GENERATIVE PDMP MODELS

1. Introduction

1.1. **Diffusion models.** Consider a data-set of realisations of an unknown probability distribution π, which is typically a distribution over images. The generative diffusion model proposed in [10] generates new samples from π as follows. For each data point x_0, one considers a diffusion that is ergodic with respect to a base distribution ρ using x_0 as initial condition and for "large" time horizon T. It follows that the law of the diffusion at time T, p_T, satisfies $p_T \approx \rho$. A typical choice of the diffusion is the OU process

$$\mathrm{d}X_t = -X_t\mathrm{d}t + \sqrt{2}\mathrm{d}W_t, \qquad X_0 \sim \pi,$$

where W_t is a Brownian motion. For this diffusion the base distribution is $\rho = N(0, I_d)$. If we reverse time of such diffusion we obtain again a diffusion which maps any initial condition drawn from p_T to a point which is a draw from π. The reverse time diffusion $(\overleftarrow{X}_t)_{t\in[0,T]} = (X_{T-t})_{t\in[0,T]}$ satisfies the SDE

$$\mathrm{d}\overleftarrow{X}_t = \left(\overleftarrow{X}_t - 2\nabla_x \log p_{T-t}(\overleftarrow{X}_t)\right)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t, \qquad \overleftarrow{X}_0 \sim p_T,$$

and satisfies $\overleftarrow{X}_T \sim \pi$. Therefore the reverse diffusion can be used to generate new data points by initialising $\overleftarrow{X}_0 \sim \rho$, as long as we obtain an approximation of the *score*, that is $\nabla \log p_t$.

Here we apply the framework above to a class of non-diffusive stochastic process known as piecewise deterministic Markov processes (PDMPs).

1.2. **Piecewise deterministic Markov processes.** A PDMP [6, 7] defined on a space $(E, \mathcal{E})$ is a Markov process with drift and jumps identified by three *characteristics*:

(1) a flow map $\varphi_t : E \times [0, \infty) \to E$ governing the deterministic motion;
(2) a jump rate $\lambda : E \to [0, \infty)$ governing the random times of jumps;
(3) a jump kernel $Q : E \times \mathcal{E} \to [0, 1]$ which is applied at event times and defines the new location of the process.

Starting at z, a PDMP with characteristics (φ_t, λ, Q) moves in space according to the flow map φ_t, which is the solution of an ODE

$$\frac{d\varphi_t(z)}{\mathrm{d}t} = \Phi(\varphi_t(z)), \quad \varphi_0(z) = z, \tag{1}$$

for random time τ, which has law

$$\mathbb{P}_z(\tau > t) = \exp\left(-\int_0^t \lambda(\varphi_u(z))\mathrm{d}u\right). \tag{2}$$

At time τ the state of the process is drawn from the jump kernel Q evaluated at $Z_{\tau-}$, that is

$$Z_\tau \sim Q(\varphi_{\tau-}(z), \cdot).$$

The generator of such PDMP is given by

$$\overrightarrow{\mathcal{L}} f(z) = \langle \Phi(z), \nabla_z f(z)\rangle + \overrightarrow{\lambda}(z)(\overrightarrow{Q} f(z) - f(z)). \tag{3}$$

where $Qf(z) = \int_E f(y)Q(z, \mathrm{d}y)$.

Date: March 29, 2024.

Example 1.1 (Bouncy particle [4]). The bouncy particle sampler (BPS) is a PDMP with state space is $E = \mathbb{R}^d \times \mathbb{R}^d$. For any $z \in E$, we write $z = (x, v)$ for $x \in \mathbb{R}^d$, $v \in \{+1, -1\}^d$, where x is interpreted as the position of the particle and v denotes the corresponding velocity. The deterministic motion is $\Phi(x, v) = (v, 0)^T$, while there are two types of random events: *reflections* and *refreshments*. Reflections enforce that $\mu(x, v) = \pi(x)\nu(v)$ is the invariant density of the process, where $\pi \propto \exp(-\psi)$ is a distribution of interest and ν is an auxiliary, rotation-invariant probability measure on $\mathbb{R}^d$ (usually the standard Gaussian measure or the uniform measure on the unit sphere $\mathbb{S}^{d-1}$). Reflections have rate $\lambda_1(x, v) = \langle v, \nabla_x \psi(x)\rangle_+$, where $a_+ = \max(0, a)$, while refreshments have rate $\lambda_2(x, v) = \lambda_r$ for $\lambda_r > 0$. The corresponding jump kernels are

$$Q_1((x, v), (\mathrm{d}y, \mathrm{d}w)) = \delta_{(x,R(x)v)}(\mathrm{d}y, \mathrm{d}w), \quad Q_2((x, v), (\mathrm{d}y, \mathrm{d}w)) = \delta_x(\mathrm{d}y)\nu(\mathrm{d}w),$$

where

$$R(x)v = v - 2\frac{\langle v, \nabla_x \psi(x)\rangle}{|\nabla_x \psi(x)|^2}\nabla_x \psi(x).$$

The operator R *reflects* the velocity v off the hyperplane that is tangent to the contour line of ψ passing though point x. The norm of the velocity is unchanged by the application of R, and this gives the interpretation that R is an elastic collision of the particle off such hyperplane. As observed in [4], a strictly positive λ_r is needed to ensure ergodicity of the BPS. The BPS has generator

$$\mathcal{L}f(x, v) = \langle v, \nabla_x f(x, v)\rangle + \lambda_1(x, v)[f(x, R(x)v) - f(x, v)] + \lambda_2 \int \big(f(x, w) - f(x, v)\big)\nu(\mathrm{d}w). \tag{4}$$

Example 1.2 (Zig-Zag process [1]). The Zig-Zag process (ZZP) is a PDMP with state space $E = \mathbb{R}^d \times \{+1, -1\}^d$. The deterministic motion is determined by $\Phi(x, v) = (v, 0)^T$, i.e. the particle travels with constant velocity v. For $i = 1, \dots, d$ we define the jump rates $\lambda_i(x, v) := (v_i \partial_i \psi(x))_+$. The corresponding (deterministic) jump kernels are given by $Q_i((x, v), (\mathrm{d}y, w)) = \delta_{(x,R_i v)}(\mathrm{d}y, w)$, where R_i is the operator that flips the sign of the i-th component of the vector it is applied to, that is $R_i v = (v_1 \dots, v_{i-1}, -v_i, v_{i+1}, \dots, v_d)$. Importantly, the ZZP does not need velocity refreshments to be ergodic [2]. The ZZP falls in our definition of PDMP taking

$$\lambda(x, v) = \sum_{i=1}^{d} \lambda_i(x, v), \quad Q((x, v), (\mathrm{d}y, w)) = \sum_{i=1}^{d} \frac{\lambda_i(x, v)}{\lambda(x, v)}\delta_{(x,R_i v)}(\mathrm{d}y, w).$$

Example 1.3 (Randomised HMC [3]). Randomised Hamiltonian Monte Carlo (HMC) refers to the PDMP with state space $E = \mathbb{R}^d \times \mathbb{R}^d$ which is characterised by Hamiltonian deterministic flow and refreshments of the velocity vector from the standard Gaussian distribution. The flow is described by $\Phi(x, v) = (v, -\nabla\psi_\rho(x))^T$, where ψ_ρ is the potential of ρ. The jump part coincides with the refreshment part of BPS, with rate λ_r and jump kernel $Q_2((x, v), (\mathrm{d}y, \mathrm{d}w)) = \delta_x(\mathrm{d}y)\nu(\mathrm{d}w)$. When the stationary distribution of the velocity part is a standard Gaussian we have deterministic dynamics of the form $X_t = X_0 \cos(\tau) + V_0 \sin(\tau)$ and $V_t = -X_0 \sin(\tau) + V_0 \cos(\tau)$, where (X_0, V_0) is the initial condition.

2. Generative models based on PDMPs

2.1. **Time reversal of PDMPs.** A formula for the time reversal of a PDMP was obtained in [5], giving that under suitable conditions the time reversed process is a non-homogeneous PDMP with generator

$$\overleftarrow{\mathcal{L}}_t f(z) = -\langle\Phi(z), \nabla_z f(z)\rangle + \overleftarrow{\lambda}_{T-t}(z)(\overleftarrow{Q}_{T-t}(z, dy)f(z) - f(z)), \tag{5}$$

where $\overleftarrow{\lambda}_t(z)$ and $\overleftarrow{Q}_t$ are the unique solutions to the balance equation

$$p_t(dy)\overleftarrow{\lambda}_t(y)\overleftarrow{Q}_t(y, dz) = p_t(dz)\overrightarrow{\lambda}_t(z)\overrightarrow{Q}_t(z, dy). \tag{6}$$

where p_t is the law of the process at time t.

Time reversal of BPS. Now let us consider BPS and derive the generator of its time reversal. The balance equation (6) can be split in the two event types in (4). Let us first consider reflections, plugging in the right hand side λ_1, Q_1 we find

$$p_t(dy, dw)\overleftarrow{\lambda}_{1,t}(y, w)\overleftarrow{Q}_{1,t}((y, w), (dx, dv)) = p_t(dx, dv)\delta_{(x,R(x)v)}(dy, dw)\lambda_1(x, v),$$

hence $\overleftarrow{\lambda}_{1,t}(y, w)\overleftarrow{Q}_{1,t}((y, w), (dx, dv)) = 0$ if $x \neq y$ or $w \neq R(x)v$, i.e. $\overleftarrow{Q}_{1,t}$ only acts on the velocity by reflecting it according to $R(x)$, that is

$$\overleftarrow{Q}_{1,t}((y, w), (dx, dv)) = \delta_{(y,R(y)w)}(dx, dv)$$

and thus $\overleftarrow{\lambda}_{1,t}$ should satisfy

$$p_t(dy, dw)\overleftarrow{\lambda}_{1,t}(y, w) = p_t(dy, dR(x)w)\lambda_1(y, R(y)w).$$

Assuming $p_t(dy, dw)$ is absolutely continuous wrt Lebesgue measure this gives

$$\overleftarrow{\lambda}_{1,t}(y, w) = \frac{p_t(y, R(x)w)}{p_t(y, w)}\lambda_1(y, R(y)w).$$

Now we focus on refreshments, in which case we look for $\overleftarrow{\lambda}_{2,t}$, $\overleftarrow{Q}_{2,t}$ that satisfy

$$p_t(dy, dw)\overleftarrow{\lambda}_{2,t}(y, w)\overleftarrow{Q}_{2,t}((y, w), (dx, dv)) = p_t(dx, dv)\lambda_r\delta_x(dy)\nu(dw)$$

which can be rewritten as

$$\overleftarrow{\lambda}_{2,t}(y, w)\overleftarrow{Q}_{2,t}((y, w), (dx, dv)) = \frac{\lambda_r\nu(dw)p_t(dy)}{p_t(dy, dw)}p_t(dv|y)\delta_y(dx)$$

from which we conclude

$$\overleftarrow{Q}_{2,t}((y, w), (dx, dv)) = p_t(v|y)\delta_y(dx)dv \tag{7}$$

$$\overleftarrow{\lambda}_{2,t}(y, w) = \lambda_r\frac{\nu(w)}{p_t(w|y)}. \tag{8}$$

Notice that to simulate the backward refreshment mechanism it is necessary to estimate the conditional distribution $p_t(w|y)$.

Time reversal of ZZP. With similar reasoning as for the case of BPS we find for $i = 1, \ldots, d$

$$\overleftarrow{Q}_{i,t}((y, w), (dx, v)) = \delta_{(y,R_iw)}(dx, v),$$

while $\overleftarrow{\lambda}_{i,t}$ should satisfy

$$p_t(dy, w)\overleftarrow{\lambda}_{i,t}(y, w) = p_t(dy, R_iw)\lambda_i(y, R_iw).$$

Assuming π is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$, one can prove that for any w, $p_t(dy, w) = p_t(y, w)dy$. This gives the condition

$$\overleftarrow{\lambda}_{i,t}(y, w) = \frac{p_t(y, R_iw)}{p_t(y, w)}\lambda_i(y, R_iw).$$

Time reversal of RHMC. RHMC only has refreshment events which coincide with the case of BPS as given by (7) and (8). In this case we can further characterise the density $p_t(v|y)$. First observe that $p_t(v|x) = \frac{p_t(x,v)}{p_t(x)}$. We can express the numerator conveniently conditioning on the amount of time that elapsed since the last velocity refreshment, which we denote by τ. We can then write $p_t(x, v) = \int_0^t p_t(x, v|\tau)p(\tau)d\tau$. Notice in particular that τ here is has law $p(d\tau) = \lambda_r \exp(-\lambda_r\tau)\mathbb{1}_{\tau<t}\, d\tau + \exp(-\lambda_r t)\delta_t(d\tau)$, since the velocity vector at time 0 is anyway "refreshed" at time 0. The conditional distribution of (x, v) given τ reads

$$p_t(x, v|\tau) = p^x_{t-\tau}(\Phi^x_{-\tau}(x, v))\nu(\Phi^v_{-\tau}(x, v))$$

where we introduced the notation $\Phi_t(x,v) = [\Phi^x_t(x,v), \Phi^v_t(x,v)]^T$ and we used the fact that at refreshment time the velocity component is drawn independently of everything else, hence the joint distribution factorises. Integrating this with respect to τ we have

$$p_t(v|x) = \frac{\int_0^t \lambda_r e^{-\lambda_r \tau} p^x_{t-\tau}(\Phi^x_{-\tau}(x,v))\nu(\Phi^v_{-\tau}(x,v))d\tau + e^{-\lambda_r t}\pi_{data}(\Phi^x_{-t}(x,v))\nu(\Phi^v_{-t}(x,v))}{p_t(x)}.$$

3. Learning the rates of time reversed PDMPs

We have seen above that the event rates of the time reversed BPS and ZZS contain the densities of the law of the process p_t. In particular, the rate of time reversed PDMPs are typically of the form

$$\overleftarrow{\lambda}_t(x,v) = \frac{p_t(x,R(x)v)}{p_t(x,v)}\lambda(x,R(x)v),$$

where λ is the rate of the forward PDMP and $R(x)$ is the deterministic operator associated to the jump kernel. We focus on learning the ratio

$$r(x,v,t) = \frac{p_t(x,R(x)v)}{p_t(x,v)} = \frac{p_t(R(x)v|x)}{p_t(v|x)}.$$

This type of problem appears when considering time reversals of jump process, which has been typically considered on discrete spaces [12, 9]. In this section we discuss how we can approximate such ratios.

3.1. **Ratio matching with Bregman divergences.**

We now describe a general approach to approximate ratios of densities based on the minimisation of Bregman divergences [11]. As we shall see, several of the available approaches can be described with this formalism. For a differentiable and strictly convex function f we define the Bregman divergence $BR_f(r,s) = f(r) - f(s) - f'(s)(r-s)$. Then we wish to approximate the true ratio r with a time dependent parametric function $s_\theta : \mathbb{R}^d \times \mathbb{R}^d \times [0,T] \to \mathbb{R}^d_+$ by solving the minimisation problem

$$\min_\theta \int_0^T q(t)\,\mathbb{E}_{p_t}\Big[BR_f(r(X_t,V_t,t), s_\theta(X_t,V_t,t))\Big]\mathrm{d}t, \tag{9}$$

where q is a probability distribution on the time variable and the expectation is with respect to the law p_t. This optimisation problem corresponds to *explicit score matching.* Ignoring terms that do not depend on θ and by a change of variable, assuming the determinant of the Jacobian corresponding to the change of variables $v' = R(x)v$ equals 1, we can rewrite the minimisation as

$$\min_\theta \int_0^T q(t)\mathbb{E}_{p_t}\Big[f'(s_\theta(X_t,V_t,t))s_\theta(X_t,V_t,t) - f(s_\theta(X_t,V_t,t)) - f'(s_\theta(X_t,R(X_t)V-t,t))\Big]\mathrm{d}t. \tag{10}$$

Notably this is independent of the true density ratio and thus it is a formulation analogous to *implicit score matching.* In the next two sections we discuss two alternatives choices for the function f. The objective function can be approximated by a Monte Carlo average, where we first draw $\tau_n \sim q$, and then simulate $(X^n_{\tau_n}, V^n_{\tau_n}) \sim p_{\tau_n}$ with initial condition X^n_0, the n-th datum, and $V^n_0 \sim \nu$. This results in the following minimisation

$$\min_\theta \sum_{n=1}^N \Big(f'(s_\theta(X_{\tau_n},V_{\tau_n},\tau_n))s_\theta(X_{\tau_n},V_{\tau_n},\tau_n) - f(s_\theta(X_{\tau_n},V_{\tau_n},\tau_n)) - f'(s_\theta(X_{\tau_n},R(X_{\tau_n})V_{\tau_n},\tau_n))\Big). \tag{11}$$

3.2. **The case of ZZS.** In the case of ZZS we wish to approximate the ratios

$$r_i(x, v, t) = \frac{p_t(x, R_i v)}{p_t(x, v)} = \frac{p_t(R_i v|x)}{p_t(v|x)}$$

for all $i = 1, \ldots, d$. We wish to learn a function $s_\theta : \mathbb{R}^d \times \{\pm 1\}^d \times [0, T] \to \mathbb{R}_+^d$ which outputs a d dimensional vector containing approximations of each of the ratios. We wish to obtain, for a fixed x, ratios of p_t evaluated at two "neighbouring" points in $v, R_i v$, which differ only in one component.

3.2.1. *Symmetrised ratio matching.* The simpler case of ratios $\frac{p(R_i v)}{p(v)}$ for $v \in \{\pm 1\}^d$ was considered in [8], in which a model $q_\theta(v)$ approximates $p(v)$ by matching these ratios, which have the advantage of being independent of the normalising constant. Here we are interested only in the ratios, hence we do not need to learn the conditional distributions $p_t(v|x)$ first. Adapting the ideas of [8] to our context, we introduce the function $g(r) = \frac{1}{1+r}$ and define the objective function

$$J_{ERM}(\theta) = \sum_{i=1}^d \mathbb{E}_{p_t}\left[(g(s_\theta(X_t, V_t, t)[i]) - g(r_i(X_t, V_t, t)))^2 + (g(s_\theta(X_t, R_i V_t, t)[i]) - g(r_i(X_t, R_i V_t, t)))^2\right] \tag{12}$$

where the subscript stands for explicit ratio matching. The choice of the function g is to improve numerical stability, since the ratios can take very large values, as well as for convenience, since this choice allows us to optimise without knowledge of the true ratios. Similarly, considering the error in the estimation of $r_i(X_t, V_t, t)$ and $r_i(X_t, R_i V_t, t)$ encourages stability of the estimates s_θ. Our approach differs from [8] in two regards: (i) we estimate the ratio directly as opposed to modeling the probabilities $p_t(v|x)$ by some parametric family. This is because we are only interested in the ratios, whereas the goal of [8] is to estimate the data distribution; (ii) we learn ratios of time dependent conditional distribution as opposed to marginal distributions $p(v)$. Then we wish to solve the minimisation problem

$$\min_\theta \int_0^T q(t) J_{ERM}(\theta) \mathrm{d}t. \tag{13}$$

where q is a user-defined distribution on the time variable. Assuming the model s_θ is expressive enough, [8, Theorem 2] ensures that the minimiser of (13) coincides with the true underlying mechanism. With simple computations we obtain the equivalent implicit formulation which minimises the objective function

$$J_{IRM}(\theta) = \sum_{i=1}^d \mathbb{E}_{p_t}\left[\frac{1}{(1 + s_\theta(x, v, t)[i])^2} + \frac{1}{(1 + s_\theta(x, R_i v, t)[i])^2} - \frac{2}{1 + s_\theta(x, v, t)[i]}\right] \tag{14}$$

Proposition 3.1. $\min_\theta J_{ERM}(\theta) = \min_\theta J_{IRM}(\theta)$.

Proof. We find

$$\begin{aligned} J_{ERM}(\theta) = C + \sum_{i=1}^d \mathbb{E}_{p_t}\Big[&g^2(s_\theta(X_t, V_t, t)[i]) + g^2(s_\theta(X_t, R_i V_t, t)[i]) \\ &- 2g(r_i(X_t, V_t, t))g(s_\theta(X_t, V_t, t)[i]) - 2g(s_\theta(X_t, R_i V_t, t)[i])g(r_i(X_t, R_i V_t, t))\Big], \end{aligned}$$

where C is a constant independent of θ. Then plugging in the expression of g we can rewrite the last term as

$$\begin{aligned} \mathbb{E}_{p_t}\Big[g(s_\theta(X_t, R_i V_t, t)[i])g(r_i(X_t, R_i V_t, t))\Big] &= \int \sum_{v \in \{\pm 1\}^d} p_t(x, v) g(s_\theta(X_t, R_i V_t, t)[i]) \frac{p_t(x, R_i v)}{p_t(x, v) + p_t(x, R_i v)} \mathrm{d}x \\ &= \mathbb{E}_{p_t}\Big[g(s_\theta(X_t, V_t, t)[i]) \frac{p_t(x, R_i v)}{p_t(x, v) + p_t(x, R_i v)}\Big]. \end{aligned}$$

Therefore we find

$$\begin{aligned}
J_{ERM}(\theta) &= C + \sum_{i=1}^{d} \mathbb{E}_{p_t}\Big[g^2(s_\theta(X_t, V_t, t)[i]) + g^2(s_\theta(X_t, R_i V_t, t)[i]) \\
&\quad - 2g(s_\theta(X_t, V_t, t)[i])\frac{p_t(x,v)}{p_t(x,v)+p_t(x,R_i v)} - 2g(s_\theta(X_t, V_t, t)[i])\frac{p_t(x,R_i v)}{p_t(x,v)+p_t(x,R_i v)}\Big] \\
&= C + \sum_{i=1}^{d} \mathbb{E}_{p_t}\Big[g^2(s_\theta(X_t, V_t, t)[i]) + g^2(s_\theta(X_t, R_i V_t, t)[i]) - 2g(s_\theta(X_t, V_t, t)[i])\Big].
\end{aligned}$$

□

Corollary 3.2. *The task* $\min_\theta J_{ERM}(\theta)$ *is equivalent to*

$$\min_\theta \sum_{i=1}^{d} \mathbb{E}_{p_t}\Big[BR_f(g(s_\theta(X_t, V_t, t)[i]), g(r(X_t, V_t, t)[i])) + BR_f(g(s_\theta(X_t, R_i V_t, t)[i]), g(r_i(X_t, R_i V_t, t)))\Big]$$

for $f = (t-1)^2/2$.

Proof. Straightforward manipulations of the Bregman divergence show that the objective function in the proposition coincides with J_{IRM} as given in (14). Hence the result follows by Proposition 3.1. □

The finite sample counterpart is

$$\min_\theta \sum_{n=1}^{N}\sum_{i=1}^{d} \left(g^2(s_\theta(X^n_{\tau_n}, V^n_{\tau_n}, \tau_n)[i]) + g^2(s_\theta(X^n_{\tau_n}, V^n_{\tau_n}, \tau_n)[i]) - 2g(s_\theta(X^n_{\tau_n}, V^n_{\tau_n}, \tau_n)[i])\right) \tag{15}$$

where $\tau_1, \ldots, \tau_N \sim q$, while (X^n_t, V^n_t) is the state obtained simulating a ZZP starting at the n-th training data point for a time t. This is a consistent estimator by [8, Theorem 2].

3.3. **KL minimisation.** An alternative approach is given in [9] and focuses on the score entropy: adapted to our case

$$J_{ERM}(\theta) = \sum_{i=1}^{d} w_i \mathbb{E}_{p_t}\left[s_\theta(X_t, V_t, t)[i] - r_i(X_t, V_t, t) - \Big(\log s_\theta(X_t, R_i V_t, t)[i]) - \log r_i(X_t, R_i V_t, t)\Big)\right]. \tag{16}$$

where $\{w_i\}$ is a sequence of weights. Also this objective function considers the error in the estimation of both $r_i(x, v, t)$ and $r_i(x, R_i v, t)$, encouraging stable solutions. Clearly the implicit formulation has objective function

$$J_{IRM}(\theta) = \sum_{i=1}^{d} w_i \mathbb{E}_{p_t}\left[s_\theta(X_t, V_t, t)[i] - \log s_\theta(X_t, R_i V_t, t)[i])\right], \tag{17}$$

which satisfies $\min_\theta J_{IRM}(\theta) = \min_\theta J_{ERM}(\theta)$. In practice we can then consider the finite sample optimisation

$$\min_\theta \sum_{n=1}^{N}\sum_{i=1}^{d} w_i \left(s_\theta(X^n_{\tau_n}, V^n_{\tau_n}, \tau_n)[i] - \log s_\theta(X^n_{\tau_n}, V^n_{\tau_n}, \tau_n)[i]\right). \tag{18}$$

3.4. **Standard approaches.**

3.4.1. *KL minimisation.* Choosing $f(r) = r\log r - r$ we find the implicit formulation

$$\min_\theta \int_0^T q(t) \int p_t(x,v)(s_\theta(x,v,t) - \log s_\theta(x,R(x)v,t))\mathrm{d}x\mathrm{d}v\mathrm{d}t. \tag{19}$$

Assuming $\int p_t(x,v)s_\theta(x,v,t) = 1$, this is equivalent to the explicit formulation

$$\min_\theta \int_0^T q(t)\mathrm{KL}(p_t(x,R(x)v)\|s_\theta(x,v,t)p_t(x,v))\mathrm{d}t. \tag{20}$$

Hence this choice has the interpretation of minimising the KL divergence between the law $p_t(x,R(x)v)$ and $s_\theta(x,v,t)p_t(x,v)$.

3.4.2. *Logistic regression.* If we choose $f(r) = r\log r - (1+r)\log(1+r)$.

3.4.3. *Least squares density ratio matching.* Choosing $f(t) = \frac{(t-1)^2}{2}$, as suggested in [11], we find that (9) writes

$$\min_\theta \int_0^T q(t)\mathbb{E}_{p_t}\left[\frac{1}{2}\left(\frac{p_t(X_t,R(X_t)V_t)}{p_t(X_t,V_t)} - s_\theta(X_t,V_t,t)\right)^2\right]\mathrm{d}t, \tag{21}$$

with implicit formulation given by

$$\min_\theta \int_0^T q(t)\mathbb{E}_{p_t}\left[\frac{1}{2}s_\theta^2(X_t,V_t,t) - s_\theta(X_t,R(X_t)V_t,t)\right]\mathrm{d}t. \tag{22}$$

Note 3.3. Similar to the case of the score estimation, we can derive an alternative formulation which mimics the case of *denoising score matching.* Denote as p_0 the density law at time 0 and as $p_{t|0}$ the conditional density at time t given the initial state at time 0. We can define the joint distribution $p_{0,t} = p_{t|0}p_0$. Then the minimisation problem

$$\min_\theta \int_0^T q(t)\mathbb{E}_{p_{0,t}}\left[\frac{1}{2}\left(\frac{p_{t|0}(X_t,R(X_t)V_t|X_0,V_0)}{p_{t|0}(X_t,V_t|X_0,V_0)} - s_\theta(X_t,V_t,t)\right)^2\right]\mathrm{d}t \tag{23}$$

is equivalent to (21). This can be seen as follows. We only need to worry about the cross product between the density ration and s_θ, since clearly $\int p_t(x,v)s_\theta^2(x,v,t)\mathrm{d}x\mathrm{d}v = \int p_{0,t}(\tilde{x},\tilde{v},x,v)s_\theta^2(x,v,t)\mathrm{d}x\mathrm{d}v\mathrm{d}\tilde{x}\mathrm{d}\tilde{v}$. Now for the cross term we find

$$\begin{aligned}
\int p_t(x,v)s_\theta(x,v,t)\frac{p_t(x,R(x)v)}{p_t(x,v)}\mathrm{d}x\mathrm{d}v &= \int s_\theta(x,v,t)p_t(x,R(x)v)\mathrm{d}x\mathrm{d}v \\
&= \int s_\theta(x,v,t)p_{t|0}(x,R(x)v|\tilde{x},\tilde{v})p_0(\tilde{x},\tilde{v})\mathrm{d}x\mathrm{d}v\mathrm{d}\tilde{x}\mathrm{d}\tilde{v} \\
&= \int s_\theta(x,v,t)\frac{p_{t|0}(x,R(x)v|\tilde{x},\tilde{v})}{p_{t|0}(x,v|\tilde{x},\tilde{v})}p_0(\tilde{x},\tilde{v})p_{t|0}(x,v|\tilde{x},\tilde{v})\mathrm{d}x\mathrm{d}v\mathrm{d}\tilde{x}\mathrm{d}\tilde{v} \\
&= \int s_\theta(x,v,t)\frac{p_{t|0}(x,R(x)v|\tilde{x},\tilde{v})}{p_{t|0}(x,v|\tilde{x},\tilde{v})}p_{0,t}(\tilde{x},\tilde{v},x,v)\mathrm{d}x\mathrm{d}v\mathrm{d}\tilde{x}\mathrm{d}\tilde{v}.
\end{aligned}$$

The requirement for these computations to hold is that $p_{t|0}(x,v|\tilde{x},\tilde{v})$ is non-zero whenever $p_{t|0}(x,R(x)v|\tilde{x},\tilde{v})$ is non-zero (again it is enough that this condition holds for points (x,v) such that $\lambda(x,R(x)v) > 0$).

3.4.4. *Constraints.* There are several constraints that can be imposed on the approximate ratio s_θ.

3.5. **An alternative framework: backward rate matching.** Alternatively we can directly match the backward rate:

$$\min_\theta \int_0^T q(t)\mathbb{E}_{p_t}\left[\frac{1}{2}\left\|\overleftarrow{\lambda}_{1,t}(X_t,V_t) - s_\theta(X_t,V_t,t)\right\|^2\right]\mathrm{d}t, \tag{24}$$

where $q(t)$ is a chosen weight distribution on the time variable. This optimisation problem corresponds to *explicit score matching.* Similarly to score estimation, we can easily show that this optimisation task is equivalent to

$$\min_\theta \int_0^T q(t)\mathbb{E}_{p_t}\left[\frac{1}{2}s_\theta^2(X_t,V_t,t) - \lambda(X_t,V_t)s_\theta(X_t,R(X_t)V_t,t)\right]\mathrm{d}t, \tag{25}$$

which is a formulation analogous to *implicit score matching.* Since we want that $s_\theta(x,v,t) = 0$ whenever $\lambda(x,R(x)v) = 0$, assuming that $\lambda(x,v) > 0$ implies $\lambda(x,R(x)v) = 0$ we can reformulate (25) as

$$\min_\theta \int_0^T q(t)\mathbb{E}_{p_t}\left[\frac{1}{2}s_\theta^2(X_t,V_t,t)\mathbb{1}_{\lambda(X_t,R(X_t)V_t)>0} - \lambda(X_t,V_t)s_\theta(X_t,R(X_t)V_t,t)\mathbb{1}_{\lambda(X_t,V_t)>0}\right]\mathrm{d}t. \tag{26}$$

Learning the rates for BPS. Let us first focus on the reflection rate. We can approximate $\overleftarrow{\lambda}_{1,t}$ by a non-negative, scalar valued function $s_{\theta_1^*}(x,v,t,1)$ where θ_1^* solves the minimisation problem

$$\theta_1^* = \arg\min_\theta \int_0^T q(t)\mathbb{E}_{p_t}\left[\|\overleftarrow{\lambda}_{1,t}(X_t,V_t) - s_\theta(X_t,V_t,t,1)\|^2\right]dt,$$

where $q(t)$ is a chosen weight distribution on the time variable. This is equivalent to

$$\theta_1^* = \arg\min_\theta \int_0^T q(t)\mathbb{E}_{p_t}\left[s_\theta(X_t,V_t,t,1)^2 - 2\lambda_1(X_t,V_t)s_\theta(X_t,R(X_t)V_t,t,1)\right]dt. \tag{27}$$

Indeed

$$\mathbb{E}_{p_t}\left[\overleftarrow{\lambda}_{1,t}(X_t,V_t)s_\theta(X_t,V_t,t,1)\right] = \int \frac{p_t(x,R(x)v)}{p_t(x,v)}\lambda_1(x,R(x)v)s_\theta(x,v,t,1)p_t(x,v)dxdv.$$

Since $\overleftarrow{\lambda}_{1,t}(x,v) = 0$ for (x,v) such that $\langle x,v\rangle > 0$, on such states we impose $s_\theta(x,v,t,1) = 0$. Thus we obtain (27) is equivalent to

$$\theta_1^* = \arg\min_\theta \int_0^T q(t)\mathbb{E}_{p_t}\left[s_\theta(X_t,V_t,t,1)^2\mathbb{1}_{\langle X_t,V_t\rangle<0} - 2\langle X_t,V_t\rangle\, s_\theta(X_t,R(X_t)V_t,t,1)\mathbb{1}_{\langle X_t,V_t\rangle>0}\right]dt. \tag{28}$$

We can then estimate (27) using the data by estimating the expectation with a Monte Carlo average.

Similarly for the refreshment rate we find

$$\theta_2^* = \arg\min_\theta \int_0^T q(t)\left(\mathbb{E}_{p_t}\left[s_\theta(X_t,V_t,t,2)^2\right] - 2\lambda_r\mathbb{E}_{p_t^x\times\nu}\left[s_\theta(X_t,V,t,2)\right]\right)dt. \tag{29}$$

where $p_t^x(x) = \int p_t(x,v)dv$ is the marginal law of the position component.

References

[1] Joris Bierkens, Paul Fearnhead, and Gareth Roberts. The zig-zag process and super-efficient sampling for bayesian analysis of big data. *Annals of Statistics*, 47, 2019.

[2] Joris Bierkens, Gareth O Roberts, and Pierre-André Zitt. Ergodicity of the zigzag process. *The Annals of Applied Probability*, 29(4):2266–2301, 2019.

[3] Nawaf Bou-Rabee and Jesús María Sanz-Serna. Randomized hamiltonian monte carlo. *The Annals of Applied Probability*, 27(4):2159–2194, 2017.

[4] Alexandre Bouchard-Côté, Sebastian J. Vollmer, and Arnaud Doucet. The Bouncy Particle Sampler: A Nonreversible Rejection-Free Markov Chain Monte Carlo Method. *Journal of the American Statistical Association*, 113(522):855–867, 2018.

[5] Giovanni Conforti and Christian Léonard. Time reversal of markov processes with jumps under a finite entropy condition. *Stochastic Processes and their Applications*, 144:85–124, 2022.
[6] M. H. A. Davis. Piecewise-Deterministic Markov Processes: A General Class of Non-Diffusion Stochastic Models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 46(3):353–388, 1984.
[7] M.H.A. Davis. *Markov Models & Optimization*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 1993.
[8] Aapo Hyvärinen. Some extensions of score matching. *Comput. Stat. Data Anal.*, 51:2499–2512, 2007.
[9] Aaron Lou, Chenlin Meng, and Stefano Ermon. Discrete diffusion modeling by estimating the ratios of the data distribution. *arXiv:2310.16834*, 2024.
[10] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021.
[11] Masashi Sugiyama, Taiji Suzuki, and Takafumi Kanamori. Density ratio matching under the bregman divergence: A unified framework of density ratio estimation. *Annals of the Institute of Statistical Mathematics*, 64, 10 2011.
[12] Haoran Sun, Lijun Yu, Bo Dai, Dale Schuurmans, and Hanjun Dai. Score-based continuous-time discrete diffusion models. In *The Eleventh International Conference on Learning Representations*, 2023.