

Irreversible Langevin MCMC on Lie Groups

Alexis Arnaudon¹, Alessandro Barp^{1,2}, and So Takao¹

¹ Imperial College London

² Alan Turing Institute

Abstract. It is well-known that irreversible MCMC algorithms converge faster to their stationary distributions than reversible ones. Using the special geometric structure of Lie groups \mathcal{G} and dissipation fields compatible with the symplectic structure, we construct an irreversible HMC-like MCMC algorithm on \mathcal{G} , where we first update the momentum by solving an OU process on the corresponding Lie algebra \mathfrak{g} , and then approximate the Hamiltonian system on $\mathcal{G} \times \mathfrak{g}$ with a reversible symplectic integrator followed by a Metropolis-Hastings correction step. In particular, when the OU process is simulated over sufficiently long times, we recover HMC as a special case. We illustrate this algorithm numerically using the example $\mathcal{G} = SO(3)$.

Keywords: Hamiltonian Monte Carlo · MCMC · Irreversible Diffusions · Lie Groups · Geometric Mechanics · Langevin Dynamics · Sampling.

1 Introduction

In this work, we construct an irreversible MCMC algorithm on Lie groups, which generalises the standard Hamiltonian Monte Carlo (HMC) algorithm on \mathbb{R}^n . The HMC method [14] generates samples from a probability density (with respect to an appropriate reference measure) known up to a constant factor by generating proposals using Hamiltonian mechanics, which is approximated by a reversible symplectic numerical integrator and followed by a Metropolis-Hastings step to correct for the bias introduced during the numerical approximations. The resulting time-homogeneous Markov chain is thus reversible, and allow distant proposals to be accepted with high probability, which decreases the correlations between samples (for a basic reference on HMC see [18], and for a geometric description see [4, 6]). However, it is well-known that ergodic irreversible diffusions converge faster to their target distributions [15, 21], and several irreversible MCMC algorithms based on Langevin dynamics have been proposed [19, 20].

From a mechanical point of view, diffusions on Lie groups are important since they form the configuration space of many interesting systems, such as the free rigid body. For example in [10] Euler-Poincaré reduction of group invariant symplectic diffusions on Lie groups are considered in view of deriving dissipative equations from a variational principle, and in [1] Langevin systems on coadjoint

orbits are constructed by adding noise and dissipation to Hamiltonian systems on Lie groups. The phase transitions of this system were analysed using a sampling method [2]. In lattice gauge theory one typically uses the HMC algorithm for semi-simple compact Lie groups which was originally presented in [16] and extended to arbitrary Lie groups in [3], see also [11, 12, 17]. In [5], it was shown how to construct HMC on homogeneous manifolds using symplectic reduction, which includes sampling on Lie groups as a special case.

To construct an irreversible algorithm on Lie groups, we first extend Langevin dynamics to general symplectic manifolds \mathcal{M} based on Bismut's symplectic diffusion process [7]. Our generalised Langevin dynamics with multiplicative noise and nonlinear dissipation has the Gibbs measure as the invariant measure, which allows us to design MCMC algorithms that sample from a Lie group \mathcal{G} when we take $\mathcal{M} = T^*\mathcal{G}$. In our Langevin system the irreversible component is determined by Hamiltonian vector fields which are compatible with the symplectic structure, thus avoiding the appearance of divergence terms associated to the volume distortion. We are then free to choose the noise-generating Hamiltonians to best suit the target distribution. Choosing Hamiltonians that only depend on position allows us to proceed with a Strang splitting of the dynamics into a position-dependent OU process in the fibres which can be solved exactly, and a Hamiltonian part which is approximated using a leapfrog scheme, followed by a Metropolis-Hastings acceptance/rejection step in a similar fashion to [8, 19, 20]. Ideally one wants to choose these Hamiltonians to achieve the fastest convergence to stationarity.

On a general manifold, it would be necessary to introduce local coordinates in order to solve the OU process on the fibres, making it difficult to implement. However, since our base manifold is a Lie group, the Maurer-Cartan form defines an isomorphism between the cotangent bundle $T^*\mathcal{G}$ and the trivial bundle $\mathcal{G} \times \mathfrak{g}^*$, which, given an inner-product on \mathfrak{g} , may further be identified with $\mathcal{G} \times \mathfrak{g}$. As a result, one may pull back the OU process on $T_g^*\mathcal{G}$ to an OU process on \mathfrak{g} for any $g \in \mathcal{G}$, thus avoiding the problem of having to choose appropriate charts. Hence on Lie groups, we obtain a practical irreversible MCMC algorithm which generalises the \mathbb{R}^n -version of the irreversible algorithm considered in [19, 20].

Finally, we simulate this algorithm in the special case $\mathcal{G} = SO(3)$ and perform a Maximum Mean Discrepancy (MMD) test to show that on average, the irreversible algorithm converges faster to the stationary measure than the corresponding reversible HMC on $SO(3)$.

2 Diffusions on Symplectic Manifolds

We consider diffusion processes on symplectic manifolds (\mathcal{M}, ω) , where we have a natural volume form ω^n , and define the canonical Poisson bracket $\{g, f\} := \omega(X_f, X_g) = X_g f$, where X_g is the Hamiltonian vector field associated to the Hamiltonian $g : \mathcal{M} \rightarrow \mathbb{R}$, i.e. $dg = \iota_{X_g} \omega$. Given arbitrary functions H and H_i

for $i = 1, \dots, m$ on \mathcal{M} , we consider the SDE

$$dZ_t = \left(X_H(Z_t) - \frac{\beta}{2} \sum_{i=1}^m \{H_i, H\} X_{H_i}(Z_t) \right) dt + \sum_{i=1}^m X_{H_i}(Z_t) \circ dW_t^i, \quad (1)$$

which has the generator, or forward Kolmogorov operator

$$\mathcal{L}f = -\{f, H\} - \frac{\beta}{2} \sum_{i=1}^m \{H, H_i\} \{f, H_i\} + \frac{1}{2} \sum_{i=1}^m \{\{f, H_i\}, H_i\}. \quad (2)$$

To show that the Gibbs measure is an invariant measure for (1), we will need the following lemma:

Lemma 1. *For a symplectic manifold (\mathcal{M}, ω) and two functions $f, g \in C^\infty(\mathcal{M})$ such that either $\partial\mathcal{M} = \emptyset$ or $g|_{\partial\mathcal{M}} = 0$, we have the following identity*

$$\int_{\mathcal{M}} \{f, g\} \omega^n = 0.$$

For the proof, see [13], Section 4.3. Hence (1) enables us to build MCMC algorithms on any symplectic manifold, and in particular the cotangent bundle of Lie groups, that converge to the Gibbs measure:

Theorem 1. *Given a symplectic manifold (\mathcal{M}, ω) without boundary, equation (1) on \mathcal{M} has the Gibbs measure*

$$\mathbb{P}_\infty(z) = p_\infty \omega^n := \frac{1}{Z} e^{-\beta H(z)} \omega^n, \quad Z = \int_{\mathcal{M}} e^{-\beta H(z)} \omega^n$$

as its stationary measure for any choice of $H_i : \mathcal{M} \rightarrow \mathbb{R}$ where $i = 1, \dots, m$.

Proof. Using the Leibniz rule $\{fg, h\} = f\{g, h\} + g\{f, h\}$, we have

$$\begin{aligned} g\{f, H_i\}\{H, H_i\} &= \{gf, H_i\}\{H, H_i\} - f\{g, H_i\}\{H, H_i\} \\ &= \dots = \{gf\{H, H_i\}, H_i\} - f\{g\{H, H_i\}, H_i\} \end{aligned}$$

and similarly

$$g\{\{f, H_i\}, H_i\} = \{g\{f, H_i\}, H_i\} - \{f\{g, H_i\}, H_i\} + f\{\{g, H_i\}, H_i\}.$$

Hence one can compute the $L^2(\mathcal{M}, \omega^n)$ -adjoint of the operator \mathcal{L} as follows

$$\begin{aligned} \int_{\mathcal{M}} g(\mathcal{L}f) \omega^n &= \int_{\mathcal{M}} g \left(-\{f, H\} - \frac{\beta}{2} \sum_{i=1}^m \{f, H_i\}\{H, H_i\} + \frac{1}{2} \sum_{i=1}^m \{\{f, H_i\}, H_i\} \right) \omega^n \\ &= \int_{\mathcal{M}} \left(-\{fg, H\} - \frac{\beta}{2} \{fg\{H, H_i\}, H_i\} + \frac{1}{2} (\{g\{f, H_i\}, H_i\} - \{f\{g, H_i\}, H_i\}) \right) \omega^n \\ &\quad + \int_{\mathcal{M}} f \left(\{g, H\} + \frac{\beta}{2} \{g\{H, H_i\}, H_i\} + \frac{1}{2} \{\{g, H_i\}, H_i\} \right) \omega^n \\ &= \int_{\mathcal{M}} f \left(\{g, H\} + \frac{\beta}{2} \{g\{H, H_i\}, H_i\} + \frac{1}{2} \{\{g, H_i\}, H_i\} \right) \omega^n = \int_{\mathcal{M}} f(\mathcal{L}^*g) \omega^n, \end{aligned}$$

where we have used Lemma 1 to integrate the Poisson brackets to 0. Hence, we obtain the Fokker-Planck operator

$$\mathcal{L}^*g = \{g, H\} + \frac{\beta}{2} \sum_{i=1}^m \{g\{H, H_i\}, H_i\} + \frac{1}{2} \sum_{i=1}^m \{\{g, H_i\}, H_i\}.$$

Now, by the derivation property of the Poisson bracket, $\{f \circ g, h\} = f' \circ g \{g, h\}$ and noting that $p'_\infty(H) = -Z^{-1} \beta e^{-\beta H} = -\beta p_\infty(H)$, one can check that

$$\begin{aligned} \mathcal{L}^*p_\infty &= p'_\infty(H)\{H, H\} + \frac{\beta}{2} \sum_{i=1}^m \{p_\infty(H)\{H, H_i\}, H_i\} \\ &\quad - \frac{\beta}{2} \sum_{i=1}^m \{p_\infty(H)\{H, H_i\}, H_i\} = 0. \end{aligned}$$

Therefore $\mathbb{P}_\infty(z) = p_\infty(z)\omega^n$ is indeed an invariant measure for (1).

If $\mathcal{M} = T^*\mathcal{Q}$ is the cotangent bundle of a manifold without boundary \mathcal{Q} , we define the marginal measure \mathbb{P}_∞^1 on \mathcal{Q} by

$$\int_A \mathbb{P}_\infty^1 = \int_{T^*A} \iota^* \mathbb{P}_\infty, \quad (3)$$

for any measurable set $A \subset \mathcal{Q}$, where $\iota : T^*A \rightarrow T^*\mathcal{Q}$ is the inclusion map. In addition, if (\mathcal{Q}, γ) is a Riemannian manifold, we can consider the Hamiltonian function $H(q, p) = \frac{1}{2}\gamma_q(p, p) + V(q)$, for $(q, p) \in T^*\mathcal{Q}$, and the marginal invariant measure $\mathbb{P}_\infty^1(dq)$ of the process (1) is simply

$$\mathbb{P}_\infty^1(dq) = \frac{1}{Z_1} e^{-V(q)} \sqrt{|g|} dq, \quad Z_1 = \int_{\mathcal{Q}} e^{-V(q)} \sqrt{|g|} dq,$$

where $\sqrt{|g|}dq$ is the Riemannian volume form.

The MCMC algorithm which we will derive in section §3 is based on a Strang splitting of the dynamics (1) into a Hamiltonian part and a Langevin part. Hereafter, we identify $T^*\mathcal{Q}$ with $T\mathcal{Q}$ through the metric and just consider the dynamics on $T\mathcal{Q}$ instead of $T^*\mathcal{Q}$.

3 Irreversible Langevin MCMC on Lie Groups

Consider a n dimensional Lie group \mathcal{G} and let $e_i, \theta^i, i = 1, \dots, n$ be an orthonormal basis of left-invariant vector fields and dual one-forms respectively. We consider $H = V \circ \pi + T : T\mathcal{G} \rightarrow \mathbb{R}$, where T is the kinetic energy associated to a bi-invariant metric on \mathcal{G} and $V \propto \log \chi : \mathcal{G} \rightarrow \mathbb{R}$ is the potential energy, where χ is the distribution we want to sample from on \mathcal{G} . We let $v^i : T\mathcal{G} \rightarrow \mathbb{R}$

be the fibre coordinate functions with respect to the left-invariant vector fields, $v^i(g, u_g) := \theta_g^i(u_g)$.

Vector fields tangent to $T\mathcal{G}$ (i.e., elements of $\Gamma(TT\mathcal{G})$) can be expanded in terms of left-invariant vector fields e_i and the fibre-coordinate vector fields ∂_{v^i} , (i.e., $\Gamma(TT\mathcal{G}) \cong \Gamma(T\mathcal{G} \oplus T\mathfrak{g})$). We consider noise Hamiltonians that depend only on position, $H_i = U_i \circ \pi$ where $U_i : \mathcal{G} \rightarrow \mathbb{R}$, so the corresponding Hamiltonian vector fields can be written as $X_{H_i} = -e_j(U_i)\partial_{v^j}$, (see [3]). Hence the stochastic process (1) on $T\mathcal{G}$ can be split up into a Langevin part

$$\begin{aligned} dZ_t &= \frac{\beta}{2} \sum_{i=1}^m X_H(H_i)X_{H_i}(Z_t)dt + \sum_{i=1}^m X_{H_i}(Z_t) \circ dW_t^i, \\ &= -\frac{\beta}{2} \sum_{i=1}^m v^k e_k(U_i)e_j(U_i)\partial_{v^j}(Z_t)dt - \sum_{i=1}^m e_j(U_i)\partial_{v^j}(Z_t) \circ dW_t^i. \end{aligned} \quad (4)$$

and a Hamiltonian part

$$\frac{dZ_t}{dt} = X_H(Z_t). \quad (5)$$

Note the geodesics are given by the one-parameter subgroups, with Hamiltonian vector field $X_T = v^k e_k$. Since X_{H_i} only has components in the fibre direction ∂_{v^i} (i.e., it has no e_i components along \mathcal{G}) the diffusion starting at any point $(g, v) \in T\mathcal{G}$ remains in $T_g\mathcal{G}$, i.e., with the same base point g . When $\mathcal{G} = \mathbb{R}^n$ and we use the standard kinetic energy $T = \frac{1}{2}\|v\|_{\mathbb{R}^n}^2$, then vector fields become gradients, i.e. $e_j = \partial_{q^j}$, and equation (4) becomes the space dependent Langevin equation for $(q, v) \in T\mathbb{R}^n$,

$$\dot{q} = 0, \quad dv_t = -\frac{\beta}{2} \nabla_q U_i(q) \nabla_q U_i(q)^T v_t dt - \nabla_q U_i(q) \circ dW_t^i. \quad (6)$$

Now let $\xi_i := e_i(1)$ be a basis of the Lie algebra \mathfrak{g} , where 1 is the identity. Then $e_i(g) = \partial_1 L_g \xi$ and we can identify $T\mathcal{G}$ with $\mathcal{G} \times \mathfrak{g}$ through the relation $(g, v^i e_i(g)) \sim (g, v^i \xi_i)$. In other words, we may now think of v^i as the Lie algebra coordinate functions $v^i : \mathcal{G} \times \mathfrak{g} \rightarrow \mathbb{R}$ with $v^i(g, u) = \theta_1^i(u)$, and since \mathfrak{g} is a vector space, we can identify $\partial_{v^i} \sim \xi_i$ and write

$$X_{H_i}(g, v) = -e_j|_g(U_i)\xi_j =: \sigma_{ji}(g)\xi_j. \quad (7)$$

for $i = 1, \dots, m$ and $j = 1, \dots, n$. For matrix Lie groups, this becomes

$$X_{H_i}(g, v) = -\text{Tr}(\nabla U_i^T g \xi_j)\xi_j = \sigma_{ji}(g)\xi_j, \quad (8)$$

where $(\nabla U_i)_{ab} := \partial_{x_{ab}} U_i$, where x_{ab} are the matrix coordinates of $g \in \mathcal{G}$. The Langevin equation (4) can then be written as

$$\dot{g} = 0, \quad dv_t = -\frac{\beta}{2} (\sigma(g)\sigma(g)^T)_{jk} v_t^k \xi_j dt + \sigma_{ji}(g)\xi_j dW_t^i, \quad (9)$$

and if we identify $\mathbf{g} \sim \mathbb{R}^n$, $v^i \xi_i \sim \mathbf{v} \in \mathbb{R}^n$, we get a standard OU process on \mathbb{R}^n

$$d\mathbf{v}_t = -\frac{\beta}{2} \sigma \sigma^T \mathbf{v}_t dt + \sigma d\mathbf{W}_t, \quad (10)$$

for the vector-valued Wiener process $\mathbf{W}_t = (W_t^1, \dots, W_t^m) \in \mathbb{R}^m$. Note that the noise term can be interpreted as an Itô integral since the diffusion coefficient σ does not depend on \mathbf{v} . We can solve this Langevin equation explicitly if the matrix $D = \sigma \sigma^T$ is invertible. This is the case if the vectors ∇U_i form an orthonormal basis of \mathbb{R}^n , or more generally if they satisfy the Hörmander condition. The explicit solution is then given by

$$v_{t+h} = e^{-\frac{\beta}{2} D h} v_t + \sigma \int_t^{t+h} e^{-\frac{\beta}{2} D(t+h-s)} d\mathbf{W}_s, \quad (11)$$

which has the transition probability

$$p(v_0, v) = \frac{1}{(2\pi)^{n/2} |\det \Sigma_h|} \exp\left(-\frac{1}{2} \left\| v - e^{-\frac{\beta}{2} D h} v_0 \right\|_{\Sigma_h^{-1}}^2\right), \quad (12)$$

where $\Sigma_h = \frac{1}{\beta} (\text{Id} - e^{-\beta D h})$.

3.1 MCMC algorithm on Lie groups

From the Langevin system considered in the previous section, we can construct the following MCMC algorithm to sample from the distribution $\chi := e^{-\frac{\beta}{2} V}$ on \mathcal{G} . Starting from $(g_0, v_0) \in T\mathcal{G}$, we iterate the following

1. Solve equation (9) exactly until time h by sampling

$$v^* \simeq \mathcal{N}\left(e^{-\frac{\beta}{2} D h} v_0, \Sigma_h\right) \quad (13)$$

to obtain $(\bar{g}_0, \bar{v}_0) = (g_0, v^*)$;

2. Approximate the Hamiltonian system (5) using N Leapfrog trajectories with step size $\delta > 0$. Starting at $(\bar{g}_0, \bar{v}_0) = (g_0, v^*)$:

For $k = 0, \dots, N - 1$:³

$$\begin{aligned} \bar{v}_{k+\frac{1}{2}} &= \bar{v}_k - \frac{\delta}{2} \text{Tr}(\partial_x V^T \bar{g}_k \xi_i) \xi_i \\ \bar{g}_{k+1} &= \bar{g}_k \exp\left(\delta \bar{v}_{k+\frac{1}{2}}\right) \\ \bar{v}_{k+1} &= \bar{v}_{k+\frac{1}{2}} - \frac{\delta}{2} \text{Tr}(\partial_x V^T \bar{g}_{k+1} \xi_i) \xi_i \end{aligned}$$

to obtain (\bar{g}_N, \bar{v}_N) . The time step δ and number of steps N are to be tuned appropriately by the users.

³ For a non-matrix group, simply replace $\text{Tr}(\partial_x V^T \bar{g}_k \xi_i)$ with $e_i|_g(V)$.

3. Accept or reject the proposal by a Metropolis-Hastings step. We accept the proposal (\bar{g}_N, \bar{v}_N) with probability

$$\alpha = \min \{1, \exp(-H(\bar{g}_N, \bar{v}_N) + H(\bar{g}_0, \bar{v}_0))\},$$

and set $(g_1, v_1) = (\bar{g}_N, \bar{v}_N)$. On the other hand, if the proposal is rejected, we set $(g_1, v_1) = (\bar{g}_0, -\bar{v}_0)$.

Notice also that in the limit $h \rightarrow \infty$, this algorithm become the standard HMC algorithm, which is reversible, but for finite h , the algorithm is irreversible (see [19] for a detailed discussion).

4 Example on $SO(3)$

As an example we will pick the rotation group $\mathcal{G} = SO(3)$, where the Lie algebra $\mathfrak{so}(3)$ consists of 3×3 anti-symmetric matrices so that the kinetic energy on the Lie algebra is $T(v) = \frac{1}{2}\text{Tr}(v^T v)$, for $v \in \mathfrak{so}(3)$. The potentials on $SO(3)$ will be defined by functions $V, U_i : \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}$ on matrices. We then pick the potential to be of the form $V(g) = e^{\alpha \text{Tr}(g)}$ and isotropic noise with

$$U_i(g) = \epsilon \text{Tr}(e^{-\xi_i} g) \quad \text{for} \quad \text{span}(\xi_i) = \mathfrak{so}(3). \quad (14)$$

We then obtain samples g_t on $SO(3)$, which we can project onto the sphere by simply letting the group act on a vector $z = (0, 0, 1)$, to get $x_t = g_t z$, see panel (a) of figure 1. From these samples g_t , we can also estimate the convergence rate of the MCMC algorithm by computing the maximum mean discrepancy (MMD) between the set of first N samples and the whole sequence, using the values on the diagonal of the matrices g_t . We see that in figure 1, small values for h give MCMC algorithms with a faster convergence rate than the HMC limit, i.e., $h \rightarrow \infty$. This is a direct consequence from the irreversibility of the algorithm, as explained in [15, 21]. Even if a faster convergence is desirable, one has to ensure that the correct distribution is sampled, and if h is taken too small, the algorithm will be close to a pure Hamiltonian dynamics, with additional reversal steps $v \rightarrow -v$ when the proposed state is rejected. We observe this effect already for $h = 0.01$ where the convergence is as fast as $h = 0.1$ for the first steps of the chain, but then later slows down, as the distribution is not sampled correctly.

Acknowledgements The authors thank A. Duncan for the useful insights that helped improve this work. AA acknowledges EPSRC funding through award EP/N014529/1 via the EPSRC Centre for Mathematics of Precision Healthcare. ST acknowledges the Schrödinger scholarship scheme for funding during this work. AB was supported by a Roth Scholarship funded by the Department of Mathematics, Imperial College London, by EPSRC Fellowship (EP/J016934/3) and by The Alan Turing Institute under the EPSRC grant [EP/N510129/1].

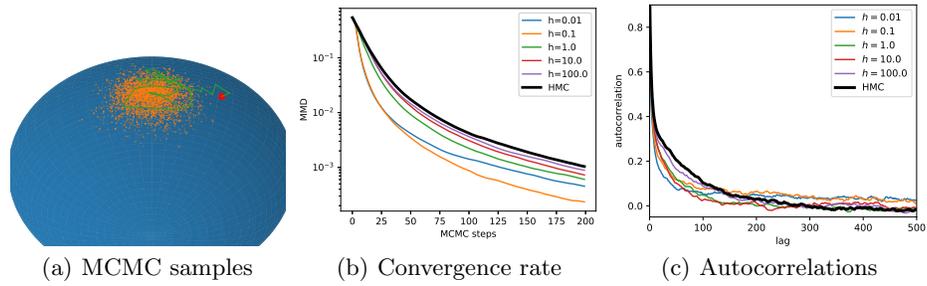


Fig. 1. This figure illustrates the irreversible MCMC algorithm on Lie group with $\mathcal{G} = SO(3)$. For each run, we sampled 5000 samples, shown in panel (a), along with the first one in red and the next 50 in green. We ran several chains with several parameters h , corresponding to the integration time of the Langevin dynamics. For large h , the MCMC algorithm converges to the HMC algorithm, also displayed in black. Each line in panel (b) and (c) are averages over 20 chains with the same parameters. For the leapfrog integrator we use 5 timesteps with a total time $\delta = 0.5$. The MMD computation on \mathcal{G} has been run with Gaussian kernel of variance $\sigma = 1$.

Bibliography

- [1] Alexis Arnaudon, Alex L De Castro, and Darryl D Holm, *Noise and dissipation on coadjoint orbits*, Journal of nonlinear science **28** (2018), no. 1, 91–145.
- [2] Alexis Arnaudon and So Takao, *Networks of coadjoint orbits: from geometric to statistical mechanics*, arXiv preprint arXiv:1804.11139 (2018).
- [3] Alessandro Barp, *Hamiltonian monte carlo on lie groups and constrained mechanics on homogeneous manifolds*, arXiv preprint arXiv:1903.04662 (2019).
- [4] Alessandro Barp, François-Xavier Briol, Anthony D Kennedy, and Mark Girolami, *Geometry and dynamics for markov chain monte carlo*, Annual Review of Statistics and Its Application **5** (2018), 451–471.
- [5] Alessandro Barp, A. D. Kennedy, and Mark Girolami, *Hamiltonian monte carlo on symmetric and homogeneous spaces via symplectic reduction*, arXiv preprint arXiv:1903.02699 (2019).
- [6] Michael Betancourt, Simon Byrne, Sam Livingstone, Mark Girolami, et al., *The geometric foundations of hamiltonian monte carlo*, Bernoulli **23** (2017), no. 4A, 2257–2298.
- [7] JM Bismut, *Mécanique aléatoire. lect*, Notes. Maths **966** (1981), 19881.
- [8] Nawaf Bou-Rabee and Eric Vanden-Eijnden, *A patch that imparts unconditional stability to certain explicit integrators for SDEs*, arXiv preprint arXiv:1010.4278 (2010).
- [9] ———, *Pathwise accuracy and ergodicity of metropolized integrators for sdes*, Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences **63** (2010), no. 5, 655–696.
- [10] Xin Chen, Ana Bela Cruzeiro, and Tudor S Ratiu, *Constrained and stochastic variational principles for dissipative equations with advected quantities*, arXiv preprint arXiv:1506.05024 (2015).
- [11] MA Clark and AD Kennedy, *Accelerating staggered-fermion dynamics with the rational hybrid monte carlo algorithm*, Physical Review D **75** (2007), no. 1, 011502.
- [12] Michael A Clark, Anthony D Kennedy, and PJ Silva, *Tuning hmc using poisson brackets*, arXiv preprint arXiv:0810.1315 (2008).
- [13] Ana Bela Cruzeiro, Darryl D Holm, and Tudor S Ratiu, *Momentum maps and stochastic clebsch action principles*, Communications in Mathematical Physics **357** (2018), no. 2, 873–912.
- [14] Simon Duane, Anthony D Kennedy, Brian J Pendleton, and Duncan Roweth, *Hybrid monte carlo*, Physics letters B **195** (1987), no. 2, 216–222.
- [15] Andrew B Duncan, Tony Lelièvre, and GA Pavliotis, *Variance reduction using nonreversible Langevin samplers*, Journal of statistical physics **163** (2016), no. 3, 457–491.
- [16] A. D. Kennedy and Pietro Rossi, *Classical Mechanics on Group Manifolds*, Nucl. Phys. **B327** (1989), 782–790.
- [17] A. D. Kennedy, P. J. Silva, and M. A. Clark, *Shadow Hamiltonians, Poisson Brackets, and Gauge Theories*, Physical Review **D87** (2013), no. 3, 034511.
- [18] Radford M Neal et al., *MCMC using Hamiltonian dynamics*, Handbook of Markov Chain Monte Carlo **2** (2011), no. 11, 2.
- [19] Michela Ottobre, *Markov Chain Monte Carlo and irreversibility*, Reports on Mathematical Physics **77** (2016), no. 3, 267–292.

- [20] Michela Ottobre, Natesh S Pillai, Frank J Pinski, Andrew M Stuart, et al., *A function space HMC algorithm with second order Langevin diffusion limit*, Bernoulli **22** (2016), no. 1, 60–106.
- [21] Luc Rey-Bellet and Konstantinos Spiliopoulos, *Irreversible langevin samplers and variance reduction: a large deviations approach*, Nonlinearity **28** (2015), no. 7, 2081.