Robust Conjugate Duality for Convex Optimization under Uncertainty with Application to Data Classification*

G. Y. Li,† V. Jeyakumar‡ and G. M. Lee§

Final Version: November 18, 2010

Abstract

In this paper we present a robust conjugate duality theory for convex programming problems in the face of data uncertainty within the framework of robust optimization, extending the powerful conjugate duality technique. We first establish robust strong duality between an uncertain primal parameterized convex programming model problem and its uncertain conjugate dual by proving strong duality between the deterministic robust counterpart of the primal model and the optimistic counterpart of its dual problem under a regularity condition. This regularity condition is not only sufficient for robust duality but also is necessary for it whenever robust duality holds for every linear perturbation of the objective function of the primal model problem. More importantly, we show that robust strong duality always holds for partially finite convex programming problems under scenario data uncertainty and that the optimistic counterpart of the dual is a tractable finite dimensional problem. As an application, we also derive a robust conjugate duality theorem for support vector machines which are a class of important convex optimization models for classifying two labelled data sets. The support vector machine has emerged as a powerful modelling tool for machine learning problems of data classification that arise in many areas of application in information and computer sciences.

*The authors are grateful to the referee and the Editor for their helpful comments and valuable suggestions which have contributed to the final preparation of the paper. Research was partially supported by a grant from the Australian Research Council and the Korea Science and Engineering Foundation NRL program grant of the Korea government (MEST)(No. ROA-2008-000-20010-0).

†Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: g.li@unsw.edu.au

‡Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: v.jeyakumar@unsw.edu.au

§Department of Applied Mathematics, Pukyong National University, Busan 608-737, Korea.

Key words. Robust convex optimization, conjugate duality under uncertainty, strong duality, partially-finite convex programs, support vector machines, robust Fenchel's duality.

AMS subject classification. 90C22,90C25,90C46

1 Introduction

Duality theory is a corner stone in the area of constrained optimization and has been studied for over a century. However, real-world problems of constrained optimization often involve input data that is noisy or uncertain due to modelling or measurement errors [3]. Consequently, how to develop mathematical approaches that are capable of treating data uncertainty in constrained optimization has become a critical question in mathematical optimization. Over the years, various deterministic as well as stochastic approaches have been developed for treating uncertainty in optimization (see [1, 2, 4, 21, 22, 31, 32] and other references therein). In this paper, we examine robust optimization framework [3] for studying conjugate duality theory for constrained optimization in the face of data uncertainty.

Consider the standard form convex optimization problem in the absence of data uncertainty

$$(P) \qquad \inf_{x\in X} f(x),$$

where $f : X \to \mathbb{R} \cup \{+\infty\}$ is a proper lower semi-continuous convex functions. This problem can be embedded into a family of parameterized problems (see [35])

$$(P_y) \qquad \inf_{x\in X} \phi(x,y),$$

where the function $\phi : X\times Y \to \mathbb{R}\cup\{+\infty\}$ satisfies $\phi(x, 0) = f(x)$. Clearly, (P_0) collapses to the original problem (P). The parameterized convex optimization problem (P_y) in the face of data uncertainty can be captured by the problem

$$(P_{y,u}) \quad \inf_{x\in X} \phi_u(x,y),$$

where $\phi_u : X \times Y \to \mathbb{R}$ is a a proper lower semi-continuous convex function and u is the uncertain parameter which belongs to the uncertainty set $\mathcal{U}$. For instance, the effect of uncertain data (a_1, a_2) on the constraint of the problem

$$\min\{h(x) \mid a_1x_1 + a_2x_2 \le b\}$$

can be captured by the problem $(P_{0,u})$, $\inf_{x\in X} \phi_u(x, 0)$, where

$$\phi_u(x,y) = h(x) + \delta_{\{x\,:\, a_1(u_1)x_1+a_2(u_2)x_2\le b+y\}}(x),$$

where $\delta_C(.)$ denotes the indicator function of a set C and the parameter $u = (u_1, u_2)$ is in an interval uncertainty set $\mathcal{U} = [c_1, d_1] \times [c_2, d_2]$.

The **robust counterpart** (RP) of problem $(P_{0,u})$ is the deterministic optimization problem

$$\inf_{x \in X} \sup_{u \in \mathcal{U}} \phi_u(x, 0).$$

On the other hand, for each fixed $u \in \mathcal{U}$, the conjugate dual problem of $(P_{0,u})$ is given by

$$\max_{y^* \in Y^*} \{-\phi_u^*(0, y^*)\}.$$

The **optimistic counterpart** (ODP) of the uncertain dual problem is also a deterministic optimization problem which is given by

$$\max_{u \in \mathcal{U}} \max_{y^* \in Y^*} \{-\phi_u^*(0, y^*)\}.$$

We say that **robust strong duality** holds whenever the values of the robust counterpart and the optimistic counterpart coincide with the dual attainment, i.e.,

$$\inf_{x \in X} \sup_{u \in \mathcal{U}} \phi_u(x, 0) = \max_{u \in \mathcal{U}} \max_{y^* \in Y^*} \{-\phi_u^*(0, y^*)\}.$$

In this paper, we first establish robust strong duality under the condition that

$$\mathrm{Pr}_{X^* \times \mathbb{R}}(\bigcup_{u \in \mathcal{U}} \mathrm{epi}\phi_u^*) \text{ is } w^*\text{-closed and convex.} \tag{1.1}$$

We then show that this condition is not only sufficient for robust duality but also is necessary for robust strong duality to hold for every linear perturbation of $\phi_u(x, y)$ in the sense that, for each linear functional x^*,

$$\inf_{x \in X} \sup_{u \in \mathcal{U}} \{\phi_u(x, 0) + \langle x^*, x \rangle\} = \max_{u \in \mathcal{U}} \max_{y^* \in Y^*} \{-\phi_u^*(-x^*, y^*)\}.$$

For related recent conjugate duality results without data uncertainties, see [11, 12, 24, 17, 19, 26]. A more recent exhaustive treatment of conjugate duality in the absence of data uncertainty can be found in [8].

We also prove that robust strong duality always holds for partially finite convex programming problems under scenario data uncertainty [3] by verifying (1.1). In this case we also see that the optimistic counterpart of the dual is a tractable finite dimensional convex problem. Partially finite convex programs covers broad classes problems including constrained approximation problems and interpolation problems. For related results for partially finite convex programs, see [7, 18].

As an application, we derive robust Fenchel's duality theorem for support vector machines [34, 30] which are a class of convex optimization models for classifying two labelled data sets. The support vector machine has emerged as a powerful modelling tool for machine learning problems of data classification that arise in many areas of information and computer sciences [16, 13, 23, 33].

The outline of the paper is as follows. Section 2 presents preliminaries of convex analysis that will be used later in the paper. Section 3 develops robust conjugate duality theorems for convex optimization including cone-constrained optimization under uncertainty. Section 4 shows that robust conjugate duality theorem always holds for partially finite convex programs under scenario uncertainty and illustrates that these problems are tractable computationally. Section 5 derives robust Fenchel's duality under uncertainty, extending the classical Fenchel's duality. Section 6 provides an application of robust Fenchel's duality to a convex optimization model that arise in data classification problems. Finally, we present additional regularity conditions ensuring robust conjugate duality in the appendix.

2 Preliminaries

We begin this section by fixing notation and preliminaries of convex analysis. Let X, Y be Banach spaces. The dual space of X (resp. Y) is denoted by X^* (resp. Y^*) which consists of all bounded linear functionals on X (resp. Y). It is known that the space X^* endowed with the weak* topology is a locally convex Hausdorff space. Let $L(X;Y)$ denote the set of all the continuous linear mappings from X to Y. Let $A \in L(X;Y)$ be a continuous linear mapping. Then the adjoint mapping of A is a continuous linear mapping from Y^* to X^* defined by

$$\langle A^*(y^*), x\rangle = \langle y^*, Ax\rangle \text{ for all } y^* \in Y^*, x \in X,$$

where $\langle \cdot, \cdot \rangle$ denotes the corresponding linear action between the dual pairs. For a set C in X, the interior (resp. closure, convex hull, conical hull) of C is denoted by int$C$ (resp. $\overline{C}$, co$C$, cone$C$). If $C \subseteq X^*$, then the weak* closure of C is denoted by $\overline{C}^{w^*}$. The indicator function $\delta_C : X \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$\delta_C(x) := \begin{cases} 0, & \text{if} \quad x \in C, \\ +\infty, & \text{otherwise}. \end{cases} \tag{2.2}$$

Let S be a closed convex cone in Y. Then the (positive) polar cone of S is defined by $S^+ := \{y^* \in Y^* : \langle y^*, y\rangle \geq 0 \text{ for all } y \in S\}$. Moreover, a map $f : X \to Y$ is said to be S-convex if for all $\mu \in [0, 1]$ and $x_1, x_2 \in X$, $(1-\mu)f(x_1) + \mu f(x_2) - f((1-\mu)x_1 + \mu x_2) \in S$. For an extended real-valued function f on X, the effective domain and the epigraph are respectively defined by dom$f := \{x \in X : f(x) < +\infty\}$ and epi$f := \{(x, r) \in X \times \mathbb{R} : f(x) \leq r\}$. We say that f is proper if $f(x) > -\infty$ for all $x \in X$ and dom$f \neq \emptyset$. Moreover, if $\liminf_{x' \to x} f(x') \geq f(x)$ for all $x \in X$, we say f is a lower semicontinuous function. A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be convex if for all $\mu \in [0, 1]$, $f((1-\mu)x_1 + \mu x_2) \leq (1-\mu)f(x_1) + \mu f(x_2)$ for all $x_1, x_2 \in X$. As usual, for any convex function f on X, its conjugate function $f^*: X \to \mathbb{R} \cup \{\pm\infty\}$ is defined by $f^*(x^*) = \sup_{x \in X}\{\langle x^*, x\rangle - f(x)\}$ for all $x^* \in X^*$. Clearly, f^* is a proper weak* lower semicontinuous convex function if f is a proper lower semicontinuous convex function and $\lambda \mathrm{epi} f^* = \mathrm{epi}(\lambda f)^*$ for any $\lambda > 0$. If one of the functions f_1, f_2 is continuous at a point in the intersection of their domain, then we have

$$\mathrm{epi}(f_1 + f_2)^* = \mathrm{epi} f_1^* + \mathrm{epi} f_2^*. \tag{2.3}$$

For details see [8, 24, 25]. The following basic results of convex analysis play an important roles later in the proofs of our duality theorems.

Lemma 2.1. *(cf. [20]) Let I be an arbitrary index set and let f_i, $i \in I$, be proper lower semicontinuous convex functions on X. Suppose that there exists $x_0 \in X$ such that $\sup_{i \in I} f_i(x_0) < \infty$. Then*

$$\mathrm{epi}(\sup_{i \in I} f_i)^* = \overline{\mathrm{co} \bigcup_{i \in I} \mathrm{epi} f_i^*}^{w^*},$$

where $\sup\limits_{i \in I} f_i : X \to \mathbb{R} \cup \{+\infty\}$ is defined by $(\sup\limits_{i \in I} f_i)(x) = \sup\limits_{i \in I} f_i(x)$ for all $x \in X$.

Lemma 2.2. *(cf. [29, Theorem 2.1]) Let $h : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a proper function. Define the marginal function $\eta : X \to \mathbb{R} \cup \{\pm\infty\}$ by $\eta(x) = \inf_{y \in Y} h(x, y)$. Define $\mathrm{Pr}_{X \times \mathbb{R}}(\mathrm{epi}\, h) = \{(x, r), \exists\, y \in Y, (x, y, r) \in \mathrm{epi}\, h\}$. Then, we have*

$$\mathrm{Pr}_{X \times \mathbb{R}}(\mathrm{epi} h) \subseteq \mathrm{epi}\, \eta \subseteq \overline{\mathrm{Pr}_{X \times \mathbb{R}}(\mathrm{epi} h)}.$$

In particular, we have $\mathrm{Pr}_{X \times \mathbb{R}}(\mathrm{epi}\, h)$ is closed if and only if the marginal function η is lower semicontinuous and the infimum in $\inf_{y \in Y} h(\overline{x}, y)$ is attained whenever $\eta(\overline{x}) > -\infty$ for $\overline{x} \in X$.

3 Robust Conjugate Duality under Uncertainty

In this Section, we establish robust conjugate duality under (1.1) and provide sufficient conditions for (1.1). We begin by noting the fact that robust weak duality always holds.

Theorem 3.1. *(robust weak duality) Let X, Y, Z be Banach spaces and let $\mathcal{U}$ be a subset of Z. Let $\phi_u : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semi-continuous and convex function for any $u \in \mathcal{U}$. Then,*

$$\inf_{x \in X} \sup_{u \in \mathcal{U}} \phi_u(x, 0) \geq \sup_{u \in \mathcal{U}} \sup_{y^* \in Y^*} \{-\phi_u^*(0, y^*)\}.$$

Proof. For each $x^* \in X^*$, define

$$\eta(x^*) = \inf_{y^* \in Y^*} (\sup_{u \in \mathcal{U}} \phi_u)^*(x^*, y^*).$$

From the definition of η, we obtain that

$$\begin{aligned}
\eta^*(x) &= \sup_{x^*\in X^*}\{\langle x^*, x\rangle - \eta(x^*)\} \\
&= \sup_{x^*\in X^*}\{\langle x^*, x\rangle - \inf_{y^*\in Y^*}(\sup_{u\in\mathcal{U}}\phi_u)^*(x^*, y^*)\} \\
&= \sup_{(x^*,y^*)\in X^*\times Y^*}\{\langle (x^*, y^*), (x, 0)\rangle - (\sup_{u\in\mathcal{U}}\phi_u)^*(x^*, y^*)\} \\
&= (\sup_{u\in\mathcal{U}}\phi_u)^{**}(x, 0) \le (\sup_{u\in\mathcal{U}}\phi_u)(x, 0). \qquad (3.4)
\end{aligned}$$

Then, we see that

$$\begin{aligned}
\eta^{**}(-x^*) &= \sup_{x\in X}\{\langle -x^*, x\rangle - \eta^*(x)\} \\
&\ge \sup_{x\in X}\{\langle -x^*, x\rangle - (\sup_{u\in\mathcal{U}}\phi_u)(x, 0)\} \\
&= -\inf_{x\in X}\sup_{u\in\mathcal{U}}\{\phi_u(x, 0) + \langle x^*, x\rangle\}. \qquad (3.5)
\end{aligned}$$

As $\eta \ge \eta^{**}$, we derive that

$$\begin{aligned}
\inf_{x\in X}\sup_{u\in\mathcal{U}}\phi_u(x, 0) \ge -\eta^{**}(0) &\ge -\eta(0) \\
&= -\inf_{y^*\in Y^*}(\sup_{u\in\mathcal{U}}\phi_u)^*(0, y^*) \\
&= \sup_{y^*\in Y^*}\{-(\sup_{u\in\mathcal{U}}\phi_u)^*(0, y^*)\} \\
&\ge \sup_{u\in\mathcal{U}}\sup_{y^*\in Y^*}\{-\phi_u^*(0, y^*)\}. \qquad (3.6)
\end{aligned}$$

□

We now establish robust strong duality under (1.1).

Theorem 3.2. *Let X, Y, Z be Banach spaces and let $\mathcal{U}$ be a subset of Z. Let $\phi_u : X\times Y \to \mathbb{R}\cup\{+\infty\}$ be a proper lower semi-continuous and convex function for any $u \in \mathcal{U}$. Suppose that $\inf_{x\in X}\sup_{u\in\mathcal{U}}\phi_u(x, 0) < +\infty$ and*

$$\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*) \text{ is } w^*\text{-closed and convex.}$$

Then,

$$\inf_{x\in X}\sup_{u\in\mathcal{U}}\phi_u(x, 0) = \max_{u\in\mathcal{U}}\max_{y^*\in Y^*}\{-\phi_u^*(0, y^*)\}.$$

Proof. Suppose that $\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*)$ is closed and convex. From the weak robust duality, we may assume that

$$\inf_{x\in X}\sup_{u\in\mathcal{U}}\phi_u(x, 0) > -\infty.$$

From the assumption $\inf_{x\in X}\sup_{u\in\mathcal{U}}\phi_u(x,0)<+\infty$, we can further assume that

$$r:=\inf_{x\in X}\sup_{u\in\mathcal{U}}\phi_u(x,0)\in\mathbb{R}.$$

For each $x^*\in X^*$, define $\eta(x^*)=\inf_{y^*\in Y^*}(\sup_{u\in\mathcal{U}}\phi_u)^*(x^*,y^*)$. As η is the marginal function of $(\sup_{u\in\mathcal{U}}\phi_u)^*$, from Lemma 2.1 and Lemma 2.2, we have

$$\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*)\subseteq\mathrm{epi}\eta\subseteq\overline{\mathrm{Pr}_{X^*\times\mathbb{R}}\big(\mathrm{epi}(\sup_{u\in\mathcal{U}}\phi_u)^*\big)}^{w^*}=\overline{\mathrm{Pr}_{X^*\times\mathbb{R}}\big(\overline{\mathrm{co}\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*}^{w^*}\big)}^{w^*}.$$

Note that

$$\mathrm{Pr}_{X^*\times\mathbb{R}}\big(\overline{\mathrm{co}\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*}^{w^*}\big)\subseteq\overline{\mathrm{co}\mathrm{Pr}_{X^*\times\mathbb{R}}\big(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*\big)}^{w^*}.$$

This together with the assumption implies that $\mathrm{epi}\,\eta$ is weak* closed and

$$\mathrm{epi}\,\eta=\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*).$$

So, η is weak* lower semicontinuous and convex. Note that $\sup_{u\in\mathcal{U}}\phi_u$ is a lower semicontinuous convex function on $X\times Y$. Since $\inf_{x\in X}\sup_{u\in\mathcal{U}}\phi_u(x,0)\in\mathbb{R}$, [35, Theorem 2.3.4] shows that $(\sup_{u\in\mathcal{U}}\phi_u)^{**}$ is a proper function on $X\times Y$. So, $\sup_{u\in\mathcal{U}}\phi_u$ is proper as $\sup_{u\in\mathcal{U}}\phi_u\geq(\sup_{u\in\mathcal{U}}\phi_u)^{**}>-\infty$ and there exists $x_0\in X$ such that $\sup_{u\in\mathcal{U}}\phi_u(x_0,0)<+\infty$. This implies that $\sup_{u\in\mathcal{U}}\phi_u$ is a proper lower semicontinuous convex function on $X\times Y$ and hence $\sup_{u\in\mathcal{U}}\phi_u=(\sup_{u\in\mathcal{U}}\phi_u)^{**}$. Then, by (3.4), $\eta^*(x)=(\sup_{u\in\mathcal{U}}\phi_u)^{**}(x,0)=(\sup_{u\in\mathcal{U}}\phi_u)(x,0)$ is also proper. Thus, η^* is proper lower semicontinuous and convex and so, η is a proper weak* lower semicontinuous convex function (e.g. see [35, Theorem 2.3.3]). In particular, we have $\eta=\eta^{**}$. Then, as

$$\eta^{**}(0)=\sup_{x\in X}\{-\eta^*(x)\}=\sup_{x\in X}\{-(\sup_{u\in\mathcal{U}}\phi_u)(x,0)\}=-\inf_{x\in X}\sup_{u\in\mathcal{U}}\{\phi_u(x,0)\},$$

we see that

$$r=\inf_{x\in X}\sup_{u\in\mathcal{U}}\{\phi_u(x,0)\}=-\eta^{**}(0)=-\eta(0),\tag{3.7}$$

and so,

$$(0,-r)\in\mathrm{epi}\eta=\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*).$$

So, there exist $y^*\in Y^*$ and $u_0\in\mathcal{U}$ such that

$$(0,y^*,-r)\in\mathrm{epi}\phi_{u_0}^*.$$

Thus, we obtain that

$$r\leq-\phi_{u_0}^*(0,y^*)\leq\max_{u\in\mathcal{U}}\max_{y^*\in Y^*}\{-\phi_u^*(0,y^*)\}.$$

Hence,, the conclusion follows from the weak robust duality. □

Some sufficient conditions ensuring $\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*)$ is w^*-closed and convex are presented in the Appendix.

Now, let us show that our main condition (1.1) is indeed a characterization for robust strong duality in the sense that it holds if and only if robust strong duality holds for each linear perturbation of the objective function.

Theorem 3.3. *Let X, Y, Z be Banach spaces and let $\mathcal{U}$ be a subset of Z. Let $\phi_u : X\times Y \to \mathbb{R}\cup\{+\infty\}$ be a proper lower semicontinuous and convex function for any $u\in\mathcal{U}$. Suppose that $\inf\limits_{x\in X}\sup\limits_{u\in\mathcal{U}}\phi_u(x,0) < +\infty$. Then, the following statements are equivalent:*

(1) $\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*)$ *is w^*-closed and convex.*

(2) *For each $x^*\in X^*$*

$$\inf_{x\in X}\sup_{u\in\mathcal{U}}\{\phi_u(x,0)+\langle x^*,x\rangle\} = \max_{u\in\mathcal{U}}\max_{y^*\in Y^*}\{-\phi_u^*(-x^*,y^*)\}.$$

Proof. [(1) ⇒ (2)] Let $x^*\in X^*$, $\overline{\phi}(x,y,u)=\phi(x,y,u)+\langle x^*,x\rangle$ and $\overline{\phi}_u(x,y)=\phi(x,y,u)$. Then, we see that $\overline{\phi}_u$ is proper, lower semicontinuous and convex for any $u\in\mathcal{U}$, and

$$\overline{\phi}_u^*(0,y^*)=\phi_u^*(-x^*,y^*).$$

Note that the assumption $\inf_{x\in X}\sup_{u\in\mathcal{U}}\phi_u(x,0)<+\infty$ implies that

$$\inf_{x\in X}\sup_{u\in\mathcal{U}}\overline{\phi}_u(x,0)<+\infty.$$

So, the conclusion follows by applying Theorem 3.2 to $\overline{\phi}_u$.

[(2) ⇒ (1)] Suppose that (2) holds. Take $(x^*,r)\in\overline{\mathrm{co}\mathrm{Pr}_{X^*\times\mathbb{R}}\big(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*\big)}^{w^*}$. For each $x^*\in X^*$, define

$$\eta(x^*)=\inf_{y^*\in Y^*}(\sup_{u\in\mathcal{U}}\phi_u)^*(x^*,y^*).$$

From (3.5), we see that for each $x^*\in X^*$

$$\begin{aligned}\inf_{x\in X}\sup_{u\in\mathcal{U}}\{\phi_u(x,0)+\langle x^*,x\rangle\}\geq -\eta^{**}(-x^*) &\geq -\eta(-x^*)\\ &= -\inf_{y^*\in Y^*}(\sup_{u\in\mathcal{U}}\phi_u)^*(-x^*,y^*)\\ &= \sup_{y^*\in Y^*}\{-(\sup_{u\in\mathcal{U}}\phi_u)^*(-x^*,y^*)\}\\ &\geq \sup_{u\in\mathcal{U}}\sup_{y^*\in Y^*}\{-\phi_u^*(-x^*,y^*)\}.\end{aligned}$$

So, the statement (2) implies that for each $x^*\in X^*$

$$\eta(-x^*)=\eta^{**}(-x^*)=-\max_{u\in\mathcal{U}}\max_{y^*\in Y^*}\{-\phi_u^*(-x^*,y^*)\}, \tag{3.8}$$

and so, η is a proper weak* lower semicontinuous convex function. Note that

$$\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*)\subseteq\mathrm{epi}\eta.$$

This implies that

$$(x^*,r)\in\overline{\mathrm{co}\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*)}^{w^*}\subseteq\mathrm{epi}\eta.$$

So, we have $\eta(x^*)\leq r$. From (3.8), we have

$$\max_{u\in\mathcal{U}}\max_{y^*\in Y^*}\{-\phi_u^*(x^*,y^*)\}=-\eta(x^*)\geq -r.$$

So, there exist $u\in\mathcal{U}$ and $y^*\in Y^*$ such that $\phi_u^*(x^*,y^*)\leq r$, i.e.,

$$(x^*,y^*,r)\in\mathrm{epi}\phi_u^*.$$

So,

$$(x^*,r)\in\mathrm{Pr}_{X^*\times\mathbb{R}}\big(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*\big)$$

Thus,

$$\overline{\mathrm{co}\mathrm{Pr}_{X^*\times\mathbb{R}}\big(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*\big)}^{w^*}=\mathrm{Pr}_{X^*\times\mathbb{R}}\big(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*\big),$$

and hence the statement (1) follows. □

Let X,Y,Z be Banach spaces and let $\mathcal{U}$ be a subset of Z. Let $f:X\to\mathbb{R}\cup\{+\infty\}$ be a proper lower semicontinuous convex function and let $g_u:X\to Y$, $u\in\mathcal{U}$, be continuous S-convex. Define $\phi_u:X\times Y\to\mathbb{R}\cup\{+\infty\}$ by $\phi_u(x,y)=f(x)+\delta_{g_u^{-1}(y-S)}(x)$ for any $u\in\mathcal{U}$, where $g_u^{-1}(y-S)=\{a\in X:-g_u(a)+y\in S\}$. For parameters $(y,u)\in Y\times\mathcal{U}$, the parameterized uncertain problem becomes

$$\begin{array}{lll}(P_{y,u}) & \inf\limits_{x\in X} & f(x)\\ & \text{s.t.} & -g_u(x)+y\in S,\end{array}$$

where y is the pertubatation parameter and u is the uncertain parameter and $u\in\mathcal{U}$. When, $y=0$, it collapses to the uncertain convex optimization problem:

$$\begin{array}{lll}(P_{0,u}) & \inf\limits_{x\in X} & f(x)\\ & \text{s.t.} & -g_u(x)\in S,\end{array}$$

where S is a closed convex cone and $u\in\mathcal{U}$. The robust counterpart of (P_u) is

$$\begin{array}{lll}(RP) & \inf\limits_{x\in X} & f(x)\\ & \text{s.t.} & -g_u(x)\in S,\ \forall\,u\in\mathcal{U}.\end{array}$$

Moreover, noting that

$$\phi_u^*(x^*, y^*) = \begin{cases} \sup_{x\in X}\{\langle x^*, x\rangle - f(x) - \langle y^*, g_u(x)\rangle\} & \text{if} \quad y^* \in S^+ \\ +\infty & \text{else,} \end{cases}$$

the uncertain dual problem (D_u) is defined by $\max_{y^*\in S^+} \inf_{x\in X}\{f(x) + \langle y^*, g_u(x)\rangle\}$. Thus, the optimistic counterpart of the uncertain dual problem is

$$(OD) \quad \max_{u\in\mathcal{U}} \max_{y^*\in S^+} \inf_{x\in X}\{f(x) + \langle y^*, g_u(x)\rangle\}.$$

Corollary 3.1. *Let $f : X \to \mathbb{R}\cup\{+\infty\}$ be a proper lower semicontinuous convex function and let $g_u : X \to Y$, $u \in \mathcal{U}$, be a continuous S-convex function. Suppose that the robust feasible set: $F := \{x \in \mathrm{dom} f : g_u(x) \in -S, \ \forall\, u \in \mathcal{U}\} \neq \emptyset$ and*

$$\mathrm{epi} f^* + \bigcup_{u\in\mathcal{U}, y^*\in S^+} \mathrm{epi}(y^* \circ g_u)^* \text{ is } w^*\text{-closed and convex.} \tag{3.9}$$

Then, we have

$$\inf_{x\in X}\{f(x) : -g_u(x) \in S, \ \forall\, u \in \mathcal{U}\} = \max_{u\in\mathcal{U}} \max_{y^*\in S^+} \inf_{x\in X}\{f(x) + \langle y^*, g_u(x)\rangle\}.$$

Proof. The conclusion follows from Theorem 3.2. □

In the case where $\mathcal{U}$ is singleton, condition (3.9) collapses to the corresponding condition in [17, 19]; extension of (3.9) to difference of convex programming problems have also been given in [14, 15].

Robust conjugate duality vs standard conjugate duality

Consider again the robust counterpart (RP) of problem $(P_{0,u})$

$$\inf_{x\in X} \sup_{u\in\mathcal{U}} \phi_u(x, 0)$$

which can be rewritten as

$$\inf_{x\in X} \tilde{\phi}(x, 0)$$

where $\tilde{\phi}(x, y) := \sup_{u\in\mathcal{U}} \phi_u(x, y)$. Then, the standard conjugate dual of (RP) is

$$(DP) \quad \max_{y^*\in Y^*}\{-\tilde{\phi}^*(0, y^*)\}.$$

The structures of the problems (RP), (DP) and (ODP) show that

$$\inf_{x\in X} \tilde{\phi}(x, 0) \geq \max_{y^*\in Y^*}\{-\tilde{\phi}^*(0, y^*)\} \geq \max_{u\in\mathcal{U}} \max_{y^*\in Y^*}\{-\phi_u^*(0, y^*)\}.$$

It then follows that, if the primal worst value equals the dual best value with dual attainment (*i.e.*, $\inf_{x\in X} \tilde{\phi}(x,0) = \max_{u\in\mathcal{U}} \max_{y^*\in Y^*}\{-\phi_u^*(0,y^*)\}$), then strong duality between the robust counterpart (RP) and its standard conjugate dual (DP) holds.

We now present an example to show that the converse of this implication is not necessarily true in general. This illustrates that the robust strong duality that primal worst equals dual best with dual attainment is a **sharp strong duality** result.

Example 3.1. *Consider the following uncertain convex programming problem*

$$(P_1) \quad \min_{x\in\mathbb{R}^2} \quad \phi_u(x,0)$$

where u is uncertain and $u \in \mathcal{U} := [0,1]$, and $\phi_u : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is defined by, for each $x = (x_1, x_2) \in \mathbb{R}^2$ and $y \in \mathbb{R}$,

$$\phi_u(x,y) \quad = \quad \begin{cases} -\frac{1}{2}x_1 - x_2 & \text{if} \quad u^2|x_1| + \max\{x_2,0\} - 2u \le y, \\ +\infty & \text{else.} \end{cases}$$

Let $\tilde{\phi}(x,y) := \max\limits_{u\in\mathcal{U}} \phi_u(x,y)$. Then, for each $x = (x_1,x_2) \in \mathbb{R}^2$ and $y \in \mathbb{R}$

$$\begin{aligned} \tilde{\phi}(x,y) \quad &= \quad \begin{cases} -\frac{1}{2}x_1 - x_2 & \text{if} \quad u^2|x_1| + \max\{x_2,0\} - 2u \le y, \ \forall u \in [0,1] \\ +\infty & \text{else,} \end{cases} \\ &= \quad \begin{cases} -\frac{1}{2}x_1 - x_2 & \text{if} \quad \max\{|x_1| + \max\{x_2,0\} - 2, \max\{x_2,0\}\} \le y \\ +\infty & \text{else.} \end{cases} \end{aligned}$$

So, we see that the value of the robust counterpart of (P_1) is

$$\begin{aligned} \inf_{x\in\mathbb{R}^2} \sup_{u\in\mathcal{U}} \phi_u(x,0) \quad &= \quad \inf_{x\in\mathbb{R}^2} \tilde{\phi}(x,0) \\ &= \quad \inf_{(x_1,x_2)\in\mathbb{R}^2} \{-\frac{1}{2}x_1 - x_2 : |x_1| \le 2, x_2 \le 0\} = -1. \end{aligned}$$

The standard conjugate dual of the robust counterpart is

$$(DP_1) \quad \max_{\lambda_1\in\mathbb{R}} \{-\tilde{\phi}^*(0,\lambda_1)\}$$

where, for each $a = (a_1,a_2) \in \mathbb{R}$ and $\lambda_1 \in \mathbb{R}$,

$$\begin{aligned} &\tilde{\phi}^*(a,\lambda_1) \\ = \quad &\sup_{x_1,x_2,y} \{(a_1 + \frac{1}{2})x_1 + (a_2+1)x_2 + \lambda_1 y : \max\{|x_1| + \max\{x_2,0\} - 2, \max\{x_2,0\}\} \le y\} \\ = \quad &\begin{cases} \sup\limits_{x_1,x_2} \{(a_1 + \frac{1}{2})x_1 + (a_2+1)x_2 + \lambda_1(\max\{|x_1| + \max\{x_2,0\} - 2, \max\{x_2,0\}\})\} & \text{if} \quad \lambda_1 \le 0 \\ +\infty & \text{else,} \end{cases} \\ = \quad &\begin{cases} \sup\limits_{x_1,x_2} \{(a_1 + \frac{1}{2})x_1 + (a_2+1)x_2 - \lambda_1(\max\{|x_1| + \max\{x_2,0\} - 2, \max\{x_2,0\}\})\} & \text{if} \quad \lambda_1 \ge 0 \\ +\infty & \text{else.} \end{cases} \end{aligned}$$

Now,

$$\begin{aligned}
\max_{\lambda_1\in\mathbb{R}}\{-\tilde{\phi}^*(0,\lambda_1)\} &= \max_{\lambda_1\geq 0}\inf_{x\in\mathbb{R}^2}\{-\frac{1}{2}x_1 - x_2 + \lambda_1\max\{|x_1| + \max\{x_2,0\} - 2, \max\{x_2,0\}\}\,\} \\
&\geq \inf_{x\in\mathbb{R}^2}\{-\frac{1}{2}x_1 - x_2 + \max\{|x_1| + \max\{x_2,0\} - 2, \max\{x_2,0\}\}\,\} = -1.
\end{aligned}$$

This together with the weak duality shows that strong duality holds between the robust counterpart and its standard conjugate dual.

On the other hand, we see that the primal worst value is not equal to the dual best value. To see this, for each $a = (a_1, a_2) \in \mathbb{R}$ *and* $\lambda_1 \in \mathbb{R}$*,*

$$\begin{aligned}
&\phi_u^*(a,\lambda_1) \\
= &\sup_{x_1,x_2,y}\{(a_1 + \frac{1}{2})x_1 + (a_2+1)x_2 + \lambda_1 y : u^2|x_1| + \max\{x_2,0\} - 2u \leq y\} \\
= &\begin{cases} \sup_{x_1,x_2}\{(a_1+\frac{1}{2})x_1 + (a_2+1)x_2 - \lambda_1(u^2|x_1| + \max\{x_2,0\} - 2u)\} & \textit{if} \quad \lambda_1 \geq 0 \\ +\infty & \textit{else.} \end{cases}
\end{aligned}$$

Then, for each $u \in [0,1]$*,*

$$\begin{aligned}
-\phi_u^*(0,\lambda_1) &= \inf\{(-\frac{1}{2}x_1 - x_2) + \lambda_1(u^2|x_1| + \max\{x_2,0\} - 2u)\} \\
&= \begin{cases} -2\lambda_1 u & \textit{if} \quad \lambda_1 u^2 \geq \frac{1}{2},\ \lambda_1 \geq 1 \\ -\infty & \textit{else.} \end{cases}
\end{aligned}$$

Note that, for each $\lambda_1 \geq 1$ *and* $u \in [0,1]$ *with* $\lambda_1 u^2 \geq 1/2$, $-2\lambda_1 u = -2\sqrt{(\lambda_1 u^2)\lambda_1} \leq -\sqrt{2}$ *and the inequality is attained when* $\lambda_1 = 1$ *and* $u = \sqrt{2}/2$*. So, the dual best value is*

$$\max_{u\in[0,1]}\max_{\lambda_1\in\mathbb{R}}\{-\phi_u^*(0,\lambda_1)\} = -\sqrt{2}.$$

Thus, the primal worst value $\neq$ *the dual best value.*

4 Tractable Partially Finite Convex Programs under Uncertainty

In this Section, we show that the robust counterpart of a partially finite convex program under component-wise scenario uncertainty is computationally tractable by establishing that the robust conjugate duality always holds and the optimistic counterpart of its conjugate dual is a finite dimensional deterministic convex programming problem. We show that the condition (3.9) is always satisfied for partially finite convex program under component-wise scenario uncertainty. This is verified by showing that the set $\bigcup_{u\in\mathcal{U},y^*\in S^+}\mathrm{epi}(y^*\circ g_u)^*$ is a finitely generated cone for partially finite convex program under component-wise scenario uncertainty.

Let $Y=\mathbb{R}^m$, $S=\mathbb{R}^m_+$ and let $f: X \to \mathbb{R}$ be a convex function. Consider the following partially finite convex programming under constraint-wise scenario data uncertainty (SP):

$$\begin{array}{lll} (SP) & \inf\limits_{x\in X} & f(x) \\ & \text{s.t.} & \langle a_i, x\rangle \le \beta_i,\ i=1,\dots,m, \end{array}$$

where the data $(a_i,\beta_i)\in X^*\times\mathbb{R}$, $i=1,\dots,m$, is uncertain and (a_i,β_i) belongs to the scenario data uncertainty set $\overline{\mathcal{U}}_i \subseteq X^*\times\mathbb{R}$ defined by

$$\overline{\mathcal{U}}_i = \{(a_i^{(0)},\beta_i^{(0)}) + \sum_{l=1}^k w_i^{(l)}(a_i^{(l)},\beta_i^{(l)}) : (w_i^{(1)},\dots,w_i^{(k)}) \in Z_i\}$$

with $Z_i := \mathrm{co}\{v_{i1},\dots,v_{iq}\}$ for some $v_{i1},\dots,v_{iq}\in\mathbb{R}^k$.

The robust counterpart of (SP) is

$$\begin{array}{lll} (RSP) & \inf\limits_{x\in X} & f(x) \\ & \text{s.t.} & \langle a_i, x\rangle \le \beta_i,\ \ \forall (a_i,\beta_i)\in\overline{\mathcal{U}}_i,\ \ i=1,\dots,m. \end{array}$$

The optimistic counterpart of its conjugate dual is given by

$$\max_{(a_i,\beta_i)\in\overline{\mathcal{U}_i}} \max_{\lambda_i\ge 0} \inf_{x\in X}\{f(x) + \sum_{i=1}^m \lambda_i(\langle a_i,x\rangle - \beta_i)\}.$$

Note that this dual problem can be rewritten as

$$\begin{array}{ll} \max\limits_{(\mu_j,\lambda_i)\in\mathbb{R}\times\mathbb{R}} & -f^*\big(-\sum\limits_{i=1}^m \lambda_i a_i\big) - \sum\limits_{i=1}^m \lambda_i\beta_i, \\ & a_i = a_i^{(0)} + \sum\limits_{l=1}^k\sum\limits_{j=1}^q \mu_{ij} v_{ij}^{(l)} a_i^{(l)},\ i=1,\dots,m \\ & \beta_i = \beta_i^{(0)} + \sum\limits_{l=1}^k\sum\limits_{j=1}^q \mu_j v_{ij}^{(l)} \beta_i^{(l)},\ i=1,\dots,m \\ & \lambda_i \ge 0, 0\le \mu_{ij}\le 1,\ \sum\limits_{j=1}^q \mu_{ij} = 1, \end{array}$$

where $(a_i^{(l)},\beta_i^{(l)})\in X^*\times\mathbb{R}$, $i=1,\dots,m$, $l=0,1,\dots,k$, and $v_{ij}(v_{ij}^{(1)},\dots,v_{ij}^{(k)})\in\mathbb{R}^k$, $i=1,\dots,m$, $j=1,\dots,q$. By letting $\lambda_{i0}=\lambda_i$ and $\lambda_{ij}=\lambda_i\mu_{ij}$, this problem can be further simplified as

$$\begin{array}{ll} \max\limits_{\lambda_{ij}\in\mathbb{R}} & -f^*\big(-\sum\limits_{i=1}^m(\lambda_{i0}a_i^{(0)} + \sum\limits_{l=1}^k\sum\limits_{j=1}^q \lambda_{ij} v_{ij}^{(l)} a_i^{(l)})\big) - \sum\limits_{i=1}^m \big(\lambda_{i0}\beta_i^{(0)} + \sum\limits_{l=1}^k\sum\limits_{j=1}^q \lambda_{ij} v_{ij}^{(l)}\beta_i^{(l)}\big) \\ & \lambda_{ij}\ge 0,\ i=1,\dots,m,\ j=0,1,\dots,q, \end{array}$$

which is a **deterministic finite dimensional** convex programming problem.

Theorem 4.1. *For the problem (SP), Suppose that the robust feasible set*

$$F := \{x \in X : \langle a_i, x\rangle \le \beta_i,\, \forall\, (a_i, \beta_i) \in \overline{\mathcal{U}_i},\, i = 1, \dots, m\} \neq \emptyset.$$

Then, robust strong duality holds:

$$\begin{aligned}
&\inf_{x\in X}\{f(x) : \langle a_i, x\rangle \le \beta_i,\, \forall\, (a_i, \beta_i) \in \overline{\mathcal{U}_i},\, i = 1, \dots, m\}\\
=\ &\max_{\lambda_{ij}\in\mathbb{R}}\{-f^*\big(-\sum_{i=1}^m(\lambda_{i0}a_i^{(0)} + \sum_{l=1}^k\sum_{j=1}^q \lambda_{ij}v_{ij}^{(l)}a_i^{(l)})\big) - \sum_{i=1}^m\big(\lambda_{i0}\beta_i^{(0)} + \sum_{l=1}^k\sum_{j=1}^q\lambda_{ij}v_{ij}^{(l)}\beta_i^{(l)}\big) :\\
&\qquad \lambda_{ij} \ge 0,\ i = 1, \dots, m,\ j = 0, 1, \dots, q\}.
\end{aligned}$$

Proof. Fix $i \in \{1, \dots, m\}$. Let $u_i = (a_i, \beta_i) \in X^* \times \mathbb{R}$ and let $g_i : X \times X^* \times \mathbb{R} \to \mathbb{R}$ by

$$g_i(x, u_i) = \langle a_i, x\rangle - \beta_i.$$

Let $\mathcal{U} = \prod_{i=1}^m \overline{\mathcal{U}_i} \subseteq \prod_{i=1}^m (X^* \times \mathbb{R})$ and define $g_u : X \to \mathbb{R}^m$, $u = (u_1, \dots, u_m) \in \mathcal{U}$ by

$$g_u(x) = (g_1(x, u_1), \dots, g_m(x, u_m))^T.$$

Clearly, $\mathcal{U}$ is convex and compact, and for each $u \in \mathcal{U}$, g_u is a continuous $\mathbb{R}^m_+$-convex function on X. Thus, the conclusion will follow from Corollary 3.1 if we show that

$$K := \mathrm{epi} f^* + \bigcup_{u_i\in\overline{\mathcal{U}}_i, \lambda_i\ge 0} \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, u_i))^* \text{ is } w^*\text{-closed and convex.}$$

[**Convexity of** K]. We first show that $\bigcup_{u_i\in\overline{\mathcal{U}}_i,\lambda_i\ge 0}\mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, u_i))^*$ is convex. To see this, we only need to show

$$\mathrm{co}\bigcup_{u_i\in\overline{\mathcal{U}}_i,\lambda_i\ge 0} \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, u_i))^* \subseteq \bigcup_{u_i\in\overline{\mathcal{U}}_i,\lambda_i\ge 0}\mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, u_i))^*.$$

Now, let $(x^*, r) \in \mathrm{co}\bigcup_{u_i\in\overline{\mathcal{U}}_i,\lambda_i\ge 0}\mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, u_i))^*$. As $\bigcup_{u_i\in\overline{\mathcal{U}}_i,\lambda_i\ge 0}\mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, u_i))^*$ is a cone, there exist $s \in \mathbb{N}$ and $(x_l^*, r_l) \in \bigcup_{u_i\in\overline{\mathcal{U}}_i,\lambda_i\ge 0}\mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, u_i))^*$, $l = 1, \dots, s$, such that $(x^*, r) = \sum_{l=1}^s (x_l^*, r_l)$. Since $(x_l^*, r_l) \in \bigcup_{u_i\in\overline{\mathcal{U}}_i,\lambda_i\ge 0}\mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, u_i))^*$, $l = 1, \dots, s$, there exist $u_{il} \in \overline{\mathcal{U}}_i$ and $\lambda_{il} \ge 0$, such that $(x_l^*, r_l) \in \mathrm{epi}(\sum_{i=1}^m \lambda_{il} g_{il}(\cdot, u_{il}))^*$, and so,

$$(x^*, r) = \sum_{l=1}^s (x_l^*, r_l) \in \sum_{l=1}^s \mathrm{epi}(\sum_{i=1}^m \lambda_{il} g_i(\cdot, u_{il}))^* = \sum_{i=1}^m \mathrm{epi}\big(\sum_{l=1}^s \lambda_{il} g_i(\cdot, u_{il})\big)^*. \tag{4.10}$$

Let $\lambda_i = \sum_{l=1}^s \lambda_{il} \ge 0$. We now show that there exist $u_i \in \overline{\mathcal{U}}$ such that, for any $x \in X$,

$$\sum_{l=1}^s \lambda_{il} g_i(x, u_{il}) = \lambda_i g_i(x, u_i). \tag{4.11}$$

If $\lambda_i = 0$, then $\lambda_{il} = 0$ for all $l = 1, \ldots, s$. So, (4.11) follows. If $\lambda_i > 0$, then let $u_i = \sum_{l=1}^{s} \frac{\lambda_{il} u_{il}}{\lambda_i}$. As $g_i(x, \cdot)$ is linear, it follows that

$$\sum_{l=1}^{s} \lambda_{il} g_i(x, u_{il}) = \lambda_i g_i(x, u_i).$$

Thus, we have $\sum_{i=1}^{m} \mathrm{epi}\big(\sum_{l=1}^{s} \lambda_{il} g_i(\cdot, u_{il})\big)^* = \sum_{i=1}^{m} \mathrm{epi}(\lambda_i g_i(\cdot, u_i))^*$, and so,

$$(x^*, r) \in \sum_{i=1}^{m} \mathrm{epi}(\lambda_i g_i(\cdot, u_i))^* \subseteq \bigcup_{u_i \in \overline{\mathcal{U}}_i, \lambda_i \geq 0} \mathrm{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, u_i))^*.$$

Therefore, $\bigcup_{u_i \in \overline{\mathcal{U}}_i, \lambda_i \geq 0} \mathrm{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, u_i))^*$ is convex, and hence K is also convex.

[**Closedness of** K]. Let $v_{ij} = (v_{ij}^{(1)}, \ldots, v_{ij}^{(k)})$, $j = 1, \ldots, q$. Note that

$$\begin{aligned} \overline{\mathcal{U}}_i &= \{(a_i^{(0)}, \beta_i^{(0)}) + \sum_{l=1}^{k} w_i^{(l)} (a_i^{(l)}, \beta_i^{(l)}) : (w_i^{(1)}, \ldots, w_i^{(k)}) \in Z_i\} \\ &= \mathrm{co}\{(a_i^{(0)}, \beta_i^{(0)}) + \sum_{l=1}^{k} v_{i1}^{(l)} (a_i^{(l)}, \beta_i^{(l)}), \ldots, (a_i^{(0)}, \beta_i^{(0)}) + \sum_{l=1}^{k} v_{iq}^{(l)} (a_i^{(l)}, \beta_i^{(l)})\} \\ &= \mathrm{co}\{(z_{i1}, \gamma_{i1}), \ldots, (z_{iq}, \gamma_{iq})\} \qquad (4.12) \end{aligned}$$

where

$$(z_{ij}, \gamma_{ij}) = (a_i^{(0)}, \beta_i^{(0)}) + \sum_{l=1}^{k} v_{ij}^{(l)} (a_i^{(l)}, \beta_i^{(l)}),\ j = 1, \ldots, q,$$

and

$$\begin{aligned} \bigcup_{u_i \in \overline{\mathcal{U}}_i, \lambda_i \geq 0} \mathrm{epi}(\sum_{i=1}^{m} \lambda_i g_i)^* &= \bigcup_{(a_i, \beta_i) \in \overline{\mathcal{U}}_i, \lambda_i \geq 0} \{\sum_{i=1}^{m} \lambda_i a_i\} \times [\sum_{i=1}^{m} \lambda_i \beta_i, +\infty) \\ &= \bigcup_{(a_i, \beta_i) \in \overline{\mathcal{U}}_i, \lambda_i \geq 0} \sum_{i=1}^{m} \lambda_i \{(a_i, \beta_i)\} + \{0\} \times [0, +\infty). \qquad (4.13) \end{aligned}$$

Letting $\mathcal{V}_i := \{(z_{i1}, \gamma_{i1}), \ldots, (z_{iq}, \gamma_{iq})\}$, we see that $\mathrm{co}\mathcal{V}_i = \overline{\mathcal{U}_i}$. From (4.12) and (4.13), we obtain that

$$\bigcup_{u_i \in \overline{\mathcal{U}}_i, \lambda_i \geq 0} \mathrm{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, u_i))^* = \bigcup_{(a_i, \beta_i) \in \mathrm{co}\mathcal{V}_i, \lambda_i \geq 0} \sum_{i=1}^{m} \lambda_i \{(a_i, \beta_i)\} + \{0\} \times [0, +\infty).$$

As $\{0\} \times [0, +\infty)$ and

$$\bigcup_{(a_i, \beta_i) \in \mathrm{co}\mathcal{V}_i, \lambda_i \geq 0} \sum_{i=1}^{m} \lambda_i \{(a_i, \beta_i)\} = \bigcup_{(a_i, \beta_i) \in \mathcal{V}_i, \lambda_{ip} \geq 0} \sum_{i=1}^{m} \sum_{p=1}^{q} \lambda_{ip} \{(a_{ip}, \beta_{ip})\}$$

are both finitely generated, their sum $\bigcup_{u_i\in\mathcal{U}_i,\lambda_i\geq 0}\mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot,u_i))^*$ is w^*-closed. So, we see that $\bigcup_{u_i\in\mathcal{U}_i,\lambda_i\geq 0}\mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot,u_i))^*$ is w^*-closed and convex which implies that

$$\bigcup_{u_i\in\mathcal{U}_i,\lambda_i\geq 0}\mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot,u_i))^* = \overline{\mathrm{co}\bigcup_{u_i\in\mathcal{U}_i,\lambda_i\geq 0}\mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot,u_i))^*}^{w^*} = \mathrm{epi}(\delta_F)^*,$$

where the second equality follows from $\delta_F = \sup_{u_i\in\mathcal{U}_i,\lambda_i\geq 0}\sum_{i=1}^m \lambda_i g_i(\cdot,u_i)$ and Lemma 2.1. Therefore, the continuity of f gives us that $K = \mathrm{epi} f^* + \mathrm{epi}(\delta_F)^*$ is w^*-closed. $\square$

5 Robust Fenchel's Duality

In this Section we derive a generalization of the celebrated Fenchel's duality theorem for convex functions under data uncertainty.

Theorem 5.1. (Robust Fenchel Duality Theorem) *Let X, Y, Z be Banach spaces and let $\mathcal{U}_1,\mathcal{U}_2$ be subsets of Z. Let $f_{u_1}: X\to\mathbb{R}\cup\{+\infty\}$, $u_1\in\mathcal{U}_1$ and $g_{u_2}: Y\to\mathbb{R}\cup\{+\infty\}$, $u_2\in\mathcal{U}_2$ be proper lower semicontinuous and convex functions. Let $A: X\to Y$ be a continuous linear mapping. Suppose that for each $(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2$,*

$$A(\mathrm{dom} f_{u_1})\cap \mathrm{dom} g_{u_2}\neq\emptyset. \tag{5.14}$$

Then, the following statements are equivalent:

(1) $\bigcup_{(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2}(\mathrm{epi} f_{u_1}^* + (A^*\times I)\mathrm{epi} g_{u_2}^*)$ *is w^*-closed and convex.*

(2) *For each $x^*\in X^*$*

$$\begin{aligned}&\inf_{x\in X}\sup_{(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2}\{f_{u_1}(x)+g_{u_2}(Ax)+\langle x^*,x\rangle\}\\ =&\max_{(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2}\max_{y^*\in Y^*}\{-f_{u_1}^*(-x^*-A^*y^*)-g_{u_2}^*(y^*)\}.\end{aligned}$$

Proof. Let $\mathcal{U}=\mathcal{U}_1\times\mathcal{U}_2$ and define $\phi_u: X\times Y\to\mathbb{R}\cup\{+\infty\}$, $u\in\mathcal{U}$ by

$$\phi_u(x,y) = f_{u_1}(x)+g_{u_2}(Ax+y)\quad \forall\, x\in X, y\in Y \text{ and } u=(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2.$$

Then, for each $u=(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2$, we have

$$\phi_u(x,0) = f_{u_1}(x)+g_{u_2}(Ax)$$

and

$$\begin{aligned}\phi_u^*(x^*,y^*) &= \sup_{x,y}\{\langle (x^*,y^*),(x,y)\rangle - \phi_u(x,y)\}\\ &= \sup_{x,y}\{\big(\langle x^*,x\rangle - f_{u_1}(x)\big)+\big(\langle y^*,y\rangle - g_{u_2}(Ax+y)\big)\}\\ &= f_{u_1}^*(x^*-A^*y^*)+g_{u_2}^*(y^*).\end{aligned}$$

Note that

$$\begin{aligned}
\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi_u^*) &= \{(x^*,r): \exists\,(y^*,u_1,u_2,s)\in Y^*\times\mathcal{U}_1\times\mathcal{U}_2\times\mathbb{R} \text{ such that} \\
&\quad (x^*-A^*y^*,s)\in \mathrm{epi}f_{u_1}^* \text{ and } (y^*,r-s)\in\mathrm{epi}g_{u_2}^*\} \\
&= \bigcup_{(u_1,u_2)\in\mathcal{U}} (\mathrm{epi}f_{u_1}^* + (A^*\times I)\mathrm{epi}g_{u_2}^*).
\end{aligned}$$

So, the conclusion follows from Theorem 3.3. □

6 Robust Duality for Data Classification Problems

In this Section, we examine robust strong duality for optimization problems that arise in machine learning problems of data classification. The support vector machine classification [3, 28] is to find a hyperplane $\{v : x^T v = \alpha\}$ that separates two classes of data sets.

Let $V \in \mathbb{R}^{n\times m}$ be the matrix of data points (totally m points in $\mathbb{R}^n$) and $w \in \{-1,1\}^m$ be the corresponding labelled vector, so that $w_i = 1$ when the ith data point belongs to the first class, and $w_i = -1$ when the ith data point belongs to the second class. Then, the support vector machine classification can be modeled as the following convex optimization problem:

$$\begin{array}{lll}
(SVM) & \min_{(x,\alpha,\xi)\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m_+} & \psi(\xi)+\phi(x) \\
& \text{s.t.} & w_i(x^T v_i-\alpha)-1+\xi_i \geq 0,
\end{array}$$

where ψ is a real-valued convex loss function and ϕ is a real-valued convex regularization function. Usually, the loss function measures the error in the separation and the regularization function is to avoid the over-fitting problem. Typical loss functions [33] include

$$\begin{cases}
\text{Hinge Loss function} & \psi_1(\xi)=\sum_{i=1}^m \xi_i \\
\text{quadratic loss function} & \psi_2(\xi)=\|\xi\|_2^2 \\
\text{log loss function} & \psi_3(\xi)=\sum_{i=1}^m \log(1+\exp(\xi_i))
\end{cases}$$

and typical regularization functions [33] include

$$\begin{cases}
L^1 \text{ penalty} & \phi_1(x)=\lambda\|x\|_1 \\
L^2 \text{ penalty} & \phi_2(x)=\lambda\|x\|_2 \\
\text{mixed penalty} & \phi_3(x)=\lambda_1\|x\|_1+\lambda_2\|x\|_2.
\end{cases}$$

When $\psi=\psi_3$ and $\phi=\phi_3$, this model has been recently successfully used in [16]. Moreover, the conjugates of loss functions and regularization functions can be found in [33].

In many real-world classification problems the data are subject to measurement or prediction errors. So, we suppose that each data v_i is uncertain and it belongs to the uncertain set $\mathcal{V}_i$. Then, the robust counterpart of our SVM problem is

$$\begin{array}{lll}(RSVM) & \min_{(x,\alpha,\xi)\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m_+} & \psi(\xi)+\phi(x) \\ & \text{s.t.} & w_i(x^T v_i-\alpha)-1+\xi_i\geq 0,\ \forall\, v_i\in\mathcal{V}_i.\end{array}$$

Let $C_u=\{(x,\alpha,\xi)\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m_+ : w_i(v_i^T x_i-\alpha)-1+\xi_i\geq 0,\ i=1,\ldots,m\}$ where $u=(v_1,\ldots,v_m)$. Let $\mathcal{U}_1=\{0\}$ and $\mathcal{U}_2=\prod_{i=1}^m \mathcal{V}_i$. Define $f_{u_1}:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$, $u_1\in\mathcal{U}_1$ and $g_{u_2}:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$, $u_2\in\mathcal{U}_1$ by $f_{u_1}(x,\alpha,\xi)=\psi(\xi)+\phi(x)$ and $g_{u_2}(x,\alpha,\xi,u)=\delta_{C_{u_2}}(x,\alpha,\xi)$. Then, we see that (RSVM) can be equivalently rewritten as

$$\inf_{(x,\alpha,\xi)\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m}\ \sup_{(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2}\{f_{u_1}(x,\alpha,\xi)+g_{u_2}(x,\alpha,\xi)\}.$$

By the robust Fenchel dual problem with $Y=X=\mathbb{R}^{n+m+1}$ and $A=I$ becomes

$$\begin{aligned}&\max_{(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2}\ \max_{(x^*,r,\zeta^*)\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m}\{-f^*_{u_1}(-x^*,-r,-\zeta^*)-g^*_{u_2}(x^*,r,\zeta^*)\}.\\ =\ &\max_{u\in\prod_{i=1}^m\mathcal{V}_i}\ \max_{(x^*,r,\zeta^*)\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m}\{-\psi^*(-\zeta^*)-\phi^*(-x^*)-\delta^*_{C_u}(x^*,r,\zeta^*)\}.\end{aligned}$$

Then, we have the following duality theorem:

Theorem 6.1. *Let $\mathcal{V}_i$ be the scenario uncertainty set given by $\mathcal{V}_i=\{v_i^{(0)}+\sum_{l=1}^k w_i^{(l)}v_i^{(l)} : (w_i^{(1)},\ldots,w_i^{(k)})\in Z_i\}$. Then,*

$$\begin{aligned}&\min_{(x,\alpha,\xi)\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m_+}\{\psi(\xi)+\phi(x): w_i(x^T v_i-\alpha)-1+\xi_i\geq 0,\ \forall\, v_i\in\mathcal{V}_i\}\\ =\ &\max_{u\in\prod_{i=1}^m\mathcal{V}_i}\ \max_{(x^*,r,\zeta^*)\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m}\{-\psi^*(-\zeta^*)-\phi^*(-x^*)-\delta^*_{C_u}(x^*,r,\zeta^*)\}\\ =\ &-\min_{\substack{(x^*,r,\zeta^*)\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m\\ u\in\prod_{i=1}^m\mathcal{V}_i}}\{\psi^*(-\zeta^*)+\phi^*(-x^*)+\delta^*_{C_u}(x^*,r,\zeta^*)\}.\end{aligned}$$

Proof. Let $\mathcal{U}_1=\{0\}$ and $\mathcal{U}_2=\prod_{i=1}^m\mathcal{V}_i$. Define $f_{u_1}:\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m\to\mathbb{R}$, $u_1\in\mathcal{U}_1$, and $g_{u_2}:\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m\to\mathbb{R}$, $u_2\in\mathcal{U}_2$, by $f_{u_1}(x,\alpha,\xi)=\psi(\xi)+\phi(x)$ and

$$g_{u_2}(x,\alpha,\xi)=\delta_{C_{u_2}}(x,\alpha,\xi).$$

Note that f_{u_1} is real valued for any $u_1\in\mathcal{U}_1$, and so, $0\in\operatorname{int}\bigcap_{(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2}\big(\mathrm{dom}g_{u_2}-\mathrm{dom}f_{u_1}\big)$ always holds. Applying Proposition 6.2 with

$$\phi_u(x,y)=f_{u_1}(x)+g_{u_2}(x+y)\quad\forall\, x\in X, y\in Y \text{ and } u=(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2,$$

this gives us that $\bigcup_{(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2}(\mathrm{epi}f^*_{u_1}+\mathrm{epi}g^*_{u_2})$ is closed. Moreover, let $h_i(x,\alpha,\xi,v_i)=-w_i(v_i^T x_i-\alpha)-\xi_i+1$, $i=1,\ldots,m$ and $h_i(x,\alpha,\xi)=-\xi_{i-m}$, $i=m+1,\ldots,2m$. Let $u_2=(v_1,\ldots,v_m)\in\mathcal{U}_2$ and define $h_{u_2}:\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m\to\mathbb{R}^{2m}$, $u_2\in\mathcal{U}_2$, by

$$h_{u_2}(x,\alpha,\xi)=(h_1(x,\alpha,\xi,v_i),\ldots,h_m(x,\alpha,\xi,v_i),h_{m+1}(x,\alpha,\xi),\ldots,h_{2m}(x,\alpha,\xi))^T.$$

Then $C_{u_2} = \{(x,\alpha,\xi) : h_{u_2}(x,\alpha,\xi) \le 0\}$, $u_2 \in \mathcal{U}_2$. As in the proof in Theorem 4.1 (see (4.13)), we have

$$\begin{aligned}\bigcup_{u_2\in\mathcal{U}_2,\lambda\in\mathbb{R}^{2m}_+} \mathrm{epi}(\lambda\circ h_{u_2})^* &= \Big(\bigcup_{v_i\in\mathcal{V}_i,\lambda_i\ge 0}\sum_{i=1}^m \lambda_i(-w_iv_i, w_i, -1, -1) + \sum_{i=m+1}^{2m}\lambda_i(0,0,-1,0)\Big)\\ &\quad +\{0\}\times\{0\}\times\{0\}\times[0,+\infty)\end{aligned}$$

which is a finitely generated cone and so is closed and convex. Note from Lemma 2.1 and $\delta_{C_{u_2}} = \sup_{\lambda\in\mathbb{R}^{2m}_+}(\lambda\circ h_{u_2})$, $u_2 \in \mathcal{U}_2$, that

$$\begin{aligned}\bigcup_{u_2\in\mathcal{U}_2,\lambda\in\mathbb{R}^{2m}_+} \mathrm{epi}(\lambda\circ h_{u_2})^* \subseteq \bigcup_{u_2\in\mathcal{U}_2}\mathrm{epi}\delta^*_{C_{u_2}} &= \bigcup_{u_2\in\mathcal{U}_2}\overline{\mathrm{co}\bigcup_{\lambda\in\mathbb{R}^{2m}_+}\mathrm{epi}(\lambda\circ h_{u_2})^*}\\ &= \bigcup_{u_2\in\mathcal{U}_2}\overline{\bigcup_{\lambda\in\mathbb{R}^{2m}_+}\mathrm{epi}(\lambda\circ h_{u_2})^*}\\ &\subseteq \overline{\bigcup_{u_2\in\mathcal{U}_2,\lambda\in\mathbb{R}^{2m}_+}\mathrm{epi}(\lambda\circ h_{u_2})^*}.\end{aligned}$$

So, $\bigcup_{u_2\in\mathcal{U}_2}\mathrm{epi}g^*_{u_2} = \bigcup_{u_2\in\mathcal{U}_2}\mathrm{epi}\delta^*_{C_{u_2}} = \bigcup_{u_2\in\mathcal{U}_2,\lambda\in\mathbb{R}^{2m}_+}\mathrm{epi}(\lambda\circ h_{u_2})^*$ is a convex cone. Hence, $\bigcup_{(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2}(\mathrm{epi}f^*_{u_1} + \mathrm{epi}g^*_{u_2})$ is convex and the conclusion follows from Theorem 5.1 with $x^* = 0$. □

Appendix: Additional Regularity Conditions

In this Section, we present conditions under which $\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi^*_u)$ is convex or w^*-closed. They provide additional regularity conditions ensuring robust conjugate duality.

Proposition 6.1. *Let X, Y, Z be Banach spaces and let $\mathcal{U}$ be a convex subset of Z. Let $\phi_u : X\times Y \to \mathbb{R}\cup\{+\infty\}$ be a proper lower semicontinuous and convex function for any $u \in \mathcal{U}$, and $u \mapsto \phi_u(x,y)$ is a concave function for any $(x,y) \in X\times Y$. Then, $\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi^*_u)$ is convex.*

Proof. Noting the fact that the projection mapping is a linear mapping, it suffices to show $\bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi^*_u$ is a convex set. Let $\lambda \in [0,1]$. Let $(x^{1*}, y^{1*}, r^1) \in \bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi^*_u$ and $(x^{2*}, y^{2*}, r^2) \in \bigcup_{u\in\mathcal{U}}\mathrm{epi}\phi^*_u$. Then, there exist $u_1 \in \mathcal{U}$ and $u_2 \in \mathcal{U}$ such that

$$(x^{1*}, y^{1*}, r^1) \in \mathrm{epi}\phi^*_{u_1} \text{ and } (x^{2*}, y^{2*}, r^2) \in \mathrm{epi}\phi^*_{u_2}.$$

So, for each $x \in X$ and $y \in Y$ we have

$$\langle (x^{1*}, y^{1*}), (x,y)\rangle - \phi_{u_1}(x,y) \le r^1 \text{ and } \langle (x^{2*}, y^{2*}), (x,y)\rangle - \phi_{u_2}(x,y) \le r^2.$$

As $u \mapsto \phi_u(x,y)$ is concave for any $(x,y) \in X \times Y$, this implies that

$$\begin{aligned} & \langle (\lambda x^{1*} + (1-\lambda)x^{2*}, \lambda y^{1*} + (1-\lambda)y^{2*}), (x,y)\rangle - \phi_{\lambda u_1 + (1-\lambda)u_2}(x,y) \\ \leq\ & \langle (\lambda x^{1*} + (1-\lambda)x^{2*}, \lambda y^{1*} + (1-\lambda)y^{2*}), (x,y)\rangle - (\lambda \phi_{u_1}(x,y) + (1-\lambda)\phi_{u_2}(x,y)) \\ \leq\ & \lambda r^1 + (1-\lambda)r^2. \end{aligned}$$

So, we see that

$$\lambda(x^{1*}, y^{1*}, r^1) + (1-\lambda)(x^{2*}, y^{2*}, r^2) \in \mathrm{epi}\phi^*_{\lambda u_1 + (1-\lambda)u_2} \subseteq \bigcup_{u \in \mathcal{U}} \mathrm{epi}\phi_u^*.$$

Hence, we see that $\bigcup_{u\in\mathcal{U}} \mathrm{epi}\phi_u^*$ is a convex set. $\square$

Proposition 6.2. *Let X, Y, Z be Banach spaces and let $\mathcal{U}$ be a convex compact subset of Z. Let $\phi_u : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous and convex function for any $u \in \mathcal{U}$, and $u \mapsto \phi_u(x,y)$ is a upper semicontinuous function for any $(x,y) \in X \times Y$. Suppose that*

$$0 \in \mathrm{int}\big(\bigcap_{u\in\mathcal{U}} \mathrm{Pr}_Y(\mathrm{dom}\phi_u)\big).$$

Then, $\mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}} \mathrm{epi}\phi_u^*)$ *is w^*-closed.*

Proof. Let $(x_t^*, r_t) \in \mathrm{Pr}_{X^*\times\mathbb{R}}(\bigcup_{u\in\mathcal{U}} \mathrm{epi}\phi_u^*)$ be a convergent net such that

$$(x_t^*, r_t) \to^{w^*} (x^*, r).$$

Then, there exist $y_t^* \in Y^*$ and $u_t \in \mathcal{U}$ such that $(x_t^*, y_t^*, r_t) \in \mathrm{epi}\phi_{u_t}^*$. So, for each $(x,y) \in X \times Y$,

$$\langle (x_t^*, y_t^*), (x,y)\rangle - \phi_{u_t}(x,y) \leq r_t. \tag{6.15}$$

As $\mathcal{U}$ is compact, by passing to a subnet if necessary, we may assume that $u_t \to u \in \mathcal{U}$. We now show that $\{y_t^*\}$ is bounded. To see this, as $0 \in \mathrm{int}\big(\bigcap_{u\in\mathcal{U}} \mathrm{Pr}_Y(\mathrm{dom}\phi_u)\big)$, there exists $\delta > 0$ such that

$$B(0,\delta) \subseteq \mathrm{Pr}_Y(\mathrm{dom}\phi_u)\ \forall\, u \in \mathcal{U}.$$

Then, for each $y \in Y$ with $\|y\| \leq \delta$, there exists $x \in X$ such that $\phi_u(x,y) < +\infty$ for all $u \in \mathcal{U}$, and so, (6.15) implies that for each t

$$\langle y_t^*, y\rangle \leq -\langle x_t^*, x\rangle + \phi_{u_t}(x,y) + r_t.$$

This implies that for each $y \in Y$ with $\|y\| \leq \delta$,

$$\sup_t \langle y_t^*, y\rangle \leq \sup_t \{-\langle x_t^*, x\rangle + r_t + \phi_{u_t}(x,y)\}.$$

Note that $(x_t^*, r_t) \to^{w^*} (x^*, r)$, $u_t \to u$ and $u \mapsto \phi_u$ is upper semicontinuous. So, for each $y \in Y$ with $\|y\| \leq \delta$, $\sup_t \langle y_t^*, y\rangle < +\infty$. This implies that for each $y \in Y$,

$$\sup_t |\langle y_t^*, y\rangle| < +\infty.$$

As Y is a Banach space, by the uniform boundedness principle, we see that

$$\sup_t \|y_t^*\| < +\infty.$$

Thus, $\{y_t^*\}$ is a bounded net. By passing to subnet if necessary, we may assume that $y_t^* \to^{w^*} y^*$. As for each $(x,y) \in X \times Y$,

$$\langle (x_t^*, y_t^*), (x,y) \rangle - \phi_{u_t}(x,y) \le r_t,$$

passing to upper limit, we have for each $(x,y) \in X \times Y$,

$$\langle (x^*, y^*), (x,y) \rangle - \phi_u(x,y) \le r,$$

which means that $\phi_u^*(x^*, y^*) \le r$. So,

$$(x^*, r) \in \mathrm{Pr}_{X^* \times \mathbb{R}}(\bigcup_{u \in \mathcal{U}} \mathrm{epi}\phi_u^*)$$

and hence $\mathrm{Pr}_{X^* \times \mathbb{R}}(\bigcup_{u \in \mathcal{U}} \mathrm{epi}\phi_u^*)$ is w^*-closed. □

Theorem 6.2. *Let X, Y, Z be Banach spaces and let $\mathcal{U}$ be a convex compact subset of Z. Let $\phi_u : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous and convex function for any $u \in \mathcal{U}$, and $u \mapsto \phi_u(x,y)$ is a upper semicontinuous concave function for any $(x,y) \in X \times Y$. Suppose that $0 \in \mathrm{int}\big(\bigcap_{u \in \mathcal{U}} \mathrm{Pr}_Y(\mathrm{dom}\phi_u)\big)$. Then,*

$$\inf_{x \in X} \sup_{u \in \mathcal{U}} \phi_u(x, 0) = \max_{u \in \mathcal{U}} \max_{y^* \in Y^*} \{-\phi_u^*(0, y^*)\}.$$

Proof. Proposition 6.1 and Proposition 6.2 imply that $\mathrm{Pr}_{X^* \times \mathbb{R}}(\bigcup_{u \in \mathcal{U}} \mathrm{epi}\phi_u^*)$ is convex and w^*-closed. Thus, the conclusion follows from Theorem 3.2. □

Remark 6.1. *Consider the special case where $\mathcal{U}$ is a singleton and $\phi_u = \phi$. In this case, the condition that $u \mapsto \phi_u(x,y)$ is an upper semicontinuous concave function for any $(x,y) \in X \times Y$ is automatically satisfied. Moreover, the condition, $0 \in \mathrm{int}\big(\bigcap_{u \in \mathcal{U}} \mathrm{Pr}_Y(\mathrm{dom}\phi_u)\big)$, collapses to the usual interior point condition, $0 \in \mathrm{int}\,\big(\mathrm{Pr}_Y(\mathrm{dom}\phi)\big)$. Thus, the preceding theorem shows that the usual interior point condition $0 \in \mathrm{int}\,\big(\mathrm{Pr}_Y(\mathrm{dom}\phi)\big)$ guarantees the standard conjugate duality. More general interior point conditions ensuring the standard conjugate duality result can be found in [8, 9, 10]. These general interior point conditions may also be extended to study robust conjugate duality. This will be examined in a forthcoming study.*

References

[1] A. Beck and A. Ben-Tal, Duality in robust optimization: primal worst equals dual best, *Oper. Res. Lett.*, **37** (2009), 1-6.

[2] A. Ben-Tal and A. Nemirovski, Robust optimization—methodology and applications, *Math. Program.*, Ser B, 92 (2002), 453-480.

[3] A. Ben-Tal, L.E. Ghaoui and A. Nemirovski, *Robust Optimization*, Princeton Series in Applied Mathematics, 2009.

[4] A. Ben-Tal and A. Nemirovski, A selected topics in robust convex optimization, *Math. Progr. B*, **112** (2008), 125-158.

[5] D. Bertsimas and D. Brown, Constructing uncertainty sets for robust linear optimization, *Oper. Res.*, 57 (2009), 1483-1495.

[6] D. Bertsimas, D. Pachamanova and M. Sim, Robust linear optimization under general norms, *Oper. Res. Lett.*, 32 (2004), 510-516.

[7] J.M. Borwein and A.S. Lewis, Partially finite convex programming. I. Quasi relative interiors and duality theory, *Math. Program.* 57(1992), Ser. B, 15-48.

[8] R.I. Bot, *Conjugate Duality in Convex Optimization*, Lecture Notes in Economics and Mathematical Systems, 637. Springer-Verlag, Berlin, 2010.

[9] R.I. Bot, E.R. Csetnek and G. Wanka, Regularity conditions via quasi-relative interior in convex programming. *SIAM J. Optim.* 19 (2008), 217-233.

[10] R.I. Bot, S.M. Grad and G. Wanka, New constraint qualification and conjugate duality for composed convex optimization problems, *J. Optim. Theory Appl.* 135 (2007), 241-255.

[11] R. S. Burachik, V. Jeyakumar and Z. Y. Wu, Necessary and sufficient conditions for stable conjugate duality, *J. Nonlinear Anal.*, Ser. A, 64 (2006), 1998-2006.

[12] R. S. Burachik and V. Jeyakumar, A new geometric condition for Fenchel duality in infinite dimensions, *Math. Program.*, Ser. B, 104(2005), 229-233.

[13] L.A. Cysiqued, J. M. Murray, M. Dunbar, V. Jeyakumar and B. J. Brew, A screening algorithm for HIV-associated neurocognitive disorders, *HIV Medicine* (2010), DOI: 10.1111/j.1468-1293.2010.00834.x

[14] N. Dinh, B.S. Mordukhovich and T.T.A. Nghia, Qualification and optimality conditions for DC programs with infinite constraints. *Acta Math. Vietnam.* 34 (2009), 125-155.

[15] N. Dinh, B.S. Mordukhovich and T.T.A. Nghia, Subdifferentials of value functions and optimality conditions for DC and bilevel infinite and semi-infinite programs. *Math. Program.*, Ser. B, 123 (2010), 101-138.

[16] M. Dunbar, J.M. Murray, L.A. Cysiqued, B.J. Brew and V. Jeyakumar, Simultaneous classification and feature selection via convex quadratic programming with application to HIV-associated neurocognitive disorder assessment, *Eur. Jour. Oper. Res.* 206 (2010), No. 2, 470-478.

[17] V. Jeyakumar, Constraint qualifications characterizing Lagrangian duality in convex optimization, *J. Optim. Theory Appl.*, 136(2008), 31-41.

[18] V. Jeyakumar, The conical hull intersection property for convex programming, *Math. Program.*, 106 (2006), 81-92.

[19] V. Jeyakumar and G. M. Lee, Complete characterizations of stable Farkas' lemma and cone-convex programming duality, *Math. Program.* 114 (2008), 335–347.

[20] V. Jeyakumar, G. M. Lee and N. Dinh, New sequential Lagrange multiplier conditions characterizing optimality without constraint qualification for convex programs, *SIAM J. Optim.* 14 (2003), 534–547.

[21] V. Jeyakumar and G. Li, Characterizing robust set containments and solutions of uncertain linear programs without qualifications, *Oper. Res. Lett.*, 38(2010), 188-194.

[22] V. Jeyakumar and G. Li, Strong duality in robust convex programming: Complete characterizations, *SIAM J. Optim*, (2010), DOI: 10.1137/100791841.

[23] V. Jeyakumar, J. Ormerod, and R.S. Womersley, Knowledge-based semidefinite linear programming classifiers. *Optim. Methods Softw.* 21 (2006), no. 5, 693–706.

[24] G. Li, and K.F. Ng, On extension of Fenchel duality and its application, *SIAM J. Optim.*, 19(2008), 1489-1509.

[25] G. Li and V. Jeyakumar, Qualification-free optimality conditions for convex programs with separable inequality constraints. *J. Convex Anal.* 16 (2009), 845-856.

[26] C. Li, D.H. Fang, G. López and M.A. López, Stable and total Fenchel duality for convex optimization problems in locally convex spaces, *SIAM J. Optim.* 20 (2009), 1032–1051.

[27] O. L. Mangasarian and David. R. Musicant Lagrangian Support Vector Machines, *Journal of Machine Learning Research*, 1(2001), 161-177.

[28] E.E. Osuna, R. Freund, and F. Girosi, Support vector machines: Training and applications. AI Memo 1602, AI Lab, Massachusetts Institute of Technology, March 1997.

[29] T. Precupanu, Closedness conditions for the optimality of a family of non-convex optimization problems, *Math. Operationsforsch. Statist. Ser. Optim.*, 15 (1984), pp. 339–346.

[30] J. Shaw-Taylor and N. Cristianini, Support Vector Machines and other kernel-based learning methods, Cambridge University Press, 2000

[31] A. Shapiro, Stochastic programming approach to optimization under uncertainty, *Math. Program.*, Ser B, 112 (2008), 183-220.

[32] A. Shapiro, D. Dentcheva and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, Philadelphia, 2009.

[33] R.M. Rifkin and R.A. Lippert, Value regularization and Fenchel duality. *J. Mach. Learn. Res.*, 8 (2007), 441–479.

[34] V.N. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, second edition, 2000.

[35] C. Zalinescu, *Convex Analysis in General Vector Spaces*, World Scientific, 2002.