

High order integration of stochastic dynamics on Riemannian manifolds with frozen flow methods

Eugen Bronasco¹, Adrien Busnot Laurent² and Baptiste Huguet²

March 27, 2025

Abstract

We present a new class of numerical methods for solving stochastic differential equations with additive noise on general Riemannian manifolds with high weak order of accuracy. In opposition to the popular approach with projection methods, the proposed methods are intrinsic: they only rely on geometric operations and avoid coordinates and embeddings. We provide a robust and general convergence analysis and an algebraic formalism of exotic planar Butcher series for the computation of order conditions at any high order. To illustrate the methodology, an explicit method of second weak order is introduced, and several numerical experiments confirm the theoretical findings and extend the approach for the sampling of the invariant measure of Riemannian Langevin dynamics.

Keywords: geometric numerical integration, stochastic differential equations, Riemannian manifolds, Riemannian Langevin, ergodicity, Lie-group methods, frozen flow, Butcher series, exotic series, post-Lie algebra, Hopf algebra, order conditions.

AMS subject classification (2020): 16T05, 41A58, 60H35, 37M25, 65L06, 70H45.

1 Introduction

Let (\mathcal{M}, g) be a smooth connected complete Riemannian manifold equipped with a metric g and its Levi-Civita connection ∇ . Let E_1, \dots, E_D be a frame, that is, D smooth vector fields with $\dim(\mathcal{M}) \leq D$ such that

$$\text{Span}_{\mathbb{R}}(E_1(y), \dots, E_D(y)) = T_y\mathcal{M}, \quad y \in \mathcal{M}.$$

Given a smooth vector field $F(x) = \sum_{d=1}^D f^d(x)E_d(x)$, we consider Stratonovich stochastic differential equations on \mathcal{M} of the form,

$$dX(t) = F(X(t))dt + \sqrt{2} \sum_{d=1}^D E_d(X(t)) \circ dW_d(t). \quad (1.1)$$

Equation (1.1) can be seen as a stochastic differential equation with additive noise w.r.t. the frame (E_d) . Under reasonable assumptions, the flow of (1.1) is well-posed on \mathcal{M} . The aim of the present paper is to introduce a new class of intrinsic one-step integrators

$$X_{n+1} = \Phi_h(X_n), \quad (1.2)$$

¹Université de Genève, Section de mathématiques, Genève, Switzerland. Eugen.Bronasco@unige.ch.

²Univ Rennes, INRIA (Research team MINGuS), IRMAR (CNRS UMR 6625) and ENS Rennes, France. Adrien.Busnot-Laurent@inria.fr, Baptiste.Huguet@math.cnrs.fr.

where Φ_h is a (random) local diffeomorphism of \mathcal{M} for a timestep h small enough, for solving dynamics of the form (1.1) with high order of accuracy in the weak sense. We take inspiration from Lie-group integrators [70, 77, 71, 72, 47, 73] and more precisely from Crouch-Grossman and commutator-free methods [29, 82, 25, 81]. In opposition to [64], we consider an approach with frames allowing us to define a new versatile class of methods that works on any smooth Riemannian manifold. An important feature of the new integrators is that they do not rely on the use of embeddings or on local coordinates. Their formulation, convergence analysis, order theory, and implementation are entirely intrinsic.

It is crucial that a numerical approximation respects the geometry of the problem and lies on \mathcal{M} , especially for the sampling of measures. In particular, our methods extend naturally for the sampling of measures interpreted as invariant measure of ergodic dynamics. Let us consider for instance the central example from molecular dynamics of Riemannian Langevin dynamics:

$$\begin{aligned} dX(t) &= - \sum_{d=1}^D (E_d[V]E_d + \nabla_{E_d}E_d)(X(t))dt + \sqrt{2} \sum_{d=1}^D E_d(X(t)) \circ dW_d(t) \\ &= -\nabla V(X(t))dt + \sqrt{2}dB_{\mathcal{M}}(t), \end{aligned} \quad (1.3)$$

where V is a smooth potential, the frame (E_d) is an orthonormal frame basis, and the vector field $F = -\nabla V - \sum \nabla_{E_d}E_d$ contains the Itô correction. Note that the choice $V = 0$ yields Brownian dynamics $B_{\mathcal{M}}(t)$ on \mathcal{M} . Riemannian Langevin dynamics (1.3) are naturally ergodic under mild assumptions, that is, the dynamics in long time behaves according to a deterministic probability measure, called the invariant measure, which takes the explicit form $d\mu_{\infty} = e^{-V}d \text{vol}$. The generator of the SDE is $\mathcal{L} = -\nabla V \cdot \nabla + \Delta_g$, where Δ_g is the Laplace-Beltrami operator. Actually, an integrator of weak order p automatically has at least order p for the invariant measure [88]. We refer for instance to [33] and references therein for appropriate assumptions to obtain the ergodicity of a numerical scheme in the context of dynamics in \mathbb{R}^d . To the best of our knowledge, there does not exist any intrinsic method of order greater than one for sampling the invariant measure of Riemannian Langevin dynamics.

There exists a handful of extrinsic stochastic integrators on manifolds, that is, integrators that rely on an embedding of \mathcal{M} in a Euclidean space of higher dimension. In the case of smooth embedded manifolds $\mathcal{M} = \{x \in \mathbb{R}^D, \zeta(x) = 0\}$ defined by constraints, the approach using projection methods is the most popular in the literature (see [60] and references therein). We cite in particular the constrained Euler method:

$$X_{n+1} = X_n + hF(X_n) + \sqrt{2h}\xi_n + \lambda_n \nabla \zeta(X_n), \quad \zeta(X_{n+1}) = 0,$$

where λ_n is a Lagrange multiplier and the ξ_n are independent standard Gaussian random vectors. While the approach using projection methods is straightforward to implement and widely used in applications, these methods rely on an embedding of \mathcal{M} into a vector space of much higher dimension and use non-intrinsic quantities, so that they often face severe timestep restrictions. Moreover, the algebraic structure associated to order theory is difficult and does not allow at the present time for the creation of methods of arbitrarily high order, nor for a systematic approach to stochastic backward error analysis [15]. Stochastic Lie group methods seldom appear in the literature [64, 31, 65, 3, 66, 28], but focus on low order approximations on specific manifolds (typically variants of Lie-Euler on the sphere), with an extrinsic convergence analysis, and no high order theory. In the recent work [10], a new approach for sampling (1.3) is developed using the following intrinsic integrator, called Riemannian Langevin method

$$X_{n+1} = \exp^{\text{Riem}}(-h\nabla V(X_n) + \sqrt{2h}\xi_{n+1})X_n,$$

where \exp^{Riem} is the geodesic exponential for the Levi-Civita connection ∇ . This method has order one for solving (1.3) in the weak sense and for the invariant measure. One could be tempted to say that such a method is exact for $V = 0$, but it still has only order one, in contrast to the Euclidean setting where it is exact in law.

We introduce the first intrinsic methods of high weak order for solving (1.1) using only evaluations of F and the frame. In particular, we highlight the following new explicit integrator of second weak order for solving (1.1), and using the minimal number of evaluations of F per step:

$$\begin{aligned} H_n &= \exp \left(\sum_{d=1}^D \left(\frac{1}{2} h f^d(X_n) + \sqrt{h} \xi_n^{d,1} \right) E_d \right) X_n \\ X_{n+1} &= \exp \left(\sum_{d=1}^D \left(\left(\frac{\sqrt{2}}{2} - 1 \right) h f^d(X_n) + (2 - \sqrt{2}) h f^d(H_n) + (1 - \sqrt{2}) \sqrt{h} \xi_n^{d,1} + \sqrt{h} \xi_n^{d,2} \right) E_d \right) \\ &\quad \exp \left(\sum_{d=1}^D \left(\left(1 - \frac{\sqrt{2}}{2} \right) h f^d(X_n) + (\sqrt{2} - 1) h f^d(H_n) + \sqrt{2h} \xi_n^{d,1} \right) E_d \right) X_n. \end{aligned}$$

where the $\xi_n^{d,l}$ are independent standard Gaussian random variables (or discrete symmetric approximations with matching moments up to order 4) and we use frozen flows (see Section 2.3). An originality of our approach is that we present a general framework for the analysis of a new class of stochastic Lie-group methods at any order on any manifold in the weak sense using a new algebraic formalism similar to the one of Butcher series. The new methods outperform in accuracy, and versatility the previous attempts.

The calculations of the order conditions are intricate and require the use of appropriate algebraic tools. We propose a new extension of exotic and Lie-Butcher series, called exotic Lie series, for the systematic computation of order conditions of the new frozen flow methods at any order in the weak sense. While Butcher trees and series [41, 21, 68, 22] represent naturally Taylor expansions of deterministic flows in \mathbb{R}^D , planar trees and Lie-Butcher series were introduced for the study of flows on manifolds [47, 81, 79] and later extended for the study of the connection algebra [78, 35, 5, 40, 74]. On the other hand, the formalism of exotic series was introduced in [55, 51, 14] for the creation of integrators for solving stochastic dynamics with additive noise with high order in the weak sense and for the invariant measure. This formalism, combined with the aromatic B-series [27, 48, 13, 53, 52], was extended in [56] into the exotic aromatic series formalism for the numerical integration of SDEs on embedded manifolds with projection methods. The geometric and algebraic properties of the exotic formalism of trees were later studied in [15, 54] (see also [26, 67, 75, 12] in the deterministic setting). The modern approach to such algebraic formalism relies extensively on Hopf algebras [26, 23, 63, 12, 14, 83, 15, 36, 20] that we identify here in the context of stochastic numerics on manifolds.

The paper is organised the following way. We present in Section 2 the main numerical results of this paper: a robust framework for the Riemannian convergence analysis of stochastic integrators on manifolds, the new stochastic frozen flow methods, and the associated order conditions for high weak order. Section 3 introduces the formalism of planar decorated and exotic forests and translates the geometric operations in the connection algebra in terms of algebraic operations on graphs. This approach is then applied for the explicit description at any order of the weak Taylor expansion of the exact flow of (1.1), of the new frozen flow methods, and of their associated order conditions. The Hopf algebra structure of exotic planar forests is further detailed in Section 4. In particular, we generalise the Munthe-Kaas-Wright Hopf algebra

[79] to the stochastic context. Numerical experiments confirm the theoretical findings in Section 5 and outlooks and future works are presented in Section 6.

2 New frozen flow methods of high weak order

We generalise the standard Euclidean framework for weak convergence in the context of Riemannian manifolds, and we introduce new frozen flow integrators and their order conditions for high weak order. We emphasize that the following convergence analysis is new, even for the underlying deterministic differential equations (see also [30]).

2.1 Stochastic differential equations on manifolds

For $v \in T\mathcal{M}$, we denote $|v| = g(v, v)^{1/2}$, its Riemannian norm. Let o an arbitrary point on \mathcal{M} . For $x \in \mathcal{M}$, we define $r(x) = d(o, x)$ the Riemannian distance function. This function is 1-Lipschitz and the squared distance r^2 is smooth on $\mathcal{M} \setminus \text{Cut}_o$. This allows us to define moments of a \mathcal{M} -valued random variable.

Let us introduce the main classes of test functions. The space $\mathcal{C}_c^\infty(\mathcal{M})$ denotes the class of smooth compactly supported functions. For $\phi \in \mathcal{C}_c^\infty(\mathcal{M})$ we denote $d\phi$ its differential and $|d\phi|$ its Riemannian norm on $T^*\mathcal{M}$, inherited from the norm on $T\mathcal{M}$. We denote $v[\phi](x)$ the differential of a test function ϕ in a direction $v \in T_x\mathcal{M}$. We shall use the wider class of test functions $\mathcal{C}_P^\infty(\mathcal{M})$, which consists in smooth functions satisfying polynomial growth estimates of the form:

$$|E_{d_q}[\dots E_{d_1}[\phi] \dots]|(x) \leq C(1 + r(x)^K), \quad q = 0, 1, \dots$$

Given a vector field F and a fixed decomposition in the frame $F = \sum f^d E_d$, we say that $F \in \mathfrak{X}_P(\mathcal{M})$ if its components are Lipschitz continuous, $|df^d| \leq C$, and satisfy: $f^d \in \mathcal{C}_P^\infty(\mathcal{M})$, $d = 1, \dots, D$. We mention that considering test functions and vector fields with polynomial growth of their derivatives up to a given order $p + 2$ would be sufficient for the creation of numerical schemes of order up to p .

Assumption 2.1. *The vector fields E_1, \dots, E_D are smooth and bounded. The vector field F is in $\mathfrak{X}_P(\mathcal{M})$. Moreover, there exist constants $\nu \geq 1$ and $\lambda \in \mathbb{R}$ such that on $\mathcal{M} \setminus \text{Cut}_o$, we have*

$$\mathcal{L}r^2 \leq \nu + \lambda r^2, \tag{2.1}$$

where \mathcal{L} is the generator

$$\mathcal{L}\phi = F[\phi] + \sum_{d=1}^D E_d[E_d[\phi]].$$

Under the smoothness assumption on F, E_1, \dots, E_D , there exists a unique solution of (1.1), up to a stopping time τ . Following [43], we recall that a stochastic process X , defined on a stochastic interval $[0, \tau]$, is a solution of (1.1) if and only if for all test functions $\phi \in \mathcal{C}_c^\infty(\mathcal{M})$, we have

$$\phi(X(t)) = \phi(X_0) + \int_0^t F[\phi](X(t))dt + \sqrt{2} \sum_{d=1}^D \int_0^t E_d[\phi](X(t)) \circ dW_d(t), \quad 0 < t < \tau.$$

Unlike the Euclidean case, where Lipschitz conditions are sufficient to avoid finite time explosion, on a manifold, a blow-up can be induced from the geometry, even for a Brownian motion (see

[43] for examples). The Lyapunov-like assumption (2.1) ensures the completeness of the SDE. This means that the solution X of (1.1) is defined on \mathbb{R}_+ , i.e $\tau = +\infty$ a.s. In addition with the completeness, it provides finite moments estimates (even finite exponential moments) for the solution of the SDE. This assumption is adapted from [89]. In the case where $(E_i)_i$ is a basis there exists a new metric \tilde{g} such that $(E_i)_i$ is an orthonormal basis. Using this metric and the associated Levi-Civita connection $\tilde{\nabla}$, we find

$$\mathcal{L}\phi = \Delta_{\tilde{g}}\phi + (F + \tilde{\nabla}_{E_i}E_i)[\phi],$$

which is the framework of [89]. The general case is treated likewise, by extracting local basis on a cover of \mathcal{M} .

There exist various simpler criteria which ensure that inequality (2.1) is satisfied. The first criterion is compactness, which can be used for classical spaces such as the sphere \mathbb{S}^n and the Lie group $\text{SO}_n(\mathbb{R})$. This condition being seldom satisfied, a second classical criterion is the Bakry-Émery criterion from [8]. This criterion applies to the case of equation (1.3), where the generator writes as $\mathcal{L} = \Delta - \nabla V \cdot \nabla$. One says that the potential V satisfies the criterion if there exists $\kappa \in \mathbb{R}$ such that

$$\text{Ric} + \text{Hess}(V) \geq \kappa.$$

The term Ric designs the Ricci tensor on (\mathcal{M}, g) and is defined as the trace of the Riemann tensor. In this context, the symmetric operator $\text{Ric} + \text{Hess}(V)$ is called the Bakry-Émery curvature and the term $\text{Hess}(V)$ can be interpreted as an additional curvature. Note that this handy criterion has been adapted in [6] for equation (1.1) with a lower bound of the operator $\text{Ric} - \nabla F$. We also cite the moment conditions from [61] which imply stochastic completeness.

By a standard induction, we obtain the following estimates.

Lemma 2.2. *Under Assumption 2.1, for all $q \geq 1$ we have $\mathcal{L}(r^{2q}) \leq C(1 + r^{2q})$.*

Proof. The proof comes from classical Γ calculus for diffusions (see [9] for instance). For all $u \in \mathcal{C}^2(\mathcal{M})$ and $\eta \in \mathcal{C}^2(\mathbb{R})$, we have

$$\mathcal{L}\phi(u) = \eta'(u)\mathcal{L}u + \eta''(u)\Gamma(u),$$

where

$$\Gamma(u) = \sum_{i=1}^D |E_i[u]|^2.$$

Applying this formula to $u = r^2$ and $\eta : x \mapsto x^q$, we obtain the result. \square

A key tool for the weak expansion of the exact flow is the Kolmogorov equation. Under Assumption 2.1, the SDE (1.1) generates a semigroup, well-defined on $\mathcal{C}_P^\infty(M)$, and Markovian. Namely, for all $\phi \in \mathcal{C}_c^\infty(M)$, the function $u(t, x) = \mathbb{E}[\phi(X(t)) | X(0) = x]$ satisfies the Kolmogorov equation

$$\partial_t u = \mathcal{L}u, \quad u(0, x) = \phi(x).$$

In order to compare an expansion of the exact solution an expansion of a numerical integrator, we need more regularity on the semigroup itself. Since the generator \mathcal{L} is elliptic, the semigroup u is \mathcal{C}^∞ on $]0, T[\times \mathcal{M}$. We further assume that the semigroup preserves the \mathcal{C}_P^∞ regularity.

Assumption 2.3 (Regularity). *For all $\phi \in \mathcal{C}_P^\infty(\mathcal{M})$, $u \in \mathcal{C}^\infty(]0, T[, \mathcal{C}_P^\infty(\mathcal{M}))$: for all $k \geq 0$, there exist $C > 0$ and $\kappa \geq 0$, depending on k , T , and ϕ such that for all $x \in \mathcal{M}$*

$$\sup_{t \in]0, T[} |\partial_t^k u(t, x)| \leq C(1 + r^\kappa(x)).$$

Note that Assumption 2.3 is satisfied automatically in \mathbb{R}^d with our regularity assumptions. For clarity, we leave the study of sufficient conditions to obtain polynomial growth estimates of the semigroup for future works.

We are now able to derive the Taylor expansion of the semigroup.

Proposition 2.4. *Under Assumption 2.3, for all $\phi \in \mathcal{C}_P^\infty(M)$ and all $h \leq h_0$ small enough, u has the Taylor expansion*

$$u(h, x) = \phi(x) + \sum_{j=1}^p \frac{h^j}{j!} \mathcal{L}^j \phi(x) + h^{p+1} R_p^h(\phi, x), \quad x \in \mathcal{M}, \quad (2.2)$$

with $|R_p^h(\phi, x)| \leq C(1 + r(x)^K)$, where C and K do not depend on x and h .

2.2 Weak order theory on Riemannian manifolds

Following the Euclidean works [88, 69], we present a general theory for achieving high order weak estimates for intrinsic methods. The approach does not rely on an embedding and does not assume compactness of \mathcal{M} . In order to obtain estimates, we use the Riemannian structure of \mathcal{M} . This structure is not needed to perform the calculations of order conditions of the new methods.

Definition 2.5. *An intrinsic one-step integrator of the form (1.2) is of local weak order p if for all $\phi \in \mathcal{C}_P^\infty(M)$, there exists $C > 0$ such that the following estimate holds for all $x \in \mathcal{M}$ and all $h \leq h_0$ small enough,*

$$|\mathbb{E}[\phi(X_1)|X_0 = x] - \mathbb{E}[\phi(X(h))|X_0 = x]| \leq C(1 + r(x)^K)h^{p+1}.$$

Similarly, the integrator is of global weak order p if for all $\phi \in \mathcal{C}_P^\infty(M)$, $T > 0$, $Nh = T$, $h \leq h_0$ small enough, and $x \in \mathcal{M}$, the following estimate is satisfied:

$$\sup_{n=0, \dots, N} |\mathbb{E}[\phi(X_n)] - \mathbb{E}[\phi(X(nh))]| \leq Ch^p.$$

Let us assume that the integrator satisfies a Taylor expansion, often called Talay-Tubaro expansion [87] in the Euclidean context, that is similar to (2.2).

Assumption 2.6. *For all $\phi \in \mathcal{C}_P^\infty(M)$ and all $h \leq h_0$ small enough, the numerical integrator (1.2) has a weak Taylor expansion of the form*

$$\mathbb{E}[\phi(X_1)|X_0 = x] = \phi(x) + \sum_{j=1}^p h^j \mathcal{A}_j \phi(x) + h^{p+1} R_p^h(\phi, x), \quad x \in \mathcal{M}, \quad (2.3)$$

where the remainder satisfies $|R_p^h(\phi, x)| \leq C(1 + r(x)^K)$ and the \mathcal{A}_j are linear differential operators.

We adapt the standard assumption of bounded moments in the Riemannian context. We shall give sufficient assumptions on the integrator in Subsection 2.3 so that this assumption is satisfied.

Assumption 2.7. *The numerical integrator (1.2) has bounded moments of any order $K \geq 0$,*

$$\sup_{n=0, \dots, N} \mathbb{E}[r(X_n)^K] \leq C(T), \quad Nh = T. \quad (2.4)$$

Under the above assumptions, we obtain the following convergence result, in the spirit of the Euclidean theory [87, 88, 69].

Theorem 2.8. *Consider a one-step integrator of the form (1.2) for solving (1.1). Under Assumptions 2.1, 2.3, 2.6, and 2.7, if the Taylor expansion (2.3) of the integrator satisfies*

$$\mathcal{A}_j = \frac{1}{j!} \mathcal{L}^j, \quad j = 1, \dots, p, \quad (2.5)$$

then the integrator has global weak order p .

Proof. Let $X^x(t_n)$ be the exact solution of (1.1) at time $t_n = nh$ starting from $x \in \mathcal{M}$ and X_n^x be the n -th step of the numerical solution. The global error satisfies, thanks to a telescopic sum argument,

$$\mathbb{E}[\phi(X_N) - \phi(X(T))] = \sum_{i=1}^N \mathbb{E}[\tilde{\phi}(X_1^{X_{N-i}^x}) - \tilde{\phi}(X^{X_{N-i}^x}(h))], \quad \tilde{\phi}(x) = \mathbb{E}[\phi(X^x(t_{i-1}))].$$

Assumption 2.3 yields that $\tilde{\phi} \in C_P^\infty(\mathcal{M})$. Then, the local order p condition, together with Assumptions 2.6 and 2.7 give

$$\mathbb{E}[\phi(X_N) - \phi(X(T))] \leq \sum_{i=1}^N Ch^{p+1} (1 + \mathbb{E}[r(X_{N-i}^x)^K]) \leq CNh^{p+1} \leq Ch^p.$$

Hence, the global weak order p . □

2.3 Stochastic frozen flow integrators

The deterministic Lie-group methods are divided into two categories that correspond to the two kinds of coordinates: the Runge-Kutta-Munthe-Kaas (RKMK) methods (see [70, 77, 71, 72]) and the frozen flow methods. We choose the latter option in this paper for the sake of generality, as the RKMK approach relies on the existence of a transitive group action on \mathcal{M} . The frozen flow methods originally appeared as Crouch-Grossmann methods [29, 82]. The generalisation presented in the papers [25, 81] allowed the creation of more efficient methods, called commutator-free Lie-group methods.

The methods rely heavily on the concept of frozen vector fields and flows: given a decomposition of a (potentially random) vector field in the frame $G = \sum_d g^d E_d$, the frozen vector field $G_x \in \mathfrak{X}(\mathcal{M})$ at point $x \in \mathcal{M}$ is

$$G_x(p) = \sum_d g^d(x) E_d(p).$$

The solution of the ODE $y'(t) = G(y(t))$ with $y(0) = p$ is denoted $\exp(tG)p$. The frozen flow $\exp(tG_x)p$ is the exact flow driven by the frozen vector field G_x .

Lemma 2.9. *If $G \in \mathfrak{X}_P(\mathcal{M})$ then, the frozen flow $\exp(tG_x)p$ is globally defined. Moreover, it satisfies*

$$d(p, \exp(tG_x)p) \leq C(1 + r(x)) \quad (2.6)$$

Proof. The coefficients of the ODE being globally Lipschitz continuous, there exists a global solution. Let us denote $\gamma : t \mapsto \exp(tG_x)p$. It is a smooth curve, joining p and $\exp(G_x)p$. As $|E_d|$ are bounded, γ has a bounded speed

$$|\dot{\gamma}_t|^2 \leq D \sum_{d=1}^D |g^d(x)|^2 |E_d(\gamma_t)|^2 \leq C(1 + r^2(x)).$$

Therefore, we have

$$d(p, \exp(tG_x)p) \leq \int_0^t |\dot{\gamma}_s| ds \leq C(1 + r(x))$$

Hence, the result. \square

Remark 2.10. *The definition of the frozen vector field and frozen flow do not depend on the choice of decomposition of G in the frame (E_d) . Indeed, let us extract a basis $(E_d)_{d \in I}$ of the frame valid in an open set $U \subset \mathcal{M}$. The vector field G decomposes uniquely as $G = \sum_{d \in I} \hat{g}^d E_d$ in U . The frozen vector field is given as $G_x = \sum_{d \in I} \alpha^d E_d$. As $G(x) = G_x(x)$ and we use a (local) basis, we obtain $\alpha^d = \hat{g}^d(x)$. The formulation of the integrators is thus independent of the choice of decomposition in the frame.*

The central assumption that motivates the use of these methods is that while computing the exact flow of a SDE is difficult, computing the flow associated to a frozen vector field often is a much simpler problem for a good choice of frame. For instance, in \mathbb{R}^D with the standard frame basis $E_i = \partial_i$, the frozen flow is the Euler method: $\exp(G_x)p = p + G(x)$. A second example is given by homogeneous manifolds. In this context, a natural choice of frame is derived from a basis (A_d) of the Lie algebra \mathfrak{g} by $E_d(y) = A_d \cdot y$, where we use for simplicity the same notation \cdot for the Lie group and Lie algebra action on \mathcal{M} . On matrix homogeneous manifolds, the frozen flow is explicitly given by

$$\exp(G_x)p = \text{Exp} \left(\sum_d g^d(x) A_d \right) p,$$

where Exp is the matrix exponential. We emphasize that a significant part of the manifolds used in numerical experiments and applications of stochastic dynamics on manifolds are homogeneous manifolds, such as Lie-groups, spheres, the Stiefel manifold, symmetric positive definite matrices, ... We further discuss the implementation of the methods in Section 5.1.

We consider the following new class of stochastic frozen flow methods, in the spirit of [82] for the deterministic part and [84, 32, 55] for the stochastic part,

$$\begin{aligned} H_n^i &= \exp \left(\sum_{d=1}^D \left(h \sum_{j=1}^s Z_{i,j,K}^0 f^d(H_n^j) + \sqrt{h} Z_{i,K}^d \right) E_d \right) \dots \\ &\quad \dots \exp \left(\sum_{d=1}^D \left(h \sum_{j=1}^s Z_{i,j,1}^0 f^d(H_n^j) + \sqrt{h} Z_{i,1}^d \right) E_d \right) X_n, \\ X_{n+1} &= \exp \left(\sum_{d=1}^D \left(h \sum_{i=1}^s z_{i,K}^0 f^d(H_n^i) + \sqrt{h} z_K^d \right) E_d \right) \dots \\ &\quad \dots \exp \left(\sum_{d=1}^D \left(h \sum_{i=1}^s z_{i,1}^0 f^d(H_n^i) + \sqrt{h} z_1^d \right) E_d \right) X_n, \end{aligned} \tag{2.7}$$

where the coefficients $z_{i,K}^0, Z_{i,j,k}^0$ are deterministic and $z_k^d, Z_{i,k}^d$ are of the form

$$z_k^d = \sum_{l=1}^L \hat{z}_{k,l} \xi_n^{d,l}, \quad Z_{i,k}^d = \sum_{l=1}^L \hat{Z}_{i,k,l} \xi_n^{d,l},$$

where the $\xi^{d,l}$ are independent standard Gaussian random variables or approximations of Gaussian random variables with the same $2p$ first moments, vanishing odd moments, and finite

moments of all order. The dependence in n of the Runge-Kutta coefficients $z_k^d, Z_{i,k}^d$ is omitted for clarity, and we enforce that the random coefficients between two steps in time are independent. The constant K is the number of exponentials used per stage. It is well-known [47] that choosing $K = 1$ does not allow to reach arbitrarily high order already in the context of ODEs. In particular, the simplest method is the frozen flow Euler method

$$X_{n+1} = \exp\left(\sum_{d=1}^D \left(hf^d(X_n) + \sqrt{2h}\xi_n^d\right)E_d\right)X_n. \quad (2.8)$$

The assumptions of Subsection 2.1 are in general not sufficient to ensure bounded moments and bounds on the Talay-Tubaro remainder of the numerical integrators. Aside from the compactness assumption, the following assumption is sufficient.

Proposition 2.11. *Assume that the frame $E_d, d = 1, \dots, D$, and the vector field F satisfy Assumption 2.1 and*

$$\left|E_{d_1}[E_{d_1}[f^d]]\right| \leq C.$$

Assume further that for an infinite number of arbitrarily large q :

$$\left|E_{d_k}[\dots E_{d_1}[r^q] \dots]\right| \leq C(1 + r^{q-k}), \quad k \leq 4.$$

Then, all consistent explicit frozen flow methods of the form (2.7) satisfy the bounded moments property and the Talay-Tubaro expansion of Assumptions 2.6 and 2.7.

Proof. Let us first show that the first remainder in the expansion (2.3) acts on the distance function by

$$\left|R_1^h(r^q, x)\right| \leq C(1 + r(x)^q), \quad (2.9)$$

for an infinite number of arbitrarily large q . Without loss of generality, we consider the following method, where $G_y = \sum_{d=1}^D g_y^d E_d$ and $g_y^d = hf^d(y) + \sqrt{2h}\xi_n^d$,

$$X_{n+1} = \exp(G_{H_n})X_n, \quad H_n = \exp(G_{X_n})X_n.$$

A Taylor expansion around $x = X_0$ yields (see Section 3 for the detailed calculations),

$$f^d(H) = f^d(x) + G_x[\phi](x) + \int_0^1 (1-t)g_x^{d_1}g_x^{d_2}E_{d_2}[E_{d_1}[f^d]](\exp(tG_x)x)dt,$$

and we obtain

$$\left|f^d(H)\right| \leq C(\xi^d)(1 + r(x)), \quad \left|\mathbb{E}[f^d(H) - f^d(x)]\right| \leq Ch(1 + r(x)),$$

where $C(\xi^d)$ is a polynomial in $|\xi^d|$. Then, a Taylor expansion for $\phi = r^q$ yields

$$\begin{aligned} \mathbb{E}[\phi(X_1)] &= \phi(x) + h\mathcal{L}\phi(x) + h^2R_1^h(\phi, x) \\ R_1^h(\phi, x) &= \mathbb{E}\left[\frac{1}{h}(f^d(H) - f^d(x))E_d[\phi](x) + \frac{1}{2}f^{d_1}(H)f^{d_2}(H)E_{d_2}[E_{d_1}[\phi]](x) \right. \\ &\quad + \frac{1}{3}f^{d_1}(H)\xi^{d_2}\xi^{d_3}(E_{d_3}[E_{d_2}[E_{d_1}[\phi]]] + E_{d_3}[E_{d_1}[E_{d_2}[\phi]]] + E_{d_1}[E_{d_3}[E_{d_2}[\phi]]])(x) \\ &\quad \left. + \frac{h}{6}f^{d_1}(H)f^{d_2}(H)f^{d_3}(H)E_{d_3}[E_{d_2}[E_{d_1}[\phi]]](x) \right] \end{aligned}$$

$$+ \frac{1}{h^2} \int_0^1 \frac{(1-t)^3}{3!} g_H^{d_1} g_H^{d_2} g_H^{d_3} g_H^{d_4} E_{d_4}[E_{d_3}[E_{d_2}[E_{d_1}[\phi]]]](\exp(tG_H)x) dt.$$

The estimate on $R_1^h(\phi, x)$ is straightforwardly obtained from (2.6). Indeed, one shows that the distance function satisfies

$$r(\exp(tG_H)x) \leq C(\xi^d)(1 + r(x)),$$

where $C(\xi^d)$ is a polynomial in $|\xi^d|$.

The expansion (2.3) is computed explicitly in Theorem 3.15. The estimate on the remainder is obtained thanks to the regularity Assumption 2.3. For Assumption 2.7, we use the Talay-Tubaro expansion (2.3) applied to $\phi = r^q$:

$$\mathbb{E}[r(X_{n+1})^q] = \mathbb{E}[r(X_n)^q] + h\mathbb{E}[\mathcal{L}r(X_n)^q] + h^2\mathbb{E}[R_1^h(r^q, X_n)],$$

where $|\mathbb{E}[\mathcal{L}\phi(X_n)]| \leq C(1 + \mathbb{E}[r(X_n)^q])$ by Proposition 2.2 and $|\mathbb{E}[R_1^h(r^q, X_n)]| \leq C(1 + \mathbb{E}[r(X_n)^q])$ by equation (2.9) (for q large enough). Thus, we find

$$\mathbb{E}[r(X_{n+1})^q] \leq C(1 + h)\mathbb{E}[r(X_n)^q] + Ch,$$

and a standard discrete Grönwall argument yields the bounded moments (2.4). \square

2.4 High-order frozen flow integrators

Let us present the order conditions of the new class of frozen flow methods (2.7) for solving the stochastic dynamics (1.1), as well as an example of second order frozen flow integrator.

We use the notation of factorial of multi-indices and sums for writing the order conditions.

Definition 2.12. *If $\mathbf{k} \in \{1, \dots, K\}^n$ is a multi-index that takes r_1 times the value k_1, \dots, r_p times the value k_p , the factorial of \mathbf{k} is*

$$\mathbf{k}! = r_1! \dots r_p!.$$

The factorial sum is given for a finite set $S \subset \{1, \dots, K\}^n$ of multi-indices and a functional f by $\sum_{\mathbf{k} \in S} \mathbf{k}! f(\mathbf{k}) = \sum_{\mathbf{k} \in S} \frac{1}{\mathbf{k}!} f(\mathbf{k})$.

The weak order conditions for stochastic Runge-Kutta methods in \mathbb{R}^d are presented in [84, 32, 55, 14] and require 4 order conditions for weak order two. On embedded manifolds, the approach with projection methods of [56] yields 28 order conditions, applies only in codimension one, and relies on an extrinsic approach with tedious calculations (see [56, App.D]). The approach proposed here is intrinsic, versatile, generalises the order theory of Lie-group methods to the stochastic setting, and yields a surprisingly low number of order conditions for weak order two.

Theorem 2.13. *Let a frozen flow integrator of the form (2.7) that satisfies the order conditions in Table 1, then, under Assumption 2.3 and the assumptions of Proposition 2.11, the integrator has local weak order two for solving (1.1). The additional order conditions for weak order three are presented in Appendix A.*

Theorem 2.13 is a direct consequence of Theorem 3.13 and Theorem 3.15, postponed to Section 3. The approach relies on a new algebraic framework of exotic planar forests and series that gives the explicit expression of the operators \mathcal{A}_j in (2.5). It allows us to provide the order conditions for arbitrarily high weak order.

Exotic forest π	Differential $\mathbb{F}(\pi)[\phi]$	Order condition $a(\pi) = e(\pi)$
\bullet	$f^i E_i[\phi]$	$z_{i,k_1}^0 = 1$
$\textcircled{1}\textcircled{1}$	$E_{d_1}[E_{d_1}[\phi]]$	$\sum_{k_1 \geq k_2}^! \mathbb{E}[z_{k_2}^{d_1} z_{k_1}^{d_1}] = 1$
\bullet	$f^j E_j[f^i] E_i[\phi]$	$z_{i,k_2}^0 z_{i,j,k_1}^0 = \frac{1}{2}$
$\textcircled{1}\textcircled{1}$	$E_{d_1}[E_{d_1}[f^i]] E_i[\phi]$	$\sum_{k_2 \geq k_3}^! \mathbb{E}[Z_{i,k_3}^{d_1} Z_{i,k_2}^{d_1}] z_{i,k_1}^0 = \frac{1}{2}$
$\bullet \bullet$	$f^j f^i E_j[E_i[\phi]]$	$\sum_{k_1 \geq k_2}^! z_{j,k_2}^0 z_{i,k_1}^0 = \frac{1}{2}$
$\textcircled{1}$	$E_{d_1}[f^i] E_i[E_{d_1}[\phi]]$	$\sum_{k_1 \geq k_2}^! z_{i,k_2}^0 \mathbb{E}[Z_{i,k_3}^{d_1} z_{k_1}^{d_1}] = 0$
$\textcircled{1}\textcircled{1}$	$E_{d_1}[f^i] E_{d_1}[E_i[\phi]]$	$\sum_{k_2 \geq k_1}^! z_{i,k_2}^0 \mathbb{E}[Z_{i,k_3}^{d_1} z_{k_1}^{d_1}] = 1$
$\bullet \textcircled{1}\textcircled{1}$	$f^i E_i[E_{d_1}[E_{d_1}[\phi]]]$	$\sum_{k_1 \geq k_2 \geq k_3}^! z_{i,k_3}^0 \mathbb{E}[z_{k_2}^{d_1} z_{k_1}^{d_1}] = \frac{1}{2}$
$\textcircled{1} \bullet \textcircled{1}$	$f^i E_{d_1}[E_i[E_{d_1}[\phi]]]$	$\sum_{k_1 \geq k_3 \geq k_2}^! z_{i,k_3}^0 \mathbb{E}[z_{k_2}^{d_1} z_{k_1}^{d_1}] = 0$
$\textcircled{1}\textcircled{1} \bullet$	$f^i E_{d_1}[E_{d_1}[E_i[\phi]]]$	$\sum_{k_3 \geq k_1 \geq k_2}^! z_{i,k_3}^0 \mathbb{E}[z_{k_2}^{d_1} z_{k_1}^{d_1}] = \frac{1}{2}$
$\textcircled{2}\textcircled{2}\textcircled{1}\textcircled{1}$	$E_{d_2}[E_{d_2}[E_{d_1}[E_{d_1}[\phi]]]]$	$\sum_{k_1 \geq k_2 \geq k_3 \geq k_4}^! \mathbb{E}[z_{k_4}^{d_2} z_{k_3}^{d_2}] \mathbb{E}[z_{k_2}^{d_1} z_{k_1}^{d_1}] = \frac{1}{2}$
$\textcircled{2}\textcircled{1}\textcircled{2}\textcircled{1}$	$E_{d_2}[E_{d_1}[E_{d_2}[E_{d_1}[\phi]]]]$	$\sum_{k_1 \geq k_2 \geq k_3 \geq k_4}^! \mathbb{E}[z_{k_4}^{d_2} z_{k_2}^{d_2}] \mathbb{E}[z_{k_3}^{d_1} z_{k_1}^{d_1}] = 0$
$\textcircled{1}\textcircled{2}\textcircled{2}\textcircled{1}$	$E_{d_1}[E_{d_2}[E_{d_2}[E_{d_1}[\phi]]]]$	$\sum_{k_1 \geq k_2 \geq k_3 \geq k_4}^! \mathbb{E}[z_{k_3}^{d_2} z_{k_2}^{d_2}] \mathbb{E}[z_{k_4}^{d_1} z_{k_1}^{d_1}] = 0$

Table 1: Order conditions of frozen flow methods up to weak order 2. The sums on the indices that do not satisfy inequalities (except the d_i) are omitted for clarity. The order conditions do not depend on the dimension of the problem.

Remark 2.14. *We shall prove in Section 3.5 that the coefficient maps a and e of the numerical and exact flow are characters of the shuffle algebra over exotic forests $(\mathcal{EF}, \sqcup, \Delta)$. Following [81], an important consequence is that the order conditions are not independent. Up to second order, the following "shuffle relations" are satisfied (and similarly for e):*

$$\begin{aligned}
a(\bullet)^2 &= 2a(\bullet \bullet), \\
a(\bullet)a(\textcircled{1}\textcircled{1}) &= a(\bullet \textcircled{1}\textcircled{1}) + a(\textcircled{1}\textcircled{1} \bullet) + a(\textcircled{1} \bullet \textcircled{1}), \\
a(\textcircled{1}\textcircled{1})^2 &= 2a(\textcircled{2}\textcircled{2}\textcircled{1}\textcircled{1}) + 2a(\textcircled{2}\textcircled{1}\textcircled{2}\textcircled{1}) + 2a(\textcircled{1}\textcircled{2}\textcircled{2}\textcircled{1}).
\end{aligned}$$

The minimal number of order conditions is not given by the number of entries in Table 1, as one has to subtract the number of shuffle relations. In particular, there are 2 conditions for order 1, 8 conditions for order 2, and 70 conditions for order 3 (against 2, 11, and 92 exotic forests of order 1, 2, and 3). This is further discussed in Remark 3.16.

Let us now provide an example of a second order method. Let $\xi_n^{d,1}, \xi_n^{d,2}, d = 1, \dots, D$, be bounded independent random variables satisfying

$$\mathbb{E}[(\xi_n^{d,l})^2] = 1, \quad \mathbb{E}[(\xi_n^{d,l})^4] = 3, \quad \mathbb{E}[(\xi_n^{d,l})^{2q+1}] = 0, \quad q = 0, 1, 2, \dots$$

For weak order two, one can use for instance

$$\mathbb{P}(\theta = 0) = \frac{2}{3}, \quad \mathbb{P}(\theta = \pm\sqrt{3}) = \frac{1}{6}$$

so that the first four moments coincide with the ones of standard Gaussians.

The new method is the following. It is explicit, intrinsic, uses two evaluations of the vector field F per step, and three frozen flow exponentials.

Method SFF2 (New explicit frozen flow integrator of weak order two for equation (1.1))

$X_0 \in \mathcal{M}$

for $n = 0, \dots, N$ **do**

$$H_n^1 = X_n$$

$$H_n^2 = \exp \left(\sum_{d=1}^D \left(\frac{1}{2} h f^d(H_n^1) + \sqrt{h} \xi_n^{d,1} \right) E_d \right) X_n \quad (2.10)$$

$$X_{n+1} = \exp \left(\sum_{d=1}^D \left(\left(\frac{\sqrt{2}}{2} - 1 \right) h f^d(H_n^1) + (2 - \sqrt{2}) h f^d(H_n^2) + (1 - \sqrt{2}) \sqrt{h} \xi_n^{d,1} + \sqrt{h} \xi_n^{d,2} \right) E_d \right) \\ \exp \left(\sum_{d=1}^D \left(\left(1 - \frac{\sqrt{2}}{2} \right) h f^d(H_n^1) + (\sqrt{2} - 1) h f^d(H_n^2) + \sqrt{2h} \xi_n^{d,1} \right) E_d \right) X_n.$$

end for

Example 2.15. *Brownian dynamics on \mathcal{M} are defined as the solution of*

$$dX(t) = -\frac{1}{2} \sum_{d=1}^D (\nabla_{E_d} E_d)(X(t)) dt + \sum_{d=1}^D E_d(X(t)) \circ dW_d(t),$$

where E_d is an orthonormal frame of \mathcal{M} . In the specific case of a Lie group equipped with a bi-invariant metric, we find $\sum_d \nabla_{E_d} E_d = 0$ and a second order integrator is

$$X_{n+1} = \exp \left(\sum_{d=1}^D \left(\left(\frac{\sqrt{2}}{2} - 1 \right) \sqrt{h} \xi_n^{d,1} + \frac{\sqrt{2}}{2} \sqrt{h} \xi_n^{d,2} \right) E_d \right) \exp \left(\sum_{d=1}^D \sqrt{h} \xi_n^{d,1} E_d \right) X_n. \quad (2.11)$$

3 Order theory of stochastic frozen flow methods

This section is devoted to the formalisation of the order theory of stochastic frozen flow methods with the help of the new algebraic formalism of exotic Lie series governing the Taylor expansions of the exact flow and the frozen flow integrators.

3.1 Connection algebra of vector fields

For a given initial value, the new integrators (with bounded random variables) evolve in one step in an open neighbourhood U of \mathcal{M} . Taking the stepsize small enough, we can extract a frame basis from the frame (E_d) . Following Remark 2.10, we do not lose any generality in assuming that (E_d) is a frame basis. We shall write our statements globally on \mathcal{M} for the sake of simplicity, and all of them are valid in a neighbourhood of the starting point of the method.

Define the product $\triangleright: \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \rightarrow \mathfrak{X}(\mathcal{M})$ by

$$F \triangleright G = \sum_{d=1}^D F[g^d] E_d, \quad G = \sum_{d=1}^D g^d E_d.$$

The product is well-posed as there is a unique decomposition of G in the frame basis. The operation \triangleright defines a new connection on $\mathfrak{X}(\mathcal{M})$, which does not coincide with the Levi-Civita connection ∇ . We extend the notation to test functions: $F \triangleright \phi = F[\phi]$ (independent of the choice of connection).

Remark 3.1. *The frozen flows $p \mapsto \exp(v)p$ are the geodesics for the connection \triangleright . Moreover, the map $x \mapsto \exp(G_x)$ is the parallel transport of the vector field G along the geodesics.*

Define the bracket $[-, -]$ by

$$[F, G] = \llbracket F, G \rrbracket - F \triangleright G + G \triangleright F,$$

where $\llbracket F, G \rrbracket$ is the Jacobi bracket of vector fields:

$$\llbracket F, G \rrbracket \triangleright \phi = F \triangleright (G \triangleright \phi) - G \triangleright (F \triangleright \phi).$$

The torsion and curvature associated to \triangleright are $T(F, G) = -[F, G]$ and $R = 0$, which is an additional difference between \triangleright and ∇ , as the Levi-Civita connection has vanishing torsion and non-vanishing curvature in general.

Let us now consider the associative non-commutative concatenation product \cdot as the differential operator

$$(F_n \cdots F_1) \triangleright \phi = \sum_{i_1, \dots, i_n} f_n^{i_n} \cdots f_1^{i_1} E_{i_n} [\dots E_{i_1} [\phi] \cdots],$$

defined on the tensor algebra $(T(\mathfrak{X}(\mathcal{M})), \cdot)$. This product is called the frozen composition of differential operators.

We shall see that the Taylor expansions of the numerical and exact flows for the equation (1.1) typically write in $T(\mathfrak{X}(\mathcal{M}))$. The number and complexity of the differential operators involved in the Taylor expansions of numerical and exact flows motivate the introduction of decorated and exotic planar forests.

3.2 Decorated planar forests

We introduce decorated planar forests and decorated Lie series.

Definition 3.2. *A decorated planar tree τ is a connected oriented graph in which every node has an outgoing edge except for a node called a root. An ordered, possibly empty, collection of decorated planar trees is called a decorated planar forest. The sets of decorated planar trees and forests are denoted by T_D and F_D , respectively.*

The vector spaces spanned by T_D and F_D are denoted by \mathcal{T}_D and \mathcal{F}_D . Decorated planar trees \mathcal{T}_D are used as a convenient formalism for representing vector fields $\mathfrak{X}(\mathcal{M})$ that appear in a Taylor expansion of a numerical or exact solution of a differential equation, while decorated planar forests \mathcal{F}_D are used to represent differential operators $T(\mathfrak{X}(\mathcal{M}))$ that appear in a Taylor expansion of a test function applied to a numerical or exact solution. We note that $\mathcal{F}_D = T(\mathcal{T}_D)$.

We use decorated planar forests \mathcal{F}_D with $D = \{\bullet, \circ^1, \dots, \circ^L\}$ to represent the Taylor expansion of a test function applied to the one-step of a frozen flow method introduced in Section 2. We consider a map $\mathbb{F}^{dec} : \mathcal{F}_D \rightarrow T(\mathfrak{X}(\mathcal{M}))$ defined as follows.

Definition 3.3. *Let the map $\mathbb{F}^{dec} : \mathcal{F}_D \rightarrow T(\mathfrak{X}(\mathcal{M}))$ be defined for $\pi \in \mathcal{F}_D$ and $\phi \in C^\infty(\mathcal{M})$ as*

$$\mathbb{F}^{dec}(\pi) \triangleright \phi := \sum_{\substack{i_v=1 \\ v \in V(\pi)}}^D \left(\prod_{v \in V} E_{I_\Pi(v)} [\mathbb{F}^{dec}(v)^{i_v}] \right) E_{I_R}[\phi],$$

with $\mathbb{F}^{dec}(\bullet) = f$ and $\mathbb{F}^{dec}(\circ^l) = \xi^{d,l} E_d$ where R is the set of roots of π , V is the set of nodes of π , $\Pi(v)$ is the set of predecessors of v ordered from right to left, $E_I[\phi] = E_{i_p}[\dots E_{i_1}[\phi] \dots]$ for $I = (i_p, \dots, i_1)$. The notation can also be frozen for $x, p \in \mathcal{M}$,

$$(\mathbb{F}_x^{dec}(\pi) \triangleright \phi)(p) := \sum_{\substack{i_v=1 \\ v \in V(\pi)}}^D \left(\prod_{v \in V} E_{I_{\Pi(v)}}[\mathbb{F}^{dec}(v)^{i_v}](x) \right) E_{I_R}[\phi](p),$$

For example, $\mathbb{F}^{dec}(\begin{smallmatrix} 1 \\ \bullet \end{smallmatrix}) \triangleright \phi = \xi^{1,i} E_i[f^j] f^k E_j[E_k[\phi]]$. Decorated Lie series are formal series of differential operators indexed by decorated planar forests. They do not converge in general, but their truncations at any order belong to $T(\mathfrak{X}(\mathcal{M}))$.

Definition 3.4. Given a one form $a \in \mathcal{F}_D^*$, called a coefficient map, a decorated Lie S -series is the following formal power series in h in $T(\mathfrak{X}(\mathcal{M}))$,

$$S_h^{dec}(a) \triangleright \phi := \sum_{\pi \in \mathcal{F}_D} h^{|\pi|} a(\pi) \mathbb{F}^{dec}(\pi) \triangleright \phi, \quad S^{dec}(a) = S_1^{dec}(a).$$

A decorated Lie series is only indexed on decorated trees and is the formal vector field,

$$B_h^{dec}(a) := \sum_{\tau \in \mathcal{T}_D} h^{|\tau|} a(\tau) \mathbb{F}^{dec}(\tau), \quad B^{dec}(a) = B_1^{dec}(a).$$

These series can also be frozen:

$$S_{h,p}^{dec}(a) \triangleright \phi := \sum_{\pi \in \mathcal{F}_D} h^{|\pi|} a(\pi) \mathbb{F}_p^{dec}(\pi) \triangleright \phi, \quad B_{h,p}^{dec}(a) := \sum_{\tau \in \mathcal{T}_D} h^{|\tau|} a(\tau) \mathbb{F}_p^{dec}(\tau).$$

We endow the space \mathcal{F}_D with the shuffle product \sqcup defined inductively for $\tau, \gamma \in \mathcal{T}_D$ and $\pi, \eta \in \mathcal{F}_D$ as

$$\tau\pi \sqcup \gamma\eta = \tau(\pi \sqcup \gamma\eta) + \gamma(\tau\pi \sqcup \eta).$$

For example,

$$\begin{aligned} \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix} \sqcup \bullet \begin{smallmatrix} 2 \\ \bullet \end{smallmatrix} &= \begin{smallmatrix} 1 \\ \bullet \bullet \end{smallmatrix} \begin{smallmatrix} 2 \\ \bullet \end{smallmatrix} + \bullet \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix} \begin{smallmatrix} 2 \\ \bullet \end{smallmatrix} + \bullet \begin{smallmatrix} 2 \\ \bullet \end{smallmatrix} \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix}, \\ \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix} \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix} \sqcup \bullet \begin{smallmatrix} 2 \\ \bullet \end{smallmatrix} &= \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix} \begin{smallmatrix} 1 \\ \bullet \bullet \end{smallmatrix} \begin{smallmatrix} 2 \\ \bullet \end{smallmatrix} + \begin{smallmatrix} 1 \\ \bullet \bullet \end{smallmatrix} \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix} \begin{smallmatrix} 2 \\ \bullet \end{smallmatrix} + \begin{smallmatrix} 1 \\ \bullet \bullet \end{smallmatrix} \begin{smallmatrix} 2 \\ \bullet \end{smallmatrix} \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix} + \bullet \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix} \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix} \begin{smallmatrix} 2 \\ \bullet \end{smallmatrix} + \bullet \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix} \begin{smallmatrix} 2 \\ \bullet \end{smallmatrix} \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix} + \bullet \begin{smallmatrix} 2 \\ \bullet \end{smallmatrix} \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix} \begin{smallmatrix} 1 \\ \bullet \end{smallmatrix}. \end{aligned}$$

A case of particular interest is when the coefficient map a is a character over (\mathcal{F}_D, \sqcup) , that is, when $a(\pi_1 \sqcup \pi_2) = a(\pi_1)a(\pi_2)$. This condition is satisfied by the frozen flow, as expressed in Proposition 3.5.

Proposition 3.5. Let $\tilde{a} \in \mathcal{T}_D^*$, the frozen flow associated to the vector field $B^{dec}(\tilde{a})$ satisfies

$$\phi(\exp(B_p^{dec}(\tilde{a}))) = \exp \cdot (B^{dec}(\tilde{a})) \triangleright \phi = S^{dec}(a) \triangleright \phi,$$

where $a \in \mathcal{F}_D^*$ extends \tilde{a} to a character of (\mathcal{F}_D, \sqcup) , that is, $a(\pi_1 \sqcup \pi_2) = a(\pi_1)a(\pi_2)$. Moreover,

$$a(\tau_1 \cdots \tau_n) = \frac{1}{n!} \tilde{a}(\tau_1) \cdots \tilde{a}(\tau_n). \quad (3.1)$$

Proof. The proof is analogous to the proof of Lemma 2.4 from [81]. The fact that $a \in \mathcal{F}_D^*$ is a character with respect to the shuffle product follows from the formula (3.1) using the fact that there are $\binom{n+m}{n}$ terms in $\pi \sqcup \eta$ with π and η having n and m trees, respectively. \square

Remark 3.6. *Representing Taylor expansions of vector fields with formal series indexed by tree-like structures is very natural from the geometric point of view. In [67, 75, 54], it is shown that exotic, aromatic and standard B-series are completely characterised by universal geometric properties of equivariance and locality. The extension of these properties to the planar case are open questions, even for standard Lie-Butcher series.*

3.3 Exotic planar forests

This section adds planarity to the exotic forests from [55, 14, 15]. We use decorated planar forests to define exotic planar forests in Definition 3.7 which are the main object of interest for representing weak stochastic Taylor expansions. Exotic planar forests are decorated planar forests in which the decoration is used to denote the pairing between certain leaves.

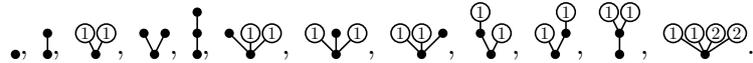
Definition 3.7. *An exotic forest is a decorated forest $(\pi, \alpha) \in F_D$ with the decorations $D = \{\bullet\} \cup \mathbb{N}$, $\mathbb{N} = \{1, 2, 3, \dots\}$, that follows the following rules. If a natural number is used as a decoration, then it must decorate exactly two leaves, that is, $|\alpha^{-1}(n)| \in \{0, 2\}$ for all $n \in \mathbb{N}$. Two exotic forests (π_1, α_1) and (π_2, α_2) are considered to be identical if $\pi_1 = \pi_2$ and there exists a map $\varphi : D \rightarrow D$ with $\phi|_{\mathbb{N}} \in \mathcal{S}_{\mathbb{N}}$ and $\varphi(\bullet) = \bullet$ such that $\alpha_1 = \varphi \circ \alpha_2$. The set of numbered leaves that correspond to the same number, that is $\alpha^{-1}(n)$ for $n \in \mathbb{N}$, is called a liana.*

We gather the numbered nodes in V_\circ and the black nodes in V_\bullet . The order of an exotic forest π is then defined as

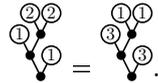
$$|\pi| = |V_\bullet| + \frac{|V_\circ|}{2}.$$

The set of exotic forests is denoted by EF . The set of all exotic forests of order p is EF^p , and the ones of order less than or equal to p is $EF^{\leq p}$. Exotic forests with one root are called exotic trees and form a set denoted ET . The corresponding vector spaces are denoted by \mathcal{EF} and \mathcal{ET} , and similarly for the graded subspaces \mathcal{EF}^p and $\mathcal{EF}^{\leq p}$.

The exotic trees of order up to three are the following. Note that the order does not coincide with the number of black nodes in general.



We also note that the choice of the number used to denote the liana does not matter, for example, the following exotic forests are identical:



Our main focus with exotic planar forests is to represent some specific elementary differentials in $\mathcal{T}(\mathfrak{X}(\mathcal{M}))$ and to translate the connection algebra to a natural algebraic structure on \mathcal{EF} .

The vector space \mathcal{EF} of exotic forests forms an algebra (\mathcal{EF}, \cdot) w.r.t. the non-commutative concatenation product. We emphasize that it does not coincide with the tensor algebra $T(\mathcal{ET})$ as for instance \circledast cannot be expressed as the concatenation of exotic trees.

We introduce the new formalism of exotic Lie series for the study of order conditions of stochastic frozen flow methods. The formalism mixes the features of the exotic formalism [55, 51, 14, 15] for the stochastic part and the planar forests formalism [47, 81, 79] for Lie-group methods.

Definition 3.8. Let the map $\mathbb{F} : \mathcal{EF} \rightarrow T(\mathcal{X}(\mathcal{M}))$ for $\pi \in EF$ and $\phi \in \mathcal{C}^\infty(\mathcal{M})$ be defined as

$$\mathbb{F}(\pi) \triangleright \phi = \sum_{\substack{i_v=1,\dots,D \\ v \in V}} \delta_{i_L} \left(\prod_{v \in V} E_{I_{\Pi(v)}}[f^{i_v}] \right) E_{I_R}[\phi],$$

where R is the set of roots, $\Pi(v)$ is the set of predecessors of v ordered from right to left, $E_I = E_{i_p} \dots E_{i_1}$ for $I = i_p, \dots, i_1$, and $\delta_{i_L} = 0$ if there exists $i_v \neq i_w$ with $\alpha(v) = \alpha(w)$ (that is, (v, w) is a liana), and is 1 else. The notation can also be frozen for $x, p \in \mathcal{M}$:

$$(\mathbb{F}_x(\pi) \triangleright \phi)(p) = \sum_{\substack{i_v=1,\dots,D \\ v \in V}} \delta_{i_L} \left(\prod_{v \in V} E_{I_{\Pi(v)}}[f^{i_v}](x) \right) E_{I_R}[\phi](p).$$

Exotic Lie S -series and exotic Lie-Butcher series are defined analogously to Definition 3.4: given a one form $a \in \mathcal{EF}^*$ (resp. $a \in \mathcal{ET}^*$), an exotic Lie S -series (resp. exotic Lie-Butcher series) is the following formal series

$$S_h(a) \triangleright \phi := \sum_{\pi \in EF} h^{|\pi|} a(\pi) \mathbb{F}(\pi) \triangleright \phi, \quad B_h(a) := \sum_{\tau \in ET} h^{|\tau|} a(\tau) \mathbb{F}(\tau),$$

and analogously for their frozen counterparts $S_{h,p}$ and $B_{h,p}$.

The exotic series arise naturally by considering the expectation of decorated series.

Proposition 3.9. Let a^{dec} be a character on (\mathcal{F}_D, \sqcup) with $D = \{\bullet, \circ^1, \dots, \circ^L\}$, then

$$\mathbb{E}[S^{\text{dec}}(a^{\text{dec}}) \triangleright \phi] = S(a) \triangleright \phi,$$

where $a = a^{\text{dec}} \circ \Phi$ is a character on \mathcal{EF} with respect to the shuffle product and $\Phi : \mathcal{EF} \rightarrow \mathcal{F}_D$ is the map that forgets the pairings of the white leaves.

For example, one finds

$$\mathbb{E} \left[\mathbb{F}^{\text{dec}} \left(\begin{array}{c} \circ^2 \quad \circ^2 \quad \circ^1 \quad \circ^1 \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \\ \bullet \quad \bullet \quad \bullet \quad \bullet \\ \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ^1 \quad \circ^1 \end{array} \right) \right] = \mathbb{F} \left(\begin{array}{c} \textcircled{3} \textcircled{3} \textcircled{2} \textcircled{2} \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \\ \textcircled{1} \quad \textcircled{1} \quad \textcircled{2} \quad \textcircled{2} \\ \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \textcircled{1} \quad \textcircled{1} \end{array} \right) + \mathbb{F} \left(\begin{array}{c} \textcircled{3} \textcircled{3} \textcircled{2} \textcircled{1} \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \\ \textcircled{1} \quad \textcircled{2} \quad \textcircled{2} \quad \textcircled{1} \\ \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \textcircled{1} \quad \textcircled{1} \end{array} \right) + \mathbb{F} \left(\begin{array}{c} \textcircled{3} \textcircled{3} \textcircled{1} \textcircled{2} \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \\ \textcircled{1} \quad \textcircled{2} \quad \textcircled{2} \quad \textcircled{1} \\ \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \textcircled{1} \quad \textcircled{1} \end{array} \right).$$

Proof. The proof is analogous to [14, Sect. 4.1] where the white leaves decorated by the same number are paired using the Isserlis formula [49]. That is,

$$\mathbb{E}[S^{\text{dec}}(a^{\text{dec}}) \triangleright \phi] = \mathbb{E} \left[\sum_{\pi \in \mathcal{F}_D} h^{|\pi|} a^{\text{dec}}(\pi) \mathbb{F}^{\text{dec}}(\pi) \triangleright \phi \right] = \sum_{\pi \in \mathcal{F}_D} h^{|\pi|} a^{\text{dec}}(\pi) \mathbb{E}[\mathbb{F}^{\text{dec}}(\pi)] \triangleright \phi.$$

We use Isserlis formula to write

$$\mathbb{E}[\mathbb{F}^{\text{dec}}(\pi, \alpha)] = \sum_{\alpha_e} \mathbb{F}(\pi, \alpha_e),$$

where (π, α) and (π, α_e) denote decorated and exotic planar forests as defined in Definitions 3.2 and 3.7 with the sum being over all α_e such that $\alpha_e^{-1}(\bullet) = \alpha^{-1}(\bullet)$ and $\alpha_e^{-1}(\mathbb{N})$ is a union of partitions into pairs of $\alpha^{-1}(\circ^l)$ for $l = 1, \dots, L$. We note that $\mathbb{E}[\mathbb{F}^{\text{dec}}(\pi)] = 0$ if π contains an odd number of nodes of color \circ^l for an $l \in \{1, \dots, L\}$. Therefore, we have,

$$\mathbb{E}[S^{\text{dec}}(a^{\text{dec}}) \triangleright \phi] = \sum_{\pi \in EF} h^{|\pi|} a^{\text{dec}}(\Phi(\pi)) \mathbb{F}(\pi) \triangleright \phi = S(a^{\text{dec}} \circ \Phi) \triangleright \phi,$$

and the statement is proved. \square

Example 3.10. Consider the Euler frozen flow method (2.8). Using Proposition 3.5, its Taylor expansion in decorated forests is given by

$$\begin{aligned} \phi(\exp(hF_p + \sqrt{2h}\xi^d E_d)p) &= \exp(hF + \sqrt{2h}\xi^d E_d) \triangleright \phi(p) = \mathbb{F}^{\text{dec}}(\exp(h\bullet + \sqrt{2h}\circ)) \triangleright \phi(p) \\ &= \left(\text{id} + h^{1/2}\sqrt{2}\xi^d E_d + h(F + \xi^{d_2}\xi^{d_1} E_{d_2} \cdot E_{d_1}) \right. \\ &\quad \left. + h^{3/2}\left(\frac{1}{\sqrt{2}}\xi^d F \cdot E_d + \frac{1}{\sqrt{2}}\xi^d E_d \cdot F + \frac{\sqrt{2}}{3}\xi^{d_3}\xi^{d_2}\xi^{d_1} E_{d_3} \cdot E_{d_2} \cdot E_{d_1}\right) + \dots \right) \triangleright \phi \\ &= \mathbb{F}^{\text{dec}}\left(\mathbf{1} + h^{1/2}\sqrt{2}\circ + h(\bullet + \circ \circ) + h^{3/2}\left(\frac{1}{\sqrt{2}}\bullet \circ + \frac{1}{\sqrt{2}}\circ \bullet + \frac{\sqrt{2}}{3}\circ \circ \circ\right) \right. \\ &\quad \left. + h^2\left(\frac{1}{2}\bullet \bullet + \frac{1}{3}\bullet \circ \circ + \frac{1}{3}\circ \bullet \circ + \frac{1}{3}\circ \circ \bullet + \frac{1}{6}\circ \circ \circ \circ\right) + \dots \right) \triangleright \phi(p). \end{aligned}$$

Then, the expectation pairs the decorated nodes together as detailed in Proposition 3.9 and yields

$$\begin{aligned} \mathbb{E}[\phi(\exp(hF_p + \sqrt{2h}\xi^d E_d)p)] &= \mathbb{F}\left(\mathbf{1} + h(\bullet + \circ \circ)\right) \\ &\quad + h^2\left(\frac{1}{2}\bullet \bullet + \frac{1}{3}\bullet \circ \circ + \frac{1}{3}\circ \bullet \circ + \frac{1}{3}\circ \circ \bullet + \frac{1}{6}\circ \circ \circ \circ + \frac{1}{6}\circ \circ \circ \circ + \frac{1}{6}\circ \circ \circ \circ\right) \triangleright \phi(p) \\ &= \left(\text{id} + h(F + E_d \cdot E_d) + h^2\left(\frac{1}{2}F \cdot F + \frac{1}{3}F \cdot E_d \cdot E_d + \frac{1}{3}E_d \cdot F \cdot E_d + \frac{1}{3}E_d \cdot E_d \cdot F \right. \right. \\ &\quad \left. \left. + \frac{1}{6}E_{d_2} \cdot E_{d_2} \cdot E_{d_1} \cdot E_{d_1} + \frac{1}{6}E_{d_2} \cdot E_{d_1} \cdot E_{d_2} \cdot E_{d_1} + \frac{1}{6}E_{d_1} \cdot E_{d_2} \cdot E_{d_2} \cdot E_{d_1}\right) + \dots \right) \triangleright \phi(p). \end{aligned}$$

3.4 Conversion of geometric operations into exotic Lie series

We saw in Proposition 3.5 that the frozen flow rewrites as a tree series operation. The following result, proved analogously to [81, Lemma 2.3], provides an algebraic description of the Taylor expansion of a frozen vector field.

Lemma 3.11. Let a smooth map $\varphi: \mathcal{M} \rightarrow \mathcal{M}$ with expansion $\phi(\varphi(p)) = (S(a) \triangleright \phi)(p)$ with $a \in \mathcal{EF}^*$, then the frozen vector field $F_\varphi: p \rightarrow F_{\varphi(p)}(p)$ satisfies

$$F_\varphi \triangleright \phi = (S(a) \triangleright hF) \triangleright \phi = B(\tilde{a}) \triangleright \phi, \quad \tilde{a}(\tau) = a(B^-(\tau)), \quad \tilde{a} \in \mathcal{ET}^*.$$

Let two smooth maps $\varphi^1, \varphi^2: \mathcal{M} \times \mathcal{M} \rightarrow \mathcal{M}$, typically $\varphi_x^k(p) = \exp(F_{\psi^k(x)}) \cdot p$ with $\psi^k: \mathcal{M} \rightarrow \mathcal{M}$. We use two possible ways to compose the φ^k . The standard composition of the maps would read

$$\varphi^1 \circ \varphi^2(p) = \varphi_{\varphi_p^2(p)}^1(\varphi_p^2(p)).$$

On the other hand, the frozen flow methods (2.7) rely on the frozen composition:

$$\varphi^1 \cdot \varphi^2(p) = \varphi_p^1 \cdot \varphi_p^2(p) = \varphi_p^1(\varphi_p^2(p)).$$

The first terms are the following

$$\exp(t\mathcal{L}) = \mathbb{F}\left(\mathbb{1} + t(\bullet + \textcircled{1}\textcircled{1}) + \frac{t^2}{2}\left(\begin{array}{c} \textcircled{1} \\ \bullet \\ \textcircled{1} \end{array} + \begin{array}{c} \textcircled{1} \\ \bullet \\ \textcircled{1} \end{array} + \dots + 2\begin{array}{c} \textcircled{1} \\ \bullet \\ \textcircled{1} \end{array} + \bullet\textcircled{1}\textcircled{1} + \textcircled{1}\textcircled{1}\bullet + \textcircled{2}\textcircled{2}\textcircled{1}\textcircled{1}\right) + \dots\right),$$

where the second order term is $\frac{t^2}{2}\mathcal{L}^2 = \frac{t^2}{2}S_t(l) * S_t(l) = \frac{t^2}{2}S_t(l * l)$.

Let us now present the expansion in exotic series of the frozen flow integrators. For the sake of clarity, the random Runge-Kutta coefficients $z_k^d, z_{i,k}^d$ are chosen as centered Gaussians. They can be replaced with bounded approximations of Gaussians with the same first moments up to order $2p$ and the expansion of the integrator will be an exotic series up to order p .

Definition 3.14. For $K > 0$, a labelling in $\text{Lab}_K(\pi)$ of a decorated forest π is a multi-index map $\mathbf{k}: V \rightarrow \{1, \dots, K\}$ such that $k_v \geq k_w$ if v and w are in the same tree, with same height and v is on the right of w .

For instance, a possible labelling in $\text{Lab}_8(\begin{array}{c} \textcircled{1} \\ \bullet \\ \textcircled{1} \end{array})$ is $\begin{array}{c} 3\textcircled{1} \\ 2\textcircled{1} \\ \bullet \\ 7 \\ 8 \end{array}$. Then, the Talay-Tubaro expansion of the new methods (2.7) is also given by an exotic Lie series.

Theorem 3.15. The Talay-Tubaro expansion of a frozen flow method of the form (2.7) is an exotic Lie series:

$$\mathbb{E}[\phi(X_1)|X_0 = p] = S_h(a) \triangleright \phi(p),$$

where a is given by

$$a(\pi) = \mathbb{E}\left[\sum_{\substack{i_v=1,\dots,s \\ v \in V}} \sum_{\mathbf{k} \in \text{Lab}_K(\pi)} \sum_{\substack{d_v=1,\dots,D \\ v \in V_0}} \delta_{d_L} \prod_{\substack{(v,w) \in E \\ v \in V_0}} Z_{i_w, i_v, k_v}^0 \prod_{\substack{(v,w) \in E \\ v \in V_0}} Z_{i_w, k_v}^{d_v} \prod_{r \in R \cap V_0} z_{i_r, k_r}^0 \prod_{r \in R \cap V_0} z_{k_r}^{d_r}\right],$$

where $\delta_{d_L} = 0$ if there exists $d_v \neq d_w$ with $(v, w) \in L$, and is 1 else. Moreover, the coefficient map a is a character on (\mathcal{EF}, \sqcup) , that is, $a(\pi_1 \sqcup \pi_2) = a(\pi_1)a(\pi_2)$.

The proof is a stochastic generalisation of [81].

Proof. Let us consider one step of the algorithm, with one Gaussian for the sake of simplicity: $Z_{i,k}^d = \hat{Z}_{i,k}\xi^d, z_k^d = \hat{z}_k\xi^d$. Let the intermediate steps

$$g_k^i = \exp\left(h \sum_{j=1}^s Z_{i,j,k}^0 f(H^j) + \sqrt{h} \hat{Z}_{i,k} \sum_d \xi_d E_d\right) \circ g_{k-1}^i, \quad g_0^i = \text{id}.$$

Theorem 3.12 yields $\phi(g_k^i) = S_h^{\text{dec}}(\alpha_k^i) \triangleright \phi$ and $\phi(H^i) = S_h^{\text{dec}}(\alpha_K^i) \triangleright \phi$ with $\alpha_0^i = \delta_{\mathbf{1}}, \alpha_k^i(\mathbf{1}) = 1$, and

$$\begin{aligned} \alpha_k^i(\tau_1 \dots \tau_n) &= \sum_{p=0}^n \frac{1}{(n-p)!} \alpha_{k-1}^i(\tau_1 \dots \tau_p) \left(\sum_{j_{p+1}=1}^s Z_{i, j_{p+1}, k}^0 \alpha_K^{j_{p+1}}(B^- \tau_{p+1}) \mathbb{1}_{\tau_{p+1} \neq \circ} + \hat{Z}_{i,k} \mathbb{1}_{\tau_{p+1} = \circ} \right) \\ &\dots \left(\sum_{j_n=1}^s Z_{i, j_n, k}^0 \alpha_K^{j_n}(B^- \tau_n) \mathbb{1}_{\tau_n \neq \circ} + \hat{Z}_{i,k} \mathbb{1}_{\tau_n = \circ} \right). \end{aligned}$$

An induction then yields

$$\alpha_K^i(\tau_1 \dots \tau_n) = \sum_{k_1 \leq \dots \leq k_n} \frac{1}{\mathbf{k}!} \left(\sum_{j_1=1}^s Z_{i, j_1, k_1}^0 \alpha_K^{j_1}(B^- \tau_1) \mathbb{1}_{\tau_1 \neq \circ} + \hat{Z}_{i, k_1} \mathbb{1}_{\tau_1 = \circ} \right)$$

$$\dots \left(\sum_{j_n=1}^s Z_{i,j_n,k_n}^0 \alpha_K^{j_n} (B^- \tau_n) \mathbb{1}_{\tau_n \neq \circ} + \hat{Z}_{i,k_n} \mathbb{1}_{\tau_n = \circ} \right).$$

Similar reasoning yields the desired expansion for the last stage of the method as $S_h^{\text{dec}}(a^{\text{dec}})$, where a^{dec} is a character by construction. Proposition 3.9 yields the expected formula for the coefficient map a , as well as the character property. \square

Examples of computations of the coefficient map a associated to frozen flow methods are presented in Table 1 and Appendix A. A more involved example of order 7 is the following:

$$\pi = \begin{array}{c} \textcircled{9} \\ \textcircled{7} \quad \textcircled{6} \quad \textcircled{5} \quad \textcircled{4} \\ \textcircled{8} \quad \textcircled{1} \quad \textcircled{1} \quad \textcircled{2} \\ \quad \quad \textcircled{3} \quad \textcircled{2} \quad \textcircled{1} \end{array}$$

$$a(\pi) = \mathbb{E} \sum_{i_1, \dots, i_9=1}^s \sum_{\substack{k_1 \geq k_2 \geq k_3 \\ k_4 \geq k_5, k_6 \geq k_7 \geq k_8}}^! \sum_{d_1, d_2=1}^D Z_{i_8, i_9, k_9}^{d_2} Z_{i_3, i_8, k_8}^0 Z_{i_3, i_7, k_7}^{d_1} Z_{i_3, i_6, k_6}^{d_1} Z_{i_1, i_5, k_5}^0 Z_{i_1, i_4, k_4}^0 z_{i_3, k_3}^0 z_{k_2}^{d_2} z_{i_1, k_1}^0$$

Remark 3.16. Order conditions for order p are obtained by requiring $(a - e)(\pi) = 0$ for all $\pi \in EF \setminus \{\mathbf{1}\}$ of order $|\pi| \leq p$. This results in the set of order conditions being indexed by the elements of $EF \setminus \{\mathbf{1}\}$. The fact that the coefficient maps a and e are characters of (\mathcal{EF}, \sqcup) induces relations between order conditions as detailed in Remark 2.14. In particular, $(a - e)(\pi \sqcup \eta) = 0$ if $(a - e)(\pi) = 0$ and $(a - e)(\eta) = 0$ using,

$$(a - e)(\pi \sqcup \eta) = a(\pi)a(\eta) - e(\pi)e(\eta) = 0.$$

Finding the minimal number of order conditions requires us to consider the quotient of \mathcal{EF} by the ideal $\mathcal{I} := \langle \pi \sqcup \eta : \pi, \eta \in \mathcal{EF} \setminus \{\mathbf{1}\} \rangle$. Using the fact that the order conditions corresponding to $\pi \sqcup \eta$ are satisfied automatically, we take $a - e$ to be a functional over \mathcal{EF}/\mathcal{I} , that is $a - e \in (\mathcal{EF}/\mathcal{I})^*$. Dualizing \mathcal{EF}/\mathcal{I} yields the space of primitive elements Prim of \mathcal{EF} with respect to the deshuffle coproduct Δ_{\sqcup} defined as

$$\text{Prim} := \text{Span}\{\pi \in \mathcal{EF} : \Delta_{\sqcup}(\pi) = \pi \otimes \mathbf{1} + \mathbf{1} \otimes \pi\}.$$

Therefore, using the identification $(a - e)(\pi) = \langle a - e, \pi \rangle$, we have $a - e \in \text{Prim}$. The detailed description of Prim is left for future work.

4 Exotic planar forests and their algebraic structure

In this section, we focus on the algebraic structures of exotic planar forests. We first recall the well-known post-Hopf structure of the connection algebra of vector fields in Section 4.1 and decorated planar forests in Section 4.2. We then explore how these structures change in the context of exotic planar forests in Section 4.3. In particular, we show that exotic planar forests form both the Grossman-Larson and Munthe-Kaas-Wright Hopf algebras. We also observe that the D-algebra structure is preserved. However, exotic planar forests do not form a post-Hopf algebra as a connected exotic forest can have multiple roots, which violates the Leibniz rule of the grafting product.

4.1 Post-Hopf algebra of vector fields

We recall the connection algebra introduced in Section 3.1 with the connection \triangleright . It is checked straightforwardly that the new connection then satisfies the following identities:

$$F \triangleright [G, H] = [F \triangleright G, H] + [G, F \triangleright H],$$

and

$$[F, G] \triangleright H = F \triangleright (G \triangleright H) - (F \triangleright G) \triangleright H - G \triangleright (F \triangleright H) + (G \triangleright F) \triangleright H.$$

Following [35, 40], we observe that $(\mathfrak{X}(\mathcal{M}), [-, -], \triangleright)$ has a structure of post-Lie algebra.

The deshuffle coproduct $\Delta_{\sqcup}: T(\mathfrak{X}(\mathcal{M})) \rightarrow T(\mathfrak{X}(\mathcal{M})) \otimes T(\mathfrak{X}(\mathcal{M}))$ is, for $A, B \in T(\mathfrak{X}(\mathcal{M}))$, $F \in \mathfrak{X}(\mathcal{M})$,

$$\Delta_{\sqcup}(AB) = \Delta_{\sqcup}A \cdot \Delta_{\sqcup}B, \quad \Delta_{\sqcup}(F) = \mathbf{1} \otimes F + F \otimes \mathbf{1},$$

where we recall $(A \otimes B) \cdot (C \otimes D) = (A \cdot C) \otimes (B \cdot D)$. Following the post-Lie extension [35] of the so-called Guin-Oudom process [80], the product \triangleright extends to $T(\mathfrak{X}(\mathcal{M}))$, for $A, B, C \in T(\mathfrak{X}(\mathcal{M}))$, $F, G \in \mathfrak{X}(\mathcal{M})$, by

$$\begin{aligned} \mathbf{1} \triangleright A &= A, & A \triangleright \mathbf{1} &= 0, \\ (F \cdot A) \triangleright G &= F \triangleright (A \triangleright G) - (F \triangleright A) \triangleright G, \\ A \triangleright (B \cdot C) &= (A_{(1)} \triangleright B) \cdot (A_{(2)} \triangleright C), \end{aligned}$$

where we use the Sweedler notation w.r.t. Δ_{\sqcup} . The triple $(T(\mathfrak{X}(\mathcal{M})), \cdot, \triangleright)$ is called a D-algebra [79]. The tensor algebra is equipped with the associative non-commutative Grossman-Larson product

$$A * B = A_{(1)} \cdot (A_{(2)} \triangleright B), \quad (A * B) \triangleright C = A \triangleright (B \triangleright C).$$

Then, $(T(\mathfrak{X}(\mathcal{M})), \cdot, \Delta_{\sqcup})$ and $(T(\mathfrak{X}(\mathcal{M})), *, \Delta_{\sqcup})$ are Hopf algebras. We note that the Hopf algebra $(T(\mathfrak{X}(\mathcal{M})), \cdot, \Delta_{\sqcup})$ equipped with the additional product \triangleright is a post-Hopf algebra with $(T(\mathfrak{X}(\mathcal{M})), *, \Delta_{\sqcup})$ being the respective subadjacent Hopf algebra [35, 62, 24].

By the Cartier-Quillen-Milnor-Moore theorem, the tensor algebra $T(\mathfrak{X}(\mathcal{M}))$ coincides with the universal enveloping algebra $\mathcal{U}(\mathfrak{X}(\mathcal{M}))$ of $(\mathfrak{X}(\mathcal{M}), [-, -])$. The concatenation and Grossman-Larson product satisfy, in particular, for $F, G \in \mathfrak{X}(\mathcal{M})$:

$$[F, G] \triangleright \phi = (F \cdot G - G \cdot F) \triangleright \phi, \quad \llbracket F, G \rrbracket \triangleright \phi = (F * G - G * F) \triangleright \phi.$$

4.2 Decorated planar forests

We recall the discussion of the decorated planar forests in Section 3.2. We present an alternative definition of the decorated planar forests in Definition 4.1.

Definition 4.1. *Let $\tau_1, \dots, \tau_n \in T_D$ be decorated planar trees, then, a decorated planar forest $\pi \in F_D$ is a concatenation $\tau_1 \cdots \tau_n$. Let $\pi \in F_D$ be a decorated planar forest, then, a decorated planar tree $\tau \in T_D$ is defined as $B_d^+(\pi)$ where the map $B_d^+: F_D \rightarrow T_D$ attaches all roots of π to a new root decorated by d .*

We define the left grafting operation $\curvearrowright: F_D \otimes F_D \rightarrow F_D$ by

$$\begin{aligned} \mathbf{1} \curvearrowright \pi &= \pi, & \pi \curvearrowright \mathbf{1} &= 0, \\ (\tau \cdot \pi_1) \curvearrowright \pi_2 &= \tau \curvearrowright (\pi_1 \curvearrowright \pi_2) - (\tau \curvearrowright \pi_1) \curvearrowright \pi_2, \end{aligned}$$

$$\pi \curvearrowright (\pi_1 \cdot \pi_2) = (\pi_{(1)} \curvearrowright \pi_1) \cdot (\pi_{(2)} \curvearrowright \pi_2),$$

with $\tau \in \mathcal{T}_D, \pi, \pi_1, \pi_2 \in \mathcal{F}_D$ and $\Delta_{\sqcup}(\pi) = \pi_{(1)} \otimes \pi_{(2)}$ using Sweedler notation. When restricted to trees, the left grafting operation is defined as

$$\tau \curvearrowright B_d^+(\pi) = B_d^+(\tau \cdot \pi) + B_d^+(\tau \curvearrowright \pi).$$

Let $\text{Lie}(\mathcal{T}_D)$ denote the space \mathcal{T} extended with the Lie bracket $[-, -]$, that is, given $\tau_1, \tau_2 \in \mathcal{T}_D$, we have $[\tau_1, \tau_2] = \tau_1 \cdot \tau_2 - \tau_2 \cdot \tau_1 \in \text{Lie}(\mathcal{T}_D)$, but $[\tau_1, \tau_2] \notin \mathcal{T}_D$. The space $(\text{Lie}(\mathcal{T}_D), [-, -], \curvearrowright)$ forms the free post-Lie algebra of primitive elements of the Hopf-algebra $(\mathcal{F}_D, \cdot, \Delta_{\sqcup})$. Consequently, $(\mathcal{F}_D, \cdot, \Delta_{\sqcup}, \curvearrowright)$ is a free post-Hopf algebra.

The universal property of $(\mathcal{F}_D, \cdot, \Delta_{\sqcup}, \curvearrowright)$ allows us to define a post-Hopf algebra homomorphism $\mathbb{F}^{\text{dec}} : \mathcal{F}_D \rightarrow \mathcal{T}(\mathfrak{X}(\mathcal{M}))$ by $\mathbb{F}^{\text{dec}}(\mathfrak{d}) = f^d$ satisfying

$$\mathbb{F}^{\text{dec}}(\pi_1 \cdot \pi_2) = \mathbb{F}^{\text{dec}}(\pi_1)\mathbb{F}^{\text{dec}}(\pi_2), \quad \mathbb{F}^{\text{dec}}(\pi_1 \curvearrowright \pi_2) = \mathbb{F}^{\text{dec}}(\pi_1) \triangleright \mathbb{F}^{\text{dec}}(\pi_2),$$

and we note that $B_d^+(\pi) = \pi \curvearrowright \mathfrak{d}$. We recall the notion of a D-algebra introduced in [79] and note that $(\mathcal{F}_D, \cdot, \curvearrowright)$ is the free D-algebra generated by the set D . That is, (\mathcal{F}_D, \cdot) forms a graded associative algebra and for all $\tau \in \mathcal{T}_D$ and $\pi, \eta \in \mathcal{F}_D$, we have, $\pi \curvearrowright \tau \in \mathcal{T}$, as well as,

$$\begin{aligned} \tau \curvearrowright (\pi \cdot \eta) &= (\tau \curvearrowright \pi) \cdot \eta + \pi \cdot (\tau \curvearrowright \eta), \\ \tau \curvearrowright (\pi \curvearrowright \eta) &= (\tau \curvearrowright \pi) \curvearrowright \eta + (\tau \cdot \pi) \curvearrowright \eta. \end{aligned}$$

We note that \mathbb{F} is also a D-algebra homomorphism.

Remark 4.2. *The space of decorated planar forests forms the tensor algebra $T(\mathcal{T}_D)$ of the space of decorated planar trees. It is the universal enveloping Lie algebra $\mathcal{F}_D = \mathcal{U}(\text{Lie}(\mathcal{T}_D))$ of the Lie algebra $\text{Lie}(\mathcal{T}_D)$ with the Lie bracket $[x, y] = x \cdot y - y \cdot x$ for $x, y \in \text{Lie}(\mathcal{T}_D)$. Moreover, $\text{Lie}(\mathcal{T}_D)$ is the space of primitive elements with respect to the deshuffle coproduct Δ_{\sqcup} , that is,*

$$\text{Lie}(\mathcal{T}_D) = \{\pi \in \mathcal{F}_D : \Delta_{\sqcup}(\pi) = \pi \otimes \mathbf{1} + \mathbf{1} \otimes \pi\}.$$

The Grossman-Larson product over decorated planar forests is defined as

$$\pi \diamond \eta := \pi_{(1)} \cdot (\pi_{(2)} \curvearrowright \eta), \quad \text{for } \pi, \eta \in \mathcal{F}_D,$$

and $\mathbb{F}^{\text{dec}}(\pi \diamond \eta) = \mathbb{F}^{\text{dec}}(\pi) * \mathbb{F}^{\text{dec}}(\eta)$ due to the fact that \mathbb{F}^{dec} is a D-algebra and post-Hopf algebra homomorphism. We generalize the concatenation and Grossman-Larson products by specifying the order in which the trees of the operands are concatenated. Let $\pi, \eta \in \mathcal{F}_D$ and let ω denote an ordering of the roots of $\pi \cdot \eta$ such that the order of roots in π and η is preserved, that is, if r_1 and r_2 are roots of π or η and $r_1 \leq r_2$, then $r_1 \leq r_2$ according to ω . We define $\cdot_{\omega} : \mathcal{F}_D \times \mathcal{F}_D \rightarrow \mathcal{F}_D$ to be the concatenation product with roots arranged according to the ordering ω . The Grossman-Larson product $\diamond_{\omega} : \mathcal{F}_D \times \mathcal{F}_D \rightarrow \mathcal{F}_D$ is defined as follows,

$$\pi \diamond_{\omega} \eta := \pi_{(1)} \cdot_{\omega} (\pi_{(2)} \curvearrowright \eta),$$

where the ordering ω is defined over the roots of $\pi \cdot \eta$. For example

$$\begin{array}{c} \bullet \\ | \\ \mathfrak{a} \\ \bullet \\ | \\ \mathfrak{b} \end{array} \cdot_{\omega} \begin{array}{c} \bullet \\ | \\ \mathfrak{c} \\ \bullet \\ | \\ \mathfrak{d} \end{array} = \begin{array}{c} \bullet \\ | \\ \mathfrak{c} \\ \bullet \\ | \\ \mathfrak{a} \end{array} \begin{array}{c} \bullet \\ | \\ \mathfrak{b} \\ \bullet \\ | \\ \mathfrak{d} \end{array}, \quad \text{for } \omega : c \leq a \leq b \leq d,$$

$$\begin{aligned}
\bullet \cdot \diamond_{\omega} \mathcal{V} \cdot \mathcal{I} &= \mathcal{V} \cdot \mathcal{I} \cdot \mathcal{I} + (\mathcal{I} \curvearrowright \mathcal{V}) \cdot \mathcal{I} + \mathcal{V} \cdot (\mathcal{I} \curvearrowright \mathcal{I}) \\
&+ (\bullet \curvearrowright \mathcal{V}) \cdot \mathcal{I} + \mathcal{V} \cdot \mathcal{I} (\bullet \curvearrowright \mathcal{I}) + \bullet \cdot \mathcal{I} \curvearrowright \mathcal{V} \cdot \mathcal{I}
\end{aligned}$$

We note that $\pi \cdot_{\omega} \eta = \pi \cdot \eta$ and $\pi \diamond_{\omega} \eta = \pi \diamond \eta$ for ω which takes the roots of π to be smaller than the roots of η . If the ordering ω in $\pi \cdot_{\omega} \eta$ does not specify the order between some roots in $\pi \cdot \eta$, then, the roots of π are taken to be smaller than the roots of η . The generalised concatenation and Grossman-Larson products are used in the context of planar exotic forests to describe the structure of the Grossman-Larson Hopf algebra.

4.3 Exotic planar forests

The algebraic structures described in this section are applied in Section 3 for the description of the composition law and the order conditions of stochastic frozen flow methods.

We construct the space \mathcal{EF} from the space of decorated planar forest \mathcal{F}_D and inherit a number of algebraic structures from it. There are two possible approaches: with an appropriate subspace or with a quotient. We use the first approach to define the algebras and the second approach to define coalgebras.

Let us consider decorated forests \mathcal{F}_D with $D = \{\bullet, \textcircled{d} : d \in \mathbb{N}\}$ and let \mathcal{F}_D^L denote its quotient by the subspace \mathcal{K} spanned by the forests in which a numbered node is found at a position which is not a leaf. It can be checked that \mathcal{K} is an ideal with respect to concatenation, shuffle, grafting, and Grossman-Larson products, and, therefore, \mathcal{F}_D^L forms an algebra when endowed with one of these products. Let $\overline{\mathcal{F}}_D^L$ denote its completion with respect to the grading given by the number of nodes, that is, $\overline{\mathcal{F}}_D^L$ is a space of formal sums over F_D^L . We define the subspace $\mathcal{F}_D^{\mathcal{E}}$ of $\overline{\mathcal{F}}_D^L$, corresponding to exotic forests, spanned by

$$\varphi(\pi, \alpha_e) := \sum_{\alpha \in P(\alpha_e)} (\pi, \alpha),$$

with $\varphi : \mathcal{EF} \rightarrow \mathcal{F}_D^{\mathcal{E}}$ an isomorphism. $P(\alpha_e)$ is the set of decorations α with $\alpha^{-1}(\bullet) = \alpha_e^{-1}(\bullet)$ and

$$\alpha(v_1) = \alpha(v_2) \quad \text{if} \quad \alpha_e(v_1) = \alpha_e(v_2), \quad \text{for } v_1, v_2 \in V(\pi).$$

Some examples of the values of φ are

$$\varphi(\textcircled{1}\textcircled{1}) = \sum_{i=1}^{\infty} \textcircled{i}\textcircled{i}, \quad \varphi(\begin{array}{c} \textcircled{2} \textcircled{2} \\ \diagup \quad \diagdown \\ \textcircled{1} \quad \textcircled{1} \end{array}) = \sum_{i,j=1}^{\infty} \begin{array}{c} \textcircled{i} \textcircled{j} \\ \diagup \quad \diagdown \\ \textcircled{i} \quad \textcircled{j} \end{array}.$$

We take $\varphi : \mathcal{EF} \rightarrow \mathcal{F}_D^{\mathcal{E}}$ to be an algebra homomorphism with respect to the generalised concatenation, shuffle, grafting, and generalised Grossman-Larson products which induces the corresponding algebraic structures over \mathcal{EF} . For instance, the shuffle product $\sqcup : \mathcal{EF} \times \mathcal{EF} \rightarrow \mathcal{EF}$ permutes the roots of forests. Let $\pi \cdot_{\omega} \eta$ be the generalised concatenation obtained by concatenating the forests π and η with the roots permuted according to the ordering ω of the roots of $\pi \cdot \eta$. The shuffle product is defined as

$$\pi \sqcup \eta = \sum_{\omega} \pi \cdot_{\omega} \eta,$$

Theorem 4.4. *The space of exotic forests \mathcal{EF} together with Grossman-Larson product and deshuffle coproduct forms a Hopf algebra.*

Proof. We note that the algebra structure is obtained by considering a subalgebra of $(\mathcal{F}_D^L, \diamond)$, while the coalgebra structure is obtained by taking a quotient of the coalgebra $(\mathcal{F}_D^L, \Delta)$. This means that the algebra and coalgebra structures are not automatically compatible. We prove the compatibility condition of the generalised Grossman-Larson product for any ordering ω ,

$$\Delta(\pi \diamond_\omega \eta) = \Delta(\pi) \diamond_\omega \Delta(\eta),$$

by induction on the number of connected components in π . Assume the compatibility is proven for all forests with the number of components less or equal to the number of components in π . Consider $\tilde{\pi} \cdot_\omega \pi$ with $\tilde{\pi} \in EF_C$ and ordering ω of roots of $\tilde{\pi} \cdot \pi \cdot \eta$ with an arbitrary order between the roots of $\tilde{\pi} \cdot \pi$ and η , then,

$$\begin{aligned} \Delta((\tilde{\pi} \cdot_\omega \pi) \diamond_\omega \eta) &= \Delta(\tilde{\pi} \diamond_\omega (\pi \diamond_\omega \eta) - (\tilde{\pi} \tilde{\curvearrowright}_\omega \pi) \diamond_\omega \eta) \\ &= \Delta(\tilde{\pi}) \diamond_\omega \Delta(\pi) \diamond_\omega \Delta(\eta) - \Delta(\tilde{\pi} \tilde{\curvearrowright}_\omega \pi) \diamond_\omega \Delta(\eta) \\ &= \Delta(\tilde{\pi} \diamond_\omega \pi - \tilde{\pi} \tilde{\curvearrowright}_\omega \pi) \diamond_\omega \Delta(\eta) = \Delta(\tilde{\pi} \cdot_\omega \pi) \diamond_\omega \Delta(\eta), \end{aligned}$$

where we use the associativity of \diamond_ω and coassociativity of Δ . It remains to show that

$$\Delta(\tilde{\pi} \diamond_\omega \eta) = \Delta(\tilde{\pi}) \diamond_\omega \Delta(\eta), \quad \text{for } \tilde{\pi} \in \text{Conn}, \eta \in \mathcal{EF}.$$

We note that $\tilde{\pi}$ has multiple roots which, when grafted onto different connected components of η , connect them into a single component. Therefore, we have

$$\Delta_\square(\tilde{\pi} \diamond_\omega \eta) = (\tilde{\pi} \diamond_\omega \eta_{(1)}) \otimes \eta_{(2)} + \eta_{(1)} \otimes (\tilde{\pi} \diamond_\omega \eta_{(2)}) = \Delta_\square(\tilde{\pi}) \diamond_\omega \Delta_\square(\eta).$$

This proves the compatibility between the Grossman-Larson product and deshuffle coproduct. Therefore, \mathcal{EF} forms a graded connected bialgebra, that is, a Hopf algebra. \square

Remark 4.5. *The space $(\mathcal{EF}, \cdot, \Delta_\square, \curvearrowright)$ is not a post-Hopf algebra, as defined in [62], since in general*

$$\pi \curvearrowright (\pi_1 \cdot \pi_2) \neq (\pi_{(1)} \curvearrowright \pi_1) \cdot (\pi_{(2)} \curvearrowright \pi_2) \text{ in } \mathcal{EF}.$$

Moreover, $(\mathcal{EF}, \cdot, \Delta_\square, \tilde{\curvearrowright})$ is not a post-Hopf algebra for the same reason. This is due to the fact that a connected planar exotic forest can have multiple roots. However, we see that $(\mathcal{EF}, \cdot, \curvearrowright)$ is a D -subalgebra of $(\mathcal{F}_D, \cdot, \curvearrowright)$, as defined in [79], with the grading of the associative algebra (\mathcal{EF}, \cdot) given by the number of roots. A straightforward extension of the arguments in [80] fails to show that the spaces $(\mathcal{EF}, \curvearrowright)$ and $(\mathcal{EF}, \tilde{\curvearrowright})$ are either brace or symmetric brace algebras for the same reason.

The Connes-Kreimer Hopf algebra structure is the dual of the Grossman-Larson one in a Euclidean setting. On homogeneous spaces, it generalises into the Munthe-Kaas-Wright Hopf algebra [79, 36].

Let an admissible cut c of an exotic planar tree τ be a subset of edges of τ such that it contains at most one edge from each path in τ from a leaf to the root. If $e = (u, v)$ is in c , then all incoming edges of v which are to the left of e are also in c . Moreover, given a liana with nodes v_1 and v_2 , if any edge in the path from v_1 to the root is in c , then the path from v_2 to the root must also contain an edge in c .

Given an admissible cut c , $\tau \setminus c$ is an exotic planar forest. Let $R^c(\tau)$ be the component of $\tau \setminus c$ which contains the root of τ . Let $P^c(\tau)$ be the linear combination of exotic planar forests obtained by taking the components of $\tau \setminus c$ that do not contain the root of τ and shuffling the roots while preserving the order of the roots cut off from the same node.

Theorem 4.6. *Let the Munthe-Kaas-Wright coproduct on \mathcal{EF} be, for $\tau \in ET$ and $\pi \in EF$, defined as*

$$\begin{aligned}\Delta_{MKW}(\tau) &:= \sum_{\text{adm. cut } c} P^c(\tau) \otimes R^c(\tau), \\ \Delta_{MKW}(\pi) &:= (id \otimes B^-) \Delta_{MKW}(B^+(\pi)),\end{aligned}$$

where the sum is over admissible cuts c . Then, $(\mathcal{EF}, \sqcup, \Delta_{MKW})$ is a Hopf algebra dual to the planar Grossman-Larson Hopf algebra $(\mathcal{EF}, \diamond, \Delta_{\sqcup})$. Its convolution product,

$$a * b = (a \otimes b) \circ \Delta_{MKW},$$

is called the composition law.

Proof. We recall that the dual of a Hopf algebra is a Hopf algebra. It is clear that \sqcup and Δ_{\sqcup} are dual. It remains to show that the Munthe-Kaas-Wright coproduct Δ_{MKW} is dual to the Grossman-Larson product \diamond . We recall that Munthe-Kaas-Wright coproduct and Grossman-Larson product are dual over the planar decorated forests. The Munthe-Kaas-Wright coproduct over exotic forests is obtained by taking the quotient over \mathcal{I} defined in (4.1) for the construction of deshuffle coproduct. This is the dual approach to the way the Grossman-Larson product over exotic forests is obtained. \square

Some examples of the values of Δ_{MKW} are

$$\begin{aligned}\Delta_{MKW}(\textcircled{1}\textcircled{1}) &= \textcircled{1}\textcircled{1} \otimes \mathbf{1} + \mathbf{1} \otimes \textcircled{1}\textcircled{1}, \\ \Delta_{MKW}(\textcircled{1}\textcircled{2}) &= \textcircled{1}\textcircled{2} \otimes \mathbf{1} + \textcircled{1}\textcircled{1} \otimes \bullet + \mathbf{1} \otimes \textcircled{1}\textcircled{2}, \\ \Delta_{MKW}(\textcircled{1}\textcircled{3}) &= \textcircled{1}\textcircled{3} \otimes \mathbf{1} + 2\textcircled{1}\textcircled{1} \otimes \bullet \bullet + (\textcircled{1}\textcircled{2} + \textcircled{2}\textcircled{1}) \otimes \bullet \bullet \bullet + (\textcircled{1}\textcircled{1} + \textcircled{1}\textcircled{1}) \otimes \bullet \bullet \bullet \\ &\quad + (\textcircled{1}\textcircled{2} + \textcircled{2}\textcircled{1}) \otimes \bullet \bullet + 2\textcircled{1}\textcircled{1} \otimes \bullet \bullet + \mathbf{1} \otimes \textcircled{1}\textcircled{3}.\end{aligned}$$

The Hopf algebra structures on \mathcal{EF} translate to the structures on $T(\mathfrak{X}(\mathcal{M}))$ via the use of the elementary differential map \mathbb{F} defined in Section 3.3. We recall that decorated planar forest form a free D-algebra with the concatenation and grafting products [79] which is generated by the set of decorations D . The elementary differential map over decorated planar forests $\mathbb{F}^{\text{dec}} : \mathcal{F}_D \rightarrow T(\mathfrak{X}(\mathcal{M}))$ is a D-algebra homomorphism that is fully characterized by its values on the nodes decorated by $D = \{\bullet, \textcircled{l} : 1 \leq l \leq L\}$,

$$\mathbb{F}^{\text{dec}}(\bullet) = f, \quad \mathbb{F}^{\text{dec}}(\textcircled{l}) = \xi^{d,l} E_d,$$

where the $\xi^{d,l}$ are independent standard Gaussian random variables. The map \mathbb{F}^{dec} thus defined is equivalent to the map \mathbb{F}^{dec} from Definition 3.3.

The space of exotic forests \mathcal{EF} is isomorphic to a D-subalgebra $\mathcal{F}_D^{\mathcal{E}}$ of the D-algebra of decorated forests with $D = \{\bullet, \textcircled{d} : d \in \mathbb{N}\}$ through the isomorphism $\varphi : \mathcal{EF} \rightarrow \mathcal{F}_D^{\mathcal{E}}$ as is described in the beginning of Section 4.3, and, thus, inherits the elementary differential homomorphism as $\mathbb{F} := \tilde{\mathbb{F}}^{\text{dec}} \circ \varphi$ with $\tilde{\mathbb{F}}^{\text{dec}} : \mathcal{F}_D^{\mathcal{E}} \rightarrow T(\mathcal{X}(\mathcal{M}))$ defined as

$$\tilde{\mathbb{F}}^{\text{dec}}(\bullet) = f, \quad \tilde{\mathbb{F}}^{\text{dec}}(\textcircled{d}) = E_d.$$

This definition is equivalent to the one given in Definition 3.8.

Proposition 4.7. *The elementary differential $\mathbb{F}: \mathcal{E}\mathcal{F} \rightarrow T(\mathfrak{X}(\mathcal{M}))$ is a Hopf algebra morphism:*

$$\mathbb{F}(\pi_1 \curvearrowright \pi_2) = \mathbb{F}(\pi_1) \triangleright \mathbb{F}(\pi_2), \quad \mathbb{F}(\pi_1 \cdot \pi_2) = \mathbb{F}(\pi_1) \cdot \mathbb{F}(\pi_2), \quad \mathbb{F}(\pi_1 \diamond \pi_2) = \mathbb{F}(\pi_1) * \mathbb{F}(\pi_2).$$

In particular, the frozen composition is represented by concatenation and the composition of differential operators is represented by the Grossman-Larson product:

$$\begin{aligned} (\mathbb{F}(\pi_1) \cdot \mathbb{F}(\pi_2)) \triangleright \phi(p) &= \mathbb{F}_p(\pi_1) \triangleright (\mathbb{F}_p(\pi_2) \triangleright \phi)(p) = \mathbb{F}(\pi_1 \cdot \pi_2) \triangleright \phi(p), \\ (\mathbb{F}(\pi_1) * \mathbb{F}(\pi_2)) \triangleright \phi(p) &= \mathbb{F}(\pi_1) \triangleright (\mathbb{F}(\pi_2) \triangleright \phi)(p) = \mathbb{F}(\pi_1 \diamond \pi_2) \triangleright \phi(p). \end{aligned}$$

5 Numerical experiments

We provide several numerical experiments to confirm the theoretical findings and the numerical properties of the new schemes. In particular, we present a practical guide for the implementation of frozen flow methods, convergence curves for the weak error on the orthogonal group, a comparison with the sampling geometric methods from the literature on the sphere, and an ergodicity test for the simulation of generalised Cauchy measures in a context where the assumptions of our convergence analysis are not satisfied.

5.1 Practical implementation of frozen flow methods

The implementation of frozen flow methods relies on the ability to find local frames and to compute the solution of the frozen equation

$$y'(t) = \sum_d \alpha_d E_d(y(t)), \quad y(0) = y_0.$$

On homogeneous manifolds, we recall that global frames can be derived from any basis (A_d) of the Lie algebra \mathfrak{g} with the derivative of Lie-group action: $E_d(p) = A_d \cdot p$. On matrix homogeneous manifolds, the solution of the frozen flow equation coincides with the matrix exponential:

$$y(t) = \text{Exp}\left(\sum_d \alpha_d A_d\right) \cdot y_0.$$

For sampling the invariant measure of Riemannian Langevin dynamics (1.3), the algorithms require at each step a local orthonormal frame basis and the solution of the frozen flow equation. While one can change the orthonormal local frame between different steps, it is crucial to rely only on one fixed frame per step. On Lie groups, choosing an orthonormal basis (A_d) of the Lie algebra \mathfrak{g} for the definition of the E_d naturally yields an orthonormal frame. On general Riemannian manifolds, one can provide an orthonormal basis of $T_{X_0}\mathcal{M}$ and parallel transport it around X_0 to obtain a local orthonormal frame basis.

In general, solving the frozen flow equation is not straightforward. It is equally difficult to solving the geodesic equation, as done in [10]. Similar to the idea of retractions (see, for instance, the textbook [4]), the frozen flow could be replaced by any approximation with sufficiently high order of accuracy and that lies on the manifold. A standard example of such an approximation is the Cayley map [47]. In our numerical experiments, we rely on exact expressions of the frozen flows using either Lie geometry or coordinates for the sake of simplicity.

We emphasize that the geometric operations needed for stochastic Lie-group methods are available on a wide collection on manifolds and in a variety of programming language (see, for instance, [37, 76, 7]).

5.2 Brownian dynamics on the special orthogonal group

Let $\mathcal{M} = \text{SO}_p(\mathbb{R})$ be the compact Lie group of special orthogonal matrices of size $p \times p$. Let (A_d) be an orthonormal basis of its Lie algebra \mathfrak{so}_d of antisymmetric matrices and for the bi-invariant metric $g(A, B) = \text{Tr}(A^T B)$. Let $E_d(y) = A_d y$ be the associated right-invariant orthonormal frame basis. As the manifold is smooth and compact, the geometric assumptions are all satisfied automatically and our analysis applies.

Following Example 2.15, Brownian dynamics on $\text{SO}_p(\mathbb{R})$ are given by

$$dX(t) = \sum_{d=1}^D E_d(X(t)) \circ dW_d(t), \quad D = \dim(\mathfrak{so}_p) = \frac{p(p-1)}{2}.$$

In this context, the implementation of the new frozen flow methods is very natural as it relies only on fast linear algebra operations. In particular, the new second order method (2.11) becomes (under a time rescaling)

$$X_{n+1} = \text{Exp} \left(\sum_{d=1}^D \left(\left(\frac{\sqrt{2}}{2} - 1 \right) \sqrt{h} \xi_n^{d,1} + \frac{\sqrt{2}}{2} \sqrt{h} \xi_n^{d,2} \right) A_d \right) \text{Exp} \left(\sum_{d=1}^D \sqrt{h} \xi_n^{d,1} A_d \right) X_n.$$

Note that there is no timestep restriction for the scheme to be well-posed. This is in striking contrast to projection methods, which would rely on a tedious non-geometric implementation, on the embedding of $\text{SO}_p(\mathbb{R})$ into $\mathbb{R}^{p \times p}$ (of more than double dimension), and would require a small enough timestep for the projection map to be well-defined, even in this simple context of a compact manifold.

We consider the weak approximation of Brownian dynamics on $\text{SO}_3(\mathbb{R})$ with the Euler frozen flow method (2.8) and the new second order method (2.11) (with discrete bounded random variables with correct first moments - see Section 2.3) for the test function

$$\phi(x) = \exp\left(-\frac{\text{Tr}((x - I_d)^T (x - I_d))}{2d}\right),$$

the initial condition $X_0 = I_d$, final time $T = 1$, and different timesteps to observe the order of convergence. We emphasize that the explicit expression of Brownian dynamics $X(t)$ on $\text{SO}_3(\mathbb{R})$ are not given by $\text{Exp}(\sum A_d W_d(t))X_0$, but rather by a stochastic Magnus expansion (see, for instance, [19, 18, 47, 92, 50]):

$$X(t) = \text{Exp}(\Omega(t))X_0, \quad d\Omega(t) = d \text{Exp}_{\Omega(t)}^{-1} \left(\sum_d A_d \circ dW_d(t) \right), \quad \Omega(0) = 0,$$

whose first terms are

$$\begin{aligned} \Omega(t) &= \sum_{d_1} A_{d_1} W_{d_1}(t) - \frac{1}{2} \int_0^t \sum_{d_1, d_2} [A_{d_2}, A_{d_1}] W_{d_2}(s) \circ dW_{d_1}(s) \\ &+ \frac{1}{12} \int_0^t \sum_{d_1, d_2, d_3} [A_{d_3}, [A_{d_2}, A_{d_1}]] W_{d_3}(s) W_{d_2}(s) \circ dW_{d_1}(s) \\ &+ \frac{1}{4} \int_0^t \int_0^s \sum_{d_1, d_2, d_3} [[A_{d_3}, A_{d_2}], A_{d_1}] W_{d_3}(u) \circ dW_{d_2}(u) \circ dW_{d_1}(s) + \dots \end{aligned}$$

For simplicity, the reference solution for $\mathbb{E}[\phi(X(T))]$ is chosen as the output of the new second order scheme with the timestep $h_{\text{ref}} = 2^{-12}$. We observe on Figure 1 the correct order of convergence of the two new methods. The second order integrator reaches the Monte-Carlo error threshold almost instantly, even for relatively large timestep.

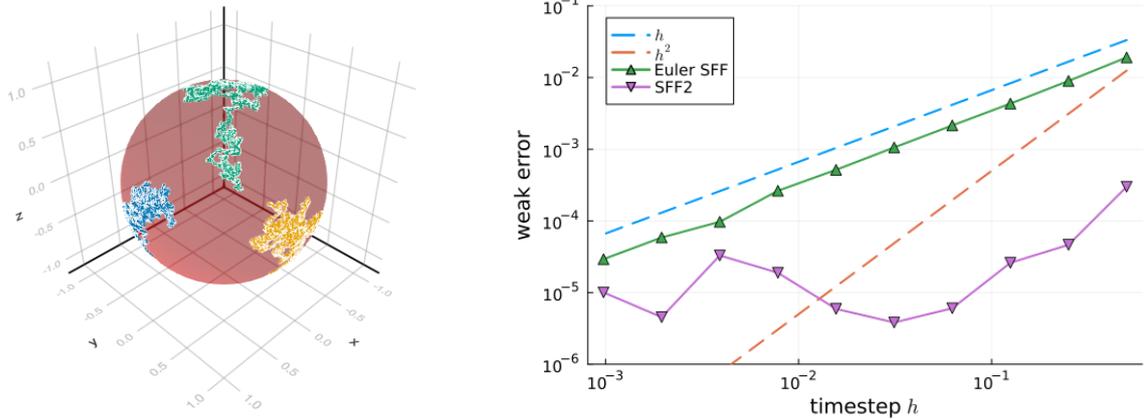


Figure 1: Brownian excursion at time $T = 1$ on $\text{SO}_3(\mathbb{R})$ represented by the trajectories of the three columns on the sphere \mathbb{S}^2 (left) and error curve for the weak approximation of Brownian dynamics on $\text{SO}_3(\mathbb{R})$ with $\phi(x) = \exp(-\text{Tr}((x - I_d)^T(x - I_d))/2d)$, $T = 1$ and $M = 10^8$ trajectories (right).

5.3 Ergodic dynamics on the sphere

Our second test case focuses on the sampling of Riemannian Langevin dynamics on the sphere \mathbb{S}^2 , following the experiment of [10]. While the approximation for the invariant measure is not the focus of the new methods implemented in this paper, this example allows us to conveniently compare the performance of the new frozen flow methods to the alternatives from the literature. The sphere \mathbb{S}^d can be seen as a homogeneous manifold with respect to the Lie group $G = \text{SO}_{d+1}(\mathbb{R})$, so that we could consider a natural frame derived from a basis of the Lie algebra \mathfrak{so}_{d+1} (which would typically fit the simulation of a stochastic rigid body). However, we need an orthonormal frame to sample from the Riemannian Langevin dynamics (1.3), so that we use a different choice of frame relying on the standard spherical coordinates. We consider the coordinates on \mathbb{S}^2 minus the north/south poles:

$$y_{\theta,\varphi} = \begin{pmatrix} \cos(\theta) \cos(\varphi) \\ \cos(\theta) \sin(\varphi) \\ \sin(\theta) \end{pmatrix} \in \mathbb{S}^2, \quad \theta \in]-\frac{\pi}{2}, \frac{\pi}{2}[, \quad \varphi \in [0, 2\pi[,$$

with the associated orthonormal frame basis:

$$E_1(y_{\theta,\varphi}) = \begin{pmatrix} -\sin(\theta) \cos(\varphi) \\ -\sin(\theta) \sin(\varphi) \\ \cos(\theta) \end{pmatrix}, \quad E_2(y_{\theta,\varphi}) = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \\ 0 \end{pmatrix},$$

and a second frame associated to the spherical coordinates on \mathbb{S}^2 minus $[\pm 1, 0, 0]$. Using two frames guarantees a bounded Lipschitz constant for the chosen frame and does not impact the sampled measure. As the manifold is compact, the assumptions of our analysis are satisfied automatically. The frozen flow is computed explicitly in coordinates. In order to obtain a fair comparison between the methods, we use Gaussian random variables.

We now compare the numerical behaviour of the frozen flow Euler method (2.8), the new second order method (2.10), the Riemannian Langevin method [10], and the extrinsic projection methods of order one [60] for sampling the invariant measure of Riemannian Langevin dynamics (1.3) on \mathbb{S}^2 . We consider the potential $V(y_{\theta,\varphi}) = -\sin(\theta)$, the associated vector field (containing the Ito correction)

$$F = -\nabla V - \nabla_{E_1} E_1 - \nabla_{E_2} E_2,$$

and the test function $\phi(x) = x_3^2$. In this context, the generator is $\mathcal{L}\phi = -\phi'(\nabla V) + \frac{1}{2}\Delta_{\mathbb{S}^2}$, with the spherical Laplacian $\Delta_{\mathbb{S}^2}$. Assuming ergodicity (which is observed numerically for the methods studied here), the numerical schemes approximate the quantity

$$\frac{\int_{\mathbb{S}^2} \phi(x) e^{-V(x)} d \text{vol}(x)}{\int_{\mathbb{S}^2} e^{-V(x)} d \text{vol}(x)} = 3 - \frac{2}{\tanh(1)}.$$

We observe the weak error curves in Figure 2, as well as the expected orders of convergence. The new second-order frozen flow method outperforms the other integrators and reaches the Monte-Carlo threshold with much larger timestep.

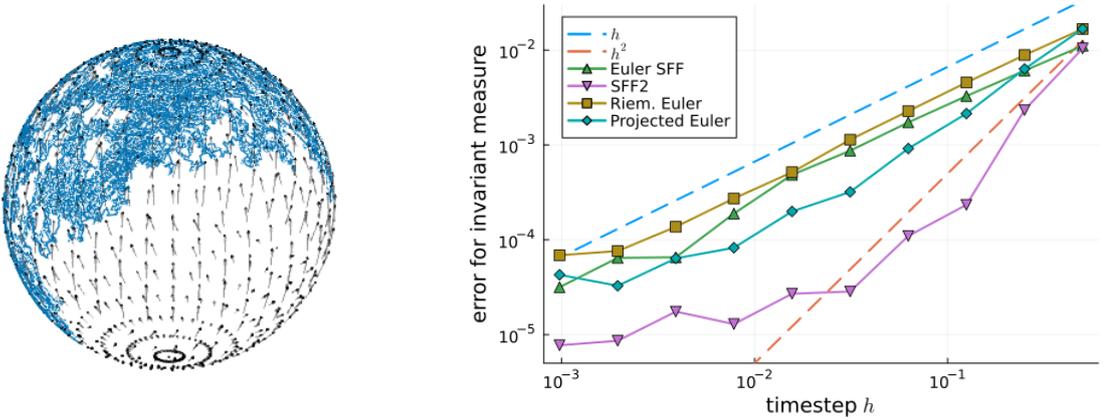


Figure 2: A trajectory of the new second order method (2.11) (left) and the convergence curve for the weak approximation of stochastic rigid body equation on the sphere (right) with $\phi(x) = x_3^2$, $T = 10$ and $M = 10^8$ trajectories.

5.4 Generalised Cauchy measures

The generalised Cauchy measures are probability measures μ_β defined on \mathbb{R}^n by

$$\mu_\beta(dx) = (1 + |x|^2)^{-\beta} \frac{dx}{Z_\beta}, \quad Z_\beta = \frac{\Gamma(\beta - n/2)}{\Gamma(\beta)},$$

with $\beta > n/2$. They can be interpreted as finite dimensional approximations of Gaussian distributions, as in [11], and are related to the fast diffusion equation for which they play the role of the heat kernel and are known as Barenblatt' solutions (see [90]). A generalised Cauchy measure is the ergodic distribution of the dynamics

$$dX_t = -2(\beta - 1)X_t dt + \sqrt{2}(1 + |X_t|^2)dB(t), \quad (5.1)$$

where $B(t)$ is a 2-dimensional Brownian motion. A direct numerical simulation of this SDE presents numerous difficulties as the multiplicative non-Lipschitz noise brings tedious stability problems, so that most standard Euclidean numerical methods fail. However, the noise can be interpreted as additive at the cost of a change of metric $g_x = (1 + |x|^2)^{-1}\langle \cdot, \cdot \rangle$. This change is not simply a transport of difficulties from noise to geometry, as it reinterprets the problem in a convenient space. The same approach has been fruitful in the study of the associated spectral gap, i.e. the speed of convergence of the semigroup to the ergodic measure, in [45]. This context

calls for methods on manifolds, as Euclidean methods are not adapted. We emphasize that the process $(X_t)_t$ does not have bounded moments at all orders. Thus, the convergence results of the standard Euclidean analysis, and of our analysis as well, do not hold. This makes the simulation of generalised Cauchy measures a convenient toy problem for testing our methods.

We are concerned with the simulation of generalised Cauchy measure on \mathbb{R}^2 . As explained in [44], the adapted manifold can be interpreted as the revolution surface

$$\mathcal{S} = \{(\tanh(r) \cos(\theta), \tanh(r) \sin(\theta), h(r)); (r, \theta) \in \mathbb{R}_+ \times \mathbb{S}^1\}, \quad (5.2)$$

with $h(r) = \operatorname{argsinh}(\cosh(r)) - \cosh^{-2}(r) \sqrt{\cosh^2(r) + 1}$. However, it is more handy to see it through a global map, as $\mathbb{R}_+ \times \mathbb{S}^1$, with coordinate (r, θ) , endowed with the metric

$$g = \begin{pmatrix} 1 & 0 \\ 0 & \tanh^2(r) \end{pmatrix}$$

in the basis $(\partial_r, \partial_\theta)$. This map comes from fitting polar coordinates on \mathbb{R}^2 ,

$$(x, y) = (\sinh(r) \cos(\theta), \sinh(r) \sin(\theta)),$$

which trivialise the metric g . We define the following orthonormal frame basis on $\mathbb{R}_+^* \times \mathbb{S}^1$

$$E_1(r, \theta) = \partial_r, \quad E_2(r, \theta) = \frac{1}{\tanh(r)} \partial_\theta.$$

In this manifold, generalised Cauchy measures correspond to Gibbs measures associated to the potential $V_\beta(r, \theta) = 2(\beta - 1) \log(\cosh(r))$ and the generator

$$\mathcal{L} = E_1^2 + E_2^2 + \left(\frac{1}{\tanh(r)} - (2\beta - 1) \tanh(r) \right) E_1.$$

Note that the frozen flow is explicit on \mathcal{S} . For all $\alpha, \beta \in \mathbb{R}$ and $(r_0, \theta_0) \in \mathbb{R}_+^* \times \mathbb{S}^1$, we find

$$\exp(t(\alpha E_1 + \beta E_2)) \cdot (r_0, \theta_0) = \left(r_0 + \alpha t, \theta_0 + \frac{\beta}{\alpha} \log \left(\frac{\sinh(r_0 + \alpha t)}{\sinh(r_0)} \right) \right),$$

for all $t \geq 0$ such that r_t stays positive. This brings an additional difficulty in this experiment. Although the potential f^1 blows up at the origin, the coefficient

$$\alpha = h f^1(r) + \sqrt{2h} \xi$$

with ξ a Gaussian random variable, can be negative and the frozen flow not defined. The use of bounded random variable is instrumental in avoiding this instability. In particular, the method (2.8) with Rademacher random variables is well-posed.

Our numerical experiment aims to illustrate the ergodicity of the simplest frozen flow method (2.8), which, to the best of our knowledge, has not yet been established in the literature. We use the first eigenfunctions of the generator $\phi^1(r, \theta) = \sinh(r)^2 - (\beta - 2)^{-1}$, $\phi^2(r, \theta) = \sinh(r) \cos(\theta)$, which both satisfy $\int \phi^k d\mu_\beta = 0$. As a consequence of the Poincaré inequality, the variance of the semigroup associated to \mathcal{L} converges at exponential speed to 0. The coefficient of the exponential is given by the spectral gap (see [46] for the spectral gap expression). The test functions ϕ^k are extremal for the Poincaré inequality. Hence, for these particular test functions, the exponential convergence can be observed not only for the variance but also pointwise. In Figure 3, we

illustrate numerically the exponential convergence to equilibrium of the numerical methods, in the spirit of [33] in the Euclidean setting. In particular, we observe a convergence in $Ce^{-4(\beta-2)t}$ and $C'e^{-2(\beta-1)t}$ for our specific choice of test functions, confirming numerically the results in [46]. For the first eigenfunction ϕ^1 , the error for the invariant measure reaches the discretisation bias with exponential speed. We observe that the bias threshold evolves linearly in h (up to Monte-Carlo error), as expected for a method of order one. For ϕ^2 , we do not observe the bias of the Euler frozen flow method (2.8), hinting that for this specific test function, the method might be of higher order of accuracy for the invariant measure. For the sake of comparison, we mention that the standard Euclidean integrators applied to (5.1) face severe instability issues and thus fail to converge, let alone be ergodic.

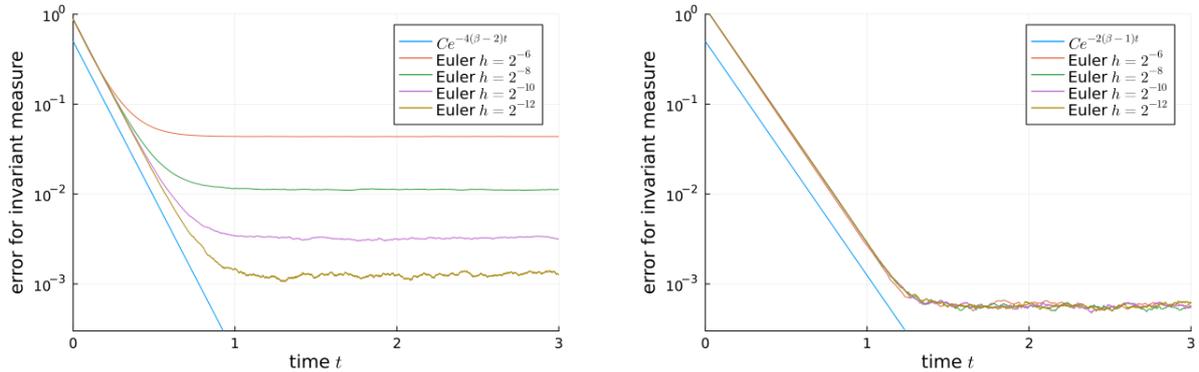


Figure 3: Evolution in time of the error for the invariant measure of the Euler frozen flow method (2.8) for the simulation of generalised Cauchy measures on the manifold (5.2) with $\beta = 4$, $M = 10^8$ trajectories, and the test functions $\phi^1(r, \theta) = \sinh(r)^2 - (\beta - 2)^{-1}$ (left) and $\phi^2(r, \theta) = \sinh(r) \cos(\theta)$ (right).

6 Conclusion

In this paper, we provided a brand-new class of intrinsic discretisations for the approximation of stochastic dynamics on Riemannian manifolds with high weak order of accuracy. We presented a new robust convergence analysis, that naturally generalises the Euclidean analysis, and the new algebraic formalism of planar exotic trees, forests, and series for the study of the order conditions. The analysis applies in particular for the creation of high-order sampling methods for Riemannian Langevin dynamics. The new methods were tested on a variety of different manifolds, where we observe that the new second order frozen flow method outperforms the other approaches from the literature in terms of accuracy.

The tools developed in the present paper open several avenues for future research. Similar to the Euclidean setting, we will extend the analysis to general SDEs with multiplicative noise and to the creation of integrators with high-order of accuracy for sampling the invariant measure, and low weak order (in the spirit of [57, 91]). This requires generalisations of the exotic forests formalism [20], the challenging extension of the integration by parts of trees in the planar case [55, 56, 14, 15], and the extension of stochastic backward error analysis for the invariant measure [86, 93, 33, 2, 15] on manifolds. While the focus of this paper is on discretisations and improvements in accuracy only, our approach could naturally be combined with popular sampling techniques, such as MLMC [38, 39], perturbations [59, 34, 1],... Then, following the work [10], one could wonder whether a purely Riemannian approach of algebraic order theory could be used. Our approach follows the one of deterministic Lie-group methods and thus relies

on a curvature-free connection, unrelated to the natural Levi-Civita connection on Riemannian manifolds. The study of such new methods is already open in the deterministic setting and relies on the challenging general understanding of the connection algebra [5]. Finally, the new algebraic formalism of planar exotic series is interesting in itself and could be studied for its universal combinatorial, algebraic, and geometric properties, but also for its potential applications in different fields, in the spirit of the use of Butcher series and their extensions in rough paths [42, 58], renormalisation theory [17], variational calculus [53, 52] or approximation of PDEs [16]. These projects will be studied in upcoming works.

Acknowledgements. The authors would like to thank R. Bergmann, E. Celledoni, H. Munthe-Kaas, and B. Owren for helpful discussions and advice on manipulating the package Manifolds.jl [7]. The authors acknowledge the support of the Swiss National Science Foundation, projects No 200020_214819, and No. 200020_192129, of the Research Council of Norway through project 302831 “Computational Dynamics and Stochastics on Manifolds” (CODYSMA), and of the French program ANR-11-LABX-0020-0 (Labex Lebesgue). The computations were performed on the Baobab cluster of the University of Geneva.

References

- [1] 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.
- [2] 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.
- [3] M. Ableidinger and E. Buckwar. Weak stochastic Runge–Kutta Munthe-Kaas methods for finite spin ensembles. *Applied Numerical Mathematics*, 118:50–63, 2017.
- [4] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, Princeton, NJ, 2008. With a foreword by Paul Van Dooren.
- [5] M. J. H. Al-Kaabi, K. Ebrahimi-Fard, D. Manchon, and H. Z. Munthe-Kaas. Algebraic aspects of connections: from torsion, curvature, and post-lie algebras to Gavrilov’s double exponential and special polynomials. *Submitted*, 2022.
- [6] A. Antonyuk and A. Antonyuk. Nonexplosion and solvability of nonlinear diffusion equations on noncompact manifolds. *Ukr. Math. J.*, 59:1632–1652, 2007.
- [7] S. D. Axen, M. Baran, R. Bergmann, and K. Rzecki. Manifolds.jl: An extensible julia framework for data analysis on manifolds. *ACM Transactions on Mathematical Software*, 49(4), dec 2023.
- [8] D. Bakry. Un critère de non-explosion pour certaines diffusions sur une variété riemannienne complète. *C.R. Acad. Sc. Paris*, 303(1):23–26, 1986.
- [9] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren Math. Wiss.* Cham: Springer, 2014.
- [10] K. Bharath, A. Lewis, A. Sharma, and M. V. Tretyakov. Sampling and estimation on manifolds using the Langevin diffusion. *arXiv:2312.14882*, 2023.
- [11] S. G. Bobkov and M. Ledoux. Weighted Poincaré-type inequalities for Cauchy and other convex measures. *Ann. Probab.*, 37(2):403–427, 2009.
- [12] G. Bogfjellmo. Algebraic structure of aromatic B-series. *J. Comput. Dyn.*, 6(2):199–222, 2019.

- [13] G. Bogfjellmo, E. Celledoni, R. I. McLachlan, B. Owren, and G. R. W. Quispel. Using aromas to search for preserved measures and integrals in Kahan’s method. *Math. Comp.*, 93(348):1633–1653, 2024.
- [14] E. Bronasco. Exotic B-series and S-series: algebraic structures and order conditions for invariant measure sampling. *Found. Comput. Math.*, pages 1–31, 2024.
- [15] E. Bronasco and A. Laurent. Hopf algebra structures for the backward error analysis of ergodic stochastic differential equations. *arXiv:2407.07451*, 2024.
- [16] Y. A. Bronsard, Y. Bruned, and K. Schratz. Approximations of dispersive PDEs in the presence of low-regularity randomness. *arXiv preprint arXiv:2205.02156*, 2022.
- [17] C. Brouder. Runge–Kutta methods and renormalization. *Eur. Phys. J. C*, 12(3):521–534, 2000.
- [18] K. Burrage and P. M. Burrage. High strong order methods for non-commutative stochastic ordinary differential equation systems and the Magnus formula. *Phys. D*, 133(1-4):34–48, 1999. Predictability: quantifying uncertainty in models of complex phenomena (Los Alamos, NM, 1998).
- [19] P. M. Burrage. *Runge-Kutta methods for stochastic differential equations*. PhD thesis, Dept. Maths., Univ. Queensland, Australia, 1999.
- [20] A. Busnot Laurent, K. Debrabant, and A. Kværnø. Optimal stochastic Runge-Kutta methods for the weak integration of stochastic dynamics. *Preprint*, 2025.
- [21] J. C. Butcher. *Numerical methods for ordinary differential equations*. John Wiley & Sons, Ltd., Chichester, third edition, 2016.
- [22] J. C. Butcher. *B-series: algebraic analysis of numerical methods*. Springer, 2021.
- [23] D. Calaque, K. Ebrahimi-Fard, and D. Manchon. Two interacting Hopf algebras of trees: a Hopf-algebraic approach to composition and substitution of B-series. *Adv. in Appl. Math.*, 47(2):282–308, 2011.
- [24] P. Catoire. The Cartier-Quillen-Milnor-Moore theorem in the Post-Hopf case. *arXiv preprint arXiv:2401.09116*, 2024.
- [25] E. Celledoni, A. Marthinsen, and B. Owren. Commutator-free Lie group methods. *Future Generation Computer Systems*, 19(3):341–352, 2003.
- [26] P. Chartier, E. Hairer, and G. Vilmart. Algebraic structures of B-series. *Found. Comput. Math.*, 10(4):407–427, 2010.
- [27] P. Chartier and A. Murua. Preserving first integrals and volume forms of additively split systems. *IMA J. Numer. Anal.*, 27(2):381–405, 2007.
- [28] X. Cheng, J. Zhang, and S. Sra. Efficient sampling on Riemannian manifolds via Langevin MCMC. *Advances in Neural Information Processing Systems*, 35:5995–6006, 2022.
- [29] P. E. Crouch and R. Grossman. Numerical integration of ordinary differential equations on manifolds. *Journal of Nonlinear Science*, 3:1–33, 1993.
- [30] C. Curry and A. Schmeding. Convergence of lie group integrators. *Numerische Mathematik*, 144:357–373, 2020.
- [31] R. L. Davidchack, T. Ouldridge, and M. Tretyakov. New Langevin and gradient thermostats for rigid body dynamics. *The Journal of chemical physics*, 142(14), 2015.
- [32] K. Debrabant. Runge-Kutta methods for third order weak approximation of SDEs with multidimensional additive noise. *BIT Numer. Math.*, 50(3):541–558, 2010.
- [33] A. Debussche and E. Faou. Weak backward error analysis for SDEs. *SIAM J. Numer. Anal.*, 50(3):1735–1752, 2012.

- [34] A. B. Duncan, T. Lelièvre, and G. A. Pavliotis. Variance reduction using nonreversible Langevin samplers. *J. Stat. Phys.*, 163(3):457–491, 2016.
- [35] K. Ebrahimi-Fard, A. Lundervold, and H. Z. Munthe-Kaas. On the Lie enveloping algebra of a post-Lie algebra. *J. Lie Theory*, 25(4):1139–1165, 2015.
- [36] K. Ebrahimi-Fard and L. Rahm. A survey on the Munthe-Kaas–Wright Hopf algebra. *Journal of Computational Dynamics*, 2024.
- [37] K. Engø, A. Marthinsen, and H. Z. Munthe-Kaas. Diffman: An object-oriented matlab toolbox for solving differential equations on manifolds. *Applied numerical mathematics*, 39(3-4):323–347, 2001.
- [38] M. B. Giles. Multilevel Monte Carlo path simulation. *Oper. Res.*, 56(3):607–617, 2008.
- [39] M. B. Giles. Multilevel Monte Carlo methods. *Acta Numer.*, 24:259–328, 2015.
- [40] E. Grong, H. Z. Munthe-Kaas, and J. Stava. Post-Lie algebra structure of manifolds with constant curvature and torsion. *arXiv preprint arXiv:2305.02688*, 2023.
- [41] 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.
- [42] M. Hairer and D. Kelly. Geometric versus non-geometric rough paths. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(1):207–251, 2015.
- [43] E. P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.
- [44] B. Huguet. *Stochastic calculus on manifold and application to functional inequalities*. PhD thesis, Université de Bordeaux, 2020.
- [45] B. Huguet. Intertwining relations for diffusions in manifolds and applications to functional inequalities. *Stoch. Proc. and their App.*, 145:1–28, 2022.
- [46] B. Huguet. Poincaré inequalities and integrated curvature-dimension criterion for generalised cauchy and convex measures. *Bernoulli*, 30(3):2207–2227, 2024.
- [47] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna. Lie-group methods. In *Acta numerica, 2000*, volume 9 of *Acta Numer.*, pages 215–365. Cambridge Univ. Press, Cambridge, 2000.
- [48] A. Iserles, G. R. W. Quispel, and P. S. P. Tse. B-series methods cannot be volume-preserving. *BIT Numer. Math.*, 47(2):351–378, 2007.
- [49] L. Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12, 11 1918.
- [50] K. Kamm, S. Pagliarani, and A. Pascucci. On the stochastic Magnus expansion and its application to SPDEs. *J. Sci. Comput.*, 89(3):Paper No. 56, 31, 2021.
- [51] A. Laurent. *Algebraic Tools and Multiscale Methods for the Numerical Integration of Stochastic Evolutionary Problems*. PhD thesis, University of Geneva, 2021.
- [52] A. Laurent. The Lie derivative and Noether’s theorem on the aromatic bicomplex for the study of volume-preserving numerical integrators. *J. Comput. Dyn.*, 11(1):10–22, 2024.
- [53] A. Laurent, R. I. McLachlan, H. Z. Munthe-Kaas, and O. Verdier. The aromatic bicomplex for the description of divergence-free aromatic forms and volume-preserving integrators. *Forum Math. Sigma*, 11:Paper No. e69, 2023.
- [54] A. Laurent and H. Z. Munthe-Kaas. The universal equivariance properties of exotic aromatic B-series. *Foundations of Computational Mathematics*, pages 1–32, 2024.
- [55] 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.

- [56] A. Laurent and G. Vilmart. Order conditions for sampling the invariant measure of ergodic stochastic differential equations on manifolds. *Found. Comput. Math.*, 22(3):649–695, 2022.
- [57] B. Leimkuhler and C. Matthews. Rational construction of stochastic numerical methods for molecular sampling. *Appl. Math. Res. Express. AMRX*, 2013(1):34–56, 2013.
- [58] A. Lejay. Constructing general rough differential equations through flow approximations. *Electron. J. Probab.*, 27:Paper No. 7, 24, 2022.
- [59] 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.
- [60] T. Lelièvre, M. Rousset, and G. Stoltz. *Free energy computations*. Imperial College Press, London, 2010. A mathematical perspective.
- [61] X.-M. Li. Stochastic differential equations on noncompact manifolds: moment stability and its topological consequences. *Probab. Theory Relat. Fields*, 100:417–428, 1994.
- [62] Y. Li, Y. Sheng, and R. Tang. Post-hopf algebras, relative rota–baxter operators and solutions to the yang–baxter equation. *Journal of Noncommutative Geometry*, 18(2):605–630, 2023.
- [63] A. Lundervold and H. Munthe-Kaas. Hopf algebras of formal diffeomorphisms and numerical integration on manifolds. In *Combinatorics and physics*, volume 539 of *Contemp. Math.*, pages 295–324. Amer. Math. Soc., Providence, RI, 2011.
- [64] S. J. A. Malham and A. Wiese. Stochastic Lie group integrators. *SIAM J. Sci. Comput.*, 30(2):597–617, 2008.
- [65] G. Marjanovic, M. J. Piggott, and V. Solo. A simple approach to numerical methods for stochastic differential equations in Lie groups. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 7143–7150. IEEE, 2015.
- [66] G. Marjanovic and V. Solo. Numerical methods for stochastic differential equations in matrix Lie groups made simple. *IEEE Trans. Automat. Control*, 63(12):4035–4050, 2018.
- [67] 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.
- [68] R. I. McLachlan, K. Modin, H. Munthe-Kaas, and O. Verdier. Butcher series: a story of rooted trees and numerical methods for evolution equations. *Asia Pac. Math. Newsl.*, 7(1):1–11, 2017.
- [69] G. N. Milstein and M. V. Tretyakov. *Stochastic numerics for mathematical physics*. Scientific Computation. Springer-Verlag, Berlin, 2004.
- [70] H. Munthe-Kaas. Lie-Butcher theory for Runge-Kutta methods. *BIT Numer. Math.*, 35:572–587, 1995.
- [71] H. Munthe-Kaas. Runge-Kutta methods on Lie groups. *BIT Numer. Math.*, 38:92–111, 1998.
- [72] H. Munthe-Kaas. High order Runge-Kutta methods on manifolds. *Applied Numerical Mathematics*, 29(1):115–127, 1999.
- [73] H. Munthe-Kaas. Geometric integration on symmetric spaces. *J. Comput. Dyn.*, 11(1):43–58, 2024.
- [74] H. Munthe-Kaas and J. Stava. Lie admissible triple algebras: The connection algebra of symmetric spaces. *Submitted*, 2023.
- [75] H. Munthe-Kaas and O. Verdier. Aromatic Butcher series. *Found. Comput. Math.*, 16(1):183–215, 2016.
- [76] H. Munthe-Kaas and O. Verdier. Integrators on homogeneous spaces: Isotropy choice and connections. *Foundations of Computational Mathematics*, 16:899–939, 2016.

- [77] H. Munthe-Kaas and A. Zanna. Numerical integration of differential equations on homogeneous manifolds. In *Foundations of Computational Mathematics: Selected Papers of a Conference Held at Rio de Janeiro, January 1997*, pages 305–315. Springer, 1997.
- [78] H. Z. Munthe-Kaas and A. Lundervold. On post-Lie algebras, Lie–Butcher series and moving frames. *Foundations of Computational Mathematics*, 13:583–613, 2013.
- [79] H. Z. Munthe-Kaas and W. M. Wright. On the Hopf algebraic structure of Lie group integrators. *Found. Comput. Math.*, 8(2):227–257, 2008.
- [80] J.-M. Oudom and D. Guin. On the Lie enveloping algebra of a pre-Lie algebra. *J. K-Theory*, 2(1):147–167, 2008.
- [81] B. Owren. Order conditions for commutator-free Lie group methods. *Journal of Physics A: Mathematical and General*, 39(19):5585, 2006.
- [82] B. Owren and A. Marthinsen. Runge-Kutta methods adapted to manifolds and based on rigid frames. *BIT Numer. Math.*, 39(1):116–142, 1999.
- [83] L. Rahm. An operadic approach to substitution in Lie-Butcher series. *Forum Math. Sigma*, 10:Paper No. e20, 29, 2022.
- [84] A. Rößler. Runge-Kutta methods for Stratonovich stochastic differential equation systems with commutative noise. In *Proceedings of the 10th International Congress on Computational and Applied Mathematics (ICCAM-2002)*, volume 164/165, pages 613–627, 2004.
- [85] A. Rößler. Stochastic Taylor expansions for the expectation of functionals of diffusion processes. *Stochastic Anal. Appl.*, 22(6):1553–1576, 2004.
- [86] T. Shardlow. Modified equations for stochastic differential equations. *BIT Numer. Math.*, 46(1):111–125, 2006.
- [87] 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.
- [88] 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.
- [89] J. Thompson. Brownian motion and the distance to a submanifold. *Potential Anal.*, 45(3):485–508, 2016.
- [90] J. L. Vázquez. *The porous medium equation. Mathematical theory*. Oxford Math. Monogr. Oxford: Oxford University Press, 2007.
- [91] G. Vilmart. Postprocessed integrators for the high order integration of ergodic SDEs. *SIAM J. Sci. Comput.*, 37(1):A201–A220, 2015.
- [92] Z. Wang, Q. Ma, Z. Yao, and X. Ding. The Magnus expansion for stochastic differential equations. *J. Nonlinear Sci.*, 30(1):419–447, 2020.
- [93] K. C. Zygalakis. On the existence and the applications of modified equations for stochastic differential equations. *SIAM J. Sci. Comput.*, 33(1):102–130, 2011.

Appendices

A Order conditions for third weak order

For the sake of completion, we detail in Table 2 the 92 additional order conditions for weak order 3 with stochastic frozen flow methods (2.7). Recall that these conditions are not independent and satisfy shuffle relations (see Remark 2.14), so that only 70 of these conditions are actually necessary.

Exotic forest π	$a(\pi)$	$e(\pi)$
	$z_{i,k_3}^0 Z_{i,j,k_2}^0 Z_{j,k,k_1}^0$	$\frac{1}{6}$
	$\sum_{k_1 \geq k_2}^! z_{i,k_3}^0 Z_{i,j,k_2}^0 Z_{i,k,k_1}^0$	$\frac{1}{6}$
	$\sum_{k_1 \geq k_2}^! z_{i,k_4}^0 Z_{i,j,k_3}^0 Z_{j,k_2}^{d_1} Z_{j,k_1}^{d_1}$	$\frac{1}{6}$
	$\sum_{k_1 \geq k_3}^! z_{i,k_4}^0 Z_{i,j,k_3}^0 Z_{j,k_2}^{d_1} Z_{i,k_1}^{d_1}$	$\frac{1}{3}$
	$\sum_{k_3 \geq k_1}^! z_{i,k_4}^0 Z_{i,j,k_3}^0 Z_{j,k_2}^{d_1} Z_{i,k_1}^{d_1}$	$\frac{1}{3}$
	$\sum_{k_1 \geq k_2 \geq k_3}^! z_{i,k_4}^0 Z_{i,j,k_3}^0 Z_{i,k_2}^{d_1} Z_{i,k_1}^{d_1}$	$\frac{1}{6}$
	$\sum_{k_1 \geq k_3 \geq k_2}^! z_{i,k_4}^0 Z_{i,j,k_3}^0 Z_{i,k_2}^{d_1} Z_{i,k_1}^{d_1}$	0
	$\sum_{k_3 \geq k_1 \geq k_2}^! z_{i,k_4}^0 Z_{i,j,k_3}^0 Z_{i,k_2}^{d_1} Z_{i,k_1}^{d_1}$	$\frac{1}{6}$
	$\sum_{k_1 \geq k_2 \geq k_3 \geq k_4}^! z_{i,k_5}^0 Z_{i,k_4}^{d_2} Z_{i,k_3}^{d_2} Z_{i,k_2}^{d_1} Z_{i,k_1}^{d_1}$	$\frac{1}{6}$
	$\sum_{k_1 \geq k_2 \geq k_3 \geq k_4}^! z_{i,k_5}^0 Z_{i,k_4}^{d_2} Z_{i,k_3}^{d_1} Z_{i,k_2}^{d_2} Z_{i,k_1}^{d_1}$	0
	$\sum_{k_1 \geq k_2 \geq k_3 \geq k_4}^! z_{i,k_5}^0 Z_{i,k_4}^{d_1} Z_{i,k_3}^{d_2} Z_{i,k_2}^{d_2} Z_{i,k_1}^{d_1}$	0
	$\sum_{k_1 \geq k_2}^! Z_{j,k,k_3}^0 Z_{j,k_2}^0 z_{i,k_1}^0$	$\frac{1}{6}$
	$\sum_{k_2 \geq k_1}^! Z_{j,k,k_3}^0 z_{j,k_2}^0 z_{i,k_1}^0$	$\frac{1}{3}$
	$\sum_{k_1 \geq k_2, k_3 \geq k_4}^! Z_{j,k_4}^{d_1} Z_{j,k_3}^{d_1} z_{j,k_2}^0 z_{i,k_1}^0$	$\frac{1}{6}$
	$\sum_{k_2 \geq k_1, k_3 \geq k_4}^! Z_{j,k_4}^{d_1} Z_{j,k_3}^{d_1} z_{j,k_2}^0 z_{i,k_1}^0$	$\frac{1}{3}$
	$\sum_{k_1 \geq k_2}^! Z_{j,k_4}^{d_1} Z_{i,k_3}^{d_1} z_{j,k_2}^0 z_{i,k_1}^0$	$\frac{1}{6}$
	$\sum_{k_1 \geq k_2}^! Z_{j,k_4}^{d_1} Z_{i,j,k_3}^0 z_{i,k_2}^0 z_{k_1}^{d_1}$	0
	$\sum_{k_1 \geq k_2, k_3 \geq k_4}^! Z_{i,j,k_4}^0 Z_{i,k_3}^{d_1} z_{i,k_2}^0 z_{k_1}^{d_1}$	0
	$\sum_{k_1 \geq k_2, k_4 \geq k_3}^! Z_{i,j,k_4}^0 Z_{i,k_3}^{d_1} z_{i,k_2}^0 z_{k_1}^{d_1}$	0
	$\sum_{k_2 \geq k_1, k_3 \geq k_4}^! Z_{i,j,k_4}^0 Z_{i,k_3}^{d_1} z_{i,k_2}^0 z_{k_1}^{d_1}$	0
	$\sum_{k_2 \geq k_1, k_4 \geq k_3}^! Z_{i,j,k_4}^0 Z_{i,k_3}^{d_1} z_{i,k_2}^0 z_{k_1}^{d_1}$	$\frac{1}{3}$
	$\sum_{k_1 \geq k_2, k_3 \geq k_4 \geq k_5}^! Z_{i,k_5}^{d_2} Z_{i,k_4}^{d_2} Z_{i,k_3}^{d_1} z_{i,k_2}^0 z_{k_1}^{d_1}$	0
	$\sum_{k_1 \geq k_2, k_3 \geq k_4 \geq k_5}^! Z_{i,k_5}^{d_2} Z_{i,k_4}^{d_1} Z_{i,k_3}^{d_2} z_{i,k_2}^0 z_{k_1}^{d_1}$	0
	$\sum_{k_1 \geq k_2, k_3 \geq k_4 \geq k_5}^! Z_{i,k_5}^{d_1} Z_{i,k_4}^{d_2} Z_{i,k_3}^{d_2} z_{i,k_2}^0 z_{k_1}^{d_1}$	0

