Variational Consistency of Robust Integral Functionals Induced by Empirical Measures

Qiaoyi Wang*
Chongqing Normal University
National Center for Applied Mathematics in Chongqing
Chongqing, China

April 15, 2026

Abstract

We study the variational convergence of robust integral functionals induced by empirical probability measures. We establish a generalized consistency framework where the ambiguity set is constructed using a probability metric. By replacing traditional uniform equicontinuity assumptions with inherent convexity and general probability metric domination, we prove the almost sure pointwise and uniform convergence of the robust objective and constraint functionals via convex analysis. Furthermore, using variational analysis, we demonstrate the Painlevé–Kuratowski set convergence of the empirical feasible regions and establish both optimal value convergence and the outer semicontinuity of the constrained minimizer sets.

1 Introduction

We consider a class of robust integral functionals of the form:

$$\min_{x\in\mathcal{X}} \sup_{P\in\mathcal{P}_N} \mathbb{E}_P[f(x,\omega)]. \tag{1}$$

Unlike classical stochastic programming which assumes full knowledge of the true probability distribution P^* [1], formulation (1) evaluates decisions against a family of plausible distributions, denoted by the ambiguity set $\mathcal{P}_N$. In our framework, motivated by data-driven distributionally robust optimization (DRO), $\mathcal{P}_N$ is constructed as a neighborhood centered at an empirical probability measure derived from data, defined via a suitable probability metric [2, 3, 4].

The primary contribution of this paper is the establishment of full asymptotic consistency for the empirical robust model. By exploiting the inherent convexity of the

*Corresponding author. Email: 723281265@qq.com

robust functionals and introducing metric domination, we bypass traditional ε-net techniques to prove almost sure uniform convergence. Using variational analysis [5], we completely characterize the asymptotic behavior of the constrained problem, establishing the Painlevé–Kuratowski set convergence of the empirical feasible regions and proving the exact consistency of the constrained minimizer sets.

2 Problem Formulation and Assumptions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an abstract probability space. The uncertainty in our model is captured by a random vector $\omega : \Omega \to \Xi$ defined on this probability space.

Assumption 2.1 (Space and measurability)**.** The uncertainty space $\Xi \subset \mathbb{R}^n$ is a compact metric space equipped with its Borel σ-algebra. The decision set $\mathcal{X} \subset \mathbb{R}^d$ is a nonempty, compact, and convex set.

To conduct a generalized consistency analysis, we assume the ambiguity set is defined via a probability distance $d(\cdot,\cdot)$ on $\mathscr{P}(\Xi)$ that dominates a specific class of test functions.

Assumption 2.2 (Probability Metric Domination)**.** Let the probability metric d metrize the weak topology on $\mathscr{P}(\Xi)$. There exists a linear class of functions $\mathfrak{H} \subset C_b(\Xi)$ and a seminorm $|\cdot|_{\mathfrak{H}}$ such that for any $f \in \mathfrak{H}$ and any $P, Q \in \mathscr{P}(\Xi)$:

$$\left| \int_{\Xi} f(\omega)\, P(d\omega) - \int_{\Xi} f(\omega)\, Q(d\omega) \right| \le |f|_{\mathfrak{H}}\, d(P, Q). \tag{2}$$

To invoke robust convergence guarantees from convex analysis without suffering from boundary discontinuities, we require the decision set to be embedded within a common open domain of the integrand.

Assumption 2.3 (Convex integrand with common domain)**.** Let $f : \mathbb{R}^d \times \Xi \to \mathbb{R} \cup \{+\infty\}$ be measurable. Assume the following:

(i) There exists a nonempty open convex set $U \subset \mathbb{R}^d$ such that

$$\mathcal{X} \subset U,$$

and for every $\omega \in \Xi$, the function $x \mapsto f(x, \omega)$ is finite and convex on U.

(ii) For every $x \in U$, the mapping $\omega \mapsto f(x, \omega)$ belongs to $\mathfrak{H}$.

(iii) For every fixed $x \in U$, the seminorm is finite:

$$|f(x, \cdot)|_{\mathfrak{H}} < \infty.$$

(iv) For every fixed $x \in U$,

$$\sup_{\omega \in \Xi} |f(x, \omega)| < \infty.$$

Let $\omega_1, \ldots, \omega_N$ be a sequence of independent and identically distributed (i.i.d.) random variables defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, sharing the true underlying probability measure $P^* \in \mathscr{P}(\Xi)$. Let $\hat{P}_N := \frac{1}{N}\sum_{i=1}^{N} \delta_{\omega_i}$ denote the empirical reference measure. The empirical ambiguity set of radius $\delta_N \geq 0$ is defined as:

$$\mathcal{P}_N := \left\{ P \in \mathscr{P}(\Xi) : d(P, \hat{P}_N) \leq \delta_N \right\}. \tag{3}$$

3 Statistical Consistency

In this section, we establish the statistical consistency of the empirical robust optimization model. We proceed by exploiting the convexity of the robust functionals, providing a cohesive framework deeply rooted in convex analysis.

3.1 Preliminaries on Set Convergence

To analyze the asymptotic behavior of the ambiguity sets and the feasible regions, we rely on the concept of set convergence in the sense of Painlevé–Kuratowski (see, e.g., [5, Chapter 4]).

Let (S, d_S) be a general metric space and $\{C_N\}_{N\in\mathbb{N}}$ be a sequence of subsets in S. The point-to-set distance from a point $x \in S$ to a set $C \subseteq S$ is defined as

$$\operatorname{dist}(x, C) := \inf_{y \in C} d_S(x, y).$$

By convention, if $C = \emptyset$, $\operatorname{dist}(x, C) = \infty$.

Definition 3.1 (Painlevé–Kuratowski Set Convergence)**.** The *inner limit* (or lower limit) of the sequence $\{C_N\}$, denoted by $\operatorname{Liminf}_{N\to\infty} C_N$, is the set of all limits of sequences $\{x_N\}$ such that $x_N \in C_N$ for all sufficiently large N. Equivalently, it can be characterized via the distance function:

$$\operatorname*{Liminf}_{N\to\infty} C_N = \left\{ x \in S : \limsup_{N\to\infty} \operatorname{dist}(x, C_N) = 0 \right\}.$$

The *outer limit* (or upper limit) of the sequence $\{C_N\}$, denoted by $\operatorname{Limsup}_{N\to\infty} C_N$, is the set of all cluster points of sequences $\{x_N\}$ such that $x_N \in C_N$. Equivalently:

$$\operatorname*{Limsup}_{N\to\infty} C_N = \left\{ x \in S : \liminf_{N\to\infty} \operatorname{dist}(x, C_N) = 0 \right\}.$$

The sequence $\{C_N\}$ is said to *converge* to a closed set C in the Painlevé–Kuratowski sense, written as $C_N \xrightarrow{\text{PK}} C$, if the inner and outer limits coincide with C, i.e.,

$$\operatorname*{Limsup}_{N\to\infty} C_N \subseteq C \subseteq \operatorname*{Liminf}_{N\to\infty} C_N.$$

Remark 3.2 (Application in the Present Context)**.** In our statistical consistency framework, the Painlevé–Kuratowski convergence is applied in two distinct metric spaces:

1. In the space of probability measures $S = \mathscr{P}(\Xi)$ equipped with the metric $d_S = d$, analyzing the collapse of the empirical ambiguity sets $\mathcal{P}_N \xrightarrow{\text{PK}} \{P^*\}$.

2. In the finite-dimensional decision space $S = \mathbb{R}^d$ equipped with the abstract metric $d_{\mathcal{X}}$, characterizing the asymptotic behavior of the empirical feasible regions $\Gamma_N \xrightarrow{\text{PK}} \Gamma_\infty$ and the minimizer sets.

The equivalent distance characterization in Definition 3.1 will be extensively utilized to construct converging sequences via the Slater points.

Assumption 3.3 (Sampling Scheme and Radius)**.** The sequence of observations $\omega_1, \omega_2, \ldots : \Omega \to \Xi$ is i.i.d. with common law $P^* \in \mathscr{P}(\Xi)$. Furthermore, under the metric assumption 2.2, the ambiguity radius is a deterministic sequence satisfying $\lim_{N\to\infty} \delta_N = 0$ [6].

Lemma 3.4 (Almost sure convergence of empirical laws and ambiguity sets)**.** *Under Assumptions 2.1 and 3.3, we have almost surely:*

$$\hat{P}_N \Rightarrow P^* \quad \textit{and} \quad \mathcal{P}_N \xrightarrow{\text{PK}} \{P^*\}, \tag{4}$$

where $\Rightarrow$ denotes weak convergence on $\mathscr{P}(\Xi)$.

Proof. Step 1: Weak and metric convergence of empirical measures. Let $\varphi \in C_b(\Xi)$ be any bounded continuous test function. By the strong law of large numbers,

$$\int_\Xi \varphi(\omega)\, \hat{P}_N(d\omega) = \frac{1}{N}\sum_{i=1}^N \varphi(\omega_i) \xrightarrow{a.s.} \int_\Xi \varphi(\omega)\, P^*(d\omega).$$

Because the uncertainty space $\Xi \subset \mathbb{R}^n$ is a compact metric space (Assumption 2.1), it is separable. Thus, the above convergence holds simultaneously for a countable convergence-determining class of functions in $C_b(\Xi)$, which implies weak convergence $\hat{P}_N \Rightarrow P^*$ almost surely [7]. Furthermore, since the distance d metrizes the weak topology on $\mathscr{P}(\Xi)$ (Assumption 2.2), this weak convergence mathematically dictates that $\lim_{N\to\infty} d(\hat{P}_N, P^*) = 0$ almost surely [7].

Step 2: Outer limit of the ambiguity sets. We now work on the almost-sure event where $d(\hat{P}_N, P^*) \to 0$. We first prove the outer inclusion $\operatorname{Limsup}_{N\to\infty} \mathcal{P}_N \subseteq \{P^*\}$. Let $P_N \in \mathcal{P}_N$ be an arbitrary sequence such that $P_N \Rightarrow P$. By the constraint definition of the ambiguity set $\mathcal{P}_N$ and the triangle inequality,

$$d(P_N, P^*) \le d(P_N, \hat{P}_N) + d(\hat{P}_N, P^*) \le \delta_N + d(\hat{P}_N, P^*).$$

Taking the upper limit as $N \to \infty$, since $\delta_N \to 0$ (Assumption 3.3) and $d(\hat{P}_N, P^*) \to 0$ from Step 1, we obtain $\limsup_{N\to\infty} d(P_N, P^*) \le 0$. Hence, $d(P_N, P^*) \to 0$, which implies $P_N \Rightarrow P^*$. Since limits in a metric space are unique, $P = P^*$, proving the outer inclusion.

Step 3: Inner limit of the ambiguity sets. We next prove the inner inclusion $\{P^*\} \subseteq \operatorname{Liminf}_{N\to\infty} \mathcal{P}_N$. It suffices to explicitly construct a sequence $Q_N \in \mathcal{P}_N$ such that $Q_N \Rightarrow P^*$. Choose the sequence of empirical centers, $Q_N := \hat{P}_N$. By definition of the metric, $d(\hat{P}_N, \hat{P}_N) = 0 \leq \delta_N$, which guarantees $\hat{P}_N \in \mathcal{P}_N$ for all N. From Step 1, we already established that $Q_N = \hat{P}_N \Rightarrow P^*$. Therefore, the sequence of centers constitutes a valid converging sequence within the sets, satisfying the inner inclusion.

Combining the inclusions from Step 2 and Step 3 concludes the proof that $\mathcal{P}_N \xrightarrow{\text{PK}} \{P^*\}$ almost surely. □

Define the empirical and population robust objectives respectively as:

$$\mathcal{R}_N(f)(x) := \sup_{P\in\mathcal{P}_N} \mathbb{E}_P[f(x,\omega)], \qquad \mathcal{R}_\infty(f)(x) := \mathbb{E}_{P^*}[f(x,\omega)].$$

3.2 Convexity and Uniform Convergence

Lemma 3.5 (Convexity of robust functionals)**.** *Under Assumption 2.3, both $\mathcal{R}_N(f)$ and $\mathcal{R}_\infty(f)$ are finite convex functions on U, hence continuous on U.*

Proof. By Assumption 2.3(i), $f(\cdot,\omega)$ is convex on U. For any $x, y \in U$ and $\lambda \in [0,1]$,

$$f(\lambda x + (1-\lambda)y, \omega) \leq \lambda f(x,\omega) + (1-\lambda) f(y,\omega).$$

For any $P \in \mathcal{P}_N$, integrating both sides with respect to P yields:

$$\mathbb{E}_P[f(\lambda x + (1-\lambda)y,\omega)] \leq \lambda \mathbb{E}_P[f(x,\omega)] + (1-\lambda)\mathbb{E}_P[f(y,\omega)].$$

Taking the supremum over $P \in \mathcal{P}_N$ preserves this inequality:

$$\mathcal{R}_N(f)(\lambda x + (1-\lambda)y) \leq \lambda \mathcal{R}_N(f)(x) + (1-\lambda)\mathcal{R}_N(f)(y),$$

which proves that $\mathcal{R}_N(f)$ is a convex function on U. Because f is uniformly bounded over $U \times \Xi$ (Assumption 2.3(iv)), $\mathcal{R}_N(f)$ is finite on U. Therefore, both $\mathcal{R}_N(f)$ and $\mathcal{R}_\infty(f)$ are finite convex functions on the open convex set U, and it follows that they are continuous on U. □

Lemma 3.6 (Pointwise convergence)**.** *Under Assumptions 2.2, 2.3, and 3.3, for every $x \in U$, we have almost sure pointwise convergence:*

$$\mathbb{P}\left(\lim_{N\to\infty} |\mathcal{R}_N(f)(x) - \mathcal{R}_\infty(f)(x)| = 0\right) = 1. \tag{5}$$

Proof. For any fixed $x \in U$ and any arbitrary measure $P \in \mathcal{P}_N$, applying the metric domination property (2) yields:

$$|\mathbb{E}_P[f(x,\omega)] - \mathbb{E}_{P^*}[f(x,\omega)]| \leq |f(x,\cdot)|_{\mathfrak{H}}\, d(P,P^*).$$

By the triangle inequality and the constraint definition of the ambiguity set $\mathcal{P}_N$:

$$d(P, P^*) \leq d(P, \hat{P}_N) + d(\hat{P}_N, P^*) \leq \delta_N + d(\hat{P}_N, P^*).$$

Because the seminorm $|f(x,\cdot)|_{\mathfrak{H}} < \infty$ is bounded for the fixed x (Assumption 2.3(iii)), taking the supremum over all $P \in \mathcal{P}_N$ yields:

$$|\mathcal{R}_N(f)(x) - \mathcal{R}_\infty(f)(x)| \leq |f(x,\cdot)|_{\mathfrak{H}} \big(\delta_N + d(\hat{P}_N, P^*)\big).$$

Dictated by Assumption 3.3 and Lemma 3.4, the terms δ_N and $d(\hat{P}_N, P^*)$ converge to zero almost surely. Thus, pointwise convergence holds for all $x \in U$ almost surely. □

Theorem 3.7 (Uniform convergence of empirical robust objectives)**.** *Under the stated assumptions, the sequence of robust objectives converges uniformly on the decision set $\mathcal{X}$ almost surely:*

$$\mathbb{P}\left(\lim_{N\to\infty} \sup_{x\in\mathcal{X}} |\mathcal{R}_N(f)(x) - \mathcal{R}_\infty(f)(x)| = 0\right) = 1. \tag{6}$$

Proof. By Lemma 3.5, almost surely, the sequence $\{\mathcal{R}_N(f)\}_{N=1}^{\infty}$ and the limit $\mathcal{R}_\infty(f)$ consist of finite convex functions on U. By Lemma 3.6, we have almost sure pointwise convergence:

$$\lim_{N\to\infty} \mathcal{R}_N(f)(x) = \mathcal{R}_\infty(f)(x) \quad \text{for all } x \in U. \tag{7}$$

According to the classical convex function convergence theorem [8, Theorem 10.8], pointwise convergence of a sequence of finite convex functions on an open convex set guarantees uniform convergence on any compact subset. Since convergence holds on U and U is an open convex set, uniform convergence holds on the compact subset $\mathcal{X} \subset U$. Therefore, it follows that:

$$\lim_{N\to\infty} \sup_{x\in\mathcal{X}} |\mathcal{R}_N(f)(x) - \mathcal{R}_\infty(f)(x)| = 0 \quad \text{almost surely.} \tag{8}$$

□

3.3 Unconstrained robust consistency

Define the unconstrained optimal values and minimizer sets:

$$\hat{v}_N := \inf_{x\in\mathcal{X}} \mathcal{R}_N(f)(x), \qquad v^* := \inf_{x\in\mathcal{X}} \mathcal{R}_\infty(f)(x),$$

$$\hat{X}_N := \arg\min_{x\in\mathcal{X}} \mathcal{R}_N(f)(x), \qquad X^* := \arg\min_{x\in\mathcal{X}} \mathcal{R}_\infty(f)(x).$$

Theorem 3.8 (Unconstrained robust consistency)**.** *Under the assumptions of Theorem 3.7, almost surely,*

$$\hat{v}_N \to v^*, \tag{9}$$

and

$$\operatorname*{Limsup}_{N\to\infty} \hat{X}_N \subseteq X^*. \tag{10}$$

Proof. For every $x \in \mathcal{X}$, the uniform error bound provides:

$$\mathcal{R}_\infty(f)(x) - \sup_{z\in\mathcal{X}} |\mathcal{R}_N(f)(z) - \mathcal{R}_\infty(f)(z)| \le \mathcal{R}_N(f)(x) \le \mathcal{R}_\infty(f)(x) + \sup_{z\in\mathcal{X}} |\mathcal{R}_N(f)(z) - \mathcal{R}_\infty(f)(z)|.$$

Taking the infimum over $x \in \mathcal{X}$ yields

$$v^* - \sup_{z\in\mathcal{X}} |\mathcal{R}_N(f)(z) - \mathcal{R}_\infty(f)(z)| \le \hat{v}_N \le v^* + \sup_{z\in\mathcal{X}} |\mathcal{R}_N(f)(z) - \mathcal{R}_\infty(f)(z)|.$$

By Theorem 3.7, the uniform error term tends to 0 almost surely. Hence $\hat{v}_N \to v^*$.

To prove the outer semicontinuity of minimizer sets, let $x_N \in \hat{X}_N$ and suppose that $x_N \to \bar{x} \in \mathcal{X}$ along a subsequence. Select $x^* \in X^*$ to be an arbitrary true optimal point. Because $x_N \in \hat{X}_N$ serves as the empirical minimizer, $\mathcal{R}_N(f)(x_N) = \hat{v}_N \le \mathcal{R}_N(f)(x^*)$. By selectively adding and subtracting evaluation terms:

$$\begin{aligned}
\mathcal{R}_\infty(f)(x_N) - \mathcal{R}_\infty(f)(x^*) &= \big(\mathcal{R}_\infty(f)(x_N) - \mathcal{R}_N(f)(x_N)\big) + \big(\mathcal{R}_N(f)(x_N) - \mathcal{R}_N(f)(x^*)\big) \\
&\quad + \big(\mathcal{R}_N(f)(x^*) - \mathcal{R}_\infty(f)(x^*)\big) \\
&\le |\mathcal{R}_\infty(f)(x_N) - \mathcal{R}_N(f)(x_N)| + |\mathcal{R}_N(f)(x^*) - \mathcal{R}_\infty(f)(x^*)| \\
&\le 2\sup_{z\in\mathcal{X}} |\mathcal{R}_N(f)(z) - \mathcal{R}_\infty(f)(z)| \,.
\end{aligned}$$

Taking the upper limit as $N \to \infty$, we obtain $\limsup_{N\to\infty}\big(\mathcal{R}_\infty(f)(x_N) - \mathcal{R}_\infty(f)(x^*)\big) \le 0$. Since $\mathcal{R}_\infty(f)$ is finite convex on the open convex set U, hence continuous on U, and therefore continuous on $\mathcal{X} \subset U$, passing to the exact limit along $x_N \to \bar{x}$ yields $\mathcal{R}_\infty(f)(\bar{x}) - \mathcal{R}_\infty(f)(x^*) \le 0$. Because $x^* \in X^*$, $\mathcal{R}_\infty(f)(x^*) = v^*$. Therefore, $\mathcal{R}_\infty(f)(\bar{x}) = v^*$, proving $\bar{x} \in X^*$. Hence $\operatorname{Limsup}_{N\to\infty} \hat{X}_N \subseteq X^*$. □

3.4 Constrained problem

We now systematically extend the generalized variational consistency to structurally constrained DRO problems. Let the population and empirical robust constraint functionals be denoted by $\Phi_\infty(x) := \mathcal{R}_\infty(g)(x)$ and $\Phi_N(x) := \mathcal{R}_N(g)(x)$ respectively.

Assumption 3.9 (Constraint Regularity and Slater Condition)**.** The constraint function $g(x, \omega)$ satisfies the conditions of Assumption 2.3 with the same open convex set U. Furthermore, there exists a strictly feasible point $x^\circ \in \mathcal{X}$ such that $\Phi_\infty(x^\circ) < 0$.

Define the true population and approximate empirical feasible sets [9]:

$$\Gamma_\infty := \{x \in \mathcal{X} : \Phi_\infty(x) \le 0\}, \qquad \Gamma_N := \{x \in \mathcal{X} : \Phi_N(x) \le 0\}. \tag{11}$$

Define the corresponding constrained optimal values and constrained minimizer sets:

$$v_\infty^c := \inf_{x\in\Gamma_\infty} \mathcal{R}_\infty(f)(x), \qquad \hat{v}_N^c := \inf_{x\in\Gamma_N} \mathcal{R}_N(f)(x),$$

$$X_c^* := \arg\min_{x\in\Gamma_\infty} \mathcal{R}_\infty(f)(x), \qquad \hat{X}_N^c := \arg\min_{x\in\Gamma_N} \mathcal{R}_N(f)(x).$$

Theorem 3.10 (Uniform convergence of empirical robust constraint functionals)**.** *Under Assumption 3.9, almost surely,*

$$\mathbb{P}\left(\lim_{N\to\infty} \sup_{x\in\mathcal{X}} |\Phi_N(x) - \Phi_\infty(x)| = 0\right) = 1. \tag{12}$$

Proof. Because $g(x,\omega)$ complies with the effective domain and convexity conditions specified in Assumption 2.3, the dependent constraint sequence $\Phi_N(x)$ preserves finite convexity on U. By mirroring the exact convex-analytical framework developed in Theorem 3.7, Theorem 10.8 in [8] guarantees that the uniform convergence holds on the compact subset $\mathcal{X} \subset U$. Thus, $\sup_{x\in\mathcal{X}} |\Phi_N(x) - \Phi_\infty(x)| \to 0$ almost surely. □

Theorem 3.11 (Outer convergence of empirical feasible sets)**.** *Under the assumptions of Theorem 3.10, almost surely [5],*

$$\operatorname*{Limsup}_{N\to\infty} \Gamma_N \subseteq \Gamma_\infty. \tag{13}$$

Proof. Let $x_N \in \Gamma_N$ and logically hypothesize that $x_N \to \bar{x} \in \mathcal{X}$ along a specific subsequence. Because $x_N \in \Gamma_N$, we evaluate $\Phi_N(x_N) \le 0$. Formally:

$$\begin{aligned}
\Phi_\infty(x_N) &= \Phi_\infty(x_N) - \Phi_N(x_N) + \Phi_N(x_N) \\
&\le |\Phi_\infty(x_N) - \Phi_N(x_N)| \\
&\le \sup_{x\in\mathcal{X}} |\Phi_\infty(x) - \Phi_N(x)|.
\end{aligned}$$

Applying the upper analytical limit as $N \to \infty$ and utilizing the derived uniform convergence in (12), we establish $\limsup_{N\to\infty} \Phi_\infty(x_N) \le 0$. Since Φ_∞ is finite convex on U, hence continuous on U, and therefore continuous on $\mathcal{X} \subset U$, taking the exact limit along $x_N \to \bar{x}$ ensures $\Phi_\infty(\bar{x}) \le 0$. Consequently, $\bar{x} \in \Gamma_\infty$. □

Theorem 3.12 (Inner convergence of empirical feasible sets)**.** *Under the assumptions above, almost surely,*

$$\Gamma_\infty \subseteq \operatorname*{Liminf}_{N\to\infty} \Gamma_N. \tag{14}$$

Proof. Let $x \in \Gamma_\infty$ be completely arbitrary. We systematically demonstrate that $x \in \operatorname{Liminf}_{N\to\infty} \Gamma_N$.

We first consider the scenario where $\Phi_\infty(x) < 0$. Set $\eta := -\Phi_\infty(x) > 0$. By the established uniform convergence in (12), there exists an integer N_0 such that for all

$N \geq N_0$, $\sup_{z\in\mathcal{X}} |\Phi_N(z) - \Phi_\infty(z)| < \frac{\eta}{2}$. Hence, for all $N \geq N_0$,

$$\Phi_N(x) \leq \Phi_\infty(x) + |\Phi_N(x) - \Phi_\infty(x)| < -\eta + \frac{\eta}{2} = -\frac{\eta}{2} < 0.$$

Therefore $x \in \Gamma_N$ holds for all sufficiently large N, which functionally means $x \in \operatorname{Liminf}_{N\to\infty} \Gamma_N$.

We subsequently evaluate the boundary scenario where $\Phi_\infty(x) = 0$. Let x° be the strictly feasible point designated in Assumption 3.9 satisfying $\Phi_\infty(x^\circ) < 0$. For any parameter $\lambda \in (0,1)$, define the point $x^\lambda := (1-\lambda)x + \lambda x^\circ$. Since $\mathcal{X}$ is convex, $x^\lambda \in \mathcal{X}$. By the convexity property of $\Phi_\infty(\cdot)$,

$$\Phi_\infty(x^\lambda) \leq (1-\lambda)\Phi_\infty(x) + \lambda\Phi_\infty(x^\circ) < 0.$$

Fix $\lambda \in (0,1)$. Since $\Phi_\infty(x^\lambda) < 0$, by deploying the initial phase of the proof, there exists N_λ such that $x^\lambda \in \Gamma_N$ holds for all $N \geq N_\lambda$. Hence, for all $N \geq N_\lambda$, the distance is bounded by $\operatorname{dist}(x, \Gamma_N) \leq d_{\mathcal{X}}(x, x^\lambda)$. Taking the formal upper limit with respect to N dictates $\limsup_{N\to\infty} \operatorname{dist}(x, \Gamma_N) \leq d_{\mathcal{X}}(x, x^\lambda)$. Letting $\lambda \downarrow 0$ forces $\lim_{N\to\infty} \operatorname{dist}(x, \Gamma_N) = 0$. Therefore $x \in \operatorname{Liminf}_{N\to\infty} \Gamma_N$. □

Theorem 3.13 (Constrained robust consistency)**.** *Suppose that the fundamental assumptions of Theorems 3.7, 3.10, 3.11, and 3.12 are fully satisfied. Then, almost surely,*

$$\hat{v}_N^c \to v_\infty^c, \tag{15}$$

and

$$\operatorname*{Limsup}_{N\to\infty} \hat{X}_N^c \subseteq X_c^*. \tag{16}$$

Proof. We first show the upper bound $\limsup_{N\to\infty} \hat{v}_N^c \leq v_\infty^c$. Extract an arbitrary optimal point $x^* \in X_c^* \subseteq \Gamma_\infty$. Dictated by Theorem 3.12, there exists a sequence $x_N \in \Gamma_N$ such that $x_N \to x^*$. Because $\hat{v}_N^c$ characterizes the absolute minimum of $\mathcal{R}_N(f)$ constrained over Γ_N, we evaluate $\hat{v}_N^c \leq \mathcal{R}_N(f)(x_N)$. Therefore

$$\begin{aligned}
\limsup_{N\to\infty} \hat{v}_N^c &\leq \limsup_{N\to\infty} \mathcal{R}_N(f)(x_N) \\
&\leq \limsup_{N\to\infty} \left(\mathcal{R}_\infty(f)(x_N) + \sup_{z\in\mathcal{X}} |\mathcal{R}_N(f)(z) - \mathcal{R}_\infty(f)(z)| \right) \\
&= \mathcal{R}_\infty(f)(x^*) = v_\infty^c.
\end{aligned}$$

This completes the confirmation that $\limsup_{N\to\infty} \hat{v}_N^c \leq v_\infty^c$.

Next, we establish the lower bound $v_\infty^c \leq \liminf_{N\to\infty} \hat{v}_N^c$. Select a subsequence $\{N_m\}$ defined to precisely capture $\lim_{m\to\infty} \hat{v}_{N_m}^c = \liminf_{N\to\infty} \hat{v}_N^c$. For each specific sequence index m, extract an optimizer $\hat{x}_{N_m} \in \hat{X}_{N_m}^c \subseteq \Gamma_{N_m}$ satisfying $\mathcal{R}_{N_m}(f)(\hat{x}_{N_m}) = \hat{v}_{N_m}^c$. Because the operational set $\mathcal{X}$ maintains compactness, passing to a further sequential subsequence if necessary, we assume that $\hat{x}_{N_m} \to \bar{x} \in \mathcal{X}$. Since $\hat{x}_{N_m} \in \Gamma_{N_m}$, Theorem 3.11

implies that $\bar{x} \in \Gamma_\infty$. Consequently, governed by the baseline definition of v^c_∞, we establish $v^c_\infty \leq \mathcal{R}_\infty(f)(\bar{x})$. Furthermore, evaluating the discrepancy limits:

$$\begin{aligned}\left|\hat{v}^c_{N_m} - \mathcal{R}_\infty(f)(\bar{x})\right| &= |\mathcal{R}_{N_m}(f)(\hat{x}_{N_m}) - \mathcal{R}_\infty(f)(\bar{x})| \\ &\leq |\mathcal{R}_{N_m}(f)(\hat{x}_{N_m}) - \mathcal{R}_\infty(f)(\hat{x}_{N_m})| + |\mathcal{R}_\infty(f)(\hat{x}_{N_m}) - \mathcal{R}_\infty(f)(\bar{x})| \\ &\leq \sup_{z\in\mathcal{X}} |\mathcal{R}_{N_m}(f)(z) - \mathcal{R}_\infty(f)(z)| + |\mathcal{R}_\infty(f)(\hat{x}_{N_m}) - \mathcal{R}_\infty(f)(\bar{x})| \,.\end{aligned}$$

The bounding right-hand side tends to 0, validating $\hat{v}^c_{N_m} \to \mathcal{R}_\infty(f)(\bar{x})$. Integrating this with $v^c_\infty \leq \mathcal{R}_\infty(f)(\bar{x})$, we obtain $v^c_\infty \leq \liminf_{N\to\infty} \hat{v}^c_N$. The amalgamation with the upper bound validates equation (15).

Finally, we verify outer semicontinuity via strict separation. We initiate the proof that $\operatorname{Limsup}_{N\to\infty} \hat{X}^c_N \subseteq X^*_c$. Suppose, to investigate for contradiction, that this specific inclusion fails. In this scenario, there exists a threshold $\varepsilon_0 > 0$, an infinite subsequence, and a sequence of optimizers $\hat{x}_N \in \hat{X}^c_N$ guaranteeing

$$\operatorname{dist}(\hat{x}_N, X^*_c) \geq \varepsilon_0 \qquad \forall N.$$

Because the overarching spatial domain $\mathcal{X}$ is compact, advancing to a refined subsequence, we assert $\hat{x}_N \to \bar{x} \in \mathcal{X}$. Since the sequence adheres to $\hat{x}_N \in \Gamma_N$, Theorem 3.11 mandates that $\bar{x} \in \Gamma_\infty$. Furthermore, driven by the absolute continuity of the minimum distance metric, $\operatorname{dist}(\bar{x}, X^*_c) \geq \varepsilon_0$, which signifies $\bar{x} \notin X^*_c$.

Because the accumulated limit $\bar{x} \in \Gamma_\infty$ remains bounded away from the theoretical optimal solution boundary X^*_c, and the functional surface $\mathcal{R}_\infty(f)$ asserts continuity across the compact operational set $\{x \in \Gamma_\infty : \operatorname{dist}(x, X^*_c) \geq \varepsilon_0\}$, the operational mechanics necessitate achieving a minimum superseding the theoretical optimal tier v^c_∞. Consequently, the structural landscape dictates the existence of a positive gap $\delta > 0$ ensuring

$$\mathcal{R}_\infty(f)(\bar{x}) \geq v^c_\infty + \delta.$$

Concurrently, since the empirical sequence inherently obeys $\hat{x}_N \in \hat{X}^c_N$, we evaluate the optimal baseline $\mathcal{R}_N(f)(\hat{x}_N) = \hat{v}^c_N$. Employing the verified framework of robust objective uniform convergence (Theorem 3.7) paired directly with the continuous structural surface of $\mathcal{R}_\infty(f)$, we establish the bounded difference:

$$\begin{aligned}|\mathcal{R}_N(f)(\hat{x}_N) - \mathcal{R}_\infty(f)(\bar{x})| &\leq |\mathcal{R}_N(f)(\hat{x}_N) - \mathcal{R}_\infty(f)(\hat{x}_N)| + |\mathcal{R}_\infty(f)(\hat{x}_N) - \mathcal{R}_\infty(f)(\bar{x})| \\ &\leq \sup_{z\in\mathcal{X}} |\mathcal{R}_N(f)(z) - \mathcal{R}_\infty(f)(z)| + |\mathcal{R}_\infty(f)(\hat{x}_N) - \mathcal{R}_\infty(f)(\bar{x})| \,.\end{aligned}$$

As $N \to \infty$, the primary deviation vector vanishes uniformly almost surely due to the baseline structural convergence algorithms, and the secondary deviation vector simultaneously converges to zero because $\hat{x}_N \to \bar{x}$ alongside an inherently continuous environment $\mathcal{R}_\infty(f)$. Thus,

$$\lim_{N\to\infty} \mathcal{R}_N(f)(\hat{x}_N) = \mathcal{R}_\infty(f)(\bar{x}).$$

However, as explicitly derived and locked in the preliminary analytical frameworks, $\lim_{N\to\infty} \mathcal{R}_N(f)(\hat{x}_N) = \lim_{N\to\infty} \hat{v}_N^c = v_\infty^c$. This intersection mandates that $v_\infty^c = \mathcal{R}_\infty(f)(\bar{x}) \geq v_\infty^c + \delta$, which introduces a logical contradiction given the established constraint $\delta > 0$.

Therefore, our foundational contradiction assumption is compromised and rendered logically invalid, cementing the conclusion that the inclusion $\mathrm{Limsup}_{N\to\infty} \hat{X}_N^c \subseteq X_c^*$ holds. □

Corollary 3.14 (One-sided distance convergence of constrained minimizers)**.** *Under the fundamental structural assumptions establishing Theorem 3.13, almost surely,*

$$\sup_{x\in\hat{X}_N^c} \mathrm{dist}(x, X_c^*) \to 0.$$

Proof. Suppose, to analyze the contrary possibility, that the theoretical framework encounters a localized anomaly establishing $\varepsilon_0 > 0$, a specific analytical subsequence, and a localized point cluster $x_N \in \hat{X}_N^c$ actively ensuring

$$\mathrm{dist}(x_N, X_c^*) \geq \varepsilon_0 \qquad \forall N.$$

Because the decision space $\mathcal{X}$ is compact, refining the analysis to a highly localized further subsequence if required, we fundamentally assert the limit $x_N \to \bar{x} \in \mathcal{X}$. By Theorem 3.13, every isolated theoretical cluster point inherent to $\hat{X}_N^c$ falls into the optimal domain X_c^*. Hence the condition asserts $\bar{x} \in X_c^*$.

Since the optimal target domain X_c^* represents a structurally closed environment and $\bar{x} \in X_c^*$, the definition of the point-to-set distance dictates that $\mathrm{dist}(x_N, X_c^*) \leq d_{\mathcal{X}}(x_N, \bar{x})$, where $d_{\mathcal{X}}$ denotes the abstract metric equipped on the decision space $\mathcal{X}$ **(cf. Section 3.1)**. Because the sequence limits to $\bar{x}$, the metric distance compresses: $d_{\mathcal{X}}(x_N, \bar{x}) \to 0$. This forces $\mathrm{dist}(x_N, X_c^*) \to 0$, which explicitly contradicts the initial bounded anomaly $\mathrm{dist}(x_N, X_c^*) \geq \varepsilon_0$. Therefore, we conclusively prove $\sup_{x\in\hat{X}_N^c} \mathrm{dist}(x, X_c^*) \to 0$. □

3.5 Extension to Multiple Constraints

The consistency results established above can be seamlessly extended to problems with multiple robust constraints. Suppose the optimization problem is subject to a finite collection of constraint functions $g_j(x, \omega)$ for $j = 1, \ldots, J$. Define the population and empirical robust constraint functionals respectively as

$$\Phi_{\infty,j}(x) := \mathcal{R}_\infty(g_j)(x) \quad \text{and} \quad \Phi_{N,j}(x) := \mathcal{R}_N(g_j)(x), \qquad j = 1, \ldots, J.$$

The corresponding population and empirical feasible sets are given by the intersection of individual constraint sets:

$$\Gamma_\infty := \bigcap_{j=1}^{J} \{x \in \mathcal{X} : \Phi_{\infty,j}(x) \le 0\}, \qquad \Gamma_N := \bigcap_{j=1}^{J} \{x \in \mathcal{X} : \Phi_{N,j}(x) \le 0\}.$$

To guarantee the topological regularity of the finite intersection, the Slater condition must be enforced jointly across all constraints.

Assumption 3.15 (Joint Slater Condition)**.** Each constraint function $g_j(x, \omega)$ independently satisfies the conditions of Assumption 2.3 with the same open convex set U. Furthermore, there exists a common strictly feasible point $x^\circ \in \mathcal{X}$ such that

$$\Phi_{\infty,j}(x^\circ) < 0, \qquad \forall j = 1, \dots, J.$$

Theorem 3.16 (Consistency under multiple constraints)**.** *Under Assumption 3.15, almost surely, the empirical feasible set converges to the population feasible set in the Painlevé–Kuratowski sense:*

$$\Gamma_N \xrightarrow{\text{PK}} \Gamma_\infty.$$

Consequently, the constrained optimal value converges ($\hat{v}_N^c \to v_\infty^c$), the empirical minimizer set is outer semicontinuous ($\operatorname{Limsup}_{N\to\infty} \hat{X}_N^c \subseteq X_c^*$), and the one-sided distance converges ($\sup_{x\in\hat{X}_N^c} \operatorname{dist}(x, X_c^*) \to 0$).

Proof. The proof for the set convergence $\Gamma_N \xrightarrow{\text{PK}} \Gamma_\infty$ follows by systematically replacing the single constraint condition in Theorems 3.11 and 3.12 with the finite intersection of multiple constraints. Specifically, for the outer limit ($\operatorname{Limsup}_{N\to\infty} \Gamma_N \subseteq \Gamma_\infty$), define the bound:

$$\epsilon_N := \max_{1\le j\le J} \sup_{x\in\mathcal{X}} |\Phi_{N,j}(x) - \Phi_{\infty,j}(x)|.$$

Since the uniform convergence holds individually for each $j = 1, \dots, J$, it follows that $\epsilon_N \to 0$ almost surely. Let $x_N \in \Gamma_N$ be a sequence such that $x_N \to \bar{x}$. For each j, we have $\Phi_{\infty,j}(x_N) \le \Phi_{\infty,j}(x_N) - \Phi_{N,j}(x_N) + \Phi_{N,j}(x_N) \le \epsilon_N$. Taking the limit ensures $\Phi_{\infty,j}(\bar{x}) \le 0$ for all j, confirming $\bar{x} \in \Gamma_\infty$. For the inner limit ($\Gamma_\infty \subseteq \operatorname{Liminf}_{N\to\infty} \Gamma_N$), the joint Slater condition guarantees that the convex combination $x^\lambda := (1-\lambda)x + \lambda x^\circ$ strictly satisfies all empirical constraints $\Phi_{N,j}(x^\lambda) < 0$ simultaneously for sufficiently large N.

Crucially, once the topological set convergence $\Gamma_N \xrightarrow{\text{PK}} \Gamma_\infty$ is established, the asymptotic properties of the optimization problem become independent of the specific functional forms defining the feasible sets. Therefore, the proofs detailed in Theorem 3.13 and Corollary 3.14 apply directly and identically to the multiple-constraint setting, yielding the convergence of the optimal value and the minimizer sets without any further modification. □

References

[1] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming.* SIAM, 2009.

[2] P. Mohajerin Esfahani and D. Kuhn. Data-driven distributionally robust optimization using the wasserstein metric. *Mathematical Programming*, 171(1-2):115–166, 2018.

[3] R. Gao and A. J. Kleywegt. Distributionally robust stochastic optimization with wasserstein distance. *Mathematics of Operations Research*, 48(2):603–655, 2023.

[4] D. Kuhn, P. M. Esfahani, V. A. Nguyen, and S. Shafieezadeh-Abadeh. Wasserstein distributionally robust optimization: Theory and applications in machine learning. *Operations Research & Management Science in the Age of Analytics*, pages 130–166, 2019.

[5] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis.* Springer, 1998.

[6] N. Fournier and A. Guillin. On the rate of convergence in wasserstein distance. *Probability Theory and Related Fields*, 162:707–738, 2015.

[7] P. Billingsley. *Convergence of Probability Measures.* Wiley, 1999.

[8] R. Tyrrell Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, NJ, 1970.

[9] J. Dupačová. Stability in stochastic programming with recourse—contaminated distributions. *Mathematical Programming*, 46(1-3):507–520, 1990.