J Optim Theory Appl (2014) 161:738–762
DOI 10.1007/s10957-013-0468-4

Unified Duality Theory for Constrained Extremum Problems. Part I: Image Space Analysis

S.K. Zhu · S.J. Li

Received: 22 September 2013 / Accepted: 26 October 2013 / Published online: 9 November 2013

Abstract This paper is concerned with a unified duality theory for a constrained extremum problem. Following along with the image space analysis, a unified duality scheme for a constrained extremum problem is proposed by virtue of the class of regular weak separation functions in the image space. Some equivalent characterizations of the zero duality property are obtained under an appropriate assumption. Moreover, some necessary and sufficient conditions for the zero duality property are also established in terms of the perturbation function. In the accompanying paper, the Lagrange-type duality, Wolfe duality and Mond–Weir duality will be discussed as special duality schemes in a unified interpretation. Simultaneously, three practical classes of regular weak separation functions will be also considered.

Keywords Image space analysis · Constrained extremum problem · Separation function · Lagrange-type duality · Perturbation function

1 Introduction

This two-part paper series is concerned with the duality theory of a constrained extremum problem. We particularly aim to establish a duality problem and study some of its relationships to the primal problem by means of the image space analysis (for short, ISA).

In the last decades, the ISA for constrained extremum problems has been of great interest in the academic and professional communities. The ISA was initiated in [1] and extensively used as a preliminary and auxiliary step for investigating some topics

S.K. Zhu · S.J. Li (✉)
College of Mathematics and Statistics, Chongqing University, Chongqing 401331, China
e-mail: lisj@cqu.edu.cn

S.K. Zhu
e-mail: zskcqu@163.com

of the optimization theory, such as optimality conditions [2–9], existence of solutions [10], dualities [11–13], variational principles [4, 14], penalty methods [3, 9, 15, 16], regularities and stabilities [17, 18], and so on.

Since the path-breaking paper [2], it has been shown that the ISA is a unified scheme for studying any kind of problem, which can be casted into the form of the impossibility of a parametric system like, for example, constrained extremum problems, variational inequalities, equilibrium problems, and so on. Moreover, many papers have been devoted to extend the ISA to investigate the constrained vector optimization problem, the vector variational inequality and the vector equilibrium problem. In [19], several theoretical aspects of vector optimization problems and vector variational inequalities, like optimality conditions, dualities, penalty methods as well as scalarization, were exploited based on the ISA. Moreover, the ISA was generalized to a vector quasi-equilibrium problem with a variable ordering relation in [20], and some scalar and vector saddle point optimality conditions, arising from the existence of a vector separation in the corresponding image space (for short, IS), were established. For more details, we refer to [21–23] and references therein.

As we know, the Lagrangian duality is an important method in constrained optimization theory. Among several crucial aspects, the zero duality gap property between the primal and dual problems plays a key role. In the past few years, the classic Lagrangian duality theory has been extensively generalized by various kinds of Lagrangian-type functions, especially the augmented Lagrangian function, introduced by Hestenes [24, 25] and Powell [26], and the nonlinear Lagrangian function, proposed by Rubinov et al. [27]. In [28], Yang and Huang established an equivalence between two types of zero duality gap properties, which were described using augmented Lagrangian dual functions and nonlinear Lagrangian dual functions. Rubinov et al. [29] examined the validity of the zero duality gap properties for two important dual schemes, and obtained some necessary and sufficient conditions for the zero duality gap property in terms of the lower semicontinuity of the perturbation functions. In [30], Wang et al. presented a unified nonlinear Lagrangian dual scheme, and established necessary and sufficient conditions for the zero duality gap property. They also derived necessary and sufficient conditions for four classes of zero duality gap properties, and obtained the equivalence among them.

Motivated by the work reported in [2, 4–6, 11, 12, 30, 31], the purpose of this paper is to establish a unified duality scheme by virtue of the ISA. To this end, we first propose a generalized Lagrange function for a constrained extremum problem based on the class of regular weak separation functions in the IS and then, establish a general dual problem. Under an appropriate assumption, we obtain some equivalent characterizations to the zero duality gap property. In addition, we get some necessary and sufficient conditions for the zero duality gap property in the form of lower semicontinuity of the perturbation function. In the accompanying paper (Part II), we will discuss the Lagrange-type duality, Wolfe duality and Mond–Weir duality in a unified interpretation. As applications, we will particularly consider three practical classes of regular weak separation functions, which are the separable function in [4], the augmented Lagrangian function in Rockafellar and Wets [32], and the nonlinear Lagrangian function in Rubinov et al. [27].

The organization of this paper is as follows. In Sect. 2, we recall some preliminaries, especially the concept of separation functions in the ISA. In Sect. 3, we establish a unified duality scheme for a constrained extremum problem based on the class of regular weak separation functions in the IS, and study the zero duality gap property.

2 Preliminaries and Separation Functions in the ISA

In this paper, we consider the following constrained extremum problem:

$$\text{(P)} \quad \min f(x), \quad \text{s.t.} \quad x \in X, \quad g_i(x) \geq 0, \quad i = 1, 2, \dots, m,$$

where X is a metric space, and $f : X \to \mathbb{R}$ and $g_i : X \to \mathbb{R}, i = 1, 2, \dots, m$, are real-valued functions. As usual, we denote the vector-valued function $g : X \to \mathbb{R}^m$ by $g(x) := (g_1(x), g_2(x), \dots, g_m(x))$ for all $x \in X$, and the feasible set by $R := \{x \in X \mid g_i(x) \geq 0, i = 1, 2, \dots, m\}$. Throughout this paper, we assume that $R \neq \varnothing$. Take arbitrary $\bar{x} \in X$, we consider the map $A : X \to \mathbb{R}^{1+m}$ with

$$A(x) := \big(f(\bar{x}) - f(x), g(x)\big), \quad \forall x \in X,$$

and the sets

$$\mathcal{K} := \big\{(u, v) \in \mathbb{R}^{1+m} \mid u = f(\bar{x}) - f(x), v = g(x), x \in X\big\},$$
$$\mathcal{H} := \big\{(u, v) \in \mathbb{R}^{1+m} \mid u > 0, v_i \geq 0, i = 1, 2, \dots, m\big\}.$$

For simplicity, we denote $\mathcal{H}$ by $\mathbb{R}_{++} \times \mathbb{R}^m_+$. Obviously, $\mathcal{K}$ is the image of the map A, i.e., $\mathcal{K} = A(X)$. We recall from [4, 11, 12] that the space $\mathbb{R}^{1+m}$ is said to be the *image space* associated with (P) and $\mathcal{K}$ is said to be the *image* of (P). It is worth noting that the image $\mathcal{K}$ is not generally convex even though the functions involved enjoy some convexity properties. To overcome this difficulty, a regularization of the image $\mathcal{K}$, namely the extension with respect to the cone $\operatorname{cl}\mathcal{H}$, denoted by

$$\begin{aligned} \mathcal{E} &:= \mathcal{K} - \operatorname{cl}\mathcal{H} \\ &= \big\{(u, v) \in \mathbb{R}^{1+m} \mid u \leq f(\bar{x}) - f(x), v_i \leq g_i(x), i = 1, 2, \dots, m, x \in X\big\} \end{aligned}$$

was introduced in [4]. As a result, the convexity of the regularization image $\mathcal{E}$ can be verified under some appropriate conditions; for example, A is $-(\operatorname{cl}\mathcal{H})$-convexlike. We refer to [4, 7–12] and references therein for more details.

Remark 2.1 It is worth mentioning that the choice of $\bar{x}$ is arbitrary. Based on this fact, we use the notation A and $\mathcal{K}$, although they seem to be dependent on the point $\bar{x}$. Just as shown in [4], even if it is a mere formal question, it is not convenient, and it is better to define u as above in $\mathcal{K}$; see more details in [4, 11, 12]. Note that $\mathcal{H} + \operatorname{cl}\mathcal{H} = \mathcal{H}$. Take especially $\bar{x} \in R$. Then, it is easy to verify that $\bar{x}$ is a global minimum point for (P) if and only if

$$\mathcal{K} \cap \mathcal{H} = \varnothing,$$

or equivalently,

$$\mathcal{E} \cap \mathcal{H} = \varnothing.$$

Throughout this paper, the choice of $\bar{x}$ is arbitrary unless otherwise specified.

Consider a function $w : \mathbb{R}^{1+m} \times \Pi \to \mathbb{R}$, where Π is a set of parameters to be specified case by case. For every $\pi \in \Pi$, the non-negative level set and the positive level set of the function $w(\bullet; \pi) : \mathbb{R}^{1+m} \to \mathbb{R}$ are respectively defined by

$$\mathrm{lev}_{\geq 0} w(\bullet; \pi) := \left\{(u, v) \in \mathbb{R}^{1+m} \mid w(u, v; \pi) \geq 0\right\}$$

and

$$\mathrm{lev}_{> 0} w(\bullet; \pi) := \left\{(u, v) \in \mathbb{R}^{1+m} \mid w(u, v; \pi) > 0\right\}.$$

In the sequel, for the sake of simplicity, we will always use the same symbol Π to denote the set of parameters, even though the family of parameters is not necessarily the same.

Definition 2.1 (See [4]) The class of all the functions $w : \mathbb{R}^{1+m} \times \Pi \to \mathbb{R}$ such that

$$\mathcal{H} \subset \mathrm{lev}_{\geq 0} w(\bullet; \pi), \quad \forall \pi \in \Pi \tag{1}$$

and

$$\bigcap_{\pi \in \Pi} \mathrm{lev}_{> 0} w(\bullet; \pi) \subset \mathcal{H} \tag{2}$$

is called the class of weak separation functions, and is denoted by $\mathcal{W}(\Pi)$.

Since the class $\mathcal{W}(\Pi)$ is too large, another subclass is introduced as follows by strengthening Definition 2.1.

Definition 2.2 (See [4]) The class of all the functions $w : \mathbb{R}^{1+m} \times \Pi \to \mathbb{R}$ such that

$$\bigcap_{\pi \in \Pi} \mathrm{lev}_{> 0} w(\bullet; \pi) = \mathcal{H} \tag{3}$$

is called the class of regular weak separation functions, and is denoted by $\mathcal{W}_R(\Pi)$.

It is worth noting that the main distinction between the classes $\mathcal{W}(\Pi)$ and $\mathcal{W}_R(\Pi)$ reflects different historical approaches, i.e., the Fritz–John and the Kuhn–Tucker approaches, to the use of the same basic scheme in optimization theories. We refer to [4, 6] for more details. In this paper, we will use the subclass $\mathcal{W}^{\ell s}(\Pi)$ consisting of all the elements in $\mathcal{W}(\Pi)$, such that the function $w(\bullet; \pi) : \mathbb{R}^{1+m} \to \mathbb{R}$ is lower semicontinuous for every $\pi \in \Pi$. Obviously, $\mathcal{W}_R(\Pi) \subset \mathcal{W}(\Pi)$ and $\mathcal{W}^{\ell s}(\Pi) \subset \mathcal{W}(\Pi)$. Throughout the paper, we set

$$\mathcal{W}_R^{\ell s}(\Pi) := \mathcal{W}_R(\Pi) \cap \mathcal{W}^{\ell s}(\Pi).$$

3 Unified Duality Scheme and Zero Duality Gap Property

Inspired by the ideas reported in [2, 4, 6, 29, 30, 33], the purpose of this section is to establish a unified duality scheme for (P) by virtue of the class $\mathcal{W}_R(\Pi)$, and to investigate the zero duality gap property.

Definition 3.1 Given the classes $\mathcal{W}(\Pi)$ and $\mathcal{W}_R(\Pi)$, the sets $\mathcal{K}$ and $\mathcal{H}$ admit a separation with respect to $w \in \mathcal{W}(\Pi)$ and $\hat{\pi} \in \Pi$ iff

$$w(u, v; \hat{\pi}) \leq 0, \quad \forall (u, v) \in \mathcal{K}. \tag{4}$$

Moreover, the separation is said to be regular iff $w \in \mathcal{W}_R(\Pi)$.

We observe that $\mathcal{K} \cap \mathcal{H} = \varnothing$ if either $\mathcal{K}$ and $\mathcal{H}$ admit a separation, and (4) is verified in strict sense (as it follows from (1) and (2)), or $\mathcal{K}$ and $\mathcal{H}$ admit a regular separation (as it follows from (3)). In addition, let $\bar{x} \in R$. Then it follows from Remark 2.1 that $\bar{x}$ is a global minimum point for (P).

3.1 A Unified Duality Via the Class $\mathcal{W}_R(\Pi)$

In this subsection, we follow the regular weak separation approach to establish a unified duality scheme for (P).

As shown in Definition 3.1, the regular separation of $\mathcal{K}$ and $\mathcal{H}$ in the IS is associated with a parameter $\hat{\pi} \in \Pi$, and moreover, gives a sufficient condition for $\bar{x}$ to be a global minimum point for (P) when $\bar{x} \in R$. Naturally, this motivates us to study the corresponding results of (P) in the primal space, that is, every $(u, v) \in \mathcal{K}$ turns to $u = f(\bar{x}) - f(x)$, $v = g(x)$ with $x \in X$. Given the class $\mathcal{W}_R(\Pi)$ and a regular weak separation function $w \in \mathcal{W}_R(\Pi)$, we consider the real-valued function $L_w : X \times \Pi \to \mathbb{R}$, defined by

$$L_w(x, \pi) := w(1, 0_{\mathbb{R}^m}; \pi) f(\bar{x}) - w\bigl(f(\bar{x}) - f(x), g(x); \pi\bigr), \quad \forall (x, \pi) \in X \times \Pi, \tag{5}$$

where $0_{\mathbb{R}^m}$ denotes the origin of $\mathbb{R}^m$. The function L_w will be called a generalized Lagrange function for (P) corresponding to w.

We observe that, for $\Pi = \mathbb{R}_+ \times \mathbb{R}^m_+$ and $w(u, v; \theta, \lambda) = \theta u + \langle \lambda, v \rangle$, for all $(u, v) \in \mathbb{R}^{1+m}$ and all $(\theta, \lambda) \in \Pi$, $L_w(x; \theta, \lambda) = \theta f(x) - \langle \lambda, g(x) \rangle$ collapses to the standard John function. In addition, if $\theta > 0$, especially, $\theta = 1$, then $L_w(x; 1, \lambda) = f(x) - \langle \lambda, g(x) \rangle$ collapses to the standard Lagrange function.

For every given $\pi \in \Pi$, consider the set

$$\mathcal{I}_\pi := \Bigl\{ \hat{x} \in X \mid L_w(\hat{x}, \pi) = \inf_{x \in X} L_w(x, \pi) \Bigr\}.$$

Let

$$\mathcal{M} := \bigl\{ (\hat{x}, \pi) \in X \times \Pi \mid \hat{x} \in \mathcal{I}_\pi \bigr\}. \tag{6}$$

Now, based on the general arguments of separations in the IS of (P), we consider the following dual problem:

$$\text{(DP)} \qquad \sup_{(x,\pi)\in\mathcal{Q}} L_w(x,\pi),$$

where $\mathcal{Q}$ is a nonempty subset of $\mathcal{M}$. For simplicity, we denote

$$\mathcal{Q}^\circ := \{\pi \in \Pi \mid \exists x \in X, \text{ s.t. } (x,\pi) \in \mathcal{Q}\}.$$

Moreover, if

$$\inf_{x\in R} f(x) = \sup_{(x,\pi)\in\mathcal{Q}} L_w(x,\pi),$$

then we say that the zero duality gap property with respect to w holds.

Analogously, for every given $x \in X$, consider the set

$$\mathcal{J}_x := \Big\{\hat{\pi} \in \Pi \mid L_w(x,\hat{\pi}) = \sup_{\pi\in\Pi} L_w(x,\pi)\Big\}.$$

Let

$$\mathcal{N} := \{(x,\hat{\pi}) \in X \times \Pi \mid \hat{\pi} \in \mathcal{J}_x\}. \tag{7}$$

We consider a new problem associated with (P) as follows:

$$(\widehat{\text{P}}) \qquad \inf_{(x,\pi)\in\mathcal{F}} L_w(x,\pi),$$

where $\mathcal{F}$ is a nonempty subset of $\mathcal{N}$. Similarly, we denote

$$\mathcal{F}^o := \{x \in X \mid \exists \pi \in \Pi, \text{s.t. } (x,\pi) \in \mathcal{F}\}.$$

Throughout this paper, we assume that $R \subset \mathcal{F}^o$. Note that this assumption is not restrictive for several major classes of generalized Lagrange-type functions, such as the Lagrange function $L_w(x;1,\lambda) := f(x) - \langle\lambda, g(x)\rangle, \forall x \in X, \forall\lambda \in \mathbb{R}^m_+$, and the augmented Lagrange function $L_w(x;1,\lambda) := f(x) - \sup_{z\in\{g(x)\}-\mathbb{R}^m_+}(\langle\lambda,z\rangle - r\sigma(z)), \forall x \in X, \forall\lambda \in \mathbb{R}^m_+$ (see more details in Sect. 4 of [34]). In fact, it is easy to verify that we have $\hat{\pi} = (1, 0_{\mathbb{R}^m}) \in \mathcal{J}_x$ for every $x \in R$, which implies $R \times \{(1, 0_{\mathbb{R}^m})\} \subset \mathcal{N}$. Thus, $R \subset \mathcal{F}^o$ holds if we take $R \times \{(1, 0_{\mathbb{R}^m})\} \subset \mathcal{F}$.

For simplicity, we denote

$$\alpha = \sup_{(x,\pi)\in\mathcal{Q}} L_w(x,\pi) \quad \text{and} \quad \beta = \inf_{(x,\pi)\in\mathcal{F}} L_w(x,\pi).$$

It is easy to verify that the inequality relationship $\alpha \le \beta$ always holds. In fact, we have from (6) and (7) that

$$\begin{aligned}\alpha = \sup_{(x,\pi)\in\mathcal{Q}} L_w(x,\pi) &= \sup_{\pi\in\mathcal{Q}^\circ}\inf_{x\in X} L_w(x,\pi)\\ &\le \sup_{\pi\in\Pi}\inf_{x\in\mathcal{F}^o} L_w(x,\pi)\\ &\le \inf_{x\in\mathcal{F}^o}\sup_{\pi\in\Pi} L_w(x,\pi) = \inf_{(x,\pi)\in\mathcal{F}} L_w(x,\pi) = \beta.\end{aligned}$$

Next, we first explain the relationships between the problems $(\widehat{\mathrm{P}})$ and (P), and then obtain the weak duality property for (P) under an appropriate condition. To this end, we shall consider regular weak separation functions satisfying the following assumption:

Assumption $\mathcal{A}$. $w \in \mathcal{W}_R(\Pi)$ satisfies $w(1, 0_{\mathbb{R}^m}; \pi) = 1$ for all $\pi \in \Pi$, and

$$\inf_{\pi \in \Pi} w(u, v; \pi) = \begin{cases} u, & \text{if } v \in \mathbb{R}^m_+, \\ -\infty, & \text{if } v \notin \mathbb{R}^m_+. \end{cases}$$

Note that Assumption $\mathcal{A}$ is essentially a reinforcement of Definition 2.2. In fact, if we make on further assumption that the infimum in Assumption $\mathcal{A}$ can be attained for every $v \in \mathbb{R}^m_+$, then every function $w : \mathbb{R}^{1+m} \times \Pi \to \mathbb{R}$ satisfying Assumption $\mathcal{A}$ must be a regular weak separation function. On the one hand, for every $(u, v) \in \mathcal{H}$, i.e., $u > 0$ and $v \in \mathbb{R}^m_+$, it follows from Assumption $\mathcal{A}$ that

$$w(u, v; \pi) \geq \inf_{\pi \in \Pi} w(u, v; \pi) = u > 0, \quad \forall \pi \in \Pi,$$

which implies

$$\mathcal{H} \subset \bigcap_{\pi \in \Pi} \mathrm{lev}_{>0} w(\bullet; \pi).$$

On the other hand, for every $(u, v) \in \mathbb{R}^{1+m}$ satisfying $w(u, v; \pi) > 0$ for all $\pi \in \Pi$, we have $\inf_{\pi \in \Pi} w(u, v; \pi) \geq 0$. Then it follows that $v \in \mathbb{R}^m_+$. Otherwise, by Assumption $\mathcal{A}$, $v \notin \mathbb{R}^m_+$ implies $\inf_{\pi \in \Pi} w(u, v; \pi) = -\infty$. This is a contradiction. Since the infimum in Assumption $\mathcal{A}$ is attained, there exists some $\hat{\pi} \in \Pi$ such that

$$u = \inf_{\pi \in \Pi} w(u, v; \pi) = w(u, v; \hat{\pi}) > 0.$$

Thus, we can conclude that $w \in \mathcal{W}_R(\Pi)$. However, it is not any regular weak separation function satisfying Assumption $\mathcal{A}$. We give the following example to illustrate the case.

Example 3.1 Let the set of parameters $\Pi = \mathbb{R}_{++} \times \mathbb{R}^m_+$ and the function $w : \mathbb{R}^{1+m} \times (\mathbb{R}_{++} \times \mathbb{R}^m_+) \to \mathbb{R}$ be defined by $w(u, v; \theta, \lambda) := \theta u + \langle \lambda, v \rangle, \forall (u, v) \in \mathbb{R}^{1+m}, \forall (\theta, \lambda) \in \Pi$. By Definition 2.2, it is easy to verify that w is a regular weak separation function. Although, Assumption $\mathcal{A}$ does not hold. In fact, by directly calculating, we have $w(1, 0_{\mathbb{R}^m}; \theta, \lambda) = \theta$ for all $(\theta, \lambda) \in \mathbb{R}_{++} \times \mathbb{R}^m_+$, and

$$\inf_{(\theta,\lambda) \in \Pi} w(u, v; \theta, \lambda) = \begin{cases} 0, & \text{if } v \in \mathbb{R}^m_+, u \geq 0, \\ -\infty, & \text{if } v \in \mathbb{R}^m_+, u < 0 \text{ or } v \notin \mathbb{R}^m_+. \end{cases}$$

Note that, if we take $\Pi = \{1\} \times \mathbb{R}^m_+$ and the function $w : \mathbb{R}^{1+m} \times (\{1\} \times \mathbb{R}^m_+) \to \mathbb{R}$ defined by $w(u, v; 1, \lambda) := u + \langle \lambda, v \rangle, \forall (u, v) \in \mathbb{R}^{1+m}, \forall (1, \lambda) \in \Pi$, then it is easy to verify that w is a regular weak separation function satisfying Assumption $\mathcal{A}$.

The following lemma is a direct consequence of Assumption $\mathcal{A}$.

Lemma 3.1 *Let the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ satisfy Assumption $\mathcal{A}$. Then it follows that $\mathcal{F}^o \subset R$. Together with $R \subset \mathcal{F}^o$, one has $R = \mathcal{F}^o$.*

Proof Take arbitrary $x \in \mathcal{F}^o$. Then there exists some $\hat{\pi} \in \Pi$ such that

$$L_w(x, \hat{\pi}) = \sup_{\pi \in \Pi} L_w(x, \pi).$$

Together with Assumption $\mathcal{A}$, it follows that

$$w\big(f(\bar{x}) - f(x), g(x); \hat{\pi}\big) = \inf_{\pi \in \Pi} w\big(f(\bar{x}) - f(x), g(x); \pi\big).$$

This implies $g(x) \in \mathbb{R}^m_+$. Otherwise, $g(x) \notin \mathbb{R}^m_+$. We have necessarily

$$\inf_{\pi \in \Pi} w\big(f(\bar{x}) - f(x), g(x); \pi\big) = -\infty$$

by Assumption $\mathcal{A}$. This is a contradiction. Therefore, we have $x \in R$, which shows $\mathcal{F}^o \subset R$ since x is arbitrary. □

Now, we give the following lemma to show the relationships between the problems $(\widehat{\mathrm{P}})$ and (P) under Assumption $\mathcal{A}$.

Lemma 3.2 *Let the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ satisfy Assumption $\mathcal{A}$. Then the extrema of problems $(\widehat{\mathrm{P}})$ and (P) are equal, that is,*

$$\inf_{x \in R} f(x) = \beta < +\infty.$$

Moreover, if $(\hat{x}, \hat{\pi}) \in \mathcal{F}$ is a global minimum point of $(\widehat{\mathrm{P}})$, then $\hat{x} \in R$ and it is a global minimum point of (P). Conversely, if $\hat{x}$ is a global minimum point of (P), then there exists some $\hat{\pi} \in \mathcal{J}_{\hat{x}}$ such that $(\hat{x}, \hat{\pi}) \in \mathcal{F}$, and it is a global minimum point of $(\widehat{\mathrm{P}})$.

Proof It follows from (5), (6), (7) and Assumption $\mathcal{A}$ that

$$\begin{aligned}
\beta = \inf_{(x,\pi) \in \mathcal{F}} L_w(x, \pi) &= \inf_{x \in \mathcal{F}^o} \sup_{\pi \in \Pi} L_w(x, \pi) \\
&= \inf_{x \in \mathcal{F}^o} \sup_{\pi \in \Pi} \big(w(1, 0_{\mathbb{R}^m}; \pi) f(\bar{x}) - w\big(f(\bar{x}) - f(x), g(x); \pi\big)\big) \\
&= f(\bar{x}) - \sup_{x \in \mathcal{F}^o} \inf_{\pi \in \Pi} w\big(f(\bar{x}) - f(x), g(x); \pi\big).
\end{aligned}$$

By Assumption $\mathcal{A}$ and Lemma 3.1, it follows that $\mathcal{F}^o = R$. Then we have

$$\beta = f(\bar{x}) - \sup_{x \in R} \inf_{\pi \in \Pi} w\big(f(\bar{x}) - f(x), g(x); \pi\big).$$

Furthermore, it follows from $x \in R$ that $g(x) \in \mathbb{R}^m_+$, and we get

$$\inf_{\pi \in \Pi} w\big(f(\bar{x}) - f(x), g(x); \pi\big) = f(\bar{x}) - f(x).$$

Thus, we can conclude that

$$\beta = \inf_{(x,\pi)\in\mathcal{F}} L_w(x,\pi) = f(\bar{x}) - \sup_{x\in R}\big(f(\bar{x}) - f(x)\big) \\ = \inf_{x\in R} f(x),$$

that is, the extrema of the problems $(\widehat{\mathrm{P}})$ and (P) are equal. Since $R \neq \varnothing$, we get

$$\beta = \inf_{x\in R} f(x) < +\infty.$$

Moreover, on the one hand, if $(\hat{x},\hat{\pi}) \in \mathcal{F}$ is a global minimum point of $(\widehat{\mathrm{P}})$, then $(\hat{x},\hat{\pi}) \in \mathcal{N}$ and $L_w(\hat{x},\hat{\pi}) = \beta$. By (7), we have $\hat{\pi} \in \mathcal{J}_{\hat{x}}$. Moreover, it follows from Assumption $\mathcal{A}$ that

$$L_w(\hat{x},\hat{\pi}) = \sup_{\pi\in\Pi} L_w(\hat{x},\pi) = f(\bar{x}) - \inf_{\pi\in\Pi} w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \pi\big).$$

Together with $L_w(\hat{x},\hat{\pi}) = \beta < +\infty$ and Assumption $\mathcal{A}$, we have $g(\hat{x}) \in \mathbb{R}^m_+$ and

$$\inf_{x\in R} f(x) = \beta = L_w(\hat{x},\hat{\pi}) = f(\bar{x}) - \big(f(\bar{x}) - f(\hat{x})\big) = f(\hat{x}),$$

which implies that $\hat{x} \in R$ and it is a global minimum point of (P). On the other hand, if $\hat{x} \in R$ is a global minimum point of (P), then

$$f(\hat{x}) = \inf_{x\in R} f(x).$$

Since $\hat{x} \in R$, we have $g(\hat{x}) \in \mathbb{R}^m_+$. It follows from Assumption $\mathcal{A}$ and Lemma 3.1 that

$$\begin{aligned} f(\hat{x}) &= f(\bar{x}) - \big(f(\bar{x}) - f(\hat{x})\big) \\ &= w(1, 0_{\mathbb{R}^m}; \pi) f(\bar{x}) - \inf_{\pi\in\Pi} w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \pi\big) \\ &= \sup_{\pi\in\Pi} L_w(\hat{x},\pi) \\ &\geq \inf_{x\in\mathcal{F}^o} \sup_{\pi\in\Pi} L_w(x,\pi) \quad \big(\text{since } \hat{x} \in R = \mathcal{F}^o\big) \\ &= \inf_{(x,\pi)\in\mathcal{F}} L_w(x,\pi) = \beta = \inf_{x\in R} f(x) = f(\hat{x}). \end{aligned}$$

Thus, we have

$$\inf_{(x,\pi)\in\mathcal{F}} L_w(x,\pi) = \sup_{\pi\in\Pi} L_w(\hat{x},\pi).$$

Furthermore, since $\hat{x} \in \mathcal{F}^o$, then there exists some $\hat{\pi} \in \Pi$ such that $(\hat{x},\hat{\pi}) \in \mathcal{F}$. Therefore, $\hat{\pi} \in \mathcal{J}_{\hat{x}}$ and

$$L_w(\hat{x},\hat{\pi}) = \sup_{\pi\in\Pi} L_w(\hat{x},\pi) = \inf_{(x,\pi)\in\mathcal{F}} L_w(x,\pi),$$

that is, $(\hat{x},\hat{\pi})$ is a global minimum point of $(\widehat{\mathrm{P}})$. □

Theorem 3.1 *Let the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ satisfy Assumption $\mathcal{A}$. Then the following results hold*:

(i) *The weak duality holds*, *that is*,

$$\inf_{x \in R} f(x) \geq \sup_{(x,\pi) \in \mathcal{Q}} L_w(x, \pi).$$

(ii) $\hat{x} \in X$ *is a global minimum point of* (P) *if and only if*

$$\sup_{\pi \in \Pi} L_w(\hat{x}, \pi) = \beta.$$

Proof (i) Since $\alpha \leq \beta$ always holds, it follows from Lemma 3.2 that

$$\inf_{x \in R} f(x) \geq \sup_{(x,\pi) \in \mathcal{Q}} L_w(x, \pi).$$

(ii) $\hat{x} \in X$ being a global minimum point of (P) implies

$$g(\hat{x}) \in \mathbb{R}^m_+, \quad \inf_{x \in R} f(x) = f(\hat{x}). \tag{8}$$

Moreover, it follows from (5) and Assumption $\mathcal{A}$ that

$$\begin{aligned} \sup_{\pi \in \Pi} L_w(\hat{x}, \pi) &= \sup_{\pi \in \Pi} \big(w(1, 0_{\mathbb{R}^m}; \pi) f(\bar{x}) - w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \pi\big)\big) \\ &= f(\bar{x}) - \inf_{\pi \in \Pi} w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \pi\big) \\ &= f(\bar{x}) - \big(f(\bar{x}) - f(\hat{x})\big) = f(\hat{x}). \end{aligned}$$

Together with Lemma 3.2 and (8), we have

$$\sup_{\pi \in \Pi} L_w(\hat{x}, \pi) = \inf_{x \in R} f(x) = \beta.$$

Conversely, if $\beta = \sup_{\pi \in \Pi} L_w(\hat{x}, \pi)$, then we get

$$\begin{aligned} \beta &= \sup_{\pi \in \Pi} \big(w(1, 0_{\mathbb{R}^m}; \pi) f(\bar{x}) - w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \pi\big)\big) \\ &= f(\bar{x}) - \inf_{\pi \in \Pi} w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \pi\big). \end{aligned}$$

Since $\beta < +\infty$, it follows from Assumption $\mathcal{A}$ that $g(\hat{x}) \in \mathbb{R}^m_+$, and then, $\hat{x} \in R$ and

$$\inf_{\pi \in \Pi} w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \pi\big) = f(\bar{x}) - f(\hat{x}).$$

Together with Lemma 3.2, we have

$$\begin{aligned} \inf_{x \in R} f(x) = \beta &= f(\bar{x}) - \inf_{\pi \in \Pi} w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \pi\big) \\ &= f(\bar{x}) - \big(f(\bar{x}) - f(\hat{x})\big) = f(\hat{x}). \end{aligned}$$

Thus, $\hat{x} \in R$ is a global minimum point of (P). This completes the proof. $\square$

Note that all the results in Lemmas 3.1 and 3.2, and Theorem 3.1, are independent of the choice of $\bar{x}$. As we know, when the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ is separable with respect to u and v (see more details in Sect. 4, Case 1, of [34]), a similar condition to Assumption $\mathcal{A}$ has been obtained by Giannessi et al. [14]. Moreover, the following assumption in Rubinov et al. [6, 33], which is weaker than Assumption $\mathcal{A}$, has been introduced:

$$(\widetilde{\mathcal{A}}) \qquad \inf_{\pi \in \Pi} w(u, v; \pi) \geq u, \quad \forall v \in \mathbb{R}^m_+,$$

and it has been shown that Assumption $(\widetilde{\mathcal{A}})$ is sufficient to assure that the weak duality holds. However, we cannot obtain all the corresponding results in Theorem 3.1, which play important roles in the sequel analysis. The following example explains the cases.

Example 3.2 Consider the following constrained extremum problem:

$$\min\left(-x^2\right), \quad \text{s.t.} \quad 1 - x^2 \geq 0, \quad -x^2 - x + 2 \geq 0.$$

Take the family of parameters $\Pi = \{2\} \times \mathbb{R}^2_+$ and the function $w : \mathbb{R}^3 \times \Pi \to \mathbb{R}$, defined by

$$\begin{aligned} w(u, v; 2, \lambda) &:= 2u + \sup_{z \in \{v\} - \mathbb{R}^2_+} \left(\langle \lambda, z\rangle - \|z\|_1\right) \\ &= 2u + \sup_{z_1 \leq v_1, z_2 \leq v_2} \left(\lambda_1 z_1 + \lambda_2 z_2 - |z_1| - |z_2|\right), \\ &\qquad \forall (u, v) \in \mathbb{R}^3, \forall (2, \lambda) \in \Pi. \end{aligned}$$

It is easy to verify that w is a regular weak separation function and

$$\inf_{(2,\lambda) \in \Pi} w(u, v; 2, \lambda) = 2u, \quad \forall v \in \mathbb{R}^2_+$$

(see more details in Sect. 4, Case 2, of [34]). Thus, Assumption $(\widetilde{\mathcal{A}})$ holds but Assumption $\mathcal{A}$ does not. Moreover, let $\bar{x} = -2$ and $\hat{x} = 1$. Then we have

$$L_w(\hat{x}; 2, \lambda) = -2 - \sup_{z_1 \leq 0, z_2 \leq 0} \left(\lambda_1 z_1 + \lambda_2 z_2 - |z_1| - |z_2|\right) = -2, \quad \forall (2, \lambda) \in \Pi,$$

which implies

$$\sup_{(2,\lambda) \in \Pi} L_w(\hat{x}; 2, \lambda) = -2.$$

Clearly, $\hat{x}$ is a global minimum point, $\beta = -1$ and $\sup_{(2,\lambda)\in\Pi} L_w(\hat{x}; 2, \lambda) < \beta$. Therefore, the weak duality holds, but the conclusion (ii) in Theorem 3.1 does not hold. However, if we take $\Pi = \{1\} \times \mathbb{R}^2_+$ and the function $w : \mathbb{R}^3 \times \Pi \to \mathbb{R}$, defined by

$$w(u, v; 1, \lambda) := u + \sup_{z \in \{v\} - \mathbb{R}^2_+} \left(\langle \lambda, z\rangle - \|z\|_1\right), \quad \forall (u, v) \in \mathbb{R}^3, \forall (1, \lambda) \in \Pi,$$

then w is a regular weak separation function satisfying Assumption $\mathcal{A}$. Moreover, we have

$$\sup_{(1,\lambda)\in\Pi} L_w(\hat{x}; 1, \lambda) = -1 = \beta,$$

that is, the weak duality and the assertion (ii) of Theorem 3.1 hold.

Remark 3.1 Note that, if Assumption $\mathcal{A}$ holds, it follows from Lemma 3.2 that the zero duality gap property with respect to w can be expressed in the form $\alpha = \beta$. Since the weak duality holds, that is, $\alpha \leq \beta$ always holds (as shown in Theorem 3.1(i)), we can equivalently express the zero duality gap property as $\alpha \geq \beta$.

3.2 Zero Duality Gap Properties

The purpose of this subsection is to investigate some equivalent characterizations of the zero duality gap property for (P) from two aspects. We first follow the regular separation in the IS and the classic approach, such as the Lagrange multiplier and the saddle point, and then apply the perturbation function of (P) to discuss the zero duality gap property. In the beginning, we introduce some standard notions associated with the generalized Lagrange function for (P).

Definition 3.2 Given the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ and the generalized Lagrange function L_w for (P) corresponding to w, then $\hat{\pi} \in \Pi$ is said to be a generalized Lagrange multiplier for (P) corresponding to w iff

$$\inf_{x\in X} L_w(x, \hat{\pi}) = \beta.$$

Remark 3.2 As we know, if we let $\bar{x} \in X$ such that $f(\bar{x})$ is the global minimum for (P), then $\mathcal{K} \cap \mathcal{H} = \varnothing$ (Note that $\bar{x}$ may not be a feasible point). In addition, if $\hat{\pi} \in \Pi$ is a generalized Lagrange multiplier for (P) corresponding to w and Assumption $\mathcal{A}$ holds, it follows from Theorem 3.1(i) that

$$\sup_{x\in X} w\big(f(\bar{x}) - f(x), g(x); \hat{\pi}\big) = 0,$$

that is, $w(u, v; \hat{\pi}) \leq 0, \forall (u, v) \in \mathcal{K}$. Thus, the sets $\mathcal{K}$ and $\mathcal{H}$ admit a regular separation with respect to $w \in \mathcal{W}_R(\Pi)$ and $\hat{\pi} \in \Pi$.

Definition 3.3 Given the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ and the generalized Lagrange function L_w for (P) corresponding to w, then $(\hat{x}, \hat{\pi}) \in X \times \Pi$ is said to be a generalized Lagrange saddle point for (P) corresponding to w iff

$$L_w(\hat{x}, \pi) \leq L_w(\hat{x}, \hat{\pi}) \leq L_w(x, \hat{\pi}), \quad \forall x \in X, \forall \pi \in \Pi.$$

Next, we establish some equivalent statements to the zero duality gap property for (P).

Lemma 3.3 *Given $\hat{x} \in X$ and $\hat{\pi} \in \Pi$, and the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ satisfying Assumption $\mathcal{A}$, then the following assertions are equivalent*:

- (i) *$\hat{x} \in R$ is a global minimum point and $\hat{\pi}$ is a generalized Lagrange multiplier for* (P) *corresponding to w.*
- (ii) *$\hat{x} \in X$ and $(\hat{x}, \hat{\pi}) \in X \times \Pi$ is a generalized Lagrange saddle point for* (P) *corresponding to w.*

In addition, we also have $(\hat{x}, \hat{\pi}) \in \mathcal{M} \cap \mathcal{N}$ and $w(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \hat{\pi}) = f(\bar{x}) - f(\hat{x})$.

Proof (i) $\Rightarrow$ (ii) Since $\hat{x} \in R$ is a global minimum point and $\hat{\pi}$ is a generalized Lagrange multiplier for (P) corresponding to w, it follows from Theorem 3.1(ii) and Definition 3.2 that

$$\sup_{\pi \in \Pi} L_w(\hat{x}, \pi) = \beta = \inf_{x \in X} L_w(x, \hat{\pi}),$$

i.e., $L_w(\hat{x}, \pi) \leq \beta \leq L_w(x, \hat{\pi}), \forall x \in X, \forall \pi \in \Pi$. Specially, let $x = \hat{x}$ and $\pi = \hat{\pi}$. Then we get $\beta = L_w(\hat{x}, \hat{\pi})$. Thus, $L_w(\hat{x}, \pi) \leq L_w(\hat{x}, \hat{\pi}) \leq L_w(x, \hat{\pi}), \forall x \in X$, $\forall \pi \in \Pi$, that is, $(\hat{x}, \hat{\pi}) \in X \times \Pi$ is a generalized Lagrange saddle point for (P) corresponding to w.

(ii) $\Rightarrow$ (i) Since $(\hat{x}, \hat{\pi}) \in X \times \Pi$ is a generalized Lagrange saddle point for (P) corresponding to w, we get

$$L_w(\hat{x}, \pi) \leq L_w(\hat{x}, \hat{\pi}) \leq L_w(x, \hat{\pi}), \quad \forall x \in X, \forall \pi \in \Pi. \tag{9}$$

First, we prove that $L_w(\hat{x}, \pi) \leq L_w(\hat{x}, \hat{\pi}), \forall \pi \in \Pi$, implies $\hat{x} \in R$. In fact, it follows from Assumption $\mathcal{A}$ that the above inequality implies $w(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \pi) \geq w(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \hat{\pi}), \forall \pi \in \Pi$. So,

$$\inf_{\pi \in \Pi} w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \pi\big) \geq w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \hat{\pi}\big). \tag{10}$$

Suppose that $\hat{x} \notin R$. Then we have $g(\hat{x}) \notin \mathbb{R}^m_+$. It follows from Assumption $\mathcal{A}$ that

$$\inf_{\pi \in \Pi} w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \pi\big) = -\infty.$$

However, this is a contradiction to (10). Second, we prove that

$$w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \hat{\pi}\big) = f(\bar{x}) - f(\hat{x}).$$

On the one hand, since $\hat{x} \in R$, we have $g(\hat{x}) \in \mathbb{R}^m_+$. By Assumption $\mathcal{A}$ and (10), we get

$$\inf_{\pi \in \Pi} w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \pi\big) = f(\bar{x}) - f(\hat{x}) \geq w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \hat{\pi}\big).$$

On the other hand, it follows from (9), Assumption $\mathcal{A}$ and Lemma 3.1 that $\hat{x} \in R = \mathcal{F}^o$ and

$$\begin{aligned}\beta = \inf_{(x,\pi)\in\mathcal{F}} L_w(x,\pi) &= \inf_{x\in\mathcal{F}^o} \sup_{\pi\in\Pi} L_w(x,\pi) \\ &\le \sup_{\pi\in\Pi} L_w(\hat{x},\pi) \le L_w(\hat{x},\hat{\pi}) \le \inf_{x\in X} L_w(x,\hat{\pi}) \\ &\le \inf_{x\in\mathcal{F}^\circ} L_w(x,\hat{\pi}) \le \sup_{\pi\in\Pi} \inf_{x\in\mathcal{F}^\circ} L_w(x,\pi) \\ &\le \inf_{x\in\mathcal{F}^\circ} \sup_{\pi\in\Pi} L_w(x,\pi) \\ &= \beta. \end{aligned} \tag{11}$$

Thus, we have $\beta = L_w(\hat{x},\hat{\pi}) = w(1, 0_{\mathbb{R}^m}; \hat{\pi}) f(\bar{x}) - w(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \hat{\pi})$. Together with Assumption $\mathcal{A}$, Lemma 3.2 and $\hat{x} \in R$, we get

$$f(\bar{x}) - w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \hat{\pi}\big) = \beta = \inf_{x\in R} f(x) \le f(\hat{x}),$$

which implies $w(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \hat{\pi}) \ge f(\bar{x}) - f(\hat{x})$. Therefore, we can conclude that

$$w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \hat{\pi}\big) = f(\bar{x}) - f(\hat{x}),$$

which implies

$$f(\hat{x}) = f(\bar{x}) - w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \hat{\pi}\big) = \inf_{x\in R} f(x),$$

that is, $\hat{x} \in R$ is a global minimum point for (P). Lastly, it follows from (11) that

$$L_w(\hat{x},\hat{\pi}) = \sup_{\pi\in\Pi} L_w(\hat{x},\pi) = \inf_{x\in X} L_w(x,\hat{\pi}) = \beta.$$

Thus, $\hat{\pi}$ is a generalized Lagrange multiplier for (P) corresponding to w, and simultaneously, $\hat{x} \in \mathcal{I}_{\hat{\pi}}$ and $\hat{\pi} \in \mathcal{J}_{\hat{x}}$, i.e., $(\hat{x},\hat{\pi}) \in \mathcal{M} \cap \mathcal{N}$. This completes the proof. □

Specially, if we take $\bar{x} \in R$, then we immediately have the following characterization in terms of a regular separation related to $\mathcal{K}$ and $\mathcal{H}$ in the IS, which generalizes and improves Theorem 5.2 in [3], Proposition 5.2.11 in [4], Theorem 2.1 in [8], Theorem 3.2 in [9], and Theorems 3.2 and 3.3 in [14].

Lemma 3.4 *Given $\bar{\pi} \in \Pi$ and the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ satisfying Assumption $\mathcal{A}$, then the following assertions are equivalent*:

(i) $\bar{x} \in R$, *and the sets $\mathcal{K}$ and $\mathcal{H}$ admit a regular separation with respect to w and $\bar{\pi}$.*
(ii) $\bar{x} \in R$ *is a global minimum point and $\bar{\pi}$ is a generalized Lagrange multiplier for* (P) *corresponding to w.*
(iii) $\bar{x} \in X$ *and $(\bar{x},\bar{\pi}) \in X \times \Pi$ is a generalized Lagrange saddle point for* (P) *corresponding to w.*

In addition, we also have $(\bar{x},\bar{\pi}) \in \mathcal{M} \cap \mathcal{N}$ and $w(0, g(\bar{x}); \bar{\pi}) = 0$.

Proof We only need to prove that (i) $\Rightarrow$ (ii) and (iii) $\Rightarrow$ (i) since the equivalence of (ii) and (iii) is fulfilled from Lemma 3.3.

(i) $\Rightarrow$ (ii) Since the sets $\mathcal{K}$ and $\mathcal{H}$ admit a regular separation with respect to w and $\bar{\pi}$, that is,

$$w\big(f(\bar{x}) - f(x), g(x); \bar{\pi}\big) \le 0, \quad \forall x \in X, \tag{12}$$

we have $\mathcal{K} \cap \mathcal{H} = \varnothing$. Together with $\bar{x} \in R$ and Remark 2.1, it follows that $\bar{x}$ is a global minimum point for (P), i.e.,

$$f(\bar{x}) = \inf_{x \in R} f(x). \tag{13}$$

Moreover, by (12) we get

$$\sup_{x \in X} w\big(f(\bar{x}) - f(x), g(x); \bar{\pi}\big) \le 0.$$

Together with (5), (13), Assumption $\mathcal{A}$ and Lemma 3.2, we get

$$\begin{aligned}
\inf_{x \in X} L_w(x, \bar{\pi}) &= w(1, 0_{\mathbb{R}^m}; \bar{\pi}) f(\bar{x}) - \sup_{x \in X} w\big(f(\bar{x}) - f(x), g(x); \bar{\pi}\big) \\
&\ge f(\bar{x}) \\
&= \beta = \inf_{(x,\pi) \in \mathcal{F}} L_w(x, \pi) = \inf_{x \in \mathcal{F}^o} \sup_{\pi \in \Pi} L_w(x, \pi) \\
&\ge \inf_{x \in \mathcal{F}^o} L_w(x, \bar{\pi}) \\
&\ge \inf_{x \in X} L_w(x, \bar{\pi}),
\end{aligned}$$

where the second inequality holds since $\bar{\pi} \in \Pi$ and $\sup_{\pi \in \Pi} L_w(x, \pi) \ge L_w(x, \bar{\pi})$, $\forall x \in \mathcal{F}^o$. Therefore, we have

$$\inf_{x \in X} L_w(x, \bar{\pi}) = \beta,$$

that is, $\bar{\pi}$ is a generalized Lagrange multiplier for (P) corresponding to w.

(iii) $\Rightarrow$ (i) Let $(\bar{x}, \bar{\pi}) \in X \times \Pi$ be a generalized Lagrange saddle point for (P) corresponding to w. Then we have

$$L_w(\bar{x}, \pi) \le L_w(\bar{x}, \bar{\pi}) \le L_w(x, \bar{\pi}), \quad \forall x \in X, \forall \pi \in \Pi.$$

By the method similar to the proof of Lemma 3.3 (ii) $\Rightarrow$ (i), we get $\bar{x} \in R$, $w(0, g(\bar{x}); \bar{\pi}) = 0$ and $(\bar{x}, \bar{\pi}) \in \mathcal{M} \cap \mathcal{N}$. Together with $L_w(\bar{x}, \bar{\pi}) \le L_w(x, \bar{\pi})$, $\forall x \in X$, it follows that

$$w\big(f(\bar{x}) - f(x), g(x); \bar{\pi}\big) \le 0, \quad \forall x \in X,$$

that is, the sets $\mathcal{K}$ and $\mathcal{H}$ admit a regular separation with respect to w and $\bar{\pi}$. This completes the proof. $\square$

Remark 3.3 Compared with some recent papers [7–9] and references therein, Lemma 3.4 establishes an equivalent characterization not only for a Lagrange saddle point, but also for a Lagrange multiplier by virtue of the existence of a regular weak separation in the IS under Assumption $\mathcal{A}$. Very recently, Luo et al. [9] showed that the existence of a regular nonlinear separation was equivalent to a saddle point for a special augmented Lagrange function by using Proposition 3.2 in [9]. Note that, if we take specially the parameter set Π and the regular separation function $w \in W_R(\Pi)$, defined by (6) (resp. (18)) in [9], then Assumption $\mathcal{A}$ reduces to (19) and (20) (resp. (9) and (10)) in [9]. Moreover, it follows from Lemma 3.4 that Theorems 3.1 and 3.2, and Corollary 3.1 in [9] are immediately held. Thus, in Lemma 3.4 we establish some more comprehensive results for (P) by virtue of a more general regular weak separation in the IS.

Theorem 3.2 *Let the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ satisfy Assumption $\mathcal{A}$. Then the following results hold*:

(i) *If $\mathcal{Q} \cap \mathcal{F} \neq \varnothing$, then the zero duality gap property with respect to w holds, that is*,

$$\sup_{(x,\pi)\in\mathcal{Q}} L_w(x,\pi) = \inf_{x\in R} f(x).$$

(ii) *For every $(\hat{x},\hat{\pi}) \in \mathcal{Q} \cap \mathcal{F}$, it follows that the zero duality gap property with respect to w holds, and $\hat{x} \in R$ is a global minimum point of* (P) *and $(\hat{x},\hat{\pi})$ is a global maximum point of* (DP). *Specially, if $\mathcal{F} = \mathcal{N}$, then the converse also holds. In addition, we have*

$$w\big(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \hat{\pi}\big) = f(\bar{x}) - f(\hat{x}).$$

Proof (i) It is easy to verify that $\mathcal{Q} \cap \mathcal{F} \neq \varnothing$ implies $\alpha = \beta$. In fact, it follows from $\mathcal{Q} \cap \mathcal{F} \neq \varnothing$ that there exists some $(\hat{x},\hat{\pi}) \in \mathcal{Q} \cap \mathcal{F}$. Then we have

$$\alpha = \sup_{(x,\pi)\in\mathcal{Q}} L_w(x,\pi) \geq L_w(\hat{x},\hat{\pi}) \geq \inf_{(x,\pi)\in\mathcal{F}} L_w(x,\pi) = \beta. \tag{14}$$

Since $\alpha \leq \beta$ always holds, we get $\alpha = \beta$. Moreover, if the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ satisfies Assumption $\mathcal{A}$, then by Lemma 3.2 we have

$$\sup_{(x,\pi)\in\mathcal{Q}} L_w(x,\pi) = \alpha = \beta = \inf_{x\in R} f(x).$$

(ii) For every $(\hat{x},\hat{\pi}) \in \mathcal{Q} \cap \mathcal{F}$, we have $\mathcal{Q} \cap \mathcal{F} \neq \varnothing$. By (i), it follows that the zero duality gap property holds. Moreover, by (14) and $\alpha \leq \beta$, we get $L_w(\hat{x},\hat{\pi}) = \alpha = \beta$. Thus, $(\hat{x},\hat{\pi}) \in \mathcal{Q} \cap \mathcal{F}$ is a maximum point of (DP), and also is a minimum point of $(\widehat{\mathrm{P}})$. Together with Lemma 3.2, we have that $\hat{x} \in R$ is a global minimum point of (P), i.e.,

$$\inf_{x\in R} f(x) = f(\hat{x}).$$

So, we get $f(\hat{x}) = \beta = L_w(\hat{x},\hat{\pi})$ and then, $w(f(\bar{x}) - f(\hat{x}), g(\hat{x}); \hat{\pi}) = f(\bar{x}) - f(\hat{x})$. Specially, let $\mathcal{F} = \mathcal{N}$. If the zero duality gap property with respect to w holds, and

$\hat{x} \in R$ is a global minimum point of (P) and $(\hat{x}, \hat{\pi})$ is a global maximum point of (DP), then we have

$$(\hat{x}, \hat{\pi}) \in \mathcal{Q} \tag{15}$$

and

$$L_w(\hat{x}, \hat{\pi}) = \sup_{(x,\pi)\in\mathcal{Q}} L_w(x, \pi) = \inf_{x\in R} f(x) = \beta.$$

Since $\hat{x} \in R$ is a global minimum point of (P), by Theorem 3.1(ii) we get

$$\sup_{\pi\in\Pi} L_w(\hat{x}, \pi) = \beta.$$

Thus,

$$L_w(\hat{x}, \hat{\pi}) = \sup_{\pi\in\Pi} L_w(\hat{x}, \pi),$$

which implies $\hat{\pi} \in \mathcal{J}_{\hat{x}}$. Then we have $(\hat{x}, \hat{\pi}) \in \mathcal{N}$. Together with (15), we can conclude that $(\hat{x}, \hat{\pi}) \in \mathcal{Q} \cap \mathcal{N}$. □

Note that, if $\mathcal{F} = \mathcal{N}$, then we have the following corollary from Lemma 3.3, and Theorems 3.1 and 3.2.

Corollary 3.1 *Let $\hat{\pi} \in \Pi$ and let the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ satisfy Assumption $\mathcal{A}$. If one of the following conditions is fulfilled,*

(i) *$\hat{x} \in R$ is a global minimum point and $\hat{\pi}$ is a generalized Lagrange multiplier for* (P) *corresponding to w,*
(ii) *$\hat{x} \in X$ and $(\hat{x}, \hat{\pi}) \in X \times \Pi$ is a generalized Lagrange saddle point for* (P) *corresponding to w,*

then, for every set $\mathcal{Q} \subset \mathcal{M}$ with $(\hat{x}, \hat{\pi}) \in \mathcal{Q}$, the zero duality gap property with respect to w holds and

$$L_w(\hat{x}, \hat{\pi}) = f(\hat{x}) = \inf_{x\in R} f(x) = \max_{(x,\pi)\in\mathcal{Q}} L_w(x, \pi).$$

Proof It follows from the proof of Lemma 3.3 that the conditions (i) and (ii) are equivalent, and any one of the conditions (i)–(ii) implies

$$L_w(\hat{x}, \hat{\pi}) = \sup_{\pi\in\Pi} L_w(\hat{x}, \pi).$$

Thus, $\hat{\pi} \in \mathcal{J}_{\hat{x}}$ and then, $(\hat{x}, \hat{\pi}) \in \mathcal{N}$. Moreover, since $(\hat{x}, \hat{\pi}) \in \mathcal{Q}$, we have $(\hat{x}, \hat{\pi}) \in \mathcal{Q} \cap \mathcal{N}$. Together with Theorem 3.2(ii) with $\mathcal{F} = \mathcal{N}$, we get that the zero duality gap property holds and

$$f(\hat{x}) = \inf_{x\in R} f(x) = \max_{(x,\pi)\in\mathcal{Q}} L_w(x, \pi) = L_w(\hat{x}, \hat{\pi}).$$

This completes the proof. □

Similarly, the following result holds when $\bar{x} \in R$ and $\mathcal{F} = \mathcal{N}$ by Lemma 3.4, and Theorems 3.1 and 3.2.

Corollary 3.2 *Let $\bar{\pi} \in \Pi$ and let the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ satisfy Assumption $\mathcal{A}$. If one of the following conditions is fulfilled,*

(i) *$\bar{x} \in R$, and the sets $\mathcal{K}$ and $\mathcal{H}$ admit a regular separation with respect to w and $\bar{\pi}$,*
(ii) *$\bar{x} \in R$ is a global minimum point and $\bar{\pi}$ is a generalized Lagrange multiplier for* (P) *corresponding to w,*
(iii) *$\bar{x} \in X$ and $(\bar{x}, \bar{\pi}) \in X \times \Pi$ is a generalized Lagrange saddle point for* (P) *corresponding to w,*

then, for every set $\mathcal{Q} \subset \mathcal{M}$ with $(\bar{x}, \bar{\pi}) \in \mathcal{Q}$, the zero duality gap property with respect to w holds and

$$L_w(\bar{x}, \bar{\pi}) = f(\bar{x}) = \inf_{x \in R} f(x) = \max_{(x,\pi) \in \mathcal{Q}} L_w(x, \pi).$$

Remark 3.4 As shown in Theorem 3.2 (resp. Corollaries 3.1 and 3.2), the condition $\mathcal{Q} \cap \mathcal{F} \neq \varnothing$ (resp. $\mathcal{Q} \cap \mathcal{N} \neq \varnothing$) is very useful to establish the zero duality gap property. In fact, Corollary 3.1 (resp. Corollary 3.2) gives a sufficient condition for $\mathcal{Q} \cap \mathcal{N} \neq \varnothing$ by means of a generalized Lagrange multiplier or a generalized Lagrange saddle point (resp. a general regular separation in the IS). Moreover, we will see in Sect. 3.2 of [34] that the condition $\mathcal{Q} \cap \mathcal{N} \neq \varnothing$ for Wolfe and Mond–Weir dualities can be verified under some appropriate assumptions. Note that $\mathcal{Q} \cap \mathcal{F} \neq \varnothing$ is only a sufficient condition, and simultaneously, if $\mathcal{Q} \cap \mathcal{F} = \varnothing$, then the zero duality gap property may still hold. We give the following example to explain these cases.

Example 3.3 Consider Example 3.2. Obviously, we have the feasible set $R = [-1, 1]$. Let $\Pi = \{1\} \times \mathbb{R}^2_+$. Next, we consider two special regular weak separation functions w, defined on $\mathbb{R}^3 \times \Pi$.

Case 1. Let $w(u, v; 1, \lambda) = u + \sup_{z \in \{v\} - \mathbb{R}^2_+} (\langle \lambda, z \rangle - \|z\|_1)$ for all $(u, v) \in \mathbb{R}^3$ and all $\lambda \in \mathbb{R}^2_+$. Then w is a regular weak separation function satisfying Assumption $\mathcal{A}$. It is easy to verify that $R \times \{(1, 0, 0)\} \subset \mathcal{N}$. Specially, we take $R \times \{(1, 0, 0)\} = \mathcal{F}$, which implies $R \subset \mathcal{F}^o$. Now, take $\bar{x} = 0$ and $\hat{\pi} = (1, 1, 1) \in \Pi$. Then we have

$$\begin{aligned} L_w(x, \hat{\pi}) &= -x^2 - \sup_{z_1 \le 1 - x^2, z_2 \le -x^2 - x + 2} \big(z_1 + z_2 - |z_1| - |z_2|\big) \\ &= \begin{cases} 3x^2 + 2x - 6, & \text{if } x \le -2 \text{ or } x \ge 1, \\ x^2 - 2, & \text{if } -2 \le x \le -1, \\ -x^2, & \text{if } -1 \le x \le 1. \end{cases} \end{aligned}$$

Thus, we have $\mathcal{I}_{\hat{\pi}} = \{-1, 1\}$. If we take $\mathcal{Q} = \{(-1, (1, 1, 1)), (1, (1, 1, 1))\}$, then it follows $\mathcal{Q} \subset \mathcal{M}$ and $\mathcal{Q} \cap \mathcal{F} = \varnothing$. Moreover, by directly calculating, we get

$$\sup_{(x,\pi) \in \mathcal{Q}} L_w(x, \pi) = -1 = \inf_{x \in R} f(x),$$

which shows that the zero duality gap property holds.

Case 2. Let $w(u, v; 1, \lambda) = u + \langle \lambda, v\rangle$ for all $(u, v) \in \mathbb{R}^3$ and all $\lambda \in \mathbb{R}^2_+$. Then w is a regular weak separation function satisfying Assumption $\mathcal{A}$, and simultaneously, $R \times \{(1, 0, 0)\} \subset \mathcal{N}$ holds. Specially, let $\mathcal{F} = \mathcal{N}$, and then we have $R \subset \mathcal{F}^o$. Take arbitrary $\tilde{\lambda} \in \mathbb{R}^2_+$ with $\tilde{\lambda}_1 \geq 2$ and $\tilde{\lambda}_2 = 0$. Then we get $L_w(x; 1, \tilde{\lambda}) = (\tilde{\lambda}_1 - 1)x^2 - \tilde{\lambda}_1$. It is easy to verify that $0 \in \mathcal{I}_{(1,\tilde{\lambda})}$, which implies the set $\mathcal{Q} := \{(0, (1, \tilde{\lambda})) \in \mathbb{R} \times \Pi \mid \tilde{\lambda}_1 \geq 2, \tilde{\lambda}_2 = 0\} \subset \mathcal{M}$. Moreover, by directly calculating, we get

$$L_w(0; 1, \lambda) = -\lambda_1 - 2\lambda_2, \quad \forall (1, \lambda) \in \Pi,$$

which implies $\mathcal{J}_0 = \{(1, 0, 0)\}$. Thus, we can conclude that $\mathcal{Q} \cap \mathcal{F} = \varnothing$. Simultaneously, we have

$$\sup_{(x,(1,\lambda))\in\mathcal{Q}} L_w(x; 1, \lambda) = -2 < -1 = \inf_{x\in R} f(x),$$

that is, the zero duality gap property does not hold.

In the following, we discuss the zero duality gap property by virtue of a perturbation function. To this end, we need to introduce some standard notions related with the perturbed optimization problem of (P).

Definition 3.4 For every $y \in \mathbb{R}^m$, let $R(y) := \{x \in X \mid g_i(x) \geq y_i, i = 1, 2, \ldots, m\}$. The problem

$$\inf_{x\in R(y)} f(x)$$

is called the perturbed optimization problem of (P). The perturbation function $p : \mathbb{R}^m \to \mathbb{R} \cup \{\pm\infty\}$, associated with (P), is defined as the optimal value map of the perturbed optimization problem, that is,

$$p(y) := \inf_{x\in R(y)} f(x), \quad \forall y \in \mathbb{R}^m.$$

Remark 3.5 Obviously, $p(0_{\mathbb{R}^m}) = \inf_{x\in R} f(x)$. Thus, if Assumption $\mathcal{A}$ holds, then it follows from Lemma 3.2 and Theorem 3.1(i) that the zero duality gap property with respect to w can be expressed in the form $p(0_{\mathbb{R}^m}) = \alpha$, and $p(0_{\mathbb{R}^m}) \geq \alpha$ always holds. Therefore, the zero duality gap property with respect to w can be reformulated as $p(0_{\mathbb{R}^m}) \leq \alpha$.

In order to establish a necessary condition for the zero duality gap property with respect to w, we consider the subclass $W_R^{\ell s}(\Pi) \subset W_R(\Pi)$ satisfying the following monotonicity assumption:

Assumption $\mathcal{B}$. $w \in W_R^{\ell s}(\Pi)$ and w is monotone increasing with respect to the first argument, that is,

$$\forall (u^1, v^1), (u^2, v^2) \in \mathbb{R}^{1+m} \quad \text{with } (u^2, v^2) \in \{(u^1, v^1)\} - \mathbb{R}^{1+m}_+$$

implies

$$w(u^2, v^2; \pi) \leq w(u^1, v^1; \pi), \quad \forall \pi \in \Pi.$$

We will show in Sect. 4 of [34] that there are some major classes of regular weak separation functions satisfying Assumption $\mathcal{B}$. By Assumptions $\mathcal{A}$ and $\mathcal{B}$, we give the following necessary condition for the zero duality gap property by means of the perturbation function.

Theorem 3.3 *Given the regular weak separation function $w \in W_R(\Pi)$ satisfying Assumptions $\mathcal{A}$ and $\mathcal{B}$, if the zero duality gap property with respect to w holds, then the perturbation function p is lower semicontinuous at $0_{\mathbb{R}^m}$.*

Proof Assume that p is not lower semicontinuous at $0_{\mathbb{R}^m}$. Then $\liminf_{y\to 0_{\mathbb{R}^m}} p(y) < p(0_{\mathbb{R}^m})$. Thus, there exist some $\varepsilon > 0$ and a sequence $\{y^k\} \subset \mathbb{R}^m$ such that $y^k \to 0_{\mathbb{R}^m}$ and

$$p\big(y^k\big) \le p(0_{\mathbb{R}^m}) - \varepsilon, \quad \forall k \in \mathbb{N}. \tag{16}$$

Since the zero duality gap property with respect to w holds, we have

$$p(0_{\mathbb{R}^m}) = \alpha = \sup_{(x,\pi)\in\mathcal{Q}} L_w(x,\pi) = \sup_{\pi\in\mathcal{Q}^\circ} \inf_{x\in X} L_w(x,\pi).$$

For $\varepsilon > 0$, there exists some $\widehat{\pi} \in \mathcal{Q}^\circ$ such that

$$\begin{aligned} p(0_{\mathbb{R}^m}) - \frac{\varepsilon}{3} &\le \inf_{x\in X} L_w(x,\widehat{\pi}) \\ &\le \inf_{x\in R(y^k)} L_w(x,\widehat{\pi}), \quad \forall k \in \mathbb{N}. \end{aligned}$$

Moreover, by the definition of the perturbation function p, we get

$$p\big(y^k\big) = \inf_{x\in R(y^k)} f(x), \quad \forall k \in \mathbb{N}.$$

Then, for every $k \in \mathbb{N}$, there exists some $x^k \in R(y^k)$ such that $f(x^k) \le p(y^k) + \frac{\varepsilon}{3}$. Together with Assumptions $\mathcal{A}$ and $\mathcal{B}$, (16) and $y^k \in \{g(x^k)\} - \mathbb{R}^m_+$, we get

$$\begin{aligned} p(0_{\mathbb{R}^m}) - \frac{\varepsilon}{3} &\le \inf_{x\in R(y^k)} L_w(x,\widehat{\pi}) \\ &\le L_w\big(x^k,\widehat{\pi}\big) = w(1,0_{\mathbb{R}^m};\widehat{\pi}) f(\bar{x}) - w\big(f(\bar{x}) - f\big(x^k\big), g\big(x^k\big);\widehat{\pi}\big) \\ &\le f(\bar{x}) - w\left(f(\bar{x}) - p\big(y^k\big) - \frac{\varepsilon}{3}, y^k;\widehat{\pi}\right) \\ &\le f(\bar{x}) - w\left(f(\bar{x}) - p(0_{\mathbb{R}^m}) + \frac{2\varepsilon}{3}, y^k;\widehat{\pi}\right), \quad \forall k \in \mathbb{N}. \end{aligned}$$

Take lim sup as $k \to +\infty$ on both sides, and we have

$$p(0_{\mathbb{R}^m}) - \frac{\varepsilon}{3} \le f(\bar{x}) - \liminf_{k\to+\infty} w\left(f(\bar{x}) - p(0_{\mathbb{R}^m}) + \frac{2\varepsilon}{3}, y^k;\widehat{\pi}\right).$$

Since $y^k \to 0_{\mathbb{R}^m}$, $w \in W_R^{\ell s}(\Pi)$ and Assumption $\mathcal{A}$ holds, we have

$$\begin{aligned}
p(0_{\mathbb{R}^m}) - \frac{\varepsilon}{3} &\le f(\bar{x}) - \liminf_{y \to 0_{\mathbb{R}^m}} w\left(f(\bar{x}) - p(0_{\mathbb{R}^m}) + \frac{2\varepsilon}{3}, y; \widehat{\pi}\right) \\
&= f(\bar{x}) - w\left(f(\bar{x}) - p(0_{\mathbb{R}^m}) + \frac{2\varepsilon}{3}, 0_{\mathbb{R}^m}; \widehat{\pi}\right) \\
&\le f(\bar{x}) - \inf_{\pi \in \Pi} w\left(f(\bar{x}) - p(0_{\mathbb{R}^m}) + \frac{2\varepsilon}{3}, 0_{\mathbb{R}^m}; \pi\right) \\
&= f(\bar{x}) - \left(f(\bar{x}) - p(0_{\mathbb{R}^m}) + \frac{2\varepsilon}{3}\right) \\
&= p(0_{\mathbb{R}^m}) - \frac{2\varepsilon}{3},
\end{aligned} \tag{17}$$

where the first inequality holds since $\{y^k\} \subset \mathbb{R}^m$, $y^k \to 0_{\mathbb{R}^m}$ and

$$\liminf_{y \to 0_{\mathbb{R}^m}} w\left(f(\bar{x}) - p(0_{\mathbb{R}^m}) + \frac{2\varepsilon}{3}, y; \widehat{\pi}\right) \le \liminf_{k \to +\infty} w\left(f(\bar{x}) - p(0_{\mathbb{R}^m}) + \frac{2\varepsilon}{3}, y^k; \widehat{\pi}\right).$$

Obviously, (17) is impossible. Thus, we complete the proof. □

Conversely, in order to obtain the sufficient condition for the zero duality gap property by virtue of the perturbation function, we need the following assumption associated with the image $\mathcal{K}$ of (P) and the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ with the family of parameters specified by $\{1\} \times \mathbb{R}^m_+$:

Assumption $\mathcal{C}$. There exists some $\alpha > 0$ such that

$$w(u, v; \pi) \le \min\{u, \alpha\pi_1 v_1, \alpha\pi_2 v_2, \dots, \alpha\pi_m v_m\}$$

for every $(u, v) \in \mathcal{K}$ and every $\pi = (1, \pi_1, \pi_2, \dots, \pi_m) \in \Pi = \{1\} \times \mathbb{R}^m_+$.

Next, we establish the following sufficient condition for the zero duality gap property by virtue of the perturbation function.

Theorem 3.4 *Assume that $\mathcal{Q} = \mathcal{M}$. Let the regular weak separation function $w \in \mathcal{W}_R(\Pi)$ satisfy Assumptions $\mathcal{A}$ and $\mathcal{C}$. If the perturbation function p is lower semicontinuous at $0_{\mathbb{R}^m}$, then the zero duality gap property with respect to w holds.*

Proof Assume that the zero duality gap property with respect to w does not hold. By Assumption $\mathcal{A}$, $\mathcal{Q} = \mathcal{M}$ and Theorem 3.1(i), we have $\inf_{x \in R} f(x) > \sup_{\pi \in \Pi} \inf_{x \in X} L_w(x, \pi)$. Then there exists some $\delta > 0$ such that

$$p(0_{\mathbb{R}^m}) > \inf_{x \in X} L_w(x, \pi) + \delta, \quad \forall \pi \in \Pi.$$

Take $\pi^k = (1, k, k, \dots, k) \in \Pi$ with $k \in \mathbb{N}$. We have

$$p(0_{\mathbb{R}^m}) - \frac{\delta}{2} > \inf_{x \in X} L_w\left(x, \pi^k\right) + \frac{\delta}{2}, \quad \forall k \in \mathbb{N}.$$

Then, for every $k \in \mathbb{N}$, there exists $x^k \in X$ such that

$$p(0_{\mathbb{R}^m}) - \frac{\delta}{2} > L_w\big(x^k, \pi^k\big)$$
$$= w\big(1, 0_{\mathbb{R}^m}; \pi^k\big) f(\bar{x}) - w\big(f(\bar{x}) - f\big(x^k\big), g\big(x^k\big); \pi^k\big), \quad \forall k \in \mathbb{N}.$$

Together with Assumptions $\mathcal{A}$ and $\mathcal{C}$, and $(f(\bar{x}) - f(x^k), g(x^k)) \in \mathcal{K}$ for all $k \in \mathbb{N}$, there exists some $\alpha > 0$ such that

$$\begin{aligned} &p(0_{\mathbb{R}^m}) - \frac{\delta}{2} \\ &\quad > f(\bar{x}) - \min\big\{f(\bar{x}) - f\big(x^k\big), \alpha k g_1\big(x^k\big), \alpha k g_2\big(x^k\big), \dots, \alpha k g_m\big(x^k\big)\big\} \\ &\quad = \max\big\{f\big(x^k\big), f(\bar{x}) - \alpha k g_1\big(x^k\big), f(\bar{x}) - \alpha k g_2\big(x^k\big), \dots, f(\bar{x}) - \alpha k g_m\big(x^k\big)\big\}. \end{aligned}$$

Thus, we have $g_i(x^k) \geq \beta^k, \forall i = 1, 2, \dots, m$, where $\beta^k := \frac{p(0_{\mathbb{R}^m}) - \frac{\delta}{2} - f(\bar{x})}{-\alpha k}, \forall k \in \mathbb{N}$. Let $y^k \in \mathbb{R}^m$ with $y_i^k = \beta^k, \forall i = 1, 2, \dots, m$. Then it follows that $x^k \in R(y^k)$, which implies $p(y^k) \leq f(x^k)$. Moreover, $y^k \to 0_{\mathbb{R}^m}$ and $p(0_{\mathbb{R}^m}) - \frac{\delta}{2} > f(x^k)$. Therefore, we can conclude that

$$p(0_{\mathbb{R}^m}) - \frac{\delta}{2} \geq \liminf_{k \to +\infty} p\big(y^k\big) \geq \liminf_{y \to 0_{\mathbb{R}^m}} p(y) = p(0_{\mathbb{R}^m})$$

since p is lower semicontinuous at $0_{\mathbb{R}^m}$. This is a contradiction. □

Now, we can immediately establish the following equivalent characterization of the zero duality gap property by virtue of the perturbation function from Theorems 3.3 and 3.4.

Corollary 3.3 *Assume that $\mathcal{Q} = \mathcal{M}$. Let the regular weak separation function $w \in W_R(\Pi)$ satisfy Assumptions $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. Then the zero duality gap property with respect to w holds if and only if the perturbation function p is lower semicontinuous at $0_{\mathbb{R}^m}$.*

We will see in Sect. 4 of [34] that there exist some kinds of regular weak separation functions satisfying Assumption $\mathcal{C}$. However, Assumption $\mathcal{C}$ is very restrictive when the regular weak separation function w, defined on $\mathbb{R}^{1+m} \times (\{1\} \times \mathbb{R}^m_+)$, reduces to some special cases; for example, just as shown in [28, 29, 31], it is not valid for the classic Lagrange function $L_w(x; 1, \lambda) = f(x) - \langle \lambda, g(x) \rangle$ when we take $w(u, v; 1, \lambda) = u + \langle \lambda, v \rangle$. Simultaneously, it is also not valid for the augmented Lagrange function $L_w(x; 1, \lambda) = f(x) - \sup_{z \in \{g(x)\} - \mathbb{R}^m_+}(\langle \lambda, z \rangle - r\sigma(z))$ when we take $w(u, v; 1, \lambda) = u + \sup_{z \in \{v\} - \mathbb{R}^m_+}(\langle \lambda, v \rangle - r\sigma(z))$ (see more details in Sect. 4, Case 2, of [34]). It is worth noting that, if there exists some $\eta \in \mathbb{R}$ such that $\inf_{x \in X} f(x) \geq \eta$, i.e., f is bounded from below on X, then Theorem 3.4 can be applied to the following penalization problem $(\widetilde{\mathrm{P}})$, which is equivalent to (P) in the sense that $(\widetilde{\mathrm{P}})$ and (P) have the same feasible sets and the same objective functions. Consider the penalization problem

$$(\widetilde{\mathrm{P}}) \quad \min f(x), \quad \text{s.t.} \quad x \in X, \quad \big[g_i(x)\big]_- \geq 0, \quad i = 1, 2, \ldots, m,$$

where $[g_i(x)]_- := \min\{0, g_i(x)\}$. Note that $\bar{x} \in X$ is arbitrary. Similarly, we can replace $f(\bar{x})$ with η in Theorem 3.4. For more details, we refer to [33, 35, 36]. Then the image of $(\widetilde{\mathrm{P}})$ has the form

$$\mathcal{K} := \big\{(u, v) \in \mathbb{R}^{1+m} \mid u = \eta - f(x), v = \big[g(x)\big]_-, x \in X\big\},$$

where $[g(x)]_- := ([g_1(x)]_-, [g_2(x)]_-, \ldots, [g_m(x)]_-) \in \mathbb{R}^{1+m}$. Obviously, we have $\mathcal{K} \subset -\mathbb{R}^{1+m}_+$. Then, for the penalization problem $(\widetilde{\mathrm{P}})$, we have the following reformulation corresponding to Assumption $\mathcal{C}$:

Assumption $\widetilde{\mathcal{C}}$. There exists some $\alpha > 0$ such that

$$w(u, v; \pi) \leq \min\{u, \alpha\pi_1 v_1, \alpha\pi_2 v_2, \ldots, \alpha\pi_m v_m\}$$

for every $(u, v) \in \mathcal{K} \subset -\mathbb{R}^{1+m}_+$ and every $\pi = (1, \pi_1, \pi_2, \ldots, \pi_m) \in \Pi = \{1\} \times \mathbb{R}^m_+$.

As a result, Assumption $\widetilde{\mathcal{C}}$ is not so restrictive for the regular weak separation function w defined on $(-\mathbb{R}^{1+m}_+) \times (\{1\} \times \mathbb{R}^m_+)$. Moreover, we will show in Sect. 4 of [34] that Assumption $\widetilde{\mathcal{C}}$ is valid for the classic Lagrange penalization function $L_w(x; 1, \lambda) = f(x) - \langle\lambda, [g(x)]_-\rangle$ when we take $w(u, v; 1, \lambda) = u + \langle\lambda, v\rangle$. At the same time, Assumption $\widetilde{\mathcal{C}}$ is valid for the classic augmented Lagrange penalization function $L_w(x; 1, \lambda) = f(x) - \sup_{z \in \{[g(x)]_-\} - \mathbb{R}^m_+}(\langle\lambda, z\rangle - r\sigma(z))$ when we take $w(u, v; 1, \lambda) = u + \sup_{z \in \{v\} - \mathbb{R}^m_+}(\langle\lambda, v\rangle - r\sigma(z))$.

4 Conclusions

Recently, the introduction of ISA has shown that the IS associated with the problem provides a natural environment for the Lagrange ideas. Moreover, the classic Lagrange duality has been redescribed by means of the regular weak separation between two sets, $\mathcal{K}$ and $\mathcal{H}$, in the IS. In this paper, we present a unified duality scheme for a constrained extremum problem by virtue of the ISA. Specially, we establish an equivalent characterization not only for a Lagrange saddle point, but also for a Lagrange multiplier by using the existence of a regular weak separation in the IS, which extend and improve some existing results. At the same time, we show that the existence of a regular weak separation in the IS is sufficient and necessary for the zero duality gap property. Finally, we propose some equivalent conditions for the zero duality gap property in terms of the lower semicontinuity of the perturbation function.

As far as we know, the Lagrangian-type (exact and inexact) penalty methods are closely related with duality theories. Just as shown in this paper, the ISA provides a unified approach, namely a regular weak separation in the IS, to analyze and deduce the general duality scheme. Thus, it is very interesting and valuable to further investigate corresponding penalty methods, especially to study the relationships between the (exact and inexact) penalty and the existences of some appropriate separations in the IS, by means of ISA, and we will make effort to do it in our future work.

Acknowledgements The authors are indebted to Professor F. Giannessi for his helpful advice and valuable discussions, especially for providing references [3, 9, 13] and suggestions on the form of the generalized Lagrange function (5). They would also like to express their gratitude to two anonymous referees for their valuable comments and suggestions on the proof of Lemma 3.1. This research was supported by the National Natural Science Foundation of China (Grant: 11171362) and the Basic and Advanced Research Project of CQCSTC (Grant: cstc2013jcyjA00003).

References

1. Castellani, G., Giannessi, F.: Decomposition of mathematical programs by means of theorems of alternative for linear and nonlinear systems. In: Proc. Ninth Internat. Math. Programming Sympos., Budapest. Survey of Mathematical Programming, pp. 423–439. North-Holland, Amsterdam (1979)
2. Giannessi, F.: Theorems of the alternative and optimality conditions. J. Optim. Theory Appl. **42**, 331–365 (1984)
3. Li, J., Feng, S.Q., Zhang, Z.: A unified approach for constrained extremum problems: Image space analysis. J. Optim. Theory Appl. **159**, 69–92 (2013)
4. Giannessi, F.: Constrained Optimization and Image Space Analysis, vol. 1: Separation of Sets and Optimality Conditions. Springer, New York (2005)
5. Giannessi, F.: Theorems of the alternative for multifunctions with applications to optimization: gengeral results. J. Optim. Theory Appl. **55**, 233–256 (1987)
6. Rubinov, A.M., Uderzo, A.: On global optimality conditions via separation functions. J. Optim. Theory Appl. **109**, 345–370 (2001)
7. Li, S.J., Xu, Y.D., Zhu, S.K.: Nonlinear separation approach to constrained extremum problems. J. Optim. Theory Appl. **154**, 842–856 (2012)
8. Luo, H.Z., Mastroeni, G., Wu, H.X.: Separation approach for augmented Lagrangians in constrained nonconvex optimization. J. Optim. Theory Appl. **144**, 275–290 (2010)
9. Luo, H.Z., Wu, H.X., Liu, J.Z.: Some results on augmented Lagrangians in constrained global optimization via image space analysis. J. Optim. Theory Appl. **159**, 360–385 (2013)
10. Tardella, F.: On the image of a constrained extremum problem and some applications to the existence of a minimum. J. Optim. Theory Appl. **60**, 93–104 (1989)
11. Giannessi, F., Mastroeni, G.: Separation of sets and Wolfe duality. J. Glob. Optim. **42**, 401–412 (2008)
12. Mastroeni, G.: Some applications of the image space analysis to the duality theory for constrained extremum problems. J. Glob. Optim. **46**, 603–614 (2010)
13. Giannessi, F.: On the theory of Lagrangian duality. Optim. Lett. **1**, 9–20 (2007)
14. Giannessi, F., Mastroeni, G., Yao, J.C.: On maximum and variational principles via image space analysis. Positivity **16**, 405–427 (2012)
15. Pappalardo, M.: Image space approach to penalty methods. J. Optim. Theory Appl. **64**, 141–152 (1990)
16. Mastroeni, G.: Nonlinear separation in the image space with applications to penalty methods. Appl. Anal. **91**, 1901–1914 (2012)
17. Mastroeni, G., Pappalardo, M., Yen, N.D.: Image of a parametric optimization problem and continuity of the perturbation function. J. Optim. Theory Appl. **81**, 193–202 (1994)
18. Dien, P.H., Mastroeni, G., Pappalardo, M., Quang, P.H.: Regularity conditions for constrained extremum problems via image space. J. Optim. Theory Appl. **80**, 19–37 (1994)
19. Giannessi, F., Mastroeni, G., Pellegrini, L.: On the theory of vector optimization and variational inequalities image space analysis and separation. In: Giannessi, F. (ed.) Vector Variational Inequalities and Vector Equilibria, Mathematical Theories, pp. 153–215. Kluwer Academic, Dordrecht (2000)
20. Mastroeni, G.: On the image space analysis for vector quasi-equilibrium problems with a variable ordering relation. J. Glob. Optim. **53**, 203–214 (2012)
21. Mastroeni, G.: Optimality conditions and image space analysis for vector optimization problems. In: Ansari, Q.H., Yao, J.-C. (eds.) Recent Developments in Vector Optimization, Vector Optimization, vol. 1, pp. 169–220. Springer, Dordrecht (2012)

22. Li, J., Huang, N.J.: Image space analysis for vector variational inequalities with matrix inequality constraints and applications. J. Optim. Theory Appl. **145**, 459–477 (2010)
23. Mastroeni, G., Panicucci, B., Passacantando, M., Yao, J.C.: A separation approach to vector quasi-equilibrium problems: saddle point and gap functions. Taiwan. J. Math. **13**, 657–673 (2009)
24. Hestenes, M.R.: Multiplier and gradient methods. In: Zadeh, L.A., Neustadt, L.W., Balakrishnan, A.V. (eds.) Computing Methods in Optimization Problems. Academic Press, New York (1969)
25. Hestenes, M.R.: Multiplier and gradient methods. J. Optim. Theory Appl. **4**, 303–320 (1969)
26. Powell, M.J.D.: A method for nonlinear constraints in minimization problems. In: Fletcher, R. (ed.) Optimization. Academic Press, New York (1972)
27. Rubinov, A.M., Glover, B.M., Yang, X.Q.: Decreasing functions with applications to penalization. SIAM J. Optim. **10**, 289–313 (1999)
28. Yang, X.Q., Huang, X.X.: A nonlinear Lagrangian approach to constrained optimization problems. SIAM J. Optim. **14**, 1119–1144 (2001)
29. Rubinov, A.M., Huang, X.X., Yang, X.Q.: The zero duality gap property and lower semicontinuity of the perturbation function. Math. Oper. Res. **27**, 775–791 (2002)
30. Wang, C.Y., Yang, X.Q., Yang, X.M.: Unified nonlinear Lagrangian approach to duality and optimal paths. J. Optim. Theory Appl. **135**, 85–100 (2007)
31. Huang, X.X., Yang, X.Q.: A unified augmented Lagrangian approach to duality and exact penalization. Math. Oper. Res. **28**, 533–552 (2003)
32. Rockafellar, R.T., Wets, R.J.B.: Variational Analysis. Springer, Berlin (1998)
33. Rubinov, A.M., Yang, X.Q.: Lagrange-Type Functions in Constrained Non-convex Optimization. Kluwer Academic, Dordrecht (2003)
34. Zhu, S.K., Li, S.J.: Unified duality theory for constrained extremum problems. Part II: Special duality schemes. J. Optim. Theory Appl. (2013). doi:10.1007/s10957-013-0467-5
35. Evtushenko, Y.G., Rubinov, A.M., Zhadan, V.G.: General Lagrange-type functions in constrained global optimization. I. Auxiliary functions and optimality conditions. Optim. Methods Softw. **16**, 193–230 (2001)
36. Evtushenko, Y.G., Rubinov, A.M., Zhadan, V.G.: General Lagrange-type functions in constrained global optimization. II. Exact auxiliary functions. Optim. Methods Softw. **16**, 231–256 (2001)