WEAK LOWER SEMI-CONTINUITY OF THE OPTIMAL VALUE FUNCTION AND APPLICATIONS TO WORST-CASE ROBUST OPTIMAL CONTROL PROBLEMS

ROLAND HERZOG AND FRANK SCHMIDT

ABSTRACT. Sufficient conditions ensuring weak lower semi-continuity of the optimal value function are presented. To this end, refined inner semi-continuity properties of set-valued maps are introduced which meet the needs of the weak topology in Banach spaces. The results are applied to prove the existence of solutions in various worst-case robust optimal control problems governed by semilinear elliptic partial differential equations.

1 Introduction

Real-life optimization problems are often subject to uncertainties. They are modeled according to

$$\min \quad f(x, y)$$

where $x \in \mathbb{R}^n$ denotes the optimization variable, and y represents uncertain data, which may originate, e.g., from unreliable model parameters. Additional constraints coupling x and y are often present, as in (1.1) below.

There are various fundamentally different approaches to take these uncertainties into account. These differ with respect to their qualifications for a particular purpose. *Stochastic* methods (Artstein and Wets [1994], Kall and Wallace [1994], Wets [1974]) require an assumption about the probability distribution of the uncertain parameters. They typically try to achieve a solution which is optimal in an average sense. By contrast, we consider here the *worst-case* approach which is more conservative but does not require a priori knowledge about the distribution of uncertainties. The worst-case approach may be viewed as a two-player game in which the optimization variable x has to be determined first. Using this knowledge, an opponent then chooses the parameter y which produces the largest value of the objective $f(x, y)$.

In this paper, we consider worst-case optimization problems of the form

$$\min_{x \in X_{\text{ad}}} \quad \max_{y \in Y(x)} \quad f(x, y), \tag{1.1}$$

with feasible sets $X_{\text{ad}} \subset \mathcal{X}$ and $Y(x) \subset \mathcal{Y}$, where $\mathcal{X}$ and $\mathcal{Y}$ are normed linear spaces. This bi-level optimization problem reflects the leader/follower nature of the game. The main purpose of this paper is to study the existence of a solution of (1.1) in the case of infinite dimensional spaces $\mathcal{X}$ and $\mathcal{Y}$. To this end, we investigate the lower semi-continuity of the value function

$$\varphi(x) = \sup_{y \in Y(x)} f(x, y). \tag{1.2}$$

This and other properties of the value function are well known for finite dimensional spaces, see for example Bank et al. [1983], Dem'janov and Malozemov [1990], Dempe

Date: February 28, 2011.

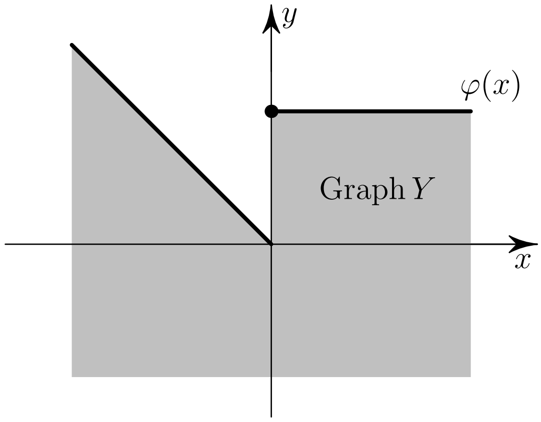

FIGURE 1.1. Graphs of Y (gray) and of the optimal value function $\varphi(x) = \max_{y \in Y(x)} y$ (bold line)

[2002] and Stein [2003]. The case of infinite dimensions is more delicate since we now have at our disposal different topologies. When using the norm (strong) topology, the proofs and results closely parallel the finite dimensional setting, see for instance Hogan [1973] and Rockafellar and Wets [1998]. However, the results available in the literature so far cannot be applied to prove the existence of solutions for worst-case optimal control problems. The reason is that X_{ad} is usually non-compact in the strong topology, and therefore *strong* lower semi-continuity of φ does not suffice to conclude the existence of a solution. We therefore resort to the stronger notion of *weak* lower semi-continuity.

The main contribution of this paper is the proof of *weak* lower semi-continuity of the value function φ under appropriate conditions on Y and f, see Theorem 2.5. These conditions accomodate the treatment of worst-case robust optimal control problems for partial differential equations (PDEs). We present in Section 3 several examples with semilinear elliptic PDEs for which we prove the existence of solutions.

We emphasize that even in finite dimensions with compact X_{ad} and continuous f, the existence of a worst-case robust solution of (1.1) may fail, as shown by the following simple example with $\mathcal{X} = \mathcal{Y} = \mathbb{R}$:

$$\min_{x \in [-1,1]} \max_{y \in Y(x)} y, \tag{1.3}$$

where the set-valued map Y is defined by

$$Y(x) = \begin{cases} \{y \in \mathbb{R} : y \le -x\} & \text{if } x < 0, \\ \{y \in \mathbb{R} : y \le 1\} & \text{otherwise.} \end{cases} \tag{1.4}$$

Both Y and the optimal value function $\varphi(x) = \max_{y \in Y(x)} y$ are illustrated in Figure 1.1. The value function is not lower semi-continuous, and problem (1.3) does not possess a solution.

2 Weak Lower Semi-Continuity of the Value Function

Throughout this section, $\mathcal{X}$ and $\mathcal{Y}$ are real normed linear spaces and f denotes a map

$$f : \mathcal{X} \times \mathcal{Y} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}.$$

Moreover, $Y : \mathcal{X} \rightrightarrows \mathcal{Y}$ is a set-valued map, i.e., $Y(x) \subset \mathcal{Y}$ for all $x \in \mathcal{X}$. For a set $X \subset \mathcal{X}$ we denote by

$$\operatorname{Graph}_X Y := \big\{(x, y) : x \in X,\ y \in Y(x)\big\} \subset \mathcal{X} \times \mathcal{Y},$$

the graph of Y relative to $X \subset \mathcal{X}$. Here we follow the notation of Rockafellar and Wets [1998]. We drop the term *relative to* if $X = \mathcal{X}$. Standard notation

$$x_n \rightharpoonup x \quad \text{and} \quad x_n \to x$$

is used to express the weak or strong convergence, respectively, of a sequence $\{x_n\}$.

In this section, we establish conditions under which the value function

$$\varphi(x) = \sup_{y \in Y(x)} f(x, y)$$

is weakly lower semi-continuous at a point x, i.e.,

$$x_n \rightharpoonup x \quad \Rightarrow \quad \varphi(x) \le \liminf_{n\to\infty} \varphi(x_n).$$

In fact, we shall investigate the weak lower semi-continuity of φ at x *relative to* a set $X \subset \mathcal{X}$, i.e.,

$$X \ni x_n \rightharpoonup x \quad \Rightarrow \quad \varphi(x) \le \liminf_{n\to\infty} \varphi(x_n).$$

(Note that the weak limit point x does not necessarily belong to X.) This property of φ requires assumptions on both, the set-valued map Y as well as the function f, which are addressed in Section 2.1 and 2.2, respectively. The main result is given in Theorem 2.5.

In the sequel, all properties concerning Y, f and φ will be characterized by sequences. That is to say, we should expressly speak in Definition 2.2 of weak-strong *sequential* inner semi-continuity, and in Definition 2.4 of weak-strong [or weak-weak] *sequential* lower semi-continuity. In order to simplify notation, however, we consistently omit the attribute *sequential*.

2.1. Inner Semi-Continuity Properties of Set-Valued Maps. Inner semi-continuity of Y is an important property. It fails to hold in the introductory example (1.4).

Definition 2.1 (see [Rockafellar and Wets, 1998, Chapter 4.A])**.** *The **inner limit** of a sequence of sets $\{Y_n\}$ with $Y_n \subset \mathcal{Y}$ is defined as*

$$\liminf_{n\to\infty} Y_n := \big\{y \in \mathcal{Y} : \limsup_{n\to\infty} d(y, Y_n) = 0\big\}.$$

Here $d(y, Y_n)$ denotes the distance of the point y to the set Y_n.

We can now define a notion of inner semi-continuity of Y which is adapted to the mere weak convergence of the argument.

Definition 2.2. *The set-valued function Y is said to be **weakly-strongly inner semi-continuous** at $\overline{x}$ relative to a set $X \subset \mathcal{X}$ if for all sequences $\{x_n\} \subset X$ satisfying $x_n \rightharpoonup \overline{x}$,*

$$Y(\overline{x}) \subset \liminf_{n\to\infty} Y(x_n).$$

The following equivalent characterization clarifies the notion of 'weak-strong' inner semi-continuity.

Proposition 2.3. *Y is weakly-strongly inner semi-continuous at $\overline{x}$ relative to X if and only if for all $\overline{y} \in Y(\overline{x})$ and all sequences $\{x_n\} \subset X$ with $x_n \rightharpoonup \overline{x}$, there exists a sequence $\{y_n\} \subset \mathcal{Y}$ with $y_n \to \overline{y}$ and $y_n \in Y(x_n)$ for sufficiently large n.*

Proof. To show necessity, consider $\overline{y} \in Y(\overline{x})$ and a sequence $\{x_n\} \subset X$ such that $x_n \rightharpoonup \overline{x}$. The weak-strong inner semi-continuity of Y at $\overline{x}$ relative to X implies the relation

$$0 \le \limsup_{n\to\infty} d(\overline{y}, Y(x_n)) = 0,$$

i.e., $\lim_{n\to\infty} d(\overline{y}, Y(x_n)) = 0$. Therefore, there exists a sequence with the properties claimed.

Conversely, let $\{x_n\} \subset X$ be a sequence converging weakly to $\overline{x}$ and let $\overline{y} \in Y(\overline{x})$. We need to show $\overline{y} \in \liminf_{n\to\infty} Y(x_n)$, or equivalently, $\limsup_{n\to\infty} d(\overline{y}, Y(x_n)) = 0$. This follows immediately from the assumption, using the sequence $\{y_n\}$. ☐

For some applications, the requirement of Y being weakly-strongly inner semi-continuous at $\overline{x}$ may still be too much to ask. To relax it further, we define in analogy to Proposition 2.3 the notion of **weak-weak inner semi-continuity** at x relative to a set $X \subset \mathcal{X}$ as follows: For all $\overline{y} \in Y(\overline{x})$ and all sequences $\{x_n\} \subset X$ with $x_n \rightharpoonup \overline{x}$, there exists a sequence $\{y_n\} \subset \mathcal{Y}$ with $y_n \rightharpoonup \overline{y}$ and $y_n \in Y(x_n)$ for sufficiently large n.

2.2. Lower-Semicontinuity Properties of Functions.

Definition 2.4. *The function f is said to be **weakly-strongly [weakly-weakly] lower semi-continuous** at $(\overline{x}, \overline{y}) \in \mathcal{X} \times \mathcal{Y}$ relative to $Z \subset \mathcal{X} \times \mathcal{Y}$ if for all sequences $\{x_n\} \subset \mathcal{X}$, $\{y_n\} \subset \mathcal{Y}$ with $(x_n, y_n) \in Z$ satisfying $x_n \rightharpoonup \overline{x}$, $y_n \to \overline{y}$ [$y_n \rightharpoonup \overline{y}$], the inequality*

$$\liminf_{n\to\infty} f(x_n, y_n) \le f(\overline{x}, \overline{y})$$

holds.

Our main result is the following.

Theorem 2.5 (Weak Lower Semi-Continuity of the Optimal Value Function)**.** *Let $X \subset \mathcal{X}$ and suppose that one of the following conditions is satisfied.*

- *(a) Y is weakly-strongly inner semi-continuous at $\overline{x}$ relative to X and f is weakly-strongly lower semi-continuous on $\overline{x} \times Y(\overline{x})$ relative to* $\operatorname{Graph}_X Y$.
- *(b) Y is weakly-weakly inner semi-continuous at $\overline{x}$ relative to X and f is weakly-weakly lower semi-continuous on $\overline{x} \times Y(\overline{x})$ relative to* $\operatorname{Graph}_X Y$.
- *(c) $Y(x) \equiv Y$ for all $x \in X$ and $f(\cdot, y)$ is weakly lower semi-continuous at $\overline{x}$ relative to X, for all $y \in Y$.*

Then the optimal value function φ is weakly lower semi-continuous at $\overline{x}$ relative to X.

Note that the weak-weak inner semi-continuity of Y is a weaker concept than weak-strong inner semi-continuity. To compensate, f needs to be weakly-weakly lower semi-continuous, which is a stronger concept than weak-strong lower semi-continuity.

Proof. We adapt the proof of Hogan [1973] which applies to the case of the strong topologies of $\mathcal{X}$ and $\mathcal{Y}$.

(a): We only need to consider the case $Y(\overline{x}) \neq \emptyset$ and $\varphi(\overline{x}) > -\infty$. Suppose that $\{x_n\} \subset X$ is a sequence converging weakly to $\overline{x}$. We need to show that $\liminf_{n\to\infty} \varphi(x_n) \ge \varphi(\overline{x})$.

Case I: $\varphi(\overline{x}) = +\infty$. The definition of φ implies that there exists a sequence $\{y_n\} \subset Y(\overline{x})$ such that $\lim_{n\to\infty} f(\overline{x}, y_n) = +\infty$. Due to the weak-strong inner semi-continuity of Y at $\overline{x}$ relative to X and Proposition 2.3 (with y_n in place of $\overline{y}$), there exists, for every $n \in \mathbb{N}$, a sequence $\{y_{n,k}\}_{k\in\mathbb{N}}$ satisfying $y_{n,k} \to y_n$ and $y_{n,k} \in Y(x_k)$ for all sufficiently large k (depending on n). The weak-strong lower semi-continuity of f on $\overline{x} \times Y(\overline{x})$ relative to $\operatorname{Graph}_X Y$ implies

$$f(\overline{x}, y_n) \le \liminf_{k\to\infty} f(x_k, y_{n,k}) \le \liminf_{k\to\infty} \varphi(x_k).$$

Passing to the limit $n \to \infty$ shows $\liminf_{k\to\infty} \varphi(x_k) = +\infty = \varphi(\overline{x})$.

Case II: $\varphi(\overline{x}) \in \mathbb{R}$. Let $\varepsilon > 0$ be given and choose $\overline{y} \in Y(\overline{x})$ such that $\varphi(\overline{x}) - \varepsilon \le f(\overline{x}, \overline{y})$ holds. Due to Proposition 2.3, there exists $\{y_n\}$ with $y_n \to \overline{y}$ and $y_n \in Y(x_n)$ for all sufficiently large n. Since f is weakly-strongly lower semi-continuous on $\overline{x} \times Y(\overline{x})$ relative to $\operatorname{Graph}_X Y$, we can conclude

$$\varphi(\overline{x}) - \varepsilon \le f(\overline{x}, \overline{y}) \le \liminf_{n\to\infty} f(x_n, y_n) \le \liminf_{n\to\infty} \varphi(x_n).$$

Since $\varepsilon > 0$ is arbitrary, the result is proved.

The proof of parts (b) and (c) proceeds along the same lines and is omitted. For part (c), we may choose $y_{n,k} \equiv y_n \in Y$ in Case I and $y_n \equiv \overline{y} \in Y$ in Case II. □

3 Application to Worst-Case Robust Optimal Control Problems

In this section, we consider three worst-case robust optimal control problems subject to semilinear elliptic PDEs. We also switch notation to comply with common variable names in optimal control. The control function u plays the role of the upper-level optimization variable (previously termed x), while the uncertain parameter p is the lower-level variable (previously termed y). In all cases, Theorem 2.5 will allow us to conclude that the value function

$$\varphi(u) = \sup_{p\in P_{\text{ad}}(u)} f(u, p)$$

is *weakly* lower semi-continuous. The existence of a worst-case robust optimal control is then a standard conclusion. The three examples differ with respect to the type of influence of the uncertain parameter, and with respect to which part of Theorem 2.5 is applicable.

We point out that results available to date concerning the *strong* semi-continuity of the value function cannot be applied here since minimizing sequences converge only *weakly*. Indeed, the set of admissible controls U_{ad} is typically not *strongly* but only *weakly* (sequentially) compact.

3.1. Uncertain Coefficient. Our first model problem is

$$\min_{u\in U_{\text{ad}}} \quad \max_{\substack{p\in P_{\text{ad}} \\ y\in H^1(\Omega)\cap C(\overline{\Omega})}} \int_\Omega \phi(x, y(x))\,\mathrm{d}x + \int_\Gamma \psi(x, u(x))\,\mathrm{d}s \tag{3.1}$$

$$\text{where} \quad \left\{ \begin{aligned} -\triangle y + d(x, y) &= 0 \quad \text{in } \Omega \\ \frac{\partial}{\partial n} y + p\,(y - u) &= 0 \quad \text{on } \Gamma. \end{aligned} \right. \tag{3.2}$$

Problem (3.1) is motivated as follows. The semilinear PDE (3.2) models, for instance, a stationary heat conduction problem with an unknown and possibly inhomogeneous heat transfer coefficient p. The control function u takes the role of the environmental or exterior temperature. We wish to find an admissible control $u \in U_{\text{ad}}$ which performs best in terms of the objective, under the worst possible realization of the heat transfer coefficient $p \in P_{\text{ad}}$.

The sets of admissible controls and potential transfer coefficients are defined as

$$U_{\text{ad}} = \{u \in L^\infty(\Gamma) : u_a \le u \le u_b \text{ a.e. on } \Gamma\} \tag{3.3a}$$

$$P_{\text{ad}} = \{p \in L^\infty(\Gamma) : |p - p_0| \le p_\delta \text{ a.e. on } \Gamma\}, \tag{3.3b}$$

where $p_0 \in L^\infty(\Gamma)$ denotes a given estimate of the uncertain heat transfer coefficient and $p_\delta \in L^\infty(\Gamma)$ is a reliability bound. In this example, P_{ad} is independent of the control.

Under the assumptions stated below, there exists, for any $(u, p) \in U_{\text{ad}} \times P_{\text{ad}}$, a unique solution $y = S(u, p)$ of the state equation (3.2). We may therefore eliminate the state variable y from the lower-level problem and consider instead the following reduced counterpart of (3.1):

$$\min_{u \in U_{\text{ad}}} \quad \max_{p \in P_{\text{ad}}} \int_\Omega \phi(x, S(u, p))\,\mathrm{d}x + \int_\Gamma \psi(x, u(x))\,\mathrm{d}s. \tag{3.4}$$

We now state our standing assumptions for this example. These coincide with the usual assumptions in the *non-robust* case for semilinear elliptic control problems, see for instance [Tröltzsch, 2010, Section 4.4].

Assumption 3.1. (E1) *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain with Lipschitz boundary Γ with $N \ge 2$.*

(E2) *Let $u_a, u_b, p_0, p_\delta \in L^\infty(\Gamma)$ and $u_a \le u_b$ hold a.e. on Γ. Moreover, $p_\delta \ge 0$ and $p_0 - p_\delta \ge \varepsilon$ for some $\varepsilon > 0$.*

(E3) *The nonlinearity $d : \Omega \times \mathbb{R} \to \mathbb{R}$ satisfies the* Carathéodory condition*, i.e.,*

$$\Omega \ni x \mapsto d(x, y) \in \mathbb{R} \quad \text{is measurable for all } y \in \mathbb{R},$$

$$\mathbb{R} \ni y \mapsto d(x, y) \in \mathbb{R} \quad \text{is continuous for almost all } x \in \Omega.$$

The same is assumed for ϕ and ψ.

(E4) *The nonlinearity $d : \Omega \times \mathbb{R} \to \mathbb{R}$ satisfies the* boundedness and local Lipschitz conditions*, i.e.,*

$$|d(x, 0)| \le K \quad \text{and} \quad |d(x, y_1) - d(x, y_2)| \le L(M)\,|y_1 - y_2|$$

holds for almost all $x \in \Omega$, and all $y_1, y_2 \in [-M, M]$.
The same is assumed for ϕ and ψ, with the obvious changes.

(E5) *Let $d(x, y)$ be monotone increasing w.r.t. y for almost all $x \in \Omega$.*

(E6) *Let $\psi(x, \cdot)$ be convex on $\mathbb{R}$ for almost all $x \in \Omega$.*

Assumption (E2) assures $p \ge \varepsilon$ and thus the well-posedness of the semilinear PDE for all $p \in P_{\text{ad}}$ follows. In order to prove the existence of a worst-case robust optimal control, we will apply Theorem 2.5. To verify its requisites, we need some properties of the control-to-state map S.

Lemma 3.2. *Suppose $s > N-1$ and $r > N/2$. For any $u \in L^s(\Gamma)$ and $p \in P_{\text{ad}}$, the semilinear state equation (3.2) has a unique solution $y = S(u, p)$ in $H^1(\Omega) \cap C(\overline{\Omega})$. The solution map has the following properties:*

(a) There exists $c_{r,s} > 0$ such that the a priori bound

$$\|S(u, p)\|_{H^1(\Omega)} + \|S(u, p)\|_{C(\overline{\Omega})} \le c_{r,s} \left(\|p\,u\|_{L^s(\Gamma)} + \|d(\cdot, 0)\|_{L^r(\Omega)}\right) \tag{3.5}$$

holds for all $u \in L^s(\Gamma)$ and $p \in P_{\text{ad}}$, with $c_{r,s}$ independent of u and p.

(b) Suppose in addition that

$$\begin{cases} 1 < s < \infty & \text{in case } N = 2, \\ (2N-2)/(N-2) \le s < \infty & \text{in case } N \ge 3. \end{cases}$$

If $\{u_n\} \subset L^s(\Gamma)$ such that $u_n \rightharpoonup u$ in $L^s(\Gamma)$, and if $\{p_n\} \subset P_{\text{ad}}$ such that $p_n \to p$ in $L^t(\Gamma)$ where

$$\begin{cases} s' \le t \le \infty & \text{in case } N = 2, \\ N-1 \le t \le \infty & \text{in case } N \ge 3, \end{cases}$$

then

$$S(u_n, p_n) \rightharpoonup S(u,p) \quad \text{in } H^1(\Omega).$$

Here s' denotes the conjugate exponent of s, i.e., $1/s + 1/s' = 1$.

Proof. (a): The proof is given in [Tröltzsch, 2010, Theorem 4.8] for the parameter independent case. It is based on the Browder-Minty theorem for monotone operators and a cut-off argument to show the boundedness of the solution. The a priori constant $c_{r,s}$ depends only on the coercivity constant associated with the linear part of the differential operator in (3.2), whose minimum is positive and attained for $p = p_0 - p_\delta$.

(b): The proof is a modification of typical arguments for semilinear equations, see, e.g., [Tröltzsch, 2010, Theorem 4.15]. Let $y_n := S(u_n, p_n)$. Since p_n is bounded in $L^\infty(\Gamma)$, $\{p_n\, u_n\}$ is bounded in $L^s(\Gamma)$, and the a priori estimate of part (a) shows that $\{y_n\}$ is bounded in $H^1(\Omega) \cap C(\overline{\Omega})$. This implies that $d_n := d(y_n)$ is bounded in $L^\infty(\Omega)$, which follows from the properties of the Nemitskii operator $d(\cdot)$, see [Tröltzsch, 2010, Lemma 4.11]. We extract a joint subsequence, still denoted by n, such that $y_n \rightharpoonup \widehat{y}$ in $H^1(\Omega)$ and $d_n \rightharpoonup \widehat{d}$ in $L^t(\Omega)$ where we fix $2N/(N+2) \le w \le 2N/(N-2)$ in case $N \ge 3$, or $1 < w < \infty$ in case $N = 2$. The upper bound on w ensures $H^1(\Omega) \hookrightarrow L^w(\Omega)$, and the lower bound implies that a function in $L^w(\Omega)$ generates an element of $H^1(\Omega)^*$.

We observe that y_n satisfies the linear state equation

$$\begin{aligned} -\triangle y_n + y_n &= -d_n + y_n && \text{in } \Omega \\ \frac{\partial}{\partial n} y_n &= p_n\,(u_n - y_n) && \text{on } \Gamma. \end{aligned}$$

The right hand side in the first equation converges weakly in $L^w(\Omega)$ to $-\widehat{d} + \widehat{y}$.

Let us consider the terms on the boundary. We have $p_n\, u_n \rightharpoonup p\, u$ in $L^v(\Gamma)$ for $v = (2N-2)/N$ by Hölder's inequality and the condition imposed on s in part (b). Note that a function in $L^v(\Gamma)$ generates an element of $H^1(\Omega)^*$. The continuity of the trace operator (see [Adams and Fournier, 2003, Theorem 5.36]) implies $y_n \rightharpoonup \widehat{y}$ in $L^q(\Gamma)$ for $q \in [1,\infty)$ in case $N = 2$ or $q = 2\,(N-1)/(N-2)$ for $N \ge 3$. Using this and $s > N-1$, $p_n\, y_n \rightharpoonup p\,\widehat{y}$ in $L^v(\Gamma)$ follows.

Consequently, the right hand side in the boundary conditions converges in $L^v(\Gamma)$ to $p\,(u - \widehat{y})$ and thus $\widehat{y}$ satisfies the equation

$$\begin{aligned} -\triangle \widehat{y} + \widehat{y} &= -\widehat{d} + \widehat{y} && \text{in } \Omega \\ \frac{\partial}{\partial n} \widehat{y} &= p\,(u - \widehat{y}) && \text{on } \Gamma. \end{aligned}$$

To conclude the proof, we can proceed as in [Tröltzsch, 2010, Theorem 4.15] to verify $\widehat{d} = d(\widehat{y})$ and thus $\widehat{y} = S(u,p)$. The convergence extends to the whole sequence since the limit $\widehat{y}$ is unique. This shows $S(u_n, p_n) \rightharpoonup S(u,p)$ in $H^1(\Omega)$. □

Remark 3.3. *The previous lemma continues to hold if P_{ad} is replaced by any other subset of $L^\infty(\Gamma)$ all of whose elements satisfy $p \ge \varepsilon$ a.e. on Γ for some $\varepsilon > 0$. This will be used in Section 3.2.*

Theorem 3.4. *The worst-case robust optimal control problem (3.4) with $U_{\rm ad}$ and $P_{\rm ad}$ as in (3.3) has at least one global solution $\overline{u} \in U_{\rm ad}$.*

Proof. We begin by verifying, for any fixed $p \in P_{\rm ad}$, the weak lower semi-continuity of

$$u \mapsto f(u,p) := \int_\Omega \phi(x, S(u,p))\,\mathrm{d}x + \int_\Gamma \psi(x, u(x))\,\mathrm{d}s$$

at every point of and relative to $U_{\rm ad}$ and w.r.t. the topology of $L^s(\Gamma)$. Suppose that $\{u_n\} \subset U_{\rm ad}$ converges weakly in $L^s(\Gamma)$ to some u. (Since $U_{\rm ad}$ is weakly closed in $L^s(\Gamma)$, $u \in U_{\rm ad}$ holds necessarily.) We need to show that $f(u,p) \le \liminf f(u_n,p)$ for any fixed $p \in P_{\rm ad}$. Using the properties of the Nemitskii operators ϕ and ψ, this follows as in [Tröltzsch, 2010, Theorem 4.15] so we can be brief. First of all, $u \mapsto \int_\Omega \phi(\cdot, S(u,p))\,\mathrm{d}x$ is weakly continuous on all of $L^s(\Gamma)$ into $\mathbb{R}$. Moreover, $u \mapsto \int_\Gamma \psi(s, u(s))\,\mathrm{d}s$ is convex and continuous on $U_{\rm ad}$ w.r.t. $L^s(\Gamma)$ and thus weakly lower semi-continuous.

Theorem 2.5 (c) now implies that the optimal value function

$$\varphi(u) = \sup_{p \in P_{\rm ad}} f(u,p)$$

is weakly lower semi-continuous at every point of and relative to $U_{\rm ad}$ w.r.t. $L^s(\Gamma)$. Moreover, φ is bounded below on $U_{\rm ad}$ since $\varphi(u) \ge f(u,p)$ for all $p \in P_{\rm ad}$ and, in turn, f is bounded below on $U_{\rm ad} \times P_{\rm ad}$, see again [Tröltzsch, 2010, Theorem 4.15]. The existence of a global optimal control $\overline{u} \in U_{\rm ad}$ now follows from standard arguments since $U_{\rm ad}$ is weakly compact in $L^s(\Gamma)$. ☐

3.2. Uncertain Temperature Dependent Coefficient. In problem (3.1)–(3.2) we assumed no a priori knowledge about the uncertain heat transfer coefficient p except for the bounds imposed in $P_{\rm ad}$, see (3.3b). Suppose now that we have available a model $m(y)$ for the heat transfer coefficient in dependence of the temperature. This model, however, is uncertain as well and it gives rise to the following definition of the set of admissible parameters:

$$P_{\rm ad}(u) = \{p \in L^\infty(\Gamma) : |p - m(S(u,p))| \le p_\delta \text{ and } p \ge \varepsilon \text{ a.e. on } \Gamma\},$$

where $y = S(u,p)$ solves (3.2). As before, $p_\delta \in L^\infty(\Gamma)$ is a given reliability bound.

Note that $p_\delta \equiv 0$ would correspond to a completely reliable model of the heat transfer coefficient's dependence on the temperature. In this case, the correct value of p satisfies the fixed point equation $p = m(S(u,p))$. With positive values of p_δ, we acknowledge the uncertainty of $m(y)$ and admit also approximate fixed points in $P_{\rm ad}(u)$. The assumption of a strictly positive heat transfer coefficient $p \ge \varepsilon$ for some $\varepsilon > 0$ assures again the well-posedness of the PDE.

The implicit structure of $P_{\rm ad}(u)$ renders the verification of its inner semi-continuity properties difficult. Therefore, we focus our attention to the following simplified variants of this set.

$$P^{(1)}_{\rm ad}(u) = \{p \in L^\infty(\Gamma) : |p - m(S(u,p_0))| \le p_\delta \text{ and } p \ge \varepsilon \text{ a.e. on } \Gamma\} \tag{3.6a}$$

$$P^{(2)}_{\rm ad}(u) = \{p \in L^\infty(\Gamma) : |p - m(S(u, m(S(u,p_0))))| \le p_\delta \text{ and } p \ge \varepsilon \text{ a.e. on } \Gamma\}. \tag{3.6b}$$

As in the previous subsection, p_0 denotes an a priori estimate of the uncertain heat transfer coefficient. And thus $m(S(u,p_0))$ is an approximation of the actual coefficient associated with the control u. It can be viewed as the first fixed point iterate for $p = m(S(u,p))$ starting from p_0. In the same way, the second fixed-point iterate gives rise to the definition of $P^{(2)}_{\rm ad}(u)$. Note that $P_{\rm ad}(u)$ defined in (3.3b) of the previous example corresponds to an approximation with zero fixed point steps.

In this subsection we consider problem (3.1)–(3.2) with U_{ad} still defined by (3.3a) and P_{ad} replaced by $P_{\mathrm{ad}}^{(2)}(u)$, see (3.6b). The results obtained below hold for the problem involving $P_{\mathrm{ad}}^{(1)}$ as well, which is simpler and its discussion therefore omitted.

Assumption 3.5. (M1) *The heat transfer coefficient model $m : \Gamma \times \mathbb{R} \to \mathbb{R}$ satisfies the Carathéodory, the boundedness, and the local Lipschitz conditions, compare Assumption 3.1, (E3) and (E4).*

(M2) *Suppose there exists $\varepsilon > 0$ such that $m(x, y) \geq \varepsilon$ all $y \in \mathbb{R}$ and a.a. $x \in \Gamma$.*

(M3) *The dimension of Ω is restricted to $N \in \{2, 3\}$.*

For the remainder of this subsection, we suppose Assumption 3.1 and 3.5 to hold. Note that (M2) implies that $p = m(S(u, m(S(u, p_0)))) \in P_{\mathrm{ad}}^{(2)}(u)$ and thus $P_{\mathrm{ad}}^{(2)}(u)$ is non-empty for every $u \in L^s(\Gamma)$ with $s > N - 1$.

Lemma 3.6. *The map $u \mapsto P_{\mathrm{ad}}^{(2)}(u)$ is weakly-strongly inner semi-continuous at every point $\overline{u} \in L^s(\Gamma)$ w.r.t. the topologies of $L^s(\Gamma)$ and $L^t(\Gamma)$, where*

$$\begin{cases} 1 < s < \infty \text{ and } 1 < t < \infty \text{ s.t. } 1/s + 1/t < 1 & \text{in case } N = 2, \\ 4 < s < \infty \text{ and } 2 < t < 4 \text{ s.t. } 1/s + 1/t < 1/2 & \text{in case } N = 3. \end{cases}$$

Proof. We present the proof for the more complicated case $N = 3$ only. Let s and t be fixed numbers according to the specification. Consider $\overline{u} \in L^s(\Gamma)$ and a sequence $\{u_n\}$, $u_n \rightharpoonup \overline{u}$ in $L^s(\Gamma)$ and $\overline{p} \in P_{\mathrm{ad}}^{(2)}(\overline{u})$. To simplify notation, we define $\widetilde{p}_n = m(S(u_n, p_0))$. We need to show that there exists a sequence $\{p_n\}$ such that $p_n \in P_{\mathrm{ad}}^{(2)}(u_n)$ for sufficiently large n and $p_n \to \overline{p}$ in $L^t(\Gamma)$.

Due to Lemma 3.2 (b), the sequence $S(u_n, p_0)$ converges weakly in $H^1(\Omega)$ and thus, by compact embedding (see [Adams and Fournier, 2003, Theorem 6.3]), strongly in $H^{1-\delta}(\Omega)$ (for all $\delta > 0$) to $S(\overline{u}, p_0)$. Applying the continuity of the trace operator, we conclude the strong convergence of $S(u_n, p_0) \to S(\overline{u}, p_0)$ in $L^t(\Gamma)$. Note that $\{S(u_n, p_0)\}$ is bounded in $L^\infty(\Gamma)$ by Lemma 3.2 (a). Since m is continuous w.r.t. $L^t(\Gamma)$ on bounded subsets of $L^\infty(\Gamma)$, we infer the strong convergence of $m(S(u_n, p_0)) = \widetilde{p}_n \to \widetilde{p} = m(S(\overline{u}, p_0))$ in $L^t(\Gamma)$.

We now define p_n as the pointwise projection of $\overline{p}$ onto $P_{\mathrm{ad}}^{(2)}(u_n)$, i.e.,

$$\begin{aligned} p_n &:= \operatorname{proj}_{P_{\mathrm{ad}}^{(2)}(u_n)}\{\overline{p}\} \\ &= \max\left\{\min\left\{\overline{p},\ m(S(u_n, \widetilde{p}_n)) + p_\delta\right\},\ m(S(u_n, \widetilde{p}_n)) - p_\delta,\ \varepsilon\right\}. \end{aligned}$$

By definition, $p_n \in P_{\mathrm{ad}}^{(2)}(u_n)$ holds. Moreover, since $m(S(u_n, p_0)) \geq \varepsilon$ holds by (M2), we may use again Lemma 3.2 (b) to conclude as above that $S(u_n, \widetilde{p}_n) \rightharpoonup S(\overline{u}, \widetilde{p})$ in $H^1(\Omega)$ and thus $S(u_n, \widetilde{p}_n) \to S(\overline{u}, \widetilde{p})$ in $L^t(\Gamma)$. Note that $\{u_n \widetilde{p}_n\}$ is bounded in $L^{2+\tau}(\Gamma)$ for some $\tau > 0$ and thus by Lemma 3.2 (a), $\{S(u_n, \widetilde{p}_n)\}$ is bounded in $L^\infty(\Gamma)$. We can thus conclude as above that $m(S(u_n, \widetilde{p}_n)) \to m(S(\bar{u}, \widetilde{p}))$ in $L^t(\Gamma)$. The continuity of the Nemytskii operators max and min now provides the conclusion $p_n \to \overline{p}$ in $L^t(\Gamma)$. □

Theorem 3.7. *The worst-case robust optimal control problem (3.1)–(3.2) with U_{ad} defined by (3.3a) and P_{ad} as in (3.6b) has at least one global solution $\overline{u} \in U_{\mathrm{ad}}$.*

Proof. We need to verify a lower semi-continuity property of f. In contrast to the proof of Theorem 3.4, P_{ad} now depends on u, and thus the dependence of f on both u and p must be considered. We define the auxiliary set $P_{\mathrm{ad}}^\varepsilon = \{p \in L^\infty(\Gamma) : p \geq \varepsilon \text{ a.e. on } \Gamma\}$. Using Lemma 3.2 (b), it follows as in the proof of Theorem 3.4 that

$$(u, p) \mapsto f(u, p) := \int_\Omega \phi(x, S(u, p))\, \mathrm{d}x + \int_\Gamma \psi(x, u(x))\, \mathrm{d}s$$

is weakly-strongly lower semi-continuous w.r.t. $L^s(\Gamma)$ and $L^r(\Gamma)$ relative to $U_{\text{ad}} \times P^\varepsilon_{\text{ad}}$ at every point of that set. Here r and s are exponents as specified in Lemma 3.2 (b). Note that $U_{\text{ad}} \times P^\varepsilon_{\text{ad}}$ contains $\text{Graph}_{U_{\text{ad}}} P^{(2)}_{\text{ad}}$.

We have shown in Lemma 3.6 that $u \mapsto P^{(2)}_{\text{ad}}(u)$ is weakly-strongly inner semi-continuous at every $\overline{u} \in L^s(\Gamma)$ w.r.t. $L^s(\Gamma)$ and $L^r(\Gamma)$ for possibly other pairs of values (r, s). However, the choice $(r, s) = (3, 2)$ in case $N = 2$ and $(r, s) = (4.5, 4.5)$ in case $N = 3$ is valid for both lemmas.

Theorem 2.5 (a) now implies that the optimal value function

$$\varphi(u) = \sup_{p \in P^{(2)}_{\text{ad}}(u)} f(u, p)$$

is weakly lower semi-continuous relative to U_{ad} w.r.t. $L^s(\Gamma)$ at every point of U_{ad}. Moreover, φ is bounded below on U_{ad} since $\varphi(u) \ge f(u, p)$ for all $p \in P^{(2)}_{\text{ad}}(u)$ and, in turn, f is bounded below even on the larger set $U_{\text{ad}} \times P^\varepsilon_{\text{ad}}$, see again [Tröltzsch, 2010, Theorem 4.15]. The existence of a global optimal control $\overline{u}$ now follows as before since U_{ad} is weakly compact in $L^s(\Gamma)$. □

3.3. Implementation Error. In this final example we address an optimal control problem with a different kind of perturbation. Here the uncertainty lies in an implementation error, i.e., the selected control function is applied only after a modification which cannot be anticipated. The model problem in this section is

$$\min_{u \in U_{\text{ad}}} \quad \max_{\substack{p \in P_{\text{ad}} \\ y \in H^1(\Omega) \cap C(\overline{\Omega})}} \int_\Omega \phi(x, y(x)) \, \mathrm{d}x + \int_\Gamma \psi(x, u(x)) \, \mathrm{d}s \tag{3.7}$$

$$\text{where} \quad \left\{ \begin{aligned} -\triangle y + d(x, y) &= 0 && \text{in } \Omega \\ \frac{\partial}{\partial n} y + y &= u + p && \text{on } \Gamma. \end{aligned} \right. \tag{3.8}$$

We wish to find an admissible control $u \in U_{\text{ad}}$ which performs best in terms of the objective, under the worst possible implementation error $p \in P_{\text{ad}}$. The admissible controls and implementation errors are defined as in (3.3a) and (3.3b).

In this section we work under the standing Assumption 3.1. (In fact, in (E2) the condition $p_0 - p_\delta \ge \varepsilon$ is not necessary here.) We observe the following properties of the control-to-state map $S(v) = S(u + p) = \widetilde{S}(u, p)$ associated with (3.8).

Lemma 3.8. *Suppose $s > N - 1$ and $r > N/2$. For any $v \in L^s(\Gamma)$, the semilinear state equation (3.8) has a unique solution $y = S(v)$ in $H^1(\Omega) \cap C(\overline{\Omega})$. The solution map has the following properties:*

(a) There exists $c_{r,s} > 0$ such that the a priori bound

$$\|S(v)\|_{H^1(\Omega)} + \|S(v)\|_{C(\overline{\Omega})} \le c_{r,s} \left(\|v\|_{L^s(\Gamma)} + \|d(\cdot, 0)\|_{L^r(\Omega)} \right) \tag{3.9}$$

holds for all $v \in L^s(\Gamma)$, with $c_{r,s}$ independent of u and p.

(b) If $\{v_n\} \subset L^s(\Gamma)$ such that $v_n \rightharpoonup v$ in $L^s(\Gamma)$, then

$$S(v_n) \rightharpoonup S(v) \quad \text{in } H^1(\Omega).$$

Proof. The result follows along the lines of Lemma 3.2 with $p \equiv 1$ and u replaced by v. □

In view of Lemma 3.8 we can eliminate the state variable y as before from the lower-level problem and consider the following reduced counterpart of (3.7)–(3.8):

$$\min_{u \in U_{\text{ad}}} \quad \max_{p \in P_{\text{ad}}} \int_\Omega \phi(x, S(u + p)) \, \mathrm{d}x + \int_\Gamma \psi(x, u(x)) \, \mathrm{d}s. \tag{3.10}$$

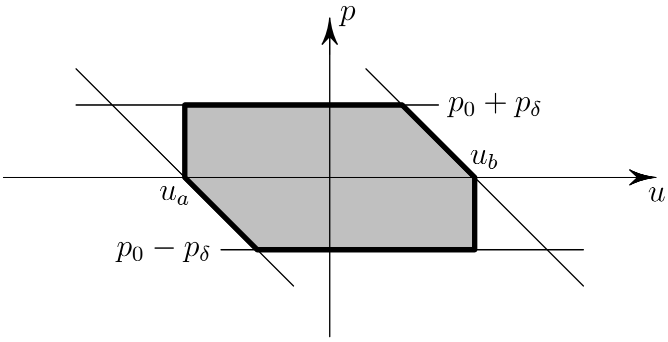

FIGURE 3.1. Pointwise graph of $u \mapsto P_{\mathrm{ad}}(u)$ in (3.11)

The proof of the following existence theorem is analogous to the proof of Theorem 3.4 with straightforward modifications and therefore omitted.

Theorem 3.9. *The worst-case robust optimal control problem (3.10) with U_{ad} and P_{ad} as in (3.3) has at least one global solution $\overline{u} \in U_{\mathrm{ad}}$.*

Remark 3.10. *With P_{ad} as in (3.3b), the perturbed control $u + p$ does not necessarily respect the bounds u_a and u_b. One might therefore instead consider the set*

$$P_{\mathrm{ad}}(u) = \{p \in L^\infty(\Gamma) : u_a \le u + p \le u_b \text{ and } |p - p_0| \le p_\delta \text{ a.e. on } \Gamma\}, \tag{3.11}$$

see Figure 3.1 for an illustration of the pointwise conditions. This set, however, does not enjoy weak-weak inner semi-continuity (nor weak-strong inner semi-continuity, of course). This is shown by means of the following example.

First we note that as before we would need to consider the inner semi-continuity properties of $u \mapsto P_{\mathrm{ad}}(u)$ only relative to U_{ad}. Suppose $\Gamma = (0, 1) \subset \mathbb{R}$, $-u_a = u_b \equiv 2$, $-p_a = p_b \equiv 1$ and consider the sequence

$$u_n(x) := 2 \operatorname{sgn}(\sin(2\pi n x)),$$

which belongs to U_{ad} and satisfies $u_n \rightharpoonup 0 = \overline{u}$ in $L^2(\Gamma)$. If p_n is a sequence in $P_{\mathrm{ad}}(u_n)$ converging weakly to $\overline{p}$ in $L^2(\Gamma)$, then it is easy to see that necessarily $-0.5 \le \int_0^1 \overline{p} \cdot 1 \, \mathrm{d}x \le 0.5$. Hence $\overline{p} \equiv 1$, which belongs to $P_{\mathrm{ad}}(\overline{u})$, cannot be attained as a weak limit.

We note that this difficulty is not present in finite dimensions where clearly,

$$P_{\mathrm{ad}}(u) = \{p \in \mathbb{R}^n : u_a \le u + p \le u_b \text{ and } |p - p_0| \le p_\delta\}$$

is inner semi-continuous relative to $U_{\mathrm{ad}} = \{u \in \mathbb{R}^n : u_a \le u \le u_b\}$.

References

R. Adams and J. Fournier. *Sobolev Spaces.* Academic Press, New York, second edition, 2003.

Zvi Artstein and Roger J.-B. Wets. Stability results for stochastic programs and sensors, allowing for discontinuous objective functions. *SIAM Journal on Optimization*, 4(3):537–550, 1994.

B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer. *Non-linear parametric optimization.* Birkhäuser Berlin, 1983.

W.F. Dem'janov and W.N. Malozemov. *Introduction to Minimax.* Dover Publication, New York, 1990.

Stephan Dempe. *Foundations of bilevel programming.* Kluwer Academic Publishers, Dordrecht, 2002.

William W. Hogan. Point-to-set maps in mathematical programming. *SIAM Review*, 15:591–603, 1973.

Peter Kall and Stein W. Wallace. *Stochastic programming.* John Wiley & Sons Ltd., Chichester, 1994.

R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998. doi: 10.1007/978-3-642-02431-3.

Oliver Stein. *Bi-level strategies in semi-infinite programming.* Nonconvex Optimization and its Applications. Kluwer Academic Publishers, Boston, MA, 2003.

F. Tröltzsch. *Optimal Control of Partial Differential Equations*, volume 112 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, 2010.

Roger J.-B. Wets. Stochastic programs with fixed recourse: the equivalent deterministic program. *SIAM Review*, 16:309–339, 1974.

Chemnitz University of Technology, Faculty of Mathematics, D–09107 Chemnitz, Germany

E-mail address: roland.herzog@mathematik.tu-chemnitz.de

URL: http://www.tu-chemnitz.de/herzog

Chemnitz University of Technology, Faculty of Mathematics, D–09107 Chemnitz, Germany

E-mail address: frank.schmidt@mathematik.tu-chemnitz.de

URL: http://www.tu-chemnitz.de/mathematik/part_dgl/people/schmidt