A Skorohod Representation Theorem
for Uniform Distance

Patrizia Berti
(Università di Modena e Reggio Emilia)

Luca Pratelli
(Accademia Navale di Livorno)

Pietro Rigo
(Università di Pavia)

# 109 (01-10)
A SKOROHOD REPRESENTATION THEOREM 
FOR UNIFORM DISTANCE 

PATRIZIA BERTI, LUCA PRATELLI, AND PIETRO RIGO

Abstract. Let $\mu_n$ be a probability measure on the Borel $\sigma$-field on $D[0,1]$ with respect to Skorohod distance, $n \geq 0$. Necessary and sufficient conditions for the following statement are provided. On some probability space, there are $D[0,1]$-valued random variables $X_n$ such that $X_n \sim \mu_n$ for all $n \geq 0$ and $\|X_n - X_0\| \to 0$ in probability, where $\|\|$ is the sup-norm. Such conditions do not require $\mu_0$ separable under $\|\|$. Applications to exchangeable empirical processes and to pure jump processes are given as well.

1. Introduction

Let $D$ be the set of real cadlag functions on $[0,1]$ and

$$\|x\| = \sup_t |x(t)|, \quad u(x, y) = \|x - y\|, \quad x, y \in D.$$ 

Also, let $d$ be Skorohod distance and $B_d, B_u$ the Borel $\sigma$-fields on $D$ with respect to (w.r.t.) $d$ and $u$, respectively.

In real problems, one usually starts with a sequence $(\mu_n : n \geq 0)$ of probabilities on $B_d$. If $\mu_n \to \mu_0$ weakly (under $d$), Skorohod representation theorem yields $d(X_n, X_0) \overset{a.s.}{\to} 0$ for some $D$-valued random variables $X_n$ such that $X_n \sim \mu_n$ for all $n \geq 0$. However, $X_n$ can fail to approximate $X_0$ uniformly. A trivial example is $\mu_n = \delta_{x_n}$, where $(x_n) \subset D$ is any sequence such that $x_n \to x_0$ according to $d$ but not according to $u$.

Lack of uniform convergence is sometimes a trouble. Thus, given a sequence $(\mu_n : n \geq 0)$ of laws on $B_d$, it is useful to have conditions for:

On some probability space $(\Omega, A, P)$, there are random variables

$$X_n : \Omega \to D \text{ such that } X_n \sim \mu_n \text{ for all } n \geq 0 \text{ and } \|X_n - X_0\| \overset{P}{\to} 0.$$ 

Convergence in probability cannot be strengthened into a.s. convergence in condition (1). In fact, it may be that (1) holds, and yet there are not $D$-valued random variables $Y_n$ such that $Y_n \sim \mu_n$ for all $n$ and $\|Y_n - Y_0\| \overset{a.s.}{\to} 0$; see Example 7.

This paper is concerned with (1). The main result is Theorem 4, which states that (1) holds if and only if

$$\limsup_{n \to \infty} \sup_{f \in L} |\mu_n(f) - \mu_0(f)| = 0,$$

2000 Mathematics Subject Classification. 60B10, 60A05, 60A10.

Key words and phrases. Cadlag function – Exchangeable empirical process – Separable probability measure – Skorohod representation theorem – Uniform distance – Weak convergence of probability measures.
where \( L \) is the set of functions \( f : D \to \mathbb{R} \) satisfying
\[
\sigma(f) \subset B_d, \quad -1 \leq f \leq 1, \quad |f(x) - f(y)| \leq \|x - y\| \quad \text{for all } x, y \in D.
\]

Theorem 4 can be commented as follows. Say that a probability \( \mu \), defined on \( B_d \) or \( B_u \), is \( u \)-separable in case \( \mu(A) = 1 \) for some \( u \)-separable \( A \in B_d \). Suppose \( \mu_0 \) is \( u \)-separable and define \( \mu_n(H) = \mu_0(A \cap H) \) for \( H \in B_u \), where \( A \in B_d \) is \( u \)-separable and \( \mu_0(A) = 1 \). Since \( \mu_n \) is defined only on \( B_d \) for \( n \geq 1 \), we adopt Hoffmann-Jorgensen’s definition of convergence in distribution for non measurable random elements; see e.g. [7] and [9]. Let \( I_0 \) be the identity map on \( (D, B_u, \mu_0) \) and \( I_n \) the identity map on \( (D, B_d, \mu_n), n \geq 1 \). Further, let \( D \) be regarded as a metric space under \( u \). Then, since \( \mu_0 \) is \( u \)-separable, one obtains:

(i) Condition (1) holds (with \( \|X_n - X_0\| \overset{a.s.}{\to} 0 \)) provided \( I_n \to I_0 \) in distribution;

(ii) \( I_n \to I_0 \) in distribution if and only if \( \lim_n \sup_{f \in L} |\mu_n(f) - \mu_0(f)| = 0 \).

Both (i) and (ii) are known facts; see Theorems 1.7.2, 1.10.3 and 1.12.1 of [9].

The spirit of Theorem 4, thus, is that one can dispense with \( u \)-separability of \( \mu_0 \) to get (1). This can look surprising, as separability of the limit law is crucial in Skorohod representation theorem; see [5]. However, \( X_n \sim \mu_n \) is asked only on \( B_d \) and not on \( B_u \). Indeed, \( X_n \) can even fail to be measurable w.r.t. \( B_u \).

Non \( u \)-separable laws on \( B_d \) are quite usual. A cadlag process \( Z \), with jumps at random time points, has typically a non \( u \)-separable distribution on \( B_d \). One example is \( Z(t) = B_{M(t)} \), where \( B \) is a standard Brownian bridge, \( M \) an independent random distribution function and the jump-points of \( M \) have a non discrete distribution. Such a \( Z \) is the limit in distribution, under \( d \), of certain exchangeable empirical processes; see [1] and [3].

In applications, unless \( \mu_0 \) is \( u \)-separable, checking condition (2) is usually difficult. In this sense, Theorem 4 can be viewed as a “negative” result, as it states that condition (1) is quite hard to reach. This is partly true. However, there are also meaningful situations where (2) can be proved with a reasonable effort. Two examples are exchangeable empirical processes, which motivated Theorem 4, and a certain class of jump processes. Both are discussed in Section 4.

Our proof of Theorem 4 is admittedly long and it is confined in a final appendix. Some preliminary results, of possible independent interest, are needed. We mention Proposition 2 and Lemma 13 in particular.

A last remark is that Theorem 4 is still valid if \( D \) is replaced by \( D([0,1], \mathcal{X}) \), the space of cadlag functions from \([0,1] \) into a separable Banach space \( \mathcal{X} \).

### 2. A PRELIMINARY RESULT

Let \( (\Omega, \mathcal{A}, P) \) be a probability space. The outer and inner measures are
\[
P^*(H) = \inf \{ P(A) : H \subset A \in \mathcal{A} \}, \quad P_*(H) = 1 - P^*(H^c), \quad H \subset \Omega.
\]

Given a metric space \((S, \rho)\) and maps \( X_n : \Omega \to S, n \geq 0 \), say that \( X_n \) converges to \( X_0 \) in \((outer)\) probability, written \( X_n \overset{P}{\rightarrow} X_0 \), in case
\[
\lim_n P^*(\rho(X_n, X_0) > \epsilon) = 0 \quad \text{for all } \epsilon > 0.
\]

In the sequel, \( d_{TV} \) denotes total variation distance between two probabilities defined on the same \( \sigma \)-field.
Proposition 1. Let $(F, \mathcal{F})$ be a measurable space and $\mu_n$ a probability on $(F, \mathcal{F})$, $n \geq 0$. Then, on some probability space $(\Omega, \mathcal{A}, P)$, there are measurable maps $X_n : (\Omega, \mathcal{A}) \rightarrow (F, \mathcal{F})$ such that

$$P_n(X_n \neq X_0) = P_n^*(X_n \neq X_0) = \nu\text{-TV}(\mu_n, \mu_0) \text{ and } X_n \sim \mu_n \text{ for all } n \geq 0.$$ 

Proposition 1 is well known, even if in a slightly different form; see Theorem 2.1 of [8]. A proof of the present version is in Section 3 of [5].

Next proposition is fundamental for proving our main result. Among other things, it can be viewed as an improvement of Proposition 1.

Proposition 2. Let $\lambda_n$ be a probability on $(F \times G, \mathcal{F} \otimes \mathcal{G})$, $n \geq 0$, where $(F, \mathcal{F})$ is a measurable space and $(G, \mathcal{G})$ a Polish space equipped with its Borel $\sigma$-field. The following conditions are equivalent:

(a) There are a probability space $(\Omega, \mathcal{A}, P)$ and measurable maps $(Y_n, Z_n) : (\Omega, \mathcal{A}) \rightarrow (F \times G, \mathcal{F} \otimes \mathcal{G})$ such that

$$(Y_n, Z_n) \sim \lambda_n \text{ for all } n \geq 0, \quad P^*(Y_n \neq Y_0) \rightarrow 0, \quad Z_n \xrightarrow{P} Z_0;$$

(b) For each bounded Lipschitz function $f : G \rightarrow \mathbb{R}$,

$$\limsup_n \sup_{A \in \mathcal{F}} \left| \int I_A(y) f(z) \lambda_n(dy, dz) - \int I_A(y) f(z) \lambda_0(dy, dz) \right| = 0.$$

To prove Proposition 2, we first recall a result of Blackwell and Dubins [6].

Theorem 3. Let $G$ be a Polish space, $\mathcal{M}$ the collection of Borel probabilities on $G$, and $m$ the Lebesgue measure on $(0,1)$. There is a Borel measurable map

$$\Phi : \mathcal{M} \times (0,1) \rightarrow G$$

such that, for every $\nu \in \mathcal{M}$,

(i) $\Phi(\nu, \cdot) \sim \nu$ under $m$;

(ii) There is a Borel set $A_\nu \subset (0,1)$ such that $m(A_\nu) = 1$ and

$$\Phi(\nu_n, t) \xrightarrow{} \Phi(\nu, t) \text{ whenever } t \in A_\nu, \nu_n \in \mathcal{M} \text{ and } \nu_n \xrightarrow{\nu} \nu \text{ weakly.}$$

We also need to recall disintegrations. Let $\lambda$ be a probability on $(F \times G, \mathcal{F} \otimes \mathcal{G})$, where $(F, \mathcal{F})$ and $(G, \mathcal{G})$ are arbitrary measurable spaces. In this paper, $\lambda$ is said to be disintegrable if there is a collection $\alpha = \{\alpha(y) : y \in F\}$ such that:

- $\alpha(y)$ is a probability on $G$ for $y \in F$;

- $y \mapsto \alpha(y)(C)$ is $\mathcal{F}$-measurable for $C \in \mathcal{G}$;

- $\lambda(A \times C) = \int_A \alpha(y)(C) \mu(dy)$ for $A \in \mathcal{F}$ and $C \in \mathcal{G}$, where $\mu(\cdot) = \lambda(\cdot \times G)$.

Such an $\alpha$ is called a disintegration for $\lambda$. For $\lambda$ to admit a disintegration, it suffices that $G$ is a Borel subset of a Polish space and $\mathcal{G}$ the Borel $\sigma$-field on $G$.

Proof of Proposition 2. ”(a) $\Rightarrow$ (b)”. Under (a), for each $A \in \mathcal{F}$ and bounded Lipschitz $f : G \rightarrow \mathbb{R}$, one obtains

$$\left| \int I_A(y) f(z) \lambda_n(dy, dz) - \int I_A(y) f(z) \lambda_0(dy, dz) \right| = \left| E_P \left( I_A(Y_n) f(Z_n) - I_A(Y_0) f(Z_0) \right) \right|$$

$$\leq E_P \left| f(Z_n) (I_A(Y_n) - I_A(Y_0)) \right| + E_P \left| I_A(Y_0) (f(Z_n) - f(Z_0)) \right|$$

$$\leq \sup |f| P^*(Y_n \neq Y_0) + E_P \left| f(Z_n) - f(Z_0) \right| \rightarrow 0.$$
"(b) ⇒ (a)". Let \( \mu_n(A) = \lambda_n(A \times G) \), \( A \in \mathcal{F} \). By (b), \( d_{TV}(\mu_n, \mu_0) \to 0 \). Hence, by Proposition 1, on a probability space \((\Theta, \mathcal{E}, Q)\) there are measurable maps \( h_n : (\Theta, \mathcal{E}) \to (F, \mathcal{F}) \) satisfying \( h_n \sim \mu_n \) for all \( n \) and \( Q^*(h_n \neq h_0) \to 0 \). Let
\[
\Omega = \Theta \times (0,1), \quad \mathcal{A} = \mathcal{E} \otimes \mathcal{B}(0,1), \quad P = Q \times m,
\]
where \( \mathcal{B}(0,1) \) is the Borel \( \sigma \)-field on \((0,1)\) and \( m \) the Lebesgue measure.

Since \( G \) is Polish, each \( \lambda_n \) admits a disintegration \( \alpha_n = \{ \alpha_n(y) : y \in F \} \). By Theorem 3, there is a map \( \Phi : \mathcal{M} \times (0,1) \to G \) satisfying conditions (i)-(ii). Let
\[
Y_n(\theta, t) = h_n(\theta) \quad \text{and} \quad Z_n(\theta, t) = \Phi\{\alpha_n(h_n(\theta)), t\}, \quad (\theta, t) \in \Theta \times (0,1).
\]
For fixed \( \theta \), condition (i) yields \( Z_n(\theta, \cdot) = \Phi\{\alpha_n(h_n(\theta)), \cdot\} \sim \alpha_n(h_n(\theta)) \) under \( m \).

Since \( \alpha_n \) is a disintegration for \( \lambda_n \), for all \( A \in \mathcal{F} \) and \( C \in \mathcal{G} \) one has
\[
P(Y_n \in A, Z_n \in C) = \int_{\Theta} I_A(h_n(\theta)) m\{t : Z_n(\theta, t) \in C\} Q(d\theta)
\]
and
\[
= \int_{\{h_n \in A\}} \alpha_n(h_n(\theta))(C) Q(d\theta) = \int_A \alpha_n(y)(C) \mu_n(dy) = \lambda_n(A \times C).
\]
Also, \( P^*(Y_n \neq Y_0) = Q^*(h_n \neq h_0) \to 0 \) by Lemma 1.2.5 of [9].

Finally, we prove \( Z_n \xrightarrow{P} Z_0 \). Write \( \alpha_n(y)(f) = \int f(z) \alpha_n(y)(dz) \) for all \( y \in F \) and \( f \in L_G \), where \( L_G \) is the set of Lipschitz functions \( f : G \to [-1,1] \). Since \( Q^*(h_n \neq h_0) \to 0 \), there are \( A_n \in \mathcal{F} \) such that \( Q(A_n^c) \to 0 \) and \( h_n = h_0 \) on \( A_n \).

Given \( f \in L_G \),
\[
E_Q[\alpha_n(h_n)(f) - \alpha_0(h_0)(f)] - 2Q(A_n^c) \leq E_Q\{I_{A_n} \alpha_n(h_0)(f) - \alpha_0(h_0)(f)\} - 2Q(A_n^c)
\]
\[
\leq E_Q[\alpha_n(h_0)(f) - \alpha_0(h_0)(f)] = \int \alpha_n(y)(f) - \alpha_0(y)(f) \mu_0(dy).
\]
Using condition (b), it is not hard to see that \( \int \alpha_n(y)(f) - \alpha_0(y)(f) \mu_0(dy) \to 0 \).

Therefore, \( \alpha_n(h_n)(f) \xrightarrow{Q} \alpha_0(h_0)(f) \) for each \( f \in L_G \), and this is equivalent to each subsequence \( (n') \) contains a further subsequence \( (n'') \) such that \( \alpha_{n''}(h_{n''}(\theta)) \xrightarrow{a.s.} \alpha_0(h_0(\theta)) \) weakly for \( Q \)-almost all \( \theta \); see Remark 2.3 and Corollary 2.4 of [2]. Thus, by property (ii) of \( \Phi \), each subsequence \( (n') \) contains a further subsequence \( (n'') \) such that \( Z_{n''} \xrightarrow{a.s.} Z_0 \). That is, \( Z_n \xrightarrow{P} Z_0 \) and this concludes the proof.

\[\square\]

3. Existence of cadlag processes, with given distributions on the Skorohod Borel \( \sigma \)-field, converging uniformly in probability

As in Section 1, \( \mathcal{B}_d \) and \( \mathcal{B}_u \) are the Borel \( \sigma \)-fields on \( D \) w.r.t. \( d \) and \( u \). Also, \( L \) is the class of functions \( f : D \to [-1,1] \) which are measurable w.r.t. \( \mathcal{B}_d \) and Lipschitz w.r.t. \( u \) with Lipschitz constant 1. We recall that, for \( x, y \in D \), the Skorohod distance \( d(x, y) \) is the infimum of those \( \epsilon > 0 \) such that
\[
\|x - y \circ \gamma\| \leq \epsilon \quad \text{and} \quad \sup_{s \neq t} \left| \log \frac{\gamma(s) - \gamma(t)}{s - t} \right| \leq \epsilon
\]
for some strictly increasing homeomorphism \( \gamma : [0,1] \to [0,1] \). The metric space \((D, d)\) is separable and complete.
We write $\mu(f) = \int f \, d\mu$ whenever $\mu$ is a probability on a $\sigma$-field and $f$ a real bounded function, measurable w.r.t. such a $\sigma$-field.

Motivations for the next result have been given in Section 1.

**Theorem 4.** Let $\mu_n$ be a probability measure on $\mathcal{B}_d$, $n \geq 0$. Then, conditions (1) and (2) are equivalent, that is,

$$\limsup_n |\mu_n(f) - \mu_0(f)| = 0$$

if and only if there are a probability space $(\Omega, \mathcal{A}, P)$ and measurable maps $X_n : (\Omega, \mathcal{A}) \to (D, \mathcal{B}_d)$ such that $X_n \sim \mu_n$ for each $n \geq 0$ and $\|X_n - X_0\| \overset{P}{\to} 0$.

The proof of Theorem 4 is given in the Appendix. Here, we state a corollary and an open problem and we make two examples.

In applications, the $\mu_n$ are often probability distributions of random variables, all defined on some probability space $(\Omega_0, \mathcal{A}_0, P_0)$. In the spirit of [4], a (minor) question is whether condition (1) holds with the $X_n$ defined on $(\Omega_0, \mathcal{A}_0, P_0)$ as well.

**Corollary 5.** Let $(\Omega_0, \mathcal{A}_0, P_0)$ be a probability space and $Z_n : (\Omega_0, \mathcal{A}_0) \to (D, \mathcal{B}_d)$ a measurable map, $n \geq 1$. Suppose $\limsup_n |E_{P_0}(f(Z_n)) - E_{P_0}(f)| = 0$ for some probability measure $\mu_0$ on $\mathcal{B}_d$. If $P_0$ is nonatomic, there are measurable maps $X_n : (\Omega_0, \mathcal{A}_0) \to (D, \mathcal{B}_d)$, $n \geq 0$, such that

$$X_0 \sim \mu_0, \quad X_n \sim Z_n \quad \text{for each } n \geq 1, \quad \|X_n - X_0\| \overset{P_0}{\to} 0.$$

Also, $P_0$ is nonatomic if $\mu_0(x) = 0$ for all $x \in D$, or if $P_0(Z_n = x) = 0$ for some $n \geq 1$ and all $x \in D$.

**Proof.** Since $(D, d)$ is separable, $P_0$ is nonatomic if $P_0(Z_n = x) = 0$ for some $n \geq 1$ and all $x \in D$. By Corollary 5.4 of [4], $(\Omega_0, \mathcal{A}_0, P_0)$ supports a $D$-valued random variable $Z_0$ with $Z_0 \sim \mu_0$. Hence, $P_0$ is nonatomic even if $\mu_0(x) = 0$ for all $x \in D$.

Next, by Theorem 4, on a probability space $(\Omega, \mathcal{A}, P)$ there are $D$-valued random variables $Y_n$ such that $Y_0 \sim \mu_0$, $Y_n \sim Z_n$ for $n \geq 1$ and $\|Y_n - Y_0\| \overset{P}{\to} 0$. Let $(D^\infty, \mathcal{B}_d^\infty)$ be the countable product of $(D, \mathcal{B}_d)$ and

$$\nu(A) = P((Y_0, Y_1, \ldots) \in A), \quad A \in \mathcal{B}_d^\infty.$$ 

Then, $\nu$ is a Borel probability on a Polish space. Thus, if $P_0$ is nonatomic, $(\Omega_0, \mathcal{A}_0, P_0)$ supports a $D^\infty$-valued random variable $X = (X_0, X_1, \ldots)$ with $X \sim \nu$; see e.g. Theorem 3.1 of [4]. Since $(X_0, X_1, \ldots) \sim (Y_0, Y_1, \ldots)$, this concludes the proof. \hfill $\square$

Let $(S, \rho)$ be a metric space such that $(x, y) \mapsto \rho(x, y)$ is measurable w.r.t. $\mathcal{E} \otimes \mathcal{E}$, where $\mathcal{E}$ is the ball $\sigma$-field on $S$. This is actually true in case $(S, \rho) = (D, \rho)$ and it is very useful to prove Theorem 4. Thus, a question is whether $(D, \rho)$ can be replaced by $(S, \rho)$ in Theorem 4. Precisely, let $(\mu_n : n \geq 0)$ be a sequence of laws on $\mathcal{E}$ and $L_\mathcal{E}$ the class of functions $f : S \to [-1, 1]$ such that $\sigma(f) \subset \mathcal{E}$ and $|f(x) - f(y)| \leq \rho(x, y)$ for all $x, y \in S$. Then,

**Conjecture:** $\lim_n \sup f \in L_\mathcal{E} |\mu_n(f) - \mu_0(f)| = 0$ if and only if $\rho(X_n, X_0) \overset{P}{\to} 0$ in probability for some $S$-valued random variables $X_n$ such that $X_n \sim \mu_n$ for all $n$.

We finally give two examples. The first shows that condition (2) cannot be weakened into $\mu_n(f) \to \mu_0(f)$ for each fixed $f \in L$. 

---

**SKOROHOD REPRESENTATION FOR UNIFORM DISTANCE**

5
Example 6. For each $n \geq 0$, let $h_n : (0, 1) \to \mathbb{[0, \infty)}$ be a Borel function such that
\[
\int_0^1 h_n(t) \, dt = 1.
\] Suppose that $h_n \to h_0$ in $\sigma(L_1, L_\infty)$ but not in $L_1$ under Lebesgue measure $m$ on $(0, 1)$, that is,
\[
\limsup_n \int_0^1 |h_n(t) - h_0(t)| \, dt > 0,
\]
(3)
\[
\lim_n \int_0^1 h_n(t) g(t) \, dt = \int_0^1 h_0(t) g(t) \, dt \quad \text{for all bounded Borel functions } g.
\]
Take a sequence $(T_n : n \geq 0)$ of $(0, 1)$-valued random variables, on a probability space $(\Theta, \mathcal{F}, Q)$, such that each $T_n$ has density $h_n$ w.r.t. $m$. Define
\[
Z_n = I_{[T_n, 1]} \quad \text{and} \quad \mu_n(A) = Q(Z_n \in A) \text{ for } A \in \mathcal{B}_d.
\]
Then $Z_n = \phi(T_n)$, with $\phi : (0, 1) \to D$ given by $\phi(t) = I_{[t, 1]}$, $t \in (0, 1)$. Hence, for fixed $f \in L$, one obtains
\[
\mu_n(f) = E_Q \{f \circ \phi(T_n)\} = \int_0^1 h_n(t) f \circ \phi(t) \, dt \longrightarrow \int_0^1 h_0(t) f \circ \phi(t) \, dt = \mu_0(f).
\]
Suppose now that $X_n \sim \mu_n$ for all $n \geq 0$, where the $X_n$ are $D$-valued random variables on some probability space $(\Omega, \mathcal{A}, P)$. Since
\[
P\{\omega : X_n(\omega)(t) \in \{0, 1\} \text{ for all } t\} = Q\{\theta : Z_n(\theta)(t) \in \{0, 1\} \text{ for all } t\} = 1,
\]
it follows that
\[
P(\|X_n - X_0\| > \frac{1}{2}) = P(X_n \neq X_0) \geq d_{TV}(\mu_n, \mu_0) = \frac{1}{2} \int_0^1 |h_n(t) - h_0(t)| \, dt.
\]
Therefore, $X_n$ fails to converge to $X_0$ in probability.

A slight change in Example 6 shows that convergence in probability cannot be strengthened into a.s. convergence in condition (1). Precisely, it may be that (1) holds, and yet there are not $D$-valued random variables $Y_n$ satisfying $Y_n \sim \mu_n$ for all $n$ and $\|Y_n - Y_0\| \overset{a.s.}{\longrightarrow} 0$.

Example 7. In the notation of Example 6, instead of (3) assume
\[
\lim_n \int_0^1 |h_n(t) - h_0(t)| \, dt = 0 \quad \text{and} \quad m(\liminf_n h_n < h_0) > 0
\]
where $m$ is Lebesgue measure on $(0, 1)$. Since
\[
d_{TV}(\mu_n, \mu_0) = \frac{1}{2} \int_0^1 |h_n(t) - h_0(t)| \, dt \longrightarrow 0,
\]
condition (1) trivially holds by Proposition 1. Suppose now that $Y_n \sim \mu_n$ for all $n \geq 0$, where the $Y_n$ are $D$-valued random variables on a probability space $(\Omega, \mathcal{A}, P)$. As $m(\liminf_n h_n < h_0) > 0$, Theorem 3.1 of [8] yields $P(Y_n = Y_0 \text{ ultimately}) < 1$. On the other hand, since $P(Y_n(t) \in \{0, 1\} \text{ for all } t) = 1$, one obtains
\[
P(\|Y_n - Y_0\| \longrightarrow 0) = P(Y_n = Y_0 \text{ ultimately}) < 1.
\]
4. Applications

Condition (2) is not always hard to be checked, even if \( \mu_0 \) is not \( u \)-separable. We illustrate this fact by two examples. To this end, we first note that conditions (1)-(2) are preserved under certain mixtures.

**Corollary 8.** Let \( G \) be the set of distribution functions on \([0, 1]\) and \( \mathcal{G} \) the \( \sigma \)-field on \( G \) generated by the maps \( g \mapsto g(t), \ 0 \leq t \leq 1 \). Let \( \pi \) be a probability on \( \mathcal{G} \) and \( \mu_n \) and \( \lambda_n \) probabilities on \( B_d \). Then, condition (1) holds provided

\[
\sup_{f \in L} |\lambda_n(f) - \lambda_0(f)| \rightarrow 0 \quad \text{and} \quad \mu_n(A) = \int \lambda_n\{x : x \circ g \in A\} \pi(dg) \quad \text{for all } n \geq 0 \text{ and } A \in B_d.
\]

**Proof.** By Theorem 4, there are a probability space \((\Theta, \mathcal{E}, Q)\) and measurable maps \( Z_0 : (\Theta, \mathcal{E}) \rightarrow (D, B_d) \) such that \( Z_n \sim \lambda_n \) for all \( n \) and \( \|Z_n - Z_0\| \xrightarrow{Q} 0 \). Define \( \Omega = \Theta \times G, A = \mathcal{E} \otimes \mathcal{G}, P = Q \times \pi, \) and \( X_n(\theta, g) = Z_n(\theta) \circ g \) for all \((\theta, g) \in \Theta \times G\). It is routine to check that\( X_n \sim \mu_n \) for all \( n \) and \( \|X_n - X_0\| \xrightarrow{P} 0 \) \( \square \)

**Example 9.** (Exchangeable empirical processes). Let \((\xi_n : n \geq 1)\) be a sequence of \([0, 1]\)-valued random variables on the probability space \((\Omega_0, \mathcal{A}_0, P_0)\). Suppose \((\xi_n)\) exchangeable and define

\[
F(t) = E_{P_0}(I_{\{t_1 \leq t\}} | \tau)
\]

where \( \tau \) is the tail \( \sigma \)-field of \((\xi_n)\). Take \( F \) to be regular, i.e., each \( F \)-path is a distribution function. Then, the \( n \)-th empirical process can be defined as

\[
Z_n(t) = \frac{\sum_{i=1}^{n} I_{\{t_1 \leq t\}} - F(t)}{\sqrt{n}}, \quad 0 \leq t \leq 1, \ n \geq 1.
\]

Since \( Z_n : (\Omega_0, \mathcal{A}_0) \rightarrow (D, B_d) \) is measurable, one can define \( \mu_n(\cdot) = P_0(Z_n \in \cdot) \). Also, let \( \mu_0 \) be the probability distribution of

\[
Z_0(t) = B_{M(t)}
\]

where \( B \) is a standard Brownian bridge on \([0, 1]\) and \( M \) an independent copy of \( F \) (with \( B \) and \( M \) defined on some probability space). Then, \( \mu_n \rightarrow \mu_0 \) weakly (under \( d \)) but \( \mu_0 \) can fail to admit any extension to \( B_n \); see [3] and Example 11 of [1]. Thus, \( Z_n \) can fail to converge in distribution, under \( u \), according to Hoffmann-Jørgensen’s definition. However, Corollaries 5 and 8 grant that:

On \((\Omega_0, \mathcal{A}_0, P_0)\), there are measurable maps \( X_n : (\Omega_0, \mathcal{A}_0) \rightarrow (D, B_d) \) such that \( X_n \sim Z_n \) for each \( n \geq 0 \) and \( \|X_n - X_0\| \xrightarrow{P_0} 0 \).

Define in fact \( B_n(t) = n^{-1/2} \sum_{i=1}^{n} I_{\{t_1 \leq t\}} - t \), where \( u_1, u_2, \ldots \) are i.i.d. random variables (on some probability space) with uniform distribution on \([0, 1]\). Then, \( B_n \rightarrow B \) in distribution, under \( u \), according to Hoffmann-Jørgensen’s definition. Let \( \lambda_n \) and \( \lambda_0 \) be the probability distributions of \( B_n \) and \( B \), respectively. Since \( \lambda_0 \) is \( u \)-separable, \( \sup_{f \in L} |\lambda_n(f) - \lambda_0(f)| \rightarrow 0 \) (see Section 1). Thus, the first condition of Corollary 8 holds. The second condition follows from de Finetti’s representation theorem, by letting \( \pi(A) = P_0(F \in A) \) for \( A \in \mathcal{G} \). Hence, condition (1) holds.

It remains to see that the \( X_n \) can be defined on \((\Omega_0, \mathcal{A}_0, P_0)\). To this end, it can be assumed \( \mathcal{A}_0 = \sigma(\xi_1, \xi_2, \ldots) \). If \( P_0 \) is nonatomic, it suffices to apply Corollary 5. Suppose \( P_0 \) has an atom \( A \). Since \( \mathcal{A}_0 = \sigma(\xi_1, \xi_2, \ldots) \), up to \( P_0 \)-null sets, \( A \) is of the
form \( A = \{ \xi_n = t_n \text{ for all } n \geq 1 \} \) for some constants \( t_n \). Let \( \sigma = (\sigma_1, \sigma_2, \ldots) \) be a permutation of \( 1, 2, \ldots \) and \( A_\sigma = \{ \xi_n = t_{\sigma_n} \text{ for all } n \geq 1 \} \). By exchangeability,
\[
P_0(A_\sigma) = P_0(A) > 0 \quad \text{for all permutations } \sigma,
\]
and this implies \( t_n = t_1 \) for all \( n \geq 1 \). Let \( H \) be the union of all \( P_0 \)-null sets. Up to \( P_0 \)-null sets, one obtains
\[
H \subset \{ \xi_n = \xi_1 \text{ for all } n \geq 1 \} \subset \{ Z_n = 0 \text{ for all } n \geq 1 \}.
\]
If \( P_0(H) = 1 \), thus, it suffices to let \( X_n = 0 \) for all \( n \geq 0 \). If \( 0 < P_0(H) < 1 \), since \( P_0(\cdot \mid H^c) \) is nonatomic and \( (\xi_n) \) is still exchangeable under \( P_0(\cdot \mid H^c) \), it is not hard to define the \( X_n \) on \((\Omega_0, A_0, P_0)\) in such a way that \( X_n \sim Z_n \) for all \( n \geq 0 \) and \( \|X_n - X_0\| \xrightarrow{P_0} 0 \).

**Example 10. (Pure jump processes).** For each \( n \geq 0 \), let
\[
C_n = (C_{n,j} : j \geq 1) \quad \text{and} \quad Y_n = (Y_{n,j} : j \geq 1)
\]
be sequences of real random variables, defined on the probability space \((\Omega_0, A_0, P_0)\), such that
\[
0 \leq Y_{n,j} \leq 1 \quad \text{and} \quad \sum_{j=1}^{\infty} |C_{n,j}| < \infty.
\]
Define
\[
Z_n(t) = \sum_{j=1}^{\infty} C_{n,j} I_{\{Y_{n,j} \leq t\}}, \quad 0 \leq t \leq 1, \ n \geq 0.
\]
Since \( Z_n : (\Omega_0, A_0) \rightarrow (D, B_d) \) is measurable, one can define \( \mu_n(\cdot) = P_0(Z_n \in \cdot) \).

Then, condition (1) holds provided
\[
C_n \text{ is independent of } Y_n \text{ for every } n \geq 0,
\]
\[
\sum_{j=1}^{\infty} |C_{n,j} - C_{0,j}| \xrightarrow{P_0} 0 \quad \text{and} \quad d_{TV}(\nu_{n,k}, \nu_{0,k}) \longrightarrow 0 \quad \text{for all } k \geq 1,
\]
where \( \nu_{n,k} \) denotes the probability distribution of \((Y_{n,1}, \ldots, Y_{n,k})\).

For instance, \( \nu_{n,k} = \nu_{0,k} \) for all \( n \) and \( k \) in case \( Y_{n,j} = V_{n+j} \) with \( V_1, V_2, \ldots \) a stationary sequence. Also, independence between \( C_n \) and \( Y_n \) can be replaced by \( \sigma(C_{n,j}) \subset \sigma(Y_{n,1}, \ldots, Y_{n,j}) \) for all \( n \geq 0 \) and \( j \geq 1 \).

To prove (1), define \( Z_{n,k}(t) = \sum_{j=1}^{k} C_{n,j} I_{\{Y_{n,j} \leq t\}} \). For each \( f \in L, \)
\[
|\mu_n(f) - \mu_0(f)| \leq |Ef(Z_n) - Ef(Z_{0,k})| + |Ef(Z_{n,k}) - Ef(Z_{0,k})| + |Ef(Z_{0,k}) - Ef(Z_0)|
\]
\[
\leq E\{2 \wedge \|Z_n - Z_{0,k}\|\} + |Ef(Z_{n,k}) - Ef(Z_{0,k})| + E\{2 \wedge \|Z_0 - Z_{0,k}\|\}
\]
\[
\leq E\{2 \wedge \sum_{j>k} |C_{n,j}|\} + |Ef(Z_{n,k}) - Ef(Z_{0,k})| + E\{2 \wedge \sum_{j>k} |C_{0,j}|\}
\]
where \( E(\cdot) = EP_0(\cdot) \). Given \( \epsilon > 0 \), take \( k \geq 1 \) such that \( E\{2 \wedge \sum_{j>k} |C_{0,j}|\} < \epsilon \).

Then,
\[
\limsup_n \sup_{f \in L} |\mu_n(f) - \mu_0(f)| < 2 \epsilon + \limsup_n \sup_{f \in L} |Ef(Z_{n,k}) - Ef(Z_{0,k})|.
\]
It remains to show that \( \sup_{f \in L} |Ef(Z_{n,k}) - Ef(Z_{0,k})| \to 0 \). Since \( \mathcal{X} \) is independent of \( \mathcal{Y}_n \), up to changing \( (\Omega, \mathcal{A}, P_0) \) with some other probability space, it can be assumed

\[
P_0(Y_{n,j} \neq Y_{0,j} \text{ for some } j \leq k) = d_{TV}(\nu_{n,k}, \nu_{0,k});
\]

see Proposition 1. The same is true if \( \sigma(C_{n,j}) \subset \sigma(Y_{n,1}, \ldots, Y_{n,j}) \) for all \( n \) and \( j \).

Then, letting \( A_{n,k} = \{Y_{n,j} = Y_{0,j} \text{ for all } j \leq k\} \), one obtains

\[
\sup_{f \in L} |Ef(Z_{n,k}) - Ef(Z_{0,k})| \leq E \{ I_{A_{n,k}} 2 \wedge \|Z_{n,k} - Z_{0,k}\| \} + 2P_0(A_{n,k})
\]

\[
\leq E \{ 2 \wedge \sum_{j=1}^{\infty} |C_{n,j} - C_{0,j}| \} + 2d_{TV}(\nu_{n,k}, \nu_{0,k}) \to 0.
\]

Thus, condition (2) holds, and an application of Theorem 4 concludes the proof.

**APPENDIX**

Three preliminary lemmas are needed to prove Theorem 4. The first is part of the folklore about Skorohod distance, and we state it without a proof. Let \( \Delta x(t) = x(t) - x(t-) \) denote the jump of \( x \in \mathcal{D} \) at \( t \in (0, 1] \).

**Lemma 11.** Fix \( \epsilon > 0 \) and \( x_n \in \mathcal{D}, n \geq 0 \). Then, \( \lim \sup \|x_n - x_0\| \leq \epsilon \) whenever \( d(x_n, x_0) \to 0 \) and

\[
|\Delta x_n(t)| > \epsilon \quad \text{for all large } n \text{ and } t \in (0, 1) \text{ such that } |\Delta x_0(t)| > \epsilon.
\]

The second lemma is a consequence of Remark 6 of [5], but we give a sketch of its proof as it is basic for Theorem 4. Let \( \mu, \nu \) be laws on \( \mathcal{B}_d \) and \( \mathcal{F}(\mu, \nu) \) the class of probabilities \( \lambda \) on \( \mathcal{B}_d \otimes \mathcal{B}_d \) such that \( \lambda(\cdot \times \mathcal{D}) = \mu(\cdot) \) and \( \lambda(\mathcal{D} \times \cdot) = \nu(\cdot) \). Since the map \((x, y) \mapsto \|x - y\|\) is measurable w.r.t. \( \mathcal{B}_d \otimes \mathcal{B}_d \), one can define

\[
W_u(\mu, \nu) = \inf_{\lambda \in \mathcal{F}(\mu, \nu)} \int 1 \wedge \|x - y\| \lambda(dx, dy).
\]

**Lemma 12.** For a sequence \((\mu_n : n \geq 0)\) of probabilities on \( \mathcal{B}_d \), condition (1) holds if and only if \( W_u(\mu_0, \mu_n) \to 0 \).

**Proof.** The "only if" part is trivial. Suppose \( W_u(\mu_0, \mu_n) \to 0 \). Let \( \Omega = \mathcal{D}^\infty \), \( \mathcal{A} = \mathcal{B}_d^\infty \) and \( X_n : \mathcal{D}^n \to \mathcal{D} \) the \( n \)-th canonical projection, \( n \geq 0 \). Take \( \lambda_n \in \mathcal{F}(\mu_0, \mu_n) \) such that \( \int 1 \wedge \|x - y\| \lambda_n(dx, dy) < \frac{1}{n} + W_u(\mu_0, \mu_n) \). Since \((\mathcal{D}, d)\) is Polish, \( \lambda_n \) admits a disintegration \( \alpha_n = \{\alpha_n(x) : x \in \mathcal{D}\} \) (see Section 2). By Ionescu-Tulcea theorem, there is a unique probability \( P \) on \( \mathcal{B}_d^\infty \) such that \( X_0 \sim \mu_0 \) and

\[
\beta_n(x_0, x_1, \ldots, x_{n-1})(A) = \alpha_n(x_0)(A), \quad (x_0, x_1, \ldots, x_{n-1}) \in \mathcal{D}^n, A \in \mathcal{B}_d,
\]

is a regular version of the conditional distribution of \( X_n \) given \((X_0, X_1, \ldots, X_{n-1})\) for all \( n \geq 1 \). Under such \( P \), one obtains \((X_0, X_n) \sim \lambda_n \) (so that \( X_n \sim \mu_n \)) and

\[
\epsilon P(\|X_0 - X_n\| > \epsilon) \leq E_P \{ 1 \wedge \|X_0 - X_n\| \} < \frac{1}{n} + W_u(\mu_0, \mu_n) \to 0 \quad \text{for all } \epsilon \in (0, 1).
\]

The third lemma needs some more effort. Let \( \phi_0(x, \epsilon) = 0 \) and

\[
\phi_{n+1}(x, \epsilon) = \inf \{ t : \phi_n(x, \epsilon) < t \leq 1, |\Delta x(t)| > \epsilon \}
\]
where \( n \geq 0, \epsilon > 0, x \in D \) and \( \inf \emptyset := 1 \). The map \( x \mapsto \phi_n(x, \epsilon) \) is universally measurable w.r.t. \( \mathcal{B}_d \) for all \( n \) and \( \epsilon \).

**Lemma 13.** Let \( \mathcal{F}_k \) be the Borel \( \sigma \)-field on \( \mathbb{R}^k \) and \( I \subset (0,1) \) a dense subset. For a sequence \( \{\mu_n : n \geq 0\} \) of probabilities on \( \mathcal{B}_d \), condition (1) holds provided

\[
\sup_{A \in \mathcal{F}_k} \left| \int f(x) I_A(\phi_1(x, \epsilon), \ldots, \phi_k(x, \epsilon)) \, \mu_n(dx) - \int f(x) I_A(\phi_1(x, \epsilon), \ldots, \phi_k(x, \epsilon)) \, \mu_0(dx) \right| \longrightarrow 0
\]

for each \( k \geq 1, \epsilon \in I \) and function \( f : D \to [-1,1] \) such that \( |f(x) - f(y)| \leq d(x, y) \) for all \( x, y \in D \).

**Proof.** Fix \( \epsilon \in I \) and write \( \phi_n(x) \) instead of \( \phi_n(x, \epsilon) \). As each \( \phi_n \) is universally measurable w.r.t. \( \mathcal{B}_d \), there is a set \( T \in \mathcal{B}_d \) such that

\[
\mu_n(T) = 1 \quad \text{and} \quad I_T \phi_n \text{ is } \mathcal{B}_d\text{-measurable for all } n \geq 0.
\]

Thus, \( \phi_n \) can be assumed \( \mathcal{B}_d \)-measurable for all \( n \). Let \( k \) be such that

\[
\mu_0 \{ x : \phi_r(x) \neq 1 \text{ for some } r > k \} < \epsilon.
\]

For such a \( k \), define \( \phi(x) = (\phi_1(x), \ldots, \phi_k(x)) \), \( x \in D \), and

\[
\lambda_n(A) = \mu_n \{ x : (\phi(x), x) \in A \}, \quad A \in \mathcal{F}_k \otimes \mathcal{B}_d.
\]

Since \( (D, d) \) is Polish, Proposition 2 applies to such \( \lambda_n \) with \( (F, \mathcal{F}) = (\mathbb{R}^k, \mathcal{F}_k) \) and \( (G, \mathcal{G}) = (D, \mathcal{B}_d) \). Condition (b) holds by the assumption of the Lemma. Thus, by Proposition 2, on a probability space \( (\Omega, \mathcal{A}, P) \) there are measurable maps \( (Y_n, Z_n) : (\Omega, \mathcal{A}) \to (\mathbb{R}^k \times D, \mathcal{F}_k \otimes \mathcal{B}_d) \) satisfying

\[
(Y_n, Z_n) \sim \lambda_n \text{ for all } n \geq 0, \quad P(Y_n \neq Y_0) \longrightarrow 0, \quad d(Z_n, Z_0) \overset{P}{\longrightarrow} 0.
\]

Since \( P(Y_n = \phi(Z_n)) = \lambda_n \{ (\phi(x), x) : x \in D \} = 1 \), one also obtains

\[
P(\phi(Z_n) = \phi(Z_0)) = 1.
\]

Next, by (4) and \( d(Z_n, Z_0) \overset{P}{\longrightarrow} 0 \), there is a subsequence \( (n_j) \) such that

\[
\lim_{n} P(\|Z_n - Z_0\| > \epsilon) = \lim_{j} P(\|Z_{n_j} - Z_0\| > \epsilon),
\]

\[
d(Z_{n_j}, Z_0) \overset{a.s.}{\longrightarrow} 0, \quad P(\phi(Z_{n_j}) = \phi(Z_0) \text{ for all } j) > 1 - \epsilon.
\]

Define \( U = \limsup_{n} \|Z_{n_j} - Z_0\| \) and

\[ H = \{ \phi_r(Z_0) = 1 \text{ for all } r > k \} \cap \{ \phi(Z_{n_j}) = \phi(Z_0) \text{ for all } j \} \cap \{ d(Z_{n_j}, Z_0) \overset{a.s.}{\longrightarrow} 0 \}.
\]

For each \( \omega \in H \), Lemma 11 applies to \( Z_0(\omega) \) and \( Z_{n_j}(\omega) \), so that \( U(\omega) \leq \epsilon \). Further,

\[
P(H^c) \leq P(\phi_r(Z_0) \neq 1 \text{ for some } r > k) + P(\phi(Z_{n_j}) \neq \phi(Z_0) \text{ for some } j) < \mu_0 \{ x : \phi_r(x) \neq 1 \text{ for some } r > k \} + \epsilon < 2 \epsilon.
\]

Since \( U \leq \epsilon \) on \( H \),

\[
\limsup_{n} P(\|Z_n - Z_0\| > \epsilon) = \lim_{j} P(\|Z_{n_j} - Z_0\| > \epsilon) \leq P(U \geq \epsilon) \leq P(U = \epsilon) + P(H^c) < P(U = \epsilon) + 2 \epsilon.
\]

On noting that \( E_P \{ 1 \wedge \|Z_0 - Z_n\| \} \leq \epsilon + P(\|Z_n - Z_0\| > \epsilon) \), one obtains

\[
\limsup_{n} W_n(\mu_0, \mu_n) \leq \limsup_{n} E_P \{ 1 \wedge \|Z_0 - Z_n\| \} < P(U = \epsilon) + 3 \epsilon.
\]
Since $I$ is dense in $(0, 1)$, then $P(U = \epsilon) + 3\epsilon$ can be made arbitrarily small for a suitable $\epsilon \in I$. Thus, $\limsup_n W_n(\mu_0, \mu_n) = 0$. An application of Lemma 12 concludes the proof.

We are now ready for the last attack to Theorem 4.

**Proof of Theorem 4.** "(1) $\Rightarrow$ (2)". Just note that

\[
|\mu_n(f) - \mu_0(f)| = |E_P\{f(X_n)\} - E_P\{f(X_0)\}| \leq E_P\{2 \wedge \|X_n - X_0\|\} \longrightarrow 0, \quad \text{for each } f \in L, \text{ under (1)}.
\]

"(2) $\Rightarrow$ (1)". Let $B_\epsilon = \{x : |\Delta x(t)| = \epsilon \text{ for some } t \in (0, 1]\}$. Then, $B_\epsilon$ is universally measurable w.r.t. $B_d$ and $\mu_0(B_\epsilon) > 0$ for at most countably many $\epsilon > 0$. Hence, $I = \{\epsilon \in (0, 1) : \mu_0(B_\epsilon) = 0\}$ is dense in $(0, 1)$.

Fix $\epsilon \in I$, $k \geq 1$, and a function $f : D \to [-1, 1]$ such that $|f(x) - f(y)| \leq d(x, y)$ for all $x, y \in D$. By Lemma 13, for condition (1) to be true, it is enough that

\[
\limsup_n A_{\phi}(\mu) = 0
\]

where $\phi(n) = \{\phi_1(x), ..., \phi_k(x)\}, x \in D$, and $\phi_j(x) = \phi_j(x, \epsilon)$ for all $j$.

In order to prove (5), given $b \in (0, \frac{1}{2})$, define

\[
F_b = \{x : |\Delta x(t)| < b, \text{ for all } t \in (0, 1]\}, \quad G_b = \{x : d(x, F_b) \geq \frac{b}{2}\}
\]

Then,

(i) $G_b \subseteq F_b/2$; (ii) $\phi(x) = \phi(y)$ whenever $x, y \in F_b$ and $\|x - y\| < b$.

Statement (ii) is straightforward. To check (i), fix $x \notin G_b$ and take $y \in F_b$ with $d(x, y) < b/2$. Let $\gamma : [0, 1] \to [0, 1]$ be a strictly increasing homeomorphism such that $\|x - y \circ \gamma\| < b/2$. For all $t \in (0, 1]$,

\[
|\Delta x(t)| \leq |\Delta y \circ \gamma(t)| + 2||x - y \circ \gamma|| < |\Delta y(\gamma(t))| + b.
\]

Similarly, $|\Delta x(t)| > |\Delta y(\gamma(t))| - b$. Since $y \in F_b$, it follows that $x \in F_b/2$.

Next, define

\[
\psi_b(x) = \frac{d(x, G_b)}{d(x, F_b) + d(x, G_b)}, \quad x \in D.
\]

Then, $\psi_b = 0$ on $G_b$ and $\psi_b$ is Lipschitz w.r.t. $d$ with Lipschitz constant $\frac{b}{d}$. Hence, $\psi_b$ is Lipschitz w.r.t. $u$ with Lipschitz constant $\frac{d}{u}$ (since $d \leq u$). Basing on (i)-(ii) and such properties of $\psi_b$, it is not hard to check that $\psi_b I_A(\phi)$ is Lipschitz w.r.t. $u$, with Lipschitz constant $\frac{b}{d}$, for every $A \in F_k$. In turn, since $d \leq u$ and $f$ is Lipschitz w.r.t. $d$ with Lipschitz constant $1$,

\[
f_A = f \psi_b I_A(\phi), \quad A \in F_k,
\]

is Lipschitz w.r.t. $u$ with Lipschitz constant $(1 + \frac{b}{d})$. Moreover,

\[
|\mu_0\{f I_A(\phi)\} - \mu_n(f_A)| \leq \mu_n|f I_A(\phi) (1 - \psi_b)| \leq \mu_n(1 - \psi_b).
\]

On noting that $(1 + \frac{b}{d})^{-1} f_A \in L$ for every $A \in F_k$, condition (2) yields

\[
\limsup_n \sup_{A \in F_k} |\mu_0\{f I_A(\phi)\} - \mu_0\{f I_A(\phi)\} | \leq \limsup_n \{\mu_n(1 - \psi_b) + \sup_{A \in F_k} |\mu_0(f_A) - \mu_0(f_A)| + \mu_0(1 - \psi_b)\} = 2 \mu_0(1 - \psi_b) \leq 2 \mu_0(F_b).$

Since $\epsilon \in I$ and $\bigcap_{b > 0} F_c^b = \{ x : |\Delta x(t)| = \epsilon \text{ for some } t \} = B_\epsilon$, one obtains
\[
\limsup_n \sup_{A \in F_k} |\mu_n\{f I_A(\phi)\} - \mu_0\{f I_A(\phi)\}| \leq 2 \lim_{b \to 0} \mu_0(F_c^b) = 2 \mu_0(B_\epsilon) = 0.
\]
Therefore, condition (5) holds and this concludes the proof. \hfill \Box

References


Patrizia Berti, Dipartimento di Matematica Pura ed Applicata "G. Vitali", Universita' di Modena e Reggio-Emilia, via Campi 213/B, 41100 Modena, Italy
E-mail address: berti.patrizia@unimore.it

Luca Pratelli, Accademia Navale, viale Italia 72, 57100 Livorno, Italy
E-mail address: pratel@mail.dm.unipi.it

Pietro Rigo (corresponding author), Dipartimento di Economia Politica e Metodi Quantitativi, Universita' di Pavia, via S. Felice 5, 27100 Pavia, Italy
E-mail address: prigo@eco.unipv.it