# VARIATIONAL APPROXIMATIONS OF A DUAL PAIR OF MATHEMATICAL PROGRAMMING PROBLEMS

HUYNH THI HONG DIEM

Abstract. We study variational approximations of a dual pair of mathematical programming problems in terms of epi/hypo-convergence and inside epi/hypo-convergence of approximating Lagrange functions of the pair. First, the Painlevé-Kuratowski convergence of approximate saddle points of approximating Lagrange functions is established under the inside epi/hypo-convergence of these approximating Lagrange functions. From this, we obtain a couple of solutions of the pair of problems and a strong duality. Under a stronger variational convergence called ancillary tight epi/hypo-convergence, we obtain the Painlevé-Kuratowski convergence of approximate minsup-points and approximate maxinf-points of approximating Lagrange functions (when approximate saddle points are not necessary to exist).

**Keywords:** Variational approximations; mathematical programming; epi/hypo-convergence; inside epi/hypo-convergence; ancillary tightness; Lagrange functions; strong duality; approximate saddle points; approximate minsup-points; approximate maxinf-points

**MSC(2010):** 90C30, 90C31, 90C47, 49J53

## 1. Introduction

The term "approximation" is common in all areas of mathematics because when facing a complicated problem, the first and popular idea is studying simpler problems approximating the original one. Here, approximating problems naturally mean that they converge in a certain sense to the original one. In optimization, most problems are related to variational properties such as being their extremum values, minimizers, maximizers, minsup-values, minsup-points, saddle points, etc. Hence, the most interesting one of the aforementioned senses of convergence is that the convergence preserves variational properties, i.e., such properties of approximating problems are guaranteed to be preserved for the approximated problem through this convergence. The term "variational convergence" is used for any type of convergence ensuring this preseration. So, this is a common terminology for types of convergence, not an exactly-defined type of convergence.

*Date*: Received: September 12, 2020, Accepted: ...., 2020.

For constrained minimization, the basic variational convergence is epi-convergence of a unifunction (called simply a function as well for convenience) $\varphi : X \to \mathbb{R} \cup \{+\infty\}$ with $X$ being a space, see, e.g., [1, 10] for main properties and applications, as well as relationships with other convergence concepts. From [7], this convergence is also modified for a finite-valued function $\varphi : A \to \mathbb{R}$ for a nonempty subset $A \subset X$ (see Definition 2.2 in Section 2 below). For the last three decades, the model of equilibrium problems has been intensively considered as a general one including most of optimization-related problems as special cases. To see how it can so contain them, consider for instance the simple minimization problem of finding $\bar{x} \in A$ to minimize $\varphi(x)$ on $A$, where $A \subset X$ is nonempty, and $\varphi : A \to \mathbb{R}$. Then, by using the bifunction $\Phi : A \times A \to \mathbb{R}$ defined by $\Phi(x, y) := \varphi(y) - \varphi(x)$ for $x, y \in A$, the above minimization problem becomes a special case of the so-called equilibrium problem: find $\bar{x} \in A$ such that $\Phi(\bar{x}, y) \geq 0$ for all $y \in A$. This bifunction was proposed in [9] and now is called the Nikaido-Isoda bifunction and the first existence result for an equilibrium problem was published in [4], but for a period beginning by that paper, the problem was called the Ky Fan inequality. After the publication [3], such a problem is often said to be an equilibrium problem.

Equilibrium problems have been also extended to the case that $\Phi$ is defined on $A \times B$ with a nonempty set $B \subset Y \neq X$. Like for various optimization-related problems in general, many results for minimization problems can be derived from the correponding ones for equilibrium problems. However, when dealing with duality properties, this approach is not useful. Namely, following the duality framework for equilibrium problems, for a primal problem (stated above), which is also called a Stampacchia equilibrium problem, its dual problem is the Minty equilibrium problem: find $\bar{y} \in A$ such that $\Phi(x, \bar{y}) \leq 0$ for all $x \in A$. Then, the two dual problems clearly coincide, i.e., this definition of dual problems has no meaning for a (scalar minimization problem). On the other hand, for such a problem, the more traditional and important model of duality is the Lagrange duality scheme with the use of a bifunction being the Lagrangian.

Motivated by the above discussions, in this paper, we study variational approximations of a mathematical programming problem (the classical model of minimization problem) and its Lagrange dual. So, we apply variational convergence of bifunctions. Its basic types are epi/hypo-convergence and lopsided convergence, including minsup-lop-convergence and maxinf-lop-convergence. However, the last two concepts are non-symmetric and not suitable for considerations of duality properties. Hence, we choose epi/hypo-convergence and show in Section 3 that our approach is really effective.

As our notations are standard, we mention only several ones. For a metric space $X$ and $A \subset X$, int$A$ and bd$A$ stand for the interior and boundary, resp, of $A$. $B(x, r)$ denotes the open ball with center $x$ and radius $r$. $\varepsilon \searrow 0$ means $\varepsilon \leq 0$ and tending to 0.

For a function $\varphi : X \to \bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$, its domain, epigraph, and hypograph are defined by $\mathrm{dom}\varphi := \{x \in X \mid \varphi(x) < +\infty\}$, $\mathrm{epi}\varphi := \{(x,r) \in X \times \mathbb{R} \mid \varphi(x) \le r\}$, and $\mathrm{hypo}\varphi := \{(x,r) \in X \times \mathbb{R} \mid \varphi(x) \ge r\}$, resp. liminf$\varphi$ and limsup$\varphi$ designate the lower and upper limits of $\varphi$, defined, resp, by

$$\mathrm{liminf}_{x\to\bar{x}}\varphi(x) := \lim_{\delta\searrow 0}[\inf_{x\in B(\bar{x},\delta)}\varphi(x)] = \sup_{\delta>0}[\inf_{x\in B(\bar{x},\delta)}\varphi(x)],$$

$$\mathrm{limsup}_{x\to\bar{x}}\varphi(x) := \lim_{\delta\searrow 0}[\sup_{x\in B(\bar{x},\delta)}\varphi(x)] = \inf_{\delta>0}[\sup_{x\in B(\bar{x},\delta)}\varphi(x)].$$

We adopt the notation

$$\mathrm{argmin}_A\varphi := \begin{cases} \{x \in X \mid \varphi(x) = \inf_A\varphi\} & \text{if} \quad \inf_A\varphi < +\infty, \\ \emptyset & \text{if} \quad \inf_A\varphi = +\infty, \end{cases}$$

$$\varepsilon\text{-}\mathrm{argmin}_A\varphi := \begin{cases} \{x \in X \mid \varphi(x) \le \inf_A\varphi + \varepsilon\} & \text{if} \quad \inf_A\varphi < +\infty, \\ \emptyset & \text{if} \quad \inf_A\varphi = +\infty, \end{cases}$$

and similarly for argmax and $\varepsilon$-argmax. For a sequence of subsets $\{A^k\}_{k\in\mathbb{N}}$ in $X$, the lower/inner limit is defined by

$$\mathrm{Liminf}_k A^k := \{x \in X \mid \exists x^k \to x \text{ with } x^k \in A^k\}.$$

The upper limit is

$$\mathrm{Limsup}_k A^k := \{x \in X \mid \exists \{k_l\}_l,\ \exists x^{k_l} \to x \text{ with } x^{k_l} \in A^{k_l}\}.$$

If $\mathrm{Liminf}_k A^k = \mathrm{Limsup}_k A^k$, one says that $A^k$ tends to $A$ in the Painlevé-Kuratowski sense. In the sequel, we usually use the abbreviations li, ls, Li, Ls for liminf, limsup, Liminf, Limsup, resp.

## 2. Preliminaries

2.1. **Dual pairs of problems in mathematical programming.** Consider the mathematical programming problem

$$\text{(NP)} \qquad \min\varphi(x) \text{ s.t. } x \in A,\ g(x) \le 0,\ h(x) = 0,$$

where $X$, $Z$ and $W$ are normed spaces, $A \subset X$ is nonempty, $C \subset Z$ is a convex ordering cone, $g : X \to Z$ and $h : X \to W$. For points in $Z$, we write $z_1 \le z_2$ if $z_2 \in z_1 + C$ and denote the feasible region of (NP) by $\Omega$ and the solution set of (NP) by Sol(NP). Consider the following Lagrangian (Lagrange function) of problem (NP)

$$\mathcal{L}(x,\mu,\nu) := \varphi(x) + \langle \mu, g(x)\rangle + \langle \nu, h(x)\rangle$$

for $x \in A$ and $(\mu,\nu) \in C^* \times W^* =: B$, with $Z^*$ and $W^*$ being the topological dual spaces of $Z$ and $W$, resp, and $C^* := \{y^* \in Z^* \mid \langle y^*, y\rangle \ge 0\}$ is the positive dual cone of $C$.

Now we clarify basic definitions and facts about (NP) and its dual. Some of them are known. However, to make the research situation clear and avoid confusions, we systematically discuss beriefly but enough detail because it may not easy to find reference material. It is convenient to write (NP) in the following equivalent minsup problem for $\mathcal{L}$

(MisP) $\min_{x\in A}\sup_{(\mu,\nu)\in B}\mathcal{L}(x,\mu,\nu).$

The solution set of (MisP) is denoted by Sol(MisP), which is also called the set of minsup points of $\mathcal{L}$, denoted by mis$\mathcal{L}$. The equivalence of (NP) and (MisP) means that the optimal values and the solution sets are equal. To see this equivalence, observe that

$$\sup_{(\mu,\nu)\in B}(\langle\mu,g(x)\rangle+\langle\nu,h(x)\rangle)=\begin{cases}0 & \text{if g(x)} \leq 0, \text{h(x)} = 0,\\ \infty & \text{otherwise.}\end{cases}$$

Hence,

$$\min_{x\in A}\sup_{(\mu,\nu)\in B}\mathcal{L}(x,\mu,\nu)=\min_{x\in\Omega}\varphi(x).$$

The function $\zeta(x):=\sup_{(\mu,\nu)\in B}\mathcal{L}(x,\mu,\nu)$ is called the sup-projection of bifunction $\mathcal{L}$ (of two components $x$ and $(\mu,\nu)$). Then, (NP) can be rewritten also as $\min_{x\in A}\zeta(x)$. From now on, we write simply $\min_A$ instead of $\min_{x\in A}$ if there is no thread of confusion and no need to emphasize $x$; and the same for infima, suprema, etc.

Define $\eta(\mu,\nu)=\inf_{x\in A}\mathcal{L}(x,\mu,\nu)$ and call it the inf-projection of $\mathcal{L}$. Then, the Lagrange dual problem (DNP) of (NP) is defined as

(DNP) $\max_{(\mu,\nu)\in B}\eta(\mu,\nu).$

To see more clearly the duality nature, we rewrite (DNP) as the maxinf problem

(MaiP) $\max_{(\mu,\nu)\in B}\inf_{x\in A}\mathcal{L}(x,\mu,\nu).$

Let Sol(DNP), Sol(MaiP), and mai$\mathcal{L}$ stand for the solution sets of (DNP), (MaiP), and the set of maxinf points of $\mathcal{L}$, resp. Of course, the three sets are equal.

For all $x\in\Omega$ and $(\mu,\nu)\in B$, one has

$$\eta(\mu,\nu)=\inf_{x'\in A}\mathcal{L}(x',\mu,\nu)\leq\mathcal{L}(x,\mu,\nu)\leq\varphi(x).$$

So, $\eta(\mu,\nu)\leq\varphi(x)$. This property (holds for all feasible points of (NP) and (DNP) in any Lagrange duality scheme) is called the weak duality. Hence,

$$\sup_{(\mu,\nu)\in B}\eta(\mu,\nu)=\sup_{(\mu,\nu)\in B}\inf_{x'\in A}\mathcal{L}(x',\mu,\nu)$$

$$\leq\inf_{x\in A}\sup_{(\mu,\nu)\in B}\mathcal{L}(x,\mu,\nu)=\inf_{x\in\Omega}\varphi(x). \tag{2.1}$$

If the two optimal objective values are attained and equal and there exist solutions of (NP) and (DNP) (i.e., mis$\mathcal{L}$ and mai$\mathcal{L}$ are nonempty): $\min_{x\in\Omega}\varphi(x)=\max_{(\mu,\nu)\in B}\eta(\mu,\nu)$, then we say that the strong duality holds or the duality gap is zero (or without duality gap). Otherwise, we say that the duality gap is nonzero (or with a duality gap).

We provide several simple nonlinear problems with different situations about duality.

*Example 2.1* (linear discrete programs with or without duality gap) Consider the problem, for $\varphi, h : \mathbb{R}^2 \to \mathbb{R}$,

$$\min \varphi(x) = -2x_1 + x_2$$
$$\text{s.t. } h(x) = x_1 + x_2 - 3 = 0,$$
$$x \in A := \{(0,0),(0,3),(1,3),(2,1)\}.$$

Substituting $-x_1 = x_2 + 3$ into $\varphi$, (NP) becomes $\min_{x\in A}(3x_2 - 6)$. So, the solution is $\bar{x} = (2,1)$ with $\varphi(\bar{x}) = -3$. If we solve (MisP), i.e., $\min_{x\in A}\zeta(x)$, we obtain the minimizer $\bar{x} = (2,1)$ with $\zeta(2,1) = -3$ because $\zeta(x)$ tends to $\infty$ at $x = (0,0)$ (when $\nu \to -\infty$) and at $x = (1,3)$ (when $\nu \to \infty$), and $\zeta(x)$ equals to 3 at $x = (0,3)$ and to $-3$ at $x = (2,1)$. Hence, $\min_{x\in A}\max_{\nu\in\mathbb{R}}\mathcal{L} = -3$ attains at $\bar{x} = (2,1)$ for any $\bar{\nu} \in \mathbb{R}$. Considering (DNP), we obtain $\eta(\nu) = \min\{-3\nu, 3, 1+\nu, -3\}$ (these values attain at $x$ equal to $(0,0)$, $(0,3)$, $(1,3)$, and $(2,1)$, resp). Hence, $\max_{\nu\in\mathbb{R}}\eta(\nu) = -3$ achieves at $\bar{x} = (2,1)$ and $\nu \in [-4,1]$. Therefore, we have a strong duality and Sol(NP)= $\{(2,1)\}$ and Sol(DNP)= $[-4,1]$ (we see that any point $(\bar{x},\bar{\nu}) \in$ Sol(NP)$\times$Sol(DNP) is a saddle point of $\mathcal{L}$), though the set of minmax points is $\{(2,1)\} \times \mathbb{R}$.

Now we modify the problem, taking the new set

$$A = \{(0,0),(0,4),(4,4),(4,0),1,2),(2,1)\}$$

and keeping the same $\varphi$ and $h$. Then similarly, we see that the solution of (NP) is $\bar{x} = (2,1)$ with $\varphi(\bar{x}) = -3$. By direct calculations, we have

$$\eta(\nu) = \begin{cases} -4+5\nu & \text{if } \nu \le -1, \\ -8+\nu & \text{if } -1 \le \nu \le 2, \\ -3\nu & \text{if } \nu \ge 2. \end{cases}$$

The solution is $\bar{\nu} = 2$ with $\eta(\bar{\nu}) = -6$. So, the primal optimal value and the dual optimal value are different, i.e., we have a duality gap.

*Example 2.2* (convex program with a duality gap) Consider the problem

$$\min x_2$$
$$\text{s.t. } g_1(x) = 1 - x_1 \le 0,\ g_2(x) = x_1^2 + x_2^2 - 1 \le 0.$$

(so, the objective is linear and the feasible region is convex). It is easy to see that the unique (optimal) solution is $(\bar{x}_1, \bar{x}_2) = (1,0)$ with the optimal objective value equal zero.

For the dual problem, we have

$$\eta(\mu) = \inf_{x\in\mathbb{R}^2}\mathcal{L}(x,\mu) =$$

$$\begin{cases} -\infty \ \text{ if } \mu_1 = \mu_2 = 0 \text{ when } \mathrm{x}_2 \to -\infty, \\ -\mu_2 \ \text{ if } \mu_1 = 0, \mu_2 > 0 \text{ (attained at } \mathrm{x}_1 = \mathrm{x}_2 = 0), \\ -\infty \ \text{ if } \mu_1 > 0, \mu_2 = 0 \text{ (when } \mathrm{x}_2 \to -\infty), \\ -(4\mu_2)^{-1} + \mu_1 - \mu_2 - \mu_1^2(4\mu_2)^{-1} \\ \qquad\qquad \text{if } \mu_1, \mu_2 > 0 \text{ (attained at } \mathrm{x}_1 = \mu_1(2\mu_2)^{-1},\ \mathrm{x}_2 = -(2\mu_2)^{-1}). \end{cases}$$

Hence,

$$\max_{\mu\in B}\eta(\mu) = \max\begin{cases}-\mu_2 & \text{if } \mu_1 = 0,\ \mu_2 > 0,\\ -(4\mu_2)^{-1} + \mu_1 - \mu_2 - \mu_1^2(4\mu_2)^{-1} & \text{if } \mu_1 > 0,\ \mu_2 > 0.\end{cases}$$

One sees that $\sup_{\mu\in B}\eta(\mu) = 0$, but the maximum in not achieved. Therefore, in (2.1) for this pair of dual problem, one has the equality supinf = infsup, but one still does not have a strong duality between (NP) and (DNP), because (DNP) does not have solutions.

Now we provide a case that the strong duality holds for a nonconvex (so nonlinear) program.

*Example 2.3* (nonconvex program with a strong duality) Consider the case $x \in \mathbb{R}$, $\varphi(x) = \sqrt{x}$, $g(x) = x - 3$ and $A = \mathbb{R}_+$. Then, clearly $\bar{x} = 0$ is the unique primal optimal solution with $\varphi(\bar{x}) = 0$. For each $\mu \geq 0$, $\eta(\mu) = \min_{x\geq 0}(\sqrt{x} + \mu(x-3)) = -3\mu$ (attained at $x = 0$). Hence, $\max_{\mu\geq 0}\eta(\mu) = 0$ attained at $\mu = 0$ which is the unique dual optimal solution, and we have a strong duality.

Strong duality in duality schemes defined via bifunctions (not only for Lagrange duality) is closely connected with saddle points. For the Lagrange function, we have the following basic concepts of points related to extrema.

**Definition 2.1.** (i) A point $\bar{x} \in A$ is called a minsup-point (maxinf-point, respectively (resp)) of $\mathcal{L}$, denoted by $\bar{x} \in \text{mis}\mathcal{L}$ ($(\bar{\mu},\bar{\nu}) \in \text{mai}\mathcal{L}$), if

$$\sup_{(\mu,\nu)\in B}\mathcal{L}(\bar{x},(\mu,\nu)) = \min_{x\in A}\sup_{(\mu,\nu)\in B}\mathcal{L}(x,(\mu,\nu))$$

$$(\inf_{x\in A}\mathcal{L}(x,(\bar{\mu},\bar{\nu})) = \max_{(\mu,\nu)\in B}\inf_{x\in A}\mathcal{L}(x,(\mu,\nu)),\ \text{resp}).$$

When $\bar{x} \in \text{mis}\mathcal{L}$, $(\bar{x},(\bar{\mu},\bar{\nu}))$ is termed a minmax-point of $\mathcal{L}$, denoted by $(\bar{x},(\bar{\mu},\bar{\nu})) \in \text{mima}\mathcal{L}$, if $(\bar{\mu},\bar{\nu}) \in \text{argmax}_{(\mu,\nu)\in B}\mathcal{L}(x,(\mu,\nu))$ (a maxmin-point $(\bar{x},(\bar{\mu},\bar{\nu})) \in \text{mami}\mathcal{L}$ is defined similarly).

(ii) A point $(\bar{x},(\bar{\mu},\bar{\nu})) \in A\times B$ is said to be a saddle point of $\mathcal{L}$, denoted by $(\bar{x},(\bar{\mu},\bar{\nu})) \in \text{sdl}\mathcal{L}$, for all $x \in A$ and $(\mu,\nu) \in B$, if $\mathcal{L}(\bar{x},(\mu,\nu)) \leq \mathcal{L}(\bar{x},(\bar{\mu},\bar{\nu})) \leq \mathcal{L}(x,(\bar{\mu},\bar{\nu}))$.

Clearly, a point $(\bar{x},(\bar{\mu},\bar{\nu}))$ is a saddle point of $\mathcal{L}$ if and only if $(\bar{x},(\bar{\mu},\bar{\nu})) \in \text{mima}\mathcal{L}\cap\text{mami}\mathcal{L}$. Observe that in Example 2.3, $(0,0) \in \text{sdl}\mathcal{L}$.

To study variational approximations of a dual pair of mathematical programming problems in the next section, we need the following concepts of approximate saddle points and approximate solutions of (NP) and (DNP). For $\varepsilon \geq 0$, a point $\bar{x}_\varepsilon \in \Omega$ is called an $\varepsilon$-solution of (NP), denoted by $\bar{x}_\varepsilon \in \varepsilon$-Sol(NP), if $\varphi(\bar{x}_\varepsilon) \leq \inf_{x\in\Omega}\varphi(x) + \varepsilon$. Then, clearly $\bar{x}_\varepsilon$ is also an $\varepsilon$-minimizer of $\zeta$ (on $A$), i.e., $\bar{x}_\varepsilon \in \varepsilon\text{-min}\zeta$, and also an $\varepsilon$-minsup point of $\mathcal{L}$, denoted by $\bar{x}_\varepsilon \in \varepsilon\text{-mis}\mathcal{L}$, in the sense that

$$\sup_B\mathcal{L}(\bar{x}_\varepsilon,\mu,\nu) \leq \min_A\sup_B\mathcal{L}(x,\mu,\nu) + \varepsilon.$$

Similarly, $\varepsilon$-Sol(DNP) $= \varepsilon$-max$\eta = \varepsilon$-mai$\mathcal{L}$. Observe that $\bar{x}_\varepsilon \in \varepsilon$-mis$\mathcal{L}$ and $(\bar{\mu}_\varepsilon, \bar{\nu}_\varepsilon) \in \varepsilon$-mai$\mathcal{L}$ at the same time mean that

$$\text{max}_B\text{inf}_A\mathcal{L} - \varepsilon \le \mathcal{L}(\bar{x}_\varepsilon, \bar{\mu}_\varepsilon, \bar{\nu}_\varepsilon) \le \text{min}_A\text{sup}_B\mathcal{L} + \varepsilon. \tag{2.2}$$

By definition, $(\bar{x}_\varepsilon, (\bar{\mu}_\varepsilon, \bar{\nu}_\varepsilon))$ is an $\varepsilon$-saddle point of $\mathcal{L}$, denoted by $(\bar{x}_\varepsilon, (\bar{\mu}_\varepsilon, \bar{\nu}_\varepsilon)) \in \varepsilon$-sdl$\mathcal{L}$, if and only if one has (stronger than (2.2))

$$\mathcal{L}(\bar{x}_\varepsilon, \mu, \nu) - \varepsilon \le \mathcal{L}(\bar{x}_\varepsilon, \bar{\mu}_\varepsilon, \bar{\nu}_\varepsilon) \le \mathcal{L}(x, \bar{\mu}_\varepsilon, \bar{\nu}_\varepsilon) + \varepsilon$$

for all $x \in A$ and $(\mu, \nu) \in B$. Evidently, $(\bar{x}_\varepsilon, (\bar{\mu}_\varepsilon, \bar{\nu}_\varepsilon)) \in \varepsilon$-sdl$\mathcal{L}$ if and only if $(\bar{x}_\varepsilon, (\bar{\mu}_\varepsilon, \bar{\nu}_\varepsilon))$ is both an $\varepsilon$-minmax point and $\varepsilon$-maxmin point of $\mathcal{L}$ in the sense that

$$\text{max}_B\mathcal{L}(\bar{x}_\varepsilon, \mu, \nu) - \varepsilon \le \mathcal{L}(\bar{x}_\varepsilon, \bar{\mu}_\varepsilon, \bar{\nu}_\varepsilon) \le \text{min}_A\mathcal{L}(x, \bar{\mu}_\varepsilon, \bar{\nu}_\varepsilon) + \varepsilon.$$

2.2. **Variational convergence.** We present briefly the definitions and properties of variational convergence needed for the subsequent sections.

In this section let $X$, $Y$ be metric spaces. First, consider $A, A^k \subset X$, $\varphi^k : A^k \to \mathbb{R}$, and $\varphi : A \to \mathbb{R}$.

**Definition 2.2.** (epi-convergence [7], inside epi-convergence [6]) $\{\varphi^k\}_k$ is called epi-convergent to $\varphi$, denoted by $\varphi^k \xrightarrow{e} \varphi$ or $\varphi =$ e-lim$_k\varphi^k$ if the following conditions are satisfied

- (a) for all $x^{k_j} \in A^{k_j} \to x$, liminf$_j\varphi^{k_j}(x^{k_j}) \ge \varphi(x)$ if $x \in A$ and $\varphi^{k_j}(x^{k_j}) \to +\infty$ if $x \notin A$;
- (b) for all $x \in A$, there exists $x^k \in A^k \to x$ such that limsup$_k\varphi^k(x^k) \le \varphi(x)$.

  Omitting the case that $x \notin A$ with the infinity condition, one has inside epi-convergence.

The above definition of the basic variational convergence was first introduced in [12] for functions $\varphi^k : X \to \mathbb{R} \cup \{+\infty\}$ and has been developed with many applications (see, e.g., [1, 10]). In [7], the above Definition 2.2 for epi-convergence together with a modification for finite-valued bifunctions $\Phi : A \times B \to \mathbb{R}$ of the concept of misup-lop-convergence defined in [2] (for $\Phi : X \times Y \to \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$) in order to apply more effectively to practical problems involving finite-valued bifunctions. The weaker notion of inside epi-convergence was recently proposed in [6] with effective applications in approximations. A notion symmetric to epi-convergence (and strongly concerned when maximization is considered) is: a sequence $\{\varphi^k\}_k$ is called hypo-convergent to $\varphi$, denoted by $\varphi^k \xrightarrow{h} \varphi$ or $\varphi =$ h-lim$_k\varphi^k$, if $\{-\varphi^k\}_k$ epi-converges to $-\varphi$. Hence, $\varphi^k \xrightarrow{h} \varphi$ means the following two conditions (symmetric to the above (a) and (b) for epi-convergence) are satisfied

(a') for all $x^{k_j} \in A^{k_j} \to x$, limsup$_j\varphi^{k_j}(x^{k_j}) \le \varphi(x)$ if $x \in A$ and $\varphi^{k_j}(x^{k_j}) \to -\infty$ if $x \notin A$;

(b') for all $x \in A$, there exists $x^k \in A^k \to x$ such that liminf$_k\varphi^k(x^k) \ge \varphi(x)$.

Omitting the case that $x \notin A$ with the infinity condition, one has inside hypo-convergence.

The basic variational property of epi-convergence is

**Proposition 2.3.** (see, e.g., [6])

(i) *If* $\{\varphi^k\}_k$ *satisfies Definition 2.2* (b) *of inside epi-convergence, then* $\mathrm{limsup}_k(\inf_{A^k}\varphi^k) \leq \inf_A \varphi$.

(ii) *If* $\varphi^k \xrightarrow{e} \varphi$, *then for any* $\varepsilon^k \to 0^+$, $\mathrm{Limsup}_k \varepsilon^k\text{-argmin}_{A^k}\varphi^k \subset \mathrm{argmin}_A \varphi$.

Now consider $A, A^k \subset X$, $B, B^k \subset Y$, and (finite-valued) bifunctions $\Phi^k : A^k \times B^k \to \mathbb{R}$ and $\Phi : A \times B \to \mathbb{R}$.

**Definition 2.4.** (epi/hypo convergence [5], inside epi/hypo-convergence [6]) Bifunctions $\Phi^k$, are called epi/hypo convergent (e/h-convergent) to a bifunction $\Phi$ if

(a) for all $x^{k_j} \in A^{k_j} \to x \in A$ and $y \in B$, there exist $y^{k_j} \in B^{k_j} \to y$ such that $\mathrm{liminf}_j \Phi^{k_j}(x^{k_j}, y^{k_j}) \geq \Phi(x, y)$, and for all $x^{k_j} \in A^{k_j} \to x \notin A$, there exist $y^{k_j} \in B^{k_j}$ such that $\Phi^{k_j}(x^{k_j}, y^{k_j}) \to +\infty$;

(b) for all $y^{k_j} \in B^{k_j} \to y \in B$ and $x \in A$, there exist $x^{k_j} \in A^{k_j} \to x$ such that $\mathrm{limsup}_j \Phi^{k_j}(x^{k_j}, y^{k_j}) \leq \Phi(x, y)$, and for all $y^{k_j} \in B^{k_j} \to y \notin B$, there exist $x^{k_j} \in A^{k_j}$ such that $\Phi^{k_j}(x^{k_j}, y^{k_j}) \to -\infty$.

The inside epi/hypo-convergence (inside e/h-convergence or i-e/h-convergence) of $\Phi^k$ to $\Phi$ means that the above (a) and (b) without the infinity conditions when $x \notin A$ and $y \notin B$ are satisfied.

Observe the crutial feature of epi/hypo-convergence and inside epi/hypo-convergence that they are completely symmetric (between $x$ and $y$, liminf and limsup, $\geq$ and $\leq$, and $+\infty$ and $-\infty$). That is why this convergence is effective for considerations of duality properties in optimization-related problems. The other basic types of variational convergence of bifunctions, minsup-lop-convergence and maxinf-lop-convergence, are non-symmetric (see [7, 8, 11]) and difficult for such consideration and we are not concerned in this paper, as we are interested in dual problems in nonlinear programming.

## 3. Variational approximations of a dual pair of problems in mathematical programming

By variational approximations of a problem, we mean a sequence of problems whose convergence (in some sense) preserves variational properties of the problems in the sequence for its limits. In this paper, we show that inside e/h-convergence and e/h-convergence guarantee this preservation for both problems in a dual pairs. We also call this sequence a sequence of approximating problems and call the primal problem the original problem. Here, variational properties mean those about saddle points, weak saddle points, minsup-points, sup-projections, optimal solutions, optimal values of objectives, etc. We call any convergence with such a preservation properties a variational convergence. So, this is a common terminology for types of convergence, not an exactly-defined type of convergence, and e/h-convergence and i-e/h-convergence are special cases of vaiational convergence, which are suitable for variational approximations of both problems in a dual pair of nonlinear programming problems. Note that, as mentioned in Section 1, basic types of variational convergence of bifunctions are epi-hypo-convergence, minsup-lop-convergence, and maxinf-lop-convergence. But, the two types of lopsided convergence are not symmetric and not suitable for considerations of duality properties in optimization. In this paper, we study variational approximations of dual problems in mathematical programming via variational convergence of Lagrange functions $\mathcal{L}^k$ of approximating problems. We need first the following definition.

**Definition 3.1.** (inside epi/hypo-inferior limit) The inside epi/hypo-inferior limit of a sequence of Lagrangians $\mathcal{L}^k$ (at $(x,(\mu,\nu)) \in A \times B$) is

$$\text{(i-e/h-li)}\mathcal{L}^k(x,(\mu,\nu)) := \\ \sup_{\{(\mu^{k_j},\nu^{k_j})\in B^{k_j}\to(\mu,\nu)\}}\inf_{\{x^{k_j}\in A^{k_j}\to x\}}\text{ls}_j\mathcal{L}^{k_j}(x^{k_j},(\mu^{k_j},\nu^{k_j})).$$

**Theorem 3.2.** (convergence of approximate saddle points) *Assume that* $\mathcal{L}^k \overset{i-e/h}{\to} \mathcal{L}$, $\varepsilon^k \searrow \varepsilon \geq 0$, $(\bar{x}^k,(\bar{\mu}^k,\bar{\nu}^k))$ *is an* $\varepsilon^k$*-saddle point of* $\mathcal{L}^k$ *for* $k \in \mathbb{N}$, *and* $(\bar{x},(\bar{\mu},\bar{\nu}))$ *belongs to* $A\times B$ *and is a cluster point of this sequence, i.e., for some sequence* $N \subset \mathbb{N}$, $(\bar{x},(\bar{\mu},\bar{\nu})) = \lim_{k\in N}(\bar{x}^k,(\bar{\mu}^k,\bar{\nu}^k))$. *Then,* $(\bar{x},(\bar{\mu},\bar{\nu}))$ *is an* $\varepsilon$*-saddle point of* $\mathcal{L}$ *and* $\mathcal{L}(\bar{x},(\bar{\mu},\bar{\nu})) = \lim_{k\in N}\mathcal{L}^k(\bar{x}^k,(\bar{\mu}^k,\bar{\nu}^k))$.

*Proof* We can assume that $N = \mathbb{N}$ for simplicity of notations. For any $(x,(\mu,\nu)) \in A \times B$, all sequences $(x^k,(\mu^k,\nu^k)) \in A^k \times B^k \to (x,(\mu,\nu))$ satisfy the $\varepsilon^k$-saddle inequalities

$$\mathcal{L}^k(\bar{x}^k,(\mu^k,\nu^k)) - \varepsilon^k \leq \mathcal{L}^k(\bar{x}^k,(\bar{\mu}^k,\bar{\nu}^k)) \leq \mathcal{L}^k(x^k,(\bar{\mu}^k,\bar{\nu}^k)) + \varepsilon^k.$$

Hence,

$$\sup_{\{(\mu^k,\nu^k)\in B^k\to(\mu,\nu))\}}\text{li}_k(\mathcal{L}^k(\bar{x}^k,(\mu^k,\nu^k)) - \varepsilon^k) \leq \text{li}_k\mathcal{L}^k(\bar{x}^k,(\bar{\mu}^k,\bar{\nu}^k)) \\ \leq \text{ls}_k\mathcal{L}^k(\bar{x}^k,(\bar{\mu}^k,\bar{\nu}^k)) \leq \inf_{\{x^k\in A^k\to x\}}\text{ls}_k(\mathcal{L}^k(x^k,(\bar{\mu}^k,\bar{\nu}^k)) + \varepsilon^k).$$

By the definition of i-e/h-convergence, $(\bar{x},\bar{y}) \in \varepsilon\text{-sdl}\mathcal{L}$ as, for all $(x,(\mu,\nu)) \in A \times B$,

$$\mathcal{L}(\bar{x},(\mu,\nu)) - \varepsilon \leq \sup_{\{(\mu^k,\nu^k)\in B^k\to(\mu,\nu)\}}\text{li}_k(\mathcal{L}^k(\bar{x}^k,(\mu^k,\nu^k)) - \varepsilon^k) \\ \leq \inf_{\{x^k\in A^k\to x\}}\text{ls}_k(\mathcal{L}^k(x^k,(\bar{\mu}^k,\bar{\nu}^k)) + \varepsilon^k) \leq \mathcal{L}(x,(\bar{\mu},\bar{\nu})) + \varepsilon.$$

To check that $\mathcal{L}(\bar{x},(\bar{\mu},\bar{\nu})) = \lim_{k\in N}\mathcal{L}^k(\bar{x}^k,(\bar{\mu}^k,\bar{\nu}^k))$, simply observe that the i-e/h-convergence and $\bar{x}^k \to \bar{x}$ ensure the existence of $(\mu^k,\nu^k) \in B^k \to (\mu,\nu)$ with

$$\mathcal{L}(\bar{x},(\bar{\mu},\bar{\nu})) \leq \text{li}_k\mathcal{L}^k(\bar{x}^k,(\mu^k,\nu^k)) \leq \text{li}_k(\mathcal{L}^k(\bar{x}^k,(\bar{\mu}^k,\bar{\nu}^k)) + \varepsilon^k) \\ = \text{li}_k\mathcal{L}^k(\bar{x}^k,(\bar{\mu}^k,\bar{\nu}^k)).$$

A similar argument gives $\mathcal{L}(\bar{x},(\bar{\mu},\bar{\nu})) \geq \text{ls}_k\mathcal{L}^k(\bar{x}^k,(\bar{\mu}^k,\bar{\nu}^k))$. □

**Proposition 3.3.** (relations between sdl$\mathcal{L}$ and Sol(NP), Sol(DNP)) $(\bar{x}, (\bar{\mu}, \bar{\nu})) \in \mathrm{sdl}\mathcal{L}$ *if* $\bar{x} \in \mathrm{Sol(NP)}$ *and* $(\bar{\mu}, \bar{\nu}) \in \mathrm{Sol(DNP)}$*, the two optimal values are equal. Conversely,* $(\bar{x}, (\bar{\mu}, \bar{\nu})) \in \mathrm{sdl}\mathcal{L}$ *implies the latter fact and also* $\langle \bar{\mu}, g(\bar{x})\rangle = 0$.

*Proof.* $\bar{x} \in \mathrm{Sol(NP)}$, i.e., $\bar{x} \in \mathrm{mis}\mathcal{L}$, means

$$\varphi(\bar{x}) + \sup_{\mu \in C^*} \langle \mu, g(\bar{x})\rangle = \inf_{x \in A}(\varphi(x) + \langle \bar{\mu}, g(x)\rangle).$$

As $\max_{\mu \in C^*} \langle \mu, g(\bar{x})\rangle = \langle \bar{\mu}, g(\bar{x})\rangle = 0$, substituting $\bar{\mu}$ into the left-hand side of the above equality, we have both saddle inequalities.

Conversely, as $(\bar{x}, (\bar{\mu}, \bar{\nu})) \in \mathrm{mima}\mathcal{L} \cap \mathrm{mami}\mathcal{L}$, $\bar{x} \in \mathrm{mis}\mathcal{L}$ and $(\bar{\mu}, \bar{\nu}) \in \mathrm{mai}\mathcal{L}$, and so is a solution of (NP) and (DNP), resp. Clearly the two optimal values being the minsup-value and maxinf-value, resp, are equal to the saddle value $\mathcal{L}(\bar{x}, (\bar{\mu}, \bar{\nu}))$. Moreover, the left-hand saddle inequality $\varphi(\bar{x}) + \langle \mu, g(\bar{x})\rangle \leq \varphi(\bar{x}) + \langle \bar{\mu}, g(\bar{x})\rangle$ for all $\mu \in C^*$ shows that $\langle \bar{\mu}, g(\bar{x})\rangle$ must be zero. □

**Corollary 3.4.** *Assume that* $\mathcal{L}^k \overset{i-e/h}{\to} \mathcal{L}$*,* $\varepsilon^k \searrow 0$*,* $(\bar{x}^k, (\bar{\mu}^k, \bar{\nu}^k))$ *is an* $\varepsilon^k$*-saddle point of* $\mathcal{L}^k$ *for* $k \in \mathbb{N}$*, and* $(\bar{x}, (\bar{\mu}, \bar{\nu}))$ *belongs to* $A \times B$ *and is a cluster point of this sequence. Then,* $\bar{x} \in \mathrm{Sol(NP)}$ *and* $(\bar{\mu}, \bar{\nu}) \in \mathrm{Sol(DNP)}$*, the two optimal values are equal, and* $\langle \bar{\mu}, g(\bar{x})\rangle = 0$.

*Proof.* Simply apply Theorem 3.2 and Proposition 3.3. □

Clearly, taking $\varepsilon^k \equiv 0$ in the above result yields a sufficient condition for the convergence of (exact) saddle points.

In the following consequence of Theorem 3.2 and Corollary 3.4, we obtain solutions of (NP) and (DNP) from limits of $\varepsilon^k$-sdl$\mathcal{L}^k$, but using assumptions direct on the data of the mathematical programming problem.

**Theorem 3.5.** (convergence of approximate saddle points under assumptions on the data of (NP)) *Assume that* $\varepsilon^k \searrow 0$*,* $(\bar{x}^k, (\bar{\mu}^k, \bar{\nu}^k))$ *is an* $\varepsilon^k$*-saddle point of* $\mathcal{L}^k$ *for* $k \in \mathbb{N}$*, and* $(\bar{x}, (\bar{\mu}, \bar{\nu})) \in A \times B$ *is a cluster point of this sequence. Assume further that*

(i) $\varphi^k$*, and for all* $(\mu, \nu) \in B$*,* $\langle \mu, g^k\rangle$ *and* $\langle \nu, h^k\rangle$ *satisfy condition* (a) *of inside epi-convergence of* $\varphi^k$*,* $\langle \mu, g^k\rangle$*, and* $\langle \nu, h^k\rangle$*, resp;*

(ii) $\varphi^k$ *satisfies* (b) *of i-epi-convergence and the bifunctions* $(x, \mu) \mapsto \langle \mu, g^k(x)\rangle$ *and* $(x, \nu) \mapsto \langle \nu, h^k(x)\rangle$ *satisfy* (b) *of i-e/h-convergence.*

*Then,* $(\bar{x}, (\bar{\mu}, \bar{\nu}))$ *is a saddle point of* $\mathcal{L}$*, and so* $\bar{x}$ *is a solution of* (NP)*,* $(\bar{\mu}, \bar{\nu})$ *is a solution of* (DNP) *and* $\langle \bar{\mu}, g(\bar{x})\rangle = 0$.

*Proof* In view of Theorem 3.2 and Corollary 3.4, we only need to verify the i-e/h-convergence of $\mathcal{L}^k$. First, we check condition (a) of i-e/h-convergence. For any $(\mu, \nu) \in B$ and $x^{k_j} \in A^{k_j} \to x \in A$, we take $(\mu^{k_j}, \nu^{k_j}) \equiv (\mu, \nu)$. Then, by (i), one has (a) of the indide e/h-convergence of $\mathcal{L}^k$:

$$\mathrm{li}_j(\mathcal{L}^{k_j}(x^{k_j}, \mu^{k_j}, \nu^{k_j}) - \mathcal{L}(x, \mu, \nu)) \geq \mathrm{li}_j \varphi^{k_j}(x^{k_j}) - \varphi(x)$$

$$+\mathrm{li}_j \langle \mu, g^{k_j}(x^{k_j}) - g(x)\rangle + \mathrm{li}_j \langle \nu, h^{k_j}(x^{k_j}) - h(x)\rangle \geq 0.$$

Using assumption (ii), we verify (b) of i-e/h-convergence of $\mathcal{L}^k$: for all $(\mu^{k_j}, \nu^{k_j}) \in B^{k_j} \to (\mu, \nu) \in B$ and $x \in A$, there exist $x^{k_j} \to x$ such that

$$\begin{aligned}\mathrm{ls}_j[\varphi^{k_j}(x^{k_j}) + \langle \mu^{k_j}, g^{k_j}(x^{k_j})\rangle + \langle \nu^{k_j}, h^{k_j}(x^{k_j})\rangle] &\le \mathrm{ls}_j\varphi^{k_j}(x^{k_j}) \\ &+\mathrm{ls}_j\langle \mu^{k_j}, g^{k_j}(x^{k_j})\rangle + \mathrm{ls}_j\langle \nu^{k_j}, h^{k_j}(x^{k_j})\rangle \\ &\le \varphi(x) + \langle \mu, g(x)\rangle + \langle \nu, h(x)\rangle.\end{aligned}$$

□

Although the assumed i-e/h-convergence in the above results is a weak condition, the existence of approximate saddle points of $\mathcal{L}^k$ is a restrictive condition, which means, roughly speaking that for all $k$, the pairs of dual problems (NP$^k$) and (DNP$^k$) satisfy an approximate strong duality. However, in many cases, only one of the two problems, say (NP$^k$) for instance, have approximate solutions, i.e., only the set approximate mis$\mathcal{L}^k$ are nonempty and we need their convergence to mis$\mathcal{L}$. Then, we can apply also e/h-convergence, but we need the following additional notions and facts as follows. We also need to recall the equivalent geometric formulation of epi-convergence (see, e.g., [1]): for a metric space $X$ (as in Section 2), $\varphi^k : X \to \mathbb{R}$ epi-converge to $\varphi^k : X \to \mathbb{R}$ if and only if epi$\varphi^k$ converge to epi$\varphi$ in the Painlevé-Kuratowski set convergence. (This is just the origin of the terminology "epi-convergence".) The corresponding statement for hypo-convergence is obtained by replacing "epi-" by "hypo-".

Now proceed to the above issue on mis$\mathcal{L}$. A function $\zeta : A \to \mathbb{R} \cup \{+\infty\}$ ($\eta : B \to \mathbb{R} \cup \{-\infty\}$, resp) is called the sup-projection (inf-projection, resp) of the Lagrange function $\mathcal{L}$ if, for $x \in A$ ($(\mu, \nu) \in B$, resp),

$$\zeta(x) := \sup_{(\mu,\nu)\in B}\mathcal{L}(x,(\mu,\nu)) \ (\eta(\mu,\nu) := \inf_{x\in A}\mathcal{L}(x,(\mu,\nu)), \text{ resp}).$$

Note that $\bar{x} \in \varepsilon$-argmin$\zeta$ if and only if $\bar{x}$ is an $\varepsilon$-minsup point of $\mathcal{L}$ and $(\bar{\mu}, \bar{\nu}) \in \varepsilon$-argmax$\eta$ if and only if $(\bar{\mu}, \bar{\nu})$ is an $\varepsilon$-maxinf point of $\mathcal{L}$.

**Definition 3.6.** (a) ($x$-ancillary e/h-convergence) $\mathcal{L}^k$ is called $x$-ancillary e/h-convergent to $\mathcal{L}$ if (a) of the e/h-convergence of $\mathcal{L}^k$ and the following condition is satisfied

(b′) $\forall x \in A$, $\forall (\mu^k, \nu^k) \in B^k \to (\mu, \nu) \in B$, $\exists x^k \in A^k \to x$, $\mathrm{ls}_k\zeta^k(x^k) \le \zeta(x)$.

(b) (($\mu, \nu$)-ancillary e/h-convergence) $\mathcal{L}^k$ is called $(\mu, \nu)$-ancillary e/h-convergent to $\mathcal{L}$ if (b) of the e/h-convergence is fulfilled together with the condition

(a′) $\forall (\mu, \nu) \in B$, $\forall x^k \in A^k \to x$, $\exists (\mu^k, \nu^k) \in B^k \to (\mu, \nu)$, $\mathrm{li}_k\eta^k(\mu^k, \nu^k) \ge \eta(\mu, \nu)$.

Observe that the two parts in Definition 3.6 are also symmetric (like in the definition of e/h-convergence) and not (directly) connected with compactness, while the known tightness conditions for minsup-lop-convergence and maxinf-lop-convergence represent types of relaxed uniform compactness conditions, see [7, 11]. The above convergence results show that i-e/h-convergence is useful for considering saddle points and duality (while quite different, tight minsup-lop-convergence is appropriate only for minsup-points and maxinf-lop-convergence is defined separately and applied to maxinf-points, see [7, 8, 11]). Even when both minsup-lop-convergence and maxinf-lop-convergence are tight (a strong assumption), we hardly have conclusions about convergence of saddle points or approximate saddle points. We will now see that applying tight e/h-convergence, we obtain also convergence results for the nonsymmetric objects minsup- and maxinf-points.

**Theorem 3.7.** (convergence of minsup-points). *Let the e/h-convergence of $\mathcal{L}^k$ to $\mathcal{L}$ be x-ancillary tight and the domains of $\zeta^k$, $\zeta$ be nonempty (except possibly for a finite number of k). Then,*

(i) $\text{e-lim}_k \zeta^k = \zeta$;

(ii)

$$\text{ls}_k[\inf_{x\in A^k}\sup_{(\mu,\nu)\in B^k}\mathcal{L}^k(x,(\mu,\nu))] \le \inf_{x\in A}\sup_{(\mu,\nu)\in B}\mathcal{L}(x,(\mu,\nu)).$$

*Moreover, if for a subsequence $\{k_j\}_j$ one has $x^{k_j}\in \text{mis}\mathcal{L}^{k_j}$ and $\lim_j x^{k_j} = \bar{x}$, then $\bar{x}\in\text{mis}\mathcal{L}$ (i.e., $\text{Ls}_k\text{mis}\mathcal{L}^k\subset\text{mis}\mathcal{L}$) and*

$$\lim_j(\sup_{(\mu,\nu)\in B^{k_j}}\mathcal{L}^{k_j}(x^{k_j},(\mu,\nu))) = \sup_{(\mu,\nu)\in B}\mathcal{L}(\bar{x},(\mu,\nu)),$$

*that is, the minsup-values of $\mathcal{L}^{k_j}$ converge to that of $\mathcal{L}$.*

*Proof* (i) Applying the geometric formulation of epi-convergence, we show first $\text{Ls}_k(\text{epi}\zeta^k)\subset\text{epi}\zeta$. Take any $(x,\alpha)$ in the left-hand side. Then, there exist $(x^{k_j},\alpha^{k_j})\in\text{epi}\zeta^{k_j}$ conversing to $(x,\alpha)$. We claim that $x\in\text{dom}\zeta$. Clearly, $x$ must be in $A$, for otherwise condition (a) of Definition 3.6 yields $(\mu^{k_j},\nu^{k_j})\in B^{k_j}$ such that $\mathcal{L}^{k_j}(x^{k_j},(\mu^{k_j},\nu^{k_j}))\to+\infty$ contradicting the fact that $\alpha^{k_j}\ge\sup_{B^{k_j}}\mathcal{L}^{k_j}(x^{k_j},\cdot)$ and $\alpha^{k_j}\to\alpha$. Suppose $x\in A\backslash\text{dom}\zeta$. Then, for any $\gamma>\alpha$, there exists $(\mu_\gamma,\nu_\gamma)\in B$ with $\mathcal{L}(x,(\mu_\gamma,\nu_\gamma))>\gamma$. In view of the aforementioned (a), there exist $(\mu_\gamma^{k_j},\nu_\gamma^{k_j})\in B^{k_j}\to(\mu_\gamma,\nu_\gamma)$ such that $\text{li}_j\mathcal{L}^{k_j}(x^{k_j},(\mu_\gamma^{k_j},\nu_\gamma^{k_j}))\ge\mathcal{L}(x,(\mu_\gamma,\nu_\gamma))$. So, we arrive at the contradiction

$$\alpha=\lim_j\alpha^{k_j}\ge\text{li}_j\zeta^{k_j}(x^{k_j})\ge\text{li}_j\mathcal{L}^{k_j}(x^{k_j},(\mu_\gamma^{k_j},\nu_\gamma^{k_j}))\ge\mathcal{L}(x,(\mu_\gamma,\nu_\gamma))>\gamma>\alpha.$$

Therefore, the claim $x\in\text{dom}\zeta$ is proved.

For any $\varepsilon>0$, there exists $(\mu_\varepsilon,\nu_\varepsilon)\in B$ with $\mathcal{L}(x,(\mu_\varepsilon,\nu_\varepsilon))\ge\zeta(x)-\varepsilon$. The condition (a) again yields $(\mu_\varepsilon^{k_j},\nu_\varepsilon^{k_j})\in B^{k_j}\to(\mu_\varepsilon,\nu_\varepsilon)$ such that

$$\text{li}_j\zeta^{k_j}(x^{k_j})\ge\text{li}_j\mathcal{L}^{k_j}(x^{k_j},(\mu_\varepsilon^{k_j},\nu_\varepsilon^{k_j}))\ge\mathcal{L}(x,(\mu_\varepsilon,\nu_\varepsilon))\ge\zeta(x)-\varepsilon.$$

Therefore, by the arbitrariness of $\varepsilon$, $\text{Ls}_k(\text{epi}\zeta^k)\subset\text{epi}\zeta$ because

$$\alpha=\lim_j\alpha^{k_j}\ge\text{li}_j\zeta^{k_j}(x^{k_j})\ge\zeta(x).$$

Next, we prove the inclusion $\text{epi}\zeta\subset\text{Li}_k(\text{epi}\zeta^k)$. By Definition 3.6, for any $(x,\alpha)\in\text{epi}\zeta$ and $(\mu^k,\nu^k)\in B^k\to(\mu,\nu)\in B$, there exist $x^k\in A^k$ such that $\text{ls}_k\zeta(x^k)\le\zeta(x)$. We find $(x_k,\gamma^k)\in\text{epi}\zeta^k$ as follows. For $k$ with $\zeta^k(x^k)\le\alpha$, take $\gamma^k=\alpha$. Consider $k$ with $\zeta^k(x^k)>\alpha$. As $\text{ls}_k\zeta^k(x^k)\le\zeta(x)$ by (b′), taking arbitrarily $\varepsilon^k\searrow 0$, one has $\zeta^k(x^k)\le\zeta(x)+\varepsilon^k$ for large $k$. Then,

$\zeta(x) \leq \alpha < \zeta^k(x^k) \leq \zeta(x)+\varepsilon^k$ with the last side tending to $\zeta(x)$. Finally, for $k$ with $\zeta^k(x^k) > \alpha$, choosing $\gamma^k = \zeta^k(x^k)$, one has $(x_k, \gamma^k) \in \text{epi}\zeta^k$ tending to $(x, \alpha)$.

(ii) Applying assertion (i), and the properties of epi-convergence recalled in Section 2, one obtains this assertion.

□

**Corollary 3.8.** (convergence of maxinf-points) *Let the e/h-convergence of $\mathcal{L}^k$ to $\mathcal{L}$ be $(\mu, \nu)$-ancillary tight and the domains of $\eta^k$, $\eta$ be nonempty (except possibly for a finite number of k). Then,*

(i) h-$\lim_k \eta^k = \eta$*;*

(ii)

$$\text{li}_k[\sup_{(\mu,\nu)\in B^k} \inf_{x\in A^k} \mathcal{L}^k(x,(\mu,\nu))] \geq \sup_{(\mu,\nu)\in B} \inf_{x\in A} \mathcal{L}(x,(\mu,\nu)).$$

*Moreover, if for a subsequence $\{k_j\}_j$ one has $(\mu^{k_j}, \nu^{k_j}) \in \text{mai}\mathcal{L}^{k_j}$ and $\lim_j(\mu^{k_j}, \nu^{k_j}) = (\bar{\mu}, \bar{\nu})$, then $(\bar{\mu}, \bar{\nu}) \in \text{mai}\mathcal{L}$ (i.e., $\text{Ls}_k \text{mai}\mathcal{L}^k \subset \text{mai}\mathcal{L}$) and*

$$\lim_j(\inf_{x\in A^{k_j}} \mathcal{L}^{k_j}(x,(\mu^{k_j}, \nu^{k_j})) = \inf_{x\in A} \mathcal{L}(x,(\bar{\mu}, \bar{\nu})),$$

*that is, the maxinf-values of $\mathcal{L}^{k_j}$ converge to that of $\mathcal{L}$.*

## References


[1] Attouch, H.: Variational Convergence for Functions and Operators. Applicable Mathematics Series, Pitman, 1984.

[2] Attouch, H., Wets, R.J-B: A convergence theory for saddle functions. Trans. Amer. Math. Soc. **280**, 1–41, 1983.

[3] Blum, E., Oettli, W.: From optimization and variational inequalities to equilibrium problems. Math. Student. **63**, 123-145, 1994.

[4] Fan, K.: A minimax inequality and applications. In: Shisha, O. (ed.) Inequality III, pp. 103113. Academic Press, New York, 1972.

[5] Diem, H.T.H., Khanh, P.Q.: Approximations of optimization-related problems in terms of variational convergence. Vietnam J. Math. **44**, 399–417, 2016.

[6] Diem, H.T.H., Khanh, P.Q.: Epi/Hypo-convergence of bifunctions on general domains and approximations of quasi-variational problems. Set-valued Var. Anal. Onlinefirst, 2019.

[7] Jofré, A., Wets, R.J-B: Variational convergence of bivariate functions: lopsided convergence. Math. Program. B **116**, 275–295, 2009.

[8] Lopez, R.: Approximations of equilibrium problems. SIAM J. Control Optim. **50**, 1038–1070, 2012.

[9] Nikaido, H., Isoda, K.: Note on noncooperative convex games. Pac. J. Math. **5** (Suppl. 1), 807–815, 1955.

[10] Rockafellar, R.T., Wets, R.J-B: Variational Analysis. Springer, Berlin, 3rd printing edition, 2009.

[11] Royset, J.O., Wets, J.B.R.: Lopsided convergence: An extention and its quantification. Math. Program. A **177**, 395–423, 2019.

[12] Wijsman, R.A.: Convergence of sequences of convex sets, cones and functions. Bull. Amer. Math. Soc. **70**, 186-188, 1964.

(Hochiminh City University of Technology) Department of Applied Mathematics, University of Technology, Ho Chi Minh City, Vietnam;
Vietnam National University, Ho Chi Minh City, Vietnam.
*Email address*: hthdiem@hcmut.edu.vn