HAL open science

Mass localization

Thibaut Le Gouic

To cite this version:

Thibaut Le Gouic. Mass localization. 2015. hal-01163389

HAL Id: hal-01163389

https://hal.science/hal-01163389v1

Preprint submitted on 12 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mass localization

Thibaut Le Gouic

June 12, 2015

Abstract

For a given class $\mathcal{F}$ of closed sets of a measured metric space (E, d, μ), we want to find the smallest element B of the class $\mathcal{F}$ such that $\mu(B) \geq 1 - \alpha$, for a given $0 < \alpha < 1$. This set B *localizes the mass* of μ. Replacing the measure μ by the empirical measure μ_n gives an empirical smallest set B_n. The article introduces a formal definition of small sets (and their size) and study the convergence of the sets B_n to B and of their size.

Contents

1 Introduction **2**
1.1 Definitions . . . 4
1.1.1 Stable set . . . 4
1.1.2 Size function . . . 6
1.2 Overview of the main result . . . 6

2 First properties **7**
2.1 Existence . . . 7
2.2 Regularity of τ . . . 8

3 Consistency **8**
3.1 τ-tightness . . . 8
3.2 τ-consistency . . . 9
3.3 Minimizer consistency . . . 10
3.4 Minimizer continuity . . . 11

4 Examples **12**
4.1 Examples of stable classes . . . 12
4.1.1 Closed sets . . . 12
4.1.2 Parametrized classes . . . 12
4.1.3 ε-separated unions . . . 12
4.2 Examples of size function . . . 12
4.2.1 Packing . . . 12
4.3 Examples of sequence of measures . . . 13

5 Proofs **14**

1 Introduction

The framework of our study is a measured metric space (E, d, μ). Mass localization intends to find in this setting a small Borel set B such that $\mu(B) \geq 1-\alpha$ for some given $0 < \alpha < 1$. The measure μ conditioned on B is a new measure that we say to be α-localized and denote μ_α. This article provides a definition of a *smallest* Borel set of probability $1-\alpha$ in order to obtain a localized version of the measure with the *smallest* support possible.

This *smallest* Borel set represents intuitively the "essential part" of the measure. However, it seems difficult to give an universal definition of "*smallest*": although a ball centered on the origin as the smallest set with standard Gaussian measure on $\mathbb{R}^d$ seems a good choice, it is not obvious to define such set if the measure is not unimodal or if it is not symmetric or if it is not even defined on an Euclidian space.

Consistency is an important property we want for our notion. In statistics, the measure μ often unknown, is usually approximated by a sequence of probability measures $(\mu_n)_{n\geq 1}$. The *smallest* closed set with μ_n -probability $1-\alpha$ should become closer to the *smallest* one of μ-probability $1-\alpha$ as n grows.

Several methods have been studied in order to define such sets.

A first method is to choose a class $\mathcal{F}$ of subsets of E partially ordered by their volume and to pick the smallest set (for this order) of this class with a μ-probability greater than $1-\alpha$. This set corresponds to the level sets of a density function f whenever μ is absolutely continuous with respect to the Lebesgue measure and the class $\mathcal{F}$ contains the level sets. An other way to define this set is to maximize

$$\mu(B) - \beta\lambda(B), \tag{1}$$

over $B \in \mathcal{F}$, where λ is the Lebesgue measure and $\mu(\{f \geq \beta\}) = 1-\alpha$. This notion is known as *excess mass*. Denote by B_β the maximizer of (1) and by B_β^n the maximizer of

$$\mu_n(B) - \beta\lambda(B),$$

for μ_n the empirical measure. It is then of interest to determine if B_β^n converges to B_β and to exhibit a rate of convergence in this case.

The article [Har87] considers the case of $\mathcal{F}$ being the set of all convex sets of $\mathbb{R}^2$ and proves that the Hausdorff distance $d_H(B_n, B)$ between B_β^n and B_β converges to 0 and satisfies

$$d_H(B_\beta^n, B_\beta) = O\left(\frac{\log n}{n}\right)^{2/7}.$$

The article [Nol91] considers sets $\mathcal{F}$ as the set of all ellipsoids. Consistency of B_β^n is proven, as well as the following limit theorem. Let c_n and c be the centers of the ellipsoids B_β^n and B_β respectively, and let σ_n and σ be the vector containing the entries of the matrix defining the ellipsoids B_β^n and B_β respectively, then, if the level sets of the measure μ are ellipsoids,

$$n^{1/3}(c_n - c, \sigma_n - \sigma)$$

is weakly converging to the maximum of a Gaussian process. [Pol97] studies a more general case, with a different notion of convergence, and showed the consistency of B_β^n for the pseudo-distance

$$d_\mu(F,G) = \mu(F \triangle G),$$

where $\triangle$ denotes the symmetric difference, whenever the class $\mathcal{F}$ is a Glivenko-Cantelli class. Under several hypotheses including that the level sets of the measure μ belongs to $\mathcal{F}$ and regularity conditions on μ, the article obtains the following rate of convergence

$$d_\mu(B_\beta, B_\beta^n) = O(n^{-\delta}),$$

for a constant δ depending on the regularity of μ. This excess mass approach leads to rather precise results in many cases. However, it comes with few drawbacks, such as the condition that $\mathcal{F}$ must contains the level sets of the unknown measure μ, which requires a certain knowledge on the μ. Requirements on the regularity of μ can also be unsatisfactory for some applications. Also, this approach is restricted to the case of spaces with finite dimension (and often $\mathbb{R}^d$).

A second method comes from the notion of trimming on $\mathbb{R}$ extended to $\mathbb{R}^d$. On $\mathbb{R}$, the smallest set of μ-probability $1-\alpha$ is defined as

$$[F^{-1}(\alpha/2); F^{-1}(1-\alpha/2)],$$

where F is the cumulative distribution function of μ. Replacing F by the empirical cumulative distribution function F_n defines the empirical smallest set. Extension to $\mathbb{R}^d$ can be done in the following way: C_α denotes the intersection of all the closed half spaces of μ-probability greater than $1-\alpha$. C_α is then a non-empty convex set for $\alpha < 1/2$, if the measure μ is regular enough. [Nol92] deals with the rate of convergence of C_n, defined similarly with the empirical measure μ_n and shows its consistency. In order to quantify the rate of convergence of C_n to C_α, the article introduce the following random functions

$$r_n(u) = \inf\{r \geq 0; ru \notin C_n\},$$

and

$$r_\alpha(u) = \inf\{r \geq 0; ru \notin C_\alpha\},$$

and establishes the weak convergence to a Gaussian process defined on the unit sphere $\mathbb{S}^{d-1}$ of the process

$$\sqrt{n}(r_n - r_\alpha),$$

under regularity conditions on the density function of μ.

The article [CAGM97] presents another method, called α-trimmed k-means, which introduces very few arbitrary parameters. This method chooses the support of the α-localized measure ν as the one minimizing the distortion to its best k-quantifier. Formally, for a given function Φ, and a given integer k, the method consists in choosing

$$B_\alpha \in \arg\min \left\{ \inf_{\{m_1,\ldots,m_k\}\subset\mathbb{R}^d} \int_B \Phi\left(\inf_{1\leq i\leq k} \|X - m_i\|\right) d\mu; \mu(B) \geq 1-\alpha \right\}.$$

After proving the existence of such minimizer, the article [CAGM97] shows the consistency of B_α: if $(\mu_n)_{n\geq 1}$ weakly converges to an absolutely continuous measure μ then, for any choice of

$$B_\alpha^n \in \arg\min \left\{ \inf_{\{m_1,\ldots,m_k\}\subset\mathbb{R}^d} \int_B \Phi\left(\inf_{1\leq i\leq k} \|X - m_i\|\right) d\mu_n; \mu_n(B) \geq 1-\alpha \right\},$$

the sequence $(B_\alpha^n)_{n\geq 1}$ converges to B_α (when unique) for the Hausdorff metric. Theses results hold on $\mathbb{R}^d$.

The main goal of our article is to provide a new definition that is intuitive and avoid usual hypotheses, that remains consistent.

1.1 Definitions

We define a notion of *smallest* closed set and introduce some properties that will help to understand its meaning. The framework of the definition aims to be fairly general. (E, d) is a Polish space (metric, separable and complete space) and μ is a Borel measure on (E, d). A *smallest* set will be defined as the minimizer of a function τ defined on a class $\mathcal{F}$ of closed subsets of E.

1.1.1 Stable set

In order to ensure the existence of the smallest set in a class $\mathcal{F}$ of sets, the class needs to be stable in some way. The following definition of such stability will be an assumption made on the class. Let first set the following notation.

For a given set B and $\varepsilon > 0$, the set B^ε is the **ε-neighborhood** of B:

$$B^\varepsilon := \{x \in E; \exists y \in B, d(x, y) < \varepsilon\}.$$

Definition 1 (Stable set)**.** *Let $(B_n)_{n\geq 1}$ be a sequence of closed sets, denote $\lim_n B_n$ the set*

$$\lim_n B_n := \bigcap_{\varepsilon>0} \bigcup_{k\geq 1} \bigcap_{n\geq k} B_n^\varepsilon.$$

*Let $\mathcal{F}$ be a class of closed sets of E. $\mathcal{F}$ is **stable** if $E \in \mathcal{F}$ and*

$$(B_n)_{n\geq 1} \subset \mathcal{F} \implies \exists (n_k)_{k\geq 1}, n_k \to \infty, \lim_k B_{n_k} \in \mathcal{F}.$$

This notion of stability is close to the completeness under Hausdorff convergence. Indeed, it is strictly equivalent if the metric space (E, d) is compact, as it will be discussed in the next remarks.

As we defined $\mathcal{F}$ as a subset of the closed sets of (E, d), we first check that our notion of stability makes sense for a class of closed sets.

Remark 2. *Given a sequence of sets $(B_n)_{n\geq 1}$, $\lim_n B_n$ is always closed. Indeed, denote $B(x, \varepsilon/2)$ the ball centered in x of radius $\varepsilon/2$,*

$$x \notin \lim_n B_n \Leftrightarrow \exists \varepsilon > 0, \forall k \geq 1, \exists n \geq k, x \notin B_n^\varepsilon \tag{2}$$

$$\Rightarrow \exists \varepsilon > 0, \forall k \geq 1, \exists n \geq k, B(x, \varepsilon/2) \cap B_n^{\varepsilon/2} = \emptyset \tag{3}$$

$$\Rightarrow \exists \varepsilon > 0, B(x, \varepsilon/2) \cap \lim_n B_n = \emptyset. \tag{4}$$

In other words, $(\lim_n B_n)^c$ is open, and $\lim_n B_n$ is thus closed.

The following remark aims to clarify stability.

Remark 3. *When $(B_n)_{n\geq 1}$ is converging to B_∞ for the Hausdorff metric, then*

$$\lim_n B_n = B_\infty.$$

Indeed, denote ε_k the smallest $\varepsilon > 0$ such that $B_n \subset B_\infty^\varepsilon$ and $B_\infty \subset B_n^\varepsilon$ for all $n \geq k$, then,

$$B_\infty = \bigcap_{\varepsilon>0} \bigcup_{k\geq 1} \bigcap_{n\geq k} B_\infty \subset \bigcap_{\varepsilon>0} \bigcup_{k\geq 1} \bigcap_{n\geq k} B_n^{\varepsilon_k} = \lim_n B_n \subset \bigcap_{\varepsilon>0} \bigcup_{k\geq 1} \bigcap_{n\geq k} B_\infty^{\varepsilon+\varepsilon_k} = B_\infty.$$

In a more general setting, given a sequence of closed balls $(K_k)_{k\geq 1}$ such that $\cup_{k\geq 1} K_k = E$, and given a sequence $(B_n)_{n\geq 1}$, if there exists B_∞ such that for any $k \geq 1$, the sequence $(B_n \cap K_k)_{n\geq 1}$ converges in Hausdorff metric to $B_\infty \cap K_k$, then

$$\lim_n B_n = B_\infty.$$

Remark 4. *In a metric space (E, d) such that every bounded closed set is compact (this is the case for instance, of locally compact length spaces), it is easier to understand the meaning of stability of a class. For any sequence $(B_n)_{n\geq 1}$ of closed sets of E, and any sequence $(K_k)_{k\geq 1}$ of increasing closed balls such that $\cup_{k\geq 1} K_k = E$, there exist a set B_∞ and a subsequence (relabeled $(B_n)_{n\geq 1}$) such that $B_n \cap K_k$ converges in Hausdorff metric to B_∞ and*

$$\lim_n B_n = B_\infty.$$

In this case, a stable class in the sense of definition 1 is just a compact class for the Hausdorff convergence on large balls. Indeed, in such spaces E, there exists an increasing sequence of compacts $(K_k)_{k\geq 1}$ such that $\cup_{k\geq 1} K_k = E$, take for instance a sequence of balls centered on the same point, with an increasing radius. The Hausdorff convergence on large balls is then equivalent to the Hausdorff convergence of $(B_n \cap K_k)_{n\geq 1}$ for any $k \in \mathbb{N}$. Since the closed sets in K_k forms a compact class for the Hausdorff convergence, there exists a subsequence of $(B_n \cap K_k)_{n\geq 1}$ converging to some B_∞^k. Using diagonal argument, we may extract a subsequence of the original sequence $(B_n)_{n\geq 1}$ such that for any $k \in \mathbb{N}$, $(B_n \cap K_k)_{n\geq 1}$ converges in Hausdorff metric to B_∞^k. It is easily checked that $B_\infty := \cup_k B_\infty^k$ is a limit of a subsequence of $(B_n)_{n\geq 1}$, in the sense of definition 1.

Let us introduce some examples of stable sets.

Example 5. *Take $\mathcal{F}$ as the set of all closed sets. Stability is then obvious since the limit considered in the definition of a stable set is always closed as shown in the remark 2.*

Example 6. *The set of all balls is generally not stable, but is does not take much to make is stable. The set of all closed balls and half spaces in $\mathbb{R}^d$ is a stable class. This assertion can be proved using parametrization of the center of the balls in spherical coordinates and using compactness of spheres.*

Example 7. *Other shapes of sets of $\mathbb{R}^d$ make stable classes. Ellipsoids, rectangles, or convex bodies with bounded diameter (by some fixed $R < \infty$) all form stable classes. And it is possible to get rid of the bounded diameter by adding some sets to the class.*

Example 8. *If $\mathcal{F}$ is a stable class of convex sets of a metric space (E, d) such that closed balls are compacts, then*

$$\mathcal{F}_\varepsilon := \{\cup_{F\in\mathcal{G}} F; \mathcal{G} \subset \mathcal{F}, \forall F, G \in \mathcal{G} \inf_{x\in F, y\in G} d(x, y) \geq \varepsilon\},$$

is also a stable class (see lemma 38).

1.1.2 Size function

As we aim to define a smallest set of the sets of $\mathcal{F}$, we need to define a notion of size. This is done using a function τ, meant to measure the *size* of a set. In order to localize the mass, we will thus minimize the size of a set, among all sets given a probability measure.

In order to express our assumptions on τ, we first define the *Hausdorff contrast*.

Definition 9 (Hausdorff contrast)**.** *Let A and B be two closed subset of a Polish space (E, d). The* ***Hausdorff contrast*** *between A and B is defined by*

$$\mathrm{Haus}(A|B) := \inf\{\varepsilon > 0 | A \subset B^{\varepsilon}\}.$$

We can then remark that the Hausdorff metric $d_H(A, B)$ between two closed sets A and B is then

$$d_H(A, B) = \mathrm{Haus}(A|B) \vee \mathrm{Haus}(B|A).$$

We now define formally a size function.

Definition 10 (Size function)**.** *Let (E, d) be a metric space. A function $\tau : \mathcal{F} \to \mathbb{R}^+$ is called a* ***size function*** *if it satisfies the three following conditions:*

(H1) *τ is increasing, i.e. $A \subset B \implies \tau(A) \leq \tau(B)$,*

(H2) *for any decreasing sequence $(A_n)_{n\geq 1} \subset \mathcal{F}$ such that $\tau(A_1) < \infty$ and $\mathrm{Haus}(A_n | \cap_k A_k) \to 0$, the following holds $\tau(A_n) \to \tau(\cap_n A_n)$,*

(H3) *for any sequence $(A_n)_{n\geq 1} \subset \mathcal{F}$, $\tau(\lim_n A_n) \leq \liminf_n \tau(A_n)$.*

Hypothesis (H2) on the size function requires some Hausdorff contrast. This particular choice make the hypothesis weaker and allow the hypothesis to hold for size function that give finite size to non compact sets. The consequences of these hypothesis will be more detailed in the sequel of the paper.

1.2 Overview of the main result

Our main result states that under the condition (H1), (H2) and (H3), for the empirical measure μ_n, and a stable class $\mathcal{F}$,

$$\tau^{\alpha} \leq \liminf_n \tau_n^{\alpha} \leq \limsup_n \tau_n^{\alpha} \leq \lim_{\varepsilon} \tau^{\alpha-\varepsilon},$$

where $\tau_n^{\alpha} = \min\{\tau(B); B \in \mathcal{F}, \mu_n(B) \geq 1 - \alpha\}$.

It implies the convergence of τ_n^{α} when $. \mapsto \tau^{\cdot}$ is continuous.

The result actually holds for a wider class of sequence of measures $(\mu_n)_{n\geq 1}$.

Moreover, simple conditions on the sequence imply the convergence of the minimizers of the τ_n^{α} for different metrics (depending on the conditions assumed). This is discussed in the next sections.

2 First properties

2.1 Existence

Let us recall the setting. (E, d) is a Polish space and μ is a Borel probability measure on (E, d). Given a size function τ, a stable class $\mathcal{F}$ of closed sets of E, and a level α, we define the support B^α of the α-localized measure μ^α of μ by - when possible:

$$B^\alpha \in \arg\min\{\tau(A); A \in \mathcal{F}, \mu(A) \geq 1 - \alpha\},$$

and set

$$\mu^\alpha = \mu(.|B^\alpha).$$

Our first concern is whether B^α exists. It is the matter of the next result.

Theorem 11 (Existence of a minimum)**.** *Let (E, d) be a Polish space, $\mathcal{F}$ a stable class and μ a probability measure on $(E, \mathcal{B}(E))$. Set $0 < \alpha < 1$. Suppose (H3). Then, there exists $B \in \mathcal{F}$ such that*

$$B \in \arg\min\left\{\tau(A); A \in \mathcal{F}(E), \mu(A) \geq 1 - \alpha\right\}.$$

Remark 12. *Hypothesis (H3) can not just be omitted. Indeed, if $\tau(B)$ is defined as the Lebesgue measure of the closure of B on $\mathbb{R}^d$, take*

$$\mu = \alpha\gamma_d + (1 - \alpha)q,$$

where q is a probability measure supported on $\mathbb{Q}^d$ and γ_d is the standard Gaussian measure on $\mathbb{R}^d$, then, the sequence $(B_n)_{n\geq 1}$ defined by

$$B_n := \{x_k\}_{1\leq k\leq n} \cup B(0, r_n),$$

with $\{x_n\}_{n\geq 1} = \mathbb{Q}^d$ and $r_n \to 0$ so that $\mu(B_n) = 1 - \alpha$, is a minimizing sequence. And $\tau(B_n) = \tau\left(B(0, r_n)\right)$ so that $\tau^\alpha = 0$ but $\tau\left(\mathbb{Q}^d\right) = +\infty$.

The minimizer is not necessarily unique. This seems natural with the following example. Take μ as the uniform law on the unit square and an isometric τ. Then any translation small enough of the minimizer will have the same size and the same measure, and will thus be another minimizer. Another result (corollary 25) will comfort us proving that minimizers form a compact set for Hausdorff metric.

The stability condition on $\mathcal{F}$ is needed for existence of the minimum. However, it can be lightly weakened.

Remark 13 (On stability of $\mathcal{F}$)**.** *Since the minimal size $\min\{\tau(A); A \in \mathcal{F}, \mu_n(A) \geq 1 - \alpha\}$ is bounded if $\tau^\alpha = \min\{\tau(A); A \in \mathcal{F}, \mu(A) \geq 1 - \alpha\} < \infty$, then we may suppose instead of stability of $\mathcal{F}$ that all the classes*

$$\mathcal{F}^M := \mathcal{F} \cap \{A; \tau(A) \leq M\}$$

for $M < \infty$ are stable. It is a weaker notion since $\tau(\lim B_n) \leq \liminf \tau(B_n)$ for any sequence $(B_n)_{n\geq 1}$ in $\mathcal{F}$, under (H3).

2.2 Regularity of τ

Denote

$$\tau^\alpha = \inf\{\tau(A); A \in \mathcal{F}, \mu(A) \geq 1-\alpha, B\}.$$

It seems natural to expect $\alpha \mapsto \tau^\alpha$ to be continuous when μ is regular enough. It also seems natural, for instance, to have B^α growing continuously when α decreases to zero, for a unimodal measure μ. This is the concern of this paragraph, the first one establishing the right continuity.

Proposition 14 (Right continuity)**.** *Let (E,d) be a Polish space, μ a probability measure on $(E,\mathcal{B}(E))$ and $\mathcal{F}$ a stable class. Let $0<\alpha<1$. Then, under (H3), $\alpha \mapsto \tau^\alpha$ is right continuous.*

The continuity will require some more hypotheses as shows the following example of discontinuity. Take $\mu = (\delta_x + \delta_y)/2$ and $\alpha = 1/2$, and it is not difficult to find some τ that is not continuous on α.

Thus, it is clear that continuity property of this function needs regularity on the measure we want to localize, with respect to the class $\mathcal{F}$. This is why we introduce the notion of $\mathcal{F}$-regularity.

Definition 15 ($\mathcal{F}$-regularity)**.** *A probability measure μ is said to be **$\mathcal{F}$-regular** if for all $B \in \mathcal{F}$, any $\delta > 0$ and any $C \in \mathcal{F}$ such that $B \subset C$ and $\mu(B) < \mu(B^\delta \cap C)$, there exists $A \in \mathcal{F}$ such that*

$$A \subset B^\delta \cap C,$$
$$\mu(B) < \mu(A).$$

The only purpose of this notion is the continuity of the application $\alpha \mapsto \tau^\alpha$. It is restrictive on μ only when $\mathcal{F}$ is not rich enough. Taking $\mathcal{F}$ as the class of all closed sets of E make any probability measure $\mathcal{F}$-regular. Indeed, since $\mu(B^\delta \cap C) = \lim_n \mu(B^{\delta-1/n} \cap C)$, there exists $n \geq 1$ such that $\mu(B) < \mu(B^{\delta-1/n} \cap C)$ and then we can choose $A := \overline{B^\delta} \cap C$. On the other hand, if $\mathcal{F}$ is not rich enough so that $\tau(\mathcal{F})$ is not even connected, it is easy to build a measure μ that is not $\mathcal{F}$-regular.

Proposition 16 (Continuity)**.** *Let μ be a probility measure on a Polish space (E,d). Suppose (H1), (H2) and (H3), and that μ is $\mathcal{F}$-regular, has a connected support and that τ^α is finite for any $\alpha > 0$ then, the mapping $\alpha \mapsto \tau^\alpha$ is continuous.*

Remark 17. *The condition $\tau^\alpha < \infty$ just avoids a degenerated case.*

This continuity condition is a first step toward the main matters of our article, the consistency.

3 Consistency

3.1 τ-tightness

In order to show the consistency of the mass localization when a sequence of measures $(\mu_n)_{n\geq 1}$ converges to a measure μ, we must make some assumptions on the sequence of measures. The first and most important hypothesis for consistency is the τ-tightness.

Definition 18 (τ-tightness)**.** *A sequence of random probability measures $(\mu_n)_{n\geq 1}$ almost surely weakly converging to a measure μ is **τ-tight** if for any $\delta > 0$ and any $B \in \mathcal{F}$ such that $\tau(B) < \infty$,*

almost surely, for any $C \in \mathcal{F}$ such that $B \subset C$ and $\mu(B) \leq \liminf_n \mu_n(C)$, there exists $A \in \mathcal{F}$ such that

$$\begin{aligned} \mu(B) &\leq \liminf_n \mu_n(A), \\ B \subset A &\subset B^\delta \cap C, \\ \tau(A) &< \infty. \end{aligned}$$

An important remark on this definition is that a τ-tight sequence of random measures does not have necessarily almost surely τ-tight realizations. This can happen to empirical measures for instance. This subtlety lies in the position of "almost surely" in the definition, that is, after the choice of B and δ made.

We can also remark the following. Inequality $\mu(B) \leq \liminf_n \mu_n(C)$ is not a consequence of $B \subset C$. Indeed, the portmanteau theorem states $\limsup_n \mu_n(C) \leq \mu(C)$ and $\limsup_n \mu_n(B) \leq \mu(B)$ since B and C are closed. The conditions for τ-tightness on $B \in \mathcal{F}$ such that $\mu(B) = \lim_n \mu_n(B)$ is clearly verified for $A := B$. The definition of τ-tightness can be understood as follows. Whenever $(\mu_n)_{n\geq 1}$ does not catch all the μ-mass of B (i.e. $\liminf_n \mu_n(B) < \mu(B)$) but some set C that contains B has its μ-mass well caught (i.e. $\mu(B) \leq \liminf_n \mu_n(C)$), then $\mathcal{F}$ must have an element A that also have its μ-mass well caught (i.e. $\mu(B) \leq \liminf_n \mu_n(A)$), of finite size (i.e. $\tau(A) < \infty$) and that is stuck between B and a δ-neighborhood of B intersected with C, for small δ.

The following proposition states that this notion is not empty, and includes the empirical measures.

Proposition 19 (τ-tightness of the empirical measure)**.** *Let μ be a probability measure on E such that $\tau^\alpha < \infty$ for $0 < \alpha < 1$. Let $(X_i)_{i\geq 1}$ be a sequence of i.i.d. random variables with common law μ. Set $\mu_n = \frac{1}{n}\sum_{1\leq i\leq n} \delta_{X_i}$. Then, $(\mu_n)_{n\geq 1}$ is τ-tight.*

The empirical measure is actually not the only simple example of τ-tight sequence. The following corollary gives a simple condition for a sequence of random probability measures to be τ-tight.

Corollary 20. *Let $(\mu_n)_{n\geq 1}$ be a sequence of random probability measure on E almost surely weakly converging to some measure μ, such that $\tau^\alpha < \infty$, for any $0 < \alpha < 1$. If for all $B \in \mathcal{F}$, almost surely, $\mu(B) \leq \lim_n \mu_n(B)$, then $(\mu_n)_{n\geq 1}$ is τ-tight.*

This corollary says that τ-tightness is implied by almost sure convergence of $\mu_n(B)$ for each $B \in \mathcal{F}$ and thus, dropping the "almost sure" makes the τ-tightness much more restrictive.

We can now state our first result on consistency.

3.2 τ-consistency

Our goal is to show that when μ_n converges to μ, the size τ_n^α of the smallest element of a given class $\mathcal{F}$ with μ_n-mass at least $1-\alpha$ converges to the size τ^α of the smallest of μ-mass at least $1-\alpha$. In other words, we want to prove consistency of the smallest size τ^α. The following theorem states conditions for this consistency to hold.

Theorem 21 (Consistency)**.** *Let (E, d) be a Polish space, $\mathcal{F}$ a stable class and $(\mu_n)_{n\geq 1}$ a τ-tight sequence of random probability measures on $(E, \mathcal{B}(E))$ almost surely weakly converging to some measure μ. Set $0 < \alpha < 1$. Choose any $B_n^\alpha \in \arg\min\{\tau(A); A \in \mathcal{F}, \mu_n(A) \geq 1-\alpha\}$ and $\mu_n^\alpha = \mu_n(.|B_n^\alpha)$, for all $n \geq 1$. Then, under hypotheses (H1), (H2) and (H3), the sequence $(\mu_n^\alpha)_{n\in\mathbb{N}}$*

is almost surely totally bounded for the weak convergence topology and $B_\infty^\alpha := \lim_k B_{n_k}$ *along any converging subsequence* $(\mu_{n_k})_{k\geq 1}$ *of* $(\mu_n)_{n\geq 1}$ *satisfies* $\mu(B_\infty^\alpha) \geq 1-\alpha$ *and almost surely*

$$\tau^\alpha \leq \tau(B_\infty^\alpha) \leq \liminf_{n\to\infty} \tau(B_n^\alpha) \leq \limsup_{n\to\infty} \tau(B_n^\alpha) \leq \lim_{\varepsilon\to 0^+} \tau^{\alpha-\varepsilon}.$$

Moreover, if μ *is* $\mathcal{F}$*-regular and its support is connected, the five terms above are equal.*

Note that the τ-tightness condition is required only for the last inequality.

It is rather clear the if $\alpha \mapsto \tau^\alpha$ is not continuous for the measure μ, we can hardly expect consistency of the smallest size τ^α. This first step of consistency brings us to consider consistency of the smallest set of the class itself.

3.3 Minimizer consistency

The smallest set in $\mathcal{F}$ with μ-mass greater than $1-\alpha$ is not always unique, and therefore consistency does not just mean that minimizer for μ_n converges to the minimizer for μ. In order to give a sense to consistency, we will consider the set of all minimizers and the Hausdorff contrast between sets of elements of $\mathcal{F}$ (for some underlying metric $d_\mathcal{F}$ on $\mathcal{F}$). We thus first recall the definition of Hausdorff contrast. Let A and B be two sets. The Hausdorff contrast between A and B is defined by

$$\mathrm{Haus}(A|B) := \inf\{\varepsilon > 0 | A \subset B^\varepsilon\}.$$

Let us denote, for $0 < \alpha < 1$, a sequence of measures $(\mu_n)_{n\geq 1}$ and a measure μ;

$$\mathcal{S}_n^\alpha = \arg\min\{\tau(A); A \in \mathcal{F}, \mu_n(A) \geq 1-\alpha\},$$

and

$$\mathcal{S}^\alpha = \arg\min\{\tau(A); A \in \mathcal{F}, \mu(A) \geq 1-\alpha\}.$$

The sets $\mathcal{S}^\alpha$ and $\mathcal{S}_n^\alpha$ are thus two subsets of $\mathcal{F}$. What we want is to find conditions under which

$$\mathrm{Haus}(\mathcal{S}_n^\alpha | \mathcal{S}^\alpha) \to 0,$$

when n tends to infinity.

We now state and comment briefly the two hypotheses that will be made for our main result.

(H4) $\cdot \mapsto \tau^\cdot$ is continuous at α,

(H5) $\forall (A_n)_{n\geq 1} \subset \mathcal{F}$ such that $\tau(\lim_n A_n) < \infty$,

$$\lim \tau(A_n) = \tau(\lim A_n) \implies d_\mathcal{F}(A_n, \lim_k A_k) \to 0$$

where $d_\mathcal{F}$ denotes a metric on $\mathcal{F}$. A typical example of such metric is the Hausdorff metric or the measure of symmetric difference. Section 4.2 is devoted to these examples and conditions that imply (H5).

The continuity condition (H4) is a consequence of the proposition 16: a connected support for an $\mathcal{F}$-regular measure suffices.

We can now state a direct consequence of theorem 21 and hypotheses (H4) and (H5).

Theorem 22 (Consistency of the minimizers)**.** *Let (E,d) be a Polish space and $(\mu_n)_{n\geq 1}$ a sequence of random τ-tight probability measures on $(E,\mathcal{B}(E))$ almost surely weakly converging to some $\mathcal{F}$-regular measure μ. Set $0<\alpha<1$. Equip $\mathcal{F}$ with a metric $d_{\mathcal{F}}$*

Suppose (H2), (H3), (H4) (for the measure μ), and (H5).

Then, almost surely,

$$\mathrm{Haus}(\mathcal{S}_n^\alpha|\mathcal{S}^\alpha)\to 0.$$

We now make some remarks on the necessity of the hypotheses.

Remark 23 (On (H4))**.** *Continuity condition on $\cdot\mapsto\tau^{\cdot}$ at α can not be dropped. Indeed, if we choose $d\mu_n=(1_{[0;1-\alpha-1/n]}+1_{[2;2+\alpha+1/n]})d\lambda$, then, taking for instance $\tau(B)=\int_0^1\mathcal{M}(B,t)tdt$, yields to the minimizers Supp $\mu_n^\alpha=[0;1-\alpha-1/n]\cup[2;2+1/n]$ and Supp $\mu_\infty^\alpha=[0;1-\alpha]$.*

The result is stated with Hausdorff contrast and not Hausdorff metric. One can wonder what happens with the Hausdorff metric.

Remark 24 (Convergence with Hausdorff metric)**.** *The convergence of* $\mathrm{Haus}(\mathcal{S}^\alpha|\mathcal{S}_n^\alpha)$ *(i.e. the other contrast), is not always true. Take for instance $\tau(B)=\int_0^1\mathcal{M}(B,t)tdt$ (see definition 28) and $d\mu_n=f_nd\lambda$ on $\mathbb{R}$, with $f_n(x)=1+x/n$ for $-1/2<x<1/2$ and $f(x)=0$ otherwise. Then, the minimizer $[-1/2;1/2-\alpha]$ for $\mu=\mathbf{1}_{[0;1]}$ is not included in any ε-neighborhood of $[-1/2+\alpha;1/2]$.*

When $\mathcal{S}^\alpha$ is a singleton, theorem 22 then states that any B_n^α converges to the minimizer for the limit μ. When $\mathcal{S}^\alpha$ is not a singleton, it states that all B_n^α in $\mathcal{S}_n^\alpha$ gets close to an element of $\mathcal{S}^\alpha$, uniformly. However, the remark 24 precises that there could be some elements of $\mathcal{S}^\alpha$ that will not be approximated.

We can derive from the proof of this result the following corollary.

Corollary 25. *Let (E,d) be a Polish space and μ a probability measure on $(E,\mathcal{B}(E))$ such that $\tau^\alpha<\infty$. Set $0<\alpha<1$. Suppose (H2), (H3) and (H5). Then* $\arg\min\{\tau(A);A\in\mathcal{F},\mu(A)\geq 1-\alpha\}$ *is compact for the Hausdorff metric topology.*

The tools developed in the proofs of these results were also effective to prove some continuity on the minimizer on the level α.

3.4 Minimizer continuity

The proof of the theorem of the minimizer consistency is based on the two lemmas 36 and 37. The same technique of proof leads to the following result.

Proposition 26. *Let (E,d) be a Polish space and a $\mathcal{F}$-regular probability measure μ on $(E,\mathbb{B}(E))$. Set $0<\alpha<1$ and $(\alpha_n)_{n\geq 1}$ converging to α. Suppose (H2), (H3), (H4), and (H5). Then,*

$$\mathrm{Haus}(\mathcal{S}^{\alpha_n}|\mathcal{S}^\alpha)\to 0.$$

In the case of a unique minimizer for all α, this result states the continuity of the function that associate α to the minimizer for α:

$$\alpha\mapsto B^\alpha.$$

4 Examples

In this section, we introduce examples of stable classes and size functions, in order to show the scope of our results.

4.1 Examples of stable classes

4.1.1 Closed sets

The simplest example of stable class is the set of all closed sets. Indeed, given a sequence of closed sets $(A_n)_{n\geq 1}$, the limit $\lim A_n$ is also closed (see remark 2).

4.1.2 Parametrized classes

On $\mathbb{R}^d$, the set of closed balls and half spaces is also a stable class. Indeed, each ball can be parametrized by (x, r), the center and radius of the ball, and half spaces are limit (in the limit of sets we defined) of balls. From this parametrization, one can show that there exists a converging subsequence to any sequence of balls and half spaces in the sense we defined for the limit of sets.

More generally, one can use the remark 13 which states that for a given size function τ, the condition on the class $\mathcal{F}$ to be stable can be weakened to the condition that $\mathcal{F}^M = \mathcal{F} \cap \{B; \tau(B) \leq M\}$ is stable for each $M \in \mathbb{R}^+$. Then, taking any parametrized class $\mathcal{F}$ such that convergence of the parameters implies convergence of the sets (in the sense defined for stable sets) and such that $\mathcal{F}^M$ is compact for the Hausdorff metric gives a stable set, since Hausdorff convergence implies convergence in the sense we defined for stable sets.

4.1.3 ε-separated unions

Another example of stable set is the one of ε-separated union of elements of a stable class of convex sets (see lemma 38). For $\mathcal{F}$ a stable class of convex sets of a metric space (E, d) such that bounded sets are compact, the following set is stable

$$\mathcal{F}_\varepsilon := \{\cup_{F\in\mathcal{G}} F; \mathcal{G} \subset \mathcal{F}, \forall F, G \in \mathcal{G} \inf_{x\in F, y\in G} d(x, y) \geq \varepsilon\}.$$

This sets give an application to classification.

4.2 Examples of size function

4.2.1 Packing

Our first example of size function is functions depending on the packing of sets. Let us recall the notion of packing and the more common one of covering (see [KT61]).

Definition 27 (t-covering)**.** *A set $\{B_i\}_{i\in I}$ of subsets of E is a **t-covering** of a set B if the diameter of any B_i does not exceed $2t$ and*

$$B \subset \bigcup_{i\in I} B_i.$$

*The cardinal of the smallest t-covering of B is then called **covering number** of B for t and is denoted $\mathcal{N}(B, t)$.*

The logarithm of $\mathcal{N}(B,t)$ is sometimes called the metric entropy of B, or its Kolmogorov entropy.

Definition 28 (t-separated)**.** *A set B is **t-separated** if the distance of every two distinct points of B is strictly greater then t. The cardinal of the greatest t-separated subset of B is called **packing number** of B for t and is denoted $\mathcal{M}(B,t)$.*

The logarithm of $\mathcal{M}(B,t)$ is sometimes called capacity of B.

These two notions are intuitively linked and carry information on the size of the set. The following proposition due to [KT61] compares $\mathcal{N}$ and $\mathcal{M}$.

Proposition 29. *For any set B of a metric space (E,d),*

$$\mathcal{M}(B,2t) \leq \mathcal{N}(\bar{B},t) = \mathcal{N}(B,t) \leq \mathcal{M}(B,t) = \mathcal{M}(\bar{B},t).$$

The idea to make the notion of size depends on the packing of the set is not random and comes from the fact that packing appears in many notions of size. For instance, the packing measure defined in [SRT88] is another definition of the Lebesgue measure (up to some constant factor) and is in the more general case of a metric space an isometric measure that coincide with Hausdorff measure on spaces with non fractional dimension as shown in [SRT88].

We will then study size function τ of the form

$$\tau(B) = \Phi(\mathcal{M}(B,.)),$$

for all $B \in \mathcal{F}$ for a function Φ on the set of packing functions.

In order to ensure that the condition (H5) is fulfilled, we define the following hypotheses.

(H'1) τ is strictly increasing (i.e. $A \subset B, A \neq B \implies \tau(A) < \tau(B)$),

(H'3) for any sequence $(B_n)_{n\geq 1} \subset \mathcal{F}$, $\Phi(\liminf \mathcal{M}(B_n,.)) \leq \liminf \Phi(\mathcal{M}(\lim B_n,.))$,

(H6) $\tau(A) < \infty \implies A$ totally bounded.

Hypotheses (H'1) and (H'3) imply respectively (H1) and (H3) and we can show that, together with (H6), they imply (H5), so that we have the following theorem.

Theorem 30 (Packing size function)**.** *Suppose that τ is a size function of the form $\tau(B) = \Phi(\mathcal{M}(B,.))$ on a stable class $\mathcal{F}$. Suppose (H'1), (H2), (H'3), (H4) and (H6).*

Then, almost surely,

$$\mathrm{Haus}(\mathcal{S}_n^\alpha | \mathcal{S}^\alpha) \to 0.$$

Remark 31 (On (H6))**.** *Since a non compact space can not be arbitrarily close to a finite set in Hausdorff distance, then for B_n^α to converge when $(\mu_n)_{n\geq 1}$ is sequence of finitely supported measures, minimizers relative to μ must be compact if $\mathcal{F}$ is rich enough to make minimizers finite sets. Minimizers relative to μ are compact when $\tau(B) < \infty$ implies that B is totally bounded.*

4.3 Examples of sequence of measures

Proposition 19 and theorem 22 can be applied to the empirical measure and then lead to the following.

Theorem 32 (Consistency for empirical measure: i.i.d. case)**.** *Let (E,d) be a Polish space, μ a probability measure on $(E,\mathbb{B}(E))$ and $\mathcal{F}$ a stable set such that $\tau^\alpha < \infty$, for any $0<\alpha<1$. Given a sample $(X_n)_{n\geq 1}$ of independent random variables with same law μ, set*

$$\mu_n = \frac{1}{n}\sum_{1\leq i\leq n}\delta_{X_i},$$

the empirical measure. Suppose (H2), (H3), (H4) and (H5). Then, almost surely,

$$\mathrm{Haus}(\mathcal{S}_n^\alpha|\mathcal{S}^\alpha)\to 0.$$

Likewise, corollary 20 leads to convergence of empirical measure for some dependent cases.

Theorem 33 (Consistency for empirical measure: dependent case)**.** *Let (E,d) be a Polish space and μ be a probability measure on $(E,\mathbb{B}(E))$ such that $\tau^\alpha<\infty$, for any $0<\alpha<1$. Given random variables $(X_n)_{n\geq 1}$ from an ergodic Markov chain with invariant measure μ, denote*

$$\mu_n = \frac{1}{n}\sum_{1\leq i\leq n}\delta_{X_i},$$

the empirical measure. Suppose (H2), (H3), (H4), (H5) and (H6). Then, almost surely,

$$\mathrm{Haus}(\mathcal{S}_n^\alpha|\mathcal{S}^\alpha)\to 0.$$

5 Proofs

The following lemma is the starting point of the existence of minimizer theorem. It is also a key lemma to most of the other results.

Lemma 34. *Let (E,d) be a Polish space. Let $(\mu_n)_{n\geq 1}$ be a sequence of probability measures weakly converging to μ_∞. Then,*

$$\textit{Supp } \mu_\infty \subset \lim_n \textit{Supp } \mu_n.$$

Moreover, for any $t>0$

$$\mathcal{M}(\textit{Supp } \mu_\infty, t) \leq \mathcal{M}(\lim_n \textit{Supp } \mu_n, t) \leq \liminf_n \mathcal{M}(\textit{Supp } \mu_n, t).$$

Proof of lemma 34. Set $t>0$ and $x\in \mathrm{Supp}\ \mu_\infty$. Then, for all $\varepsilon>0$, there exists $\eta>0$ such that $\mu_\infty(B(x,\varepsilon/2))>\eta$. Then, using the portmanteau theorem, for n large enough, $\mu_n(B(x,\varepsilon/2))>\eta$ and thus, $x\in(\mathrm{Supp}\ \mu_n)^\varepsilon$ for any n large enough, which proves the first point and first inequality.

For the second inequality, choose $m\leq \mathcal{M}(\lim_n \mathrm{Supp}\ \mu_n, t)$. Let $\{x_i\}_{1\leq i\leq m}$ be a t-separated subset of $\lim_n \mathrm{Supp}\ \mu_n$. Then, there exists $\delta>0$ such that

$$\inf_{i\neq j} d(x_i,x_j) > t+\delta.$$

From definition of $\lim_n \mathrm{Supp}\ \mu_n$ and since m is finite, there exists $k\geq 1$ such that $(x_i)_{1\geq i\geq m}\subset \cap_{n\geq k}(\mathrm{Supp}\ \mu_n)^{\delta/2}$

Choose now $x_i^n \in B(x_i, \delta/2) \cap \text{Supp } \mu_n$ for $1 \leq i \leq m$. $\{x_i^n\}_{1 \leq i \leq m}$ forms a t-separated set included in Supp μ_n since for $i \neq j$

$$\begin{aligned} d(x_i^n, x_j^n) &\geq d(x_i, x_j) - d(x_i, x_i^n) - d(x_j, x_j^n) \\ &> t + \delta - \delta/2 - \delta/2 = t. \end{aligned}$$

Thus, $m \leq \liminf_n \mathcal{M}(\text{Supp } \mu_n, t)$, for any $m \leq \mathcal{M}(\lim_n \text{Supp } \mu_n, t)$, which ends the proof. □

Proof of theorem 11. Let $(B_n)_{n \geq 1} \subset \mathcal{F}$ be a minimizing sequence such that $\mu(B_n) \geq 1 - \alpha$ for all $n \geq 1$. Set $\mu_n = \mu(.|B_n)$. Let first show that $(\mu_n)_{n \geq 1}$ is tight. Let K be a compact set such that $\mu(K) \geq 1 - \varepsilon(1 - \alpha)$, then

$$\begin{aligned} \mu_n(K) &= \mu(K \cap B_n)/\mu(B_n) \\ &= (1 - \mu(K^c \cup B_n^c))/\mu(B_n) \\ &\geq 1 - \mu(K^c)/\mu(B_n) \\ &\geq 1 - \varepsilon. \end{aligned}$$

Thus, up to a subsequence, $(\mu_n)_{n \geq 1}$ weakly converges to μ_∞. Set $B = \lim_n \text{Supp } \mu_n$. B^ε is an open set and contains Supp μ_∞ from lemma 34. Then, portmanteau theorem gives

$$\begin{aligned} 1 = \mu_\infty(B^\varepsilon) \longleftarrow \mu_n(B^\varepsilon) &= \mu(B^\varepsilon \cap B_n)/\mu(B_n) \\ &\leq \mu(B^\varepsilon)/(1 - \alpha). \end{aligned}$$

Letting ε tends to zero and using $\bigcap_{\varepsilon > 0} B^\varepsilon = B$, it comes $\mu(B) \geq 1 - \alpha$.

Proof ends using (H3) and stability of $\mathcal{F}$. □

Proof of proposition 14. Clearly, $\alpha \mapsto \tau^\alpha$ decreasing, and we thus only need to show $\tau^\alpha \geq \lim_{\varepsilon \to 0^+} \tau^{\alpha + \varepsilon}$. Set $B_\alpha \in \arg\min\{\tau(A); A \in \mathcal{F}, \mu(A) \geq 1 - \alpha\}$ and $\mu^\alpha = \mu(.|B_\alpha)$. For any $(\alpha_n)_{n \geq 1} \subset (0; 1)$ decreasing and converging to α, $(\mu^{\alpha_n})_{n \geq 1}$ is tight. Indeed, for a compact set K such that $\mu(K) \geq 1 - \varepsilon(1 - \alpha_1)$,

$$\begin{aligned} \mu^{\alpha_n}(K) &= \mu(K \cup B_{\alpha_n})/\mu(B_{\alpha_n}) \\ &\geq (1 - \mu(K^c) - \mu(B_{\alpha_n}^c)/\mu(B_{\alpha_n}) \\ &= 1 - \mu(K^c)/\mu(B_{\alpha_n}) \\ &\geq 1 - \varepsilon. \end{aligned}$$

Thus, up to a subsequence, μ^{α_n} converges to some probability measure μ_∞. Set $B_\infty = \lim_n \text{Supp } \mu_n$. B_∞^ε is an open set that contains Supp μ_∞, and thus portmanteau theorem yields,

$$\begin{aligned} 1 = \mu_\infty(B_\infty^\varepsilon) \longleftarrow & \mu^{\alpha_n}(B_\infty^\varepsilon) \\ = & \mu(B_\infty^\varepsilon \cap B_{\alpha_n})/\mu(B_{\alpha_n}) \\ \leq & \mu(B_\infty^\varepsilon)/(1 - \alpha_n) \longrightarrow \mu(B_\infty^\varepsilon)/(1 - \alpha). \end{aligned}$$

Letting ε tends to zero shows $\mu(B_\infty) \geq 1 - \alpha$. Hypothesis (H3) and stability of $\mathcal{F}$ let us conclude

$$\tau^\alpha \leq \tau(B_\infty) \leq \liminf_n \tau(B_{\alpha_n}) = \lim_n \tau(B_{\alpha_n}).$$

□

In order to show continuity of the size function (proposition 16), we establish a lemma that states that for a connected measure, any ε-neighborhood of Borel set has strictly more mass than the original Borel set.

Lemma 35. *Let μ be a probability measure with a connected support, then for any Borel set A such that $\mu(A) < 1$ and any $\varepsilon > 0$, the ε-neighborhood of A satisfies*

$$\mu(A^\varepsilon) > \mu(A).$$

Proof of lemma 35. In order to show a contradiction, suppose $\mu(A^\varepsilon) = \mu(A)$ for some $\varepsilon > 0$. Then, $\mu(A^{\varepsilon/3}) = \mu(A)$ and $\mu(((A^\varepsilon)^c)^{\varepsilon/3}) \geq \mu((A^\varepsilon)^c) = 1 - \mu(A)$ so that $\mu(A^{\varepsilon/3} \cup ((A^\varepsilon)^c)^{\varepsilon/3}) = 1$. Since $\overline{A^{\varepsilon/3}} \cap \overline{((A^\varepsilon)^c)^{\varepsilon/3}} = \emptyset$ and $\overline{A^{\varepsilon/3}}$ and $\overline{((A^\varepsilon)^c)^{\varepsilon/3}}$ are closed, the support of μ not connected, which contradicts the hypothesis. □

We can now prove continuity of the smallest size function.

Proof of proposition 16. By proposition 14 it is enough to show that $\alpha \mapsto \tau^\alpha$ is left continuous. Moreover, since $\alpha \mapsto \tau^\alpha$ is decreasing, it is enough to prove $\lim_{\varepsilon \to 0^+} \tau^{\alpha-\varepsilon} \leq \tau^\alpha$.

Let $\alpha > 0$ and $B \in \mathcal{F}$ such that $\tau^\alpha = \tau(B)$. If $\mu(B) > 1 - \alpha$ then the result is obvious. Suppose then $\mu(B) = 1 - \alpha$. By lemma 35, for any $\delta > 0$ there exists $\varepsilon > 0$ such that $\mu(B^\delta) \geq 1 - \alpha + \varepsilon$. Define a decreasing sequence $(\delta_n)_{n\geq 1}$ converging to zero, such that $\mu(B^{\delta_n}) > 1 - \alpha$ for any $n \geq 1$. Since μ is $\mathcal{F}$-regular, there exists $(K_n)_{n\geq 1} \subset \mathcal{F}$ such that for any $n \geq 1$

$$\begin{aligned}
&\tau(K_1) < \infty, \text{ since } \tau^{\alpha-\varepsilon} < \infty,\\
&K_n \subset B^{\delta_n} \cap K_{n-1},\\
&1 - \alpha < \mu(K_n) =: 1 - \alpha_n.
\end{aligned}$$

Clearly, $\alpha_n \to \alpha$. Then, using (H1) and (H2), it yields

$$\lim_{\varepsilon \to 0^+} \tau^{\alpha-\varepsilon} \longleftarrow \tau^{\alpha_n} \leq \tau(K_n) \longrightarrow \tau(\cap_n K_n) = \tau^\alpha.$$

□

Proof of proposition 19. Since τ^α is finite, for any $0 < \alpha < 1$, there exists $B \in \mathcal{F}$ such that $\tau(B) < \infty$ and $\mu(B) \geq 1 - \alpha$. For such B, the law of large number states that almost surely $\lim_n \mu_n(B) = \mu(B)$. Set $\delta > 0$. Then for any $C \in \mathcal{F}$ such that $B \subset C$ and $\mu(B) \leq \liminf_n \mu_n(C)$, for $A := B \in \mathcal{F}$ the three following conditions are clearly true

$$\begin{aligned}
\mu(B) &\leq \liminf_n \mu_n(A),\\
B \subset A &\subset B^\delta \cap C,\\
\tau(A) &< \infty.
\end{aligned}$$

□

Proof of theorem 21. Let K be a compact set such that $\mu_n(K) \geq 1 - \varepsilon(1 - \alpha)$ for any $n \geq 1$ and $\mu(K) \geq 1 - \varepsilon(1 - \alpha)$. Then,

$$\begin{aligned}
\mu_n^\alpha(K) &= \mu_n(K \cap B_n^\alpha)/\mu_n(B_n^\alpha)\\
&\geq 1 - \varepsilon,
\end{aligned}$$

showing tightness of $(\mu_n^\alpha)_{n\geq 1}$. Denote μ_∞^α a limit measure and $B_\infty^\alpha = \lim_n \text{Supp}\, \mu_n^\alpha$. Set $\varepsilon = \inf\{\eta > 0; \mu_n(B) \leq \mu(B^\eta) + \eta, \forall B \in \mathbb{B}(E)\}$ the Prokhorov distance between μ and μ_n. Then using together the portmanteau theorem and the fact that the topology induced by the Prokhorov metric coincides with the weak convergence topology, for any $\eta > 0$ there exists n_η such that for all $n > n_\eta$ - since Supp $\mu_\infty^\alpha \subset B_\infty^\alpha$ from lemma 34,

$$\begin{aligned}
1 &= \mu_\infty^\alpha(B_\infty^\alpha) \\
&\leq \mu_n^\alpha((B_\infty^\alpha)^\varepsilon | B_n^\alpha) + \eta \\
&= \mu_n((B_\infty^\alpha)^\varepsilon \cap B_n^\alpha)/\mu_n(B_n^\alpha) + \eta \\
&\leq \mu_n((B_\infty^\alpha)^\varepsilon)/(1-\alpha) + \eta \\
&\leq (\mu((B_\infty^\alpha)^{2\varepsilon}) + \varepsilon)/(1-\alpha) + \eta \longrightarrow \mu(B_\infty^\alpha)/(1-\alpha).
\end{aligned}$$

Thus, $\mu(B_\infty^\alpha) \geq 1 - \alpha$ and thus by definition of τ^α,

$$\tau^\alpha \leq \tau(B_\infty^\alpha).$$

The second inequality is a direct application of (H3).

Let us relabel a subsequence of $(\mu_n)_{n\geq 1}$ so that $\limsup_{n\to\infty} \tau(B_n^\alpha) = \lim_{n\to\infty} \tau(B_n^\alpha)$. Choose $\varepsilon > 0$, and $B_\varepsilon \in \arg\min\{\tau(A); A \in \mathcal{F}, \mu(A) \geq 1 - \alpha + \varepsilon\}$. Without loss of generality, we can suppose that $\tau(B_\varepsilon) < \infty$. Since the sequence $(\mu_n)_{n\geq 1}$ converges to μ, the portmanteau theorem states that for any $\delta > 0$,

$$\liminf_n \mu_n(B_\varepsilon^\delta) \geq \mu(B_\varepsilon).$$

Set $\delta_1 > 0$. Then, since $E \in \mathcal{F}$, τ-tightness implies that there exists $A_1 \in \mathcal{F}$ such that

$$\begin{aligned}
\tau(A_1) &< \infty, \\
B_\varepsilon \subset A_1 &\subset B_\varepsilon^{\delta_1}, \\
1 - \alpha + \varepsilon \leq \mu(B_\varepsilon) &\leq \liminf_n \mu_n(A_1), \text{ a.s..}
\end{aligned}$$

Let $(\delta_k)_{k\geq 1}$ be a decreasing sequence converging to zero. Using recursively the τ-tightness property of the sequence $(\mu_n)_{n\geq 1}$, with sets A_{k-1}, define a decreasing sequence $(A_k)_{k\geq 1}$ of elements of $\mathcal{F}$, such that $A_k \subset B_\varepsilon^{\delta_k}$ for any $k \geq 1$, and

$$\begin{aligned}
B_\varepsilon \subset A_k &\subset B_\varepsilon^{\delta_k}, \\
1 - \alpha + \varepsilon \leq \mu(B_\varepsilon) &\leq \liminf_n \mu_n(A_k) \text{ a.s..}
\end{aligned}$$

Then, for any $k \geq 1$, there exists $n_k \geq 1$ such that $\mu_n(A_k) \geq 1 - \alpha$ for any $n > n_k$. Then, since $A_k \in \mathcal{F}$, $\tau(B_n^\alpha) \leq \tau(A_k)$ for all $n \geq n_k$. By construction and since B_ε is closed, $\cap_k A_k = B_\varepsilon$. Using (H1) and (H2) yields

$$\tau(B_n^\alpha) \leq \tau(A_k) \longrightarrow_k \tau(\cap_k A_k) = \tau(B_\varepsilon),$$

which shows that for any $\varepsilon > 0$,

$$\limsup_n \tau(B_n^\alpha) \leq \tau^{\alpha-\varepsilon}.$$

The inequality of the five terms is a direct application of proposition 16. □

Proof of theorem 22. Choose $B_n^\alpha \in \mathcal{S}_n^\alpha$. Set $\mu_n^\alpha = \mu_{n|B_n^\alpha}$. Theorem 21, lemma 34 and continuity condition (H4) of $\cdot \mapsto \tau^\cdot$ shows that $(\mu_n^\alpha)_{n\geq 1}$ is totally bounded for the weak convergence topology and that $B_\infty^\alpha := \lim_k B_{n_k} \in \mathcal{S}^\alpha$ where $(n_k)_{k\geq 1}$ is a subsequence along which μ_n^α converges. Theorem 21 and (H5) then yield

$$d_H(B_\infty^\alpha, B_n^\alpha) \to 0, \tag{5}$$

for the subsequence of $(B_n)_{n\geq 1}$. Since (5) holds for a subsequence of any subsequence of the original sequence $(B_n^\alpha)_{n\geq 1}$, taking $B_n^\alpha \in \mathcal{S}_n^\alpha$ so that $1/n + d_H(B_n^\alpha, B_\infty^\alpha) \geq \text{Haus}(\mathcal{S}_n^\alpha|\mathcal{S}^\alpha)$ for any $B_\infty^\alpha \in \mathcal{S}^\alpha$, it follows

$$\text{Haus}(\mathcal{S}_n^\alpha|\mathcal{S}^\alpha) \to 0.$$

□

Proof of proposition 26. Choose a sequence $(\alpha_n)_{n\geq 1}$ converging to α. From continuity of the size function, for any n large enough, $\tau^{\alpha_n} < \infty$. Define $\mu_n = \mu(.|B_n)$ for any $B_n \in \mathcal{S}^{\alpha_n}$. Sequence $(\mu_n)_{n\geq 1}$ is tight. Indeed, let K be a compact set such that $\mu(K) \geq 1 - \varepsilon(1 - \max_{n\geq 1} \alpha_n)$, then

$$\begin{aligned}
\mu_n(K) &= \mu(K \cap B_n)/\mu(B_n) \\
&= (1 - \mu(K^c \cup B_n^c))/\mu(B_n) \\
&\geq 1 - \mu(K^c)/\mu(B_n) \\
&\geq 1 - \varepsilon.
\end{aligned}$$

Set $B_\infty = \lim_k B_{n_k}$ with $(n_k)_{k\geq 1}$ a subsequence along which $(\mu_n)_{n\geq 1}$ converges (such subsequence exists since $(\mu_n)_{n\geq 1}$ is tight). Set $\mu_\infty = \lim_k \mu_{n_k}$. Using similar arguments to those in theorem 11, we can show that $\mu(B_\infty) \geq 1 - \alpha$. Thus hypothesis (H3) yields

$$\tau^\alpha \leq \tau(B_\infty) \leq \liminf \tau(B_n) = \lim \tau^{\alpha_n} = \tau^\alpha, \tag{6}$$

where the last equality is due to the continuity condition (H4). (H5) implies

$$d_{\mathcal{F}}(B_\infty, B_n) \to 0. \tag{7}$$

Remark then that corollary 25 establishes that $\mathcal{S}^{\alpha_n}$ is compact for the Hausdorff metric topology, with underlying metric $d_{\mathcal{F}}$. Since (7) holds for a subsequence of any subsequence of the original sequence $(B_n)_{n\geq 1}$, taking

$$B_n \in \arg\max\{\min_{B\in\mathcal{S}^\alpha} d_H(B, A)|A \in \mathcal{S}^{\alpha_n}\},$$

so that $d_H(B_n, B_\infty) \geq \text{Haus}(\mathcal{S}^{\alpha_n}|\mathcal{S}^\alpha)$ for any $B_\infty \in \mathcal{S}^\alpha$ yields

$$\text{Haus}(\mathcal{S}^{\alpha_n}|\mathcal{S}^\alpha) \to 0.$$

□

Proof of corollary 25. Set $(\mu_n)_{n\geq 1} = (\mu)_{n\geq 1}$ in the theorem 22 and the result is then a direct application of (5) from the proof of the proposition. □

Proof of corollary 30. In order to prove this result, we show that (H'1), (H'3) and (H6) all together imply (H1), (H3) and (H5). Thus, suppose (H'1), (H'3) and (H6). Then, (H1) and (H3) are obviously true.

Now, in order to show (H5), choose a sequence $(A_n)_{n\geq 1} \subset \mathcal{F}$ such that $\tau(\lim_n A_n) < \infty$, and $\lim \tau(A_n) = \tau(\lim A_n)$. Then, lemma 36 implies that

$$H(\lim_k A_k | A_n) \to 0.$$

Moreover,

$$\begin{aligned}\lim_n \tau(A_n) &= \lim_n \Phi(\mathcal{M}(A_n, .)) \\ &\geq \Phi(\liminf \mathcal{M}(A_n, .) \text{ using (H'3)} \\ &\geq \Phi(\mathcal{M}(\lim_n A_n, .)) \text{ using lemma 34} \\ &= \tau(\lim A_n).\end{aligned}$$

Thus, since we supposed $\lim \tau(A_n) = \tau(\lim A_n)$, and because Φ is supposed to be strictely increasing by (H'1), lemma 34 shows

$$\liminf \mathcal{M}(A_n, t) = \mathcal{M}(\lim A_n, t), \forall t > 0.$$

Lemma 37 can thus be applied to show that, along a subsequence,

$$d_H(A_n, \lim_k A_k) \to 0.$$

Since this is true for any subsequence, it holds for the sequence itself, i.e. (H5) holds.

The result is then an application of theorem 22. □

Lemma 36. *Let (E, d) be a Polish space and $(B_n)_{n\geq 1}$ be a sequence of set of E. Then, for any $\varepsilon > 0$ such that $\mathcal{M}(\lim_k B_k, \varepsilon) < \infty$, there exists $n_0 \in \mathbb{N}$ such that for all $n > n_0$,*

$$\lim_k B_k \subset B_n^{2\varepsilon}.$$

Proof of lemma 36. Set $\varepsilon > 0$ and $m = \mathcal{M}(\lim_k B_k, \varepsilon)$. Let $\{x_i\}_{1\leq i\leq m}$ be a maximal ε-separated subset of $\lim_n B_n$. Then, by definition of $\lim_k B_k$ and since m is finite, for any $\delta > 0$, there exists k_δ such that for all $n \geq k_\delta$,

$$\{x_i\}_{1\leq i\leq m} \subset B_n^{\delta}.$$

Maximality of $\{x_i\}_{1\leq i\leq m}$ ensures

$$\lim_k B_k \subset \bigcup_{1\leq i\leq m} \overline{B(x_i, \varepsilon)} \subset B_n^{\varepsilon+\delta}.$$

Proofs ends when choosing $\delta \leq \varepsilon$. □

Lemma 37. *Let (E, d) be a metric space. Let $(B_n)_{n\geq 1}$ be a sequence of closed sets such that* $\mathrm{Haus}(B|B_n) \to 0$, *for some closed set B. Suppose that for any $s > 0$, there exists $0 < t < s$ such that*

$$\liminf \mathcal{M}(B_n, t) \leq \mathcal{M}(B, t) < \infty. \tag{8}$$

Then along a (relabeled) subsequence $(B_n)_{n\geq 1}$,

$$d_H(B_n, B) \to 0.$$

Proof of lemma 37. It suffices to show

$$\forall \varepsilon > 0, \forall n_0 \in \mathbb{N}, \exists n \geq n_0, B_n \subset (B)^{\varepsilon}.$$

In order to prove a contradiction, suppose

$$\exists \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, B_n \not\subset (B)^{\varepsilon}. \tag{9}$$

Then,

$$\exists \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n > n_0, \exists x_n \in B_n, \forall y \in B, d(x_n, y) > \varepsilon.$$

Set $t < \varepsilon$ and $m = \mathcal{M}(B, t)$. Since $\mathcal{M}(B, .)$ is right continuous, there exists $\delta > 0$ such that

$$m = \mathcal{M}(B, t) = \mathcal{M}(B, t + \delta).$$

Let $\{z_i\}_{i \leq m}$ be a maximal $(t + \delta)$-separated subset of B. Then, since for any $\eta > 0$ and n large enough,

$$B \subset (B_n)^{\eta},$$

we can construct $\{y_i\}_{i \leq m} \subset B_n$ such that for all $1 \leq i \leq m$, $d(y_i, z_i) \leq \eta$. Thus, for any $i \neq j$,

$$d(y_i, y_j) \geq d(z_i, z_j) - d(z_i, y_i) - d(z_j, y_j) > t + \delta - 2\eta.$$

Choose then $\eta < \delta/2 \wedge (\varepsilon - t)$ such that $\{y_i\}_{i \leq m}$ is a t-separated subset of B_n. Then, $\{x_n\} \cup \{y_i\}_{i \leq m}$ is also a t-separated subset of B_n since $x_n \in B_n$ and

$$d(x_n, y_i) \geq d(x_n, z_i) - d(z_i, y_i) > \varepsilon - \eta > t,$$

showing thus $\mathcal{M}(B, t) + 1 \leq \mathcal{M}(B_n, t)$, for any n large enough. This contradicts (8) and thus shows that (9) is false. □

Lemma 38. *Let $\mathcal{F}$ be a stable class of connected sets of a metric space (E, d) such that bounded sets are compact. Suppose that there exists an increasing union $(K_k)_{k \geq 1}$ of balls verifying $\cup_{k \geq 1} K_k = E$ such that for all $k \geq 1$ and any $F \in \mathcal{F}$, $K_k \cap F$ is still connected then*

$$\mathcal{F}_{\varepsilon} := \{\cup_{F \in \mathcal{G}} F; \mathcal{G} \subset \mathcal{F}, \forall F, G \in \mathcal{G} \inf_{x \in F, y \in G} d(x, y) \geq \varepsilon\},$$

is also a stable class.

Proof. Set $(B_n)_{n \geq 1} \subset \mathcal{F}_{\varepsilon}$. Remark 4 states that there exists B_{∞} such that a subsequence of $(B_n)_{n \geq 1}$ converges to B_{∞} for the generalized Hausdorff convergence. Suppose that there exist two connected components $F_1, F_2 \subset B_{\infty}$ such that

$$\underline{d}(F_1, F_2) := \inf\{d(x, y); x \in F_1, y \in F_2\} < \varepsilon.$$

Then, the hypothesis on E implies that for any ball K for which $K \cap F_1 \neq \emptyset$ and $K \cap F_2 \neq \emptyset$, there exists $x_1 \in F_1 \cap K$ and $x_2 \in F_2 \cap K$ such that $\underline{d}(F_1 \cap K, F_2 \cap K) = d(x, y)$. One can check that if x and y belong to the interior of K (which we can assume without loss of generality), it holds

$$d_H(B_n \cap K, B_{\infty} \cap K) \to 0.$$

Taking then $K \in \{K_k\}_{k\geq 1}$ large enough so that

$$\underline{d}(F_1 \cap K, F_2 \cap K) = \eta < \varepsilon.$$

and setting

$$\delta = \underline{d}(K \cap F_1, K \cap B_\infty \setminus F_1) \wedge \underline{d}(K \cap F_1, K \cap B_\infty \setminus F_1) > 0,$$

yields for all n large enough

$$F_1 \cap K \subset B_n^{((\varepsilon-\eta)\wedge\delta)/3} \cap K \text{ and } F_2 \cap K \subset B_n^{((\varepsilon-\eta)\wedge\delta)/3}, \tag{10}$$

$$B_n \cap K \subset B_\infty^{\delta/2} \cap K. \tag{11}$$

(10) shows then that there exist $y_1, y_2 \in B_n \cap K$ such that $d(x_1, y_1) \leq ((\varepsilon-\eta)\wedge\delta)/3$ and $d(x_2, y_2) \leq ((\varepsilon-\eta)\wedge\delta)/3$ and so

$$\begin{aligned} d(y_1, y_2) &\leq d(y_1, x_1) + d(x_1, x_2) + d(x_2, y_2) \\ &\leq 2(\varepsilon-\eta)/3 + \eta \\ &\leq 2\varepsilon/3 + \eta/3 < \varepsilon. \end{aligned} \tag{12}$$

And (11) shows that $B_n \cap K \cap F_1^{\delta/2} \subset F_1^{\delta/4} \cap K$ and $B_n \cap K \cap F_2^{\delta/2} \subset F_2^{\delta/4} \cap K$, since $K \cap B_\infty^{\delta/4} \cap F_i^{\delta/2} \subset K \cap F_i^{\delta/4}$ for $i = 1, 2$ by definition of δ. It follows that $y_1 \in B_n \cap K \cap F_1^{\delta/2}$ and $y_2 \in B_n \cap K \cap F_2^{\delta/2}$ are included in two distinct connected components $B_n \cap K$, which contradicts (12) and hypothesis that $B_n \in \mathcal{F}_\varepsilon$.

It is then easily checked that $F_1, F_2 \in \mathcal{F}$. □

6 Conclusion

We have thus defined a new way to localize mass of measure that is consistent for empirical measures, in Hausdorff metric. It is rather intuitive and applies to any Polish space, including infinite dimensional spaces. It thus provides an analogy to level sets in theses spaces. The major drawbacks of our methods lie in the computability of the size function for rich classes $\mathcal{F}$ and the lack of rate of convergence for now.

References

[CAGM97] J. A. Cuesta-Albertos, A. Gordaliza, and C. Matran. Trimmed k-means: An attempt to robustify quantizers. The Annals of Statistics, 25(2):pp. 553–576, 1997.

[Har87] J. A. Hartigan. Estimation of a convex density contour in two dimensions. Journal of the American Statistical Association, 82(397):267–270, 1987.

[KT61] A. N. Kolmogorov and V. M. Tihomirov. E-Entropy and E-Capacity of Sets in Functional Spaces. American Mathematical Society, 1961.

[Nol91] D Nolan. The excess-mass ellipsoid. Journal of Multivariate Analysis, 39(2):348 – 371, 1991.

[Nol92] D. Nolan. Asymptotics for multivariate trimming. Stochastic Processes and their Applications, 42(1):157 – 169, 1992.

[Pol97] Wolfgang Polonik. Minimum volume sets and generalized quantile processes. Stochastic Processes and their Applications, 69(1):1 – 24, 1997.

[SRT88] X. Saint Raymond and C. Tricot. Packing regularity of sets in n-space. In Proc. Cambridge Philos. Soc, volume 103, pages 133–145. Cambridge Univ Press, 1988.