Level sets and drift estimation for reflected Brownian motion with drift

Alejandro Cholaquidis*, Ricardo Fraiman*, Ernesto Mordecki*, Cecilia Papalardo*
* Universidad de la República

Abstract

We consider the estimation of the drift and the level sets of the stationary distribution of a Brownian motion with drift, reflected in the boundary of a compact set $S \subset \mathbb{R}^d$, departing from the observation of a trajectory of this process. We obtain the uniform consistency and rates of convergence for the proposed kernel based estimators. This problem has relevant applications in ecology, in estimating the home-range and the core-area of an animal based on tracking data. Recently, the problem of estimating the domain of a reflected Brownian motion was considered in [12], when the stationary distribution is uniform and the estimation of the core-area, defined as a level set of the stationary distribution, is meaningless. As a by-product of our results, we obtain an estimation of the drift function. In order to prove our results, some new theoretical properties of the reflected Brownian motion with drift are obtained, under fairly general assumptions. These properties allow us to perform the estimation for flexible regions close to reality. The theoretical findings are illustrated in simulated and real data examples.

1 Introduction

Level set estimation is a well studied problem with many applications in practice. One important field of applications is ecology, where the interest is in the estimation of the home-range and the core-area of the habitat of an animal, among other parameters. For this case the data is the continuous time trajectory of the animal (obtained for instance from a GPS), therefore it is not reasonable to assume independence of the observations. We will show how the results obtained in this manuscript can be applied to that problem through an illustration with real data.

The problem of the estimation of a level set $G = G_f(\lambda) = \{x \in \mathbb{R}^d : f(x) > \lambda\}$ of an unknown density $f$, is usually considered departing from a sample of independent random vectors, identically distributed according to $f$. See for instance [35], [22] and [14], where consistency is proved with respect to the Hausdorff distance and the distance in measure, assuming certain smoothness properties of the set $G$.

In the present paper we are interested in the problem of estimating the unknown level sets $G_f(\lambda)$ of the density $f$ of the stationary distribution and the drift of a reflected Brownian motion with drift (RBMD) when observing a trajectory of this process. Roughly speaking, the process under observation behaves in the interior of a compact set $S \subset \mathbb{R}^d$ like an ordinary Brownian motion with drift, and reflects (normally) at the boundary $\partial S$. More precisely, the RBMD is defined as the solution of the Skorokhod stochastic differential equation (1).
Besides the classical estimator $\hat{G} = G_f(\lambda) = \{ x \in \mathbb{R}^d : \hat{f}(x) > \lambda \}$, we also consider an estimator with a geometric shape restriction. More precisely, we consider an estimator within the class of $r$-convex sets. An $r$-convex set is characterized as the intersection of complements of open balls of radius $r$ that do not intersect the set. The family of $r$-convex sets provides a quite general and flexible class of sets allowing holes and smooth inlets in the set.

First we need some theoretical results related to the RBMD, such as the existence and uniqueness of the solution of a stochastic differential equation with normal reflection, with the corresponding conditions on the boundary $\partial S$, and verify the strong Markov property for this solution. Then we extend the non-trap property for the reflected Brownian motion introduced in [6], and its equivalent forms, to our RBMD case. Although the extension of these results from the classical reflected Brownian motion to reflected diffusions corresponding to uniformly elliptic operators in divergence form in domains with smooth boundary are expected, there is no formal proof of these facts. Moreover, the case of a convex domain has been recently addressed in [5]. For these reasons we include the formal proofs. Based on these results we derive the uniform convergence of kernel based estimators of the stationary distribution, and the convergence of the level sets with respect to the Hausdorff distance. We also derive from the stationary distribution estimator a simple estimator of the drift function. Recently, it was proved in [5] that if the set $S$ is convex and the drift is a convex function, the distance in total variation between the distribution of the process at time $t$ and the stationary distribution converges to zero at an exponential rate as $t$ goes to infinity. With a different approach, an exponential rate has been also obtained for ergodic diffusions in unbounded domains (see [16]). See also [11], where a similar problem is considered. The estimation of the stationary distribution of a stochastic differential equation with drift, without reflection, has been studied by several authors, some earlier results are given in [42] who adapt the results of [9] to this setting. More recently, [17] estimated the drift and the stationary distribution for the same model but without reflection, whereas [20] considered estimation problems for one dimensional diffusions with and without reflection.

Before introducing the formal framework, we discuss briefly the application of the proposed method to the estimation from animal tracking data of the core-area and drift. For a description of home-range estimation see for instance [12] and the references therein.

### 1.1 Estimation of home-range, core-area, and dynamics

The estimation of the home-range is an important problem in animal ecology. It was first defined by [7] as ‘the area traversed by the individual in its normal activities of food gathering, mating, and caring for young’. Besides the size and shape of the home-range, it is also important to have information about its structure and dynamics. It is well known that animals occupy certain regions with more frequency than others. Researchers have been interested in the areas where animals spend most of their time (see for instance [23]). These areas are called core-areas. In our setup, core-areas can be modelled by the level sets of the stationary distribution, while the drift function pro-
vides information about the dynamics of the movement of the animal. As mentioned in [12], since these first definitions, the concept of home-range and core-area have evolved, giving rise to a considerable amount of literature on the subject (see for instance the reviews in [38] and [29]). Core-areas have been defined through the utilization distribution, defined as the density function that describes the probability of finding the animal at a particular location (see [39]), and kernel methods have been used for estimating this utility distribution. These methods generally treat the recorded locations as independent observations. However, recent advances in animal tracking technology (VHF radio transmission, Argos system, and especially GPS) allow almost continuously recording the movements of animals. In this context, the independence of the observations cannot be assumed and new mathematical models are needed. Modeling the movement of an animal in its home-range as a continuous stochastic process provides a theoretical framework in which tracking data can be analysed.

1.2 Roadmap

This paper is organized as follows. In Section 2, we discuss conditions for the existence and uniqueness of reflected Brownian motion with drift and its stationary distribution. The main results in this section are given in Propositions 2, 3 and 4. Proposition 2 gives conditions for the domain to be non-trap for the RBMD process \( \{X_t\} \) (a condition introduced in [6], which we describe in Section 2). Proposition 3 provides conditions for the existence of a unique stationary measure for the RBMD. In Proposition 4 we show that if the domain is non-trap, we have an exponential rate of convergence to the stationary distribution for the total variation norm. All the proofs for this section are given in the Appendix. In Section 3 we obtain in Theorem 1 strong uniform convergence rates for kernel estimators of the stationary distribution based on a trajectory of the RBMD. In Corollary 1 and Theorem 3 we prove the strong consistency of two different families of level sets estimators with respect to the Hausdorff distance. The case of the estimation of level sets with a given content is considered in Theorem 4. We also derive consistent estimates of the drift function. Lastly, in Section 4 we consider some simulated and real data examples to illustrate the behaviour of the estimation methods described in the paper.

2 Reflected Brownian motion with drift

In this section we establish conditions for the existence of a reflected Brownian motion with drift and its stationary distribution, and study the connections between these conditions and some geometric constraints on its support.

2.1 Topological notation

Given a set \( S \subset \mathbb{R}^d \), we denote by \( \partial S \), \( \text{int}(S) \), and \( \overline{S} \) the boundary, interior, and closure of \( S \), respectively. The Borel sigma algebra in \( S \) will be denoted by \( \mathcal{B}(S) \). We denote by \( \langle \cdot, \cdot \rangle \) the usual inner product in \( \mathbb{R}^d \) and by \( \| \cdot \| \) the Euclidean norm. The
closed ball of radius \( \varepsilon \) centred at \( x \) is denoted by \( B(x, \varepsilon) \), while the open ball is denoted by \( B_\varepsilon(x) \). Given \( \varepsilon > 0 \) and a bounded set \( A \subset \mathbb{R}^d \), \( B(A, \varepsilon) \) denotes the parallel set \( \{ x \in \mathbb{R}^d : d(x, A) \leq \varepsilon \} \) where \( d(x, A) = \inf \{ \| x - a \| : a \in A \} \). The \( d \)-dimensional Lebesgue measure on \( \mathbb{R}^d \) will be denoted by \( \mu_L \). We will denote by \( C^2_c(D) \) the set of twice continuously differentiable functions with compact support in some domain containing \( D \).

2.2 Skorokhod SDE

Let \( D \) be a bounded domain in \( \mathbb{R}^d \) (that is, a bounded connected open set) such that \( \partial D \) is \( C^2 \). Given a \( d \)-dimensional Brownian motion \( \{ B_t \}_{t \geq 0} \) departing from \( B_0 = 0 \) defined on a filtered probability space \( (\Omega, \mathcal{F}, \{ \mathcal{F}_t \}_{t \geq 0}, \mathbb{P}_x) \), and a function \( f : D \to \mathbb{R} \), we are concerned with the problem of the existence and uniqueness of the solution of a reflected stochastic differential equation on \( D \) whose drift is given by the gradient of a function \( f : \mathbb{R}^d \to \mathbb{R} \):

\[
X_t = X_0 + B_t - \frac{1}{2} \int_0^t \nabla f(X_s) ds + \int_0^t n(X_s) \xi(ds), \quad \text{where } X_t \in \overline{D}, \forall t \geq 0. \tag{1}
\]

Here we assume that \( \nabla f \) is Lipschitz, while \( n(x) \) denotes the inner unit vector at the boundary point \( x \in \partial D \); this boundary satisfies some regularity conditions (to be specified later). This equation is called a Skorokhod stochastic differential equation. Its solution is a pair of stochastic processes \( \{ X_t, \xi_t \}_{t \geq 0} \), the first coordinate \( \{ X_t \}_{t \geq 0} \) is a reflected diffusion, which we call a reflected Brownian motion with drift (RBMD), and \( \{ \xi_t \}_{t \geq 0} \) is the corresponding local time, that is, a one-dimensional continuous non-decreasing process with \( \xi_0 = 0 \) that satisfies

\[
\xi_t = \int_0^t 1_{\{ X_s \in \partial D \}} d\xi_s.
\]

Since we have assumed that \( \partial D \) is \( C^2 \), we know that a ball of positive radius rolls freely inside and outside \( \overline{D} \) (see [37]). Then, by using the same arguments used to prove proposition 3 in [12], we can ensure that the geometric conditions for the existence of a solution of Equation (1), as required in [34], are satisfied. We then get by theorem 5.1 in [34] that there exists a unique strong solution of the Skorokhod stochastic differential equation (1). The solution is a strong solution in the sense of definition 1.6 in [24]. The following proposition provides a constructive proof of the uniqueness and strong Markov property, which is based on some results in [32]. It also gives that, for all \( h \in C^2_c(D) \), the process

\[
h(X_t) - \int_0^t \mathcal{L}(h(X_u)) du,
\]

is a submartingale, where \( \{ X_t \}_{t \geq 0} \) denotes the solution of (1) with \( X_0 = x \) a.s., and we write \( \mathcal{L} \) for the infinitesimal generator of the process, i.e.

\[
\mathcal{L}(h)(x) = \lim_{t \downarrow 0} \frac{\mathbb{E}_{x}(h(X_t)) - h(x)}{t}.
\]
Proposition 1. Let $D$ be a bounded domain such that $\partial D$ is $C^2$, and assume that $\nabla f$ is Lipschitz on $\overline{D}$. Then there exists a unique strong solution $\{X_t\}_{t \geq 0}$ of (1) which is a strong Markov process, whose generator is

$$\mathcal{L} h = \frac{1}{2} \Delta h - \frac{1}{2} \langle \nabla f, \nabla h \rangle,$$

where the test function $h \in C^2_c(D)$, with Neumann boundary condition $\langle n(x), \nabla h(x) \rangle = 0$.

Remark 1. There exists a unique positive function $p(s, x, t, y)$ satisfying

$$P(X_t \in \Gamma | X_s = x) := P(s, x, t, \Gamma) = \int_{\Gamma} p(s, x, t, y) dy$$

and, by theorem 3.2.1 of [33], the function $p$ satisfies the forward equation $\partial_s p + \mathcal{L}^* p = 0$ and $\lim_{s \to t^-} p(s, \cdot, t, y) = \delta_y$, where $\delta_y$ is the point-mass at $y$ and $\mathcal{L}^*$ is the dual of $\mathcal{L}$, that is,

$$\mathcal{L}^* h = \frac{1}{2} \Delta h + \frac{1}{2} \langle \nabla f, \nabla h \rangle.$$

2.3 Ergodic properties

In this subsection we prove that there is a unique stationary distribution given by

$$\pi(dx) = ce^{-f(x)} I_D dx,$$

where $c$ is the normalization constant. For this purpose, we give conditions for the domain to be non-trap (see Definition (5) below), which implies Harris recurrence. Then the uniqueness follows, see for instance [2].

We now define the notions of invariant measure and ergodic process, following [41].

Definition 1. A probability measure $\pi$ on $S$ is said to be an invariant measure for a time-homogeneous Markov process $\{Z_t\}_{t \geq 0}$ if

$$\int_S P_x(Z_t \in A) \pi(dx) = \pi(A), \quad \text{for all } t > 0 \text{ and all } A \in \mathcal{B}(S).$$

Remark 2. It is well known (see for instance [41]) that if the process $\{Z_t\}_{t \geq 0}$ is weakly Feller (i.e. $\forall t \geq 0$, $E_x(h(Z_t))$ is a continuous function of $x$ for all $h \in C^2_c(\overline{D})$) then there exists an invariant measure.

Definition 2. A Markov process $\{Z_t\}_{t \geq 0}$ with state space $S$ is ergodic if there exists an invariant probability measure $\pi$ such that

$$\lim_{t \to +\infty} \|P_x(Z_t \in \cdot) - \pi(\cdot)\|_{TV} = 0 \quad \forall x \in S.$$

Here $\|\mu\|_{TV}$ stands for the total variation norm of the measure $\mu$. In this case $\pi$ is called a stationary distribution.

Definition 3. A Markov process $\{Z_n\}_{n \in \mathbb{N}}$ with state space $S$ is called geometrically ergodic if there exists an invariant probability $\pi$ and real numbers $0 < \rho < 1$ and $\gamma > 0$ such that

$$P_x(Z_n \in B) - \pi(B) \leq \gamma \rho^n \quad \text{for all } x \in S \text{ and all } B \in \mathcal{B}(S).$$
2.4 Harris recurrence and the trap condition.

Let $D \subset \mathbb{R}^d$ be an open bounded set and $\mathcal{B} \subset D$. Consider the first hitting time of $\mathcal{B}$ by a stochastic process $\{Z_t\}_{t \geq 0}$ defined by $T_\mathcal{B} = \inf\{t > 0 : Z_t \in \mathcal{B}\}$.

**Definition 4.** A Markov process $\{Z_t\}_{t \geq 0}$ is called Harris recurrent if for some $\sigma$-finite measure $\mu$, we have $\mathbb{P}_x(T_A < \infty) = 1$ whenever $\mu(A) > 0$, $A \in \mathcal{B}(\overline{D})$.

Under Harris recurrence there exists a unique (up to a multiplicative constant) invariant measure (see [2]). For the RBMD we prove a stronger condition called non-trap (see [6]).

**Definition 5.** We say that $D$ is a trap domain for the stochastic process $\{Z_t\}_{t \geq 0}$ if there exists a closed ball $\mathcal{B} \subset D$ with positive radius such that
\[
\sup_{x \in D} \mathbb{E}_x T_\mathcal{B} = \infty, \tag{7}
\]
where $\mathbb{E}_x$ denotes the expectation w.r.t. $\mathbb{P}_x$. Otherwise $D$ is called a non-trap domain.

It is proved in lemma 3.2 in [6] that if $\{X_t\}_{t \geq 0}$ is a reflected Brownian motion (without drift) in a connected open set $D$ with finite volume and $\mathcal{B}_1, \mathcal{B}_2$ are closed non-degenerate balls in $D$, then $\sup_{x \in D} \mathbb{E}_x T_{\mathcal{B}_1} < \infty$ if and only if $\sup_{x \in D} \mathbb{E}_x T_{\mathcal{B}_2} < \infty$. In the Appendix we will prove the proof that the RBMD is also non-trap.

**Proposition 2.** Let $D \subset \mathbb{R}^d$ be a bounded domain such that $\partial D$ is $C^2$. Let $\{X_t\}_{t \geq 0}$ be the solution of (1), with $f$ a $C^2$ function on $\overline{D}$. Then $D$ is a non-trap domain for $\{X_t\}_{t \geq 0}$.

Observe that since we can prove that the RBMD is non-trap, it is Harris recurrent and so there exists a unique invariant measure. In order to obtain the stationary distribution, we will use the following lemma, whose proof is accomplished using Proposition 1 and reasoning as in the proof of lemma 2.1 (i) in [21].

**Lemma 1.** Let $D$ be a bounded domain such that $\partial D$ is $C^2$. Suppose that $p : \overline{D} \to \mathbb{R}$ is $C^2$, that $p$ is positive on $D$, and that $\int_D p(x)dx = 1$. Then $p$ is the density of the (unique) invariant distribution for (1) if and only if
\[
\int_D p(x)\mathcal{L}h(x)dx = 0 \quad \text{for all } h \in C_c^2(\overline{D}) \text{ satisfying } \langle \nabla h(x), \mathbf{n}(x) \rangle = 0 \text{ on } \partial D, \tag{8}
\]
where $\mathbf{n}(x)$ denotes the inner normal vector at $x \in \partial D$.

**Proposition 3.** Let $D$ be a bounded domain such that $\partial D$ is $C^2$. Assume that $\nabla f$ is Lipschitz on $\overline{D}$. Then the measure
\[
\pi(dx) = ce^{-f(x)}\|D\|dx, \tag{9}
\]
(where $c$ is a normalization constant) is the unique stationary measure of $\{X_t\}_{t \geq 0}$.

**Proposition 4.** Let $D \subset \mathbb{R}^d$ be a bounded domain such that $\partial D$ is $C^2$. Denote by $\pi$ the stationary distribution of $\{X_t\}_{t \geq 0}$ given by (9), and assume that its density is $C^2$. If $D$ is a non-trap domain for $\{X_t\}_{t \geq 0}$, then there exist positive constants $\alpha$ and $\beta$ such that
\[
\sup_{x \in D} \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{TV} \leq \beta e^{-\alpha t}. \tag{10}
\]
3 Estimation of the drift and stationary distribution

In this section we first obtain in Theorem 1 strong uniform convergence rates for the classical kernel density estimator \( \hat{g}_n \) of the density \( g \) of the stationary distribution of a Markov chain. Next, in Corollary 1 and Theorem 3 we show the convergence of two families of estimator of the level sets. We consider the case of the estimation of level sets with a given content in Theorem 4. Lastly, we derive the strong consistency of the drift estimator.

The proof of Theorem 1 is based following some ideas in [8]. The main difference from [8] being that we aim to obtain uniform convergence, to be able to estimate the level sets. In order to do so, we will introduce some notation.

Let \( \{X_n\}_{n \in \mathbb{N}} \) be a Markov process with state space \( S \subset \mathbb{R}^d \) and let \( \mu_0(dy) \) be an arbitrary initial distribution. Let \( \mu_n(dy) \) denote the distribution of \( X_n \), that is,

\[
P_{\mu_0}(X_n \in A) = \int_A \mu_n(dy) \quad \forall A \in \mathcal{B}(S),
\]

where \( P_{\mu_0} \) indicates that the initial distribution is \( \mu_0 \). Similarly, we write \( E_{\mu_0} \).

Let \( K : \mathbb{R}^d \to \mathbb{R} \) be a nonnegative bounded density function. Consider the classical kernel estimator \( \hat{g}_n \) based on \( \{X_1, \ldots, X_n\} \), given by

\[
\hat{g}_n(x) = \frac{1}{nh_n^d} \sum_{i=1}^{n} K \left( \frac{x - X_i}{h_n} \right) = \frac{1}{n} \sum_{i=1}^{n} K_h(x - y),
\]

where \( h_n \to 0 \) and \( K_h(x) = K(x/h)/h^d \).

The following generalization of the Bernstein inequality obtained in [13], will be useful throughout the present discussion. Some sharper bounds were obtained more recently (see for instance [18]). However the same rates of convergence are obtained from Collob's inequality.

Lemma 2. (Bernstein inequality for \( \varphi \)-mixing processes). Let \( Y_i \) be a sequence of \( \varphi \)-mixing random variables such that \( E(Y_i) = 0 \), \( |Y_i| \leq C_1 \), \( E|Y_i| \leq \eta \), and \( E(Y_i^2) \leq D \). Write \( \bar{\varphi}(m) = \varphi(1) + \cdots + \varphi(m) \) for each \( m \in \mathbb{N} \). Then for each \( \varepsilon > 0 \) and \( n \in \mathbb{N} \), we have

\[
P \left( \left| \sum_{k=1}^{n} Y_k \right| < \varepsilon \right) \leq 2 \exp \left( 3e^{1/2} \frac{n \varphi(m)}{m} - \alpha \varepsilon + \alpha^2 n C_2 \right),
\]

where \( C_2 = 6(D + 4\eta C_1 \bar{\varphi}(m)) \) and \( \alpha, m \) are respectively any positive real number and any positive integer less than or equal to \( n \) and satisfying \( \alpha m C_1 \leq 1/4 \). The numbers \( \alpha \) and \( m \) may also depend on \( n \).

Theorem 1. Let \( S \subset \mathbb{R}^d \) be a compact set and \( \{X_n\}_{n \in \mathbb{N}} \) a geometrically ergodic Markov chain with state space \( S \), whose stationary distribution, \( \pi \), has a Lipschitz density \( g \)
w.r.t. to Lebesgue measure. Let $\hat{g}_n(x) = \frac{1}{n} \sum_{i=1}^{n} K_h(x - X_i)$ with $K$ a Lipschitz, bounded, density such that $\int |u|K(u)du < \infty$. Let $h = h_n \to 0$ such that $nh^{d+2}/\log(n) \to 0$. Then

$$\beta_n \sup_{x \in S} |\hat{g}_n(x) - g(x)| \to 0 \quad a.s.,$$

where $\beta_n \to \infty$ fulfils $\frac{nh^d}{\log(n)\beta_n} \to \infty$ and $\log(n)/\beta_n \to 0$.

**Proof.** We will deal separately with each term on the right hand side of the following inequality:

$$\beta_n \sup_{x} |g(x) - \hat{g}_n(x)| \leq \beta_n \sup_{x} |g(x) - \mathbb{E}_x(\hat{g}_n(x))| + \beta_n \sup_{x} |\hat{g}_n(x) - \mathbb{E}_\pi(\hat{g}_n(x))|.$$  \hspace{1cm} (14)

First we bound the bias term. Let $C_g$ such that $|g(x) - g(y)| \leq C_g \|x - y\|$. Then

$$|\mathbb{E}_x(K_h(x - X_k)) - g(x)| \leq \left| \int_S K_h(x - y)g(y)dy - g(x) \right|$$

$$\leq \int_S K_h(x - y)|g(y) - g(x)|dy \leq hC_g \int_{\mathbb{R}^d} \|u\|K(u)du = O(h).$$  \hspace{1cm} (15)

Now observe that $\mathbb{E}_{\mu_0}(K_h(x - X_k)) = \int_S K_h(x - y)\mu_k(dy)$. Put $k_1 = \max_x K(x)$. Now, by (6),

$$\left| \int_S K_h(x - y)\mu_k(dy) - \int_S K_h(x - y)d\pi(y) \right| \leq \|K_h(x - y)\|_\infty \|\mu_k - \pi\|_{TV} \leq \frac{k_1}{h^d} \gamma \rho^k.$$  \hspace{1cm} (16)

Observe that $\mathbb{E}(\hat{g}_n(x)) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}_{\mu_0}(K_h(x - X_k))$, and $\mathbb{E}_\pi(K_h(x - X_k)) = \int_S K_h(x - y)g(y)dy$. Hence (16) implies

$$\beta_n \sup_x \left| \frac{1}{n} \sum_{k=1}^{n} \left[ \mathbb{E}_{\mu_0}(K_h(x - X_k)) - \mathbb{E}_\pi(K_h(x - X_k)) \right] \right| \leq \beta_n \frac{k_1 \gamma}{h^d \log(n)} \sum_{k=1}^{n} \rho^k.$$ \hspace{1cm} (17)

which, together with (15), implies that

$$\beta_n \sup_x \left| \mathbb{E}(\hat{g}_n(x)) - g(x) \right| \leq C \frac{\beta_n}{nh^d} + C' h \beta_n.$$  \hspace{1cm}

where $C$ and $C'$ are positive constants. Since $nh^{d+2}/\log(n) \to 0$, we get,

$$\beta_n \sup_x \left| \mathbb{E}(\hat{g}_n(x)) - g(x) \right| \to 0.$$  \hspace{1cm} (18)

It remains to prove that $\beta_n \sup_x \left| \hat{g}_n(x) - \mathbb{E}(\hat{g}_n(x)) \right| \to 0$. Since $S$ is compact, we can cover $S$ with $\nu \leq \frac{c}{h^{d+2}}$ balls of radius $h^{d+2}$ centred at some fixed points $\{x_1, \ldots, x_\nu\} \subset S$, $c$ being a positive constant depending on $d$ and $\mu_L(S)$. For $i = 1, \ldots, \nu$,

$$\mathbb{P}\left( |\hat{g}_n(x_i) - \mathbb{E}(\hat{g}_n(x_i))| > \varepsilon \right) = \mathbb{P}\left( \sum_{j=1}^{n} \left[ K_h(x_i - X_j) - \mathbb{E}_{\mu_0}(K_h(x_i - X_j)) \right] > n\varepsilon \right) > n\varepsilon.$$ \hspace{1cm} (19)
By proposition 4.1 of [8], the sequence \( \{X_n\}_{n \in \mathbb{N}} \) is \( \varphi \) mixing with \( \varphi(n) = 2 \gamma \rho^n \) (\( \gamma \) as in (6)). Let \( x \in S \) and \( x_i \) be such that \( \|x - x_i\| < h^{d+2} \). Then, since \( K \) is Lipschitz, there exists a constant \( R \) such that
\[
|\hat{g}_n(x) - \hat{g}_n(x_i)| \leq \frac{1}{nh^d} \sum_{j=1}^{n} K \left( \frac{x - X_j}{h} \right) - K \left( \frac{x_i - X_j}{h} \right) \leq \frac{1}{h^{d+1}} R \|x - x_i\| \leq Rh.
\]

Hence, \( |\hat{g}_n(x) - \mathbb{E}(\hat{g}_n(x))| \leq |\hat{g}_n(x_i) - \mathbb{E}(\hat{g}_n(x_i))| + 2Rh \). If we take \( n \) so large that \( 2Rh < \varepsilon/(2\beta_n) \), we get
\[
P \left( \beta_n \sup_{x \in B(x, h^{d+2})} |\hat{g}_n(x) - \mathbb{E}(\hat{g}_n(x))| > \varepsilon \right) \leq P \left( |\hat{g}_n(x_i) - \mathbb{E}(\hat{g}_n(x_i))| > \frac{\varepsilon}{2\beta_n} \right). \tag{20}
\]

Now use the Bernstein inequality (12) with \( Y_j = K((x - X_j)/h) - \mathbb{E}_{\mu_0}(K(x - X_j)/h) \) and \( C_1 = 2k_1 \). Observe that by (16), for \( n \) large enough,
\[
\mathbb{E}_{\mu_0} \left( K \left( \frac{x - X_j}{h} \right) \right) \leq k_1 \gamma \rho^n + \int_{S} K \left( \frac{x - y}{h} \right) g(y)dy \leq k_1 \gamma \rho^n + k_1 h^d g_1 \mu_L(S) \leq C'' h^d,
\]
where \( g_1 = \max_{x \in S} g(x) \), \( C'' \) being a positive constant. Hence \( \eta = 2C'' h^d \), \( D \leq C_1 C'' h^d \), and \( \hat{\varphi}(m) \leq \sum_{i=1}^{\infty} 2^i \gamma \rho^i = 2 \gamma \rho/(1 - \rho) \), so \( C_2 = 0(h^d) \). For \( \alpha = \alpha_n = o(1/\beta_n) \) such that \( \alpha_n h^d / (\beta_n \log(n)) \to \infty \), if \( m = \lfloor \beta_n \rfloor \), then \( \alpha_n C_2 h^{-d} < \varepsilon/(4\beta_n) \) and \( \alpha_n m C_1 < 1/4 \) for \( n \) large enough. On the other hand, since \( \log(n)/\beta_n \to 0 \),
\[
3e^{1/2} \frac{\varphi(m)}{m} \to 0, \text{ as } n \to \infty.
\]

Now the Bernstein inequality implies that
\[
P \left( |\hat{g}_n(x_i) - \mathbb{E}(\hat{g}_n(x_i))| > \frac{\varepsilon}{2\beta_n} \right) \leq 2 \exp \left( 3e^{1/2} \frac{\varphi(m)}{m} \right) \exp \left( -\frac{\alpha_n \varepsilon h^d}{2\beta_n} + \alpha_n^2 C_2 n \right)
\]
\[
= 2 \exp \left( 3e^{1/2} \frac{\varphi(m)}{m} \right) \exp \left( -\frac{\alpha_n h^d}{\beta_n} \left( \frac{\varepsilon}{2} - \beta_n \alpha_n C_2 h^{-d} \right) \right)
\]
\[
\leq C_3 \exp \left( -\frac{C_4 C_2 n h^d}{\beta_n} \right), \tag{21}
\]
for some positive constants \( C_3, C_4 \). Finally,
\[
P \left( \sup_x |\hat{g}_n(x) - \mathbb{E}(\hat{g}_n(x))| > \frac{\varepsilon}{\beta_n} \right) \leq \sum_{i=1}^{\nu} P \left( |\hat{g}_n(x_i) - \mathbb{E}(\hat{g}_n(x_i))| > \frac{\varepsilon}{2\beta_n} \right)
\]
\[
\leq \frac{C_5}{h^{d(d+2)}} \exp \left( -\frac{C_4 C_2 n h^d}{\beta_n} \right), \tag{22}
\]
\[
\leq C_5 \exp \left( -\frac{C_4 C_2 n h^d}{\beta_n} - C_6 \log(h) \right) \tag{23}
\]
with \(C_5\) and \(C_6\) being positive constants. Since
\[
\frac{C_4\alpha_n n h^d}{\beta_n} + C_6 \log(h) \rightarrow \infty.
\]
Now by the Borel–Cantelli Lemma we obtain that
\[
\lim_{n \to \infty} \beta_n \sup_x |\hat{g}_n(x) - \mathbb{E}(\hat{g}_n(x))| = 0 \text{ a.s.},
\]
which, together with (14) and (18), implies (13).

\[\Box\]

**Remark 3.**

i) Taking \(h = n^{-1/\nu}\) and \(\beta_n = n^\gamma\), then the best attainable rate is for
\(\gamma = \frac{1}{2}(1 - d/(d + 2))\), i.e., \(\beta_n = O(n^{1/2(1-d/(d+2))})\).

ii) If we only want uniform convergence, the conditions in \(h_n\) can be relaxed, and
replaced by \(h^d \log(\text{nh}^d) \to 0\) and \(\text{nh}^d / \log(n) \to \infty\).

Now using theorem 1 of [14] we get the following direct corollary, which establishes
the consistency in Hausdorff distance of the boundary of the estimated level sets \(\partial \hat{G}_n(\lambda)\)
(where \(G_g(\beta) = \{ x : g(x) > \beta \}\)). Recall that given two non-empty compact sets \(A, C \subset \mathbb{R}^d\), the Hausdorff distance between \(A\) and \(C\) is defined as
\[
d_H(A, C) = \max \left\{ \max_{a \in A} d(a, C), \max_{c \in C} d(c, A) \right\}, \quad (24)
\]
where \(d(a, C) = \inf_{c \in C} d(a, c)\).

**Corollary 1.** Under the the hypotheses of Theorem 1, suppose in addition that there
exists \(\lambda > 0\) such that

i) \(\partial G_g(\lambda) \neq \emptyset\);

ii) \(\forall x \in S\) with \(g(x) = \lambda\) there exists \(\{u_n\}_{n \in \mathbb{N}}\) and \(\{l_n\}_{n \in \mathbb{N}}\) such that \(u_n, l_n \to x\) and \(g(u_n) > \lambda, g(l_n) < \lambda\).

Then
\[
d_H(\partial G_g(\lambda), \partial \hat{G}_n(\lambda)) \rightarrow 0 \text{ a.s.}
\]

### 3.1 Level set estimation under shape restrictions

In this subsection we propose another estimator of the level sets, assuming a quite general
shape condition on the level sets. We assume that there exists an \(r > 0\) such that \(\overline{G}_g(\lambda)\)
is compact and \(r\)-convex, i.e. \(\overline{G}_g(\lambda) = C_r(G_g(\lambda))\), where
\[
C_r(G_g(\lambda)) = \bigcap_{\{ \mathcal{B}(x,r) : \mathcal{B}(x,r) \cap \overline{G}_g(\lambda) = \emptyset \}} \mathcal{B}(x,r)^c
\]
is the $r$-convex hull of $G_g(\lambda)$.

This condition has been extensively studied in set estimation, see for instance [15], [27] and [30]. It is also related to the level set estimation problem, see [36]. Although $r$-convexity is much less restrictive than convexity, inlets that are too sharp are not allowed, see Figure 1.

![Figure 1: A general $r$-convex set](image)

Following the notation in [19], let $\text{Unp}(S)$ be the set of points $x \in \mathbb{R}^d$ with a unique projection on $S$, denoted by $\xi_S(x)$. That is, for $x \in \text{Unp}(S)$, $\xi_S(x)$ is the unique point that attains the minimum of $\|x - y\|$ for $y \in S$. We write $\delta_S(x) = \inf\{\|x - y\| : y \in S\}$.

**Definition 6.** For $x \in S$, let $\text{reach}(S, x) = \sup\{r > 0 : \mathbb{B}(x, r) \subset \text{Unp}(S)\}$. The reach of $S$ is defined by $\text{reach}(S) = \inf \{\text{reach}(S, x) : x \in S\}$, and $S$ is said to be of positive reach if $\text{reach}(S) > 0$.

The relation between $r$-convexity, reach, and rolling type conditions, has been studied in [15].

**Definition 7.** The outer Minkowski content of $S \subset \mathbb{R}^d$ is given by

$$L_0(\partial S) = \lim_{\epsilon \to 0} \frac{\mu_L(B(S, \epsilon) \setminus S)}{\epsilon},$$

provided that the limit exists and is finite.

**Definition 8.** Let $S \subset \mathbb{R}^d$ be a closed set. A ball of radius $r$ is said to roll freely in $S$ if for each boundary point $s \in \partial S$ there exists some $x \in S$ such that $s \in \mathbb{B}(x, r) \subset S$. The set $S$ is said to satisfy the outside $r$-rolling condition if a ball of radius $r$ rolls freely in $S^c$.

We will also assume the following condition.

**HR:** A level set $\{g > \lambda\}$ fulfils HR if

(i) there exists $\delta > 0$ and $r > 0$ such that $\overline{G_g(\lambda + \varepsilon)}$ is $r$-convex for all $-\delta < \varepsilon < \delta,$ and

(ii) $d_H(\overline{G_g(\lambda + \varepsilon)}, \overline{G_g(\lambda - \varepsilon)}) \to 0$ as $\varepsilon \to 0.$
Theorem 2 in [36] gives sufficient conditions for HR (i) to hold, expressed in terms of the gradient of \( g \). More precisely, it is shown the following result.

**Theorem 2.** Let \( g : \mathbb{R}^d \to \mathbb{R} \) and \( -\infty < l \leq u < \sup g \). Assume

- \( g \in C^1(U) \) where \( U \) is a bounded open set that contains \( G_g([l-\eta]) \setminus G_g(u+\eta) \) for some \( \eta > 0 \);
- \( \nabla g \) satisfies \( \| \nabla g \| \geq m > 0 \) on \( U \) as well as a Lipschitz condition on \( U \) (or on \( \partial G_g(\lambda) \)) for all \( \lambda \in (l,u) \)

\[
\| \nabla g(x) - \nabla g(y) \| \leq k\| x - y \|,
\]

for \( x, y \in U \) (or in \( \partial G_g(\lambda) \)).

Then, for each \( \lambda \in (l,u) \), \( \overline{G_g(\lambda)} \) and \( G_g(\lambda)^c \) are \( r_0 \)-convex with \( r_0 = m/k \).

**Remark 4.** Part b) of the following Lemma can be derived from Remark 1, Section 2.4 of [14]. However, a short proof is included since similar ideas will be used to prove Lemma 5.

**Lemma 3.** Let \( g : S \to \mathbb{R} \) with \( S \subset \mathbb{R}^d \) a compact set. Then each of the following conditions implies HR (ii).

a) there exists \( \delta > 0 \) such that for all \( 0 \leq \varepsilon < \delta \) and all \( x \in G_g(\lambda - \varepsilon) \) there exists \( y = y(x) \in G_g(\lambda + \varepsilon) \) such that \( \| x - y \| \to 0 \) as \( \varepsilon \to 0 \).

b) \( g \in C^2(S) \) and \( \lambda \) is such that there exists \( 0 < \delta < \lambda \) for which \( \| \nabla g(x) \| \neq 0 \) for all \( x \in G_g(\lambda - \delta) \setminus G_g(\lambda + \delta) \).

**Proof.** (a) Since \( G_g(\lambda + \varepsilon) \subset G_g(\lambda - \varepsilon) \), it is enough to prove that for all \( \gamma > 0 \), \( G_g(\lambda - \varepsilon) \subset B(G_g(\lambda + \varepsilon), \gamma) \). Let \( \varepsilon < \min\{\gamma, \delta\} \) and \( x \in G_g(\lambda - \varepsilon) \). Then by HR (i) there exists a \( y(x) \in G_g(\lambda + \varepsilon) \) such that \( d(x, y) < \varepsilon < \gamma \), which implies \( x \in B(y, \gamma) \) and therefore \( x \in B(G_g(\lambda + \varepsilon), \gamma) \).

(b) We prove that (b) implies (a): let \( m = \min\{x \in G_g(\lambda - \delta) \setminus G_g(\lambda + \delta) \} \| \nabla g(x) \| \). Let \( x \in G_g(\lambda + \varepsilon) \) and \( y_t = x + t\nabla g(x) \). If \( t = 3\varepsilon/m^2 \), then \( \| y_t - x \| < \frac{3\varepsilon}{m} \varepsilon M \), where \( M = \max\{x \in G_g(\lambda - \delta) \setminus G_g(\lambda + \delta) \} \| \nabla g(x) \| \). Then from a Taylor expansion at \( x \), we get that

\[
g(y_t) = \lambda - \varepsilon + \frac{3\varepsilon}{m^2} \| \nabla g(x) \|^2 + \frac{9\varepsilon^2}{2m^4} \nabla g(x)^T H_\theta \nabla g(x),
\]

where \( H_\theta \) is the Hessian matrix of \( g \) at \( \theta \in [x, y_t] \), the segment joining \( x \) and \( y_t \). Since \( g \) is \( C^2 \), there exists a \( C > 0 \) such that \( \| \nabla g(x)^T H_\theta (\nabla g(x)) \| \leq C \| \nabla g(x) \|^2 \), from which it follows that for \( \varepsilon < 2m^4/(9M^2C) \),

\[
g(y_t) \geq \lambda + 2\varepsilon - \frac{9M^2C}{2m^4} \varepsilon^2 \geq \lambda + \varepsilon.
\]

Then \( y_t \in G_g(\lambda + \varepsilon) \). \( \square \)
Figure 2: If \( g'(x) = 0 \) for \( x \in G_g(\lambda) \), condition HR (ii) is not necessarily satisfied.

Consider \( \hat{g}_n \) as before. Assume that \( g \) fulfills HR. We will study the convergence in the Hausdorff distance of the following estimator:

\[
A_n(\lambda) = C_r \left( \{ X_i : \hat{g}_n(X_i) > \lambda \} \right).
\]

The rates of convergence for the estimator (25) in the independent case were obtained in [31], where an estimator of the parameter \( r \) was included. Observe that in our case it is not necessary to compute the whole set \( G_{\hat{g}_n}(\lambda) \) (which in practice is not feasible in most cases), as the estimator proposed in Corollary 1 is based just on the sample points which belong to the set \( G_{\hat{g}_n}(\lambda) \). Moreover, for the two dimensional case, the \( r \)-convex hull can be easily computed using the R software package alphahull (see [26]).

**Theorem 3.** Let \( S \subset \mathbb{R}^d \) be a compact set and \( \{ X_n \}_{n \in \mathbb{N}} \) a geometrically ergodic Markov chain with state space \( S \). Assume that its stationary distribution \( \pi \) has a continuous density \( g \) which satisfies HR. Let \( \hat{g}_n(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - X_i) \), where \( K \) is a Lipschitz density satisfying \( K(x) \leq k_1 I_{B(0,c_1)}(x) \) for some \( k_1, c_1 > 0 \). Let \( h = h_n \) be such that \( nh^d / \log(n) \to \infty \) and \( h^d \log(nh^d) \to 0 \). Then, with probability one,

\[
\lim_{n \to \infty} d_H(A_n(\lambda), G_g(\lambda)) = 0 \quad \text{and} \quad \lim_{n \to \infty} d_H(\partial A_n(\lambda), \partial G_g(\lambda)) = 0.
\]

**Proof.** From HR together with Theorem 3 in [15] it follows that it is enough to prove \( d_H(A_n(\lambda), G_g(\lambda)) \to 0 \) a.s. By Remark 1 and condition HR, for all \( \varepsilon < \delta \), \( A_n(\lambda) \subset G_g(\lambda - \varepsilon) \) for \( n \) large enough. Again by condition HR it is enough to prove that for \( n \) large enough, \( G_g(\lambda + \varepsilon) \subset A_n(\lambda) \) for \( \varepsilon < \delta \) a.s. Suppose by contradiction that there exists \( x \in G_g(\lambda + \varepsilon) \) such that \( x \notin A_n(\lambda) \) for infinitely many \( n \). Then \( x \in B(y_n, r) \cap \{ X_1, \ldots, X_n \} = \emptyset \) for infinitely many \( n \). Since \( x \in G_g(\lambda + \varepsilon) \), Theorem 1 implies that \( \hat{g}_n(x) > \lambda \) for \( n \) large enough, independent of \( x \),

\[
\lambda < \hat{g}_n(x) \leq \frac{k_1}{h^d} \frac{\# \{ B(x, c_1 h) \cap \{ X_1, \ldots, X_n \} \} }{n}.
\]
For all $x$, using Bernstein’s inequality for $Y_i = I_{B(0,h) \cap B(y_n,r)}(X_k) - \mathbb{E}_{\mu_0}(I_{B(0,h) \cap B(y_n,r)}(X_k))$, we get
\[
\sup_x \left| \frac{1}{n} \sum_{k=1}^{n} \left[ I_{B(x,h) \cap B(y_n,r)}(X_k) - \mathbb{E}_{\mu_0}(I_{B(x,h) \cap B(y_n,r)}(X_k)) \right] \right| \to 0 \text{ a.s.}
\]
Proceeding as in (17),
\[
\sup_x \left| \frac{1}{n} \sum_{k=1}^{n} \left[ \mathbb{E}_{\mu_0}(I_{B(x,h) \cap B(y_n,r)}(X_k)) - \mathbb{E}_{\pi}(I_{B(x,h) \cap B(y_n,r)}(X_k)) \right] \right| \to 0 \text{ a.s.,}
\]
and hence there are $a_n \to 0$ such that
\[
\frac{\# \{ B(x,c_1h) \cap \{ X_1, \ldots, X_n \} \cap B(y_n,r) \}}{n} \geq \int_{B(x,c_1h) \cap B(y_n,r)} g(t)dt - a_n \geq \frac{\lambda}{2} \mu_L(B(x,c_1h) \cap B(y_n,r)) \geq \frac{\lambda}{4} \mu_L(B(x,c_1h)).
\]
In the same way, for $n$ large enough and all $x \in G_g(\lambda + \varepsilon)$,
\[
\frac{\# \{ B(x,c_1h) \cap \{ X_1, \ldots, X_n \} \}}{n} \leq 2\lambda \mu_L(B(x,c_1h)).
\]
Hence
\[
\frac{\# \{ B(x,c_1h) \cap \{ X_1, \ldots, X_n \} \cap B(y_n,r) \}}{n} \geq \frac{1}{8} \frac{\# \{ B(x,c_1h) \cap \{ X_1, \ldots, X_n \} \}}{n} > \frac{\lambda h^d}{8k_1} > 0.
\]
Now $\# \{ B(x,c_1h) \cap \{ X_1, \ldots, X_n \} \cap B(y_n,r) \} > 0$ for $n$ large enough. \hfill \Box

### 3.2 Estimation of level sets with a fixed content

**Theorem 4.** Let $S \subset \mathbb{R}^d$ be a compact set and $\{ X_n \}_{n \in \mathbb{N}}$ a geometrically ergodic Markov chain with state space $S$. For $\tau \in (0,1)$, define $l_\tau = \inf \{ \lambda > 0 : \pi(G_g(\lambda)) \leq 1 - \tau \}$, $\pi$ being the stationary distribution. Assume that $\pi$ has a $C^2$ density $g$ such that $\| \nabla g(x) \| \neq 0$ for all $x \in U$, where $U$ is an open set containing $B_g(l_\tau - \varepsilon) \setminus B_g(l_\tau + \varepsilon)$ for some $\tau > 0$ and $0 < \varepsilon_0 < l_\tau$. Let $\hat{g}_n(x) = \frac{1}{n} \sum_{i=1}^{n} K_h(x - X_i)$ with $K$ a bounded Lipschitz density. Let $h = h_n$ be such that $h^d \log(nh^d) \rightarrow 0$ and $nh^d/\log(n) \rightarrow \infty$. If we define
\[
\hat{l}_\tau = \inf \{ \lambda > 0 : \frac{1}{n} \# \{ i : X_i \in G_{\hat{g}_n}(\lambda) \} \leq 1 - \tau \},
\]
then, with probability one,
\[
d_H(G_{\hat{g}_n}(\hat{l}_\tau), G_g(l_\tau)) \rightarrow 0. \quad (27)
\]

In order to prove Theorem 4 we will need two lemmas. For the first, recall that given a probability distribution $P$, $A$ is a $P$-uniformity class if $\sup_{A \in A} |P_n(A) - P(A)| \to 0$ whenever $P_n \to P$ weakly. Theorem 5 in [15] proves that the class of sets with reach bounded from below by a positive constant included in a compact set is a $P$-uniformity class.

14
Lemma 4. Let $S \subset \mathbb{R}^d$ be a compact set and $g : S \to \mathbb{R}$ a $C^2$ function such that there exists an $\varepsilon_0 > 0$ and a $c > 0$ such that $\| \nabla g(x) \| > c$ for all $x \in U$, where $U$ is an open set containing $G_g(l_r - \varepsilon_0) \setminus G_g(l_r + \varepsilon_0)$. Then $\{ G_g(\lambda) : l_r - \varepsilon_0/2 \leq \lambda \leq l_r + \varepsilon_0/2 \}$ is a $P$-uniformity class for all probability distributions $P$ on $S$ absolutely continuous w.r.t. Lebesgue measure.

Proof. It is enough to prove that there exists an $r > 0$ such that for all $l_r - \varepsilon_0 < \lambda < l_r + \varepsilon_0$, $\text{reach}(G_g(\lambda)) > r > 0$. By Theorem 2 and theorem 1 of [37], there exists an $r > 0$ such that for all $l_r - \varepsilon_0 < \lambda < l_r + \varepsilon_0$, $G_g(\lambda)$ satisfies the inner and outer $r$-rolling conditions. This together with lemma 2.3 in [27] implies that $\text{reach}(G_g(\lambda)) > r > 0$ for all $l_r - \varepsilon_0/2 \leq \lambda \leq l_r + \varepsilon_0/2$. $\square$

Lemma 5. Under the hypotheses of Lemma 4, for all $0 \leq \varepsilon < \varepsilon_0/2$ and all $l_r - \varepsilon < \lambda < l_r + \varepsilon$,

$$G_g(\lambda - \varepsilon) \setminus G_g(\lambda + \varepsilon) \subset B \left( \partial G_g(\lambda), \frac{3\varepsilon M}{c^2} \right),$$

where $M = \max_{\{x \in G_g(l_r - \varepsilon_0) \setminus G_g(l_r + \varepsilon_0)\}} \| \nabla g(x) \|$.

Proof. We first prove that for all $\varepsilon < \varepsilon_0/2$ and all $l_r - \varepsilon < \lambda < l_r + \varepsilon$,

$$d_H(G_g(\lambda + \varepsilon), G_g(\lambda - \varepsilon)) \leq \frac{3\varepsilon M}{c^2}.$$  (29)

In Lemma 3 we have proved that $y = x + \frac{3\varepsilon}{c^2} \nabla g(x) \in G_g(\lambda + \varepsilon)$ for all $x \in G_g(\lambda - \varepsilon)$ and then $d_H(G_g(\lambda - \varepsilon), G_g(\lambda + \varepsilon)) = \max_{x \in G_g(\lambda - \varepsilon)} d(x, G_g(\lambda + \varepsilon)) \leq \|x - y\| = \frac{3\varepsilon M}{c^2}$, from which there follows (29). If we take $x \in G_g(\lambda - \varepsilon)$ with $g(x) \leq \lambda$ and $y \in G_g(\lambda + \varepsilon)$, then there exists a $t \in [x, y]$ (the segment joining $x$ and $y$) such that $g(t) = \lambda$, and so $t \in \partial G_g(\lambda)$, which concludes the proof. $\square$

Proof of Theorem 4
By Remark 3 we have that $\sup_{x \in S} |g_n(x) - g(x)| \to 0$ a.s. We will prove that $\hat{l}_r \to l_r$ a.s.

Define $L(\lambda) = \pi(G_g(\lambda))$, $\hat{L}(\lambda) = \frac{1}{n} \# \{ i : X_i \in G_{g_n}(\lambda) \}$ and $\tilde{L}(\lambda) = \frac{1}{n} \# \{ i : X_i \in G_g(\lambda) \}$.

Write

$$\sup_{\frac{1}{2} l_r - \frac{\varepsilon}{2} \leq \lambda \leq \frac{1}{2} l_r + \frac{\varepsilon}{2}} |L(\lambda) - \hat{L}(\lambda)| \leq \sup_{\frac{1}{2} l_r - \frac{\varepsilon}{2} \leq \lambda \leq \frac{1}{2} l_r + \frac{\varepsilon}{2}} |L(\lambda) - \tilde{L}(\lambda)| + \sup_{\frac{1}{2} l_r - \frac{\varepsilon}{2} \leq \lambda \leq \frac{1}{2} l_r + \frac{\varepsilon}{2}} |\tilde{L}(\lambda) - \hat{L}(\lambda)|.$$  

$$|\tilde{L}(\lambda) - \hat{L}(\lambda)| = \frac{1}{n} \# \{ i : X_i \in G_g(\lambda) \} - \# \{ i : X_i \in G_{g_n}(\lambda) \} = \sup_{\frac{1}{2} l_r - \frac{\varepsilon}{2} \leq \lambda \leq \frac{1}{2} l_r + \frac{\varepsilon}{2}} \left( \# \{ i : X_i \in G_g(\lambda) \} - \# \{ i : X_i \in G_{g_n}(\lambda) \} \right).$$

Since $\sup_{x \in S} |g_n(x) - g(x)| \to 0$ a.s., we have that for all $\lambda$ and $\varepsilon$, $G_{g_n}(\lambda + \varepsilon) \subset G_g(\lambda - \varepsilon)$ with probability one, for $n$ large enough. Then, with probability one, for $n$ large enough, for all $0 \leq \varepsilon < \varepsilon_0/2$,

$$\sup_{\frac{1}{2} l_r - \frac{\varepsilon}{2} \leq \lambda \leq \frac{1}{2} l_r + \frac{\varepsilon}{2}} |\tilde{L}(\lambda) - \hat{L}(\lambda)| \leq \sup_{\frac{1}{2} l_r - \frac{\varepsilon}{2} \leq \lambda \leq \frac{1}{2} l_r + \frac{\varepsilon}{2}} \frac{2}{n} \# \{ i : X_i \in G_g(\lambda - \varepsilon) \} \setminus G_g(\lambda + \varepsilon) \}.$$
By Lemma 4, \( G_g(\lambda) \) is a P-uniformity class. Hence,

\[
\sup_{t_r - \frac{a_0}{2} \leq \lambda \leq t_r + \frac{a_0}{2}} \frac{1}{n} \# \left\{ i : X_i \in G_g(\lambda - \varepsilon) \setminus G_g(\lambda + \varepsilon) \right\} - \pi(G_g(\lambda - \varepsilon) \setminus G_g(\lambda + \varepsilon)) \rightarrow 0
\]

and

\[
\pi(G_g(\lambda - \varepsilon) \setminus G_g(\lambda + \varepsilon)) \leq g_1 \mu_L(G_g(\lambda - \varepsilon) \setminus G_g(\lambda + \varepsilon)),
\]

where \( g_1 = \max_{\varepsilon \in \mathbb{S}} g(\varepsilon) \). By Lemma 5,

\[
\sup_{t_r - \frac{a_0}{2} \leq \lambda \leq t_r + \frac{a_0}{2}} \mu_L(G_g(\lambda - \varepsilon) \setminus G_g(\lambda + \varepsilon)) \leq \sup_{t_r - \frac{a_0}{2} \leq \lambda \leq t_r + \frac{a_0}{2}} \mu_L(B(\partial G_g(\lambda), \frac{3\varepsilon M}{c^2})).
\]

For a fixed \( \varepsilon > 0, \mu_L(B(\partial G_g(\lambda), \frac{3\varepsilon M}{c^2})) \) is a continuous function of \( \lambda \), and so its maximum is attained in some \( \lambda_0 \in [l_r - \varepsilon_0/2, l_r + \varepsilon_0/2] \). Since \( \text{reach}(\partial G_g(\lambda_0)) > 0 \), the outer Minkowski content of \( G_g(\lambda_0) \) and \( G_g(\lambda_0)^c \) exist, and so by corollary 3 of [1],

\[
\sup_{t_r - \frac{a_0}{2} \leq \lambda \leq t_r + \frac{a_0}{2}} \mu_L(G_g(\lambda - \varepsilon) \setminus G_g(\lambda + \varepsilon)) = O(\varepsilon),
\]

from which it follows that

\[
\sup_{t_r - \frac{a_0}{2} \leq \lambda \leq t_r + \frac{a_0}{2}} |\hat{L}(\lambda) - \tilde{L}(\lambda)| \rightarrow 0.
\]

Using Lemma 4 it follows in the same way that

\[
\sup_{t_r - \frac{a_0}{2} \leq \lambda \leq t_r + \frac{a_0}{2}} |L(\lambda) - \hat{L}(\lambda)| \rightarrow 0,
\]

which implies that

\[
\sup_{t_r - \frac{a_0}{2} \leq \lambda \leq t_r + \frac{a_0}{2}} |L(\lambda) - \tilde{L}(\lambda)| \rightarrow 0.
\]

To prove that \( \tilde{L} \rightarrow L_r \) a.s., let \( 0 < \varepsilon < \varepsilon_0/2 \) and \( \gamma = \min\{1 - \tau - L(l_r + \varepsilon/2), L(l_r - \varepsilon/2) - (1 - \tau)\} \). Now observe that \( \gamma > 0 \) since \( L \) is decreasing in \( l_r - \varepsilon_0 \leq \lambda \leq l_r + \varepsilon_0 \). Let \( n \) be so large that \( \sup_{l_r - \varepsilon_0/2 \leq \lambda \leq l_r + \varepsilon_0/2} |L(\lambda) - \tilde{L}(\lambda)| < \gamma/2 \). Then \( l_r - \varepsilon/2 < l_r < l_r + \varepsilon/2 \).

To conclude the proof, observe that since \( \|\nabla g(x)\| > c \) for all \( x \in U \), where \( U \) is an open set containing \( \overline{G_g(l_r - \varepsilon_0)} \setminus G_g(l_r + \varepsilon_0) \), it follows that \( \{x : g(x) < \lambda\} = \{x : g(x) \leq \lambda\} \) for all \( l_r - \varepsilon_0 < \lambda < l_r + \varepsilon_0 \). Now we apply theorem 2.1 of [25], which implies that, with probability one,

\[
\sup_{t_r - \frac{a_0}{2} \leq \lambda \leq t_r + \frac{a_0}{2}} d_H(G_{g_n}(\lambda), G_g(\lambda)) \rightarrow 0.
\]

(30)

Finally the result follows since

\[
d_H(G_{g_n}(\hat{l}_r), G_g(l_r)) \leq d_H(G_{g_n}(\hat{l}_r), G_g(\hat{l}_r)) + d_H(G_g(\hat{l}_r), G_g(l_r))
\]

while (30) implies that the first term converges to zero, and the second one converges to zero by Lemma 5.
3.3 Drift estimation

A drift estimator can be easily derived from the stationary density estimator. Indeed, by Proposition 3, we have that the density of the stationary distribution is given by
\[ g(x) = ce^{-f(x)}1_D(x), \]
while the drift function is given by \(-\frac{1}{2}\nabla f(x)\). Therefore, we can define the plug-in estimator of the drift just replacing \(g\) by its estimate \(\hat{g}_n\) obtained from Theorem 1. Then we get that
\[ \hat{\mu}_1(x) = \frac{1}{2} \nabla \log(\hat{g}_n(x)). \] (31)
is an estimator of the drift. This estimate can be improved using Theorem 3 to estimate the set \(D\) by \(A_n(0)\). The case where there is no reflection was studied in [3], where an estimator of the drift function was introduced, based on a discretized version of the stochastic equation, for more general diffusion processes.

4 Examples

In this section we first assess through a simulation study, the performance of the r-convex hull of the sample points belonging to the level set of the estimator, proposed in (25). Then we show the results of applying this method to real data.

4.1 Simulations

The discrete version of the RBMD (1) is produced using the Euler scheme proposed in [4], in the following way. We first choose a step \(\delta > 0\), and denote by \(\text{sym}(z)\) the symmetric of the point \(z\) with respect to \(\partial S\). We start with \(X_0 = x\) and suppose that we have obtained \(X_i \in S\). To produce the following point, set
\[ Y_{i+1} = X_i + Z_i - \frac{1}{2} \nabla f(X_i), \]
where \(Z_i\) is a centred Gaussian random vector, independent w.r.t. \(Z_1, \ldots, Z_{i-1}\), with covariance matrix \(\delta(I_d)_{\mathbb{R}^2}\). Then
1. If \(Y_{i+1} \in S\), set \(X_{i+1} = Y_{i+1}\).
2. If \(Y_{i+1} \notin S\) and \(\text{sym}(Y_{i+1}) \in S\), set \(X_{i+1} = \text{sym}(Y_{i+1})\).
3. If \(Y_{i+1} \notin S\) and \(\text{sym}(Y_{i+1}) \notin S\), set \(X_{i+1} = X_i\).

In our example, we consider an RBMD in the set \(S = E \setminus B((4/5, 0), 1/2)\), where \(E = \{(x, y) \in \mathbb{R}^2: 4x^2/9 + y^2 \leq 1\}\), with drift function given by \(\nabla f(x, y) = (2x, 2y)\). The stationary density is
\[ g(x) = \frac{1}{c} \exp \left[-(x^2 + y^2)\right] 1_S(x, y) \quad \text{where} \quad c = \int_S \exp \left[-(x^2 + y^2)\right] dx dy. \] (32)
The trajectory is shown in Figure 3 for $\delta = 0.001$ in the first row, and $\delta = 0.003$ in the second row. The values for $N$ are 10,000; 50,000 and 100,000 in the first, second and third columns, respectively.

Figure 3: The trajectory of the RBMD, for different values of $\delta$ and $N$, in a), b) and c) $\delta = 0.001$ and $N = 10,000$, $N = 50,000$ and $N = 100,000$, respectively. In d), e) and f), $\delta = 0.003$ and $N = 10,000$, $N = 50,000$ and $N = 100,000$, respectively.

The function (32) is shown in Figure 4 a), while in b) there is shown the estimated density using a Gaussian kernel with bandwidth $h = 0.2$; in c) there is shown the estimated density using an Epanechnikov kernel with bandwidth $h = 0.4$. In both cases we have used the trajectory shown in Figure 3, with $\delta = 0.003$ and $N = 100,000$. Since we can estimate the support, we have forced the estimation to be 0 outside the estimation of the support.

For the level sets, we have considered the levels $\lambda = 0.44, 0.41, 0.34, 0.27$ and 0.03. Figure 5 a) shows the theoretical level sets for the considered values of $\lambda$, while in b) there are shown the corresponding estimated level sets. The estimation is based on the trajectory with $T = 0.003$ and $N = 100,000$ using (25) with $r = 0.4$. We have used the Gaussian kernel with $h = 0.1$. It is clear that the hole in the domain will produce border effects for the density estimation, and therefore for the level sets. A way to overcome this problem (which is computationally very expensive) is to first estimate the support using the $r$-convex hull of the trajectory and then use a variable bandwidth kernel estimate where the bandwidth is given by the lesser of a fixed $h$ and the distance from the point.
4.2 Real data examples

We considered a dataset from the Movebank database, where a natural barrier acts as a boundary of the animal’s movement. GPS collars were placed on elephants in Loango National Park in western Gabon. The area is protected by the Atlantic Ocean on the west and by Lagoon Iguéla on the east. Figure 7 a) shows in red the movement of an elephant with estimator $N = 1633$ for recorded positions. In blue we represent the boundary of the $r$-convex hull estimator for $r = 0.02$. The estimated density is shown in b), using the Gaussian kernel with bandwidth $h = 0.01$. The $r$-convex hulls of the level sets are shown in c) for $\lambda_1 = 100, \lambda_2 = 600, \lambda_3 = 1100, \lambda_4 = 1600$, and $r = 0.02$. In d) we represent the estimation of the drift, using (31) with $h = 0.5$.

5 Appendix

Here we include the proofs of the propositions stated in Section 2.

Proof of Proposition 1.

Let $\Omega = C([0, \infty), \mathbb{D})$ be the set of all continuous functions from $[0, \infty)$ to $\mathbb{D}$, equipped with the topology of uniform convergence on finite time intervals, and denote by $\mathcal{F}$ the corresponding Borel $\sigma$-algebra. For each $t \geq 0$ we denote by $Z: [0, \infty) \times \Omega \to \mathbb{D}$ the projection map defined by

$$Z_t(\omega) = \omega(t), \quad \omega \in \Omega.$$ 

From [32] p. 147 and 148 it follows, taking $\Phi(x) = d(x, \partial S)(\mathbb{I}_S(x) - \exp(-d(x, \partial S))\mathbb{I}_{S^c}(x))$ (using the fact that $\nabla d(x, \partial S) = (x - \xi_S(x))/d(x, \partial S)$ if $x \in S$, and $\nabla d(x, \partial S) =$
\[-(x - \xi_S(x))/d(x, \partial S)\text{ if } x \in S^c, \text{ see theorem 4.8 of [19]), } \rho = 0, f(t, u) = h(u) \text{ and } \gamma(x) = \mathbf{n}(x), \text{ that we can associate to each starting state } x \in \mathcal{D} \text{ a unique probability measure } \mathbb{P}_x \text{ on } (\Omega, \mathcal{F}) \text{ such that } \mathbb{P}_x(Z_0 = x) = 1 \text{ and } h(Z_t) - \int_0^t \mathcal{L}(h(Z_u))du \]
is a \mathbb{P}_x\text{-submartingale for all } h \in C^2_c(\mathcal{D}) \text{ such that } \langle \nabla h(x), \mathbf{n}(x) \rangle > 0. \text{ Since we are considering a time-independent drift and reflection, (see [32] p. 196 or theorem 12.1 of [28]), we obtain the uniqueness of the solution of (1) that satisfies the strong Markov property. To prove that there is a strong solution, it is enough to see that conditions A and B in [34] are satisfied. Since } \partial S \text{ is } C^2 \text{ it is in the hypotheses of Theorem 1 in [37] and then a ball of positive radius rolls freely in } S \text{ and } S^c. \text{ Reasoning as in proposition 3 in [12] there follow conditions A and B in [34].}

**Proof of Proposition 2.**
The proof is based on the ideas used to prove Proposition 1.4 (ii) in [6] and the following result (whose proof can be found in [10] 610–613):

\[
\inf_{(x,y) \in \mathcal{D} \times \mathcal{D}} p(0, x, t, y) = c_t > 0,
\]
where \( p(0, x, t, y) \) is the density function introduced in Remark 1. Let \( r > 0 \) and let \( x \) be such that \( \mathcal{B} = \mathcal{B}(x, r) \subset \mathcal{D} \). Then for all \( t \geq 1, \)

\[
\mathbb{P}_x(T_{\mathcal{B}} \leq t) \geq \mathbb{P}_x(T_{\mathcal{B}} \leq 1) \geq \int_{\mathcal{B}} p(0, x, 1, y) dy \geq c_1 \mu_L(\mathcal{B}) = c' > 0.
\]
(33)

By the Markov property, for every \( x \in \mathcal{D}, \mathbb{P}_x(T_{\mathcal{B}} \geq k) \leq (1 - c')^k, \) for all \( k \geq 1, \) which implies that

\[
\sup_{x \in \mathcal{D}} \mathbb{E}_x(T_{\mathcal{B}}) \leq \sup_{x \in \mathcal{D}} \sum_{k=0}^{\infty} \mathbb{P}_x(T_{\mathcal{B}} \geq k) < \infty.
\]

Observe that (33) holds for any ball \( \mathcal{B} \subset \mathcal{D} \), so we have proved that the RBMD is non-trap.

**Proof of Proposition 3**
By Lemma 1, the measure \( \pi \) is the stationary distribution if and only for all \( h \in C^2_c(\mathcal{D}) \) with \( \langle \mathbf{n}(x), \nabla h(x) \rangle = 0, \) for all \( x \in \partial G, \) one has that

\[ 0 = \int_D c e^{-f(x)} \mathcal{L} h(x) dx = \int_D c e^{-f(x)} \left[ \frac{1}{2} \Delta h(x) - \frac{1}{2} \langle \nabla f(x), \nabla h(x) \rangle \right] dx. \]

But this is a direct consequence of Green’s first identity:

\[- \int_D e^{-f(x)} \Delta h(x) = \text{amp:} \int_{\partial D} e^{-f(x)} \langle \nabla h(x), \mathbf{n}(x) \rangle d\sigma(x) + \int_D e^{-f(x)} \langle \nabla f(x), \nabla h(x) \rangle dx = \text{amp:} \int_D e^{-f(x)} \langle \nabla f(x), \nabla h(x) \rangle dx, \]

20
with $\sigma$ being the surface measure on $\partial D$.

**Proof of Proposition 4**

Proof. Let $x_0 \in D$ and $\eta > 0$ be such that $B(x_0, 3\eta) \subset D$. Since $\sup_{x \in D} E_x T_{B(x_0, \eta)} < \infty$, by the Markov inequality there exists an $n_1$ such that $\inf_{x \in D} P_x (T_{B(x_0, \eta)} \leq n_1) > 1/2$. Let $Z_t = x + B_t - \frac{1}{2} \int_0^t \nabla f(X_s) ds$ be the $d$-dimensional Brownian motion with drift given by $\frac{1}{2} \nabla f$. Observe that, since $|\nabla f(x)| < L$, by lemma 12 of [5], we have

$$P_x \left( \sup_{s \in [0,t]} \left| Z_s \right| < \eta \right) \geq 1 - \frac{\sqrt{dt} + Lt}{\eta}.$$ 

Now take $t_0$ small enough so that $1 - \frac{\sqrt{dt_0} + Lt_0}{\eta} =: p_0 > 0$. By the strong Markov property,

$$\inf_{x \in D} P_x (T_{B(x_0, \eta)} \leq n_1 \text{ and } X_t \in B(x_0, 2\eta) \text{ for } t \in [T_{B(x_0, \eta)}, T_{B(x_0, \eta)} + t_0]) > \frac{1}{2} p_0.$$ 

Let $Y = \inf \{ n \in \mathbb{N} : X_n \in B(x_0, 2\eta) \}$, then $\inf_{x \in D} P_x (Y \leq n_1 + t_0) > p_0/2$. Applying the Markov property at times $k[(n_1 + t_0)]$, $\sup_{x \in D} P_x (Y \geq k[(n_1 + t_0)]) \leq (1 - p_0/2)^k$, from which it follows that

$$\sup_{x \in D} E_x(Y) \leq \sup_{x \in D} \sum_{k=0}^{\infty} k[(n_1 + t_0)] P_x (Y \geq k[(n_1 + t_0)]) < \infty.$$ 

Applying theorem 16.0.2 of [40], we obtain, for every $n > 0$, that

$$\sup_{x \in D} \| P_x (X_n \in \cdot) - \pi(\cdot) \|_{TV} \leq c_3 e^{-c_4 n},$$

where $c_3, c_4$ are positive finite constants. Using the semigroup property of $\{X_t\}_{t \geq 0}$ and the fact that $\pi$ is invariant,

$$\sup_{x \in D} \| P_x (X_t \in \cdot) - \pi(\cdot) \|_{TV} = \sup_{x \in D} \left| \int_S P_y (X_{t-n} \in \cdot) dP_x (X_n \in dy) - \int_S P_y (X_{t-n} \in \cdot) \pi(y) \right| \leq \text{amp} \sup_{x \in D} \| P_x (X_n \in \cdot) - \pi(\cdot) \|_{TV},$$

for all $t$ and $n$, with $t \geq n$. 

**Acknowledgements**

We thank Professor K. Burdzy for helpful comments during the preparation of this manuscript. We also thank Dr. Stephen Blake, of the Max Planck Institute for Ornithology, for facilitating access to the data set that was used in this manuscript. All possible errors are the responsibility of the authors. We also thanks the associate editor and the referees for helpful suggestions that improved the manuscript.
References

[1] Ambrosio, L., Colesanti, A. and Villa, E. (2008). Outer Minkowski content for some classes of closed sets. *Math. Ann.* **342**, 727–748.

[2] Azéma, J., Kaplan-Duflo, M. and Revuz, D. (1967). Mesure invariante sur les classes récurrentes des processus de Markov. *Z. Wahrscheinlichkeit* **8**(3), 157–181.

[3] Bandi, F. M. and Moloche, G. (2008) On the functional estimation of multivariate diffusion processes. MPRA Paper. url: https://mpra.ub.uni-muenchen.de/43681/

[4] Bossy, M., Gobet, E., and Talay, D. (2004) Symmetrized Euler scheme for an efficient approximation of reflected diffusions. *J. Appl. Probab.*, **41**, 877–889.

[5] Bubeck, S., Eldan, R. and Lehec, J. Sampling from a log-concave distribution with projected Langevin Monte Carlo. (2015) Manuscript, http://arxiv.org/abs/1507.02564

[6] Burdzy, K., Chen, Z.-Q. and Marshall, D. E. (2006) Traps for reflected Brownian motion. *Math. Z.*, **252**, 103–132.

[7] Burt, W. H. (1943) Territoriality and Home Range Concepts as Applied to Mammals. *J. Mammal.*, **24**, 346–352.

[8] Campos, V. S. M. and Dorea, C. C. Y. (2005) Kernel estimation for stationary density of Markov chains with general state space. *Ann. Inst. Statist. Math.* **57**(3), 443–453.

[9] Castellana, J. V. and Leadbetter, M. R. (1986) On smoothed probability density estimation for stationary processes. *Stochastic Processes and their Applications* **21**(2), 179–193

[10] Cattiaux, P. (1992) Stochastic calculus and degenerate boundary value problems. *Ann. Inst. Fourier* **43**(3), 541–624.

[11] Cattiaux, P., León, J. R., Prieur, C. (2016) Invariant density estimation for a reflected diffusion using an Euler scheme *Preprint*.

[12] Cholaquidis, A., Fraiman, R., Lugosi, G. and Pateiro-López, B. Set estimation from reflected Brownian motion. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78**(5), 1057–1078.

[13] Collomb, G. (1984) Uniform complete convergence of the kernel predictor. *Wahrsch. Verw. Gebiete* **57**(3), 441–460.

[14] Cuevas, A., González-Manteiga, W., and Rodríguez-Casal, A. (2006) Plug-in estimation of general level sets. *Aust. N. Z. J. Stat.* **48**(1), 7–19.
[15] Cuevas, A., Fraiman, R. and Pateiro-López, B. (2012) On statistical properties of sets fulfilling rolling-type conditions. Adv. in Appl. Probab., 44, 311–239.

[16] Dalalyan, A. (2005) Sharp adaptative estimation of the drift function for ergodic diffusions. Ann. Statist. 33(6), 2507–2528.

[17] Dalalya, A. and Reiss, M. (2007) Asymptotic statistical equivalence for ergodic diffusions: The multidimensional case. Probability Theory and Related Fields 137(1–2), 25–47.

[18] Doukhan, P. and Neumann, M. (2007) Probability and moment inequalities for sums of weakly dependent random variables, with applications. Stochastic Processes and their Applications 117, 878–903.

[19] Federer, H. (1959) Curvature measures. Trans. Amer. Math. Soc., 93, 418–491.

[20] Gobet, E., Hoffmann, M., and Reiβ(2004) Nonparametric estimation of scalar diffusions based on low-frequency data. Ann. Stat. 32(5), 2223–2253.

[21] Harrison, J. M. and Williams, R. J. (1987) Multidimensional reflected Brownian motions having exponential stationary distributions. Ann. Probab., 15, 115–137.

[22] Härddie, W., Park, B. U. and Tsybakov, A. B. (1995) Estimation of non-sharp support boundaries J. Multivariate Anal. 55, 205–218.

[23] Hayne, D. W. (1949) Calculation of Size of Home Range. J. Mammal., 30, 1–18.

[24] Ikeda, N., and Watanabe, S. (1981) Stochastic Differential Equations and Diffusion Processes. Amsterdam: North-Holland, 1981.

[25] Molchanov, I. (1998) A limit theorem for solutions of inequalities. Scandinavian Journal of Statistics 25, 235–242.

[26] Pateiro-López, B. and Rodríguez-Casal, A. (2010) Generalizing the Convex Hull of a Sample: The R Package alphahull. J. Stat. Softw., 34(5), 1–28.

[27] Pateiro-López, B. and Rodríguez-Casal, A. (2009) Surface area estimation under convexity type assumptions. J. Nonparametr. Stat., 21(6), 729–741.

[28] Pinsky, R. G. (1995) Positive Harmonic Functions and Diffusion. Cambridge Univ. Press.

[29] Powell, R. A. (2000) Animal home ranges and territories and home range estimators. In Research Techniques in Animal Ecology: Controversies and Consequences (eds Boitani, L. & Fuller, T.), 65–110. New York: Columbia University Press.

[30] Rodríguez-Casal, A. (2007) Set estimation under convexity-type assumptions. Ann. Inst. Henri Poincaré Probab. Stat., 43, 763–774.
[31] Rodríguez-Casal, A. and Saavedra-Nieves, P. (2014) A fully data-driven method for estimating density level sets. https://arxiv.org/pdf/1411.7687v1.pdf.

[32] Stroock, D. W. and Varadhan, S. R. S. (1971) Diffusion Processes with Boundary Conditions. Comm. Pure Appl. Math. 24, 147–225.

[33] Stroock, D. W. and Varadhan, S. R. S. (1997) Multidimensional Diffusion Processes. Berlin: Springer-Verlag.

[34] Saisho, Y. (1987) Stochastic differential equations for multi-dimensional domain with reflecting boundary. Probab. Theory Related Fields, 74, 455–477.

[35] Tsybakov, A. B. (1997) On nonparametric estimation of density level sets. Ann. Statist. 25(3), 948–969.

[36] Walther, G. (1997) Granulometric smoothing. Ann. Statist., 25, 2273–2299.

[37] Walther, G. (1999) On a generalization of Blaschke’s Rolling Theorem and the Smoothing of Surfaces. Math. Methods Appl. Sci., 22, 301–316.

[38] Worton, B. J. (1987) A review of models of home range for animal movement. Ecol. Model., 38, 277–298.

[39] Worton, B. J. (1989) Kernel Methods for Estimating the Utilization Distribution in Home-Range Studies. Ecology, 70, 164–168.

[40] Meyn, S. P. and Tweedie, R. L. (1993a) Markov Chains and Stochastic Stability. London: Springer-Verlag.

[41] Meyn, S. P. and Tweedie, R. L. (1993b) Stability of Markovian processes II: Continuous-time processes and sampled chains Adv. Appl. Prob. 25, 487–517.

[42] Veretnikov, A. Yu. (1999) On Castellana–Leadbetter’s condition for diffusion density estimation. Statistical Inference for Stochastic Processes 2(1), 1–9.
Figure 5: a) Theoretical level sets. b) Estimation using (25) for $r = 0.4$, with Gaussian kernel and a bandwidth $h = 0.1$. In red the core-area.
Figure 6: (Left) Theoretical drift (Right) estimation with (31)
Figure 7: a) Trajectory and $r$-convex hull for $r = 0.02$ b) Estimation of the density using Gaussian Kernel with $h = 0.01$ c) $r$-convex hull of the level sets for $r = 0.02$ d) Estimation of the drift