SUPPLEMENT TO “FULLY ADAPTIVE DENSITY-BASED CLUSTERING”

By Ingo Steinwart

University of Stuttgart

In this supplement several auxiliary results, which are partially taken from [9], are presented and the assumptions made in the paper are discussed in more detail. This material is contained in the sections A.1 to A.10. In addition, we present a couple of two-dimensional examples that show that the assumptions imposed in the paper are not only met by many discontinuous densities, but also by many continuous densities. This material is contained in the sections B.1 and B.2.

Appendix A. Remaining Proofs and Additional Material. In this appendix, the auxiliary results from [9] are presented and the assumptions are discussed in more detail than it was possible in the main paper.

A.1. Material Related to Level Sets. In this section we present some additional results from [9] related to the definition of M_ρ.

To begin with, we note that using the definition of the support of a measure it becomes obvious that M_ρ can be expressed by

$$M_\rho = \bigl\{x \in X : \mu_\rho(U) > 0 \text{ for all open neighborhoods } U \text{ of } x\bigr\}\,. \tag{A.1.1}$$

Furthermore, if $\operatorname{supp}\mu = X$, we actually have $M_\rho = X$ for all $\rho \le 0$, but typically we are, of course, interested in the case $\rho > 0$, only. The next lemma shows that the sets M_ρ are ordered in the usual way.

Lemma A.1.1. *Let (X, d) be a complete separable metric space, μ be a σ-finite measure on X, and P be a μ-absolutely continuous distribution on X. Then, for all $\rho_1 \le \rho_2$, we have*

$$M_{\rho_2} \subset M_{\rho_1}\,.$$

Proof of Lemma A.1.1. We fix an $x \in M_{\rho_2}$ and an open set $U \subset X$ with $x \in U$. Moreover, we fix a μ-density h of P. Then we obtain

$$\mu_{\rho_1}(U) = \mu\bigl(\{h \ge \rho_1\} \cap U\bigr) \ge \mu\bigl(\{h \ge \rho_2\} \cap U\bigr) = \mu_{\rho_2}(U) > 0\,,$$

and hence we obtain $x \in M_{\rho_1}$ by (A.1.1). $\square$

The following lemma describes the relationship between M_ρ and $\{h \ge \rho\}$.

Lemma A.1.2. *Let (X,d) be a complete separable metric space, μ be a σ-finite measure on X with $\operatorname{supp}\mu = X$, and P be a μ-absolutely continuous distribution on X. Then, for all μ-densities h of P and all $\rho \in \mathbb{R}$, we have*

$$\{h \overset{\circ}{\geq} \rho\} \subset M_\rho \subset \overline{\{h \geq \rho\}}\,.$$

If h is continuous, we even have $\{h>\rho\} \subset M_\rho \subset \{h\geq\rho\}$ and $\partial M_\rho \subset \{h=\rho\}$.

Proof of Lemma A.1.2. By definition, M_ρ is the smallest closed set A satisfying $\mu(\{h \geq \rho\} \setminus A) = 0$. Moreover, we obviously have

$$\mu\bigl(\{h \geq \rho\} \setminus \overline{\{h \geq \rho\}}\bigr) = 0\,,$$

and hence we obtain $M_\rho \subset \overline{\{h \geq \rho\}}$. To show the other inclusion, we fix an $x \in \{h \overset{\circ}{\geq} \rho\}$ and an open set $U \subset X$ with $x \in U$. Then $\{h \overset{\circ}{\geq} \rho\} \cap U$ is open and non-empty, and hence $\operatorname{supp}\mu = X$ yields

$$\mu_\rho(U) = \mu\bigl(\{h \geq \rho\} \cap U\bigr) \geq \mu\bigl(\{h \overset{\circ}{\geq} \rho\} \cap U\bigr) > 0\,.$$

By (A.1.1) we conclude that $x \in M_\rho$, that is, we have shown $\{h \overset{\circ}{\geq} \rho\} \subset M_\rho$.

Now assume that h is continuous. Clearly, we have $\{h > \rho\} \subset \{h \geq \rho\}$ and since $\{h > \rho\}$ is open, we conclude that $\{h > \rho\} \subset \{h \overset{\circ}{\geq} \rho\} \subset M_\rho$ by the previously shown inclusion. Moreover, since $\{h \geq \rho\}$ is closed, we find $M_\rho \subset \overline{\{h \geq \rho\}} = \{h \geq \rho\}$. Recalling that M_ρ is closed by definition, we further find $\partial M_\rho \subset M_\rho \subset \{h \geq \rho\}$, and thus it remains to show $\partial M_\rho \subset \{h \leq \rho\}$. Let us assume the converse, i.e., that there exists an $x \in \partial M_\rho$ such that $h(x) > \rho$. By the continuity we then find an open neighborhood U of x with $U \subset \{h > \rho\}$. Since $x \in \partial M_\rho$, we further find an $y \in U \setminus M_\rho$, while our construction together with the previously shown $\{h > \rho\} \subset M_\rho$ yields the contradicting statement $U \setminus M_\rho \subset \{h > \rho\} \setminus M_\rho = \emptyset$. □

The next lemma provides some simple sufficient conditions for normality.

Lemma A.1.3. *Let (X,d) be a complete separable metric space, μ be a σ-finite measure on X with $\operatorname{supp}\mu = X$, and P be a μ-absolutely continuous distribution on X. Then the following statements hold:*

- *i) If P has an upper semi-continuous μ-density, then it is upper normal at every level.*
- *ii) If P has a lower semi-continuous μ-density, then it is lower normal at every level.*
- *iii) If, for some $\rho \geq 0$, P has a μ-density h such that $\mu(\partial\{h \geq \rho\}) = 0$, then P is normal at level ρ.*

Proof of Lemma A.1.3. *i).* Let us fix an upper semi-continuous μ-density h of P. Then $\{h \geq \rho\}$ is closed, and hence Lemma A.1.2 shows $M_\rho \subset \overline{\{h \geq \rho\}} = \{h \geq \rho\}$. Thus, P is upper normal at level ρ.

ii). Let h be a lower semi-continuous μ-density of P. By Lemma A.1.2 we then know $\{h > \rho\} = \{h \overset{\circ}{>} \rho\} \subset \{h \overset{\circ}{\geq} \rho\} \subset \mathring{M}_\rho$. This yields the assertion.

iii). The upper normality follows from (2.3). To see that P is lower normal, we use the inclusion $\{h > \rho\} \setminus \mathring{M}_\rho \subset \overline{\{h \geq \rho\}} \setminus \{h \overset{\circ}{\geq} \rho\} = \partial\{h \geq \rho\}$ which follows from Lemma A.1.2. $\square$

Let us now assume that P is upper normal at some level ρ. By (2.2) we then immediately see that

$$\mu(M_\rho \vartriangle \{h \geq \rho\}) = 0 \tag{A.1.2}$$

for *all* μ-densities h of P. In other words, up to μ-zero measures, M_ρ equals the ρ-level set of *all* μ-densities h of P. Moreover, if for some $\rho^* > 0$ and $\rho^{**} > \rho^*$, the distribution P is upper normal at every level $\rho \in (\rho^*, \rho^{**}]$, then using the monotonicity of the sets M_ρ and $\{h \geq \rho\}$ in ρ as well as $(\cup_{i\in I} A_i) \vartriangle (\cup_{i\in I} B_i) \subset \cup_{i\in I}(A_i \vartriangle B_i)$, we find

$$\mu\biggl(\{h > \rho^*\} \vartriangle \bigcup_{\rho > \rho^*} M_\rho\biggr) \leq \mu\biggl(\bigcup_{n\in\mathbb{N}} \bigl(\{h \geq \rho^* + 1/n\} \vartriangle M_{\rho^*+1/n}\bigr)\biggr) = 0 \tag{A.1.3}$$

for all μ-densities h of P, and if P has a continuous density h, we even have $\bigcup_{\rho>\rho^*} M_\rho = \{h > \rho^*\}$ by an easy consequence of Lemma A.1.2. Similarly, if P is lower normal at every level $\rho \in (\rho^*, \rho^{**}]$, we find

$$\mu\biggl(\{h > \rho^*\} \setminus \bigcup_{\rho > \rho^*} \mathring{M}_\rho\biggr) \leq \mu\biggl(\bigcup_{n\in\mathbb{N}} \bigl(\{h > \rho^* + 1/n\} \setminus \mathring{M}_{\rho^*+1/n}\bigr)\biggr) = 0\,, \tag{A.1.4}$$

and if in addition, (A.1.3) holds, we obtain $\mu(\bigcup_{\rho>\rho^*} M_\rho \vartriangle \bigcup_{\rho>\rho^*} \mathring{M}_\rho) = 0$.

A.2. Proofs and Material on Connected Components. This section contains the proofs related to Subsection 2.2. In addition, we recall several additional results on connected components from [9].

Lemma A.2.1. *Let $A \subset B$ be two non-empty sets with partitions $\mathcal{P}(A)$ and $\mathcal{P}(B)$, respectively. Then the following statements are equivalent:*

i) $\mathcal{P}(A)$ is comparable to $\mathcal{P}(B)$.

ii) There exists a $\zeta : \mathcal{P}(A) \to \mathcal{P}(B)$ such that, for all $A' \in \mathcal{P}(A)$, we have

$$A' \subset \zeta(A')\,. \tag{A.2.1}$$

Moreover, if one these statements are true, the map ζ is uniquely determined by (A.2.1). We call ζ the cell relating map (CRM) between A and B.

Proof of Lemma A.2.1. *ii)* $\Rightarrow$ *i)*. Trivial.

i) $\Rightarrow$ *ii)*. For $A' \in \mathcal{P}(A)$ we find a $B' \in \mathcal{P}(B)$ such that $A' \subset B'$. Defining $\zeta(A') := B'$ then gives the desired Property (A.2.1).

Finally, assume that *ii)* is true but ζ is not unique. Then there exist $A' \in \mathcal{P}(A)$ and $B', B'' \in \mathcal{P}(B)$ with $B' \neq B''$ and both $A' \subset B'$ and $A' \subset B''$. Since $A' \neq \emptyset$, this yields $B' \cap B'' \neq \emptyset$, which in turn implies $B' = B''$ as $\mathcal{P}(B)$ is a partition, i.e. we have found a contradiction. $\square$

Proof of Lemma 2.4. Clearly, $\zeta := \zeta_{B,C} \circ \zeta_{A,B}$ maps from $\mathcal{P}(A)$ to $\mathcal{P}(C)$. Moreover, for $A' \in \mathcal{P}(A)$ we have $A' \subset \zeta_{A,B}(A')$ and for $B' := \zeta_{A,B}(A') \in \mathcal{P}(B)$ we have $B' \subset \zeta_{B,C}(B')$. Combining these inclusions we find

$$A' \subset \zeta_{A,B}(A') \subset \zeta_{B,C}(\zeta_{A,B}(A')) = \zeta_{B,C} \circ \zeta_{A,B}(A') = \zeta(A')$$

for all $A' \in \mathcal{P}(A)$. Consequently, $\mathcal{P}(A)$ is comparable to $\mathcal{P}(C)$ and by Lemma A.2.1 we see that ζ is the CRM $\zeta_{A,C}$, that is $\zeta_{A,C} = \zeta = \zeta_{B,C} \circ \zeta_{A,B}$. $\square$

Lemma A.2.2. *Let (X, d) be a metric space, $A \subset X$ be a non-empty subset and $\tau > 0$. Then every τ-connected component of A is τ-connected.*

Proof of Lemma A.2.2. Let A' be a τ-connected component of A and $x, x' \in A'$. Then x and x' are τ-connected in A, and hence there exist $x_1, \dots, x_n \in A$ such that $x_1 = x$, $x_n = x'$ and $d(x_i, x_{i+1}) < \tau$ for all $i = 1, \dots, n-1$. Now, $d(x_1, x_2) < \tau$ shows that x_1 and x_2 are τ-connected in A, and hence they belong to the same τ-connected component, i.e. we have found $x_2 \in A'$. Iterating this argument, we find $x_i \in A'$ for all $i = 1, \dots, n$. Consequently, x and x' are not only τ-connected in A, but also τ-connected in A'. This shows that A' is τ-connected. $\square$

Lemma A.2.3. *Let (X, d) be a metric space and $A \subset B$ be two closed non-empty subsets of X with $|\mathcal{C}(B)| < \infty$. Then $\mathcal{C}(A)$ is comparable to $\mathcal{C}(B)$.*

Proof of Lemma A.2.3. Let us fix an $A' \in \mathcal{C}(A)$. Since $A \subset B$ and $|\mathcal{C}(B)| < \infty$ there then exist an $m \geq 1$ and mutually distinct $B_1, \dots, B_m \in \mathcal{C}(B)$ with $A' \subset B_1 \cup \dots \cup B_m$ and $A' \cap B_i \neq \emptyset$ for all $i = 1, \dots, m$. Since A and B are closed, A' and the sets $A' \cap B_i$ are also closed. Consequently, the sets $A' \cap B_i$ are also closed in A' with respect to the relative topology of A'. Let us now assume that $m > 1$. Then $A' \cap B_1$ and $(A' \cap B_2) \cup \dots \cup (A' \cap B_m)$ are two disjoint relatively closed non-empty subsets of A' whose union equals

A'. Consequently A' is not connected, which contradicts $A' \in \mathcal{C}(A)$. In other words, we have $m = 1$, that is, $\mathcal{C}(A)$ is comparable to $\mathcal{C}(B)$. $\square$

LEMMA A.2.4. *Let (X, d) be a metric space, $A \subset X$ be non-empty and $\tau > 0$. Then we have $d(A', A'') \geq \tau$ for all $A', A'' \in \mathcal{C}_\tau(A)$ with $A' \neq A''$. Moreover, if A is closed, all $A' \in \mathcal{C}_\tau(A)$ are closed, and if X is compact we have $|\mathcal{C}_\tau(A)| < \infty$.*

PROOF OF LEMMA A.2.4. Let $A' \neq A''$ be two τ-connected components of A. Then we have $d(x', x'') \geq \tau$ for all $x' \in A'$ and $x'' \in A''$, since otherwise x' and x'' would be τ-connected in A. Thus, we have $d(A', A'') \geq \tau$, and from the latter and the compactness of X, we conclude that $|\mathcal{C}_\tau(A)| < \infty$. Finally, let $(x_i) \subset A'$ be a sequence in some component $A' \in \mathcal{C}_\tau(A)$ such that $x_i \to x$ for some $x \in X$. Since A is closed, we have $x \in A$, and hence $x \in A''$ for some $A'' \in \mathcal{C}_\tau(A)$. By construction we find $d(A', A'') = 0$, and hence we obtain $A' = A''$ by the assertion that has been shown first. $\square$

LEMMA A.2.5. *Let (X, d) be a metric space, $A \subset X$ be a non-empty subset and $\tau > 0$. Then the following statements are equivalent:*

i) A is τ-connected.
ii) For all non-empty subsets A^+ and A^- of A with $A^+ \cup A^- = A$ and $A^+ \cap A^- = \emptyset$ we have $d(A^+, A^-) < \tau$.

PROOF OF LEMMA A.2.5. *i) $\Rightarrow$ ii).* We fix two subsets A^+ and A^- of A with $A^+ \cup A^- = A$ and $A^+ \cap A^- = \emptyset$. Let us further fix two points $x^+ \in A^+$ and $x^- \in A^-$. Since A is τ-connected, there then exist $x_1, \ldots, x_n \in A$ such that $x_1 = x^-$, $x_n = x^+$ and $d(x_i, x_{i+1}) < \tau$ for all $i = 1, \ldots, n-1$. Then, $x^+ \in A^+$ and $x^- \in A^-$ imply the existence of an $i \in \{1, \ldots, n-1\}$ with $x_i \in A^-$ and $x_{i+1} \in A^+$. This yields $d(A^+, A^-) \leq d(x_i, x_{i+1}) < \tau$.

ii) $\Rightarrow$ i). Assume that A is not τ-connected, that is $|\mathcal{C}_\tau(A)| > 1$. We pick an $A^+ \in \mathcal{C}_\tau(A)$ and write $A^- := A \setminus A^+$. Since $|\mathcal{C}_\tau(A)| > 1$, both sets are non-empty, and our construction ensures that they are also disjoint and satisfy $A^+ \cup A^- = A$. Moreover, for every $A' \in \mathcal{C}_\tau(A)$ with $A' \neq A^+$ we know $d(A^+, A') \geq \tau$ by Lemma A.2.4 and since A^- is the union of such A', we conclude $d(A^+, A^-) \geq \tau$. $\square$

COROLLARY A.2.6. *Let (X, d) be a metric space, $A \subset B \subset X$ be non-empty subsets and $\tau > 0$. If A is τ-connected, then there exists exactly one τ-connected component B' of B with $A \cap B' \neq \emptyset$. Moreover, B' is the only τ-connected component B'' of B that satisfies $A \subset B''$.*

Proof of Corollary A.2.6. The second assertion is a direct consequence of the first, and hence it suffice to show the first assertion. Let us assume the first is not true. Since $A \subset B$ there then exist $B', B'' \in \mathcal{C}_\tau(B)$ with $B' \neq B''$, $A \cap B' \neq \emptyset$, and $A \cap B'' \neq \emptyset$. We write $A^- := A \cap B'$ and $A^+ := A \cap (B \setminus B')$. Since $B'' \subset B \setminus B'$, we obtain $A^+ \neq \emptyset$, and therefore, Lemma A.2.5 shows $d(A^-, A^+) < \tau$. Consequently, there exist $x^- \in A^-$ and $x^+ \in A^+$ with $d(x^+, x^-) < \tau$. Now we obviously have $x^- \in B'$, and by construction, we also find a $B''' \in \mathcal{C}_\tau(B)$ with $x^+ \in B'''$. Our previous inequality then yields $d(B', B''') < \tau$, while Lemma A.2.4 shows $d(B', B''') \geq \tau$, that is, we have found a contradiction. □

Lemma A.2.7. *Let (X, d) be a metric space, $A \subset B$ be two non-empty subsets of X and $\tau > 0$. Then $\mathcal{C}_\tau(A)$ is comparable to $\mathcal{C}_\tau(B)$.*

Proof of Lemma A.2.7. For $A' \in \mathcal{C}_\tau(A)$, Corollary A.2.6 shows that there is exactly $B' \in \mathcal{C}_\tau(B)$ with $A' \subset B'$. Thus, $\mathcal{C}_\tau(A)$ is comparable to $\mathcal{C}_\tau(B)$. □

Lemma A.2.8. *Let (X, d) be a metric space, $A \subset X$ be a non-empty subset and $\tau > 0$. Then, for a partition $A_1, \ldots, A_m$ of A, the following statements are equivalent:*

i) $\mathcal{C}_\tau(A) = \{A_1, \ldots, A_m\}$.
ii) *A_i is τ-connected for all $i = 1, \ldots, m$, and $d(A_i, A_j) \geq \tau$ for all $i \neq j$.*

Proof of Lemma A.2.8. *i)* $\Rightarrow$ *ii)*. Follows from Lemma A.2.4.

ii) $\Rightarrow$ *i)*. Let us fix an $A' \in \mathcal{C}_\tau(A)$ and an A_i with $A_i \cap A' \neq \emptyset$. Since A_i is τ-connected and $A' \in \mathcal{C}_\tau(A)$, Corollary A.2.6 applied to the sets $A_i \subset A \subset X$ yields $A_i \subset A'$. Moreover, $A_1, \ldots, A_m$ is a partition of A, and thus we conclude that

$$A' = \bigcup_{i \in I} A_i \,,$$

where $I := \{i : A_i \cap A' \neq \emptyset\}$. Now let us assume that $|I| \geq 2$. We fix an $i_0 \in I$ and write $A^+ := A_{i_0}$ and $A^- := \bigcup_{i \in I \setminus \{i_0\}} A_i$. Since $|I| \geq 2$, we obtain $A^- \neq \emptyset$, and Lemma A.2.5 thus shows $d(A^+, A^-) < \tau$. On the other hand, our assumption ensures $d(A^+, A^-) \geq \tau$, and hence $|I| \geq 2$ cannot be true. Consequently, there exists a unique index i with $A' = A_i$. □

Lemma A.2.9. *Let (X, d) be a compact metric space and $A \subset X$ be a non-empty closed subset. Then the following statements are equivalent:*

i) A is connected.

ii) A is τ-connected for all $\tau > 0$.

Proof of Lemma A.2.9. *i) $\Rightarrow$ ii).* Assume that A is not τ-connected for some $\tau > 0$. Then, by Lemma A.2.4, there are finitely many τ-connected components $A_1, \dots, A_m$ of A with $m > 1$. We write $A' := A_1$ and $A'' := A_2 \cup \dots \cup A_m$. Then A' and A'' are non-empty, disjoint and $A' \cup A'' = A$ by construction. Moreover, Lemma A.2.4 shows that A' and A'' are closed since A is closed, and hence A cannot be connected.

ii) $\Rightarrow$ i). Let us assume that A is not connected. Then there exist two non-empty closed disjoint subsets of A with $A' \cup A'' = A$. Since X is compact, A' and A'' are also compact, and hence $A' \cap A'' = \emptyset$ implies $\tau := d(A', A'') > 0$. Lemma A.2.5 then shows that A is not τ-connected. $\square$

The next proposition investigates the relation between $\mathcal{C}_\tau(A)$ and $\mathcal{C}(A)$.

Proposition A.2.10. *Let (X, d) be a compact metric space and $A \subset X$ be a non-empty closed subset. Then the following statements hold:*

i) For all $\tau > 0$, $\mathcal{C}(A)$ is comparable to $\mathcal{C}_\tau(A)$ and the CRM $\zeta : \mathcal{C}(A) \to \mathcal{C}_\tau(A)$ is surjective.

ii) If $|\mathcal{C}(A)| < \infty$, we have

$$\tau_A^* := \min\{d(A', A'') : A', A'' \in \mathcal{C}(A) \text{ with } A' \neq A''\} > 0\,,$$

where $\min \emptyset := \infty$. Moreover, for all $\tau \in (0, \tau_A^*] \cap (0, \infty)$, we have $\mathcal{C}(A) = \mathcal{C}_\tau(A)$ and, for such τ, the CRM $\zeta : \mathcal{C}(A) \to \mathcal{C}_\tau(A)$ is bijective. Finally, if $\tau_A^* < \infty$, that is, $|\mathcal{C}(A)| > 1$, we have

$$\tau_A^* = \max\{\tau > 0 : \mathcal{C}(A) = \mathcal{C}_\tau(A)\}\,.$$

Note that, in general, a closed subset of A may have infinitely many topologically connected components as, e.g., the Cantor set shows. In this case, the second assertion of the lemma above is, in general, no longer true.

Proof of Proposition A.2.10. *i).* Let $A' \in \mathcal{C}(A)$ and $\tau > 0$. Since A is closed, so is A', and hence A' is τ-connected by Lemma A.2.9. Consequently, Corollary A.2.6 shows that there exists an $A'' \in \mathcal{C}_\tau(A)$ with $A' \subset A''$, i.e. $\mathcal{C}(A)$ is comparable to $\mathcal{C}_\tau(A)$. Now we fix an $A'' \in \mathcal{C}_\tau(A)$. Then there exists an $x \in A''$, and to this x, there exists an $A' \in \mathcal{C}(A)$ with $x \in A'$. This yields $A' \cap A'' \neq \emptyset$, and since A' is τ-connected by Lemma A.2.9, Corollary A.2.6 shows $A' \subset A''$, i.e. we obtain $\zeta(A') = A''$.

ii). Let $A_1, \dots, A_m$ be the topologically connected components of A. Then the components are closed, and since A is a closed and thus compact subset of

X, the components are compact, too. This shows $d(A_i, A_j) > 0$ for all $i \neq j$, and consequently we obtain $\tau_A^* > 0$. Let us fix a $\tau \in (0, \tau_A^*] \cap (0, \infty)$. Then, Lemma A.2.9 shows that each A_i is τ-connected, and therefore Lemma A.2.8 together with $d(A_i, A_j) \geq \tau_A^* \geq \tau$ for all $i \neq j$ yields $\mathcal{C}_\tau(A) = \{A_1, \ldots, A_m\}$. Consequently, we have proved $\mathcal{C}(A) = \mathcal{C}_\tau(A)$. The bijectivity of ζ now follows from its surjectivity. For the proof of the last equation, we define $\tau^* := \sup\{\tau > 0 : \mathcal{C}(A) = \mathcal{C}_\tau(A)\}$. Then we have already seen that $\tau_A^* \leq \tau^*$. Now suppose that $\tau_A^* < \tau^*$. Then there exists a $\tau \in (\tau_A^*, \tau^*)$ with $\mathcal{C}(A) = \mathcal{C}_\tau(A)$. On the one hand, we then find $d(A_i, A_j) \geq \tau$ for all $i \neq j$ by Lemma A.2.4, while on the other hand $\tau > \tau_A^*$ shows that there exist $i_0 \neq j_0$ with $d(A_{i_0}, A_{j_0}) < \tau$. In other words, the assumption $\tau_A^* < \tau^*$ leads to a contradiction, and hence we have $\tau_A^* = \tau^*$. □

The last lemma in this subsection shows the monotonicity of τ_A^*.

Lemma A.2.11. *Let (X, d) be a compact metric space and $A \subset B$ be two non-empty closed subsets of X with $|\mathcal{C}(A)| < \infty$ and $|\mathcal{C}(B)| < \infty$. If the CRM $\zeta : \mathcal{C}(A) \to \mathcal{C}(B)$ is injective, then we have $\tau_A^* \geq \tau_B^*$.*

Proof of Lemma A.2.11. Let us fix some $A', A'' \in \mathcal{C}(A)$ with $A' \neq A''$. Since ζ is injective, we then obtain $\zeta(A') \neq \zeta(A'')$. Combining this with $A' \subset \zeta(A')$ and $A'' \subset \zeta(A'')$, we find

$$d(A', A'') \geq d(\zeta(A'), \zeta(A'')) \geq \tau_B^*,$$

where the last inequality follows from the definition of τ_B^*. Taking the infimum over all A' and A'' with $A' \neq A''$ yields the assertion. □

A.3. Additional Material Related to Tubes around Sets. This section contains additional material on the operations $A^{+\delta}$ and $A^{-\delta}$.

Let us begin by noting that in the literature there is another, closely related concept for adding and cutting off δ-tubes, which is based on the Minkowski addition. Namely, in generic metric spaces (X, d), we can define

$$A^{\oplus\delta} := \{x \in X : \exists y \in A \text{ with } d(x, y) \leq \delta\}$$
$$A^{\ominus\delta} := \{x \in X : B(x, \delta) \subset A\}$$

for $A \subset X$ and $\delta > 0$, where $B(x, \delta) := \{y \in X : d(x, y) \leq \delta\}$ denotes the closed ball with radius δ and center x. Some simple considerations then show $A^{\ominus(\delta+\epsilon)} \subset A^{-\delta} \subset A^{\ominus\delta}$ and $A^{\oplus\delta} \subset A^{+\delta} \subset A^{\oplus(\delta+\epsilon)}$ for all $\epsilon, \delta > 0$, that is, the operations of both concepts almost coincide. In addition, it is straightforward to check that $A^{\ominus\delta} = X \setminus (X \setminus A)^{\oplus\delta}$.

Usually, the operations $\oplus\delta$ and $\ominus\delta$ are considered for the special case $X := \mathbb{R}^d$ equipped with the Euclidean norm. In this case, we immediately obtain the more common expressions

$$A^{\oplus\delta} = \{x + y : x \in A \text{ and } y \in \delta B_{\ell_2^d}\}$$
$$A^{\ominus\delta} = \{x \in \mathbb{R}^d : x + \delta B_{\ell_2^d} \subset A\},$$

where $B_{\ell_2^d}$ denotes the closed unit Euclidean ball at the origin. Note that the latter formulas remain true for sufficiently small $\delta > 0$, if we consider the "relative case" $X \subset \mathbb{R}^d$ and subsets $A \subset X$ satisfying $d(A, \mathbb{R}^d \setminus X) \in (0, \infty)$.

In general, it is cumbersome to determine the exact forms of $A^{+\delta}$ and $A^{-\delta}$, respectively $A^{\oplus\delta}$ and $A^{\ominus\delta}$ for a given A. For a particular class of sets $A \subset \mathbb{R}^2$, Example B.1.1 illustrates this by providing both $A^{\oplus\delta}$ and $A^{\ominus\delta}$.

The next lemma establishes some basic properties of the introduced operations.

Lemma A.3.1. *Let (X, d) be a metric space and $A, B \subset X$ be two subsets. Then the following statements hold:*

i) If A is compact, then $A^{+\delta} = A^{\oplus\delta}$.
ii) We have $d(A, B) \leq d(A^{+\delta}, B^{+\delta}) + 2\delta$.
iii) We have

$$\bigcap_{\delta>0} A^{+\delta} = \overline{A}\,. \tag{A.3.1}$$

iv) We have $(A \cup B)^{+\delta} = A^{+\delta} \cup B^{+\delta}$ and $(A \cap B)^{+\delta} \subset A^{+\delta} \cap B^{+\delta}$.
v) We have $A^{-\delta} \cup B^{-\delta} \subset (A \cup B)^{-\delta}$ and, if $d(A, B) > \delta$, we actually have $A^{-\delta} \cup B^{-\delta} = (A \cup B)^{-\delta}$.
vi) For $A_1, A_2 \subset X$ with $A_1 \cap A_2 = \emptyset$ and $B_i \subset A_i$ with $d(B_1, B_2) > \delta$, we have

$$(A_1^{-\delta} \setminus B_1^{-\delta}) \cup (A_2^{-\delta} \setminus B_2^{-\delta}) \subset (A_1 \cup A_2)^{-\delta} \setminus (B_1 \cup B_2)^{-\delta},$$

and equality holds, if $d(A_1, A_2) > \delta$.
vii) For all $\delta > 0$ and $\epsilon > 0$, we have $A \subset (A^{+\delta+\epsilon})^{-\delta}$ and $(A^{-\delta-\epsilon})^{+\delta} \subset A$.
viii) For all $\delta > 0$ and $\epsilon > 0$, we have $(\partial A)^{+\delta} \subset A^{+\delta+\epsilon} \setminus A^{-\delta-\epsilon}$.

Proof of Lemma A.3.1. *i)*. Clearly, it suffices to prove $A^{+\delta} \subset A^{\oplus\delta}$. To prove this inclusion, we fix an $x \in A^{+\delta}$. Then there exists a sequence $(x_n) \subset A$ with $d(x, x_n) \leq \delta + 1/n$ for all $n \geq 1$. Since A is compact, we may assume without loss of generality that (x_n) converges to some $x' \in A$. Now we easily obtain the assertion from $d(x, x') \leq d(x, x_n) + d(x_n, x')$.

ii). Let us fix an $x \in A^{+\delta}$ and an $y \in B^{+\delta}$. Then there exist two sequences $(x_n) \subset A$ and $(y_n) \subset B$ such that $d(x, x_n) \leq \delta + 1/n$ and $d(y, y_n) \leq \delta + 1/n$ for all $n \geq 1$. For $n \geq 1$, this construction now yields

$$d(A,B) \leq d(x_n, y_n) \leq d(x_n, x) + d(x,y) + d(y, y_n) \leq d(x,y) + 2\delta + 2/n\,,$$

and by first letting $n \to \infty$ and then taking the infimum over all $x \in A^{+\delta}$ and $y \in B^{+\delta}$, we obtain the assertion.

iii). To show the inclusion $\supset$, we fix an $x \in \overline{A}$. Then there exists a sequence $(x_n) \subset A$ with $x_n \to x$ for $n \to \infty$. For $\delta > 0$ there then exists an n_δ such that $d(x, x_n) \leq \delta$ for all $n \geq n_\delta$. This shows $d(x, A) \leq \delta$, i.e. $x \in A^{+\delta}$. To show the converse inclusion $\subset$, we fix an $x \in X$ that satisfies $x \in A^{+1/n}$ for all $n \geq 1$. Then there exists a sequence $(x_n) \subset A$ with $d(x, x_n) \leq 1/n$, and hence we find $x_n \to x$ for $n \to \infty$. This shows $x \in \overline{A}$.

iv). If $x \in (A \cup B)^{+\delta}$, there exists a sequence $(x_n) \subset A \cup B$ with $d(x, x_n) \leq \delta + 1/n$. Without loss of generality we may assume that $(x_n) \subset A$, which immediately yields $x \in A^{+\delta}$. The converse inclusion $A^{+\delta} \cup B^{+\delta} \subset (A \cup B)^{+\delta}$ and the inclusion $(A \cap B)^{+\delta} \subset A^{+\delta} \cap B^{+\delta}$ are trivial.

v). The first inclusion follows from part *iv)* and simple set algebra, namely

$$\begin{aligned} A^{-\delta} \cup B^{-\delta} = X \setminus \bigl((X \setminus A)^{+\delta} \cap (X \setminus B)^{+\delta}\bigr) &\subset X \setminus \bigl((X \setminus A) \cap (X \setminus B)\bigr)^{+\delta} \\ &= X \setminus \bigl(X \setminus (A \cup B)\bigr)^{+\delta} \\ &= (A \cup B)^{-\delta}\,. \end{aligned}$$

To show the converse inclusion, we fix an $x \in (A \cup B)^{-\delta}$. Since $(A \cup B)^{-\delta} \subset A \cup B$, we may assume without loss of generality that $x \in A$. It then remains to show that $x \in A^{-\delta}$, that is $d(x, X \setminus A) > \delta$. Obviously, $A \cap B = \emptyset$, which follows from $d(A, B) > \delta$, implies

$$X \setminus A = ((X \setminus A) \cap (X \setminus B)) \cup ((X \setminus A) \cap B) = (X \setminus (A \cup B)) \cup B\,,$$

and hence we obtain $d(x, X \setminus A) = d(x, X \setminus (A \cup B)) \wedge d(x, B) > \delta \wedge \delta = \delta$ where we used both $x \in (A \cup B)^{-\delta}$ and $d(A, B) > \delta$.

vi). Using the formula $(A_1 \cup A_2) \setminus (B_1 \cup B_2) = (A_1 \setminus B_1) \cup (A_2 \setminus B_2)$, which easily follows from $A_i \setminus B_j = A_i$ for $i \neq j$, we obtain

$$\begin{aligned} (A_1^{-\delta} \setminus B_1^{-\delta}) \cup (A_2^{-\delta} \setminus B_2^{-\delta}) &= (A_1^{-\delta} \cup A_2^{-\delta}) \setminus (B_1^{-\delta} \cup B_2^{-\delta}) \\ &\subset (A_1 \cup A_2)^{-\delta} \setminus (B_1 \cup B_2)^{-\delta}\,, \end{aligned}$$

where in the last step we used *v)*. The second assertion also follows from *v)*.

vii). Obviously, $A \subset (A^{+\delta+\epsilon})^{-\delta}$ is equivalent to $(X \setminus A^{+\delta+\epsilon})^{+\delta} \subset X \setminus A$. To prove the latter, we fix an $x \in (X \setminus A^{+\delta+\epsilon})^{+\delta}$. Then there exists a sequence $(x_n) \subset X \setminus A^{+\delta+\epsilon}$ with $d(x, x_n) \leq \delta + 1/n$ for all $n \geq 1$. Moreover, $(x_n) \subset X \setminus A^{+\delta+\epsilon}$ implies $d(x_n, x') > \delta + \epsilon$ for all $n \geq 1$ and $x' \in A$. Now assume that we had $x \in A$. For an index n with $1/n \leq \epsilon$, we would then obtain $\delta + \epsilon < d(x_n, x) \leq \delta + \epsilon$, and hence $x \in A$ cannot be true.

To show the second inclusion we fix an $x \in (A^{-\delta-\epsilon})^{+\delta}$. Then there exists a sequence $(x_n) \subset A^{-\delta-\epsilon}$ such that $d(x, x_n) \leq \delta + 1/n$ for all $n \geq 1$. This time, $x_n \in A^{-\delta-\epsilon}$ implies $x_n \notin (X \setminus A)^{+\delta+\epsilon}$, that is $d(x_n, x') > \delta + \varepsilon$ for all $n \geq 1$ and $x' \in X \setminus A$. Choosing an n with $1/n \leq \epsilon$, we then find $x \in A$.

viii). We fix an $x \in (\partial A)^{\oplus\delta}$. By definition, there then exists an $x' \in \partial A$ with $d(x, x') \leq \delta$. Moreover, by the definition of the boundary, there exists an $x'' \in A$ with $d(x', x'') \leq \epsilon$, and hence we find $d(x, x'') \leq \delta + \epsilon$, i.e. $x \in A^{+\delta+\epsilon}$. Since $\partial A = \partial(X \setminus A)$, the same argument yields $x \in (X \setminus A)^{+\delta+\epsilon}$, i.e. $x \notin A^{-\delta-\epsilon}$. Thus, we have shown $(\partial A)^{\oplus\delta} \subset A^{+\delta+\epsilon} \setminus A^{-\delta-\epsilon}$. Using $(\partial A)^{+\delta} \subset (\partial A)^{\oplus(\delta+\epsilon)}$ and a simple change of variables then yields the assertion. □

A.4. Additional Material Related to Persistence. In this section we recall and prove two results of [9] that extend Theorem 2.7.

We begin with the following lemma, which shows that $\mathcal{C}_\tau(A)$ is persistent in $\mathcal{C}_\tau(A^{+\delta})$, if $\tau > 0$ and $\delta > 0$ are sufficiently small.

Lemma A.4.1. *Let (X, d) be a compact metric space, and $A \subset X$ be non-empty. Then, for all $\delta > 0$ and $\tau > \delta$, the following statements hold:*

- *i) The set $(A')^{+\delta}$ is τ-connected for all $A' \in \mathcal{C}_\tau(A)$.*
- *ii) The CRM $\zeta : \mathcal{C}_\tau(A) \to \mathcal{C}_\tau(A^{+\delta})$ is surjective.*
- *iii) If A is closed, $|\mathcal{C}(A)| < \infty$, and $\tau \leq \tau_A^*/3$, then the CRM $\zeta : \mathcal{C}_\tau(A) \to \mathcal{C}_\tau(A^{+\delta})$ is bijective and satisfies*

$$\zeta(A') = (A')^{+\delta}, \qquad A' \in \mathcal{C}_\tau(A). \tag{A.4.1}$$

Proof of Lemma A.4.1. *i).* Since $\tau > \delta$, there exist an $\varepsilon > 0$ with $\delta + \varepsilon < \tau$. For $x \in (A')^{+\delta}$, there thus exists an $x' \in A'$ with $d(x, x') \leq \delta + \varepsilon < \tau$, i.e. x and x' are τ-connected. Since A' is τ-connected, it is then easy to show that every pair $x, x'' \in (A')^{+\delta}$ is τ-connected.

ii). Let us fix an $A' \in \mathcal{C}_\tau(A^{+\delta})$ and an $x \in A'$. For $n \geq 1$ there then exists an $x_n \in A$ with $d(x, x_n) \leq \delta + 1/n$ and since by Lemma A.2.4 there only exist finitely many τ-connected components of A, we may assume without loss of generality that there exists an $A'' \in \mathcal{C}_\tau(A)$ with $x_n \in A''$ for all $n \geq 1$. This yields $d(x, A'') \leq \delta + 1/n$ for all $n \geq 1$, and hence $d(x, A'') \leq \delta$. Consequently, we obtain $x \in (A'')^{+\delta}$, i.e. we have $(A'')^{+\delta} \cap A' \neq \emptyset$. Since $(A'')^{+\delta} \subset A^{+\delta}$, we then conclude that $(A'')^{+\delta} \subset A'$ by Corollary A.2.6 and part *i).* Furthermore, we clearly have $A'' \subset (A'')^{+\delta}$, and hence $\zeta(A'') = A'$.

iii). Let us first consider the case $|\mathcal{C}(A)| = 1$. In this case, part *i)* of Proposition A.2.10 shows $|\mathcal{C}_\tau(A)| = 1$, and thus $|\mathcal{C}_\tau(A^{+\delta})| = 1$ by the already established part *ii)*. This makes the assertion obvious.

In the case $|\mathcal{C}(A)| > 1$ we write $A_1, \dots, A_m$ for the τ-connected components of A. By part *iv)* of Lemma A.3.1 we then obtain

$$A^{+\delta} = \bigcup_{i=1}^{m} A_i^{+\delta}\,. \tag{A.4.2}$$

Since $|\mathcal{C}(A)| > 1$, we further have $\tau_A^* < \infty$, and hence part *ii)* of Proposition A.2.10 yields $\mathcal{C}(A) = \mathcal{C}_\tau(A)$. The definition of τ_A^* thus gives $d(A_i, A_j) \geq \tau_A^* \geq 3\tau$ for all $i \neq j$. Our first goal is to show that

$$d(A_i^{+\delta}, A_j^{+\delta}) \geq \tau\,, \qquad\qquad i \neq j\,. \tag{A.4.3}$$

To this end, we fix $i \neq j$ and both an $x_i \in A_i^{+\delta}$ and an $x_j \in A_j^{+\delta}$. Now, the compactness of X yields the compactness of A_i and A_j by Lemma A.2.4, and hence part *i)* of Lemma A.3.1 shows that there exist $x_i' \in A_i$ and $x_j' \in A_j$ with $d(x_i, x_i') \leq \delta$ and $d(x_j, x_j') \leq \delta$. This yields

$$3\tau \leq d(x_i', x_j') \leq d(x_i', x_i) + d(x_i, x_j) + d(x_j, x_j') \leq 2\delta + d(x_i, x_j)\,,$$

and the latter together with $\delta < \tau$ implies (A.4.3).

Now *i)* showed that each $A_i^{+\delta}$, $i = 1, \dots, m$, is τ-connected. Combining this with (A.4.2), (A.4.3), and Lemma A.2.8, we see that $A_1^{+\delta}, \dots, A_m^{+\delta}$ are the τ-connected components of $A^{+\delta}$. The bijectivity of ζ then follows from the surjectivity and a cardinality argument, and (A.4.1) is obvious. $\square$

The following theorem is an extended version of the statements of Theorem 2.7 that deal with $\mathcal{C}_\tau(M_\rho^{+\delta})$.

THEOREM A.4.2. *Let (X, d) be a compact metric space, μ be a finite Borel measure on X and P be a μ-absolutely continuous distribution on X that can be clustered between ρ^* and ρ^{**}. Then the function τ^* defined by (2.6) is monotonically increasing. Moreover, for all $\varepsilon^* \in (0, \rho^{**} - \rho^*]$, $\delta > 0$, $\tau \in (\delta, \tau^*(\varepsilon^*)]$, and all $\rho \in [0, \rho^{**}]$, the following statements hold:*

i) *We have $1 \leq |\mathcal{C}_\tau(M_\rho^{+\delta})| \leq 2$.*
ii) *If $\rho \geq \rho^* + \varepsilon^*$, then $|\mathcal{C}_\tau(M_\rho^{+\delta})| = 2$ and $\mathcal{C}(M_\rho) \sqsubseteq \mathcal{C}_\tau(M_\rho^{+\delta})$.*
iii) *If $|\mathcal{C}_\tau(M_\rho^{+\delta})| = 2$, then $\rho \geq \rho^*$ and $\mathcal{C}_\tau(M_{\rho^{**}}^{+\delta}) \sqsubseteq \mathcal{C}_\tau(M_\rho^{+\delta})$.*
iv) *If $\mathcal{C}_\tau(M_{\rho^{**}}^{-\delta}) \sqsubseteq \mathcal{C}_\tau(M_{\rho^{**}}^{+\delta})$ and $|\mathcal{C}_\tau(M_\rho^{-\delta})| = 1$, then $\rho < \rho^* + \varepsilon^*$.*

Proof of Theorem A.4.2. Let us first show the assertions related to the function τ^*. To this end, we first observe that for $\varepsilon \in (0, \rho^{**} - \rho^*]$ we have $|\mathcal{C}(M_{\rho^*+\varepsilon})| = |\mathcal{C}(M_{\rho^{**}})| = 2$ by Definition 2.5. This shows $\tau^*(\varepsilon) < \infty$.

Let us now fix $\varepsilon_1, \varepsilon_2 \in (0, \rho^{**} - \rho^*]$ with $\varepsilon_1 \leq \varepsilon_2$. Then Definition 2.5 guarantees that both $M_{\rho^*+\varepsilon_1}$ and $M_{\rho^*+\varepsilon_2}$ have two topologically connected components and that the CRM $\zeta : \mathcal{C}(M_{\rho^*+\varepsilon_2}) \to \mathcal{C}(M_{\rho^*+\varepsilon_1})$ is bijective. From Lemma A.2.11 we thus obtain

$$\tau^*(\varepsilon_2) = \frac{1}{3}\tau^*_{M_{\rho^*+\varepsilon_2}} \geq \frac{1}{3}\tau^*_{M_{\rho^*+\varepsilon_1}} = \tau^*(\varepsilon_1)\,.$$

i). Since $\emptyset \neq M_\rho \subset M_\rho^{+\delta}$, we find $|\mathcal{C}_\tau(M_\rho^{+\delta})| \geq 1$. On the other hand, since $\tau > \delta$, part *ii)* of Lemma A.4.1 and part *i)* of Proposition A.2.10 yield

$$|\mathcal{C}_\tau(M_\rho^{+\delta})| \leq |\mathcal{C}_\tau(M_\rho)| \leq |\mathcal{C}(M_\rho)| \leq 2\,. \tag{A.4.4}$$

ii). Let us fix a $\rho \in [\rho^* + \varepsilon^*, \rho^{**}]$. For $\varepsilon := \rho - \rho^*$, the monotonicity of τ^* then gives $\tau^*(\varepsilon^*) \leq \tau^*(\varepsilon)$, and hence we obtain

$$\tau \leq \frac{1}{3}\tau^*_{M_{\rho^*+\varepsilon^*}} \leq \frac{1}{3}\tau^*_{M_\rho} < \infty\,.$$

Part *ii)* of Proposition A.2.10 thus shows that the CRM $\zeta_\rho : \mathcal{C}(M_\rho) \to \mathcal{C}_\tau(M_\rho)$ is bijective. Furthermore, $\delta < \tau \leq \tau^*_{M_\rho}/3$ together with part *iii)* of Lemma A.4.1 shows that the CRM $\zeta_\delta : \mathcal{C}_\tau(M_\rho) \to \mathcal{C}_\tau(M_\rho^{+\delta})$ is bijective. Consequently, the CRM $\zeta = \zeta_\delta \circ \zeta_\rho : \mathcal{C}(M_\rho) \to \mathcal{C}_\tau(M_\rho^{+\delta})$ is bijective, and from the latter we conclude that $|\mathcal{C}_\tau(M_\rho^{+\delta})| = |\mathcal{C}(M_\rho)| = 2$.

iii). Since $|\mathcal{C}_\tau(M_\rho^{+\delta})| = 2$, the already established (A.4.4) yields $|\mathcal{C}(M_\rho)| = 2$, and hence Definition 2.5 implies both $\rho \geq \rho^*$ and the bijectivity of the CRM $\zeta^{**} : \mathcal{C}(M_{\rho^{**}}) \to \mathcal{C}(M_\rho)$. Moreover, for ρ^{**}, the already established part *ii)* shows that the CRM $\zeta_M : \mathcal{C}_\tau(M_{\rho^{**}}) \to \mathcal{C}_\tau(M_{\rho^{**}}^{+\delta})$ is bijective, and the proof of *ii)* further showed $\mathcal{C}(M_{\rho^{**}}) = \mathcal{C}_\tau(M_{\rho^{**}})$. Consequently, ζ_M equals the CRM $\mathcal{C}(M_{\rho^{**}}) \to \mathcal{C}_\tau(M_{\rho^{**}}^{+\delta})$. In addition, $\delta < \tau$ together with part *ii)* of Lemma A.4.1 and part *i)* of Proposition A.2.10 shows that the CRM $\zeta_\rho : \mathcal{C}(M_\rho) \to \mathcal{C}_\tau(M_\rho^{+\delta})$ is surjective. Now, by Lemma 2.4 these maps commute in the sense of the following diagram

$$\begin{tikzcd}
\mathcal{C}(M_{\rho^{**}}) \arrow[r, "\zeta^{**}"] \arrow[d, "\zeta_M"'] & \mathcal{C}(M_\rho) \arrow[d, "\zeta_\rho"] \\
\mathcal{C}_\tau(M_{\rho^{**}}^{+\delta}) \arrow[r, "\zeta"'] & \mathcal{C}_\tau(M_\rho^{+\delta})
\end{tikzcd}$$

and consequently, the CRM ζ is surjective. Since $|\mathcal{C}_\tau(M^{+\delta}_{\rho^{**}})| = |\mathcal{C}(M_{\rho^{**}})| = 2$ and $|\mathcal{C}_\tau(M^{+\delta}_{\rho})| = 2$, we then conclude that ζ is bijective.

iv). We proceed by contraposition. To this end, we fix an $\rho \in [\rho^* + \varepsilon^*, \rho^{**}]$. By the already established part *ii)* we then find $|\mathcal{C}_\tau(M^{+\delta}_{\rho})| = 2$, and part *iii)* thus shows that the CRM $\zeta_M : \mathcal{C}_\tau(M^{+\delta}_{\rho^{**}}) \to \mathcal{C}_\tau(M^{+\delta}_{\rho})$ is bijective. Moreover, Lemma 2.4 yields the following diagram

$$\begin{tikzcd}
\mathcal{C}_\tau(M^{-\delta}_{\rho^{**}}) \arrow[r, "\zeta"] \arrow[d, "\zeta_V"'] & \mathcal{C}_\tau(M^{+\delta}_{\rho^{**}}) \arrow[d, "\zeta_M"] \\
\mathcal{C}_\tau(M^{-\delta}_{\rho}) \arrow[r, "\zeta_{V,M}"'] & \mathcal{C}_\tau(M^{+\delta}_{\rho})
\end{tikzcd}$$

where ζ, ζ_V, and $\zeta_{V,M}$ are the corresponding CRMs. Now our assumption guarantees that ζ is bijective, and hence the diagram shows that $\zeta_{V,M} \circ \zeta_V$ is bijective. Consequently, ζ_V is injective, and from the latter we obtain $2 = |\mathcal{C}_\tau(M^{+\delta}_{\rho})| = |\mathcal{C}_\tau(M^{-\delta}_{\rho^{**}})| \leq |\mathcal{C}_\tau(M^{-\delta}_{\rho})|$. □

The next lemma investigates situations in which $\mathcal{C}_\tau(A^{-\delta})$ is persistent in $\mathcal{C}(A)$. In particular, it shows that if τ is sufficiently large compared to δ and $|\mathcal{C}_\tau(A^{-\delta})| = |\mathcal{C}(A)|$, then we obtain persistence. Informally speaking this means that gluing δ-cuts by τ-connectivity may preserve the component structure.

Lemma A.4.3. *Let (X,d) be a compact metric space, and $A \subset X$ be non-empty and closed with $|\mathcal{C}(A)| < \infty$. We define $\psi^*_A : (0,\infty) \to [0,\infty]$ by*

$$\psi^*_A(\delta) := \sup_{x \in A} d(x, A^{-\delta})\,, \qquad\qquad \delta > 0.$$

Then, for all $\delta > 0$ and all $\tau > 2\psi^*_A(\delta)$, the following statements hold:

- *i) For all $B' \in \mathcal{C}(A)$, there is at most one $A' \in \mathcal{C}_\tau(A^{-\delta})$ with $A' \cap B' \neq \emptyset$.*
- *ii) We have $|\mathcal{C}_\tau(A^{-\delta})| \leq |\mathcal{C}(A)|$.*
- *iii) If $|\mathcal{C}_\tau(A^{-\delta})| = |\mathcal{C}(A)|$, then $\mathcal{C}_\tau(A^{-\delta})$ is persistent in $\mathcal{C}(A)$. Moreover, for all $B', B'' \in \mathcal{C}(A)$ with $B' \neq B''$ we have*

$$d(B', B'') \geq \tau - 2\psi^*_A(\delta)\,. \tag{A.4.5}$$

Proof of Lemma A.4.3. *i).* Let us fix a $\psi > 2\psi^*_A(\delta)$ with $\psi < \tau$ and a $\tau' \in (0, \tau^*_A)$ such that $\psi + \tau' < \tau$, where τ^*_A is the constant defined in Proposition A.2.10. Moreover, we fix a $B' \in \mathcal{C}(A)$. By Proposition A.2.10

we then see that $\mathcal{C}(A) = \mathcal{C}_{\tau'}(A)$, and hence B' is τ'-connected. Now let $A_1, \dots, A_m$ be the τ-connected components of $A^{-\delta}$. Clearly, Lemma A.2.4 yields $d(A_i, A_j) \geq \tau$ for all $i \neq j$. Assume that *i)* is not true, that is, there exist indices i_0, j_0 with $i_0 \neq j_0$ such that $A_{i_0} \cap B' \neq \emptyset$ and $A_{j_0} \cap B' \neq \emptyset$. Thus, there exist $x' \in A_{i_0} \cap B'$ and $x'' \in A_{j_0} \cap B'$, and since B' is τ'-connected, there further exist $x_0, \dots, x_{n+1} \in B' \subset A$ with $x_0 = x'$, $x_{n+1} = x''$ and $d(x_i, x_{i+1}) < \tau'$ for all $i = 0, \dots, n$. Moreover, our assumptions guarantee $d(x_i, A^{-\delta}) < \psi/2$ for all $i = 0, \dots, n+1$. For all $i = 0, \dots, n+1$, there thus exists an index ℓ_i with

$$d(x_i, A_{\ell_i}) < \psi/2\,.$$

In addition, we have $x_0 \in A_{i_0}$ and $x_{n+1} \in A_{j_0}$ by construction, and hence we may actually choose $\ell_0 = i_0$ and $\ell_{n+1} = j_0$. Since we assumed $\ell_0 \neq \ell_{n+1}$, there then exists an $i \in \{0, \dots, n\}$ with $\ell_i \neq \ell_{i+1}$. For this index, our construction now yields

$$d(A_{\ell_i}, A_{\ell_{i+1}}) \leq d(x_i, A_{\ell_i}) + d(x_i, x_{i+1}) + d(x_{i+1}, A_{\ell_{i+1}}) < \psi + \tau' < \tau\,,$$

which contradicts the earlier established $d(A_{\ell_i}, A_{\ell_{i+1}}) \geq \tau$.

ii). Since $A^{-\delta} \subset A$, there exists, for every $A' \in \mathcal{C}_\tau(A^{-\delta})$, a $B' \in \mathcal{C}(A)$ with $A' \cap B' \neq \emptyset$. We pick one such B' and define $\zeta(A') := B'$. Now part *i)* shows that $\zeta : \mathcal{C}_\tau(A^{-\delta}) \to \mathcal{C}(A)$ is injective, and hence we find $|\mathcal{C}_\tau(A^{-\delta})| \leq |\mathcal{C}(A)|$.

iii). As mentioned in part *ii)*, we have an injective map $\zeta : \mathcal{C}_\tau(A^{-\delta}) \to \mathcal{C}(A)$ that satisfies

$$A' \cap \zeta(A') \neq \emptyset\,, \qquad A' \in \mathcal{C}_\tau(A^{-\delta})\,. \tag{A.4.6}$$

Now, $|\mathcal{C}_\tau(A^{-\delta})| = |\mathcal{C}(A)|$ together with the assumed $|\mathcal{C}(A)| < \infty$ implies that ζ is actually bijective. Let us first show that ζ is the only map that satisfies (A.4.6). To this end, assume the converse, that is, for some $A' \in \mathcal{C}_\tau(A^{-\delta})$, there exists an $B' \in \mathcal{C}(A)$ with $B' \neq \zeta(A')$ and $A' \cap B' \neq \emptyset$. Since ζ is bijective, there then exists an $A'' \in \mathcal{C}_\tau(A^{-\delta})$ with $\zeta(A'') = B'$, and hence we have $A'' \cap B' \neq \emptyset$ by (A.4.6). By part *i)*, we conclude that $A' = A''$, which in turn yields $\zeta(A') = \zeta(A'') = B'$. In other words, we have found a contradiction, and hence ζ is indeed the only map that satisfies (A.4.6).

Let us now show that $\mathcal{C}_\tau(A^{-\delta})$ is persistent in $\mathcal{C}(A)$. Since we assumed $|\mathcal{C}_\tau(A^{-\delta})| = |\mathcal{C}(A)|$, it suffices to prove that the injective map $\zeta : \mathcal{C}_\tau(A^{-\delta}) \to \mathcal{C}(A)$ defined by (A.4.6) is a CRM, i.e. it satisfies

$$A' \subset \zeta(A')\,, \qquad A' \in \mathcal{C}_\tau(A^{-\delta})\,. \tag{A.4.7}$$

To show (A.4.7), we pick an $A' \in \mathcal{C}_\tau(A^{-\delta})$ and write $B_1, \dots, B_m$ for the topologically connected components of A. Since $A^{-\delta} \subset A$, we then have

$A' \subset B_1 \cup \cdots \cup B_m$, where the latter union is disjoint. Now, we have just seen that $\zeta(A') \in \{B_1, \ldots, B_m\}$ is the only component satisfying $A' \cap \zeta(A') \neq \emptyset$, and therefore we can conclude $A' \subset \zeta(A')$.

Finally, let us show (A.4.5). To this end, we first prove that, for all $A' \in \mathcal{C}_\tau(A^{-\delta})$ and $x \in \zeta(A')$ we have

$$d(x, A') \leq \psi_A^*(\delta)\,, \tag{A.4.8}$$

where $\zeta : \mathcal{C}_\tau(A^{-\delta}) \to \mathcal{C}(A)$ is the bijective CRM considered above. Let us assume that (A.4.8) is not true, that is, there exist an $A' \in C_\tau(A^{-\delta})$ and an $x \in \zeta(A')$ such that $d(x, A') > \psi_A^*(\delta)$. Since $d(x, A^{-\delta}) \leq \psi_A^*(\delta)$, there further exists an $A'' \in C_\tau(A^{-\delta})$ with $d(x, A'') \leq \psi_A^*(\delta)$. Obviously, this yields $A' \neq A''$. Let us fix a $\tau' \in (0, \tau_A^*)$ with $2\psi_A^*(\delta) + \tau' < \tau$, and an $x' \in A'$. For $B' := \zeta(A')$, we then have $x' \in B'$ by (A.4.7), and our construction guarantees $x \in B'$. Now, the rest of the proof is similar to that of *i)*. Namely, since B' is τ'-connected, there exist $x_0, \ldots, x_{n+1} \in B'$ with $x_0 = x$, $x_{n+1} = x'$ and $d(x_i, x_{i+1}) < \tau'$ for all $i = 0, \ldots, n$. Let $A_1, \ldots, A_m$ be the τ-connected components of $A^{-\delta}$. Then, for all $i = 0, \ldots, n+1$, there exists an index ℓ_i with

$$d(x_i, A_{\ell_i}) \leq \psi_A^*(\delta)\,,$$

where we may choose $A_{\ell_0} = A''$ and $A_{\ell_{n+1}} = A'$. Since $\ell_0 \neq \ell_{n+1}$, there then exists an $i \in \{0, \ldots, n\}$ with $\ell_i \neq \ell_{i+1}$, and our construction yields

$$\tau \leq d(A_{\ell_i}, A_{\ell_{i+1}}) \leq d(x_i, A_{\ell_i}) + d(x_i, x_{i+1}) + d(x_{i+1}, A_{\ell_{i+1}}) < 2\psi_A^*(\delta) + \tau' < \tau\,.$$

To prove (A.4.5), we again assume the converse, that is, that there exist $B', B'' \in \mathcal{C}(A)$ with $B' \neq B''$ and $d(B', B'') < \tau - 2\psi_A^*(\delta)$. Then there exist $x' \in B'$ and $x'' \in B''$ such that $d(x', x'') < \tau - 2\psi_A^*(\delta)$. Now, since ζ is bijective, there exists $A', A'' \in C_\tau(A^{-\delta})$ with $A' \neq A''$, $B' = \zeta(A')$, and $B'' = \zeta(A'')$. Using (A.4.8), we then obtain

$$\tau \leq d(A', A'') \leq d(x', A') + d(x', x'') + d(x'', A'') < 2\psi_A^*(\delta) + \tau - 2\psi_A^*(\delta) = \tau\,,$$

i.e. we again have found a contradiction. □

The following theorem provides an extended version of the statements of Theorem 2.7 that deal with $\mathcal{C}_\tau(M_\rho^{-\delta})$.

Theorem A.4.4. *Let Assumption C be satisfied and* $\varepsilon^* \in (0, \rho^{**} - \rho^*]$, $\delta \in (0, \delta_{\mathrm{thick}}]$, $\tau \in (\psi(\delta), \tau^*(\varepsilon^*)]$, *and* $\rho \in [0, \rho^{**}]$. *Then, we have:*

i) We have $1 \leq |\mathcal{C}_\tau(M_\rho^{-\delta})| \leq 2$.

ii) We have $\mathcal{C}_\tau(M_{\rho^{**}}^{-\delta}) \sqsubseteq \mathcal{C}_\tau(M_{\rho^{**}}^{+\delta})$.

iii) If $|\mathcal{C}_\tau(M_\rho^{-\delta})| = 2$, *then* $\rho \geq \rho^*$ *and* $\mathcal{C}_\tau(M_{\rho^{**}}^{-\delta}) \sqsubseteq \mathcal{C}_\tau(M_\rho^{-\delta}) \sqsubseteq \mathcal{C}(M_\rho)$.

Proof of Theorem A.4.4. *i).* We first observe that $\delta \leq \delta_{\text{thick}}$ implies

$$\sup_{x\in M_\rho} d(x, M_\rho^{-\delta}) = \psi^*_{M_\rho}(\delta) \leq c_{\text{thick}}\delta^\gamma < \infty\,,$$

and thus $M_\rho^{-\delta} \neq \emptyset$, i.e. $|\mathcal{C}_\tau(M_\rho^{-\delta})| \geq 1$. Conversely, we have $|\mathcal{C}_\tau(M_\rho^{-\delta})| \leq |\mathcal{C}(M_\rho)| \leq 2$, where the first inequality was established in part *ii)* of Lemma A.4.3 and the second is ensured by Definition 2.5.

ii). The monotonicity of τ^* established in Theorem A.4.2 yields $\delta < \psi(\delta) < \tau \leq \tau^*(\varepsilon^*) \leq \tau^*_{M_{\rho^{**}}}/3$. By part *iii)* of Lemma A.4.1 we then conclude that the CRM $\mathcal{C}_\tau(M_{\rho^{**}}) \to \mathcal{C}_\tau(M_{\rho^{**}}^{+\delta})$ is bijective, and part *ii)* of Theorem A.4.2 shows $|\mathcal{C}_\tau(M_{\rho^{**}})| = |\mathcal{C}_\tau(M_{\rho^{**}}^{+\delta})| = 2$. By Lemma 2.4 it thus suffices to show that the CRM $\zeta : \mathcal{C}_\tau(M_{\rho^{**}}^{-\delta}) \to \mathcal{C}_\tau(M_{\rho^{**}})$ is bijective. Furthermore, if $|\mathcal{C}_\tau(M_{\rho^{**}}^{-\delta})| = 1$, this map is automatically injective, and if $|\mathcal{C}_\tau(M_{\rho^{**}}^{-\delta})| = 2$, the injectivity follows from the surjectivity and the above proven $|\mathcal{C}_\tau(M_{\rho^{**}})| = 2$. Consequently, it actually suffices to show that ζ is surjective. To this end, we fix a $B' \in \mathcal{C}_\tau(M_{\rho^{**}})$ and an $x \in B'$. Then our assumption ensures $d(x, M_{\rho^{**}}^{-\delta}) < \psi(\delta)$, and hence there exists an $A' \in \mathcal{C}_\tau(M_{\rho^{**}}^{-\delta})$ with $d(x, A') < \psi(\delta)$. Therefore, $\psi(\delta) < \tau$ implies that x and A' are τ-connected, which yields $x \in A'$. In other words, we have shown $A' \cap B' \neq \emptyset$. By Lemma A.2.6 and the definition of ζ, we conclude that $\zeta(A') = B'$.

iii). We have $2 = |\mathcal{C}_\tau(M_\rho^{-\delta})| \leq |\mathcal{C}(M_\rho)| \leq 2$, where the first inequality was shown in part *ii)* of Lemma A.4.3 and the second is guaranteed by Definition 2.5. We conclude that $|\mathcal{C}(M_\rho)| = 2$, and hence Definition 2.5 ensures both $\rho \geq \rho^*$ and the bijectivity of the CRM $\zeta_{\text{top}} : \mathcal{C}(M_{\rho^{**}}) \to \mathcal{C}(M_\rho)$. Furthermore, $|\mathcal{C}_\tau(M_\rho^{-\delta})| = |\mathcal{C}(M_\rho)|$, which has been shown above, together with part *iii)* of Lemma A.4.3 yields a bijective CRM $\zeta_\rho : \mathcal{C}_\tau(M_\rho^{-\delta}) \to \mathcal{C}(M_\rho)$, i.e. the second persistence $C_\tau(M_\rho^{-\delta}) \sqsubseteq \mathcal{C}(M_\rho)$ is shown. Moreover, part *ii)* of Theorem A.4.2 shows $|\mathcal{C}_\tau(M_{\rho^{**}}^{+\delta})| = 2$, and hence the already established bijectivity of $\zeta^{**} : \mathcal{C}_\tau(M_{\rho^{**}}^{-\delta}) \to \mathcal{C}_\tau(M_{\rho^{**}}^{+\delta})$ gives $|\mathcal{C}_\tau(M_{\rho^{**}}^{-\delta})| = |\mathcal{C}_\tau(M_{\rho^{**}}^{+\delta})| = 2 = |\mathcal{C}(M_{\rho^{**}})|$. Consequently, part *iii)* of Lemma A.4.3 yields a bijective CRM $\zeta_{\rho^{**}} : \mathcal{C}_\tau(M_{\rho^{**}}^{-\delta}) \to \mathcal{C}(M_{\rho^{**}})$. Then the CRM $\zeta : \mathcal{C}_\tau(M_{\rho^{**}}^{-\delta}) \to \mathcal{C}_\tau(M_\rho^{-\delta})$ enjoys the following diagram

$$\begin{tikzcd} \mathcal{C}_\tau(M_{\rho^{**}}^{-\delta}) \arrow[r, "\zeta_{\rho^{**}}"] \arrow[d, "\zeta"'] & \mathcal{C}(M_{\rho^{**}}) \arrow[d, "\zeta_{\mathrm{top}}"] \\ \mathcal{C}_\tau(M_\rho^{-\delta}) \arrow[r, "\zeta_\rho"'] & \mathcal{C}(M_\rho) \end{tikzcd}$$

whose commutativity follows from Lemma 2.4. Then the bijectivity of $\zeta_{\rho^{**}}$, ζ_{top}, and ζ_ρ yields the bijectivity of ζ, which completes the proof. □

A.5. Additional Material Related to Thickness. In this section we discuss some aspects related to the thickness assumption introduced in Definition 2.6.

To this end, let (X, d) be an arbitrary metric spaces and $A \subset X$. We then define the function $\psi_A^* : (0, \infty) \to [0, \infty]$ by

$$\psi_A^*(\delta) := \sup_{x \in M_\rho} d(x, A^{-\delta}) , \qquad\qquad \delta > 0.$$

Obviously, $\psi_{M_\rho}^*$ coincides with the left-hand side of (2.5).

Our first observation is that the definition of ψ_A^* immediately yields $A \subset (A^{-\delta})^{+\psi_A^*(\delta)}$ for all $\delta > 0$ with $\psi_A^*(\delta) < \infty$, and it is also straightforward to see that $\psi_A^*(\delta)$ is the smallest $\psi > 0$, for which this inclusion holds, that is

$$\psi_A^*(\delta) = \min\bigl\{\psi \geq 0 : A \subset (A^{-\delta})^{+\psi}\bigr\}$$

for all $\delta > 0$. In other words, $\psi_A^*(\delta)$ gives the size of the smallest tube needed to recover a superset of A from $A^{-\delta}$. In particular, if δ is too large, that is $A^{-\delta} = \emptyset$, we obviously have $\psi_A^*(\delta) = \infty$ and no recovery is possible.

Intuitively it is not surprising that ψ_A^* grows at least linearly, that is

$$\psi_A^*(\delta) \geq \delta \tag{A.5.1}$$

for all $\delta > 0$ provided that $d(A, X \setminus A) = 0$. Indeed, $\psi_A^*(\delta) < \delta$ for some $\delta > 0$ gives us an $\epsilon > 0$ such that $d(x, A^{-\delta}) < \delta - \epsilon$ for all $x \in A$. Since $d(A, X \setminus A) = 0$ there then exists an $x \in A$ with $d(x, X \setminus A) < \epsilon$, and for this x there exists an $x' \in A^{-\delta}$ with $d(x, x') < \delta - \epsilon$. Now the definition of $A^{-\delta}$ gives $d(x', X \setminus A) > \delta$, and hence we find a contradiction by

$$\delta < d(x', X \setminus A) \leq d(x', x) + d(x, X \setminus A) < \delta .$$

For generic sets A, the function ψ_A^* is usually hard to bound, but for some classes of sets, ψ_A^* can be computed precisely. For example, for an interval

$I = [a,b]$, we have $\psi_I^*(\delta) = \delta$ for all $\delta \in (0,(b-a)/2]$, and $\psi_I^*(\delta) = \infty$, otherwise. Clearly, this example can be extended to finite unions of such intervals and for intervals that are not closed, the only difference occurs at $\delta = (b-a)/2$. In higher dimensions, an interesting class of sets A with linear behavior of ψ_A^* is described by Serra's model, see [7, p. 144], that consist of all compact sets $A \subset \mathbb{R}^d$ for which there is a $\delta_0 > 0$ with

$$A = (A^{\ominus\delta_0})^{\oplus\delta_0} = (A^{\oplus\delta_0})^{\ominus\delta_0} .$$

If, in addition, A is path-connected, then [11, Theorem 1] shows that this relation also holds for all $\delta \in (0, \delta_0]$. In this case, we then obtain

$$A = (A^{\ominus(\delta+\epsilon)})^{\oplus(\delta+\epsilon)} \subset (A^{\ominus(\delta+\epsilon)})^{+\delta+\epsilon} \subset (A^{-\delta})^{+\delta+\epsilon}$$

for all $\delta \in (0, \delta_0)$ and $0 < \epsilon \le \delta_0 - \delta$. In other words, we have $\psi_A^*(\delta) \le \delta + \epsilon$, and letting $\epsilon \to 0$, we thus conclude $\psi_A^*(\delta) = \delta$ for all $\delta \in (0, \delta_0)$. With the help of Lemma A.3.1, it is not hard to see that this result generalizes to finite unions of compact, path-connected sets, which has already been observed in [11]. Finally, note that [11, Theorem 1] also provides some useful characterizations of (path-connected) compact sets belonging to Serra's model. In a nutshell, these are the sets whose boundary is a $(d-1)$-dimensional sub-manifold of $\mathbb{R}^d$ with outward pointing unit normal vectors satisfying a Lipschitz condition.

Fortunately, our analysis does not require the exact form of ψ_A^*, but only its asymptotic behavior for $\delta \to 0$. Therefore, it is interesting to note that ψ_A^* is also asymptotically invariant against bi-Lipschitz transformations. To be more precise, let (X, d) and (Y, e) be two metric spaces and $I : X \to Y$ be a bijective map for which there exists a constant $C > 0$ such that

$$C^{-1} e(I(x), I(x')) \le d(x, x') \le C e(I(x), I(x'))$$

for all $x, x' \in X$. For $A \subset X$ and $\delta > 0$, we then have $I(A^{+\delta/C}) \subset (I(A))^{+\delta} \subset I(A^{+C\delta})$, which in turn implies

$$C^{-1} \psi_A^*(\delta/C) \le \psi_{I(A)}^*(\delta) \le C \psi_A^*(C\delta)$$

for all $\delta > 0$. In particular, we have $\psi_A^*(\delta) \preceq \delta^\gamma$ for some $\gamma \in (0,1]$ if and only if $\psi_{I(A)}^*(\delta) \preceq \delta^\gamma$.

Last but not least we like to mention that based on the sets $A \subset \mathbb{R}^2$ considered in Example B.1.1, Example B.1.2 estimates ψ_A^*. In particular, this example provides various sets A with $\psi_A^*(\delta) \sim \delta$ that do not belong to Serra's model, and this class of sets can be further expanded by using bi-Lipschitz transformations as discussed above.

Now consider Definition 2.6, which excludes thin cusps and bridges, where the thinness and length of both is controlled by γ. Such assumptions have been widely used in the literature on level set estimation and density-based clustering. For example, a basically identical assumption has been made in [8] for the exponent $\gamma = 1$, which can be taken, if, e.g., the level sets belong to Serra's model. Moreover, level sets belonging to Serra's model have been investigated in [10]. In particular, [10, Theorem 2] shows that most level sets of a C^1-density with Lipschitz continuous gradient belong to Serra's model. Unfortunately, however, levels at which the density has a saddle point are excluded in this theorem, and some other elementary sets such as cubes in $\mathbb{R}^d$ do not belong to Serra's model, either. For this reason, we allow constants $c_{\text{thick}} > 1$ in Definition 2.6. Moreover, the exponent $\gamma < 1$ is allowed to provide more flexibility in situations, in which very thin bridges are expected. However, based on the discussion on ψ_A^* as well as the examples provided in Section B.2, we strongly believe, that in most cases assuming $\gamma = 1$ is reasonable. With the help of the discussion on ψ_A^* it is also easy to see that we have $M_\rho \subset (M^{-\delta})^{+\psi(\delta)/2}$ for all $\delta \in (0, \delta_{\text{thick}}]$ and all $\rho \in (0, \rho^{**}]$. In addition, it becomes clear that exponents $\gamma > 1$ are impossible as soon as $d(M_\rho, X \setminus M_\rho) = 0$ for some $\rho \in (0, \rho^{**}]$. Finally, recall that a less geometric assumption excluding thin features has been used by various authors, see e.g. [3, 2, 6] and the references therein, and an overview of these and similar assumptions can be found in [1].

Understanding (2.5) in the one-dimensional case is very simple. Indeed, if $X \subset \mathbb{R}$ is an interval and P can be topologically clustered between ρ^* and ρ^{**}, then, for all $\rho \in [0, \rho^{**}]$, the level set M_ρ consists of either one or two closed intervals. Using this, the discussion on ψ_A^* shows that P actually has thick levels of order $\gamma = 1$ up to the level ρ^{**}. Moreover, a possible thickness function is $\psi(\delta) = 3\delta$ for all $\delta \in (0, \delta_{\text{thick}}]$, where δ_{thick} equals the smaller radius of the two intervals at level ρ^{**}.

Finally, using the discussion on ψ_A^* it is not hard to construct distributions with discontinuous densities that have thick levels of order, e.g. $\gamma = 1$. For continuous densities, however, this task is significantly harder due to the above mentioned saddle point effects at the critical level ρ^*. Therefore, we have added Example B.2.1, which provides a large class of such densities in the case $X \subset \mathbb{R}^2$.

A.6. Proofs and Results Related to Algorithm 2.1. The main goals of this section is to prove Theorem 2.8 and to provide background material from [9] for the proof of Theorem 2.9.

Lemma A.6.1. *Let (X,d) be a compact metric space and μ be a finite Borel measure on X with $\operatorname{supp}\mu = X$. Moreover, let P be a μ-absolutely continuous distribution on X, and $(L_\rho)_{\rho\geq 0}$ be a decreasing family of sets $L_\rho \subset X$ such that*

$$M^{-\delta}_{\rho+\varepsilon} \subset L_\rho \subset M^{+\delta}_{\rho-\varepsilon}$$

for some fixed $\delta > 0$, $\varepsilon \geq 0$, and all $\rho \geq 0$. For some fixed $\rho \geq 0$ and $\tau > 0$, let $\zeta : \mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}) \to \mathcal{C}_\tau(L_\rho)$ be the CRM. Then we have:

i) For all $A' \in \mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})$ with $A' \cap M^{-\delta}_{\rho+3\varepsilon} \neq \emptyset$ we have $\zeta(A') \cap L_{\rho+2\varepsilon} \neq \emptyset$.

ii) For all $B' \in \mathcal{C}_\tau(L_\rho)$ with $B' \notin \zeta(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}))$, we have

$$B' \subset (X \setminus M_{\rho+\varepsilon})^{+\delta} \cap M^{+\delta}_{\rho-\varepsilon} \tag{A.6.1}$$

$$B' \cap L_{\rho+2\varepsilon} \subset (X \setminus M_{\rho+\varepsilon})^{+\delta} \cap M^{+\delta}_{\rho+\varepsilon}\,. \tag{A.6.2}$$

Proof of Lemma A.6.1. *i).* Using the CRM property $A' \subset \zeta(A')$ and the inclusion $M^{-\delta}_{\rho+3\varepsilon} \subset L_{\rho+2\varepsilon}$, we obtain

$$\emptyset \neq A' \cap M^{-\delta}_{\rho+3\varepsilon} \subset \xi(A') \cap L_{\rho+2\varepsilon}\,.$$

ii). We fix a $B' \in \mathcal{C}_\tau(L_\rho) \setminus \zeta(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}))$. For $x \in B'$ we then have

$$x \notin \bigcup_{A' \in \mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})} \zeta(A')\,,$$

and hence the CRM property yields

$$x \notin \bigcup_{A' \in \mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})} A' = M^{-\delta}_{\rho+\varepsilon}\,.$$

This shows $x \in (X \setminus M_{\rho+\varepsilon})^{+\delta}$, i.e. we have proved $B' \subset (X \setminus M_{\rho+\varepsilon})^{+\delta}$. Now, (A.6.1) follows from $B' \subset L_\rho \subset M^{+\delta}_{\rho-\varepsilon}$, and (A.6.2) follows from $B' \cap L_{\rho+2\varepsilon} \subset L_{\rho+2\varepsilon} \subset M^{+\delta}_{\rho+\varepsilon}$. □

Proof of Theorem 2.8. We first establish the following *disjoint* union:

$$\begin{aligned}\mathcal{C}_\tau(L_\rho) = \zeta(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})) &\cup \bigl\{B' \in \mathcal{C}_\tau(L_\rho) \setminus \zeta(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})) : B' \cap L_{\rho+2\varepsilon} \neq \emptyset\bigr\} \\ &\cup \bigl\{B' \in \mathcal{C}_\tau(L_\rho) : B' \cap L_{\rho+2\varepsilon} = \emptyset\bigr\}\,. \end{aligned} \tag{A.6.3}$$

We begin by showing the auxiliary result

$$A' \cap M^{-\delta}_{\rho+3\varepsilon} \neq \emptyset\,, \qquad A' \in \mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}). \tag{A.6.4}$$

To this end, we observe that *i)* and *ii)* of Theorem A.4.2 yield $|\mathcal{C}_\tau(M^{+\delta}_{\rho^{**}})| = 2$, and hence part *ii)* of Theorem A.4.4 implies $|\mathcal{C}_\tau(M^{-\delta}_{\rho^{**}})| = 2$. Let W' and W'' be the two τ-connected components of $M^{-\delta}_{\rho^{**}}$. We first assume that $M^{-\delta}_{\rho+\varepsilon}$ has exactly one τ-connected component A', i.e. $A' = M^{-\delta}_{\rho+\varepsilon}$. Then $\rho + 3\varepsilon \leq \rho^{**}$ and $\rho + \varepsilon \leq \rho + 3\varepsilon$ imply

$$\emptyset \neq M^{-\delta}_{\rho^{**}} \subset M^{-\delta}_{\rho+3\varepsilon} = M^{-\delta}_{\rho+\varepsilon} \cap M^{-\delta}_{\rho+3\varepsilon} = A' \cap M^{-\delta}_{\rho+3\varepsilon}\,,$$

i.e. we have shown (A.6.4). Let us now assume that $M^{-\delta}_{\rho+\varepsilon}$ has more than one τ-component. Then it has exactly two such components A' and A'' by $\rho + \varepsilon < \rho^{**}$ and part *i)* of Theorem A.4.4. By part *iii)* of Theorem A.4.4 we may then assume without loss of generality that we have $W' \subset A'$ and $W'' \subset A''$. Since $\rho + 3\varepsilon \leq \rho^{**}$ implies $M^{-\delta}_{\rho^{**}} \subset M^{-\delta}_{\rho+3\varepsilon}$, these inclusions yield $\emptyset \neq W' = W' \cap M^{-\delta}_{\rho^{**}} \subset A' \cap M^{-\delta}_{\rho+3\varepsilon}$ and $\emptyset \neq W'' = W'' \cap M^{-\delta}_{\rho^{**}} \subset A'' \cap M^{-\delta}_{\rho+3\varepsilon}$. Consequently, we have proved (A.6.4) in this case, too.

Now, from (A.6.4) we conclude by part *i)* of Lemma A.6.1 that $B' \cap L_{\rho+2\varepsilon} \neq \emptyset$ for all $B' \in \zeta(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}))$. This yields

$$\begin{aligned}&\bigl\{B' \in \mathcal{C}_\tau(L_\rho) \setminus \zeta(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})) : B' \cap L_{\rho+2\varepsilon} = \emptyset\bigr\}\\ &= \bigl\{B' \in \mathcal{C}_\tau(L_\rho) : B' \cap L_{\rho+2\varepsilon} = \emptyset\bigr\}\,,\end{aligned}$$

which in turn implies (A.6.3).

Let us now show (2.8). Clearly, by (A.6.3) it remains to show

$$B' \cap L_{\rho+2\varepsilon} = \emptyset\,,$$

for all $B' \in \mathcal{C}_\tau(L_\rho) \setminus \zeta(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}))$. Let us assume the converse, that is, there exists a $B' \in \mathcal{C}_\tau(L_\rho) \setminus \zeta(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}))$ with $B' \cap L_{\rho+2\varepsilon} \neq \emptyset$. Since $L_{\rho+2\varepsilon} \subset M^{+\delta}_{\rho+\varepsilon}$, there then exists an $x \in B' \cap M^{+\delta}_{\rho+\varepsilon}$. By part *i)* of Lemma A.3.1 this gives an $x' \in M_{\rho+\varepsilon}$ with $d(x, x') \leq \delta$, and hence we obtain

$$d(x', M^{-\delta}_{\rho+\varepsilon}) \leq \psi^*_{M_{\rho+\varepsilon}}(\delta) \leq c_{\mathrm{thick}}\delta^\gamma < 2c_{\mathrm{thick}}\delta^\gamma\,.$$

From this inequality we conclude that there exists an $x'' \in M^{-\delta}_{\rho+\varepsilon}$ satisfying $d(x', x'') < 2c_{\mathrm{thick}}\delta^\gamma$. Let $A'' \in \mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})$ be the unique τ-connected component satisfying $x'' \in A''$. The CRM property then yields $x'' \in A'' \subset \zeta(A'') =: B''$, and thus, using $c \geq 1$, we find

$$d(B', B'') \leq d(x, x'') \leq d(x, x') + d(x', x'') < \delta + 2c_{\mathrm{thick}}\delta^\gamma \leq 3c_{\mathrm{thick}}\delta^\gamma < \tau\,.$$

However, since $B' \notin \zeta(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}))$ and $B'' \in \zeta(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}))$ we obtain $B' \neq B''$, and hence Lemma A.2.4 yields $d(B', B'') \geq \tau$. $\square$

THEOREM A.6.2. *Let Assumption C be satisfied. Furthermore, let* $\varepsilon^* \leq (\rho^{**} - \rho^*)/9$, $\delta \in (0, \delta_{\text{thick}}]$, $\tau \in (\psi(\delta), \tau^*(\varepsilon^*)]$, *and* $\varepsilon \in (0, \varepsilon^*]$. *In addition, let* D *be a data set and* $(L_{D,\rho})_{\rho \geq 0}$ *be a decreasing family satisfying*

$$M_{\rho+\varepsilon}^{-\delta} \subset L_{D,\rho} \subset M_{\rho-\varepsilon}^{+\delta}$$

for all $\rho \geq 0$. *Furthermore, assume that Algorithm 2.1 receives the parameters* τ, ε, *and* $(L_{D,\rho})_{\rho \geq 0}$. *Then, the following statements are true:*

i) The returned level ρ_D^* *satisfies* $\rho_D^* \in [\rho^* + 2\varepsilon, \rho^* + \varepsilon^* + 5\varepsilon]$.

ii) We have $|\mathcal{C}_\tau(M_{\rho_D^*+\varepsilon}^{-\delta})| = 2$ *and the CRM* $\zeta : \mathcal{C}_\tau(M_{\rho_D^*+\varepsilon}^{-\delta}) \to \mathcal{C}_\tau(L_{D,\rho_D^*})$ *is injective.*

iii) Algorithm 2.1 returns the two τ*-connected components of* $\zeta(C_\tau(M_{\rho_D^*+\varepsilon}^{-\delta}))$.

iv) There exist CRMs $\zeta_{\rho^{**}} : \mathcal{C}_\tau(M_{\rho^{**}}^{-\delta}) \to \mathcal{C}(M_{\rho^{**}})$ *and* $\zeta_{\rho_D^*+\varepsilon} : \mathcal{C}_\tau(M_{\rho_D^*+\varepsilon}^{-\delta}) \to \mathcal{C}(M_{\rho_D^*+\varepsilon})$ *such that we have a commutative diagram of bijective CRMs:*

$$\begin{tikzcd}
\mathcal{C}_\tau(M_{\rho^{**}}^{-\delta}) \arrow[r, "\zeta_{\rho^{**}}"] \arrow[d, "\zeta_{\rho^{**},\rho_D^*+\varepsilon}"'] & \mathcal{C}(M_{\rho^{**}}) \arrow[d, "\tilde{\zeta}"] \\
\mathcal{C}_\tau(M_{\rho_D^*+\varepsilon}^{-\delta}) \arrow[r, "\zeta_{\rho_D^*+\varepsilon}"'] & \mathcal{C}(M_{\rho_D^*+\varepsilon})
\end{tikzcd}$$

PROOF OF THEOREM A.6.2. We begin with some general observations. To this end, let $\rho \in [0, \rho^{**} - 4\varepsilon]$ be the level that is currently considered in Line 3 of Algorithm 2.1. Then, Theorem 2.8 shows that Algorithm 2.1 identifies exactly the τ-connected components of $L_{D,\rho}$ that belong to the set $\zeta(\mathcal{C}_\tau(M_{\rho+\varepsilon}^{-\delta}))$, where $\zeta : \mathcal{C}_\tau(M_{\rho+\varepsilon}^{-\delta}) \to \mathcal{C}_\tau(L_{D,\rho})$ is the CRM. In the following, we thus consider the set $\zeta(\mathcal{C}_\tau(M_{\rho+\varepsilon}^{-\delta}))$. Moreover, we note that the returned level ρ_D^* always satisfies $\rho_D^* \geq \rho + 3\varepsilon$ by Line 4 and Line 6, and equality holds if and only if $|\zeta(\mathcal{C}_\tau(M_{\rho+\varepsilon}^{-\delta}))| \neq 1$.

i). Let us first consider the case $\rho \in [0, \rho^* - \varepsilon)$. Then $\rho + \varepsilon < \rho^*$ together with part *i)* and *iii)* of Theorem A.4.4 shows $|\mathcal{C}_\tau(M_{\rho+\varepsilon}^{-\delta})| = 1$, and hence $|\zeta(\mathcal{C}_\tau(M_{\rho+\varepsilon}^{-\delta}))| = 1$. Our initial consideration then show, that Algorithm 2.1 does not leave its loop, and thus $\rho_D^* \geq \rho^* + 2\varepsilon$.

Let us now consider the case $\rho \in [\rho^* + \varepsilon^* + \varepsilon, \rho^* + \varepsilon^* + 2\varepsilon]$. Here we first note that Algorithm 2.1 actually inspects such an ρ, since it iteratively inspects all $\rho = i\varepsilon$, $i = 0, 1, \ldots$, and the width of the interval above is ε. Moreover, our assumptions on ε^* and ε guarantee $\rho^* + \varepsilon^* + 2\varepsilon \leq \rho^{**} - 4\varepsilon$, and hence we have $\rho \in [\rho^* + \varepsilon^* + \varepsilon, \rho^{**} - 4\varepsilon]$, i.e., we are in the situation described at the beginning of the proof. We write $\zeta_V : \mathcal{C}_\tau(M^{-\delta}_{\rho^{**}}) \to \mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})$, $\zeta_M : \mathcal{C}_\tau(M^{+\delta}_{\rho^{**}}) \to \mathcal{C}_\tau(M^{+\delta}_{\rho-\varepsilon})$, and $\zeta_{V,M} : \mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}) \to \mathcal{C}_\tau(M^{+\delta}_{\rho-\varepsilon})$ for the CRMs between the involved sets. We then obtain the commutative diagram

$$\begin{tikzcd}
\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}) \arrow[r, "\zeta_{V,M}"] & \mathcal{C}_\tau(M^{+\delta}_{\rho-\varepsilon}) \\
\mathcal{C}_\tau(M^{-\delta}_{\rho^{**}}) \arrow[u, "\zeta_V"] \arrow[r, "\zeta^{**}"'] & \mathcal{C}_\tau(M^{+\delta}_{\rho^{**}}) \arrow[u, "\zeta_M"']
\end{tikzcd}$$

where the CRM ζ^{**} is bijective by part *ii)* of Theorem A.4.4. Moreover, $\rho-\varepsilon \geq \rho^*+\varepsilon^*$ together with part *ii)* of Theorem A.4.2 shows $|\mathcal{C}_\tau(M^{+\delta}_{\rho-\varepsilon})| = 2$, and by *iii)* of Theorem A.4.2 we conclude that ζ_M is bijective. Similarly, $\rho + \varepsilon \geq \rho^* + \varepsilon^*$ and the bijectivity of ζ^{**} show by *iv)* of Theorem A.4.2 that $|\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})| = 2$, and thus ζ_V is bijective by part *iii)* of Theorem A.4.4. Consequently, $\zeta_{V,M}$ is bijective. Let us further consider the CRM $\zeta' : \mathcal{C}_\tau(L_{D,\rho}) \to \mathcal{C}_\tau(M^{+\delta}_{\rho-\varepsilon})$. Then Lemma 2.4 yields another diagram:

$$\begin{tikzcd}
\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}) \arrow[rr, "\zeta_{V,M}"] \arrow[dr, "\zeta"'] & & \mathcal{C}_\tau(M^{+\delta}_{\rho-\varepsilon}) \\
& \mathcal{C}_\tau(L_{D,\rho}) \arrow[ur, "\zeta'"'] &
\end{tikzcd}$$

Since $\zeta_{V,M}$ is bijective, we then find that ζ is injective, and since we have already seen that $M^{-\delta}_{\rho+\varepsilon}$ has two τ-connected components, we conclude that $\zeta(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}))$ contains two elements. Consequently, the stopping criterion of Algorithm 2.1 is satisfied, that is, $\rho^*_D = \rho + 3\varepsilon \leq \rho^* + \varepsilon^* + 5\varepsilon$.

ii). Theorem 2.8 shows that in its last run through the loop Algorithm 2.1 identifies exactly the τ-connected components of $L_{D,\rho}$ that belong to the set $\zeta_{-3\varepsilon}(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})$, where $\rho := \rho^*_D - 3\varepsilon$ and $\zeta_{-3\varepsilon} : \mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}) \to \mathcal{C}_\tau(L_{D,\rho})$ is the CRM. Moreover, since Algorithm 2.1 stops at $\rho^*_D - 3\varepsilon$, we have $|\zeta_{-3\varepsilon}(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}))| \neq 1$ by our remarks at the beginning of the proof, and thus $|\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})| \neq 1$. From the already proven part *i)* we further know that $\rho + \varepsilon = \rho^*_D - 2\varepsilon \leq \rho^* + \varepsilon^* + 3\varepsilon \leq \rho^* + 4\varepsilon^* \leq \rho^{**}$, and part *i)* of Theorem A.4.4 hence gives $|\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})| = 2$. For later purposes, note that the latter together with $|\zeta_{-3\varepsilon}(\mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}))| \neq 1$ implies the injectivity of $\zeta_{-3\varepsilon}$. Now, part *iii)* of Theorem A.4.4 shows that the CRM $\zeta_{\rho^{**},\rho+\varepsilon} : \mathcal{C}_\tau(M^{-\delta}_{\rho^{**}}) \to \mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon})$ is bijective. Let us consider the following commutative diagram:

$$\begin{tikzcd}
\mathcal{C}_\tau(M^{-\delta}_{\rho^{**}}) \arrow[rr, "\zeta_{\rho^{**},\rho+\varepsilon}"] \arrow[dr, "\zeta_{\rho^{**},\rho^*_D+\varepsilon}"'] & & \mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}) \\
& \mathcal{C}_\tau(M^{-\delta}_{\rho^*_D+\varepsilon}) \arrow[ur, "\tilde{\zeta}"'] &
\end{tikzcd}$$

where the remaining two maps are the corresponding CRMs, whose existence is guaranteed by $\rho^*_D + \varepsilon \leq \rho^*_D + 7\varepsilon^* \leq \rho^{**}$ and $\rho + \varepsilon \leq \rho^*_D + \varepsilon$, respectively. Now the bijectivity of $\zeta_{\rho^{**},\rho+\varepsilon}$ shows that $\zeta_{\rho^{**},\rho^*_D+\varepsilon}$ is injective. Moreover, $\rho^*_D + \varepsilon \leq \rho^{**}$ implies $|\mathcal{C}_\tau(M^{-\delta}_{\rho^*_D+\varepsilon})| \leq 2$ by part *i)* of Theorem A.4.4, while $\rho^{**} \geq \rho^* + \varepsilon^*$ implies $|\mathcal{C}_\tau(M^{-\delta}_{\rho^{**}})| = 2$ by part *iv)* of Theorem A.4.2 and part *ii)* of Theorem A.4.4. Therefore, $\zeta_{\rho^{**},\rho^*_D+\varepsilon}$ is actually bijective. This yields both $|\mathcal{C}_\tau(M^{-\delta}_{\rho^*_D+\varepsilon})| = 2$, which is the first assertion, and the bijectivity of $\tilde{\zeta}$. Let us consider yet another commutative diagram

$$\begin{tikzcd}
\mathcal{C}_\tau(M^{-\delta}_{\rho^*_D+\varepsilon}) \arrow[r, "\tilde{\zeta}"] \arrow[d, "\zeta"'] & \mathcal{C}_\tau(M^{-\delta}_{\rho+\varepsilon}) \arrow[d, "\zeta_{-3\varepsilon}"] \\
\mathcal{C}_\tau((L_{D,\rho^*_D}) \arrow[r, "\zeta'"'] & \mathcal{C}_\tau(L_{D,\rho})
\end{tikzcd}$$

where again, all occurring maps are the CRMs between the respective sets. Now we have already shown that $\zeta_{-3\varepsilon}$ is injective and that $\tilde{\zeta}$ is bijective. Consequently, ζ is injective.

iii). This assertions follows from Theorem 2.8 and the inequality $\rho^*_D \leq \rho^{**} - 3\varepsilon$, which follows from part *i)*.

iv). We have already seen in the proof of part *ii)* that $|\mathcal{C}_\tau(M^{-\delta}_{\rho^{**}})| = 2$, and consequently part *iii)* of Lemma A.4.3 shows that there exists a bijective CRM $\zeta_{\rho^{**}} : \mathcal{C}_\tau(M^{-\delta}_{\rho^{**}}) \to \mathcal{C}(M_{\rho^{**}})$. Moreover, part *ii)* shows $|\mathcal{C}_\tau(M^{-\delta}_{\rho^*_D+\varepsilon})| = 2$, thus part *iii)* of Lemma A.4.3 yields another bijective CRM $\zeta_{\rho^*_D+\varepsilon} : \mathcal{C}_\tau(M^{-\delta}_{\rho^*_D+\varepsilon}) \to \mathcal{C}(M_{\rho^*_D+\varepsilon})$. Furthermore, in the proof of part *ii)* we have already seen that CRM $\zeta_{\rho^{**},\rho^*_D+\varepsilon}$ is bijective. This gives the diagram. □

A.7. Additional Material Related to Assumption A. In this section we discuss Assumption A, which describes the partitions needed for our histogram approach, in more detail.

We begin with an example of partitions satisfying Assumption A.

Example A.7.1. *Let $X := [0,1]^{\mathrm{d}}$ be equipped with the metric defined by the supremum norm $\|\cdot\|_{\ell_\infty^{\mathrm{d}}}$, and λ^{d} be the Lebesgue measure. For $\delta \in (0,1]$, there then exists a unique $\ell \in \mathbb{N}$ with $\frac{1}{\ell+1} < \delta \le \frac{1}{\ell}$. We define $h := \frac{1}{1+\ell}$ and write $\mathcal{A}_\delta$ for the usual partition of $[0,1]^{\mathrm{d}}$ into hypercubes of side-length h. Then, for each $A_i \in \mathcal{A}_\delta$, we have $\operatorname{diam} A_i = h \le \delta$ and $\lambda^{\mathrm{d}}(A_i) = h^{\mathrm{d}} \ge 2^{-d}\delta^{\mathrm{d}}$. Moreover, we obviously have $|\mathcal{A}_\delta| = h^{-\mathrm{d}} \le 2^{\mathrm{d}}\delta^{-\mathrm{d}}$, and hence $(\mathcal{A}_\delta)_{\delta\in(0,1]}$ satisfies Assumption A with $c_{\mathrm{part}} := 2^{\mathrm{d}}$.*

The next lemma describes a general situation in which there exist partitions satisfying Assumption A. For its formulation, recall that the covering numbers of a compact metric space (X,d) are defined by

$$\mathcal{N}(X,d,\delta) := \min\Bigl\{ n \ge 1 : \exists x_1,\dots,x_n \in X \text{ with } X \subset \bigcup_{i=1}^n B(x_i,\delta)\Bigr\},\ \delta > 0,$$

where again $B(x,\delta)$ denotes the closed ball with center x and radius δ.

Lemma A.7.2. *Let (X,d) be a compact metric space for which there exist constants $c > 0$ and $\mathrm{d} > 0$ such that*

$$\mathcal{N}(X,d,\delta) \le c\delta^{-\mathrm{d}}, \qquad \delta \in (0,1/4].$$

Moreover, assume that there exists a finite measure μ on X such that

$$\mu(B(x,\delta)) \ge c^{-1}\delta^{\mathrm{d}}$$

for all $x \in X$ and $\delta \in (0,1/4]$. Then Assumption A is satisfied for d and $c_{\mathrm{part}} = 4^{\mathrm{d}}c$.

Note that the unit spheres $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ together with their surface measures satisfy the assumptions for $\mathrm{d} = d-1$, see also Corollary A.7.3.

Proof of Lemma A.7.2. Let us recall that a δ-packing in X is a family $y_1,\dots,y_m \in X$ with $d(y_i,y_j) > 2\delta$ for all $i \ne j$. Let us write

$$\mathcal{M}(X,d,\delta) := \max\Bigl\{ m \ge 1 : \exists\delta\text{-packing } y_1,\dots,y_m \text{ in } X\Bigr\}$$

for the size of the largest possible δ-packing in X. Then it is well-known that we have the following inequalities between these packing numbers and the covering numbers:

$$\mathcal{M}(X,d,\delta) \le \mathcal{N}(X,d,\delta) \le \mathcal{M}(X,d,\delta/2), \qquad \delta > 0. \tag{A.7.1}$$

Let us now fix a $\delta \in (0,1]$ and a maximal $\delta/4$-packing $y_1, \dots, y_m$ in X. By (A.7.1) we conclude that

$$m = \mathcal{M}(X, d, \delta/4) \leq \mathcal{N}(X, d, \delta/4) \leq 4^{\mathrm{d}} c \delta^{-\mathrm{d}}\,.$$

To construct the partition $\mathcal{A}_\delta$, we consider a Voronoi partition $A_1, \dots, A_m$ that corresponds to the points $y_1, \dots, y_m$, where the behavior of the cells on their boundary may be arbitrary, i.e. ties may be arbitrarily resolved. Our next goal is to show

$$B(y_i, \delta/4) \subset A_i \subset B(y_i, \delta/2)\,, \qquad\qquad i = 1, \dots, m. \tag{A.7.2}$$

To prove the left inclusion, we fix an $x \in B(y_i, \delta/4)$. For $j \neq i$, we then find

$$\delta/2 < d(y_i, y_j) \leq d(y_i, x) + d(x, y_j) \leq \delta/4 + d(x, y_j)\,,$$

and hence $d(x, y_j) > \delta/4 \geq d(x, y_i)$. From the latter we conclude that $x \in A_i$.

For the proof of the right inclusion, we assume that it does not hold for some index $i \in \{1, \dots, m\}$. Then there exists an $x \in A_i$ such that $d(x, y_i) > \delta/2$. On the hand, since $y_1, \dots, y_m$ is a maximal $\delta/4$-packing in X, there exists a $j \in \{1, \dots, m\}$ with $d(x, y_j) \leq 2\delta/4 = \delta/2$, and hence we have $d(x, y_j) \leq \delta/2 < d(x, x_i)$. This implies $x \notin A_i$, i.e. we have found a contradiction.

Now, using (A.7.2), we obtain both $\mu(A_i) \geq \mu(B(y_i, \delta/4)) \geq 4^{-\mathrm{d}} c^{-1} \delta^{\mathrm{d}}$ and $\operatorname{diam} A_i \leq \operatorname{diam} B(y_i, \delta/2) \leq \delta$. $\square$

The next corollary in particular shows that one of the assumptions made in Lemma A.7.2 can be omitted if the measure behaves regularly on balls.

Corollary A.7.3. *Let (X, d) be a compact metric space and μ be a finite measure on X for which there exists a constant $K \geq 1$ such hat*

$$K^{-1} \leq \frac{\mu(B(y, \delta))}{\mu(B(x, \delta))} \leq K\,, \qquad\qquad x, y \in X,\ \delta \in (0, 1/4].$$

If there exist constants $c > 0$ and $\mathrm{d} > 0$ such that

$$\mathcal{N}(X, d, \delta) \leq c\delta^{-\mathrm{d}}\,, \qquad\qquad \delta \in (0, 1/4],$$

then Assumption A is satisfied for d and $c_{\mathrm{part}} = 4^{\mathrm{d}} cK$. Similarly, if

$$\mu(B(x, \delta)) \geq c^{-1} \delta^{\mathrm{d}}\,, \qquad\qquad \delta \in (0, 1/8],$$

holds true, then Assumption A is satisfied for d and $c_{\mathrm{part}} = 8^{\mathrm{d}} cK$.

If $(X, d, \cdot)$ is a compact group with invariant metric d and μ is its Haar measure, then we have $K = 1$. Moreover, if $X \subset \mathbb{R}^d$ is a sufficiently smooth manifold and μ is its surface measure, then the corollary is also applicable.

Proof of Corollary A.7.3. To show the first assertion, we fix a $\delta \in (0, 1/4]$ and a minimal δ-net $x_1, \ldots, x_n$ of X. For an $x \in X$ we then obtain

$$1 = \mu(X) \le \sum_{i=1}^{n} \mu(B(x_i, \delta)) \le nK\mu(B(x, \delta)) \le cK\delta^{-\mathrm{d}}\mu(B(x, \delta)) \, .$$

By Lemma A.7.2 we thus obtain the first assertion.

To prove the second assertion we fix a $\delta \in (0, 1/4]$ and a maximal $\delta/2$-packing $y_1, \ldots, y_m$ of X. Then $B(y_i, \delta/2) \cap B(y_j, \delta/2) = \emptyset$ for $i \ne j$ implies

$$1 = \mu(X) \ge \sum_{i=1}^{m} \mu(B(y_i, \delta/2)) \ge mK^{-1}\mu(B(x, \delta/2)) \ge m2^{-\mathrm{d}}c^{-1}K^{-1}\delta^{\mathrm{d}} \, ,$$

and hence $\mathcal{N}(X, d, \delta) \le \mathcal{M}(X, d, \delta/2) = m \le 2^{\mathrm{d}}cK\delta^{-\mathrm{d}}$ by (A.7.1). Lemma A.7.2 then yields the second assertion. □

A.8. Material Related to Basic Properties of Histograms. The goal of this section is to establish the key inclusion (2.7) for our histogram-based approach. The material of this section is taken from [9].

Our first result shows that $h_{D,\delta}$ *uniformly* approximates its infinite-sample counterpart

$$h_{P,\delta}(x) := \sum_{j=1}^{m} \frac{P(A_j)}{\mu(A_j)} \cdot \mathbf{1}_{A_j}(x) \, , \qquad\qquad x \in X,$$

with high probability, where $\mathcal{A}_\delta = (A_1, \ldots, A_m)$ for a fixed $\delta > 0$.

Theorem A.8.1. *Let Assumption A be satisfied and P be a distribution on X. Then, for all $n \ge 1$, $\varepsilon > 0$, and $\delta > 0$, we have*

$$P^n\Big(\big\{D \in X^n : \|h_{D,\delta} - h_{P,\delta}\|_\infty \ge \varepsilon\big\}\Big) \le 2c_{\mathrm{part}} \exp\Big(-\mathrm{d}\ln\delta - \frac{2n\varepsilon^2\delta^{2\mathrm{d}}}{c_{\mathrm{part}}^2}\Big) \, .$$

In addition, if P is μ-absolutely continuous and there exists a bounded μ-density h of P, then, for all $n \ge 1$, $\varepsilon > 0$, and $\delta > 0$, we have

$$P^n\Big(D \in X^n : \|h_{D,\delta} - h_{P,\delta}\|_\infty \ge \varepsilon\Big) \le 2c_{\mathrm{part}} \exp\bigg(\ln\delta^{-\mathrm{d}} - \frac{3n\varepsilon^2\delta^{\mathrm{d}}}{c_{\mathrm{part}}(6\|h\|_\infty + 2\varepsilon)}\bigg) .$$

Proof of Theorem A.8.1. We fix an $A \in \mathcal{A}_\delta$ and write $f := \mu(A)^{-1}\mathbf{1}_A$. Then f is non-negative and our assumptions ensure $\|f\|_\infty \leq c_{\mathrm{part}}\delta^{-\mathrm{d}}$. Consequently, Hoeffding's inequality, see e.g. [4, Theorem 8.1], yields

$$P^n\biggl(\Bigl\{D \in X^n : \Bigl|\frac{1}{n}\sum_{i=1}^n f(x_i) - \mathbb{E}_P f\Bigr| < \varepsilon\Bigr\}\biggr) \geq 1 - 2\exp\Bigl(-\frac{2n\varepsilon^2\delta^{2\mathrm{d}}}{c_{\mathrm{part}}^2}\Bigr)$$

for all $n \geq 1$ and $\varepsilon > 0$, where we assumed $D = (x_1, \ldots, x_n)$. Furthermore, we have $\frac{1}{n}\sum_{i=1}^n f(x_i) = \mu(A)^{-1}D(A)$ and $\mathbb{E}_P f = \mu(A)^{-1}P(A)$. By a union bound argument and $|\mathcal{A}_\delta| \leq c_{\mathrm{part}}\delta^{-\mathrm{d}}$, we thus obtain

$$P^n\biggl(\Bigl\{D \in X^n : \sup_{A \in \mathcal{A}_\delta}\Bigl|\frac{D(A)}{\mu(A)} - \frac{P(A)}{\mu(A)}\Bigr| < \varepsilon\Bigr\}\biggr) \geq 1 - 2c_{\mathrm{part}}\delta^{-\mathrm{d}}\exp\Bigl(-\frac{2n\varepsilon^2\delta^{2\mathrm{d}}}{c_{\mathrm{part}}^2}\Bigr).$$

Since, for $x \in X$ and $A \in \mathcal{A}_\delta$ with $x \in A$, we have $h_{D,\delta}(x) = \mu(A)^{-1}D(A)$ and $h_{P,\delta}(x) = \mu(A)^{-1}P(A)$, we then find the first assertion.

To show the second inequality, we write $f := \mu(A)^{-1}(\mathbf{1}_A - P(A))$ for a fixed $A \in \mathcal{A}_\delta$. This yields $\mathbb{E}_P f = 0$, $\|f\|_\infty \leq c_{\mathrm{part}}\delta^{-\mathrm{d}}$, and

$$\mathbb{E}_P f^2 \leq \mu(A)^{-2}P(A) \leq \mu(A)^{-1}\|h\|_\infty \leq c_{\mathrm{part}}\delta^{-\mathrm{d}}\|h\|_\infty .$$

Consequently, Bernstein's inequality, see e.g. [4, Theorem 8.2], yields

$$P^n\biggl(\Bigl\{D \in X^n : \Bigl|\frac{1}{n}\sum_{i=1}^n f(x_i)\Bigr| < \varepsilon\Bigr\}\biggr) \geq 1 - 2\exp\biggl(-\frac{3n\varepsilon^2\delta^{\mathrm{d}}}{c_{\mathrm{part}}(6\|h\|_\infty + 2\varepsilon)}\biggr).$$

Using $\frac{1}{n}\sum_{i=1}^n f(x_i) = (D(A) - P(A))\mu(A)^{-1}$, the rest of the proof follows the lines of the proof of the first inequality. □

The next result specifies the vertical and horizontal uncertainty of a plug-in level set estimate $\{\hat{h} \geq \rho\}$, if $\hat{h}$ is a uniform approximation of $h_{P,\delta}$.

Lemma A.8.2. *Let Assumption A be satisfied, P be a μ-absolutely continuous distribution on X, and $\hat{h} : X \to \mathbb{R}$ be a function with $\|\hat{h} - h_{P,\delta}\|_\infty \leq \varepsilon$ for some $\varepsilon \geq 0$. Then, for all $\rho \geq 0$, the following statements hold:*

i) If P is upper normal at the level $\rho + \varepsilon$, then we have $M_{\rho+\varepsilon}^{-\delta} \subset \{\hat{h} \geq \rho\}$.
ii) If P is upper normal at the level $\rho - \varepsilon$, then we have $\{\hat{h} \geq \rho\} \subset M_{\rho-\varepsilon}^{+\delta}$.

Proof of Lemma A.8.2. *i).* We will show the equivalent inclusion $\{\hat{h} < \rho\} \subset (X \setminus M_{\rho+\varepsilon})^{+\delta}$. To this end, we fix an $x \in X$ with $\hat{h}(x) < \rho$. If $x \in X \setminus M_{\rho+\varepsilon}$, we immediately obtain $x \in (X \setminus M_{\rho+\varepsilon})^{+\delta}$, and hence we may restrict our considerations to the case $x \in M_{\rho+\varepsilon}$. Then, $\hat{h}(x) < \rho$ together with $\|\hat{h} - h_{P,\delta}\|_\infty \leq \varepsilon$ implies $h_{P,\delta}(x) \leq \hat{h}(x) + \varepsilon < \rho + \varepsilon$. Now let A be the unique cell of the partition $\mathcal{A}_\delta$ satisfying $x \in A$. The definition of $h_{P,\delta}$ together with the assumed $0 < \mu(A) < \infty$ then yields

$$\int_A h \, d\mu = P(A) = h_{P,\delta}(x)\mu(A) < (\rho + \varepsilon)\mu(A) \,, \tag{A.8.1}$$

where $h : X \to [0, \infty)$ is an arbitrary μ-density of P. Our next goal is to show that there exists an $x' \in (X \setminus M_{\rho+\varepsilon}) \cap A$. Suppose the converse, that is $A \subset M_{\rho+\varepsilon}$. Then the upper normality of P at the level $\rho + \varepsilon$ yields $\mu(A \setminus \{h \geq \rho + \varepsilon\}) \leq \mu(M_{\rho+\varepsilon} \setminus \{h \geq \rho + \varepsilon\}) = 0$, and hence we conclude that $\mu(A \cap \{h \geq \rho + \varepsilon\}) = \mu(A)$. This leads to

$$\int\limits_A h \, d\mu = \int\limits_{A \cap \{h \geq \rho+\varepsilon\}} h \, d\mu + \int\limits_{A \setminus \{h \geq \rho+\varepsilon\}} h \, d\mu = \int\limits_{A \cap \{h \geq \rho+\varepsilon\}} h \, d\mu \geq (\rho + \varepsilon)\mu(A) \,.$$

However, this inequality contradicts (A.8.1), and hence there does exist an $x' \in (X \setminus M_{\rho+\varepsilon}) \cap A$. This implies $d(x, X \setminus M_{\rho+\varepsilon}) \leq d(x, x') \leq \operatorname{diam} A \leq \delta$, i.e. we have shown $x \in (X \setminus M_{\rho+\varepsilon})^{+\delta}$.

ii). Let us fix an $x \in X$ with $\hat{h}(x) \geq \rho$. If $x \in M_{\rho-\varepsilon}$, we immediately obtain $x \in M_{\rho-\varepsilon}^{+\delta}$, and hence it remains to consider the case $x \in X \setminus M_{\rho-\varepsilon}$. Clearly, if $\rho - \varepsilon \leq 0$, this case is impossible, and hence we may additionally assume $\rho - \varepsilon > 0$. Then, $\hat{h}(x) \geq \rho$ together with $\|\hat{h} - h_{P,\delta}\|_\infty \leq \varepsilon$ yields $h_{P,\delta}(x) \geq \hat{h}(x) - \varepsilon \geq \rho - \varepsilon$. Now let A be the unique cell of the partition $\mathcal{A}_\delta$ satisfying $x \in A$. By the definition of $h_{P,\delta}$ and $\mu(A) > 0$ we then obtain

$$\int_A h \, d\mu = P(A) = h_{P,\delta}(x)\mu(A) \geq (\rho - \varepsilon)\mu(A) \,, \tag{A.8.2}$$

where $h : X \to [0, \infty)$ is an arbitrary μ-density of P. Next we show that there exists an $x' \in M_{\rho-\varepsilon} \cap A$. Suppose the converse holds, that is $A \subset X \setminus M_{\rho-\varepsilon}$. Then the assumed upper normality of P at the level $\rho - \varepsilon$ yields

$$\mu(M_{\rho-\varepsilon} \vartriangle \{h \geq \rho - \varepsilon\}) = 0 \,,$$

and thus we find $\mu((X \setminus M_{\rho-\varepsilon}) \vartriangle \{h < \rho - \varepsilon\}) = 0$ by $A \vartriangle B = (X \setminus A) \vartriangle (X \setminus B)$. Combining this with the assumed $A \subset X \setminus M_{\rho-\varepsilon}$, we obtain

$$\mu\big(A \setminus \{h < \rho - \varepsilon\}\big) \leq \mu\big((X \setminus M_{\rho-\varepsilon}) \setminus \{h < \rho - \varepsilon\}\big) = 0 \,,$$

and this implies

$$\int\limits_A h \, d\mu = \int\limits_{A \cap \{h < \rho-\varepsilon\}} h \, d\mu + \int\limits_{A \setminus \{h < \rho-\varepsilon\}} h \, d\mu = \int\limits_{A \cap \{h < \rho-\varepsilon\}} h \, d\mu < (\rho - \varepsilon)\mu(A) \,.$$

This contradicts (A.8.2), and hence there does exist an $x' \in M_{\rho-\varepsilon} \cap A$. This yields $d(x, M_{\rho-\varepsilon}) \leq d(x, x') \leq \operatorname{diam} A \leq \delta$, i.e. we have shown $x \in M^{+\delta}_{\rho-\varepsilon}$. $\square$

A.9. Proofs and Additional Material Related to the Consistency. In this section we prove Theorem 4.1. Furthermore, it contains additional material related to the assumptions made in that theorem.

LEMMA A.9.1. *Let (X, d) be a metric space, μ be a finite Borel measure on X, and $(A_\rho)_{\rho\in\mathbb{R}}$ be a decreasing family of closed subsets of X. For $\rho^* \in \mathbb{R}$, we write*

$$\dot{A}_{\rho^*} := \bigcup_{\rho>\rho^*} \mathring{A}_\rho \qquad \text{and} \qquad \hat{A}_{\rho^*} := \bigcup_{\rho>\rho^*} A_\rho\,.$$

Then we have

$$\dot{A}_{\rho^*} = \bigcup_{\rho>\rho^*} \bigcup_{\varepsilon>0} \bigcup_{\delta>0} A^{-\delta}_{\rho+\varepsilon}\,.$$

Moreover, the following statements are equivalent:

i) $\mu(\hat{A}_{\rho^*} \setminus \dot{A}_{\rho^*}) = 0$.
ii) *For all $\varepsilon > 0$, there exists a $\rho_\varepsilon > \rho^*$ such that, for all $\rho \in (\rho^*, \rho_\epsilon]$, we have $\mu(A_\rho \setminus \mathring{A}_\rho) \leq \varepsilon$.*

PROOF OF LEMMA A.9.1. To show the first equality, we observe that (A.3.1) implies

$$\bigcap_{\rho>\rho^*} \bigcap_{\varepsilon>0} \bigcap_{\delta>0} (X \setminus A_{\rho+\varepsilon})^{+\delta} = \bigcap_{\varepsilon>0} \bigcap_{\rho>\rho^*} \overline{X \setminus A_{\rho+\varepsilon}} = \bigcap_{\rho>\rho^*} \overline{X \setminus A_\rho}\,.$$

Moreover, every set $A \subset X$ satisfies $\overline{X \setminus A} = X \setminus \mathring{A}$, and hence we obtain

$$\bigcap_{\rho>\rho^*} \bigcap_{\varepsilon>0} \bigcap_{\delta>0} (X \setminus A_{\rho+\varepsilon})^{+\delta} = \bigcap_{\rho>\rho^*} \overline{X \setminus A_\rho} = \bigcap_{\rho>\rho^*} (X \setminus \mathring{A}_\rho) = X \setminus \bigcup_{\rho>\rho^*} \mathring{A}_\rho\,.$$

Therefore, by taking the complement we find

$$\begin{aligned}\bigcup_{\rho>\rho^*} \mathring{A}_\rho = X \setminus \Bigl(\bigcap_{\rho>\rho^*} \bigcap_{\varepsilon>0} \bigcap_{\delta>0} (X \setminus A_{\rho+\varepsilon})^{+\delta} \Bigr) &= \bigcup_{\rho>\rho^*} \bigcup_{\varepsilon>0} \bigcup_{\delta>0} \bigl(X \setminus (X \setminus A_{\rho+\varepsilon})^{+\delta}\bigr) \\ &= \bigcup_{\rho>\rho^*} \bigcup_{\varepsilon>0} \bigcup_{\delta>0} A^{-\delta}_{\rho+\varepsilon}\,.\end{aligned}$$

i) $\Rightarrow$ *ii)*. Let us fix an $\varepsilon > 0$. Since $\mathring{A}_\rho = \bigcup_{\rho' \geq \rho} \mathring{A}_{\rho'} \nearrow \dot{A}_{\rho^*}$ for $\rho \searrow \rho^*$, the σ-continuity of finite measures yields a $\rho_\varepsilon > \rho^*$ such that $\mu(\hat{A}_{\rho^*} \setminus \mathring{A}_\rho) \leq \varepsilon$ for all $\rho \in (\rho^*, \rho_\varepsilon]$. Using $A_\rho \subset \hat{A}_{\rho^*}$ for $\rho > \rho^*$, we then obtain the assertion $\mu(A_\rho \setminus \mathring{A}_\rho) \leq \mu(\hat{A}_{\rho^*} \setminus \mathring{A}_\rho) \leq \varepsilon$.

ii) $\Rightarrow$ *i)*. Let us fix an $\varepsilon > 0$. For $\rho \in (\rho^*, \rho_\varepsilon]$, we then have $\mathring{A}_\rho \subset \dot{A}_{\rho^*}$, and hence our assumption yields $\mu(A_\rho \setminus \dot{A}_{\rho^*}) \leq \varepsilon$. In other words, we have $\lim_{\rho \searrow \rho^*} \mu(A_\rho \setminus \dot{A}_{\rho^*}) = 0$. Moreover, we have $A_\rho \nearrow \hat{A}_{\rho^*}$ for $\rho \searrow \rho^*$, and hence the σ-continuity of μ yields $\lim_{\rho \searrow \rho^*} \mu(A_\rho \setminus \dot{A}_{\rho^*}) = \mu(\hat{A}_{\rho^*} \setminus \dot{A}_{\rho^*})$. $\square$

Lemma A.9.2. *Let $f : (0,1] \to (0,\infty)$ be a monotonously increasing function and $g : (0, f(1)] \to [0,1]$ be its generalized inverse, that is*

$$g(y) := \inf\bigl\{x \in (0,1] : f(x) \geq y\bigr\}, \qquad y \in (0,1].$$

Then we have $lim_{y \to 0^+} g(y) = 0$.

Proof of Lemma A.9.2. Let $(y_n) \subset (0, f(1)]$ be a sequence with $y_n \to 0$. For $n \geq 1$, we write $E_n := \{x \in (0,1] : f(x) \geq y_n\}$. Let us fix an $\varepsilon \in (0,1]$. Since f is strictly positive, we then find $f(\varepsilon) > 0$, and hence there exists an $n_0 \geq 1$ such that $f(\varepsilon) \geq y_n$ for all $n \geq n_0$. Thus, we have $\varepsilon \in E_n$ for all $n \geq n_0$, and from the latter we obtain $g(y_n) = \inf E_n \leq \varepsilon$ for such n. $\square$

Before we prove Theorem 4.1, let us briefly illustrate the additional assumption $\mu(\overline{A_i^* \cup A_2^*} \setminus (A_1^* \cup A_2^*)) = 0$. To this end, we fix a μ-density h of P. Then Lemma A.1.2 tells us that

$$A_i^* \cup A_2^* = \bigcup_{\rho > \rho^*} M_\rho \subset \bigcup_{\rho > \rho^*} \overline{\{h \geq \rho\}} \subset \overline{\bigcup_{\rho > \rho^*} \{h \geq \rho\}} = \overline{\{h > \rho^*\}}\,.$$

Using the normality in Assumption C, which implies (A.1.3), we then obtain

$$\begin{aligned}\mu\bigl(\overline{A_i^* \cup A_2^*} \setminus (A_1^* \cup A_2^*)\bigr) \leq \mu\bigl(\overline{\{h > \rho^*\}} \setminus \{h > \rho^*\}\bigr) &\leq \mu(\partial\{h > \rho^*\}) \\ &= \mu(\partial\{h \leq \rho^*\})\,.\end{aligned}$$

Consequently, the additional assumption is satisfied, if there exists a μ-density h of P such that $\mu(\partial\{h \leq \rho^*\}) = 0$. In this respect recall, that Lemma A.1.3 showed that P is normal, if, for all $\rho \in \mathbb{R}$, we have a μ-density h of P with $\mu(\partial\{h \geq \rho\}) = 0$.

Proof of Theorem 4.1. We fix an $\epsilon > 0$. For $n \geq 1$, $\tau := \tau_n$, and $\varepsilon := \varepsilon_n$, we define ε_n^* by the right hand-side of (3.4). Then, Lemma A.9.2 shows $0 < \varepsilon_n^* \leq \epsilon \wedge (\rho^{**} - \rho^*)/9$ for sufficiently large n. In addition, δ_n and

ε_n satisfy (3.2) for sufficiently large n by (4.1), and we also have $\delta_n \leq \delta_{\text{thick}}$ for sufficiently large n. Thus, there is an $n_0 \geq 1$ such that, for all $n \geq n_0$, the values ε_n, δ_n, τ_n and ε_n^* satisfy the assumptions of Theorem 3.1 and $\varepsilon_n^* \leq \epsilon$.

Let us now consider an $n \geq n_0$ and a data set $D \in X^n$ satisfying both the assertions *i) - v)* of Theorem A.6.2 and (2.10). By Theorem 3.1 and our previous considerations we then know that the probability P^n of D is not less than $1 - e^{-\varsigma}$. Now, part *i)* of Theorem A.6.2 yields $\rho_D^* - \rho^* \geq 2\varepsilon_n > 0$ and

$$\rho_D^* - \rho^* \leq \varepsilon_n^* + 5\varepsilon_n \leq 6\varepsilon_n^* \leq 6\epsilon\,,$$

i.e. we have shown the first convergence.

To prove the second convergence, we write A_i, $i = 1, 2$, for the two topologically connected components of $M_{\rho^{**}}$. For $\rho \in (\rho^*, \rho^{**}]$, we further define $A_\rho^i := \zeta_\rho(A_i)$, where $\zeta_\rho : \mathcal{C}(M_{\rho^{**}}) \to \mathcal{C}(M_\rho)$ is the CRM. In addition, we write $A_\rho^i := \emptyset$ for $\rho > \rho^{**}$ and $A_\rho^i := X$ for $\rho \leq \rho^*$. Let us first show

$$\mu(\hat{A}_{\rho^*}^i \setminus \dot{A}_{\rho^*}^i) = 0 \tag{A.9.1}$$

for $i = 1, 2$, where we used the notation of Lemma A.9.1. To this end, we fix an $\epsilon > 0$. Since P is lower and upper normal at every level $\rho \in [\rho^*, \rho^{**}]$ we find, for an arbitrary μ-density h of P,

$$\mu(\hat{M}_{\rho^*} \setminus \dot{M}_{\rho^*}) = \mu\big(\{h > \rho^*\} \setminus \dot{M}_{\rho^*}\big) = 0\,,$$

where we used (A.1.3), (A.1.4), and the notation of Lemma A.9.1. Lemma A.9.1 then shows that there exists a $\rho_\epsilon > \rho^*$ such that

$$\mu(M_\rho \setminus \mathring{M}_\rho) \leq \epsilon \tag{A.9.2}$$

for all $\rho \in (\rho^*, \rho_\epsilon]$, where we may assume without loss of generality that $\rho_\epsilon \leq \rho^{**}$. Let us now fix a $\rho \in (\rho^*, \rho_\epsilon]$. Then we obviously have $\mathring{A}_\rho^1 \cup \mathring{A}_\rho^2 \subset \mathring{M}_\rho$. To prove that the converse inclusion also holds, we pick an $x \in \mathring{M}_\rho$. Without loss of generality we may assume that $x \in A_\rho^1$. Since A_ρ^2 is closed and thus compact, we then have $\varepsilon := d(x, A_\rho^2) > 0$. Moreover, since $\mathring{M}_\rho$ is open, there exists a $\delta \in (0, \varepsilon)$ such that $B(x, \delta) \subset \mathring{M}_\rho$. This yields $B(x, \delta) \subset A_\rho^1 \cup A_\rho^2$, and by $d(x, A_\rho^2) > \delta$, we conclude that $B(x, \delta) \subset A_\rho^1$. This shows $x \in \mathring{A}_\rho^1$, and hence we indeed have $\mathring{M}_\rho = \mathring{A}_\rho^1 \cup \mathring{A}_\rho^2$. Now we use this equality to obtain

$$M_\rho \setminus \mathring{M}_\rho = \big(A_\rho^1 \setminus (\mathring{A}_\rho^1 \cup \mathring{A}_\rho^2)\big) \cup \big(A_\rho^2 \setminus (\mathring{A}_\rho^1 \cup \mathring{A}_\rho^2)\big) = (A_\rho^1 \setminus \mathring{A}_\rho^1) \cup (A_\rho^2 \setminus \mathring{A}_\rho^2)\,.$$

By (A.9.2), this implies $\mu(A_\rho^i \setminus \mathring{A}_\rho^i) \leq \epsilon$, and thus Lemma A.9.1 shows (A.9.1).

Let us now fix an $\epsilon > 0$ and a $\varsigma \geq 1$. By the equality of Lemma A.9.1 and the σ-continuity of finite measures there then exist $\delta_\epsilon > 0$, $\varepsilon_\epsilon > 0$, and $\rho_\epsilon \in (\rho^*, \rho^{**}]$ such that, for all $\varepsilon \in (0, \varepsilon_\epsilon]$, $\delta \in (0, \delta_\epsilon]$, $\rho \in (\rho^*, \rho_\epsilon]$, and $i = 1, 2$, we have $\mu(\dot{A}^i_{\rho^*} \setminus (A^i_{\rho+\varepsilon})^{-\delta}) \leq \epsilon$. Combining this with $A^*_i = \hat{A}^i_{\rho^*}$, which holds by the definition of the clusters A^*_i, and Equation (A.9.1) we then obtain

$$\mu\big(A^*_i \setminus (A^i_{\rho+\varepsilon})^{-\delta}\big) = \mu\big(\hat{A}^i_{\rho^*} \setminus (A^i_{\rho+\varepsilon})^{-\delta}\big) = \mu\big(\dot{A}^i_{\rho^*} \setminus (A^i_{\rho+\varepsilon})^{-\delta}\big) \leq \epsilon\,. \tag{A.9.3}$$

Moreover, our assumption $\mu(\overline{A^*_i \cup A^*_2} \setminus (A^*_1 \cup A^*_2)) = 0$ means $\mu(\overline{\hat{M}_{\rho^*}} \setminus \hat{M}_{\rho^*}) = 0$, and since by part *iii)* of Lemma A.3.1 we know that

$$\bigcap_{\delta>0}\Big(\bigcup_{\rho>\rho^*} M_\rho\Big)^{+\delta} = \overline{\bigcup_{\rho>\rho^*} M_\rho} = \overline{\hat{M}_{\rho^*}}$$

we find

$$\mu\Big(\Big(\bigcup_{\rho>\rho^*} M_\rho\Big)^{+\delta} \setminus \hat{M}_{\rho^*}\Big) \leq \epsilon$$

for all sufficiently small $\delta > 0$. From this it is easy to conclude that

$$\mu(M^{+\delta}_{\rho-\varepsilon} \setminus \hat{M}_{\rho^*}) \leq \epsilon \tag{A.9.4}$$

for all sufficiently small $\varepsilon > 0$, $\delta > 0$ and all $\rho > \rho^* + \varepsilon$. Without loss of generality, we may thus assume that (A.9.4) also holds for all $\varepsilon \in (0, \varepsilon_\epsilon]$, $\delta \in (0, \delta_\epsilon]$ and all $\rho > \rho^* + \varepsilon$.

For given $\tau := \tau_n$ and $\varepsilon := \varepsilon_n$ we now define ε^*_n by the right hand-side of (3.4). Then, Lemma A.9.2 shows $\varepsilon^*_n \to 0$, and hence we obtain $\varepsilon^*_n \leq \min\{\frac{\rho_\epsilon - \rho^*}{9}, \epsilon, \varepsilon_\epsilon\}$ for all sufficiently large n. In addition, δ_n and ε_n satisfy (3.2) for sufficiently large n by (4.1), and we also have $\varepsilon_n \leq \epsilon \wedge \varepsilon_\epsilon$ and $\delta_n \leq \delta_\epsilon \wedge \delta_{\mathrm{thick}}$ for sufficiently large n. Consequently, there exists an $n_0 \geq 1$ such that, for all $n \geq n_0$, the values ε_n, δ_n, τ_n and ε^*_n satisfy the assumptions of Theorem 3.1 as well as $\varepsilon_n \leq \epsilon \wedge \varepsilon_\epsilon$ and $\delta_n \leq \delta_\epsilon$.

Let us now consider an $n \geq n_0$ and a data set $D \in X^n$ satisfying both the assertions *i) - v)* of Theorem A.6.2 and (2.10). By Theorem 3.1 and our previous considerations we then know that the probability P^n of D is not less than $1 - e^{-\varsigma}$. Now, part *i)* of Theorem A.6.2 gives both $\rho^*_D \geq \rho^* + 2\varepsilon_n > \rho^* + \varepsilon_n$ and $\rho^*_D \leq \rho^* + \varepsilon^*_n + 5\varepsilon_n \leq \rho^* + 6\varepsilon^*_n \leq \rho_\epsilon$, and hence (A.9.3) and (A.9.4) hold for $\varepsilon := \varepsilon_n$, $\delta := \delta_n$, and $\rho := \rho^*_D$. Consequently, (2.10) shows

$$\begin{aligned}
\mu(B_1(D) \vartriangle A^*_1) + \mu(B_2(D) \vartriangle A^*_2) &\leq 2\mu\big(A^*_1 \setminus (A^1_{\rho+\varepsilon})^{-\delta}\big) + 2\mu\big(A^*_2 \setminus (A^2_{\rho+\varepsilon})^{-\delta}\big) \\
&\qquad + \mu\big(M^{+\delta}_{\rho-\varepsilon} \setminus \{h > \rho^*\}\big) \\
&\leq 4\epsilon + \mu\big(M^{+\delta}_{\rho-\varepsilon} \setminus \hat{M}_{\rho^*}\big) \\
&\leq 5\epsilon\,,
\end{aligned}$$

where in the second to last step we also used (A.1.4). □

A.10. Additional Material Related to Rates. In this section, the assumption made in Section 4 are discussed in some more detail.

Let us begin with the following lemma, which gives a sufficient condition for a non-trivial separation exponent.

LEMMA A.10.1. *Let $X \subset \mathbb{R}^d$ be compact and convex, $\|\cdot\|$ be some norm on $\mathbb{R}^d$, and P be a Lebesgue absolutely continuous distribution on X that can be clustered between the levels ρ^* and ρ^{**}. Assume that P has a continuous density h and that there exist constants $c > 0$ and $\theta \in (0, \infty)$ such that*

$$|h(x) - h(x')| \leq c\,\|x - x'\|^\theta \tag{A.10.1}$$

for all $x \in \{h \leq \rho^*\}$, $\rho \in (\rho^*, \rho^{**}]$, and $x' \in \partial_X M_\rho$, where $\partial_X M_\rho$ denotes the boundary of M_ρ in X. Then the clusters of P have separation exponent θ.

PROOF OF LEMMA A.10.1. Let $\varepsilon \in (0, \rho^{**} - \rho^*]$ and A_1 and A_2 be the connected components of $M_{\rho^*+\varepsilon}$. Since A_1 and A_2 are closed, they are compact, and hence there exist $x_1 \in A_1$ and $x_2 \in A_2$ with

$$a := \|x_1 - x_2\| = d(A_1, A_2)\,, \tag{A.10.2}$$

where we note that $A_1 \cap A_2 = \emptyset$ implies $a > 0$. For $t \in [0, 1]$, we now consider

$$x(t) := tx_1 + (1 - t)x_2\,.$$

Since X is convex, we note that $x(t) \in X$ for all $t \in [0, 1]$. Our first goal is to show that $x_i \in \partial_X M_{\rho^*+\varepsilon}$ for $i = 1, 2$. To this end, we assume the converse, e.g. $x_2 \in \mathring{M}_{\rho^*+\varepsilon}$. Then there exists an $\epsilon \in (0, a)$ with $B_X(x_2, \epsilon) \subset \mathring{A}_2$, where $B_X(x_2, \epsilon) := \{x \in X : \|x - x_2\| \leq \epsilon\}$. Now $\|x(\epsilon/a) - x_2\| = \epsilon$ implies $x(\epsilon/a) \in A_2$, while $\|x(\epsilon/a) - x_1\| = a - \epsilon$ shows $\|x(\epsilon/a) - x_1\| < d(A_1, A_2)$. Together this contradicts (A.10.2).

For what follows, let us now observe that $t \mapsto x(t)$ is a continuous map on $[0, 1]$, and since h is continuous, there exists a $t^* \in [0, 1]$ with $h(x(t^*)) = \min_{t\in[0,1]} h(x(t))$. Our next goal is to show that

$$h(x(t^*)) \leq \rho^*\,. \tag{A.10.3}$$

To this end, we assume the converse, that is $h(x(t^*)) > \rho^*$. Then there exists a $\delta \in (0, \varepsilon]$ such that $h(x(t)) > \rho^* + \delta$ for all $t \in [0, 1]$, and therefore an application of Lemma A.1.2 using the continuity of h yields $x(t) \in M_{\rho^*+\delta}$ for all $t \in [0, 1]$. In other words, x_1 and x_2 are path-connected in $M_{\rho^*+\delta}$, and since the connecting path is a straight line, it is easy to see that x_1 and x_2 are τ-connected for all $\tau > 0$. Let us pick a $\tau \le 3\tau^*(\delta) = \tau^*_{M_{\rho^*+\delta}}$. Since $|\mathcal{C}(M_{\rho^*+\delta})| = 2$, part *ii)* of Proposition A.2.10 then shows $\mathcal{C}(M_{\rho^*+\delta}) = \mathcal{C}_\tau(M_{\rho^*+\delta})$. Let $\tilde{A}_1$ and $\tilde{A}_2$ be the two topologically connected components of $M_{\rho^*+\delta}$. Our previous considerations then showed that $\tilde{A}_1$ and $\tilde{A}_2$ are also the two τ-connected components of $M_{\rho^*+\delta}$. Now, $\delta \le \varepsilon$ gives a CRM $\zeta : \mathcal{C}(M_{\rho^*+\varepsilon}) \to \mathcal{C}(M_{\rho^*+\delta})$, which is bijective, since P can be clustered between ρ^* and ρ^{**}. Without loss of generality we may thus assume that $\zeta(A_i) = \tilde{A}_i$ for $i = 1, 2$. This yields $x_i \in A_i \subset \tilde{A}_i$, i.e. x_1 and x_2 do not belong to the same τ-connected component of $M_{\rho^*+\delta}$. Clearly, this contradicts our observation that x_1 and x_2 are τ-connected, and hence (A.10.3) is proven.

Now assume without loss of generality that $t^* \in [1/2, 1)$. Since we have already seen that $x_1 \in \partial_X M_{\rho^*+\varepsilon}$, our assumption (A.10.1) and (A.10.3) yield

$$\big|h(x(t^*)) - h(x_1)\big| \le c \, \|x(t^*) - x_1\|^\theta .$$

In addition, Lemma A.1.2 shows $x_1 \in M_{\rho^*+\varepsilon} \subset \{h \ge \rho^* + \varepsilon\}$. Combining these estimates with (A.10.2) and $d(A_1, A_2) = \tau^*_{M_{\rho^*+\varepsilon}} = 3\tau^*(\varepsilon)$, we find

$$\begin{aligned} \rho^* + \varepsilon \le h(x_1) \le h(x(t^*)) + c \, \|x(t^*) - x_1\|^\theta &\le \rho^* + c \, \|x(t^*) - x_1\|^\theta \\ &\le \rho^* + c \, 2^{-\theta} d^\theta(A_1, A_2) \\ &= \rho^* + c \, (3/2)^{-\theta} \tau^*(\varepsilon)^\theta , \end{aligned}$$

and from the latter the assertion easily follows. $\square$

Note that (A.10.1) holds, if the density h in Lemma A.10.1 is actually θ-Hölder-continuous, and it is easy to see that the converse is, in general, not true. Moreover, using the inclusion $\partial_X M_\rho \subset \{h = \rho\}$ established in Lemma A.1.2, it is easy to check that (A.10.1) is *equivalent* to

$$\big|h(x) - \rho\big| \le c \, d(x, \partial_X M_\rho)^\theta \tag{A.10.4}$$

for all $x \in \{h \le \rho^*\}$ and $\rho \in (\rho^*, \rho^{**}]$. Note that a localized but two-sided version of this condition has been used in [8] for a level set estimator that is adaptive with respect to the Hausdorff metric.

Our next goal is to discuss the assumptions made in Theorem 4.7 in more detail. To this end, we need a couple of technical lemmata.

Lemma A.10.2. *Let $X \subset \mathbb{R}^d$ be compact and convex and d be a metric on X that is defined by a norm on $\mathbb{R}^d$. Then, we have*

$$d(x, \partial_X A) \le d(x, X \setminus A)$$

for all $A \subset X$ and $x \in \overline{A}$, where $\partial_X A$ denotes the boundary of A in X.

Proof of Lemma A.10.2. Before we begin with the proof we note that $\overline{B}^X = \overline{B}^{\mathbb{R}^d}$ for all $B \subset X$ since X is closed, i.e., taking the closure with respect to X or $\mathbb{R}^d$ is the same. Like in the statement of the lemma, we will thus omit the superscript. Let us now write $\delta := d(x, X \setminus A)$. Then there exists a sequence $(x_n) \subset X \setminus A$ such that $d(x, x_n) \to \delta$. Since X is assumed to be compact, so is $\overline{X \setminus A}$, and thus there exists an $x_\infty \in \overline{X \setminus A}$ such that $d(x, x_\infty) \leq \delta$. Obviously, it suffices to show $x_\infty \in \partial_X A$. Let us assume the converse. Since $\partial_X A = \overline{A} \cap \overline{X \setminus A}$, we then have $x_\infty \notin \overline{A}$, that is $x_\infty \in X \setminus \overline{A}$. Now, the latter set is open in X, and hence there exists an $\varepsilon > 0$ such that $B_X(x_\infty, \varepsilon) \subset X \setminus \overline{A}$, where $B_X(x_\infty, \varepsilon)$ denotes the closed ball in X that has center x_∞ and radius ε. This ε must satisfy $\varepsilon < \delta$, since otherwise we would find a contradiction to $x \in \overline{A}$ by $x \in B_X(x_\infty, \delta) \subset B_X(x_\infty, \varepsilon) \subset X \setminus \overline{A}$. For $t := \varepsilon/\delta \in (0, 1)$ we now define $x' := tx + (1-t)x_\infty$. The convexity of X implies $x' \in X$, and since d is defined by a norm, we have $d(x_\infty, x') = td(x, x_\infty) \leq \varepsilon$. Together, this yields $x' \in B_X(x_\infty, \varepsilon) \subset X \setminus \overline{A} \subset X \setminus A$. Consequently, $d(x, x') = (1-t)d(x, x_\infty) \leq (1-t)\delta < \delta$ implies $d(x, X \setminus A) < \delta$, which contradicts the definition of δ. $\square$

Lemma A.10.3. *Let $X \subset \mathbb{R}^d$ be compact and convex and d be a metric on X that is defined by a norm on $\mathbb{R}^d$. Then, for all $A \subset X$ and $\delta > 0$, we have*

$$A^{+\delta} \setminus A^{-\delta} \subset (\partial_X A)^{+\delta},$$

where the operations $A^{+\delta}$ and $A^{-\delta}$ as well as the boundary $\partial_X A$ are with respect to the metric space (X, d).

Proof of Lemma A.10.3. Let us fix an $x \in A^{+\delta} \setminus A^{-\delta} = A^{+\delta} \cap (X \setminus A)^{+\delta}$. If $x \in \overline{A}$, then Lemma A.10.2 immediately yields $d(x, \partial_X A) \leq d(x, X \setminus A) \leq \delta$, that is $x \in (\partial_X A)^{+\delta}$. It thus suffices to consider the case $x \notin \overline{A}$. Then we find $x \in X \setminus \overline{A} \subset X \setminus A \subset \overline{X \setminus A}$, and hence another application of Lemma A.10.2 yields $d(x, \partial_X(X \setminus A)) \leq d(x, A) \leq \delta$. Now the assertion easily follows from $\partial_X(X \setminus A) = \overline{X \setminus A} \cap \overline{X \setminus (X \setminus A)} = \overline{X \setminus A} \cap \overline{A} = \partial_X A$. $\square$

The next lemma shows that assuming an α-smooth boundary with $\alpha > 1$ does not make sense. It further shows that, for each level set with rectifiable boundary in the sense of [5, 3.2.14], the bound (4.9) holds with $\alpha = 1$.

Lemma A.10.4. *Let λ^d be the d-dimensional Lebesgue measure, $\mathcal{H}^{d-1}$ be the $(d-1)$-dimensional Hausdorff measure on $\mathbb{R}^d$, and σ_d be the volume of the d-dimensional unit Euclidean ball in $\mathbb{R}^d$. Then, for every non-empty, bounded, and measurable subset $A \subset \mathbb{R}^d$ the following statements hold:*

i) There exists a $\delta_A > 0$*, such that for* $\underline{c}_A := d\sigma_d^{1/d}\lambda^d(\overline{A})^{1-1/d}/2$ *and all* $\delta \in (0, \delta_A]$*, we have*

$$\lambda^d(A^{+\delta} \setminus A^{-\delta}) \geq \underline{c}_A \cdot \delta\,.$$

ii) If ∂A *is* $(d-1)$*-rectifiable and* $\mathcal{H}^{d-1}(\partial A) > 0$*, there exists a* $\delta_A > 0$*, such that, for all* $\delta \in (0, \delta_A]$*, we have*

$$\lambda^d\big(A^{+\delta} \setminus A^{-\delta}\big) \leq 4\mathcal{H}^{d-1}(\partial A) \cdot \delta\,.$$

Proof of Lemma A.10.4. Let us first recall that, for an integer $0 \leq m \leq d$, the upper and lower Minkowski content of a $B \subset \mathbb{R}^d$ is defined by

$$\mathcal{M}^{*m}(B) := \limsup_{\delta\to 0^+} \frac{\lambda^d(B^{+\delta})}{\sigma_{d-m}\delta^{d-m}}$$
$$\mathcal{M}^{m}_{*}(B) := \liminf_{\delta\to 0^+} \frac{\lambda^d(B^{+\delta})}{\sigma_{d-m}\delta^{d-m}}\,,$$

where σ_{d-m} denotes the λ^{d-m}-volume of the unit Euclidean ball in $\mathbb{R}^{d-m}$. It is easy to check that these definitions coincide with those in [5, 3.2.37].

i). Since in the case $\lambda^d(\overline{A}) = 0$ there is nothing to prove, we restrict our considerations to the case $\lambda^d(\overline{A}) > 0$. Now, A is bounded, and hence we have $\lambda^d(\overline{A}) < \infty$. The isoperimetric inequality [5, 3.2.43] thus yields

$$d\sigma_d^{1/d}\lambda^d(\overline{A})^{1-1/d} \leq \mathcal{M}^{d-1}_{*}(\partial A)\,,$$

and hence, there exists a $\delta_A > 0$, such that, for all $\delta \in (0, \delta_A]$, we have

$$\frac{d\sigma_d^{1/d}\lambda^d(\overline{A})^{1-1/d}}{2} \leq \frac{\lambda^d\big((\partial A)^{+\delta}\big)}{\sigma_1\delta} \leq \frac{\lambda^d\big(A^{+2\delta} \setminus A^{-2\delta}\big)}{2\delta}\,,$$

where in the last estimate we used part *viii)* of Lemma A.3.1 and $\sigma_1 = 2$.

ii). Since ∂A is closed and $(d-1)$-rectifiable in the sense of [5, 3.2.14], we find

$$\mathcal{M}^{*(d-1)}(\partial A) = \mathcal{H}^{d-1}(\partial A)$$

by [5, 3.2.39]. Moreover, since ∂A is bounded, the boundary is contained in a compact set $X \subset \mathbb{R}^d$ such that the relative boundary $\partial_X A$ of A in X equals ∂A and the sets $A^{+\delta}$ and $A^{-\delta}$ considered in X equal the sets $A^{+\delta}$ and $A^{-\delta}$ when considered in $\mathbb{R}^d$ for all $\delta \in (0, 1]$. By Lemma A.10.3 there thus exists a $\delta_A > 0$ such that

$$\frac{\lambda^d\big(A^{+\delta} \setminus A^{-\delta}\big)}{2\delta} \leq \frac{\lambda^d\big((\partial A)^{+\delta}\big)}{\sigma_1\delta} \leq 2\mathcal{H}^{d-1}(\partial A)$$

for all $\delta \in (0, \delta_A]$. □

The next lemma shows that a bound (4.9) together with a regular behavior of h around the level of interest ensures a non-trivial flatness exponent.

LEMMA A.10.5. *Let (X,d) be a complete, separable metric space, μ be a finite Borel measure on X with $\operatorname{supp}\mu = X$, and P be a μ-absolutely continuous distribution on X. Furthermore, let $\rho \geq 0$ be a level and h be a μ-density of P for which there exist constants $c > 0$, $\alpha \in (0,1]$, $\delta_0 > 0$, and $\theta \in (0,\infty)$ such that*

$$\mu(M_\rho^{+\delta} \setminus M_\rho^{-\delta}) \leq c\delta^\alpha \tag{A.10.5}$$

for all $\delta \in (0,\delta_0]$ and

$$d(x,\partial M_\rho)^\theta \leq c\,\big|h(x) - \rho\big| \tag{A.10.6}$$

for all $x \in \{h > \rho\}$. Then P has flatness exponent α/θ at level ρ.

PROOF OF LEMMA A.10.5. Let us fix an $s > 0$. For $x \in \{0 < h - \rho < s\}$ we then find $d(x,\partial M_\rho)^\theta \leq cs$ by (A.10.6), that is $x \in (\partial M_\rho)^{+\delta}$ for $\delta := (cs)^{1/\theta}$. Using part *viii)* of Lemma A.3.1, we conclude that $x \in M_\rho^{+2\delta} \setminus M_\rho^{-2\delta}$. In the case $2\delta \leq \delta_0$, we thus obtain

$$\mu\big(\{0 < h - \rho < s\}\big) \leq \mu\big(M_\rho^{+2\delta} \setminus M_\rho^{-2\delta}\big) \leq 2^\alpha c\,\delta^\alpha = 2^\alpha c^{1+\alpha/\theta} s^{\alpha/\theta}\,,$$

and since μ is a finite measure, it is then easy to see that we can increase the constant on the right-hand side so that it holds for all $s > 0$. □

Appendix B. Continuous Densities in two Dimensions. In this appendix, we present a couple of two-dimensional examples that show that the assumptions imposed in the paper are not only met by many discontinuous densities, but also by many continuous densities.

B.1. Single Two-Dimensional Sets. In this section we consider the operations $\oplus\delta$ and $\ominus\delta$ for a specific class of sets $A \subset \mathbb{R}^2$.

We begin with an example of a set $A \subset \mathbb{R}^2$, for which we can compute $A^{\oplus\delta}$ and $A^{\ominus\delta}$ explicitly. This example will be the base of all further examples.

EXAMPLE B.1.1. *Let $X := [-1,1] \times [-2,2]$ be equipped with the metric defined by the supremums norm. Furthermore, for $x_-^\pm \in (-0.6,-0.4)$ and $x_+^\pm \in (0.4,0.6)$ we fix two continuous functions $f^-, f^+ : [-1,1] \to [-1,1]$ such that f^+ is increasing on $[-1,x_-^+] \cup [0,x_+^+]$ and decreasing on $[x_-^+,0] \cup$*

$[x_+^+, 1]$, while f^- is decreasing on $[-1, x_-^-] \cup [0, x_+^-]$ and increasing on $[x_-^-, 0] \cup [x_+^-, 1]$. In addition, assume that $\{f^- < 0\} = \{f^+ > 0\}$ and $\{f^- = 0\} = \{f^+ = 0\}$ as well as $f^-(\pm 0.5) < 0$ and $f^+(\pm 0.5) > 0$. Now consider the (non-empty) set A enveloped by $f^\pm$, that is

$$A := \big\{(x, y) \in X : x \in [-1, 1] \text{ and } f^-(x) \le y \le f^+(x)\big\}\,.$$

To describe $A^{\ominus\delta}$ for $\delta \in (0, 0.1]$, we define $f^\pm_{-\delta} : [-1, 1] \to [-1, 1]$ by

$$f^\pm_{-\delta}(x) := \begin{cases} f^\pm(-1) & \text{if } x \in [-1, -1+\delta] \\ f^\pm(0) & \text{if } x \in [-\delta, +\delta] \\ f^\pm(1) & \text{if } x \in [1-\delta, 1] \end{cases}$$

and $f^-_{-\delta}(x) := f^-(x-\delta) \vee f^-(x+\delta)$, respectively $f^+_{-\delta}(x) := f^+(x-\delta) \wedge f^+(x+\delta)$ for the remaining $x \in [-1, 1]$. Then we have

$$A^{\ominus\delta} = \big\{(x, y) \in X : x \in [-1, 1] \text{ and } f^-_{-\delta}(x) + \delta \le y \le f^+_{-\delta}(x) - \delta\big\}\,.$$

Moreover, to describe $A^{\oplus\delta}$, we define

$$\begin{aligned} x_{0,-1} &:= \min\big\{x \in [-1, -0.5] : f^+(x) - f^-(x) \ge 0\big\} \\ x_{0,-0} &:= \max\big\{x \in [-0.5, 0] : f^+(x) - f^-(x) \ge 0\big\} \\ x_{0,+0} &:= \min\big\{x \in [0, 0.5] : f^+(x) - f^-(x) \ge 0\big\} \\ x_{0,+1} &:= \max\big\{x \in [0.5, 1] : f^+(x) - f^-(x) \ge 0\big\}\,, \end{aligned}$$

where the minima are attained by the continuity of $f^\pm$ and the fact that all sets are non-empty. Furthermore, we define $f^\pm_{+\delta} : [-1, 1] \to [-1, 1]$ by

$$f^\pm_{+\delta}(x) := \begin{cases} f^\pm(x+\delta) & \text{if } x \in [-1 \vee (x_{0,-1} - \delta), x_-^\pm - \delta] \\ f^\pm(x_-^\pm) & \text{if } x \in [x_-^\pm - \delta, x_-^\pm + \delta] \\ f^\pm(x_+^\pm) & \text{if } x \in [x_+^\pm - \delta, x_+^\pm + \delta] \\ f^\pm(x-\delta) & \text{if } x \in [x_+^\pm + \delta, (x_{0,+1} + \delta) \wedge 1] \end{cases}$$

as well as $f^-_{+\delta}(x) := f^-(x-\delta) \wedge f^-(x+\delta)$ and $f^+_{+\delta}(x) := f^+(x-\delta) \vee f^+(x+\delta)$ for $x \in [x_-^\pm + \delta, x_+^\pm - \delta] \setminus (x_{0,-0} + \delta, x_{0,+0} - \delta)$ and $f^\pm_{+\delta}(x) := -2\delta$ for the remaining $x \in [-1, 1]$. Then we have

$$A^{\oplus\delta} = \big\{(x, y) \in X : x \in [-1, 1] \text{ and } f^-_{+\delta}(x) - \delta \le y \le f^+_{+\delta}(x) + \delta\big\}\,.$$

Finally, we have $|\mathcal{C}(A)| \le 2$ with $|\mathcal{C}(A)| = 2$ if and only if $x_{0,-0} < x_{0,+0}$, and in the latter case we further have $\tau_A^* = x_{0,+0} - x_{0,-0}$.

Proof of Example B.1.1. Let us fix a $\delta \in (0, 1/10]$. To simplify notations, we further write $g^- := f^-_{-\delta} + \delta$ and $g^+ := f^+_{-\delta} - \delta$.

Proof of "$A^{\ominus\delta} \subset \dots$". By $A^{\ominus\delta} = X \setminus (X \setminus A)^{\oplus\delta}$ it suffices to show that

$$\big\{(x, y) \in X : x \in [-1, 1] \text{ and } \big(y < g^-(x) \text{ or } y > g^+(x)\big)\big\} \subset (X \setminus A)^{\oplus\delta} .$$

By symmetry, it further suffices to consider the case $x \geq 0$ and $y > g^+(x)$. Moreover, to show the inclusion above, it finally suffices to find $x' \in [-1, 1]$ and $y' \in [-2, 2]$ with $|x - x'| \leq \delta$, $|y - y'| \leq \delta$ and $y' > f^+(x')$. However, this task is straightforward. Indeed, we can always set $y' := (y + \delta) \wedge 2$, and if $x \in [0, \delta]$ then $x' := 0$ works, since $y' = (y + \delta) \wedge 2 > g^+(x) + \delta = f^+(0) = f^+(x')$, while for $x \in [1 - \delta, 1]$, the choice $x' := 1$ does by an analogous argument. Finally, if $x \in (\delta, 1 - \delta)$, we set $x' := x - \delta$ if $g^+(x) = f^+(x - \delta) - \delta$ and $x' := x + \delta$ if $g^+(x) = f^+(x + \delta) - \delta$.

Proof of "$A^{\ominus\delta} \supset \dots$". Again, it suffices to consider $x \geq 0$. Let us fix a y with $g^-(x) \leq y \leq g^+(x)$. Then, our goal is to show $(x, y) \notin (X \setminus A)^{\oplus\delta}$, i.e.,

$$\|(x, y) - (x', y')\|_\infty > \delta \tag{B.1.1}$$

for all $(x', y') \in X \setminus A$. In the following, we thus fix a pair $(x', y') \in X \setminus A$ for which (B.1.1) is not true and show that this leads to a contradiction. We begin by considering the case $x \in [0, \delta]$. Since (B.1.1) is not true, we find $|x - x'| \leq \delta$, and hence $x^\pm_- \leq x' \leq x^\pm_+$. Then, if $y' > f^+(x')$, this leads to

$$y \leq g^+(x) = f^+(0) - \delta \leq f^+(x') - \delta < y' - \delta ,$$

which contradicts the assumed $|y - y'| \leq \delta$. The case $y' < f^-(x')$ analogously leads to a contradiction. Now consider the case $x \in [1 - \delta, 1]$. Then $|x - x'| \leq \delta$ implies $x' \geq x^\pm_+$. Thus, $y' > f^+(x')$ leads to another contradiction by

$$y \leq g^+(x) = f^+(1) - \delta \leq f^+(x') - \delta < y' - \delta ,$$

and the case $y' < f^-(x')$ can be treated analogously. It thus remains to consider the case $x \in [\delta, 1 - \delta]$. Then $|x - x'| \leq \delta$ implies $x - \delta \leq x' \leq x + \delta$. For $x' \leq x^+_+$ we thus find $f^+(x - \delta) \leq f^+(x')$, while for $x' \geq x^+_+$ we find $f^+(x + \delta) \leq f^+(x')$. For $y' > f^+(x')$ we hence obtain a contradiction by

$$y \leq g^+(x) = (f^+(x - \delta) \wedge f^+(x + \delta)) - \delta \leq f^+(x') - \delta < y' - \delta ,$$

and, again, the case $y' < f^-(x')$ can be shown similarly.

Proof of "$A^{\oplus\delta} \subset \dots$". Let us fix a pair $(x, y) \in A^{\oplus\delta}$. Without loss of generality we restrict our considerations to the case $y \geq 0$ and $x \in [-1, 0]$. To show that $y \leq f^+_{+\delta}(x) + \delta$ we assume the converse, that is $y > f^+_{+\delta}(x) + \delta$.

Since $(x,y) \in A^{\oplus\delta}$ we then find $(x',y') \in A$ with $\|(x,y)-(x',y')\|_\infty \le \delta$. From the latter we infer that both $x-\delta \le x' \le x+\delta$ and

$$y' \ge y - \delta > f^+_{+\delta}(x)\,. \tag{B.1.2}$$

If $x \in [-1, -1 \vee (x_{0,-1}-\delta))$ we get a contradiction, since $(x',y') \in A$ implies $x \ge x'-\delta \ge x_{0,-1}-\delta$. Moreover, for $x \in [-1\vee(x_{0,-1}-\delta), x^+_- - \delta]$, we obtain

$$f^+_{+\delta}(x) = f^+(x+\delta) \ge f^+(x') \ge y'\,,$$

which contradicts (B.1.2). If $x \in [x^+_- - \delta, x^+_- + \delta]$ we get a contradiction from $f^+_{+\delta}(x) = f^+(x^+_-) \ge f^+(x') \ge y'$, and if $x \in [x^+_- + \delta, 0 \wedge (x_{0,-0}+\delta)]$ we have

$$f^+_{+\delta}(x) = f^+(x-\delta) \vee f^+(x+\delta) \ge f^+(x-\delta) \ge f^+(x') \ge y'$$

which again contradicts (B.1.2). Finally, if $x \in (0 \wedge x_{0,-0}+\delta, 0]$ we obtain a contradiction from $x > x_{0,-0}+\delta \ge x'+\delta$.

Proof of "$A^{\oplus\delta} \supset \dots$". Let us fix a pair $(x,y) \in X$ with $f^-_{+\delta}(x) - \delta \le y \le f^+_{+\delta}(x)+\delta$. Without loss of generality we again consider the case $y \ge 0$ and $x \in [-1,0]$, only. To show $(x,y) \in A^{\oplus\delta}$ we need to find a pair $(x',y') \in A$ with $\|(x,y)-(x',y')\|_\infty \le \delta$. Let us assume that we have found an x' with $|x-x'| \le \delta$ and $f(x') \ge y-\delta$. For y' defined by

$$y' := f(x') \wedge (y+\delta)$$

we then immediately obtain $y' \le y+\delta$. Moreover, if we actually have $y' = y+\delta$, then we obtain $|y-y'| \le \delta$, while in the case $y' < y+\delta$ we find $y' = f(x') \ge y-\delta$, that is again $|y-y'| \le \delta$. Thus, it suffices to find an x' with the properties above. To this end, we first observe that we can exclude the case $x \in [-1, -1\vee(x_{0,-1}-\delta))$, since for such x we have $0 \le y \le f^+_{+\delta}(x)+\delta = -\delta$. Analogously, we can exclude the case $x \in (0 \wedge (x_{0,-0}+\delta), 0]$. Now consider the case $x \in [-1 \vee (x_{0,-1}-\delta), x^+_- - \delta]$. For $x' := x+\delta$ we then have

$$f(x') = f(x+\delta) = f^+_{+\delta}(x) \ge y - \delta\,,$$

and hence x' satisfies the desired properties. Moreover, for $x \in [x^+_- - \delta, x^+_- + \delta]$ we define $x' := x^+_-$, which gives $|x-x'| \le \delta$. In addition, we again have $f(x') = f(x^+_-) = f^+_{+\delta}(x) \ge y-\delta$. Finally, let us consider the case $x \in [x^+_- + \delta, 0 \wedge (x_{0,-0}+\delta)]$. Let us first assume that $f(x-\delta) \ge f(x+\delta)$. For $x' := x-\delta$ we then obtain $f(x') = f(x-\delta) = f^+_{+\delta}(x) \ge y-\delta$. Analogously, if $f(x-\delta) \le f(x+\delta)$, then $x' := x+\delta$ has the desired properties.

Finally, $|\mathcal{C}(A)| \le 2$ is obvious, and so is the equivalence between $|\mathcal{C}(A)| = 2$ and $x_{0,-0} < x_{0,+0}$. In the latter case, $A_1 := \{(x,y) \in A : x \le x_{0,-0}\}$ and $A_2 := \{(x,y) \in A : x \ge x_{0,+0}\}$ are the two components of A, and from this it is easy to conclude that $\tau^*_A = x_{0,+0} - x_{0,-0}$. $\square$

Our next example shows how to estimate the function ψ_A^* for the sets considered in Example B.1.1

Example B.1.2. *Let us consider the situation of Example B.1.1. To simplify the presentation, let us additionally assume that the monotonicity of f^+ and f^- is actually strict and that A has sufficiently thick parts on both sides of the y-axis in the sense of*

$$[-0.8,-0.2]\cup[0.2,0.8]\subset\{f^-\leq -0.2\}\cap\{f^+\geq 0.2\}\,. \tag{B.1.3}$$

Note that, for all $\delta\in(0,0.1]$, this condition in particular ensures that $A^{\ominus\delta}$ contains open neighborhoods around the points $(-0.5,0)$ and $(0,0.5)$. Moreover, for $\delta\in[0,0.1]$ we define

$$\begin{aligned}
x_{\delta,-1} &:= \min\bigl\{x\in[-1,-0.8] : f^+(x)-f^-(x)\geq 2\delta\bigr\}\\
x_{\delta,-0} &:= \max\bigl\{x\in[-0.2,0] : f^+(x)-f^-(x)\geq 2\delta\bigr\}\\
x_{\delta,+0} &:= \min\bigl\{x\in[0,0.2] : f^+(x)-f^-(x)\geq 2\delta\bigr\}\\
x_{\delta,+1} &:= \max\bigl\{x\in[0.8,1] : f^+(x)-f^-(x)\geq 2\delta\bigr\}\,,
\end{aligned}$$

where we note that the minima and maxima are attained by (B.1.3) and the continuity of $f^\pm$. For the same reason we further have $x_{\delta,-1}<-0.8$, $x_{\delta,-0}>-0.2$, $x_{\delta,+0}<0.2$, and $x_{\delta,+1}>0.8$. Then, $f^+_{-\delta}$ has exactly two local maxima $x^+_{\delta,-}$ and $x^+_{\delta,+}$, satisfying $x^+_{\delta,-}\in[-1,0]$ and $x^+_{\delta,+}\in[0,1]$, and $f^-_{-\delta}$ has exactly two local minima $x^-_{\delta,-}$ and $x^-_{\delta,+}$, satisfying $x^-_{\delta,-}\in[-1,0]$ and $x^-_{\delta,+}\in[0,1]$. Moreover, for all $\delta\in(0,0.1]$ we have

$$\begin{aligned}
\psi_A^*(\delta)\leq \delta + \Bigl(&\max\bigl\{|x_{\delta,i}-x_{0,i}| : i\in\{-1,-0,+0,+1\}\bigr\}\\
&\vee \max\bigl\{|f^i(x^i_j)-f^i_{-\delta}(x^i_{\delta,j})| : i,j\in\{-,+\}\bigr\}\Bigr)\,.
\end{aligned}$$

The right hand-side of this inequality can be further estimated under some regularity assumptions. Indeed, if there exist $c>0$ and $\gamma\in(0,1]$ such that

$$|f^\pm(x^\pm_\pm)-f^\pm(x)|\leq c|x^\pm_\pm - x|^\gamma\,,\qquad\qquad x\in[x^\pm_\pm-0.1, x^\pm_\pm+0.1]\,, \tag{B.1.4}$$

then, for all $\delta\in(0,0.1]$, we can bound the second maximum by

$$\max\bigl\{|f^i(x^i_j)-f^i_{-\delta}(x^i_{\delta,j})| : i,j\in\{-,+\}\bigr\}\leq c\delta^\gamma\,.$$

In addition, if, for some $i\in\{-1,-0,+0,+1\}$, we write $2\delta_0:=f^+(x_{0,i})-f^-(x_{0,i})$, then $|x_{\delta,i}-x_{0,i}|=0$ for all $\delta\in(0,\delta_0]$, i.e. the corresponding term *in the first maximum disappears for these δ. If $\delta_0 < 0.1$, and we additionally assume, for example, that*

$$|f^{\pm}(x)| \geq c^{-1/\gamma}|x_{0,-1} - x|^{1/\gamma} \tag{B.1.5}$$

for all $x \in [x_{0,-1}, -0.8]$, then we have $|x_{\delta,-1} - x_{0,-1}| \leq c\delta^\gamma$ for all $\delta \in (\delta_0, 0.1]$. Combining these assumptions we obtain a variety of sets A satisfying $\psi_A^*(\delta) \leq (c+1)\delta^\gamma$ for all $\delta \in (0, 0.1]$, and these examples of sets can be even further extended by considering bi-Lipschitz transformations of X.

Before we can prove the assertions made in the example above, we need to establish the following technical lemma.

LEMMA B.1.3. *Let $x^* \in [2/5, 3/5]$ and $f : [0,1] :\to \mathbb{R}$ be a continuous function that is strictly increasing on $[0, x^*]$ and strictly decreasing on $[x^*, 1]$. For $\delta \in (0, 1/8]$ we define $f_{-\delta} : [0,1] \to \mathbb{R}$ by*

$$f_{-\delta}(x) := \begin{cases} f(0) & \text{if } x \in [0, \delta] \\ f(x-\delta) \wedge f(x+\delta) & \text{if } x \in [\delta, 1-\delta] \\ f(1) & \text{if } x \in [1-\delta, 1]\,. \end{cases}$$

Then there exists exactly one $x_\delta^* \in [0,1]$ such that $f_{-\delta}(x_\delta^*) \geq f_{-\delta}(x)$ for all $x \in [0,1]$. Moreover, we have $x_\delta^* \in (x^* - \delta, x^* + \delta)$ and x_δ^* is the only element $x \in [\delta, 1-\delta]$ that satisfies $f(x-\delta) = f(x+\delta)$. Finally, we have

$$f_{-\delta}(x) = \begin{cases} f(x-\delta) & \text{if } x \in [\delta, x_\delta^*] \\ f(x+\delta) & \text{if } x \in [x_\delta^*, 1-\delta]\,. \end{cases}$$

PROOF OF LEMMA B.1.3. We first show that there is an $x_0 \in (x^* - \delta, x^* + \delta)$ such that $f(x_0 - \delta) = f(x_0 + \delta)$. To this end, we observe $g : [x^* - \delta, x^* + \delta] \to \mathbb{R}$ defined by $g := f(\,\cdot\, - \delta) - f(\,\cdot\, + \delta)$ is continuous, and since $g(x^* - \delta) = f(x^* - 2\delta) - f(x^*) < 0$ and $g(x^* + \delta) = f(x^*) - f(x^* + 2\delta) > 0$, we find an $x_0 \in (x^* - \delta, x^* + \delta)$ such that $g(x_0) = 0$ by the intermediate value theorem.

Let us now show that $f(x-\delta) < f(x+\delta)$ for all $x \in [\delta, x_0]$ and $f(x-\delta) > f(x+\delta)$ for all $x \in [x_0, 1-\delta]$. Clearly, for $x \in [\delta, x^* - \delta]$, the strict monotonicity of f on $[0, x^*]$ yields $f(x-\delta) < f(x+\delta)$. Moreover, for $x \in (x^* - \delta, x_0)$, we have $f(x-\delta) < f(x_0 - \delta) = f(x_0 + \delta) < f(x+\delta)$ since $f(\,\cdot\, - \delta) : [x^* - \delta, x^* + \delta] \to \mathbb{R}$ is strictly increasing, while $f(\,\cdot\, + \delta) : [x^* - \delta, x^* + \delta] \to \mathbb{R}$ is strictly decreasing. This shows the assertion for $x \in [\delta, x_0]$, and the assertion for $x \in [x_0, 1-\delta]$ can be shown analogously.

Combining the two results above, we find that there exists exactly one $x_0 \in [\delta, 1-\delta]$ satisfying $f(x_0-\delta) = f(x_0+\delta)$, and for this x_0 we further know $x_0 \in (x^* - \delta, x^* + \delta)$. In addition, these results show

$$f_{-\delta}(x) = \begin{cases} f(x-\delta) & \text{if } x \in [\delta, x_0] \\ f(x+\delta) & \text{if } x \in [x_0, 1-\delta]\,. \end{cases}$$

Let us now return to global maximizers of $f_{-\delta}$. To this end, we first observe that the existence of a global maximum of $f_{-\delta}$ follows from the continuity of $f_{-\delta}$ and the compactness of $[0,1]$. Let us now fix an $x_\delta \in [0,1]$ at which this global maximum is attained by $f_{-\delta}$. We first observe that $x_\delta \in (\delta, 1-\delta)$. Indeed, if, e.g., we had $x_\delta \geq 1-\delta$, we would obtain $f(1) = f_{-\delta}(x_\delta) \geq f_{-\delta}(1-2\delta) = f(1-3\delta) \wedge f(1-\delta) = f(1-\delta) > f(1)$ using $1 - 3\delta > x^*$, and $x_\delta \leq \delta$ would similarly lead to a contradiction. We next show that we actually have $x_\delta \in [x^* - \delta, x^* + \delta]$. To this end, it suffices to show

$$x_\delta \geq x^* - \delta \qquad \Longleftrightarrow \qquad x_\delta \leq x^* + \delta\,. \tag{B.1.6}$$

To show one implication, assume that $x_\delta \geq x^* - \delta$. Since $f_{-\delta}$ attains its maximum at x_δ, we then obtain

$$f(x_\delta+\delta) \geq f(x_\delta-\delta) \wedge f(x_\delta+\delta) = f_{-\delta}(x_\delta) \geq f_{-\delta}(x^*+\delta) = f(x^*+2\delta)\,.$$

Now $x_\delta + \delta \leq x^* + 2\delta$ follows from the assumed $x_\delta + \delta \geq x^*$ and the strict monotonicity of f on $[x^*, 1]$. Analogously, $x_\delta \leq x^* + \delta \Rightarrow x_\delta \geq x^* - \delta$ can be shown, and hence (B.1.6) is indeed true.

Finally, we can prove the remaining assertion. To this end, we pick again an x_δ at which $f_{-\delta}$ attains its maximum. Then we have already seen that $x_\delta \in [x^*-\delta, x^*+\delta]$. Now observe that assuming $x_\delta < x_0$ leads to $f(x_\delta - \delta) < f(x_0-\delta) = f(x_0+\delta) < f(x_\delta+\delta)$ using $x_0, x_\delta \in [x^*-\delta, x^*+\delta]$, which in turn yields the contradiction

$$f_{-\delta}(x_\delta) = f(x_\delta-\delta) \wedge f(x_\delta+\delta) = f(x_\delta-\delta) < f(x_0-\delta) \wedge f(x_0+\delta) = f_{-\delta}(x_0)\,.$$

Analogously, we find a contradiction assuming $x_\delta > x_0$, and hence we have $x_\delta = x_0$. Consequently, x_δ is unique and solves $f(x-\delta) = f(x+\delta)$. $\square$

Proof of Example B.1.2. We first note that the existence and uniqueness of the local extrema is guaranteed by Lemma B.1.3. In addition, this lemma actually shows $x^+_{\delta,-} \in (x^+_- - \delta, x^+_- + \delta)$, $x^-_{\delta,-} \in (x^-_- - \delta, x^-_- + \delta)$, $x^+_{\delta,+} \in (x^+_+ - \delta, x^+_+ + \delta)$, and $x^-_{\delta,+} \in (x^-_+ - \delta, x^-_+ + \delta)$. Moreover, we have

$$\psi^*_A(\delta) = \sup_{z\in A} d(z, A^{-\delta}) \leq \sup_{z\in A} d(z, A^{\ominus\delta})$$

by $A^{-\delta} \subset A^{\ominus\delta}$. We will thus estimate $d(z, A^{\ominus\delta})$ for $z := (x, y) \in A$.

We begin with the case $x \in [-1, x_{\delta,-1}]$. For later purposes, note that the definition of A yields $x \geq x_{0,-1}$. By the monotonicity of $f^{\pm}$ on $[-1, -0.8+\delta]$ we further know $f^{\pm}_{\delta}(x+\delta) = f^{\pm}(x)$. We write $x' := x_{\delta,-1} + \delta$ and

$$y' := \begin{cases} f^-(x_{\delta,-1}) + \delta & \text{if } y \leq f^-(x_{\delta,-1}) + \delta \\ y & \text{if } y \in [f^-(x_{\delta,-1}) + \delta, f^+(x_{\delta,-1}) - \delta] \\ f^+(x_{\delta,-1}) - \delta & \text{if } y \geq f^+(x_{\delta,-1}) - \delta\,. \end{cases}$$

If $y \leq f^-(x_{\delta,-1})+\delta$, we then obtain $y \leq y'$ and $y' = f^-(x_{\delta,-1})+\delta \leq f^-(x)+\delta \leq y+\delta$, that is $|y-y'| \leq \delta$, and it is easy to check that the same is true in the two other cases. Consequently, we have $\|(x, y)-(x', y')\|_\infty = x_{\delta,-1}+\delta-x$, and our construction further ensures

$$y' \in [f^-(x_{\delta,-1}) + \delta, f^+(x_{\delta,-1}) - \delta] = [f^-_{-\delta}(x') + \delta, f^+_{-\delta}(x') - \delta]\,.$$

By Example B.1.1 we conclude $(x', y') \in A^{\ominus\delta}$, and from this we easily find

$$d(z, A^{\ominus\delta}) \leq \delta + x_{\delta,-1} - x \leq \delta + x_{\delta,-1} - x_{0,-1}\,. \tag{B.1.7}$$

To show that (B.1.7) is also true in the case $x \in [x_{\delta,-1}, -0.8+\delta]$, we first observe that the monotonicity of $f^{\pm}$ on $[-1, -0.8+2\delta]$ yields

$$f^+(x) - f^-(x) \geq f^+(x_{\delta,-1}) - f^-(x_{\delta,-1}) \geq 2\delta\,,$$

and consequently, we can define

$$y' := \begin{cases} f^-(x) + \delta & \text{if } y \leq f^-(x) + \delta \\ y & \text{if } y \in [f^-(x) + \delta, f^+(x) - \delta] \\ f^+(x) - \delta & \text{if } y \geq f^+(x) - \delta\,. \end{cases}$$

If $y \leq f^-(x) + \delta$ we then obtain $y \leq y'$ and $y' = f^-(x) + \delta \leq y + \delta$, that is $|y - y'| \leq \delta$, and again it is easy to check that the same is true in the two other cases. Writing $x' := x + \delta$, we thus have $\|(x, y) - (x', y')\|_\infty = \delta$. Moreover, the construction together with $f^{\pm}_{\delta}(x+\delta) = f^{\pm}(x)$ ensures

$$y' \in [f^-(x) + \delta, f^+(x) - \delta] = [f^-_{-\delta}(x') + \delta, f^+_{-\delta}(x') - \delta]\,,$$

and hence we find $(x', y') \in A^{\ominus\delta}$ by Example B.1.1. Thus, we have shown $d(z, A^{\ominus\delta}) \leq \delta \leq \delta+x_{\delta,-1}-x_{0,-1}$, i.e. (B.1.7) is true for all $x \in [-1, -0.8+\delta]$.

Now consider the case $x \in [-0.8+\delta, -0.2-\delta]$. Here, we will focus on the sub-case $y \geq 0$, since the subcase $y \leq 0$ can be treated analogously. For later purposes, note that we have $f^-(x\pm\delta) \le -2\delta$. Now, if $x \in [-0.8+\delta, x^+_{\delta,-}-\delta]$, we set $x' := x+\delta$ and $y' := y \wedge (f^+(x)-\delta)$. This gives $y' \le y$ and $y-\delta \le f^+(x)-\delta \le y'$, and hence we again have $\|(x,y)-(x',y')\|_\infty = \delta$. Moreover, our constructions together with Lemma B.1.3 ensures

$$y' \in [-\delta, f^+(x)-\delta] = [-\delta, f^+_{-\delta}(x')-\delta] \subset [f^-_{-\delta}(x')+\delta, f^+_{-\delta}(x')-\delta]\,,$$

that is $(x',y') \in A^{\ominus\delta}$, and hence (B.1.7) is true in this case, too. The next case, we consider, is $x \in [x^+_{\delta,-}-\delta, x^+_{\delta,-}+\delta]$. In this case we set $x' := x^+_{\delta,-}$ and $y' := y \wedge (f^+_{-\delta}(x^+_{\delta,-})-\delta)$. This implies

$$y' \in [-\delta, f^+_{-\delta}(x^+_{\delta,-})-\delta] \subset [f^-_{-\delta}(x')+\delta, f^+_{-\delta}(x')-\delta]\,,$$

and hence $(x',y') \in A^{\ominus\delta}$. We further have $|x-x'| \le \delta$ and, if $y \le f^+_{-\delta}(x^+_{\delta,-})-\delta$, we also have $|y-y'| = 0$. Conversely, if $y \ge f^+_{-\delta}(x^+_{\delta,-})-\delta$, we find

$$y \le f^+(x) \le f^+(x^+_-) = f^+(x^+_-) - (f^+_{-\delta}(x^+_{\delta,-})-\delta) + y'\,,$$

that is $|y-y'| \le \delta + f^+(x^+_-) - f^+_{-\delta}(x^+_{\delta,-})$. Combining the latter two cases, we therefore obtain $\|(x,y)-(x',y')\|_\infty \le \delta + f^+(x^+_-) - f^+_{-\delta}(x^+_{\delta,-})$, that is $d(z, A^{\ominus\delta}) \le \delta + f^+(x^+_-) - f^+_{-\delta}(x^+_{\delta,-})$. Since all remaining cases can be treated analogously, the proof of the general estimate of $\psi^*_A(\delta)$ is finished.

Now consider the additional assumptions of $f^\pm$. For example, assume

$$|f^+(x^+_-) - f^+(x)| \le c|x^+_- - x|^\gamma$$

for all $x \in [x^+_- - 0.1, x^+_- + 0.1]$. Lemma B.1.3 shows $x^+_{\delta,-} \in (x^+_- - \delta, x^+_- + \delta)$. Without loss of generality, we assume $x^+_{\delta,-} \in [x^+_-, x^+_- + \delta)$. Using Lemma B.1.3 and $x^+_{\delta,-} - \delta \in [x^+_- - \delta, x^+_-) \subset [x^+_- - 0.1, x^+_- + 0.1]$, we then obtain

$$\big|f^+(x^+_-) - f^+_{-\delta}(x^+_{\delta,-})\big| = \big|f^+(x^+_-) - f^+(x^+_{\delta,-}-\delta)\big| \le c\big|x^+_- - x^+_{\delta,-} + \delta\big|^\gamma \le c\delta^\gamma\,.$$

Now assume that, for e.g. $i := -1$, we have $\delta_0 > 0$. For $\delta \in (0, \delta_0]$ we then find $f^+(x_{0,-1}) - f^-(x_{0,-1}) \ge 2\delta$, and thus $x_{0,-1} = x_{\delta,-1} = -1$. Conversely, let $\delta \in (\delta_0, 0.1]$. Then we have $f^+(x_{0,-1}) - f^-(x_{0,-1}) < 2\delta$ and a simple application of the intermediate value theorem thus yields $f^+(x_{\delta,-1}) - f^-(x_{\delta,-1}) = 2\delta$. Using the additional assumption on $f^\pm$ around the point $x_{0,-1}$, we then find

$$\begin{aligned} 2c^{-1/\gamma}|x_{\delta,-1} - x_{0,-1}|^{1/\gamma} \le |f^-(x_{\delta,-1})| + |f^+(x_{\delta,-1})| &= f^+(x_{\delta,-1}) - f^-(x_{\delta,-1}) \\ &= 2\delta\,, \end{aligned}$$

that is $|x_{\delta,-1} - x_{0,-1}| \le c\delta^\gamma$. □

B.2. Continuous Densities. In this section we present a class of continuous densities on $\mathbb{R}^2$ that meet the assumptions made in the paper. The first example, which represents the main result of this supplement, shows that many continuous distributions satisfy our thickness assumption.

EXAMPLE B.2.1. *Let $X := [-1,1] \times [-2,2]$ be equipped with the metric defined by the supremums norm. Moreover, let P be a Lebesgue absolutely continuous distribution that has a continuous density h. Furthermore, assume that there exists a $\rho^{**} > 0$, such that, for all $\rho \in (0, \rho^{**}]$, the level set M_ρ is of the form considered in Example B.1.2. In addition, we assume that there is a constant $K \in (0,1)$ such that*

$$\big|h(x,y) - \rho^* - x^2 + y^2\big| \leq K(x^2 + y^2) \tag{B.2.1}$$

for some $\rho^* \in [0, \rho^{**})$ and all $(x,y) \in \{h > 0\} \cap \big([-0.2, 0.2] \times (-1.1, 1.1)\big)$. Moreover, assume that h is continuously differentiable on the sets

$$A_1 := \{h > 0\} \cap \Big(\big((-0.7,-0.3) \cup (0.3, 0.7)\big) \times \big((-1.1,-0.2) \cup (0.2, 1.1)\big)\Big)$$

$$A_2 := \{h > 0\} \cap \Big(\big((-1,-0.8) \cup (0.8, 1)\big) \times \big((-1.1, 0) \cup (0.2, 1.1)\big)\Big)$$

$$A_3 := \{h > 0\} \cap \left\{(x,y) \in X : x \in (-0.2, 0) \cup (0, 0.2) \text{ and } |y| < \sqrt{\frac{1+K}{1-K}}|x|\right\}$$

with $h_y := \frac{\partial h}{\partial y} \neq 0$ on A_1 and $h_x := \frac{\partial h}{\partial x} \neq 0$ on $A_2 \cup A_3$. Finally, assume that there is a constant $C > 0$ such that $|h_x| \leq C|h_y|$ on A_1 and $|h_y| \leq C|h_x|$ on $A_2 \cup A_3$. Then P has thick levels of order $\gamma = 1$ with $\delta_{\text{thick}} = 0.1$ and

$$c_{\text{thick}} = 1 + \max\left\{C, \sqrt{\frac{1+K}{1-K}}\right\}.$$

Moreover, P can be clustered between ρ^* and ρ^{**} and we have

$$\frac{2}{\sqrt{1-K}}\sqrt{\varepsilon} \leq \tau^*_{M_{\rho^*+\varepsilon}} \leq \frac{2}{\sqrt{1+K}}\sqrt{\varepsilon}\,, \qquad \varepsilon \in (0, \rho^{**} - \rho]. \tag{B.2.2}$$

PROOF OF EXAMPLE B.2.1. Since we consider the Lebesgue measure on X, we have $M_0 = X$. Moreover, we have $X^{-\delta} = X$ since we consider the operation in X, and from this, we immediately see $\psi^*_X(\delta) = 0$ for all $\delta > 0$. Consequently, there is nothing to prove for $\rho = 0$.

Let us now fix some $\rho \in (0, \rho^{**}]$. Moreover, let $f^\pm : [-1,1] \to [-1,1]$ be the two functions satisfying the assumptions of Example B.1.2 and

$$M_\rho = \big\{(x,y) \in X : x \in [-1,1] \text{ and } f^-(x) \leq y \leq f^+(x)\big\}.$$

We pick an $(x,y) \in M_\rho$ with $y = f^+(x)$ or $y = f^-(x)$. Then we find $(x,y) \in \partial M_\rho$, and thus we have $h(x,y) = \rho$ by Lemma A.1.2, that is $h(x, f^\pm(x)) = \rho$.

Our first goal is to verify (B.1.4). To this end, we solely focus without loss of generality to the case x_+^+ and f^+, since the other cases can be treated analogously. Let us fix an $x \in [x_+^+ - 0.1, x_+^+ + 0.1]$. Then we have $x \in (0.3, 0.7)$ and thus $f^+(x) \in (0.2, 1.1)$ by (B.1.3). Consequently, h is continuously differentiable in $(x, f^+(x))$. By the implicit function theorem and the previously shown $h(x', f^+(x')) = \rho$ for all $x' \in (0.3, 0.7)$ we then conclude that f^+ is continuously differentiable at x and

$$(f^+(x))' = -\left(\frac{\partial h}{\partial y}(x, f^+(x))\right)^{-1} \cdot \frac{\partial h}{\partial x}(x, f^+(x)) = \frac{h_x(x, f^+(x))}{h_y(x, f^+(x))}\,. \tag{B.2.3}$$

Using $|h_x| \le C|h_y|$ on A_1, we thus find $|(f^+(x))'| \le C$, and hence f^+ is Lipschitz continuous on $(0.3, 0.7)$ with Lipschitz constant smaller than or equal to C. This implies (B.1.4) with constant C and exponent $\gamma = 1$.

Now consider the endpoints $x_{0,\pm1}$, where again it suffices to consider one case, say $x_{0,-1}$, due to symmetry. Let us write $2\delta_0 := f^+(x_{0,-1}) - f^-(x_{0,-1})$. Then, if $\delta_0 \ge 0.1$, we have $|x_{\delta,-1} - x_{0,-1}| = 0$ for all $\delta \in (0, 0.1]$ by Example B.1.2, and hence it suffice to show (B.1.5) in the case $\delta_0 < 0.1$. Observing that it actually suffices to show (B.1.5) for all $x \in (x_{0,-1}, -0.8)$ by continuity, we begin by fixing such an x. By monotonicity we then have $0 < f^+(x) < f^+(0.8) < 1.1$, and hence h is continuously differentiable at $(x, f^+(x))$. The implicit function theorem and the previously shown $h(x', f^+(x')) = \rho$ for all $x' \in (x_{0,-1}, -0.8)$, then shows that f^+ is continuously differentiable at x and (B.2.3) holds. Using $|h_y| \le C|h_x|$ on A_2, we then find $|(f^+(x))'| \ge 1/C$, and the fundamental theorem of calculus thus yields

$$\big|f^+(x') - f^+(x)\big| = \left|\int_x^{x'} (f^+(t))' dt\right| \ge C^{-1}|x' - x|$$

for all $x, x' \in (x_{0,-1}, -0.8)$. Now, letting $x' \to x_{0,-1}$, we obtain

$$|f^+(x)| \ge f^+(x) - f^+(x_{0,-1}) = \big|f^+(x) - f^+(x_{0,-1})\big| \ge C^{-1}|x_{0,-1} - x|$$

for all $x \in (x_{0,-1}, -0.8)$, i.e. (B.1.5) holds with constant C and $\gamma = 1$.

Finally, let us consider the points $x_{0,\pm0}$, where yet another time, we only focus on one case, say $x_{0,+0}$. For $x \in [x_{0,+0}, 0.2]$, we then have

$$\rho = h(x, f^+(x)) \le \rho^* + (1+K)x^2 + (K-1)(f^+(x))^2\,, \tag{B.2.4}$$

that is $(f^+(x))^2 \le \frac{\rho^*-\rho}{1-K} + \frac{1+K}{1-K}x^2$. Analogously, we can find a lower bound on $(f^+(x))^2$, so that we end up having

$$(f^+(x))^2 \in \left[\frac{\rho^* - \rho}{1+K} + \frac{1-K}{1+K}x^2, \frac{\rho^* - \rho}{1-K} + \frac{1+K}{1-K}x^2\right], \tag{B.2.5}$$

and an analogue result holds for $(f^-(x))^2$. Again, our goal is to show an analogue of (B.1.5). To this end, we first consider the case $\rho \in (0, \rho^*]$. By (B.2.1), we then know that $h(0,0) = \rho^* \geq \rho$, and hence $f^+(0) \geq 0$. Analogously, we find $f^-(0) \leq 0$, which together implies $x_{0,+0} = 0$. Furthermore, for $x \in [x_{0,+0}, 0.2]$, (B.2.5) gives

$$f^+(x) \geq \sqrt{\frac{\rho^* - \rho}{1+K} + \frac{1-K}{1+K}x^2} \geq \sqrt{\frac{1-K}{1+K}}|x| = \sqrt{\frac{1-K}{1+K}}\,|x_{0,+0} - x|\,,$$

that is (B.1.5) holds with constant $\sqrt{\frac{1+K}{1-K}}$ and exponent $\gamma = 1$. Let us now consider the case $\rho \in (\rho^*, \rho^{**}]$. For $x \in (x_{0,+0}, 0.2)$, (B.2.5) then yields

$$f^+(x) \leq \sqrt{\frac{\rho^* - \rho}{1-K} + \frac{1+K}{1-K}x^2} < \sqrt{\frac{1+K}{1-K}}|x|\,,$$

and thus we find $(x, f^+(x)) \in A_3$. Consequently, h is continuously differentiable at $(x, f^+(x))$, and (B.2.3) holds. As for $x_{0,-1}$, we can then show that (B.1.5) holds with constant C and exponent $\gamma = 1$.

In order to show that P can be clustered between the levels ρ^* and ρ^{**}, we first note that the assumed continuity of h guarantees that P is normal by Lemma A.1.3. Let us now fix a $\rho \in (\rho^*, \rho^{**}]$. Since from (B.2.1) we infer that $h(0,0) = \rho^*$, we then obtain $(0,0) \notin M_\rho$. The latter implies $x_{0,-0} < 0 < x_{0,+0}$, where $x_{0,-0}$ and $x_{0,+0}$ are the points defined in Example B.1.2 for the set M_ρ. By Example B.1.1 we then see that $\mathcal{C}(M_\rho)| = 2$. Analogously, for $\rho \in [0, \rho^*]$, the equality $h(0,0) = \rho^*$ implies $x_{0,-0} = 0 = x_{0,+0}$, which shows $\mathcal{C}(M_\rho)| = 1$. Finally, the bijectivity of $\zeta : \mathcal{C}(M_{\rho^{**}}) \to \mathcal{C}(M_\rho)$ follows from the form of the connected components described in Example B.1.1.

Let us finally prove (B.2.2). To this end, we fix an $\varepsilon \in (0, \rho^{**} - \rho]$ and define $\rho := \rho^* + \varepsilon$. Then we have already observed that $x_{0,-0} < 0 < x_{0,+0}$, and hence $f^\pm(x_{0,\pm 0}) = 0$. For $x := x_{0,+0}$ we then obtain

$$\rho = h(x, f^+(x)) \leq \rho^* + (1+K)x^2$$

by (B.2.4), and applying some simple transformations we thus find $x_{0,+0} = x \geq \sqrt{\frac{\rho - \rho^*}{1+K}} = \sqrt{\frac{\varepsilon}{1+K}}$. For $x := x_{0,+0}$ we further have

$$\rho = h(x, f^+(x)) \geq \rho^* + (1-K)x^2\,,$$

and thus $x_{0,+0} \leq \sqrt{\frac{\varepsilon}{1-K}}$. Since analogous estimates can be derived for $x_{0,-0}$, the formula $\tau^*_{M_{\rho^*+\varepsilon}} = x_{0,+0} - x_{0,-0}$ found in Example B.1.1 gives (B.2.2). $\square$

The last example of this appendix shows that the distributions from the previous example have a smooth boundary.

Example B.2.2. *Let X and P be as in Example B.2.1. Then the clusters have an α-smooth boundary for $\alpha = 1$ and*

$$c_{\text{bound}} = 8\left(10 + C + \sqrt{\frac{1+K}{1-K}}\right).$$

Proof of Example B.2.2. Let us first consider the case $0 < \delta \leq 0.1$. To this end, we fix a $\rho \in (\rho^*, \rho^{**}]$. Without loss of generality, we only consider the connected component A with $x < 0$ for all $(x, y) \in A$. We know that $A^{+\delta/2} \setminus A^{-\delta/2} \subset A^{\oplus\delta} \setminus A^{\ominus\delta}$ and the latter two sets have been calculated in Example B.1.1. In the following, we will only estimate $\lambda^2(\{(x,y) : y \geq 0\} \cap A^{\oplus\delta} \setminus A^{\ominus\delta})$, the case $y \leq 0$ can be treated analogously. Our first intermediate result towards the desired estimate is

$$\begin{aligned}
\lambda^2\big([-1 \vee (x_{0,-1} - \delta), x_{\delta,-1}] \times [0,2] \cap A^{\oplus\delta} \setminus A^{\ominus\delta}\big) &\leq 2|(x_{0,-1} - \delta) - x_{\delta,-1}| \\
&\leq 2\delta + 2|x_{0,-1} - x_{\delta,-1}| \\
&\leq 2(1+C)\delta\,,
\end{aligned}$$

where in the last step we used that the proof of Example B.2.1 showed (B.1.5) for $c = C$ and $\gamma = 1$. Moreover, we have

$$\begin{aligned}
\lambda^2\big([x_{\delta,-1}, x_-^+ - \delta] \times [0,2] \cap A^{\oplus\delta} \setminus A^{\ominus\delta}\big) &= \int_{x_{\delta,-1}}^{x_-^+ - \delta} f^+(x+\delta) - f^+(x-\delta) + 2\delta \, dx \\
&\leq 2\delta + \int_{x_-^+ - \delta}^{x_-^+ + \delta} f(x)\, dx \\
&\leq 4\delta
\end{aligned}$$

and analogously we obtain $\lambda^2\big([x_-^+ + \delta, x_{\delta,-0}] \times [0,2] \cap A^{\oplus\delta} \setminus A^{\ominus\delta}\big) \leq 4\delta$. In addition, we easily find $\lambda^2\big([x_-^+ - \delta, x_-^+ + \delta] \times [0,2] \cap A^{\oplus\delta} \setminus A^{\ominus\delta}\big) \leq 4\delta$ and finally, we have

$$\begin{aligned}
\lambda^2\big([x_{\delta,-0}, 0 \wedge (x_{0,-0} + \delta)] \times [0,2] \cap A^{\oplus\delta} \setminus A^{\ominus\delta}\big) &\leq 2\big|x_{\delta,-0} - x_{0,-0} - \delta\big| \\
&\leq 2\delta + 2\sqrt{\frac{1+K}{1-K}}\delta\,,
\end{aligned}$$

where we used that the proof of Example B.2.1 showed (B.1.5) for $c = \sqrt{\frac{1+K}{1-K}}$ and $\gamma = 1$. Combining all these estimates we obtain

$$\lambda^2\big([-1,0] \times [0,2] \cap A^{\oplus\delta} \setminus A^{\ominus\delta}\big) \leq 4\left(6 + C + \sqrt{\frac{1+K}{1-K}}\right)\delta$$

for all $\delta \in (0, 0.05]$. Moreover, for $\delta \in [0.05, 1]$ we easily obtain

$$\lambda^2\big([-1,0] \times [0,2] \cap A^{\oplus\delta} \setminus A^{\ominus\delta}\big) \leq 2 \leq 40\delta\,.$$

Combining both estimates and adding the case $y \leq 0$, we then obtain the assertion. □

References.

[1] A. Cuevas. Set estimation: another bridge between statistics and geometry. *Boletín de Estadística e Investigación Operativa*, 25:71–85, 2009.

[2] A. Cuevas, M. Febrero, and R. Fraiman. Estimating the number of clusters. *Canad. J. Statist.*, 28:367–382, 2000.

[3] A. Cuevas and R. Fraiman. A plug-in approach to support estimation. *Ann. Statist.*, 25:2300–2312, 1997.

[4] L. Devroye, L. Györfi, and G. Lugosi. *A Probabilistic Theory of Pattern Recognition.* Springer, New York, 1996.

[5] H. Federer. *Geometric Measure Theory.* Springer-Verlag, Berlin, 1969.

[6] P. Rigollet. Generalization error bounds in semi-supervised classification under the cluster assumption. *J. Mach. Learn. Res.*, 8:1369–1392, 2007.

[7] J. Serra. *Image Analysis and Mathematical Morphology*, volume 1. Academic Press, San Diego, 1982.

[8] A. Singh, C. Scott, and R. Nowak. Adaptive Hausdorff estimation of density level sets. *Ann. Statist.*, 37:2760–2782, 2009.

[9] I. Steinwart. Adaptive density level set clustering. In S. Kakade and U. von Luxburg, editors, *JMLR Workshop and Conference Proceedings Volume 19: Proceedings of the 24th Conference on Learning Theory 2011*, pages 703–738, 2011.

[10] G. Walther. Granulometric smoothing. *Ann. Statist.*, 25:2273–2299, 1997.

[11] G. Walther. On a generalization of Blaschke's rolling theorem and the smoothing of surfaces. *Math. Methods Appl. Sci.*, 22:301–316, 1999.