Separation of Nonconvex Sets with General Augmenting Functions

Angelia Nedić* and Asuman Ozdaglar†

January 8, 2007

Abstract

In this paper, we consider two geometric optimization problems that are dual to each other and characterize conditions under which the optimal values of the two problems are equal. This characterization relies on establishing separation results for nonconvex sets using general concave surfaces defined in terms of convex augmenting functions. We prove separation results for bounded-below augmenting functions, unbounded augmenting functions, and asymptotic augmenting functions.

Key words: Augmenting functions, separation of nonconvex sets, asymptotic cone, asymptotic function.

MSC2000 subject classification: Primary: 90C30, 90C46.

OR/MS subject classification: Primary: Programming/nonlinear.

1 Introduction

Convex optimization duality has been traditionally established using hyperplanes that can separate the epigraph of the primal (or perturbation) function of the constrained optimization problem from a point that does not belong to the closure of the epigraph (see, for example, [15], [9], [8], [5], [6], [2]). For a nonconvex optimization problem, the epigraph of the primal function is typically nonconvex and the separation using linear surfaces may not be possible. Recent literature has considered dual problems that are constructed using augmented Lagrangian functions, where the augmenting functions are nonnegative and satisfy either coercivity assumptions ([16], [10]), or peak-at-zero-type assumptions ([17], [18]).

*Department of Industrial and Enterprise Systems Engineering, University of Illinois, angelia@uiuc.edu

†Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, asuman@mit.edu

In our earlier work [12], we presented a geometric approach that provides a unified framework to study augmented optimization duality and penalty methods. We considered dual problems constructed using convex *nonnegative augmenting functions*, and provided necessary and sufficient conditions for zero duality gap without explicitly imposing coercivity assumptions.

In this paper, we study a similar geometric framework using more general classes of augmented Lagrangian functions. In particular, given a nonempty set $V \subset \mathbb{R}^m \times \mathbb{R}$ that intersects the w-axis, $\{(0, w) \mid w \in \mathbb{R}\}$, we consider two simple geometric optimization problems that are dual to each other. Contrary to our earlier work [12], the augmenting functions that are used to define the geometric dual problem need not be nonnegative or even bounded from below. More specifically, the geometric problems are defined as follows:

Geometric Primal Problem: The geometric primal problem consists of determining the minimum value intercept of V and the w-axis, i.e.,

$$\inf_{(0,w)\in V} w.$$

We denote the primal optimal value by w^*.

Geometric Dual Problem: To define the geometric dual problem, we consider an *augmenting function*, $\sigma : \mathbb{R}^m \mapsto (-\infty, \infty]$ defined as a function that is convex, not identically equal to 0, and satisfies $\sigma(0) = 0$ [1].

The geometric dual problem considers concave surfaces $\{(u, \phi_c(u)) \mid u \in \mathbb{R}^m\}$ that lie below the set V and $\phi_c : \mathbb{R}^m \mapsto \mathbb{R}$ has the following form

$$\phi_c(u) = -\frac{1}{c}\,\sigma(cu) + \xi, \tag{1}$$

where σ is an augmenting function, $c > 0$ is a scaling (or penalty) parameter, and $\xi \in \mathbb{R}$. This surface can be expressed as $\{(u, w) \in \mathbb{R}^m \times \mathbb{R} \mid w + \sigma(cu)/c = \xi\}$, and thus it intercepts the vertical axis $\{(0, w) \mid w \in \mathbb{R}\}$ at the level ξ. Furthermore, it is below the set V if and only if

$$w + \frac{1}{c}\,\sigma(cu) \geq \xi \quad \text{for all } (u, w) \in V.$$

Therefore, among all surfaces that are defined by an augmenting function σ and a scalar $c > 0$, and that support the set V from below, the maximum intercept with the w-axis is given by

$$d(c) = \inf_{(u,w)\in V} \left\{ w + \frac{1}{c}\,\sigma(cu) \right\}.$$

The dual problem consists of determining the maximum of these intercepts over $c > 0$, i.e.,

$$\sup_{c>0} d(c).$$

[1] This definition of augmenting function is motivated by the one introduced by Rockafellar and Wets [16] (see Definition 11.55).

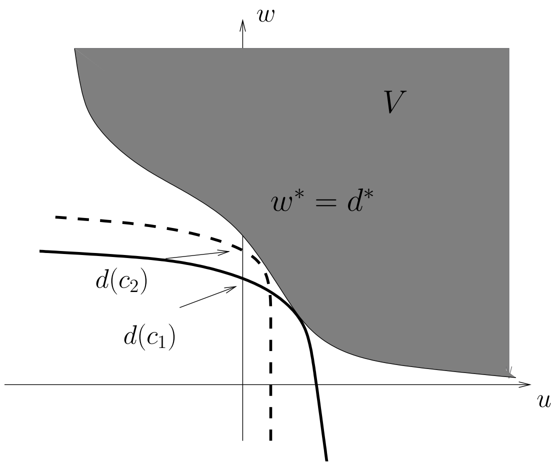

Figure 1: Geometric primal and dual problems. The figure illustrates the maximum intercept of a concave surface supporting the set V, $d(c)$ for $c_2 \geq c_1$.

We denote the dual optimal value by d^*. From the construction of the dual problem, it can be seen that the dual optimal value d^* does not exceed w^* (see Figure 1), i.e., the *weak duality* relation

$$d^* \leq w^*$$

holds (see also Proposition 1 in [12]). We say that *there is zero duality gap* when $d^* = w^*$, and *there is a duality gap* when $d^* < w^*$.

Our goal is to provide necessary and sufficient conditions for zero duality gap between the geometric primal and dual problems. To present sufficient conditions for zero duality gap, we first establish conditions that guarantee the strong separation of the set V from a point $(0, w_0)$ that does not belong to the closure of the set V. This separation is realized through the use of concave surfaces defined by augmenting functions [cf. Eq. (1)]. In particular, we say that the augmenting function σ *strongly separates* the set V and the vector $(0, w_0) \notin \mathrm{cl}(V)$ when for some $c > 0$ and $\xi \in \mathbb{R}$,

$$w + \frac{1}{c}\,\sigma(cu) \geq \xi > w_0 \quad \text{for all } (u, w) \in V.$$

We consider separating surfaces constructed by using bounded-below augmenting functions, unbounded augmenting functions, and asymptotic augmenting functions (see Figure 2). Our zero duality gap results for the two geometric problems can be used to establish zero duality gap for constrained (augmented) optimization duality and to establish convergence of general classes of penalty methods without imposing compactness assumptions. This can be done using a line of analysis similar to that of our paper [12], where we established such results for a class of nonnegative augmenting functions. We do not provide this analysis here, and a reader interested in this should refer to [12].

We further note that, the geometric dual problem considered in [12] involves the separation of the set V from a point not belonging to its closure using concave surfaces of the form

$$\phi_c(u) = -c\sigma(u) + \xi,$$

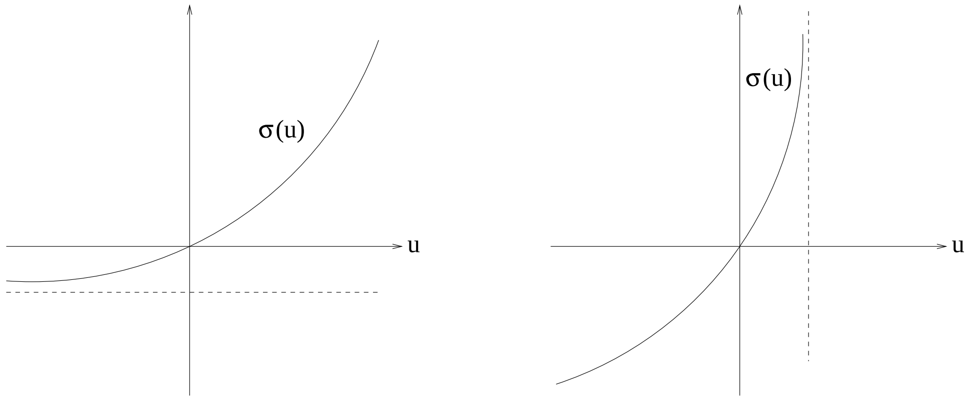

Figure 2: General augmenting functions $\sigma(u)$ for $u \in \mathbb{R}$. Part (a) illustrates a bounded-below augmenting function, e.g., $\sigma(u) = a(e^u - 1)$. Part (b) illustrates an unbounded augmenting function, e.g., $\sigma(u) = -\log(1-u)$, which is also an asymptotic augmenting function, i.e, $\sigma(u)$ goes to plus infinity along a finite asymptote.

i.e., we considered scaling the augmenting functions linearly by a parameter $c \geq 0$. Under some conditions, we have shown that for nonnegative augmenting functions σ and for any possible "dent in the bottom shape of the set V", there exists a large enough scalar $c > 0$ such that the desired separation can be achieved using linear scaling. However, for augmenting functions that can take negative values, such a scaling may not achieve the desired separation as shown in the following example: Consider the set V and the exponential augmenting function $\sigma(u) = e^u - 1$, illustrated in Figure 3. The point $(0, w_0)$ with $w_0 < 0$ does not belong to the closure of the set V. However, it can be seen that there does not exist a finite $c \geq 0$ for which the set V and the point $(0, w_0)$ can be separated using a concave surface of the form

$$\phi_c(u) = -c\sigma(u) + \xi$$

(cf. Figure 3). In this paper, we will show that for augmenting functions that are not necessarily nonnegative, a scaling of the form $\frac{1}{c}\,\sigma(cu)$ will achieve the desired separation with a large enough value of the scaling parameter c.

The paper is organized as follows: In Section 2, we present some properties of the set V and the augmenting functions, and we analyze the implications related to separation properties. In Section 3, we establish our separation theorems. In particular, we discuss some sufficient conditions on augmenting functions and the set V that guarantee the separation of this set and a vector $(0, w_0)$ that does not belong to the closure of the set V. We present our separation results for exponential-type, unbounded, and asymptotic augmenting functions. In Section 4, we provide necessary and sufficient conditions for strong duality between the geometric primal and the geometric dual problems. In Section 5, we give some concluding comments.

1.1 Notation, Terminology, and Basics

Here, we introduce our notation and basic terminology. We view a vector as a column vector, and we denote the inner product of two vectors x and y by $x'y$. For a scalar

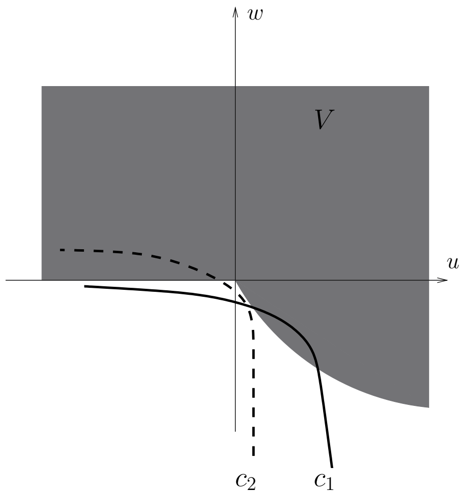

Figure 3: Separating the point $(0, w_0)$, with $w_0 < 0$, from the set V using concave surfaces of the form $-c_i(e^u - 1) + \xi$, with $c_2 > c_1$.

sequence $\{\gamma_k\}$ approaching the zero value monotonically from above, we write $\gamma_k \downarrow 0$.

For any vector $u \in \mathbb{R}^n$, we can write

$$u = u^+ + u^- \qquad \text{with} \;\; u^+ \geq 0 \;\; \text{and} \;\; u^- \leq 0,$$

where the vector u^+ is the component-wise maximum of u and the zero vector, i.e.,

$$u^+ = (\max\{0, u_1\}, ..., \max\{0, u_n\})',$$

and the vector u^- is the component-wise minimum of u and the zero vector, i.e.,

$$u^- = (\min\{0, u_1\}, ..., \min\{0, u_n\})'.$$

For a function $f : \mathbb{R}^n \mapsto [-\infty, \infty]$, we denote the domain of f by $\text{dom}(f)$, i.e.,

$$\text{dom}(f) = \{x \in \mathbb{R}^n \mid f(x) < \infty\}.$$

We denote the epigraph of f by $\text{epi}(f)$, i.e.,

$$\text{epi}(f) = \{(x, w) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq w\}.$$

For any scalar γ, we denote the (lower) γ-level set of f by $L_f(\gamma)$, i.e.,

$$L_f(\gamma) = \{x \in \mathbb{R}^n \mid f(x) \leq \gamma\}.$$

We say that the function f is *level-bounded* when the set $L_f(\gamma)$ is bounded for every scalar γ.

We denote the closure of a set X by $\text{cl}(X)$. We define a cone K as a set of all vectors x such that $\lambda x \in K$ whenever $x \in K$ and $\lambda \geq 0$. For a given nonempty set X, the cone generated by the set X is denoted by $\text{cone}(X)$ and is given by

$$\text{cone}(X) = \{y \mid y = \lambda x \text{ for some } x \in X \text{ and } \lambda \geq 0\}.$$

When establishing the separation results, we use the notion of an asymptotic cone of a set. In particular, the *asymptotic cone* of a set C (see Sec. 2.1 of [2]), also called the *horizon cone* (see Sec. 3.B of [16]), is denoted by C^∞ and is defined as follows.

Definition 1 (*Asymptotic Cone*) The asymptotic cone C^∞ of a nonempty set C is given by

$$C^\infty = \{d \mid \lambda_k x_k \to d \text{ for some } \{x_k\} \subset C \text{ and } \{\lambda_k\} \subset \mathbb{R} \text{ with } \lambda_k \geq 0, \lambda_k \to 0\}.$$

A direction $d \in C^\infty$ is referred to as an *asymptotic direction* of the set C.

We also use the notion of an asymptotic function in the subsequent analysis. The *asymptotic function* of a function f (see Sec. 2.5 of [2]), also called the *horizon cone* (see Sec. 3.C of [16]), denoted by f^∞, is defined through the asymptotic cone of the epigraph $\text{epi}(f)$ of the function f, as follows.

Definition 2 (*Asymptotic Function*) The asymptotic function f^∞ of a proper function $f : \mathbb{R}^n \mapsto (-\infty, \infty]$ is defined in terms of its epigraph as

$$\text{epi}(f^\infty) = (\text{epi}(f))^\infty.$$

Some basic properties of an asymptotic function are given in the following lemma (cf. Rockafellar and Wets [16], Theorem 3.21).

Lemma 1 (*Asymptotic Function Properties*) Let f^∞ be the asymptotic function of a proper function $f : \mathbb{R}^n \mapsto (-\infty, \infty]$. Then, we have:

(a) The following relation holds

$$f^\infty(y) = \inf \left\{ \liminf_{k\to\infty} \frac{f(c_k y_k)}{c_k} \text{ for some } c_k \to \infty \text{ and } y_k \to y \right\},$$

where the infimum is taken over all sequences $\{y_k\}$ with $y_k \to y$ and all sequences $\{c_k\}$ with $c_k \to \infty$.

(b) The function f^∞ is closed and positively homogeneous, i.e.,

$$f^\infty(\lambda y) = \lambda f^\infty(y) \qquad \text{for all } y \text{ and all } \lambda > 0.$$

(c) When f is proper, closed, and convex, the function f^∞ is also convex, and we have

$$f^\infty(y) = \lim_{c\to\infty} \frac{f(\bar{y} + cy) - f(\bar{y})}{c} \qquad \text{for all } \bar{y} \in \text{dom}(f).$$

2 Preliminary Results

In this section, we establish some properties of the set V and the augmenting function σ, which will be essential in our subsequent separation results.

2.1 Properties of the set V

We first present some properties of the set V and establish their implications.

Definition 3 We say that a set $V \subset \mathbb{R}^m \times \mathbb{R}$ is *extending upward in w-space* (*extending upward in u-space*) if for every vector $(\bar{u}, \bar{w}) \in V$, the half-line $\{(\bar{u}, w) \mid w \geq \bar{w}\}$ (the cone $\{(u, \bar{w}) \mid u \geq \bar{u}\}$) is contained in V.

The preceding properties just defined are satisfied, for example, when the set V is the epigraph of a nonincreasing function. For the sets that are extending upward in w-space and u-space, we have the following lemma, which will be used in the subsequent analysis. The proofs of these results can be found in our earlier work [12].

Lemma 2 Let $V \subset \mathbb{R}^m \times \mathbb{R}$ be a nonempty set.

(a) When the set V is extending upward in w-space, the closure $\text{cl}(V)$ of the set V is also extending upward in w-space. Furthermore, for any vector $(0, w_0) \notin \text{cl}(V)$, we have

$$w_0 < \bar{w}^* = \inf_{(0,w) \in \text{cl}(V)} w \leq w^*.$$

(b) When the set V is extending upward in u-space, the closure $\text{cl}(V)$ of the set V is also extending upward in u-space.

(c) When the set V is extending upward in u-space, the closure of the cone generated by the set V is also extending upward in u-space, i.e., for any $(\bar{u}, \bar{w}) \in \text{cl}(\text{cone}(V))$, the cone $\{(u, \bar{w}) \mid u \geq \bar{u}\}$ is contained in $\text{cl}(\text{cone}(V))$.

The proofs of some of our separation results require the separation of the half-line $\{(0, w) \mid w \leq \bar{w}\}$ for some $\bar{w} < 0$ and the cone generated by the set V. To avoid complications that may arise when the set V has an infinite slope around the origin, we consider the cone generated by a slightly upward translation of the set V in w-space, a set we denote by $\tilde{V}$.

For the separation results, another important characteristic of the set V is the "bottom-shape" of V. In particular, it is desirable that the changes in w are commensurate with the changes in u for $(u, w) \in V$, i.e., the ratio of $\|u\|$ and w-values is asymptotically finite, as w decreases. To characterize this, we use the notion of asymptotic directions and asymptotic cone of a nonempty set (see Section 1.1).

In the next lemma, we study the implications of the asymptotic directions of set V on the properties of the cone generated by the set $\tilde{V}$. This result plays a key role in the subsequent development. The result is an extension of the similar result established in our paper [12], therefore the proof is omitted.

Lemma 3 Let $V \subset \mathbb{R}^m \times \mathbb{R}$ be a nonempty set. Assume that $w^* = \inf_{(0,w) \in V} w$ is finite, and that V extends upward in u-space and w-space. Consider a nonzero vector $(\bar{u}, \bar{w}) \in \mathbb{R}^m \times \mathbb{R}$ that satisfies

$$\text{either} \quad \bar{u} = 0, \ \bar{w} < 0, \qquad \text{or} \quad \bar{u} \leq 0, \ \bar{w} = 0.$$

Assume that $(\bar{u}, \bar{w})$ is not an asymptotic direction of V, i.e.,

$$(\bar{u}, \bar{w}) \notin V^\infty.$$

For a given $\epsilon > 0$, consider the set $\tilde{V}$ given by

$$\tilde{V} = \{(u, w) \mid (u, w - \epsilon) \in V\}, \tag{2}$$

and the cone generated by $\tilde{V}$, denoted by K. Then, the vector $(\bar{u}, \bar{w})$ does not belong to the closure of the cone K generated by $\tilde{V}$, i.e.,

$$(\bar{u}, \bar{w}) \notin \mathrm{cl}(K).$$

2.2 Separation Properties of Augmenting Functions

In this section, we analyze the properties of augmenting functions under some conditions. These properties are crucial for our proofs of the separation results in Section 3.

We start with augmenting functions that are bounded from below. Here, we consider a nonempty cone $C \subset \mathbb{R}^m \times \mathbb{R}$ and a family of nonempty and convex sets parametrized by a scalar $c > 0$. Each of these convex sets is related to a level set of the function σ_c, a scaled version of the augmenting function σ. Under some conditions on the function σ, we show that if the family of the convex sets and the cone C have no vectors in common then there is a large enough scalar c and a convex set $X_c \subset \mathbb{R}^m \times \mathbb{R}$ defined in terms of σ_c having no vector in common with the cone C. In particular, we have the following lemma.

Lemma 4 Assume that:

(a) The function $\sigma : \mathbb{R}^m \mapsto (-\infty, \infty]$ is an augmenting function satisfying the following two conditions:

  (i) The function σ is bounded-below, i.e.,

$$\sigma(u) \geq \sigma_0 \quad \text{for some scalar } \sigma_0 \text{ and for all } u.$$

  (ii) For any vector sequence $\{u_k\} \subset \mathbb{R}^m$ and any positive scalar sequence $\{c_k\}$ with $c_k \to \infty$, if the relation $\limsup_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} < \infty$ holds, then the nonnegative part of the sequence $\{u_k\}$ converges to zero, i.e.,

$$\limsup_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} < \infty \text{ with } \{u_k\} \subset \mathbb{R}^m \text{ and } c_k \to \infty \quad \Rightarrow \quad u_k^+ \to 0,$$

  where $u^+ = (\max\{0, u_1\}, \ldots, \max\{0, u_m\})$.

(b) The cone $C \subset \mathbb{R}^m \times \mathbb{R}$ is nonempty and closed. Furthermore, it extends upward in u-space and does not contain any vector of the form $(0, w)$ with $w < 0$.

(c) For all $\gamma > 0$, there exists $\bar{c} > 0$ such that

$$\{(u, \tilde{w}) \mid u \in L_{\sigma_c}(\gamma)\} \cap C = \emptyset \qquad \text{for all } c \geq \bar{c},$$

where $\tilde{w} < 0$ and $L_{\sigma_c}(\gamma)$ is a γ-level set of the function $\sigma_c(u) = \sigma(cu)/c$ for all u.

Then, there exists a large enough $c_0 > 0$ such that, for all $c \geq c_0$, the set X_c defined by

$$X_c = \left\{ (u, w) \in \mathbb{R}^m \times \mathbb{R} \;\middle|\; w \leq -\sigma_c(u) + \tilde{w} \right\} \tag{3}$$

has no vector in common with the cone C.

Proof. To arrive at a contradiction assume that there is no $c_0 > 0$ such that $X_c \cap C = \emptyset$ for all $c \geq c_0$. Then, there exist a positive scalar sequence $\{c_k\}$ with $c_k \to \infty$ and a vector sequence $\{(u_k, w_k)\} \subset \mathbb{R}^m \times \mathbb{R}$ such that

$$(u_k, w_k) \in X_{c_k} \cap C \quad \text{for all } k. \tag{4}$$

In particular, by the definition of X_c in Eq. (3), we have that

$$w_k \leq -\frac{\sigma(c_k u_k)}{c_k} + \tilde{w} \quad \text{for all } k, \tag{5}$$

where $\tilde{w} < 0$.

We now consider separately the two possible cases: $w_k \geq \tilde{w}$ for infinitely many indices k, and $w_k < \tilde{w}$ for infinitely many indices k.

Case 1: $w_k \geq \tilde{w}$ infinitely often.
By taking an appropriate subsequence if necessary, we may assume without loss of generality that $w_k \geq \tilde{w}$ for all k. Then, from relation (5) we have that

$$\tilde{w} \leq w_k \leq -\frac{\sigma(c_k u_k)}{c_k} + \tilde{w} \quad \text{for all } k. \tag{6}$$

This relation implies that $\sigma(c_k u_k)/c_k \leq 0$ for all k, and therefore

$$\limsup_{k \to \infty} \frac{\sigma(c_k u_k)}{c_k} \leq 0. \tag{7}$$

Since $\sigma(u) \geq \sigma_0$ [cf. condition (i) in part (a)] and $c_k \to \infty$, it follows that

$$\liminf_{k \to \infty} \frac{\sigma(c_k u_k)}{c_k} \geq 0.$$

From the preceding inequality and relation (7) we see that $\frac{1}{c_k}\,\sigma(c_k u_k) \to 0$, and by using this in relation (6), we obtain

$$\lim_{k \to \infty} w_k = \tilde{w}.$$

We have that $(u_k, w_k) \in C$ [cf. Eq. (4)] and $u_k \leq u_k^+$ for all k, and because the cone C extends upward in u-space [cf. part (b) of the assumptions], it follows that

$$(u_k^+, w_k) \in C \qquad \text{for all } k.$$

In view of relation (7) and condition (ii) in part (a) of the assumptions, we have $u_k^+ \to 0$. Because C is closed, the relations $u_k^+ \to 0$ and $w_k \to \tilde{w}$ imply that $(0, \tilde{w}) \in C$ with $\tilde{w} < 0$.

This, however, contradicts the assumption that C does not contain any vector of the form $(0, w)$ with $w < 0$ [cf. part (b) of the assumptions].

Case 2: $w_k < \tilde{w}$ infinitely often.
Again, by taking an appropriate subsequence, we may assume without loss of generality that

$$w_k < \tilde{w} \quad \text{for all } k, \tag{8}$$

with $\tilde{w} < 0$. Note that the set X_c can be viewed as the zero-level set of the function $F_c(u, w) = w - \tilde{w} + \sigma_c(u)$ with $\sigma_c(u) = \sigma(cu)/c$ for all u, i.e.,

$$X_c = \{(u, w) \in \mathbb{R}^m \times \mathbb{R} \ \mid F_c(u, w) \leq 0\}.$$

By the definition of the augmenting function, we have that σ is convex and $\sigma(0) = 0$. Therefore, for any $c > 0$, the function $F_c(u, w)$ is convex and $F_c(0, \tilde{w}) = 0$. Consequently, for any $c > 0$, the set X_c is convex and contains the vector $(0, \tilde{w})$.

Consider now the scalars

$$\alpha_k = \frac{\tilde{w}}{w_k},$$

for which we have $\alpha_k \in (0, 1)$ for all k by Eq. (8) and $\tilde{w} < 0$. From the convexity of the sets X_{c_k}, and the relations $(0, \tilde{w}) \in X_{c_k}$ and $(u_k, w_k) \in X_{c_k}$ for all k [cf. Eq. (4)], it follows that

$$\alpha_k(u_k, w_k) + (1 - \alpha_k)(0, \tilde{w}) = (\alpha_k u_k, \tilde{w} + (1 - \alpha_k)\tilde{w}) \in X_{c_k} \quad \text{for all } k.$$

Using the definition of the set X_c [cf. Eq. (3)], we obtain

$$\tilde{w} + (1 - \alpha_k)\tilde{w} \leq -\sigma_{c_k}(\alpha_k u_k) + \tilde{w},$$

implying that

$$(1 - \alpha_k)\tilde{w} \leq -\sigma_{c_k}(\alpha_k u_k).$$

Since $\tilde{w} < 0$ and $\alpha_k \in (0, 1)$ for all k, it follows that

$$\sigma_{c_k}(\alpha_k u_k) \leq -(1 - \alpha_k)\tilde{w} < -\tilde{w} = |\tilde{w}| \quad \text{for all } k.$$

Hence, for all k, the vector $\alpha_k u_k$ belongs to level set $L_{\sigma_{c_k}}(\gamma)$ with $\gamma = |\tilde{w}| > 0$, and therefore

$$(\alpha_k u_k, \tilde{w}) \in \{(u, \tilde{w}) \mid u \in L_{\sigma_{c_k}}(\gamma)\} \quad \text{for } \gamma = |\tilde{w}| \ \text{ and for all } \ k.$$

On the other hand, the vector (u_k, w_k) lies in the cone C for all k [cf. Eq. (4)], and because $\alpha_k > 0$, the vector $\alpha_k(u_k, w_k) = (\alpha_k u_k, \tilde{w})$ also lies in the cone C. Therefore

$$(\alpha_k u_k, \tilde{w}) \in \{(u, \tilde{w}) \mid u \in L_{\sigma_{c_k}}(\gamma)\} \cap C \quad \text{for all } k,$$

with $c_k \to \infty$. However, this contradicts the assumption in part (b) that the set $\{(u, \tilde{w}) \mid u \in L_{\sigma_c}(\gamma)\}$ and the cone C have no vector in common for any $\gamma > 0$ and for all large enough c. **Q.E.D.**

We note here that the analysis of Case 2 in the preceding proof does not use any of the assumptions on the augmenting function given in part (a). In particular, it only requires the convexity of the function σ and the relation $\sigma(0) = 0$, which hold for any augmenting function.

The implication of Lemma 4 is that there exists a large enough $c_0 > 0$ such that, for all $c \geq c_0$, the set $S_c = \{(u, 2\tilde{w}) \mid u \in L_{\sigma_c}(|\tilde{w}|)\}$ and the cone C are separated by the concave function

$$\phi_c(u) = -\frac{1}{c}\,\sigma(cu) + \tilde{w}, \quad u \in \mathbb{R}^m.$$

In particular, the lemma asserts that for all $c \geq c_0$,

$$w \leq \phi_c(u) < z \qquad \text{for all } (u, w) \in X_c \text{ and } (u, z) \in C.$$

Moreover, one can verify that the set S_c is contained in the set X_c when $\tilde{w} < 0$ and $c > 0$. Therefore, for all $c \geq c_0$,

$$w \leq \phi_c(u) < z \qquad \text{for all } (u, w) \in S_c \text{ and } (u, z) \in C,$$

thus showing that ϕ_c separates the set S_c and the cone C.

We, now, consider the augmenting functions σ that may be unbounded from below. Under some conditions on σ, we show a result similar to that of Lemma 4.

Lemma 5 Assume that:

(a) The function $\sigma : \mathbb{R}^m \mapsto (-\infty, \infty]$ is an augmenting function satisfying the following two conditions:

  (i) For any vector sequence $\{u_k\} \subset \mathbb{R}^m$ with $u_k \to \bar{u}$ for some $\bar{u} \in \mathbb{R}^m$, and for any positive scalar sequence $\{c_k\}$ with $c_k \to \infty$, if the relation $\limsup_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} < \infty$ holds, then the vector $\bar{u}$ is nonpositive i.e.,

$$\limsup_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} < \infty \ \text{ with } u_k \to \bar{u} \text{ and } c_k \to \infty \quad \Rightarrow \quad \bar{u} \leq 0.$$

  (ii) For any vector sequence $\{u_k\} \subset \mathbb{R}^m$ with $u_k \to \bar{u}$ and $\bar{u} \leq 0$, and for any positive scalar sequence $\{c_k\}$ with $c_k \to \infty$, we have

$$\liminf_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} \geq 0.$$

(b) The cone $C \subset \mathbb{R}^m \times \mathbb{R}$ is nonempty and closed. It extends upward in u-space, and does not contain any vector of the form $(0, w)$ with $w < 0$ or a vector of the form $(u, 0)$ with $u \leq 0$, $u \neq 0$.

(c) For all $\gamma > 0$, there exists $\bar{c} > 0$ such that

$$\{(u, \tilde{w}) \mid u \in L_{\sigma_c}(\gamma)\} \cap C = \emptyset \qquad \text{for all } c \geq \bar{c},$$

where $\tilde{w} < 0$ and $L_{\sigma_c}(\gamma)$ is a γ-level set of the function $\sigma_c(u) = \sigma(cu)/c$ for all u.

Then, there exists a large enough $c_0 > 0$ such that, for all $c \geq c_0$, the set X_c defined by

$$X_c = \left\{ (u, w) \in \mathbb{R}^m \times \mathbb{R} \;\middle|\; w \leq -\sigma_c(u) + \tilde{w} \right\} \tag{9}$$

has no vector in common with the cone C.

Proof. To arrive at a contradiction assume that there is no c_0 such that $X_c \cap C = \emptyset$ for all $c \geq c_0$. Then, there exist a scalar sequence $\{c_k\}$ with $c_k \to \infty$ and a vector sequence $\{(u_k, w_k)\} \subset \mathbb{R}^m \times \mathbb{R}$ such that

$$(u_k, w_k) \in X_{c_k} \cap C \quad \text{for all } k. \tag{10}$$

In particular, by the definition of X_c in Eq. (9), we have that

$$w_k \leq -\frac{\sigma(c_k u_k)}{c_k} + \tilde{w} \quad \text{for all } k, \tag{11}$$

where $\tilde{w} < 0$.

We now consider separately the two cases: $w_k \geq \tilde{w}$ for infinitely many indices k and $w_k < \tilde{w}$ for infinitely many indices k.

Case 1: $w_k \geq \tilde{w}$ infinitely often.
By taking an appropriate subsequence if necessary, we may assume without loss of generality that $w_k \geq \tilde{w}$ for all k. Then, from relation (11) we have that

$$\tilde{w} \leq w_k \leq -\frac{\sigma(c_k u_k)}{c_k} + \tilde{w} \quad \text{for all } k. \tag{12}$$

Suppose that $\{u_k\}$ is bounded, and without loss of generality we may assume that $u_k \to \bar{u}$. From the preceding relation it follows that

$$\limsup_{k \to \infty} \frac{\sigma(c_k u_k)}{c_k} \leq 0.$$

Since $c_k \to \infty$ and $u_k \to \bar{u}$, by condition (i) in part (a), we have that $\bar{u} \leq 0$. Then, by condition (ii) in part (a), we further have

$$\liminf_{k \to \infty} \frac{\sigma(c_k u_k)}{c_k} \geq 0.$$

Therefore, $\frac{1}{c_k} \sigma(c_k u_k) \to 0$, and by using this in Eq. (12), we obtain $w_k \to \tilde{w}$. Because C is closed [see assumptions in part (b)] and $(u_k, w_k) \in C$ for all k [cf. Eq. (10)] with $u_k \to \bar{u}$ and $w_k \to \tilde{w}$, it follows that $(\bar{u}, \tilde{w}) \in C$ where $\bar{u} \leq 0$ and $\tilde{w} < 0$. By assumptions in part (b), the cone C is extending upward in u-space, and therefore

$$(0, \tilde{w}) \in C \quad \text{with } \tilde{w} < 0.$$

This however, contradicts the assumption that the cone C does not contain any vector of the form $(0, w)$ with $w < 0$ [cf. assumptions in part (b)].

Suppose now that $\{u_k\}$ is unbounded, and let us assume without loss of generality that $\|u_k\| \to \infty$. By dividing with $\|u_k\|$ in Eq. (12), we obtain

$$\frac{\tilde{w}}{\|u_k\|} \le \frac{w_k}{\|u_k\|} \le -\frac{\sigma(\lambda_k v_k)}{\lambda_k} + \frac{\tilde{w}}{\|u_k\|} \quad \text{for all } k, \tag{13}$$

where

$$\lambda_k = c_k \|u_k\|, \qquad v_k = \frac{u_k}{\|u_k\|}.$$

Note that v_k is bounded, and again, let us assume without loss of generality that $v_k \to \bar{v}$ for some vector $\bar{v}$ with $\bar{v} \neq 0$.

Relation (13) implies that

$$\limsup_{k\to\infty} \frac{\sigma(\lambda_k v_k)}{\lambda_k} \le 0.$$

Since $\lambda_k \to \infty$ and $v_k \to \bar{v}$, by condition (i) in part (a), we have that $\bar{v} \le 0$. Then, by condition (ii) in part (a), we also have

$$\liminf_{k\to\infty} \frac{\sigma(\lambda_k v_k)}{\lambda_k} \ge 0.$$

In view of the preceding two relations, it follows that $\frac{\sigma(\lambda_k v_k)}{\lambda_k} \to 0$. By using this in Eq. (13), together with $\|u_k\| \to \infty$, we obtain

$$\lim_{k\to\infty} \frac{w_k}{\|u_k\|} = 0.$$

Since (u_k, w_k) belongs to the cone C [cf. Eq. (10)], it follows that $(v_k, w_k/\|u_k\|) \in C$. In view of the relations $v_k \to \bar{v}$ with $\bar{v} \le 0$, $\bar{v} \neq 0$ and $w_k/\|u_k\| \to 0$, and the closedness of C [cf. assumptions in part (b)], it follows that

$$(\bar{v}, 0) \in C \qquad \text{with } \bar{v} \le 0 \text{ and } \bar{v} \neq 0.$$

This, however, contradicts the assumption in part (b) that the cone C does not contain any vector of the form $(u, 0)$ with $u \le 0$ and $u \neq 0$.

Case 2: $w_k < \tilde{w}$ infinitely often.
The proof is the same as for *Case 2* in the proof of Lemma 4. **Q.E.D.**

3 Separation Theorems

In this section, we discuss some sufficient conditions on augmenting functions and the set V that guarantee the separation of this set and a vector $(0, w_0)$ not belonging to the closure of the set V. In Section 3.1, we focus on separation results for a set V that does not have $(0, -1)$ as its asymptotic direction. In Section 3.2, we relax this condition and establish separation results using local conditions around the origin for the set V.

3.1 Global conditions on the set V

Throughout this section, we consider a set V that has a nonempty intersection with w-axis, extends upward both in u-space and w-space, and does not have $(0,-1)$ as its asymptotic direction. These properties of V are formally imposed in the following assumption.

Assumption 1 Let $V \subset \mathbb{R}^m \times \mathbb{R}$ be a nonempty set that satisfies the following:

(a) The primal optimal value is finite, i.e., $w^* = \inf_{(0,w)\in V} w$ is finite.

(b) The set V extends upward in u-space and w-space.

(c) The vector $(0,-1)$ is not an asymptotic direction of V, i.e.,

$$(0,-1) \notin V^\infty.$$

Assumption 1(a) formally states the requirement that the set V intersects the w-axis. Assumption 1(b) is satisfied, for example, when V is the epigraph of a nonincreasing function. Assumption 1(c) formalizes the requirement that the changes in w are commensurate with the changes in u for $(u,w) \in V$. It can be viewed as a requirement on the rate of decrease of w with respect to u. As an example, suppose that V is the epigraph of the function $f(u) = -\|u\|^\beta$ for $u \in \mathbb{R}^m$ and some scalar β, i.e., $V = \mathrm{epi}(f)$. Then, we have

$$\frac{1}{|f(u)|}(u, f(u)) = \left(\|u\|^{1-\beta}, -1\right).$$

For $\beta < 0$, as $u \to 0$, we have $|f(u)| \to \infty$ and $\|u\|^{1-\beta} \to 0$. Hence, for $\beta < 0$, $\frac{1}{|f(u)|}(u, f(u)) \to (0,-1)$, implying that the direction $(0,-1)$ is an asymptotic direction of the set V. However, for $\beta > 0$, we have $|f(u)| \to \infty$ as $\|u\| \to \infty$. Furthermore, as $\|u\| \to \infty$, we have $\frac{1}{|f(u)|}(u, f(u)) \to (0,-1)$ if and only if $1-\beta < 0$. Thus, the vector $(0,-1)$ is again an asymptotic direction of V for $\beta > 1$, and it is not an asymptotic direction of V for $0 \le \beta \le 1$.

3.1.1 Bounded-Below Augmenting Functions

In this section, we establish a separation result for an augmenting function σ that is bounded from below. In particular, we consider a class of augmenting functions σ satisfying the following assumption.

Assumption 2 Let σ be an augmenting function with the following properties:

(a) The function σ is bounded-below, i.e.,

$$\sigma(u) \ge \sigma_0 \quad \text{for some scalar } \sigma_0 \text{ and for all } u.$$

(b) For any sequence $\{u_k\} \subset \mathbb{R}^m$ and any positive scalar sequence $\{c_k\}$ with $c_k \to \infty$, if the relation $\limsup_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} < \infty$ holds, then the nonnegative part of the sequence $\{u_k\}$ converges to zero, i.e.,

$$\limsup_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} < \infty \ \text{ with } \{u_k\} \subset \mathbb{R}^m \text{ and } c_k \to \infty \quad \Rightarrow \quad u_k^+ \to 0,$$

where $u^+ = (\max\{0, u_1\}, \ldots, \max\{0, u_m\})$.

Note that any nonnegative augmenting function [i.e., $\sigma_0 = 0$] satisfies Assumption 2(a). The following are some examples of the nonnegative augmenting functions that satisfy Assumption 2(b):

$$\sigma(u) = \sum_{i=1}^{m} (\max\{0, u_i\})^\beta \quad \text{with } \beta > 1$$

(cf. Luenberger [11]), where $\beta = 2$ is a popular choice (for example, see Polyak [13]);

$$\sigma(u) = (u^+)' Q u^+ \qquad \text{with } u_i^+ = \max\{0, u_i\}$$

(cf. Luenberger [11]), where Q is a symmetric positive definite matrix.

The following is an example of an augmenting function that satisfies Assumption 2 with $\sigma_0 < 0$:

$$\sigma(u) = a_1(e^{u_1} - 1) + ... + a_m(e^{u_m} - 1) \qquad \text{for } u = (u_1, ..., u_m) \in \mathbb{R}^m \text{ and } a_i > 0 \text{ for all } i.$$

This function has been considered by Tseng and Bertsekas [19] in constructing penalty and multiplier methods for convex optimization problems.

We next state our separation result for an augmenting function bounded from below.

Proposition 1 (*Bounded-Below Augmenting Function*) Let $V \subset \mathbb{R}^m \times \mathbb{R}$ be a nonempty set satisfying Assumption 1. Let σ be an augmenting function satisfying Assumption 2. Then, the set V and a vector $(0, w_0)$ that does not belong to the closure of V can be strongly separated by the function σ, i.e., there exist scalars $c_0 > 0$ and ξ_0 such that for all $c \geq c_0$,

$$w + \frac{1}{c}\sigma(cu) \geq \xi_0 > w_0 \qquad \text{for all } (u, w) \in V.$$

Proof. By Lemma 2(a), we have that

$$w_0 < \bar{w}^* = \inf_{(0,w) \in \mathrm{cl}(V)} w \leq w^*.$$

By this relation and Assumption 1(a), it follows that $\bar{w}^*$ is finite. By using the translation of space along w-axis if necessary, we may assume without loss of generality that $\bar{w}^* = 0$, so that $w_0 < 0$.

Consider the set $\tilde{V}$ given by

$$\tilde{V} = \left\{ (u, w) \;\middle|\; \left(u, w + \frac{w_0}{4}\right) \in V \right\}, \tag{14}$$

and the cone generated by $\tilde{V}$, denoted by K. The proof relies on constructing a family of convex sets $\{X_c \mid c > 0\}$, where for each c, the set X_c is defined in terms of the function $\sigma_c(u) = \sigma(cu)/c$. Moreover, each of these sets contains the vector $(0, w_0/2)$ and extends downward along the w-axis. We show that at least one of these sets does not have any vector in common with the closure of the cone K. That set is actually a surface that separates $(0, w_0/2)$ and the closure of K, and therefore, it also separates $(0, w_0)$ and V.

In particular, the proof is given in the following steps:

Step 1: We first show that for all $\gamma > 0$, there exists some $\bar{c} > 0$ such that

$$\{(u, w_0/2) \mid u \in L_{\sigma_c}(\gamma)\} \cap \mathrm{cl}(K) = \emptyset \quad \text{for all } c \geq \bar{c}, \tag{15}$$

where σ_c is a function given by

$$\sigma_c(u) = \frac{1}{c}\,\sigma(cu) \quad \text{for all } u.$$

To arrive at a contradiction, suppose that relation (15) does not hold. Then, given some $\gamma > 0$, there exist a scalar sequence $\{c_k\}$ with $c_k \to \infty$ and $c_k > 0$ for all k, and a vector sequence $\{u_k\}$ such that

$$\frac{\sigma(c_k u_k)}{c_k} \leq \gamma, \qquad (u_k, w_0/2) \in \mathrm{cl}(K). \tag{16}$$

By taking the limit superior in relation (16), we obtain

$$\limsup_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} \leq \gamma < \infty.$$

Because $c_k \to \infty$, by Assumption 2(b) on the augmenting function, we have that

$$u_k^+ \to 0.$$

The set V extends upward in u-space, and so does the set $\tilde{V}$, an upward translation of the set V along the w-axis [cf. Eq. (14)]. Then, by Lemma 2(c), the closure of the cone K generated by $\tilde{V}$ also extends upward in u-space. Since $(u_k, w_0/2) \in \mathrm{cl}(K)$ [cf. Eq. (16)] and $u_k \leq u_k^+$ for all k, it follows that

$$(u_k^+, w_0/2) \in \mathrm{cl}(K) \qquad \text{for all } k.$$

Furthermore, because $u_k^+ \to 0$, we obtain $(0, w_0/2) \in \mathrm{cl}(K)$, and therefore

$$(0, w_0) \in \mathrm{cl}(K).$$

On the other hand, since $(0, w_0) \notin \mathrm{cl}(V)$, by Lemma 3 we have that $(0, w_0) \notin \mathrm{cl}(K)$, contradicting the preceding relation. Thus, relation (15) must hold.

Step 2: We consider the sets X_c given by

$$X_c = \left\{(u, w) \in \mathbb{R}^m \times \mathbb{R} \;\middle|\; w \leq -\sigma_c(u) + \frac{w_0}{2}\right\} \qquad \text{for } c > 0. \tag{17}$$

We apply Lemma 4 to assert that there exists some sufficiently large $c_0 > 0$ such that

$$X_c \cap \mathrm{cl}(K) = \emptyset \qquad \text{for all } c \geq c_0. \tag{18}$$

In particular, we use Lemma 4 with the following identification:

$$C = \mathrm{cl}(K) \quad \text{and} \quad \tilde{w} = \frac{w_0}{2}.$$

To verify that the assumptions of Lemma 4 are satisfied, note that Assumption 2 is identical to the assumptions in part (a) of Lemma 4. Furthermore, since the set V extends upward in u-space, so does the cone $\mathrm{cl}(K)$ by Lemma 2(c). Thus, the cone $C = \mathrm{cl}(K)$ satisfies the assumptions in part (b) of Lemma 4. In view of relation (15), the condition (c) of Lemma 4 is also satisfied. Hence, by this lemma, relation (18) holds.

By using relation (18) and the definition of the set X_c in Eq. (17), it follows that for all $c \geq c_0$,

$$w > -\frac{\sigma(cu)}{c} + \frac{w_0}{2} \quad \text{for all } (u, w) \in \mathrm{cl}(K).$$

The cone K is generated by the set $\tilde{V}$, so that the preceding relation holds for all $(u, w) \in \tilde{V}$, implying that for all $c \geq c_0$,

$$w - \frac{w_0}{4} + \frac{\sigma(cu)}{c} > \frac{w_0}{2} \qquad \text{for all } (u, w) \in V.$$

Furthermore, since $w_0 < 0$, it follows that for all $c \geq c_0$,

$$w + \frac{\sigma(cu)}{c} \geq \xi_0 > w_0 \qquad \text{for all } (u, w) \in V \quad \text{and} \quad \xi_0 = \frac{3w_0}{4},$$

thus completing the proof. **Q.E.D.**

3.1.2 Unbounded Augmenting Functions

In this section, we extend the separation result of the preceding section to augmenting functions that are not necessarily bounded from below. In particular, we consider augmenting functions σ that satisfy the following assumption.

Assumption 3 Let σ be an augmenting function with the following properties:

(a) For any sequence $\{u_k\} \subset \mathbb{R}^m$ with $u_k \to \bar{u}$ and for any positive scalar sequence $\{c_k\}$ with $c_k \to \infty$, the relation $\limsup_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} < \infty$ implies that the vector $\bar{u}$ is nonpositive, i.e.,

$$\limsup_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} < \infty \;\text{ with } u_k \to \bar{u} \text{ and } c_k \to \infty \qquad \Rightarrow \qquad \bar{u} \leq 0.$$

(b) For any sequence $\{u_k\} \subset \mathbb{R}^m$ with $u_k \to \bar{u}$ and $\bar{u} \leq 0$, and for any positive scalar sequence $\{c_k\}$ with $c_k \to \infty$, we have

$$\liminf_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} \geq 0.$$

For example, Assumption 3 is satisfied for an augmenting function σ of the form

$$\sigma(u) = \sum_{i=1}^{m} \theta(u_i)$$

with the following choices of the scalar function θ:

$$\theta(x) = \begin{cases} -\log(1-x) & x < 1, \\ +\infty & x \geq 1, \end{cases}$$

(cf. modified barrier method of Polyak [14]),

$$\theta(x) = \begin{cases} \frac{x}{1-x} & x < 1, \\ +\infty & x \geq 1, \end{cases}$$

(cf. hyperbolic MBF method of Polyak [14]),

$$\theta(x) = \begin{cases} x + \frac{1}{2}x^2 & x \geq -\frac{1}{2}, \\ -\frac{1}{4}\log(-2x) - \frac{3}{8} & x < -\frac{1}{2}, \end{cases}$$

(cf. quadratic logarithmic method of Ben-Tal and Zibulevski [3]).

More generally, we have the following result:

Lemma 6 Let $\theta : \mathbb{R} \mapsto (-\infty, \infty]$ be a closed convex function with $\theta(0) = 0$, and having the asymptotic function θ^∞ that satisfies $\theta^\infty(-1) = 0$ and $\theta^\infty(1) = +\infty$. Then, the function $\sigma : \mathbb{R}^m \mapsto (-\infty, \infty]$ defined by

$$\sigma(u) = \sum_{i=1}^{m} \theta(u_i),$$

is an augmenting function that satisfies Assumption 3.

Proof. Since θ is a convex function, not identically equal to 0 [because $\theta^\infty(1) = +\infty$], with $\theta(0) = 0$, it follows immediately that σ is an augmenting function.

By the assumption that $\theta^\infty(-1) = 0$ and $\theta^\infty(1) = +\infty$, and by the positive homogeneity of the function θ^∞ [cf. Lemma 1(b)], we have

$$\theta^\infty(x) = 0 \quad \text{for all } x < 0, \tag{19}$$

$$\theta^\infty(x) = +\infty \quad \text{for all } x > 0. \tag{20}$$

The function θ is proper, closed, and convex. Moreover, any scalar x with $x < 0$ is in the domain of θ^∞ [cf. Eq. 19]. Then, by using Lemma 1(c) with $y = 0$ and $\bar{y} < 0$, we obtain

$$\theta^\infty(0) = 0. \tag{21}$$

Now, consider a vector sequence $\{u_k\} \subset \mathbb{R}^m$ with $u_k \to \bar{u}$, and a positive scalar sequence $\{c_k\}$ with $c_k \to \infty$. Furthermore, let $\{u_{ki}\}$ denote the scalar sequence formed by the i^{th} components of the vectors u_k. Note that $\theta^\infty(x) \geq 0$ for all scalars x, and by using Lemma 1(a), we can see that

$$0 \leq \theta^\infty(\bar{u}_i) \leq \liminf_{k\to\infty} \frac{\theta(c_k u_{ki})}{c_k} \qquad \text{for all } i = 1, \ldots, m. \tag{22}$$

By summing over all $i = 1, \ldots, m$, we obtain

$$0 \leq \sum_{i=1}^{m} \liminf_{k \to \infty} \frac{\theta(c_k u_{ki})}{c_k} \leq \liminf_{k \to \infty} \sum_{i=1}^{m} \frac{\theta(c_k u_{ki})}{c_k} = \liminf_{k \to \infty} \frac{\sigma(c_k u_k)}{c_k}. \tag{23}$$

This relation implies that σ satisfies Assumption 3(b).

Suppose now that

$$\limsup_{k \to \infty} \frac{\sigma(c_k u_k)}{c_k} < +\infty.$$

Then, in view of relation (23), we have that

$$\sum_{i=1}^{m} \liminf_{k \to \infty} \frac{\theta(c_k u_{ki})}{c_k} \leq \liminf_{k \to \infty} \frac{\sigma(c_k u_k)}{c_k} \leq \limsup_{k \to \infty} \frac{\sigma(c_k u_k)}{c_k} < +\infty,$$

implying by Eq. (22) that $\theta^\infty(\bar{u}_i) < +\infty$ for all i. Since $\theta^\infty(x)$ is finite only for $x \leq 0$ [see Eqs. (19)–(21)], it follows that $\bar{u}_i \leq 0$ for all i. Thus, σ satisfies Assumption 3(a). **Q.E.D.**

The convergence behavior of penalty functions that satisfy the assumptions of Lemma 6 was studied by Auslender, Cominetti, and Haddou [1] for convex optimization problems (see the second class of penalty methods studied there).

To establish a separation result for an augmenting function with possibly unbounded values, we use an additional assumption on the set V.

Assumption 4 For any $\bar{u} \in \mathbb{R}^m$ with $\bar{u} \leq 0$ and $\bar{u} \neq 0$, the vector $(\bar{u}, 0)$ is not an asymptotic direction of the set V, i.e.,

$$(\bar{u}, 0) \notin V^\infty \qquad \text{for any } \bar{u} \leq 0,\ \bar{u} \neq 0.$$

We next state our separation result for unbounded augmenting functions. The proof uses a similar construction as in the proof of Proposition 1. However, here, the proof is more involved since we deal with the augmenting functions that need not be bounded from below.

Proposition 2 (*Unbounded Augmenting Function*) Let $V \subset \mathbb{R}^m \times \mathbb{R}$ be a nonempty set satisfying Assumption 1 and Assumption 4. Let σ be an augmenting function satisfying Assumption 3. Then, the set V and a vector $(0, w_0)$ that does not belong to the closure of V can be strongly separated by the augmenting function σ, i.e., there exist scalars $c_0 > 0$ and ξ_0 such that for all $c \geq c_0$,

$$w + \frac{1}{c}\,\sigma(cu) \geq \xi_0 > w_0 \quad \text{for all } (u, w) \in V.$$

Proof. By Lemma 2(a), we have that

$$w_0 < \bar{w}^* = \inf_{(0,w) \in \mathrm{cl}(V)} w \leq w^*.$$

This relation and Assumption 1(a) imply that $\bar{w}^*$ is finite. By using the translation of space along w-axis if necessary, we may assume without loss of generality that $\bar{w}^* = 0$, so that $w_0 < 0$. Again, we consider the set $\tilde{V}$ given by

$$\tilde{V} = \left\{ (u, w) \;\Big|\; \left(u, w + \frac{w_0}{4} \right) \in V \right\}, \tag{24}$$

and the cone K generated by $\tilde{V}$. The proof is given in the following steps.

Step 1: We first show that for all $\gamma > 0$, there exists some $\bar{c} > 0$ such that

$$\{(u, w_0/2) \mid u \in L_{\sigma_c}(\gamma)\} \cap \mathrm{cl}(K) = \emptyset \quad \text{for all } c \geq \bar{c}, \tag{25}$$

where σ_c is a function given by $\sigma_c(u) = \sigma(cu)/c$ for all u. To arrive at a contradiction, suppose that relation (25) does not hold. Then, for some $\gamma > 0$, there exist a sequence $\{c_k\}$ with $c_k \to \infty$ and $c_k > 0$ for all k, and a sequence $\{u_k\}$ such that

$$\frac{\sigma(c_k u_k)}{c_k} \leq \gamma, \qquad (u_k, w_0/2) \in \mathrm{cl}(K) \quad \text{for all } k. \tag{26}$$

Suppose that $\{u_k\}$ is bounded. Then, $\{u_k\}$ has a limit point $\bar{u} \in \mathbb{R}^m$, and without loss of generality we may assume that $u_k \to \bar{u}$. By taking the limit superior in the first relation of Eq. (26), we obtain

$$\limsup_{k \to \infty} \frac{\sigma(c_k u_k)}{c_k} \leq \gamma < \infty.$$

Because $c_k \to \infty$ and $u_k \to \bar{u}$, by Assumption 3(a), it follows that $\bar{u} \leq 0$. In view of the relations $(u_k, w_0/2) \in \mathrm{cl}(K)$ for all k [cf. Eq. (26)] and $u_k \to \bar{u}$, we have $(\bar{u}, w_0/2) \in \mathrm{cl}(K)$. The set V extends upward in u-space and so does the set $\tilde{V}$ [cf. Eq. (24)]. Then, by Lemma 2(c), the closure of the cone K generated by the set $\tilde{V}$ also extends upward in u-space. Since $(\bar{u}, w_0/2) \in \mathrm{cl}(K)$ and $\bar{u} \leq 0$, it follows that $(0, w_0/2) \in \mathrm{cl}(K)$, implying that

$$(0, w_0) \in \mathrm{cl}(K).$$

On the other hand, since $(0, w_0) \notin \mathrm{cl}(V)$, by Lemma 3 we have that $(0, w_0) \notin \mathrm{cl}(K)$, contradicting the preceding relation.

Suppose now that $\{u_k\}$ is unbounded. Again, by taking an appropriate subsequence, without loss of generality we may assume that $\|u_k\| \to \infty$. Then, from Eq. (26) we obtain

$$\frac{\sigma(c_k u_k)}{c_k \|u_k\|} \leq \frac{\gamma}{\|u_k\|}, \qquad \left(\frac{u_k}{\|u_k\|}, \frac{w_0}{2\|u_k\|} \right) \in \mathrm{cl}(K) \quad \text{for all } k. \tag{27}$$

We can write

$$\frac{\sigma(\lambda_k v_k)}{\lambda_k} \leq \frac{\gamma}{\|u_k\|},$$

for $\lambda_k = c_k \|u_k\| \to \infty$ and $v_k = u_k / \|u_k\|$. Note that v_k is bounded and, therefore, it has a limit $\bar{v} \in \mathbb{R}^m$ with $\bar{v} \neq 0$. Without loss of generality, we may assume that $v_k \to \bar{v}$. Furthermore, from the preceding relation and $\|u_k\| \to \infty$, it follows that

$$\limsup_{k \to \infty} \frac{\sigma(\lambda_k v_k)}{\lambda_k} \leq 0.$$

Because $c_k \to \infty$ and $v_k \to \bar{v}$, by Assumption 3(a), it follows that $\bar{v} \le 0$.

In view of the relations

$$\left(\frac{u_k}{\|u_k\|}, \frac{w_0}{2\|u_k\|} \right) \in \mathrm{cl}(K) \quad \text{for all } k$$

[cf. Eq. (27)], and $u_k/\|u_k\| \to \bar{v}$ and $\|u_k\| \to \infty$, we obtain

$$(\bar{v}, 0) \in \mathrm{cl}(K) \quad \text{with } \bar{v} \le 0 \text{ and } \bar{v} \ne 0. \tag{28}$$

At the same time, according to Assumption 4, the set V has no asymptotic direction of the form $(\bar{u}, 0)$ with $\bar{u} \le 0$ and $\bar{u} \ne 0$. Then, according to Lemma 3, the cone $\mathrm{cl}(K)$ does not have asymptotic directions of the form $(\bar{u}, 0)$ with $\bar{u} \le 0$ and $\bar{u} \ne 0$, contradicting relation (28). Thus, relation (25) must hold.

Step 2: We consider the sets X_c given by

$$X_c = \left\{ (u, w) \in \mathbb{R}^m \times \mathbb{R} \;\middle|\; w \le -\sigma_c(u) + \frac{w_0}{2} \right\} \qquad \text{for } c > 0. \tag{29}$$

We apply Lemma 5 to assert that there exists some sufficiently large c_0 such that

$$X_c \cap \mathrm{cl}(K) = \emptyset \qquad \text{for all } c \ge c_0. \tag{30}$$

In particular, we use Lemma 5 with the following identification:

$$C = \mathrm{cl}(K) \quad \text{and} \quad \tilde{w} = \frac{w_0}{2}.$$

To verify that the assumptions of Lemma 5 are satisfied, note that Assumption 3 is identical to the assumptions in part (a) of Lemma 5. We have also argued that the closure of the cone K generated by the set $\tilde{V}$ extends upward in u-space and does not have asymptotic directions of the form $(0, w)$ with $w < 0$ or $(u, 0)$ with $u \le 0$ and $u \ne 0$. Thus, the cone $C = \mathrm{cl}(K)$ satisfies the assumptions in part (b) of Lemma 5. In view of relation (25), the condition (c) in Lemma 5 is also satisfied. Hence, by this lemma, relation (30) holds.

From relation (30) and the definition of the set X_c in Eq. (29), it follows that for all $c \ge c_0$,

$$w > -\frac{\sigma(cu)}{c} + \frac{w_0}{2} \quad \text{for all } (u, w) \in \mathrm{cl}(K).$$

The cone K is generated by the set $\tilde{V}$, so that the preceding relation holds for all $(u, w) \in \tilde{V}$, implying that for all $c \ge c_0$,

$$w - \frac{w_0}{4} + \frac{\sigma(cu)}{c} > \frac{w_0}{2} \qquad \text{for all } (u, w) \in V.$$

Furthermore, since $w_0 < 0$, it follows that for all $c \ge c_0$,

$$w + \frac{\sigma(cu)}{c} \ge \xi_0 > w_0 \qquad \text{for all } (u, w) \in V \quad \text{and} \quad \xi_0 = \frac{3w_0}{4},$$

thus completing the proof. **Q.E.D.**

3.2 Local Conditions on the set V

In this section, we present separation results when the condition $(0, -1) \notin V^\infty$ is relaxed in Assumption 1. In particular, we consider sufficient conditions for the separation of a vector $(0, w_0) \notin \mathrm{cl}(V)$ and the set V by an augmenting function for the case when the direction $(0, -1)$ may be an asymptotic direction of V. However, we still assume that the set V extends upward both in u-space and w-space. Furthermore, instead of assuming that the value $w^* = \inf_{(0,w)\in V} w$ is finite, we assume that the w-components of the set V are bounded from below in some neighborhood of the origin. These properties of V are formally imposed in the following assumption.

Assumption 5 Let $V \subset \mathbb{R}^m \times \mathbb{R}$ be a nonempty set that satisfies the following:

(a) There exists a neighborhood of $u = 0$ such that the w-components of the vectors $(u, w) \in V$ are bounded for all u in that neighborhood, i.e., there exists $\delta > 0$ such that

$$\inf_{\substack{(u,w)\in V \\ \|u\|\leq\delta}} w > -\infty.$$

(b) The set V extends upward in u-space and w-space.

3.2.1 Asymptotic Augmenting Functions

Here, we establish separation results for the sets that satisfy Assumption 5 by using augmenting functions that go to infinity along a finite positive asymptote. In particular, we consider the augmenting functions that satisfy the following assumption.

Assumption 6 Let σ be an augmenting function such that:

(a) For some scalars $a_i > 0$, $i = 1, \ldots, m$, we have

$$\sigma(u) = +\infty \qquad \text{for } u \in \mathbb{R}^m \text{ with } u_i \geq a_i \text{ for some } i \in \{1, \ldots, m\}.$$

(b) For any vector sequence $\{u_k\} \subset \mathbb{R}^m$ with $u_k \to \bar{u}$ and $\bar{u} \leq 0$, and for any positive scalar sequence $\{c_k\}$ with $c_k \to \infty$, we have

$$\liminf_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} \geq 0.$$

Note that the unbounded augmenting functions of the form modified barrier method and hyperbolic MBF method of Polyak satisfy this assumption. Note also that any augmenting function σ that satisfies Assumption 6(a) also satisfies Assumption 2(a). To see this, let $\{u_k\}$ be a sequence with $u_k \to \bar{u}$ and $\{c_k\}$ be a sequence with $c_k \to \infty$, and assume that

$$\limsup_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} < \infty. \tag{31}$$

Assume to arrive at a contradiction that $\bar{u}_i > 0$ for some i. Then, since $c_k \to \infty$, we have $c_k u_{ki} > a_i$ for all sufficiently large k, implying that $\sigma(c_k u_k) = +\infty$ for all sufficiently large k, and contradicting Eq. (31).

We now present sufficient conditions for the separation of a vector $(0, w_0) \notin \mathrm{cl}(V)$ and the set V.

Proposition 3 (*Asymptotic Augmenting Function*) Let $V \subset \mathbb{R}^m \times \mathbb{R}$ be a nonempty set and σ be an augmenting function. Assume also that one of the following holds:

(a) The set V satisfies Assumption 5, and the augmenting function σ is nonnegative and satisfies Assumption 6(a).

(b) The set V satisfies Assumption 4 and Assumption 5, and the augmenting function σ satisfies Assumption 6.

Then, the set V and a vector $(0, w_0) \in \mathbb{R}^m \times \mathbb{R}$ that does not belong to the closure of V can be separated by the function σ, i.e., there exist scalars $c_0 > 0$ and ξ_0 such that for all $c \geq c_0$,

$$w + \frac{1}{c}\,\sigma(cu) \geq \xi_0 > w_0 \quad \text{for all } (u, w) \in V.$$

Proof. To arrive at a contradiction, assume that there are no scalars $c_0 > 0$ and ξ_0 such that the preceding relation holds. Then, there exist a positive scalar sequence $\{c_k\}$ with $c_k \to \infty$ and a vector sequence $\{(u_k, w_k)\} \subset V$ such that

$$w_k + \frac{\sigma(c_k u_k)}{c_k} \leq w_0 \quad \text{for all } k. \tag{32}$$

We show that, under Assumption 6(a) on the augmenting function σ, Eq. (32) implies that $u = 0$ is a limit point of the sequence $\{u_k^+\}$, i.e., $\{u_k^+\}$ converges to $u = 0$ along some subsequence. Suppose that this is not the case. Then, because $u_k^+ \geq 0$ for all k, there exist a coordinate index i and a scalar $\epsilon >$ such that $u_{ki} \geq \epsilon$ for all sufficiently large k. Then, since $c_k \to \infty$, we have

$$c_k u_{ki} \geq c_k \epsilon > a_i \qquad \text{for all sufficiently large } k,$$

where the scalar a_i is such that $\sigma(u) = +\infty$ for any vector u with $u_i \geq a_i$ [cf. Assumption 6(a)]. Hence, $\sigma(c_k u_k) = +\infty$ for all sufficiently large k, thus contradicting Eq. (32). Therefore $u = 0$ is a limit point of the sequence $\{u_k^+\}$. By taking an appropriate subsequence if necessary, we may assume without loss of generality that

$$u_k^+ \to 0 \qquad \text{and} \qquad \|u_k^+\| \leq \delta \ \text{ for all } k, \tag{33}$$

where δ is the scalar defining the local property of V, as given in Assumption 5(a).

The set V extends upward in u-space by Assumption 5(b), and since $u \leq u^+$ for all u, we have

$$(u_k^+, w_k) \in V \qquad \text{for all } k. \tag{34}$$

By Assumption 5(a), the components w are bounded from below for all $(u, w) \in V$ with $\|u\| \leq \delta$. In view of this, from the preceding two relations, we conclude that the sequence $\{w_k\}$ is bounded from below, i.e., for some w_δ we have

$$w_k \geq w_\delta \qquad \text{for all } k.$$

From now on, the arguments for parts (a) and (b) are different.

(a) Since the function σ is nonnegative [i.e., $\sigma(u) \geq 0$ for all u], from Eq. (32) it follows that

$$w_k \leq w_0 \quad \text{for all } k.$$

The preceding two relations imply that $\{w_k\}$ is bounded, and therefore it has a limit point $\bar{w}$ with $\bar{w} \leq w_0$. We may assume that $w_k \to \bar{w}$. In view of relations $u_k^+ \to 0$ and $(u_k^+, w_k) \in V$ for all k [cf. Eqs. (33) and (34)], it follows that $(0, \bar{w}) \in \mathrm{cl}(V)$. The set V extends upward in w-space by Assumption 5(b), and so does the closure of V by Lemma 2(a). Thus, since $(0, \bar{w}) \in \mathrm{cl}(V)$ and $\bar{w} \leq w_0$, we have $(0, w_0) \in \mathrm{cl}(V)$- a contradiction.

(b) Suppose that the sequence $\{u_k\}$ is bounded. Then, it has a limit point $\bar{u}$, and we may assume that $u_k \to \bar{u}$. Since $u_k^+ \to 0$ [cf. Eq. (33)], it follows that $\bar{u} \leq 0$. Therefore, by Assumption 6(b), we have that

$$\liminf_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} \geq 0.$$

By taking the limit inferior in Eq. (32) and using the preceding relation, we obtain

$$\liminf_{k\to\infty} w_k \leq \liminf_{k\to\infty} w_k + \liminf_{k\to\infty} \frac{\sigma(c_k u_k)}{c_k} \leq \liminf_{k\to\infty} \left\{ w_k + \frac{\sigma(c_k u_k)}{c_k} \right\} \leq w_0.$$

Because $w_k \geq w_\delta$ for all k, we see that the value $\tilde{w} = \liminf_{k\to\infty} w_k$ is finite, and in view of the preceding relation, we have that $\tilde{w} \leq w_0$. Thus, $w_k \to \tilde{w}$ along some subsequence, and since $u_k^+ \to 0$ and $(u_k^+, w_k) \in V$ for all k [cf. Eqs. (33) and (34)], it follows that $(0, \tilde{w}) \in \mathrm{cl}(V)$ with $\tilde{w} \leq w_0$. The set V extends upward in w-space by Assumption 5(b), and so does the closure of V by Lemma 2(a), implying that $(0, w_0) \in \mathrm{cl}(V)$- a contradiction.

Suppose now that the sequence $\{u_k\}$ is unbounded, and without loss of generality assume that $\|u_k\| \to \infty$. Consider the sequence $\{v_k\}$ where $v_k = u_k/\|u_k\|$ for all k. Note that the sequence $\{v_k\}$ is bounded, and we may assume that it is convergent, i.e., $v_k \to \bar{v}$ with $\bar{v} \neq 0$. In view of relations $v_k^+ = u_k^+/\|u_k\|$, $u_k^+ \to 0$ and $\|u_k\| \to \infty$, it follows that $v_k^+ \to 0$. Because $v_k^- \to \bar{v}^-$, and $v_k^+ \to \bar{v}^+$, we see that $\bar{v} \leq 0$. Hence,

$$v_k = \frac{u_k}{\|u_k\|} \to \bar{v} \qquad \text{with } \bar{v} \leq 0 \text{ and } \bar{v} \neq 0.$$

By dividing Eq. (32) with $\|u_k\|$, we obtain

$$\frac{w_k}{\|u_k\|} + \frac{\sigma(\lambda_k v_k)}{\lambda_k} \leq \frac{w_0}{\|u_k\|} \qquad \text{for all } k, \tag{35}$$

where $\lambda_k = c_k\|u_k\|$. Since $\lambda_k \to \infty$ and $v_k \to \bar{v}$ with $\bar{v} \leq 0$, by Assumption 6(b), we have that $\liminf_{k\to\infty} \frac{\sigma(\lambda_k v_k)}{\lambda_k} \geq 0$. Thus, by taking the limit inferior in Eq. (35), we obtain

$$\liminf_{k\to\infty} \frac{w_k}{\|u_k\|} \leq 0.$$

This relation implies that there exists a sequence $\{\bar{w}_k\}$ with $\bar{w}_k \geq w_k$ such that

$$\liminf_{k\to\infty} \frac{\bar{w}_k}{\|u_k\|} = 0.$$

Since the set V is extending upward in w-space, and $(u_k, w_k) \in V$ and $\bar{w}_k \geq w_k$ for all k, it follows that the sequence $\{(u_k, \bar{w}_k)\}$ is contained in the set V. Therefore, along some subsequence we have

$$\frac{1}{\|u_k\|}(u_k, \bar{w}_k) \to (\bar{v}, 0)$$

with $\{(u_k, \bar{w}_k)\} \subset V$, $\bar{v} \leq 0$, and $\bar{v} \neq 0$. This, however, contradicts the assumption that the set V has no asymptotic direction of the form $(u, 0)$ with $u \leq 0$ and $u \neq 0$ [cf. Assumption 4]. **Q.E.D.**

4 Necessary and Sufficient Conditions for Geometric Zero Duality Gap

In this section, we use the separation results of Section 3 to provide necessary and sufficient conditions guaranteeing that the optimal values of the geometric primal and dual problems are equal. Recall that for a given nonempty set $V \subset \mathbb{R}^m \times \mathbb{R}$, we define the geometric primal problem as

$$\inf_{(0,w)\in V} w, \tag{36}$$

and denote the primal optimal value by w^*. For a given augmenting function σ, and a scalar $c \geq 0$, we define a dual function $d(c)$ as

$$d(c) = \inf_{(u,w)\in V} \left\{ w + \frac{1}{c}\,\sigma(cu) \right\}.$$

We consider the geometric dual problem

$$\sup_{c\geq 0} d(c), \tag{37}$$

and denote the dual optimal value by d^*.

In our analysis of the necessary conditions for zero duality gap, we use the following assumption on the augmenting function.

Assumption 7 (*Continuity at the origin*) Let σ be an augmenting function that is continuous at $u = 0$.

Since an augmenting function σ is convex by definition, the assumption that σ is continuous at the origin $u = 0$ holds, for example, when the origin is in the relative interior of the domain of σ (see [6], or [15]). Note that all examples of the augmenting functions that we have considered in Section 3 satisfy this condition.

We now present a necessary condition for zero duality gap.

Proposition 4 (*Necessary Conditions for Zero Duality Gap*) Let $V \subset \mathbb{R}^m \times \mathbb{R}$ be a nonempty set. Let σ be an augmenting function that satisfies Assumption 7. Consider the geometric primal and dual problems defined in Eqs. (36) and (37). Assume that there is zero duality gap, i.e., $d^* = w^*$. Then, for any sequence $\{(u_k, w_k)\} \subset V$ with $u_k \to 0$, we have

$$\liminf_{k\to\infty} w_k \geq w^*.$$

Proof. Let $\{(u_k, w_k)\} \subset V$ be a sequence such that $u_k \to 0$. By definition, the augmenting function $\sigma(u)$ satisfies $\sigma(0) = 0$. Using this and the continuity of $\sigma(u)$ at $u = 0$ (cf. Assumption 7), we obtain for all $c > 0$,

$$\begin{aligned}
\liminf_{k\to\infty} w_k &= \liminf_{k\to\infty} w_k + \frac{1}{c}\,\sigma(c\,0) \\
&= \liminf_{k\to\infty} \left\{ w_k + \frac{1}{c}\,\sigma(cu_k) \right\} \\
&\geq \inf_{(u,w)\in V} \left\{ w + \frac{1}{c}\,\sigma(cu) \right\} \\
&= d(c).
\end{aligned}$$

Hence

$$\liminf_{k\to\infty} w_k \geq \sup_{c\geq 0} d(c) = d^*,$$

and since $d^* = w^*$, it follows that $\liminf_{k\to\infty} w_k \geq w^*$. **Q.E.D.**

We next provide sufficient conditions for zero duality gap using the assumptions of Section 3 on the set V and the augmenting function σ.

Proposition 5 (*Sufficient Conditions for Zero Duality Gap*) Let $V \subset \mathbb{R}^m \times \mathbb{R}$ be a nonempty set. Consider the geometric primal and dual problems defined in Eqs. (36) and (37). Assume that for any sequence $\{(u_k, w_k)\} \subset V$ with $u_k \to 0$, we have

$$\liminf_{k\to\infty} w_k \geq w^*.$$

Assume further that one of the following holds:

(a) The set V satisfies Assumption 1 and the augmenting function σ satisfies Assumption 2.

(b) The set V satisfies Assumption 1 and Assumption 4, and the augmenting function σ satisfies Assumption 3.

(c) The set V satisfies Assumption 5, and the augmenting function σ is nonnegative and satisfies Assumption 6(a).

(d) The set V satisfies Assumption 4 and Assumption 5, and the augmenting function σ satisfies Assumption 6.

Then, there is zero duality gap, i.e., $d^* = w^*$.

Proof. Under either Assumption 1(a) or Assumption 5(a), we have that w^* is finite. Let $\epsilon > 0$ be arbitrary, and consider the vector $(0, w^* - \epsilon)$. We show that $(0, w^* - \epsilon)$

does not belong to the closure of the set V. To obtain a contradiction, assume that $(0, w^* - \epsilon) \in \mathrm{cl}(V)$. Then, there exists a sequence $\{(u_k, w_k)\} \subset V$ with $u_k \to 0$ and $w_k \to w^* - \epsilon$, contradicting the assumption that $\liminf_{k\to\infty} w_k \geq w^*$. Hence, $(0, w^* - \epsilon)$ does not belong to the closure of the set V.

Consider the set V and the augmenting function σ that satisfy part (a) of the assumptions [alternatively, part (b), (c), or (d) of the assumptions, respectively]. Since $(0, w^* - \epsilon) \notin \mathrm{cl}(V)$, by Proposition 1 [Proposition 2, Proposition 3(a), or Proposition 3(b), respectively], it follows that there exists a scalar $c > 0$ such that

$$\inf_{(u,w)\in V} \left\{ w + \frac{1}{c}\sigma(cu) \right\} \geq w^* - \epsilon.$$

Therefore

$$d(c) \geq w^* - \epsilon,$$

implying that

$$d^* = \sup_{c\geq 0} d(c) \geq w^* - \epsilon.$$

By letting $\epsilon \to 0$, we obtain

$$d^* \geq w^*,$$

which together with the weak duality relation ($d^* \leq w^*$) implies that $d^* = w^*$. **Q.E.D.**

5 Conclusions

In this paper, we introduced two geometric optimization problems that are dual to each other and studied conditions under which the optimal values of these problems are equal. To establish this, we showed that we can use general concave surfaces to separate nonconvex sets with certain properties.

A number of concluding comments are useful:

- We provide conditions on the (nonconvex) set V under which we can strongly separate V from a point of the form $(0, w)$ that does not belong to its closure. The separation is realized by concave surfaces which are constructed using the augmenting functions. The condition $(0, -1) \notin V^\infty$ plays a key role in the separation results of Section 3.1. When applying these results to constrained optimization duality, this condition translates into assumptions on the asymptotic directions of $\mathrm{epi}(p)$. Providing general sufficient conditions on optimization problems that guarantee these assumptions is an interesting research question, which we are currently investigating.

- Here, our focus is on convex augmenting functions. In particular, the convexity assumption on the augmenting functions is used to establish separation with a few assumptions on the set V. With some additional assumptions on the set V, it is possible to extend our analysis to nonconvex augmenting functions, which is ongoing work.

- Another important issue arising in the context of constrained optimization problems is establishing conditions under which there exists a finite value for the parameter c such that $d(c) = w^*$. This will allow us to show the existence of exact penalty methods, which we plan to address in our future work.

References

[1] Auslender A., Cominetti R., and Haddou M., "Asymptotic Analysis for Penalty and Barrier Methods in Convex and Linear Programming," *Mathematics of Operations Research*, vol. 22, no. 1, pp. 43-62, 1997.

[2] Auslender A. and Teboulle, M., Asymptotic Cones and Functions in Optimization and Variational Inequalities, Springer-Verlag, New York, 2003.

[3] Ben-Tal A. and Zibulevski M., "Penalty /barrier multiplier methods: A new class of augmented Lagrangian algorithms for large scale convex prorgramming problems," Research report 4/93, Optimization Laboratory, Faculty of Industrial Engineering and Management, Technion-Israel Institute of Technology, 1993.

[4] Bertsekas D. P., Constrained Optimization and Lagrange Multiplier Methods, Athena Scientific, Belmont, MA, 1996.

[5] Bertsekas D. P., Nedić A., and Ozdaglar A., "Min Common/Max Crossing Duality: A Simple Geometric Framework for Convex Optimization and Minimax Theory," Report LIDS-P-2536, Jan. 2002.

[6] Bertsekas D. P., Nedić A., and Ozdaglar A., Convex Analysis and Optimization, Athena Scientific, Belmont, MA, 2003.

[7] Bonnans J. F. and Shapiro A., Perturbation Analysis of Optimization Problems, Springer-Verlag, New York, NY, 2000.

[8] Borwein J. M. and Lewis A. S., Convex Analysis and Nonlinear Optimization, Springer-Verlag, New York, NY, 2000.

[9] Hiriart-Urruty J.-B. Lemarechal C., Convex Analysis and Minimization Algorithms, vols. I and II, Springer-Verlag, Berlin and NY, 1993.

[10] Huang X. X. and Yang X. Q., "A unified augmented Lagrangian approach to duality and exact penalization," *Mathematics of Operations Research*, vol. 28, no. 3, pp. 533-552, 2003.

[11] Luenberger D. G., Linear and Nonlinear Programming, Kluwer Academic Publishers, MA, 2-nd edition, 2004.

[12] Nedić A. and Ozdaglar A., "A Geometric Framework for Nonconvex Optimization Duality using Augmented Lagrangian Functions," forthcoming in *Journal of Global Optimization*, 2006.

[13] Polyak B. T., Introduction to Optimization, Optimization Software Inc., N.Y., 1987.

[14] Polyak R., "Modified barrier functions: theory and methods," *Mathematical Programming*, vol. 54, pp. 177-222, 1992.

[15] Rockafellar R. T., Convex Analysis, Princeton University Press, Princeton, N.J., 1970.

[16] Rockafellar R. T. and Wets R. J.-B., Variational Analysis, Springer-Verlag, New York, 1998.

[17] Rubinov A. M., Huang X. X., and Yang X. Q., "The zero duality gap property and lower semicontinuity of the perturbation function," *Mathematics of Operations Research*, vol. 27, no. 4, pp. 775-791, 2002.

[18] Rubinov A. M. and Yang X. Q., Lagrange-type Functions in Constrained Nonconvex Optimization, Kluwer Academic Publishers, MA, 2003.

[19] Tseng P. and Bertsekas D. P., "On the convergence of the exponential multiplier method for convex programming," *Mathematical Programming*, vol. 60, pp. 1-19, 1993.