Journal of Convex Analysis
Volume 30 (2023), No. 3, 917–936
https://doi.org/10.68381/jca30043

Epi-Convergence of Expectation Functions under Varying Measures and Integrands

Eugene A. Feinberg

Dept. of Applied Mathematics and Statistics, Stony Brook University, New York, U.S.A.

Pavlo O. Kasyanov

Institute for Applied System Analysis, National Technical University of Ukraine "Igor Sikorsky Polytechnic Institute", Kyiv, Ukraine

Johannes O. Royset

Operations Research Department, Naval Postgraduate School, Monterey, U.S.A.
joroyset@nps.edu

Dedicated to Roger J-B Wets on the occasion of his 85th birthday.

Received: August 7, 2022
Accepted: August 17, 2022

For expectation functions on metric spaces, we provide sufficient conditions for epi-convergence under varying probability measures and integrands, and examine applications in the area of sieve estimators, mollifier smoothing, PDE-constrained optimization, and stochastic optimization with expectation constraints. As a stepping stone to epi-convergence of independent interest, we develop parametric Fatou's lemmas under mild integrability assumptions. In the setting of Suslin metric spaces, the assumptions are expressed in terms of Pasch-Hausdorff envelopes. For general metric spaces, the assumptions shift to semicontinuity of integrands also on the sample space, which then is assumed to be a metric space.

Keywords: Epi-convergence, expectation function, stochastic optimization, sieve estimators, mollifers.

2020 Mathematics Subject Classification: 90C15, 60F17.

1. Introduction

It is well known that epi-convergence is a suitable framework for analyzing approximating optimization problems. It furnishes guarantees about near-minimizers, level-sets, and minimum values for the problems as compared to those of a limiting, actual problem; see for example [2, 4, 22]. A recent summary of properties for functions defined on metric spaces is given by [25]. Epi-convergence is therefore central in the development of optimization algorithms, which often rely on approximations, and in stability analysis of optimization problems. It also emerges as a means to establish strong duality for dual problems obtained via general perturbation schemes [27, Section 5.F]. In this paper, we provide sufficient conditions for epi-convergence of expectation functions under varying measures and integrands.

The study of expectation functions is critical for stochastic optimization and statistical estimation. The first systematic examination of epi-convergence for such functions appears to be [9], which also includes a review of the earlier literature.

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

The setting is limited to expectation functions defined on $\mathbb{R}^n$, a fixed integrand, and probabilities converging weakly under a tightness assumption. An extension to separable reflexive Banach spaces, with a focus on Mosco-epi-convergence, is developed in [20] under similar assumptions. For separable Banach spaces, [3] establishes epi-convergence for empirical distributions (generated by pairwise independent sampling) and a fixed integrand, and also includes extensions to convergence in the stronger Attouch-Wets topology. A further extension to complete separable metric spaces is given by [1], again under a fixed integrand and empirical distributions obtained by independent sampling. A nearly identical setting is considered in [16], which also develops statistical applications. A specialization to the case of convex integrands is provided by [18]; see also [29]. Going beyond empirical distributions, [7, 19] establish epi-convergence of expectation functions for ergodic processes. Recently, [8] reexamines epi-convergence under weakly converging probabilities as in [9], but now in the general setting of Hausdorff topological spaces. Without making explicit assumptions on the sequence of probabilities, but instead expressing assumptions in terms of Pasch-Hausdorff envelopes, [17] develops a general framework for establishing epi-convergence of expectation functions defined on separable metric spaces.

While we make precise comparison below, the present paper is distinct by considering *varying integrands*; the above studies keep the integrand fixed as the probabilities change. Varying integrands appear in applications, for example, as the result of computationally motivated approximations that need to be consider in conjunction with the varying probabilities. Sieves in statistical estimation [15, 26] as well as discretization of an infinite-dimensional space of controls and numerical solution of differential equations [6, 21, 28] can be viewed from this angle. Moreover, we establish conditions under which a lower limit of approximating expectation functions is bounded from below by the actual expectation function in the sense of a "parametric" Fatou's lemma with varying measure. This property furnishes the "first half" of epi-convergence (a limit gives the "second half"), but we argue that it stands on its own as a useful property, especially in difficult cases when epi-convergence might be beyond reach.

To formalize the setting, we let $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ and denote by $\mathcal{B}(\overline{\mathbb{R}})$ its Borel σ-algebra. For probability space $(\Omega, \mathcal{A}, \mu)$ and $(\mathcal{A}, \mathcal{B}(\overline{\mathbb{R}}))$-measurable function $g : \Omega \to \overline{\mathbb{R}}$, the integral

$$\int_\Omega g(\omega)\mu(d\omega) := \int_\Omega g_+(\omega)\mu(d\omega) - \int_\Omega g_-(\omega)\mu(d\omega),$$

where $g_+(\omega) := \max\{g(\omega), 0\}$, $g_-(\omega) := -\min\{g(\omega), 0\}$, and we use the conventions $+\infty - \alpha = +\infty$ and $\beta - (+\infty) = -\infty$ hold for $\alpha \in \overline{\mathbb{R}}$ and $\beta \in \mathbb{R}$ everywhere in the paper. This integral is an expectation, and we often use the notation

$$\mathbb{E}^\mu\big[g(\omega)\big] := \int_\Omega g(\omega)\mu(d\omega).$$

Let $\hat{\mathcal{A}}$ be the μ-completion of the σ-algebra $\mathcal{A}$. Then the integral is always defined for an $\hat{\mathcal{A}}$-measurable function $g\colon (\Omega, \hat{\mathcal{A}}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, and, if both integrals of g_+ and g_- are equal to $+\infty$, then the integral of g is also equal to $+\infty$.

Let $(\mathbb{X}, d_{\mathbb{X}})$ be a metric space and $(\Xi, \mathcal{F})$ be a measurable space. Then, a function $f\colon \Xi \times \mathbb{X} \to \overline{\mathbb{R}}$ is an *integrand* if $f(\cdot, x)$ is $(\mathcal{F}, \mathcal{B}(\overline{\mathbb{R}}))$-measurable for each $x \in \mathbb{X}$. An integrand f together with a probability P on $(\Xi, \mathcal{F})$ define an *expectation function* $E^P[f] : \mathbb{X} \to \overline{\mathbb{R}}$ given by

$$E^P\big[f\big](x) := \mathbb{E}^P\big[f(\xi, x)\big], \quad x \in \mathbb{X}.$$

If Ξ is a metric space, we always consider $\mathcal{F}$ being its Borel σ-algebra $\mathcal{B}(\Xi)$.

Let $\mathbb{N} = \{1, 2, \dots\}$. We consider sequences of integrands $\{f, f^\nu : \Xi \times \mathbb{X} \to \overline{\mathbb{R}},\ \nu \in \mathbb{N}\}$ and probabilities $\{P, P^\nu, \nu \in \mathbb{N}\}$, all defined on $(\Xi, \mathcal{F})$, and examine sufficient conditions for epi-convergence of the resulting sequence of expectation functions $\{E^{P^\nu}[f^\nu] : \mathbb{X} \to \overline{\mathbb{R}}, \nu \in \mathbb{N}\}$ to the expectation function $E^P[f] : \mathbb{X} \to \overline{\mathbb{R}}$. We also examine when $x^\nu \to x$ implies $\liminf_{\nu\to+\infty} E^{P^\nu}[f^\nu](x^\nu) \ge E^P[f](x)$, which can be viewed as a "functional" Fatou's lemma under varying measure.

Section 2 presents preliminary facts. The main theorems appear in Section 3. Section 4 provides motivating applications. An appendix provides an additional proof.

2. Preliminaries

We recall that $\{h^\nu : \mathbb{X} \to \overline{\mathbb{R}}, \nu \in \mathbb{N}\}$ *epi-converges* to $h : \mathbb{X} \to \overline{\mathbb{R}}$, written as $h = \text{e-lim}_{\nu\to+\infty} h^\nu$, if at each $x \in \mathbb{X}$, the following two conditions hold:

(i) if $\{x^\nu\}_{\nu\in\mathbb{N}}$ converges to x, then $h(x) \le \liminf\limits_{\nu\to+\infty} h^\nu(x^\nu)$;

(ii) there exists a sequence $\{x^\nu\}_{\nu\in\mathbb{N}}$ convergent to x such that $h(x) = \lim\limits_{\nu\to+\infty} h^\nu(x^\nu)$.

For a function $h\colon \mathbb{X} \to \overline{\mathbb{R}}$, its *epigraph* is given by $\operatorname{epi} h := \{(x, \alpha) \in \mathbb{X} \times \mathbb{R} \mid h(x) \le \alpha\}$. The function is *lower semicontinuous* (lsc) if $\operatorname{epi} h$ is a closed subset of $\mathbb{X} \times \mathbb{R}$ in the product topology. It is *upper semicontinuous* (usc) if $-h$ is lsc. The *lower regularization* of h, denoted by $\underline{h} : \mathbb{X} \to \overline{\mathbb{R}}$, is defined as

$$\underline{h}(x) := \liminf_{x'\to x} h(x'), \quad x \in \mathbb{X}.$$

Thus, $\underline{h} : \mathbb{X} \to \overline{\mathbb{R}}$ is lsc. If h is lsc, then $\underline{h} = h$.

The *Pasch-Hausdorff envelope* of $h : \mathbb{X} \to \overline{\mathbb{R}}$ with parameter $\kappa \in [0, +\infty)$ is the function $h_\kappa : \mathbb{X} \to \overline{\mathbb{R}}$ defined as

$$h_\kappa(x) := \inf_{x'\in\mathbb{X}} \big\{h(x') + \kappa d_{\mathbb{X}}(x, x')\big\}, \quad x \in \mathbb{X}. \tag{1}$$

Trivially, when $0 \le \kappa_1 \le \kappa_2 < +\infty$, one has

$$h_{\kappa_1}(x) \le h_{\kappa_2}(x) \le h(x) \quad \forall x \in \mathbb{X}.$$

The definition of an infimum enables us to conclude that $h_\kappa(x) = \underline{h}_\kappa(x)$ for each $\kappa \in [0, +\infty)$ and each $x \in \mathbb{X}$, where $\underline{h}_\kappa$ is the Pasch-Hausdorff envelope of $\underline{h}$. If $\operatorname{epi} h \ne \emptyset$ and there are $\alpha > 0$, $\beta \in \mathbb{R}$, and $\bar{x} \in \mathbb{X}$ such that $h(x) + \alpha d_{\mathbb{X}}(x, \bar{x}) + \beta \ge 0$ for all $x \in \mathbb{X}$, then (see for example [16, Proposition 3.3])

$$\forall x \in \mathbb{X} :\ h_\kappa(x) \uparrow \underline{h}(x) \text{ as } \kappa \to +\infty \tag{2}$$

and $h_\alpha(\bar{x}) > -\infty$.

In fact, for $\kappa > \alpha$, h_κ is real-valued and Lipschitz continuous with modulus κ, i.e.,

$$\left|h_\kappa(x) - h_\kappa(x')\right| \le \kappa d_{\mathbb{X}}(x, x') \quad \forall x, x' \in \mathbb{X}. \tag{3}$$

Furthermore, if $\{h^\nu : \mathbb{X} \to \overline{\mathbb{R}}, \nu \in \mathbb{N}\}$ is a sequence of functions such that for some $\kappa_0 \in \mathbb{N}$ and $x_0 \in \mathbb{X}$,

$$\liminf_{\nu \to +\infty} h^\nu_{\kappa_0}(x_0) > -\infty, \tag{4}$$

then [16, Proposition 3.4] implies that for each $x \in \mathbb{X}$

$$\liminf_{(\nu,y)\to(+\infty,x)} h^\nu(y) = \sup_{\kappa\in\mathbb{N}} \liminf_{\nu\to+\infty} h^\nu_\kappa(x). \tag{5}$$

For $f : \Xi \times \mathbb{X} \to \overline{\mathbb{R}}$, $\xi \in \Xi$, and $\kappa \in [0, +\infty)$, we abuse notation slightly and write f_κ for the Pasch-Hausdorff envelope of $f(\xi, \cdot)$, i.e., $f_\kappa : \Xi \times \mathbb{X} \to \overline{\mathbb{R}}$ is given by $f_\kappa(\xi, x) = \inf_{x' \in \mathbb{X}}\{f(\xi, x') + \kappa d_{\mathbb{X}}(x', x)\}$. In view of [16, Propositions 4.2], if $\mathbb{X}$ is a Suslin space and the function f is measurable, then for each $x \in \mathbb{X}$ the function $\xi \mapsto f_\kappa(\xi, x)$ is $\hat{\mathcal{F}}$-measurable, where $\hat{\mathcal{F}}$ is the P-completion of the sigma-algebra $\mathcal{F}$. Likewise, f^ν_κ stems from $f^\nu : \Xi \times \mathbb{X} \to \overline{\mathbb{R}}$.

For probability spaces $(\Xi, \mathcal{F}, P^\nu)$, $\nu \in \mathbb{N}$, let us consider the probability spaces $(\Xi, \hat{\mathcal{F}}^\nu, P^\nu)$, where $\hat{\mathcal{F}}^\nu$ are P^ν-completions of the σ-algebra $\mathcal{F}$. As explained in the previous paragraph, if $f^\nu : \Xi \times \mathbb{X} \to \overline{\mathbb{R}}$ are measurable functions, $\nu \in \mathbb{N}$, then for each $\kappa, \nu \in \mathbb{N}$ and for each $x \in \mathbb{X}$, the function $\xi \mapsto f^\nu_\kappa(\xi, x)$ is $\hat{\mathcal{F}}^\nu$-measurable on Ξ if $\mathbb{X}$ is a Suslin space.

3. Main Results

This section provides two sufficient conditions for epi-convergence of expectation functions. Theorem 3.2 extends [17, Theorem 3.1] by considering varying integrands. As in [17], the main assumptions are expressed in terms of Pasch-Hausdorff envelopes. Theorem 3.7 establishes epi-convergence in the setting with weakly converging probabilities and extends [8, Theorem 1] to varying integrands. (While we restrict the attention to metric spaces, [8] deals with Hausdorff topological spaces; see below for a detailed comparison.) Here, we dispense of Pasch-Hausdorff envelopes and instead rely on Fatou's lemmas for weakly convergent probabilities [11, 12, 14].

In contrast to [17], [8], and earlier studies, we also bring forth bounds on lower epi-limits, which emerge as useful in their own right as demonstrated in Section 4. These bounds appear as Lemma 3.1 and Theorem 3.3.

Lemma 3.1. (Parametric Fatou's lemma with varying probabilities) *Let $(\Xi, \mathcal{F})$ be a measurable space, and $\mathbb{X}$ be a Suslin space. For probability measures $\{P, P^\nu, \nu \in \mathbb{N}\}$ on $(\Xi, \mathcal{F})$ and functions $\{f, f^\nu : \Xi \times \mathbb{X} \to \overline{\mathbb{R}}, \nu \in \mathbb{N}\}$, suppose that:*

(i) *$\{f, f^\nu, \nu \in \mathbb{N}\}$ are measurable functions, and $f(\xi, \cdot)$ is lsc for P-a.e. $\xi \in \Xi$;*

(ii) *there exist a countable dense subset $\mathbb{X}_0$ of $\mathbb{X}$ and an integer $\kappa_0 \in \mathbb{N}$ such that*

$$\liminf_{\nu\to+\infty} E^{P^\nu}\left[f^\nu_\kappa\right](x_0) \ge E^P\left[f_\kappa\right](x_0) > -\infty \tag{6}$$

for each $x_0 \in \mathbb{X}_0$ and each $\kappa \in [\kappa_0, +\infty)$.

Then for every $x \in \mathbb{X}$ we have

$$\liminf_{(\nu,y)\to(+\infty,x)} E^{P^\nu}\big[f^\nu\big](y) \ge E^{P}\big[f\big](x),$$

with the right-hand side also exceeding $-\infty$ when there exists $\tilde{x} \in \mathbb{X}$ such that

$$\liminf_{(\nu,y)\to(+\infty,\tilde{x})} E^{P^\nu}\big[f^\nu\big](y) < +\infty.$$

Proof. The proof of the lemma rethinks the proof of [17, Theorem 3.1]; see Appendix A. □

We next turn to epi-convergence.

Theorem 3.2. (Epi-convergence of expectation functions) *Let $(\Xi, \mathcal{F})$ be a measurable space, and $\mathbb{X}$ be a Suslin space. For probability measures $\{P, P^\nu, \nu \in \mathbb{N}\}$ on $(\Xi, \mathcal{F})$ and functions $\{f, f^\nu : \Xi \times \mathbb{X} \to \overline{\mathbb{R}}\}$, suppose that assumptions* (i)–(ii) *of Lemma* 3.1 *hold. Moreover, suppose that there is a countable dense subset $\mathbb{X}^0$ of $\mathbb{X}$ such that for each $x^0 \in \mathbb{X}^0$ we have*

$$\limsup_{\nu\to+\infty} E^{P^\nu}\big[f^\nu\big](x^0) \le E^{P}\big[f\big](x^0).$$

Then

$$\operatorname*{e-lim}_{\nu\to+\infty} E^{P^\nu}\big[f^\nu\big] = E^{P}\big[f\big].$$

Proof. The liminf-condition in the definition of epi-convergence is immediately furnished by Lemma 3.1. The limit-condition of the definition is established in the second part of the proof of [17, Theorem 3.1] with minor technical clarifications. □

Next, we turn to the special case when the probabilities P^ν converge weakly to P, which is common in applications. As is well known, weak convergence may arise for empirical distributions, in the context of central limit theorems, and for ergodic processes. But, weak convergence may also define a topology on spaces of distribution functions for the analysis of distributionally-robust optimization problems [24].

We denote by $\mathbf{I}B$ the indicator of the event B, that is $\mathbf{I}B = 1$ if B is True, and $\mathbf{I}B = 0$ if B is False.

Theorem 3.3. (Parametric Fatou's lemma for weakly convergent probabilities) *Let $\mathbb{X}$ and Ξ be metric spaces. For probability measures $\{P, P^\nu, \nu \in \mathbb{N}\}$ on $(\Xi, \mathcal{B}(\Xi))$ and functions $\{f, f^\nu : \Xi \times \mathbb{X} \to \overline{\mathbb{R}}\}$, suppose that P^ν converges weakly to P and the following hold for each $x \in \mathbb{X}$:*

(i) $\{f(\,\cdot\,, x), f^\nu(\,\cdot\,, x), \nu \in \mathbb{N}\}$ *are measurable;*

(ii) *one has*
$$\liminf_{K\to+\infty} \liminf_{(\nu,y)\to(+\infty,x)} \mathbb{E}^{P^\nu}\big[f^\nu(\xi, y)\mathbf{I}\{\xi \,:\, f^\nu(\xi, y) \le -K\}\big] = 0; \tag{7}$$

(iii) *for P-a.e. $\xi \in \Xi$:*
$$\liminf_{(\nu,y,\zeta)\to(+\infty,x,\xi)} f^\nu(\zeta, y) \ge f(\xi, x). \tag{8}$$

Then for each $x \in \mathbb{X}$:
$$\liminf_{(\nu,y)\to(+\infty,x)} E^{P^\nu}\big[f^\nu\big](y) \ge E^{P}\big[f\big](x). \tag{9}$$

We observe that the first lower limit in (7) can be replaced by a limit because the subsequent expression is nondecreasing in K. The same observation applies to formula (10) below.

Before proving Theorem 3.3, we formulate and prove Theorem 3.4 which is a particular case of Theorem 3.3 when $\mathbb{X}$ is a singleton. Theorem 3.4 is an extension

of Fatou's lemma for weakly converging probabilities and unbounded functions [12, Theorem 2.4] to the convention $\mathbb{E}^P[g(\xi)] = +\infty$ if $\mathbb{E}^P[g_+(\xi)] = \mathbb{E}^P[g_-(\xi)] = +\infty$. In [12, Theorem 2.4], the generalized Fatou's inequality (11) is proved under the assumption that each expectation in the inequality has its positive part and/or its negative part being finite. Another formal difference between Theorem 3.4 and [12, Theorem 2.4] is that the former deals with probability measures while the latter deals with finite measures. However, the case of finite measures follows from the case of probability measures by normalization, if the limiting finite measure of the entire space Ξ is positive. If the limiting finite measure is 0, then the right-hand side of inequality (11) is 0, and assumption (10) implies that the lower limit on the left-hand side of (11) is nonnegative. Thus, the extension of Theorem 3.4 from probabilities to finite measures is routine.

Theorem 3.4. (Extended Fatou's lemma for weakly convergent probabilities; cp. [12, Theorem 2.4]) *Let Ξ be a metric space, $\{P^\nu\}_{\nu\in\mathbb{N}}$ be a sequence of probabilities on Ξ converging weakly to a probability P on Ξ, and $\{h^\nu\}_{\nu\in\mathbb{N}}$ be a sequence of measurable $\overline{\mathbb{R}}$-valued functions on Ξ such that*

$$\liminf_{K\to+\infty} \liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\mathbf{I}\{\xi \,:\, h^\nu(\xi) \le -K\}\big] = 0. \tag{10}$$

Then

$$\liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\big] \ge \mathbb{E}^{P}\big[\liminf_{(\nu,\zeta)\to(+\infty,\xi)} h^\nu(\zeta)\big]. \tag{11}$$

The proof of Theorem 3.4 uses the following lemma stating additional properties of expectations of functions satisfying assumption (10), which amounts to asymptotic uniform integrability from below.

Lemma 3.5. *Under the assumptions of Theorem* 3.4*, either*

$$\liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\big] = +\infty \quad \text{or} \quad \mathbb{E}^{P}\big[\big(\liminf_{(\nu,\zeta)\to(+\infty,\xi)} h^\nu(\zeta)\big)_+\big] < +\infty.$$

Proof. In view of (10), $\liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\mathbf{I}\{\xi \,:\, h^\nu(\xi) \le -K_0\}\big] > -\infty$ for some real $K_0 > 0$. Therefore,

$$\begin{aligned}
&\limsup_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu_-(\xi)\big] \\
&\le \limsup_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu_-(\xi)\mathbf{I}\{\xi \,:\, h^\nu_-(\xi) \ge K_0\}\big] + \limsup_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu_-(\xi)\mathbf{I}\{\xi \,:\, h^\nu_-(\xi) < K_0\}\big] \\
&\le -\liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\mathbf{I}\{\xi \,:\, h^\nu(\xi) \le -K_0\}\big] + K_0 < +\infty.
\end{aligned}$$

Thus, $\limsup_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu_-(\xi)\big] < +\infty$, and this inequality implies

$$\liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu_+(\xi)\big] \le \limsup_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu_-(\xi)\big] + \liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\big]. \tag{12}$$

Therefore, due to (12), if we have $\liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\big] < +\infty$, then we obtain $\liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu_+(\xi)\big] < +\infty$. Thus, according to Fatou's lemma for weakly convergent probabilities [12, Theorem 2.4], we get

$$\mathbb{E}^{P}\big[\liminf_{(\nu,\zeta)\to(+\infty,\xi)} h^\nu_+(\zeta)\big] \le \liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu_+(\xi)\big] < +\infty.$$

Hence, $$\mathbb{E}^P\big[\big(\liminf_{(\nu,\zeta)\to(+\infty,\xi)} h^\nu(\zeta)\big)_+\big] = \mathbb{E}^P\big[\liminf_{(\nu,\zeta)\to(+\infty,\xi)} h^\nu_+(\zeta)\big] < +\infty$$

because $\big(\liminf_{k\to\infty} a^k\big)_+ = \liminf_{k\to\infty} a^k_+$ for each sequence $\{a^k\}_{k\in\mathbb{N}} \subset \overline{\mathbb{R}}$ since we have $\big(\inf_{k\ge n} a^k\big)_+ = \inf_{k\ge n} a^k_+$ for all $n \in \mathbb{N}$ and, therefore,

$$\big(\sup_{n\in\mathbb{N}} \inf_{k\ge n} a^k\big)_+ = \sup_{n\in\mathbb{N}} \inf_{k\ge n} a^k_+.$$ □

Proof of Theorem 3.4. If $\liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\big] = +\infty$ then (11) holds due to the convention $+\infty - \alpha = +\infty$ for each $\alpha \in \overline{\mathbb{R}}$.
Otherwise, if $\liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\big] < +\infty$, then Lemma 3.5 implies that

$$\mathbb{E}^P\big[\big(\liminf_{(\nu,\zeta)\to(+\infty,\xi)} h^\nu(\zeta)\big)_+\big] < +\infty. \tag{13}$$

Fatou's lemma for weakly converging probabilities [12, Theorem 2.4], (10), and (13) imply (11). □

Let us consider the following corollary to Theorem 3.4.

Corollary 3.6. *Let Ξ be a metric space, $\{P^\nu\}_{\nu\in\mathbb{N}}$ be a sequence of probabilities on Ξ converging weakly to a probability P on Ξ, and $\{h^\nu\}_{\nu\in\mathbb{N}}$ be a sequence of measurable $\overline{\mathbb{R}}$-valued functions on Ξ such that*

$$\limsup_{K\to+\infty} \limsup_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\mathbf{I}\{\xi : h^\nu(\xi) \ge K\}\big] = 0. \tag{14}$$

Then
$$\limsup_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\big] \le \mathbb{E}^P\big[\limsup_{(\nu,\zeta)\to(+\infty,\xi)} h^\nu(\zeta)\big]. \tag{15}$$

Proof. The proof repeats the proof of Theorem 3.4 with minor technical modifications. If

$$\limsup_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\big] = -\infty,$$

then (15) holds trivially. Otherwise, if

$$\limsup_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[h^\nu(\xi)\big] > -\infty,$$

then Lemma 3.5 applied to the sequence $\{-h^\nu\}_{\nu\in\mathbb{N}}$ implies that

$$\mathbb{E}^P\big[\big(\limsup_{(\nu,\zeta)\to(+\infty,\xi)} h^\nu(\zeta)\big)_-\big] > -\infty. \tag{16}$$

Fatou's lemma for weakly converging probabilities [12, Theorem 2.4], (14), and (16) imply (15). □

Proof of Theorem 3.3. Fix an arbitrary $x \in \mathbb{X}$. Consider a sequence $\{x^\nu\}_{\nu\in\mathbb{N}}$ converging to x such that

$$\liminf_{(\nu,y)\to(+\infty,x)} E^{P^\nu}\big[f^\nu\big](y) = \liminf_{(\nu,x^\nu)\to(+\infty,x)} E^{P^\nu}\big[f^\nu\big](x^\nu) = \liminf_{\nu\to+\infty} E^{P^\nu}\big[f^\nu\big](x^\nu). \tag{17}$$

Let us set $h^\nu(\,\cdot\,) := f^\nu(\,\cdot\,, x^\nu)$, $\nu \in \mathbb{N}$. We observe that (7) implies (10). Therefore, since

$$\liminf_{(\nu,\zeta)\to(+\infty,\xi)} h^\nu(\zeta) \ge \liminf_{(\nu,y,\zeta)\to(+\infty,x,\xi)} f^\nu(\zeta, y),$$

(17), Theorem 3.4, and (8) imply (9). □

Assumption (iii) in Theorem 3.3 relates to lsc as follows. For each $x \in \mathbb{X}$, suppose that the functions $\{f^\nu(\,\cdot\,, x), \nu \in \mathbb{N}\}$ are *equi-lsc*, i.e., for P-a.e. $\xi \in \Xi$, and for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$f^\nu(\zeta, x) > f^\nu(\xi, x) - \varepsilon \qquad \text{for all } \zeta \in \mathbb{B}_\Xi(\xi;\delta) \text{ and for sufficiently large } \nu \in \mathbb{N},$$

where $\mathbb{B}_\Xi(\xi;\delta) := \{\zeta \in \Xi : d_\Xi(\zeta,\xi) < \delta\}$. Then (8) can be replaced with *lower semiconvergence* of $\{f^\nu(\,\cdot\,, x), \nu \in \mathbb{N}\}$ to $f(\,\cdot\,, x)$ in probability P, that is, for each $\varepsilon > 0$, one has

$$P\big(\{\xi \in \Xi : f^\nu(\xi, x) \le f(\xi, x) - \varepsilon\}\big) \to 0 \quad \text{as } \nu \to +\infty. \tag{18}$$

This follows from [11, Theorem 4.1]. A sufficient condition for (18) is

$$\liminf_{\nu\to+\infty} f^\nu(\xi, x) \ge f(\xi, x) \quad \text{for } P\text{-a.e. } \xi \in \Xi. \tag{19}$$

Moreover, (18) implies that

$$\liminf_{k\to+\infty} f^{\nu_\kappa}(\xi, x) \ge f(\xi, x) \quad \text{for } P\text{-a.e. } \xi \in \Xi, \tag{20}$$

for a subsequence $\{f^{\nu_\kappa}\}_{k\in\mathbb{N}} \subset \{f^\nu\}_{\nu\in\mathbb{N}}$.

According to [12, Theorem 2.6], a sufficient condition for (7) is that for each $\bar{x} \in \mathbb{X}$ there exists $\rho > 0$ and a sequence of measurable $\overline{\mathbb{R}}$-valued functions $\{g^\nu\}_{\nu\in\mathbb{N}}$ on Ξ such that

$$\min\Big\{0, \inf_{x\in\mathbb{B}_\mathbb{X}(\bar{x};\rho)} f^\nu(\xi, x)\Big\} \ge g^\nu(\xi)$$

for sufficiently large $\nu \in \mathbb{N}$ and $\xi \in \Xi$, where $\mathbb{B}_\mathbb{X}(\bar{x};\rho) := \{x \in \mathbb{X} : d_\mathbb{X}(\bar{x}, x) \le \rho\}$, and

$$-\infty < \mathbb{E}^P\big[\limsup_{(\nu,\zeta)\to(+\infty,\xi)} g^\nu(\zeta)\big] \le \liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[g^\nu(\xi)\big]. \tag{21}$$

As demonstrated in [12, Example 3.3], condition (7) cannot be replaced in Theorem 3.3 with the following inequalities, which are weaker than (21):

$$-\infty < \mathbb{E}^P\big[\liminf_{(\nu,\zeta)\to(+\infty,\xi)} g^\nu(\zeta)\big] \le \liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[g^\nu(\xi)\big]. \tag{22}$$

If $\{P^\nu\}_{\nu\in\mathbb{N}}$ converges to P setwise, then the following inequalities

$$-\infty < \mathbb{E}^P\big[\limsup_{\nu\to+\infty} g^\nu(\xi)\big] \le \liminf_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[g^\nu(\xi)\big], \tag{23}$$

which are similar to (21), imply the validity of Fatou's inequality [13, Theorem 4.1].

If $P^\nu = P$, then $P^\nu \to P$ setwise trivially. If $g^\nu = g$, then (23) becomes obviously $\mathbb{E}^P\big[g(\xi)\big] > -\infty$, which is the existence of an integrable minorant. This is a stronger

condition than the uniform integrability condition (7) even in the classic case, when $\mathbb{X}$ is a singleton. In this special, case when $P^\nu = P$ and $f^\nu = f$, [27, Proposition 8.55] states that if f is random lsc and locally inf-integrable, then f is lsc. In terms of the present paper, this is the special case when $g = f$. If $P^\nu \to P$, in total variation, then additional results on the convergence of expectations can be obtained from the uniform Fatou's lemma [14, Theorem 2.1].

Theorem 3.7. (Epi-convergence under weakly convergent probabilities) *Let $\mathbb{X}$ and Ξ be metric spaces. For probability measures $\{P, P^\nu, \nu \in \mathbb{N}\}$ on $(\Xi, \mathcal{B}(\Xi))$ and functions $\{f, f^\nu : \Xi \times \mathbb{X} \to \overline{\mathbb{R}}\}$, suppose that P^ν converges weakly to P, assumptions* (i)–(iii) *of Theorem* 3.3 *hold, and for each $x \in \mathbb{X}$ either $E^P[f](x) = +\infty$ or there exists a sequence $\{x^\nu\}_{\nu\in\mathbb{N}} \subset \mathbb{X}$ converging to x with the following properties:*

(i) *one has*
$$\limsup_{K\to+\infty} \limsup_{\nu\to+\infty} \mathbb{E}^{P^\nu}\big[f^\nu(\xi, x^\nu)\mathbf{I}\{\xi \,:\, f^\nu(\xi, x^\nu) \ge K\}\big] = 0; \tag{24}$$

(ii) *for P-a.e. $\xi \in \Xi$:*
$$\limsup_{(\nu,\zeta)\to(+\infty,\xi)} f^\nu(\zeta, x^\nu) \le f(\xi, x). \tag{25}$$

Then
$$\operatorname*{e-lim}_{\nu\to+\infty} E^{P^\nu}\big[f^\nu\big] = E^P\big[f\big]. \tag{26}$$

Proof. Choose an arbitrary $x \in \mathbb{X}$. If $E^P\big[f\big](x) = +\infty$, then Theorem 3.3 implies $\operatorname*{e-lim}_{\nu\to+\infty} E^{P^\nu}\big[f^\nu\big](x) = E^P\big[f\big](x)$.

Let $E^P\big[f\big](x) < +\infty$. Fix a sequence $\{x^\nu\}_{\nu\in\mathbb{N}} \subset \mathbb{X}$ converging to x and satisfying (24) and (25). Inequality (9) follows from Theorem 3.3. Therefore, to prove epi-convergence it is sufficient to establish the inequality

$$\limsup_{\nu\to+\infty} E^{P^\nu}\big[f^\nu\big](x^\nu) \le E^P\big[f\big](x), \tag{27}$$

which follows from Corollary 3.6 with $h^\nu(\,\cdot\,) := f^\nu(\,\cdot\,, x^\nu)$, $\nu \in \mathbb{N}$. We observe that (24) implies (14), and (15) and (25) imply (27). □

Theorem 3.7 strengthens the results of [9, 8] by considering approximating functions f^ν rather than a fixed integrand f. Specifically, [8, Theorem 1] allows $\mathbb{X}$ to be a Hausdorff topological space and confirms epi-convergence of $E^{P^\nu}\big[f\big]$ to $E^P\big[f\big]$ under the assumptions that f is lsc and $f(\cdot, x)$ is usc for each x. Moreover, P^ν converges weakly to P and one has the following tightness condition: for each $x \in \mathbb{X}$, for each neighborhood V of x, and for each $\epsilon > 0$, there exists a compact $K_\epsilon \subset \mathbb{X}$ such that

$$\int_{\Xi\setminus K_\epsilon} \big|f(\xi, x)\big| P^\nu(d\xi) + \int_{\Xi\setminus K_\epsilon} \big|f_V(\xi)\big| P^\nu(d\xi) < \epsilon \quad \forall \nu, \tag{28}$$

where
$$f_V(\xi) = \liminf_{\zeta\to\xi} [\inf_{y\in V} f(\zeta, y)], \quad \xi \in \Xi.$$

These assumptions of [8, Theorem 1], when $\mathbb{X}$ is a metrizable space, imply the conditions of Theorem 3.7. Indeed, suppose that f is lsc and $f(\cdot, x)$ is usc for each x, and suppose that for a fixed $x \in \mathbb{X}$, for each neighborhood V of x, and for each $\epsilon > 0$, there exists compact K_ϵ such that (28) hold. Then, we can verify (7) as follows. According to the proof of Theorem 1 from [8], the function f_V is continuous.

Its absolute value is bounded by some $\gamma > 0$ on the compact set K_ϵ. Therefore,

$$\begin{aligned}\liminf_{y\to x} \mathbb{E}^{P^\nu}\big[f(\xi,y)\mathbf{I}\{\xi \,:\, f(\xi,y)\le -\gamma\}\big] \\ \ge \inf_{y\in V} \mathbb{E}^{P^\nu}\big[f(\xi,y)\mathbf{I}\{\xi \,:\, f(\xi,y)\le -\gamma\}\big] \\ \ge \mathbb{E}^{P^\nu}\big[f_V(\xi)\mathbf{I}\{\xi \,:\, f_V(\xi)\le -\gamma\}\big] \ge -\varepsilon,\end{aligned} \tag{29}$$

where the first inequality follows from the definition of lower limits, the second inequality holds because $f_V(\xi) \le f(\xi,y) < 0$ if $f(\xi,y) \le -\gamma$, and the last inequality follows from (28) because if $f_V(\xi) < -\gamma$, then $\xi \in \Xi \setminus K_\epsilon$. Thus, (7) holds. Similarly to these arguments, we can verify (24) with $x^\nu := x$, $\nu \in \mathbb{N}$. Inequality (8) for $f^\nu = f$, $\nu \in \mathbb{N}$, follows from lsc of f. Inequality (25) for $f^\nu = f$, $\nu \in \mathbb{N}$, and $x^\nu := x$, $\nu \in \mathbb{N}$, follows from usc of $f(\,\cdot\,, x)$.

If the family of functions $\{f^\nu(\,\cdot\,, x^\nu), \nu \in \mathbb{N}\}$ is equi-usc, i.e., if the family of functions $\{-f^\nu(\,\cdot\,, x^\nu),\ \nu \in \mathbb{N}\}$ is equi-lsc, then (25) can be replaced with *upper semiconvergence* of $\{f^\nu(\,\cdot\,, x^\nu), \nu \in \mathbb{N}\}$ to $f(\,\cdot\,, x)$ in measure P, i.e., for each $\varepsilon > 0$

$$P\big(\{\xi \in \Xi : f^\nu(\xi, x^\nu) \ge f(\xi, x) + \varepsilon\}\big) \to 0 \text{ as } \nu \to +\infty. \tag{30}$$

A sufficient condition for (30) is

$$\limsup_{\nu\to+\infty} f^\nu(\xi, x^\nu) \le f(\xi, x) \quad \text{for } P\text{-a.e. } \xi \in \Xi. \tag{31}$$

4. Applications

As illustrations of the possibilities emerging from the above results, we present two applications of Theorem 3.3 (parametric Fatou's lemma) and two examples of Theorem 3.7 (epi-convergence). In the following, for any set C, we let $\iota_C(x) = 0$ if $x \in C$ and $\iota_C(x) = +\infty$ otherwise.

4.1. Sieves in Nonparametric Statistics

Consistency analysis of parametric and nonparametric estimation problems in statistics is supported by the prior results. Here, we consider M-estimators with sieves [15, 26] that are distinguished by being selected from a collection of sets indexed by sample size and other varying quantities. Specifically, let $(\mathbb{X}, d_{\mathbb{X}})$ be a metric space of extended real-valued functions defined on a closed subset S of $\mathbb{R}^d$, let $(\Xi, \mathcal{B}(\Xi))$ be a measurable space, and let $(\mathcal{P}, d_{\mathcal{P}})$ be a metric space of probability measures defined on $(\Xi, \mathcal{B}(\Xi))$. For given collections $\{f, f^\nu : \Xi \times \mathbb{X} \to \overline{\mathbb{R}}, \nu \in \mathbb{N}\}$, $\{X, X^\nu \subset \mathbb{X}, \nu \in \mathbb{N}\}$, and $\{P, P^\nu \in \mathcal{P}, \nu \in \mathbb{N}\}$, a minimizer of the optimization problem

$$\operatorname*{minimize}_{x\in X^\nu} \mathbb{E}^{P^\nu}\big[f^\nu(\xi, x)\big] \tag{32}$$

is an estimator of "true" functions in $\mathbb{X}$ that, in turn, solve

$$\operatorname*{minimize}_{x\in X} \mathbb{E}^{P}\big[f(\xi, x)\big]. \tag{33}$$

Often, P^ν is an empirical distribution produced by ν samples as an approximation of the actual distribution P, but many other possibilities exist; see for example [16, 17].

The discrepancy between X^ν and X stems from the need to consider restrictions and approximations of the function class $\mathbb{X}$, especially when ν is small and/or d is large. The "loss function" f is approximated by f^ν to accommodate computations. For example, a nonsmooth hinge-loss might be replaced by an approximating soft-max loss. The consistency of estimators obtained from (32) can be established through epi-convergence; see [9, 16, 17]. However, the focus has been on the special case $f^\nu = f$ and $X^\nu = X$ for all ν; an exception is [26] where an adjustment of (32) is proposed for $X^\nu \neq X$ under the assumption that $X = \mathbb{X}$ and $f^\nu = f$. The source of the difficulty caused by $X^\nu \neq X$ is gleaned from the assumptions in Theorem 3.7, which essentially insist on $f^\nu(\xi^\nu, x^\nu) \to f(\xi, x)$ when $\xi^\nu \to \xi$ and $x^\nu \to x$. We discuss a different approach leveraging Theorem 3.3.

We equip $\mathbb{X} \times \mathcal{P}$ with the product topology. The actual problem (33) is equivalently expressed as

$$\underset{(x,Q)\in\mathbb{X}\times\mathcal{P}}{\text{minimize}}\ \phi(x,Q) := \mathbb{E}^Q\big[f(\xi,x) + \iota_X(x)\big] + \iota_{\{P\}}(Q). \tag{34}$$

This reformulation motivates an alternative to (32) involving a nonnegative parameter $\theta^\nu \in [0, +\infty)$:

$$\underset{(x,Q)\in\mathbb{X}\times\mathcal{P}}{\text{minimize}}\ \phi^\nu(x,Q) := \mathbb{E}^Q\big[f^\nu(\xi,x) + \iota_{X^\nu}(x)\big] + \theta^\nu d_{\mathcal{P}}(Q,P^\nu); \tag{35}$$

see [5] for related reformulations in finite dimensions. We show via Theorem 3.3 that estimators obtained as the x-component of minimizers of (35) are consistent, in the sense that ϕ^ν epi-converges to ϕ (as functions on $\mathbb{X} \times \mathcal{P}$), under broad conditions. In particular, the analysis is unencumbered by $X^\nu \neq X$. These sets may even depend on ξ, but the details are omitted below. Regardless, we observe that it is advantageous to keep $\iota_{X^\nu}(x)$ inside the expectation defining ϕ^ν because it helps with assumption (ii) in Proposition 4.1.

We recall that the *outer limit* of a sequence of sets $\{A^\nu, \nu \in \mathbb{N}\}$ in a topological space, denoted by LimOut A^ν, is the collection of points to which a subsequence of $\{a^\nu \in A^\nu, \nu \in \mathbb{N}\}$ converges. The *inner limit*, denoted by LimInn A^ν, is the collection of points to which a sequence $\{a^\nu \in A^\nu, \nu \in \mathbb{N}\}$ converges. If both the inner limit and the outer limit are equal to A, we say that $\{A^\nu, \nu \in \mathbb{N}\}$ *set-converges* to A.

Proposition 4.1. (consistency of sieved M-estimator) *In the notation of this subsection, suppose that $\{X^\nu, \nu \in \mathbb{N}\}$ are closed and set-converge to X, $d_{\mathcal{P}}$ metrizes weak convergence, $\theta^\nu \to +\infty$, $\theta^\nu d_{\mathcal{P}}(P^\nu, P) \to 0$, and the following hold for each $x \in \mathbb{X}$:*

(i) *$f(\cdot, x)$ and $f^\nu(\cdot, x)$ are measurable for all $\nu \in \mathbb{N}$ and $f(\xi, x) > -\infty$ for all $\xi \in \Xi$;*

(ii) *for $Q^\nu \to P$*
$$\liminf_{K\to+\infty}\ \liminf_{(\nu,y)\to(+\infty,x)}\ \mathbb{E}^{Q^\nu}\big[\big(f^\nu(\xi,y) + \iota_{X^\nu}(y)\big)\mathbf{I}\big\{\xi : f^\nu(\xi,y) + \iota_{X^\nu}(y) \le -K\big\}\big] = 0; \tag{36}$$

(iii) *for P-a.e. $\xi \in \Xi$*
$$\liminf_{(\nu,y,\zeta)\to(+\infty,x,\xi)} f^\nu(\zeta,y) \ge f(\xi,x); \tag{37}$$

(iv) *for* $\{x^\nu \in X^\nu, \nu \in \mathbb{N}\}$ *converging to* x *and* $\{Q^\nu \in \mathcal{P}, \nu \in \mathbb{N}\}$ *converging to* $Q \in \mathcal{P}$,
$$\liminf_{\nu\to+\infty} \mathbb{E}^{Q^\nu}\big[f^\nu(\xi, x^\nu)\big] > -\infty;$$

(v) *for* $\{x^\nu \in X^\nu, \nu \in \mathbb{N}\}$ *converging to* x, *there exists a* P*-integrable function* $g\colon \Xi \to [0, +\infty)$ *such that, for* P*-a.e.* $\xi \in \Xi$, *one has* $f^\nu(\xi, x^\nu) \le g(\xi)$ *for all* $\nu \in \mathbb{N}$ *and* $f^\nu(\xi, x^\nu) \to f(\xi, x)$ *as* $\nu \to +\infty$.

Then, ϕ^ν *epi-converges to* ϕ.

Proof. First, we consider the liminf-condition required for epi-convergence. Let $\{(x^\nu, Q^\nu) \in \mathbb{X} \times \mathcal{P}, \nu \in \mathbb{N}\}$ converge to $(\bar{x}, Q) \in \mathbb{X} \times \mathcal{P}$. We consider three cases.

(a) Suppose that $\bar{x} \in X$ and $Q = P$. Without loss of generality, we assume that $x^\nu \in X^\nu$ because otherwise $\phi^\nu(x^\nu, Q^\nu) = +\infty$. Let $\hat{f}, \hat{f}^\nu : \Xi \times \mathbb{X} \to \overline{\mathbb{R}}$ be defined by $\hat{f}(\xi, x) := f(\xi, x) + \iota_X(x)$ and $\hat{f}^\nu(\xi, x) := f^\nu(\xi, x) + \iota_{X^\nu}(x)$. Then, for each $x \in \mathbb{X}$, $\hat{f}(\cdot, x)$ and $\hat{f}^\nu(\cdot, x)$ are measurable; recall that X^ν is closed and X as well by virtue of being a set-limit. Let $x \in \mathbb{X}$. For P-a.e. $\xi \in \Xi$, one has

$$\begin{aligned}\liminf_{(\nu,y,\zeta)\to(+\infty,x,\xi)} \hat{f}^\nu(\zeta, y) &\ge \liminf_{(\nu,y,\zeta)\to(+\infty,x,\xi)} f^\nu(\zeta, y) + \liminf_{(\nu,y)\to(+\infty,x)} \iota_{X^\nu}(y)\\ &\ge f(\xi, x) + \iota_X(x) = \hat{f}(\xi, x).\end{aligned}$$

The last inequality follows by (iii) and the fact that X^ν set-converges to X. Therefore, we can invoke Theorem 3.3 for the integrands $\{\hat{f}, \hat{f}^\nu, \nu \in \mathbb{N}\}$ and probabilities $\{P, Q^\nu, \nu \in \mathbb{N}\}$ and conclude that

$$\begin{aligned}\liminf_{\nu\to+\infty} \phi^\nu(x^\nu, Q^\nu) &\ge \liminf_{\nu\to+\infty} \mathbb{E}^{Q^\nu}\big[f^\nu(\xi, x^\nu) + \iota_{X^\nu}(x^\nu)\big] = \liminf_{\nu\to+\infty} E^{Q^\nu}[\hat{f}^\nu](x^\nu)\\ &\ge E^P[\hat{f}](\bar{x}) = \phi(\bar{x}, P).\end{aligned}$$

(b) If $\bar{x} \notin X$, then $\phi(\bar{x}, Q) = +\infty$ and we seek to establish that $\phi^\nu(x^\nu, Q^\nu) \to +\infty$. Since X^ν set-converges to X, $x^\nu \notin X^\nu$ for sufficiently large ν. This implies that $\phi^\nu(x^\nu, Q^\nu) = +\infty$ for sufficiently large ν because $\iota_{X^\nu}(x^\nu) = +\infty$.

(c) If $\bar{x} \in X$ and $Q \ne P$, then $\phi(\bar{x}, Q) = +\infty$. Since $\{d_{\mathcal{P}}(Q^\nu, P^\nu), \nu \in \mathbb{N}\}$ is bounded away from zero as $\nu \to +\infty$ because $d_{\mathcal{P}}(Q^\nu, P^\nu) \ge d_{\mathcal{P}}(Q, P) - d_{\mathcal{P}}(Q, Q^\nu) - d_{\mathcal{P}}(P, P^\nu)$, one has

$$\liminf_{\nu\to+\infty} \phi^\nu(x^\nu, Q^\nu) \ge \liminf_{\nu\to+\infty} \mathbb{E}^{Q^\nu}\big[f^\nu(\xi, x^\nu) + \iota_{X^\nu}(x^\nu)\big] + \liminf_{\nu\to+\infty} \theta^\nu d_{\mathcal{P}}(Q^\nu, P^\nu) = +\infty.$$

Here, the inequality follows from (iv).

Second, we consider the limit-condition in the definition of epi-convergence. Let $(\bar{x}, Q) \in \mathbb{X} \times \mathcal{P}$. In light of the above liminf-condition, it suffices to show that $\limsup_{\nu\to+\infty} \phi^\nu(x^\nu, Q^\nu) \le \phi(x, Q)$ for some $x^\nu \to x$ and $Q^\nu \to Q$. Without loss of generality, we assume that $\bar{x} \in X$ and $Q = P$ because otherwise the limsup-condition holds trivially. Since X^ν set-converges to X, there exists $\{x^\nu \in X^\nu, \nu \in \mathbb{N}\}$ converging to $\bar{x}$. We construct $\{Q^\nu = P, \nu \in \mathbb{N}\}$. Then,

$$\begin{aligned}\limsup_{\nu\to+\infty} \phi^\nu(x^\nu, Q^\nu) &\le \liminf_{\nu\to+\infty} \mathbb{E}^P\big[f^\nu(\xi, x^\nu)\big] + \limsup_{\nu\to+\infty} \theta^\nu d_{\mathcal{P}}(P, P^\nu)\\ &\le \mathbb{E}^P\big[f(\xi, \bar{x})\big] = \phi(\bar{x}, P)\end{aligned}$$

because $\theta^\nu d_{\mathcal{P}}(P, P^\nu) \to 0$. The last inequality follows from (v), which in conjunction with (ii), permits us to invoke the dominated convergence theorem. □

As a concrete example, we let $\mathbb{X}$ be the space of extended real-valued functions on $S \subset \mathbb{R}^d$, excluding the function identical to $-\infty$. This is a Polish space under the hypo-distance; see [26]. Let $\Xi = S$ and $f^\nu(\xi, x) = f(\xi, x) = -\log x(\xi)$, which corresponds to a maximum likelihood estimator of a probability density on S. In this case, the requirement in (iii) holds because x^ν converging to x in the hypo-distance implies that $\limsup_{\nu\to+\infty} x^\nu(\xi^\nu) \le x(\xi)$ for any $\xi^\nu \in S \to \xi$. The requirement in (v) about $f(\xi, x^\nu) \to f(\xi, x)$ when x^ν converges to x in the hypo-distance now translates into having $x^\nu(\xi) \to x(\xi)$, which holds if $\mathbb{X}$ is equi-usc; see [26] for details and examples.

4.2. Mollifiers for discontinuous functions

For a normed linear space $\mathbb{X}$, it can be challenging to minimize a function $g : \mathbb{X} \to \overline{\mathbb{R}}$ if it is not continuous. A standard approach is to consider mollifiers to "smooth" the function sufficiently, i.e., to replace g by the approximation $x \mapsto \mathbb{E}^{P^\nu}[g(x+\xi)]$, where the probabilities P^ν converge weakly to P, with $P(\{0\}) = 1$. Typically, these probabilities are assumed to be absolutely continuous relative to a Lebesgue measure; see [10]. One would establish epi-convergence of the approximations to g as $\nu \to +\infty$ to justify the approach. We show that Theorem 3.3 can be brought in to justify approaches of this kind under general conditions including approximating $g^\nu : \mathbb{X} \to \overline{\mathbb{R}}$ and discontinuous g. In contrast, [10] assumes that for each $x \in \mathbb{X}$ there exist points $x^\nu \to x$ such that $g(x^\nu) \to g(x)$ and then also in the special case of $\mathbb{X} = \mathbb{R}^n$.

Suppose that $\{g^\nu : \mathbb{X} \to \overline{\mathbb{R}}, \nu \in \mathbb{N}\}$ and $(\mathcal{P}, d_{\mathcal{P}})$ is a metric space of probabilities defined on $(\mathbb{X}, \mathcal{B}_{\mathbb{X}})$, where $\mathcal{B}_{\mathbb{X}}$ is the Borel sigma-algebra on $\mathbb{X}$. For a collection $\{P^\nu \in \mathcal{P}, \nu \in \mathbb{N}\}$ and $\theta^\nu \in [0, +\infty)$, we consider the problem

$$\underset{(x,Q)\in\mathbb{X}\times\mathcal{P}}{\text{minimize}}\ \phi^\nu(x, Q) := \mathbb{E}^Q\big[g^\nu(x+\xi)\big] + \theta^\nu d_{\mathcal{P}}(Q, P^\nu) \tag{38}$$

as a substitute for the actual problem of minimizing g over $\mathbb{X}$ or, equivalently, solving

$$\underset{(x,Q)\in\mathbb{X}\times\mathcal{P}}{\text{minimize}}\ \phi(x, Q) := \mathbb{E}^Q\big[g(x+\xi)\big] + \iota_{\{P\}}(Q). \tag{39}$$

We show via Theorem 3.3 that ϕ^ν epi-converges to ϕ (as functions on $\mathbb{X}\times\mathcal{P}$), which justifies the solution of (38) in lieu of (39).

Proposition 4.2. (Mollifiers) *In the notation of this subsection, suppose that $d_{\mathcal{P}}$ metrizes weak convergence, $\theta^\nu \to +\infty$, $\theta^\nu d_{\mathcal{P}}(P^\nu, P) \to 0$, and the following hold for each $x \in \mathbb{X}$:*

(i) *$\{g, g^\nu, \nu \in \mathbb{N}\}$ are measurable;*

(ii) *for $Q^\nu \to Q \in \mathcal{P}$*

$$\liminf_{K\to+\infty} \liminf_{(\nu,y)\to(+\infty,x)} \mathbb{E}^{Q^\nu}\big[\big(g^\nu(y+\xi)\big)\mathbf{I}\{\xi : g^\nu(y+\xi) \le -K\}\big] = 0; \tag{40}$$

(iii) *one has*

$$\liminf_{(\nu,y)\to(+\infty,x)} g^\nu(y) \ge g(x); \tag{41}$$

(iv) *for $Q \in \mathcal{P}$, $\mathbb{E}^Q[g(x+\xi)] > -\infty$;*

(v) *there exists a P-integrable function $h:\Xi \to [0,+\infty)$ such that, for P-a.e. $\xi\in\Xi$, one has $g^\nu(x+\xi) \le h(\xi)$ for all $\nu \in \mathbb{N}$ and $g^\nu(x) \to g(x)$ as $\nu \to +\infty$.*

Then, ϕ^ν epi-converges to ϕ.

Proof. First, we consider the liminf-condition required for epi-convergence. Let $\{(x^\nu, Q^\nu) \in \mathbb{X} \times \mathcal{P}, \nu \in \mathbb{N}\}$ converge to $(\bar{x}, Q) \in \mathbb{X} \times \mathcal{P}$. Let $f, f^\nu : \Xi \times \mathbb{X} \to \overline{\mathbb{R}}$ be defined by $f(\xi, x) := g(x+\xi)$ and $f^\nu(\xi, x) := g^\nu(x+\xi)$. Then, for each $x \in \mathbb{X}$, $f(\cdot, x)$ and $f^\nu(\cdot, x)$ are measurable. Let $x \in \mathbb{X}$. For all $\xi \in \Xi$, one has

$$\liminf_{(\nu,y,\zeta)\to(+\infty,x,\xi)} f^\nu(\zeta, y) = \liminf_{(\nu,y,\zeta)\to(+\infty,x,\xi)} g^\nu(y+\zeta) \ge g(x+\xi) = f(\xi, x)$$

by assumption (iii). In consequence we can invoke Theorem 3.3 for the integrands $\{f, f^\nu, \nu \in \mathbb{N}\}$ and the probabilities $\{Q, Q^\nu, \nu \in \mathbb{N}\}$ and conclude that

$$\liminf_{\nu\to+\infty} E^{Q^\nu}[f^\nu](x^\nu) \ge E^Q[f](\bar{x}).$$

The right-hand side is not $-\infty$ by assumption (iv). We consider two cases. (a) Suppose that $Q = P$. Then,

$$\begin{aligned}\liminf_{\nu\to+\infty} \phi^\nu(x^\nu, Q^\nu) &\ge \liminf_{\nu\to+\infty} \mathbb{E}^{Q^\nu}\big[g^\nu(x^\nu+\xi)\big] = \liminf_{\nu\to+\infty} E^{Q^\nu}[f^\nu](x^\nu) \\ &\ge E^P[f](\bar{x}) = \phi(\bar{x}, P).\end{aligned}$$

(b) If $Q \ne P$, then $\phi(\bar{x}, Q) = +\infty$. Since $\{d_{\mathcal{P}}(Q^\nu, P^\nu), \nu \in \mathbb{N}\}$ is bounded away from zero as $\nu \to +\infty$, one has

$$\liminf_{\nu\to+\infty} \phi^\nu(x^\nu, Q^\nu) \ge \liminf_{\nu\to+\infty} \mathbb{E}^{Q^\nu}\big[g^\nu(x^\nu+\xi)\big] + \liminf_{\nu\to+\infty} \theta^\nu d_{\mathcal{P}}(Q^\nu, P^\nu) = +\infty.$$

Here, the first inequality follows from (iv).

Second, we consider the limit-condition in the definition of epi-convergence. Let $(\bar{x}, Q) \in \mathbb{X} \times \mathcal{P}$. It suffices to show that $\limsup_{\nu\to+\infty} \phi^\nu(x^\nu, Q^\nu) \le \phi(x, Q)$ for some $x^\nu \to x$ and $Q^\nu \to Q$. Without loss of generality, we assume that $Q = P$. We construct $\{x^\nu = \bar{x}, Q^\nu = P, \nu \in \mathbb{N}\}$. Then,

$$\begin{aligned}\limsup_{\nu\to+\infty} \phi^\nu(x^\nu, Q^\nu) &\le \liminf_{\nu\to+\infty} \mathbb{E}^P\big[g^\nu(\bar{x}+\xi)\big] + \limsup_{\nu\to+\infty} \theta^\nu d_{\mathcal{P}}(P, P^\nu) \\ &\le \mathbb{E}^P\big[g(\bar{x}+\xi)\big] = \phi(\bar{x}, P).\end{aligned}$$

The last inequality follows from (v), which in conjunction with (ii) permits us to invoke the dominated convergence theorem. □

4.3. PDE-constrained optimization

We next turn to the application of Theorem 3.7 in the context of PDE-constrained optimization and again assume that $\mathbb{X}$ and Ξ are metric spaces. Suppose that we are faced with a PDE parameterized by $\xi \in \Xi$, representing parameters subject to uncertainty, and $x \in \mathbb{X}$ representing control input. We assume that for each $\xi \in \Xi$ and $x \in \mathbb{X}$, there is a unique solution of the PDE denoted by $s(\xi, x)$ in a metric space $(\mathbb{U}, d_{\mathbb{U}})$, i.e., there is a well defined mapping

$$s : \Xi \times \mathbb{X} \to \mathbb{U}$$

furnishing solutions of the PDE for the various input parameters.

Suppose that $g : \mathbb{U} \times \mathbb{X} \to \overline{\mathbb{R}}$ represents a quantity of interest and P is a probability on $(\Xi, \mathcal{B}(\Xi))$ modeling the randomness in the parameter ξ, with Ξ being a metric space and $\mathcal{B}(\Xi)$ its Borel sigma-algebra. Thus, a PDE-constrained optimization problem tends to involve an expectation function

$$E^P[f] : \mathbb{X} \to \overline{\mathbb{R}}, \quad \text{where the integrand } f : \Xi \times \mathbb{X} \to \overline{\mathbb{R}} \text{ has } f(\xi, x) = g\big(s(\xi, x), x\big).$$

For computational and/or modeling reasons, the probability P needs to be replaced by probabilities P^ν on $(\Xi, \mathcal{B}(\Xi))$, g replaced by an approximation $g^\nu : \mathbb{U} \times \mathbb{X} \to \overline{\mathbb{R}}$, and the solution mapping s by $s^\nu : \Xi \times \mathbb{X} \to \mathbb{U}$. The latter usually represents numerical solutions of the PDE. If $\mathbb{X}$ is infinite-dimensional, then it might need to be approximated by a set $X^\nu \subset \mathbb{X}$. The approximation g^ν could then simply be defined by $g^\nu(u, x) = g(u, x) + \iota_{X^\nu}(x)$. Regardless of the specifics, the setting produces an approximating expectation function

$$E^{P^\nu}[f^\nu] : \mathbb{X} \to \overline{\mathbb{R}}, \quad \text{where the integrand } f^\nu : \Xi \times \mathbb{X} \to \overline{\mathbb{R}} \text{ has } f^\nu(\xi, x) = g^\nu\big(s^\nu(\xi, x), x\big).$$

A PDE-constrained optimization problem involving $E^P[f]$ can now be approximated using $E^{P^\nu}[f^\nu]$, with this approach being justified through epi-convergence. Theorem 3.7 provides sufficient conditions for such epi-convergence. Prior work do not consider epi-convergence in this general setting; see for example [6], which leverages continuous convergence of expectations and thus needs relatively strong assumptions.

We show via Theorem 3.7 that $E^{P^\nu}[f^\nu]$ epi-converges to $E^P[f]$ under mild assumptions.

Proposition 4.3. (PDE-constrained optimization) *In the notation of this subsection, suppose that P^ν converges weakly to P and the following hold for each $x \in \mathbb{X}$:*

(i) $\{s(\cdot, x), s^\nu(\cdot, x), g(\cdot, x), g^\nu(\cdot, x), \nu \in \mathbb{N}\}$ *are continuous;*

(ii) $$\liminf_{K \to +\infty} \liminf_{(\nu, y) \to (+\infty, x)} \mathbb{E}^{P^\nu}\big[\big(g^\nu(s^\nu(\xi, y), y)\big)\mathbf{I}\{\xi : g^\nu(s^\nu(\xi, y), y) \le -K\}\big] = 0; \quad (42)$$

(iii) *for P-a.e. $\xi \in \Xi$,*
$$\lim_{(\nu, y, \zeta) \to (+\infty, x, \xi)} s^\nu(\zeta, y) = s(\xi, x),$$
$$\lim_{(\nu, v) \to (+\infty, u)} g^\nu(v, x) = g(u, x), \qquad \liminf_{(\nu, y, v) \to (+\infty, x, u)} g^\nu(v, y) \ge g(u, x);$$

(iv) *if $\mathbb{E}^P[g(s(\xi, x), x)] < \infty$, then*
$$\limsup_{K \to +\infty} \limsup_{\nu \to +\infty} \mathbb{E}^{P^\nu}\big[\big(g^\nu(s^\nu(\xi, x), x)\big)\mathbf{I}\{\xi : g^\nu(s^\nu(\xi, x), x) \ge K\}\big] = 0. \quad (43)$$

Then, $E^{P^\nu}[f^\nu]$ epi-converges to $E^P[f]$.

Proof. We first verify the assumptions of Theorem 3.3 for the integrands given by $f(\xi, x) = g(s(\xi, x), x)$ and $f^\nu(\xi, x) = g^\nu(s^\nu(\xi, x), x)$. The continuity properties from assumption (i) confirm assumption (i) in Theorem 3.3. Assumptions (ii) and (iii) of Theorem 3.3 hold by our assumptions (ii) and (iii), respectively.

Second, we consider the remaining assumptions of Theorem 3.7 with $x^\nu = x$ and $E^P[g(s(\xi, x), x)] < +\infty$. Assumption (i) of Theorem 3.7 holds by our assumption (iv).

For P-a.e. $\xi \in \Xi$, our assumption (iii) implies that

$$\limsup_{(\nu,\zeta)\to(+\infty,\xi)} g^\nu\big(s^\nu(\zeta,x),x\big) \ge g\big(s(\xi,x),x\big).$$

Thus, Assumption (ii) of Theorem 3.7 holds and the theorem applies. □

4.4. Expectation constraints

While an expectation function might arise as the (effective) objective function in an optimization problem, and then the study of epi-convergence of its approximations becomes directly relevant, we may also have expectation functions arising as constraint functions. Still, epi-convergence remains a key tool and we again leverage Theorem 3.7.

For continuous $f_0 : \mathbb{X} \to \mathbb{R}$, integrands $\{f, f^\nu : \Xi \times \mathbb{X} \to \overline{\mathbb{R}}, \nu \in \mathbb{N}\}$, and probabilities $\{P, P^\nu, \nu \in \mathbb{N}\}$ on $(\Xi, \mathcal{B}(\Xi))$, suppose that we are faced with the problem

$$\operatorname*{minimize}_{x\in\mathbb{X}} \phi(x) := f_0(x) + \iota_{(-\infty,0]}\big(E^P[f](x)\big).$$

The expectation function needs to be approximated by $E^{P^\nu}[f^\nu]$. Suppose that the assumptions of Theorem 3.7 holds so that $E^{P^\nu}[f^\nu]$ epi-converges to $E^P[f]$. Then, the approximating problem

$$\operatorname*{minimize}_{x\in\mathbb{X}} \phi^\nu(x) := f_0(x) + \theta^\nu \max\big\{0, E^{P^\nu}[f^\nu](x)\big\}$$

is justified because ϕ^ν epi-converges to ϕ under the assumption that the scalar $\theta^\nu \to +\infty$ and the following constraint qualification holds: For each $x \in \mathbb{X}$ satisfying $E^P[f](x) = 0$, there exist points $x^\nu \in \mathbb{X} \to x$ such that $E^P[f](x^\nu) < 0$.

We see this as follows. It is straightforward to confirm $\liminf_{\nu\to+\infty} \phi^\nu(x^\nu) \ge \phi$ whenever $x^\nu \to x$. The limit condition in the definition of epi-convergence is confirmed by the following argument. Let $x \in \mathbb{X}$. Without loss of generality, we assume that $E^P[f](x) \le 0$. There are two cases.

(a) Suppose that $E^P[f](x) < 0$. Since $E^{P^\nu}[f^\nu]$ epi-converges to $E^P[f]$, there exist points $x^\nu \to x$ such that $\limsup_{\nu\to+\infty} E^{P^\nu}[f^\nu](x^\nu) \le E^P[f](x)$.

Thus, for sufficiently large ν, $E^{P^\nu}[f^\nu](x^\nu) \le 0$. This means that

$$\begin{aligned}\limsup_{\nu\to+\infty} \phi^\nu(x^\nu) &= \limsup_{\nu\to+\infty} f_0(x^\nu) + \limsup_{\nu\to+\infty} \theta^\nu \max\big\{0, E^{P^\nu}[f^\nu](x^\nu)\big\} \\ &\le f_0(x) = \phi(x). \end{aligned} \tag{44}$$

(b) Suppose that $E^P[f](x) = 0$. By assumption, there exist points $\bar{x}_k \in \mathbb{X} \to x$ such that $E^P[f](\bar{x}_k) < 0$ for $k \in \mathbb{N}$. Fix k. Since $E^{P^\nu}[f^\nu]$ epi-converges to $E^P[f]$, there exist points $x_k^\nu \to \bar{x}_k$ as $\nu \to +\infty$ such that

$$\limsup_{\nu\to+\infty} E^{P^\nu}[f^\nu](x_k^\nu) \le E^P[f](\bar{x}_k).$$

Thus, there is ν_k such that for all $\nu \ge \nu_k$, $E^{P^\nu}[f^\nu](x_k^\nu) \le 0$. Without loss of generality, we assume that $\nu_k > \nu_l$ for $k > l$. Construct for $k \in \mathbb{N}$: for $\nu_{k+1} > \nu \ge \nu_k$, set $x^\nu = x_k^{\nu_k}$. This produces $\{x^\nu\}_{\nu=\nu_1}^\infty$ with the property $x^\nu \to x$. Moreover, we have $E^{P^\nu}[f^\nu](x^\nu) = E^{P^\nu}[f^\nu](x_k^{\nu_k}) \le 0$ for some k. Thus, (44) holds again.

A. Proof of Lemma 3.1

The proof of Lemma 3.1 uses the following properties of the Pasch-Hausdorff envelope.

Lemma A.1. *For each $\kappa \in [0,+\infty)$ and $h : \mathbb{X} \to \overline{\mathbb{R}}$, let h_κ be its Pasch-Hausdorff envelope. Then the following properties are equivalent:*

(a) $h(x) = +\infty$ *for all* $x \in \mathbb{X}$;

(b) $h_\kappa(\bar{x}) = +\infty$ *for some* $\bar{x} \in \mathbb{X}$;

(c) $h_\kappa(x) = +\infty$ *for all* $x \in \mathbb{X}$.

Proof. The proof follows directly from the definition of the Pasch-Hausdorff envelope in (1). □

Proof of Lemma 3.1. Let us set, for $\nu \in \mathbb{N}$ and for $x \in \mathbb{X}$,

$$h^\nu(x) := E^{P^\nu}[f^\nu](x). \tag{45}$$

If there exists $\nu^0 \in \mathbb{N}$ such that $h^\nu(y) = +\infty$ for all $\nu \ge \nu_0$ and for all $y \in \mathbb{X}$, then, according to Lemma A.1(a,c), $h^\nu_\kappa(y) = +\infty$ for all $y \in \mathbb{X}$, where h^ν_κ is the Pasch-Hausdorff envelope of h^ν. In this case, the conclusions of Lemma 3.1 hold.

Otherwise, for each $\nu^0 \in \mathbb{N}$ there exists $\nu > \nu_0$ such that $h^\nu(y^\nu) < +\infty$ for some $y^\nu \in \mathbb{X}$. In this case, due to Lemma A.1(a,b), $h^\nu_\kappa(x) < +\infty$ for all $x \in \mathbb{X}$ and for all $\kappa \in (0,+\infty)$. Therefore, for each $x \in \mathbb{X}$

$$\begin{aligned} E^{P^\nu}\big[f^\nu_\kappa\big](x) &= \mathbb{E}^{P^\nu}\Big[\inf_{y\in\mathbb{X}}\big\{f^\nu(\xi,y)+\kappa d_{\mathbb{X}}(y,x)\big\}\Big] \le \inf_{y\in\mathbb{X}}\mathbb{E}^{P^\nu}\big[f^\nu(\xi,y)+\kappa d_{\mathbb{X}}(y,x)\big] \\ &= \inf_{y\in\mathbb{X}}\{h^\nu(y)+\kappa d_{\mathbb{X}}(y,x)\} = h^\nu_\kappa(x), \end{aligned} \tag{46}$$

where the first equality follows from the definition of

$$f^\nu_\kappa(\xi,x) = \inf_{y\in\mathbb{X}}\{f^\nu(\xi,y)+\kappa d_{\mathbb{X}}(y,x)\},$$

which is $\hat{\mathcal{B}}^\nu(\Xi)$-measurable in ξ in view of [16, Proposition 4.2], the inequality holds because the expectation of the measurable infimum is not greater than the infimum of the expectation, the second equality follows from (45), and the last equality follows from (1).

Fix an arbitrary $x_0 \in \mathbb{X}$. Inequalities (46) and (6) imply that for each $\kappa \in [\kappa_0,+\infty)$

$$-\infty < E^P\big[f_\kappa\big](x_0) \le \liminf_{\nu\to+\infty} h^\nu_\kappa(x_0). \tag{47}$$

Therefore, there exist $\nu(x_0) \in \mathbb{N}$ and $\gamma(x_0) \in (0,+\infty)$ such that

$$-\infty < -\gamma(x_0) \le h^\nu_\kappa(x_0) \quad \text{for each } \nu \ge \nu(x_0) \text{ and for each } \kappa \ge \kappa_0. \tag{48}$$

Since (47) implies (4), then, according to (5),

$$\liminf_{(\nu,y)\to(+\infty,x)} h^\nu(y) = \sup_{\kappa\in\mathbb{N}} \liminf_{\nu\to+\infty} h^\nu_\kappa(x) \quad \text{for each } x \in \mathbb{X}. \tag{49}$$

We fix an arbitrary $\kappa > \kappa_0$ and prove that either the function $x \mapsto \liminf_{\nu\to+\infty} h^\nu_\kappa(x)$ is identically equals to $+\infty$ or it is real-valued and Lipschitz continuous with modulus κ. Indeed, let

$$\liminf_{\nu\to+\infty} h^\nu_\kappa(\tilde{x}) < +\infty \text{ for some } \tilde{x} \in \mathbb{X}. \tag{50}$$

For $x \in \mathbb{X}$ we set $I(x) := \{\nu \in \mathbb{N} : \nu \geq \nu(x_0) \text{ and } h^\nu_\kappa(x) < +\infty\}$.

Due to Lemma A.1(b,c), $I := I(x) = I(y)$ for all $x, y \in \mathbb{X}$. Moreover, (50) implies that I is an infinite set. Therefore, there exists a strictly increasing sequence $\{\nu_m\}_{m\in\mathbb{N}} \subset \mathbb{N}$ such that $I = \{\nu_m : m \in \mathbb{N}\}$ and

$$\liminf_{\nu\to+\infty} h^\nu_\kappa(x) = \liminf_{m\to+\infty} h^{\nu_m}_\kappa(x) \tag{51}$$

for each $x \in \mathbb{X}$. Therefore, the definition of I, (48), and (3) imply that each function $h^{\nu_m}_\kappa$, $m \in \mathbb{N}$, is real-valued and Lipschitz continuous with modulus κ. Therefore,

$$-\infty < h^{\nu_m}_\kappa(y) - \kappa d_{\mathbb{X}}(x, y) \leq h^{\nu_m}_\kappa(x) \leq h^{\nu_m}_\kappa(y) + \kappa d_{\mathbb{X}}(x, y) < +\infty \tag{52}$$

for each $x, y \in \mathbb{X}$. Setting $y = x_0$ in (52) and passing to the lower limit as $m \to +\infty$ in the second inequality, we obtain that

$$-\infty < \liminf_{m\to+\infty} h^{\nu_m}_\kappa(x_0) - \kappa d_{\mathbb{X}}(x, x_0) \leq \liminf_{m\to+\infty} h^{\nu_m}_\kappa(x) \tag{53}$$

for each $x \in \mathbb{X}$, where the first inequality follows from (47). Similarly, setting $y = \tilde{x}$ in (52) and passing to the lower limit as $m \to +\infty$ in the third inequality, we obtain that

$$\liminf_{m\to+\infty} h^{\nu_m}_\kappa(x) \leq \liminf_{m\to+\infty} h^{\nu_m}_\kappa(\tilde{x}) + \kappa d_{\mathbb{X}}(x, \tilde{x}) < +\infty \tag{54}$$

for each $x \in \mathbb{X}$, where the second inequality follows from (50). Finally, (51), (53) and (54) imply that, under (50), the function $x \mapsto \liminf_{\nu\to+\infty} h^\nu_\kappa(x)$ is real-valued. Moreover, if we pass to the lower limit as $m \to +\infty$ in the second and the third inequalities of (52) for each $x, y \in \mathbb{X}$ and use equality (51) in each summand, then we obtain that the function $x \mapsto \liminf_{\nu\to+\infty} h^\nu_\kappa(x)$ is Lipschitz continuous with modulus κ.

If the function $x \mapsto \liminf_{\nu\to+\infty} h^\nu_\kappa(x)$ is identically equal to $+\infty$, then, according to (49), the conclusions of Lemma 3.1 hold. To finish the proof let us consider the case, when the function $x \mapsto \liminf_{\nu\to+\infty} h^\nu_\kappa(x)$ is real-valued and Lipschitz continuous with modulus κ.

Let us prove that the function $E^P[f_\kappa]$ is real-valued and Lipschitz continuous with modulus κ. Indeed, since $\liminf_{\nu\to+\infty} h^\nu_\kappa(x_0) < +\infty$, then inequalities (47) imply that $E^P\big[f_\kappa\big](x_0)$ is a real number. Thus,

$$f_\kappa(\xi, x_0) \text{ is finite for } P\text{-a.e. } \xi \in \Xi. \tag{55}$$

Therefore, according to (3), for P-a.e. $\xi \in \Xi$ the function $x \mapsto f_\kappa(\xi, x)$ is real-valued and Lipschitz continuous with modulus κ, that is,

$$f_\kappa(\xi, y) - \kappa d_{\mathbb{X}}(x, y) \leq f_\kappa(\xi, x) \leq f_\kappa(\xi, y) + \kappa d_{\mathbb{X}}(x, y) \tag{56}$$

for each $x, y \in \mathbb{X}$.

The function $E^P[f_\kappa]$ is real-valued because

$$-\infty < E^P\big[f_\kappa\big](x_0) - \kappa d_{\mathbb{X}}(x, x_0) \le E^P\big[f_\kappa\big](x) \le E^P\big[f_\kappa\big](x_0) + \kappa d_{\mathbb{X}}(x, x_0) < +\infty,$$

where the first and the last inequalities hold because $E^P[f_\kappa](x_0)$ is a real number, and the second and the third inequalities follow from inequalities (56) with $y = x_0$ if we take the expectation $\mathbb{E}^P$ of them. Now let us take the expectation $\mathbb{E}^P$ of (56) for each $x, y \in \mathbb{X}$. We obtain that the function $E^P[f_\kappa](x)$ is Lipschitz continuous with modulus κ.

The second inequality in (47) holds for each $x \in \mathbb{X}$ because both sides of this inequality are real-valued and Lipschitz with modulus κ. Therefore,

$$-\infty < \sup_{\kappa\in\mathbb{N}} E^P\big[f_\kappa\big](x) \le \sup_{\kappa\in\mathbb{N}} E^P\big[f_\kappa\big](x) = \liminf_{(\nu,y)\to(+\infty,x)} E^{P^\nu}[f^\nu](x) \tag{57}$$

for each $x \in \mathbb{X}$, where the equality follows from (49) and (45). Moreover, (55), (2) with $h := f(\xi, \cdot\,)$, and lsc of $f(\xi, \cdot\,)$ for P-a.e. $\xi \in \Xi$ imply that $f_\kappa(\xi, x) \uparrow f(\xi, x)$ for each $x \in \mathbb{X}$ and for P-a.e. $\xi \in \Xi$. Therefore, the first inequality in (57) and the monotone convergence theorem imply

$$\sup_{\kappa\in\mathbb{N}} E^P\big[f_\kappa\big](x) = E^P\big[f\big](x) \quad \text{for each } x \in \mathbb{X}. \tag{58}$$

Finally, (57)–(58) imply that $\liminf_{(\nu,y)\to(+\infty,x)} E^{P^\nu}[f^\nu](y) \ge E^P\big[f\big](x) > -\infty$. □

Acknowledgements. The research of the third author is supported in part by the Office of Naval Research, Science of Autonomy (N0001421WX00142) and the Air Force Office of Scientific Research, Mathematical Optimization (21RT0484).

References

[1] Z. Artstein, R. J-B Wets: *Consistency of minimizers and the SLLN for stochastic programs*, J. Convex Analysis 2/1-2 (1995) 1–17.

[2] H. Attouch: *Variational Convergence of Functions and Operators*, Applicable Mathematics Series, Pitman, London (1984).

[3] H. Attouch, R. J-B Wets: *Epigraphical processes: Laws of large numbers for random lsc functions*, Séminaire d'Analyse Convexe, Montpellier 13 (1990) 1–29.

[4] G. Beer, R. T. Rockafellar, R. J-B Wets: *A characterization of epi-convergence in terms of convergence of level sets*, Proc. Amer. Math. Soc. 116/3 (1992) 753–761.

[5] L. L. Chen, J. O. Royset: *Rockafellian relaxation in optimization under uncertainty: Asymptotically exact formulations*, arXiv: 2204.04762 (2022).

[6] P. Chen, J. O. Royset: *Performance bounds in PDE-constrained optimization under uncertainty*, SIAM J. Optimization 33/3 (2023).

[7] C. Choirat, C. Hess, R. Seri: *A functional version of the Birkhoff ergodic theorem for a normal integrand: A variational approach*, Ann. Probability 31/1 (2003) 63–92.

[8] H. T. H. Diem: *Consistency of statistical estimators of solutions to stochastic optimization problems*, J. Global Optimization 83 (2022) 825–842.

[9] J. Dupacova, R. J-B Wets: *Asymptotic behavior of statistical estimators and optimal solutions of stochastic optimization problems*, Ann. Statistics 16 (1988) 1517–1549.

[10] Yu. M. Ermoliev, V. I. Norkin, R. J-B Wets: *The minimization of semi-continuous functions: Mollifier subgradients*, SIAM J. Control Optimization 1 (1995) 149–167.

[11] E. A. Feinberg, P. O. Kasyanov, Y. Liang: *Fatou's lemma in its classical form and Lebesgue's convergence theorems for varying measures with applications to Markov decision processes*, Theory Prob. Appl. 65/2 (2020) 270–291.

[12] E. A. Feinberg, P. O. Kasyanov, Y. Liang: *Fatou's lemma for weakly converging measures under the uniform integrability condition*, Theory Prob. Appl. 64/4 (2020) 615–631.

[13] E. A. Feinberg, P. O. Kasyanov, N. V. Zadoianchuk: *Fatou's lemma for weakly converging probabilities*, Theory Prob. Appl. 58/4 (2014) 683–689.

[14] E. A. Feinberg, P. O. Kasyanov, M. Z. Zgurovsky: *Uniform Fatou's lemma*, J. Math. Analysis Appl. 444/1 (2016) 550–567.

[15] S. Geman, C.-R. Hwang: *Nonparametric maximum likelihood estimation by the method of sieves*, Ann. Statistics 10 (1982) 401–414.

[16] C. Hess: *Epi-convergence of sequences of normal integrands and strong consistency of the maximum likelihood estimator*, Ann. Statistics 24 (1996) 1298–1315.

[17] C. Hess, R. Seri: *Generic consistency for approximate stochastic programming and statistical problems*, SIAM J. Optimization 29/1 (2019) 290–317.

[18] A. J. King, R. J-B Wets: *Epi-consistency of convex stochastic programs*, Stoch. Stoch. Rep. 34/1-2 (1991) 83–92.

[19] L. A. Korf, R. J-B Wets: *Random lsc functions: An ergodic theorem*, Math. Oper. Res. 26/2 (2001) 421–445.

[20] R. Lucchetti, R. J-B Wets: *Convergence of minima of integral functionals, with applications to optimal control and stochastic optimization*, Stat. Decisions 11 (1993) 69–84.

[21] C. Phelps, J. O. Royset, Q. Gong: *Optimal control of uncertain systems using sample average approximations*, SIAM J. Control Optimization 54/1 (2016) 1–29.

[22] R. T. Rockafellar, R. J-B Wets: *Variational Analysis*, Springer, Berlin (1998).

[23] J. O. Royset: *Approximations of semicontinuous functions with applications to stochastic optimization and statistical estimation*, Math. Programming 184 (2020) 289–318.

[24] J. O. Royset, R. J-B Wets: *Variational theory for optimization under stochastic ambiguity*, SIAM J. Optimization 27/2 (2017) 1118–1149.

[25] J. O. Royset, R. J-B Wets: *Lopsided convergence: An extension and its quantification*, Math. Programming 177/1 (2019) 395–423.

[26] J. O. Royset, R. J-B Wets: *Variational analysis of constrained M-estimators*, Ann. Statistics 48/5 (2020) 2759–2790.

[27] J. O. Royset, R. J-B Wets: *An Optimization Primer*, Springer, Berlin (2021).

[28] A. Schwartz, E. Polak: *Consistent approximations for optimal control problems based on Runge-Kutta integration*, SIAM J. Control Optimization 34/4 (1996) 1235–1269.

[29] A. Sur, J. R. Birge: *Asymptotic behavior of solutions: An application to stochastic NLP*, Math. Programming 191 (2022) 281–306.