Existence of Augmented Lagrange Multipliers for Semi-infinite Programming Problems[1]

R.S. Burachik

School of Information Technology and Mathematical Sciences, University of South

Australia

e-mail: Regina.Burachik@unisa.edu.au

X.Q. Yang

Department of Applied Mathematics, The Hong Kong Polytechnic University,

Kowloon, Hong Kong

e-mail: mayangxq@polyu.edu.hk

Y.Y. Zhou

Department of Mathematics, Soochow University, Suzhou 215006, China

e-mail: yuyingz@suda.edu.cn

This paper is dedicated to the memory of the late Professor Vladimir Demyanov.

Abstract. Using an augmented Lagrangian approach, we study the existence of augmented Lagrange multipliers of a semi-infinite programming problem (SIP, in short) and discuss their characterizations in terms of saddle points. In the case of a sharp Lagrangian, we obtain a first-order necessary condition for the existence of an augmented Lagrange multiplier for (SIP) and some first-order sufficient conditions by assuming inf-compactness of the data functions and the MFCQ. Using a valley at 0 augmenting function and assuming suitable second-order sufficient conditions, we obtain the existence of an augmented Lagrange multiplier for (SIP).

Key words: Semi-infinite programming; augmented Lagrange multiplier; optimality conditions; sharp Lagrangian; a valley at 0 augmenting function.

AMS Subject Classification: 90C26, 90C34, 90C46

1 Introduction

We consider a semi-infinite programming problem (SIP, in short) in the following form

$$(\text{SIP}) \min_{x \in \mathbb{R}^n} f(x) \ \text{ subject to } g(x, \omega) \leq 0, \omega \in \Omega,$$

where Ω is a (possibly infinite) nonempty set, $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \times \Omega \to \mathbb{R}$ are real-valued functions. (SIP) is said to be convex if f and $g(\cdot, \omega)(\omega \in \Omega)$ are convex.

(SIP) has attracted much attention due to its various applications in engineering design, optimal control, economic equilibria etc. It has become an active field of research in applied mathematics. For example, see Fiacco and Kortanek [1], Goberna and López [2], Polak [3], Polak and Wuu [4], Polak [5] and the references therein. The main difficulty in solving (SIP) comes from the fact that it has an infinite number of constraints, and hence it is hard to design algorithms for solving it. However, in recent years, many effective methods have been proposed for solving semi-infinite programming, such as discretization methods (see Hettich [6], Still [7]), reduction methods (see Hettich and Kortanek [8], Reemsten and Görner [9]) and exchange methods (see Goberna and López [2], Hettich and Kortanek [8], Reemsten and Görner [9], Zhang et al. [10]). For studying duality properties of (SIP), we will use an augmented Lagrangian approach, which is crucial for dealing with nonconvex and constrained optimization problems. The augmented Lagrangian function was first proposed independently by Hestenes [11] and Powell [12] to solve an optimization problem with equality constraints. This was extended to solve an inequality constrained optimization problem by Buys [13]. The augmented Lagrangian method has been used to solve an optimization problem with both equality and inequality constraints in Rockafellar [14]. Under a growth condition, one can relate the value of the dual problem to the behavior of the perturbation function. In Rockafellar [14], the definition of a quadratic growth condition was introduced and a necessary and sufficient condition for a zero duality gap between a constrained optimization problem and its augmented Lagrangian dual problem was obtained. In Rockafellar and Wets [15, 11K*], an augmented Lagrangian with a convex augmenting function was introduced and the corresponding zero duality property was established. The definitions of several growth conditions were delineated by Penot [16]. By assuming that the augmenting function satisfies the level coercivity and the perturbation function satisfies some growth condition, several zero duality gap results between the primal problem and its augmented Lagrangian dual problem were obtained in Penot [16]. A level-bounded augmenting function was given in Huang and Yang [17], where the convexity of augmenting functions in Rockafellar and Wets [15] was replaced by a level-boundedness condition. A valley at 0 augmenting function

was given in Zhou and Yang [18], where the level-boundedness condition of the augmenting functions in Huang and Yang [17] was replaced by the valley at 0 property. The latter property allows one to introduce a duality scheme with zero duality gap. More general frameworks for establishing zero duality gap and the existence of exact penalty parameters were introduced in [19, 20, 21]. Optimization problems in which the constraints are given in terms of a set inclusion were considered by Shapiro and Sun [22]. Using a quadratic growth condition (a property originally introduced by Rockafellar [14]), and a quadratic augmented Lagrangian, Shapiro and Sun established strong duality and second-order conditions ensuring existence of Lagrange multipliers (see [22, Theorem 2.4, Theorem 3.4]). The concept of augmented Lagrange multiplier is related with that of a global saddle point of augmented Lagrangian functions. This connection was exploited by Sun et al. [23], who considered four classes of augmented Lagrangian functions. Under second-order sufficient conditions and certain additional conditions, Sun et al. [23] established the existence of local and global saddle points of these four kinds of augmented Lagrangian functions. Rückmann and Shapiro [24] studied the (SIP) via an augmented Lagrangian approach with a nonnegative convex augmenting function and a reduction approach method. They obtained some necessary and sufficient conditions for the existence of an augmented Lagrange multiplier. Furthermore, they discussed two particular cases: the proximal Lagrangian and the sharp Lagrangian.

Sufficient conditions for the existence of an augmented multiplier of an augmented Lagrangian with a valley at 0 augmenting function of a cone constrained optimization problem were obtained in Zhou et al [25] by using local saddle point conditions and bounded conditions of a perturbed constraint set. In the present paper we will use the augmented Lagrangian to study (SIP). Firstly, we use the sharp Lagrangian to derive first-order necessary conditions and first-order sufficient conditions for the existence of an augmented Lagrange multiplier for (SIP) (see Theorems 3.1 and 3.2). Secondly, we consider an augmented Lagrangian with an augmenting function having a valley at 0 property. This allows us to obtain the existence of an augmented Lagrange multiplier for (SIP) under two possible assumptions: (i) that (SIP) has a unique global optimal solution, and (ii) that the objective and the constraint functions satisfy a generalized representation condition originally introduced in [26].

Since the augmenting function used in this paper may be nonconvex and nonquadratic, standard techniques such as those used in Rückmann and Shapiro [24] and in Shapiro and Sun [22] to show the existence of augmented Lagrange multipliers cannot be applied here. For the same reasons, the saddle point results we derive cannot make use of the techniques used in Sun et al [23]. As a result, both the results and the

proof techniques we use in the present paper are different from those in the last three references.

The structure of the paper is outlined as follows: In Section 2, we introduce some notations and preliminary results needed in the sequel of this article. In particular, we show that the augmented Lagrangian is lower semicontinuous by following a technique developed in [19]. In Section 3, we obtain some first-order optimality conditions for the existence of an augmented Lagrange multiplier of the sharp Lagrangian problem. In Section 4, we obtain some second-order conditions for the existence of an augmented Lagrange multiplier of a valley at 0 augmented Lagrangian for (SIP).

2 Preliminaries

Let $\mathbb{R}_+ := [0, +\infty), \mathbb{R}_{++} := (0, +\infty[, \mathbb{R}_{\pm\infty} :=] - \infty, +\infty[$ and the indicator function of a set $Z \subset \mathbf{Y}$ be defined by $\delta_Z(y) = 0$, if $y \in Z$, and $\delta_Z(y) = +\infty$, otherwise.

Let Ω be a nonempty set and call $\mathbb{R}^\Omega$ be the set of all functions from Ω to $\mathbb{R}$. Denote by $\mathbb{R}^{(\Omega)}$ the subset of $\mathbb{R}^\Omega$ consisting of all functions from Ω to $\mathbb{R}$ which are nonzero in at most a finite subset of Ω. More precisely, $\mathbb{R}^{(\Omega)}$ is defined as follows.

$$\mathbb{R}^{(\Omega)} := \{\lambda : \Omega \to \mathbb{R} : \ \lambda(\omega) = 0, \text{ for all } \omega \in \Omega \text{ except for a finite number}\}, \qquad (1)$$

and $\mathbb{R}_+^{(\Omega)} := \{\lambda \in \mathbb{R}^{(\Omega)} : \lambda(\omega) \geq 0, \forall \omega \in \Omega\}$. For $\lambda \in \mathbb{R}^{(\Omega)}$, let

$$\mathrm{supp}(\lambda) := \{\omega \in \Omega : \lambda(\omega) \neq 0\}.$$

Therefore, we say that $\mathbb{R}^{(\Omega)}$ is the subset of $\mathbb{R}^\Omega$ consisting of all functions with *finite support*. The set $\mathbb{R}^{(\Omega)}$ will play the role of the dual space, and hence will contain the Lagrange multipliers.

For $\lambda \in \mathbb{R}^{(\Omega)}$ and $y \in \mathbb{R}^\Omega$, we can define a scalar product in a natural way:

$$\langle \lambda, y \rangle := \sum_{\omega \in \mathrm{supp}(\lambda)} \lambda(\omega) y(\omega).$$

For a finite set $\Gamma = \{\omega_1, \ldots, \omega_m\} \subset \Omega$ and $y \in \mathbb{R}^\Omega$, denote by $y^\Gamma(\cdot)$ the restriction of function $y(\cdot)$ to Γ, i.e.,

$$y^\Gamma(\omega) := \begin{cases} y(\omega), & \text{if } \omega \in \Gamma, \\ 0, & \text{otherwise.} \end{cases} \qquad (2)$$

By abuse of notation, we may use the same symbol y^Γ to denote the finite dimensional vector that has for coordinates $y^\Gamma = (y(\omega_1), \ldots, y(\omega_m))^T$.

Let $\sigma : \mathbb{R}^{(\Omega)} \to \mathbb{R}_+$ be a function with $\sigma(0) = 0$. We refer to the function σ as an *augmenting function.*

For a given $x \in \mathbb{R}^n$, we define the set

$$D(x) := \{y \in \mathbb{R}^\Omega \ : \ g(x, \omega) + y(\omega) \le 0 \, \forall \, \omega \in \Omega\}, \tag{3}$$

which is the set of "admissible" perturbations at the point x.

We will now use all these elements to define the augmented Lagrangian for (SIP).

Definition 2.1 *Fix $\lambda \in \mathbb{R}^{(\Omega)}$, and its corresponding support $\Gamma = supp(\lambda)$. Fix also $r > 0$, a penalty parameter. The* augmented Lagrangian *for the problem (SIP) is the function $L^\Omega : \mathbb{R}^n \times \mathbb{R}^{(\Omega)} \times \mathbb{R}_{++} \to \mathbb{R}_{\pm\infty}$ defined by*

$$
\begin{aligned}
L^\Omega(x, \lambda, r) \ &:= \ \inf_{y \in \mathbb{R}^\Omega} \{f(x) - \langle \lambda, y \rangle + r\sigma(y^\Gamma) : g(x, \omega) + y(\omega) \le 0, \omega \in \Omega\} \\
&= \ \inf_{y \in \mathbb{R}^\Omega} \{f(x) - \langle \lambda, y \rangle + r\sigma(y^\Gamma) + \delta_{D(x)}(y)\}.
\end{aligned}
\tag{4}
$$

Denote by val(SIP) the optimal value of the problem (SIP), and denote by X_0 the set of feasible solutions of (SIP), i.e.,

$$X_0 := \{x \in \mathbb{R}^n : g(x, \omega) \le 0, \omega \in \Omega\}. \tag{5}$$

Let $x^* \in X_0$. Denote the active constraint index set at x^* by

$$I(x^*) := \{\omega \in \Omega : g(x^*, \omega) = 0\}. \tag{6}$$

We assume throughout the paper that the optimal value of problem (SIP) is finite.

Fix $y \in \mathbb{R}^\Omega$ and denote by val(P_y) the optimal value of the following parameterized problem of (SIP):

$$(P_y) \quad \min f(x) \ \text{subject to} \quad x \in \mathbb{R}^n, \quad g(x, \omega) + y(\omega) \le 0, \ \omega \in \Omega. \tag{7}$$

With this notation, for $y = 0$ we naturally obtain val(P_0) = val(SIP).

Let $\lambda \in \mathbb{R}^{(\Omega)}$, $\Gamma = supp(\lambda)$ and $r > 0$. Define $v_r^\Omega : \mathbb{R}^\Omega \to \mathbb{R}_{\pm\infty}$ as

$$v_r^\Omega(y) := \text{val}(P_y) + r\sigma(y^\Gamma), \quad y \in \mathbb{R}^\Omega. \tag{8}$$

The assumption that $\sigma(0) = 0$ implies that $v_r^\Omega(0) = $ val(SIP) for every $r > 0$. By (4)-(8), we always have that

$$
\begin{aligned}
v_r^\Omega(0) \ge \ &\inf_{y \in \mathbb{R}^\Omega} \{v_r^\Omega(y) - \langle \lambda, y \rangle\} = \inf_{y \in \mathbb{R}^\Omega} \{\text{val}(P_y) - \langle \lambda, y \rangle + r\sigma(y^\Gamma)\} \\
&= \inf_{y \in \mathbb{R}^\Omega} \inf_{x \in \mathbb{R}^n} \{f(x) - \langle \lambda, y \rangle + r\sigma(y^\Gamma) : g(x, \omega) + y(\omega) \le 0, \ \omega \in \Omega\} \\
&= \inf_{x \in \mathbb{R}^n} \inf_{y \in \mathbb{R}^\Omega} \{f(x) - \langle \lambda, y \rangle + r\sigma(y^\Gamma) : g(x, \omega) + y(\omega) \le 0, \omega \in \Omega\} \\
&= \inf_{x \in \mathbb{R}^n} L^\Omega(x, \lambda, r).
\end{aligned}
\tag{9}
$$

When there exists $r > 0$ such that the above expression holds as an equality, we say that λ is an augmented Lagrange multiplier with penalty parameter r. We make this precise in the next definition.

Definition 2.2 $\lambda \in \mathbb{R}^{(\Omega)}$ *is said to be an* augmented Lagrange multiplier *for* (SIP) *if* val(SIP) *is finite and there exists* $r \geq 0$ *such that*

$$v_r^\Omega(y) \geq v_r^\Omega(0) + \langle \lambda, y \rangle, \quad \forall y \in \mathbb{R}^\Omega. \tag{10}$$

If (10) *holds for* r, *then we say that the augmented Lagrange multiplier* λ *has a* penalty parameter r.

Remark 2.1 *(i) If* (10) *holds for a given* $r > 0$, *then* λ *is an augmented Lagrange multiplier for* (SIP) *for every* $r' > r$. *This is a direct consequence of* (8) *and the fact that* σ *is nonnegative.*

(ii) It follows from (10) *that* λ *is an augmented Lagrange multiplier for* (SIP) *if and only if it satisfies an inequality resembling the subgradient inequality for convex functions. Indeed, recall that* $\lambda \in \partial v_r^\Omega(y_1)$ *if and only if* $v_r^\Omega(y) - v_r^\Omega(y_1) \geq \langle \lambda, y - y_1 \rangle, \forall y \in \mathbb{R}^\Omega$. *Note that* $\partial v_r^\Omega(0)$ *is well-defined only if* $v_r^\Omega(0) = $ *val*(SIP) *is finite.*

2.1 Discretizations and Lagrangian Functions for (SIP)

A finite set $\Gamma := \{\omega_1, \ldots, \omega_m\} \subset \Omega$ naturally induces a discretized version of (SIP), in which only the constraints corresponding to the elements of Γ are taken into account. This discretized problem is denoted (P^Γ), and defined as:

$$(P^\Gamma) \quad \min f(x) \text{ subject to } \quad x \in \mathbb{R}^n, \quad g(x, \omega_j) \leq 0, j = 1, \ldots, m.$$

Denote by S^Γ the set of optimal solutions of (P^Γ). As in (7), we can perturb the discretized problem (P^Γ), this time with $\bar{y} = (\bar{y}_1, \ldots, \bar{y}_m)^T \in \mathbb{R}^m$. This parameterized problem becomes:

$$(P_{\bar{y}}^\Gamma) \quad \min f(x) \text{ subject to } \quad x \in \mathbb{R}^n, \quad g(x, \omega_j) + \bar{y}_j \leq 0, j = 1, \ldots, m. \tag{11}$$

We denote by $v^\Gamma(\bar{y})$ the optimal value of $(P_{\bar{y}}^\Gamma)$. Recall that $y^\Gamma(\cdot)$ is the restriction of the function $y(\cdot)$ to Γ (cf. (2)). Define $\sigma^\Gamma(\bar{y}) := \sigma(y^\Gamma)$. This allows us to define, as in (8):

$$v_r^\Gamma(\bar{y}) := v^\Gamma(\bar{y}) + r\sigma^\Gamma(\bar{y}).$$

We define next the Lagrangians associated with the discretized problem $(P_{\bar{y}}^\Gamma)$.

Definition 2.3 *Given a finite set* $\Gamma := \{\omega_1, \ldots, \omega_m\} \subset \Omega$, *the* augmented Lagrangian *and the* classical Lagrangian *for problem* (P^Γ) *are defined respectively by*

$$L^\Gamma(x, \bar{\lambda}, r) := \inf_{\bar{y} \in \mathbb{R}^m} \{f(x) - \langle \bar{\lambda}, \bar{y} \rangle + r\sigma^\Gamma(\bar{y}) : g(x, \omega_j) + \bar{y}_j \leq 0, \ \omega_j \in \Gamma\}, \tag{12}$$

$$L_0^\Gamma(x, \bar{\lambda}) := f(x) + \sum_{w \in \Gamma} \bar{\lambda}(\omega) g(x, \omega), \tag{13}$$

where $x \in \mathbb{R}^n, \bar{\lambda} \in \mathbb{R}^m$ *and* $r \in \mathbb{R}_+$.

Remark 2.2 *Consider Problem* (11), *and let* $y \in \mathbb{R}^{(\Omega)}$ *be such that* $y(\omega_j) =: \bar{y}_j$, *for all* $\omega_j \in \Gamma$. *In this situation, the discretized problem* $(\mathrm{P_y})$ *in* (7) *is a relaxation of problem* $(\mathrm{P_y})$, *in the sense that the constraint set in* (11) *is larger than the one in* (7). *Namely,*

$$\text{If } y^{\Gamma} = \bar{y} \in \mathbb{R}^m, \quad \text{then} \quad \text{val}(P_y) \geq v^{\Gamma}(\bar{y}),$$

This fact and (8) *yield*

$$v_r^{\Omega}(y) = \text{val}(P_y) + r\sigma(y^{\Gamma}) \geq v_r^{\Gamma}(\bar{y}), \tag{14}$$

where $\bar{y}$ *is the restriction of* $y \in \mathbb{R}^{(\Omega)}$ *to the set* Γ.

We define next an augmented Lagrangian for the discretized problems.

Definition 2.4 *Fix* $\Gamma := \{\omega_1, \ldots, \omega_m\} \subset \Omega$, *a vector* $\lambda \in \mathbb{R}^m$ *is said to be an augmented Lagrange multiplier for* (P^{Γ}) *if* $\text{val}(\mathrm{P}^{\Gamma})$ *is finite and there exists* $r \geq 0$ *such that*

$$v_r^{\Gamma}(y) \geq v^{\Omega}(0) + \langle \lambda, y \rangle, \quad \forall y \in \mathbb{R}^m. \tag{15}$$

If (15) *holds for* r, *then we say that the augmented Lagrange multiplier* λ *has a* penalty parameter r.

The following is an example where the augmented Lagrangian for (P^{Γ}) with a separable augmenting function is computed explicitly.

Example 2.1 *If* σ^{Γ} *is separable, i.e.,* $\sigma^{\Gamma}(\bar{y}) = \sum_{j=1}^m \sigma_j(\bar{y}_j)$, *then the expression of the augmented Lagrangian can be simplified to*

$$
\begin{aligned}
L^{\Gamma}(x, \bar{\lambda}, r) &= \inf_{\bar{y} \in \mathbb{R}^m} \{f(x) - \langle \bar{\lambda}, \bar{y} \rangle + r \sum_{j=1}^m \sigma_j(\bar{y}_j) : g(x, \omega_j) + \bar{y}_i \leq 0, j = 1, \cdots, m\} \\
&= f(x) + r \sum_{j=1}^m \inf_{\bar{y}_j \in \mathbb{R}} \{-\frac{\bar{\lambda}_j}{r} \bar{y}_j + \sigma_j(\bar{y}_j) : \bar{y}_j \leq -g(x, \omega_j)\}.
\end{aligned}
$$

Assume furthermore that $\sigma_j(\bar{y}_j) = \max\{\bar{y}_j^2, \sqrt{|\bar{y}_j|}\}$ *for all* $j = 1, \ldots, m$, *and let* $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *be defined as* $\psi_j(a, b) = \inf\{-bz + \max\{z^2, \sqrt{|z|}\} : z \leq -a\}$, *then*

$$
\psi_j(a,b) = \begin{cases}
0, & \text{if } 0 \leq -a, -1 \leq b \leq 1 \text{ or } 0 \leq -a \leq \frac{1}{b^2}, 1 \leq b; \\
-|b| + 1, & \text{if } -1 \leq -a, b \leq -1 \text{ or } 1 \leq -a, 1 \leq b \\
& \quad \text{or } -1 \leq -a \leq -\left(1 + \frac{1}{b}\right)^2, -1 \leq b \leq -\frac{1}{2}; \\
ab + \max\{a^2, \sqrt{|a|}\}, & \text{if } -a \leq -1, b \leq -1 \text{ or } -a \leq 0, -\frac{1}{2} \leq b \leq 1 \\
& \quad \text{or } -a \in (-\infty, -1] \cup \left[-\left(1 + \frac{1}{b}\right)^2, 0\right], \\
& \quad -1 \leq b \leq -\frac{1}{2} \\
& \quad \text{or } -a \in (-\infty, 0] \cup \left[\frac{1}{b^2}, 1\right], 1 \leq b.
\end{cases}
$$

Thus, we obtain

$$L^{\Gamma}(x, \lambda, r) = f(x) + r \sum_{j=1}^{m} \psi_j(g(x, \omega_j), \frac{\lambda_j}{r}).$$

The following lemma shows that the optimal value of the discretized augmented Lagrangian is always equal to the augmented Lagrangian for (SIP). It also relates the discretized Lagrangian with the classical one. In the sequel, we will use the following notation: given a finite set $\Gamma := \{\omega_1, \ldots, \omega_m\} \subset \Omega$, and $(x, u) \in \mathbb{R}^n \times \mathbb{R}^m_+$, call $\gamma(x, u) \in \mathbb{R}^m$ the vector defined by

$$[\gamma(x, u)]_j = -g(x, \omega_j) - u_j, \text{ for } j = 1, \ldots, m. \tag{16}$$

Lemma 2.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \times \Omega \to \mathbb{R}$ be real-valued functions, and let $\lambda \in \mathbb{R}^{(\Omega)}$, $\sigma : \mathbb{R}^\Omega \to \mathbb{R}_+$ be a real-valued function with $\sigma(0) = 0$, $\Gamma = supp(\lambda) = \{\omega_1, \ldots, \omega_m\} \subset \Omega$ and $r > 0$. Then*

(i)

$$L^{\Omega}(x, \lambda, r) = L^{\Gamma}(x, \bar{\lambda}, r), \quad x \in \mathbb{R}^n, \tag{17}$$

where $\bar{\lambda} = (\lambda(\omega_1), \cdots, \lambda(\omega_m))^T$.

(ii) For every $x \in \mathbb{R}^n$, we have

$$\begin{aligned} L^{\Gamma}(x, \bar{\lambda}, r) &= f(x) + \sum_{j=1}^{m} \lambda(\omega_j) g(x, \omega_j) + \inf_{u \geq 0} \{\sum_{j=1}^{m} \lambda(\omega_j) u_j + r\sigma^{\Gamma}(\gamma(x, u))\} \\ &= L_0^{\Gamma}(x, \bar{\lambda}) + \inf_{u \geq 0} \{\sum_{j=1}^{m} \lambda(\omega_j) u_j + r\sigma^{\Gamma}(\gamma(x, u))\}, \end{aligned} \tag{18}$$

where $\gamma(x, u)$ is as in (16).

Proof: (i) Let $x \in \mathbb{R}^n$. For any $y \in \mathbb{R}^\Omega$ such that $g(x, \omega) + y(\omega) \leq 0$, $\omega \in \Gamma$, set $\bar{y}_j = y(\omega_j), j = 1, \cdots, m$, so $\bar{y} = (\bar{y}_1, \cdots, \bar{y}_m)^T$, and we have $g(x, \omega_j) + \bar{y}_j \leq 0$ for all $\omega_j \in \Gamma$. The definitions then yield

$$\begin{aligned} f(x) - \langle \lambda, y \rangle + r\sigma(y^{\Gamma}) &= f(x) - \langle \bar{\lambda}, \bar{y} \rangle + r\sigma^{\Gamma}(\bar{y}) \\ &\geq L^{\Gamma}(x, \bar{\lambda}, r), \end{aligned}$$

where we used (12) in the last inequality. Taking now infimum over $y \in D(x)$ and using (4) we obtain

$$L^{\Omega}(x, \lambda, r) \geq L^{\Gamma}(x, \bar{\lambda}, r).$$

Therefore, (17) holds. To prove the opposite inequality, define the set

$$D^{\Gamma}(x) := \{y \in D(x) : supp(y) \subset \Gamma\}.$$

The set $D^\Gamma(x)$ can be identified with the subset of $\mathbb{R}^m$ determined by all the images $y^\Gamma = (y(\omega_1), \ldots, y(\omega_m))^T \in \mathbb{R}^m$ for $y \in D^\Gamma(x)$. Namely, the set $D^\Gamma(x)$ can be identified with the set

$$D_m^\Gamma(x) := \{\bar{y} \in \mathbb{R}^m \ : \ g(x, \omega_j) + \bar{y}_j \leq 0, \ \forall \omega_j \in \Gamma\}.$$

Given any function $y \in D^\Gamma(x)$, call $\bar{y} \in \mathbb{R}^m$ the vector of images of y. Since $supp(y) \subset \Gamma$ and $y^\Gamma = \bar{y}$ we have that $\sigma(y^\Gamma) = \sigma^\Gamma(\bar{y})$ and $\langle \lambda, y \rangle = \langle \bar{\lambda}, y^\Gamma \rangle = \langle \bar{\lambda}, \bar{y} \rangle$, where we used the definition of $\bar{\lambda}$ in the latter equality. Altogether, we can write

$$
\begin{aligned}
L^\Omega(x, \lambda, r) &= f(x) + \inf_{y \in D(x)} -\langle \lambda, y \rangle + r\sigma(y^\Gamma) \\
&\leq f(x) + \inf_{y \in D^\Gamma(x)} -\langle \lambda, y \rangle + r\sigma(y^\Gamma) \\
&= f(x) + \inf_{\bar{y} \in D_m^\Gamma(x)} -\langle \bar{\lambda}, \bar{y} \rangle + r\sigma^\Gamma(\bar{y}) = L^\Gamma(x, \bar{\lambda}, r),
\end{aligned}
$$

which proves the opposite inequality.

(ii) Fix $x \in \mathbb{R}^n$ and take any $y \in \mathbb{R}^\Omega$ such that $g(x, \omega_j) + y(\omega_j) \leq 0$ for all $j = 1, \ldots, m$. Define $u_j := -(g(x, \omega_j) + y(\omega_j))$ for all $j = 1, \ldots, m$. Using (12) we can write

$$
\begin{aligned}
L^\Gamma(x, \bar{\lambda}, r) &= \inf_{\bar{y} \in \mathbb{R}^m} \{f(x) - \langle \bar{\lambda}, \bar{y} \rangle + r\sigma(\bar{y}) : g(x, \omega_j) + \bar{y}_j \leq 0, \ \omega_j \in \Gamma\} \\
&= \inf_{u \geq 0} \{f(x) + \sum_{j=1}^m \lambda(\omega_j)(g(x, \omega_j) + u_j) \\
&\quad + r\sigma(-g(x, \omega_1) - u_1, \cdots, -g(x, \omega_m) - u_m)\} \\
&= f(x) + \sum_{j=1}^m \lambda(\omega_j)g(x, \omega_j) + \inf_{u \geq 0} \{\sum_{j=1}^m \lambda(\omega_j)u_j + r\sigma(\gamma(x, u))\}.
\end{aligned}
$$

That is, the augmented Lagrangian $L^\Gamma(x, \bar{\lambda}, r)$ of (P^Γ) can be expressed as (18). $\quad\square$

The next lemma proves that (x^*, λ^*) is a saddle point of $L^\Omega(\cdot, \cdot, r^*)$ when (x^*, λ^*) is a Karush-Kuhn-Tucker (KKT in short) point such that x^* is optimal for (SIP) and λ^* is an augmented Lagrange multiplier with penalty parameter r^*.

Lemma 2.2 Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $g : \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}$ be real-valued functions, and let $\sigma : \mathbb{R}^\Omega \rightarrow \mathbb{R}_+$ be a real-valued function with $\sigma(0) = 0$, $x^* \in X_0$, $\lambda^* \in \mathbb{R}_+^{(\Omega)}$, $\Gamma = supp(\lambda^*) = \{\omega_1, \ldots, \omega_m\}$. Assume that the KKT conditions for (SIP) is satisfied at x^* and λ^*, i.e.,

$$\nabla_x L_0^\Gamma(x^*, \lambda^*) = 0, \tag{19}$$

and

$$\lambda^*(\omega) \geq 0, \lambda^*(\omega)g(x^*, \omega) = 0, \forall \omega \in \Omega. \tag{20}$$

Then

$$L^{\Omega}(x^*, \lambda^*, r) = L^{\Gamma}(x^*, \bar{\lambda}^*, r) = f(x^*), \quad \forall r > 0, \tag{21}$$

where $\bar{\lambda}^* = (\lambda^*(\omega_1), \cdots, \lambda^*(\omega_m))$.

Proof: Since $x^* \in X_0$, we have $g(x^*, \omega) \leq 0, \omega \in \Omega$. It follows from (4) that

$$L^{\Omega}(x^*, \lambda^*, r) \leq f(x^*). \tag{22}$$

Letting $\tilde{u}_j = -g(x^*, \omega_j) \geq 0$ for all $j = 1, \ldots, m$, we obtain

$$0 \leq \inf_{u \geq 0} \sigma^{\Gamma}(\gamma(x^*, u)) \leq \sigma^{\Gamma}(\gamma(x^*, \tilde{u})) = \sigma(0) = 0. \tag{23}$$

By using (20), we have

$$\inf_{u \geq 0} \sum_{j=1}^{m} \lambda^*(\omega_j) u_j = 0. \tag{24}$$

Noticing the KKT conditions and combining (23) and (24) with (18), we have

$$
\begin{aligned}
& L^{\Gamma}(x^*, \bar{\lambda}^*, r) \\
= \ & f(x^*) + \sum_{j=1}^{m} \lambda^*(\omega_j) g(x^*, \omega_j) + \inf_{u \geq 0}\{\sum_{j=1}^{m} \lambda^*(\omega_j) u_j + r\sigma^{\Gamma}(\gamma(x^*, u))\} \\
\geq \ & f(x^*) + \inf_{u \geq 0}\{\sum_{j=1}^{m} \lambda^*(\omega_j) u_j\} + r \inf_{u \geq 0}\{\sigma^{\Gamma}(\gamma(x^*, u))\} = f(x^*).
\end{aligned}
$$

This, together with (17) and (22), implies that (21) holds. $\qquad \square$

2.2 Duality Properties of (SIP)

In this section we establish some properties and strong duality results for (SIP), including first order optimality conditions and a global saddle property.

Lemma 2.3 *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *and* $g : \mathbb{R}^n \times \Omega \to \mathbb{R}$ *be real-valued functions, and let* $\sigma : \mathbb{R}^{\Omega} \to \mathbb{R}_+$ *be a real-valued function with* $\sigma(0) = 0$, $x^* \in X_0$, $\lambda^* \in \mathbb{R}_+^{(\Omega)}$, $\Gamma = supp(\lambda^*) = \{\omega_1, \ldots, \omega_m\}$. *Assume that the KKT conditions (19) and (20) for (SIP) are satisfied at* x^* *and* λ^*. *If there exists* $r^* > 0$, *such that* λ^* *is an augmented Lagrange multiplier of the problem (SIP) with the penalty parameter* r^* *and* $f(x^*) = val(SIP)$, *then*

$$L^{\Omega}(x, \lambda^*, r^*) \geq L^{\Omega}(x^*, \lambda^*, r^*) \geq L^{\Omega}(x^*, \lambda, r^*), \quad \forall x \in \mathbb{R}^n, \ \lambda \in \mathbb{R}^{(\Omega)}. \tag{25}$$

Conversely, if

$$L^{\Omega}(x, \lambda^*, r^*) \geq L^{\Omega}(x^*, \lambda^*, r^*), \quad \forall x \in \mathbb{R}^n, \tag{26}$$

then λ^* *is an augmented Lagrange multiplier of (SIP) with the penalty parameter* r^*.

Proof: As λ^* is an augmented Lagrange multiplier of (SIP) with the penalty parameter r^*, we have

$$v_{r^*}^\Omega(y) \geq v_{r^*}^\Omega(0) + \langle \lambda^*, y \rangle, \quad \forall y \in \mathbb{R}^\Omega.$$

Therefore

$$\inf_{y \in \mathbb{R}^\Omega} \{v_{r^*}^\Omega(y) - \langle \lambda^*, y \rangle\} \geq v_{r^*}^\Omega(0) = \text{val(SIP)}. \tag{27}$$

It follows from (21), (9) and (27) that

$$
\begin{aligned}
f(x^*) &= L^\Omega(x^*, \lambda^*, r^*) \geq \inf_{x \in \mathbb{R}^n} L^\Omega(x, \lambda^*, r^*) \\
&= \inf_{y \in \mathbb{R}^\Omega} \{v_{r^*}^\Omega(y) - \langle \lambda^*, y \rangle\} \\
&\geq \text{val(SIP)} = f(x^*).
\end{aligned} \tag{28}
$$

The above expression implies that

$$L^\Omega(x, \lambda^*, r^*) \geq L^\Omega(x^*, \lambda^*, r^*), \quad \forall x \in \mathbb{R}^n. \tag{29}$$

Let $\lambda \in \mathbb{R}^{(\Omega)}$. Since $x^* \in X_0$ and $0 \in \mathbb{R}^\Omega$, it follows from (4) that

$$L^\Omega(x^*, \lambda, r^*) \leq f(x^*) = L^\Omega(x^*, \lambda^*, r^*). \tag{30}$$

By (29) and (30), we obtain both inequalities in (25).

Assume that there exists $r^* > 0$, such that (26) holds. Then we obtain

$$\inf_{x \in \mathbb{R}^n} L^\Omega(x, \lambda^*, r^*) \geq L^\Omega(x^*, \lambda^*, r^*) = f(x^*) = \text{val(SIP)}. \tag{31}$$

It follows from (9) and (31) that

$$\inf_{y \in \mathbb{R}^\Omega} \{v_{r^*}^\Omega(y) - \langle \lambda^*, y \rangle\} = \inf_{x \in \mathbb{R}^n} L^\Omega(x, \lambda^*, r^*) \geq \text{val(SIP)}.$$

This implies that

$$v_{r^*}^\Omega(y) \geq v_{r^*}^\Omega(0) + \langle \lambda^*, y \rangle, \quad \forall y \in \mathbb{R}^\Omega.$$

That is, λ^* is an augmented Lagrange multiplier of (SIP) with penalty parameter r^*.

$\square$

Remark 2.3 By Lemma 2.3, if $\lambda^* \in \mathbb{R}^{(\Omega)}$ is an augmented Lagrange multiplier of (SIP) with the penalty parameter r^*, and x^* is an optimal solution of (SIP) and satisfies (20), then (x^*, λ^*) is a saddle point of $L^\Omega(\cdot, \cdot, r^*)$. Conversely, if for some $r^* > 0$, (x^*, λ^*) is a saddle point of $L(\cdot, \cdot, r^*)$, then λ^* is an augmented Lagrange multiplier of (SIP) with the penalty parameter r^*.

3   First-order Conditions for Augmented Lagrange Multipliers

In this section, we focus on the augmented Lagrangian in which the augmenting function is given by the norm of the function y^Γ, the restriction of $y \in \mathbb{R}^\Omega$ to a fixed finite subset $\Gamma \subset \Omega$. Namely, given $\Gamma := \{\omega_1, \ldots, \omega_m\}$ we consider $\sigma^\Gamma : \mathbb{R}^{(\Omega)} \to \mathbb{R}_+$, defined as

$$\sigma^\Gamma(y) := \sigma(y^\Gamma) = \|y^\Gamma\| = \sum_{\omega \in \Gamma} |y(\omega)| = \sum_{j=1}^m |y(\omega_j)|. \tag{32}$$

The resulting augmented Lagrangian, inspired in [15, Example 11.58], is called the *sharp Lagrangian*. It is a classical fact from convex programming that the sharp Lagrangian admits an exact penalty parameter (see, e.g., [27, Theorem 9.3.1]). We will show that this can also be achieved in our more general context. First, we will obtain necessary optimality conditions for the existence of an augmented Lagrange multiplier of the sharp Lagrangian problem. Second, we establish sufficient conditions for the existence of an augmented Lagrange multiplier, by assuming inf-compactness of the functions involved and a Mangasarian-Fromovitz constraint qualification. We start by establishing necessary first-order necessary conditions for the existence of an augmented Lagrange multiplier of the semi-infinite programming problem (SIP).

Given a fixed finite subset $\Gamma \subset \Omega$, we will use the following notation for the set of active constraints of (SIP) and the set of active constraints of (P^Γ) at a given $x \in X_0$, respectively.

$$I(x) := \{w \in \Omega \, : \, g(x, w) = 0\}, \quad \Gamma(x) := \{w \in \Gamma \, : \, g(x, w) = 0\} = I(x) \cap \Gamma. \tag{33}$$

Theorem 3.1 *Let $f(\cdot)$ and $g(\cdot, \omega), \omega \in \Omega$, be real-valued continuously differentiable functions. Fix $\lambda^* \in \mathbb{R}^{(\Omega)}$, and call $\Gamma = supp(\lambda^*) = \{\omega_1, \ldots, \omega_m\}$. Consider σ^Γ as in (32). Assume that*

(a) λ^* is an augmented Lagrange multiplier of (SIP) *with penalty parameter r^*,*

(b) x^* is a local minimum for (SIP), *and*

(c) val(SIP) $= f(x^*)$.

In this situation, we have that

$$\nabla f(x^*)^\top d \geq 0, \quad \forall d \in \Psi(x^*),$$

where

$$\Psi(x^*) := \{d \in \mathbb{R}^n : \nabla_x g(x^*, \omega)^\top d \leq 0, \omega \in I(x^*)\},$$

is the cone of feasible directions from x^*.

Proof: Assumption (a) means that

$$v_{r^*}^{\Omega}(y) - \langle \lambda^*, y \rangle \geq v_{r^*}^{\Omega}(0), \quad \forall y \in \mathbb{R}^{\Omega}.$$

Noting that $v_{r^*}^{\Omega}(y) = \mathrm{val}(P_y) + r^* \sigma(y^{\Gamma}) = \mathrm{val}(P_y) + r^* \sum_{\omega \in \Gamma} |y(\omega)|$ and $v_{r^*}^{\Omega}(0) = \mathrm{val}(\mathrm{SIP})$, the last inequality is rewritten as

$$\mathrm{val}(P_y) - \langle \lambda^*, y \rangle + r^* \sum_{\omega \in \Gamma} |y(\omega)| \geq \mathrm{val}(\mathrm{SIP}), \quad \forall y \in \mathbb{R}^{\Omega}. \tag{34}$$

Taking infimum in the above inequality and using (c), we obtain

$$\inf_{y \in \mathbb{R}^{\Omega}} \left\{ \mathrm{val}(P_y) - \langle \lambda^*, y \rangle + r^* \sum_{\omega \in \Gamma} |y(\omega)|\right\} \geq \mathrm{val}(\mathrm{SIP}) = f(x^*).$$

Using now the definition of $\mathrm{val}(P_y)$ and the fact (b) that x^* is a strict local minimum for (SIP) we obtain

$$
\begin{aligned}
f(x^*) &\leq \inf_{y \in \mathbb{R}^{\Omega}} \{\mathrm{val}(P_y) - \langle \lambda^*, y \rangle + r^* \sum_{\omega \in \Gamma} |y(\omega)|\} \\
&= \inf_{y \in \mathbb{R}^{\Omega}} \inf_{x \in \mathbb{R}^n} \{f(x) - \langle \lambda^*, y \rangle + r^* \sum_{\omega \in \Gamma} |y(\omega)| \; : \; g(x, \omega) + y(\omega) \leq 0, \omega \in \Omega\} \\
&= \inf_{x \in \mathbb{R}^n} \inf_{y \in \mathbb{R}^{\Omega}} \{f(x) + \sum_{j=1}^{m} \lambda^*(\omega_j)(-y(\omega_j)) \\
&\qquad + r^* \sum_{j=1}^{m} |y(\omega_j)| \; : \; g(x, \omega) + y(\omega) \leq 0, \omega \in \Omega\} \\
&= \inf_{x \in \mathbb{R}^n} L^{\Omega}(x, \lambda^*, r^*),
\end{aligned}
$$

where we used the definition of L^{Ω} (see (4)) in the last equality, and the fact that $\Gamma = \mathrm{supp}\,\lambda^*$ in the second-to-last equality. This expression readily gives

$$
\begin{aligned}
f(x^*) &\leq \inf_{y \in \mathbb{R}^{\Omega}} \{f(x) + \sum_{j=1}^{m} \lambda^*(\omega_j)(-y(\omega_j)) \\
&\qquad + r^* \sum_{j=1}^{m} |y(\omega_j)| \; : \; g(x, \omega) + y(\omega) \leq 0, \omega \in \Omega\} = L^{\Omega}(x, \lambda^*, r^*),
\end{aligned}
\tag{35}
$$

for all $x \in \mathbb{R}^n$. Fix $x \in \mathbb{R}^n$, for each $\omega_j \in \Gamma$, we have two possibilities. Either $g(x, \omega_j) \leq 0$, or $g(x, \omega_j) > 0$. If $g(x, \omega_j) \leq 0$, take $\tilde{y}(\omega_j) = 0$, otherwise, take $\tilde{y}(\omega_j) = -g(x, \omega_j)$. Namely, take $\tilde{y}(\omega) := -\max\{0, g(x, \omega)\}$ for $\omega \in \Gamma$ and $\tilde{y}(\omega) = 0$ for $\omega \notin \Gamma$. Using this choice of $\tilde{y} \in \mathbb{R}^{\Omega}$ in (35) we can write

$$
\begin{aligned}
f(x^*) \leq L^{\Omega}(x, \lambda^*, r^*) &= \inf_{y \in \mathbb{R}^{\Omega}} \{f(x) + \sum_{j=1}^{m} \lambda^*(\omega_j)(-y(\omega_j)) \\
&\qquad + r^* \sum_{j=1}^{m} |y(\omega_j)| \; : \; g(x, \omega) + y(\omega) \leq 0, \omega \in \Omega\} \\
&\leq f(x) + \sum_{j=1}^{m} \lambda^*(\omega_j) g^+(x, \omega_j) + r^* \sum_{j=1}^{m} g^+(x, \omega_j),
\end{aligned}
\tag{36}
$$

where $(a)^+ = \max\{a, 0\}$ for any $a \in \mathbb{R}$. Inequality (36) implies that

$$f(x^*) \leq f(x) + \sum_{j=1}^{m} (\lambda^*(\omega_j) + r^*) g^+(x, \omega_j), \forall x \in \mathbb{R}^n,$$

i.e., x^* is an unconstrained minimum of $H(x) := f(x) + \sum_{j=1}^{m}(\lambda^*(\omega_j) + r^*)g^+(x, \omega_j)$. Thus it must satisfy the necessary first order optimality conditions

$$H^\circ(x^*, d) = \nabla f(x^*)^T d + \sum_{j \in I(x^*)} (\lambda^*(\omega_j) + r^*) \max\{\nabla_x g(x^*, \omega_j)^T d, 0\} \geq 0, \ \forall d \in \mathbb{R}^n,$$

where $H^\circ(x^*, d)$ is the Clarke generalized directional derivative of H at x^* in the direction d, see [29], and we can restrict the sum to $j \in I(x^*)$ because x^* is feasible, and hence $\nabla_x g^+(x^*, \omega_j)^\top d = 0$ for all $j \notin I(x^*)$ and all $d \in \mathbb{R}^n$. Fix now $d \in \Psi(x^*)$. By definition of $\Psi(x^*)$ we must have $\max\{\nabla_x g(x^*, \omega_j)^T d, 0\} = 0$, and the above inequality yields

$$\nabla f(x^*)^\top d \geq 0, \quad \forall d \in \Psi(x^*).$$

This completes the proof. $\qquad\square$

The last result in this section establishes sufficient first order conditions for the existence of an augmented Lagrange multiplier. Given $\Gamma = \{\omega_1, \dots, \omega_m\} \subset \Omega$, define

$$X_0^\Gamma := \{x \in \mathbb{R}^n : g(x, \omega) \leq 0, \omega \in \Gamma\}, \tag{37}$$

the subset of points which are feasible for problem (P^Γ). We will make the following assumptions on the objective $f(\cdot)$ and the subset of constraint functions $g(\cdot, \omega)$ for $\omega \in \Gamma$.

Assumption 3.1 Let $f(x)$, and $g(x, \omega)$ ($\omega \in \Gamma$) be *inf-compact*, that is, there exist $\alpha \in \mathbb{R}$, $\omega \in \Gamma$, such that the sets $\{x \in \mathbb{R}^n : f(x) \leq \alpha\}$ and $\{x \in \mathbb{R}^n : g(x, \omega) \leq \alpha\}$ are nonempty and compact.

Remark 3.1 *If Assumption 3.1 holds, then the inf-compactness of f implies that the solution set S^Γ of the discretized problem (P^Γ) is nonempty and compact.*

The *extended Mangasarian-Fromovitz Constraint Qualification* (EMFCQ in short) is said to hold for (P^Γ) at a feasible solution x_0, see Bonnans and Shairo [28, p. 510], if there exists $z \in \mathbb{R}^n$ such that

$$\nabla_x g(x_0, \omega_j)^\top z < 0, \forall \omega_j \in \Gamma \text{ such that } g(x_0, w_j) = 0.$$

Assumption 3.2 The (EMFCQ) holds for the reduced problem (P^Γ) at every $x \in S^\Gamma$.

Let $x \in X_0$ and define the set of (classical) Lagrange multipliers associated with x:

$$\Lambda^\Gamma(x) := \left\{\lambda \in \mathbb{R}^m : \nabla_x L_0^\Gamma(x, \lambda) = 0, \lambda \geq 0, \lambda_j g(x, \omega_j) = 0, \ j = 1, \dots, m\right\}.$$

If Assumption 3.2 holds, then $\Lambda^\Gamma(x)$ is bounded (cf. Gauvin [30], see also Bonnans and Shairo [28, p. 510]). A general result on the boundedness of the set of Lagrange multipliers for an optimization problem with a closed set inclusion constraint is given in Burke [31].

Theorem 3.2 *We assume that (SIP) is convex. Let $\lambda^* \in \mathbb{R}^{(\Omega)}$, and $\Gamma = supp(\lambda^*) = \{\omega_1, \ldots, \omega_m\}$. Denote by $\lambda^*(\omega_j) = \bar{\lambda}_j$, $j = 1, \ldots, m$. Suppose that $val(SIP) = val(P^\Gamma)$. Let $f(\cdot)$ and $g(\cdot, \omega), \omega \in \Omega$, be real-valued continuously differentiable functions, and let σ^Γ be as in (32). Suppose that* Assumption 3.1 *and* Assumption 3.2 *are satisfied. Then, there exists $r^* > 0$, such that*

(i) *$\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m)^T$ is an augmented Lagrange multiplier of (P^Γ), with penalty parameter r^*, and*

(ii) *$\lambda^* \in \mathbb{R}^{(\Omega)}$ is an augmented Lagrange multiplier of (SIP), with the same penalty parameter r^*.*

Proof: (i) First, we claim that there exists $\alpha_0 \in \mathbb{R}$ such that $f(x) \geq \alpha_0, \forall x \in \mathbb{R}^n$. Indeed, if this is not the case, there exists a sequence $\{x_k\}$ such that $f(x_k) \to -\infty$ as $k \to \infty$. By Assumption 3.1, there exists $\alpha \in \mathbb{R}$, such that the set

$$\mathbb{E} = \{x \in \mathbb{R}^n : f(x) \leq \alpha\}$$

is nonempty and compact. So, $x_k \in \mathbb{E}$ for large enough k. Thus, there exists a subsequence $\{x_{k_j}\}$ of $\{x_k\}$, such that $x_{k_j} \to x_0$ as $j \to \infty$. Since f is continuous, $x_0 \in \mathbb{E}$ and $f(x_{k_j}) \to f(x_0)$ as $j \to \infty$, which is impossible.

Recall that $v^\Gamma(\bar{y})$ is the optimal value of $(P_{\bar{y}}^\Gamma)$. Thus

$$v^\Gamma(\bar{y}) \geq \alpha_0, \tag{38}$$

since $f(x) \geq \alpha_0, \forall x \in \mathbb{R}^n$.

To complete our proof, we will use similar arguments to the ones used in Rückmann and Shapiro [24, Theorem 3]. It follows from Assumption 3.1 that the set of optimal solutions of (P^Γ) is compact. By Assumption 3.2, the sets $\Lambda^\Gamma(x)$ are uniformly bounded for all x in a neighborhood of any point $\bar{x} \in S^\Gamma$ (cf. Proposition 4.43 of Bonnans and Shapiro [28]). Thus, we obtain that the set $\bigcup_{x \in S^\Gamma} \Lambda^\Gamma(x)$ is bounded, as S^Γ is compact, and then $\sup_{\lambda \in \bigcup_{x \in S^\Gamma} \Lambda^\Gamma(x)} \|\lambda - \bar{\lambda}\|$ is finite. As for (P^Γ), Robinson's constraint qualification is equivalent to the EMFCQ, by Theorem 4.27 of Bonnans and Shapiro [28], we have

$$(v^\Gamma)'_-(0, d) \geq \inf_{x \in S^\Gamma} \inf_{\lambda \in \Lambda^\Gamma(x)} \nabla_y \{f(x) + \langle \lambda, g(x) + y \rangle\}^\top d = \inf_{x \in S^\Gamma} \inf_{\lambda \in \Lambda^\Gamma(x)} \langle \lambda, d \rangle, \tag{39}$$

where $(v^\Gamma)'_-(0, d)$ is the lower Hadamard directional derivative of $v^\Gamma(\cdot)$ at 0 in a unit direction $d \in \mathbb{R}^m$, more precisely,

$$(v^\Gamma)'_-(0, d) := \sup_{r > 0} \left[\inf_{t \in (0, r), \|d - d'\| < r} \left[\frac{v^\Gamma(td') - v^\Gamma(0)}{t} \right] \right].$$

Using the above definition and (39), we can write, for a fixed $\varepsilon > 0$

$$\sup_{r>0} \left[\inf_{t \in [0,r], \, \|d-d'\|<r} \left[\frac{v^\Gamma(td') - v^\Gamma(0)}{t} \right] \right] > \inf_{x \in S^\Gamma} \inf_{\lambda \in \Lambda^\Gamma(x)} \langle \lambda, d \rangle - \varepsilon \, \|d\|.$$

Hence by definition of the supremum in the left-hand side there exists $\tau > 0$ such that for all $d' \in \tau B_{\mathbb{R}^m}$, which is the unit ball in $\mathbb{R}^m$, and every $t \in [0, \tau]$ we can write

$$\frac{v^\Gamma(td') - v^\Gamma(0)}{t} \;>\; \left[\inf_{x \in S^\Gamma} \inf_{\lambda \in \Lambda^\Gamma(x)} \langle \lambda, d \rangle \right] - \varepsilon \|d\|. \tag{40}$$

Multiplying by t the expression above and restricting the quotient in the left hand side to $d' = d$, we can write (40) in terms of $\bar{y} := td \in \tau B_{\mathbb{R}^m}$, to obtain

$$v^\Gamma(\bar{y}) - v^\Gamma(0) \;>\; \left[\inf_{x \in S^\Gamma} \inf_{\lambda \in \Lambda^\Gamma(x)} \langle \lambda, \bar{y} \rangle \right] - \varepsilon \|\bar{y}\|. \tag{41}$$

Let us find a suitable lower bound for the first term in the right-hand side of (41). We can write

$$\langle \lambda, \bar{y} \rangle \geq \langle \lambda - \bar{\lambda}, \bar{y} \rangle + \langle \bar{\lambda}, \bar{y} \rangle \geq \langle \bar{\lambda}, \bar{y} \rangle - \|\lambda - \bar{\lambda}\| \|\bar{y}\|.$$

Taking now the infimum in this expression we obtain

$$\begin{aligned} \inf_{x \in S^\Gamma} \inf_{\lambda \in \Lambda^\Gamma(x)} \langle \lambda, \bar{y} \rangle \;\geq\;& \langle \bar{\lambda}, \bar{y} \rangle + \|\bar{y}\| \inf_{x \in S^\Gamma} \inf_{\lambda \in \Lambda^\Gamma(x)} \left[-\|\lambda - \bar{\lambda}\| \right] \\ =\;& \langle \bar{\lambda}, \bar{y} \rangle - \|\bar{y}\| \sup_{x \in S^\Gamma} \sup_{\lambda \in \Lambda^\Gamma(x)} \left[\|\lambda - \bar{\lambda}\| \right]. \end{aligned} \tag{42}$$

The boundedness of the set $\bigcup_{x \in S^\Gamma} \Lambda^\Gamma(x)$ implies the existence of $r_1 > 0$ such that

$$r_1 > \sup_{x \in S^\Gamma} \sup_{\lambda \in \Lambda^\Gamma(x)} \|\lambda - \bar{\lambda}\|.$$

Using this fact, and (42)-(41) we obtain

$$\begin{aligned} v^\Gamma(\bar{y}) \;>\;& v^\Gamma(0) + \inf_{x \in S^\Gamma} \inf_{\lambda \in \Lambda^\Gamma(x)} \langle \lambda, \bar{y} \rangle - \varepsilon \|\bar{y}\| \\ =\;& v^\Gamma(0) + \langle \bar{\lambda}, \bar{y} \rangle - \|\bar{y}\| \sup_{x \in S^\Gamma} \sup_{\lambda \in \Lambda^\Gamma(x)} \left[\|\lambda - \bar{\lambda}\| \right] - \varepsilon \|\bar{y}\| \\ >\;& v^\Gamma(0) + \langle \bar{\lambda}, \bar{y} \rangle - \|\bar{y}\| \, r_1 - \varepsilon \|\bar{y}\| > v^\Gamma(0) + \langle \bar{\lambda}, \bar{y} \rangle - \|\bar{y}\| (r_1 + \varepsilon). \end{aligned} \tag{43}$$

Noting that every $\bar{y} \in \tau B_{\mathbb{R}^m}$ can be written as $\bar{y} := td$ for $t \in [0, \tau]$ and d a unit vector, we conclude that the inequality in (43) holds for every $\bar{y} \in \tau B_{\mathbb{R}^m}$. Altogether, we can write for all $\bar{y} \in \tau B_{\mathbb{R}^m}$

$$\begin{aligned} v_r^\Gamma(\bar{y}) \;=\;& v^\Gamma(\bar{y}) + r\sigma^\Gamma(\bar{y}) = v^\Gamma(\bar{y}) + r\|\bar{y}\| \\ \geq\;& v^\Gamma(0) + \|\bar{y}\|(r - \varepsilon - r_1) + \langle \bar{\lambda}, \bar{y} \rangle, \end{aligned} \tag{44}$$

where we used the definition of v_r^Γ in the first equality, the definition of σ^Γ in the second equality, and (43) in the inequality. This implies that for $r > (\varepsilon + r_1)$, we obtain the inequality in the definition of an augmented Lagrangian for (P^Γ), for every $\bar{y} \in \tau B_{\mathbb{R}^m}$.

Now we need to establish this inequality for $\bar{y} \notin \tau B_{\mathbb{R}^m}$. Recalling that $v^\Gamma(0) = \mathrm{val}(P^\Gamma)$, and using also (38) we can write

$$
\begin{aligned}
v_r^\Gamma(\bar{y}) - v^\Gamma(0) - \langle \bar{\lambda}, \bar{y} \rangle &= v^\Gamma(\bar{y}) + r\sigma^\Gamma(\bar{y}) - v^\Gamma(0) - \langle \bar{\lambda}, \bar{y} \rangle \\
&= v^\Gamma(\bar{y}) + r\|\bar{y}\| - v^\Gamma(0) - \langle \bar{\lambda}, \bar{y} \rangle \\
&\geq \alpha_0 - \mathrm{val}(P^\Gamma) + (r - \|\bar{\lambda}\|)\|\bar{y}\| \\
&\geq \alpha_0 - \mathrm{val}(P^\Gamma) + (r - \|\bar{\lambda}\|)\tau,
\end{aligned}
\tag{45}
$$

where the first inequality comes from (38) and the last inequality follows from choosing $r > \|\bar{\lambda}\|$ and the fact that $\bar{y} \notin \tau B_{\mathbb{R}^m}$. Given these two inequalities, we can choose

$$
r^* > \max\{r_1 + \varepsilon, \|\bar{\lambda}\| + \frac{\mathrm{val}(P^\Gamma) - \alpha_0}{\tau}\} > 0.
\tag{46}
$$

With this choice, (44), (45) and (46) yield

$$
v_{r^*}^\Gamma(y) \geq v^\Gamma(0) + \langle \bar{\lambda}, y \rangle, \quad \forall y \in \mathbb{R}^m.
\tag{47}
$$

From (15) $\bar{\lambda}$ is an augmented Lagrange multiplier of (P^Γ), with penalty parameter r^*.

(ii) Since $\mathrm{val}(SIP) = \mathrm{val}(P^\Gamma)$, we have that $v_{r^*}^\Omega(0) = v_{r^*}^\Gamma(0) = v^\Gamma(0) = \mathrm{val}(SIP)$. Fix an arbitrary $y \in \mathbb{R}^{(\Omega)}$, we can take its restriction y^Γ to Γ, namely set

$$
(y(\omega_1), \ldots, y(\omega_m))^T =: y^\Gamma.
$$

By Remark 2.2 and (14) we have $v_{r^*}^\Omega(y) \geq v_{r^*}^\Gamma(y^\Gamma)$. So, it follows from (47) applied to $y := y^\Gamma$ that

$$
v_{r^*}^\Omega(y) \geq v_{r^*}^\Gamma(y^\Gamma) \geq v_{r^*}^\Omega(0) + \langle \bar{\lambda}, y \rangle, \quad \forall y \in \mathbb{R}^{(\Omega)}.
$$

Therefore, $\bar{\lambda} \in \mathbb{R}^{(\Omega)}$ is an augmented Lagrange multiplier of (SIP), with penalty parameter r^*. $\qquad \square$

Remark 3.2 Note that the existence of an augmented Lagrange multiplier is a global property. The proof of Theorem 3.2 shows that Assumption 3.1 is essential to obtain the existence of an augmented Lagrange multiplier of (SIP). Indeed, by Assumption 3.1 and the continuity property of f, we obtain that (38) and (45) hold.

4 Second-order Conditions for Augmented Lagrange Multipliers

In this section, we investigate second-order conditions for the existence of an augmented Lagrange multiplier for the semi-infinite programming problem (SIP). Recall that for

$y \in \mathbb{R}^\Omega$, the norm of the restriction y^Γ to a finite set $\Gamma = \{\omega_1, \ldots, \omega_m\}$ is defined as $\|y^\Gamma\| = \sum_{\omega \in \Gamma} |y(\omega)|$, i.e., the ℓ_1-norm in $\mathbb{R}^m$ of the vector $y^\Gamma = (y(\omega_1), \ldots, y(\omega_m))^T$. We will also use $\|\cdot\|_\infty$ for denoting the *infinity norm* in $\mathbb{R}^m$. Namely, given $z \in \mathbb{R}^m$, we write $\|z\|_\infty := \max_{j=1,\ldots,m} |z_j|$. Next we state a crucial assumption we make on σ.

Definition 4.1 *Let $\sigma : \mathbb{R}^{(\Omega)} \to \mathbb{R}_+$ be a real-valued function and Γ be a finite subset of Ω.*

(a) *We say that σ has a valley at 0 with respect to Γ in $\mathbb{R}^\Omega$ if σ is continuous at 0 with $\sigma(0) = 0$, and for every $\delta > 0$ we have $c_\delta = \inf_{y \in \mathbb{R}^\Omega, \|y^\Gamma\| \geq \delta} \sigma(y^\Gamma) > 0$.*

(b) *Fix $p \in (0, 1]$. We say that σ satisfies a p-growth condition at 0 with respect to Γ if there exist $\rho_0 > 0$ such that*

$$\lim_{y \in \mathbb{R}^\Omega, y \to 0} \frac{\sigma(y^\Gamma)}{\|y^\Gamma\|^p} = \rho_0 > 0. \tag{48}$$

Remark 4.1 *Under assumption (a) in Definition 4.1, there exists $\alpha > 0$ such that the level set*

$$V_\alpha := \{z \in \mathbb{R}^m \ : \ \sigma^\Gamma(z) \leq \alpha\},$$

is bounded. Indeed, if this is not true then for all $\alpha > 0$ the set V_α is unbounded. Take $0 < \alpha < c_1$, where $c_1 > 0$ is as in Definition 4.1(a) for $\delta = 1$. Since V_α is unbounded there exists $z \in V_\alpha$ such that $\|z\| > 1$. Altogether, we have

$$\alpha \geq \sigma^\Gamma(z) \geq \inf_{y \in \mathbb{R}^\Omega, \|y^\Gamma\| \geq 1} \sigma(y^\Gamma) = c_1,$$

contradicting the choice of α. Hence, the claim is true and there exists $\alpha_0 > 0$ such that V_{α_0} is bounded. For future use, denote by M_{α_0} the positive constant such that

$$V_{\alpha_0} \subset B[0, M_{\alpha_0}] \subset \mathbb{R}^m,$$

where $B[0, M_{\alpha_0}]$ is the ball centered at 0 with radius M_{α_0}.

The properties described in Definition 4.1 are used for establishing a lower bound on the augmented Lagrangian. The next technical lemma shows why this assumption is instrumental in establishing the lower bound.

Lemma 4.1 *Fix a pair $(\hat{x}, \hat{\lambda}) \in \mathbb{R}^n \times \mathbb{R}_+^{(\Omega)}$ and denote by $\Gamma = supp(\hat{\lambda}) = \{\omega_1, \ldots, \omega_m\}$. Assume that*

(i) *For every $\omega \in \Gamma$, the function $g(\cdot, \omega) : \mathbb{R}^n \to \mathbb{R}$ is continuous at $\hat{x}$, and*

(ii) $\sigma : \mathbb{R}^{(\Omega)} \to \mathbb{R}^+$ *has a valley at 0 respect to* Γ, *and it satisfies p-growth condition at 0 with respect to* Γ.

In this situation, there exists an open neighborhood $N(\hat{x})$ *of* $\hat{x}$ *and* $r' > 0$ *such that for every* $x \in N(\hat{x})$ *and every* $r > r'$ *we have*

$$\inf_{u \geq 0} \left\{ \sum_{j=1}^m \hat{\lambda}(w_j) \left(g(x, \omega_j) + u_j \right) + r\sigma(-g(x, \omega_j) - u_j) \right\} \geq 0.$$

Proof: The statement of the lemma is clearly true when $\hat{\lambda} = 0$, so it is enough to prove the lemma for $\hat{\lambda} \neq 0$. As in (16) define $\gamma(x, u) \in \mathbb{R}^m$ as $[\gamma(x, u)]_j := -g(x, \omega_j) - u_j$ for all $j = 1, \ldots, m$. With this notation, we have that $[\gamma(x, u)]_j \leq -g(x, \omega_j)$ for all $u \geq 0$. The infimum in the statement of the lemma thus re-writes as

$$\inf_{u \geq 0} \left\{ \sum_{j=1}^m \hat{\lambda}(w_j)([-\gamma(x, u)]_j) + r\sigma(\gamma(x, u)) \right\}. \tag{49}$$

Fix $r_0 > 0$. The continuity of g implies that there exists a constant $C_1 > 0$ such that

$$\max_{x \in B[\hat{x}, r_0]} \left\{ \sum_{j=1}^m |g(x, \omega_j)| \right\} \leq C_1. \tag{50}$$

This allows us to find a lower bound for the first term in the argument of (49). Indeed,

$$\begin{aligned}
\sum_{j=1}^m \hat{\lambda}(w_j)([-\gamma(x, u)]_j) &\geq \sum_{j=1}^m \hat{\lambda}(w_j) g(x, \omega_j) \\
&\geq -\|\hat{\lambda}\|_\infty \sum_{j=1}^m g(x, \omega_j) \\
&\geq -C_1 \|\hat{\lambda}\|_\infty,
\end{aligned} \tag{51}$$

where we used the fact that $u_j \geq 0$ in the first inequality. The p-growth condition assumed in part (ii) implies that

$$\sup_{r > 0} \inf_{|z| < r} \frac{\sigma(z)}{\|z\|^p} = \rho_0 > \rho_0 / 2 > 0.$$

Therefore, there exists $r_1 > 0$ such that

$$\frac{\sigma(z)}{\|z\|^p} > \rho_0 / 2,$$

for every $z \in B(0, r_1) \subset \mathbb{R}^m$. Take $r_2 < \min\{1, r_1\}$ This yields

$$\sigma(z) \geq (\rho_0 / 2)\|z\|^p \geq (\rho_0 / 2)\|z\|, \tag{52}$$

where we used the fact that $p \in (0, 1]$ in the last inequality. Call $A(r, x, u) := \sum_{j=1}^m \hat{\lambda}(w_j)([-\gamma(x, u)]_j) + r\sigma(\gamma(x, u))$, the expression between the curly brackets in (49). We will show that

$$\inf_{u \geq 0, \, x \in B(\hat{x}, r_0)} A(r, x, u) \geq 0, \tag{53}$$

for large enough $r > 0$. We consider first the case in which $u \geq 0$ is such that $\|\gamma(x,u)\| < r_2$. In this case, using (52) for $z := \gamma(x,u)$ we obtain

$$
\begin{aligned}
A(r,x,u) &= \sum_{j=1}^{m} \hat{\lambda}(w_j)([-\gamma(x,u)]_j) + r\sigma(\gamma(x,u)) \\
&\geq -\|\hat{\lambda}\|_\infty \|\gamma(x,u)\| + r\sigma(\gamma(x,u)) \\
&\geq -\|\hat{\lambda}\|_\infty \|\gamma(x,u)\| + r\,(\rho_0/2)\|\gamma(x,u)\| \\
&= \|\gamma(x,u)\| \left(r\,(\rho_0/2) - \|\hat{\lambda}\|_\infty \right),
\end{aligned}
$$

which is positive for $r > \frac{2\|\hat{\lambda}\|_\infty}{\rho_0} =: r_3$. Hence, for all $x \in B(\hat{x}, r_0)$ we can write

$$
\inf_{r > r_3} \{ A(r,x,u) \ : \ u \geq 0 \text{ and } \|\gamma(x,u)\| < r_2 \} \geq 0. \tag{54}
$$

To complete the proof, we consider now those $u \geq 0$ such that $\|\gamma(x,u)\| \geq r_2$. Using the valley at 0 property and (51) have

$$
\begin{aligned}
A(r,x,u) &= \sum_{\omega_j \in \Gamma(\hat{x})} \hat{\lambda}(w_j)([-\gamma(x,u)]_j) + r\sigma(\gamma(x,u)) \\
&\geq -C_1\|\hat{\lambda}\|_\infty + r\sigma(\gamma(x,u)) \\
&\geq -C_1\|\hat{\lambda}\|_\infty + r\,c_{r_2},
\end{aligned}
$$

which is positive when $r > \frac{C_1\|\hat{\lambda}\|_\infty}{c_{r_2}} =: r_4$. Hence, for all $x \in B(\hat{x}, r_0)$ we have

$$
\inf_{r > r_4} \{ A(r,x,u) \ : \ u \geq 0 \text{ and } \|\gamma(x,u)\| \geq r_2 \} \geq 0. \tag{55}
$$

From (55) and (54) we deduce that, for $r > \max\{r_3, r_4\} =: r'$, the statement of the lemma is true for $N(\hat{x}) := B(\hat{x}, r_0)$. $\qquad\square$

The next result will be used to show that the Lagrangian L^Γ is lsc when σ is a valley at zero with respect to Γ.

Lemma 4.2 *Fix a pair* $(\hat{x}, \hat{\lambda}) \in \mathbb{R}^n \times \mathbb{R}_+^{(\Omega)}$, *and denote by* $\Gamma = supp(\hat{\lambda}) = \{\omega_1, \ldots, \omega_m\}$. *Assume that*

(i) $g(\cdot, \omega), \omega \in \Gamma$, *is real-valued and continuous,*

(ii) $\sigma : \mathbb{R}^{(\Omega)} \to \mathbb{R}^+$ *is lsc and it has a valley at 0 with respect to* Γ.

For a fixed $r > 0$, *define* $q(\cdot, r) : \mathbb{R}^n \to \mathbb{R}_{\pm\infty}$ *as*

$$
q(x,r) := \inf_{u \geq 0} \left\{ \sum_{j=1}^{m} \hat{\lambda}(w_j)([-\gamma(x,u)]_j) + r\sigma(\gamma(x,u)) \right\}.
$$

Then, for $r > 0$ *large enough, the function* $q(\cdot, r)$ *is lsc in* $\mathbb{R}^n$.

Proof: Fix $r > 0$. Set $Q(r, x, u) := \sum_{j=1}^m \hat{\lambda}(w_j)([-\gamma(x, u)]_j) + r\sigma(\gamma(x, u)) + \delta_{\mathbb{R}_+^m}(u)$.

Then, we can write

$$q(x, r) = \inf_{u \in \mathbb{R}^m} Q(r, x, u).$$

Theorem 1.17 in [15] states that $q(\cdot, r)$ is lsc if (a) $Q(r, \cdot, \cdot)$ is lsc, and (b) $Q(r, \cdot, \cdot)$ is level-bounded in u locally uniformly in x. Condition (b) can be re-stated as follows. For every $\alpha \in \mathbb{R}$ there exists a ball $B(\hat{x}, r_0)$ such that the set

$$W_\alpha(r) := \{(x, u) \ : \ x \in B(\hat{x}, r_0), \ Q(r, x, u) \leq \alpha\} \quad \text{is bounded in } \mathbb{R}^n \times \mathbb{R}^m. \tag{56}$$

Since the $\mathbb{R}^n$-coordinate of $W_\alpha(r)$ is bounded by definition, it is enough to show that the $\mathbb{R}^m$-coordinate of $W_\alpha(r)$ is bounded. We start by noting that $Q(r, \cdot, \cdot)$ is lsc. This is a consequence of the fact that $Q(r, \cdot, \cdot)$ is the sum of lower semicontinuous functions. Indeed, the continuity assumption (i) on g, together with the lower semicontinuity of σ assumed in (ii) implies that $\sigma \circ \gamma(\cdot, \cdot)$ is lsc. The first term in the expression of $Q(r, \cdot, \cdot)$ is continuous on u and hence lsc. The third term in the expression of $Q(r, \cdot, \cdot)$ is $\delta_{\mathbb{R}_+^m}$, which is lsc because the set $\mathbb{R}_+^m$ is closed. This proves that $Q(r, \cdot, \cdot)$ is lsc for every $r > 0$.

Let us show now that, for $r > 0$ and large enough, $Q(r, \cdot, \cdot)$ is level-bounded in u locally uniformly in x. Namely, we will prove that (56) holds for $r_0 = 1$ and for $r > 0$ large enough. The continuity of g implies that there exists $C_1 > 0$ such that

$$\max_{x \in B[\hat{x}, 1]} \{\sum_{j=1}^m |g(x, \omega_j)|\} \leq C_1. \tag{57}$$

Note that, for every fixed ω_{j_0} we have

$$C_1 \geq \sum_{j=1}^m |g(x, \omega_j)| \geq |g(x, \omega_{j_0})| \geq -g(x, \omega_{j_0}). \tag{58}$$

Assume that $(x, u) \in W_\alpha(r)$ (and hence $x \in B[\hat{x}, 1]$), then we must have $u \geq 0$ and we can write a lower bound for the first term in the expression of Q as follows:

$$\begin{aligned}
\sum_{j=1}^m \hat{\lambda}(w_j)([-\gamma(x, u)]_j) &= \sum_{j=1}^m \hat{\lambda}(w_j)(g(x, \omega_j) + u_j) \\
&\geq \sum_{j=1}^m \hat{\lambda}(w_j)g(x, \omega_j) \\
&\geq -C_1\|\hat{\lambda}\|,
\end{aligned}$$

where we used the fact that $u_j \geq 0$ in the fist inequality and (58) in the last one. Altogether, we can write, for every $(x, u) \in W_\alpha(r) := \{(x, u) : x \in B(\hat{x}, 1), Q(r, x, u) \leq \alpha\}$,

$$\begin{aligned}
\alpha \geq Q(r, x, u) &= \sum_{j=1}^m \hat{\lambda}(w_j)([-\gamma(x, u)]_j) + r\sigma^\Gamma(\gamma(x, u)) \\
&\geq -C_1\|\hat{\lambda}\| + r\sigma^\Gamma(\gamma(x, u)).
\end{aligned}$$

Therefore, we deduce that

$$W_\alpha(r) \subset \{(x, u) \ : \ \sigma^\Gamma(\gamma(x, u)) \leq \frac{\alpha + C_1 \|\hat{\lambda}\|}{r}\}.$$

By Remark 4.1 there exists α_0 such that the set

$$V_{\alpha_0} := \{z \in \mathbb{R}^m \ : \ \sigma^\Gamma(z) \leq \alpha_0\},$$

is bounded. Take $r > 0$ such that $\frac{\alpha + C_1 \|\hat{\lambda}\|}{r} \leq \alpha_0$. For this choice of r, we have that, if $(x, u) \in W_\alpha(r)$, then $\gamma(x, u) \in V_{\alpha_0}$. Let M_{α_0} be as in Remark 4.1. We can thus write, for every u such that $(x, u) \in W_\alpha(r)$,

$$
\begin{aligned}
\|u\| &= \textstyle\sum_{j=1}^m |u_j| = \sum_{j=1}^m |-g(x, \omega_j) + g(x, \omega_j) + u_j| \\
&\leq \textstyle\sum_{j=1}^m |-g(x, \omega_j)| + |g(x, \omega_j) + u_j| \leq C_1 + \|\gamma(x, u)\| \leq C_1 + M_{\alpha_0},
\end{aligned}
$$

showing that, for this choice of r, we have

$$W_\alpha(r) \subset B[\hat{x}, 1] \times B[0, C_1 + M_{\alpha_0}].$$

Hence (56) holds for $r > 0$ large enough. By Theorem 1.17 in [15], the function $q(\cdot, r)$ is lsc for $r > 0$ large enough. $\qquad\square$

This lemma readily gives the lower semicontinuity of $L^\Gamma(\cdot, \lambda, r)$ for large enough r.

Proposition 4.1 *Fix a pair $(\hat{x}, \hat{\lambda}) \in \mathbb{R}^n \times \mathbb{R}_+^{(\Omega)}$, and denote by $\Gamma = supp(\hat{\lambda}) = \{\omega_1, \ldots, \omega_m\}$. Assume that*

(i) $f(\cdot)$ is real-valued and lsc and $g(\cdot, \omega), \omega \in \Gamma$, is real-valued and continuous,

(ii) $\sigma : \mathbb{R}^{(\Omega)} \to \mathbb{R}^+$ is lsc and it has a valley at 0 with respect to Γ.

In this situation, $L^\Gamma(\cdot, \lambda, r)$ is lsc for $r > 0$ large enough.

Proof: By definition we know that

$$
\begin{aligned}
L^\Gamma(x, \hat{\lambda}, r) &= f(x) + \inf_{u \geq 0}\{\textstyle\sum_{j=1}^m \hat{\lambda}(\omega_j)[-\gamma(x, u)]_j + r\sigma(\gamma(x, u))\} \\
&= f(x) + q(x, r),
\end{aligned}
$$

where we are using the notation of Lemma 4.2. By assumption (i) the first term is lsc. Since we are in conditions of Lemma 4.2, the second term is lsc for r large enough. Hence the whole expression is lsc for r large enough. $\qquad\square$

Remark 4.2 *With the notation of Definition 4.1, we have that the sharp augmenting function $\sigma^\Gamma(y) := \|y^\Gamma\|$ used in Section 3 has a valley at 0 with $c_\delta = \delta$, and it satisfies a p-growth condition at 0 with respect to Γ for $p = 1$. More generally, the augmenting function $\sigma_p^\Gamma(y) := \|y^\Gamma\|^p$ for $p \in (0, 1]$ has a valley at 0 with $c_\delta = \delta^p$, and it satisfies a p-growth condition at 0.*

The following lemma uses an augmenting function as in Definition 4.1(a)(b).

Lemma 4.3 *Let $f(\cdot)$ and $g(\cdot, \omega), \omega \in \Omega$, be real-valued twice continuously differentiable functions, and let $x^* \in X_0$, $\lambda^* \in \mathbb{R}^{(\Omega)}$, $\Gamma = supp(\lambda^*) = \{\omega_1, \ldots, \omega_m\}$, $\sigma : \mathbb{R}^{(\Omega)} \to \mathbb{R}^+$ is lsc, it has a valley at 0 respect to Γ, and it satisfies p-growth condition at 0 with respect to Γ.*

Assume that the KKT conditions (19) *and* (20) *for (SIP) are satisfied, and that the second-order sufficient condition given by*

$$z^T \nabla_{xx}^2 L_0^\Gamma(x^*, \lambda^*) z > 0, \ \forall z \in V(x^*), \tag{59}$$

holds, where $V(x^*)$ is defined by

$$V(x^*) := \left\{ 0 \neq d \in \mathbb{R}^n \ \middle| \ \begin{array}{l} \nabla f(x^*)^T d \leq 0, \\ \nabla_x g(x^*, \omega)^T d \leq 0, \ \ \forall \ \omega \in I(x^*) \bigcap \Gamma \end{array} \right\}. \tag{60}$$

Then there exist $r^* > 0$ and a neighborhood $N(x^*)$ of x^* such that

$$L^\Omega(x, \lambda^*, r^*) \geq L^\Omega(x^*, \lambda^*, r^*), \quad \forall x \in N(x^*). \tag{61}$$

Proof: By contradiction, suppose that there exist two sequences $\{r_k\}$ and $\{x_k\}$ with $r_k > 0$, $x_k \neq x^*$, such that $r_k \to \infty$, $x_k \to x^*$, as $k \to \infty$ and

$$L^\Omega(x_k, \lambda^*, r_k) < L^\Omega(x^*, \lambda^*, r_k). \tag{62}$$

We have

$$L_0^\Gamma(\lambda^*, r^*) = f(x_k) + \sum_{j=1}^m \lambda^*(\omega_j) g(x_k, \omega_j) \leq L^\Gamma(x_k, \lambda^*, r_k) = L^\Omega(x_k, \lambda^*, r_k) < f(x^*), \tag{63}$$

where the first inequality follows from (18) and the nonnegativity of $\lambda^*(\omega)$ ($\forall \omega \in \Omega$) (since (20) holds), the second equality is due to (17) and the last inequality comes from (62) and (21). Define $v_k := x_k - x^*$ and $s_k := \frac{v_k}{\|v_k\|}$. By compactness of the ball, there exists a subsequence of $\{s_k\}$ converging to a vector s with $\|s\| = 1$. Without loss of generality, suppose that $s = \lim_{k \to \infty} s_k$. We will show first that $s \in V(x^*)$. We can write

$$0 > \frac{L^\Omega(x_k, \lambda^*, r_k) - L^\Omega(x^*, \lambda^*, r_k)}{\|v_k\|} \geq \frac{f(x_k) - f(x^*)}{\|v_k\|} + \frac{\inf_{u \geq 0} A(r_k, x_k, u)}{\|v_k\|}, \tag{64}$$

where $A(r, x, u)$ is as in Lemma 4.1. Since we are under conditions of Lemma 4.1, there exist $r', r_0 > 0$ such that the last term is nonnegative for $r_k > r'$ and $x_k \in B(x^*, r_0)$. Since r_k tends to infinity and x_k tends to x^* we can find k_0 such that $r_k > r'$ and

$x_k \in B(x^*, r_0)$ for $k \geq k_0$. We assume from now on that $k \geq k_0$. Using this fact and the Taylor development of f we can write

$$0 \; > \; \frac{f(x_k) - f(x^*)}{\|v_k\|} + \frac{\inf_{u \geq 0} A(r_k, x_k, u)}{\|v_k\|} \geq \frac{f(x_k) - f(x^*)}{\|v_k\|} \tag{65}$$

$$= \nabla f(x^*)^T s_k + \frac{o(\|v_k\|)}{\|v_k\|}, \tag{66}$$

where $\lim_k \frac{o(\|v_k\|)}{\|v_k\|} = 0$. Taking limit yields

$$\nabla f(x^*)^T s \leq 0, \tag{67}$$

which is one of the inequalities defining $V(x^*)$. To show that the remaining inequalities in $V(x^*)$ hold for s, write again

$$0 \; > \; \frac{f(x_k) - f(x^*)}{\|v_k\|} + \frac{\inf_{u \geq 0} A(r_k, x_k, u)}{\|v_k\|} \tag{68}$$

$$= \; \nabla_x L_0^\Gamma(x^*, \lambda^*)^T s_k + \frac{o(\|v_k\|)}{\|v_k\|} + r_k \frac{\inf_{u \geq 0} \sigma(\gamma(x_k, u))}{\|v_k\|}$$

$$= \; \frac{o(\|v_k\|)}{\|v_k\|} + r_k \frac{\inf_{u \geq 0} \sigma(\gamma(x_k, u))}{\|v_k\|}, \tag{69}$$

where the first inequality follows from (62), the first equality follows from the definitions of $A(x_k, r_k, u)$ and L_0^Γ, and the second equality comes from (19) in Lemma 2.2. Thus, we have

$$\lim_{k \to \infty} \frac{\inf_{u \geq 0} \sigma(\gamma(x_k, u))}{\|v_k\|} = 0,$$

since $r_k \to \infty$, as $k \to \infty$. Therefore, there exists $u^k \in \mathbb{R}_+^m$ such that

$$\lim_{k \to \infty} \frac{\sigma(\gamma(x_k, u^k))}{\|v_k\|} = 0, \tag{70}$$

and then

$$\lim_{k \to \infty} \sigma(\gamma(x_k, u^k)) = 0.$$

Since σ has a valley at 0, this can only happen if

$$\lim_{k \to \infty} \|\gamma(x_k, u^k)\| = 0. \tag{71}$$

The p-growth condition now implies that

$$\lim_{k \to \infty} \frac{\sigma(\gamma(x_k, u^k))}{\|\gamma(x_k, u^k)\|^p} = \rho_0. \tag{72}$$

Note that $0 < p \leq 1$ and $|g(x_k, \omega_j) + u_j^k| \geq g^+(x_k, \omega_j)$ for all $\omega_j \in \Gamma(x^*)$. Proceeding as in the proof of Lemma 4.1, and taking k large enough so that $\|\gamma(x_k, u^k)\| < 1$ we obtain

$$\sum_{\omega_j \in \Gamma(x^*)} g^+(x_k, \omega_j) \leq \|\gamma(x_k, u^k)\| \leq \|\gamma(x_k, u^k)\|^p. \tag{73}$$

From (72) we can take also k large enough so that $\frac{\sigma(\gamma(x_k, u^k))}{\|\gamma(x_k, u^k)\|^p} > \rho_0/2$. Combine this with (73) to write

$$0 \le \frac{\sum_{\omega_j \in \Gamma(x^*)} g^+(x_k, \omega_j)}{\|v_k\|} \le \frac{\|\gamma(x_k, u^k)\|^p}{\|v_k\|} \le \frac{2}{\rho_0} \frac{\sigma(\gamma(x_k, u^k))}{\|v_k\|}.$$

Using the above expression and (70) we deduce that

$$\lim_{k \to \infty} \frac{\sum_{\omega_j \in \Gamma(x^*)} g^+(x_k, \omega_j)}{\|v_k\|} = 0. \tag{74}$$

Since the functions $g(\cdot, \omega_j)$, $\omega_j \in \Gamma(x^*)$ are continuously differentiable at x^*, and $g(x^*, \omega_j) = 0$ for all $\omega_j \in \Gamma(x^*)$, we use the Taylor development of $g(\cdot, \omega_j)$ to write

$$\frac{g^+(x_k, \omega_j)}{\|v_k\|} = \max\{0, \frac{g(x_k, \omega_j)}{\|v_k\|}\} \tag{75}$$

$$= \max\left\{0, \nabla_x g(x^*, \omega_j)^T s_k + \frac{o(\|v_k\|)}{\|v_k\|}\right\} \tag{76}$$

for all $\omega_j \in \Gamma(x^*)$. Since each term in the sum of (74) is nonnegative, we deduce that

$$\lim_{k \to \infty} \frac{g^+(x_k, \omega_j)}{\|v_k\|} = 0.$$

for all $\omega_j \in \Gamma(x^*)$. Use this fact in (75) to obtain

$$0 = \lim_{k \to +\infty} \max\left\{0, \nabla_x g(x^*, \omega_j)^T s_k + \frac{o(\|v_k\|)}{\|v_k\|}\right\} = \max\left\{0, \nabla_x g(x^*, \omega_j)^T s\right\}, \tag{77}$$

for all $\omega_j \in \Gamma(x^*)$. This readily gives

$$\nabla_x g(x^*, \omega_j)^T s \le 0, \tag{78}$$

for all $\omega_j \in \Gamma(x^*)$. Combining this fact and (67) we conclude that $s \in V(x^*)$. The Taylor development of L_0^Γ and the first order conditions yield

$$f(x_k) - f(x^*) + \sum_{j=1}^m \lambda^*(\omega_j) g(x_k, \omega_j)$$

$$= f(x_k) - f(x^*) + \sum_{j=1}^m \left[\lambda^*(\omega_j) g(x_k, \omega_j) - \lambda^*(\omega_j) g(x^*, \omega_j)\right]$$

$$= \frac{1}{2} s_k^T \left[\nabla^2 f(x^*) + \sum_{j=1}^m \lambda^*(\omega_j) \nabla_{xx}^2 g(x^*, \omega_j)\right] s_k \|v_k\|^2 + o(\|v_k\|^2)$$

$$= \|v_k\|^2 \left(s_k^T \left[\nabla^2 f(x^*) + \sum_{j=1}^m \lambda^*(\omega_j) \nabla_{xx}^2 g(x^*, \omega_j)\right] s_k + \frac{o(\|v_k\|^2)}{\|v_k\|^2}\right)$$

$$= \|v_k\|^2 \left(s_k^T \nabla_{xx}^2 L_0^\Gamma(x^*, \lambda^*) s_k + \frac{o(\|v_k\|^2)}{\|v_k\|^2}\right)$$

Assumption (59) and the fact that $s \in V(x^*)$, imply that, for k large enough, the expression between parentheses is positive. Hence, we have that

$$f(x_k) - f(x^*) + \sum_{j=1}^{m} \lambda^*(\omega_j) g(x_k, \omega_j) > 0,$$

contradicting (63) for large enough k. This implies that (62) cannot hold, and therefore (61) must hold. $\qquad \square$

The following result establishes existence of an augmented Lagrange multiplier for (SIP), under the conditions of Lemma 4.3 and some boundedness assumptions on f and $g(\cdot, \omega)$.

Theorem 4.1 *Let x^* be a unique global optimal solution of (P^Γ). Suppose that all conditions in Lemma 4.3 are satisfied. Moreover, if*

    (i)   *f and $g(x, \omega)$ $(\omega \in \Gamma)$ are bounded below,*

    (ii)   *there exists $\mu_0 > 0$ such that the set below is bounded*

$$\Theta := \{x \in \mathbb{R}^n : g(x, \omega) \leq \mu_0, \omega \in \Gamma\}, \tag{79}$$

then λ^* is an augmented Lagrange multiplier of (SIP).

Proof: By condition (i), there exists $\xi_0 > -\infty$ such that

$$\min_{x \in \mathbb{R}^n, \omega \in \Omega} \{f(x), g(x, \omega)\} \geq \xi_0. \tag{80}$$

By Lemma 2.2 (ii), it is enough to prove that there exists $r_0^* > 0$ such that (26) holds for every $x \in \mathbb{R}^n$. We establish this fact in two steps, according to whether x belongs or not to the set Θ defined in (79). We use in this proof the notation of Lemma 4.1.

Step I. In this step we prove that there exists $r_1^* > 0$ such that (26) holds on $\mathbb{R}^n \backslash \Theta$. If $x \in \mathbb{R}^n \backslash \Theta$, then $g(x, \omega_{j_0}) > \mu_0$ for some $\omega_{j_0} \in \Gamma$. Thus, for every $u \geq$ we have

$$\|\gamma(x, u)\| \geq g(x, \omega_{j_0}) + u_{j_0} > \mu_0.$$

Since σ has a valley at 0, there exists $c_{\mu_0} > 0$ such that

$$\sigma(\gamma(x, u)) \geq c_{\mu_0}. \tag{81}$$

By (18), (80) and (81), we have

$$L^\Gamma(x, \lambda^*, r) = f(x) + \sum_{j=1}^{m} \lambda^*(\omega_j) g(x, \omega_j) + \inf_{u \geq 0} \{\sum_{j=1}^{m} \lambda^*(\omega_j) u_j + r\sigma(\gamma(x, u))\}$$

$$\geq \xi_0 (1 + \sum_{j=1}^{m} \lambda^*(\omega_j)) + r c_{\mu_0}, \quad \forall x \in \mathbb{R}^n \backslash \Theta.$$

Then there exists $r_1^* > 0$ such that

$$L^\Gamma(x, \lambda^*, r_1^*) \geq f(x^*), \quad \forall x \in \mathbb{R}^n \backslash \Theta.$$

This inequality, combined with (17) in Lemma 2.1 and (21) in Lemma 2.2, yields that

$$L^\Omega(x, \lambda^*, r_1^*) \geq L^\Omega(x^*, \lambda^*, r_1^*), \quad \forall x \in \mathbb{R}^n \backslash \Theta. \tag{82}$$

That is, (26) holds on $\mathbb{R}^n \backslash \Theta$.

Step II. We prove in this step that there exists $r_2^* > 0$ such that (26) holds for $x \in \Theta$. By contradiction, suppose that there exist sequences $\{r_k\}$ and $\{x_k\}$ with $r_k > 0$, $x_k \in \Theta$, such that $r_k \to +\infty$ and

$$L^\Gamma(x_k, \lambda^*, r_k) = L^\Omega(x_k, \lambda^*, r_k) < L^\Omega(x^*, \lambda^*, r_k) = f(x^*), \tag{83}$$

where we used (21) in the equalities. Set $\Gamma = \{\omega_1, \ldots, \omega_m\}$. It follows from (18) and (83) that

$$f(x^*) > f(x_k) + \sum_{j=1}^m \lambda^*(\omega_j) g(\omega_j, x_k) + r_k \inf_{u \geq 0} \sigma(\gamma(x_k, u)). \tag{84}$$

Since $r_k \to +\infty$ the above inequality implies that

$$\lim_{k \to \infty} \inf_{u \geq 0} \sigma(\gamma(x_k, u)) = 0.$$

Therefore, we can take a sequence $\{u^k\} \subset \mathbb{R}_+^m$ such that

$$\lim_{k \to \infty} \sigma(\gamma(x_k, u^k)) = 0.$$

By condition (ii), Θ is bounded, and the sequence $\{x_k\}$ has at least a cluster point $\bar{x}$. Without loss of generality, we assume that $x_k \to \bar{x}$. We claim that $\bar{x} \in X_0^\Gamma$, where X_0^Γ is defined by (37). Indeed, the fact that σ has a valley at zero and the last expression imply that, for every $\omega_j \in \Gamma$ we have

$$0 = \lim_{k \to \infty} \|\gamma(x_k, u^k)\| \geq \lim_{k \to \infty} |g(x_k, \omega_j) + u_j^k| = 0.$$

Since g is continuous, the sequence $\{g(x_k, \omega_j)\}$ is bounded for every j, and hence we conclude that the sequence $\{u^k\} \subset \mathbb{R}_+^m$ is bounded too. Therefore, they both have convergent subsequences. We may and do denote these still by $\{g(x_k, \omega_j)\}$ and $\{u^k\}$. Hence, we can write, for every $\omega_j \in \Gamma$,

$$g(\bar{x}, \omega_j) = \lim_{k \to \infty} g(x_k, \omega_j) = \lim_{k \to \infty} -u_j^k \leq 0,$$

because $\{u^k\} \subset \mathbb{R}_+^m$. Therefore $\bar{x} \in X_0^\Gamma$. Since we are under the conditions of Lemma 4.3, there exist $r^* > 0$ and a neighborhood $N(x^*)$ of x^* such that (61) holds, i.e.,

$$L^\Omega(x, \lambda^*, r^*) > L^\Omega(x^*, \lambda^*, r^*) = f(x^*), \quad \forall x \in N(x^*). \tag{85}$$

It follows from (83) and (85) that $x_k \notin N(x^*)$. Indeed, for large enough k we must have $r_k > r^*$ and hence (83) gives

$$L^\Gamma(x_k, \lambda^*, r^*) \le L^\Gamma(x_k, \lambda^*, r_k) < L^\Omega(x^*, \lambda^*, r_k) = f(x^*),$$

where we used the fact that $r_k > r^*$ in the first inequality. This expression and (85) yield $x_k \notin N(x^*)$. Since $x_k \to \bar{x}$, we deduce that $\bar{x} \ne x^*$. By assumption, x^* is a unique global optimal solution of (P^Γ) and $\bar{x}$ is feasible for (P^Γ), hence we must have $f(\bar{x}) > f(x^*)$. Let $\beta = f(\bar{x}) - f(x^*)$. By using (17) and (18), we have

$$
\begin{aligned}
L^\Omega(\bar{x}, \lambda^*, r) =& f(\bar{x}) + \inf_{u \ge 0}\{\sum_{j=1}^m \lambda^*(\omega_j)[-\gamma(\bar{x}, u)]_j + r\sigma^\Gamma(\gamma(\bar{x}, u))\} \\
\ge& f(x^*) + \frac{\beta}{2} + \inf_{u \ge 0}\{\sum_{j=1}^m \lambda^*(\omega_j)[-\gamma(\bar{x}, u)]_j + r\sigma^\Gamma(\gamma(\bar{x}, u))\}.
\end{aligned}
\tag{86}
$$

Since we are under the conditions of Lemma 4.1, the last term in (86) is nonnegative for r large enough. Using this in (86) gives

$$L^\Omega(\bar{x}, \lambda^*, r) \ge f(x^*) + \frac{\beta}{2}, \quad \forall r \ge r'. \tag{87}$$

Fix k_0 such that $r_k \ge r'$ for all $k \ge k_0$, where $r' > 0$ is as in (87). Using Proposition 4.1, we may and do assume that r' is large enough such that $L^\Gamma(\cdot, \lambda^*, r')$ is lsc. Then by (83) we obtain

$$
\begin{aligned}
f(x^*) \ge& \liminf_k L^\Gamma(x_k, \lambda^*, r_k) \ge \liminf_k L^\Gamma(x_k, \lambda^*, r') \\
\ge& L^\Gamma(\bar{x}, \lambda^*, r') = L^\Omega(\bar{x}, \lambda^*, r) \ge f(x^*) + \frac{\beta}{2},
\end{aligned}
$$

where we used (83) in the first inequality, the fact that $L^\Gamma(x_k, \lambda^*, r_k) \ge L^\Gamma(\bar{x}, \lambda^*, r')$ if $r_k \ge r'$ in the second inequality, lower semicontinuity of $L^\Gamma(\cdot, \lambda^*, r')$ in the third inequality, (17) in the equality, and (87) in the last inequality. This expression entails a contradiction and hence there exists $r_2^* > 0$ such that

$$L^\Omega(x, \lambda^*, r_2^*) \ge L^\Omega(x^*, \lambda^*, r_2^*), \quad \forall x \in \Theta. \tag{88}$$

Equivalently, (26) holds on Θ. Let $r_0^* = \max\{r_1^*, r_2^*\}$. It follows from (82) and (88) that (26) holds on $\mathbb{R}^n$. The proof is complete. $\qquad\square$

Remark 4.3 Shapiro and Sun [22] established second-order conditions ensuring existence of an augmented Lagrange multiplier when the augmenting function is assumed to be a quadratic function and condition (R) (called *quadratic growth condition* in Rockafellar[14]), is satisfied. Under some second-order sufficient conditions and by assuming that the global optimal solution of the primal problem is unique, Sun et al

[23] obtained existence results of global saddle points for four classes of augmented Lagrangian functions (by Remark 2.3, this implies existence of augmented Lagrange multipliers). In the latter paper, the augmenting function σ is twice differentiable and convex, while in the other three classes of augmented Lagrangian functions the augmenting function is assumed to be twice differentiable. In our analysis (see Theorem 4.1), the augmenting function is not necessarily convex nor differentiable. Hence, the result and the proof techniques in Theorem 4.1 are different than those in Shapiro and Sun [22] and Sun et al [23]. Moreover it is known that some nonconvex and nondifferentiable penalty functions, which are the special cases of the augmented Lagrangians with nonconvex and nondifferentiable augmenting functions, are able to provide a smaller exact penalty parameter than that of the classical l_1 penalty function, see [33, Theorem 4.9], and indeed it has been shown that nonconvex and nondifferentiable penalty functions need only smaller penalty parameters to achieve the required accuracy than the classical l_1 penalty function in some practical applications, see[34].

Next we apply a global second-order sufficient condition [26] to obtain the existence of an augmented Lagrange multiplier.

Definition 4.2 *Let X be a subset of $\mathbb{R}^n$, $x^* \in X$, $W(x^*)$ be a subset of $\mathbb{R}^n$ and $f : X \to \mathbb{R}$ be twice continuously differentiable at x^*. We say that a generalized representation condition holds for f at x^* on X with respect to $\eta(x, x^*) \in W(x^*)$ if, for every $x \in X$,*

$$f(x) = f(x^*) + \nabla f(x^*)^\top (x - x^*) + \frac{1}{2}\eta(x, x^*)^\top \nabla^2 f(x^*)\eta(x, x^*).$$

Remark 4.4 There are some functions that satisfy the generalized representation condition in Definition 4.2 (cf. Yang [26]). For example, for a linear fractional function

$$f(x) = \frac{b^\top x + c}{a^\top x + s}, \quad x \in X$$

where $a, b \in \mathbb{R}^n$, $c, s \in \mathbb{R}$ and X is a convex set such that $a^\top x + s > 0, \forall x \in X$. It is easy to show that

$$f(x) = f(x^*) + \nabla f(x^*)^\top (x - x^*) + \frac{1}{2}\frac{a^\top x^* + s}{a^\top x + s}(x - x^*)^\top \nabla^2 f(x^*)(x - x^*).$$

Then $f(x)$ satisfies the generalized representation condition on X with

$$W(x^*) = \{\eta(x, x^*) \in \mathbb{R}^n : \eta(x, x^*) = \sqrt{\frac{a^\top x^* + s}{a^\top x + s}}(x - x^*)\}.$$

Theorem 4.2 *Let $f(\cdot)$ and $g(\cdot, \omega), \omega \in \Omega$, be real-valued twice continuously differentiable functions, and let $x^* \in X_0$, $\lambda^* \in \mathbb{R}^{(\Omega)}$, $\Gamma = \mathrm{supp}(\lambda^*) = \{\omega_1, \ldots, \omega_m\}$, $V(x^*)$ be defined by (60), $\sigma : \mathbb{R}^\Omega \to \mathbb{R}^+$ have a valley at 0 respect to Γ and (48) hold, Assume that the KKT conditions (19) and (20) for problem (SIP) and conditions (i), (ii) in Theorem 4.1 are satisfied. Moreover, if there exists a set $W(x^*) \subset \mathbb{R}^n \setminus \{0\}$ such that $V(x^*) \subset W(x^*)$ and the generalized representation condition holds for $f(x)$ and $g(x, \omega_j), j = 1, \cdots, m$, at x^* on X_0^Γ with respect to the same $\eta(x, x^*) \in W(x^*)$, $\eta(\cdot, x^*) : \mathbb{R} \to W(x^*)$ is continuous,*

$$y^\top \nabla_{xx}^2 L_0^\Gamma(x^*, \lambda^*) y > 0, \quad \forall y \in W(x^*). \tag{89}$$

Then λ^* is an augmented Lagrange multiplier of (SIP).

Proof: As in the proof of Theorem 4.1, we will prove that there exists $r^* > 0$ such that (26) holds on $\mathbb{R}^n = \Theta \cup (\mathbb{R}^n \setminus \Theta)$, where Θ is defined by (79). The proof for the case in which $x \notin \Theta$ follows the same line of argument as that in Step 1 of the proof of Theorem 4.1 and hence it is omitted. Namely, Step 1 of the proof of Theorem 4.1 can be quoted to conclude that there exists $r_1^* > 0$ such that (26) holds on $\mathbb{R}^n \setminus \Theta$. To complete the proof, we prove that there exists $r_2^* > 0$ such that (26) holds on Θ. By contradiction, suppose that there exist sequences $\{r_k\}$ and $\{x_k\}$ with $r_k > 0$, $x_k \in \Theta$, such that $r_k \to +\infty$ and (83) holds. Again, similar steps as those in Theorem 4.1, Step II establish (84) and proves that $\bar{x} \in X_0^\Gamma$, where X_0^Γ is defined by (37). Because $V(x^*) \subset W(x^*)$, it follows from Lemma 4.3 that there are $r^* > 0$ and a neighborhood $N(x^*)$ of x^* such that (61) holds, i.e.,

$$L^\Omega(x, \lambda^*, r^*) > L^\Omega(x^*, \lambda^*, r^*), \quad \forall x \in N(x^*).$$

Thus, it follows from (83) that $\bar{x} \neq x^*$, since $x_k \to \bar{x}$, as $k \to \infty$. By the assumption, there is $\eta(\bar{x}, x^*) \in W(x^*)$ such that the generalized representation conditions of $f(x)$ and $g(x, \omega_j), j = 1, \cdots, m$, hold. By (89), we have

$$\eta(\bar{x}, x^*)^\top \nabla_{xx}^2 L_0^\Gamma(x^*, \lambda^*) \eta(\bar{x}, x^*) > 0.$$

Thus, for k sufficiently large, and by continuity of $\eta(\cdot, x^*)$,

$$\eta(x_k, x^*)^\top \nabla_{xx}^2 L_0^\Gamma(x^*, \lambda^*) \eta(x_k, x^*) > 0. \tag{90}$$

Thus, by the KKT conditions and (90), we have

$$f(x_k) + \sum_{j=1}^{m} \lambda^*(\omega_j) g(\omega_j, x_k)$$

$$= f(x^*) + \nabla f(x_k)^\top (x_k - x^*) + \frac{1}{2}\eta(x_k, x^*)^\top \nabla^2 f(x_k)^\top \eta(x_k, x^*) + \sum_{j=1}^{m} \lambda^*(\omega_j) g(\omega_j, x_k)$$

$$+ \sum_{j=1}^{m} \lambda^*(\omega_j) \nabla g(\omega_j, x_k)^\top (x_k - x^*) + \frac{1}{2}\sum_{j=1}^{m} \lambda^*(\omega_j)\eta(x_k, x^*)^\top \nabla_{xx}^2 g(\omega_j, x_k)\eta(x_k, x^*)$$

$$= \frac{1}{2}\eta(x_k, x^*)^\top \nabla_{xx}^2 L^\Omega(x^*, \lambda^*)\eta(x_k, x^*)$$

$$> f(x^*),$$

which contradicts to (84). Therefore, there exists $r_2^* > r^* > 0$ such that (26) holds on Θ.

Let $r^* = \max\{r_1^*, r_2^*\}$. Therefore, (26) holds on $\mathbb{R}^n$. By Lemma 2.2, λ^* is an augmented Lagrange multiplier of (SIP) with penalty parameter r^*. $\square$

Remark 4.5 If Assumption 3.1 holds, then conditions (i) and (ii) of Theorem 4.1 are satisfied (see the proof of Theorem 3.2).

5   Conclusions

In the case of a sharp Lagrangian, we obtained first-order necessary or sufficient conditions for the existence of an augmented Lagrange multiplier for (SIP). Using a valley at 0 augmenting function, we obtained second-order sufficient conditions for the existence of an augmented Lagrange multiplier for (SIP). We employed the discretization technique in our work and provided some characterizations of augmented Lagrange multipliers for (SIP) in terms of saddle points. However, we could not establish the equivalence of the (SIP) and its discretized optimization problem when an augmented Lagrange multiplier exists.

References

[1] Fiacco, A.V. and Kortanek, K.O. (1983). Semi-Infinite Programming and Applications, Lecture Notes in Econ. and Math. Systems 215, Springer, New York.

[2] Goberna, M.A. and López, M.A. (1998). Linear Semi-Infinite Optimization, Wiley, New-York.

[3] Polak, E. (1987). On the mathematical foundations of nondifferentiable optimization in engineering design. SIAM Review, 29, 21–89.

[4] Polak, E. and Wuu, T.L. (1989). On the design of stabilizing compensators via semi-infinite optimization. IEEE Transactions on Automatic Control, 34, 196–200.

[5] Polak, E. (1993). On the use of consistent approximations in the solution of semi-infinite optimization and optimal control problems, Mathematical Programming, 62, 385–414.

[6] Hettich, R. (1986). An implementation of a discretization method for semi-infinite programming, Mathematical Programming, 34, 354–361.

[7] Still, G. (2001). Discretization in semi-infinite programming: The rate of convergence, Mathematical Programming, 91, 53–69.

[8] Hettich, R. and Kortanek, K.O. (1993). Semi-infinite programming: Theory, methods and applications, SIAM Review, 35, 380–429.

[9] Reemsten, R. and Görner, S. (1998). Numerical methods for semi-infinite programming: A survey, in Semi-Infinite Programming, R. Reemsten and J. Rückmann, eds., Kluwer Academic Publishers, Boston, 195-275.

[10] Zhang, L.P., Wu, S.Y. and López, M.A. (2010). A new exchange method for convex semi-infinite programming, SIAM Journal on Optimization, 20, 2959-2977.

[11] Hestenes, M.R. (1969). Multiplier and gradient methods, Journal of Optimization Theory and Applications, 4, 303-320.

[12] Powell, M.J.D. (1969). A method for nonlinear constraints in minimization problems. In "Optimization" (R. Fletcher, ed.), Academic Press, New York, 283-298.

[13] Buys, J.D. (1972). Dual algorithms for constrained optimization problems, Ph.D. thesis, University of Leiden, Leiden, The Netherlands.

[14] Rockafellar, R.T. (1974). Augmented Lagrange multiplier functions and duality in nonconvex programming, SIAM Journal on Control and Optimization, 12, 268-285.

[15] Rockafellar, R.T. and Wets, R.J.B. (1998). Variational Analysis, Springer Verlag, Berlin, Germany.

[16] Penot, J. P. (2002). Augmented Lagrangians, duality and growth conditions, Journal of Nonlinear and Convex Analysis, 3, 283-302.

[17] Huang, X.X. and Yang, X.Q. (2003). A unified augmented Lagrangian approach to duality and exact penalization, Mathematics of Operations Research, 28, 533-552.

[18] Zhou, Y.Y. and Yang, X.Q. (2004). Some results about duality and exact penalization, Journal of Global Optimization, 29, 497-509.

[19] Burachik, R.S. and Rubinov, A. (2007). Abstract convexity and augmented Lagrangians. SIAM J. Optim. Vol. 18, No. 2, pp. 413-436.

[20] Burachik, R. S. (2011). On primal convergence for augmented Lagrangian duality. Optimization, Vol. 60, No. 8-9, pp. 979–990.

[21] Burachik, R. S., Iusem, A. N., and Melo, J. G. (2015) The exact penalty map for nonsmooth and nonconvex optimization. Optimization, Vol. 64, No. 4, 717–738.

[22] Shapiro, A. and Sun, J. (2004). Some properties of the augmented Lagrangian in cone constrained optimization, Mathematics of Operations Research, 29, 479-491.

[23] Sun, X.L., Li, D. and McKinnon, K.I.M. (2005). On saddle points of augmented Lagrangians for constrained nonconvex optimization, SIAM Journal on Optimization, 15, 1128-1146.

[24] Rückmann, J.J. and Shapiro, A. (2009). Augmented Lagrangians in semi-infinite programming, Mathematical Programming, Ser. B 116, 499-512.

[25] Zhou, Y.Y., Zhou, J.C. and Yang, X.Q. (2014). Existence of augmented Lagrange multipliers for cone constrained optimization problems. Journal of Global Optimization 58, no. 2, 243-260.

[26] Yang, X.Q. (1998). Second-order global optimality conditions for convex composite optimization, Mathematical Programming, 81, 327-347.

[27] Bazaraa, M. S., Sherali, H. D. and Shetty, C. M. (2006). Nonlinear Programming: Theory and Algorithms, third edition, Wiley-Interscience, New Jersey.

[28] Bonnans, J.F. and Shapiro, A. (2000). Perturbation Analysis of Optimization Problems, Springer, New York.

[29] Clarke, F.H. (1983). Optimization and Nonsmooth Analysis, Classics in Applied Mathematics, SIAM.

[30] Gauvin, J. (1977). A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming, Mathematical Programming, 12, 136-138.

[31] Burke, J.V. (1991). An exact penalization viewpoint of constrained optimization, SIAM Journal on Control and Optimization, 29, 968-998.

[32] Han, S.P. and Mangasarian, O.L. (1979). Exact penalty functions in nonlinear programming, Mathematical Programming, 17, 251-269.

[33] Rubinov, A. M. and Yang, X. Q. (2003). Lagrange-type Functions in Constrained Non-convex Optimization, Kluwer Academic Publishers.

[34] Wang, S., Yang, X.Q. and Teo. K.L. (2006). Power penalty method for a linear complementarity problem arising from American option valuation. Journal of Optimization Theory and Applications, 129(2), 227-254.