GENERALIZED WASSERSTEIN DISTANCE AND WEAK CONVERGENCE OF SUBLINEAR EXPECTATIONS

XINPENG LI AND YIQING LIN

Abstract. In this paper, we define the generalized Wasserstein distance for sets of Borel probability measures and demonstrate that the weak convergence of sublinear expectations can be characterized by means of this distance.

1. Introduction

In classical probability theory, the Wasserstein distance of order $p$ between two Borel probability measures $\mu$ and $\nu$ in the Polish metric space $(\Omega, d)$ is defined by

$$W_p(\mu, \nu) = \inf \{ E[d(X, Y)^p]^{1/p} : \text{law}(X) = \mu, \text{law}(Y) = \nu \},$$

which is related to optimal transport problems and can also be used to characterize the weak convergence of probability measures in the Wasserstein space $P_p(\Omega)$ (see Definition 6.4 in Villani [10]), roughly speaking, under some moment condition, that $\mu_k$ converges weakly to $\mu$ is equivalent to $W_p(\mu_k, \mu) \to 0$. In particular, $W_1$ is commonly called Kantorovich-Rubinstein distance and the following duality formula is well-known:

$$(1.1) \quad W_1(\mu, \nu) = \sup_{||\varphi||_{\text{Lip}} \leq 1} \{ E_\mu[\varphi] - E_\nu[\varphi] \},$$

where $\Phi := \{ \varphi : ||\varphi||_{\text{Lip}} \leq 1 \}$ is the collection of all Lipschitz functions on $(\Omega, d)$ whose Lipschitz constants are at most 1 (see Villani [9, 10]).

Recently, Peng establishes a theory of time-consistent sublinear expectations in [5, 6], i.e., the $G$-expectation theory. This new theory provides with several important models and tools for studying robust hedging and utility maximization problems in the financial markets under Knightian uncertainty (see eg. Epstein and Ji [2] and Vorbrink [11]). According to Denis et al. [1], $G$-expectation can be expressed as an upper expectation introduced in Huber [3], i.e., $E_G[\cdot] := \sup_{\mu \in \mathcal{P}} E_\mu[\cdot]$, where $\mathcal{P}$ is a weakly compact set of martingale measures. Compared with the sublinear functionals studied by Lebedev in [4], $G$-expectation is non-dominated due to the mutual singularity of elements in $\mathcal{P}$.

In this paper, the classical notion of distance between two probability measures is extended, more precisely, we define a generalized Wasserstein metric for two weakly compact sets of probability measures, which could be regarded as generators of (not necessarily dominated) sublinear expectations. This metric is given by a Hausdorff type distance function, whose value is the “greatest” of all Wasserstein distances from a probability measure in one set to the “closest” probability measure in the other set. Furthermore, we adopt the notion of weak convergence for sublinear expectations from Peng [5, 6, 7].

2000 Mathematics Subject Classification. 60A10.

Key words and phrases. sublinear expectations; weak convergence; Kantorovich-Rubinstein duality formula; Wasserstein distance.
and then characterize this type of convergence by the generalized Wasserstein metric. We notice that the main techniques in this paper could be applied to considering transport inequality in the sublinear expectation context, for example, the transport-entropy inequality and the related large deviation principle. Also, the results may provide with some useful tools for the further study of robust optimal transport problems on sublinear expectation spaces.

This paper is organized as follows: in the next section, we recall some notations in the framework of sublinear expectations and define the generalized Wasserstein distance. In Section 3, the duality formula for generalized Kantorovich-Rubinstein distance are given and characterization of the weak convergence of sublinear expectations is discussed.

2. Preliminaries

We first recall preliminaries in the framework of sublinear expectations. Let \((\Omega, d)\) be a Polish space equipped with the Borel \(\sigma\)-algebra \(\mathcal{B}(\Omega)\). In the sequel, we only consider the Borel probability measures on this space and denote by \(\mathcal{P}(\Omega)\) the collection of all sets of these measures. For each \(P \in \mathcal{P}(\Omega)\), one can define a sublinear expectation \(E^P[\cdot]\) in the following form:

\[
E^P[\varphi] := \sup_{\mu \in P} E_{\mu}[\varphi], \quad \forall \varphi \in \mathcal{L}^0(\Omega),
\]

where \(\mathcal{L}^0(\Omega)\) is the space of all \(\mathcal{B}(\Omega)\)-measurable real functions.

It is easy to check that \(E^P[\cdot]\) satisfies the following properties: for \(X, Y \in \mathcal{L}^0(\Omega)\),

1. Monotonicity: If \(X \geq Y\), then \(E^P[X] \geq E^P[Y]\);
2. Constant preserving: \(E^P[c] = c, \forall c \in \mathbb{R}\);
3. Sub-additivity: \(E^P[X + Y] \leq E^P[X] + E^P[Y]\);
4. Positive homogeneity: \(E^P[\lambda X] = \lambda E^P[X], \forall \lambda \geq 0\),

where the usual rule \(0 \cdot \infty = 0\) is applied in (4).

In Peng \cite{5, 6}, the properties of sublinear expectations are systematically studied and some limit theorems are proved. In particular, Peng \cite{7} introduces the following notion of weak convergence for sublinear expectations:

**Definition 2.1.** We say that a sequence \(\{E^P_n[\cdot]\}_{n=1}^\infty\) of sublinear expectations converges weakly to some \(E^P[\cdot]\), if for each bounded and continuous function \(\varphi \in C_b(\Omega)\),

\[
\lim_{n \to \infty} E^{P_n}[\varphi] = E^P[\varphi].
\]

Now we define the generalized Wasserstein metric as follows:

**Definition 2.2.** Let \(\mathcal{P}_1, \mathcal{P}_2 \in \mathcal{P}(\Omega)\). For \(p \geq 1\), we define the \(p\)-order Wasserstein metric between \(\mathcal{P}_1\) and \(\mathcal{P}_2\) in the following form:

\[
W_p(\mathcal{P}_1, \mathcal{P}_2) = \max \left( \sup_{\mu \in \mathcal{P}_1} \inf_{\nu \in \mathcal{P}_2} W_p(\mu, \nu), \sup_{\nu \in \mathcal{P}_2} \inf_{\mu \in \mathcal{P}_1} W_p(\mu, \nu) \right),
\]

where \(W_p(\mu, \nu)\) is the classical \(p\)-order Wasserstein distance between two probability measures \(\mu\) and \(\nu\).

It is evident that the function \(W_p\) is ordered in \(p\) as its classical counterpart \(W_p\), namely, \(W_p \leq W_q\), if \(1 \leq p \leq q\).
**Remark 2.3.** It is easy to see that $W_p$ is a semi-metric and it will be a metric if we only consider weakly compact elements in $P(\Omega)$.

Similarly to Definition 6.4 in Villani [10], we can define a space $P_p(\Omega)$, on which $W_p$ defines a finite distance.

**Definition 2.4.** (Wasserstein space) For $p \geq 1$, we denote by $P_p(\Omega)$ a subset of $P(\Omega)$, which is the collection of all $P$ such that

$$E_{\mu}P \text{ is associated with such a set of probability measures, according to Corollary 3.3 in [8].}$$

**(iii)** We remark that a typical sublinear expectation, the $G$-expectation defined in Peng [5, 6], is associated with such a set of probability measures, according to [1].

**(iv)** The sublinear expectation admits an unique representation on $P_p(\Omega)$, i.e., let $P_1, P_2 \in P_p(\Omega)$, if $E^{P_1}[\varphi] = E^{P_2}[\varphi], \forall \varphi \in C_b(\Omega)$, then $P_1 = P_2$. The proof can be easily derived by Corollary 2.8.

The main technical tool in this paper is the minimax theorem in [8] as follows:

**Theorem 2.6** (Corollary 3.3 in [8]). Let $M$ and $N$ be convex spaces, one of which is compact, and $f$ a function on $M \times N$, quasi-concave-convex and upper semi-continuous-lower semi-continuous. Then $\sup f = \inf \sup f$.

In the rest of this paper, the following special case of the above theorem will be quoted, where $f$ is a linear and continuous function.

**Theorem 2.7.** Let $P \in P_1(\Omega)$. Fixing a singleton $\{\mu^*\} \in P_1(\Omega)$, we have

$$\inf_{\mu \in P} \sup_{||\varphi||_{L^p} \leq 1} \{E_{\mu^*}[\varphi] - E_{\mu}[\varphi]\} = \sup_{||\varphi||_{L^p} \leq 1} \inf_{\mu \in P} \{E_{\mu^*}[\varphi] - E_{\mu}[\varphi]\}.$$  

**Corollary 2.8.** Let $P \in P_1(\Omega)$. If there exists a singleton $\{\mu^*\} \in P_1(\Omega)$ such that

$$E_{\mu^*}[\varphi] \leq E_{\mu}[\varphi], \forall \varphi \in C_b(\Omega),$$

then $\mu^* \in P$.

**Proof.** By (b) in Definition 2.4, we can use $\Phi := \{\varphi : ||\varphi||_{L^p} \leq 1\}$ instead of $C_b(\Omega)$ in (2.2). Then we can rewrite (2.2) as

$$\sup_{\varphi \in \Phi} \inf_{\mu \in P} \{E_{\mu^*}[\varphi] - E_{\mu}[\varphi]\} = 0.$$  

By Theorem 2.7, we have

$$\inf_{\mu \in P} \sup_{\varphi \in \Phi} \{E_{\mu^*}[\varphi] - E_{\mu}[\varphi]\} = 0.$$
Since $\mathcal{P}$ is weakly compact, we can choose $\bar{\mu} \in \mathcal{P}$ such that $\sup_{\phi \in \Phi} \{ E_{\mu^*}[\phi] - E_{\bar{\mu}}[\phi] \} = 0$, which implies that $\mu^* = \bar{\mu} \in \mathcal{P}$. \hfill $\square$

3. Main results

In this section, we first study the duality formula for the generalized Kantorovich-Rubinstein distance $W_1$ on $\mathcal{P}_1(\Omega)$ defined in the previous section.

**Theorem 3.1** (Duality formula for the generalized Kantorovich-Rubinstein distance). Let $\mathcal{P}_1, \mathcal{P}_2 \in \mathcal{P}_1(\Omega)$. Then, we have

$$W_1(\mathcal{P}_1, \mathcal{P}_2) = \sup_{||\phi||_{Lip} \leq 1} \{ E_{\mathcal{P}_1}[\phi] - E_{\mathcal{P}_2}[\phi] \}.$$ 

**Proof.** By the classical duality formula for the Kantorovich-Rubinstein distance and Theorem 2.7, we have

$$\sup_{\mu \in \mathcal{P}_1} \inf_{\nu \in \mathcal{P}_2} W_1(\mu, \nu) = \sup_{\mu \in \mathcal{P}_1} \inf_{||\phi||_{Lip} \leq 1} \{ E_{\mu}[\phi] - E_{\nu}[\phi] \} \leq \sup_{\|\phi\|_{Lip} \leq 1} \{ E_{\mathcal{P}_1}[\phi] - E_{\mathcal{P}_2}[\phi] \}.$$

Similarly, we can obtain

$$\sup_{\nu \in \mathcal{P}_2} \inf_{\mu \in \mathcal{P}_1} W_1(\mu, \nu) = \sup_{||\phi||_{Lip} \leq 1} \{ E_{\mathcal{P}_2}[\phi] - E_{\mathcal{P}_1}[\phi] \}.$$

Therefore,

$$W_1(\mathcal{P}_1, \mathcal{P}_2) = \max \{ \sup_{||\phi||_{Lip} \leq 1} \{ E_{\mathcal{P}_1}[\phi] - E_{\mathcal{P}_2}[\phi] \}, \sup_{||\phi||_{Lip} \leq 1} \{ E_{\mathcal{P}_2}[\phi] - E_{\mathcal{P}_1}[\phi] \} \}.$$

It is well-known that Wasserstein distances can metrize weak convergence in the classical probability theory. Correspondingly, we can prove a similar theorem for sublinear expectations. Before presenting this result, we shall discuss some properties of the weak convergence of sublinear expectations.

Given a collection of probability measures $\mathcal{P}$, define

$$\overline{\mathcal{P}}(A) := \sup_{P \in \mathcal{P}} P(A); \underline{\mathcal{P}}(A) := \inf_{P \in \mathcal{P}} P(A).$$

Then, we have the following proposition:

**Proposition 3.2.** Let $p \geq 1$. Suppose that $\{\mathcal{P}_n\}_{n=1}^\infty$ and $\mathcal{P}$ are weakly compact sets of probability measures on $(\Omega, d)$. If $\{E_{\mathcal{P}_n}[\cdot]\}_{n=1}^\infty$ converges weakly to $E_{\mathcal{P}}[\cdot]$, then we have

(i) For each closed set $F$, 

$$\limsup_{n \to \infty} \overline{\mathcal{P}_n}(F) \leq \overline{\mathcal{P}}(F).$$
Or equivalently, for each open set $G$,

$$\liminf_{n \to \infty} P_n(G) \geq P(G).$$

(ii) The set $P^* = \bigcup_{n=1}^{\infty} P_n$ is tight.

Proof. (i) For a closed set $F$, there exists a sequence of bounded continuous functions $\{\psi_k\}_{k=1}^{\infty}$ such that $\psi_k \downarrow 1_F$, thus from Theorem 31 in Denis et al. [1], $E^P[\psi_k] \downarrow P(F)$. On the other hand,

$$\limsup_{n \to \infty} P_n(F) \leq \limsup_{n \to \infty} E^P_n[\psi_k] = E^P[\psi_k], \quad \forall k \in \mathbb{N},$$

which implies that (i) holds.

(ii) Since $\Omega$ is Polish, we can choose a subset $\{\omega_1, \cdots, \omega_n, \cdots\}$ dense in $\Omega$ and for a fixed $m \in \mathbb{N}$, we consider the cover $\{B(\omega_i, \frac{1}{m}) := \{\omega : d(\omega, \omega) < \frac{1}{m}\}\}_{i=1}^{\infty}$.

Fixing $\varepsilon > 0$, we first prove that for each $m$, there exists $k_m$ such that

$$(3.1) \quad P_n \left( \bigcup_{i=1}^{k_m} B(\omega_i, \frac{1}{m}) \right) > 1 - \varepsilon/2^m, \quad \forall n \in \mathbb{N}.$$  

Otherwise, there exists $m_0$, such that for each $k$, there exists $n_k$,

$$P_{n_k} \left( \bigcup_{i=1}^{k} B(\omega_i, \frac{1}{m_0}) \right) \leq 1 - \varepsilon/2^{m_0}.$$  

For each $n \in \mathbb{N}$, from the weak compactness of $P_n$ and Lemma 8 in [1], we know

$$P_n \left( \bigcup_{i=1}^{k} B(\omega_i, \frac{1}{m_0}) \right) = 1 - P_n \left( \left( \bigcup_{i=1}^{k} B(\omega_i, \frac{1}{m_0}) \right)^c \right) \uparrow 1, \quad \text{as } k \to \infty.$$  

Thus, we can prove by contradiction that $\lim_{k \to \infty} n_k = \infty$. Then, by (i), $\forall j \in \mathbb{N}$,

$$P \left( \bigcup_{i=1}^{j} B(\omega_i, \frac{1}{m_0}) \right) \leq \liminf_{k \to \infty} P_{n_k} \left( \bigcup_{i=1}^{j} B(\omega_i, \frac{1}{m_0}) \right) \leq \liminf_{k \to \infty} P_{n_k} \left( \bigcup_{i=1}^{k} B(\omega_i, \frac{1}{m_0}) \right) \leq 1 - \varepsilon/2^{m_0},$$

which implies that

$$P \left( \left( \bigcup_{i=1}^{j} B(\omega_i, \frac{1}{m_0}) \right)^c \right) \geq \varepsilon/2^{m_0}, \quad \forall j \in \mathbb{N}.$$  

This is a contradiction to the weak compactness of $P$.  

Choosing $k_m$ as in (3.1), we can verify that $K = \bigcap_{m=1}^{\infty} \bigcup_{i=1}^{k_m} B(\omega_i, \frac{1}{m})$ is compact. Then, for each $n$, we have
\[
\overline{\mathcal{P}}_n(K^c) = \overline{\mathcal{P}}_n \left( \bigcup_{m=1}^{\infty} \left( \bigcup_{i=1}^{k_m} B(\omega_i, \frac{1}{m}) \right)^c \right)
\leq \sum_{m=1}^{\infty} \overline{\mathcal{P}}_n \left( \bigcup_{i=1}^{k_m} B(\omega_i, \frac{1}{m})^c \right)
= \sum_{m=1}^{\infty} \left( 1 - \mathcal{P}_n \left( \bigcup_{i=1}^{k_m} B(\omega_i, \frac{1}{m}) \right) \right)
< \sum_{m=1}^{\infty} \varepsilon / 2^m = \varepsilon.
\]
Therefore, the set $\mathcal{P}^* = \bigcup_{n=1}^{\infty} \mathcal{P}_n$ is tight. \hfill \Box

**Proposition 3.3.** Let $p \geq 1$. Suppose that $\{\mathcal{P}_n\}_{n=1}^{\infty}$ and $\mathcal{P}$ are elements in $\mathcal{P}_p(\Omega)$. If $\{\mathcal{P}_n\}_{n=1}^{\infty}$ satisfies
\[
\lim_{K \to \infty} \lim_{n \to \infty} \sup \mathbb{E}^{\mathcal{P}_n} [d(\omega, \cdot)^p 1_{\{\omega \in \Omega: d(\omega, \omega_k) \geq K\} \{\cdot\}}] = 0,
\]
then the following statements are equivalent:
(i) $W_p(\mathcal{P}_n, \mathcal{P}) \to 0$.
(ii) $\hat{W}_1(\mathcal{P}_n, \mathcal{P}) \to 0$.

**Proof.** We only need to prove that “(ii) $\Rightarrow$ (i)”.” Let $\tilde{d} = \min(d, 1)$, and denote by $\hat{W}_p$ the classical Wasserstein distances defined with $\tilde{d}$ and $\hat{W}_p$ the generalized one defined in terms of $\hat{W}_p$ (see Definition 2.2). We first show that
\[
W_p(\mathcal{P}_n, \mathcal{P}) \to 0 \iff \hat{W}_p(\mathcal{P}_n, \mathcal{P}) \to 0,
\]
where the necessity is directly deduced from $W_p \geq \hat{W}_p$.

To prove the sufficiency, we first see that the distance function $d$ is dominated by the sum of three parts, that is, for $K \geq 1$, $\omega$, $\tilde{\omega}$ and $\omega_0 \in \Omega$:
\[
d(\omega, \tilde{\omega}) \leq d(\omega, \omega) \land K + 2d(\omega, \omega_0) 1_{\{d(\omega, \omega_0) \geq 2K\}} + 2d(\tilde{\omega}, \omega_0) 1_{\{d(\tilde{\omega}, \omega_0) \geq 2K\}}.
\]
Then, there exists a constant $C_p$ only depending on $p$ such that
\[
d(\omega, \tilde{\omega})^p \leq C_p (K^p \tilde{d}(\omega, \tilde{\omega})^p + d(\omega, \omega_0)^p 1_{\{d(\omega, \omega_0) \geq 2K\}} + d(\tilde{\omega}, \omega_0)^p 1_{\{d(\tilde{\omega}, \omega_0) \geq 2K\}}).
\]
By the same argument in the proof of Theorem 7.12 in Villani [9], we have for two probability measures $\mu$ and $\nu$,
\[
W_p(\mu, \nu)^p \leq C_p K^p \hat{W}_p(\mu, \nu)^p + C_p E_{\mu}[d(\omega_0, \cdot)^p 1_{\{d(\omega_0, \omega) \geq 2K\} \{\cdot\}}] + C_p E_{\nu}[d(\omega_0, \cdot)^p 1_{\{d(\omega_0, \omega) \geq 2K\} \{\cdot\}}],
\]
which implies that
\[
W_p(\mathcal{P}_n, \mathcal{P})^p \leq C_p K^p \hat{W}_p(\mathcal{P}_n, \mathcal{P})^p + C_p \mathbb{E}^{\mathcal{P}_n} [d(\omega_0, \cdot)^p 1_{\{d(\omega_0, \omega) \geq 2K\} \{\cdot\}}] + C_p \mathbb{E}^{\mathcal{P}} [d(\omega_0, \cdot)^p 1_{\{d(\omega_0, \omega) \geq 2K\} \{\cdot\}}].
\]
First letting \( n \to \infty \) and then letting \( K \to \infty \), we have \( \mathcal{W}_p(\mathcal{P}_n, \mathcal{P}) \to 0 \).

Since the distance function \( \tilde{d} \) is bounded by 1, all the distances \( \tilde{W}_p \) are equivalent (see §7.1.2 in [9]), so does \( \tilde{W}_p \). Therefore, we have

\[
\mathcal{W}_p(\mathcal{P}_n, \mathcal{P}) \to 0 \iff \tilde{\mathcal{W}}_p(\mathcal{P}_n, \mathcal{P}) \to 0 \iff \tilde{\mathcal{W}}_1(\mathcal{P}_n, \mathcal{P}) \to 0 \iff \mathcal{W}_1(\mathcal{P}_n, \mathcal{P}) \to 0.
\]

\[\square\]

**Theorem 3.4** (Wasserstein distance metrize weak convergence). Let \( p \geq 1 \). Suppose \( \{\mathcal{P}_n\}_{n=1}^{\infty} \) and \( \mathcal{P} \) are elements in \( \mathcal{P}_p(\Omega) \). Then the following statements are equivalent:

(i) \( \mathcal{W}_p(\mathcal{P}_n, \mathcal{P}) \to 0 \).

(ii) For each \( \varphi \in C(\Omega) \) with the growth condition: for some \( \omega_0 \in \Omega, |\varphi(\omega)| \leq C(1 + d(\omega_0, \omega)^p), \forall \omega \in \Omega \), we have

\[
\lim_{n \to \infty} \mathbb{E}^{\mathcal{P}_n}[\varphi] = \mathbb{E}^{\mathcal{P}}[\varphi].
\]

(iii) \( \{\mathbb{E}^{\mathcal{P}_n}[\cdot]\}_{n=1}^{\infty} \) converges weakly to \( \mathbb{E}^{\mathcal{P}}[\cdot] \) and (3.2) holds.

**Proof.** First, we prove the equivalence of (ii) and (iii).

“(ii) \(\implies\) (iii)” : It is easy to see that \( \{\mathbb{E}^{\mathcal{P}_n}[\cdot]\}_{n=1}^{\infty} \) converges weakly to \( \mathbb{E}^{\mathcal{P}}[\cdot] \), thus we only need to prove (3.2) holds. Indeed, the function \( f_K(\omega) := d(\omega_0, \omega)^p \mathbf{1}_{\{d(\omega_0, \omega) \leq K\}} \) is upper semi-continuous and thus, can be approximated by \( \varphi_m \downarrow f_K \) with \( |\varphi_m(\omega)| \leq C(1 + d(\omega_0, \omega)^p) \). Then we have

\[
\limsup_{n \to \infty} \mathbb{E}^{\mathcal{P}_n}[f_K] \leq \limsup_{n \to \infty} \mathbb{E}^{\mathcal{P}_n}[\varphi_m] = \mathbb{E}^{\mathcal{P}}[\varphi_m], \forall m \in \mathbb{N}.
\]

Letting \( m \to \infty \),

\[
\limsup_{n \to \infty} \mathbb{E}^{\mathcal{P}_n}[f_K] \leq \mathbb{E}^{\mathcal{P}}[f_K].
\]

Since \( \mathcal{P} \in \mathcal{P}_p(\Omega) \), from Definition 2.4 (b), we conclude (3.2).

“(iii) \(\implies\) (ii)” : Given \( \varphi \in C(\Omega) \) with the growth condition \( |\varphi(\omega)| \leq C(1 + d(\omega_0, \omega)^p) \), let

\[
\varphi_K = (-C(1 + K^p) \vee \varphi) \wedge C(1 + K^p)
\]

and \( \psi_K = \varphi - \varphi_K \). Then,

\[
|\mathbb{E}^{\mathcal{P}_n}[\varphi] - \mathbb{E}^{\mathcal{P}}[\varphi]| \leq |\mathbb{E}^{\mathcal{P}_n}[\varphi_K] - \mathbb{E}^{\mathcal{P}}[\varphi_K]| + |\mathbb{E}^{\mathcal{P}_n}[\psi_K]| + |\mathbb{E}^{\mathcal{P}}[\psi_K]|
\]

\[
\leq |\mathbb{E}^{\mathcal{P}_n}[\varphi_K] - \mathbb{E}^{\mathcal{P}}[\varphi_K]| + C\mathbb{E}^{\mathcal{P}_n}[d(\omega_0, \cdot)^p \mathbf{1}_{\{d(\omega_0, \omega) \geq K\}}(\cdot)]
\]

\[
+ C\mathbb{E}^{\mathcal{P}}[d(\omega_0, \cdot)^p \mathbf{1}_{\{d(\omega_0, \omega) \geq K\}}(\cdot)]
\]

Therefore,

\[
\limsup_{n \to \infty} |\mathbb{E}^{\mathcal{P}_n}[\varphi] - \mathbb{E}^{\mathcal{P}}[\varphi]| \leq C\limsup_{n \to \infty} \mathbb{E}^{\mathcal{P}_n}[d(\omega_0, \cdot)^p \mathbf{1}_{\{d(\omega_0, \omega) \geq K\}}(\cdot)]
\]

\[
+ C\mathbb{E}^{\mathcal{P}}[d(\omega_0, \cdot)^p \mathbf{1}_{\{d(\omega_0, \omega) \geq K\}}(\cdot)].
\]

Letting \( K \to \infty \), the first term of the right-hand side goes to 0 by (3.2), the second one goes to 0 since \( \mathcal{P} \in \mathcal{P}_p(\Omega) \). Thus (ii) holds.

Now, we prove “(i) \(\implies\) (iii)” . If \( \mathcal{W}_p(\mathcal{P}_n, \mathcal{P}) \to 0 \), then \( \mathcal{W}_1(\mathcal{P}_n, \mathcal{P}) \to 0 \). By Theorem 3.1 we have

\[
\sup_{\|\varphi\|_{\text{Lip}} \leq 1} \{ |\mathbb{E}^{\mathcal{P}_n}[\varphi] - \mathbb{E}^{\mathcal{P}}[\varphi]| \} \to 0,
\]
which implies that \( \lim_{n \to \infty} \mathbb{E}^{P_n}[\varphi] = \mathbb{E}^{P}[\varphi] \) holds for all Lipschitz functions. For each bounded continuous function \( \varphi \), there exist sequences of bounded Lipschitz functions \( \{\varphi_k\}_{k=1}^\infty \) and \( \{\overline{\varphi}_k\}_{k=1}^\infty \) such that
\[
\varphi_k \uparrow \varphi \quad \text{and} \quad \overline{\varphi}_k \downarrow \varphi.
\]
Then, we can deduce
\[
\lim_{n \to \infty} \sup_{k \to \infty} \mathbb{E}^{P_n}[\varphi] \leq \lim_{k \to \infty} \inf_{n \to \infty} \sup_{k \to \infty} \mathbb{E}^{P_n}[\overline{\varphi}_k] = \lim_{k \to \infty} \inf_{n \to \infty} \mathbb{E}^{P}[\overline{\varphi}_k] = \mathbb{E}^{P}[\varphi].
\]
Similarly, we have
\[
\lim_{n \to \infty} \inf_{k \to \infty} \sup_{n \to \infty} \mathbb{E}^{P_n}[\varphi] \geq \lim_{k \to \infty} \sup_{n \to \infty} \mathbb{E}^{P_n}[\varphi] = \lim_{k \to \infty} \sup_{n \to \infty} \mathbb{E}^{P}[\varphi] = \mathbb{E}^{P}[\varphi].
\]
In conclusion, \( \mathbb{E}^{P_n}[] \) converges weakly to \( \mathbb{E}^{P}[] \).

It only remains to be seen whether (3.2) holds. Indeed, one can first verify that for each \( \overline{\varphi} \in \Omega \),
\[
d(\omega, \omega)1_{\{d(\omega, \omega) \geq K\}} \leq d(\omega, \overline{\varphi})1_{\{d(\omega, \overline{\varphi}) \geq \frac{K}{2}\}} + 2d(\omega, \overline{\varphi}),
\]
and immediately obtain
\[
d(\omega, \omega)P1_{\{d(\omega, \omega) \geq K\}} \leq 2^{p-1}d(\omega, \overline{\varphi})P1_{\{d(\omega, \overline{\varphi}) \geq \frac{K}{2}\}} + 2^{p-1}d(\omega, \overline{\varphi})^p.
\]
Then we can deduce that for any \( \mu_n \in \mathcal{P}_n \) and \( \mu \in \mathcal{P} \),
\[
E_{\mu_n}[d(\omega, \cdot)P1_{\{d(\omega, \omega) \geq K\}}(\cdot)] \leq 2^{p-1}E_{\mu}[d(\omega, \cdot)P1_{\{d(\omega, \omega) \geq \frac{K}{2}\}}(\cdot)] + 2^{p-1}W_p(\mu_n, \mu)^p,
\]
which implies that
\[
\mathbb{E}^{P_n}[d(\omega, \cdot)P1_{\{d(\omega, \omega) \geq K\}}(\cdot)] \leq 2^{p-1}\mathbb{E}^{P}[d(\omega, \cdot)P1_{\{d(\omega, \omega) \geq \frac{K}{2}\}}(\cdot)] + 2^{p-1}W_p(\mathcal{P}_n, \mathcal{P})^p,
\]
where the first term of the right-hand side goes to 0 since \( \mathcal{P} \in \mathcal{P}_p(\Omega) \), and the second one vanishes as a result of (i). Consequently, (3.2) holds.

Finally, we show (iii) \( \Rightarrow \) (i). Thanks to Proposition 3.3, we only need to prove the case when \( p = 1 \). By the definition of \( \mathcal{W}_1 \), it suffices to prove that
\[
\sup_{\mu_n \in \mathcal{P}_n} \inf_{\mu \in \mathcal{P}} W_1(\mu_n, \mu) \to 0 \quad \text{and} \quad \sup_{\mu \in \mathcal{P}} \inf_{\mu_n \in \mathcal{P}_n} W_1(\mu_n, \mu) \to 0.
\]
(a) Suppose \( \sup_{\mu_n \in \mathcal{P}_n} \inf_{\mu \in \mathcal{P}} W_1(\mu_n, \mu) \to 0 \), then there exists an \( \varepsilon > 0 \) and a subsequence \( \{\mathcal{P}_{n_k}\}_{k=1}^\infty \subset \{\mathcal{P}_n\}_{n=1}^\infty \), such that for each \( n_k \), there exists a \( \mu_{n_k} \in \mathcal{P}_{n_k} \), with
\[
\inf_{\mu \in \mathcal{P}} W_1(\mu_{n_k}, \mu) \geq \varepsilon,
\]
which implies that for all \( \mu \in \mathcal{P} \), \( W_1(\mu_{n_k}, \mu) \geq \varepsilon \). By Proposition 3.2, the set \( \bigcup_{n=1}^{\infty} \mathcal{P}_n \) is tight, so it is relatively weakly compact by Prokhorov’s theorem. Then, there exists a subsequence of \( \{\mu_{n_k}\}_{k=1}^\infty \), still denoted by \( \{\mu_n\}_{n=1}^\infty \), which converges weakly to a probability measure \( \mu^* \). In this case, we can apply the classical result for Wasserstein distance to obtain \( W_1(\mu_{n_k}, \mu^*) \to 0 \). On the other hand, for each \( \varphi \in C_b(\Omega) \),
\[
E_{\mu^*}[\varphi] = \lim_{k \to \infty} E_{\mu_{n_k}}[\varphi] \leq \lim_{k \to \infty} \mathbb{E}^{P_{n_k}}[\varphi] = \mathbb{E}^{P}[\varphi],
\]
which implies from Corollary 2.8 that \( \mu^* \in \mathcal{P} \). Therefore, we obtain immediately that
\[
\inf_{\mu \in \mathcal{P}} W_1(\mu_{n_k}, \mu) \leq W_1(\mu_{n_k}, \mu^*) \to 0, \quad \text{as} \quad k \to \infty,
\]
which is in contradiction to (3.3).
(b) Suppose sup\(\mu \in \mathcal{P}\) inf\(\mu_n \in \mathcal{P}_n\) \(W_1(\mu_n, \mu) \rightarrow 0\), then there exist an \(\varepsilon > 0\), a sequence \(\{\mu_{nk}^*\}_{k=1}^{\infty} \subset \mathcal{P}\) and a subsequence \(\{\mathcal{P}_{nk}\}_{k=1}^{\infty} \subset \{\mathcal{P}_n\}_{n=1}^{\infty}\), such that
\[
\inf_{\mu_{nk} \in \mathcal{P}_{nk}} W_1(\mu_{nk}, \mu) \geq 4\varepsilon.
\]

Similarly to (a), one can find a subsequence of \(\{\mu_{nk}^*\}_{k=1}^{\infty}\), still denoted by \(\{\mu_{nk}^*\}_{k=1}^{\infty}\), which converges weakly to a probability measure \(\mu^* \in \mathcal{P}\). For sufficient large \(k\),
\[
\inf_{\mu_{nk} \in \mathcal{P}_{nk}} W_1(\mu_{nk}, \mu^*) \geq 3\varepsilon.
\]

From the classical duality formula for the Kantorovich-Rubinstein distance \((1.1)\), we have
\[
\inf_{\mu_{nk} \in \mathcal{P}_{nk}} W_1(\mu_{nk}, \mu^*) = \inf_{\mu_{nk} \in \mathcal{P}_{nk}} \sup_{|\varphi|_{\text{Lip}} \leq 1} \{E_{\mu^*}[\varphi] - E_{\mu_{nk}}[\varphi]\}.
\]

Since \(\mathcal{P}_{nk}\) and \(\Phi := \{\varphi : |\varphi|_{\text{Lip}} \leq 1\}\) are convex as well as \(\mathcal{P}_{nk}\) is weakly compact, the minimax theorem implies
\[
\inf_{\mu_{nk} \in \mathcal{P}_{nk}} \sup_{|\varphi|_{\text{Lip}} \leq 1} \{E_{\mu^*}[\varphi] - E_{\mu_{nk}}[\varphi]\} = \sup_{|\varphi|_{\text{Lip}} \leq 1} \inf_{\mu_{nk} \in \mathcal{P}_{nk}} \{E_{\mu^*}[\varphi] - E_{\mu_{nk}}[\varphi]\}
\]
\[
= \sup_{|\varphi|_{\text{Lip}} \leq 1} \{E_{\mu^*}[\varphi] - \mathbb{E}^{P_{nk}}[\varphi]\} \geq 3\varepsilon.
\]

Therefore, thanks to \((3.2)\), one can find a \(\varphi \in C_b(\Omega)\), such that for sufficient large \(k\),
\[
E_{\mu^*}[\varphi] \geq \mathbb{E}^{P_{nk}}[\varphi] + 2\varepsilon \geq \mathbb{E}^{P}[\varphi] + \varepsilon,
\]
which is in contradiction to \(\mu^* \in \mathcal{P}\).

\[\square\]

**Corollary 3.5.** Let \(p \geq 1\). Suppose that \(\mathcal{P}_{nk}\) and \(\mathcal{P}\) be sets of probability measures on \((\Omega, d)\) satisfying all conditions in Definition \(2.4\) expect the assumption of convexity. If \(\mathcal{W}_p(\mathcal{P}_n, \mathcal{P}) \rightarrow 0\), then (ii) and (iii) in Theorem \(3.4\) still hold.

**Proof.** Let \(\mathcal{P}_{nk}^*\) and \(\mathcal{P}^*\) be the convex hull of \(\mathcal{P}_n\) and \(\mathcal{P}\) respectively. We first observe that \(\mathbb{E}^{P^*}[\varphi] = \mathbb{E}^{P}[\varphi]\) for all \(\varphi\), which implies that \(\mathcal{P}_{nk}^*, \mathcal{P}^* \subset \mathcal{P}_p(\Omega)\). As in the proof of above theorem, we also have
\[
\mathbb{E}^{P^*} [d(\omega, \cdot)^p 1_{\{\omega: d(\omega, \omega) \geq K\}} (\cdot)] = \mathbb{E}^{P^*} [d(\omega, \cdot)^p 1_{\{\omega: d(\omega, \omega) \geq K\}} (\cdot)]
\]
\[
\leq 2^{p-1} \mathbb{E}^{P} [d(\omega, \cdot)^p 1_{\{\omega: d(\omega, \omega) \geq K\}} (\cdot)] + 2^{p-1} \mathbb{W}_p(\mathcal{P}_n, \mathcal{P})^p,
\]
thus \(\{\mathcal{P}_{nk}^*\}_{n=1}^{\infty}\) also satisfies \((3.2)\). So it suffice to prove that \(\mathcal{W}_1(\mathcal{P}_{nk}^*, \mathcal{P}^*) \rightarrow 0\), which yields \(\mathcal{W}_p(\mathcal{P}_{nk}^*, \mathcal{P}^*) \rightarrow 0\) by Proposition \(3.3\).

Without the assumption of convexity on the sets of probability measures, we still have for each \(n\),
\[
\sup_{\mu \in \mathcal{P}_{nk}} \inf_{\nu \in \mathcal{P}} W_1(\mu, \nu) = \sup_{\mu \in \mathcal{P}_{nk}} \inf_{\nu \in \mathcal{P}} \sup_{|\varphi|_{\text{Lip}} \leq 1} \{E_{\mu}[\varphi] - E_{\nu}[\varphi]\}
\]
\[
\geq \sup_{\mu \in \mathcal{P}_{nk}} \sup_{|\varphi|_{\text{Lip}} \leq 1} \inf_{\nu \in \mathcal{P}} \{E_{\mu}[\varphi] - E_{\nu}[\varphi]\}
\]
\[
= \sup_{|\varphi|_{\text{Lip}} \leq 1} \inf_{\mu \in \mathcal{P}_{nk}} \{E_{\mu}[\varphi] - E_{\nu}[\varphi]\}
\]
\[
= \sup_{|\varphi|_{\text{Lip}} \leq 1} \{\mathbb{E}^{P_{nk}}[\varphi] - \mathbb{E}^{P}[\varphi]\}.
\]
Similarly, 
\[ \sup_{\nu \in \mathcal{P}} \inf_{\mu \in \mathcal{P}_n} W_1(\mu, \nu) \geq \sup_{||\varphi||_{\text{Lip}} \leq 1} \{ E^\mu[\varphi] - E^{P_n}[\varphi] \}. \]

Therefore, 
\[ W_1(\mathcal{P}_n, \mathcal{P}) \geq \sup_{||\varphi||_{\text{Lip}} \leq 1} \{ ||E^{P_n}[\varphi] - E^\mu[\varphi]|| \} \]
\[ = \sup_{||\varphi||_{\text{Lip}} \leq 1} \{ ||E^{P_n^*}[\varphi] - E^{P^*}[\varphi]|| \} = W_1(\mathcal{P}_n^*, \mathcal{P}^*). \]

From which we can imply \( W_1(\mathcal{P}_n^*, \mathcal{P}^*) \to 0 \). By Theorem 3.4, (ii) and (iii) hold for \( E^{P_n}[\cdot] \) and \( E^{P^*}[\cdot] \), so do \( E^{P_n^*}[\cdot] \) and \( E^P[\cdot] \). \( \square \)

Remark 3.6. In general, without the assumption of convexity, the statement (ii) and (iii) in Theorem 3.4 may not imply \( W_p(\mathcal{P}_n, \mathcal{P}) \to 0 \). Here is a simple counterexample: consider a set \( \mathcal{P} \) satisfying Definition 2.4 except for the convexity and such that \( W_1(\mathcal{P}, \mathcal{P}^*) > 0 \), where \( \mathcal{P}^* \) is the convex hull of \( \mathcal{P} \). Let \( \mathcal{P}_n = \mathcal{P}^* \), for each \( n \in \mathbb{N} \). Obviously, (ii) and (iii) in Theorem 3.4 hold since \( E^{P_n}[\varphi] = E^{P^*}[\varphi] = E^P[\varphi] \), but \( W_1(\mathcal{P}_n, \mathcal{P}) \nrightarrow 0 \).

Acknowledgements The authors gratefully acknowledge the helpful suggestions and comments from both the associate editor and the anonymous reviewer. X. Li is supported by the China Postdoctoral Science Foundation (No.2014M561907) and the Fundamental Research Funds of Shandong University (No.2014GN007); Y. Lin is supported by the European Research Council (ERC) under grant FA506041 and by the Austrian Science Fund (FWF) under grant P25815.

References

[1] Denis, L., Hu, M., Peng, S. (2011) Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion paths, Potential Anal. 34(2): 139-161.
[2] Epstein, L. and Ji, S. (2013) Ambiguous volatility and asset pricing in continuous time, The Review of Financial Studies, 26(7): 1740-1786.
[3] Huber, P. Robust Statistics, John Wiley & Sons, 1981.
[4] Lebedev, A. (1991) On monotone dominated sublinear functionals on the space of measurable functions, Sibirskii Matematicheskii Zhurnal 33(6): 94-105.
[5] Peng, S. (2008) A new central limit theorem under sublinear expectations, arXiv:0803.2656v1.
[6] Peng, S. (2010) Nonlinear expectations and stochastic calculus under uncertainty, arXiv:1002.4546.
[7] Peng, S. (2010) Tightness, weak compactness of nonlinear expectations and application to CLT, arXiv:1006.2541.
[8] Sion, M. (1958) On general minimax theorems. Pacific Journal of Mathematics, 8(1): 171-176.
[9] Villani, C. Topics in optimal transpotation, American Mathematical Society, 2003
[10] Villani, C. Optimal transport: old and new, Springer, 2008
[11] Vorbrink, J. (2014) Financial markets with volatility uncertainty, Journal of Mathematical Economics, 53: 64-78.

Institute for Advanced Research and School of Mathematics
Shandong University
250100, Jinan, China
E-mail address: lixinpeng@sdu.edu.cn

Fakultät für Mathematik
Universität Wien
1090 Wien, Austria
E-mail address: yiqing.lin@univie.ac.at