

Infinite dimensional Piecewise Deterministic Markov Processes

Dobson, Paul; Bierkens, Joris

DOI

[10.1016/j.spa.2023.08.010](https://doi.org/10.1016/j.spa.2023.08.010)

Publication date

2023

Document Version

Final published version

Published in

Stochastic Processes and their Applications

Citation (APA)

Dobson, P., & Bierkens, J. (2023). Infinite dimensional Piecewise Deterministic Markov Processes. *Stochastic Processes and their Applications*, 165, 337-396. <https://doi.org/10.1016/j.spa.2023.08.010>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



Infinite dimensional Piecewise Deterministic Markov Processes

Paul Dobson^a, Joris Bierkens^{b,*}

^a School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, Scotland, UK

^b Delft Institute of Applied Mathematics, TU Delft, Mekelweg 4, 2628 CD Delft, The Netherlands

Received 23 May 2022; received in revised form 11 April 2023; accepted 30 August 2023

Available online 9 September 2023

Abstract

In this paper we aim to construct infinite dimensional versions of well established Piecewise Deterministic Monte Carlo methods, such as the Bouncy Particle Sampler, the Zig-Zag Sampler and the Boomerang Sampler. In order to do so we provide an abstract infinite dimensional framework for Piecewise Deterministic Markov Processes (PDMPs) with unbounded event intensities. We further develop exponential convergence to equilibrium of the infinite dimensional Boomerang Sampler, using hypocoercivity techniques. Furthermore we establish how the infinite dimensional Boomerang Sampler admits a finite dimensional approximation, rendering it suitable for computer simulation.

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

MSC: primary 60G53; secondary 65C05; 46N30

Keywords: Piecewise Deterministic Markov Processes; Infinite Dimensional Stochastic Process; Hypocoercivity; Uniform in time approximation

1. Introduction

A *piecewise deterministic Markov process (PDMP)* in a topological space \mathcal{Z} is a Markov process with deterministic trajectories, with random jumps at random times. The deterministic trajectories are described by a deterministic semigroup flow $\varphi_t : \mathcal{Z} \rightarrow \mathcal{Z}$, $t \geq 0$. The random times are distributed according to a space-dependent event rate $\lambda : \mathcal{Z} \rightarrow [0, \infty)$. At these random times the process jumps according to a Markov transition operator $Q : \mathcal{Z} \times \mathcal{B}(\mathcal{Z}) \rightarrow [0, 1]$. Together the flow $(\varphi_t)_{t \geq 0}$, the jump intensity λ and the transition operator fully describe the dynamics of the piecewise deterministic Markov process.

* Corresponding author.

E-mail addresses: p.dobson_1@hw.ac.uk (P. Dobson), joris.bierkens@tudelft.nl (J. Bierkens).

PDMPs have been introduced in the probability literature in the work by Davis [14,15], motivated as providing a versatile probabilistic model in operations research. Later PDMPs have found many other applications in, e.g., mathematical biology [22], finance and statistical survival analysis [25]. Recent years have seen a strongly emerging interest in PDMPs from the field of computational statistics [8,10], where PDMPs are considered as viable alternatives for classical Markov Chain Monte Carlo methods such as the Metropolis–Hastings algorithm and the Gibbs sampler.

Traditionally PDMPs have been considered on (subsets of) a finite dimensional space \mathcal{Z} . However PDMPs can naturally be defined on, e.g., nonlinear manifolds or infinite dimensional spaces. In this work we carry out the latter task of defining and analysing PDMPs on infinite dimensional Banach spaces. To our knowledge the only work concerning PDMPs on infinite dimensional spaces so far [27,28], restricts the jump intensities to be globally bounded. One of the key aspects in this work is therefore the extension to unbounded intensities, which is highly relevant from a practical perspective, especially for applications in computational statistics. For much of this work we have in mind an infinite dimensional extension of piecewise deterministic Monte Carlo methods; in particular extensions of the Zig-Zag process [8], the Bouncy Particle Sampler [10] and the Boomerang Sampler [9]. As it turns out these processes are not all easily extendable to infinite dimensions.

1.1. Structure of this work

In Section 2 we provide the general construction of a PDMP assuming values in a Banach space. Also, we construct the accompanying extended generator and establish the Feller property under the stated assumptions.

Next in Section 3 we determine a general condition for an infinite dimensional PDMP to have an invariant measure with a Gibbs density relative to a reference measure.

In Section 4 we discuss several examples of infinite dimensional extensions of (traditionally) finite dimensional PDMPs: the Zig-Zag Sampler [8], Bouncy Particle Sampler [10], and the Boomerang Sampler [9]. As it turns out the Zig-Zag Sampler can be put into an infinite dimensional framework, but the Bouncy Particle Sampler does not satisfy the conditions of the general theory: a natural bound to use for contour reflections is the Cameron–Martin norm however this is only finite on a space of measure zero. The Boomerang Sampler is particularly well suited for extension to infinite dimensions, as it allows a Gaussian reference measure which remains well-defined in an infinite dimensional space. In the remainder of this work, we will therefore focus on the Boomerang Sampler. In Section 5 we establish that cylindrical functions (see (3) for a definition) are a core for the generator of the Boomerang Sampler. In Section 6 we provide conditions under which the Boomerang Sampler converges exponentially fast in $L^2(\mu)$, with μ the stationary measure, based on the hypocoercivity approach initiated by [16], while being careful about its extension to an infinite dimensional setting.

Finally in Section 7 we consider how infinite dimensional PDMPs may be approximated in finite dimensions, allowing for direct computer simulation. This may be motivated from the viewpoint of e.g. [29], where the approach in Bayesian inverse problems is to formulate these in an infinite dimensional setting, design a suitable infinite dimensional sampling algorithm, and only at the time of implementation consider a suitable finite dimensional approximation. In Section 8 we have collected the remaining proofs.

1.2. Notation

For a given topological space, measurability is always considered relative to the Borel σ -algebra, unless stated otherwise.

Given two normed vector spaces H_1, H_2 we define the following function spaces:

- The set of all bounded and measurable (respectively continuous) functions $f : H_1 \rightarrow H_2$ we denote $B_b(H_1; H_2)$ (resp. $C_b(H_1; H_2)$) and we endow this space with the supremum norm, if $H_2 = \mathbb{R}$ then we write $B_b(H_1)$ (resp. $C_b(H_1)$);
- For $k \geq 1$ or $k = \infty$ we denote by $C_b^k(H_1)$ the space of k times differentiable bounded functions with bounded derivatives up to order k (here and throughout when we say differentiable we mean in the sense of the Fréchet derivative). Similarly we denote by $C_c^k(H_1)$ to be the set of k times continuously differentiable functions with compact support.
- Given a measure μ we denote the space of square integrable functions $f : H_1 \rightarrow H_2$ by $L_\mu^2(H_1; H_2)$, if $H_2 = \mathbb{R}$ write $L_\mu^2(H_1)$ and will often abbreviate this to L_μ^2 ;

As explained in the introduction, we will mostly be working in appropriate Hilbert spaces. However, the results of Section 2 are stated in the setting of more general Banach spaces. Throughout the paper we denote by $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$ a (possibly) infinite dimensional Banach space and let $(\mathcal{H}, \langle \cdot, \cdot \rangle, \|\cdot\|)$ be an infinite dimensional separable Hilbert space. Consider the probability measure π on \mathcal{H} , defined as follows:

$$\frac{d\pi}{d\pi_0} \propto \exp(-\Phi), \quad \pi_0 := \mathcal{N}(0, \Sigma). \tag{1}$$

That is, π is absolutely continuous with respect to a Gaussian measure π_0 with mean zero and covariance operator Σ . Here Φ is a real valued functional defined on \mathcal{H} . Measures of the form (1) naturally arise in Bayesian nonparametric statistics and in the study of conditioned diffusions [24,29]. The covariance operator Σ is a positive definite, self-adjoint, trace class operator on \mathcal{H} , with eigenbasis $\{\gamma_j^2, e_j\}$:

$$\Sigma e_j = \gamma_j^2 e_j, \quad \forall j \in \mathbb{N}, \tag{2}$$

and we assume that the set $\{e_j\}_{j \in \mathbb{N}}$ is an orthonormal basis for \mathcal{H} , so that every $x \in \mathcal{H}$ can be expressed as $x = \sum_{j \geq 1} x_j e_j$, where $x_j := \langle x, e_j \rangle$ and the norm of $x \in \mathcal{H}$ is

$$\|x\| = \left(\sum_{j=1}^{\infty} |x_j|^2 \right)^{1/2}.$$

For a function $\phi : \mathcal{H} \rightarrow \mathbb{R}$ we will write $\partial_{x_i} \phi(x) := \langle \nabla \phi(x), e_i \rangle = \lim_{h \rightarrow 0} \frac{\phi(x + h e_i) - \phi(x)}{h}$, whenever the limit exists.

If $A, B \in \mathbb{R}$ we write $A \lesssim B$ if there exists a constant $K > 0$ such that $A \leq KB$. Finally, for any $a \in \mathbb{R}$, $a^+ = \max(a, 0)$ and $a^- = \max(-a, 0)$ denote the positive and negative part of a , respectively.

For a Hilbert space \mathcal{H} let $\mathcal{FC}_b^\infty(H)$ (respectively \mathcal{FC}_c^∞) denote the set of all functions F of the form

$$F(z) = f(\langle h_1, z \rangle_H, \dots, \langle h_n, z \rangle_H) \tag{3}$$

for some $n \in \mathbb{N}$, $h_1, \dots, h_n \in H$ and $f \in C_b^\infty(\mathbb{R}^n)$ (respectively $f \in C_c^\infty(\mathbb{R}^n)$).

2. Construction of infinite dimensional PDMPs

Let us first describe how we can construct a PDMP on an infinite dimensional space. We will follow the construction given by [15]. Let \mathcal{Z} be a Banach space, equipped with its Borel σ -algebra. We can define a PDMP by its characteristics $(\mathfrak{X}, \lambda, Q)$, where \mathfrak{X} denotes the vector field of the deterministic flow, λ denotes the jump intensity and Q denotes the jump kernel. More specifically, $\mathfrak{X} : \mathcal{Z} \rightarrow \mathcal{Z}$ is a vector field on \mathcal{Z} which generates a flow map $\{\varphi_t\}_{t \geq 0}$, i.e., $\varphi_t : \mathcal{Z} \rightarrow \mathcal{Z}$ for each $t \geq 0$, and for each $t \geq 0$,

$$\frac{d}{dt} \varphi_t(z) = \mathfrak{X} \varphi_t(z), \quad \varphi_0(z) = z.$$

Note that we are assuming that φ_t is well-defined for all $t \geq 0$, that is, there is no finite time explosion of the flow $\varphi_t(z)$ for any $z \in \mathcal{Z}$. The function λ maps from \mathcal{Z} into $[0, \infty)$ and has the property that for each $z \in \mathcal{Z}$ there exists some $\varepsilon(z) > 0$ such that $s \mapsto \lambda(\varphi_s(z))$ is integrable on $[0, \varepsilon(z))$. Q is a transition kernel such that

- (1) For every Borel measurable set $A \subseteq \mathcal{Z}$ the map $z \mapsto Q(z, A)$ is measurable.
- (2) There are no phantom jumps, i.e. $Q(z, \{z\}) = 0$ for all $z \in H$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote the Hilbert cube, that is the canonical space for a sequence of i.i.d. uniformly distributed random variables $\{U_n\}_{n=1}^\infty$ taking values on $[0, 1]$. Fix an initial condition $z \in \mathcal{Z}$, and set

$$F(t, z) = \exp \left(- \int_0^t \lambda(\varphi_s(z)) ds \right).$$

That is the survivor function of the first jump time T_1 . Set

$$\psi_1(u, z) = \inf\{t : F(t, z) \leq u\}$$

where we use the convention that $\inf \emptyset = +\infty$. Define $T_1(\omega) = \psi_1(U_1(\omega), z)$. Now by [27, Lemma 2.1.1] there exists a Borel measurable function $\psi_2^z : [0, 1] \rightarrow \mathcal{Z}$ such that the pushforward of the Lebesgue measure with respect to ψ on H equals $Q(z, \cdot)$. We define the sample path up to the first jump time by

$$Z_t(\omega) = \begin{cases} \varphi_t(z), & 0 \leq t < T_1(\omega) \\ \psi_2^{\varphi_{T_1(\omega)}(z)}(U_2(\omega)), & t = T_1(\omega). \end{cases}$$

We now iterate this procedure to construct $Z_t(\omega)$ for all $t \geq 0$ provided $\sup_n T_n = +\infty$, that is the process is *non-explosive*. If λ is bounded then the process is non-explosive. In order to consider unbounded λ we will split the transition kernel Q into two parts Q_r and Q_b which will later correspond to a refreshment kernel and a reflection kernel respectively.

We will associate with this PDMP (at least formally) a generator

$$\begin{aligned} \mathcal{L}f(z) &= L_{\mathfrak{X}}f(z) + \lambda_r(z) \int (f(y) - f(z))Q_r(z, dy) + \lambda_b(z) \int (f(y) - f(z))Q_b(z, dy), \\ &z \in \mathcal{Z}. \end{aligned} \tag{4}$$

Here $L_{\mathfrak{X}}$ is a linear differential operator on a space of functions from \mathcal{Z} to \mathbb{R} defined by

$$L_{\mathfrak{X}}f(z) = \left. \frac{d}{dt} f(\varphi_t(z)) \right|_{t=0} = df(z; \mathfrak{X}(z)),$$

where $df(z, \cdot)$ denotes the Fréchet derivative of f at z , whenever this is well-defined.

Assumption 2.1 (Wellposedness of Abstract PDMP). With the notation introduced so far, we make the following assumptions.

- (i) The intensities $\lambda_r : \mathcal{Z} \rightarrow [0, \infty)$ and $\lambda_b : \mathcal{Z} \rightarrow [0, \infty)$ are measurable.
- (ii) for any fixed $z \in \mathcal{Z}$, $Q_b(z, \cdot), Q_r(z, \cdot)$ are Markov transition kernels on \mathcal{Z} , such that $z \mapsto Q_b f(z)$ and $z \mapsto Q_r f(z)$ are measurable for every bounded measurable function f on \mathcal{Z} ;
- (iii) the reflection Markov kernel Q_b is non-expansive with respect to the norm $\|\cdot\|_{\mathcal{Z}}$ i.e.

$$Q_b(z, \{y \in \mathcal{Z} : \|y\|_{\mathcal{Z}} \leq \|z\|_{\mathcal{Z}}\}) = 1 \quad \text{for all } z \in \mathcal{Z};$$

- (iv) the refreshment Markov kernel Q_r is locally uniformly bounded in probability, i.e., for all $\varepsilon > 0$ there exists a $R > 0$ such that

$$Q_r(z, B_R) > 1 - \varepsilon \quad \text{for all } z \in B_R,$$

where B_R denotes the open ball with radius R in \mathcal{Z} with respect to the $\|\cdot\|_{\mathcal{Z}}$ -norm,

- (v) the total reflection intensity $\lambda_b(z)$ is bounded on bounded sets in \mathcal{Z} , i.e. for all $R > 0$

$$\sup_{z \in B_R} \lambda_b(z) < \infty;$$

- (vi) the total refreshment intensity $\lambda_r(z)$ is globally bounded: there is a constant $M > 0$ such that $\lambda_r(z) \leq M$ for all $z \in \mathcal{Z}$;
- (vii) there is a continuous function $t \mapsto c_t \in (0, \infty)$ such that the flow $\varphi_t(z)$ grows at a rate c_t with respect to the $\|\cdot\|_{\mathcal{Z}}$ -norm, i.e. for $z \in \mathcal{Z}$,

$$\|\varphi_t(z)\|_{\mathcal{Z}} \leq c_t(1 + \|z\|_{\mathcal{Z}}).$$

- (viii) the functions λ_r, λ_b are continuous and for any continuous function f we have that $Q_b f, Q_r f$ are both continuous.

Remark 2.2 (On Assumption 2.1). These conditions are intended to be as general as possible. From a practical viewpoint it seems that (iii) is the most difficult condition to satisfy. In particular, we will see in Section 4 that this condition is not satisfied by the Bouncy Particle Sampler. \triangle

The purpose of this section is to prove the following statement.

Theorem 2.3. Suppose Assumption 2.1 is satisfied. Then there exists a cadlag Markov process $(Z_t)_{t \geq 0}$ in \mathcal{Z} with characteristics $(\mathfrak{X}, \lambda, Q)$.

Proof. At the beginning of this section we gave a construction of the PDMP; however this construction is only valid for $t \leq \sup_n T_n$, therefore it remains to show that $\sup_n T_n = \infty$. This is a consequence of Proposition 2.4. The proof that this construction of a PDMP is a strong Markov process is analogous to [15, Theorem 25.5]. \square

Proposition 2.4. Suppose Assumption 2.1 is satisfied. Write $T_{\text{explode}} = \sup_{i \geq 1} T_i$. Then

- (i) For all $\varepsilon > 0, t \geq 0$ and initial condition $z_0 \in \mathcal{Z}$ there exists an $R > 0$ such that $\mathbb{P}_{z_0}(T_{\text{explode}} > t \text{ and } Z_s \in B_R \text{ for all } s \in [0, t]) > 1 - \varepsilon$.

- (ii) $T_{\text{explode}} = \infty$ almost surely.
- (iii) For all $t \geq 0, r \geq 0$

$$\lim_{n \rightarrow \infty} \sup_{z_0 \in B_r} \mathbb{P}_{z_0}(T_n > t) = 1.$$

Statement (ii) of Proposition 2.4 is equivalent to saying that only finitely many events can occur in any finite time interval.

Proof. Suppose $t \geq 0, n \in \mathbb{N}$ and $\varepsilon > 0$, and let $r > 0$ sufficiently large so that $z_0 \in B_r$. Let M be an upper bound for the total refreshment intensity, using Assumption 2.1(vi). Let N denote a Poisson random variable with parameter Mt and let $k_r \in \mathbb{N}$ such that $\mathbb{P}(N \geq k_r) < \varepsilon/3$. By Assumption 2.1(vi) with probability at least $1 - \varepsilon/3$ there will be fewer than k_r jumps due to the refreshment kernels. Using Assumption 2.1(iv), let R' be sufficiently large so that $r < R'$ and $Q_r(z, B_{R'}) > (1 - \varepsilon/3)^{1/k_r}$ for all $z \in B_{R'}$, so that with probability at least $1 - \varepsilon/3$ all the observed refreshments map into $B_{R'}$. Conditional on the event that all refreshments map into $B_{R'}$, since all reflections are norm preserving (using Assumption 2.1(iii)) it follows from Assumption 2.1(vii) that $\|Z_t\| \leq c_t(1 + R')$. Since λ_b is bounded on bounded sets (Assumption 2.1(v)), it follows that there is a constant k_b such that with probability at least $1 - \varepsilon/3$ there are at most k_b reflection events. We conclude that, for all $\varepsilon > 0$ and $t \geq 0$ we have

$$\mathbb{P}_{z_0}(T_{k_b+k_r} > t \text{ and } Z_s \in B_{c_t(1+R')} \text{ for all } s \in [0, t]) > 1 - \varepsilon,$$

establishing the first statement of the lemma. The second statement follows from the first since for all ε there exists $R > 0$ such that

$$\mathbb{P}_{z_0}(T_{\text{explode}} > t) \geq \mathbb{P}_{z_0}(T_{\text{explode}} > t \text{ and } Z_t \in B_R) > 1 - \varepsilon$$

where t is arbitrary and ε can be chosen arbitrarily small. Moreover, as $R = c_t(1 + R')$ is independent of the choice of z_0 in B_r we have that

$$\sup_{z_0 \in B_r} \mathbb{P}_{z_0}(T_{k_b+k_r} > t) > 1 - \varepsilon.$$

As k_r and k_b are deterministic and depend only on ε, r, R' for each $\varepsilon > 0$ we may take n sufficiently large that $n \geq k_r + k_b$ which gives

$$\lim_{n \rightarrow \infty} \sup_{z_0 \in B_r} \mathbb{P}_{z_0}(T_n > t) > 1 - \varepsilon.$$

Now since ε was arbitrary we get the third statement of the lemma. \square

Define the semigroup $\{\mathcal{P}_t\}_{t \geq 0}$ on the space $B_b(\mathcal{Z})$, the set of all bounded and measurable functions $f : \mathcal{Z} \rightarrow \mathbb{R}$ endowed with the supremum norm, by setting

$$\mathcal{P}_t f(z) = \mathbb{E}_z[f(Z_t^z)]. \tag{5}$$

Lemma 2.5. *If Assumption 2.1 holds, then $\{\mathcal{P}_t\}_{t \geq 0}$ defined by (5) is a Feller semigroup, i.e. $\mathcal{P}_t(C_b(\mathcal{Z})) \subseteq C_b(\mathcal{Z})$.*

Proof of Lemma 2.5. Note that \mathcal{P}_t is a contraction with respect to the supremum norm so it is sufficient to show that $\mathcal{P}_t f$ is continuous for any $f \in C_b(\mathcal{Z})$. We will follow the strategy of [15, Theorem 27.6]. Fix $f \in C_b(\mathcal{Z})$ and define for any $g \in C_b([0, \infty) \times \mathcal{Z})$

$$Gg(t, z) = \mathbb{E}_z[f(Z_t)\mathbb{1}_{t < T_1} + g(t - T_1, Z_{T_1})\mathbb{1}_{t \geq T_1}].$$

Here $\mathbb{1}_A$ denotes the characteristic function on the set A . With this notation we can also express $Gg(t, z)$ as

$$Gg(t, z) = f(\varphi_t(z))e^{-\int_0^t \lambda(\varphi_s(z))ds} + \int_0^t Qg(t - s, \cdot)(\varphi_s(z))\lambda(\varphi_s(z))e^{-\int_0^s \lambda(\varphi_r(z))dr} ds.$$

Notice that $Gg(t, x, v)$ is continuous in (x, v) provided f, g are both continuous. Following the proof of [15, Lemma 27.3] we have

$$|G^n g(t, z) - \mathcal{P}_t f(z)| \leq 2c\mathbb{P}_z(T_n \leq t)$$

where $c = \max(\|f\|_\infty, \|g\|_\infty)$. By Proposition 2.4 we have that $G^n g(t, z)$ converges to $\mathcal{P}_t f(z)$ uniformly on bounded sets; therefore $\mathcal{P}_t f(z)$ is continuous. \square

2.1. Extended generator

For PDMPs the domain of the extended generator can be explicitly characterised as shown in [15,27]

Definition 2.6. Let D_{ext} denote the set of measurable functions $f : H \rightarrow \mathbb{R}$ with the property that there exists a measurable function $h : \mathcal{Z} \rightarrow \mathbb{R}$ such that $t \mapsto h(Z_t)$ is integrable almost surely and the process

$$C_t^f = f(Z_t) - f(Z_0) - \int_0^t h(Z_s)ds \tag{6}$$

is a \mathbb{P}_z -local martingale. Then we set $\mathcal{L}f = h$ and call $(\mathcal{L}, D_{\text{ext}})$ the extended generator.

Theorem 2.7. Let Z_t denote the PDMP with characteristics $(\mathfrak{X}, \lambda, Q)$ and assume that Assumption 2.1 holds. Then the domain of the extended generator D_{ext} is given by all measurable functions f such that

- (i) For each $z \in \mathcal{Z}$ the function $t \mapsto f(\varphi_t(z))$ is differentiable for almost every t ;
- (ii) $(z, t, \omega) \mapsto f(z) - f(Z_{t-}(\omega))$ is a valid integrand for the compensating measure \tilde{p} . Here $\tilde{p}(t, A) = \int_0^t Q(Z_s, A)\lambda(Z_s)ds$.

Then, for $f \in D_{\text{ext}}$ the extended generator is given by

$$\mathcal{L}f(z) = L_{\mathfrak{X}}f(z) + \lambda(z) \int_{\mathcal{Z}} (f(y) - f(z))Q(z, dy). \tag{7}$$

This follows from the proof of [27, Theorem 2.2.1]. We note that [27] assumes that λ is bounded, however for characterising the extended generator they only use that λ is bounded to ensure the process is well-defined. Since we have established the PDMP is well-defined the proof of [27, Theorem 2.2.1] holds. From [15, Remark 26.16] if f is bounded and measurable then condition (ii) is satisfied.

Recall the semigroup $\{\mathcal{P}_t\}_{t \geq 0}$ defined by (5) defined on the space $B_b(\mathcal{Z})$. Note that $\{\mathcal{P}_t\}_{t \geq 0}$ is a contraction semigroup however it need not be a strongly continuous semigroup, therefore we shall consider a smaller space on which $\{\mathcal{P}_t\}_{t \geq 0}$ is strongly continuous. By Lemma 2.5 if Assumption 2.1 holds then $\{\mathcal{P}_t\}_{t \geq 0}$ is a Feller semigroup, i.e. for every $t \geq 0$ we have that $\mathcal{P}_t(C_b(\mathcal{Z})) \subseteq C_b(\mathcal{Z})$. Set $C_0(\mathcal{Z}) = \{f \in C_b(\mathcal{Z}) : \lim_{t \rightarrow 0} \|\mathcal{P}_t f - f\|_\infty = 0\}$. Note this space is closed by the same arguments used in [15, pg 28–29] and as $\{\mathcal{P}_t\}_{t \geq 0}$ is a Feller semigroup $\mathcal{P}_t(C_0(\mathcal{Z})) \subseteq C_0(\mathcal{Z})$ for every $t \geq 0$. On the space $C_0(\mathcal{Z})$ the semigroup $\{\mathcal{P}_t\}_{t \geq 0}$ is strongly

continuous. Let (\mathcal{L}, D_{C_0}) be the generator of $\{\mathcal{P}_t\}_{t \geq 0}$ on the space $C_0(\mathcal{Z})$. We can see from the definition of the extended generator that \mathcal{L} is the restriction¹ of the extended generator L to the set D_{C_0} .

3. Invariant measure of PDMP

In this section we give conditions to ensure a PDMP has an invariant measure μ . The measures we consider will be absolutely continuous with respect to a probability measure μ_0 and we will assume there exists a lower semicontinuous function $\Phi : \mathcal{Z} \rightarrow \mathbb{R}$ which is bounded from below such that

$$\frac{d\mu}{d\mu_0}(z) = \frac{e^{-\Phi(z)}}{\int e^{-\Phi(z)}\mu_0(dz)}. \tag{8}$$

Example 3.1 (*Inverse Problem for a Diffusion Coefficient*). This example is based on [29, Section 3.3]. Consider the inverse problem of determining the diffusion coefficient from observations of the solutions of the PDE

$$\frac{d}{dt} \left(\kappa(t) \frac{dv}{dt} \right) = 0, \tag{9}$$

$$v(0) = 0, v(1) = 1. \tag{10}$$

We make observations $\{v(t_k)\}_{k=1}^q$ subject to Gaussian measurement error. Write observations as

$$y_j = v(t_j) + \eta_j, j = 1, \dots, q$$

where η_j are i.i.d. and distributed according to $N(0, 1)$. To ensure that κ is strictly positive we set $u(x) = \ln(\kappa(x))$ and view $u \in L^2((0, 1))$. Now define $J_x : L^\infty((0, 1)) \rightarrow \mathbb{R}$ by

$$J_t(w) = \int_0^t \exp(-w(z))dz.$$

Then the solution of (9) may be written as

$$v(x) = \frac{J_t(u)}{J_1(u)}.$$

Set

$$\Phi(u) = \frac{1}{2} \sum_{j=1}^q \left| y_j - \frac{J_{t_j}(u)}{J_1(u)} \right|^2$$

Now we place a Gaussian prior π_0 on the space $L^2((0, 1))$ with mean zero and covariance operator $\beta(-\Delta)^{-\alpha}$. Here we set the domain of Δ to be

$$\left\{ \Delta u = u'', D(\Delta) = \left\{ u \in H^2((0, 1)) : u(0) = u(1), \int_0^1 u(s)ds = 0 \right\} \right\}.$$

¹ Let $f \in D_{C_0}$ and set C_t^f to be defined as in (6). Then we have

$$\mathbb{E}[C_t^f | \mathcal{F}_s] = \mathcal{P}_{t-s} f(Z_s) - f(Z_0) - \int_s^t \mathcal{P}_{r-s} \mathcal{L} f(Z_s) dr - \int_0^s \mathcal{L} f(Z_r) dr.$$

As $f \in D_{C_0}$ we can write $\int_s^t \mathcal{P}_{r-s} \mathcal{L} f(Z_s) dr = \mathcal{P}_{t-s} f(Z_s) - f(Z_s)$ which gives

$$\mathbb{E}[C_t^f | \mathcal{F}_s] = f(Z_s) - f(Z_0) - \int_0^s \mathcal{L} f(Z_r) dr = C_s^f.$$

Therefore $f \in D_{\text{ext}}$.

From [29, Lemma 6.25] we have that $\pi_0(C([0, 1])) = 1$ provided $\alpha > 1/2$. Then by [29, Theorem 3.4] the posterior distribution $\pi^y(du) = \mathbb{P}(du|y)$ given the data y is absolutely continuous with respect to π_0 and

$$\frac{d\pi^y}{d\pi_0}(u) \propto \exp(-\Phi(u)). \quad \Delta$$

Theorem 3.1. Assume that Assumption 2.1 holds. Suppose \mathcal{Z} is equipped with a probability measure μ_0 . Define a transition kernel Q^* by

$$\int_{\mathcal{Z}} \int_{\mathcal{Z}} g(y, z) Q^*(y, dz) \mu_0(dy) = \int_{\mathcal{Z}} \int_{\mathcal{Z}} g(y, z) Q(z, dy) \mu_0(dz) \tag{11}$$

for all $g \in C_b(\mathcal{Z} \times \mathcal{Z})$. Let μ denote a probability measure satisfying $\mu \propto e^{-\Phi} \mu_0$. Assume that

$$e^{\Phi(z)} L_{\mathfrak{X}}^*(e^{-\Phi(\cdot)})(z) + \int_{\mathcal{Z}} \lambda(y) e^{\Phi(z)-\Phi(y)} Q^*(z, dy) - \lambda(z) = 0. \tag{12}$$

Here $L_{\mathfrak{X}}^*$ denotes the adjoint of $(L_{\mathfrak{X}}, D(L_{\mathfrak{X}}))$ as an operator acting on the space $L^2_{\mu_0}$. Then μ is formally an invariant measure for the semigroup \mathcal{P}_t in the sense that

$$\int_{\mathcal{Z}} \mathcal{L}f(z) \mu(dz) = 0 \tag{13}$$

for all $f \in D(L_{\mathfrak{X}}) \cap D_{C_0}$.

In order to verify that μ is an invariant measure it is necessary that (13) holds for all f in a core for (\mathcal{L}, D_{C_0}) . In general for PDMP it is extremely difficult to determine a core for the generator, in Section 5 we prove that sufficiently smooth functions are a core for the Boomerang Sampler and find the invariant measure.

Proof of Theorem 3.1. Fix $f \in D(L_{\mathfrak{X}}) \cap D_{C_0}$ then we have

$$\begin{aligned} \int_{\mathcal{Z}} \mathcal{L}f(z) \mu(dz) &= \int_{\mathcal{Z}} f(z) L_{\mathfrak{X}}^*(e^{-\Phi(\cdot)})(z) \mu_0(dz) + \int_{\mathcal{Z}} \int_{\mathcal{Z}} \lambda(z) f(y) Q(z, dy) e^{-\Phi(z)} \mu_0(dz) \\ &\quad - \int_{\mathcal{Z}} \lambda(z) f(z) e^{-\Phi(z)} \mu_0(dz) \\ &= \int_{\mathcal{Z}} f(z) \left[e^{\Phi(z)} L_{\mathfrak{X}}^*(e^{-\Phi(\cdot)})(z) + \int_{\mathcal{Z}} \lambda(y) e^{\Phi(z)-\Phi(y)} Q^*(z, dy) - \lambda(z) \right] \\ &\quad \times \mu(dz) \\ &= 0. \quad \square \end{aligned}$$

Example 3.2. Suppose the transition kernel Q corresponds to a deterministic transition, i.e. $Q(z, dy) = \delta_{R(z)}(dy)$ for some invertible function R . Assume that

- (1) The measure μ_0 is invariant under R , i.e. $\mu_0 \circ R = \mu_0$;
- (2) The mapping Φ is invariant under R^{-1} , i.e. $\Phi \circ R^{-1} = \Phi$;
- (3) The map $z \mapsto e^{-\Phi(z)}$ belongs to the domain of $L_{\mathfrak{X}}^*$, and

$$e^{\Phi} L_{\mathfrak{X}}^* e^{-\Phi} + \lambda \circ R^{-1} - \lambda = 0. \tag{14}$$

Then $\mu(\mathcal{L}f) = 0$ for all $f \in D_{C_0}$. Indeed, in this case $Q^*(y, dz) = \delta_{R(y)}$ and we can rewrite the left hand side of (12) as

$$e^{\Phi(z)} L_{\mathfrak{X}}^*(e^{-\Phi(\cdot)})(z) + \lambda(R^{-1}(z))e^{\Phi(z)-\Phi(R^{-1}(z))} - \lambda(z).$$

Using that Φ is invariant under R^{-1} this simplifies to

$$e^{\Phi(z)} L_{\mathfrak{X}}^*(e^{-\Phi(\cdot)})(z) + \lambda(R^{-1}(z)) - \lambda(z)$$

which is equal to zero by (14) therefore the measure $e^{-\Phi} \mu_0$ is an invariant measure. Δ

Example 3.3 (Finite Dimensional Zig-Zag Process). We now consider a finite dimensional example, set $\mathcal{Z} = \mathbb{R}^d \times \mathbb{R}^d$, and for any $z \in \mathbb{R}^d \times \mathbb{R}^d$ we write $z = (x, v)$ for some $x \in \mathbb{R}^d, v \in \mathbb{R}^d$. Suppose we wish to sample from the measure $\pi(dx) = e^{-\Phi(x)} dx$, which we extend to a measure μ on $\mathbb{R}^d \times \mathbb{R}^d$ by $\mu(dx, dv) = \pi(dx) \nu_0(dv)$ where ν_0 is the uniform distribution on $\{1, -1\}^d$. Let

$$\mathcal{X} \begin{pmatrix} x \\ v \end{pmatrix} = \begin{pmatrix} v \\ 0 \end{pmatrix}$$

and set

$$Q(z, dy) = \sum_{i=1}^d \frac{\lambda_i(z)}{\lambda(z)} \delta_{F_i z}$$

where $F_i(x, v) = (x, v_1, \dots, v_{i-1}, -v_i, v_{i+1}, \dots, v_d)$, $\lambda(z) = \sum_{i=1}^d \lambda_i(z)$ and $\lambda_i(x, v) = (v_i \partial_{x_i} \Phi(x))^+$. Note that $L_{\mathfrak{X}}$ has an invariant measure $\mu_0 = \nu_0(dv) dx$ which is not a probability measure. By [15, Theorem 27.6] the semigroup associated to this process is a Feller semigroup.

For this example we have that

$$Q^*(z, dy) = \sum_{i=1}^d \frac{\lambda_i(F_i z)}{\lambda(F_i z)} \delta_{F_i z}$$

in which case, showing that (12) holds reduces to

$$\partial_{x_i} \Phi(x) v_i + \lambda_i(F_i z) - \lambda_i(z) = 0.$$

It is immediate to check that this holds for our choice of λ_i . Δ

4. Examples of infinite dimensional PDMPs

4.1. Zig-Zag sampler in infinite dimensional Hilbert spaces

In this section we present an infinite dimensional version of the Zig-Zag Sampler, which we refer to as IDZZ, evolving in the Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$. This algorithm is designed to target the measure π as defined in (1), however as in the finite dimensional ZZ we extend the space to be of the form (x, v) . That is, we set the state space to be

$$\mathcal{Z} = \mathcal{H} \times \mathcal{V}_{ZZS} \tag{15}$$

where

$$\mathcal{V}_{ZZS} = \{v \in \mathcal{H} : \Sigma^{-1} v \in \mathcal{H}\},$$

and we recall (from Section 1.2) that Σ is a self-adjoint trace class operator on \mathcal{H} . We equip \mathcal{V}_{ZZS} with the inner product $(v, w) \mapsto \langle \Sigma^{-1} v, \Sigma^{-1} w \rangle$, making it into a Hilbert space. Note that

in order to ensure well-posedness of the IDZZ we require that the auxiliary variable v belongs to a smoother space than \mathcal{H} , indeed the space \mathcal{V}_{ZZS} is also contained within the Cameron–Martin space of π_0 and hence is a null set with respect to π_0 . On this extended state space \mathcal{Z} we design the IDZZ to have invariant measure μ defined by (8) where

$$\mu_0 = \pi_0 \times \nu_a, \quad \text{where } \pi_0 \sim \mathcal{N}(0, \Sigma), \quad \nu_a = \bigotimes_{i=1}^{\infty} \left(\frac{1}{2} \delta_{-a_i} + \frac{1}{2} \delta_{a_i} \right), \quad a \in \mathcal{V}_{ZZS}. \quad (16)$$

Note here that different choices of a can lead to different invariant measures so the process we define is not ergodic on the space $\mathcal{H} \times \mathcal{V}_{ZZS}$. This could be resolved by restricting \mathcal{V}_{ZZS} to the space

$$\{v \in \mathcal{H} : v = (v_i)_{i=1}^{\infty}, v_i \in \{a_i, -a_i\}\}$$

for some fixed $a = (a_i)_{i=1}^{\infty} \in \mathcal{V}_{ZZS}$. However to remain consistent with the framework of this paper it is more convenient to have that \mathcal{V}_{ZZS} is a Hilbert space.

In this section we will make the following assumptions on the functional Φ . Suppose Φ is everywhere Fréchet differentiable. For each $x \in \mathcal{H}$ the Fréchet derivative $d\Phi(x)$ is identified with an element $\nabla\Phi(x)$ of \mathcal{H} .

Assumption 4.1. The functional Φ satisfies the following:

- (1) **Domain of Φ :** the functional Φ is defined everywhere on \mathcal{H} and

$$0 \leq \Phi(x) \lesssim 1 + \|x\|^2.$$

- (2) **Derivatives of Φ :** The function Φ is differentiable and $\nabla_x \Phi$ is locally (on bounded sets) Lipschitz and grows at most linearly:

$$\|\nabla_x \Phi(x)\| \lesssim 1 + \|x\|. \quad (17)$$

- (3) **Growth of Φ :** The functional Φ can be written as $\Phi = \Phi_1 + \Phi_2$ where Φ_1 is convex and bounded from below, and Φ_2 is bounded with bounded first and second order derivatives.

The generator of the IDZZ is then defined as follows:

$$\mathcal{L}f(x, v) = \langle \nabla f(x), v \rangle + \sum_{i=1}^{\infty} \lambda_i(x, v) (f(x, F_i v) - f(x, v)), \quad (x, v) \in \mathcal{H} \times \mathcal{V}_{ZZS} \quad (18)$$

where f is in a class of suitably smooth functions. The switching intensities are given by

$$\lambda_i(x, v) = \left(v_i \left(\frac{x_i}{\gamma_i^2} + \partial_{x_i} \Phi(x) \right) \right)^+.$$

In terms of the abstract framework in Section 2, we have

$$\begin{aligned} \lambda(x, v) &= \sum_{i=1}^{\infty} \lambda_i(x, v), \quad Q_i(x, v, dw) = \delta_{F_i v}(dw), \\ Q(x, v, dy, dw) &= \frac{1}{\lambda(x, v)} \sum_{i=1}^{\infty} \lambda_i(x, v) \delta_x(dy) Q_i(x, v, dw). \end{aligned}$$

Remark 4.2. As λ is defined by an infinite series we must verify that it is well defined for all $(x, v) \in \mathcal{H} \times \mathcal{V}_{ZZS}$. Note that

$$\sum_{i=1}^{\infty} \lambda_i(x, v) \leq \sum_{i=1}^{\infty} \frac{|v_i|}{\gamma_i^2} |x_i| + \sum_{i=1}^{\infty} |v_i| |\partial_{x_i} \Phi(x)| \leq \|\Sigma^{-1}v\| \|x\| + \|v\| \|\nabla \Phi(x)\|. \tag{19}$$

Therefore λ is well defined. \triangle

Theorem 4.3. *Let Assumption 4.1 hold. Then there exists a piecewise deterministic Markov process in $\mathcal{H} \times \mathcal{V}$ with (formal) generator (18). Moreover, for any $a \in \mathcal{V}_{ZZS}$ the product measure μ given by (8) with μ_0 given by (16) is invariant for such a process.*

Proof. The proof is deferred to Section 8.1.1. \square

4.1.1. Choice of velocities

In the following example we discuss how to choose optimal velocity magnitudes (a_i) for the infinite dimensional zig-zag process, in order to minimise the computational complexity for estimation of the functional $x \mapsto \|x\|_r^2$, where for $r \geq 0$

$$\|x\|_r^2 := \langle \Sigma^{-r}x, x \rangle.$$

I.e., we will derive below the limiting normal distribution of

$$\frac{1}{\sqrt{T}} \int_0^T \{ \|X(s)\|_r^2 - \pi_0(\| \cdot \|_r^2) \} ds,$$

as $T \rightarrow \infty$, depending on the choice of (a_i) , and determine how to choose (a_i) in order to minimise the asymptotic mean squared error given a fixed computational budget (i.e., for a fixed maximum number of switches).

Let σ_f^2 denote the asymptotic variance for a one-dimensional canonical (i.e., no refreshments) Zig-Zag process with position $(X(t))_{t \geq 0}$ and velocities $V(t) \in \{-1, 1\}$ targeting the standard normal distribution, evaluated with respect to the function f , i.e.

$$\frac{1}{\sqrt{T}} \int_0^T \{ f(X(s)) ds - \pi(f) \} ds \xrightarrow{d} \mathcal{N}(0, \sigma_f^2),$$

where π denotes the standard normal distribution. Write π_γ for the distribution $\mathcal{N}(0, \gamma^2)$.

Lemma 4.4. *A one-dimensional Zig-Zag process with velocities $\{-a, a\}$ with stationary distribution $\mu_{\gamma,a} = \pi_\gamma \otimes \text{Uniform}(\{-a, +a\})$ and canonical switching intensities, evaluated with respect to the function $\tilde{f}(x) = cf(x/\gamma)$ has asymptotic variance $c^2\gamma\sigma_f^2/a$. The expected number of switches compared to the canonical Zig-Zag process with unit speed velocities is proportional to a/γ (per unit time, in stationarity).*

Proof. Let $\tilde{X}(t)$ denote the Zig-Zag process as specified. The stationary distribution of $\tilde{X}(t)/\gamma$ is then a standard normal distribution, and the process $\tilde{X}(t)/\gamma$ has velocities $\pm a/\gamma$. Slowing down the clock by a factor a/γ , the process $\Xi(t) := \tilde{X}(\gamma t/a)/\gamma$ is a Zig-Zag process at unit

speed targetting the standard normal distribution. We compute

$$\begin{aligned} \frac{1}{\sqrt{T}} \int_0^T \{cf(\tilde{X}(s)/\gamma) - \mathbb{E}_{\pi_\gamma} f(\tilde{X}/\gamma)\} ds &= \frac{1}{\sqrt{T}} \int_0^T \{cf(\Xi(as/\gamma)) - c\pi(f)\} ds \\ &= \frac{\gamma}{a\sqrt{T}} \int_0^{aT/\gamma} \{cf(\Xi(r)) - c\pi(f)\} dr \\ &= \frac{c\sqrt{\gamma}}{\sqrt{a}} \frac{1}{\sqrt{T}} \int_0^{\tilde{T}} \{f(\Xi(r)) - \pi(f)\} dr \\ &\xrightarrow{d} \mathcal{N}(0, c^2\gamma\sigma_f^2/a). \end{aligned}$$

The number of switches follows directly from the slowing down factor. \square

Example 4.1. Define v^2 to be the asymptotic variance of the canonical Zig-Zag process with unit velocities targetting the standard normal distribution, evaluated with respect to the function $x \mapsto x^2$. By [7, Example 1], we have $v^2 = 4\sqrt{2/\pi}$, but the exact value is of no importance here. If we consider now a zigzag process targetting $\mathcal{N}(0, \gamma^2)$ with speeds $\{-a, a\}$ evaluated with respect to the function $x \mapsto x^2/\gamma^{2r} = cx^2/\gamma^2$ for $c = \gamma^{2(1-r)}$, we obtain from Lemma 4.4 that the associated asymptotic variance is $\gamma^{5-4r}v^2/a$. \triangle

We will now consider an infinite dimensional Zig-Zag process targetting the function $x \mapsto \|x\|_r^2$, for $r > 0$. To ensure that all the required operations are welldefined, we introduce the following assumption.

Assumption 4.5. Suppose $r \in [0, \frac{1}{2})$ is such that $\sum_{i=1}^\infty \gamma_i^{2-4r} < \infty$.

Example 4.2. In the case of a Wiener process in $L^2[0, 1]$, we have $\gamma_i^2 = (i - 1/2)^{-2}\pi^{-2} \sim i^{-2}$. Consequently, for any $r < 1/4$ we have $\sum_{i=1}^\infty \gamma_i^{2-4r} < \infty$, so that Assumption 4.5 is satisfied. \triangle

Proposition 4.6. Suppose Assumption 4.5 is satisfied. Consider the zigzag process of Theorem 4.3 with $\Phi(x) = 0$ for all x . The asymptotic variance σ_r^2 of the Zig-Zag process with respect to the function $x \mapsto \|x\|_r^2$ is given by $\sigma_r^2 = v^2 \sum_{i=1}^\infty \gamma_i^{5-4r}/a_i$, where v^2 is as defined in Example 4.1. The expected number of switches per unit time interval is proportional to $\sum_{i=1}^\infty a_i/\gamma_i$.

Proof. Note that the generator of the Zig-Zag process can be interpreted as a sum of infinitely many one-dimensional Zig-Zag processes. By this factorisation property and Example 4.1, the asymptotic variance decomposes as

$$\sigma_r^2 = \sum_{i=1}^\infty \gamma_i^{5-4r} v^2/a_i.$$

The proportionality factor of the number of switches is a direct consequence of the second statement of Lemma 4.4. \square

Suppose now that we have a fixed computational budget of (approximately) N operations, and we wish to minimise the approximate standard error $\epsilon_{r,T} = \sigma_r/\sqrt{T}$, where T is the length

of the time interval we will simulate within N operations. We may then phrase the question of choosing optimal velocities (a_i) as follows:

$$\text{minimise } \frac{1}{T} v^2 \sum_{i=1}^{\infty} \gamma_i^{5-4r} \quad \text{subject to } T \sum_{i=1}^{\infty} a_i/\gamma_i \leq N.$$

Resolving the constraint by setting T at its maximal possible value, $T = \frac{N}{\sum_{i=1}^{\infty} a_i/\gamma_i}$, the problem transforms into the minimisation problem

$$\text{minimise } \frac{v^2}{N} \left(\sum_{i=1}^{\infty} \gamma_i^{5-4r}/a_i \right) \left(\sum_{i=1}^{\infty} a_i/\gamma_i \right).$$

Minimisation with respect to a_i (using a Cauchy–Schwarz argument) yields

$$a_i = C \gamma_i^{3-2r}$$

for some constant C , where the choice of the constant C does not affect the minimisation objective.

We find that, for this choice of (a_i) , the expected number of switches is finite,

$$N = T \sum_{i=1}^{\infty} a_i/\gamma_i = T \sum_{i=1}^{\infty} \gamma_i^{2-2r} \leq T \left(\max_j \gamma_j^{2r} \right) \sum_{i=1}^{\infty} \gamma_i^{2-4r} < \infty$$

by [Assumption 4.5](#). Furthermore the Zig-Zag process is well-defined using [Theorem 4.3](#) since

$$\|\Sigma^{-1}v\|^2 = \sum_{i=1}^{\infty} \frac{a_i^2}{\gamma_i^4} = \sum_{i=1}^{\infty} \gamma_i^{2-4r} < \infty$$

by [Assumption 4.5](#).

Note that, for large r , the functional $x \mapsto \|x\|_r^2$ is sensitive to the high frequency components of x (i.e. the large indices in the expansion $x = \sum_{i=1}^{\infty} x_i e_i$). Also, for large r , the high frequency velocities $a_i = \gamma_i^{3-2r}$ are relatively larger, so that the high frequency components are explored more.

We have shown how, for a particular example of a functional $x \mapsto \|x\|_r^2$ relative to a Gaussian target we may choose velocities to minimise the statistical error when using the IDZZ as a sampling method, for a fixed computational budget. It is an interesting research problem, but beyond the scope of this work, how this may be extended to other functionals, general target distributions and other sampling methods.

4.2. Bouncy particle sampler in infinite dimensional Hilbert spaces

In this section we introduce an extension of the BP algorithm which is well-defined in the infinite dimensional Hilbert space

$$\mathcal{Z} = \mathcal{H} \times \mathcal{V}_{\text{BPS}}.$$

As with IDZZ the velocity component needs to be smoother than the position component, set

$$\mathcal{V}_{\text{BPS}} = \{v \in \mathcal{H} : \Sigma^{-1}v \in \mathcal{H}\}.$$

We will design the IDBP to have invariant measure μ defined by [\(8\)](#) where

$$\mu_0 = \pi_0 \times \nu_0, \quad \text{where } \pi_0 \sim \mathcal{N}(0, \Sigma), \nu_0 \sim \mathcal{N}(0, \Sigma^\xi) \tag{20}$$

and $\zeta \geq 2$. Note that as $\zeta \geq 2$ then $\nu_0(\mathcal{V}_{\text{BPS}}) = 1$ therefore μ can be viewed as a probability measure on $\mathcal{H} \times \mathcal{V}_{\text{BPS}}$.

The generator of IDBP is defined, on a set of suitably regular functions, as

$$\begin{aligned} \mathcal{L}f(x, v) &= \langle v, \nabla_x f(x, v) \rangle + \lambda(x, v)[f(x, R(x)v) - f(x, v)] \\ &\quad + \lambda_{\text{refr}} \int_{\mathcal{V}_{\text{BPS}}} [f(x, w) - f(x, v)] \nu_0(dw), \end{aligned} \tag{21}$$

where

$$\lambda(x, v) := \langle v, \nabla \Psi \rangle^+, \quad R(x)v := v - \frac{2\langle v, \nabla \Psi(x) \rangle}{\|\Sigma^{\zeta/2} \nabla \Psi(x)\|^2} \Sigma^{\zeta} \nabla \Psi(x), \tag{22}$$

$\lambda_{\text{refr}} > 0$ is a positive constant, Ψ is been defined by

$$\Psi(x) := \Phi(x) + \frac{1}{2} \langle \Sigma^{-1}x, x \rangle. \tag{23}$$

Remark 4.7. Note that $\Psi(x)$ is not well defined for all $x \in \mathcal{H}$ since Σ^{-1} is an unbounded operator however this is to be viewed as just a formal definition. Indeed, the IDBPS does not rely on Ψ itself but only on $\nabla \Psi$, which always appears with a term to smooth it. In [Proposition 4.10](#) we show that all the terms that appear are well defined for all $x \in \mathcal{H}$ and $v \in \mathcal{V}_{\text{BPS}}$. \triangle

We will show in [Proposition 4.10](#) that the resulting algorithm generates a continuous-time dynamics $(X_t, V_t) \in \mathcal{H} \times \mathcal{V}_{\text{BPS}}$ with invariant distribution μ . Here well-posedness is in the sense of [Propositions 4.8](#) and [4.10](#).

Proposition 4.8. *Let [Assumption 4.1](#) hold. For any $\zeta \geq 4$ and for every $x \in \mathcal{H}$, $v \in \mathcal{V}_{\text{BPS}}$, we have*

- (1) *the intensity $\lambda(x, v)$ is well defined (in the sense that the scalar product $\langle v, \nabla \Psi(x) \rangle$ is finite);*
- (2) *$R(x)v \in \mathcal{V}_{\text{BPS}}$.*

Proof. The proof is deferred to [Section 8.1.2](#). \square

Remark 4.9. To explain the choice of scaling of the covariance operator in [\(21\)–\(22\)](#) and at the same time show properties of IDBP, let us consider the following (formal) generator

$$\begin{aligned} \mathcal{L}f(x, v) &= \langle v, \nabla_x f(x, v) \rangle + \lambda(x, v)[f(x, R(x)v) - f(x, v)] \\ &\quad + \lambda_{\text{ref}} [(Qf)(x, v) - f(x, v)], \end{aligned} \tag{24}$$

with

$$\lambda(x, v) := \langle \Sigma^\eta v, \nabla \Psi \rangle^+, \quad R(x)v := v - \frac{2\langle v, \Sigma^\epsilon \nabla \Psi \rangle}{\|\Sigma^\gamma \nabla \Psi\|^2} \Sigma^\beta \nabla \Psi, \tag{25}$$

where $\delta, \eta, \epsilon, \gamma, \beta$ are arbitrary positive parameters. Then in order to have that $R(x)$ is an involution, the measure ν_0 is invariant under $R(x)$ and that μ is an invariant measure it is necessary that $\beta = 2\gamma, \epsilon = \eta = 0$, which gives [\(21\)](#). \triangle

Proposition 4.10. *With the notation introduced above,*

- (1) The reflection operator is involutive, i.e. $R(x)[R(x)v] = v$, for every $x \in \mathcal{H}$, $v \in \mathcal{V}_{\text{BPS}}$;
- (2) The reflection operator satisfies the following property

$$\langle R(x)v, \nabla \Psi \rangle = -\langle v, \nabla \Psi \rangle. \tag{26}$$

- (3) The Gaussian measure is invariant under reflections, i.e., the centred measure ν_0 with covariance Σ^ζ and the measure $\nu_R := \nu_0 \circ R(x)$ coincide on \mathcal{H} ;
- (4) The measure μ is invariant for the operator \mathcal{L} in (24)–(25).

Proof of Proposition 4.10. The proof is deferred to Section 8.1.2. \square

Lemma 4.11. With the notation introduced above, for each $\alpha \in \mathbb{R}$ define $\|x\|_\alpha = \|\Sigma^\alpha x\|$. Set $\Phi = 0$. Then

$$\|R(x)v\|_\alpha \leq \|v\| \quad \text{for any } (x, v) \in \mathcal{H} \times \mathcal{V}_{\text{BPS}} \tag{27}$$

holds if and only if $\alpha = -\zeta/2$. Therefore Assumption 2.1 cannot be satisfied for any $\alpha > -1$.

Remark 4.12. We have shown that if IDBPS is well posed then it has the correct invariant measure. However as shown in 4.11 even for the quadratic problem where $\Phi = 0$ in order to satisfy Assumption 2.1 we need to work in a space that has measure zero with respect to the measure μ . Hence we cannot apply Theorem 2.3 in this case. Therefore it remains to show that IDBPS is well posed, in particular one needs to show that it is not possible to have an infinite number of reflections in a finite time interval. This problem can be alleviated by replacing the reflection operator $R(x)$ by $R(x)v = -v$, i.e. to introduce pure reflections in the IDBPS. Establishing wellposedness for $R(x)$ as given by (22) is beyond the scope of this work; however we will give here a heuristic argument to show why we can expect IDBPS to exist under Assumption 4.1.

We estimate the expected number of reflections per unit time by considering the rate in stationarity, that is we will bound

$$\mathbb{E}_\mu[\lambda(X, V)] = \mathbb{E}_\mu[\langle V, \nabla \Phi(X) + \Sigma^{-1}X \rangle^+].$$

By Assumption 4.1 we have that $\nabla \Phi$ grows at most linearly and therefore we have the bound

$$\mathbb{E}_\mu[\lambda(X, V)] \leq \mathbb{E}_\mu[\|V\|(1 + \|X\|) + \|\Sigma^{-1}V\|\|X\|].$$

By independence of X, V and using that $\Sigma^{-1}V$ is Gaussian with covariance $\Sigma^{\zeta-2}$, these terms are bounded. \triangle

4.3. Boomerang Sampler

We now introduce the Boomerang Sampler which differs from IDZZ and IDBP by having circular deterministic dynamics instead of linear motion. As with IDZZ and IDBP we will extend the state space \mathcal{H} to include a velocity component and we define the IDB on the infinite dimensional Hilbert space

$$\mathcal{Z} = \mathcal{H} \times \mathcal{H}.$$

In contrast to the IDZZ and IDBP settings we may allow the velocity to belong to the same space as the position component. We design the IDB to have invariant measure given by (8) where $\mu_0 = \pi_0 \times \nu_0$ for some Gaussian measures $\pi_0 = \mathcal{N}(0, \Sigma_x)$, $\nu_0 = \mathcal{N}(0, \Sigma_v)$.

4.3.1. Definition and well posedness of the Boomerang Sampler

The dynamics of the Boomerang Sampler are composed of three different mechanisms: the deterministic flow which is chosen to be a Hamiltonian flow, designed to preserve the reference measure μ_0 ; reflections which will preserve the same Hamiltonian as the deterministic flow; and refreshment which will happen at a constant rate. First we describe the deterministic dynamics. Fix two bounded positive self-adjoint linear operators \mathcal{X}, \mathcal{P} and take φ to be the flow map associated to the Hamiltonian dynamics described by

$$\frac{d}{dt}\varphi_t(x, v) = \begin{pmatrix} 0 & \mathcal{P} \\ -\mathcal{X} & 0 \end{pmatrix} \varphi_t(x, v). \tag{28}$$

Recall that φ_t takes values in $\mathcal{H} \times \mathcal{H}$ so we can view $\varphi_t(x, v)$ as two dimensional vector whose components belong to \mathcal{H} . These dynamics correspond to the Hamiltonian

$$E^{\mathcal{X}, \mathcal{P}}(x, v) := \langle \mathcal{X}x, x \rangle_{\mathcal{H}} + \langle \mathcal{P}v, v \rangle_{\mathcal{H}}. \tag{29}$$

There are two natural choices for $E^{\mathcal{X}, \mathcal{P}}(x, v)$, one is the $\mathcal{H} \times \mathcal{H}$ norm which corresponds to taking $\mathcal{X} = \mathcal{P} = 1$ and we will see that this requires that $\Sigma_x = \Sigma_v$. Another natural choice is to take $\mathcal{X} = \Sigma_x^{-1}$ and $\mathcal{P} = \Sigma_v^{-1}$ however the Hamiltonian is then only defined on the Cameron Martin Space which has measure zero with respect to μ . This gives many technical difficulties and even well-posedness of the process may not hold (see [Example 4.3](#)) so we shall restrict ourselves to the case where \mathcal{X}, \mathcal{P} are bounded operators which implies that the Hamiltonian is everywhere defined. Define the operators

$$L_{\mathfrak{X}}f(x, v) = \frac{d}{dt} \Big|_{t=0} f(\varphi_t(x, v)) = \left\langle \begin{pmatrix} \mathcal{P}v \\ -\mathcal{X}x \end{pmatrix}, \begin{pmatrix} \nabla_x f(x, v) \\ \nabla_v f(x, v) \end{pmatrix} \right\rangle_{\mathcal{H}^2} \tag{30}$$

$$L_r f(x, v) = \lambda_r \int_{\mathcal{H}} [f(x, w) - f(x, v)] \nu_0(dw) \tag{31}$$

$$L_b f(x, v) = \sum_{i=1}^{\infty} \lambda_i(x, v) [f(x, R_i(x)v) - f(x, v)]. \tag{32}$$

Here λ_r is a positive constant, $\lambda_i : \mathcal{H} \times \mathcal{H} \rightarrow [0, \infty)$ for each $i \in \mathbb{N}$ and for each $i \in \mathbb{N}, x \in \mathcal{H}, R_i(x)$ is a bounded linear operator. We formally summarise the Markov process (X_t, V_t) by the generator

$$\mathcal{L} = L_{\mathfrak{X}} + L_r + L_b.$$

Such a process exists by [Proposition 2.4](#) provided the following [Hypothesis 4.13](#) holds. This process is a PDMP defined on $\mathcal{Z} = \mathcal{H} \times \mathcal{H}$ with characteristics

$$\mathfrak{X} : \begin{pmatrix} x \\ v \end{pmatrix} \mapsto \begin{pmatrix} \mathcal{P}v \\ -\mathcal{X}x \end{pmatrix}, \tag{33}$$

$$\lambda(x, v) := \sum_{i=1}^{\infty} \lambda_i(x, v) + \lambda_r,$$

$$\begin{aligned} Q((x, v), A) &:= \frac{1}{\sum_j \lambda_j(x, v) + \lambda_r} \sum_{i=1}^{\infty} \lambda_i(x, v) \delta_{(x, R_i(x)v)}(A) \\ &+ \frac{\lambda_r}{\sum_j \lambda_j(x, v) + \lambda_r} (\delta_x \times \nu_0)(A). \end{aligned}$$

From now on we will assume the following hypothesis holds.

Hypothesis 4.13. Assume that:

- (1) The intensities λ_n and λ are continuous and bounded by the Hamiltonian $E^{\mathcal{X}, \mathcal{P}}(x, v)$ (recall this is defined by (29)), that is

$$\sum_{n=1}^{\infty} \lambda_n(x, v) \leq C(1 + E^{\mathcal{X}, \mathcal{P}}(x, v)). \tag{34}$$

- (2) The flow map φ_t is given by the solution of (28) for some bounded linear operators \mathcal{X}, \mathcal{P} which are self-adjoint and commute;
- (3) Each reflection operator $R_i(x)$ is a linear bounded operator for any $x \in \mathcal{H}$, and satisfies

$$\langle \mathcal{P}R_i(x)v, R_i(x)v \rangle_{\mathcal{H}} \leq \langle \mathcal{P}v, v \rangle_{\mathcal{H}}, \text{ for all } x, v \in \mathcal{H}. \tag{35}$$

Moreover, assume that $(x, v) \mapsto R_i(x)v$ is continuous.

Remark 4.14. We shall now comment on each of these assumptions in turn.

- (1) Since \mathcal{X}, \mathcal{P} are bounded φ is well defined and for each initial condition (x, v) the map $t \mapsto \varphi_t(x, v)$ is smooth.
- (2) Requiring that the reflections R_i do not increase the Hamiltonian allows us to prove that the process is non-explosive, i.e. $\sup_i T^i = \infty$ almost surely. The strategy for showing this relies on the fact that except for refreshment jumps the dynamics is bounded by the Hamiltonian $E^{\mathcal{X}, \mathcal{P}}(x, v)$ therefore the jump rates are bounded since (34) holds.
- (3) Under Hypothesis 4.13 we have that Assumption 2.1 also holds, in particular the Boomerang Sampler is well-defined by Theorem 2.3. Since \mathcal{P}, \mathcal{X} are bounded we can work with the seminorm $E^{\mathcal{X}, \mathcal{P}}$ in which case Assumption 2.1(iii) follows from (35) and Assumption 2.1(vii) is satisfied with c constant since the flow φ_t preserves the Hamiltonian. \triangle

Definition 4.15. Given a Hilbert space \mathcal{H} , a covariance operator Σ_x and a function Φ we say that the process $\{X_t, V_t\}_{t \geq 0}$ is a *Boomerang Sampler with characteristics* $(\Sigma_v, \mathcal{X}, \mathcal{P}, \{\lambda_i\}_i, \lambda_r, \{R_i\}_i)$ if Hypothesis 4.13 is satisfied and (X_t, V_t) is the PDMP constructed in Section 2 with characteristics given by (33). The generator of this process acts on sufficiently smooth functions by

$$\begin{aligned} \mathcal{L}f(x, v) &= \langle \mathcal{P}v, \nabla_x f(x, v) \rangle_{\mathcal{H}} - \langle \mathcal{X}x, \nabla_v f(x, v) \rangle_{\mathcal{H}} + \lambda_r \int_{\mathcal{H}} f(x, w) - f(x, v) \nu_0(dw) \\ &+ \sum_{i=1}^{\infty} \lambda_i(x, v)[f(x, R_i(x)v) - f(x, v)] \end{aligned}$$

We call the Boomerang Sampler with characteristics $(\Sigma_x, 1, 1, (\langle \nabla_x \Phi(x), v \rangle_{\mathcal{H}})_+, \lambda_r, v \mapsto -v)$ the *pure reflection Boomerang Sampler*, here $\lambda_i = 0$ for $i > 1$ so we have not included them in the notation.

The following example shows that if we drop the assumption that \mathcal{X}, \mathcal{P} are bounded then we can still define a weak solution to the dynamics. For simplicity we will take $\mathcal{X} = \mathcal{P} = \Sigma_x^{-1}$.

Example 4.3. Suppose that \mathcal{X} is an (unbounded) self-adjoint positive operator which is diagonalisable, that is there exists an orthonormal basis of eigenvectors. Then there is a process (X_t, V_t) such that

$$X_t - X_0 = \mathcal{X} \int_0^t V_s ds \tag{36}$$

$$V_t - V_0 = -\mathcal{X} \int_0^t X_s ds. \tag{37}$$

We will construct this process by converting the problem into an infinite system of two dimensional ODEs. Fix an orthonormal basis $\{e_k\}_{k=1}^\infty$ of eigenvectors of \mathcal{X} and let γ_k^2 be the eigenvalue associated to e_k .

Consider the system

$$\frac{d}{dt} X_t^i = \gamma_i^2 V_t^i, \quad \frac{d}{dt} V_t^i = -\gamma_i^2 X_t^i$$

with initial condition $X_0^i = x^i, V_0^i = v^i$. The solution of this system is given by

$$\begin{pmatrix} X_t^i \\ V_t^i \end{pmatrix} = \begin{pmatrix} x \cos(\gamma_i^2 t) + v \sin(\gamma_i^2 t) \\ -x \sin(\gamma_i^2 t) + v \cos(\gamma_i^2 t) \end{pmatrix}. \tag{38}$$

We also have that

$$|X_t^i|^2 + |V_t^i|^2 = |x^i|^2 + |v^i|^2$$

In particular, this gives us that $\sum_{i=1}^N X_t^i e_i$ and $\sum_{i=1}^N V_t^i e_i$ converge to some $X_t, V_t \in \mathcal{H}$ for each $t \geq 0$. So we have

$$X_t - X_0 = \lim_{N \rightarrow \infty} \sum_{i=1}^N (X_t^i - x^i) e_i = \lim_{N \rightarrow \infty} \sum_{i=1}^N \int_0^t V_s^i \gamma_i e_i ds = \lim_{N \rightarrow \infty} \mathcal{X} \left(\sum_{i=1}^N \int_0^t V_s^i e_i ds \right)$$

Since \mathcal{X} is a closed operator $\int_0^t V_s ds, \int_0^t X_s ds \in D(\mathcal{X})$ and

$$X_t - X_0 = \mathcal{X} \int_0^t V_s ds, \quad V_t - V_0 = \mathcal{X} \int_0^t X_s ds.$$

Therefore we have weak solution to (28) but a weak solution is the most we can expect. To have a strong solution we require that $X_t \in D(\mathcal{X})$ and $V_t \in D(\mathcal{P})$. However $X_t, V_t \in D(\mathcal{X})$ if and only if

$$\sum_{i=1}^N \gamma_i^2 (|X_t^i|^2 + |V_t^i|^2) < \infty.$$

Using the expressions for X_t^i, V_t^i given by (38) we can rewrite this as

$$\sum_{i=1}^N \gamma_i^2 (|X_t^i|^2 + |V_t^i|^2) = \sum_{i=1}^N \gamma_i^2 (|x^i|^2 + |v^i|^2).$$

Therefore $X_t, V_t \in D(\mathcal{X})$ if and only if $x, v \in D(\mathcal{X})$. Now as \mathcal{X} is a closed densely defined operator we have that $D(\mathcal{X}) = \mathcal{H}$ if and only if \mathcal{X} is bounded. \triangle

4.3.2. Invariant measure

In the previous section we constructed the PDMP and showed that the construction was well defined for all $t \geq 0$, now we show that it has the desired invariant measure. Recall the definition of the semigroup was given by (5), here we take $\mathcal{Z} = \mathcal{H} \times \mathcal{H}$ and $Z_t = (X_t, V_t)$.

Hypothesis 4.16. We shall assume that:

- (1) Hypothesis 4.13 holds;
- (2) The operators \mathcal{X}, \mathcal{P} satisfy the consistency condition:

$$\mathcal{P}\Sigma_v = \Sigma_x \mathcal{X}. \tag{39}$$

- (3) The function Φ satisfies 4.1 and the following relation is satisfied:

$$\sum_{n=1}^{\infty} (\lambda_n(x, v) - \lambda_n(x, R_n(x)v)) = \langle \nabla_x \Phi(x), \mathcal{P}v \rangle_{\mathcal{H}}; \tag{40}$$

- (4) The reflection operators $R_i(x)$ are involutive (i.e. $R_i(x)R_i(x)v = v$ for all $x, v \in \mathcal{H}$) and invariant under the measure ν_0 , that is

$$\int_{\mathcal{H}} f(R_i(x)v)\nu_0(dv) = \int_{\mathcal{H}} f(v)\nu_0(dv), \quad \text{for all } f \in L^2_{\nu_0}, x \in \mathcal{H}, i \geq 1. \tag{41}$$

Proposition 4.17. Fix a Hilbert space \mathcal{H} , a covariance operator Σ_x and a potential function Φ and let $\{X_t, V_t\}_{t \geq 0}$ be a Boomerang Sampler with characteristics $(\Sigma_v, \mathcal{X}, \mathcal{P}, \{\lambda_i\}_i, \lambda_r, \{R_i\}_i)$. Assume that Hypothesis 4.16 holds then μ is formally an invariant measure for $\{X_t, V_t\}_{t \geq 0}$ in the sense that (13) for all f sufficiently smooth.

Proof of Proposition 4.17. The proof is deferred to Section 8.1.3. \square

4.3.3. Pure reflection Boomerang Sampler

In this section we take $\lambda_n = 0$ for $n > 1$ and consider choices of $\lambda = \lambda_1$ and $R = R_1$. Observe that we may always choose $R(x)v = -v$, due to the symmetry of ν_0 and the Hamiltonian $E^{\mathcal{X}, \mathcal{P}}$ we have that this choice preserves both ν_0 and $E^{\mathcal{X}, \mathcal{P}}$. With this choice of R we must take λ to be

$$\lambda(x, v) = (\langle \nabla_x \Phi(x), \mathcal{P}v \rangle_{\mathcal{H}})^+ + \gamma(x, v)$$

where $\gamma(x, v) = \gamma(x, -v)$ for all $x, v \in \mathcal{H}$ and is non-negative. Let us consider another choice of $R(x)$. To ensure that $R(x)$ preserves ν_0 it is sufficient to test (41) for any f_q of the form $f_q(v) = \exp(i\langle v, q \rangle_{\mathcal{H}})$ for every $q \in \mathcal{H}$. In which case, $f_q(R(x)v) = \exp(i\langle R(x)v, q \rangle_{\mathcal{H}}) = f_{R(x)^*q}(v)$, i.e.

$$\hat{\nu}_0(q) = \hat{\nu}_0(R(x)^*q)$$

where $\hat{\nu}_0$ denotes the Fourier transform of ν_0 . Using that ν_0 is normally distributed with covariance operator Σ_v we can rewrite this as

$$\exp\left(-\frac{1}{2}\langle \Sigma_v q, q \rangle_{\mathcal{H}}\right) = \exp\left(-\frac{1}{2}\langle \Sigma_v R(x)^*q, R(x)^*q \rangle_{\mathcal{H}}\right).$$

Therefore the measure ν_0 is invariant under $R(x)$ if and only if

$$\langle \Sigma_v q, q \rangle_{\mathcal{H}} = \langle R(x)\Sigma_v R(x)^*q, q \rangle_{\mathcal{H}} \tag{42}$$

Remark 4.18. In finite dimensions we may take

$$R(x)v = v - 2 \frac{\langle \nabla_x \Phi(x), \mathcal{P}v \rangle}{\langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), \mathcal{P} \nabla_x \Phi(x) \rangle} \Sigma_v \mathcal{P} \nabla_x \Phi(x). \tag{43}$$

Indeed, for this choice we have

$$R(x)^*v = v - 2 \frac{\langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), v \rangle}{\langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), \mathcal{P} \nabla_x \Phi(x) \rangle} \mathcal{P} \nabla_x \Phi(x).$$

Now we can check that (42) holds.

$$\begin{aligned} \langle \Sigma_v R(x)^*v, R(x)^*v \rangle_{\mathcal{H}} &= \left\langle \Sigma_v v - 2 \frac{\langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), v \rangle}{\langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), \mathcal{P} \nabla_x \Phi(x) \rangle} \Sigma_v \mathcal{P} \nabla_x \Phi(x), \right. \\ &\quad \left. v - 2 \frac{\langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), v \rangle}{\langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), \mathcal{P} \nabla_x \Phi(x) \rangle} \mathcal{P} \nabla_x \Phi(x) \right\rangle_{\mathcal{H}} \\ &= \langle \Sigma_v v, v \rangle_{\mathcal{H}} - 4 \frac{\langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), v \rangle}{\langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), \mathcal{P} \nabla_x \Phi(x) \rangle} \langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), v \rangle_{\mathcal{H}} \\ &\quad + 4 \frac{\langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), v \rangle^2}{\langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), \mathcal{P} \nabla_x \Phi(x) \rangle^2} \langle \Sigma_v \mathcal{P} \nabla_x \Phi(x), \mathcal{P} \nabla_x \Phi(x) \rangle_{\mathcal{H}} \\ &= \langle \Sigma_v v, v \rangle_{\mathcal{H}}. \end{aligned}$$

However in infinite dimensional space (35) does not hold. Instead (35) holds in finite dimensions when the inner product is taken to be the Cameron Martin space inner product which is equivalent to the original inner product since in this case $\Sigma_x^{-1}, \Sigma_v^{-1}$ are bounded. Δ

4.3.4. Factorised Boomerang Sampler

For this section we consider an alternate choice for λ_n, R_n . As in the pure reflection Boomerang Sampler we take $\mathcal{X} = \mathcal{P} = 1, \Sigma_x = \Sigma_v = \Sigma$. Let $\{e_i\}_{i=1}^\infty$ be an orthonormal basis of eigenvectors of Σ and set $R_i(x)v = v - 2\langle v, e_i \rangle_{\mathcal{H}} e_i$, and set

$$\lambda_i(x, v) = (v_i \partial_{x_i} \Phi(x))^+ + \gamma_i(x, v)$$

where $v_i = \langle v, e_i \rangle_{\mathcal{H}}$ and $\gamma : \mathcal{H} \times \mathcal{H} \rightarrow [0, \infty)$ satisfies $\gamma_i(x, v) = \gamma_i(x, R_i v)$.

Assume that $\Phi \in L^1_{\pi_0}$ is smooth and $\nabla_x \Phi$ is globally Lipschitz then Hypothesis 4.16 is satisfied and we have that μ is an invariant measure by Proposition 4.17.

5. Core for the generator of Boomerang Sampler

In this section we will give conditions under which we have a core for the generator \mathcal{L} in L^2_{μ} . Even in finite dimensional situations it is difficult to determine when a set is core for the generator of a PDMP, conditions are given in [17] for the set C_c^1 to be a core in $C_0(\mathbb{R}^d)$. In particular they verify these conditions for BPS and in [3] they verify the conditions of ZZS. In both cases, the strategy is to first consider the case of a smooth intensity rate λ and show that $\mathcal{P}_t(C_c^1(\mathbb{R}^d)) \subseteq C_c^1(\mathbb{R}^d)$ then extend these results to the canonical intensity rate. See [1, Section 4.1] for a discussion of cores for PDMP in a finite dimensional setting.

For the remaining sections we will only consider the Boomerang Sampler and will concentrate on the Pure Reflection Boomerang Sampler and Factorised Boomerang Sampler. Let us first introduce a framework which will include both Pure Reflection Boomerang Sampler and Factorised Boomerang Sampler. For this section we restrict to the Boomerang Sampler with characteristics $(\Sigma, 1, 1, \{\lambda_n\}_n, \lambda_r, \{R_n\}_n)$ where

$$\lambda_n(x, v) = \frac{1}{2} (\langle \nabla_x \Phi(x), v - R_n v \rangle_{\mathcal{H}})^+. \tag{44}$$

Here R_i is a linear operator on \mathcal{H} , with $R_n^2 = 1$, $\|R_n v\| = \|v\|$ for any $v \in \mathcal{H}$, satisfies (42) and

$$\sum_{n=1}^{\infty} \|(1 - R_n)v\|^2 = 4\|v\|^2. \tag{45}$$

In this setting the generator acts on smooth functions by

$$\begin{aligned} \mathcal{L}f(x, v) &= \langle v, \nabla_x f(x, v) \rangle_{\mathcal{H}} - \langle x, \nabla_v f(x, v) \rangle_{\mathcal{H}} + \lambda_r \int_{\mathcal{H}} [f(x, w) - f(x, v)] v_0(dw) \\ &+ \sum_{n=1}^{\infty} \lambda_n(x, v) [f(x, R_n v) - f(x, v)]. \end{aligned} \tag{46}$$

Observe that here λ is the canonical rate, that is there is no γ_i terms, this is to simplify the exposition however the same results hold for any bounded function γ_i with $\gamma(x, R_i(v)) = \gamma_i(x, v)$.

Lemma 5.1. *Suppose that Φ satisfies Assumption 4.1. Let (X_t, V_t) be the Boomerang Sampler with characteristics $(\Sigma, 1, 1, \{\lambda_i\}_i, \lambda_r, \{R_i\}_i)$ where λ_i are given by (44), and let R_i satisfy the above conditions. Then Hypothesis 4.16 holds, hence (X_t, V_t) is well-defined.*

In this section we aim to show the set $\mathcal{F}C_c^\infty(\mathcal{H} \times \mathcal{H})$ is a core for the generator \mathcal{L} as given by (46). The proof that $\mathcal{F}C_c^\infty(\mathcal{H} \times \mathcal{H})$ is a core is split into 4 steps:

- **Step 1:** We consider a PDMP $(\tilde{X}_t, \tilde{V}_t)$ with smooth intensities and no refreshments i.e. $(\tilde{X}_t, \tilde{V}_t)$ is a Boomerang Sampler with characteristics $(\Sigma, 1, 1, \{\tilde{\lambda}_i\}_i, 0, \{R_i\}_i)$ where

$$\tilde{\lambda}_n(x, v) = -\log(\phi(\exp(-\langle \nabla_x \Psi(x), v - R_n v \rangle_{\mathcal{H}}))), \quad \phi(r) = r(1 + r)^{-1}. \tag{47}$$

It is a simple calculation to verify that Hypothesis 4.13 is satisfied so we have that $(\tilde{X}_t, \tilde{V}_t)$ admits a strongly continuous semigroup $\tilde{\mathcal{P}}_t$ on the space $C_0(\mathcal{H} \times \mathcal{H})$ and with generator given by $(\tilde{\mathcal{L}}, D(\tilde{\mathcal{L}}))$. Moreover for sufficiently smooth $f \in D(\tilde{\mathcal{L}})$ we have

$$\tilde{\mathcal{L}}f(x, v) = L_{\tilde{x}}f(x, v) + \sum_{n=1}^{\infty} \tilde{\lambda}_n(x, v) [f(x, R_n v) - f(x, v)]. \tag{48}$$

This choice of $\tilde{\lambda}_n$ is motivated by [3] where they use an analogous choice to obtain a smooth and strictly positive intensity for a finite dimensional ZZS. The advantage of working with $(\tilde{X}_t, \tilde{V}_t)$ is that the semigroup is differentiable since the intensity is differentiable, and as there are no refreshments the process remains on a sphere in $\mathcal{H} \times \mathcal{H}$. For this PDMP we prove that the set

$$\mathcal{C} = \{f \in C_b^1(\mathcal{H} \times \mathcal{H}) : \exists r > 0 \text{ with } \text{supp}(f) \subseteq B_r\} \tag{49}$$

is a core for $\tilde{\mathcal{L}}$, see Theorem 5.5.

- **Step 2:** We show that since \mathcal{C} is a core for $(\tilde{\mathcal{L}}, D(\tilde{\mathcal{L}}))$ we also have that \mathcal{C} is a core for (\mathcal{L}, D_{C_0}) . This is proven in Corollary 5.3 and the proof follows from showing that the difference between $\tilde{\mathcal{L}}$ and \mathcal{L} is a bounded operator.
- **Step 3:** Prove that μ is an invariant measure for (X_t, V_t) and that we can extend \mathcal{P}_t to a strongly continuous semigroup (which as an abuse of notation we still denote as \mathcal{P}_t) on L^2_μ , and we denote its generator by $(\mathcal{L}, D_{L^2_\mu})$.
- **Step 4:** We show that $\mathcal{F}C_c^\infty(\mathcal{H} \times \mathcal{H})$ is a core for $(\mathcal{L}, D_{L^2_\mu})$.

Theorem 5.2. Let $\{\tilde{\mathcal{P}}_t\}$ be the semigroup corresponding to the generator (48), assume that Φ satisfies Assumption 4.1 and that ϕ is twice continuously differentiable with $\nabla_x^2 \Phi$ bounded on bounded sets. For any $f \in C_b^1(\mathcal{H} \times \mathcal{H})$ with $\text{supp}(f) \subseteq B_r$ for some $r > 0$ then there exists a constant $C(r)$ which depends on r such that

$$\|\nabla \tilde{\mathcal{P}}_t f\|_\infty \leq \|\nabla f\|_\infty + \|f\|_\infty C(r)t^{\frac{1}{2}}. \tag{50}$$

Moreover, \mathcal{C} is a core for $(\tilde{\mathcal{L}}, D(\tilde{\mathcal{L}}))$.

Proof of Theorem 5.2. First we show that (50) holds. Fix $f \in \mathcal{C}$. In order to prove that $\tilde{\mathcal{P}}_t f$ is differentiable we will condition on the number of jumps. We can then express the law of (X_t, V_t) explicitly and differentiate this directly. For this approach we need to be able to differentiate the intensity rate λ_i and the reflection operator $R_i v$. For this reason we replaced λ_i with the differentiable rate $\tilde{\lambda}_i$. Since R_i and \mathfrak{X} are linear operators we have that $(x, v) \mapsto R_i v$ and $(x, v) \mapsto \varphi_i(x, v)$ are both differentiable. Define the operator \mathcal{B}_i to be the reflection operation and \mathcal{B}'_i as the derivative of \mathcal{B}_i , that is,

$$\begin{aligned} \mathcal{B}_i g(t, x, v) &= g(t, x, R_i v) \\ \mathcal{B}'_i &= \begin{pmatrix} 1 & 0 \\ 0 & R_i \end{pmatrix}. \end{aligned}$$

Recall we denote by T_i the i th event time. We shall set $J_i = n$ if at T_i the velocity changes according to R_n . Note that the probability of more than one event occurs at some time T_i is zero so we do not allow this event to occur. Define the processes ξ_t, C_t by

$$\xi_t = \sum_{m=1}^{N_t} C_{T_m} \nabla \log(\lambda_{J_m})(X_{T_m}, V_{T_m-}) - \sum_{i=1}^{\infty} \left(\int_0^t C_s \nabla_x \tilde{\lambda}_i(X_s, V_s) ds \right) \tag{51}$$

$$C_t = \nabla \varphi_{T_1}(x, v) \mathcal{B}'_{J_1} \cdots \nabla \varphi_{T_{N_t}-T_{N_t-1}}(X_{T_{N_t-1}}, V_{T_{N_t-1}}) \mathcal{B}'_{J_{N_t}} \nabla \varphi_{t-T_{N_t}}(X_{T_{N_t}}, V_{T_{N_t}}) \tag{52}$$

and note that C_t is a random orthogonal matrix. Note here if $t < T_1$ then we set $C_t = \nabla \varphi_t(x, v)$. With this notation we can write the derivative of the semigroup as

$$\nabla \tilde{\mathcal{P}}_t f(x, v) = \mathbb{E}[C_t \nabla f(X_t, V_t)] + \mathbb{E}[\xi_t f(X_t, V_t)]. \tag{53}$$

The proof of (53) is deferred to Proposition 8.1. Observe that since C_t is an orthogonal matrix we can bound the first term by the supremum-norm of ∇f . Therefore to show (50) it remains to find a bound for ξ_t , i.e. it suffices to show

$$\sup_{x, v \in B_r} \mathbb{E}[\|\xi_t\|_{\mathcal{H} \times \mathcal{H}}^2] \leq C(r)^2 t \tag{54}$$

where $r > 0$ is such that $\text{supp}(f) \subseteq B_r$. The proof of this bound is deferred to Proposition 8.2.

In order to prove \mathcal{C} is a core for $(\tilde{\mathcal{L}}, D(\tilde{\mathcal{L}}))$ it is sufficient to show that the semigroup $\tilde{\mathcal{P}}_t$ preserves \mathcal{C} then by [21, Proposition 3.3] we have that \mathcal{C} is a core for $\tilde{\mathcal{L}}$. We have that $\tilde{\mathcal{P}}_t$ preserves $C_b(\mathcal{H} \times \mathcal{H})$ by Lemma 2.5, and since there are no refreshments the process preserves the norm $\|\cdot\|_{\mathcal{H}}$ hence the semigroup also preserves the property of having bounded support. It remains to show that $\tilde{\mathcal{P}}_t f \in C_b^1(\mathcal{H} \times \mathcal{H})$ for any $f \in \mathcal{C}$. From (53), C_t is orthogonal, the support of f is contained in B_r , and Cauchy–Schwarz we have

$$\|\nabla \tilde{\mathcal{P}}_t f\|_\infty \leq \|\nabla f\|_\infty + \|f\|_\infty \sup_{x, v \in B_r} \mathbb{E}_{x, v}[\|\xi_t\|_{\mathcal{H} \times \mathcal{H}}^2]^{\frac{1}{2}} \leq \|\nabla f\|_\infty + \|f\|_\infty C(r)t^{\frac{1}{2}}. \tag{55}$$

Therefore \mathcal{C} is a core for $(\tilde{\mathcal{L}}, D_{C_0})$. \square

Corollary 5.3. *Suppose the assumptions of Theorem 5.2 hold. Then \mathcal{C} is a core for (\mathcal{L}, D_{C_0}) .*

Proof. Note that

$$\begin{aligned}
 (\tilde{\mathcal{L}} - \mathcal{L})f(x, v) &= \lambda_r \int_{\mathcal{H}} [f(x, w) - f(x, v)]v_0(dw) \\
 &\quad + \sum_{i=1}^{\infty} (\tilde{\lambda}_i(x, v) - \lambda_i(x, v))[f(x, R_i v) - f(x, v)].
 \end{aligned}$$

Observe that $(\tilde{\mathcal{L}} - \mathcal{L})f$ corresponds to a Boomerang Sampler with characteristics $(\Sigma, 0, 0, \{\tilde{\lambda}_i - \lambda_i\}_i, 0, \{R_i\}_i)$. Note that this is a pure jump process with bounded rates and hence corresponds to a strongly continuous contraction semigroup in $C_b(\mathcal{H} \times \mathcal{H})$. Indeed this follows from Theorem 2.7: since there is no deterministic motion we have that $C_b(\mathcal{H} \times \mathcal{H})$ is contained within the domain of the extended generator. It is then straightforward to check that the semigroup corresponding to a Boomerang Sampler with characteristics $(\Sigma, 0, 0, \{\tilde{\lambda}_i - \lambda_i\}_i, 0, \{R_i\}_i)$ is strongly continuous and contractive in $C_b(\mathcal{H} \times \mathcal{H})$. In particular, by the Hille–Yosida Theorem $\tilde{\mathcal{L}} - \mathcal{L}$ is dissipative in $C_b(\mathcal{H} \times \mathcal{H})$. Now since $\tilde{\mathcal{L}} - \mathcal{L}$ is a bounded and dissipative operator in $C_b(\mathcal{H} \times \mathcal{H})$, by [20, Theorem 2.7] we have that $(\tilde{\mathcal{L}}, D_{C_0})$ generates a contraction semigroup on the space $C_0(\mathcal{H} \times \mathcal{H})$. Since $\mathcal{C} \subseteq D(\tilde{\mathcal{L}})$ is a core for $(\tilde{\mathcal{L}}, D(\tilde{\mathcal{L}}))$ and $\mathcal{C} \subseteq D_{C_0}$ then using that \mathcal{L} is closed we see that $D(\tilde{\mathcal{L}}) \subseteq D(\mathcal{L})$. Therefore the semigroup generated by $(\tilde{\mathcal{L}}, D_{C_0})$ is an extension of the semigroup generated by $(\tilde{\mathcal{L}}, D(\tilde{\mathcal{L}}))$ and therefore must coincide and in particular we have $D(\mathcal{L}) = D(\tilde{\mathcal{L}})$. Therefore we have that \mathcal{C} is a core for (\mathcal{L}, D_{C_0}) . \square

Corollary 5.4. *Suppose that the assumptions of Theorem 5.2 hold, the assumptions of Theorem 3.1 hold then μ is an invariant measure for \mathcal{P}_t . Moreover, there is an extension of \mathcal{P}_t to L^2_{μ} which is strongly continuous and has generator given by $(\mathcal{L}, D_{L^2_{\mu}})$.*

Proof of Corollary 5.4. By Proposition 4.17 if Hypothesis 4.16 holds then (13) for all $f \in D_{C_0} \cap D(L_{\mathfrak{X}})$, therefore it remains to show that (X_t, V_t) satisfies Hypothesis 4.16 and that $\mathcal{C} \subseteq D(L_{\mathfrak{X}})$. Hypothesis 4.16 is verified in Lemma 5.1. To see that $\mathcal{C} \subseteq D(L_{\mathfrak{X}})$, fix $f \in \mathcal{C}$ and observe that by Taylor’s theorem for any $t > 0$ there exists some $s \in (0, t)$

$$\begin{aligned}
 \frac{1}{t}(f(e^{t\mathfrak{X}}x) - f(x)) - L_{\mathfrak{X}}f(x) &= \frac{\partial}{\partial t}f(e^{t\mathfrak{X}}x)|_{t=s} - L_{\mathfrak{X}}f(x) \\
 &= \langle \mathfrak{X}(e^{s\mathfrak{X}}x), \nabla f(e^{s\mathfrak{X}}x) \rangle - \langle \mathfrak{X}(x), \nabla f(x) \rangle
 \end{aligned}$$

Note that the right hand side is bounded and converges to zero as $t \rightarrow 0$ therefore by the dominated convergence theorem

$$\lim_{t \rightarrow 0} \left\| \frac{1}{t}(f(e^{t\mathfrak{X}}x) - f(x)) - L_{\mathfrak{X}}f(x) \right\|_{L^2_{\mu_0}} = 0.$$

Hence μ is an invariant measure for \mathcal{P}_t .

To see that \mathcal{P}_t extends to a C_0 -semigroup on L^2_{μ} follows from the proof of [13, Theorem 5.8]. The only difference is that [13, Theorem 5.8] is stated for semigroups defined on $C_b(\mathcal{H} \times \mathcal{H})$. However it is clear that the proof still holds for semigroups defined on $C_0(\mathcal{H} \times \mathcal{H})$ provided $C_0(\mathcal{H} \times \mathcal{H})$ is dense in L^2_{μ} . We show that $C_0(\mathcal{H} \times \mathcal{H})$ by showing that $\mathcal{C} \subseteq D_{C_0}$. Fix $f \in \mathcal{C}$ then by Theorem 2.7 we have

$$\mathcal{P}_t f(x, v) - f(x, v) = \int_0^t \mathcal{P}_s \mathcal{L} f(x, v) ds.$$

Therefore to show that $\mathcal{P}_t f - f$ converges to zero as $t \rightarrow 0$ it remains to show $\mathcal{L}f(x, v)$ is bounded uniformly in (x, v) . Let $K > 0$ be such that $\text{supp}(f) \subseteq B_K$ then

$$\|\mathcal{L}f\|_\infty \leq 2K\|\nabla f\|_\infty + 2\lambda_r\|f\|_\infty + 2\|f\|_\infty \sup_{(x,v) \in B_K} \sum_{n=1}^\infty \lambda_n(x, v).$$

By Hypothesis 4.13 which holds by Lemma 5.1 we have that $\sum \lambda_n$ is bounded on B_K so we have that $\|\mathcal{L}f\|_\infty < \infty$. Therefore $\mathcal{C} \subseteq C_0(\mathcal{H} \times \mathcal{H})$. Since \mathcal{C} is dense in L^2_μ we have that $C_0(\mathcal{H} \times \mathcal{H})$ is dense in L^2_μ . As $\{\mathcal{P}_t\}_{t \geq 0}$ is strongly continuous on L^2_μ it has a generator given by $(\mathcal{L}, D_{L^2_\mu})$ which is an extension of (\mathcal{L}, D_{C_0}) . \square

Theorem 5.5. *Suppose that the assumptions of Corollary 5.4 hold then $\mathcal{F}C_c^\infty(\mathcal{H} \times \mathcal{H})$ is a core for $(\mathcal{L}, D_{L^2_\mu})$.*

Proof of Theorem 5.5. Note that $\mathcal{F}C_b^\infty(\mathcal{H})$ and $\mathcal{F}C_c^\infty(\mathcal{H})$ are both dense in L^2_μ (and $L^4_{\mu_0}$), see [26, Lemma 2.2].

$$e^{tL_{\mathfrak{X}}} f(x, v) = f(\varphi_t(x, v))$$

Since $e^{tL_{\mathfrak{X}}}$ preserves the set $\mathcal{F}C_c^\infty(\mathcal{H} \times \mathcal{H})$ we have by [21, Proposition 3.3] that this set is a core for $L_{\mathfrak{X}}$.

Fix $f \in \mathcal{C}$, it is immediate to check that $f \in D_{L^4_{\mu_0}}(L_{\mathfrak{X}})$ so there exists $f_n \in \mathcal{F}C_c^\infty(\mathcal{H} \times \mathcal{H})$ such that $f_n \rightarrow f$ and $L_{\mathfrak{X}} f_n \rightarrow L_{\mathfrak{X}} f$ in $L^4_{\mu_0}$. As Φ is bounded from below we also have that $f_n \rightarrow f$ and $L_{\mathfrak{X}} f_n \rightarrow L_{\mathfrak{X}} f$ in L^2_μ and that $f_n \rightarrow f$ in L^4_μ . Now

$$\begin{aligned} \|\mathcal{L}f - \mathcal{L}f_n\|_{L^2_\mu} &\leq \|L_{\mathfrak{X}} f - L_{\mathfrak{X}} f_n\|_{L^2_\mu} + \left\| \sum_{i=1}^\infty \lambda_i(f(\cdot, R_i \cdot) - f_n(\cdot, R_i \cdot)) \right\|_{L^2_\mu} \\ &\quad + \left\| \sum_{i=1}^\infty \lambda_i(f - f_n) \right\|_{L^2_\mu} + 2\lambda_r\|f - f_n\|_{L^4_\mu}. \end{aligned}$$

Using Cauchy–Schwarz, and that μ_0 is invariant under R_i we can bound this by

$$\|\mathcal{L}f - \mathcal{L}f_n\|_{L^2_\mu} \leq \|L_{\mathfrak{X}} f - L_{\mathfrak{X}} f_n\|_{L^2_\mu} + 2\left\| \sum_{i=1}^\infty \lambda_i \right\|_{L^4_\mu}^{\frac{1}{2}} \|f - f_n\|_{L^4_\mu}^{\frac{1}{2}} + 2\lambda_r\|f - f_n\|_{L^4_\mu}.$$

Note here $\lambda \in L^4_\mu$ by (100). Letting n tend to ∞ we have that $\mathcal{L}f_n \rightarrow \mathcal{L}f$ in L^2_μ . That is the closure $\overline{\mathcal{F}C_c^\infty(\mathcal{H} \times \mathcal{H})}$ of $\mathcal{F}C_c^\infty(\mathcal{H} \times \mathcal{H})$ in the graph norm of \mathcal{L} with respect to the L^2_μ topology contains \mathcal{C} . Now \mathcal{C} is a core for (\mathcal{L}, D_{C_0}) and since convergence with the supremum norm implies convergence in L^2_μ -norm we have that $D_{C_0} \subseteq \overline{\mathcal{F}C_c^\infty(\mathcal{H} \times \mathcal{H})}$. Again by [21, Proposition 3.3] we have that D_{C_0} is a core for $(\mathcal{L}, D_{L^2_\mu})$ and hence $\mathcal{F}C_c^\infty(\mathcal{H} \times \mathcal{H})$ is a core for $(\mathcal{L}, D_{L^2_\mu})$. \square

6. Exponential convergence to equilibria

In this section we will employ the strategy of [2,16] to give conditions under which we have exponential convergence to equilibria. Let (X_t, V_t) be the Boomerang Sampler with characteristics $(\Sigma, 1, 1, \{\lambda_i\}_i, \lambda_r, \{R_i\}_i)$ where λ_i are given by (44), and let R_i satisfy the following conditions.

Hypothesis 6.1. Suppose R_i is a linear operator on \mathcal{H} , with $R_n^2 = 1$, $\|R_n v\| = \|v\|$ for any $v \in \mathcal{H}$, satisfies (42) and (45). Moreover suppose that for any $v \in \mathcal{H}$

$$\sum_{n=1}^{\infty} (1 - R_n)v = 2v, \tag{56}$$

$$\frac{1}{4} \sum_{n=1}^{\infty} (1 - R_n)^* \Sigma (1 - R_n) = \Sigma, \tag{57}$$

$$(1 - R_n)^* \Sigma (1 - R_m) = 0 \text{ for } m \neq n, \tag{58}$$

Here $(1 - R_n)^*$ denotes the adjoint of $(1 - R_n)$ as a linear operator from \mathcal{H} to \mathcal{H} .

For simplicity we have canonical rates, that is we have set $\gamma_i = 0$ where γ_i is as in Sections 4.3.3 and 4.3.4. Note that if Φ satisfies Assumption 4.1 then Hypothesis 4.16 is satisfied and hence Hypothesis 4.13 also, see Lemma 5.1.

Example 6.1 (Pure Reflection Boomerang Sampler). We can express the Pure Reflection Boomerang Sampler in the above notation by setting $R_1 = -1$, $R_n = 1$ for any $n > 1$. \triangle

Example 6.2 (Factorised Boomerang Sampler). Alternatively we can also express the Factorised Boomerang Sampler by taking $R_n v = v - 2\langle v, e_n \rangle_{\mathcal{H}} e_n$. Indeed in this case (58) follows immediately as e_n, e_m are orthogonal for $n \neq m$ and diagonalise Σ , moreover we have

$$(1 - R_n)^* \Sigma (1 - R_m)v = 4v_n \gamma_n^2 e_n \delta_{n,m}.$$

Taking $n = m$ and summing over n gives

$$\sum_{n=1}^{\infty} (1 - R_n)^* \Sigma (1 - R_n)v = 4 \sum_{n=1}^{\infty} v_n e_n \gamma_n^2 = 4 \Sigma v.$$

Finally checking (56)

$$\sum_{n=1}^{\infty} (1 - R_n)v = 2 \sum_{n=1}^{\infty} v_n e_n = 2v. \quad \triangle$$

The conditions we will require for exponential convergence can be expressed in terms of the second order differential operator defined on smooth functions $g : \mathcal{H} \rightarrow \mathbb{R}$ by

$$\mathcal{A}g(x) = \text{Tr}(\Sigma \nabla_x^2 g(x)) - \langle x, \nabla_x g(x) \rangle_{\mathcal{H}} - \langle \nabla_x \Phi(x), \Sigma \nabla_x g(x) \rangle_{\mathcal{H}}. \tag{59}$$

Note that this is an operator only on the x variable.

Theorem 6.2. Let \mathcal{P}_t be the semigroup corresponding to the Boomerang Sampler which is described by the generator (46). Assume that Assumption 4.1 is satisfied, and that Φ is twice continuously differentiable with bounded Hessian, and there exist $c_2 \geq 0, c_1 > 0$ such that

$$\mathcal{A}\Phi(x) \leq c_2 - c_1 \|\Sigma^{\frac{1}{2}} \nabla_x \Phi(x)\|_{\mathcal{H}}^2. \tag{60}$$

Also assume that $e^{-\Phi} \in L^1(\pi_0)$ where $\pi_0 = \nu_0 = \mathcal{N}(0, \Sigma)$, and set π to be the measure defined by (1) and $\mu = \pi \times \nu_0$. Then there exist constants $C, \kappa > 0$ such that

$$\|\mathcal{P}_t f - \mu(f)\|_{L^2_\mu} \leq C e^{-\kappa t} \|f - \mu(f)\|_{L^2_\mu} \tag{61}$$

for all $f \in L^2_\mu$.

Corollary 6.3. *Let Φ be a twice continuously differentiable function satisfying Assumption 4.1, has bounded Hessian and is such that $\langle \nabla_x \Phi(x), x \rangle_{\mathcal{H}} \geq C$ for some $C \in \mathbb{R}$. Then Φ satisfies the assumptions of Theorem 6.2. In particular, this is satisfied if Φ is convex, bounded from below and has bounded Hessian.*

Proof of Corollary 6.3. To show that (60) holds, observe that

$$A\Phi(x) = \text{Tr}(\Sigma \nabla_x^2 \Phi(x)) - \langle x, \nabla_x \Phi(x) \rangle_{\mathcal{H}} - \langle \nabla_x \Phi(x), \Sigma \nabla_x \Phi(x) \rangle_{\mathcal{H}}.$$

since Φ has bounded Hessian we have that $\text{Tr}(\Sigma \nabla_x^2 \Phi(x))$ is bounded so it is sufficient that $\langle x, \nabla_x \Phi(x) \rangle_{\mathcal{H}} \geq C$ for some $C \in \mathbb{R}$. We now show this holds if Φ is convex, and bounded from below. Fix $x \in \mathcal{H}$, by Taylor’s theorem there exists $z(x)$ such that

$$\Phi(0) = \Phi(x) - \langle \nabla_x \Phi(x), x \rangle_{\mathcal{H}} + \frac{1}{2} \langle \nabla_x^2 \Phi(z(x))x, x \rangle_{\mathcal{H}}.$$

Using that Φ is convex

$$\Phi(0) \geq \Phi(x) - \langle \nabla_x \Phi(x), x \rangle_{\mathcal{H}}.$$

Finally rearranging we have

$$\langle \nabla_x \Phi(x), x \rangle_{\mathcal{H}} \geq \Phi(x) - \Phi(0)$$

which is bounded from below as Φ is bounded from below. \square

Example 6.3. Fix $T > 0$ and consider the one-dimensional SDE with initial and terminal condition

$$\begin{cases} dY_t = b(Y_t)dt + dW_t, & t \in (0, T) \\ Y_0 = Y_T = 0. \end{cases} \tag{62}$$

Here W_t is a one-dimensional Brownian motion, $b : \mathbb{R} \rightarrow \mathbb{R}$ is a smooth function and σ is some positive constant. We shall assume that $b \in C_b^3(\mathbb{R})$, i.e. that b is bounded and has bounded first and second order derivatives. Denote by π the law of Y defined on the space $L^2([0, T])$, the aim is to sample from π . We shall do so by using a Girsanov transformation to write the measure π as a Radon–Nikodym derivative of the measure π_0 , where π_0 is the law of the one-dimensional Brownian Bridge on $[0, T]$. Set $\mathcal{H} = L^2([0, T])$ and define the operator

$$\begin{cases} Ax(\xi) = \frac{1}{2}x''(\xi), & x \in D(A), \\ D(A) = H^2(0, 1) \cap H_0^1(0, 1), \end{cases} \tag{63}$$

where H^k are the usual Sobolev spaces and

$$H_0^1(0, T) = \{x \in H^1(0, T) : x(0) = x(T) = 0\}.$$

Let X_t be a one-dimensional Brownian bridge, then set

$$\mathcal{E}_t = \exp \left(\int_0^t b(X_s)dX_s - \frac{1}{2} \int_0^t b(X_s)^2 ds \right).$$

By Girsanov’s theorem the process

$$W_t := X_t - \int_0^t b(X_s)ds$$

is a Brownian motion under π where $\frac{d\pi}{d\pi_0} = \mathcal{E}_T$. Note that by rearranging we have that X_t solves (62) under π . Now let $B = \int_0^t b(y)dy$, then by Itô’s formula we have

$$B(X_T) - B(X_0) - \frac{1}{2} \int_0^T b'(X_s)ds = \int_0^T b(X_s)dX_s.$$

Recall that $X_T = X_0 = 0$ so we can now write \mathcal{E}_T as

$$\mathcal{E}_T = \exp\left(-\frac{1}{2} \int_0^T b'(X_s)ds - \frac{1}{2} \int_0^T b(X_s)^2 ds\right).$$

Therefore we have that

$$\begin{aligned} \frac{d\pi}{d\pi_0}(x) &= \exp(-\Phi(x)) \\ \Phi(x) &= \frac{1}{2} \int_0^T b'(X_s)ds + \frac{1}{2} \int_0^T b(X_s)^2 ds \end{aligned}$$

Note that Φ is smooth and

$$\begin{aligned} \nabla_x \Phi(x)(t) &= \frac{1}{2} b''(x_t) + b(x_t)b'(x_t) \\ \nabla_x^2 \Phi(x)(t) &= \frac{1}{2} b'''(x_t) + b(x_t)b''(x_t) + 2b'(x_t)^2. \end{aligned}$$

As $b \in C_b^3(\mathbb{R})$ we have that Φ is bounded with bounded first and second order derivatives, therefore by Remark 6.10 all the assumptions of Theorem 6.2 are satisfied so the Boomerang Sampler converges exponentially in the sense of (61).

To keep calculations short we are only working in 1-d, but similar arguments can be applied in higher dimensions. In [6, Equation 1.5] it is shown that if we consider the SDE

$$dX_t = f(X_t)dt + BdW_t$$

where $f = -BB^T \nabla_x V$ then the corresponding Φ is given by

$$\Phi(x) = \int_0^t \frac{1}{2} |B^{-1} f(x_t)|^2 + \frac{1}{2} (\nabla_x \cdot f)(x_t) dt. \quad \Delta$$

6.1. Hypocoercivity

We shall prove Theorem 6.2 by applying the Abstract Hypocoercivity Theorem introduced in [16]. We shall use the formulation from [19] which has been developed for infinite dimensional stochastic differential equations.

Assumption 6.4.

- (D1) Set $H = L^2_\mu$ and let $(L, D(L))$ be a linear operator on H generating a C_0 -semigroup $\{P_t\}_{t \geq 0}$. Where $\{P_t\}_{t \geq 0}$ is conservative, i.e. $P_t 1 = 1$ for all $t \geq 0$ and has invariant measure μ .
- (D2) Let $\mathcal{C} \subseteq D(L)$ be a dense subspace of H which is a core for the operator $(L, D(L))$.

- (D3) Let $(S, D(S))$ be symmetric and let $(A, D(A))$ be a closed and antisymmetric operator on H such that $\mathcal{C} \subseteq D(S) \cap D(A)$ as well as $L|_{\mathcal{C}} = S - A$.
- (D4) Let $\Pi : H \rightarrow H$ be an orthogonal projection which satisfies $\Pi(H) \subseteq D(S)$, $S\Pi = 0$ as well as $\Pi(\mathcal{C}) \subseteq D(A)$, $A\Pi(\mathcal{C}) \subseteq D(A)$.

Assumption 6.5.

- (H1) Assume that $\Pi A \Pi|_{\mathcal{C}} = 0$.
- (H2) There exists $\lambda_m > 0$ such that

$$-\langle Sf, f \rangle_H \geq \lambda_m \|(1 - \Pi)f\|^2 \text{ for all } f \in \mathcal{C}.$$

- (H3) Define $G = \Pi A^2 \Pi$ on \mathcal{C} . Assume that (G, D) is essentially self-adjoint on H and assume that there exists λ_M such that

$$\|\Pi A \Pi f\|_H^2 \geq \lambda_M \|\Pi f\|_H^2 \text{ for all } f \in \mathcal{C}.$$

- (H4) Define B to be the unique extension to a bounded linear operator on H of the operator $(B, D(A\Pi)^*)$ where

$$B = (I + (A\Pi)^*(A\Pi))^{-1}(A\Pi)^* \text{ on } D((A\Pi)^*).$$

Assume that the operators (BS, \mathcal{C}) and $(BA(1 - \Pi), \mathcal{C})$ are bounded and there exist constants $c_1, c_2 < \infty$ such that for all $f \in \mathcal{C}$

$$\|BSf\|_H \leq c_1 \|(1 - \Pi)f\|_H, \quad \text{and} \quad \|BA(1 - \Pi)f\|_H \leq c_2 \|(1 - \Pi)f\|_H.$$

Theorem 6.6. Assume that Assumptions 6.4 and 6.5 hold. Then there exist positive constants $C, \kappa > 0$ such that for any function $f \in L^2_\mu$, and $t \geq 0$,

$$\|P_t f - \mu(f)\|_{\mathcal{H}} \leq C e^{-\kappa t} \|f - \mu(f)\|_{\mathcal{H}}$$

where $\mu(f) = \int f d\mu$.

Then Theorem 6.2 follows from Theorem 6.6 once we verify these conditions for the Boomerang Sampler. Before we verify these conditions we state the following lemma which is quoted from [19, Lemma 1].

Lemma 6.7. For the Gaussian measure ν_0 and $v_1, v_2, v_3, v_4 \in \mathcal{H}$ we have that

$$\begin{aligned} \int_{\mathcal{H}} \langle u, v_1 \rangle_{\mathcal{H}} \langle u, v_2 \rangle_{\mathcal{H}} \nu_0(du) &= \langle \Sigma v_1, v_2 \rangle_{\mathcal{H}}, \\ \int_{\mathcal{H}} \langle u, v_1 \rangle_{\mathcal{H}} \langle u, v_2 \rangle_{\mathcal{H}} \langle u, v_3 \rangle_{\mathcal{H}} \langle u, v_4 \rangle_{\mathcal{H}} \nu_0(du) &= \langle \Sigma v_1, v_2 \rangle_{\mathcal{H}} \langle \Sigma v_3, v_4 \rangle_{\mathcal{H}} \\ &\quad + \langle \Sigma v_1, v_3 \rangle_{\mathcal{H}} \langle \Sigma v_2, v_4 \rangle_{\mathcal{H}} \\ &\quad + \langle \Sigma v_1, v_4 \rangle_{\mathcal{H}} \langle \Sigma v_2, v_3 \rangle_{\mathcal{H}}. \end{aligned}$$

Proof of Theorem 6.2. To simplify notation we will only consider f such that $\mu(f) = 0$, the result then follows by setting $\tilde{f} = f - \mu(f)$. Let us first consider Assumption 6.4, note that (D1) is immediate, and (D2) follows from Theorem 5.5.

Define $(S, D(S))$ and $(A, D(A))$ as the closure of the operators $(S, \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H}))$, $(A, \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H}))$ where

$$Sf(x, v) := \frac{1}{2}(\mathcal{L} + \mathcal{L}^*)f(x, v) = \sum_{n=1}^\infty \lambda_n^e(x, v)[f(x, R_n v) - f(x, v)] + L_r f(x, v),$$

$$f \in \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H}) \tag{64}$$

$$Af(x, v) := \frac{1}{2}(\mathcal{L} - \mathcal{L}^*)f(x, v) = L_x f(x, v) + \sum_{n=1}^\infty \lambda_n^o(x, v)[f(x, R_n v) - f(x, v)],$$

$$f \in \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H}). \tag{65}$$

Here

$$\lambda_n^e(x, v) = \frac{1}{2}(\lambda_n(x, v) + \lambda_n(x, R_n v)) = \frac{1}{4}|\langle \nabla_x \Phi(x), v - R_n v \rangle_{\mathcal{H}}| + \gamma_n(x, v), \tag{66}$$

$$\lambda_n^o(x, v) = \frac{1}{2}(\lambda_n(x, v) - \lambda_n(x, R_n v)) = \frac{1}{4}\langle \nabla_x \Phi(x), v - R_n v \rangle_{\mathcal{H}}. \tag{67}$$

For $f \in L_\mu^2$ set

$$\Pi(f)(x) = \int_{\mathcal{H}} f(x, v) \nu_0(dv). \tag{68}$$

Note that since $(S, \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H}))$ and $(A, \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H}))$ are the restriction of the operators $(S^*, D(S^*))$ and $(-A^*, D(A^*))$ to the set $\mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H})$ which is dense we have that $(S, \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H}))$ and $(A, \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H}))$ are closable. It is immediate to verify that Π is an orthogonal projection on L_μ^2 therefore all of these operators are well-defined and we have that **(D3)** is satisfied.

Let us now verify **(D4)**. Fix $g \in \Pi(L_\mu^2)$ and take $g_n \in \mathcal{FC}_c^\infty(\mathcal{H} \times \mathcal{H})$ such that $g_n \rightarrow g$ which exists since $\mathcal{FC}_c^\infty(\mathcal{H} \times \mathcal{H})$ is dense in L_μ^2 . Then $\Pi g_n \in \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H})$, and we have $S(g_n) = 0$. Since $(S, D(S))$ is closed we have that $\Pi(L_\mu^2) \subseteq D(S)$ and $S\Pi = 0$. Fix $f \in \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H})$ then $\Pi f \in \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H}) \subseteq D(A)$ and

$$A\Pi f = \langle v, \nabla_x \Pi f(x, v) \rangle_{\mathcal{H}}. \tag{69}$$

We can establish that $A\Pi f \in D(A)$ by following the same argument as in [19, Lemma 6]. Since the details are the same we will simply sketch the strategy here. It is sufficient to construct a sequence $f_m \in \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H})$ such that $f_m \rightarrow A\Pi f$ in L_μ^2 and Af_m is convergent in L_μ^2 . Since $f \in \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H})$ there is $n \in \mathbb{N}$ and a projection P_n such that f only depends on the variables $(P_n(x), P_n(v))$, we can construct the sequence $\{f_m\}_m$ by setting $f_m(x, v) = \varphi_m^n(P_n(v))A\Pi f(x, v)$ for a suitable cut-off function $\varphi_m^n \in C_c^\infty(\mathbb{R}^n)$. This sequence of functions converge pointwise to $A\Pi f$ and one can show convergence in L_μ^2 by the Dominated Convergence Theorem.

Let us now verify **Assumption 6.5**. Note that **Assumption 6.5(H1)** is verified immediately. **Assumption 6.5(H2)** follows from **Lemma 6.8**.

In order to verify **Assumption 6.5(H3)** we must first derive an expression for the operator G . Fix $f \in \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H})$ and let $g = \Pi f$ then we by applying A to (69) we have

$$(A^2g)(x, v) = \langle v, \nabla_x \langle v, \nabla_x g(x) \rangle_{\mathcal{H}} \rangle_{\mathcal{H}} - \langle x, \nabla_x g(x) \rangle_{\mathcal{H}} - \sum_{n=1}^{\infty} \lambda_n^o(x, v) \langle (1 - R_n)v, \nabla_x g(x) \rangle_{\mathcal{H}} \tag{70}$$

$$= \langle v, \nabla_x^2 g(x)v \rangle_{\mathcal{H}} - \langle x, \nabla_x g(x) \rangle_{\mathcal{H}} - \sum_{n=1}^{\infty} \lambda_n^o(x, v) \langle (1 - R_n)v, \nabla_x g(x) \rangle_{\mathcal{H}} \tag{71}$$

Using (67) and **Lemma 6.7** we have

$$\begin{aligned} \Pi[\lambda_n^o(x, v) \langle (1 - R_n)v, \nabla_x \Pi f(x) \rangle_{\mathcal{H}}] &= \frac{1}{4} \Pi[\langle \nabla_x \Phi, (1 - R_n)v \rangle_{\mathcal{H}} \\ &\quad \times \langle (1 - R_n)v, \nabla_x \Pi f(x) \rangle_{\mathcal{H}}] \\ &= \frac{1}{4} \langle (1 - R_n) \nabla_x \Phi(x), \Sigma (1 - R_n) \nabla_x \Pi f(x) \rangle_{\mathcal{H}}. \end{aligned}$$

Here we have used that Π denotes integration with respect to ν_0 and ν_0 has mean zero and covariance operator Σ . Summing over n and using (57) we have

$$\sum_{n=1}^{\infty} \Pi[\lambda_n^o(x, v) \langle (1 - R_n)v, \nabla_x \Pi f(x) \rangle_{\mathcal{H}}] = \langle \nabla_x \Phi(x), \Sigma \nabla_x \Pi f(x) \rangle_{\mathcal{H}}.$$

By applying Π to the first term in (71) we find

$$\Pi \langle v, \nabla_x^2 \Pi f(x)v \rangle_{\mathcal{H}} = \text{Tr}(\Sigma \nabla_x^2 \Pi f(x)).$$

Therefore

$$\Pi A^2 \Pi f = \text{Tr}(\Sigma \nabla_x^2 \Pi f(x)) - \langle x, \nabla_x \Pi f(x) \rangle_{\mathcal{H}} - \langle \nabla_x \Phi(x), \Sigma \nabla_x \Pi f(x) \rangle_{\mathcal{H}}.$$

Therefore, we have $Gf = \Pi T^2 \Pi f = \mathcal{A}g$ for any $f \in \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H})$.

By [19, Theorem 1] G is essentially m -dissipative in L_μ^2 and hence essentially self-adjoint in L_μ^2 . From [19, Proposition 4] we have that $(G, \mathcal{FC}_b^\infty(\mathcal{H}))$ satisfies the Poincaré inequality for all $g \in \mathcal{FC}_b^\infty(\mathcal{H})$

$$\int_{\mathcal{H}} (g(x) - \pi(g))^2 \pi(dx) \leq \lambda_1^{-1} \int_{\mathcal{H}} \langle \Sigma \nabla_x g, \nabla_x g \rangle_{\mathcal{H}} \pi(dx) = -\langle \mathcal{A}g, g \rangle_{L_\pi^2}.$$

Note that [19] requires that the function Φ is convex however if $\Phi = \Phi_1 + \Phi_2$ with Φ_1 convex and Φ_2 bounded then the Poincaré inequality still holds, this can be seen by first applying [19, Proposition 4] for Φ_1 and then by the same argument as in [4, Proposition 4.2.7] the Poincaré inequality holds for Φ .

Therefore taking $g = \Pi f$ for $f \in L_\mu^2$ we have that **Assumption 6.5(H3)** holds.

As explained in [23, Remark 2.17] we can replace **Assumption 6.5(H4)** with

$$\langle (B(S - A(1 - \Pi))f, \Pi f) \rangle \leq c_3 \| (1 - \Pi)f \|^2 \|\Pi f\| \tag{72}$$

for some constant $c_3 \geq 0$ and for all $f \in \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H})$. Therefore it suffices to prove that (72) holds which we defer to **Proposition 6.11**. \square

Lemma 6.8. Assume that the assumptions of [Theorem 6.2](#) hold and let S and Π be the operators defined by [\(64\)](#) and [\(68\)](#) respectively. Then for any $f \in D(S)$ we have

$$-\langle Sf, f \rangle_{L^2_\mu} \geq \lambda_r \|(1 - \Pi)f\|_{L^2_\mu}^2$$

In particular, we have that [Assumption 6.5\(H2\)](#) of [Theorem 6.6](#) holds.

Proof of Lemma 6.8. Note that

$$-\langle Sf, f \rangle_{L^2_\mu} = - \sum_{n=1}^\infty \int_{\mathcal{H}} \int_{\mathcal{H}} (\lambda_n^e(x, v)[f(x, R_n v) - f(x, v)] + L_r f(x, v)) f(x, v) \nu_0(dv) \pi(dx).$$

Since ν_0 and λ_n^e are both invariant under R_n we have

$$\begin{aligned} \int_{\mathcal{H}} \lambda_n^e(x, v)[f(x, R_n v) - f(x, v)]^2 \nu_0(dv) &= \int_{\mathcal{H}} 2\lambda_n^e(x, v) f(x, v)^2 \\ &\quad - 2\lambda_n^e(x, v) f(x, R_n v) f(x, v) \nu_0(dv) \\ &= -2 \int_{\mathcal{H}} \lambda_n^e(x, v)[f(x, R_n v) - f(x, v)] \\ &\quad \times f(x, v) \nu_0(dv). \end{aligned}$$

Therefore we have for any $n \in \mathbb{N}$

$$\langle \lambda^e(x, v)[f(x, R_n v) - f(x, v)], f \rangle_{L^2_\mu} \leq 0.$$

Which gives

$$\begin{aligned} -\langle Sf, f \rangle_{L^2_\mu} &\geq -\langle L_r f, f \rangle_{L^2_\mu} \\ &= \lambda_r \langle (1 - \Pi)f, f \rangle_{L^2_\mu} \end{aligned}$$

The result follows since Π is an orthogonal projection. \square

Proposition 6.9. Let \mathcal{P}_t be the semigroup corresponding to the Boomerang Sampler which is described by the generator [\(46\)](#). Assume that Φ is twice continuously differentiable with bounded Hessian and there exist $c_2 \geq 0, c_1 > 0$ such that [\(60\)](#) holds. Also assume that $e^{-\Phi} \in L^1(\pi_0)$ where $\pi_0 = \nu_0 = \mathcal{N}(0, \Sigma)$, and set π to be the measure defined by [\(1\)](#) and $\mu = \pi \times \nu_0$. Fix $g \in \mathcal{FC}_b^\infty(\mathcal{H})$ and set $h = g - \mathcal{A}g$ then we have

$$\|\Sigma^{\frac{1}{2}} \nabla_x g\|_{L^2_\pi} \leq \|h\|_{L^2_\pi}, \tag{73}$$

$$\|\nabla_x \Sigma \nabla_x g\|_{L^2_\pi} \leq \kappa_1 \|h\|_{L^2_\pi}, \tag{74}$$

$$\left\| \langle \Sigma^{\frac{1}{2}} \nabla_x g, \Sigma^{\frac{1}{2}} \nabla_x \Phi \rangle_{\mathcal{H}} \right\|_{L^2_\pi} \leq \left\| \|\Sigma^{\frac{1}{2}} \nabla_x g\|_{\mathcal{H}} \Sigma^{\frac{1}{2}} \nabla_x \Phi \right\|_{L^2_\pi(\mathcal{H}; \mathcal{H})} \leq \kappa_2 \|h\|_{L^2_\pi}. \tag{75}$$

Here $\kappa_1^2 = 2 + c_\Phi, -c_\Phi \leq \nabla_x^2 \Phi(x)$ for all $x \in \mathcal{H}$ and $\kappa_2^2 = \frac{4\kappa_1^2}{c_1^2} + \frac{2c_2}{c_1}$.

We defer the proof of [Proposition 6.9](#) to [Section 8.3](#).

Remark 6.10. In the context of [Proposition 6.9](#) if we assume that $\nabla_x \Phi$ is bounded then [\(75\)](#) follows immediately from [\(73\)](#). Therefore we can remove the assumption that [\(60\)](#) holds in [Theorem 6.2](#) if $\nabla_x \Phi$ is bounded. \triangle

With these estimates to hand we can show that Assumption 6.5(H4) of Theorem 6.6 holds. These are based on [2, Lemma 12 & 13].

Proposition 6.11. *Let S, A, Π be defined by (64), (65) and (68). Define B as in Theorem 6.6. Assume that estimates (73), (74) and (75) hold then we have for all $f \in \mathcal{FC}_b^\infty(\mathcal{H} \times \mathcal{H})$*

$$|\langle BSf, \Pi f \rangle_{L_\mu^2}| \leq \left(\sqrt{3}\kappa_2 + \frac{1}{\sqrt{2}}\lambda_r \right) \|(1 - \Pi)f\|_{L_\mu^2} \|f\|_{L_\mu^2}, \tag{76}$$

$$|\langle BA(1 - \Pi)f, \Pi f \rangle_{L_\mu^2}| \leq \sqrt{2}(\kappa_1 + \kappa_2) \|(1 - \Pi)f\|_{L_\mu^2} \|f\|_{L_\mu^2}. \tag{77}$$

Here $\kappa_2 = \sqrt{\frac{4\kappa_1^2}{c_1^2} + \frac{c_2}{c_1}}$ and κ_1 is defined as in Proposition 6.9.

Proof of Proposition 6.11. The proof is deferred to Section 8.3. \square

7. Finite dimensional approximation

In this section we wish to construct a finite dimensional approximation to the Boomerang Sampler. We shall achieve this by taking a finite dimensional approximation of Φ and then constructing a Boomerang sampler to sample from this measure. Let us assume we have a potential function $\Phi : \mathcal{H} \rightarrow \mathbb{R}$ continuously differentiable and measures $\pi_0, \nu_0, \pi, \mu_0, \mu$ as in Section 3.

Fix an orthonormal basis, $\{e_i\}_{i=1}^\infty$, of eigenvectors of Σ_x . Denote by Proj_N the orthogonal projection onto $\mathcal{H}_N := \text{span}\{e_1, \dots, e_N\}$, and define

$$\Phi_N(x) = \Phi(\text{Proj}_N(x)).$$

Note that the projection of π_0 (respectively ν_0) to the space \mathcal{H}_N is still a Gaussian measure as linear transformations preserve Gaussianity and is centred with covariance operator $\Sigma_{x,N} = \text{Proj}_N \circ \Sigma_x$ (resp. $\Sigma_{v,N} = \text{Proj}_N \circ \Sigma_v$). Now we can consider the problem of how to sample from the measure μ_N defined by

$$\frac{d\mu_N}{d\mu_{0,N}}(dx, dv) = \frac{\exp(-\Phi_N(x))}{\int \exp(-\Phi_N(y))\mu_{0,N}(dy, dw)}$$

where $\mu_{0,N}(dx, dv) = \pi_{0,N}(dx)\nu_{0,N}(dv)$, $\pi_{0,N} = \mathcal{N}(0, \Sigma_{x,N})$ and $\nu_{0,N} = \mathcal{N}(0, \Sigma_{v,N})$. Note that although these measures are supported on the finite dimensional space \mathcal{H}_N^2 we can also view them as measures on \mathcal{H}^2 .

Let us introduce some more notation for any $x \in \mathcal{H}$ we denote by

$$x^N := \text{Proj}_N x = \sum_{n=1}^N x_n e_n, \quad x_n := \langle x, e_n \rangle,$$

$$x_\perp^N := x - x^N = \sum_{n>N} x_n e_n,$$

and similarly we define $v^N := \text{Proj}_N v$, $v_\perp^N := v - v^N$.

Let $\{X_t, V_t\}$ denote the Boomerang Sampler whose generator is given by (46). We shall construct an N -dimensional process $\{X_t^N, V_t^N\}_{t \geq 0}$ which takes values in $\mathcal{H}_N \times \mathcal{H}_N$, although this process takes values in the space $\mathcal{H}_N \times \mathcal{H}_N$ it is convenient for the analysis to view it as a process taking values in $\mathcal{H} \times \mathcal{H}$. We first consider two examples before giving a more abstract framework, which is the pure reflection Boomerang Sampler and the Factorised Boomerang Sampler.

Example 7.1. Let $\{X_t, V_t\}_{t \geq 0}$ denote the pure reflection Boomerang Sampler whose generator is given by (46)

$$\begin{aligned} \mathcal{L}_{PR} f(x, v) &= \langle v, \nabla_x f(x, v) \rangle_{\mathcal{H}} - \langle x, \nabla_v f(x, v) \rangle_{\mathcal{H}} + \lambda_r \int_{\mathcal{H}} [f(x, w) - f(x, v)] \nu_0(dw) \\ &\quad + (\langle \nabla_x \Phi(x), v \rangle_{\mathcal{H}})^+ [f(x, -v) - f(x, v)]. \end{aligned}$$

To construct a finite dimensional approximation we need to specify the deterministic dynamics, the reflections and the refreshments. For the deterministic dynamics act on the first N components in the same way as the full process but leave all other components fixed. Reflections will only reflect the first N components instead of the entire velocity and refreshments will refresh with the measure $\nu_{0,N}$. That is the process X_t^N, V_t^N will have generator given by

$$\begin{aligned} \mathcal{L}_{PR}^N f(x, v) &= \langle v^N, \nabla_x f(x, v) \rangle_{\mathcal{H}} - \langle x^N, \nabla_v f(x, v) \rangle_{\mathcal{H}} \\ &\quad + \lambda_r \int_{\mathcal{H}} [f(x, w) - f(x, v)] \nu_{0,N}(dw) \\ &\quad + (\langle \nabla_x \Phi_N(x), v^N \rangle_{\mathcal{H}})^+ [f(x, -v^N + v_{\perp}^N) - f(x, v)]. \end{aligned}$$

It is immediate to check that Hypothesis 4.16 so this process is well posed and has μ_N as an invariant measure. Δ

Example 7.2. Let $\{X_t, V_t\}_{t \geq 0}$ denote the Factorised Boomerang Sampler whose generator is given by (46)

$$\begin{aligned} \mathcal{L}_F f(x, v) &= \langle v, \nabla_x f(x, v) \rangle_{\mathcal{H}} - \langle x, \nabla_v f(x, v) \rangle_{\mathcal{H}} + \lambda_r \int_{\mathcal{H}} [f(x, w) - f(x, v)] \nu_0(dw) \\ &\quad + \sum_{n=1}^{\infty} (\partial_{x_n} \Phi(x) v_n)^+ [f(x, R_n v) - f(x, v)]. \end{aligned}$$

Here $R_n v = v - 2\langle v, e_n \rangle_{\mathcal{H}} e_n$. The finite dimensional approximation is constructed analogously to Example 7.1 except for the reflections. In this case the first N reflections are the same as the Factorised Boomerang, that is the reflection operator R_n^N changes the sign of the n th component for $n \leq N$. All other reflections are set to the identity, that is we only have N reflection operators. The process X_t^N, V_t^N will have generator given by

$$\begin{aligned} \mathcal{L}_{PR}^N f(x, v) &= \langle v^N, \nabla_x f(x, v) \rangle_{\mathcal{H}} - \langle x^N, \nabla_v f(x, v) \rangle_{\mathcal{H}} \\ &\quad + \lambda_r \int_{\mathcal{H}} [f(x, w) - f(x, v)] \nu_{0,N}(dw) \\ &\quad + \sum_{n=1}^N (\partial_{x_n} \Phi_N(x) v_n)^+ [f(x, R_n v) - f(x, v)]. \end{aligned}$$

It is immediate to check that Hypothesis 4.16 is satisfied so that this process is well posed and has μ_N as an invariant measure. Δ

Now we introduce a framework which will allow us to tackle both the situations of Examples 7.1 and 7.2 simultaneously. Let $\{X_t, V_t\}_{t \geq 0}$ denote the Boomerang Sampler with generator (46). We shall assume throughout this section that the assumptions and notation stated at the start of Section 6 hold. Let $\{X_t^N, V_t^N\}_{t \geq 0}$ be the Boomerang Sampler with characteristics $(\Sigma_{x,N}, \text{Proj}_N, \text{Proj}_N, \{\lambda_N^n\}, \lambda_r, \{R_N^n\}_n)$ which samples from the measure μ_N on the space \mathcal{H}^2 .

Here λ_n^N and R_n^N are approximations of λ_n and R_n respectively and we shall require that

$$\left\| \sum_{n=1}^{\infty} (\lambda_n - \lambda_n^N) \right\|^2 \leq C(\Phi) \sum_{n>N} \gamma_n^2 \tag{78}$$

$$\left\| \sum_{n=1}^{\infty} \lambda_n(x, v) \|R_n v - R_n^N v\|_{\mathcal{H}} \right\|_{L_{\mu}^2}^2 \leq C \|\Sigma \nabla_x \Phi\|_{L_{\pi}^2}^2 \sum_{n>N} \gamma_n^2. \tag{79}$$

Moreover we shall assume that $R_n^N v_{\perp}^N = v_{\perp}^N$ and that $\lambda_n^N(x, v) = \lambda_n^N(x^N, v^N)$ so that the process $\{X_t^N, V_t^N\}$ only changes the first N components. As in Section 5, Hypothesis 4.16 is satisfied for these choices with respect to potential function Φ_N and by Proposition 4.17 we have that μ_N is an invariant measure for this process. Note this process corresponds to the generator

$$\begin{aligned} \mathcal{L}_N f(x, v) &= \langle v^N, \nabla_x f(x, v) \rangle_{\mathcal{H}} - \langle x^N, \nabla_v f(x, v) \rangle_{\mathcal{H}} \\ &\quad + \lambda_{\tau} \int_{\mathcal{H}} f(x, w^N + v_{\perp}^N) - f(x, v) \nu_{0,N}(dw) \\ &\quad + \sum_{n=1}^{\infty} \lambda_n^N(x, v) [f(x, R_n^N v) - f(x, v)] \end{aligned} \tag{80}$$

for $x, v \in \mathcal{H}$ and $f \in D(\mathcal{L}_N)$. Define \mathcal{P}_t^N as follows

$$\mathcal{P}_t^N f(x, v) = \mathbb{E}[f(X_t^{N,x,v}, V_t^{N,x,v})] \tag{81}$$

and recall that \mathcal{P}_t is the semigroup associated to the process $\{X_t, V_t\}_{t \geq 0}$.

In order to prove convergence of \mathcal{P}_t^N to \mathcal{P}_t in a suitable topology we shall rely on the following estimate.

Lemma 7.1. *Let $\{\mathcal{P}_t\}$ be semigroup corresponding to the generator (46). Assume that Φ is twice continuously differentiable, bounded from below and has bounded Hessian. Let $\{\mathcal{P}_t^N\}_{t \geq 0}$ be the semigroup defined by (81) with generator \mathcal{L}_N defined by (80). Then there exists a constant C which may depend on Φ but is independent of t and N such for each $f \in C_b^1(\mathcal{H}^2)$, $N \in \mathbb{N}$, $t \geq 0$ we have*

$$\|(\mathcal{L}_N - \mathcal{L})\mathcal{P}_t^N f\|_{L_{\mu}^2} \leq C (\|\nabla_x f\|_{\infty} + \|\nabla_v f\|_{\infty} + \|f\|_{\infty}) \left(\sum_{i=N+1}^{\infty} \gamma_i^2 \right)^{\frac{1}{2}}.$$

Proof. To simplify notation let us set $f_t^N(x, v) = \mathcal{P}_t^N f(x, v)$ and notice that as the process (X_t^N, V_t^N) only lives in \mathcal{H}_N^2 and does not depend on x_{\perp}^N or v_{\perp}^N we have for $i > N$

$$\partial_{x_i} f_t^N(x, v) = \langle \nabla_x f_t^N, e_i \rangle_{\mathcal{H}} = \mathcal{P}_t^N (\partial_{x_i} f)(x, v), \tag{82}$$

$$\partial_{v_i} f_t^N(x, v) = \langle \nabla_v f_t^N, e_i \rangle_{\mathcal{H}} = \mathcal{P}_t^N (\partial_{v_i} f)(x, v). \tag{83}$$

In particular, for f with bounded derivatives we have the estimates

$$\begin{aligned} \|(1 - \text{Proj}_N) \nabla_x f_t^N(x, v)\|_{L_{\mu}^2(\mathcal{H}^2; \mathcal{H})} &\leq \|\nabla_x f\|_{\infty}, \\ \|(1 - \text{Proj}_N) \nabla_v f_t^N(x, v)\|_{L_{\mu}^2(\mathcal{H}^2; \mathcal{H})} &\leq \|\nabla_v f\|_{\infty} \end{aligned} \tag{84}$$

Fix $g \in C_b^1(\mathcal{H}^2)$ then

$$|\mathcal{L}g(x, v) - \mathcal{L}_N g(x, v)| \leq |\langle v_\perp^N, \nabla_x g(x, v) \rangle_{\mathcal{H}}| + |\langle x_\perp^N, \nabla_v g(x, v) \rangle_{\mathcal{H}}| \tag{*}$$

$$+ \lambda_r \left| \int_{\mathcal{H}} g(x, w) v_0(dw) - \int_{\mathcal{H}} g(x, w^N + v_\perp^N) v_{0,N}(dw_N) \right| \tag{**}$$

$$+ \sum_{n=1}^\infty |\lambda_n(x, v)[g(x, R_n v) - g(x, v)] - \lambda_n^N(x, v)[g(x, R_n^N v) - g(x, v)]|. \tag{***}$$

Replacing g with f_t^N in (*)-(***), taking the L_μ^2 -norm and using the triangle inequality we have

$$\|\mathcal{L}f_t^N(x, v) - \mathcal{L}_N f_t^N(x, v)\|_{L_\mu^2} \leq \|\langle v_\perp^N, \nabla_x f_t^N(x, v) \rangle_{\mathcal{H}}\|_{L_\mu^2} + \|\langle x_\perp^N, \nabla_v f_t^N(x, v) \rangle_{\mathcal{H}}\|_{L_\mu^2} \tag{†}$$

$$+ \lambda_r \left\| \int_{\mathcal{H}} f_t^N(x, w) v_0(dw) - \int_{\mathcal{H}} f_t^N(x, w^N + v_\perp^N) v_{0,N}(dw) \right\|_{L_\mu^2} \tag{††}$$

$$+ \left\| \sum_{n=1}^\infty (\lambda_n(x, v)[f_t^N(x, R_n v) - f_t^N(x, v)] - \lambda_n^N(x, v)[f_t^N(x, R_n^N v) - f_t^N(x, v)]) \right\|_{L_\mu^2}. \tag{†††}$$

First let us consider (†), by expanding in terms of $\{e_i\}_{i=1}^\infty$ and using (84) we have

$$\|\langle v_\perp^N, \nabla_x f_t^N(x, v) \rangle_{\mathcal{H}}\|_{L_\mu^2}^2 = \int_{\mathcal{H}^2} \sum_{i=N+1}^\infty v_i^2 (\partial_{x_i} f_t^N)^2 \mu(dx, dv) \leq \|\nabla_x f\|_\infty^2 \sum_{i=N+1}^\infty \gamma_i^2, \tag{85}$$

$$\|\langle x_\perp^N, \nabla_v f_t^N(x, v) \rangle_{\mathcal{H}}\|_{L_\mu^2}^2 = \int_{\mathcal{H}^2} \sum_{i=N+1}^\infty x_i^2 (\partial_{v_i} f_t^N)^2 \mu(dx, dv) \leq \|\nabla_v f\|_\infty^2 \|x_\perp^N\|_{L_\pi^2}^2, \tag{86}$$

Since we are assuming that Φ is bounded from below we can estimate $\|x_\perp^N\|_{L_\mu^2(\mathcal{H}; \mathcal{H})}$, let $C_\Phi = \inf_{x \in \mathcal{H}} \Phi(x)$ then

$$\|x_\perp^N\|_{L_\mu^2(\mathcal{H}; \mathcal{H})}^2 \leq e^{-C_\Phi} \sum_{i=N+1}^\infty \gamma_i^2. \tag{87}$$

Therefore applying (87) to (86) we have

$$\|\langle v_\perp^N, \nabla_x f_t^N(x, v) \rangle_{\mathcal{H}}\|_{L_\mu^2} + \|\langle x_\perp^N, \nabla_v f_t^N(x, v) \rangle_{\mathcal{H}}\|_{L_\mu^2} \tag{88}$$

$$\leq \left(\|\nabla_x f\|_\infty + e^{-\frac{1}{2}C_\Phi} \|\nabla_v f\|_\infty \right) \left(\sum_{i=N+1}^\infty \gamma_i^2 \right)^{\frac{1}{2}}. \tag{89}$$

By the Fundamental Theorem of Calculus for any $g \in C_b^1(\mathcal{H})$ we have

$$|g(x, w) - g(x, w^N + v_\perp^N)| \leq \int_0^1 |\langle \nabla_v g(x, s w_\perp^N + (1-s)v_\perp^N + w^N), w_\perp^N - v_\perp^N \rangle| ds \leq \|(1 - \text{Proj}_N) \nabla_v g\|_\infty \|w_\perp^N - v_\perp^N\|_{\mathcal{H}}. \tag{90}$$

If we apply this with the function g replaced by f_t^N then

$$|f_t^N(x, w) - f_t^N(x, w_\perp^N + v_\perp^N)| \leq \|\nabla_v f\|_\infty \|w_\perp^N - v_\perp^N\|_{\mathcal{H}}.$$

Now integrating with respect to μ we obtain an estimate for $(\dagger\dagger)$

$$\begin{aligned} |(\dagger\dagger)|^2 &\leq \lambda_r^2 \int_{\mathcal{H}} \int_{\mathcal{H}} |f_t^N(x, w) - f_t^N(x, w_\perp^N + v_\perp^N)|^2 \mu(dx, dv) \nu_0(dw) \\ &\leq \lambda_r^2 \|\nabla_v f\|_\infty^2 \int_{\mathcal{H}} \int_{\mathcal{H}} |w_\perp^N - v_\perp^N|^2 \nu_0(dw) \nu_0(dv) \leq 2\lambda_r^2 \|\nabla_v f\|_\infty^2 \sum_{i=N+1}^\infty \gamma_i^2. \end{aligned}$$

It remains to consider $(\dagger \dagger \dagger)$.

$$\begin{aligned} &\left\| \sum_{n=1}^\infty (\lambda_n(x, v)[f_t^N(x, R_n v) - f_t^N(x, v)] - \lambda_n^N(x, v)[f_t^N(x, R_n^N v) - f_t^N(x, v)]) \right\|_{L_\mu^2} \\ &\leq \left\| \sum_{n=1}^\infty (\lambda_n(x, v) - \lambda_n^N(x, v))[f_t^N(x, R_n v) - f_t^N(x, v)] \right\|_{L_\mu^2} \\ &+ \left\| \sum_{n=1}^\infty \lambda_n(x, v)[f_t^N(x, R_n v) - f_t^N(x, R_n^N v)] \right\|_{L_\mu^2} \\ &\leq 2\|f\|_\infty \left\| \sum_{n=1}^\infty (\lambda_n - \lambda_n^N) \right\|_{L_\mu^2} \tag{91} \end{aligned}$$

$$+ \left\| \sum_{n=1}^\infty \lambda_n(x, v)[f_t^N(x, R_n v) - f_t^N(x, R_n^N v)] \right\|_{L_\mu^2}. \tag{92}$$

We can bound these terms using (78) and (79).

So combining all these estimates we have

$$\begin{aligned} &\|\mathcal{L}f_t^N(x, v) - \mathcal{L}_N f_t^N(x, v)\|_{L_\mu^2} \\ &\leq C (\|f\|_\infty + \|\nabla_x f\|_\infty + \|\nabla_v f\|_\infty + \|\nabla_v f\|_\infty) \left(\sum_{i=N+1}^\infty \gamma_i^2 \right)^{\frac{1}{2}}. \quad \square \end{aligned}$$

Now we will prove for every $f \in C_b^1(\mathcal{H} \times \mathcal{H})$ that $\mathcal{P}_t^N f$ converges to $\mathcal{P}_t f$ in L_μ^2 , uniformly for t in compact intervals.

Theorem 7.2. *Let $\{\mathcal{P}_t\}$ be semigroup corresponding to the generator (46). Assume that Φ is twice continuously differentiable, convex, bounded from below and has bounded Hessian. Let $\{\mathcal{P}_t^N\}_{t \geq 0}$ be defined by (81). Then for each $f \in C_b^1(\mathcal{H}^2)$, $\mathcal{P}_t^N f$ converges to $\mathcal{P}_t f$ uniformly in time, that is for any $T > 0$ and $f \in C_b^1(\mathcal{H}^2)$*

$$\lim_{N \rightarrow \infty} \sup_{t \in [0, T]} \|\mathcal{P}_t f - \mathcal{P}_t^N f\|_{L_\mu^2} = 0.$$

Moreover, fix $T > 0$ then for any $f \in C_b^1(\mathcal{H}^2)$ there is a constant $C = C(f, \Phi)$ such that for N sufficiently large

$$\sup_{0 \leq t \leq T} \|\mathcal{P}_t f - \mathcal{P}_t^N f\|_{L_\mu^2} \leq C(f, \Phi) T \left(\sum_{i=N+1}^\infty \gamma_i^2 \right)^{\frac{1}{2}}$$

Proof of Theorem 7.2. Fix $f \in C_b^1(\mathcal{H}^2)$, we can write the difference of the semigroups in terms of their generators using as follows:

$$\begin{aligned} \|\mathcal{P}_t f - \mathcal{P}_t^N f\|_{L_\mu^2} &= \left\| \int_0^t \partial_s \mathcal{P}_{t-s} \mathcal{P}_s^N f ds \right\|_{L_\mu^2} \\ &\leq \int_0^t \|\mathcal{P}_{t-s}(\mathcal{L}_N - \mathcal{L})\mathcal{P}_s^N f\|_{L_\mu^2} ds \end{aligned}$$

Now since \mathcal{P}_t is a contraction semigroup on L_μ^2 we have

$$\|\mathcal{P}_t f - \mathcal{P}_t^N f\|_{L_\mu^2} \leq \int_0^t \|(\mathcal{L}_N - \mathcal{L})\mathcal{P}_s^N f\|_{L_\mu^2} ds$$

By Lemma 7.1 we bound this with

$$\|\mathcal{P}_t f - \mathcal{P}_t^N f\|_{L_\mu^2} \leq Ct (\|\nabla_x f\|_\infty + \|\nabla_p f\|_\infty + \|f\|_\infty) \left(\sum_{i=N+1}^\infty \gamma_i^2 \right)^{\frac{1}{2}}. \quad \square$$

Example 7.3. Set $\mathcal{H} = \ell^2$ to be the space of square integrable sequences and fix $s > 1$ and define the operator

$$\Sigma = (n^{-s} x_n)_{n=1}^\infty.$$

Note this is a bounded, self-adjoint, positive operator of Trace class. Set $\pi_0 = \mathcal{N}(0, \Sigma)$ and define $\Phi(x) = \|x\|_{\ell^2}^2/2$. Then the measure π is given by

$$\frac{d\pi}{d\pi_0}(x) = e^{-\frac{1}{2}\|x\|_{\ell^2}^2} \prod_{n=1}^\infty (1 + n^{-s}).$$

It is immediate to check that Φ is smooth, bounded below and has bounded Hessian so the conditions of Theorem 7.2 are satisfied. Therefore we can construct the pure reflection Boomerang sampler for this setting and a finite dimensional approximation with the canonical basis which converges on compact time intervals. As Φ is convex we will also obtain convergence uniform in time by Theorem 7.3. The rate of convergence is determined by

$$\sum_{i=N+1}^\infty \frac{1}{i^s} =: \zeta(s, N + 1).$$

Here ζ is the Hurwitz zeta function. To leading order $\zeta(s, N + 1) \approx N^{1-s}/(1 - s)$ so we may choose s to obtain any polynomial rate. \triangle

So far we have proven convergence for any finite time, but as we are interested in convergence to equilibria uniform in time estimates are more helpful. In order to prove these estimates we shall use the Hypocoercivity of Section 6. The following theorem is inspired by [11] in which exponential decay of the derivatives of a semigroup are used to prove uniform in time convergence of the weak error between an Euler scheme approximation and an SDE. This technique is also used in [5] to show uniform in time convergence between a SDE on a fast dynamical network and an averaged SDE. In both of these papers the results rely upon obtaining exponential decay of the derivatives of the semigroup, conditions for such an estimate are given in [12]. In contrast for a PDMP the semigroup need not be even differentiable in all directions so we cannot hope to have an estimate which decays exponentially. Instead of the

derivative estimate we shall use the exponential convergence from the Hypocoercivity theory established in Section 6. This has the advantage that we do not require derivative estimates however we instead need control on the error between the approximate invariant measure μ_N and the invariant measure for the infinite dimensional process μ .

Theorem 7.3. *Let $\{\mathcal{P}_t\}$ be semigroup corresponding to the generator (46). Assume that Φ is twice continuously differentiable, convex (or bounded with bounded derivative), bounded from below and has bounded Hessian. Let $\{\mathcal{P}_t^N\}_{t \geq 0}$ be the corresponding to the finite dimensional approximation constructed and suppose that*

$$\lim_{N \rightarrow \infty} \left\| \frac{d\mu}{d\mu_N} - 1 \right\|_{L^2_{\mu_N}} = 0. \tag{93}$$

Then for each $f \in C_b^1(\mathcal{H}^2)$, $\mathcal{P}_t^N f$ converges to $\mathcal{P}_t f$ uniformly in time, that is for any f bounded and measurable

$$\lim_{N \rightarrow \infty} \sup_{t \geq 0} \|\mathcal{P}_t f - \mathcal{P}_t^N f\|_{L^2_\mu} = 0.$$

Moreover, for any $f \in C_b^1(\mathcal{H}^2)$ there is a constant $C = C(f, \Phi)$ such that for N sufficiently large

$$\|\mathcal{P}_t f - \mathcal{P}_t^N f\|_{L^2_\mu} \leq C(f, \Phi) \left(\sum_{i=N+1}^\infty \gamma_i^2 \right)^{\frac{1}{2}} + \left\| \frac{d\mu}{d\mu_N} - 1 \right\|_{L^2_{\mu_N}}.$$

Proof of Theorem 7.3. Fix $f \in C_b^1(\mathcal{H}^2)$, we can write the difference of the semigroups in terms of their generators as follows:

$$\begin{aligned} \|\mathcal{P}_t f - \mathcal{P}_t^N f\|_{L^2_\mu} &= \left\| \int_0^t \partial_s \mathcal{P}_{t-s} \mathcal{P}_s^N f ds \right\|_{L^2_\mu} \\ &\leq \int_0^t \|\mathcal{P}_{t-s} (\mathcal{L}_N - \mathcal{L}) \mathcal{P}_s^N f\|_{L^2_\mu} ds \end{aligned}$$

We wish to use hypocoercivity to control this, to which end we add and subtract $\mu((\mathcal{L}_N - \mathcal{L})\mathcal{P}_s^N f)$. Note that $\mu(\mathcal{L}\mathcal{P}_s^N f) = 0$ since μ is an invariant measure for \mathcal{P}_t .

$$\|\mathcal{P}_t f - \mathcal{P}_t^N f\|_{L^2_\mu} \leq \int_0^t \|\mathcal{P}_{t-s} (\mathcal{L}_N - \mathcal{L}) \mathcal{P}_s^N f - \mu(\mathcal{L}_N \mathcal{P}_s^N f)\|_{L^2_\mu} + |\mu(\mathcal{L}_N \mathcal{P}_s^N f)| ds$$

By Theorem 6.2 we have

$$\begin{aligned} \|\mathcal{P}_t f - \mathcal{P}_t^N f\|_{L^2_\mu} &\leq \int_0^t C e^{-\kappa(t-s)} \|(\mathcal{L}_N - \mathcal{L}) \mathcal{P}_s^N f - \mu(\mathcal{L}_N \mathcal{P}_s^N f)\|_{L^2_\mu} + |\mu(\mathcal{L}_N \mathcal{P}_s^N f)| ds \\ &\leq \int_0^t C e^{-\kappa(t-s)} \|(\mathcal{L}_N - \mathcal{L}) \mathcal{P}_s^N f\|_{L^2_\mu} + (C + 1) |\mu(\mathcal{L}_N \mathcal{P}_s^N f)| ds. \end{aligned}$$

Let us consider the term $|\mu(\mathcal{L}_N \mathcal{P}_s^N f)|$.

$$\begin{aligned} |\mu(\mathcal{L}_N \mathcal{P}_s^N f)| &= |\mu(\mathcal{L}_N \mathcal{P}_s^N f) - \mu_N(\mathcal{L}_N \mathcal{P}_s^N f)| \\ &= \mu_N \left((\mathcal{L}_N \mathcal{P}_s^N f) \left(\frac{d\mu}{d\mu_N} - 1 \right) \right) \\ &\leq \|\mathcal{L}_N \mathcal{P}_s^N f\|_{L^2_{\mu_N}} \left\| \frac{d\mu}{d\mu_N} - 1 \right\|_{L^2_{\mu_N}} \end{aligned}$$

Under our assumptions we have that \mathcal{P}_t^N is hypocoercive with constants independent of N , therefore

$$|\mu(\mathcal{L}_N \mathcal{P}_s^N f)| \leq C e^{-\rho t} \|\mathcal{L}_N f\|_{L^2_{\mu_N}} \left\| \frac{d\mu}{d\mu_N} - 1 \right\|_{L^2_{\mu_N}}.$$

Hence

$$\begin{aligned} & \|\mathcal{P}_t f - \mathcal{P}_t^N f\|_{L^2_{\mu}} \\ & \leq \int_0^t C e^{-\kappa(t-s)} \|(\mathcal{L}_N - \mathcal{L})\mathcal{P}_s^N f\|_{L^2_{\mu}} ds + \frac{C(C+1)}{\rho} \|\mathcal{L}_N f\|_{L^2_{\mu_N}} \left\| \frac{d\mu}{d\mu_N} - 1 \right\|_{L^2_{\mu_N}}. \end{aligned}$$

It remains to show that

$$\lim_{N \rightarrow \infty} \sup_{t \geq 0} \int_0^t e^{-\kappa(t-s)} \|(\mathcal{L}_N - \mathcal{L})\mathcal{P}_s^N f\|_{L^2_{\mu}} ds = 0.$$

This follows since we can bound the integrand using [Lemma 7.1](#) which gives

$$\begin{aligned} \int_0^t e^{-\kappa(t-s)} \|(\mathcal{L}_N - \mathcal{L})\mathcal{P}_s^N f\|_{L^2_{\mu}} ds & \leq \int_0^t e^{-\kappa(t-s)} ds C (\|\nabla_x f\|_{\infty} + \|\nabla_v f\|_{\infty} + \|f\|_{\infty}) \\ & \quad \times \left(\sum_{i=N+1}^{\infty} \gamma_i^2 \right)^{\frac{1}{2}} \\ & = \frac{1}{\kappa} (1 - e^{-\kappa t}) C (\|\nabla_x f\|_{\infty} + \|\nabla_v f\|_{\infty} + \|f\|_{\infty}) \\ & \quad \times \left(\sum_{i=N+1}^{\infty} \gamma_i^2 \right)^{\frac{1}{2}} \quad \square \end{aligned}$$

Example 7.4 (*Example 7.3 continued*). We showed in [Example 7.3](#) how to construct an example with an arbitrarily chosen polynomial rate of convergence. We wish to continue this example to show in this case that we get the same order of convergence uniform in time. Note that as Φ is convex we have that all the conditions of [Theorem 7.3](#) are satisfied provided [\(93\)](#) holds. If we show that

$$\left\| \frac{d\mu}{d\mu_N} - 1 \right\|_{L^2_{\mu_N}}^2 \leq C \sum_{i=N+1}^{\infty} \gamma_i^2$$

holds then [\(93\)](#) is satisfied and the rate of convergence of $\mathcal{P}_t^N f$ is the same as in [Example 7.3](#). Note that if we do not require the rate of convergence then it is immediate to check that [\(93\)](#) holds since $d\mu/d\mu_N$ is bounded uniformly in N by 2 so the results follows by the dominated convergence theorem. For this example

$$\frac{d\mu}{d\mu_N} = \frac{d\pi}{d\pi_N} = \frac{\exp(-\|x_{\perp}^N\|_{\mathcal{H}}^2/2)}{\pi_0(\exp(-\|x_{\perp}^N\|_{\mathcal{H}}^2/2))}.$$

Then we can write

$$\left\| \frac{d\mu}{d\mu_N} - 1 \right\|_{L^2_{\mu_N}}^2 = \int_{\mathcal{H}} \left| \frac{\exp(-\|x_{\perp}^N\|_{\mathcal{H}}^2/2)}{\pi_0(\exp(-\|x_{\perp}^N\|_{\mathcal{H}}^2/2))} - 1 \right|^2 \frac{\exp(-\|x^N\|_{\mathcal{H}}^2/2)}{\pi_0(\exp(-\|x^N\|_{\mathcal{H}}^2/2))} \pi_0(dx).$$

Under π_0 x^N is independent of x_{\perp}^N so this simplifies to

$$\begin{aligned} \left\| \frac{d\mu}{d\mu_N} - 1 \right\|_{L^2_{\mu_N}}^2 &= \int_{\mathcal{H}} \left| \frac{\exp(-\|x_{\perp}^N\|_{\mathcal{H}}^2/2)}{\pi_0(\exp(-\|x_{\perp}^N\|_{\mathcal{H}}^2/2))} - 1 \right|^2 \pi_0(dx) \\ &= \int_{\mathcal{H}} \frac{\exp(-\|x_{\perp}^N\|_{\mathcal{H}}^2)}{\pi_0(\exp(-\|x_{\perp}^N\|_{\mathcal{H}}^2/2))^2} \pi_0(dx) - 1. \end{aligned}$$

Again using the independence structure of π_0 we can rewrite this as

$$\left\| \frac{d\mu}{d\mu_N} - 1 \right\|_{L^2_{\mu_N}}^2 = \prod_{i>N} \left[\frac{\pi_0(\exp(-x_i^2))}{\pi_0(\exp(-x_i^2/2))^2} \right] - 1.$$

As π_0 is a Gaussian measure we can evaluate this integrals and obtain

$$\left\| \frac{d\mu}{d\mu_N} - 1 \right\|_{L^2_{\mu_N}}^2 = \prod_{i>N} \left[\frac{\gamma_i^2 + 1}{\sqrt{2\gamma_i^2 + 1}} \right] - 1 \sim \frac{1}{2} \sum_{i>N} \gamma_i^4. \quad \Delta$$

7.1. Finite dimensional approximation of pure reflection Boomerang Sampler

In this section we verify that (78) and (79) both hold for the approximation of the Pure reflection Boomerang Sampler given in Example 7.1.

Lemma 7.4. *Let R_N and λ_N be as in Example 7.1 then (78) and (79) hold.*

Proof. First consider (78). In this case the left hand side of (78) simplifies to

$$\begin{aligned} \left\| \sum_{n=1}^{\infty} (\lambda_n - \lambda_n^N) \right\| &= \|(\langle \nabla_x \Phi(x), v \rangle_{\mathcal{H}})_+ - (\langle \nabla_x \Phi(x_N), v_N \rangle_{\mathcal{H}})_+\| \\ &\leq \| \langle \nabla_x \Phi(x) - \nabla_x \Phi(x_N), v \rangle_{\mathcal{H}} \| \\ &\leq \sqrt{\text{Tr}(\Sigma)} \| \nabla_x \Phi(x) - \nabla_x \Phi(x_N) \| \\ &\leq \sqrt{\text{Tr}(\Sigma)} \| \nabla_x^2 \Phi \|_{\infty} \| x - x_N \|. \end{aligned}$$

Then (78) follows by (87).

Now consider (79) the left hand side of which is

$$\begin{aligned} \sum_{n=1}^{\infty} \| \lambda_n(x, v) \| R_n v - R_n^N v \|_{\mathcal{H}} \|_{L^2_{\mu}} &= 2 \| (\langle \nabla_x \Phi(x), v \rangle_{\mathcal{H}})_+ \| v_{\perp}^N \|_{\mathcal{H}} \|_{L^2_{\mu}} \\ &\leq 2 \left(\sum_{i=1}^{\infty} \sum_{j>N} \int_{\mathcal{H}^2} (\partial_{x_i} \Phi(x))^2 v_i^2 v_j^2 \mu(dx, dv) \right)^{\frac{1}{2}}. \end{aligned}$$

Now since e_i are eigenvectors of Σ we have that $\int v_i^2 v_j^2 dv = \delta_{ij} 3\gamma_j^4 + \gamma_i^2 \gamma_j^2 (1 - \delta_{ij})$ so

$$\begin{aligned} & \sum_{n=1}^{\infty} \|\lambda_n(x, v)\| R_n v - R_n^N v \|_{\mathcal{H}} \|_{L^2_{\mu}} \\ &= 2 \left(\sum_{j>N} \gamma_j^2 \sum_{\substack{i=1 \\ i \neq j}}^{\infty} \gamma_i^2 \int_{\mathcal{H}^2} (\partial_{x_i} \Phi(x))^2 \pi(dx) + \sum_{j>N} 3\gamma_j^4 \int_{\mathcal{H}^2} (\partial_{x_j} \Phi(x))^2 \pi(dx) \right)^{\frac{1}{2}} \\ &\leq 4 \|\Sigma \nabla_x \Phi\|_{L^2_{\pi}} \left(\sum_{j>N} \gamma_j^2 \right)^{\frac{1}{2}} \quad \square \end{aligned}$$

7.2. Finite dimensional approximation of Factorised Boomerang Sampler

In this section we verify that (78) and (79) both hold for the approximation of the Factorised Boomerang Sampler given in Example 7.2.

Lemma 7.5. *Let R_N and λ_N be as in Example 7.2 then (78) and (79) hold.*

Proof. First consider (78). In this case the left hand side of (78) simplifies to

$$\begin{aligned} \left\| \sum_{n=1}^{\infty} |\lambda_n - \lambda_n^N| \right\|^2 &= \int_{\mathcal{H}^2} \sum_{n,m} |\lambda_n(x, v) - \lambda_n^N(x, v)| |\lambda_m(x, v) - \lambda_m^N(x, v)| \mu(dx, dv) \\ &= \int_{\mathcal{H}^2} \sum_{n,m \leq N} |\lambda_n - \lambda_n^N| |\lambda_m - \lambda_m^N| d\mu \\ &\quad + \int_{\mathcal{H}^2} \sum_{n,m > N} |\lambda_n - \lambda_n^N| |\lambda_m - \lambda_m^N| d\mu \\ &\quad + 2 \int_{\mathcal{H}^2} \sum_{n \leq N, m > N} |\lambda_n - \lambda_n^N| |\lambda_m - \lambda_m^N| d\mu \\ &\leq \int_{\mathcal{H}^2} \sum_{n,m \leq N} |v_n v_m| |\partial_{x_n} \Phi(x) - \partial_{x_n} \Phi(x_N)| |\partial_{x_m} \Phi(x) - \partial_{x_m} \Phi(x_N)| d\mu \\ &\quad + \int_{\mathcal{H}^2} \sum_{n,m > N} |v_n v_m| |\partial_{x_n} \Phi(x)| |\partial_{x_m} \Phi(x)| d\mu \\ &\quad + 2 \int_{\mathcal{H}^2} \sum_{n \leq N, m > N} |v_n v_m| |\partial_{x_n} \Phi(x) - \partial_{x_n} \Phi(x_N)| |\partial_{x_m} \Phi(x)| d\mu \end{aligned}$$

Note that v is independent of x under μ and $v_0(|v_n v_m|) \leq \sqrt{\gamma_n^2 \gamma_m^2}$.

$$\begin{aligned} \left\| \sum_{n=1}^{\infty} |\lambda_n - \lambda_n^N| \right\|^2 &\leq \int_{\mathcal{H}} \sum_{n,m \leq N} \sqrt{\gamma_n^2 \gamma_m^2} |\partial_{x_n} \Phi(x) - \partial_{x_n} \Phi(x_N)| |\partial_{x_m} \Phi(x) - \partial_{x_m} \Phi(x_N)| d\pi \\ &\quad + \int_{\mathcal{H}} \sum_{n,m > N} \sqrt{\gamma_n^2 \gamma_m^2} |\partial_{x_n} \Phi(x)| |\partial_{x_m} \Phi(x)| d\pi \end{aligned}$$

$$\begin{aligned}
 &+ 2 \int_{\mathcal{H}} \sum_{n \leq N, m > N} \sqrt{\gamma_n^2 \gamma_m^2} |\partial_{x_n} \Phi(x) - \partial_{x_n} \Phi(x_N)| |\partial_{x_m} \Phi(x)| d\pi \\
 &\leq \left(\sum_{n \leq N} \sqrt{\gamma_n^2} \|\partial_{x_n} \Phi(x) - \partial_{x_n} \Phi(x_N)\|_{L^2_{\pi}} \right)^2 \\
 &+ \left(\sum_{n > N} \sqrt{\gamma_n^2} \|\partial_{x_n} \Phi(x)\|_{L^2_{\pi}} \right)^2 + 2 \sum_{n \leq N, m > N} \sqrt{\gamma_n^2 \gamma_m^2} \|\partial_{x_n} \Phi(x) - \partial_{x_n} \Phi(x_N)\|_{L^2_{\pi}} \\
 &\times \|\partial_{x_m} \Phi(x)\|_{L^2_{\pi}}
 \end{aligned}$$

Now use the Mean value theorem

$$\begin{aligned}
 \left\| \sum_{n=1}^{\infty} |\lambda_n - \lambda_n^N| \right\|^2 &\leq \left(\sum_{n \leq N} \sqrt{\gamma_n^2} \|\nabla_x \partial_{x_n} \Phi\|_{\infty} \|x - x_N\|_{L^2_{\pi}} \right)^2 + \left(\sum_{n > N} \sqrt{\gamma_n^2} \|\partial_{x_n} \Phi(x)\|_{L^2_{\pi}} \right)^2 \\
 &+ 2 \sum_{n \leq N, m > N} \sqrt{\gamma_n^2 \gamma_m^2} \|\nabla_x \partial_{x_n} \Phi\|_{\infty} \|x - x_N\|_{L^2_{\pi}} \|\partial_{x_m} \Phi(x)\|_{L^2_{\pi}}
 \end{aligned}$$

Using Cauchy–Schwarz

$$\begin{aligned}
 \left\| \sum_{n=1}^{\infty} |\lambda_n - \lambda_n^N| \right\|^2 &\leq \left(\sum_{n \leq N} \gamma_n^2 \right) \left(\sum_{n \leq N} \|\nabla_x \partial_{x_n} \Phi\|_{\infty}^2 \right) \|x - x_N\|_{L^2_{\pi}}^2 \\
 &+ \left(\sum_{n > N} \gamma_n^2 \right) \left(\sum_{n > N} \|\partial_{x_n} \Phi(x)\|_{L^2_{\pi}}^2 \right) \\
 &+ 2 \|x - x_N\|_{L^2_{\pi}} \sqrt{\sum_{n \leq N} \gamma_n^2} \sqrt{\sum_{n \leq N} \|\nabla_x \partial_{x_n} \Phi\|_{\infty}^2} \sqrt{\sum_{m > N} \gamma_m^2} \\
 &\times \sqrt{\sum_{m > N} \|\partial_{x_m} \Phi(x)\|_{L^2_{\pi}}^2} \\
 &\leq \text{Tr}(\Sigma) \|\nabla_x^2 \Phi\|_{\infty}^2 \|x - x_N\|_{L^2_{\pi}}^2 + \left(\sum_{n > N} \gamma_n^2 \right) \|\nabla \Phi(x)\|_{L^2_{\pi}}^2 \\
 &+ 2 \|x - x_N\|_{L^2_{\pi}} \sqrt{\text{Tr}(\Sigma)} \|\nabla_x^2 \Phi\|_{\infty} \sqrt{\sum_{m > N} \gamma_m^2} \|\nabla_x \Phi(x)\|_{L^2_{\pi}}
 \end{aligned}$$

Then (78) follows by (87).

Now consider (79) the left hand side of which is

$$\begin{aligned}
 &\left\| \sum_{n=1}^{\infty} \lambda_n(x, v) \|R_n v - R_n^N v\|_{\mathcal{H}} \right\|_{L^2_{\mu}}^2 \\
 &= \int_{\mathcal{H}^2} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \lambda_n(x, v) \lambda_m(x, v) \|R_n v - R_n^N v\|_{\mathcal{H}} \|R_m v - R_m^N v\|_{\mathcal{H}} \pi(dx) \nu_0(dv).
 \end{aligned}$$

Now if $n \leq N$ then $R_n^N = R_n$ so all terms vanish except when $n > N$ and $m > N$. In that case $R_n^N = 1$ so $R_n v - R_n^N v = -2v_n e_n$, thus we have

$$\begin{aligned} \left\| \sum_{n=1}^{\infty} \lambda_n(x, v) \|R_n v - R_n^N v\|_{\mathcal{H}} \right\|_{L^2_{\mu}}^2 &= 4 \int_{\mathcal{H}^2} \sum_{n>N} \sum_{m>N} (\partial_{x_n} \Phi(x) v_n)^+ (\partial_{x_m} \Phi(x) v_m)^+ \\ &\quad \times |v_n| |v_m| \pi(dx) \nu_0(dv), \\ &\leq 4 \int_{\mathcal{H}^2} \sum_{n>N} \sum_{m>N} |\partial_{x_n} \Phi(x)| |\partial_{x_m} \Phi(x)| \\ &\quad \times |v_n|^2 |v_m|^2 \pi(dx) \nu_0(dv). \end{aligned}$$

Now $\int v_i^2 v_j^2 dv = \delta_{ij} 3\gamma_j^4 + \gamma_i^2 \gamma_j^2 (1 - \delta_{ij})$, in particular $\int v_i^2 v_j^2 dv \leq 3\gamma_i^2 \gamma_j^2$.

$$\begin{aligned} \left\| \sum_{n>N} \lambda_n(x, v) \|R_n v - R_n^N v\|_{\mathcal{H}} \right\|_{L^2_{\mu}}^2 &\leq 12 \int_{\mathcal{H}} \sum_{n>N} \sum_{m>N} \gamma_n^2 \gamma_m^2 |\partial_{x_n} \Phi(x)| |\partial_{x_m} \Phi(x)| \pi(dx) \\ &\leq 12 \int_{\mathcal{H}} \left(\sum_{n>N} \gamma_n^2 |\partial_{x_n} \Phi(x)| \right)^2 \pi(dx) \\ &\leq 12 \left(\sum_{n>N} \gamma_n^2 \right) \int_{\mathcal{H}} \left(\sum_{n>N} \gamma_n^2 |\partial_{x_n} \Phi(x)|^2 \right) \pi(dx) \\ &\leq 12 \|\Sigma \nabla_x \Phi(x)\|_{L^2_{\pi}}^2 \left(\sum_{n>N} \gamma_n^2 \right). \quad \square \end{aligned}$$

8. Proofs

8.1. Proofs of Section 4

8.1.1. Proofs of Section 4.1

Proof of Theorem 4.3. To prove well-posedness of the piecewise deterministic Markov process we shall apply Theorem 2.3, to this end we show that Assumption 2.1 holds.

(i,viii) Using Fréchet differentiability of Φ , this follows from the estimate

$$|\lambda_i(x, v) - \lambda_i(y, v)| \leq \frac{|v_i|}{\gamma_i^2} |x_i - y_i| + |v_i| |\partial_{x_i} \Phi(x) - \partial_{y_i} \Phi(y)|.$$

Now summing over i and using the Cauchy–Schwarz inequality we have

$$|\lambda(x, v) - \lambda(y, v)| \leq \|\Sigma^{-1} v\| \|x - y\| + \|v\| \|\nabla_x \phi(x) - \nabla_x \phi(y)\|.$$

Therefore $x \mapsto \lambda(x, v)$ is continuous, continuity in the v component is similar.

- (ii) This is immediate from the definition of (Q_i) .
- (iii) Is immediate since the absolute value of velocities is preserved by reflections.
- (iv) There are no refreshments.
- (v) This follows from (19) and Assumption 4.1(3).
- (vi) There are no refreshments.
- (vii) The flow is given by $\varphi_t(x, v) = (x + tv, v)$ so we have that (vii) holds with c_t growing linearly.

To show that μ is an invariant measure we shall appeal to [Theorem 3.1](#), so we must show that [\(12\)](#) holds. In this case $L_x f = \langle v, \nabla_x f \rangle_{\mathcal{H}}$ then by [[13](#), Proposition 10.20] the adjoint of L_x in $L^2_{\pi_0}$ is given by

$$L^*_x f = -\langle \nabla_x f, v \rangle_{\mathcal{H}} + \langle \Sigma^{-1}v, x \rangle_{\mathcal{H}} f.$$

Note that $\Sigma^{-1}v \in \mathcal{H}$ so the above expression is well defined. Therefore we have

$$e^{\Phi(x)} L^*_x (e^{-\Phi})(x, v) = (\langle \nabla_x \Phi(x), v \rangle_{\mathcal{H}} + \langle \Sigma^{-1}v, x \rangle_{\mathcal{H}}). \tag{94}$$

Next we have

$$Q^*((x, v), dw) = \sum_{i=1}^{\infty} \frac{\lambda_i(x, F_i v)}{\lambda(x, F_i v)} \delta_{F_i v}(dw).$$

Indeed we can verify [\(11\)](#) holds, the left hand side of [\(11\)](#) is

$$\int_{\mathcal{H}} \int_{\mathcal{H}} g(v, w) Q^*((x, v), dw) \mu_0(dx, dv) = \sum_{i=1}^{\infty} \int_{\mathcal{H}} \int_{\mathcal{H}} \frac{\lambda_i(x, F_i v)}{\lambda(x, F_i v)} g(v, F_i v) \mu_0(dx, dv).$$

As F_i leave μ_0 invariant we can do a change of variables $v \rightarrow F_i v$ for each i .

$$\begin{aligned} \int_{\mathcal{H}} \int_{\mathcal{H}} g(v, w) Q^*((x, v), dw) \mu_0(dx, dv) &= \sum_{i=1}^{\infty} \int_{\mathcal{H}} \int_{\mathcal{H}} \frac{\lambda_i(x, v)}{\lambda(x, v)} g(F_i v, v) \mu_0(dx, dv) \\ &= \int_{\mathcal{H}} \int_{\mathcal{H}} g(w, v) Q((x, v), dw) \mu_0(dx, dv). \end{aligned}$$

Thus [\(11\)](#) holds. Now

$$\int_{\mathcal{H}} \lambda(x, w) Q^*((x, v), dw) - \lambda(x, v) = \sum_{i=1}^{\infty} \frac{\lambda_i(x, F_i v)}{\lambda(x, F_i v)} \lambda(x, F_i v) - \lambda(x, v)$$

Now using the definition of λ_i we have

$$\int_{\mathcal{H}} \lambda(x, w) Q^*((x, v), dw) - \lambda(x, v) = -\langle v, \nabla_x \Phi(x) \rangle_{\mathcal{H}} - \langle \Sigma^{-1}v, x \rangle_{\mathcal{H}}. \tag{95}$$

Finally combining [\(94\)](#) and [\(95\)](#) we have [\(12\)](#) holds. Then by [Theorem 3.1](#) we have μ is formally an invariant measure for the Zig Zag process. \square

8.1.2. Proofs of Section 4.2

Proof of Proposition 4.8. Fix $x \in \mathcal{H}$ with $\Sigma^{-1}x \in \mathcal{H}$. Under our assumptions, $\nabla \Psi$ is

$$\nabla \Psi = \nabla \Phi(x) + \Sigma^{-1}x.$$

Therefore

$$\langle v, \nabla \Psi \rangle = \langle v, \nabla \Phi \rangle + \langle \Sigma^{-1}v, x \rangle. \tag{96}$$

By [Assumption 4.1](#) the first addend on the right hand side of the above is well defined and since $\Sigma^{-1}v \in \mathcal{H}$ the second addend is also well defined for all $x \in \mathcal{H}, v \in \mathcal{V}_{\text{BPS}}$. Now by continuity we have that $\langle v, \nabla \Psi \rangle$ is well defined for all $x \in \mathcal{H}, v \in \mathcal{V}_{\text{BPS}}$, and hence λ is also.

In order to prove the statement (2) it suffices to show that $\Sigma^{-1} \Sigma^{\zeta} \nabla \Psi(x) \in \mathcal{H}$ for every $x \in \mathcal{H}$. Indeed, if this is the case, then it is easy to see that also $\|\Sigma^{\zeta/2} \nabla \Psi\|$ is finite. To show

that $\Sigma^{\zeta-1}\nabla\Psi(x) \in \mathcal{H}$ we act as above and write for every $x \in \mathcal{H}, v \in \mathcal{V}_{\text{BPS}}$ with $\Sigma^{-1}x \in \mathcal{H}$ that

$$\Sigma^{\zeta-1}\nabla\Psi = \Sigma^{\zeta-1}\nabla\Phi(x) + \Sigma^{\zeta-2}x$$

Imposing that $\zeta \geq 2$ we ensure this is well defined. \square

Proof of Proposition 4.10. To show (1),

$$\begin{aligned} R(x)[R(x)v] &= R(x) \left[v - \frac{2\langle v, \nabla\Psi \rangle}{\|\Sigma^{\zeta/2}\nabla\Psi\|^2} \Sigma^{\zeta}\nabla\Psi \right] \\ &= v - \frac{2\langle v, \nabla\Psi \rangle}{\|\Sigma^{\zeta/2}\nabla\Psi\|^2} \Sigma^{\zeta}\nabla\Psi - 2 \frac{\Sigma^{\zeta}\nabla\Psi}{\|\Sigma^{\zeta/2}\nabla\Psi\|^2} \\ &\quad \times \left(\langle v, \nabla\Psi \rangle - 2 \frac{\langle v, \nabla\Psi \rangle}{\|\Sigma^{\zeta/2}\nabla\Psi\|^2} \langle \Sigma^{\zeta}\nabla\Psi, \nabla\Psi \rangle \right) \\ &= v + \left[-4 \frac{\langle v, \nabla\Psi \rangle}{\|\Sigma^{\zeta/2}\nabla\Psi\|^2} + 4 \frac{\langle v, \nabla\Psi \rangle}{\|\Sigma^{\zeta/2}\nabla\Psi\|^4} \langle \Sigma^{\zeta}\nabla\Psi, \nabla\Psi \rangle \right] \Sigma^{\zeta}\nabla\Psi = v. \end{aligned}$$

Therefore the claim follows. As for point (2):

$$\langle R(x)v, \nabla\Psi \rangle = \langle v, \nabla\Psi \rangle - \frac{2\langle v, \nabla\Psi \rangle}{\|\Sigma^{\zeta/2}\nabla\Psi\|^2} \langle \Sigma^{\zeta}\nabla\Psi, \nabla\Psi \rangle.$$

Hence, under the constraints on the parameters stated in point (2), (26) holds. To prove the statement in point (3), recall that the Fourier transform $\hat{\nu}$ of a measure ν on \mathcal{H}

$$\hat{\nu}(\xi) = \int_{\mathcal{H}^{\alpha}} e^{i\langle z, \xi \rangle} d\nu(z)$$

and that a measure ν_0 is a centred Gaussian with covariance operator Σ^{ζ} on \mathcal{H} if and only if

$$\hat{\nu}_0(\xi) = e^{-\frac{1}{2}\langle \Sigma^{\zeta}\xi, \xi \rangle}.$$

Because Fourier transforms characterise measures (see [13, Proposition 1.7]), we just need to impose $\hat{\nu} = \hat{\nu}_R$. To this end, start by defining the formal adjoint operator R'_x of $R(x)$, namely the operator

$$R'_x\xi := \xi - 2 \frac{\langle \xi, \Sigma^{\zeta}\nabla\Psi \rangle}{\|\Sigma^{\zeta/2}\nabla\Psi\|^2} \nabla\Psi.$$

With this definition one has

$$\langle R(x)z, \xi \rangle = \langle z, R(x)'\xi \rangle.$$

Let us now calculate $\hat{\nu}_R$:

$$\begin{aligned} \hat{\nu}_R(\xi) &= \int_{\mathcal{H}} e^{i\langle z, \xi \rangle} d\nu_R(z) = \int_{\mathcal{H}} e^{i\langle R(x)z, \xi \rangle} d\nu_0(z) \\ &= \int_{\mathcal{H}} e^{i\langle z, R'_x\xi \rangle} d\nu_0(z) = e^{-\frac{1}{2}\langle \Sigma^{\zeta}(R'_x\xi), (R'_x\xi) \rangle} \end{aligned}$$

where in the first equality we have used the involutivity of R_x , which gives the standard change of variables formula

$$\int g(z) d(\nu \circ R_x)(z) = \int (g \circ R_x)(z) d\nu(z).$$

In order to have $\hat{v}_R = \hat{v}_0$, we then need to have

$$\langle \Sigma^\zeta (R'_x \xi), (R'_x \xi) \rangle = \langle \Sigma^\zeta \xi, \xi \rangle, \text{ for every } \xi \in \mathcal{H}.$$

Because

$$\begin{aligned} \langle \Sigma^\zeta (R'_x \xi), (R'_x \xi) \rangle &= \langle \Sigma^\zeta \xi, \xi \rangle - 4 \frac{\langle \xi, \Sigma^\zeta \nabla \Psi \rangle}{\|\Sigma^{\zeta/2} \nabla \Psi\|^2} \langle \Sigma^\zeta \xi, \nabla \Psi \rangle \\ &\quad + 4 \frac{|\langle \xi, \Sigma^\zeta \nabla \Psi \rangle|^2}{\|\Sigma^{\zeta/2} \nabla \Psi\|^4} \langle \Sigma^\zeta \nabla \Psi, \nabla \Psi \rangle = \langle \Sigma^\zeta \xi, \xi \rangle, \end{aligned}$$

thus the reflection invariance of ν_0 does hold.

To prove statement (4) we show that (12) holds. Since IDBPS fits in the setting of Example 3.2 it is sufficient to verify the conditions given in that example, note that the first two conditions are satisfied since the reflection operator only acts on the v component and by statement (3) of this proposition so it suffices to show that (14) holds. As in the proof of invariance for the IDZZS we have that (94) holds so it remains to show that

$$\langle \nabla_x \Phi(x), v \rangle + \langle \Sigma^{-1} v, x \rangle + \lambda(x, R(x)v) - \lambda(x, v) = 0 \tag{97}$$

By (26) we have that

$$\lambda(x, R(x)v) - \lambda(x, v) = -\langle \Sigma^\eta v, \nabla \Psi \rangle = -\langle \nabla_x \Phi(x), v \rangle - \langle \Sigma^{-1} v, x \rangle.$$

Therefore (97) holds. \square

Proof of Lemma 4.11. Since $\Phi = 0$ we can write

$$R(x)v = v - \frac{2\langle v, \Sigma^{-1}x \rangle}{\|\Sigma^{\zeta/2-1}x\|^2} \Sigma^{\zeta-1}x.$$

Then we have

$$\|R(x)v\|_\alpha^2 = \|v\|_\alpha^2 + \frac{4\langle \Sigma^{-1}v, x \rangle^2}{\|\Sigma^{\zeta/2-1}x\|^4} \|\Sigma^{\zeta-1+\alpha}x\|^2 - \frac{4\langle \Sigma^{-1}v, x \rangle}{\|\Sigma^{\zeta/2-1}x\|^2} \langle \Sigma^{\zeta-1+\alpha}x, \Sigma^\alpha v \rangle.$$

Therefore (27) holds only if

$$\frac{\langle \Sigma^{-1}v, x \rangle}{\|\Sigma^{\zeta/2-1}x\|^2} \|\Sigma^{\zeta-1+\alpha}x\|^2 \leq \langle \Sigma^{\zeta-1+\alpha}x, \Sigma^\alpha v \rangle.$$

This is equivalent to requiring

$$\left\langle v, \frac{\|\Sigma^{\zeta-1+\alpha}x\|^2}{\|\Sigma^{\zeta/2-1}x\|^2} \Sigma^{-1}x - \Sigma^{\zeta+2\alpha-1}x \right\rangle \leq 0.$$

As this must hold for all $v \in \mathcal{V}_{\text{BPS}}$ it is necessary and sufficient to have

$$\frac{\|\Sigma^{\zeta-1+\alpha}x\|^2}{\|\Sigma^{\zeta/2-1}x\|^2} \Sigma^{-1}x = \Sigma^{\zeta+2\alpha-1}x.$$

This holds if and only if $\alpha = -\zeta/2$. \square

8.1.3. Proof of Section 4.3

Proof of Proposition 4.17. Note that refreshment clearly preserves the invariant measure so without loss of generality we may set $\lambda_r = 0$. We first prove that (13) holds by applying

Theorem 3.1. To do so we must determine $L_{\mathfrak{X}}^*$ and Q^* . First let us consider $L_{\mathfrak{X}}$. By Hypothesis 4.16 and Lemma A.2 we have that $e^{-\Phi} \in D(L^*)$ and hence $L_{\mathfrak{X}}^*e^{-\Phi} = -L_{\mathfrak{X}}e^{-\Phi}$.

Similar to Example 3.3 we have that

$$Q^*(z, dy) = \sum_{i=1}^{\infty} \frac{\lambda_i(R_i z)}{\tilde{\lambda}(R_i z)} \delta_{R_i z}.$$

Now we can write the left hand side of (12) as

$$e^{\Phi(z)} L_{\mathfrak{X}}^*(e^{-\Phi(\cdot)})(z) + \int_H \tilde{\lambda}(y) e^{\Phi(z) - \Phi(y)} Q^*(z, dy) - \tilde{\lambda}(z) \tag{98}$$

$$= L_{\mathfrak{X}}(\Phi)(z) + \sum_{i=1}^{\infty} \lambda_i(R_i z) e^{\Phi(z) - \Phi(R_i z)} - \tilde{\lambda}(z) \tag{99}$$

Recall $\Phi(R_i z) = \Phi(x, R_i v) = \Phi(x)$ as R_i only acts on the p component. Therefore (99) simplifies to

$$L_{\mathfrak{X}}(\Phi)(z) + \tilde{\lambda}(R_i z) - \tilde{\lambda}(z).$$

This is equal to zero by (40) since $L_{\mathfrak{X}}\Phi = \langle \nabla_x \Phi, \mathcal{P}p \rangle_{\mathcal{H}}$. By Theorem 3.1 we have that (13) holds for all $f \in C_0(\mathcal{H}^2)$. We can extend this to hold for all $f \in L^2_{\mu}$ by density. \square

8.2. Proofs of Section 5

Proof of Lemma 5.1. In order to verify that Hypothesis 4.16 holds. The only assumption which does not follow immediately is (34). By Cauchy–Schwarz and (45) we have

$$\begin{aligned} \lambda(x) &= \sum_{n=1}^{\infty} \frac{1}{2} ((\nabla_x \Phi(x), v - R_n v)_{\mathcal{H}})^+ \\ &\leq \frac{1}{2} \|\nabla_x \Phi(x)\|_{\mathcal{H}} \sum_{n=1}^{\infty} \|(1 - R_n)v\|_{\mathcal{H}} \\ &\leq \frac{1}{2} \|\nabla_x \Phi(x)\|_{\mathcal{H}} \sqrt{\sum_{n=1}^{\infty} \|(1 - R_n)v\|_{\mathcal{H}}^2} \leq \|\nabla_x \Phi(x)\|_{\mathcal{H}} \|v\|_{\mathcal{H}}. \end{aligned}$$

Then by Assumption 4.1

$$\lambda(x) \lesssim (1 + \|x\|_{\mathcal{H}}) \|v\|_{\mathcal{H}} \tag{100}$$

and (34) follows. \square

Proposition 8.1. Let $\{\tilde{\mathcal{P}}_t\}$ be the semigroup corresponding to the generator (48) and assume that $\tilde{\lambda}_i$ is continuously differentiable for all i . Then for any $t > 0, f \in \mathcal{C}, x, v \in \mathcal{H}$ we have that (53) holds.

Proof of Proposition 8.1. Define the total rate of events to be

$$A_t(x, v) = \sum_{i=1}^{\infty} \int_0^t \tilde{\lambda}_i(\varphi_s(x, v)) ds.$$

Set

$$F_{n,t} = F_{n,t}(x, v) = \mathbb{E}_{x,v}[\mathbb{1}_{[T_n, T_{n+1})}(t)f(X_t, V_t)]$$

then we have

$$\tilde{P}_t f(x, v) = \sum_{n=0}^{\infty} F_{n,t}.$$

Next we observe that for $n \geq 1$ $F_{n,t} = QF_{n-1, \cdot}(t, x, v)$ where

$$\begin{aligned} Qg(t, x, v) &= \mathbb{E}_{x,v}[\mathbb{1}_{T_1 < t}g(t - T_1, X_t, V_t)] \\ &= \sum_{i=1}^{\infty} \int_0^t e^{-A_s(x,v)} \tilde{\lambda}_i(\varphi_s(x, v)) \mathcal{B}_i g(t - s, \varphi_s(x, v)) ds. \end{aligned} \tag{101}$$

We shall make the following inductive hypothesis.

$$\begin{aligned} \nabla F_{n,t} &= \mathbb{E}[\mathbb{1}_{[T_n, T_{n+1})}(t)C_t \nabla f(X_t, V_t)] \\ &\quad - \mathbb{E}\left[\left(\int_0^t C_s \nabla \lambda(X_s, V_s) ds\right) \mathbb{1}_{[T_n, T_{n+1})}(t)f(X_t, V_t)\right] \end{aligned} \tag{102}$$

$$+ \mathbb{E}\left[\sum_{m=1}^n C_{T_m-} \nabla \log(\lambda_{J_m})(X_{T_m}, V_{T_m-}) \mathbb{1}_{[T_n, T_{n+1})}(t)f(X_t, V_t)\right] \tag{103}$$

We shall first show that (102) holds for $n = 0$. For $n = 0$ we can write

$$F_{0,t} = f(\varphi_t(x, v))e^{-A_t(x,v)}.$$

We can differentiate this to obtain

$$\nabla F_{0,t} = \nabla \varphi_t(x, v)(\nabla f)(\varphi_t(x, v))e^{-A_t(x,v)} - f(\varphi_t(x, v))\nabla A_t(x, v)e^{-A_t(x,v)}.$$

Rewriting this in terms of expectations we have

$$\nabla F_{0,t} = \mathbb{E}[\mathbb{1}_{[T_0, T_1)}(t)(\nabla \varphi_t(x, v)(\nabla f)(X_t, V_t) - f(X_t, V_t)\nabla A_t(x, v))].$$

Let us assume that (102) holds for some $n \geq 0$. We can differentiate (101) to find

$$\begin{aligned} \nabla Qg(t, x, v) &= \sum_{i=1}^{\infty} \int_0^t e^{-A_s(x,v)} \tilde{\lambda}_i(\varphi_s(x, v)) \nabla \varphi_s(x, v)(\nabla \mathcal{B}_i g)(t - s, \varphi_s(x, v)) ds \\ &\quad - \sum_{i=1}^{\infty} \int_0^t \nabla A_s(x, v) e^{-A_s(x,v)} \tilde{\lambda}_i(\varphi_s(x, v)) \mathcal{B}_i g(t - s, \varphi_s(x, v)) ds \\ &\quad + \sum_{i=1}^{\infty} \int_0^t e^{-A_s(x,v)} \nabla \varphi_s(x, v) \nabla \tilde{\lambda}_i(\varphi_s(x, v)) \mathcal{B}_i g(t - s, \varphi_s(x, v)) ds \end{aligned}$$

We can rewrite these in terms of expectations using the definition of Q, J_1 .

$$\begin{aligned} \nabla Qg(t, x, v) &= \mathbb{E}_{x,v}[\mathbb{1}_{T_1 < t} \nabla \varphi_{T_1}(x, v)(\nabla \mathcal{B}_{J_1} g)(t - T_1, X_{T_1}, V_{T_1-})] \\ &\quad - \mathbb{E}_{x,v}[\mathbb{1}_{T_1 < t} \nabla A_{T_1}(x, v)g(t - T_1, X_{T_1}, V_{T_1})] \\ &\quad + \mathbb{E}_{x,v}[\mathbb{1}_{T_1 < t} \nabla \varphi_{T_1}(x, v)(\nabla \log \tilde{\lambda}_{J_1})(X_{T_1}, V_{T_1-})g(t - T_1, X_{T_1}, V_{T_1})]. \end{aligned}$$

Now if $J_1 = i$ for some $i \geq 1$ we have

$$(\nabla \mathcal{B}_i g)(s, y, w) = \nabla [g(s, y, R_i w)] = \mathcal{B}'_i(\nabla g)(s, y, R_i w) = \mathcal{B}'_i \mathcal{B}_i(\nabla g)(s, y, w)$$

where

$$B'_i = \begin{pmatrix} 1 & 0 \\ 0 & R_i \end{pmatrix}.$$

With this notation in mind we can write

$$\begin{aligned} \nabla Qg(t, x, v) &= \mathbb{E}_{x,v} \left[\mathbb{1}_{T_1 < t} \nabla \varphi_{T_1}(x, v) B'_{J_1}(\nabla g)(t - T_1, X_{T_1}, V_{T_1}) \right] \\ &\quad - \mathbb{E}_{x,v} \left[\mathbb{1}_{T_1 < t} \nabla A_{T_1}(x, v) g(t - T_1, X_{T_1}, V_{T_1}) \right] \\ &\quad + \mathbb{E}_{x,v} \left[\mathbb{1}_{T_1 < t} \nabla \varphi_{T_1}(x, v) (\nabla \log \tilde{\lambda}_{J_1})(X_{T_1}, V_{T_1-}) g(t - T_1, X_{T_1}, V_{T_1}) \right]. \end{aligned}$$

Now setting $g(t, x, v) = F_{n,t}(x, v)$ and using the inductive hypothesis (102) we have

$$\begin{aligned} \nabla Qg(t, x, v) &= \mathbb{E}_{x,v} \left[\mathbb{1}_{[T_n, T_{n+1})}(t) C_t \nabla f(X_t, V_t) \right] \\ &\quad + \mathbb{E}_{x,v} \left[\sum_{m=2}^n C_{T_m-} \nabla \log(\tilde{\lambda}_{J_m})(X_{T_m}, V_{T_m-}) \mathbb{1}_{[T_n, T_{n+1})}(t) f(X_t, V_t) \right] \\ &\quad - \mathbb{E}_{x,v} \left[\mathbb{1}_{[T_{n+1}, T_{n+2})}(t) \nabla A_{T_1}(x, v) f(X_t, V_t) \right] \\ &\quad - \mathbb{E}_{x,v} \left[\left(\int_{T_1}^t C_{T_1} \nabla \varphi_s \nabla \tilde{\lambda}(X_s, V_s) ds \right) \mathbb{1}_{[T_{n+1}, T_{n+2})}(t) f(X_t, V_t) \right] \\ &\quad + \mathbb{E}_{x,v} \left[\mathbb{1}_{[T_1, T_2)}(t) \nabla \varphi_{T_1}(x, v) (\nabla \log \tilde{\lambda}_{J_1})(X_{T_1}, V_{T_1-}) g(t - T_1, X_{T_1}, V_{T_1}) \right]. \end{aligned}$$

Combining these terms we have that (102) holds for $n + 1$, therefore holds for all $n \geq 0$. \square

Proposition 8.2. *Let $\{\tilde{\mathcal{P}}_t\}$ be the semigroup corresponding to the generator (48) and assume that Φ is twice continuously differentiable with both $\nabla_x \Phi, \nabla_x^2 \Phi$ bounded on bounded sets. For any $f \in C_b^1(\mathcal{H} \times \mathcal{H})$ with $\text{supp}(f) \subseteq B_r$ for some $r > 0$ then there exists a constant $C(r)$ which depends on r and Φ such that (54) holds.*

Proof of Proposition 8.2. We shall define an extended PDMP which will also keep track of component of the latest jump and the pre-multiplier C_t ,

$$\tilde{Z}_t = (X_t, V_t, C_t, J_{T_{N_t}})$$

this is a PDMP with rate $\tilde{\lambda}$, deterministic motion $\tilde{\varphi}_t$ and jumps according to the Markov kernel \tilde{Q} where

$$\begin{aligned} \tilde{\varphi}_t(x, v, c, j) &= (\varphi_t(x, v), c \nabla \varphi_t(x, v), j) \\ \tilde{Q}((x, v, c, j), (dy, dw, d\bar{c}, di)) &= \sum_{n=1}^{\infty} \frac{\tilde{\lambda}_n(x, v)}{\sum_{m=1}^{\infty} \tilde{\lambda}_m(x, v)} \delta_{x, R_n v, c B'_{n,m}}(dy, dw, d\bar{c}, di). \end{aligned}$$

The random counting measure, \hat{p} , determined by \tilde{Z}_t is

$$\hat{p} = \sum_{n=1}^{\infty} \delta_{T_n, \tilde{Z}_{T_n}},$$

where δ_z here represents the Dirac Delta measure which has $\delta_z(\{z\}) = 1$. By [25, Equation (7.32)] the random counting measure \hat{p} has compensating measure \tilde{p} which is given by

$$\tilde{p}(dt, d\tilde{z}) = \tilde{\lambda}(\tilde{\mathcal{Z}}_{t-})\tilde{Q}(\tilde{\mathcal{Z}}_{t-}, d\tilde{z})dt.$$

That is $p - \tilde{p}$ is a local martingale (with filtration given by the natural filtration of $\tilde{\mathcal{Z}}_t$).

Notice that for integrable $g : [0, \infty) \times \mathcal{H} \times \mathcal{H} \times \mathcal{H}^{2 \times 2} \times \mathbb{N} \rightarrow \mathbb{R}$ we have

$$\int gdp = \sum_{n=1}^{\infty} g(T_n, X_{T_n}, V_{T_n}, C_{T_n}, J_n) - \int_0^{\infty} \sum_{m=1}^{\infty} \tilde{\lambda}_m(X_s, V_{s-})g(s, X_s, R_m V_{s-}, C_{s-}B'_m, m)ds.$$

Fix $t > 0$ and set g_t to be

$$g_t(s, x, v, c, j) = cB'_j \nabla \log \tilde{\lambda}_j(x, R_j v) \mathbb{1}_{s \leq t}.$$

With this choice of g we have that

$$\begin{aligned} \int g_t dp &= \sum_{n=1}^{N_t} C_{T_n} B'_n \nabla \log \tilde{\lambda}_{J_n}(X_{T_n}, R_{J_n} V_{T_n}) \\ &\quad - \int_0^t \sum_{m=1}^{\infty} \tilde{\lambda}_m(X_s, V_{s-}) C_{s-} B'_m B'_m \nabla \log \tilde{\lambda}_m(X_t, R_m^2 V_{s-}) ds \\ &= \sum_{m=1}^{N_t} C_{T_m-} \nabla \log \tilde{\lambda}_{J_m}(X_{T_n}, V_{T_n-}) - \int_0^t C_s \nabla \tilde{\lambda}(X_s, V_s) ds = \xi_t \end{aligned}$$

is a local martingale. Moreover, by [25, Proposition 4.6.2] we have

$$\mathbb{E} \left[\left\| \int g dp \right\|_{\mathcal{H} \times \mathcal{H}}^2 \right] = \mathbb{E} \left[\int \|g\|_{\mathcal{H} \times \mathcal{H}}^2 d\tilde{p} \right].$$

Therefore, for this choice of g we have

$$\begin{aligned} \mathbb{E} \left[\|\xi_t\|_{\mathcal{H} \times \mathcal{H}}^2 \right] &= \mathbb{E} \left[\int_0^t \sum_{j=1}^{\infty} \tilde{\lambda}_j(X_s, V_s) |C_s \nabla \log \tilde{\lambda}_j(X_s, V_s)|^2 ds \right] \\ &\leq \mathbb{E} \left[\int_0^t \sum_{j=1}^{\infty} \frac{|\nabla \tilde{\lambda}_j(X_s, V_s)|^2}{\tilde{\lambda}_j(X_s, V_s)} ds \right] \end{aligned} \tag{104}$$

Recall $\tilde{\lambda}_j$ is given by (47) so

$$\frac{|\nabla \tilde{\lambda}_j(x, v)|^2}{\tilde{\lambda}_j(x, v)} = \frac{\psi'(\langle \nabla_x \Phi(x), v - R_j v \rangle_{\mathcal{H}})^2 |\nabla \langle \nabla_x \Phi(x), v - R_j v \rangle_{\mathcal{H}}|^2}{-\log(\phi(\exp(-\langle \nabla_x \Phi(x), v - R_j v \rangle_{\mathcal{H}})))}$$

where

$$\psi(s) = \log(\phi(\exp(-s))).$$

Note that for the choice of $\phi(r) = \frac{r}{1+r}$ we have that for all $s \in \mathbb{R}$

$$0 \leq \frac{\psi'(s)^2}{-\log(\phi(\exp(-s)))} \leq \frac{1}{2}.$$

Using this bound we obtain

$$\frac{|\nabla \tilde{\lambda}_j(x, v)|^2}{\tilde{\lambda}_j(x, v)} \leq \frac{1}{2} |\nabla \langle \nabla_x \Phi(x), v - R_j v \rangle_{\mathcal{H}}|^2 = \frac{1}{2} (|\nabla_x^2 \Phi(x)(v - R_j v)|^2 + |(1 - R_j) \nabla_x \Phi(x)|^2).$$

Summing over j we have and using (45) we have

$$\sum_{j=1}^{\infty} \frac{|\nabla \tilde{\lambda}_j(x, v)|^2}{\tilde{\lambda}_j(x, v)} \leq 2(\|\nabla_x^2 \Phi(x)\|^2 \|v\|^2 + \|\nabla_x \Phi(x)\|^2). \tag{105}$$

Therefore

$$\mathbb{E} [\|\xi_t\|_{\mathcal{H} \times \mathcal{H}}^2] \leq 2\mathbb{E} \left[\int_0^t (\|\nabla_x^2 \Phi(X_s)\|^2 \|V_s\|^2 + \|\nabla_x \Phi(X_s)\|^2) ds \right].$$

Since there are no refreshments, plus both φ_t and R_t preserve the norm we have that $\|(X_s, V_s)\|_{\mathcal{H} \times \mathcal{H}} = \|(x, v)\|_{\mathcal{H} \times \mathcal{H}}$. That is, the process always remains on the sphere centred at zero with radius $\|(x, v)\|_{\mathcal{H} \times \mathcal{H}}$. Moreover as f has support contained within B_r we need only consider $(x, v) \in B_r$ so $(X_s, V_s) \in B_r$. Since we have that $\nabla_x \Phi$, $\nabla_x^2 \Phi$ and V_s are bounded on bounded sets there exists a positive constant $C(r)$ such that

$$\mathbb{E} [\|\xi_t\|_{\mathcal{H} \times \mathcal{H}}^2] \leq C(r)^2 t. \quad \square \tag{106}$$

8.3. Proofs of Section 6.1

Proof of Proposition 6.9. By [18, Theorem 3.7] we have for any $g \in \mathcal{FC}_b^\infty(\mathcal{H})$ and $h = g - \mathcal{A}g$ it holds

$$\int_{\mathcal{H}} g^2 + \langle \Sigma \nabla_x g, \nabla_x g \rangle_{\mathcal{H}} d\pi = \int_{\mathcal{H}} g h d\pi \tag{107}$$

$$\int_{\mathcal{H}} \|\Sigma^{\frac{1}{2}} \nabla_x g\|_{\mathcal{H}}^2 + \text{Tr}[(\Sigma \nabla_x^2 g)^2] + \langle \nabla_x^2 \Phi \Sigma \nabla_x g, \Sigma \nabla_x g \rangle_{\mathcal{H}} d\pi = \int_{\mathcal{H}} (\mathcal{A}g)^2 d\pi. \tag{108}$$

As shown in [18, Remark 3.8] (107) implies

$$\int_{\mathcal{H}} g^2 + \langle \Sigma \nabla_x g, \nabla_x g \rangle_{\mathcal{H}} d\pi \leq \int_{\mathcal{H}} h^2 d\pi. \tag{109}$$

Note that although [18] typically assumes Φ is convex the results we used above do not require convexity. Therefore (73) follows.

Since $\nabla_x^2 \Phi$ is bounded there exists $c_\Phi \geq 0$ such that $\Sigma^{\frac{1}{2}} \nabla_x^2 \Phi(x) \Sigma^{\frac{1}{2}} \geq -c_\Phi$ for all $x \in \mathcal{H}$. Then we can use this bound in (108) to obtain

$$\int_{\mathcal{H}} \|\Sigma^{\frac{1}{2}} \nabla_x g\|_{\mathcal{H}}^2 + \text{Tr}[(\Sigma \nabla_x^2 g)^2] d\pi \leq \int_{\mathcal{H}} (\mathcal{A}g)^2 d\pi + c_\Phi \int_{\mathcal{H}} \langle \nabla_x g, \Sigma \nabla_x g \rangle_{\mathcal{H}} d\pi.$$

It remains to bound the right hand side. For the first term we use that $\mathcal{A}g = g - h$ and $\|g\|_{L^2_\pi} \leq \|h\|_{L^2_\pi}$. For the second term we use (109). These give

$$\int_{\mathcal{H}} \|\Sigma^{\frac{1}{2}} \nabla_x g\|_{\mathcal{H}}^2 + \text{Tr}[(\Sigma \nabla_x^2 g)^2] d\pi \leq (2 + c_\Phi) \int_{\mathcal{H}} h^2 d\pi.$$

It remains to show that (75) holds.

Fix $\varphi \in C^2(\mathcal{H})$, and note that by Young’s inequality we have

$$2\langle \varphi(\Sigma^{\frac{1}{2}} \nabla_x) \Phi, (\Sigma^{\frac{1}{2}} \nabla_x) \varphi \rangle_{L^2_\pi} \leq \varepsilon^{-1} \|(\Sigma^{\frac{1}{2}} \nabla_x) \varphi\|_{L^2_\pi}^2 + \varepsilon \|\varphi(\Sigma^{\frac{1}{2}} \nabla_x) \Phi\|_{L^2_\pi}^2$$

where $\varepsilon > 0$ will be chosen later. Also we have

$$\begin{aligned} 2\langle \varphi(\Sigma^{\frac{1}{2}} \nabla_x) \Phi, (\Sigma^{\frac{1}{2}} \nabla_x) \varphi \rangle_{L^2_\pi} &= \langle (\Sigma^{\frac{1}{2}} \nabla_x) \Phi, (\Sigma^{\frac{1}{2}} \nabla_x)(\varphi^2) \rangle_{L^2_\pi} \\ &= \langle -\mathcal{A} \Phi, \varphi^2 \rangle_{L^2_\pi}. \end{aligned}$$

Now using (60) we get

$$2\langle \varphi(\Sigma^{\frac{1}{2}} \nabla_x) \Phi, (\Sigma^{\frac{1}{2}} \nabla_x) \varphi \rangle_{L^2_\pi} \geq C_1 \|\varphi(\Sigma^{\frac{1}{2}} \nabla_x) \Phi\|_{L^2_\pi}^2 - C_2 \|\varphi\|_{L^2_\pi}^2.$$

Now by setting $\varepsilon = C_1/2$ we get

$$\|\varphi(\Sigma^{\frac{1}{2}} \nabla_x) \Phi\|_{L^2_\pi}^2 \leq \frac{4}{C_1^2} \|(\Sigma^{\frac{1}{2}} \nabla_x) \varphi\|_{L^2_\pi}^2 + \frac{2C_2}{C_1} \|\varphi\|_{L^2_\pi}^2. \tag{110}$$

Choose $\varphi(x) = \sqrt{\delta + \langle (\Sigma^{\frac{1}{2}} \nabla_x)g, (\Sigma^{\frac{1}{2}} \nabla_x)g \rangle_{\mathcal{H}}}$ and observe that

$$\begin{aligned} \|\Sigma^{\frac{1}{2}} \nabla_x \varphi(x)\|_{\mathcal{H}} &= \frac{\| \langle (\nabla_x \Sigma \nabla_x)g, (\Sigma^{\frac{1}{2}} \nabla_x)g \rangle_{\mathcal{H}} \|_{\mathcal{H}}}{\sqrt{\delta + \langle (\Sigma^{\frac{1}{2}} \nabla_x)g, (\Sigma^{\frac{1}{2}} \nabla_x)g \rangle_{\mathcal{H}}}} \\ &\leq \|\nabla_x \Sigma \nabla_x g\|_{\mathcal{H}} \frac{\|\Sigma^{\frac{1}{2}} \nabla_x g\|_{\mathcal{H}}}{\sqrt{\delta + \langle (\Sigma^{\frac{1}{2}} \nabla_x)g, (\Sigma^{\frac{1}{2}} \nabla_x)g \rangle_{\mathcal{H}}}} \\ &\leq \|\nabla_x \Sigma \nabla_x g\|_{\mathcal{H}}. \end{aligned}$$

With this choice of φ we have that (110) gives

$$\left\| \langle (\Sigma^{\frac{1}{2}} \nabla_x)g \rangle_{\mathcal{H}} (\Sigma^{\frac{1}{2}} \nabla_x) \Phi \right\|_{L^2_\pi}^2 \leq \frac{4}{C_1^2} \|\nabla_x \Sigma \nabla_x g\|_{L^2_\pi}^2 + \frac{2C_2}{C_1} \left\| \langle (\Sigma^{\frac{1}{2}} \nabla_x)g \right\|_{L^2_\pi}^2 + \frac{2C_2}{C_1} \delta.$$

Finally by (73),(74) and taking $\delta \rightarrow 0$ we have

$$\left\| \langle (\Sigma^{\frac{1}{2}} \nabla_x)g \rangle_{\mathcal{H}} (\Sigma^{\frac{1}{2}} \nabla_x) \Phi \right\|_{L^2_\pi}^2 \leq \left(\frac{4\kappa_1^2}{C_1^2} + \frac{2C_2}{C_1} \right) \|h\|_{L^2_\pi}^2. \quad \square$$

Proof of Proposition 6.11. Let us first show that (76) holds. Fix $f \in \mathcal{F}C_b^\infty(\mathcal{H} \times \mathcal{H})$ and observe that

$$|\langle BSf, \Pi f \rangle_{L^2_\mu}| = |\langle f, SB^* \Pi f \rangle_{L^2_\mu}|.$$

Since Π is an orthogonal projection onto the kernel of S we can rewrite the right hand side of the above expression as

$$|\langle BSf, \Pi f \rangle_{L^2_\mu}| = |\langle (1 - \Pi)f, SB^* \Pi f \rangle_{L^2_\mu}| \leq \|(1 - \Pi)f\| \|SB^* f\|_{L^2_\mu}.$$

Therefore (76) follows provided we have

$$\|(BS)^* g\|_{L^2_\mu} \leq c_4 \|g\|_{L^2_\mu}. \tag{111}$$

For $f \in \mathcal{F}C_b^\infty(\mathcal{H} \times \mathcal{H})$ let $u_f \in L^2_\pi(\mathcal{H})$ be the function that satisfies $(1 - G)u_f = \Pi f$. This is an abuse of notation as we defined u_g above for $g \in L^2_\pi$, as Πf can be viewed as belonging to L^2_π we should use the notation $u_{\Pi f}$ however to simplify the exposition we drop the Π in

the notation. Then $SB^*f = SAu_f$. As u_f does not depend on v we can write SAu_f as

$$SAu_f = S \langle v, \nabla_x u_f(x) \rangle_{\mathcal{H}} = - \sum_{n=1}^{\infty} \lambda_n^e(x, v) \langle (1 - R_n)v, \nabla_x u_f(x) \rangle_{\mathcal{H}} + \lambda_r(\Pi - 1) \langle v, \nabla_x u_f(x) \rangle_{\mathcal{H}}.$$

Recall that $\lambda_n^e(x, v) = \frac{1}{4} | \langle (1 - R_n)v, \nabla_x \Phi(x) \rangle_{\mathcal{H}} | + \gamma_n(x, v)$ and Π acts by integration with respect to a centred Gaussian measure. This gives

$$SAu_f = -\frac{1}{4} \sum_{n=1}^{\infty} | \langle (1 - R_n)v, \nabla_x \Phi(x) \rangle_{\mathcal{H}} | \langle (1 - R_n)v, \nabla_x u_f(x) \rangle_{\mathcal{H}} - \lambda_r \langle v, \nabla_x u_f(x) \rangle_{\mathcal{H}}.$$

Taking the L_{μ}^2 norm we have

$$\|SAu_f\|_{L_{\mu}^2} \leq \left\| \frac{1}{4} \sum_{n=1}^{\infty} \langle (1 - R_n)v, \nabla_x \Phi(x) \rangle_{\mathcal{H}} \langle (1 - R_n)v, \nabla_x u_f(x) \rangle_{\mathcal{H}} \right\|_{L_{\mu}^2} + \lambda_r \| \langle v, \nabla_x u_f(x) \rangle_{\mathcal{H}} \|_{L_{\mu}^2}.$$

Using that ν_0 is a Gaussian measure with covariance Σ we can write the second term (by Lemma 6.7) as

$$\| \langle v, \nabla_x u_f(x) \rangle_{\mathcal{H}} \|_{L_{\mu}^2}^2 = \int_{\mathcal{H}} \langle \Sigma \nabla_x u_f(x), \nabla_x u_f(x) \rangle_{\mathcal{H}} \pi(dx) = \| \Sigma^{\frac{1}{2}} \nabla_x u_f \|_{L_{\Sigma}^2}^2.$$

By Lemma 6.7 the first term can be written as

$$\begin{aligned} & \left\| \frac{1}{4} \sum_{n=1}^{\infty} \langle v, (1 - R_n)^* \nabla_x \Phi(x) \rangle_{\mathcal{H}} \langle v, (1 - R_n)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \right\|_{L_{\mu}^2}^2 \\ &= \frac{1}{16} \sum_{n,m=1}^{\infty} \int_{\mathcal{H}^2} [\langle v, (1 - R_n)^* \nabla_x \Phi(x) \rangle_{\mathcal{H}} \langle v, (1 - R_m)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \\ & \quad \times \langle v, (1 - R_m)^* \nabla_x \Phi(x) \rangle_{\mathcal{H}} \langle v, (1 - R_n)^* \nabla_x u_f(x) \rangle_{\mathcal{H}}] \mu(dx, dv) \\ &= \frac{1}{16} \sum_{n,m=1}^{\infty} \int_{\mathcal{H}} \langle \Sigma(1 - R_n)^* \nabla_x \Phi(x), (1 - R_n)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \\ & \quad \times \langle \Sigma(1 - R_m)^* \nabla_x \Phi(x), (1 - R_m)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \pi(dx) \\ &+ \frac{1}{16} \sum_{n,m=1}^{\infty} \int_{\mathcal{H}} \langle \Sigma(1 - R_n)^* \nabla_x \Phi(x), (1 - R_m)^* \nabla_x \Phi(x) \rangle_{\mathcal{H}} \\ & \quad \times \langle \Sigma(1 - R_n)^* \nabla_x u_f(x), (1 - R_m)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \pi(dx) \\ &+ \frac{1}{16} \sum_{n,m=1}^{\infty} \int_{\mathcal{H}} \langle \Sigma(1 - R_n)^* \nabla_x \Phi(x), (1 - R_m)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \\ & \quad \times \langle \Sigma(1 - R_m)^* \nabla_x u_f(x), (1 - R_n)^* \nabla_x \Phi(x) \rangle_{\mathcal{H}} \pi(dx) \\ &= \int_{\mathcal{H}} \langle \Sigma \nabla_x \Phi(x), \nabla_x u_f(x) \rangle_{\mathcal{H}}^2 \pi(dx) \\ &+ \frac{1}{16} \sum_{n=1}^{\infty} \int_{\mathcal{H}} \langle \Sigma(1 - R_n)^* \nabla_x \Phi(x), (1 - R_n)^* \nabla_x \Phi(x) \rangle_{\mathcal{H}} \end{aligned}$$

$$\begin{aligned} & \times \langle \Sigma(1 - R_n)^* \nabla_x u_f(x), (1 - R_n)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \pi(dx) \\ & + \frac{1}{16} \sum_{n=1}^{\infty} \int_{\mathcal{H}} \langle \Sigma(1 - R_n)^* \nabla_x \Phi(x), (1 - R_n)^* \nabla_x u_f(x) \rangle_{\mathcal{H}}^2 \pi(dx) \end{aligned}$$

Now using Cauchy–Schwarz we have

$$\begin{aligned} & \left\| \frac{1}{4} \sum_{n=1}^{\infty} \langle v, (1 - R_n)^* \nabla_x \Phi(x) \rangle_{\mathcal{H}} \langle v, (1 - R_n)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \right\|_{L^2_{\mu}} \\ & \leq \int_{\mathcal{H}} \|\Sigma^{\frac{1}{2}} \nabla_x \Phi(x)\|_{\mathcal{H}}^2 \|\Sigma^{\frac{1}{2}} \nabla_x u_f(x)\|_{\mathcal{H}}^2 \pi(dx) \\ & + \frac{1}{8} \sum_{n=1}^{\infty} \int_{\mathcal{H}} \|\Sigma^{\frac{1}{2}}(1 - R_n)^* \nabla_x \Phi(x)\|_{\mathcal{H}}^2 \|\Sigma^{\frac{1}{2}}(1 - R_n)^* \nabla_x u_f(x)\|_{\mathcal{H}}^2 \pi(dx). \end{aligned}$$

Fix a function $F : \mathcal{H} \rightarrow \mathcal{H}$ then by (57) we have the following bound

$$\|\Sigma^{\frac{1}{2}}(1 - R_n)^* F\|_{\mathcal{H}}^2 \leq \sum_{n=1}^{\infty} \|\Sigma^{\frac{1}{2}}(1 - R_n)^* F\|_{\mathcal{H}}^2 \tag{112}$$

$$= \sum_{n=1}^{\infty} \langle (1 - R_n) \Sigma(1 - R_n)^* F, F \rangle_{\mathcal{H}} \tag{113}$$

$$= 4 \|\Sigma^{\frac{1}{2}} F\|_{\mathcal{H}}^2 \tag{114}$$

By setting $F(x) = \nabla_x \Phi(x)$ in (112)–(114) we have $\|\Sigma^{\frac{1}{2}}(1 - R_n)^* \nabla_x \Phi(x)\|_{\mathcal{H}}^2 \leq 4 \|\Sigma^{\frac{1}{2}} \nabla_x \Phi(x)\|_{\mathcal{H}}^2$ which gives

$$\begin{aligned} & \left\| \frac{1}{4} \sum_{n=1}^{\infty} \langle v, (1 - R_n)^* \nabla_x \Phi(x) \rangle_{\mathcal{H}} \langle v, (1 - R_n)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \right\|_{L^2_{\mu}} \\ & \leq \int_{\mathcal{H}} \|\Sigma^{\frac{1}{2}} \nabla_x \Phi(x)\|_{\mathcal{H}}^2 \|\Sigma^{\frac{1}{2}} \nabla_x u_f(x)\|_{\mathcal{H}}^2 \pi(dx) \\ & + \frac{1}{2} \sum_{n=1}^{\infty} \int_{\mathcal{H}} \|\Sigma^{\frac{1}{2}} \nabla_x \Phi(x)\|_{\mathcal{H}}^2 \|\Sigma^{\frac{1}{2}}(1 - R_n)^* \nabla_x u_f(x)\|_{\mathcal{H}}^2 \pi(dx) \\ & \leq 3 \int_{\mathcal{H}} \|\Sigma^{\frac{1}{2}} \nabla_x \Phi(x)\|_{\mathcal{H}}^2 \|\Sigma^{\frac{1}{2}} \nabla_x u_f(x)\|_{\mathcal{H}}^2 \pi(dx). \end{aligned}$$

Therefore we have

$$\|SAu_f\|_{L^2_{\mu}} \leq \sqrt{3} \left\| \left\| \Sigma^{\frac{1}{2}} \nabla_x u_f \right\|_{\mathcal{H}} \left\| \Sigma^{\frac{1}{2}} \nabla_x \Phi(x) \right\|_{L^2_{\pi}(\mathcal{H}; \mathcal{H})} \right\| + \lambda_r \|\Sigma^{\frac{1}{2}} \nabla_x u_f\|_{L^2_{\pi}(\mathcal{H}; \mathcal{H})}.$$

Now by (73) and (75) we can bound these terms by

$$\|SAu_f\|_{L^2_{\mu}}^2 \leq \left(\sqrt{3} \left(\frac{4}{c_1^2} + \frac{c_2}{c_1} \right)^{\frac{1}{2}} + \frac{1}{\sqrt{2}} \lambda_r \right) \|f\|_{L^2_{\mu}}^2.$$

That is (111) holds.

Now we prove that (77) holds. Observe that

$$\begin{aligned} |\langle BA(1 - \Pi)f, \Pi f \rangle_{L^2_\mu}| &= |\langle (1 - \Pi)f, (1 - \Pi)AB^* \Pi f \rangle_{L^2_\mu}| \\ &= |\langle (1 - \Pi)f, (1 - \Pi)A^2 u_f \rangle_{L^2_\mu}| \\ &\leq \|(1 - \Pi)f\|_{L^2_\mu} \|(1 - \Pi)A^2 u_f\|_{L^2_\mu}. \end{aligned}$$

Recall from (71) that

$$\begin{aligned} A^2 u_f &= \langle v, \nabla_x^2 u_f(x)v \rangle_{\mathcal{H}} - \langle x, \nabla_x u_f(x) \rangle_{\mathcal{H}} \\ &\quad - \frac{1}{4} \sum_{n=1}^\infty \langle \nabla_x \Phi(x), (1 - R_n)v \rangle_{\mathcal{H}} \langle (1 - R_n)v, \nabla_x u_f(x) \rangle_{\mathcal{H}}. \end{aligned}$$

Applying $(1 - \Pi)$ gives

$$\begin{aligned} (1 - \Pi)A^2 u_f &= \langle v, \nabla_x^2 u_f(x)v \rangle_{\mathcal{H}} - \text{Tr}(\Sigma \nabla_x^2 u_f(x)) \\ &\quad - \frac{1}{4} \sum_{n=1}^\infty \langle \nabla_x \Phi(x), (1 - R_n)v \rangle_{\mathcal{H}} \langle (1 - R_n)v, \nabla_x u_f(x) \rangle_{\mathcal{H}} \\ &\quad + \frac{1}{4} \sum_{n=1}^\infty \langle \Sigma(1 - R_n)^* \nabla_x \Phi(x), (1 - R_n)^* \nabla_x u_f(x) \rangle_{\mathcal{H}}. \end{aligned}$$

We can simplify this expression using (57)

$$\begin{aligned} (1 - \Pi)A^2 u_f &= \langle v, \nabla_x^2 u_f(x)v \rangle_{\mathcal{H}} - \text{Tr}(\Sigma \nabla_x^2 u_f(x)) \\ &\quad - \frac{1}{4} \sum_{n=1}^\infty \langle \nabla_x \Phi(x), (1 - R_n)v \rangle_{\mathcal{H}} \langle (1 - R_n)v, \nabla_x u_f(x) \rangle_{\mathcal{H}} + \langle \Sigma \nabla_x \Phi(x), \nabla_x u_f(x) \rangle_{\mathcal{H}}. \end{aligned}$$

Using that Π is an orthogonal projection we can write

$$\begin{aligned} &\| \langle v, \nabla_x^2 u_f(x)v \rangle_{\mathcal{H}} - \text{Tr}(\Sigma \nabla_x^2 u_f(x)) \|_{L^2_\mu}^2 \\ &= \int_{\mathcal{H}^2} \langle v, \nabla_x^2 u_f(x)v \rangle_{\mathcal{H}}^2 - \text{Tr}(\Sigma \nabla_x^2 u_f(x))^2 \mu(dx, dv). \end{aligned}$$

By Lemma 6.7 we have

$$\begin{aligned} \int_{\mathcal{H}^2} \langle v, \nabla_x^2 u_f(x)v \rangle_{\mathcal{H}}^2 \mu(dx, dp) &= 2 \int_{\mathcal{H}} \text{Tr}(\nabla_x^2 u_f(x) \Sigma^2 \nabla_x^2 u_f(x)) \pi(dx) \\ &\quad + \int_{\mathcal{H}} \text{Tr}(\Sigma \nabla_x^2 u_f(x))^2 \pi(dx). \end{aligned}$$

Thus

$$\| \langle v, \nabla_x^2 u_f(x)v \rangle_{\mathcal{H}} - \text{Tr}(\Sigma \nabla_x^2 u_f(x)) \|_{L^2_\mu}^2 = 2 \| \Sigma \nabla_x^2 u_f(x) \|_{L^2_\pi(\mathcal{H}; \mathcal{H} \otimes \mathcal{H})}.$$

We can treat the second term similarly, by Lemma 6.7 we write the norm of this term as

$$\begin{aligned} &\| \frac{1}{4} \sum_{n=1}^\infty \langle \nabla_x \Phi(x), (1 - R_n)v \rangle_{\mathcal{H}} \langle (1 - R_n)v, \nabla_x u_f(x) \rangle_{\mathcal{H}} - \langle \Sigma \nabla_x \Phi(x), \nabla_x u_f(x) \rangle_{\mathcal{H}} \|_{L^2_\mu}^2 \\ &= \int_{\mathcal{H}} \int_{\mathcal{H}} \frac{1}{16} \sum_{n,m=1}^\infty \langle \nabla_x \Phi(x), (1 - R_n)v \rangle_{\mathcal{H}} \langle (1 - R_n)v, \nabla_x u_f(x) \rangle_{\mathcal{H}} \\ &\quad \times \langle \nabla_x \Phi(x), (1 - R_m)v \rangle_{\mathcal{H}} \langle (1 - R_m)v, \nabla_x u_f(x) \rangle_{\mathcal{H}} \end{aligned}$$

$$\begin{aligned}
 & - \langle \Sigma \nabla_x \Phi(x), \nabla_x u_f(x) \rangle_{\mathcal{H}}^2 v_0(dv) \pi(dx) \\
 & = \int_{\mathcal{H}} \frac{1}{16} \sum_{n,m=1}^{\infty} \langle \Sigma(1 - R_n)^* \nabla_x \Phi(x), (1 - R_n)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \\
 & \times \langle \Sigma(1 - R_m)^* \nabla_x \Phi(x), (1 - R_m)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \pi(dx) \\
 & + \int_{\mathcal{H}} \frac{1}{16} \sum_{n,m=1}^{\infty} \langle \Sigma(1 - R_n)^* \nabla_x \Phi(x), (1 - R_m)^* \nabla_x \Phi(x) \rangle_{\mathcal{H}} \\
 & \times \langle \Sigma(1 - R_n)^* \nabla_x u_f(x), (1 - R_m)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \pi(dx) \\
 & + \int_{\mathcal{H}} \frac{1}{16} \sum_{n,m=1}^{\infty} \langle \Sigma(1 - R_n)^* \nabla_x \Phi(x), (1 - R_m)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \\
 & \times \langle \Sigma(1 - R_n)^* \nabla_x u_f(x), (1 - R_m)^* \nabla_x \Phi(x) \rangle_{\mathcal{H}} \pi(dx) \\
 & - \int_{\mathcal{H}} \langle \Sigma \nabla_x \Phi(x), \nabla_x u_f(x) \rangle_{\mathcal{H}}^2 \pi(dx) \\
 & = \int_{\mathcal{H}} \frac{1}{16} \sum_{n=1}^{\infty} \langle \Sigma(1 - R_n)^* \nabla_x \Phi(x), (1 - R_n)^* \nabla_x \Phi(x) \rangle_{\mathcal{H}} \\
 & \times \langle \Sigma(1 - R_n)^* \nabla_x u_f(x), (1 - R_n)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \pi(dx) \\
 & + \int_{\mathcal{H}} \frac{1}{16} \sum_{n=1}^{\infty} \langle \Sigma(1 - R_n)^* \nabla_x \Phi(x), (1 - R_n)^* \nabla_x u_f(x) \rangle_{\mathcal{H}} \\
 & \times \langle \Sigma(1 - R_n)^* \nabla_x u_f(x), (1 - R_n)^* \nabla_x \Phi(x) \rangle_{\mathcal{H}} \pi(dx) \\
 & \leq 2 \int_{\mathcal{H}} \|\Sigma^{\frac{1}{2}} \nabla_x \Phi(x)\|_{\mathcal{H}}^2 \|\Sigma^{\frac{1}{2}} \nabla_x u_f(x)\|_{\mathcal{H}}^2 \pi(dx).
 \end{aligned}$$

Thus,

$$\|(1 - \Pi)A^2 u_f\|_{L_{\mu}^2} \leq \sqrt{2} \|\Sigma \nabla_x^2 u_f\|_{L_{\pi}^2} + \sqrt{2} \left\| \|\Sigma^{\frac{1}{2}} \nabla_x u_f(x)\|_{\mathcal{H}} \Sigma^{\frac{1}{2}} \nabla_x \Phi(x) \right\|_{L_{\pi}^2}.$$

By (74) and (75) we can bound this by

$$\|(1 - \Pi)A^2 u_f\|_{L_{\mu}^2} \leq \sqrt{2}(\kappa_1 + \kappa_2) \|\Pi f\|_{L_{\mu}^2}. \quad \square$$

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We wish to thank Andrew Duncan, Michela Ottobre and Andreas Eberle for helpful discussions.

This work is part of the research programme ‘Zigzagging through computational barriers’ with project number 016.Vidi.189.043, which is financed by the Dutch Research Council (NWO).

Appendix. Differentiability in the direction \mathcal{X}

Define the set C to be the set of all bounded $f : \mathcal{H}^2 \rightarrow \mathbb{R}$ for which the following limit is well-defined and bounded:

$$L_{\mathcal{X}} f(x, v) := \lim_{t \rightarrow 0} \frac{1}{t} [f(\varphi_t(x, v)) - f(x, v)] = \frac{d}{dt} \Big|_{t=0} f(\varphi_t(x, v)).$$

Lemma A.1. Assume that (39) holds, then flow φ_t is invariant under the probability measure μ_0 .

Proof of Lemma A.1. Fix $y, w \in \mathcal{H}$ and define the function

$$\psi(x, v) = \exp(i \langle x, y \rangle_{\mathcal{H}} + i \langle v, w \rangle_{\mathcal{H}}).$$

Then since φ_t is a bounded linear operator on the space \mathcal{H}^2 we have

$$\psi(\varphi_t(x, v)) = \exp(i \langle \varphi_t^*(y, w), (x, v) \rangle_{\mathcal{H}^2}).$$

Here φ_t^* denotes the \mathcal{H}^2 adjoint of φ_t , when viewed as a linear operator. Now integrating with respect to μ_0 we have

$$\int_{\mathcal{H}^2} \psi(\varphi_t(x, v)) \mu_0(dx, dv) = \hat{\mu}_0(\varphi_t^*(y, w))$$

where $\hat{\mu}_0$ denotes the Fourier transform of the measure μ_0 . As μ_0 is a Gaussian measure we have

$$\int_{\mathcal{H}^2} \psi(\varphi_t(x, v)) \mu_0(dx, dv) = \exp\left(-\frac{1}{2} \left\langle \begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_v \end{pmatrix} \varphi_t^*(y, w), \varphi_t^*(y, w) \right\rangle_{\mathcal{H}^2}\right).$$

Differentiating this expression we have

$$\begin{aligned} \frac{d}{dt} \int_{\mathcal{H}^2} \psi(\varphi_t(x, v)) \mu_0(dx, dv) &= -\frac{1}{2} \left\langle \left(\begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_v \end{pmatrix} \frac{d}{dt} \varphi_t^*(y, w), \varphi_t^*(y, w) \right) \right\rangle_{\mathcal{H}^2} \\ &+ \left\langle \left(\begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_v \end{pmatrix} \varphi_t^*(y, w), \frac{d}{dt} \varphi_t^*(y, w) \right) \right\rangle_{\mathcal{H}^2} \\ &\times \exp\left(-\frac{1}{2} \left\langle \begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_v \end{pmatrix} \varphi_t^*(y, w), \varphi_t^*(y, w) \right\rangle_{\mathcal{H}^2}\right) \\ &= -\frac{1}{2} \left\langle \left(\begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_v \end{pmatrix} \begin{pmatrix} 0 & -\mathcal{X} \\ \mathcal{P} & 0 \end{pmatrix} \varphi_t^*(y, w), \varphi_t^*(y, w) \right) \right\rangle_{\mathcal{H}^2} \\ &+ \left\langle \left(\begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_v \end{pmatrix} \varphi_t^*(y, w), \begin{pmatrix} 0 & -\mathcal{X} \\ \mathcal{P} & 0 \end{pmatrix} \varphi_t^*(y, w) \right) \right\rangle_{\mathcal{H}^2} \\ &\times \exp\left(-\frac{1}{2} \left\langle \begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_v \end{pmatrix} \varphi_t^*(y, w), \varphi_t^*(y, w) \right\rangle_{\mathcal{H}^2}\right) \\ &= -\frac{1}{2} \left\langle \left(\begin{pmatrix} 0 & -\Sigma_x \mathcal{X} \\ \Sigma_v \mathcal{P} & 0 \end{pmatrix} \varphi_t^*(y, w), \varphi_t^*(y, w) \right) \right\rangle_{\mathcal{H}^2} \\ &+ \left\langle \left(\begin{pmatrix} 0 & \mathcal{P} \Sigma_v \\ -\mathcal{X} \Sigma_x & 0 \end{pmatrix} \varphi_t^*(y, w), \varphi_t^*(y, w) \right) \right\rangle_{\mathcal{H}^2} \\ &\times \exp\left(-\frac{1}{2} \left\langle \begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_v \end{pmatrix} \varphi_t^*(y, w), \varphi_t^*(y, w) \right\rangle_{\mathcal{H}^2}\right) \end{aligned}$$

$$= - \left(\left\langle \begin{pmatrix} 0 & -\Sigma_x \mathcal{X} \\ \Sigma_v \mathcal{P} & 0 \end{pmatrix} \varphi_t^*(y, w), \varphi_t^*(y, w) \right\rangle_{\mathcal{H}^2} \right) \\ \times \exp \left(-\frac{1}{2} \left\langle \begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_v \end{pmatrix} \varphi_t^*(y, w), \varphi_t^*(y, w) \right\rangle_{\mathcal{H}^2} \right)$$

Now for any $x', v' \in \mathcal{H}$ we have

$$\left\langle \begin{pmatrix} 0 & -\Sigma_x \mathcal{X} \\ \Sigma_v \mathcal{P} & 0 \end{pmatrix} \begin{pmatrix} x' \\ v' \end{pmatrix}, \begin{pmatrix} x' \\ v' \end{pmatrix} \right\rangle_{\mathcal{H}^2} = \langle (\mathcal{P} \Sigma_v - \Sigma_x \mathcal{X}) v', x' \rangle_{\mathcal{H}}$$

which is equal to zero by (39). \square

Lemma A.2. Assume (39) holds, then for any $f, g \in C$ then

$$\int_{\mathcal{H}^2} L_{\tilde{x}} f(x, v) g(x, v) \mu_0(dx, dv) = - \int_{\mathcal{H}^2} f(x, v) L_{\tilde{x}} g(x, v) \mu_0(dx, dv).$$

Proof. Recall that $L_{\tilde{x}}$ is the directional derivative with respect to the integral curve φ_t so we have

$$\int_{\mathcal{H}^2} L_{\tilde{x}} f(x, v) g(x, v) \mu_0(dx, dv) = \lim_{t \rightarrow 0} \frac{1}{t} \left(\int_{\mathcal{H}^2} f(\varphi_t(x, v)) g(x, v) \mu_0(dx, dv) - \int_{\mathcal{H}^2} f(x, v) g(x, v) \mu_0(dx, dv) \right).$$

Note that exchanging the limit and the integral is justified by the Dominated Convergence Theorem which is justified since $L_{\tilde{x}} f$ is bounded. Using Lemma A.1

$$\int_{\mathcal{H}^2} L_{\tilde{x}} f(x, v) g(x, v) \mu_0(dx, dv) = \lim_{t \rightarrow 0} \frac{1}{t} \left(\int_{\mathcal{H}^2} f(x, v) g(\varphi_{-t}(x, v)) \mu_0(dx, dv) - \int_{\mathcal{H}^2} f(x, v) g(x, v) \mu_0(dx, dv) \right) \\ = - \int_{\mathcal{H}^2} f(x, v) L_{\tilde{x}} g(x, v) \mu_0(dx, dv).$$

As above exchanging the limit and the integral is justified by the Dominated Convergence Theorem which is justified since $L_{\tilde{x}} g$ is bounded. \square

Corollary A.3. Assume Hypothesis 4.16 holds. Then for any f, g such that $t \mapsto f(\varphi_t(x, v))$ and $t \mapsto g(\varphi_t(x, v))$ are differentiable μ_0 -a.e. then

$$\int_{\mathcal{H}^2} L_{\tilde{x}} f(x, v) g(x, v) \mu(dx, dv) = - \int_{\mathcal{H}^2} f(x, v) L_{\tilde{x}} g(x, v) \mu(dx, dv) \\ + \int \langle \nabla_x \Phi(x), \mathcal{P} p \rangle f(x, v) g(x, v) \mu(dx, dv).$$

References

[1] Christophe Andrieu, Paul Dobson, Andi Q. Wang, Subgeometric hypocoercivity for piecewise-deterministic Markov process Monte Carlo methods, Electron. J. Probab. 26 (2021) 1–26.
 [2] Christophe Andrieu, Alain Durmus, Nikolas Nüsken, Julien Roussel, Hypocoercivity of piecewise deterministic Markov process-Monte Carlo, Ann. Appl. Probab. 31 (5) (2021) 2478–2517.
 [3] Christophe Andrieu, Samuel Livingstone, Peskun–Tierney ordering for Markovian Monte Carlo: Beyond the reversible scenario, Ann. Statist. 49 (4) (2021) 1958–1981.

- [4] Dominique Bakry, Franck Barthe, Patrick Cattiaux, Arnaud Guillin, et al., A simple proof of the Poincaré inequality for a large class of probability measures, *Electron. Commun. Probab.* 13 (2008) 60–66.
- [5] Julien Barré, Paul Dobson, Michela Ottobre, Ewelina Zatorska, Fast non-mean-field networks: Uniform in time averaging, *SIAM J. Math. Anal.* 53 (1) (2021) 937–972.
- [6] Alexandros Beskos, Gareth Roberts, Andrew Stuart, Jochen Voss, MCMC methods for diffusion bridges, *Stoch. Dyn.* 8 (03) (2008) 319–350.
- [7] J. Bierkens, A. Duncan, Limit theorems for the Zig-Zag process, *Adv. Appl. Probab.* 49 (3) (2017).
- [8] Joris Bierkens, Paul Fearnhead, Gareth Roberts, The zig-zag process and super-efficient sampling for Bayesian analysis of big data, *Ann. Statist.* 47 (3) (2019) 1288–1320.
- [9] Joris Bierkens, Sebastiano Grazi, Kengo Kamatani, Gareth Roberts, The boomerang sampler, in: *International Conference on Machine Learning*, PMLR, 2020, pp. 908–918.
- [10] Alexandre Bouchard-Côté, Sebastian J. Vollmer, Arnaud Doucet, The bouncy particle sampler: A nonreversible rejection-free Markov chain Monte Carlo method, *J. Amer. Statist. Assoc.* 113 (522) (2018) 855–867.
- [11] Dan Crisan, Paul Dobson, Michela Ottobre, Uniform in time estimates for the weak error of the Euler method for SDEs and a pathwise approach to derivative estimates for diffusion semigroups, *Trans. Amer. Math. Soc.* 374 (5) (2021) 3289–3330.
- [12] Dan Crisan, Michela Ottobre, Pointwise gradient bounds for degenerate semigroups (of UFG type), in: *Proc. R. Soc. a*, Vol. 472, The Royal Society, 2016, 20160442.
- [13] G. Da Prato, *An Introduction To Infinite-Dimensional Analysis*, in: *Universitext*, Springer Berlin Heidelberg, 2006.
- [14] M.H.A. Davis, Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 46 (3) (1984) 353–388.
- [15] Mark H.A. Davis, *Markov Models & Optimization*, Vol. 49, CRC Press, 1993.
- [16] Jean Dolbeault, Clément Mouhot, Christian Schmeiser, Hypocoercivity for linear kinetic equations conserving mass, *Trans. Amer. Math. Soc.* 367 (6) (2015) 3807–3828.
- [17] Alain Durmus, Arnaud Guillin, Pierre Monmarché, Piecewise deterministic Markov processes and their invariant measures, *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques* 57 (3) (2021) 1442 – 1475.
- [18] Benedikt Eisenhuth, Martin Grothaus, Essential M-dissipativity for Possibly Degenerate Generators of Infinite-dimensional Diffusion Processes, *Integral Equations and Operator Theory* 94 (3) (2022) 28.
- [19] Benedikt Eisenhuth, Martin Grothaus, Hypocoercivity for non-linear infinite-dimensional degenerate stochastic differential equations, *Stochastics and Partial Differential Equations: Analysis and Computations* (2023).
- [20] Klaus-Jochen Engel, Rainer Nagel, One-parameter semigroups for linear evolution equations, in: *Semigroup Forum*, Vol. 63, Springer, 2001, pp. 278–280.
- [21] Stewart N. Ethier, Thomas G. Kurtz, *Markov Processes: Characterization and Convergence*, Vol. 282, John Wiley & Sons, 2009.
- [22] J. Fontbona, H. Guérin, F. Malrieu, Quantitative estimates for the long-time behavior of an ergodic variant of the telegraph process, *Adv. Appl. Probab.* 44 (4) (2012) 977–994.
- [23] Martin Grothaus, Patrik Stilgenbauer, Hypocoercivity for Kolmogorov backward evolution equations and applications, *J. Funct. Anal.* 267 (10) (2014) 3515–3556.
- [24] M. Hairer, A.M. Stuart, J. Voss, A Bayesian approach to data assimilation, *Physica D* (2005).
- [25] Martin Jacobsen, *Point Process Theory and Applications*, in: *Probability and its Applications*, Birkhäuser Boston, Inc., Boston, MA, 2006, p. xii+328.
- [26] Giuseppe Da Prato, Alessandra Lunardi, Sobolev regularity for a class of second order elliptic pde’s in infinite dimension, *Ann. Probab.* 42 (5) (2014) 2113–2160.
- [27] Martin Georg Riedler, *Spatio-Temporal Stochastic Hybrid Models of Biological Excitable Membranes* (Ph.D. thesis), Heriot-Watt University, 2011.
- [28] Martin Riedler, Michèle Thieullen, Gilles Wainrib, Limit theorems for infinite-dimensional piecewise deterministic Markov processes. Applications to stochastic excitable membrane models, *Electron. J. Probab.* 17 (2012).
- [29] Andrew M. Stuart, Inverse problems: a Bayesian perspective, *Acta Numer.* 19 (2010) 451.