Robust Best Approximation with Interpolation Constraints under Ellipsoidal Uncertainty: Strong Duality and Nonsmooth Newton Methods *

V. Jeyakumar† G. Y. Li‡ B. S. Mordukhovich § and J. H. Wang¶

Revised Version: February 5, 2013

Abstract

In this paper we present a duality approach for finding a robust best approximation from a set involving interpolation constraints and uncertain inequality constraints in a Hilbert space that is immunized against the data uncertainty using a nonsmooth Newton method. Following the framework of robust optimization, we assume that the input data of the inequality constraints are not known exactly while they belong to an ellipsoidal data uncertainty set. We first show that finding a robust best approximation is equivalent to solving a second-order cone complementarity problem by establishing a strong duality theorem under a strict feasibility condition. We then examine a nonsmooth version of Newton's method and present their convergence analysis in terms of the metric regularity condition.

Keywords: Robust optimization, nonsmooth Newton methods, ellipsoidal uncertainty, best approximations, duality.

---

*The authors are grateful to the referees for their valuable suggestions and helpful comments 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 Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: g.li@unsw.edu.au.

§Department of Mathematics, Wayne State University, Detroit, MI 48202, USA. Email: boris@math.wayne.edu.

¶Department of Mathematics, Zhejiang University of Technology, Hangzhou 310032, P. R. China, and Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: wjh@zjut.edu.cn.

1 Introduction

This paper is devoted to robust best approximation under uncertain conditions. More specifically, it deals with *uncertain linear inequality constraints*, *ellipsoid uncertainty* and *exact interpolation constraints*. The ellipsoid uncertainty model covers a broad and tractable class of infinite-dimensional robust optimization problems, which naturally arise in various aspects of operations research and its applications; see, e.g., [2, 3, 8, 11, 23] and the references therein. Let us start with formulating the best *approximation/optimization problem* in the Hilbert space $L^2[0,1]$ described as follows:

$$
\begin{aligned}
\text{(SP)} \quad & \min_{x \in L^2[0,1]} && \frac{1}{2}\int_0^1 x(t)^2 dt \\
& \text{s.t.} && \int_0^1 a_i(t)x(t)dt \le \beta_i,\ i=1,\ldots,m, \\
& && \int_0^1 h_j(t)x(t)dt = d_j,\ j=1,\ldots,s,
\end{aligned}
$$

where the data $a_i \in L^2[0,1]$ and $\beta_i \in \mathbb{R}$ as $i=1,\ldots,m$ of the inequality constraints are uncertain and belong to the *ellipsoidal uncertainty sets*

$$
\overline{\mathcal{U}}_i := \left\{ (\overline{a}_i^{(0)}, \overline{\beta}_i^{(0)}) + \sum_{l=1}^{k} w_i^{(l)} (\overline{a}_i^{(l)}, \overline{\beta}_i^{(l)}) \,\Big|\, \|w_i\|_{\mathbb{R}^k} \le 1 \right\},\ i=1,\ldots,m,
$$

with $h_j \in L^2[0,1]$, $d_j \in \mathbb{R}$ as $j=1,\ldots,s$, and $k \in \mathbb{N}$. A comprehensive account of solvable cases of uncertainty sets in convex optimization can be found in [3, 4]. In the setting where the data (a_i, β_i) are assumed to be known exactly, problems of the above type have been systematically studied and efficiently solved by various numerical techniques; see [6, 7, 24, 19, 20, 35]. However, in reality the data of the constraints are often uncertain (i.e., they are not known exactly) due to estimation and prediction errors, lack of information, etc. We examine the model problem (SP) in the face of data uncertainty by employing the robust optimization framework. It is based on the description of uncertainty via sets as opposed to probability distributions, which are commonly used in stochastic approaches; see, e.g., the books [5, 32] and their bibliographies. This paper develops the *deterministic approach* to optimization under uncertainty and the corresponding *robust duality theory* elements of which have been successfully employed in [1, 14, 15, 17, 16]. It allows us to derive and investigate appropriate versions of the *nonsmooth Newton method* for finding a robust best approximation that is immunized against the data uncertainty. Our basic framework is to solve the following *robust counterpart* of (SP):

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

where the uncertain constraints are enforced for all possible uncertainties within the uncertainty sets $\overline{\mathcal{U}}_i$, $i=1,\ldots,m$, and where at least one feasible solution is assumed to exist. In this vein we make the following *contributions* to constrained best approximation under input data uncertainty. **(i)** Under a strict feasibility condition, we prove that *strong duality* between the infinite-dimensional robust optimization problem (RSP) and its *finite-dimensional* (optimistic) dual problem

$$
\begin{aligned}
\text{(ODSP)} \quad \max_{z_i^{(l)}\in\mathbb{R},\mu_j\in\mathbb{R}} \quad & -\frac{1}{2}\int_0^1 \Big(\sum_{i=1}^m\sum_{l=0}^k z_i^{(l)}\overline{a}_i^{(l)}(t) + \sum_{j=1}^s \mu_j h_j\Big)^2 dt - \sum_{i=1}^m\sum_{l=0}^k z_i^{(l)}\overline{\beta}_i^{(l)} - \sum_{j=1}^s \mu_j d_j \\
\text{s.t.} \quad & \|(z_i^{(1)},\ldots,z_i^{(k)})\|_{\mathbb{R}^k} \le z_i^{(0)}, \ i=1,\ldots,m.
\end{aligned}
$$

holds. We also provide a precise representation of the (unique) robust best approximation via the corresponding optimal solution of the optimistic dual problem. This is illustrated by a numerical example. **(ii)** By reformulating (ODSP) in terms of an equivalent *second-order cone complementarity problem*, we establish that finding a robust best approximation is equivalent to finding a solution to a *strongly semismooth equation* in finite dimensions. This makes it possible for us to present a version of the nonsmooth Newton method in our robust optimization framework and provide its convergence analysis.

The rest of the paper is organized as follows. In Section 2 we show how a robust solution to the original infinite-dimensional (SP), which is also a solution to the robust optimization problem (RSP), can be obtained via solving the finite-dimensional dual problem (ODSP). A numerical example is given to illustrate this strong duality result. In Section 3 we establish that solving (ODSP) is equivalent to solving a nonsmooth equation, which is proved to be strongly semismooth. In Section 4, we present well-posedness and local convergence analysis of the basic (B-differentiable) version of the nonsmooth Newton algorithm developed in this paper for robust optimization. We also briefly discuss its other semismooth counterparts. We conclude the paper by discussing some future research directions in the final Section 5.

2 Strong Duality in Robust Best Approximation

In this section we construct a robust dual problem, known as the optimistic counterpart of the uncertain Lagrange dual problem, and establish a strong duality theorem. It is shown then how the robust best approximation can be derived from a dual solution. Let $X = L^2[0,1]$. Recall that the robust counterpart of (SP) is formulated as

$$
\begin{aligned}
\text{(RSP)} \quad \min_{x\in X} \quad & \frac{1}{2}\int_0^1 x(t)^2 dt \\
\text{s.t.} \quad & \int_0^1 a_i(t)x(t)dt \le \beta_i,\ \forall (a_i,\beta_i) \in \overline{\mathcal{U}}_i,\ i=1,\dots,m, \\
& \int_0^1 h_j(t)x(t)dt = d_j,\ j=1,\dots,s.
\end{aligned}
$$

Define $f : L^2[0,1] \to \mathbb{R}$ by $f(x) := \frac{1}{2}\|x\|^2 = \frac{1}{2}\langle x,x\rangle$, where $\langle\cdot,\cdot\rangle$ is the L^2 inner product given by $\langle x_1,x_2\rangle := \int_0^1 x_1(t)x_2(t)\mathrm{d}t$ for all $x_1,x_2 \in L^2[0,1]$. The *uncertain Lagrangian dual problem* of (SP) is defined by

$$
\text{(DSP)} \quad \max_{\substack{\lambda_i\ge 0\\ \mu_j\in\mathbb{R}}} \inf_{x\in X} \Big\{ f(x) + \sum_{i=1}^m \lambda_i(\langle a_i,x\rangle - \beta_i) + \sum_{j=1}^s \mu_j(\langle h_j,x\rangle - d_j) \Big\}
$$

for each uncertain parameter (a_i,β_i). The *optimistic counterpart* of the uncertain Lagrangian dual problem is given by

$$
\text{(ODSP)} \qquad \max_{\substack{(a_i,\beta_i)\in\overline{\mathcal{U}_i}\\ \lambda_i\ge 0,\mu_j\in\mathbb{R}}} \inf_{x\in X} \Big\{ f(x) + \sum_{i=1}^m \lambda_i(\langle a_i,x\rangle - \beta_i) + \sum_{j=1}^s \mu_j(\langle h_j,x\rangle - d_j) \Big\},
$$

where the maximum is taken over all the pairs $(a_i,\beta_i) \in \overline{\mathcal{U}_i}$. We begin with establishing the robust strong duality under a strict feasibility condition.

Theorem 2.1 (Robust Strong Duality). *For the robust problem* (RSP)*, assume the fulfillment of the following* STRICT FEASIBILITY CONDITION*: there is $x_0 \in L^2[0,1]$ such that*

$$
\langle a_i,x_0\rangle < \beta_i\ \forall\,(a_i,\beta_i)\in\overline{\mathcal{U}}_i \ \ \textit{and}\ \ \langle h_j,x_0\rangle = d_j \ \ \textit{as}\ \ i=1,\dots,m,\ j=1,\dots,s. \tag{2.1}
$$

Then we have the ROBUST STRONG DUALITY $\min\text{(RSP)} = \max\text{(ODSP)}$*, where the minimum in* (RSP) *is uniquely realized.*

Proof. Since the objective function in (RSP) is *strongly convex*, this problem admits a unique optimal solution. Observe first that problem (RSP) can be equivalently written as

$$
\inf_{x\in X} \Big\{ f(x) \Big|\ g_i(x) \le 0 \ \ \text{as}\ \ i=1,\dots,m \ \ \text{and}\ \ \langle h_j,x\rangle = d_j \ \ \text{as}\ \ j=1,\dots,s \Big\},
$$

where $g_i(x) := \sup_{(a_i,\beta_i)\in\overline{\mathcal{U}}_i} \{\langle a_i, x\rangle - \beta_i\}$, $i = 1,\dots,m$, are real-valued convex functions on X. From the strong duality result of convex programming [13, Theorem 3.1] we get that

$$\begin{aligned}\min\,(\text{RSP}) &= \max_{\substack{\lambda_i\geq 0\\ \mu_j\in\mathbb{R}}} \inf_{x\in X} \Big\{ f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^s \mu_j(\langle h_j, x\rangle - d_j)\Big\}\\ &= \max_{\substack{\lambda_i\geq 0\\ \mu_j\in\mathbb{R}}} \inf_{x\in X} \sup_{(a_i,\beta_i)\in\overline{\mathcal{U}}_i} \Big\{ f(x) + \sum_{i=1}^m \lambda_i(\langle a_i, x\rangle - \beta_i) + \sum_{j=1}^s \mu_j(\langle h_j, x\rangle - d_j)\Big\}.\end{aligned}$$

Since the function $x \mapsto f(x) + \sum_{i=1}^m \lambda_i(\langle a_i, x\rangle - \beta_i)$ is convex, the function $(a_i, \beta_i) \mapsto f(x) + \sum_{i=1}^m \lambda_i(\langle a_i, x\rangle - \beta_i)$ is linear (hence concave), and the set $\overline{\mathcal{U}}_i$ is compact, the convex-concave minimax result of [35, Theorem 2.10.2] ensures that

$$\begin{aligned}&\max_{\substack{\lambda_i\geq 0\\ \mu_j\in\mathbb{R}}} \inf_{x\in X} \sup_{(a_i,\beta_i)\in\overline{\mathcal{U}}_i} \Big\{ f(x) + \sum_{i=1}^m \lambda_i(\langle a_i, x\rangle - \beta_i) + \sum_{j=1}^s \mu_j(\langle h_j, x\rangle - d_j)\Big\}\\ =&\max_{\substack{\lambda_i\geq 0\\ \mu_j\in\mathbb{R}}} \max_{(a_i,\beta_i)\in\overline{\mathcal{U}}_i} \inf_{x\in X} \Big\{ f(x) + \sum_{i=1}^m \lambda_i(\langle a_i, x\rangle - \beta_i) + \sum_{j=1}^s \mu_j(\langle h_j, x\rangle - d_j)\Big\},\end{aligned}$$

which yields therefore the conclusion of this theorem. □

Next we show that the optimistic dual counterpart (ODSP) can be equivalently rewritten as a *second-order cone programming problem.*

Theorem 2.2 (Second-order Cone Programming Model of the Dual Problem). *Let*

$$\text{SOC}_k := \Big\{(x_0, x_1, \dots, x_k) \in \mathbb{R}^{k+1} \Big|\; x_0 \geq \|(x_1, \dots, x_k)\|_{\mathbb{R}^k}\Big\}$$

be the second-order cone in $\mathbb{R}^{k+1}$*. Then the dual problem* (ODSP) *is equivalent to the following problem of second-order cone programming:*

$$\begin{aligned}\max\Big\{ &-\frac{1}{2}\int_0^1 \Big(\sum_{i=1}^m\sum_{l=0}^k z_i^{(l)}\overline{a}_i^{(l)}(t) + \sum_{j=1}^s \mu_j h_j\Big)^2 dt - \sum_{i=1}^m\sum_{l=0}^k z_i^{(l)}\overline{\beta}_i^{(l)}\\ &-\sum_{j=1}^s \mu_j d_j \Big|\; (z,\mu) \in \prod_{i=1}^m \text{SOC}_k \times \mathbb{R}^s\Big\},\end{aligned} \tag{2.2}$$

where $z := (z_1, \dots, z_m) \in \mathbb{R}^{(k+1)m}$ *with* $z_i := (z_i^{(0)}, z_i^{(1)}, \dots, z_i^{(k)}) \in \mathbb{R}^{(k+1)}$ *as* $i = 1, \dots, m$*, and where* $\mu := (\mu_1, \dots, \mu_s) \in \mathbb{R}^s$.

Proof. First observe the relationships

$$
\begin{aligned}
&\max_{\substack{(a_i,\beta_i)\in\overline{\mathcal{U}_i}\\ \lambda_i\ge 0}} \inf_{x\in L^2[0,1]} \Big\{ f(x) + \sum_{i=1}^m \lambda_i(\langle a_i, x\rangle - \beta_i) + \sum_{j=1}^s \mu_j(\langle h_j, x\rangle - d_j)\Big\}\\
=\ &\max_{\substack{(a_i,\beta_i)\in\overline{\mathcal{U}_i}\\ \lambda_i\ge 0}} \Big\{ -f^*\Big(-\sum_{i=1}^m \lambda_i a_i - \sum_{j=1}^s \mu_j h_j\Big) - \sum_{i=1}^m \lambda_i\beta_i - \sum_{j=1}^s \mu_j d_j\Big\}\\
=\ &\max_{\substack{\|(w_i^{(1)},\ldots,w_i^{(k)})\|\le 1\\ \lambda_i\ge 0}} \Big\{ -f^*\Big(-\sum_{i=1}^m \lambda_i\big(\overline{a}_i^{(0)} + \sum_{l=1}^k w_i^{(l)}\overline{a}_i^{(l)}\big) - \sum_{j=1}^s \mu_j h_j\Big)\\
-\ &\sum_{i=1}^m \lambda_i\big(\overline{\beta}_i^{(0)} + \sum_{l=1}^k w_i^{(l)}\overline{\beta}_i^{(l)}\big) - \sum_{j=1}^s \mu_j d_j\Big\},
\end{aligned}
$$

where f^* denotes the conjugate function of a convex function f, i.e.,

$$
f^*(x^*) := \sup_{x\in L^2[0,1]} \big\{\langle x^*, x\rangle - f(x)\big\},\quad x^*\in L^2[0,1].
$$

Letting $z_i := (z_i^{(1)},\ldots,z_i^{(k)})$ with $z_i^{(l)} := \lambda_i w_i^{(l)}$ as $l = 1,\ldots,k$ and $z_i^{(0)} := \lambda_i \ge 0$ as $i = 1,\ldots,m$, we get the equivalence

$$
\big[\|(w_i^{(1)},\ldots,w_i^{(k)})\|_{\mathbb{R}^k} \le 1,\ \lambda_i\ge 0\big] \iff \|(z_i^{(1)},\ldots,z_i^{(k)})\|_{\mathbb{R}^k} \le z_i^{(0)}.
$$

Thus the optimistic counterpart of the dual problem (ODSP) can be expressed as the following convex program in the finite-dimensional space $\mathbb{R}^{m(k+1)}$:

$$
\begin{aligned}
\max_{z_i^{(l)}\in\mathbb{R},\mu_j\in\mathbb{R}}\quad & -f^*\Big(-\sum_{i=1}^m\sum_{l=0}^k z_i^{(l)}\overline{a}_i^{(l)} - \sum_{j=1}^s \mu_j h_j\Big) - \sum_{i=1}^m\sum_{l=0}^k z_i^{(l)}\overline{\beta}_i^{(l)} - \sum_{j=1}^s \mu_j d_j\\
\text{s.t.}\quad & \|(z_i^{(1)},\ldots,z_i^{(k)})\|_{\mathbb{R}^k} \le z_i^{(0)},\ i = 1,\ldots,m.
\end{aligned}
$$

Noting that $f^*(y) = \frac{1}{2}\|y\|^2$ for each $y\in L^2[0,1]$, we can describe this problem as

$$
\begin{aligned}
\max_{z_i^{(l)}\in\mathbb{R},\mu_j\in\mathbb{R}}\quad & -\frac{1}{2}\int_0^1 \Big(\sum_{i=1}^m\sum_{l=0}^k z_i^{(l)}\overline{a}_i^{(l)}(t) + \sum_{j=1}^s \mu_j h_j\Big)^2 dt - \sum_{i=1}^m\sum_{l=0}^k z_i^{(l)}\overline{\beta}_i^{(l)} - \sum_{j=1}^s \mu_j d_j\\
\text{s.t.}\quad & \|(z_i^{(1)},\ldots,z_i^{(k)})\|_{\mathbb{R}^k} \le z_i^{(0)},\ i = 1,\ldots,m,
\end{aligned}
$$

which is the second-order cone programming problem (2.2) stated in the theorem. □

The next result provides a precise representation of the (unique) solution to the original robust optimization problem (RSP) via the corresponding optimal solution to the optimistic dual counterpart (ODSP).

Theorem 2.3 (Representation of Robust Optimal Solution). *Consider the dual pair* (RSP) *and* (ODSP). *Assume that the strict feasibility condition of Theorem* 2.1 *is satisfied. Then the unique robust solution* $x^* \in L^2[0,1]$ *to problem* (RSP) *is given by*

$$x^* = -\sum_{i=1}^{m}\sum_{l=0}^{k} \overline{z}_i^{(l)} \overline{a}_i^{(l)} - \sum_{j=1}^{s} \overline{\mu}_j h_j, \tag{2.3}$$

where $(\overline{z}, \overline{\mu}) \in \mathbb{R}^{m(k+1)} \times \mathbb{R}^s$ *is the corresponding optimal solution to the optimistic dual problem* (ODSP) *in the equivalent second-order programming form* (2.2).

Proof. By the *strong duality* of Theorem 2.1 and the equivalent description of the dual problem (ODSP) from Theorem 2.2, it remains to show that the function x^* from (2.3) is *feasible* to (RSP) and that

$$\frac{1}{2}\|x^*\|^2 = \max(\text{ODSP}). \tag{2.4}$$

To verify feasibility, denote by

$$g(z,\mu) := -\frac{1}{2}\int_0^1 \Big(\sum_{i=1}^{m}\sum_{l=0}^{k} z_i^{(l)}\overline{a}_i^{(l)}(t) + \sum_{j=1}^{s}\mu_j h_j\Big)^2 dt - \sum_{i=1}^{m}\sum_{l=0}^{k} z_i^{(l)}\overline{\beta}_i^{(l)} - \sum_{j=1}^{s}\mu_j d_j$$

the objective function of problem (ODSP) in the equivalent form (2.2) with the notation of Theorem 2.2. Denoting further by $(\overline{z},\overline{\mu})$ an optimal solution to problem (ODSP) in form (2.2) and using well-known necessary and sufficient optimality conditions in convex programming under the constraint qualification/strict feasibility condition (2.1) imposed in the theorem (see. e.g., [35, 13]), we have the relationships

$$-\nabla_z g(\overline{z},\overline{\mu}) \in \prod_{i=1}^{m} \text{SOC}_k,\ \nabla_\mu g(\overline{z},\overline{\mu}) = 0, \text{ and } \langle \overline{z}, \nabla_z g(\overline{z},\overline{\mu})\rangle = 0, \tag{2.5}$$

where $\nabla_z g$ (resp. $\nabla_\mu g$) denotes the partial derivative of g with respect to z (resp. μ). Observe also that for each $(a_i,\beta_i) \in \overline{\mathcal{U}}_i$ there is $w_i \in \mathbb{R}^k$ with $\|w_i\|_{\mathbb{R}^k} \le 1$ such that

$$(a_i,\beta_i) = (\overline{a}_i^{(0)}, \overline{\beta}_i^{(0)}) + \sum_{l=1}^{k} w_i^{(l)}(\overline{a}_i^{(l)}, \overline{\beta}_i^{(l)}).$$

Furthermore, the inclusion in (2.5) is equivalent to the inequalities

$$-\big(\langle \overline{a}_i^{(0)}, x^*\rangle - \overline{\beta}_i^{(0)}\big) \ge \| - \big(\langle \overline{a}_i^{(1)}, x^*\rangle - \overline{\beta}_i^{(1)}, \ldots, \langle \overline{a}_i^{(k)}, x^*\rangle - \overline{\beta}_i^{(k)}\big)\|_{\mathbb{R}^k}$$

for each $i = 1,\ldots,m$. Thus we get that

$$\begin{aligned}\langle a_i, x^*\rangle - \beta_i &= \big(\langle \overline{a}_i^{(0)}, x^*\rangle - \overline{\beta}_i^{(0)}\big) + \sum_{l=1}^{k} w_i^{(l)}\big(\langle \overline{a}_i^{(l)}, x^*\rangle - \overline{\beta}_i^{(l)}\big) \\ &\le \big(\langle \overline{a}_i^{(0)}, x^*\rangle - \overline{\beta}_i^{(0)}\big) + \|\big(\langle \overline{a}_i^{(1)}, x^*\rangle - \overline{\beta}_i^{(1)}, \ldots, \langle \overline{a}_i^{(k)}, x^*\rangle - \overline{\beta}_i^{(k)}\big)\|_{\mathbb{R}^k} \le 0.\end{aligned}$$

It follows from $\nabla_\mu g(\overline{z},\overline{\mu}) = 0$ in (2.5) that $\langle h_j, x^*\rangle - d_j = 0$ as $j = 1,\ldots,s$, which ensures that the element x^* from (2.3) is feasible to the robust optimization problem (RSP). Finally, we verify that formula (2.4) holds. Indeed, the complementary slackness condition $\langle \overline{z}, \nabla_z g(\overline{z},\overline{\mu})\rangle = 0$ in (2.5) gives us that $\sum_{i=1}^{m}\sum_{l=0}^{k} \overline{z}_i^l (\langle \overline{a}_i^{(l)}, x^*\rangle - \overline{\beta}_i^{(l)}) = 0$. Together with the equalities $\langle h_j, x^*\rangle = d_j$ for all $j = 1,\ldots,s$, this implies that

$$
\begin{aligned}
\max(\text{ODSP}) &= -\frac{1}{2}\int_0^1 \Big(\sum_{i=1}^{m}\sum_{l=0}^{k} \overline{z}_i^{(l)}\overline{a}_i^{(l)}(t) + \sum_{j=1}^{s}\overline{\mu}_j h_j\Big)^2 dt - \sum_{i=1}^{m}\sum_{l=0}^{k} \overline{z}_i^{(l)}\overline{\beta}_i^{(l)} - \sum_{j=1}^{s}\overline{\mu}_j d_j \\
&= -\frac{1}{2}\|x^*\|^2 + \langle -\sum_{i=1}^{m}\sum_{l=0}^{k} \overline{z}_i^{(l)}\overline{a}_i^{(l)}, x^*\rangle - \sum_{j=1}^{s}\overline{\mu}_j\langle h_j, x^*\rangle \\
&= -\frac{1}{2}\|x^*\|^2 + \langle x^*, x^*\rangle = \frac{1}{2}\|x^*\|^2,
\end{aligned}
$$

which thus completes the proof of the theorem. □

To illustrate the application of our robust duality result and its realization in the solution representation of Theorem 2.3, we consider the following numerical example.

Example 2.4 (Realization of Duality Relationships). Consider the best approximation problem of the (SP) type under ellipsoidal uncertainty:

$$
\min_{x\in L^2[0,1]} \Big\{\frac{1}{2}\int_0^1 x(t)^2 dt \ \text{ s.t. } \int_0^1 (\alpha t + t^2)x(t)dt \le -1,\ \int_0^1 x(t)dt = 1\Big\},
$$

where the equality $\int_0^1 x(t)dt = 1$ is a normalization condition and so is free of uncertainty while the parameter $\alpha \in \mathbb{R}$ is uncertain with $\alpha \in [-0.5, 0.5]$. In other words, the data $a_1(t) = \alpha t + t^2$ is uncertain and $a_1(t)$ belongs to the ellipsoidal uncertainty set

$$
\overline{\mathcal{U}}_1 := \Big\{(t^2, -1) + w_1^{(1)}\Big(\frac{t}{2}, 0\Big)\Big|\ |w_1^{(1)}| \le 1\Big\}.
$$

The corresponding robust best approximation problem (RSP) is as follows:

$$
\min_{x\in L^2[0,1]} \Big\{\frac{1}{2}\int_0^1 x(t)^2 dt\Big|\ \int_0^1 a_1(t)x(t)dt \le \beta_1,\ \forall (a_1(t),\beta_1) \in \overline{\mathcal{U}}_1,\ \int_0^1 x(t)dt = 1\Big\}.
$$

Its dual counterpart (ODSP) is given, by Theorem 2.2, in the form:

$$
\begin{aligned}
&\max_{\substack{(z_1^{(0)}, z_1^{(1)})\in\mathbb{R}^2 \\ \mu\in\mathbb{R}}} \Big\{ -\frac{1}{2}\int_0^1 \Big(z_1^{(0)}t^2 + z_1^{(1)}\frac{t}{2} + \mu\Big)^2 dt + z_1^{(0)} - \mu\Big|\ |z_1^{(1)}| \le z_1^{(0)}\Big\} \\
= &\max_{\substack{(z_1^{(0)}, z_1^{(1)})\in\mathbb{R}^2 \\ \mu\in\mathbb{R}}} \Big\{ -\frac{1}{10}\Big(z_1^{(0)}\Big)^2 - \frac{1}{24}\Big(z_1^{(1)}\Big)^2 - \frac{1}{8}z_1^{(0)}z_1^{(1)} - \frac{1}{4}z_1^{(1)}\mu \\
&-\frac{1}{3}z_1^{(0)}\mu + z_1^{(0)} - \frac{1}{2}\mu^2 - \mu\Big|\ |z_1^{(1)}| \le z_1^{(0)}\Big\}.
\end{aligned}
$$

Let $z^* = (z_0^*, z_1^*) = (\frac{780}{19}, -\frac{780}{19})$, $\mu^* = -\frac{84}{19}$, and

$$g(z,\mu) := -\frac{1}{10}(z_1^{(0)})^2 - \frac{1}{24}(z_1^{(1)})^2 - \frac{1}{8}z_1^{(0)}z_1^{(1)} - \frac{1}{4}z_1^{(1)}\mu - \frac{1}{3}z_1^{(0)}\mu + z_1^{(0)} - \frac{1}{2}\mu^2 - \mu,$$

where $z = (z_1^{(0)}, z_1^{(1)}) \in \mathbb{R}^2$ and $\mu \in \mathbb{R}$. It is easy to check that $\nabla_z g(z^*, \mu^*) = (-\frac{23}{38}, -\frac{23}{38})$ and $\nabla_\mu g(z^*, \mu^*) = 0$; so we have

$$\langle \nabla_z g(z^*, \mu^*), z^* \rangle = 0 \ \text{ and } \ -\nabla_z g(z^*, \mu^*) \in \mathrm{SOC}_1 \times \{0\},$$

i.e., the KKT conditions (2.5) are satisfied. This confirms by the optimality conditions in convex programming that (z^*, μ^*) is an optimal solution to (ODSP) with the maximum value

$$\max\,(\mathrm{ODSP}) = g(z^*, \mu^*) = \frac{432}{19}.$$

Now we are going to apply Theorem 2.1 on strong duality in order to determine an optimal solution to the robust optimization problem (RSP). Check first that the strict feasibility condition of this theorem is satisfied. Indeed, it holds for the function $x_0(t) = 61 - 120t$ as $t \in [0,1]$. To proceed with applying Theorem 2.1, it is sufficient to find a function $x^* \in L^2[0,1]$, which is feasible to (RSP) and such that

$$\frac{1}{2}\int_0^1 \big(x^*(t)\big)^2 dt = \frac{432}{19} = \big(\max\,(\mathrm{ODSP})\big). \tag{2.6}$$

Define the following one

$$x^*(t) := -\Big(z_0^* t^2 + z_1^* \frac{t}{2}\Big) - \mu^* = -\frac{780}{19}t^2 + \frac{390}{19}t + \frac{84}{19} \ \text{ for all } \ t \in [0,1]$$

and calculate directly that it satisfies the first equality in (2.6). The latter ensures that

$$\int_0^1 (\alpha t + t^2) x^*(t) dt = -\frac{23\alpha}{19} - \frac{61}{38} \le -1$$

whenever $\alpha \in [-0.5, 0.5]$. Observing finally that $\int_0^1 x^*(t)dt = 1$, we conclude that x^* is feasible to (RSP) and thus it is optimal for this problem by Theorem 2.1.

3 Solving the Dual Problem

Example 2.4 illustrates, in a specific setting, how to use the strong duality of Theorem 2.1 for determining the optimal value and even the optimal solution to the infinite-dimensional robust optimization/best approximation problem via solving its dual counterpart, which is a standard problem of convex optimization in finite dimensions. However, the above example does not provide any hint on how to proceed in the general setting. From now on we employ the established strong duality to find the optimal value of the primal infinite-dimensional problem (RSP) by developing constructive *numerical algorithms* for solving the dual finite-dimensional optimization problem (ODSP). This section concerns the first step of our approach and shows that solving the dual problem (ODSP) is equivalent to solving a certain *nonsmooth equation* (in fact a *strongly semismooth* one). To proceed, consider the second-order cone programming form of (ODSP) described in Theorem 2.2 and introduce, via the initial data of this problem, the linear mappings $A_i : L^2[0,1] \mapsto \mathbb{R}^{k+1}$ as $i = 1, \ldots, m$ and $B : L^2[0,1] \mapsto \mathbb{R}^s$ by

$$A_i(x) := \left(\int_0^1 \overline{a}_i^{(0)}(t)x(t)dt, \ldots, \int_0^1 \overline{a}_i^{(k)}(t)x(t)dt \right),$$

$$B(x) := \left(\int_0^1 h_1(t)x(t)dt, \ldots, \int_0^1 h_s(t)x(t)dt \right).$$

Let $b_i = (\overline{\beta}_i^{(0)}, \overline{\beta}_i^{(1)}, \ldots, \overline{\beta}_i^{(k)}) \in \mathbb{R}^{k+1}$ as $i = 1, \ldots, m$, and let $d = (d_1, \ldots, d_s) \in \mathbb{R}^s$. Define the linear mappings $F_i : \mathbb{R}^{m(k+1)} \times \mathbb{R}^s \to \mathbb{R}^{k+1}$, $i = 1, \ldots, m$, and $H : \mathbb{R}^{m(k+1)} \times \mathbb{R}^s \to \mathbb{R}^s$ by

$$F_i(z, \mu) := A_i \Bigg(\sum_{r=1}^{m} \sum_{l=0}^{k} z_r^{(l)} \overline{a}_r^{(l)} + \sum_{j=1}^{s} \mu_j h_j \Bigg) + b_i,$$

$$H(z, \mu) := B \Bigg(\sum_{r=1}^{m} \sum_{l=0}^{k} z_r^{(l)} \overline{a}_r^{(l)} + \sum_{j=1}^{s} \mu_j h_j \Bigg) + d \tag{3.1}$$

and then $G : \mathbb{R}^{(k+1)m} \times \mathbb{R}^s \times \mathbb{R}^{(k+1)m} \to \mathbb{R}^{(k+1)m} \times \mathbb{R}^s \times \mathbb{R}^{(k+1)m}$ by

$$G(z, \mu, v) = \begin{pmatrix} F_1(z, \mu) - v_1 \\ \cdots \\ F_m(z, \mu) - v_m \\ H(z, \mu) \\ z_1 - \mathrm{P}_{\mathrm{SOC_k}}(z_1 - v_1) \\ \cdots \\ z_m - \mathrm{P}_{\mathrm{SOC_k}}(z_m - v_m) \end{pmatrix}, \tag{3.2}$$

where $z_i = (z_i^{(0)}, z_i^{(1)}, \ldots, z_i^{(k)}) \in \mathbb{R}^{k+1}$ as $i = 1, \ldots, m$, and where $\mathrm{P}_{\mathrm{SOC_k}}(z)$ stands for the unique Euclidean projection of a vector z on the second-order cone $\mathrm{SOC_k}$. We show below that the mapping G in (3.2) enjoys the strong semismooth property that plays a crucial role in establishing the quadratic convergence rate for the Newton-type algorithms developed below. Recall (see [31] and the books [9, 18] for more details on the constructions involved) that a mapping $F : \mathbb{R}^p \to \mathbb{R}^q$ between finite-dimensional spaces is *semismooth* at $u \in \mathbb{R}^p$

if the following conditions hold: **(i)** F is locally Lipschitzian around u and directionally differentiable at this point; **(ii)** $F(u+h)-F(u)-V^T h = o(\|h\|)$ for each $V \in \partial_B F(u+h)$ and $h \in \mathbb{R}^p$ with the norm $\|h\|$ sufficiently small, where $\partial_B F(\overline{x})$ is the so-called *B-subdifferential* of F at $\overline{x}$ defined by

$$\partial_B F(\overline{x}) := \Big\{ \lim_{\nu\to\infty} \nabla F(x_\nu) \Big|\, F \;\text{ is differentiable at }\; x_\nu \to \overline{x} \Big\}, \tag{3.3}$$

and where the symbol ^{T} stands as usual for the matrix transposition. It is well-known by the classical Rademacher theorem that the set (3.3) is a nonempty compact set in $\mathbb{R}^{p\times q}$. Furthermore, if we have $F(u+h)-F(u)-V^T h = O(\|h\|^2)$ for each $V \in \partial_B F(u+h)$ and $h \in \mathbb{R}^p$ with $\|h\|$ sufficiently small, then F is *strongly semismooth* at u. Finally, if F is semismooth (resp. strongly semismooth) everywhere on $\mathbb{R}^p$, then F is simply called semismooth (resp. strongly semismooth) with no point indication. In what follow, we consider the *nonsmooth equation* given by

$$G(z,\mu,v) = 0 \tag{3.4}$$

where G is defined as in (3.2), which is clearly nondifferentiable due to the projection operator involved. Let us show that solving the dual problem (ODSP) is in fact *equivalent* to solving the nonsmooth equation (3.4).

Proposition 3.1 (Reducing (ODSP) to Nonsmooth Equations). *Let (z^*,μ^*) be a feasible point of the dual problem* (ODSP)*. Then (z^*,μ^*) is an optimal solution to* (ODSP) *if and only if the triple $(z^*,\mu^*,F(z^*,\mu^*))$ is a solution to the nonsmooth equation* (3.4)*, where $F(z,\mu) := (F_1(z,\mu),\ldots,F_m(z,\mu))$.*

Proof. Let $z := (z_1^{(0)},\ldots,z_1^{(k)},\ldots,z_m^{(0)},\ldots,z_m^{(k)}) \in \mathbb{R}^{m(k+1)}$ and $\mu := (\mu_1,\ldots,\mu_s) \in \mathbb{R}^s$. Then the objective function of (ODSP) is given by

$$g(z,\mu) \;=\; -\frac{1}{2}\int_0^1 \Big(\sum_{i=1}^m \sum_{l=0}^k z_i^{(l)} \overline{a}_i^{(l)}(t) + \sum_{j=1}^s \mu_j h_j \Big)^2 dt - \sum_{i=1}^m \sum_{l=0}^k z_i^{(l)} \overline{\beta}_i^{(l)} - \sum_{j=1}^s \mu_j d_j. \tag{3.5}$$

The pair (z^*,μ^*) is a solution to (ODSP) if and only if (z^*,μ^*) maximizes the concave function g over the convex set $\prod_{i=1}^m \mathrm{SOC}_k \times \mathbb{R}^s$. Employing the well-known optimality conditions in convex programming (see, e.g., [35]) in the setting under consideration, we get that (z^*,μ^*) is an optimal solution to (ODSP) if and only if it satisfies the relationships

$$\begin{cases} z_i \in \mathrm{SOC}_k, \; i = 1,\ldots,m, \\ -\nabla_z g(z,\mu) \in \prod_{i=1}^m \mathrm{SOC}_k, \\ \nabla_z g(z,\mu)^T z = 0, \; \nabla_\mu g(z,\mu) = 0, \end{cases} \tag{3.6}$$

where $z_i := (z_i^{(0)}, z_i^{(1)},\ldots,z_i^{(k)}) \in \mathbb{R}^{k+1}$. Note that

$$\nabla_z g(z,\mu) = -\big(F_1(z,\mu)^T,\ldots,F_m(z,\mu)^T\big)^T \;\text{ and }\; \nabla_\mu g(z,\mu) = -H(z,\mu)$$

for each $(z, \mu) \in \mathbb{R}^{m(k+1)} \times \mathbb{R}^s$. Taking into account the constructions above, relationships (3.6) can be equivalently be rewritten as

$$\begin{cases} F_i(z,\mu) - v_i = 0, \quad i = 1, \ldots, m, \\ H(z,\mu) = 0, \\ z_i \in \mathrm{SOC}_k, \quad i = 1, \ldots, m, \\ v_i \in \mathrm{SOC}_k, \quad i = 1, \ldots, m, \\ \langle v_i, z_i \rangle = 0, \quad i = 1, \ldots, m, \end{cases} \tag{3.7}$$

Furthermore, by [10, Lemma 2.1] we have the equivalence

$$u \in \mathrm{SOC}_k,\ v \in \mathrm{SOC} \text{ and } \langle u, v \rangle = 0 \iff u - \mathrm{P}_{\mathrm{SOC}_k}(u - v) = 0,$$

which allows us to represent system (3.7) in the equivalent form of the nonsmooth equation (3.4) with $v_i = F_i(z, \mu)$. Thus (z^*, μ^*) is an optimal solution to (ODSP) if an only if the triple $(z^*, \mu^*, F(z^*, \mu^*))$ solves (3.4). □

Let us finally show that the mapping G in (3.2) defining the nonsmooth equation (3.4) is actually *strongly semismooth* on the whole space.

Theorem 3.2 (Strong Semismoothness of the Nonsmooth Equation). *The mapping G defined in* (3.2) *is strongly semismooth.*

Proof. Observe first that any linear mapping is strongly semismooth. Thus the mappings $(z, \mu, v) \mapsto F_i(z, \mu) - v_i$ and $(z, \mu, v) \mapsto H(z, \mu)$ are all strongly semismooth. Furthermore, it is known that the projection mapping $\mathrm{P}_{\mathrm{SOC_k}}$ on the second-order cone is strongly semismooth; see [10, Proposition 4.5]. This ensures that the mapping $(z, \mu, v) \mapsto z_i - \mathrm{P}_{SOC_k}(z_i - v_i)$ is strongly semismooth for each $i = 1, \ldots, m$. Since the strong semismoothness of a vector-valued mapping is obviously equivalent to this property for each of its components, we conclude that G in (3.2) is strongly semismooth. □

4 Robust Newton Methods: Well-Posedness and Local Convergence

We now present an approach for solving the dual problem (ODSP), and hence the primal robust optimization problem (RSP), by developing an appropriate version of the generalized Newton method to find a root of the equivalent nonsmooth equation (3.4). We pay main attention to the development of the so-called *B-differentiable* version of the nonsmooth Newton method initiated in [27] and essentially improved later in [30] and recently in [12] in the framework of abstract nonsmooth equations. Our goal here is to design a Newton-type algorithm for solving the robust problem (RSP) via its dual counterpart (ODSP) and the equivalent nonsmooth equation (3.4) and to establish suitable conditions for its local quadratic convergence. We also briefly discuss some other Newton-type algorithms for solving the problems under consideration. Recall that the general scheme of the B-differentiable Newton method to solve the nonsmooth equation $G(u) = 0$ described by an arbitrary directionally differentiable mapping $G\colon \mathbb{R}^p \to \mathbb{R}^p$ is given in the iterative form

$$u_{n+1} := u_n + d_n, \ n = 0, 1, \ldots, \tag{4.1}$$

where $d = d_n$ is a solution of the *generalized Newton equation*

$$-G(u_n) = G'(u_n; d). \tag{4.2}$$

We refer the reader to [9, 27, 28] for the origin and history of the B-differentiable Newton method and the terminology behind. As discussed in [30, pp. 243–244], the generalized Newton equation (4.2) *may not be solvable* for strongly semismooth (and hence directionally differentiable and Lipschitz continuous) mappings G even when each matrix belonging to the B-subdifferential (3.3) of G around the solution point is nonsingular; see the example therein. This means that algorithm (4.1) with (4.2) may *not be well defined*. General conditions for the solvability of (4.2) and its far-going extension via graphical derivatives have been recently obtained in [12] by using the Mordukhovich coderivative criterion for the *metric regularity* of the underlying mapping G in (4.2); see [25, 29]. Recall that a mapping $F : \mathbb{R}^p \to \mathbb{R}^p$ is said to be metrically regular at $\overline{x}$ if there are neighborhoods U of $\overline{x}$ and V of $\overline{y} = F(\overline{x})$ as well as a number $\mu > 0$ such that

$$\text{dist}(x, F^{-1}(y)) \le \mu \text{dist}(y, F(x)) \text{ for all } x \in U \text{ and } y \in V.$$

An equivalent condition ensuring the above metrically regular condition is that

$$0 \in \partial_M \langle v, F \rangle(\overline{x}) \Longrightarrow v = 0 \tag{4.3}$$

where ∂_M is the *basic/limiting subdifferential* introduced by Mordukhovich; see [26]. It is known that

$$\partial_B F(\overline{x})^T v \subseteq \partial_M \langle v, F \rangle(\overline{x}) \subseteq \partial_C F(\overline{x})^T v \text{ for all } v \in \mathbb{R}^p, \tag{4.4}$$

where ∂_C is the *generalized Jacobian* introduced by Clarke (defined as the convex hull of the B-subdifferential (3.3) for Lipschitz continuous mappings). In general, both inclusions in (4.4) can be strict (see [12]). Therefore, metric regularity of F at $\overline{x}$ is stronger than the requirement that each matrix belonging to the B-subdifferential (3.3) of G around the solution point is nonsingular (which is not sufficient for the solvability of (4.2)) and is weaker than the requirement that each matrix belonging to the Clarke subdifferential (3.3) of G around the solution point is nonsingular. We first recall a recent convergence result of generalized Newton method in terms of metric regularity produced in [12].

Lemma 4.1. *Let $F : \mathbb{R}^p \to \mathbb{R}^p$ be strongly semismooth, one-to-one, and metrically regular around a solution $\overline{x}$. Then the generalized Newton method (4.1), (4.2) is well defined (meaning that equation (4.2) is solvable) and converges quadratically to the solution $\overline{x}$.*

We now derive the solvability and quadratic convergence of the generalized Newton method.

Theorem 4.2 (Well-Posedness & Quadratic Convergence). *Let $\overline{u} = (\overline{z}, \overline{\mu}, \overline{v}) \in \mathbb{R}^{(k+1)m} \times \mathbb{R}^s \times \mathbb{R}^{(k+1)m}$ be a solution of the nonsmooth equation $G(u) = 0$ with G defined in* (3.2)*, and suppose that G is one-to-one, and metrically regular around $\overline{u}$. Then there exists a neighborhood Ω of $\overline{u}$ such that for any starting point $u_0 \in \Omega$ the sequence $u_n = (z^n, \mu^n, v^n)$ generated by* (4.1) *is well defined and quadratically converges to $\overline{u}$. Moreover, the sequence of functions*

$$x_n(t) := -\sum_{i=1}^{m}\sum_{l=0}^{k}(z^n)_i^l \overline{a}_i^{(l)}(t) - \sum_{j=1}^{s}(\mu^n)_j h_j(t), \quad t \in [0,1], \tag{4.5}$$

with $z^n = \big((z^n)_1^0, \dots, (z^n)_1^k, \dots, (z^n)_m^0, \dots, (z^n)_m^k\big)$ and $\mu^n = \big((\mu^n)_1, \dots, (\mu^n)_s\big)$ as $n \in \{0\} \cup \mathbb{N}$ strongly L^2-converges to a robust solution $x^* \in L^2[0,1]$ of problem (RSP).

Proof. As proved in Theorem 3.2, the mapping G from (3.2) is strongly semismooth. Thus the well defineness and the quadratic convergence of the generalized Newton method follows from the preceding lemma. Moreover, due to $u_n \to \overline{u}$ we have $z^n \to \overline{z}$ and $\mu^n \to \overline{\mu}$ as $n \to \infty$. Define

$$x^*(t) := -\sum_{i=1}^{m}\sum_{l=0}^{k}\overline{z}_i^l \overline{a}_i^{(l)}(t) - \sum_{j=1}^{s}\overline{\mu}_j h_j(t), \quad t \in [0,1], \tag{4.6}$$

with $\overline{z} = \big(\overline{z}_1^0, \dots, \overline{z}_1^k, \dots, \overline{z}_m^0, \dots, \overline{z}_m^k\big) \in \mathbb{R}^{m(k+1)}$ and $\overline{\mu} = (\overline{\mu}_1, \dots, \overline{\mu}_s) \in \mathbb{R}^s$. It follows from (4.5) and (4.6) that

$$\|x_n - x^*\| \le \|\sum_{i=1}^{m}\sum_{l=0}^{k}\big((z^n)_i^l - \overline{z}_i^l\big)\overline{a}_i^{(l)}\| + \|\sum_{j=1}^{s}((\mu^n)_j - \overline{\mu}_j)h_j\| \to 0,$$

where the norm is taken in $L^2[0,1]$. This allows us to conclude by Theorem 2.3 that x^* is an optimal solution to the robust optimization problem (RSP). □

In general, our condition for quadratic convergence that G is one-to-one and metric regular near $\overline{x}$ is strictly weaker than the requirement that each matrix belonging to the Clarke subdifferential (3.3) of G around the solution point is nonsingular. Note that our dual problem (2.2) automatically satisfies the Slater condition (and so also satisfies the Robinson constraint qualification). As shown in [33], the requirement that each matrix belonging to the Clarke subdifferential (3.3) of G around the solution point is nonsingular is equivalent to the primal constraint nondegeneracy and strong second order sufficient condition (which was used as a sufficient condition ensuring the quadratic convergence of a sequential quadratic programming method for solving a nonlinear second-order cone programming problem in [33]).

Remark 4.3 (**Other Versions of the Robust Newton Method).** The above constructions and results, employing iterations (4.1), (4.2) of the B-differentiable Newton algorithm to solve the robust optimization problem (RSP), can be developed in other frameworks of the generalized Newton method for nonsmooth equations. The so-called *semismooth Newton method* for nonsmooth equations $G(u) = 0$ is written in the form

$$u_{n+1} := u_n - V_n^{-1} G(u_n),\ V_n \in \partial G(u_n), \quad n = 0, 1, \ldots, \tag{4.7}$$

where ∂G stands for either B-subdifferential $\partial_B G$ defined in (3.3) or for the generalized Jacobian $\partial_C G$ provided that all the matrices from the corresponding sets are nonsingular; see [9, 30, 31] for more details on (4.7) and [12] for recent extensions. Moreover, it is also possible to develop smoothing type Newton method for solving the dual problem as in [34].

Remark 4.4. (Global Convergence of the Kantorovich type) We conclude this section by pointing out that, for the robust Newton algorithms of solving (RSP) discussed above, we can also establish the so-called *global convergence* of the *Kantorovich type* [12, 22, 9, 27, 28] in terms of the metric regularity assumption.

As an illustration, we briefly outline how our robust Newton method may be applied to solve the dual problem of the robust best approximation problem considered in Example 2.4. Further numerical experiment will be carried out in a forthcoming study.

Example 4.5. Consider the best approximation problem under ellipsoidal uncertainty that was considered in Example 2.4:

$$\min_{x \in L^2[0,1]} \Big\{ \frac{1}{2} \int_0^1 x(t)^2 dt \ \text{ s.t. } \ \int_0^1 a_1(t) x(t) dt \le -1,\ \int_0^1 x(t) dt = 1 \Big\}.$$

The corresponding robust best approximation problem (RSP) is as follows:

$$\min_{x \in L^2[0,1]} \Big\{ \frac{1}{2} \int_0^1 x(t)^2 dt \Big| \ \int_0^1 a_1(t) x(t) dt \le \beta_1,\ \forall (a_1(t), \beta_1) \in \overline{\mathcal{U}}_1,\ \ \int_0^1 x(t) dt = 1 \Big\},$$

where $\overline{\mathcal{U}}_1$ is an ellipsoidal uncertainty set given by

$$\overline{\mathcal{U}}_1 := \Big\{ (t^2, -1) + w_1^{(1)} \Big(\frac{t}{2}, 0 \Big) \Big| \ |w_1^{(1)}| \le 1 \Big\}.$$

Its dual counterpart (ODSP) is given, by Theorem 2.2, in the form:

$$
\begin{aligned}
& \max_{\substack{z=(z_1^{(0)},z_1^{(1)})\in\mathbb{R}^2 \\ \mu\in\mathbb{R}}} \Big\{ -\frac{1}{2}\int_0^1 \Big(z_1^{(0)}t^2 + z_1^{(1)}\frac{t}{2} + \mu\Big)^2 dt + z_1^{(0)} - \mu \Big| \, |z_1^{(1)}| \le z_1^{(0)} \Big\} \\
= \; & \max_{\substack{z=(z_1^{(0)},z_1^{(1)})\in\mathbb{R}^2 \\ \mu\in\mathbb{R}}} \Big\{ g(z,\mu) \Big| \, |z_1^{(1)}| \le z_1^{(0)} \Big\}.
\end{aligned}
$$

where

$$
g(z,\mu) := -\frac{1}{10}(z_1^{(0)})^2 - \frac{1}{24}(z_1^{(1)})^2 - \frac{1}{8}z_1^{(0)}z_1^{(1)} - \frac{1}{4}z_1^{(1)}\mu - \frac{1}{3}z_1^{(0)}\mu + z_1^{(0)} - \frac{1}{2}\mu^2 - \mu.
$$

It can be verifed that

$$
\nabla_z g(z,\mu) = (-\frac{1}{5}z_1^{(0)} - \frac{1}{8}z_1^{(1)} - \frac{1}{3}\mu + 1, -\frac{1}{12}z_1^{(1)} - \frac{1}{8}z_1^{(0)} - \frac{1}{4}\mu) \text{ and } \nabla_\mu g(z,\mu) = -\frac{1}{4}z_1^{(1)} - \frac{1}{3}z_1^{(0)} - \mu - 1.
$$

Then, the corresponding equivalent nonsmooth equation for the optimistic counterpart (ODSP) is $G(z,\mu,v) = 0$ where

$$
G(z,\mu,v) = \begin{pmatrix} -\frac{1}{5}z_1^{(0)} - \frac{1}{8}z_1^{(1)} - \frac{1}{3}\mu + 1 - v_1^{(0)} \\ -\frac{1}{12}z_1^{(1)} - \frac{1}{8}z_1^{(0)} - \frac{1}{4}\mu - v_1^{(1)} \\ -\frac{1}{4}z_1^{(1)} - \frac{1}{3}z_1^{(0)} - \mu - 1 \\ z - \mathrm{Pr}_{SOC_1}(z - v) \end{pmatrix},
$$

$z = (z_1^{(0)}, z_1^{(1)})$, $v = (v_1^{(0)}, v_1^{(1)})$ and $\mathrm{SOC}_1 = \{(x_0, x_1) \in \mathbb{R}^2 : |x_1| \le x_0\}$. Then, one could apply the Newton method (4.1)-(4.2) to this nonsmooth equation and calcuate a solution for the dual problem.

5 Conclusion

In this paper we examined best approximation problems with uncertain linear inequality constraints and exact interpolation constraints by studying robust best approximation problems where the uncertain constraints are enforced for all possible uncertainties within prescribed ellipsoidal uncertainty sets. We established that finding a robust best approximation is equivalent to solving a second-order cone complementarity problem by proving a strong duality theorem under a strict feasibility condition. We provided a version of nonsmooth Newton's method and its convergence analysis in terms of the metric regularity condition. It would be of interest to investigate whether our duality approach can be employed to efficiently solve robust best approximation problems with shape-preserving interpolation constraints [21]. Computer implementation of our robust nonsmooth Newton methods for finding robust best approximations will also be of some interest and will be carried out in a forthcoming study.

Acknowledgements. This research was partially supported by the Australian Research Council under grant DP-1292508. The third author also acknowledges a partial support of the USA National Science Foundation under grant DMS-1007132. The work of the third author was partially carried out while he was visiting the School of Mathematics and Statistics at the University of New South Wales, Sydney, Australia.

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 solutions of uncertain linear programs, *Oper. Res. Lett.*, **25** (1999), 1–13.

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

[4] D. Bertsimas, D. B. Brown and C. Caramanis, Theory and applications of robust optimization, *SIAM Rev.*, **53** (2011), 464-501.

[5] J.R. Birge and F. Louveaux, *Stochastic Programming*, Springer, Berlin, 1997.

[6] F. Deutsch, *Best Approximation in Inner Product Spaces,* CMS Books in Mathematics, 7, Springer, New York, 2001.

[7] 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.

[8] L. El Ghaoui and H. Lebret, Robust solutions to least-squares problems with uncertain data, *SIAM J. Matrix Anal. Appl.*, **18** (1997), 1035–1064.

[9] F. Facchinei and J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, I/II, Springer, New York, 2003.

[10] 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), 297–320.

[11] D. Goldfarb and G. Iyengar, Robust convex quadratically constrained programs. New trends in optimization and computational algorithms, *Math. Program.*, **97** (2003), 495–515.

[12] T. Hoheisel, C. Kanzow, B.S. Mordukhovich and H. Phan, Generalized Newton's method for nonsmooth equations based on graphical derivatives, *Nonlinear Anal.*, **74** (2011) 1324–1340.

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

[14] V. Jeyakumar and G. Li, Strong duality in robust convex programming: Complete characterizations, *SIAM J. Optim.*, **20** (2010), 3384–3407.

[15] Jeyakumar and G. Li, Robust duality for fractional programming problems with constraint-wise data uncertainty, *J. Optim. Theory Appl.* (2011) 151, 292-302.

[16] 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.

[17] V. Jeyakumar, G. Li and G.M. Lee, Robust duality for generalized convex programming problems under data uncertainty, *Nonlinear Analysis.*, (2011) 74 (6), 2327-2341.

[18] V. Jeyakumar and D. T. Luc, *Nonsmooth Vector Functions and Continuous Optimization*, SOIA 10, Springer, New York, 2008.

[19] V. Jeyakumar and H. Mohebi, A global approach to nonlinearly constrained best approximation, *Numer. Funct. Anal. Optim.*, **26** (2005), 205–227.

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

[21] V. Jeyakumar, J.H. Wang and G. Li, Lagrange multiplier characterizations of robust best approximations under constraint data uncertainty, *J. Math. Anal. & Appl.* **393** (2012), 285–297.

[22] L.V. Kantorovich and G.P. Akilov, *Functional Analysis in Normed Spaces*, Macmillan, New York, 1964.

[23] Z. Lu, A computational study on robust portfolio selection based on a joint ellipsoidal uncertainty set, *Math. Program.*, **126** (2011), 193–201.

[24] C.A. Micchelli and F. Utreras, Smoothing and interpolation in a convex subset of a Hilbert space. *SIAM J. Sci. Statist. Comput.*, **9** (1988), 728–746.

[25] B.S. Mordukhovich, Complete characterizations of openness, metric regularity and Lipschitzian properties of multifubnctions, *Trans. Amer. Math. Soc.*, **340** (1993), 1–35.

[26] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 330, Springer, Berlin, 2006.

[27] J.S. Pang, Newton's method for B-differentiable functions, *Math. Oper. Res.*, **15** (1990), 311–341.

[28] S.M. Robinson, Newton's method for a class of nonsmooth functions, *Set-Valued Anal.*, **2** (1994), 291–305.

[29] R.T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 317, Springer, Berlin, 1998.

[30] L.Q. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, *Math. Oper. Res.*, **18** (1993), 227–244.

[31] L.Q. Qi and J. Sun, A nonsmooth version of Newton's method, *Math. Program.*, **58** (1993), 353–367.

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

[33] Y. Wang and L.W. Zhang, Properties of equation reformulation of the Karush-Kuhn-Tucker condition for nonlinear second order cone optimization problems. *Math. Methods Oper. Res.* **70** (2009), 195–218.

[34] Y. Wang and L.W. Zhang, Nonsingularity in second-order cone programming via the smoothing metric projector. *Sci. China Math.* **53** (2010), 1025–1038.

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