

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/162640>

Copyright and reuse:

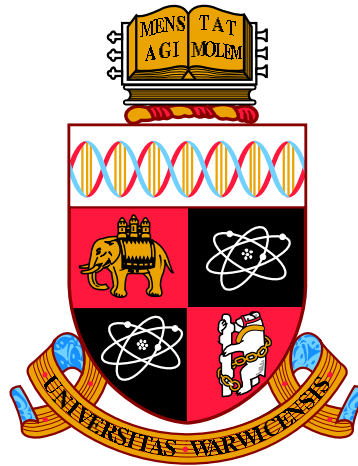
This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk



Stochastic processes on curved spaces

by

Veno Mramor

A thesis submitted for the degree of
Doctor of Philosophy in Statistics

University of Warwick, Department of Statistics
March 2021

Contents

List of Figures	iii
Acknowledgments	iv
Declarations	v
Abstract	vi
Introduction and overview	1
Chapter 1 Continuous semimartingales on manifolds	3
1.1 Smooth manifolds, smooth maps and vector fields	4
1.2 Semimartingales and SDEs on manifolds	5
1.3 Connection, frame bundle, horizontal lift and anti-development	7
1.4 Stochastic horizontal lift and stochastic anti-development of continuous semimartingales	12
1.5 Brownian motion on Riemannian manifolds	13
Chapter 2 Spherical Brownian motion and its projections	16
2.1 Introduction and main results	16
2.2 Characterization of the projected process	20
2.2.1 Skew-product decomposition for SDE (2.2)	21
2.2.2 Pathwise uniqueness for SDE (2.2)	28
Chapter 3 Simulation of spherical Brownian motion	33
3.1 Introduction	33
3.2 The skew-product decomposition	37
3.3 Simulation of Wright-Fisher diffusion processes	40

3.4	General spherical Brownian motion and new representation of transition density	42
3.5	Computer simulations	45
3.6	Brownian motion on projective spaces	49
Chapter 4	Lévy processes on smooth manifolds with a connection	52
4.1	Introduction	52
4.2	Construction and characterization of Lévy processes on manifolds with a connection	55
4.2.1	Marcus SDE	56
4.2.2	Horizontal Lévy processes and Lévy processes on a smooth manifold with a connection	59
4.2.3	Generators of Lévy processes on manifolds with a connection	62
4.2.4	Construction of horizontal lifts and anti-developments for discontinuous processes	65
4.3	Examples	69
4.3.1	Riemannian manifolds	69
4.3.2	Lie groups	72
4.3.3	Further examples	75
4.4	Proofs	77
4.4.1	Marcus SDEs on manifolds	77
4.4.2	Proofs of Proposition 4.2.9, Lemma 4.2.11 and Proposition 4.2.12	80
4.4.3	Proof of Theorem 4.2.19	82
4.4.4	Proof of Theorem 4.2.14	90
4.4.5	Proofs of results in Section 4.3.2	97
Appendix A		100
A.1	Semimartingales	100
A.2	Lévy processes and Brownian motion	102

List of Figures

1.1	Developing a curve on the sphere	12
3.1	Absolute and relative errors of the new representation	46
3.2	Plots of values of $b_{m+i}^{(\tau, d-1)}(m)$ and partial sums $\sum_{k=m}^{m+i} (-1)^{k-m} b_k^{(\tau, d-1)}(m)$.	47
3.3	Plots of values $q_m^{d-1}(\tau)$	48
4.1	Relation between the solution X and the integrator W of the Marcus SDE	58
4.2	Parallel transport and non-trivial holonomy group of the sphere	61

Acknowledgments

I want to thank my supervisor Aleksandar Mijatović for all the support and knowledge he passed on to me. Our regular meetings were always insightful and provided a gentle yet thorough induction into an academic life. I would like to thank Gerónimo Uribe Bravo for all the help while working on joint papers and for hosting me during my academic visit to UNAM in Mexico City in September 2019. I am grateful to the Department of Statistics for funding my studies.

I want to thank the examiners Andreas Kyprianou and Kenneth David Elworthy for their careful reading of the thesis and helpful comments which greatly improved the final version of the thesis.

Finally, I would like to thank my family, and especially my partner Tjaša, for all the love and support during my studies.

Declarations

I declare that I have written a PhD thesis entitled **Stochastic processes on curved spaces** completely by myself under the supervision of Aleksandar Mijatović. I confirm that the thesis has not been submitted for a degree at any other university.

During my PhD I have written papers [MMU18], [MMU20] and [MMU21] in collaboration with Aleksandar Mijatović and Gerónimo Uribe Bravo. These papers had appeared in peer-reviewed journals. Chapter 2 is based on [MMU18], while the results from [MMU20] and [MMU21] are presented in Chapter 3. Additionally, I have written a paper [MM20] in collaboration with Aleksandar Mijatović. It has been submitted to a peer-reviewed journal and the preprint is available on ArXiv. This paper forms the basis of Chapter 4.

Abstract

We investigate semimartingales and other classes of stochastic processes on smooth manifolds. First, we introduce continuous semimartingales on smooth manifolds and review a well-established theory surrounding it, including the notion of Brownian motion on Riemannian manifolds. Second, we focus on the Brownian motion on a unit sphere of arbitrary dimension and consider the projected process consisting of several of the components of the original spherical Brownian motion. We obtain a stochastic differential equation the projection satisfies and find its invariant measure. We study a more general stochastic differential equation on the unit ball and show the existence and path-wise uniqueness of the solution and we show that the solution enjoys a skew-product decomposition analogous to the classical skew-product decomposition of Euclidean Brownian motion. Third, a variant of the skew-product decomposition exists for the full spherical Brownian motion and features Wright-Fisher diffusion process as the radial process. Since Wright-Fisher diffusion processes can be simulated exactly as in [JS17], we utilise the rapid-spinning phenomenon of the skew-product decomposition and obtain a novel algorithm for the exact simulation of the increments of spherical Brownian motion. We analyse the algorithm numerically and obtain the range of parameters where it remains stable and show that it then outperforms previous simulation algorithms from the literature. Fourth, we define a Lévy process on a smooth manifold with a connection as a projection of a solution of a Marcus stochastic differential equation on the holonomy bundle, driven by the holonomy-invariant Euclidean Lévy process. We show that Lévy processes on manifolds are Markov processes and characterize them via their infinitesimal generators. In order to prove this characterization we obtain the path-wise construction of the stochastic horizontal lift and stochastic anti-development of a general discontinuous semimartingale, which generalizes the result in [PE92] to manifolds with non-unique geodesics between distinct points. We compare our definition with previous definitions of Lévy processes on Lie groups and Riemannian manifolds and provide several examples.

Introduction and overview

Stochastic processes are one of the building blocks of probability theory. The most well-studied stochastic processes are those with state space equal to either the real line \mathbb{R} or the Euclidean space \mathbb{R}^d for some $d \in \mathbb{N}$. However, even considering this reduction of possible state spaces, the class of all such processes is still too broad to study effectively and several subclasses with additional “pleasant” properties are typically considered.

A class of Markov processes consists of processes whose future evolution depends on the past only through the present value. Its subclass of Lévy processes comprises processes with stationary and independent increments. For martingales the conditional expectation of the future values is equal to the current value and a class of semimartingales, i.e. processes which can be represented as a sum of a local martingale and a finite variation process, turns out to be the most general class of processes against which a meaningful notion of stochastic integration can be defined. While Markov processes can be easily generalized to more general state spaces, the definitions of the rest of the aforementioned classes depend deeply on the structure of the Euclidean space. Therefore, their corresponding generalizations cannot be achieved as trivially. However, on certain suitable state spaces some of these generalizations can indeed be obtained by utilising alternative approaches.

The main goal of this thesis is to study some particular classes of stochastic process whose state space is no longer a (flat) Euclidean space but a (curved) smooth manifold and the thesis is structured as follows.

In Chapter 1 we introduce all the prerequisites from differential geometry required to study semimartingales and stochastic differential equation on manifolds. The chapter focuses on continuous processes on manifolds and we review some of the well-known results about their stochastic horizontal lifts and stochastic anti-developments. Moreover, we introduce a notion of a Brownian motion on a Riemannian manifold.

Chapter 2 studies Brownian motion on the unit sphere of arbitrary dimension. More precisely, we study the projected process consisting of several components of the spher-

ical Brownian motion. We characterize the projected process in terms of the stochastic differential equation it satisfies and find its invariant measure. We also study a more general stochastic differential equation and show the existence and path-wise uniqueness of the solution and show that it enjoys a skew-product decomposition. The results in Chapter 2 are taken from [MMU18].

In Chapter 3 we present an algorithm for exact simulation of increments of spherical Brownian motion. The algorithm is based on the skew-product decomposition which features Wright-Fisher diffusion processes, which can be simulated exactly. We also analyse the algorithm numerically and compare it to the other available algorithms from the literature. The content of this chapter is based on [MMU20] and [MMU21].

In Chapter 4 we show how a notion of Lévy processes can be defined on a manifold equipped with a connection. As we need to consider discontinuous processes, we introduce a notion of a Marcus SDE on a manifold, which allows that both integrators and solutions have jumps. The definition of Lévy processes on a manifold is inspired by the Eells-Elworthy-Malliavin construction of the Brownian motion on a Riemannian manifold. Thus, we have to generalize the results on stochastic horizontal lifts and stochastic anti-developments from Chapter 1 to discontinuous semimartingales. We show that Lévy processes are Markov and we characterize them by their infinitesimal generator. Finally, we present some examples of Lévy processes on manifolds and additionally, compare our new definition with previous ones on Lie groups and Riemannian manifolds. The exposition is based on [MM20].

In Appendix A we briefly review the notions of semimartingales, Lévy processes and Brownian motion and some of the well-known results about them.

Chapter 1

Continuous semimartingales on manifolds

The goal of this chapter is to show how the notion of (continuous) semimartingales can be extended to smooth manifolds and to consider stochastic differential equations (SDEs) on manifolds. Continuity assumption streamlines the exposition and allows us to establish a nice relation between continuous semimartingales on a manifold and continuous semimartingales on the Euclidean space of the same dimension as the manifold. The theory of discontinuous semimartingales on manifolds has additional intricacies, making it less-studied in the literature. We tackle some of its aspects in Chapter 4.

This chapter reviews main facts from differential geometry and stochastic analysis on the manifolds which will be used in the rest of the thesis. For a classical differential geometry the main references are [KN63, Lee12]. Our exposition is based mostly on [Hsu02, Chapters 1 and 2] since it nicely explains the interplay between differential geometry and stochastic analysis.

Finally, we introduce the notion of a Brownian motion on a Riemannian manifold and in particular consider the Brownian motion on sphere which will be a central object of study in Chapters 2 and 3.

We introduce most of the notions in plain text and emphasise them in boldface. Some more important notions are further emphasised in their own definitions.

In this chapter, as well as in the rest of the thesis we always use Einstein's convention of implicitly summing over repeated indices, when they appear once as an upper and once as a lower index.

1.1 Smooth manifolds, smooth maps and vector fields

We consider a **smooth manifold** M of dimension d , i.e. a topological manifold which locally looks like \mathbb{R}^d along with a smooth structure given by some countable atlas (consisting of local charts). A **local chart** (O, φ) is a homeomorphism $\varphi: O \rightarrow \varphi(O) \subseteq \mathbb{R}^d$ from an open set $O \subseteq M$ onto an open set in \mathbb{R}^d and gives rise to **local coordinates** by $\varphi = (x^1, \dots, x^d)$. A function $f: M \rightarrow \mathbb{R}$ is called **smooth**, if it is smooth as a function of local coordinates, more precisely, if a function $f \circ \varphi^{-1}: \varphi(O) \rightarrow \mathbb{R}$ is smooth. The space of smooth functions on M is denoted by $C^\infty(M)$. A map $F: M \rightarrow N$ between manifolds M and N is called **smooth**, if for any point $p \in M$ we can find a local chart (O, φ) containing p and a local chart (\tilde{O}, ψ) containing $F(p)$, such that $F(O) \subseteq \tilde{O}$ and $\psi \circ F \circ \varphi^{-1}: \varphi(O) \rightarrow \psi(\tilde{O})$ is smooth.

A **tangent vector** v at a point $p \in M$ is a linear mapping $v: C^\infty(M) \rightarrow \mathbb{R}$ which is a derivation, i.e. $v(fg) = f(p)v(g) + g(p)v(f)$. The **tangent space** T_pM consists of all tangent vectors at the point $p \in M$ and is a vector space of dimension d . All tangent spaces can be combined into a smooth manifold TM of dimension $2d$ called the **tangent bundle**. The **vector field** \mathcal{X} on M is given as a smooth section of the tangent bundle, i.e. it is a smooth prescription of a tangent vector¹ \mathcal{X}_p to each point $p \in M$. The **space of all vector fields** is denoted by $\Gamma(TM)$ and can be given a structure of a $C^\infty(M)$ -module, where multiplication by functions is defined pointwise. Additionally, each vector field acts pointwise on smooth functions and the resulting function $\mathcal{X}f: p \mapsto \mathcal{X}_p(f)$ is smooth again. Therefore, we can alternatively think of vector fields as those operators on the space of smooth functions which are derivations, i.e. operators satisfying the product rule $\mathcal{X}(fg) = f\mathcal{X}g + g\mathcal{X}f$. On each local chart (O, φ) we can locally define **coordinate vector fields** $\frac{\partial}{\partial x^j}$ by $\frac{\partial}{\partial x^j} \Big|_p f := \frac{\partial f \circ \varphi^{-1}}{\partial x^j}(\varphi(p))$. When considering \mathbb{R} as a smooth manifold with a single chart (\mathbb{R}, id) , we also write $\frac{d}{dt}$ for the single coordinate vector field.

A smooth map $F: M \rightarrow N$ induces a linear map between tangent spaces. More precisely, for any $p \in M$ there is a linear map $dF_p: T_pM \rightarrow T_{F(p)}N$ called the **differential of F at p** defined by $dF_p(v)(g) := v(g \circ F)$ for any $g \in C^\infty(N)$. Furthermore, the differential can be thought of as a smooth map $dF: TM \rightarrow TN$ between the tangent bundles.

To each vector field we associate its integral curves. A smooth curve $\gamma: I \subseteq \mathbb{R} \rightarrow M$ is an **integral curve of \mathcal{X}** if $\dot{\gamma}_s := d\gamma_s \left(\frac{d}{dt} \Big|_s \right) = \mathcal{X}_{\gamma_s}$ holds for every $s \in I$. As a consequence of local existence and uniqueness theorem for ordinary differential equations (ODEs), there exists a unique integral curve on the maximal time interval once the

¹We also write $\mathcal{X}|_p$ or $\mathcal{X}(p)$ when typographically convenient.

initial point is fixed. We say that a vector field is **complete** if all its integral curves are defined for all \mathbb{R} . This allows us to talk about a flow of a vector field and a flow exponential map. A **flow of a complete vector field** \mathcal{X} is a one-parameter group of diffeomorphisms $\Phi_s^{\mathcal{X}}: M \rightarrow M$ for every $s \in \mathbb{R}$, such that for any $p \in M$ the curve $s \mapsto \gamma_{\mathcal{X}}^p(s) := \Phi_s^{\mathcal{X}}(p)$ is the unique integral curve of \mathcal{X} with $\gamma_{\mathcal{X}}^p(0) = p$. We then have the flow property $\Phi_s^{\mathcal{X}} \circ \Phi_t^{\mathcal{X}} = \Phi_{s+t}^{\mathcal{X}}$. We additionally define the **flow exponential map** $\exp(\mathcal{X}) := \Phi_1^{\mathcal{X}}: M \rightarrow M$ which represents a diffeomorphism which moves points a unit time along the integral curves of a vector field \mathcal{X} .

1.2 Semimartingales and SDEs on manifolds

A notion of continuous semimartingales can be generalized to manifold valued processes. A continuous M -valued stochastic process X is a **semimartingale** if $f(X)$ is a continuous real-valued semimartingale for every $f \in C^\infty(M)$. We want to consider manifold valued SDEs. The coefficient will be smooth, hence the solution will be uniquely determined, but typically only defined on a stochastic interval $[0, \tau)$, where τ will be a (predictable) stopping time denoting the lifetime of the solution. Hence we need to extend the notion of a semimartingale to a process defined on a stochastic interval. For a predictable stopping time τ we say that a continuous M -valued process $X = (X_t)_{t \in [0, \tau)}$ is a **semimartingale on a stochastic interval** $[0, \tau)$ if $f(X^{\tau_n})$ is a real-valued semimartingale for any $f \in C^\infty(M)$ and $n \in \mathbb{N}$, where $X^{\tau_n} = (X_t^{\tau_n})_{t \in \mathbb{R}_+}$ with $X_t^{\tau_n} = X_{\min\{\tau_n, t\}}$ for $t \in \mathbb{R}_+$ denotes a stopped process and stopping times $\{\tau_n; n \in \mathbb{N}\}$ form an announcing sequence of τ , i.e. $\tau_n \uparrow \tau$ as $n \rightarrow \infty$ and $\tau_n < \tau$ a.s. for every $n \in \mathbb{N}$. Such a notion of a (continuous) semimartingale was first defined in [Sch80] and then studied in [Sch82] and [Mey81].

We can now define what it means for a process to be a solution of an SDE on a manifold.

Definition 1.2.1. Let $V_1, \dots, V_\ell \in \Gamma(TM)$ be smooth vector fields on M and W a continuous \mathbb{R}^ℓ -valued semimartingale. Then we call a continuous M -valued semimartingale X on $[0, \tau)$ a **solution of an SDE**

$$dX_t = V_i(X_t) \circ dW_t^i \tag{1.1}$$

if for any $f \in C^\infty(M)$ it satisfies the equation

$$f(X_t) - f(X_0) = \int_0^t V_i f(X_s) \circ dW_s^i \quad \text{for } 0 \leq t < \tau. \tag{1.2}$$

Remark 1.2.2. (i) Since $V_i f \in C^\infty(M)$, the integrals in (1.2) are just ordinary Stratonovich integrals of continuous real semimartingales $V_i f(X_t)$ against continuous real semimartingales W_t^i .

(ii) A more precise meaning of (1.2) is as follows. For an announcing sequence $\{\tau_n; n \in \mathbb{N}\}$ of τ we require that for every $n \in \mathbb{N}$ the stopped process X^{τ_n} is a semimartingale and that equation (1.2) with X exchanged by X^{τ_n} holds for every $t \in \mathbb{R}_+$.

It is crucial to use the Stratonovich integral in the definition of a solution and not the more common Itô integral, otherwise the solutions of SDEs would typically not exist. The Stratonovich integral obeys a classical chain rule formula which can be used to show that if vector fields V_i are tangent to some submanifold in \mathbb{R}^n , then a solution will stay on that submanifold when started there. This is important since any manifold M can be embedded in a high-dimensional Euclidean space by Whitney's theorem and one can then combine above facts with classical results on SDEs in Euclidean spaces to prove the following uniqueness and existence results for manifold valued SDEs.

Proposition 1.2.3 (Theorem 1.2.9 from [Hsu02]). *There exists a unique solution X of SDE (1.1) and it is a semimartingale defined on a maximal stochastic interval $[0, \tau)$, where τ is a predictable stopping time.*

Every continuous M -valued semimartingale can be realised as a solution of some SDE on the d -dimensional manifold M (see [Sch82] or [Hsu02, Lemma 2.3.3]). However, this typically requires more than d vector fields and the same number of continuous real semimartingales driving the SDE. Since d is the dimension of the manifold, it seems intuitively plausible that there should be some relation between \mathbb{R}^d -valued continuous semimartingales and continuous M -valued semimartingales. Such a correspondence typically cannot be achieved directly on the manifold and we need an intermediate step where we consider the so-called horizontal processes on the frame bundle of the manifold. In the next section we introduce all the necessary notions from differential geometry required for such a construction. Additionally, we show the correspondence between smooth curves on the manifold M and smooth curves on the Euclidean space \mathbb{R}^d which we then generalize in Section 1.4 to the aforementioned correspondence between continuous semimartingales.

1.3 Connection, frame bundle, horizontal lift and anti-derivation

A **(linear) connection** is given in terms of a covariant derivative, which prescribes how to derive vector fields along each other. A **covariant derivative** is a bilinear map $\nabla: \Gamma(TM) \times \Gamma(TM) \rightarrow \Gamma(TM), (X, Y) \mapsto \nabla_X Y$ which is additionally $C^\infty(M)$ -linear in the first argument and obeys the product rule in the second argument i.e. $\nabla_X(fY) = f\nabla_X Y + (Xf)Y$. Covariant derivatives act locally; to compute $(\nabla_X Y)_p$ it is enough to know only the value X_p and the values of Y along a curve through p which has a tangent vector X_p at the point p . In local coordinates x^i , the vector fields $\partial_i := \frac{\partial}{\partial x^i}$ form a local frame, i.e. the tangent vectors $\partial_i|_p$ form a basis of the tangent space $T_p M$ at any point p in the local chart. Hence we can locally express $\nabla_{\partial_i} \partial_j = \Gamma_{ij}^k \partial_k$ for some smooth functions Γ_{ij}^k , which are known as Christoffel symbols. Christoffel symbols uniquely determine the covariant derivative.

A vector field \mathcal{X} along a curve γ , is said to be **parallel (along the curve γ)** if $\nabla_{\dot{\gamma}} \mathcal{X} = 0$ holds at every point of the curve. In a local chart a vector field along γ can be expressed as $\mathcal{X}_{\gamma_t} = \mathcal{X}^k(t) \partial_k$ for some smooth functions \mathcal{X}^k and a parallel vector field is then given as a solution of the following system of ordinary differential equations (ODEs):

$$\dot{\mathcal{X}}^k(t) + \Gamma_{ij}^k(\gamma_t) \dot{\gamma}_t^i \mathcal{X}^j(t) = 0.$$

Once a curve γ with $\gamma_0 = p$ and $v \in T_p M$ are fixed, the local existence and uniqueness theorem for ODEs guarantees that there is a unique parallel vector field \mathcal{X} along γ with $\mathcal{X}_p = v$. Since the system of ODEs is (non-autonomous) linear, the solution cannot explode and is defined along the whole curve γ .

Actually, such a construction determines a unique (parallel) lift of a curve onto a tangent bundle (once the initial point is fixed). Additionally, using parallel vector fields along the curve γ with $\gamma_0 = p$ and $\gamma_1 = q$ we can connect the tangent spaces $T_p M$ and $T_q M$. More precisely, for each $v \in T_p M$ we define $\tau_\gamma(v) \in T_q M$ as a value at q of the unique parallel vector field \mathcal{X} (along γ) with $\mathcal{X}_p = v$. The map τ_γ is called the **parallel transport along γ** and can be shown to be a linear isomorphism.

Covariant derivative allows us to define **geodesics**, which are special curves γ such that their velocity vector field $\dot{\gamma}$ is parallel along γ . In local coordinates a geodesic $\gamma = (\gamma^i)$ is given as a solution of the system of second order ODEs

$$\ddot{\gamma}_t^k + \Gamma_{ij}^k(\gamma_t) \dot{\gamma}_t^i \dot{\gamma}_t^j = 0$$

for $k = 1, \dots, d$ and therefore a geodesic is uniquely determined once we fix the initial point and the initial velocity. This allows us to define the exponential map in essentially the same manner as on Riemannian manifolds. A **geodesic exponential map** is a map $\text{Exp}_p: O \subseteq T_p M \rightarrow M$, $\text{Exp}_p(X) := \gamma(1)$, where γ is a unique geodesic with $\gamma(0) = p$ and $\dot{\gamma}(0) = X$. Set O is an open star-shaped neighbourhood of $0 \in T_p M$ where the geodesic exponential map is well defined². If the domain of the geodesic exponential map at p is equal to T_p for any $p \in M$ we say that M is **geodesically complete**. Note that by Hopf-Rinow theorem compact Riemannian manifolds are always geodesically complete (see Section 1.5 for more details on Riemannian manifolds), but there exist non-metric connections on compact manifolds which are not geodesically complete [O’N83, Ex. 7.16].

Let $F(M)$ denote a **frame bundle over** M (defined as a bundle of linear frames in [KN63, p. 56]) with a bundle projection $\pi: F(M) \rightarrow M$. The bundle $F(M)$ is a smooth manifold of dimension $d + d^2$ and the projection π is a smooth map. A frame $u \in F(M)$, such that $\pi(u) = p$, is an ordered basis u_1, \dots, u_d of the tangent space $T_p M$. We may regard it as a linear isomorphism $u: \mathbb{R}^d \rightarrow T_p M$ given by

$$u(x) := x^i u_i \quad \text{for each } x = x^i e_i \in \mathbb{R}^d, \quad (1.3)$$

where e_1, \dots, e_d is a standard basis of \mathbb{R}^d . The general linear group $\text{GL}(d)$, consisting of invertible matrices in $\mathbb{R}^d \otimes \mathbb{R}^d$, acts on $F(M)$ on the right by $R_g(u) = ug := u \circ g$, where $g \in \text{GL}(d)$ is considered as an automorphism of \mathbb{R}^d . This turns $F(M)$ into a principal fibre bundle (over M) with a structure group $\text{GL}(d)$ (see definition in [KN63, §5, Ch. I]).

On the frame bundle a notion of **vertical** vectors is naturally defined. More precisely, the vertical space is defined by $V_u F(M) := \ker d\pi_u$. We now show how the covariant derivative can be used to define horizontal vectors. Let u_t be a smooth curve in $F(M)$ so that $\pi(u_t)$ is a smooth curve in M . A curve u_t is called **horizontal** if the vector field $u_t e$ is parallel along the curve $\pi(u_t)$ for each $e \in \mathbb{R}^d$. Additionally, a vector in $T_u F(M)$ is called **horizontal** if it is a velocity vector of a horizontal curve through u . The space of horizontal vectors at u is denoted by $H_u F(M)$. It is easily seen that the following equivariance property holds: $d(R_g)_u(H_u F(M)) = H_{ug} F(M)$ for all $g \in \text{GL}(d)$, where $d(R_g)_u: T_u F(M) \rightarrow T_{ug} F(M)$ is a differential at u of a (right) group action map $R_g: u \mapsto ug$. The vertical and horizontal space intersect trivially and furthermore the equality $T_u F(M) = V_u F(M) \oplus H_u F(M)$ holds, giving us the splitting of tangent spaces

²The differential equation for a geodesic has a unique solution defined on a maximal time interval. However, this interval might not be all of \mathbb{R} . Hence geodesics might explode in finite time and the domain of the geodesic exponential map might have to be restricted.

of the frame bundle³.

Since we have the splitting, dimensionality argument shows that the differential $d\pi_u$ restricted to the horizontal space is an isomorphism between $H_u F(M)$ and $T_{\pi(u)}M$. The inverse of this (restricted) differential defines a **horizontal lift** of tangent vectors in M . Moreover, horizontal lifts can be obtained for (piecewise) smooth curves. Let γ be a smooth curve in M . We call a curve u_t in $F(M)$ a **horizontal lift of the curve** γ if $\pi(u_t) = \gamma_t$ and if it is horizontal, i.e. $\dot{u}_s \in H_{u_s}F(M)$ for all s . Once the initial frame $u_0 \in \pi^{-1}(\{\gamma_0\})$ is fixed, the horizontal lift is unique and exists for all times for which the underlying curve γ is defined. This can be seen by the following computation in a local chart. A local chart for $O \subseteq M$ around p induces a local chart for $\pi^{-1}(O) \subseteq F(M)$ around $\pi^{-1}(p)$. A frame u above p can be uniquely represented as $u = (x^i, r_m^k)$ where $ue_m = r_m^k \partial_k$. Then the horizontal lift of γ is given as $(\gamma_t^i, r_m^k(t))$, where smooth functions r_m^k satisfy a system of linear ODEs

$$\dot{r}_m^k(t) + r_m^l(t)\Gamma_{jl}^k(\gamma_t)\dot{\gamma}_t^j = 0. \quad (1.4)$$

After fixing the initial frame the uniqueness and existence of the horizontal lift thus follow from the theory of ODEs and linearity of the system implies non-explosion. Furthermore, if u_t is a horizontal lift of γ with an initial frame u_0 , then $u_t g$ is a horizontal lift of γ with an initial frame $u_0 g$ for any $g \in \text{GL}(d)$. In particular, we recover the parallel transport as $\tau_\gamma = u_1 u_0^{-1}$ for any horizontal lift u_t of γ , since the quantity on the right hand side does not depend on the initial frame (and hence on the particular choice of the horizontal lift).

Horizontal spaces can also be described by a family of horizontal vector fields. For any $e \in \mathbb{R}^d$ a **fundamental horizontal vector field** H_e is defined by $H_e(u) := (d\pi_u)^{-1}(ue) \in H_u F(M)$ for any frame u , where $d\pi_u: H_u F(M) \rightarrow T_{\pi(u)}M$ is an isomorphism. Equivalently, the vector field H_e is uniquely determined by

$$(H1) \quad H_e(u) \in H_u F(M),$$

$$(H2) \quad d\pi_u(H_e(u)) = ue$$

for each $u \in F(M)$.

Since the definition is pointwise, smoothness is not immediate, but quickly follows from the following representation in local coordinates taken from [Hsu02, Prop. 2.1.3].

³This kind of splitting is also known as a principal Ehresmann connection on a principal fibre bundle, whereas covariant derivative is also known as a linear connection and can be considered as a structure on the tangent bundle TM . Since TM is an associated bundle of $F(M)$, (principal) connections on $F(M)$ are naturally induced by linear connections on TM and vice versa [KN63, Thm. 7.5 in Ch. III].

Lemma 1.3.1. *In terms of a local chart on $F(M)$ where $u = (x, r) = (x^i, r_m^k) \in F(M)$, we have*

$$H_e(u) = e^i r_i^j \frac{\partial}{\partial x^j} - e^i r_i^j r_m^l \Gamma_{jl}^k(x) \frac{\partial}{\partial r_m^k},$$

where $e = e^i e_i$.

It is trivially seen that the map $e \mapsto H_e$ is linear. We denote $H_i := H_{e_i}$ for $i = 1, \dots, d$ and then the vector fields H_1, \dots, H_d span the whole horizontal space $H_u F(M)$ at any frame u . The following lemma relates the fundamental horizontal fields at different frames above the same point in the manifold.

Lemma 1.3.2 (Prop. III.2.2 from [KN63]). *Let $e \in \mathbb{R}^d$ and $g \in \text{GL}(d)$. Then*

$$H_e(ug) = d(R_g)_u(H_{ge}(u))$$

for every $u \in F(M)$, where $d(R_g)_u: T_u F(M) \rightarrow T_{ug} F(M)$ is a differential at u of a (right) group action map $R_g: u \mapsto ug$.

The following two lemmas are easy corollaries.

Lemma 1.3.3. *For any $u \in F(M)$, $e \in \mathbb{R}^d$, $g \in \text{GL}(d)$ we have*

$$\exp(H_e)(ug) = \exp(H_{ge})(u) \cdot g.$$

Proof. If ψ is an integral curve of H_{ge} with $\psi(0) = u$, then Lemma 1.3.2 implies that $\tilde{\psi} := \psi g$ is an integral curve of H_e with $\tilde{\psi}(0) = ug$ and the claim follows from the uniqueness of integral curves and the definition of the flow exponential map. \square

Lemma 1.3.4. *For every smooth function $f \in C^\infty(F(M))$, $x_1, \dots, x_m \in \mathbb{R}^d$ and $g \in \text{GL}(d)$ we have*

$$(H_{x_1} \cdots H_{x_m} f) \circ R_g = H_{gx_1} \cdots H_{gx_m} (f \circ R_g).$$

Proof. We prove the lemma by induction on m . The base case of $m = 1$ is just a restatement of Lemma 1.3.2. For an inductive step let the claim hold for m vectors in \mathbb{R}^d and take an additional vector x_{m+1} . Denote $g = H_{x_{m+1}} f$ and then

$$\begin{aligned} (H_{x_1} \cdots H_{x_{m+1}} f) \circ R_g &= (H_{x_1} \cdots H_{x_m} g) \circ R_g \\ &= H_{gx_1} \cdots H_{gx_m} (g \circ R_g) \\ &= H_{gx_1} \cdots H_{gx_{m+1}} (f \circ R_g), \end{aligned}$$

where the second equality holds by the induction hypothesis and the last equality by the base case $m = 1$. \square

So far we were able to start with a curve γ in M and after fixing an initial frame u_0 we could construct a unique horizontal lift u_t in the frame bundle $F(M)$. We can go one step further and define an **anti-development of a curve** γ (or of its lift u_t) by

$$w_t := \int_0^t u_s^{-1}(\dot{\gamma}_s) ds,$$

which is a smooth curve in \mathbb{R}^d started at 0. One could also reverse the construction and start with a curve w_t in \mathbb{R}^d and construct a horizontal curve u_t on $F(M)$ (and a curve $\gamma_t := \pi(u_t)$ on M) whose anti-development is w_t . Once we fix the initial frame this correspondence is unique and u_t is given as a solution of the ODE

$$\dot{u}_t = H_i(u_t)\dot{w}_t^i. \tag{1.5}$$

Therefore, we obtain a one-to-one correspondence between a (piecewise) smooth curves in M , their horizontal lifts in $F(M)$ and their anti-developments in \mathbb{R}^d once the initial frame is fixed.

The process of starting with a curve in \mathbb{R}^d and obtaining the corresponding developed curve on the manifold is also known as “rolling without slipping.” Intuitively, we “roll” the tangent space along the curve and the point of contact traces the developed curve on the manifold. An example on the sphere is shown in Figure 1.1.

It is interesting to check how the correspondence is expressed for special curves – geodesics. One quickly finds that the anti-development of a geodesic is a uniform motion, i.e. a constant speed straight line. This simplifies the ODE that the horizontal lift u_t is a solution of and actually shows that u_t is an integral curve of some fundamental horizontal vector field i.e. there exists some $e \in \mathbb{R}^d$ such that $\dot{u}_t = H_e(u_t)$. Therefore, uniform motion in \mathbb{R}^d corresponds to moving along an integral curve of some fundamental horizontal vector field on $F(M)$ or equivalently moving along a geodesic in M . This also relates flow exponential map⁴ on $F(M)$ to a geodesic exponential map on M so that the equality

$$\pi(\exp(H_e)(u)) = \text{Exp}_p(d\pi_u(H_e)) = \text{Exp}_p(ue) \tag{1.6}$$

holds for any $e \in \mathbb{R}^d$ and any frame u with $\pi(u) = p$.

⁴We have only defined flows of complete fields, but analogously we could define locals flows of (non-complete) fields.

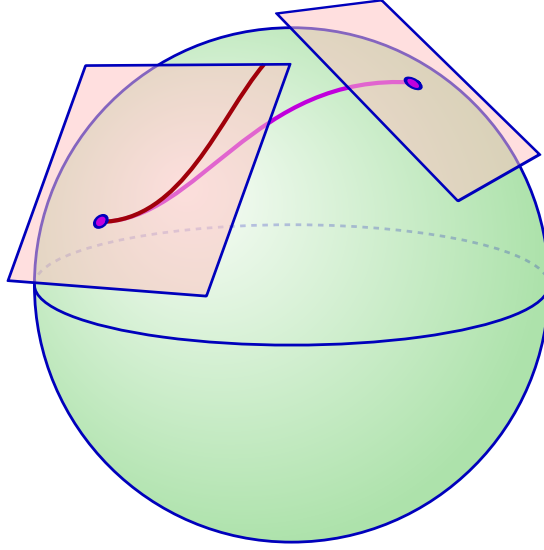


Figure 1.1: Developing a curve on the sphere (from [Wik20a])

1.4 Stochastic horizontal lift and stochastic anti-development of continuous semimartingales

The correspondence for smooth curves from the previous section can be generalized to continuous semimartingales once we use the notion of SDEs on manifolds from Section 1.2.

We wish to define horizontal semimartingales on a frame bundle. However, since sample paths of semimartingales are typically very irregular, there is no direct way of generalizing the notion of a parallel vector field along a semimartingale and we rather use the following definition inspired by equation (1.5) and the aforementioned correspondence between smooth curves, their horizontal lifts and anti-developments.

Definition 1.4.1. A continuous $F(M)$ -valued semimartingale U on $[0, \tau)$ is said to be **horizontal** if there exists an \mathbb{R}^d -valued continuous semimartingale $W = (W_t)_{t \in \mathbb{R}_+}$ with $W_0 = 0$, such that U is a solution of an SDE

$$dU_t = H_i(U_t) \circ dW_t^i \quad (1.7)$$

on the stochastic interval $[0, \tau)$. The semimartingale W is then called a **stochastic anti-development** of U (or of its projection $X = \pi(U)$). Additionally, if X is a continuous M -valued semimartingale on $[0, \tau)$, then a horizontal semimartingale U is called a **stochastic horizontal lift** of X if $\pi(U) = X$.

It is known that a stochastic horizontal lift of X exists and is uniquely determined on $[0, \tau)$ once the initial frame is chosen (see e.g. [Mey81], [Dar82] and [Shi82] for original results or [Hsu02, Thm. 2.3.5] for a more recent approach). Moreover, [Hsu02, Thm. 2.3.4] shows that the stochastic anti-development exists and is unique after choosing the initial frame. Thus, we have a unique correspondence between continuous M -valued semimartingales, their horizontal lifts, and continuous \mathbb{R}^d -valued semimartingales – their anti-developments.

1.5 Brownian motion on Riemannian manifolds

We start by defining what it means for a continuous process to be a diffusion generated by a second order differential operator on M .

Definition 1.5.1. Let L be a smooth second order differential operator on a manifold M . A continuous adapted process $X = (X_t)_{t \in [0, \tau)}$ is called a **diffusion generated by L** if it is an M -valued semimartingale on $[0, \tau)$ such that

$$M^f(X)_t := f(X_t) - f(X_0) - \int_0^t Lf(X_s) ds$$

is a local martingale for all $f \in C^\infty(M)$.

Remark 1.5.2. Process $M^f(X)$ is only defined on $[0, \tau)$, so we actually require that for each $n \in \mathbb{N}$ the stopped processes $M^f(X)^{\tau_n}$ are classical local martingales, where $\{\tau_n ; n \in \mathbb{N}\}$ form an announcing sequence of τ .

Recall that the Brownian motion on the Euclidean space \mathbb{R}^d has infinite lifetime and is generated by $\Delta/2$, where $\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x^i}$ is the Laplace operator on \mathbb{R}^d . We wish to generalize this operator to a Riemannian manifold and use it generate a process which can then be called a Brownian motion on a Riemannian manifold. The easiest way to generalize the operator is to recall the formula $\Delta f = \operatorname{div} \operatorname{grad} f$ for $f = (f^1, \dots, f^d) \in C^\infty(\mathbb{R}^d)$, where the gradient of a smooth function is a vector field given by $\operatorname{grad} f = \left(\frac{\partial f^1}{\partial x^1}, \dots, \frac{\partial f^d}{\partial x^d} \right)$ and the divergence of a vector field $X = (X^1, \dots, X^d)$ is a smooth function $\operatorname{div} X = \sum_{i=1}^d \frac{\partial X^i}{\partial x^i}$. Thus, it suffices to generalize the gradient and divergence operators to Riemannian manifolds.

We start with some general definitions. A manifold M is a **Riemannian manifold** if it is equipped with a metric $\langle \cdot, \cdot \rangle$ which is a symmetric positive definite bilinear 2-form i.e. a bilinear map $\langle \cdot, \cdot \rangle_p : T_p M \times T_p M \rightarrow \mathbb{R}$ which is positive definite and varies smoothly in p .

Recall that for any $f \in C^\infty(M)$ we can define its differential df , which is a differential 1-form, i.e. a section of the cotangent bundle. Furthermore, the Riemannian metric allows us to identify tangent and cotangent vectors (hence vector fields and differential 1-forms), so we may consider a vector field induced by df . It is known as the **gradient (vector field)** and is uniquely determined by $\langle \text{grad } f, X \rangle = df(X)$ for any $X \in \Gamma(TM)$.

An arbitrary connection ∇ on M is called **compatible with the metric** if $X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle$ holds for any vector fields X, Y, Z . This condition is natural, since it is equivalent to parallel transport preserving the inner product. In general, there are many connections on a Riemannian manifold M which are compatible with the given metric, but it is possible to single one out.

Let **torsion** be defined as $T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$, where $[X, Y]$ is a Lie bracket of vector fields, i.e. $[X, Y]$ is a vector field defined via prescription $[X, Y]f := X(Yf) - Y(Xf)$ for any $f \in C^\infty(M)$. Then there exists a unique connection on M compatible with the metric and with vanishing torsion. It is called a **Levi-Civita connection**. From now on we assume that M is equipped with a Levi-Civita connection and we denote it simply by ∇ .

For a vector field $X \in \Gamma(TM)$ the differential ∇X is a (1,1)-tensor determined by $\nabla X(Y) = \nabla_Y X$ for any vector field $Y \in \Gamma(TM)$ (see [Lee97, Ch. 2] for definitions of general tensors and operations on them). Thus, we may define a **divergence** of a vector field X as a trace (contraction) of the (1,1)-tensor ∇X , i.e. $\text{div } X := \text{tr}(\nabla X)$

Hence we may consider the **Laplace-Beltrami operator** Δ_M on a Riemannian manifold M , which is determined by $\Delta_M f := \text{div grad } f$ for $f \in C^\infty(M)$. A process X is called a **Brownian motion on a Riemannian manifold** M if it is a diffusion generated by $\Delta_M/2$. We now present a particularly nice description of Riemannian Brownian motion, known as the Eells-Elworthy-Malliavin construction, which can be achieved using the theory from Sections 1.3 and 1.4.

The metric induces the inner product on each tangent space and allows us to consider **orthonormal frames** i.e. frames $u: \mathbb{R}^d \rightarrow T_p M$ which are linear isometries, where \mathbb{R}^d is equipped with the standard Euclidean metric. Analogously to the frame bundle we construct an **orthonormal frame bundle** $O(M)$ which is a principal frame bundle with a structure group $O(d)$, the group of orthogonal matrices in $\mathbb{R}^{d \times d}$. An orthonormal frame bundle is a subbundle of a full frame bundle. Compatibility with the metric implies that the fundamental horizontal vector fields corresponding to the Levi-Civita connection are tangent to the orthonormal frame bundle $O(M)$, which allows us to interpret them also as vector fields on $O(M)$. Therefore, a horizontal process with an initial point on the orthonormal frame bundle stays on the orthonormal frame bundle for all its lifetime. Addition-

ally, the anti-development of a manifold valued semimartingale can be considered and the anti-development of the Riemannian Brownian motion is exactly the standard Euclidean Brownian motion [Hsu02, Prop. 3.2.1]. More precisely, the Riemannian Brownian motion is constructed by first considering a process U solving SDE (1.7) with $U_0 \in O(M)$ where W is the Euclidean Brownian motion. Such a process U is known as a **horizontal Brownian motion** and its projection $X := \pi(U)$ is the Brownian motion on a Riemannian manifold M .

The above construction is pleasing due to its intrinsic nature, but it is still quite indirect. Moreover, manifolds are often introduced extrinsically as submanifolds in a Euclidean space \mathbb{R}^N for some dimension N . If the Riemannian manifold M is isometrically embedded in \mathbb{R}^N , i.e. the metric on M coincides with the one induced by the standard metric on the ambient space \mathbb{R}^N , then the Brownian motion on M can be constructed in a more direct manner. Let e_1, \dots, e_N be the standard basis of \mathbb{R}^N . For each $x \in M$ let $P_\alpha(x)$ denote the orthogonal projection of the vector e_α onto the tangent space $T_x M \subseteq \mathbb{R}^N$, making P_α a vector field on M for each $\alpha = 1, \dots, N$. Then [Hsu02, Thm. 3.1.4] shows that $\sum_{\alpha=1}^N P_\alpha^2 = \Delta_M$. Therefore, if we consider a solution X of the SDE

$$dX_t = P_\alpha(X_t) \circ dB_t^\alpha, \quad (1.8)$$

where $X_0 \in M$ and B is an \mathbb{R}^N -valued Brownian motion, it is exactly the Brownian motion on the Riemannian manifold M .

In the following two chapters the central object of study will be the Brownian motion on a unit sphere $\mathbb{S}^{d-1} := \{z \in \mathbb{R}^d ; |z| = 1\} \subseteq \mathbb{R}^d$ equipped with a metric induced by the ambient space. In this setting SDE (1.8) simplifies significantly since the orthogonal projection onto $T_z \mathbb{S}^{d-1}$ is given via multiplication by the matrix $I_d - zz^\top$, where I_d is the identity matrix of dimension d and $^\top$ denotes transposition. Hence the Brownian motion Z on a sphere \mathbb{S}^{d-1} is given as a solution of a Stratonovich SDE

$$dZ_t = (I_d - Z_t Z_t^\top) \circ dB_t. \quad (1.9)$$

This is known as a Stroock representation of the spherical Brownian motion.

Chapter 2

Spherical Brownian motion and its projections

In this chapter based on [MMU18] we study a process consisting of the first n coordinates of a Brownian motion on the unit sphere in $\mathbb{R}^{n+\ell}$ and obtain the SDE it satisfies. The SDE has non-Lipschitz coefficients but we are able to provide an analysis of existence and pathwise uniqueness and show that they always hold. The square of the radial component turns out to be a Wright-Fisher diffusion process with mutation and it features in a skew-product decomposition of the projected spherical Brownian motion. We also study a more general SDE on the unit ball in \mathbb{R}^n which allows us to geometrically realize the Wright-Fisher diffusion with general non-negative parameters as the radial component of its solution.

2.1 Introduction and main results

The Theorem of Archimedes [ArcBC] states that the projection π_1 from the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ to any coordinate (in \mathbb{R}^3) preserves the uniform distribution; see [AM13] and the references therein for a very modern account or [PR12] for one with more probabilistic insight. In probabilistic language, if $U^{(2)}$ is a uniform random vector on \mathbb{S}^2 , then $\pi_1(U^{(2)})$ is uniform on $[-1,1]$. In fact, this holds in any dimension $d \geq 3$: for a uniform random vector $U^{(d-1)}$ on the Euclidean unit sphere $\mathbb{S}^{d-1} := \{z \in \mathbb{R}^d ; |z| = 1\}$, its projection $\pi_{d-2}(U^{(d-1)})$ onto any $d - 2$ coordinates is uniform on the unit ball $\mathbb{B}^{d-2} := \{z \in \mathbb{R}^{d-2} ; |z| \leq 1\}$. A more general version of this result for spheres in the p -norm can be found in [BGMN05].

Since the uniform distribution on \mathbb{S}^{d-1} is the invariant measure for the Brownian

motion on the sphere, it is natural to investigate the process obtained by projecting it to the ball \mathbb{B}^{d-2} . Such a process ought to have a uniform distribution on \mathbb{B}^{d-2} as its invariant measure. The aim of this chapter is to give a complete characterization of such processes in terms of SDEs they satisfy and to deduce certain structural consequences of this characterization.

Let Z be the Brownian motion on the sphere $\mathbb{S}^{n+\ell-1}$ ($n, \ell \in \mathbb{N}$). We will use the Stroock representation in Itô form (see also (1.9)) and consider Z as the solution of an equation

$$dZ_t = (I_{n+\ell} - Z_t Z_t^\top) d\tilde{B}_t - \frac{n+\ell-1}{2} Z_t dt, \quad Z_0 \in \mathbb{S}^{n+\ell-1}, \quad (2.1)$$

where \tilde{B} is a Brownian motion on $\mathbb{R}^{n+\ell}$, $I_{n+\ell}$ denotes the identity matrix of dimension $n+\ell$ and $^\top$ denotes transposition.

Proposition 2.1.1. *Let X denote the first n coordinates of Z . Then, there exists a Brownian motion B on \mathbb{R}^n such that the pair (X, B) satisfies the SDE*

$$dX_t = \sigma(X_t) dB_t - \frac{n+\ell-1}{2} X_t dt, \quad X_0 \in \mathbb{B}^n,$$

where the volatility matrix $\sigma(x) \in \mathbb{R}^n \otimes \mathbb{R}^n$ takes the form

$$\sigma(x) = I_n - \left(1 - \sqrt{1 - |x|^2}\right) \frac{xx^\top}{|x|^2} \mathbf{1}(|x| > 0), \quad x \in \mathbb{B}^n.$$

Pathwise uniqueness holds for this SDE and X is a strong Markov process with a unique invariant measure which admits the density

$$h(x) = \frac{\Gamma((n+\ell)/2)}{\pi^{n/2} \Gamma(\ell/2)} \left(1 - |x|^2\right)^{(\ell-2)/2} \mathbf{1}(|x| \leq 1).$$

Furthermore, $U = |X|^2$ is a Wright-Fisher diffusion, i.e. there exists a scalar Brownian motion β such that the pair (U, β) satisfies the SDE

$$dU_t = 2\sqrt{U_t(1-U_t)} d\beta_t + [n(1-U_t) - \ell U_t] dt.$$

Remark 2.1.2. In the case $\ell = 2$, the invariant measure of X in Proposition 2.1.1 is uniform measure on the unit ball \mathbb{B}^n , as expected from the theorem of Archimedes.

Remark 2.1.3. The process X enjoys a skew-product decomposition analogous to the one of Brownian motion in \mathbb{R}^n ; it is a particular case of Theorem 2.1.5 below.

Remark 2.1.4. Since $\sigma(x)$ is the unique non-negative definite square root of the matrix

$I_n - xx^\top$ for $x \in \mathbb{B}^n$, the SDE for the projected process X implies that its infinitesimal generator equals

$$\frac{1}{2} \sum_{i,j=1}^n (\delta_{ij} - x^i x^j) \frac{\partial^2}{\partial x^i \partial x^j} - \sum_{i=1}^n \frac{n-1+\ell}{2} \frac{\partial}{\partial x^i}.$$

This operator equals the generator $\frac{1}{2}\Delta_{n-1+\ell,n}$ of a process considered in [Bak96]. Our results show that the martingale problem arising from this generator is well-posed, at least for initial measures with support contained in \mathbb{B}^n .

More generally, and in order to state the skew-product decomposition, we will consider the SDE on the unit ball \mathbb{B}^n given by

$$dX_t = \gamma(|X_t|)\sigma(X_t)dB_t - g(|X_t|)X_t dt, \quad (2.2)$$

with a starting point $X_0 = x_0 \in \mathbb{B}^n$. Assume that $\gamma: [0,1] \rightarrow (0,\infty)$ and $g: [0,1] \rightarrow \mathbb{R}$ are Lipschitz continuous¹ and satisfy $\frac{g(1)}{\gamma^2(1)} \geq \frac{n-1}{2}$. In particular, the above SDE extends the projected process X of Proposition 2.1.1 to non-integer dimensions. The coefficients of SDE (2.2) are bounded and continuous implying that weak existence of the solutions holds [IW89, Ch. IV, Thm 2.2]. Technically, the theorem can only be applied to SDEs on the whole Euclidean space \mathbb{R}^n , so we temporarily consider continuous extension of functions γ, g, σ to the whole \mathbb{R}^n by prescribing $\gamma(r) = \gamma(1), g(r) = g(1)$ for $r \geq 1$ and $\sigma(x) = \sigma(x/|x|) = I_n - xx^\top / |x|^2$ for x with $|x| \geq 1$. However, the condition $\frac{g(1)}{\gamma^2(1)} \geq \frac{n-1}{2}$ turns out to be necessary and sufficient for a solution to stay in the unit ball² (see Lemma 2.2.1). Therefore, when a solution is started in the unit ball, it stays inside. Hence the extensions of functions γ, g, σ were not actually necessary and we can indeed study SDE (2.2) on the unit ball.

The function σ is locally Lipschitz only on the interior of the ball \mathbb{B}^n , making it impossible to apply the classical theory for the uniqueness of solutions of SDEs (note that we do not exclude the cases when either $X_0 \in \mathbb{S}^{n-1}$ or the boundary sphere is reached in finite time).

The radial component $R := |X|$ of a solution X of SDE (2.2) and its square $U := |X|^2$ are the unique strong solutions of the respective SDEs in (2.6) and (2.4) below; the latter reduces to the SDE of a Wright-Fisher diffusion with mutation in the setting of Proposition 2.1.1. In particular, both processes are strong Markov. After a time-change, a pathwise comparison of U with a Wright-Fisher diffusion implies that for $n \geq 2$ the

¹Note that by Lipschitz continuity of the euclidean norm $|\cdot|$, functions $\gamma(|\cdot|)$ and $g(|\cdot|)$ are also Lipschitz.

²In fact, if $\frac{g(1)}{\gamma^2(1)} = \frac{n-1}{2}$ and $X_0 \in \mathbb{S}^{n-1}$, the solution X is a Brownian motion on \mathbb{S}^{n-1} time-changed by $t \mapsto \gamma^2(1)t$.

process X never hits 0 (see Lemma 2.2.1 in Section 2.2 below). This enables us to define a time-change process $S_s(t) := \int_s^t \frac{\gamma^2(R_u)}{R_u^2} du$, satisfying $\lim_{t \rightarrow \infty} S_s(t) = \infty$, and its inverse $T_s: [0, \infty) \rightarrow [s, \infty)$ (see Lemma 2.2.3 below). Moreover, it turns out that X possesses a skew-product decomposition analogous to the one of Brownian motion on \mathbb{R}^n .

Theorem 2.1.5 (Skew-product decomposition). *Let $n \geq 2$ and X be a solution of SDE (2.2). Pick $s \in \mathbb{R}_+ := [0, \infty)$ and assume that either $s > 0$ or $s = 0$ and $X_0 \neq 0$. Then the process $\widehat{V} = (\widehat{V}_t)_{t \in \mathbb{R}_+}$, given by $\widehat{V}_t := X_{T_s(t)}/R_{T_s(t)}$, is a Brownian motion on \mathbb{S}^{n-1} (started at $\widehat{V}_0 = X_s/R_s$) independent of R . Hence we obtain the skew-product decomposition $X_t = R_t \widehat{V}_{S_s(t)}$ for $t \geq s$. Furthermore, if $X_0 = 0$, then \widehat{V}_t is uniformly distributed on \mathbb{S}^{n-1} for any $t > 0$ and subsequently evolves as a stationary Brownian motion on the sphere.*

Since the skew-product decomposition expresses X as a measurable functional of a pair of independent processes (R, \widehat{V}) with given distributions, the following result holds.

Corollary 2.1.6. *Uniqueness in law holds for SDE (2.2).*

The pathwise uniqueness of SDE (2.2) is more delicate because σ is not Lipschitz at the boundary of the ball \mathbb{B}^n . Since σ is locally Lipschitz, pathwise uniqueness holds up to the first hitting time of the boundary by well established arguments. Hence, if $\frac{g(u)}{\gamma^2(u)} \geq \frac{n-1}{2} + 1$ holds for u sufficiently close to one, $|X_0| < 1$, then X never visits the boundary of \mathbb{B}^n and pathwise uniqueness holds. If $\frac{g(1)}{\gamma^2(1)} = \frac{n-1}{2}$, X behaves as time-changed Brownian motion on \mathbb{S}^{n-1} after the first time it hits the boundary and hence pathwise uniqueness also holds in this case. If $n = 1$, the SDE (2.2) simplifies to $dX_t^1 = \gamma(|X_t^1|) \sqrt{1 - (X_t^1)^2} dB_t - g(|X_t^1|) X_t^1 dt$ and pathwise uniqueness holds by a theorem of Yamada and Watanabe [RY99, IX.3.5]. Pathwise uniqueness of the similar looking equation $dX_t = \gamma(|X_t|)(1 - |X_t|^2)^{1/2} dB_t - g(|X_t|) X_t dt$ on \mathbb{B}^n , where γ and g are positive Lipschitz functions and $\frac{g(1)}{\gamma^2(1)}$ is sufficiently large, was established by DeBlassie [DeB04] by a clever generalization of the idea in [Swa02]. Note that the diffusion coefficient in the SDE (2.2) depends on x and not just on its length $|x|$, making it impossible to apply the result of [DeB04]. However, it is possible to adapt the method of [DeB04] to our setting and obtain:

Theorem 2.1.7. *If $\frac{g(1)}{\gamma^2(1)} - \frac{n-1}{2} > \sqrt{2} - 1 \doteq 0.4142$, then pathwise uniqueness holds for the SDE (2.2).*

The remaining cases, when $n \geq 2$ and $\frac{g(1)}{\gamma^2(1)} - \frac{n-1}{2} \in (0, \sqrt{2} - 1]$, are left open.

2.2 Characterization of the projected process

We are interested in the process consisting of the first n coordinates of Brownian motion on the sphere $\mathbb{S}^{n-1+\ell}$. One way of constructing such a Brownian motion is via the Stroock representation, i.e. a solution to the SDE (2.1). Note that the coefficients of SDE (2.1) are locally Lipschitz continuous, so pathwise uniqueness holds.

Recall that superscripts denote components of vectors, e.g. X^i means i -th component of process X . When we wish to express powers we will enclose variables in additional sets of parentheses.

Proof of Proposition 2.1.1. Let X' denote last ℓ coordinates of Z and similarly split $\tilde{B} = (\tilde{B}^1, \tilde{B}^2)$. We claim that the process B given by the equation

$$B_t = \int_0^t \sigma(X_s) d\tilde{B}_s^1 + \int_0^t \left(-\frac{X_s X_s'^\top}{(1 - |X_s|^2)^{1/2}} \mathbf{1}(|X_s| < 1) + X_s z^\top \mathbf{1}(|X_s| = 1) \right) d\tilde{B}_s^2,$$

where $z \in \mathbb{S}^{\ell-1}$ is an arbitrary (fixed) unit vector, is an n -dimensional Brownian motion.

Note that $|X_t'|^2 = |Z_t|^2 - |X_t|^2 = 1 - |X_t|^2$. Furthermore, let us consider the $n \times (n + \ell)$ matrix A_t given in block form by

$$A_t := \left[\sigma(X_t), D_t \right],$$

where the $n \times \ell$ matrix D_t is given by

$$D_t := -X_t X_t'^\top (1 - |X_t|^2)^{-1/2} \mathbf{1}(|X_t| < 1) + X_t z^\top \mathbf{1}(|X_t| = 1).$$

We can compute $\sigma(X_t)^2 = \sigma(X_t)\sigma(X_t)^\top = I_n - X_t X_t^\top$ and $D_t D_t^\top = X_t X_t^\top$, so it follows that $A_t A_t^\top = \sigma(X_t)\sigma(X_t)^\top + D_t D_t^\top = I_n$. Since B is defined by $B_t = \int_0^t A_s d\tilde{B}_s$ and it is a continuous local martingale with quadratic variation $\langle B^i, B^j \rangle_t = \int_0^t (A_s A_s^\top)_{ij} ds = \delta_{ij} t$, it is an n -dimensional Brownian motion by Levy's characterization theorem (see Theorem A.2.5). Further calculations show that $\sigma(X_t) D_t = -X_t X_t'^\top$. Combining the aforementioned facts along with the definition of Z , we get the SDE satisfied by X :

$$\begin{aligned} dX_t &= (I_n - X_t X_t^\top) d\tilde{B}_t^1 - X_t X_t'^\top d\tilde{B}_t^2 - \frac{n-1+\ell}{2} X_t dt \\ &= \sigma(X_t)^2 d\tilde{B}_t^1 + \sigma(X_t) D_t d\tilde{B}_t^2 - \frac{n-1+\ell}{2} X_t dt \\ &= \sigma(X_t) A_t d\tilde{B}_t - \frac{n-1+\ell}{2} X_t dt = \sigma(X_t) dB_t - \frac{n-1+\ell}{2} X_t dt. \end{aligned}$$

The above SDE is just a special case of SDE (2.2) with $\gamma \equiv 1$ and $g \equiv \frac{n-1+\ell}{2}$, therefore pathwise uniqueness holds immediately by Theorem 2.1.7 since $\frac{g(1)}{\gamma^2(1)} - \frac{n-1}{2} = \frac{\ell}{2} > \sqrt{2} - 1$ for $\ell \in \mathbb{N}$. Consequently X is a strong Markov process. Furthermore, Lemma 2.2.1 shows that $U = |X|^2$ is a Wright-Fisher diffusion with mutation rates n and ℓ . When $n \geq 2$ this also helps us find invariant measure for the process X since we can use the skew-product decomposition from Theorem 2.1.5. The invariant measure for the Wright-Fisher diffusion U is given by Beta($n/2, \ell/2$) distribution. Hence $g(r) = B_{n,\ell} r^{n-1} (1-r^2)^{(\ell-2)/2} \mathbf{1}(r \in [0,1])$ is the density of the invariant measure of R . The invariant measure for the Brownian motion on a sphere is a normalised uniform measure. This continues to hold for the time-changed Brownian motion on a sphere as long as the time change is independent of the Brownian motion. So let us suppose that the initial distribution of the process X has the density h from Proposition 2.1.1. Then, using polar coordinates and the skew-product decomposition, the density of R_0 is g . Since this density is invariant for R , R_t has density g for all $t \geq 0$. The time-changed Brownian motion on a sphere also remains uniformly distributed and reversing polar coordinates we get that X_t has density h for any t .

In the case $n = 1$ the process X^1 satisfies the SDE $dX_t^1 = \sqrt{1 - (X_t^1)^2} dB_t - \frac{\ell}{2} X_t^1 dt$, and by the forward Kolmogorov equation the invariant density can easily be seen to be equal to

$$h(x) = \frac{\Gamma((1+\ell)/2)}{\pi^{1/2} \Gamma(\ell/2)} (1-x^2)^{(\ell-2)/2} \mathbf{1}(|x| \leq 1). \quad \square$$

2.2.1 Skew-product decomposition for SDE (2.2)

The solution of (2.2) naturally lives on the closed unit ball \mathbb{B}^n . To better understand such a process it is crucial to understand its radial component (or its square) and in particular if and when it hits the boundary (or zero). The square of the radial component will be shown to be closely related to a Wright-Fisher diffusion so let us first recall some known facts about the latter (see also Section 3.3). Fix non-negative parameters α and β . The SDE

$$dZ_t = 2\sqrt{Z_t(1-Z_t)} dB_t + (\alpha(1-Z_t) - \beta Z_t) dt, \quad Z_0 = x \in [0,1] \quad (2.3)$$

has a unique strong solution which is called the Wright-Fisher diffusion; denote it by $WF_x(\alpha, \beta)$. It is possible to define $WF_x(\alpha, \beta)$ for negative α (resp. β), but then the solution has a finite lifetime equal to the first hitting time of 0 (resp. 1). Pathwise uniqueness is a consequence of 1/2-Hölder continuity of the diffusion coefficient and applying theorem of Yamada and Watanabe [RY99, IX.3.5]. Furthermore, it is well-

known³ that $\text{WF}_x(\alpha, \beta)$ never hits 0 for $\alpha \geq 2$ when $x > 0$ and never hits 1 for $\beta \geq 2$ when $x < 1$.

Some of the facts about the square of the radial component of the solution to SDE (2.2) are summarized in the following lemma.

Lemma 2.2.1. *Let X be a solution of (2.2) where $\frac{g(1)}{\gamma^2(1)} \geq \frac{n-1}{2}$. Then the process $U = |X|^2$ satisfies the SDE*

$$\begin{aligned} dU_t &= 2\tilde{\gamma}(U_t)\sqrt{U_t(1-U_t)}d\theta_t \\ &\quad + \tilde{\gamma}^2(U_t) \left(n(1-U_t) - \left(\frac{2\tilde{g}(U_t)}{\tilde{\gamma}^2(U_t)} - (n-1) \right) U_t \right) dt, \end{aligned} \quad (2.4)$$

where $\tilde{g}(u) := g(\sqrt{u})$, $\tilde{\gamma}(u) := \gamma(\sqrt{u})$, and the \mathcal{F}_t -Brownian motion θ is defined by

$$\theta_t = \sum_{i=1}^n \int_0^t \frac{X_s^i}{\sqrt{U_s}} \mathbf{1}(U_s > 0) dB_s^i + \int_0^t \mathbf{1}(U_s = 0) d\chi_s,$$

where the scalar Brownian motion χ is independent of B .

For $n \geq 2$, the process U never hits 0 and if $\frac{g(u)}{\gamma^2(u)} \geq \frac{n-1}{2} + 1$ holds near 1, then U never hits 1.

Proof. Since $U = |X|^2$ and $d\langle X^i \rangle_t = \gamma^2(|X_t|)(1 - (X_t^i)^2)dt$ we can apply Itô's formula and use $x^\top \sigma(x) = \sqrt{1 - |x|^2}x^\top$ to get

$$\begin{aligned} dU_t &= 2 \sum_{i=1}^n X_t^i dX_t^i + \frac{1}{2} \cdot 2 \sum_{i=1}^n d\langle X^i \rangle_t \\ &= 2\gamma(|X_t|) \sum_{i=1}^n X_t^i \sqrt{1 - U_t} dB_t^i - 2g(|X_t|) \sum_{i=1}^n (X_t^i)^2 dt + \gamma^2(|X_t|) \sum_{i=1}^n (1 - (X_t^i)^2) dt \end{aligned}$$

which in turn yields (2.4). Moreover, since $\langle \theta \rangle_t = t$, the continuous local martingale θ is a Brownian motion by Levy's characterization theorem.

We can slightly simplify the equation (2.4) by time-change without affecting the boundary hitting properties. Define q by

$$q_t := \int_0^t \gamma^2(|X_u|) du \quad \text{and its inverse} \quad \tilde{q}_t := \inf \{ u \geq 0 ; q(u) = t \}.$$

³It can for example be seen from Lemma 2.2.2 and well-known facts about hitting of 0 of Bessel processes.

Then $\widehat{U}_t := U_{\widetilde{q}_t}$ satisfies the SDE

$$d\widehat{U}_t = 2\sqrt{\widehat{U}_t(1-\widehat{U}_t)}d\widetilde{\theta}_t + \left(n(1-\widehat{U}_t) - \left(\frac{2\widetilde{g}(\widehat{U}_t)}{\widetilde{\gamma}^2(\widehat{U}_t)} - (n-1) \right) \widehat{U}_t \right) dt \quad (2.5)$$

where $\widetilde{\theta}_t = \int_0^{\widetilde{q}_t} \gamma(|X_u|)d\theta_u$ is also a Brownian motion. This is almost the same equation as that of a Wright-Fisher diffusion. The volatility term is exactly the same so that the difference appears only in the drift term, which is, nevertheless, still Lipschitz continuous⁴. Again we can use the Yamada-Watanabe Theorem [RY99, Theorem 3.5 in Ch. IX] and conclude that pathwise uniqueness holds for the SDE (2.5). Since $\frac{g(1)}{\gamma^2(1)} \geq \frac{n-1}{2}$ we can consider SDE (2.5) with the increased drift

$$n(1-u) - \left(\frac{2\widetilde{g}(u)}{\widetilde{\gamma}^2(u)} - \frac{2\widetilde{g}(1)}{\widetilde{\gamma}^2(1)} \right) u$$

and for $\widehat{U}_0 = 1$ such an equation has the unique solution $\widehat{U}_t \equiv 1$. By using the comparison theorem for SDEs, as in [RY99, Ch. IX, (3.7)], we deduce that the inequality $\widehat{U}_t \leq 1$ holds for the solution of (2.5) as long as $\widehat{U}_0 \leq 1$ a.s. By denoting

$$M := \max_{u \in [0,1]} \left(\frac{2\widetilde{g}(u)}{\widetilde{\gamma}^2(u)} - (n-1) \right) \geq 0,$$

another application of the comparison theorem shows that the solution of (2.5) is always larger than $WF_x(n, M)$ started at the same point, so it is always non-negative and therefore $U_t \in [0,1]$ for $U_0 \in [0,1]$ a.s. which also means it has infinite lifetime. This second use of comparison theorem also shows, that for $n \geq 2$, the processes \widehat{U} and U never hit 0 unless they start there. Similarly, if $\frac{g(u)}{\gamma^2(u)} \geq \frac{n-1}{2} + 1$ holds near 1, we could at least locally (near 1) use the comparison theorem to show that \widehat{U} is smaller than $WF_x(n, 2)$ started at the same point and therefore under such conditions \widehat{U} (and U) never hits 1. \square

Before we can prove Theorem 2.1.5 we need two additional lemmas. The first one expresses a Wright-Fisher diffusion as a certain skew-product of two independent squared Bessel processes. For notation and basic facts about (squared) Bessel processes see [RY99, Chapter XI].

Lemma 2.2.2. *Let $\alpha \geq 0, \beta \in \mathbb{R}$ and let \mathcal{X} be a $BESQ^\alpha$ process started at $x_0 \geq 0$ and \mathcal{Y} be an independent $BESQ^\beta$ process started at $y_0 \geq 0$ such that $x_0 + y_0 > 0$. Let*

⁴The function $u \mapsto uf(\sqrt{u})$ is Lipschitz if f is. Hence, $u \mapsto u\frac{\widetilde{g}(u)}{\widetilde{\gamma}^2(u)}$ is Lipschitz continuous.

$T_0(\mathcal{Y}) := \inf \{t \geq 0 ; \mathcal{Y}_t = 0\}$. Define the continuous additive functional ρ as

$$\rho_t = \int_0^t \frac{1}{\mathcal{X}_u + \mathcal{Y}_u} du \text{ and its inverse } \zeta_t = \inf \{u \geq 0 ; \rho_u = t\}.$$

Let $U_t := \frac{\mathcal{X}_t}{\mathcal{X}_t + \mathcal{Y}_t}$ for $t < T_0(\mathcal{Y})$. Then $\widehat{U}_t := U_{\zeta_t}$ for $t < \zeta_{T_0(\mathcal{Y})}$ is a Wright-Fisher diffusion $\text{WF}_{x_0/(x_0+y_0)}(\alpha, \beta)$ and it is independent of $\mathcal{X} + \mathcal{Y}$.

Proof. The lemma is proved in [WY98, Proposition 8] for non-negative coefficients α, β and a larger stopping time $T_0(\mathcal{X} + \mathcal{Y}) = \inf \{t \geq 0 ; \mathcal{X}_t + \mathcal{Y}_t = 0\}$, where they use the term Jacobi diffusion for particular Wright-Fisher diffusions. The same proof works for $\beta < 0$, since $T_0(\mathcal{Y}) < T_0(\mathcal{X} + \mathcal{Y})$ and hence on the stochastic interval $[0, T_0(\mathcal{Y}))$ all the processes in the proof of [WY98, Proposition 8] are well defined and all the calculations stay exactly the same. \square

Our next lemma summarizes facts about the time-change used in Theorem 2.1.5.

Lemma 2.2.3. *Let $n \geq 2$ and either $s > 0$ or $s = 0$ with $X_0 \neq 0$. Define*

$$S_s(t) := \int_s^t \frac{\gamma^2(R_u)}{R_u^2} du.$$

Then $S_s: [s, \infty) \rightarrow \mathbb{R}_+$ is continuous, strictly increasing, and $\lim_{t \rightarrow \infty} S_s(t) = \infty$. Its right continuous inverse $T_s: \mathbb{R}_+ \rightarrow [s, \infty)$ is also continuous and strictly increasing with $T_s(0) = s$. Furthermore, if $X_0 = 0$, then $\lim_{s \downarrow 0} S_s(t) = \infty$ holds for any $t > 0$.

Proof. Since a.s. $R_t^2 \in (0, 1]$ for any $t \geq s$, R_t is continuous in t , and also $\gamma(R_t)$ is continuous in t and bounded away from 0, everything but the last claim follows from classical results on inverses of strictly increasing continuous functions. Since $\lim_{s \downarrow 0} \int_s^t \frac{\gamma^2(R_u)}{R_u^2} du = \int_0^t \frac{\gamma^2(R_u)}{R_u^2} du$, in order to prove the last claim it is enough to prove that $\int_0^t \frac{\gamma^2(R_u)}{R_u^2} du = \infty$ for any $t > 0$. Changing variables with the time change q_t from Lemma 2.2.1 we get

$$\int_0^t \frac{\gamma^2(R_u)}{R_u^2} du = \int_0^t \widehat{U}_{q_u}^{-1} dq_u = \int_0^{q_t} \widehat{U}_u^{-1} du.$$

By the comparison theorem for SDEs and strict positivity of q_t it is therefore enough to prove that $\int_0^{t \wedge T_1(\mathcal{U})} \mathcal{U}_u^{-1} du = \infty$ for any $t > 0$, where \mathcal{U} is Wright-Fisher(n, m) diffusion started at 0,

$$m := \min_{u \in [0, 1]} \left(\frac{2\tilde{g}(u)}{\tilde{\gamma}^2(u)} - (n - 1) \right)$$

is a possibly negative constant, and $T_1(\mathcal{U}) = \inf \{t \geq 0 ; \mathcal{U}_t = 1\}$ is strictly positive. By Lemma 2.2.2 we can find a BESQ n process \mathcal{X} started at 0 and a BESQ m process \mathcal{Y} not started at 0 such that $\mathcal{U}_t = \frac{\mathcal{X}_{\zeta_t}}{\mathcal{X}_{\zeta_t} + \mathcal{Y}_{\zeta_t}}$ holds for $t < T_1(\mathcal{U}) = \rho_{T_0(\mathcal{Y})}$ with the time changes ζ and ρ defined as in Lemma 2.2.2. Using the time change formula for the Lebesgue-Stieltjes integral we get

$$\int_0^{t \wedge T_1(\mathcal{U})} \mathcal{U}_t^{-1} du = \int_0^{t \wedge T_1(\mathcal{U})} \frac{\mathcal{X}_{\zeta_u} + \mathcal{Y}_{\zeta_u}}{\mathcal{X}_{\zeta_u}} du = \int_0^{t \wedge T_1(\mathcal{U})} \mathcal{X}_{\zeta_u}^{-1} d\zeta_u = \int_0^{\zeta_t \wedge \zeta_{T_1(\mathcal{U})}} \mathcal{X}_u^{-1} du.$$

This corresponds to the integral $\int_0^{\zeta_t \wedge \zeta_{T_1(\mathcal{U})}} h(\sqrt{\mathcal{X}_u}) du$ where $h(x) = x^{-2}$ and $\sqrt{\mathcal{X}}$ is a Bessel process started at 0 with parameter $n \geq 2$. Since $\zeta_{T_1(\mathcal{U})} = T_0(\mathcal{Y}) > 0$, $\zeta_t > 0$ for any $t > 0$ and $-2 \leq -(2 \wedge n)$, Corollary 2.4 in [Che01] implies that the integral $S_0(t)$ diverges. \square

Proof of Theorem 2.1.5. Let us define a function $\kappa(x^1, \dots, x^n) := (\sum_{i=1}^n (x^i)^2)^{1/2}$. Then $R_t = \kappa(X_t)$, $V_t^i = \frac{\partial \kappa}{\partial x^i}(X_t)$ and a straightforward computation shows that

$$\begin{aligned} \frac{\partial^2 \kappa}{\partial x^i \partial x^j}(x) &= r^{-3}(\delta_{ij} r^2 - x^i x^j), \\ \frac{\partial^3 \kappa}{\partial x^i \partial x^j \partial x^k}(x) &= 3r^{-5} x^i x^j x^k - r^{-3} (x^i \delta_{jk} + x^j \delta_{ik} + x^k \delta_{ij}), \end{aligned}$$

where $r := \kappa(x)$. Note that also $d\langle X^i, X^j \rangle_t = \gamma^2(|X_t|)(\delta_{ij} - X_t^i X_t^j) dt$. Now we can use the Itô's formula to get equations for R_t and V_t . First,

$$\begin{aligned} dR_t &= \sum_{i=1}^n \frac{X_t^i}{R_t} dX_t^i + \frac{1}{2} \sum_{i,j=1}^n \frac{\delta_{ij} R_t^2 - X_t^i X_t^j}{R_t^3} d\langle X^i, X^j \rangle_t \\ &= \gamma(|X_t|) \sum_{j=1}^n \frac{X_t^j}{R_t} \sqrt{1 - R_t^2} dB_t^j - g(|X_t|) R_t dt \\ &\quad + \frac{\gamma^2(|X_t|)}{2} \sum_{i,j=1}^n \frac{\delta_{ij} R_t^2 - X_t^i X_t^j}{R_t^3} (\delta_{ij} - X_t^i X_t^j) dt \\ &= \gamma(R_t) \sum_{j=1}^n V_t^j \sqrt{1 - R_t^2} dB_t^j + \frac{(n-1)\gamma^2(R_t) - 2g(R_t)R_t^2}{2R_t} dt. \end{aligned}$$

Since $d\hat{t} = \sum_{j=1}^n V_t^j dB_t^j$ is also 1-dimensional Brownian motion (actually, it is the same

process as in Lemma 2.2.1) we write the above equation in a more compact way as

$$dR_t = \gamma(R_t) \sqrt{1 - R_t^2} d\theta_t + ((n-1)\gamma^2(R_t) - 2g(R_t)R_t^2) / (2R_t) dt. \quad (2.6)$$

We can also write the equation for V_t^i .

$$dV_t^i = \sum_{j=1}^n \frac{\delta_{ij}R_t^2 - X_t^i X_t^j}{R_t^3} dX_t^j + \frac{1}{2} \sum_{j,k=1}^n \left(\frac{\partial^3 \kappa}{\partial x^i \partial x^j \partial x^k} (X_t) \right) d\langle X^j, X^k \rangle_t$$

Similar calculations as above show that the first sum is equal to

$$\gamma(R_t) \sum_{j=1}^n \frac{\delta_{ij} - V_t^i V_t^j}{R_t} dB_t^j$$

where the drift part vanishes and we compute the second sum as

$$\begin{aligned} & \frac{\gamma^2(R_t)}{2} R_t^{-5} \sum_{j,k=1}^n \left(3X_t^i X_t^j X_t^k - (X_t^i \delta_{jk} + X_t^j \delta_{ik} + X_t^k \delta_{ij}) R_t^2 \right) (\delta_{jk} - X_t^j X_t^k) dt = \\ & = \frac{\gamma^2(R_t)}{2} R_t^{-5} \left(3X_t^i R_t^2 - (nX_t^i + X_t^i + X_t^i) R_t^2 - 3X_t^i R_t^4 + (X_t^i + X_t^i + X_t^i) R_t^4 \right) dt \\ & = -\gamma(R_t) \frac{n-1}{2} \frac{X_t^i}{R_t^3} dt = -\frac{\gamma(R_t)}{R_t^2} \frac{n-1}{2} V_t^i dt. \end{aligned}$$

Altogether, in vector form, we get $dV_t = \frac{\gamma(R_t)}{R_t} (I_n - V_t V_t^\top) dB_t - \frac{\gamma^2(R_t)}{R_t^2} \cdot \frac{n-1}{2} V_t dt$. To show that the process $\widehat{V}_t = V_{T_t}$ is Brownian motion on a sphere it is sufficient to show that it satisfies SDE (2.1). The hard part is actually showing independence of R , as the identification of Brownian motion will simply follow from the change of time formula for the Itô and the Lebesgue-Stieltjes integrals. To show independence let us enlarge the probability space to accommodate another scalar Brownian motion ξ which is independent of B and define a continuous local martingale $W_t = \int_s^{T_s(t)} \frac{\gamma(R_u)}{R_u} (I_n - V_u V_u^\top) dB_u + \int_s^{T_s(t)} \frac{\gamma(R_u)}{R_u} V_u d\xi_u$. Then we can compute $\langle W^i, W^j \rangle_t = \int_s^{T_s(t)} \delta_{ij} \frac{\gamma^2(R_u)}{R_u^2} du = \delta_{ij} t$, which shows that W is \mathcal{G}_t -Brownian motion, where $\mathcal{G}_t := \mathcal{F}_{T_s(t)}$. Since

$$(I_n - V_u V_u^\top) \cdot \left[\frac{\gamma(R_u)}{R_u} (I_n - V_u V_u^\top), \quad \frac{\gamma(R_u)}{R_u} V_u \right] = \left[\frac{\gamma(R_u)}{R_u} (I_n - V_u V_u^\top), \quad 0 \right]$$

we can use the change of time formula for stochastic and Lebesgue-Stieltjes integral [RY99, Chapter V, §1] and we get $\widehat{V}_t - \widehat{V}_0 = \int_0^t (I_n - \widehat{V}_u \widehat{V}_u^\top) dW_u - \int_0^t \frac{n-1}{2} \widehat{V}_u du$ so that \widehat{V} is a Brownian motion on \mathbb{S}^{n-1} .

To prove independence of \widehat{V} and R it is enough to prove independence of Brownian motions W and θ which are driving the respective SDEs. Then \widehat{V} and R are strong solutions of their corresponding SDEs which are driven by independent Brownian motions, so they are also independent. This holds since we note that for a strong solution X of any SDE there exists a measurable map Φ , such that $X = \Phi(\widetilde{B})$, where \widetilde{B} is Brownian motion driving the SDE c.f. [Che00, YW71]. Therefore we can find measurable maps Φ_1, Φ_2 , such that $\widehat{V} = \Phi_1(W), R = \Phi_2(\theta)$ and independence does indeed follow from the independence of θ and W . The Markov property implies that W depends on $\mathcal{G}_0 = \mathcal{F}_s$ only through $W_0 = 0$, so W is independent of \mathcal{F}_s . Hence W is independent of $(\theta_t)_{t \in [0, s]}$. Therefore, it is enough to prove that W is independent of $(\theta_t - \theta_s)_{t \geq s}$. Define $\eta_t := \theta_{T_s(t)} - \theta_s = \int_s^{T_s(t)} V_u^\top dB_u$ so that η is a \mathcal{G}_t -local martingale. Simple calculation shows that $\langle W_i, \eta \rangle_t = 0$ and $\langle \eta \rangle_t = T_s(t) - s$ with inverse $S_s(t + s)$. We then use Knight's Theorem (also known as the multidimensional Dambis-Dubins-Schwarz Theorem found in [RY99, Chapter V, Theorem 1.9]) to show that W and $(\eta_{S_s(t+s)})_{s \geq t} = (\theta_{s+t} - \theta_s)_{s \geq t}$ are independent Brownian motions.

To address the last statement of the theorem we need to consider the situation when the solution is started from 0. The evolution of such a process is given by $(R_t \phi_{S_s(t)}, t \geq s)$ where R is a square root of a solution to SDE (2.4) and ϕ is an independent Brownian motion on the sphere started at $\phi_0 = X_s/R_s$. Due to *rapid spinning* (i.e. $\lim_{s \downarrow 0} S_s(t) = \infty$), the initial point $\phi_0 = X_s/R_s$ will be forced to be uniformly distributed on the sphere. This follows from the properties of the skew-product decomposition established in this proof, Lemma 2.2.3 above and [GMW19, Lemma 3.12]. \square

Proof of Corollary 2.1.6 . For $n = 1$, pathwise uniqueness holds so uniqueness in law follows trivially. Now let $n \geq 2$. When $x_0 \neq 0$, by Theorem 2.1.5, the solution X is a measurable functional of two independent processes R and \widehat{V} with given laws. Hence the law of X is unique.

Finally, we consider the case of $X_0 = 0$. What follows is an almost direct application of the proof of Theorem 1.1 in [GMW19]. For any $k \in \mathbb{N}$ and open set $U \subseteq \mathbb{R}^k$, define a measurable function $F_U : (0, \infty)^k \rightarrow [0, 1], F_U(t_1, \dots, t_k) := \mathbb{P}_\Psi[(\psi_{t_1}, \dots, \psi_{t_k}) \in U]$, where law $\mathbb{P}_\Psi[\cdot]$ is defined in [GMW19, Lemma 3.7]. Letting $\mathcal{F}_\infty^R := \sigma(R_u, u \in \mathbb{R}_+)$, we apply Lemma 2.2.3 and Theorem 2.1.5 above and [GMW19, Lemma 3.12] to get

$$\mathbb{P}[(X_{t_1}/R_{t_1}, \dots, X_{t_k}/R_{t_k}) \in U | \mathcal{F}_\infty^R] = F_U(S_s(t_1), \dots, S_s(t_k))$$

a.s. for $0 < s < t_1 < \dots < t_k$. Hence

$$\mathbb{P}[(X_{t_1}/R_{t_1}, \dots, X_{t_k}/R_{t_k}) \in U] = \mathbb{E}[F_U(S_s(t_1), \dots, S_s(t_k))]$$

and therefore the finite-dimensional distributions of $(X_t/R_t, t > 0)$ are determined uniquely by $\mathbb{P}_\Psi[\cdot]$ and the law of R . Moreover, law of X is determined uniquely by the law of $(R, X/R)$, therefore uniquely by $\mathbb{P}_\Psi[\cdot]$ and the law of process R solving SDE (2.6) started at 0. \square

2.2.2 Pathwise uniqueness for SDE (2.2)

Let X and \tilde{X} be solutions of SDE (2.2) driven by the same Brownian motion B and started at the same point $x_0 \in \mathbb{B}^n$. Pathwise uniqueness clearly holds up to the hitting time of the boundary, so by restarting argument it is enough to prove pathwise uniqueness for starting points x_0 on the boundary. Furthermore, it is enough to prove that $X_t = \tilde{X}_t$ for $t \leq \tau_\varepsilon$ where $\tau_\varepsilon = \inf \left\{ t \geq 0 ; |X_t|^2 \wedge |\tilde{X}_t|^2 \leq 1 - \varepsilon \right\}$ for some $\varepsilon > 0$ and without loss of generality we can clearly assume that $\varepsilon < \frac{1}{2}$. To prove equality of processes we will apply the method of DeBlassie [DeB04]. Namely, we wish to use Gronwall's lemma, but due to non-Lipschitzness we cannot apply it directly to $\mathbb{E}[|X_t - \tilde{X}_t|^2]$. The idea of DeBlassie (and Swart before with $p = \frac{1}{2}$) is to denote $Y := 1 - |X|^2$, $\tilde{Y} := 1 - |\tilde{X}|^2$ and look at the process $W := |X - \tilde{X}|^2 + (Y^p - \tilde{Y}^p)^2$ for some $p \in (\frac{1}{2}, 1)$. We then have

$$\begin{aligned} dY_t &= -2\gamma(|X_t|)Y_t^{1/2} \sum_{i=1}^n X_t^i dB_t^i - n\gamma^2(|X_t|)Y_t dt \\ &\quad + (2g(|X_t|) - (n-1)\gamma^2(|X_t|)) |X_t|^2 dt. \end{aligned}$$

A slight modification of [DeB04, Lemma 2.1] implies that for $p > 1 + \frac{n-1}{2} - \frac{g(1)}{\gamma^2(1)}$ a formal application of Itô's formula for the mapping $x \mapsto x^p$ is justified. Defining

$$G(u) := g(u) - \frac{n-1}{2}\gamma^2(u) + (p-1)\gamma^2(u),$$

we get

$$\begin{aligned} dY_t^p &= -2p\gamma(|X_t|)Y_t^{p-1/2} \sum_{i=1}^n X_t^i dB_t^i + 2pY_t^{p-1} |X_t|^2 \mathbf{1}(Y > 0)G(|X_t|)dt \\ &\quad - np\gamma^2(|X_t|)Y_t^p dt, \end{aligned}$$

where $t \leq \tau_\varepsilon$, $|X_0|^2 > 1 - \varepsilon$, and $\varepsilon = \varepsilon(p)$ is chosen in such a way that $p > 1 + \frac{n-1}{2} - \frac{g(u)}{\gamma^2(u)}$ holds for $u \in (1 - \varepsilon(p), 1]$. The latter condition is necessary to keep the second term on the right hand side negative to allow the use of Fatou's lemma. Furthermore, $\int_0^t \mathbf{1}(Y_s = 0) ds = 0$ holds. Note that essentially all necessary calculations and results are the same as in [DeB04] if we change their g for $g - \frac{n-1}{2}\gamma^2$. Subtracting the equations for Y^p and \tilde{Y}^p we get

$$\begin{aligned} d(Y^p - \tilde{Y}^p)_t &= -2p \sum_{i=1}^n \left(\gamma(|X_t|) Y_t^{p-1/2} X_t^i - \gamma(|\tilde{X}_t|) \tilde{Y}_t^{p-1/2} \tilde{X}_t^i \right) dB_t^i \\ &\quad + 2p \left(Y_t^{p-1} |X_t|^2 \mathbf{1}(Y > 0) G(|X_t|) - \tilde{Y}_t^{p-1} |\tilde{X}_t|^2 \mathbf{1}(\tilde{Y} > 0) G(|\tilde{X}_t|) \right) dt \\ &\quad - np \left(\gamma^2(|X_t|) Y_t^p - \gamma^2(|\tilde{X}_t|) \tilde{Y}_t^p \right) dt \\ &=: dM_t + \mathbf{I}_1 dt + \mathbf{I}_2 dt \end{aligned}$$

and Itô's formula yields

$$\begin{aligned} d(Y^p - \tilde{Y}^p)_t^2 &= 2(Y_t^p - \tilde{Y}_t^p)(dM_t + \mathbf{I}_1 dt + \mathbf{I}_2 dt) \\ &\quad + 4p^2 \sum_{i=1}^n \left(\gamma(|X_t|) Y_t^{p-1/2} X_t^i - \gamma(|\tilde{X}_t|) \tilde{Y}_t^{p-1/2} \tilde{X}_t^i \right) dt \\ &=: d\tilde{M}_t + 2(Y_t^p - \tilde{Y}_t^p)(\mathbf{I}_1 dt + \mathbf{I}_2 dt) + \mathbf{I}_3 dt. \end{aligned}$$

We can also compute

$$\begin{aligned} d|X_t - \tilde{X}_t|^2 &= 2 \sum_{i=1}^n (X_t^i - \tilde{X}_t^i) \sum_{j=1}^n \left(\gamma(|X_t|) \sigma_{ij}(X_t) - \gamma(|\tilde{X}_t|) \sigma_{ij}(\tilde{X}_t) \right) dB_t^j \\ &\quad - 2 \sum_{i=1}^n (X_t^i - \tilde{X}_t^i) \left(g(|X_t|) X_t^i - g(|\tilde{X}_t|) \tilde{X}_t^i \right) dt \\ &\quad + \sum_{i,j=1}^n (\gamma(|X_t|) \sigma_{ij}(X_t) - \gamma(|\tilde{X}_t|) \sigma_{ij}(\tilde{X}_t))^2 dt \\ &=: dN_t + \mathbf{I}_4 dt + \mathbf{I}_5 dt. \end{aligned}$$

The term \mathbf{I}_5 is the one which disallows direct use of Gronwall's lemma. It is singular in the sense that $\frac{\mathbf{I}_5}{W}$ can be arbitrarily large. Another singular term is \mathbf{I}_3 , but fortunately we also have a negative singular term $2(Y_t^p - \tilde{Y}_t^p)\mathbf{I}_1$, which will ensure that altogether we stay non-singular. We will bound all the terms \mathbf{I}_k and since $\mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3$ and \mathbf{I}_4 are exactly

the same⁵ as in [DeB04] in Lemmas 3.1, 3.2, 3.4, and 3.5, respectively, we will not do the calculations but only summarize final results. Let us introduce a non-negative process $Z := (Y^p - \tilde{Y}^p)(\tilde{Y}^{p-1} - Y^{p-1})$. To make sense of Z_t we implicitly multiply everything by $\mathbf{1}(Y_t > 0, \tilde{Y}_t > 0)$. We will use this convention until the end of the proof. Then we have

$$\begin{aligned} (Y_t^p - \tilde{Y}_t^p)\mathbf{I}_1 &\leq -2pZ_t |X_t|^2 G(|X_t|) + C_1 \varepsilon Z_t, \\ |\mathbf{I}_2| &\leq C_2 \left(|Y_t^p - \tilde{Y}_t^p| + |X_t - \tilde{X}_t| \right), \\ \mathbf{I}_3 &\leq \frac{p(2p-1)^2}{1-p} \gamma^2(|X_t|) |X_t|^2 Z_t + C_3 |X_t - \tilde{X}_t|^2 + C_3 \varepsilon Z_t, \\ \mathbf{I}_4 &\leq C_4 |X_t - \tilde{X}_t|^2, \end{aligned}$$

where constants C_1, C_2, C_3 , and C_4 are independent of ε . The bound for \mathbf{I}_5 has to be done differently due to non-diagonal nature of our SDE. By straightforward computation e.g. by computing the Frobenius norm of the matrix $\gamma(|X_t|)\sigma(X_t) - \gamma(|\tilde{X}_t|)\sigma(\tilde{X}_t)$, we see that

$$\begin{aligned} \mathbf{I}_5 &= \gamma^2(|X_t|) \left(1 - \sqrt{1 - |X_t|^2} \right)^2 + \gamma^2(|\tilde{X}_t|) \left(1 - \sqrt{1 - |\tilde{X}_t|^2} \right)^2 \\ &\quad - 2\gamma(|X_t|)\gamma(|\tilde{X}_t|) \left(1 - \sqrt{1 - |X_t|^2} \right) \left(1 - \sqrt{1 - |\tilde{X}_t|^2} \right) \frac{(X_t \cdot \tilde{X}_t)^2}{|X_t|^2 |\tilde{X}_t|^2} \\ &= \left(\gamma(|X_t|)Y_t^{1/2} - \gamma(|\tilde{X}_t|)\tilde{Y}_t^{1/2} + \gamma(|\tilde{X}_t|) - \gamma(|X_t|) \right)^2 \\ &\quad + 2\gamma(|X_t|)\gamma(|\tilde{X}_t|) \left(1 - \sqrt{1 - |X_t|^2} \right) \left(1 - \sqrt{1 - |\tilde{X}_t|^2} \right) \frac{|X_t|^2 |\tilde{X}_t|^2 - (X_t \cdot \tilde{X}_t)^2}{|X_t|^2 |\tilde{X}_t|^2}. \end{aligned}$$

For the first term we can use Cauchy-Schwarz inequality to bound it from above by

$$2 \left(\gamma(|X_t|)Y_t^{1/2} - \gamma(|\tilde{X}_t|)\tilde{Y}_t^{1/2} \right)^2 + 2 \left(\gamma(|\tilde{X}_t|) - \gamma(|X_t|) \right)^2 \leq C \left(\varepsilon^{2-p} Z_t + |X_t - \tilde{X}_t|^2 \right),$$

where the inequality follows from the proof of [DeB04, Lemma 3.6] and the Lipschitz continuity of γ . The second term can by our assumptions be bounded by

$$8(\sup \gamma)^2 \left(|X_t|^2 |\tilde{X}_t|^2 - (X_t \cdot \tilde{X}_t)^2 \right).$$

⁵Our G is defined slightly differently but it is still Lipschitz, so everything works.

Using $2X_t \cdot \tilde{X}_t = |X_t|^2 + |\tilde{X}_t|^2 - |X_t - \tilde{X}_t|^2$ we get $|X_t|^2 |\tilde{X}_t|^2 - (X_t \cdot \tilde{X}_t)^2 \leq |X_t - \tilde{X}_t|^2$. Hence we have shown $\mathbf{I}_5 \leq C_5(\varepsilon^{2-p}Z_t + |X_t - \tilde{X}_t|^2)$ with C_5 independent of ε . Using above facts we get

$$\begin{aligned}
2(Y_t^p - \tilde{Y}_t^p)\mathbf{I}_1 + \mathbf{I}_3 + \mathbf{I}_5 &\leq -4pZ_t |X_t|^2 G(|X_t|) + C\varepsilon Z_t \\
&+ \frac{p(2p-1)^2}{1-p} |X_t|^2 Z_t + C|X_t - \tilde{X}_t|^2 + C\varepsilon Z_t \\
&+ C(\varepsilon^{2-p}Z_t + |X_t - \tilde{X}_t|^2) \\
&= 4pZ_t \gamma^2(|X_t|) |X_t|^2 \left(1 - p + \frac{(2p-1)^2}{4(1-p)} + \frac{n-1}{2} - \frac{g(|X_t|)}{\gamma^2(|X_t|)}\right) \\
&+ C(2\varepsilon + \varepsilon^{2-p})Z_t + 2C|X_t - \tilde{X}_t|^2.
\end{aligned}$$

Note that the expression $1 - p + \frac{(2p-1)^2}{4(1-p)}$ is minimized at $p = 1 - \frac{\sqrt{2}}{4}$ and the value is then $\sqrt{2} - 1$. Therefore, we use the initial assumption that $\frac{g(1)}{\gamma^2(1)} - \frac{n-1}{2} > \sqrt{2} - 1$ to ensure the whole bracket is negative. Note also that such a choice of p implies $p = 1 - \frac{\sqrt{2}}{4} > 1 - (\sqrt{2} - 1) > 1 + \frac{n-1}{2} - \frac{g(1)}{\gamma^2(1)}$ so all previous calculations are justifiable since we have the necessary condition for our use of Lemma 2.1 from [DeB04]. Fixing $p = 1 - \frac{\sqrt{2}}{4}$ we then let ε be possibly even smaller to ensure that $\frac{g(u)}{\gamma^2(u)} - \frac{n-1}{2} > \sqrt{2} - 1 + \delta$ holds on $(1 - \varepsilon, 1]$ for some small fixed $\delta > 0$. The coefficient in front of Z_t then equals

$$\begin{aligned}
4p\gamma(|X_t|) |X_t|^2 \left(\sqrt{2} - 1 + \frac{n-1}{2} - \frac{g(|X_t|)}{\gamma^2(|X_t|)}\right) + C(2\varepsilon + \varepsilon^{2-p}) \\
\leq -4p(\inf \gamma^2)(1 - \varepsilon)\delta + C(2\varepsilon + \varepsilon^{2-p}).
\end{aligned}$$

Therefore, by letting ε be small enough we ensure that this coefficient in front of non-negative Z_t is negative and the bound $2(Y_t^p - \tilde{Y}_t^p)\mathbf{I}_1 + \mathbf{I}_3 + \mathbf{I}_5 \leq C|X_t - \tilde{X}_t|^2$ follows. Recall that

$$dW_t = d\tilde{M}_t + dN_t + 2(Y_t^p - \tilde{Y}_t^p)\mathbf{I}_1 dt + 2(Y_t^p - \tilde{Y}_t^p)\mathbf{I}_2 dt + \mathbf{I}_3 dt + \mathbf{I}_4 dt + \mathbf{I}_5 dt$$

and let $\tilde{\tau}_m$ be a localizing sequence of stopping times for the local martingale $\tilde{M} + N$. Then using above bounds and the fact that $\int_0^t \mathbf{1}(Y_s = 0 \text{ or } \tilde{Y}_s = 0) ds = 0$ yields

$$\begin{aligned} \mathbb{E}[W_{t \wedge \tau_\varepsilon \wedge \tilde{\tau}_m}] &= \mathbb{E} \left[\int_0^{t \wedge \tau_\varepsilon \wedge \tilde{\tau}_m} \left(2(Y_s^p - \tilde{Y}_s^p)(\mathbf{I}_1 + \mathbf{I}_2) + \mathbf{I}_3 + \mathbf{I}_4 + \mathbf{I}_5 \right) \mathbf{1}(Y_s > 0, \tilde{Y}_s > 0) ds \right] \\ &\leq C \mathbb{E} \left[\int_0^{t \wedge \tau_\varepsilon \wedge \tilde{\tau}_m} \left(|X_s - \tilde{X}_s|^2 + 2(Y_s^p - \tilde{Y}_s^p)^2 + 2|Y_s^p - \tilde{Y}_s^p| |X_s - \tilde{X}_s| \right) ds \right] \\ &\leq 3C \int_0^{t \wedge \tau_\varepsilon \wedge \tilde{\tau}_m} \mathbb{E}[W_s] ds \end{aligned}$$

and Gronwall's lemma implies that $\mathbb{E}[W_t] = 0$ and by non-negativity also $W_t = 0$ for $t \leq \tau_\varepsilon \wedge \tilde{\tau}_m$. Letting $m \rightarrow \infty$ we get $W_t = 0$ for $t < \tau_\varepsilon$. Therefore $X_t = \tilde{X}_t$ for $t < \tau_\varepsilon$ and pathwise uniqueness in Theorem 2.1.7 follows.

Chapter 3

Simulation of spherical Brownian motion

In this chapter we collect results of [MMU20] and [MMU21] on simulation of spherical Brownian motion. We describe an exact simulation algorithm for the increments of Brownian motion on a sphere of arbitrary dimension, based on the skew-product decomposition of the process with respect to the standard geodesic distance. The radial process turns out to be closely related to a Wright-Fisher diffusion, increments of which can be simulated exactly using the work of Jenkins & Spanò [JS17]. The rapid spinning phenomenon of the skew-product decomposition then yields the algorithm for the increments of the process on the sphere. Additionally, we consider a more general Brownian motion on a sphere of arbitrary radius and with a general diffusion coefficient and present a novel formula for its transition density. We also analyse the algorithm numerically and show that it remains stable when the time-step parameter is not too small.

3.1 Introduction

Since Brownian motion (BM) can be interpreted as a continuous time random walk with symmetric (Gaussian) increments, it is of fundamental importance in physics and other natural sciences. In applications one is often interested in the BM on curved surfaces and other manifolds, see e.g. [KDPN00] and [LTT08] for the modelling of fluorescent marker molecules in cell membranes and the motion of bacteria or any other diffusing particles, respectively. Often Monte Carlo simulation algorithms for such models on curved spaces are constructed using approximate tangent plane methods, which are accurate only for very small time steps, making the algorithms computationally expensive. Algorithms

allowing simulation over longer time steps are hence of particular interest in the physics literature. For example, [NEE03] gives a simulation algorithm for the BM on the three-dimensional sphere, based on its quaternionic structure. This algorithm is neither exact nor does it generalize easily to other dimensions. In [CEE10] it is applied to design an approximate simulation algorithm for the BM on the two-dimensional sphere. A further approximate algorithm for the simulation of BM on the two-sphere is given in [GSS12]. These improved methods not only give good results for small time-steps, but continue to provide a good approximation also for medium-sized time-steps (see also Section 3.4). Nevertheless, all these algorithms are just approximate and as the time-step increases the results become less and less accurate.

In contrast, for any dimension $d \geq 3$, our Algorithm 1 below simulates exactly the increments of the BM Z on the sphere \mathbb{S}^{d-1} . It is based on two facts established in Section 3.2: the radial part of Z (with respect to the standard metric on \mathbb{S}^{d-1}) can be transformed to a Wright-Fisher diffusions and, due to the rapid spinning of the skew-product decomposition of Z , its angular component is uniform on \mathbb{S}^{d-2} .

Algorithm 1 Simulation of the increment of Brownian motion Z on \mathbb{S}^{d-1} over any time interval

Require: Starting point $z \in \mathbb{S}^{d-1}$ and time horizon $t > 0$

- 1: Simulate the radial component: $X \sim \widetilde{\text{WF}}_{0,t}(\frac{d-1}{2}, \frac{d-1}{2})$ ▷ Algorithm 2 in Section 3.3
 - 2: Simulate the angular component: Y uniform on \mathbb{S}^{d-2}
 - 3: Set $O(z) := I - 2uu^\top$, where $u := (e_d - z)/|e_d - z|$
 - 4: **return** $O(z)(2\sqrt{X(1-X)}Y^\top, 1 - 2X)^\top$
-

The vectors in Algorithm 1 are column vectors of appropriate dimension, $e_d := (0, \dots, 0, 1)^\top \in \mathbb{S}^{d-1}$ denotes the north pole of the sphere and $O(z)$ is the reflection of \mathbb{R}^d across the hyperplane through the origin with the normal u . Algorithm 2 is the exact simulation algorithm for the increment of the Wright-Fisher diffusion given in [JS17, Alg. 1] (see Eq. (3.4) in the Section 3.3 for an appropriate definition of Wright-Fisher diffusions). Step 2 in Algorithm 1 consists of simulating a vector N in \mathbb{R}^{d-1} with independent standard normal components and setting $Y = N/|N|$ (we denote the standard Euclidean norm by $|\cdot|$ throughout).

The key property of the orthogonal matrix $O(z) \in \mathbb{R}^d \otimes \mathbb{R}^d$ in Algorithm 1 is $O(z)e_d = z$. In fact, any orthogonal matrix in $\mathbb{R}^d \otimes \mathbb{R}^d$ with this property would lead to an exact sample from the increment of BM on \mathbb{S}^{d-1} . Indeed, if $O_1(z), O_2(z) \in \mathbb{R}^d \otimes \mathbb{R}^d$ are two such orthogonal matrices, then the product $O_1^{-1}(z)O_2(z)$ fixes e_d and is given by an orthogonal transformation $\tilde{O}(z) \in \mathbb{R}^{d-1} \otimes \mathbb{R}^{d-1}$ on the orthogonal complement $\{e_d\}^\perp$ in \mathbb{R}^d . Hence

$O_2(z)(2\sqrt{X(1-X)}Y^\top, 1-2X)^\top = O_1(z)(2\sqrt{X(1-X)}(\tilde{O}(z)Y)^\top, 1-2X)^\top$ implying

$$O_2(z)(2\sqrt{X(1-X)}Y^\top, 1-2X)^\top \stackrel{d}{=} O_1(z)(2\sqrt{X(1-X)}Y^\top, 1-2X)^\top,$$

where $\stackrel{d}{=}$ denotes equality in law. The formula for $O(z)$ in Algorithm 1 is chosen due to its simplicity.

Algorithm 1 exploits the symmetry of both the geometry of \mathbb{S}^{d-1} and the law of spherical BM to reduce the simulation problem in any dimension to the simulation of an increment of a one-dimensional diffusion. Unlike the simulation methods in [NEE03, CEE10, GSS12], Algorithm 1 depends on the dimension d only through the value of the mutation parameters in the Wright-Fisher diffusion. As will be shown in Section 3.5, Algorithm 2, which simulates exactly from the law of the increment of the Wright-Fisher diffusion, is numerically stable for the time intervals of length $t \geq 0.1$. This is essentially because, when t is very small, the modulus of the summands in the alternating series (3.5) is increasing as a function of the index for large values of the index, before it starts to decrease monotonically. Conversely, Algorithm 2 (and hence Algorithm 1) becomes more efficient with increasing time horizon t (for small t , a normal approximation can be used to obtain a fast approximate algorithm substituting Algorithm 2, see [JS17, Thm 1] and the comments therein, or some of the aforementioned approximate algorithms can be used). Dimension d does not affect the performance of Algorithm 1.

Algorithm 1 above and the Markov property of the BM Z on the sphere \mathbb{S}^{d-1} yield an exact sample from any finite-dimensional marginal of Z . Moreover, as the BM on $\mathbb{S}^1 \subseteq \mathbb{C}$ can be represented as $Z = \exp(iW)$, where $i^2 = -1$ and W denotes a standard BM on \mathbb{R} , the assumption $d \geq 3$ in Algorithm 1 is not restrictive.

The main technical result (Proposition 3.2.1 below) states that the spherical BM Z enjoys a skew-product decomposition in which the last component of Z is independent of the normalized and suitably time-changed remaining components. Its main application in the present chapter consists of establishing the validity of Algorithm 1. The proof of Proposition 3.2.1 is analogous to that of Theorem 2.1.5. However, Proposition 3.2.1 is not a corollary of Theorem 2.1.5, as the latter implies only the independence of the modulus of the last component (but not of its sign). This difference makes the claim of Proposition 3.2.1 more general than Theorem 2.1.5 and possibly of independent interest.

The classical skew-product decomposition of Brownian motion ζ in \mathbb{R}^d , started away from the origin, features the BM Z on \mathbb{S}^{d-1} as follows: $Z_t = \zeta_{C_t}/|\zeta_{C_t}|$, where the increasing process $(C_t)_{t \geq 0}$ is the inverse of the additive functional $\int_0^\cdot |\zeta_u|^{-2} du$ of ζ . Moreover, the process Z and $|\zeta|$ are independent of each other. It is tempting to try to use the above

representation to obtain a sample from Z_t . One would first need to simulate C_t , then get a sample from the conditional law of ζ_{C_t} , given C_t , and finally normalise this sample to get an \mathbb{S}^{d-1} -valued random element identically distributed as Z_t . However, we stress that this does not yield an approximate simulation algorithm for the increments of Z , let alone an exact one. Even if one is able to simulate C_t , which is in itself non-trivial and can to the best of our knowledge only be done approximately, the key obstacle is the fact that the process ζ is not independent of the time C_t . More precisely, since C_t depends on the trajectory of ζ up to C_t , the conditional law of ζ_{C_t} , given C_t , is intractable, so the second step of the proposed sampling procedure cannot be achieved.

A more general notion of a Brownian motion on a sphere of arbitrary radius and a general diffusion coefficient is considered in Section 3.4. The main result of that section is a new representation of a transition density of spherical Brownian motion given by equation (3.9). The representation is a consequence of the skew-product decomposition of the spherical Brownian motion from Theorem 3.2.1 and the representation of transition densities of Wright-Fisher diffusion processes obtained in [GL83, JS17].

In Section 3.5 we analyse our algorithm numerically. In contrast to the existing literature in which algorithms produce samples from an approximate distribution and are hence forced to use smaller time-steps for accurate results, our algorithm is exact, so that it produces samples directly from the required distribution of the increments, and can handle arbitrarily large time-steps. In fact, its performance improves with increasing time-steps. The limiting factor of our algorithm is actually imperfect floating point arithmetic on computers. Since several quantities in the algorithm have to be computed by complicated expressions which become numerically unstable for smaller time-steps, we are actually limited with how small a time-step we are allowed to take. As such, our algorithm complements the other available methods. Moreover, in the numerical examples of Section 3.5 it appears to outperform them when the parameters are in the correct domain so we can use our algorithm.

Finally, we note that Algorithm 1 yields an algorithm for the exact simulation of the increments of the BM on the real $\mathbb{R}P^n$, complex $\mathbb{C}P^n$ and quaternionic $\mathbb{H}P^n$ projective spaces (for any integer n). These are Riemannian manifolds of the (real) dimension n , $2n$ and $4n$, respectively, with canonical Riemannian metrics described in Section 3.6. The Riemannian submersion π (mapping $\mathbb{S}^n \rightarrow \mathbb{R}P^n$, $\mathbb{S}^{2n+1} \rightarrow \mathbb{C}P^n$ and $\mathbb{S}^{4n+3} \rightarrow \mathbb{H}P^n$) by Lemma 3.6.1 below projects the BM Z on the relevant sphere to the BM $\pi(Z)$ on the projective space. Since $\pi(z) = [z]$ only converts the standard coordinates of the point z on the sphere to the homogeneous coordinates $[z]$ in the projective space, the random

element

$$\pi \left(O(z)(2\sqrt{X(1-X)}Y^\top, 1-2X)^\top \right),$$

where matrix $O(z)$ and random variables X, Y are the same as in Algorithm 1, gives the exact sample of the increment of the BM on the projective space, started at $[z]$, over the time horizon $t > 0$.

3.2 The skew-product decomposition

Brownian motion $Z = (Z_t)_{t \in \mathbb{R}_+}$ on the sphere \mathbb{S}^{d-1} will be once again introduced via the Stroock representation, i.e. it is given (in Itô form) by SDE

$$dZ_t = (I_d - Z_t Z_t^\top) dB_t - \frac{d-1}{2} Z_t dt, \quad Z_0 \in \mathbb{S}^{d-1}, \quad (3.1)$$

which possesses a unique strong solution, where B is a BM on \mathbb{R}^d . However, as (3.1) is non-constructive, we cannot use it directly for the exact simulation of the increments of Z in dimension $d \geq 3$.

As explained in the introduction, the key idea is to identify the skew-product decomposition of Z and exploit the rapid spinning phenomenon of the angular component at the starting point Z_0 , together with the symmetries of the sphere \mathbb{S}^{d-1} and the law of the spherical BM Z to obtain Algorithm 1. Let $D = (D_t)_{t \in \mathbb{R}_+}$ be the geodesic distance $D_t := \text{dis}(Z_t) \in [0, \pi]$ between Z_t and some fixed point $w \in \mathbb{S}^{d-1}$. In [IM96, p. 269] and [PR88], the authors show that D satisfies the SDE $dD_t = d\beta_t + \frac{d-2}{2} \cot(D_t) dt$, where β is a scalar BM. Since $\text{dis}(z) = \arccos(\langle z, w \rangle)$ for any $z \in \mathbb{S}^{d-1}$, the natural transformation $\tilde{\mathcal{X}} := \cos(D)$ leads to the Jacobi diffusion (see e.g. [WY98] and the proof of Proposition 3.2.1 below) satisfying the SDE

$$d\tilde{\mathcal{X}}_t = \sqrt{1 - \tilde{\mathcal{X}}_t^2} d\tilde{\beta}_t - \frac{d-1}{2} \tilde{\mathcal{X}}_t dt, \quad \tilde{\mathcal{X}}_0 = \langle Z_0, w \rangle, \quad (3.2)$$

where $\tilde{\beta} = -\beta$ is a standard scalar BM. This simple (and in our context crucial) observation has to the best of our knowledge not been made in the literature and was for the first time noticed in [MMU20]. Note that $\tilde{\mathcal{X}}_t = \langle Z_t, w \rangle$. We henceforth fix $w = e_d$ and obtain $\tilde{\mathcal{X}}_t = Z_t^d$, making $\tilde{\mathcal{X}}$ the process considered in Proposition 2.1.1 for $n = 1$ and $\ell = d - 1$. A linear transformation $\mathcal{X}_t := \frac{1 - \tilde{\mathcal{X}}_t}{2}$ yields the SDE

$$d\mathcal{X}_t = \sqrt{\mathcal{X}_t(1 - \mathcal{X}_t)} d\beta_t + \left(\frac{d-1}{4}(1 - \mathcal{X}_t) - \frac{d-1}{4}\mathcal{X}_t \right) dt, \quad \mathcal{X}_0 = \frac{1 - Z_0^d}{2}, \quad (3.3)$$

making \mathcal{X} a Wright-Fisher diffusion with mutation parameters $\theta_1 = \frac{d-1}{2} = \theta_2$ (see also section 3.3 where we additionally explain the discrepancy between (3.3) and (2.3) by a factor 2). A weak form of the skew-product decomposition (on the level of generators) of the BM on the sphere has been established in [PR88], see also [IM96, p. 269]. In order to give a pathwise skew-product representation of Z , first note that for $d \geq 3$ the Wright-Fisher diffusion \mathcal{X} visits neither 0 nor 1. We may thus introduce the following time-changes: for $0 \leq s \leq t$ define $S_s(t) := \int_s^t \frac{1}{4\mathcal{X}_u(1-\mathcal{X}_u)} du$, satisfying $\lim_{t \rightarrow \infty} S_s(t) = \infty$ (see proof of Proposition 3.2.1 below), and its inverse $T_s: [0, \infty) \rightarrow [s, \infty)$. We now state our main result.

Proposition 3.2.1 (Skew-product decomposition of the BM on \mathbb{S}^{d-1}). *Let $d \geq 3$ and Z be a solution of SDE (3.1). Pick $s \in [0, \infty)$ and assume that either $s > 0$ or $s = 0$ and $Z_0 \neq e_d$. Let $U = (Z^1, \dots, Z^{d-1})^\top$ denote the first $d-1$ components of Z . Then $|U| = 2\sqrt{\mathcal{X}(1-\mathcal{X})}$ and the process $\widehat{V} = (\widehat{V}_t)_{t \geq 0}$, given by $\widehat{V}_t := U_{T_s(t)} / |U_{T_s(t)}|$ is a BM on the sphere \mathbb{S}^{d-2} , started at $\widehat{V}_0 = U_s / |U_s|$, independent of \mathcal{X} . Hence we obtain the skew-product decomposition $Z_t = (2\sqrt{\mathcal{X}_t(1-\mathcal{X}_t)} \widehat{V}_{S_s(t)}^\top, 1 - 2\mathcal{X}_t)^\top$ for $t \geq s$. Furthermore, if $Z_0 = e_d$ (i.e. $U_0 = 0$), then \widehat{V}_t is uniformly distributed on \mathbb{S}^{d-2} for any $t > 0$ and subsequently evolves as a stationary BM on the sphere.*

Since time-changing a stationary process by an independent time-change does not affect the marginal distributions of the process, the second part of the theorem immediately yields Corollary 3.2.2, which in turn justifies Algorithm 1 for the exact simulation of the increments of BM on the sphere \mathbb{S}^{d-1} .

Corollary 3.2.2. *Let $d \geq 3$ and Z be the BM on \mathbb{S}^{d-1} started at $Z_0 = e_d$. Additionally, let \mathcal{W} be the unique strong solution of SDE (3.4) with mutation parameters $\theta_1 = \frac{d-1}{2} = \theta_2$, started at $\mathcal{W}_0 = 0$, and Y a random vector uniformly distributed on \mathbb{S}^{d-2} , independent of \mathcal{W} . Then for each $t \geq 0$, the random vectors Z_t and $(2\sqrt{\mathcal{W}_t(1-\mathcal{W}_t)} Y^\top, 1 - 2\mathcal{W}_t)^\top$ are identically distributed.*

It may appear that the skew-product decomposition in Proposition 3.2.1 can be applied directly to simulate the increments of the spherical BM started at any point in \mathbb{S}^{d-1} . Since the increment of the Wright-Fisher diffusion can be simulated exactly using Algorithm 2 below, Proposition 3.2.1 appears to reduce the problem to the simulation of the increment of the BM on \mathbb{S}^{d-2} . Recursively, this would reduce the problem to the simulation of the BM on \mathbb{S}^1 , where the algorithm is trivial. Unfortunately, for this direct approach to work we would need to obtain a sample from law of the pair $(\mathcal{X}_t, \int_0^t \frac{1}{4\mathcal{X}_u(1-\mathcal{X}_u)} du)$, which at the time of writing we do not know how to do. As stated

in Corollary 3.2.2, this problem disappears when the BM Z starts at the north pole e_d . Since for any orthogonal matrix $A \in \mathbb{R}^d \otimes \mathbb{R}^d$ the process AZ solves the SDE in (3.1) started at AZ_0 , Corollary 3.2.2 and the orthogonal matrix $O(z)$ in Algorithm 1 circumvent the need to simulate from the law of the pair $(\mathcal{X}_t, \int_0^t \frac{1}{4\mathcal{X}_u(1-\mathcal{X}_u)} du)$ for an arbitrary starting point $z \in \mathbb{S}^{d-1}$.

Proof of Proposition 3.2.1. In order to apply Theorem 2.1.5, let $n = d - 1$, $\ell = 1$ and set $\gamma \equiv 1$, $g \equiv \frac{d-1}{2}$. Then the process U in Proposition 3.2.1 equals the process X considered in Theorem 2.1.5. In particular, the processes $|U|$ and R are equal implying that the time-changes $S_s(t)$ and $T_s(t)$ used in both Proposition 3.2.1 and Theorem 2.1.5 coincide. Moreover, since Lemma 2.2.3 holds in the current setting the time-changes are well-defined. We finally note that $|U| = 2\sqrt{\mathcal{X}(1-\mathcal{X})}$.

As mentioned in the introduction, the only thing left to prove is the independence of \widehat{V} and $\mathcal{X} = (1-Z^d)/2$. The argument is analogous to the one that yielded the independence of \widehat{V} and R in the proof of Theorem 2.1.5, but the statement does not follow directly from Theorem 2.1.5. Let $V := U/|U|$ and note that, Itô's formula applied to V yields the following $dV_t = |U_t|^{-1} (I_{d-1} - V_t V_t^\top) d\widetilde{B}_t - |U_t|^{-2} \frac{d-2}{2} V_t dt$, where $\widetilde{B} = (B^1, \dots, B^{d-1})^\top$. By Proposition 2.1.1 (with $n = 1, \ell = d - 1$) we have that $\widetilde{\mathcal{X}} = Z^d$ satisfies SDE (3.2) above with $\widetilde{\mathcal{X}}_0 = Z_0^d$. Moreover, as in the proof of Proposition 2.1.1, the scalar BM $\widetilde{\beta}$ in (3.2) is given by $d\widetilde{\beta}_t = -(1 - 2\mathcal{X}_t) V_t^\top d\widetilde{B}_t + |U_t| dB_t^d$. Consequently, \mathcal{X} satisfies SDE (3.3) with $\mathcal{X}_0 = \frac{1-Z_0^d}{2}$ and the scalar BM $\beta = -\widetilde{\beta}$. Since $\widehat{V} = V_{T_s(\cdot)}$, the change-of-time formulae [RY99, Ch. V, §1] for the Itô and the Lebesgue-Stieltjes integrals imply that $\widehat{V}_t - \widehat{V}_0 = \int_0^t (I_{d-1} - \widehat{V}_u \widehat{V}_u^\top) d\widetilde{W}_u - \int_0^t \frac{d-2}{2} \widehat{V}_u du$, where $\widetilde{W}_t = \int_s^{T_s(t)} |U_u|^{-1} d\widetilde{B}_u$ is a standard BM on \mathbb{R}^{d-1} by Levy's characterization (see Theorem A.2.5). This implies that \widehat{V} solves the SDE in (3.1) in dimension $d - 1$, making it a BM on \mathbb{S}^{d-2} .

In order to conclude that \mathcal{X} and \widehat{V} are independent, we assume without loss of generality that the underlying probability space with filtration $(\mathcal{F}_t)_{t \geq 0}$ supports a further scalar (\mathcal{F}_t) -BM ξ , independent of B . Define the continuous local martingale $W = (W_t)_{t \geq 0}$ in \mathbb{R}^{d-1} by

$$W_t = \int_s^{T_s(t)} \frac{1}{|U_u|} (I_{d-1} - V_u V_u^\top) d\widetilde{B}_u + \int_s^{T_s(t)} \frac{V_u}{|U_u|} d\xi_u.$$

Since $(I_{d-1} - V_u V_u^\top)(I_{d-1} - V_u V_u^\top) = I_{d-1} - V_u V_u^\top$, it follows easily from the independence of B and ξ that the components of its quadratic covariation W are given by $\langle W^i, W^j \rangle_t = \int_s^{T_s(t)} \delta_{ij} |U_u|^{-2} du = \delta_{ij} t$ for any $i, j \in \{1, \dots, d-1\}$. Hence, by Levy's characterization, the process W is a standard (\mathcal{G}_t) -BM, where the filtration $(\mathcal{G}_t)_{t \geq 0}$ is defined by $\mathcal{G}_t := \mathcal{F}_{T_s(t)}$

for all $t \geq 0$. Moreover, since

$$(I_{d-1} - V_u V_u^\top) \cdot \left[\frac{1}{|U_u|} (I_{d-1} - V_u V_u^\top), \frac{V_u}{|U_u|} \right] = \left[\frac{1}{|U_u|} (I_{d-1} - V_u V_u^\top), 0 \right]$$

we can apply the change-of-time formulae for the stochastic and Lebesgue-Stieltjes integral [RY99, Ch. V, §1] to obtain $\widehat{V}_t - \widehat{V}_0 = \int_0^t (I_{d-1} - \widehat{V}_u \widehat{V}_u^\top) dW_u - \int_0^t \frac{d-2}{2} \widehat{V}_u du$.

Since SDE in (3.1) in dimension $d-1$ has a unique strong solution, the independence of \widehat{V} and \mathcal{X} follows from the independence of the driving BMs W and β . Since W and β run on different time scales, we first note that the Markov property of W implies that W depends on $\mathcal{G}_0 = \mathcal{F}_s$ only through its position at time zero, i.e. $W_0 = 0$, making it independent of \mathcal{F}_s . In particular, W is independent of $(\beta_t)_{t \in [0, s]}$. Therefore it is sufficient to prove that W is independent of $(\beta_{u+s} - \beta_s)_{u \geq 0}$. For any $t \in [0, \infty)$, define $\eta_t := \beta_{T_s(t)} - \beta_s = \int_s^{T_s(t)} (1 - 2\mathcal{X}_u) V_u^\top d\widetilde{B}_u - \int_s^{T_s(t)} |U_u| dB_u^d$ so that η is a (\mathcal{G}_t) -local martingale. Simple calculations show that $\langle W^i, \eta \rangle_t = 0$ for any component W^i of W and the quadratic variation $\langle \eta \rangle_t = T_s(t) - s$ satisfies $\langle \eta \rangle_{S_s(u+s)} = u$ for all $u \geq 0$. By Knight's Theorem (also known as the multidimensional Dambis-Dubins-Schwarz Theorem [RY99, Ch. V, Thm 1.9]) we have that W and $(\eta_{S_s(t+s)})_{t \geq 0} = (\beta_{s+t} - \beta_s)_{t \geq 0}$ are independent BMs. This concludes the proof of the proposition. \square

3.3 Simulation of Wright-Fisher diffusion processes

Fix non-negative parameters θ_1 and θ_2 and consider the SDE

$$d\mathcal{W}_t = \sqrt{\mathcal{W}_t(1 - \mathcal{W}_t)} dB_t + \frac{1}{2}(\theta_1(1 - \mathcal{W}_t) - \theta_2\mathcal{W}_t)dt, \quad \mathcal{W}_0 = x \in [0, 1]. \quad (3.4)$$

Its unique strong solution is known as the Wright-Fisher diffusion with mutation parameters θ_1 and θ_2 ; denote it by $\widetilde{\text{WF}}_x(\theta_1, \theta_2)$ and its law at time t by $\widetilde{\text{WF}}_{x,t}(\theta_1, \theta_2)$. We have already considered a closely related SDE (2.3) with non-negative parameters α and β and also called its solution a Wright-Fisher diffusion and denoted its law by $\text{WF}_x(\alpha, \beta)$.

The two notions really are related as we can easily show that if the process \mathcal{Z} is $\text{WF}_x(\alpha, \beta)$, then the time-changed process \mathcal{W} given by $\mathcal{W}_t := \mathcal{Z}_{t/4}$ is $\widetilde{\text{WF}}_x(\alpha/2, \beta/2)$. Therefore, we may choose the definition which suits us better for the application at hand and all the results can be transferred to an alternative definition by a simple time-change and by suitably transforming the parameters.

For simulation we use the definition given by SDE (3.4) which is incidentally the one used in [JS17] from which the following algorithm is taken.

Algorithm 2 Exact simulation from the law $\widetilde{\text{WF}}_{x,t}(\theta_1, \theta_2)$

Require: Mutation parameters θ_1 and θ_2 , starting point $x \in [0,1]$ and time horizon $t > 0$

- 1: Simulate $M \stackrel{d}{=} A_\infty^\theta(t)$ ▷ Use [JS17, Alg. 2]
 - 2: Simulate $L \sim \text{Binomial}(M, x)$
 - 3: Simulate $Y \sim \text{Beta}(\theta_1 + L, \theta_2 + M - L)$
 - 4: **return** Y
-

The random variable $A_\infty^\theta(t)$ in step 1 is integer-valued with the mass function $\{q_m^\theta(t); m = 0, 1, \dots\}$, where $\theta := \theta_1 + \theta_2$, that can be described as follows. Let $(A_n^\theta(t))_{t \in \mathbb{R}_+}$ be a pure death process on the non-negative integers, started at $A_n^\theta(0) = n$, where the only transitions are of the form $m \mapsto m - 1$ and occur at the rate $m(m + \theta - 1)/2$ for each $m \in \{1, \dots, n\}$. Then $q_m^\theta(t) = \mathbb{P}[A_\infty^\theta(t) = m]$ can be expressed as the limit $q_m^\theta(t) = \lim_{n \rightarrow \infty} \mathbb{P}[A_n^\theta(t) = m]$.

The coefficients $q_m^\theta(t)$ are known in terms of the following alternating series [JS17, Eq. (5)]

$$q_m^{d-1}(t) := \sum_{k=m}^{\infty} (-1)^{k-m} b_k^{(t, d-1)}(m), \quad (3.5)$$

with the summands

$$b_k^{(t, d-1)}(m) := \frac{(d + 2k - 2)(d - 1 + m)_{(k-1)}}{m!(k - m)!} e^{-k(k+d-2)t/2},$$

where the Pochhammer symbol is given by $a_{(x)} = \frac{\Gamma(a+x)}{\Gamma(a)}$ for $a > 0, x \geq -1$.

The absolute values of summands $b_k^{(t, d-1)}(m)$ are ultimately monotonically decreasing and this property is exploited in [JS17, Alg. 2] for the simulation of $A_\infty^\theta(t)$. In theory, [JS17, Alg. 2] is exact for all values of parameters. In practice, when the time horizon t is very small, the algorithm runs into running time and floating number precision problems, as it requires multiplication of very large and very small numbers. Further analysis of these phenomena is considered in Section 3.5 where we show that the algorithm works well for $t \geq 0.1$ (the values of θ_1, θ_2 do not seem to affect the performance). Additionally, for $t \geq 0.1$ almost all coefficients $q_m^{d-1}(t)$ are very small and the larger ones can be efficiently computed to a machine precision. Hence, in practical applications it turns out that essentially exact sample from $A_\infty^\theta(t)$ can be quickly and efficiently obtained by a simple inversion sampling instead of using a more complicated and computationally expensive [JS17, Alg. 2].

3.4 General spherical Brownian motion and new representation of transition density

Especially in physics, one is often interested in a slightly more general notion of a Brownian motion on the sphere

$$\mathbb{S}^{d-1}(R) := \{x \in \mathbb{R}^d ; |x| = R\},$$

of radius $R > 0$. More precisely, if we denote by $\rho_y^{(D,R)}(x,t)$ the transition density of the general spherical Brownian motion started at $y \in \mathbb{S}^{d-1}(R)$, we are interested in the diffusion equation:

$$\frac{\partial \rho_y^{(D,R)}}{\partial t}(x,t) = D \nabla_{\mathbb{S}^{d-1}(R)}^2 \rho_y^{(D,R)}(x,t), \quad \rho_y^{(D,R)}(x,0) = \delta(x,y), \quad (3.6)$$

where $\nabla_{\mathbb{S}^{d-1}(R)}^2$ is the Laplace-Beltrami operator on the sphere $\mathbb{S}^{d-1}(R)$, $D > 0$ is a diffusion coefficient and $y \in \mathbb{S}^{d-1}(R)$ is the initial point. Note that the standard spherical Brownian motion, which we considered in previous sections, is such a process which diffuses on a sphere $\mathbb{S}^{d-1} = \mathbb{S}^{d-1}(1)$ of radius 1 and corresponds to $D = 1/2$.

Symmetry of the spherical Brownian motion (or, more precisely, of the Laplace-Beltrami operator on the sphere) allows us to deduce the following fact: if $A \in \mathbb{R}^d \otimes \mathbb{R}^d$ is any orthogonal matrix, then $\rho_{Ay}^{(D,R)}(Ax,t) = \rho_y^{(D,R)}(x,t)$. This in particular shows that it is enough to consider a single initial point – we are going to use the north pole $R \cdot e_d \in \mathbb{S}^{d-1}(R)$, where $e_d := (0, \dots, 0, 1)^\top$ represents the north pole of the unit sphere. Furthermore, symmetry allows us to deduce that $\rho_y^{(D,R)}(x,t)$ does not depend on the whole vector x but only depends on the geodesic distance between the vectors x and y .

Recall that we are using the standard round metric on the sphere which is induced by the standard Euclidean scalar product $\langle \cdot, \cdot \rangle$. The geodesic distance measures the shortest distance between two points (i.e. along great circles which are intersections of the sphere with two dimensional planes through the origin) and is given as a function $\text{dis}: \mathbb{S}^{d-1}(R) \times \mathbb{S}^{d-1}(R) \rightarrow [0, R\pi]$, $\text{dis}(x,y) = R \arccos(\langle x,y \rangle / R^2) =: R\theta$, i.e. the distance is equal to R times the angle between the two vectors (here denoted by θ). Since in unit sphere coordinates relation $\nabla_{\mathbb{S}^{d-1}(R)}^2 = \nabla_{\mathbb{S}^{d-1}}^2 / R^2$ holds, we can define a new parameter $\tau := 2Dt/R^2$, which we use instead of time and then $\rho_{Re_d}^{(D,R)}(x,t) = \rho_{e_d}^{(1/2,1)}(x/R, \tau)$ holds. This allows us to focus on the standard spherical Brownian motion and we can then extend results to arbitrary radius and diffusion coefficient by rescaling and linear time-change.

Such rescaling leads to the following simulation algorithm which is a generalization

of Algorithm 1.

Algorithm 3 Simulating from the transition density $\rho_y^{(D,R)}(\cdot; t)$ of the spherical Brownian motion

Require: Starting point $y \in \mathbb{S}^{d-1}(R)$ and time horizon $t > 0$

- 1: Set $\tau = 2Dt/R^2$
 - 2: Simulate the radial component: $X \sim \widetilde{\text{WF}}_{0,\tau}(\frac{d-1}{2}, \frac{d-1}{2})$ ▷ Algorithm 2 in Section 3.3
 - 3: Simulate the angular component: Y uniform on \mathbb{S}^{d-2}
 - 4: Set $O(y) = I - 2uu^\top$, where $u = (e_d - y/R)/|e_d - y/R|$
 - 5: **return** $RO(y)(2\sqrt{X(1-X)}Y^\top, 1 - 2X)^\top$
-

We also want to analyse the transition density so from now on we fix the parameters $D = 1/2$ and $R = 1$ and omit them from the notation and simply write $\rho_{e_d}(z, \tau)$ for the transition density at time τ of the standard spherical Brownian motion started at e_d . Any point $z \in \mathbb{S}^{d-1}$ can be represented as $z = \sin \theta w + \cos \theta e_d$, where $w \in \mathbb{S}^{d-2} \subseteq \mathbb{R}^{d-1} \times \{0\}$ and $\theta \in [0, \pi]$ is the angle between the vectors z and e_d . Due to symmetry we see that $\rho_{e_d}(z, \tau) = \rho_{e_d}(\sin \theta w + \cos \theta e_d, \tau)$ does not depend on the vector w but only depends on the angle θ . Hence we can and will consider the one-dimensional density

$$\begin{aligned} \rho(\theta, \tau) &:= \text{Vol}(\mathbb{S}^{d-2}(\sin \theta)) \rho_{e_d}(\sin \theta w + \cos \theta e_d, \tau) \\ &= A_{d-2} \sin^{d-2} \theta \cdot \rho_{e_d}(\sin \theta w + \cos \theta e_d, \tau), \end{aligned}$$

where $A_n := 2\pi^{(n+1)/2}/\Gamma(\frac{n+1}{2})$ is the volume of an n -dimensional sphere \mathbb{S}^n . The additional multiplicative factor $A_{d-2} \sin^{d-2} \theta$ is included due to only considering a one-dimensional process (and its transition density) instead of the whole process on the sphere and we have $\int_0^\pi \rho(\theta, \tau) d\theta = 1$. For $\tau = 0$ the initial distribution $\rho(\theta, 0)$ of this density is equal to the delta function with support at 0.

Using (ultra)spherical harmonics and their addition formula we can get the following formal solution of the diffusion equation at time τ (see [KT81, p. 339], where their Jacobi polynomials are normalised differently to our Gegenbauer polynomials):

$$\rho(\theta, \tau) = A_{d-2} \sin^{d-2} \theta \sum_{l=0}^{\infty} \frac{h(l, d)}{A_{d-1}} \frac{C_l^{(d/2-1)}(\cos \theta)}{C_l^{(d/2-1)}(1)} e^{\lambda_l \tau / 2}, \quad (3.7)$$

where $C_n^{(\alpha)}$ are Gegenbauer polynomials given by their generating series $(1 - 2xt + t^2)^{-\alpha} = \sum_{n=0}^{\infty} C_n^{(\alpha)}(x)t^n$ and $h(\ell, d) = \binom{d+\ell-1}{d-1} - \binom{d+\ell-3}{d-1}$ is the dimension of the space of spherical

harmonics of degree ℓ , corresponding to the eigenvalue $\lambda_\ell := -\ell(\ell + d - 2)$. equation (3.7) without the additional factor $A_{d-2} \sin^{d-2} \theta$ (i.e. the density for the process on the sphere) has also been derived in [Cai04] where a further simplification $h(\ell, d)/C_\ell^{(d/2-1)}(1) = (2\ell + d - 2)/(d - 2)$ is used.

In the special case $d = 3$, i.e. for the standard unit sphere \mathbb{S}^2 , equation (3.7) specializes to the well-known formula

$$\rho(\theta, \tau) = \frac{\sin \theta}{2} \sum_{\ell=0}^{\infty} (2\ell + 1) P_\ell(\cos \theta) e^{-\ell(\ell+1)\tau/2}, \quad (3.8)$$

where P_ℓ is the ℓ -th Legendre polynomial. Symmetry arguments show that being able to simulate from the density $\rho(\cdot, \tau)$ in turn allows us to simulate increments of the spherical Brownian motion. Unfortunately, formal solutions given by formulas (3.7) or (3.8) oscillate a lot even when large amount of terms are taken into account and are therefore not suitable for simulation. One is then usually forced to use some approximative density and sampling from it.

The simplest option is to locally approximate a sphere with its tangent space and use a standard Brownian (Gaussian) increment which is then translated to an increment on a sphere using the exponential map, i.e. moving an appropriate distance along the geodesics (great circles) of a sphere. Such an approximation corresponds to density $Q_{Tang}(\theta, \tau) = \mathcal{N}_1(\tau) \theta^{d-2} \exp(-\theta^2/(2\tau))$, where $\mathcal{N}_1(\tau)$ is a normalization constant such that $\int_0^\pi Q_{Tang}(\theta, \tau) d\theta = 1$. This approximation largely ignores the curvature of the sphere so it is a good approximation only for very small values of parameter τ . Therefore, [GSS12] proposed an improved approximation. It is given by $Q_{Approx}(\theta, \tau) = \mathcal{N}_2(\tau) (\theta \sin \theta)^{(d-2)/2} \exp(-\theta^2/(2\tau))$, where $\mathcal{N}_2(\tau)$ is a normalization constant. This new approximation continues to give good results even for intermediate values of τ . In the case of $d = 4$ the same approximative density was derived in [NEE03] by using a different method.

One of the main result in this chapter is the following alternative representation of a solution of the diffusion equation, which is much more suitable for simulation. It is the backbone of Algorithms 1 and 3 and the reason why we are able to sample directly from the transition density and not just from an approximative density as in the other previously mentioned methods. The alternative representation of the density (3.7) is given by

$$\rho(\theta, \tau) = \frac{\sin \theta}{2} \sum_{m=0}^{\infty} q_m^{d-1}(\tau) g_{(\frac{d-1}{2}, \frac{d-1}{2}+m)} \left(\frac{1 - \cos \theta}{2} \right), \quad (3.9)$$

where

$$g_{(\alpha,\beta)}(x) := \frac{1}{\text{B}(\alpha,\beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad x \in [0,1], \alpha, \beta > 0$$

in which B represents the Beta function so that $g_{(\alpha,\beta)}$ is a density of a Beta distribution and $q_m^{d-1}(\tau)$ are given by (3.5).

To show that $\rho(\theta, t)$ is indeed represented as in (3.9) we have to notice that the density $\rho(\theta, t)$ is actually the transition density of the one-dimensional $[0, \pi]$ -valued process $D_t = \text{dis}(Z_t, e_d)$, where Z_t denotes the standard Brownian motion on the unit sphere \mathbb{S}^{d-1} started at e_d . Since $\cos(D_t)$ is nothing but the last component of the spherical Brownian motion Z_t , it is quite natural to further transform the process D_t and instead look at $\mathcal{X}_t = \frac{1 - \cos(D_t)}{2}$. This process is $[0, 1]$ -valued and we have shown in Section 3.2 show that the process \mathcal{X}_t is a Wright-Fisher diffusion process with parameters $\theta_1 = \frac{d-1}{2} = \theta_2$ and an initial point $y = 0$ (see (3.3)). Its transition density $f(x; t)$ is a solution of a Fokker-Planck equation

$$\begin{aligned} \frac{\partial f(x; t)}{\partial t} &= \frac{\partial^2}{\partial x^2} \left(\frac{x(1-x)}{2} f(x; t) \right) - \frac{\partial}{\partial x} \left(\left(\frac{d-1}{4}(1-x) - \frac{d-1}{4}x \right) f(x; t) \right), \\ f(x; 0) &= \delta(x, 0). \end{aligned}$$

The transition densities of Wright-Fisher diffusion processes have a spectral decomposition given by the associated Jacobi polynomials [GS10], but such a representation is essentially equivalent to representation (3.7). However, there exists another representation of their transition densities, which arises from the moment duality between Wright-Fisher diffusion processes and certain coalescent processes. This representation and its use for simulation was given in [GL83] and it is given explicitly in [JS17, Eq. (4)]. The transformation formula for densities then immediately yields $\rho(\theta, \tau) = \frac{\sin \theta}{2} f\left(\frac{1 - \cos \theta}{2}; \tau\right)$ and combining it with [JS17, Eq. (4)] we immediately get the representation (3.9).

3.5 Computer simulations

First, we check numerically that our alternative representation (3.9) is actually equivalent to a more well-known representation (3.7). To see that we are going to let $\tau = 0.5$ and denote by ρ_1^d the density computed by the first 60 terms in (3.7) and by ρ_2^d density computed by the first 40 terms in (3.9) where each coefficient $q_m^{d-1}(\tau)$ is calculated using the first 60 terms in (3.5). All the densities have been calculated in Mathematica for $d = 3, 4$ and on Figure 3.1 we plot the absolute and relative error. Both errors are extremely small across the interval $[0, \pi]$, which confirms that both densities are equal.

Additionally, increasing the dimension d seems to improve results.

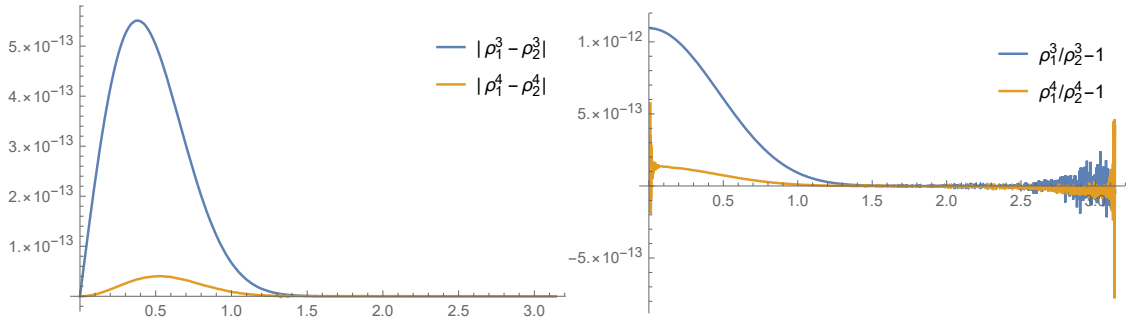


Figure 3.1: Absolute errors $|\rho_1^d - \rho_2^d|$ on the left plot and relative error $\rho_1^d / \rho_2^d - 1$ on the right plot for $d = 3, 4$.

Typically, when using algorithms for the simulation of the spherical Brownian motion, we need to repeatedly get samples for a fixed dimension d and time τ . The most computationally expensive part of our simulation Algorithms 1 and 3 is step 1 in Algorithm 2 which returns a sample from an integer-valued random variable $A_\infty^{d-1}(\tau)$.

The first step is to pre-compute coefficients $q_m^{d-1}(\tau)$ to a prescribed precision and afterwards use them in an inversion sampling of $A_\infty^{d-1}(\tau)$. One could also use an exact algorithm [JS17, Alg. 2], which is essentially just a modification of inversion sampling where we never have to know the coefficients exactly. However, the exact algorithm is much more computationally expensive and as we will see below the basic inversion sampling already gives essentially exact sample anyway.

Since the algorithm depends deeply on the representation (3.7), it suffers from numerical instability as $\tau \rightarrow 0$. In theory the algorithm works correctly for all values of τ , but in practice we are limited by running time and by imperfect floating number precision, and whence the potential incorrect computations of the relevant coefficients $q_m^{d-1}(\tau)$. When τ is small we have to compute products of the form $\frac{(d+2k-2)(d-1+m)_{(k-1)}}{m!(k-m)!} e^{-k(k+d-2)\tau/2}$ for large values of k and m , which means multiplying very large and very small numbers. This quickly accumulates numerical errors which cause the algorithm to fail. For example, using direct calculations done in Mathematica via partial sums and the equation (3.5) we calculate $q_{26}^2(0.05)$ to be equal to -0.00427 and even more extremely $q_{26}^2(0.03)$ to be equal to 29741.98 . Both of these results are clearly wrong and will cause the algorithm to fail.

In order to understand better how computation of coefficients $q_m^{d-1}(\tau)$ can go wrong, we have to look at the terms $b_k^{(\tau, d-1)}(m)$ from equation (3.5). We need to be able to quantify the error in our calculations, meaning we need to know how well the partial

sums $\sum_{k=m}^{m+K} (-1)^{k-m} b_k^{(\tau, d-1)}(m)$ approximate the coefficients $q_m^{d-1}(\tau)$. The partial sums form an alternating series and one is tempted to say that the partial sums with even K are always an upper bound for $q_m^{d-1}(\tau)$, whereas the partial sums with odd K are always a lower bound. Unfortunately, for this to hold, we would need $b_k^{(\tau, d-1)}(m) \downarrow 0$ as $k \rightarrow \infty$, which is not the case, although it is true once k is large enough. To see this we notice that

$$\frac{b_{k+1}^{(\tau, d-1)}(m)}{b_k^{(\tau, d-1)}(m)} = \frac{m+k+d-2}{k-m+1} \frac{2k+d}{2k+d-2} e^{-(2k+d-1)\tau/2},$$

which is a product of three decreasing functions in k and the limit as $k \rightarrow \infty$ is equal to 0 (recall that only terms with $k \geq m$ are relevant). This means that as soon as the quotient is for the first smaller than 1 then the terms $b_k^{(\tau, d-1)}(m)$ start decaying indefinitely. Therefore, for a large enough number of terms taken, the exact value $q_m^{d-1}(\tau)$ always lies between the two consecutive partial sums and they differ by just $b_k^{(\tau, d-1)}(m)$, which is hence also the upper bound for the error. This means that in order to compute coefficients $q_m^{d-1}(\tau)$ to a fixed precision $\varepsilon > 0$ we simply start computing each coefficient via partial sums in an infinite series representation (3.5) and we stop once the terms $b_k^{(\tau, d-1)}(m)$ start decaying and additionally $b_k^{(\tau, d-1)}(m) < \varepsilon$ holds.

As we have seen above, the partial sums $\sum_{k=m}^{m+K} (-1)^{k-m} b_k^{(\tau, d-1)}(m)$ eventually become arbitrarily close to the precise value $q_m^{d-1}(\tau) \in [0, 1]$. However, before reaching this value they can sometimes be very far off, in some cases the difference can be of several orders of magnitude. To see this we plot in Figure 3.2 some of the terms $b_{m+i}^{(\tau, d-1)}(m)$ and some of the partial sums, which are used in the computation of $q_m^{d-1}(\tau)$.

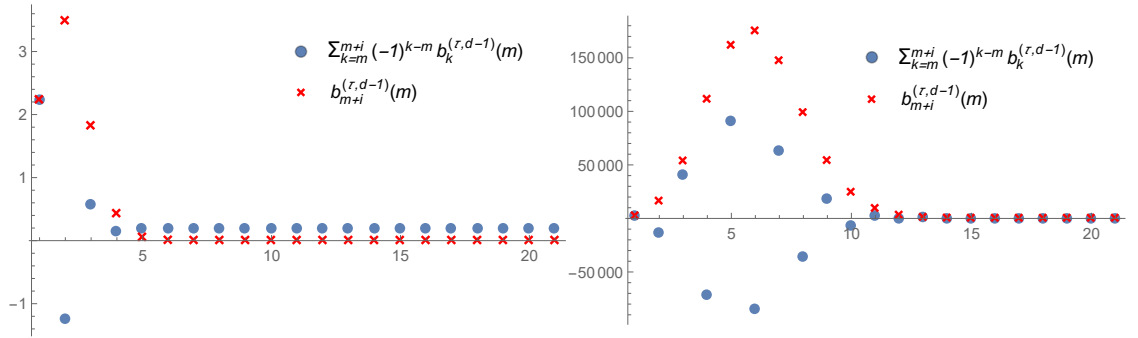


Figure 3.2: Plots of values $b_{m+i}^{(\tau, d-1)}(m)$ and of partial sums $\sum_{k=m}^{m+i} (-1)^{k-m} b_k^{(\tau, d-1)}(m)$ for $i = 0, \dots, 20$ and parameters for the left plot: $d = 3, \tau = 0.5, m = 2$ and for the right plot: $d = 3, \tau = 0.1, m = 13$.

One of the things immediately recognisable on Figure 3.2 is unimodality of terms

$b_k^{(\tau, d-1)}(m)$, which is obvious from calculations in the previous paragraph. Another noticeable thing is the size of the terms. Whereas on the left plot, where the time τ is not too small, values of terms $b_k^{(\tau, d-1)}(m)$ and partial sums are relatively small, we can see that on the right plot, where the time τ is smaller, values become really large, in our particular case they are of order 10^5 . We have to contrast this to the fact, that for the values of parameters in the right plot the corresponding coefficient $q_{13}^2(0.1)$ is approximately equal to 0.00688 which is of several orders of magnitudes smaller than the terms (and the partial sums) used to calculate it. When the difference between the size of terms and the correct value becomes too large, numerical inaccuracies quickly add up and numerically computed coefficients $q_m^{d-1}(\tau)$ are far from the correct values. This problem becomes even more apparent as the time τ goes to zero. Inspecting coefficients $q_m^{d-1}(\tau)$ suggests that the smallest time for which calculations of coefficients still seem to go through without major numerical errors is around $\tau = 0.1$ and the dimension d doesn't seem to affect this as much (increasing dimension actually extends possible times τ for which the algorithm works correctly).

We also need to analyse coefficients $q_m^{d-1}(\tau)$ for a fixed time τ and dimension d and see which coefficients are relevant for the simulation (i.e. which coefficients are not negligible). To get a better understanding we therefore plot on Figure 3.3 some of the coefficients for various times and dimensions.

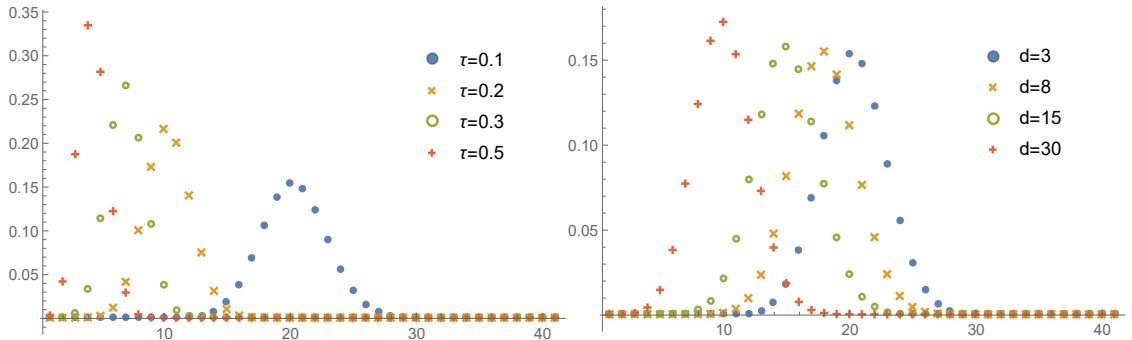


Figure 3.3: Plots of values $q_m^{d-1}(\tau)$ for $m = 0, \dots, 40$ and parameters for the left plot: $d = 3$ and times $\tau = 0.1, 0.2, 0.3, 0.5$ and for the right plot: $\tau = 0.1$ and dimensions $d = 3, 8, 15, 30$.

Again we notice unimodality, which in this case is not directly apparent from the representation (3.5). Additionally, when the time τ increases, the distribution moves to the left, which is intuitively clear from a description via death processes. Similarly, when the dimension d increases, distribution moves to the left, but the effect is less profound.

Since we are limited with how small time τ we can take i.e. $\tau \geq 0.1$, numerical calculations quickly show that the only relevant coefficients $q_m^{d-1}(\tau)$ are those with $m \leq 40$. All other coefficients will be essentially equal to zero. Of course for larger values of τ and a larger dimension d there might be even more essentially zero coefficients and there is no harm in setting them all equal to 0. Therefore, for a fixed time τ we will only need to pre-calculate each coefficient $q_m^{d-1}(\tau)$ for $m \leq 40$. We do this calculations by taking sufficiently many terms in representation (3.5) to get the desired accuracy. Additionally, we can stop calculations if we reach a coefficient $q_m^{d-1}(\tau)$ which is already small enough and late enough so that the coefficients have already started decreasing. Again, we can set all succeeding coefficients to 0. Once all relevant coefficients are calculated, we can then use the standard inversion sampling to simulate a random variable $A_\infty^{d-1}(\tau)$ and use it in Algorithm 2.

As we have seen our simulation algorithms do not work well for small values of τ . The bottleneck is always the simulation of $A_\infty^{d-1}(\tau)$. Hence, if we still want to use a variant of the algorithms, we need to resort to some kind of an approximation and one option is using the following normal approximation for $A_\infty^{d-1}(\tau)$, which was originally given in [Gri84]. Let $\beta = (d-2)\tau/2$ and $\eta = \frac{\beta}{e^\beta - 1}$. Then $A_\infty^{d-1}(\tau)$ is approximately distributed as a normal $N\left(\mu^{(d-1,\tau)}, (\sigma^{(d-1,\tau)})^2\right)$ random variable, where $\mu^{(d-1,\tau)} = \frac{2\eta}{t}$ and $(\sigma^{(d-1,\tau)})^2 = \frac{2\eta}{\tau}(\eta + \beta)^2 \left(1 + \frac{\eta}{\eta + \beta} - 2\eta\right) \beta^{-2}$. Alternatively, one can resort to using the approximation $\Theta \sim Q_{Approx}(\theta, \tau)$ and setting $X = \frac{1 - \cos(\Theta)}{2}$ instead of step 1 in Algorithms 1 and 3. Such an approximation actually seems to give better results than the normal approximation.

3.6 Brownian motion on projective spaces

Pick a field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$, where \mathbb{C} are the complex numbers and \mathbb{H} denotes the quaternions, consider \mathbb{F}^{n+1} as an $(n+1)$ -dimensional vector space over \mathbb{F} and recall that the n -dimensional projective space $\mathbb{F}\mathbb{P}^n$ is defined as a space of all 1-dimensional subspaces of \mathbb{F}^{n+1} . More precisely, define $\mathbb{F}\mathbb{P}^n$ to be the set of all equivalence classes $[x_0 : \dots : x_n]$, where the $(n+1)$ -tuple $(x_0, \dots, x_n)^\top$ is in $\mathbb{F}^{n+1} \setminus \{0\}$ and $[x_0 : \dots : x_n] = [y_0 : \dots : y_n]$ if and only if there exists a scalar $\lambda \in \mathbb{F} \setminus \{0\}$ such that $x_i = \lambda y_i$ for each $i = 0, \dots, n$. For any $x \in \mathbb{F}^{n+1} \setminus \{0\}$, $[x] := [x_0 : \dots : x_n]$ denotes the *homogeneous coordinates* of the corresponding point in $\mathbb{F}\mathbb{P}^n$.

Let $\pi: \mathbb{F}^{n+1} \setminus \{0\} \rightarrow \mathbb{F}\mathbb{P}^n$ be the quotient map given by $\pi(x) := [x]$. $\mathbb{F}\mathbb{P}^n$ is topologised by the quotient topology i.e. a subset $U \subseteq \mathbb{F}\mathbb{P}^n$ is open if and only if $\pi^{-1}(U)$ is open

in $\mathbb{F}^{n+1} \setminus \{0\}$. It is easy to see that $\mathbb{F}\mathbb{P}^n$ is Hausdorff, second-countable and compact since the restriction of π to the sphere \mathbb{S}^n (resp. \mathbb{S}^{2n+1} , \mathbb{S}^{4n+3}) if $\mathbb{F} = \mathbb{R}$ (resp. \mathbb{C} , \mathbb{H}), is surjective. Moreover, $\mathbb{F}\mathbb{P}^n$ is a smooth manifold. Indeed, for any $i = 0, \dots, n$, let $\tilde{U}_i = \{x \in \mathbb{F}^{n+1} ; x_i \neq 0\}$ and note that $\tilde{U}_i = \pi^{-1}(\pi(\tilde{U}_i))$, making the set $U_i = \pi(\tilde{U}_i)$ open in $\mathbb{F}\mathbb{P}^n$. Define the chart $\varphi_i: U_i \rightarrow \mathbb{F}^n$ by $\varphi_i([x]) = (\frac{x_0}{x_i}, \dots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \dots, \frac{x_n}{x_i})^\top$. The map φ_i is well-defined (since it assigns the same value to each $(n+1)$ -tuple in a given equivalence class in U_i) and a homeomorphism with the inverse $\varphi_i^{-1}(y_1, \dots, y_n) = [y_1 : \dots : y_{i-1} : 1 : y_i : \dots : y_n]$. Charts (U_i, φ_i) , $i = 0, \dots, n$ are smoothly compatible. Thus $\mathbb{F}\mathbb{P}^n$ is a compact smooth manifold of dimension n (resp. $2n$, $4n$) if $\mathbb{F} = \mathbb{R}$ (resp. \mathbb{C} , \mathbb{H}).

An equivalent way of defining the smooth structure on $\mathbb{F}\mathbb{P}^n$ is via the action on \mathbb{S}^n (resp. \mathbb{S}^{2n+1} , \mathbb{S}^{4n+3}) if $\mathbb{F} = \mathbb{R}$ (resp. \mathbb{C} , \mathbb{H}) of the multiplicative group of unit elements in \mathbb{F} , see proof of Lemma 3.6.1 below. This action is free, proper and smooth, so by the Quotient Manifold Theorem [Lee12, Thm 21.10] there exists a unique smooth structure on the projective spaces $\mathbb{R}\mathbb{P}^n = \mathbb{S}^n/\mathbb{S}^0$, $\mathbb{C}\mathbb{P}^n = \mathbb{S}^{2n+1}/\mathbb{S}^1$ and $\mathbb{H}\mathbb{P}^n = \mathbb{S}^{4n+3}/\mathbb{S}^3$ that makes the quotient projection π from a sphere onto $\mathbb{F}\mathbb{P}^n$ a smooth submersion i.e. its differential $d\pi_x$ is surjective at each point x in the sphere.

The spheres \mathbb{S}^n , \mathbb{S}^{2n+1} and \mathbb{S}^{4n+3} are Riemannian manifolds with their respective metrics induced by the ambient Euclidean spaces and the groups of unit elements $\mathbb{S}^0, \mathbb{S}^1$ and \mathbb{S}^3 act via isometries. Thus, the push-forward Riemannian metric on $\mathbb{F}\mathbb{P}^n$ is well-defined and turns the projection π into a Riemannian submersion. Put differently, the sphere \mathbb{S}^n (resp. \mathbb{S}^{2n+1} , \mathbb{S}^{4n+3}) is a smooth fibre bundle over $\mathbb{R}\mathbb{P}^n$ (resp. $\mathbb{C}\mathbb{P}^n$, $\mathbb{H}\mathbb{P}^n$) with the fibre diffeomorphic to \mathbb{S}^0 (resp. \mathbb{S}^1 , \mathbb{S}^3) and the differential $d\pi_x$ is an isometry when restricted to the space of horizontal tangent vectors at any x in the sphere and the tangent space at $\pi(x)$ in the projective space. This statement is trivial when $\mathbb{F} = \mathbb{R}$, since in that case we have $\mathbb{S}^0 = \{-1, 1\}$ and π is a local diffeomorphism. The other cases are given in [Lee97, Problem 3-8]. If $\mathbb{F} = \mathbb{C}$ this construction yields the well-known Fubini-Study metric on $\mathbb{C}\mathbb{P}^n$ [Jos17, Ch. 7.1].

The BM on $\mathbb{F}\mathbb{P}^n$ can be defined as a strong Markov process with the generator equal to $\frac{1}{2}\Delta_{\mathbb{F}\mathbb{P}^n}$, where $\Delta_{\mathbb{F}\mathbb{P}^n}$ is the Laplace-Beltrami operator corresponding to the Fubini-Study metric on $\mathbb{F}\mathbb{P}^n$ [Hsu02, p.74]. Finally, we establish a connection between the spherical BM and the BM on projective spaces.

Lemma 3.6.1. *Let $n \geq 1$ and let Z be the BM on the sphere \mathbb{S}^n (resp. \mathbb{S}^{2n+1} , \mathbb{S}^{4n+3}) started at z . Then the process $\pi(Z)$ is the BM on $\mathbb{R}\mathbb{P}^n$ (resp. $\mathbb{C}\mathbb{P}^n$, $\mathbb{H}\mathbb{P}^n$) started at $\pi(z)$.*

Proof. Fibres of the Riemannian submersion π are orbits in $\mathbb{S}^n \subseteq \mathbb{R}^{n+1}$ (resp. $\mathbb{S}^{2n+1} \subseteq \mathbb{C}^{n+1}$, $\mathbb{S}^{4n+3} \subseteq \mathbb{H}^{n+1}$) of the (right) group action $((z_0, \dots, z_n)^\top, \lambda) \mapsto (z_0\lambda, \dots, z_n\lambda)^\top$,

where $(z_0, \dots, z_n)^\top \in \mathbb{S}^n$ (resp. $\mathbb{S}^{2n+1}, \mathbb{S}^{4n+3}$) and λ is an element of the multiplicative group of unit elements in \mathbb{R} (resp. \mathbb{C}, \mathbb{H}). This group is isomorphic to \mathbb{S}^0 (resp. $\mathbb{S}^1, \mathbb{S}^3$) and isometric to the submanifold $\{(z_0, \dots, z_n)^\top \in \mathbb{S}^n$ (resp. $\mathbb{S}^{2n+1}, \mathbb{S}^{4n+3}$); $z_1 = \dots = z_n = 0\}$ in \mathbb{S}^n (resp. $\mathbb{S}^{2n+1}, \mathbb{S}^{4n+3}$). This submanifold is easily seen to be totally geodesic i.e. any geodesic in the submanifold is also a geodesic in the ambient manifold. Moreover, each fibre (i.e. orbit) can be isometrically mapped onto this submanifold by multiplication from the left by a suitable element of special orthogonal group (for any element x in the fibre pick any matrix $O(x)$ in the special orthogonal group such that $O(x)x = (1, 0, \dots, 0)^\top$). Hence all fibres are totally geodesic submanifolds.

Let (M, g) be a k -dimensional Riemannian submanifold of a Riemannian manifold $(\widetilde{M}, \widetilde{g})$, i.e. the metric g is a restriction of \widetilde{g} to the tangent bundle of M . For any two vector fields X, Y on M , the second fundamental form II is given by $II(X, Y) = \widetilde{\Delta}_{\widetilde{X}}\widetilde{Y} - \Delta_X Y$, where Δ (resp. $\widetilde{\Delta}$) represent the metric connections on M (resp. \widetilde{M}) and the vector fields $\widetilde{X}, \widetilde{Y}$ are arbitrary extensions of the vector fields X, Y to \widetilde{M} . The second fundamental form is well-defined and takes values in the normal bundle of the submanifold M (see [Lee97, Thm 8.2]). By [Lee97, Exercise 8.4], the submanifold M is totally geodesic if and only if $II(X, Y) = 0$ for any vector fields X, Y on M . The mean curvature H_x at any $x \in M$ is proportional to the metric trace of II , i.e. for any orthonormal frame u_1, \dots, u_k in the neighbourhood of x we have $H_x = \frac{1}{k} \sum_{i=1}^k II(u_i, u_i)_x$. Clearly, H_x is equal to 0 for each $x \in M$ if M is totally geodesic in \widetilde{M} .

By [Pau90, Thm 1], the projection $\pi(Z)$ in $\mathbb{R}\mathbb{P}^n$ (resp. $\mathbb{C}\mathbb{P}^n, \mathbb{H}\mathbb{P}^n$) of the BM Z on \mathbb{S}^n (resp. $\mathbb{S}^{2n+1}, \mathbb{S}^{4n+3}$) is a BM with the drift given by $V_x = -\frac{n}{2}d\pi_x(H_x)$ (resp. $-\frac{2n}{2}d\pi_x(H_x), -\frac{4n}{2}d\pi_x(H_x)$), where H_x is the mean curvature of the fibre of x in the ambient sphere. Since all the fibres are totally geodesic, the drift vanishes and the lemma follows.

Note that once we have shown that the mean curvature is equal to zero, we could have alternatively deduced the final result more directly by using known results on projections of Brownian motions via Riemannian submersions (see e.g. [Elw82, Thm. 10E]). \square

Chapter 4

Lévy processes on smooth manifolds with a connection

In this chapter based on the preprint [MM20] we define a Lévy process on a smooth manifold M with a connection as a projection of a solution of a Marcus stochastic differential equation on the holonomy bundle of M , driven by a holonomy-invariant Lévy process on the Euclidean space. On a Riemannian manifold, our definition (with Levi-Civita connection) generalizes the Eells-Elworthy-Malliavin construction of Riemannian Brownian motion and extends the class of isotropic Lévy process introduced in Applebaum and Estrade [AE00]. On a Lie group with a surjective exponential map, our definition (with left-invariant connection) coincides with the classical definition of a (left) Lévy process given in terms of its increments.

The main theorem characterizes the class of Lévy processes via their generators on M , generalizing the fact that the Laplace-Beltrami operator generates Brownian motion on a Riemannian manifold. Its proof requires a pathwise construction of a stochastic horizontal lift and anti-development of a discontinuous semimartingale, leading to a generalization of results in [PE92] to smooth manifolds with non-unique geodesics between distinct points.

4.1 Introduction

Lévy processes on the Euclidean space \mathbb{R}^d are fundamental and well-studied stochastic processes with càdlàg (right continuous with left limits) paths, characterized by the independence and stationarity of their increments (see e.g. monographs [Ber96, Sat99, Kyp14] and the references therein). Lévy processes have a natural generalization to

Lie groups introduced in [Hun56] and have been studied extensively over the last half a century (see monograph [Lia04] for a modern account). The definition relies on the group structure of the state space, which allows to define a natural notion of the increment of a process. Since there is no natural notion of an increment on a Riemannian manifold M , [AE00] uses a geodesics-based approach to define isotropic Lévy processes on M . However, these two definitions are not entirely compatible: for example, if M is a d -dimensional torus with a flat metric, isotropic Lévy processes form a strict subclass of Lévy processes on M , defined via its Lie group structure. The same phenomenon may be observed on the Euclidean space \mathbb{R}^d or indeed any Lie group equipped with a left-invariant Riemannian metric.

The main contribution of this chapter is twofold. **(I)** By relaxing key restrictions from the definition of isotropic Lévy process on Riemannian manifolds from [AE00], we extend the definition of Lévy processes to smooth manifolds equipped with a connection. A fundamental property of Lévy processes is that they evolve and jump in the “same way” from any point in the state space. Our generalization is natural, since it is a connection (and not a metric) that is a required geometric structure for the comparison of jumps along geodesics at different points on the manifold. Moreover, a canonical connection is often available when there is no canonical metric. For example the class of Lévy processes induced by the left-invariant connection on a Lie group (with a surjective exponential map) coincides with the class of the classical (left) Lévy process. Furthermore, our definition with the Levi-Civita connection on a Riemannian manifold extends the class of isotropic Lévy processes, closing the aforementioned compatibility gap. **(II)** Our main theorem (Theorem 4.2.14 below) characterizes the law of a Lévy process on a smooth manifold with a connection via its infinitesimal generator, generalizing the classical result that the law of a Brownian motion on a Riemannian manifold is characterized by the Laplace-Beltrami operator.

Our definition generalizes the celebrated Eells-Elworthy-Malliavin construction of a Brownian motion on a d -dimensional Riemannian manifold, given as a projection of a horizontal Brownian motion in the orthonormal frame bundle. The horizontal Brownian motion is defined as a solution of a Stratonovich stochastic differential equation (SDE) on the orthonormal frame bundle, with the integrator a standard Brownian motion on \mathbb{R}^d and the coefficients given by the fundamental horizontal vector fields. Consider a smooth manifold M with a connection and a corresponding holonomy bundle (a subbundle of the frame bundle of M). We define a Lévy process X on M to be the projection of a horizontal Lévy process U , defined via an SDE on a holonomy bundle with the integrator a Lévy process Y on \mathbb{R}^d and the coefficients given by the fundamental horizontal vector

fields in the tangent bundle of the holonomy bundle (see Section 4.2 below). Since Y is typically discontinuous, we use the theory of Marcus SDEs [Mar81, KPP95, Fuj91, AK93] in order to obtain the correct transformation properties under the change of coordinates on a smooth manifold, not possessed by the Stratonovich integral in the presence of jumps.

A Lévy process on \mathbb{R}^d is a time-homogeneous Markov process. In order to ensure that the time-homogeneous Markov property continues to hold for our process X on M , the horizontal process U on the holonomy bundle has to be (a) time-homogeneous Markov and (b) invariant under the natural action of the corresponding holonomy group. Since U is a strong solution of a Marcus SDE with the integrator Y , requirement (a) forces the process Y to be a Lévy process on \mathbb{R}^d [JP91, Est97]. Condition (b) is equivalent to the generating triplet of Y being invariant under the action of a holonomy group on \mathbb{R}^d (see (4.5) below).

Holonomy bundle and the corresponding holonomy group are not uniquely determined by the connection on a connected smooth manifold M . However, all possible choices of holonomy bundles are diffeomorphic and the corresponding holonomy groups are conjugate Lie subgroups of the group of invertible matrices in $\mathbb{R}^d \otimes \mathbb{R}^d$. Thus, the choice of the holonomy bundle and the corresponding holonomy group does not affect the class of all Lévy processes on M described by our definition (see Proposition 4.2.9 below).

The characterization of the law of a Lévy process X via its infinitesimal generator requires a construction of a stochastic horizontal lift U in the holonomy bundle and a stochastic anti-development Y in \mathbb{R}^d , such that U solves the Marcus SDE driven by Y and is mapped into X by the bundle projection. The jumps of the solution U of the Marcus SDE are given via the flow of the horizontal vector fields, making X jump along the geodesics in M . The main technical difficulty in constructing a stochastic horizontal lift U from the generator of X is that a jump of X may be induced by many different jumps of U . This issue disappears when a manifold M possesses a unique geodesic between any two points, as is assumed in [PE92], since then the structure of the Marcus SDE implies that the jumps of X and U are in one-to-one correspondence. However, this geometric assumption is quite restrictive as it excludes for example all compact Riemannian manifolds. More precisely, compact Riemannian manifolds are bounded as metric spaces and on the other hand geodesically complete by the Hopf-Rinow theorem, which would imply unboundedness if the unique geodesic assumption were in place. To circumvent this issue in a general setting, we first append compatible “jump data” to the process X . We can then prove that X and the “jump data” together uniquely determine the horizontal lift U and the anti-development Y , see Theorem 4.2.19 below. Moreover,

it is possible to construct suitable “jump data” so that the anti-development Y of the process X is precisely a Lévy process on \mathbb{R}^d with the generating triplet induced by the generator of X .

On Riemannian manifolds it is natural to consider the Levi-Civita connection which is uniquely determined by the metric. A holonomy group of this connection is a Lie subgroup of the orthogonal group implying that our new definition of Lévy processes is a genuine generalization of isotropic Lévy processes defined in [AE00] (see Section 4.3.1 below for details). In particular, since the holonomy group of the flat connection on \mathbb{R}^d is trivial, our definition coincides with the classical definition of Lévy processes on \mathbb{R}^d . Lie groups do not necessarily possess a natural Riemannian metric, but have a natural choice of a left-invariant connection. In Section 4.3.2 below we show that this choice of a connection implies that a Lévy process satisfying our definition is a classical (left) Lévy process on a Lie group. If the exponential map of the Lie group is not surjective, then jumps which are not in the image of the exponential map cannot be expressed via the geodesics of the left-invariant connection, so (left) Lévy process having such jumps cannot be expressed via our definition. However, as soon as all jumps can be expressed in such a way, the (left) Lévy process can be expressed via our new definition (see Proposition 4.3.2). In particular, on the Lie group with surjective exponential map the two notions of Lévy processes coincide.

The remainder of the chapter is structured as follows. In Section 4.2 we apply the theory of Marcus SDEs on manifolds to construct a Lévy process on a smooth manifold with a connection and state our main results. In Section 4.3 we discuss examples of Lévy processes on manifolds with a connection, including special cases of Riemannian manifolds and Lie groups as well as new examples with exceptional holonomy groups. The proofs of our main results are given in Section 4.4.

4.2 Construction and characterization of Lévy processes on manifolds with a connection

Let M be a connected smooth d -dimensional manifold without a boundary. Recall the following notations from Chapter 1: $C^\infty(M)$ denotes the space of smooth real valued functions on M , TM is the tangent bundle on M , $\Gamma(TM)$ denotes the vector fields on M (i.e. smooth sections of the tangent bundle). For a predictable stopping time τ a càdlàg M -valued stochastic process $X = (X_t)_{t \in [0, \tau]}$ is called a **semimartingale on a stochastic interval** $[0, \tau]$ if $f(X^{\tau_n})$ is a real-valued semimartingale for any $f \in C^\infty(M)$

and $n \in \mathbb{N}$, where $X^{\tau_n} = (X_t^{\tau_n})_{t \in \mathbb{R}_+}$ with $X_t^{\tau_n} = X_{\min\{\tau_n, t\}}$ for $t \in \mathbb{R}_+$ denotes a stopped process and stopping times $\{\tau_n ; n \in \mathbb{N}\}$ form an announcing sequence of τ , i.e. $\tau_n \uparrow \tau$ as $n \rightarrow \infty$ and $\tau_n < \tau$ a.s. for every $n \in \mathbb{N}$. See [Pro92, JS03] as a general reference for standard (real) semimartingales and predictable stopping times. One easily checks using Itô's formula that if $M = \mathbb{R}^d$, this definition (with $\tau \equiv \infty$) is equivalent to the standard definition of a semimartingale.

The corresponding generalization of Lévy processes is not as straightforward since on a general smooth manifold there is no suitable notion of an increment. An alternative approach, also used in [AE00], exploits the fact that points of a smooth manifold possess tangent spaces, which are real vector spaces providing a local linearization of the manifold. Each tangent space is isomorphic to \mathbb{R}^d , making it feasible to “develop” a Lévy process on a tangent space onto the manifold. Since at different points in M the tangent spaces are not canonically isomorphic, for such a construction to work we need to be able to use curves in M to “connect & compare” different tangent spaces. This highlights the need for a (linear) connection on M . Our main aim is to show that this geometric structure is sufficient for a Lévy process on M to be defined, circumventing the need for a Riemannian metric used in [AE00]. There is an intermediate step – we need to consider a horizontal process on the holonomy bundle, which is a solution of a specific Marcus SDE. Thus, we essentially need to generalize results from Section 1.4 to discontinuous semimartingales and the construction of Lévy processes will then be analogous to the Eells-Elworthy-Malliavin construction of the Riemannian Brownian motion from Section 1.5.

In Section 4.2.1 we recall basic properties of Marcus SDEs on manifolds. In Section 4.2.2 we apply Marcus SDEs to define a horizontal Lévy process on the holonomy bundle, which is then projected to a Lévy process on a manifold with a connection. Section 4.2.3 describes our class of Lévy processes via generators on M .

4.2.1 Marcus SDE

The Stratonovich integral which is used for SDEs on smooth manifolds with continuous semimartingale integrators does not obey the classical chain rule formula when the integrator has jumps. The first theory of SDEs with jumps on manifolds was introduced in [Rog81] and allowed integration against Brownian motion and Poisson random measures with quite general jump coefficients. A more modern theory with the same class of integrators is studied in [Kun19] and also deals with stochastic flows of diffeomorphisms. Technically, such an approach with a suitable choice of a jump coefficient could be used for our Definition 4.2.8 of Lévy processes on manifold where the integrator is a Euclidean

Lévy process. However, we also wish to consider more general càdlàg semimartingale integrators, especially in the context of stochastic horizontal lifts and stochastic anti-developments (see Definition 4.2.15 and Theorem 4.2.19).

Hence, we will consider a unified approach and utilise a more general notion of a Marcus (or canonical) SDE which was introduced in the Euclidean spaces in [Mar81] and was later generalized (still in Euclidean spaces) in [KPP95] and allows integration against arbitrary càdlàg semimartingales. Due to correct change-of-variable properties Marcus SDE arises naturally in the context of smooth manifolds [Fuj91, AK93, AE00]. In order to describe the generalization of Marcus SDE to smooth manifolds required in this chapter, we need to use the notions of (complete) vector fields and flow exponential map as described in Section 1.1.

Definition 4.2.1. Let M be a smooth manifold, let $V_1, \dots, V_\ell \in \Gamma(TM)$ be smooth vector fields on M and let W be a càdlàg \mathbb{R}^ℓ -valued semimartingale. We call an M -valued semimartingale X defined on a stochastic interval $[0, \tau)$ a **solution of a Marcus SDE**

$$dX_t = V_i(X_{t-}) \diamond dW_t^i \quad (4.1)$$

if for any $f \in C^\infty(M)$ it satisfies the equation

$$\begin{aligned} f(X_t) - f(X_0) &= \int_0^t V_i f(X_{s-}) \circ dW_s^i \\ &+ \sum_{0 < s \leq t} \left(f(\exp(V_i \Delta W_s^i)(X_{s-})) - f(X_{s-}) - V_i f(X_{s-}) \Delta W_s^i \right) \end{aligned} \quad (4.2)$$

for $0 \leq t < \tau$, where Itô's circle \circ denotes the Stratonovich integral, $\Delta W_s = W_s - W_{s-}$ and \exp is the flow exponential map.

Remark 4.2.2. (i) A more precise meaning of (4.2) is as follows. Consider an announcing sequence $\{\tau_n ; n \in \mathbb{N}\}$ of τ . We require that for every $n \in \mathbb{N}$ the stopped process X^{τ_n} is a semimartingale and that equation (4.2) with X exchanged by X^{τ_n} holds for every $t \in \mathbb{R}_+$.

(ii) The sum in the defining property (4.2) of a Marcus SDE can be interpreted as follows: the process X jumps only if the integrator W jumps and then the jump of X occurs along an integral curve of a certain vector field, constructed as a linear combination of the driving vector fields with the coefficients given by the components of the jump of the integrator (see Figure 4.1 below).

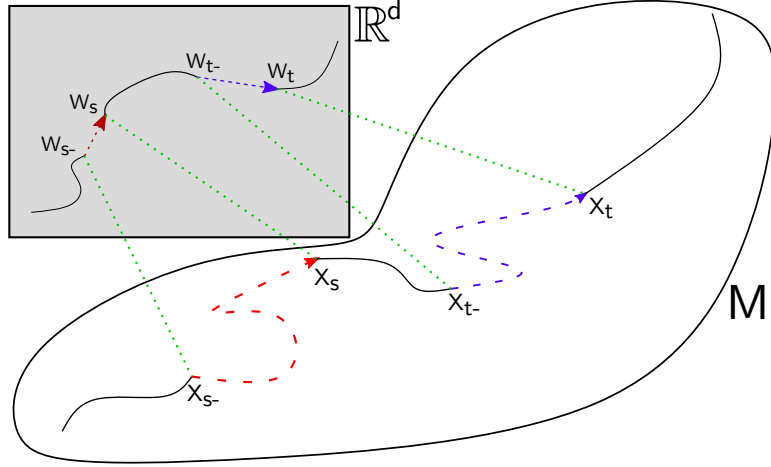


Figure 4.1: Relation between the solution X and the integrator W of the Marcus SDE

Proposition 4.2.3. *Suppose that for every $a = a^i e_i \in \mathbb{R}^d$ the vector field $a^i V_i$ is complete. Then there exists a unique solution X of Marcus SDE (4.1) and it is a semimartingale defined on a maximal stochastic interval $[0, \tau)$, where τ is a predictable stopping time which is an explosion time¹.*

Remark 4.2.4. A solution process X of Marcus SDE (4.1) exists even when vector fields are not complete. However, then the stopping time τ is not necessarily predictable and it might happen that X “jumps” to the cemetery point ∂ and does not reach it continuously. To avoid these issues we will insist on the completeness assumption.

Since all vector fields on compact manifolds are complete, we immediately obtain the following corollary.

Corollary 4.2.5. *If M is compact, then the solution X of Marcus SDE (4.1) is a semimartingale defined on the whole interval $[0, \infty)$.*

A more general type of SDEs on manifolds, which includes (4.1), was considered in [Coh96a, Coh96b]. But for the sake of completeness and due to its simplicity, we give a direct proof of Proposition 4.2.3 in Section 4.4.1 below.

¹If we consider a one-point compactification $\widehat{M} := M \cup \{\partial\}$ this means that on the event $\{\tau < \infty\}$ as $t \rightarrow \tau$ we have $X_t \rightarrow \partial$ or equivalently X exits all compacts in M .

4.2.2 Horizontal Lévy processes and Lévy processes on a smooth manifold with a connection

Recall from Section 1.3 the definition of the frame bundle $F(M)$ and the horizontal vector fields $H_e, e \in \mathbb{R}^d$ induced by the connection. In order to be able to use Proposition 4.2.3 and avoid the non-predictability issues outlined Remark 4.2.4, we will assume that all the vector fields $H_e, e \in \mathbb{R}^d$ are complete, which is equivalent to geodesic completeness of the manifold M (see paragraph preceding equation (1.6) for relation between geodesics and integral curves of fundamental horizontal vector fields). Recall also the notation $H_i = H_{e_i}$ for $i = 1, \dots, d$, where e_1, \dots, e_d is the standard basis of \mathbb{R}^d .

Definition 4.2.6. Let M be equipped with a connection which turns it into a geodesically complete manifold. An $F(M)$ -valued semimartingale U defined on a stochastic interval $[0, \tau)$ is called a **horizontal Lévy process** if there exists an \mathbb{R}^d -valued Lévy process Y (see [Sat99, Def. 1.6] for a definition of \mathbb{R}^d -valued Lévy processes) such that U is a solution of a Marcus SDE

$$dU_t = H_i(U_{t-}) \diamond dY_t^i, \quad U_0 = u \in F(M), \quad (4.3)$$

on the maximal (stochastic) interval $[0, \tau)$.

Remark 4.2.7. (i) Horizontal Lévy processes were originally introduced in [App95] on the orthonormal bundle of a Riemannian manifold.

(ii) We require geodesic completeness so that the flow exponential map is well defined on complete vector fields $H_e, e \in \mathbb{R}^d$. Proposition 4.2.3 then shows that a horizontal Lévy process U is uniquely determined by the Lévy process Y and is defined on the maximal stochastic interval $[0, \tau)$, where the assumption on geodesic completeness guarantees that τ is predictable.

(iii) If M is a Riemannian manifold, then horizontal Brownian motion introduced in Section 1.5 is also a horizontal Lévy process.

In the next step, we use the bundle projection π and consider the projected process $X := \pi(U)$ which is a candidate for a Lévy process on M . Since the integrator Y of Marcus SDE (4.3) is a Lévy process, the horizontal Lévy process U is always a strong Markov process [KPP95, Thm. 5.1], but X in general need not be. More precisely, if $X_0 = p$ we only know that U_0 is a frame in the fibre above p , i.e. $U_0 = u$ such that $\pi(u) = p$. But U_0 might be a different frame v in the same fibre. The future evolution of

U , and hence of X , could be vastly different in these two cases, $U_0 = u$ or v . However, since u and v are in the same fibre, there exists $g \in \text{GL}(d)$ such that $u = vg$. With this in mind, consider the process $\tilde{U} := Ug$, which by Lemma 4.4.3 below solves the Marcus SDE

$$d\tilde{U}_t = (R_g)_* H_i(\tilde{U}_{t-}) \diamond dY_t^i,$$

where push-forward vector fields $(R_g)_* H_i$ are given in (4.21) below. Moreover, by Lemma 1.3.2, we have $(R_g)_* H_e = H_{g^{-1}e}$ for every $e \in \mathbb{R}^d$. Thus, linearity of $e \mapsto H_e$ and Lemma 4.4.1 imply that \tilde{U} is also a solution of Marcus SDE

$$d\tilde{U}_t = H_i(\tilde{U}_{t-}) \diamond d\tilde{Y}_t^i$$

with $\tilde{U}_0 = ug = v$ where $\tilde{Y} := g^{-1}Y$. Consider a solution V of Marcus SDE (4.3) with the initial condition V_0 , which is driven by Y . Therefore, if Y and \tilde{Y} were to have the same law, then (weak) uniqueness of solutions of Marcus SDEs shows that \tilde{U} and V also have the same law, thus implying that the processes $\pi(U) = \pi(\tilde{U})$ and $\pi(V)$ are equal in law. Hence by assuming the appropriate invariance of the integrator Y we may use the result of Dynkin on transformations of Markov processes [Dyn65, Thm. 10.13] which ensures the Markovianity of the projected process X . Fortunately, as we will see below, we do not have to check such an invariance for every $g \in \text{GL}(d)$, but only for elements of a specific subgroup – the holonomy group, which we now describe.

For any frame $u \in F(M)$, a connection on M allows us to reduce the frame bundle $F(M)$ to a subbundle $P(u)$ known as the holonomy bundle. The **holonomy bundle** $P(u)$ is a principal fibre bundle over M with a structure group given by the **holonomy group** $\text{Hol}(u)$, a Lie subgroup of $\text{GL}(d)$ [KN63, Thm. 4.2 in Ch. II]. More precisely, a holonomy bundle $P(u)$ consists of all the frames in $F(M)$ connected to u via horizontal curves in $F(M)$ (i.e. curves with velocity vectors in horizontal spaces). The holonomy group measures how tangent vectors transform after they are parallel transported along loops. An example of a manifold with a non-trivial holonomy group is a sphere and it is shown on Figure 4.2 how a tangent vector transforms (rotates) into a different vector after parallel transport along geodesics from A to N to B to A .

The essential property for our definition of Lévy processes is that the fundamental horizontal vector fields H_e , $e \in \mathbb{R}^d$, are tangent to the subbundle $P(u)$ (this follows by the definition of the holonomy bundle via horizontal curves), hence they can be interpreted as vector fields on $P(u)$. By Proposition 4.2.3, any horizontal Lévy process U (a solution of Marcus SDE (4.3)) with $U_0 \in P(u)$ stays in $P(u)$ for all times for which it is defined.

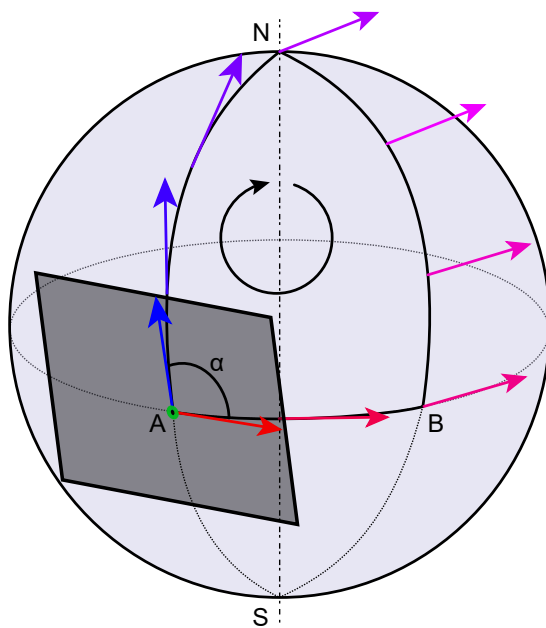


Figure 4.2: Parallel transport and non-trivial holonomy group of the sphere (adjusted from [Wik20b])

Hence to use [Dyn65, Thm 10.13] which ensures that the projection $X = \pi(U)$ is Markov, it is sufficient to check that the integrator Y is $\text{Hol}(u)$ -invariant, i.e. the processes Y and gY are equal in law for every $g \in \text{Hol}(u)$ or equivalently the characteristics of Y satisfy (4.5) below.

Definition 4.2.8. Let Y be a $\text{Hol}(u)$ -invariant Lévy process on \mathbb{R}^d and let $U = (U_t)_{t \in [0, \tau]}$ be a horizontal Lévy process on the holonomy bundle $P(u)$ that solves Marcus SDE (4.3) with $U_0 \in P(u)$ a.s. Then with π the bundle projection as above, $X = \pi(U)$ is said to be a **Lévy process on M** .

Proposition 4.2.9. *A process X on M satisfying Definition 4.2.8 is a Markov process and is a semimartingale on $[0, \tau)$. The class of all Lévy processes on a connected smooth manifold M is invariant under the choice of the frame $u \in F(M)$ and the corresponding holonomy bundle $P(u)$.*

Remark 4.2.10. (i) The holonomy bundle is a reduction of the frame bundle (see [KN63] for definition of a reduction). In fact, it is the smallest possible reduction compatible with the connection [RS17, Remark 1.7.14]. Hence we cannot in general reduce the holonomy group invariance of the Lévy integrator any further and the class of

Lévy processes on the smooth manifold M with a connection from Definition 4.2.8 is the largest possible.

- (ii) Due to a particular form of Marcus SDE (4.3), we know that for any $t < \tau$ the equality $U_t = \exp(H_i \Delta Y_t^i)(U_{t-})$ holds, where \exp is a flow exponential map (see Section 4.2.1). Then equation (1.6) shows that $X_t = \text{Exp}_{X_{t-}}(U_{t-} \Delta Y_t)$ holds for any $t < \tau$, where Exp_p denotes a geodesic exponential map based at a point $p \in M$. Therefore, jumps of X occur only along geodesics.
- (iii) In the special case of Riemannian manifolds, [Moh04, Sec. 3] considered a holonomy bundle reduction and obtained similar results to Proposition 4.2.9. More details about the Riemannian case are given in Section 4.3.1.
- (iv) Note that if a Lévy process Y is adapted to filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$, then also the horizontal Lévy process U and the Lévy process X on M are adapted to the same filtration. However, even when we consider the natural filtration $(\mathcal{F}_t^Y)_{t \in \mathbb{R}_+}$ of Y , it could happen that the natural filtrations of U and X are strictly smaller than $(\mathcal{F}_t^Y)_{t \in \mathbb{R}_+}$.

4.2.3 Generators of Lévy processes on manifolds with a connection

In this section we show a characterisation of Lévy processes on smooth manifolds via their generators. The first step is to analyse the generators of horizontal Lévy processes.

Lemma 4.2.11. *The (extended) generator \mathcal{L}_U of a horizontal Lévy process U , which solves Marcus SDE (4.3), is given by*

$$\begin{aligned} \mathcal{L}_U(f)(v) &= H_b f(v) + \frac{1}{2} a^{ij} H_i H_j f(v) \\ &\quad + \int_{\mathbb{R}^d \setminus \{0\}} (f(\exp(H_x)(v)) - f(v) - \mathbf{1}(|x| < 1) x^i H_i f(v)) \nu(dx) \end{aligned} \quad (4.4)$$

for any frame $v \in P(u)$, where (a, ν, b) is a generating triplet [Sat99, p. 65] of a Lévy integrator Y associated with a cut-off function $\chi(x) = \mathbf{1}(|x| < 1)$ and the domain of \mathcal{L}_U includes smooth functions on the holonomy bundle $P(u)$.

Recall that a Lévy process Y is $\text{Hol}(u)$ -invariant if the laws of processes Y and gY are the same for every $g \in \text{Hol}(u)$. This holds if and only if the generating triplet (a, ν, b)

of Y satisfies

$$\begin{aligned}
& \nu \text{ is } g\text{-invariant (i.e. } g_*\nu = \nu); \\
& gag^\top = a; \\
& gb = b + \int_{\mathbb{R}^d \setminus \{0\}} (\mathbf{1}(|g^{-1}x| < 1) - \mathbf{1}(|x| < 1)) \nu(dx)
\end{aligned} \tag{4.5}$$

for every $g \in \text{Hol}(u)$, where the push-forward measure $g_*\nu$ is defined by $g_*\nu(A) := \nu(g^{-1}(A))$ for every Borel measurable set $A \in \mathfrak{B}(\mathbb{R}^d)$, a and g are interpreted as matrices in $\mathbb{R}^d \otimes \mathbb{R}^d$ and $^\top$ denotes matrix transpose. Note that when $\text{Hol}(u)$ is a subgroup of the orthogonal group $O(d)$, the condition on the drift reduces to $gb = b$.

Since Lévy process X is given as a projection of a horizontal Lévy process U , its generator is a “push-forward” of the generator given by (4.4) and invariance conditions (4.5) ensure it is indeed a well-defined operator.

Proposition 4.2.12. *The generator \mathcal{L}_X of a Lévy process X on a smooth manifold M with a connection is given by*

$$\begin{aligned}
\mathcal{L}_X(f)(p) &= H_b(f \circ \pi)(v) + \frac{1}{2} a^{ij} H_i H_j (f \circ \pi)(v) \\
&+ \int_{\mathbb{R}^d \setminus \{0\}} (f(\text{Exp}_p(vx)) - f(p) - \mathbf{1}(|x| < 1) x^i H_i (f \circ \pi)(v)) \nu(dx), \tag{4.6}
\end{aligned}$$

where v is any frame (in the holonomy bundle $P(u)$) with $\pi(v) = p$ and Exp_p is a geodesic exponential map based at p . In particular, (4.6) does not depend on the choice of v . The domain of the generator \mathcal{L}_X comprises functions f on M such that $f \circ \pi$ is in the domain of the generator \mathcal{L}_U .

Remark 4.2.13. In some special cases the formula in (4.6) can be simplified.

- (i) Suppose that $-I_d \in \text{Hol}(u)$, where I_d is the identity matrix in $\mathbb{R}^d \otimes \mathbb{R}^d$. Then $-v \in P(u)$, and hence $H_i(f \circ \pi)(-v) = -H_i(f \circ \pi)(v)$, and (4.5) implies $b = 0$.

Assume further that $a = 0$. Then by (4.6) applied with v and $-v$ yields

$$\begin{aligned}\mathcal{L}_X(f)(p) &= \frac{1}{2} \int_{\mathbb{R}^d \setminus \{0\}} (f(\text{Exp}_p(vx)) - f(p) - \mathbf{1}(|x| < 1)x^i H_i(f \circ \pi)(v)) \nu(dx) \\ &\quad + \frac{1}{2} \int_{\mathbb{R}^d \setminus \{0\}} (f(\text{Exp}_p(-vx)) - f(p) + \mathbf{1}(|x| < 1)x^i H_i(f \circ \pi)(v)) \nu(dx) \\ &= \frac{1}{2} \int_{T_p M} (f(\text{Exp}_p(y)) - 2f(p) + f(\text{Exp}_p(-y))) \nu_p(dy),\end{aligned}\tag{4.7}$$

where the push-forward measure $\nu_p := v_*\nu$ is independent of v by (4.5), implying the expression for the jump part of the generator obtained in [ASB20]. There are many examples of Riemannian manifolds with a Levi-Civita connection where $-I_d \in \text{Hol}(u)$, e.g. when d is even and $\text{Hol}(u) = \text{SO}(d)$ is the special orthogonal group. There also exist manifolds with a connection not induced by the Riemannian metric such that $-I_d \in \text{Hol}(u)$, see Section 4.3.3.

- (ii) In the case the stochastic anti-development Lévy process Y is $\text{O}(d)$ -invariant in the sense of (4.5), the generating triplet equals $(\alpha I_d, \nu, 0)$ for some $\alpha \geq 0$ and an $\text{O}(d)$ -invariant Lévy measure ν . If M is a Riemannian manifold equipped with the Levi-Civita connection, the holonomy bundle is a subbundle of the orthonormal frame bundle and the horizontal Laplace operator satisfies $\sum_{i=1}^d H_i^2(f \circ \pi)(v) = \Delta_M f(p)$ for any orthonormal frame v in the fibre above p , where Δ_M is the Laplace-Beltrami operator on M . Thus (4.6) and (4.7) imply

$$\mathcal{L}_X(f)(p) = \frac{\alpha}{2} \Delta_M f(p) + \frac{1}{2} \int_{T_p M} (f(\text{Exp}_p(y)) - 2f(p) + f(\text{Exp}_p(-y))) \nu_p(dy),$$

which equals the generator of an isotropic Lévy process in [AE00], where the improper integrals in the jump part were used implicitly in [AE00, Eq. (3.6)].

For the formula (4.6) to be of use, it is crucial to prove that any Markov process X on M with the generator \mathcal{L}_X is indeed a Lévy process on M satisfying Definition 4.2.8. Put differently, we need to construct an \mathbb{R}^d -valued Lévy process Y such that X can be represented via the Marcus SDE in Definition 4.2.8. Typically, we will not be able to construct Y on the same probability space on which X is defined and we will need to consider an extended probability space. Additionally, the filtration $(\mathcal{H}_t)_{t \in \mathbb{R}_+}$, to which X is adapted, will also have to be extended to a larger filtration $(\mathcal{G}_t)_{t \in \mathbb{R}_+}$ and this will constitute a strong Markov extension (see [CJ81, Def. 2.47]). Existence of the appropriate

process Y on the extended probability space is the content of the following theorem.

Theorem 4.2.14. *Let X be a Markov process on M adapted to $(\mathcal{H}_t)_{t \in \mathbb{R}_+}$ with a generator \mathcal{L}_X given by (4.6). Then there exists an extended probability space and on it an \mathbb{R}^d -valued Lévy process Y adapted to an extended filtration $(\mathcal{G}_t)_{t \in \mathbb{R}_+}$ which forms a strong Markov extension, such that $X = \pi(U)$ where U is a horizontal Lévy process solving Marcus SDE (4.3).*

4.2.4 Construction of horizontal lifts and anti-developments for discontinuous processes

As mentioned above, the proof of Theorem 4.2.14 requires us to construct an integrator of Marcus SDE (4.3), such that the solution of this Marcus SDE is projected into a given process. In order to highlight the difficulties arising due to the presence of jumps, it is natural to consider such a construction in a more general setting of horizontal lifts and stochastic anti-developments of general càdlàg semimartingales. Theorem 4.2.19 below, which resolves these problems and we hope is of independent interest, is a key step in the proof of Theorem 4.2.14.

Definition 4.2.15. An $F(M)$ -valued semimartingale U on $[0, \tau)$ is said to be **horizontal** if there exists an \mathbb{R}^d -valued semimartingale $W = (W_t)_{t \in \mathbb{R}_+}$ with $W_0 = 0$, such that U is a solution of a Marcus SDE

$$dU_t = H_i(U_{t-}) \diamond dW_t^i \quad (4.8)$$

on the stochastic interval $[0, \tau)$. The semimartingale W is then called a **stochastic anti-development** of U (or of its projection $X = \pi(U)$). Additionally, if X is an M -valued semimartingale on $[0, \tau)$, then a horizontal semimartingale U is called a **stochastic horizontal lift** of X if $\pi(U) = X$.

Remark 4.2.16. (i) If W in Definition 4.2.15 is in addition a Lévy process on \mathbb{R}^d , we obtain Definition 4.2.6.

(ii) Only times up to τ are used in Marcus SDE (4.8) so it suffices for W to be semimartingale defined only on $[0, \tau)$.

If X is a continuous M -valued semimartingale on $[0, \tau)$, then the stochastic anti-development and the stochastic horizontal lift of X always exist and are unique (on $[0, \tau)$) once the initial frame U_0 satisfying $\pi(U_0) = X_0$ has been chosen [Hsu02, Sec. 2.3]. However, in the discontinuous case, the uniqueness of the stochastic anti-development and the existence of the stochastic horizontal lift may fail. To illustrate some potential

issues posed by jumps, we first focus on the stochastic anti-development of a horizontal semimartingale U , which exists by definition. Uniqueness, however, is related to whether at a jump time s of U there exists a unique fundamental horizontal vector field (and hence a unique jump of the integrator corresponding to it), whose flow exponential map connects the two endpoints U_{s-} and U_s of the jump. A simple example where uniqueness fails is when $M = \mathbb{S}^1 := \{z \in \mathbb{R}^2 ; |z| = 1\}$ is a unit circle equipped with the Levi-Civita connection induced by the standard Riemannian metric of the ambient space \mathbb{R}^2 . Then one easily sees that $F(\mathbb{S}^1)$ is actually diffeomorphic to $\mathbb{S}^1 \times (\mathbb{R} \setminus \{0\})$ and the fundamental horizontal vector field is $H_1 = (\frac{d}{d\theta}, 0)$. Hence $\exp(2\pi k H_1)(u) = u$ for $k \in \mathbb{Z}$, implying that no jump occurs in the case where the integrator has a jump of a size $2\pi k$. We can then take an arbitrary real semimartingale W and let U be a horizontal semimartingale satisfying $dU_t = H_1(U_{t-}) \diamond dW_t$. But then U is also a solution of a Marcus SDE $dU_t = H_1(U_t) \diamond d\widetilde{W}_t$, where $\widetilde{W} = W + 2\pi N$ and N is a Poisson process. In particular, this shows that a continuous horizontal process may have a discontinuous stochastic anti-development (its continuous stochastic anti-development is unique).

On the other hand, the stochastic horizontal lift of a semimartingale X need not exist. Jumps of a horizontal semimartingale occur only along integral curves of fundamental horizontal vector fields, hence jumps of its projection can only occur along geodesics i.e. the end-point of a jump has to be given as a point in the image of a geodesic exponential map based at the start point of the jump. But it could be the case that between the two endpoints X_{t-} and X_t of the jump there does not exist a geodesic. For example, we can take as our manifold a punctured plane $\mathbb{R}^2 \setminus \{0\}$, which is an open subspace of the Euclidean plane, and since there are no geodesics (straight lines) between x and $-x$, a process making such jumps cannot have a stochastic horizontal lift.

Furthermore, even when suitable geodesics exist, they might not be unique, implying the non-uniqueness of horizontal lifts. A typical example is given by the two antipodal points on a sphere and there exists a whole one parameter family of possible geodesics (along great circles). Even between two arbitrary points on a sphere one could travel several times around the sphere along a great circle before stopping at the endpoint, thus breaking uniqueness. Another similar issue related to non-uniqueness of geodesics is that even when we observe no jump of a semimartingale, we could still sometimes force its horizontal lift to jump in a way that would not manifest itself as a jump after projection.

Under suitable assumptions, these issues can be dispensed with. In the setting of Riemannian manifolds, [PE92] considered semimartingales such that there exists a unique geodesic between the two end points of any jump. They were able to show that for such processes the stochastic horizontal lift exists and is unique once the initial frame

is fixed, and the same is true for the stochastic anti-development. As mentioned in the introduction of this chapter, the assumption of unique geodesics is quite strict, excluding compact Riemannian manifolds, such as spheres and projective spaces. In order to deal with general smooth manifolds we provide an alternative approach based on Theorem 4.2.19 below.

The main issues when constructing horizontal lift and anti-development of a semimartingale X concern the jumps of X . It is not sufficient to simply know that a jump occurred at some time s , but we also need to know how it happened, i.e. which specific geodesic was used to induce the jump. Geodesics can be parametrized via the geodesic exponential map, hence we need to specify a tangent vector $v \in T_{X_{s-}}M$ such that $X_s = \text{Exp}_{X_{s-}}(v)$ holds. Typically, we cannot uniquely determine such a vector from X alone, hence this “jump data” needs to be appended to the semimartingale X . Note that such “jump data” will typically have to be defined on an extended probability space and will be adapted to an extended filtration $(\mathcal{G}_t)_{t \in \mathbb{R}_+}$ to which X will still be adapted.

In order to record the necessary “jump data”, fix a holonomy bundle $P(u)$ with a holonomy group $\text{Hol}(u)$ and consider some (global) section $q: M \rightarrow P(u)$, i.e. a map q satisfying $\pi \circ q = \text{id}_M$. We require that q is measurable and that q maps any compact set in M into a pre-compact set in $P(u)$. These requirements would hold trivially if q were smooth, but in general it is not possible to construct even a continuous section. However, the fibre bundle structure allows us to find smooth local sections. Combining countably many such local sections we can always find a measurable (global) section q satisfying compactness condition. Fixing one such section allows us to record the “jump data” as an \mathbb{R}^d -valued process.

Definition 4.2.17. An \mathbb{R}^d -valued process J (possibly) on an extended probability space and adapted to filtration $(\mathcal{G}_t)_{t \in \mathbb{R}_+}$ is said to **represent the jumps of a semimartingale X on $[0, \tau)$** if

$$X_s = \text{Exp}_{X_{s-}}(q(X_{s-})J_s) \text{ for } s < \tau \quad \text{and} \quad \sum_{u \leq t} |J_u|^2 < \infty \text{ for } t \in \mathbb{R}_+. \quad (4.9)$$

Remark 4.2.18. (i) It would suffice for a process J to be defined only on $[0, \tau)$, but it turns out to be convenient in proofs to have it defined for all times in $[0, \infty)$.

(ii) If we start with an anti-development W and use it to construct the processes U and X , we may then represent the jumps of X by a process $J_s := q(X_{s-})^{-1}U_{s-}\Delta W_s$. Processes U and X are càdlàg, hence locally bounded, meaning that for every $t \in \mathbb{R}_+$ and a.e. ω the paths $X(\omega)$ and $U(\omega)$ are included in compact sets for all times in

$[0, t]$, hence by compactness condition on q also matrices $q(X_{s-}(\omega))^{-1}U_{s-}(\omega)$ are included in a compact set for $s \in [0, t]$, so that these matrices are bounded in operator norm. Since jumps of a semimartingale W are square summable we thus immediately obtain the summability condition in (4.9). Hence it is natural to require this summability condition for any process J in Definition 4.2.17. Note also that any such J may be different from 0 for at most countably many times s .

- (iii) For a general semimartingale X there might not exist any process J which represents its jumps. However, (ii) above shows that if X has an anti-development, such a process always exists. Actually, Theorem 4.2.19 below shows that being able to represent jumps of X by some process J is equivalent to X having a stochastic anti-development. The property of X having a stochastic anti-development does not depend on the choice of a section. Moreover, this means that if a pair (q, J) satisfies (4.9), then for any other measurable section \tilde{q} satisfying compactness condition we can define a process $\tilde{J}_s := \tilde{q}(X_{s-})^{-1}U_{s-}\Delta W_s$ (where U and W are as in (ii)) so that (\tilde{q}, \tilde{J}) satisfies (4.9), including the summability condition. This shows that for a semimartingale X the property of being able to represent its jumps as in Definition 4.2.17 does not depend on the choice of a measurable section satisfying compactness condition. Furthermore, the results of Theorem 4.2.19 can be obtained for any choice of such section.

When constructing a horizontal lift U and an anti-development W of $X = (X_t)_{t \in [0, \tau]}$, the process J will enable us to uniquely determine the corresponding jumps of U and W . More precisely, by Definition 4.2.17, Remark 4.2.18 and equation (1.6) we should have

$$U_s = \exp\left(H_{U_{s-}^{-1}q(X_{s-})J_s}\right)(U_{s-}) \quad (4.10)$$

and

$$\Delta W_s = U_{s-}^{-1}q(X_{s-})J_s, \quad (4.11)$$

for all (jump) times $s < \tau$.

Thus, we are able to start with X and J and construct a unique horizontal lift and anti-development.

Theorem 4.2.19. *Let X be a M -valued semimartingale defined on $[0, \tau]$ along with a process J which represents its jumps. Let $u_0 \in \pi^{-1}(\{X_0\}) \cap P(u)$. Then there exists a unique horizontal lift $U = (U_t)_{t \in [0, \tau]}$ of X with $U_0 = u_0$ and values in $P(u)$ satisfying (4.10). It has a unique anti-development $W = (W_t)_{t \in [0, \tau]}$ satisfying (4.11).*

Remark 4.2.20. Both U and W are adapted to filtration $(\mathcal{G}_t)_{t \in \mathbb{R}_+}$ to which X and J were adapted to start with.

As a corollary we may recover the results of [PE92, Thm. 3.2 and Prop. 4.3]. When there exists a unique geodesic between any two endpoints of a jump of a semimartingale X , then there exists a unique² process J which represents its jumps and hence we obtain a unique horizontal lift and a unique anti-development.

4.3 Examples

In this section we show how Definition 4.2.8 relates to previous definitions of Lévy processes on Riemannian manifolds and on Lie groups. We also provide some concrete examples which illustrate different definitions and (slight) differences between them.

4.3.1 Riemannian manifolds

Recall the definition and some basic fact about Riemannian manifolds from Section 1.5. On a Riemannian manifold M the metric uniquely determines a canonical Levi-Civita connection. It is easy to see that the associated fundamental horizontal vector fields are then also vector fields on the orthonormal frame bundle $O(M)$ which consists of orthonormal frames $u: \mathbb{R}^d \rightarrow T_{\pi(u)}M$, i.e. frames which are isometries. Thus, the horizontal Lévy process – a solution of a Marcus SDE (4.3) – stays on the orthonormal frame bundle if it is started there. In [AE00] they considered such horizontal Lévy processes and then their projections to M are Markov if the integrator Y is a Lévy process invariant by the full orthogonal group $O(d)$. In this case they call the projections isotropic Lévy processes on the Riemannian manifold M .

However, it is often possible to reduce the orthonormal frame bundle even further, and hence require less invariance. In the Riemannian case such a reduction was first considered in [Moh04, Sec. 3]. Namely, compatibility of the Levi-Civita connection with the metric means that if u is some (any) fixed orthonormal frame, then the holonomy bundle $P(u)$ is a subbundle of $O(M)$ and the holonomy group $\text{Hol}(u)$ is a Lie subgroup of $O(d)$. Thus, for a fixed orthonormal frame u the Lévy processes from Definition 4.2.8 require less invariance of the integrator than isotropic Lévy processes, and hence the latter are just a special case of the former.

²Process J is uniquely determined only on $[0, \tau)$, but those are the only values used in construction of the horizontal lift and anti-development in Theorem 4.2.19.

Nevertheless, there are still examples where using a holonomy bundle reduction does not extend the class of processes. For a simple example we may consider a d -dimensional sphere $\mathbb{S}^d := \{x \in \mathbb{R}^{d+1} ; |x| = 1\}$ equipped with a standard round metric, i.e. the metric inherited by the standard Euclidean metric of the ambient space \mathbb{R}^{d+1} . Then it is not hard to see that the holonomy group is the special orthogonal group $\text{SO}(d)$, hence isotropic Lévy processes exhaust all Lévy processes on a sphere.

Still, holonomy bundle reduction often extends the class of Lévy process and to see the difference between only isotropic Lévy processes and all possible Lévy processes from Definition 4.2.8 we now analyse a couple of examples on flat Riemannian manifolds (meaning the metric is locally isometric to the standard Euclidean one) where isotropic Lévy processes form a strictly smaller class of processes than all the possible processes from Definition 4.2.8.

Tori

Let $n \in \mathbb{N}$ and consider the action of a discrete group \mathbb{Z}^n on \mathbb{R}^n given by

$$\varphi: \mathbb{R}^n \times \mathbb{Z}^n \rightarrow \mathbb{R}^n, \varphi(x, z) = x + z.$$

The action is free and proper so the Quotient manifold theorem [Lee12, Thm. 21.10] implies that the quotient $\mathbb{T}^n := \mathbb{R}^n / \mathbb{Z}^n$ has a unique smooth structure making it into a smooth manifold of dimension n , such that the quotient projection $\pi: \mathbb{R}^n \rightarrow \mathbb{T}^n$ is a smooth submersion. The smooth manifold \mathbb{T}^n is known as an n -dimensional torus. Furthermore, the projection π is a local diffeomorphism and \mathbb{R}^n is a universal covering space of \mathbb{T}^n . We consider a standard Euclidean metric g on \mathbb{R}^n and since φ acts by isometries, there is a canonical metric \tilde{g} on \mathbb{T}^n defined by

$$\tilde{g}_p(u, v) := g_x((d\pi_x)^{-1}u, (d\pi_x)^{-1}v), \quad (4.12)$$

where $x \in \mathbb{R}^n$ is any point with $\pi(x) = p \in \mathbb{T}^n$ and $u, v \in T_p\mathbb{T}^n$.

We now compute the holonomy group. Let $\tilde{\gamma}$ be a loop based at $p \in \mathbb{T}^n$. Covering space theory tells us that for any $x \in \pi^{-1}(p)$ there exists a unique curve γ in \mathbb{R}^n with $\gamma_0 = x$, such that $\pi \circ \gamma = \tilde{\gamma}$. The endpoint of γ is not necessarily equal to x , but it is always some element of the fibre $\pi^{-1}(p)$, so it can be written as $\varphi(x, z)$ for a unique $z \in \mathbb{Z}^n$. It is now clear that the parallel transport $\tilde{\tau}_{\tilde{\gamma}}$ along $\tilde{\gamma}$ of tangent vectors in $T_p\mathbb{T}^n$

can be expressed in terms of the parallel transport τ_γ along γ in \mathbb{R}^n and it is given by

$$\tilde{\tau}_\gamma = d\pi_{\varphi(x,z)} \circ \tau_\gamma \circ (d\pi_x)^{-1} = d\pi_x \circ (d\varphi(\cdot, z)_x)^{-1} \circ \tau_\gamma \circ (d\pi_x)^{-1} = \text{id}_{T_p\mathbb{T}^n}.$$

In the last equality we used that the parallel transport of the flat Euclidean metric, when we identify $T\mathbb{R}^n$ by $\mathbb{R}^n \times \mathbb{R}^n$, is given exactly by the translation $\varphi(\cdot, z)$. Hence, the holonomy group is trivial and the holonomy bundle is diffeomorphic to \mathbb{T}^n . It also follows that each horizontal vector field H_i on the holonomy bundle is π -related (see (4.15) below for the definition) to the standard vector field ∂_i on \mathbb{R}^d for $i = 1, \dots, d$. This implies the geodesic completeness of the torus and shows that every Lévy process on \mathbb{T}^n is given as $\pi(X)$, where X is an arbitrary Lévy process on \mathbb{R}^n . On the other hand isotropic Lévy processes on \mathbb{T}^n are exactly those given as $\pi(X)$ for an isotropic Lévy process X on \mathbb{R}^n and thus form a strictly smaller class. Actually, \mathbb{Z}^n is a closed subgroup of \mathbb{R}^n and the action we considered is given via the group operation, so torus \mathbb{T}^n can actually be given a structure of a Lie group and we can alternatively proceed as in Section 4.3.2 below to recover the same results.

Generalized Klein bottles

The second example generalizes the Klein bottle. Let $n \in \mathbb{N}$ and let $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$. The group we consider is a semi-direct product $\mathbb{Z}^n \rtimes \mathbb{Z}$, so that the group multiplication is given by $(a,b) \cdot (z,w) := ((-1)^w a + z, b + w)$. Note that this is not a subgroup of \mathbb{R}^{n+1} with group operation of addition. The group action we consider is

$$\varphi: (\mathbb{R}^n \times \mathbb{R}) \times (\mathbb{Z}^n \rtimes \mathbb{Z}) \rightarrow \mathbb{R}^n \times \mathbb{R}, \varphi((x,y), (z,w)) = ((-1)^w x + z, y + w).$$

Again, the action is free and proper so the quotient $\mathbb{K}^{n+1} := (\mathbb{R}^n \times \mathbb{R}) / (\mathbb{Z}^n \rtimes \mathbb{Z})$ has a unique smooth structure making it into a smooth manifold of dimension n . We call the manifold \mathbb{K}^{n+1} an $(n+1)$ -dimensional Klein bottle. The projection $\pi: \mathbb{R}^{n+1} \rightarrow \mathbb{K}^{n+1}$ is a local diffeomorphism and \mathbb{R}^{n+1} is a universal covering space of \mathbb{K}^{n+1} . As in (4.12) we can endow \mathbb{K}^{n+1} with a flat Riemannian metric turning in into a geodesically complete manifold.

To compute the holonomy group we proceed as above. Let $\tilde{\gamma}$ be a loop based at $p \in \mathbb{K}^{n+1}$ and γ a unique curve in \mathbb{R}^{n+1} starting at $(x,y) \in \pi^{-1}(p)$, such that $\pi \circ \gamma = \tilde{\gamma}$. The endpoint of γ can be written as $\varphi((x,y), (z,w))$ for a unique element $(z,w) \in \mathbb{Z}^n \rtimes \mathbb{Z}$.

Again, we may compute the parallel transport as

$$\tilde{\tau}_\gamma = d\pi_{\varphi((x,y),(z,w))} \circ \tau_\gamma \circ (d\pi_{(x,y)})^{-1} = d\pi_{(x,y)} \circ (d\varphi(\cdot, (z,w))_{(x,y)})^{-1} \circ \tau_\gamma \circ (d\pi_{(x,y)})^{-1}.$$

We then easily see that the map $(d\varphi(\cdot, (z,w))_{(x,y)})^{-1} \circ \tau_\gamma$ can be expressed in terms of standard coordinates on \mathbb{R}^{n+1} as

$$\begin{bmatrix} (-1)^w I_n & 0 \\ 0^\top & 1 \end{bmatrix},$$

where I_n is an identity matrix of dimension n and $0 \in \mathbb{R}^n$ is a zero column vector. Thus, the holonomy group is isomorphic to \mathbb{Z}_2 and all the Lévy processes on \mathbb{K}^{n+1} are given by $\pi(X)$, where X is a Lévy process on \mathbb{R}^{n+1} with the law invariant under the multiplication by the matrix

$$\begin{bmatrix} -I_n & 0 \\ 0^\top & 1 \end{bmatrix}.$$

Again, this is less restrictive than only using isotropic Lévy processes on \mathbb{R}^{n+1} which induce isotropic Lévy processes on \mathbb{K}^{n+1} .

4.3.2 Lie groups

We will use Definition 4.2.8 to construct Lévy processes on a connected d -dimensional Lie group G and compare them to a classical definition of (left) Lévy processes on Lie groups defined via independent and stationary increments (see [Lia04] or [Hun56]).

First, we need to select a suitable connection on G . Lie groups are parallelizable since they possess a global frame. To get a global frame we consider a basis V_1, \dots, V_d of a Lie algebra $\mathfrak{g} := T_e G$ of G , where e denotes the identity element. Given an element $V \in \mathfrak{g}$ we define a left-invariant vector field V^L by $V^L|_g := d(L_g)_e V$, where $d(L_g)_e$ is a differential of left multiplication $L_g: h \mapsto gh$. Then a global frame is given by V_1^L, \dots, V_d^L . Analogously, we could define right-invariant vector fields V_1^R, \dots, V_d^R , which also form a global frame. Having a global frame allows us to define a connection (covariant derivative) on elements of the global frame and then extending it to all vector fields by $C^\infty(G)$ -linearity in the first argument and by the product rule in the second argument.

We want to select an appropriate connection so that the parallel transport induced by any curve γ is given exactly by the differential of the left multiplication $d(L_{\gamma_1 \gamma_0^{-1}})_{\gamma_0}$ where the left multiplication takes the initial point γ_0 of the curve to its end point γ_1 . Since we want that this equality holds for any curve γ , we quickly see that the only

possible such connection is given via

$$\nabla_{V_i^L} V_j^L = 0 \quad \text{for all } i, j = 1, \dots, d$$

and we call this connection a left-invariant connection. It is easily seen that the left-invariant connection turns G into a geodesically complete manifold. If we wanted that the parallel transport of a connection is given by the differential of right translations, we could use the same prescription, but the global frame would have to consist of right-invariant vector fields V_1^R, \dots, V_d^R .

It is obvious that the holonomy group associated with the left-invariant connection is trivial, so for any frame $u \in F(G)$ the restricted projection $\pi: P(u) \rightarrow G$ is actually a diffeomorphism between the holonomy bundle $P(u)$ and the Lie group G . For definiteness, we fix u to be the frame with $\pi(u) = e$ for which $u(e_i) = V_i$ holds for every $i = 1, \dots, d$.

In order to get a Lévy process X on the Lie group G we first construct a horizontal Lévy process U on $P(u)$ as a solution of Marcus SDE (4.3), where H_i are fundamental horizontal vector fields of the left-invariant connection and then $X = \pi(U)$. Since $\pi: P(u) \rightarrow G$ is a diffeomorphism, Lemma 4.4.3 shows that the process X is actually a solution of a Marcus SDE

$$X_t = \pi_* H_i(X_{t-}) \diamond dY_t^i,$$

and due to our particular choice of frame u and the associated holonomy bundle $P(u)$, the push-forward vector fields $\pi_* H_i$ are equal to V_i^L for each $i = 1, \dots, d$, so X is a solution of a Marcus SDE

$$dX_t = V_i^L(X_{t-}) \diamond dY_t^i. \quad (4.13)$$

Since V_i^L are left-invariant vector fields and since Y is a Lévy process it seems obvious the process X has independent and stationary increments, making it a (left) Lévy process on a Lie group G as per classical definition. However, we would also like to relate some characteristics of the Lévy process Y to those of the process X . In particular, we would like to relate their Lévy measures, which govern their jumps and are therefore measures on $\mathbb{R}^d \setminus \{0\}$ and $G \setminus \{e\}$, respectively, which satisfy certain integrability conditions.

Process X is a solution of (4.13) and by the definition of the Marcus SDE when the process Y jumps at time t , then X also jumps simultaneously and³ $X_t = X_{t-} \exp(V_i \Delta Y_t^i)$, where $\exp: \mathfrak{g} \rightarrow G$ is the exponential map of the Lie group G . The Lie algebra \mathfrak{g} is isomorphic to \mathbb{R}^d so given a Lévy measure ν of Y we can construct a push-forward measure $\tilde{\mu} := \widetilde{\exp}_* \nu$ on G where $\widetilde{\exp}: \mathbb{R}^d \ni x = x^i e_i \mapsto \exp(x^i V_i) \in G$ and then the restricted

³Recall that we are using Einstein's convention.

measure $\mu := \tilde{\mu}|_{G \setminus \{e\}}$ should be a Lévy measure of the process X . We summarize these facts in the next proposition and provide the proof in Section 4.4.5.

Proposition 4.3.1. *Let X be a solution of a Marcus SDE (4.13), where Y is an \mathbb{R}^d -valued Lévy process with a Lévy measure ν . Then X is a (left) Lévy process on G with a Lévy measure $\mu := \widetilde{\exp}_* \nu|_{G \setminus \{e\}}$.*

Proposition 4.3.1 shows that every process X constructed via Definition 4.2.8 is a classical (left) Lévy process in G , but the converse need not be true. Being able to represent a Lévy process X in G as a solution of a Marcus SDE (4.13) is equivalent to the existence of a stochastic horizontal lift U of X , such that its anti-development Y is a Euclidean Lévy process. By Proposition 4.3.1 we know that Lévy measures of processes X and Y then need to be related by $\mu = \widetilde{\exp}_* \nu|_{G \setminus \{e\}}$. But the exponential map need not be surjective, so there can exist Lévy measures on G , and hence (left) Lévy processes on G , which cannot be represented by this construction. Obviously, to even have a chance of finding an anti-development Y , we must be able to represent all the jumps of X via the exponential map. Actually, as soon as this is the case, we are able to construct the process Y .

Proposition 4.3.2. *Let X be a left Lévy process in G with a Lévy measure μ . Then there exists an \mathbb{R}^d -valued Lévy process Y such that X satisfies Marcus SDE (4.13) if and only if $\mu(\text{Im}(\exp)^c) = 0$.*

The proof of the proposition is given in Section 4.4.5 and we can also deduce the following corollary.

Corollary 4.3.3. *Let G be a Lie group with a surjective exponential map. Then a process X is a (left) Lévy process in G if and only if it is a solution of a Marcus SDE (4.13) for some \mathbb{R}^d -valued Lévy process Y .*

Remark 4.3.4. (i) A statement analogous to the *if* part of Corollary 4.3.3 is stated in the review article [AL, p. 8]. In fact, Proposition 4.3.1 shows that the *if* part holds without the assumption on the surjectivity of the exponential map.

(ii) In the *only if* part, for a process X to satisfy Marcus SDE (4.13), one needs to construct a Lévy process Y on the extended probability space such that (4.13) holds. It is not sufficient to construct a Lévy process Y on some other unrelated probability space, such that a different solution \tilde{X} of (4.13) driven by Y has the same law as X . The construction of the Lévy process Y from X is given in the proof of our main result, Theorem 4.2.14.

Conditions of the corollary often hold, for example if the group is connected and compact or if it is simply connected and nilpotent. On the other hand, we trivially observe non-surjectivity of the exponential map when a Lie group is not connected, since then the image of the (connected) Lie algebra under the (continuous) exponential map obviously cannot be the whole (disconnected) group. A non-trivial example of a connected (and non-compact) Lie group with a non-surjective exponential map is given by the group $\mathrm{GL}(2)_+$, i.e. the group of matrices in $\mathbb{R}^2 \otimes \mathbb{R}^2$ with strictly positive determinants.

4.3.3 Further examples

Definition 4.2.8 allows us to define Lévy processes on an arbitrary smooth manifold as soon as we specify a connection on it. On a general smooth manifold there are (infinitely) many possible choices of a connection and in some cases, as is seen in previous sections, there is a canonical choice. The particular connection we use is intimately related to the Definition 4.2.8 and leads to often non-trivial conditions (4.5). We will now show how these conditions restrict and may sometimes even severely limit the class of Lévy processes on a manifold.

Recall that the holonomy group $\mathrm{Hol}(u)$ of the holonomy bundle $P(u)$ (associated with a frame $u \in F(M)$) is a Lie subgroup of the general linear group $\mathrm{GL}(d)$, where d is the dimension of the underlying manifold M . The holonomy group is determined by the connection and it is natural to consider the question of whether any Lie subgroup of the general linear group $\mathrm{GL}(d)$ can be realized as a holonomy group of some manifold with a connection. A partial answer has been known for over half a century and [HO56] shows that for any connected Lie subgroup of $\mathrm{GL}(d)$ there exist some manifold with a connection, such that the induced holonomy group is exactly the selected subgroup.

Hence it is possible to find a manifold with a connection, such that its holonomy group is equal to $\mathrm{GL}(d)_+$, i.e. invertible matrices with strictly positive determinants. In this case the Lévy integrator in the definition of a horizontal Lévy process has to satisfy conditions (4.5) and it is quickly seen that having such a large holonomy group forces the Lévy integrator to be a trivial process constantly equal to 0. Hence all Lévy processes on such a manifold have to be constant as well. This is quite an extreme example, but it shows that the choice of the connection can have a profound effect on the size of the class of Lévy processes.

For a less extreme example let us consider a group consisting of matrices of the form

$$\begin{bmatrix} I_{d-1} & 0 \\ 0^\top & \beta \end{bmatrix},$$

where I_{d-1} is an identity matrix of dimension $d - 1$, $0 \in \mathbb{R}^{d-1}$ is a zero column vector and $\beta > 0$. Such matrices for strictly positive values of parameter β form a connected Lie subgroup of $\text{GL}(d)$, hence it may once again be realized as a holonomy group of some manifold with a connection. In this case conditions (4.5) force the Lévy integrator to stay in the $(d - 1)$ -dimensional hyperplane perpendicular to e_d . There are no further restrictions, so we may take any Lévy integrator, which is given as an arbitrary $(d - 1)$ -dimensional Euclidean Lévy process in the first $d - 1$ components and with vanishing last component. This essentially means that the Lévy process on the manifold “loses one direction” in which it may move, but this is a geometric restriction required to keep our definition well-posed.

Cylinder with a family of connections

We now present an example of a 2-dimensional manifold with a 1-parameter family of connections on it. Depending on the value of the parameter, the holonomy group turns out to be a cyclic subgroup of $\text{SO}(2)$, which is either finite or an infinite dense subgroup isomorphic to \mathbb{Z} . In particular, if the rank of the holonomy group is even, it contains the element $-I_2$, providing a class of examples where we may simplify the jump part of the generator of Lévy processes on the manifold as in Remark 4.2.13(i). Moreover, the connection introduced below is not torsion-free (for $\alpha \neq 0$) and thus cannot be the Levi-Civita connection of a Riemannian manifold.

The 2-dimensional manifold $M = \mathbb{S}^1 \times \mathbb{R}$ is paralellizable with a global frame given by the vector fields V_θ, V_z defined as follows: under the natural inclusion $\mathbb{S}^1 \times \mathbb{R} \rightarrow \mathbb{R}^2 \times \mathbb{R}$ let $V_\theta = (-y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y})|_M$ and $V_z = (\frac{\partial}{\partial z})|_M$. We can now specify a connection (covariant derivative) via the global frame:

$$\nabla_{V_z} V_\theta = \nabla_{V_z} V_z = 0, \quad \nabla_{V_\theta} V_\theta = -\alpha V_z, \quad \nabla_{V_\theta} V_z = \alpha V_\theta,$$

where $\alpha \in \mathbb{R}$ is a real parameter. By using (1.3) we can bound the derivatives of geodesics (in the ambient space) which shows that M equipped with this connection is geodesically complete. Furthermore, it is easy to describe the parallel transport of such a connection as we may use a global frame to identify different tangent spaces. If we are moving along V_z , the tangent vectors remain unchanged, while moving along V_θ rotates the tangent vectors by an angle expressed in radians as α times the directed length of the horizontal displacement. Hence the holonomy group is generated by the rotation around the origin through the angle $2\pi\alpha$. The holonomy group is thus isomorphic to \mathbb{Z} if α is irrational and to $(\mathbb{Z}_q, +)$ (for some $q \in \mathbb{N}$) if α is a rational. In particular, by taking $\alpha = \frac{1}{2}$, the

holonomy group equals $\{I_2, -I_2\}$.

4.4 Proofs

4.4.1 Marcus SDEs on manifolds

We start with the proof of our main existence and uniqueness result for Marcus SDEs.

Proof of Proposition 4.2.3. In [KPP95] Marcus SDE was studied in Euclidean spaces. They use a component-wise definition of a solution; however, the chain rule [KPP95, Prop. 4.2] shows that it is equivalent to Definition 4.2.1 when the smooth manifold M is taken to be the Euclidean space and the vector fields are Lipschitz continuous, so that the solution is defined for all times in \mathbb{R}_+ . To prove the proposition for a general smooth manifold M , we use Whitney's embedding theorem to embed M as a closed submanifold of the Euclidean space \mathbb{R}^n for a sufficiently large⁴ n . In particular, this allows us to identify the point ∂ of compactification \widehat{M} with the point ∞ of compactification of \mathbb{R}^d . We can then smoothly extend vector fields V_i on M to vector fields V_i^* on the whole \mathbb{R}^n . Being smooth, these vector fields are locally Lipschitz continuous. We consider closed balls $\mathbb{B}_k := \{x \in \mathbb{R}^n ; |x| \leq k\}$ for $k \in \mathbb{N}$ and for each $k \in \mathbb{N}$ we take smooth functions φ_k such that $\varphi_k = 1$ on \mathbb{B}_k and $\text{supp } \varphi_k \subseteq \mathbb{B}_{k+1}$. By denoting $V_i^{(k)} := \varphi_k V_i^*$ for $i = 1, \dots, \ell$ we see that vector fields $V_i^{(k)}$ are compactly supported, hence they are globally Lipschitz continuous and any linear combination of them is a complete vector field.

We consider a Marcus SDE

$$dX_t^{(k)} = V_i^{(k)}(X_{t-}^{(k)}) \diamond dW_t^i \quad (4.14)$$

on \mathbb{R}^n with the initial condition $X_0^{(k)} \in M$ and without loss of generality (by translation) we may assume that $X_0^{(k)} = 0 \in \mathbb{B}_k$. Then [KPP95, Thm. 3.2] shows that there exists a unique solution $(X_t^{(k)})_{t \in \mathbb{R}_+}$ of (4.14) and since vector fields $V_i^{(k)}$ are tangent to M , [KPP95, Prop 4.2] shows that the solution $X^{(k)}$ stays on M . We define

$$\tau_k := \inf \left\{ s \in \mathbb{R}_+ ; \exists \theta \in [0,1] \text{ such that } \exp(\theta V_i \Delta W_s^i)(X_{s-}^{(k)}) \notin \mathbb{B}_k \right\},$$

i.e. τ_k is the first time the process $X^{(k)}$ or the "hidden" trajectory of its jump exits \mathbb{B}_k . For $k \leq \ell$ consider the processes $X^{(k)}$ and $X^{(\ell)}$. Restricted to the interval $[0, \tau_k)$, they depend on their corresponding vector fields $V_i^{(k)}$ and $V_i^{(\ell)}$ restricted to \mathbb{B}_n , but the vector

⁴Twice the dimension of the manifold M suffices.

fields coincide there. Therefore, uniqueness of solutions from [KPP95, Thm. 3.2] means that the processes $X^{(k)}$ and $X^{(\ell)}$ coincide on $[0, \tau_k)$, so we may unambiguously define a process $(X_t)_{t \in [0, \tau)}$ where $\tau = \lim_{k \rightarrow \infty} \tau_k$ so that $X_t = X_t^{(k)}$ on $[0, \tau_k)$. It easily follows that the process X is exactly the unique solution of Marcus SDE (4.1) on $[0, \tau)$. Finally, we have to prove that τ is an explosion time and that $\{\tau_n ; n \in \mathbb{N}\}$ form its announcing sequence. This immediately follows once we show that $\tau_n < \tau$ a.s. for every $n \in \mathbb{N}$.

In general, it is possible that $X^{(k)}$ exits \mathbb{B}_k at a time τ_k by a jump – such a jump being induced by the jump ΔW_{τ_k} of the integrator. If the vector field $V_i \Delta W_{\tau_k}^i$ were not complete then $\exp(V_i \Delta W_{\tau_k}^i)(X_{\tau_k-})$ might not be defined and in this case $\tau_\ell = \tau_k$ for all $\ell \geq k$, and hence $\tau_k = \tau$. However, this cannot happen if as assumed all the vector fields $a^i V_i$ are complete, since then the flow exponential map is always defined, thus the norm of $\exp(V_i \Delta W_{\tau_k}^i)(X_{\tau_k-})$ is smaller than L for some $L \in \mathbb{N}$ larger than k and we have $\tau_k < \tau_L \leq \tau$ for any $k \in \mathbb{N}$ and this finishes the proof. \square

We now prove a couple of technical lemmas which elucidate the transformation rules of Marcus SDEs. First, we may consider a Marcus SDE and linearly transform the vector fields appearing in the SDE. Next lemma shows that this is the same as appropriately linearly transforming the integrator.

Lemma 4.4.1. *Let $\ell, m \in \mathbb{N}$ and let $V_1, \dots, V_\ell \in \Gamma(TM)$ be smooth vector fields on M and W be a càdlàg \mathbb{R}^m -valued semimartingale. Suppose that $\tilde{V}_1, \dots, \tilde{V}_m \in \Gamma(TM)$ are smooth vector fields given as $\tilde{V}_i := A_i^j V_j$ for some matrix $A \in \mathbb{R}^\ell \otimes \mathbb{R}^m$. Then an M -valued semimartingale X defined on $[0, \tau)$ is a solution of Marcus SDE*

$$dX_t = \tilde{V}_i(X_{t-}) \diamond dW_t^i,$$

on $[0, \tau)$ if and only if it is a solution of Marcus SDE

$$dX_t = V_j(X_{t-}) \diamond d\tilde{W}_t^j$$

on $[0, \tau)$, where an \mathbb{R}^ℓ -valued semimartingale \tilde{W} is given by $\tilde{W} := AW$.

Proof. The defining property (4.2) for the first Marcus SDE is the equation

$$\begin{aligned} f(X_t) - f(X_0) &= \int_0^t A_i^j V_j f(X_{s-}) \circ dW_s^i \\ &+ \sum_{0 < s \leq t} \left(f(\exp(A_i^j V_j \Delta W_s^i)(X_{s-})) - f(X_{s-}) - A_i^j V_j f(X_{s-}) \Delta W_s^i \right) \end{aligned}$$

for any $f \in C^\infty(M)$, and simple algebra using $\widetilde{W}^j = A_i^j W^i$ shows that this is also exactly the defining property for the second Marcus SDE. \square

Another important property of the Marcus SDE is that it behaves nicely under diffeomorphisms. We will first prove a slightly more general result. Let $\Phi: M \rightarrow N$ be a smooth map. We say that vector fields $V \in \Gamma(TM)$ and $V' \in \Gamma(TN)$ are Φ -**related** if

$$d\Phi_p V_p = V'_{\Phi(p)} \quad (4.15)$$

for every $p \in M$, where $d\Phi_p$ represents the differential of smooth map Φ at a point p .

Lemma 4.4.2. *Let M and N be smooth manifolds and let $\Phi: M \rightarrow N$ be a smooth map. Let $V_1, \dots, V_\ell \in \Gamma(TM)$ be vector fields and let a semimartingale X defined on $[0, \tau)$ be an M -valued solution of a Marcus SDE*

$$dX_t = V_i(X_{t-}) \diamond dW_t^i \quad (4.16)$$

on $[0, \tau)$. Suppose that $V'_1, \dots, V'_\ell \in \Gamma(TN)$ are vector fields, such that V_i and V'_i are Φ -related for every $i = 1, \dots, \ell$. Then the N -valued semimartingale $\widetilde{X} := \Phi(X)$ is a solution of a Marcus SDE

$$d\widetilde{X}_t = V'_i(\widetilde{X}_{t-}) \diamond dW_t^i \quad (4.17)$$

on $[0, \tau)$.

Proof. We need to show that the defining property (4.2) of Marcus SDE (4.17) holds for \widetilde{X} when we take an arbitrary smooth function $f \in C^\infty(N)$. We can take a function $f \circ \Phi \in C^\infty(M)$ and use the defining property (4.2) of Marcus SDE (4.16) for X to get

$$\begin{aligned} f(\widetilde{X}_t) - f(\widetilde{X}_0) &= \int_0^t V_i(f \circ \Phi)(X_{s-}) \diamond dW_s^i \\ &+ \sum_{0 < s \leq t} \left(f \circ \Phi(\exp(V_i \Delta W_s^i)(X_{s-})) - f(\widetilde{X}_{s-}) - V_i(f \circ \Phi)(X_{s-}) \Delta W_s^i \right). \end{aligned} \quad (4.18)$$

Equation (4.15) implies that $V'_i f(\Phi(p)) = V_i(f \circ \Phi)(p)$ holds for any $f \in C^\infty(N)$ and $p \in M$, so we have

$$V'_i f(\widetilde{X}_{t-}) = V_i(f \circ \Phi)(X_{t-}) \quad (4.19)$$

for any $i = 1, \dots, \ell$.

Using linearity in (4.15) shows that vector fields $V_x := x^i V_i$ and $V'_x := x^i V'_i$ are Φ -related for any $x = x^i e_i \in \mathbb{R}^d$. By (4.15), if a tangent vector of a curve γ is equal to $V_x|_p$

at p , then the tangent vector of a curve $\Phi \circ \gamma$ is equal to $V'_x|_{\Phi(p)}$ at $\Phi(p)$. In particular, Φ maps integral curves of V_x to integral curves of V'_x , hence

$$\Phi(\exp(V_x)(p)) = \exp(V'_x)(\Phi(p)) \quad (4.20)$$

holds for all $p \in M$.

Inserting (4.19) and (4.20) into equation (4.18) yields exactly the defining property for Marcus SDE (4.17) and \tilde{X} is its solution. \square

If a map $\Phi: M \rightarrow N$ is a diffeomorphism, we can define a **push-forward** $\Phi_*: \Gamma(TM) \rightarrow \Gamma(TN)$ so that for a vector field $V \in \Gamma(TM)$ the **push-forward vector field** $\Phi_*V \in \Gamma(TN)$ is defined by the prescription

$$(\Phi_*V)_q := d\Phi_{\Phi^{-1}(q)}V_{\Phi^{-1}(q)} \quad (4.21)$$

for any $q \in N$. It is clear from this definition that the vector fields V and Φ_*V are Φ -related, so the following lemma is an immediate corollary of Lemma 4.4.2.

Lemma 4.4.3. *Let $\Phi: M \rightarrow N$ be a diffeomorphism and let semimartingale X defined on $[0, \tau)$ be an M -valued solution of a Marcus SDE*

$$dX_t = V_i(X_{t-}) \diamond dW_t^i$$

on $[0, \tau)$. Then the N -valued semimartingale $\tilde{X} := \Phi(X)$ is a solution of a Marcus SDE

$$d\tilde{X}_t = \Phi_*V_i(\tilde{X}_{t-}) \diamond dW_t^i$$

on $[0, \tau)$.

4.4.2 Proofs of Proposition 4.2.9, Lemma 4.2.11 and Proposition 4.2.12

Proof of Proposition 4.2.9. In the discussion prior to the Definition 4.2.8 it was already shown that the invariance of the integrator under the holonomy group implies Markovianity of the projected process. We now show that changing the frame and the associated holonomy bundle does not alter the class of Lévy processes on a connected smooth manifold M .

Consider a different frame $v \in F(M)$. Connectedness of M implies that $\pi: P(v) \rightarrow M$ is surjective and if $v' \in P(v)$ is another frame, then $P(v') = P(v)$ and $\text{Hol}(v') = \text{Hol}(v)$, hence we may assume that $v = ug$ for some $g \in \text{GL}(d)$ so that $\text{Hol}(v) = g^{-1}\text{Hol}(u)g$

holds [KN63, Prop. II.4.1]. We take a Lévy process $X = \pi(U)$, where U is a horizontal Lévy process and a solution of a Marcus SDE (4.3) with a $\text{Hol}(u)$ -invariant integrator Y and $U_0 \in P(u)$. Hence, if we denote $\tilde{U} := Ug$ and $\tilde{Y} := g^{-1}Y$, then Lemmas 4.4.3, 1.3.2 and 4.4.1 imply that \tilde{U} is a $P(v)$ -valued solution of the Marcus SDE

$$d\tilde{U}_t = H_i(\tilde{U}_{t-}) \diamond d\tilde{Y}_t^i$$

with an integrator \tilde{Y} , which is $\text{Hol}(v)$ -invariant. Since $X = \pi(U) = \pi(\tilde{U})$, X is also a Lévy process when defined through the holonomy bundle $P(v)$. This shows that the same processes are called Lévy processes on M no matter which holonomy bundle we use for their definition. \square

Proof of Lemma 4.2.11. Recall that a Euclidean Lévy process Y with the generating triplet (a, ν, b) can be expressed via the Lévy-Itô decomposition (Theorem A.2.4)

$$Y_t^i = b^i t + \sigma_j^i B_t^j + \int_0^t \int_{0 < |x| < 1} x^i \tilde{N}(ds, dx) + \int_0^t \int_{|x| \geq 1} x^i N(ds, dx), \quad (4.22)$$

where $\sum_{k=1}^d \sigma_k^i \sigma_k^j = a^{ij}$, B is a standard d -dimensional Brownian motion, N is an independent Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d$ with mean measure $\mathcal{L}eb \otimes \nu$, i.e. $\mathbb{E}[N([0, t], A)] = t\nu(A)$, and \tilde{N} is the compensated random measure defined by $\tilde{N}([0, t], A) = N([0, t], A) - t\nu(A)$. Since U is a solution of Marcus SDE (4.3), we can take $f \in C^\infty(P(u))$ and substitute (4.22) into the defining property of Marcus SDEs to get

$$\begin{aligned} f(U_t) &= f(U_0) + b^i \int_0^t H_i f(U_{s-}) ds + \sigma_j^i \int_0^t H_i f(U_{s-}) \circ dB_s^j \\ &\quad + \int_0^t \int_{0 < |x| < 1} x^i H_i f(U_{s-}) \tilde{N}(ds, dx) + \int_0^t \int_{|x| \geq 1} x^i H_i f(U_{s-}) N(ds, dx) \\ &\quad + \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} (f(\exp(H_x)(U_{s-})) - f(U_{s-}) - x^i H_i f(U_{s-})) N(ds, dx). \end{aligned}$$

After recalling the definition of Stratonovich integral and some algebraic manipulation

using $\tilde{N}(ds, dx) = N(ds, dx) - ds \nu(dx)$ we can rewrite the equation as

$$\begin{aligned} f(U_t) &= f(U_0) + \sigma_j^i \int_0^t H_i f(U_{s-}) dB_s^j \\ &+ \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} (f(\exp(H_x)(U_{s-})) - f(U_{s-})) \tilde{N}(ds, dx) \\ &+ \int_0^t \mathcal{L}_U(f)(U_s) ds, \end{aligned} \tag{4.23}$$

where \mathcal{L}_U is given by (4.4). Since the first two integrals in the right hand side of equation (4.23) are (local) martingales, \mathcal{L}_U is the (extended) generator (see[Dyn65, ÇJPS80]) of the horizontal Lévy process U , at least on smooth functions. \square

Proof of Proposition 4.2.12. Since $f(X) = f \circ \pi(U)$, the only possible prescription for the generator of X is given by

$$\mathcal{L}_X(f)(p) = \mathcal{L}_U(f \circ \pi)(v), \tag{4.24}$$

where v is any frame in $P(u)$ with $\pi(v) = p$ and \mathcal{L}_U is a generator of a horizontal Lévy process obtained in Lemma 4.2.11. We use invariance conditions (4.5) to apply [Dyn65, Thm. 10.13] from which it follows that right hand side of (4.24) is indeed well-defined and equal to the generator of X . The theorem also identifies the domain of the generator.

In order to finish the proof and get the explicit expression (4.6) for the generator we simply use in the above expression the equality (1.6) relating flow and geodesic exponential maps. \square

4.4.3 Proof of Theorem 4.2.19

For the proof of the theorem we will need the following estimate for the flow exponential map of the fundamental horizontal vector fields in local coordinates on the frame bundle.

Lemma 4.4.4. *Consider local coordinates on $O \subseteq M$ and the associated local coordinates on $\pi^{-1}(O) \subseteq F(M)$. Let $u = (x^i, r_m^k) = (x, r) \in \pi^{-1}(O)$, $c = c^n e_n \in \mathbb{R}^d$ and consider $\exp(H_c)(u) = \psi(1)$, where $\psi(t) = \psi(t, u, c)$ is the integral curve of H_c starting at $\psi(0) = u$, which we assume stays in $\pi^{-1}(O)$ for all $t \in [0, 1]$. Then we have*

$$\exp(H_c)(u) = (x^i, r_m^k) + \left(c^n r_n^i, -c^n r_n^j r_m^l \Gamma_{jl}^k(x) \right) + c^n c^{n'} F_{nn'}(\psi(\theta)) \tag{4.25}$$

for some smooth functions $F_{nn'}$ with values in \mathbb{R}^{d+d^2} and some $\theta \in (0, 1)$.

Proof. Via local coordinates we may consider ψ as a smooth curve in \mathbb{R}^{d+d^2} and we can use Taylor approximation up to order 2 to write:

$$\exp(H_c)(u) = \psi(1) = \psi(0) + \dot{\psi}(0) + \frac{1}{2}\ddot{\psi}(\theta)$$

for some $\theta \in (0,1)$. We obviously have $\psi(0) = u = (x^i, r_m^k)$. Since ψ is an integral curve of a fundamental horizontal vector field, we use the representation in local coordinates from Lemma 1.3.1 to write

$$\dot{\psi}(t) = H_c(\psi(t)) = \left(c^n \psi_n^i(t), -c^n \psi_n^j(t) \psi_m^l(t) \Gamma_{jl}^k(\pi \circ \psi(t)) \right), \quad (4.26)$$

from which $\dot{\psi}(0) = \left(c^n r_n^i, -c^n r_n^j r_m^l \Gamma_{jl}^k(x) \right)$ easily follows. Hence, we have already obtained the first two terms in (4.25) and we only need to obtain the final one. We compute further time derivative of (4.26) in which we can use (4.26) once more to get

$$\begin{aligned} \ddot{\psi}^i(t) &= c^n \dot{\psi}_n^i(t) = -c^n c^{n'} \psi_{n'}^j(t) \psi_n^l(t) \Gamma_{jl}^i(\pi \circ \psi(t)) \\ \ddot{\psi}_m^k(t) &= c^n c^{n'} \psi_{n'}^\alpha(t) \psi_n^\beta(t) \Gamma_{\alpha\beta}^j(\pi \circ \psi(t)) \psi_m^l(t) \Gamma_{jl}^k(\pi \circ \psi(t)) \\ &\quad + c^n c^{n'} \psi_n^j(t) \psi_{n'}^\alpha(t) \psi_m^\beta(t) \Gamma_{\alpha\beta}^l(\pi \circ \psi(t)) \Gamma_{jl}^k(\pi \circ \psi(t)) \\ &\quad - c^n c^{n'} \psi_n^j(t) \psi_m^l(t) \partial_\alpha \Gamma_{jl}^k(\pi \circ \psi(t)) \psi_{n'}^\alpha(t), \end{aligned}$$

so we can indeed write $\frac{1}{2}\ddot{\psi}(\theta) = c^n c^{n'} F_{nn'}(\psi(\theta))$ for some smooth functions $F_{nn'}$, and thus establish the results of the lemma. \square

Proof of Theorem 4.2.19. We will construct a horizontal lift and an anti-development using local charts and then patch the local solutions together. We consider local coordinates φ^α on a precompact set $O \subseteq M$ and the associated local coordinates ϕ^α, ϕ_m^k on $\pi^{-1}(O) \subseteq F(M)$. Inspired by the deterministic horizontal lift equation (1.4) and condition (4.10) we claim that the horizontal lift $U = (\varphi^i(X), R_m^k)$ solves the following SDE written in local coordinates:

$$\begin{aligned} R_m^k(t) &= R_m^k(0) - \int_0^t R_m^l(s-) \Gamma_{jl}^k(X_{s-}) \circ d\varphi^j(X_s) \\ &\quad + \sum_{0 < s \leq t} \left(\exp \left(H_{U_{s-}^{-1}q(X_{s-})J_s} \right) (U_{s-})_m^k - R_m^k(s-) + R_m^l(s-) \Gamma_{jl}^k(X_{s-}) \Delta\varphi^j(X_s) \right), \end{aligned} \quad (4.27)$$

for every $t < \tilde{\tau}$, where $\tilde{\tau} = \inf \{s \in \mathbb{R}_+ ; X_s \notin O\}$. Its anti-development W is then

defined by

$$W_t^i = \int_0^t \bar{R}_j^i(s-) \circ d\varphi^j(X_s) + \sum_{0 < s \leq t} \left((U_{s-}^{-1} q(X_{s-}) J_s)^i - \bar{R}_j^i(s-) \Delta \varphi^j(X_s) \right) \quad (4.28)$$

for $t < \tilde{\tau}$, where \bar{R} is an inverse of R , i.e. $\bar{R}_j^i R_l^k = \delta_l^i \delta_j^k$. Note that such U and W will be adapted to $(\mathcal{G}_t)_{t \in \mathbb{R}_+}$.

We first analyse the equality (4.27). For it to even make sense, we need to show that the sum on the right hand side is a.s. finite. We assume that we already have such a process U which is càdlàg, hence U_- is locally bounded and we denote its trajectory up to time t by $\kappa_t(\omega)$, which is a bounded set a.s. Hence there exists $\varepsilon(\omega) > 0$ and an a.s. bounded set $D_t(\omega) \subseteq \pi^{-1}(O)$ such that every integral curve of a horizontal vector fields H_c with initial point in κ_t and $|c| < \varepsilon$ stays⁵ in D_t up to time 1. We denote $c_s = U_{s-}^{-1} q(X_{s-}) J_s$, where $U_{s-}^{-1} q(X_{s-})$ is a locally bounded holonomy group valued process. This local boundedness together with condition (4.10) implies that there are at most finitely many times $s \in [0, t]$ with $|c_s| \geq \varepsilon$, so the sum of absolute values of terms corresponding to such times is a finite random variable which we denote by C_t . Since $X_{s-} = \pi(U_{s-})$ and $X_s = \pi(\exp(H_{c_s})(U_{s-}))$, we can use Lemma 4.4.4 and write each of the remaining terms corresponding to $|c_s| < \varepsilon$ as

$$\begin{aligned} & R_m^k(s-) - c_s^n R_n^j(s-) R_m^l(s-) \Gamma_{jl}^k(X_{s-}) + c_s^n c_s^{n'} (F_{nn'})_m^k(\psi(\theta, U_{s-})) - R_m^k(s-) \\ & + R_m^l(s-) \Gamma_{jl}^k(X_{s-}) \left(\varphi^j(X_{s-}) + c_s^n R_n^j(s-) + c_s^n c_s^{n'} (F_{nn'})^j(\psi(\theta, U_{s-})) - \varphi^j(X_{s-}) \right) \\ & = c_s^n c_s^{n'} \left(\tilde{F}_{nn'}(\psi(\theta, U_{s-})) \right)_m^k, \end{aligned}$$

for some smooth functions $\tilde{F}_{nn'}$ and some $\theta = \theta(s, \omega) \in (0, 1)$. Triangle and Cauchy-Schwarz inequalities then imply that the absolute value of the whole sum is smaller or

⁵This follows from results about systems of ODEs in Euclidean setting and uses only that the set of initial points is bounded and that the vector field is locally Lipschitz in the argument as well as in the parameter.

equal to

$$\begin{aligned}
& C_t + \sum_{0 < s \leq t, |c_s| < \varepsilon} \left| c_s^n c_s^{n'} \left(\tilde{F}_{nn'}(\psi(\theta, U_{s-})) \right)_m^k \right| \\
& \leq C_t + d^{-1} \sup_{1 \leq n, n' \leq d, \theta \in [0, 1], s \in [0, t], |c_s| < \varepsilon} \left| \left(\tilde{F}_{nn'}(\psi(\theta, U_{s-})) \right)_m^k \right| \sum_{0 < s \leq t, |c_s| < \varepsilon} |c_s|^2 \\
& \leq C_t + K_t L_t^2 d^{-1} \sum_{s \leq t} |J_s|^2 < \infty,
\end{aligned}$$

where L_t is the (local) bound for the holonomy group valued process $U_{s-}^{-1}q(X_{s-})$ and we have used that the final sum is finite by (4.9). We obtain the bound $K_t(\omega)$ for the supremum as a consequence of taking a supremum of finitely many continuous functions over a subset of a bounded set D_t .

Next, we need to show that there exists a solution of SDE (4.27) which is a semimartingale. We wish to use the general theory of SDEs against semimartingales, so we need to rewrite the equation into a correct form. We start by considering SDE (4.27) in the matrix form so that the first integral can be written as

$$\int_0^t F(R(s-)) \circ d\varphi(X_s) = \int_0^t F_j(R(s-)) \circ d\varphi^j(X_s),$$

where the random coefficients $F_j: (\mathbb{R}^d \otimes \mathbb{R}^d) \times \Omega \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d$ are given by

$$F_j(r, \omega)_{mj}^k := r_m^l \Gamma_{jl}^k(X_{s-}(\omega))$$

for $j = 1, \dots, d$ and $\varphi(X)$ is an \mathbb{R}^d -valued semimartingale $\varphi(X) := (\varphi^1(X), \dots, \varphi^d(X))^\top$. Note that as a function of r , the coefficient F is linear. Next, we need to appropriately transform the sum. We view the sum as an $\mathbb{R}^d \otimes \mathbb{R}^d$ -valued processes and we may use Lemma 1.3.3 to express the sum as a component-wise integral

$$\int_0^t A(X_{s-}, J_s) R(s-) dZ_s,$$

where $Z_t = \sum_{0 < s \leq t} |J_s|^2$ is a (finite variation) real semimartingale. The $\mathbb{R}^d \otimes \mathbb{R}^d$ -valued process $A(X_{t-}, J_t)$ has components given by

$$A(X_{t-}, J_t)_m^k := \left(\exp \left(H_{\tilde{U}_{t-}^{-1}q(X_{t-})J_t} \right) (\tilde{U}_{t-})_m^k - \delta_m^k + B(X_{t-}, J_t)_m^k \right) / |J_t|^2,$$

where $\delta_m^k = \mathbb{1}(k = m)$ is Dirac delta, the process \tilde{U} is given by $\tilde{U} := UR^{-1}$, and the $\mathbb{R}^d \otimes \mathbb{R}^d$ -valued process $B(X_{t-}, J_t)$ has components given by

$$B(X_{t-}, J_t)_m^k := \Gamma_{jm}^k(X_{t-}) \Delta \varphi^j(X_t) = \Gamma_{jm}^k(X_{t-}) (\varphi^j(\text{Exp}_{X_{t-}}(q(X_{t-})J_t)) - \varphi^j(X_{t-})).$$

Notice that \tilde{U} has a simple representation in local coordinates as $\tilde{U}_t = (X_t, I_d)$, so that $A(X_{t-}, J_t)$ is indeed just a function of X_{t-} and J_t and there is no dependence on the process R . Therefore, we can rewrite SDE (4.27) in matrix form as

$$R(t) = R(0) + \int_0^t F(R(s-)) \circ d\varphi(X_s) + \int_0^t A(X_{s-}, J_s) R(s-) dZ_s. \quad (4.29)$$

Note that the integrands in (4.29) are linear in $R(s-)$ with random linear coefficients given by $\Gamma_{jl}^k(X_{s-})$ and $A(X_{s-}, J_s)$. Thus (4.29) is a standard form of an SDE against semimartingales with the integrands being linear, hence functional Lipschitz (in the terminology of [Pro92, Ch. V]). More precisely, the integrands are obtained via an action of a functional Lipschitz operator on the process R . Functional Lipschitzness is a consequence of linearity and a.s. boundedness of linear coefficients on finite intervals. Indeed, for a.e. ω and every $t > 0$ we have $\sup_{s \leq t} |\Gamma_{jl}^k(X_{s-}(\omega))| < \infty$, which follows from the càdlàg property of X and the smoothness of Christoffel symbols. Moreover, it also holds that $\sup_{s \leq t} |A(X_{s-}(\omega), J_s(\omega))| < \infty$. Observe first that A is non-zero only when J is non-zero, which holds for at most countably many times, and further for every $\varepsilon > 0$ we have $|J_s(\omega)| > \varepsilon$ for at most finitely times s . Thus, the finiteness of the supremum holds trivially at finitely many large-jump times. We may assume that all of the (possibly infinitely many) remaining small jumps are bounded by a constant ε . By the boundedness of $\tilde{U}_{s-}^{-1}(\omega)q(X_{s-}(\omega))$ by some constant K (see Remark 4.2.18 (ii)) we can therefore consider ε small enough, such that the transformed jumps $\tilde{U}_{s-}^{-1}(\omega)q(X_{s-}(\omega))J_s(\omega)$, which are bounded by εK , are small enough, so that we may apply Lemma 4.4.4 with $c = \tilde{U}_{s-}^{-1}(\omega)q(X_{s-}(\omega))J_s(\omega)$ and $u = \tilde{U}_{s-}(\omega) = (x^j, r_m^k) = (\varphi^j(X_{s-}(\omega)), \delta_m^k)$. Using (4.25) we compute

$$\begin{aligned} A(X_{s-}, J_s)_m^k &= \left(\delta_m^k - c^j \Gamma_{jm}^k(X_{s-}(\omega)) + c^n c^{n'} F_{nn'}(\psi(\theta))_m^k - \delta_m^k \right) / |J_s(\omega)|^2 \\ &\quad + \Gamma_{jm}^k(X_{s-}(\omega)) \left(\varphi^j(X_{s-}(\omega)) + c^j + c^n c^{n'} F_{nn'}(\psi(\theta))^j - \varphi^j(X_{s-}(\omega)) \right) / |J_s(\omega)|^2 \\ &= c^n c^{n'} \left(F_{nn'}(\psi(\theta))_m^k + \Gamma_{jm}^k(X_{s-}(\omega)) F_{nn'}(\psi(\theta))^j \right) / |J_s(\omega)|^2 \\ &\leq K^2 \sum_{n, n'=1}^d \left(F_{nn'}(\psi(\theta))_m^k + \Gamma_{jm}^k(X_{s-}(\omega)) F_{nn'}(\psi(\theta))^j \right), \end{aligned}$$

from which boundedness of supremum of A follows by using the càdlàg property of X , smoothness of maps Γ_{jm}^k and $F_{nn'}$, and the fact that integral curves ψ stay inside a compact set due to small sizes of jumps and the càdlàg property of X .

Thus, the coefficients of the SDE are indeed functional Lipschitz and [Pro92, Thm. 7 in Ch. V] implies that there exists a unique solution R of the SDE (4.29) and the linearity of coefficients implies non-explosion of the solution, i.e. process R is defined until the stopping time $\tilde{\tau}$. Actually, there is a potential issue as the coefficient $A(X_{s-}, J_s)R(s-)$ is not predictable, but it is still optionally measurable w.r.t. $(\mathcal{G}_t)_{t \in \mathbb{R}_+}$ and the integral is against an increasing finite variation process so we may still use a trivial extension of [Pro92, Thm. 7 in Ch. V] (see also remarks in the proof of [KPP95, Thm. 3.2]).

We claim that the frame bundle valued process U given in local coordinates as $U = (X, R)$ is exactly the horizontal lift of X . It is clear that U is a lift of X , i.e. it projects to X , but we still need to check that it is horizontal, more precisely, we need to identify its anti-development. The candidate for the anti-development is given by equation (4.28), initially only for $t < \tilde{\tau}$, and similar calculation as above shows that the sum in (4.28) is almost surely convergent so that W is a semimartingale. We need to check that W is an anti-development of U , i.e. we need to show that U solves Marcus SDE (4.1) when the integrator W is given by (4.28). It is enough to check the defining property of Marcus SDE (4.1) for functions which constitute a local chart on $\pi^{-1}(O)$.

We first take $f = \phi^\alpha$, so that $\phi^\alpha(U_t) = \varphi^\alpha(X_t)$ and note that by the representation in local coordinates of fundamental horizontal vector fields from Lemma 1.3.1 we have $H_i \phi^\alpha(u) = r_i^\alpha$ for $u = (x^j, r_m^k)$. Therefore,

$$\begin{aligned}
& \int_0^t H_i \phi^\alpha(U_{s-}) \circ dW_s^i \\
& + \sum_{0 < s \leq t} (\phi^\alpha(\exp(H_i \Delta W_s^i)(U_{s-})) - \phi^\alpha(U_{s-}) - H_i \phi^\alpha(U_{s-}) \Delta W_s^i) \\
& = \int_0^t R_i^\alpha(s-) \bar{R}_j^i(s-) \circ d\varphi^j(X_s) \\
& + \sum_{0 < s \leq t} (R_i^\alpha(s-)(U_{s-}^{-1} q(X_{s-}) J_s)^i - R_i^\alpha(s-) \bar{R}_j^i(s-) \Delta \varphi^j(X_s)) \\
& + \sum_{0 < s \leq t} (\varphi^\alpha(X_s) - \varphi^\alpha(X_{s-}) - R_i^\alpha(s-) \Delta W_s^i) \\
& = \int_0^t 1 \circ d\varphi^\alpha(X_s) + \sum_{0 < s \leq t} (R_i^\alpha(s-) \Delta W_s^i - \Delta \varphi^\alpha(X_s) + \Delta \varphi^\alpha(X_s) - R_i^\alpha(s-) \Delta W_s^i) \\
& = \varphi^\alpha(X_s) - \varphi^\alpha(X_0) = \phi^\alpha(U_t) - \phi^\alpha(U_0)
\end{aligned}$$

holds, where we used (4.28) in the first equality and (4.11) in the second inequality. This establishes the defining property of Marcus SDE (4.1) for $f = \phi^\alpha$.

We now take $f = \phi_m^k$, so that $\phi_m^k(U_s) = R_m^k(s)$ and note that Lemma 1.3.1 implies the equality $H_i \phi_m^k(u) = -r_i^\beta r_m^l \Gamma_{\beta l}^k(p)$ for $u = (p^i, r_m^k)$. Then

$$\begin{aligned}
& \int_0^t H_i \phi_m^k(U_{s-}) \circ dW_s^i \\
& + \sum_{0 < s \leq t} \left(\phi_m^k(\exp(H_i \Delta W_s^i)(U_{s-})) - \phi_m^k(U_{s-}) - H_i \phi_m^k(U_{s-}) \Delta W_s^i \right) \\
& = - \int_0^t R_i^\beta(s-) R_m^l(s-) \Gamma_{\beta l}^k(X_{s-}) \bar{R}_j^i(s-) \circ d\varphi^j(X_s) \\
& + \sum_{0 < s \leq t} \left(-R_i^\beta(s-) R_m^l(s-) \Gamma_{\beta l}^k(X_{s-}) \Delta W_s^i + R_i^\beta(s-) R_m^l(s-) \Gamma_{\beta l}^k(X_{s-}) \bar{R}_j^i(s-) \Delta \varphi^j(X_s) \right) \\
& + \sum_{0 < s \leq t} \left(\phi_m^k(\exp(H_{U_{s-}^{-1} q(X_{s-}) J_s})(U_{s-})) - \phi_m^k(U_{s-}) + R_i^\beta(s-) R_m^l(s-) \Gamma_{\beta l}^k(X_{s-}) \Delta W_s^i \right) \\
& = - \int_0^t R_m^l(s-) \Gamma_{jl}^k(X_{s-}) \circ d\varphi^j(X_s) \\
& + \sum_{0 < s \leq t} \left(\phi_m^k(\exp(H_{U_{s-}^{-1} q(X_{s-}) J_s})(U_{s-})) - \phi_m^k(U_{s-}) + R_m^l(s-) \Gamma_{jl}^k(X_{s-}) \Delta \varphi^j(X_s) \right) \\
& = R_m^k(t) - R_m^k(0) = \phi_m^k(U_t) - \phi_m^k(U_0)
\end{aligned}$$

holds, where we used (4.28) in the first equality and (4.27) in the penultimate equality and this establishes the defining property of Marcus SDE (4.1) for $f = \phi_m^k$. Thus, we have established that W is an anti-development of U (and X), hence U is indeed a horizontal process and a horizontal lift of X .

We want to extend processes U and W to a larger time interval. Note that we can find a (random) sequence of precompact charts O_n such that $X_{\tau_{n-1}} \in O_n$ for $n \in \mathbb{N}$ where $\tau_0 = 0$ and $\tau_n = \inf \{s > \tau_{n-1} ; X_s \notin O_n\}$ for $n \in \mathbb{N}$ and such that $\tau_n \rightarrow \tau$ (the lifetime of X) as $n \rightarrow \infty$. Construction from above first defines horizontal lift and anti-development on $[0, \tau_1)$. At time τ_1 we then define $U_{\tau_1} = \exp\left(H_{U_{\tau_1-}^{-1} q(X_{\tau_1-}) J_{\tau_1}}\right)(U_{\tau_1-})$ and $W_{\tau_1} = W_{\tau_1-} + U_{\tau_1-}^{-1} q(X_{\tau_1-}) J_{\tau_1}$ and this makes U horizontal lift and W an anti-development of X on $[0, \tau_1]$. Since $U_{\tau_1} \in \pi^{-1}(O_2)$ we can reuse the construction from above, where now the initial point is U_{τ_1} and this defines processes U and W also on $[\tau_1, \tau_2)$. Iteratively, this defines the horizontal lift and anti-development on $[0, \tau_n)$ so by sending $n \rightarrow \infty$, the horizontal lift and anti-development are semimartingales defined until the lifetime of X .

We still need to prove the uniqueness of horizontal lift and anti-development and hence independence of selection of local charts used in the construction. We will proceed similarly to [PE92, Thm. 3.2]. Suppose that U, W are a horizontal lift and anti-development, respectively, induced by X, J and suppose V, \widetilde{W} are another such pair with $U_0 = V_0$. Since $\pi(U) = \pi(V) = X$, there is a unique càdlàg process g_s with values in $\text{GL}(d)$ such that $V_s = U_s g_s = \rho(U_s, g_s)$, where $\rho: F(M) \times \text{GL}(d) \rightarrow M$ is the right group action of $\text{GL}(d)$ on the frame bundle. In particular, $g_0 = e$ and the equality

$$\Delta \widetilde{W}_t = V_{t-}^{-1} q(X_{t-}) J_t = g_{t-}^{-1} U_{t-}^{-1} q(X_{t-}) J_t = g_{t-}^{-1} \Delta W_t$$

holds. Actually, it is not hard to see that

$$\widetilde{W}_t = \int_0^t g_s^{-1} \circ dW_t \quad (4.30)$$

holds. Since W (resp. \widetilde{W}) is an anti-development of U (resp. V), we can use the appropriate versions of (4.10) to compute

$$U_t g_t = V_t = \exp(H_{\Delta \widetilde{W}_t})(V_{t-}) = \exp(H_{g_{t-}^{-1} \Delta W_t})(U_{t-} g_{t-}) = \exp(H_{\Delta W_t})(U_{t-}) g_{t-} = U_t g_{t-},$$

where we used Lemma 1.3.3 in the penultimate equality. It follows that $g_t = g_{t-}$ and hence the process g_t is continuous.

To prove uniqueness it is equivalent to show that the process g_t from above is constantly equal to the identity. We consider local coordinates u^β on $F(M)$ (these are all ϕ^α and ϕ_m^k from above combined) and local coordinates h^k on $\text{GL}(d)$. For ease of notation we denote $V_t^\alpha = u^\alpha(V_t)$, $U_t^\beta = u^\beta(U_t)$ and $g_t^k = h^k(g_t)$. By the defining property of Marcus SDEs we have

$$\begin{aligned} V_t^\alpha &= V_0^\alpha + \int_0^t H_i u^\alpha(V_{s-}) \circ d\widetilde{W}_s^i \\ &\quad + \sum_{0 < s \leq t} \left(u^\alpha \left(\exp(H_i \Delta \widetilde{W}_s^i)(V_{s-}) \right) - V_{s-}^\alpha - H_i u^\alpha(U_{s-}) \Delta \widetilde{W}_s^i \right) \\ &= V_0^\alpha + \int_0^t H_i (u^\alpha \circ R_{g_s})(U_{s-}) \circ dW_s^i \\ &\quad + \sum_{0 < s \leq t} (V_s^\alpha - V_{s-}^\alpha - H_i (u^\alpha \circ R_{g_s})(U_{s-}) \Delta W_s^i), \end{aligned} \quad (4.31)$$

where we have used (4.30) and Lemma 1.3.4. We also define $\rho^\alpha = u^\alpha \circ \rho$ so that we can

alternatively write $V_t^\alpha = \rho^\alpha(U_t, g_t)$ and we use Itô's formula to compute

$$\begin{aligned}
V_t^\alpha &= V_0^\alpha + \int_0^t \frac{\partial \rho^\alpha}{\partial u^\beta}(U_{s-}, g_s) \circ dU_s^\beta + \int_0^t \frac{\partial \phi^\alpha}{\partial h^k}(U_{s-}, g_s) \circ dg_s^k \\
&\quad + \sum_{0 < s \leq t} \left(\rho^\alpha(U_s, g_s) - \rho^\alpha(U_{s-}, g_s) - \frac{\partial \rho^\alpha}{\partial u^\beta}(U_{s-}, g_s) \Delta U_s^\beta \right) \\
&= V_0^\alpha + \int_0^t \frac{\partial \rho^\alpha}{\partial u^\beta}(U_{s-}, g_s) H_i u^\beta(U_{s-}) \circ W_s^i \\
&\quad + \sum_{0 < s \leq t} \frac{\partial \rho^\alpha}{\partial u^\beta}(U_{s-}, g_s) \left(U_s^\beta - U_{s-}^\beta - H_i u^\beta(U_{s-}) \Delta W_s^i \right) \\
&\quad + \int_0^t \frac{\partial \phi^\alpha}{\partial h^k}(U_{s-}, g_s) \circ dg_s^k + \sum_{0 < s \leq t} \left(V_s^\alpha - V_{s-}^\alpha - \frac{\partial \rho^\alpha}{\partial u^\beta}(U_{s-}, g_s) \Delta U_s^\beta \right),
\end{aligned}$$

where we used that g_t is continuous and an analogous version of (4.31) for U . Since

$$\frac{\partial \rho^\alpha}{\partial u^\beta}(U_{s-}, g_s) H_i u^\beta(U_{s-}) = H_i(u^\alpha \circ R_{g_s})(U_{s-})$$

holds by the chain rule, we see by comparing the above expressions for V_t^α that

$$\int_0^t \frac{\partial \phi^\alpha}{\partial h^k}(U_{s-}, g_s) \circ dg_s^k = 0$$

holds for all α . Since the matrix $[\frac{\partial \phi^\alpha}{\partial h^k}]_k^\alpha$ is of full rank (this is true since $g \mapsto \rho(u, g)$ is an embedding), we can conclude that

$$\int_0^t 1 \circ dg_s^k = 0$$

holds for all k . Therefore, the process g_t is constant and hence constantly equal to identity. This proves that the horizontal lift and anti-development are unique and this completes the proof of the theorem. \square

4.4.4 Proof of Theorem 4.2.14

We already know from Theorem 4.2.19 that the anti-development (when it exists) is uniquely determined on $[0, \tau)$ by the pair of processes X and J . Due to a particular jump structure of X inferred from its generator, we will see that there will always exist a process J satisfying Definition 4.2.17, but we need to find an appropriate one such that the anti-development will be a Euclidean Lévy process.

Hence, for the proof of Theorem 4.2.14 we need to construct a suitable jump process J (on the extended probability space) and show that the associated anti-development is a Lévy process with given characteristics. Since the anti-development is a priori only defined on $[0, \tau)$, we also need to show that it can be extended past τ .

We will first focus on the construction of the appropriate jump process J , or more precisely, on the construction of the associated random measure on $\mathbb{R}^d \times \mathbb{R}_+$. This simple random measure, which we will also denote by J has (random) unit atoms at (s, J_s) when $J_s \neq 0$. Before constructing it, we want to characterize it. The following lemma will be useful.

Lemma 4.4.5. *Let G be a subgroup of $GL(d)$ and let N be a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d$ with a mean measure $\mathcal{L}eb \otimes \nu$, where ν is a G -invariant Lévy measure. Further, let $(g_t)_{t \in \mathbb{R}_+}$ be a predictable process with values in G . Denote by Φ a random transformation of $\mathbb{R}_+ \times \mathbb{R}^d$ given by*

$$\Phi(\omega, t, z) = (t, g_t(\omega)z).$$

*Then the transformed (push-forward) measure $\Phi_*N := N\Phi^{-1}$ is also a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d$ with a mean measure $\mathcal{L}eb \otimes \nu$.*

Proof. The compensator measure of N is its mean measure $\mu := \mathcal{L}eb \otimes \nu$. It is enough to show that $\Phi_*\mu = \mu$, as we can then use [Kal90, Thm. 2.1] to deduce the claim. Let us take an arbitrary $I \in \mathfrak{B}(\mathbb{R}_+)$, $A \in \mathfrak{B}(\mathbb{R}^d)$ and then

$$\Phi_*\mu(\omega)(I \times A) = \mu(\Phi^{-1}(\omega)(I \times A)) = \int_I \nu(g_t(\omega)A) dt = \int_I \nu(A) dt = \mathcal{L}eb(I)\nu(A)$$

holds for almost every ω , where we used Fubini-Tonelli theorem in the second equality and a G -invariance of the measure ν in the third equality. Since sets $I \times A$ generate $\mathfrak{B}(\mathbb{R}_+ \times \mathbb{R}^d)$, we have shown that $\Phi_*\mu$ is non-random and equal to μ , so the claim of the lemma follows. \square

Therefore, if Y is a Lévy process, with jumps given by a Poisson random measure N , which is an anti-development of X , then we know that J needs to be given as a transformation of N where each atom at (t, z) with $t < \tau$ is transformed to the atom at $(t, q^{-1}(X_{t-})U_{t-}z)$, where U is solution of Marcus SDE (4.3) and q is a section of the holonomy bundle as in Section 4.2.4. We do not transform the atoms at (t, z) with $t \geq \tau$. But $q^{-1}(X_-)U_-$ is a predictable $\text{Hol}(u)$ -valued process (we trivially extend it past τ) and Lévy measure ν of Y is $\text{Hol}(u)$ -invariant, hence Lemma 4.4.5 implies that J is also a

Poisson random measure with mean measure $\mathcal{L}eb \otimes \nu$. Note that the condition equivalent to Definition 4.2.17 which relates the random measure J and a process X is

$$\begin{aligned} J(\{(s,z)\}) = 1 &\implies X_s = \text{Exp}_{X_{s-}}(q(X_{s-})z) \\ J(\{s\} \times \mathbb{R}^d) = 0 &\implies X_s = X_{s-} \end{aligned} \quad \text{for } s < \tau. \quad (4.32)$$

Hence the goal is to find suitable J which is a Poisson random measure and we have the following result.

Proposition 4.4.6. *Let X be a Markov process on M adapted to $(\mathcal{H}_t)_{t \in \mathbb{R}_+}$ with the generator \mathcal{L}_X given by (4.6). Then on the extended probability space there exists a Poisson random measure J on $\mathbb{R}_+ \times \mathbb{R}^d$ with a mean measure $\mathcal{L}eb \otimes \nu$ and adapted to an extended filtration $(\mathcal{G}_t)_{t \in \mathbb{R}_+}$ such that (4.32) holds.*

Before proceeding with the proof, we explain the phenomena that may occur when constructing J , and they illustrate why typically the process X alone does not possess sufficient information and randomness (even on $[0, \tau)$) to carry out the construction .

- (i) Let $X_s = y \neq X_{s-} = x$, such that $B_{xy} := \{z \in \mathbb{R}^d \setminus \{0\} ; \text{Exp}_x(q(x)z) = y\}$ contains exactly one element z' . Then it is necessary to put $J(\{(s,z')\}) = 1$.
- (ii) Again let $X_s = y \neq X_{s-} = x$, but now the set B_{xy} contains more than one element⁶. Of course, it is clear that we should have $J(\{s\} \times B_{xy}) = 1$, but it is not clear which element should be chosen as the atom. If we simply fixed one possible element of B_{xy} for every pair (x,y) , this would correctly prescribe the jumps of X and satisfy (4.32), but J might not have the correct intensity measure $\mathcal{L}eb \otimes \nu$. Therefore, in this case we need some additional randomness to appropriately randomly assign the atom on the set B_{xy} .
- (iii) Finally, it is possible that $X_s = X_{s-} = x$ and there can still be jumps (atoms) in J (see also the discussion following Definition 4.2.15). This can happen since the set $B_x := \{z \in \mathbb{R}^d \setminus \{0\} ; \text{Exp}_x(q(x)z) = x\}$ may be non-empty. Thus, even if $J(\{s\} \times B_x) = 1$, actually no jump of X occurs at time s . Naively, we might simply ignore these jumps, but then again the intensity measure of J might not be equal to $\mathcal{L}eb \otimes \nu$. Therefore, if we wish to get a correct Poisson random measure, we need to add some of these jumps (atoms) as well, even though they will not manifest themselves as jumps of X . They will have to be taken from an additional

⁶These sets are a.s. non-empty due to the form of a generator of X .

(independent) Poisson random measure on the extended probability space which we will also have to use to define J after time τ .

This illustrates that for a general process X we will need to extend our probability space (and filtration) to allow for this additional randomness. This is done in [ÇJ81, Section 4], albeit they only consider a measure ν on a real line, but essentially every step generalizes to our case of a multidimensional measure. The only step which needs some adjustment due to a slightly different setup is the following. For ease of notation, put $k(x,z) = \text{Exp}_x(q(x)z)$ and note that we could have written the jumps in the integral part of generator \mathcal{L}_X in (4.6) via the kernel $K: M \times \mathcal{B}(M) \rightarrow \mathbb{R}_+$ defined by

$$K(x,A) = \int_{\mathbb{R}^d} \mathbf{1}(k(x,z) \in A) \nu(dz) \text{ for } A \in \mathcal{B}(M \setminus \{x\}),$$

and this means that the equivalent of [ÇJ81, Lemma 3.4] holds. Furthermore, the following equivalent of [ÇJ81, Lemma 4.32] holds.

Lemma 4.4.7. *Let K be a kernel as above. Then there exists a measurable function $\tilde{k}: M \times M \times (0,1] \rightarrow \mathbb{R}^d$, such that*

$$\int_0^1 \int_A \mathbf{1}(\tilde{k}(x,y,u) \in B) K(x,dy) du = \int_B \mathbf{1}(k(x,z) \in A) \nu(dz)$$

for every $A \in \mathcal{B}(M), B \in \mathcal{B}(\mathbb{R}^d)$ and $x \in M$. Moreover, for every $x \in M$ and $K(x,dy) du$ almost every (y,u) we have

$$k(x, \tilde{k}(x,y,u)) = y.$$

Proof of Proposition 4.4.6. By using Lemma 4.4.7 we may deduce as in [ÇJ81, §4c, Eq. (4.41) and Prop. 4.42] that we can construct a random measure \bar{J} on the extended probability space, such that (4.32) holds (with \bar{J} taking the role of J) and such that the compensator measure of \bar{J} is given by $\mathbf{1}(k(X_{t-}, z) \neq X_{t-}, t < \tau) dt \nu(dz)$. This is almost the required random measure. To add the missing part, we consider on a further extended probability space an independent Poisson random measure N with an intensity measure $\mathcal{L}eb \otimes \nu$ and define a random measure

$$\begin{aligned} J(dt, dz) &= \bar{J}(dt, dz) \\ &+ \left(\mathbf{1}(k(X_{t-}, z) = X_{t-}, \bar{J}(\{t\} \times \mathbb{R}^d) = 0, t < \tau) + \mathbf{1}(t \geq \tau) \right) N(dt, dz). \end{aligned}$$

It is easy to see that its compensator measure is a non-random measure $\mathcal{L}eb \otimes \nu$, which means that J is a Poisson random measure with mean measure $\mathcal{L}eb \otimes \nu$. Furthermore,

since added jumps (atoms) from the measure N do not manifest themselves as jumps of X , the condition (4.32) still holds and this finishes the proof of proposition. \square

Proof of Theorem 4.2.14. Proposition 4.4.6 allows us to construct a Poisson random measure J and hence the associated jump process J which represents jumps of X . Therefore, by Theorem 4.2.19 there exists a uniquely determined anti-development on $[0, \tau)$ induced by X and J and we denote it by Y . The only thing left to do is to show that Y can be extended past τ so that it is a Lévy process with correct characteristics.

First, note that the jump measure of Y is given as a transformation of J where each atom at (t, z) is transformed to the atom at $(t, U_t^{-1}q(X_{t-})z)$, and again Lemma 4.4.5 implies that the jump measure of Y is exactly a Poisson random measure with mean measure $\mathcal{L}eb \otimes \nu$ restricted to times smaller than τ .

To determine the remaining characteristics, we recall that Y is a semimartingale on a stochastic interval $[0, \tau)$. We wish to use the notion of semimartingale characteristics which are classically defined for semimartingales with infinite lifetime. However, as the Appendix in [Sch12] shows, essentially every result about semimartingales can be extended to semimartingales defined up to a predictable stopping time.

Recall that every semimartingale Y has a representation $Y = Y_0 + M + A$, where M is a local martingale started at 0 and A is a process of locally finite variation, but this decomposition is not unique. To get uniqueness we define a process $Y_t^e = \sum_{s \leq t} \Delta Y_s \mathbf{1}(|\Delta Y_s| \geq 1)$ and then a process $Y - Y^e$ with bounded jumps is a special semimartingale, meaning that it has a unique representation $Y - Y^e = M + A$, where A is a predictable finite variation process and M is a local martingale. Then the local characteristics $(A, C, \tilde{\Gamma})$ of a semimartingale Y as in [JS03, Chapter II] are given by:

- (i) $A = (A^j)_{j=1, \dots, n}$ is a unique predictable finite variation process in the decomposition of a special semimartingale $Y - Y^e$,
- (ii) $C = (C^{ij})_{i, j=1, \dots, n}$ is a quadratic variation of the continuous local martingale part Y^c of Y (or equivalently of M) and is given by $C^{ij} = [Y^{i,c}, Y^{j,c}]$,
- (iii) $\tilde{\Gamma}$ is a predictable compensator of the jump measure Γ of the process Y where $\Gamma(dt, dy) = \sum_{s \leq t} \mathbf{1}(|\Delta Y_s| > 0) \delta_{(s, \Delta Y_s)}(dt, dy)$.

We have already established, that the jump measure of Y is a Poisson random measure, hence the third characteristic is equal to a non-random measure $\mathcal{L}eb \otimes \nu$.

The semimartingale Y can be written explicitly via its characteristics as in [JS03, Thm. II.2.34] and plugging this decomposition into the defining property (4.2) of Marcus

SDE (4.3) for a function $f \circ \pi$, where $f \in C^\infty(M)$, we see that

$$\begin{aligned} & f(X_t) - f(X_0) - \int_0^t \left(H_i(f \circ \pi)(U_{s-}) dA^i + \frac{1}{2} H_i H_j (f \circ \pi)(U_{s-}) dC_s^{ij} \right) \\ & - \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} (f(\text{Exp}_{X_{s-}}(U_{s-}x)) - f(X_{s-}) - \mathbf{1}(|x| < 1) x^i H_i(f \circ \pi)(U_{s-})) \tilde{\Gamma}(ds, dx) \end{aligned}$$

is a local martingale. We can compare the above expression to the generator \mathcal{L}_X from (4.6) and we get

$$A_t^j = b^j t, \quad C_t^{ij} = a^{ij} t, \quad \tilde{\Gamma}(dt, dz) = dt \nu(dz), \quad (4.33)$$

which are the characteristics of Y for $t < \tau$. Note that by observing the generator we can only uniquely determine A and C , and there can be more than one measure $\tilde{\Gamma}$ for which the above equality holds (this can happen since functions $z \mapsto k(x, z)$ from above are not necessarily injective). However, we have already identified the third characteristic without directly referring to the generator.

By [JS03, Thm. II.4.19] we know that (“constant”) characteristics as in (4.33) imply that the semimartingale is actually a Lévy process. However, Y is only defined for times in $[0, \tau)$ and we need to extend it to all times in \mathbb{R}_+ . If we can find a Lévy process $\tilde{Y} = (\tilde{Y}_t)_{t \in \mathbb{R}_+}$ with characteristics (4.33) such that $\tilde{Y}_t = Y_t$ for $t \in [0, \tau)$, then \tilde{Y} is also an anti-development of X and it is exactly the Lévy process we needed to construct to prove the theorem. Existence of such a process \tilde{Y} is guaranteed by the following lemma and this concludes the proof of the theorem. \square

Lemma 4.4.8. *Let Y be a semimartingale on $[0, \tau)$ with characteristics (4.33) for $t < \tau$, where τ is a predictable stopping time. Then there exists a Lévy process $\tilde{Y} = (\tilde{Y}_t)_{t \in \mathbb{R}_+}$ with characteristics (4.33) such that $\tilde{Y}_t = Y_t$ for $t \in [0, \tau)$.*

Proof. Let $\{\tau_n ; n \in \mathbb{N}\}$ be an announcing sequence of τ . Consider an additional Lévy process Z with characteristics (4.33) which is independent of Y . For every $n \in \mathbb{N}$ we may define a process $Z^{(n)} = (Z_t^{(n)})_{t \in \mathbb{R}_+}$ by

$$Z_t^{(n)} := Y_t^{\tau_n} + \mathbf{1}(t \geq \tau_n)(Z_t - Z_{\tau_n})$$

for $t \in \mathbb{R}_+$. Since characteristics of $Z^{(n)}$ are also given by (4.33), it is a Lévy process and

obviously $Z_t^{(n)} = Y_t$ holds for $t \in [0, \tau_n]$. Lévy process $Z^{(n)}$ has the Lévy-Itô decomposition

$$Z_t^{(n),i} = b^i t + \sigma_j^i B_t^{(n),j} + \int_0^t \int_{0 < |x| < 1} x^i \widetilde{N}^{(n)}(ds, dx) + \int_0^t \int_{|x| \geq 1} x^i N^{(n)}(ds, dx)$$

for $i = 1, \dots, d$, where $B^{(n)}$ is a standard Brownian motion and $N^{(n)}$ a Poisson random measure with mean measure $\mathcal{L}eb \otimes \nu$. Note that $N^{(n)}$ is equal to the jump measure Γ of Y on $[0, \tau_n]$. Furthermore, for $m < n$ we deduce from $Z_t^{(n)} = Z_t^{(m)}$ for $t \in [0, \tau_m]$ that also $B_t^{(n)} = B_t^{(m)}$ for $t \in [0, \tau_m]$, so that we may unambiguously define a continuous process $B = (B_t)_{t \in [0, \tau]}$ by $B_t = B_t^{(n)}$ for $t \in [0, \tau_n]$.

Let us assume that there exists a Brownian motion $\widetilde{B} = (\widetilde{B}_t)_{t \in \mathbb{R}_+}$ such that $\widetilde{B}_t = B_t$ for $t \in [0, \tau)$. We may then consider a Poisson random measure N' with mean measure $\mathcal{L}eb \otimes \nu$ which is independent of Y and \widetilde{B} and define a random measure

$$N(dt, dx) := \mathbf{1}(t < \tau) \Gamma(dt, dx) + \mathbf{1}(t \geq \tau) N'(dt, dx),$$

which is also a Poisson random measure with mean measure $\mathcal{L}eb \otimes \nu$ since $\mathcal{L}eb \otimes \nu$ is its compensator measure. Hence we may define a process $\widetilde{Y} = (\widetilde{Y}_t)_{t \in \mathbb{R}_+}$ by

$$\widetilde{Y}_t^i := b^i t + \sigma_j^i \widetilde{B}_t^j + \int_0^t \int_{0 < |x| < 1} x^i \widetilde{N}(ds, dx) + \int_0^t \int_{|x| \geq 1} x^i N(ds, dx)$$

for $i = 1, \dots, d$. It is clear that \widetilde{Y} is a Lévy process with characteristics (4.33) and since $\widetilde{B}_t = B_t$ for $t \in [0, \tau)$ and $N = \Gamma$ on $[0, \tau)$ we also have $\widetilde{Y}_t = Y_t$ for $t \in [0, \tau)$, so \widetilde{Y} is the required process.

Therefore, in order to finish the proof we only have to show the existence of Brownian motion \widetilde{B} as above. Note that since $B_t = B_t^{(n)}$ for $t \in [0, \tau_n]$, the process B is a continuous local martingale on $[0, \tau)$ (see the definition in [GS72, Section 3] or in [RY99, Exercise (1.48) in Ch. IV]) and has a quadratic covariation $\langle B^i, B^j \rangle_t = \delta^{ij} t$ on $[0, \tau)$ for $i, j = 1, \dots, d$. For $m \in \mathbb{N}$ consider a constant stopping time $T_m \equiv m$ and a stopped process B^{T_m} , which is also a continuous local martingale on $[0, \tau)$. We have a trivial bound $\langle B^{T_m, i} \rangle_\tau \leq m$ for $i = 1, \dots, d$, so we can conclude from the Burkholder-Davis-Gundy inequality [RY99, Thm. (4.1) in Ch. IV] that $\mathbb{E}[\sup_{t < \tau} |B_t^{T_m}|] < \infty$ and it then follows that the limit $\lim_{t \rightarrow \tau} B_t^{T_m}$ exists (see the discussion after [GS72, Def. 3.1] and [RY99, Exercise (1.48) in Ch. IV]). On $\{\tau < \infty\}$ we have that $T_m \geq \tau$ from some sufficiently large m onward, so also the limit $\lim_{t \rightarrow \tau} B_t$ exists on the same event. As we may define $B_\tau := \lim_{t \rightarrow \tau} B_t$ on $\{\tau < \infty\}$ it becomes meaningful to consider a continuous stopped

process B^τ and we can define a continuous process $(\tilde{B}_t)_{t \in \mathbb{R}_+}$ by

$$\tilde{B}_t = B_t^\tau + \mathbf{1}(t \geq \tau)(B_t' - B_\tau'),$$

where $(B_t')_{t \in \mathbb{R}_+}$ is another Brownian motion independent of B . It is easily shown by Lévy's characterization that \tilde{B} is also a Brownian motion and obviously $\tilde{B}_t = B_t$ holds for $t \in [0, \tau)$, so it is the required process needed to finish the proof of the lemma. \square

4.4.5 Proofs of results in Section 4.3.2

Proof of Proposition 4.3.1. We recall the Lévy-Itô decomposition of a Lévy process Y as in (4.22). Since X solves Marcus SDE (4.13) we can take $f \in C^\infty(G)$ and substitute (4.22) into the defining property of the Marcus SDE to get

$$\begin{aligned} f(X_t) &= f(X_0) + b^i \int_0^t V_i^L f(X_{s-}) ds + \sigma_j^i \int_0^t V_i^L f(X_{s-}) \circ dB_s^j \\ &\quad + \int_0^t \int_{0 < |x| < 1} x^i V_i^L f(X_{s-}) \tilde{N}(ds, dx) + \int_0^t \int_{|x| \geq 1} x^i V_i^L f(X_{s-}) N(ds, dx) \\ &\quad + \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} (f(X_{s-} \widetilde{\exp}(x)) - f(X_{s-}) - x^i V_i^L f(X_{s-})) N(ds, dx). \end{aligned}$$

After splitting the domain in the last integral and some algebraic manipulation using $\tilde{N}(ds, dx) = N(ds, dx) - ds \nu(dx)$ we get

$$\begin{aligned} f(X_t) &= f(X_0) + b^i \int_0^t V_i^L f(X_{s-}) ds + \sigma_j^i \int_0^t V_i^L f(X_{s-}) \circ dB_s^j \tag{4.34} \\ &\quad + \int_0^t \int_{0 < |x| < 1} (f(X_{s-} \widetilde{\exp}(x)) - f(X_{s-})) \tilde{N}(ds, dx) \\ &\quad + \int_0^t \int_{|x| \geq 1} (f(X_{s-} \widetilde{\exp}(x)) - f(X_{s-})) N(ds, dx) \\ &\quad + \int_0^t \int_{0 < |x| < 1} (f(X_{s-} \widetilde{\exp}(x)) - f(X_{s-}) - x^i V_i^L f(X_{s-})) ds \nu(dx). \end{aligned}$$

We recall that $\widetilde{\exp}$ is a diffeomorphism from some neighbourhood O of $0 \in \mathbb{R}^d$ to some neighbourhood U of the unit $e \in G$ and we may assume that $O \subseteq \{x \in \mathbb{R}^d ; |x| < 1\}$. We take a smaller open set \tilde{O} with $0 \in O$ and $\text{cl}(\tilde{O}) \subseteq O$ and let $\tilde{U} := \widetilde{\exp}(\tilde{O})$. Then we can find $l \in C^\infty(\mathbb{R}^d, \mathbb{R}^d)$ such that $l(x) = x$ on $\text{cl}(\tilde{O})$ and $\text{supp } l \subseteq O$. Given such a

function l , we define $\xi \in C^\infty(G, \mathbb{R}^d)$ by

$$\xi(h) = \begin{cases} 0, & h \in \widetilde{\text{exp}}(\text{supp } l)^c \\ l(\widetilde{\text{exp}}^{-1}(h)), & h \in U. \end{cases}$$

Function ξ is smooth since it is smooth on both open domains and agrees on their intersection. We can check that $l(x) = \xi(\widetilde{\text{exp}}(x))$ holds. Moreover, simple calculations show that $\xi(e) = 0$ and $V_j^L \xi^i(e) = \delta_j^i$. This means that functions ξ^i are suitable coordinate functions on the neighbourhood of e and since $\int_G \sum_{i=1}^d (\xi^i(h))^2 \mu(dh) = \int_{\mathbb{R}^d} \sum_{i=1}^d (l^i(x))^2 \nu(dx) < \infty$, $\mu(\{e\}) = 0$ and $\mu(A^c) = \nu(\widetilde{\text{exp}}^{-1}(A^c)) < \infty$ for any open $A \ni e$, the measure μ is a indeed Lévy measure on G (see [Lia04, Ch. 1] for definitions).

Then the last integral in (4.34) can be expressed as

$$\begin{aligned} & \int_0^t \int_{0 < |x| < 1} (f(X_{s-} \widetilde{\text{exp}}(x)) - f(X_{s-}) - \xi^i(\widetilde{\text{exp}}(x)) V_i^L f(X_{s-})) ds \nu(dx) \\ & + c^i \int_0^t V_i^L f(X_{s-}) ds \end{aligned} \quad (4.35)$$

where $c^i = \int_{0 < |x| < 1} (l^i(x) - x^i) \nu(dx)$ for $i = 1, \dots, d$ are finite since $l(x) = x$ near 0. We then plug (4.35) back into (4.34), use the fact that $l(x) = \xi(\widetilde{\text{exp}}(x)) = 0$ for $|x| > 1$ and manipulate the integral in the third line with $\widetilde{N}(ds, dx) = N(ds, dx) - ds \nu(dx)$ to get

$$\begin{aligned} f(X_t) &= f(X_0) + \widetilde{b}^i \int_0^t V_i^L f(X_{s-}) ds + \sigma_j^i \int_0^t V_i^L f(X_{s-}) \circ dB_s^j \\ &+ \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} (f(X_{s-} \widetilde{\text{exp}}(x)) - f(X_{s-})) \widetilde{N}(ds, dx) \\ &+ \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} (f(X_{s-} \widetilde{\text{exp}}(x)) - f(X_{s-}) - \xi^i(\widetilde{\text{exp}}(x)) V_i^L f(X_{s-})) ds \nu(dx) \end{aligned} \quad (4.36)$$

where $\widetilde{b}^i = b^i + c^i$. Additionally, we define a Poisson random measure N' on $\mathbb{R}_+ \times (G \setminus \{e\})$ by $N'([0, t], A') = N([0, t], \widetilde{\text{exp}}^{-1}(A'))$, where $A' \subseteq G \setminus \{e\}$ is any measurable set. This means that mean measure of N' is $\mathcal{L}eb \otimes \mu$ and we also define a compensator $\widetilde{N}'([0, t], A') = N'([0, t], A') - t\mu(A)$. We then use a change-of-variables formula for the last two integrals

in (4.36) to get

$$\begin{aligned}
f(X_t) &= f(X_0) + \widetilde{b}^i \int_0^t V_i^L f(X_{s-}) ds + \sigma_j^i \int_0^t V_i^L f(X_{s-}) \circ dB_s^j \\
&+ \int_0^t \int_{G \setminus \{e\}} (f(X_{s-h}) - f(X_{s-})) \widetilde{N}'(ds, dh) \\
&+ \int_0^t \int_{G \setminus \{e\}} (f(X_{s-h}) - f(X_{s-}) - \xi^i(h) V_i^L f(X_{s-})) ds \mu(dh).
\end{aligned}$$

Technically, after using a change-of-variables formula we get integration against measure $\widetilde{\mu} := \widetilde{\exp}_* \nu$ and the Poisson random measure corresponding to that measure with the domain of the integration being the whole group G , but since both integrands vanish at e , we can instead use μ and \widetilde{N}' , and the domain of integration reduces to $G \setminus \{e\}$. This last stochastic integral equation for X is also known as the Lévy-Itô decomposition of a Lévy process in a Lie group (see [Lia04, Thm. 1.2]) and it indeed follows that X is a (left) Lévy process with a Lévy measure μ . \square

Proof of Proposition 4.3.2. The *only if* part is a consequence of Proposition 4.3.1 so we now focus on the *if* part. The generator of a (left) Lévy process is identified in [Lia04, Thm. 1.1]. If we are able to find a Lévy measure ν on \mathbb{R}^d such that $\mu = \widetilde{\exp}_* \nu|_{G \setminus \{e\}}$ holds, we may then reverse the computations in the proof of Proposition 4.3.1 to express the generator of a left Lévy process in a form on which we may use the characterization from Theorem 4.2.14 and thus find a required process Y .

To construct such a suitable measure ν we first take open sets $U \subseteq G$ and $O \subseteq \mathbb{R}^d$ such that $\widetilde{\exp}: O \rightarrow U$ is a diffeomorphism. This allows us to define a measure ρ_O on O by $\rho_O = \widetilde{\exp}_*^{-1}(\mu|_U)$. Since $\mu(\text{Im}(\exp)^c) = 0$, we can treat μ as a measure on $\text{Im}(\exp)$. We denote $F = \widetilde{\exp}^{-1}(U^c)$ and then the restricted map $\widetilde{\exp}: F \rightarrow \text{Im}(\exp) \cap U^c$ is still surjective. Since μ is a finite measure on $\text{Im}(\exp) \cap U^c$ we can then use [Dob07, Prop. 1.101] to show that there exists a measure ρ_F on F such that $\widetilde{\exp}_* \rho_F = \mu|_{\text{Im}(\exp) \cap U^c}$. Actually, [Dob07] deals with sub-probability measures but all of the results extend to finite measures. Finally, the measure ν on \mathbb{R}^d is defined by $\nu(A) = \rho_O(A \cap O) + \rho_F(A \cap F)$ for a measurable set A . Clearly $\mu = \widetilde{\exp}_* \nu|_{G \setminus \{e\}}$ holds and, moreover, ν is a Lévy measure since $\nu(\{0\}) = \mu(\{e\}) = 0$, $\nu(O^c) = \mu(U^c) < \infty$ and $\int_O \sum_{i=1}^d (x^i)^2 \nu(dx) = \int_U \sum_{i=1}^d (\xi^i(h))^2 \mu(dh) < \infty$, where ξ^i are coordinate functions introduced in the proof of Proposition 4.3.1. \square

Appendix A

A.1 Semimartingales

In this section we review some basic facts about semimartingales on Euclidean spaces. The exposition is based on [Pro92].

We consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, \infty]}, \mathbb{P})$. The **filtration** $(\mathcal{F}_t)_{t \in [0, \infty]}$ is an increasing collection of σ -algebras, i.e. $\mathcal{F}_s \subseteq \mathcal{F}_t$ if $s \leq t$. A **stochastic process** X (on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, \infty]}, \mathbb{P})$) with values in a state space S is a collection $(X_t)_{t \in \mathbb{R}_+}$ of S -valued random variables. A stochastic process X is **(\mathcal{F}_t) -adapted** (or simply **adapted**, when the filtration is obvious from the context) if X_t is \mathcal{F}_t -measurable for every $t \in \mathbb{R}_+$. In the rest of this section we will only consider stochastic process with values in either the real line \mathbb{R} or the Euclidean space \mathbb{R}^d for some $d \in \mathbb{N}$. For each $\omega \in \Omega$ we can look at the map $t \mapsto X_t(\omega)$ and call it a **sample path**. The process X is called **càdlàg** if it a.s. has sample paths that are right-continuous and have left limits.

A process M is a **martingale** if the following two conditions hold:

- (i) M_t is integrable, i.e. $\mathbb{E}[|M_t|] < \infty$ for all $t \in \mathbb{R}_+$;
- (ii) the martingale property holds, i.e. if $s \leq t$, then $\mathbb{E}[M_t | \mathcal{F}_s] = M_s$ a.s.

In order to introduce local martingales we need to consider stopping times and stopped processes. A random time T with values in $\mathbb{R}_+ \cup \{\infty\}$ is called an **(\mathcal{F}_t) -stopping time** (or simply a stopping time) if $\{T \leq t\} \in \mathcal{F}_t$ for every $t \in \mathbb{R}_+$. Given a stochastic process X and a stopping time T we may form a **stopped process** X^T given by $X_t^T = X_{\min\{t, T\}}$. If X is càdlàg and adapted, then so is X^T . Additionally, if M is a càdlàg martingale, then so is M^T . We say that a càdlàg adapted process M is a **local martingale**, if there exists a sequence $\{T_n\}_{n \in \mathbb{N}}$ of stopping times increasing to ∞ , such that $M^{T_n} \mathbf{1}(T_n > 0)$ is a martingale for every $n \in \mathbb{N}$.

A **variation** of a real valued function f on the interval $[a,b]$ is $V(f,[a,b]) := \sup_{\pi_n} \sum_{i=1}^n |f(t_i) - f(t_{i-1})|$, where the supremum is taken over all partitions $\pi_n : (a = t_0 < t_1 < \dots < t_{n-1} < t_n = b)$. We say that a càdlàg process A is of **finite variation** if a.s. its sample paths are of finite variation on compacts. It can be shown that one could alternatively define finite variation processes as processes which can be expressed as a difference of two increasing processes.

A càdlàg adapted stochastic process $X = (X_t)_{t \in \mathbb{R}_+}$ is called a **semimartingale** if it can be represented as:

$$X_t = X_0 + M_t + A_t,$$

where M is a local martingale and A a finite variation process, both null at 0. The class of semimartingales turns out to be the largest class of processes against which stochastic integration is possible [Pro92, Thm. 43 in Ch. III]. Another important feature of (real) semimartingales is that they have finite quadratic variation. In order to introduce it we require some further definitions.

We say that a sequence $(X^n)_{n \in \mathbb{N}}$ of processes **converges ucp** (uniformly on compact sets in probability) **to** X if

$$\sup_{0 \leq s \leq t} |X_s^n - X_s| \xrightarrow{\mathbb{P}} 0$$

as $n \rightarrow \infty$ for every $t \in \mathbb{R}_+$, where $\xrightarrow{\mathbb{P}}$ denotes convergence in probability. A **random partition** σ is a sequence

$$\sigma : 0 \leq T_0 \leq T_1 \leq \dots \leq T_k < \infty$$

of finite stopping times. A sequence of random partitions σ^n

$$\sigma^n : 0 \leq T_0^n \leq T_1^n \leq \dots \leq T_{k_n}^n < \infty$$

is said to **tend to identity** if

- (i) $\lim_{n \rightarrow \infty} T_{k_n}^n = \infty$ a.s.; and
- (ii) $\sup_k |T_k^n - T_{k-1}^n| \xrightarrow{n \rightarrow \infty} 0$ a.s.

The **quadratic variation process** of a real semimartingale X is an increasing càdlàg adapted process $[X, X] = ([X, X]_t)_{t \in \mathbb{R}_+}$ given by

$$X_0^2 + \sum_k (X^{T_{k+1}^n} - X^{T_k^n})^2 \xrightarrow{n \rightarrow \infty} [X, X]$$

with ucp convergence, where $\sigma^n : 0 \leq T_0^n \leq T_1^n \leq \dots \leq T_{k_n}^n < \infty$ is a sequence of random partitions tending to identity and $X^{T_k^n}$ denotes stopped processes. By polarization we may also obtain a **quadratic covariation** $[X, Y]$ of two semimartingales X and Y by the formula

$$[X, Y] := \frac{1}{4}([X + Y, X + Y] - [X - Y, X - Y]).$$

Alternatively, quadratic variation (and covariation) can be defined via stochastic integration [Pro92, Sec. 6 in Ch. II] and [Pro92, Thm. 22 in Ch. II] shows that the two definitions are equivalent. In particular, the quadratic variation does not depend on the choice of the random partition tending to identity used in its definition.

A.2 Lévy processes and Brownian motion

We now recall the definitions of Lévy processes and Brownian motion and collect some well-known facts about them.

Let B be an \mathbb{R}^d -valued process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, \infty)}, \mathbb{P})$. We say that B is an **(\mathcal{F}_t) -Brownian motion** if:

- (i) B is adapted and has continuous paths a.s.,
- (ii) $B_0 = 0$ a.s.,
- (iii) for $0 \leq s \leq t$, $B_t - B_s$ is independent of \mathcal{F}_s (independent increments),
- (iv) for $0 \leq s \leq t$, $B_t - B_s$ is equal in distribution to B_{t-s} (stationary increments),
- (v) for $t > 0$, the distribution of B_t is multivariate normal with mean 0 and covariance matrix tI_d .

Brownian motion is a fundamental continuous time stochastic process with well-studied properties. It is a martingale and a strong Markov process. However, we often need to study a more general class of processes which can be obtained by relaxing some of the conditions in the definition of Brownian motion.

Let Y be an \mathbb{R}^d -valued process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, \infty)}, \mathbb{P})$. We say that Y is an **(\mathcal{F}_t) -Lévy process** if:

- (i) Y is adapted and has càdlàg paths a.s.,
- (ii) $Y_0 = 0$ a.s.,
- (iii) for $0 \leq s \leq t$, $Y_t - Y_s$ is independent of \mathcal{F}_s (independent increments),

(iv) for $0 \leq s \leq t$, $Y_t - Y_s$ is equal in distribution to Y_{t-s} (stationary increments).

The law of the Lévy process is characterized by its characteristic function.

Theorem A.2.1 (Lévy-Khintchine formula (Theorems 7.10 and 8.1 in [Sat99])). *Let Y be a Lévy process in \mathbb{R}^d . For $t \geq 0$ the characteristic function of Y_t is given by:*

$$\mathbb{E}[e^{i\theta^\top Y_t}] = e^{-t\psi(\theta)},$$

where $\theta \in \mathbb{R}^d$, $i^2 = -1$ and the characteristic exponent $\psi: \mathbb{R}^d \rightarrow \mathbb{C}$ is given by

$$\psi(\theta) = -i b^\top \theta + \frac{1}{2} \theta^\top a \theta + \int_{\mathbb{R}^d} \left(1 - e^{i\theta^\top x} + i\theta^\top x \chi(x)\right) \nu(dx),$$

where $\chi(x) = \mathbf{1}(|x| < 1)$ is a cut-off function, $b = (b^i) \in \mathbb{R}^d$, $a = (a^{ij}) \in \mathbb{R}^d \otimes \mathbb{R}^d$ is a positive semi-definite matrix and the measure ν on \mathbb{R}^d satisfies $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}^d} \min(1, |x|^2) \nu(dx) < \infty$. Triplet (a, ν, b) is called the **generating triplet** associated with the cut-off function χ of the Lévy process Y .

Remark A.2.2. The generating triplet of the Brownian motion is $(I_d, 0, 0)$.

Lévy processes are strong Markov processes [Sat99, Thm. 40.10] and the generating triplet also features in the infinitesimal generator.

Theorem A.2.3 (Theorem 31.5 in [Sat99]). *Let Y be a Lévy process in \mathbb{R}^d with the generating triplet (a, ν, b) . Then the generator \mathcal{L}_Y is given by*

$$\mathcal{L}_Y f(y) = \frac{1}{2} a^{ij} \partial_i \partial_j f(y) + b^i \partial_i f(y) + \int_{\mathbb{R}^d} (f(x+y) - f(y) - x^i \partial_i f(y) \chi(x)) \nu(dx)$$

on smooth functions f with compact support, where $\partial_i := \frac{\partial}{\partial x^i}$.

Every Lévy process can be expressed in terms of a Brownian motion and a Poisson random measure.

Theorem A.2.4 (Lévy-Itô decomposition, Theorem 2.4.16 in [App04]). *Let Y be a Lévy process in \mathbb{R}^d with the generating triplet (a, ν, b) . Then we can express it as*

$$Y_t^i = b^i t + \sigma_j^i B_t^j + \int_0^t \int_{0 < |x| < 1} x^i \tilde{N}(ds, dx) + \int_0^t \int_{|x| \geq 1} x^i N(ds, dx), \quad (\text{A.1})$$

where the matrix $(\sigma_k^i) \in \mathbb{R}^d \otimes \mathbb{R}^d$ satisfies $\sum_{k=1}^d \sigma_k^i \sigma_k^j = a^{ij}$, B is a standard d -dimensional Brownian motion, N is an independent Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d$ with mean

measure $\mathcal{L}eb \otimes \nu$, i.e. $\mathbb{E}[N([0,t],A)] = t\nu(A)$, and \tilde{N} is the compensated random measure defined by $\tilde{N}([0,t],A) = N([0,t],A) - t\nu(A)$.

Lévy processes are semimartingales and we can compute their quadratic variation. This is particularly useful for a Brownian motion, which can be uniquely characterized by its quadratic variation among all continuous local martingales.

Theorem A.2.5 (Lévy's characterization theorem, Theorem 40 in Chapter II in [Pro92]). *Let X be an continuous \mathbb{R}^d -valued local martingale. Then X is a standard Brownian motion if and only if*

$$\langle X^i, X^j \rangle_t = \delta_{ij}t$$

for every $i, j = 1, \dots, d$ and $t \in \mathbb{R}_+$.

Remark A.2.6. In the theorem we used the notation $\langle \cdot, \cdot \rangle$ which usually stands for a predictable quadratic variation. For general càdlàg semimartingales the quadratic variation and the predictable quadratic variation do not necessarily coincide, but they are the same if we consider continuous semimartingales. It is more common to use the notation $\langle \cdot, \cdot \rangle$ when dealing with continuous semimartingales, hence we used it in the theorem (and in the main text).

Bibliography

- [AE00] D. Applebaum and A. Estrade, *Isotropic Lévy processes on Riemannian manifolds*, The Annals of Probability **28** (2000), no. 1, 166–184.
- [AK93] D. Applebaum and H. Kunita, *Lévy flows on manifolds and Lévy processes on Lie groups*, J. Math. Kyoto Univ. **33** (1993), no. 4, 1103–1123.
- [AL] D. Applebaum and M. Liao, *Markov Processes with Jumps on Manifolds and Lie Groups*, Geometry and Invariance in Stochastic Dynamics, to appear.
- [AM13] T. M. Apostol and M. A. Mnatsakanian, *New balancing principles applied to circumsolids of revolution, and to n -dimensional spheres, cylindroids, and cylindrical wedges*, Amer. Math. Monthly **120** (2013), no. 4, 298–321. MR 3035124
- [App95] D. Applebaum, *A horizontal Lévy process on the bundle of orthonormal frames over a complete Riemannian manifold*, Séminaire de probabilités de Strasbourg **29** (1995), 166–180 (en). MR 1459458
- [App04] ———, *Lévy Processes and Stochastic Calculus*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2004.
- [ArcBC] Archimedes, *On the sphere and cylinder*, ca. 225 BC.
- [ASB20] D. Applebaum and R. Shewell Brockway, *L^2 properties of Lévy Generators on Compact Riemannian Manifolds*, Journal of Theoretical Probability (2020).
- [Bak96] D. Bakry, *Remarques sur les semigroupes de Jacobi*, Astérisque **236** (1996), 23–40.
- [Ber96] J. Bertoin, *Lévy Processes*, Cambridge Tracts in Mathematics 121, Cambridge University Press, 1996.

- [BGMN05] F. Barthe, O. Guédon, S. Mendelson, and A. Naor, *A probabilistic approach to the geometry of the l_p^n -ball*, Ann. Probab. **33** (2005), no. 2, 480–513. MR 2123199
- [Cai04] J.-M. Caillol, *Random walks on hyperspheres of arbitrary dimensions*, Journal of Physics A: Mathematical and General **37** (2004), no. 9, 3077.
- [CEE10] Tobias Carlsson, Tobias Ekholm, and Christer Elvingsson, *Algorithm for generating a Brownian motion on a sphere*, Journal of Physics A: Mathematical and Theoretical **43** (2010), no. 50, 505001.
- [Che00] A.S. Cherny, *On the strong and weak solutions of stochastic differential equations governing Bessel processes*, Stochastics and Stochastic Reports **70** (2000), no. 3-4, 213–219.
- [Che01] A. S. Cherny, *Convergence of some integrals associated with Bessel processes*, Theory of Probability & Its Applications **45** (2001), no. 2, 195–209.
- [ÇJ81] E. Çinlar and J. Jacod, *Representation of Semimartingale Markov Processes in Terms of Wiener Processes and Poisson Random Measures*, Seminar on Stochastic Processes, 1981 (E. Çinlar, K. L. Chung, and R. K. Gettoor, eds.), Birkhäuser Boston, 1981, pp. 159–242.
- [ÇJPS80] E. Çinlar, J. Jacod, Philip Protter, and M. Sharpe, *Semimartingales and Markov processes*, Probability Theory and Related Fields **54** (1980), 161–219.
- [Coh96a] S. Cohen, *Géométrie différentielle stochastique avec sauts 1*, Stochastics and Stochastic Reports **56** (1996), no. 3-4, 179–203.
- [Coh96b] ———, *Géométrie différentielle stochastique avec sauts 2: discrétisation et applications des eds avec sauts*, Stochastics and Stochastic Reports **56** (1996), no. 3-4, 205–225.
- [Dar82] R. W. R. Darling, *Martingales on manifolds and geometric ito calculus*, Ph.D. thesis, University of Warwick, 1982.
- [DeB04] D. DeBlassie, *Uniqueness for diffusions degenerating at the boundary of a smooth bounded set*, Ann. Probab. **32** (2004), no. 4, 3167–3190.
- [Dob07] E.E. Doberkat, *Stochastic Relations: Foundations for Markov Transition Systems*, Chapman & Hall/CRC Studies in Informatics Series, CRC Press, 2007.

- [Dyn65] E. B. Dynkin, *Markov processes: Volume 1*, 1 ed., Die Grundlehren der Mathematischen Wissenschaften 121/122, Springer-Verlag Berlin Heidelberg, 1965.
- [Elw82] K. D. Elworthy, *Stochastic differential equations on manifolds*, London Mathematical Society Lecture Note Series, Cambridge University Press, 1982.
- [Est97] A. Estrade, *A characterization of Markov solutions for stochastic differential equations with jumps*, Séminaire de probabilités de Strasbourg **31** (1997), 315–321 (en). MR 1478740
- [Fuj91] T. Fujiwara, *Stochastic differential equations of jump type on manifolds and Lévy flows*, J. Math. Kyoto Univ. **31** (1991), no. 1, 99–119.
- [GL83] R. C. Griffiths and W.-H. Li, *Simulating allele frequencies in a population and the genetic differentiation of populations under mutation pressure*, Theoretical Population Biology **23** (1983), no. 1, 19 – 33.
- [GMW19] N. Georgiou, A. Mijatović, and A. R. Wade, *Invariance principle for non-homogeneous random walks*, Electron. J. Probab. **24** (2019), 38 pp.
- [Gri84] R. C. Griffiths, *Asymptotic line-of-descent distributions*, Journal of Mathematical Biology **21** (1984), no. 1, 67–75.
- [GS72] R. K. Gettoor and M. J. Sharpe, *Conformal martingales*, Inventiones Mathematicae **16** (1972), no. 4, 271–308.
- [GS10] R. C. Griffiths and D. Spanò, *Diffusion processes and coalescent trees*, ArXiv e-prints (2010).
- [GSS12] A. Ghosh, J. Samuel, and S. Sinha, *A "Gaussian" for diffusion on the sphere*, EPL (Europhysics Letters) **98** (2012), no. 3, 30003.
- [HO56] J. Hano and H. Ozeki, *On the holonomy groups of linear connections*, Nagoya Math. J. **10** (1956), 97–100.
- [Hsu02] E.P. Hsu, *Stochastic Analysis on Manifolds*, Graduate Studies in Mathematics 38, vol. 38, American Mathematical Society, 2002.
- [Hun56] G. A. Hunt, *Semi-Groups of Measures on Lie Groups*, Transactions of the American Mathematical Society **81** (1956), no. 2, 264–293.

- [IM96] K. Ito and H.P. McKean, *Diffusion Processes and their Sample Paths*, Classics in Mathematics, Springer, 1996.
- [IW89] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, Kodansha scientific books, North-Holland, 1989.
- [Jos17] J. Jost, *Riemannian Geometry and Geometric Analysis*, 7 ed., Universitext, Springer International Publishing, 2017.
- [JP91] J. Jacod and P. Protter, *Une remarque sur les equations differentielles stochastiques a solutions markoviennes*, Séminaire de Probabilités XXV (Berlin, Heidelberg) (Jaques Azéma, Marc Yor, and Paul André Meyer, eds.), Springer Berlin Heidelberg, 1991, pp. 138–139.
- [JS03] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, 2 ed., Grundlehren der mathematischen Wissenschaften 288, Springer-Verlag Berlin Heidelberg, 2003.
- [JS17] P.A. Jenkins and D. Spanò, *Exact simulation of the Wright–Fisher diffusion*, Ann. Appl. Probab. **27** (2017), no. 3, 1478–1509.
- [Kal90] O. Kallenberg, *Random time change and an integral representation for marked stopping times*, Probability Theory and Related Fields **86** (1990), no. 2, 167–202.
- [KDPN00] M. M. G. Krishna, Ranjan Das, N. Periasamy, and Rajaram Nityananda, *Translational diffusion of fluorescent probes on a sphere: Monte Carlo simulations, theory, and fluorescence anisotropy experiment*, The Journal of Chemical Physics **112** (2000), no. 19, 8502–8514.
- [KN63] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Wiley Classics Library, vol. 1, Wiley-Interscience, 1963.
- [KPP95] T. G. Kurtz, É. Pardoux, and P. Protter, *Stratonovich stochastic differential equations driven by general semimartingales*, Annales de l’I.H.P. Probabilités et Statistiques **31** (1995), no. 2, 351–377 (en). MR 1324812
- [KT81] S. Karlin and H. M. Taylor, *A Second Course in Stochastic Processes*, Academic Press, 1981.

- [Kun19] H. Kunita, *Stochastic flows and jump-diffusions*, 1 ed., Probability Theory and Stochastic Modelling, Springer, 2019.
- [Kyp14] A.E. Kyprianou, *Fluctuations of Lévy Processes with Applications: Introductory Lectures*, Universitext, Springer Berlin Heidelberg, 2014.
- [Lee97] J.M. Lee, *Riemannian Manifolds: An Introduction to Curvature*, 1st ed., Graduate Texts in Mathematics, Springer, 1997.
- [Lee12] ———, *Introduction to Smooth Manifolds*, 2nd ed., Graduate Texts in Mathematics 218, Springer-Verlag New York, 2012.
- [Lia04] Ming Liao, *Lévy processes in Lie groups*, Cambridge Tracts in Mathematics, Cambridge University Press, 2004.
- [LTT08] Guanglai Li, Lick-Kong Tam, and Jay X. Tang, *Amplified effect of Brownian motion in bacterial near-surface swimming*, Proceedings of the National Academy of Sciences **105** (2008), no. 47, 18355–18359.
- [Mar81] S. I. Marcus, *Modeling and approximation of stochastic differential equations driven by semimartingales*, Stochastics **4** (1981), no. 3, 223–245.
- [Mey81] P. A. Meyer, *Geometrie stochastique sans larmes*, Séminaire de Probabilités XV 1979/80 (Berlin, Heidelberg) (Jacques Azéma and Marc Yor, eds.), Springer Berlin Heidelberg, 1981, pp. 44–102.
- [MM20] A. Mijatović and V. Mramor, *Lévy processes on smooth manifolds with a connection*, ArXiv e-prints, 2012.11633 (2020).
- [MMU18] A. Mijatović, V. Mramor, and G. Uribe Bravo, *Projections of spherical Brownian motion*, Electron. Commun. Probab. **23** (2018), 12 pp.
- [MMU20] A. Mijatović, V. Mramor, and G. Uribe Bravo, *A note on the exact simulation of spherical Brownian motion*, Statistics & Probability Letters **165** (2020), 108836.
- [MMU21] A. Mijatović, V. Mramor, and G. Uribe Bravo, *An algorithm for simulating Brownian increments on a sphere*, Journal of Physics A: Mathematical and Theoretical (2021).
- [Moh04] A. Mohari, *Ergodicity of Lévy flows*, Stochastic Processes and their Applications **112** (2004), no. 2, 245–259.

- [NEE03] J. Nissfolk, T. Ekholm, and C. Elvingson, *Brownian dynamics simulations on a hypersphere in 4-space*, The Journal of Chemical Physics **119** (2003), no. 13, 6423–6432.
- [O’N83] B. O’Neill, *Semi-Riemannian geometry: with applications to relativity*, Pure and applied mathematics 103, Academic Press, 1983.
- [Pau90] J.E Pauwels, *Riemannian submersions of Brownian motions*, Stochastics and Stochastic Reports **29** (1990), no. 4, 425–436.
- [PE92] M. Pontier and A. Estrade, *Relèvement horizontal d’une semimartingale càdlàg*, Séminaire de probabilités de Strasbourg **26** (1992), 127–145 (fr). MR 1231989
- [PR88] E.J. Pauwels and L.C.G. Rogers, *Skew-Product Decompositions of Brownian Motions*, Contemporary Math. **73** (1988), 237–262.
- [PR12] J. Pitman and N. Ross, *Archimedes, Gauss, and Stein*, Notices Amer. Math. Soc. **59** (2012), no. 10, 1416–1421. MR 3025901
- [Pro92] P. Protter, *Stochastic Integration and Differential Equation*, second ed., Springer-Verlag, Berlin, Heidelberg, 1992.
- [Rog81] S. J. Rogerson, *Stochastic dynamical systems and processes with discontinuous sample paths*, Ph.D. thesis, University of Warwick, 1981.
- [RS17] G. Rudolph and M. Schmidt, *Differential Geometry and Mathematical Physics: Part II. Fibre Bundles, Topology and Gauge Fields*, 1st ed., Theoretical and Mathematical Physics, Springer, 2017.
- [RY99] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3rd ed., Springer, 1999.
- [Sat99] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 1999.
- [Sch80] L. Schwartz, *Semi-Martingales sur des Varietes, et Martingales Conformes sur des Varietes Analytiques Complexes*, 1980.
- [Sch82] ———, *Geometrie differentielle du 2ème ordre, semi-martingales et equations differentielles stochastiques sur une variete differentielle*, Séminaire de

- Probabilités XVI, 1980/81 Supplément: Géométrie Différentielle Stochastique (Berlin, Heidelberg) (Jacques Azéma and Marc Yor, eds.), Springer Berlin Heidelberg, 1982, pp. 1–148.
- [Sch12] Alexander Schnurr, *On the semimartingale nature of Feller processes with killing*, Stochastic Processes and their Applications **122** (2012), no. 7, 2758–2780.
- [Shi82] I. Shigekawa, *On stochastic horizontal lifts*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **59** (1982), 211–221.
- [Swa02] J. M. Swart, *Pathwise uniqueness for a SDE with non-Lipschitz coefficients*, Stochastic Processes and their Applications **98** (2002), no. 1, 131–149.
- [Wik20a] Wikipedia contributors, *Parallel transport sphere.svg* — *Wikipedia, the free encyclopedia*, 2020, https://en.wikipedia.org/wiki/File:Parallel_transport_sphere.svg, [Online; accessed 15. 8. 2020].
- [Wik20b] ———, *Parallel transport.svg* — *Wikipedia, the free encyclopedia*, 2020, https://en.wikipedia.org/wiki/File:Parallel_Transport.svg, [Online; accessed 15. 8. 2020].
- [WY98] J. Warren and M. Yor, *The Brownian burglar: conditioning Brownian motion by its local time process*, Séminaire de Probabilités XXXII (Berlin, Heidelberg) (Jacques Azéma, Marc Yor, Michel Émery, and Michel Ledoux, eds.), Springer Berlin Heidelberg, 1998, pp. 328–342.
- [YW71] T. Yamada and S. Watanabe, *On the uniqueness of solutions of stochastic differential equations*, J. Math. Kyoto Univ. **11** (1971), no. 1, 155–167.