www.csgb.dk

RESEARCH REPORT

2019

Mathias Højgaard Jensen, Anton Mallasto and Stefan Sommer

Simulation of Conditioned Diffusions on the Flat Torus

Simulation of Conditioned Diffusions on the Flat Torus

Mathias Højgaard Jensen, Anton Mallasto and Stefan Sommer

Department of Computer Science, University of Copenhagen, Denmark
{matje,mallasto,sommer}@di.ku.dk

Abstract

Diffusion processes are fundamental in modelling stochastic dynamics in natural sciences. Recently, simulating such processes on complicated geometries has found applications for example in biology, where toroidal data arises naturally when studying the backbone of protein sequences, creating a demand for efficient sampling methods. In this paper, we propose a method for simulating diffusions on the flat torus, conditioned on hitting a terminal point after a fixed time, by considering a diffusion process in $\mathbb{R}^2$ which we project onto the torus. We contribute a convergence result for this diffusion process, translating into convergence of the projected process to the terminal point on the torus. We also show that under a suitable change of measure, the Euclidean diffusion is locally a Brownian motion.

Keywords: Simulation, Conditioned diffusion, Manifold diffusion, Flat Torus

1   Introduction

Stochastic differential equations are ubiquitous in models describing evolution of dynamical systems with, e.g. in modelling the evolution of DNA or protein structure, in pricing financial derivatives, or for modelling changes in landmark configurations which are essential in shape analysis and computational anatomy. In settings where the beginning and end values are known on some fixed time interval, the use of Brownian bridges becomes natural to evaluate the uncertainty on the intermediate time interval.

When the data elements are elements of non-linear spaces, here differentiable manifolds, methodology for simulating bridge processes is lacking. In particular, in cases where the transition probability densities are intractable, it is of interest to use simulation schemes that can numerically approximate the true densities. In this paper we propose a method for simulating diffusion bridges on the flat torus, $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, i.e. we propose a process that can easily be simulated and satisfies that the distribution of the true bridge of interest is absolutely continuous with respect to the distribution of this proposal process. This specific case will serve

as an example of the more general setting of simulating diffusion bridge processes on Riemannian manifolds. Because of the non-trivial topology of the torus $\mathbb{T}^2$, the conditioned process will be equivalent to a process in $\mathbb{R}^2$ that is conditioned on ending up in a set of points. Therefore, we will address the question of conditioning a process on infinitely many points. Secondly, we will handle the case when the process crosses the cut locus of the target point, i.e. the set of points with no unique distance minimizing geodesic.

It is a basic consequence of Doob's h-transform that the distribution of a conditioned diffusion process is the same as another diffusion process with the drift depending on the transition density. However, as mentioned in [1], using this transform directly is undesirable for simulation purposes as the transition density is often intractable. Instead, the authors introduce a diffusion process which can easily be simulated and with the property that the distribution of the true conditioned diffusion is absolutely continuous wrt. the diffusion used for simulation. We here use this approach that in [1] covers the Euclidean case as the starting point for developing a simulation scheme on the torus.

Recent papers have considered diffusion processes on the torus, for example, Langevin diffusions on the torus were studied in [3] and [4], in the latter to describe protein evolution. In this paper, we introduce a diffusion process in $\mathbb{R}^2$ which can easily be simulated and projects onto a bridge process on the torus. More generally, Brownian bridges on manifolds have been studied for example in the context of landmark manifolds [9] and used for approximating the transition density of the Brownian motion. The present paper uses bridges on the flat torus to exemplify how some of the challenges of bridge simulation on Riemannian manifolds can be addressed, here in particular non-trivial topology of the manifold.

We begin in Section 2 with a short introduction to Brownian bridge processes in the standard Eucliden case and how it relates to the definition of a Brownian bridge process on the flat torus. At the end we introduce the stochastic differential equation (SDE) which will be used for simulating the bridge process. In Section 3 we argue that a strong solution of our proposed SDE exist. We show results about convergence and absolute continuity in Section 4. Numerical examples are presented in Section 5.

2 Theoretical Setup

This section will briefly review some Brownian bridges theory and discuss the torus case. A more general theory of diffusion bridges can be found in [1], constituting the main reference for this work. At the end, we introduce our proposal process.

Consider a Brownian motion $W = (W_t)_{t \geq 0}$ in $\mathbb{R}^n$. By conditioning, it can be shown that W will end up at a given point at a given time. For example, the process given by $B_t = W_t - \frac{t}{T} W_T$ defines a Brownian bridge conditioned to return to 0 at time T. It can be shown that the diffusion process given by

$$dX_t = \frac{b - X_t}{T - t} dt + dW_t; \qquad 0 \leq t < T \text{and} \quad X_0 = a, \tag{2.1}$$

for given $a, b \in \mathbb{R}^d$ and W a d-dimensional standard Brownian motion, is a d-dimensional Brownian bridge from a to b on $[0, T]$ (see e.g. [6, sec. 5.6]). More generally, diffusion bridges can be defined through Doob's h-transform, that is, the distribution of a diffusion

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \qquad X_0 = a,$$

conditioned on $X_T = b$ is the same as that of

$$dY_t = \tilde{b}(t, Y_t)dt + \sigma(t, Y_t)dW_t,$$
$$\tilde{b}(t, x) = b(t, x) + \sigma(t, x)\sigma^T(t, x)\nabla_x \log(p(t, x; T, b)),$$

where $p(t, x; T, b)$ denotes the transition density of the process X. In the usual setting where p is the transition density of a Brownian motion it has the form

$$p(s, x; t, y) = \frac{1}{\sqrt{2\pi(t - s)}} \exp\left(-\frac{||x - y||^2}{2(t - s)}\right), \qquad s < t,$$

which yields (2.1).

We propose a method similar to the Euclidean scheme [1] for simulating Brownian bridges on the flat torus, which is of the form

$$dX_t = b(t, X_t)dt + \sigma dW_t; \quad 0 \le t < T \qquad \text{and} \qquad X_0 = a \quad \text{a.s.,} \qquad (2.2)$$

where $\sigma > 0$, $a \in \mathbb{T}^2$ is given, and W is a two-dimensional standard Brownian motion. The exact form of $b(t, x)$ will become apparent below. It is important here to note that in the particular case of the flat torus the transition density for the Brownian motion is known and therefore it is possible to simulate from the distribution of the true Brownian bridge on $\mathbb{T}^2$, however, it requires the calculation of the distance to infinitely many points which the proposed model does not. In Figure 4 is shown paths of the proposed model and the corresponding paths of the true bridge process.

Let $\pi \colon \mathbb{R}^2 \to \mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ denote the canonical projection onto the torus. The standard two-dimensional Brownian motion $W = (W^1, W^2)$, for two independent one-dimensional Brownian motions W^1 and W^2, is mapped to a Brownian motion $B = (B_t)_{t \ge 0}$ on the flat torus $\mathbb{T}^2$ by the projection map π. Indeed, we can identify the torus $\mathbb{T}^2$ with the unit cube $Q = \{x \in \mathbb{R}^2 : -\frac{1}{2} \le x_k < \frac{1}{2}, k = 1, 2\}$. Then for $g \in C^\infty(\mathbb{T}^2)$ the Laplace-Beltrami operator, $\Delta_{\mathbb{T}^2}$, on $\mathbb{T}^2$ corresponds to the restriction to Q of the usual Euclidean Laplacian, $\Delta_{\mathbb{R}^2}\tilde{g}$, where $\tilde{g}$ denotes the periodic extension of g, i.e. $\tilde{g} = g \circ \pi$ (see [8, Sec. 3.5]). Since W is a Brownian motion in $\mathbb{R}^2$ if and only if it satisfies the diffusion equation

$$h(W_t) \overset{m}{=} h(W_0) - \frac{1}{2}\int_0^t \Delta_{\mathbb{R}^2} h(W_s)ds,$$

for all smooth functions h, where $X \overset{m}{=} Y$ means that the difference $X - Y$ is a local martingale (see e.g. [2, Sec. 1.5]), it follows that, for $h = \tilde{g}$,

$$\tilde{g}(W_t) \overset{m}{=} \tilde{g}(W_0) - \frac{1}{2}\int_0^t \Delta_{\mathbb{R}^2}\tilde{g}(W_s)ds = g(B_0) - \frac{1}{2}\int_0^t \Delta_{\mathbb{T}^2} g(B_s)ds \overset{m}{=} g(B_t).$$

3

As this holds for all smooth functions g on $\mathbb{T}^2$, we get that B is a Brownian motion on $\mathbb{T}^2$ in agreement with the definition of a manifold-valued Brownian motion given in [5, Sec. 3.2].

By conditioning B on $\mathbb{T}^2$ to hit a given point $a \in \mathbb{T}^2$, at some fixed time $0 \leq T < \infty$, it is seen that

$$\{\omega \in \Omega : B_T(\omega) = a\} = \{\omega \in \Omega : W_T(\omega) \in \pi^{-1}(a)\},$$

and so simulating a Brownian bridge on the flat torus $\mathbb{T}^2$ is equivalent to simulating a two-dimensional standard Brownian motion conditioned to end up in the set $\pi^{-1}(a)$ at time T. The diffusion given by (2.1) will not suffice as it is constructed to hit exactly one point. It will, however, provide one subset of sample paths of the Brownian bridge on $\mathbb{T}^2$, corresponding to subset of paths that will "unwrap" the same number of times that it "wraps" around the cut locus. This is illustrated in Figure 1. To give a precise meaning to this statement we consider the h-transform

$$h(t, z) = \sum_{y \in \pi^{-1}(a)} \frac{p(t, z; T, y)}{p(0, z_0; T, y)},$$

with p denoting the transition density of the two-dimensional Brownian motion, which by Doob's h-transform implies that the distribution of W conditioned on $W_T \in \pi^{-1}(a)$ is the same as the distribution of the diffusion

$$\begin{aligned}
dZ_t &= \sigma^2 \nabla_z \log \Big(\sum_{y \in \pi^{-1}(a)} p(t, z; T, y) \Big) \Big|_{x=Z_t} dt + \sigma dW_t \\
&= \sum_{y \in \pi^{-1}(a)} g_y(t, Z_t) \frac{y - Z_t}{T - t} dt + \sigma dW_t, \qquad Z_0 = z_0,
\end{aligned} \tag{2.3}$$

where

$$g_y(t, x) = \frac{\exp \big(- \frac{\|y-z\|^2}{2\sigma^2(T-t)} \big)}{\sum_{y \in \pi^{-1}(a)} \exp \big(- \frac{\|y-z\|^2}{2\sigma^2(T-t)} \big)}.$$

Instead, we propose to consider the diffusion process on $[0, T)$, for some fixed positive T, defined by

$$dX_t = 1_{G^c}(X_t) \frac{\alpha(X_t) - X_t}{T - t} dt + \sigma dW_t, \qquad X_0 = x_0 \tag{2.4}$$

where $\sigma > 0$ and α is defined by

$$\alpha(X_t) = \arg\min_{y \in \pi^{-1}(a)} \|y - X_t\|,$$

with $a \in \mathbb{T}^2$, and where G is the set of "straigt lines" of the form $\mathbb{R} \times \{x\}$ (resp. $\{x\} \times \mathbb{R}$) in $\mathbb{R}^2$ where $\alpha(X_t)$ is not unique (see Figure 1). The indicator function removes the drift when the process does not have a natural attraction point.
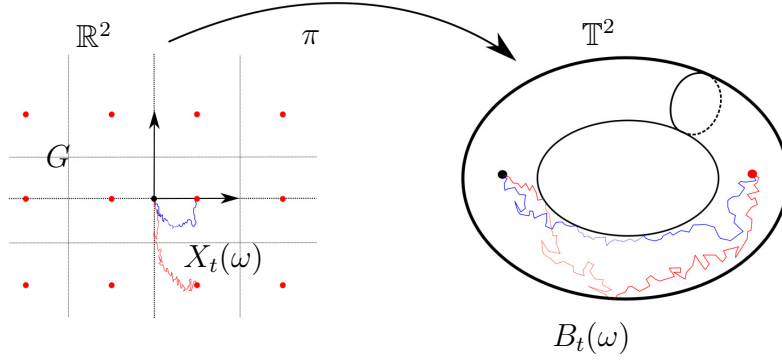
4

Figure 1: The figure illustrates the possibility of the diffusion path going an arbitrary number of times around the torus, starting at the black dot and ending in the red. This is illustrated by the red path. The conditioning on single point in $\mathbb{T}^2$ therefore leads to conditioning on multiple points in $\mathbb{R}^2$. Left: Two paths from the same two-dimensional process with multiple endpoints. Right: The projection of the two paths onto the torus.

3   Existence of Strong Solution

The drift term in equation (2.4) is discontinuous. However, we below show that it posses certain regularity conditions and use this to show that a strong solution to the SDE exist.

In order to ensure the existence of a solution to the diffusion in (2.4), we need some regularity of the drift term. The drift coefficient is given by

$$1_{G^c}(X_t)\frac{\alpha(X_t) - X_t}{T-t} = \begin{cases} \frac{\alpha(X_t) - X_t}{T-t}, & \text{if } X_t \in G^c \\ 0, & \text{otherwise,} \end{cases} \tag{3.1}$$

for every $0 \leq t < T$, where the superscript c denotes the complement. It is a discontinuous process with the set of discontinuities being the set G consisting of the set of straight lines in $\mathbb{R}^2$ where the argmin process is non-unique. It is not even clear that the drift term is suitably measurable as the argmin map in general is not.

Lemma 3.1. *Let $b\colon [0,T)\times\mathbb{R}^2 \to \mathbb{R}^2$ be the map given by (3.1). Then b is $\mathcal{B}([0,T))\otimes \mathcal{B}(\mathbb{R}^2) - \mathcal{B}(\mathbb{R}^2)$ measurable. Furthermore, the map $(s,\omega) \mapsto b(s, X_s(\omega))$ is $\mathcal{B}([0,t])\otimes \mathcal{F}_t^0$ measurable, for every $0 \leq t < T$, where $(\mathcal{F}_t^0)$ denotes the natural filtration generated by X. This is called progressive measurability.*

Proof. First note that G^c is a Borel measurable set as we can write write it as a countable union of open sets, i.e., for $y = (y_1, y_2)$ we have

$$G^c = \bigcup_{y\in\pi^{-1}(a)} \left(y_1 - \tfrac{1}{2}, y_1 + \tfrac{1}{2}\right) \times \left(y_2 - \tfrac{1}{2}, y_2 + \tfrac{1}{2}\right) =: \bigcup_{y\in\pi^{-1}(a)} V_y.$$

Now, we need to show that for all $A \in \mathcal{B}(\mathbb{R}^2)$, the set $b^{-1}(A)$ is an element of $\mathcal{B}([0,T)) \otimes \mathcal{B}(\mathbb{R}^2)$. It is enough to consider all open subsets $U \subseteq \mathbb{R}^2$ as these sets generate the Borel algebra on $\mathbb{R}^2$. So let U be an arbitrary open subset, then we have that

$$b^{-1}(U) = b^{-1}(U) \cap \left([0,T) \times G^c\right) \cup b^{-1}(U) \cap \left([0,T) \times G\right).$$

5

As b is continuous on each of the sets $[0, T) \times V_y$ we have that $b^{-1}(U) \cap \left([0, T) \times G^c\right)$ is a countable union of open sets and therefore an element of $\mathcal{B}([0, T)) \otimes \mathcal{B}(\mathbb{R}^2)$. For the second part we see that

$$b^{-1}(U) \cap ([0, T) \times G) = \begin{cases} [0, T) \times G, & \text{if } (0, 0) \in U \\ \emptyset, & \text{otherwise,} \end{cases}$$

where both are elements of $\mathcal{B}([0, T)) \otimes \mathcal{B}(\mathbb{R}^2)$. This shows that b is Borel measurable.

Progressive measurability follows by a very similar argument. $\qquad\square$

Usually, global or local Lipschitz conditions are imposed on the drift and diffusion coefficients in order to secure global (resp. local) strong solutions to an SDE. This is a too strong condition for the drift term in this case, however, it is bounded in the following sense.

Lemma 3.2. *The drift coefficient in* (3.1) *is uniformly bounded in x and in t on* $[0, S]$, *for any* $0 \le S < T$.

Proof. The first assertion is clear. Let $S \in [0, T)$ be arbitrary and $0 \le t \le S$. For every $x \in G^c$ there exist a $y \in \pi^{-1}(a)$ such that we have

$$\left\| 1_{G^c}(x) \frac{\alpha(x) - x}{T - t} \right\|^2 = \left\| \frac{y - x}{T - t} \right\|^2 \le \frac{C}{(T - S)^2} = C_S,$$

for some positive constants $C > 0$. $\qquad\square$

We now come to the main result of this section.

Proposition 3.3. *There exist a strong solution of* (2.4) *on* $[0, T)$, *which is strongly unique.*

Proof. The drift term is Borel measurable and bounded on $[0, S]$ by Lemma 3.1 and Lemma 3.2. As indicated in [10, Thm. 2] and [11, Thm.1] (2.4) has a strong solution which is strongly unique. $\qquad\square$

Remark 3.4. The assumption in [10, Thm. 2] can be verified by using smooth bump functions.

4 Convergence and Absolute Continuity

The considerations above make the solution of (2.4) into a continuous semimartingale. If a semimartingale X takes its values in an open set U of $\mathbb{R}^2$ then Itô's formula holds true for any $C^{1,2}([0, T) \times U)$ functions as well.

Proposition 4.1. *Let X be a solution to* (2.4) *on the filtered probability space* $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$. *For every $\omega \in \Omega$ for which there exist an $S < T$ such that $X_t(\omega)$ stays in G^c on $[S, T)$, then X converges pointwise almost surely to $\pi^{-1}(a)$.*

Proof. Assume that for some $\omega \in \Omega$ there exist some $S < T$ such that on $[S, T]$ the process $X_t(\omega)$ takes its values in G^c. By continuity of the process it will take it its values in some open neighborhood V_y of the point $y \in \pi^{-1}$. The proof is then identical to the proof in [1, Lemma 4]. $\qquad\square$

Remark 4.2. It is of course of interest to show that for almost every path the process will converge. This can be obtained by showing that the process will not intersect G infinitely many times close to T.

Consider the stochastic process $\mathcal{E}$ on $0 \le t \le S$ defined by

$$\mathcal{E}(L)_t = \exp\Big(-\int_0^t b(s, X_s) dW_s - \frac{1}{2}\int_0^t \|b(s, X_s)\|^2 ds\Big), \qquad (4.1)$$

where L is the local martingale in the exponential. This is known as the Doléans-Dade exponential. From Lemma 3.2 it follows that, for all $t \le S$,

$$\mathbb{E}\Big[\exp\Big(\int_0^t \|b(s, X_s)\|^2 ds\Big)\Big] \le \exp(t C_S) < \infty.$$

The above is known as the Novikov condition (cf. [7]) which ensures that (4.1) is a martingale on $[0, T)$. Girsanov's theorem ([6, Thm. 5.1 Chap. 3]) then provides that the process defined by

$$\widetilde{W}_t = W_t + \int_0^t b(s, X_s) ds$$

is a Brownian motion under the new measure Q introduced below.

Theorem 4.3. *Let X defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a solution of (2.4) on $[0, S]$ for $S < T$. The process in (4.1) defined on $0 \le t \le S$ $(S < T)$ is a true martingale and so there exists a measure Q which is absolutely continuous wrt. P such that X is Q-Brownian motion.*

Proof. The martingale property of (4.1) on $[0, S]$ is a consequence of the Novikov condition. Then Girsanov's theorem gives us that X is a Q-Brownian motion on $[0, S]$. $\qquad\square$

From the (perhaps obvious) fact that the distribution of the true Brownian bridge is locally equivalent to the distribution of the Brownian motion up to time $t < T$, it follows that the distribution of the Brownian bridge is absolutely continuous wrt. the proposed process up to time $t < T$.

Remark 4.4. A bit of extra work is needed to obtain the correction term as in [1]. There are indications that it is possible to simulate from the true distribution of the Brownian bridge on the torus, however, Theorem 1 shows that (2.4) can approximate it.
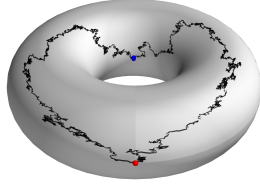
(a) Paths visualized on an embedded torus.
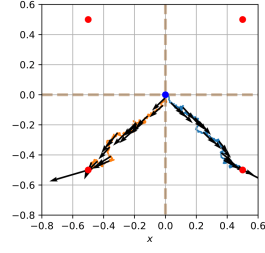
(b) The two Euclidean paths that are mapped onto the torus.

Figure 2: Two different paths visualized both on the torus and in Eucliden space. The blue dot represents the starting point and the red represents the end point.

5   Numerical Experiments.

For the numerical implementation of the proposed SDE in equation (2.4) we implemented the Euler-Maruyama scheme, i.e. taking n equidistant discretization points of the time interval $t_1, \ldots, t_n$, with $t_{i+1} - t_i = \Delta t$, the numerical equation becomes

$$x_{t_{i+1}} = x_{t_i} + \frac{\arg\min_{y \in \pi^{-1}(a)}(\|y - x_{t_i}\|) - x_{t_i}}{T - t_i} \Delta t + \sigma \Delta W_{t_i},$$

where $\Delta W_{t_{i+1}} = W_{t_{i+1}} - W_{t_i}$ is equal in distribution to a normal random variable with mean zero and variance Δt.

Figure 2a shows the implementation of the numerical scheme on an embedded torus and Figure 2b its Euclidean counterpart. Figure 3a shows the behaviour of the drift term along a given path, illustrating that the attraction becomes stronger as time approaches the terminal time. The vector fields in Figure 3b shows the constant attraction to the center of the open subsets.
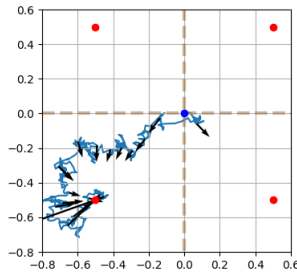
(a) Drift term
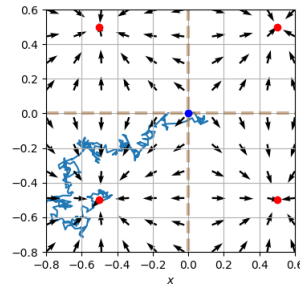
(b) Vector field

Figure 3: Figure 3a depicts the evolution of the drift term. It shows how the pull from the drift becomes stronger near the end. Figure 3b shows the underlying vector field.
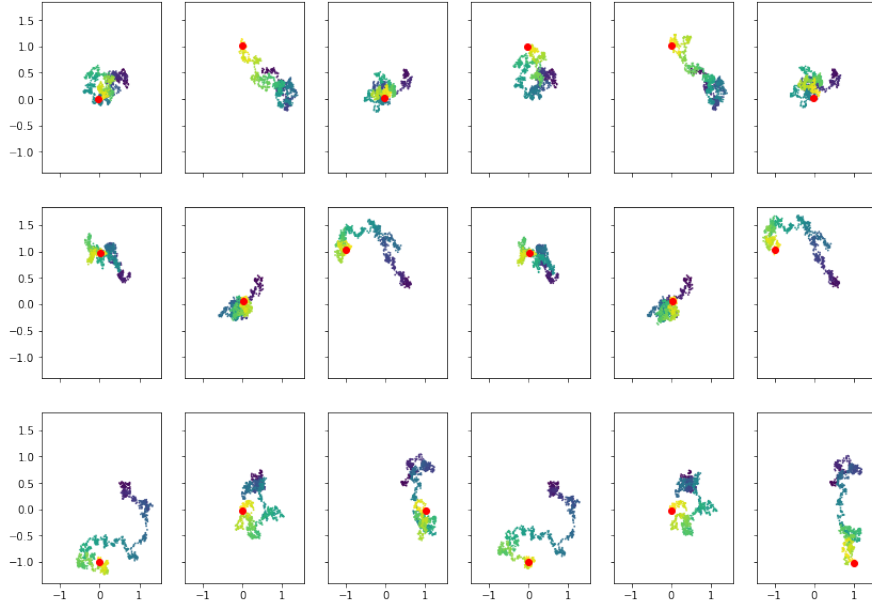
8

Figure 4: Figure 4 shows 9 paths from the proposed model (2.4) on the left and the corresponding paths from the true bridge (2.3) on the right. It is seen that the first and the last paths disagree on the limiting point, whereas the rest looks fairly similar. The picture agree with the fact that roughly four in five have the same limiting point. Here $\sigma = 0.8$ and the conditioning points being the integers in $[-2, 2] \times [-2, 2]$.

Acknowledgements

References

[1] Delyon, B., Hu, Y.: Simulation of conditioned diffusion and application to parameter estimation. Stochastic Processes and their Applications **116**(11), 1660–1675 (2006)

[2] Emery, M.: Stochastic calculus in manifolds. Springer (1989)

[3] García-Portugués, E., Sørensen, M., Mardia, K.V., Hamelryck, T.: Langevin diffusions on the torus: estimation and applications. Statistics and Computing pp. 1–22 (2017)

[4] Golden, M., García-Portugués, E., Sørensen, M., Mardia, K.V., Hamelryck, T., Hein, J.: A Generative Angular Model of Protein Structure Evolution. Molecular Biology and Evolution **34**(8), 2085–2100 (2017)

[5] Hsu, E.P.: Stochastic analysis on manifolds, vol. 38. American Mathematical Soc. (2002)

[6] Karatzas, I., Shreve, S.: Brownian Motion and Stochastic Calculus. Springer, New York, 2nd edition edn. (Aug 1991)

[7] Novikov, A.A.: On an Identity for Stochastic Integrals. Theory of Probability & Its Applications **17**(4), 717–720 (Sep 1973)

[8] Sogge, C.D.: Hangzhou Lectures on Eigenfunctions of the Laplacian (AM-188). Princeton University Press (Mar 2014)

[9] Sommer, S., Arnaudon, A., Kuhnel, L., Joshi, S.: Bridge simulation and metric estimation on landmark manifolds. In: Graphs in Biomedical Image Analysis, Computational Anatomy and Imaging Genetics, pp. 79–91. Springer (2017)

[10] Veretennikov, A.Y.: On the strong solutions of stochastic differential equations. Theory of Probability & Its Applications **24**(2), 354–366 (1980)

[11] Veretennikov, A.J.: On strong solutions and explicit formulas for solutions of stochastic integral equations. Sbornik: Mathematics **39**, 387–403 (1981)