Weierstraß-Institut
für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Optimality conditions for convex stochastic optimization problems in Banach spaces with almost sure state constraints

Caroline Geiersbach[1], Winnifried Wollner[2]

submitted: August 31, 2020

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: caroline.geiersbach@wias-berlin.de

[2] Fachbereich Mathematik
Technische Universität Darmstadt
Dolivostr. 15
64293 Darmstadt
Germany
E-Mail: wollner@mathematik.tu-darmstadt.de

No. 2755
Berlin 2020

2010 *Mathematics Subject Classification.* 49K20, 49K21, 49K45, 49N15, 49J53.

Key words and phrases. PDE-constrained optimization under uncertainty, optimization in Banach spaces, optimality conditions, convex stochastic optimization in Banach spaces, two-stage stochastic optimization, regular Lagrange multipliers, duality.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Optimality conditions for convex stochastic optimization problems in Banach spaces with almost sure state constraints

Caroline Geiersbach, Winnifried Wollner

Abstract

We analyze a convex stochastic optimization problem where the state is assumed to belong to the Bochner space of essentially bounded random variables with images in a reflexive and separable Banach space. For this problem, we obtain optimality conditions that are, with an appropriate model, necessary and sufficient. Additionally, the Lagrange multipliers associated with optimality conditions are integrable vector-valued functions and not only measures. A model problem is given demonstrating the application to PDE-constrained optimization under uncertainty.

1 Introduction

Let X_1 and X_2 be real, reflexive, and separable Banach spaces. $(\Omega, \mathcal{F}, \mathbb{P})$ denotes a complete probability space, where Ω represents the sample space, $\mathcal{F} \subset 2^\Omega$ is the σ-algebra of events on the power set of Ω, and $\mathbb{P} \colon \Omega \to [0,1]$ is a probability measure. We assume $C_1 \subset X_1$ is nonempty, closed, and convex; $X_{2,\text{ad}}(x_1,\omega) \subset X_2$ is assumed to be nonempty, closed, and convex for all $x_1 \in C_1$ and almost all $\omega \in \Omega$. We are interested in a convex stochastic optimization problem of the form

$$\begin{aligned} \min_{x_1, x_2(\cdot)} \quad & \left\{ \mathbb{E}[J(x_1, x_2(\cdot))] = \int_\Omega J(x_1, x_2(\omega)) \, \mathrm{d}\mathbb{P}(\omega) \right\} \\ \text{s.t.} \quad & \begin{cases} x_1 \in C_1, \\ x_2(\omega) \in X_{2,\text{ad}}(x_1,\omega) \quad \text{a.s.}, \end{cases} \end{aligned} \tag{1}$$

where J is a convex real-valued mapping. In this model, the variable x_1, unlike x_2, is independent of the random data. As such, this problem can be interpreted as a static two-stage stochastic optimization problem. By "static," we mean to differentiate the problem from a stochastic optimization problem with recourse, where the second-stage "decision" x_2 is made only *after* observing a random element ω. Here, the function $\omega \mapsto x_2(\omega)$ is provided at the onset, which gives all possible decisions for each ω.

Such problems are of interest for applications to optimization with partial differential equations (PDEs) under uncertainty, where the set to which $x_2(\omega)$ belongs includes those states solving a PDE. This field is a rapidly developing one, with many developments in understanding the modeling, theory, and design of efficient algorithms; see, e.g., [9, 21, 27, 18, 1, 33, 14, 8, 12] and the references therein. So far, research has mostly been limited to the case where the control (in our notation, the first-stage variable x_1) has been subject to additional constraints. In this case, optimality conditions have already been established for risk-averse problems in [19, 20]. However, additional constraints on the state (here, x_2), beyond a uniquely solvable equation, have yet to be investigated thoroughly. Although chance constraints have been handled in such applications, cf. [11], the treatment of pointwise almost sure constraints on the state appear to be missing from the literature.

DOI 10.20347/WIAS.PREPRINT.2755 Berlin 2020

As a first step in this treatment, optimality conditions play a central role, and we pursue this in the current paper. Pointwise state constraints, without uncertainty, have received some attention over the last years. In general, optimality conditions require Lagrange multipliers coming from the non-separable space of regular Borel measures, see, e.g., [5, 6]. Only in rare circumstances it can be shown that multipliers can be found in more a regular, separable, space, see [7].

In this paper, we are focused on obtaining optimality conditions in the case where x_2 belongs to the Bochner space $L^\infty(\Omega, X_2)$. This choice is motivated by the goal of including problems where there is an almost sure bound such as

$$x_2(\omega) \leq_K \psi(\omega),$$

where $\psi \in L^\infty(\Omega, X_2)$ and $\leq_K$ represents a partial order on X_2. An example with this type of inequality is given in Subsection 4.1. The choice of $L^p(\Omega, X_2)$ for $p < \infty$ is not appropriate, as the cone $\{v \in L^p(\Omega, X_2) : v(\omega) \leq_K 0\}$ contains no interior points; this property is especially important in the establishment of Lagrange multipliers for our application. Therefore, we will view the problem presented in (1) in the framework of two-stage stochastic optimization (for an introduction, see [32, 23]). This framework allows us to generalize results from a series of papers by Rockafellar and Wets [28, 29, 30, 31], who established optimality theory of general convex stochastic optimization problems with states belonging to the space $L^\infty(\Omega, \mathbb{R}^n)$. As the class of problems we are treating involve equality constraints, we include that theory here, which is not covered by the papers [28, 29, 30, 31]. Additionally, we emphasize that care must be taken in our setting, where the random variables are vector-valued.

We will proceed by introducing our notation and proving essential results about subdifferentiability of convex integral functionals on the space $L^\infty(\Omega, X)$ in Section 2. The core of the paper is contained in Section 3, where we use the perturbation approach to show the existence of saddle points for a suitably tailored generalized Lagrangian. This approach allows us to look for Lagrange multipliers in the space $L^1(\Omega, X^*)$, instead of $(L^\infty(\Omega, X))^*$, and provide Karush–Kuhn–Tucker conditions for our problem. In Section 4, we show an application to PDE-constrained optimization under uncertainty. We close with some remarks in Section 5.

2 Background and Notation

Throughout, we shall employ the following notation. We assume that X is a real, reflexive, and separable space; the dual is denoted by X^* and the canonical dual pairing is written as $\langle \cdot, \cdot \rangle_{X^*,X}$. Given a set $C \subset X$, δ_C denotes the indicator function, where $\delta_C(x) = 0$ if $x \in C$ and $\delta_C(x) = \infty$ otherwise. The interior of a set C is denoted by $\operatorname{int} C$. The sum of two sets A and B with $\lambda \in \mathbb{R}$ is given by $A + \lambda B := \{a + \lambda b : a \in A, b \in B\}$. We recall that for a proper function $h : X \to (-\infty, \infty]$, the subdifferential (in the sense of convex analysis) is the set-valued operator defined by

$$\partial h : X \rightrightarrows X^* : x \mapsto \{q \in X^* : \langle q, y - x \rangle_{X^*,X} + h(x) \leq h(y) \quad \forall y \in X\}.$$

The domain of h is denoted by $\operatorname{dom}(h) := \{x \in X : h(x) < \infty\}$. Given $K \subset X$, the support function of K is denoted by

$$\sigma(K, v) := \sup_{x \in K} \langle v, x \rangle_{X^*,X} \qquad \forall v \in X^*.$$

A strongly $\mathbb{P}$-measurable mapping from Ω to a Banach space X is referred to as an X-valued random variable. As the underlying probability space is considered fixed, we will frequently write simply "measurable" instead of "$\mathbb{P}$-measurable." Additionally, since we only consider separable spaces, weak and strong measurability coincide, in which case we can simply refer to measurability of a random variable.[1]

Given a Banach space X equipped with the norm $\|\cdot\|_X$, the Bochner space $L^r(\Omega, X)$ is the set of all (equivalence classes of) X-valued random variables having finite norm, where the norm is given by

$$\|y\|_{L^r(\Omega,X)} := \begin{cases} (\int_\Omega \|y(\omega)\|_X^r \, \mathrm{d}\mathbb{P}(\omega))^{1/r}, & 1 \le r < \infty, \\ \operatorname{ess\,sup}_{\omega\in\Omega} \|y(\omega)\|_X, & r = \infty. \end{cases}$$

An X-valued random variable x is Bochner integrable if there exists a sequence $\{x_n\}$ of $\mathbb{P}$-simple functions $x_n : \Omega \to X$ such that $\lim_{n\to\infty} \int_\Omega \|x_n(\omega) - x(\omega)\|_X \, \mathrm{d}\mathbb{P}(\omega) = 0$. The limit of the integrals of x_n gives the Bochner integral (the expectation), i.e.,

$$\mathbb{E}[x] := \int_\Omega x(\omega) \, \mathrm{d}\mathbb{P}(\omega) = \lim_{n\to\infty} \int_\Omega x_n(\omega) \, \mathrm{d}\mathbb{P}(\omega).$$

Clearly, this expectation is an element of X.

Recall that a property is said to hold *almost surely* (a.s.) provided that the set (in Ω) where the property does not hold is a set of measure zero. As an example, two random variables ξ, ξ' are said to be equal almost surely, $\xi = \xi'$ a.s., if and only if $\mathbb{P}(\{\omega \in \Omega : \xi(\omega) \neq \xi'(\omega)\}) = 0$, or equivalently, $\mathbb{P}(\{\omega \in \Omega : \xi(\omega) = \xi'(\omega)\}) = 1$.

2.1 Subdifferentiability of convex integral functionals on $L^\infty(\Omega, X)$

In order to obtain optimality conditions for a problem of the form (1), we will first provide some background on convex integral functionals defined on the space $L^\infty(\Omega, X)$, where X is assumed to be a real, reflexive, and separable Banach space.[2] We denote the σ-algebra of Borel sets on X by $\mathcal{B}$. We study convex functionals of the form

$$I_f(x) := \int_\Omega f(x(\omega), \omega) \, \mathrm{d}\mathbb{P}(\omega), \tag{2}$$

where $x : \Omega \to X$ and $f : X \times \Omega \to (-\infty, \infty]$. The function f is called a convex integrand if $f_\omega := f(\cdot, \omega)$ is convex for every ω (it is no loss of generality to redefine a functional that is only convex for almost every ω). This integrand is called *normal* if it is not identically infinity, it is $(\mathcal{B} \times \mathcal{F})$-measurable, and f_ω is lower semicontinuous in X for each $\omega \in \Omega$. An example of a function that is normal is one that is finite everywhere and Carathéodory, meaning f measurable in ω for fixed x and continuous in x for fixed ω. Normality of f makes it superpositionally measurable, meaning $\omega \mapsto f(x(\omega), \omega)$ is measurable if $x : \Omega \to X$ is measurable; see, e.g., [3, Lemma 8.2.3].

If $\omega \mapsto f(x(\omega), \omega)$ is majorized by an integrable function, then the integral functional (2) is finite; if no such majorant exists, by convention, we set $I_f(x) = \infty$. The conjugate of the normal convex integrand f_ω is the function f_ω^* defined on X^* by

$$f_\omega^*(x^*) := \sup_{x\in X} \{\langle x^*, x\rangle_{X^*,X} - f_\omega(x)\}.$$

[1]More precisely, for $y : \Omega \to X$, the following assertions are equivalent: 1) y is strongly measurable and 2) y is separably-valued and measurable [16, Corollary 1.1.10].

[2]While we continue using the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the results of this section also hold for more general σ-finite complete measure spaces.

By [22, Proposition 6.1], f_ω^* is a normal convex integrand and $(f_\omega^*)^* = f_\omega$. We recall, see, e.g., [3, Proposition 6.5.4] that if f_ω is convex,

$$x^* \in \partial f_\omega(x) \quad \text{if and only if} \quad \langle x^*, x\rangle_{X^*,X} = f_\omega(x) + f_\omega^*(x^*). \tag{3}$$

Even if the Radon–Nikodym property is satisfied for X, there is not generally an isometry between $(L^\infty(\Omega, X))^*$ and $L^1(\Omega, X^*)$. However, there is a useful decomposition on this dual space; namely, elements can be decomposed into absolutely continuous and singular parts. A continuous linear functional $v \in (L^\infty(\Omega, X))^*$ of the form

$$v(x) = \int_\Omega \langle x^*(\omega), x(\omega)\rangle_{X^*,X} \,\mathrm{d}\mathbb{P}(\omega)$$

for some $x^* \in L^1(\Omega, X^*)$ is said to be absolutely continuous. These functionals form a closed subspace of $(L^\infty(\Omega, X))^*$ that is isometric to $L^1(\Omega, X^*)$. This subspace has a complement consisting of singular functionals, defined next.

Definition 2.1. A functional $v^\circ \in (L^\infty(\Omega, X))^*$ is called *singular (relative to $\mathbb{P}$)* if there exists a sequence $\{F_n\} \subset \mathcal{F}$ with $F_{n+1} \subset F_n$ for all n, $\mathbb{P}(F_n) \to 0$ as $n \to \infty$, and $v^\circ(x) = 0$ for all $x \in L^\infty(\Omega, X)$ satisfying $x(\omega) \equiv 0$ for almost all $\omega \in F_n$ for some n.

The following decomposition result was proven in [17, Appendix 1, Theorem 3].

Theorem 2.2 (Ioffe and Levin)**.** Each functional $v^* \in (L^\infty(\Omega, X))^*$ has a unique decomposition

$$v^* = v + v^\circ, \tag{4}$$

where v is absolutely continuous, v° is singular relative to $\mathbb{P}$, and

$$\|v^*\|_{(L^\infty(\Omega,X))^*} = \|v\|_{(L^\infty(\Omega,X))^*} + \|v^\circ\|_{(L^\infty(\Omega,X))^*}.$$

The next result characterizes the convex conjugate of a functional I_f defined on $L^\infty(\Omega, X)$. By definition, the convex functional on $(L^\infty(\Omega, X))^*$ that is conjugate to I_f is given by

$$I_f^*(v^*) := \sup_{z \in L^\infty(\Omega,X)} \{v^*(z) - I_f(z)\}. \tag{5}$$

This functional is closely related to the integral functional I_{f^*}, where f^* denotes the conjugate of the normal convex integrand f as before. The following theorem relates I_f^* to I_{f^*} and was proven for $X = \mathbb{R}^n$ in [25, Theorem 1] and later for separable (generally non-reflexive) Banach spaces in [22, Theorem 6.4].

Theorem 2.3 (Levin)**.** Assume f is a normal convex integrand and $I_f(x) < \infty$ for some $x \in L^\infty(\Omega, X)$. Then the functional I_f^* can be represented by the decomposition

$$I_f^*(v^*) = I_{f^*}(x^*) + \sigma(\mathrm{dom}(I_f), v^\circ), \tag{6}$$

where $x^* \in L^1(\Omega, X^*)$ corresponds to the absolutely continuous part of v^* and $v^\circ \in (L^\infty(\Omega, X))^*$ corresponds to the singular part of v^*, and $\sigma(\mathrm{dom}(I_f), v^\circ)$ denotes the support functional of $\mathrm{dom}(I_f)$ in v°.

Remark 2.4. The assumption that $I_f(x) < \infty$ for some $x \in L^\infty(\Omega, X)$ implies that I_{f^*} is a well-defined convex functional on $L^1(\Omega, X^*)$ with values in $(-\infty, \infty]$. Indeed, since f_ω^* and f_ω are conjugate to each other, we have for all ω and all $x^* \in L^1(\Omega, X^*)$ that

$$f_\omega^*(x^*(\omega)) \geq \langle x^*(\omega), x(\omega)\rangle_{X^*,X} - f_\omega(x(\omega)). \tag{7}$$

The right side is integrable by assumption, so $I_{f^*} > -\infty$ on $L^1(\Omega, X^*)$. If one additionally has $I_{f^*}(x^*) < \infty$ for some $x^* \in L^1(\Omega, X^*)$, then one shows in the same way that I_f is well-defined on $L^\infty(\Omega, X)$ with values in $(-\infty, \infty]$.

The following result gives a bound on the singular element v°.

Theorem 2.5. Let f be a normal convex integrand. Let $\bar{x} \in L^\infty(\Omega, X)$ be such that there exists $r > 0$ and an integrable function k_r of ω satisfying $f_\omega(x(\omega)) \leq k_r(\omega)$ as long as $\|x - \bar{x}\|_{L^\infty(\Omega,X)} < r$. Then the conjugate integrand $f_\omega^*(x^*(\omega))$ is majorized by an integrable function of ω for at least one $x^* \in L^1(\Omega, X^*)$. Additionally, I_f is continuous at x as long as $\|x - \bar{x}\|_{L^\infty(\Omega,X)} < r$ and the function $\sigma(\mathrm{dom}(I_f), \cdot)$ given in (6) can be bounded as follows:

$$\sigma(\mathrm{dom}(I_f), v^\circ) \geq v^\circ(\bar{x}) + r\|v^\circ\|_{(L^\infty(\Omega,X))^*}. \tag{8}$$

Proof. We proceed as in [25, Theorem 2], making modifications for the infinite-dimensional setting. Using (3), we have

$$\partial f_\omega(\bar{x}(\omega)) = \{q \in X^* \colon \langle q, \bar{x}(\omega)\rangle_{X^*,X} = f_\omega(\bar{x}(\omega)) + f_\omega^*(q)\}.$$

We show that the set-valued map $\omega \mapsto \partial f_\omega(\bar{x}(\omega))$ is measurable by first proving that the support function of $\partial f_\omega(\bar{x}(\omega))$ is measurable. Since f_ω is convex and finite on a neighborhood of $\bar{x}(\omega)$, it is continuous at $\bar{x}(\omega)$, so the set $\partial f_\omega(\bar{x}(\omega))$ is a nonempty, convex, and weakly* compact subset of X^* and f_ω is Hadamard directionally differentiable in $\bar{x}(\omega)$ [4, Proposition 2.126]. Since X is reflexive, the support function of $\partial f_\omega(\bar{x}(\omega))$ in x is given by

$$\sigma(\partial f_\omega(\bar{x}(\omega)), x) = \sup_{q \in \partial f_\omega(\bar{x}(\omega))} \langle x, q\rangle_{X,X^*}.$$

Thus, since f_ω is convex, we have

$$\begin{aligned}
\sigma(\partial f_\omega(\bar{x}(\omega)), x) &= f'_\omega(\bar{x}(\omega); x) \\
&= \lim_{t \to 0^+} \frac{1}{t}\Big(f_\omega(\bar{x}(\omega) + tx) - f_\omega(\bar{x}(\omega))\Big) \\
&= \inf_{t \geq 0} \frac{1}{t}\Big(f_\omega(\bar{x}(\omega) + tx) - f_\omega(\bar{x}(\omega))\Big) \\
&\leq f_\omega(\bar{x}(\omega) + x) - f_\omega(\bar{x}(\omega)).
\end{aligned} \tag{9}$$

Measurability of $\omega \mapsto \sigma(\partial f_\omega(\bar{x}(\omega)), x)$ follows from the fact that the limit of a sequence of measurable functions is measurable [3, p. 307]. Since X is reflexive and separable, we obtain from [3, Theorem 8.2.14] that $\omega \mapsto \partial f_\omega(\bar{x}(\omega))$ is measurable. The measurable selection theorem [3, Theorem 8.1.3] guarantees the existence of a measurable function $x^* : \Omega \to X^*$ such that $x^*(\omega) \in \partial f_\omega(\bar{x}(\omega))$ for every $\omega \in \Omega$. From (9) it follows for this x^* that

$$\langle x, x^*(\omega)\rangle_{X,X^*} \leq \sigma(\partial f_\omega(\bar{x}(\omega)), x) \leq f_\omega(\bar{x}(\omega) + x) - f_\omega(\bar{x}(\omega)).$$

As long as $x \in X$ satisfies $\|x\|_X < r$, we obtain by assumption that

$$r\|x^*(\omega)\|_{X^*} = \sup_{x:\|x\|_X \le r} \langle x, x^*(\omega)\rangle_{X,X^*} \le k_r(\omega) - f_\omega(\bar{x}(\omega)). \tag{10}$$

The right-hand side of (10) is integrable, thus x^* must also be integrable, i.e., $x^* \in L^1(\Omega, X^*)$.

Now, by (3), we have for this $x^* \in L^1(\Omega, X^*)$

$$f_\omega^*(x^*(\omega)) = \langle x^*(\omega), \bar{x}(\omega)\rangle_{X^*,X} - f_\omega(\bar{x}(\omega)),$$

from which we immediately obtain that $f_\omega^*(x^*(\omega))$ is majorizable.

For any $x \in L^\infty(\Omega, X)$ satisfying $\|x - \bar{x}\|_{L^\infty(\Omega,X)} < r$, we get

$$I_f(x) \le \int_\Omega k_r(\omega)\, \mathrm{d}\mathbb{P}(\omega) < \infty,$$

implying $I_f(x)$ is bounded above and continuous at x. Of course, this means that $x \in \mathrm{dom}(I_f)$, so

$$\begin{aligned} \sigma(\mathrm{dom}(I_f), v^\circ) = \sup_{x \in \mathrm{dom}(I_f)} v^\circ(x) &\ge \sup_{x:\|x-\bar{x}\|_{L^\infty(\Omega,X)}<r} v^\circ(x) \\ &= v^\circ(\bar{x}) + r\|v^\circ\|_{(L^\infty(\Omega,X))^*}. \end{aligned}$$

This is the expression (8), so the proof is complete. □

The next two results can be obtained as in [25, Corollary 2A, 2C].

Corollary 2.6. Assume f is a normal convex integrand and $f(x(\omega), \omega)$ is an integrable function of ω for every $x \in L^\infty(\Omega, X)$. Then I_f and I_{f^*} are well-defined convex functionals on $L^\infty(\Omega, X)$ and $L^1(\Omega, X^*)$, respectively, that are conjugate to each other in the sense that

$$\begin{aligned} I_{f^*}(x^*) &= \sup_{x \in L^\infty(\Omega,X)} \left\{ \int_\Omega \langle x^*(\omega), x(\omega)\rangle_{X^*,X}\, \mathrm{d}\mathbb{P}(\omega) - I_f(x) \right\}, \\ I_f(x) &= \sup_{x^* \in L^1(\Omega,X^*)} \left\{ \int_\Omega \langle x^*(\omega), x(\omega)\rangle_{X^*,X}\, \mathrm{d}\mathbb{P}(\omega) - I_{f^*}(x^*) \right\}. \end{aligned}$$

Furthermore, if v^* is an absolutely continuous functional corresponding to a function $x^* \in L^1(\Omega, X^*)$, then $I_f^*(v^*) = I_{f^*}(x^*)$, while $I_f^*(v^*) = \infty$ for any v^* that is not absolutely continuous.

Proof. Since $f(x(\omega), \omega)$ is integrable for all x, it is also integrable for $x \equiv 0$. Now, by [22, Theorem 5.1], this implies the existence of a $r > 0$ and integrable function k_r such that $f_\omega(0 + x) \le k_r(\omega)$ a.s. for all $x \in X$ such that $\|x\|_X \le r$. Theorem 2.5 gives the bound (8), which in combination with (6) gives the conclusion with $r = \infty$. □

Corollary 2.7. Let f and $\bar{x}$ satisfy the assumptions of Theorem 2.5. Then $v^* \in (L^\infty(\Omega, X))^*$ is an element of $\partial I_f(\bar{x})$ if and only if

$$x^*(\omega) \in \partial f_\omega(\bar{x}(\omega)) \quad \text{a.s.}, \tag{11}$$

where $x^* \in L^1(\Omega, X^*)$ corresponds to the absolutely continuous part v of v^* and the singular part v° of v^* satisfies $\sigma(\mathrm{dom}(I_f), v^\circ) = v^\circ(\bar{x})$. Moreover, $\partial I_f(\bar{x})$ can be identified with a nonempty, weakly compact subset of $L^1(\Omega, X^*)$. In particular, v^* belongs to $\partial I_f(\bar{x})$ if and only if $v^\circ \equiv 0$ and $v = x^*$ satisfies (11).

Proof. By Theorem 2.5, I_f is finite on a neighborhood of $\bar{x}$ and is continuous at $\bar{x}$; it is naturally convex by convexity of f. In particular $\partial I_f(\bar{x})$ is a nonempty, weakly* compact subset of $(L^\infty(\Omega, X))^*$.

Using (4), notice that by (3) $v^* \in \partial I_f(\bar{x})$ if and only if

$$\begin{aligned} 0 &= I_f^*(v^*) + I_f(\bar{x}) - v^*(\bar{x}) \\ &= \sup_{z \in L^\infty(\Omega,X)} \{v^\circ(z) + v(z) - I_f(z)\} + I_f(\bar{x}) - v^\circ(\bar{x}) - v(\bar{x}), \end{aligned}$$

i.e., the supremum is attained in $z = \bar{x}$. Now, by Theorem 2.3 and (7) it follows that

$$\begin{aligned} v^\circ(\bar{x}) + v(\bar{x}) - I_f(\bar{x}) &= I_f^*(v^*) \\ &= I_{f^*}(x^*) + \sigma(\mathsf{dom}(I_f), v^\circ) \\ &\geq v(\bar{x}) - I_f(\bar{x}) + \sigma(\mathsf{dom}(I_f), v^\circ) \end{aligned}$$

and thus $v^\circ(\bar{x}) \geq \sigma(\mathsf{dom}(I_f), v^\circ)$. By (8), this can be the case if and only if $v^\circ \equiv 0$. Thus using (6), we have that

$$\begin{aligned} 0 &= I_f^*(v^*) + I_f(\bar{x}) - \sigma(\mathsf{dom}(I_f), v^\circ) - \int_\Omega \langle x^*(\omega), \bar{x}(\omega)\rangle_{X^*,X} \, \mathsf{d}\mathbb{P}(\omega) \\ &= I_{f^*}(x^*) + I_f(\bar{x}) - \int_\Omega \langle x^*(\omega), \bar{x}(\omega)\rangle_{X^*,X} \, \mathsf{d}\mathbb{P}(\omega). \\ &= \int_\Omega f_\omega^*(x^*(\omega)) + f_\omega(\bar{x}(\omega)) - \langle x^*(\omega), \bar{x}(\omega)\rangle_{X^*,X} \, \mathsf{d}\mathbb{P}(\omega). \end{aligned} \tag{12}$$

Notice that the integrand in (12) is non-negative by definition of the conjugate f_ω^*, i.e., (7). We obtain that the integrand (12) is almost surely equal to zero and, recalling the equivalent expression for the subdifferential (3), (11) follows.

For the second claim, since X is reflexive and separable, we have the isometric isomorphism [16, Corollary 1.3.22]

$$(L^1(\Omega, X^*))^* \simeq L^\infty(\Omega, X^{**}) = L^\infty(\Omega, X). \tag{13}$$

Since all elements of the subdifferential in fact belong to $L^1(\Omega, X^*)$, $\partial I_f(\bar{x})$ can be identified with a subset of $L^1(\Omega, X^*)$. The fact that this subset is weakly compact follows from (13). □

3 Lagrangian Duality and Optimality Conditions

In everything that follows, we will consider the case where the admissible set of states from (1) contains both an equality and inequality (cone) constraint. Let W and R be real, reflexive, and separable Banach spaces. The equality and inequality constraint are defined by the mappings $e : X_1 \times X_2 \times \Omega \to W$ and $i : X_1 \times X_2 \times \Omega \to R$, respectively. Given a cone $K \subset R$, the partial order $\leq_K$ is defined by

$$r \leq_K 0 \quad \Leftrightarrow \quad -r \in K,$$

or equivalently, $r \geq_K 0$ if and only if $r \in K$. The corresponding dual cone is denoted by $K^\oplus := \{r^* \in R^* : \langle r^*, r\rangle_{R^*,R} \geq 0 \, \forall r \in K\}$. With that, the admissible set takes the form

$$X_{2,\mathsf{ad}}(x_1, \omega) := \{x_2 \in C_2 : e(x_1, x_2, \omega) = 0, i(x_1, x_2, \omega) \leq_K 0\}.$$

Additionally, we assume that the integrand takes the form

$$J(x_1, x_2) := J_1(x_1) + J_2(x_1, x_2). \tag{14}$$

The problem introduced in (1) is now defined over $x := (x_1, x_2) \in X := X_1 \times L^\infty(\Omega, X_2)$ by

$$\min_{x\in X} \quad \{j(x) := J_1(x_1) + \mathbb{E}[J_2(x_1, x_2(\cdot))]\}$$
$$\text{s.t.} \quad \begin{cases} x_1 \in C_1, \\ x_2(\omega) \in C_2 \text{ a.s.}, \\ e(x_1, x_2(\omega), \omega) = 0 \text{ a.s.}, \\ i(x_1, x_2(\omega), \omega) \leq_K 0 \text{ a.s.} \end{cases} \tag{P}$$

We make the following assumptions about Problem (P).

Assumption 3.1. Let $C_1 \subset X_1$ and $C_2 \subset X_2$ be nonempty, closed, and convex sets and let $K \subset R$ be a nonempty, closed, and convex cone. Assume that the integrand $(x_1, x_2) \mapsto J(x_1, x_2)$ is convex on $X_1 \times X_2$ and is everywhere defined and finite. Moreover, assume that there exist functions $a_r \colon X_1 \to \mathbb{R}$ such that

$$|J_2(x_1, x_2)| \leq a_r(x_1)$$

for all $\|x_2\|_{X_2} \leq r$. Assume $e(x_1, x_2, \omega)$ is continuous and linear in (x_1, x_2) and $i(x_1, x_2, \omega)$ is continuous and convex in (x_1, x_2); $e(x_1, x_2, \omega)$ and $i(x_1, x_2, \omega)$ are measurable and there exist functions $b_r \colon X_1 \to \mathbb{R}$ such that for all $\|x_2\|_{X_2} \leq r$ it is

$$\|e(x_1, x_2, \omega)\|_W \leq b_r(x_1),$$
$$\|i(x_1, x_2, \omega)\|_R \leq b_r(x_1).$$

Remark 3.2. By Assumption 3.1 the mappings J_2, e, and i are Carathéodory and thus for any $x_2 \in L^\infty(\Omega; X_2)$ and $x_1 \in X_1$ the mappings

$$\omega \mapsto J_2(x_1, x_2(\omega)), \quad \omega \mapsto e(x_1, x_2(\omega), \omega), \quad \omega \mapsto i(x_1, x_2(\omega), \omega)$$

are measurable, see, [3, Corollary 8.2.3]. The respective growth conditions assert that

$$J_2(x_1, x_2(\cdot)) \in L^\infty(\Omega),$$
$$e(x_1, x_2(\cdot), \cdot) \in L^\infty(\Omega, W), \quad i(x_1, x_2(\cdot), \cdot) \in L^\infty(\Omega, R).$$

For more on growth conditions, see, e.g., [2, Section 3.7].

To obtain optimality conditions, it is natural to define the Lagrangian

$$\begin{aligned}\mathbb{L}(x, \tilde\lambda) = j(x) &+ \langle \tilde\lambda_e, e(x_1, x_2(\cdot), \cdot)\rangle_{(L^\infty(\Omega,W))^*, L^\infty(\Omega,W)} \\ &+ \langle \tilde\lambda_i, i(x_1, x_2(\cdot), \cdot)\rangle_{(L^\infty(\Omega,R))^*, L^\infty(\Omega,R)}.\end{aligned}$$

However, $\tilde\lambda_e$ and $\tilde\lambda_i$ do not have natural representations in their corresponding dual spaces. We will show that under certain conditions, Lagrange multipliers can be found in the space $L^1(\Omega, W^*)$ for the equality constraint and $L^1(\Omega, R^*)$ for the inequality constraint. To this end, we will show when saddle points of a (generalized) Lagrangian exist in Subsection 3.1. This will allow us to formulate Karush–Kuhn–Tucker (KKT) conditions for Problem (P) in Subsection 3.2.

3.1 The Generalized Lagrangian and Existence of Saddle Points

In this section, we define a generalized Lagrangian and discuss the existence of saddle points for Problem (P). We will use the perturbation approach, meaning that we first introduce the perturbed problem

$$\min_{x\in X} \quad \varphi(x,u) \qquad \text{s.t.} \begin{cases} x_1 \in C_1, \\ x_2(\omega) \in C_2 \text{ a.s.}, \\ e(x_1, x_2(\omega), \omega) = u_e(\omega) \text{ a.s.}, \\ i(x_1, x_2(\omega), \omega) \leq_K u_i(\omega) \text{ a.s.} \end{cases} \tag{P^u}$$

where $\varphi(x,u) = j(x)$ if all constraints of (P^u) are fulfilled, and $\varphi(x,u) = \infty$ otherwise. We define the space of perturbations by

$$U := L^\infty(\Omega, W) \times L^\infty(\Omega, R)$$

and the space of Lagrange multipliers by

$$\Lambda := L^1(\Omega, W^*) \times L^1(\Omega, R^*).$$

These spaces can be paired for $u = (u_e, u_i) \in U$ and $\lambda = (\lambda_e, \lambda_i)$ with the bilinear form

$$\langle u, \lambda\rangle_{U,\Lambda} := \int_\Omega \langle u_e(\omega), \lambda_e(\omega)\rangle_{W,W^*} + \langle u_i(\omega), \lambda_i(\omega)\rangle_{R,R^*} \, \mathrm{d}\mathbb{P}(\omega). \tag{15}$$

The generalized Lagrangian on $X \times \Lambda$ is defined by

$$L(x,\lambda) := \inf_{u\in U} \left\{\langle u, \lambda\rangle_{U,\Lambda} + \varphi(x,u)\right\}. \tag{16}$$

Given the sets

$$\begin{aligned} X_0 &:= \{x = (x_1, x_2) \in X : x_1 \in C_1 \text{ and } x_2(\omega) \in C_2 \text{ a.s.}\}, \\ \Lambda_0 &:= \{\lambda = (\lambda_e, \lambda_i) \in \Lambda : \lambda_i(\omega) \in K^\oplus \text{ a.s.}\}, \end{aligned}$$

it is possible to show (see Appendix) that the Lagrangian takes the form

$$L(x,\lambda) = \begin{cases} J_1(x_1) + \mathbb{E}[\bar{J}_2(x_1, x_2(\cdot), \lambda(\cdot), \cdot)], & \text{if } x \in X_0, \lambda \in \Lambda_0 \\ -\infty, & \text{if } x \in X_0, \lambda \notin \Lambda_0, \\ \infty, & \text{if } x \notin X_0, \end{cases} \tag{17}$$

where

$$\begin{aligned} \bar{J}_2(x_1, x_2, \lambda, \omega) := J_2(x_1, x_2) &+ \langle \lambda_e, e(x_1, x_2, \omega)\rangle_{W^*,W} \\ &+ \langle \lambda_i, i(x_1, x_2, \omega)\rangle_{R^*,R}. \end{aligned}$$

A saddle point of L is by definition a point $(\bar{x}, \bar{\lambda}) \in X \times \Lambda$ such that

$$L(\bar{x}, \lambda) \leq L(\bar{x}, \bar{\lambda}) \leq L(x, \bar{\lambda}) \quad \forall (x,\lambda) \in X \times \Lambda. \tag{18}$$

Now, we define the dual problem

$$\max_{\lambda\in\Lambda} \left\{ g(\lambda) := \inf_{x\in X} L(x,\lambda) \right\}. \tag{D}$$

By basic duality, the question of the existence of saddle points is the same as identifying those $(\bar{x}, \bar{\lambda})$ for which the minimum of Problem (P) and maximum of Problem (D) is attained, i.e.,

$$\inf \mathsf{P} = \inf_{x\in X} \sup_{\lambda\in\Lambda} L(x,\lambda) = \sup_{\lambda\in\Lambda} \inf_{x\in X} L(x,\lambda) = \sup \mathsf{D}.$$

By the above definitions, it is clear that for all $x \in X_0$, $j(x) = \sup_{\lambda\in\Lambda} L(x,\lambda)$ and $\varphi(x,0) = j(x)$, from which we get

$$\varphi(x,u) = \sup_{\lambda\in\Lambda_0} \{L(x,\lambda) - \langle u,\lambda\rangle_{U,\Lambda}\}.$$

It is straightforward to show that L is convex in x for given $\lambda \in \Lambda_0$ and concave in λ and that φ is convex in (x,u). Moreover, $\varphi \not\equiv \infty$. It will be convenient to define $X' = X_1^* \times L^1(\Omega, X_2^*)$ and the pairing

$$\langle x, x'\rangle_{X,X'} = \langle x_1, x_1'\rangle_{X_1,X_1^*} + \int_\Omega \langle x_2(\omega), x_2'(\omega))\rangle_{X_2,X_2^*} \, \mathrm{d}\mathbb{P}(\omega). \tag{19}$$

Lemma 3.3. Let Assumption 3.1 be satisfied. Then the function $\varphi : X \times U \to \mathbb{R} \cup \{\infty\}$ is weak* lower semicontinuous.

Proof. We argue as in [29, Proposition 3]. Let $Y := X \times U$ and denote the pairing on $Z := X' \times \Lambda$ by

$$\langle y, z\rangle_{Y,Z} := \langle x, x'\rangle_{X,X'} + \langle u, \lambda\rangle_{U,\Lambda}. \tag{20}$$

Since $Y = Z^*$, the topology induced by the pairing (20) coincides with the weak* topology on Y. We define $\varphi_1(x_1) = J_1(x_1)$, if $x_1 \in C_1$ and $\varphi_1(x_1) = \infty$ if $x_1 \notin C_1$ and

$$\varphi_2(x_1,x_2,u,\omega) = \begin{cases} J_2(x_1,x_2), & \text{if } x_2(\omega) \in C_2, e(x_1,x_2,\omega) = u_e, \\ & \quad i(x_1,x_2,\omega) \le_K u_i, \\ \infty, & \text{otherwise.} \end{cases}$$

Obviously, $\varphi(x,u) = \varphi_1(x_1) + \int_\Omega \varphi_2(x_1, x_2(\omega), u(\omega), \omega) \, \mathrm{d}\mathbb{P}(\omega)$. Let $\langle \cdot,\cdot\rangle_{Y',Z'}$ denote the pairing of $Y' := X_1 \times X_2 \times (W \times R)$ with $Z' := X_1^* \times X_2^* \times (W^* \times R^*)$; then the conjugate integrand to φ_2 is given by

$$\varphi_2^*(z',\omega) = \sup_{y'\in Y'} \{\langle y', z'\rangle_{Y',Z'} - \varphi_2(y',\omega)\}.$$

Defining $h(y',\omega) = J_2(x_1,x_2)$ for $y' = (x_1,x_2,u)$ we have $h(y',\omega) \le \varphi_2(y',\omega)$ a.s. The function h is a normal convex integrand and is integrable on $X_1 \times L^\infty(\Omega, X_1) \times (L^\infty(\Omega, W) \times L^\infty(\Omega, R))$ by Assumption 3.1. Thus with the conjugate integrand h^*, I_h and I_{h^*} are conjugate to each other by Corollary 2.6, meaning that $I_{h^*} \not\equiv \infty$.

Since, $h \le \varphi_2$ we have $h^* \ge \varphi_2^*$, and hence there exists a point $z \in Z$ such that $I_{\varphi_2^*}(z) < \infty$. Since there clearly exists a point such that I_{φ_2} is finite, it follows that I_{φ_2} and $I_{\varphi_2^*}$ are conjugate to one another and are weak* lower semicontinuous, see [24, p. 227]. Since φ_1 is also weakly lower semicontinuous with respect to the natural pairing on the reflexive space X_1, φ_1 and hence φ are also weak* lower semicontinuous. □

The following result is based on [29, Theorem 3]. We define the value function

$$v(u) := \inf_{x\in X} \varphi(x,u). \tag{21}$$

Obviously, $v(0) = \inf \mathsf{P}$. For the next result, we define the second-stage admissible set by

$$X_{2,0} = \{x_2 \in L^\infty(\Omega, X_2) : x_2(\omega) \in C_2 \text{ a.s.}\}. \tag{22}$$

Theorem 3.4. Let Assumption 3.1 be satisfied. Supposing C_1 and C_2 are bounded sets, then

$$-\infty < \min \mathsf{P} = \sup \mathsf{D}.$$

Proof. We first show that $X_{2,0}$ is compact with respect to the weak* topology on $L^\infty(\Omega, X_2)$. This is argued by showing that I_h and I_{h^*} are conjugate to each other, where

$$h(x_2, \omega) := \delta_{C_2}(x_2)$$

and h^* denotes the conjugate of h. Since $C_2 \neq \emptyset$ is convex and closed, h is a normal convex integrand. It is easy to see that $h^*(0, \omega) = 0$, so in particular $I_{h^*}(0) < \infty$, meaning there exists a point where I_{h^*} is finite. Note I_h is also finite in at least one point since C_2 is nonempty. It follows that I_h and I_{h^*} are conjugate to one another, meaning that I_h is lower semicontinuous with respect to the weak* topology on $L^\infty(\Omega, X_2)$. In particular, for a weak* convergent sequence $\{y_n\} \subset X_{2,0}' := \{x_2 \in L^\infty(\Omega, X_2) : I_h(x_2) \le 0\}$ such that $y_n \rightharpoonup^* \bar{y}$ it follows that

$$\liminf_{n\to\infty} I_h(y_n) \ge I_h(\bar{y}),$$

so $\bar{y} \in X_{2,0}'$; hence, $X_{2,0}'$ is closed with respect to to the weak* topology. By definition of h, we deduce that $\bar{y}(\omega) \in C_2$ a.s. and therefore $X_{2,0}$ is also closed. Of course, $X_{2,0}$ is bounded, so $X_{2,0}$ is weak* compact, see, e.g., [10, Corollary V.4.3]. It is clear that the set C_1 is compact in X_1 with respect to the weak topology on X_1. It therefore follows that X_0 is weak* compact.

Since X_0 is weak* compact and by Lemma 3.3, φ is weak* lower semicontinuous on $X \times U$, we have for all $u \in U$ that

$$\inf_{x\in X} \varphi(x,u) = \inf_{x\in X_0} \varphi(x,u) = \min_{x\in X_0} \varphi(x,u) = v(u) > -\infty.$$

It is easy to verify $-v^*(-\lambda) = g(\lambda)$ and hence $v^{**}(u) = \sup_{\lambda\in\Lambda}\{g(\lambda) - \langle \lambda, u\rangle_{\Lambda,U}\}$. It follows that

$$v^{**}(0) = \sup_{\lambda\in\Lambda} g(\lambda) = \sup \mathsf{D}.$$

To conclude the proof, we show that v is weak* lower semicontinuous in U. Notice that the level set

$$\mathsf{lev}_\alpha \varphi = \{(x,u) \in X \times U : \varphi(x,u) \le \alpha\}$$

is weak*-closed by weak* lower semicontinuity of φ, see Lemma 3.3. Additionally, φ is finite only if $x \in X_0$, so the projection of $\mathsf{lev}_\alpha\varphi$ onto X is contained in X_0. Thus the projection of $\mathsf{lev}_\alpha\varphi$ onto U, which corresponds to the level set $\{u \in U : v(u) \le \alpha\}$, is closed in the weak* topology, from which we conclude that v is weak* and weak lower semicontinuous. Since $v > -\infty$ and v is convex and lower semicontinuous, we have that $v^{**} = v$ (cf. [4, Theorem 2.113]) and therefore

$$-\infty < \min \mathsf{P} = v(0) = v^{**}(0) = \sup \mathsf{D}.$$

□

Corollary 3.5. Let Assumption 3.1 be satisfied and j be radially unbounded, i.e., $j(x) \to \infty$ as $\|x\| \to \infty$ then

$$-\infty < \min \mathsf{P} = \sup \mathsf{D}.$$

Proof. Inspection of the proof of Theorem 3.4 shows that the only place where boundedness of C_1 and C_2 comes into play is the weak* compactness of X_0. However, if $x_0 \in X$ is an arbitrary feasible point of (P^u) then the set $N_0 := \{x \in X \mid j(x) \le j(x_0)\}$ is bounded due to radial unboundedness of j. Hence, clearly,

$$\inf_{x\in X} \varphi(x,u) = \inf_{x\in X_0\cap N_0} \varphi(x,u) = \min_{x\in X_0\cap N_0} \varphi(x,u) = v(u) > -\infty$$

holds and the proof of Theorem 3.4 can be repeated. □

Theorem 3.4 has shown that a necessary condition for the minimum to be obtained in Problem (P) is for C_1 and C_2 to be bounded sets. We will now focus on establishing sufficient conditions. Recalling Definition 2.1, let $\mathcal{S}_e$ and $\mathcal{S}_i$ denote the sets of singular functionals defined on $L^\infty(\Omega, W)$ and $L^\infty(\Omega, R)$, respectively. We define

$$\Lambda^\circ = \{\lambda^\circ = (\lambda_e^\circ, \lambda_i^\circ) \in \mathcal{S}_e \times \mathcal{S}_i\},$$
$$\Lambda_0^\circ = \{\lambda^\circ = (\lambda_e^\circ, \lambda_i^\circ) \in \Lambda^\circ : \lambda_i^\circ(y) \ge 0 \,\forall y \in L^\infty(\Omega, R) : y \ge_K 0 \text{ a.s.}\},$$

as well as

$$L^\circ(x, \lambda^\circ) = \lambda_e^\circ(e(x_1, x_2(\cdot), \cdot)) + \lambda_i^\circ(i(x_1, x_2(\cdot), \cdot)).$$

Given $\lambda^\circ \in \Lambda_0^\circ$, notice the implication

$$e(x_1, x_2(\omega), \omega) = 0, i(x_1, x_2(\omega), \omega) \le_K 0 \text{ a.s.} \Rightarrow L^\circ(x, \lambda^\circ) \le 0. \tag{23}$$

Also, from the results in Subsection 2.1, we have $(\lambda_e, \lambda_e^\circ) \in L^1(\Omega, W) \times \mathcal{S}_e \cong (L^\infty(\Omega, W))^*$ and $(\lambda_i, \lambda_i^\circ) \in L^1(\Omega, R^*) \times \mathcal{S}_i \cong (L^\infty(\Omega, R))^*$. This means that $\Lambda \times \Lambda^\circ$ characterizes the dual space $(L^\infty(\Omega, W) \times L^\infty(\Omega, R))^*$. Here, we are interested in finding conditions under which the singular part Λ° vanishes in the optimum.

With that goal in mind, we define an extension of the Lagrangian (17) for Problem (P) on the space $X \times \Lambda \times \Lambda^\circ$ via

$$\bar{L}(x, \lambda, \lambda^\circ) = \begin{cases} L(x,\lambda) + L^\circ(x, \lambda^\circ) & \text{if } x \in X_0, (\lambda, \lambda^\circ) \in \Lambda_0 \times \Lambda_0^\circ, \\ -\infty, & \text{if } x \in X_0, (\lambda, \lambda^\circ) \notin \Lambda_0 \times \Lambda_0^\circ, \\ \ \infty, & \text{if } x \notin X_0. \end{cases} \tag{24}$$

The corresponding extended dual problem is given by

$$\max_{(\lambda,\lambda^\circ)\in\Lambda\times\Lambda^\circ} \left\{ \bar{g}(\lambda, \lambda^\circ) := \inf_{x\in X} \bar{L}(x, \lambda, \lambda^\circ) \right\}. \tag{$\bar{\mathsf{D}}$}$$

Clearly, $\bar{g}(\lambda, 0) = g(\lambda)$ and thus $\sup \mathsf{D} \le \sup \bar{\mathsf{D}}$. Additionally, $\sup \bar{\mathsf{D}} \le \inf \mathsf{P}$, since by (23), we have

$$\begin{aligned} \sup_{(\lambda,\lambda^\circ)} \bar{g}(\lambda, \lambda^\circ) &= \sup_{(\lambda,\lambda^\circ)} \inf_{x\in X} \{L(x,\lambda) + L^\circ(x,\lambda^\circ)\} \\ &\le \inf_{x\in X} \sup_{(\lambda,\lambda^\circ)} \{L(x,\lambda) + L^\circ(x,\lambda^\circ)\}. \end{aligned}$$

For a sufficient condition, we introduce the induced feasible set for the first-stage variable x_1:

$$\begin{aligned} \tilde{C}_1 := \{x_1 \in X_1 \,:\, \exists x_2 \in L^\infty(\Omega, X_2) \text{ s.t. } & e(x_1, x_2(\omega), \omega) = 0 \text{ a.s.}, \\ & i(x_1, x_2(\omega), \omega) \le_K 0 \text{ a.s.}, \ x_2(\omega) \in C_2 \text{ a.s.}\} \end{aligned}$$

Problem (P) is said to satisfy the *relatively complete recourse* condition if and only if

$$C_1 \subset \tilde{C}_1. \tag{25}$$

Remark 3.6. In fact, it is possible to relax this assumption to ri $C_1 \subset \tilde{C}_1^\circ$, where ri C_1 denotes the relative interior of C_1 and $\tilde{C}_1^\circ$ represents the *singularly induced feasible set*; see [31] for more details.

Additionally, we will require a regularity condition. We call the problem *strictly feasible* if the value function v, defined in (21), satisfies

$$0 \in \operatorname{int} \operatorname{dom} v. \tag{26}$$

Remark 3.7. The condition (26) implies by [26, Theorem 18] that v is bounded above in a neighborhood of zero and is continuous at zero. Notice that $v(u) = \inf_{x \in X} \varphi(x,u)$ is only finite (and equal to $j(x)$) if the constraints are satisfied, meaning $x_1 \in C_1$ and almost surely $x_2(\omega) \in C_2$, $e(x_1, x_2(\omega), \omega) = u_e$, $i(x_1, x_2(\omega), \omega) \leq_K u_i$. This condition can therefore be thought of as an “almost sure” Slater condition.

Theorem 3.8. Let Assumption 3.1 be satisfied. Suppose the relatively complete recourse condition (25) is satisfied and Problem (P) is strictly feasible, i.e., (26) holds. Then

$$\inf \mathsf{P} = \max \mathsf{D} < \infty.$$

Proof. We modify the arguments from [30, Theorem 3] to fit our setting. By Remark 3.7, v is bounded above on a neighborhood of zero, so we have by [26, Theorem 17] that

$$\inf \mathsf{P} = \max \bar{\mathsf{D}} < \infty. \tag{27}$$

In the next step, we prove that condition (25) implies

$$\bar{g}(\lambda, \lambda^\circ) \leq g(\lambda) \quad \forall (\lambda, \lambda^\circ) \in \Lambda_0 \times \Lambda_0^\circ. \tag{28}$$

With this the proof will be complete since now,

$$\max \bar{\mathsf{D}} \leq \sup \mathsf{D} \leq \max \bar{\mathsf{D}}$$

is asserted and a solution (λ, λ°) of ($\bar{\mathsf{D}}$) gives a solution λ of (D).

To show (28), let $(\lambda, \lambda^\circ) \in \Lambda_0 \times \Lambda_0^\circ$ be arbitrary. Recalling the feasible set (22), we define

$$\ell(x_1, \lambda^\circ) = \inf_{x_2 \in X_{2,0}} L^\circ(x, \lambda^0).$$

We skip the trivial case $\bar{g}(\lambda, \lambda^\circ) = -\infty$ and now show that

$$\bar{g}(\lambda, \lambda^\circ) = \inf_{x \in X_0} \{L(x, \lambda) + \ell(x_1, \lambda^0)\}. \tag{29}$$

It is obvious that

$$\begin{aligned} &\inf_{x_2 \in X_{2,0}} \left\{ \mathbb{E}[\bar{J}_2(x_1, x_2(\cdot), \lambda(\cdot), \cdot)] + L^\circ(x, \lambda^0) \right\} \\ &\qquad \geq \inf_{x_2 \in X_{2,0}} \mathbb{E}[\bar{J}_2(x_1, x_2(\cdot), \lambda(\cdot), \cdot)] + \inf_{x_2 \in X_{2,0}} L^\circ(x, \lambda^0). \end{aligned}$$

By definition, for the functional λ_e° there exists a decreasing sequence of sets $\{F_{e,n}\} \subset \mathcal{F}_n$ such that $\mathbb{P}(F_{e,n}) \to 0$ as $n \to \infty$ and $\lambda_e^\circ(w) = 0$ for all $w \in L^\infty(\Omega, W)$ such that $w = 0$ a.s. on $F_{e,n}$. The sets $F_{i,n}$ corresponding to λ_i° are defined analogously. We define $F_n = F_{e,n} \cup F_{i,n}$ and

$$y_n(\omega) = \begin{cases} y'(\omega), & \omega \in F_n \\ y''(\omega), & \omega \notin F_n \end{cases}$$

for arbitrary $y', y'' \in X_{2,0}$. If $\omega \in F_n$, then $e(x_1, y_n(\omega), \omega) = e(x_1, y'(\omega), \omega)$ and $i(x_1, y_n(\omega), \omega) = i(x_1, y'(\omega), \omega)$, meaning that

$$\lambda_e^\circ(e(x_1, y_n(\omega), \omega)) = \lambda_e^\circ(e(x_1, y'(\omega), \omega))$$

and

$$\lambda_i^\circ(i(x_1, y_n(\omega), \omega)) = \lambda_i^\circ(i(x_1, y'(\omega), \omega)).$$

Thus, for any y', y'', and $\varepsilon > 0$, there exists an n_0 such that for $n \geq n_0$ and $x_2 = y_n$ it holds that

$$\begin{aligned}\mathbb{E}[\bar{J}_2(x_1, x_2(\cdot), \lambda(\cdot), \omega)] &+ \lambda_e^\circ(e(x_1, x_2(\cdot), \cdot)) + \lambda_i^\circ(i(x_1, x_2(\cdot), \cdot)) \\ &\leq \mathbb{E}[\bar{J}_2(x_1, y''(\cdot), \lambda(\cdot), \cdot)] + \lambda_e^\circ(e(x_1, y'(\cdot), \cdot)) + \lambda_i^\circ(i(x_1, y'(\cdot), \cdot)) + \varepsilon.\end{aligned}$$

With that, we have shown (29). We now define

$$h(x_1) = \begin{cases} \inf_{x_2 \in X_{2,0}} L(x, \lambda), & \text{if } x_1 \in C_1, \\ \infty, & \text{else} \end{cases}$$

and

$$k(x_1) = -\ell(x_1, \lambda^\circ).$$

Notice that $\bar{g}(\lambda, \lambda^\circ) = \inf_{x_1 \in X_1}\{h(x_1) - k(x_1)\}$. Additionally, $h \not\equiv \infty$ is convex and $k > -\infty$ is concave. In fact, since $\bar{g}$ is finite, k cannot be identical to ∞ and h must be proper. Therefore by Fenchel's duality theorem (cf. [3, Theorem 6.5.6]), with $h^*(v) = \sup_{x_1 \in X_1}\{\langle v, x_1\rangle_{X_1^*, X_1} - h(x_1)\}$ and $k^*(v) = \inf_{x_1 \in X_1}\{\langle v, x_1\rangle_{X_1^*, X_1} - k(x_1)\}$, we have

$$\bar{g}(\lambda, \lambda^\circ) = \max_{x_1^* \in X_1^*}\{k^*(x_1^*) - h^*(x_1^*)\}. \tag{30}$$

Let x_1^* denote the maximizer of (30), meaning $\bar{g}(\lambda, \lambda^\circ) = k^*(x_1^*) - h^*(x_1^*)$. Then by definition of h^*, we have for all $x_1 \in X_1$ that

$$h(x_1) - \langle x_1^*, x_1\rangle_{X_1^*, X_1} \geq \bar{g}(\lambda, \lambda^\circ) - k^*(x_1^*). \tag{31}$$

Likewise by definition of k and k^*, we get

$$\ell(x_1, \lambda^\circ) + \langle x_1^*, x_1\rangle_{X_1^*, X_1} \geq k^*(x_1^*).$$

It is straightforward to see that $\ell(x_1, \lambda^\circ) \leq 0$ for all $x_1 \in \tilde{C}_1$. Indeed, $x_1 \in \tilde{C}_1$ implies that there exists a $x_2 \in X_{2,0}$ satisfying $e(x_1, x_2(\omega), \omega) = 0$ and $i(x_1, x_2(\omega), \omega) \leq_K 0$ a.s. Recalling (23), we get

$$\langle x_1^*, x_1\rangle_{X_1^*, X_1} \geq k^*(x_1^*)$$

for all $x_1 \in \tilde{C}_1 \supset C_1$. From (31) we thus have for all $x_1 \in C_1$ that $h(x_1) \geq \bar{g}(\lambda, \lambda^\circ)$ holds, and hence

$$L(x, \lambda) \geq h(x_1) \geq \bar{g}(\lambda, \lambda^\circ)$$

for all $x \in X_0$ and all $(\lambda, \lambda^\circ) \in \Lambda \times \Lambda^\circ$. It follows that $g(\lambda) \geq \inf_{x \in X_0} L(x, \lambda) \geq \bar{g}(\lambda, \lambda^\circ)$ and we have shown (28) finishing the proof. □

3.2 Karush–Kuhn–Tucker Conditions

In Section 3.1, we showed that saddle points of the generalized Lagrangian exist under relatively mild assumptions. We require that the constraint sets C_1 and C_2 are bounded. Additionally, the problem must satisfy an almost sure strict feasibility condition in addition to a standard assumption in stochastic models known as a relative recourse assumption. We now turn to obtaining optimality conditions under the assumption that a saddle point exists. This leads us to the following central result.

Theorem 3.9. Let Assumption 3.1 be satisfied. Then a point $(\bar{x}, \bar{\lambda}) \in (X_1 \times L^\infty(\Omega, X_2)) \times (L^1(\Omega, W^*) \times L^1(\Omega, R^*))$ is a saddle point of the Lagrangian (17) if and only if the following conditions are satisfied:

(i) There exists a function $\rho \in L^1(\Omega, X_1^*)$ such that

$$x_1 \mapsto J_1(x_1) + \langle \mathbb{E}[\rho], x_1 \rangle_{X_1^*, X_1}$$

attains its minimum over C_1 at $\bar{x}_1$.

(ii) The function

$$\begin{aligned}(x_1, x_2) \mapsto & J_2(x_1, x_2) + \langle \bar{\lambda}_e(\omega), e(x_1, x_2, \omega) \rangle_{W^*, W} \\ & + \langle \bar{\lambda}_i(\omega), i(x_1, x_2, \omega) \rangle_{R^*, R} - \langle \rho(\omega), x_1 \rangle_{X_1^*, X_1}\end{aligned}$$

attains its minimum in $X_1 \times C_2$ at $(\bar{x}_1, \bar{x}_2(\omega))$ for almost every $\omega \in \Omega$.

(iii) It holds that $\bar{x}_1 \in C_1$ and the following conditions hold almost surely:

$$\begin{aligned} & e(\bar{x}_1, \bar{x}_2(\omega), \omega) = 0, \quad \bar{x}_2(\omega) \in C_2, \quad \bar{\lambda}_i(\omega) \in K^\oplus, \\ & i(\bar{x}_1, \bar{x}_2(\omega), \omega) \leq_K 0, \quad \langle \bar{\lambda}_i(\omega), i(\bar{x}_1, \bar{x}_2(\omega), \omega) \rangle_{R^*, R} = 0. \end{aligned}$$

The appearance of this extra Lagrange multiplier ρ in Theorem 3.9 might seem surprising; however, it is standard in two-stage stochastic optimization. It is known as a “nonanticipativity” constraint and comes from this particular setting, where the first stage variable x_1 is deterministic and the second-stage variable x_2 is random.

Proof of Theorem 3.9. We follow the arguments from [28, Section 3]. We first show that the existence of a saddle point implies condition (iii). Notice that $(\bar{x}, \bar{\lambda})$ can only be a saddle point if $(\bar{x}, \bar{\lambda}) \in X_0 \times \Lambda_0$, which immediately implies

$$\bar{x}_1 \in C_1, \quad \bar{x}_2(\omega) \in C_2 \text{ a.s.}, \quad \bar{\lambda}_i(\omega) \in K^\oplus \text{ a.s.}$$

For $\bar{x} = (\bar{x}_1, \bar{x}_2)$, we have by definition of the Lagrangian (17) that

$$\sup_{\lambda \in \Lambda_0} L(\bar{x}, \lambda) = \sup_{\lambda \in \Lambda_0} \Big\{ J_1(\bar{x}_1) + \int_\Omega \bar{J}_2(\bar{x}_1, \bar{x}_2(\omega), \lambda(\omega), \omega) \, \mathrm{d}\mathbb{P}(\omega) \Big\}.$$

We show that $\sup_{\lambda \in \Lambda_0} L(\bar{x}, \lambda) = \infty$ unless $e(\bar{x}_1, \bar{x}_2(\omega), \omega) = 0$ and $i(\bar{x}_1, \bar{x}_2(\omega), \omega)) \leq_K 0$ a.s. Indeed, suppose that the set

$$E := \{\omega \in \Omega : -i(\bar{x}_1, \bar{x}_2(\omega), \omega) \notin K\}$$

has positive probability, meaning $\mathbb{P}(E) > 0$. Then defining $\lambda_n \equiv n$ on E and $\lambda_n \equiv 0$ on $\Omega \backslash E$, one gets $\mathbb{E}[\langle \lambda_n, i(x_1, x_2(\cdot), \cdot)\rangle_{R^*,R}] \to \infty$ as $n \to \infty$. An analogous argument can be applied to the equality constraint. Now, since $\bar{\lambda}_i(\omega) \in K^{\oplus}$ and $i(\bar{x}_1, \bar{x}_2(\omega), \omega) \leq_K 0$ a.s., we have that $\langle \bar{\lambda}_i(\omega), i(\bar{x}_1, \bar{x}_2(\omega), \omega)\rangle_{R^*,R} \leq 0$ a.s. The supremum of $L(\bar{x}, \lambda)$ can therefore only be attained at $\bar{\lambda}$ if and only if $\langle \bar{\lambda}_i(\omega), i(\bar{x}_1, \bar{x}_2(\omega), \omega)\rangle_{R^*,R} = 0$ a.s. We have shown that if $(\bar{x}, \bar{\lambda})$ is a saddle point, then condition (iii) is fulfilled.

It is easy to see that conditions (i)–(iii) imply that $(\bar{x}, \bar{\lambda})$ is a saddle point. Indeed, for every $x = (x_1, x_2) \in X$, conditions (i)–(ii) imply

$$\begin{aligned}
&L(\bar{x}, \bar{\lambda}) \\
&\quad = J_1(\bar{x}_1) + \langle \mathbb{E}[\rho], \bar{x}_1\rangle_{X_1^*,X_1} + \mathbb{E}[\bar{J}_2(\bar{x}_1, \bar{x}_2(\cdot), \bar{\lambda}(\cdot), \omega) - \langle \rho(\cdot), \bar{x}_1\rangle_{X_1^*,X_1}] \\
&\quad \leq J_1(x_1) + \langle \mathbb{E}[\rho], x_1\rangle_{X_1^*,X_1} + \mathbb{E}[\bar{J}_2(x_1, x_2(\cdot), \bar{\lambda}(\cdot), \omega) - \langle \rho(\cdot), x_1\rangle_{X_1^*,X_1}] \\
&\quad = L(x, \bar{\lambda}).
\end{aligned}$$

To show that $L(\bar{x}, \lambda) \leq L(\bar{x}, \bar{\lambda})$ for all $\lambda \in \Lambda$, it is enough to show that

$$\mathbb{E}[\bar{J}_2(\bar{x}_1, \bar{x}_2(\cdot), \lambda(\cdot), \omega)] \leq \mathbb{E}[\bar{J}_2(\bar{x}_1, \bar{x}_2(\cdot), \bar{\lambda}(\cdot), \omega)] \quad \forall \lambda \in \Lambda. \tag{32}$$

Since $e(\bar{x}_1, \bar{x}_2(\omega), \omega) = 0$ and $\langle \lambda_i(\omega), i(\bar{x}_1, \bar{x}_2(\omega), \omega)\rangle_{R^*,R} \leq 0$ a.s., (32) must certainly be satisfied, since (as we argued before) the maximum of $L(\bar{x}, \lambda)$ can only be attained if

$$\langle \bar{\lambda}_i(\omega), i(\bar{x}_1, \bar{x}_2(\omega), \omega)\rangle_{R^*,R} = 0$$

a.s.

Now, for the most involved part of the proof, we show that if $(\bar{x}, \bar{\lambda})$ is a saddle point, then conditions (i) and (ii) must be satisfied. To simplify, we redefine $\bar{\lambda}_i$ so that $\bar{\lambda}_i(\omega) \geq 0$ for *all* $\omega \in \Omega$. We define

$$\begin{aligned}
&h_2(x_1, x_2, \omega) \\
&\quad = J_2(x_1, x_2) + \langle \bar{\lambda}_e, e(x_1, x_2, \omega)\rangle_{W^*,W} + \langle \bar{\lambda}_i, i(x_1, x_2, \omega)\rangle_{R^*,R}.
\end{aligned} \tag{33}$$

The function h_2 is clearly convex in X; $h_2(x_1(\omega), x_2(\omega), \omega)$ is integrable by Assumption 3.1 and the fact that $\bar{\lambda}_e \in L^1(\Omega, W^*)$ and $\bar{\lambda}_i \in L^1(\Omega, R^*)$. In particular, we get by by Corollary 2.6 that

$$H_2(x_1, x_2) := \int_\Omega h_2(x_1(\omega), x_2(\omega), \omega) \, \mathrm{d}\mathbb{P}(\omega)$$

is well-defined and finite on $L^\infty(\Omega, X_1) \times L^\infty(\Omega, X_2)$ as well as convex and continuous.

Let $\iota : X_1 \times L^\infty(\Omega, X_2) \to L^\infty(\Omega, X_1) \times L^\infty(\Omega, X_2)$ be the continuous injection, which maps elements of X_1 to the corresponding constant in $L^\infty(\Omega, X_1)$ and maps each element of $L^\infty(\Omega, X_2)$ to itself. Setting $H_1(x_1, x_2) = J_1(x_1)$ if $x \in X_0$ and $H_1(x_1, x_2) = \infty$ otherwise, we have

$$L(x, \bar{\lambda}) = H_1(x_1, x_2) + H_2(\iota(x_1, x_2)) \quad \forall x \in X_0.$$

From $L(\bar{x}, \bar{\lambda}) = \min_{x \in X_0} L(x, \bar{\lambda})$ it follows that

$$H_1(\bar{x}_1, \bar{x}_2) + H_2(\iota(\bar{x}_1, \bar{x}_2)) = \min_{(x_1,x_2)\in X_0} H_1(x_1, x_2) + H_2(\iota(x_1, x_2)).$$

By the Moreau–Rockafellar theorem (cf., e.g., [4, Theorem 2.168]) we have, where ι^* maps $(L^\infty(\Omega, X_1) \times L^\infty(\Omega, X_2))^*$ to $(X_1 \times L^\infty(\Omega, X_2))^*$,

$$0 \in \partial H_1(\bar{x}_1, \bar{x}_2) + \iota^* \partial H_2(\iota(\bar{x}_1, \bar{x}_2)).$$

In particular, there exists $q \in (L^\infty(\Omega, X_1) \times L^\infty(\Omega, X_2))^*$ such that

$$-\iota^* q \in \partial H_1(\bar{x}_1, \bar{x}_2) \quad \text{and} \quad q \in \partial H_2(\iota(\bar{x}_1, \bar{x}_2)).$$

Since h_2 satisfies the conditions of Corollary 2.7, it follows that $\partial H_2(\bar{x}_1, \bar{x}_2) \subset (L^\infty(\Omega, X_1) \times L^\infty(\Omega, X_2))^*$ consists of continuous linear functionals on $L^\infty(\Omega, X_1) \times L^\infty(\Omega, X_2)$, which can be identified with pairs $(q_1, q_2) \in L^1(\Omega, X_1^*) \times L^1(\Omega, X_2^*)$ such that

$$q(\omega) = (q_1(\omega), q_2(\omega)) \in \partial h_2(\bar{x}_1, \bar{x}_2(\omega), \omega) \quad \text{a.s.} \tag{34}$$

Notice that for $q_1^* \in L^1(\Omega, X_1^*)$, the adjoint $\iota_1^* : (L^\infty(\Omega, X_1))^* \to X_1^*$ satisfies, for any $x_1 \in X_1$,

$$\langle \iota_1^* q_1^*, x_1 \rangle_{X_1^*, X_1} = \langle q_1, \iota_1 x_1 \rangle_{L^1(\Omega, X_1^*), L^\infty(\Omega, X_1)} = \mathbb{E}[\langle q_1(\cdot), x_1 \rangle_{X_1^*, X_1}].$$

Hence $\iota^* q = (\mathbb{E}[q_1], q_2) \in X_1^* \times L^1(\Omega, X_2^*)$. Thus $-\iota^* q \in \partial H_1(\bar{x}_1, \bar{x}_2)$ can be written as

$$H_1(x_1, x_2) \geq H_1(\bar{x}_1, \bar{x}_2) - \langle \mathbb{E}[q_1], x_1 - \bar{x}_1 \rangle_{X_1^*, X_1} - \mathbb{E}[\langle q_2, x_2 - \bar{x}_2 \rangle_{X_2^*, X_2}]$$

for all $(x_1, x_2) \in X$. Recalling $H_1(x_1, x_2) = J_1(x_1)$ if $x \in X_0$, we get

$$J_1(x_1) \geq J_1(\bar{x}_1) - \langle \mathbb{E}[q_1], x_1 - \bar{x}_1 \rangle_{X_1^*, X_1} \quad \forall x_1 \in C_1 \tag{35}$$

and

$$\mathbb{E}[\langle q_2, x_2 - \bar{x}_2 \rangle_{X_2^*, X_2}] \geq 0 \quad \forall x_2 \in L^\infty(\Omega, X_2) : x_2(\omega) \in C_2 \text{ a.s.} \tag{36}$$

The expression (35) is clearly equivalent to condition (i).

We claim that (36) implies

$$\langle q_2(\omega), x_2 - \bar{x}_2(\omega) \rangle_{X_2^*, X_2} \geq 0 \quad \forall x_2 \in C_2 \text{ a.s.} \tag{37}$$

Let $\hat{C}_2$ be a countable dense subset of C_2. For $x_2 \in \hat{C}_2$, we define

$$\tilde{x}_2(\omega) := \begin{cases} x_2, & \text{if } \langle q_2(\omega), x_2 - \bar{x}_2(\omega) \rangle_{X_2^*, X_2} < 0 \\ \bar{x}_2(\omega), & \text{otherwise} \end{cases}.$$

The function $\tilde{x}_2$ is clearly in $L^\infty(\Omega, X_2)$. Since (36) holds we have

$$0 \leq \mathbb{E}[\langle q_2(\cdot), x_2 - \bar{x}_2(\cdot) \rangle_{X_2^*, X_2}] = \mathbb{E}[\min(0, \langle q_2(\cdot), x_2 - \bar{x}_2(\cdot) \rangle_{X_2^*, X_2})],$$

which gives $\langle q_2(\omega), x_2 - \bar{x}_2(\omega) \rangle_{X_2^*, X_2} \geq 0$ a.s. Since this is true for all $x_2 \in \hat{C}_2$ and $\hat{C}_2$ is countable, there exists a set $\Omega' \subset \Omega$ such that $\mathbb{P}(\Omega') = 1$ and

$$\langle q_2(\omega), x_2 - \bar{x}_2(\omega) \rangle_{X_2^*, X_2} \geq 0 \quad \forall x_2 \in \hat{C}_2 \text{ and } \forall \omega \in \Omega'.$$

Passing to the closure of $\hat{C}_2$, we get

$$\langle q_2(\omega), x_2 - \bar{x}_2(\omega) \rangle_{X_2^*, X_2} \geq 0 \quad \forall x_2 \in C_2 \text{ and } \forall \omega \in \Omega',$$

and hence we have shown (37).

Finally, (34) implies with (37) that for all $(x_1, x_2) \in X_1 \times C_2$,

$$h_2(x_1, x_2, \omega) \geq h_2(\bar{x}_1, \bar{x}_2(\omega), \omega) + \langle q_1(\omega), x_1 - \bar{x}_1 \rangle_{X_1^*, X_1} \quad \text{a.s.}$$

With the definition of h_2 given in (33), it follows that

$$\begin{aligned} &J_2(x_1, x_2) + \langle \bar{\lambda}_e, e(x_1, x_2, \omega) \rangle_{W^*, W} \\ &\qquad + \langle \bar{\lambda}_i, i(x_1, x_2, \omega) \rangle_{R^*, R} - \langle q_1(\omega), x_1 \rangle_{X_1^*, X_1} \\ &\quad \geq J_2(\bar{x}_1, \bar{x}_2(\omega)) + \langle \bar{\lambda}_e, e(\bar{x}_1, \bar{x}_2(\omega), \omega) \rangle_{W^*, W} \\ &\qquad\quad + \langle \bar{\lambda}_i, i(\bar{x}_1, \bar{x}_2(\omega), \omega) \rangle_{R^*, R} - \langle q_1(\omega), \bar{x}_1 \rangle_{X_1^*, X_1} \end{aligned} \tag{38}$$

for all $(x_1, x_2) \in X_1 \times C_2$. The inequality (38) is clearly equivalent to condition (ii) with $\rho(\omega) := q_1(\omega)$.

4 Model Problem with Almost Sure State Constraints

Before we proceed to a concrete example, we will discuss a particular class of problems that will help us in verifying the measurability requirements posed in Assumption 3.1. Let $\mathcal{L}(Y,W)$ denote the space of all bounded linear operators from Y to W. A random linear operator $\mathcal{A}:\Omega\to\mathcal{L}(Y,W)$ is called strongly measurable if for all $y\in Y$ the W-valued random variable $\omega\mapsto\mathcal{A}(\omega)y$ is strongly measurable. Let $\mathcal{A}:\Omega\to\mathcal{L}(Y,W)$, $\mathcal{B}:\Omega\to\mathcal{L}(X_1,W)$, and $g:\Omega\to W$ be (strongly) measurable random operators. We consider the random linear operator equation

$$\mathcal{A}(\omega)y=\mathcal{B}(\omega)x_1+g(\omega). \tag{39}$$

The inverse and adjoint operators are to be understood in the "almost sure" sense; e.g., for $\mathcal{B}$, the adjoint operator is the random operator $\mathcal{B}^*$ such that for all $(x_1,w^*)\in X_1\times W^*$,

$$\mathbb{P}(\{\omega\in\Omega:\langle w^*,\mathcal{B}(\omega)x_1\rangle_{W^*,W}=\langle\mathcal{B}^*(\omega)w^*,x_1\rangle_{X_1^*,X_1}\})=1.$$

The following theorem will help us verify measurability in the application.

Theorem 4.1 (Hans [15])**.** Let $\mathcal{A}:\Omega\to\mathcal{L}(Y,W)$. Then $\mathcal{A}(\omega)$ is invertible a.s if and only if ran$(\mathcal{A}^*(\omega))=Y^*$ a.s. If these conditions are satisfied, then $\mathcal{A}^*(\omega)$ is invertible and $(\mathcal{A}^*(\omega))^{-1}=(\mathcal{A}^{-1}(\omega))^*$. Moreover, if any of the operators $\mathcal{A}(\omega)$, $\mathcal{A}^{-1}(\omega)$, $\mathcal{A}^*(\omega)$, $(\mathcal{A}^{-1}(\omega))^*$ is measurable, then all four operators are measurable.

If $\mathcal{A}(\omega)\in\mathcal{L}(Y,W)$ is a linear isomorphism for almost every ω, then $\mathcal{A}(\omega)$ is invertible and $\mathcal{A}^{-1}(\omega)\in\mathcal{L}(W,Y)$. The existence and uniqueness of the solution to (39), given by

$$y(\omega)=\mathcal{A}^{-1}(\omega)(\mathcal{B}(\omega)u+g(\omega))\in Y,$$

follows. By Theorem 4.1, $\mathcal{A}^{-1}(\omega)$ is measurable, hence y is strongly measurable as a product of strongly measurable functions; see [16, Proposition 1.1.28, Corollary 1.1.28].

4.1 Example

Let $D\subset\mathbb{R}^2$ be a bounded Lipschitz domain. $W^{1,p}(D)$ denotes the (reflexive and separable) Sobolev space on D consisting of functions in $L^p(D)$ having first-order distributional derivatives also in $L^p(D)$. $W_0^{1,p}(D)$ is the subset of functions in $W_0^{1,p}(D)$ that vanish on the boundary ∂D. Additionally, $W^{-1,p}(D)$ denotes the dual space of $W_0^{1,p'}(D)$, where $1/p+1/p'=1$.

We set $X_1=L^2(D)$, $Y=W_0^{1,p}(D)$, for some suitable $p>2$, and let $C_1\subset X_1$ and $C_2\subset Y$ be nonempty, convex, and closed sets. The inner product on X_1 is denoted by $(\cdot,\cdot)_{X_1}$. Given a target $y_D\in X_1$, a constant $\alpha>0$, and a constraint $\psi\in L^\infty(\Omega,Y)$, the problem is

$$\min_{(x_1,y)\in X_1\times L^\infty(\Omega,Y)}\quad \frac{1}{2}\mathbb{E}\left[\|y-y_D\|_{X_1}^2\right]+\frac{\alpha}{2}\|x_1\|_{X_1}^2$$
$$\text{s.t.}\begin{cases} x_1\in C_1, & \\ y(\cdot,\omega)\in C_2 \text{ a.s.}, & \\ -\nabla\cdot(a(s,\omega)\nabla y(s,\omega))=x_1(s)+g(s,\omega) & \text{on } D\times\Omega \text{ a.e.},\\ y(s,\omega)=0 & \text{on } \partial D\times\Omega \text{ a.e.},\\ y(s,\omega)\le\psi(s,\omega) & \text{on } D\times\Omega \text{ a.e.},\end{cases} \tag{P'}$$

where "a.e." signifies almost everywhere in D and almost surely in Ω. We note that the solution to the PDE is a random field $y : \Omega \times D \to \mathbb{R}$; we use the shorthand $y_\omega := y(\cdot, \omega)$ to denote a single realization. The random fields $a : D \times \Omega \to \mathbb{R}$ and $g : D \times \Omega \to \mathbb{R}$ are subject to the following assumption.

Assumption 4.2. The function g satisfies $g \in L^\infty(\Omega, L^2(D))$. There exist $a_{\min}, a_{\max}$ such that $0 < a_{\min} \le a(s,\omega) \le a_{\max} < \infty$ a.e. on $D \times \Omega$. Additionally, $a \in L^\infty(\Omega, C^t(D))$ for some $t \in (0, 1]$.

It will be useful to define the (self-adjoint) operators

$$\mathcal{A}(\omega)y := b_\omega(y, \cdot) \quad \text{for} \quad b_\omega(y, \phi) := \int_D a(\cdot, \omega)\nabla y \cdot \nabla \phi \,\mathrm{d}s$$

and $\mathcal{B}(\omega) := \mathrm{id}_{X_1}$. We first address the solvability of the random PDE in Problem (P$'$).

Lemma 4.3. Under Assumption 4.2, there exists $p > 2$ such that for all $x_1 \in X_1$ and almost every $\omega \in \Omega$, there exists a unique $y_\omega = y(\cdot, \omega) \in Y$. Furthermore, $y \in L^\infty(\Omega, Y)$.

Proof. Due to Assumption 4.2 and [13] there exists some $p > 2$ such that, a.s., $\mathcal{A}(\omega)\colon Y = W_0^{1,p}(D) \to W^{-1,p}(D)$ is an isomorphism and $\|\mathcal{A}^{-1}(\omega)\|_{\mathcal{L}(W_0^{1,p}(D), W^{-1,p}(D))} \le c$ for a constant c independent of ω.

Now, since $D \subset \mathbb{R}^2$, $L^2(D) \subset W^{-1,p}(D)$ for all $p < \infty$ and thus

$$y_\omega = \mathcal{A}(\omega)^{-1}(\mathcal{B}(\omega)x_1 + g(\cdot, \omega)) \in Y$$

is well-defined with $\mathcal{B}\colon L^2(D) \to L^\infty(\Omega; L^2(D))$ being the mapping to constant functions in Ω.

Clearly, it holds a.s.

$$\begin{aligned} \|y_\omega\|_Y &\le \|\mathcal{A}^{-1}(\omega)\|_{\mathcal{L}(W_0^{1,p}(D), W^{-1,p}(D))} \|\mathcal{B}(\omega)x_1 + g(\omega)\|_{W^{-1,p}(D)} \\ &\le c(\|x_1\|_{W^{-1,p}(D)} + \|g(\omega)\|_{W^{-1,p}(D)}) \\ &\le c(\|x_1\|_{L^2(D)} + \|g\|_{L^\infty(\Omega; L^2(D))}) \end{aligned}$$

Strong measurability of y follows as argued after Theorem 4.1 and thus the assertion follows. □

To obtain necessary and sufficient KKT conditions, we first note that unless the constraint $x_2(s,\omega) \le \psi(s,\omega)$ is trivially satisfied almost surely, Problem (P$'$) does not satisfy the relatively complete recourse condition (25). It therefore makes sense to modify the model to ensure that the second-stage problem is always feasible. We introduce a slack variable $z \in Y$ and constant $\alpha' > 0$; the second-stage variable is then defined by $x_2 = (y, z) \in X_2 := L^\infty(\Omega, Y) \times L^\infty(\Omega, Y)$. This modified problem is

$$\min_{x_1, x_2 \in X_1 \times X_2} \quad \frac{1}{2}\mathbb{E}\left[\|y - y_D\|_{X_1}^2 + \alpha' \|z\|_{X_1}^2\right] + \frac{\alpha}{2}\|x_1\|_{X_1}^2$$

$$\text{s.t.} \begin{cases} x_1 \in C_1, & \\ y(\cdot, \omega) \in C_2 \text{ a.s.}, & \\ z(\cdot, \omega) \in C_2 \text{ a.s.}, & \\ -\nabla \cdot (a(s,\omega)\nabla y(s,\omega)) = x_1(s) + g(s,\omega) & \text{on } D \times \Omega \text{ a.e.}, \\ y(s,\omega) = 0 & \text{on } \partial D \times \Omega \text{ a.e.}, \\ y(s,\omega) \le \psi(s,\omega) + z(s,\omega) & \text{on } D \times \Omega \text{ a.e.} \end{cases} \tag{P$'_s$}$$

It is clear that Problem (P$'_s$) now satisfies the condition (25) of relatively complete recourse if C_2 and ψ are chosen appropriately. For example, by Lemma 4.3, one immediately obtains a unique solution y to the PDE constraint where $\|y\|_{L^\infty(\Omega;X_2)} \le c$ whenever C_1 is bounded in $L^2(D)$. Then, if C_2 is a sufficiently large ball $z := \psi - y$ is again in C_2 and thus the pair (y,z) is feasible.

In this model, we have

$$\begin{aligned} J_1(x_1) &= \frac{\alpha}{2}\|x_1\|^2_{X_1}, \\ J_2(x_1,x_2) &= \frac{1}{2}\|y-y_D\|^2_{X_1} + \frac{\alpha'}{2}\|z\|^2_{X_1}, \\ e(x_1,x_2,\omega) &= \mathcal{A}(\omega)y - \mathcal{B}(\omega)x_1 - g(\cdot,\omega) \in Y^*, \\ i(x_1,x_2,\omega) &= y - \psi(\cdot,\omega) - z \in Y, \\ K &= \{y \in Y : y(s) \ge 0 \quad \text{on } D \text{ a.e.}\}. \end{aligned}$$

It is clear that Assumption 3.1 is satisfied here. Indeed, $J(x_1,x_2) = J_1(x_1) + J_2(x_1,x_2)$ is convex, everywhere defined, and continuous in $X_1 \times X_2$. The function $e(x_1,x_2,\omega)$ is linear and continuous in (x_1,x_2); measurability follows from the assumed measurability of the underlying operators. Additionally, $i(x_1,x_2,\omega)$ is linear and continuous in x_2 as well as measurable since $\psi \in L^\infty(\Omega,Y)$.

Now, we can formulate KKT conditions for Problem (P$'_s$).

Lemma 4.4. Suppose Assumption 4.2 is satisfied and C_1, C_2 are bounded. Then $(\bar{x},\bar{\lambda})$ is a saddle point of the Lagrangian (17) for Problem (P$'_s$) if and only if there exist $\rho \in L^1(\Omega,X_1^*)$, $\bar{\lambda}_e \in L^1(\Omega,Y)$, and $\bar{\lambda}_i \in L^1(\Omega,Y^*)$ such that for all $x_1 \in C_1$ and all $(y,z) \in C_2 \times C_2$,

$$(\alpha\bar{x}_1 + \mathbb{E}[\rho], x_1 - \bar{x}_1)_{X_1} \ge 0 \tag{40a}$$

$$\mathcal{B}^*(\omega)\bar{\lambda}_{e,\omega} + \rho_\omega = 0, \tag{40b}$$

$$(\bar{y}_\omega - y_D, y - \bar{y}_\omega)_{X_1} + \langle \mathcal{A}^*(\omega)\bar{\lambda}_{e,\omega} + \bar{\lambda}_{i,\omega}, y - \bar{y}_\omega\rangle_{Y^*,Y} \ge 0, \tag{40c}$$

$$(\alpha'\bar{z}_\omega, z - \bar{z}_\omega)_{X_1} - \langle \bar{\lambda}_{i,\omega}, z - \bar{z}_\omega\rangle_{Y^*,Y} \ge 0, \tag{40d}$$

$$\mathcal{A}(\omega)\bar{y}_\omega - \mathcal{B}(\omega)\bar{x}_1 - g_\omega = 0, \tag{40e}$$

$$\bar{\lambda}_{i,\omega} \in K^\oplus, \quad \bar{y}_\omega - \bar{z}_\omega \le_K \psi_\omega, \quad \langle \bar{\lambda}_{i,\omega}, \bar{y}_\omega - \bar{z}_\omega - \psi_\omega\rangle_{Y^*,Y} = 0, \tag{40f}$$

where (40b)–(40f) hold for almost all $\omega \in \Omega$. These conditions are necessary and sufficient for optimality.

Proof. We apply the optimality conditions (i)–(iii) from Theorem 3.9.

Let $f_1(x_1) := J_1(x_1) + \langle \mathbb{E}[\rho], x_1\rangle_{X_1^*,X_1}$. We recall that the optimum x_1 over C_1 is attained if and only if $\langle f_1'(\bar{x}_1), x_1 - \bar{x}_1\rangle_{X_1^*,X_1} \ge 0$ for all $x_1 \in C_1$. Hence condition (i) is equivalent to (40a). Now, we define

$$\begin{aligned} f_2(x_1,x_2,\omega) := J_2(x_1,x_2) &+ \langle \bar{\lambda}_{e,\omega}, e(x_1,x_2,\omega)\rangle_{Y,Y^*} \\ &+ \langle \bar{\lambda}_{i,\omega}, i(x_1,x_2,\omega)\rangle_{Y^*,Y} - \langle \rho_\omega, x_1\rangle_{X_1^*,X_1}. \end{aligned}$$

Now, (ii) is equivalent to stationarity of f_2 yielding (40b)–(40d). To see this, we compute

$$D_{x_1}f_2(x_1,x_2(\omega),\omega)[h] = \langle -\mathcal{B}^*(\omega)\bar{\lambda}_{e,\omega} - \rho_\omega, h\rangle_{X_1^*,X_1},$$

so $D_{x_1}f_2(\bar{x}_1,\bar{x}_2(\omega),\omega) = 0$ a.s. if and only if (40b) holds. Recalling that $x_2 = (y,z)$, we compute

$$\begin{aligned} D_y f_2(x_1,x_2(\omega),\omega)[k_1] &= (y_\omega - y_D, k_1)_{X_1} + \langle \mathcal{A}^*(\omega)\bar{\lambda}_{e,\omega} + \bar{\lambda}_{i,\omega}, k_1\rangle_{Y^*,Y}, \\ D_z f_2(x_1,x_2(\omega),\omega)[k_2] &= (\alpha' z_\omega, k_2)_{X_1} - \langle \bar{\lambda}_{i,\omega}, k_2\rangle_{Y^*,Y}, \end{aligned}$$

which at the optimum $\bar{x}_2 = (\bar{y}, \bar{z})$ over $C_2 \times C_2$ is equivalent to (40c)–(40d). Condition (iii) is clearly equivalent to (40e) and (40f).

For the final statement, it suffices to verify that Problem (P'_s) is strictly feasible. Since $p > 2$ and $D \subset \mathbb{R}^2$ is bounded, $W^{1,p}(D)$ is compactly embedded in $C(\bar{D})$. Note that $y_\omega, z_\omega \in W^{1,p}(D)$ satisfying

$$i(x_1, x_2(\omega), \omega) = y_\omega - \psi(\cdot, \omega) - z_\omega <_K 0$$

means $\eta_\omega(s) := y_\omega(s) - \psi(s, \omega) - z_\omega(s) < 0 \quad$ a.e. on $\bar{D}$. Now, the continuous function η_ω must take its maximum on the compact set $\bar{D}$, so there exists a $\varepsilon = \varepsilon(\omega)$ such that $\eta_\omega = i(x_1, x_2(\omega), \omega) < -\varepsilon$ a.e. on $\bar{D}$. If $v_\omega \in W^{1,p}(D)$ is chosen such that $\|v_\omega\|_\infty \leq \delta(\omega)$, then

$$\begin{aligned} i(x_1, x_2(\omega) + v_\omega, \omega) &= i(x_1, x_2(\omega), \omega) + v_\omega \leq -\varepsilon + \|v_\omega\|_\infty \\ &\leq -\varepsilon + \delta(\omega) \end{aligned}$$

and therefore $i(x_1, x_2(\omega), \omega) < v_\omega$ if $\delta(\omega) < \epsilon$. By Theorem 3.4 and Theorem 3.8, these conditions are necessary and sufficient. □

5 Conclusion

In this paper, we focused on obtaining necessary and sufficient first-order optimality conditions for a class of stochastic convex optimization problems. The first stage variable x_1 was assumed to belong to a reflexive and separable Banach space, and the second-stage variable x_2 was assumed to be an essentially bounded random variable having an image in a reflexive and separable Banach space. While the study of such problems in finite dimensions is classical, going back to a series of papers from the 1970s by Rockafellar and Wets, its treatment in Bochner spaces, although cursorily handled in [24, 26], was not complete enough to handle a class of problems of increasing interest, namely PDE-constrained optimization under uncertainty. In such problems, it is desirable to find a control x_1 such that a partial differential equation depending on the control is satisfied. The additional pointwise constraints on the solution to the PDE presented surprising difficulties. In order to obtain necessary and sufficient conditions for optimality, we built on the decomposition result provided by Ioffe and Levin [17], in which the Bochner space $L^\infty(\Omega, X)$ is decomposed into its absolutely continuous part and a singular part. We find that the singular part vanishes in the optimality conditions if strict feasibility and relatively complete recourse conditions are satisfied. This provides necessary and sufficient conditions for optimality with integrable Lagrange multipliers. While the example model problem we chose to illustrate the theory involved smooth functions, we remark that the optimality conditions do not require smoothness of the objective functions. Therefore we believe our theory to be applicable to more general risk-averse problems.

A Appendix

Expansion of generalized Lagrangian (17). If $x \notin X_0$, then $x \notin \operatorname{dom} \varphi(\cdot, u)$ and therefore $L(x, \lambda) = \infty$ by definition of (16).

Now we observe the case $x \in X_0$. The constraint $i(x_1, x_2(\omega), \omega) \leq_K u_i$ is equivalent to $u_i - i(x_1, x_2(\omega), \omega) \in K$. Since $x \in X_0$, φ can be redefined equivalently by

$$\varphi(x, u) := j(x) + \mathbb{E}[\delta_{\{u_e(\cdot)\}}(e, x_1, x_2(\cdot), \cdot)] + \mathbb{E}[\delta_K(u_i(\cdot) - i(x_1, x_2(\cdot), \cdot))].$$

(The equivalence is clear after one notices that the indicator function is non-negative.) Expanding (16), we get

$$L(x,\lambda) = j(x) + \inf_{u\in U} \big\{ \mathbb{E}[\delta_{\{u_e(\cdot)\}}(e(x_1,x_2(\cdot),\cdot))] \\ + \mathbb{E}[\delta_K(u_i(\cdot) - i(x_1,x_2(\cdot),\cdot))] + \langle u,\lambda\rangle_{U,\Lambda} \big\}.$$

Recalling the definition of the pairing (15), we first see that

$$\begin{aligned} \inf_{u_e\in L^\infty(\Omega,W)} \int_\Omega & \delta_{\{u_e(\omega)\}}(e(x_1,x_2(\omega),\omega)) \\ &+ \langle u_e(\omega),\lambda_e(\omega)\rangle_{W^*,W}\, \mathsf{d}\mathbb{P}(\omega) \\ = &\int_\Omega \langle e(x_1,x_2(\omega),\omega),\lambda_e(\omega)\rangle_{W^*,W}\, \mathsf{d}\mathbb{P}(\omega) \\ &+ \inf_{z\in L^\infty(\Omega,W)} \int_\Omega \delta_{\{0\}}(z(\omega)) - \langle z(\omega),\lambda_e(\omega)\rangle_{W^*,W}\, \mathsf{d}\mathbb{P}(\omega) \\ = &\int_\Omega \langle e(x_1,x_2(\omega),\omega),\lambda_e(\omega)\rangle_{W^*,W}\, \mathsf{d}\mathbb{P}(\omega) \\ &- \int_\Omega \delta^*_{\{0\}}(\lambda_e(\omega))\, \mathsf{d}\mathbb{P}(\omega) \\ = &\int_\Omega \langle e(x_1,x_2(\omega),\omega),\lambda_e(\omega)\rangle_{W^*,W}\, \mathsf{d}\mathbb{P}(\omega), \end{aligned} \tag{41}$$

where in the last step, we used that the conjugate of the indicator function is equal to the support function.

Similarly,

$$\begin{aligned} \inf_{u_i\in L^\infty(\Omega,R)} \int_\Omega & \delta_K\big(u_i(\omega) - i(x_1,x_2(\omega),\omega)\big) \\ &+ \langle u_i(\omega),\lambda_i(\omega)\rangle_{R,R^*}\, \mathsf{d}\mathbb{P}(\omega) \\ = &\int_\Omega \langle i(x_1,x_2(\omega),\omega),\lambda_i(\omega)\rangle_{R,R^*}\, \mathsf{d}\mathbb{P}(\omega) \\ &- \sup_{z\in L^\infty(\Omega,R)} \int_\Omega \delta_K(-z(\omega)) - \langle z(\omega),\lambda_i(\omega)\rangle_{R,R^*}\, \mathsf{d}\mathbb{P}(\omega) \\ = &\int_\Omega \langle i(x_1,x_2(\omega),\omega),\lambda_i(\omega)\rangle_{R,R^*} - \sup_{z'\in -K} \langle z',\lambda_i(\omega)\rangle_{R,R^*}\, \mathsf{d}\mathbb{P}(\omega) \\ = &\int_\Omega \langle i(x_1,x_2(\omega),\omega),\lambda_i(\omega)\rangle_{R,R^*} - \delta_{K^\oplus}(\lambda_i(\omega))\, \mathsf{d}\mathbb{P}(\omega). \end{aligned} \tag{42}$$

If $\lambda_i \notin K^\oplus$, then the integral is equal to $-\infty$. Otherwise, if $\lambda \in \Lambda_0$ (and $x \in X_0$), we get after combining (41) and (42) the expression

$$L(x,\lambda) = j(x) + \mathbb{E}[e(x_1,x_2(\cdot),\cdot),\lambda_e(\cdot)\rangle_{W^*,W}] + \mathbb{E}[\langle i(x_1,x_2(\cdot),\cdot),\lambda_i(\cdot)\rangle_{R,R^*}].$$

DOI 10.20347/WIAS.PREPRINT.2755

References

[1] Alen Alexanderian, Noemi Petra, Georg Stadler, and Omar Ghattas. Mean-variance risk-averse optimal control of systems governed by PDEs with random parameter fields using quadratic approximations. *SIAM/ASA J. Uncertain. Quantif.*, 5(1):1166–1192, 2017. doi: 10.1137/16M106306X.

[2] Jürgen Appell and Petr P. Zabrejko. *Nonlinear Superposition Operators*. Cambridge University Press, 1990.

[3] Jean-Pierre Aubin and Hélène Frankowska. *Set-valued analysis*. Modern Birkhäuser Classics. Birkhäuser Boston Inc., Boston, MA, 2009. doi: 10.1007/978-0-8176-4848-0. Reprint of the 1990 edition.

[4] J Frédéric Bonnans and Alexander Shapiro. *Perturbation analysis of optimization problems*. Springer Series in Operations Research. Springer-Verlag, New York, 2000. doi: 10.1007/978-1-4612-1394-9.

[5] Eduardo Casas. Control of an elliptic problem with pointwise state constraints. *SIAM J. Control Optim.*, 24(6):1309–1318, 1986. doi: 10.1137/0324078.

[6] Eduardo Casas and Joseph Frédéric Bonnans. Contrôle de systèmes elliptiques semilinéares comportant des contraintes sur l'état. In *Nonlinear Partial Differential Equations and their Applications 8*, Pitman Res. Notes Math. Ser., pages 69–86. Longman, New York, 1988.

[7] Eduardo Casas, Mariano Mateos, and Boris Vexler. New regularity results and improved error estimates for optimal control problems with state constraints. *ESAIM: Control Optim. Calc. Var.*, 20(3):803–822, 2014. doi: 10.1051/cocv/2013084.

[8] Peng Chen, Umberto Villa, and Omar Ghattas. Taylor approximation and variance reduction for PDE-constrained optimal control under uncertainty. *J. Comput. Phys.*, 385:163–186, 2019. doi: 10.1016/j.jcp.2019.01.047.

[9] Sergio Conti, Harald Held, Martin Pach, Martin Rumpf, and Rüdiger Schultz. Shape optimization under uncertainty—a stochastic programming perspective. *SIAM J. Optim.*, 19(4):1610–1632, 2008. doi: 10.1137/070702059.

[10] Nelson Dunford and Jacob T. Schwartz. *Linear Operators Part I: General Theory*, volume 7 of *Pure and Applied Mathematics*. Interscience Publishers, Inc., New York, 1957.

[11] M Hassan Farshbaf-Shaker, René Henrion, and Dietmar Hömberg. Properties of chance constraints in infinite dimensions with an application to PDE constrained optimization. *Set-Valued Var. Anal.*, 26(4):821–841, 2018. doi: 10.1007/s11228-017-0452-5.

[12] Caroline Geiersbach and Winnifried Wollner. A stochastic gradient method with mesh refinement for PDE-constrained optimization under uncertainty. Preprint 1905.08650, arXiv, 2019. URL `https://arxiv.org/pdf/1905.08650`.

[13] Konrad Gröger. A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations. *Math. Ann.*, 283(4):679–687, 1989. doi: 10.1007/BF01442860.

[14] Philipp A Guth, Vesa Kaarnioja, Frances Y Kuo, Claudia Schillings, and Ian H Sloan. A quasi-Monte Carlo method for an optimal control problem under uncertainty. Preprint 1910.10022, arXiv, 2019. URL https://arxiv.org/pdf/1910.10022.

[15] Otto Hans. Inverse and adjoint transforms of linear bounded random transforms. In *Trans. First Prague Conf. on Information Theory, Statist. Decis. Fct. and Random Processes*, pages 127–133. Publishing House of the Czechoslovak Academy of Sciences, Prague, 1957.

[16] Tuomas Hytönen, Jan Van Neerven, Mark Veraar, and Lutz Weis. *Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory*. Springer.

[17] Aleksandr Davidovich Ioffe and Vladimir L'vovich Levin. Subdifferentials of convex functions. *Trudy Moskov. Mat. Obšč.*, 26:3–73, 1972.

[18] Drew Kouri and Thomas Surowiec. Risk-averse PDE-constrained optimization using the conditional value-at-risk. *SIAM J. Optim.*, 26(1):365–396, 2016. doi: 10.1137/140954556.

[19] Drew Kouri and Thomas Surowiec. Existence and optimality conditions for risk-averse PDE-constrained optimization. *SIAM/ASA J. Uncertain. Quantif.*, 6(2):787–815, 2018. doi: 10.1137/16M1086613.

[20] Drew Kouri and Thomas Surowiec. Risk-averse optimal control of semilinear elliptic PDEs. *ESAIM: Control Optim. Calc. Var.*, 2019. doi: 10.1051/cocv/2019061.

[21] Drew Kouri, Matthias Heinkenschloss, Denis Ridzal, and Bart Van Bloemen Waanders. A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty. *SIAM J. Sci. Comput.*, 35(4):A1847–A1879, 2013. doi: 10.1137/120892362.

[22] Vladimir L'vovich Levin. Convex integral functionals and the theory of lifting. *Russian Mathematical Surveys*, 30(2):119–184, 1975. doi: 10.1070/rm1975v030n02abeh001408.

[23] Georg Ch. Pflug and Alois Pichler. *Multistage Stochastic Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, Cham, 2014. doi: 10.1007/978-3-319-08843-3.

[24] Ralph Tyrrell Rockafellar. Convex integral functionals and duality. In *Contributions to nonlinear functional analysis*, pages 215–236. Elsevier, 1971.

[25] Ralph Tyrrell Rockafellar. Integrals which are convex functionals. II. *Pacific J. Math*, 39(2):439–469, 1971. URL http://projecteuclid.org/euclid.pjm/1102969571.

[26] Ralph Tyrrell Rockafellar. *Conjugate duality and optimization*. SIAM, 1974.

[27] Ralph Tyrrell Rockafellar and Johannes Royset. Engineering decisions under risk averseness. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 1(2):04015003, 2015. doi: 10.1061/AJRUA6.0000816.

[28] Ralph Tyrrell Rockafellar and Roger J-B Wets. Stochastic convex programming: Kuhn-Tucker conditions. *J. Math. Econom.*, 2(3):349–370, 1975. doi: 10.1016/0304-4068(75)90003-8.

[29] Ralph Tyrrell Rockafellar and Roger J-B Wets. Stochastic convex programming: basic duality. *Pacific J. Math.*, 62(1):173–195, 1976. URL http://projecteuclid.org/euclid.pjm/1102867873.

[30] Ralph Tyrrell Rockafellar and Roger J-B Wets. Stochastic convex programming: singular multipliers and extended duality singular multipliers and duality. *Pacific J. Math.*, 62(2):507–522, 1976. URL http://projecteuclid.org/euclid.pjm/1102867735.

[31] Ralph Tyrrell Rockafellar and Roger J-B Wets. Stochastic convex programming: relatively complete recourse and induced feasibility. *SIAM J. Control Optim.*, 14(3):574–589, 1976. doi: 10.1137/0314038.

[32] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory*, volume 9 of *MPS/SIAM Series on Optimization*. SIAM, Philadelphia, 2009. doi: 10.1137/1.9780898718751.

[33] Andreas Van Barel and Stefan Vandewalle. Robust optimization of PDEs with random coefficients using a multilevel Monte Carlo method. In *SIAM/ASA J. Uncertain. Quantif.*, volume 7, pages 174–202, 2019. doi: 10.1137/17M1155892.