Math. Program., Ser. A (2019) 177:395–423
https://doi.org/10.1007/s10107-018-1275-3

FULL LENGTH PAPER

Lopsided convergence: an extension and its quantification

Johannes O. Royset[1] · **Roger J-B Wets**[2]

Received: 23 August 2016 / Accepted: 6 April 2018 / Published online: 26 May 2018

Abstract Much of the development of lopsided convergence for bifunctions defined on product spaces was in response to motivating applications. A wider class of applications requires an extension that would allow for arbitrary domains, not only product spaces. This leads to an extension of the definition and its implications that include the convergence of solutions and optimal values of a broad class of minsup problems. In the process we relax the definition of lopsided convergence even for the classical situation of product spaces. We now capture applications in optimization under stochastic ambiguity, Generalized Nash games, and many others. We also introduce the lop-distance between bifunctions, which leads to the first quantification of lopsided convergence. This quantification facilitates the study of convergence rates of methods for solving a series of problems including minsup problems, Generalized Nash games, and various equilibrium problems.

Keywords Lopsided convergence · Lop-convergence · Lop-distance · Attouch–Wets distance · Epi-convergence · Hypo-convergence · Minsup problems · Generalized Nash games

Mathematics Subject Classification 49M99 · 65K10 · 90C15 · 90C33 · 90C47

✉ Johannes O. Royset
joroyset@nps.edu

Roger J-B Wets
rjbwets@ucdavis.edu

1 Operations Research Department, Naval Postgraduate School, Monterey, CA, USA

2 Department of Mathematics, University of California, Davis, Davis, CA, USA

1 Introduction

The notion of *lopsided convergence* of bifunctions (= bivariate functions defined on a product space) was introduced in [3] for extended real-valued bifunctions. The focus on extended real-valued functions was motivated by the incentive to keep the development in concordance with the elegant "duality" results of Rockafellar [25, Chapters 33–37] and the subsequent convergence theory for saddle functions [2,4,19]. But this paradigm turned out to become unmanageable when confronted with a series of applications that required dealing with bifunctions that were not of the convex-concave type. Eventually, this led to restricting the convergence theory for bifunctions to real-valued bivariate functions (only) defined on specific subsets of the product space, cf. [21,28] and especially [20]. Lopsided convergence is emerging as a central tool in the study of linear and nonlinear complementarity problems, fixed points, variational inequalities, inclusions, noncooperative games, mathematical programs with equilibrium constraints, optimality conditions, Walras and Nash equilibrium problems, optimization under stochastic ambiguity, and robust optimization; see the recent developments in [21,28,29]. Already, Aubin and Ekeland [7, Chapter 6] brought to the fore the ineluctable connections between some of these applications when dealing with existence issues.

Prior studies deal exclusively with bifunctions defined on a product space, but applications in Generalized Nash games, robust optimization, stochastic optimization with decision-dependent measures, and Generalized semi-infinite programming require extensions to bifunctions for which the second variable's domain depend on the first variable. For example, in a minsup problem[1] this corresponds to the situation when the inner maximization has a feasible region that depends on the outer minimization variable, in a Generalized Nash game, the need arises when the set of feasible actions for any given agent depends on the actions of the other agents.

We extend the definition of lopsided convergence to deal with these situations and establish an array of results addressing this wider setting. Specifically, we show that for this extended notion of lopsided convergence, optimal solutions and optimal values of approximating minsup problems tend to those of an original minsup problem. We also relax the definition of lopsided convergence for bifunctions defined on a product space and, therefore, broaden the area of application even in this classical situation. In the process, we recast, and in a couple of instances refine, the fundamental implications of epi-convergence in the present framework, i.e., for finite-valued functions defined on a subset of a metric space.

For the first time we quantify lopsided convergence by defining the *lop-distance*. The lop-distance between two bifunctions is given in terms of the Attouch–Wets distance [5] between the sup-projections of the bifunctions with respect to the second variable. Thus, we place firmly the emphasis on the outer minimization in a minsup problem at the expense of the inner maximization. This imbalance indeed motivated the terminology "lopsided." In the context of minsup problems, Generalized Nash game, and many other situations, this perspective is reasonable as the inner problem is

[1] We prefer "minsup problem" to "minimax problem" as the inner maximization may not be attained in much of our development.

certainly secondary as illustrated below; the solution of the outer minimization being primary. For example, this leads to estimates of the rate of convergence of that outer solution as demonstrated in [29]. We note that the point of view differs from that of epi/hypo-convergence [4] and analysis of the convex/concave case [25, Chapters 33–37]. There the focus is on finding saddle-point *pairs*, which implies a certain balance between the inner and outer problems and indeed symmetry in the convex/concave case. Of course, our new viewpoint remains applicable in the convex/concave case and it yields somewhat sharper results not discussed here.

The article proceeds in Sect. 2 with a couple of motivating examples. In Sect. 3, we give the new definition, provide sufficiency conditions, and also discuss foundations related to epi-convergence. Consequences of lopsided convergence, with illustrations from Generalized Nash games, are established in Sect. 4 and the lop-distance is introduced in Sect. 5.

Throughout, we let (X, d_X) and (Y, d_Y) be two metric spaces and consider finite-valued bifunctions defined on nonempty subsets of $X \times Y$.

2 Motivation

In contrast to this article predecessors', e.g., [28], that deal with bifunctions of the form $F : C \times D \subset X \times Y \to \mathbb{R}$ with D a (fixed) subset of Y, we consider bifunctions with arbitrary domains. That is, the set $D \subset Y$ of permissible values for the second variable might depend on $x \in X$, the first variable. This extension is essential to deal with two main application areas as illustrated next, but also a wide range of generalizations of models in [21].

2.1 Optimization under stochastic ambiguity

Consider the minsup problem

$$\min_{x \in C} \sup_{y \in D(x)} F(x, y),$$

where $D : C \rightrightarrows Y$ is a *set-valued mapping*, $C = \operatorname{dom} D = \{x \;:\; D(x) \neq \emptyset\} \subset X$, and Y is the space of probability distributions defined on $\mathbb{R}^m$, with an appropriate metric. It identifies an optimization model with stochastic ambiguity where C is a collection of feasible decisions, $D(x)$ an *ambiguity set* of probability distributions for every $x \in C$, and F a bifunction that depends both on the decision and the distribution. For example, $F(x, P) = \mathbb{E}^P[\varphi(x, \boldsymbol{\xi})]$, where $\boldsymbol{\xi}$ is a random vector with distribution function P and the expectation is therefore taken with respect to a distribution function that is determined by the inner maximization problem. In applications, it is sometimes crucial to allow the ambiguity set to depend in a nontrivial manner on the decision x to capture situations where the decision maker affects the uncertainty as modeled here via the set $D(x)$; see [29,33] as well as [11,13,30,31] for related models. In [29], we leverage the results of the present paper to establish convergence of solutions of approximating optimization problems on $\mathbb{R}^n$ under stochastic ambiguity to those of an actual problem and also illustrate the use of the lop-distance defined in Sect. 5 for that context.

2.2 Generalized Nash games

As we shall see, the study of Generalized Nash games naturally leads to the study of bifunctions that are defined on a (proper) subset of a product space. An *equilibrium* of a Generalized Nash game with a finite set A of agents, is a solution $\bar{x} = (\bar{x}_a,\ a \in A)$ that satisfies

$$\bar{x}_a \in \operatorname*{argmin}_{x_a \in D_a(\bar{x}_{-a})} c_a(x_a, \bar{x}_{-a}), \quad \text{for all } a \in A,$$

where $x_{-a} = (x_{a'} : a' \in A\backslash\{a\})$, c_a is the cost function for agent a, and $D_a(\bar{x}_{-a})$ is the set of available strategies for agent a, which depends on the choices of strategies by the other agents.

The idea of using bifunctions to characterize equilibria of such games goes back at least to [24]; see also the review [14]. One approach leverages the *Nikaido–Isoda bifunction*, which is given by

$$F(x, y) = \sum_{a \in A} \Big[c_a(x_a, x_{-a}) - c_a(y_a, x_{-a})\Big] \quad \text{for } x \in C, \quad y \in D(x)$$

with

$$C = \Big\{x \ :\ x_a \in D_a(x_{-a}) \quad \text{for all } a \in A\Big\} \text{ and } D(x) = \prod_{a \in A} D_a(x_{-a}).$$

Clearly, the (effective) domain of this bifunction might be rather involved.

To align with the notation elsewhere, we think of $D_a(x_{-a})$ as a subset of a metric space X_a of conceivable strategies for agent a and the underlying metric space for the first variable in the bifunction is $X = \prod_{a \in A} X_a$, equipped with the product metric. Thus, $C \subset X$. Certainly, $D_a : X_{-a} \rightrightarrows X_a$, where $X_{-a} = \prod_{a' \in A\backslash a} X_{a'}$, but the mapping D can be restricted to C as other points are irrelevant, i.e., $D : C \rightrightarrows X$. We therefore have that the other underlying metric space for the second variable in the bifunction $Y = X$ in this case. Moreover, $c_a : X_a \times X_{-a} \to \mathbb{R}$. In our notation, highlighting the role of minsup-points, we then obtain the following characterization (see, e.g., [14]).

Proposition 2.1 (Characterization of equilibrium) *In the notation of this subsection, $\bar{x}$ is an equilibrium if and only if it is a minsup-point of F with nonpositive minsup-value, i.e.,*

$$\bar{x} \in \operatorname*{argmin}_{x \in C} \sup_{y \in D(x)} F(x, y) \textit{ and } \sup_{y \in D(\bar{x})} F(\bar{x}, y) \le 0.$$

Proof If $\bar{x}$ is a Nash equilibrium, then

$$c_a(\bar{x}_a, \bar{x}_{-a}) \leq c_a(y_a, \bar{x}_{-a}) \quad \text{for all } y_a \in D_a(\bar{x}_{-a}), a \in A.$$

Thus, $F(\bar{x}, y) \leq 0$ for all $y \in D(\bar{x})$. For any $x \in C$, $x \in D(x)$ and therefore $\sup_{y \in D(x)} F(x, y) \geq 0$. In particular,

$$\sup_{y \in D(\bar{x})} F(\bar{x}, y) = F(\bar{x}, \bar{x}) = 0.$$

Consequently, $\bar{x} \in \operatorname{argmin}_{x \in C} \sup_{y \in D(x)} F(x, y)$. For the converse, let $\bar{x}$ be a minsup-point of F and $\sup_{y \in D(\bar{x})} F(\bar{x}, y) \leq 0$. Then,

$$0 \geq \sup_{y \in D(\bar{x})} F(\bar{x}, y) = \sum_{a \in A} \Big[c_a(\bar{x}_a, \bar{x}_{-a}) - \inf_{y_a \in D_a(\bar{x}_{-a})} c_a(y_a, \bar{x}_{-a}) \Big].$$

The upper bound of zero and the fact that $\bar{x}_a \in D_a(\bar{x}_{-a})$ for all $a \in A$ imply that each term in the sum must be zero and the conclusion follows. □

In Sect. 4.3, we show how stability and approximation of the solutions of Generalized Nash games can be examined through the lense of lopsided convergence. The results are widely applicable as the strategy spaces of the agents are only required to be metric spaces and an agent's constraint set may neither be convex nor compact and could very well depend on the other agents' choices of strategies—a centerpiece of the new definition of lopsided convergence. Thus, we extend results in [18,21] that both consider finite-dimensional strategies and constraint sets independent of other agents' choices of strategies.

3 Lopsided convergence

As already mentioned in the Introduction, to encompass the family of applications sketched out in Sect. 2 and others, we need to extend the definition of lopsided convergence, and the resulting theory, to a larger class of bifunctions than those considered in earlier work. The *domain*, $\operatorname{dom} F \neq \emptyset$, of a finite-valued bifunction F will no longer be restricted to a product subset of $X \times Y$ but could be *any* subset. The family of all such bifunctions is denoted

$$\operatorname{bfcns}(X, Y) := \big\{ F : \operatorname{dom} F \to \mathbb{R} \; : \; \emptyset \neq \operatorname{dom} F \subset X \times Y \big\}.$$

It is on this family we introduce the new definition of lopsided convergence.

3.1 Definition

The (first) x-variable of a bifunction takes a primary role in our development leading us to the following description of the domain of a bifunction. We associate with a bifunction F, the set

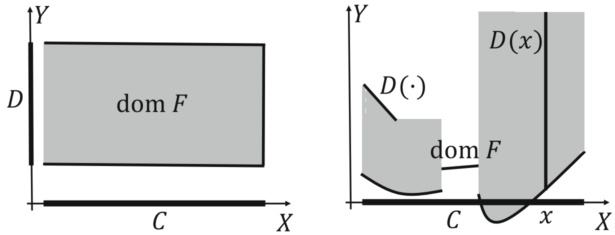

Fig. 1 Domains of bifunctions: product set (left) and general (right)

$$C := \big\{x \in X \ : \ \exists\, y \in Y \ \text{ such that } (x, y) \in \operatorname{dom} F\big\}$$

and the set-valued mapping $D : C \rightrightarrows Y$ such that

$$D(x) := \big\{y \in Y \ : \ (x, y) \in \operatorname{dom} F\big\} \quad \text{for } x \in C.$$

Thus, $\operatorname{dom} F = \big\{(x, y) \in X \times Y \ : \ x \in C, \ \ y \in D(x)\big\}$, with $\operatorname{dom} D = \big\{x \in X \ : \ D(x) \neq \emptyset\big\} = C$. When $\operatorname{dom} F$ is a product set it agrees with having D a constant mapping; the output of that mapping is then also denoted by D (instead of $D(x)$). Figure 1 illustrates the case with a product set (left portion) and the general case (right portion). Throughout, $\operatorname{dom} F$ is described in terms of such C and D.

In applications, a bifunction of interest might be defined on a "large" subset of $X \times Y$, possibly everywhere, but the context requires restrictions to some "smaller" subset for example dictated by constraints imposed on the variables. If these constraints are that $x \in C \subset X$ and $y \in D(x)$ for some $D : C \rightrightarrows Y$, then F in our notation would become the original bifunction restricted to the set $\{(x, y) \in X \times Y \ : \ x \in C, y \in D(x)\}$, which becomes $\operatorname{dom} F$. In other applications, a bifunction F might be the only problem data given. These differences are immaterial to the following development as both are captured by considering $\operatorname{bfcns}(X, Y)$.

With $\mathbb{N} := \{1, 2, 3, \ldots\}$, collections of bifunctions are often denoted by $\{F^\nu, \nu \in \mathbb{N}\} \subset \operatorname{bfcns}(X, Y)$. Analogously to the notation above, $\operatorname{dom} F^\nu$ is described by a set $C^\nu \subset X$ and a set-valued mapping $D^\nu : C^\nu \rightrightarrows Y$ such that $\operatorname{dom} F^\nu = \big\{(x, y) \in X \times Y \ : \ x \in C^\nu, \ \ y \in D^\nu(x)\big\}$. If not specified otherwise, the index ν runs over $\mathbb{N}$ so that $x^\nu \to x$ means that the whole sequence $\{x^\nu, \nu \in \mathbb{N}\}$ converges to x. Let

$$\mathcal{N}_\infty^\# \text{ be all the subsets of } \mathbb{N} \text{ determined by subsequences,}$$

i.e., $N \in \mathcal{N}_\infty^\#$ is an infinite collection of strictly increasing natural numbers. Thus, $\{x^\nu, \nu \in N\}$ is a subsequence of $\{x^\nu, \nu \in \mathbb{N}\}$; its convergence to x is noted by $x^\nu \overset{N}{\to} x$.

The extended definition of lopsided convergence takes the following form.

Definition 3.1 (*Lopsided convergence*) Let $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \operatorname{bfcns}(X, Y)$ with domains described by (C, D) and (C^ν, D^ν), respectively. Then, F^ν *converges lopsided, or lop-converges,* to F, written $F^\nu \xrightarrow{lop} F$, when

(a) $\forall N \in \mathcal{N}_\infty^\#$, $x^\nu \in C^\nu \xrightarrow{N} x \in C$, and $y \in D(x)$, $\exists\, y^\nu \in D^\nu(x^\nu) \xrightarrow{N} y$ such that $\liminf_{\nu\in N} F^\nu(x^\nu, y^\nu) \geq F(x, y)$ and
$\forall N \in \mathcal{N}_\infty^\#$ and $x^\nu \in C^\nu \xrightarrow{N} x \notin C$, $\exists\, y^\nu \in D^\nu(x^\nu)$ such that $F^\nu(x^\nu, y^\nu) \xrightarrow{N} \infty$;
(b) $\forall\, x \in C$, $\exists\, x^\nu \in C^\nu \to x$ such that $\forall N \in \mathcal{N}_\infty^\#$ and $y^\nu \in D^\nu(x^\nu) \xrightarrow{N} y \in Y$, $\limsup_{\nu\in N} F^\nu(x^\nu, y^\nu) \leq F(x, y)$ if $y \in D(x)$ and $F^\nu(x^\nu, y^\nu) \xrightarrow{N} -\infty$ otherwise.

Lop-convergence does not have a direct geometric interpretation. However, as discussed in Sect. 4, it is intimately tied to epi- and hypo-convergence, which are easily understood in terms of the convergence of epigraphs and hypographs; see Sect. 3.4. We can therefore understand, in part, lop-convergence through these geometric interpretations. A preview of the conclusions reached in Sect. 4 builds intuition at this stage: If the bifunctions F^ν and F do not depend on y, then lop-convergence "collapses" to epi-convergence. Under mild assumptions, $F^\nu \xrightarrow{lop} F$ implies that $\sup_{y\in D^\nu(\cdot)} F^\nu(\cdot, y)$ epi-converges to $\sup_{y\in D(\cdot)} F(\cdot, y)$. We also have that at every $x \in C$, $F^\nu(x^\nu, \cdot)$ hypo-converges to $F(x, \cdot)$ for some $x^\nu \to x$.

We note that even in the case when dom F and dom F^ν are product sets for all ν, our definition represents a relaxation of the requirements in prior definitions as $\{y^\nu, \nu \in \mathbb{N}\}$ in condition (a) of Definition 3.1 for the case $x \notin C$ is not required to converge to y.

Contrast with earlier definition for product sets. Suppose that $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \text{bfcns}(X, Y)$ have dom $F = C \times D$ and dom $F^\nu = C^\nu \times D^\nu$ for sets $D, D^\nu \subset Y$. In [20,28], lop-convergence is defined to take place when[2] (a') $\forall N \in \mathcal{N}_\infty^\#$, $x^\nu \in C^\nu \xrightarrow{N} x \in X$, and $y \in D$, $\exists\, y^\nu \in D^\nu \xrightarrow{N} y$ such that $\liminf_{\nu\in N} F^\nu(x^\nu, y^\nu) \geq F(x, y)$ if $x \in C$ and $F^\nu(x^\nu, y^\nu) \xrightarrow{N} \infty$ otherwise; and (b') $\forall\, x \in C$, $\exists\, x^\nu \in C^\nu \to x$ such that $\forall N \in \mathcal{N}_\infty^\#$ and $y^\nu \in D^\nu \to y \in Y$, $\limsup_{\nu\in N} F^\nu(x^\nu, y^\nu) \leq F(x, y)$ if $y \in D$ and $F^\nu(x^\nu, y^\nu) \xrightarrow{N} -\infty$ otherwise. The condition (b') is exactly condition (b) of Definition 3.1 for the case of product sets. However, condition (a') is stronger than condition (a) of Definition 3.1 as illustrated by the following trivial example. Let $X = Y = \mathbb{R}$, $C^\nu = C = (0, 1]$, $D^\nu = D = [0, 1]$, and $F^\nu(x, y) = F(x, y) = 1/(x + y)$ for (x, y) in the their domains. It is clear that (a') fails for $y = 1$, $x = 0$, and $x^\nu = 1/\nu$ as there is no $y^\nu \to y$ such that $F^\nu(x^\nu, y^\nu) \to \infty$. However, condition (a) of Definition 3.1 holds as one can take $y^\nu = 1/\nu$ in the case $x \notin C$. Then, $F^\nu(x^\nu, y^\nu) = \nu/2 \to \infty$. If $x \in C$, then one can take $y^\nu = y$ and obtain that $F^\nu(x^\nu, y^\nu) = F(x^\nu, y) \to F(x, y)$ by continuity of F. For condition (b) of Definition 3.1, one can take $x^\nu = x$ and only be concerned with $y \in Y$ due to the closedness of Y. Continuity of F then allows us to conclude that $F^\nu \xrightarrow{lop} F$. In this case, with $F^\nu = F$, convergence is indeed "natural" and Definition 3.1 addresses this situation.

3.2 About sufficiency

One can come up with a wide collection of sufficient conditions for lop-convergence in terms of the way the "components" of a sequence of bifunctions F^ν converge to those

[2] The need to check subsequences $N \in \mathcal{N}_\infty^\#$ is understood, but not explicitly stated in these references.

of the limiting bifunction F; note that it will be necessary to broaden convergence notions for functions and mappings to take into account the fact that the domains of these bifunctions are generally not identical. The following results are only meant to illustrate the possibilities and might not be as taut as possible and certainly not necessary. The more interesting ones come from specific applications such as those laid out in [12,21] involving bifunctions whose domains are product sets and the more general family, when the domains are not restricted to product sets, such as the examples in Sect. 4.3 and those described in [29].

In the following, convergence of sequences in $X \times Y$ are always in the sense of the product topology, i.e., $(x^\nu, y^\nu) \subset X \times Y \to (x, y) \in X \times Y$ if $\max\{d_X(x^\nu, x), d_Y(y^\nu, y)\} \to 0$. Convergence of sets are always in the sense of Painlevé–Kuratowski. Specifically, in a metric space, the *outer limit* of a sequence of sets $\{A^\nu, \nu \in \mathbb{N}\}$, denoted by OutLim A^ν, is the collection of points x to which a subsequence of $\{x^\nu \in A^\nu, \nu \in \mathbb{N}\}$ converges. The *inner limit*, denote by InnLim A^ν, is the points to which a sequence of $\{x^\nu \in A^\nu, \nu \in \mathbb{N}\}$ converges. If both limits exist and are identical to A, we say that A^ν (set-)converges to A, which is denoted by $A^\nu \to A$; see [9,26].

Theorem 3.2 (Sufficiency when $C = C^\nu$) *For bifunctions* $\{F, F^\nu, \nu \in \mathbb{N}\} \subset$ bfcns(X, Y) *with domains described by* (C, D) *and* (C^ν, D^ν)*, respectively,* $F^\nu \xrightarrow{lop} F$ *when*

(a) $C = C^\nu$, $\nu \in \mathbb{N}$, *are closed;*
(b) *the mappings D^ν continuously converge to D, relative to C, i.e.,* $\forall x^\nu \in C \to x \in C$, $D^\nu(x^\nu) \to D(x)$*; and*
(c) *the bifunctions F^ν continuously converge to F, relative to their domains, i.e.,* $\forall (x^\nu, y^\nu) \in \operatorname{dom} F^\nu \to (x, y) \in \operatorname{dom} F$, $F^\nu(x^\nu, y^\nu) \to F(x, y)$.

Proof Since C is closed, given any $x^\nu \in C \to x$ it always entails $x \in C$. Moreover, continuous convergence of the mappings D^ν to D, relative to C, implies $D^\nu(x^\nu) \to D(x)$ which, in turn, implies that for any $y \in D(x)$ one can find $y^\nu \in D^\nu(x^\nu)$ converging to y. From (c), the continuous convergence of the bifunctions F^ν to F, implies $F(x^\nu, y^\nu) \to F(x, y)$ which immediately yields condition (a) of Definition 3.1.

To verify condition (b) of Definition 3.1, given any $x \in C$, by choosing the sequence $\{x^\nu = x, \nu \in \mathbb{N}\}$, $D^\nu(x) \to D(x)$ follows from continuous convergence of the mappings. Thus, whenever $y^\nu \in D^\nu(x) \to y$, $y \in D(x)$. In turn, this means that we only have to check if $\limsup F^\nu(x, y^\nu) \le F(x, y)$ which, of course, is satisfied since $F^\nu(x, y^\nu) \to F(x, y)$ in view of assumption (c). □

Next, we deal with the situation when the domains of the bifunctions are product sets. In fact, the next statement can be viewed as a refinement of the earlier theorem. We recall that a function $f : C \to \mathbb{R}$, with $C \subset X$ and X any metric space, is lower semicontinuous (lsc) when, for all $x^\nu \in C \to x \in X$, $\liminf f(x^\nu) \ge f(x)$ if $x \in C$ and $f(x^\nu) \to \infty$ otherwise.

Proposition 3.3 (Sufficiency under product sets) *Let* $\{F : C \times D \to \mathbb{R}, F^\nu : C \times D^\nu \to \mathbb{R}, \nu \in \mathbb{N}\} \subset$ bfcns(X, Y) *with C closed and* $D^\nu \subset Y \to D \subset Y$*. If in terms of some lsc bifunction* $\widetilde{F} : X \times Y \to \mathbb{R}$,

$$F = \widetilde{F} \text{ on } C \times D \text{ and } F^\nu = \widetilde{F} \text{ on } C \times D^\nu,$$

then $F^\nu \xrightarrow{lop} F$*, provided that, for all* $x \in C$*,* $\widetilde{F}(x, \cdot)$ *is usc.*

Proof First, consider condition (a) of Definition 3.1, which now simplifies to finding, for every $y \in D$ and $x^\nu \in C \to x \in C$, a sequence $y^\nu \in D^\nu \to y$ such that $\liminf \widetilde{F}(x^\nu, y^\nu) \geq \widetilde{F}(x, y)$. This condition follows from the set-convergence of D^ν to D and lower semicontinuity of $\widetilde{F}$. Second, for condition (b) of Definition 3.1, we select $x^\nu = x \in C$ for all ν. Since $D^\nu \to D$, any sequence $y^\nu \in D^\nu \to y$, implies $y \in D$. Thus, the condition simplifies to $\limsup \widetilde{F}(x, y^\nu) \leq \widetilde{F}(x, y)$ which holds in view of the last assumption of the proposition. □

Finally, we record a result for the case when there are approximations in the set controlling the (first) x-variable, which may occur, for example, in sensitivity analysis, semi-infinite optimization, stochastic programming, and problems with an infinite-dimensional space X that all may trigger approximations.

Proposition 3.4 (Sufficiency when $C \neq C^\nu$) *For bifunctions* $\{F, F^\nu, \nu \in \mathbb{N}\} \subset$ bfcns(X, Y) *with domains described by* (C, D) *and* (C^ν, D^ν)*, respectively,* $F^\nu \xrightarrow{lop} F$ *when*

(a) $C^\nu \to C$*;*
(b) *for some continuous set-valued mapping* $\widetilde{D} : X \rightrightarrows Y$*,*

$$D = \widetilde{D} \text{ on } C \text{ and } D^\nu = \widetilde{D} \text{ on } C^\nu;$$

(c) *for some continuous bifunction* $\widetilde{F} : X \times Y \to \mathbb{R}$*,*

$$F = \widetilde{F} \text{ on } \operatorname{dom} F \text{ and } F^\nu = \widetilde{F} \text{ on } \operatorname{dom} F^\nu.$$

Proof Since $C^\nu \to C$, every $x^\nu \in C^\nu \to x$ must have $x \in C$. Thus, condition (a) of Definition 3.1 simplifies to finding, for every $x^\nu \in C^\nu \to x \in C$ and $y \in \widetilde{D}(x)$, a sequence $y^\nu \in \widetilde{D}(x^\nu) \to y$ with $\liminf \widetilde{F}(x^\nu, y^\nu) \geq \widetilde{F}(x, y)$. Since D is continuous, such a sequence exists and the inequality therefore follows from the continuity of $\widetilde{F}$. We next turn to condition (b) of Definition 3.1. Since $\widetilde{D}$ is continuous, $y^\nu \in \widetilde{D}(x^\nu) \to y$ implies that $y \in \widetilde{D}(x)$ whenever $x^\nu \to x$. Thus, the condition simplifies to finding, for every $x \in C$, a sequence $x^\nu \in C^\nu \to x$ such that $\limsup \widetilde{F}(x^\nu, y^\nu) \leq \widetilde{F}(x, y)$ for all $y^\nu \in \widetilde{D}(x^\nu) \to y \in \widetilde{D}(x)$. Since $C^\nu \to C$, there is certainly a sequence $x^\nu \in C^\nu \to x$ for all $x \in C$. The inequality then holds in view of the continuity of $\widetilde{F}$. □

It is easy to find generalizations of the preceding results, for example rather than requiring continuous convergence of the F^ν in Theorem 3.2 one could be satisfied with some "semicontinuous" convergence complemented with a pointwise upper semicontinuity condition. At this point, we shall not get involved in all the possibilities as eventually one is bound to be mostly interested in conditions that apply in specific applications. However, we caution that certain "natural" conditions are *not* sufficient as exemplified next.

Failure of lop-convergence under graphical convergence. Consider the following situation where $F^\nu = F$ for all $\nu \in \mathbb{N}$, with domains described by $C = C^\nu = \mathbb{R}$ and $D^\nu(x) = D(x) = [-1, 1]$ if $x \leq 0$, and $\{0\}$ otherwise. Certainly the mappings D^ν graphically converge to D since they are identical. However, when considering condition (a) of Definition 3.1 with $x^\nu > 0 \to x = 0$ and $y = 1/2 \in D(x)$, there are no $y^\nu \in D^\nu(x^\nu) \to y$ and lop-convergence fails. We note that in this case pointwise convergence $D^\nu(x) \to D(x)$ holds for all $x \in \mathbb{R}$ and, consequently, the mappings are equi-osc [26, Theorem 5.40]. We next give a more involved example where pointwise convergence again holds, but now for problems with different solutions.

Failure of lop-convergence under pointwise set-convergence. Suppose that $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \text{bfcns}(\mathbb{R}, \mathbb{R})$ with domains described by $C = C^\nu = [0, 1]$, $D(x) = D^\nu(x) = \{0\}$ for $x \in [0, 1)$, $D^\nu(1) = [0, 1 + 1/\nu]$, and $D(1) = [0, 1]$. Moreover, let $F(x, y) = F^\nu(x, y) = 0$ if $x \in [0, 1)$, and $F(1, y) = F^\nu(1, y) = -2 + y$ if $y \leq 1$ and $F^\nu(1, y) = 1$ if $y > 1$. Clearly, for every $x \in [0, 1]$, $D^\nu(x) \to D(x)$. However, lop-convergence of F^ν to F fails as for $x = 1$, $x^\nu = 1 - 1/\nu$, and $y = 1$, there exists no sequence $\{y^\nu, \nu \in \mathbb{N}\}$, with $y^\nu \in D^\nu(x^\nu) = \{0\}$ that converges to y as required by condition (a) of Definition 3.1. Here, $\sup_{y \in D^\nu(x)} F^\nu(x, y) = 0$ if $x \in [0, 1)$ and $\sup_{y \in D^\nu(x)} F^\nu(x, y) = 1$ if $x = 1$, and $\sup_{y \in D(x)} F(x, y) = 0$ if $x \in [0, 1)$ and $\sup_{y \in D(x)} F(1, y) = -1$. Thus, the optimal value of $\min_{x \in C^\nu} \sup_{y \in D^\nu(x)} F^\nu(x, y)$, which is 0, does not converge to the optimal value of $\min_{x \in C} \sup_{y \in D(x)} F(x, y)$, which is -1. Since a main purpose of a notion of variational convergence of bifunctions is to ensure convergence of such optimal values, it is clear that pointwise set-convergence is not strong enough.

3.3 Tightness

A slight strengthening of lop-convergence that amounts to a relaxed compactness assumption becomes beneficial in Sect. 4 when deriving consequences.

Definition 3.5 (*Ancillary-tight lop-convergence*) The lop-convergence of $\{F^\nu, \nu \in \mathbb{N}\} \subset \text{bfcns}(X, Y)$ to $F \in \text{bfcns}(X, Y)$ is ancillary-tight when for every $\varepsilon > 0$ and sequence $x^\nu \to x$ selected in condition (b) of Definition 3.1, there exists a compact set $B_\varepsilon \subset Y$ and an integer ν_ε such that

$$\sup_{y \in D^\nu(x^\nu) \cap B_\varepsilon} F^\nu(x^\nu, y) \geq \sup_{y \in D^\nu(x^\nu)} F^\nu(x^\nu, y) - \varepsilon \quad \text{for all } \nu \geq \nu_\varepsilon,$$

where D^ν describes dom F^ν.

As usual, we interpret the supremum over an empty subset of $\mathbb{R}$ as $-\infty$. The added requirement for ancillary-tightness is satisfied if all $D^\nu(x^\nu)$ are contained in a compact set, but many other possibilities exist. For example, in [29] ancillary-tightness is connected to "tightness" in the sense of probability theory when Y is a space of distribution functions with a metric corresponding to convergence in distribution. Applications in optimization under stochastic ambiguity can be of this form; see Sect. 2.1. In this case, if $\{D^\nu(x^\nu), \nu \in \mathbb{N}\}$ is tight in the sense of probability theory for every sequence

$x^\nu \to x$ selected in condition (b) of Definition 3.1, then lop-convergence implies ancillary-tight lop-convergence, i.e., ancillary-tightness is "automatic."

If ancillary-tightness is combined with a similar condition for the outer minimization, we obtain a further strengthening of the notion.

Definition 3.6 (*Tight lop-convergence*) The ancillary-tight lop-convergence of bifunctions $\{F^\nu, \nu \in \mathbb{N}\} \subset \mathrm{bfcns}(X, Y)$ to $F \in \mathrm{bfcns}(X, Y)$ is tight when for any $\varepsilon > 0$ one can find a compact set $A_\varepsilon \subset X$ and an integer ν_ε such that

$$\inf_{x\in C^\nu \cap A_\varepsilon} \sup_{y\in D^\nu(x)} F^\nu(x, y) \le \inf_{x\in C^\nu} \sup_{y\in D^\nu(x)} F^\nu(x, y) + \varepsilon \quad \text{for all } \nu \ge \nu_\varepsilon,$$

where (C^ν, D^ν) describes dom F^ν.

The infimum of an empty subset of $\mathbb{R}$ is interpreted as ∞. This further strengthening of the requirements would be satisfied if all C^ν are contained in a compact set, but this is certainly not a necessity.

3.4 Epi- and hypo-convergence: a summary

Before we develop consequences of lop-convergence, we give some background facts about epi- and hypo-convergence for (univariate) functions; see [1,8,26] for comprehensive treatments. We present results for real-valued functions defined on (nonempty) subsets of (X, d_X); in many ways, this is just a variant of the more traditional framework that considers extended real-valued functions, cf. [8,26], but slight sharpening of some results become possible. Our focus will, thus, be on

$$\mathrm{fcns}(X) = \left\{ f : C \to \mathbb{R} \ : \quad \text{for some } \emptyset \ne C \subset X \right\}.$$

In this section, for functions $f, f^\nu \in \mathrm{fcns}(X)$, we denote by C and C^ν their domains, respectively. The *epigraph* of f, epi $f \subset X \times \mathbb{R}$, consists of all points that lie on or above the graph of f; it is lsc if epi f is a closed subset of $X \times \mathbb{R}$ (with respect to the product topology generated by d_X and the usual metric on $\mathbb{R}$) and, provided X is a linear space,[3] it is convex if its epigraph is convex. The *hypograph* of f, hypo $f \subset X \times \mathbb{R}$, consists of all points that lie on or below the graph of f; it is upper semicontinuous (usc) if hypo f is closed and, provided X is a linear space, it is concave if its hypograph is convex.

A sequence of functions $\{f^\nu, \nu \in \mathbb{N}\} \subset \mathrm{fcns}(X)$ *epi-converges* to a function $f \in \mathrm{fcns}(X)$, written $f^\nu \xrightarrow{e} f$, when the epigraphs epi f^ν set-converge to epi f; similarly, they *hypo-converge*, written $f^\nu \xrightarrow{h} f$, if the hypographs hypo f^ν set-converge to hypo f. Equivalently, epi-convergence can also be defined as follows:

Definition 3.7 (*Epi- and hypo-convergence*) For $\{f, f^\nu, \nu \in \mathbb{N}\} \subset \mathrm{fcns}(X)$ with domains C and C^ν, respectively, we have that $f^\nu \xrightarrow{e} f$ if and only if

[3] Statements about convexity/concavity are the only ones that require a linear space in this paper.

(a) $\forall N \in \mathcal{N}_\infty^\#$ and $x^\nu \in C^\nu \xrightarrow{N} x$, $\liminf_{\nu\in N} f^\nu(x^\nu) \geq f(x)$ if $x \in C$ and $f^\nu(x^\nu) \xrightarrow{N} \infty$ otherwise,
(b) $\forall x \in C$, $\exists\, x^\nu \in C^\nu \to x$ such that $\limsup f^\nu(x^\nu) \leq f(x)$.

The functions f^ν are said to *epi-converge tightly* to f when $f^\nu \xrightarrow{e} f$ and for all $\varepsilon > 0$, one can find a compact set $B_\varepsilon \subset X$ and an index ν_ε such that

$$\forall\, \nu \geq \nu_\varepsilon : \quad \inf_{x\in C^\nu\cap B_\varepsilon} f^\nu(x) \leq \inf_{x\in C^\nu} f^\nu(x) + \varepsilon.$$

Moreover, $f^\nu \xrightarrow{h} f$ if and only if $-f^\nu \xrightarrow{e} -f$ and they hypo-converge tightly if the functions $-f^\nu$ epi-converge tightly to $-f$.

As follows immediately from the properties of set-limits, an epi-limit is always lsc and, provided that X is linear, it is convex whenever the functions f^ν are convex. Moreover, a hypo-limit is always usc and, provided that X is linear, it is concave whenever the functions f^ν are concave. The topology induced by epi-convergence is metrizable, a property we leverage in Sect. 5.

For $f \in \mathrm{fcns}(X)$ with domain C, optimal values are denoted by

$$\inf f := \inf\{f(x) \ : \ x \in C\} \ \text{ and } \ \sup f := \sup\{f(x) \ : \ x \in C\},$$

and, with $\varepsilon \geq 0$, (near-)optimal solutions by

$$\begin{aligned}\varepsilon\text{-}\operatorname{argmin} f &:= \{x \in C \ : \ f(x) \leq \inf f + \varepsilon\} \text{ and}\\ \varepsilon\text{-}\operatorname{argmax} f &:= \{x \in C \ : \ f(x) \geq \sup f - \varepsilon\}.\end{aligned}$$

Since C is nonempty because of $f \in \mathrm{fcns}(X)$, $\inf f < \infty$. Moreover, $\inf f = -\infty$ implies that $\operatorname{argmin} f = \emptyset$. Convergence of optimal solutions and optimal values are summarized in the next theorem. The result and proof are mostly the same as those of [26, Theorem 7.31] and [20, Theorems 2.5 and 2.8], which consider $X = \mathbb{R}^n$, but stated here for completeness with some clarification and improvements, especially regarding the role of finiteness of $\inf f$ and convergence of near-optimal solutions. We refer to [4] for early results of this kind.

Theorem 3.8 (Epi- and hypo-convergence: basic properties) *Consider* $\{f, f^\nu, \ \nu \in \mathbb{N}\} \subset \mathrm{fcns}(X)$. *If* $f^\nu \xrightarrow{e} f$, *then the following hold:*

(a) $\limsup\big(\inf f^\nu\big) \leq \inf f$ *and* $\forall\{\varepsilon^\nu \downarrow 0, \nu \in \mathbb{N}\}$, $\mathrm{OutLim}\big(\varepsilon^\nu\text{-}\operatorname{argmin} f^\nu\big) \subset \operatorname{argmin} f$.
(b) *If* $\{x^\nu \in \operatorname{argmin} f^\nu, \nu \in N\}$ *converges for some* $N \in \mathcal{N}_\infty^\#$, *then* $\lim_{\nu\in N}(\inf f^\nu) = \inf f$.
(c) $\inf f^\nu \to \inf f > -\infty \Longleftrightarrow f^\nu \xrightarrow{e} f$ *tightly.*
(d) $\inf f^\nu \to \inf f$ *and* $\varepsilon > 0 \Longrightarrow \mathrm{InnLim}(\varepsilon\text{-}\operatorname{argmin} f^\nu) \supset \operatorname{argmin} f$.
(e) $\inf f^\nu \to \inf f$ *and* X *is separable*[4] $\Longrightarrow \exists\{\varepsilon^\nu \downarrow 0, \nu \in \mathbb{N}\}$ *such that* $\varepsilon^\nu\text{–}\operatorname{argmin} f^\nu \to \operatorname{argmin} f$.

[4] We deduce from a counterexample in [10] that the separability assumption cannot be relaxed.

(f) $\exists\{\varepsilon^\nu \downarrow 0, \nu \in \mathbb{N}\}$ *such that* $\varepsilon^\nu\text{-}\operatorname{argmin} f^\nu \to \operatorname{argmin} f \neq \emptyset \Longrightarrow \inf f^\nu \to \inf f > -\infty$.

If $f^\nu \xrightarrow{h} f$, *then* $\operatorname{liminf}_\nu(\sup f^\nu) \geq \sup f$ *and (a)–(f) hold with min/inf replaced by max/sup,* $> -\infty$ *by* $< \infty$*, and tight epi-convergence by tight hypo-convergence.*

Proof Let C and C^ν be the domains of f and f^ν, respectively. For part (a), we first suppose that $\inf f$ is finite and let $\varepsilon > 0$. There exists $x \in C$ such that $f(x) \leq \inf f + \varepsilon$ and also, by condition (b) of Definition 3.7, $x^\nu \in C^\nu \to x$ such that $\operatorname{limsup} f^\nu(x^\nu) \leq f(x)$. Thus, $\operatorname{limsup}(\inf f^\nu) \leq \operatorname{limsup} f^\nu(x^\nu) \leq f(x) \leq \inf f + \varepsilon$. Second, suppose that $\inf f = -\infty$ and let $\delta > 0$. Then, there exists $x \in C$ such that $f(x) \leq -\delta$ and also, by condition (b) of Definition 3.7, $x^\nu \in C^\nu \to x$ such that $\operatorname{limsup} f^\nu(x^\nu) \leq f(x)$. Thus, $\operatorname{limsup}(\inf f^\nu) \leq \operatorname{limsup} f^\nu(x^\nu) \leq f(x) \leq -\delta$. Since ε and δ are arbitrary, the first result of part (a) is established. For the second result, suppose that $\bar{x} \in \operatorname{OutLim}(\varepsilon^\nu\text{-}\operatorname{argmin} f^\nu)$. Then, there exist $N \in \mathcal{N}_\infty^\#$ and $x^\nu \in \varepsilon^\nu\text{-}\operatorname{argmin} f^\nu \xrightarrow{N} \bar{x}$. Thus,

$$\operatorname*{limsup}_{\nu\in N} f^\nu(x^\nu) \leq \operatorname*{limsup}_{\nu\in N}(\inf f^\nu + \varepsilon^\nu) \leq \inf f < \infty.$$

In view of condition (a) of Definition 3.7, this implies that $\bar{x} \in C$ and also

$$f(\bar{x}) \leq \operatorname*{liminf}_{\nu\in N} f^\nu(x^\nu) \leq \operatorname*{limsup}_{\nu\in N} f^\nu(x^\nu) \leq \inf f.$$

Hence, $\bar{x} \in \operatorname{argmin} f$ and the proof of part (a) is complete.

For part (b), let $\bar{x}$ be the limit of $\{x^\nu \in \operatorname{argmin} f^\nu, \nu \in N\}$. In view of part (a), $\bar{x} \in \operatorname{argmin} f$ and also $\operatorname{limsup}(\inf f^\nu) \leq \inf f$. Condition (a) of Definition 3.7 implies that $\operatorname{liminf}_{\nu\in N}(\inf f^\nu) = \operatorname{liminf}_{\nu\in N} f^\nu(x^\nu) \geq f(\bar{x}) = \inf f$ and the conclusion holds.

Part (c), necessity. Suppose that $\inf f^\nu \to \inf f > -\infty$ and let $\varepsilon > 0$. Then, there exist $\nu_1 \in \mathbb{N}$ such that $\inf f \leq \inf f^\nu + \varepsilon/3$ for all $\nu \geq \nu_1$ and also $\bar{x} \in C$ such that $f(\bar{x}) \leq \inf f + \varepsilon/3$. In view of condition (b) of Definition 3.7, there exist $x^\nu \in C^\nu \to \bar{x}$ and $\nu_2 \geq \nu_1$ such that $f^\nu(x^\nu) \leq f(\bar{x}) + \varepsilon/3$ for all $\nu \geq \nu_2$. Let $B \subset X$ be a compact set containing $\{x^\nu, \nu \in \mathbb{N}\}$. Thus, for $\nu \geq \nu_2$,

$$\inf_{x\in C^\nu\cap B} f^\nu(x) \leq f^\nu(x^\nu) \leq f(\bar{x}) + \varepsilon/3 \leq \inf f + 2\varepsilon/3 \leq \inf f^\nu + \varepsilon.$$

Part (c), sufficiency. For the sake of a contradiction let $\inf f = -\infty$. Then, (a) implies that $\inf f^\nu \to -\infty$. Since the epi-convergence is tight, there exists a compact set $B \subset X$ such that $\inf_{x\in C^\nu\cap B} f^\nu(x) \to -\infty$ and therefore also a sequence $\{x^\nu \in C^\nu \cap B, \nu \in \mathbb{N}\}$ such that $f^\nu(x^\nu) \to -\infty$. The compactness of B implies that for some $N \in \mathcal{N}_\infty^\#$ and $x \in B$, $\lim_{\nu\in N} x^\nu = x$. In view of condition (a) of Definition 3.7, $\operatorname{liminf}_{\nu\in N} f^\nu(x^\nu) \geq f(x) \in \mathbb{R}$ if $x \in C$ and $f^\nu(x^\nu) \xrightarrow{N} \infty$ otherwise. However, both cases contradict $f^\nu(x^\nu) \to -\infty$ and thus $\inf f > -\infty$. We next show that $\operatorname{liminf}(\inf f^\nu) \geq \inf f$, which, together with (a) completes the proof of (c). We start by showing that for any compact set $B \subset X$, $\{\inf_{x\in C^\nu\cap B} f^\nu(x), \nu \in \mathbb{N}\}$, except possibly for a finite number of indexes, is bounded away from $-\infty$. For the sake of a contradiction, suppose that for $N \in \mathcal{N}_\infty^\#$ we have $\inf_{x\in C^\nu\cap B} f^\nu(x) < \inf f - 1$ for all $\nu \in N$. Then, there exists $\{x^\nu \in C^\nu \cap B, \nu \in N\}$ such that $f^\nu(x^\nu) < \inf f - 1$ for all $\nu \in N$. Since B is compact, there is a cluster point $\bar{x} \in B$ of $\{x^\nu, \nu \in N\}$. By condition (a) of Definition 3.7 and the boundedness of $\{f^\nu(x^\nu), \nu \in N\}$, $\bar{x} \in C$. Then, the same condition gives that $\liminf_{\nu\in N} f^\nu(x^\nu) \geq f(\bar{x}) \geq \inf f$, which is a contradiction. Hence, $\{\inf_{x\in C^\nu\cap B} f^\nu(x), \nu \in \mathbb{N}\}$, except possibly for a finite number of indexes, is bounded away from $-\infty$. Next, for $\varepsilon > 0$, let the compact set $B_\varepsilon \subset X$ and ν_ε be such that $\inf_{x\in C^\nu\cap B_\varepsilon} f^\nu(x) \leq \inf f^\nu + \varepsilon$ for all $\nu \geq \nu_\varepsilon$, which holds by the tightness assumption. Since $\{\inf_{x\in C^\nu\cap B} f^\nu(x), \nu \in \mathbb{N}\}$ is eventually bounded away from $-\infty$, there exist $\bar{\nu}_\varepsilon \geq \nu_\varepsilon$ and $x^\nu \in C^\nu \cap B_\varepsilon$ such that $f^\nu(x^\nu) \leq \inf_{x\in C^\nu\cap B_\varepsilon} f^\nu(x) + \varepsilon$ for all $\nu \geq \bar{\nu}_\varepsilon$. Thus, for $\nu \geq \bar{\nu}_\varepsilon$,

$$f^\nu(x^\nu) \leq \inf_{x\in C^\nu\cap B_\varepsilon} f^\nu(x) + \varepsilon \leq \inf f^\nu + 2\varepsilon.$$

In view of (a), we conclude that $\{f^\nu(x^\nu), \nu \in \mathbb{N}\}$ is bounded from above. The compactness of B_ε implies that there exist $N \in \mathcal{N}_\infty^\#$ and $\bar{x} \in B_\varepsilon$ such that $\lim_{\nu\in N} x^\nu = \bar{x}$. In view of condition (a) of Definition 3.7, $\bar{x} \in C$ because $\{f^\nu(x^\nu), \nu \in N\}$ is bounded from above. The same condition then implies that

$$\liminf_{\nu\in N}(\inf f^\nu) + 2\varepsilon \geq \liminf_{\nu\in N}(\inf_{x\in C^\nu\cap B_\varepsilon} f^\nu(x)) + \varepsilon \geq \liminf_{\nu\in N} f^\nu(x^\nu) \geq f(\bar{x}) \geq \inf f.$$

Since this argument holds not only for $\{f^\nu, \nu \in \mathbb{N}\}$ but also for all subsequences, $\liminf(\inf f^\nu) + 2\varepsilon \geq \inf f$. We reach the conclusion of part (c) after recognizing that ε is arbitrary.

For part (d), let $\bar{x} \in \operatorname{argmin} f$. Condition (b) of Definition 3.7 implies that there exist $x^\nu \in C^\nu \to \bar{x}$ and $\nu_1 \in \mathbb{N}$ such that $f^\nu(x^\nu) \leq f(\bar{x}) + \varepsilon/2$ for all $\nu \geq \nu_1$. Since $\inf f$ is finite, there is also $\nu_2 \geq \nu_1$ such that $\inf f \leq \inf f^\nu + \varepsilon/2$ for all $\nu \geq \nu_2$. Thus,

$$f^\nu(x^\nu) \leq f(\bar{x}) + \varepsilon/2 = \inf f + \varepsilon/2 \leq \inf f^\nu + \varepsilon$$

for all $\nu \geq \nu_2$ and we conclude that $\bar{x} \in \operatorname{InnLim}(\varepsilon\text{-}\operatorname{argmin} f^\nu)$.

Part (e). From (a), $\operatorname{OutLim}(\varepsilon^\nu\text{-}\operatorname{argmin} f^\nu) \subset \operatorname{argmin} f$ for any $\varepsilon^\nu \downarrow 0$. Thus it suffices to show that $\operatorname{InnLim}(\varepsilon^\nu\text{-}\operatorname{argmin} f^\nu) \supset \operatorname{argmin} f$ for some $\varepsilon^\nu \downarrow 0$. If $\operatorname{argmin} f = \emptyset$, then the inclusion is automatic. Thus, we can assume that $\operatorname{argmin} f \neq \emptyset$ and $\inf f$ is finite. Theorem 3.1 in [10] states the following result: For any collection $\{h, h^\nu, \nu \in \mathbb{N}\}$ of extended real-valued lsc functions on a separable metric space, if $h^\nu \xrightarrow{e} h$ and $\alpha \in \mathbb{R}$, then there exists $\alpha^\nu \to \alpha$ such that $\{x : h^\nu(x) \leq \alpha^\nu\} \to \{x : h(x) \leq \alpha\}$. To apply this theorem here, let lsc f^ν be the lsc-regularization of f^ν on X, i.e., the highest lsc function on X not exceeding f^ν extended to the whole of X by assigning it the value ∞ outside its domain. Since lsc $f^\nu \xrightarrow{e} f$, which already is lsc and can be extended in the same manner, it follows that there exists $\alpha^\nu \to \inf f$ such that

$$\{x \in X \;:\; (\operatorname{lsc} f^\nu)(x) \leq \alpha^\nu\} \to \{x \in X \;:\; f(x) \leq \inf f\} = \operatorname{argmin} f.$$

Set $\{\varepsilon^\nu = 1/\nu + \max[0, \alpha^\nu - \inf f^\nu], \nu \in \mathbb{N}\}$ and $\bar{x} \in \operatorname{argmin} f$. In view of the above set convergence, there exists a sequence $\{x^\nu, \nu \in \mathbb{N}\}$, with $(\operatorname{lsc} f^\nu)(x^\nu) \leq \alpha^\nu$, converging to $\bar{x}$. The definition of lsc-regularization implies that there exists $\{y^\nu \in C^\nu, \nu \in \mathbb{N}\}$ such that $d_X(x^\nu, y^\nu) \le 1/\nu$ and $f^\nu(y^\nu) \le (\text{lsc}\, f^\nu)(x^\nu) + 1/\nu$. Thus,

$$f^\nu(y^\nu) \le (\text{lsc}\, f^\nu)(x^\nu) + 1/\nu \le \alpha^\nu + 1/\nu \le \varepsilon^\nu + \inf f^\nu$$

and therefore $y^\nu \in \varepsilon^\nu\text{-}\operatorname{argmin} f^\nu$. Consequently, $\bar{x} \in \operatorname{InnLim}(\varepsilon^\nu\text{-}\operatorname{argmin} f^\nu)$, which implies that $\operatorname{InnLim}(\varepsilon^\nu\text{-}\operatorname{argmin} f^\nu) \supset \operatorname{argmin} f$. Since $\inf f^\nu \to \inf f$, $\varepsilon^\nu \downarrow 0$ and the conclusion holds.

For part (f), let $\bar{x} \in \operatorname{argmin} f$. The assumption permits the selection of $\bar{\nu} \in \mathbb{N}$ and a sequence $\{x^\nu \in \varepsilon^\nu\text{-}\operatorname{argmin} f^\nu, \nu \ge \bar{\nu}\} \to \bar{x}$. Thus, by condition (a) of Definition 3.7,

$$\inf f = f(\bar{x}) \le \operatorname{liminf} f^\nu(x^\nu) \le \operatorname{liminf}(\inf f^\nu + \varepsilon^\nu) = \operatorname{liminf}(\inf f^\nu).$$

In view of part (a), the conclusion follows. □

4 Consequences of lopsided convergence

The main consequence of lop-convergence of bifunctions F^ν to F is the convergence of solutions of the

$$\text{approximating minsup problems} \quad \min_{x \in C^\nu} \sup_{y \in D^\nu(x)} F^\nu(x, y)$$

to those of an

$$\text{actual minsup problem} \quad \min_{x \in C} \sup_{y \in D(x)} F(x, y),$$

where (C, D) and (C^ν, D^ν) describe $\operatorname{dom} F$ and $\operatorname{dom} F^\nu$, respectively. We give detailed results below, but start with fundamental properties associated with lopsided convergence.

4.1 Basic properties

For "degenerate" bifunctions that only depend on their first argument, lopsided convergence is equivalent to epi-convergence, a direct consequence of the definitions, as stated next.

Proposition 4.1 (Lop-convergence subsumes epi-convergence) *Suppose that the bifunctions* $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \operatorname{bfcns}(X, Y)$ *has* $\operatorname{dom} F = C \times Y$, $\operatorname{dom} F^\nu = C^\nu \times Y$, $F(x, y) = F(x, y')$ *for all* $x \in C$ *and* $y, y' \in Y$, *and* $F^\nu(x, y) = F^\nu(x, y')$ *for all* $x \in C^\nu$ *and* $y, y' \in Y$. *Then, for any* $y \in Y$,

$$F^\nu \text{ lop-converges to } F \iff F^\nu(\cdot, y) : C^\nu \to \mathbb{R} \text{ epi-converges to } F(\cdot, y) : C \to \mathbb{R}.$$

Lop-convergence implies hypo-convergence for certain functions as seen next.

Proposition 4.2 (Hypo-convergence of slices) *For* $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \text{bfcns}(X, Y)$ *with domains described by* (C, D) *and* (C^ν, D^ν)*, respectively, we have that* $F^\nu \xrightarrow{lop} F$ *implies that for all* $x \in C$*, there exists* $x^\nu \in C^\nu \to x$ *such that the functions* $F^\nu(x^\nu, \cdot) : D^\nu(x^\nu) \to \mathbb{R}$ *hypo-converge to* $F(x, \cdot) : D(x) \to \mathbb{R}$.

Proof From condition (b) of Definition 3.1 there exists $x^\nu \in C^\nu \to x$ such that the functions $\{-F^\nu(x^\nu, \cdot) : D^\nu(x^\nu) \to \mathbb{R}, \nu \in \mathbb{N}\}$ and $-F(x, \cdot) : D(x) \to \mathbb{R}$ satisfy condition (a) of Definition 3.7. From condition (a) of Definition 3.1, for any $y \in D(x)$ one can find $y^\nu \in D^\nu(x^\nu) \to y$ such that condition (b) of Definition 3.7 holds for $\{-F^\nu(x^\nu, \cdot) : D^\nu(x^\nu) \to \mathbb{R}, \nu \in \mathbb{N}\}$ and $-F(x, \cdot) : D(x) \to \mathbb{R}$. Thus, the former functions epi-converge to the latter. □

The *sup-projection*[5] (in y) of $F \in \text{bfcns}(X, Y)$ with domain described by (C, D) exists if $\sup_{y \in D(x)} F(x, y) < \infty$ for some $x \in C$. When it exists, the sup-projection of F is the real-valued function f given by

$$f(x) = \sup_{y \in D(x)} F(x, y) \text{ whenever } x \in C \text{ and } \sup_{y \in D(x)} F(x, y) < \infty.$$

This means that the domain of the sup-projection is a subset of C, which could be strict. Bifunctions without sup-projections are somewhat pathological and in the case of minsup problems correspond to infeasibility. Since $D(x) \neq \emptyset$ for $x \in C$, $\sup_{y \in D(x)} F(x, y) > -\infty$ and, thus, sup-projections are indeed real-valued. We denote by f and f^ν sup-projections of F and F^ν, respectively.

It is clear that the sup-projection of F^ν with domain defined by (C^ν, D^ν) might not exist, i.e., $\sup_{y \in D^\nu(x)} F^\nu(x, y) = \infty$ for all $x \in C^\nu$, even if that of F does and $F^\nu \xrightarrow{lop} F$ as the following example shows.

Absence of sup-projection. Consider the bifunctions $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \text{bfcns}(\mathbb{R}, \mathbb{R})$ with domains described by $C = C^\nu = [0, 1]$, $D = D^\nu = \{y \in \mathbb{R} : y \geq 0\}$, and values $F(x, y) = yx$ and $F^\nu(x, y) = y(x + 1/\nu)$ when defined. Clearly, the sup-projection of F has domain $\{0\}$. However, there is no ν for which F^ν has a sup-projection. This situation takes place even though one can show that $F^\nu \xrightarrow{lop} F$.

A consequence of lopsided convergence for the epigraphs of sup-projections is given next.

Theorem 4.3 (Containment of sup-projections) *For bifunctions* $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \text{bfcns}(X, Y)$*, with corresponding sup-projections* $\{f, f^\nu, \nu \in \mathbb{N}\}$*,* $F^\nu \xrightarrow{lop} F$ *implies that*

$$\text{OutLim}\left(\text{epi } f^\nu\right) \subset \text{epi } f.$$

Proof Let (C, D) and (C^ν, D^ν) describe the domains of F and F^ν, respectively. Suppose that $(x, \alpha) \in \text{OutLim}(\text{epi } f^\nu)$. Then there exist $N \in \mathcal{N}_\infty^\#$ and $\{(x^\nu, \alpha^\nu), \nu \in N\}$, with $x^\nu \in C^\nu$, $\sup_{y \in D^\nu(x^\nu)} F^\nu(x^\nu, y) \leq \alpha^\nu$, $x^\nu \xrightarrow{N} x$, and $\alpha^\nu \xrightarrow{N} \alpha$. If $x \notin C$, then

[5] In parametric optimization and elsewhere sup-projections of bifunctions are sometimes called optimal value functions.

we can construct by condition (a) of Definition 3.1 a sequence $y^\nu \in D^\nu(x^\nu)$ such that $F^\nu(x^\nu, y^\nu) \xrightarrow{N} \infty$. However,

$$\alpha^\nu \geq \sup_{y \in D^\nu(x^\nu)} F^\nu(x^\nu, y) \geq F^\nu(x^\nu, y^\nu),\ \nu \in N,$$

implies a contradiction since $\alpha^\nu \xrightarrow{N} \alpha \in \mathbb{R}$. Thus, $x \in C$. If $\sup_{y \in D(x)} F(x, y) = \infty$, then there exists $y \in D(x)$ such that $F(x, y) \geq \alpha + 1$. Condition (a) of Definition 3.1 ensures that there exists a sequence $y^\nu \in D^\nu(x^\nu) \xrightarrow{N} y$ such that $\operatorname{liminf}_{\nu \in N} F^\nu(x^\nu, y^\nu) \geq F(x, y)$. Consequently,

$$\alpha = \operatorname*{liminf}_{\nu \in N} \alpha^\nu \geq \operatorname*{liminf}_{\nu \in N} \sup_{y \in D^\nu(x^\nu)} F^\nu(x^\nu, y) \geq \operatorname*{liminf}_{\nu \in N} F^\nu(x^\nu, y^\nu) \geq F(x, y) \geq \alpha + 1,$$

which is a contradiction. Hence, it suffices to consider the case with $\sup_{y \in D(x)} F(x, y)$ finite. Given any $\varepsilon > 0$ arbitrarily small, pick $y_\varepsilon \in D(x)$ such that $F(x, y_\varepsilon) \geq \sup_{y \in D(x)} F(x, y) - \varepsilon$. Then condition (a) of Definition 3.1 again yields $y^\nu \in D^\nu(x^\nu) \xrightarrow{N} y_\varepsilon$ such that

$$\operatorname*{liminf}_{\nu \in N} \sup_{y \in D^\nu(x^\nu)} F^\nu(x^\nu, y) \geq \operatorname*{liminf}_{\nu \in N} F^\nu(x^\nu, y^\nu) \geq F(x, y_\varepsilon) \geq \sup_{y \in D(x)} F(x, y) - \varepsilon,$$

implying $\operatorname{liminf}_{\nu \in N} \sup_{y \in D^\nu(x^\nu)} F^\nu(x^\nu, y) \geq \sup_{y \in D(x)} F(x, y)$. Since

$$\alpha = \operatorname*{liminf}_{\nu \in N} \alpha^\nu \geq \operatorname*{liminf}_{\nu \in N} \sup_{y \in D^\nu(x^\nu)} F^\nu(x^\nu, y) \geq \sup_{y \in D(x)} F(x, y),$$

the conclusion follows. □

It is clear from the following example that lop-convergence of bifunctions does not guarantee epi-convergence of the corresponding sup-projections.

Absence of tightness. For $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \text{bfcns}(\mathbb{R}, \mathbb{R})$ with domains $\mathbb{R} \times \mathbb{R}$ and values $F(x, y) = 0$ for $(x, y) \in \mathbb{R} \times \mathbb{R}$, and $F^\nu(x, y) = 1$ if $x \in \mathbb{R}$ and $y = \nu$, and zero otherwise. It is easy to show that $F^\nu \xrightarrow{lop} F$. The sup-projection of F^ν has value 1 for all $x \in \mathbb{R}$ and that of F has value 0 for all $x \in \mathbb{R}$. Thus, the inclusion in Theorem 4.3 is strict. We observe that ancillary-tightness fails in this instance and that property is indeed key to eliminating such possibilities. When Y is a space of distribution functions, these questions translates into those about tightness in the sense of probability theory; see the remark after Definition 3.5 and [29].

Theorem 4.4 (Epi-convergence of sup-projections) *Suppose the bifunctions* $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \text{bfcns}(X, Y)$ *have corresponding sup-projections* $\{f, f^\nu, \nu \in \mathbb{N}\}$. *If* F^ν *lop-converges ancillary-tightly to* F, *then*

$$\text{epi}\, f^\nu \to \text{epi}\, f, \ \textit{also written}\ f^\nu \xrightarrow{e} f.$$

Proof Let (C, D) and (C^ν, D^ν) describe the domains of F and F^ν, respectively, and $x \in \text{dom}\, f$. Now, choose $x^\nu \in C^\nu \to x$ such that $F^\nu(x^\nu, \cdot) : D^\nu(x^\nu) \to \mathbb{R}$

hypo-converge to $F(x,\cdot) : D(x) \to \mathbb{R}$, cf. Proposition 4.2. In fact, these functions hypo-converge tightly as an immediate consequence of ancillary-tightness. Thus,

$$\sup_{y\in D^\nu(x^\nu)} F^\nu(x^\nu, y^\nu) \to \sup_{y\in D(x)} F(x, y),$$

via Theorem 3.8. This fact together with Theorem 4.3 establish the conclusion. □

It is clear that ancillary-tight lop-convergence is a stronger condition than epi-convergence of the corresponding sup-projections, i.e., the converse of Theorem 4.4 fails. For example, the bifunctions $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \text{bfcns}(\mathbb{R}, \mathbb{R})$ with domains described by $C = C^\nu = [0, 1]$ and $D = D^\nu = [0, 1]$, and values $F(x, y) = 1$, $F^\nu(x, y) = y$ for all $(x, y) \in C \times D$ have sup-projections that are identical and, certainly, $f^\nu \xrightarrow{e} f$. However, condition (a) of Definition 3.1 fails and therefore F^ν does not lop-converge to F.

We end this subsection by listing consequences for semicontinuity, concavity, and convexity.

Proposition 4.5 (usc and concavity) *For bifunctions* $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \text{bfcns}(X, Y)$ *with domains described by* (C, D) *and* (C^ν, D^ν)*, respectively, and any* $x \in C$*, the lop-convergence* $F^\nu \xrightarrow{lop} F$ *implies that the (univariate) function* $F(x,\cdot) : D(x) \to \mathbb{R}$ *is usc.*

When Y *is a linear space and for all* $x^\nu \in C^\nu \to x$, $F^\nu(x^\nu,\cdot) : D^\nu(x^\nu) \to \mathbb{R}$ *is concave, it also implies that* $F(x,\cdot)$ *is concave.*

Proof In view of Proposition 4.2, for every $x \in C$, there exists $x^\nu \in C^\nu \to x$ such that $F^\nu(x^\nu,\cdot)$ hypo-converges to $F(x,\cdot)$, which therefore must be usc. If $F^\nu(x^\nu,\cdot)$ are concave, its hypo-limit must also be concave, which establishes the conclusions. □

Proposition 4.6 (lsc and convexity) *Suppose the bifunctions* $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \text{bfcns}(X, Y)$ *have domains described by* (C, D) *and* (C^ν, D^ν)*, respectively, and corresponding sup-projections* $\{f, f^\nu, \nu \in \mathbb{N}\}$*. If* F^ν *lop-converges ancillary-tightly to* F*, then* f *is lsc and also convex provided that* X *is a linear space,* D^ν *is constant on* C^ν*, and* $F^\nu(\cdot, y)$ *is convex for all* $y \in D^\nu$*.*

Proof In view of Theorem 4.4, f is an epi-limit and thus lsc. Convexity is guaranteed if $f^\nu = \sup_{y\in D^\nu(\cdot)} F^\nu(\cdot, y)$, $\nu \in \mathbb{N}$, are convex, which holds by the stated assumptions. □

4.2 Consequence for minsup problems

We have now reached the presentation of the main results of the paper. The *minsup-value* of a bifunction $F \in \text{bfcns}(X, Y)$ is defined as

$$\text{infsup}\, F := \inf_{x\in C} \sup_{y\in D(x)} F(x, y),$$

which clearly is the optimal value of the minsup problem $\min_{x\in C} \sup_{y\in D(x)} F(x, y)$. The corresponding ε-optimal solutions, referred to as ε*-minsup-points* of F, are given by

$$\varepsilon\text{-}\operatorname{argminsup} F := \left\{ x \in C \ : \ \sup_{y \in D(x)} F(x, y) \leq \operatorname{infsup} F + \varepsilon \right\}, \quad \text{for } \varepsilon \geq 0.$$

If $\varepsilon = 0$, we simply refer to such points as *minsup-points*.

Theorem 4.7 (Bounds on minsup-value) *Suppose that* $\{F, F^\nu, \ \nu \in \mathbb{N}\} \subset \operatorname{bfcns}(X, Y)$ *have sup-projections,* $F^\nu \xrightarrow{lop} F$, *and* $\{x^\nu \in \operatorname{argminsup} F^\nu, \nu \in \mathbb{N}\}$ *exist. Then, the following hold:*

(a) *If* $\{x^\nu, \nu \in N\}$ *converges for some* $N \in \mathcal{N}_\infty^\#$, *then* $\liminf_{\nu \in N}(\operatorname{infsup} F^\nu) \geq \operatorname{infsup} F$.
(b) *If lop-convergence is ancillary-tight, then* $\limsup(\operatorname{infsup} F^\nu) \leq \operatorname{infsup} F$ *and*

$$\forall \{\varepsilon^\nu \downarrow 0, \nu \in \mathbb{N}\}, \operatorname{OutLim}(\varepsilon^\nu\text{-}\operatorname{argminsup} F^\nu) \subset \operatorname{argminsup} F.$$

(c) *If lop-convergence is ancillary-tight and* $\{x^\nu, \nu \in N\}$ *converges for some* $N \in \mathcal{N}_\infty^\#$, *then* $\lim_{\nu \in N} (\operatorname{infsup} F^\nu) = \operatorname{infsup} F$.

Proof Let f and f^ν be the sup-projections of F and F^ν, respectively. We first consider (a): Let $N \in \mathcal{N}_\infty^\#$ and $\bar{x}$ be such that $x^\nu \xrightarrow{N} \bar{x}$. Theorem 4.3 applies so that $\operatorname{OutLim}(\operatorname{epi} f^\nu) \subset \operatorname{epi} f$, which in turn ensures that condition (a) of Definition 3.7 holds. Thus, if $\bar{x} \in \operatorname{dom} f$, then $\liminf_{\nu \in N}(\operatorname{infsup} F^\nu) = \liminf_{\nu \in N} f^\nu(x^\nu) \geq f(\bar{x}) \geq \operatorname{infsup} F$. If $\bar{x} \notin \operatorname{dom} f$, then $\liminf_{\nu \in N}(\operatorname{infsup} F^\nu) = \liminf_{\nu \in N} f^\nu(x^\nu) \to \infty$ and the conclusion holds. Second, consider (b) and (c): Theorem 4.4 implies that $f^\nu \xrightarrow{e} f$ and thus the conclusion is a direct application of Theorem 3.8. □

Tight lop-convergence implies the following strengthening of the result for minsup-points and values.

Theorem 4.8 (Approximating minsup-points) *Suppose that* $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \operatorname{bfcns}(X, Y)$ *have sup-projections and* $F^\nu \xrightarrow{lop} F$ *tightly. Then,*

(a) $\operatorname{infsup} F^\nu \to \operatorname{infsup} F$, *which is finite;*
(b) *for* $\varepsilon > 0$, $\operatorname{InnLim}(\varepsilon\text{-}\operatorname{argminsup} F^\nu) \supset \operatorname{argminsup} F$;
(c) *for* X *separable,* $\exists \{\varepsilon^\nu \downarrow 0, \nu \in \mathbb{N}\}$ *such that* $\varepsilon^\nu\text{-}\operatorname{argminsup} F^\nu \to \operatorname{argminsup} F$.

Proof The assumptions imply those of Theorem 4.4 and thus the sup-projection of F^ν epi-converges to that of F. In view of Definition 3.6, the epi-convergence is tight; see Definition 3.7. The conclusion is then a direct application of Theorem 3.8. □

It is well-known that the infimum of a lsc function with a domain contained in a compact set is attained. Consequently, if the sup-projection f of some bifunction $F \in \operatorname{bfcns}(X, Y)$ is lsc, which holds under mild assumptions (cf. Proposition 5.1), and $\operatorname{dom} f$ is contained in some compact set, then there exists a minimizer of f and thus also a minsup-point of F. We next state a result that relaxes the compactness requirement.

Theorem 4.9 (Existence of minsup-points) *Suppose that* $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \operatorname{bfcns}(X, Y)$ *have sup-projections, those of* F^ν, *denoted by* f^ν, *are lsc,* $F^\nu \xrightarrow{lop} F$

ancillary-tightly, and there are compact sets $\{B^\nu \subset X, \nu \in \mathbb{N}\}$ *such that* $\operatorname{dom} f^\nu \subset B^\nu$. *Then,* $\{x^\nu \in \operatorname{argminsup} F^\nu, \nu \in \mathbb{N}\}$ *exist and every cluster point of* $\{x^\nu, \nu \in \mathbb{N}\}$ *is a minsup-point of* F.

Proof The discussion prior to the theorem ensures the existence of minsup-points of F^ν for every ν. The result is then a consequence of Theorem 4.7. □

Theorem 4.9 does not ensure the existence of a cluster point of $\{x^\nu, \nu \in \mathbb{N}\}$. Still, it provides an approach for establishing the existence of a minsup-point of F: first construct a sequence $\{F^\nu, \nu \in \mathbb{N}\}$, with the required properties, that lop-converges ancillary-tightly to F and, second, prove that $\{x^\nu, \nu \in \mathbb{N}\}$ has a cluster point.

4.3 Application to generalized Nash games

From Sect. 2.2, we know that the solutions of a Generalized Nash game are fully characterized by the associated Nikaido-Isoda bifunction; see Proposition 2.1. Here, we illustrate the application of lopsided convergence in this context and in the process obtain stability results for Generalized Nash games in metric spaces with agent-constraints that depend on the other agents' choices of strategies.

We consider the

$$\text{actual game} \quad \bar{x}_a \in \operatorname*{argmin}_{x_a \in D_a(\bar{x}_{-a})} c_a(x_a, \bar{x}_{-a}), \quad \text{for all } a \in A,$$

where the notation follows that of Sect. 2.2, and the

$$\text{approximating game} \quad \bar{x}_a \in \operatorname*{argmin}_{x_a \in D_a^\nu(\bar{x}_{-a})} c_a^\nu(x_a, \bar{x}_{-a}), \quad \text{for all } a \in A,$$

where $c_a^\nu : X_a \times X_{-a} \to \mathbb{R}$ is the approximating cost function and $D_a^\nu : X_{-a} \rightrightarrows X_a$ is the approximating set-valued mapping that defines the constraints for agent $a \in A$. The Nikaido–Isoda bifunction associated with the actual game is defined in Sect. 2.2. Similarly, the one for the approximating game is

$$F^\nu(x, y) = \sum_{a \in A} \Big[c_a^\nu(x_a, x_{-a}) - c_a^\nu(y_a, x_{-a})\Big] \quad \text{for } x \in C^\nu \subset X, \quad y \in D^\nu(x) \subset X$$

with

$$C^\nu = \Big\{x \in X \ : \ x_a \in D_a^\nu(x_{-a}) \text{ for all } a \in A\Big\} \text{ and } D^\nu(x) = \prod_{a \in A} D_a^\nu(x_{-a}).$$

The set-valued mapping $D^\nu : C^\nu \rightrightarrows X$. We assume throughout this subsection that there exist $x, x^\nu \in X$ such that $x_a \in D_a(x_{-a})$ and $x_a^\nu \in D_a^\nu(x_{-a}^\nu)$ for all $a \in A$ and $\nu \in \mathbb{N}$, i.e., there are feasible solutions. Thus, $F, F^\nu \in \operatorname{bfcns}(X, X)$. The approximating Nikaido–Isoda bifunction lop-converges to the actual bifunction under natural assumptions as seen next. This result and the associated propositions extend in many ways Corollary 7.5 of [21], which assumes convex and compact agent-constraint sets that are independent of other agents' strategies as well as other assumptions, and also [18], which also adopts the independence assumption and, in addition, constraint sets that are not approximated.

Theorem 4.10 (Lop-convergence of Nikaido–Isoda bifunctions) *For Nikaido–Isoda bifunctions* $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \mathrm{bfcns}(X, X)$ *defined in terms of cost functions* $\{c_a, c_a^\nu, a \in A, \nu \in \mathbb{N}\}$ *and constraint mappings* $\{D_a, D_a^\nu, a \in A, \nu \in \mathbb{N}\}$*, we have* $F^\nu \xrightarrow{lop} F$ *provided that*

(a) *the mappings* D_a^ν *continuously converge to* D_a*, i.e.,* $\forall x^\nu \in X \to x \in X$, $D_a^\nu(x_{-a}^\nu) \to D_a(x_{-a})\ \forall a \in A$*;*
(b) *the cost functions* c_a^ν *continuously converge to* c_a*, i.e.,* $\forall x^\nu \in X \to x \in X$, $c_a^\nu(x_a^\nu, x_{-a}^\nu) \to c_a(x_a, x_{-a})\ \forall a \in A$*; and*
(c) *either* $D_a^\nu(x_{-a}) \supset D_a(x_{-a})\ \ \forall x_{-a} \in X_{-a}$*,* $a \in A$*, and* $\nu \in \mathbb{N}$ *or the mappings* D_a^ν *and* D_a *are constant for all* $\nu \in \mathbb{N}$ *and* $a \in A$*.*

Proof We obtain lop-convergence directly from Definition 3.1. Let C, C^ν, D, D^ν be as defined in this subsection and Sect. 2.2. Suppose that $x^\nu \in C^\nu \to x \in C$ and $y \in D(x)$. Since $D_a^\nu(x_{-a}^\nu) \to D_a(x_{-a}^\nu)$ for all $a \in A$, $D^\nu(x^\nu) \to D(x)$. This set-convergence guarantees that there exists $y^\nu \in D^\nu(x^\nu) \to y$. The continuous convergence of c_a^ν to c_a carries over to the bifunctions and consequently $F^\nu(x^\nu, y^\nu) \to F(x, y)$. The first part of Definition 3.1(a) is satisfied. Next, since $C^\nu = \cap_{a\in A}\{x \in X : x_a \in D_a^\nu(x_{-a})\}$, it follows by assumption (a) that

$$\begin{aligned}\mathrm{OutLim}\, C^\nu &\subset \bigcap_{a\in A} \mathrm{OutLim}\{x \in X : x_a \in D_a^\nu(x_{-a})\}\\ &\subset \bigcap_{a\in A}\{x \in X : x_a \in D_a(x_{-a})\} = C.\end{aligned}$$

Thus, every convergent sequence $x^\nu \in C^\nu$ must have a limit in C and the second part of Definition 3.1(a) is automatically satisfied. We then turn to Definition 3.1(b) and let $x \in C$. First, suppose that $D_a^\nu(x_{-a}) \supset D_a(x_{-a})\ \ \forall x_{-a} \in X_{-a}$, $a \in A$, and $\nu \in \mathbb{N}$; see assumption (c). Then, set $x^\nu = x$, which will be in C^ν because $C^\nu \supset C$. Suppose that $y^\nu \in D^\nu(x) \to y \in D(x)$. Again, assumption (b) ensures that $F^\nu(x^\nu, y^\nu) \to F(x, y)$. This concludes the proof for the first assumption in (c); the possibility $y^\nu \in D^\nu(x) \to y \notin D(x)$ can be ruled out by assumption (a). Second, suppose that the mappings D_a^ν and D_a are constant for all $\nu \in \mathbb{N}$ and $a \in A$; these constant sets are also denoted by D_a^ν and D_a. Since in this case $C = \prod_{a\in A} D_a$ and $C^\nu = \prod_{a\in A} D_a^\nu$, $C^\nu \to C$ and there exists $x^\nu \in C^\nu \to x$. Suppose that $y^\nu \in D^\nu(x^\nu) \to y \in D(x)$. Again, assumption (b) ensures that $F^\nu(x^\nu, y^\nu) \to F(x, y)$. This concludes the proof as the possibility $y^\nu \in D^\nu(x) \to y \notin D(x)$ can again be ruled out. □

It is possible to relax the assumptions of Theorem 4.10. For example in (a), the continuous convergence may be only relative to smaller sets such as for $x^\nu \in C^\nu \to x \in C$. However, since C^ν and C are not easily characterized from D_a^ν and D_a, we settled on the stated, more easily checked, assumptions. The continuous convergence of the cost functions (assumption (b)) is usually unavoidable (see the discussion in [21, Section 5] for the context of variational inequalities), but fortunately it is often satisfied in the applications we have in mind.

Although other options exist, it is generally impossible to completely eliminate some "constraint qualification" of the kind stated in assumption (c) of Theorem 4.10. The source of the difficulty is that assumption (a) fails to guarantee that $C^\nu \to C$. Although $\mathrm{OutLim}\, C^\nu \subset C$, as leveraged in the above proof, the inclusion may be strict.

Absence of constraint qualification. Suppose that $A = \{1, 2\}$, $X = \mathbb{R} \times \mathbb{R}$, and for all $x \in X$, $a \in A$, and $\nu \in \mathbb{N}$,

$$D_a(x_{-a}) = \{0\} \cup \{x_{-a}\} \text{ and } D_a^\nu(x_{-a}) = \{0\} \cup \{x_{-a} + 1/\nu\}.$$

In this case, $C = \{x \in \mathbb{R} \times \mathbb{R} \ : \ x_1 = x_2\}$ and $C^\nu = \{(0, 0), (0, 1/\nu), (1/\nu, 0)\}$ for all ν. Clearly, $\mathrm{OutLim}\, C^\nu = \{(0, 0)\}$ is a strict subset of C. This takes place even though assumption (a) of Theorem 4.10 holds. However, as indicated in the second alternative in assumption (c) of Theorem 4.10, this difficulty evaporates when D^ν and D are constant set-valued mappings, i.e., an agent's constraint set is independent of the other agents' strategies.

Theorem 4.10 provides a main step in sensitivity analysis and study of approximations and issues related to existence of solutions for Generalized Nash games. Leveraging Theorems 4.7, 4.8, and 4.9, it is easy to develop a variety of results, possibly bringing in ancillary-tightness and tightness. We give two results that also permit arguments directly from the definition of lopsided convergence.

Proposition 4.11 (Convergence of equilibria) *For Nikaido–Isoda bifunctions* $\{F, F^\nu, \nu \in \mathbb{N}\} \subset \mathrm{bfcns}(X, X)$ *associated with the actual and approximating games, respectively, suppose that* $F^\nu \xrightarrow{lop} F$ *and* $\{\bar{x}^\nu, \nu \in \mathbb{N}\}$ *are equilibria of the approximating games. Then, every cluster point* $\bar{x}$ *of* $\{\bar{x}^\nu, \nu \in \mathbb{N}\}$ *is an equilibrium of the actual game.*

Proof By Proposition 2.1, $\bar{x}^\nu \in \mathrm{argminsup}\, F^\nu$ and $f^\nu(\bar{x}^\nu) \leq 0$, where f^ν is the sup-projection of F^ν. Let $N \in \mathcal{N}_\infty^\#$ be such that $\bar{x}^\nu \xrightarrow{N} \bar{x}$ and C, C^ν, D, D^ν be as defined in this subsection and Sect. 2.2. Suppose that $\bar{x} \notin C$. Then, by part (a) of Definition 3.1, there exists $y^\nu \in D^\nu(\bar{x}^\nu)$ such that $F^\nu(\bar{x}^\nu, y^\nu) \xrightarrow{N} \infty$. However, $0 \geq f^\nu(\bar{x}^\nu) \geq F^\nu(\bar{x}^\nu, y^\nu)$ for $\nu \in N$, which is a contradiction. Thus, $\bar{x} \in C$. Next, suppose that $\sup_{y \in D(\bar{x})} F(\bar{x}, y) > 0$. Then, there exists $\bar{y} \in D(\bar{x})$ such that $F(\bar{x}, \bar{y}) > 0$. Again by part (a) of Definition 3.1, there exists $y^\nu \in D^\nu(\bar{x}^\nu) \xrightarrow{N} \bar{y}$ such that $\liminf_{\nu \in N} F^\nu(\bar{x}^\nu, y^\nu) \geq F(\bar{x}, \bar{y})$. As argued above, the left-hand side is nonpositive and another contradiction is reached. Therefore, $\sup_{y \in D(\bar{x})} F(\bar{x}, y) \leq 0$. For all $x \in C$, $x \in D(x)$ and $F(x, x) = 0$. Consequently, $\sup_{y \in D(x)} F(x, y) \geq 0$, which establishes that $\bar{x} \in \mathrm{argminsup}\, F$. By Proposition 2.1, $\bar{x}$ is an equilibrium of the actual game. $\square$

The next proposition resembles a result in [22].

Proposition 4.12 (Convergence of equilibria; specifics) *Suppose that the actual and approximating games are given by cost functions* $\{c_a, c_a^\nu, a \in A, \nu \in \mathbb{N}\}$ *and constraint mappings* $\{D_a, D_a^\nu, a \in A, \nu \in \mathbb{N}\}$, *with*

(a) D_a^ν *continuously converge to* D_a, *i.e.,* $\forall x^\nu \in X \to x \in X$, $D_a^\nu(x_{-a}^\nu) \to D_a(x_{-a})$ $\forall a \in A$; *and*
(b) c_a^ν *continuously converge to* c_a, *i.e.,* $\forall x^\nu \in X \to x \in X$, $c_a^\nu(x_a^\nu, x_{-a}^\nu) \to c_a(x_a, x_{-a})$ $\forall a \in A$.

If $\{\bar{x}^\nu, \nu \in \mathbb{N}\}$ *are equilibria of the approximating games, then every cluster point* $\bar{x}$ *of* $\{\bar{x}^\nu, \nu \in \mathbb{N}\}$ *is an equilibrium of the actual game.*

Proof Examining the proof of Theorem 4.10, we find that the present assumptions suffice to establish part (a) of Definition 3.1. Since part (b) of that definition is not utilized in the proof of Proposition 4.11, the arguments there carry over. □

The existence of equilibria for the approximating games are assumed in these proposition. Although nontrivial in general, Theorem 4.9 provides a path to such existence results; see also [21] for approaches based on Ky Fan bifunctions.

For computational reasons, an agent's constraints may be removed and replaced with a penalty in the objective function. This natural idea (see for example [15–17]) also leads to lop-convergence under mild assumptions. Specifically, for some $g_a : X_a \times X_{-a} \to \mathbb{R}^m$, $a \in A$, suppose that

$$D_a(x_{-a}) = \{x_a \in X_a \ : \ g_a(x_a, x_{-a}) \le 0\} \quad \text{for } x_{-a} \in X_{-a},$$

which together with the cost functions c_a define the actual game in this case. For $\rho^\nu \ge 0$, the approximating game has for all $a \in A$,

$$c_a^\nu(x_a, x_{-a}) = c_a(x_a, x_{-a}) + \rho^\nu g_a(x_a, x_{-a})^+,$$

where $v^+ = \max\{0, v_1, \ldots, v_m\}$ for $v \in \mathbb{R}^m$, and $D_a^\nu(x_{-a}) = X_a$, i.e., in the approximating game, the agents are solving unconstrained problems, but with penalized cost functions. The corresponding Nikaido–Isoda bifunctions are given as

$$\begin{aligned}
&F(x, y) = \sum_{a\in A} \Big[c_a(x_a, x_{-a}) - c_a(y_a, x_{-a})\Big] \\
&\text{for } x \in C = \{x \in X \ : \ g_a(x_a, x_{-a}) \le 0 \ \forall a \in A\} \text{ and} \\
&y \in D(x) = \prod_{a\in A} \{y_a \in X_a \ : \ g_a(y_a, x_{-a}) \le 0\}; \\
&F^\nu(x, y) = \sum_{a\in A} \Big[c_a(x_a, x_{-a}) + \rho^\nu g_a(x_a, x_{-a})^+ - c_a(y_a, x_{-a}) - \rho^\nu g_a(y_a, x_{-a})^+\Big] \\
&\text{for } x \in C^\nu = X \text{ and } y \in D^\nu(x) = X.
\end{aligned}$$

The latter bifunctions lop-converge to the former under mild assumption.

Proposition 4.13 (Lop-convergence in penalty methods) *For continuous* $c_a : X_a \times X_{-a} \to \mathbb{R}$ *and* $g_a : X_a \times X_{-a} \to \mathbb{R}^m$*,* $a \in A$*, suppose that*

(a) *for all* $x \in X$ *there exists* $y \in X$ *with* $g_a(y_a, x_{-a}) \leq 0\ \forall a \in A$*; and*
(b) *for all* $x, y \in X$ *with* $g_a(y_a, x_{-a}) \leq 0\ \forall a \in A$*, there exists* $y^\nu \in X \to y$ *such that* $g_a(y^\nu_a, x_{-a}) < 0\ \forall a \in A$ *and* $\nu \in \mathbb{N}$*.*

If F *is the Nikaido–Isoda bifunction defined by* $\{(c_a, D_a), a \in A\}$*, with* $D_a(\cdot) = \{x_a \in X_a\ :\ g_a(x_a, \cdot) \leq 0\}$*, and* F^ν *is the Nikaido–Isoda bifunction defined by* $\{(c_a + \rho^\nu g_a^+, X_a), a \in A\}$*, then* $F^\nu \xrightarrow{lop} F$ *provided that* $\rho^\nu \to \infty$*.*

Proof We start by establishing that assumption (b), in conjunction with continuity, implies the seemingly stronger property that for all $x^\nu \in X \to x \in X$ and $y \in X$ with $g_a(y_a, x_{-a}) \leq 0\ \forall a \in A$, there exists $y^\nu \in X \to y$ such that $g_a(y^\nu_a, x^\nu_{-a}) \leq 0\ \forall a \in A$. Let $x, y \in X$ with $g_a(y_a, x_{-a}) \leq 0\ \forall a \in A$ and $x^\nu \in X \to x$. By assumption (b), there exists $\bar{y}^k \in X \to y$ such that $g_a(\bar{y}^k_a, x_{-a}) < 0\ \forall a \in A$ and $k \in \mathbb{N}$. By the continuity of g_a, there exists $\{\varepsilon^k > 0,\ k \in \mathbb{N}\}$ such that

$$g_a(\bar{y}^k_a, \bar{x}_{-a}) \leq 0 \text{ for all } a \in A \text{ and } \bar{x} \in \mathbb{B}(x, \varepsilon^k) = \{\bar{x} \in X\ :\ d_X(\bar{x}, x) \leq \varepsilon^k\},$$

where d_X is the metric on X. Since $x^\nu \to x$, there exists $\nu_1 \in \mathbb{N}$ such that $x^\nu \in \mathbb{B}(x, \varepsilon^1)$ for all $\nu \geq \nu_1$. Moreover, for $k = 2, 3, \ldots$, there exists $\nu_k > \nu_{k-1}$ such that $x^\nu \in \mathbb{B}(x, \varepsilon^k)$ for all $\nu \geq \nu_k$. We then construct the sequence $\{y^\nu, \nu \in \mathbb{N}, \nu \geq \nu_1\}$ by setting

$$y^\nu = \bar{y}^k \quad \text{for } \nu_k \leq \nu < \nu_{k+1}, \quad k \in \mathbb{N}.$$

Since $\bar{y}^k \to y$, $y^\nu \to y$. Thus, for all $k \in \mathbb{N}$, $a \in A$, and $\nu_k \leq \nu < \nu_{k+1}$,

$$g_a(y^\nu_a, x^\nu_{-a}) = g_a(\bar{y}^k_a, x^\nu_{-a}) \leq 0$$

because $x^\nu \in \mathbb{B}(x, \varepsilon^k)$. This establishes the claimed property.

We proceed by showing lop-convergence directly from Definition 3.1. Suppose that $x^\nu \in C^\nu \to x \in C$ and $y \in D(x)$. By the property just established, there exists $y^\nu \in X \to y$ such that $g_a(y^\nu_a, x^\nu_{-a})^+ = 0\ \forall a \in A$. Hence,

$$\begin{aligned}\liminf F^\nu(x^\nu, y^\nu) &= \liminf \sum_{a \in A} \Big[c_a(x^\nu_a, x^\nu_{-a}) + \rho^\nu g_a(x^\nu_a, x^\nu_{-a})^+ - c_a(y^\nu_a, x^\nu_{-a})\Big] \\ &\geq \liminf \sum_{a \in A} \Big[c_a(x^\nu_a, x^\nu_{-a}) - c_a(y^\nu_a, x^\nu_{-a})\Big] = F(x, y).\end{aligned}$$

Next, suppose that $x^\nu \in C^\nu \to x \notin C$. By assumption (a) and the established property, there exist $y \in X$ and $y^\nu \in X \to y$ with $g_a(y^\nu_a, x^\nu_{-a})^+ = 0\ \forall a \in A$. We then have that

$$F^\nu(x^\nu, y^\nu) = \sum_{a \in A} \Big[c_a(x^\nu_a, x^\nu_{-a}) + \rho^\nu g_a(x^\nu_a, x^\nu_{-a})^+ - c_a(y^\nu_a, x^\nu_{-a})\Big] \to \infty$$

because $\rho^\nu \to \infty$ and $\sum_{a\in A} g_a(x_a^\nu, x_{-a}^\nu)^+$ remains bounded away from zero as $\nu \to \infty$ by virtue of $x \notin C$ and the continuity of g_a for all $a \in A$. Consequently, part (a) of Definition 3.1 holds. Next, suppose that $x \in C$ and set $x^\nu = x$, which implies that $g_a(x_a, x_{-a})^+ = 0$ for all $a \in A$. If $y^\nu \in X \to y \in D(x)$, then

$$\begin{aligned}\limsup F^\nu(x^\nu, y^\nu) &= \limsup \sum_{a\in A} \Big[c_a(x_a, x_{-a}) - c_a(y_a^\nu, x_{-a}) - \rho^\nu g_a(y_a^\nu, x_{-a})^+\Big] \\ &\le \limsup \sum_{a\in A} \Big[c_a(x_a, x_{-a}) - c_a(y_a^\nu, x_{-a})\Big] = F(x, y).\end{aligned}$$

Otherwise, if $y^\nu \in X \to y \notin D(x)$, then we have that

$$F^\nu(x^\nu, y^\nu) = \sum_{a\in A} \Big[c_a(x_a, x_{-a}) - c_a(y_a^\nu, x_{-a}) - \rho^\nu g_a(y_a^\nu, x_{-a})^+\Big] \to -\infty$$

because $\rho^\nu \to \infty$ and $\sum_{a\in A} g_a(y_a^\nu, x_{-a})^+$ remains bounded away from zero as $\nu \to \infty$ due to the fact that $y \notin D(x)$. This establishes part (b) of Definition 3.1 and the proof is complete. □

Again, a constraint qualification enters; see assumption (b) in Proposition 4.13. In essence, we assume that given strategies of others for which an agent has a feasible response, the agent can unilaterally make its constraints inactive by moving slightly away from the response. The condition resembles the familiar Slater condition and other ones related to the existence of nonempty interiors of feasible sets. It is apparent that Proposition 4.13 in conjunction with Proposition 4.11 can be used to establish convergence of algorithmic approaches that rely on penalization of constraints and the solution of the resulting unconstrained game; see [15–17] for results in this direction, though under somewhat different assumptions. The proposition is an example of lop-convergence without the approximating constraints ever approaching those of the actual game.

5 Quantification of lopsided convergence

We next turn to a definition of "distance" between bifunctions that in some sense characterizes lop-convergence. We would like the distance between F^ν and F to tend to zero if and only if F^ν lop-converges to F. In this section, we develop for the first time such a distance, which essentially characterizes lop-convergence after passing to certain equivalence classes of bifunctions. We begin with establishing a foundation regarding a distance between (univariate) functions.

5.1 Attouch–Wets distance between functions

We limit the scope to the subset of fcns(X) consisting of lsc functions and let

$$\text{lsc-fcns}(X) := \{h \in \text{fcns}(X) \ : \ h \text{ lsc}\}.$$

This set is equipped with the *Attouch–Wets (aw) distance* $d\!l^{\text{aw}}$, which is given for a fixed but arbitrary point[6] $\bar{x} \in X$. Specifically, the aw-distance is defined for any $h, g \in \text{lsc-fcns}(X)$ as

$$d\!l^{\text{aw}}(h, g) := \int_0^\infty d\!l_\rho^{\text{aw}}(h, g)e^{-\rho}d\rho,$$

where the *ρ-aw-distance*, for any $\rho \geq 0$ is given by

$$d\!l_\rho^{\text{aw}}(h, g) := \sup\left\{\left|\text{dist}\left((x, \alpha), \text{epi}\, h\right) - \text{dist}\left((x, \alpha), \text{epi}\, g\right)\right| \;:\; d_X(x, \bar{x}) \leq \rho, |\alpha| \leq \rho\right\}$$

and

$$\text{dist}\left((x, \alpha), \text{epi}\, h\right) := \inf\left\{\max\{d_X(x, x'), |\alpha - \alpha'|\} \;:\; (x', \alpha') \in \text{epi}\, h\right\},$$

with a similar expression pertaining to epi g. We observe that $\bar{x}$ serves as the center of a ball on which $d\!l_\rho^{\text{aw}}$ is computed. The aw-distance defined here is essentially that pioneered in [5], which in turn can trace its origin from the study of convex cones in [32] and convex sets in [23]; see also [9,26].

It is immediate that $d\!l^{\text{aw}}$ is a metric on lsc-fcns(X). We deduce from [9, Theorem 3.1.7] that for $h^\nu, h \in \text{lsc-fcns}(X)$,

$$d\!l^{\text{aw}}(h^\nu, h) \to 0 \Longrightarrow h^\nu \xrightarrow{e} h.$$

If (X, d_X) is a finitely compact[7] metric space, then the converse also holds in view of [9, Theorems 3.1.4, 5.1.10, 5.2.10]; see also [27].

5.2 Lop-distance

Parallel to the restriction to lsc-fcns(X) in the definition of $d\!l^{\text{aw}}$, we focus on bifunctions with lsc sup-projections. Specifically, let

$$\text{lsc-bfcns}(X, Y) := \left\{F \in \text{bfcns}(X, Y) \text{ with lsc sup-projection}\right\}.$$

In view of Proposition 4.6, the sup-projection of a bifunction to which a sequence of bifunctions lop-converge ancillary-tightly must be lsc. In fact, it is well-known that sup-projections are lsc under mild assumptions. Sufficient conditions are given next, a result essentially in [6, Theorem 2, Section 2.5.2] but proved here for completeness.

Proposition 5.1 (lsc sup-projection) *When it exists, the sup-projection f of a bifunction $F \in$ bfcns(X, Y), with domain described (C, D), is lsc if either*

(a) *D is inner semicontinuous*[8] *and F is lsc; or*

[6] Although the topology induced by the aw-distance is the same for any point selected, the value of the distance will depend on this choice; see [5,27] and [26, Section 7.J] for details.

[7] A metric space is finitely compact if all its closed balls are compact.

[8] $D : C \rightrightarrows Y$ is inner semicontinuous if for every $x^\nu \in C \to x \in C$, InnLim $D(x^\nu) \supset D(x)$.

(b) *$D \subset Y$, i.e., the set-valued mapping is constant, and $F(\cdot, y)$ is lsc for all $y \in D$.*

Proof Let $x^\nu \in \operatorname{dom} f \to x \in \operatorname{dom} f$. Then $\sup_{y \in D(x)} F(x, y)$ is finite and for every $\varepsilon > 0$ there is a $y_\varepsilon \in D(x)$ such that $\sup_{y \in D(x)} F(x, y) \le F(x, y_\varepsilon) + \varepsilon$. Moreover, the inner semicontinuity of D implies that there exists a sequence $y_\varepsilon^\nu \in D(x^\nu) \to y_\varepsilon$. Since F is lsc, there exists $\bar\nu$ such that for all $\nu \ge \bar\nu$, $F(x^\nu, y_\varepsilon^\nu) \ge F(x, y_\varepsilon) - \varepsilon$. Thus, for $\nu \ge \bar\nu$,

$$\sup_{y \in D(x^\nu)} F(x^\nu, y) \ge F(x^\nu, y_\varepsilon^\nu) \ge F(x, y_\varepsilon) - \varepsilon \ge \sup_{y \in D(x)} F(x, y) - 2\varepsilon.$$

Since ε is arbitrary, $\liminf \sup_{y \in D(x^\nu)} F(x^\nu, y) \ge \sup_{y \in D(x)} F(x, y)$. Next, suppose that $x^\nu \in \operatorname{dom} f \to x \notin \operatorname{dom} f$. We consider two cases. First, let $x \notin C$. Then, there exist $y^\nu \in D(x^\nu)$ such that $F(x^\nu, y^\nu) \to \infty$. Thus, $\sup_{y \in D(x^\nu)} F(x^\nu, y) \to \infty$. Second, let $x \in C$, but $\sup_{y \in D(x)} F(x, y) = \infty$. Then, for every $M < \infty$ there is a $y_M \in D(x)$ such that $F(x, y_M) > M$. Following a similar argument as earlier, we find that $\liminf \sup_{y \in D(x^\nu)} F(x^\nu, y) > M$. Since M is arbitrary, the first conclusion follows. In the second case with D constant, the sequence $\{y_\varepsilon^\nu\}$ can be selected to coincide with y_ε for all ν and lsc in x only suffices. □

Definition 5.2 (*Lop-distance*) For any bifunctions $F, F' \in$ lsc-bfcns(X, Y), with sup-projections f, f', the lop-distance

$$d\!l^{\text{lop}}(F, F') := d\!l^{\text{aw}}(f, f').$$

Applications and estimates of the lop-distances are in their infancy. A first attempt in the context of minsup problems with $X = \mathbb{R}^n$, especially for optimization under stochastic ambiguity, can be found in [29], which leverages the present paper. There we see that the lop-distance between two bifunctions directly relates to the difference between their minsup-values.

It is immediately clear that the lop-distance is only a pseudometric on lsc-bfcns(X, Y) as there are nonidentical bifunctions with the same sup-projection. Thus, we pass to *equivalence classes*.

Definition 5.3 (*Equivalence classes of bifunctions*) Bifunctions $F, F' \in$ lsc-bfcns(X, Y) are equivalent, denoted by $F \sim F'$, if their sup-projections exist and are identical.

For bifunctions $F, F' \in$ lsc-bfcns(X, Y) with domains described by (C, D) and (C', D'), respectively, we might have $F \sim F'$ even if $C \ne C'$ as long as every $x \in C$ that is not in C' has $\sup_{y \in D(x)} F(x, y) = \infty$ and every $x \in C'$ that is not in C has $\sup_{y \in D'(x)} F'(x, y) = \infty$. Thus, the domains of the sup-projections can still coincide. Of course, the values of the sup-projections must also be the same. It is clear that $d\!l^{\text{lop}}$ is a metric on the quotient set of lsc-bfcns(X, Y).

Theorem 5.4 (Quantification of lop-convergence) *For $\{F, F^\nu, \nu \in \mathbb{N}\} \subset$ lsc-bfcns(X, Y), the following hold:*

(a) *If* $d\!l^{\text{lop}}(F^\nu, F) \to 0$, *then there exist* $\widetilde{F}^\nu, \widetilde{F} \in$ *lsc-bfcns*(X, Y) *such that*

$$\widetilde{F}^\nu \sim F^\nu, \widetilde{F} \sim F, \text{ and } \widetilde{F}^\nu \xrightarrow{lop} \widetilde{F}.$$

(b) *If* $F^\nu \xrightarrow{lop} F$ *ancillary-tightly and* (X, d_X) *is finitely compact, then* $d\!l^{\text{lop}}(F^\nu, F) \to 0$.

Proof Since $F, F^\nu \in$ lsc-bfcns(X, Y), the corresponding sup-projections f, f^ν exist. For part (a), set $\widetilde{C} = \text{dom } f$, which is nonempty, and $\widetilde{D} = Y$. We then define the bifunction $\widetilde{F} : \widetilde{C} \times Y \to \mathbb{R}$ have values $\widetilde{F}(x, y) = f(x)$ for $x \in \widetilde{C}$ and $y \in Y$. By construction, for $x \in \widetilde{C}$,

$$\sup_{y \in Y} \widetilde{F}(x, y) = \sup_{y \in Y} \sup_{y' \in D(x)} F(x, y') = f(x) < \infty,$$

where D describes the domain of F. Consequently, $\widetilde{f}$, the sup-projection of $\widetilde{F}$, exists and dom $\widetilde{f} = \widetilde{C}$. Moreover, $\widetilde{f} = f$ on $\widetilde{C}$. Thus, $\widetilde{F} \in$ lsc-bfcns(X, Y) and $\widetilde{F} \sim F$. An identical construction using the sup-projection f^ν of F^ν and setting $\widetilde{C}^\nu = \text{dom } f^\nu$, $\widetilde{D}^\nu = Y$, and $\widetilde{F}^\nu : \widetilde{C}^\nu \times Y \to \mathbb{R}$ having $\widetilde{F}^\nu(x, y) = f^\nu(x)$ for $x \in \widetilde{C}^\nu$ and $y \in Y$ results in $\widetilde{F}^\nu \in$ lsc-bfcns(X, Y) and $\widetilde{F}^\nu \sim F^\nu$. We then obtain for any $y \in Y$ that

$$d\!l^{\text{aw}}\left(\widetilde{F}^\nu(\cdot, y), \widetilde{F}(\cdot, y)\right) = d\!l^{\text{aw}}(f^\nu, f) = d\!l^{\text{lop}}\Big(F^\nu, F\Big) \to 0.$$

Since $d\!l^{\text{aw}}(\widetilde{F}^\nu(\cdot, y), \widetilde{F}(\cdot, y)) \to 0$ implies that $\widetilde{F}^\nu(\cdot, y) : \widetilde{C}^\nu \to \mathbb{R}$ epi-converges to $\widetilde{F}(\cdot, y) : \widetilde{C} \to \mathbb{R}$, it follows from Proposition 4.1 that $\widetilde{F}^\nu$ lop-converges to $\widetilde{F}$.

For part (b) we conclude from Theorem 4.4 that f^ν epi-converges to f. Thus, $d\!l^{\text{aw}}(f^\nu, f) \to 0$ under the additional assumption that (X, d_X) is finitely compact and the result follows from the definition of the lop-distance. □

We recall that ancillary-tightness is ensured, for example, by some compactness property on the sets over which the inner maximization is taking place. Thus, a restriction to bifunctions satisfying such properties would ensure that the lop-distance fully characterizes lopsided convergence on this class provided that (X, d_X) is finitely compact. The lop-distance generates what we define as the *lop-topology* on the space of equivalence classes of such bifunctions.

Acknowledgements This material is based upon work supported in part by DARPA under Grant HR0011-14-1-0060 and the U.S. Office of Naval Research under Grant No. N00014-17-2372.

References

1. Attouch, H.: Variational Convergence for Functions and Operators, Applicable Mathematics Sciences. Pitman, Totowa (1984)
2. Attouch, H., Azé, D., Wets, R.J-B: Convergence of convex-concave saddle functions: continuity properties of the Legendre-Fenchel transform with applications to convex programming and mechanics. Ann. l'Inst. H. Poincaré Anal. Nonlinéaire 5, 537–572 (1988)
3. Attouch, H., Wets, R.J-B: Convergence des points min/sup et de points fixes. C. R. l'Acad. Sci. Paris **296**, 657–660 (1983)

4. Attouch, H., Wets, R.J-B: A convergence theory for saddle functions. Trans. Am. Math. Soc. **280**(1), 1–41 (1983)
5. Attouch, H., Wets, R.J-B: Quantitative stability of variational systems: I. The epigraphical distance. Trans. Am. Math. Soc. **328**(2), 695–729 (1991)
6. Aubin, J.-P.: Mathematical Methods of Game and Economic Theory. North-Holland, Amsterdam (1982)
7. Aubin, J.-P., Ekeland, I.: Applied Nonlinear Analysis. Issue 1237 of Pure and Applied Mathematics. Wiley, New York (1984)
8. Aubin, J.-P., Frankowska, H.: Set-Valued Analysis. Birkhäuser, Basel (1990)
9. Beer, G.: Topologies on Closed and Closed Convex Sets, Volume 268 of Mathematics and Its Applications. Kluwer, Dordrecht (1992)
10. Beer, G., Rockafellar, R.T., Wets, R.J-B: A characterization of epi-convergence in terms of convergence of level sets. Proc. Am. Math. Soc. **116**(3), 753–761 (1992)
11. Ben-Tal, A., El Ghaoui, L., Nemirovski, A.: Robust Optimization. Princeton University Press, Princeton (2009)
12. Deride, J., Jofré, A., Wets, R.J-B: Solving deterministic and stochastic equilibrium problems via augmented Walrasian. Comput. Econ. (2018, to appear)
13. El Ghaoui, L., Calafiore, G.C.: Optimization Models. Cambridge University Press, Cambridge (2014)
14. Facchinei, F., Kanzow, C.: Generalized Nash equilibrium problems. Ann. Oper. Res. **175**(1), 177–221 (2010)
15. Facchinei, F., Kanzow, C.: Penalty methods for the solution of generalized Nash equilibrium problems. SIAM J. Optim. **20**(5), 2228–2253 (2010)
16. Facchinei, F., Pang, J.-S.: Exact penalty functions for generalized Nash problems. In: Di Pillo, G., Roma, M. (eds.) Large-Scale Nonlinear Optimization, pp. 115–126. Springer, Berlin (2006)
17. Fukushima, M., Pang, J.-S.: Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games. CMS **2**, 21–56 (2005)
18. Gürkan, G., Pang, J.-S.: Approximations of Nash equilibria. Math. Program. B **117**, 223–253 (2009)
19. Jofré, A., Wets, R.J-B: Continuity properties of Walras equilibrium points. Ann. Oper. Res. **114**, 229–243 (2002)
20. Jofré, A., Wets, R.J-B: Variational convergence of bivariate functions: lopsided convergence. Math. Program. B **116**, 275–295 (2009)
21. Jofré, A., Wets, R.J-B: Variational convergence of bifunctions: motivating applications. SIAM J. Optim. **24**(4), 1952–1979 (2014)
22. Morgan, J., Scalzo, V.: Variational stability of social Nash equilibria. Int. Game Theory Rev. **10**(1), 17–24 (2008)
23. Mosco, U.: Convergence of convex sets and of solutions of variational inequalities. Adv. Math. **3**(4), 510–585 (1969)
24. Nikaido, H., Isoda, K.: Note on noncooperative convex games. Pac. J. Math. **5**, 807–815 (1955)
25. Rockafellar, R.T.: Convex Analysis. Vol. 28 of Princeton Math. Series. Princeton University Press, Princeton (1970)
26. Rockafellar, R.T., Wets, R.J-B: Variational Analysis, Volume 317 of Grundlehren der Mathematischen Wissenschaft, 3rd printing-2009 edition. Springer, Berlin (1998)
27. Royset, J.O.: Approximations and solution estimates in optimization. Math. Program. (2018). https://doi.org/10.1007/s10107-017-1165-0
28. Royset, J.O., Wets, R.J-B: Optimality functions and lopsided convergence. J. Optim. Theory Appl. **169**(3), 965–983 (2016)
29. Royset, J.O., Wets, R.J-B: Variational theory for optimization under stochastic ambiguity. SIAM J. Optim. **27**(2), 1118–1149 (2017)
30. Stein, O.: Bi-level Strategies in Semi-Infinite Programming. Kluwer, Dordrecht (2003)
31. Sun, H., Xu, H.: Convergence analysis for distributionally robust optimization and equilibrium problems. Math. Oper. Res. **41**, 377–401 (2015)
32. Walkup, D., Wets, R.J-B: Continuity of some convex-cone-valued mappings. Proc. Am. Math. Soc. **18**, 229–235 (1967)
33. Zhang, J., Xu, H., Zhang, L.: Quantitative stability analysis for distributionally robust optimization with moment constraints. SIAM J. Optim. **26**(3), 1855–1882 (2016)