JOTA manuscript No.
(will be inserted by the editor)

A Characterization of Probabilities with Full Support and the Laplace Method

Simone Cerreia-Vioglio ·

Fabio Maccheroni ·

Massimo Marinacci

Received: date / Accepted: date

Communicated by Nizar Touzi

Abstract We show that a probability measure on a metric space has full support, if, and only if, the set of all probability measures, that are absolutely continuous with respect to it, is dense in the set of all Borel probability measures. We illustrate the result through a general version of Laplace's method, which in turn leads to general stochastic convergence to global maxima.

Keywords Absolute Continuity · Support of a Measure · Laplace Method

Mathematics Subject Classification (2000) 46N30 · 46N10 · 60B05

Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci

Bocconi University and IGIER

Milano, Italy

Corresponding author: fabio.maccheroni@unibocconi.it

1 Introduction

Intuitively, a probability measure on a metric space has full support, if "anything is possible"; formally, if every nonempty open set has positive probability. These measures are important in global optimization, because they allow algorithms – such as Simulated Annealing – to explore the entire space,[1] and in game theory they capture the "anything is possible" notion – when a player is reasoning about opponents. See, e.g., [3,4].

Clearly, when the space is finite, a probability measure has full support, if, and only if, the set of all probability measures, that are absolutely continuous with respect to it, *coincides with* the set of all (Borel) probability measures. We show that, on a metric space, a probability measure has full support, if, and only if, the set of all probability measures, that are absolutely continuous with respect to it, *is dense* in the set of all Borel probability measures.

Since the assumption of full support amounts to (strict) positivity on nonempty open sets, our result sheds light on the notion of strict positivity of a probability measure in the infinite case. In a functional analysis perspective, our result can be regarded as a characterization of strictly positive continuous linear functionals within the duality of bounded continuous functions and signed Borel measures.

To illustrate this result, we prove a general version of Laplace's method.

[1] See, e.g., [1,2].

2 Setup and Preliminaries

We adopt the notation of [5] to which we refer for general background. Let X be a topological space. We denote by $C(X)$ (resp., $C_b(X)$) the vector space of all continuous (resp., continuous and bounded) functions $f : X \to \mathbb{R}$, by $\mathcal{B}(X)$ the Borel sigma-algebra of X, and by $\mathcal{P}(X)$ the set of all Borel probability measures on $\mathcal{B}(X)$ with the topology $\sigma(\mathcal{P}(X), C_b(X))$ of weak convergence.

Given any $\lambda \in \mathcal{P}(X)$, we denote by $\mathcal{P}_\lambda(X)$ (resp., $\mathcal{P}_\lambda^\star(X)$) the collection of all $\mu \in \mathcal{P}(X)$ that are absolutely continuous with respect to λ (resp., that have continuous and bounded density with respect to λ), and by $\ell_\lambda : C_b(X) \to \mathbb{R}$ the positive linear functional $\ell_\lambda(f) = \int_X f d\lambda$.

Definition 2.1 The *support* of $\lambda \in \mathcal{P}(X)$, denoted by $\operatorname{supp}\lambda$, is (if it exists) a closed subset of X with λ-null complement, such that $\lambda(G) > 0$ for all open subsets G of X that have a nonempty intersection with it.

The probability measure λ has *full support* if $\operatorname{supp}\lambda = X$, that is, $\lambda(G) > 0$ for all nonempty open subsets G of X.

If X is the dual of a separable normed space (for example, a reflexive and separable normed space), we endow it with the weak* topology and consider the Borel sigma-algebra generated by this topology. With this topology, compact sets are metrizable and their closed and convex hulls are compact, by Alaoglu's Theorem. The next basic result is a slight modification of Proposition 1.1 in [6].

Proposition 2.1 *Let X be the dual of a separable normed space. If $\mu \in \mathcal{P}(X)$ has bounded support, then there exists a unique element $m \in X$ such that*

$$\langle \phi, m \rangle = \int_X \langle \phi, x \rangle \, d\mu(x) \tag{1}$$

for all linear and continuous functionals $\phi : X \to \mathbb{R}$.

The element m, called *barycenter* of μ, belongs to the closed and convex hull of $\operatorname{supp} \mu$. When X is $\mathbb{R}^n$, the barycenter of a Borel probability measure μ on $\mathbb{R}^n$ that has bounded support is easily seen to be the vector $m = \int_X x d\mu(x)$ obtained by component-wise integration.

Finally, we recall the definition of Γ-convergence (see, e.g., [11]). In reading it, recall that, for each $x \in X$, $\mathcal{N}(x)$ denotes the collection of all open neighborhoods of x.

Definition 2.2 Let $\{F_n\}$ be a sequence of functions from X to $[-\infty, \infty]$. If, for all $x \in X$,

$$\sup_{U \in \mathcal{N}(x)} \liminf_{n \to \infty} \inf_{y \in U} F_n(y) = F(x) = \sup_{U \in \mathcal{N}(x)} \limsup_{n \to \infty} \inf_{y \in U} F_n(y)$$

then $F : X \to [-\infty, \infty]$ is said to be the *Γ-limit* of $\{F_n\}$.

3 Main Result

We state and prove our main result. The equivalence between points (i) and (iv), i.e., between the strict positivity of λ and ℓ_λ, is essentially known and reported here for completeness and perspective.

Theorem 3.1 *Let X be a metric space. The following conditions are equivalent for $\lambda \in \mathcal{P}(X)$:*

(i) λ *has full support* X*;*

(ii) $\mathrm{cl}(\mathcal{P}_\lambda^\star(X)) = \mathcal{P}(X)$*;*

(iii) $\mathrm{cl}(\mathcal{P}_\lambda(X)) = \mathcal{P}(X)$*;*

(iv) ℓ_λ *is strictly positive, i.e.,* $\int_X f d\lambda > 0$ *for all* $0 \neq f \in C_b^+(X)$*.*

Proof If X is a singleton, the statement is trivial. Let us assume that X contains more than one point.

(i) implies (ii). We first show that $\delta_{\bar{x}} \in \mathrm{cl}(\mathcal{P}_\lambda^\star(X))$ for all $\bar{x} \in X$. Let $\bar{x} \in X$, and, for each $n \in \mathbb{N}$, consider the sets B_n and C_n defined by

$$B_n = \left\{x \in X : d(x, \bar{x}) \leq \frac{1}{n}\right\} \text{ and } C_n = \left\{x \in X : d(x, \bar{x}) \geq \frac{2}{n}\right\}.$$

Both sets are closed and clearly $B_n \cap C_n = \varnothing$. If n is large enough, say for all $n \geq \bar{n}$, both sets are nonempty because there exists $x \neq \bar{x}$ in X. By the Urysohn Lemma (e.g., Theorem 2.46 in [5]), it follows that for each $n \geq \bar{n}$ there exists $\varphi_n \in C_b(X)$ such that $\varphi_n(X) \subseteq [0,1]$, $\varphi_n(B_n) = 1$, and $\varphi_n(C_n) = 0$. Since $\bar{x} \in \mathrm{supp}\,\lambda$ and $\varphi_n(\bar{x}) = 1$ (see, e.g., Lemma 12.16 in [5]), it follows that

$$k_n = \int_X \varphi_n d\lambda > 0 \qquad \forall n \geq \bar{n}.$$

Now, for each $n \geq \bar{n}$, set $\psi_n = \varphi_n / k_n$ and define the measure $\lambda_n : \mathcal{B} \to \mathbb{R}$ by $\lambda_n(B) = \int_B \psi_n d\lambda$. Notice that $\lambda_n \in \mathcal{P}_\lambda^\star(X)$ because $\psi_n \in C_b(X)$.

We next show that $\lambda_n \to \delta_{\bar{x}}$. Define $S_n = \{x \in X : d(x, \bar{x}) \leq 2/n\}$ for all $n \geq \bar{n}$. Notice that $S_n^c \subseteq C_n$ so that $1 = \int_{S_n} \psi_n d\lambda + \int_{S_n^c} \psi_n d\lambda = \int_{S_n} \psi_n d\lambda = \lambda_n(S_n)$ for all $n \geq \bar{n}$. Consider an open subset G of X. We have two cases:

1. $\bar{x} \notin G$. It follows that $\liminf \lambda_n(G) \geq 0 = \delta_{\bar{x}}(G)$.
2. $\bar{x} \in G$. For $n \geq \bar{n}$ large enough, say $n \geq \bar{m}$, we have that $S_n \subseteq G$. Then, for all $n \geq \bar{m}$, $\lambda_n(G) \geq \lambda_n(S_n) \geq 1$, yielding that $\liminf \lambda_n(G) \geq 1 = \delta_{\bar{x}}(G)$.

In both cases, $\liminf \lambda_n(G) \geq \delta_{\bar{x}}(G)$ holds. Since G was an arbitrarily chosen open subset of X, by the Portmanteau Theorem (e.g., Theorem 15.3 in [5]) it follows that $\lambda_n \to \delta_{\bar{x}}$.

Since $\bar{x}$ was arbitrarily chosen in X, we have that $\{\delta_x\}_{x\in X} \subseteq \mathrm{cl}(\mathcal{P}_\lambda^\star(X))$. Since $\mathcal{P}_\lambda^\star(X)$ is convex, then $\mathrm{cl}(\mathcal{P}_\lambda^\star(X))$ is closed and convex, it follows that $\mathrm{cl}(\mathcal{P}_\lambda^\star(X)) \supseteq \mathrm{cl}\left(\mathrm{co}\left(\{\delta_x\}_{x\in X}\right)\right)$. But $\mathrm{co}\left(\{\delta_x\}_{x\in X}\right)$ is dense in $\mathcal{P}(X)$, we conclude that $\mathcal{P}(X) \supseteq \mathrm{cl}(\mathcal{P}_\lambda^\star(X)) \supseteq \mathrm{cl}\left(\mathrm{co}\left(\{\delta_x\}_{x\in X}\right)\right) = \mathcal{P}(X)$.

(ii) implies (iii). This follows from $\mathcal{P}_\lambda^\star(X) \subseteq \mathcal{P}_\lambda(X)$.

(iii) implies (iv). By contradiction, assume that $\mathrm{cl}(\mathcal{P}_\lambda(X)) = \mathcal{P}(X)$ and ℓ_λ is not strictly positive. In this case, there exists $g \in C_b^+(X) \setminus \{0\}$ such that $\int_X g d\lambda = 0$. Consider the open set $G = \{x \in X : g(x) > 0\} \neq \varnothing$. Since $\int_X g d\lambda = 0$, then $\lambda(\{x \in X : g(x) > 0\}) = 0$, that is, $\lambda(G) = 0$. Consider $\bar{x} \in G$. Since $\mathrm{cl}(\mathcal{P}_\lambda(X)) = \mathcal{P}(X)$, there exists a net $\{\lambda_\alpha\} \subseteq \mathcal{P}_\lambda(X)$ such that $\lambda_\alpha \to \delta_{\bar{x}}$. For each α, since λ_α is absolutely continuous with respect to λ, we have that $\lambda_\alpha(G) = 0$. Since $\lambda_\alpha \to \delta_{\bar{x}}$, by the Portmanteau Theorem, we have that $0 = \liminf \lambda_\alpha(G) \geq \delta_{\bar{x}}(G) = 1$, a contradiction.

(iv) implies (i). By contradiction, assume that ℓ_λ is strictly positive and there exists a nonempty open subset G of X with $\lambda(G) = 0$. Consider $\bar{x} \in G$. By the Urysohn Lemma, and since G^c is closed and nonempty, there exists

$\varphi \in C_b(X)$ such that $\varphi(X) \subseteq [0,1]$, $\varphi(\bar{x}) = 1$, and $\varphi(x) = 0$ for all $x \in G^c$. Since $\varphi \in C_b^+(X) \setminus \{0\}$, it follows that

$$0 < \ell_\lambda(\varphi) = \int_X \varphi d\lambda = \int_G \varphi d\lambda + \int_{G^c} \varphi d\lambda = 0$$

a contradiction. □

Finally, observe that the result depends only on the topology of X, so we could have used the term metrizable, rather than metric, throughout.

3.1 **Feasibility and Arbitrages**

Lemma 5.6 of [7] shows that, if λ is a Borel full support measure on a Polish space, then

$$\mathrm{cl}\left(\mathcal{P}_\lambda^e(X)\right) = \mathcal{P}(X)$$

where $\mathcal{P}_\lambda^e(X)$ is the collection of all $\mu \in \mathcal{P}(X)$ that are *equivalent* to λ. In turn, this allows them to show that, whenever there exists a full support martingale measure (the market is feasible), the set of these measures is dense in $\mathcal{P}(X)$.

Notice that $\mathrm{cl}\left(\mathcal{P}_\lambda^e(X)\right) = \mathrm{cl}\left(\mathcal{P}_\lambda(X)\right)$ on any metric space X. Therefore, point (iii) of Theorem 3.1 can be replaced by

$$\mathrm{cl}\left(\mathcal{P}_\lambda^e(X)\right) = \mathcal{P}(X).$$

In this sense, our implication (i) $\Longrightarrow$ (iii) extends Lemma 5.6 in [7] to general metric spaces and extends their feasibility results. Another difference is that our result also yields the opposite implication (iii) $\Longrightarrow$ (i).

Similarly, $\mathcal{P}_\lambda^\star(X)$ can be replaced with $\mathcal{P}_\lambda^{e\star}(X) = \mathcal{P}_\lambda^e(X) \cap \mathcal{P}_\lambda^\star(X)$ in point (ii) above.

4 Illustration: Laplace Method

Consider the optimization problem

$$\max_x u(x) \quad \text{sub } x \in K \tag{2}$$

where $u : X \to \mathbb{R}$ is a continuous function and K is a compact and metrizable set.

Laplace's method is a fundamental method to find maximum values and maximizers of this general optimization problem. For this reason, it plays an important role in many applications (see, e.g., [8] for an introductory overview and some relevant references). To illustrate the scope of our main result, here we establish a general abstract version of this classic method. A related result appears in [9], though in a different setup and with an altogether different approach.

In the statement, we denote by $\overset{w}{\Longrightarrow}$ the $\sigma(\mathcal{P}(X), C_b(X))$-convergence and by δ_x the Dirac probability measure concentrated on a point $x \in X$.

Theorem 4.1 *Let X be a topological space, $u : X \to \mathbb{R}$ a continuous function, λ a Borel probability measure with compact and metrizable support K, and $\{s_n\} \subseteq]0, \infty[$ a divergent sequence. Then*

$$\frac{1}{s_n} \log \int_X e^{s_n u} d\lambda \to \max_K u \qquad \text{as } n \to \infty. \tag{3}$$

Moreover, if u has a unique maximizer x^u in K, then

$$\mu_n \overset{w}{\Longrightarrow} \delta_{x^u} \qquad \text{as } n \to \infty \tag{4}$$

where μ_n *is, for each* $n \in \mathbb{N}$*, defined by*

$$\mu_n(B) = \frac{\int_B e^{s_n u} d\lambda}{\int_X e^{s_n u} d\lambda} \qquad \forall B \in \mathcal{B}(X). \tag{5}$$

Proof It is sufficient to prove our result when $\{s_n\}$ is increasing. First, assume $K = X$, that is, X is compact and metrizable, and λ has full support. In this case, $\sigma(\mathcal{P}(X), C_b(X)) = \sigma(\mathcal{P}(X), C(X))$ is the relative weak* topology on $\mathcal{P}(X)$, and $\mathcal{P}(X)$ is compact and metrizable with respect to it (see Corollary 14.15 and Theorem 15.11 in [5]). Denote by

$$R(\mu\|\lambda) = \begin{cases} \int_X \frac{d\mu}{d\lambda} \log\left(\frac{d\mu}{d\lambda}\right) d\lambda, & \text{if } \mu \ll \lambda, \\ \infty, & \text{else,} \end{cases}$$

the relative entropy of any μ in $\mathcal{P}(X)$ with respect to λ (see Section 1.4 in [10]).

For each $n \in \mathbb{N}$, set $f_n = -s_n u$ and observe that, by Proposition 1.4.2 in [10],

$$-\log \int_X e^{-f_n} d\lambda = \min_{\mu \in \mathcal{P}(X)} \left\{ R(\mu\|\lambda) + \int_X f_n d\mu \right\}$$

and the minimum of this variational formula is uniquely attained at the element μ_n of $\mathcal{P}(X)$ given by

$$\mu_n(B) = \frac{\int_B e^{-f_n(x)} d\lambda(x)}{\int_X e^{-f_n(y)} d\lambda(y)}$$

for all Borel subsets B of X. Recalling our substitution

$$\begin{aligned} -\frac{1}{s_n} \log \int_X e^{s_n u} d\lambda &= \frac{1}{s_n}\left[-\log \int_X e^{-f_n} d\lambda\right] = \frac{1}{s_n} \min_{\mu \in \mathcal{P}(X)} \left\{ R(\mu\|\lambda) - s_n \int_X u d\mu \right\} \\ &= \min_{\mu \in \mathcal{P}(X)} \left\{ \frac{1}{s_n} R(\mu\|\lambda) - \int_X u d\mu \right\}. \end{aligned}$$

For each $n \in \mathbb{N}$, the function $F_n : \mathcal{P}(X) \to]-\infty, \infty]$ defined by

$$F_n(\mu) = \frac{1}{s_n} R(\mu \| \lambda) - \int_X u d\mu \qquad \forall \mu \in \mathcal{P}(X)$$

is weak* lower semicontinuous on $\mathcal{P}(X)$ (see Lemma 1.4.3 in [10] and Proposition 1.9 in [11]). Moreover, the sequence $\{F_n\}$ is decreasing and pointwise converges to

$$F_\infty(\mu) = \chi_{\operatorname{dom} R(\cdot\|\lambda)}(\mu) - \int_X u d\mu \qquad \forall \mu \in \mathcal{P}(X). \tag{6}$$

By Proposition 5.7 in [11], this sequence Γ-converges to the weak* lower semicontinuous envelope $\mathrm{sc}^- F_\infty$ of F_∞. Since $U : \mu \mapsto \int_X u d\mu$ is continuous and everywhere finite on $\mathcal{P}(X)$, by Proposition 3.7 and Example 3.4 in [11]

$$\left(\mathrm{sc}^- F_\infty\right)(\mu) = \left(\mathrm{sc}^- \chi_{\operatorname{dom} R(\cdot\|\lambda)}\right)(\mu) - \int_X u d\mu = \chi_{\operatorname{cl}(\operatorname{dom} R(\cdot\|\lambda))}(\mu) - \int_X u d\mu.$$

For each $\mu \in \mathcal{P}^\star_\lambda(X)$, $d\mu/d\lambda$ is bounded and continuous, hence there exists $k \geq 0$ such that $0 \leq d\mu/d\lambda \leq k$ and so

$$-\frac{1}{e} \leq \frac{d\mu}{d\lambda} \log\left(\frac{d\mu}{d\lambda}\right) \leq k^2 \implies R(\mu\|\lambda) < \infty \implies \mu \in \operatorname{dom} R(\cdot\|\lambda).$$

Therefore, $\mathcal{P}^\star_\lambda(X) \subseteq \operatorname{dom} R(\cdot\|\lambda)$ and so, by Theorem 3.1,

$$\mathcal{P}(X) = \operatorname{cl}(\mathcal{P}^\star_\lambda(X)) \subseteq \operatorname{cl}(\operatorname{dom} R(\cdot\|\lambda)) = \mathcal{P}(X).$$

Summing up, F_n Γ-converges to $-\int_X u d\mu$. By Theorem 7.4 in [11], this implies

$$\begin{aligned}
\lim_{n\to\infty} \min_{\mu \in \mathcal{P}(X)} \left\{\frac{1}{s_n} R(\mu\|\lambda) - \int_X u d\mu\right\} &= \min_{\mu \in \mathcal{P}(X)} \left\{-\int_X u d\mu\right\} \\
&= -\max_{\mu \in \mathcal{P}(X)} \left\{\int_X u d\mu\right\} \\
&= -\max_{x \in X} u(x).
\end{aligned}$$

But, for all $n \in \mathbb{N}$ we have

$$\min_{\mu \in \mathcal{P}(X)} \left\{ \frac{1}{s_n} R(\mu \| \lambda) - \int_X u d\mu \right\} = -\frac{1}{s_n} \log \int_X e^{s_n u} d\lambda.$$

So, (3) holds.

Moreover, if u has a unique maximizer x^u in X, then U has δ_{x^u} as its unique maximizer. In fact, if $\mu \in \mathcal{P}(X) \setminus \{\delta_{x^u}\}$, then $\mu(X \setminus \{x^u\}) > 0$, and so

$$\begin{aligned}
\int_X u d\delta_{x^u} - \int_X u d\mu &= \int_X (u(x^u) - u(x)) d\mu(x) \\
&= \int_{\{x^u\}} (u(x^u) - u(x)) d\mu(x) \\
&\quad + \int_{X \setminus \{x^u\}} (u(x^u) - u(x)) d\mu(x) > 0
\end{aligned}$$

since the first summand is null, the second is strictly positive.[2] Since $\mathcal{P}(X)$ is compact, the sequence F_n is equi-coercive (see Definition 7.6 in [11]); in addition, it Γ-converges to $-U$ with unique minimum point δ_{x^u} in $\mathcal{P}(X)$. For each n, the probability measure μ_n is a minimizer for F_n in $\mathcal{P}(X)$. By Corollary 7.24 in [11], μ_n weak* converges to δ_{x^u}.

In the general case, consider the compact and metrizable space K, the continuous function $w = u_{|K}$, and the Borel probability measure $\nu = \lambda_{|K}$. It is easy to show that ν has full support on K. In fact, if O is a nonempty open subset of K, there exists an open subset G of X such that $\varnothing \neq O = G \cap K = G \cap \operatorname{supp} \lambda$; by definition of support, it follows $\lambda(G) > 0$, but then

[2] $\int_{X \setminus \{x^u\}} (u(x^u) - u(x)) d\mu(x) = 0$ would imply $\mu(\{x \in X \setminus \{x^u\} : u(x^u) - u(x) > 0\}) = 0$, a contradiction because $u(x^u) - u(x) > 0$ for all $x \in X \setminus \{x^u\}$, thus $\{x \in X \setminus \{x^u\} : u(x^u) - u(x) > 0\} = X \setminus \{x^u\}$.

$\nu(O) = \lambda(G \cap \operatorname{supp} \lambda) = \lambda(G \cap \operatorname{supp} \lambda) + \lambda(G \cap (\operatorname{supp} \lambda)^c) = \lambda(G) > 0$. The previous part of the proof implies

$$\frac{1}{s_n} \log \int_K e^{s_n w} d\nu \to \max_K w \qquad \text{as } n \to \infty.$$

But $s_n^{-1} \log \int_X e^{s_n u} d\lambda = s_n^{-1} \log \int_K e^{s_n w} d\nu$ for all $n \in \mathbb{N}$ and $\max_K u = \max_K w$, thus (3) holds.

Moreover, if u has a unique maximizer x^u in K, again by the previous part of the proof we can consider the sequence $\{\rho_n\}$ of probability measures defined by

$$\rho_n(L) = \frac{\int_L e^{s_n w} d\nu}{\int_K e^{s_n w} d\nu} \qquad \forall n \in \mathbb{N}$$

for all Borel subsets L of K, and have that, given any $g \in C(K)$,

$$\int_K g d\rho_n \to g(x^u) \qquad \text{as } n \to \infty.$$

But for each $f \in C_b(X)$, $f_{|K} \in C(K)$ and $\int_X f d\mu_n = \int_K f_{|K} d\rho_n$ for all $n \in \mathbb{N}$, then the sequence $\{\mu_n\}$, defined by (5), $\sigma(\mathcal{P}(X), C_b(X))$-converges to δ_{x^u}.

□

If X is the dual of a separable normed space and is endowed with the weak* topology, then the boundedness of the support of λ is equivalent to its compactness being the support closed by definition, and – as we observed in the previous section – each μ_n has a barycenter m_n in the weak* closed and convex hull of $K = \operatorname{supp} \lambda$. Next proposition shows that these barycenters weakly*-converge to the maximizer. Here $\xrightharpoonup{w^*}$ denotes weak*-convergence. The simple proof is omitted.

Proposition 4.1 *Let X be the dual of a separable normed space endowed with the weak* topology, $u : X \to \mathbb{R}$ a continuous function, λ a Borel probability measure with compact support K, $\{s_n\} \subseteq]0, \infty[$ a divergent sequence, $\{\mu_n\}$ the sequence of Borel probability measures defined in (5), and $\{m_n\} \subseteq X$ the sequence of their barycenters. If u has a unique maximizer x^u in K, then*

$$m_n \overset{w^*}{\rightharpoonup} x^u \qquad \text{as } n \to \infty. \tag{7}$$

In particular, if X is a separable and reflexive Banach space, then its weak and weak* topologies coincide, therefore m_n weakly converges to x^u. Clearly, the sequence of barycenters is included in K if this set is convex.

When X is $\mathbb{R}^n$, γ is a Borel measure, and $K \neq \varnothing$ is a compact set such that $\gamma(G \cap K) \in]0, \infty[$ for all open subsets G of $\mathbb{R}^n$ having nonempty intersection with it, we have

$$\frac{1}{s_n} \log \frac{1}{\gamma(K)} \int_K e^{s_n u(x)} d\gamma(x) \to \max_K u \qquad \text{as } n \to \infty \tag{8}$$

and, if x^u is the unique maximizer of u on K,

$$m_n = \frac{\int_K e^{s_n u(x)} x d\gamma(x)}{\int_K e^{s_n u(y)} d\gamma(y)} \to x^u. \tag{9}$$

This convergence in $\mathbb{R}^n$ has been first established by [12,13] (see p. 22 in [2]). The weak* convergence (7) thus substantially generalizes his results.

5 Conclusions

We have characterized fully supported probability measures on *metric spaces* and derived a general stochastic convergence result to global maxima. This result is especially powerful for functions with a *unique* maximizer on the constraint set. Natural directions of investigation include extending our results to general topological spaces or, more interestingly perhaps, extending Theorem 4.1 and Proposition 4.1 to functions u that admit more than one maximizer.

Acknowledgements We thank Isabella Morgan Wolfskeil for excellent research assistance. Simone Cerreia-Vioglio, and Fabio Maccheroni and Massimo Marinacci gratefully acknowledge the financial support of ERC (grants SDDM-TEA and INDIMACRO, respectively).

References

1. Romeijn, H.E., Smith, R.L.: Simulated Annealing for constrained global optimization. Journal of Global Optimization, 5, 101-126, 1994.
2. Hiriart-Urruty, J.-B.: Conditions for global optimality. In: Horst, R., Pardalos, P.M. (eds.): Handbook of Global Optimization, pp. 1-26. Springer, New York, 1995.
3. De Bruin, B.: Explaining Games: the Epistemic Programme in Game Theory. Springer, New York, 2010.
4. Dekel, E., Siniscalchi, M., Epistemic game theory. In: Young, P., Zamir, S. (eds.): Handbook of Game Theory, vol. 4, pp. 629-702. Elsevier, New York, 2014.
5. Aliprantis, C.D., Border, K.C.: Infinite Dimensional Analysis, 3^{rd} ed.. Springer, New York, 2006.
6. Phelps, R.R.: Lectures on Choquet's Theorem, 2^{nd} ed.. Springer, Heidelberg, 2001.
7. Burzoni, M., Frittelli, M., Maggis, M.: Universal arbitrage aggregator in discrete-time markets under uncertainty. Finance and Stochastics, 20, 1-50, 2016.
8. Parpas, P., Rustem, B.: Laplace method and applications to optimization problems. In Floudas, C.A., Pardalos, P.M. (eds.): Encyclopedia of Optimization, pp. 1818-1822. Springer, New York, 2009.
9. Hwang, C.R.: Laplace's method revisited: weak convergence of probability measures. Annals of Probability, 8, 1177-1182, 1980.

10. Dupuis, P., Ellis, R.S.: A Weak Convergence Approach to the Theory of Large Deviations. John Wiley & Sons, New York, 2011.

11. Dal Maso, G.: An Introduction to Γ-convergence. Birkhäuser, Boston, 1993.

12. Pincus, M.: A closed form solution of certain programming problems. Operations Research, 16, 690-694, 1968.

13. Pincus, M.: A Monte Carlo method for the approximate solution of certain types of constrained optimization problems. *Operations Research*, 18, 1225-1228, 1970.