KERNEL ESTIMATION OF DENSITY LEVEL SETS

Benoît CADRE[1]

UMR CNRS 5149, Equipe de Probabilités et Statistique
Université Montpellier II, CC 051, Place E. Bataillon
34095 Montpellier cedex 5, FRANCE

Abstract. Let f be a multivariate density and f_n be a kernel estimate of f drawn from the n-sample $X_1, \cdots, X_n$ of i.i.d. random variables with density f. We compute the asymptotic rate of convergence towards 0 of the volume of the symmetric difference between the t-level set $\{f \geq t\}$ and its plug-in estimator $\{f_n \geq t\}$. As a corollary, we obtain the exact rate of convergence of a plug-in type estimate of the density level set corresponding to a fixed probability for the law induced by f.

Key-words : Kernel estimate, Density level sets, Hausdorff measure.

2000 Mathematics Subject Classification : 62H12, 62H30.

Introduction. Stimulated by a growing demand from applied results, the theory of density level sets estimation has developed significantly over the last few years. One of the most important application of density level sets estimation is in unsupervised cluster analysis (see for instance, Hartigan [1], Cuevas, Febrero and Fraiman [2,3]). Here, one tries to break a complex data set into a series of piecewise similar groups or structures, each of which may then be regarded as a separate class of data, thus reducing overall data complexity. However, there are many other fields where the knowledge of density level sets is of great interest. For example, Devroye and Wise [4], Grenander [5], Cuevas [6], Cuevas and Fraiman [7] and Baíllo, Cuevas and Justel [8] studied the related problem of density support estimation for pattern recognition and for detection of the abnormal behavior of a system.

In this paper, we consider the problem of estimating the t-level set $\mathcal{L}(t)$ of a multivariate probability density f with support in $I\!R^k$ from independent random variables $X_1, \cdots, X_n$ with density f. Recall that for $t \geq 0$, the t-level set of the density f is defined as follows :

$$\mathcal{L}(t) = \{x \in I\!R^k \,:\, f(x) \geq t\}.$$

[1] cadre@math.univ-montp2.fr

The question is how to define the estimates of $\mathcal{L}(t)$ from the n-sample $X_1, \cdots, X_n$. Even in a nonparametric framework, there are many possible answers to this question, depending on the restrictions one can impose on the level set and the density under study (for a survey in set estimation, see Cuevas and Rodríguez-Casal [9]). One may find in the literature many types of estimators, such as the plug-in estimators (Baíllo, Cuesta-Albertos and Cuevas [10], Baíllo [11], Cuevas and Fraiman [7], Molchanov [12,13]), the estimators defined by an excess mass approach (Hartigan [14], Müller [15], Müller and Sawitzki [16], Nolan [17], Polonik [18], Tsybakhov [19]), the "naive" estimators (Devroye and Wise [4], Cuevas and Rodríguez-Casal [20], Walther [21] for a more sophisticated version of this idea), or the estimators constructed using a convex hull of the sample (Bräker, Hsing and Bingham [22], Dümbgen and Walther [23]). However, most of these techniques have several disadvantages when evaluated against the two main criteria, namely statistical performance and computational feasibility.

In this paper, we study a plug-in type estimator of the density level set $\mathcal{L}(t)$ which, regarding the statistical performance and computational feasibility, does not care about the specific shape of the level set and leads to easily computable sets estimators. Our estimator uses a kernel density estimate of f (Rosenblatt [24]) : given a kernel K on $\mathbb{R}^k$ (*i.e.*, a probability density on $\mathbb{R}^k$) and a bandwidth $h = h(n) > 0$ such that $h \to 0$ as n grows to infinity, the kernel estimate of f is given by

$$f_n(x) = \frac{1}{nh^k} \sum_{i=1}^{n} K\Big(\frac{x - X_i}{h}\Big), \ x \in \mathbb{R}^k.$$

We let the plug-in estimate $\mathcal{L}_n(t)$ of $\mathcal{L}(t)$ be defined as

$$\mathcal{L}_n(t) = \{x \in \mathbb{R}^k \ : \ f_n(x) \geq t\}.$$

Under various assumptions, Baíllo [11], Baíllo, Cuesta-Albertos and Cuevas [10], and Cuevas and Fraiman [7] obtained some rates of convergence. Let us also mention the general study by Molchanov [13] who obtained the exact rate of convergence for the Hausdorff metric. But, Molchanov's result may not be applied in our context since the basic assumption that the stochastic process $\xi_n(.) = \sqrt{nh^k}(f_n(.) - f(.))$ converges in law (for the topology induced by the uniform metric) is not satisfied. Indeed, the sequence of stochastic processes $(\xi_n(.))_n$ is not tight because, as easily seen, under the suitable assumptions, $\xi_n(x)$ and $\xi_n(y)$ are asymptotically normal and independent for all $x \neq y$ (see the necessary and sufficient condition for tightness in Billingsley [25], Theorem 8.2).

In this paper, the distance between two Borel sets in $I\!R^k$ is defined as a measure -in particular the volume or Lebesgue measure λ on $I\!R^k$- of the symmetric difference denoted Δ (*i.e.*, $A\Delta B = (A \cap B^c) \cup (A^c \cap B)$ for all sets A, B). Our main result (Theorem 2.1) deals with the limit law of

$$\sqrt{nh^k}\,\lambda\Big(\mathcal{L}_n(t)\Delta\mathcal{L}(t)\Big),$$

which is proved to be degenerate.

Consider now the following problem. In many statistical analyses, it is of interest to estimate the density level set corresponding to a fixed probability $p \in [0, 1]$ for the law induced by f. The data contained in this level set can then be regarded as the most important data if p is far enough from 0. Since f is unknown, the level t of this density level set is unknown as well. The natural estimate of the target density level set $\mathcal{L}(t)$ becomes $\mathcal{L}_n(t_n)$, where t_n is such that

$$\int_{\mathcal{L}_n(t_n)} f_n d\lambda = p.$$

As a consequence of our main result, we obtain in Corollary 2.1 the exact asymptotic rate of convergence of $\mathcal{L}_n(t_n)$ to $\mathcal{L}(t)$. More precisely, we prove that for some β_n which only depends on the data, one has :

$$\beta_n\sqrt{nh^k}\,\lambda\Big(\mathcal{L}_n(t_n)\Delta\mathcal{L}(t)\Big) \to \sqrt{\frac{2}{\pi}\int K^2 d\lambda}$$

in probability. This result improves some of the results in Baíllo [11] and Baíllo, Cuesta-Albertos and Cuevas [10] in which only a rate of convergence is obtained.

The precise formulations of Theorem 2.1 and Corollary 2.1 are given in Section 2. Section 3 is devoted to the proof of Theorem 2.1 while the proof of Corollary 2.1 is given in Section 4. Appendix A is dedicated to some technical points involving the (k-1)-dimensional Hausdorff measure, and Appendix B is devoted to the existence of t_n as defined above.

2. The main results.

2.1 Estimation of t-level sets. In the following, $\Theta \subset (0, \sup_{\mathbb{R}^k} f)$ denotes an open interval and $\|.\|$ stands for the Euclidean norm over any finite dimensional space. Let us introduce the hypotheses on the density f :
H1. f is twice continuously differentiable and $f(x) \to 0$ as $\|x\| \to \infty$;
H2. For all $t \in \Theta$,

$$\inf_{f^{-1}(\{t\})} \|\nabla f\| > 0,$$

where, here and in the following, $\nabla \psi(x)$ denotes the gradient at $x \in \mathbb{R}^k$ of the differentiable function $\psi : \mathbb{R}^k \to \mathbb{R}$. To understand the previous assumptions, let us mention the fact that under the conditions **H1**, **H2**, we have (see Proposition A2, Appendix A) :

$$\forall t \in \Theta : \quad \lambda(f^{-1}[t-\varepsilon, t+\varepsilon]) \to 0 \text{ as } \varepsilon \to 0,$$

a condition first used by Polonik [18]. Roughly speaking, this property means that the set where f is constant do not charge the Lebesgue measure on $\mathbb{R}^k$. (Note the importance of the assumption $0 \notin \Theta$: if the level $t = 0$, then $\lambda(f^{-1}[t-\varepsilon, t+\varepsilon]) = \lambda(f^{-1}[0, \varepsilon]) = \infty$.) Next, we introduce the assumptions on the kernel K :
H3. K is a continuously differentiable and compactly supported function. Moreover, there exists a monotone non increasing function $\mu : \mathbb{R}_+ \to \mathbb{R}$ such that $K(x) = \mu(\|x\|)$ for all $x \in \mathbb{R}^k$.
The assumption on the support of K is only provided for simplicity of the proofs. As a matter of fact, one could consider a more general class of kernels, such as the gaussian kernel for instance. Moreover, as we will use Pollard's results [26], K is assumed to be of the form $\mu(\|.\|)$.

Throughout the paper, $\mathcal{H}$ denotes the (k-1)-dimensional Hausdorff measure on $\mathbb{R}^k$ (*cf.* Evans and Gariepy [27]). Recall that $\mathcal{H}$ agrees with ordinary "(k-1)-dimensional surface area" on nice sets. Moreover, ∂A is the boundary of the set $A \subset \mathbb{R}^k$,

$$\alpha(k) = \begin{cases} 3 & \text{if } k = 1\,; \\ k+4 & \text{if } k \geq 2. \end{cases}$$

and for any bounded Borel function $g : \mathbb{R}^k \to \mathbb{R}_+$, λ_g stands for the measure defined for each Borel set $A \subset \mathbb{R}^k$ by

$$\lambda_g(A) = \int_A g \, d\lambda.$$

Finally, the notation $\stackrel{\mathrm{P}}{\to}$ denotes the convergence in probability.

It can be proved that if **H1**, **H3** hold and if $\lambda(\partial\mathcal{L}(t)) = 0$, one has :

$$\lambda\Big(\mathcal{L}_n(t)\Delta\mathcal{L}(t)\Big) \stackrel{\mathrm{P}}{\to} 0.$$

The aim of Theorem 2.1 below is to obtain the exact rate of convergence.

Theorem 2.1. *Let* $g : I\!R^k \to I\!R_+$ *be a bounded Borel function and assume that* **H1**-**H3** *hold. If* $nh^k/(\log n)^{16} \to \infty$ *and* $nh^{\alpha(k)}(\log n)^2 \to 0$*, then for almost all (a.a.)* $t \in \Theta$ *:*

$$\sqrt{nh^k}\,\lambda_g\Big(\mathcal{L}_n(t)\Delta\mathcal{L}(t)\Big) \stackrel{\mathrm{P}}{\to} \sqrt{\frac{2t}{\pi}\int K^2 d\lambda}\int_{\partial\mathcal{L}(t)} \frac{g}{\|\nabla f\|} d\mathcal{H}.$$

Remarks • Up to the log factors, our conditions on the bandwidth are quite similar (except for the univariate case) to the classical conditions $nh^k/\log n \to \infty$ and $nh^{k+4} \to 0$ in density estimation (see for instance Prakasa Rao [28]).
• Under the assumptions $nh^k/(\log n)^{16} \to \infty$ and $nh^{\alpha(k)}(\log n)^2 \to 0$, the convergence rate established in this result is of type $(nh^k)^{-1/2}$. If we restrict ourselves to choices of bandwidth such that $h = O(n^{-s})$ with $s > 0$, then the conditions on the bandwidth now read as : $1/(k+4) < s < 1/k$. The best convergence rate one can get (corresponding to values of s close to $1/(k+4)$) must be necessarily slower than $O(n^{-2/(k+4)})$, which is in turn strictly slower than the rates of type $O((\log n/n)^{2/(k+1)})$ obtained when $k > 3$ by Walther ([21], Theorem 4), using another estimator of the target level set.
• Notice that the rightmost integral is defined because g is bounded and $\mathcal{L}(t)$ is a compact set for all $t > 0$ according to **H1**.
• We which to emphasize that, from a statistical point of view, this result is essentially useful when $g \equiv 1$. Indeed, we then have the asymptotic behavior of the volume of the symmetric difference between the two level sets. The general case is provided for the proof of Corollary 2.1 below.
• If we only assume f to be Lipschitz instead of **H1**, then f is an almost everywhere continuously differentiable function by Rademacher's theorem and Theorem 2.1 holds under the additional assumption on the bandwidth : $nh^{k+2}(\log n)^2 \to 0$.

2.2 Estimation of level sets with fixed probability. Now, we will derive a corollary for the case where the level t of the target level set is unknown.

We shall assume that $\inf \Theta > 0$ –so that $\lambda(f^{-1}(\Theta)) < \infty$ under **H1**– and, for coherence of the assumptions, we shall consider the levels in an open interval Θ_s, which is a strict subset of Θ, *i.e.*,

$$\inf \Theta < \inf \Theta_s \text{ and } \sup \Theta > \sup \Theta_s.$$

Let us now denote by $\mathcal{P}$ the mapping defined on Θ_s by :

$$\mathcal{P}(t) = \lambda_f(\mathcal{L}(t)),\ t \in \Theta_s.$$

Observe that by Proposition A2, $\mathcal{P}$ is one-to-one if f satisfies **H1**, **H2**. Fix $p \in \mathcal{P}(\Theta_s)$. We denote by $t^{(p)} \in \Theta_s$ the unique real number such that $\lambda_f(\mathcal{L}(t^{(p)})) = p$. Let us now turn to the construction of a natural estimator of $t^{(p)}$. It is proved in Appendix B that under the assumptions of Corollary 2.1 below, with probability tending to 1, there exists $t_n^{(p)}$ with $\lambda_{f_n}(\mathcal{L}_n(t_n^{(p)})) = p$. For simplicity, we shall assume throughout that such a $t_n^{(p)}$ always exists. It is naturally defined as an estimator of $t^{(p)}$.

The aim of Corollary 2.1 below is to obtain the exact rate of convergence of $\mathcal{L}_n(t_n^{(p)})$ to $\mathcal{L}(t^{(p)})$. We also introduce an estimator of the unknown integral in Theorem 2.1.

Corollary 2.1. *Let $k \geq 2$, $(\alpha_n)_n$ be a sequence of positive real numbers such that $\alpha_n \to 0$ and assume that* **H1**-**H3** *hold. If $nh^{k+2}/\log n \to \infty$, $nh^{k+4}(\log n)^2 \to 0$ and $\alpha_n^2 nh^k/(\log n)^2 \to \infty$ then, for a.a. $p \in \mathcal{P}(\Theta_s)$:*

$$\sqrt{nh^k}\frac{\beta_n}{\sqrt{t_n^{(p)}}}\lambda\Big(\mathcal{L}_n(t_n^{(p)})\Delta\mathcal{L}(t^{(p)})\Big) \xrightarrow{\mathrm{P}} \sqrt{\frac{2}{\pi}\int K^2 d\lambda},$$

where $\beta_n = \alpha_n/\lambda(\mathcal{L}_n(t_n^{(p)}) - \mathcal{L}_n(t_n^{(p)} + \alpha_n))$.

Remarks • From a statistical point of view, it is interesting to mention the fact that under the assumptions of the corollary, we have for all $p \in \mathcal{P}(\Theta_s)$: $t_n^{(p)} \to t^{(p)}$ with probability 1 (see Lemma 4.3).

• When $k = 1$, the conditions of Theorem 2.1 on the bandwidth h do not permit to derive Corollary 2.1. In practice, estimations of density level sets and their applications to cluster analysis for instance are mainly interesting in high-dimensional problems.

3. Proof of Theorem 2.1.

3.1. Auxiliary results and proof of Theorem 2.1. For all $t > 0$, let

$$\mathcal{V}_n^t = f^{-1}\Big[t - \frac{(\log n)^\beta}{\sqrt{nh^k}}, t\Big] \quad \text{and} \quad \overline{\mathcal{V}}_n^t = f^{-1}\Big[t, t + \frac{(\log n)^\beta}{\sqrt{nh^k}}\Big],$$

where $\beta > 1/2$ is fixed. Moreover, $\tilde{K}$ stands for the real number :

$$\tilde{K} = \int K^2 d\lambda.$$

Proposition 3.1. *Let $g : \mathbb{R}^k \to \mathbb{R}_+$ be a bounded Borel function and assume that* **H1**-**H3** *hold. If $nh^k/(\log n)^{31\beta} \to \infty$ and $nh^{\alpha(k)}(\log n)^{2\beta} \to 0$, then for a.a. $t \in \Theta$:*

$$\begin{aligned} \lim_n \sqrt{nh^k} \int_{\mathcal{V}_n^t} P(f_n(x) \geq t) d\lambda_g(x) &= \lim_n \sqrt{nh^k} \int_{\overline{\mathcal{V}}_n^t} P(f_n(x) < t) d\lambda_g(x) \\ &= \sqrt{\frac{t\tilde{K}}{2\pi}} \int_{\partial\mathcal{L}(t)} \frac{g}{\|\nabla f\|} d\mathcal{H}. \end{aligned}$$

Proposition 3.2. *Let $g : \mathbb{R}^k \to \mathbb{R}_+$ be a bounded Borel function and assume that* **H1**-**H3** *hold. If $nh^k/(\log n)^{5\beta} \to \infty$ and $nh^{\alpha(k)}(\log n)^{2\beta} \to 0$, then for a.a. $t \in \Theta$:*

$$\lim_n nh^k \text{var}\Big[\lambda_g\Big(\mathcal{V}_n^t \cap \mathcal{L}_n(t)\Big)\Big] = 0 = \lim_n nh^k \text{var}\Big[\lambda_g\Big(\overline{\mathcal{V}}_n^t \cap \mathcal{L}_n(t)^c\Big)\Big].$$

Proof of Theorem 2.1. Let $t \in \Theta$ be such that both conclusions of Propositions 3.1 and 3.2 hold. According to **H3** and Pollard ([26], Theorem 37 and Problem 28, Chapter II), we have almost surely (a.s.) :

$$\sup_{\mathbb{R}^k} |f_n - Ef_n| \to 0.$$

Moreover, since both $\sup_n Ef_n(x)$ and $f(x)$ vanish as $\|x\| \to \infty$ by **H1**, **H3**, we have :

$$\sup_{\mathbb{R}^k} |Ef_n - f| \to 0.$$

Thus, a.s. and for n large enough :

$$\sup_{\mathbb{R}^k} |f_n - f| \leq \frac{t}{2}.$$

Consequently, $\mathcal{L}_n(t) \subset \mathcal{L}(t/2)$ and since $\mathcal{L}(t) \subset \mathcal{L}(t/2)$, we get :

$$\lambda_g\Big(\mathcal{L}_n(t)\Delta\mathcal{L}(t)\Big) = \int_{\mathcal{L}(t/2)} \mathbf{1}_{\{f_n<t, f\geq t\}} d\lambda_g + \int_{\mathcal{L}(t/2)} \mathbf{1}_{\{f_n\geq t, f<t\}} d\lambda_g. \qquad (3.1)$$

Let

$$A_n = \Big\{\sqrt{nh^k} \sup_{\mathcal{L}(t/2)} |f_n - f| \leq (\log n)^\beta\Big\}.$$

Since $\mathcal{L}(t/2)$ is a compact set by **H1**, it is a classical exercise to prove that $P(A_n) \to 1$ under the assumptions of the theorem. Hence, one only needs to prove that the result of Theorem 2.1 holds on the event A_n. But on A_n, one has according to (3.1) : $\lambda_g(\mathcal{L}_n(t)\Delta\mathcal{L}(t)) = J_n^1 + J_n^2$, where :

$$J_n^1 = \lambda_g\Big(\overline{\mathcal{V}}_n^t \cap \mathcal{L}_n(t)^c\Big) \text{ and } J_n^2 = \lambda_g\Big(\mathcal{V}_n^t \cap \mathcal{L}_n(t)\Big).$$

By Propositions 3.1 and 3.2, if $j = 1$ or $j = 2$:

$$\sqrt{nh^k} J_n^j \xrightarrow{\text{P}} \sqrt{\frac{t\tilde{K}}{2\pi}} \int_{\partial\mathcal{L}(t)} \frac{g}{\|\nabla f\|} d\mathcal{H}, \qquad (3.2)$$

if the bandwidth h satisfies $nh^{\alpha(k)}(\log n)^{2\beta} \to 0$ and $nh^k/(\log n)^{31\beta} \to \infty$. Letting $\beta = 16/31$, the theorem is proved •

3.2. Proof of Proposition 3.1. Let X be a random variable with density f,

$$V_n(x) = \operatorname{var} K\Big(\frac{x - X}{h}\Big) \text{ and } Z_n(x) = \frac{h^k\sqrt{n}}{\sqrt{V_n(x)}}(f_n(x) - Ef_n(x)),$$

for all $x \in \mathbb{R}^k$ such that $V_n(x) \neq 0$. Moreover, Φ denotes the distribution function of the $\mathcal{N}(0, 1)$ law.
In the proofs, c denotes a positive constant whose value may vary from line to line.

Lemma 3.1. *Assume that* **H1**, **H3** *hold and let* $\mathcal{C} \subset \mathbb{R}^k$ *be a compact set such that* $\inf_{\mathcal{C}} f > 0$. *Then, there exists* $c > 0$ *such that for all* $n \geq 1$, $x \in \mathcal{C}$ *and* $u \in \mathbb{R}$ *:*

$$|P(Z_n(x) \leq u) - \Phi(u)| \leq \frac{c}{\sqrt{nh^k}}.$$

Proof. By the Berry-Essèen inequality (*cf.* Feller [29]), one has for all $n \geq 1$, $u \in I\!\!R$ and $x \in I\!\!R^k$ such that $V_n(x) \neq 0$:

$$|P(Z_n(x) \leq u) - \Phi(u)| \leq \frac{3}{\sqrt{nV_n(x)^3}} E\Big|K\Big(\frac{x-X}{h}\Big) - EK\Big(\frac{x-X}{h}\Big)\Big|^3.$$

It is a classical exercise to deduce from **H1**, **H3** that

$$\sup_{x\in\mathcal{C}} E\Big|K\Big(\frac{x-X}{h}\Big) - EK\Big(\frac{x-X}{h}\Big)\Big|^3 \leq c\,h^k \text{ and } \inf_{x\in\mathcal{C}} V_n(x) \geq c\,h^k,$$

hence the lemma •

For all Borel bounded function $g : I\!\!R^k \to I\!\!R_+$, we let $\Theta_0(g)$ to be the set of $t \in \Theta$ such that :

$$\lim_{\varepsilon\searrow 0} \frac{1}{\varepsilon}\lambda_g\Big(f^{-1}[t-\varepsilon,t]\Big) = \lim_{\varepsilon\searrow 0} \frac{1}{\varepsilon}\lambda_g\Big(f^{-1}[t,t+\varepsilon]\Big) = \int_{\partial\mathcal{L}(t)} \frac{g}{\|\nabla f\|} d\mathcal{H}.$$

Lemma 3.2. *Let $g : I\!\!R^k \to I\!\!R_+$ be a bounded Borel function and assume that* **H1***,* **H2** *hold. Then we have $\Theta_0(g) = \Theta$ up to a λ-null set.*

Proof. According to **H1**, **H2**, for all $t \in \Theta$, there exists $\eta > 0$ such that :

$$\inf_{f^{-1}[t-\eta,t+\eta]} \|\nabla f\| > 0.$$

We deduce from Proposition A1 that for all $t \in \Theta$ and $\varepsilon > 0$ small enough :

$$\frac{1}{\varepsilon}\lambda_g\Big(f^{-1}[t-\varepsilon,t]\Big) = \frac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\partial\mathcal{L}(s)} \frac{g}{\|\nabla f\|} d\mathcal{H}\, ds.$$

Using the Lebesgue-Besicovitch theorem (*cf.* Evans and Gariepy [27], Theorem 1, Chapter I), we then have for a.a. $t \in \Theta$:

$$\lim_{\varepsilon\searrow 0} \frac{1}{\varepsilon}\lambda_g\Big(f^{-1}[t-\varepsilon,t]\Big) = \int_{\partial\mathcal{L}(t)} \frac{g}{\|\nabla f\|} d\mathcal{H},$$

and the same result holds for $\lambda_g(f^{-1}[t,t+\varepsilon])$ instead of $\lambda_g(f^{-1}[t-\varepsilon,t])$, hence the lemma •

Note that a straightforward consequence of Proposition A2 is $\lambda(\partial\mathcal{L}(t)) = 0$ for all $t \in \Theta$, since $0 \notin \Theta$. In particular,

$$\lambda\Big(f^{-1}[t-\varepsilon, t+\varepsilon]\Big) = \lambda\Big(f^{-1}(t-\varepsilon, t+\varepsilon)\Big),$$

for all $t \in \Theta$ and $\varepsilon > 0$ small enough.

We now let for $t \in \Theta$ and $x \in I\!R^k$ such that $f(x)V_n(x) \neq 0$:

$$t_n(x) = \sqrt{\frac{nh^k}{\tilde{K} f(x)}}(t - f(x)) \text{ and } \bar{t}_n(x) = \frac{h^k\sqrt{n}}{\sqrt{V_n(x)}}(t - Ef_n(x)),$$

and finally, $\overline{\Phi}(u) = 1 - \Phi(u)$ for all $u \in I\!R$.

Lemma 3.3. *Let $g : I\!R^k \to I\!R_+$ be a bounded Borel function and assume that* **H1**, **H2** *hold. If $nh^k/(\log n)^{2\beta} \to \infty$ and $nh^{k+4}(\log n)^{2\beta} \to 0$, then for all $t \in \Theta_0(g)$:*

$$\lim_n \sqrt{nh^k}\Big[\int_{\mathcal{V}_n^t} P(f_n(x) \geq t)d\lambda_g(x) - \int_{\mathcal{V}_n^t} \overline{\Phi}(t_n(x))d\lambda_g(x)\Big] = 0$$
$$\text{and} \quad \lim_n \sqrt{nh^k}\Big[\int_{\overline{\mathcal{V}}_n^t} P(f_n(x) < t)d\lambda_g(x) - \int_{\overline{\mathcal{V}}_n^t} \Phi(t_n(x))d\lambda_g(x)\Big] = 0.$$

Proof. We only prove the first equality. Let $t \in \Theta_0(g)$. First note that for all $x \in I\!R^k$ such that $V_n(x) \neq 0$:

$$P(f_n(x) \geq t) = P(Z_n(x) \geq \bar{t}_n(x)).$$

There exists a compact set $\mathcal{C} \subset I\!R^k$ such that $\inf_{\mathcal{C}} f > 0$ and $\mathcal{V}_n^t \subset \mathcal{C}$ for all n. Observe that by Lemma 3.1 and the above remarks,

$$\sqrt{nh^k}\Big[\int_{\mathcal{V}_n^t} P(f_n(x) \geq t)d\lambda_g(x) - \int_{\mathcal{V}_n^t} \overline{\Phi}(\bar{t}_n(x))d\lambda_g(x)\Big] \leq c\,\lambda_g(\mathcal{V}_n^t).$$

Since $\lambda_g(\mathcal{V}_n^t) \to 0$ by Lemma 3.2, one only needs now to prove that :

$$E_n := \sqrt{nh^k}\int_{\mathcal{V}_n^t} |\overline{\Phi}(\bar{t}_n(x)) - \overline{\Phi}(t_n(x))|d\lambda_g(x) \to 0.$$

One deduces from the Lipschitz property of Φ that

$$E_n \leq c\sqrt{nh^k}\lambda_g(\mathcal{V}_n^t) \sup_{x\in\mathcal{V}_n^t} |\bar{t}_n(x) - t_n(x)|. \qquad (3.3)$$

However, by definitions of $\bar{t}_n(x)$ and $t_n(x)$, we have for all $x \in \mathcal{V}_n^t$:

$$\begin{aligned}
&\frac{1}{\sqrt{nh^k}}|\bar{t}_n(x) - t_n(x)| \\
\leq\ & \left(|t - f(x)|\left|\frac{1}{\sqrt{\tilde{K}f(x)}} - \frac{1}{\sqrt{V_n(x)h^{-k}}}\right| + \sqrt{\frac{h^k}{V_n(x)}}|Ef_n(x) - f(x)|\right) \\
\leq\ & \left(\frac{(\log n)^\beta}{\sqrt{nh^k}}\sqrt{\frac{|\tilde{K}f(x) - V_n(x)h^{-k}|}{\tilde{K}f(x)V_n(x)h^{-k}}} + \sqrt{\frac{h^k}{V_n(x)}}|Ef_n(x) - f(x)|\right). \quad (3.4)
\end{aligned}$$

It is a classical exercise to deduce from **H1**, **H3** that, since $\mathcal{V}_n^t$ is contained in $\mathcal{C}$,

$$\sup_{x \in \mathcal{V}_n^t} |Ef_n(x) - f(x)| \leq c\,h^2,$$

and similarly, that

$$\sup_{x \in \mathcal{V}_n^t} |\tilde{K}f(x) - V_n(x)h^{-k}| \leq c\,h.$$

One deduces from (3.4) and above that

$$\sup_{x \in \mathcal{V}_n^t} |\bar{t}_n(x) - t_n(x)| \leq c\,(\sqrt{h}\,(\log n)^\beta + \sqrt{nh^{k+4}}).$$

Thus, by (3.3) and since $t \in \Theta_0(g)$, one has for all n large enough :

$$E_n \leq c\,(\log n)^\beta(\sqrt{h}\,(\log n)^\beta + \sqrt{nh^{k+4}}),$$

and the latter term vanishes by assumptions on h, hence the lemma •

Proof of Proposition 3.1. By Lemma 3.2, one only needs to prove Proposition 3.1 for all $t \in \Theta_0(g)$. Fix $t \in \Theta_0(g)$, and let

$$I_n := \int_{\mathcal{V}_n^t} \overline{\Phi}(t_n(x))d\lambda_g(x) \text{ and } \overline{I}_n := \int_{\overline{\mathcal{V}}_n^t} \Phi(t_n(x))d\lambda_g(x).$$

By Lemma 3.3, the task is now to prove that

$$\lim_n \sqrt{nh^k}\,I_n = \sqrt{\frac{t\tilde{K}}{2\pi}}\int_{\partial\mathcal{L}(t)} \frac{g}{\|\nabla f\|}d\mathcal{H} = \lim_n \sqrt{nh^k}\,\overline{I}_n.$$

We only show the first equality. One has

$$I_n = \frac{1}{\sqrt{2\pi\tilde{K}}}\int_{\mathcal{V}_n^t}\int_{b_n(x)}^{\infty} \exp\Big(-\frac{u^2}{2\tilde{K}}\Big)du\,d\lambda_g(x),$$

where for all $x \in I\!\!R^k$ such that $f(x) > 0$, $b_n(x) = \sqrt{nh^k}(t - f(x))/f(x)^{1/2}$. By Fubini's theorem :

$$I_n = \frac{1}{\sqrt{2\pi\tilde{K}}} \int_0^\infty \exp\Big(-\frac{u^2}{2\tilde{K}}\Big) \lambda_g\Big(f^{-1}\Big[\max\Big(t - \frac{(\log n)^\beta}{\sqrt{nh^k}}, \chi\Big(\frac{u}{\sqrt{nh^k}}\Big)^2\Big), t\Big]\Big) du,$$

where for all $v \geq 0$, $\chi(v) = -v/2 + (1/2)\sqrt{v^2 + 4t}$. It is straightforward to prove the equivalence :

$$u \in [0, r_n] \Leftrightarrow \chi\Big(\frac{u}{\sqrt{nh^k}}\Big)^2 \geq t - \frac{(\log n)^\beta}{\sqrt{nh^k}},$$

where $r_n = (\log n)^\beta / \sqrt{t - (\log n)^\beta (nh^k)^{-1/2}}$, so that one can split I_n into two terms, *i.e.*, $I_n = I_n^1 + I_n^2$, where

$$\begin{aligned} I_n^1 &= \frac{1}{\sqrt{2\pi\tilde{K}}} \int_0^{r_n} \exp\Big(-\frac{u^2}{2\tilde{K}}\Big) \lambda_g\Big(f^{-1}\Big[\chi\Big(\frac{u}{\sqrt{nh^k}}\Big)^2, t\Big]\Big) du \\ \text{and } I_n^2 &= \frac{1}{\sqrt{2\pi\tilde{K}}} \int_{r_n}^\infty \exp\Big(-\frac{u^2}{2\tilde{K}}\Big) \lambda_g\Big(f^{-1}\Big[t - \frac{(\log n)^\beta}{\sqrt{nh^k}}, t\Big]\Big) du. \end{aligned}$$

Since $t \in \Theta_0(g)$, one has for all n large enough :

$$\sqrt{nh^k}\, I_n^2 \leq c\,(\log n)^\beta \int_{r_n}^\infty \exp\Big(-\frac{u^2}{2\tilde{K}}\Big) du, \qquad (3.5)$$

and the rightmost term vanishes. Thus, it remains to compute the limit of $\sqrt{nh^k} I_n^1$. Using an expansion of χ in a neighborhood of the origin, we get

$$\lim_n \sqrt{nh^k}\, \lambda_g\Big(f^{-1}\Big[\chi\Big(\frac{u}{\sqrt{nh^k}}\Big)^2, t\Big]\Big) = u\sqrt{t} \int_{\partial\mathcal{L}(t)} \frac{g}{\|\nabla f\|} d\mathcal{H}, \qquad (3.6)$$

for all $u \geq 0$, since $t \in \Theta_0(g)$. Moreover, one deduces from Lemma 3.2 that for all n large enough and for all $u \in [0, r_n]$:

$$\begin{aligned} \sqrt{nh^k}\, \lambda_g\Big(f^{-1}\Big[\chi\Big(\frac{u}{\sqrt{nh^k}}\Big)^2, t\Big]\Big) &\leq c\sqrt{nh^k}\Big(t - \chi\Big(\frac{u}{\sqrt{nh^k}}\Big)^2\Big) \\ &\leq c\,u, \qquad (3.7) \end{aligned}$$

because $r_n/\sqrt{nh^k} \to 0$. Thus, according to (3.5)-(3.7) and the Lebesgue theorem :

$$\begin{aligned} \lim_n \sqrt{nh^k}\, I_n &= \lim_n \sqrt{nh^k}\, I_n^1 \\ &= \frac{1}{\sqrt{2\pi\tilde{K}}} \int_0^\infty \exp\Big(-\frac{u^2}{2\tilde{K}}\Big) u\sqrt{t} \int_{\partial\mathcal{L}(t)} \frac{g}{\|\nabla f\|} d\mathcal{H}\, du \\ &= \sqrt{\frac{t\tilde{K}}{2\pi}} \int_{\partial\mathcal{L}(t)} \frac{g}{\|\nabla f\|} d\mathcal{H}, \end{aligned}$$

hence the proposition •

3.3. Proof of Proposition 3.2. From now on, we introduce two random variables N_1, N_2 with law $\mathcal{N}(0,1)$ such that $N_1, N_2, X_1, X_2, \cdots$ are independent. We let

$$\sigma_n = \frac{1}{(\log n)^{2\beta} \log\log n}, \ \forall n \geq 2.$$

(As we will see later, the random variable $Z_n(x) + \sigma_n N_1$ -for instance- has a density with respect to the Lebesgue measure.) For simplicity, we assume in the following that under **H3**, the support of K is contained in the Euclidean unit ball of $I\!R^k$.

Lemma 3.4. *Let $g : I\!R^k \to I\!R_+$ be a bounded Borel function and assume that* **H2** *holds. If $nh^k/(\log n)^{2\beta} \to \infty$, then for all $t \in \Theta_0(g)$ there exists $c > 0$ such that for n large enough :*

$$\int_{\mathcal{V}_n^t} P\Big(\Big\{Z_n(x) \geq \bar{t}_n(x)\Big\}\Delta\Big\{Z_n(x) + \sigma_n N_1 \geq \bar{t}_n(x)\Big\}\Big) d\lambda_g(x) \leq c\, w_n;$$

$$\textit{and} \quad \int_{\overline{\mathcal{V}}_n^t} P\Big(\Big\{Z_n(x) < \bar{t}_n(x)\Big\}\Delta\Big\{Z_n(x) + \sigma_n N_1 < \bar{t}_n(x)\Big\}\Big) d\lambda_g(x) \leq c\, w_n,$$

where $w_n = (\log n)^\beta/(nh^k) + \sigma_n (\log n)^\beta/\sqrt{nh^k}$.

Proof. We only prove the first inequality. Let $t \in \Theta_0(g)$ and

$$P_n := \int_{\mathcal{V}_n^t} P\Big(\Big\{Z_n(x) \geq \bar{t}_n(x)\Big\}\Delta\Big\{Z_n(x) + \sigma_n N_1 \geq \bar{t}_n(x)\Big\}\Big) d\lambda_g(x).$$

By independence of N_1 and $Z_n(x)$, P_n is smaller than

$$\int_{\mathcal{V}_n^t} \int \exp\Big(-\frac{z^2}{2}\Big) P\Big(\Big\{Z_n(x) \geq \bar{t}_n(x)\Big\}\Delta\Big\{Z_n(x) + \sigma_n z \geq \bar{t}_n(x)\Big\}\Big) dz\, d\lambda_g(x),$$

and consequently,

$$P_n \leq \int_{\mathcal{V}_n^t} \int \exp\Big(-\frac{z^2}{2}\Big) P\Big(|Z_n(x) - \bar{t}_n(x)| \leq \sigma_n |z|\Big) dz\, d\lambda_g(x).$$

Since $t \in \Theta_0(g)$, one deduces from Lemma 3.1 that for n large enough :

$$\begin{aligned} P_n &\leq c\,\frac{\lambda_g(\mathcal{V}_n^t)}{\sqrt{nh^k}} + \int_{\mathcal{V}_n^t} \int \exp\Big(-\frac{z^2}{2}\Big) P\Big(|N_1 - \bar{t}_n(x)| \leq \sigma_n |z|\Big) dz\, d\lambda_g(x) \\ &\leq c\Big(\frac{(\log n)^\beta}{nh^k} + \frac{\sigma_n (\log n)^\beta}{\sqrt{nh^k}}\Big), \end{aligned}$$

hence the lemma •

Lemma 3.5. *Fix $t \in \Theta$ and assume that* **H1**, **H3** *hold. Then, there exists a polynomial function Q of degree 5 defined on $\mathbb{R}^2$ such that for all $(u_1, u_2) \in \mathbb{R}^2$ and n large enough :*

$$\Big| E \exp\Big(i\Big(u_1 Z_n(x) + u_2 Z_n(y)\Big)\Big) - E\exp\Big(iu_1 Z_n(x)\Big) E\exp\Big(iu_2 Z_n(y)\Big)\Big| \leq \frac{Q(|u_1|, |u_2|)}{\sqrt{nh^k}},$$

if $x, y \in \mathcal{V}_n^t \cup \overline{\mathcal{V}}_n^t$ are such that $\|x - y\| \geq 2h$.

Proof. First of all, fix $u_1, u_2 \in \mathbb{R}$, $x, y \in \mathcal{V}_n^t \cup \overline{\mathcal{V}}_n^t$ and consider the following quantities :

$$\begin{aligned} M_1 &:= \frac{u_1}{\sqrt{nV_n(x)}}\Big[K\Big(\frac{x-X}{h}\Big) - EK\Big(\frac{x-X}{h}\Big)\Big] \\ \text{and } M_2 &:= \frac{u_2}{\sqrt{nV_n(y)}}\Big[K\Big(\frac{y-X}{h}\Big) - EK\Big(\frac{y-X}{h}\Big)\Big]. \end{aligned}$$

One deduces from the inequality $|\exp(iw) - 1 - iw + w^2/2| \leq |w|\ \forall w \in \mathbb{R}$ that

$$\Big| E\exp\Big(i\Big(M_1 + M_2\Big)\Big) - 1 + \frac{1}{2}E(M_1 + M_2)^2\Big|$$

$$= \Big|E\Big[\exp\Big(i\Big(M_1 + M_2\Big)\Big) - 1 - i(M_1 + M_2) + \frac{1}{2}(M_1 + M_2)^2\Big]\Big| \leq E|M_1 + M_2|^3.$$

In a similar fashion, if $j = 1$ or $j = 2$:

$$\Big|E\exp(iM_j) - 1 + \frac{1}{2}EM_j^2\Big| = \Big|E\Big[\exp(iM_j) - 1 - iM_j + \frac{1}{2}M_j^2\Big]\Big| \leq E|M_j|^3.$$

Consequently,

$$\begin{aligned} &\Big|E\exp\Big(i\Big(M_1 + M_2\Big)\Big) - E\exp\Big(iM_1\Big)E\exp\Big(iM_2\Big)\Big| \\ \leq\ & E|M_1 + M_2|^3 + \Big|\Big(1 - \frac{1}{2}E|M_1 + M_2|^2\Big) - \Big(1 - \frac{1}{2}EM_1^2\Big)\Big(1 - \frac{1}{2}EM_2^2\Big)\Big| \\ &+ \Big|1 - \frac{1}{2}EM_1^2\Big|E|M_2|^3 + \Big|1 - \frac{1}{2}EM_2^2\Big|E|M_1|^3. \qquad (3.8) \end{aligned}$$

It is an easy exercise to prove that for all n large enough, one has $\inf V_n(x) \geq ch^k$, the infimum being taken over all $x \in \mathcal{V}_n^t \cup \overline{\mathcal{V}}_n^t$. Consequently, if $j = 1$ or $j = 2$:

$$E|M_j|^3 \leq c\,\frac{|u_j|^3}{\sqrt{n^3h^k}},$$

from which we deduce that :

$$E|M_1+M_2|^3 \le c\,\frac{|u_1|^3+|u_2|^3}{\sqrt{n^3h^k}}.$$

Moreover, $EM_1^2 = u_1^2/n$, $EM_2^2 = u_2^2/n$ and for all $x,y\in\mathcal{V}_n^t\cup\overline{\mathcal{V}}_n^t$ such that $\|x-y\|\ge 2h$:

$$E(M_1+M_2)^2 = EM_1^2+EM_2^2-\frac{u_1u_2}{n\sqrt{V_n(x)V_n(y)}}EK\Big(\frac{x-X}{h}\Big)EK\Big(\frac{y-X}{h}\Big),$$

because the support of K is contained in the unit ball and hence

$$EK\Big(\frac{x-X}{h}\Big)K\Big(\frac{y-X}{h}\Big)=0.$$

One deduces from above and (3.8) that for all $x,y\in\mathcal{V}_n^t\cup\overline{\mathcal{V}}_n^t$ such that $\|x-y\|\ge 2h$:

$$\begin{aligned}&\Big|E\exp\Big(i\Big(M_1+M_2\Big)\Big)-E\exp\Big(iM_1\Big)E\exp\Big(iM_2\Big)\Big|\\ \le\ & c\,\frac{|u_1|^3+|u_2|^3}{\sqrt{n^3h^k}}+\frac{(u_1u_2)^2}{n^2}+c\,\frac{|u_2|^3(1+u_1^2)+|u_1|^3(1+u_2^2)}{\sqrt{n^3h^k}}+c\,\frac{|u_1u_2|h^k}{n}.\end{aligned}$$

By assumption, $nh^{3k}\to 0$ so that for n large enough : $h^k\le 1/\sqrt{nh^k}$. Consequently,

$$\Big|E\exp\Big(i\Big(M_1+M_2\Big)\Big)-E\exp\Big(iM_1\Big)E\exp\Big(iM_2\Big)\Big|\le\frac{Q(|u_1|,|u_2|)}{\sqrt{nh^k}},$$

where Q is defined for all $u_1,u_2\in I\!R$ by :

$$Q(u_1,u_2)=c(u_1^3+u_2^3+(u_1u_2)^2+u_1u_2+u_2^2u_1^3+u_1^3u_2^2).$$

Consequently, for all $u_1,u_2\in I\!R$ and $x,y\in\mathcal{V}_n^t\cup\overline{\mathcal{V}}_n^t$ such that $\|x-y\|\ge 2h$:

$$\begin{aligned}&\Big|E\exp\Big(i\Big(u_1Z_n(x)+u_2Z_n(y)\Big)\Big)-E\exp\Big(iu_1Z_n(x)\Big)E\exp\Big(iu_2Z_n(y)\Big)\Big|\\ =\ &\Big|\Big(E\exp\Big(i\Big(M_1+M_2\Big)\Big)\Big)^n-\Big(E\exp\Big(iM_1\Big)E\exp\Big(iM_2\Big)\Big)^n\Big|\\ \le\ & n\Big|E\exp\Big(i\Big(M_1+M_2\Big)\Big)-E\exp\Big(iM_1\Big)E\exp\Big(iM_2\Big)\Big|\\ \le\ &\frac{Q(|u_1|,|u_2|)}{\sqrt{nh^k}},\end{aligned}$$

hence the lemma •

In the following, uv stands for the usual inner product of $u, v \in I\!R^2$.

Lemma 3.6. *Let $x, y \in I\!R^k$ be such that $V_n(x)V_n(y) \neq 0$. Then, the bivariate random variable*

$$\begin{pmatrix} Z_n(x) + \sigma_n N_1 \\ Z_n(y) + \sigma_n N_2 \end{pmatrix}$$

has a density $\varphi_n^{x,y}$ defined for all $u \in I\!R^2$ by

$$\varphi_n^{x,y}(u) = \frac{1}{4\pi^2} \int E\Big[\exp\Big(i\Big(v_1 Z_n(x) + v_2 Z_n(y)\Big)\Big)\Big] \exp\Big(-i\,uv - \frac{1}{2}\sigma_n^2 \|v\|^2\Big) dv.$$

Proof. By independence of $X_1, \cdots, X_n, N_1$ and N_2, the random variable

$$\begin{pmatrix} Z_n(x) \\ Z_n(y) \end{pmatrix} + \sigma_n \begin{pmatrix} N_1 \\ N_2 \end{pmatrix}$$

has a density $\varphi_n^{x,y}$ defined for all $u = (u_1, u_2) \in I\!R^2$ by

$$\varphi_n^{x,y}(u) = \frac{1}{2\pi\sigma_n^2} E\Big[\exp\Big(-\frac{(u_1 - Z_n(x))^2}{2\sigma_n^2}\Big) \exp\Big(-\frac{(u_2 - Z_n(y))^2}{2\sigma_n^2}\Big)\Big].$$

Using the equality

$$\frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\Big(-\frac{z^2}{2\sigma_n^2}\Big) = \frac{1}{2\pi} \int \exp\Big(-izw - \frac{1}{2}\sigma_n^2 w^2\Big) dw \;\forall z \in I\!R,$$

we deduce from the Fubini theorem that

$$\varphi_n^{x,y}(u) = \frac{1}{4\pi^2} \int E\Big[\exp\Big(i\Big(v_1 Z_n(x) + v_2 Z_n(y)\Big)\Big)\Big] \exp\Big(-iuv - \frac{1}{2}\sigma_n^2 \|v\|^2\Big) dv,$$

hence the lemma •

Proof of Proposition 3.2. We only prove the first equality of Proposition 3.2. According to Lemma 3.2, one only needs to prove the result for each $t \in \Theta_0(g)$. Hence we fix $t \in \Theta_0(g)$ and we put :

$$A_n(x) = \Big\{Z_n(x) \geq \bar{t}_n(x)\Big\},\; A_n^j(x) = \Big\{Z_n(x) + \sigma_n N_j \geq \bar{t}_n(x)\Big\},\; j = 1, 2,$$

for all $x \in I\!R^k$ such that $V_n(x) \neq 0$. First note that since the events $A_n(x)$ and $\{f_n(x) \geq t\}$ are equal, one has

$$\begin{aligned} &\mathrm{var}\Big[\lambda_g\Big(\mathcal{V}_n^t \cap \mathcal{L}_n(t)\Big)\Big] \\ = &\int_{(\mathcal{V}_n^t)^{\times 2}} \Big(P(A_n(x) \cap A_n(y)) - P(A_n(x))P(A_n(y))\Big) d\lambda_g^{\otimes 2}(x, y). \quad (3.9) \end{aligned}$$

But, by Lemma 3.4 and since $t \in \Theta_0(g)$, one has for all n large enough :

$$\begin{aligned}
& nh^k \int_{(\mathcal{V}_n^t)^{\times 2}} \Big(P(A_n(x) \cap A_n(y)) - P(A_n^1(x) \cap A_n^2(y))\Big) d\lambda_g^{\otimes 2}(x,y) \\
\leq\ & 2nh^k \lambda_g(\mathcal{V}_n^t) \int_{\mathcal{V}_n^t} P(A_n(x) \Delta A_n^1(x)) d\lambda_g(x) \\
\leq\ & c\,(\log n)^\beta \sqrt{nh^k} \Big(\frac{(\log n)^\beta}{nh^k} + \frac{\sigma_n (\log n)^\beta}{\sqrt{nh^k}} \Big) \\
\leq\ & c \Big(\frac{(\log n)^{2\beta}}{\sqrt{nh^k}} + \sigma_n (\log n)^{2\beta} \Big),
\end{aligned}$$

and the latter term tends to 0 by assumption. In a similar fashion, one can prove that

$$nh^k \int_{(\mathcal{V}_n^t)^{\times 2}} \Big(P(A_n(x)) P(A_n(y)) - P(A_n^1(x)) P(A_n^2(y)) \Big) d\lambda_g^{\otimes 2}(x,y) \to 0.$$

By the above results and (3.9), it remains to show that

$$nh^k \int_{(\mathcal{V}_n^t)^{\times 2}} \Big(P(A_n^1(x) \cap A_n^2(y)) - P(A_n^1(x)) P(A_n^2(y)) \Big) d\lambda_g^{\otimes 2}(x,y) \to 0. \quad (3.10)$$

Let $T(h) = \{(x,y) \in (\mathbb{R}^k)^{\times 2} : \|x-y\| \leq 2h\}$. According to the Fubini theorem,

$$\begin{aligned}
nh^k \lambda_g^{\otimes 2}\Big((\mathcal{V}_n^t)^{\times 2} \cap T(h) \Big) &= nh^k \int_{\mathcal{V}_n^t} \lambda_g \Big(\mathcal{V}_n^t \cap B(x, 2h) \Big) d\lambda_g(x) \\
&\leq nh^k \int_{\mathcal{V}_n^t} \lambda_g(B(x,2h)) d\lambda_g(x),
\end{aligned}$$

where $B(z,r)$ stands for the Euclidean closed ball with center at $z \in \mathbb{R}^k$ and radius $r > 0$. Since $t \in \Theta_0(g)$, one deduces that

$$\begin{aligned}
nh^k \lambda_g^{\otimes 2}\Big((\mathcal{V}_n^t)^{\times 2} \cap T(h) \Big) &\leq c\, nh^k \frac{(\log n)^\beta}{\sqrt{nh^k}} h^k \\
&\leq c \sqrt{nh^{3k} (\log n)^{2\beta}},
\end{aligned}$$

so that, by assumption on the bandwidth h :

$$\lim_n nh^k \lambda_g^{\otimes 2} \Big((\mathcal{V}_n^t)^{\times 2} \cap T(h) \Big) = 0.$$

Let now $\mathcal{S}_n = (\mathcal{V}_n^t)^{\times 2} \cap T(h)^c$. According to (3.10) and the above result, one only needs now to prove that :

$$nh^k \int_{\mathcal{S}_n} \Big(P(A_n^1(x) \cap A_n^2(y)) - P(A_n^1(x)) P(A_n^2(y)) \Big) d\lambda_g^{\otimes 2}(x,y) \to 0. \quad (3.11)$$

By Lemmas 3.5 and 3.6, one has for all $x, y \in \mathcal{S}_n$:

$$\begin{aligned}
& \Big| P(A_n^1(x) \cap A_n^2(y)) - P(A_n^1(x)) P(A_n^2(y)) \Big| \\
\leq \ & \int \Big| E \exp \Big(i \Big(u_1 Z_n(x) + u_2 Z_n(y) \Big) \Big) \\
& \quad - E \exp \Big(i u_1 Z_n(x) \Big) E \exp \Big(i u_2 Z_n(y) \Big) \Big| \exp \Big(- \frac{1}{2} \sigma_n^2 \|u\|^2 \Big) du_1 du_2 \\
\leq \ & \frac{1}{\sqrt{nh^k}} \int Q(|u_1|, |u_2|) \exp \Big(- \frac{1}{2} \sigma_n^2 \|u\|^2 \Big) du_1 du_2 \\
\leq \ & \frac{c}{\sigma_n^7 \sqrt{nh^k}},
\end{aligned}$$

where Q is the polynomial function defined in Lemma 3.5. Consequently, one has for all n large enough :

$$\begin{aligned}
& nh^k \int_{\mathcal{S}_n} \Big(P(A_n^1(x) \cap A_n^2(y)) - P(A_n^1(x)) P(A_n^2(y)) \Big) d\lambda_g^{\otimes 2}(x,y) \\
\leq \ & c \frac{\sqrt{nh^k}}{\sigma_n^7} \lambda_g^{\otimes 2}(\mathcal{S}_n) \\
\leq \ & c \frac{\sqrt{nh^k}}{\sigma_n^7} \lambda_g(\mathcal{V}_n^t)^2 \\
\leq \ & c \frac{(\log n)^{2\beta}}{\sigma_n^7 \sqrt{nh^k}},
\end{aligned}$$

which tends to 0 by assumption, hence (3.11) •

4. Proof of Corollary 2.1.

Lemma 4.1. *Let $k \geq 2$ and assume that* **H1**-**H3** *hold. If $nh^{k+4}(\log n)^2 \to 0$ and $nh^k/(\log n)^{16} \to \infty$, then for a.a. $t \in \Theta_s$:*

$$\sqrt{nh^k} \Big(\lambda_{f_n}(\mathcal{L}(t)) - \lambda_{f_n}(\mathcal{L}_n(t)) \Big) \xrightarrow{\mathrm{P}} 0.$$

Proof. Let $t \in \Theta_s$ be such that the conclusion of Theorem 2.1 holds both for $g \equiv f$ and $g \equiv 1$. Notice that

$$\begin{aligned}\lambda_{f_n}(\mathcal{L}(t)) - \lambda_{f_n}(\mathcal{L}_n(t)) &= \int f_n\Big(\mathbf{1}_{\{f\geq t\}} - \mathbf{1}_{\{f_n\geq t\}}\Big)d\lambda \\ &= \int_{\mathcal{L}(t)} f_n\mathbf{1}_{\{f_n<t\}}d\lambda - \int_{\mathcal{L}(t)^c} f_n\mathbf{1}_{\{f_n\geq t\}}d\lambda.\end{aligned}$$

As in the proof of Theorem 2.1, we see that the result of the lemma will hold if we show that $\sqrt{nh^k}K_n \xrightarrow{\text{P}} 0$, where

$$K_n := \int_{\overline{\mathcal{V}}_n^t} f_n\mathbf{1}_{\{f_n<t\}}d\lambda - \int_{\mathcal{V}_n^t} f_n\mathbf{1}_{\{f_n\geq t\}}d\lambda.$$

Split K_n into four terms as follows :

$$\begin{aligned}K_n &= \int_{\overline{\mathcal{V}}_n^t}(f_n - f)\mathbf{1}_{\{f_n<t\}}d\lambda - \int_{\mathcal{V}_n^t}(f_n - f)\mathbf{1}_{\{f_n\geq t\}}d\lambda \\ &\quad + \int_{\overline{\mathcal{V}}_n^t}\mathbf{1}_{\{f_n<t\}}d\lambda_f - \int_{\mathcal{V}_n^t}\mathbf{1}_{\{f_n\geq t\}}d\lambda_f. \qquad (4.1)\end{aligned}$$

On one hand, it is a classical exercise to deduce from **H1**, **H3** that

$$\sup_{\overline{\mathcal{V}}_n^t}|f_n - f| \xrightarrow{\text{P}} 0.$$

Thus, using (3.2),

$$\sqrt{nh^k}\int_{\overline{\mathcal{V}}_n^t}(f_n - f)\mathbf{1}_{\{f_n<t\}}d\lambda \xrightarrow{\text{P}} 0.$$

In a similar fashion :

$$\sqrt{nh^k}\int_{\mathcal{V}_n^t}(f_n - f)\mathbf{1}_{\{f_n\geq t\}}d\lambda \xrightarrow{\text{P}} 0.$$

On the other hand, we get from (3.2) that :

$$\lim_n \sqrt{nh^k}\int_{\mathcal{V}_n^t}\mathbf{1}_{\{f_n\geq t\}}d\lambda_f = \lim_n \sqrt{nh^k}\int_{\overline{\mathcal{V}}_n^t}\mathbf{1}_{\{f_n<t\}}d\lambda_f,$$

where the limits are in probability. By the above results and (4.1), $\sqrt{nh^k}K_n$ tends to 0 in probability, hence the lemma •

Lemma 4.2. *Let $k \geq 2$, $t \in \Theta_s$ and assume that* **H1**, **H3** *hold. If $nh^{k+4} \to 0$, then :*

$$\sqrt{nh^k}\Big(\lambda_f(\mathcal{L}(t)) - \lambda_{f_n}(\mathcal{L}(t))\Big) \xrightarrow{\text{P}} 0.$$

Proof. Observe that

$$\lambda_f(\mathcal{L}(t)) - \lambda_{f_n}(\mathcal{L}(t)) = \int_{\mathcal{L}(t)} (f - Ef_n) d\lambda + \int_{\mathcal{L}(t)} (Ef_n - f_n) d\lambda.$$

According to **H1**, **H3**, we have :

$$\int_{\mathcal{L}(t)} |f - Ef_n| d\lambda \leq ch^2,$$

and since $nh^{k+4} \to 0$, we only need to prove that

$$\sqrt{nh^k} \int_{\mathcal{L}(t)} (Ef_n - f_n) d\lambda \xrightarrow{\mathrm{P}} 0.$$

We prove that this convergence holds in quadratic mean. We have :

$$\begin{aligned} E\Big(\sqrt{nh^k} \int_{\mathcal{L}(t)} (Ef_n - f_n) d\lambda\Big)^2 &\leq \frac{1}{h^k} E\Big(\int_{\mathcal{L}(t)} K\Big(\frac{x - X}{h}\Big) dx\Big)^2 \\ &\leq \frac{1}{h^k} \int_{\mathcal{L}(t)^{\times 2}} EK\Big(\frac{x - X}{h}\Big) K\Big(\frac{y - X}{h}\Big) dx dy. \end{aligned}$$

Recall that we assume in Section 3.3 that the support of K is contained in the unit ball so that if $\|x - y\| \geq 2h$,

$$EK\Big(\frac{x - X}{h}\Big) K\Big(\frac{y - X}{h}\Big) = 0.$$

Letting $R(h) = \{(x, y) \in \mathcal{L}(t)^{\times 2} : \|x - y\| \leq 2h\}$, one deduces from above that

$$\begin{aligned} E\Big(\sqrt{nh^k} \int_{\mathcal{L}(t)} (Ef_n - f_n) d\lambda\Big)^2 &\leq \frac{c}{h^k} \int_{R(h)} \int K\Big(\frac{x - u}{h}\Big) f(u) du dx dy \\ &\leq c \int_{R(h)} \int K(v) f(x - hv) dv dx dy \\ &\leq c\, \lambda^{\otimes 2}(R(h)) \\ &\leq c \int_{\mathcal{L}(t)} \lambda\Big(\mathcal{L}(t) \cap B(x, 2h)\Big) dx, \end{aligned}$$

according to the Fubini theorem. Thus, we get :

$$E\Big(\sqrt{nh^k} \int_{\mathcal{L}(t)} (Ef_n - f_n) d\lambda\Big)^2 \leq ch^k,$$

hence the lemma •

Lemma 4.3. *Let $p \in \mathcal{P}(\Theta_s)$ and assume that* **H1-H3** *hold. If $nh^k/\log n \to \infty$, then $t_n^{(p)} \to t^{(p)}$ a.s.*

Proof. Let $t = t^{(p)}$ and $t_n = t_n^{(p)}$. As seen in the proof of Theorem 2.1, $\sup_{\mathbb{R}^k} |f_n - f| \to 0$ a.s. Hence, one can fix

$$\omega \in \Big\{ \sup_{\mathbb{R}^k} |f_n - f| \to 0 \Big\}.$$

For notational convenience, we omit ω until the end of this proof. Since for large values of n, t_n is contained in the closed interval $I \subset \Theta$ (see Appendix B), from each sequence of integers, one can extract a subsequence $(n_k)_k$ such that $t_{n_k} \to t^*$, where $t^* \in I$. On one hand, according to Scheffé's theorem,

$$\lim_n \Big(\lambda_{f_{n_k}}(\mathcal{L}_{n_k}(t_{n_k})) - \lambda_f(\mathcal{L}_{n_k}(t_{n_k})) \Big) = 0, \qquad (4.2)$$

since both f and f_{n_k} are density functions on $\mathbb{R}^k$ and

$$\Big| \lambda_{f_{n_k}}(\mathcal{L}_{n_k}(t_{n_k})) - \lambda_f(\mathcal{L}_{n_k}(t_{n_k})) \Big| \leq \int |f_{n_k} - f| d\lambda.$$

On the other hand, letting $\varepsilon_k = \sup_{\mathbb{R}^k} |f_{n_k} - f|$, one observes that

$$\begin{aligned} \Big| \lambda_f(\mathcal{L}(t_{n_k})) - \lambda_f(\mathcal{L}_{n_k}(t_{n_k})) \Big| &= \int f \Big| \mathbf{1}_{\{f \geq t_{n_k}\}} - \mathbf{1}_{\{f_{n_k} \geq t_{n_k}\}} \Big| d\lambda \\ &\leq \int f \mathbf{1}_{\{t_{n_k} - \varepsilon_k \leq f \leq t_{n_k} + \varepsilon_k\}} d\lambda \\ &\leq c\, \lambda\Big(f^{-1}([t_{n_k} - \varepsilon_k, t_{n_k} + \varepsilon_k] \cap (0, \sup_{\mathbb{R}^k} f]) \Big), \end{aligned}$$

and the latter term tends to 0 as $k \to \infty$ according to Proposition A2. One deduces from (4.2) that :

$$\begin{aligned} \lim_n \Big(\lambda_f(\mathcal{L}(t)) - \lambda_f(\mathcal{L}(t_{n_k})) \Big) &= \lim_n \Big(p - \lambda_f(\mathcal{L}(t_{n_k})) \Big) \\ &= \lim_n \Big(\lambda_{f_{n_k}}(\mathcal{L}_{n_k}(t_{n_k})) - \lambda_f(\mathcal{L}_{n_k}(t_{n_k})) \Big) \\ &\quad + \lim_n \Big(\lambda_f(\mathcal{L}_{n_k}(t_{n_k})) - \lambda_f(\mathcal{L}(t_{n_k})) \Big) \\ &= 0. \qquad (4.3) \end{aligned}$$

Moreover, the map $s \mapsto \lambda_f(\mathcal{L}(s))$ defined on I is continuous according to Proposition A2. Consequently, one has

$$\lim_n \lambda_f(\mathcal{L}(t_{n_k})) = \lambda_f(\mathcal{L}(t^*)),$$

and thus, by (4.3), $\lambda_f(\mathcal{L}(t)) = \lambda_f(\mathcal{L}(t^*))$ and hence $t = t^*$ because $\mathcal{P}$ is one-to-one. One conclude $t_n \to t$ since we proved that from each sequence of integers, one can extract a subsequence $(n_k)_k$ such that $t_{n_k} \to t$. The lemma is proved •

Lemma 4.4. *Let $k \geq 2$ and assume that* **H1**-**H3** *hold. If $nh^{k+4}(\log n)^2 \to 0$ and $nh^{k+2}/\log n \to \infty$, then for a.a. $p \in \mathcal{P}(\Theta_s)$:*

$$\sqrt{nh^k} \int_{t_n^{(p)}}^{t^{(p)}} \int_{\partial \mathcal{L}_n(s)} \frac{1}{\|\nabla f_n\|} d\mathcal{H}\, ds \xrightarrow{\mathrm{P}} 0.$$

Proof. One only needs to choose $p \in \mathcal{P}(\Theta_s)$ such that the conclusion of Lemma 4.1 holds for $t^{(p)}$. For simplicity, let $t = t^{(p)}$ and $t_n = t_n^{(p)}$. It is a classical exercise to prove that since $nh^{k+2}/\log n \to \infty$ and $nh^{k+4} \to 0$,

$$\|\nabla f_n\| \to \|\nabla f\| \text{ a.s.,}$$

uniformly over the compact sets. Thus, by Lemma 4.3 and **H2**, we have a.s. and for n large enough :

$$\inf_{f^{-1}[\min(t_n,t),\max(t_n,t)]} \|\nabla f_n\| > 0. \qquad (4.4)$$

We deduce from Proposition A1 that a.s. and for n large enough :

$$\begin{aligned}
\lambda_{f_n}(\mathcal{L}_n(t_n)) - \lambda_{f_n}(\mathcal{L}_n(t)) &= \int \Big(\mathbf{1}_{\{f_n \geq t_n\}} - \mathbf{1}_{\{f_n \geq t\}}\Big) d\lambda_{f_n} \\
&= \int \mathbf{1}_{\{t_n \leq f_n < t\}} d\lambda_{f_n} - \int \mathbf{1}_{\{t \leq f_n < t_n\}} d\lambda_{f_n} \\
&= \int_{t_n}^{t} \int_{\partial \mathcal{L}_n(s)} \frac{f_n}{\|\nabla f_n\|} d\mathcal{H}\, ds,
\end{aligned}$$

where the latter integral is defined according to (4.4). Consequently,

$$\Big|\lambda_{f_n}(\mathcal{L}_n(t_n)) - \lambda_{f_n}(\mathcal{L}_n(t))\Big| = \int_{\min(t_n,t)}^{\max(t_n,t)} s \int_{\partial \mathcal{L}_n(s)} \frac{1}{\|\nabla f_n\|} d\mathcal{H}\, ds.$$

By Lemma 4.3, one has a.s. and for n large enough : $t_n \geq t/2$. Since $\lambda_{f_n}(\mathcal{L}_n(t_n)) = p = \lambda_f(\mathcal{L}(t))$, one deduces that :

$$\Big|\lambda_f(\mathcal{L}(t)) - \lambda_{f_n}(\mathcal{L}_n(t))\Big| \geq \frac{t}{2} \int_{\min(t_n,t)}^{\max(t_n,t)} \int_{\partial \mathcal{L}_n(s)} \frac{1}{\|\nabla f_n\|} d\mathcal{H}\, ds.$$

We can now conclude the proof of the lemma because

$$\sqrt{nh^k}\Big|\lambda_f(\mathcal{L}(t)) - \lambda_{f_n}(\mathcal{L}_n(t))\Big| \xrightarrow{\mathrm{P}} 0,$$

by Lemmas 4.1 and 4.2 •

Lemma 4.5 *Assume that* **H1**-**H3** *hold. If* $nh^k/(\log n)^2 \to \infty$*, then for a.a.* $p \in \mathcal{P}(\Theta_s)$ *:*

$$\frac{\sqrt{nh^k}}{\log n}\,|t_n^{(p)} - t^{(p)}| \xrightarrow{\mathrm{P}} 0.$$

Proof. By **H2** and the Lebesgue-Besicovitch theorem (Evans and Gariepy [27], Theorem 1, Chapter I), we have for a.a. $p \in \mathcal{P}(\Theta_s)$:

$$\frac{1}{\varepsilon}\int_{t^{(p)}-\varepsilon}^{t^{(p)}}\int_{\partial\mathcal{L}(s)} \frac{f}{\|\nabla f\|} d\mathcal{H}\, ds \to \int_{\partial\mathcal{L}(t^{(p)})} \frac{f}{\|\nabla f\|} d\mathcal{H},$$

as $\varepsilon \searrow 0$. Thus, one only needs to prove the lemma for $p \in \mathcal{P}(\Theta_s)$ such that the above result holds. For convenience, let $t = t^{(p)}$ and $t_n = t_n^{(p)}$. It suffices to show that

$$\frac{\sqrt{nh^k}}{\log n}\,|t_n^{(p)} - t^{(p)}| \xrightarrow{\mathrm{P}} 0$$

on the event A_n defined by

$$A_n = \Big\{\sup_{\mathcal{L}(t/2)} |f_n - f| \leq r_n\Big\},$$

where $r_n = (\log n)^{3/4}/\sqrt{nh^k}$, because $P(A_n) \to 1$ (see the proof of Theorem 2.1). According to Lemma 4.3, one has a.s. and for n large enough : $\mathcal{L}(t_n) \cup \mathcal{L}_n(t_n) \subset \mathcal{L}(t/2)$ on the event A_n. Then,

$$\begin{aligned}
|\lambda_f(\mathcal{L}(t_n)) - \lambda_{f_n}(\mathcal{L}_n(t_n))| &= \Big|\int_{\mathcal{L}(t_n)} f d\lambda - \int_{\mathcal{L}_n(t_n)} f_n d\lambda\Big| \\
&\leq \int_{\mathcal{L}(t/2)} |f_n - f| d\lambda + \int f\Big|\mathbf{1}_{\mathcal{L}(t_n)} - \mathbf{1}_{\mathcal{L}_n(t_n)}\Big| d\lambda \\
&\leq c\, r_n + c\,\lambda\Big(\mathcal{L}(t_n)\Delta\mathcal{L}_n(t_n)\Big). \qquad (4.5)
\end{aligned}$$

But, on A_n :

$$\lambda\Big(\mathcal{L}(t_n)\Delta\mathcal{L}_n(t_n)\Big) \leq \lambda\Big(\Big\{t_n - r_n \leq f \leq t_n + r_n\Big\}\Big).$$

By **H1**, **H2**, there exists a neighborhood V of t such that

$$\inf_{f^{-1}(V)} \|\nabla f\| > 0,$$

thus, by Lemma 4.3, one has a.s. and for n large enough :

$$\begin{aligned}\lambda\Big(\mathcal{L}(t_n)\Delta\mathcal{L}_n(t_n)\Big) &\leq \sup_{s\in V}\lambda\Big(\Big\{s-r_n \leq f \leq s+r_n\Big\}\Big) \\ &\leq c\,r_n,\end{aligned}$$

where the latter inequality is a consequence of Proposition A1. According to (4.5), one has on A_n and for n large enough :

$$|\lambda_f(\mathcal{L}(t_n)) - \lambda_f(\mathcal{L}(t))| = |\lambda_f(\mathcal{L}(t_n)) - \lambda_{f_n}(\mathcal{L}_n(t_n))| \leq c\,r_n.$$

Observe now that by Proposition A1 and our choice of t, one has a.s. :

$$\frac{\lambda_f(\mathcal{L}(t_n)) - \lambda_f(\mathcal{L}(t))}{t_n - t} \to \int_{\partial\mathcal{L}(t)} \frac{f}{\|\nabla f\|} d\mathcal{H} \neq 0,$$

thus on A_n,

$$|t_n - t| \leq c\,r_n,$$

for n large enough, hence the lemma •

Lemma 4.6. *Assume that* **H1**-**H3** *hold and let* $(\alpha_n)_n$ *be a sequence of positive real numbers. If* $\alpha_n \to 0$*,* $\alpha_n^2 nh^k/(\log n)^2 \to \infty$ *and* $nh^k/(\log n)^2 \to \infty$*, then for a.a.* $p \in \mathcal{P}(\Theta_s)$ *:*

$$\frac{1}{\alpha_n}\lambda\Big(\mathcal{L}_n(t_n^{(p)}) - \mathcal{L}_n(t_n^{(p)} + \alpha_n)\Big) \xrightarrow{\mathrm{P}} \int_{\partial\mathcal{L}(t^{(p)})} \frac{1}{\|\nabla f\|} d\mathcal{H}.$$

Proof. According to Proposition A1 and **H1**, **H2**, one has for a.a. $t \in \Theta_s$:

$$\frac{1}{\varepsilon}\lambda\Big(\mathcal{L}(t) - \mathcal{L}(t+\varepsilon)\Big) = \frac{1}{\varepsilon}\lambda\Big(\Big\{t \leq f \leq t+\varepsilon\Big\}\Big) \to \int_{\partial\mathcal{L}(t^{(p)})} \frac{1}{\|\nabla f\|} d\mathcal{H},$$

as $\varepsilon \searrow 0$. Hence, it suffices to prove the lemma for all $p \in \mathcal{P}(\Theta_s)$ such that the above result holds with $t = t^{(p)}$. For convenience, let $t = t^{(p)}$ and $t_n = t_n^{(p)}$. By Lemma 4.5, one only needs to prove that

$$\frac{1}{\alpha_n}\lambda\Big(\mathcal{L}_n(t_n) - \mathcal{L}_n(t_n+\alpha_n)\Big) = \frac{1}{\alpha_n}\lambda\Big(\Big\{t_n \leq f_n < t_n + \alpha_n\Big\}\Big)$$

converges in probability to

$$\int_{\partial\mathcal{L}(t^{(p)})} \frac{1}{\|\nabla f\|} d\mathcal{H},$$

on the event B_n defined by

$$B_n = \Big\{ \sup_{\mathcal{L}(t/2)} |f_n - f| \leq v_n,\ |t_n - t| \leq v_n \Big\},$$

where $v_n = \log n/\sqrt{nh^k}$, because $P(B_n) \to 1$. But, for n large enough, one has $\mathcal{L}_n(t_n) \cup \mathcal{L}(t) \subset \mathcal{L}(t/2)$ on B_n. Consequently,

$$\frac{1}{\alpha_n}\Big|\lambda\Big(\Big\{t_n \leq f_n < t_n + \alpha_n\Big\}\Big) - \lambda\Big(\Big\{t \leq f \leq t + \alpha_n\Big\}\Big)\Big|$$

$$\leq \frac{1}{\alpha_n}\lambda\Big(\Big\{t - 2v_n \leq f \leq t + 2v_n\Big\}\Big) \leq c\,\frac{v_n}{\alpha_n},$$

and the latter term tends to 0 by assumption on α_n. Finally, the choice of t implies that

$$\frac{1}{\alpha_n}\lambda\Big(\Big\{t \leq f_n \leq t + \alpha_n\Big\}\Big) \to \int_{\partial\mathcal{L}(t^{(p)})} \frac{1}{\|\nabla f\|} d\mathcal{H},$$

so that on B_n :

$$\frac{1}{\alpha_n}\lambda\Big(\Big\{t_n \leq f_n < t_n + \alpha_n\Big\}\Big) \overset{\mathrm{P}}{\to} \int_{\partial\mathcal{L}(t^{(p)})} \frac{1}{\|\nabla f\|} d\mathcal{H},$$

hence the lemma •

Proof of Corollary 2.1. According to Lemma 4.3, Lemma 4.6 and Theorem 2.1, one only needs to prove that for a.a. $p \in \mathcal{P}(\Theta_s)$:

$$\sqrt{nh^k}\Big[\lambda\Big(\mathcal{L}_n(t_n^{(p)})\Delta\mathcal{L}(t^{(p)})\Big) - \lambda\Big(\mathcal{L}_n(t^{(p)})\Delta\mathcal{L}(t^{(p)})\Big)\Big] \overset{\mathrm{P}}{\to} 0.$$

Moreover, it suffices to show the above result for each $p \in \mathcal{P}(\Theta_s)$ such that the conclusion of Lemma 4.4 holds. Fix such a $p \in \mathcal{P}(\Theta_s)$ and, for simplicity, let $t = t^{(p)}$ and $t_n = t_n^{(p)}$. A straightforward computation gives the relation :

$$D_n := \lambda\Big(\mathcal{L}_n(t_n)\Delta\mathcal{L}(t)\Big) - \lambda\Big(\mathcal{L}_n(t)\Delta\mathcal{L}(t)\Big) = \int \Big(\mathbf{1}_{\{f_n \geq t_n\}} - \mathbf{1}_{\{f_n \geq t\}}\Big)\eta\, d\lambda,$$

where $\eta = 1 - 2\mathbf{1}_{\{f \geq t\}}$. Then,

$$D_n = \int \mathbf{1}_{\{t_n \leq f_n < t\}}\eta\, d\lambda - \int \mathbf{1}_{\{t \leq f_n < t_n\}}\eta\, d\lambda.$$

By (4.4) and **H3**, one can now apply Proposition A1, which gives :

$$D_n = \int_{t_n}^{t} \int_{\partial \mathcal{L}_n(s)} \frac{\eta}{\|\nabla f_n\|} d\mathcal{H}\, ds.$$

Consequently,

$$|D_n| \le \int_{\min(t_n,t)}^{\max(t_n,t)} \int_{\partial \mathcal{L}_n(s)} \frac{1}{\|\nabla f_n\|} d\mathcal{H}\, ds,$$

so that by Lemma 4.4 :

$$\sqrt{nh^k} D_n = \sqrt{nh^k} \Big[\lambda\Big(\mathcal{L}_n(t_n) \Delta \mathcal{L}(t)\Big) - \lambda\Big(\mathcal{L}_n(t) \Delta \mathcal{L}(t)\Big) \Big] \xrightarrow{\mathrm{P}} 0,$$

hence the corollary •

Appendix A : Change of variable formula and an application. Proposition A1 below is a consequence of the change of variables formula given in Evans and Gariepy ([27], Chapter III, Theorem 2). For a similar proof, see also Chapter III, Proposition 3 in the same book.

Proposition A1. *Let $\varphi : \mathbb{R}^k \to \mathbb{R}_+$ be a continuously differentiable function such that $\varphi(x) \to 0$ as $\|x\| \to \infty$, and $J \subset \mathbb{R}_+$ be an interval such that* $\inf J > 0$ *and*

$$\inf_{\varphi^{-1}(J)} \|\nabla \varphi\| > 0.$$

Then, for all bounded Borel function $g : \mathbb{R}^k \to \mathbb{R}$:

$$\int_{\varphi^{-1}(J)} g d\lambda = \int_J \int_{\varphi^{-1}(\{s\})} \frac{g}{\|\nabla \varphi\|} d\mathcal{H}\, ds.$$

Proof. Notice that φ is a locally Lipschitz function and

$$g\mathbf{1}_{\varphi^{-1}(J)}$$

is integrable because $\varphi^{-1}(J)$ is bounded. Proposition A1 is then an easy consequence of Theorem 2 in Evans and Gariepy ([27], Chapter III) •

Proposition A2. *Assume that* **H1, H2** *hold. Then, for all $t \in \Theta$,*

$$\lambda(f^{-1}[t - \varepsilon, t + \varepsilon]) \to 0, \ \textit{as}\ \varepsilon \to 0.$$

Proof. One can find a closed interval F such that $t \in F \subset \Theta$. Since $0 \notin F$, $\lambda(f^{-1}(F)) < \infty$. By Proposition A1, we then deduce that the map

$$s \mapsto \int_{f^{-1}(\{s\})} \frac{1}{\|\nabla f\|} d\mathcal{H}$$

defined on F is integrable. Proposition A2 is then an easy consequence of Proposition A1 and the Lebesgue theorem •

Appendix B : On the existence of $t_n^{(p)}$. Assume that the hypotheses of Corollary 2.1 hold. Let $t \in \Theta$. Then, $\lambda(f^{-1}(\{t\})) = 0$ by Proposition A2, and hence $\lambda_f(\partial\mathcal{L}(t)) = 0$. Since, as seen in the proof of Theorem 2.1, $\sup_{I\!R^k} |f_n - f| \to 0$ a.s., we deduce that a.s. :

$$\lambda_{f_n}(\mathcal{L}_n(t)) \to \lambda_f(\mathcal{L}(t))$$

(use for instance Theorem 5.1 in Billingsley [25], with the Lebesgue measure restricted to some compact set instead of a probability measure). Observe that this convergence holds uniformly over the compact sets of Θ since the map $s \mapsto \lambda_{f_n}(\mathcal{L}_n(s))$ is non increasing. Let now $I \subset \Theta$ be a closed interval such that

$$\inf I < \inf \Theta_s \text{ and } \sup I > \sup \Theta_s.$$

Using again the monotonicity argument, we deduce that

$$P\Big(\lambda_{f_n}(\mathcal{L}_n(\inf I)) \geq \lambda_f(\mathcal{L}(\inf \Theta_s)), \lambda_{f_n}(\mathcal{L}_n(\sup I)) \leq \lambda_f(\mathcal{L}(\sup \Theta_s))\Big) \to 1,$$

and hence, since a.s., $(\lambda_{f_n}(\mathcal{L}_n(s)))_{s\in I}$ uniformly converges to $(\lambda_f(\mathcal{L}(s)))_{s\in I}$, that

$$P\Big(\exists t \in I : \ \lambda_{f_n}(\mathcal{L}_n(t)) = p\Big) \to 1,$$

for all $p \in \mathcal{P}(\Theta_s)$. As a conclusion, for all $p \in \mathcal{P}(\Theta_s)$, with probability tending to 1, there exists $t_n^{(p)} \in I$ such that $\lambda_{f_n}(\mathcal{L}_n(t_n^{(p)})) = p$ •

Acknowledgements. The author thank André Mas and Nicolas Molinari for many helpful comments, and the Associate Editor and the Referee for their careful reading of the paper.

REFERENCES

[1] J.A. Hartigan, Clustering Algorithms, Wiley, New-York, 1975.

[2] A. Cuevas, M. Febrero and R. Fraiman, Estimating the Number of Clusters, Can. J. Stat. 28 (2000) 367-382.

[3] A. Cuevas, M. Febrero and R. Fraiman, Cluster analysis : a Further Approach Based on Density Estimation, Comput. Stat. Data Anal. 36 (2001) 441-459.

[4] L. Devroye and G.L. Wise, Detection of Abnormal Behavior via Nonparametric Estimation of the Support, SIAM J. Appl. Math. 38 (1980) 480-488.

[5] U. Grenander, Abstract Inference, Wiley, New-York, 1981.

[6] A. Cuevas, On Pattern Analysis in the Non-convex Case, Kybernetes 19 (1990) 26-33.

[7] A. Cuevas and R. Fraiman, A Plug-in Approach to Support Estimation, Ann. Statist. 25 (1997) 2300-2312.

[8] A. Baíllo, A. Cuevas and A. Justel, Set Estimation and Nonparametric Detection, Can. J. Stat. 28 (2000) 765-782.

[9] A. Cuevas and A. Rodríguez-Casal, Set Estimation : an Overview and some Recent Developpements. In Recent Advances and Trends in Nonparametric Statistics, Elsevier, North Holland, 2003.

[10] A. Baíllo, J.A. Cuesta-Albertos and A. Cuevas, Convergence Rates in Nonparametric Estimation of Level Sets, Stat. Probab. Lett. 53 (2001) 27-35.

[11] A. Baíllo, Total Error in a Plug-in Estimator of Level Sets, Stat. Probab. Lett. 65 (2003) 411-417.

[12] I.S. Molchanov, Empirical Estimation of Distribution Quantiles of Random Closed Sets, Theory Probab. Appl. 35 (1990) 594-600.

[13] I.S. Molchanov, A Limit Theorem for Solutions of Inequalities, Scand. J. Stat. 25 (1998) 235-242.

[14] J.A. Hartigan, Estimation of a Convex Density Contour in Two Dimensions, J. Amer. Statist. Assoc. 82 (1987) 267-270.

[15] D.W. Müller, The Excess Mass Approach in Statistics, Beiträge zur Statistik, Univ. Heidelberg, 1993.

[16] D.W. Müller and G. Sawitzki, Excess Mass Estimates and Tests of Multimodality, J. Amer. Statist. Assoc. 86 (1991) 738-746.

[17] D. Nolan, The Excess-mass Ellipsoid, J. Multivariate Anal. 39 (1991) 348-371.

[18] W. Polonik, Measuring Mass Concentration and Estimating Density Contour Clusters - an Excess Mass Approach, Ann. Statist. 23 (1995) 855-881.

[19] A.B. Tsybakov, On Nonparametric Estimation of Density Level Sets, Ann. Statist. 25 (1997) 948-969.

[20] A. Cuevas and A. Rodríguez-Casal, On Boundary Estimation, Adv. Appl. Probab. 36 (2004) 340-354.

[21] G. Walther, Granulometric Smoothing, Ann. Statist. 25 (1997) 2273-2299.

[22] H. Bräker, T. Hsing and N.H. Bingham, On the Hausdorff Distance Between a Convex Set and an Interior Random Convex Hull, Adv. Appl. Probab. 30 (1998) 295-316.

[23] L. Dümbgen and G. Walther, Rates of Convergence for Random Approximations of Convex Sets, Adv. Appl. Probab. 28 (1996) 384-393.

[24] M. Rosenblatt, Remarks on some Nonparametric Estimates of a Density Function, Ann. Math. Statist. 27 (1956) 832-837.

[25] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.

[26] D. Pollard, Convergence of Stochastic Processes, Springer, New York, 1984.

[27] L.C. Evans and R.F. Gariepy, Measure Theory and Fine Properties of Functions, CRC Press, Boca Raton, 1992.

[28] B.L.S Prakasa Rao, Nonparametric Functional Estimation, Academic Press, Orlando, 1983.

[29] W. Feller, An Introduction to Probability Theory and its Applications, Wiley, New-York, 1992.