Lagrange Multiplier Characterizations of Robust Best Approximations under Constraint Data Uncertainty*

V. Jeyakumar[†], J. H. Wang[‡] and G. Li [§]

Second Revised Version: March 25, 2012

Abstract

In this paper we explain how to characterize the best approximation to any x in a Hilbert space X from the set $C \cap \{x \in X : g_i(x) \leq 0, i = 1, 2, \cdots, m\}$ in the face of data uncertainty in the convex constraints, $g_i(x) \leq 0, i = 1, 2, \cdots, m$, where C is a closed convex subset of X. Following the robust optimization approach, we establish Lagrange multiplier characterizations of the robust constrained best approximation that is immunized against data uncertainty. This is done by characterizing the best approximation to any x from the robust counterpart of the constraints where the constraints are satisfied for all possible uncertainties within the prescribed uncertainty sets. Unlike the traditional Lagrange multiplier characterizations without data uncertainty, for constrained best approximation problems in the face uncertainty, we show that the strong conical hull intersection property (strong CHIP) alone is not sufficient to guarantee the Lagrange multiplier characterizations. We present conditions which guarantee that the strong CHIP is necessary and sufficient for the multiplier characterization. We also establish that the strong CHIP is automatically satisfied for the cases of polyhedral constraints with polytope uncertainty, and linear constraints with interval uncertainty. As an application, we show how robust solutions of shape preserving interpolation problems under ellipsoidal and box uncertainty cases can be obtained in terms of Lagrange multipliers under strict robust feasibility conditions.

Key words. Robust optimization, best approximation, strong conical hull intersection property, ellipsoidal uncertainty, shape-preserving interpolation

---

*Research of the first and the third authors was partially supported by a grant from the Australian Research Council. The work of the second author was carried out while she was at the University of New South Wales, Sydney, Australia. The authors are grateful to the referees and the editor for their constructive comments and helpful suggestions which have contributed to the final preparation of the paper.

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

[‡]Department of Mathematics, Zhejiang University of Technology, Hangzhou 310032, P. R. China. E-mail: wjh@zjut.edu.cn

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

1 Introduction

Studies of determining the best approximation [5, 7, 8, 17] to any x in a Hilbert space X from the set $C \cap \{x \in X : g_i(x) \le 0, i = 1, 2, \cdots, m\}$ commonly assume accurate values for the data or parameters in the constraints $g_i(x) \le 0,\ i = 1, \cdots, m$, where $g_i : X \to \mathbb{R}, i = 1, 2 \cdots, m$, are continuous convex functions [6, 7, 8, 9, 14, 18]. However, such precise information is rarely available in practice because of estimation errors or lack of complete information. An effective approach to dealing with the data uncertainty is to treat uncertainty as deterministic and describes it in terms of bounded sets. This approach in optimization is known as robust optimization [1, 2, 11] and is a complementary approach to stochastic optimization [20] which describes uncertainty in terms of probability distributions.

Following the framework of robust optimization [1], the constrained best approximation problems in the face of constraint data uncertainty can be captured by examining the best approximation from the set $C \cap \{x \in X : g_i(x, v_i) \le 0, i = 1, 2, \cdots, m\}$, where $g_i(\cdot, v_i)$ is convex and v_i is the uncertain parameter which belongs to an uncertainty set $\mathcal{V}_i \subseteq \mathbb{R}^{n_i}$. For instance, the constrained interpolation problems in $L^2[0, 1]$ with uncertain linear inequality constraints can be examined within this framework where $g_i(x) = \langle a_i, x\rangle - \beta_i$, $C := \{x \in L^2[0, 1] : \langle h_j, x\rangle = d_j,\ j = 1, \dots, r\}$ and the data (a_i, β_i), $i = 1, \dots, m$, are uncertain and $(h_j, d_j) \in L^2[0, 1] \times \mathbb{R}$ where $\langle x_1, x_2\rangle$ is the inner product defined in $L^2[0, 1]$. In the case of ellipsoidal uncertainty, $g_i(x, v_i) := \langle a_i(v_i), x\rangle - \beta_i(v_i)$, $a_i(v_i) := a_i^{(0)} + \sum_{l=1}^k v_i^{(l)} a_i^{(l)}$, $\beta_i(v_i) := \beta_i^{(0)} + \sum_{l=1}^k v_i^{(l)} \beta_i^{(l)}$, $\mathcal{V}_i := \{(v_i^{(1)}, \dots, v_i^{(k)}) \in \mathbb{R}^k : \|(v_i^{(1)}, \dots, v_i^{(k)})\|_{\mathbb{R}^k} \le 1\}$ is an ellipsoidal uncertainty set and $(a_i^{(l)}, \beta_i^{(l)}) \in L^2[0, 1] \times \mathbb{R}$. For recent work on robust convex optimization duality, see [10, 11].

In this paper we study the problem of characterizing the *robust best approximation* to any x from the set $C \cap \{x \in X : g_i(x) \le 0, i = 1, 2, \cdots, m\}$ that is immunized against constraint data uncertainty. This is done by examining the best approximation to any x from the robust counterpart of the convex inequality constraints

$$K := C \cap \{x \in X : g_i(x, v_i) \le 0, \forall\, v_i \in \mathcal{V}_i, i = 1, 2, \cdots, m\},$$

where the uncertain constraints are enforced for every possible value of the data within the prescribed uncertainty sets.

It is known that if $\mathcal{V}_i$ is a singleton (i.e. there is no data uncertainty), for each $i = 1, 2, \dots, m$, and if each $g_i(\cdot, v_i)$ is a polyhedral function (i.e maximum of finitely many affine functions), then the strong conical hull intersection property (strong CHIP) [5, 9, 14, 15] guarantees that x_0 is the best approximation to x from K (i.e. $x_0 = P_K(x)$) if and only if

$$x_0 = P_C(x - \sum_{i=1}^m \lambda_i l_i), \text{ for some } \lambda_i \ge 0,\ l_i \in \partial g_i(\cdot, v_i))(x_0),\ v_i \in \mathcal{V}_i,\ \sum_{i=1}^m \lambda_i g_i(x_0, v_i) = 0. \quad (1.1)$$

However, this Lagrange multiplier characterization may fail (see Example 3.1) while the strong CHIP holds for problems in the face of data uncertainty where $\mathcal{V}_i$ is not a singleton and $g_i(x, \cdot)$ is not concave.

The purpose of this work is to establish the above Lagrange multiplier characterization for robust best approximation under additional conditions, and provide classes of functions and uncertainty sets for which the strong CHIP alone is sufficient for the characterization. We also present characterizations in the cases of polyhedral constraints with polytope uncertainty, and linear constraints with interval uncertainty where the strong CHIP is automatically satisfied. As an application, we show how the robust shape preserving interpolation under ellipsoidal and box uncertainty cases can be characterized in terms of Lagrange multipliers under strict robust feasibility condition.

The outline of the paper is as follows. Section 2 presents preliminary results on conjugate functions, strong CHIP and best approximations. Section 3 describes characterizations for the robust best approximation using the strong CHIP and illustrates these characterizations for special classes of constraints and uncertainty sets. Section 4 provides characterizations for the robust shape-preserving interpolation problems under uncertainty. Section 5 concludes with a discussion on further research.

2 Preliminaries

We begin this section by fixing the notation and definitions that will be used later in the paper. Let X and Y be Banach spaces. For $C \subseteq X$, the closure of C and the interior of C are denoted by $\mathrm{cl}(C)$ and $\mathrm{int}(C)$ respectively. The quasi-relative interior of C ([21]) is denoted by $\mathrm{qri}(C)$ and is defined by

$$\mathrm{qri}(C) = \{x \in C : \mathrm{cl}(\bigcup_{\lambda \geq 0} \lambda(C - x)) \text{ is a linear subspace } \}.$$

A set C is called a polyhedral set if there exist $m \in \mathbb{N}$, $a_1, \ldots, a_m \in X^*$ and $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ such that $C = \{x : a_i(x) \leq \alpha_i, i = 1, \ldots, m\}$. Moreover, C is called a polytope if it is a compact polyhedral.

The continuous dual space of X will be denoted by X^*. For a set $W \subset X^*$, the weak*-closure of W will be denoted by $w^*\text{-cl}\, W$. For a subset D of X, the indicator function δ_D is defined by $\delta_D(x) = 0$ if $x \in D$ and $\delta_D(x) = +\infty$ if $x \notin D$. The support function σ_D is defined by $\sigma_D(u) = \sup_{x \in D} u(x) \quad (u \in X^*)$. The epigraph of f, epi$f$, is defined by

$$\mathrm{epi} f = \{(x, r) \in X \times \mathbb{R} : x \in \mathrm{dom} f, f(x) \leq r\},$$

where the effective domain of f, dom$f$, is given by $\mathrm{dom} f = \{x \in X : f(x) < +\infty\}$.

Let $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a proper lower semi-continuous convex function. Then the conjugate function of f is defined by $f^* : X^* \longrightarrow \mathbb{R} \cup \{+\infty\}$ such that $f^*(u) = \sup_{x \in \mathrm{dom} f}\{u(x) - f(x)\} \quad (u \in X^*)$. Clearly, f^* is a proper lower semicontinuous convex function and $\lambda \mathrm{epi} f^* = \mathrm{epi}(\lambda f)^*$ for any $\lambda > 0$.

Lemma 2.1. *(cf. [19]) 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*

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

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

For a subset W of X, define the polar cone of W by

$$W^\circ := \{f \in X^* : f(w) \le 0 \quad \forall\, w \in W\},$$

and the dual cone of W by $W^+ := -W^\circ$. For a non-empty subset W of X, the convex hull (resp. conical hull) of W denoted by co$W$ (resp. cone$W$) which is the intersection of all convex sets (resp. convex cones) containing W. The nonnegative orthant of $\mathbb{R}^n$ is denoted by $\mathbb{R}^n_+$ and is defined by $\mathbb{R}^n_+ := \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_i \ge 0\}$.

For a function $f : X \longrightarrow \mathrm{I\!R}$, the subdifferential of f at $x \in X$, denoted by $\partial f(x)$, is defined by

$$\partial f(x) := \{u \in X^* : f(y) \ge f(x) + u(y - x) \quad \forall\, y \in X\}.$$

It is well known that $\partial f(x) \ne \emptyset$ for all $x \in X$ if f is a continuous convex function.

For a non-empty subset W of X and $x \in X$, we define $d(x, W) := \inf_{w \in W} \|x - w\|$. A point $w_0 \in W$ is called a best approximation to $x \in X$ (i.e. $w_0 \in P_W(x)$), if $d(x, W) = \|x - w_0\|$. If for each $x \in X$ there exists a unique best approximation $w_0 \in W$, then W is called a Chebyshev subset of X. Recall (see [5]) that every closed convex set in a Hilbert space is Chebyshev.

The following characterization of best approximation in Hilbert spaces is well known (see [5]).

Lemma 2.2. *Let W be a closed convex subset of a Hilbert space X, $x \in X$, and $w_0 \in W$. Then $w_0 = P_W(x)$ if and only if $x - w_0 \in (W - w_0)^\circ$.*

Definition 2.1. (Strong CHIP)*[5] Let C_1 and C_2 be closed convex sets in X and let $x \in C_1 \cap C_2$. Then, the pair $\{C_1,\ C_2\}$ is said to have the strong CHIP at x, if*

$$(C_1 \cap C_2 - x)^\circ = (C_1 - x)^\circ + (C_2 - x)^\circ.$$

The pair $\{C_1,\ C_2\}$ is said to have the strong CHIP if it has the strong CHIP at each $x \in C_1 \cap C_2$.

Lemma 2.3. *[14] Let C_1 and C_2 be closed convex subsets of X such that* $\text{epi}\sigma_{C_1} + \text{epi}\sigma_{C_2}$ *is w^*-closed. Then the pair $\{C_1,\ C_2\}$ has the strong CHIP.*

In particular, if one of the following two conditions is satisfied: (1) $\text{int}(C_1) \cap C_2 \ne \emptyset$ (2) C_1 is a polyhedral and $C_1 \cap \text{qri}(C_2) \ne \emptyset$, then $\text{epi}\sigma_{C_1} + \text{epi}\sigma_{C_2}$ is w*-closed (and hence the pair $\{C_1,\ C_2\}$ has the strong CHIP). For related conditions for strong CHIP, see [3, 4, 5, 14].

3 Robust Best Approximations under Uncertainty

In this Section, we provide conditions characterizing the robust best approximation x_0 to x from K in terms of the best approximation to a perturbation $(x - \sum_{i=1}^m \lambda_i l_i)$ of x from the set C for some multipliers $\lambda_i \ge 0$, $l_i \in \partial g_i(\cdot, v_i)(x_0)$, $v_i \in \mathcal{V}_i$ with $\sum_{i=1}^m \lambda_i g_i(x_0, v_i) = 0$.

We first note that if

$$x_0 = P_C(x - \sum_{i=1}^m \lambda_i l_i), \text{ for some } \lambda_i \ge 0,\ l_i \in \partial g_i(\cdot, v_i)(x_0),\ v_i \in \mathcal{V}_i \text{ with } \sum_{i=1}^m \lambda_i g_i(x_0, v_i) = 0$$

then it follows easily from Lemma 2.1 and the definitions of the subdifferential and the polar cone that $x_0 = P_K(x)$. On the other hand, it is known that if $\mathcal{V}_i$ is a singleton (i.e. there is no data uncertainty), for each $i = 1, 2, \ldots, m$, and if each $g_i(\cdot, v_i)$ is a polyhedral function (i.e maximum of finitely many affine functions), then the strong CHIP guarantees that $x_0 = P_K(x)$ if and only if

$$x_0 = P_C(x - \sum_{i=1}^{m} \lambda_i l_i), \text{ for some } \lambda_i \geq 0,\ l_i \in \partial g_i(\cdot, v_i))(x_0),\ v_i \in \mathcal{V}_i \text{ with } \sum_{i=1}^{m} \lambda_i g_i(x_0, v_i) = 0.$$

For details see, e.g. [9] and other references therein.

The following example illustrates that the above Lagrange multiplier characterization may fail in the face of uncertainty where $\mathcal{V}_i$ is not a singleton and the strong CHIP holds.

Example 3.1. (Failure of Multiplier Characterization under Uncertainty) Let $\mathcal{V}_1 = [0, 1]$ and let $g_1 : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ be defined by

$$g_1(x, v_1) = v_1^2 |x_1| + \max\{x_2, 0\} - 2v_1.$$

Clearly, $g_1(\cdot, v_1)$ is a polyhedral function for each $v_1 \in \mathbb{R}$. Consider $C = \mathbb{R}^2$, $D = \{x : g_1(x, v_1) \leq 0, \forall v_1 \in \mathcal{V}_1\}$. It can be verified that

$$K = C \cap D = D = \{(x_1, x_2) : |x_1| \leq 2 \text{ and } x_2 \leq 0\}$$

and $x_0 = (2, 0) \in K$. Clearly, $\{C, D\}$ has strong CHIP at x_0. We now show that the Lagrange multiplier characterization fails. To see this, we first observe that

$$\begin{aligned} M(x_0) &:= \bigcup_{\substack{\lambda_1 \geq 0 \\ v_1 \in [0,1]}} \{\partial(\lambda_1 g_1(\cdot, v_1))(x_0) : \lambda_1 g_1(x_0, v_1) = 0\} \\ &= \bigcup_{\substack{\lambda_1 \geq 0 \\ v_1 \in [0,1]}} \lambda_1(\{v_1^2\} \times [0, 1]) : \lambda_1(2v_1^2 - 2v_1) = 0\} \\ &= \{(x_1, x_2) : x_1 \geq x_2 \geq 0\} \cup (\{0\} \times \mathbb{R}_+). \end{aligned}$$

Consider $x = (3, 2)$. Then, $P_K(x) = x_0$ and $u := x - x_0 = (1, 2) \notin M(x_0)$. Thus, the Lagrange multiplier characterization fails.

We now show that the strong CHIP of $\{C, D\}$ is necessary and sufficient for the Lagrange multiplier characterization, under a regularity condition that, the *characteristic cone*, $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \mathrm{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, v_i))^*$, is closed and convex.

Theorem 3.1. (Multiplier Characterization). *Let X be a Hilbert space. Let C, D be closed convex sets in X where $D = \{x \in X : g_i(x, v_i) \leq 0, \forall v_i \in \mathcal{V}_i, i = 1, \cdots, m\}$. Let $x_0 \in K := C \cap D$. Suppose that $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \mathrm{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, v_i))^*$ is closed and convex. Then the following assertions are equivalent.*

(i) *$\{C, D\}$ has the strong CHIP at x_0.*

(ii) *For any $x \in X$, $x_0 = P_K(x)$ if and only if*

$$x_0 = P_C(x - \sum_{i=1}^{m} \lambda_i l_i), \text{ for some } \lambda_i \geq 0,\ l_i \in \partial g_i(\cdot, v_i)(x_0),\ v_i \in \mathcal{V}_i \text{ with } \sum_{i=1}^{m} \lambda_i g_i(x_0, v_i) = 0.$$

Proof. [(ii) $\Longrightarrow$ (i)]. Since $(C-x_0)^\circ + (D-x_0)^\circ \subset (K-x_0)^\circ$, it is sufficient to show that

$$(K-x_0)^\circ \subset (C-x_0)^\circ + (D-x_0)^\circ. \tag{3.2}$$

To see this, let $u \in (K-x_0)^\circ$. Then, by Lemma 2.2, we have $x_0 = P_K(u+x_0)$. So, by our assumption, there exists $l \in \sum_{i=1}^m \lambda_i \partial g_i(\cdot, v_i)(x_0)$ for some $\lambda_i \geq 0, v_i \in \mathcal{V}_i$ with $\sum_{i=1}^m \lambda_i g_i(x_0, v_i) = 0$ such that $x_0 = P_C(u + x_0 - l)$, which implies that $u - l \in (C-x_0)^\circ$ because of Lemma 2.2. Now, for each $y \in X$,

$$l(y-x_0) \leq \sum_{i=1}^m \lambda_i g_i(y, v_i) - \sum_{i=1}^m \lambda_i g_i(x_0, v_i) = \sum_{i=1}^m \lambda_i g_i(y, v_i).$$

This gives us that, for each $y \in D$, $l(y-x_0) \leq 0$. So, $l \in (D-x_0)^\circ$. Consequently, we get that $u = u - l + l \in (C-x_0)^\circ + (D-x_0)^\circ$. Therefore (i) holds.

[(i)$\Longrightarrow$(ii)]. Let

$$M(x_0) := \bigcup_{\substack{\lambda \in \mathbb{R}^m_+ \\ v \in \Pi_{i=1}^m \mathcal{V}_i}} \{\partial(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))(x_0) : \sum_{i=1}^m \lambda_i g_i(x_0, v_i) = 0\}.$$

We first show that $M(x_0) = (D-x_0)^\circ$. Let $u \in M(x_0)$. Then, there exist $\lambda = (\lambda_1, \cdots, \lambda_m) \in \mathbb{R}^m_+$ and $v = (v_1, \cdots, v_m)$ such that $\sum_{i=1}^m \lambda_i g_i(x_0, v_i) = 0$ and $u \in \partial(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))(x_0)$. So, for each $y \in X$,

$$u(y-x_0) \leq \sum_{i=1}^m \lambda_i g_i(y, v_i) - \sum_{i=1}^m \lambda_i g_i(x_0, v_i) = \sum_{i=1}^m \lambda_i g_i(y, v_i).$$

This gives us that, for each $y \in D$, $u(y-x_0) \leq 0$. So, $u \in (D-x_0)^\circ$ and hence $M(x_0) \subseteq (D-x_0)^\circ$.

Conversely, let $u \in (D-x_0)^\circ$. Note that $\delta_D(\cdot) = \sup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \sum_{i=1}^m \lambda_i g_i(\cdot, v_i)$ and $\delta_D^* = \sigma_D$. It follows from Lemma 2.1 that

$$\mathrm{epi}\sigma_D = w^*\text{-cl}(\mathrm{co} \bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))^*). \tag{3.3}$$

Note that $u \in (D-x_0)^\circ$ if and only if $(u, u(x_0)) \in \mathrm{epi}\sigma_D(u)$ which is, in turn equivalent to

$$(u, u(x_0)) \in w^*\text{-cl}(\mathrm{co} \bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))^*) = \bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))^*)$$

where the second equality follows by our assumption and fact that, in a Hilbert space, a convex set is closed if and only if it is weakly convex. Then, there exist $\lambda = (\lambda_1, \cdots, \lambda_m) \in \mathbb{R}^m_+$ and $v = (v_1, \cdots, v_m) \in \Pi_{i=1}^m \mathcal{V}_i$ such that $(u, u(x_0)) \in \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))^*$. Thus,

$$\sup_{y \in X}\{u(y) - \sum_{i=1}^m \lambda_i g_i(y, v_i)\} \leq u(x_0), \tag{3.4}$$

which gives us that $\sum_{i=1}^m \lambda_i g_i(x_0, v_i) = 0$ because $x_0 \in D$. Moreover, (3.4) yields that for each $y \in X$, $u(y - x_0) \leq \sum_{i=1}^m \lambda_i g_i(y, v_i)$. This shows that $u \in \partial(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))(x_0)$ and so $u \in M(x_0)$.

Suppose that (i) holds. It is sufficient show that $x_0 = P_K(x)$ implies that

$$x_0 = P_C(x - \sum_{i=1}^m \lambda_i \partial g_i(\cdot, v_i)(x_0)) \text{ for some } \lambda_i \geq 0, v_i \in \mathcal{V}_i \text{ with } \sum_{i=1}^m \lambda_i g_i(x_0, v_i) = 0$$

as the converse implication always holds. To see this, let $x_0 = P_K(x)$. Applying Lemma 2.2, we obtain that

$$x - x_0 \in (K - x_0)^\circ. \tag{3.5}$$

On the other hand, it follows from (i) and $M(x_0) = (D - x_0)^\circ$ that

$$(K - x_0)^\circ = (C - x_0)^\circ + (D - x_0)^\circ \quad = \quad (C - x_0)^\circ + M(x_0).$$

This together with (3.5) implies that there exists $u \in M(x_0)$ such that $x - u - x_0 \in (C - x_0)^\circ$. Applying Lemma 2.2 again, we obtain that $x_0 \in P_C(x - u)$. □

Remark 3.1. (A Comparison with the Uncertainty-free Case) *In the case where there is no uncertainty, i.e., each $\mathcal{V}_i$ is a singleton, $i = 1, \ldots, m$, the characteristic cone*

$$\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))^*$$

is always convex. So, our result collapses to the known result that the strong CHIP and the multiplier characterization (1.1), which is also known as the perturbation property, are equivalent, under a closed cone condition (e.g. see [9]). However, the characteristic cone

$$\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))^*$$

may not be a convex cone for approximation problems in the face of uncertainty. Indeed, as in Example 3.1, let $\mathcal{V}_1 = [0, 1]$ and $g_1 : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ be defined by $g_1(x, v_1) = v_1^2 |x_1| + \max\{x_2, 0\} - 2v_1$. Routine calculation gives us that So,

$$\begin{aligned} K := \bigcup_{v_1 \in \mathcal{V}_1, \lambda_1 \geq 0} \mathrm{epi}\big(\lambda_1 g_1(\cdot, v_1)\big)^* &= \bigcup_{v_1 \in [0,1], \lambda_1 \geq 0} [-\lambda_1 v_1^2, \lambda_1 v_1^2] \times [0, \lambda_1] \times [2\lambda_1 v_1, +\infty) \\ &= \bigcup_{s \geq r \geq 0} [-r, r] \times [0, s] \times [2\sqrt{rs}, +\infty). \end{aligned}$$

which is a closed cone. However, as the muliplier condition fails (as shown in Example 3.1), the preceding theorem gives us that K cannot be convex. In fact, letting $a = (0, 1, 0)$, $b = (1, 1, 2)$ and $c := \frac{a+b}{2} = (0.5, 1, 1)$. Direct verification gives that $a, b \in K$ while $c \notin K$.

Remark 3.2. (Conditions for Closed Convex Characteristic Cone). *Note that if, for each $i = 1, \ldots, m$, $\mathcal{V}_i$ is convex and compact, $g_i(x, \cdot)$ is concave and if the robust Slater condition holds, i.e. $g_i(x^*, v_i) < 0$, $i = 1, \ldots, m$, for each $v_i \in \mathcal{V}_i$ for some $x^* \in X$, then $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))^*$ is w^*-closed and convex. This follows from the fact that (see Propositions 2.3 and 3.2) [10]) the set $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))^*$ is convex whenever $\mathcal{V}_i$ is convex and compact and $g(x, \cdot)$ is concave and that $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))^*$ is closed by the robust Slater condition. For details see [10].*

Let us now examine special classes of constraints and uncertainty sets for which the main characterizations are simplified and strengthened. We see that the strong CHIP alone characterizes the perturbation property for constrained best approximation problems with polyhedral constraints with polytope uncertainty. In particular, in the case of linear constraints with interval uncertainty, the perturbation property is explicitly given simply in terms of of the original data.

Corollary 3.1. (Polyhedral Constraints with Polytope Uncertainty). *Let X be a Hilbert space. Let C, D be closed convex sets in X where $D = \{x \in X : \; g_i(x, v_i) \leq 0, \forall\, v_i \in \mathcal{V}_i, \; i = 1, \cdots, m\}$. Let $x_0 \in K := C \cap D$. Suppose that for each $i = 1, \ldots, m$, $\mathcal{V}_i$ is a polytope, $g_i(\cdot, v_i)$ is polyhedral and $g_i(x, \cdot)$ is affine. Then the following assertions are equivalent.*

(i) *$\{C, D\}$ has the strong CHIP at x_0.*

(ii) *For any $x \in X$, $x_0 = P_K(x)$ if and only if*

$$x_0 = P_C(x - \sum_{i=1}^m \lambda_i l_i), \text{ for some } \lambda_i \geq 0, \; l_i \in \partial g_i(\cdot, v_i))(x_0), \; v_i \in \mathcal{V}_i \text{ with } \sum_{i=1}^m \lambda_i g_i(x_0, v_i) = 0.$$

Proof. Let $M = \bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))^*$. Similar to the proof of Theorem 4.1 [10], one can show that M is closed if $g(\cdot, v_i)$ is polyhedral and $\mathcal{V}_i$ is a polytope, and M is convex if $g(x, \cdot)$ is affine. The conclusion then follows from Theorem 3.1. ☐

Corollary 3.2. *Let X be a Hilbert space. Let C be a polyhedral set and let $D = \{x \in X | \; g_i(x, v_i) \leq 0, \forall\, v_i \in \mathcal{V}_i, \; i = 1, \cdots, m\}$. Let $x_0 \in K := C \cap D$. Suppose that for each $i = 1, \ldots, m$, $\mathcal{V}_i$ is a polytope, $g_i(\cdot, v_i)$ is polyhedral and $g_i(x, \cdot)$ is affine. Then, for any $x \in X$, $x_0 = P_K(x)$ if and only if*

$$x_0 = P_C(x - \sum_{i=1}^m \lambda_i l_i), \text{ for some } \lambda_i \geq 0, \; l_i \in \partial g_i(\cdot, v_i))(x_0), \; v_i \in \mathcal{V}_i \text{ with } \sum_{i=1}^m \lambda_i g_i(x_0, v_i) = 0.$$

Proof. The conclusion will follow from the preceding corollary if we show that $\{C, D\}$ has the strong CHIP. To see this, we first observe that

$$D = \{x \in X | \max_{v_i \in \mathcal{V}_i} g_i(x, v_i) \leq 0, \; i = 1, \cdots, m\}.$$

As $\mathcal{V}_i$ is a polytope, we can write $\mathcal{V}_i = \mathrm{co}\{v_i^1, \ldots, v_i^{s_i}\}$ where v_i^j are extreme points of $\mathcal{V}_i$, $j = 1, \ldots, r_i$ and $s_i \in \mathbb{N}$. As $g_i(\cdot, v_i)$ is polyhedral, for each $x \in X$, $\max_{v_i \in \mathcal{V}_i} g_i(x, v_i)$ attains its maximum on some extreme points of $\mathcal{V}_i$, and so, $\max_{v_i \in \mathcal{V}_i} g_i(x, v_i) = \max_{1 \leq j \leq s_i} g_i(x, v_i^j)$ for each $x \in X$. Therefore, $D = \{x \in X | g_i(x, v_i^j) \leq 0, i = 1, \ldots, m, j = 1, \ldots, r_i\}$ is also a polyhedral set. Therefore, $\{C, D\}$ has the strong CHIP and the conclusion follows. ☐

Corollary 3.3. (Linear Constraints with Interval Uncertainty). *Let* $D := \{x \in \mathbb{R}^n |\ (v_i^j)^T x + b_i^j \leq 0,\ \forall v_i^j \in \mathcal{V}_{ij}, i = 1, \ldots, m, j = 1, \cdots, p,\}$ *and let* $\mathcal{V}_{ij} = [\underline{v}_i^j, \overline{v}_i^j] \subseteq \mathbb{R}^n$, $i = 1, \ldots, m$ *and* $j = 1, \ldots, p$. *Let* $x_0 \in K$. *Then the following assertions are equivalent.*

(i) $\{C, D\}$ *has the strong CHIP at* x_0.

(ii) *For any* $x \in \mathbb{R}^n$, $x_0 = P_K(x)$ *if and only if* $x_0 = P_C(x - \sum_{i,j} \lambda_{ij} v_i^j)$ *for some* $v_i^j \in \mathcal{V}_{ij}$, $\lambda_{ij} \geq 0$ *such that* $\lambda_{ij}[(v_i^j)^T x_0 - b_i^j] = 0$, $i = 1, \cdots, m,\ j = 1, \cdots, p$.

Proof. Define $g_{ij} : \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}$, $i = 1, \ldots, m$ and $j = 1, \ldots, p$, by $g_{ij}(x, v_{ij}) = (v_i^j)^T x + b_i^j$. Then, each $g_{ij}(\cdot, v_{ij})$ is polyhedral and each $g_{ij}(x, \cdot)$ is affine. Moreover,

$$\begin{aligned} &\bigcup_{\substack{\lambda \in \mathbb{R}^m_+ \\ v \in \Pi_{i=1}^m \mathcal{V}_i}} \{\partial(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))(x_0) : \sum_{i=1}^m \lambda_i g_i(x_0, v_i) = 0\} \\ = &\bigcup_{\substack{\lambda_{ij} \geq 0 \\ v_i^j \in \mathcal{V}_{ij}}} \{\sum_{i,j} \lambda_{ij} v_i^j :\ \sum_{i,j} \lambda_{ij}((v_i^j)^T x + b_i^j) = 0\}. \end{aligned}$$

Thus, the conclusion follows from Corollary 3.1. □

Corollary 3.4. *Let* $C = \{x \in \mathbb{R}^n | a^T x = c\}$ *with* $a \in \mathbb{R}^n$ *and* $c \in \mathbb{R}$ *and* $D := \{x \in \mathbb{R}^n |\ (v_i^j)^T x + b_i^j \leq 0,\ \forall v_i^j \in \mathcal{V}_{ij}, i = 1, \ldots, m, j = 1, \cdots, p,\}$ *and let* $\mathcal{V}_{ij} = [\underline{v}_i^j, \overline{v}_i^j] \subseteq \mathbb{R}^n$, $i = 1, \ldots, m$ *and* $j = 1, \ldots, p$. *Let* $x_0 \in K := C \cap D$. *Then, for any* $x \in \mathbb{R}^n$, $x_0 = P_K(x)$ *if and only if*

$$x_0 = \big(x - \sum_{i,j} \lambda_{ij} v_i^j\big) - \big(a^T x - \sum_{i,j} \lambda_{ij}\, a^T v_i^j - c\big) \frac{a}{\|a\|^2},$$

for some $v_i^j \in \mathcal{V}_{ij}$, $\lambda_{ij} \geq 0$ *such that* $\lambda_{ij}[(v_i^j)^T x_0 - b_i^j] = 0$, $i = 1, \cdots, m,\ j = 1, \cdots, p$.

Proof. Similar to the proof of Corollary 3.2, we see that D is a polyhedral set and so, $\{C, D\}$ has the strong CHIP. Thus, the preceding corollary implies that: for any $x \in \mathbb{R}^n$, $x_0 = P_K(x)$ if and only if $x_0 = P_C(x - \sum_{i,j} \lambda_{ij} v_i^j)$ for some $v_i^j \in \mathcal{V}_{ij}$, $\lambda_{ij} \geq 0$ such that $\lambda_{ij}[(v_i^j)^T x_0 - b_i^j] = 0$, $i = 1, \cdots, m,\ j = 1, \cdots, p$. Therefore, the conclusion follows by the fact that (cf. [5, Theorem 6.17])

$$P_C(u) = u - (a^T u - c) \frac{a}{\|a\|^2}.$$

□

Remark 3.3. (An Alternative Approach under an Additional Assumption) *In Theorem 3.1, we provided a self-contained and direct proof for the Lagrange multiplier characterization of best approximation. However, it is worthwhile noting an alternative method of proof for Theorem 3.1. Under the additional assumption that, for each* $x \in X$, $v_i \mapsto g_i(x, v_i)$ *is upper semicontinuous,* $i = 1, \ldots, m$, *one could also prove the conclusion of Theorem 3.1 whenever* $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \mathrm{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))^*$ *is closed and convex, by using the* *classical Ioffe-Tikhomirov theorem [21, Theorem 2.4.18] on subdifferential of a supremum of convex functions. To see this, note that, for each $x \in X$,*

$$\delta_D(x) = \sup_{\lambda_i \geq 0, v_i \in \mathcal{V}_i, i=1,\ldots,m} \sum_{i=1}^{m} \lambda_i g_i(x, v_i).$$

Let $x_0 \in D$. As each $\mathcal{V}_i$ is compact, $v_i \mapsto g_i(x,\ v_i)$ is upper semicontinuous for any $x \in X$ and each $x \mapsto g_i(x, v_i)$ is a real-valued convex function (and so, is continuous) for each $v_i \in \mathcal{V}_i$, Ioffe-Tikhomirov theorem gives us that

$$(D - x_0)^\circ = \partial\delta_D(x_0) = w^*\text{-cl}\big(\text{co} \bigcup \{\sum_{i=1}^{m} \lambda_i \partial g(\cdot, v_i)(x_0) : \lambda_i \geq 0,\ v_i \in \mathcal{V}_i, \lambda_i g_i(x_0, v_i) = 0\}\big).$$

As $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \text{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, v_i))^*$ is closed and convex, it can be verified that

$$\bigcup \{\sum_{i=1}^{m} \lambda_i \partial g(\cdot, v_i)(x_0) : \lambda_i \geq 0,\ v_i \in \mathcal{V}_i, \lambda_i g_i(x_0, v_i) = 0\}$$

is also closed and convex (and so is weak closed and convex). Thus,*

$$(D - x_0)^\circ = \bigcup \{\sum_{i=1}^{m} \lambda_i \partial g(\cdot, v_i)(x_0) : \lambda_i \geq 0,\ v_i \in \mathcal{V}_i, \lambda_i g_i(x_0, v_i) = 0\}.$$

So, we see that strong CHIP is equivalent to the Lagrange multiplier characterization.

Now, we provide an asymptotic multiplier characterization of best approximations under a relaxed assumption that $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \text{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, v_i))^*$ is only convex.

Theorem 3.2. (Asymptotic Multiplier Characterization under Uncertainty). *Let X be a Hilbert space. Let C, D be closed convex sets in X where $D = \{x \in X : g_i(x, v_i) \leq 0, \forall\, v_i \in \mathcal{V}_i,\ i = 1, \cdots, m\}$. Let $x_0 \in K := C \cap D$. Suppose that $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \text{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, v_i))^*$ is convex. Then the following assertions are equivalent.*
(i) *$\{C, D\}$ has the strong CHIP at x_0.*
(ii) *For any $x \in X$, $x_0 = P_K(x)$ if and only if $x_0 = P_C(x - l)$, for some $l \in \tilde{M}(x_0)$, where*

$$\tilde{M}(x_0) = \{x^* \in X^* : (x^*, x^*(x_0)) \in \text{cl}\big(\bigcup_{\lambda_i \geq 0, v_i \in \mathcal{V}_i} \text{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, v_i))^*\big)\}.$$

Proof. As before, it can be verified that (ii) $\Rightarrow$ (i) is always true. To see (i) $\Rightarrow$ (ii), we only need to show $(D - x_0)^\circ = \tilde{M}(x_0)$. From (3.3) and our assumption, we see that

$$\text{epi}\sigma_D \;=\; w^*\text{-cl}(\text{co} \bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \text{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, v_i))^*) = w^*\text{-cl}(\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \text{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, v_i))^*).$$

Note that $u \in (D - x_0)^\circ$ if and only if $(u, u(x_0)) \in \text{epi}\sigma_D(u)$ which is, in turn equivalent to

$$(u, u(x_0)) \in w^*\text{-cl}(\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \text{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, v_i))^*) = \tilde{M}(x_0).$$

where the second equality follows by our assumption and fact that, in a Hilbert space, a convex set is closed if and only if it is weakly convex. Hence, the conclusion follows. $\square$

In the case where there is no uncertainty, i.e., each $\mathcal{V}_i$ is a singleton, $i = 1, \ldots, m$, the characteristic cone

$$\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \text{epi}(\sum_{i=1}^{m} \lambda_i g_i(\cdot, v_i))^*$$

is always convex. So, in the uncertainty free case, Theorem 3.2 collapses to the known result of the sequential Lagrange multiplier characterization of the strong CHIP in [15]. It is worth noting from Theorem 3.2 that even asymptotic multiplier characterization in terms of strong CHIP holds under additional condition in the face of uncertainty.

4 Shape-Preserving Interpolations under Uncertainty

In this Section we derive the Lagrange multiplier characterization for the robust solution of a shape-preserving best approximation problem under ellipsoidal data uncertainty set under a robust strict feasibility condition.

Consider the constrained approximation problem under data uncertainty in the inequality constraints:

$$\begin{aligned}(SP) \quad & \min_{x \in L^2[0,1]} \quad \left(\int_0^1 (x(t) - \overline{x}(t))^2 dt \right)^{\frac{1}{2}} \\ & \text{s.t.} \quad \int_0^1 a_i(t)x(t)dt \leq \beta_i, \; i = 1, \ldots, m, \\ & \qquad \int_0^1 h_j(t)x(t)dt = d_j, \; j = 1, \ldots, r, \\ & \qquad x(t) \geq 0 \text{ a.e.,}\end{aligned}$$

where the data $(a_i, \beta_i) \in L^2[0,1] \times \mathbb{R}$ is uncertain and $\bar{x} \in L^2[0,1]$. We assume that the data $(a_i, \beta_i) \in L^2[0,1] \times \mathbb{R}$ in (SP) is uncertain and belongs to the ellipsoidal data uncertainty set $\overline{\mathcal{U}}_i^e \subseteq L^2[0,1] \times \mathbb{R}$ defined by

$$\overline{\mathcal{U}}_i^e = \{(a_i^{(0)}, \beta_i^{(0)}) + \sum_{l=1}^{k} w_i^{(l)}(a_i^{(l)}, \beta_i^{(l)}) : \|(w_i^{(1)}, \ldots, w_i^{(k)})\|_{\mathbb{R}^k} \leq 1\},$$

where $\|\cdot\|_{\mathbb{R}^k}$ denotes the usual Euclidean norm in $\mathbb{R}^k$,

Then, the robust counterpart of (SP) is

$$\begin{aligned}(RSP_e) \quad & \min_{x \in L^2[0,1]} \quad \left(\int_0^1 (\overline{x}(t) - x(t))^2 dt \right)^{\frac{1}{2}} \\ & \text{s.t.} \quad \int_0^1 a_i(t)x(t)dt \leq \beta_i, \; \forall (a_i, \beta_i) \in \overline{\mathcal{U}}_i^e, \; i = 1, \ldots, m, \\ & \qquad \int_0^1 h_j(t)x(t)dt = d_j, \; j = 1, \ldots, r, \\ & \qquad x(t) \in L_+^2[0,1],\end{aligned}$$

where $L^2_+[0,1] := \{x(t) \in L^2[0,1] : x(t) \geq 0 \text{ a.e.}\}$.

Let $\langle x_1, x_2\rangle$ be the inner product in $L^2[0,1]$ defined by $\langle x_1, x_2\rangle = \int_0^1 x_1(t)x_2(t)dt$. Define $C = L^2_+[0,1] := \{x \in L^2[0,1] : x(t) \geq 0 \text{ a.e.}\}$,

$$D = \{x \in L^2[0,1] : \langle u_i, x\rangle - r_i \leq 0,\ \forall (u_i, r_i) \in \mathcal{V}_i^e,\ i = 1, \dots, m+2r\},$$

where

$$\mathcal{V}_i^e = \begin{cases} \overline{\mathcal{U}}_i^e, & \text{if} \quad i = 1, \dots, m, \\ \{(h_{i-m}, d_{i-m})\}, & \text{if} \quad i = m+1, \dots, m+r, \\ \{-(h_{i-m-s}, d_{i-m-s})\}, & \text{if} \quad i = m+r+1, \dots, m+2r. \end{cases}$$

We now provide Lagrange multiplier characterization for the robust solution of the shape-preserving approximation problem under ellipsoidal data uncertainty set. Recall that $[z(t)]_+ = \max\{z(t), 0\}$ for each $z \in L^2[0,1]$, $L^2_{++}[0,1] := \{x \in L^2[0,1] : x(t) > 0\, a.e.\}$ and the second order cone SOC_k is defined by $\text{SOC}_k = \{(z_0, z_1, \dots, z_k) \in \mathbb{R}^{k+1} : z_0 \geq \|(z_1, \dots, z_k)\|_{\mathbb{R}^k}\}$.

Theorem 4.1. (Ellipsoidal data uncertainty) *Suppose that there exists $x^* \in L^2_{++}[0,1]$ such that $\int_0^1 a_i(t)x^*(t)dt < \beta_i$, $\forall (a_i, \beta_i) \in \overline{\mathcal{U}}_i^e$, $i = 1, \dots, m$, and $\int_0^1 h_j(t)x^*(t)dt = d_j$, $j = 1, \dots, r$. Then, for any $\overline{x} \in L^2[0,1]$, $x_0 \in K := C \cap D$ is a solution of (RSP_e) if and only if*

$$x_0 = [\overline{x} - \sum_{i=1}^m \sum_{l=0}^k s_i^{(l)} a_i^{(l)} - \sum_{j=1}^r \mu_j h_j]_+$$

for some $s_i = (s_i^{(0)}, \dots, s_i^{(k)}) \in \text{SOC}_k$, $\mu_j \in \mathbb{R}$ with $\sum_{l=0}^k s_i^{(l)}(\langle a_i^{(l)}, x_0\rangle - \beta_i^{(l)}) = 0$, $i = 1, \dots, m$.

Proof. Let $g_i(x, v_i) = \langle u_i, x\rangle - r_i \leq 0$ where $v_i = (u_i, r_i) \in \mathcal{V}_i^e$. It is clear that each $g_i(x, \cdot)$ is affine, and so is concave. Thus, $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i^e}} \text{epi}(\sum_{i=1}^{m+2r} \lambda_i g_i(\cdot, v_i))^*$ is convex. We now show that $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i^e}} \text{epi}(\sum_{i=1}^{m+2r} \lambda_i g_i(\cdot, v_i))^*$ is closed. Let $D_1 = \{x \in L^2[0,1] : \int_0^1 a_i(t)x(t)dt \leq \beta_i,\ \forall (a_i, \beta_i) \in \overline{\mathcal{U}}_i^e,\ i = 1, \dots, m\}$, and $D_2 = \{x \in L^2[0,1] : \int_0^1 h_j(t)x(t)dt = d_j,\ j = 1, \dots, r\}$. Define

$$G_i(x) = \sup_{(a_i, \beta_i) \in \overline{\mathcal{U}}_i^b} \{\int_0^1 a_i(t)x(t)dt - \beta_i\},\ i = 1, \dots, m.$$

Then $D_1 = \{x : G_i(x) \leq 0, i = 1, \dots, m\}$. As $x^* \in \text{int}(D_1) \cap D_2$,

$$\text{epi}\sigma_{D_1 \cap D_2} = \text{epi}\sigma_{D_1} + \text{epi}\sigma_{D_2} = \text{co} \bigcup_{\lambda_i \geq 0} \text{epi}(\sum_{i=1}^m \lambda_i G_i)^* + \bigcup_{\mu_j \in \mathbb{R}} \mu_j\{(h_j, d_j)\}.$$

where the second equality follows by $\sigma_{D_1} = \sup_{\lambda_i \geq 0}(\sum_{i=1}^m \lambda_i G_i)$, Lemma 2.1 and the fact that $\bigcup_{\lambda_i \geq 0} \text{epi}(\sum_{i=1}^m \lambda_i q_i)^*$ is closed if all q_i are continuous convex functions satisfying $\{x : q_i(x) < 0, i = 1, \dots, m\} \neq \emptyset$. By the definition of G_i and Lemma 2.1, $\text{epi}(\sum_{i=1}^m \lambda_i G_i)^* =$

$\mathrm{co}\bigcup_{\substack{\lambda_i\geq 0\\(a_i,\beta_i)\in\mathcal{U}_i^e}}\sum_{i=1}^m\{\lambda_i(a_i,\beta_i)\}$. So,

$$\begin{aligned}
\mathrm{epi}\sigma_{D_1\cap D_2} &= \mathrm{co}\bigcup_{\lambda_i\geq 0}\mathrm{co}\bigcup_{(a_i,\beta_i)\in\mathcal{U}_i^e}\sum_{i=1}^m\{\lambda_i(a_i,\beta_i)\}+\bigcup_{\mu_j\in\mathbb{R}}\mu_j\{(h_j,d_j)\}\\
&\subseteq \mathrm{co}\bigcup_{\substack{\lambda_i\geq 0\\(a_i,\beta_i)\in\mathcal{U}_i^e}}\{\lambda_i(a_i,\beta_i)\}+\bigcup_{\mu_j\in\mathbb{R}}\mu_j\{(h_j,d_j)\}\\
&\subseteq \mathrm{co}\Big(\bigcup_{\substack{\lambda_i\geq 0\\(u_i,r_i)\in\mathcal{V}_i}}\sum_{i=1}^{m+2r}\{\lambda_i(u_i,r_i)\}\Big)\\
&= \mathrm{co}\bigcup_{\substack{\lambda_i\geq 0\\v_i\in\mathcal{V}_i}}\mathrm{epi}(\sum_{i=1}^{m+2r}\lambda_i g_i(\cdot,v_i))^*=\bigcup_{\substack{\lambda_i\geq 0\\v_i\in\mathcal{V}_i}}\mathrm{epi}(\sum_{i=1}^{m+2r}\lambda_i g_i(\cdot,v_i))^*,
\end{aligned}$$

where the last equality follows as $\bigcup_{\substack{\lambda_i\geq 0\\v_i\in\mathcal{V}_i}}\mathrm{epi}(\sum_{i=1}^{m+2r}\lambda_i g_i(\cdot,v_i))^*$ is convex. The reverse inclusion

$$\bigcup_{\substack{\lambda_i\geq 0\\v_i\in\mathcal{V}_i}}\mathrm{epi}(\sum_{i=1}^{m+2r}\lambda_i g_i(\cdot,v_i))^*\subseteq\mathrm{epi}\sigma_{D_1\cap D_2}$$

always holds. So, $\bigcup_{\substack{\lambda_i\geq 0\\v_i\in\mathcal{V}_i}}\mathrm{epi}(\sum_{i=1}^{m+2r}\lambda_i g_i(\cdot,v_i))^*=\mathrm{epi}\sigma_{D_1\cap D_2}$ which is weakly closed and convex (and so, is closed). It is easy to see that $x^*\in\mathrm{qri}(C\cap D_1)\cap D_2$. Note that $\{A_1,A_2\}$ has strong CHIP at any $a\in A_1\cap A_2$ if either $\mathrm{int}A_1\cap A_2\neq\emptyset$ or $\mathrm{qri}(A_1)\cap A_2\neq\emptyset$ and A_2 is a polyhedral, where for a set A, $\mathrm{qri}A$ denotes the quasi relative interior of A. Now, for any $x_0\in K=C\cap D$,

$$(K-x_0)^\circ=(C\cap D_1-x_0)^\circ+(D_2-x_0)^\circ=(C-x_0)^\circ+(D_1-x_0)^\circ+(D_2-x_0)^\circ=(C-x_0)^\circ+(D-x_0)^\circ.$$

So, $\{C,D\}$ has the strong CHIP at x_0 for any $x_0\in K:=C\cap D$. Thus, by Theorem 3.1, for any $\overline{x}\in L^2[0,1]$, $x_0=P_K(\overline{x})$ if and only if

$$x_0=P_C(\overline{x}-\sum_{i=1}^{m+2r}\lambda_i\partial g_i(\cdot,v_i)(x_0))\text{ for some }\lambda_i\geq 0,v_i\in\mathcal{V}_i^e\text{ with }\sum_{i=1}^{m+2r}\lambda_i g_i(x_0,v_i)=0.$$

Observe that, for any $x\in D$ (and so, $\langle h_j,x\rangle=d_j$, $j=1,\ldots,r$)

$$\begin{aligned}
&\bigcup_{\substack{\lambda\in\mathbb{R}_+^{m+2r}\\v\in\Pi_{i=1}^{m+2r}\mathcal{V}_i}}\{\partial(\sum_{i=1}^{m+2r}\lambda_i g_i(\cdot,v_i))(x):\sum_{i=1}^{m+2r}\lambda_i g_i(x,v_i)=0\}\\
=&\bigcup_{\substack{\lambda\in\mathbb{R}_+^m,\mu_j\in\mathbb{R}\\\|(w_i^{(1)},\ldots,w_i^{(k)})\|_{\mathbb{R}^k}\leq 1}}\{\sum_{i=1}^m\lambda_i(a_i^{(0)}+\sum_{l=1}^k w_i^{(l)}a_i^{(l)})+\sum_{j=1}^r\mu_j h_j\\
&\qquad:\lambda_i(\langle a_i^{(0)}+\sum_{l=1}^k w_i^{(l)}a_i^{(l)},x\rangle-(\beta_i^{(0)}+\sum_{l=1}^k w_i^{(l)}\beta_i^{(l)}))=0\}.
\end{aligned}$$

So, for any $\overline{x} \in L^2[0,1]$, $x_0 = P_K(\overline{x})$ if and only if

$$x_0 = P_C(\overline{x} - \sum_{i=1}^{m} \lambda_i (a_i^{(0)} + \sum_{l=1}^{k} w_i^{(l)} a_i^{(l)}) - \sum_{j=1}^{r} \mu_j h_j)$$

for some $\lambda_i \geq 0$, $\mu_j \in \mathbb{R}$ and $\|(w_i^{(1)}, \ldots, w_i^{(k)})\|_{\mathbb{R}^k} \leq 1$ with $\lambda_i(\langle a_i^{(0)} + \sum_{l=1}^{k} w_i^{(l)} a_i^{(l)}, x_0\rangle - (\beta_i^{(0)} + \sum_{l=1}^{k} w_i^{(l)} \beta_i^{(l)})) = 0$. Let $s_i = (s_i^{(1)}, \ldots, s_i^{(k)})$, $s_i^{(l)} = \lambda_i w_i^{(l)}$, $l = 1, \ldots, k$ and $s_i^{(0)} = \lambda_i \geq 0$. Note that

$$\|(w_i^{(1)}, \ldots, w_i^{(k)})\|_{\mathbb{R}^k} \leq 1,\ \lambda_i \geq 0 \ \Leftrightarrow\ \|(s_i^{(1)}, \ldots, s_i^{(k)})\|_{\mathbb{R}^k} \leq s_i^{(0)}.$$

So, for any $\overline{x} \in L^2[0,1]$, $x_0 = P_K(\overline{x})$ if and only if

$$x_0 = [\overline{x} - \sum_{i=1}^{m} \sum_{l=0}^{k} s_i^{(l)} a_i^{(l)} - \sum_{j=1}^{r} \mu_j h_j]_+$$

for some $s_i = (s_i^{(0)}, \ldots, s_i^{(k)}) \in \mathrm{SOC}_k$, $\mu_j \in \mathbb{R}$ with $\sum_{l=0}^{k} s_i^{(l)}(\langle a_i^{(l)}, x_0\rangle - \beta_i^{(l)}) = 0$, $i = 1, \ldots, m$. Therefore, the conclusion follows as $x_0 = P_K(\overline{x})$ means x_0 is a solution of (RSP_e) and $P_C(a) = [a]_+$ for any $a \in L^2[0,1]$. $\square$

Remark 4.1. (Reducing the multiplier conditions to solving convex second-order cone programs) *We now see how the Lagrange multiplier characterization for the robust solution of the shape-preserving interpolation problem results from a solution of convex second-order cone programming problem. To see this, let* $s := (s_1^{(0)}, \ldots, s_1^{(k)}, \ldots, s_m^{(0)}, \ldots, s_m^{(k)}) \in \mathbb{R}^{m(k+1)}$ *and* $\mu := (\mu_1, \ldots, \mu_r) \in \mathbb{R}^r$*. Let* $u_{s,\mu}(t) := \sum_{i=1}^{m} \sum_{l=0}^{k} s_i^{(l)} \overline{a}_i^{(l)}(t) + \sum_{j=1}^{r} \mu_j h_j(t)$ *and let*

$$g(s,\mu) \ := \ \frac{1}{2} \int_0^1 \Big(\overline{x}(t) - u_{s,\mu}(t) - [\overline{x}(t) - u_{s,\mu}(t)]_+\Big)^2 dt + \sum_{i=1}^{m} \sum_{l=0}^{k} s_i^{(l)} \overline{\beta}_i^{(l)} + \sum_{j=1}^{r} \mu_j d_j.$$

As $(s,\mu) \mapsto u_{s,\mu}$ *is linear and* $f : L^2[0,1] \mapsto \mathbb{R}$*, defined by* $f(x) = \frac{1}{2}\int_0^1 (x(t) - [x(t)]_+)^2 dt$*, is a continuously differentiable convex function with* $\nabla f(x) = x(t) - [x(t)]_+$*, we see that* g *is a continuously differentiable convex on* $\mathbb{R}^{m(k+1)+r}$*. So,* x_0 *is a solution of* (RSP_e) *if and only if* $x_0 = [\overline{x} - \sum_{i=1}^{m} \sum_{l=0}^{k} \overline{s}_i^{(l)} \overline{a}_i^{(l)} - \sum_{j=1}^{r} \overline{\mu}_j h_j]_+$ *where* $(\overline{s}, \overline{\mu})$ *is a solution of the following convex second order cone program*

$$(P) \qquad \min_{s \in \mathbb{R}^{m(k+1)}, \mu \in \mathbb{R}^s} \quad g(s,\mu)$$

$$s.t. \qquad (s,\mu) \in \prod_{i=1}^{m} \mathrm{SOC}_k \times \mathbb{R}^s.$$

Indeed, employing the well-known optimality conditions in convex programming (see, e.g., [21]) in the setting under consideration, we obtain that $(\overline{s}, \overline{\mu})$ *is an optimal solution to (P) if and only if it satisfies the following conditions*

$$\begin{cases} s \in \prod_{i=1}^{m} \mathrm{SOC}_k, \\ \nabla_s g(s,\mu) \in \prod_{i=1}^{m} \mathrm{SOC}_k, \\ \nabla_s g(s,\mu)^T s = 0,\ \nabla_\mu g(s,\mu) = 0. \end{cases} \tag{4.6}$$

Let $x_0 = [\overline{x} - \sum_{i=1}^m \sum_{l=0}^k \overline{s}_i^{(l)} \overline{a}_i^{(l)} - \sum_{j=1}^r \overline{\mu}_j h_j]_+$. *Clearly,* $x_0 \in C := L^2_+[0,1]$. *Note that*

$$\nabla_\mu g(s,\mu) = 0 \;\Leftrightarrow\; \langle h_j, x_0\rangle = d_j,\; j = 1,\ldots,r$$

and

$$\begin{aligned}\nabla_s g(s,\mu) \in \prod_{i=1}^m \mathrm{SOC}_k \;\;&\Leftrightarrow\;\; \big(-\langle a_i^{(0)}, x_0\rangle + \beta_i^{(0)}, \ldots, -\langle a_i^{(k)}, x_0\rangle + \beta_i^{(k)}\big) \in \mathrm{SOC}_k,\; i = 1,\ldots,m \\ &\Leftrightarrow\;\; \int_0^1 a_i(t)x_0(t)dt \le \beta_i,\, \forall (a_i,\beta_i) \in \overline{\mathcal{U}}_i^e,\, i = 1,\ldots,m.\end{aligned}$$

Hence, $x_0 \in D$ *and so,* $x_0 \in K := C \cap D$. *Finally, as*

$$s \in \prod_{i=1}^m \mathrm{SOC}_k \text{ and } \nabla_s g(s,\mu)^T s = 0 \;\Leftrightarrow\; \sum_{l=0}^k s_i^{(l)}(\langle a_i^{(l)}, x_0\rangle - \beta_i^{(l)}) = 0,\; i = 1,\ldots,m,$$

the preceding theorem shows that x_0 *is a solution of* (RSP_e) *if and only if* $x_0 = [\overline{x} - \sum_{i=1}^m \sum_{l=0}^k \overline{s}_i^{(l)} \overline{a}_i^{(l)} - \sum_{j=1}^r \overline{\mu}_j h_j]_+$, *where* $(\overline{s},\overline{\mu})$ *is a solution of (P).*

Now assume that the data $(a_i,\beta_i) \in L^2[0,1] \times \mathbb{R}$ in (SP) is uncertain and belongs to the box data uncertainty set $\overline{\mathcal{U}}_i^b \subseteq L^2[0,1] \times \mathbb{R}$ defined by

$$\overline{\mathcal{U}}_i^b = \{(a_i^{(0)}, \beta_i^{(0)}) + \sum_{l=1}^k w_i^{(l)}(a_i^{(l)}, \beta_i^{(l)}) : \|(w_i^{(1)}, \ldots, w_i^{(k)})\|_\infty \le 1\},$$

where $\|x\|_\infty = \max\{|x_i| : 1 \le i \le k\}$ denotes the usual supremum norm in $\mathbb{R}^k$,

Then, the robust counterpart of (SP) is

$$\begin{aligned}(RSP_b)\quad \min_{x\in L^2[0,1]}\quad & \left(\int_0^1 (\overline{x}(t) - x(t))^2 dt\right)^{\frac{1}{2}} \\ \text{s.t.}\quad & \int_0^1 a_i(t)x(t)dt \le \beta_i,\, \forall (a_i,\beta_i) \in \overline{\mathcal{U}}_i^b,\, i = 1,\ldots,m, \\ & \int_0^1 h_j(t)x(t)dt = d_j,\, j = 1,\ldots,r, \\ & x(t) \in L^2_+[0,1],\end{aligned}$$

where $L^2_+[0,1] := \{x(t) \in L^2[0,1] : x(t) \ge 0 \text{ a.e.}\}$. Define $C = L^2_+[0,1] := \{x \in L^2[0,1] : x(t) \ge 0 \text{ a.e.}\}$, $D = \{x \in L^2[0,1] : \langle u_i, x\rangle - r_i \le 0, \forall (u_i, r_i) \in \mathcal{V}_i^b,\, i = 1,\ldots,m+2r\}$ and

$$\mathcal{V}_i^b = \begin{cases} \overline{\mathcal{U}}_i^b, & \text{if } \; i = 1,\ldots,m, \\ \{(h_{i-m}, d_{i-m})\}, & \text{if } \; i = m+1,\ldots,m+r, \\ \{-(h_{i-m-s}, d_{i-m-s})\}, & \text{if } \; i = m+r+1,\ldots,m+2r.\end{cases}$$

Then, finding a solution of (RSP_b) is equivalent to find the projection of $\overline{x}$ from the set $K = C \cap D$.

Theorem 4.2. (Box data uncertainty). *For the problem (RSP_b), the following two statements are equivalent:*

(i) *$\{C, D\}$ has the strong CHIP at x_0;*

(ii) *For any $\overline{x} \in L^2[0,1]$, $x_0 \in K := C \cap D$ is a solution of (RSP_b) if and only if*

$$x_0 = [\overline{x} - \sum_{i=1}^{m} \sum_{l=0}^{k} s_i^{(l)} a_i^{(l)} - \sum_{j=1}^{r} \mu_j h_j]_+$$

for some $s_i = (s_i^{(0)}, \ldots, s_i^{(k)}) \in \{(z_0, \ldots, z_k) : z_0 \geq |z_l|, \, l = 1, \ldots, k\}$, $\mu_j \in \mathbb{R}$ with

$$\sum_{l=0}^{k} s_i^{(l)} (\langle a_i^{(l)}, x_0 \rangle - \beta_i^{(l)}) = 0, \; i = 1, \ldots, m.$$

Proof. Let $g_i(x, v_i) = \langle u_i, x \rangle - r_i \leq 0$ where $v_i = (u_i, r_i) \in \mathcal{V}_i^b$. It is clear that each $g_i(x, \cdot)$ is affine, and so is concave. Thus,

$$\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i^b}} \text{epi}(\sum_{i=1}^{m+2r} \lambda_i g_i(\cdot, v_i))^* = \bigcup_{\substack{\lambda_i \geq 0 \\ (a_i, \beta_i) \in \mathcal{V}_i^b}} \sum_{i=1}^{m+2r} \{(u_i, r_i)\},$$

is convex. Moreover, as $\mathcal{V}_i^b$ is a polytope and each $g_i(x, \cdot)$ is affine, $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i^b}} \text{epi}(\sum_{i=1}^{m+2r} \lambda_i g_i(\cdot, v_i))^*$ is closed. Thus, from Theorem 3.1, $\{C, D\}$ has the strong CHIP at x_0 is equivalent to: for any $\overline{x} \in L^2[0,1]$, $x_0 = P_K(\overline{x})$ if and only if

$$x_0 = P_C(\overline{x} - \sum_{i=1}^{m+2r} \lambda_i \partial g_i(\cdot, v_i)(x_0)) \text{ for some } \lambda_i \geq 0, v_i \in \mathcal{V}_i^b \text{ with } \sum_{i=1}^{m+2r} \lambda_i g_i(x_0, v_i) = 0.$$

Similar to the proof of Theorem 4.1, we see that $\{C, D\}$ has the strong CHIP at x_0 is equivalent to: for any $\overline{x} \in L^2[0,1]$, $x_0 = P_K(\overline{x})$ if and only if

$$x_0 = P_C(\overline{x} - \sum_{i=1}^{m} \lambda_i (a_i^{(0)} + \sum_{l=1}^{k} w_i^{(l)} a_i^{(l)}) - \sum_{j=1}^{r} \mu_j h_j)$$

for some $\lambda_i \geq 0$, $\mu_j \in \mathbb{R}$ and $\|(w_i^{(1)}, \ldots, w_i^{(k)})\|_\infty \leq 1$ with $\lambda_i(\langle a_i^{(0)} + \sum_{l=1}^{k} w_i^{(l)} a_i^{(l)}, x_0 \rangle - (\beta_i^{(0)} + \sum_{l=1}^{k} w_i^{(l)} \beta_i^{(l)})) = 0$. Let $s_i = (s_i^{(1)}, \ldots, s_i^{(k)})$, $s_i^{(l)} = \lambda_i w_i^l$, $l = 1, \ldots, k$ and $s_i^{(0)} = \lambda_i \geq 0$. Note that

$$\|(w_i^{(1)}, \ldots, w_i^{(k)})\|_\infty \leq 1, \, \lambda_i \geq 0 \; \Leftrightarrow \; s_i^{(0)} \geq |s_i^{(l)}|, l = 1, \ldots, k.$$

So, for any $\overline{x} \in L^2[0,1]$, $x_0 = P_K(\overline{x})$ if and only if

$$x_0 = [\overline{x} - \sum_{i=1}^{m} \sum_{l=0}^{k} s_i^{(l)} a_i^{(l)} - \sum_{j=1}^{r} \mu_j h_j]_+$$

for some $s_i = (s_i^{(0)}, \ldots, s_i^{(k)})$ with $s_i^{(0)} \geq |s_i^{(l)}|$, $l = 1, \ldots, k$, $\mu_j \in \mathbb{R}$ with

$$\sum_{l=0}^{k} s_i^{(l)}(\langle a_i^{(l)}, x_0\rangle - \beta_i^{(l)}) = 0, \; i = 1, \ldots, m.$$

Therefore, the conclusion follows as $x_0 = P_K(\overline{x})$ means x_0 is a solution of (RSP_b) and $P_C(a) = [a]_+$ for any $a \in L^2[0,1]$. □

As a corollary, we obtain a Lagrange multiplier characterization of the solution of (RSP_b), under a weaker version of strict feasibility condition.

Corollary 4.1. *Suppose that there exists $x^* \in L^2_{++}[0,1]$ such that $\int_0^1 a_i(t)x^*(t)dt \leq \beta_i$, $\forall (a_i, \beta_i) \in \overline{\mathcal{U}}_i^b$, $i = 1, \ldots, m$, and $\int_0^1 h_j(t)x^*(t)dt = d_j$, $j = 1, \ldots, r$. Then, for any $\overline{x} \in L^2[0,1]$, $x_0 \in K := C \cap D$ is a solution of (RSP_b) if and only if*

$$x_0 = [\overline{x} - \sum_{i=1}^{m}\sum_{l=0}^{k} s_i^{(l)} a_i^{(l)} - \sum_{j=1}^{r} \mu_j h_j]_+$$

for some $s_i = (s_i^{(0)}, \ldots, s_i^{(k)}) \in \{(z_0, \ldots, z_k) : z_0 \geq |z_l|, l = 1, \ldots, k\}$, $\mu_j \in \mathbb{R}$ with

$$\sum_{l=0}^{k} s_i^{(l)}(\langle a_i^{(l)}, x_0\rangle - \beta_i^{(l)}) = 0, \; i = 1, \ldots, m.$$

Proof. As there exists $x^* \in L^2_{++}[0,1]$ such that $\int_0^1 a_i(t)x^*(t)dt \leq \beta_i$, $\forall (a_i, \beta_i) \in \overline{\mathcal{U}}_i^b$, $i = 1, \ldots, m$, and $\int_0^1 h_j(t)x^*(t)dt = d_j$, $j = 1, \ldots, r$, we have $x^* \in \text{qri}(C) \cap D$ where $C = L^2_+[0,1] := \{x \in L^2[0,1] : x(t) \geq 0 \text{ a.e.}\}$ and $D = \{x \in L^2[0,1] : \langle u_i, x\rangle - r_i \leq 0, \forall (u_i, r_i) \in \mathcal{V}_i^b, i = 1, \ldots, m + 2r\}$. Note that $D = \{x \in L^2[0,1] : \sup_{(u_i, r_i) \in \mathcal{V}_i^b}\{\langle u_i, x\rangle - r_i\} \leq 0, i = 1, \ldots, m\}$ and the supremum of a linear function over a polytope is attained at one of the finitely many extreme points of the polytope. So, D is a polyhedral. Thus, $\{C, D\}$ has strong CHIP, and hence, the conclusion follows from the preceding theorem. □

5 Conclusion and Further Research

In this paper we have shown that Lagrange multiplier characterization of robust best approximation in terms of strong CHIP depends both on convexity and closure of the characteristic cone $\bigcup_{\substack{\lambda_i \geq 0 \\ v_i \in \mathcal{V}_i}} \text{epi}(\sum_{i=1}^m \lambda_i g_i(\cdot, v_i))^*$. We have seen, in particular, that the convexity of the characteristic cone relies upon the geometry of the function $g(x, \cdot)$ whereas the closure of the cone may depend on the geometric structure of the uncertainty set $\mathcal{V}_i$. As an application, we have also established a Lagrange multiplier condition characterizing robust shape-preserving interpolation under ellipsoidal uncertainty. One approach to solving such robust problems is to reformulate the Lagrange multiplier condition as a nonsmooth equation using the second order cone complementary function and then design semismooth Newton methods. On the other hand, it would also be of interest to study other Robust Optimization approaches to solve best approximation problems under data uncertainty. They will be investigated in a forthcoming study.

References

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

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

[3] R. S. Burachik and V. Jeyakumar, A simple closure condition for the normal cone intersection formula, *Proc. Amer. Math. Soc,* **133**(6) (2004), 1741-1748.

[4] F. Deutsch, The role of conical hull intersection property in convex optimization and approximation, in Approximation Theory IX, C.K. Chui and L. L. Schumaker, eds. (1998), Vanderbilt University Press, Nashville, TN.

[5] F. Deutsch, Best Approximation in Inner Product Spaces, Springer-Verlag, New York, 2000.

[6] F. Deutsch, W. Li and J. Swetits, Fenchel duality and the strong conical hull intersection property, *J. Optim. Theory Appl.,* **102** (1999), 681-695.

[7] F. Deutsch, W. Li and J. Ward, A dual approach to constrained interpolation from a convex subset of Hilbert space, *J. Approx. Theory,* **90** (1997), 385-444.

[8] F. Deutsch, W. Li and J. D. Ward, Best approximation from the intersection of a closed convex set and a polyhedron in Hilbert space, weak Slater conditions, and the strong conical hull intersection property, *SIAM J. Optim.,* **10** (1999), 252-268.

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

[10] V. Jeyakumar and G. Li, Strong duality in robust convex programming: complete characterizations , *SIAM J. Optim.,* **6** (2010), 3384-3407.

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

[12] V. Jeyakumar, G. Li and G.M. Lee, A Robust von Neumann Minimax Theorem for Zero-Sum Games under Bounded Payoff Uncertainty, *Oper. Res. Lett.*, **39** (2011), 109-114.

[13] 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(2)** (2003), 534-547.

[14] V. Jeyakumar and H. Mohebi, A global approach to nonlinearly constrained best approximation, *Numer. Funct. Anal. Optim.*, **26** (2005), no. 2, 205-227.

[15] V. Jeyakumar and H. Mohebi, Limiting $\epsilon-$subgradient characterizations of constrained best approximation, *J. Approx. Theory* **135(2)** (2005), 145-159.

[16] C. Kanzow, I. Ferenczi and M. Fukushima, On the local convergence of semismooth Newton methods for linear and nonlinear second-order cone programs without strict complementarity, *SIAM J. Optim.*, **20** (2009), no. 1, 297-320.

[17] C. Li and X. Jin, Nonlinearly constrained best approximation in Hilbert spaces: the strong CHIP, and the basic constraint qualification, *SIAM J. Optim.*, **13(1)** (2002), 228-239.

[18] C. Li and K. F. Ng, Constraint qualification, the strong CHIP and best approximation with convex constraints in Banach spaces, *SIAM J. Optim.,* **14** (2003), 584-607.

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

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

[21] C. Zălinescu, Convex Analysis in General Vector Spaces, World Scientific, London, 2002.