\phi) \quad \textit{and} \quad F(\circ)(\mathcal{L}\phi) = F(\text{•}\!\downarrow\!\circ + \circ\!\downarrow\!\text{•} + 2\,\text{•}\!\downarrow\!\text{•} + \frac{1}{2}\,\circ\!\downarrow\!\circ)(\phi).$$

Combining these equalities, we deduce

$$\mathcal{L}^2\phi = F(\text{•}\!\downarrow\!\text{•} + \text{•}\!\diagup\!\diagdown\!\text{•} + \text{•}\!\downarrow\!\circ + \frac{1}{2}\,\circ\!\downarrow\!\text{•} + \text{•}\!\downarrow\!\text{•} + \frac{1}{4}\,\circ\!\downarrow\!\circ)(\phi).$$

In the manifold case, that is, when $\mathcal{L}$ *is given by* (2.2.3)*, the same methodology applies and the complete decomposition of* $\mathcal{L}^2$ *in exotic aromatic forests is given in Table C.1.*

CHAPTER 4

High order methods for sampling the invariant measure of ergodic SDEs

Note: This chapter is based on the articles [83] ($\mathbb{R}^d$ case) and [85] (manifold case), both in collaboration with G. Vilmart.

4.1 Introduction

We propose in this chapter a new methodology for the construction of high order integrators for sampling the invariant measure of ergodic stochastic differential equations with dynamics in $\mathbb{R}^d$, or constrained on a manifold $\mathcal{M}$. In particular, we obtain the order conditions for sampling the invariant measure for the classes of Runge-Kutta methods (3.1.1)-(3.1.2) applied to the Langevin dynamics (2.1.3) in $\mathbb{R}^d$ and to the constrained Langevin dynamics (2.1.4) on a manifold $\mathcal{M}$. The methodology is extended in $\mathbb{R}^d$ to allow the use of postprocessors, and is also applied in the context of modified SDEs, partitioned problems and perturbed dynamics. The analysis is valid for arbitrarily high order and relies on the new formalism of exotic aromatic B-series introduced in Chapter 3. We mention that the numerical results presented in this chapter can be understood independently from the formalism of Chapter 3. To illustrate the methodology, several examples of methods of order two for constrained problems are introduced, and numerical experiments on the sphere, the torus and the special linear group confirm the theoretical findings.

As seen in Chapter 2, a natural way to achieve high order p for the invariant measure is to consider a numerical scheme with high standard weak order. Several examples in the litterature (as, for instance, [14, 86, 87, 6]) show that it is possible to reach high order p for the invariant measure while keeping a low weak order (typically weak order one). In [5, 6], a methodology for the analysis and design of high order integrators for the invariant measure in $\mathbb{R}^d$ is introduced and serves as a crucial ingredient in this work. The approach combines the usual Talay-Tubaro approach [121] and recent developments of the theory of backward error analysis and modified differential equations in the stochastic context [126, 2, 48, 79, 80], a major tool in the area of deterministic geometric numerical integration [65]. In [123] for finite dimensions and in [15] in the context of parabolic SPDEs, this approach is combined with the idea of processing from

Butcher [23], to design efficient postprocessed integrators with high order for the invariant measure at a negligible overcost compared to standard low order schemes. The postprocessor methodology is extended in [1] for a class of explicit stabilized schemes of order two for the invariant measure and with optimally large stability domains.

For Langevin dynamics constrained on a manifold (2.1.4), one is interested in integrators with high order for the invariant measure that stay on the manifold (that is, $X_n \in \mathcal{M}$ for all n). A widely used and simple numerical scheme for sampling the invariant measure distribution on manifolds is the Euler scheme (see [37, 88, 90, 91] for instance). Two variants exist for the overdamped Langevin equation (2.1.4), both of order one in the weak sense, and for sampling the invariant measure: the Euler integrator with explicit projection direction

$$X_{n+1} = X_n + hf(X_n) + \sigma\sqrt{h}\xi_n + \lambda g(X_n), \quad \zeta(X_{n+1}) = 0, \tag{4.1.1}$$

and alternatively the Euler integrator with implicit projection direction

$$X_{n+1} = X_n + hf(X_n) + \sigma\sqrt{h}\xi_n + \lambda g(X_{n+1}), \quad \zeta(X_{n+1}) = 0. \tag{4.1.2}$$

To the best of our knowledge, no high order numerical integrators for sampling the invariant measure of the overdamped Langevin equation with constraints (2.1.4) and that do not reduce to a splitting method have been proposed in the literature. In [92], an order two discretization based on the RATTLE integrator (see [115, 9, 67]) is applied to the underdamped Langevin equation, rather than to the overdamped Langevin dynamic (2.1.4). The previously described discretizations can be combined with Metropolis-Hastings rejection procedures [103, 69]. We quote in particular the Markov-Chain Monte-Carlo (MCMC) methods [61, 18, 91] and the Hybrid Monte-Carlo methods [124, 92], where the need for a reverse projection check is shown to be a key step. We also mention the integrators in [125, 93] that are based on an Euler discretization and present new approaches for projecting on the manifold. The alternative approach of using Metropolis-Hastings rejection procedure allows us to fully remove the bias on the invariant measure. Analogous to the Euclidean case, this procedure does not make high order discretizations obsolete because, in particular, the rejection rate depends on the quality of the discretization and the dimension of the problem in general, and in the case of stiff problems or problems in high dimension, it suffers from timestep restrictions. Note also that in the specific case where $\mathcal{M}$ is a Lie group, high order integrators can be naturally obtained using splitting methods, that are, however, typically limited to weak order two of accuracy due to the necessity of negative time steps, that are not suitable for non-reversible problems (see [12] for further details in the context of ODEs). In this chapter, we propose the first method of order two for sampling the invariant measure of constrained Langevin dynamics (2.1.4) that does not reduce to a splitting method.

This chapter is organized as follows. In Section 4.2, we explain the methodology for the analysis of the accuracy of integrators for sampling the invariant measure on a manifold $\mathcal{M}$ by using Theorem 2.3.2. We also give a proof of this theorem in the context of compact manifolds. In Section 4.3, we apply this methodology in the context of $\mathbb{R}^d$. We build high order approximations for sampling the invariant measure of Langevin dynamics using modified equations and we derive the order conditions of order two and three of the class of Runge-Kutta methods (3.1.1). We generalize the results to include postprocessors, and adapt the methodology in the context of partitioned problems and perturbed systems. In Section 4.4, we apply the methodology in the context of manifolds to the class of Runge-Kutta methods (3.1.2) for solving the constrained overdamped Langevin equation (2.1.4), to derive arbitrary high order conditions for the invariant measure, with special emphasis on order two conditions, and

to introduce a new order two scheme that uses only a few evaluations of f per step. This new order two scheme is then compared to the Euler scheme (4.1.2) in Section 4.5 in numerical experiments on a sphere, a torus and the special linear group $\mathrm{SL}(m)$ to confirm its order of convergence for sampling the invariant measure.

4.2 A criterion for high order approximation for the invariant measure

In this section, we give a quick overview of the methodology for the analysis of the accuracy for the invariant measure of ergodic integrators using Theorem 2.3.2. Then, we prove Theorem 2.3.2 in the context of compact manifolds.

The methodology is the following. Given a consistent numerical integrator or a class of integrators such as the Runge-Kutta methods (3.1.1)-(3.1.2), we compute the Talay-Tubaro expansion (2.2.6). In the context of Langevin dynamics (2.1.3)-(2.1.4), the exotic aromatic B-series formalism proves to be a convenient tool for the computation of the operators $\mathcal{A}_j$, as explained in Subsection 3.3.1. Then, we integrate by parts multiple times the integrals $\int \mathcal{A}_j \phi d\mu_\infty$ in order to transform a high order differential operator $\mathcal{A}_j : \phi \to F(\gamma)(\phi)$ into a differential operator $\mathcal{A}_j^0$ of lower order in ϕ such that $\int \mathcal{A}_j \phi d\mu_\infty = \int \mathcal{A}_j^0 \phi d\mu_\infty$. We simplify the integration by parts process in Subsection 3.3.2 with the exotic aromatic B-series formalism. If $\mathcal{A}_j^0 = 0$ for all $j \leqslant p$, then Theorem 2.3.2 ensures that the numerical integrator has at least order p for sampling the invariant measure of (2.1.1).

Let us now prove Theorem 2.3.2. The proof relies on Talay-Tubaro expansions [121], backward error analysis and modified differential equations for SDEs [126, 2, 48, 79, 80]. It is similar to [5, Thm. 3.2] in the context of smooth compact manifolds. In the spirit of backward error analysis for differential equations (see [65, 126, 2, 48]), we build a modified generator $\mathcal{L}^h$ such that $U(x,h) = \mathbb{E}[\phi(X_1)|X_0 = x]$ formally satisfies

$$U(x,h) = \sum_{j \geqslant 0} \frac{h^j}{j!} (\mathcal{L}^h)^j \phi(x). \tag{4.2.1}$$

Truncating this formal series yields an estimate of the form

$$U(x,h) = \phi(x) + \sum_{j=1}^{N} \frac{h^j}{j!} (\mathcal{L}^h)^j \phi(x) + h^{N+1} R_N^h(\phi, x), \quad x \in N_{\mathcal{M}},$$

where $N_{\mathcal{M}}$ is an open neighbourhood of $\mathcal{M}$ in $\mathbb{R}^d$ and the remainder satisfies $\left|R_N^h(\phi,x)\right| \leqslant C_N(\phi)$. For this, we write formally $\mathcal{L}^h = \mathcal{L} + \sum_{n \geqslant 1} h^n L_n$ and compare the series expression in (2.2.6) and (4.2.1). By formally identifying the powers of h, we deduce the following rigorous definition of the L_n on an open neighbourhood of $\mathcal{M}$ in $\mathbb{R}^d$,

$$L_0 = \mathcal{L}, \quad L_n = \mathcal{A}_n + \sum_{l=1}^{n} \frac{B_l}{l!} \sum_{n_1 + \cdots + n_{l+1} = n-l} L_{n_1} \cdots L_{n_l} \mathcal{A}_{n_{l+1}}, \quad n \geqslant 1, \tag{4.2.2}$$

where the B_l are the Bernoulli numbers (see [48, 126, 5] for similar expansions in $\mathbb{T}^d$ or $\mathbb{R}^d$). Using Assumption 2.3.1, we build recursively a sequence of functions (ρ_n) such that

$$\mathcal{L}^* \rho_n = -\sum_{l=1}^{n} L_l^* \rho_{n-l} \quad \text{and} \quad \rho_0 = \rho_\infty, \tag{4.2.3}$$

where $\int_{\mathcal{M}} \rho_n d\sigma_{\mathcal{M}} = 0$ for $n \geqslant 1$. We denote $\rho_r^h = \sum_{n=0}^r h^n \rho_n$ and $d\mu_r^h = \rho_r^h d\sigma_{\mathcal{M}}$ and adapt on the manifold $\mathcal{M}$ the following result from [48, Thm. 2.1] in the context of $\mathbb{R}^d$.

Lemma 4.2.1. *Under Assumptions 2.2.2, 2.3.1 and 2.2.5, for all $\phi \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$, for every positive integer r, there exists a constant $C_r(\phi)$ independent of h such that, for all h small enough,*

$$\left| \int_{\mathcal{M}} \phi d\mu^h - \int_{\mathcal{M}} \phi d\mu_r^h \right| \leqslant C_r(\phi) h^{r+1}. \tag{4.2.4}$$

We omit the proof of Lemma 4.2.1 as it is exactly the same as in [48, Thm. 2.1] by replacing dx by $d\sigma_{\mathcal{M}}$ and $\mathbb{T}^d$ by $\mathcal{M}$. We are now able to prove Theorem 2.3.2.

Proof of Theorem 2.3.2. As $\mathcal{A}_j^* \rho_\infty = 0$ for $j = 1, \dots, r-1$, we deduce recursively from (4.2.2) and (4.2.3) that $\rho_j = 0$ for $j = 1, \dots, r-1$, which yields $\rho_r^h = \rho_\infty + h^r \rho_r$. Using the definition of the error for the invariant measure (2.3.3) and the ergodicity of the integrator (2.3.2), equation (4.2.4) becomes

$$\left| e(\phi, h) - h^r \int_{\mathcal{M}} \phi(x) \rho_r(x) d\sigma_{\mathcal{M}}(x) \right| \leqslant C h^{r+1}.$$

We are left to prove that

$$\int_{\mathcal{M}} \phi(x) \rho_r(x) d\sigma_{\mathcal{M}}(x) = \int_0^\infty \int_{\mathcal{M}} u(x,t) \mathcal{A}_r^* \rho_\infty(x) d\sigma_{\mathcal{M}}(x) dt.$$

By the backward Kolmogorov equation and ergodicity, u satisfies

$$\lim_{T \to \infty} u(x,T) = \phi(x) + \int_0^\infty \mathcal{L} u(x,t) dt = \int_{\mathcal{M}} \phi(y) d\mu_\infty(y).$$

Using $\mathcal{L}^* \rho_r = -L_r^* \rho_\infty = -\mathcal{A}_r^* \rho_\infty$, we deduce

$$\begin{aligned}
\int_{\mathcal{M}} \phi(x) \rho_r(x) d\sigma_{\mathcal{M}}(x) &= -\int_0^\infty \int_{\mathcal{M}} \mathcal{L} u(x,t) \rho_r(x) d\sigma_{\mathcal{M}}(x) dt + \int_{\mathcal{M}} \phi(y) d\mu_\infty(y) \int_{\mathcal{M}} \rho_r(x) d\sigma_{\mathcal{M}}(x) \\
&= -\int_0^\infty \int_{\mathcal{M}} u(x,t) \mathcal{L}^* \rho_r(x) d\sigma_{\mathcal{M}}(x) dt \\
&= \int_0^\infty \int_{\mathcal{M}} u(x,t) \mathcal{A}_r^* \rho_\infty(x) d\sigma_{\mathcal{M}}(x) dt,
\end{aligned}$$

where we used that $\int_{\mathcal{M}} \rho_r(x) d\sigma_{\mathcal{M}}(x) = 0$. This concludes the proof of Theorem 2.3.2. □

Remark 4.2.2. *One can consider possible generalisations of Theorem 2.3.2 in the case where the manifold $\mathcal{M}$ is not compact, or if $\mathcal{M}$ is a manifold of any dimension. We refer to [5] for the non-compact extension of Theorem 2.3.2 in the context of $\mathbb{R}^d$.*

4.3 High order sampling integrators in $\mathbb{R}^d$

In this section, we apply the formalism of B-series presented in Chapter 3 to obtain high order integrators for sampling the invariant measure of Langevin dynamics (2.1.3) in the context of $\mathbb{R}^d$.

4.3.1 Improvement of a method order via a modified equation

In [5], a recursive method to obtain integrators of any order for the convergence to the invariant measure is presented. Let us suppose we have an integrator of order exactly $p \geqslant 1$ for the invariant measure, that is, for all $j < p$, $\mathcal{A}_j^* \rho_\infty = 0$ and $\mathcal{A}_p^* \rho_\infty \neq 0$. By integrating by parts, we can write $\int_{\mathbb{R}^d} \mathcal{A}_p \phi \rho_\infty dx$ as $\int_{\mathbb{R}^d} \phi' f_p \rho_\infty dx$. We then consider the same numerical integrator but for the modified equation where we replaced f by $f - h^p f_p$. Applying Theorem 2.3.2 to the new context, we see that this integrator is at least of order $p+1$ for the original equation.

In this section, we give tools to simplify the computation of those modified integrators, in particular to calculate simply the operators $\mathcal{A}_j$, and to find the function f_p.

Example. *For the θ-method* (3.1.3)*, we have $\mathcal{A}_1 = F(\gamma)$ where γ is given by* (3.3.3)*. Applying integration by parts as described in Section 3.3.2, we obtain*

$$\gamma \sim (\frac{1}{2} - \theta)\Big(\quad + \quad \Big).$$

Then f_1 is given by

$$f_1 = \left(\frac{1}{2} - \theta\right)(f'f + \frac{\sigma^2}{2}\Delta f).$$

Application. *Let us calculate the modified vector field f_2 for the θ-method* (3.1.3)*. First we write the differential operator $\mathcal{A}_2$ for the modified equation where we replaced f by $f - hf_1$. We call it $\mathcal{A}_2^{(1)}$. We find $\mathcal{A}_2^{(1)} = F(\gamma)$ where*

$$\gamma = \theta(3\theta-1) \quad + \frac{\theta(3\theta-1)}{2} \quad + \frac{\theta(4\theta-1)}{2} \quad + \theta^2 \quad + \frac{\theta^2}{4} \quad + \theta^2 \quad + \frac{\theta(5\theta-1)}{2} \quad$$
$$+ \frac{\theta(2\theta+1)}{2} \quad + \frac{\theta(2\theta+1)}{4} \quad + \frac{4\theta-1}{2} \quad + \frac{(4\theta-1)}{4} \quad + \frac{(4\theta-1)}{4} \quad + \frac{(4\theta-1)}{8} \quad$$
$$+ \frac{\theta}{2} \quad + \theta \quad + \frac{1}{6} \quad + \frac{\theta}{2} \quad + \frac{1}{4} \quad + \frac{1}{8} \quad + \frac{1}{48} \quad .$$

Applying integration by parts, we find

$$\gamma \sim \left(-2\theta^2 + 2\theta - \frac{1}{2}\right) \quad + \left(-\theta^2 + \theta - \frac{1}{4}\right) \quad + \left(-\frac{3\theta^2}{2} + \frac{3\theta}{2} - \frac{1}{3}\right) \quad$$
$$+ \left(-\theta^2 + \theta - \frac{1}{6}\right) \quad + \left(-\frac{\theta^2}{4} + \frac{\theta}{4} - \frac{1}{24}\right) \quad + \left(-\theta^2 + \theta - \frac{1}{6}\right) \quad .$$

Thus we define

$$f_2 = \left(-2\theta^2 + 2\theta - \frac{1}{2}\right) f'f'f + \left(-\theta^2 + \theta - \frac{1}{4}\right)\sigma^2 f'\Delta f$$
$$+ \left(-\frac{3\theta^2}{2} + \frac{3\theta}{2} - \frac{1}{3}\right)\sigma^2 \sum_i f''(e_i, f'(e_i)) + \left(-\theta^2 + \theta - \frac{1}{6}\right) f''(f,f)$$
$$+ \left(-\frac{\theta^2}{4} + \frac{\theta}{4} - \frac{1}{24}\right)\sigma^4 \Delta^2 f + \left(-\theta^2 + \theta - \frac{1}{6}\right)\sigma^2 (\Delta f)'(f),$$

and, if the θ-scheme applied to $dX = (f - hf_1 - h^2 f_2)dt + \sigma dW$ is ergodic, then it has order 3 for the invariant measure.

For f_1, we recover the formula of [5, Prop. 5.1]. The computation of f_2 was first done for $\theta = 0$ in [5, Prop. 5.2], which reveals a typographical error. Note that for f linear and $\theta = \frac{1}{2}$, we get $f_2 = 0$, as the θ-method with $\theta = \frac{1}{2}$ samples the invariant measure exactly in the Gaussian case [3].

This method can give numerical integrators of any order, but it comes with a high computing cost if the partial derivatives of f are difficult to compute. In the following sections, we present order conditions for certain classes of numerical schemes, in order to obtain high order methods avoiding derivatives and unnecessary evaluations of f.

4.3.2 Order conditions for stochastic Runge-Kutta schemes

We consider stochastic Runge-Kutta schemes (3.1.1) for solving the overdamped Langevin equation (2.1.3). Using the proposed framework, our goal is to find algebraic conditions on the coefficients $A = (a_{ij})$, $b = (b_i)$ and $d = (d_i)$ to achieve a given order condition for the invariant measure.

First, we suppose $\sum b_i = 1$ in order for $\mathcal{A}_0 = \mathcal{L}$ in Assumption 2.2.5 to be satisfied. Then, we have $\mathcal{A}_1\phi = F(\gamma_1)(\phi)$ where

$$\gamma_1 = \sum b_i c_i$$  $$+ \frac{1}{2}\sum b_i d_i^2$$  $$+ \sum b_i d_i$$  $$+ \frac{1}{2}$$  $$+ \frac{1}{2}$$  $$+ \frac{1}{8}$$  .

Lemma 3.3.6 yields

$$\gamma_1 \sim \left(\sum b_i c_i + \frac{1}{2} - 2\sum b_i d_i\right)$$  $$+ \frac{1}{2}\left(\sum b_i d_i^2 + \frac{1}{2} - 2\sum b_i d_i\right)$$  .

Thus if the method satisfies

$$\sum b_i = 1, \quad \sum b_i c_i + \frac{1}{2} - 2\sum b_i d_i = 0, \quad \sum b_i d_i^2 + \frac{1}{2} - 2\sum b_i d_i = 0,$$

then $\gamma_1 \sim 0$ and, according to Theorem 3.3.8, we get a Runge-Kutta scheme of order two for sampling the invariant measure of (2.1.3).

By continuing this approach, we obtain the order conditions of order three, and our analysis allows us to obtain the conditions for any order. The following result states the order conditions for Runge-Kutta methods.

Theorem 4.3.1. *Consider an ergodic Runge-Kutta method of the form* (3.1.1) *with* $\sum b_i = 1$ *for solving* (2.1.3). *If* $\mathcal{A}_i = F(\gamma_i)$ *and* $\gamma_i \sim \gamma_i^0$, *where* $F(\gamma_i^0)$ *is an operator of order one, if* A, b *and* d *are chosen such that* $\gamma_i^0 = 0$ *for all* $1 \leqslant i < p$, *then the method has at least order* p *for the invariant measure. In particular, Table 4.1 gives sufficient conditions to have consistency and order two or three for the invariant measure for the Runge-Kutta schemes* (3.1.1).

Remark 4.3.2. *As presented in Theorem 3.3.3, the exotic aromatic B-series can also be used to derive weak order conditions. We collect in Table 4.2 the weak order conditions of the class of Runge-Kutta schemes* (3.1.1) *up to order* $p \leqslant 3$. *We recover exactly the same order conditions as first derived in [45, Thm. 4] using different types of trees and B-series, and in the specific case where* f *is a gradient. We recall that the conditions for weak order 3 have no solution for a method of the form* (3.1.1) *with only* $l = 1$ *noise. Indeed, fixing* $d^{(2)} = 0$ *in Table 4.2, we obtain the incompatible order conditions* $\sum b_i \big(d_i^{(1)}\big)^2 = \frac{1}{2}$ *(third line of Table 4.2) and* $\sum b_i \big(d_i^{(1)}\big)^2 = \frac{1}{3}$ *(last line of Table 4.2). Taking* $l = 2$ *noises in the method* (3.1.1) *is sufficient for reaching weak order 3.*

Order	Tree τ	$F(\tau)(\phi)$	Order condition
1		$\phi' f$	$\sum b_i = 1$
2		$\phi' f' f$	$\sum b_i c_i - 2\sum b_i d_i = -\frac{1}{2}$
		$\sigma^2 \phi' \Delta f$	$\sum b_i d_i^2 - 2\sum b_i d_i = -\frac{1}{2}$
3		$\phi' f' f' f$	$\sum b_i a_{ij} c_j - 2\sum b_i a_{ij} d_j + \sum b_i c_i - \left(\sum b_i d_i\right)^2 = 0$
		$\sigma^2 \phi' f' \Delta f$	$\sum b_i a_{ij} d_j^2 - 2\sum b_i a_{ij} d_j + \sum b_i c_i - \left(\sum b_i d_i\right)^2 = 0$
		$\phi' f''(f,f)$	$\frac{1}{2}\sum b_i c_i^2 - 2\sum b_i d_i c_i$ $-2\sum b_i d_i + 2\sum b_i d_i^2 + \sum b_i c_i = -\frac{1}{3}$
		$\sigma^2 \sum \phi' f''(f'(e_i), e_i)$	$\sum b_i d_i a_{ij} d_j - \sum b_i c_i d_i - \sum b_i d_i + \sum b_i d_i^2$ $+\sum b_i c_i - \sum b_i a_{ij} d_j - \frac{1}{2}\left(\sum b_i d_i\right)^2 = -\frac{1}{6}$
		$\sigma^2 \phi'((\Delta f)'(f))$	$\frac{1}{2}\sum b_i c_i d_i^2 - \sum b_i d_i^3 - 2\sum b_i d_i$ $+\frac{5}{2}\sum b_i d_i^2 - \sum b_i c_i d_i + \frac{1}{2}\sum b_i c_i = -\frac{1}{3}$
		$\sigma^4 \phi' \Delta^2 f$	$\frac{1}{8}\sum b_i d_i^4 - \frac{1}{2}\sum b_i d_i^3$ $-\frac{1}{2}\sum b_i d_i + \frac{3}{4}\sum b_i d_i^2 = -\frac{1}{12}$

Table 4.1: Runge-Kutta order conditions for the invariant measure (See Theorem 4.3.1). The sums are over all involved indices.

Remark 4.3.3. *As explained in the introduction, the study of weak order conditions in finite time using rooted trees is already well documented in the literature, but the framework of exotic aromatic B-series has the advantage to involve rooted forests that do not depend on the dimension d, which permits us to compute integration by parts and hence derive the order conditions for the invariant measure. Since weak convergence implies convergence with at least the same order for the invariant measure, the weak order conditions (Table 4.2) imply the order conditions for the invariant measure (Table 4.1). In particular, comparing Table 4.1 and Table 4.2, we observe that there is a strictly lower number of order conditions for the convergence to the invariant measure compared to the standard weak convergence.*

4.3.3 Order conditions for postprocessed integrators

In this section, we extend our analysis to the case of integrators combined with postprocessors [123]. As stated in Theorem 2.3.3, it permits us to increase the order for the invariant measure of a given method while maintaining a low number of function evaluations per time step. Applying Theorem 2.3.3 and Theorem 3.3.10 (for the computation of the operator $\mathcal{A}_p + [\mathcal{L}, \overline{\mathcal{A}_p}]$), we obtain the order conditions on postprocessors to increase by one the order of a given method.

Order	Tree τ	$F(\tau)(\phi)$	Order condition
1		$\phi' f$	$\sum b_i = 1$
2		$\phi' f' f$	$\sum b_i c_i = \frac{1}{2}$
		$\sigma^2 \phi' \Delta f$	$\sum b_i \big(d_i^{(1)}\big)^2 + \sum b_i \big(d_i^{(2)}\big)^2 = \frac{1}{2}$
		$\sigma^2 \sum \phi''(e_i, f'(e_i))$	$\sum b_i d_i^{(1)} = \frac{1}{2}$
3		$\phi' f' f' f$	$\sum b_i a_{ij} c_j = \frac{1}{6}$
		$\sigma^2 \phi' f' \Delta f$	$\sum b_i a_{ij} \big(d_j^{(1)}\big)^2 + \sum b_i a_{ij} \big(d_j^{(2)}\big)^2 = \frac{1}{6}$
		$\phi' f''(f, f)$	$\sum b_i c_i^2 = \frac{1}{3}$
		$\sigma^2 \sum \phi' f''(f'(e_i), e_i)$	$\sum b_i d_i^{(1)} a_{ij} d_j^{(1)} + \sum b_i d_i^{(2)} a_{ij} d_j^{(2)} = \frac{1}{6}$
		$\sigma^2 \phi'((\Delta f)'(f))$	$\sum b_i c_i \big(d_i^{(1)}\big)^2 + \sum b_i c_i \big(d_i^{(2)}\big)^2 = \frac{1}{3}$
		$\sigma^4 \phi' \Delta^2 f$	$\sum b_i \Big(\big(d_i^{(1)}\big)^2 + \big(d_i^{(2)}\big)^2\Big)^2 = \frac{1}{3}$
		$\sigma^2 \sum \phi''(e_i, f' f'(e_i))$	$\sum b_i a_{ij} d_j^{(1)} + \frac{1}{2}\Big(\sum b_i d_i^{(2)}\Big)^2 = \frac{5}{24}$
		$\sigma^2 \sum \phi''(e_i, f''(e_i, f))$	$\sum b_i c_i d_i^{(1)} = \frac{1}{3}$
		$\sigma^2 \sum \phi''(e_i, (\Delta f)'(e_i))$	$\sum b_i \big(d_i^{(1)}\big)^3 + \sum b_i d_i^{(1)} \big(d_i^{(2)}\big)^2 = \frac{1}{3}$
		$\sigma^2 \sum \phi'''(e_i, e_j, f''(e_i, e_j))$	$\sum b_i \big(d_i^{(1)}\big)^2 = \frac{1}{3}$

Table 4.2: Runge-Kutta standard weak order conditions for $l = 2$ noises. The sums are over all involved indices. We recover the same conditions as in [45].

Theorem 4.3.4. *Consider an ergodic Runge-Kutta method of the form* (3.1.1) *and of order* $p \geqslant 1$ *for sampling the invariant measure of* (2.1.3)*, and the following associated postprocessor*

$$\overline{Y_i} = X_n + h \sum_{j=1}^{s} \overline{a_{ij}} f(\overline{Y_j}) + \overline{d_i} \sigma \sqrt{h}\, \overline{\xi_n}, \qquad i = 1, \dots, s,$$
$$\overline{X_n} = X_n + h \sum_{i=1}^{s} \overline{b_i} f(\overline{Y_i}) + \overline{d_0} \sigma \sqrt{h}\, \overline{\xi_n}.$$

If the postprocessor satisfies an expansion of the form (2.3.4)*, if* γ *is the exotic aromatic B-series such that* $F(\gamma) = (\mathcal{A}_p + [\mathcal{L}, \overline{\mathcal{A}_p}])$ *and if* A*,* b*,* d*,* $\overline{A}$*,* $\overline{b}$*,* $\overline{d}$*,* $\overline{d_0}$ *are chosen such that* $\gamma \sim 0$*, then the postprocessed method* $\overline{X_n}$ *has at least order* $p + 1$ *for the invariant measure. In particular, if the conditions of order two in Table 4.3 are verified, then the postprocessed integrator has*

order two. If the Runge-Kutta method has order two for the invariant measure (see Table 4.1), if $\sum \overline{b_i} = \overline{d_0}^2$ *and if the conditions of order three in Table 4.3 are verified, then the method has order three.*

Order	Tree τ	Order condition
2		$\sum b_i c_i - 2\sum b_i d_i - 2\sum \overline{b_i} + 2\overline{d_0}^2 = -\frac{1}{2}$
		$\sum b_i d_i^2 - 2\sum b_i d_i - \sum \overline{b_i} + \overline{d_0}^2 = -\frac{1}{2}$
3		$\sum b_i a_{ij} c_j - 2\sum b_i a_{ij} d_j + \sum b_i c_i$ $-(\sum b_i d_i)^2 - 2\sum \overline{b_i}\overline{c_i} + 4\overline{d_0}\sum \overline{b_i d_i} - \overline{d_0}^4 = 0$
		$\sum b_i a_{ij} d_j^2 - 2\sum b_i a_{ij} d_j + \sum b_i c_i$ $-(\sum b_i d_i)^2 - \sum \overline{b_i}\overline{c_i} - \sum \overline{b_i d_i}^2 + 4\overline{d_0}\sum \overline{b_i d_i} - \overline{d_0}^4 = 0$
		$\frac{1}{2}\sum b_i c_i^2 - 2\sum b_i d_i c_i - 2\sum b_i d_i + 2\sum b_i d_i^2$ $+\sum b_i c_i - \sum \overline{b_i}\overline{c_i} + 2\overline{d_0}\sum \overline{b_i d_i} - \frac{\overline{d_0}^4}{2} = -\frac{1}{3}$
		$\sum b_i d_i a_{ij} d_j - \sum b_i c_i d_i - \sum b_i d_i + \sum b_i d_i^2 + \sum b_i c_i$ $-\sum b_i a_{ij} d_j - \frac{1}{2}(\sum b_i d_i)^2 - \sum \overline{b_i}\overline{c_i} + 2\overline{d_0}\sum \overline{b_i d_i} - \frac{\overline{d_0}^4}{2} = -\frac{1}{6}$
		$\frac{1}{2}\sum b_i c_i d_i^2 - \sum b_i d_i^3 - 2\sum b_i d_i + \frac{5}{2}\sum b_i d_i^2 - \sum b_i c_i d_i$ $+\frac{1}{2}\sum b_i c_i - \frac{1}{2}\sum \overline{b_i}\overline{c_i} - \frac{1}{2}\sum \overline{b_i d_i}^2 + 2\overline{d_0}\sum \overline{b_i d_i} - \frac{\overline{d_0}^4}{2} = -\frac{1}{3}$
		$\frac{1}{8}\sum b_i d_i^4 - \frac{1}{2}\sum b_i d_i^3 - \frac{1}{2}\sum b_i d_i$ $+\frac{3}{4}\sum b_i d_i^2 - \frac{1}{4}\sum \overline{b_i d_i}^2 + \frac{1}{2}\overline{d_0}\sum \overline{b_i d_i} - \frac{\overline{d_0}^4}{8} = -\frac{1}{12}$

Table 4.3: Order conditions for Runge-Kutta method with Runge-Kutta postprocessor (See Theorem 4.3.4). The sums are over all involved indices.

Example. *The following Runge-Kutta method, introduced in [123], is of order 2 for the invariant measure of* (2.1.3) *(if it is ergodic):*

$$\begin{aligned} X_{n+1} &= X_n + hf(X_{n+1} + \tfrac{-1+\sqrt{2}}{2}\sigma\sqrt{h}\xi_n) + \sigma\sqrt{h}\xi_n, \\ \overline{X_n} &= X_n + h\tfrac{\sqrt{2}}{2}f(\overline{X_n}) + \tfrac{\sqrt{4\sqrt{2}-1}}{2}\sigma\sqrt{h}\,\overline{\xi_n}. \end{aligned}$$

Indeed, its coefficients, placed in the following Butcher tableau, fulfill the conditions of order 2 of Theorem 4.3.4 (See Table 4.3):

$$\begin{array}{c|c|c|c|c|c} c & A & d & \overline{c} & \overline{A} & \overline{d} \\ \hline & b & & & \overline{b} & \overline{d}_0 \end{array} = \begin{array}{c|c|c|c|c} 1 & 1 & \frac{1+\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{4\sqrt{2}-1}}{2} \\ \hline & 1 & & & \frac{\sqrt{2}}{2} & \frac{\sqrt{4\sqrt{2}-1}}{2} \end{array}.$$

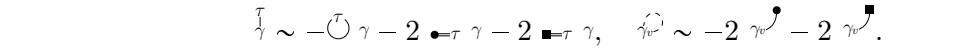
4.3.4 Order conditions for partitioned methods

We now assume that $f = f_1 + f_2$ and we consider partitioned integrators that apply different numerical treatments to each f_i. We explain in this section how to extend the exotic aromatic

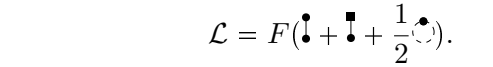
B-series formalism to compute order conditions for such partitioned integrators. The advantage is to treat each part of f differently according to their properties. For example, if f_1 is stiff and f_2 is non-stiff, one would like to apply an implicit method to f_1 and an explicit method to f_2 (IMEX methods).

We follow the formalism of [65, Sect. III.2] for bicoloured B-series, called P-series. We introduce square nodes which represent the function f_2. Circle nodes now correspond to f_1 but the root still corresponds to ϕ. We call these new forests exotic aromatic P-forests. There are two slight changes in the computation rules compared to the non-partitioned case:

- The integration by parts presented in Lemma 3.3.6 can be written as



- The operator $\mathcal{L}$ is now written as



In addition to the partitioning of the method, one can also add a postprocessor. The results of Section 4.3.2 and Section 4.3.3 are straightforwardly adapted to the P-forests.

Theorem 4.3.5. *Consider a Runge-Kutta method of order p for sampling the invariant measure of* (2.1.3) *of the form*

$$\begin{aligned} Y_i &= X_n + h\sum_{j=1}^{s} a_{ij} f_1(Y_j) + \widehat{a_{ij}} f_2(Y_j) + d_i\sigma\sqrt{h}\xi_n, \qquad i = 1,\dots,s,\\ X_{n+1} &= X_n + h\sum_{i=1}^{s} b_i f_1(Y_i) + \widehat{b_i} f_2(Y_i) + \sigma\sqrt{h}\xi_n, \end{aligned}$$

together with the following Runge-Kutta postprocessor

$$\begin{aligned} \overline{Y_i} &= X_n + h\sum_{j=1}^{s} \overline{a_{ij}} f_1(\overline{Y_j}) + \widehat{\overline{a_{ij}}} f_2(\overline{Y_j}) + \overline{d_i}\sigma\sqrt{h}\,\overline{\xi_n}, \qquad i = 1,\dots,s,\\ \overline{X_n} &= X_n + h\sum_{i=1}^{s} \overline{b_i} f_1(\overline{Y_i}) + \widehat{\overline{b_i}} f_2(\overline{Y_i}) + \overline{d_0}\sigma\sqrt{h}\,\overline{\xi_n}. \end{aligned}$$

If the postprocessor satisfies an expansion of the form (2.3.4)*, if γ is the exotic aromatic B-series such that $F(\gamma) = (\mathcal{A}_p + [\mathcal{L}, \overline{\mathcal{A}_p}])$ and if the coefficients of the method are chosen such that $\gamma \sim 0$, then the postprocessed method $\overline{X_n}$ has at least order $p+1$ for the invariant measure. In particular, the conditions for consistency and order 2 are in Table 4.4.*

Example. *Using the previously introduced formalism, we see that the following method, adapted from [15, Lemma 2.9], is of order 2 for sampling the invariant measure of* (2.1.3) *(if it is ergodic):*

$$\begin{aligned} X_{n+1} &= X_n + \tfrac{h}{2} f_1(X_{n+1} + \tfrac{1}{2}\sigma\sqrt{h}\xi_n) + \tfrac{h}{2} f_1(X_{n+1} + \tfrac{3}{2}\sigma\sqrt{h}\xi_n)\\ &\quad + h f_2(X_n + \tfrac{1}{2}\sigma\sqrt{h}\xi_n) + \sigma\sqrt{h}\xi_n,\\ \overline{X_n} &= X_n + \tfrac{1}{2}\sigma\sqrt{h}\,\overline{\xi_n}. \end{aligned}$$

It can be put in Runge-Kutta form with the coefficients below:

$$\begin{array}{c|c|c|c|c} c & A & \widehat{c} & \widehat{A} & d \\ \hline & b & & \widehat{b} & \end{array} = \begin{array}{c|ccc|c|ccc|c} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 \\ 1 & 0 & 1/2 & 1/2 & 1 & 1 & 0 & 0 & 1/2 \\ 1 & 0 & 1/2 & 1/2 & 1 & 1 & 0 & 0 & 3/2 \\ \hline & 0 & 1/2 & 1/2 & & 1 & 0 & 0 & \end{array}$$

with the parameters $s = 0$ and $\overline{d_0} = \frac{1}{2}$ for the postprocessor.

Order	Tree τ	$F(\tau)(\phi)$	Order condition
1		$\phi' f_1$	$\sum b_i = 1$
		$\phi' f_2$	$\sum \widehat{b}_i = 1$
2		$\phi' f_1' f_1$	$\sum b_i c_i - 2\sum b_i d_i - 2\sum \overline{b_i} + 2\overline{d_0}^2 = -\frac{1}{2}$
		$\phi' f_1' f_2$	$\sum b_i \widehat{c}_i - 2\sum b_i d_i - \sum \overline{b_i} - \sum \widehat{\overline{b_i}} + 2\overline{d_0}^2 = -\frac{1}{2}$
		$\phi' f_2' f_1$	$\sum \widehat{b}_i c_i - 2\sum \widehat{b}_i d_i - \sum \overline{b_i} - \sum \widehat{\overline{b_i}} + 2\overline{d_0}^2 = -\frac{1}{2}$
		$\phi' f_2' f_2$	$\sum \widehat{b}_i \widehat{c}_i - 2\sum \widehat{b}_i d_i - 2\sum \widehat{\overline{b_i}} + 2\overline{d_0}^2 = -\frac{1}{2}$
		$\sigma^2 \phi' \Delta f_1$	$\sum b_i d_i^2 - 2\sum b_i d_i - \sum \overline{b_i} + \overline{d_0}^2 = -\frac{1}{2}$
		$\sigma^2 \phi' \Delta f_2$	$\sum \widehat{b}_i d_i^2 - 2\sum \widehat{b}_i d_i - \sum \widehat{\overline{b_i}} + \overline{d_0}^2 = -\frac{1}{2}$

Table 4.4: Order conditions for partitioned Runge-Kutta method with postprocessor (See Theorem 4.3.5). The sums are over all involved indices.

If we add a family of independent noises $(\chi_n)_n$ that are also independent of the $(\xi_n)_n$, then by extending Theorem 4.3.5, we can show that the following IMEX method has order 2 for the invariant measure of (2.1.3) *(if it is ergodic):*

$$\begin{aligned} X_{n+1} &= X_n + h f_1(X_{n+1} + \tfrac{1}{2}\sigma\sqrt{h}\chi_n) + h f_2(X_n + \tfrac{1}{2}\sigma\sqrt{h}\xi_n) + \sigma\sqrt{h}\xi_n, \\ \overline{X_n} &= X_n + \tfrac{1}{2}\sigma\sqrt{h}\,\overline{\xi_n}. \end{aligned}$$

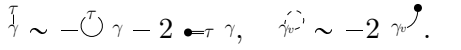
4.3.5 Non-reversible perturbation

An interesting modification of (2.1.3) is to introduce a non gradient perturbation that preserves the invariant measure. It permits for some classes of problems to improve the rate of convergence to equilibrium [89], and it can also reduce the variance [51]. As in Section 4.3.4, we write $f = f_1 + f_2$ and we use bicoloured forests. We suppose $f_1 = -\nabla V$, and f_2 is a perturbation of f_1 that satisfies

$$\operatorname{div}\left(f_2 e^{-\frac{2}{\sigma^2}V}\right) = 0. \tag{4.3.1}$$

The perturbation f_2 does not modify the invariant measure. Indeed equation (4.3.1) implies that the adjoint of $\mathcal{B}\phi = \phi'(f_2)$ satisfies $\mathcal{B}^*\rho_\infty = 0$, and thus the invariant measure is preserved. A simple example of such non gradient perturbations is $f_2 = J\nabla V$, with J a fixed antisymmetric matrix. All the results of Section 3.3 carry over to this setting with the following modifications:

- The integration by parts presented in Lemma 3.3.6 can be written as



- The generator reads $\mathcal{L} = F($$)$.

- We have $F($$) = -2F($$)$, and these differentials vanish if $f_2 = J\nabla V$.

With similar calculations as we did in Section 4.3.4, we obtain the following result.

Theorem 4.3.6. *Consider an ergodic Runge-Kutta method and a postprocessor as in Theorem 4.3.5, where $f_1 = -\nabla V$ and f_2 satisfies* (4.3.1)*. If the postprocessor satisfies an expansion of the form* (2.3.4)*, if γ is the exotic aromatic B-series such that $F(\gamma) = (\mathcal{A}_p + [\mathcal{L}, \overline{\mathcal{A}_p}])$ and if the coefficients of the method are chosen such that $\gamma \sim 0$, then the postprocessed method $\overline{X_n}$ has at least order $p+1$ for the invariant measure of* (2.1.3)*. In particular, the conditions for order 1 and order 2 are in Table 4.5.*

Order	Tree τ	$F(\tau)(\phi)$	Order condition
1		$\phi' f_1$	$\sum b_i = 1$
2		$\phi' f_1' f_1$	$\sum b_i c_i - 2\sum b_i d_i - 2\sum \overline{b_i} + 2\overline{d_0}^2 = -\frac{1}{2}$
		$\phi' f_1' f_2$	$\sum b_i \widehat{c_i} + \sum \overline{b_i} - \sum \widehat{\overline{b_i}} = 0$
		$\phi' f_2' f_1$	$\sum \widehat{b_i} c_i - 2\sum \widehat{b_i} d_i + \sum \widehat{b_i} - \sum \overline{b_i} - \sum \widehat{\overline{b_i}} + 2\overline{d_0}^2 = 0$
		$\phi' f_2' f_2$	$\sum \widehat{b_i}\widehat{c_i} - \frac{1}{2}\left(\sum \widehat{b_i}\right)^2 = 0$
		$\sigma^2 \phi' \Delta f_1$	$\sum b_i d_i^2 - 2\sum b_i d_i - \sum \overline{b_i} + \overline{d_0}^2 = -\frac{1}{2}$
		$\sigma^2 \phi' \Delta f_2$	$\sum \widehat{b_i} d_i^2 - 2\sum \widehat{b_i} d_i + \sum \widehat{b_i} - \sum \widehat{\overline{b_i}} + \overline{d_0}^2 = 0$

Table 4.5: Order conditions for partitioned Runge-Kutta method with postprocessor for the perturbed equation (See Theorem 4.3.6). The sums are over all involved indices.

Note that if $f_2 = J\nabla V$, we have  ~ -2  $- 2$  . Thus, if $f_2 = J\nabla V$, the order condition for  can be omitted and the two conditions of  and  are respectively replaced by

$$2\sum \widehat{b_i} d_i - \sum \widehat{b_i} + \sum b_i \widehat{c_i} - \sum \widehat{b_i} d_i^2 + \sum \overline{b_i} - \overline{d_0}^2 = 0,$$
$$\sum \widehat{b_i} c_i - \sum \widehat{b_i} d_i^2 - \sum \overline{b_i} + \overline{d_0}^2 = 0.$$

In order for the method to satisfy $\mathcal{A}_0 = \mathcal{L}$, the condition $\sum \widehat{b_i} = 1$ should be added in Table 4.5, but it is not necessary to achieve order 1 for the invariant measure.

Example. *If f_1 is a gradient and f_2 satisfies* (4.3.1)*, the following consistent postprocessed scheme has order 2 for the invariant measure (if it is ergodic):*

$$\begin{aligned} X_{n+1} &= X_n + hf_1(X_n + \tfrac{\sigma}{2}\sqrt{h}\xi_n) + \tfrac{5}{4}hf_2(X_n + \tfrac{\sigma}{2}\sqrt{h}\xi_n) \\ &\quad -\tfrac{1}{4}hf_2(X_n - 2hf_2(X_n + \tfrac{\sigma}{2}\sqrt{h}\xi_n) - \tfrac{\sigma}{2}\sqrt{h}\xi_n) + \sigma\sqrt{h}\xi_n, \\ \overline{X_n} &= X_n + \tfrac{\sigma}{2}\sqrt{h}\,\overline{\xi_n}. \end{aligned} \tag{4.3.2}$$

If $f_2 = J\nabla V$, it needs two evaluations of ∇V per timestep similarly to a standard Runge-Kutta weak order 2 method. For $f_2 = 0$, the scheme (4.3.2) *coincides with the one proposed in [86], formulated in a different manner (See [123]).*

Remark 4.3.7. *In the recent work [4], the convergence to equilibrium is improved and the variance is reduced by applying a reversible stochastic perturbation that does not modify the invariant measure, instead of a deterministic perturbation. The B-series formalism presented in this thesis allows us to study and build high order integrators for these kind of problems. This is matter for future work.*

4.4 High order sampling integrators on manifolds

In this section, we propose a new class of Runge-Kutta methods for sampling the invariant measure of equation (2.1.4), and present the methodology for deriving the conditions of any order for the invariant measure using Theorem 2.3.2. In particular, we compute exactly the consistency and order two conditions for the invariant measure as they are the most relevant for applications.

4.4.1 Runge-Kutta methods for constrained overdamped Langevin

When discretizing naively equation (2.1.4), one cannot ensure in general that the integrator stays on $\mathcal{M}$. It is natural to discretize instead the equivalent formulation with Lagrange multipliers

$$dX(t) = f(X(t))dt + \sigma dW(t) + g(X(t))d\lambda_t, \quad \zeta(X) = 0, \quad X(0) = X_0 \in \mathcal{M}.$$

The class of numerical schemes we obtain is in the spirit of deterministic Runge-Kutta methods for differential algebraic problems such as the methods SHAKE and RATTLE (see [115, 9, 67]), introduced in the context of constrained Hamiltonian dynamics, or the SPARK class of methods for general DAEs (see [74]). Since evaluating f is in practical applications the most expensive part of the algorithm compared to evaluating g, we propose high order integrators that are implicit in g and explicit in f in the spirit of implicit-explicit (IMEX) integrators (see, e.g., [67]), so that there are only a few evaluations of f per step. We thus consider the class of Runge-Kutta integrators (3.1.2). Ideally, one aims for IMEX integrators with a low number of evaluations of f, we hence assume in addition that $\widehat{A}$ is a lower triangular matrix and A is a strictly lower triangular matrix (in the spirit of DIRK methods). We represent the numerical integrators with their associated Butcher tableau, where $b = (a_{s,i})_i$, $\widehat{b} = (\widehat{a}_{s,i})_i$, $c = A\mathbb{1}$ and $\mathbb{1} = (1, \dots, 1)^T$,

$$\begin{array}{c|c||c|c||c} c & A & \delta & \widehat{A} & d \\ \hline & b^T & & \widehat{b}^T & \end{array}.$$

For instance, the Euler schemes can be written as Runge-Kutta methods of the form (3.1.2) with $s = 2$ and the following Butcher tableaux:

$$\text{Euler (4.1.1)}: \begin{array}{c|cc||c|cc||c} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ \hline & 1 & 0 & & 1 & 0 & \end{array}, \qquad \text{Euler (4.1.2)}: \begin{array}{c|cc||c|cc||c} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ \hline & 1 & 0 & & 0 & 1 & \end{array}.$$

Note that the class of methods (3.1.2) satisfies automatically Assumption 2.2.5.

Remark 4.4.1. *The class of Runge-Kutta methods* (3.1.2) *can be straightforwardly generalized (as done in [83] in the Euclidean case $\mathbb{R}^d$) to study partitioned problems where $f = f_1 + f_2$ and, for example, to create IMEX schemes. In order to improve the order of the method without increasing its cost, one could also apply a postprocessor (in the spirit of [123] in $\mathbb{R}^d$) or use multiple independent noises in* (3.1.2) *instead of only one random variable $\xi_n \sim \mathcal{N}(0, I_d)$. This last extension can increase the number of conditions but may also increase the set of solutions. We refer in particular to [45, 83] in the context of $\mathbb{R}^d$, where it is shown for a class of stochastic Runge-Kutta method that the order conditions for weak order 3 cannot be satisfied in general, unless we use at least two independent noises. In addition, if we write the internal stages of* (3.1.2) *as*

$$Y_i = X_n + h\sum_{j=1}^{s} a_{ij} f(Y_j) + \sigma\sqrt{h} d_i \xi_n + \left(\sum_{j=1}^{s} \widehat{a}_{ij} g(Y_j)\right)\lambda_i,$$

where $g : \mathbb{R}^d \to \mathbb{R}^{d\times q}$ and $\lambda_i \in \mathbb{R}^q$, then the same class of methods is also fit for solving (2.1.4) *with a multidimensional constraint $\zeta : \mathbb{R}^d \to \mathbb{R}^q$. Note that the coefficients of the method do not depend on the dimension of the space d or the codimension q of the manifold. This will be studied in future work.*

Remark 4.4.2. *If ξ_n is a Gaussian random variable, its realisations can be arbitrarily large, and the existence and uniqueness of the solution of the system* (3.1.2) *does not hold in general. A standard remedy to ensure that the projection on $\mathcal{M}$ always exists for $h \leqslant h_0$ small enough is to replace the standard Gaussian random vectors ξ in* (3.1.2) *by bounded discrete random vectors $\widehat{\xi}$ that have the same first moments in the spirit of [104, Chap. 2]. This way, the order of the method is preserved both in the weak sense and for the invariant measure, and the method is well-posed for all h small enough. For weak/ergodic order two, one can consider, for instance, the random vectors $\widehat{\xi}$ with independent components $\widehat{\xi}_i$ that satisfy*

$$\mathbb{P}(\widehat{\xi}_i = 0) = \frac{2}{3} \quad \textit{and} \quad \mathbb{P}(\widehat{\xi}_i = \pm\sqrt{3}) = \frac{1}{6}, \quad i = 1, \dots, d. \tag{4.4.1}$$

The following lemma guarantees the well-posedness of a method of the form (3.1.2) with bounded random variables $\widehat{\xi}_n$. The result is still true when A and $\widehat{A}$ are general matrices, but we consider only the lower triangular case for the sake of brevity. This result is in the spirit of [67, Chap. VII] for deterministic DAEs.

Lemma 4.4.3. *For Runge-Kutta methods of the form* (3.1.2) *where the ξ_n are replaced by bounded random variables $\widehat{\xi}_n$, there exists $h_0 > 0$ such that for all $h \leqslant h_0$, for any initial condition $X_n \in \mathcal{M}$, there exists a unique solution X_{n+1} of* (3.1.2) *in a neighbourhood of X_n. Furthermore, the internal stages satisfy $Y_i = X_n + \mathcal{O}(\sqrt{h})$ and $\lambda_i = \mathcal{O}(\sqrt{h})$ for $i = 1, \dots, s$.*

Proof. We proceed by induction on i. We assume that for $j < i$, the Y_j are already defined and satisfy $Y_j = X_n + \mathcal{O}(\sqrt{h})$. The result is straightforward if $\delta_i = 0$. We thus assume that $\delta_i = 1$ and prove the existence of a unique solution to the equations of the internal stage i:

$$Y_i = X_n + h\sum_{j=1}^{i-1} a_{ij} f(Y_j) + \sigma\sqrt{h} d_i \widehat{\xi}_n + \lambda_i \sum_{j=1}^{i} \widehat{a}_{ij} g(Y_j), \tag{4.4.2}$$

$$\zeta(Y_i) = 0. \tag{4.4.3}$$

Using $\zeta(X_n) = 0$, we write equation (4.4.3) as

$$\zeta(Y_i) - \zeta(X_n) = \int_0^1 g^T(X_n + \tau(Y_i - X_n)) d\tau (Y_i - X_n) = 0. \tag{4.4.4}$$

Inserting (4.4.2) in (4.4.4) yields

$$\int_0^1 g^T(X_n+\tau(Y_i-X_n))d\tau\Big[h\sum_{j=1}^{i-1}a_{ij}f(Y_j)+\sigma\sqrt{h}d_i\widehat{\xi}_n+\lambda_i\sum_{j=1}^{i}\widehat{a}_{ij}g(Y_j)\Big]=0. \tag{4.4.5}$$

Multiplying both sides of equation (4.4.2) by $\int_0^1 g^T(X_n+\tau(Y_i-X_n))d\tau\Big(\sum_{j=1}^{i}\widehat{a}_{ij}g(Y_j)\Big)$, and substituting λ_i in (4.4.2) with its value from (4.4.5), we deduce that $F(Y_i,h)=0$, where the function $F:\mathbb{R}^d\times\mathbb{R}\to\mathbb{R}^d$ is given by

$$\begin{aligned}F(y,t)=\int_0^1 g^T(X_n+\tau(y-X_n))d\tau\Big[&\Big(t\sum_{j=1}^{i-1}a_{ij}f(Y_j)+\sigma\sqrt{t}d_i\widehat{\xi}_n\Big)\Big(\sum_{j=1}^{i-1}\widehat{a}_{ij}g(Y_j)+\widehat{a}_{ii}g(y)\Big)\\&+\Big(\sum_{j=1}^{i-1}\widehat{a}_{ij}g(Y_j)+\widehat{a}_{ii}g(y)\Big)\Big(y-X_n-t\sum_{j=1}^{i-1}a_{ij}f(Y_j)-\sigma\sqrt{t}d_i\widehat{\xi}_n\Big)\Big].\end{aligned}$$

As $F(X_n,0)=0$ and the partial differential $\partial_y F(X_n,0)=G(X_n)I_d$ is invertible, the implicit function theorem yields the existence and uniqueness of Y_i in a ball of center X_n for $h\leqslant h_0$ small enough. As $\widehat{\xi}_n$ is bounded and $\mathcal{M}$ is compact, there exists a deterministic h_0 that works for every initial condition $X_n\in\mathcal{M}$. Now that Y_i is well-posed, we deduce from the identity $F(Y_i,h)=0$ that $Y_i=X_n+\mathcal{O}(\sqrt{h})$ and we derive from (4.4.5) that λ_i is well-posed for h small enough and satisfies $\lambda_i=\mathcal{O}(\sqrt{h})$. Finally we observe that (Y_i,λ_i) is indeed a solution to (4.4.2)-(4.4.3). □

Remark 4.4.4. *In practice, one can solve numerically each internal stage of the set of equations* (3.1.2) *with a fixed point iteration or a Newton method starting from $Y_i=X_n$ and $\lambda_i=0$. As $\mathcal{M}$ is compact, if the ξ_n are replaced by bounded random variables, these two methods converge for $h\leqslant h_0$ where h_0 is small enough and independent of the initial condition. It is crucial to initialize the Y_i in a neighbourhood of X_n as* (3.1.2) *has multiple solutions in general. For example, the Euler scheme* (4.1.2) *always has two solutions if $\mathcal{M}$ is a sphere (the two intersections of $\mathcal{M}$ and a straight line going through the center of $\mathcal{M}$).*

Before looking at the consistency and the order conditions of the class of methods (3.1.2), we introduce a concise notation for multiplying vectors component-wise.

Definition 4.4.5. *For $y,y^{(1)},\dots,y^{(n)}\in\mathbb{R}^d$ and $m\geqslant 0$, we define the diamond product and the diamond power as the vectors in $\mathbb{R}^d$,*

$$(y^{(1)}\diamond\ldots\diamond y^{(n)})_i=\prod_{k=1}^{n}y_i^{(k)}\quad\text{and}\quad(y^{\diamond m})_i=y_i^m.$$

We present below the detailed calculation of the consistency conditions of the class of methods (3.1.2) for the constrained overdamped Langevin equation (2.1.4). A key ingredient of this result is an expansion of the Lagrange multipliers λ_i in powers of h in the spirit of [90, Lemma 3.25]. Similar proofs can be found in [90, Prop. 3.24] for the Euler schemes (4.1.1)-(4.1.2), and in [5, 83] for Runge-Kutta methods in $\mathbb{R}^d$.

Proposition 4.4.6. *For a Runge-Kutta method of the form* (3.1.2)*, the operator* $\mathcal{A}_0$ *in* (2.2.6) *is given for* $\phi \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$ *by*

$$\mathcal{A}_0\phi = b^T\mathbb{1}\phi' f - b^T\mathbb{1}G^{-1}(g,f)\phi' g - \frac{\sigma^2}{2}d_s^2 G^{-1}\operatorname{div}(g)\phi' g + \frac{\sigma^2}{2}d_s^2\Delta\phi - \frac{\sigma^2}{2}d_s^2 G^{-1}\phi''(g,g)$$
$$+ \sigma^2 d_s\left(\widehat{b}^T d - \widehat{b}^T(\delta \bullet d) + \frac{1}{2}d_s\right)G^{-2}(g,g'g)\phi' g + \sigma^2 d_s\left(\widehat{b}^T(\delta \bullet d) - \widehat{b}^T d\right)G^{-1}\phi' g' g.$$

In particular, if

$$b^T\mathbb{1} = d_s = 1 \quad \textit{and} \quad \widehat{b}^T d = \widehat{b}^T(\delta \bullet d), \tag{4.4.6}$$

then the method is consistent, that is, $\mathcal{A}_0 = \mathcal{L}$.

Proof. If we apply one step of a method of the form (3.1.2) with the initial condition $X_0 = x$, then the internal stages Y_i satisfy the following expansion

$$Y_i = x + \sigma\sqrt{h}d_i\xi + hc_i f(x) + R^h, \quad \text{if} \quad \delta_i = 0,$$
$$Y_i = x + \sqrt{h}\left[\sigma d_i\xi + \lambda_{1/2,i}(x)g(x)\right] + h\Big[c_i f(x) + \lambda_{1,i}(x)g(x) + \sigma\lambda_{1/2,i}(x)\sum_{j=1}^s \widehat{a}_{ij}d_j g'(x)\xi$$
$$+ \lambda_{1/2,i}(x)\sum_{j=1}^s \widehat{a}_{ij}\lambda_{1/2,j}(x)\delta_j g'(x)g(x)\Big] + R^h, \quad \text{if} \quad \delta_i = 1,$$

where the remainder satisfies $\left|R^h\right| \leqslant Ch^{3/2}$, and where we used that λ_i can be developed in powers of $\sqrt{h}$ as $\lambda_i = \sqrt{h}\lambda_{1/2,i} + h\lambda_{1,i} + \dots$ in the spirit of [90, Lemma 3.25]. If $\delta_i = 1$, $\zeta(Y_i)$ can also be expanded as

$$\zeta(Y_i) = \zeta(x) + \sqrt{h}\left[\sigma d_i(g,\xi) + \lambda_{1/2,i}G\right] + h\Big[c_i(g,f) + \lambda_{1,i}G + \lambda_{1/2,i}\sum_{j=1}^s \widehat{a}_{ij}\lambda_{1/2,j}\delta_j(g,g'g)$$
$$+ \sigma\lambda_{1/2,i}\sum_{j=1}^s \widehat{a}_{ij}d_j(g,g'\xi) + \frac{1}{2}\sigma^2 d_i^2(\xi,g'\xi) + \sigma\lambda_{1/2,i}d_i(g,g'\xi) + \frac{1}{2}\lambda_{1/2,i}^2(g,g'g)\Big] + \dots$$

where we omitted the dependency in x of G, g, g' and the $\lambda_{k/2,j}$'s for brevity. We have $\zeta(Y_i) = \zeta(x) = 0$ (since $x \in \mathcal{M}$), thus by identifying each term of the expansion with zero, we get

$$\lambda_{1/2,i} = -\sigma\delta_i d_i G^{-1}(g,\xi),$$
$$\lambda_{1,i} = -\delta_i c_i G^{-1}(g,f) + \sigma^2\delta_i\left(\sum_{j=1}^s \widehat{a}_{ij}d_i d_j + d_i^2\right)G^{-2}(g,\xi)(g,g'\xi)$$
$$- \sigma^2\delta_i\left(\sum_{j=1}^s \widehat{a}_{ij}\delta_j d_i d_j + \frac{1}{2}d_i^2\right)G^{-3}(g,\xi)^2(g,g'g) - \frac{\sigma^2}{2}\delta_i d_i^2 G^{-1}(\xi,g'\xi).$$

For ϕ a test function, the operator $\mathcal{A}_0\phi$ satisfies

$$\mathbb{E}[\phi(X_1)] = \mathbb{E}[\phi(Y_s)] = \phi(x) + h\mathcal{A}_0\phi(x) + h^2\mathcal{A}_1\phi(x) + \dots$$

By replacing Y_s with its expansion in powers of $h^{1/2}$, and by identifying the first terms, we deduce that

$$\mathcal{A}_0\phi = \mathbb{E}\Big[c_s\phi' f - c_s G^{-1}(g,f)\phi' g - \frac{\sigma^2}{2} d_s^2 G^{-1}(\xi, g'\xi)\phi' g + \sigma^2 d_s(\widehat{b}^T d + d_s) G^{-2}(g,\xi)(g,g'\xi)\phi' g$$
$$- \sigma^2 d_s \left(\widehat{b}^T(\delta \bullet d) + \frac{1}{2} d_s\right) G^{-3}(g,\xi)^2 (g,g'g)\phi' g + \frac{\sigma^2}{2} d_s^2 \phi''(\xi,\xi) - \sigma^2 d_s^2 G^{-1}(g,\xi)\phi''(g,\xi)$$
$$+ \frac{\sigma^2}{2} d_s^2 G^{-2}(g,\xi)^2 \phi''(g,g) + \sigma^2 d_s \widehat{b}^T(\delta \bullet d) G^{-2}(g,\xi)^2 \phi' g' g - \sigma^2 d_s \widehat{b}^T d G^{-1}(g,\xi)\phi' g'\xi\Big],$$

where we used that $\delta_s = 1$ and that all the terms containing an odd number of ξ vanish since odd moments of ξ are zero. Distributing the expectation on each term and using $c_s = b^T \mathbb{1}$ yield the desired expression of $\mathcal{A}_0\phi$. We deduce the consistency conditions $b^T\mathbb{1} = d_s = 1$ and $\widehat{b}^T d = \widehat{b}^T(\delta \bullet d)$ in order to get $\mathcal{A}_0 = \mathcal{L}$. □

Remark 4.4.7. *The analysis presented in Section 4.4.1 is conducted for the overdamped Langevin dynamics* (2.1.4)*. It would be interesting to consider extensions with multiplicative noise or a non-gradient vector field f. The calculations would likely become more involved and we may get more order conditions (see, for instance, [5, Thm. 3.3] and [83, Remark 5.1 and Sect. 5.5] in the context of $\mathbb{R}^d$, where many additional terms arise, in particular for the integration by parts calculations). This will be studied in future work.*

4.4.2 Order conditions for the invariant measure on manifolds

We now derive the methodology for getting the conditions of arbitrary high order for sampling the invariant measure of the constrained overdamped Langevin equation (2.1.4). In particular, the following theorem presents the Runge-Kutta conditions for order two for the invariant measure on $\mathcal{M}$. Note that the number of conditions does not depend on the dimension of the space d.

Theorem 4.4.8 (Runge-Kutta conditions for order two for the invariant measure)**.** *We consider a Runge-Kutta method of the form* (3.1.1) *and assume the consistency condition* (4.4.6)*. If the method is ergodic and if the following conditions are satisfied, then the integrator has order two for the invariant measure of* (2.1.4)*:*

$$\widehat{b}^T d = b^T d,$$
$$b^T c = b^T(\delta \bullet c) = b^T d^{\bullet 2} = b^T(\delta \bullet d^{\bullet 2}) = 2\widehat{b}^T d - \tfrac{1}{2},$$
$$\widehat{b}^T c = \widehat{b}^T(\delta \bullet c) = \widehat{b}^T d^{\bullet 2} = \widehat{b}^T(\delta \bullet d^{\bullet 2}) = \widehat{b}^T d^{\bullet 3} = \widehat{b}^T(\delta \bullet d^{\bullet 3}) = 2\widehat{b}^T d - \tfrac{1}{2},$$
$$\widehat{b}^T(c \bullet d) = \widehat{b}^T(\delta \bullet c \bullet d),$$
$$b^T(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))) = 0,$$
$$\widehat{b}^T A((\delta - \mathbb{1}) \bullet d)) = \widehat{b}^T(\delta \bullet A((\delta - \mathbb{1}) \bullet d))) = (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \tfrac{1}{2},$$
$$\widehat{b}^T(d \bullet \widehat{A}c) = \widehat{b}^T(d \bullet \widehat{A}d^{\bullet 2}) = \widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) = 2\widehat{b}^T(d \bullet \widehat{A}d) + (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \tfrac{1}{2},$$
$$\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) = \widehat{b}^T(d \bullet \widehat{A}d) + \tfrac{1}{2}(\widehat{b}^T d)^2,$$
$$\widehat{b}^T(c \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d) + \widehat{b}^T(d \bullet \widehat{A}((\delta - 3 \cdot \mathbb{1}) \bullet d) + \widehat{b}^T(d \bullet \widehat{A}(\delta \bullet c) = 2(\widehat{b}^T d)^2 - 4\widehat{b}^T d + 1,$$
$$\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}(\delta \bullet d)) + \widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d)) = 2\widehat{b}^T(d \bullet \widehat{A}d) + \tfrac{3}{2}(\widehat{b}^T d)^2 - 2\widehat{b}^T d + \tfrac{1}{2},$$
$$\widehat{b}^T(d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) = 0,$$
$$\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) + 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))) = (4 - 2\widehat{b}^T d)\widehat{b}^T(d \bullet \widehat{A}d) + 3(\widehat{b}^T d)^2 - 4\widehat{b}^T d + 1.$$

In the particular case where we set $\delta = \mathbb{1}$, the order two conditions reduce to the following:

$$
\begin{aligned}
&(\widehat{b}^T d)^2 - 2\widehat{b}^T d + \tfrac{1}{2} = 0,\\
&\widehat{b}^T d = b^T d,\\
&b^T c = b^T d^{\bullet 2} = \widehat{b}^T c = \widehat{b}^T d^{\bullet 2} = \widehat{b}^T d^{\bullet 3} = 2\widehat{b}^T d - \tfrac{1}{2},\\
&\widehat{b}^T (d \bullet \widehat{A} c) = \widehat{b}^T (d \bullet \widehat{A} d^{\bullet 2}) = 2\widehat{b}^T (d \bullet \widehat{A} d),\\
&\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A} d) = \widehat{b}^T (d \bullet \widehat{A} d) + \widehat{b}^T d - \tfrac{1}{4},\\
&\widehat{b}^T (d \bullet (\widehat{A} d)^{\bullet 2}) = (4 - 2\widehat{b}^T d)\widehat{b}^T (d \bullet \widehat{A} d) + 2\widehat{b}^T d - \tfrac{1}{2}.
\end{aligned}
$$

For simplicity, we used in Theorem 4.4.8 the notation $\bullet$ of Definition 4.4.5. For instance, the condition $\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A} d) = \widehat{b}^T (d \bullet \widehat{A} d) + \frac{1}{2}(\widehat{b}^T d)^2$ rewrites into

$$
\sum_{i,j=1}^{s} \widehat{b}_i d_i^2 \widehat{a}_{ij} d_j = \sum_{i,j=1}^{s} \widehat{b}_i d_i \widehat{a}_{ij} d_j + \frac{1}{2}\Big(\sum_{i=1}^{s} \widehat{b}_i d_i \Big)^2 .
$$

The order conditions of Theorem 4.4.8 can be obtained from straightforward calculations with the following methodology. We compute the operator $\mathcal{A}_1$ with the same method used for $\mathcal{A}_0$ in Proposition 4.4.6. It is a differential operator of order four, and we present the complete expansion of $\mathcal{A}_1$ in exotic aromatic forests in Table C.1. We use the process of integration by parts as presented in Section 3.3.2 to transform $\int_{\mathcal{M}} \mathcal{A}_1 \phi d\mu_\infty$ into an integral of the form $\int_{\mathcal{M}} \mathcal{A}_1^0 \phi d\mu_\infty$ where $\mathcal{A}_1^0 \phi$ is a differential operator of lower order in ϕ. The integrations by parts of the order four and three terms are in Appendix B, and the complete expansion of $\mathcal{A}_1^0$ in exotic aromatic forests is in Table C.2. Then, we find sufficient conditions such that $\mathcal{A}_1^0 = 0$. This implies that $\mathcal{A}_1^* \rho_\infty = 0$, and Theorem 2.3.2 gives the order two for the invariant measure.

Although constructing methods of high weak order is not the main focus of this work, considering the explicit formula for $\mathcal{A}_1$ and comparing with $\mathcal{L}^2/2$ (see Appendix C for their detailed expansion in B-series), one immediately obtains the following result for weak order two of accuracy.

Theorem 4.4.9 (Runge-Kutta conditions for weak order two)**.** *We consider a Runge-Kutta method of the form* (3.1.1) *and assume that it satisfies* (4.4.6)*. If the following conditions are satisfied, then the integrator has weak order two for solving* (2.1.4)*:*

$$
\begin{aligned}
&b^T d = b^T c = b^T (\delta \bullet c) = b^T d^{\bullet 2} = b^T (\delta \bullet d^{\bullet 2}) = \tfrac{1}{2},\\
&\widehat{b}^T d = \widehat{b}^T c = \widehat{b}^T (\delta \bullet c) = \widehat{b}^T d^{\bullet 2} = \widehat{b}^T (\delta \bullet d^{\bullet 2}) = \widehat{b}^T d^{\bullet 3} = \widehat{b}^T (\delta \bullet d^{\bullet 3}) = \tfrac{1}{2},\\
&\widehat{b}^T (c \bullet d) = \widehat{b}^T (\delta \bullet c \bullet d),\\
&\widehat{b}^T (d \bullet \widehat{A} d) = \tfrac{1}{8},\\
&b^T (d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) = 0,\\
&\widehat{b}^T A((\mathbb{1} - \delta) \bullet d) = \widehat{b}^T (\delta \bullet A((\mathbb{1} - \delta) \bullet d)) = \tfrac{1}{4},\\
&\widehat{b}^T (d \bullet \widehat{A} c) = \widehat{b}^T (d \bullet \widehat{A} d^{\bullet 2}) = \widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) = 0,\\
&\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A} d) = \tfrac{1}{4},\\
&\widehat{b}^T (c \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) - \widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d)) - \widehat{b}^T (d \bullet \widehat{A}(\delta \bullet c)) = \tfrac{1}{8},\\
&\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}(\delta \bullet d)) + \widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d)) = \tfrac{1}{8},\\
&\widehat{b}^T (d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) = 0,\\
&\widehat{b}^T (d \bullet (\widehat{A} d)^{\bullet 2}) + 3\widehat{b}^T (d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))) = \tfrac{1}{8}.
\end{aligned}
$$

Remark 4.4.10. *For $\delta = \mathbb{1}$, the weak order two conditions of Theorem 4.4.9 have no solution, which is in contrast with the invariant measure case presented in Theorem 4.4.8. Indeed, the condition $\widehat{b}^T A((\mathbb{1} - \delta) \bullet d) = \frac{1}{4}$ cannot be fulfilled if we fix $\delta = \mathbb{1}$.*

4.4.3 Illustrative examples of high order Runge-Kutta methods on manifolds

In this section, we give several examples of high order Runge-Kutta methods of the form (3.1.2). The purpose of these examples is to illustrate our analysis, and deriving new integrators with small error constant, favourable stability properties, small variance and fast convergence to equilibrium is a challenging open question which is not addressed in the present work. First, we introduce a method that has order two for sampling the invariant measure of the constrained Langevin dynamics (2.1.4). Since there are many solutions to the order conditions, we obtain this integrator by solving numerically an optimization problem: we minimize the absolute values of the coefficients of the method under the constraints given by the order conditions of Theorem 4.4.8. This method is explicit in f and uses only three evaluations of f per step. It is defined by the following Butcher tableau

$$\begin{array}{c|cccc||c|cccc||c}
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & d_1 \\
c_2 & c_2 & 0 & 0 & 0 & 1 & \widehat{a}_{21} & \widehat{a}_{22} & 0 & 0 & d_2 \\
c_3 & 0 & c_3 & 0 & 0 & 1 & \widehat{a}_{31} & \widehat{a}_{32} & \widehat{a}_{33} & 0 & d_3 \\
1 & \widehat{a}_{41} & \widehat{a}_{42} & \widehat{a}_{43} & 0 & 1 & \widehat{a}_{41} & \widehat{a}_{42} & \widehat{a}_{43} & 0 & 1 \\
\hline
 & \widehat{a}_{41} & \widehat{a}_{42} & \widehat{a}_{43} & 0 & & \widehat{a}_{41} & \widehat{a}_{42} & \widehat{a}_{43} & 0 &
\end{array}$$

or by the associated set of equations

$$\begin{aligned}
Y_1 &= X_n + \sigma\sqrt{h} d_1 \xi_n + \lambda_1 g(Y_1), \\
Y_2 &= X_n + h c_2 f(Y_1) + \sigma\sqrt{h} d_2 \xi_n + \lambda_2 \left[\widehat{a}_{21} g(Y_1) + \widehat{a}_{22} g(Y_2)\right], \\
Y_3 &= X_n + h c_3 f(Y_2) + \sigma\sqrt{h} d_3 \xi_n + \lambda_3 \left[\widehat{a}_{31} g(Y_1) + \widehat{a}_{32} g(Y_2) + \widehat{a}_{33} g(Y_3)\right], \\
X_{n+1} &= X_n + h \sum_{j=1}^{3} \widehat{a}_{4j} f(Y_j) + \sigma\sqrt{h}\xi_n + \lambda_4 \sum_{j=1}^{3} \widehat{a}_{4j} g(Y_j), \qquad (4.4.7) \\
&\text{where } \lambda_1, \lambda_2, \lambda_3, \lambda_4 \text{ are such that } \zeta(Y_1) = \zeta(Y_2) = \zeta(Y_3) = \zeta(X_{n+1}) = 0,
\end{aligned}$$

and with the values of c_i, d_i, $\widehat{a}_{ij}$ given in Appendix A. To implement one step of this scheme, we apply a few iterations of the Newton method to find the projections onto $\mathcal{M}$. We emphasize that if the stepsize h is not small enough, the fixed point problems of finding λ_i such that $\zeta(Y_i) = 0$ may not be well defined, leading to diverging Newton iterations. Following Remark 4.4.2, we replace the standard Gaussian random vectors ξ in (3.1.2) by independent bounded discrete random vectors $\widehat{\xi}$ that satisfy (4.4.1). This way, the order two for the invariant measure is preserved and the method is well-posed for h small enough.

With the same methodology we used to obtain the order conditions of Theorem 4.4.8 and Theorem 4.4.9, and with the expressions of $\mathcal{A}_1\phi$ and $\mathcal{A}_1^0\phi$ (see Section 3.3.2 for further details), we also get classes of Runge-Kutta integrators and their order conditions for the following specific subproblems.

Euclidean case $\mathbb{R}^d$. Fixing $g = 0$ in the expressions of $\mathcal{A}_1\phi$ and $\mathcal{A}_1^0\phi$ yields the order two conditions in the weak sense and for the invariant measure in $\mathbb{R}^d$ as given in [83, Tables 1-2].

Deterministic case. Fixing $\sigma = 0$ in the expression of $\mathcal{A}_1\phi$ yields the order conditions for approximating the solution of ODEs of the form $\dot{x} = \Pi_{\mathcal{M}}(x)f(x)$, where f is a gradient. Note that this equation can be written as the following differential algebraic equation (DAE) of index two (see [67, Chap. VII]):

$$\begin{aligned} \dot{x} &= f(x) + \lambda g(x), \\ 0 &= \zeta(x). \end{aligned} \tag{4.4.8}$$

We obtain a class of deterministic Runge-Kutta methods for solving DAEs of the form (4.4.8) by setting $\sigma = 0$ in (3.1.2). A Runge-Kutta method of this form is consistent if $b^T\mathbb{1} = 1$, and has order two if $\widehat{b}^T c = \widehat{b}^T(\delta \bullet c) = b^T c = b^T(\delta \bullet c) = 1/2$. For instance, an order two method for solving ODEs of the form (4.4.8) is

$$X_{n+1} = X_n + h\frac{f(X_n) + f(X_{n+1})}{2} + \lambda\frac{g(X_n) + g(X_{n+1})}{2}, \quad \zeta(X_{n+1}) = 0.$$

Spherical case. In the simple case where $\mathcal{M}$ is the unit sphere in $\mathbb{R}^d$ (that is, when the constraint is of the form $\zeta(x) = (|x|^2 - 1)/2$ and $g(x) = x$), the consistency conditions (4.4.6) reduce to $b^T\mathbb{1} = d_s = 1$. The weak order two conditions of Theorem 4.4.9 reduce to the following conditions:

$$\begin{aligned} &b^T d = b^T c = b^T(\delta \bullet c) = b^T d^{\bullet 2} = b^T(\delta \bullet d^{\bullet 2}) = \widehat{b}^T d = \widehat{b}^T c = \tfrac{1}{2}, \\ &\widehat{b}^T(d \bullet \widehat{A}d) = \tfrac{1}{8}, \\ &\widehat{b}^T A((\mathbb{1} - \delta) \bullet d) = \tfrac{1}{4}, \\ &\widehat{b}^T(d \bullet \widehat{A}c) = 0. \end{aligned}$$

On the other hand, the order two conditions for the invariant measure of Theorem 4.4.8 on the sphere are the following:

$$\begin{aligned} &\widehat{b}^T d = b^T d, \\ &b^T c = b^T(\delta \bullet c) = b^T d^{\bullet 2} = b^T(\delta \bullet d^{\bullet 2}) = \widehat{b}^T c = 2\widehat{b}^T d - \tfrac{1}{2}, \\ &\widehat{b}^T(d \bullet \widehat{A}c) = 2\widehat{b}^T(d \bullet \widehat{A}d) + (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \tfrac{1}{2}, \\ &\widehat{b}^T A((\delta - \mathbb{1}) \bullet d) = (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \tfrac{1}{2}. \end{aligned}$$

For example, the following integrator has order two for the invariant measure if $\mathcal{M}$ is a sphere as its coefficients satisfy the order conditions on the sphere:

$$\begin{aligned} Y_1 &= X_n + h\left(\frac{3}{2} - \sqrt{2}\right) f(Y_2) + \sigma\sqrt{h}\left(1 - \frac{\sqrt{2}}{2}\right)\xi_n + \lambda_1(2Y_1 - Y_2), \quad \zeta(Y_1) = 0, \\ Y_2 &= X_n + hf(Y_1) + \sigma\sqrt{h}\xi_n + \lambda_2 Y_1, \quad \zeta(Y_2) = 0, \\ X_{n+1} &= Y_2. \end{aligned}$$

Brownian motions on manifolds. Runge-Kutta methods of the form (3.1.2) can also be used for simulating a Brownian motion on a manifold (see [71, Chap. III]) by solving numerically

$$dX(t) = \Pi_{\mathcal{M}}(X(t)) \circ dW(t), \quad X(0) = X_0 \in \mathcal{M}. \tag{4.4.9}$$

We recall that in the context of $\mathbb{R}^d$, the Euler-Maruyama integrator is exact for approximating a Brownian motion in law. However, in the context of manifolds, there are no exact Runge-Kutta integrators for simulating a Brownian motion on $\mathcal{M}$ in general. In particular, the Euler scheme (4.1.2) only has weak order one for solving (4.4.9) in general. Fixing $f = 0$ in (3.1.2) yields a class of Runge-Kutta methods for solving (4.4.9). The consistency conditions are $d_s = 1$ and $\widehat{b}^T d = \widehat{b}^T(\delta \bullet d)$. The conditions for order two for the invariant measure (respectively for weak order two) of such a Runge-Kutta method are obtained by deleting the order conditions in Theorem 4.4.8 (respectively in Theorem 4.4.9) that involve A, b or c. In the specific case where $\mathcal{M}$ is a sphere, the consistency conditions (4.4.6) become $d_s = 1$ and the weak order two conditions of Theorem 4.4.9 reduce to the two following conditions

$$\widehat{b}^T d = \frac{1}{2}, \quad \widehat{b}^T(d \bullet \widehat{A}d) = \frac{1}{8}.$$

For example, a weak order two method for simulating a Brownian motion on a sphere is

$$X_{n+1} = X_n + \sqrt{h}\xi_n + \lambda \frac{3X_n + \sqrt{h}\xi_n + X_{n+1}}{4}, \quad \zeta(X_{n+1}) = 0.$$

In addition, there are no additional order two conditions for the invariant measure, that is, any consistent integrator, such as the Euler scheme (4.1.2), has at least order two for the invariant measure on the sphere.

4.5 Numerical experiments

In this section, we perform numerical experiments to confirm the theoretical findings of Section 4.4, first on a sphere and a torus in $\mathbb{R}^3$, and then on the special linear group. We refer to [5] for further numerical experiments on the class of Runge-Kutta methods (3.1.1) in the context of $\mathbb{R}^d$.

4.5.1 Invariant measure approximation on a sphere and a torus

To check the numerical order two of the Runge-Kutta integrator (4.4.7) presented in Section 4.4.3, we first compare it with the Euler scheme (4.1.2) on the unit sphere in $\mathbb{R}^3$, where the constraint is given by $\zeta(x) = (x_1^2+x_2^2+x_3^2-1)/2$. We choose the potential $V(x) = 25(1-x_1^2-x_2^2)$, with $\sigma = \sqrt{2}$, $\phi(x) = x_3^2$, $f = -\nabla V$, $g = \nabla\zeta$, $M = 10^7$ independent trajectories to have a small Monte-Carlo error and a final time $T = 20$. Observe that for the smaller final time $T = 10$ (not included in the figures for conciseness), the convergence curves reveal to be nearly identical to the case $T = 20$ considered in Figure 4.1, which suggests that the numerical solutions are already very close to equilibrium at these final times. Following Remark 4.4.2 and Lemma 4.4.3, we use discrete bounded random variables satisfying (4.4.1) in the implementation of the integrators. For both integrators, we compute the Monte-Carlo estimator $\bar{J} = \frac{1}{M}\sum_{m=1}^{M} \phi(X_N^{(m)}) \simeq \mathbb{E}[\phi(X_N)]$, where $X_n^{(m)}$ is the m-th realisation of the integrator at time $t_n = nh$, and N is an integer satisfying $Nh = T$. We compare this approximation with a reference value of $\int_{\mathcal{M}} \phi d\mu_\infty$ computed via a standard quadrature formula, and we plot the error for the invariant measure (2.3.3) versus different timestep h. We also plot an estimate of the Monte-Carlo error by using the standard error of the mean estimator $\big(\sum_{m=1}^{M}(\phi(X_N^{(m)}) - \bar{J})^2\big)^{1/2}/\sqrt{M(M-1)}$. We observe in all convergence plots that the Monte-Carlo error prevails for small values of the timestep h. In Figure 4.1, we observe as expected order one for the Euler scheme (4.1.2) and order two for the Runge-Kutta scheme (4.4.7).

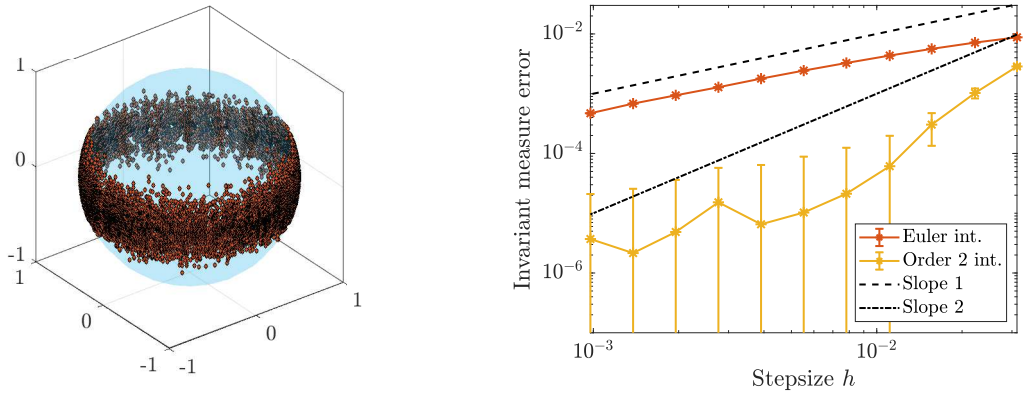


Figure 4.1: A trajectory of the order two method (left) and the convergence curve for the sphere for the invariant measure (right) with the potential $V(x) = 25(1 - x_1^2 - x_2^2)$, $\phi(x) = x_3^2$, a final time $T = 20$ and $M = 10^7$ trajectories.

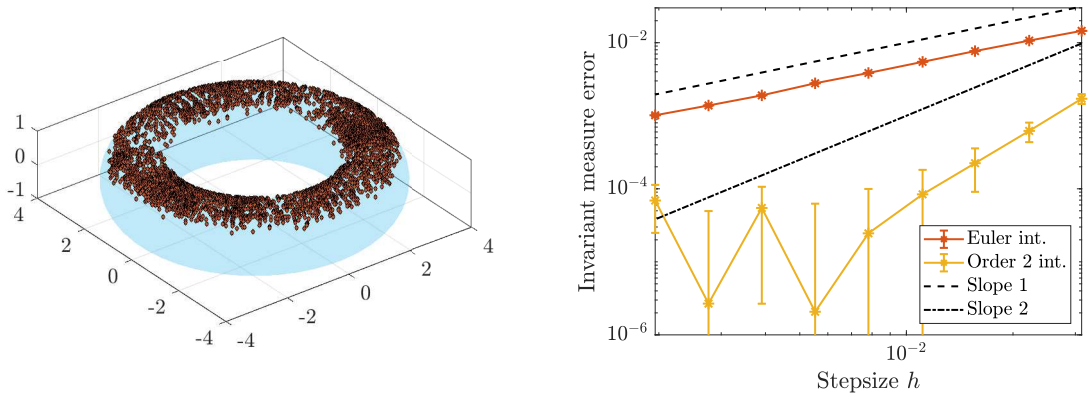


Figure 4.2: A trajectory of the order two method (left) and the convergence curve for the torus for the invariant measure (right) with the potential $V(x) = 25(x_3 - r)^2$, $\phi(x) = x_3^2$, a final time $T = 20$ and $M = 10^7$ trajectories.

We then apply the Euler scheme (4.1.2) and the Runge-Kutta integrator (4.4.7) on a torus defined by the constraint $\zeta(x) = (x_1^2 + x_2^2 + x_3^2 + R^2 - r^2)^2 - 4R^2(x_1^2 + x_2^2)$ with $R = 3$ and $r = 1$. The potential is $V(x) = 25(x_3 - r)^2$ and we choose $\sigma = \sqrt{2}$, $\phi(x) = x_3^2$, $f = -\nabla V$, $g = \nabla\zeta$, a final time $T = 20$ and $M = 10^7$ independent trajectories. On Figure 4.2, we plot the error for the invariant measure versus the timestep h, by using a reference value for $\int_{\mathcal{M}} \phi d\mu_\infty$ obtained with a standard quadrature formula. As expected, we observe order two for the proposed integrator. These curves confirm the theoretical findings presented in Section 4.4. In particular, the scheme (4.4.7) has order two of accuracy for the invariant measure on manifolds, according to Theorem 4.4.8. Note that if we had chosen a very short final time T, we would have observed the weak order one instead of the order two for the invariant measure as we would not have reached equilibrium.

4.5.2 Invariant measure approximation on the special linear group

Sampling on a manifold $\mathcal{M}$ is especially useful to compute integrals of the form $\int_{\mathcal{M}} \phi(x) d\mu_\infty$ when $\mathcal{M}$ is a manifold of high dimension. The class of methods (3.1.2) is convenient as the number of order conditions does not increase with the dimension of the space increasing. We apply Method (4.4.7) on a Lie group (in the spirit of [124, 125]) to see how it performs in high dimension. We choose the special linear group $\mathrm{SL}(m) = \{M \in \mathbb{R}^{m\times m}, \det(M) = 1\}$, seen as a submanifold of $\mathbb{R}^{m^2}$ of codimension 1. As explained in Remark 4.2.2, our analysis still applies to $\mathrm{SL}(m)$ if we choose a potential V with appropriate growth assumptions, even if it is not a compact manifold. We compare the Euler scheme (4.1.2) and the Runge-Kutta integrator (4.4.7) on $\mathcal{M} = \mathrm{SL}(m)$ for different m (that is, with the constraint $\zeta(x) = \det(x) - 1$), where we use in the implementation discrete random variables satisfying (4.4.1). We choose the potential

$$V(x) = 25\,\mathrm{Tr}((x - I_{m^2})^T (x - I_{m^2})) \tag{4.5.1}$$

and the parameters $\sigma = \sqrt{2}$, $\phi(x) = \mathrm{Tr}(x)$ and $M = 10^6$ trajectories. Each trajectory is an approximation of the solution of equation (2.1.4) at time $T = 10$ with a timestep $h = T/N$ and $N = 2^{12}$ steps. With this timestep h, the Newton method used in the Euler scheme (4.1.2) does not converge for approximately 0.005% of the trajectories for $m = 4$. We choose to discard these trajectories, which induces a negligible bias in the expectation. This does not occur for the Runge-Kutta integrator (4.4.7). We recall that for a small enough timestep h, the Newton method would always converge (see also Remark 4.4.4). The reference solution for $J(m) = \int_{\mathrm{SL}(m)} \phi(x) d\mu_\infty(x)$ is computed with the Runge-Kutta method (4.4.7) with $h_{\mathrm{ref}} = 2^{-14}T$. With the factor 25 in the potential (4.5.1), the solution of (2.1.4) stays close to I_{m^2}, and $J(m)$ is close to $\phi(I_{m^2}) = m$. This choice of factor permits to explore a reasonably small area of $\mathrm{SL}(m)$ with moderate manifold curvature. We observe numerically that replacing the factor 25 by 1 in (4.5.1) induces a severe timestep restriction (results not included for conciseness). The computation of $J(m)$ could also be done via the parametrization given by the Iwasawa decomposition for $\mathrm{SL}(m)$ (see, for instance, [62, Chap. 1]) and the use of standard quadrature methods, but these methods have prohibitive costs in high dimension. We put together the numerical results in Table 4.11 and observe that the Runge-Kutta method (4.4.7) performs significantly better than the Euler scheme (4.1.2).

m	$\dim(\mathrm{SL}(m))$	$J(m)$	$\bar{J}_{\mathrm{Euler}}$	error for $\bar{J}_{\mathrm{Euler}}$	$\bar{J}_2$	error for $\bar{J}_2$
2	3	2.00967	2.01031	$6.4 \cdot 10^{-4}$	2.00962	$4.4 \cdot 10^{-5}$
3	8	3.01954	3.02068	$1.1 \cdot 10^{-3}$	3.01934	$2.0 \cdot 10^{-4}$
4	15	4.02930	4.03095	$1.6 \cdot 10^{-3}$	4.02907	$2.3 \cdot 10^{-4}$

Table 4.11: Numerical approximation of the integral $J(m) = \int_{\mathrm{SL}(m)} \phi(x) d\mu_\infty$ for $2 \leqslant m \leqslant 4$ with the estimator $\bar{J} = M^{-1} \sum_{k=1}^{M} \phi(X_N^{(k)})$ where (X_n) is given by the Euler scheme (4.1.2) for $\bar{J}_{\mathrm{Euler}}$ and by the Runge-Kutta integrator (4.4.7) for $\bar{J}_2$, with their respective errors. The average is taken over $M = 10^6$ trajectories, with the potential (4.5.1), $\phi(x) = \mathrm{Tr}(x)$, a final time $T = 10$ and a timestep $h = 2^{-12}T$.

CHAPTER 5

Multirevolution integrators for SDEs with fast stochastic oscillations

Note: This chapter is almost identical to the article [84] in collaboration with G. Vilmart.

5.1 Introduction

In this chapter, we develop invariant-preserving integrators of second weak order that are robust with respect to the stiffness ε both in accuracy and cost for the class of highly-oscillatory d-dimensional SDEs (2.1.5). This class of SDEs includes in particular highly-oscillatory Kubo oscillators (see [39])

$$dX(t) = \frac{2\pi}{\sqrt{\varepsilon}} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} X(t) \circ dW(t) + \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix} X(t)dt, \ a \in \mathbb{R}, \tag{5.1.1}$$

or equivalently, $dY = 2i\pi\varepsilon^{-1/2} Y \circ dW + iaYdt$ in the complex setting where $Y = X_1 + iX_2$.

Applying standard SDE integrators to solve equation (2.1.5) requires in general a time stepsize $h \leqslant \varepsilon$ to be accurate, which makes these methods dramatically expensive when ε is small. The goal of this work is to create robust numerical methods, i.e., numerical integrators whose cost and accuracy do not deteriorate when ε becomes small. Several classes of methods have already been developed for highly-oscillatory SDEs with a deterministic fast oscillation (see for instance [41, 122]), but not in the case where the stiff oscillatory part is applied to the noise itself. To numerically face this challenge, we introduce a new methodology to develop robust methods of any high weak order to approximate the solution of equation (2.1.5). In particular, we propose a method of weak order two, and a geometric modification of this algorithm that preserves quadratic invariants.

Stochastic oscillations as defined in (2.1.5) typically arise in fiber optics models (see [7, 8, 58]) with a spatial discretization of the highly-oscillatory nonlinear Schrödinger equation (NLS) with white noise dispersion

$$du(t) = \frac{i}{\sqrt{\varepsilon}} \Delta u(t) \circ dW(t) + F(u(t))dt, \ u(t=0) = u_0. \tag{5.1.2}$$

As described for instance in [58], in the case $\varepsilon = 1$, the NLS equation (5.1.2) with a cubic nonlinearity $F(u) = |u|^2 u$ is a model in dimension $d = 1$ describing the propagation of a signal in optical fibers where x corresponds to the retarded time, while t corresponds to the distance along the fiber. Taking into account the inevitable chromatic dispersion effects of the signal, modeled by a random centered stationary process m with a coefficient $\nu > 0$, yields the following random PDE,

$$\frac{\partial v}{\partial x}(x,t) = \nu i m(x) \frac{\partial^2 v}{\partial t^2}(x,t) + \nu^2 F(v(x,t)),\ v(x=0,t) = u_0(t).$$

The perfect fiber would satisfy $m = 0$, but in practice, engineers build fibers with a small varying dispersion coefficient. To limit the pulse broadening induced by random dispersion, specialists use a wide range of dispersion management techniques (see for instance [58] and references therein). In [97, 44], the authors show that if we denote $u^\nu(x,t) = v(x/\nu^2, t)$, then as ν tends to 0 and under some ergodicity assumptions on m, u^ν converges to the solution u of equation (5.1.2) with $\varepsilon = 1$. The non-stiff counterpart of equation (5.1.2), i.e., for $\varepsilon = 1$, has also been studied theorically in [49] for a particular nonlinearity. The highly-oscillatory behavior ($\varepsilon \ll 1$) appears naturally when observing the propagation in long time with a small nonlinearity (via the change of variable $t \leftarrow \varepsilon t$) or the propagation of a small initial data in an optical fiber with a polynomial nonlinearity (via the change of variable $u \leftarrow u/\varepsilon$). A goal of this chapter is to develop efficient and cheap numerical methods that can model the propagation of pulses in this context, in order to observe some specific behaviors and, ultimately, to build enhanced fibers. Models of the form (5.1.2) also appear in the recent work [54] in the context of stochastic three-wave semi-linear systems. We emphasize that there is a growing interest in the recent litterature for stochastic models involving a fast Stratonovitch noise in the context of ergodic stochastic dynamics. In [4], it is shown for a class of overdamped Langevin equations that adding an appropriate fast Stratnovitch noise permits to increase the convergence rate to equilibrium, while reducing the asymptotic variance at infinity. This suggests that new efficient samplers for the invariant distribution of Langevin type models in context of large dimensional molecular dynamics models could be developed. We also mention the recent homogenization results on stochastic dynamics with fast Stratonovitch noises in [94] where our periodicity assumption is replaced by an ergodicity assumption on the fast component of the dynamics posed on manifolds.

Numerous possibilities exist for numerically integrating equations (2.1.5) or (5.1.2). We highlight in particular the exponential integrators [39, 52] for the SDE (2.1.5), and the exponential integrators [40], the Fourier split-step method [97] or the Crank-Nicholson scheme [11] for the SPDE (5.1.2). These methods have the advantage that they preserve the L^2 invariant of the equation (that is, $\|u(t)\|_{L^2} = \|u_0\|_{L^2}$ for all $t \geqslant 0$) for a class of polynomial nonlinearities. However, they face a severe timestep restriction $h \leqslant \varepsilon$ when the stiff parameter ε is small. Even in the case of deterministic oscillations, there are restrictions in general, though some robust algorithms exist (see [41] for instance). The methods presented in this chapter solve this issue of stepsize restriction. The idea is to approximate the solution of equation (2.1.5) at random times called revolution times because they correspond to complete revolutions of the oscillatory part $dX = \varepsilon^{-1/2} AX \circ dW$. This is in the spirit of [70] which also approximates the solution of SDEs at random times.

The chapter is organized as follows. Section 5.2 is devoted to the presentation of the new integrators. In Section 5.3, we build an asymptotic expansion of the solution of (2.1.5) and evaluate it at revolution times to derive the new integrators and a limit model for equation (2.1.5). Section 5.4 is devoted to the weak convergence theorems and their proofs. In Section 5.5, we

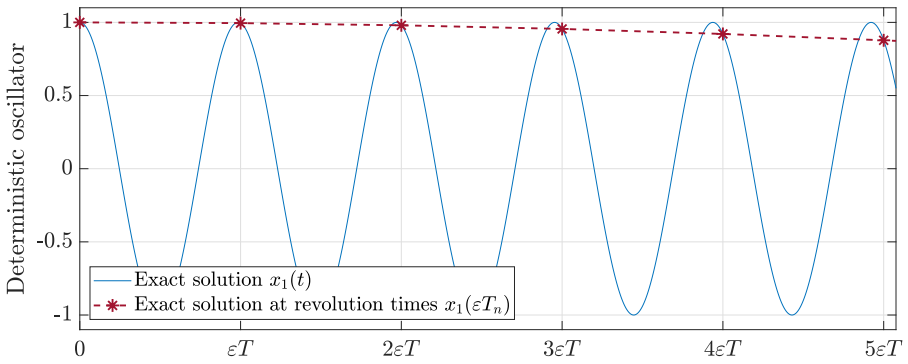


Figure 5.1: Exact solution evaluated at revolution times for the deterministic oscillator (5.2.1) with $F(y) = iy$ and $\varepsilon = 10^{-1}$.

present numerical experiments to confirm our theoretical error estimates, and we apply the new methods to solve numerically the Schrödinger equation (5.1.2).

5.2 Multirevolution integrators for stochastic oscillators

Initially created in [102, 27] in the context of celestial mechanics and later extended using geometric integration (see for instance [106, 28, 35]), multirevolution methods represent a class of numerical methods used for solving highly-oscillatory differential equations while reducing the cost of computation. In particular, they can approximate the solution of highly-oscillatory ODEs of the following form at stroboscopic times εNT, where $T = 1$ is the period of $\frac{dx}{dt} = Ax$, and N is an integer,

$$\frac{dx}{dt} = \frac{1}{\varepsilon}Ax + F(x), \ x(0) = x_0. \tag{5.2.1}$$

The solution x of this equation at times εNT is a perturbation of identity, that is, x satifies $x(\varepsilon t) = x_0 + \mathcal{O}(\varepsilon t)$, thus the solution loses its highly-oscillatory feature when evaluated at stroboscopic times, as shown in Figure 5.1 for the first component of the solution of equation (5.2.1) with $F(x) = ix$ (respectively $F(y) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y$ in the real setting). The idea of multirevolution is to approximate $x(\varepsilon N)$ with $N = \mathcal{O}(\varepsilon^{-1})$ with a cost independent of ε.

For stochastic oscillations, the solution $X(t) = e^{\varepsilon^{-1/2}AW(t)}X_0$ of $dX = \varepsilon^{-1/2}AX \circ dW$ is not periodic, but satisfies $X(\varepsilon T_N) = X_0$ where the T_N are random variables called revolution times and defined by

$$\begin{aligned} T_0 &= 0, \\ T_{N+1} &= \inf\left\{t > T_N, \varepsilon^{-1/2}\left|W(\varepsilon t) - W(\varepsilon T_N)\right| \geqslant 1\right\}, \ N = 0, 1, 2, \dots \end{aligned} \tag{5.2.2}$$

If X is the solution of (2.1.5), we show in Subsection 5.3.1 that X evaluated at times εT_N is a perturbation of the identity (in a strong and weak sense). Figure 5.2 illustrates the definition of revolution times and shows the perturbation of identity property on the first component of a Kubo oscillator (5.1.1) with $a = 1$. We highlight that the revolutions times T_N can be simulated without simulating the exact path W. Also we emphasize that the proposed algorithms do not

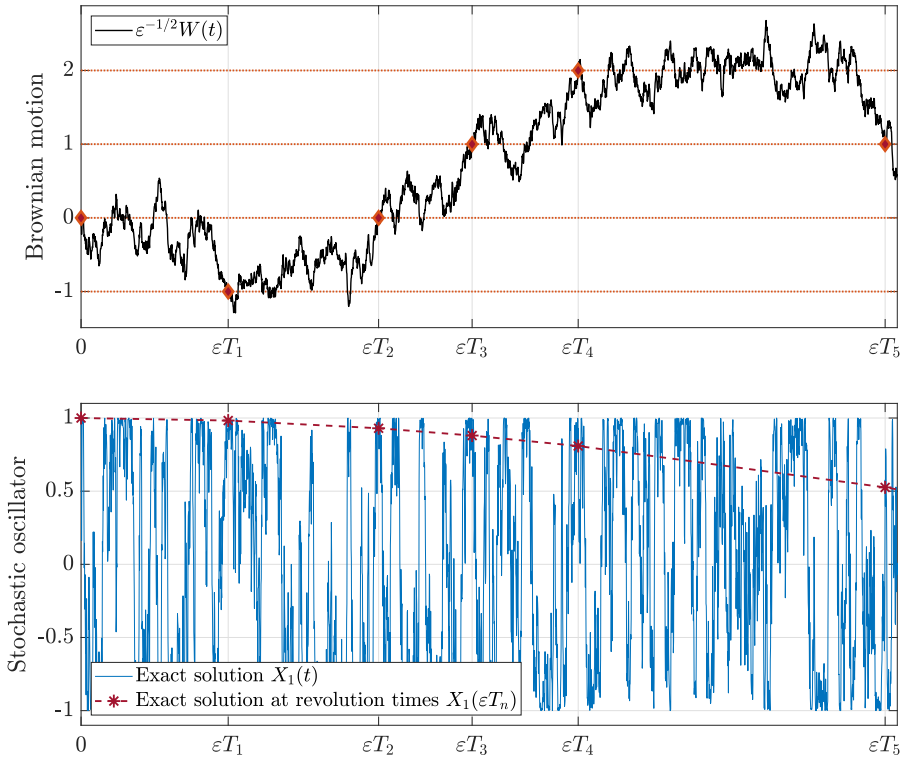


Figure 5.2: Revolution times (5.2.2) of a Brownian path (top) and exact solution evaluated at revolution times for the Kubo oscillator (5.1.1) with $a = 1$ and $\varepsilon = 10^{-1}$ (bottom).

require to simulate W due to the use of appropriate discrete random variables. This will be explained in Section 5.3.4.

We show in Section 5.3.3 that the solution X of (2.1.5) evaluated at times $\varepsilon T_{t\varepsilon^{-1}}$ (when t/ε is an integer) converges weakly when $\varepsilon \to 0$ to the solution y_t of the deterministic ODE

$$\frac{dy_t}{dt} = \langle g^0 \rangle(y_t), \; y_0 = X_0,$$

where $g^0_\theta(y) = e^{-A\theta}F(e^{A\theta}y)$ and $\langle g^0 \rangle := \int_0^1 g^0_\theta d\theta$. This ODE is exactly the same as the asymptotic model for deterministic oscillators of the form (5.2.1). This asymptotic model naturally yields a weak order 1 deterministic integrator. We propose the two following new multirevolution methods of second weak order for integrating equation (2.1.5) at the revolution times εT_{Nm} for $m = 0, 1, 2, \ldots$ with cost in $H = N\varepsilon = \mathcal{O}(1)$ independent of ε. Method B is a geometric modification of Method A to preserve quadratic invariants of the form $Q(y) = \frac{1}{2}y^T S y$ where $S \in \mathbb{R}^{d\times d}$ is a given symmetric matrix. Methods A and B involve a Fourier decomposition of the following functions that are 1-periodic with respect to θ,

$$g^0_\theta(y) = e^{-A\theta}F(e^{A\theta}y) = \sum_{k\in\mathbb{Z}} c^0_k(y)e^{2i\pi k\theta} \tag{5.2.3}$$
$$g^1_\theta(y)(z) = e^{-A\theta}F'(e^{A\theta}y)(e^{A\theta}z) = \sum_{p\in\mathbb{Z}} c^1_p(y)(z)e^{2i\pi p\theta}$$

with Fourier coefficients $(c_k^0(y))_{k\in\mathbb{Z}}$ and $(c_p^1(y))_{p\in\mathbb{Z}}$, respectively. The series appearing in (5.2.3) have an infinite number of terms in general. For a practical implementation of the new methods, we truncate these series up to an even number of modes K_t, while inducing an exponentially small error (see Remark 5.4.3). For each timestep, we also introduce the bounded discrete random variables $(\widehat{\alpha}_k^N)_k$, and deterministic sequences $(\widehat{\beta}_{p,k}^N)_{p,k}$ and $(\widehat{\widehat{\beta}}_{p,k}^N)_{p,k}$ that satisfy

$$\mathbb{E}[\widehat{\alpha}_k^N] = \begin{cases} 1 \text{ if } k=0, \\ 0 \text{ else,} \end{cases} \qquad \mathbb{E}[\widehat{\alpha}_p^N \widehat{\alpha}_k^N] = \begin{cases} 1+\frac{2}{3N} \text{ if } p=k=0, \\ \frac{1}{\pi^2 p^2 N} \text{ if } p+k=0,\ p,k\neq 0, \\ 0 \text{ else,} \end{cases}$$

$$\widehat{\beta}_{p,k}^N = \begin{cases} \frac{1}{2}+\frac{1}{3N} \text{ if } p=k=0, \\ \frac{1}{2\pi^2 k^2 N} \text{ if } p=0,\ k\neq 0, \\ \frac{-1}{2\pi^2 p^2 N} \text{ if } p\neq 0,\ k=0, \\ \frac{1}{2\pi^2 p^2 N} \text{ if } p+k=0,\ p,k\neq 0, \\ 0 \text{ else,} \end{cases} \qquad \widehat{\widehat{\beta}}_{p,k}^N = \begin{cases} \frac{1}{2\pi^2 k^2 N} \text{ if } p=0,\ k\neq 0, \\ \frac{-1}{2\pi^2 p^2 N} \text{ if } p\neq 0,\ k=0, \\ 0 \text{ else.} \end{cases}$$

The definition and construction of these random variables are further discussed in Sections 5.3.2 and 5.3.4.

Method A (Explicit integrator of weak order two in $H = N\varepsilon$ to approximate the solution of equation (2.1.5) at times εT_{Nm} for $m = 0, 1, 2, \ldots$)

$Y_0 = X_0$
for $m \geqslant 0$ **do**

$$Y_{m+1} = Y_m + H \sum_{k=-K_t/2}^{K_t/2-1} c_k^0(Y_m)\widehat{\alpha}_k^N + H^2 \sum_{p,k=-K_t/2}^{K_t/2-1} c_p^1(Y_m)(c_k^0(Y_m))\widehat{\beta}_{p,k}^N \tag{5.2.4}$$

end for

Method B (Geometric integrator of weak order two in $H = N\varepsilon$ to approximate the solution of equation (2.1.5) at times εT_{Nm} for $m = 0, 1, 2, \ldots$ while preserving quadratic invariants)

$Y_0 = X_0$
for $m \geqslant 0$ **do**

$$\begin{aligned} Y_{m+1} = Y_m &+ H \sum_{k=-K_t/2}^{K_t/2-1} c_k^0\left(\frac{Y_m+Y_{m+1}}{2}\right)\widehat{\alpha}_k^N \\ &+ H^2 \sum_{p,k=-K_t/2}^{K_t/2-1} c_p^1\left(\frac{Y_m+Y_{m+1}}{2}\right)\left(c_k^0\left(\frac{Y_m+Y_{m+1}}{2}\right)\right)\widehat{\widehat{\beta}}_{p,k}^N \end{aligned} \tag{5.2.5}$$

end for

Remark 5.2.1. *One could apply a Newton iteration to solve the implicit equation* (5.2.5) *in Method B. However, a few fixed point iterations are sufficient (see discussion in [65, Chap. VIII] for non-stiff implicit methods). Indeed, since the Lipschitz constant of the iterated map has size $\mathcal{O}(H)$, the convergence rate of the fixed point iterations is independent of the smallness of ε.*

Remark 5.2.2. *We observe that* $\widehat{\beta}^N_{p,k}$ *and* $\widehat{\widetilde{\beta}}^N_{p,k}$ *are always zero except when* $p = 0$, $k = 0$ *or* $p + k = 0$. *Thus the computational cost of one step of Methods A and B grows linearly in the number of modes in* (5.2.3).

5.3 Analysis and asymptotic expansion of the exact solution

In this section, we first obtain a local expansion of the solution of (2.1.5) and then evaluate it at particular random times to deal with the highly-oscillatory patterns of the exact solution. Finally we derive from this expansion an asymptotic limit for equation (2.1.5) when $\varepsilon \to 0$.

5.3.1 Asymptotic expansion of the exact solution

Instead of studying directly equation (2.1.5), we apply the change of variable $t \leftarrow \varepsilon^{-1}t$ to obtain the following equation, whose solution satisfies $Y(t) = X(\varepsilon t)$ with X solution of (2.1.5),

$$dY(t) = AY(t) \circ d\widetilde{W}(t) + \varepsilon F(Y(t))dt,\ Y(0) = X_0, \tag{5.3.1}$$

where we denote for simplicity the Brownian motion $\widetilde{W}(t) = \varepsilon^{-1/2}W(\varepsilon t)$ again by W. We introduce the following assumption which guarantees in particular global existence and uniqueness of the solution.

Assumption 5.3.1. *The function* F *is globally Lipschitz continuous and lies in* $\mathcal{C}^3_P$, *i.e., there exists constants* L, C, $K > 0$ *such that for all* y, y_1, $y_2 \in \mathbb{R}^d$

$$|F(y_1) - F(y_2)| \leqslant L\,|y_1 - y_2| \qquad \left|F^{(i)}(y)\right| \leqslant C(1 + |y|^K),\ i \in \{0, 1, 2, 3\}. \tag{5.3.2}$$

Also the initial condition X_0 *has bounded moments, that is,* $\mathbb{E}[|X_0|^p] < \infty$ *for* $p \geqslant 0$.

Therefore we denote $\varphi_{\varepsilon,t}(X_0) = Y(t)$ the solution of equation (5.3.1) and focus in the rest of the chapter on the approximation of $\varphi_{\varepsilon,t}(y)$ at times $t = \mathcal{O}(\varepsilon^{-1})$. The variation of constants formula yields

$$\varphi_{\varepsilon,t}(y) = e^{AW(t)}y + \varepsilon \int_0^t e^{A(W(t)-W(s))}F(\varphi_{\varepsilon,s}(y))ds. \tag{5.3.3}$$

We deduce the following regularity properties.

Lemma 5.3.2. *Under Assumption 5.3.1, the following estimates hold for all* y, y_1, $y_2 \in \mathbb{R}^d$, *where* C *and* K *are independent of* ε *and* t,

1. $|\varphi_{\varepsilon,t}(y_1) - \varphi_{\varepsilon,t}(y_2)| \leqslant C\,|y_1 - y_2|\,e^{C\varepsilon t}$,
2. $|\varphi_{\varepsilon,t}(y)| \leqslant C(1 + |y|)e^{C\varepsilon t}$,
3. $\varphi_{\varepsilon,t}(y)$ *is* $\mathcal{C}^3$ *in* y *and* $\left|\varphi^{(i)}_{\varepsilon,t}(y)\right| \leqslant C(\varepsilon t)^{i-1}(1 + |y|^K)e^{C\varepsilon t}$ *for* $i \in 1, 2, 3$.

Proof. 1. First, $\varphi_{\varepsilon,t}(y)$ is the solution of

$$\varphi_{\varepsilon,t}(y) = e^{AW(t)}y + \varepsilon e^{AW(t)} \int_0^t e^{-AW(s)}F(\varphi_{\varepsilon,s}(y))ds.$$

Using the boundedness of the continuous periodic function $\theta \to e^{\theta A}$ and Assumption 5.3.1, we get

$$|\varphi_{\varepsilon,t}(y_1) - \varphi_{\varepsilon,t}(y_2)| \leqslant |y| + L\varepsilon \int_0^t |\varphi_{\varepsilon,s}(y_1) - \varphi_{\varepsilon,s}(y_2)|\, ds.$$

The Gronwall lemma yields the desired bound.

2. Straightforward using previous statement.

3. Differentiating the integral formulation defining $\varphi_{\varepsilon,t}(y)$ gives

$$\partial_y \varphi_{\varepsilon,t}(y)(h) = e^{AW(t)}h + \varepsilon e^{AW(t)} \int_0^t e^{-AW(s)} F'(\varphi_{\varepsilon,s}(y))(\partial_y \varphi_{\varepsilon,s}(y)(h))ds.$$

Then Assumption 5.3.1 yields

$$|\partial_y \varphi_{\varepsilon,t}(y)(h)| \leqslant |h| + L\varepsilon \int_0^t |\partial_y \varphi_{\varepsilon,t}(y)(h)|\, ds.$$

The Gronwall lemma allows us to obtain

$$|\partial_y \varphi_{\varepsilon,t}(y)| \leqslant e^{L\varepsilon t}.$$

For the second derivative, we get

$$\begin{aligned}\partial_y^2 \varphi_{\varepsilon,t}(y)(h,k) = \varepsilon e^{AW(t)} \int_0^t e^{-AW(s)} [&F'(\varphi_{\varepsilon,s}(y))(\partial_y^2 \varphi_{\varepsilon,s}(y)(h,k)) \\ &+ F''(\varphi_{\varepsilon,s}(y))(\partial_y \varphi_{\varepsilon,s}(y)(h), \partial_y \varphi_{\varepsilon,s}(y)(k))]ds.\end{aligned}$$

Then

$$\begin{aligned}\left|\partial_y^2 \varphi_{\varepsilon,t}(y)(h,k)\right| &\leqslant C\varepsilon \int_0^t [(1 + |\varphi_{\varepsilon,s}(y)|^K)\, |\partial_y \varphi_{\varepsilon,s}(y)(h)|\, |\partial_y \varphi_{\varepsilon,s}(y)(k)| \\ &\quad + \left|\partial_y^2 \varphi_{\varepsilon,s}(y)(h,k)\right|]ds \\ &\leqslant C\varepsilon t(1 + |y|^K) e^{C\varepsilon t}\, |h|\, |k| + C\varepsilon \int_0^t \left|\partial_y^2 \varphi_{\varepsilon,s}(y)(h,k)\right|.\end{aligned}$$

Then the Gronwall lemma allows us to conclude. The proof is similar for the third derivative.

□

Using a local expansion of the solution of (5.3.1) in ε, we define the following first and second order approximations of $\varphi_{\varepsilon,t}(y)$,

$$\psi^1_{\varepsilon,t}(y) = e^{AW(t)}y + \varepsilon e^{AW(t)} \int_0^t e^{-AW(s)} F(e^{AW(s)}y)ds \tag{5.3.4}$$

$$\begin{aligned}\psi^2_{\varepsilon,t}(y) = \psi^1_{\varepsilon,t}(y) &+ \varepsilon^2 e^{AW(t)} \int_0^t e^{-AW(s)} \\ &\cdot F'(e^{AW(s)}y) \left(e^{AW(s)} \int_0^s e^{-AW(r)} F(e^{AW(r)}y)dr \right) ds.\end{aligned}$$

Proposition 5.3.3 (local expansion)**.** *Under Assumption 5.3.1, for all $y \in \mathbb{R}^d$, $j \in \{1,2\}$ and $t \geqslant 0$, there exist C and K two positive constants independent of ε and t such that*

$$\left|\varphi_{\varepsilon,t}(y) - \psi^j_{\varepsilon,t}(y)\right| \leqslant C(1+|y|^K)e^{C\varepsilon t}(\varepsilon t)^{j+1}.$$

The functions $\psi^j_{\varepsilon,t}$ satisfy the following straightforward inequalities proved with similar arguments as for Lemma 5.3.2.

Lemma 5.3.4. *With the assumptions and notations of Proposition 5.3.3, the following estimates hold for all $y \in \mathbb{R}^d$, where C and K are independent of ε and t,*

$$\left|\psi^1_{\varepsilon,t}(y)\right| \leqslant C(1+|y|)e^{C\varepsilon t}, \tag{5.3.5}$$

$$\left|\psi^2_{\varepsilon,t}(y)\right| \leqslant C(1+|y|^K)e^{C\varepsilon t}, \tag{5.3.6}$$

$$\left|\psi^2_{\varepsilon,t}(y) - e^{AW(t)}y\right| \leqslant C(1+|y|^K)(\varepsilon t)e^{C\varepsilon t}. \tag{5.3.7}$$

Proof of Proposition 5.3.3. Using Assumption 5.3.1, we get

$$\left|\varphi_{\varepsilon,t}(y) - \psi^1_{\varepsilon,t}(y)\right| \leqslant L\varepsilon \int_0^t \left|\varphi_{\varepsilon,s}(y) - e^{AW(s)}y\right| ds.$$

Then Lemma 5.3.2 yields

$$\begin{aligned}\left|\varphi_{\varepsilon,s}(y) - e^{AW(s)}y\right| &\leqslant C\varepsilon \int_0^s |F(\varphi_{\varepsilon,r}(y))|\, dr \leqslant C\varepsilon \int_0^s (1+|\varphi_{\varepsilon,r}(y)|)dr \\ &\leqslant C\varepsilon \int_0^s (1+C(1+|y|)e^{C\varepsilon r})dr \leqslant C(1+|y|)e^{C\varepsilon s}(\varepsilon s).\end{aligned}$$

We deduce $\left|\varphi_{\varepsilon,t}(y) - \psi^1_{\varepsilon,t}(y)\right| \leqslant C(1+|y|)e^{C\varepsilon t}(\varepsilon t)^2$.

For $j=2$, we first denote

$$\widetilde{\psi}^2_{\varepsilon,t}(y) = e^{AW(t)}y + \varepsilon e^{AW(t)} \int_0^t e^{-AW(s)}F(\psi^1_{\varepsilon,s}(y))ds.$$

With the same arguments we used for $j=1$ and inequality (5.3.5), we have

$$\left|\varphi_{\varepsilon,t}(y) - \widetilde{\psi}^2_{\varepsilon,t}(y)\right| \leqslant C(1+|y|^K)e^{C\varepsilon t}(\varepsilon t)^3.$$

It is sufficient to prove that $\left|\psi^2_{\varepsilon,t}(y) - \widetilde{\psi}^2_{\varepsilon,t}(y)\right| \leqslant C(1+|y|^K)e^{C\varepsilon t}(\varepsilon t)^3$. A Taylor expansion of $F(\psi^1_{\varepsilon,s}(y))$ in ε gives

$$F(\psi^1_{\varepsilon,s}(y)) = F(e^{AW(s)}y) + \varepsilon F'(e^{AW(s)}y)\left(e^{AW(s)}\int_0^s e^{-AW(r)}F(e^{AW(r)}y)dr\right) + R_{\varepsilon,s}.$$

The remainder $R_{\varepsilon,s}$ satisfies

$$|R_{\varepsilon,s}| \leqslant C\varepsilon^2 \sup_{x\in[e^{AW(s)}y,\psi^1_{\varepsilon,s}(y)]} \left\|F''(x)\right\| \left|e^{AW(s)}\int_0^s e^{-AW(r)}F(e^{AW(r)}y)dr\right|^2.$$

Then, using the polynomial growth of F'' and inequality (5.3.5), we get

$$|R_{\varepsilon,s}| \leqslant C(1+\left|e^{AW(s)}y\right|^K + \left|\psi^1_{\varepsilon,s}(y)\right|^K)(\varepsilon s)^2 e^{C\varepsilon s} \leqslant C(1+|y|^K)(\varepsilon s)^2 e^{C\varepsilon s}.$$

Hence the result. □

We shall prove in Section 5.3.3 that the function $\psi^2_{\varepsilon,t}$ in (5.3.4) evaluated at the revolution times T_N (defined in (5.2.2)) yields a strong order 2 approximation in $H = \varepsilon N$.

Remark 5.3.5. *If we replace the Brownian motion W in* (5.3.4) *by a piecewise linear function W_τ defined by*

$$W_\tau = \left(1 - \frac{t}{\tau} + i\right) W_i + \left(\frac{t}{\tau} - i\right) W_{i+1} \text{ for } i\tau \leqslant t \leqslant (i+1)\tau, \tag{5.3.8}$$

where $W_0 = 0$ and $W_{i+1} = Wi + \sqrt{\tau}\xi_i$ with $(\xi_i)_i$ a family of independent standard Gaussian random variables, then it can be shown that we obtain an integrator of strong order two in εt. However, the cost of a standard method computing an approximation of the integrals of equation (5.3.4) *by replacing W with W_τ is in $\mathcal{O}(t^2/\tau^2)$, which makes this method tremendously expensive for $t = \mathcal{O}(\varepsilon^{-1})$. This is why we develop in Section 5.3.4 weak integrators based on a weak approximation of equation* (5.3.4) *with a cost independent of t. We shall replace stochastic integrals by appropriate discrete random variables in order not to simulate any expensive Brownian path W.*

5.3.2 Properties of the revolution times

In this section, we study some properties linked to the revolution times T_N that will be useful for the analysis.

Proposition 5.3.6. *The revolution times T_N defined in* (5.2.2) *are positive and finite almost surely. Their differences $(T_{N+1} - T_N)_{N\geqslant 0}$ are independent identically distributed random variables (same law as T_1). The Laplace transform $\mathbb{E}[e^{zT_1}]$ of T_1 exists and is analytic for $\mathrm{Re}(z) < \frac{\pi^2}{8}$. In addition, for $x \in \left[0, \frac{\pi^2}{8}\right[$, $\mathbb{E}[e^{xT_1}] = \frac{1}{\cos(\sqrt{2x})}$. The variable T_1 has bounded moments, and they are given by*

$$\mathbb{E}[T_1^k] = \frac{(-2)^k k!}{(2k)!} \sum_{j=1}^{p} (-1)^j \sum_{\substack{n_1+\cdots+n_j=p \\ n_i \in \mathbb{N}^*}} \binom{2p}{2n_1, \ldots, 2n_j}. \tag{5.3.9}$$

In particular, $\mathbb{E}[T_1] = 1$, $\mathbb{E}[T_1^2] = \frac{5}{3}$, and $\mathrm{Var}(T_1) = \frac{2}{3}$. Finally, for a fixed $\varepsilon_0 \in \left]0, \frac{\pi^2}{16}\right[$, for all $\varepsilon \in]0, \varepsilon_0]$ and $p \geqslant 0$, we have the estimate

$$\mathbb{E}[e^{\varepsilon T_N} (\varepsilon T_N)^p] \leqslant C e^{C\varepsilon N} (\varepsilon N)^p. \tag{5.3.10}$$

The law of the first revolution time T_1 has an analytic density, but there is no closed formula for it. It can be numerically approximated accurately by inverting the Laplace transform. In Figure 5.3, we observe the convergence in law of T_N to a Gaussian variable according to the central limit theorem.

Proof of Proposition 5.3.6. The first properties can be deduced from [109, Chap. 2.3], where the Laplace transform formula is obtained with an analytic continuation of the equality $\mathbb{E}[e^{-xT_1}] = \frac{1}{\cosh(\sqrt{2x})}$ for $x > 0$. Comparing the Taylor expansions of $\mathbb{E}[e^{xT_1}]$ and $\frac{1}{\cos(\sqrt{2x})}$ yields (5.3.9). The estimate (5.3.10) is proved as follows

$$\begin{aligned}\mathbb{E}[e^{\varepsilon T_N} (\varepsilon T_N)^p] &\leqslant \mathbb{E}[e^{2\varepsilon T_N}]^{1/2} \mathbb{E}[(\varepsilon T_N)^{2p}]^{1/2} = \mathbb{E}[e^{2\varepsilon T_1}]^{N/2} \varepsilon^p \mathbb{E}[T_N^{2p}]^{1/2} \\ &\leqslant \mathbb{E}[e^{2\varepsilon_0 T_1}]^{\varepsilon N/2\varepsilon_0} (\varepsilon N)^p \mathbb{E}[T_1^{2p}]^{1/2} \leqslant C e^{C\varepsilon N} (\varepsilon N)^p,\end{aligned}$$

where we used first the Cauchy-Schwarz inequality and then twice the Jensen inequality. □

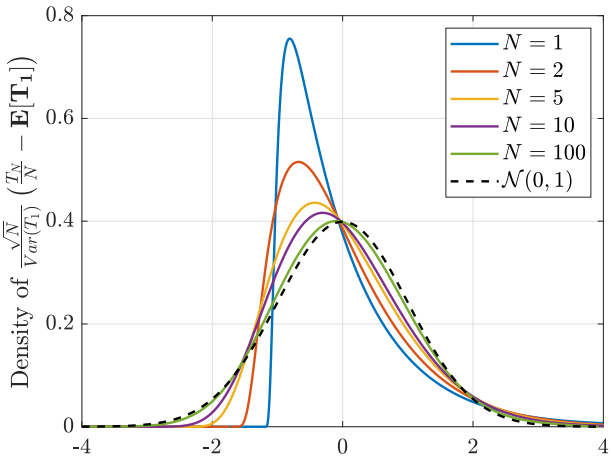


Figure 5.3: Convergence in law of $\frac{\sqrt{N}}{Var(T_1)}\left(\frac{T_N}{N} - \mathbb{E}[T_1]\right)$ to a standard Gaussian random variable.

For developing an algorithm for the weak error, it is useful to know the moments of the random variables appearing in the discretization, that are costly to simulate numerically, in order to replace them with cheap discrete approximations with the same first and second moments. This is the goal of the following proposition.

Proposition 5.3.7. *The random variables*

$$\begin{aligned}
\alpha_k^N &= \tfrac{1}{N}\textstyle\int_0^{T_N} e^{2i\pi kW(s)}ds,\\
\beta_{p,k}^N &= \tfrac{1}{N^2}\textstyle\int_0^{T_N} e^{2i\pi pW(s)}\int_0^s e^{2i\pi kW(r)}drds,\\
\widetilde{\beta}_{p,k}^N &= \tfrac{1}{N^2}\textstyle\int_0^{T_N} e^{2i\pi pW(s)}\left(\int_0^s e^{2i\pi kW(r)} - \tfrac{1}{2}\int_0^{T_N} e^{2i\pi kW(r)}dr\right)ds = \beta_{p,k}^N - \frac{\alpha_p^N\alpha_k^N}{2}
\end{aligned}$$

satisfy

$$\mathbb{E}[\alpha_k^N] = \delta_k = \begin{cases} 1 \text{ if } k = 0,\\ 0 \text{ else,}\end{cases}$$

$$\mathbb{E}[\beta_{p,k}^N] = \begin{cases} \frac{1}{2} + \frac{1}{3N} \text{ if } p = k = 0,\\ \frac{1}{2\pi^2k^2N} \text{ if } p = 0,\ k \neq 0,\\ \frac{-1}{2\pi^2p^2N} \text{ if } p \neq 0,\ k = 0,\\ \frac{1}{2\pi^2p^2N} \text{ if } p + k = 0,\ p,k \neq 0,\\ 0 \text{ else,}\end{cases}$$

$$\mathbb{E}[\alpha_p^N\alpha_k^N] = \begin{cases} 1 + \frac{2}{3N} \text{ if } p = k = 0,\\ \frac{1}{\pi^2p^2N} \text{ if } p + k = 0,\ p,k \neq 0,\\ 0 \text{ else,}\end{cases}$$

$$\mathbb{E}[\widetilde{\beta}_{p,k}^N] = \begin{cases} \frac{1}{2\pi^2k^2N} \text{ if } p = 0,\ k \neq 0,\\ \frac{-1}{2\pi^2p^2N} \text{ if } p \neq 0,\ k = 0,\\ 0 \text{ else.}\end{cases}$$

Proof. Let $k \neq 0$ (the case $k = 0$ is straightforward using Proposition 5.3.6), then the Itô formula applied to $e^{2i\pi kW(s)}$ gives

$$\frac{1}{2\pi^2k^2N}(e^{2i\pi kW(t)} - 1) = \frac{i}{\pi kN}\int_0^t e^{2i\pi kW(s)}dW(s) - \frac{1}{N}\int_0^t e^{2i\pi kW(s)}ds,$$

which yields at time $t = T_N$,

$$\alpha_k^N = \frac{i}{\pi kN}\int_0^{T_N} e^{2i\pi kW(s)}dW(s).$$

Then $t \mapsto \int_0^t e^{2i\pi kW(s)}dW(s)$ is a martingale, so by the Doob theorem, as $t \wedge T_N$ is finite, $\mathbb{E}[\int_0^{t\wedge T_N} e^{2i\pi kW(s)}dW(s)] = 0$ for all t. The dominated convergence theorem for stochastic integrals allows us to take the limit $t \to \infty$ and yields $\mathbb{E}[\alpha_k^N] = 0$.

For the coefficients $\beta_{p,k}^N$, let $(p,k) \neq (0,0)$ (the case $p = k = 0$ is obtained straightforwardly using Proposition 5.3.6), we use the Itô formula on $e^{2i\pi pW(s)}$ and we integrate from s to T_N,

$$1 - e^{2i\pi pW(s)} = 2i\pi p \int_s^{T_N} e^{2i\pi pW(r)}dW(r) - 2\pi^2 p^2 \int_s^{T_N} e^{2i\pi pW(r)}dr.$$

Then, multiplying by $\frac{1}{N^2}e^{2i\pi kW(s)}$ and integrating from 0 to T_N yields

$$\begin{aligned}\frac{\alpha_k^N - \alpha_{p+k}^N}{N} &= \frac{2i\pi p}{N^2}\int_0^{T_N} e^{2i\pi kW(s)}\int_s^{T_N} e^{2i\pi pW(r)}dW(r)ds \\ &\quad - \frac{2\pi^2 p^2}{N^2}\int_0^{T_N} e^{2i\pi kW(s)}\int_s^{T_N} e^{2i\pi pW(r)}drds.\end{aligned}$$

Using the stochastic Fubini theorem, we deduce

$$\int_0^{T_N} e^{2i\pi kW(s)}\int_s^{T_N} e^{2i\pi pW(r)}dW(r)ds = \int_0^{T_N} e^{2i\pi pW(r)}\int_0^r e^{2i\pi kW(s)}dsdW(r),$$

which has zero average by the same arguments as before. Also by the Fubini theorem for stochastic integrals, $\beta_{p,k}^N = \frac{1}{N^2}\int_0^{T_N} e^{2i\pi kW(s)}\int_s^{T_N} e^{2i\pi pW(r)}drds$, so that we get if $p \neq 0$,

$$\mathbb{E}[\beta_{p,k}^N] = \frac{\delta_{p+k} - \delta_k}{2\pi^2 p^2 N}.$$

The case $p = 0$ is obtained by integrating by parts and using the same arguments. Indeed, we find

$$\beta_{0,k}^N = \frac{T_N}{N}\alpha_k^N - \beta_{k,0}^N = \frac{i}{\pi k N^2}\int_0^{T_N} T_N e^{2i\pi kW(s)}dW(s) - \beta_{k,0}^N,$$

and $\mathbb{E}[\beta_{0,k}^N] = -\mathbb{E}[\beta_{k,0}^N]$. Finally, the moments $\mathbb{E}[\alpha_p^N\alpha_k^N]$ are computed via the equality $\beta_{p,k}^N + \beta_{k,p}^N = \alpha_p^N\alpha_k^N$. Then we obtain $\mathbb{E}[\widetilde{\beta}_{p,k}^N]$ from the formula $\widetilde{\beta}_{p,k}^N = \beta_{p,k}^N - \frac{\alpha_p^N\alpha_k^N}{2}$. □

Remark 5.3.8 (stochastic Fourier series)**.** *Let f be a L^2 function on $]0,1[$ extended on $\mathbb{R}$ by 1-periodicity, whose Fourier coefficients are denoted as $(c_k)_{k\in\mathbb{Z}}$. Then we deduce from Proposition 5.3.7 the following equalities, where the second one is the stochastic version of the Bessel-Parseval theorem,*

$$\begin{aligned}\mathbb{E}\left[\int_0^{T_1} f(W(s))ds\right] &= c_0 = \int_0^1 f(\theta)d\theta, \\ \mathbb{E}\left[\int_0^{T_1} |f(W(s))|^2 ds\right] &= \sum_{k\in\mathbb{Z}} |c_k|^2, \\ \mathbb{E}\left[\left|\int_0^{T_1} f(W(s))ds\right|^2\right] &= \frac{5|c_0|^2}{3} + \sum_{k\in\mathbb{Z}^*}\frac{|c_k|^2}{\pi^2 k^2}.\end{aligned}$$

5.3.3 Asymptotic expansion at revolution times and limit model

With the results of Subsection 5.3.2, it is now possible to evaluate the local expansions (5.3.4) at revolution times. To approximate numerically the integrals appearing in equation (5.3.4) without evaluating F and F' too many times, we first replace the 1-periodic functions $g^0_\theta(y)$ and $g^1_\theta(y)(z)$ defined in equation (5.2.3) by their associated Fourier series with the Fourier coefficients $(c^0_k(y))_{k\in\mathbb{Z}}$ and $(c^1_p(y))_{p\in\mathbb{Z}}$. We define the following approximation of $\varphi_{\varepsilon,t}(y)$,

$$\begin{aligned}\psi_{\varepsilon,t}(y) &= e^{AW(t)}y + \varepsilon\sum_{k\in\mathbb{Z}} e^{AW(t)}c^0_k(y)\int_0^t e^{2i\pi kW(s)}ds \\ &\quad + \varepsilon^2\sum_{p,k\in\mathbb{Z}} e^{AW(t)}c^1_p(y)\left(c^0_k(y)\int_0^t\int_0^s e^{2i\pi pW(s)}e^{2i\pi kW(r)}drds\right).\end{aligned} \tag{5.3.11}$$

Notice that $c^0_k(y)\in\mathbb{C}^d$ and $c^1_p(y) = (c^0_k)'(y)\in\mathbb{C}^{d\times d}$ but $\psi_{\varepsilon,t}(y)\in\mathbb{R}^d$. We now evaluate this function $\psi_{\varepsilon,t}(y)$ at time $t = T_N$ to get a second order strong approximation.

Proposition 5.3.9. *We define the following quantity*

$$\psi_{\varepsilon,N}(y) = y + H\sum_{k\in\mathbb{Z}} c^0_k(y)\alpha^N_k + H^2\sum_{p,k\in\mathbb{Z}} c^1_p(y)(c^0_k(y))\beta^N_{p,k},$$

where $(c^0_k(y))_{k\in\mathbb{Z}}$ and $(c^1_p(y))_{p\in\mathbb{Z}}$ are the Fourier coefficients of the 1-periodic functions $g^0_\theta(y)$ and $g^1_\theta(y)$ defined in (5.2.3)*, α^N_k, $\beta^N_{p,k}$ are the random variables defined in Proposition 5.3.7 and $y\in\mathbb{R}^d$ is deterministic. Under Assumption 5.3.1, for all test functions $\phi\in\mathcal{C}^3_P$, there exists $H_0>0$ such that for all $H = N\varepsilon\leqslant H_0$, the following estimates hold, where C and K are independent of ε and N,*

$$\mathbb{E}\left[|\varphi_{\varepsilon,T_N}(y) - \psi_{\varepsilon,N}(y)|^2\right]^{1/2} \leqslant C(1+|y|^K)H^3, \tag{5.3.12}$$

$$|\mathbb{E}[\phi(\varphi_{\varepsilon,T_N}(y))] - \mathbb{E}[\phi(\psi_{\varepsilon,N}(y))]| \leqslant C(1+|y|^K)H^3, \tag{5.3.13}$$

that is, $\psi_{\varepsilon,N}(y)$ is a numerical approximation of $\varphi_{\varepsilon,T_N}(y)$ of strong/weak local order two.

Proof. Inequality (5.3.12) is a straightforward consequence of Proposition 5.3.6 when evaluating the estimates of Proposition 5.3.3 at time T_N. For the weak local estimate (5.3.13), using inequality (5.3.12), the mean value inequality, Lemma 5.3.2 and equations (5.3.5) and (5.3.6), we get

$$\begin{aligned}&\mathbb{E}\Big[\Big|\phi(\varphi_{\varepsilon,T_N}(y)) - \phi(\psi^j_{\varepsilon,T_N}(y))\Big|\Big] \\ &\qquad\leqslant \mathbb{E}\left[\sup_{x\in[\varphi_{\varepsilon,T_N}(y),\psi^j_{\varepsilon,T_N}(y)]}\big|\phi'(x)\big|\Big|\varphi_{\varepsilon,T_N}(y) - \psi^j_{\varepsilon,T_N}(y)\Big|\right] \\ &\qquad\leqslant C(1+|y|^K)\mathbb{E}\left[e^{C\varepsilon T_N}(\varepsilon T_N)^{j+1}\sup_{x\in[\varphi_{\varepsilon,T_N}(y),\psi^j_{\varepsilon,T_N}(y)]}(1+|x|^p)\right] \\ &\qquad\leqslant C(1+|y|^K)\mathbb{E}\left[(\varepsilon T_N)^{j+1}e^{C\varepsilon T_N}\left(1+|\varphi_{\varepsilon,T_N}(y)|^p + \Big|\psi^j_{\varepsilon,T_N}(y)\Big|^p\right)\right] \\ &\qquad\leqslant C(1+|y|^K)\mathbb{E}\left[(\varepsilon T_N)^{j+1}e^{C\varepsilon T_N}\right].\end{aligned}$$

Finally we obtain inequality (5.3.13) by taking H small enough so that we can apply Proposition 5.3.6. □

For a fixed $T = N\varepsilon$, when $\varepsilon \to 0$ (or equivalently $N \to \infty$), the solution of (5.3.1) evaluated at stroboscopic times $T_N = T_{T\varepsilon^{-1}}$ converges weakly to the solution of a deterministic ODE that involves only the first mode $c_0^0 = \langle g^0 \rangle = \int_0^1 g_\theta^0 d\theta$ of g^0. This asymptotic model is the same one as for the deterministic equation (5.2.1). The proof is postponed to Subsection 5.4.3.

Proposition 5.3.10 (asymptotic model)**.** *Under Assumption 5.3.1, for $T > 0$, the solution $\varphi_{\varepsilon,T_{T\varepsilon^{-1}}}(X_0)$ (for ε such that $T\varepsilon^{-1}$ is an integer) of equation* (2.1.5) *converges weakly when $\varepsilon \to 0$ to the solution at time T of*

$$\frac{dy_t}{dt} = \langle g^0 \rangle(y_t),\ y_0 = X_0, \tag{5.3.14}$$

that is, for all test function $\phi \in \mathcal{C}_P^3$,

$$\lim_{\varepsilon \to 0} \left| \mathbb{E}[\phi(\varphi_{\varepsilon,T_{T\varepsilon^{-1}}}(X_0))] - \mathbb{E}[\phi(y_T)] \right| = 0.$$

Remark 5.3.11. *It can be proven using the results of Section 5.4 that the solution of the asymptotic model (Proposition 5.3.10) is an order one weak approximation of $X(\varepsilon T_{Nm})$ for $m \geqslant 0$ and X solution of equation* (2.1.5)*. We deduce the following simple one-step explicit deterministic integrator that corresponds to the Euler method applied to equation* (5.3.14)*,*

$$y_0 = X_0,\ y_{m+1} = y_m + H c_0^0(y_m). \tag{5.3.15}$$

Its cost is independent of ε and N, and it has weak order one w.r.t. H, that is, for all $m \geqslant 0$, we have $\mathbb{E}[\phi(\varphi_{\varepsilon,T_{Nm}}(X_0))] - \mathbb{E}[\phi(y_m)] = \mathcal{O}(H)$.

5.3.4 Construction of the second order integrators

To obtain an integrator of weak order two with a cost independent of ε and N, we truncate the local expansion of Proposition 5.3.9. We also replace the involved random variables with cheap discrete random variables with the same first and second moments. To simulate the random variable α_k^N with discrete random variables $\widehat{\alpha}_k^N$ with the same first and second moments, we introduce a set $(\xi_k)_{k\in\mathbb{N}}$ of independent random variables such that $\mathbb{P}(\xi_k = \pm 1) = \frac{1}{2}$, the covariance matrix $(C_\alpha^N)_{p,k}$ such that

$$(C_\alpha^N)_{2p-1:2p,2k-1:2k} = \begin{pmatrix} \mathrm{Cov}(\mathrm{Re}(\alpha_p^N), \mathrm{Re}(\alpha_k^N)) & \mathrm{Cov}(\mathrm{Re}(\alpha_p^N), \mathrm{Im}(\alpha_k^N)) \\ \mathrm{Cov}(\mathrm{Im}(\alpha_p^N), \mathrm{Re}(\alpha_k^N)) & \mathrm{Cov}(\mathrm{Im}(\alpha_p^N), \mathrm{Im}(\alpha_k^N)) \end{pmatrix},$$

and Γ^N its square root. Then, $\widehat{\alpha}_k^N$ is defined for $k \geqslant 0$ as

$$\widehat{\alpha}_k^N = \delta_k + \sum_{l\in\mathbb{N}} (\Gamma_{2k-1,l}^N + i\Gamma_{2k,l}^N)\xi_l \quad \text{with} \quad \delta_k = \begin{cases} 1 \text{ if } k = 0 \\ 0 \text{ else} \end{cases}$$

and we fix $\widehat{\alpha}_k^N = \overline{\widehat{\alpha}_{-k}^N}$ for $k < 0$ (so that the solution stays real while still having the good moments). We also define $\widehat{\beta}_{p,k}^N = \mathbb{E}[\beta_{p,k}^N]$ with the values of Proposition 5.3.7. Doing so yields Method A.

For Method B, we adapt Method A in the spirit of the implicit midpoint scheme for ODEs (see [65, Chap. IV]) so that it preserves any quadratic invariant. We also replace $\widehat{\beta}_{p,k}^N$ by $\widetilde{\widehat{\beta}}_{p,k}^N = \mathbb{E}[\widetilde{\beta}_{p,k}^N]$, using the values of Proposition 5.3.7.

Remark 5.3.12. *The methodology presented in Section 5.3 can be generalized to any order. Thus, under more regularity assumptions on F, it is possible to build algorithms similar to Method A of any weak order and that are still robust with respect to the stiffness ε. For order 3, Method A becomes*

$$\begin{aligned} Y_{m+1} = Y_m &+ H\sum_{k\in\mathbb{Z}} c_k^0(Y_m)\widehat{\alpha}_k^N + H^2\sum_{p,k\in\mathbb{Z}} c_p^1(Y_m)(c_k^0(Y_m))\widehat{\beta}_{p,k}^N \\ &+ H^3\sum_{l,p,k\in\mathbb{Z}} c_l^1(c_p^1(c_k^0))(Y_m)\widehat{\gamma}_{l,p,k}^{(1),N} + c_l^2(c_p^0,c_k^0)(Y_m)\widehat{\gamma}_{l,p,k}^{(2),N} \end{aligned}$$

with the new random variables

$$\begin{aligned} \gamma_{l,p,k}^{(1),N} &= \frac{1}{N^2}\int_0^{T_N} e^{2i\pi qW(s)}\int_0^s e^{2i\pi pW(r)}\int_0^r e^{2i\pi kW(q)}dqdrds, \\ \gamma_{l,p,k}^{(2),N} &= \frac{1}{2N^2}\int_0^{T_N} e^{2i\pi qW(s)}\int_0^s e^{2i\pi pW(r)}\int_0^s e^{2i\pi kW(q)}dqdrds, \end{aligned}$$

and where the discrete random variables share the same moments up to order 3 for the $\widehat{\alpha}_k^N$, order 2 for the $\widehat{\beta}_{p,k}^N$, and order 1 for the $\widehat{\gamma}_{l,p,k}^{(i),N}$. It is also possible to generalize Method B up to any order in the spirit of the implicit midpoint scheme, but the construction of discrete random variables allowing the preservation of quadratic invariants is not obvious for higher orders (although backward error analysis guarantees the preservation of quadratic invariants for the exact random variables based on W).

5.4 Weak convergence analysis

We focus in this section on the proofs of the following two theorems, showing the order two convergence of Methods A and B.

Theorem 5.4.1. *Assume that the Fourier coefficients c_k^0, c_p^1 of g_θ^0 and g_θ^1 in (5.2.3) are non-zero only for $-K_t/2 \leqslant k,p < K_t/2$. Then, under Assumption 5.3.1, Method A has weak order two for solving (2.1.5), that is, for all $T>0$, for all test functions $\phi \in \mathcal{C}_P^3$, there exists $H_0>0$ such that for all $H \leqslant H_0$, for all $m \geqslant 0$ such that $mN\varepsilon = mH \leqslant T$, there exist two positive constants K and C both independent of ε, N and K_t such that*

$$|\mathbb{E}[\phi(\varphi_{\varepsilon,T_{Nm}}(X_0))] - \mathbb{E}[\phi(Y_m)]| \leqslant CH^2(1+\mathbb{E}[|X_0|^K]). \tag{5.4.1}$$

Theorem 5.4.2. *Assume that the Fourier coefficients c_k^0, c_p^1 of g_θ^0 and g_θ^1 in (5.2.3) are non-zero only for $-K_t/2 \leqslant k,p < K_t/2$. Then, under Assumption 5.3.1, if $c_0^1(c_k^0)$ and $c_p^1(c_0^0)$ are Lipschitz continuous uniformly in k and p, Method B is well defined and is of weak order two for solving (2.1.5) (i.e., it satisfies an estimate of the form (5.4.1)). In addition, if for a fixed symmetric matrix $S \in \mathbb{R}^{d\times d}$, the quantity $Q(y) = \frac{1}{2}y^TSy$ is preserved by equation (2.1.5), then Method B also preserves the invariant $Q(y) = \frac{1}{2}y^TSy$, that is, the solution Y_{m+1} of equation (5.2.5) satisfies $Q(Y_{m+1}) = Q(Y_m)$.*

These two theorems focus on approximating the exact solution of equation (2.1.5) at the revolution times εT_{Nm}, $m = 0,1,\ldots$, but one could compute an approximation at different times by composing with other methods at the end of the integration.

Remark 5.4.3. *Since the error constant C in* (5.4.1) *is independent of the number K_t of Fourier modes, we emphasize that Theorems 5.4.1 and 5.4.2 remain valid for infinitely many modes ($K_t \to \infty$). In addition, assuming that F is of class C_P^{s+1} yields a truncation error of the Fourier series in* (5.2.3) *of size $\mathcal{O}((1+|y|^K)K_t^{-s})$ (see, e.g., [96, Sect. III.1.3]), and if g_θ^0 is assumed analytic in θ (for example if F is a polynomial), this error becomes exponentially small as $\mathcal{O}((1+|y|)e^{-cK_t})$. For simplicity of the analysis, we thus assumed in Theorems 5.4.1 and 5.4.2 that g_θ^0 and g_θ^1 have a finite number K_t of non-zero Fourier modes in* (5.2.3). *If this assumption does not hold, the truncation errors $\mathcal{O}((1+|y|^K)K_t^{-s})$ or $\mathcal{O}((1+|y|)e^{-cK_t})$ should be added in the right-hand side of the error estimate* (5.4.1). *Let us prove it in the analytic case. We first apply the change of variable $\widetilde{\varphi}_{\varepsilon,t}(y) = e^{-AW(t)}\varphi_{\varepsilon,t}(y)$ that has no effect at time $t = T_{Nm}$. We now have to compare the two solutions of the following integral formulations*

$$\widetilde{\varphi}_{\varepsilon,t}(y) = y + \varepsilon\int_0^t g^0_{W(s)}(\widetilde{\varphi}_{\varepsilon,s}(y))ds,$$
$$\widetilde{\varphi}^{(K_t)}_{\varepsilon,t}(y) = y + \varepsilon\int_0^t \sum_{k=-K_t/2}^{K_t/2-1} c_k^0(\widetilde{\varphi}^{(K_t)}_{\varepsilon,s}(y))e^{2i\pi kW(s)}ds.$$

Using the truncation estimates that we previously discussed and the Lipschitz property of g_θ^0, one gets

$$\left|\widetilde{\varphi}_{\varepsilon,t} - \widetilde{\varphi}^{(K_t)}_{\varepsilon,t}\right|(y) \leqslant \varepsilon\int_0^t \left|g^0_{W(s)}(\widetilde{\varphi}_{\varepsilon,s}(y)) - g^0_{W(s)}(\widetilde{\varphi}^{(K_t)}_{\varepsilon,s}(y))\right|ds + C\varepsilon t e^{-cK_t}\sup_{[0,1]}|g^0(y)|$$
$$\leqslant C\varepsilon\int_0^t \left|\widetilde{\varphi}_{\varepsilon,s}(y) - \widetilde{\varphi}^{(K_t)}_{\varepsilon,s}(y)\right|ds + C\varepsilon t(1+|y|)e^{-cK_t}.$$

The Gronwall lemma and Proposition 5.3.6 yield for $mN\varepsilon \leqslant T$,

$$\mathbb{E}\left[\left|\widetilde{\varphi}_{\varepsilon,T_{Nm}}(y) - \widetilde{\varphi}^{(K_t)}_{\varepsilon,T_{Nm}}(y)\right|^2\right]^{1/2} \leqslant C(1+|y|)e^{-cK_t}.$$

The structure of the convergence proof is similar to the one for standard SDE integrators, see e.g. [104, Chap. 2], but one has to be cautious because our solution is evaluated at stochastic times and the error constants should not depend on ε or N.

5.4.1 Boundedness of the numerical moments

Proposition 5.4.4 (bounded moments for the integrator (5.2.4))**.** *Assume that for $y \in \mathbb{R}^d$, the numerical integrator $\widehat{\psi}_{\varepsilon,N}(y)$ is given by*

$$\widehat{\psi}_{\varepsilon,N}(y) = y + H\sum_{k\in\mathbb{Z}} c_k^0(y)\widehat{\alpha}_k^N + H^2\sum_{p,k\in\mathbb{Z}} c_p^1(y)(c_k^0(y))\widehat{\beta}^N_{p,k}, \tag{5.4.2}$$

where $\widehat{\alpha}_k^N$, $\widehat{\beta}^N_{p,k}$ are random variables defined such that for all $q > 0$, $\mathbb{E}\left[\left(\sum_k \frac{|\widehat{\alpha}_k^N|^2}{k^2}\right)^q\right]$ and $\mathbb{E}\left[\left(\sum_{p,k} \frac{|\widehat{\beta}^N_{p,k}|^2}{k^2}\right)^q\right]$ are bounded uniformly in N. Then, under Assumption 5.3.1 and if $|y|$ has bounded moments, for any $T > 0$, for all m, H such that $m\varepsilon N = mH \leqslant T$, for all $q > 0$, we have $\mathbb{E}[|\widehat{\psi}^m_{\varepsilon,N}(y)|^{2q}] \leqslant C_q(1+\mathbb{E}[|y|^{2q}])$, where C_q is independent of m, ε and N.

Proof. We first prove

$$\left|\widehat{\psi}_{\varepsilon,N}(y) - y\right| \leqslant CH(1+|y|)M_N, \tag{5.4.3}$$

where $\mathbb{E}[(M_N)^{2q}] \leqslant C_q$ for all $q > 0$. We have

$$\begin{aligned}\left|\widehat{\psi}_{\varepsilon,N}(y) - y\right| &= H\left|\sum_{k\in\mathbb{Z}} c_k^0(y)\widehat{\alpha}_k^N + H\sum_{p,k\in\mathbb{Z}} c_p^1(y)(c_k^0(y))\widehat{\beta}_{p,k}^N\right| \\ &\leqslant CH\left(M_N^{(0)}\sqrt{\sum_{k\in\mathbb{Z}} k^2\left|c_k^0(y)\right|^2} + M_N^{(1)}\sqrt{\sum_{p\in\mathbb{Z}}\left|c_p^1(y)\right|^2}\sqrt{\sum_{k\in\mathbb{Z}} k^2\left|c_k^0(y)\right|^2}\right)\end{aligned}$$

where $M_N^{(0)} = \sqrt{\sum_k \frac{\left|\widehat{\alpha}_k^N\right|^2}{k^2}}$ and $M_N^{(l)} = \sqrt{\sum_{p,k} \frac{\left|\widehat{\beta}_{p,k}^N\right|^2}{k^2}}$ have moments bounded uniformly in N. Then using the Bessel-Parseval theorem, we get $\sum_k k^2\left|c_k^0(y)\right|^2 = \int_0^1 \left|\partial_\theta g_\theta(y)^0\right|^2 d\theta$. Assumption 5.3.1 yields $\left|\partial_\theta g_\theta^0(y)\right| \leqslant C(1+|y|)$. Then, the Bessel-Parseval theorem applied on g_θ^1 gives $\sqrt{\sum_p \left|c_p^1(y)\right|^2} \leqslant C$, hence the result.

We define $\Delta\psi_m = \widehat{\psi}_{\varepsilon,N}^{m+1}(y) - \widehat{\psi}_{\varepsilon,N}^m(y) = (\widehat{\psi}_{\varepsilon,N} - \mathrm{Id})(\widehat{\psi}_{\varepsilon,N}^m(y))$, then

$$(\widehat{\psi}_{\varepsilon,N}^{m+1}(y))^{2q} = (\widehat{\psi}_{\varepsilon,N}^m(y))^{2q} + \sum_{j=1}^{2q}\binom{2l}{j}(\widehat{\psi}_{\varepsilon,N}^m(y))^{2q-j}\Delta\psi_m^j.$$

Equation (5.4.3) and the bounded moments of M_N give

$$\begin{aligned}\left|\mathbb{E}\left[(\widehat{\psi}_{\varepsilon,N}^m(y))^{2q-j}\Delta\psi_m^j\right]\right| &\leqslant \mathbb{E}\left[\left|\widehat{\psi}_{\varepsilon,N}^m(y)\right|^{2q-j} CH^j(1+\left|\widehat{\psi}_{\varepsilon,N}^m(y)\right|^j)M_N^j\right] \\ &\leqslant C_q H\left(1+\mathbb{E}\left[\left|\widehat{\psi}_{\varepsilon,N}^m(y)\right|^{2q}\right]\right).\end{aligned}$$

We deduce

$$1+\mathbb{E}\left[\left|\widehat{\psi}_{\varepsilon,N}^{m+1}(y)\right|^{2q}\right] \leqslant e^{C_q H}\left(1+\mathbb{E}\left[\left|\widehat{\psi}_{\varepsilon,N}^m(y)\right|^{2q}\right]\right)$$

and by induction $\mathbb{E}\left[\left|\widehat{\psi}_{\varepsilon,N}^m(y)\right|^{2q}\right] \leqslant e^{C_q mH}(1+\mathbb{E}[|y|^{2q}]) \leqslant e^{C_q T}(1+\mathbb{E}[|y|^{2q}])$. □

Proposition 5.4.5 (bounded moments for the integrator (5.2.5))**.** *Assume that for $y \in \mathbb{R}^d$, the numerical scheme $\widehat{\psi}_{\varepsilon,N}(y)$ satisfies*

$$\begin{aligned}\widehat{\psi}_{\varepsilon,N}(y) = y &+ H\sum_{k\in\mathbb{Z}} c_k^0\left(\frac{y+\widehat{\psi}_{\varepsilon,N}(y)}{2}\right)\widehat{\alpha}_k^N \\ &+ H^2\sum_{p,k\in\mathbb{Z}} c_p^1\left(\frac{y+\widehat{\psi}_{\varepsilon,N}(y)}{2}\right)\left(c_k^0\left(\frac{y+\widehat{\psi}_{\varepsilon,N}(y)}{2}\right)\right)\widehat{\widehat{\beta}}_{p,k}^N,\end{aligned} \tag{5.4.4}$$

where $\widehat{\alpha}_k^N$, $\widehat{\widehat{\beta}}_{p,k}^N$ are random variables defined such that for all $q > 0$, $\sum_k \left|\widehat{\alpha}_k^N\right|$, $\sum_{p,k}\left|\widehat{\widehat{\beta}}_{p,k}^N\right|$, $\mathbb{E}\left[\left(\sum_k \frac{\left|\widehat{\alpha}_k^N\right|^2}{k^2}\right)^q\right]$ and $\mathbb{E}\left[\left(\sum_{p,k}\frac{\left|\widehat{\widehat{\beta}}_{p,k}^N\right|^2}{k^2}\right)^q\right]$ are bounded uniformly in N. Then, under Assumption 5.3.1 and if $|y|$ has bounded moments, for H_0 small enough and any $T > 0$, for

all m, H such that $m\varepsilon N = mH \leqslant T$ and $H \leqslant H_0$, for all $q > 0$, we have $\mathbb{E}[\left|\widehat{\psi}^m_{\varepsilon,N}(y)\right|^{2q}] \leqslant C_q(1 + \mathbb{E}[|y|^{2q}])$, where C_q is independent of m, ε and N.

Proof. We prove an equivalent of the estimate (5.4.3) for $\widehat{\psi}_{\varepsilon,N}(y)$. The growth properties of the Fourier coefficients yield

$$\begin{aligned}\left|\widehat{\psi}_{\varepsilon,N}(y) - y\right| \leqslant CH &\left(\sum_k \left|c_k^0\right|(y)\left|\widehat{\alpha}_k^N\right| + \sum_{p,k}\left|c_k^0\right|(y)\left|\widehat{\widetilde{\beta}}_{p,k}^N\right|\right)\\ &+ CH\left(\sum_k \left|\widehat{\alpha}_k^N\right| + \sum_{p,k}\left|\widehat{\widetilde{\beta}}_{p,k}^N\right|\right)\left|\widehat{\psi}_{\varepsilon,N}(y) - y\right|,\end{aligned}$$

As $\sum_k \left|\widehat{\alpha}_k^N\right| + \sum_{p,k}\left|\widehat{\widetilde{\beta}}_{p,k}^N\right|$ is bounded, using the same estimates as in the proof of Proposition 5.4.4, we get for all $H \leqslant H_0$ small enough,

$$\left|\widehat{\psi}_{\varepsilon,N}(y) - y\right| \leqslant CH(1 + |y|)M_N, \tag{5.4.5}$$

where M_N has bounded moments. The remaining of the proof is the same as in the proof of Proposition 5.4.4. □

5.4.2 Local weak error

Proposition 5.4.6 (local error estimate)**.** *Assume that for $y \in \mathbb{R}^d$ deterministic, the numerical scheme can be written as*

$$\widehat{\psi}_{\varepsilon,N}(y) = y + H\sum_{k\in\mathbb{Z}} c_k^0(y)\widehat{\alpha}_k^N + H^2\sum_{p,k\in\mathbb{Z}} c_p^1(y)(c_k^0(y))\widehat{\beta}_{p,k}^N + R,$$

where $\mathbb{E}[|R|] \leqslant C(1 + |y|^K)H^3$ and $\widehat{\alpha}_k^N \in \mathbb{C}$, $\widehat{\beta}_{p,k}^N \in \mathbb{R}$ are random variables such that $\widehat{\alpha}_k^N = \overline{\widehat{\alpha}_{-k}^N}$ and

$$\mathbb{E}[\widehat{\alpha}_k^N] = \mathbb{E}[\alpha_k^N],\ \mathbb{E}[\widehat{\beta}_{p,k}^N] = \mathbb{E}[\beta_{p,k}^N],\ \mathbb{E}[\widehat{\alpha}_{k_1}^N\widehat{\alpha}_{k_2}^N] = \mathbb{E}[\alpha_{k_1}^N\alpha_{k_2}^N].$$

Under Assumption 5.3.1, if $\widehat{\psi}_{\varepsilon,N}(y)$ satisfies the assumptions of Proposition 5.4.4 (or Proposition 5.4.5), for all test function $\phi \in \mathcal{C}_P^3$, there exists $H_0 > 0$ such that for all $H = N\varepsilon \leqslant H_0$, the following estimate holds, where C and K are independent of ε and N,

$$\left|\mathbb{E}[\phi(\varphi_{\varepsilon,T_N}(y))] - \mathbb{E}[\phi(\widehat{\psi}_{\varepsilon,N}(y))]\right| \leqslant C(1 + |y|^K)H^3,$$

that is, the numerical scheme has weak local order two.

Proof. Using Proposition 5.3.9 and its notation $\psi_{\varepsilon,N}(y)$, it is enough to prove that

$$\left|\mathbb{E}[\phi(\psi_{\varepsilon,N}(y))] - \mathbb{E}[\phi(\widehat{\psi}_{\varepsilon,N}(y))]\right| \leqslant C(1 + |y|^K)H^3.$$

A local expansion gives

$$\phi(\psi_{\varepsilon,N}(y)) = \phi(y) + \phi'(y)(\psi_{\varepsilon,N}(y) - y) + \phi''(y)(\psi_{\varepsilon,N}(y) - y, \psi_{\varepsilon,N}(y) - y) + R_1.$$

As $\psi_{\varepsilon,N}(y) = \psi^2_{\varepsilon,T_N}(y)$ (see equation (5.3.4)), using inequalities (5.3.6), (5.3.7) evaluated at T_N and Proposition 5.3.6 yields

$$\begin{aligned}
\mathbb{E}[|R_1|] &\leqslant \mathbb{E}\left[\sup_{x\in[y,\psi_{\varepsilon,N}(y)]}\left|\phi^{(3)}(x)\right||\psi_{\varepsilon,N}(y)-y|^3\right]\\
&\leqslant \mathbb{E}\left[C(1+|y|^K+|\psi_{\varepsilon,N}(y)|^K)(1+|y|^K)(\varepsilon T_N)^3 e^{C\varepsilon T_N}\right]\\
&\leqslant \mathbb{E}\left[C(1+|y|^K)(\varepsilon T_N)^3 e^{C\varepsilon T_N}\right]\\
&\leqslant C(1+|y|^K)H^3.
\end{aligned}$$

We obtain a similar expansion for $\phi(\widehat{\psi}_{\varepsilon,N}(y))$:

$$\phi(\widehat{\psi}_{\varepsilon,N}(y)) = \phi(y) + \phi'(y)(\widehat{\psi}_{\varepsilon,N}(y)-y) + \phi''(y)(\widehat{\psi}_{\varepsilon,N}(y)-y, \widehat{\psi}_{\varepsilon,N}(y)-y) + \widehat{R_1},$$

where, using inequality (5.4.3) (or (5.4.5)),

$$\begin{aligned}
\mathbb{E}[|\widehat{R_1}|] &\leqslant \mathbb{E}\left[\sup_{x\in[y,\widehat{\psi}_{\varepsilon,N}(y)]}\left|\phi^{(3)}(x)\right|\left|\widehat{\psi}_{\varepsilon,N}(y)-y\right|^3\right]\\
&\leqslant C\mathbb{E}\left[(1+|y|^K+\left|\widehat{\psi}_{\varepsilon,N}(y)\right|^K)(1+|y|^K)H^3M_N^3\right]\\
&\leqslant C(1+|y|^K)H^3\mathbb{E}\left[(1+M_N^K)\right]\\
&\leqslant C(1+|y|^K)H^3.
\end{aligned}$$

Making the difference of both equations gives

$$\begin{aligned}
\phi(\psi_{\varepsilon,N}(y)) - \phi(\widehat{\psi}_{\varepsilon,N}(y)) &= \phi'(y)(\psi_{\varepsilon,N}(y)-\widehat{\psi}_{\varepsilon,N}(y)) - \phi''(y)(\psi_{\varepsilon,N}(y)-y)^2\\
&\quad + \phi''(y)(\widehat{\psi}_{\varepsilon,N}(y)-y)^2 + R,
\end{aligned} \tag{5.4.6}$$

where $\mathbb{E}[|R|] \leqslant C(1+|y|^K)H^3$. For the first term of (5.4.6), we have

$$\begin{aligned}
\mathbb{E}[\phi'(y)(\psi_{\varepsilon,N}(y)-\widehat{\psi}_{\varepsilon,N}(y))] &= H\sum_{k\in\mathbb{Z}}\mathbb{E}[\phi'(y)(c_k^0(y)(\alpha_k^N-\widehat{\alpha}_k^N))]\\
&\quad + H^2\sum_{p,k\in\mathbb{Z}}\mathbb{E}[\phi'(y)(c_p^1(y)(c_k^0(y))(\beta_{p,k}^N-\widehat{\beta}_{p,k}^N))].
\end{aligned}$$

Then, we get

$$\mathbb{E}[\phi'(y)(c_k^0(y)(\alpha_k^N-\widehat{\alpha}_k^N))] = \mathbb{E}[\alpha_k^N-\widehat{\alpha}_k^N]\phi'(y)(c_k^0(y)) = 0.$$

We can do the same thing with the term in $\beta_{p,k}^N$ and obtain

$$\mathbb{E}[\phi'(y)(\psi_{\varepsilon,N}(y)-\widehat{\psi}_{\varepsilon,N}(y))] = 0.$$

Let us now study the second order term $Z = \phi''(y)(\widehat{\psi}_{\varepsilon,N}(y)-y)^2 - \phi''(y)(\psi_{\varepsilon,N}(y)-y)^2$ that appears in (5.4.6). We develop this expression and keep only the order one and two terms to obtain $Z = H^2Y + R$ where $\mathbb{E}[|R|] \leqslant C(1+|y|^K)H^3$ (by the same arguments as before) and

$$\begin{aligned} Y &= \sum_{k_1,k_2} \left[\phi''(y)(c^0_{k_1}(y)\widehat{\alpha}^N_{k_1}, c^0_{k_2}(y)\widehat{\alpha}^N_{k_2}) - \phi''(y)(c^0_{k_1}(y)\alpha^N_{k_1}, c^0_{k_2}(y)\alpha^N_{k_2})\right] \\ &= \sum_{k_1,k_2} (\widehat{\alpha}^N_{k_1}\widehat{\alpha}^N_{k_2} - \alpha^N_{k_1}\alpha^N_{k_2})\phi''(y)(c^0_{k_1}(y), c^0_{k_2}(y)) \end{aligned}$$

The condition on the moments of the $\widehat{\alpha}^N_k$ yields $\mathbb{E}[Y] = 0$.

Putting all these arguments together in (5.4.6), we finally get that

$$\left|\mathbb{E}[\phi(\psi_{\varepsilon,N}(y))] - \mathbb{E}[\phi(\widehat{\psi}_{\varepsilon,N}(y))]\right| \leqslant C(1+|y|^K)H^3.$$

We deduce the local order two of the proposed numerical scheme. □

Remark. *The constant H_0 in Proposition 5.4.6 depends on F, but also on the polynomial growth power of ϕ and its first three derivatives. This dependence is expected when trying to evaluate the solution of SDEs at random times. To make H_0 independent of the test functions, one can consider the following sets of test functions*

$$\mathcal{C}^3_{P,K} = \{\phi \in \mathcal{C}^3, \exists C > 0, \exists k \leqslant K, \forall y, \left|\phi^{(i)}(y)\right| \leqslant C(1+|y|^k), i \in \{0,1,2,3\}\}.$$

5.4.3 Global error

Theorem 5.4.7 (global convergence)**.** *Assume that the numerical scheme $\widehat{\psi}_{\varepsilon,N}$ satisfies equation* (5.4.2) *(respectively equation* (5.4.4)*) where $\widehat{\alpha}^N_k \in \mathbb{C}$, $\widehat{\beta}^N_{p,k} \in \mathbb{R}$ (respectively $\widehat{\widetilde{\beta}}^N_{p,k} \in \mathbb{R}$) are random variables such that $\widehat{\alpha}^N_k = \overline{\widehat{\alpha}^N_{-k}}$ and*

$$\mathbb{E}[\widehat{\alpha}^N_k] = \mathbb{E}[\alpha^N_k],\ \mathbb{E}[\widehat{\beta}^N_{p,k}] = \mathbb{E}[\beta^N_{p,k}],\ \mathbb{E}[\widehat{\alpha}^N_{k_1}\widehat{\alpha}^N_{k_2}] = \mathbb{E}[\alpha^N_{k_1}\alpha^N_{k_2}].$$

(respectively $\widehat{\alpha}^N_k$ satisfies the same conditions and $\widehat{\widetilde{\beta}}^N_{p,k}$ satisfies $\mathbb{E}[\widehat{\widetilde{\beta}}^N_{p,k}] = \mathbb{E}[\widetilde{\beta}^N_{p,k}]$). Under Assumption 5.3.1, if for all $q > 0$, $\mathbb{E}\left[\left(\sum_k \frac{|\widehat{\alpha}^N_k|^2}{k^2}\right)^q\right]$ and $\mathbb{E}\left[\left(\sum_{p,k} \frac{\left|\widehat{\beta}^N_{p,k}\right|^2}{k^2}\right)^q\right]$ are bounded uniformly in N (respectively $\sum_k \left|\widehat{\alpha}^N_k\right|$, $\sum_{p,k}\left|\widehat{\widetilde{\beta}}^N_{p,k}\right|$, $\mathbb{E}\left[\left(\sum_k \frac{|\widehat{\alpha}^N_k|^2}{k^2}\right)^q\right]$ and $\mathbb{E}\left[\left(\sum_{p,k} \frac{\left|\widehat{\widetilde{\beta}}^N_{p,k}\right|^2}{k^2}\right)^q\right]$ are bounded uniformly in N), for all $T > 0$, for all test function $\phi \in \mathcal{C}^3_P$, there exists $H_0 > 0$ such that for all $H \leqslant H_0$, for all $M \geqslant 0$ such that $MN\varepsilon = MH \leqslant T$, there exist two positive constants K and C both independent of ε and N such that

$$\left|\mathbb{E}[\phi(\varphi_{\varepsilon,T_{NM}}(X_0))] - \mathbb{E}[\phi(\widehat{\psi}^M_{\varepsilon,N}(X_0))]\right| \leqslant CH^2(1+\mathbb{E}[|X_0|^K]).$$

Proof. We denote

$$e_M = \mathbb{E}[\phi(\varphi_{\varepsilon,T_{NM}}(X_0))] - \mathbb{E}[\phi(\widehat{\psi}^M_{\varepsilon,N}(X_0))]$$

and write it with a telescopic sum

$$\begin{aligned} e_M &= \sum_{m=1}^{M} \mathbb{E}[\phi(\varphi_{\varepsilon,T_{N(m-1)}}(\widehat{\psi}^{M-m+1}_{\varepsilon,N}(X_0)))] - \mathbb{E}[\phi(\varphi_{\varepsilon,T_{Nm}}(\widehat{\psi}^{M-m}_{\varepsilon,N}(X_0)))] \\ &= \sum_{m=1}^{M} \mathbb{E}[\widetilde{\phi}_{m-1}(\widehat{\psi}_{\varepsilon,N}(\widehat{\psi}^{M-m}_{\varepsilon,N}(X_0)))] - \mathbb{E}[\widetilde{\phi}_{m-1}(\varphi_{\varepsilon,T_N}(\widehat{\psi}^{M-m}_{\varepsilon,N}(X_0)))] \end{aligned}$$

where $\widetilde{\phi}_{m-1} = \phi \circ \varphi_{\varepsilon,T_{N(m-1)}}$. Using Lemma 5.3.2 and $\phi \in \mathcal{C}_P^3$, we obtain for $0 \leqslant i \leqslant 3$,

$$\left|\widetilde{\phi}_m^{(i)}(y)\right| \leqslant Ce^{C\varepsilon T_{Nm}}(1+|y|^K).$$

Thus, knowing the hitting times involved, $\widetilde{\phi}_m \in \mathcal{C}_P^3$. Using Assumption 5.3.1, $(c_p^0)' = c_p^1$ and $\beta_{p,k}^N = \widetilde{\beta}_{p,k}^N + \frac{\alpha_p \alpha_k}{2}$, we deduce that $\widehat{\psi}_{\varepsilon,N}$ satisfies the assumptions of Proposition 5.4.6. Applying Proposition 5.4.6 to each term of e_M gives

$$|e_M| \leqslant \sum_{m=1}^{M} C\mathbb{E}\left[e^{C\varepsilon T_{Nm}}\right] H^3 \left(1 + \mathbb{E}\left[\left|\widehat{\psi}_{\varepsilon,N}^{M-m}(X_0)\right|^K\right]\right).$$

Finally, the moments of $\widehat{\psi}_{\varepsilon,N}^m(X_0)$ are all bounded uniformly in ε, N and m according to Proposition 5.4.4 (respectively 5.4.5). Thus

$$|e_M| \leqslant \sum_{m=1}^{M} CH^3(1+\mathbb{E}[|X_0|^K]) \leqslant CH^2(1+\mathbb{E}[|X_0|^K]).$$

We deduce the global weak order two. □

With the help of Theorem 5.4.7, we prove Proposition 5.3.10 and the convergence of Methods A and B.

Proof of Proposition 5.3.10. Writing Theorem 5.4.7 for order one yields for all $H = N\varepsilon$ small enough and all $M \geqslant 0$,

$$|\mathbb{E}[\phi(\varphi_{\varepsilon,T_{NM}}(X_0))] - \mathbb{E}[\phi(y_{\varepsilon NM})]| \leqslant C(\varepsilon N)^2(1+\mathbb{E}[|X_0|^K]).$$

Evaluating in $N = 1$, $M = \frac{T}{\varepsilon}$ and taking the limit $\varepsilon \to 0$ yield the result. □

Proof of Theorem 5.4.1. As $\widehat{\alpha}_k^N \leqslant C$ and $\sum_{p,k} \frac{\left|\mathbb{E}[\beta_{p,k}^N]\right|^2}{k^2}$ converges by Proposition 5.3.7, Theorem 5.4.7 applies and concludes the proof. □

Proof of Theorem 5.4.2. The regularity assumptions yield the Lipschitzness of the $c_k^0(y)$ and the involved $c_p^1(y)(c_k^0(y))$ with constants independent of k and p. As $\sum_k \left|\widehat{\alpha}_k^N\right|$ and $\sum_{p,k} \left|\widehat{\widetilde{\beta}}_{p,k}^N\right|$ are bounded, the right-hand side of equation (5.2.5) is a contraction for all $H \leqslant H_0$ small enough and the constant does not depend on Y_m, so H_0 depends only on F and F'. Thus, the integrator is well-posed for all $H \leqslant H_0$.

The weak order two is obtained using Theorem 5.4.7. Indeed the use of discrete random variables and Proposition 5.3.7 gives the convergence of the series involved.

For showing that Method B preserves quadratic invariants, it is sufficient to prove that $Q'(y)(\sum_k c_k^0(y)\widehat{\alpha}_k^N) = 0$ and $Q'(y)(\sum_{p,k} c_p^1(y)(c_k^0(y))\widehat{\widetilde{\beta}}_{p,k}^N) = 0$ (see [65, Chap. IV]). The preservation of Q by equation (2.1.5) yields $Q'(y)(Ay) = 0$ and $Q'(y)(F(y)) = 0$. We deduce the following two equations, valid for all $y \in \mathbb{R}^d$,

$$y^T S g_\theta^0(y) = 0, \tag{5.4.7}$$

$$y^T S g_\theta^1(y)(g_\nu^0(y)) = -(g_\nu^0(y))^T S g_\theta^0(y), \tag{5.4.8}$$

where equation (5.4.8) is obtained by differentiating equation (5.4.7) in the direction g_ν^0. Using equation (5.4.7), we have

$$Q'(y)\left(\sum_k c_k^0(y)\widehat{\alpha}_k^N\right) = \int_0^1 Q'(y)(g_\theta^0(y))\sum_k e^{-2i\pi k\theta}\widehat{\alpha}_k^N d\theta = 0.$$

For the second order term, equation (5.4.8) and the values of Proposition 5.3.7 yield

$$\begin{aligned}
&Q'(y)\left(\sum_{p,k} c_p^1(y)(c_k^0(y))\widehat{\widehat{\beta}}_{p,k}^N\right)\\
&\quad= \int_0^1\int_0^1 y^T S g_\theta^1(y)(g_\nu^0(y))\sum_{p,k} e^{-2i\pi p\theta}e^{-2i\pi k\nu}\widehat{\widehat{\beta}}_{p,k}^N d\nu d\theta\\
&\quad= -\int_0^1\int_0^1 (g_\nu^0(y))^T S g_\theta^0(y)\sum_{p,k} e^{-2i\pi p\theta}e^{-2i\pi k\nu}\widehat{\widehat{\beta}}_{p,k}^N d\nu d\theta\\
&\quad= -\frac{1}{2}\int_0^1\int_0^1 (g_\nu^0(y))^T S g_\theta^0(y)\sum_{p,k}[e^{-2i\pi p\theta}e^{-2i\pi k\nu} + e^{-2i\pi p\nu}e^{-2i\pi k\theta}]\widehat{\widehat{\beta}}_{p,k}^N d\nu d\theta\\
&\quad= 0.
\end{aligned}$$

Hence Method B is well-posed, has weak order 2 and preserves the invariant Q. □

5.5 Numerical experiments

In this section, we first illustrate numerically the weak order two of Methods A and B with convergence curves. Then, we apply the new algorithms to solve the nonlinear Schrödinger equation with highly-oscillatory white noise dispersion (5.1.2).

5.5.1 Weak order of convergence

To confirm the results of Theorem 5.4.1 and Theorem 5.4.2, we check numerically if Methods A and B have weak order two of accuracy w.r.t. H uniformly in ε and N. As the Euler-Maruyama method and the algorithms presented in [39, 11, 40] are completely inaccurate if they do not satisfy the severe timestep restriction $h \ll \varepsilon$, we compare the performance of Methods A and B to the performance of the Euler method (5.3.15). We first apply the algorithms to equation (2.1.5) with the linearity $F(y) = iy$, $A = 2i\pi$, $X_0 = 1$ and $\varepsilon = 10^{-3}$. Equivalently we can write it in the real setting as

$$dX = \frac{2\pi}{\sqrt{\varepsilon}}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} X \circ dW + \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} X dt,\ X_0 = \begin{pmatrix}1\\ 0\end{pmatrix}.$$

We plot on a logarithmic scale an estimate of the weak error for approximating X at time $T = 10^{-3}T_{2^8}$ where $\mathbb{E}[T] = 0.256$. The exact solution $X(T)$ is approximated by the output of Method B for $H = \varepsilon$. The parameters N and m are varying under the condition that $Nm = 2^8$. The test function is $\phi(y) = 2y_1 + 4y_2$ and the average is taken over 10^7 trajectories. We choose the tolerance 10^{-13} for the fixed point. On the right picture of Figure 5.4, we use a modification

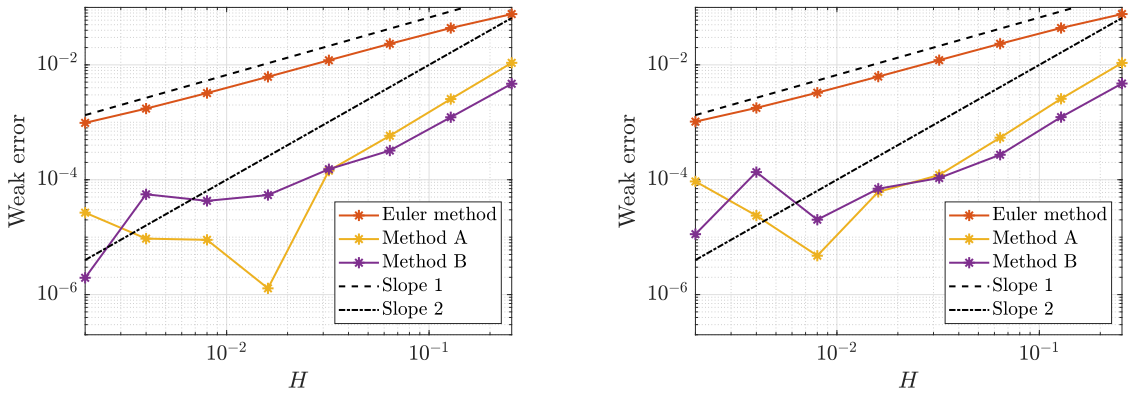


Figure 5.4: Weak error versus the stepsize $H = N\varepsilon$ for approximating the solution of equation (2.1.5) at time εT_{2^8} for the linear $F(y) = iy$ (left) and the non-linear $F(y) = i(1 + \mathrm{Re}(y)^3 + \mathrm{Im}(y)^5)y$ (right) with $A = 2i\pi$, $X_0 = 1$, $\varepsilon = 10^{-3}$ and the test function $\phi(y) = 2\,\mathrm{Re}(y) + 4\,\mathrm{Im}(y)$.

of a Kubo oscillator introduced in [39] with the nonlinearity $F(y) = i(1 + \mathrm{Re}(y)^3 + \mathrm{Im}(y)^5)y$. In the real setting, it yields the following two-dimensional SDE

$$dX = \frac{2\pi}{\sqrt{\varepsilon}} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} X \circ dW + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} (1 + X_1^3 + X_2^5) X dt, \; X_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

We take 8 modes for the Fourier decomposition and the same other parameters as before. The average is taken over 10^6 trajectories.

In both cases, we observe the weak order two of Methods A and B. The irregularities of the curve for a small H come from Monte-Carlo errors. We repeated the same experiment on many other examples and we always observe the desired order two as long as H is small enough.

5.5.2 Numerical experiments on the NLS equation with white noise dispersion

We now apply the algorithms to solve the following SPDE of the form (5.1.2) on the torus $\mathbb{T} = [-\pi, \pi]$, with a polynomial nonlinearity and the stiffness parameter $\varepsilon = 10^{-2}$,

$$du = \frac{2\pi}{\sqrt{\varepsilon}} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \Delta u \circ dW + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} |u|^{2\sigma} u dt, \; x \in \mathbb{T}, \; t > 0, \tag{5.5.1}$$

where the unknown u is a random process depending on $x \in \mathbb{T}$ and $t \geqslant 0$. We consider a spectral discretization in space of this equation with $K_x = 2^7$ modes $u(x,t) \approx \sum_{|l| \leqslant K_x} Y_l(t) e^{ilx}$. We obtain an equation of the desired form (2.1.5) with a truncated nonlinearity and the block-diagonal matrix

$$A = \mathrm{diag}(-2\pi l^2 \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, |l| \leqslant K_x).$$

Beginning with the initial condition $u_0(x) = \exp(-3x^4 + x^2)$ on $\mathbb{T}$ that decreases fast enough, we apply Methods A and B in the two cases $\sigma = 2$ and $\sigma = 4$ with $K_t = 2^6$ modes, $N = 10$ revolutions, $m = 150$ iterations and a tolerance of 10^{-13} for the fixed point iteration. Figure 5.5 shows the evolution in time of one trajectory given by Method B (with a 300 points evaluation grid in space).

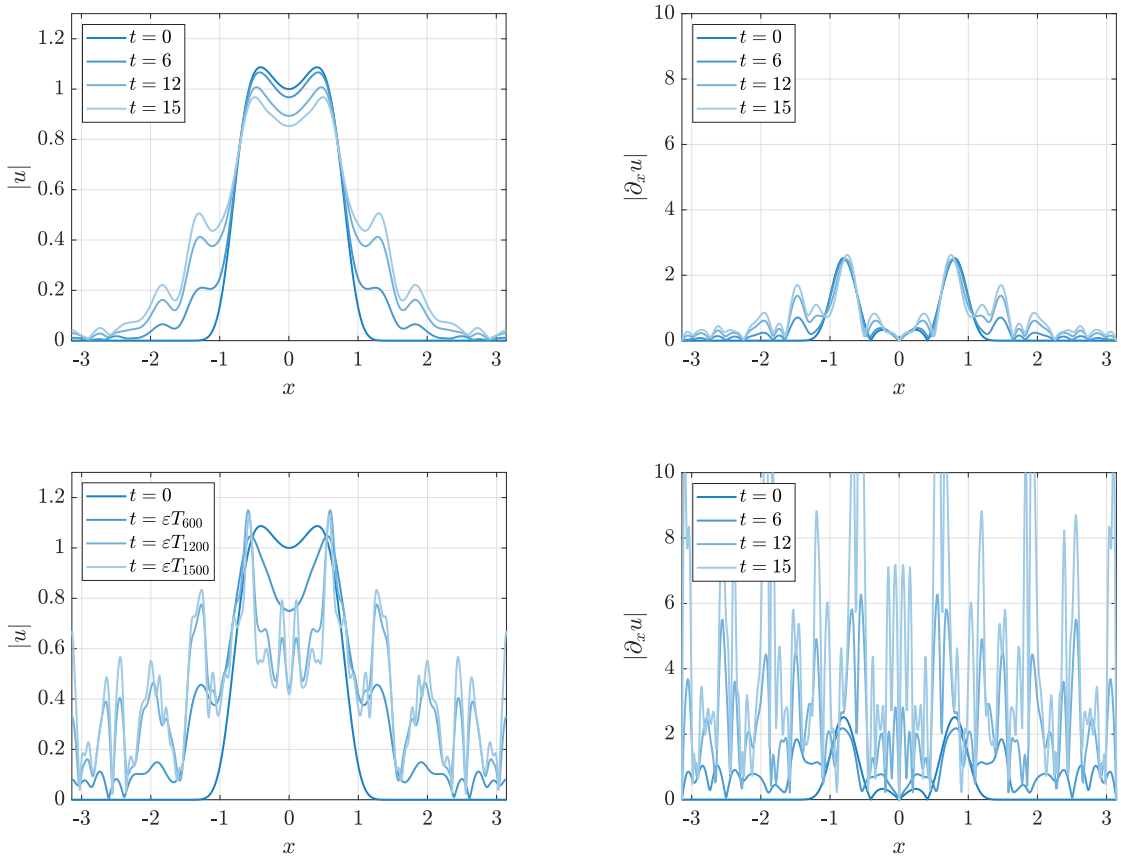


Figure 5.5: Approximation by Method B of $|u|$ and $|\partial_x u|$ with u solution of a spatial discretization with $K_x = 2^7$ modes of the nonlinear Schrödinger equation with white noise dispersion (5.5.1) on the torus $\mathbb{T} = [-\pi, \pi]$ with the parameters $\varepsilon = 10^{-2}$, $\sigma = 2$ (top) and $\sigma = 4$ (bottom).

In Figure 5.6, we observe the behavior of the discrete L^2 and H^1 norms of one trajectory given by our two algorithms and the Euler method (5.3.15) (the simulated $(\alpha_k)_k$ are the same for Methods A and B). The Euler method quickly blows up in both norms. The L^2 norm of Method A is not conserved. In contrast, Method B preserves the L^2 norm according to Theorem 5.4.2. When $\sigma = 4$, numerical simulations hint that a blow-up in the H^1 norm always happens for all considered methods at a certain time that increases as ε goes to zero. We recall that in the optic fiber model (5.1.2), t represents the distance along the optic fiber and a cubic nonlinearity ($\sigma = 2$) is typically considered [58]. For $\sigma = 2$, we do not observe any blow-up in the H^1 norm in Figure 5.6, suggesting the well-posedness of the model for all optic fiber distance. Also, the larger σ is, the sooner the blow-up happens. These behaviors agree with the blow-up conjecture for $\varepsilon = 1$ and $\sigma \geqslant 4$ presented in [11], and suggest that the conjecture persists in the highly-oscillatory regime $\varepsilon \ll 1$.

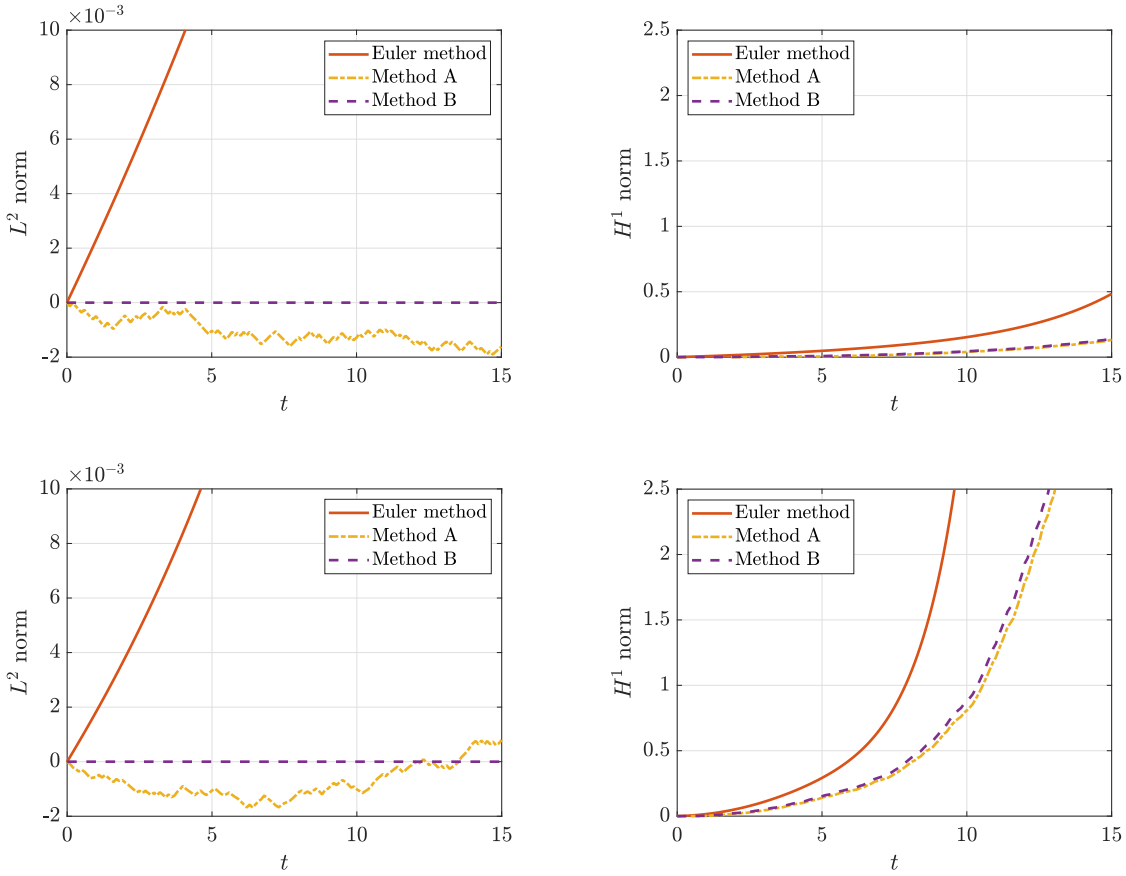


Figure 5.6: Evolution in long time of the quantities $\|U_t\|_{L^2} - \|u_0\|_{L^2}$ (left) and $\|U_t\|_{H^1} - \|u_0\|_{H^1}$ (right) with U_t the approximation computed with Euler method and Methods A and B for one trajectory of equation (5.5.1) with $\varepsilon = 10^{-2}$, $\sigma = 2$ (top) and $\sigma = 4$ (bottom).

CHAPTER 6

Conclusion and outlook

In this thesis, we introduced new methods and tools for the numerical integration of stochastic evolutionary problems.

First, in Chapter 3 and in the articles [83, 85], we introduced a new formalism of B-series, called exotic aromatic B-series, that is designed specifically for computing order conditions for sampling the invariant measure of Langevin dynamics in $\mathbb{R}^d$ or on manifolds. This formalism also applies to partitioned methods, to the use of postprocessors, or perturbations that improve the rate of convergence to infinity. The large number of terms in the Taylor expansions involved in the considered schemes shows the necessity of a strong formalism to manage the tedious calculations of order conditions already for relatively small orders for the invariant measure (order two or three already).

In Chapter 4 and in the articles [83, 85], we applied the B-series formalism to find explicitly the order conditions for a general class of Runge-Kutta methods for solving Langevin dynamics in $\mathbb{R}^d$ and on manifolds, in the weak sense and for sampling directly the invariant measure. In particular, we introduced the first order two scheme that does not reduce to a splitting method for sampling the invariant measure of constrained Langevin dynamics.

In Chapter 5 and in the article [84], we designed a robust numerical methods for solving stochastic (partial) differential equations with fast stochastic oscillations. We introduced a method of weak order two with computational cost and accuracy both independent of the stiffness of the oscillations, and we built a geometric modification of this method that preserves quadratic invariants. We proved that these algorithms have weak order two, and illustrated their convergence and behavior with the help of numerical experiments, by solving in particular highly-oscillatory Kubo oscillators and the nonlinear Schrödinger equation with white noise dispersion discretized in space with a Fourier expansion.

The new tools and methods presented in this work give rise to many new interesting research ideas, that we present in the following sections.

6.1 Future work

Even for order two on a manifold, the computations of the order conditions for sampling the invariant measure are barely feasible by hand, and their complexity is much higher without

the help of the B-series formalism (see Appendix C). With E. Bronasco and G. Vilmart, we are implementing the exotic aromatic B-series and the graph operations discussed in Chapter 3 to make the calculations of the order conditions automatic on a symbolic manipulation package (see [16]). The exotic aromatic B-series are also interesting for their algebraic and geometric properties. It would be interesting to find a link between isometric equivariance and exotic aromatic B-series methods as indicated in Section 3.2.4 and in the spirit of [100, 105, 56], and to study the algebraic structure of the exotic aromatic B-series in the spirit of [13, 31].

The methodology we presented for creating high order integrators Langevin dynamics can be generalized in various ways by creating partitioned methods, adding postprocessors or perturbations or considering a multiplicative noise. Combining high order integration with efficient Metropolis-Hastings rejection procedures [103, 69] is also a crucial question where the tools presented here could bring insight. More generally, the work presented in Chapter 3 and Chapter 4 could lead to the development of new integrators with small error constant, favorable stability properties, small variance and fast convergence to equilibrium, which is a challenging open question.

We presented in Chapter 5 a high order method for solving stochastic evolutionary problems with a dispersion term that is led by a highly-oscillatory white noise. The multirevolution methods A and B are robust in the regime where $\varepsilon \ll 1$. It would be interesting to find a uniformly accurate (UA) method that has a cost and accuracy independent of the (possibly not small) parameter ε, in the spirit of the recent works [33, 34]. We could try to generalize the analysis directly to SPDEs with fast white noise dispersion, or to highly-oscillatory SDEs led, for instance, by two fast white noise terms of different frequencies, or by one fast white noise and one fast deterministic oscillation, in the spirit of [32] in the deterministic context.

The variety of tools we developed in this work recently led us to the study of a uniformly accurate discretization for penalized Langevin dynamics. We give more details on this work in Section 6.2.

6.2 A uniformly accurate integrator for penalized Langevin dynamics

In this section, we present a few results of the work [82] in preparation. We saw in Chapter 2 the overdamped Langevin dynamic (2.1.3) in $\mathbb{R}^d$, and its constrained counterpart (2.1.4) on a manifold. In a variety of physical applications, the constraints are fixed up to a small parameter ε. The trajectory of the solution lies in the vicinity of the manifold $\mathcal{M}$. One is then interested in classes of SDEs in $\mathbb{R}^d$ indexed by a possibly stiff parameter ε that is related to the distance to the manifold. With the same notations as in Chapter 2, we introduce the following penalized dynamics in $\mathbb{R}^d$ to simulate trajectories in a vicinity of the manifold $\mathcal{M}$,

$$dX^\varepsilon(t) = f(X^\varepsilon(t))dt + \sigma dW(t) + \frac{\sigma^2}{4}\nabla \ln(\det(G))(X^\varepsilon(t))dt - \frac{1}{\varepsilon}(gG^{-1}\zeta)(X^\varepsilon(t))dt, \quad (6.2.1)$$

where $X^\varepsilon(0) = X_0$ and the parameter ε can be of arbitrary size. It was shown in [38, Appendix C] for similar penalized dynamics with a stiff term of the form $\frac{1}{\varepsilon}g\zeta$ instead of $\frac{1}{\varepsilon}gG^{-1}\zeta$, that the solution X^ε converges strongly to the solution X^0 of (2.1.4) if $X_0 \in \mathcal{M}$. This strong convergence still holds for the dynamic (6.2.1), and we applied a renormalization in the stiff term as it simplifies the analysis.

A commonly used discretization for solving (6.2.1) is the explicit Euler scheme,

$$X_{n+1} = X_n + \sqrt{h}\sigma\xi + hf(X_n) + h\frac{\sigma^2}{4}\nabla \ln(\det(G))(X_n) - \frac{h}{\varepsilon}(gG^{-1}\zeta)(X_n). \tag{6.2.2}$$

This integrator has weak order one of accuracy, but it faces some severe stepsize restriction, typically of the form $h \ll \varepsilon$, in order to be accurate in the regime $\varepsilon \ll 1$. The alternative of the explicit Euler scheme (6.2.2) on the manifold is the constrained Euler scheme (4.1.1). This integrator has weak order one for solving (2.1.4) and it lies on the manifold $\mathcal{M}$. It is a consistent approximation of (6.2.1) if ε tends to zero. This integrator is, however, not fit for solving (6.2.1) if ε is of size $\mathcal{O}(1)$, since the exact solution $X^\varepsilon(t)$ is not constrained on the manifold in this regime.

Inspired by the results of Chapter 4, Chapter 5 and [38, Appendix C], we propose a new integrator for solving penalized Langevin dynamics. In the context of a compact manifold $\mathcal{M}$ of codimension one, the algorithm is the following:

New method: Uniform Discretization of Penalized Langevin Overdamped Dynamics (6.2.1)

$$\begin{aligned}
X^\varepsilon_{n+1} &= X^\varepsilon_n + \sqrt{h}\sigma\xi_n + hf(X^\varepsilon_n) + \frac{(1-e^{-h/\varepsilon})^2}{2}(\zeta^2 G^{-2} g' g)(X^\varepsilon_n) \\
&\quad + \frac{\sigma^2\varepsilon}{4}(1-e^{-2h/\varepsilon})(G^{-1}g'g)(X^\varepsilon_n) + g(X^\varepsilon_n)\lambda \\
\zeta(X^\varepsilon_{n+1}) &= e^{-h/\varepsilon}\zeta(X^\varepsilon_n) + \sigma\sqrt{\frac{\varepsilon}{2}(1-e^{-2h/\varepsilon})}g^T(X^\varepsilon_n)\xi_n \\
&\quad + \varepsilon(1-e^{-h/\varepsilon})(g^T f + \frac{\sigma^2}{2}G^{-1}g^T g' g + \frac{\sigma^2}{2}\operatorname{div}(g))(X^\varepsilon_n)
\end{aligned} \tag{6.2.3}$$

We would like to prove in the future work [82] that the method (6.2.3) is uniformly accurate (UA), that is, it has an accuracy and a cost that do not depend on the stiff parameter ε (see, for instance, [30, 32, 33, 34] and references therein). In particular, it would imply that the method (6.2.3) is asymptotic-preserving (AP) [76, 75, 77], that is, it converges in both regimes $\varepsilon \ll 1$ and ε of the size of $\mathcal{O}(1)$. Note that, in this latter case, the convergence can be proved straightforwardly with a Taylor expansion. Using robust discretizations with respect to the stiff parameter ε allows us to take relatively large timesteps h, even if ε is very small. It plays an important role in multilevel Monte-Carlo methods (see [59, 60] and references therein), or parallel in time methods [57] such as Parareal [95].

The idea of the algorithm (6.2.3) is the following. Instead of evaluating the possibly stiff term $\frac{h}{\varepsilon}gG^{-1}\zeta$ as in the Euler scheme (6.2.2), we project in (6.2.3) a modified step of the explicit Euler scheme in $\mathbb{R}^d$ on a manifold that is close to $\mathcal{M}$ when $\varepsilon \ll 1$, and whose constraint is given by a truncation of a uniform expansion of $\zeta(X^\varepsilon)$. The discretization (6.2.3) is in the spirit of the class of Runge-Kutta projection methods (3.1.2) of Chapter 3 and Chapter 4, and uses truncated uniform expansions in h that are uniform with respect to the parameter ε in the spirit of Chapter 5. Note that the integrator (6.2.3) has a cost similar to the cost of the constrained Euler scheme (4.1.1) in terms of the number of evaluation of the functions f, ζ, g and g'.

To observe numerically the uniform accuracy property of the method (6.2.3), we apply it on a torus in $\mathbb{R}^3$ for sampling the invariant measure of (6.2.1) for different time steps h and parameters ε, and we compare it with the Euler integrators (6.2.2) in $\mathbb{R}^d$ and (4.1.1) on the

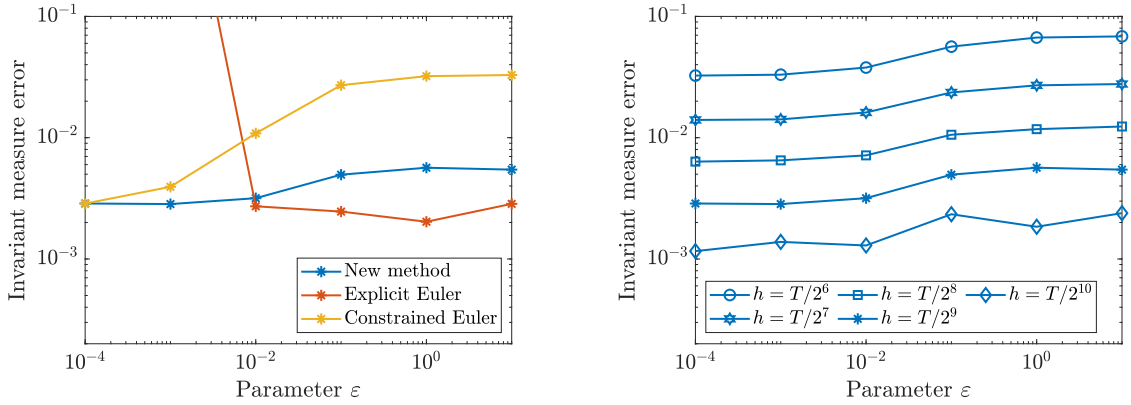


Figure 6.1: Error for sampling the invariant measure of the penalized Langevin dynamics (6.2.1) on a torus in $\mathbb{R}^3$ of the discretization (6.2.3) and the Euler integrators (6.2.2) and (4.1.1) for different values of ε with $h = 2^{-9}T$ (left), and error curves versus ε of the discretization (6.2.3) for different timesteps $h = 2^{-i}T$ and $i = 6, \dots, 10$ (right), with the parameters $T = 10$, $f(x) = -25(x_1 - R + r, x_2, x_3)$, $\phi(x) = |x|^2$ and $M = 10^7$ trajectories.

manifold $\mathcal{M}$. The constraint is given by $\zeta(x) = (x_1^2 + x_2^2 + x_3^2 + R^2 - r^2)^2 - 4R^2(x_1^2 + x_2^2)$ with $R = 3$ and $r = 1$, and we choose $f(x) = -25(x_1 - R + r, x_2, x_3)$, with the parameters $\sigma = \sqrt{2}$, $\phi(x) = |x|^2$ and the final time $T = 10$. Increasing the value of T does not modify the computed averages, which hints that we reached the equilibrium. The factor 25 in f confines the solution in a reasonably small neighborhood of the torus, which allows a faster convergence to equilibrium and to take less trajectories. With the help of the Baobab cluster of the University of Geneva, we compute the Monte-Carlo estimator $\bar{J} = \frac{1}{M}\sum_{m=1}^{M} \phi(X_N^{(m)}) \simeq \mathbb{E}[\phi(X_N)]$ with $M = 10^7$ trajectories, where $X_n^{(m)}$ is the m-th realisation of the integrator at time $t_n = nh$, and N is an integer satisfying $Nh = T$. We compare this approximation with a reference value of $\int_{\mathcal{M}} \phi d\mu_\infty$ computed with the method (6.2.3), by using a timestep $h = 2^{-12}$. We also plot an estimate of the Monte-Carlo error by using the standard error of the mean estimator $\left(\sum_{m=1}^{M}(\phi(X_N^{(m)}) - \bar{J})^2\right)^{1/2}/\sqrt{M(M-1)}$. We observe in Figure 6.1 that the explicit Euler scheme (6.2.2) blows up in the regime $\varepsilon \ll 1$, and that the constrained integrator (4.1.1) is accurate only in the regime $\varepsilon \ll 1$. In contrast, the accuracy of the method (6.2.3) does not deteriorate depending on ε.

Bibliography

[1] A. Abdulle, I. Almuslimani, and G. Vilmart. Optimal explicit stabilized integrator of weak order 1 for stiff and ergodic stochastic differential equations. *SIAM/ASA J. Uncertain. Quantif.*, 6(2):937–964, 2018.

[2] A. Abdulle, D. Cohen, G. Vilmart, and K. C. Zygalakis. High weak order methods for stochastic differential equations based on modified equations. *SIAM J. Sci. Comput.*, 34(3):A1800–A1823, 2012.

[3] A. Abdulle, T. Li, and W. E. Effectiveness of implicit methods for stiff stochastic differential equations. *Commun. Comput. Phys.*, 3(2):295–307, 2008.

[4] A. Abdulle, G. A. Pavliotis, and G. Vilmart. Accelerated convergence to equilibrium and reduced asymptotic variance for Langevin dynamics using Stratonovich perturbations. *C. R. Math. Acad. Sci. Paris*, 357(4):349–354, 2019.

[5] A. Abdulle, G. Vilmart, and K. C. Zygalakis. High order numerical approximation of the invariant measure of ergodic SDEs. *SIAM J. Numer. Anal.*, 52(4):1600–1622, 2014.

[6] A. Abdulle, G. Vilmart, and K. C. Zygalakis. Long time accuracy of Lie-Trotter splitting methods for Langevin dynamics. *SIAM J. Numer. Anal.*, 53(1):1–16, 2015.

[7] G. Agrawal. *Nonlinear Fiber Optics*. Electronics & Electrical. Academic Press, 2007.

[8] G. Agrawal. *Applications of Nonlinear Fiber Optics*. Number vol. 10 in Applications of nonlinear fiber optics. Elsevier, 2008.

[9] H. C. Andersen. Rattle: A "velocity" version of the Shake algorithm for molecular dynamics calculations. *J. Comput. Phys.*, 52(1):24–34, 1983.

[10] S. Anmarkrud and A. Kværnø. Order conditions for stochastic Runge-Kutta methods preserving quadratic invariants of Stratonovich SDEs. *J. Comput. Appl. Math.*, 316:40–46, 2017.

[11] R. Belaouar, A. de Bouard, and A. Debussche. Numerical analysis of the nonlinear Schrödinger equation with white noise dispersion. *Stoch. Partial Differ. Equ. Anal. Comput.*, 3(1):103–132, 2015.

[12] S. Blanes and F. Casas. On the necessity of negative coefficients for operator splitting schemes of order higher than two. *Appl. Numer. Math.*, 54(1):23–37, 2005.

[13] G. Bogfjellmo. Algebraic structure of aromatic B-series. *J. Comput. Dyn.*, 6(2):199–222, 2019.

[14] N. Bou-Rabee and H. Owhadi. Long-run accuracy of variational integrators in the stochastic context. *SIAM J. Numer. Anal.*, 48(1):278–297, 2010.

[15] C.-E. Bréhier and G. Vilmart. High order integrator for sampling the invariant distribution of a class of parabolic stochastic PDEs with additive space-time noise. *SIAM J. Sci. Comput.*, 38(4):A2283–A2306, 2016.

[16] E. Bronasco. Algebraic tools and symbolic package for the study of order conditions for sampling the invariant measure of ergodic stochastic differential equations. Master's thesis, University of Geneva, 2021. Supervisors: A. Laurent and G. Vilmart.

[17] C. Brouder. Runge–Kutta methods and renormalization. *Eur. Phys. J. C*, 12(3):521–534, 2000.

[18] M. Brubaker, M. Salzmann, and R. Urtasun. A family of MCMC methods on implicitly defined manifolds. In *Artificial intelligence and statistics*, pages 161–172, 2012.

[19] Y. Bruned, M. Hairer, and L. Zambotti. Algebraic renormalisation of regularity structures. *Invent. Math.*, 215(3):1039–1156, 2019.

[20] K. Burrage and P. M. Burrage. High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations. *Appl. Numer. Math.*, 22(1-3):81–101, 1996. Special issue celebrating the centenary of Runge-Kutta methods.

[21] K. Burrage and P. M. Burrage. Order conditions of stochastic Runge-Kutta methods by B-series. *SIAM J. Numer. Anal.*, 38(5):1626–1646, 2000.

[22] J. C. Butcher. Coefficients for the study of Runge-Kutta integration processes. *J. Austral. Math. Soc.*, 3:185–201, 1963.

[23] J. C. Butcher. The effective order of Runge-Kutta methods. In *Conf. on Numerical Solution of Differential Equations (Dundee, 1969)*, pages 133–139. Springer, Berlin, 1969.

[24] J. C. Butcher. An algebraic theory of integration methods. *Math. Comp.*, 26:79–106, 1972.

[25] J. C. Butcher. *Numerical methods for ordinary differential equations*. John Wiley & Sons, Ltd., Chichester, third edition, 2016.

[26] J. C. Butcher. *B-series: algebraic analysis of numerical methods*. Springer, 2021.

[27] M. Calvo, L. O. Jay, J. I. Montijano, and L. Rández. Approximate compositions of a near identity map by multi-revolution Runge-Kutta methods. *Numer. Math.*, 97(4):635–666, 2004.

[28] M. Calvo, J. I. Montijano, and L. Rández. On explicit multi-revolution Runge-Kutta schemes. *Adv. Comput. Math.*, 26(1-3):105–120, 2007.

[29] M. P. Calvo, A. Murua, and J. M. Sanz-Serna. Modified equations for ODEs. In *Chaotic numerics (Geelong, 1993)*, volume 172 of *Contemp. Math.*, pages 63–74. Amer. Math. Soc., Providence, RI, 1994.

[30] P. Chartier, N. Crouseilles, M. Lemou, and F. Méhats. Uniformly accurate numerical schemes for highly oscillatory Klein-Gordon and nonlinear Schrödinger equations. *Numer. Math.*, 129(2):211–250, 2015.

[31] P. Chartier, E. Hairer, and G. Vilmart. Algebraic structures of B-series. *Found. Comput. Math.*, 10(4):407–427, 2010.

[32] P. Chartier, M. Lemou, and F. Méhats. Highly-oscillatory evolution equations with multiple frequencies: averaging and numerics. *Numer. Math.*, 136(4):907–939, 2017.

[33] P. Chartier, M. Lemou, F. Méhats, and G. Vilmart. A new class of uniformly accurate numerical schemes for highly oscillatory evolution equations. *Found. Comput. Math.*, 20(1):1–33, 2020.

[34] P. Chartier, M. Lemou, and L. Trémant. Uniformly accurate numerical schemes for a class of dissipative systems. *Submitted*, 2020.

[35] P. Chartier, J. Makazaga, A. Murua, and G. Vilmart. Multi-revolution composition methods for highly oscillatory differential equations. *Numer. Math.*, 128(1):167–192, 2014.

[36] P. Chartier and A. Murua. Preserving first integrals and volume forms of additively split systems. *IMA J. Numer. Anal.*, 27(2):381–405, 2007.

[37] G. Ciccotti, R. Kapral, and E. Vanden-Eijnden. Blue moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics. *ChemPhysChem*, 6(9):1809–1814, 2005.

[38] G. Ciccotti, T. Lelievre, and E. Vanden-Eijnden. Projection of diffusions on submanifolds: application to mean force computation. *Comm. Pure Appl. Math.*, 61(3):371–408, 2008.

[39] D. Cohen. On the numerical discretisation of stochastic oscillators. *Math. Comput. Simulation*, 82(8):1478–1495, 2012.

[40] D. Cohen and G. Dujardin. Exponential integrators for nonlinear Schrödinger equations with white noise dispersion. *Stoch. Partial Differ. Equ. Anal. Comput.*, 5(4):592–613, 2017.

[41] D. Cohen and M. Sigg. Convergence analysis of trigonometric methods for stiff second-order stochastic differential equations. *Numer. Math.*, 121(1):1–29, 2012.

[42] A. Connes and D. Kreimer. Hopf algebras, renormalization and noncommutative geometry. *Comm. Math. Phys.*, 199(1):203–242, 1998.

[43] G. J. Cooper. Stability of Runge-Kutta methods for trajectory problems. *IMA J. Numer. Anal.*, 7(1):1–13, 1987.

[44] A. de Bouard and A. Debussche. The nonlinear Schrödinger equation with white noise dispersion. *J. Funct. Anal.*, 259(5):1300–1321, 2010.

[45] K. Debrabant. Runge-Kutta methods for third order weak approximation of SDEs with multidimensional additive noise. *BIT Numer. Math.*, 50(3):541–558, 2010.

[46] K. Debrabant and A. Kværnø. B-series analysis of stochastic Runge-Kutta methods that use an iterative scheme to compute their internal stage values. *SIAM J. Numer. Anal.*, 47(1):181–203, 2008/09.

[47] K. Debrabant and A. Kværnø. Composition of stochastic B-series with applications to implicit Taylor methods. *Appl. Numer. Math.*, 61(4):501–511, 2011.

[48] A. Debussche and E. Faou. Weak backward error analysis for SDEs. *SIAM J. Numer. Anal.*, 50(3):1735–1752, 2012.

[49] A. Debussche and Y. Tsutsumi. 1D quintic nonlinear Schrödinger equation with white noise dispersion. *J. Math. Pures Appl. (9)*, 96(4):363–376, 2011.

[50] P. Diaconis, S. Holmes, and M. Shahshahani. Sampling from a manifold. In *Advances in modern statistical theory and applications: a Festschrift in honor of Morris L. Eaton*, volume 10 of *Inst. Math. Stat. (IMS) Collect.*, pages 102–125. Inst. Math. Statist., Beachwood, OH, 2013.

[51] A. B. Duncan, T. Lelièvre, and G. A. Pavliotis. Variance reduction using nonreversible Langevin samplers. *J. Stat. Phys.*, 163(3):457–491, 2016.

[52] U. Erdoğan and G. J. Lord. A new class of exponential integrators for SDEs with multiplicative noise. *IMA J. Numer. Anal.*, 39(2):820–846, 2019.

[53] L. C. Evans. *An introduction to stochastic differential equations.* American Mathematical Society, Providence, RI, 2013.

[54] E. Faou. Linearized wave turbulence convergence results for three-wave systems. *Comm. Math. Phys.*, 378(2):807–849, 2020.

[55] E. Faou and T. Lelièvre. Conservative stochastic differential equations: mathematical and numerical analysis. *Math. Comp.*, 78(268):2047–2074, 2009.

[56] G. Fløystad, D. Manchon, and H. Z. Munthe-Kaas. The universal pre-Lie-Rinehart algebras of aromatic trees. *Submitted*, arXiv:2002.05718, 2020.

[57] M. J. Gander. 50 years of time parallel time integration. In *Multiple shooting and time domain decomposition methods*, volume 9 of *Contrib. Math. Comput. Sci.*, pages 69–113. Springer, Cham, 2015.

[58] J. Garnier. Stabilization of dispersion-managed solitons in random optical fibers by strong dispersion management. *Opt. Commun.*, 206(4-6):411–438, 2002.

[59] M. B. Giles. Multilevel Monte Carlo path simulation. *Oper. Res.*, 56(3):607–617, 2008.

[60] M. B. Giles. Multilevel Monte Carlo methods. *Acta Numer.*, 24:259–328, 2015.

[61] M. Girolami and B. Calderhead. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 73(2):123–214, 2011. With discussion and a reply by the authors.

[62] D. Goldfeld. *Automorphic forms and L-functions for the group* GL($n, \boldsymbol{R}$), volume 99 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006. With an appendix by Kevin A. Broughan.

[63] E. Hairer. Backward analysis of numerical integrators and symplectic methods. *Ann. Numer. Math.*, 1(1-4):107–132, 1994. Scientific computation and differential equations (Auckland, 1993).

[64] E. Hairer. Long-time energy conservation of numerical integrators. In *Foundations of computational mathematics, Santander 2005*, volume 331 of *London Math. Soc. Lecture Note Ser.*, pages 162–180. Cambridge Univ. Press, Cambridge, 2006.

[65] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006. Structure-preserving algorithms for ordinary differential equations.

[66] E. Hairer and G. Wanner. On the Butcher group and general multi-value methods. *Computing (Arch. Elektron. Rechnen)*, 13(1):1–15, 1974.

[67] E. Hairer and G. Wanner. *Solving ordinary differential equations. II*, volume 14 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2010. Stiff and differential-algebraic problems, Second revised edition, paperback.

[68] R. Z. Hasminskii. *Stochastic stability of differential equations*, volume 7 of *Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis*. Sijthoff & Noordhoff, Alphen aan den Rijn—Germantown, Md., 1980. Translated from the Russian by D. Louvish.

[69] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[70] N. Hofmann, T. Müller-Gronbach, and K. Ritter. Optimal approximation of stochastic differential equations by adaptive step-size control. *Math. Comp.*, 69(231):1017–1034, 2000.

[71] E. P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.

[72] A. Iserles, G. R. W. Quispel, and P. S. P. Tse. B-series methods cannot be volume-preserving. *BIT Numer. Math.*, 47(2):351–378, 2007.

[73] L. Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12, 11 1918.

[74] L. O. Jay. Structure preservation for constrained dynamics with super partitioned additive Runge–Kutta methods. *SIAM J. Sci. Comput.*, 20(2):416–446, 1998.

[75] S. Jin. Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM J. Sci. Comput.*, 21(2):441–454, 1999.

[76] A. Klar. An asymptotic-induced scheme for nonstationary transport equations in the diffusive limit. *SIAM J. Numer. Anal.*, 35(3):1073–1094, 1998.

[77] A. Klar. An asymptotic preserving numerical scheme for kinetic equations in the low Mach number limit. *SIAM J. Numer. Anal.*, 36(5):1507–1527, 1999.

[78] Y. Komori, T. Mitsui, and H. Sugiura. Rooted tree analysis of the order conditions of ROW-type scheme for stochastic differential equations. *BIT Numer. Math.*, 37(1):43–66, 1997.

[79] M. Kopec. Weak backward error analysis for Langevin process. *BIT Numer. Math.*, 55(4):1057–1103, 2015.

[80] M. Kopec. Weak backward error analysis for overdamped Langevin processes. *IMA J. Numer. Anal.*, 35(2):583–614, 2015.

[81] D. Küpper, A. Kværnø, and A. Rößler. A Runge-Kutta method for index 1 stochastic differential-algebraic equations with scalar noise. *BIT Numer. Math.*, 52(2):437–455, 2012.

[82] A. Laurent. A uniformly accurate integrator for penalized overdamped Langevin dynamics. *In preparation*, 2021.

[83] A. Laurent and G. Vilmart. Exotic aromatic B-series for the study of long time integrators for a class of ergodic SDEs. *Math. Comp.*, 89(321):169–202, 2020.

[84] A. Laurent and G. Vilmart. Multirevolution integrators for differential equations with fast stochastic oscillations. *SIAM J. Sci. Comput.*, 42(1):A115–A139, 2020.

[85] A. Laurent and G. Vilmart. Order conditions for sampling the invariant measure of ergodic stochastic differential equations on manifolds. *To appear in Found. Comput. Math.*, 2021.

[86] B. Leimkuhler and C. Matthews. Rational construction of stochastic numerical methods for molecular sampling. *Appl. Math. Res. Express. AMRX*, 2013(1):34–56, 2013.

[87] B. Leimkuhler, C. Matthews, and G. Stoltz. The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics. *IMA J. Numer. Anal.*, 36(1):13–79, 2016.

[88] T. Lelièvre, C. Le Bris, and E. Vanden-Eijnden. Analyse de certains schémas de discrétisation pour des équations différentielles stochastiques contraintes. *C. R. Math. Acad. Sci. Paris*, 346(7-8):471–476, 2008.

[89] T. Lelièvre, F. Nier, and G. A. Pavliotis. Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion. *J. Stat. Phys.*, 152(2):237–274, 2013.

[90] T. Lelièvre, M. Rousset, and G. Stoltz. *Free energy computations*. Imperial College Press, London, 2010. A mathematical perspective.

[91] T. Lelièvre, M. Rousset, and G. Stoltz. Langevin dynamics with constraints and computation of free energy differences. *Math. Comp.*, 81(280):2071–2125, 2012.

[92] T. Lelièvre, M. Rousset, and G. Stoltz. Hybrid Monte Carlo methods for sampling probability measures on submanifolds. *Numer. Math.*, 143(2):379–421, 2019.

[93] T. Lelièvre, G. Stoltz, and W. Zhang. Multiple projection MCMC algorithms on submanifolds. *Submitted*, arXiv:2003.09402, 2020.

[94] X.-M. Li. Homogenisation on homogeneous spaces. *J. Math. Soc. Japan*, 70(2):519–572, 2018. With an appendix by Dmitriy Rumynin.

[95] J.-L. Lions, Y. Maday, and G. Turinici. Résolution d'EDP par un schéma en temps "pararéel". *C. R. Acad. Sci. Paris Sér. I Math.*, 332(7):661–668, 2001.

[96] C. Lubich. *From quantum to classical molecular dynamics: reduced models and numerical analysis*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2008.

[97] R. Marty. On a splitting scheme for the nonlinear Schrödinger equation in a random medium. *Commun. Math. Sci.*, 4(4):679–705, 2006.

[98] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2):185–232, 2002.

[99] J. C. Mattingly, A. M. Stuart, and M. V. Tretyakov. Convergence of numerical time-averaging and stationary measures via Poisson equations. *SIAM J. Numer. Anal.*, 48(2):552–577, 2010.

[100] R. I. McLachlan, K. Modin, H. Munthe-Kaas, and O. Verdier. B-series methods are exactly the affine equivariant methods. *Numer. Math.*, 133(3):599–622, 2016.

[101] R. I. McLachlan, K. Modin, H. Munthe-Kaas, and O. Verdier. Butcher series: a story of rooted trees and numerical methods for evolution equations. *Asia Pac. Math. Newsl.*, 7(1):1–11, 2017.

[102] B. Melendo and M. Palacios. A new approach to the construction of multirevolution methods and their implementation. *Appl. Numer. Math.*, 23(2):259–274, 1997.

[103] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, 1953.

[104] G. N. Milstein and M. V. Tretyakov. *Stochastic numerics for mathematical physics*. Scientific Computation. Springer-Verlag, Berlin, 2004.

[105] H. Munthe-Kaas and O. Verdier. Aromatic Butcher series. *Found. Comput. Math.*, 16(1):183–215, 2016.

[106] A. Murua and J. M. Sanz-Serna. Order conditions for numerical integrators obtained by composing simpler integrators. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.*, 357(1754):1079–1100, 1999.

[107] N. Nüsken and G. A. Pavliotis. Constructing sampling schemes via coupling: Markov semigroups and optimal transport. *SIAM/ASA J. Uncertain. Quantif.*, 7(1):324–382, 2019.

[108] B. Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.

[109] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1994.

[110] A. Rößler. Stochastic Taylor expansions for the expectation of functionals of diffusion processes. *Stochastic Anal. Appl.*, 22(6):1553–1576, 2004.

[111] A. Rößler. Rooted tree analysis for order conditions of stochastic Runge-Kutta methods for the weak approximation of stochastic differential equations. *Stoch. Anal. Appl.*, 24(1):97–134, 2006.

[112] A. Rößler. Runge-Kutta methods for Itô stochastic differential equations with scalar noise. *BIT Numer. Math.*, 46(1):97–110, 2006.

[113] A. Rößler. Stochastic Taylor expansions for functionals of diffusion processes. *Stoch. Anal. Appl.*, 28(3):415–429, 2010.

[114] A. Rößler. Strong and weak approximation methods for stochastic differential equations—some recent developments. In *Recent developments in applied probability and statistics*, pages 127–153. Physica, Heidelberg, 2010.

[115] J.-P. Ryckaert, G. Ciccotti, and H. J. Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.*, 23(3):327–341, 1977.

[116] T. Sakai. *Riemannian geometry*, volume 149 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1996. Translated from the 1992 Japanese original by the author.

[117] U. Sharma and W. Zhang. Non-reversible sampling schemes on submanifolds. *Submitted*, 2020.

[118] D. Talay. Second-order discretization schemes of stochastic differential systems for the computation of the invariant law. *Stochastics: An International Journal of Probability and Stochastic Processes*, 29(1):13–36, 1990.

[119] D. Talay. Probabilistic numerical methods for partial differential equations: elements of analysis. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 148–196. Springer, Berlin, 1996.

[120] D. Talay. Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields*, 8(2):163–198, 2002. Inhomogeneous random systems (Cergy-Pontoise, 2001).

[121] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.

[122] G. Vilmart. Weak second order multirevolution composition methods for highly oscillatory stochastic differential equations with additive or multiplicative noise. *SIAM J. Sci. Comput.*, 36(4):A1770–A1796, 2014.

[123] G. Vilmart. Postprocessed integrators for the high order integration of ergodic SDEs. *SIAM J. Sci. Comput.*, 37(1):A201–A220, 2015.

[124] E. Zappa, M. Holmes-Cerfon, and J. Goodman. Monte Carlo on manifolds: sampling densities and integrating functions. *Comm. Pure Appl. Math.*, 71(12):2609–2647, 2018.

[125] W. Zhang. Ergodic SDEs on submanifolds and related numerical sampling schemes. *ESAIM Math. Model. Numer. Anal.*, 2019.

[126] K. C. Zygalakis. On the existence and the applications of modified equations for stochastic differential equations. *SIAM J. Sci. Comput.*, 33(1):102–130, 2011.

Appendices

APPENDIX A

Coefficients of the order two Runge-Kutta method

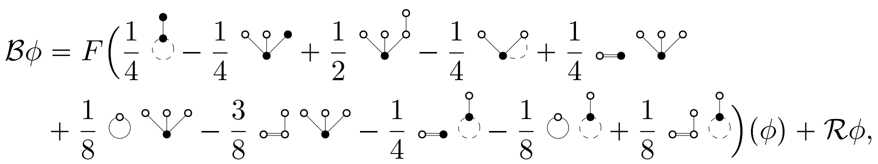
The coefficients of the Runge-Kutta method (4.4.7) used in Section 4.4.3 are

$$
\begin{aligned}
c_2 &= 0.621729189582953540, & c_3 &= 0.102032386582165330,\\
d_1 &= -0.898931652839146019, & d_2 &= -1.66233102561284629,\\
d_3 &= 0.318924515019668897, & \widehat{a}_{21} &= 0.584372887990673524,\\
\widehat{a}_{31} &= 0.887706593835748395, & \widehat{a}_{32} &= -0.345018694936693742,\\
\widehat{a}_{41} &= 0.0547449506054026516, & \widehat{a}_{42} &= -0.0205123070437693053,\\
\widehat{a}_{22} &= 1 - \widehat{a}_{21}, & \widehat{a}_{33} &= 1 - \widehat{a}_{31} - \widehat{a}_{32},\\
\widehat{a}_{43} &= 1 - \widehat{a}_{41} - \widehat{a}_{42}. & &
\end{aligned}
$$

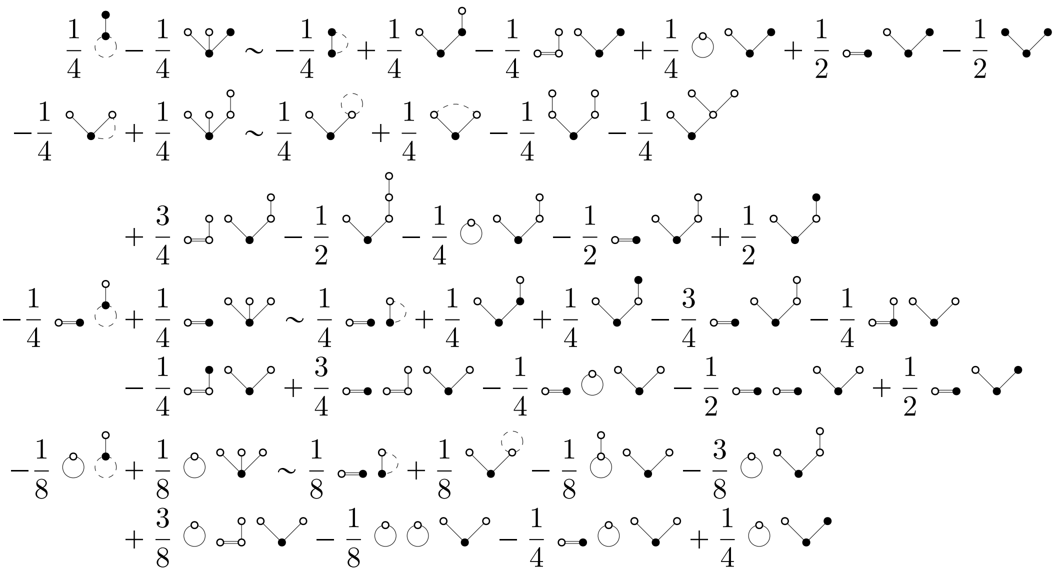
APPENDIX B

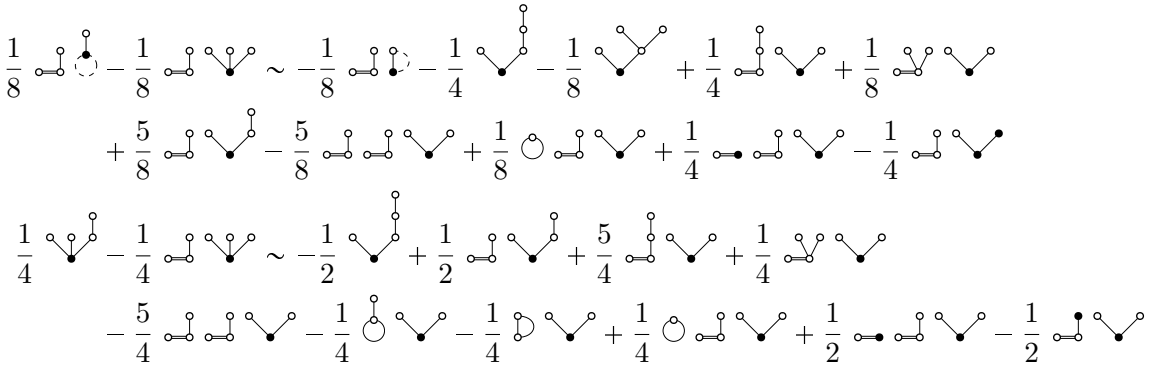
Integration by parts using the tree formalism

We provide here the detailed calculations of the integrations by parts of the order four and three terms, that are needed for the proof of Theorem 4.4.8. The integrations by parts of the order four terms are presented in the examples (3.3.6) and (3.3.9), and they allow us to rewrite $\int_{\mathcal{M}} \mathcal{A}_1 \phi d\mu_\infty$ as $\int_{\mathcal{M}} \mathcal{B} \phi d\mu_\infty$ where $\mathcal{B}$ is a differential operator of order three given by



and $\mathcal{R}$ is a differential operator of order two. Using Lemma 3.3.6 multiple times, we get the following integrations by parts of the order three terms of $\mathcal{B}\phi$.





APPENDIX C

Decomposition of the operators in exotic aromatic forests

Forest γ	Differential $F(\gamma)(\phi)$	Exact $e(\gamma)$	Numerical approximation $a(\gamma)$
Terms of order 4 w.r.t. ϕ			
	$\sigma^4\Delta^2\phi$	$\frac{1}{8}$	$\frac{1}{8}$
	$\sigma^4 G^{-1}\Delta\phi''(g,g)$	$-\frac{1}{4}$	$-\frac{1}{4}$
	$\sigma^4 G^{-2}\phi^{(4)}(g,g,g,g)$	$\frac{1}{8}$	$\frac{1}{8}$
Terms of order 3 w.r.t. ϕ			
	$\sigma^2\Delta\phi' f$	$\frac{1}{2}$	$\frac{1}{2}$
	$\sigma^2 G^{-1}\phi^{(3)}(g,g,f)$	$-\frac{1}{2}$	$-\frac{1}{2}$
	$\sigma^4 G^{-2}\phi^{(3)}(g,g,g'g)$	1	1
	$\sigma^4 G^{-1}\sum\phi^{(3)}(g,g'e_i,e_i)$	$-\frac{1}{2}$	$-\frac{1}{2}$
	$\sigma^2 G^{-2}(g,f)\phi^{(3)}(g,g,g)$	$\frac{1}{2}$	$\frac{1}{2}$
	$\sigma^4 G^{-2}\operatorname{div}(g)\phi^{(3)}(g,g,g)$	$\frac{1}{4}$	$\frac{1}{4}$
	$\sigma^4 G^{-2}(g,g'g)\phi^{(3)}(g,g,g)$	$-\frac{3}{4}$	$-\frac{3}{4}$
	$\sigma^2 G^{-1}(g,f)\Delta\phi'(g)$	$-\frac{1}{2}$	$-\frac{1}{2}$

Table C.1 (Part 1/8): Coefficients in exotic aromatic B-series of the operators $\mathcal{L}^2\phi/2 = \sum e(\gamma)F(\gamma)(\phi)$ and $\mathcal{A}_1\phi = \sum a(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).

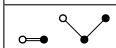
γ	$F(\gamma)(\phi)$	$e(\gamma)$	$a(\gamma)$
	$\sigma^4 G^{-1}\operatorname{div}(g)\Delta\phi'(g)$	$-\frac{1}{4}$	$-\frac{1}{4}$
	$\sigma^4 G^{-2}(g,g'g)\Delta\phi'(g)$	$\frac{1}{4}$	$\frac{1}{4}$
Terms of order 2 w.r.t. ϕ			
	$\phi''(f,f)$	$\frac{1}{2}$	$\frac{1}{2}$
	$\sigma^2\sum\phi''(f'e_i,e_i)$	$\frac{1}{2}$	$b^T d$
	$\sigma^2 G^{-1}\phi''(g,g'f)$	-1	-1
	$\sigma^2 G^{-1}\phi''(g,f'g)$	-1	$-b^T d-\widehat{b}^T d$
	$\sigma^4 G^{-2}\phi''(g,g'g'g)$	$\frac{3}{2}$	$-2\widehat{b}^T(d\blacklozenge\widehat{A}d)-(\widehat{b}^T d)^2+2\widehat{b}^T d+1$
	$\sigma^2 G^{-2}\phi''(g'g,g'g)$	$\frac{1}{4}$	$-\widehat{b}^T(d\blacklozenge\widehat{A}d)-\frac{1}{2}(\widehat{b}^T d)^2+\widehat{b}^T d$
	$\sigma^4 G^{-2}\phi''(g,g''(g,g))$	$\frac{1}{2}$	$2\widehat{b}^T d^{\blacklozenge 2}-2\widehat{b}^T(\delta\blacklozenge d^{\blacklozenge 2})+\frac{1}{2}$
	$\sigma^4 G^{-1}\phi''(g,\Delta g)$	$-\frac{1}{2}$	$-\frac{1}{2}$
	$\sigma^4 G^{-1}\sum\phi''(g'e_i,g'e_i)$	$-\frac{1}{4}$	$\widehat{b}^T(d\blacklozenge\widehat{A}d)+\frac{1}{2}(\widehat{b}^T d)^2-\widehat{b}^T d$
	$\sigma^4 G^{-1}\sum\phi''(g''(g,e_i),e_i)$	0	$\widehat{b}^T(\delta\blacklozenge d^{\blacklozenge 2})-\widehat{b}^T d^{\blacklozenge 2}$
	$G^{-1}(g,f)\phi''(g,f)$	-1	-1
	$\sigma^2 G^{-1}\operatorname{div}(g)\phi''(g,f)$	$-\frac{1}{2}$	$-\frac{1}{2}$
	$\sigma^2 G^{-2}(g,g'g)\phi''(g,f)$	$\frac{1}{2}$	$\frac{1}{2}$
	$\sigma^2 G^{-1}(g,f)\sum\phi''(e_i,g'e_i)$	$-\frac{1}{2}$	$-\widehat{b}^T d$
	$\sigma^4 G^{-1}\operatorname{div}(g)\sum\phi''(e_i,g'e_i)$	$-\frac{1}{4}$	$-\frac{1}{2}\widehat{b}^T d$
	$\sigma^4 G^{-2}(g,g'g)\sum\phi''(e_i,g'e_i)$	$\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T d$
	$\sigma^2 G^{-2}(g,f)\phi''(g,g'g)$	2	$2\widehat{b}^T d+1$
	$\sigma^4 G^{-2}\operatorname{div}(g)\phi''(g,g'g)$	1	$\widehat{b}^T d+\frac{1}{2}$
	$\sigma^4 G^{-3}(g,g'g)\phi''(g,g'g)$	$-\frac{5}{2}$	$2\widehat{b}^T(d\blacklozenge\widehat{A}d)+(\widehat{b}^T d)^2-3\widehat{b}^T d-\frac{3}{2}$

Table C.1 (Part 2/8): Coefficients in exotic aromatic B-series of the operators $\mathcal{L}^2\phi/2=\sum e(\gamma)F(\gamma)(\phi)$ and $\mathcal{A}_1\phi=\sum a(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).

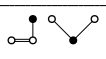
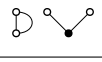
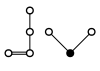
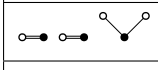
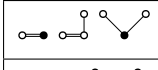
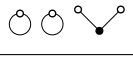
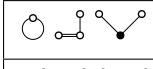
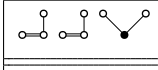
γ	$e(\gamma)$	$a(\gamma)$
	1	1
	$\frac{1}{2}$	$b^T d$
	$\frac{1}{2}$	$\frac{1}{2}$
	$\frac{1}{4}$	$\frac{1}{4}$
	$-\frac{7}{4}$	$\widehat{b}^T(d \blacklozenge \widehat{A}d) + \frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d - \frac{3}{2}$
	$-\frac{1}{2}$	$\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) - \widehat{b}^T d^{\blacklozenge 2} - \frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$
	-2	$-\widehat{b}^T d - \frac{3}{2}$
	$\frac{1}{8}$	$\frac{1}{8}$
	-1	$-\frac{1}{2}\widehat{b}^T d - \frac{3}{4}$
	$\frac{19}{8}$	$-\widehat{b}^T(d \blacklozenge \widehat{A}d) - \frac{1}{2}(\widehat{b}^T d)^2 + \frac{3}{2}\widehat{b}^T d + \frac{15}{8}$
Terms of order 1 w.r.t. ϕ		
	$\frac{1}{2}$	$b^T c$
	$\frac{1}{4}$	$\frac{1}{2} b^T d^{\blacklozenge 2}$
	$-\frac{1}{2}$	$\widehat{b}^T(d \blacklozenge \widehat{A}c) - \widehat{b}^T d$
	$-\frac{1}{2}$	$\widehat{b}^T A((\delta - \mathbb{1}) \blacklozenge d) - \widehat{b}^T d b^T d$
	0	$b^T(d \blacklozenge \widehat{A}((\delta - \mathbb{1}) \blacklozenge d))$
	$-\frac{1}{4}$	$-\frac{1}{2} b^T(\delta \blacklozenge d^{\blacklozenge 2})$
	0	$\widehat{b}^T(\delta \blacklozenge c \blacklozenge d) - \widehat{b}^T(c \blacklozenge d)$
	$-\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}d^{\blacklozenge 2}) - \frac{1}{2}\widehat{b}^T d$

Table C.1 (Part 3/8): Coefficients in exotic aromatic B-series of the operators $\mathcal{L}^2\phi/2 = \sum e(\gamma)F(\gamma)(\phi)$ and $\mathcal{A}_1\phi = \sum a(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).

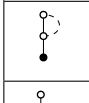
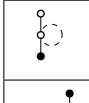
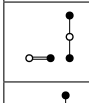
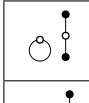
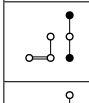
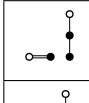
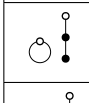
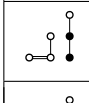
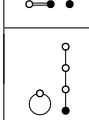
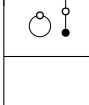
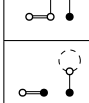
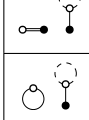
γ	$e(\gamma)$	$a(\gamma)$
	$\frac{1}{2}$	$-\widehat{b}^T(d \blacklozenge (\widehat{A}d)^{\blacklozenge 2}) - 3\widehat{b}^T(d \blacklozenge \widehat{A}(d \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)))$ $-2\widehat{b}^T d\widehat{b}^T(d \blacklozenge \widehat{A}d) + (\widehat{b}^T d)^2 + \widehat{b}^T d$
	$\frac{1}{4}$	$\widehat{b}^T(d \blacklozenge \widehat{A}d^{\blacklozenge 2}) - \frac{3}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d^{\blacklozenge 2})) + \widehat{b}^T d(\widehat{b}^T d^{\blacklozenge 2} - \widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2})) + \frac{1}{2}\widehat{b}^T d$
	0	$-2\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}d) + (\widehat{b}^T d + 1)\widehat{b}^T d^{\blacklozenge 2} - \widehat{b}^T d\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2})$
	0	$\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}d) - \frac{1}{2}\widehat{b}^T d^{\blacklozenge 2}$
	0	$\frac{1}{2}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 3}) - \frac{1}{2}\widehat{b}^T d^{\blacklozenge 3}$
	$-\frac{1}{2}$	$-\widehat{b}^T c$
	$-\frac{1}{4}$	$-\frac{1}{2}\widehat{b}^T c$
	$\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T c$
	$-\frac{1}{2}$	$-b^T(\delta \blacklozenge c)$
	$-\frac{1}{4}$	$-\frac{1}{2}b^T(\delta \blacklozenge d^{\blacklozenge 2})$
	$\frac{1}{4}$	$b^T(d \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d))) + \frac{1}{2}b^T(\delta \blacklozenge d^{\blacklozenge 2})$
	1	$\widehat{b}^T(c \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)) + (\widehat{b}^T d)^2 + \widehat{b}^T d$ $+\widehat{b}^T(d \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)) - \widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge c))$
	$\frac{1}{2}$	$\frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)) + \frac{1}{2}(\widehat{b}^T d)^2 + \frac{1}{2}\widehat{b}^T d$ $+\frac{1}{2}\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)) - \frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d^{\blacklozenge 2}))$
	-1	$3\widehat{b}^T(d \blacklozenge \widehat{A}(d \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d))) + (2\widehat{b}^T d - \frac{1}{2})\widehat{b}^T(d \blacklozenge \widehat{A}d)$ $-3\widehat{b}^T(d \blacklozenge (\widehat{A}((\mathbb{1} - \delta) \blacklozenge d))^{\blacklozenge 2}) + \frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d)) + \widehat{b}^T(d \blacklozenge (\widehat{A}d)^{\blacklozenge 2})$ $+\frac{1}{2}\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}((\delta - \mathbb{1}) \blacklozenge d)) + \frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d^{\blacklozenge 2})) - \frac{3}{2}(\widehat{b}^T d)^2 - \frac{3}{2}\widehat{b}^T d$
	$-\frac{1}{4}$	$-\frac{1}{2}\widehat{b}^T d^{\blacklozenge 2}$
	$-\frac{1}{8}$	$-\frac{1}{4}\widehat{b}^T d^{\blacklozenge 2}$

Table C.1 (Part 4/8): Coefficients in exotic aromatic B-series of the operators $\mathcal{L}^2\phi/2 = \sum e(\gamma)F(\gamma)(\phi)$ and $\mathcal{A}_1\phi = \sum a(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).

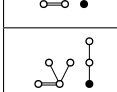
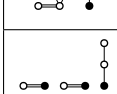
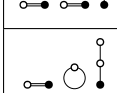
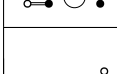
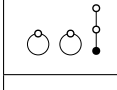
γ	$e(\gamma)$	$a(\gamma)$
	$\frac{1}{8}$	$\frac{1}{4}\widehat{b}^T d^{\bullet 2}$
	$\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$\frac{1}{8}$	$\frac{1}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$-\frac{1}{8}$	$\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) - (\widehat{b}^T d + \frac{1}{2})\widehat{b}^T d^{\bullet 2} + (\widehat{b}^T d - \frac{1}{4})\widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$\frac{1}{2}$	$-\widehat{b}^T(d \bullet \widehat{A}c) + \widehat{b}^T d$
	$\frac{1}{2}$	$\widehat{b}^T(\delta \bullet A((\mathbb{1} - \delta) \bullet d)) + \widehat{b}^T d b^T d$
	$\frac{1}{4}$	$-\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}d^{\bullet 2}) + \frac{1}{2}\widehat{b}^T d$
	0	$-\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) + \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$-\frac{1}{2}$	$\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) + 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)))$ $+2\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) + 2\widehat{b}^T d\widehat{b}^T(d \bullet \widehat{A}d) - \widehat{b}^T(\delta \bullet d^{\bullet 2}) - (\widehat{b}^T d)^2 - \widehat{b}^T d$
	$-\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((3\delta - 2 \cdot \mathbb{1}) \bullet d^{\bullet 2})) + \widehat{b}^T d(\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \widehat{b}^T d^{\bullet 2}) - \frac{1}{2}\widehat{b}^T d$
	$\frac{1}{2}$	$\widehat{b}^T(\delta \bullet c)$
	$\frac{1}{2}$	$\frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{1}{2}\widehat{b}^T(\delta \bullet c)$
	$-\frac{3}{2}$	$\widehat{b}^T(c \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \widehat{b}^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d))$ $+\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet c)) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}\widehat{b}^T(\delta \bullet c) - (\widehat{b}^T d)^2 - \widehat{b}^T d$
	$\frac{1}{8}$	$\frac{1}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$-\frac{3}{4}$	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2}))$ $+\frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}(\widehat{b}^T d)^2 - \frac{1}{2}\widehat{b}^T d$
	$\frac{9}{8}$	$3\widehat{b}^T(d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) - \frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\delta + \mathbb{1}) \bullet d))$ $-\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) - 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)))$ $+(\frac{1}{2} - 2\widehat{b}^T d)\widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d)) - \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2}))$ $+\frac{9}{2}\widehat{b}^T(\delta \bullet d^{\bullet 3}) - \frac{15}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{3}{2}(\widehat{b}^T d)^2 + \frac{3}{2}\widehat{b}^T d$
	$-\frac{1}{2}$	$-b^T c$

Table C.1 (Part 5/8): Coefficients in exotic aromatic B-series of the operators $\mathcal{L}^2\phi/2 = \sum e(\gamma)F(\gamma)(\phi)$ and $\mathcal{A}_1\phi = \sum a(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).

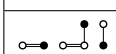
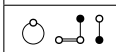
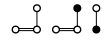
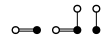
γ	$e(\gamma)$	$a(\gamma)$
	$-\frac{1}{2}$	$-\frac{1}{2}$
	$-\frac{1}{2}$	$-\frac{1}{2}$
	$-\frac{1}{4}$	$-\frac{1}{2}b^T d^{\blacklozenge 2}$
	$-\frac{1}{2}$	$-b^T d$
	$\frac{3}{2}$	$-\widehat{b}^T(d \blacklozenge \widehat{A}c) + \widehat{b}^T d + 1$
	$\frac{3}{2}$	$b^T(d \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)) + \widehat{b}^T A((\mathbb{1} - \delta) \blacklozenge d) + (\widehat{b}^T d + 2)b^T d$
	$\frac{1}{2}$	$\widehat{b}^T(c \blacklozenge d) - \widehat{b}^T(\delta \blacklozenge c \blacklozenge d) + \frac{1}{2}$
	$\frac{1}{4}$	$\frac{1}{2}b^T(\delta \blacklozenge d^{\blacklozenge 2})$
	$\frac{3}{4}$	$-\frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}d^{\blacklozenge 2}) + \frac{1}{2}\widehat{b}^T d + \frac{1}{2}$
	$\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T d^{\blacklozenge 3} - \frac{1}{2}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 3}) + \frac{1}{4}$
	$\frac{1}{2}$	$-\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}d) + \frac{3}{2}\widehat{b}^T d^{\blacklozenge 2} - \widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) + \frac{1}{2}$
	$\frac{1}{4}$	$-\widehat{b}^T(d \blacklozenge \widehat{A}d) - \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d$
	$-\frac{9}{4}$	$\widehat{b}^T(d \blacklozenge (\widehat{A}d)^{\blacklozenge 2}) + 3\widehat{b}^T(d \blacklozenge \widehat{A}(d \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)))$ $+(2\widehat{b}^T d + 3)\widehat{b}^T(d \blacklozenge \widehat{A}d) + \frac{1}{2}(\widehat{b}^T d)^2 - 4\widehat{b}^T d - 1$
	$-\frac{7}{4}$	$\frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}((3\delta - 2 \cdot \mathbb{1}) \blacklozenge d^{\blacklozenge 2})) + 2\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}d)$ $-(2\widehat{b}^T d + 3)\widehat{b}^T d^{\blacklozenge 2} + 2(\widehat{b}^T d + 1)\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) - \frac{1}{2}\widehat{b}^T d - \frac{3}{2}$
	$-\frac{1}{8}$	$-\frac{1}{8}$
	$-\frac{1}{8}$	$-\frac{1}{8}$
	$\frac{3}{2}$	$\widehat{b}^T c + 1$
	$\frac{3}{4}$	$\frac{1}{2}\widehat{b}^T c + \frac{1}{2}$
	$-\frac{9}{4}$	$\widehat{b}^T(d \blacklozenge \widehat{A}c) - \frac{1}{2}\widehat{b}^T c - \widehat{b}^T d - \frac{3}{2}$
	$\frac{1}{2}$	$b^T(\delta \blacklozenge c)$
	$\frac{1}{4}$	$\frac{1}{2}b^T(\delta \blacklozenge d^{\blacklozenge 2})$

Table C.1 (Part 6/8): Coefficients in exotic aromatic B-series of the operators $\mathcal{L}^2\phi/2 = \sum e(\gamma)F(\gamma)(\phi)$ and $\mathcal{A}_1\phi = \sum a(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).

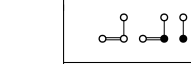
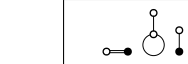
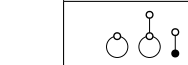
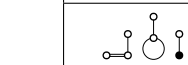
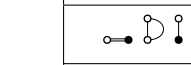
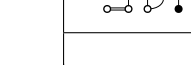
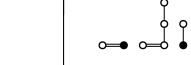
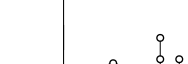
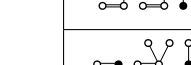
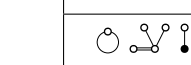
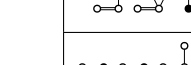
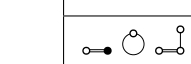
γ	$e(\gamma)$	$a(\gamma)$
	$-\frac{5}{4}$	$\widehat{b}^T(\delta \bullet A((\delta - \mathbb{1}) \bullet d))$ $+b^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) - \frac{1}{2}b^T(\delta \bullet d^{\bullet 2}) - (\widehat{b}^T d + 1)b^T d$
	$\frac{3}{4}$	$\frac{1}{2}\widehat{b}^T d^{\bullet 2} + \frac{1}{2}$
	$\frac{3}{8}$	$\frac{1}{4}\widehat{b}^T d^{\bullet 2} + \frac{1}{4}$
	$-\frac{9}{8}$	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}d^{\bullet 2}) - \frac{1}{4}\widehat{b}^T d^{\bullet 2} - \frac{1}{2}\widehat{b}^T d - \frac{3}{4}$
	$\frac{1}{2}$	$\widehat{b}^T d$
	$\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T d$
	$-\frac{1}{2}$	$\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}\widehat{b}^T d - \frac{1}{4}$
	-3	$\widehat{b}^T(c \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \widehat{b}^T(d \bullet \widehat{A}(\delta \bullet c))$ $+\widehat{b}^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) - (\widehat{b}^T d)^2 - 3\widehat{b}^T d - 1$
	$-\frac{3}{2}$	$\frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2}))$ $+\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) - \frac{1}{2}(\widehat{b}^T d)^2 - \frac{3}{2}\widehat{b}^T d - \frac{1}{2}$
	$\frac{23}{4}$	$3\widehat{b}^T(d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) - \frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((3 \cdot \mathbb{1} + \delta) \bullet d))$ $-2\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) - 6\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)))$ $-\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) - \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d))$ $-(4\widehat{b}^T d + \frac{5}{2})\widehat{b}^T(d \bullet \widehat{A}d) + \widehat{b}^T(\delta \bullet d^{\bullet 2}) + (\widehat{b}^T d)^2 + \frac{13}{2}\widehat{b}^T d + 3$
	$-\frac{3}{4}$	$-\frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}$
	$-\frac{3}{8}$	$-\frac{1}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{4}$
	$\frac{13}{8}$	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((2 \cdot \mathbb{1} - 3\delta) \bullet d^{\bullet 2})) - \widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d)$ $+(2\widehat{b}^T d + \frac{3}{2})\widehat{b}^T d^{\bullet 2} - (2\widehat{b}^T d + \frac{3}{4})\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{1}{2}\widehat{b}^T d + \frac{5}{4}$
	-1	$-\widehat{b}^T(\delta \bullet c) - \frac{1}{2}$
	-1	$-\frac{1}{2}\widehat{b}^T(\delta \bullet c) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}$
	$\frac{7}{2}$	$\widehat{b}^T(c \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) + \widehat{b}^T(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))$ $-\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet c)) + \frac{1}{2}\widehat{b}^T(\delta \bullet c) + \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + (\widehat{b}^T d)^2 + 2\widehat{b}^T d + \frac{3}{2}$
	$-\frac{1}{4}$	$-\frac{1}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{8}$
	$\frac{7}{4}$	$\frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) - \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2}))$ $+\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) + \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d + \frac{3}{4}$

Table C.1 (Part 7/8): Coefficients in exotic aromatic B-series of the operators $\mathcal{L}^2\phi/2 = \sum e(\gamma)F(\gamma)(\phi)$ and $\mathcal{A}_1\phi = \sum a(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).

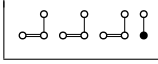
γ	$e(\gamma)$	$a(\gamma)$
	$-\frac{7}{2}$	$\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) + 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)))$ $-3\widehat{b}^T(d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) + (2\widehat{b}^T d + \frac{1}{2})\widehat{b}^T(d \bullet \widehat{A}d)$ $+\frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\mathbb{1} + \delta) \bullet d)) + \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) - \frac{9}{2}\widehat{b}^T(\delta \bullet d^{\bullet 3})$ $+\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d)) + \frac{15}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - (\widehat{b}^T d)^2 - 3\widehat{b}^T d - \frac{15}{8}$

Table C.1 (Part 8/8): Coefficients in exotic aromatic B-series of the operators $\mathcal{L}^2\phi/2 = \sum e(\gamma)F(\gamma)(\phi)$ and $\mathcal{A}_1\phi = \sum a(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).

γ	$a^0(\gamma)$
	$b^T d - \widehat{b}^T d$
	$b^T c - 2b^T d + \frac{1}{2}$
	$\frac{1}{2}b^T d^{\bullet 2} - b^T d + \frac{1}{4}$
	$\widehat{b}^T(d \bullet \widehat{A}c) - 2\widehat{b}^T(d \bullet \widehat{A}d) - (\widehat{b}^T d)^2 + 2\widehat{b}^T d - \frac{1}{2}$
	$\widehat{b}^T A((\delta - \mathbb{1}) \bullet d)) - \widehat{b}^T d b^T d + 2\widehat{b}^T d - \frac{1}{2}$
	$b^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d))$
	$-\frac{1}{2}b^T(\delta \bullet d^{\bullet 2}) + b^T d - \frac{1}{4}$
	$\widehat{b}^T(\delta \bullet c \bullet d) - \widehat{b}^T(c \bullet d) + 2\widehat{b}^T d^{\bullet 2} - 2\widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}d^{\bullet 2}) - \widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d - \frac{1}{4}$
	$-\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) - 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)))$ $+(4 - 2\widehat{b}^T d)\widehat{b}^T(d \bullet \widehat{A}d) + 3(\widehat{b}^T d)^2 - 4\widehat{b}^T d + 1$
	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((2 \cdot \mathbb{1} - 3\delta) \bullet d^{\bullet 2})) + \widehat{b}^T(d \bullet \widehat{A}d)$ $+(\widehat{b}^T d - 1)(\widehat{b}^T d^{\bullet 2} - \widehat{b}^T(\delta \bullet d^{\bullet 2})) + \frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d + \frac{1}{4}$
	$-2\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) + 2\widehat{b}^T(d \bullet \widehat{A}d)$ $+(\widehat{b}^T d - 2)\widehat{b}^T d^{\bullet 2} + (3 - \widehat{b}^T d)\widehat{b}^T(\delta \bullet d^{\bullet 2}) + (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2}$

Table C.2 (Part 1/5): Coefficients in exotic aromatic B-series of the operator $\mathcal{A}_1^0\phi = \sum a^0(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).

γ	$a^0(\gamma)$
	$\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}d) - \widehat{b}^T(d \blacklozenge \widehat{A}d) + \frac{1}{2}\widehat{b}^T d^{\blacklozenge 2} - \widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) - \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d - \frac{1}{4}$
	$\frac{1}{2}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 3}) - \frac{1}{2}\widehat{b}^T d^{\blacklozenge 3} + \widehat{b}^T d^{\blacklozenge 2} - \widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2})$
	$-\widehat{b}^T d^{\blacklozenge 2} + \widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2})$
	$-\widehat{b}^T c + 2\widehat{b}^T d - \frac{1}{2}$
	$-\frac{1}{2}\widehat{b}^T c + \widehat{b}^T d - \frac{1}{4}$
	$\frac{1}{2}\widehat{b}^T c - \widehat{b}^T d + \frac{1}{4}$
	$-b^T(\delta \blacklozenge c) + 2b^T d - \frac{1}{2}$
	$-\frac{1}{2}b^T(\delta \blacklozenge d^{\blacklozenge 2}) + b^T d - \frac{1}{4}$
	$b^T(d \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)) + \frac{1}{2}b^T(\delta \blacklozenge d^{\blacklozenge 2}) - b^T d + \frac{1}{4}$
	$\widehat{b}^T(d \blacklozenge \widehat{A}((3 \cdot \mathbb{1} - \delta) \blacklozenge d)) - \widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge c))$ $+\widehat{b}^T(c \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)) + 2(\widehat{b}^T d)^2 - 4\widehat{b}^T d + 1$
	$\frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}((3 \cdot \mathbb{1} - \delta) \blacklozenge d)) - \frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d^{\blacklozenge 2}))$ $+\frac{1}{2}\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)) + (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2}$
	$\widehat{b}^T(d \blacklozenge (\widehat{A}d)^{\blacklozenge 2}) + 3\widehat{b}^T(d \blacklozenge \widehat{A}(d \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d))) + (2\widehat{b}^T d - \frac{11}{2})\widehat{b}^T(d \blacklozenge \widehat{A}d)$ $-3\widehat{b}^T(d \blacklozenge (\widehat{A}((\mathbb{1} - \delta) \blacklozenge d))^{\blacklozenge 2}) - \frac{1}{2}\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d))$ $+\frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d)) + \frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d^{\blacklozenge 2})) - 4(\widehat{b}^T d)^2 + 6\widehat{b}^T d - \frac{3}{2}$
	$-\frac{1}{2}\widehat{b}^T d^{\blacklozenge 2} + \widehat{b}^T d - \frac{1}{4}$
	$-\frac{1}{4}\widehat{b}^T d^{\blacklozenge 2} + \frac{1}{2}\widehat{b}^T d - \frac{1}{8}$
	$\frac{1}{4}\widehat{b}^T d^{\blacklozenge 2} - \frac{1}{2}\widehat{b}^T d + \frac{1}{8}$
	$-2\widehat{b}^T d^{\blacklozenge 2} + \frac{5}{2}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) - \widehat{b}^T d + \frac{1}{4}$
	$\frac{5}{4}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) - \widehat{b}^T d^{\blacklozenge 2} - \frac{1}{2}\widehat{b}^T d + \frac{1}{8}$

Table C.2 (Part 2/5): Coefficients in exotic aromatic B-series of the operator $\mathcal{A}_1^0\phi = \sum a^0(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).

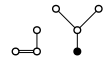
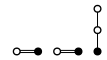
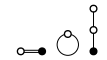
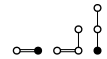
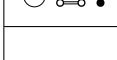
γ	$a^0(\gamma)$
	$(\frac{5}{2} - \widehat{b}^T d)\widehat{b}^T d^{\bullet 2} + (\widehat{b}^T d - \frac{13}{4})\widehat{b}^T(\delta \bullet d^{\bullet 2})$ $+\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) - \widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}(\widehat{b}^T d)^2 + \frac{3}{2}\widehat{b}^T d - \frac{3}{8}$
	$-\widehat{b}^T(d \bullet \widehat{A}c) + 2\widehat{b}^T(d \bullet \widehat{A}d) + (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2}$
	$\widehat{b}^T(\delta \bullet A((\mathbb{1} - \delta) \bullet d)) + \widehat{b}^T d b^T d - 2\widehat{b}^T d + \frac{1}{2}$
	$-\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}d^{\bullet 2}) + \widehat{b}^T(d \bullet \widehat{A}d) + \frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d + \frac{1}{4}$
	$-\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) + \widehat{b}^T(d \bullet \widehat{A}d) + \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d + \frac{1}{4}$
	$\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) + 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))) + 2\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d)$ $+(2\widehat{b}^T d - 6)\widehat{b}^T(d \bullet \widehat{A}d) - \widehat{b}^T(\delta \bullet d^{\bullet 2}) - 4(\widehat{b}^T d)^2 + 6\widehat{b}^T d - \frac{3}{2}$
	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((3\delta - 2 \cdot \mathbb{1}) \bullet d^{\bullet 2})) - \widehat{b}^T(d \bullet \widehat{A}d)$ $+(1 - \widehat{b}^T d)(\widehat{b}^T d^{\bullet 2} - \widehat{b}^T(\delta \bullet d^{\bullet 2})) - \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d - \frac{1}{4}$
	$\widehat{b}^T(\delta \bullet c) - 2\widehat{b}^T d + \frac{1}{2}$
	$\frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{1}{2}\widehat{b}^T(\delta \bullet c) - 2\widehat{b}^T d + \frac{1}{2}$
	$\widehat{b}^T(c \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \widehat{b}^T(d \bullet \widehat{A}((\delta - 3 \cdot \mathbb{1}) \bullet d))$ $+\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet c)) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}\widehat{b}^T(\delta \bullet c) - 2(\widehat{b}^T d)^2 + 6\widehat{b}^T d - \frac{3}{2}$
	$\frac{1}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}\widehat{b}^T d + \frac{1}{8}$
	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((\delta - 3 \cdot \mathbb{1}) \bullet d)) + \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2}))$ $+\frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - (\widehat{b}^T d)^2 + 3\widehat{b}^T d - \frac{3}{4}$
	$3\widehat{b}^T(d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) + (\frac{13}{2} - 2\widehat{b}^T d)\widehat{b}^T(d \bullet \widehat{A}d)$ $-\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) - 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)))$ $-\frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\delta + \mathbb{1}) \bullet d)) - \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) + \frac{9}{2}\widehat{b}^T(\delta \bullet d^{\bullet 3})$ $-\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d)) - \frac{15}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{9}{2}(\widehat{b}^T d)^2 - \frac{15}{2}\widehat{b}^T d + \frac{15}{8}$
	$-b^T c + 2\widehat{b}^T d - \frac{1}{2}$
	$-\frac{1}{2}b^T d^{\bullet 2} + \widehat{b}^T d - \frac{1}{4}$
	$-b^T d + \widehat{b}^T d$

Table C.2 (Part 3/5): Coefficients in exotic aromatic B-series of the operator $\mathcal{A}_1^0\phi = \sum a^0(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).

γ	$a^0(\gamma)$
	$-\widehat{b}^T(d \bullet \widehat{A}c) + 2\widehat{b}^T(d \bullet \widehat{A}d) + (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2}$
	$b^T(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) + \widehat{b}^T A((\mathbb{1} - \delta) \bullet d) + (\widehat{b}^T d + 2)b^T d - 4\widehat{b}^T d + \frac{1}{2}$
	$\widehat{b}^T(c \bullet d) - \widehat{b}^T(\delta \bullet c \bullet d) - 2\widehat{b}^T d^{\bullet 2} + 2\widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$\frac{1}{2}b^T(\delta \bullet d^{\bullet 2}) - \widehat{b}^T d + \frac{1}{4}$
	$-\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}d^{\bullet 2}) + \widehat{b}^T(d \bullet \widehat{A}d) + \frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d + \frac{1}{4}$
	$\frac{1}{2}\widehat{b}^T d^{\bullet 3} - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 3}) - \widehat{b}^T d^{\bullet 2} + \widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$-\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) + \widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}\widehat{b}^T d^{\bullet 2} + \widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d + \frac{1}{4}$
	$\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) + 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)))$ $+(2\widehat{b}^T d - 4)\widehat{b}^T(d \bullet \widehat{A}d) - 3(\widehat{b}^T d)^2 + 4\widehat{b}^T d - 1$
	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((3\delta - 2 \cdot \mathbb{1}) \bullet d^{\bullet 2})) + 2\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) - 3\widehat{b}^T(d \bullet \widehat{A}d)$ $+(3 - 2\widehat{b}^T d)\widehat{b}^T d^{\bullet 2} + (2\widehat{b}^T d - 4)\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{3}{2}(\widehat{b}^T d)^2 + 3\widehat{b}^T d - \frac{3}{4}$
	$\widehat{b}^T d^{\bullet 2} - \widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$\widehat{b}^T c - 2\widehat{b}^T d + \frac{1}{2}$
	$\frac{1}{2}\widehat{b}^T c - \widehat{b}^T d + \frac{1}{4}$
	$\widehat{b}^T(d \bullet \widehat{A}c) - 2\widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}\widehat{b}^T c - (\widehat{b}^T d)^2 + 3\widehat{b}^T d - \frac{3}{4}$
	$b^T(\delta \bullet c) - 2\widehat{b}^T d + \frac{1}{2}$
	$\frac{1}{2}b^T(\delta \bullet d^{\bullet 2}) - \widehat{b}^T d + \frac{1}{4}$
	$b^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \widehat{b}^T(\delta \bullet A((\delta - \mathbb{1}) \bullet d))$ $-\frac{1}{2}b^T(\delta \bullet d^{\bullet 2}) - (\widehat{b}^T d + 1)b^T d + 4\widehat{b}^T d - \frac{3}{4}$
	$\frac{1}{2}\widehat{b}^T d^{\bullet 2} - \widehat{b}^T d + \frac{1}{4}$
	$\frac{1}{4}\widehat{b}^T d^{\bullet 2} - \frac{1}{2}\widehat{b}^T d + \frac{1}{8}$
	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}d^{\bullet 2}) - \widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{4}\widehat{b}^T d^{\bullet 2} - \frac{1}{2}(\widehat{b}^T d)^2 + \frac{3}{2}\widehat{b}^T d - \frac{3}{8}$
	$\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) - \widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d - \frac{1}{4}$

Table C.2 (Part 4/5): Coefficients in exotic aromatic B-series of the operator $\mathcal{A}_1^0\phi = \sum a^0(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).

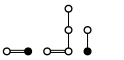
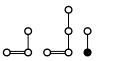
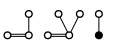
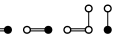
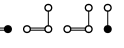
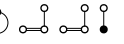
γ	$a^0(\gamma)$
	$\widehat{b}^T(c \blacklozenge \widehat{A}((\delta - \mathbb{1}) \blacklozenge d)) + \widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge c))$ $+\widehat{b}^T(d \blacklozenge \widehat{A}((\delta - 3 \cdot \mathbb{1}) \blacklozenge d)) - 2(\widehat{b}^T d)^2 + 4\widehat{b}^T d - 1$
	$\frac{1}{2}\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}((\delta - \mathbb{1}) \blacklozenge d)) + \frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d^{\blacklozenge 2}))$ $+\frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}((\delta - 3 \cdot \mathbb{1}) \blacklozenge d)) - (\widehat{b}^T d)^2 + 2\widehat{b}^T d - \frac{1}{2}$
	$3\widehat{b}^T(d \blacklozenge (\widehat{A}((\mathbb{1} - \delta) \blacklozenge d))^{\blacklozenge 2}) - \frac{1}{2}\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}((3 \cdot \mathbb{1} + \delta) \blacklozenge d))$ $-2\widehat{b}^T(d \blacklozenge (\widehat{A}d)^{\blacklozenge 2}) - 6\widehat{b}^T(d \blacklozenge \widehat{A}(d \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d))) + (\frac{23}{2} - 4\widehat{b}^T d)\widehat{b}^T(d \blacklozenge \widehat{A}d)$ $-\frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d)) - \frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d^{\blacklozenge 2})) + \widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) + 8(\widehat{b}^T d)^2 - 12\widehat{b}^T d + 3$
	$2\widehat{b}^T d^{\blacklozenge 2} - \frac{5}{2}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) + \widehat{b}^T d - \frac{1}{4}$
	$\widehat{b}^T d^{\blacklozenge 2} - \frac{5}{4}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) + \frac{1}{2}\widehat{b}^T d - \frac{1}{8}$
	$\frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}((2 \cdot \mathbb{1} - 3\delta) \blacklozenge d^{\blacklozenge 2})) - \widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}d) + 2\widehat{b}^T(d \blacklozenge \widehat{A}d)$ $+(2\widehat{b}^T d - \frac{7}{2})\widehat{b}^T d^{\blacklozenge 2} + (\frac{17}{4} - 2\widehat{b}^T d)\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) + (\widehat{b}^T d)^2 - \frac{5}{2}\widehat{b}^T d + \frac{5}{8}$
	$-\widehat{b}^T(\delta \blacklozenge c) + 2\widehat{b}^T d - \frac{1}{2}$
	$-\frac{1}{2}\widehat{b}^T(\delta \blacklozenge c) - \frac{1}{2}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) + 2\widehat{b}^T d - \frac{1}{2}$
	$\widehat{b}^T(c \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)) + \widehat{b}^T(d \blacklozenge \widehat{A}((3 \cdot \mathbb{1} - \delta) \blacklozenge d))$ $-\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge c)) + \frac{1}{2}\widehat{b}^T(\delta \blacklozenge c) + \frac{1}{2}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) + 2(\widehat{b}^T d)^2 - 6\widehat{b}^T d + \frac{3}{2}$
	$-\frac{1}{4}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) + \frac{1}{2}\widehat{b}^T d - \frac{1}{8}$
	$\frac{1}{2}\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)) - \frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d^{\blacklozenge 2}))$ $+\frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}((3 \cdot \mathbb{1} - \delta) \blacklozenge d)) + \frac{1}{2}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) + (\widehat{b}^T d)^2 - 3\widehat{b}^T d + \frac{3}{4}$
	$\widehat{b}^T(d \blacklozenge (\widehat{A}d)^{\blacklozenge 2}) + 3\widehat{b}^T(d \blacklozenge \widehat{A}(d \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d)))$ $-3\widehat{b}^T(d \blacklozenge (\widehat{A}((\mathbb{1} - \delta) \blacklozenge d))^{\blacklozenge 2}) + \frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d))$ $+(2\widehat{b}^T d - \frac{13}{2})\widehat{b}^T(d \blacklozenge \widehat{A}d) + \frac{1}{2}\widehat{b}^T(d^{\blacklozenge 2} \blacklozenge \widehat{A}((\mathbb{1} + \delta) \blacklozenge d))$ $+\frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}(\delta \blacklozenge d^{\blacklozenge 2})) - \frac{9}{2}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 3}) + \frac{15}{4}\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) - \frac{9}{2}(\widehat{b}^T d)^2 + \frac{15}{2}\widehat{b}^T d - \frac{15}{8}$

Table C.2 (Part 5/5): Coefficients in exotic aromatic B-series of the operator $\mathcal{A}_1^0\phi = \sum a^0(\gamma)F(\gamma)(\phi)$ for consistent Runge-Kutta methods of the form (3.1.2).