Order Conditions for Sampling the Invariant Measure of Ergodic Stochastic Differential Equations on Manifolds

Adrien Laurent[1] · Gilles Vilmart[1]

© The Author(s) 2021

Abstract

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 constrained overdamped Langevin equation. The analysis is valid for arbitrarily high order and relies on an extension of the exotic aromatic Butcher-series formalism. To illustrate the methodology, a method of order two is introduced, and numerical experiments on the sphere, the torus and the special linear group confirm the theoretical findings.

Keywords Constrained stochastic differential equations · Manifolds · Invariant measure · Ergodicity · Exotic aromatic B-series · Order conditions

Mathematics Subject Classification 60H35 · 70H45 · 37M25 · 65L06

1 Introduction

We consider systems of stochastic differential equations (SDEs) in $\mathbb{R}^d$ subject to a smooth scalar constraint and a Stratonovich noise of the form

$$
\mathrm{d}X(t) = \Pi_{\mathcal{M}}(X(t))f(X(t))\mathrm{d}t + \Pi_{\mathcal{M}}(X(t))\Sigma(X(t)) \circ \mathrm{d}W(t), \quad X(0) = X_0 \in \mathcal{M},
\tag{1.1}
$$

Communicated by Christian Lubich.

✉ Adrien Laurent
   Adrien.Laurent@unige.ch

   Gilles Vilmart
   Gilles.Vilmart@unige.ch

[1]   Section de Mathématiques, Université de Genève, UNI DUFOUR, 24 rue du Général Dufour, CP 64, 1211 Geneva 4, Switzerland

where $\Pi_{\mathcal{M}} \colon \mathbb{R}^d \to \mathbb{R}^{d \times d}$ is the orthogonal projection on the tangent bundle of the manifold $\mathcal{M} = \{x \in \mathbb{R}^d, \zeta(x) = 0\}$ of codimension q, $\zeta \colon \mathbb{R}^d \to \mathbb{R}^q$ is a given constraint, $f \colon \mathbb{R}^d \to \mathbb{R}^d$ is a smooth drift, $\Sigma \colon \mathbb{R}^d \to \mathbb{R}^{d \times d}$ is a smooth diffusion coefficient, and W is a standard d-dimensional Brownian motion in $\mathbb{R}^d$ on a probability space equipped with a filtration and fulfilling the usual assumptions. For simplicity of the analysis, we assume that $\mathcal{M}$ is a compact smooth manifold of codimension $q = 1$. The smoothness and compactness of $\mathcal{M}$ guarantee in particular the existence and uniqueness of a solution to (1.1) with bounded moments for all times $t > 0$.[1] In addition, thanks to the projection operator $\Pi_{\mathcal{M}}$, the solution $X(t)$ lies on $\mathcal{M}$ for all $t > 0$. In the additive noise case where $\Sigma(x) = \sigma I_d$ with $\sigma > 0$, Eq. (1.1) can also be rewritten equivalently with a Lagrange multiplier (see [44, Sect. 3.2.4.1] or [45, Sect. 3.3]) as:

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

where $g = \nabla \zeta$ and λ is an adapted stochastic process determined by the equation $\zeta(X) = 0$.

A major motivation of model (1.1) appears in computational problems in molecular dynamics with the constrained overdamped Langevin equation (obtained in the particular case where $\Sigma(x) = \sigma I_d$ is a constant homothety),

$$dX(t) = \Pi_{\mathcal{M}}(X(t))f(X(t))dt + \sigma \Pi_{\mathcal{M}}(X(t)) \circ dW(t), \quad X(0) = X_0 \in \mathcal{M}, \tag{1.2}$$

with $\sigma > 0$, $f = -\nabla V$ and $V \colon \mathbb{R}^d \to \mathbb{R}$ is a smooth potential. 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. The possible constraints can be induced for example by strong covalent bonds between atoms, or fixed angles in molecules. Sampling from the constrained overdamped Langevin equation allows to compute the so-called free energy, which is a key quantity in thermodynamic (see, for instance, [18,44,45] and references therein). Equations of the form (1.1) appear naturally when studying conservative SDEs, that is, SDEs possessing an invariant H conserved almost surely by all realizations of (1.1). The solution of conservative SDEs is 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 [10,23] and references therein).

Under regularity conditions on the generator of the SDE and on $\mathcal{M}$, it was shown in [18,25] that the solution $X(t)$ of the SDE (1.1) is ergodic, 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 $\mathbb{R}^d$, such that for all test functions $\phi \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$,

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \phi(X(t))dt = \int_{\mathcal{M}} \phi(x)d\mu_\infty(x) \quad \text{almost surely.} \tag{1.3}$$

---

[1] The extension to manifolds of any codimension $q \geq 1$ will be further discussed in Remarks 2.6 and 3.1.

In the case of the overdamped Langevin Eq. (1.2) on $\mathcal{M}$, the process is naturally ergodic and the invariant measure is given by $\mathrm{d}\mu_\infty = \rho_\infty \mathrm{d}\sigma_\mathcal{M} = \frac{1}{Z} \exp\left(-\frac{2}{\sigma^2} V\right) \mathrm{d}\sigma_\mathcal{M}$ with $Z = \int_\mathcal{M} \exp\left(-\frac{2}{\sigma^2} V\right) \mathrm{d}\sigma_\mathcal{M}$. Approximating the quantity $\int_\mathcal{M} \phi(x) \mathrm{d}\mu_\infty(x)$ is a computational challenge when the dimension d is high, which is the case in the context of molecular dynamics where the dimension is proportional to the number of particles, because a standard quadrature formula becomes prohibitively expensive. We emphasize that μ_∞ is singular with respect to the Lebesgue measure on $\mathbb{R}^d$. In addition, the integrator samples should remain on the manifold $\mathcal{M}$. Hence, the order conditions for sampling the invariant measure in the Euclidean context of $\mathbb{R}^d$ do not generalize straightforwardly to the manifold case. The main goal of this article is to build and analyse high-order one-step integrators for approximating $\int_\mathcal{M} \phi(x) \mathrm{d}\mu_\infty(x)$ that lie on the manifold $\mathcal{M}$ and that have the form

$$X_{n+1} = \Phi(X_n, h, \xi_n), \tag{1.4}$$

where the ξ_n are standard independent random vectors and h is the numerical step.

There are different ways to approximate the solution of the SDE problem (1.1). A strong approximation focuses on approaching the realization of a single trajectory of (1.1) for a given realization of the Wiener process W. A weak approximation approaches the average of functionals of the solution. We focus here on the approximation for the invariant measure, that is, approaching averages of functionals of the solution in the stationary state. This convergence is the numerical equivalent of (1.3). The integrator (1.4) is said to have order p for the invariant measure if for all $\phi \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$, there exists a positive constant $C(\phi)$ independent of the initial condition X_0 such that

$$e(\phi, h) \leq C(\phi) h^p \quad \text{where} \quad e(\phi, h) = \left| \lim_{N \to \infty} \frac{1}{N+1} \sum_{n=0}^{N} \phi(X_n) - \int_\mathcal{M} \phi \mathrm{d}\mu_\infty \right|. \tag{1.5}$$

We recall that a scheme of weak order r immediately has order $p \geq r$ for the invariant measure. For the underdamped and overdamped Langevin dynamics in $\mathbb{R}^d$, the articles [3,4,9,40,41] proposed multiple schemes of high order for the invariant measure with low weak order (typically $r = 1$). We mention in particular the work [3] that introduced a methodology for the analysis and design of high-order integrators for the invariant measure. This methodology, which relies on Talay–Tubaro expansions [63], backward error analysis and modified differential equations for SDEs [1,22,36,37,67], is generalized in the context of manifolds in the present paper.

A widely used and simple numerical scheme for sampling the invariant measure distribution on manifolds is the Euler scheme (see [17,42,44,45] for instance). Two variants exist for the overdamped Langevin Eq. (1.2), both of order one in the weak sense, or for sampling the invariant measure: the Euler integrator with explicit projection direction

$$X_{n+1} = X_n + hf(X_n) + \sigma\sqrt{h}\xi_n + \lambda g(X_n), \quad \zeta(X_{n+1}) = 0, \tag{1.6}$$

and alternatively the Euler integrator with implicit projection direction

$$X_{n+1} = X_n + hf(X_n) + \sigma\sqrt{h}\xi_n + \lambda g(X_{n+1}), \quad \zeta(X_{n+1}) = 0. \tag{1.7}$$

To the best of our knowledge, no high-order numerical integrators for sampling the invariant measure of the overdamped Langevin equation with constraints (1.2) have been proposed in the literature. In [46], an order two discretization based on the RATTLE integrator (see [5,31,58]) is applied to the underdamped Langevin equation, rather than to the overdamped Langevin dynamic (1.2). The previously described discretizations can be combined with Metropolis–Hastings rejection procedures [32, 50]. We quote in particular the Markov chain Monte Carlo (MCMC) methods [10, 27,45] and the hybrid Monte Carlo methods [46,65], where the need for a reverse projection check is shown to be a key step. We also mention the integrators in [47,66] 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 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 (see [7] for further details in the context of ODEs).

This article proposes new tools for constructing integrators of any high order for sampling the invariant measure of constrained SDEs of the form (1.1) and relies on the formalism of trees and Butcher-series. Originally introduced by Hairer and Wanner in [30], and based on the work of Butcher [13], 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 [14,29]. In the last decades, several works extended B-series to the stochastic context. We mention in particular Burrage and Burrage [11,12] and Komori et al. [35] who first introduced stochastic trees and B-series for studying the order conditions of strong convergence of SDEs, Rößler [53–57] and Debrabant and Kværnø [19–21] for the design and analysis of high-order weak and strong integrators on a finite time interval, [6] for creating schemes preserving quadratic invariants, and [38], where tree series were applied to a class of stochastic differential algebraic equations (SDAEs) for the computation of strong order conditions. Finally, we mention the recent work [39], which introduced the exotic aromatic B-series for the computation of order conditions for sampling the invariant measure of ergodic SDEs in $\mathbb{R}^d$, and that we extend in this paper to the context of SDEs on manifolds.

This article is organized as follows. Section 2 is devoted to the analysis of the accuracy of integrators for sampling the invariant measure on a manifold $\mathcal{M}$. In Sect. 3, we apply this methodology on a class of Runge–Kutta methods for solving the constrained overdamped Langevin equation (1.2), 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. The detailed calculations of the order conditions for the invariant measure are done in Sect. 4 with the help of an extension of the exotic aromatic B-series formalism [39]. We compare in Sect. 5 the new order two scheme with the Euler scheme (1.7) in numerical experiments on a sphere, a torus and the special linear group SL(m) to confirm its order of convergence for sampling the invariant measure.

2 High Order Ergodic Approximation on a Manifold

In this section, we present a new criterion for building integrators of any order for the invariant measure by extending the $\mathbb{R}^d$ results in [3,22] to the context of manifolds. We first settle down a few notations and assumptions, before we recall the standard weak expansions of the exact and numerical solution using the backward Kolmogorov equation. For $\zeta \colon \mathbb{R}^d \to \mathbb{R}$ a smooth map, 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 $\mathbb{R}^d$. We assume in the rest of the article that $\mathcal{M}$ is a compact and smooth manifold of codimension one embedded in $\mathbb{R}^d$. We suppose in addition that the Gram function G is strictly positive on $\mathcal{M}$, $G(x) \geq \alpha > 0$ for all $x \in \mathcal{M}$. With these notations, the projection $\Pi_{\mathcal{M}}$ on the tangent bundle is given by $\Pi_{\mathcal{M}}(x) = I_d - G(x)^{-1}g(x)g(x)^T$. We denote $\mathcal{L}$ the generator of the SDE (1.1). It is given, for $\phi \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$, by

$$\mathcal{L}\phi = \phi'(\Pi_{\mathcal{M}}f) + \frac{1}{2}\sum_{i=1}^{d}\phi'((\Pi_{\mathcal{M}}\Sigma e_i)'(\Pi_{\mathcal{M}}\Sigma e_i)) + \frac{1}{2}\sum_{i=1}^{d}\phi''(\Pi_{\mathcal{M}}\Sigma e_i, \Pi_{\mathcal{M}}\Sigma e_i),$$

$$(2.1)$$

where $(e_i)_{i=1,\ldots,d}$ is the canonical basis of $\mathbb{R}^d$ and, for all vectors $a^1, \ldots, a^m \in \mathbb{R}^d$, we use the following notation for differentials in $\mathbb{R}^d$,

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

For the overdamped Langevin equation (1.2), the generator (2.1) reduces to

$$\mathcal{L}\phi = \phi'f - G^{-1}(g, f)\phi'g - \frac{\sigma^2}{2}G^{-1}\operatorname{div}(g)\phi'g + \frac{\sigma^2}{2}G^{-2}(g, g'g)\phi'g + \frac{\sigma^2}{2}\Delta\phi$$

$$- \frac{\sigma^2}{2}G^{-1}\phi''(g, g) = \frac{\sigma^2}{2}\exp\left(\frac{2}{\sigma^2}V\right)\operatorname{div}_{\mathcal{M}}\left(\exp\left(-\frac{2}{\sigma^2}V\right)\nabla_{\mathcal{M}}\phi\right), \quad (2.2)$$

where $\nabla_{\mathcal{M}}\psi := \Pi_{\mathcal{M}}\nabla\psi$ and $\operatorname{div}_{\mathcal{M}}(H) := \operatorname{div}(H) - G^{-1}(g, H'(g))$. The adjoint $\mathcal{L}^*$ of the generator (2.1) in $L^2(\mathrm{d}\sigma_{\mathcal{M}})$ for the SDE (1.1), i.e. the operator that satisfies for all test functions $\phi, \psi \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$,

$$\int_{\mathcal{M}} (\mathcal{L}\phi)\psi \, d\sigma_{\mathcal{M}} = \int_{\mathcal{M}} \phi(\mathcal{L}^*\psi) \, d\sigma_{\mathcal{M}},$$

is given by

$$\mathcal{L}^*\phi = -\operatorname{div}_{\mathcal{M}}(\phi f) + \frac{1}{2} \sum_{i=1}^{d} \operatorname{div}_{\mathcal{M}}(\operatorname{div}_{\mathcal{M}}(\phi \Sigma e_i) \Sigma e_i).$$

Remark 2.1 As $\mathcal{L}$ is a self-adjoint operator in $L^2(\mathrm{d}\mu_\infty)$, but not in $L^2(\mathrm{d}\sigma_{\mathcal{M}})$ in general, it could be more natural to perform the analysis in the space $L^2(\mathrm{d}\mu_\infty)$. However, as we allow the substages of our numerical integrators to explore the open neighbourhood of $\mathcal{M}$ in $\mathbb{R}^d$, we shall work in this paper with differential operators that cannot be rewritten in general with intrinsic derivatives on the manifold $\mathcal{M}$. In addition, performing directly the integration by parts calculations in $L^2(\mathrm{d}\mu_\infty)$ with such operators is not straightforward, and this motivated the choice of $L^2(\mathrm{d}\sigma_{\mathcal{M}})$ for the analysis. A similar choice was done in [39] in the context of $\mathbb{R}^d$.

We follow the framework of [25]. 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 $\psi : U \subset N_{\mathcal{M}} \to (V \subset \mathbb{R}^{d-1}) \times (-\varepsilon, \varepsilon)$, such that if $\psi(x) = (y, z)$, then $z = \zeta(x)$. We make the following regularity assumption on the generator $\mathcal{L}$.

Assumption 2.2 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 $v : T\mathcal{M} \to \mathbb{R}$ on $\mathcal{M}$ of norm one, we have

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

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 $\mathcal{M}$, rewritten in the local orthogonal coordinate system.

This assumption is a variant in the manifold case of the uniform ellipticity property of the generator $\mathcal{L}$ in the Euclidean context of $\mathbb{R}^d$. In addition, Assumption 2.2 is automatically satisfied for the constrained overdamped Langevin equation (1.2) and yields that the function $u(x, t) = \mathbb{E}[\phi(X(t))|X(0) = x]$ satisfies the backward Kolmogorov equation (see [25]):

$$\frac{\partial u}{\partial t}(x, t) = \mathcal{L}u(x, t), \quad u(x, 0) = \phi(x), \quad x \in N_{\mathcal{M}}, \quad t > 0. \tag{2.3}$$

We refer to [36,37] for similar results in the context of $\mathbb{R}^d$. The backward Kolmogorov equation (2.3) allows us to write the following expansion of $u(x, h) = \mathbb{E}[\phi(X(h))|X_0 = x]$ for h small enough

$$u(x, h) = \phi(x) + \sum_{j=1}^{N} \frac{h^j}{j!} \mathcal{L}^j \phi(x) + h^{N+1} R_N^h(\phi, x), \quad x \in N_{\mathcal{M}}, \qquad (2.4)$$

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

We now assume the existence and uniqueness of an invariant measure, as well as an additional regularity property on $\mathcal{L}$, in the spirit of [22, Hypotheses H1-H2] in the context of $\mathbb{R}^d$.

Assumption 2.3 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 constrained overdamped Langevin equation (1.2) (see [25, Sect. 2.3] for further details). Assumption 2.3 yields the ergodicity of the process $X(t)$ solution of (1.1) with the unique invariant measure $d\mu_\infty = \rho_\infty d\sigma_{\mathcal{M}}$ on $\mathcal{M}$. To proceed further, we shall assume that the integrator (1.4) is ergodic, that is, there exists a measure $d\mu_h$ that has a density with respect to $d\sigma_{\mathcal{M}}$ such that

$$\lim_{N \to \infty} \frac{1}{N+1} \sum_{n=0}^{N} \phi(X_n) = \int_{\mathcal{M}} \phi d\mu^h \quad \text{almost surely.} \qquad (2.5)$$

We refer to [48,60–62] in the Euclidean case, and to [25] in the manifold case, and references therein, for further details on the ergodicity of numerical integrators. In addition, we suppose that $\mathbb{E}[\phi(X_1)|X_0 = x]$, the numerical analog of $u(x, h)$, can be developed in powers of h as was done, for instance, in [3,63] in the context of $\mathbb{R}^d$.

Assumption 2.4 For all $\phi \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$, the numerical integrator (1.4) has a weak Taylor expansion of the form

$$\mathbb{E}[\phi(X_1)|X_0 = x] = \phi(x) + \sum_{j=1}^{N} h^j \mathcal{A}_{j-1} \phi(x) + h^{N+1} R_N^h(\phi, x), \quad x \in N_{\mathcal{M}}, (2.6)$$

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 $\left| R_N^h(\phi, x) \right| \leq C_N(\phi)$ where the constant $C_N(\phi)$ is independent of h and x. 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).

Under Assumptions 2.2 and 2.4, by comparing the expansions (2.6) and (2.4), the integrator has at least weak order p if $\mathcal{A}_{j-1} = \mathcal{L}^j/j!$ for $j = 1, \ldots, p$. However, as observed already in $\mathbb{R}^d$, high order for the invariant measure can be achieved in spite of a low weak order. This is the purpose of Theorem 2.5 where we present a new sufficient condition for a scheme to have order r for the invariant measure. This result, which relies on the powerful tool of backward error analysis for SDEs, is similar to [3, Thm. 3.3] in the context of smooth compact manifolds.

Theorem 2.5 *Under Assumptions* 2.2, 2.3 *and* 2.4, *if the numerical scheme is consistent (that is, if $\mathcal{A}_0 = \mathcal{L}$) and ergodic, and if it satisfies in $L^2(\mathrm{d}\sigma_{\mathcal{M}})$*

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

then it has order r for the invariant measure and the numerical error (1.5) *satisfies, for $h \to 0$,*

$$e(\phi, h) = h^r \int_{\mathcal{M}} \phi(x) \rho_r(x) \mathrm{d}\sigma_{\mathcal{M}}(x) + \mathcal{O}(h^{r+1})$$

$$= h^r \int_0^\infty \int_{\mathcal{M}} u(x, t) \mathcal{A}_r^* \rho_\infty(x) \mathrm{d}\sigma_{\mathcal{M}}(x) \mathrm{d}t + \mathcal{O}(h^{r+1}),$$

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 $\int_{\mathcal{M}} \rho_r \mathrm{d}\sigma_{\mathcal{M}} = 0$, with $N_{\mathcal{M}}$ an open neighbourhood of $\mathcal{M}$ in $\mathbb{R}^d$.

The proof of Theorem 2.5 is detailed in Appendix A for the sake of completeness. The idea is to write an expansion of the error in the spirit of [63], and to generalize the analysis in [3,22] on $\mathbb{T}^d$ and in [25] to the context of smooth compact manifolds.

Theorem 2.5 states the result for times $t \to \infty$. A bound of the error at finite time $t_n = nh$ is typically given by the following exponential estimate (see [22,25])

$$\left| \mathbb{E}[\phi(X_n)] - \int_{\mathcal{M}} \phi(x) \mathrm{d}\mu_\infty(x) \right| \le K e^{-\mu t_n} + C h^p,$$

where the constant $\mu > 0$ is in practice the spectral gap of a certain operator that depends on the numerical integrator. Reducing the error term $K e^{-\mu t_n}$ is out of the scope of this paper, though the recent works [2,24,43] proposed numerical methods in $\mathbb{R}^d$ that improve the rate of convergence at infinity, while sometimes also reducing the variance.

Remark 2.6 One can consider possible generalizations of Theorem 2.5 in the case where $\mathcal{M}$ is not compact, or if $\mathcal{M}$ is a manifold of any dimension. We refer to [3] for the non-compact extension of Theorem 2.5 in the context of $\mathbb{R}^d$.

3 High-Order Integrators for Constrained Langevin Dynamics

In this section, we propose a new class of Runge–Kutta methods for sampling the invariant measure of Eq. (1.2), and present the methodology for deriving the conditions of any order for the invariant measure using Theorem 2.5. In particular, we compute exactly the consistency and order two conditions for the invariant measure as they are the most relevant for the applications.

3.1 Runge–Kutta Methods for Constrained Overdamped Langevin

When discretizing naively Eq. (1.2), one cannot ensure in general that the integrator stays on $\mathcal{M}$. It is natural to discretize instead the equivalent formulation with Lagrange multipliers

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

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 RAT-TLE (see [5,31,58]), introduced in the context of constrained Hamiltonian dynamics, or the SPARK class of methods for general DAEs (see [34]). 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., [31]), so that there are only a few evaluations of f per step. We thus consider the following class of Runge–Kutta integrators

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

$$\zeta(Y_i) = 0 \quad \text{if} \quad \delta_i = 1, \quad i = 1,\ldots,s,$$
$$X_{n+1} = Y_s, \tag{3.1}$$

where $A = (a_{ij})$, $\widehat{A} = (\widehat{a}_{ij}) \in \mathbb{R}^{s\times s}$ and $\delta_i = \sum_{j=1}^{s} \widehat{a}_{ij} \in \{0, 1\}$ are the given Runge–Kutta coefficients, and the $\xi_n \sim \mathcal{N}(0, I_d)$ are independent standard Gaussian random vectors in $\mathbb{R}^d$. (An alternative with discrete bounded random variables is discussed in Remark 3.2.) 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 $j = 1,\ldots,s$ (internal stages without projection, $Y_i \notin \mathcal{M}$ a.s.). 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 $\mathbb{1} = (1,\ldots,1)^T$.

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

For instance, the Euler schemes can be written as Runge–Kutta methods of the form (3.1) with $s = 2$ and the following Butcher tableaux.

$$
\text{Euler (1.6)}: \begin{array}{cc|cc|c} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ \hline & 1 & 0 & & 1 & 0 & \end{array} \qquad \text{Euler (1.7)}: \begin{array}{cc|cc|c} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ \hline & 1 & 0 & & 0 & 1 & \end{array}
$$

Note that the class of methods (3.1) satisfies automatically Assumption 2.4.

Remark 3.1 The class of Runge–Kutta methods (3.1) can be straightforwardly generalized (as done in [39] 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 [64] in $\mathbb{R}^d$) or use multiple independent noises in (3.1) 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 [19,39] in the context of $\mathbb{R}^d$, where it is shown for a class of stochastic Runge–Kutta methods 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 rewrite the internal stages of (3.1) 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 (1.2) 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 3.2 If ξ_n is a Gaussian random variable, its realizations can be arbitrarily large, and the existence and uniqueness of the solution of the system (3.1) does not hold in general. A standard remedy to ensure that the projection on $\mathcal{M}$ always exists for $h \leq h_0$ small enough is to replace the standard Gaussian random vectors ξ in (3.1) by bounded discrete random vectors $\widehat{\xi}$ that have the same first moments in the spirit of [51, 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

$$
\mathbb{P}(\widehat{\xi}_i = 0) = \frac{2}{3} \quad \text{and} \quad \mathbb{P}(\widehat{\xi}_i = \pm\sqrt{3}) = \frac{1}{6}, \quad i = 1, \ldots, d. \tag{3.2}
$$

The following lemma guarantees the well-posedness of a method of the form (3.1) 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 [31, Chap. VII] for deterministic DAEs.

Lemma 3.3 *For Runge–Kutta methods of the form* (3.1) *where the ξ_n are replaced by bounded random variables $\widehat{\xi}_n$, there exists $h_0 > 0$ such that for all $h \leq h_0$, for any initial condition $X_n \in \mathcal{M}$, there exists a unique solution X_{n+1} of* (3.1) *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 $i = 1, \ldots, s$.*

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:

$$Y_i = X_n + h \sum_{j=1}^{i-1} a_{ij} f(Y_j) + \sigma \sqrt{h} d_i \widehat{\xi}_n + \lambda_i \sum_{j=1}^{i} \widehat{a}_{ij} g(Y_j), \qquad (3.3)$$

$$\zeta(Y_i) = 0. \qquad (3.4)$$

Using $\zeta(X_n) = 0$, we rewrite Eq. (3.4) as

$$\zeta(Y_i) - \zeta(X_n) = \int_0^1 g^T(X_n + \tau(Y_i - X_n)) d\tau (Y_i - X_n) = 0. \qquad (3.5)$$

Inserting (3.3) in (3.5) yields

$$\int_0^1 g^T(X_n + \tau(Y_i - X_n)) d\tau \Big[h \sum_{j=1}^{i-1} a_{ij} f(Y_j) + \sigma \sqrt{h} d_i \widehat{\xi}_n + \lambda_i \sum_{j=1}^{i} \widehat{a}_{ij} g(Y_j) \Big] = 0. \qquad (3.6)$$

Multiplying both sides of Eq. (3.3) 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 (3.3) with its value from (3.6), 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

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

$$\Big(\sum_{j=1}^{i-1} \widehat{a}_{ij} g(Y_j) + \widehat{a}_{ii} g(y) \Big)$$

$$+ \Big(\sum_{j=1}^{i-1} \widehat{a}_{ij} g(Y_j) + \widehat{a}_{ii} g(y) \Big) \Big(y - X_n - t \sum_{j=1}^{i-1} a_{ij} f(Y_j) - \sigma \sqrt{t} d_i \widehat{\xi}_n \Big) \Big].$$

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 centre X_n for $h \leq h_0$ 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 (3.6) 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 (3.3) and (3.4).   □

Remark 3.4 In practice, one can solve numerically each internal stage of the set of equations (3.1) with a fixed point iterations 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 $h \leq h_0$ 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) has multiple solutions in general. For example, the Euler scheme (1.7) always has two solutions if $\mathcal{M}$ is a sphere (the two intersections of $\mathcal{M}$ and a straight line going through the centre of $\mathcal{M}$).

Before looking at the consistency and the order conditions of the class of methods (3.1), we introduce a concise notation for multiplying vectors component-wise.

Definition 3.5 For $y, y^{(1)}, \ldots, y^{(n)} \in \mathbb{R}^d$ and $m \geq 0$, we define the diamond product and the diamond power as the vectors in $\mathbb{R}^d$,

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

We present below the detailed calculation of the consistency conditions of the class of methods (3.1) for the constrained overdamped Langevin equation (1.2). Similar proofs can be found in [44, Prop. 3.24] for the Euler schemes (1.6) and (1.7), and in [3,39] for Runge–Kutta methods in $\mathbb{R}^d$.

Proposition 3.6 *For a Runge–Kutta method of the form* (3.1)*, the operator* $\mathcal{A}_0$ *in* (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

$$b^T \mathbb{1} = d_s = 1 \quad \text{and} \quad \widehat{b}^T d = \widehat{b}^T (\delta \bullet d), \tag{3.7}$$

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) with the initial condition $X_0 = x$, then the internal stages Y_i satisfy the following expansion

$$Y_i = x + \sigma \sqrt{h} d_i \xi + h c_i f(x) + R^h, \quad \text{if} \quad \delta_i = 0,$$

$$Y_i = x + \sqrt{h} \left[\sigma d_i \xi + \lambda_{1/2,i}(x) g(x) \right]$$

$$+ h \left[c_i f(x) + \lambda_{1,i}(x) g(x) + \sigma \lambda_{1/2,i}(x) \sum_{j=1}^{s} \widehat{a}_{ij} d_j g'(x) \xi \right.$$

$$\left. + \lambda_{1/2,i}(x) \sum_{j=1}^{s} \widehat{a}_{ij} \lambda_{1/2,j}(x) \delta_j g'(x) g(x) \right] + R^h, \quad \text{if} \quad \delta_i = 1,$$

where the remainder satisfies $\left| R^h \right| \leq Ch^{3/2}$, 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 [44, Lemma 3.25]. If $\delta_i = 1$, $\zeta(Y_i)$ can also be expanded as

$$\zeta(Y_i) = \zeta(x) + \sqrt{h} \left[\sigma d_i(g, \xi) + \lambda_{1/2,i} G \right]$$

$$+ h \left[c_i(g, f) + \lambda_{1,i} G + \lambda_{1/2,i} \sum_{j=1}^{s} \widehat{a}_{ij} \lambda_{1/2,j} \delta_j(g, g'g) \right.$$

$$+ \sigma \lambda_{1/2,i} \sum_{j=1}^{s} \widehat{a}_{ij} d_j(g, g'\xi) + \frac{1}{2} \sigma^2 d_i^2(\xi, g'\xi)$$

$$\left. + \sigma \lambda_{1/2,i} d_i(g, g'\xi) + \frac{1}{2} \lambda_{1/2,i}^2(g, g'g) \right] + \dots$$

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

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

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

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

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 $\mathcal{A}_0 = \mathcal{L}$. $\qquad\square$

Remark 3.7 The analysis presented in Sect. 3.1 is conducted for the overdamped Langevin dynamics (1.2). 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, [3, Thm. 3.3] and [39, 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.

3.2 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 (1.2). 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 3.8 (Runge–Kutta conditions for order two for the invariant measure) *We consider a Runge–Kutta method of the form* (3.1) *and assume the consistency condition* (3.7). *If the method is ergodic and if the following conditions are satisfied, then the integrator has order two for the invariant measure:*

$$
\begin{aligned}
& \widehat{b}^T d = b^T d, \\
& b^T c = b^T (\delta \bullet c) = b^T d^{\bullet 2} = b^T (\delta \bullet d^{\bullet 2}) = 2\widehat{b}^T d - \frac{1}{2}, \\
& \widehat{b}^T c = \widehat{b}^T (\delta \bullet c) = \widehat{b}^T d^{\bullet 2} = \widehat{b}^T (\delta \bullet d^{\bullet 2}) = \widehat{b}^T d^{\bullet 3} = \widehat{b}^T (\delta \bullet d^{\bullet 3}) = 2\widehat{b}^T d - \frac{1}{2},
\end{aligned}
$$

$$\widehat{b}^T (c \bullet d) = \widehat{b}^T (\delta \bullet c \bullet d),$$

$$b^T (d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))) = 0,$$

$$\widehat{b}^T A((\delta - \mathbb{1}) \bullet d)) = \widehat{b}^T (\delta \bullet A((\delta - \mathbb{1}) \bullet d))) = (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2},$$

$$\widehat{b}^T (d \bullet \widehat{A}c) = \widehat{b}^T (d \bullet \widehat{A}d^{\bullet 2}) = \widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) = 2\widehat{b}^T (d \bullet \widehat{A}d) + (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2},$$

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

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

$$\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}(\delta \bullet d)) + \widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d)) = 2\widehat{b}^T (d \bullet \widehat{A}d) + \frac{3}{2}(\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2},$$

$$\widehat{b}^T (d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) = 0,$$

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

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

$$(\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2} = 0,$$

$$\widehat{b}^T d = b^T d,$$

$$b^T c = b^T d^{\bullet 2} = \widehat{b}^T c = \widehat{b}^T d^{\bullet 2} = \widehat{b}^T d^{\bullet 3} = 2\widehat{b}^T d - \frac{1}{2},$$

$$\widehat{b}^T (d \bullet \widehat{A}c) = \widehat{b}^T (d \bullet \widehat{A}d^{\bullet 2}) = 2\widehat{b}^T (d \bullet \widehat{A}d),$$

$$\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}d) = \widehat{b}^T (d \bullet \widehat{A}d) + \widehat{b}^T d - \frac{1}{4},$$

$$\widehat{b}^T (d \bullet (\widehat{A}d)^{\bullet 2}) = (4 - 2\widehat{b}^T d)\widehat{b}^T (d \bullet \widehat{A}d) + 2\widehat{b}^T d - \frac{1}{2}.$$

For simplicity, we used in Theorem 3.8 the notation $\bullet$ of Definition 3.5. 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

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

The order conditions of Theorem 3.8 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 3.6. It is a differential operator of order four with the following first terms

$$\mathcal{A}_1\phi = \frac{\sigma^4}{8}\Delta^2\phi - \frac{\sigma^4}{4}G^{-1}\Delta\phi''(g, g) + \frac{\sigma^4}{8}G^{-2}\phi^{(4)}(g, g, g, g) + \mathcal{B}\phi, \quad (3.8)$$

where $\mathcal{B}$ is a differential operator of order three. We present the complete expansion of $\mathcal{A}_1$ in Sect. 4 by using a B-series approach. If we assume that $\widehat{b}^T d = b^T d$, then we can integrate by parts 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 order one in ϕ (in the spirit of [3,39]). On a manifold, the integration by parts is a corollary of the Green theorem (see, for instance, [59, Chap. II]). As we shall see below, it reveals a crucial tool for

deriving order conditions for the invariant measure. To perform the calculations in a systematic manner, a formalization of the integration by parts process with trees and B-series is presented in Sect. 4.

Lemma 3.9 (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*

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

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

$$\int_{\mathcal{M}} \left[G^{-k}\psi'H - G^{-(k+1)}(g, H)\psi'g \right] d\mu_{\infty}$$
$$= \int_{\mathcal{M}} \left[G^{-(k+1)}(g, H'g)\psi \right.$$
$$- (2k+1)G^{-(k+2)}(g, g'g)(g, H)\psi$$
$$- G^{-k}\,\mathrm{div}(H)\psi + 2kG^{-(k+1)}(g, g'H)\psi$$
$$+ G^{-(k+1)}\,\mathrm{div}(g)(g, H)\psi + \frac{2}{\sigma^2}G^{-(k+1)}(g, f)(g, H)\psi$$
$$\left. - \frac{2}{\sigma^2}G^{-k}(f, H)\psi \right] d\mu_{\infty}. \tag{3.9}$$

For example, let us integrate by parts the terms of order four w.r.t. ϕ of the operator $\mathcal{A}_1\phi$ in Eq. (3.8). Applying identity (3.9) with $\psi = \frac{\sigma^4}{8}\Delta\phi'(e_i)$, $H = e_i$ and $k = 0$, and then summing on $i = 1, \ldots, d$ yields

$$\int_{\mathcal{M}} \left[\frac{\sigma^4}{8}\Delta^2\phi - \frac{\sigma^4}{8}G^{-1}\Delta\phi''(g, g) \right] d\mu_{\infty} = \int_{\mathcal{M}} \left[-\frac{\sigma^4}{8}G^{-2}(g, g'g)\Delta\phi'g \right.$$
$$+ \frac{\sigma^4}{8}G^{-1}\,\mathrm{div}(g)\Delta\phi'g + \frac{\sigma^2}{4}G^{-1}(g, f)\Delta\phi'g - \frac{\sigma^2}{4}\Delta\phi'f \right] d\mu_{\infty}. \tag{3.10}$$

We apply again (3.9) with $\psi = \frac{\sigma^4}{8}\phi^{(3)}(g, g, e_i)$, $H = e_i$ and $k = 1$, and then sum on $i = 1, \ldots, d$ to get

$$\int_{\mathcal{M}} \left[\frac{\sigma^4}{8}G^{-1}\Delta\phi''(g, g) - \frac{\sigma^4}{8}G^{-2}\phi^{(4)}(g, g, g, g) \right] d\mu_{\infty}$$
$$= \int_{\mathcal{M}} \left[-\frac{\sigma^4}{4}G^{-1}\sum_i \phi^{(3)}(g, \partial_i g, ei) + \frac{\sigma^4}{2}G^{-2}\phi^{(3)}(g, g, g'g) \right.$$
$$- \frac{3\sigma^4}{8}G^{-3}(g, g'g)\phi^{(3)}(g, g, g) + \frac{\sigma^4}{8}G^{-2}\,\mathrm{div}(g)\phi^{(3)}(g, g, g)$$
$$+ \frac{\sigma^2}{4}G^{-2}(g, f)\phi^{(3)}(g, g, g) - \frac{\sigma^2}{4}G^{-1}\phi^{(3)}(g, g, f) \right] d\mu_{\infty}. \tag{3.11}$$

Subtracting (3.11) from (3.10) allows to express $\int_{\mathcal{M}} \mathcal{A}_1 \phi \mathrm{d}\mu_\infty$ with derivatives of ϕ of order strictly less than 4. We iterate this method to obtain $\int_{\mathcal{M}} \mathcal{A}_1 \phi \mathrm{d}\mu_\infty = \int_{\mathcal{M}} \mathcal{A}_1^0 \phi \mathrm{d}\mu_\infty$ where $\mathcal{A}_1^0$ is an operator of order one in ϕ and then find sufficient conditions such that $\mathcal{A}_1^0 = 0$. This implies that $\mathcal{A}_1^* \rho_\infty = 0$, and Theorem 2.5 then gives the order two for the invariant measure. The computation of $\mathcal{A}_1^0$ is further detailed in Sect. 4.

Although constructing methods of high weak order is not the main focus of this paper, considering the explicit formula for $\mathcal{A}_1$ and comparing with $\mathcal{L}^2/2$ (see Sect. 4 for their detailed expansion in B-series), one immediately obtains the following theorem for weak order two of accuracy.

Theorem 3.10 (Runge–Kutta conditions for weak order two) *We consider a Runge–Kutta method of the form* (3.1) *and assume that it satisfies* (3.7). *If the following conditions are satisfied, then the integrator has weak order two:*

$$b^T d = b^T c = b^T (\delta \bullet c) = b^T d^{\bullet 2} = b^T (\delta \bullet d^{\bullet 2}) = \frac{1}{2},$$

$$\widehat{b}^T d = \widehat{b}^T c = \widehat{b}^T (\delta \bullet c) = \widehat{b}^T d^{\bullet 2} = \widehat{b}^T (\delta \bullet d^{\bullet 2}) = \widehat{b}^T d^{\bullet 3} = \widehat{b}^T (\delta \bullet d^{\bullet 3}) = \frac{1}{2},$$

$$\widehat{b}^T (c \bullet d) = \widehat{b}^T (\delta \bullet c \bullet d),$$

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

$$b^T (d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) = 0,$$

$$\widehat{b}^T A((\mathbb{1} - \delta) \bullet d) = \widehat{b}^T (\delta \bullet A((\mathbb{1} - \delta) \bullet d)) = \frac{1}{4},$$

$$\widehat{b}^T (d \bullet \widehat{A} c) = \widehat{b}^T (d \bullet \widehat{A} d^{\bullet 2}) = \widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) = 0,$$

$$\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A} d) = \frac{1}{4},$$

$$\widehat{b}^T (c \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) - \widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d)) - \widehat{b}^T (d \bullet \widehat{A}(\delta \bullet c)) = \frac{1}{8},$$

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

$$\widehat{b}^T (d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) = 0,$$

$$\widehat{b}^T (d \bullet (\widehat{A} d)^{\bullet 2}) + 3\widehat{b}^T (d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))) = \frac{1}{8}.$$

Remark 3.11 For $\delta = \mathbb{1}$, the weak order two conditions of Theorem 3.10 have no solution, which is in contrast with the invariant measure case presented in Theorem 3.8. Indeed, the condition $\widehat{b}^T A((\mathbb{1} - \delta) \bullet d) = \frac{1}{4}$ cannot be fulfilled if we fix $\delta = \mathbb{1}$.

3.3 Illustrative Examples of High Order Runge–Kutta Methods on Manifolds

In this section, we present several examples of high-order Runge–Kutta methods of the form (3.1). 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 paper. First, we introduce a method that has order two for sampling the invariant measure of the constrained Langevin dynamics (1.2). 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 3.8. This method is explicit in f and uses only three evaluations of f per step. It is defined by the following Butcher tableau

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

or by the associated set of equations

$$
\begin{aligned}
Y_1 &= X_n + \sigma\sqrt{h}d_1\xi_n + \lambda_1 g(Y_1), \\
Y_2 &= X_n + hc_2 f(Y_1) + \sigma\sqrt{h}d_2\xi_n + \lambda_2\left[\widehat{a}_{21}g(Y_1) + \widehat{a}_{22}g(Y_2)\right], \\
Y_3 &= X_n + hc_3 f(Y_2) + \sigma\sqrt{h}d_3\xi_n + \lambda_3\left[\widehat{a}_{31}g(Y_1) + \widehat{a}_{32}g(Y_2) + \widehat{a}_{33}g(Y_3)\right], \\
X_{n+1} &= X_n + h\sum_{j=1}^{3}\widehat{a}_{4j}f(Y_j) + \sigma\sqrt{h}\xi_n + \lambda_4\sum_{j=1}^{3}\widehat{a}_{4j}g(Y_j),
\end{aligned}
\tag{3.12}
$$

where $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are such that $\zeta(Y_1) = \zeta(Y_2) = \zeta(Y_3) = \zeta(X_{n+1}) = 0$,

and with the values of $c_i, d_i, \widehat{a}_{ij}$ given in Appendix C. To implement one step of this scheme, we apply a few iterations of the Newton method to find the projections on $\mathcal{M}$. 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 3.2, we replace the standard Gaussian random vectors ξ in (3.1) by independent bounded discrete random vectors $\widehat{\xi}$ that satisfy (3.2). 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 3.8 and Theorem 3.10, and with the expressions of $\mathcal{A}_1\phi$ and $\mathcal{A}_1^0\phi$ (see Sect. 4 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 [39, 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 rewritten as the following differential algebraic equation (DAE) of index two (see [31, Chap. VII]):

$$
\begin{aligned}
\dot{x} &= f(x) + \lambda g(x), \\
0 &= \zeta(x).
\end{aligned}
\tag{3.13}
$$

We obtain a class of deterministic Runge–Kutta methods for solving DAEs of the form (3.13) by setting $\sigma = 0$ in (3.1). 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 (3.13) 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 (3.7) reduce to $b^T \mathbb{1} = d_s = 1$. The weak order two conditions of Theorem 3.10 reduce to the following conditions:

$$b^T d = b^T c = b^T(\delta \bullet c) = b^T d^{\bullet 2} = b^T(\delta \bullet d^{\bullet 2}) = \widehat{b}^T d = \widehat{b}^T c = \frac{1}{2},$$

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

$$\widehat{b}^T A((\mathbb{1} - \delta) \bullet d) = \frac{1}{4},$$

$$\widehat{b}^T(d \bullet \widehat{A}c) = 0.$$

On the other hand, the order two conditions for the invariant measure of Theorem 3.8 on the sphere are the following:

$$\widehat{b}^T d = b^T d,$$

$$b^T c = b^T(\delta \bullet c) = b^T d^{\bullet 2} = b^T(\delta \bullet d^{\bullet 2}) = \widehat{b}^T c = 2\widehat{b}^T d - \frac{1}{2},$$

$$\widehat{b}^T(d \bullet \widehat{A}c) = 2\widehat{b}^T(d \bullet \widehat{A}d) + (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2},$$

$$\widehat{b}^T A((\delta - \mathbb{1}) \bullet d) = (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2}.$$

For example, the following integrator has order two for the invariant measure if $\mathcal{M}$ is a sphere:

$$Y_1 = X_n + h\left(\frac{3}{2} - \sqrt{2}\right)f(Y_2) + \sigma\sqrt{h}\left(1 - \frac{\sqrt{2}}{2}\right)\xi_n + \lambda_1(2Y_1 - Y_2),$$

$$Y_2 = X_n + hf(Y_1) + \sigma\sqrt{h}\xi_n + \lambda_2 Y_1, \quad \zeta(Y_1) = \zeta(Y_2) = 0,$$

$$X_{n+1} = Y_2.$$

Brownian motions on manifolds. Runge–Kutta methods of the form (3.1) can also be used for simulating a Brownian motion on a manifold (see [33, Chap. III]) by solving numerically

$$dX = \Pi_{\mathcal{M}}(X) \circ dW, \quad X(0) = X_0 \in \mathcal{M}. \tag{3.14}$$

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 (1.7) only has weak order one for solving (3.14) in general. Fixing $f = 0$ in (3.1) yields a class of Runge–Kutta methods for solving (3.14). 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 3.8 (respectively, in Theorem 3.10) that involve A, b or c. In the specific case where $\mathcal{M}$ is a sphere, the consistency conditions (3.7) become $d_s = 1$ and the weak order two conditions of Theorem 3.10 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 (1.7), has at least order two for the invariant measure on the sphere.

4 Exotic Aromatic B-Series for Computing Order Conditions

As described in the introduction, B-series were introduced to tackle the calculations of order conditions of ODEs by representing Taylor expansions with trees. In [16], an extension of the original B-series, called aromatic B-series, was used to study volume-preserving integrators. It allowed in particular to represent the divergence of a B-series. B-series and aromatic B-series were also studied later in [26,49,52] for their geometric properties, and in [8,15] for their algebraic structure of Hopf algebras. In [39], a new formalism of B-series, called exotic aromatic B-series, was introduced for computing order conditions for sampling the invariant measure of SDEs in $\mathbb{R}^d$. It added a new kind of edge, called *liana*, to the aromatic trees in order to represent new terms such as the Laplacian of an aromatic B-series. In this section, we extend the formalism of exotic aromatic B-series by allowing the representation of scalar products, and show that the operators $\mathcal{L}^j$ and $\mathcal{A}_j$ can be represented conveniently in the form of B-series. We also rewrite the integration by parts formula (3.9) as a straightforward process on graphs, and apply it to compute $\mathcal{A}_1^0$.

We consider graphs $\gamma = (V, E, L)$ where V is the set of nodes, E the set of edges and L the set of lianas. We split the set of edges into $E = E_0 \cup E_S$ where E_0 are the standard oriented edges as defined in [39], 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 consider graphs where each node has exactly one outgoing edge, except exactly one node, called the root r, that has none. If we consider only the graph (V, E), where we erase the lianas, it can be decomposed in 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 decompose the set of nodes in $V = V_f \cup V_g \cup \{r\}$ where V_f are the nodes representing the function f and are represented with black disks (respectively, V_g represent the function g and are drawn with white disks). We write $N_f(\gamma)$ the number of elements of V_f (respectively, $N_g(\gamma)$ the number of elements of V_g) and $N_l(\gamma)$ the number of lianas. The order of a directed graph $\gamma = (V, E, L)$ is defined as:

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

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

$$V_f = \{v_2, v_5, v_6\}, \quad V_g = \{v_1, v_3, v_4, v_7\}, \quad E_S = \{(v_6, v_7)\},$$
$$E_0 = \{(v_1, r), (v_2, v_1), (v_3, r), (v_4, v_4), (v_5, v_4)\}, \quad L = \{(v_2, v_2), (v_3, v_5), (v_5, v_6)\},$$

$$(4.1)$$

satisfies $|\gamma| = 7$ and is represented as
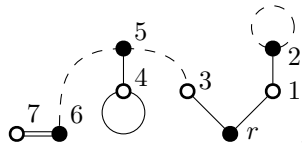
We say that two directed graphs (V^1, E^1, L^1) and (V^2, E^2, L^2) are equivalent if there exists a bijection $\varphi : V^1 \to V^2$ such that

$$\varphi(V_f^1) = V_f^2, \quad \varphi(V_g^1) = V_g^2, \quad (\varphi \times \varphi)(E^1) = E^2,$$
$$(\varphi \times \varphi)(E_S^1) = E_S^2, \quad (\varphi \times \varphi)(L^1) = L^2.$$

We call exotic aromatic forests the equivalence classes of these directed graphs $\gamma = (V, E, L)$, and we denote $\mathcal{EAT}$ the set of exotic aromatic forests. 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 4.1 Let $\gamma = (V, E, L) \in \mathcal{EAT}$, and let $f, g : \mathbb{R}^d \to \mathbb{R}^d$ and $\phi : \mathbb{R}^d \to \mathbb{R}$ be smooth functions. We denote $l_1, \ldots, l_s$ the elements of L, $v_1, \ldots, v_m$ the elements of $V \smallsetminus \{r\}$ and $\delta_{i,j}$ the Kronecker symbol ($\delta_{i,j} = 1$ if $i = j$, $\delta_{i,j} = 0$ else). We use the notation for $v \in V$, $I_{\pi(v)} = (i_{q_1}, \ldots, i_{q_s})$ where $\pi(v) = \{q_1, \ldots, q_s\}$ are the

predecessors of v, and $J_{\Gamma(v)} = (j_{l_{x_1}}, \ldots, j_{l_{x_t}})$ where $\Gamma(v) = \{l_{x_1}, \ldots, l_{x_t}\}$ are the lianas linked to v. Then, $F(\gamma)$ is defined as

$$F(\gamma)(f, g, \phi) = \sigma^{2(|\gamma| - N_f(\gamma))} G^{-N_g(\gamma)/2}$$

$$\sum_{i_{v_1}, \ldots, i_{v_m} = 1}^{d} \sum_{j_{l_1}, \ldots, j_{l_s} = 1}^{d} \left(\prod_{v \in V_f} \delta_{i_v, i_{S(v)}} \partial_{I_{\pi(v)}} \partial_{J_{\Gamma(v)}} f_{i_v} \right)$$

$$\cdot \left(\prod_{v \in V_g} \delta_{i_v, i_{S(v)}} \partial_{I_{\pi(v)}} \partial_{J_{\Gamma(v)}} g_{i_v} \right) \partial_{I_{\pi(r)}} \partial_{J_{\Gamma(r)}} \phi.$$

For example, the differential associated with the exotic aromatic forest γ given by (4.1) is

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

$$\cdot \partial_{i_{v_2}} g_{i_{v_1}} \partial_{j_{l_2}} g_{i_{v_3}} \partial_{i_{v_4} i_{v_5}} g_{i_{v_4}} \delta_{i_{v_7}, i_{v_6}} g_{i_{v_7}} \partial_{i_{v_1} i_{v_3}} \phi.$$

We extend the definition of F on $\mathrm{Span}(\mathcal{EAT})$ by linearity and write, for the sake of simplicity, $F(\gamma)(\phi)$ instead of $F(\gamma)(f, g, \phi)$. An exotic aromatic B-series is a formal series indexed over $\mathcal{EAT}$ of the form

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

Remark 4.2 As we assumed that the functions f and g are gradients, multiple exotic aromatic forests can represent the same differential. We do not detail here the method to identify two such forests as it is similar to [39, Prop. 4.7] in the context of $\mathbb{R}^d$.

The following result states that the operators $\mathcal{L}^j/j!$ and $\mathcal{A}_j$ can be written with exotic aromatic forests. We omit the proof for the sake of brevity as it is similar to [39, Thm. 4.1].

Proposition 4.3 *Take a Runge–Kutta method of the form* (3.1), *then the expansions* (2.4) *and* (2.6) *can be formally written with exotic aromatic B-series, that is, there exists 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

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

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

For example, the operator $\mathcal{L}$ in (2.2) can be rewritten with exotic aromatic forests as:

$$\mathcal{L}\phi = \phi' f - G^{-1}(g, f)\phi' g - \frac{\sigma^2}{2}G^{-1}\,\mathrm{div}(g)\phi' g$$

$$+ \frac{\sigma^2}{2}G^{-2}(g, g'g)\phi' g + \frac{\sigma^2}{2}\Delta\phi - \frac{\sigma^2}{2}G^{-1}\phi''(g, g)$$

$$= F\left(\vcenter{\hbox{\includegraphics{}}} - \vcenter{\hbox{}} - \frac{1}{2}\,\vcenter{\hbox{}} + \frac{1}{2}\,\vcenter{\hbox{}} + \frac{1}{2}\,\vcenter{\hbox{}} - \frac{1}{2}\,\vcenter{\hbox{}}\right)(\phi).$$

We present in Table 2 (see Appendix D) the decomposition in exotic aromatic forests of the operators $\mathcal{L}^2\phi/2 = \mathcal{L}(\mathcal{L}\phi)/2$ and $\mathcal{A}_1\phi$ under the consistency condition (3.7).

Remark 4.4 If we replace the functions g and ϕ by f and fix $\sigma = G = 1$, the newly obtained exotic aromatic B-series satisfy an isometric equivariance property, that is, they stay unchanged when applying an isometric coordinate transformation. It was proved in [52] that under a condition of locality, aromatic B-series are exactly the affine equivariant methods, that is, the maps that stay unchanged when applying an affine coordinate transformation. Analogously, it would be interesting to make a link between the isometric equivariant maps and the exotic aromatic B-series.

In the spirit of the Butcher product on trees [29, Chap. III], we introduce a few notations for writing with ease different operations on forests.

Notation *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. $\overset{\tau}{\gamma}$: *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. $\circ\!\!-\!\tau$ *(resp. $\bullet\!\!-\!\tau$): aroma obtained by linking the root of τ to a white node (resp. a black node) with a new edge in E_S*
3. $\circlearrowleft$: *sum of all aromas obtained by linking the root of τ to a node of τ with a new edge in E_0*
4. $\overset{\frown}{\gamma_v}$: *sum of all exotic aromatic forests/vector fields obtained by linking the node v to a node of γ with a new liana*
5. $\gamma_v{}^{\mathcal{J}}$: *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 $\overset{\frown}{\gamma_v}{}^{\mathcal{J}}$ and $\circ\!\!-\!\overset{\tau}{\gamma}$, where operation 1 is always applied first.

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

$$\overset{\tau}{\gamma} = \vcenter{\hbox{}} + 2\,\vcenter{\hbox{}}, \quad \circ\!\!-\!\tau = \vcenter{\hbox{}}, \quad \circlearrowleft = \vcenter{\hbox{}} + \vcenter{\hbox{}}, \quad \overset{\frown}{\gamma_v} = \vcenter{\hbox{}} + 2\,\vcenter{\hbox{}}, \quad \gamma_v{}^{\mathcal{J}} = \vcenter{\hbox{}}.$$

The integration by parts (3.9) can be rewritten conveniently with exotic aromatic forests.

Lemma 4.5 *Let $\gamma \in \mathcal{EAT}$ and $\tau \in \mathcal{EAV}$, then the process of integration by parts rewrites into*

$$\int_{\mathcal{M}} F\left(\overset{\top}{\gamma} - \text{⊸}_\tau \overset{\circ}{\gamma}\right)(\phi)d\mu_\infty = \int_{\mathcal{M}} F\left(\text{⊸}_\top \gamma - (N_g(\gamma) + N_g(\tau) + 1)\,\text{⊸}\,\text{⊸}_\tau\gamma - \bigcirc\gamma\right.$$
$$+ (N_g(\gamma) + N_g(\tau))\,\text{⊸}\overset{\top}{}\gamma + \bigcirc\text{⊸}_\tau\gamma + 2\,\text{⊸•}\,\text{⊸}_\tau\gamma - 2\,\text{•⊸}_\tau\gamma\Big)$$
$$(\phi)d\mu_\infty, \tag{4.2}$$

$$\int_{\mathcal{M}} F\left(\overset{\frown}{\gamma_v} - \overset{\circ}{\gamma_v}\right)(\phi)d\mu_\infty = \int_{\mathcal{M}} F\left(-(N_g(\gamma) + 1)\,\text{⊸}\gamma_v\right.$$
$$+ N_g(\gamma)\,\gamma_v + \bigcirc\gamma_v + 2\,\text{⊸•}\,\gamma_v - 2\,\gamma_v\Big)(\phi)d\mu_\infty. \tag{4.3}$$

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 (4.2) or (4.3). 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 integrations by parts (3.10) and (3.11) can be rewritten with exotic aromatic B-series by using (4.3) with $\gamma = \bigcirc$ and $\gamma = \vee$. It yields

$$\frac{1}{8}\,(\!\cdot\!) - \frac{1}{8}\,\vee \sim -\frac{1}{8}\,\text{⊸}\!\cdot + \frac{1}{8}\,\bigcirc\!\cdot + \frac{1}{4}\,\text{⊸•}\!\cdot - \frac{1}{4}\,\cdot,$$

$$\frac{1}{8}\,\vee - \frac{1}{8}\,\psi \sim -\frac{1}{4}\,\vee + \frac{1}{2}\,\psi$$
$$\qquad - \frac{3}{8}\,\text{⊸}\psi + \frac{1}{8}\,\bigcirc\psi + \frac{1}{4}\,\text{⊸•}\,\psi - \frac{1}{4}\,\psi.$$

For the sake of completeness, we present in Appendix B the integrations by parts for the order 3 terms of $\mathcal{A}_1\phi$. The computations are similar for the terms of order two in ϕ.

Remark 4.6 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 4.5 reduces to the two following equations:

$$\overset{\top}{\gamma} \sim -\bigcirc\gamma - 2\,\text{•⊸}_\tau\gamma, \quad \overset{\frown}{\gamma_v} \sim -2\,\gamma_v.$$

We recover the process of integration by parts described in [39, Thm. 4.4] in the context of exotic aromatic B-series in $\mathbb{R}^d$.

We can now revisit the statement of Theorem 2.5 in terms of B-series.

Theorem 4.7 *Take a consistent ergodic Runge–Kutta method of the form (3.1). 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 $1 \leq i < r$, then the method has at least order r for the invariant measure.*

By applying repeatedly the process of integrations by parts described in Lemma 4.5, one can simplify the operator $\mathcal{A}_i = F(\gamma_i)$ into an operator of the form $\mathcal{A}_i^0 = F(\gamma_i^0)$ such that $\gamma_i \sim \gamma_i^0$. The complete decomposition of $\mathcal{A}_1^0$ into exotic aromatic forests is detailed in Table 3 (see Appendix D). According to Theorem 4.7, choosing the coefficients of the Runge–Kutta method such that $\gamma_1^0 = 0$ yields the order two conditions for the invariant measure, as stated in Theorem 3.8.

Remark 4.8 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 $\vee$, $\vee$, $\vee$, …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), the operator $\mathcal{A}_1^0 = F(\gamma_1^0)$ has the form

$$\gamma_1^0 = (b^T d - \widehat{b}^T d) \, \bullet\!\!-\!\!\bullet \quad \vee \quad + \sum_{\substack{|\gamma|=2 \\ |\pi(r)|=1}} a^0(\gamma)\gamma,$$

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.

5 Numerical Experiments

In this section, we perform numerical experiments to confirm the theoretical findings, first on a sphere and a torus in $\mathbb{R}^3$, and then on the special linear group.

5.1 Invariant Measure Approximation on a Sphere and a Torus

To check the numerical order two of the Runge–Kutta integrator (3.12) presented in Sect. 3.3, we first compare it with the Euler scheme (1.7) 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 nearly identical to the case $T = 20$ considered in Fig. 1, which suggests that the numerical solutions are already very close to equilibrium at these final times. Following Remark 3.2 and Lemma 3.3, we use discrete bounded random variables satisfying (3.2) in the implementation of the integrators. For both integrators, we compute the Monte Carlo estimator $\overline{J} = \frac{1}{M}\sum_{m=1}^M \phi(X_N^{(m)}) \simeq \mathbb{E}[\phi(X_N)]$, where $X_n^{(m)}$ is the m-th realization of the integrator at
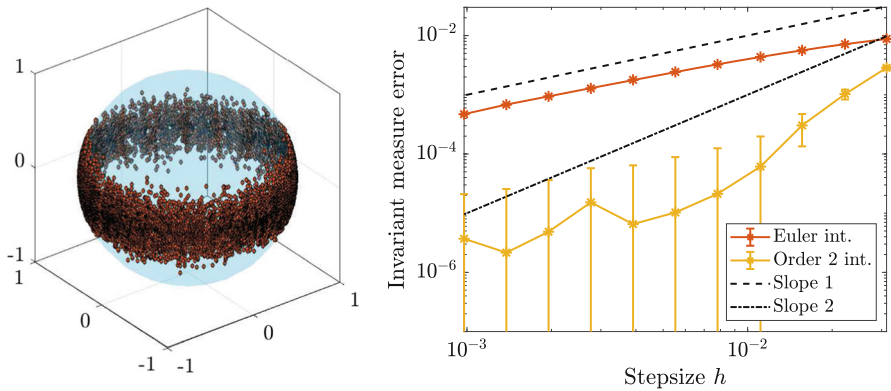
Fig. 1 A trajectory of the order two method (left) and the convergence curve for the sphere for the invariant measure (right) with the potential $V(x) = 25(1 - x_1^2 - x_2^2)$, $\phi(x) = x_3^2$, a final time $T = 20$ and $M = 10^7$ trajectories

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 (1.5) versus different timestep h. We also plot an estimate of the Monte Carlo error by using the standard error of the mean estimator $\left(\sum_{m=1}^{M} (\phi(X_N^{(m)}) - \overline{J})^2 \right)^{1/2} / \sqrt{M(M-1)}$. We observe in all convergence plots that the Monte Carlo error prevails for small values of the timestep h. On Fig. 1, we observe as expected order one for the Euler scheme (1.7) and order two for the Runge–Kutta scheme (3.12).

We then apply the Euler scheme (1.7) and the Runge–Kutta integrator (3.12) 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 Fig. 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 Sect. 3. In particular, the scheme (3.12) has order two of accuracy for the invariant measure on manifolds, according to Theorem 3.8. 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.

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) is convenient as the number of order conditions does not increase with the dimension of the space increasing. We apply Method (3.12) on a Lie group (in the spirit of [65,66]) to see how it performs in high dimension. We choose the special
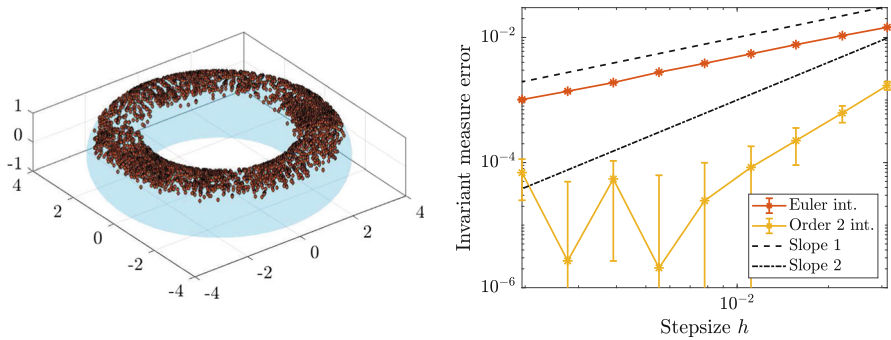
Fig. 2 A trajectory of the order two method (left) and the convergence curve for the torus for the invariant measure (right) with the potential $V(x) = 25(x_3 - r)^2$, $\phi(x) = x_3^2$, a final time $T = 20$ and $M = 10^7$ trajectories

Table 1 Numerical approximation of the integral $J(m) = \int_{\mathrm{SL}(m)} \phi(x) \mathrm{d}\mu_\infty$ for $2 \leq m \leq 4$ with the estimator $\overline{J} = M^{-1} \sum_{k=1}^{M} \phi(X_N^{(k)})$ where (X_n) is given by the Euler scheme (1.7) for $\overline{J}_{\mathrm{Euler}}$ and by the Runge–Kutta integrator (3.12) for $\overline{J}_2$, with their respective errors

m	$\dim(\mathrm{SL}(m))$	$J(m)$	$\overline{J}_{\mathrm{Euler}}$	error for $\overline{J}_{\mathrm{Euler}}$	$\overline{J}_2$	error for $\overline{J}_2$
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	$2.3 \cdot 10^{-4}$

The average is taken over $M = 10^6$ trajectories, with the potential (5.1), $\phi(x) = \mathrm{Tr}(x)$, a final time $T = 10$ and a timestep $h = 2^{-12}T$

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 2.6, 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 (1.7) and the Runge–Kutta integrator (3.12) 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 (3.2). We choose the potential

$$V(x) = 25 \, \mathrm{Tr}((x - I_{m^2})^T (x - I_{m^2})) \tag{5.1}$$

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 Eq. (1.2) 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 (1.7) 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 (3.12). We recall that for a small enough timestep h, the Newton method would always converge (see also Remark 3.4). The reference solution for $J(m) = \int_{\mathrm{SL}(m)} \phi(x) \mathrm{d}\mu_\infty(x)$ is

computed with the Runge–Kutta method (3.12) with $h_{\text{ref}} = 2^{-14}T$. With the factor 25 in the potential (5.1), the solution of (1.2) 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 $\text{SL}(m)$ with moderate manifold curvature. We observe numerically that replacing the factor 25 by 1 in (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 $\text{SL}(m)$ (see, for instance, [28, 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 1 and observe that the Runge–Kutta method (3.12) performs significantly better than the Euler scheme (1.7).

Appendices

A Proof of Theorem 2.5

In the spirit of backward error analysis for differential equations (see [1,22,29,67]), we build a modified generator $\mathcal{L}^h$ such that $U(x, h) = \mathbb{E}[\phi(X_1)|X_0 = x]$ formally satisfies

$$U(x, h) = \sum_{j \geq 0} \frac{h^j}{j!} (\mathcal{L}^h)^j \phi(x). \tag{A.1}$$

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 $\left| R_N^h(\phi, x) \right| \leq C_N(\phi)$. For this, we write formally $\mathcal{L}^h = \mathcal{L} + \sum_{n \geq 1} h^n L_n$ and compare the series expression in (2.6) and (A.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 $\mathbb{R}^d$,

$$L_0 = \mathcal{L}, \quad L_n = \mathcal{A}_n + \sum_{l=1}^{n} \frac{B_l}{l!} \sum_{n_1+\cdots+n_{l+1}=n-l} L_{n_1} \cdots L_{n_l} \mathcal{A}_{n_{l+1}}, \quad n \geq 1, \quad (A.2)$$

where the B_l are the Bernoulli numbers (see [3,22,67] for similar expansions in $\mathbb{T}^d$ or $\mathbb{R}^d$). Using Assumption 2.3, we build recursively a sequence of functions (ρ_n) such that

$$\mathcal{L}^* \rho_n = -\sum_{l=1}^{n} L_l^* \rho_{n-l} \quad \text{and} \quad \rho_0 = \rho_\infty, \tag{A.3}$$

where $\int_{\mathcal{M}} \rho_n d\sigma_{\mathcal{M}} = 0$ for $n \geq 1$. 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 [22, Thm. 2.1] in the context of $\mathbb{R}^d$.

Lemma A.1 *Under Assumptions 2.2, 2.3 and 2.4, 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,*

$$\left| \int_{\mathcal{M}} \phi d\mu^h - \int_{\mathcal{M}} \phi d\mu_r^h \right| \leq C_r(\phi) h^{r+1}. \tag{A.4}$$

We omit the proof of Lemma A.1 as it is exactly the same as in [22, 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 2.5.

Proof of Theorem 2.5 As $\mathcal{A}_j^* \rho_\infty = 0$ for $j = 1, \ldots, r-1$, we deduce recursively from (A.2) and (A.3) that $\rho_j = 0$ for $j = 1, \ldots, r-1$, which yields $\rho_r^h = \rho_\infty + h^r \rho_r$. Using the definition of the error for the invariant measure (1.5) and the ergodicity of the integrator (2.5), Eq. (A.4) becomes

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

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

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

where we used that $\int_{\mathcal{M}}\rho_r(x)\mathrm{d}\sigma_{\mathcal{M}}(x)=0$. This concludes the proof of Theorem 2.5.
□

B Integration by Parts Using the Tree Formalism

We provide here the detailed calculations of the integrations by parts of the order three terms that are needed for the proof of Theorem 3.8. After applying the operations (3.10) and (3.11), the integral $\int_{\mathcal{M}}\mathcal{A}_1\phi\mathrm{d}\mu_\infty$ is transformed into $\int_{\mathcal{M}}\mathcal{B}\phi\mathrm{d}\mu_\infty$ where $\mathcal{B}$ is a differential operator of order three given by

$$
\mathcal{B}\phi = F\Big(\frac{1}{4}\,\mathord{\text{⟨tree⟩}} - \frac{1}{4}\,\mathord{\text{⟨tree⟩}} + \frac{1}{2}\,\mathord{\text{⟨tree⟩}} - \frac{1}{4}\,\mathord{\text{⟨tree⟩}} + \frac{1}{4}\,\mathord{\text{⟨tree⟩}}
$$
$$
+ \frac{1}{8}\,\mathord{\text{⟨tree⟩}} - \frac{3}{8}\,\mathord{\text{⟨tree⟩}} - \frac{1}{4}\,\mathord{\text{⟨tree⟩}} - \frac{1}{8}\,\mathord{\text{⟨tree⟩}} + \frac{1}{8}\,\mathord{\text{⟨tree⟩}}\Big)(\phi) + \mathcal{R}\phi,
$$

and $\mathcal{R}$ is a differential operator of order two. Using Lemma 4.5 multiple times, we get the following integrations by parts of the order three terms of $\mathcal{B}\phi$.

$$
\frac{1}{4}\,\mathord{\text{⟨tree⟩}} - \frac{1}{4}\,\mathord{\text{⟨tree⟩}} \sim -\frac{1}{4}\,\mathord{\text{⟨tree⟩}} + \frac{1}{4}\,\mathord{\text{⟨tree⟩}} - \frac{1}{4}\,\mathord{\text{⟨tree⟩}} + \frac{1}{4}\,\mathord{\text{⟨tree⟩}} + \frac{1}{2}\,\mathord{\text{⟨tree⟩}} - \frac{1}{2}\,\mathord{\text{⟨tree⟩}}
$$

$$
-\frac{1}{4}\,\mathord{\text{⟨tree⟩}} + \frac{1}{4}\,\mathord{\text{⟨tree⟩}} \sim \frac{1}{4}\,\mathord{\text{⟨tree⟩}} + \frac{1}{4}\,\mathord{\text{⟨tree⟩}} - \frac{1}{4}\,\mathord{\text{⟨tree⟩}} - \frac{1}{4}\,\mathord{\text{⟨tree⟩}}
$$

$$
+\frac{3}{4}\,\mathord{\text{⟨tree⟩}} - \frac{1}{2}\,\mathord{\text{⟨tree⟩}} - \frac{1}{4}\,\mathord{\text{⟨tree⟩}} - \frac{1}{2}\,\mathord{\text{⟨tree⟩}} + \frac{1}{2}\,\mathord{\text{⟨tree⟩}}
$$

$$
-\frac{1}{4}\,\mathord{\text{⟨tree⟩}} + \frac{1}{4}\,\mathord{\text{⟨tree⟩}} \sim \frac{1}{4}\,\mathord{\text{⟨tree⟩}} + \frac{1}{4}\,\mathord{\text{⟨tree⟩}} + \frac{1}{4}\,\mathord{\text{⟨tree⟩}} - \frac{3}{4}\,\mathord{\text{⟨tree⟩}} - \frac{1}{4}\,\mathord{\text{⟨tree⟩}}
$$

$$
-\frac{1}{4}\,\mathord{\text{⟨tree⟩}} + \frac{3}{4}\,\mathord{\text{⟨tree⟩}} - \frac{1}{4}\,\mathord{\text{⟨tree⟩}} - \frac{1}{2}\,\mathord{\text{⟨tree⟩}} + \frac{1}{2}\,\mathord{\text{⟨tree⟩}}
$$

$$
-\frac{1}{8}\,\mathord{\text{⟨tree⟩}} + \frac{1}{8}\,\mathord{\text{⟨tree⟩}} \sim \frac{1}{8}\,\mathord{\text{⟨tree⟩}} + \frac{1}{8}\,\mathord{\text{⟨tree⟩}} - \frac{1}{8}\,\mathord{\text{⟨tree⟩}} - \frac{3}{8}\,\mathord{\text{⟨tree⟩}}
$$

$$+ \frac{3}{8} \quad - \frac{1}{8} \quad - \frac{1}{4} \quad + \frac{1}{4}$$

$$\frac{1}{8} \quad - \frac{1}{8} \quad \sim \quad - \frac{1}{8} \quad - \frac{1}{4} \quad - \frac{1}{8} \quad + \frac{1}{4} \quad + \frac{1}{8}$$

$$+ \frac{5}{8} \quad - \frac{5}{8} \quad + \frac{1}{8} \quad + \frac{1}{4} \quad - \frac{1}{4}$$

$$\frac{1}{4} \quad - \frac{1}{4} \quad \sim \quad - \frac{1}{2} \quad + \frac{1}{2} \quad + \frac{5}{4} \quad + \frac{1}{4} \quad - \frac{5}{4}$$

$$- \frac{1}{4} \quad - \frac{1}{4} \quad + \frac{1}{4} \quad + \frac{1}{2} \quad - \frac{1}{2}$$

C Coefficients of the Order Two Runge–Kutta Method

The coefficients of the Runge–Kutta method (3.12) used in Sect. 5 are

$$c_2 = 0.621729189582953540, \qquad c_3 = 0.102032386582165330,$$
$$d_1 = -0.898931652839146019, \qquad d_2 = -1.66233102561284629,$$
$$d_3 = 0.318924515019668897, \qquad \widehat{a}_{21} = 0.584372887990673524,$$
$$\widehat{a}_{31} = 0.887706593835748395, \qquad \widehat{a}_{32} = -0.345018694936693742,$$
$$\widehat{a}_{41} = 0.054744950605 4026516, \qquad \widehat{a}_{42} = -0.0205123070437693053,$$
$$\widehat{a}_{22} = 1 - \widehat{a}_{21}, \qquad \widehat{a}_{33} = 1 - \widehat{a}_{31} - \widehat{a}_{32},$$
$$\widehat{a}_{43} = 1 - \widehat{a}_{41} - \widehat{a}_{42}.$$

D Decomposition of the Operators in Exotic Aromatic Forests

See Tables 2 and 3.

Table 2 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)

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
	$\sigma^4 G^{-1}\sum \phi^{(3)}(g,g'e_i,e_i)$	$-\frac{1}{2}$	$-\frac{1}{2}$
	$\sigma^2 G^{-2}(g,f)\phi^{(3)}(g,g,g)$	$\frac{1}{2}$	$\frac{1}{2}$
	$\sigma^4 G^{-2}\operatorname{div}(g)\phi^{(3)}(g,g,g)$	$\frac{1}{4}$	$\frac{1}{4}$
	$\sigma^4 G^{-2}(g,g'g)\phi^{(3)}(g,g,g)$	$-\frac{3}{4}$	$-\frac{3}{4}$
	$\sigma^2 G^{-1}(g,f)\Delta\phi'(g)$	$-\frac{1}{2}$	$-\frac{1}{2}$
	$\sigma^4 G^{-1}\operatorname{div}(g)\Delta\phi'(g)$	$-\frac{1}{4}$	$-\frac{1}{4}$
	$\sigma^4 G^{-2}(g,g'g)\Delta\phi'(g)$	$\frac{1}{4}$	$\frac{1}{4}$
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	$-b^T d - \widehat{b}^T d$
	$\sigma^4 G^{-2}\phi''(g,g'g'g)$	$\frac{3}{2}$	$-2\widehat{b}^T(d \bullet \widehat{A}d) - (\widehat{b}^T d)^2 + 2\widehat{b}^T d + 1$

Table 2 continued

Forest γ	Differential $F(\gamma)(\phi)$	Exact $e(\gamma)$	Numerical approximation $a(\gamma)$
	$\sigma^2 G^{-2} \phi''(g'g, g'g)$	$\frac{1}{4}$	$-\widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d$
	$\sigma^4 G^{-2} \phi''(g, g''(g, g))$	$\frac{1}{2}$	$2\widehat{b}^T d^{\bullet 2} - 2\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{1}{2}$
	$\sigma^4 G^{-1} \phi''(g, \Delta g)$	$-\frac{1}{2}$	$-\frac{1}{2}$
	$\sigma^4 G^{-1} \sum \phi''(g'e_i, g'e_i)$	$-\frac{1}{4}$	$\widehat{b}^T(d \bullet \widehat{A}d) + \frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d$
	$\sigma^4 G^{-1} \sum \phi''(g''(g, e_i), e_i)$	0	$\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \widehat{b}^T d^{\bullet 2}$
	$G^{-1}(g, f)\phi''(g, f)$	-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	$\widehat{b}^T d + \frac{1}{2}$
	$\sigma^4 G^{-3}(g, g'g)\phi''(g, g'g)$	$-\frac{5}{2}$	$2\widehat{b}^T(d \bullet \widehat{A}d) + (\widehat{b}^T d)^2 - 3\widehat{b}^T d - \frac{3}{2}$
	$\sigma^2 G^{-2}(g, g'f)\phi''(g, g)$	1	1
	$\sigma^2 G^{-2}(g, f'g)\phi''(g, g)$	$\frac{1}{2}$	$b^T d$
	$\sigma^4 G^{-2} \sum \partial_{ij} g_i g_j \phi''(g, g)$	$\frac{1}{2}$	$\frac{1}{2}$
	$\sigma^4 G^{-2} \sum \partial_j g_i \partial_i g_j \phi''(g, g)$	$\frac{1}{4}$	$\frac{1}{4}$
	$\sigma^4 G^{-3}(g, g'g'g)\phi''(g, g)$	$-\frac{7}{4}$	$\widehat{b}^T(d \bullet \widehat{A}d) + \frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d - \frac{3}{2}$

Table 2 continued

Forest γ	Differential $F(\gamma)(\phi)$	Exact $e(\gamma)$	Numerical approximation $a(\gamma)$
	$\sigma^4 G^{-3}(g, g''(g,g))\phi''(g,g)$	$-\frac{1}{2}$	$\widehat{b}^T(\delta \blacklozenge d^{\blacklozenge 2}) - \widehat{b}^T d^{\blacklozenge 2} - \frac{1}{2}$
	$G^{-2}(g, f)^2\phi''(g,g)$	$\frac{1}{2}$	$\frac{1}{2}$
	$\sigma^2 G^{-2}(g,f)\,\mathrm{div}(g)\phi''(g,g)$	$\frac{1}{2}$	$\frac{1}{2}$

γ	$e(\gamma)$	$a(\gamma)$
	-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	$\widehat{b}^T(\delta \blacklozenge c \blacklozenge d) - \widehat{b}^T(c \blacklozenge d)$
	$-\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T(d \blacklozenge \widehat{A}d^{\blacklozenge 2}) - \frac{1}{2}\widehat{b}^T d$
	$\frac{1}{2}$	$-\widehat{b}^T(d \blacklozenge (\widehat{A}d)^{\blacklozenge 2}) - 3\widehat{b}^T(d \blacklozenge \widehat{A}(d \blacklozenge \widehat{A}((\mathbb{1} - \delta) \blacklozenge d))) - 2\widehat{b}^T d\widehat{b}^T(d \blacklozenge \widehat{A}d) + (\widehat{b}^T d)^2 + \widehat{b}^T d$

Table 2 continued

γ	$e(\gamma)$	$a(\gamma)$
	$\frac{1}{4}$	$\widehat{b}^T(d \diamond \widehat{A}d^{\diamond 2}) - \frac{3}{2}\widehat{b}^T(d \diamond \widehat{A}(\delta \diamond d^{\diamond 2})) + \widehat{b}^T d(\widehat{b}^T d^{\diamond 2} - \widehat{b}^T(\delta \diamond d^{\diamond 2})) + \frac{1}{2}\widehat{b}^T d$
	0	$-2\widehat{b}^T(d^{\diamond 2} \diamond \widehat{A}d) + (\widehat{b}^T d + 1)\widehat{b}^T d^{\diamond 2} - \widehat{b}^T d\widehat{b}^T(\delta \diamond d^{\diamond 2})$
	0	$\widehat{b}^T(d^{\diamond 2} \diamond \widehat{A}d) - \frac{1}{2}\widehat{b}^T d^{\diamond 2}$
	0	$\frac{1}{2}\widehat{b}^T(\delta \diamond d^{\diamond 3}) - \frac{1}{2}\widehat{b}^T d^{\diamond 3}$
	$-\frac{1}{2}$	$-\widehat{b}^T c$
	$-\frac{1}{4}$	$-\frac{1}{2}\widehat{b}^T c$
	$\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T c$
	$-\frac{1}{2}$	$-b^T(\delta \diamond c)$
	$-\frac{1}{4}$	$-\frac{1}{2}b^T(\delta \diamond d^{\diamond 2})$
	$\frac{1}{4}$	$b^T(d \diamond \widehat{A}((\mathbb{1} - \delta) \diamond d))) + \frac{1}{2}b^T(\delta \diamond d^{\diamond 2})$
	1	$\widehat{b}^T(c \diamond \widehat{A}((\mathbb{1} - \delta) \diamond d)) + (\widehat{b}^T d)^2 + \widehat{b}^T d + \widehat{b}^T(d \diamond \widehat{A}((\mathbb{1} - \delta) \diamond d)) - \widehat{b}^T(d \diamond \widehat{A}(\delta \diamond c))$
	$\frac{1}{2}$	$\frac{1}{2}\widehat{b}^T(d \diamond \widehat{A}((\mathbb{1} - \delta) \diamond d)) + \frac{1}{2}(\widehat{b}^T d)^2 + \frac{1}{2}\widehat{b}^T d + \frac{1}{2}\widehat{b}^T(d^{\diamond 2} \diamond \widehat{A}((\mathbb{1} - \delta) \diamond d)) - \frac{1}{2}\widehat{b}^T(d \diamond \widehat{A}(\delta \diamond d^{\diamond 2}))$
	-1	$3\widehat{b}^T(d \diamond \widehat{A}(d \diamond \widehat{A}((\mathbb{1} - \delta) \diamond d))) + (2\widehat{b}^T d - \frac{1}{2})\widehat{b}^T(d \diamond \widehat{A}d) - 3\widehat{b}^T(d \diamond (\widehat{A}((\mathbb{1} - \delta) \diamond d))^{\diamond 2}) + \frac{1}{2}\widehat{b}^T(d \diamond \widehat{A}(\delta \diamond d)) + \widehat{b}^T(d \diamond (\widehat{A}d)^{\diamond 2}) + \frac{1}{2}\widehat{b}^T(d^{\diamond 2} \diamond \widehat{A}((\delta - \mathbb{1}) \diamond d))) + \frac{1}{2}\widehat{b}^T(d \diamond \widehat{A}(\delta \diamond d^{\diamond 2})) - \frac{3}{2}(\widehat{b}^T d)^2 - \frac{3}{2}\widehat{b}^T d$

Table 2 continued

γ	$e(\gamma)$	$a(\gamma)$
	$-\frac{1}{4}$	$-\frac{1}{2}\widehat{b}^T d^{\bullet 2}$
	$-\frac{1}{8}$	$-\frac{1}{4}\widehat{b}^T d^{\bullet 2}$
	$\frac{1}{8}$	$\frac{1}{4}\widehat{b}^T d^{\bullet 2}$
	$\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T (\delta \bullet d^{\bullet 2})$
	$\frac{1}{8}$	$\frac{1}{4}\widehat{b}^T (\delta \bullet d^{\bullet 2})$
	$-\frac{1}{8}$	$\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}d) - (\widehat{b}^T d + \frac{1}{2})\widehat{b}^T d^{\bullet 2} + (\widehat{b}^T d - \frac{1}{4})\widehat{b}^T (\delta \bullet d^{\bullet 2})$
	$\frac{1}{2}$	$-\widehat{b}^T (d \bullet \widehat{A}c) + \widehat{b}^T d$
	$\frac{1}{2}$	$\widehat{b}^T (\delta \bullet A((\mathbb{1} - \delta) \bullet d)) + \widehat{b}^T d b^T d$
	$\frac{1}{4}$	$-\frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}d^{\bullet 2}) + \frac{1}{2}\widehat{b}^T d$
	0	$-\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}d) + \frac{1}{2}\widehat{b}^T (\delta \bullet d^{\bullet 2})$
	$-\frac{1}{2}$	$\widehat{b}^T (d \bullet (\widehat{A}d)^{\bullet 2}) + 3\widehat{b}^T (d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)))$ $+2\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}d) + 2\widehat{b}^T d\widehat{b}^T (d \bullet \widehat{A}d) - \widehat{b}^T (\delta \bullet d^{\bullet 2}) - (\widehat{b}^T d)^2 - \widehat{b}^T d$
	$-\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}((3\delta - 2 \cdot \mathbb{1}) \bullet d^{\bullet 2})) + \widehat{b}^T d(\widehat{b}^T (\delta \bullet d^{\bullet 2}) - \widehat{b}^T d^{\bullet 2}) - \frac{1}{2}\widehat{b}^T d$
	$\frac{1}{2}$	$\widehat{b}^T (\delta \bullet c)$
	$\frac{1}{2}$	$\frac{1}{2}\widehat{b}^T (\delta \bullet d^{\bullet 2}) + \frac{1}{2}\widehat{b}^T (\delta \bullet c)$
	$-\frac{3}{2}$	$\widehat{b}^T (c \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \widehat{b}^T (d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d))$ $+\widehat{b}^T (d \bullet \widehat{A}(\delta \bullet c)) - \frac{1}{2}\widehat{b}^T (\delta \bullet d^{\bullet 2}) - \frac{1}{2}\widehat{b}^T (\delta \bullet c) - (\widehat{b}^T d)^2 - \widehat{b}^T d$
	$\frac{1}{8}$	$\frac{1}{4}\widehat{b}^T (\delta \bullet d^{\bullet 2})$

Table 2 continued

γ	$e(\gamma)$	$a(\gamma)$
	$-\frac{3}{4}$	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2}))$ $+\frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}(\widehat{b}^T d)^2 - \frac{1}{2}\widehat{b}^T d$
	$\frac{9}{8}$	$3\widehat{b}^T(d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) - \frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\delta + \mathbb{1}) \bullet d))$ $-\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) - 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)))$ $+(\frac{1}{2} - 2\widehat{b}^T d)\widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d)) - \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2}))$ $+\frac{9}{2}\widehat{b}^T(\delta \bullet d^{\bullet 3}) - \frac{15}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{3}{2}(\widehat{b}^T d)^2 + \frac{3}{2}\widehat{b}^T d$
	$-\frac{1}{2}$	$-b^T c$
	$-\frac{1}{2}$	$-\frac{1}{2}$
	$-\frac{1}{2}$	$-\frac{1}{2}$
	$-\frac{1}{4}$	$-\frac{1}{2}b^T d^{\bullet 2}$
	$-\frac{1}{2}$	$-b^T d$
	$\frac{3}{2}$	$-\widehat{b}^T(d \bullet \widehat{A}c) + \widehat{b}^T d + 1$
	$\frac{3}{2}$	$b^T(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) + \widehat{b}^T A((\mathbb{1} - \delta) \bullet d) + (\widehat{b}^T d + 2)b^T d$
	$\frac{1}{2}$	$\widehat{b}^T(c \bullet d) - \widehat{b}^T(\delta \bullet c \bullet d) + \frac{1}{2}$
	$\frac{1}{4}$	$\frac{1}{2}b^T(\delta \bullet d^{\bullet 2})$
	$\frac{3}{4}$	$-\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}d^{\bullet 2}) + \frac{1}{2}\widehat{b}^T d + \frac{1}{2}$
	$\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T d^{\bullet 3} - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 3}) + \frac{1}{4}$
	$\frac{1}{2}$	$-\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) + \frac{3}{2}\widehat{b}^T d^{\bullet 2} - \widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{1}{2}$
	$\frac{1}{4}$	$-\widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d$
	$-\frac{9}{4}$	$\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) + 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))) +$ $(2\widehat{b}^T d + 3)\widehat{b}^T(d \bullet \widehat{A}d) + \frac{1}{2}(\widehat{b}^T d)^2 - 4\widehat{b}^T d - 1$
	$-\frac{7}{4}$	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((3\delta - 2 \cdot \mathbb{1}) \bullet d^{\bullet 2})) + 2\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d)$ $-(2\widehat{b}^T d + 3)\widehat{b}^T d^{\bullet 2} + 2(\widehat{b}^T d + 1)\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}\widehat{b}^T d - \frac{3}{2}$

Table 2 continued

γ	$e(\gamma)$	$a(\gamma)$
	$-\frac{1}{8}$	$-\frac{1}{8}$
	$-\frac{1}{8}$	$-\frac{1}{8}$
	$\frac{3}{2}$	$\widehat{b}^T c + 1$
	$\frac{3}{4}$	$\frac{1}{2}\widehat{b}^T c + \frac{1}{2}$
	$-\frac{9}{4}$	$\widehat{b}^T(d \bullet \widehat{A}c) - \frac{1}{2}\widehat{b}^T c - \widehat{b}^T d - \frac{3}{2}$
	$\frac{1}{2}$	$b^T(\delta \bullet c)$
	$\frac{1}{4}$	$\frac{1}{2}b^T(\delta \bullet d^{\bullet 2})$
	$-\frac{5}{4}$	$\widehat{b}^T(\delta \bullet A((\delta - \mathbb{1}) \bullet d)) + b^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) - \frac{1}{2}b^T(\delta \bullet d^{\bullet 2}) - (\widehat{b}^T d + 1)b^T d$
	$\frac{3}{4}$	$\frac{1}{2}\widehat{b}^T d^{\bullet 2} + \frac{1}{2}$
	$\frac{3}{8}$	$\frac{1}{4}\widehat{b}^T d^{\bullet 2} + \frac{1}{4}$
	$-\frac{9}{8}$	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}d^{\bullet 2}) - \frac{1}{4}\widehat{b}^T d^{\bullet 2} - \frac{1}{2}\widehat{b}^T d - \frac{3}{4}$
	$\frac{1}{2}$	$\widehat{b}^T d$
	$\frac{1}{4}$	$\frac{1}{2}\widehat{b}^T d$
	$-\frac{1}{2}$	$\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}\widehat{b}^T d - \frac{1}{4}$
	-3	$\widehat{b}^T(c \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \widehat{b}^T(d \bullet \widehat{A}(\delta \bullet c)) + \widehat{b}^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) - (\widehat{b}^T d)^2 - 3\widehat{b}^T d - 1$
	$-\frac{3}{2}$	$\frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) + \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) - \frac{1}{2}(\widehat{b}^T d)^2 - \frac{3}{2}\widehat{b}^T d - \frac{1}{2}$
	$\frac{23}{4}$	$3\widehat{b}^T(d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) - \frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((3 \cdot \mathbb{1} + \delta) \bullet d)) - 2\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) - 6\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))) - \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) - \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d)) - (4\widehat{b}^T d + \frac{5}{2})\widehat{b}^T(d \bullet \widehat{A}d) + \widehat{b}^T(\delta \bullet d^{\bullet 2}) + (\widehat{b}^T d)^2 + \frac{13}{2}\widehat{b}^T d + 3$
	$-\frac{3}{4}$	$-\frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}$
	$-\frac{3}{8}$	$-\frac{1}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{4}$

Table 2 continued
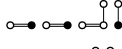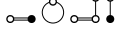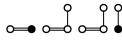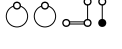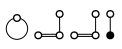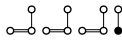
γ	$e(\gamma)$	$a(\gamma)$
		$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((2 \cdot \mathbb{1} - 3\delta) \bullet d^{\bullet 2})) - \widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d)$
	$\frac{13}{8}$	$+(2\widehat{b}^T d + \frac{3}{2})\widehat{b}^T d^{\bullet 2} - (2\widehat{b}^T d + \frac{3}{4})\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{1}{2}\widehat{b}^T d + \frac{5}{4}$
	-1	$-\widehat{b}^T(\delta \bullet c) - \frac{1}{2}$
	-1	$-\frac{1}{2}\widehat{b}^T(\delta \bullet c) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}$
		$\widehat{b}^T(c \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) + \widehat{b}^T(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))$
	$\frac{7}{2}$	$-\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet c)) + \frac{1}{2}\widehat{b}^T(\delta \bullet c) + \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + (\widehat{b}^T d)^2 + 2\widehat{b}^T d + \frac{3}{2}$
	$-\frac{1}{4}$	$-\frac{1}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{8}$
		$\frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) - \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2}))$
	$\frac{7}{4}$	$+\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) + \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d + \frac{3}{4}$
		$\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) + 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)))$
		$-3\widehat{b}^T(d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) + (2\widehat{b}^T d + \frac{1}{2})\widehat{b}^T(d \bullet \widehat{A}d)$
		$+\frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\mathbb{1} + \delta) \bullet d)) + \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) - \frac{9}{2}\widehat{b}^T(\delta \bullet d^{\bullet 3})$
	$-\frac{7}{2}$	$+\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d)) + \frac{15}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - (\widehat{b}^T d)^2 - 3\widehat{b}^T d - \frac{15}{8}$

Table 3 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)

γ	$a^0(\gamma)$
	$b^T d - \widehat{b}^T d$
	$b^T c - 2b^T d + \frac{1}{2}$
	$\frac{1}{2}b^T d^{\bullet 2} - b^T d + \frac{1}{4}$
	$\widehat{b}^T(d \bullet \widehat{A}c) - 2\widehat{b}^T(d \bullet \widehat{A}d) - (\widehat{b}^T d)^2 + 2\widehat{b}^T d - \frac{1}{2}$
	$\widehat{b}^T A((\delta - \mathbb{1}) \bullet d)) - \widehat{b}^T d b^T d + 2\widehat{b}^T d - \frac{1}{2}$
	$b^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d))$
	$-\frac{1}{2}b^T(\delta \bullet d^{\bullet 2}) + b^T d - \frac{1}{4}$
	$\widehat{b}^T(\delta \bullet c \bullet d) - \widehat{b}^T(c \bullet d) + 2\widehat{b}^T d^{\bullet 2} - 2\widehat{b}^T(\delta \bullet d^{\bullet 2})$

Table 3 continued

| γ	$a^0(\gamma)$
	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}d^{\bullet 2}) - \widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d - \frac{1}{4}$
	$-\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) - 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1}-\delta) \bullet d))) + (4 - 2\widehat{b}^T d)\widehat{b}^T(d \bullet \widehat{A}d) + 3(\widehat{b}^T d)^2 - 4\widehat{b}^T d + 1$
	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((2 \cdot \mathbb{1} - 3\delta) \bullet d^{\bullet 2})) + \widehat{b}^T(d \bullet \widehat{A}d) + (\widehat{b}^T d - 1)(\widehat{b}^T d^{\bullet 2} - \widehat{b}^T(\delta \bullet d^{\bullet 2})) + \frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d + \frac{1}{4}$
	$-2\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) + 2\widehat{b}^T(d \bullet \widehat{A}d) + (\widehat{b}^T d - 2)\widehat{b}^T d^{\bullet 2} + (3 - \widehat{b}^T d)\widehat{b}^T(\delta \bullet d^{\bullet 2}) + (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2}$
	$\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) - \widehat{b}^T(d \bullet \widehat{A}d) + \frac{1}{2}\widehat{b}^T d^{\bullet 2} - \widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d - \frac{1}{4}$
	$\frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 3}) - \frac{1}{2}\widehat{b}^T d^{\bullet 3} + \widehat{b}^T d^{\bullet 2} - \widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$-\widehat{b}^T d^{\bullet 2} + \widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$-\widehat{b}^T c + 2\widehat{b}^T d - \frac{1}{2}$
	$-\frac{1}{2}\widehat{b}^T c + \widehat{b}^T d - \frac{1}{4}$
	$\frac{1}{2}\widehat{b}^T c - \widehat{b}^T d + \frac{1}{4}$
	$-b^T(\delta \bullet c) + 2b^T d - \frac{1}{2}$
	$-\frac{1}{2}b^T(\delta \bullet d^{\bullet 2}) + b^T d - \frac{1}{4}$
	$b^T(d \bullet \widehat{A}((\mathbb{1}-\delta) \bullet d)) + \frac{1}{2}b^T(\delta \bullet d^{\bullet 2}) - b^T d + \frac{1}{4}$
	$\widehat{b}^T(d \bullet \widehat{A}((3 \cdot \mathbb{1} - \delta) \bullet d)) - \widehat{b}^T(d \bullet \widehat{A}(\delta \bullet c)) + \widehat{b}^T(c \bullet \widehat{A}((\mathbb{1}-\delta) \bullet d)) + 2(\widehat{b}^T d)^2 - 4\widehat{b}^T d + 1$

Table 3 continued

γ	$a^0(\gamma)$

$\frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}((3 \cdot \mathbb{1} - \delta) \bullet d)) - \frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) + \frac{1}{2}\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) + (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2}$

$\widehat{b}^T (d \bullet (\widehat{A}d)^{\bullet 2}) + 3\widehat{b}^T (d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))) + (2\widehat{b}^T d - \frac{11}{2})\widehat{b}^T (d \bullet \widehat{A}d)$

$-3\widehat{b}^T (d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) - \frac{1}{2}\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))$
$+ \frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d)) + \frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) - 4(\widehat{b}^T d)^2 + 6\widehat{b}^T d - \frac{3}{2}$

$-\frac{1}{2}\widehat{b}^T d^{\bullet 2} + \widehat{b}^T d - \frac{1}{4}$

$-\frac{1}{4}\widehat{b}^T d^{\bullet 2} + \frac{1}{2}\widehat{b}^T d - \frac{1}{8}$

$\frac{1}{4}\widehat{b}^T d^{\bullet 2} - \frac{1}{2}\widehat{b}^T d + \frac{1}{8}$

$-2\widehat{b}^T d^{\bullet 2} + \frac{5}{2}\widehat{b}^T (\delta \bullet d^{\bullet 2}) - \widehat{b}^T d + \frac{1}{4}$

$\frac{5}{4}\widehat{b}^T (\delta \bullet d^{\bullet 2}) - \widehat{b}^T d^{\bullet 2} - \frac{1}{2}\widehat{b}^T d + \frac{1}{8}$

$(\frac{5}{2} - \widehat{b}^T d)\widehat{b}^T d^{\bullet 2} + (\widehat{b}^T d - \frac{13}{4})\widehat{b}^T (\delta \bullet d^{\bullet 2}) + \widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}d) - \widehat{b}^T (d \bullet \widehat{A}d) - \frac{1}{2}(\widehat{b}^T d)^2 + \frac{3}{2}\widehat{b}^T d - \frac{3}{8}$

$-\widehat{b}^T (d \bullet \widehat{A}c) + 2\widehat{b}^T (d \bullet \widehat{A}d) + (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2}$

$\widehat{b}^T (\delta \bullet A((\mathbb{1} - \delta) \bullet d)) + \widehat{b}^T d b^T d - 2\widehat{b}^T d + \frac{1}{2}$

$-\frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}d^{\bullet 2}) + \widehat{b}^T (d \bullet \widehat{A}d) + \frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d + \frac{1}{4}$

$-\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}d) + \widehat{b}^T (d \bullet \widehat{A}d) + \frac{1}{2}\widehat{b}^T (\delta \bullet d^{\bullet 2}) + \frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d + \frac{1}{4}$

$\widehat{b}^T (d \bullet (\widehat{A}d)^{\bullet 2}) + 3\widehat{b}^T (d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))) + 2\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}d)$

$+(2\widehat{b}^T d - 6)\widehat{b}^T (d \bullet \widehat{A}d) - \widehat{b}^T (\delta \bullet d^{\bullet 2}) - 4(\widehat{b}^T d)^2 + 6\widehat{b}^T d - \frac{3}{2}$

$\frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}((3\delta - 2 \cdot \mathbb{1}) \bullet d^{\bullet 2})) - \widehat{b}^T (d \bullet \widehat{A}d) + (1 - \widehat{b}^T d)(\widehat{b}^T d^{\bullet 2} - \widehat{b}^T (\delta \bullet d^{\bullet 2})) - \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d - \frac{1}{4}$

Table 3 continued

| γ	$a^0(\gamma)$
	$\widehat{b}^T(\delta \bullet c) - 2\widehat{b}^T d + \frac{1}{2}$
	$\frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{1}{2}\widehat{b}^T(\delta \bullet c) - 2\widehat{b}^T d + \frac{1}{2}$
	$\widehat{b}^T(c \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \widehat{b}^T(d \bullet \widehat{A}((\delta - 3 \cdot \mathbb{1}) \bullet d))$
	$+\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet c)) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}\widehat{b}^T(\delta \bullet c) - 2(\widehat{b}^T d)^2 + 6\widehat{b}^T d - \frac{3}{2}$
	$\frac{1}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}\widehat{b}^T d + \frac{1}{8}$
	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((\delta - 3 \cdot \mathbb{1}) \bullet d)) + \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2}))$
	$+\frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - (\widehat{b}^T d)^2 + 3\widehat{b}^T d - \frac{3}{4}$
	$3\widehat{b}^T(d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) + (\frac{13}{2} - 2\widehat{b}^T d)\widehat{b}^T(d \bullet \widehat{A}d)$
	$-\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) - 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))) - \frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\delta + \mathbb{1}) \bullet d))$
	$-\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) + \frac{9}{2}\widehat{b}^T(\delta \bullet d^{\bullet 3}) - \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d)) -$
	$\frac{15}{4}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \frac{9}{2}(\widehat{b}^T d)^2 - \frac{15}{2}\widehat{b}^T d + \frac{15}{8}$
	$-b^T c + 2\widehat{b}^T d - \frac{1}{2}$
	$-\frac{1}{2}b^T d^{\bullet 2} + \widehat{b}^T d - \frac{1}{4}$
	$-b^T d + \widehat{b}^T d$
	$-\widehat{b}^T(d \bullet \widehat{A}c) + 2\widehat{b}^T(d \bullet \widehat{A}d) + (\widehat{b}^T d)^2 - 2\widehat{b}^T d + \frac{1}{2}$
	$b^T(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) + \widehat{b}^T A((\mathbb{1} - \delta) \bullet d) + (\widehat{b}^T d + 2)b^T d - 4\widehat{b}^T d + \frac{1}{2}$
	$\widehat{b}^T(c \bullet d) - \widehat{b}^T(\delta \bullet c \bullet d) - 2\widehat{b}^T d^{\bullet 2} + 2\widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$\frac{1}{2}b^T(\delta \bullet d^{\bullet 2}) - \widehat{b}^T d + \frac{1}{4}$
	$-\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}d^{\bullet 2}) + \widehat{b}^T(d \bullet \widehat{A}d) + \frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d + \frac{1}{4}$
	$\frac{1}{2}\widehat{b}^T d^{\bullet 3} - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 3}) - \widehat{b}^T d^{\bullet 2} + \widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$-\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) + \widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}\widehat{b}^T d^{\bullet 2} + \widehat{b}^T(\delta \bullet d^{\bullet 2}) +$ $\frac{1}{2}(\widehat{b}^T d)^2 - \widehat{b}^T d + \frac{1}{4}$

Table 3 continued
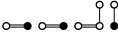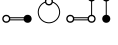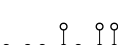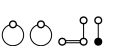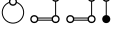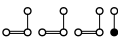
γ	$a^0(\gamma)$
	$\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) + 3\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1}-\delta) \bullet d))) + (2\widehat{b}^T d - 4)\widehat{b}^T(d \bullet \widehat{A}d) - 3(\widehat{b}^T d)^2 + 4\widehat{b}^T d - 1$
	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((3\delta - 2 \cdot \mathbb{1}) \bullet d^{\bullet 2})) + 2\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) - 3\widehat{b}^T(d \bullet \widehat{A}d)$ $+(3-2\widehat{b}^T d)\widehat{b}^T d^{\bullet 2} + (2\widehat{b}^T d - 4)\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{3}{2}(\widehat{b}^T d)^2 + 3\widehat{b}^T d - \frac{3}{4}$
	$\widehat{b}^T d^{\bullet 2} - \widehat{b}^T(\delta \bullet d^{\bullet 2})$
	$\widehat{b}^T c - 2\widehat{b}^T d + \frac{1}{2}$
	$\frac{1}{2}\widehat{b}^T c - \widehat{b}^T d + \frac{1}{4}$
	$\widehat{b}^T(d \bullet \widehat{A}c) - 2\widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}\widehat{b}^T c - (\widehat{b}^T d)^2 + 3\widehat{b}^T d - \frac{3}{4}$
	$b^T(\delta \bullet c) - 2\widehat{b}^T d + \frac{1}{2}$
	$\frac{1}{2}b^T(\delta \bullet d^{\bullet 2}) - \widehat{b}^T d + \frac{1}{4}$
	$b^T(d \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \widehat{b}^T(\delta \bullet A((\delta - \mathbb{1}) \bullet d)) - \frac{1}{2}b^T(\delta \bullet d^{\bullet 2}) - (\widehat{b}^T d + 1)b^T d + 4\widehat{b}^T d - \frac{3}{4}$
	$\frac{1}{2}\widehat{b}^T d^{\bullet 2} - \widehat{b}^T d + \frac{1}{4}$
	$\frac{1}{4}\widehat{b}^T d^{\bullet 2} - \frac{1}{2}\widehat{b}^T d + \frac{1}{8}$
	$\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}d^{\bullet 2}) - \widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{4}\widehat{b}^T d^{\bullet 2} - \frac{1}{2}(\widehat{b}^T d)^2 + \frac{3}{2}\widehat{b}^T d - \frac{3}{8}$
	$\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}d) - \widehat{b}^T(d \bullet \widehat{A}d) - \frac{1}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) - \frac{1}{2}(\widehat{b}^T d)^2 + \widehat{b}^T d - \frac{1}{4}$
	$\widehat{b}^T(c \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \widehat{b}^T(d \bullet \widehat{A}(\delta \bullet c)) + \widehat{b}^T(d \bullet \widehat{A}((\delta - 3 \cdot \mathbb{1}) \bullet d)) - 2(\widehat{b}^T d)^2 + 4\widehat{b}^T d - 1$
	$\frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((\delta - \mathbb{1}) \bullet d)) + \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) + \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}((\delta - 3 \cdot \mathbb{1}) \bullet d)) - (\widehat{b}^T d)^2 + 2\widehat{b}^T d - \frac{1}{2}$
	$3\widehat{b}^T(d \bullet (\widehat{A}((\mathbb{1}-\delta) \bullet d))^{\bullet 2}) - \frac{1}{2}\widehat{b}^T(d^{\bullet 2} \bullet \widehat{A}((3 \cdot \mathbb{1} + \delta) \bullet d)) -2\widehat{b}^T(d \bullet (\widehat{A}d)^{\bullet 2}) - 6\widehat{b}^T(d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1}-\delta) \bullet d))) + (\frac{23}{2} - 4\widehat{b}^T d)\widehat{b}^T(d \bullet \widehat{A}d)$ $-\frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d)) - \frac{1}{2}\widehat{b}^T(d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) + \widehat{b}^T(\delta \bullet d^{\bullet 2}) + 8(\widehat{b}^T d)^2 - 12\widehat{b}^T d + 3$
	$2\widehat{b}^T d^{\bullet 2} - \frac{5}{2}\widehat{b}^T(\delta \bullet d^{\bullet 2}) + \widehat{b}^T d - \frac{1}{4}$

Table 3 continued

γ	$a^0(\gamma)$
	$\widehat{b}^T d^{\bullet 2} - \frac{5}{4}\widehat{b}^T (\delta \bullet d^{\bullet 2}) + \frac{1}{2}\widehat{b}^T d - \frac{1}{8}$
	$\frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}((2 \cdot \mathbb{1} - 3\delta) \bullet d^{\bullet 2})) - \widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}d) + 2\widehat{b}^T (d \bullet \widehat{A}d)$
	$+ (2\widehat{b}^T d - \frac{7}{2})\widehat{b}^T d^{\bullet 2} + (\frac{17}{4} - 2\widehat{b}^T d)\widehat{b}^T (\delta \bullet d^{\bullet 2}) + (\widehat{b}^T d)^2 - \frac{5}{2}\widehat{b}^T d + \frac{5}{8}$
	$-\widehat{b}^T (\delta \bullet c) + 2\widehat{b}^T d - \frac{1}{2}$
	$-\frac{1}{2}\widehat{b}^T (\delta \bullet c) - \frac{1}{2}\widehat{b}^T (\delta \bullet d^{\bullet 2}) + 2\widehat{b}^T d - \frac{1}{2}$
	$\widehat{b}^T (c \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) + \widehat{b}^T (d \bullet \widehat{A}((3 \cdot \mathbb{1} - \delta) \bullet d))$
	$-\widehat{b}^T (d \bullet \widehat{A}(\delta \bullet c)) + \frac{1}{2}\widehat{b}^T (\delta \bullet c) + \frac{1}{2}\widehat{b}^T (\delta \bullet d^{\bullet 2}) + 2(\widehat{b}^T d)^2 - 6\widehat{b}^T d + \frac{3}{2}$
	$-\frac{1}{4}\widehat{b}^T (\delta \bullet d^{\bullet 2}) + \frac{1}{2}\widehat{b}^T d - \frac{1}{8}$
	$\frac{1}{2}\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d)) - \frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d^{\bullet 2}))$
	$+ \frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}((3 \cdot \mathbb{1} - \delta) \bullet d)) + \frac{1}{2}\widehat{b}^T (\delta \bullet d^{\bullet 2}) + (\widehat{b}^T d)^2 - 3\widehat{b}^T d + \frac{3}{4}$
	$\widehat{b}^T (d \bullet (\widehat{A}d)^{\bullet 2}) + 3\widehat{b}^T (d \bullet \widehat{A}(d \bullet \widehat{A}((\mathbb{1} - \delta) \bullet d))) + (2\widehat{b}^T d - \frac{13}{2})\widehat{b}^T (d \bullet \widehat{A}d)$
	$-3\widehat{b}^T (d \bullet (\widehat{A}((\mathbb{1} - \delta) \bullet d))^{\bullet 2}) + \frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d)) + \frac{1}{2}\widehat{b}^T (d^{\bullet 2} \bullet \widehat{A}((\mathbb{1} + \delta) \bullet d))$
	$+ \frac{1}{2}\widehat{b}^T (d \bullet \widehat{A}(\delta \bullet d^{\bullet 2})) - \frac{9}{2}\widehat{b}^T (\delta \bullet d^{\bullet 3}) + \frac{15}{4}\widehat{b}^T (\delta \bullet d^{\bullet 2}) - \frac{9}{2}(\widehat{b}^T d)^2 + \frac{15}{2}\widehat{b}^T d - \frac{15}{8}$

References

1. 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.
2. 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.
3. 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.
4. 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.
5. H. C. Andersen. Rattle: A "velocity" version of the Shake algorithm for molecular dynamics calculations. *J. Comput. Phys.*, 52(1):24–34, 1983.
6. 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.
7. 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.
8. G. Bogfjellmo. Algebraic structure of aromatic B-series. *J. Comput. Dyn.*, 6(2):199–222, 2019.
9. 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.
10. 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.
11. 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.

12. 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.
13. J. C. Butcher. An algebraic theory of integration methods. *Math. Comp.*, 26:79–106, 1972.
14. J. C. Butcher. *Numerical methods for ordinary differential equations*. John Wiley & Sons, Ltd., Chichester, third edition, 2016.
15. P. Chartier, E. Hairer, and G. Vilmart. Algebraic structures of B-series. *Found. Comput. Math.*, 10(4):407–427, 2010.
16. P. Chartier and A. Murua. Preserving first integrals and volume forms of additively split systems. *IMA J. Numer. Anal.*, 27(2):381–405, 2007.
17. G. Ciccotti, R. Kapral, and E. Vanden-Eijnden. Blue moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics. *ChemPhysChem*, 6(9):1809–1814, 2005.
18. 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.
19. K. Debrabant. Runge–Kutta methods for third order weak approximation of SDEs with multidimensional additive noise. *BIT Numer. Math.*, 50(3):541–558, 2010.
20. 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.
21. K. Debrabant and A. Kværnø. Composition of stochastic B-series with applications to implicit Taylor methods. *Appl. Numer. Math.*, 61(4):501–511, 2011.
22. A. Debussche and E. Faou. Weak backward error analysis for SDEs. *SIAM J. Numer. Anal.*, 50(3):1735–1752, 2012.
23. 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.
24. A. B. Duncan, T. Lelièvre, and G. A. Pavliotis. Variance reduction using nonreversible Langevin samplers. *J. Stat. Phys.*, 163(3):457–491, 2016.
25. E. Faou and T. Lelièvre. Conservative stochastic differential equations: mathematical and numerical analysis. *Math. Comp.*, 78(268):2047–2074, 2009.
26. G. Fløystad, D. Manchon, and H. Z. Munthe-Kaas. The universal pre-Lie–Rinehart algebras of aromatic trees. *Submitted*, arXiv:2002.05718, 2020.
27. 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.
28. D. Goldfeld. Automorphic forms and L-functions for the group $GL(n, \mathbf{R})$, volume 99 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2006. With an appendix by Kevin A. Broughan.
29. 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.
30. E. Hairer and G. Wanner. On the Butcher group and general multi-value methods. *Computing (Arch. Elektron. Rechnen)*, 13(1):1–15, 1974.
31. 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.
32. W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
33. E. P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.
34. L. O. Jay. Structure preservation for constrained dynamics with super partitioned additive Runge–Kutta methods. *SIAM J. Sci. Comput.*, 20(2):416–446, 1998.
35. 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.
36. M. Kopec. Weak backward error analysis for Langevin process. *BIT Numer. Math.*, 55(4):1057–1103, 2015.
37. M. Kopec. Weak backward error analysis for overdamped Langevin processes. *IMA J. Numer. Anal.*, 35(2):583–614, 2015.
38. 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.

39. 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.
40. B. Leimkuhler and C. Matthews. Rational construction of stochastic numerical methods for molecular sampling. *Appl. Math. Res. Express. AMRX*, 2013(1):34–56, 2013.
41. 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.
42. 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.
43. 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.
44. T. Lelièvre, M. Rousset, and G. Stoltz. *Free energy computations*. Imperial College Press, London, 2010. A mathematical perspective.
45. T. Lelièvre, M. Rousset, and G. Stoltz. Langevin dynamics with constraints and computation of free energy differences. *Math. Comp.*, 81(280):2071–2125, 2012.
46. T. Lelièvre, M. Rousset, and G. Stoltz. Hybrid Monte Carlo methods for sampling probability measures on submanifolds. *Numer. Math.*, 143(2):379–421, 2019.
47. T. Lelièvre, G. Stoltz, and W. Zhang. Multiple projection MCMC algorithms on submanifolds. *Submitted*, arXiv:2003.09402, 2020.
48. 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.
49. 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.
50. 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.
51. G. N. Milstein and M. V. Tretyakov. *Stochastic numerics for mathematical physics*. Scientific Computation. Springer-Verlag, Berlin, 2004.
52. H. Munthe-Kaas and O. Verdier. Aromatic Butcher series. *Found. Comput. Math.*, 16(1):183–215, 2016.
53. A. Rößler. Stochastic Taylor expansions for the expectation of functionals of diffusion processes. *Stochastic Anal. Appl.*, 22(6):1553–1576, 2004.
54. 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.
55. A. Rößler. Runge–Kutta methods for Itô stochastic differential equations with scalar noise. *BIT Numer. Math.*, 46(1):97–110, 2006.
56. A. Rößler. Stochastic Taylor expansions for functionals of diffusion processes. *Stoch. Anal. Appl.*, 28(3):415–429, 2010.
57. 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.
58. 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.
59. 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.
60. 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.
61. 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.
62. D. Talay. Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. volume 8, pages 163–198. 2002. Inhomogeneous random systems (Cergy-Pontoise, 2001).
63. 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.

64. G. Vilmart. Postprocessed integrators for the high order integration of ergodic SDEs. *SIAM J. Sci. Comput.*, 37(1):A201–A220, 2015.

65. 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.

66. W. Zhang. Ergodic SDEs on submanifolds and related numerical sampling schemes. *ESAIM Math. Model. Numer. Anal.*, 2019.

67. 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.